



Published in final edited form as:

Nat Neurosci. 2018 April ; 21(4): 589–597. doi:10.1038/s41593-018-0092-6.

A common neural circuit mechanism for internally guided and externally reinforced forms of motor learning

Erin Hisey¹, Matthew Gene Kearney¹, and Richard Mooney¹

¹Department of Neurobiology, Duke University, Durham, North Carolina, USA

Abstract

The complex skills underlying verbal and musical expression can be learned without external punishment or reward, indicating their learning is internally guided. The neural mechanisms that mediate internally guided learning are poorly understood, but a circuit comprising dopamine-releasing neurons in the midbrain ventral tegmental area (VTA) and their targets in the basal ganglia (BG) are important to externally reinforced learning. Juvenile zebra finches copy a tutor song in a process that is internally guided and, in adulthood, can learn to modify the fundamental frequency (pitch) of a target syllable in response to external reinforcement with white noise. Here we combined intersectional genetic ablation of VTA neurons, reversible blockade of dopamine receptors in the BG, and singing-triggered optogenetic stimulation of VTA terminals to establish that a common VTA – BG circuit enables internally-guided song copying and externally reinforced syllable pitch learning.

Introduction

Some of the most complex and expressive behaviors, such as speaking or singing, depend on rapid and precise motor sequences that are learned with reference to internal guides and without reinforcement by external reward or punishment^{1,2}. The neural circuit mechanisms that underlie the internally guided learning of rapid and precise motor sequences are not well understood, but external reinforcement can drive the learning of relatively slow and simple behaviors by modulating the activity of ventral tegmental area (VTA) neurons that release dopamine in the basal ganglia (BG)^{3–10}. The rarity of well-documented forms of internally guided learning in non-human animals has complicated the analysis of its underlying neural mechanisms. Vocal learning in songbirds shares many parallels with human speech learning^{11,12}, including a developmental sensitive period when juvenile songbirds copy the

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use: http://www.nature.com/authors/editorial_policies/license.html#terms

*Corresponding Author: Richard Mooney, mooney@neuro.duke.edu.

Data and Code Availability Statement

The data and software code that support the findings of this study are available from the corresponding author upon reasonable request.

Author contributions: E.H. and R.M. designed experiments. E.H. and M.G.K. executed experiments and analyzed data. E.H. and R.M. wrote the manuscript.

A Reporting Summary is available online.

The authors declare no competing financial interests.

song of an adult tutor, a process that is internally guided^{13–16}. Moreover, adult songbirds can modify the fundamental frequency of individual song syllables when subjected to external reinforcement with singing-triggered noise, a process referred to as pitch learning^{17–19}. Consequently, studies in songbirds provide a unique opportunity for testing whether a common circuit mechanism involving dopamine-dependent signaling from the VTA to the BG is important to both internally guided and externally reinforced forms of motor learning.

The songbird brain is distinguished by a neural network for singing and song learning^{20,21}, including a specialized basal ganglia region (Area X) that is important to juvenile and adult forms of vocal plasticity^{22–24} (Figure 1A). Similar to the mammalian BG, Area X is densely innervated by neurons in the VTA and the substantia nigra pars compacta (SNc) that are positive for tyrosine hydroxylase (TH)²³, the synthetic enzyme for dopamine (the songbird VTA and SNc form a continuous group of cells that do not differ in their projection targets or innervation²⁵; here we refer to both VTA and SNc cells that project to Area X collectively as VTA_X neurons). Intriguingly, recent studies show that VTA_X neurons in adult zebra finches subjected to singing-contingent noise function to encode reward prediction error²⁶, a well-described property of mammalian VTA neurons that serves as an essential component in reinforcement learning^{7–10}. These parallels support a model in which dopamine release from VTA_X neurons is crucial to adult pitch learning, an idea that has gained support from a recent study showing that pharmacological lesions of TH⁺ terminals in Area X interfere with this form of song plasticity²⁷.

Despite these important advances in understanding the singing-related properties of VTA_X neurons and the role of TH⁺ terminals in Area X in adult pitch learning, several critical steps are necessary to better understand the cellular effectors and circuit mechanisms underlying internally guided and externally reinforced forms of vocal learning. One step is to confirm that VTA_X neurons and the dopamine receptors they activate in Area X are necessary to adult pitch learning, because the 6-OHDA treatment as applied in Area X may have damaged both dopaminergic and noradrenergic fibers and in any event did not identify the parent cell group of the affected fibers. Another critical step is to establish whether trial-by-trial variations in VTA_X terminal activity, as might arise in response to singing-contingent noise, are sufficient to drive adult pitch learning, as expected in a reinforcement learning framework. Finally, beyond testing the necessity and sufficiency of VTA_X neurons and dopamine signaling in Area X in adult pitch learning, whether these same cells and signaling pathways are important to juvenile song copying remains unexplored.

Results

VTA_X neurons are necessary for externally reinforced learning

We used an intersectional genetic method to selectively ablate VTA_X neurons (Figure 1B), allowing us to test their role in an adult form of vocal learning in which white noise is used to drive changes in the fundamental frequency (i.e., pitch) of a target syllable (“pitch learning”¹⁷, Figure 1C). We confirmed previous results²³ demonstrating that nearly 95% of VTA_X cells are likely to release dopamine by injecting retrograde tracer into Area X and finding extensive overlap between retrogradely labeled neurons in the VTA and cells positive for tyrosine hydroxylase (345/367 retrogradely VTA_X cells also positive for TH⁺; n = 3

example sections through VTA from 3 birds). To ablate VTA_X neurons, we injected young adult (n = 6; 95 ± 5 days post hatch (dph); mean ± S.E.M. unless otherwise noted) male zebra finches with a virally encoded Cre-dependent caspase (AAV2/1.EF1α.FLEX-Casp3-2A-TEV)²⁸ in the VTA and injected a retrogradely traveling virally encoded Cre in Area X (AAV2/9.CMV.HI.GFP-Cre.SV40; Figure 1B, D). Several (5 ± 1) days after the viral injections, prior to high levels of viral expression, we targeted a syllable in each bird's motif with pitch-contingent noise. Briefly, we measured the baseline variation in the target syllable's pitch and set a threshold within this distribution where pitch variants falling below the threshold triggered a brief, intense noise burst (Figure 1C; threshold was set at the 70th percentile of the pitch distribution (i.e., syllables that fell below the 70th percentile of the pitch distribution triggered white noise playback)). During this early period following viral injections, birds rapidly shifted the pitch of the target syllable to "escape" noise playback. Following four days of noise exposure, during which the threshold was adjusted upwards each day to drive continued pitch learning away from the baseline value, we discontinued noise playback and measured the rate and magnitude of recovery of pitch to the pre-treatment baseline. We then repeated these behavioral experiments one month later, after viral expression had ablated a proportion of VTA_X neurons, as determined by tracer injection into Area X and post hoc quantification of retrogradely labeled neurons in the VTA (Figure 1B).

Ablating VTA_X neurons significantly impaired adult pitch learning. Within bird comparisons revealed that the maximum amount of pitch learning one month after viral injections was significantly less than the maximum amount measured within the first week following these injections (Figure 1E, F; n = 6 adult birds; p = 0.042). Post hoc histological analysis revealed that the reduction in the rate of pitch learning measured at one month versus one week after viral injections was inversely correlated with the number of surviving VTA_X neurons (Figure S1A) and, in another adult bird, that intersectional VTA_X lesions reduced TH+ immunoreactivity in Area X (Figure S1B). Notably, genetically ablating VTA_X neurons did not affect the amount of recovery following pitch learning (Figure 1G) and did not significantly alter the trial-to-trial variability of the target syllable (Figure 1H). Moreover, genetically ablating VTA neurons that project to a region of the striatum medial to Area X, a manipulation that spared VTA_X neurons, had no effect on pitch learning (Figure S1C; n = 5 adult male zebra finches; mean of 3957 ± 169 VTA_X cells from four VTA_{MSl} birds versus a mean of 2795 ± 92 VTA_X cells for the six VTA_X – lesioned birds; two-tailed t-test: p = 0.0003). Furthermore, injections of either only AAV-Cre to Area X or only AAV-caspase to VTA also had no effect on pitch learning (Figure S1C; n = 2 adult male zebra finches). Therefore, a full complement of VTA_X neurons is necessary to enable normal levels of noise-induced pitch learning in adult male zebra finches.

Optogenetic stimulation of VTA terminals in Area X is sufficient to drive vocal learning

The behavioral effects of targeted VTA_X ablation indicate that these neurons are necessary to support normal levels of adult pitch learning but do not establish whether VTA_X activity by itself is sufficient to drive pitch learning in the absence of external reinforcement. To begin to resolve this issue, we bilaterally injected an adeno-associated virus containing a humanized Channelrhodopsin gene^{29–31} (AAV2/9.CAGChR2.mCherry, n = 2, or

AAV2/9.CAG-NeurexinChr2.YFP, $n = 6$) in the VTA of adult male zebra finches. After waiting several months to achieve robust expression of ChR2 in VTA terminals within Area X (VTA_X terminals), we bilaterally implanted optical fibers in Area X (Figure 2A, B; $n = 8$ animals; mean interval between viral injections in the VTA and fiber implantation in Area X: 137 ± 18 days; mean age at implantation: 252 ± 24 dph). In a subset of these birds (6/8), we used optrode recordings in Area X prior to fiber optic implantation to verify that brief light pulses (50–100 milliseconds; 473 nm) delivered in Area X evoked an increase in multiunit activity (Figure 2A; following behavioral experiments in all birds, histological methods were used to confirm ChR2 expression in the VTA and Area X (Figure 2B) and cannula placement over Area X). We then adapted the pitch learning protocol to optogenetically activate VTA_X terminals when the pitch of a target syllable fell either above or below a specified threshold. Analysis of unstimulated “catch” trials indicated that such pitch-contingent stimulation of VTA_X terminals applied over several days was sufficient to drive lasting changes in the pitch of the target syllable (Figure 2C; pulse duration, 50 milliseconds, 473 nm; threshold was set to apply stimulation to either the upper or lower 70% of the syllable distribution).

