



NEUROSCIENCE

Expansion of learning capacity elicited by interspecific hybridization

Yukino Shibata^{1,2}, Noriyuki Toji^{2,3}, Hongdi Wang⁴, Yasuhiro Go^{5,6,7}, Kazuhiro Wada^{1,3,8*}

Learned behavior, a fundamental adaptive trait in fluctuating environments, is shaped by species-specific constraints. This phenomenon is evident in songbirds, which acquire their species-specific songs through vocal learning. To explore the neurogenetic mechanisms underlying species-specific song learning, we generated F₁ hybrid songbirds by crossing *Taeniopygia guttata* with *Aidemosyne modesta*. These F₁ hybrids demonstrate expanded learning capacities, adeptly mimicking songs from both parental species and other heterospecific songs more extensively than their parental counterparts. Despite the conserved size of brain regions and neuron numbers in the neural circuits for song learning and production, single-cell transcriptomics reveals distinctive transcriptional characteristics in the F₁ hybrids, especially in vocal-motor projection neurons. These neurons exhibit enrichment for nonadditively expressed genes, particularly those related to ion channel activity and cell adhesion, which are associated with the degree of song learning among F₁ individuals. Our findings provide insights into the emergence of altered learning capabilities through hybridization, linked to cell type-specific transcriptional changes.

INTRODUCTION

Learned behavior is crucial for animals' adaptive responses and survival under fluctuating environments (1, 2). Learning capacities, such as memory retention and cognitive bias, which can affect the acquisition and manifestation of learned behaviors, are not only influenced by experience but also by species-specific genetic features (3, 4). However, the neurogenetic mechanisms underlying species-specific learning capacity remain unclear.

Songbirds acquire their songs through vocal learning, influenced by species-specific learning capacities. These capacities are constrained by various factors, including learning biases to conspecific songs, the ability to memorize and produce multiple song repertoires, and the specific developmental timing at which new songs can be learned (5–8). These species-specific songs are essential acoustic signals for mating and territorial defense in social interactions within and between species (9, 10). For juvenile songbirds, auditory exposure to conspecific songs is crucial for memorizing a reference model for song learning (11, 12). However, certain species-specific song features still manifest even without song tutoring (6, 13, 14), indicating a synergistic interaction between experience and genetic predispositions during song learning (5, 7, 15, 16). While songbird species exhibit diverse species-specific song features, such as unique syllables and complex syllable sequences (8), the neural circuits involved in vocal learning and production, called the “song circuits,” are specialized and conserved across species. However, the mechanisms by which such conserved neural substrates regulate species-specific song learning remain unclear.

Interspecific F₁ hybrids have a set of polymorphic alleles in heterozygous genomes inherited from both parental species. Thus, they exhibit various phenotypic traits, ranging from an average of both parental species to dominant or recessive effects resembling either parental species. In certain cases, F₁ hybrids display phenotypic novelty, such as exceptional phenotypic expression (heterosis/hybrid vigor) or broader variability compared to either parent species (17–20). Therefore, hybridization serves as a notable evolutionary driver, generating phenotypic diversity across many species (19, 21). While numerous studies have focused on morphological and physiological traits in F₁ hybrids in comparison with their parental species, a few studies have examined the potential alterations in latent cognitive abilities, such as memory and learning (22–25). Specifically, research on vocal learning in interspecies hybrids remains scarce. Some studies have suggested that hybrid songbirds, both in the wild and in laboratory settings, produce songs that exhibit intermediate characteristics between those of the two parental species, along with individual diversity. However, these studies often fall short of comprehensively assessing the quality of song learning and development or elucidating the neurogenetic mechanisms in F₁ hybrids (26–29).

In this study, we generated interspecific F₁ hybrid songbirds to investigate the potential phenotypic novelty in vocal learning during song acquisition. We crossed *Taeniopygia guttata* (zebra finch: ZF) with *Aidemosyne modesta* (cherry finch: CF, or plum-headed finch) (Fig. 1A), which are partially sympatric in Australia (30). The males of both parental species produce species-specific songs characterized by acoustic units (syllables), temporal structure (sequence), and prosody (rhythm) (fig. S1) (30–32). ZFs exhibit a learning bias toward conspecific songs and have a limited capacity to learn heterospecific songs (27, 33, 34). In contrast, CFs have not been studied regarding their learning traits for song acquisition. Since ZF and CF belong to different genera, we hypothesized that their genetic divergence might increase the heterozygosity, thereby potentially enhancing the phenotypic novelties and diversities of song learning in the F₁ hybrids (19, 20).

Using F₁ hybrids and their parental species, ZFs and CFs, we integrated multiple approaches—including an experimentally controlled

¹Graduate School of Life Science, Hokkaido University, Sapporo 060-0810, Japan.

²Research Fellowship for Young Scientists of the Japan Society for the Promotion of Science, Sapporo 060-0810, Japan. ³Faculty of Science, Hokkaido University, Sapporo 060-0810, Japan. ⁴Evolutionary Neurobiology Unit, Okinawa Institute of Science and Technology Graduate University, Okinawa 904-0497, Japan. ⁵Graduate School of Information Science, University of Hyogo, Kobe 650-0047, Japan. ⁶Exploratory Research Center on Life and Living Systems (ExCELLS), National Institutes of Natural Sciences (NINS), Okazaki 444-8585, Japan. ⁷National Institute for Physiological Sciences (NIPS), National Institutes of Natural Sciences (NINS), Okazaki 444-8585, Japan. ⁸Research and Education Center for Brain Science, Hokkaido University, Sapporo 060-8638, Japan.

*Corresponding author. Email: wada@sci.hokudai.ac.jp

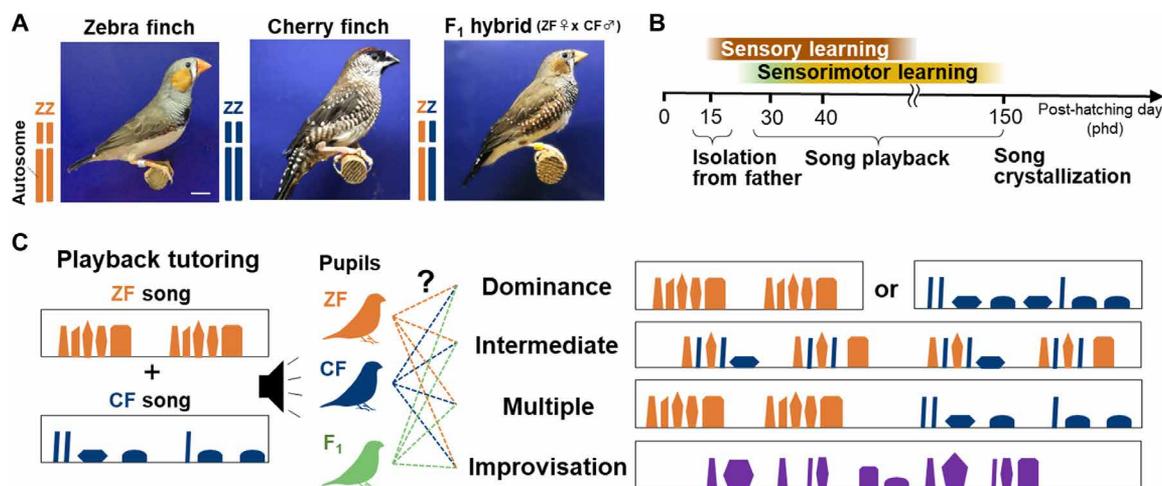


Fig. 1. Song tutoring of F₁ hybrids crossing zebra and cherry finches. (A) Photographs of a ZF, CF, and an F₁ hybrid (ZF♀ x CF♂). Scale bar, 1 cm. Vertical bars represent the genome composition of F₁ hybrids generated by crossing ZF with CF. Male F₁ hybrids have identical sets of autosomal and sex chromosomes from ZF (orange) and CF (blue). (B) Timeline illustrating the process of song tutoring and learning. (C) Predicting the potential forms of song acquisition through playback tutoring with songs from two parent species: Dominance, where the pupil selectively learns either ZF or CF tutor song; Intermediate, involving a mix of syllables or sequential features from both parental species; Multiple, where the pupil learns both species' tutor songs correctly, retaining species-specific syllable acoustics and sequence patterns; and Improvisation, where the pupil acquires a novel song dissimilar to either tutor song.

song tutoring system, mesoscopic anatomical analysis, and single-nuclei RNA sequencing (snRNA-seq)—to explore the neurogenetic mechanisms underlying the learning capabilities of F₁ hybrids compared to their parental species. We found that the F₁ hybrids have the learning capacity to acquire songs from both parental species, exhibiting heterotic traits in song learning not observed in the parental species. Moreover, unlike their parental counterparts, the F₁ hybrids have an expanded ability to learn genetically unrelated heterospecific songs. The snRNA-seq profiling reveals distinct transcriptional signatures with the unique expression of nonadditive genes in the song circuits of F₁ hybrids compared to their parental species.

RESULTS

Heterosis in vocal learning in the F₁ hybrids

To investigate vocal learning characteristics in interspecific F₁ hybrids and their parental species, we tutored juveniles from three groups—ZFs, CFs, and F₁ hybrids—during the sensitive period for vocal learning, using playback of recorded songs from both parental species (Fig. 1 and fig. S2). Following the song tutoring conditions, ZF pupils predominantly exhibited learning biases toward conspecific songs, while most CF pupils showed very limited learning of tutored songs, including their conspecific songs (Fig. 2A; see fig. S2 for tutoring with the single parental species songs). In contrast, the F₁ hybrids acquired songs from both parental species, manifesting as independent song repertoires (Fig. 2A and movie S1). Notably, approximately 30% of F₁ hybrids exhibited near-perfect imitation of both the prosody and song sequence from the songs of both parental species. This learning tendency was similarly observed among F₁ hybrids from reciprocal parental combinations (fig. S2). However, because of challenges in obtaining a sufficient number of F₁ hybrids from CF♀ and ZF♂ pairing, our analysis exclusively focused on F₁ hybrids derived from ZF♀ and CF♂ for comprehensive investigations.

Through a quantitative analysis comparing tutored songs and songs acquired by pupils (see Materials and Methods), we found that the F₁ hybrids learned more syllables from the tutored songs than either of the parental species (Steel-Dwass test, **P* < 0.05, ***P* < 0.01, ****P* < 0.001) (Fig. 2B). Furthermore, the F₁ hybrids developed a broader syllable repertoire size (unique syllable types) compared to ZFs and CFs (Steel-Dwass test, **P* < 0.05, ***P* < 0.01) (Fig. 2C). The correlation between these two song attributes in the F₁ hybrids was statistically significant (Spearman's rank correlation, *P* = 0.0001, *r* = 0.77) (Fig. 2D), suggesting that the F₁ hybrids had an expanded learning capacity for song acquisition compared to those of the parental species.

We then examined whether F₁ hybrids could modify their inherent song characteristics, specifically their syllable repertoire size, to boost song learning capacity. Using socially isolated rearing, we ensured that F₁ hybrid pupils developed songs without the influence of tutor model songs (Fig. 2E). These socially isolated F₁ hybrids had significantly smaller syllable repertoire sizes compared to F₁ pupils tutored with songs from both parental species (*n* = 6; Steel-Dwass test, **P* < 0.05) (Fig. 2C). This result underscores the importance of providing tutor song models during vocal learning to facilitate song acquisition in the F₁ hybrids, suggesting a heterotic trait for song learning in these hybrids.

Vocal mimicry of heterospecific songs by F₁ hybrids

Previous studies have indicated the ability of ZFs to learn heterospecific songs (33, 34). However, when tutored with heterospecific songs, the songs acquired by ZFs often exhibit limitations in terms of the number of syllables learned and their sequential transitions compared to the original tutored heterospecific songs. The finding that the F₁ hybrids demonstrate a capacity to learn songs from two distinct species led us to investigate whether they could mimic the songs of heterospecifics with greater accuracy than their parental species. To explore this hypothesis, F₁ hybrid individuals were

