Identification of a motor-to-auditory pathway important for vocal learning

Todd F Roberts1,5, Erin Hisey2,5, Masashi Tanaka2, Matthew G Kearney2, Gaurav Chattree1, Cindy F Yang3,4, Nirao M Shah4 & Richard Mooney2

Learning to vocalize depends on the ability to adaptively modify the temporal and spectral features of vocal elements. Neurons that convey motor-related signals to the auditory system are theorized to facilitate vocal learning, but the identity and function of such neurons remain unknown. Here we identify a previously unknown neuron type in the songbird brain that transmits vocal motor signals to the auditory cortex. Genetically ablating these neurons in juveniles disrupted their ability to imitate features of an adult tutor's song. Ablating these neurons in adults had little effect on previously learned songs but interfered with their ability to adaptively modify the duration of vocal elements and largely prevented the degradation of songs' temporal features that is normally caused by deafening. These findings identify a motor to auditory circuit essential to vocal imitation and to the adaptive modification of vocal timing.

Learning complex motor sequences, such as those required for speech or musical expression, depends on rapid and precise sensorimotor interactions. Auditory-to-motor interactions are evident in the important role that auditory experience plays in speech and musical learning. Motor-to-sensory interactions are manifest in the motor-related signals that are readily detected in many sensory brain regions, including the auditory system1–4. Such motor-to-sensory (i.e., forward) signals are thought to facilitate skill-learning, including speech and musical learning, through various processes including cancelation of motor-related sensory feedback and predictive computations that give rise to learning-related error signals5–13. The neurons and circuits that convey forward signals important to vocal learning remain unknown.

The zebra finch is an oscine songbird that exhibits many striking parallels to humans in its capacity for vocal learning14. Zebra finches copy the temporal and spectral features of a tutor’s song during a juvenile sensitive period, use auditory feedback to maintain stable adult song and can adaptively modify the timing and frequency of individual song elements in response to aversive auditory cues15–18. The zebra finch’s brain contains a specialized network for singing, including the auditory system1–4. Such motor-to-sensory forward interactions appear to facilitate learning, including speech and musical learning, through various processes including cancelation of motor-related sensory feedback and predictive computations that give rise to learning-related error signals5–13. The neurons and circuits that convey forward signals important to vocal learning remain unknown.

Here we show that PNs that link HVC to Av are distinct from HVC RA and HVC X cells. These HVC RA and HVC X cells convey motor-related information to the auditory system, play an essential role in juvenile song copying and are necessary to the adaptive modification of temporal but not spectral features of the adult’s song. These findings provide, to our knowledge, the first neuronal and circuit candidates for a forward mechanism important to vocal learning.

RESULTS
A distinct neuron type projects from HVC to the auditory system
To identify the HVC neurons projecting to Av, we injected variously colored fluorescent retrograde tracers into Av, RA and Area X (Fig. 1a,b; n = 7 adult birds (>100 d after hatching (dph)), n = 3 juvenile birds (60 dph); Av, dextran, Alexa Fluor 594; RA, dextran, Alexa Fluor 488; and Area X, fast blue; all three regions were accurately targeted in 70% of cases, 9 of 14 hemispheres in adult birds and 5 of 6 hemispheres in juvenile birds). All injections that successfully
targeted Av, as determined by retrograde labeling in two of its other known afferents (the thalamic nucleus uvaeformis (Uva) and telencephalic nucleus interface (NIf)), also retrogradely labeled a sparsely distributed population of HVC neurons. In contrast, injections targeting regions immediately dorsal or ventral to Av did not label HVC neurons in either juvenile or adult animals (n = 6 hemispheres from 6 animals). Notably, HVC neurons that were labeled by tracer injections in Av were not co-labeled by tracers injected in either Area X or RA (Fig. 1b; n = 14 hemispheres from 10 animals). These findings suggest that the axonal projections from HVC to Av arise from a previously unrecognized PN type (i.e., the HVCAv neuron) and not from the axon collaterals of HVCRA or HVCX neurons.

To better characterize HVCAv neuronal morphology, we injected fluorescent retrograde tracers or a retrogradely transported GFP-expressing virus (AAV2/9-GFP) into Av in a separate cohort of male zebra finches (n = 20 birds). Confocal analysis in fixed tissue revealed that HVCAv neurons possessed stellate dendrites with sparsely distributed spines, oblong cell bodies and an axon that exited along HVC’s rostroventral border (Fig. 1c; dendritic radial extent: 140.21 ± 4.98 μm, n = 29 cells from 3 birds; 0.16 spines per μm, n = 460 spines across 2.929 μm of dendrite, n = 10 cells from 2 birds; soma diameter: 18.61 ± 0.87 μm, 29 cells from 3 birds). Although the extent of dendritic labeling using retrograde expression of GFP could be variable and should be interpreted cautiously, the mean radius of HVCAv cell dendrites (140.21 ± 4.98 μm) was intermediate between reported values for HVCX and HVCRA cells (168 ± 4 μm and 123 ± 0.9 μm, respectively)30. Analysis of GFP-expressing HVCAv neurons revealed that the long axes of their cell bodies overlapped with the mean diameters of HVCX and HVCRA cell bodies, whereas the mean somatic areas of HVCAv cell bodies overlapped with those of HVCRA but not HVCX.

**Figure 1** A distinct class of projection neurons links the song premotor nucleus HVC to the auditory region Av. (a) Schematic showing the design of retrograde-labeling experiments. Different fluorescent retrograde tracers were injected into Av (dextran, Alexa Fluor 594), RA (dextran, Alexa Fluor 488) and Area X (fast blue). (b) Parasagittal section through retrogradely labeled HVC (left) reveals that neurons projecting into Av (red) are distinct from neurons projecting to Area X (blue) or RA (green). Scale bar, 100 μm. (c) Representative images of individual HVCAv neurons retrogradely labeled by AAV2/9 GFP injection into Av. Scale bar, 15 μm. (d) Visualized whole-cell current-clamp recordings from retrogradely labeled HVC projection neurons in brain slices in response to 500-ms current injection (−50 and +200 pA; dashed line = −80 mV). (e–g) HVCAv neurons differ from HVCX neurons in their intrinsic properties (*P = 0.05; **P = 0.01), including their spike adaptation rate (two-sample t(10) = 3.5, P = 0.005 for n = 5 HVCAv neurons vs. n = 9 HVCX neurons; two-sample t(8) = 2.6, P = 0.02 for n = 9 HVCAv neurons vs. n = 6 HVCRA neurons), (f) spike gain (two-sample t(10) = 3.3, P = 0.007 for n = 7 HVCAv neurons vs. n = 9 HVCX neurons; two-sample t(6) = 2.7, P = 0.03 for n = 9 HVCX neurons vs. n = 7 HVCRA neurons) and (g) input resistance (two-sample t(10) = 2.3, P = 0.04 for n = 7 HVCAv neurons vs. n = 9 HVCX neurons). (h) HVCAv neurons differ from HVCRA neurons in their input resistance (two-sample t(8) = 2.5, P = 0.04 for n = 7 HVCAv neurons vs. n = 7 HVCRA neurons; two-sample t(6) = 3.2, P = 0.01 for n = 9 HVCX neurons vs. n = 7 HVCRA neurons). (i) HVCAv neurons differ from HVCX and HVCRA neurons in the coefficient of variation (CV) of their DC-evoked interspike intervals (two-sample t(9) = 2.9, P = 0.03 for n = 7 HVCAv neurons and n = 6 HVCRA neurons; two-sample t(12) = 3.0, P = 0.01 for n = 7 HVCRA neurons and n = 9 HVCX neurons). Box plot elements in e–i represent mean and standard error.
cells, consistent with an oblong shape (Supplementary Fig. 1; GFP-expressing HVC<sub>C</sub> and HVC<sub>RA</sub> cells from Tschida and Mooney<sup>31</sup> were reanalyzed here).