In contrast to experiments that used pitch-contingent noise to drive pitch learning^{17,19} (Figure 1C), the pitch of the target syllable shifted toward the frequency range that received optogenetic stimulation ($n = 10$ syllables from 8 birds, Figure 2C–G, Figure S2A–C). Similar to noise-driven pitch learning, the change in pitch occurred gradually during the first day of exposure, and the absolute change from baseline continued to increase following daily adjustments of the pitch threshold (Figure 2C; Figure S2C). The pitch distribution and mean pitch of the target syllables were significantly shifted from baseline following several days of optogenetic stimulation (Figure 2H–J; 5 ± 1 days, range: 4–10 days), whereas other syllables in the birds’ motifs were unaffected, regardless of their proximity to the target syllable (Figure S3A–C; $n = 7$ syllables from 5 birds). We also compared the pitch values of the first, middle and last third of the target syllables. We found that the pitch contours were modified differently across birds with a slight trend towards the largest changes in pitch occurring in the middle and the last third of the syllable (Figure S4A, B). In contrast to these effects on syllable pitch, optogenetic stimulation of VTA_X terminals had no acute effects on the pitch or trial-to-trial variability of the target syllable (Figure S4C–E). Moreover, almost all (6/8) birds subjected to syllable-triggered optogenetic stimulation of VTA_X terminals sang significantly more on the last day of stimulation than on the day before the beginning of light stimulation (Figure S4F). Therefore, pitch-contingent optogenetic stimulation of VTA_X terminals is sufficient to drive pitch learning in adult male zebra finches and also appears to positively reinforce singing more generally.

Although VTA_X neurons are TH+ and thus likely to release dopamine, they may also release other transmitters, as described for mammalian VTA terminals in the BG³². Therefore, we combined microdialysis methods to reversibly block D1-type dopamine receptors in Area X with pitch-contingent optogenetic stimulation of VTA_X terminals ($n = 3$ adult male zebra finches). We found that when a D1R antagonist (SCH22390) was infused into Area X, optogenetic stimulation of VTA_X terminals induced little or no pitch learning, whereas the same stimulation could drive robust pitch learning when saline was infused into Area X either before or after this drug treatment day (Figure S2D–G). Therefore, the pitch-

contingent optogenetic stimulation of VTA_X terminals is sufficient to drive pitch learning in adult zebra finches, and microdialysis experiments performed here in a small number of animals suggests that this form of adult learning depends on D1 receptor signaling in Area X.

Notably, the pitch distribution and mean pitch of a target syllable did not shift when VTA_X terminals were optogenetically stimulated regardless of the target syllable's pitch, consistent with the idea that performance-contingent variations in VTA_X terminal activity are necessary to drive pitch learning (Figure 2I, J; 100% contingency, $n = 2$ syllables from 2 birds previously described that displayed pitch learning in response to a 70% stimulation contingency). In contrast to birds injected with AAV-ChR2 constructs, syllable-triggered pitch-contingent illumination of GFP-expressing VTA_X terminals or of Area X in birds that had not been injected with any virus had no effect on the pitch of the target syllable (Figure 2I, J; $n = 3$ syllables from 2 birds injected in the VTA with AAV2/9.CAG-GFP and $n = 2$ syllables from 2 birds that had not been injected with virus; 70% contingency).

VTA_X neurons project almost exclusively to Area X

Taken together, the intersectional cell ablation and optogenetic experiments strongly implicate VTA_X terminals in Area X as a critical component of learning-related vocal plasticity. Although VTA_X neurons do not provide appreciable input to surrounding striatal regions²³, one potential confound is that they may extend collaterals to other song-related brain nuclei, the inadvertent destruction or stimulation of which might account for the learning-related effects we observed. To explore this possibility, we used dual retrograde tracing methods to determine whether VTA_X neurons also innervate other forebrain song nuclei that are densely innervated by TH⁺ fibers^{33,34} (HVC (used here as a proper name), nucleus interface of the nidopallium (Nif), and the lateral magnocellular nucleus of anterior nidopallium (LMAN); Figure S5). We detected only a small percentage of double-labeled VTA_X neurons following these dual tracer injections (percentage of VTA_X cells that also project to: HVC: 1.7% (23/1323), 3 hemispheres, 2 birds; Nif: 5.6% (95/1696), 3 hemispheres, 2 birds; LMAN, 4.8% (74/1549), 3 hemispheres, 2 birds). Thus, VTA_X neurons likely influence adult pitch learning through their terminals in Area X.

D1-type receptors in Area X are necessary for externally reinforced learning

In mammals, the VTA mediates reinforcement learning by activating dopamine receptors in the BG^{8,9}, and we showed that pitch-contingent optogenetic stimulation of VTA_X terminals in Area X can drive pitch learning through a D1-receptor-dependent mechanism. To determine whether the VTA influences adult pitch learning through dopamine receptors, we used microdialysis methods^{19,35} to reversibly block different dopamine receptor types in Area X of adult male zebra finches while targeting syllables with pitch-contingent noise (Figure 3A, $n = 6$, 99 ± 9 dph). Bilateral infusion of a D1R receptor antagonist (SCH23390)^{36,37} into Area X prevented pitch learning (Figure 3B–E) without affecting trial-to-trial variability of the target syllable's pitch (Figure S6). Similar treatment with sulpiride, a D2R antagonist³⁶, exerted variable effects on the daily amount of pitch learning but also strongly reduced the total amount of singing, without affecting trial-to-trial song variability (Figure 3F, G; Figure S6; $n = 6$ birds; one of these birds sang too infrequently (<20 times per

day) to support pitch learning experiments). When we corrected for this reduced amount of singing by estimating the amount of pitch learning per rendition of the target syllable, we found that the rate of pitch learning during sulpiride treatment was either enhanced or unchanged in four birds and reduced in the other bird (Figure 3H), suggesting that VTA terminals may act selectively through D1 receptors in Area X to drive pitch learning in the adult zebra finch. Moreover, although D1 and D2 receptors can be co-expressed in single medium spiny neurons within Area X^{38,39}, the current study indicates that they mediate distinct behavioral functions, reminiscent of the functional segregation observed in the mammalian striatum^{40–42}.

VTA_X cells and D1 receptors in Area X are necessary for internally reinforced learning

Whereas adult pitch learning is driven by exposure to loud noise, an extrinsic cue, juvenile song copying progresses without any external reinforcement¹³. Therefore, a remaining issue is whether the mechanisms that underlie adult pitch learning identified here are similar to those that are necessary to juvenile song copying. Specifically, we tested the importance of VTA_X neurons and D1 receptors in Area X to juvenile song copying. To test the role of VTA_X neurons in juvenile song copying, we used intersectional genetic methods to ablate these neurons during the second month after hatching, a period when juvenile zebra finches are actively modifying their own songs to match those of a tutor^{14,43} (Figure 4A). Juveniles were housed from 0–60 dph with an adult male tutor, providing them with abundant auditory experience of a suitable vocal model. Between 20–30 dph, we injected these juveniles ($n = 12$, 26 ± 1 dph) with AAV2/1.EF1 α .FLEX-Casp3-2A-TEV in the VTA and AAV2/9.CMV.HI.GFP-Cre.SV40 in Area X and recorded their songs at monthly intervals (Figure 4A; songs were recorded at 60, 90 and 120 dph). We also tracked the song development of another cohort of similarly housed juveniles that were siblings of the experimental animals and that were injected either with AAV2/1.EF1 α .FLEX-Casp3-2A-TEV in the VTA, AAV2/9.CMV.HI.GFP-Cre.SV40 in Area X, or no virus ($n = 3$ virally injected animals, 20 ± 1 dph at the time of injections; $n = 2$ animals that were not injected with any virus).

As adults, birds that had been injected with both viruses produced significantly worse copies of their tutor's song than did control animals (Figure 4B–D; see Figure S7 for sonograms of all juveniles). Post hoc histological analysis revealed a strong correlation between the number of surviving VTA_X neurons in adulthood and the similarity of the experimental bird's song to his tutor's song (Figure 4C, $R^2 = 0.689$, $p < 0.001$). Moreover, adults ($n = 6$) with the lowest number (<3500) of surviving VTA_X neurons produced poor copies of their tutors' songs at all time points (Figure 4E), although they sang similar amounts as control animals (Figure S8C, D), displayed normal levels of song stereotypy (Figure 4D, right), and progressed to their final songs in a manner similar to controls (Figure 4F; Figure S8A, B; Figure S9). Finally, when a subset of these adult males with reduced numbers of VTA_X cells were presented with a female, they sang a similar amount as did control males, suggesting that learning deficits did not simply reflect reduced motivation to sing (two-tailed t-test: number of motifs sung on first presentation of a female: VTA_X cell lesioned birds: 8 ± 2 motifs, $n = 3$ birds; control birds: 7 ± 1 motifs, $n = 3$ birds; $p = 0.8$). Therefore, intersectional ablation of VTA_X neurons disrupts juvenile song copying, underscoring that these neurons

provide a common cellular foundation for internally guided and externally reinforced forms of vocal learning.

We then used microdialysis methods to infuse a D1 receptor antagonist (SCH23390) into Area X in a cohort of juveniles for a ten-day period during the height of sensorimotor learning (~50–60 dph), when much of the tutor song is normally copied by juvenile zebra finches (Figure 4G; $n = 5$ juvenile male zebra finches, 45 ± 1 dph at time of infusion; all juveniles were raised in the presence of an adult tutor prior to this period). We recorded their songs continuously during this drug treatment and at the end of this period we flushed the probes with saline and then returned the birds to the colony until they reached early adulthood, at which time we recorded their songs once again ($n = 4$, 94 ± 4 dph; one animal was sacrificed at the end of the drug treatment period). We also recorded the songs of another cohort of juveniles that either received saline infusions in Area X ($n = 3$; 45 ± 2 dph at time of infusion) or were not manipulated ($n = 3$; 47 ± 0 dph at time of first recording). Compared to these control animals, the songs of juveniles infused with D1 antagonists in Area X showed little or no increase in similarity to the tutor song during the infusion period (Figure 4H). Despite the lack of copying during the juvenile treatment period, adult experimental and control birds ultimately displayed similar levels of tutor song copying, because juveniles treated with SCH23390 could subsequently compensate for their copying deficit during the post-treatment period (Figure 4I). In summary, similar to adult pitch learning, juvenile song copying depends on VTA_X neurons and the activation of D1 receptors in Area X.

