Corticobasal ganglia projecting neurons are required for juvenile vocal learning but not for adult vocal plasticity in songbirds

Miguel Sánchez-Valpuesta[a], Yumeno Suzuki[a], Yukino Shibata[a], Noriyuki Toji[b], Yu Ji[a], Nasiba Afrin[a], Chinweike Norman Asogwa[a], Ippei Kojima[a], Daisuke Mizuguchi[c], Satoshi Kojima[c], Kazuo Okanoya[d], Haruo Okado[e], Kenta Kobayashi[f], and Kazuhiro Wada (和多和宏)[a,b,g,1]

[a]Graduate School of Life Science, Hokkaido University, Sapporo, Hokkaido 060-0810, Japan; [b]Faculty of Science, Hokkaido University, Sapporo, Hokkaido 060-0810, Japan; [c]Department of Structure and Function of Neural Networks, Korea Brain Research Institute, Daegu 41068, South Korea; [d]Department of Cognitive and Behavioral Sciences, The University of Tokyo, Meguro, Tokyo 153-8902, Japan; [e]Department of Brain Development and Neural Regeneration, Tokyo Metropolitan Institute of Medical Science, Setagaya, Tokyo 156-8506, Japan; [f]Center for Genetic Analysis of Behavior, National Institute for Physiological Sciences, Okazaki, Aichi 444-8585, Japan; and [g]Department of Biological Sciences, Hokkaido University, Sapporo, Hokkaido 060-0810, Japan

Edited by Eric I. Knudsen, Stanford University School of Medicine, Stanford, CA, and approved October 1, 2019 (received for review August 6, 2019)

Birdsong, like human speech, consists of a sequence of temporally precise movements acquired through vocal learning. The learning of such sequential vocalizations depends on the neural function of the motor cortex and basal ganglia. However, it is unknown how the connections between cortical and basal ganglia components contribute to vocal motor skill learning, as mammalian motor cortices serve multiple types of motor action and most experimentally tractable animals do not exhibit vocal learning. Here, we leveraged the zebra finch, a songbird, as an animal model to explore the function of the connectivity between cortex-like (HVC) and basal ganglia (area X), connected by $HVC_{(X)}$ projection neurons with temporally precise firing during singing. By specifically ablating $HVC_{(X)}$ neurons, juvenile zebra finches failed to copy tutored syllable acoustics and developed temporally unstable songs with less sequence consistency. In contrast, $HVC_{(X)}$-ablated adults did not alter their learned song structure, but generated acoustic fluctuations and responded to auditory feedback disruption by the introduction of song deterioration, as did normal adults. These results indicate that the corticobasal ganglia input is important for learning the acoustic and temporal aspects of song structure, but not for generating vocal fluctuations that contribute to the maintenance of an already learned vocal pattern.

critical period | sensorimotor learning | time-locked firing | zebra finch | sensory feedback

Complex motor skills are composed of a series of sequential movements acquired through learning with repetitive practice (1, 2). Neural activity coding for temporal information is considered to play an important role in the coordination of motor exploration and performance evaluation during the learning and execution of sequential movements (3, 4). General temporal information for externally reinforced motor sequence learning, such as start or stop timing, has been shown to be transferred from the cortical areas involved in cognitive control to the basal ganglia (5–7). In addition, premotor and motor cortical areas have the potential to carry into the basal ganglia fine-grained temporal information more suited for the precise control of motor learning (8–10). Indeed, cortical-basal ganglia synaptic plasticity is necessary for the acquisition of motor sequences (11), implying a potential link between impairments of connectivity from the cortex to basal ganglia and motor control pathologies (12, 13). However, how the cortical-basal ganglia connection functionally and causally contributes to learning and maintenance of sequential motor skills remains largely unknown.

Birdsong is produced as a sequence of skilled vocal movements, which are acquired through vocal learning (14–16). Songbirds memorize and copy the acoustic and sequential features of a tutor's song during a critical period of vocal learning (Fig. 1*A*). In the songbird brain, a distinct group of brain nuclei, called the song system, contributes to song learning and production (17, 18) (Fig. 1*B*). The song system is composed of 2 major circuits: the posterior vocal motor pathway and the anterior forebrain pathway (AFP). Although the vocal motor pathway participates in song production (19–21), the AFP, which is homologous to the mammalian cortical–basal ganglia–thalamic loop, plays a crucial role in vocal motor learning (22–24).

Song nucleus HVC (used as a proper name) in nidopallium, which is analogous to the mammalian premotor cortex, stands on top of the hierarchy of the song system and projects to both the vocal motor pathway and the AFP (25). HVC is a critical site for both the production and learning of song (26–28) and contains 2 major subpopulations of projection neurons: $HVC_{(RA)}$ neurons that project to the nucleus robustus of the arcopallium (RA), which is the telencephalic output locus connecting to the tracheosyringeal part of the hypoglossal nucleus (nXIIts) (29, 30), and $HVC_{(X)}$ neurons projecting to the basal ganglia nucleus area X in the AFP (31, 32). Both types of projection neurons are active at specific time points during singing renditions (8, 20, 33). It

Significance

We addressed the question, "How do corticobasal ganglia projecting neurons contribute to vocal learning?" We performed specific ablation of the vocal cortical neurons projecting to the basal ganglia, $HVC_{(X)}$ neurons in a songbird, which generate temporally precise firing during singing. Specific ablation of $HVC_{(X)}$ neurons in juveniles caused deficits in learning the tutor song's acoustics and less consistency of song sequence. In contrast, adult $HVC_{(X)}$ neuron ablation did not affect the degree of vocal fluctuations or cause alteration in song structure by auditory feedback inhibition. These results support the hypothesis that $HVC_{(X)}$ neurons are a neural substrate for transferring temporal signals, but not for regulating vocal fluctuations or conveying auditory feedback, to the basal ganglia for vocal learning and maintenance.

Author contributions: M.S.-V., K.O., and K.W. designed research; M.S.-V., Y. Suzuki, Y. Shibata, Y.J., N.A., C.N.A., and K.W. performed research; M.S.-V., I.K., H.O., K.K., and K.W. contributed new reagents/analytic tools; M.S.-V., Y. Shibata, N.T., D.M., S.K., and K.W. analyzed data; and M.S.-V. and K.W. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: wada@sci.hokudai.ac.jp.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1913575116/-/DCSupplemental.

First published October 21, 2019.

