Autism-linked gene FoxP1 selectively regulates the cultural transmission of learned vocalizations

Francisco Garcia-Oscos[1*], Therese Koch[1*], Harshida Pancholi[1*], Massimo Trusel[1*], Vamsi Daliparthi[1], Fatma Ayhan[1], Marissa Co[1], Danyal H. Alam[1], Jennifer E. Holdway[1], Genevieve Konopka[1], Todd F. Roberts[1#]
[1]Department of Neuroscience, UT Southwestern Medical Center, Dallas, TX, USA.
*Equal contribution
[#]Correspondence should be addressed to TFR (Todd.Roberts@utsouthwestern.edu)

Autism spectrum disorders (ASD) are characterized by impaired learning of culturally transmitted behaviors like social skills, speech, and language[1-3]. These behaviors are learned by copying parents and other social models during development, a two-stage process that involves forming memories of appropriate behaviors during social experiences and then using those memories to guide imitation. How ASD-linked genes impair these often-intertwined aspects of learning is not known, thereby limiting our understanding of the developmental progression of ASD and the targeting of therapeutic interventions. Here we show that these aspects of learning are dissociable and that the ASD-linked gene *FoxP1* selectively impairs learning from social experience, but not behavioral imitation. Haploinsufficiency of *FOXP1* in humans causes FOXP1 syndrome, a neurodevelopmental disorder typified by severe disruptions in speech and language development, and other ASD-associated symptoms[4,5]. We tested how knockdown of *FoxP1* (FP1-KD) affects the cultural transmission of vocal behaviors in zebra finches, a songbird that learns by memorizing and vocally copying the song of an adult 'song-tutor'. We find that FP1-KD blocks song learning in juvenile birds by selectively impairing their ability to encode a memory during social experiences with a song-tutor. These learning deficits are linked to disruptions in experience-driven structural and functional plasticity. However, if birds are exposed to tutor-song prior to FP1-KD, their ability to imitate that song during development is unaffected. Thus, FP1-KD impairs cultural transmission of vocalizations by disrupting the ability to form appropriate vocal memories, yet spares the ability to use previously acquired memories to guide vocal learning. This indicates that learning from social experience may be particularly vulnerable in FOXP1 syndrome.

Humans and other animals learn many of their complex and socially oriented behaviors by imitating more experienced individuals in their environment. For example, development of spoken language is rooted in a child's ability to imitate the speech patterns of their parent(s) and other adults [6-8]. This cultural transmission of behavior is impaired in many neurodevelopmental disorders, most notably ASD[1-3]. However, how ASD risk genes impact behavioral imitation is still not known. We sought to examine this issue by testing the role of FoxP1 in the cultural transmission of song between adult and juvenile zebra finches (Fig. 1a-d, Extended Data Fig. 1).