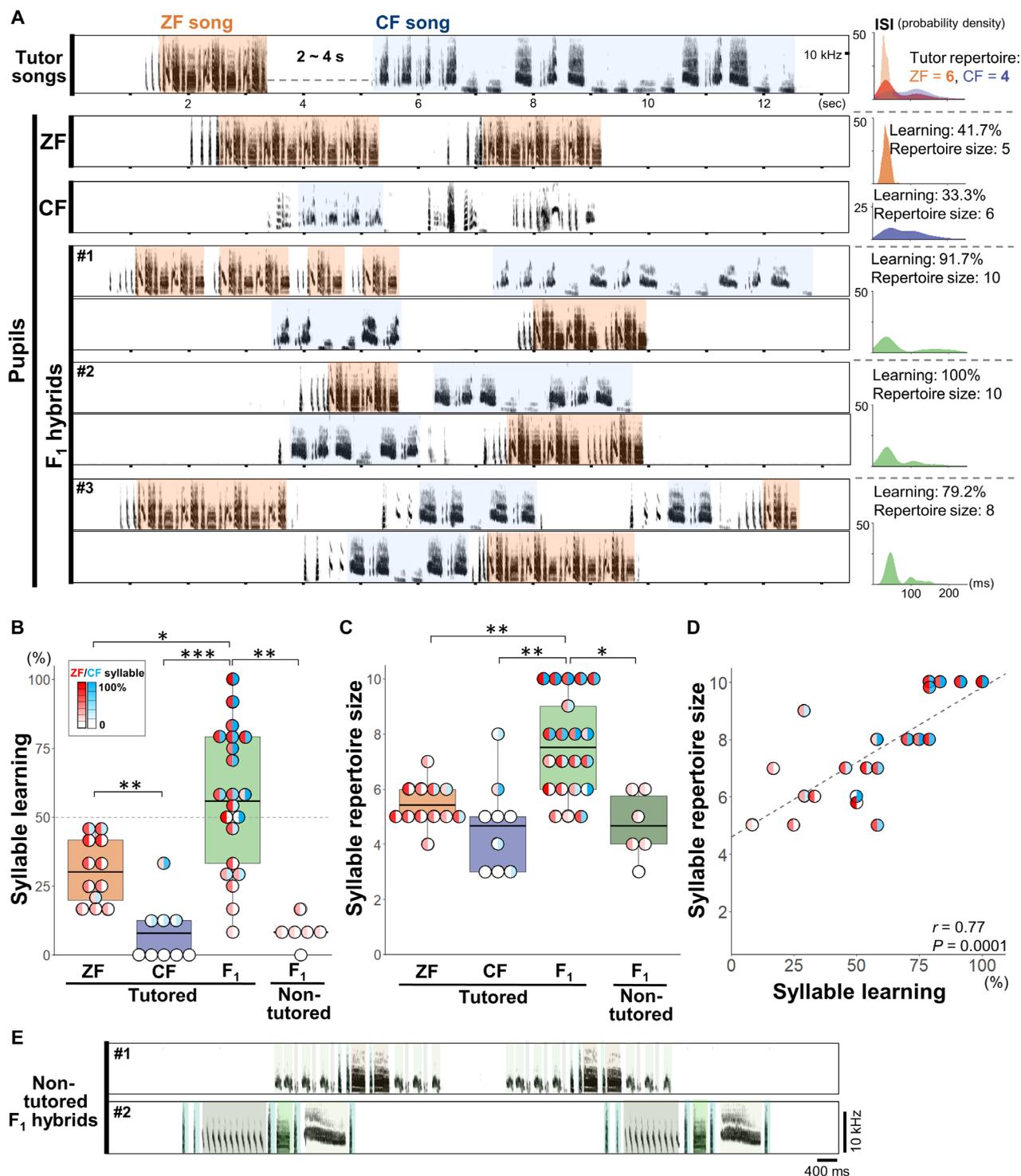


Fig. 2. Heterosis in song learning of F₁ hybrids crossing ZF with CF. (A) (Left) Examples of songs acquired by ZF, CF, and F₁ hybrids after song tutoring with playback of both ZF and CF songs. ZF- and CF-like song structures are colored orange and blue, respectively. (Right) Syllable learning rate, syllable repertoire size, and distribution pattern of inter-syllable intervals (ISI) of the respective birds. (B) Syllable learning rate of ZFs, CFs, and F₁ hybrids tutored with both ZF and CF songs. The intensity of red and blue colors in each circle indicates the syllable learning rate from ZF and CF tutor songs, respectively, for each pupil bird. The dotted line at 50% indicates cases where birds may exclusively learn from one parental species or a mix of syllables from both species (ZF = 12, CF = 9, F₁ = 21, and 6 birds tutored and non-tutored, respectively; * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, Steel-Dwass test). (C) Syllable repertoire size (the number of unique syllables in songs) acquired by ZFs, CFs, and F₁ hybrids tutored with both ZF and CF songs. The dark green box plot represents the syllable repertoire size of non-tutored F₁ hybrids (* $P < 0.05$, ** $P < 0.01$, Steel-Dwass test). (D) Correlation between syllable learning rate and syllable repertoire size of F₁ hybrids ($r = 0.77$, $P = 0.0001$, Spearman's rank correlation). (E) Examples of songs of the socially isolated, non-tutored F₁ hybrids ($n = 2$ birds).

exposed to songs of the owl finch (OF; *Taeniopygia bichenovii*), Bengalese finch (BF; *Lonchura striata* var. *Domestica*), and canary (CN; *Serinus canaria*) (fig. S3). These species are classified within the same, proximal, and distal genera in relation to the ZF, respectively (Fig. 3A) (30, 35). The acoustic characteristics and sequential features of the songs from these three species not only differ among themselves but also from the songs of the F₁ hybrids' parental species (Fig. 3, B to D) (27, 36).

Even with passive exposure to song playback through a speaker alone, which is considered less effective for song learning than tutoring under social interaction with live birds (37, 38), we observed that the F₁ hybrids exhibited clear learning outcomes across multiple

song features, such as song acoustics, sequence, and duration (Fig. 3, B to D). Notably, all the F₁ hybrids successfully learned OF songs, reproducing them near-perfectly to the human ear ($n = 3$) (Fig. 3B, fig. S4, and movie S2). While the F₁ hybrids tutored with BF or CN songs struggled to mimic sections with high-frequency syllable repetitions, these pupils still acquired the songs, showing notable similarity in both syllable acoustics and sequence to the tutored heterospecific songs (BF and CN songs tutored $n = 4$ and 2, respectively) (Fig. 3B, fig. S4, and movie S2). Through a comprehensive evaluation of song similarity, including syllable acoustics, syllable transition patterns, and song-bout duration, between the tutored songs and pupils' acquired songs, we found that most F₁ hybrids

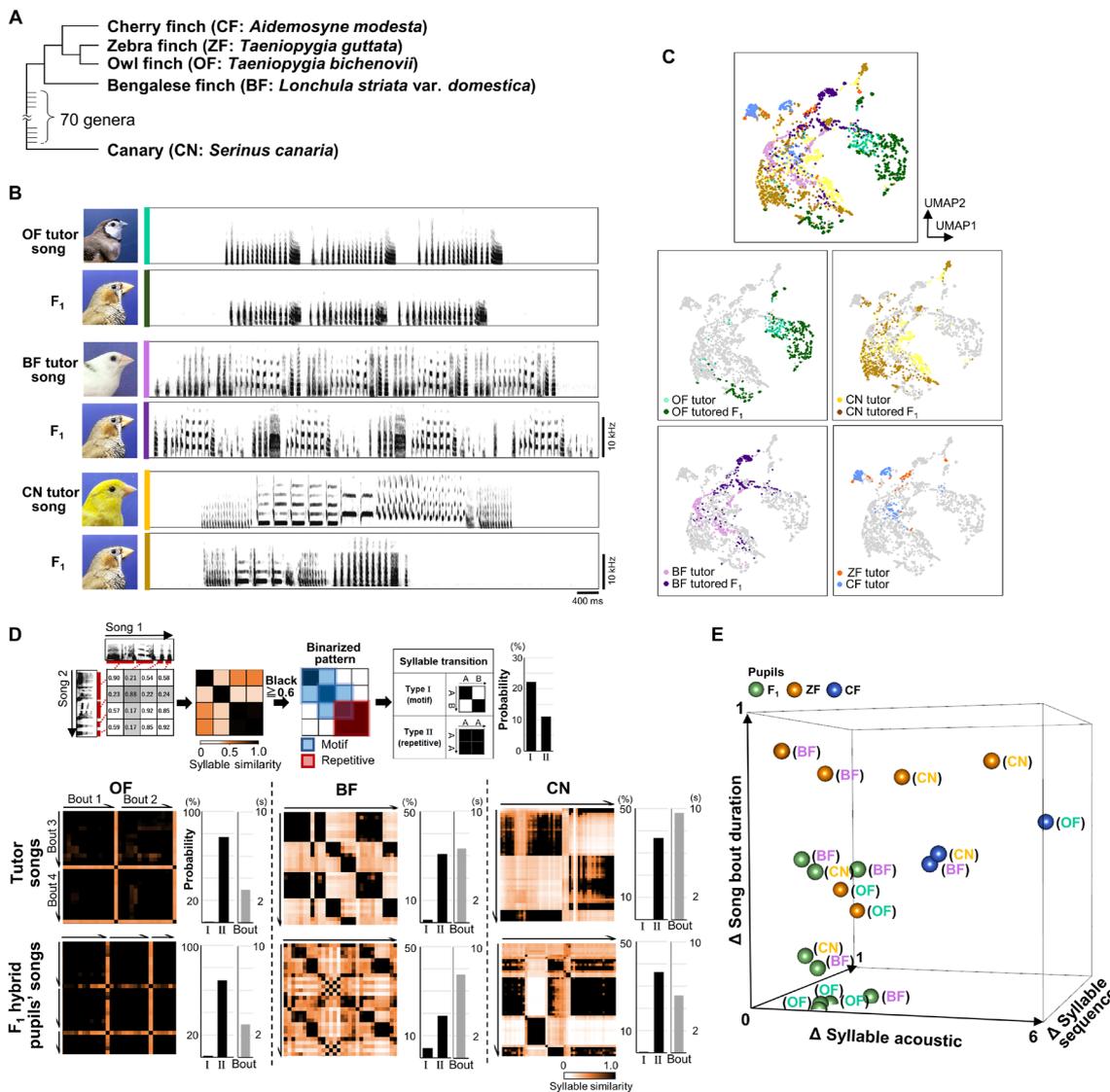


Fig. 3. Songs of F₁ hybrids tutored with songs from genetically unrelated species. (A) Phylogenetic relationships among ZF, CF, OF, BF, and CN. (B) Examples of songs acquired by F₁ hybrids tutored with playback of OF, BF, or CN songs. (C) UMAPs displaying song syllables from F₁ pupils and playback tutor songs (289 to 300 syllables from F₁ pupils' songs; 66 to 244 syllables from playback tutor songs). (D) (Top) Syllable similarity matrix (SSM) analysis used to identify syllable transition types: types I and II as motif- and repetition-syllable transitions, respectively. (Bottom left) Representative SSMs of the playback tutor songs and songs acquired by F₁ pupils. Each SSM was created from one to three song bouts of one representative bird. (Bottom right bar graphs) Occurrence rates of syllable transition types I (motif) and II (repetition) for playback tutor and pupil songs (black bars), along with the average duration of song bouts (gray bars). (E) Three-dimensional plot showing differences in syllable acoustics, syllable sequence, and song-bout duration between tutor and pupil songs. The species of pupils and tutored songs are represented by the dot colors and the characters in parentheses, respectively. A smaller distance from the coordinate origin indicates a higher similarity in the mimicry of the tutor songs by pupils.

demonstrated better learning in acquiring heterospecific songs compared to their parental species counterparts, regardless of the tutored species' songs (Fig. 3E and fig. S4). These findings indicate that despite individual variation in learning outcomes among F₁ hybrids, they have an expanded learning capability to imitate songs from hetero-specific species with greater fidelity than their parental species.

Consistent brain size between F₁ hybrids and their parental species

Past studies have demonstrated a positive correlation between syllable repertoire size and the song nuclei size or neuron number,

especially in the posterior vocal-motor pathway (VMP), across songbird species (Fig. 4A) (39–41). On this basis, we speculated whether the heterotic trait of F₁ hybrids in song learning, which allows them to acquire a more extensive syllable repertoire, might be attributed to an enlarged size or increased neuron number of certain song nuclei in the song circuits, relative to their parental species (Fig. 4B). To delineate the targeted song nuclei precisely, we used gene markers, *androgen receptor* (*AR*) for HVC (used as a proper name) and the lateral magnocellular nucleus of the anterior nidopallium (LMAN), *GRIK1* for the robust nucleus of the arcopallium (RA), *GRIA1* for the basal ganglia nucleus Area X, and *GRM2* for