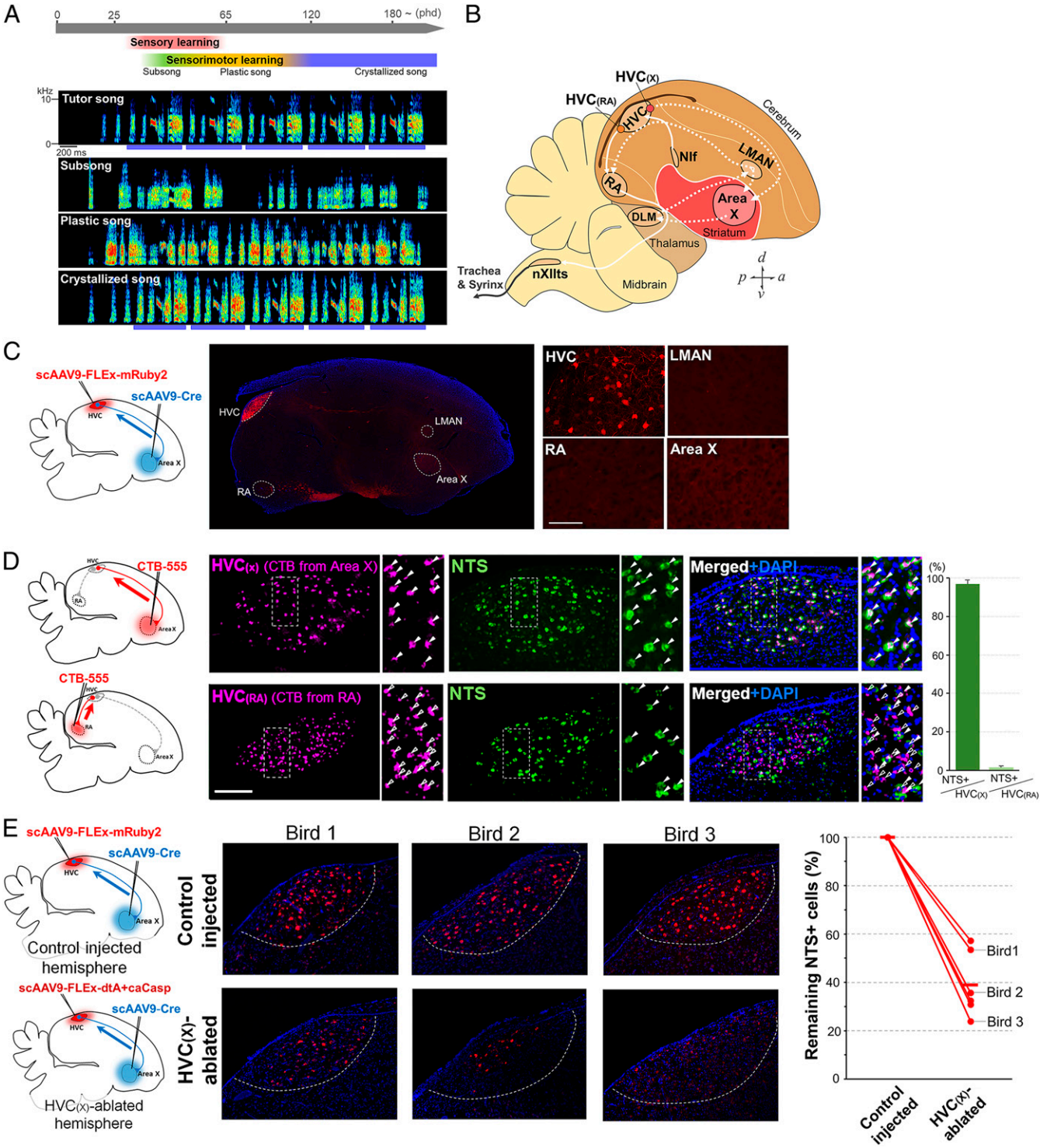

Fig. 1. Specific ablation of $HVC_{(X)}$ neurons projecting to the basal ganglia area X. (*A*, *Top*) Time course of song learning in the zebra finch. (*A*, *Bottom*) Spectrograms illustrating the progression of song learning. The blue bars represent the motif structure of the crystallized song. (*B*) Diagram showing selected song-control regions and connections in the zebra finch brain. The posterior motor pathway and the AFP (cortical–basal ganglia–thalamic circuit) are represented as solid and dotted white lines, respectively. Area X, area X of the striatum; DLM, dorsal lateral nucleus of the medial thalamus; HVC (used as a proper name); LMAN, lateral magnocellular nucleus of the anterior nidopallium; NIf, interfacial nucleus of the nidopallium; nXIIts, tracheosyringeal part of the hypoglossal nucleus; RA, robust nucleus of the arcopallium. (*C*, *Left* diagram) $HVC_{(X)}$ projection neurons were targeted using a combination of retrograding AAV-Cre injected in basal ganglia nucleus area X and AAV-FLEx-mRuby injected in HVC. (*C*, *Right*) Restricted expression of FLEx-inverted mRuby2 fluorescent protein in the $HVC_{(X)}$ cell population. (Scale bar, 100 μm.) (*D*) Selective expression of NTS in $HVC_{(X)}$ neurons (green). $HVC_{(X)}$ and $HVC_{(RA)}$ neurons were backfilled with the retrograde tracer CTB-555 from area X and RA, respectively (magenta). DAPI (blue). (*E*) Normalized decreased amount of $HVC_{(X)}$ neurons between control (*Left*) and lesioned HVC. The control hemisphere was injected with scAAV9-Cre in area X and with scAAV9-FLEx-mRuby2 in HVC. The lesioned hemisphere was injected with scAAV9-Cre in area X and with a mixture of scAAV9-FLEx-dtA and -caCasp in HVC.

has been proposed that the firing of $HVC_{(X)}$ neurons collectively represents temporal sequence information during song renditions but does not convey the properties of constitutive vocal gestures nor sensory feedback signals (8, 9, 34–38). Although the ablation of $HVC_{(X)}$ neurons in adults does not affect the execution of learned vocalization (39), the potential functional contribution of $HVC_{(X)}$ neurons in vocal learning remains unclear. In addition, the AFP is a crucial neural site for the generation of vocal exploration and the refinement of vocal performance using auditory feedback information (40–43). However, it is unknown how the temporally precise firing inputs from $HVC_{(X)}$ neurons to area X relate to regulation of vocal variability and auditory-dependent song maintenance.

Here, we performed cell type-specific ablation of $HVC_{(X)}$ neurons [$HVC_{(X)}$ ablation] to disrupt the transfer of sequential and temporally precise firing from a cortical-like region to the basal ganglia. To elucidate the cellular functions of $HVC_{(X)}$ neurons on song learning and maintenance, we ablated $HVC_{(X)}$ neurons in juvenile zebra finches before the initiation of vocal motor learning and analyzed their acquired songs. In addition, we examined the effect of $HVC_{(X)}$ ablation on the regulation of vocal variability and change in song structure after deprivation of auditory feedback in adults.

Results

$HVC_{(X)}$ Neuron-Specific Ablation in Zebra Finches. To achieve sufficient and specific ablation of $HVC_{(X)}$ neurons in vivo, we used self-complementary adeno-associated virus (scAAV) vectors to ensure a stronger and faster induction than normal single-stranded AAV genomes (44). We used AAV serotype 9 capsid (AAV9), which allows retrograde transport from the site of injection, and a Cre/FLEx switch system for the conditional induction of gene expression. We first injected scAAV9-Cre and scAAV9-FLEx-mRuby2 into area X and HVC, respectively, to test the induction rate and timing of expression of a targeted gene (i.e., mRuby2) in $HVC_{(X)}$ neurons. One week after the injections, we observed the selective expression of mRuby2 protein in $HVC_{(X)}$ neurons (Fig. 1*C* and *SI Appendix*, Fig. S1*A*). Neurotensin (NTS) mRNA was used as a marker of $HVC_{(X)}$ neurons. We confirmed that NTS mRNA was specifically expressed in $HVC_{(X)}$ neurons labeled with retrograde cholera toxin B (CTB) injected in area X [$n = 4$; 96.9 ± 2.0% of total $HVC_{(X)}$ neurons], but not in $HVC_{(RA)}$ neurons (1.7 ± 0.7%) (Fig. 1*D*). Using NTS labeling, a reliable estimation of residual $HVC_{(X)}$ neuron number after ablation could be performed without additional retrograde labeling from area X to HVC. We then evaluated the ablation efficiency of $HVC_{(X)}$ neurons by the dual induction of diphtheria toxin A (dtA) (45, 46) and constitutively active caspase 3 (caCasp) (47) into the same cells using the Cre/FLEx switch system. Both dtA and caspase 3 have been shown to synergistically potentiate caspase 3-dependent apoptotic cell death (48, 49). For this purpose, we injected scAAV9-Cre into area X and a mixture of scAAV9-FLEx-dtA and -caCasp into HVC in a test hemisphere. In the other hemisphere of the same animal, scAAV9-Cre and scAAV9-FLEx-mRuby2 were injected into area X and HVC, respectively, to serve as the control hemisphere. The number of $HVC_{(X)}$ neurons was compared between the control and $HVC_{(X)}$-ablated hemispheres (Fig. 1*D* and *SI Appendix*, Fig. S1*B*). As a result, the residual number of $HVC_{(X)}$ neurons was reduced to 23.9 to 57.2% (mean ± SD = 38.8 ± 13.5%) in the ablated hemisphere when compared against the control HVC of the same individual ($n = 6$; $P < 0.001$, unpaired *t* test). Although the cell ablation procedure did not achieve complete removal of $HVC_{(X)}$ neurons, our method using a mixture of dtA and caCasp showed a similar or higher effective reduction of target cells compared with previous efforts at cell ablation in songbirds (39, 42, 50).

Deficits in Song Learning and Development by Ablation of $HVC_{(X)}$ Neurons. To examine the cell type-specific function of $HVC_{(X)}$ neurons in song learning and development, we bilaterally injected scAAV9-Cre and scAAV9-FLEx-dtA/-caCasp into area X and HVC, respectively, to ablate $HVC_{(X)}$ neurons before the initiation of sensorimotor learning. The injected juveniles were tutored using playback songs (posthatching day [phd] ~35) (Fig. 2*A*). We continuously recorded their songs daily and later evaluated the residual number of $HVC_{(X)}$ neurons in the adult stage (phd 180 to 200) by fluorescence in situ hybridization (FISH) using an NTS probe. We used birds possessing $HVC_{(X)}$ neuron densities lower than 130 NTS^+ cells per mm^2 in both hemispheres for further analyses [NTS^+ cells per mm^2 in HVC (mean ± SD); $HVC_{(X)}$ ablation: 104.1 ± 10.8, control injection: 386.2 ± 33.7] (*SI Appendix*, Fig. S2). The degree of $HVC_{(X)}$ ablation in individual birds ranged from 68.6 to 76.3% (mean ± SD, 73.0 ± 2.8%) in $HVC_{(X)}$-ablated birds when compared with the average density of NTS^+ cells in the HVC of control birds. $HVC_{(X)}$-ablated birds developed their songs from subsongs with unstable syllable acoustics into a more stable and consistent song structure through the critical period of song acquisition. The timing of song stabilization of $HVC_{(X)}$-ablated birds tended to be delayed compared with control birds for both syllable acoustic

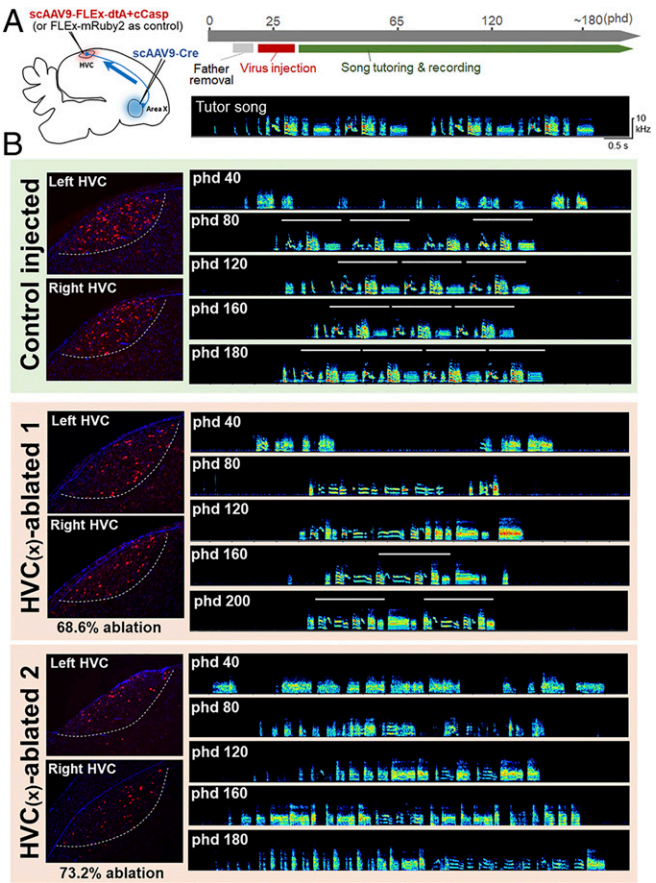

Fig. 2. Ablation of $HVC_{(X)}$ neurons in juveniles induces deficits in song learning and development. (*A*) Experimental timeline for $HVC_{(X)}$ ablation and song tutoring. (*B*) Examples of song development in a control injected (green-colored background) and 2 $HVC_{(X)}$-ablated (brown-colored background) birds. $HVC_{(X)}$-ablated birds 1 and 2 had decreases of 68.6% and 73.2% of $HVC_{(X)}$ neurons, respectively, compared with the average of $HVC_{(X)}$ neurons in the control birds. The white lines in the song spectrograms represent the motif structure of songs. The remaining $HVC_{(X)}$ neurons were labeled with NTS (red). DAPI (blue). The white dotted lines represent HVC borders.

and sequential features (Fig. 2*B*). Thus, we used the syllable acoustics and sequences of acquired songs as behavioral parameters to evaluate the successfulness of song learning.