36 FOXP1 (forkhead-box protein 1) is one of the top ASD-associated genes, and its haploinsufficiency causes specific
37 language impairment and intellectual disability in children[5]. *FoxP1* is expressed in many of the same areas of the pallium
38 and basal ganglia in mammals and songbirds[9-11]. In zebra finches, *FoxP1* expression is enriched in many brain regions
39 that are known to be important for song learning[10-12] (Fig. 1e, Extended Data Fig. 2). Here we focus on the role of FoxP1
40 in the pallial region HVC, which is involved in the formation of song memories, the vocal imitation process, and in
41 production of the learned song[13-19]. We first characterized FoxP1 expression in HVC projection cell-types, then tested
42 its necessity in different aspects of song learning. Lastly, we identified anatomical, electrophysiological and circuit
43 abnormalities associated with disrupted expression of FoxP1 during vocal learning.
44
45 **FoxP1 is Expressed in Striatal Projecting HVC_X Neurons**
46 HVC has three non-overlapping classes of projection neurons[20] (Fig. 1e). HVC_X and HVC_{Av} neurons transmit vocal motor-
47 related signals to the striatum or auditory system, respectively. These two pathways are important for behavioral
48 imitation of tutor-song in juvenile birds, but are not essential for song production in adult birds[20-22]. HVC_{RA} neurons
49 provide descending motor commands to the pallial nucleus RA, a connection which is necessary for the production of
50 learned song at all stages of life[17,21,23]. We used anatomical tracing and immunolabeling to examine FoxP1 expression
51 in these different classes of projection neurons. We found that FoxP1 is expressed in the majority of HVC neurons
52 projecting to the striatum (74.41 ± 2.17% of HVC_X neurons) and in a much smaller percentage of neurons projecting to
53 the auditory system or to RA (24.27 ± 2.54% of HVC_{Av} neurons, 29.98 ± 0.65% of HVC_{RA} neurons, Fig. 1f, Extended Data
54 Fig. 2a-f). The widespread expression of FoxP1 in striatal projecting HVC_X neurons raises speculation that FoxP1 could
55 function in sensorimotor aspects of song learning. HVC_X neurons are thought to carry timing information about song
56 to the basal ganglia, which facilitate accurate reinforcement-based imitation of syllables and sub-syllable elements[24-
57 27]. In support of this view, genetic lesions of HVC_X neurons have recently been shown to significantly disrupt behavioral
58 imitation of song in juvenile male zebra finches[22].
59
60
61 **Knockdown of FoxP1 in HVC Blocks Learning from Social Experiences but not Behavioral Imitation of Song**
62 To test the function of FoxP1 in the social transmission of birdsong we developed a short hairpin RNA (shRNA) against
63 FoxP1 and showed that it can significantly knockdown expression of FoxP1 *in vivo* and *in vitro* (Fig. 1g-h, Extended Data
64 Fig. 2h). Using an adeno associated virus to express this shFoxP1, we knocked-down expression of FoxP1 (FP1-KD) in
65 age-matched juvenile birds prior to different aspects of song learning.
66
67 Juvenile male zebra finches memorize the song of their father or other adult song-tutor(s) during social interactions.
68 They then use auditory feedback and extensive practice to learn how to accurately imitate this memorized song by 90-
69 100 days post hatching (dph)[18]. Juvenile birds can memorize the song of a tutor at any time between 20-60 dph, but

they don't start to practice singing until approximately 35-40 dph (Fig 1a). This developmental progression and the ability to raise birds in groups without a song-tutor – hereby referred to as 'isolates' – allowed us to knock down *FoxP1* expression prior to behavioral imitation of song, either before or after birds had an opportunity to form a memory of a tutor-song[16] (Fig. 2a: tutoring pre-FP1-KD, referred to as 'behavioral imitation' group (**FP1-KD BI**), Fig.2e: tutoring post-FP1-KD, referred to as 'social experience' group (**FP1-KD SE**)). Given the widespread expression of FoxP1 in striatal projecting HVC_X neurons, we hypothesized that FP1-KD might disrupt behavioral imitation of the tutors' song[22]. However, we found that FP1-KD after birds had an opportunity to form a memory of the tutor-song (FP1-KD BI birds) did not disrupt their ability to then copy that song during development (Fig. 2a-d). In fact, the adult songs of the FP1-KD BI group were stereotyped and indistinguishable from normal zebra finch song, suggesting that FoxP1 in HVC is not necessary for behavioral imitation in juvenile birds or in more basic aspects of song production.

In contrast, FP1-KD prior to social experience with a tutor severely disrupted subsequent song learning (FP1-KD SE birds, Fig. 2e-h). We found that birds with FP1-KD prior to social experience learned little from their song-tutor and statistically less than birds in the FP1-KD BI or control birds (Fig. 2a-l, q-r). As adults, FP1-KD SE birds sang songs that were highly variable from trial-to-trial and with entropy rates (Methods) higher than either the FP1-KD BI or control birds (Fig. 2a-l, s-t). Rather, their songs had entropy rates that were statistically indistinguishable from birds that were never tutored during development (full isolates, Fig. 2 m-p, u). Together, these findings suggest that FP1-KD critically impairs the cultural transmission of vocal behavior, potentially by selectively disrupting the ability to form appropriate vocal memories during social experiences.