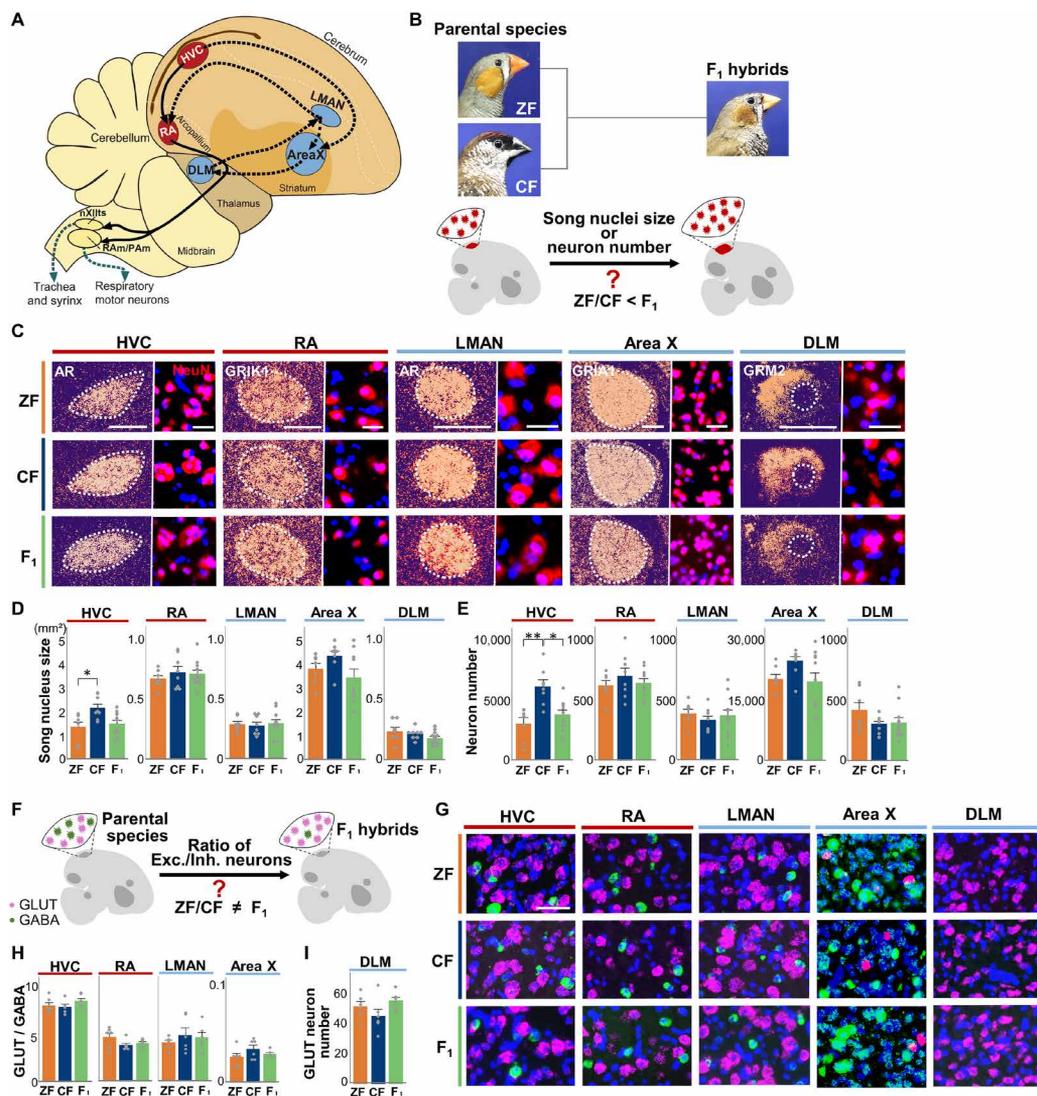


Fig. 4. The mesoscopic similarity of song circuits between F₁ hybrids and their parental species. (A) A diagram of the song circuits for song learning and production. Solid and dashed black lines denote the VMP and the AFP, respectively. (B) Schematic image of the potential mesoscopic changes in the size and/or neuron number of the song nuclei in F₁ hybrids compared to the parental species. (C) (Left) Boundary of song nuclei depicted by expressions of marker genes. Scale bar, 0.5 mm. (Right) Neurons labeled with anti-*NeuN* antibody (red) and 4',6-diamidino-2-phenylindole (DAPI, blue). Scale bar, 25 μm. (D) Size of song nuclei (HVC, RA, LMAN, Area X, and DLM) in ZFs, CFs, and F₁ hybrids (ZF = 7, CF = 8, F₁ = 11 birds; **P* < 0.05, Steel-Dwss test). (E) Neuron number of the song nuclei of ZFs, CFs, and F₁ hybrids (ZF = 7, CF = 8, and F₁ = 11 birds; **P* < 0.05, Steel-Dwss test). (F) Schematic image depicting the potential change in the ratio of excitatory and inhibitory neurons in the song nuclei of F₁ hybrids compared to the parental species. (G) *vGluT2*-positive excitatory neurons (magenta), *Gad1*-positive inhibitory neurons (green), and DAPI cell nuclear staining (blue) in HVC, RA, LMAN, Area X, and DLM. Scale bar, 50 μm. (H) Quantification of the ratio of excitatory and inhibitory neurons in the song nuclei (HVC, RA, LMAN, and Area X) of ZFs, CFs, and F₁ hybrids (*n* = 6 birds per group; ^{NS}*P* > 0.05, Steel-Dwss test). (I) The number of excitatory neurons in DLM of ZFs, CFs, and F₁ hybrids (*n* = 6 birds per group; ^{NS}*P* > 0.05, Steel-Dwss test). In DLM, only excitatory neurons were observed as GABA inhibitory neurons were absent.

the dorsal lateral nucleus of the medial thalamus (DLM) as delineated by less *GRM2* expression in DLM than the surrounding thalamic region (Fig. 4C) (42, 43). However, we observed no heterotic alterations in the size and neuron number of the examined song nuclei among F₁ hybrids, ZFs, and CFs (Fig. 4, D and E). Furthermore, no significant differences were found in the ratio of excitatory (*SLC17A6/vGlut*+) to inhibitory (*Gad1*+) neurons in these song nuclei between the F₁ hybrids and the parental species (Fig. 4, F to I).

Cell type-specific transcriptional divergence in song nuclei of F₁ hybrids

We then used snRNA-seq to explore the potential transcriptional divergence in specific cell types in the song circuits of F₁ hybrids relative to their parental species. To this end, we used the brain tissues from the F₁ hybrids tutored with both parental species' songs ($n = 4$ birds with average score of syllable learning: 82.3%, range 58.3 to 100% in Fig. 2B) and compared their transcriptional characteristics with their parental species counterparts. Our investigation focused on four distinct brain regions: HVC, RA, Area X, and the caudomedial nidopallium (NCM), while excluding the smaller song nuclei LMAN and DLM, to ensure an adequate collection of cells from individual birds for conducting snRNA-seq (Figs. 4A and 5A). HVC and RA, located in the VMP, are crucial for acquiring and producing song sequences and acoustics (44–48). Contrarily, Area X, a component of the anterior forebrain pathway, modulates vocal fluctuation to maintain song quality during and after song learning (49). NCM was selected as a higher auditory site involved in song perception and memorization (50–54).

In the Uniform Manifold Approximation and Projection (UMAP) plots set with the same input principal component dimensions (dim = 1:15), we assorted cell clusters by cell type-specific gene markers for glutamatergic (GLUT) and GABAergic (GABA) neurons, medium spiny neurons (MSN), astrocytes, oligodendrocytes, oligodendrocyte precursor cells (OPC), and microglia (Fig. 5B). For each brain region (HVC, RA, Area X, and NCM), we adjusted to use every 1200 cells per bird group (i.e., ZFs, CFs, and F₁ hybrids) for constructing UMAPs and then colored each cell differently to fit their genotypes (Fig. 5C). Across all the examined brain regions, the distributions of GABA neurons, microglia, oligodendrocytes, OPCs, and MSN subtype 3 (MSN3) were almost indistinguishable among the three groups despite the presence of multiple diverse subtypes of GABA neurons (blue-shaded areas in Fig. 5C). This indicates similar transcriptional features of GABA neurons across the brain regions of the three groups. In addition, astrocytes in all areas and oligodendrocytes in RA exhibited intermediate distributions in the F₁ hybrids compared to the counterparts in the two parental species within the same clusters of each cell type (orange-shaded areas in Fig. 5C).

We found that GLUT neurons in both HVC and RA, which primarily represent the projection neurons from HVC to RA [HVC_(RA)-PNs] and RA projection neurons (RA-PNs) that lead to the downstream peripheral vocal nuclei, along with MSN1 and MSN2 in Area X, were compartmentalized into three discrete cell clusters corresponding to ZFs, CFs, and F₁ hybrids (red-shaded areas in Fig. 5C). Conversely, in the auditory area NCM, GLUT neuron distribution patterns were integrated across the three groups. To confirm these findings, we constructed a unified UMAP that included all cells from HVC, RA, Area X, and NCM. This further validated the distinct cell type-specific distribution patterns of

transcriptional traits among F₁ hybrids, ZFs, and CFs (Fig. 5D). These results showed consistent cell distribution patterns, where GLUT-projecting neurons in the vocal-motor song nuclei HVC and RA, as well as MSN1 in Area X, exhibited distinct transcriptional signatures across the three groups. Notably, HVC_(RA)-PNs and RA-PNs are distributed as completely separated cell clusters among F₁ hybrids, ZFs, and CFs. In contrast, for all other remaining cell types, the F₁ hybrids displayed transcriptional trends that were either intermediate or similar to those of the parental species. This finding was further supported by the observation of a larger number of differentially expressed genes (DEGs) in GLUT HVC_(RA)-PNs and RA-PNs compared to other cell types across the three groups (Fig. 5E).

Enrichment of nonadditively expressed genes in GLUT projection neurons in the vocal-motor circuits

During our analysis of DEGs between the F₁ hybrids and their parental species, we identified a certain number of nonadditive (transgressive) genes present in each cell type of the examined brain regions (Fig. 6, A and B, and fig. S5). Nonadditively expressed genes are those that exhibit preferential expression toward one parental species, manifesting as either a dominance in gene expression (i.e., ZF or CF bias) or as over- or underdominance compared to both parental species (Fig. 6A). The expression of such nonadditive genes has been suggested to be one of the biological mechanisms underlying heterosis (55–57). Nonetheless, a comprehensive understanding of the cellular mechanisms contributing to heterosis, particularly at the single-cell level, remains elusive.

Consistent with the observation of the most distinctive transcriptional signatures in GLUT HVC_(RA)-PNs and RA-PNs of F₁ hybrids compared to ZFs and CFs, we found that these GLUT-projecting neurons in the VMP expressed more nonadditive genes (189 and 406 genes, respectively) than other cell types (Fig. 6B). In contrast, 47 genes were identified as nonadditively expressed genes in GLUT neurons in NCM, indicating that not only GLUT neuron type-specific but also brain region-selective regulation of nonadditive genes. GABA neurons expressed fewer nonadditive genes across the examined brain regions (6 to 40 genes). MSNs and glial cell types also exhibited relatively moderate expression of nonadditive genes. To examine the impact of nonadditive gene expression on the transcriptional characteristics in each cell type of F₁, ZF, and CF, we reconstructed the UMAP after excluding the nonadditively expressed genes (Fig. 6C and fig. S5). As a result, the distinct compartmentalized distribution patterns of the cell clusters of HVC_(RA)-PNs and RA-PNs from the F₁ hybrids, in contrast to ZFs and CFs, were no longer evident; they appeared more intermediary or merged. In contrast, such cell-distribution pattern changes were not observed when additive genes that were randomly selected in the same number as nonadditive genes were excluded (Fig. 6C and fig. S5). This suggests that the expression of nonadditive genes, particularly enriched in GLUT-projecting neurons in HVC and RA, could play a pivotal role in shaping the altered transcriptional characteristics in the F₁ hybrids compared to the parental species.

In addition, functional enrichment analysis of the nonadditively expressed genes identified in the brain regions of the F₁ hybrids revealed statistical significance for a variety of Gene Ontology (GO) terms (Fig. 6D). Consistent with the number of nonadditive genes identified in each cell type, only a limited number of GO terms (zero to five terms) were associated with nonadditive genes in MSN1 in

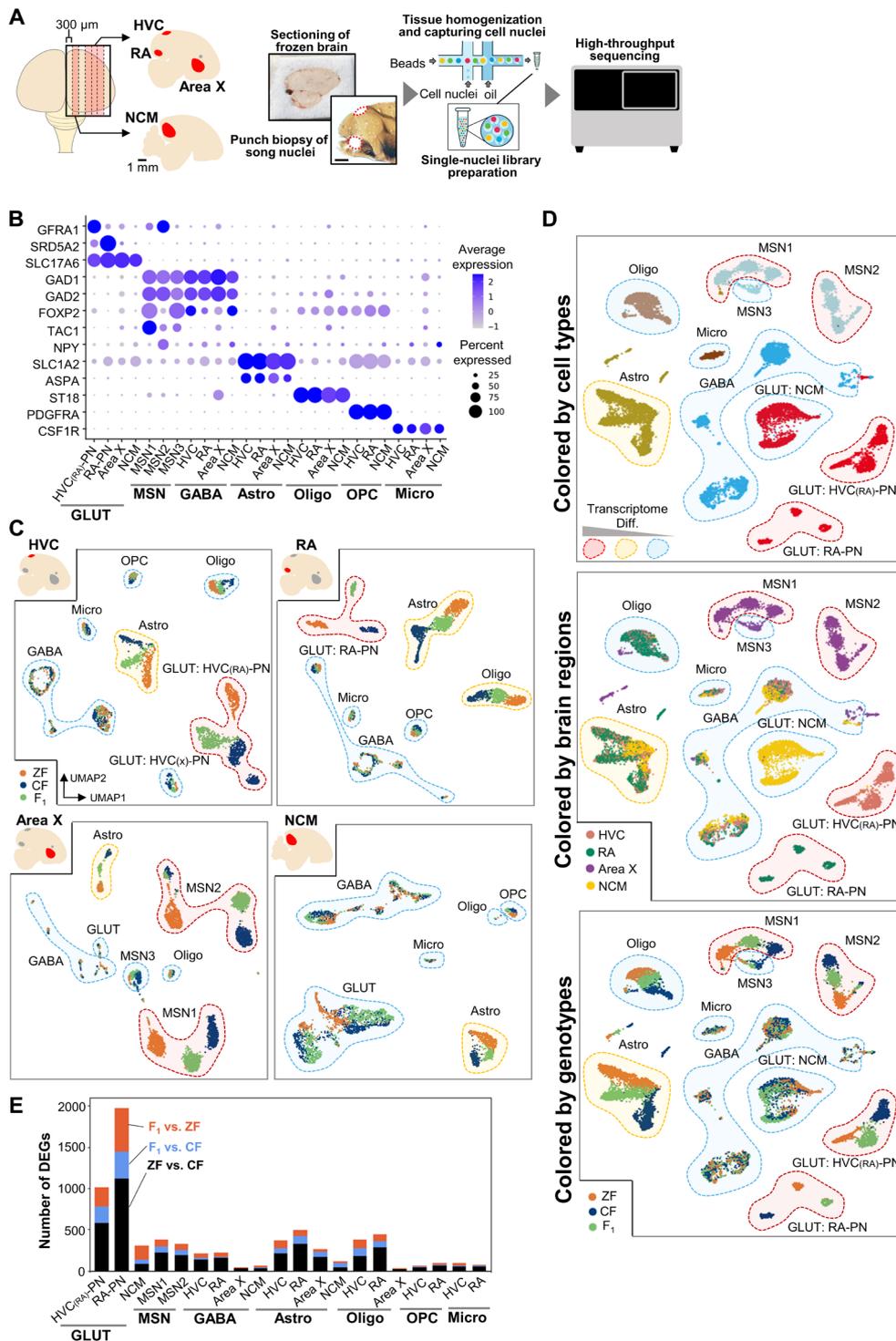


Fig. 5. Distinct transcriptional signatures in the song nuclei of F₁ hybrids. (A) Experimental procedure for single-nucleus RNA sequencing (snRNA-seq) of HVC, RA, Area X, and NCM. (B) Combination of marker gene expressions for identifying various cell types in HVC, RA, Area X, and NCM, including GLUT, PN, GABA, astrocytes (Astro), oligodendrocytes (Oligo), OPC, and microglial cells (Micro). (C) UMAP mapping of mRNA transcriptomes in HVC, RA, Area X, and NCM of ZFs, CFs, and F₁ hybrids. Dot colors represent the respective bird groups (orange for ZFs, navy for CFs, and green for F₁ hybrids). Blue-shaded clusters indicate similar cell distributions among the three bird groups, while orange-shaded clusters denote intermediate cell distribution patterns of F₁ hybrids between the parental species. Red-shaded clusters represent distinctly separated three cell clusters corresponding to ZFs, CFs, and F₁ hybrids. (D) Integrated UMAPs displaying all cells from HVC, RA, Area X, and NCM, with distinct colors indicating cell types (top), brain regions (middle), and species (bottom). (E) The number of DEGs between ZFs, CFs, and F₁ hybrids in each cell type of HVC, RA, Area X, and NCM.