$HVC_{(X)}$-ablated birds showed deficits in copying the acoustic features of syllables from the tutored songs (Fig. 3*A* and *SI Appendix*, Fig. S3). Although there was no significant difference in the number of unique syllables between control and $HVC_{(X)}$-ablated birds in adults (phd 150 to 160) ($P = 0.12$, Student's t test), $HVC_{(X)}$-ablated birds showed a significant decrease in the average of each syllable similarity score toward the original tutored syllables compared with those of control birds (**$P < 0.01$, Student's t test) (Fig. 3*B*). $HVC_{(X)}$-ablated birds did not completely fail to mimic the syllables of the tutored song. Rather, they copied some populations of syllables from the tutor song, although other populations of acquired syllables did not belong to the tutor song (Fig. 3*A*). This mixture of copied and noncopied syllables caused a significantly higher coefficient of variation (CV) in the syllable similarity scores in the $HVC_{(X)}$-ablated birds vs. control birds (*$P < 0.05$, Student's t test) (Fig. 3*B*). We further tested whether there was an association between the existing number of $HVC_{(X)}$ neurons and the syllable similarity score of individual birds (Fig. 3*C*). We found a significant correlation between the 2 factors ($P < 0.027$, $r = 0.758$, Pearson's correlation coefficient), indicating the potential contribution of $HVC_{(X)}$ neurons to the accurate learning of syllable acoustics.

We further noticed that 2 of 5 $HVC_{(X)}$-ablated birds produced acoustically unstable syllables with variable entropy variances and durations even in a mature adult stage (phd >150) (Fig. 3*D*). In addition, although there was no significant difference in the median of intersyllable gap duration between control and $HVC_{(X)}$-ablated birds' groups, 2 of the $HVC_{(X)}$-ablated birds produced strikingly short intersyllable gaps (median, <25 ms) (Fig. 3 *E* and *F*). A certain number of $HVC_{(X)}$-ablated birds had defects not only in song learning but also in producing stable acoustic structures of song.

Ablation of $HVC_{(X)}$ Neurons in Juveniles Increased the Instability of the Adult Song Syllable Sequence. We next examined the developmental effects of $HVC_{(X)}$ ablation on the syllable sequence in songs. We used the syllable similarity matrix (SSM) method, which allowed quantitative analysis of the frequency of characteristic syllable transitions in songs without using human-biased procedures for syllable identification (51). In this analysis, 2 successive paired- and repetitive-syllable transitions were respectively measured as motif or repetitive indices (Fig. 4*A*, *SI Appendix*, Fig. S4, and *Materials and Methods*). The analysis indicated that $HVC_{(X)}$-ablated birds displayed a significant decrease in the frequency of paired-syllable transitions, forming motif structures when compared with control birds (*$P < 0.05$, Student's t test) (Fig. 4*B*). Two of the $HVC_{(X)}$-ablated birds produced a relatively higher degree of repetitive-syllable transitions in their songs compared with the songs of control birds, despite the existence of a variety of individual differences among the ablated birds (Fig. 4*B*). In addition, we calculated song consistency to examine the sequence variability of their songs (23). In agreement with SSM analysis, $HVC_{(X)}$-ablated birds showed significantly decreased song consistency when

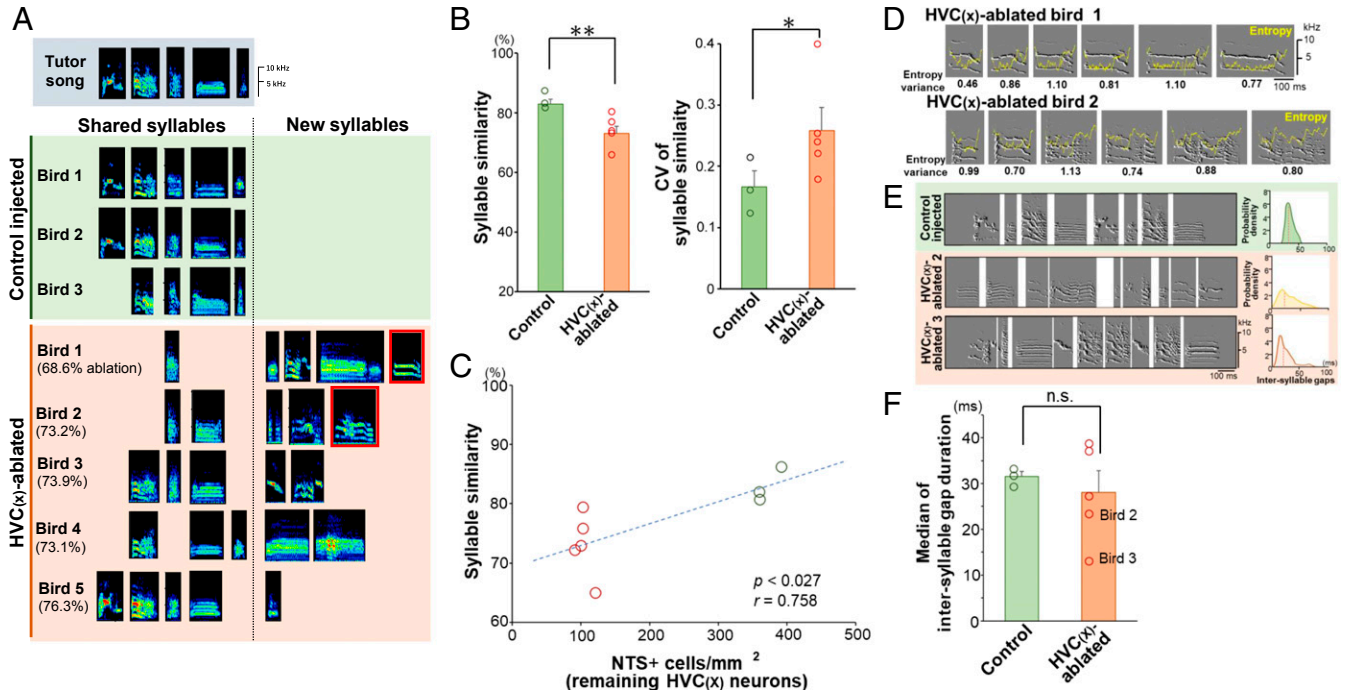

Fig. 3. Ablation of $HVC_{(X)}$ neurons in the juvenile stage caused abnormality in syllable acoustics and intersyllable gap duration in adult songs. (*A*) Examples of acquired syllables in control (green background) and $HVC_{(X)}$-ablated (brown-colored background) birds. Syllables outlined with red lines were further analyzed in *C*. (*B*) Differences between control and $HVC_{(X)}$-ablated birds in the syllable similarity between syllables of each pupil and the tutor song (*Left*) and its CV (*Right*) (n = 3 controls, $n = 5$ ablated birds; Student's t test: *$P < 0.05$, **$P < 0.01$). Mean + SEM for bar graphs. (*Left*) Each point represents the average similarity score of all syllable types for individual birds. (*Right*) Each point represents the CV of the similarity scores of all syllable types for individual birds. (*C*) Correlation between NTS^+ cell density in HVC [i.e., degree of residual $HVC_{(X)}$ neurons] and syllable similarity between syllables of each pupil and the tutor song ($P < 0.027$, $r = 0.758$, Pearson's correlation coefficient). The green and red circles represent control and $HVC_{(X)}$-ablated birds, respectively. (*D*) High variability in duration and acoustics in syllables in the adult stage (phd 150) for birds whose $HVC_{(X)}$ neurons were ablated in the juvenile stage. The yellow lines represent acoustic entropy, and numerical values show entropy variance. (*E*) Examples of abnormal intersyllable gaps in the adult stage for birds whose $HVC_{(X)}$ neurons were ablated in the juvenile stage. (*Left*) Variability and shortening of intersyllable gaps in $HVC_{(X)}$-ablated birds in the juvenile stage. (*Right*) Probability density of intersyllable gaps from each bird ($n = 100$ gaps). The red dotted lines indicate median values. (*F*) Median of intersyllable gap duration between control and $HVC_{(X)}$-ablated birds (100 intersyllable gaps per bird). Each dot represents an individual bird's value. Bird ID numbers are consistent between *A* and *C*–*F*.

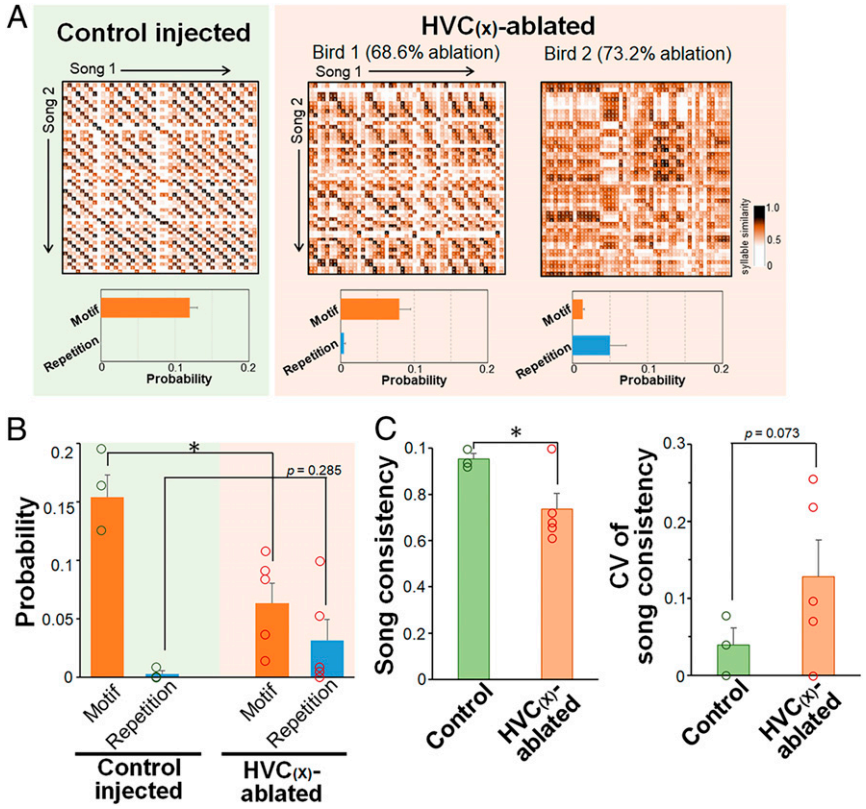

Fig. 4. Altered sequential properties of the adult songs of $HVC_{(x)}$-ablated birds in the juvenile stage. (*A*, *Top*) Representative syllable similarity matrices (SSMs) for adult songs (phd 150) in control (green background) and 2 $HVC_{(x)}$-ablated (brown background) birds. (*A*, *Bottom*) Probabilities of motif and repetition indices for each bird. (*B*) Probabilities of motif and repetition indices in the adult stage (phd 150) in control and $HVC_{(x)}$-ablated birds ($n = 3$ controls, $n = 5$ ablated birds; Student's *t* test, $^*P < 0.05$). The dots indicate individual bird's values. (*C*) Song sequence consistency and its CV at phd 150 in control and $HVC_{(x)}$-ablated birds ($n = 3$ controls, $n = 5$ ablated birds; Student's *t* test, $^*P < 0.05$). Mean + SEM for all graphs. The dots indicate individual bird's values.

compared with control birds ($^*P < 0.05$, Student's *t* test) (Fig. 4*C*). Furthermore, the scores for song consistency showed a large variation among the $HVC_{(X)}$-ablated birds, which is reflected as a large degree in the CV in song consistency. Taken together, these results indicated that the reduction of $HVC_{(X)}$ neurons not only caused deficits in copying tutor song acoustics but also led to the abnormal development of song phonology and sequence.