Knockdown of FoxP1 Inhibits Dendritic Spine Turnover and Decreases the Intrinsic Excitability of HVC_X Neurons

We next sought to identify the consequences of FP1-KD on structural and functional plasticity in HVC. Previous research has shown that plasticity in HVC is predictive of a young bird's ability to form tutor-song memories during social interactions[13,14]. Birds with high levels of dendritic spine turnover are better learners than birds with low levels of spine turnover[14]. Therefore, we first used longitudinal *in vivo* two-photon imaging to track FP1-KD mediated changes to dendritic spines in our social experience groups and in separate cohorts of juvenile isolates (Fig. 3a). We imaged spines on virally transduced, retrogradely labeled HVC_X neurons, both in our FP1-KD cohorts and in control birds. We found that HVC_X neurons had significantly higher spine density in FP1-KD adults and lower spine density in FP1-KD juveniles than in their respective age-matched controls (Fig. 3b-c). This suggests that FP1-KD disrupts normal changes in spine density that occur during song learning. Perhaps more pertinently, we found that FP1-KD signifcantly reduced spine turnover in both juvenile and adult birds compared to age-matched control birds (Fig. 3d-e). We then verified these results with independent analysis by 4 additional scorers on a randomized subset of this data ($P < 0.05$, Paired t-test). Therefore, plasticity in the striatal projecting HVC_X neuronal population may be necessary for young birds to learn from social experiences and to form the vocal memories used to guide song imitation.

This decreased spine turnover in FP1-KD birds could be tied to differences in the intrinsic excitability of HVC_X neurons. For example, deafening of adult zebra finches causes increased intrinsic excitability and increased dendritic spine turnover on HVC_X neurons[28]. To examine this, we conducted whole-cell current-clamp recordings of HVC_X neurons from our social experience groups (FP1-KD and control) and from cohorts of juvenile isolates (Fig. 3f). We found that neurons expressing shFoxP1 had significantly decreased intrinsic excitability compared to HVC_X neurons recorded from age-matched controls (Fig. 3g-h).

Knockdown of FoxP1 Blocks Synaptic and Network Hallmarks of Learning during Social Experience

We next explored if social experience with a tutor, and presumably formation of a vocal memory, is sufficient to elicit changes in the intrinsic excitability and synaptic physiology of HVC_X neurons. We made comparisons between juvenile isolates and age-matched birds that were housed with a song tutor for two consecutive days (Fig. 4a). We found that this acute social experience with a tutor drove a significant increase in the intrinsic excitability of HVC_X neurons in control birds. FP1-KD prevented these changes in HVC_X neurons (Fig. 4b). With regards to synaptic properties, we found that neither FP1-KD nor social experience affected the excitation-to-inhibition ratio (E/I ratio) in HVC_X neurons (Fig. 4c). However, social experience led to a significant increase in AMPA/NMDA receptor ratios in HVC_X neurons from control birds (Fig. 4d), a result consistent with synaptic strengthening following tutor experience and the known dependence on NMDA receptor activity in HVC for encoding of tutor song memories[13]. In contrast, FP1-KD prevented this AMPA/NMDA increase following social experience. These results indicate that FP1-KD blocks key synaptic and intrinsic electrophysiological signatures of learning following social experiences with a song-tutor.

Lastly, we tested whether these cellular and synaptic effects of FP1-KD were sufficient to block *in vivo*, network-level hallmarks of song memory. Acquisition of tutor song memories following brief social experiences is correlated with the emergence of prolonged patterns of bursting activity in HVC and in RA[14,15,29]. Taking advantage of the lack of a corpus callosum connecting HVC from the right and left hemispheres[30], we conducted experiments in juvenile isolates in which we knocked down FoxP1 in only one, pseudo-randomly assigned, brain hemisphere and expressed a control virus in the other. Birds were either maintained in social isolation from a song-tutor (Extended Data Fig. 4a) or given two days of social experience with a song-tutor (Fig. 4e), before making bilateral extracellular recordings from HVC to assess baseline and learning related changes in network activity. In isolate birds (baseline condition), we did not detect any network level differences in spontaneous neuronal activity between the FP1-KD and control hemispheres (Extended Data Fig. 4 and 5a-b). This indicates that FP1-KD does not cause large scale changes in the excitability or bursting properties of HVC neurons in the absence of prior tutor-song experience.