We observed that HVC<sub>Av</sub> neurons constitute only a small percentage of HVC PN cells (130 ± 69.8 neurons per hemisphere; 26 hemispheres from 17 birds (i.e., ≤ 1%); estimates of HVC neuron numbers vary between 50,000 and 100,000 per hemisphere<sup>26,32</sup>) and appeared to be evenly distributed throughout HVC, resulting in a tiled appearance (Fig. 1b). In contrast to HVC<sub>X</sub> and HVC<sub>RA</sub> axons, HVC<sub>Av</sub> axons did not appear to extend local collaterals within HVC. Therefore, HVC<sub>Av</sub> neurons may function primarily to transmit signals to Av.

Neurons with different postsynaptic targets often exhibit physiological differences that correspond to different functional roles<sup>30,33–35</sup>. We used visually targeted whole-cell current-clamp recordings in brain slices to record from HVC<sub>Av</sub>, HVC<sub>X</sub> or HVC<sub>RA</sub> neurons retrogradely labeled with fluorescent tracers. We found that HVC<sub>Av</sub> neurons exhibited no spontaneous action potential activity but fired at high rates with little spike rate accommodation when depolarized (Fig. 1d). The intrinsic properties of HVC<sub>Av</sub> neurons differed from those of HVC<sub>X</sub> or HVC<sub>RA</sub> neurons with respect to spike adaptation, spike gain, input resistance, the amplitude of hyperpolarization-activated cation currents (<i>h</i>) and the coefficient of variation in direct current (DC)-evoked interspike intervals (Fig. 1d–i). Therefore, HVC<sub>Av</sub> neurons can be distinguished from the other two HVC PN types by their axonal projections to the auditory system and by their intrinsic electrical properties.

**HVC<sub>Av</sub> neurons transmit song motor-related signals**

Electrical and optical recordings in freely singing birds indicate that HVC<sub>C</sub> and HVC<sub>RA</sub> neurons transmit motor-related activity<sup>24,36–39</sup>. To begin to characterize the functional properties of HVC<sub>Av</sub> neurons, we monitored their activity in freely singing adult zebra finches using in vivo calcium imaging methods. We expressed genetically encoded calcium indicators in HVC<sub>Av</sub> neurons by injecting a retrogradely transported virus (AAV2/9-CAG-GCaMP6s; n = 3 birds) into Av or, in one additional bird, by injecting a Cre-dependent virus (AAV2/1-FLEX-CAG-GCaMP6s) into Av and a retrogradely transported Cre-recombinase containing virus (AAV2/9-CMV-Cre) in Av. A 1-mm diameter gradient-index (GRIN) lens was implanted over HVC and calcium transients were detected using a miniaturized head-mounted fluorescent microscope (Fig. 2a,b; n = 4 adult birds). In all four birds, bulk signals were collected across the entire imaging field. In one of these four birds, putative cell bodies were discernible in the raw images (Supplementary Movie 1), which we subsequently subjected to automated region-of-interest (ROI) analysis using a modified constrained non-negative factorization (CNMF) algorithm adapted for single-photon data.<sup>40</sup>

Singing was accompanied by substantial increases in bulk calcium signals in HVC, whereas small changes or no changes were detected in HVC when birds were not vocalizing (49 of 49 songs in 4 birds showed > 2 standard deviations (s.d.) increase in fluorescence relative to a presong baseline, Fig. 2c and Supplementary Fig. 2). Calcium signals increased before song onset, consistent with a premotor signal (135.74 ± 54.93 ms before song onset for 49 songs from 4 birds, where signal onsets were defined as ∆F/F increases 2 s.d. above baseline). In the one bird in which calcium transients in putative cell bodies could be detected, different ROIs displayed different singing-related activity patterns, whereas activity patterns of individual ROIs exhibited consistent temporal profiles from one rendition of the song motif to the next (Fig. 2c–e and Supplementary Fig. 2). Although all ROIs showed increased activity during singing, increases in some ROIs

**Figure 2** HVC<sub>Av</sub> neurons transmit motor-related signals during song production. (a) Schematic of imaging experiments employing a miniature epifluorescence microscope (Inscopix) for in vivo calcium imaging of HVC<sub>Av</sub> neurons expressing GCaMP6s. (b) Top: bulk fluorescent signal measured in HVC at rest (left) and during singing (right). Bottom: bulk change in fluorescent signal during singing of multiple song bouts (song onset to song offset for each bout indicated by blue shading). Scale bars, 100 µm × 100 µm (top) and 1% and 2 s (bottom). (c) Average change in bulk fluorescence signals, measured during singing, in birds expressing GCaMP6s in HVC<sub>Av</sub> neurons with their hearing intact (top, n = 4), after deafening (middle, n = 2) and in a bird not injected with GCaMP (bottom, n = 1). Onset and offset of song bout indicated by blue shading. Black line represents mean change in fluorescence and gray shading represents standard error. Scale bars, 2% and 1 s. (d) Example field of GCaMP-labeled HVC<sub>Av</sub> neurons imaged through miniature microscope in a singing zebra finch. Putative individual cells, which were identified with CNMF analysis, are indicated by dashed colored outlines. Scale bars, 100 µm × 100 µm. (e) Activity of 5 putative HVC<sub>Av</sub> neurons indicated in d during a single song bout comprising two motifs (sound spectrogram at top; color in the sound spectrogram indicates relative power with black indicating the lowest power (during silent intervals) and red indicating the highest relative power within the vocal range). Trace colors correspond to colored dashed circles in d. (f) Singing-related activity of 5 different putative HVC<sub>Av</sub> neurons from the same bird from which d and e were collected, several days after deafening; activity is aligned to a single bout comprising two motifs. Scale bars in e and f, 20 arbitrary activity units and 500 ms.
preceded song onset whereas increases in others occurred during the motif (Fig. 2e and Supplementary Fig. 2). The temporal profiles of all ROIs isolated using CNMF were prolonged and in many cases persisted following song offset (Fig. 2e). The prolonged transients measured in individual ROIs are likely a characteristic of the indicator rather than a special property of HVCAv cells, because the distribution of decay-time constants of calcium transients measured in putative HVCAv neurons was indistinguishable from the distribution of decay-time constants measured from a mixed population of GCaMP6s-expressing HVCAv neurons imaged during singing in another bird (Supplementary Fig. 2). Finally, the summed activity of putative HVCAv cell bodies in the fourth bird resembled the bulk signals collected in the other three birds (Supplementary Fig. 2), suggesting that the bulk signals largely reflected somatic activity of retrogradely labeled HVCAv neurons.

The increased activity before song onset is consistent with the idea that HVCAv neurons convey a motor-related signal. To further test this idea, we imaged HVCAv neuron activity in a subset of birds before and after deafening (n = 2 birds from which bulk signals were recorded, including one in which putative cell bodies were also detected; deafening was achieved by bilateral removal of the cochlea). We found that singing-related activity of HVCAv neurons persisted after deafening (Fig. 2f and Supplementary Movie 2; 29 of 30 songs in 2 birds showed > 2 s.d. increase in fluorescence over a presong baseline). Furthermore, the singing-related activity patterns observed after deafening were qualitatively similar to those measured previously with intact hearing, including transients that occurred before song onset and that were reproducible across motifs (Fig. 2f and Supplementary Fig. 2; bulk signal activity increased 112.11 ± 61.72 ms before song onset, n = 30 songs from 2 birds). These findings support the idea that singing-related activity in HVCAv neurons is motor-related.