Ablation of $HVC_{(X)}$ Neurons in Adults Did Not Alter Crystallized Song Structure or Vocal Fluctuation. In a previous study, chromophore-laser ablation of $HVC_{(X)}$ neurons in the adult stage does not alter the learned song structure (39). Similarly, lesions of adult area X cause no apparent alterations in syllable acoustics and sequence (23, 24). However, area X lesions alter the within-syllable variability in fundamental frequency (FF) and cause a transient effect on cross-rendition variability in syllable FF even in the adult stage, suggesting that area X activity is related to the role of the AFP in generating exploratory vocal variability (43, 52, 53). However, it remains unknown whether $HVC_{(X)}$ ablation also influences vocal variability in a similar manner to area X lesions.

To examine this possibility, we injected scAAV9-Cre and scAAV9-FLEx-dtA/-caCasp into area X and HVC, respectively, of adult zebra finches (phd <120), causing a similar degree of ablation, ranging from 68.3 to 86.1% (mean ± SD, 79.1 ± 8.3%), as observed in $HVC_{(X)}$ ablation in juveniles (*SI Appendix*, Fig. S2). Owing to the absence of any changes in song structure at a few days after virus injection [i.e., $HVC_{(X)}$ ablation had yet to occur yet due to the time lag of gene induction by AAVs], we considered that bilateral AAV injections did not cause direct physical damage (Fig. 5 *A* and *B*). For quantitative comparisons of song structure changes between preablation and 2 to 3 wk postablation, we used the motif index based on the SSM for the syllable sequence (51), Kullback–Leibler (K–L) distance based on 2D syllable scatter plots for syllable acoustics (duration and mean frequency modulation [FM]) (54, 55) (Fig. 5*C*), and motif duration for song tempo (43, 56). Similar to a previous study using the laser ablation of adult $HVC_{(X)}$ neurons (39), our $HVC_{(X)}$ ablation in the adult stage did not induce changes in these parameters of song structure (Fig. 5 *D*–*F*), indicating learning state-dependent effects of $HVC_{(X)}$ ablation on the production of structured songs. We then examined the potential contribution of $HVC_{(X)}$ neurons to the generation of song variability by focusing on "within-syllable variability" and "cross-rendition variability" in syllable FF between control and $HVC_{(X)}$-ablated adult birds. We found no obvious alterations in pre/post changes in both within- and cross-rendition syllable variability in FF between control and $HVC_{(X)}$-ablated birds (Fig. 5 *G* and *H*). In line with this finding, there were no significant differences in both the within- and cross-rendition variability in FF of the syllables between the pre-injection and postinjection states of $HVC_{(X)}$-ablated birds (*SI Appendix*, Fig. S5). These results indicate that $HVC_{(X)}$-ablated birds generate vocal fluctuations to a similar degree compared with the control birds.

Auditory Feedback-Dependent Song Changes after $HVC_{(X)}$ Ablation. Although we found a consistent generation of vocal fluctuations after adult $HVC_{(X)}$ ablation, such $HVC_{(X)}$-ablated birds should transmit deteriorated temporal information to the AFP. Therefore, we hypothesized that $HVC_{(X)}$ ablation may have a different effect on auditory feedback-dependent vocal plasticity

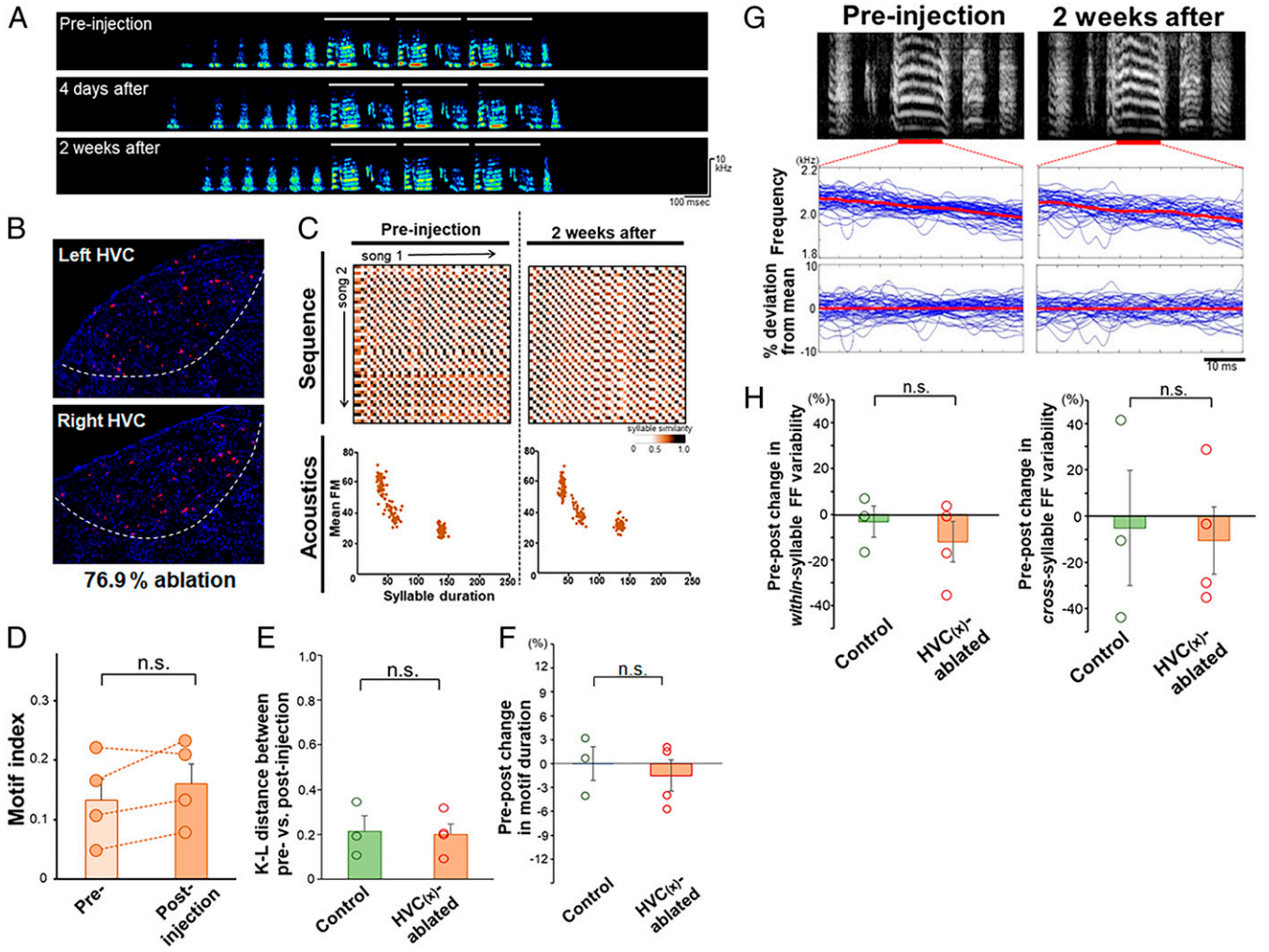

Fig. 5. Nonobvious change in song structure by ablation of $HVC_{(X)}$ neurons in adults. (*A*) Representative spectrogram of birds that were ablated in $HVC_{(X)}$ neurons in adults. The white bars represent the motif structure of songs. (*B*) Example of the extent of $HVC_{(X)}$ ablation in an ablated adult (with 76.9% ablation) as shown in *A* and *C*. NTS (red) and DAPI (blue). (*C*) Syllable sequence and acoustic stability before and after ablation of $HVC_{(X)}$ neurons. Sequential patterns are shown as SSMs and acoustics as a scatter density plot of syllable duration vs. mean FM ($n = 150$ syllables). (*D*) No effect of $HVC_{(X)}$ ablation on the song motif index of adult zebra finches. Each point corresponds to an individual bird. (*E*) No effect of $HVC_{(X)}$ ablation on syllable acoustics measured by the K–L distance of syllable scatter density plots (duration vs. mean FM) between preinjection and postinjection time points ($n = 3$ controls, $n = 4$ ablated birds; Student's *t* test: n.s., $P > 0.05$). (*F*) Pre–post change in motif duration between control and $HVC_{(X)}$-ablated birds (Student's *t* test: n.s., $P > 0.05$). (*G*) Example of an FF trajectory of a syllable in preinjection (*Top Left*) and 2 wk postinjection (*Top Right*) of songs from an $HVC_{(X)}$-ablated adult, expressed as raw frequency traces (*Middle*) and percent deviation from the within-rendition mean (*Bottom*). The blue and red lines indicate each rendition and the mean across renditions, respectively. (*H*) Pre–post changes in within- and cross-rendition syllable variability in FF between control and $HVC_{(X)}$-ablated birds (Student's *t* test: n.s., $P > 0.05$). Mean + SEM for all graphs.

than area X lesions. The neural activity output from the AFP to the song nucleus RA plays an important role in the regulation of vocal fluctuations (43, 57, 58), which could, in turn, be a driving force for the induction of song degradation after the disruption of auditory feedback (52, 57–59). Therefore, to examine whether adult $HVC_{(X)}$-ablated birds undergo a degradation of song structure by auditory deprivation, we prepared adult $HVC_{(X)}$-ablated and deafened birds by bilateral cochlear extirpation after 3 wk after virus injection (Fig. 6*A*). Ablation procedures caused a similar degree of $HVC_{(X)}$ ablation, ranging from 66.2 to 78.8% (mean ± SD, 72.2 ± 5.3%), as observed in the juveniles and adults without deafening (*SI Appendix*, Fig. S2). We then compared the degree of song degradation after deafening with age-matched deafened-alone birds. We found that deafened birds after $HVC_{(X)}$ ablation had a similar trajectory and variation of degradation of both sequence and acoustic features as the degree of degradation shown in deafened-alone birds (Fig. 6 *B*–*D*). A comparison of the motif indices and K–L distances indicated significant differences between the predeafening and postdeafening time points in both $HVC_{(X)}$-ablated and deafened and control deafened-alone birds (Fig. 6 *C* and *D*). However, there were no significant differences in the motif indices and K–L distances between the 2 groups after 1 and 2 mo. These results indicate that AFP output activity generated under a severely reduced number of $HVC_{(X)}$ neurons is still sufficient to induce auditory-dependent song structural change.