Discussion

Here we show that the same VTA – BG circuits and dopamine signaling pathways are necessary to internally guided vocal copying in juvenile songbirds and externally reinforced forms of vocal learning in adults, highlighting a common and developmentally conserved mechanism for these different types of learning. Our findings extend the prior observation that 6-OHDA lesions of TH+ terminals in Area X can impair adult pitch learning²⁷ by localizing the likely source of these terminals to the VTA, highlighting the necessity of D1 receptors in Area X for this form of learning, and by defining a role for the VTA – BG/D1 pathway in juvenile song copying. Beyond establishing the necessity of this pathway to both forms of vocal learning, we found that pitch-contingent optogenetic stimulation of VTA_X terminals is sufficient to drive pitch learning, supporting a model in which trial by trial variations in VTA activity are the critical reinforcing signal for vocal learning⁴⁴.

The strong parallels between the avian and mammalian basal ganglia include a robust dopaminergic projection from the VTA/SNc, bolstering speculation that these inputs play an essential role in birdsong learning^{26,44,45}. Indeed, the recent finding that using 6-OHDA to lesion TH+ terminals in Area X interferes with pitch learning in adult birds provided critical experimental support for this idea. The current findings that intersectional genetic ablation of VTA_X neurons interferes with adult pitch learning extend this prior observation by localizing the cell bodies that provide the TH+ fibers in Area X to the VTA and thus help to inform an anatomically-grounded circuit model. Moreover, by showing that intersectional VTA_X lesions disrupt juvenile song copying, the current study supports the idea that a

common mechanism serves both internally-guided song learning in juveniles and externally reinforced vocal plasticity in adults. In both juveniles and adults, the residual learning capacity correlated with the number of surviving VTA neurons, raising the possibility that the relatively large endowment of VTA_X neurons (relative to the smaller number of VTA neurons that project to the medial striatum, e.g.²³) reflects strong selective pressures on vocal learning and its underlying neural circuitry. A broader implication is that evolutionarily ancient circuitry that first arose to enable reinforcement learning in response to aversive and appetitive cues was later co-opted to guide forms of motor learning that are internally guided and do not depend on external reinforcement.

Beyond merely extending the earlier study using 6-OHDA lesions in Area X of adult birds, the microdialysis methods used here show that D1-receptor signaling in Area X is a critical effector of both juvenile song copying and adult pitch learning. The current observations further underscore that a common circuit and signaling pathway mediates these two different forms of learning. Moreover, the strong and reversible effects of D1- receptor blockade on song copying in juveniles with substantial prior tutor song exposure help assign these deleterious effects to disruptions of sensorimotor learning rather than the preceding epoch of sensory learning. This distinction is less readily made with intersectional genetic ablation of VTA_X neurons, the onset and time course of which is variable, relatively slow, and only subject to post hoc monitoring. A further novel insight provided by microdialysis applied in adult birds is that D1- and D2- receptor blockers exert distinct effects on singing: Whereas D1- receptor blockade in Area X completely and reversibly abolished adult pitch learning without significantly affecting the amount of singing, D2-receptor blockade strongly and consistently suppressed singing while exerting variable effects on pitch learning. These differential effects raise the possibility that song production is differentially regulated through different dopamine receptor subtypes in Area X, reminiscent of the functionally distinct effects of D1- and D2-signaling in the mammalian striatum on locomotion⁴¹.

We found that pitch-contingent optogenetic stimulation of VTA terminals in Area X over the course of hours could induce frequency shifts in target syllable without affecting other syllables in the motif, highly similar to noise-driven pitch learning^{17,19}. One important difference is that optogenetic stimulation of VTA_X terminals positively reinforced the pitch of the target syllable, opposite in sign of pitch changes driven by noise. In fact, because singing-triggered noise can depress VTA_X neuron activity²⁶, whereas “escapes” from noise transiently elevates activity in these neurons, the current study provides a causal link between auditory feedback-dependent differences in VTA_X activity and long-lasting changes to vocal performance. Given that the caudal auditory forebrain contains neurons that respond selectively to singing-triggered noise and that project to the VTA⁴⁶, these findings further advance an error detection circuit that harnesses singing-related auditory feedback information to modulate VTA_X neuron activity on a trial-by-trial basis to affect vocal motor learning. One major goal is to dissect the underlying circuitry that converts singing-triggered excitation in the auditory forebrain into transient suppression of VTA_X neuron activity. Another critical step will be to determine whether and how auditory-related afferents to VTA_X neurons function during juvenile sensorimotor learning, when singing-related auditory feedback is compared to the memory of the tutor song, and motor learning proceeds in the absence of aversive external auditory cues. A distinct possibility is that VTA_X neurons

supply the instructive signals used to guide this form of imitative learning, although the approaches used here cannot distinguish between permissive and instructive roles for these neurons in juvenile song copying.

One of the many remarkable parallels between birdsong and human speech is that both are acoustical signals where fine temporal modulations on the timescale of 10–100 milliseconds are salient to their communicative functions. In contrast to slower forms of motor learning involving lever-pressing or licking^{3–5,7}, juvenile copying and adult pitch learning require similarly precise modification of vocal structure. Our findings advance dopamine-dependent synaptic plasticity in the BG as a likely cellular effector of vocal learning⁴⁴, while raising the question of how relatively slow signaling through G-protein coupled receptors underlies a form of motor learning that exhibits millisecond precision⁴⁷. One attractive model involves short-lived synaptic tags on medium spiny neurons (MSNs) in Area X, which are hypothesized to be formed by temporally coincident glutamatergic activity from song premotor regions⁴⁴. When these patterns of premotor activity produce vocalization-related feedback that stimulates dopamine release from VTA_X neurons, subsequent dopamine receptor activation on these same MSNs stabilize these synaptic tags and strengthens the relevant premotor synapses, resulting in an adaptive bias that the BG supplies to the song motor system⁴⁴. More broadly, by providing evidence of a role for the VTA in avian motor learning, the current study suggests that an evolutionarily conserved circuit mechanism supports different forms of learning across both birds and mammals⁴⁸, raising the possibility that this mechanism is also central to speech and musical learning in humans, as well as playing a critical role in birdsong learning.

Materials and Methods

Juvenile (18–49 dph) and adult (82–436 dph) male zebra finches were obtained from the Mooney lab breeding colony within the Duke University Medical Center animal facility. Experimental procedures were conducted in accordance with the National Institutes of Health guidelines and were reviewed and approved by the Duke University Medical Center Animal Care and Use Committee. Viral vectors were acquired from University of Pennsylvania Vector Core and University of North Carolina, Chapel Hill Vector Core.

Genetic ablation of VTA_X neurons

Male zebra finches (20–25 days post hatch (dph) for juvenile experiments, 100–110 dph for adult experiments) were food deprived for 30 minutes and then anesthetized with 2% isoflurane gas before being placed on top of a small heating pad in a custom stereotaxic apparatus. Rate of breathing and stability of surgical plane were monitored throughout surgery. The feathers over the skull were trimmed and topical anesthetic (0.25% bupivacaine) was applied before an incision was made in the skin from anterior to posterior with a scalpel. After pushing skin from the center of the skull with a cotton swab doused in 70% ethanol, craniotomies were made with a smaller scalpel at a predetermined distance from the bifurcation of the midsagittal sinus (the ‘y-sinus’; coordinates measured from y-sinus: VTA: head angle 37 degrees, 1.65 mm anterior, 0.5 and 1.8 mm lateral, 6.2 mm ventral; Area X: head angle 43 degrees, 5.3 mm anterior, 1.6 mm lateral, 3.2, 2.9 and 2.7

mm ventral). To selectively ablate VTA_X cells, a pressure injection system (Drummond Nanoject II) was used to make bilateral injections of a retrogradely transported Cre construct (AAV2/9.CMV.HI.GFP-Cre.SV40; Penn Vector, a total of 15 injections of 32.2 nl of Cre per hemisphere) into Area X at 3 different depths. A locally expressed Cre-dependent caspase construct was then injected into the VTA at 2 different locations along the medial-lateral axis (AAV2/1.Ef1 α .FLEXCasp3-2A-TEV; construct courtesy of Nirao Shah, UCSF, 15 injections of 32.2 nl of Casp3 per site per hemisphere i.e. a total of 4 caspase injection sites per bird). After these viral injections, the craniotomies were sealed with bone wax, the incision site was closed with tissue adhesive, and the bird was allowed to recover from anesthesia under a heat lamp. At the endpoint of each experiment and 5 days prior to perfusion, birds were injected with AlexaFluor 594 in Area X to retrogradely label VTA_X neurons. Five days after these tracer injections, birds were deeply anesthetized with an intraperitoneal injection of pentobarbital solution (Euthasol) and then perfused through the heart with 0.025 M phosphate-buffered saline followed by 4% paraformaldehyde. The brain was then removed from the skull and placed in a cryoprotective formalin sucrose solution (30% sucrose in 4% paraformaldehyde) overnight. The next day consecutive sagittal sections of the cryoprotected brain were cut on a freezing microtome and alternate sections were mounted on glass slides. A subset of alternate sections were treated with an antibody against tyrosine hydroxylase (α TH, 1:1000, Abcam) overnight at 4°C and reacted with secondary antibody (1:500, Abcam) at RT for 1 hour then mounted on slides to visualize TH+ cells in the VTA. A similar process was used to visualize TH+ fibers in Area X. Sections containing VTA (or VTA terminals in Area X) were visualized and imaged under a confocal microscope (Zeiss Axioskop 2). The images were then examined in an image-processing program and the number of fluorescent retrogradely labeled cells in VTA was counted in a semi-automated manner (ImageJ, CellCounter plug-in).