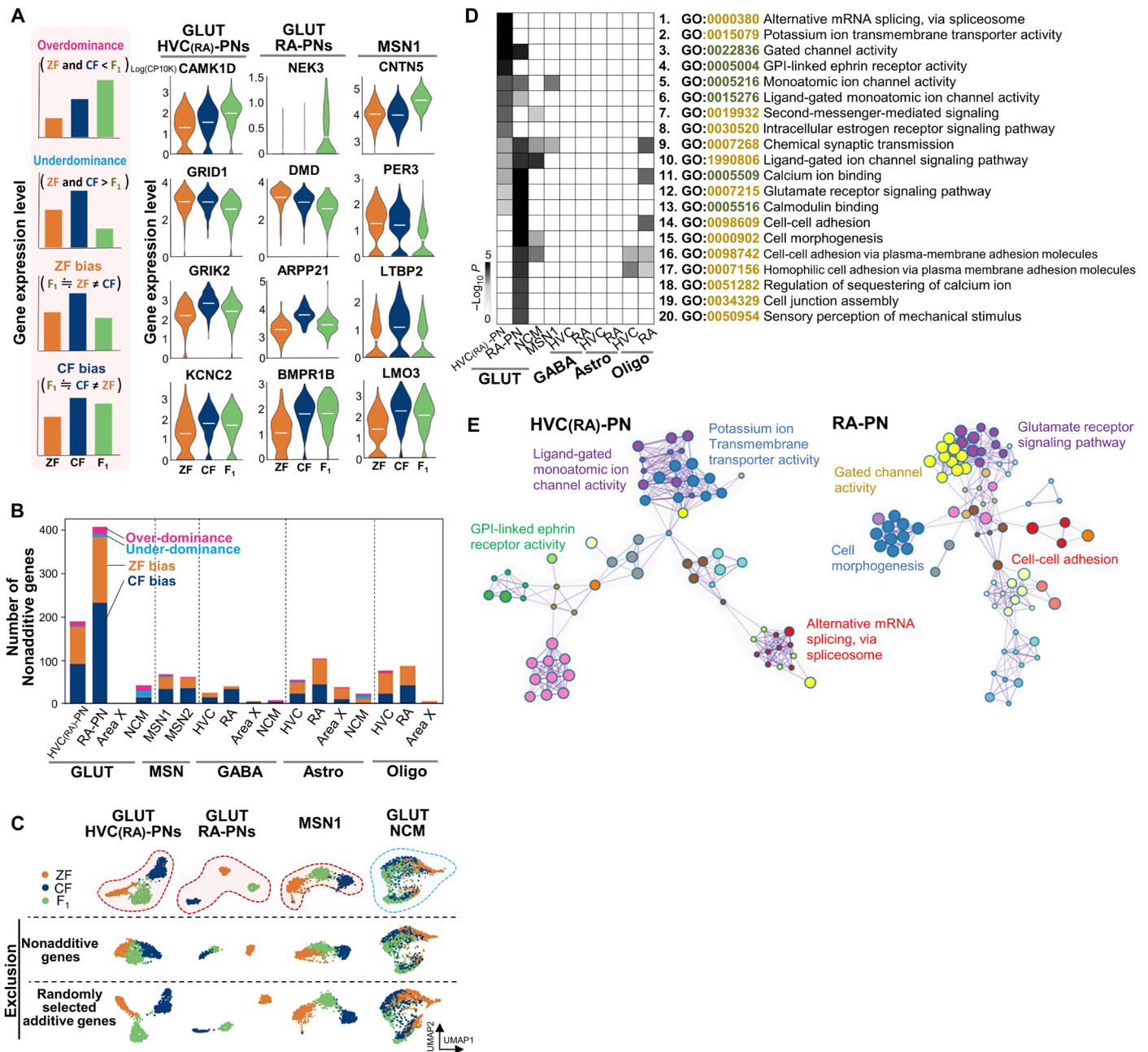


Fig. 6. Nonadditive genes enriched in GLUT projection neurons in the VMP of F₁ hybrids. (A) Examples of nonadditive genes in GLUT HVC(RA)-PNs, GLUT RA-PNs, and Area X MSN1 of F₁ hybrids. *CAMK1D*, calcium/calmodulin-dependent protein kinase ID; *GRID1*, glutamate ionotropic receptor delta type subunit 1; *GRIK2*, glutamate ionotropic receptor kainate type subunit 2; *KCNC2*, potassium voltage-gated channel subfamily C member 2; *NEK3*, NIMA related kinase 3; *DMD*, dystrophin; *ARPP21*, cAMP regulated phosphoprotein 21; *BMPR1B*, bone morphogenetic protein receptor type 1B; *CNTN5*, contactin 5; *PER3*, period circadian regulator 3; *LTBP2*, latent transforming growth factor Beta binding protein 2; *LMO3*, LIM domain only 3. The units of gene expression level: Counts per ten thousand [log (CP10K)]. (B) The number of nonadditive genes in each cell type in HVC, RA, Area X, and NCM. (C) UMAP distributions for GLUT HVC(RA)-PNs, GLUT RA-PNs, MSN1, and GLUT NCM neurons when gene expression values of the top 2000 highly variable genes are included (top); when nonadditive genes are excluded (middle); and when an equal number of random subset of additive genes are excluded (bottom). (D) Top 20 significantly enriched GO terms identified through GO enrichment analysis of nonadditive genes, along with their corresponding enrichment levels in each cell type. Green and brown colors represent the molecular function and biological process GO terms, respectively. (E) Similarity networks of enriched GO terms for GLUT HVC(RA)-PNs and GLUT RA-PNs. Each node represents an enriched term, colored according to a cluster identifier. Node size is proportional to the number of nonadditive genes in the terms.

Area X, as well as in GABA neurons, astrocytes, and oligodendrocytes in HVC and RA (Fig. 6D). In contrast, GLUT HVC_(RA)-PNs and RA-PNs emerged as the primary cell types accumulating non-additive genes with significantly enriched GO terms, such as “alternative splicing factor,” “ion channel activity,” “synaptic transmission,” “cell-cell adhesion,” and “signal transducers” (Fig. 6, D and E). This suggests that F₁ hybrids may have unique physiological properties in the GLUT-projecting neurons within the VMP, potentially leading to alterations in the generation and transmission of species-specific neural activity patterns.

In addition, we found that a certain number of the nonadditive genes composing the enriched GO terms exhibited significant correlations with the degree of song learning, in terms of syllable learning rate, syllable repertoire size, or both, across F₁ hybrids (8.5% of 82 genes and 9.1% of 164 genes in HVC_(RA)-PNs and RA-PNs, respectively) ($n = 7$ birds) (Fig. 7A). This occurrence rate exceeds that of correlated genes randomly selected in the same numbers from additive genes in each projecting neuron type (4.4 and 2.6% in HVC_(RA)-PNs and RA-PNs, respectively) (Fig. 7B). In summary, neural molecular changes potentially associated with the expanded learning capacity of the F₁ hybrids were represented as distinct cell type-specific transcriptional features, characterized by the expression of nonadditive genes in the song circuits.

DISCUSSION

While the importance of learned behaviors in evolutionary adaptation is widely recognized, the neurogenetic mechanisms underlying species-specific learning capacities remain unknown. In this study, we used automated passive song playback to tutor interspecies F₁

hybrid songbirds and found that the genetic heterogeneity in these F₁ hybrids could influence species-specific vocal learning traits. In prior research, we reported that, following tutoring with both parental species' songs, F₁ hybrids between ZF and OF develop a broad array of songs, ranging from ZF- to OF-like songs across individuals (27, 58). This varied song repertoire differed from what was acquired in F₁ hybrids crossing ZFs with CFs. The most substantial finding of the present study is that F₁ hybrids from ZFs and CFs pairings exhibited an expanded ability to learn songs from both parental species, as well as from genetically unrelated species. This suggests that their modified vocal mimicry skill is linked to a heterotic trait in vocal learning.

Evolutionary changes in species-specific song learning are thought to be driven by genetic and epigenetic changes that modify the learning constraints during song acquisition (8, 59, 60). In our study, using a tutoring scheme with songs from both parental species, we found that the F₁ hybrids do not exhibit a specific dominant bias toward the songs of either parental species. Instead, they learn and vocalize songs from both species, each characterized by distinct acoustic, temporal, and prosodic features. This expansion of song learning ability in the F₁ hybrids may result from altered learning constraints affecting both sensory and sensorimotor learning processes: the auditory sensory recognition and memorization of tutored songs, as well as the vocal-motor execution of these memorized songs. Such modifications would likely require alterations in the physiological properties of the neural circuits involved in song learning and production in the F₁ hybrids. Variations in gene expression levels and/or patterns in anatomically hardwired neural circuits, conserved across species, play a pivotal role in influencing the divergence of behavioral phenotypes (61–64). In line with this,

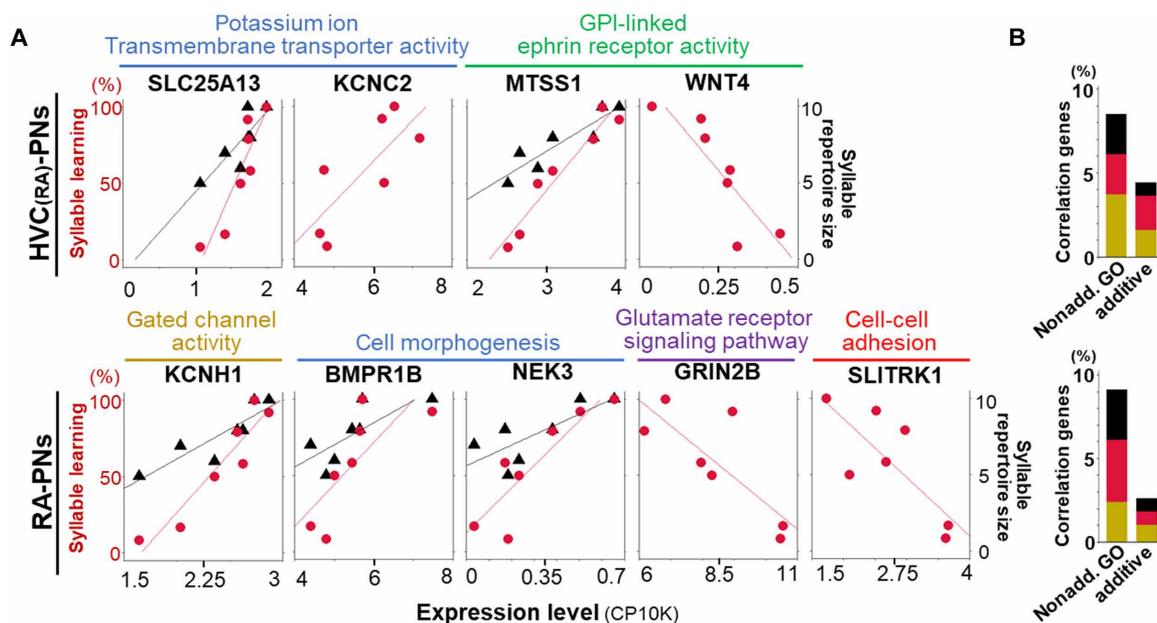


Fig. 7. Association between the expression of nonadditive genes and the degree of song learning in F₁ hybrid individuals. (A) Correlation between the expression of GO-enriched nonadditive genes and syllable learning rate (represented in red) and/or syllable repertoire size (represented in black) among F₁ hybrids ($n = 7$ birds). *SLC25A13*, solute carrier family 25 member 13; *MTSS1*, *MTSS1*-BAR domain containing 1; *WNT4*, *Wnt* family member 4; *KCNH1*, potassium voltage-gated channel subfamily H member 1; *GRIN2B*, glutamate ionotropic receptor NMDA type subunit 2B; *SLITRK1*, *SLIT* and *NTRK*-like family member 1 ($P < 0.05$ for all genes shown, Pearson's product moment correlation). (B) Comparison of the number of genes associated with song learning, either in syllable learning rate (red), syllable repertoire size (black), or both (brown), between nonadditive genes composing the enriched GO terms and randomly selected additive genes.

our snRNA-seq results from the vocal-motor and basal ganglia song nuclei of the F₁ hybrids revealed distinctive transcriptional signatures in the song nuclei, especially in GLUT-projecting neurons of HVC and RA, when compared to their parental species. This implies that the transcriptional changes in these projecting neurons in the neural circuits could influence species-specific learning constraints. Compared to the GLUT-projecting neurons in the VMP, fewer transcriptional differences were observed across all cell types in the NCM, a region contributing to higher auditory recognition and processing. However, the potential significance of transcriptional characteristics in auditory areas for the learning outcomes observed in the F₁ hybrids and species-specific learning capacities should not be overlooked. Although a certain number of NCM neurons exhibit varying neural activities in response to conspecific and heterospecific songs (51, 53, 65), only specific subsets of these neurons may play a role in shaping the memory of a song model to be learned (52). Sampling from multiple auditory regions, including field L (a primary cortical auditory region) and the caudomedial mesopallium (another higher auditory region), could be useful to examine whether some observed differences in song learning capacity are due to variations in the auditory system of F₁ hybrids. Further research is necessary to explore latent changes in physiological properties, cellular transcription, and microcircuitry connections in the focused areas of this study, as well as in other brain regions engaged in song learning and production throughout the song acquisition process.