Discussion

We utilized songbirds as a model system to investigate the cell type-specific function of cortical neurons projecting to the basal ganglia on motor skill learning, motor fluctuation, and sensory-feedback–dependent alterations of learned motor skills. The song system, which includes the AFP, is a discrete neural circuit that shares a number of similarities with mammalian motor circuits (17, 18, 60). Unlike mammalian motor circuits, the song system is specialized for a well-defined behavior, singing, which

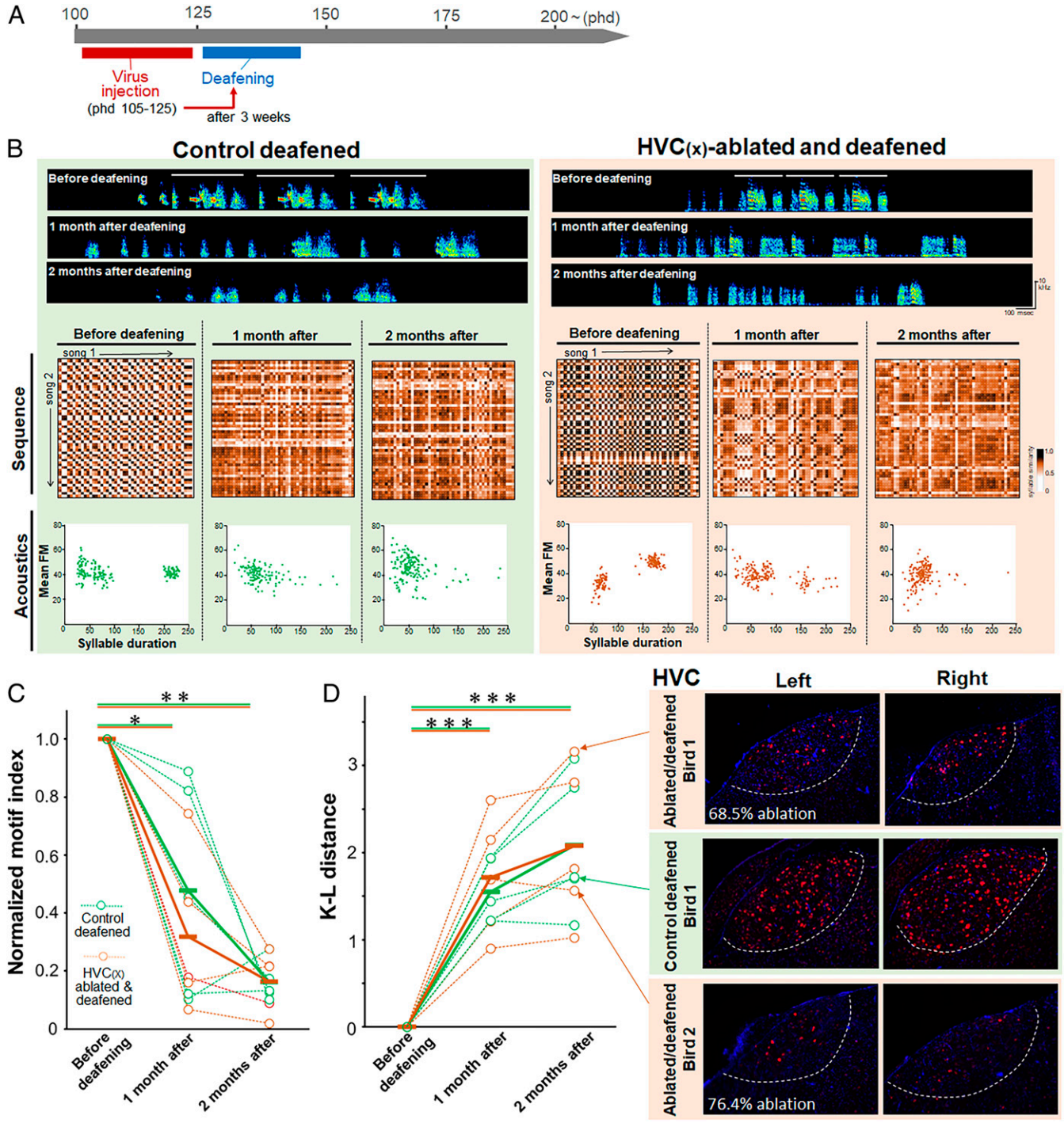

Fig. 6. Ablation of $HVC_{(X)}$ neurons does not alter song degradation after deafening in adult zebra finches. (*A*) Timeline of $HVC_{(X)}$ ablation and deafening in the adult stage. (*B*) Deafening-induced degradation of the syllable sequence and acoustics in a control (green background) and $HVC_{(X)}$-ablated (brown background) adult birds. (*C*) Similar rates of deafening-induced degradation of song motif structure between control and $HVC_{(X)}$-ablated adult birds ($n = 5$ for each group; paired *t* test: $^{*}P < 0.05$, $^{**}P < 0.01$). The green and brown lines represent control and $HVC_{(X)}$-ablated birds, respectively. The dotted and solid lines represent individual and average values, respectively. (*D*, *Left*) Similar rates of acoustic degradation after deafening between control and $HVC_{(X)}$-ablated birds, as calculated by the K–L distance ($n = 5$ for each group; paired *t* test: $^{***}P < 0.001$). (*D*, *Right*) Remaining $HVC_{(X)}$ neurons in 3 representative birds [a control and 2 $HVC_{(X)}$-ablated birds], visualized by NTS (red) and DAPI (blue). The white dotted lines represent the border of HVC.

therefore allows us to investigate cell type-specific functions in the circuits through quantitative behavioral measurements. Using AAV-induced genetic ablation of $HVC_{(X)}$ neurons in juvenile and adult zebra finches, we found a functional contribution of $HVC_{(X)}$ neurons to learning the acoustic and temporal aspects of song structure. In contrast, we did not observe effects of $HVC_{(X)}$ ablation on the generation of vocal fluctuations and auditory-dependent changes in already learned songs. These results broadly support the hypothesis that the temporally precise activity of $HVC_{(X)}$ neurons is crucial for vocal motor learning, but is not

involved in the generation of vocal fluctuation or transfer sensory feedback signals to the basal ganglia nucleus area X (3, 34, 36–38). As a potential future research direction, an experiment combining $HVC_{(X)}$ ablation and song-triggered microdisruption of auditory feedback, such as pitch-shift learning manipulation in juvenile and adult stages, would be valuable to elucidate the potential contribution of "temporally precise firing" of $HVC_{(X)}$ neurons to real-time modulation of song acoustics and sequence (40, 53, 57, 61).

The ablation of $HVC_{(X)}$ neurons indicated similarities and differences to lesions of the basal ganglia song nucleus area X (23, 52). Both $HVC_{(X)}$ ablation and area X lesions occurring before the initiation of vocal motor learning induced similar effects on song acquisition and execution. Both lesions affect the ability to copy tutor songs and lead to the production of sequentially unstable songs with inconsistency of syllable and intersyllable gap durations, which are different from the effects of LMAN (lateral magnocellular nucleus of the anterior nidopallium) lesion (22–24). In addition, neither $HVC_{(X)}$ neurons nor area X are required for the rendition of learned structured songs (23, 39). However, auditory-feedback–driven song changes in adults were strikingly different between 2 cases. Like LMAN lesions, area X lesions prevent deafening-induced song degradation (52, 59). In contrast, deafened birds after $HVC_{(X)}$ ablation showed a very similar trajectory of song structural changes compared with $HVC_{(X)}$-intact deafened birds at both the phonological and sequential levels (Fig. 6). The degree of $HVC_{(X)}$ ablation in adult deafened birds was 66 to 78% in this study. In contrast, area X lesions ranged from 75 to 100% in a previous study (52). Therefore, it is important to consider the possibility that the difference in enabling auditory-feedback–dependent vocal plasticity between area X lesions and those in our study is not caused by the irrelevance of the connection between HVC and area X, but may reflect the qualitative and quantitative differences in the magnitude of disruption of area X activity between the 2 studies. A previous electrophysiological study demonstrated that lesions of area X diminished song-locked burst firing tendency in LMAN, but did not affect the firing rate during singing (52), suggesting that the generation and transmission of temporally biased burst firing signals from LMAN to RA is a crucial factor for the induction of auditory-feedback–dependent song plasticity. In this scenario, $HVC_{(X)}$-ablated birds may show dampened song-locked firing patterns but may maintain burst firing in LMAN, which still induces vocal variability via the AFP outflow to RA during singing. If so, our results may support the hypothesis that functional AFP-driven vocal fluctuation is generated by area X, DLM (dorsal lateral nucleus of the medial thalamus), and LMAN independently from $HVC_{(X)}$-derived temporal information. In addition, our results support recent studies showing that auditory feedback signals are not transferred into the AFP via $HVC_{(X)}$ neurons (37, 38), but rather from other areas such as the ventral tegmental area, a region implicated in reinforcement learning (62, 63). However, we cannot rule out the possibility that other HVC cell populations may be a locus for the transfer of auditory feedback signal from the auditory forebrain to the song system.