Genetic ablation experiments

Pitch contingent learning—Young adult male birds (100–110 dph) were screened for syllables with clear tonal components and for the amount of song produced. Birds that fit both these criteria were then bilaterally injected with viruses to ablate VTA_X cells. After the birds recovered from surgery and began singing readily, their songs were recorded and a template to detect the fundamental frequency (i.e., pitch) of a tonal syllable was made in a custom software program (EvTAF, Tumer and Brainard, 2007). The template was designed to detect no less than 75% of the renditions of the targeted syllable with no more than a 5 millisecond jitter in detection onset. After collecting two days of “baseline” song, a threshold at the upper 70th percentile of the target syllable’s pitch distribution was set and a 50 millisecond white noise burst (~70 dB) was played through a nearby speaker to the bird whenever the program template detected that the pitch of the targeted syllable was below this threshold; over hours and days, this manipulation results in an adaptive shift in the pitch of the target syllable. The bird’s pitch for the targeted syllable was measured in the late morning and early evening for the next four days and the threshold was adjusted each morning and early evening to the upper 70th percentile of the bird’s pitch distribution to promote more rapid learning. After four days of pitch-contingent white noise experience, the white noise was discontinued and the bird’s song was recorded for the next three to four days as the pitch of the targeted syllable recovered toward its baseline value. This entire

process was repeated one month after viral injections, when VTA_X neurons had been ablated by the intersectional viral treatment. At the end of the recovery period from this second pitch learning experiment, birds were injected with dextran in Area X in order to allow for the number of VTA_X cells remaining to be quantified, as previously described.

Juvenile song copying—Juvenile male zebra finches (20–25 dph) were injected with viruses as previously described to ablate VTA_X cells, then isolated with their siblings and father until 60 dph. At 60 dph juveniles were housed with other virally injected birds and isolated temporarily for recording at 60, 90, and 120 dph using a custom song recording program (SOUND ANALYSIS PRO 2011 (SAP)). We relied on percent similarity, a measure that combines measures of pitch, amplitude modulation, frequency modulation, Weiner entropy and goodness of pitch, to gauge the similarity of song elements between two sets of songs (i.e., the pupil's song and that of his tutor). We chose representative motifs (>30 ~200–500 ms long motifs) from pupils and used the asymmetric time-courses setting to compare the pupil motif to a representative tutor motif that we confirmed was highly similar to the tutor's other motifs. SAP was also used to measure spectral features of single syllables, such as entropy and entropy variance, over development in a subset of birds. Once birds reached 120 dph, they were injected with dextran in Area X and the number of VTA_X cells was quantified as previously described.

Singing-triggered optogenetic stimulation of VTA_X terminals

Using surgical methods previously described, young adult male birds (60–90 dph) were bilaterally injected in the VTA with a virus containing a channelrhodopsin construct (2/9.AAV-CAG-ChR2-mCherry or 2/9.AAV-CAG-ChR2-YFP-neurexin) at 4 different sites (50 injections of 9.2 nl of ChR2 per site, 2 sites per hemisphere). After waiting 3 to 6 months to allow for optimal viral expression, birds were anesthetized and placed in a stereotaxic apparatus and craniotomies were made over Area X bilaterally. Six of the eight birds used for these experiments were tested for terminal field optogenetic responses in Area X with a 500 kOhm tungsten electrode (MicroProbes Inc.) coupled to a fiberoptic cable (ThorLabs, 200um diameter core) through which 50–100 millisecond pulses of light were delivered and neural activity was recorded simultaneously (Differential A–C Amplifier 1700, A–M Systems). All birds were then implanted bilaterally over Area X with fiberoptic ferrules at an anterior angle to avoid passing through LMAN (coordinates: 43 degree head angle, mark 5.3 mm rostral; adjust head angle to 72 degrees and move 1.2 mm rostral from previous “mark”, 1.6 mm lateral, 2.7–3.0 mm ventral). Craniotomies were then sealed with melted bone wax and ferrules were secured in place with MetaBond and then covered with a layer of VetBond. After birds recovered from anesthesia under a heat lamp, fiberoptic cables (ThorLabs, 200 um core, 0.37 NA) were connected to the newly implanted ferrules by ferrule sleeves. The other ends of the fiberoptic cables were attached to a two-channel optical commutator (FRJ_1x2i_FC-2FC, Doric), allowing the bird to move about its cage freely. The commutator was then connected by a patch cable (ThorLabs) to a DPSS laser (BL473T3-100, Shanghai Lasers).

As described above for adult pitch learning experiments, we created a template that detected no less than 75% of the renditions of the targeted syllable with no more than a 5 millisecond

jitter in detection onset. After collecting two days of “baseline” song (i.e., produced when the bird was connected to the fiberoptic cables but the laser remained off), a threshold at the upper (or lower) 70th percentile of the target syllable’s pitch distribution was set and a 50 millisecond pulse of blue light (473 nm, 5–8 mW emitted at each ferrule) was delivered to Area X whenever the program detected that the pitch of the targeted syllable was below (or above) this threshold. The bird’s pitch for the targeted syllable was measured in the late morning and early evening for the next four days (for 8 out of 10 syllables, see below) and the threshold was adjusted to the upper 70th percentile of their pitch accordingly. Out of the ten syllables targeted, eight were exposed to pitch-contingent optogenetic stimulation for four days, one for six days, and another for ten days. Light stimulation was then ended and song in the absence of stimulation was recorded for up to four days. Birds were then uncoupled from the fiberoptic cables and returned to the colony. 3–5 months after stimulation ended, birds were again recorded for 4–5 days before being perfused. Histology was performed as described above, with alternate sections stained against mCherry or GFP (Abcam) for visualization of the terminal field in Area X. Only birds that had accurate placement of ferrules in the center of Area X and robust labeling of cell bodies in VTA and of axon terminals in Area X were included in our analysis. Exclusion of birds was blind to behavioral results. One bird was excluded from our analysis as both ferrule implantation and viral injection localization were incorrect (i.e. placement over LMAN rather than Area X, viral injection caudal and dorsal to VTA). In sample sections, we counted the numbers of ChR2-YFP+ cells. Comparing the average number of ChR2-YFP+ cells per section to the average number of retrogradely labeled cells from other sections (from other tissue), we can provide a rough estimate that ~ 40% of VTA_X neurons are ChR2+.

Microdialysis experiments

Adult pitch learning experiments—Young adult birds (> 80 dph) with clear tonal elements in their song were chosen for implantation of microdialysis probes. Probes were constructed in house from plastic tubing which served as a drug reservoir fitted at the end with a 0.7–1.0 mm-long semipermeable membrane which allowed drug to slowly diffuse throughout the day (see Hamaguchi and Mooney, 2012 for probe design). Using surgical procedures and stereotaxic coordinates described above, craniotomies were made over Area X and neural recordings were made to confirm its depth (Differential A-C Amplifier 1700, A-M Systems). We approached Area X rostrally as to avoid LMAN (anterior Area X coordinates: initial head angle 43 degrees, 5.3 mm anterior marked with scalpel on skull, then adjusted head angle to 72 degrees, 1.2 mm anterior from scalpel mark, 1.7 mm lateral, 2.9–3.2 mm ventral). Probes were then implanted with the tip of the semipermeable membrane placed at the most ventral part of Area X so that the membrane extended through the dorsal-ventral extent of Area X. The surgical site was covered with melted bone wax, and the probes were secured in place first using MetaBond and then a coating of VetBond. Birds were then removed from the apparatus and recovered under a heat lamp. After recovery birds were placed in a sound isolation box and their first full day of song was recorded and used to make an EvTAF template to target a tonal syllable as described above. Birds were recorded in the absence of white noise for two hours the morning after their first full day of song then infused with saline and recorded in the presence of pitch-contingent

white noise for the next 8 hours, after which they were again infused with saline and white noise was turned off (“learning day”, day 1 on figure 2A).

This protocol was repeated the following day with the white noise remaining off (“recovery day,” day 2) The next day the protocol from the saline “learning day” (day 3) was repeated but either 5 mM SCH23390 or 0.5 µg/ml sulpiride were infused after two hours of recording instead of saline and washed out with saline after 8 hours of recording. After this drug “learning day,” the bird underwent a “recovery day” (day 4) after which followed another “learning day” (day 5) with saline only. Before perfusion, birds were infused with fluorescent muscimol-BODIPY for 2 hours to allow for post-hoc visualization of drug diffusion through the semi-permeable membrane into Area X. Birds were then perfused and histology was performed as described above to assess correct placement of the dialysis probes in Area X and the extent of drug diffusion. This manner of quantifying the spread of drug from our microdialysis probes may underestimate the amount of drug spread as muscimol-BODIPY is of a higher molecular weight than SCH23390 or sulpiride and was infused for a shorter period of time than the SCH23390 or sulpiride.

Juvenile song copying experiments—Young (40–49 dph) juvenile male zebra finches that had recognizable syllables but had not yet developed a stereotyped motif were implanted bilaterally with microdialysis probes in Area X in the manner previously described. After they recovered from surgery they were infused with saline until they began singing again. After recording at least one day of singing with saline infusion, the birds were infused in the morning ~10–20 minutes before ‘lights on’ with either SCH23390 (experimental birds) or saline (3 of 6 of the control birds were implanted and infused with saline only, the 3 other birds were not implanted and were recorded continuously for 12 days) each morning for the next 10 days. After 10 days of drug or saline treatment, juveniles were infused with saline and recorded for 2–3 more days. They were then placed back in the colony until they reached early adulthood (90–110 dph), when they were again isolated and recorded for 1–2 days. Following perfusion, as described above, the fixed tissue was examined for correct placement of microdialysis probes in Area X. Because the probes clogged 2–3 weeks after implantation, we were not able to infuse tracers to estimate drug diffusion in these birds.

Analysis of song data

Coefficient of variation—The pitch of a small component of the syllable (5–10 ms) was measured for no fewer than 50 “catch” syllables. The standard deviation of these small syllable components was then divided by the mean of the small syllable components.

Percent change in pitch—For adult VTA_X ablation experiments, channelrhodopsin experiments, and adult microdialysis experiments the percent change in pitch of the targeted syllable after learning was calculated. The pitch of the entire tonal component of the syllable was measured for no fewer than 50 “catch” syllables on the day before the manipulation began and then again on the last day of the manipulation. The experimental pitch was subtracted from the baseline pitch, divided by the baseline pitch, and then multiplied by 100 to calculate the percent change in pitch.