Neurons important to song timing synapse on HVCAv neurons

Various findings indicate that HVCAv neurons are a source of premotor signals that contribute to song timing and indicate that HVCAv neurons convey song-motor-related information to the basal ganglia.24–36,39,41 As HVCRA and HVCAx axon collaterals form excitatory synapses with other HVCAv neurons,42 either or both PNs could transmit motor-related signals to HVCAv neurons. To clarify which of these PN types provide input to HVCAv neurons, we made visually targeted whole-cell recordings from retrogradely labeled HVCAv neurons in brain slices and antidromically stimulated HVCAa or HVCAx axon collaterals (Fig. 3). Antidromic stimulation of HVCAa axon collaterals routinely elicited excitatory and inhibitory synaptic currents in HVCAv neurons (Fig. 3a; excitatory postsynaptic current (EPSC), 10 of 11 cells; inhibitory postsynaptic current (IPSC), 9 of 11 cells; Supplementary Fig. 3; no input, 1 of 11 cells; n = 5 birds). In contrast, antidromic stimulation of HVCAx cells never evoked a synaptic response in HVCAv neurons (Fig. 3a; 0 of 8 cells from 4 birds), even though the same stimulation could elicit antidromic spikes in HVCAv neurons (Fig. 3b; n = 2 of 5 cells from 2 birds) and also evoked synaptic responses in both HVCAa (Fig. 3b; EPSC, 2 of 5 cells; IPSC, 2 of 5 cells; no input, 1 of 5 cells; n = 3 birds) and HVCAx neurons (Supplementary Fig. 3; IPSC, 3 of 5 cells; no input, 0 of 5 cells; n = 2 birds). Furthermore, although axons from two of HVC's afferents, the thalamic nucleus Uva and the NIf, enter HVC along its rostroventral border, electrical stimulation along this border failed to evoke synaptic responses in HVCAv neurons (Supplementary Fig. 3; n = 0 of 8 cells from 4 birds). Therefore, HVCAv neurons were specifically positioned within the HVC microcircuit to convey song-timing information from HVCAa neurons to the auditory system.

Genetically ablating HVCAv neurons disrupts juvenile song copying

Juvenile male zebra finches memorize a tutor song between 20 and 60 dph and then copy this memorized song model in a process of sensorimotor learning that spans 45–90 d (refs. 15,17). To determine the role of HVCAv neurons in song copying, we used an intersectional genetic method43 to selectively ablate these cells in juvenile birds after they had an opportunity to memorize a tutor song but before they had formed an accurate copy of the tutor song model (Fig. 4a,b). We injected a virus expressing a Cre-dependent form of caspase 3 (AAV2/1-FLEX-taCas3p–TEVp), which triggers cellular apoptosis43, into HVC and injected a retrogradely transported virus expressing Cre-recombinase (AAV2/9-CMV-Cre) into Av of juvenile male zebra finches (Fig. 4a,b; n = 5 male zebra finches; 35–40 dph when injected with viruses; viral expression required 2–4 weeks (see Online Methods)). Siblings that received either no virus or that were injected only with AAV2/1-FLEX-taCas3p–TEVp into HVC served as control groups. To quantify the effects of this cell-killing method, we injected a retrograde tracer into Av in adult birds that had been subjected to intersectional HVCAv lesions as juveniles and quantified the number of surviving HVCAv neurons. Intersectional expression of caspase 3 (Cas3) in HVCAv neurons of juveniles significantly reduced the adult number of HVCAv neurons compared to either of two control groups (Fig. 4c; the two control groups had similar numbers of HVCAv neurons and were subsequently treated as a single group).

All birds were raised to adulthood and their adult songs were compared to their tutors' songs to quantify song copying (Online Methods). Genetically ablating HVCAv neurons in juveniles severely disrupted their ability to copy a tutor song (Fig. 4d,e; n = 5 experimental birds, 46.3% similarity to tutor song; n = 8 control siblings, 72.9% similarity to tutor song).
Av neurons in juvenile zebra finches impairs their ability to copy a tutor song. (a) Schematic showing viral strategy for genetically ablating HVC_{Av} neurons with a Cre-dependent form of caspase 3 (Cas3). (b) Timeline of experiments. Viral injections of AAV2/1-FLEX-taCas3-TEVp into HVC and AAV2/9-Cre-GFP into Av were made in juvenile male zebra finches between 35 and 40 dph. Birds were separated from their parents and other siblings starting at 60 dph and the quality of song copying was measured at >90 dph, when song copying is complete. (c) The number of retrogradely labeled HVC_{Av} neurons (HVC_{Av} neurons per hemisphere) in adult birds that were injected as juveniles with both AAV2/1-FLEX-taCas3-TEVp into HVC and AAV2/9-Cre-GFP into Av (red, HVC_{Av}) was significantly lower than the number of HVC_{Av} neurons in found in adult siblings injected as juveniles only with AAV2/1-FLEX-taCas3-TEVp into HVC (blue, Sib-Cas3 Ctrl) or adult siblings that were not injected with any virus (green, Sib Ctrl; Mann-Whitney U_8 = 86, z = −3.31, P = 0.0009 for HVC_{Av} lesioned (red) vs. Sib-Cas3 (blue); Mann-Whitney U_{16} = 16, z = 3.44, P = 0.0006 for HVC_{Av} lesioned (red) vs. Sib Ctrl (green)). (d) Adult birds injected with both Cas3 and Cre at 35–40 dph sang poorer copies of their tutors’ songs compared to their control siblings (combined group of Sib-Cas3 and Sib Ctrl birds), as measured by song similarity-percentage scores (Online Methods) but showed no difference from control siblings in measures of percent accuracy and sequential match percentage (Online Methods). Tukey boxplots in c and d show the first and third quartiles of the data; the median, mean and 1.5x interquartile range are shown by bands, dots and whiskers, respectively ( * P = 0.05; ** P = 0.01). (e) Example sound spectrograms from the tutor (top), from one adult pupil (control) with a normal complement of HVC_{Av} neurons and from two adult siblings (siblings 1 and 2) in which HVC_{Av} neurons were genetically ablated during juvenile life. Color in the sound spectrogram indicates relative power, with black indicating the lowest power (during silent intervals) and red indicating the highest relative power within the vocal range. Individual syllables in the tutor’s song and the copied versions of these syllables in the pupil’s adult songs, defined by >70% accuracy scores, are color-coded by lines under each of the tutor’s syllables. The control sibling’s song was 78.5% similar to the tutor; this bird had 106.5 HVC_{Av} neurons per hemisphere (neurons averaged from the two hemispheres). Sibling 1’s song was 50.6% similar to the tutor; this bird had 42.5 HVC_{Av} neurons. Sibling 2’s song was 39.5% similar to tutor; this bird had 41.5 HVC_{Av} neurons. Scale bar, 100 ms; ordinate spans 0–10 kHz. Colored bars correspond to individual syllables copied from the tutor’s song. (f) Similarity of the adult bird’s copy to its tutor’s song plotted as a function of the average surviving numbers of HVC_{Av} neurons counted in the two hemispheres in each bird (R^2 = 0.5, P = 0.004, n = 13 birds, data are mean ± s.e.m.). (g) Accuracy of the adult song copy as a function of average number of HVC_{Av} neurons counted in the two hemispheres from each bird (R^2 = 0.09, P = 0.9, n = 13 birds, data are mean ± s.e.m.). (h) Sequential match of the adult copy to its tutor song as a function of average surviving number of HVC_{Av} neurons counted from the two hemispheres in each bird (R^2 = 0.03, P = 0.25, n = 13 birds; data are mean ± s.e.m.).