In general, virus injection-based cell ablation is technically limited in terms of the removal efficiency of targeted cells, often not achieving complete ablation (in this study, ~85% ablation was the maximum efficiency) when compared with transgenesis-based cell ablation. However, this technique still has benefits for exploring the function of $HVC_{(X)}$ neurons. In the zebra finch, individual $HVC_{(X)}$ neurons generate temporally precise sparse and brief bursts of spikes during each song rendition (8), with bursts of different $HVC_{(X)}$ neurons being generated at different time points in the song motif, covering both syllables and intersyllable gaps. In addition, the axon terminal arborization of $HVC_{(X)}$ neurons in area X lacks topographical organization between the 2 song nuclei, indicating that each $HVC_{(X)}$ neuron projects to most of area X (64). These issues suggest that, during singing, cell assemblies of $HVC_{(X)}$ neurons transmit information about the successive current song-locked time to area X as a continuous-temporal code that allows temporal specificity for song learning (3). Thus, the virus injection-based $HVC_{(X)}$-ablated birds can be thought of as a model system for motor learning displaying deteriorated internal temporal firing information. Therefore, in the ablated birds left with significantly decreased numbers of $HVC_{(X)}$ neurons, area X should receive a temporally incomplete (but not completely extinct) code as a sequentially inconsistent time representation for the learning process. This may be one of the reasons why the ablated juveniles still retained the ability to copy a few syllables from the tutor songs and developed relatively structured songs, despite showing phonological and sequential instability (Fig. 3), instead of producing completely unstructured songs. It is necessary to point out a potential abnormality in the HVC microcircuit resulting from depletion of an $HVC_{(X)}$ subpopulation. Although we did not examine the cell number of other HVC neuron subpopulations, such as $HVC_{(RA)}$ neurons and interneurons, a previous study indicated that ablation of $HVC_{(X)}$ neurons in juveniles up-regulates the incorporation of new $HVC_{(RA)}$ neurons in HVC, although $HVC_{(X)}$-ablated adults do not increase $HVC_{(RA)}$ neurons (39). The increased number of $HVC_{(RA)}$ neurons in $HVC_{(X)}$-ablated birds in the juvenile stage might cause the deficit in song learning and development through induction of an imbalance in synaptic connections between HVC and LMAN to RA and/or potential circuitry disruption within HVC. The same applies to other described functions of $HVC_{(X)}$ neurons in the HVC nucleus, such as retinoic acid synthesis (65). Here, the mRNA of its synthesizing enzyme is only expressed by $HVC_{(X)}$ neurons, but its proteins are found in neighboring $HVC_{(RA)}$ neurons (66) or the guidance of newly born $HVC_{(RA)}$ neurons by interactions with $HVC_{(X)}$ neurons (67). All of these hypothesized functions of $HVC_{(X)}$ neurons in the HVC microcircuit would be affected by $HVC_{(X)}$ neuron ablation.

We showed the different learning state-dependent effects of $HVC_{(X)}$ ablation on song production, finding no apparent effects of $HVC_{(X)}$ neurons on the execution and maintenance of learned songs in the zebra finch, consistent with a previous study (39). However, owing to incomplete $HVC_{(X)}$ ablation, it is necessary to consider the possibility that the residual population of $HVC_{(X)}$ neurons could still fulfill their role in generating song fluctuation and regulating auditory-dependent song deterioration in adults. Hence, adult birds with ablated $HVC_{(X)}$ neurons could execute and maintain a learned song in a similar way to the control birds. Therefore, it remains crucial to perform transgenesis-based $HVC_{(X)}$ neuron-specific ablation, although transgenic songbirds have yet to become a widely used experimental approach.

In summary, our results shed light on how cortical neurons projecting to the basal ganglia contribute to motor skill learning, thus confirming the importance of the inputs from the song motor nucleus HVC input to the AFP in song learning. Furthermore, our data portray the learning state-specific role of cortical-basal ganglia projection neurons for vocal skill learning. In general, motor cortex in mammals is believed to play a role in learning and execution of skilled motor patterns, although it is not fully understood how the motor cortex contributes to these processes. Interestingly, in rats, lesions of the primary and secondary forelimb motor cortical areas induce no recognizable change in already acquired motor skills, i.e., grasp movements. In contrast, lesions of the motor cortex prior to training alter the acquisition of skilled motor patterns (68). Although the results are similar to our findings on the specific ablation of $HVC_{(X)}$ neurons at the learning state-specific aspect, lesions of whole HVC (non–cell-specific ablation) disrupt learned song in adulthood (27). This finding is similar with speech apraxia observed after lesions of the premotor/motor cortex in adult humans (69, 70). A possible explanation for this apparently paradoxical result between species may be caused

by different behavioral tasks, which, in turn, would depend on different projecting neuronal pathways or subcortical structures for their learning and maintenance. Thus, our data show a functional resemblance and difference in the motor control network for motor skill learning and execution between avian and mammalian species. Furthermore, considering the important role of the cortical–basal ganglia circuit for speech learning and production in humans and the paucity of animal models for vocal learning (17, 18, 71–73), studies of the cortical–basal ganglia circuit in songbirds may enhance our understanding of the neural basis of vocal developmental and vocal communication disorders in humans.

Materials and Methods

Ethics Statement. All animal experiments were performed according to the guidelines of the Committee on Animal Experiments of Hokkaido University from whom permission for this study was obtained (Approval No. 13-0061). The guidelines are based on national regulations for animal welfare in Japan (Law for the Humane Treatment and Management of Animals; after Partial Amendment No. 105, 2011). For brain sampling, the birds were humanely killed by decapitation after being injected with a lethal dose of pentobarbital.

Animals. Male zebra finches were obtained from our breeding colony at Hokkaido University. The photoperiod was constantly maintained at a 13/11-h light/dark cycle with food and water provided ad libitum. The sex of the birds was checked by PCR to select male juveniles before experimental manipulation. Birds for the song developmental study were raised in individual breeding cages with their parents and siblings until phd 5 to 15. Juveniles (along with their siblings) were then raised in a sound-attenuation box by their mother with their siblings after removal of their father (removed before phd 15) until they could feed themselves (phd ∼35). Juvenile birds were subsequently separated from their mother and siblings and housed in individual isolation boxes for song playback, with the same tutor song being played back from phd 30 to all developing juveniles until at least phd 140. Birds for the adult experiments were raised in individual breeding cages with their parents and siblings until phd 60 to 100, and then housed in common cages with other male birds.

Song Recording, Tutoring, and Analysis. Songs were recorded using a unidirectional microphone (SM57; Shure) connected to a computer with the sound event triggered by recording software Sound Analysis Pro (sap v2011.089; http://soundanalysispro.com/) (74). Each song bout was saved as a sound file (.wav file), including time information. Low-frequency noise (<0.5 kHz) and mechanical noise were filtered out using Avisoft-SASLab's (Avisoft Bioacoustics) bandpass filter. With respect to song tutoring, birds were individually housed in a sound-attenuating box containing a mirror to reduce social isolation. Tutor songs were played 5 times in the morning and 5 times in the afternoon at 55 to 75 dB from a speaker (SRS-M30; Sony) passively controlled by Sound Analysis Pro.

For the analysis of similarity between pupil and tutor songs, the comparison of tutor and pupil syllable acoustic features was performed using the SAP program's similarity module. The score was calculated using the "symmetric" and "time courses" comparison settings after manually adjusting the thresholds for every syllable. Overall, 80 to 120 syllables, including multiple syllable types, were compared with syllables from tutor songs to obtain each similarity score between syllables in the pupil and tutor songs. The mean values of the similarity score for each syllable type in pupil songs against each syllable type in tutor songs were calculated, and the highest mean values were used as the similarity scores of each syllable type. We used the total mean value of the similarity scores of all syllable types for each individual bird. For the CV of syllable similarity (shown in Fig. 3*B*), the CV using the similarity scores of each syllable type was calculated for individual birds.

To analyze the syllable transitions, song similarity matrix (SSM) analysis was performed (51). In all, 250 syllables from songs chosen randomly at phd 150 were used. Introductory notes in a song were not included in the song rendition. A total of 100 serially separated ".son"-converted syllable files were transferred to the Avisoft CORRELATOR program to calculate the similarity scores between the syllables of 2 songs by the round-robin comparison. The score was calculated as the peak correlation coefficient between 2 syllables according to the following formula:

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

where m_a and m_b are the mean values of the spectrograms a and b, respectively. a_{xy} and b_{xy} are the intensities of the spectrogram points at the locations x and y, respectively. The syllable similarity score is a value ranging from 0 to 1. Similarity scores between the syllables in 2 song renditions were exported into cells in the Microsoft Excel spreadsheet by maintaining the syllable sequence order in the original songs. The spreadsheet was named with the information of the similarity scores between the syllables as an SSM. In this study, 10 SSMs per bird were prepared by the round-robin comparison of 250 syllables. To qualitatively visualize the information of syllable temporal sequences in songs, each cell in the SSM was color-coded according to the value of the similarity score. A similarity score of 0.595 was used as the threshold to distinguish similar or different syllables. For the quantitative analysis of syllable temporal structures, the occurrence rate of characteristic patterns of binarized "2 row × 2 column" cells in the SSMs was calculated. For the binarized patterning of 2 × 2 cells in the SSMs, the R software program was used to find the most similar binarized pattern for each 2 × 2 cell in the SSM from 12 possible patterns. The "motif" pattern was defined as a "paired-syllables transition," indicating the existence of 2 successive syllables that were different but with the same sequential order in 2 songs. This can be illustrated by "song 1 [··A B·····] vs. song 2 [····A B····]" (in this case, A and B represent 2 different syllables). The "repetition" pattern was a case of the existence of the "repetitive-syllables transition" by 2 successive identical or very similar syllables in 2 songs: for instance, "song 1 [·····A A··] vs. song 2 [···A A·····]." The mean of the occurrence rate of the motif and repetition patterns and their CVs from 10 total SSMs per an individual animal were used for statistical analyses.

For song sequence analysis, song consistency was measured (23). Sequence consistency is calculated as the sum of typical transitions per bout divided by the sum of total transitions per bout. This measures how consistently the bird sings the same transitions over several bouts. Syllable identification was performed and aligned by 2 different researchers without information on individual birds. For highly variable syllables, on the basis of acoustic morphology on the spectrogram and sequential position between singing rendition, we categorized them as identical syllables.