The manipulation of specific gene expression through transgenic modification has been demonstrated to induce either learning deficits or enhancements in various animal species (66–68). For instance, transgenic manipulations that alter synaptic transmissions in the hippocampus of mice, such as knocking down the nociceptin receptor or overexpressing the *N*-methyl-D-aspartate-type glutamate receptor 2B, facilitate neural plasticity. This leads to long-term potentiation, subsequently enhancing learning in spatial memory and object recognition (68–70). Heterozygous loci in F₁ hybrids can induce phenotypic novelty, including heterosis, through de novo allelic interactions between the parental species (19, 20). The expression of nonadditive genes is considered a key factor in generating phenotypic novelty in F₁ hybrids across species (55–57). We identified a series of nonadditively expressed genes enriched in the neural excitation and signaling machinery of GLUT-projecting neurons in the vocal-motor nuclei and associated with the degree of song learning. However, it remains unclear whether such nonadditive genes causally facilitate motor learning capabilities beyond species-specific constraints. To gain deeper insight into the potential impact of the enriched expression of nonadditive genes on learning capacity, conducting neurophysiological studies that explore the intrinsic neural excitation properties and developmental changes in neural plasticity in these neurons using cell type-selective gene manipulation techniques is crucial. In addition, single-cell transcriptional analysis using juvenile brain tissues during the sensitive period of song learning is pivotal to understanding neural molecular basis, including the transcriptional regulation mechanisms underlying the expression of nonadditive genes, of heterosis in vocal learning as a phenotypic novelty.

While vocal mimicry has arisen repeatedly within the songbird clade (60), the underlying evolutionary mechanisms remain unknown. Vocal mimicry tends to be exhibited by songbirds that are continuous singers with large repertoires instead of being limited to

particular habitats, geographical regions, or mating systems (71, 72). Recent genome sequencing studies have revealed that interspecific hybridization is more prevalent than previously speculated and can contribute to introgression and speciation. Our investigation found that F₁ hybrids between ZFs and CFs exhibited altered learning capacities, enabling them to acquire larger syllable repertoires and learn genetically unrelated heterospecific songs, which are foundational traits for vocal mimicry (60, 72). Our findings suggest that interspecies hybridization could shape the modified learning capacity for species-specific learning constraints, allowing vocal mimicry during speciation in songbird species. Comparative genome analysis may reveal a genomic vestige of interspecies hybridization in some species having vocal mimicry ability compared to the closest species constrained to conspecific song learning.

Although the evolutionary significance of interspecies hybridization on learning and cognitive functions remains elusive, genetic incompatibilities in F₁ hybrids might lead to cognitive impairments (23). As a result, interspecies hybridization does not always positively influence the adaptation of cognitive and learning capacities in F₁ individuals (24, 73). Moreover, it is important to recognize that “expanded learning” in interspecies hybrids does not necessarily denote adaptive and improved functionality across all domains. In particular, species-specific constraints on song preference and learning are crucial for facilitating assortative mating with conspecifics and preventing mating attempts with heterospecifics. Therefore, the expanded learning flexibility observed in hybrids might be considered maladaptive, reflecting a more “leaky” or permissive song learning filter in terms of song learning and production. This does not serve as an effective behavioral signal for conspecific recognition and assortative mating. Nonetheless, the heterogenetic backgrounds resulting from interspecies hybridization, encompassing a variety of parental species combinations, can diversely affect learning and cognitive capabilities. This may affect hybrid fitness both positively and negatively, influenced by both natural and sexual selection. Further studies using F₁ hybrids provide a promising research avenue to explore the biological significance of altered learning phenotypes that arise by modifying learning constraints.

MATERIALS AND METHODS

Animals

ZF (*T. guttata*) and CF (*A. modesta*) were bred from our breeding colony established at Hokkaido University to obtain conspecific and F₁ hybrid offspring. Because of the inherent challenges associated with obtaining a sufficient number of F₁ hybrids derived from CF♀ and ZF♂, we exclusively used F₁ hybrids crossing ZF♀ with CF♂ for comprehensive analysis. The photoperiod was maintained at a 13-hour light/11-hour dark cycle, with free access to food and water. Polymerase chain reaction (PCR)-based genotyping techniques were used to determine the sex of the juvenile birds. All experimental procedures were conducted under the guidelines and approval of the Committee on Animal Experiments of Hokkaido University (#18-0053 and 23-0110). These guidelines adhere to the national regulations for animal welfare in Japan.

Song tutoring and recording

Juvenile song tutoring was conducted according to previously described methods (27, 36, 74). Passive song playback was implemented

within sound attenuation boxes using recorded songs. To record songs, a microphone (XM8500-ULTRAVOICE; Behringer) was suspended near a perch from the top of the cage and connected to a sound amplifier (US-16-08; TASCAM). Songs were recorded at a sampling rate of 44 kHz and a 16-bit amplitude resolution using Sound Analysis Pro (version SAP2011.211; <http://soundanalysispro.com/>). By post-hatching day (phd) 15, when hatchlings remained in the nest, they were separated from their male parents to prevent exposure to paternal songs. This is a standard procedure to extend the sensitive period (75, 76). Subsequently, the juveniles were raised solely by their mothers. After fledging, mothers and other siblings were removed, and male juveniles were individually housed in sound-attenuation boxes with acoustic absorbents (box size, 65 cm by 27 cm by 30 cm; cage size, 54 cm by 22 cm by 23 cm) (74, 77).

Forty-two birds (ZF = 12, CF = 9, and F₁ = 21 birds) were used for song tutoring with playback of recorded songs of the two parental species (i.e., both ZF and CF songs). In addition, nine F₁ hybrids were tutored using songs from genetically unrelated species (three by OF songs, four by BF songs, and two by CN songs). Similarly, six ZFs and three CFs ($n = 2$ each and 1 each, respectively) were tutored using OF, BF, and CN songs. Song playback sessions commenced at 34 ± 7 phd (means \pm SD) and continued until adulthood (>147 phd). Because of the lack of studies testing the duration of sensory and sensorimotor learning phases of CFs, we based the timeline of song playback on past song tutoring experiments conducted with ZFs, BFs, and OFs (27, 36, 74, 77). When multiple male chicks were obtained from the same clutch, we allocated individual male chicks to receive tutoring with different types/species songs once they reached fledging age (typically around 30 phd). Fourteen song playback sessions were scheduled daily, with seven sessions occurring during the morning (8 a.m. to 12 p.m.) and afternoon (1 p.m. to 6 p.m.) (27, 74, 77). A song file, randomly chosen from a collection of four stocked files with a total duration of up to 15 s, was played through a speaker (MM-SPL11UBK; SANWA) at a volume of 55 to 75 dB. The timing of song playback was controlled using Sound Analysis Pro, with a playback probability of 0.0025/s and intervals exceeding 20 s.

During tutoring with the songs of the two parental species (ZF and CF), an interval of silence ranging from 2.3 to 4.2 s was interjected between renditions of ZF and CF songs within a single stock file (fig. S2). The ZF and CF song models used for playback were originally recorded from a single ZF and CF male within our breeding colony, representing typical ZF and CF song vocalizations. For juvenile song tutoring using songs of genetically unrelated species (OF, BF, and CN), each recorded song file was randomly broadcasted from a collection of four to five prepared song files (fig. S3). In the case of non-song tutoring, male F₁ juveniles were individually housed without exposure to songs within a sound attenuation box from fledging until adulthood (>150 phd) ($n = 6$).

Analysis of syllable acoustics

An individual syllable was defined as a vocal element within a song, characterized by a silent interval exceeding the segmentation threshold (set as one-third times the median of 500 silent gap durations). Repetitive introductory notes were excluded from the analysis. High- and low-frequency background noise (below 0.50 and above 18.0 kHz) was removed by the SASlab (version 5.3.00) “Frequency Domain Transformation” tool. The Audacity “noise reduction” tool was used for noise reduction in the middle-frequency range. The

song files were segmented into individual syllables using SASlab batch processing, labeling, and segmentation.

To analyze species differences in syllable acoustics between ZFs and CFs, adult songs from 10 birds of each species were used. Fifteen acoustic parameters were analyzed, including the following: unique syllable number (syllable repertoire size), syllable duration, inter-syllable gap duration, the ratio of syllable duration to inter-syllable gap duration, mean pitch, mean frequency modulation (mean FM), the square of mean amplitude modulation (mean AM²), mean entropy, mean pitch goodness, mean frequency, variance frequency modulation, variance entropy, variance pitch goodness, variance mean frequency, and variance amplitude modulation (variance AM). Syllable identification was performed on randomly selected song recordings from each bird using Avisoft SAS Lab Pro or Song browser. The identified syllables were saved as separate WAV files, and 500 syllables per individual were used for the analysis (27, 77). Sound Analysis Pro was used for calculating acoustic parameters, with the exception of inter-syllable gap duration and the ratio of syllable duration to inter-syllable gap duration, which were manually calculated. The median value of 500 syllables was used as a representative value for each bird across the acoustic parameters.

Visualization of song sequential property

The syllable similarity matrix (SSM) method was used (Fig. 3D) (27, 36, 77). Briefly, spectrograms of two different songs from a bird were aligned along with columns and rows of a matrix. Two sets of song renditions, each comprising a minimum of 50 syllables, were prepared. In cases where a song bout contained fewer than 50 syllables, two or more song bouts produced by the same individual bird were merged to create a combined rendition with over 50 syllables. For all column versus row syllable combinations, the similarity of the syllable spectrogram was calculated as the peak correlation coefficient by the round-robin comparison using the Avisoft CORRELATOR application. Similarity score (Φ_{XY}) between syllables a and b was calculated according to the following formula:

$$\Phi_{XY} = \frac{\sum_X \sum_Y [(a_{xy} - m_a) \times (b_{xy} - m_b)]}{\sqrt{\sum_X \sum_Y (a_{xy} - m_a)^2 \times \sum_X \sum_Y (b_{xy} - m_b)^2}}$$

where m_a and m_b are the mean values of the spectrograms a and b , respectively. a_{xy} and b_{xy} denote the intensities of the spectrogram points at the locations x and y , respectively. $\Phi_{XY} = 1$ indicates that syllables a and b are identical, whereas $\Phi_{XY} = 0$ means that the two syllables are not similar at all. After generating SSM on a spreadsheet, the matrix was converted into a binarized pattern consisting of black and white cells. The similarity score threshold of 0.6 was applied (36), wherein cells with a high similarity score (≥ 0.6) were colored black, while those with a low similarity score (< 0.6) were colored white. The prevalence of characteristic patterns within the binarized “2 rows \times 2 columns” cells in the SSMs was quantitatively analyzed to examine syllable temporal structures. The R software program was used to identify the most similar binarized pattern for each 2 \times 2 cell in the SSM from a total of 12 possible patterns. Subsequently, the percentage of the following three transition types was calculated. Syllable transition type I referred to as a “paired-syllables transition,” indicating the presence of two consecutive dissimilar syllables with identical sequential order in a pair of songs (e.g., “song bout 1 [...A B.....] vs. song bout 2 [...A B.....]”). Syllable transition

type II encompassed the “repetitive-syllables transition,” occurring when two successive identical or highly similar syllables appeared in two songs (e.g., “song bout 1 [.....A A..] vs. song bout 2 [...A A.....]”). A and B represented distinct syllables in the given examples.

Unsupervised identification and labeling of unique syllables

The unsupervised clustering method of T-distributed Stochastic Neighbor Embedding (t-SNE) was used to identify unique syllable types in each bird, 500 to 1384 syllables per bird used to plot approximately 100 syllables per syllable cluster. This aimed to represent a unique syllable cluster while maintaining clear separations from other clusters. Nearly twice the number of syllables (about 1000) were required for analysis in the F₁ hybrids that learned syllables from both ZF and CF tutor songs compared to ZF and CF pupils. The syllables were sampled from multiple song bouts randomly selected from a day during the adult stage (phd 145 to 184 for ZFs, phd 150 to 215 for CFs, and phd 130 to 181 for F₁ hybrids). For each syllable, 12 syllable acoustic parameters were measured, including syllable duration, mean pitch, mean frequency modulation, mean amplitude modulation, mean entropy, mean pitch goodness, mean frequency, variance frequency, variance entropy, variance pitch goodness, variance mean frequency, and variance amplitude modulation. Sound Analysis Pro was used to obtain these measurements. The resulting data files, containing the values of 12 acoustic parameters, were used for generating t-SNE plots. The cluster boundary was determined on the basis of the density-based spatial clustering of application with noise. Syllables that could not be assigned to any cluster or belonged to small clusters comprising fewer than 30 syllables (i.e., representing a less than 3 to 6% rate of the examined 500 to 1000 syllables) were excluded from further evaluation because these syllables could be considered very rare across song bouts. The remaining clusters in the t-SNE plots were identified as unique syllable types.