Ablating HVC_{Av} neurons protects song’s temporal features in deafened adults

Adult zebra finches use auditory feedback to maintain stable songs, as deafening triggers the spectral degradation of individual syllables and also destabilizes syllable sequences over a period of days to weeks. Furthermore, much of the neural circuitry implicated in song copying in juveniles also plays a role in feedback-dependent song maintenance in adults. For example, lesioning LMAN, the premotor output of a corticobasal ganglia pathway important to juvenile song learning, has little effect on adult song but largely prevents deafening-induced song degradation. To test whether HVC_{Av} neurons are also important to song maintenance, we first selectively ablated HVC_{Av} neurons in adult zebra finches and recorded their songs for 6 to 10 weeks. All

similarly; Mann-Whitney U = −3.8, z = 2.56, P = 0.01). Within-bird comparisons indicated that similarity of the adult pupil’s song to the tutor song correlated with the number of surviving HVC_{Av} neurons in the pupil’s brain (Fig. 4f-h). Most pupils were able to copy at least some syllables from their tutor (Fig. 4c,g and Supplementary Fig. 4); 4 of 5 birds copied at least one syllable from their tutor, with syllable copying determined by >70% accuracy; Online Methods), suggesting that reducing the numbers of HVC_{Av} neurons does not block the capacity or motivation to copy individual syllables. Nonetheless, lesioned birds as a group failed to faithfully copy their tutors, indicating that HVC_{Av} neurons are necessary to normal levels of song copying.

A remaining issue is whether the song copying deficits resulted from ablating HVC_{Av} neurons and not from the ablation of other HVC neurons. In a separate group of birds, we tested whether the intersectional methods used to target HVC_{Av} cells also kill HVC_{X} cells, some of which extend axons through or near Av and might plausibly take up the AAV2/9-CMV-Cre virus from the injection site in Av. However, we found that the number of HVC_{X} neurons was unaffected by the intersectional method used to kill HVC_{Av} neurons (Supplementary Fig. 5 and Online Methods; n = 2 birds). Therefore, deficits in song copying most likely arose from the selective loss of HVC_{Av} neurons.
birds continued to sing highly stereotyped songs following ablation of HVC_{Av} neurons (Fig. 5a,b; two-sample \( t_{12} = -1.4, P = 0.16 \) for \( n = 12 \) HVC_{Av}-lesioned birds versus 12 HVC_{Av}-intact birds). Although there was a slight—yet significant—decrease in song self-similarity and motif duration (Supplementary Fig. 6; two-sample \( t_{11} = 4.34, P = 0.001 \) comparing percent self-similarity before lesioning to percent self-similarity after lesioning in experimental birds.), the trial-to-trial variability in spectral features of individual syllables and the timing of intersyllable gaps were unaffected by bilateral ablation of HVC_{Av} neurons (syllable frequency: two-sample paired \( t_{4} = 0.2, P = 0.8 \); intersyllable gaps: two-sample paired \( t_{4} = 1.3, P = 0.27 \); \( n = 5 \) birds in each group). In contrast, unilateral HVC lesions, made by injecting AAV2/1-FLEX-taCas3-TEVp and AAV2/9-CMV-Cre into HVC, disrupted song production in a manner similar to the effects of unilateral HVC lesions (Supplementary Fig. 4). In summary, in adult birds with intact hearing, selectively ablating HVC_{Av} neurons showed little effect on acute song performance, trial-to-trial song variability or longer-term song maintenance.

We then deafened a cohort of adult male zebra finches in which we had previously bilaterally ablated HVC_{Av} neurons and also deafened another cohort of animals in which HVC was left intact (\( n = 7 \) animals in each group). Visual inspection of sonograms suggested that the songs of birds with HVC_{Av} lesions remained more stable following deafening than the songs of deafened birds with an intact HVC (Fig. 5c). To better characterize these contrasting effects, we quantified temporal changes by calculating differences in syllable transition matrices generated before and 10 weeks after deafening, and we quantified the spectral changes using a self-similarity score over the same period (Fig. 5d,e and Online Methods). Selectively ablating HVC_{Av} neurons largely prevented deafening-induced destabilization of syllable sequences, as revealed by lower differences in matrix scores in deafened birds with HVC_{Av} lesions (Fig. 5d,e; two-sample \( t_{12} = -4.8, P = 0.0004 \) for changes in song syllable transition matrix before and after deafening from \( n = 7 \) lesioned versus 7 intact birds). In contrast, HVC_{Av} lesions did not prevent spectral degradation of syllables following deafening, as self-similarity scores before and 10 weeks after deafening were similar in the experimental and control groups (Fig. 5e; two-sample \( t_{12} = 1.9, P = 0.07 \) comparing self-similarity scores from lesioned versus intact birds). (f) Within-syllable comparisons reveal a significant increase in entropy before and 10 weeks after deafening in HVC_{Av}-lesioned birds (two-sample \( t_{12} = -4.0, **P = 0.0019 \) ). Tukey boxplots show the first and third quartiles of the data; the median and 1.5x interquartile range are shown by bands and whiskers, respectively.

Figure 5 Intersectional ablation of HVC_{Av} neurons in adult birds does not disrupt song production but does attenuate deafening-induced degradation of song’s temporal features. (a) Schematic showing viral strategy for genetically ablating HVC_{Av} neurons with a Cre-dependent caspase 3 in adult birds. (b) Sonograms recorded from an adult zebra finch before and 10 weeks after bilateral ablation of HVC_{Av} neurons reveal no evidence of song degradation (two-sample \( t_{12} = -1.4, P = 0.16 \) for comparisons in song self-similarity from \( n = 12 \) HVC_{Av} lesioned birds vs. \( n = 12 \) control adult birds). Color in the sound spectrogram indicates relative power, with black indicating the lowest power (during silent intervals) and red indicating the highest relative power within the vocal range. (c) Example sonograms recorded before and 10 weeks after deafening in adult zebra finches in which HVC_{Av} neurons had been previously ablated (left) or were intact (right). Scale bars, 50 msec. (d) Top: song transition matrix scores, which serve as a primary measure of the linearity and stability of syllable sequences, are calculated by summing probability scores along the boxed region (yellow oval, core syllable transition sequence before deafening) within the matrix. Scale bar, difference matrix score (Online Methods). Bottom: subtracting the song matrix at 10 weeks after deafening from the predeafening song matrix revealed little change in syllable sequence linearity and stability in birds with bilateral HVC_{Av} lesions, whereas the difference matrix from deafened but not lesioned controls showed marked changes. (e) Ablating HVC_{Av} neurons reduced the magnitude of changes to syllable sequences but not to syllable features following hearing loss, as revealed by group data plots of difference matrix scores and self-similarity scores for syllables before and after 10 weeks after deafening (\( n = 7 \) HVC_{Av}-lesioned birds, \( n = 7 \) HVC_{Av} intact birds; two-sample \( t_{12} = -4.8, P = 0.0004 \) comparing difference matrix scores from lesioned vs. intact birds; two-sample \( t_{12} = 1.9, P = 0.07 \) comparing self-similarity scores from lesioned versus intact birds). (f) Within-syllable comparisons reveal a significant increase in entropy before and 10 weeks after deafening in HVC_{Av}-lesioned birds (two-sample \( t_{12} = -4.0, **P = 0.0019 \) ). Tukey boxplots show the first and third quartiles of the data; the median and 1.5x interquartile range are shown by bands and whiskers, respectively.
consistent with the idea that these neurons serve a specialized role in mediating feedback-dependent changes to song timing.