To measure the dynamics of syllable acoustic changes between 2 time points, we quantified changes in syllable acoustic features and syllable populations as 2D scatter density plots. Syllable duration (milliseconds; denoted as m) and mean FM (denoted as n) were measured to calculate the K–L distance (54, 55), which was adapted as a way to measure the distance between 2 sets of syllable populations by comparing their probability density distributions. Syllable segmentation was performed manually for all syllables on a SASLab spectrogram after turning the amplitude intensity to the maximum in order to clarify any continuities/discontinuities in syllable boundaries. In all, 150 syllables were used to generate a 2D density scatter plot. The probability density functions of each set of syllables were estimated at 2 different time points a and b, as Qa and Qb for the 2 time points, and the K–L distance score was then calculated to compare the density functions. If we let $qa(m, n)$ and $qb(m, n)$ denote the estimated probabilities for bin ($m = 20$, $n = 5$) for time points a and b, respectively, then the K–L distance between Qa and Qb is defined as follows:

$$DKL(Qa||Qb) = \sum_{m=a}^{M} \sum_{n=a}^{N} qa(m,n) \log_2 \frac{qa(m,n)}{qb(m,n)}.$$

A larger value for the K–L distance corresponds to a lower similarity between the distributions of 2 sets of syllable populations at different time points. Thus, a K–L distance of 0 indicates a perfect match between 2 sets of syllable populations. These behavioral analyses were performed as blind, without information of the residual number of $HVC_{(X)}$ neurons of each individual.

Effects of $HVC_{(X)}$ lesions on variability of song acoustic structure was calculated by the 2 measures, "within-syllable variability" and "cross-rendition variability" of the FF in song syllables (43). We randomly chose ∼50 song motifs recorded on the prelesion day and those recorded on the postlesion day, and extracted only syllables that had clear and flat harmonic structure. For each syllable rendition, a trajectory of FF was obtained in a sound segment of harmonic structure as in a previous study (41). Briefly, spectrograms were calculated using a Gaussian-windowed short-time Fourier transform (σ = 1 ms) sampled at 8 kHz, and a trajectory of the FF (the first harmonic frequency) was obtained by calculating the FF in individual time bins. For a subset of syllables that exhibits relatively low signal-to-noise ratios in the first harmonic

frequency, the second or upper harmonic frequency was used to quantify the FF trajectory. For each syllable, FF trajectories of all renditions were aligned by the onset of the syllables, based on amplitude-threshold crossings, and flat portions (≥25 ms) of FF trajectories were used for further analysis. We first removed the modulation of FF trajectories that was consistent across renditions by calculating residual FF trajectories as percent deviation from the mean trajectory across renditions. We then obtained within-syllable variability by calculating the SD of FF within each FF trajectory and averaging it across all renditions. To obtain cross-rendition variability, mean FF in each FF trajectory was calculated, and then the SD of mean FF across all renditions was computed. All raw data related with song analyses can be found in Dataset S1.

In Situ Hybridization. NTS cDNA fragments used for the synthesis of in situ hybridization probes were cloned from a whole-brain cDNA mixture of a male zebra finch. Total RNA was transcribed to cDNA using SuperScript Reverse Transcriptase (Invitrogen) with oligo-dT primers. The cDNAs were amplified by PCR using oligo DNA primers directed to conserved regions of the coding sequence from the National Center for Biotechnology Information (NCBI) cDNA database (accession no. NM_001245684). PCR products were ligated into the pGEM-T Easy plasmid (Promega). The cloned sequences were searched using NCBI BLAST/BLASTX to compare with homologous genes with other species, and genome loci were identified using BLAT of the University of California, Santa Cruz, Genome Browser. For FISH, digoxigenin-labeled riboprobes were used. The number of $HVC_{(X)}$ neurons was estimated as the average NTS^+ cells per mm^2 in both hemispheres of individuals. On the basis of the value of NTS^+ cells per mm^2, the degree of ablation of $HVC_{(X)}$ neurons in individual birds was calculated as a normalized value (in percentage) with the average of NTS^+ cells per mm^2 of control birds. See *SI Appendix, Supplementary Text*, for details.

AAV Construction. All of the viral ITR-flanked genomes used in this study were of the scAAV vector type (44). The pscAAV-GFP vector containing a CMV promoter was obtained from Addgene (#32396). AAV plasmids containing Cre and DIO (double-floxed inverted ORF)/FLEx (Flip excision) inserts were obtained from Dr. Kenta Kobayashi from the National Institute of Physiological Sciences (Okazaki, Japan) and subsequently cloned into the pscAAV vector plasmids after amplification of the Cre and DIO/FLEx sequences by primers containing the corresponding restriction enzymes in the target plasmid. To cell-specifically ablate the $HVC_{(X)}$ cells, a combination of dtA and constitutively active caspase 3 was used (45–47). Diphtheria toxin was cloned from pAAV-mCherry-FLEx-dtA (Addgene; #58536) by primers with specific enzyme sites and inserted into the previously constructed scAAV-DIO/FLEx. Owing to the restricted carrying capacity of the pscAAV vector, it became necessary to generate a constitutively active caspase 3 (47) by insertional mutagenesis of rAAV-flex-taCasp3-Tevp obtained from Gene Therapy Center Vector Core at the University of North Carolina at Chapel Hill. This insertion consisted of the substitution of valine with glutamic acid at residue 266 of the protein, with subsequent amplification and cloning into an scAAV-DIO/FLEx vector. AAVs were produced in-house using AAVpro 293T (Takara) cells transfected with a polyethyleneimine-condensed recombinant DNA mixture, based on a protocol kindly provided to us by the V. Gradinaru Laboratory (California Institute of Technology, Pasadena, CA). See *SI Appendix, Supplementary Text*, for details.

Surgery. Virus injection surgeries were performed on a custom-modified stereotaxic apparatus under 0.6 to 2.0% isoflurane anesthesia. To locate HVC and area X, both stereotaxic coordinates from the midsagittal sinus "y point" (0 mm rostral–caudal and 2.0 to 2.2 mm medial–lateral from the y point for HVC, 7.8 mm rostral–caudal and 1.5 mm medial–lateral from the y point for area X) and electrophysiological measurements using 1 M NaCl backfilled glass capillaries attached to a recording-capable Nanoject II (Drummond) were used. The location of injection sites for juvenile birds was slightly different (roughly 0.3 mm shallower for area X and closer to the midsagittal sinus for HVC), and special care was taken to shorten the surgery time as much as possible. The viral solution (virus titer 5.0×10^{12} to 5.1×10^{13} Vg/mL, a total of 1 μL in each area X, and 800 nL in each HVC) was injected with a pressure Nanojector II. Deafening surgery was performed on the birds by cochlear extirpation after crystallization at phd 104 to 110 for the adult deaf $HVC_{(X)}$ ablation experiment. The birds were anesthetized with pentobarbital (6.48 mg/mL; 60 μL/10 g body weight) by i.p. injection. After fixing the head in a custom-made stereotaxic apparatus with ear bars, a small window was made through the neck muscle and the skull near the end of the elastic extension of the hyoid bone. A small hole was then made in the cochlear dome. The cochlea was pulled out with a fine hooked wire. The removed cochleae were confirmed by visual inspection under a dissecting microscope. After cochlear removal, the birds recovered on a heat pad before being put back in their cages.

ACKNOWLEDGMENTS. We thank Masahiko Kobayashi for initial AAV production and Keiko Sumida for her excellent bird care and breeding. This work was supported by Ministry of Education, Culture, Sports, Science and Technology/Japan Society for the Promotion of Science KAKENHI Grant JP17H01015 (to K.O. and K.W.) and by 4903-JP17H06380, JP19H04888, JP17K19629, and JP18H02520 (to K.W.).

1. O. Hikosaka, K. Nakamura, K. Sakai, H. Nakahara, Central mechanisms of motor skill learning. *Curr. Opin. Neurobiol.* **12**, 217–222 (2002).
2. J. Tanji, Sequential organization of multiple movements: Involvement of cortical motor areas. *Annu. Rev. Neurosci.* **24**, 631–651 (2001).
3. M. S. Fee, J. H. Goldberg, A hypothesis for basal ganglia-dependent reinforcement learning in the songbird. *Neuroscience* **198**, 152–170 (2011).
4. T. Teşileanu, B. Ölveczky, V. Balasubramanian, Rules and mechanisms for efficient two-stage learning in neural circuits. *eLife* **6**, e20944 (2017).
5. N. Fujii, A. M. Graybiel, Representation of action sequence boundaries by macaque prefrontal cortical neurons. *Science* **301**, 1246–1249 (2003).
6. X. Jin, R. M. Costa, Shaping action sequences in basal ganglia circuits. *Curr. Opin. Neurobiol.* **33**, 188–196 (2015).
7. T. D. Barnes, Y. Kubota, D. Hu, D. Z. Jin, A. M. Graybiel, Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories. *Nature* **437**, 1158–1161 (2005).
8. A. A. Kozhevnikov, M. S. Fee, Singing-related activity of identified HVC neurons in the zebra finch. *J. Neurophysiol.* **97**, 4271–4283 (2007).
9. H. Fujimoto, T. Hasegawa, D. Watanabe, Neural coding of syntactic structure in learned vocalizations in the songbird. *J. Neurosci.* **31**, 10023–10033 (2011).
10. Q. Li *et al.*, Refinement of learned skilled movement representation in motor cortex deep output layer. *Nat. Commun.* **8**, 15834 (2017).
11. X. Jin, R. M. Costa, Start/stop signals emerge in nigrostriatal circuits during sequence learning. *Nature* **466**, 457–462 (2010).
12. E. D. Stefanova, V. S. Kostic, L. Ziropadja, M. Markovic, G. G. Ocic, Visuomotor skill learning on serial reaction time task in patients with early Parkinson's disease. *Mov. Disord.* **15**, 1095–1103 (2000).
13. D. B. Willingham, W. J. Koroshetz, Evidence for dissociable motor skills in Huntington's disease patients. *Psychobiology* **21**, 173–182 (1993).
14. P. Marler, Birdsong and speech development: Could there be parallels? *Am. Sci.* **58**, 669–673 (1970).
15. M. Konishi, The role of auditory feedback in the control of vocalization in the white-crowned sparrow. *Z. Tierpsychol.* **22**, 770–783 (1965).
16. M. S. Brainard, A. J. Doupe, Auditory feedback in learning and maintenance of vocal behaviour. *Nat. Rev. Neurosci.* **1**, 31–40 (2000).
17. A. J. Doupe, P. K. Kuhl, Birdsong and human speech: Common themes and mechanisms. *Annu. Rev. Neurosci.* **22**, 567–631 (1999).
18. E. D. Jarvis, Learned birdsong and the neurobiology of human language. *Ann. N. Y. Acad. Sci.* **1016**, 749–777 (2004).
19. A. C. Yu, D. Margoliash, Temporal hierarchical control of singing in birds. *Science* **273**, 1871–1875 (1996).
20. 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).
21. M. S. Fee, A. A. Kozhevnikov, R. H. Hahnloser, Neural mechanisms of vocal sequence generation in the songbird. *Ann. N. Y. Acad. Sci.* **1016**, 153–170 (2004).
22. S. W. Bottjer, E. A. Miesner, A. P. Arnold, Forebrain lesions disrupt development but not maintenance of song in passerine birds. *Science* **224**, 901–903 (1984).
23. C. Scharff, F. Nottebohm, A comparative study of the behavioral deficits following lesions of various parts of the zebra finch song system: Implications for vocal learning. *J. Neurosci.* **11**, 2896–2913 (1991).
24. F. Sohrabji, E. J. Nordeen, K. W. Nordeen, Selective impairment of song learning following lesions of a forebrain nucleus in the juvenile zebra finch. *Behav. Neural Biol.* **53**, 51–63 (1990).
25. A. R. Pfenning *et al.*, Convergent transcriptional specializations in the brains of humans and song-learning birds. *Science* **346**, 1256846 (2014).
26. F. Nottebohm, T. M. Stokes, C. M. Leonard, Central control of song in the canary, *Serinus canarius*. *J. Comp. Neurol.* **165**, 457–486 (1976).
27. D. Aronov, A. S. Andalman, M. S. Fee, A specialized forebrain circuit for vocal babbling in the juvenile songbird. *Science* **320**, 630–634 (2008).
28. 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).
29. D. S. Vicario, F. Nottebohm, Organization of the zebra finch song control system: I. Representation of syringeal muscles in the hypoglossal nucleus. *J. Comp. Neurol.* **271**, 346–354 (1988).
30. J. M. Wild, The avian nucleus retroambigualis: A nucleus for breathing, singing and calling. *Brain Res.* **606**, 319–324 (1993).