Evaluation of syllable learning achievement

Unique syllables from both tutored model songs (using all syllables) and songs produced by pupil birds (10 syllables per each unique syllable type) were used (77). In the model tutor songs of the parental species, the ZF and CF songs, there were six and four distinct syllable types, respectively. The similarity score for each tutor-pupil syllable pairing was calculated using the previously described SSM methods. The 75th percentile values of the similarity scores for each syllable type pairing—comparing 10 spectrograms of an identical syllable type in a pupil song to all syllables of a specific type in the tutor songs—were used as the “representative similarity score.” If the representative similarity scores exceeded 0.6, the syllable type was classified as a “learned syllable.” Conversely, if the representative value was below 0.6, such a syllable of pupils was designated as a “novel (not learned)” syllable. If the similarity score between a pupil syllable exceeded 0.6 for multiple tutor syllables, the highest similarity score was used to determine which tutor syllable was most similar.

The proficiency of syllables learning for each bird was determined using the following formula:

$$\text{Syllable learning (\%)} = \frac{[ZF(n) / ZF(N)] + [CF(n) / CF(N)]}{2} \times 100$$

where $ZF(n)$ and $CF(n)$ represent the numbers of learned syllables from the tutored songs of ZFs or CFs, respectively. $ZF(N)$ and $CF(N)$ denote the total number of unique syllable types contained in tutor

songs of ZFs and CFs [in the case of tutoring from both parental species’ songs, $ZF(N)$ was set at 6, and $CF(N)$ was set at 4]. The dotted line at 50% indicates cases where birds may learn syllables exclusively from one parental species, or alternatively, a mixture of syllables from both species.

Evaluation of the quality of heterospecific song mimicry

To evaluate the comprehensive quality of heterospecific song mimicry, syllable acoustics, syllable sequence patterns, and song-bout duration were compared between playback tutor songs and pupil songs. To assess syllable acoustic similarity, the Euclidean distance of syllable acoustics between pupils’ songs and their corresponding tutor songs (Δ syllable acoustics) was calculated, selecting an equal number of syllables from each pupil and the corresponding tutor songs (110 to 250 syllables, depending on the abundance of playback tutor song syllables). This Euclidean distance, calculated on the basis of 12 acoustic parameters also used for t-SNE plots, indicates higher acoustic similarity with smaller values. Syllable sequence pattern similarity involves comparing the proportion of syllable transition types I and II (indicating motif and repetitive, respectively) in 50×50 syllable SSMs of pupil and tutor songs. The differences in syllable sequence patterns were quantified as “ Δ syllable sequence” using the formula: Δ Syllable sequence = $(|tutor\ P1 - pupil\ P1| + |tutor\ P2 - pupil\ P2|)/2$, where “tutor P1” and “pupil P1” represent the proportion of syllable transition pattern type I in the SSM of tutor and pupil songs, respectively. Song-bout duration differences (Δ song bout duration) were assessed by calculating the average duration from 10 randomly selected song bouts for each pupil and using all available playback tutor song files for comparison. Differences in song-bout duration were calculated using the formula: Δ song bout duration = $|(\text{average tutor bout duration}) - (\text{average pupil bout duration})| / (\text{average tutor bout duration})$. The quality of heterospecific song mimicry among ZF, CF, and F₁ pupils was visualized on a three-dimensional plot of these three measurements. In this plot, a smaller distance from the origin (0, 0, 0) indicates a higher comprehensive similarity between tutor and pupil songs.

Evaluation of song nuclei size and constituent neural cell density

To evaluate song nuclei size, in situ hybridization was performed using genes exhibiting different expression patterns within and outside the song nuclei (HVC, RA, LMAN, Area X, and DLM). Twenty-six adult birds were used, comprising 7 ZFs, 8 CFs, and 11 F₁ hybrids. Brain sampling was performed after placing the bird in a sound-attenuating box overnight under silence, ensuring silent and dark conditions, and the samples were stocked at -80°C until use. Fresh-frozen brain slices were sectioned sagittally with a thickness of 12 μm from a hemisphere of each bird’s brain and mounted on glass slides. A total of eight brain sections, each separated approximately by 400 μm , were used as a set per probe/antibody for in situ hybridization and immunohistochemistry.

Four genes were used as in situ hybridization probes: *androgen receptor (AR)* for labeling HVC and LMAN (43), *glutamate receptor ionotropic AMPA subunit 1 (GRIA1)* for labeling Area X, *glutamate receptor metabotropic type 2 (GRM2)* to delineate lower GRM2 expression in DLM than the surrounding thalamic region, and *glutamate receptor ionotropic kainate type subunit 1 (GRIK1)* for labeling RA (42). The cDNA fragments of AR [1032 base pairs (bp)], GRIA1 (1340 bp), GRIK1 (500 bp), and GRM2 (1782 bp) were PCR-amplified

using oligo-primers from pGEM-T easy cloning vectors. The amplified cDNA fragments were purified and used as templates for synthesizing antisense ^{35}S -riboprobes through in vitro transcription using SP6 or T7 RNA polymerase. During the prehybridization step, frozen brain sections were fixed in 4% paraformaldehyde (PFA)/1× phosphate-buffered saline (PBS), washed in 1× PBS, acetylated, dehydrated in an ascending ethanol series, and air-dried. A total of 1×10^6 cpm of the ^{35}S -riboprobe was added into 140 μl of hybridization solution [50% formamide, 10% dextran, 1× Denhardt's solution, 12 mM EDTA (pH 8.0), 10 mM tris-HCl (pH 8.0), 300 mM NaCl, yeast tRNA (0.5 mg/ml), 10 mM dithiothreitol]. Hybridization was carried out at 65°C for 14 hours in an oil bath. Subsequently, the slides were washed in chloroform, 2× saline sodium phosphate EDTA (SSPE), 50% formamide in 2× SSPE, 0.1% SSPE, and ascending ethanol series. Following drying, the slides were exposed to x-ray film (Biomax MR, Kodak) for appropriate intervals and developed.

To count the number of neural cells within the song nuclei, immunohistochemistry using anti-NeuN antibody was performed on the same set of 26 birds used for in situ hybridization. Frozen sections mounted on slide glasses were fixed in 4% PFA/1× PBS for 10 min and washed in 1× PBS. Slide glasses were incubated with anti-NeuN antibody (dilution 1:800, Millipore, no. MAB377) in blocking solution (0.3% Triton X-100/4% normal goat serum/1% bovine serum albumin/1× PBS) at 4°C for 16 hours. The primary antibodies were detected using anti-mouse Alexa Fluor 555-conjugated antibodies (Thermo Fisher Scientific, no. A-21422). Last, the slides were mounted with Prolong gold with 4',6-diamidino-2-phenylindole (DAPI, Invitrogen).

Each sample from 26 birds used for in situ hybridization and immunohistochemistry was randomly labeled with letters A to Z, without any species or bird identification information, to ensure blinding. The area size of each song nucleus was quantified by measuring the summed area (square millimeter) in eight brain sections collected at 400- μm intervals, spanning from the lateral to medial direction. Brain images exposed in x-ray films of in situ hybridization were captured using a stereo microscope with a charge-coupled device camera (Leica Z16 AP0 and Leica DFC490). The boundaries of song nuclei were traced in the digital images, and the sizes of those areas were measured using Adobe Photoshop software. For neural cell counting, fluorescent signals from the anti-NeuN antibody were acquired using a fluorescence microscope (Keyence, BZ-X810). The NeuN signals, along with cellular nuclei stained with DAPI, were counted within sampling windows measuring 150 by 150 μm^2 (for HVC, RA, and Area X) or 100 by 100 μm^2 (for LMAN and DLM). These counts were used to assess neural cell density (per square millimeter).

Signal amplification by exchange reaction-fluorescent in situ hybridization

To assess the ratio of excitatory GLUT neurons and inhibitory GABA neurons in song nuclei, signal amplification by exchange reaction-fluorescent in situ hybridization (SABER-FISH) was conducted following the protocol provided at <http://saber.fish/>. A total of six birds from each of ZFs, CFs, and F₁ hybrids that were used to compare song nuclei size and neuron density were also used to calculate the ratio between excitatory and inhibitory neurons. The labeling of GLUT and GABA neurons was performed using SABER-FISH. mRNA sequences for *SLC17A6* (*vGlut*, GLUT neuron marker)

and *GAD1* (GABA neuron marker) were obtained from the National Center for Biotechnology Information database, and probes were designed using the OligoMiner pipeline. Fresh-frozen brain sections on slide glasses were fixed in cold 4% PFA. The slides were washed in 1× PBS three times at room temperature, followed by washing in buffer 1 at 45°C. The first original probe of *SLC17A6* and *GAD1* was hybridized in hybridization solution 1 (1 μg per gene per slide) overnight at 42°C. After the first hybridization, the slides were washed in wash buffer 1 at 42°C, 2× SSCT (2×SSC, 0.1% Tween 20), and wash buffer 2. The second branching probe was hybridized in the hybridization solution 2 (1 μg per glass) for 6 hours to overnight at 37°C. After the second hybridization, the slides were washed in wash buffer 2, 2× SSCT, and 1× PBST (1× PBS, 0.1% Tween 20) at 37°C. The fluorescent primer was then hybridized in a fluorescence solution containing Hoechst to stain cell nuclei for 1 hour to overnight at 37°C. Subsequently, the slides were washed in 1× PBST and 1× PBS and treated with an autofluorescence quenching solution (Vector Trueview). Last, the slides were washed in 1× PBS and mounted. Single-plane images of cell distribution for each song nucleus were obtained using a fluorescence microscope (Keyence, BZ-X810) at a magnification of $\times 40$. Cells exhibiting signals from both SABER-FISH probes and DAPI-stained cell nuclei were counted using the counting tool in Adobe Photoshop. Sampling windows of 150 by 150 μm^2 for HVC, RA, and Area X, and 100 by 100 μm^2 for LMAN and DLM were set per section. Images from two sections for HVC/RA/LMAN/Area X were captured as counting windows, and the average cell number from these counting windows for each song nucleus was calculated. For DLM, because of its relatively small size, cell counting was performed on one section.

Single-nuclei RNA sequencing

For snRNA-seq of HVC and RA, adult males of F₁ hybrids crossed ZF female with CF male ($n = 7$, 162 to 369 phd), ZFs ($n = 3$, 320 to 814 phd), and CFs ($n = 4$, 202 to 267 phd) were used in total. For Area X snRNA-seq, we used two birds for F₁ hybrids, ZFs, and CFs, which were used for HVC and RA snRNA-seq. Except for one ZF (phd 145), all birds were the same individuals used for HVC and RA sampling. For NCM snRNA-seq, two F₁ hybrids, one ZF, and two CFs were used. The sampled F₁ hybrids and CFs were tutored with the playback of recorded songs of the two parental species (i.e., both ZF and CF songs), while ZFs were tutored with playback of conspecific ZF songs (refer to fig. S6 for the effect of auditory inputs on species-specific cellular distributions in UMAP). For brain sampling, the bird was placed in a sound-attenuating box overnight under silent and dark conditions. Before light onset, brain tissues were sampled at non-singing conditions the next morning and embedded in OCT Compound (Sakura Finetek Japan). Frozen brain sections were cut at a thickness of 300 μm in the sagittal plane using a cryostat microtome (Leica Biosystems). HVC, RA, and Area X tissues were punched out with a Rapid-Core sampling tool (0.5- to 1.2-mm diameter; EMS) and stored at -80°C until nuclei isolation. For sampling NCM tissues, a brain section was collected with a thickness of 300 μm located 200 to 500 μm from the midline. A punch biopsy with a diameter of 1.2 mm was performed at the center of the caudal part of the nidopallium, covering approximately 40 to 60% of the NCM region in the brain section.

The punched tissue samples were separated into seven tubes to make 12 snRNA-seq libraries [Library #1: F₁-HVC I ($n = 4$); #2: F₁-RA I ($n = 4$); #3: CF-HVC I ($n = 3$); #4 CF-RA I ($n = 2$); #5 CF-RA

II ($n = 1$); #6: ZF-HVC ($n = 3$); #7: ZF-RA ($n = 3$); #8: F₁-HVC II ($n = 3$); #9 F₁-RA II ($n = 3$); #10: Area X ($n = 6$); #11: F₁-NCM ($n = 2$); and #12: ZF/CF-NCM (ZF = 1 and CF = 2)] based on brain region (HVC, RA, Area X, and NCM) and sampling timings. Punched tissues were homogenized in 750 μ l of ice-cold Nuclei PURE lysis buffer by 40 to 60 strokes of a glass Dounce tissue grinder (Wheaton), incubated for 5 min on ice, and transferred into a 2-ml centrifuge tube. Dounce tissue grinder was washed with 300 μ l of lysis buffer three times to wash off the remaining cell nuclei, and the wash solution was added to the homogenate in a 2-ml tube. The homogenate sample was pipetted 10 times with 1-ml micropipettes for further separation of cell nuclei and centrifuged at 400g for 10 min at 4°C. After removing the supernatant, the pellet was resuspended in 1 ml of nuclei wash buffer by pipetting five times with wide-bore 1-ml tip and then 15 times with normal 1-ml tip. The suspension was centrifuged at 400g for 10 min at 4°C to pellet cell nuclei again to remove the wash buffer as supernatant. Cell nuclei pellets were finally suspended in 100 μ l of nuclei wash and resuspension buffers containing DAPI and then filtered with 40- μ m cell strainers.