A role for HVCAv neurons in song-timing plasticity
To further explore the role of HVCAv neurons in feedback-dependent song plasticity, we used singing-contingent playback of white noise to induce adult birds to modify either the spectral (i.e., fundamental frequency, or ‘pitch’) or temporal (i.e., song element duration) features of their songs, before and after bilaterally ablating HVCAv neurons. These experiments exploit slight trial-to-trial variations in syllable pitch and song element timing: targeting white noise playback to pitch or timing variants that fall above or below a user-set threshold drives the bird to adaptively modify pitch or timing, subsequently reducing the number of renditions that trigger noise playback. Notably, adult finches can independently modify either syllable pitch or song element duration using these protocols and can also rapidly recover their original pitch and timing once noise playback is discontinued.16,18

We measured rates of pitch and song element duration learning and subsequent recovery before and after genetic ablation of HVCAv neurons. Rates of pitch learning and recovery did not differ before and after genetic ablation of HVCAv neurons (Fig. 6a–c; n = 5 birds). In contrast, almost all (4 of 5) birds in which HVCAv neurons were ablated displayed much slower rates of song element duration learning and all (5 of 5) were severely impaired in their ability to recover their original song element timings after noise playback was discontinued (Fig. 6d–f). Taken together with the adult deafening studies, these findings indicate that HVCAv neurons played a specialized role in the feedback-dependent modification and recovery of song timing in adult zebra finches.

DISCUSSION
This study identifies a previously unknown neuron type in the song nucleus HVC that transmits vocal motor-related signals to the auditory system and that plays an important role in juvenile and adult forms of vocal learning. These HVCAv neurons form the cellular linchpin of a vocal premotor-to-auditory microcircuit: they are privileged synaptic targets of HVCRA cells, which generate premotor signals important to song timing,20,24,25,36, and they transmit information to Av, a discrete nucleus in the CM, part of the avian secondary auditory cortex.29 Because CM is reciprocally connected with other auditory regions implicated in song learning,29,47 the HVC-to-Av synapse may afford a crucial site where vocal motor-related signals interact with the auditory system to affect song learning and maintenance.

Figure 6 Intersectional ablation of HVCAv neurons in adult zebra finches interferes with feedback-dependent plasticity of song element timing but not syllable pitch. (a) Schematic of pitch-learning model. (b) Syllable pitch (i.e., fundamental frequency) measured during pitch-contingent playback of white noise (WN, red) and subsequent recovery of original pitch after noise was discontinued, measured in an adult male zebra finch before (unfilled circles) and after (gray filled circles) genetic ablation of HVCAv neurons (average s.d. of pitch during baseline, WN and recovery days was 9.2 Hz, 9.2 Hz and 10.6 Hz, respectively, before lesions; and 13.0 Hz, 10.9 Hz and 11.3 Hz, respectively, after lesioning). (c) Change in syllable pitch per day (Hz per d) before, during and after WN playback across birds before (black) and after (red) genetic ablation of HVCAv neurons. (d) Schematic of song-element duration learning model. (e) Change in song-element duration in response to song-element duration-contingent WN playback (red) and subsequent recovery of timing measured in an adult male zebra finch before (unfilled circles) and after (gray filled circles) genetic ablation of HVCAv neurons (average s.d. of song element duration during baseline, WN and recovery days was 3.9 ms, 4.1 ms and 3.7 ms, respectively, before lesions; and 3.3 ms, 3.8 ms and 5.2 ms, respectively, after lesioning). (f) Change in song element duration per day (ms per d) before, during and after WN playback across birds before (black) and after (red) genetic ablation of HVCAv neurons reveals a significant disruption in recovery of song timing (n = 5 birds, paired t test, t2 = 4.8, **P = 0.008). Tukey boxplots in c and f show the first and third quartiles of the data; the medians, means and 1.5× interquartile ranges are shown by bands, dots and whiskers, respectively.
The anatomical placement of the HVCAv neuron in the sensorimotor hierarchy of the songbird brain is reminiscent of motor-to-auditory cortical projections in the mammalian brain. In the mammalian brain, these motor cortical inputs provide a source of movement-related signals that can suppress and modulate auditory processing in a state- and movement-dependent manner. In humans, pro-}

The current findings support the idea that HVCAv cells transmit motion-related signals to the auditory system that are important for learning and modifying song timing. First, HVCAv cells receive input from HVCX cells, which are a source of premotor signals for song timing, but they do not appear to receive input from HVCX or cells from either Uva or NIf, two afferents that convey motor- and auditory-related signals to HVC. This privileged organization could enable HVCAv cells to provide a direct source of song-timing information to the auditory system. Second, we found that selectively ablating HVCAv neurons in juveniles markedly interferes with song imitation while partially sparing their ability to copy individual syllables. Similarly, killing these cells in adults strongly impeded deafening-induced degradation of songs' temporal features, even though spectral degradation following deafening was still evident. Finally, most adults with reduced numbers of HVCAv cells displayed slower rates of song-element duration learning, and all of them failed to recover normal song timing following vocalization-contingent feedback perturbation. Together, these observations support the idea that HVCAv cells are important to the learning and modification of song timing but not to pitch learning, two processes that prior studies indicate are mediated by distinct neural pathways. Therefore, functional dissection of circuits downstream of the HVCAv cell may help identify mechanisms that underlie the learning and feedback-dependent modification of vocal timing.

The region immediately downstream of HVCAv cells, Av, is embedded in CM, a part of the avian secondary auditory cortex that contains a variety of distinct neuron types, including some that exhibit evidence of song-related corollary discharge. The placement of HVCAv cells in the motor-to-sensory hierarchy and their motor-related activity during singing makes them a likely source of these corollary discharge signals. Notably, a key function of corollary discharge signals in forward models of sensorimotor integration is to help generate error signals important to learning and maintaining skilled behaviors. In fact, neurons that detect errors in singing-related auditory feedback exist in CM and other regions of the songbird auditory forebrain, and they may receive indirect input from HVCAv cells. Finally, some neurons in these downstream auditory regions innervate midbrain dopaminergic neurons, which are speculated to generate the error-correction signals that drive adaptive vocal plasticity in both juvenile and adult songbirds. Therefore, a pathway involving relatively few synapses may enable motor-related signals emanating from HVC to influence error-detection and error-correction processes necessary to vocal learning.

In this light, it is worthwhile to contrast the effects of selective HVCAv cell ablation with lesions made in LMAN, the output nucleus of the corticalosubicular ganglia pathway. Although both types of lesions disrupt juvenile song copying and slightly shorten motif duration, LMAN lesions in adults block pitch learning but do not severely interfere with noise-contingent modulation or recovery of song timing, whereas HVCAv cell lesions interfere with the modification and recovery of song timing but leave pitch learning and recovery unaffected. Another notable difference is that HVCAv cell lesions selectively preserve song's temporal features following adult deafening, whereas LMAN lesions block deafening-induced degradation of song's spectral and temporal features. Thus, although both HVCAv and LMAN lesions provide some protection from deafening-induced song degradation, these protective effects may result from disruptions at different stages of sensorimotor processing. Specifically, we speculate that HVCAv lesions attenuate deafening-induced song degradation by reducing the strength of motor-based signals necessary to error detection, thus minimizing the strength of error signals that would normally arise in the absence of singing-related auditory feedback. In contrast, LMAN lesions occlude deafening-induced changes by preventing these error signals from engaging the error correction machinery in the corticobasal ganglia network.

An important future will be to analyze the circuit computations that result from interactions between HVCAv neurons and downstream neurons in the auditory forebrain. One possibility is that motor-related signals from HVCAv neurons gate transmission from the auditory forebrain to HVC in a vocalization-dependent manner. Consistent with this idea, auditory inputs to HVC are rapidly gated off during singing, and some auditory neurons in CM display reduced sensitivity to auditory stimulation during vocalization. Another related possibility is that vocal motor-related signals from HVC to the auditory system function as part of a predictive mechanism, as theorized to occur in forward models of vocal learning.