31. P. Dutar, H. M. Vu, D. J. Perkel, Multiple cell types distinguished by physiological, pharmacological, and anatomic properties in nucleus HVc of the adult zebra finch. *J. Neurophysiol.* **80**, 1828–1838 (1998).
32. M. Kubota, I. Taniguchi, Electrophysiological characteristics of classes of neuron in the HVc of the zebra finch. *J. Neurophysiol.* **80**, 914–923 (1998).
33. J. F. Prather, S. Peters, S. Nowicki, R. Mooney, Precise auditory-vocal mirroring in neurons for learned vocal communication. *Nature* **451**, 305–310 (2008).
34. G. F. Lynch, T. S. Okubo, A. Hanuschkin, R. H. Hahnloser, M. S. Fee, Rhythmic continuous-time coding in the songbird analog of vocal motor cortex. *Neuron* **90**, 877–892 (2016).
35. M. A. Long, D. Z. Jin, M. S. Fee, Support for a synaptic chain model of neuronal sequence generation. *Nature* **468**, 394–399 (2010).
36. M. A. Picardo *et al*., Population-level representation of a temporal sequence underlying song production in the zebra finch. *Neuron* **90**, 866–876 (2016).
37. K. Hamaguchi, K. A. Tschida, I. Yoon, B. R. Donald, R. Mooney, Auditory synapses to song premotor neurons are gated off during vocalization in zebra finches. *eLife* **3**, e01833 (2014).
38. D. Vallentin, M. A. Long, Motor origin of precise synaptic inputs onto forebrain neurons driving a skilled behavior. *J. Neurosci.* **35**, 299–307 (2015).
39. C. Scharff, J. R. Kirn, M. Grossman, J. D. Macklis, F. Nottebohm, Targeted neuronal death affects neuronal replacement and vocal behavior in adult songbirds. *Neuron* **25**, 481–492 (2000).
40. A. S. Andalman, M. S. Fee, A basal ganglia-forebrain circuit in the songbird biases motor output to avoid vocal errors. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 12518–12523 (2009).
41. J. D. Charlesworth, T. L. Warren, M. S. Brainard, Covert skill learning in a cortical-basal ganglia circuit. *Nature* **486**, 251–255 (2012).
42. E. Hisey, M. G. Kearney, R. Mooney, A common neural circuit mechanism for internally guided and externally reinforced forms of motor learning. *Nat. Neurosci.* **21**, 589–597 (2018).
43. 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).
44. D. M. McCarty, P. E. Monahan, R. J. Samulski, Self-complementary recombinant adeno-associated virus (scAAV) vectors promote efficient transduction independently of DNA synthesis. *Gene Ther.* **8**, 1248–1254 (2001).
45. D. M. Lin, V. J. Auld, C. S. Goodman, Targeted neuronal cell ablation in the *Drosophila* embryo: Pathfinding by follower growth cones in the absence of pioneers. *Neuron* **14**, 707–715 (1995).
46. E. Foster *et al*., Targeted ablation, silencing, and activation establish glycinergic dorsal horn neurons as key components of a spinal gate for pain and itch. *Neuron* **85**, 1289–1304 (2015).
47. J. Walters *et al*., A constitutively active and uninhibitable caspase-3 zymogen efficiently induces apoptosis. *Biochem. J.* **424**, 335–345 (2009).
48. A. Kageyama, I. Kusano, T. Tamura, T. Oda, T. Muramatsu, Comparison of the apoptosis-inducing abilities of various protein synthesis inhibitors in U937 cells. *Biosci. Biotechnol. Biochem.* **66**, 835–839 (2002).
49. N. Komatsu, T. Oda, T. Muramatsu, Involvement of both caspase-like proteases and serine proteases in apoptotic cell death induced by ricin, modeccin, diphtheria toxin, and pseudomonas toxin. *J. Biochem.* **124**, 1038–1044 (1998).
50. T. F. Roberts *et al*., Identification of a motor-to-auditory pathway important for vocal learning. *Nat. Neurosci.* **20**, 978–986 (2017).
51. R. Imai *et al*., A quantitative method for analyzing species-specific vocal sequence pattern and its developmental dynamics. *J. Neurosci. Methods* **271**, 25–33 (2016).
52. S. Kojima, M. H. Kao, A. J. Doupe, Task-related "cortical" bursting depends critically on basal ganglia input and is linked to vocal plasticity. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 4756–4761 (2013).
53. F. Ali *et al*., The basal ganglia is necessary for learning spectral, but not temporal, features of birdsong. *Neuron* **80**, 494–506 (2013).
54. W. Wu, J. A. Thompson, R. Bertram, F. Johnson, A statistical method for quantifying songbird phonology and syntax. *J. Neurosci. Methods* **174**, 147–154 (2008).
55. E. Ohgushi, C. Mori, K. Wada, Diurnal oscillation of vocal development associated with clustered singing by juvenile songbirds. *J. Exp. Biol.* **218**, 2260–2268 (2015).
56. L. Kubikova *et al*., Basal ganglia function, stuttering, sequencing, and repair in adult songbirds. *Sci. Rep.* **4**, 6590 (2014).
57. M. H. Kao, A. J. Doupe, M. S. Brainard, Contributions of an avian basal ganglia-forebrain circuit to real-time modulation of song. *Nature* **433**, 638–643 (2005).
58. B. P. Ölveczky, A. S. Andalman, M. S. Fee, Vocal experimentation in the juvenile songbird requires a basal ganglia circuit. *PLoS Biol.* **3**, e153 (2005).
59. M. S. Brainard, A. J. Doupe, Interruption of a basal ganglia-forebrain circuit prevents plasticity of learned vocalizations. *Nature* **404**, 762–766 (2000).
60. M. S. Fee, C. Scharff, The songbird as a model for the generation and learning of complex sequential. *ILAR J.* **51**, 362–377 (2010).
61. J. T. Sakata, M. S. Brainard, Online contributions of auditory feedback to neural activity in avian song control circuitry. *J. Neurosci.* **28**, 11378–11390 (2008).
62. L. Xiao *et al*., A basal ganglia circuit sufficient to guide birdsong learning. *Neuron* **98**, 208–221.e5 (2018).
63. V. Gadagkar *et al*., Dopamine neurons encode performance error in singing birds. *Science* **354**, 1278–1282 (2016).
64. M. Luo, L. Ding, D. J. Perkel, An avian basal ganglia pathway essential for vocal learning forms a closed topographic loop. *J. Neurosci.* **21**, 6836–6845 (2001).
65. N. I. Denisenko-Nehrbass, E. Jarvis, C. Scharff, F. Nottebohm, C. V. Mello, Site-specific retinoic acid production in the brain of adult songbirds. *Neuron* **27**, 359–370 (2000).
66. T. C. Roeske, C. Scharff, C. R. Olson, A. Nshdejan, C. V. Mello, Long-distance retinoid signaling in the zebra finch brain. *PLoS One* **9**, e111722 (2014).
67. B. B. Scott, T. Gardner, N. Ji, M. S. Fee, C. Lois, Wandering neuronal migration in the postnatal vertebrate forebrain. *J. Neurosci.* **32**, 1436–1446 (2012).
68. R. Kawai *et al*., Motor cortex is required for learning but not for executing a motor skill. *Neuron* **86**, 800–812 (2015).
69. J. Graff-Radford *et al*., The neuroanatomy of pure apraxia of speech in stroke. *Brain Lang.* **129**, 43–46 (2014).
70. D. Moser, A. Basilakos, P. Fillmore, J. Fridriksson, Brain damage associated with apraxia of speech: Evidence from case studies. *Neurocase* **22**, 346–356 (2016).
71. P. A. Alm, Stuttering and the basal ganglia circuits: A critical review of possible relations. *J. Commun. Disord.* **37**, 325–369 (2004).
72. S. Krishnan, K. E. Watkins, D. V. M. Bishop, Neurobiological basis of language learning difficulties. *Trends Cogn. Sci.* **20**, 701–714 (2016).
73. P. Deriziotis, S. E. Fisher, Speech and language: Translating the genome. *Trends Genet.* **33**, 642–656 (2017).
74. O. Tchernichovski, F. Nottebohm, C. E. Ho, B. Pesaran, P. P. Mitra, A procedure for an automated measurement of song similarity. *Anim. Behav.* **59**, 1167–1176 (2000).