Isolated cell nuclei were sorted using DAPI fluorescence with a cell sorter (SH800, Sony). Following the manufacturer's protocol, 10 \times Chromium libraries were prepared using Chromium NEXT GEM single-cell library kit v3.1 (PN-1000269, 10x Genomics). The cDNA with cell barcode identifiers was PCR-amplified, and sequencing libraries were prepared. The constructed library was sequenced on MGI DNBSEQ-G400 (150 bp paired-end) or Illumina Nova-Seq6000 (28/91PE) platform. The Cell Ranger Software Suite (v6.0.0) was used for sample demultiplexing, barcode processing, single-cell 3' unique molecular identifier counting, and mapping on the reference genome. The ZF genome (bTaeGut1.4.pri, GCF_003957565.2) was used as a mapping reference genome. For identifying transcripts from each bird, the Souporecell singularity container was used for individual cell demultiplexing. Cell indexes were clustered on the basis of single-nucleotide polymorphism (SNP) information to match the number of individuals mixed in the snRNA-seq libraries using "souporecell_pipeline.py". Individual-specific SNPs were identified using the vcf file information in the Souporecell output and a custom Perl script. For library#12: ZF/CF-NCM, species-specific SNPs were used to identify the genotypes of transcripts. To determine which cluster of cell index belonged to which individual, the genomic DNA of the sample birds was sequenced with Sanger sequencing to check for individual-specific SNPs.

Cell cluster analysis

The R package Seurat 4.2.0 was used for data filtering and analyses. The mapped data were filtered using the "CreateSeuratObject" function with $\text{min.cells} = 3$ and $\text{min.features} = 200$. Cells containing multiple nuclei in one droplet were removed on the basis of the souporecell output. The filtered gene-barcode matrix was then normalized using the "LogNormalize" method with default parameters, and the top 2000 variable genes were identified using the "vst" method in the Seurat FindVariableFeatures function. Subsequently, the data were scaled for each gene, shifting the mean expression across cells to 0 and adjusting the expression variance among cells to 1. Principal components analysis (PCA) was performed using the top 2000 variable genes. To visualize the data distribution, UMAP was performed on the top 15 to 50 principal components to display the cells. Graph-based clustering was performed on the PCA-reduced

data using the "FindNeighbors" function (with top 50 principal components) and "FindClusters" function (with a resolution of 0.5) to differentiate between different cell types.

To identify the cell type of each cell cluster, the expression of established marker genes was used (Fig. 5B) (58, 78, 79). The marker genes used were as follows: *GFRA1* for HVC GLUT neurons projecting to RA [HVC_(RA) neuron]; *SRD5A2* for HVC GLUT neurons projecting to Area X [HVC_(X) neuron]; *ADCYAP1* for GLUT neurons in arcopallium surrounding RA; *GAD1* and *GAD2* for GABA neurons; *SLC1A2* and *ASPA* for astrocytes; *PDGFRA* for OPCs; *ST18* for oligodendrocytes; *CSF1R* for microglia; and *SRD5A2* and *SCUBE1* for the GLUT neurons in RA projecting to the nucleus of cranial nerve XII. MSN in Area X were identified on the basis of another previous report (58), *FOXP2*, *TAC1*, and *NPY* to identify and classify three clusters of MSN. Clusters without specific marker genes were filtered out as unknown clusters. Subsequently, to integrate data from multiple libraries and account for differences in cell number among bird groups, Seurat objects from multiple libraries were combined using "IntegrateData" (anchor weighting dimension: 50) with 5000 anchor genes identified by "FindIntegrationAnchors" (Canonical correlation analysis dimension to specify the neighbor search space: top 50). After data integration, 15,958 in HVC, 20,129 cells in RA, 5509 cells in Area X, and 6943 cells in NCM were obtained for further analysis. To ensure equal representation among species groups and avoid the influence of cell number differences, we randomly selected 1200 cells per species per song nuclei for generating UMAPs. For categorizing cell cluster distributions in UMAPs (as "similar," "different," or "intermediate" between the ZFs, CFs, and F₁ hybrids), the genotypes of 10 neighboring cells for each cell were first examined, considering the UMAP 1 and 2 axes (Fig. 5C). The percentage of cells that share the same genotype with all 10 neighboring cells was then calculated for each cell type. We then conveniently classified the cell types as having different distributions when more than 85% of the cells were located with all 10 neighbors of the same genotype. Conversely, if less than 50% of the cells shared the same genotype with their 10 neighbors, such cell types were classified as similar distributions. An intermediate distribution was defined as any categorization not meeting the above criteria.

DEGs between two groups were identified using "Findmarkers." According to the population size, equal number of cells (40 to 150 cells per group) was randomly selected from each species group (ZF, CF, and F₁) for each cell type. Cell types with fewer than 40 cells due to a shortage of successfully sequenced cells were excluded from this analysis, as they had insufficient cell numbers in each group. Significance was determined using the Wilcoxon rank-sum test (default setting) in the Findmarkers function, and *P* values were adjusted by Bonferroni correction using all genes in the dataset.

Nonadditively expressed genes were defined as genes whose expression level significantly differed from the mean expression level of the two parental species groups. In this study, we used FindMarkers function in Seurat (with a minimum FC difference setting of 0 and a minimum fraction of cells detected in either population of 0) to identify genes whose expression level significantly differed from at least one of the parental species as nonadditively expressed genes. The analysis of nonadditively expressed genes was performed for each cell type independently. Furthermore, nonadditively expressed genes were classified into four types based on the relationship in gene expression levels among the F₁ hybrid and the two parental

species: Type I represented overdominance ($F_1 > ZF$ and CF), type II represented underdominance ($F_1 < ZF$ and CF), type III represented ZF bias (significant difference in gene expression level between F_1 and CF but not F_1 and ZF), and type IV represented CF bias (the opposite case to type III).

GO analysis was performed to examine gene enrichment of nonadditively expressed genes in each cell type using Metascape v3.5.20230501 (<https://metascape.org/gp/index.html#/main/step1>). All the input entrez gene IDs were converted to human entrez gene ID to perform enrichment analysis.

For the correlation analysis between gene expression levels and both syllable learning rate and syllable repertoire size, seven F_1 hybrids with varied syllable learning rates and repertoire sizes were examined. Gene expression levels in each individual were represented as average values among 95 cells per bird of $HVC_{(RA)}$ -PNs or 71 cells per bird of RA-PNs, calculated using the “AverageExpression” function. The correlation between the average expression level of nonadditive genes and both syllable learning rate and syllable repertoire size was examined by Pearson’s product-moment correlation in R.

Statistical analysis

In song acoustic analysis, the significance of the difference in syllable learning achievement and syllable repertoire size among ZFs, CFs, and the F_1 hybrids was assessed using the Steel-Dwass test. The correlation between the number of learned syllables and syllable repertoire size was evaluated on the basis of Spearman’s rank correlation. The significance of the acoustic differences between ZFs and CFs was examined using the exact Wilcoxon rank-sum test. Data on neuron numbers and the ratio of GLUT to GABA neurons in song nuclei were compared using the Steel-Dwass test. For all these cases, the appropriate statistical test was determined after conducting the Shapiro-Wilk test to examine data normality. snRNA-seq data were analyzed using Seurat 4.2.0, including statistic tests. Specifically, to examine significantly different gene expression levels between the two groups, the Wilcoxon rank-sum test and Bonferroni correction were applied. The correlation between gene expression levels and either syllable learning achievement or syllable repertoire size was analyzed using Pearson’s product-moment correlation coefficient. GO enrichment analysis for nonadditively expressed genes was performed with Metascape v3.5. The method for P value calculation is explained at <https://metascape.org/blog/?p=122>.

Supplementary Materials

This PDF file includes:

Figs. S1 to S6

Legends for movies S1 and S2

Other Supplementary Material for this manuscript includes the following:

Movies S1 and S2

REFERENCES AND NOTES

1. A. J. Tierney, The evolution of learned and innate behavior: Contributions from genetics and neurobiology to a theory of behavioral evolution. *Anim. Learn. Behav.* **14**, 339–348 (1986).
2. B. R. Moore, The evolution of learning. *Biol. Rev. Camb. Philos. Soc.* **79**, 301–335 (2004).
3. R. C. Berwick, A. D. Friederici, N. Chomsky, J. J. Bolhuis, Evolution, brain, and the nature of language. *Trends Cogn. Sci.* **17**, 89–98 (2013).
4. P. S. Katz, R. M. Harris-Warrick, The evolution of neuronal circuits underlying species-specific behavior. *Curr. Opin. Neurobiol.* **9**, 628–633 (1999).
5. T. J. Gardner, F. Naef, F. Nottebohm, Freedom and rules: The acquisition and reprogramming of a bird’s learned song. *Science* **308**, 1046–1049 (2005).
6. P. Price, Developmental determinants of structure in zebra finch song. *J. Comp. Physiol. Psychol.* **93**, 260–277 (1979).
7. P. Marler, V. Sherman, Song structure without auditory feedback: Emendations of the auditory template hypothesis. *J. Neurosci.* **3**, 517–531 (1983).
8. E. A. Brenowitz, M. D. Beecher, Song learning in birds: Diversity and plasticity, opportunities and challenges. *Trends Neurosci.* **28**, 127–132 (2005).
9. J. Krebs, R. Ashcroft, M. Webber, Song repertoires and territory defence in the great tit. *Nature* **271**, 539–542 (1978).
10. S. Nowicki, W. A. Searcy, Song function and the evolution of female preferences: Why birds sing, why brains matter. *Ann. N. Y. Acad. Sci.* **1016**, 704–723 (2004).
11. M. Konishi, The role of auditory feedback in the control of vocalization in the white-crowned sparrow. *Z. Tierpsychol.* **22**, 770–783 (1965).
12. A. J. Doupe, P. K. Kuhl, Birdsong and human speech: Common themes and mechanisms. *Annu. Rev. Neurosci.* **22**, 567–631 (1999).
13. C. Mori, K. Wada, Audition-independent vocal crystallization associated with intrinsic developmental gene expression dynamics. *J. Neurosci.* **35**, 878–889 (2015).
14. M. Konishi, Effects of deafening on song development in two species of Juncos. *The Condor* **66**, 85–102 (1964).
15. C. Mori, W. C. Liu, K. Wada, Recurrent development of song idiosyncrasy without auditory inputs in the canary, an open-ended vocal learner. *Sci. Rep.* **8**, 8732 (2018).
16. O. Feher, H. Wang, S. Saar, P. P. Mitra, O. Tchernichovski, De novo establishment of wild-type song culture in the zebra finch. *Nature* **459**, 564–568 (2009).
17. K. Atsumi, M. Lagisz, S. Nakagawa, Nonadditive genetic effects induce novel phenotypic distributions in male mating traits of F_1 hybrids. *Evolution* **75**, 1304–1315 (2021).
18. S. Edmands, Heterosis and outbreeding depression in interpopulation crosses spanning a wide range of divergence. *Evolution* **53**, 1757–1768 (1999).
19. R. Abbott, D. Albach, S. Ansell, J. W. Arntzen, S. J. Baird, N. Bierne, J. Boughman, A. Brelsford, C. A. Buerkle, R. Buggs, R. K. Butlin, U. Dieckmann, F. Eroukhmanoff, A. Grill, S. H. Cahan, J. S. Hermansen, G. Hewitt, A. G. Hudson, C. Jiggins, J. Jones, B. Keller, T. Marczewski, J. Mallet, P. Martinez-Rodriguez, M. Most, S. Mullen, R. Nichols, A. W. Nolte, C. Parisod, K. Pfennig, A. M. Rice, M. G. Ritchie, B. Seifert, C. M. Smadja, R. Stelkens, J. M. Szymura, R. Vainola, J. B. Wolf, D. Zinner, Hybridization and speciation. *J. Evol. Biol.* **26**, 229–246 (2013).
20. R. Stelkens, O. Seehausen, Genetic distance between species predicts novel trait expression in their hybrids. *Evolution* **63**, 884–897 (2009).
21. S. Lamichaney, F. Han, M. T. Webster, L. Andersson, B. R. Grant, P. R. Grant, Rapid hybrid speciation in Darwin’s finches. *Science* **359**, 224–228 (2018).
22. B. Osthaus, L. Proops, I. Hocking, F. Burden, Spatial cognition and perseveration by horses, donkeys and mules in a simple A-not-B detour task. *Anim. Cogn.* **16**, 301–305 (2013).
23. M. A. McQuillan, T. C. Roth II, A. V. Huynh, A. M. Rice, Hybrid chickadees are deficient in learning and memory. *Evolution* **72**, 1155–1164 (2018).
24. C. Vila-Pouca, H. De Waele, A. Kotrschal, The effect of experimental hybridization on cognition and brain anatomy: Limited phenotypic variation and transgression in Poeciliidae. *Evolution* **76**, 2864–2878 (2022).
25. R. R. Hoy, J. Hahn, R. C. Paul, Hybrid cricket auditory behavior: Evidence for genetic coupling in animal communication. *Science* **195**, 82–84 (1977).
26. M. Takahasi, H. Kagawa, M. Ikebuchi, K. Okanoya, Case studies of song and call learning by a hybrid Bengalese–Zebra Finch and Bengalese-fostered Zebra Finches: Assessing innate factors in vocal learning. *Ornithol. Sci.* **5**, 85–93 (2006).
27. H. Wang, A. Sawai, N. Toji, R. Sugioka, Y. Shibata, Y. Suzuki, Y. Ji, S. Hayase, S. Akama, J. Sese, K. Wada, Transcriptional regulatory divergence underpinning species-specific learned vocalization in songbirds. *PLoS Biol.* **17**, e3000476 (2019).
28. R. B. Payne, Behavior and songs of hybrid parasitic finches. *The Auk* **97**, 118–134 (1980).
29. D. Wheatcroft, A. Qvarnström, Genetic divergence of early song discrimination between two young songbird species. *Nat. Ecol. Evol.* **1**, 0192 (2017).
30. J. M. Forshaw, M. Shephard, *Grassfinches in Australia* (CSIRO Publishing, 2012).
31. R. Zann, *The Zebra Finch: A Synthesis of Field and Laboratory Studies* (Oxford Univ. Press, 1996).
32. C. J. O. Harrison, Song and display of the Plum-headed Finch. *Emu* **74**, 254–255 (1974).
33. N. S. Clayton, The effects of cross-fostering on selective song learning in Estrildid finches. *Behaviour* **109**, 163–174 (1989).
34. A. Gehrold, S. Leitner, S. Laucht, S. Derognaucourt, Heterospecific exposure affects the development of secondary sexual traits in male zebra finches (*Taeniopygia guttata*). *Behav. Processes* **94**, 67–75 (2013).
35. D. M. Hooper, T. D. Price, Rates of karyotypic evolution in Estrildid finches differ between island and continental clades. *Evolution* **69**, 890–903 (2015).
36. R. Imai, A. Sawai, S. Hayase, H. Furukawa, C. N. Asogwa, M. Sanchez, H. Wang, C. Mori, K. Wada, A quantitative method for analyzing species-specific vocal sequence pattern and its developmental dynamics. *J. Neurosci. Methods* **271**, 25–33 (2016).