In this scenario, HVCAv neurons transmit a corollary discharge signal to the auditory system that is shaped by prior vocal experience to predict features of auditory feedback that are themselves a consequence of HVC’s premotor activity. Circuits downstream of HVC that integrate such predictive signals with vocalization-related feedback could give rise to the putative error signals that have been detected in CM and other parts of the auditory forebrain. Of course, song learning depends not only on the individual's experience of his vocalizations but also on auditory experience of an appropriate song model, the latter of which is stored as a long-term auditory memory. Thus, a distinct possibility is that vocalization-dependent signals transmitted by HVCAv neurons to the auditory system reactivate this song memory in a context-dependent manner and temporally register it with vocalization-related auditory feedback, facilitating the comparisons between song memories and song performance that are necessary to vocal learning.

METHODS

Methods, including statements of data availability and any associated accession codes and references, are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

The authors thank D. Schneider, K. Tschida, T. Warren and S. Lisberger for reading and commenting on the manuscript; K. Hamaguchi (Duke University Medical Center) for software support; W.Y. Peh (Duke University Medical Center) for help with calcium imaging experiments; B.P. Olveczky (Harvard University) for providing Conditional Auditory Feedback software; J. Baltzegar, M. Booze (Duke University Medical Center), J. Holdway and A. Guerrero (UT Southwestern Medical Center) for animal husbandry and laboratory support. This research was supported by grants from the National Science Foundation (R.M. (IOS-1354962),
22. Keller, G.B. & Hahnloser, R.H. Neural processing of auditory feedback during vocal practice in a songbird. Nature 457, 187–190 (2009).
23. Heinke-Maldonado, T.H., Mathalon, D.H., Gray, M. & Ford, J.M. Fine-tuning of auditory cortex during speech production. Psychophysiology 42, 180–190 (2005).
24. Hahnloser, R.H., Kozhevnikov, A.A. & Fee, M.S. An ultra-sparse code underlies the generation of neural sequences in a songbird. Nature 419, 65–70 (2002).
25. Hamaguchi, K., Tanaka, M. & Mooney, R. A distributed recurrent network contributes to temporally precise vocalizations. Neuron 91, 680–693 (2016).
26. Alvarez-Buylla, A., Kinn, J.R. & Nottebohm, F. Birth of projection neurons in adult avian brain may be related to perceptual or motor learning. Science 249, 1444–1446 (1990).
27. Sotiradj, F., Nordeen, E.J. & Nordeen, K.W. Selective impairment of song learning following lesions of a forebrain nucleus in the juvenile zebra finch. Behav. Neural Biol. 53, 51–63 (1993).
28. Scharff, C. & Nottebohm, F. 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).
29. Akutagawa, E. & Konishi, M. New brain pathways found in the vocal control system of a songbird. J. Comp. Neurol. 518, 3086–3110 (2010).
30. Mooney, R. Different substrates underlie song selectivity in identified HVC neurons of the zebra finch. J. Neurosci. 20, 5420–5436 (2000).
31. Takeda, K.A. & Mooney, R. Defeasing drives cell-type-specific changes to dendritic spines in a sensorimotor nucleus important to learned vocalizations. Neuron 73, 1028–1039 (2012).
32. Wild, J.M., Williams, M.N., Howie, G.J. & Mooney, R. Calcium-binding proteins define interneurons in HVC of the zebra finch (Taenopygia guttata). J. Comp. Neurol. 483, 76–90 (2005).
33. Durat, P., Vu, H.M. & Perkel, D.J. Multiple cell types distinguished by physiological, pharmacological, and anatomic properties in nucleus HVC of the adult zebra finch. J. Neurophysiol. 80, 1828–1838 (1998).
34. Hattori, A.M. & Nelson, S.B. Layer V neurons in mouse cortex projecting to different targets have distinct physiological properties. J. Neurophysiol. 98, 3330–3430 (2007).
35. Kim, E.J., Juvinnett, A.L., Kyuwa, E.M., Jacobs, M.W. & Callaway, E.M. Three types of cortical layer 5 neurons that differ in brain-wide connectivity and function. Neuron 88, 1253–1267 (2015).
36. Long, M.A., Jin, D.Z. & Fee, M.S. Support for a synaptic chain model of neuronal sequence generation. Nature 468, 394–399 (2010).
37. Liberti, W.A., Ill et al. Unstable neurons underlie a stable learned behavior. Nature Neurosci. 19, 1665–1671 (2016).
38. Picard, M.A. et al. Population-level representation of a temporal sequence underlying song production in the zebra finch. Neuron 13, 1509–1522 (2008).
39. Hamaguchi, K., Tsodica, J.A., Yoon, I., Donald, B.R. & Mooney, R. Auditory synapses to song premotor neurons are gated off during vocalization in zebra finches. Elife 3, e01833 (2014).
40. Zhou, P., Resendez, S.L., Stuber, G.D., Kass, R.E. & Paninski, L. Efficient and accurate extraction of in vivo calcium signals from microendoscope video data. Preprint at https://arxiv.org/abs/1605.07266 (2016).
41. Prather, J.F., Peters, S., Nowicki, S. & Mooney, R. Precise auditory-vocal mirroring in premotor neurons for learned vocal communication. Nature 451, 305–310 (2008).
42. Meroney, R. & Prather, J.F. The HVC microcircuit: the synaptic basis for interactions between song motor and vocal plasticity pathways. J. Neurosci. 25, 1952–1964 (2005).
43. Yang, C.F. et al. Sexually dimorphic neurons in the ventromedial hypothalamus govern mating in male mice with sexes and aggression in males. Cell 153, 896–909 (2013).
44. Nordeen, K.W. & Nordeen, E.J. Auditory feedback is necessary for the maintenance of stereotyped song in adult zebra finches. Behav. Neural Biol. 57, 58–66 (1992).
45. Brainard, M.S. & Doupe, A.J. Interruption of a basal ganglia-fornix circuit prevents plasticity of learned vocalizations. Nature 404, 762–766 (2000).
46. Williams, H., Crane, L.A., Hale, T.K., Esposito, M.A. & Nottebohm, F. Right-side dominance for song control in the zebra finch. J. Neurobiol. 23, 1006–1020 (1992).
47. Baster, E.E. et al. A synaptic basis for auditory-vocal integration in the songbird. J. Neurosci. 28, 1509–1522 (2008).
48. Nelson, A. et al. A circuit for motor cortical modulation of auditory cortical activity. J. Neurosci. 33, 14342–14353 (2013).
49. Mandelblat-Cerf, Y., Las, L., Denisenko, N. & Fee, M.S. A role for descending pathways in fine-tuning auditory cortical projections in songbird vocal learning. Elife 3 (2014).
50. Gale, S.D., Person, A.L. & Perkel, D.J. A novel basal ganglia pathway forms a loop linking a vocal learning circuit with its dopaminergic input. J. Comp. Neurol. 508, 824–839 (2008).
ONLINE METHODS

Animals. Juvenile and adult male zebra finches were obtained from the Duke University Medical Center or the UT Southwestern Medical Center breeding facilities. Experimental procedures were conducted in accordance with the National Institutes of Health guidelines and were reviewed by the Duke University Medical Center and UT Southwestern Medical Center Animal Care and Use Committees.