Change in auROC—To better gauge the change in birds' pitch distributions after manipulation, the change in the area under the receiver operator characteristic (auROC) of the pitch of the targeted syllable after learning was measured for adult VTA_X ablation experiments, channelrhodopsin experiments, and adult microdialysis experiments. The pitch of the entire tonal component of the syllable was measured using no fewer than 50 “catch” syllables on the day before the manipulation began and then again on the last day of the manipulation. The auROC was calculated by taking the integral between the proportion of baseline pitches correctly considered baseline pitches and the proportion of experimental pitches incorrectly considered baseline pitches. The data were then bootstrapped to ensure an unbiased measurement.

Similarity scores—For juvenile VTA_X and microdialysis experiments, SAP2011 was used to calculate percent similarity either to the pupil's tutor's song (tutor similarity) or to the pupil's own song (self-similarity). A representative motif from the tutor (for tutor similarity) or the pupil (for self-similarity) was selected then compared to no fewer than 30 of the pupil's motifs in SAP2011 using asymmetrical time-courses under the ‘Similarity’ tab.

Statistics

Data are presented as mean \pm S.E.M. unless otherwise noted. Error bars in all figures indicate the standard error of the mean. All groups with greater than 8 samples were tested for and passed the Kolmogorov-Smirnov test for normality. Groups with fewer than 8 samples were not large enough to detect normality but parametric tests were still used in order to detect differences in small samples. P values were calculated from two-tailed t-tests between only two groups and listed in the figure legends. For groups of more than two, ANOVAs were performed first to check for interaction term significance before t-tests were performed. P values of 0.05 or below were considered significant. No statistical methods were used to pre-determine sample sizes, but our sample sizes are similar to those reported in previous publications (Tumer and Brainard, 2007; Andalman and Fee, 2009). The experimenters were not blinded to allocation of subjects and allocation of subjects was not randomized. Automatic detection and calculation of syllable frequencies allowed the experimenter to be blind to conditions before and after viral expression. All data were analyzed with Matlab software.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank F. Wang, S. Soderling, H. Greenside, H. Yin, T. Warren and K. Tschida for reading versions of this manuscript, and T. Warren for his help with implementing pitch learning techniques. We also thank M. Booze for his management of the bird colony and members of the Mooney lab for helpful discussion. This research was supported by NIH R01DC002524 and NIH R01NS099288 to R.M. and NIH F30NS096871 to M.G.K.

References

1. Pinker, S. *The Language Instinct*. William Morrow and Company; 1994.

2. Locke, JL. *The Child's Path to Spoken Language*. Harvard University Press; 1995.
3. Tsai HC, et al. Phasic firing in dopaminergic neurons is sufficient for behavioral conditioning. *Science*. 2009; 324:1080–1084. DOI: 10.1126/science.1168878 [PubMed: 19389999]
4. Adamantidis AR, et al. Optogenetic interrogation of dopaminergic modulation of the multiple phases of reward-seeking behavior. *J Neurosci*. 2011; 31:10829–10835. DOI: 10.1523/JNEUROSCI.2246-11.2011 [PubMed: 21795535]
5. Zweifel LS, et al. Disruption of NMDAR-dependent burst firing by dopamine neurons provides selective assessment of phasic dopamine-dependent behavior. *Proc Natl Acad Sci U S A*. 2009; 106:7281–7288. DOI: 10.1073/pnas.0813415106 [PubMed: 19342487]
6. Witten IB, et al. Recombinase-driver rat lines: tools, techniques, and optogenetic application to dopamine-mediated reinforcement. *Neuron*. 2011; 72:721–733. DOI: 10.1016/j.neuron.2011.10.028 [PubMed: 22153370]
7. Steinberg EE, et al. A causal link between prediction errors, dopamine neurons and learning. *Nat Neurosci*. 2013; 16:966–973. DOI: 10.1038/nn.3413 [PubMed: 23708143]
8. Surmeier DJ, Plotkin J, Shen W. Dopamine and synaptic plasticity in dorsal striatal circuits controlling action selection. *Curr Opin Neurobiol*. 2009; 19:621–628. DOI: 10.1016/j.conb.2009.10.003 [PubMed: 19896832]
9. Gerfen CR, Surmeier DJ. Modulation of striatal projection systems by dopamine. *Annu Rev Neurosci*. 2011; 34:441–466. DOI: 10.1146/annurev-neuro-061010-113641 [PubMed: 21469956]
10. Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. *Science*. 1997; 275:1593–1599. [PubMed: 9054347]
11. Prather J, Okanoya K, Bolhuis JJ. Brains for birds and babies: Neural parallels between birdsong and speech acquisition. *Neurosci Biobehav Rev*. 2017
12. Brainard MS, Doupe AJ. Translating birdsong: songbirds as a model for basic and applied medical research. *Annu Rev Neurosci*. 2013; 36:489–517. DOI: 10.1146/annurev-neuro-060909-152826 [PubMed: 23750515]
13. Eales LA. Song learning in zebra finches: some effects of song model availability on what is learnt and when. *Animal Behav*. 1985; 33:1293–1300.
14. Immelman, K. *Bird Vocalisations*. Cambridge University Press; 1969.
15. Price PH. Developmental determinants of structure in zebra finch song. *J Comp Physiol Psychol*. 1979; 93:260–277.
16. Marler P, Peters S. Sparrows learn adult song and more from memory. *Science*. 1981; 213:780–782. DOI: 10.1126/science.213.4509.780 [PubMed: 17834587]
17. Tumer EC, Brainard MS. Performance variability enables adaptive plasticity of 'crystallized' adult birdsong. *Nature*. 2007; 450:1240–1244. DOI: 10.1038/nature06390 [PubMed: 18097411]
18. Ali F, et al. The basal ganglia is necessary for learning spectral, but not temporal, features of birdsong. *Neuron*. 2013; 80:494–506. DOI: 10.1016/j.neuron.2013.07.049 [PubMed: 24075977]
19. Andalman AS, Fee MS. A basal ganglia-forebrain circuit in the songbird biases motor output to avoid vocal errors. *Proc Natl Acad Sci U S A*. 2009; 106:12518–12523. DOI: 10.1073/pnas.0903214106 [PubMed: 19597157]
20. Nottebohm F, Arnold AP. Sexual dimorphism in vocal control areas of the songbird brain. *Science*. 1976; 194:211–213. [PubMed: 959852]
21. Nottebohm F, Kelley DB, Paton JA. Connections of vocal control nuclei in the canary telencephalon. *J Comp Neurol*. 1982; 207:344–357. DOI: 10.1002/cne.902070406 [PubMed: 7119147]
22. Sohrabji F, Nordeen EJ, Nordeen KW. Selective impairment of song learning following lesions of a forebrain nucleus in the juvenile zebra finch. *Behav Neural Biol*. 1990; 53:51–63. [PubMed: 2302141]
23. Person AL, Gale SD, Farries MA, Perkel DJ. Organization of the songbird basal ganglia, including area X. *J Comp Neurol*. 2008; 508:840–866. DOI: 10.1002/cne.21699 [PubMed: 18398825]
24. Bottjer SW, Brady JD, Cribbs B. Connections of a motor cortical region in zebra finches: relation to pathways for vocal learning. *J Comp Neurol*. 2000; 420:244–260. [PubMed: 10753310]

25. Gale SD, Person AL, Perkel DJ. A novel basal ganglia pathway forms a loop linking a vocal learning circuit with its dopaminergic input. *J Comp Neurol.* 2008; 508:824–839. DOI: 10.1002/cne.21700 [PubMed: 18398824]
26. Gadagkar V, et al. Dopamine neurons encode performance error in singing birds. *Science.* 2016; 354:1278–1282. DOI: 10.1126/science.aah6837 [PubMed: 27940871]
27. Hoffmann LA, Saravanan V, Wood AN, He L, Sober SJ. Dopaminergic Contributions to Vocal Learning. *J Neurosci.* 2016; 36:2176–2189. DOI: 10.1523/JNEUROSCI.3883-15.2016 [PubMed: 26888928]
28. Yang CF, et al. Sexually dimorphic neurons in the ventromedial hypothalamus govern mating in both sexes and aggression in males. *Cell.* 2013; 153:896–909. DOI: 10.1016/j.cell.2013.04.017 [PubMed: 23663785]
29. Tye KM, Deisseroth K. Optogenetic investigation of neural circuits underlying brain disease in animal models. *Nat Rev Neurosci.* 2012; 13:251–266. DOI: 10.1038/nrn3171 [PubMed: 22430017]
30. Deisseroth K. Optogenetics: 10 years of microbial opsins in neuroscience. *Nat Neurosci.* 2015; 18:1213–1225. DOI: 10.1038/nn.4091 [PubMed: 26308982]
31. Boyden ES. Optogenetics and the future of neuroscience. *Nat Neurosci.* 2015; 18:1200–1201. DOI: 10.1038/nn.4094 [PubMed: 26308980]
32. Tritsch NX, Ding JB, Sabatini BL. Dopaminergic neurons inhibit striatal output through non-canonical release of GABA. *Nature.* 2012; 490:262–266. DOI: 10.1038/nature11466 [PubMed: 23034651]
33. Lewis JW, Ryan SM, Arnold AP, Butcher LL. Evidence for a catecholaminergic projection to area X in the zebra finch. *J Comp Neurol.* 1981; 196:347–354. DOI: 10.1002/cne.901960212 [PubMed: 7217361]
34. Soha JA, Shimizu T, Doupe AJ. Development of the catecholaminergic innervation of the song system of the male zebra finch. *J Neurobiol.* 1996; 29:473–489. DOI: 10.1002/(SICI)1097-4695(199604)29:4<473::AID-NEU5>3.0.CO;2-5 [PubMed: 8656212]
35. Hamaguchi K, Mooney R. Recurrent interactions between the input and output of a songbird cortico-basal ganglia pathway are implicated in vocal sequence variability. *J Neurosci.* 2012; 32:11671–11687. DOI: 10.1523/JNEUROSCI.1666-12.2012 [PubMed: 22915110]
36. Ding L, Perkel DJ, Farries MA. Presynaptic depression of glutamatergic synaptic transmission by D1-like dopamine receptor activation in the avian basal ganglia. *J Neurosci.* 2003; 23:6086–6095. [PubMed: 12853427]
37. Leblois A, Wendel BJ, Perkel DJ. Striatal dopamine modulates basal ganglia output and regulates social context-dependent behavioral variability through D1 receptors. *J Neurosci.* 2010; 30:5730–5743. DOI: 10.1523/JNEUROSCI.5974-09.2010 [PubMed: 20410125]
38. Ding L, Perkel DJ. Dopamine modulates excitability of spiny neurons in the avian basal ganglia. *J Neurosci.* 2002; 22:5210–5218. [PubMed: 12077216]
39. Kubikova L, Wada K, Jarvis ED. Dopamine receptors in a songbird brain. *J Comp Neurol.* 2010; 518:741–769. DOI: 10.1002/cne.22255 [PubMed: 20058221]
40. Kravitz AV, et al. Regulation of parkinsonian motor behaviours by optogenetic control of basal ganglia circuitry. *Nature.* 2010; 466:622–626. DOI: 10.1038/nature09159 [PubMed: 20613723]
41. Kravitz AV, Tye LD, Kreitzer AC. Distinct roles for direct and indirect pathway striatal neurons in reinforcement. *Nat Neurosci.* 2012; 15:816–818. DOI: 10.1038/nn.3100 [PubMed: 22544310]
42. Albin RL, Young AB, Penney JB. The functional anatomy of basal ganglia disorders. *Trends Neurosci.* 1989; 12:366–375. [PubMed: 2479133]
43. Deregnaucourt S, Mitra PP, Feher O, Pytte C, Tchernichovski O. How sleep affects the developmental learning of bird song. *Nature.* 2005; 433:710–716. DOI: 10.1038/nature03275 [PubMed: 15716944]
44. Fee MS, Goldberg JH. A hypothesis for basal ganglia-dependent reinforcement learning in the songbird. *Neuroscience.* 2011; 198:152–170. DOI: 10.1016/j.neuroscience.2011.09.069 [PubMed: 22015923]
45. Doya K, Sejnowski TJ. A Novel Reinforcement Model of Birdsong Vocalization Learning. *Advances in neural information processing systems.* 1995:101–108.