37. S. Deregnacourt, C. Poirier, A. V. Kant, A. V. Linden, M. Gahr, Comparisons of different methods to train a young zebra finch (*Taeniopygia guttata*) to learn a song. *J. Physiol. Paris* **107**, 210–218 (2013).
38. Y. Chen, L. E. Matheson, J. T. Sakata, Mechanisms underlying the social enhancement of vocal learning in songbirds. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 6641–6646 (2016).
39. F. Nottebohm, S. Kasparian, C. Pandazis, Brain space for a learned task. *Brain Res.* **213**, 99–109 (1981).
40. T. J. Devoogd, J. R. Krebs, S. D. Healy, A. Purvis, Relations between song repertoire size and the volume of brain nuclei related to song: Comparative evolutionary analyses amongst oscine birds. *Proc. R. Soc. Lond. B Biol. Sci.* **254**, 75–82 (1993).
41. B. Ward, E. Nordeen, K. Nordeen, Individual variation in neuron number predicts differences in the propensity for avian vocal imitation. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 1277–1282 (1998).
42. K. Wada, H. Sakaguchi, E. D. Jarvis, M. Hagiwara, Differential expression of glutamate receptors in avian neural pathways for learned vocalization. *J. Comp. Neurol.* **476**, 44–64 (2004).
43. K. Wada, S. Hayase, R. Imai, C. Mori, M. Kobayashi, W. C. Liu, M. Takahashi, K. Okanoya, Differential androgen receptor expression and DNA methylation state in striatum song nucleus Area X between wild and domesticated songbird strains. *Eur. J. Neurosci.* **38**, 2600–2610 (2013).
44. B. P. Ölveczky, T. M. Otchy, J. H. Goldberg, D. Aronov, M. S. Fee, Changes in the neural control of a complex motor sequence during learning. *J. Neurophysiol.* **106**, 386–397 (2011).
45. A. C. Yu, D. Margoliash, Temporal hierarchical control of singing in birds. *Science* **273**, 1871–1875 (1996).
46. S. J. Sober, M. J. Wohlgemuth, M. S. Brainard, Central contributions to acoustic variation in birdsong. *J. Neurosci.* **28**, 10370–10379 (2008).
47. R. H. Hahnloser, A. A. Kozhevnikov, M. S. Fee, An ultra-sparse code underlies the generation of neural sequences in a songbird. *Nature* **419**, 65–70 (2002).
48. T. F. Roberts, S. M. Gobes, M. Murugan, B. P. Ölveczky, R. Mooney, Motor circuits are required to encode a sensory model for imitative learning. *Nat. Neurosci.* **15**, 1454–1459 (2012).
49. S. Kojima, M. H. Kao, A. J. Doupe, M. S. Brainard, The avian basal ganglia are a source of rapid behavioral variation that enables vocal motor exploration. *J. Neurosci.* **38**, 9635–9647 (2018).
50. S. E. London, D. F. Clayton, Functional identification of sensory mechanisms required for developmental song learning. *Nat. Neurosci.* **11**, 579–586 (2008).
51. C. V. Mello, D. S. Vicario, D. F. Clayton, Song presentation induces gene expression in the songbird forebrain. *Proc. Natl. Acad. Sci. U.S.A.* **89**, 6818–6822 (1992).
52. S. Yanagihara, Y. Yazaki-Sugiyama, Auditory experience-dependent cortical circuit shaping for memory formation in bird song learning. *Nat. Commun.* **7**, 11946 (2016).
53. J. M. Moore, S. M. N. Woolley, Emergent tuning for learned vocalizations in auditory cortex. *Nat. Neurosci.* **22**, 1469–1476 (2019).
54. J. J. Bolhuis, G. G. O. Zijlstra, A. M. den Boer-Visser, E. A. Van der Zee, Localized neuronal activation in the zebra finch brain is related to the strength of song learning. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 2282–2285 (2000).
55. R. Fujimoto, K. Uezono, S. Ishikura, K. Osabe, W. J. Peacock, E. S. Dennis, Recent research on the mechanism of heterosis is important for crop and vegetable breeding systems. *Breed. Sci.* **68**, 145–158 (2018).
56. A. Zhu, I. K. Greaves, P. C. Liu, L. Wu, E. S. Dennis, W. J. Peacock, Early changes of gene activity in developing seedlings of Arabidopsis hybrids relative to parents may contribute to hybrid vigour. *Plant J.* **88**, 597–607 (2016).
57. Z. J. Chen, Genomic and epigenetic insights into the molecular bases of heterosis. *Nat. Rev. Genet.* **14**, 471–482 (2013).
58. N. Toji, A. Sawai, H. Wang, Y. Ji, R. Sugioka, Y. Go, K. Wada, A predisposed motor bias shapes individuality in vocal learning. *Proc. Natl. Acad. Sci. U.S.A.* **121**, e2308837121 (2024).
59. D. Gil, M. Gahr, The honesty of bird song: Multiple constraints for multiple traits. *Trends Ecol. Evol.* **17**, 133–141 (2002).
60. M. Goller, D. Shizuka, Evolutionary origins of vocal mimicry in songbirds. *Evol. Lett.* **2**, 417–426 (2018).
61. S. Stern, C. Kirst, C. I. Bargmann, Neuromodulatory control of long-term behavioral patterns and individuality across development. *Cell* **171**, 1649–1662.e10 (2017).
62. C. Pantoja, A. Hoagland, E. C. Carroll, V. Karalis, A. Conner, E. Y. Isacoff, Neuromodulatory regulation of behavioral individuality in zebrafish. *Neuron* **91**, 587–601 (2016).
63. E. Marder, Neuromodulation of neuronal circuits: Back to the future. *Neuron* **76**, 1–11 (2012).
64. P. S. Katz, Neural mechanisms underlying the evolvability of behaviour. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **366**, 2086–2099 (2011).
65. K. M. Schroeder, L. Remage-Healey, Adult-like neural representation of species-specific songs in the auditory forebrain of zebra finch nestlings. *Dev. Neurobiol.* **81**, 123–138 (2021).
66. E. A. Drier, M. K. Tello, M. Cowan, P. Wu, N. Blace, T. C. Sacktor, J. C. P. Yin, Memory enhancement and formation by atypical PKM activity in *Drosophila melanogaster*. *Nat. Neurosci.* **5**, 316–324 (2002).
67. F. Mery, T. J. Kawecki, Experimental evolution of learning ability in fruit flies. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 14274–14279 (2002).
68. Y. S. Lee, A. J. Silva, The molecular and cellular biology of enhanced cognition. *Nat. Rev. Neurosci.* **10**, 126–140 (2009).
69. Y.-P. Tang, E. Shimizu, G. R. Dube, C. Rampon, G. A. Kerchner, M. Zhuo, G. Liu, J. Z. Tsien, Genetic enhancement of learning and memory in mice. *Nature* **401**, 63–69 (1999).
70. T. Manabe, Y. Noda, T. Mamiya, H. Katagiri, T. Houtani, M. Nishi, T. Noda, T. Takahashi, T. Sugimoto, T. Nabeshima, H. Takeshima, Facilitation of long-term potentiation and memory in mice lacking nociceptin receptors. *Nature* **394**, 577–581 (1998).
71. D. Kroodsma, The diversity and plasticity of birdsong, in *Nature's Music: The Science of Birdsong*, (Elsevier Academic Press, 2004), pp. 108–131.
72. L. A. Kelley, R. L. Coe, J. R. Madden, S. D. Healy, Vocal mimicry in songbirds. *Anim. Behav.* **76**, 521–528 (2008).
73. A. M. Rice, M. A. McQuillan, Maladaptive learning and memory in hybrids as a reproductive isolating barrier. *Proc. R. Soc. B* **285**, 20180542 (2018).
74. S. Hayase, H. Wang, E. Ohgushi, M. Kobayashi, C. Mori, H. Horita, K. Mineta, W. C. Liu, K. Wada, Vocal practice regulates singing activity-dependent genes underlying age-independent vocal learning in songbirds. *PLoS Biol.* **16**, e2006537 (2018).
75. O. Tchernichovski, P. P. Mitra, T. Lints, F. Nottebohm, Dynamics of the vocal imitation process: How a zebra finch learns its song. *Science* **291**, 2564–2569 (2001).
76. D. G. Mets, M. S. Brainard, Genetic variation interacts with experience to determine interindividual differences in learned song. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 421–426 (2017).
77. M. Sanchez-Valpuesta, Y. Suzuki, Y. Shibata, N. Toji, Y. Ji, N. Afrin, C. N. Asogwa, I. Kojima, D. Mizuguchi, S. Kojima, K. Okanoya, H. Okado, K. Kobayashi, K. Wada, Corticobasal ganglia projecting neurons are required for juvenile vocal learning but not for adult vocal plasticity in songbirds. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 22833–22843 (2019).
78. B. M. Colquitt, D. P. Merullo, G. Konopka, T. F. Roberts, M. S. Brainard, Cellular transcriptomics reveals evolutionary identities of songbird vocal circuits. *Science* **371**, eabd9704 (2021).
79. N. C. Asogwa, N. Toji, Z. He, C. Shao, Y. Shibata, S. Tatsumoto, H. Ishikawa, Y. Go, K. Wada, Nicotinic acetylcholine receptors in a songbird brain. *J. Comp. Neurol.* **530**, 1966–1991 (2022).

Acknowledgments: We thank S. C. Woolley and T. J. Gardner for comments and discussion; M. Suga and K. Sumida for great contribution toward breeding our experimental birds; and H. Ishikawa and S. Tatsumoto for experimental and technical assistance with gene expression analysis. Computations were partially performed on the NIG supercomputer at ROIS National Institute of Genetics. **Funding:** This work was supported by JSPS KAKENHI JP20J20319 (Y.S.), JSPS KAKENHI JP16H06279-PAGS (Y.S.), JSPS KAKENHI JP20K15891 (N.T.), JSPS KAKENHI JP22J00446 (N.T.), JSPS KAKENHI JP22H04925-PAGS (N.T.), JSPS KAKENHI JP16H06524 (Y.G.), JSPS KAKENHI JP16H06531 (Y.G.), JSPS KAKENHI JP21H05238 (Y.G.), JSPS KAKENHI JP21H05245 (Y.G.), JSPS KAKENHI 4903- Evolvinguistics-JP17H06383 (K.W.), JSPS KAKENHI JP16H06279-PAGS (K.W.), JSPS KAKENHI JP21H02456 (K.W.), JSPS KAKENHI JP21K18265 (K.W.), the Joint Research of the Exploratory Research Center on Life and Living Systems (ExCELLS) (program no. 21-406, 22EXC313) (K.W.), and the Takeda Science Foundation (K.W.). **Author contributions:** Conceptualization: Y.S., N.T., and K.W. Methodology: Y.S., N.T., Y.G., and K.W. Resources: Y.S., N.T., Y.G., and K.W. Investigation: Y.S., N.T., H.W., Y.G., and K.W. Visualization: Y.S., N.T., and K.W. Funding acquisition: Y.S., N.T., Y.G., and K.W. Data curation: Y.S., N.T., H.W., and K.W. Validation: N.T. and K.W. Software: N.T. and H.W. Formal analysis: Y.S., N.T., and H.W. Project administration: Y.S. and K.W. Supervision: K.W. Writing—original draft: Y.S. and K.W. Writing—review and editing: Y.S., N.T., H.W., Y.G., and K.W. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. The single-cell RNA-seq data used in this study are available in DDBJ Sequence Read Archive, bioproject accession number PRJDB16179.

Submitted 5 December 2023

Accepted 14 May 2024

Published 19 June 2024

10.1126/sciadv.adn3409