Anatomical pathway tracing and cell counts. Juvenile and young adult zebra finches (60–140 d) were anesthetized with 2% isoflurane gas before being placed in a custom stereotaxic apparatus. After applying a topical anesthetic (0.25% bupivacaine) and making a vertical incision in the skin over the skull, we made ∼1-mm craniotomies in the skull at a predetermined distance from the bifurcation of a major blood vessel ('Y sinus'; Av: 1.75 mm anterior, 2.0 mm lateral, 1 mm deep; RA: 1.0 mm posterior, 2.4 mm lateral, 2.0–2.5 mm deep; Area X: 5.1 mm anterior, 1.6 mm lateral, 3.0 mm deep). Using a glass pipette attached to a pressure injection system (Drummond Nanoject II), we gave bilateral injections of Alexa Fluor dextran 594 or 488 (10,000 MW; Life Technologies), fast blue (Sigma or Polysciences; three injections of 32.2 nL) or virally encoded GFP (AAV2/9. CMV-scGF; PennVector; 30 injections of 32.2 nL) into Av, RA and/or Area X. After tracer injections, craniotomies were sealed with bone wax and the incision site in the skin closed with a tissue adhesive (VetBond). After 6 d, or 1 month in the case of viral injections, birds were deeply anesthetized with an I.P. injection of pentobarbital solution (Euthasol) and perfused with 0.025 M phosphate-buffered saline followed by 0.05 M Alexa Fluor 488 or Alexa Fluor 594 solution contained (µM): 124 potassium-gluconate, 4 NaCl, 10 HEPES, 2 EGTA, 1.1 MgCl2 and 0.05 Alexa Fluor 594 dextran (pH 7.4, adjusted with KOH). For recordings of spikes in HVC neurons, the pipette solution contained (mM): 120 cesium-methanesulfonate, 4 NaCl, 10 HEPES, 2 EGTA, 2 MgCl2, 2 ATP-Mg, 0.3 GTP-Na3, 10 phosphocreatine-Na2 and 0.05 Alexa Fluor 594 hydroxide (pH 7.3, adjusted with KOH). For recordings of EPSCs in HVCAv neurons, the pipette solution contained (mM): 120 cesium-methanesulfonate, 4 NaCl, 10 HEPES, 2 EGTA, 2 MgCl2, 2 ATP-Mg, 0.3 GTP-Na3, 10 phosphocreatine-Na2, 5 QX-314 bromide and 0.05 Alexa Fluor 594 hydroxide (pH 7.3, adjusted with KOH). The pipettes were fabricated from borosilicate glass and the pipette resistance was 5–8 MΩ. Membrane potentials were corrected for measured liquid-junction potentials just before each recording.

Data were acquired with a Multichip 700B, low-pass filtered at 2.8 kHz and digitized at 20 kHz. Recordings with high series resistances of > 30 MΩ or with a large leak current (the resting current of < −100 pA at −70 mV) were excluded from the analysis. The series resistance was compensated for by 30%. At the beginning of recordings, 500-ms voltage pulses from −70 mV to −90 mV were applied to measure input resistance and Ith amplitude. The Ith current amplitude was calculated as the 480–500 ms value subtracted from the 80–100 ms value. To characterize the firing properties of HVC neurons, 500-ms current pulses were applied from the resting current of 0 pA. Spike adaptation was quantified as the difference between the first and last interspike interval at the mean firing rate of ∼40 Hz. Spike gain was calculated as a slope of a line fitting to the firing-rate–current relation (Hz/µA) in a range < 40 Hz. The CV interspike interval was calculated at the mean firing rate of ∼40 Hz. To examine synaptic connections from HVCAv to HVC neurons, HVCAv axons were antidromically stimulated with a concentric bipolar electrode (CBBRC75; FHC) placed midway between HVC and RA, and evoked EPSCs were recorded in HVC neurons voltage-clamped at −70 mV.

Genetic ablation of HVCAv neurons. Using a similar procedure to that described above for tracer injections, we anesthetized zebra finches (35–40 d for juvenile experiments, 105–110 d for adult experiments) and placed them in a stereotaxic apparatus. After making an incision in the skin, we made craniotomies over Av and HVC at predetermined distances from the Y sinus (Av: 1.75 mm anterior, 2.0 mm lateral, 1 mm deep; HVC: 2.4 mm lateral, 0.375 mm deep). To selectively ablate HVCAv cells, we gave birds bilateral injections of a retrogradely transported Cre construct (AAV2/9.CMV13.GFP::Cre.SV40; PennVector) into Av and a locally expressed Cre-dependent caspase construct (AAV2/1.IEcre.FLEX-Casp3-2A-TEV; Nirao Shah, UCSF) into HVC (13 injections of 32.2 nL of Cre, 28 injections of 32.2 nL of CS3). After injections, craniotomies were sealed with bone wax, the incision site closed with VetBond and birds were allowed to recover from anesthesia under a heat lamp. Birds were injected with Alexa Fluor 594 dextran into Av 6 d before perfusion, and HVCAv cell numbers were quantified by counting all retrogradely labeled HVCAv neurons in HVC. To examine whether viral HVCAv lesions inadvertently killed HVCX cells, some of which extend axons through or near Av, we made bilateral injections of retrogradely transported Cre into Av and unilateral injections of locally expressed Cre-dependent caspase virus into HVC. Birds were then injected with Alexa Fluor 594 dextran into Area X 6 d before perfusion, and HVCAv cell numbers were quantified by counting all retrogradely labeled HVCAv neurons in HVC.

Bilateral removal of cochlea. Adult male birds (100–140 d) were anesthetized with Equithesin and secured to a custom-made metal platform. The tympanic membrane and inner ear bones including the oval window were removed with forceps and the cochlea was removed with a small wire hook. Cochleae were then examined under the dissecting microscope to ensure complete removal. The incision site in the skin was closed with VetBond and birds recovered from anesthesia under a heat lamp.

Calcium imaging of HVCAv neurons. Using similar surgical procedures as those described in the ablation experiments above, we made bilateral injections of either a retrogradely transported GCaMP (AAV2/9.CAG.GCaMP6s; Penn Vector) into Av (54 injections of 9.2 nL of GCaMP, n = 3 birds) or a retrogradely transported Cre (AAV2/9.CMV.HL.GFP::Cre.SV40; Penn Vector) into Av (43 injections of 9.2 nL of Cre) and a locally expressed Cre-dependent GGaMP6s construct (AAV2/1.FLEX.CAG.GCaMP6s; Penn Vector) into HVC (54 injections of 9.2 nL of GCaMP, n = 1 bird). Three weeks after viral injection, birds were injected with Alexa Fluor 594 dextran into Area X to help visualize HVC. Four weeks after viral injection, the skull and dura over HVC were removed and Kwk-Sil (World Precision Instruments) was placed over the putative site of HVC. The bird was then imaged under an epifluorescence microscope (Zeiss 510) to determine the borders of HVC by the presence of Alexa Fluor 594 dextran in HVCAv cells. A 1-mm diameter gradient-index (GRIN) lens was then implanted on the dorsal surface of HVC within the identified boundaries. Next, a baseplate was implanted on the bird's skull to hold a miniature microscope (Inscopix, nVista HD) for imaging. After a period of recovery (~3 d), birds were placed in a recording chamber and the miniature microscope was attached to the baseplate for imaging. The activity of HVCAv neurons was then imaged, and data was collected using nVista software (Inscopix) (fluorescence) and a custom written acquisition program in LabVIEW (sound and frame times). The LED power was maintained between 10 and 20%, corresponding to 0.12 and 0.24 mW/mm2. The maximum field of view was 1,440 pixels × 1,080 pixels or 900 µm × 900 µm. Pixel size was 0.625 × 0.625 µm. Sound was sampled at 44 kHZ and imaging data were sampled between 10 Hz and 30 Hz corresponding to exposure times of 99.84 ms to 33.25 ms. Imaging data acquired above 10 Hz were down-sampled such that all analyses were performed on 10-Hz signals. Imaging sessions were conducted in 30-s trials in which LED power was on and data were recorded. Offline analysis of imaging data was performed using Mosaic (Inscopix) and custom software written in Matlab. Changes in fluorescence in the raw imaging data were quantified with ∆F/F0 calculated as (F − F0)/F0. Next, regions of interest (ROIs) were manually drawn, containing a large fraction of the brain surface (defined as 'bulk signal'), for cases in which signals resembling cells could be identified, a modified CNMF algorithm adapted for single-photon microendoscope data was used to automatically identify ROIs corresponding to putative individual neurons. Fluorescence traces for both bulk signals and putative single neurons were computed and then aligned to audio recordings of song. Song bouts were identified by visual inspection and were defined as sequences of introductory notes and syllables separated by gaps of less than 0.5 s. Song onset was defined as the first frame after
the onset of the first element of song and song offset as the first frame after the offset of the last element of song. Mean fluorescence responses related to singing were computed by averaging the bulk signal responses of every song bout before (n = 49 songs from 4 birds) and after deafening (n = 30 songs from 2 birds). In one control bird, the same procedures were performed in the absence of viral injection (n = 10 songs from 1 bird).