46. Mandelblat-Cerf Y, Las L, Denisenko N, Fee MS. A role for descending auditory cortical projections in songbird vocal learning. *Elife*. 2014; 3
47. Charlesworth JD, Tumer EC, Warren TL, Brainard MS. Learning the microstructure of successful behavior. *Nat Neurosci*. 2011; 14:373–380. DOI: 10.1038/nn.2748 [PubMed: 21278732]
48. Reiner A, Medina L, Veenman CL. Structural and functional evolution of the basal ganglia in vertebrates. *Brain Res Brain Res Rev*. 1998; 28:235–285. [PubMed: 9858740]
49. Tumer EC, Brainard MS. Performance variability enables adaptive plasticity of 'crystallized' adult birdsong. *Nature*. 2007; 450:1240–1244. [PubMed: 18097411]
50. Hamaguchi K, Mooney R. Recurrent interactions between the input and output of a songbird cortico-basal ganglia pathway are implicated in vocal sequence variability. *Journal of Neuroscience*. 2012; 32(34):11671–11687. [PubMed: 22915110]
51. Andalman AS, Fee MS. A basal ganglia-forebrain circuit in the songbird biases motor output to avoid vocal errors. *Proc Natl Acad Sci U S A*. 2009; 28;106(30):12518–23.

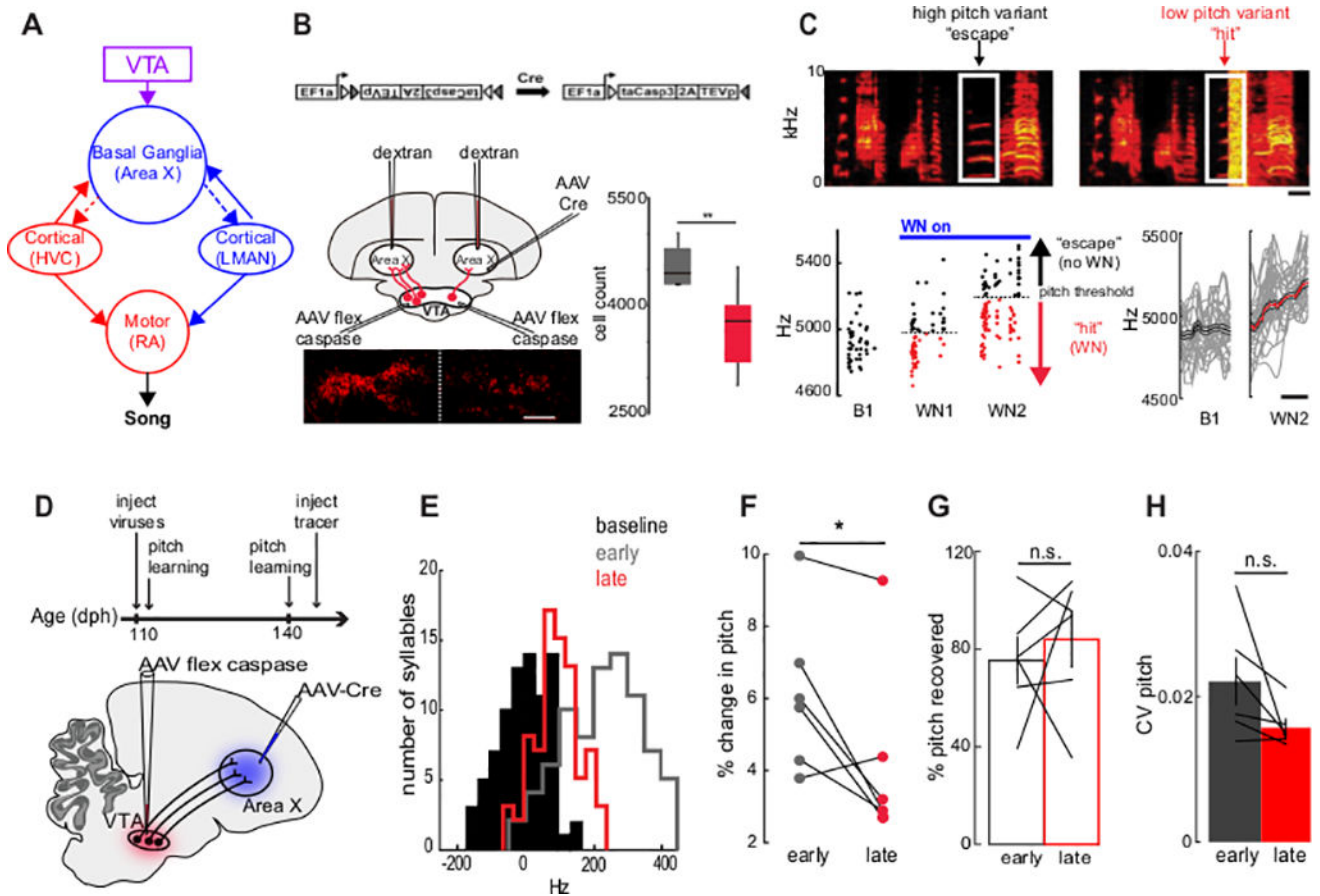


Figure 1. Genetically ablating VTA_X neurons in adult birds reduces pitch learning
 (A) Schematic of neural circuit for song, highlighting the VTA and BG (Area X). (B) *Top*, Cre-dependent caspase construct. *Bottom left*, Unilateral VTA_X ablation with VTA histology (an exemplar brain section from one of 2 birds subjected to unilateral ablation is shown). *Bottom right*, VTA_X cell counts for control (grey) and bilateral VTA_X ablation (red) birds. Unpaired two-tailed t-test: control: 4423 ± 143 cells, n = 10 birds; experimental: 3669 ± 156 cells, n = 10 birds; P=0.003. $t(18) = -3.38$. Horizontal lines of box plots represent the first quartile, median, and third quartile; whiskers of box plot represent the minimum and maximum. Scale bar, 500 μm. All values mean ± S.E.M throughout (C) *Top*, Example sonograms during pitch learning. White boxes indicate the targeted syllable. *Bottom left*, pitch of targeted syllable before (baseline, B1) pitch learning, during the first day of pitch learning (White Noise day 1 (WN1)) and during the second day of pitch learning (White Noise day 2; WN2)). Black dots, “escapes”; red dots, “hits.” *Bottom right*, frequency contours and mean of target syllable before (B1) (n = 50 syllables) and two days after (WN2) WN (n = 50 syllables). Scale bars, 20 milliseconds. (D) Experimental design for adult pitch learning with VTA_X ablation. (E) Pitch distribution of a target syllable before WN (black), after WN early (grey), and after WN late in the viral expression window (red) normalized to the pitch at baseline. (F) Percent change in pitch of target syllables. Paired two tailed t-test: early: 6.11 ± 0.84%; late: 4.18 ± 0.96%, n = 6 syllables from 6 birds, P = 0.042. $t(5) = 2.710$. (G) Percent of pitch recovered three days after discontinuing WN.