137 In contrast, following two-days of social experience with a tutor, we observed large scale differences in the
138 spontaneous bursting activity recorded in the control and FP1-KD hemispheres. We observed significantly fewer bursts
139 in the FP1-KD hemispheres, and these bursts had fewer spikes and were shorter than bursts recorded in the control
140 brain hemispheres (Fig. 4g-m). The total number of spikes, inter-spike interval distribution, and inter-spike interval
141 within bursts were not affected by FP1-KD (Fig. 4f and Extended Data Fig. 5c-d). This suggests that while overall activity
142 levels were preserved in the two hemispheres, the tutor-song experience drove redistribution of activity into sustained
143 bursting patterns in the control, but not in the FP1-KD hemisphere. This indicates that knockdown of FoxP1 is sufficient
144 to block circuit level hallmarks of learning from social experiences and encoding of vocal memories that guide the
145 cultural transmission of song.
146
147
148 **Discussion**
149 How ASD risk genes differentially impact learning of culturally transmitted behaviors is difficult to study in traditional
150 model organisms; many don't robustly imitate social behaviors nor transmit behavioral repertoires from one
151 generation to the next[31]. Our results demonstrate that FP1-KD produces a series of synaptic, cellular and network
152 deficits, some of which resemble those recently described in mouse models[32-34]. However, the characteristics of the
153 songbird system allowed us to test the function of FoxP1 in social experience-dependent imitation of behavior. We
154 show that FoxP1 in HVC is essential for young birds to encode memories that are used to guide imitation, but not for
155 the imitation of a previously acquired song memory. Identified roles of HVC include motor control of song, vocal
156 imitation, and encoding of tutor-song memories[13-20]. Therefore, FP1-KD mediated selective disruption of learning from
157 social experiences is surprising and indicates a specialized and essential role for the network of FoxP1-expressing HVC
158 neurons in the cultural transmission of vocal behaviors.
159
160 Notably, we find that FoxP1 is widely expressed in striatal projecting HVC_X neurons. These neurons have previously
161 been shown to have properties resembling 'mirror' neurons[35-38], but their role in the cultural transmission of vocal
162 behaviors was not well understood. The functional significance of mirror neurons in behavioral imitation and ASD is a
163 topic of much debate[39-44]. We show that knockdown of FoxP1 in striatal projecting HVC_X neurons inhibits their spine
164 plasticity, dampens their intrinsic excitability, and blocks the cellular and network level signatures associated with
165 tutor-song memory formation following social experience. Although it is not yet clear whether the disruptions in
166 learning shown here depend exclusively on HVC_X neurons, as they are not the only cell-type in HVC expressing FoxP1,
167 our research provides proof that a detailed examination of how ASD-linked genes affect different aspects of vocal
168 imitation can offer novel and important insights into their roles in neurodevelopmental disorders.
169
170

Acknowledgments: The authors thank members of the Roberts and Konopka laboratories for discussion and comments on the manuscript, Andrea Guerrero and Matthew Harper for laboratory support. **Funding:** This research was supported by grants from the US National Institutes of Health R21DC016340 to TFR and GK, R01NS108424 & R01DC014364 to TFR, R01MH102603 to GK, and from the National Science Foundation IOS-1457206 to TFR. FA and DHA were supported by T32HL139438. **Author contributions:** TFR and GK conceived the project. TFR supervised the research. FG-O, TK, HP, MT and TFR designed the experiments and wrote the manuscript with input from all authors. FG-O, TK, HP, MT, VD, MC, DHA and JEH collected and analyzed the data for the project. GK, MC and FA designed and tested the knockdown viral construct. All authors read and commented on the manuscript. **Competing interests:** Authors declare no competing interests. **Data and materials availability:** All data is available in the main text or the supplementary materials.

Supplemental Information:

Materials and Methods

Extended Data Figures 1 to 5

References

1 Mody, M. & Belliveau, J. W. Speech and Language Impairments in Autism: Insights from Behavior and Neuroimaging. *N Am J Med Sci (Boston)* **5**, 157-161 (2013).