Behavioral analysis of song learning and song stability. Percentage-similarity accuracy and sequential match scores in Sound Analysis Pro were used to quantify the similarity of the pupil's song to that of its tutor in our juvenile CS3 experiments (http://soundanalysispro.com/manual-1/chapter-10-similarity-measurements/the-similarity-score). The percent similarity score compares large scale features of songs measured over 50- to 70-ms time windows by calculating Euclidean distances across mean values of song spectral features (pitch, FM, AM, Wiener entropy and goodness of pitch). This score reflects the percentage of the tutor's song that has been copied by the pupil. The accuracy score provides a measure of the local similarity of the pupil's song to the tutor's song as calculated over 5- to 10-ms time windows. The sequential match score provides a measure of how well the sequence of the tutor's song was copied by comparing the temporal ordering of the 50- to 70-ms time windows in the tutor's song to those in the pupil's song with an 80-ms alignment tolerance. We also used the percent similarity to quantify the similarity of the bird's song after CS3/Cre viral injection to its song before injection in our adult CS3 experiments.

In song-learning experiments, juveniles injected with AAV2/9.CMV.HI.GFP-Cre.SV40 in Av and AAV2/2.Ef1α.FLEX-Casp3-2A-TEV in HVC, as well as control siblings, including those injected with AAV2/1.Ef1α.FLEX-Casp3-2A-TEV alone in HVC and those not receiving any viral injections, were housed with their father until 60 d of age, at which point birds were raised in isolation until adulthood (90–100 d). To determine the similarity of the pupil's song to the tutor's song, we calculated similarity scores by comparing the adult pupil's song (90–100 d) to the song of their tutor (n > 30).

In adult HVC, ablation experiments, we recorded songs 1–2 d before viral injections then again 4 weeks after injection with AAV2/9.CMV.HI.GFP-Cre.SV40 in Av and AAV2/2.Ef1α.FLEX-Casp3-2A-TEV in HVC. A subset of these birds was subsequently deafened and their songs recorded 10 weeks later. We calculated baseline similarity scores to establish that the bird sang a relatively stable song before injection by comparing pairs of motifs before injection (n > 30 motif pairs). We then compared pairs of motifs before and after injection (n > 30 motif pairs) to determine whether the bird's song changed after HVC lesion. For the subset of deafened birds, motifs were compared immediately before and 10 weeks following deafening (n > 30).

To examine sequential changes in song syllable structure of deafened birds, we quantified changes in transitions from one song syllable to the next. Using custom-written scripts (Matlab) we classified individual syllables and calculated the probability of transition from one syllable to the next. After assembling a matrix of all syllable transitions for songs before deafening and after deafening, we subtracted the postdeafening matrix from the predeafening matrix. The difference matrix score was calculated by taking the average of the transition probabilities along the top-left-to-bottom-right diagonal of the difference matrix, which corresponds to the core syllable transition sequence of the bird's song. The resulting ‘difference matrix’ was used to examine the change in transition probability after deafening. For example, before deafening a bird sings 'A B C' with a 100% probability that syllable B follows syllable A, a 100% probability that syllable C follows B and a 100% probability that the end of the sequence follows C. These three probabilities lie along the left-top-to-bottom-right diagonal of the transition matrix and correspond to the core syllable transition sequence of the bird's song. If after deafening the probability that B would follow A was reduced from 100% to 50%, we would take the difference between the probability that B would follow A before deafening and the probability that B would follow A before deafening (i.e., P(A|B) before − P(A|B) after). We would then average the difference matrix scores from each transition to arrive at the overall difference matrix score (i.e., mean P(A|B) difference matrix score P(B|C) difference matrix score P(C|end) difference matrix score). The lower the overall difference matrix score, the less the bird's core syllable sequence changed after deafening.

Song feature modulation. A modified version of the Conditional Auditory Feedback software (LabVIEW) by Ali et al. was used to implement the song feature modulation protocol16. Real-time auditory input from the bird's cage was input to the program, which then took measurements (for example, duration or pitch) on a targeted region of the bird's song as it was sung. Detection was done by comparing the auditory input to a template, a sound file containing a recording of a portion of the bird's song (150–500 ms in length) that immediately preceded the desired target region. For pitch modulation, targeted regions were 5-ms components of the bird's song that had a well-defined pitch (i.e., harmonic stack). For timing modulation, targets were the region between the onsets of two syllables, i.e., a syllable + an intersyllable gap and ended at the start of the next syllable. Targets were chosen where onsets of the two syllables were well-defined with clear rapid rises in amplitude. This is the ‘song element’ referred to in the in Targets. For both pitch and timing modulation, also needed to be preceded by a sufficient amount of song to create a template allowing for detection with high sensitivity and specificity. Throughout all experiments we estimated that the true detection rate (number of measurements of the intended target + number of times the target was produced) remained above 75%. We also estimate that the false measurement rate (number of measurements of unintended regions / number of measurements made) remained below 25%. Modulation of the bird's song was then achieved by playing white noise into the bird's chamber whenever the measurement of the target region fell below a set threshold value (in all experiments we drove the measurement up). Initial threshold values were set at the average value of the measurement recorded over a 3-day baseline period. Threshold values were then dynamically adjusted by the program as the bird altered its song: whenever 80% of the last 200 measurements were above the threshold, the threshold was then set to the mean of the last 200 measurements, though only if this mean was above the current threshold.

The experimental protocol was as follows. Adult birds were initially injected bilaterally into Av with AAV2/9-Cre-GFP. Birds were then run through either pitch or timing modulation protocols. Before starting white noise, baseline measurements of the target were calculated over at least a 3-d period. White noise was then turned on to drive the value of the measurement up and was kept on until the daily average value of the measurement was at least 2 s.d. higher than the average value of the measurement over the baseline period. The s.d. used here was the s.d. of the measurements over the baseline period. Birds were then allowed to return passively (white noise turned off). Three birds were then run through the protocol again, this time to modulate the other feature (i.e., if they had just run through pitch modulation they were now run through timing modulation). Two birds only did timing modulation and two birds only did pitch modulation. After returning to baseline, birds were then injected with AAV2/1-flex-taCas3-3-TEV into HVC bilaterally. Birds were then given at least 3 weeks before starting the next arm of the experiment. Birds were then again run through whichever feature modulation protocols they had done previously. Caspase-injected birds did not have clear returns to baseline for timing features, so they were allowed to return (passive measurements) for as many days as they required before caspase injection plus an extra 5–10 d.

Statistical analysis. Male zebra finches were randomly assigned to experimental groups. No statistical methods were used to predetermine sample sizes, but our sample sizes are similar to those reported in previous publications. Data sets used for statistical comparisons were tested for normality and none of the data were excluded from analysis. Parametric (two-tailed two-sample t tests and paired t tests) or nonparametric tests (Mann-Whitney U tests) were used accordingly for statistical comparisons. Data collection and analysis were not performed blind to the conditions of the experiments. A Supplementary Methods Checklist is available.

Data availability. The data that support the findings of this study and any associated custom software used in data analysis are available from the corresponding authors upon reasonable request.

Code availability. Song-feature modulation code for pitch and temporal learning of song and analysis code for population calcium imaging data have been previously published. LabVIEW codes for acquiring imaging frames and sound for calcium imaging experiments are included in Supplementary Software.