Paired two tailed t-test: early: $75.30 \pm 8.67\%$; late: $83.89 \pm 10.24\%$; $n = 6$, $p = 0.577$; $t(5) = -0.596$. (H) Coefficient of variation of target syllable pitch. Paired two tailed t-test: early: 0.022 ± 0.003 ; late: 0.016 ± 0.004 ; $n = 6$, $p = 0.099$; $t(5) = 2.02$. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. All values shown as mean \pm standard error of the mean (S.E.M.).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

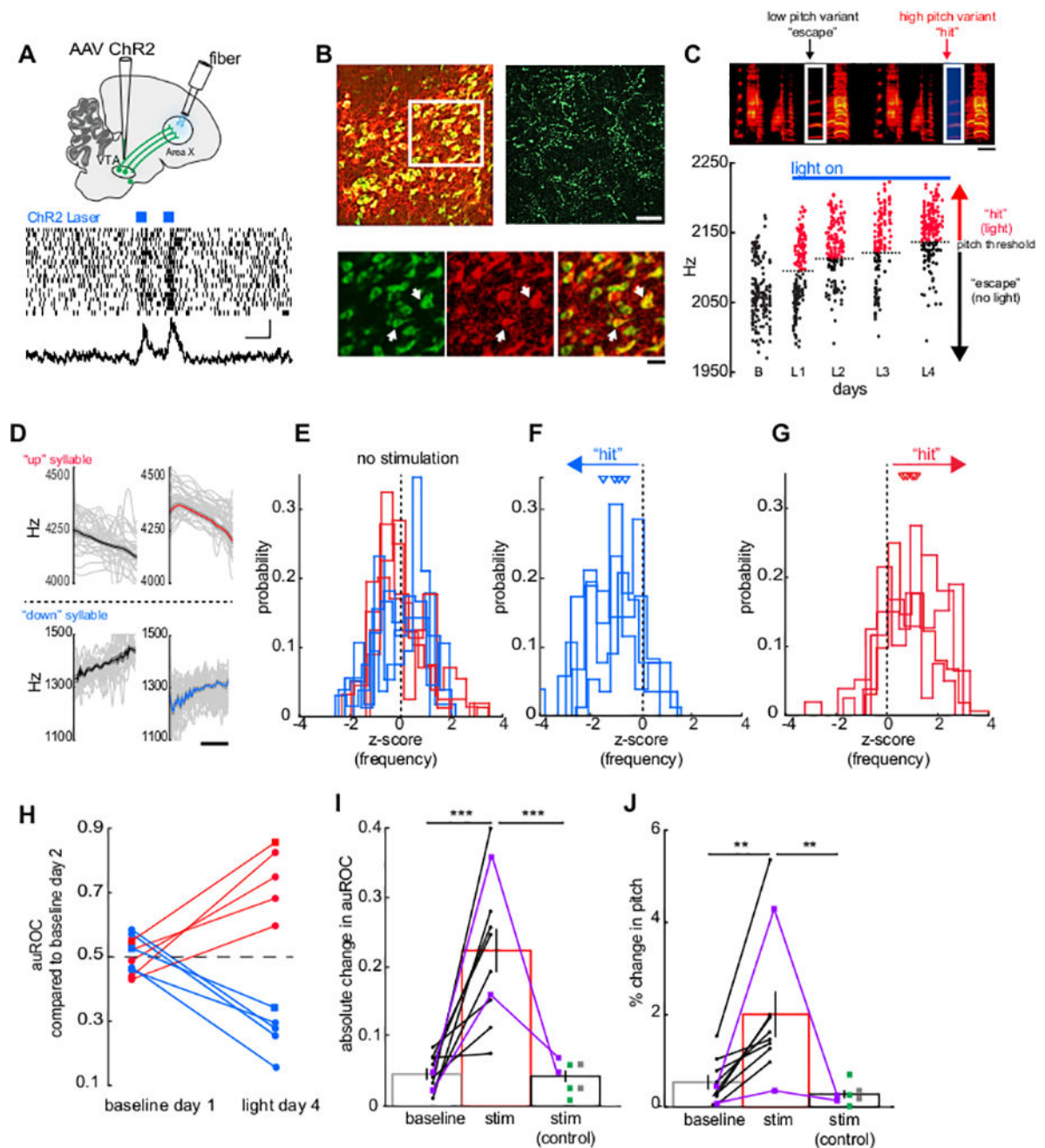


Figure 2. Pitch-contingent stimulation of VTA_X terminals is sufficient to drive pitch learning (A) *Top*, experimental design. *Bottom*, activity in Area X during optogenetic stimulation of VTA terminals. Scale bars, 500 milliseconds, 50 spikes/sec, n = 20 trial. (B) *Top left*, merged image of TH+ VTA cells (red) with ChR2 (green) expression (shown is an exemplar brain section from one of 3 birds in which co-staining was performed). *Top right*, ChR2 terminals in Area X (an exemplar brain section from one of 3 birds in which immunostaining was performed is shown). Scale bar, 100 μ m. *Bottom row*, inset from *top left*. Scale bar, 50 μ m. White arrows indicate co-labeled cells. (C) Pitch-contingent optogenetic stimulation of VTA_X terminals. B, baseline day; L1, first day with light

stimulation; L2, second day with light stimulation. Scale bar, 50 milliseconds. (D) Frequency contours and mean of target syllables before (left) and after (right) stimulation ($n = 50$ syllables). “Up” (red) and “down” (blue) syllables receive stimulation when bird sings above or below threshold, respectively. (E) Z-scored frequency of all syllables before stimulation ($n = 8$ syllables, 6 birds). (F) Z-scored frequency of “down” syllables ($n = 4$ syllables from 4 birds) after stimulation. Triangles indicate mean z-score for each syllable. (G) Z-scored frequency of “up” ($n = 4$ syllables from 4 birds) syllables after stimulation. Triangles indicate mean z-score for each syllable. (H) auROC for “up” and “down” syllables ($n = 10$ syllables, 8 birds). Squares correspond to birds that were tested with both 70% (experimental) and 100% (control) contingency stimulation. (I) Mean change in auROC of target syllable frequency between the last day of baseline and: the first day of baseline (0.046 ± 0.009); the last day of pitch-contingent optogenetic stimulation of VTA_X terminals in experimental birds (0.211 ± 0.039); and the last day of pitch-contingent optical stimulation in Area X of the various control birds (0.043 ± 0.008). Paired two-tailed t-test: absolute change in auROC from baseline day 2 versus light day 4: $n = 10$ syllables from 8 birds, $p = 0.00005$; $t(18) = -5.247$; unpaired two-tailed t-test: absolute change in auROC from baseline day 2 versus light day 4 for control birds: $n = 7$ syllables from 6 birds, $p = 0.639$; $t(15) = -0.479$; unpaired t-test: absolute change in auROC from light day 4 experimental birds versus light day 4 for control birds: $n = 7$ syllables from 6 birds; $p = 0.0004$; $t(15) = -4.537$; green, GFP; grey, no injection; purple, 100% contingency. (J) Mean absolute percent change in pitch frequency between the last day of baseline and: the first day of baseline ($0.596 \pm 0.176\%$); the last day of pitch-contingent optogenetic stimulation of VTA_X terminals in experimental birds ($1.929 \pm 0.512\%$); and the last day of pitch-contingent optogenetic stimulation in Area X of control birds ($0.302 \pm 0.083\%$). Paired two tailed t-test: absolute percent change in pitch from baseline day 2 versus light day 4: $n = 10$ syllables from 8 birds; $p = 0.008$; $t(18) = -2.989$; unpaired t-test: absolute percent change in pitch from baseline day 2 versus light day 4 for control birds: $n = 7$ syllables from 6 birds; $p = 0.198$; $t(15) = -1.348$; unpaired t-test: absolute percent change in pitch from light day 4 experimental birds versus light day 4 for control birds, $n = 7$ syllables from 6 birds; $p = 0.009$; $t(15) = -2.990$. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. All values shown as mean \pm standard error of the mean (S.E.M.).

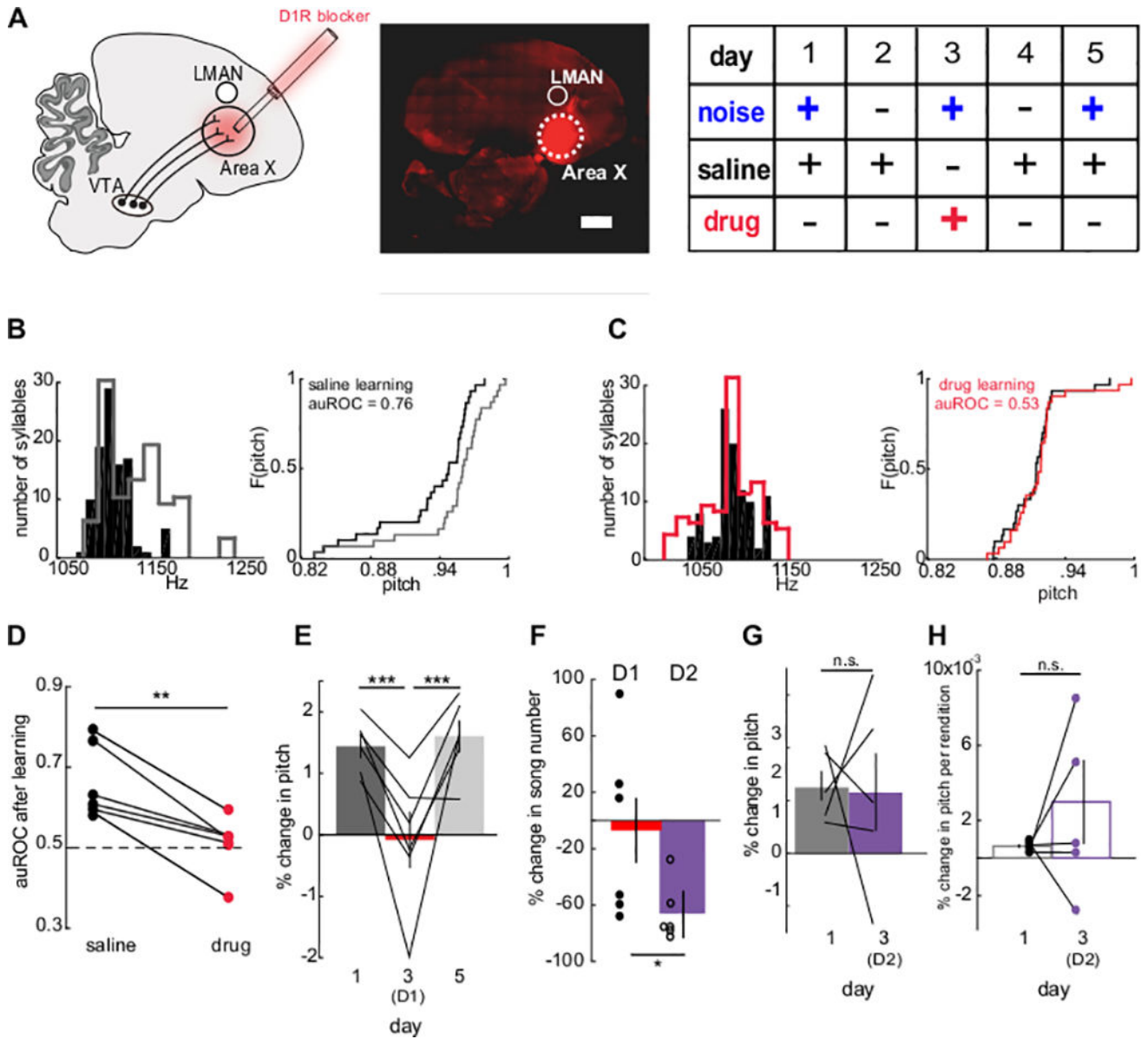


Figure 3. Adult pitch learning requires activation of D1 receptors in Area X

(A) *Left*, schematic. *Middle*, infusion of muscimol-BODIPY through microdialysis probes to visualize drug spread (shown is an exemplar brain section from one of 6 birds in which BODIPY labeling was visualized). Scale bar, 1 mm. *Right*, experimental design. (B) *Left*, target syllable pitch distribution before (black) and after WN with saline infusion (grey) from one bird on day 1. *Right*, cumulative distribution of normalized target syllable pitch during first (black) and last third (grey) of day 1. (C) As in B) but with infusion of SCH23390 (red) instead of saline. (D) Change in auROC of target syllable pitch after WN during saline (black) or SCH23390 (red) infusion. Paired two tailed t-test: saline learning (day 1): 0.663 ± 0.034 ; SCH23390 learning (day 3): 0.512 ± 0.027 ; $n = 6$ syllables from 6 birds, $p = 0.004$; $t(5) = 5.091$. (E) Mean percent change in target syllable pitch after WN during saline (dark grey (day 1, $1.443 \pm 0.164\%$) and light grey (day 5, $1.609 \pm 0.225\%$)) or

SCH23390 (red, day 3, $-0.085 \pm 0.410\%$) infusion. Paired two tailed t-test for learning: $n = 6$, for saline learning vs. SCH23390 learning (day 1 vs. day 3): $p = 0.006$; $t(5) = 4.626$; for SCH23390 learning vs. saline learning after SCH23390 learning (day 3 vs. day 5): $p = 0.018$; $t(5) = -3.468$; for saline learning vs. saline learning after SCH23390 learning (day 1 vs. day 5): $p = 0.591$; $t(5) = -0.573$. (F) Mean percent change from baseline number of songs sung during SCH23390 (red) or D2 antagonist sulpiride infusion (purple). Unpaired two-tailed t-test: percent change in number of songs: D1 antagonist: $-7.810 \pm 23.175\%$, $n = 6$ birds; D2 antagonist: $-66.12 \pm 17.38\%$, $n = 6$ birds; $p = 0.05$; $t(10) = 2.181$. (G) Percent change in target syllable pitch after WN during saline (grey) or sulpiride (purple). Paired two tailed t-test for learning: saline learning (day 1): $1.282 \pm 0.266\%$; sulpiride learning (day 3 (D2)): $1.159 \pm 0.725\%$; $n = 5$, $p = 0.91$; $t(4) = 0.121$. (H) Mean percent change in target syllable pitch per rendition after WN with saline (grey) or sulpiride (purple). Paired two tailed t-test for learning per rendition: saline learning (day 1): $0.00058 \pm 0.00008\%$; sulpiride learning (day 3 (D2)): $0.00246 \pm 0.001834\%$; $n = 5$, $p = 0.42$; $t(4) = -0.902$. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. All values shown as mean \pm standard error of the mean (S.E.M.).

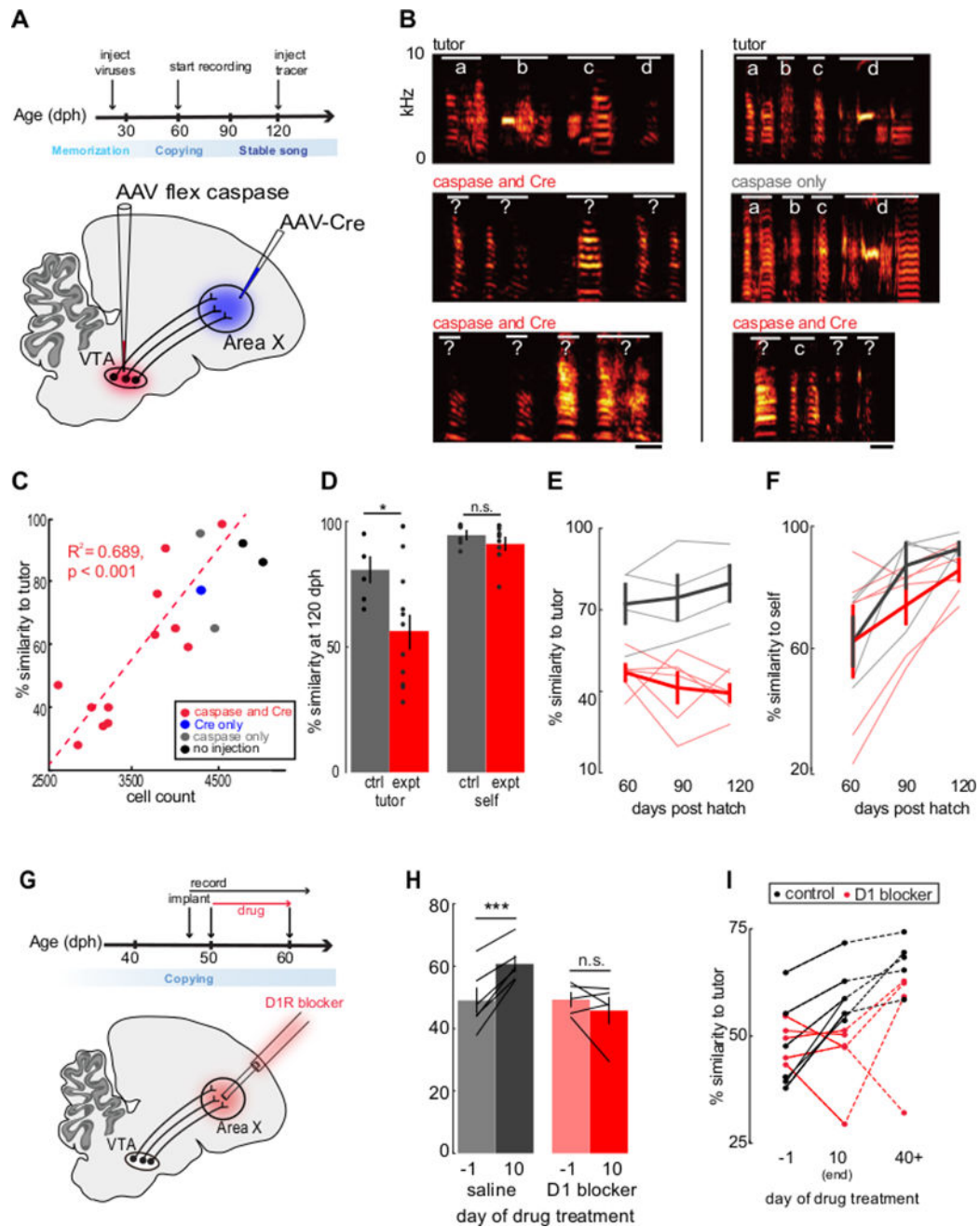


Figure 4. VTA_X neurons and D1 receptor activity in Area X are necessary for accurate song copying in juvenile zebra finches

(A) Experimental design. (B) Example sonograms, red indicates VTA_X ablation birds, grey indicates control bird injected with only caspase in VTA. Scale bar, 50 milliseconds. (C) Mean percent similarity to tutor as a function of remaining VTA_X neuron number. Linear regression: $R^2 = 0.689$, $n = 12$ experimental birds, $p = 0.0008$; $F(11) = 22.2$; degrees of freedom: 11. (D) Mean percent similarity to tutor (*left*) and to adult self (*right*). Paired two-tailed t-test: percent tutor similarity: control: $83.00 \pm 4.87\%$, $n = 5$; experimental: $56.25 \pm 6.61\%$, $n = 12$; $p = 0.027$; $t(15) = -2.442$; paired two-tailed t-test: percent self-similarity: control: $95.07 \pm 1.82\%$, $n = 5$; experimental: $90.98 \pm 2.76\%$, $n = 12$; $p = 0.34$; $t(15) =$

–0.977. (E) Mean percent similarity to tutor over the course of sensorimotor learning. Red, experimental birds (n = 6). Grey, control birds (n = 4). (F) Mean percent similarity to adult self over the course of sensorimotor learning. Red, experimental birds (n = 6). Grey, control birds (n = 4). (G) Experimental design. (H) Percent similarity to tutor one day before and 10 days after beginning of daily microdialysis for control (grey) (n = 3 implanted birds infused with saline and n = 3 intact birds) and SCH23390 birds (red). 2-factor repeated measures ANOVA: A × B: p = 0.002, F(1): 19.45; paired two-tailed t-test: control birds: saline day –1: 49.04 ± 3.79%; saline day 10: 60.75 ± 2.46%, n = 6, p = 0.001; t(5) = –6.787; paired two-tailed t-test: experimental birds: drug day –1: 49.29 ± 1.72%; drug day 10: 45.78 ± 1.03%, n = 5, p = 0.33; t(4) = 1.101. (I) Percent similarity to tutor one day before, 10 days after, and >40 days after beginning of microdialysis for SCH23390 (red) and control birds (black) Unpaired two-tailed t-test: control birds: 40+ days after beginning of treatment: 67.08 ± 2.30%, n = 5; experimental birds: 40+ days after beginning of treatment: 54.19 ± 6.40%, n = 4, p = 0.17; t(6) = –1.549. *P < 0.05, **P < 0.01, ***P < 0.001. All values shown as mean ± standard error of the mean (S.E.M.).