2 Sperdin, H. F. & Schaer, M. Aberrant Development of Speech Processing in Young Children with Autism: New Insights from Neuroimaging Biomarkers. *Frontiers in neuroscience* **10**, 393, doi:10.3389/fnins.2016.00393 (2016).

3 Tager-Flusberg, H., Paul, R. & Lord, C. E. in *Handbook of Autism and Pervasive Developmental Disorders, and Behavior* Vol. Vol. 1. Diagnosis, Development, Neurobiology (ed R. Paul F. Volkmar, A. Klin, & D.J. Cohen (Eds.)) 335–364 (Wiley, 2005).

4 Meerschaut, I. *et al.* FOXP1-related intellectual disability syndrome: a recognisable entity. *J Med Genet* **54**, 613-623, doi:10.1136/jmedgenet-2017-104579 (2017).

5 Siper, P. M. *et al.* Prospective investigation of FOXP1 syndrome. *Molecular autism* **8**, 57, doi:10.1186/s13229-017-0172-6 (2017).

6 Bolhuis, J. J., Okanoya, K. & Scharff, C. Twitter evolution: converging mechanisms in birdsong and human speech. *Nat Rev Neurosci* **11**, 747-759, doi:10.1038/nrn2931 (2010).

7 Doupe, A. J. & Kuhl, P. K. Birdsong and human speech: common themes and mechanisms. *Annu Rev Neurosci* **22**, 567-631 (1999).

8 Konopka, G. & Roberts, T. F. Insights into the Neural and Genetic Basis of Vocal Communication. *Cell* **164**, 1269-1276, doi:10.1016/j.cell.2016.02.039 (2016).

9 Tamura, S., Morikawa, Y., Iwanishi, H., Hisaoka, T. & Senba, E. Foxp1 gene expression in projection neurons of the mouse striatum. *Neuroscience* **124**, 261-267, doi:10.1016/j.neuroscience.2003.11.036 (2004).

10 Teramitsu, I., Kudo, L. C., London, S. E., Geschwind, D. H. & White, S. A. Parallel FoxP1 and FoxP2 expression in songbird and human brain predicts functional interaction. *J Neurosci* **24**, 3152-3163, doi:10.1523/JNEUROSCI.5589-03.2004 (2004).

11 Mendoza, E. *et al.* Differential coexpression of FoxP1, FoxP2, and FoxP4 in the Zebra Finch (Taeniopygia guttata) song system. *J Comp Neurol* **523**, 1318-1340, doi:10.1002/cne.23731 (2015).

12 Norton, P., Barschke, P., Scharff, C. & Mendoza, E. Differential Song Deficits after Lentivirus-Mediated Knockdown of FoxP1, FoxP2, or FoxP4 in Area X of Juvenile Zebra Finches. *J Neurosci* **39**, 9782-9796, doi:10.1523/JNEUROSCI.1250-19.2019 (2019).
13 Roberts, T. F., Gobes, S. M., Murugan, M., Olveczky, B. P. & Mooney, R. Motor circuits are required to encode a sensory model for imitative learning. *Nat Neurosci* **15**, 1454-1459, doi:10.1038/nn.3206 (2012).
14 Roberts, T. F., Tschida, K. A., Klein, M. E. & Mooney, R. Rapid spine stabilization and synaptic enhancement at the onset of behavioural learning. *Nature* **463**, 948-952, doi:10.1038/nature08759 (2010).
15 Tanaka, M., Sun, F., Li, Y. & Mooney, R. A mesocortical dopamine circuit enables the cultural transmission of vocal behaviour. *Nature* **563**, 117-120, doi:10.1038/s41586-018-0636-7 (2018).
16 Zhao, W., Garcia-Oscos, F., Dinh, D. & Roberts, T. F. Inception of memories that guide vocal learning in the songbird. *Science* **366**, 83 - 89 (2019).
17 Aronov, D., Andalman, A. S. & Fee, M. S. A specialized forebrain circuit for vocal babbling in the juvenile songbird. *Science* **320**, 630-634 (2008).
18 Ikeda, M. Z., Trusel, M. & Roberts, T. F. Memory circuits for vocal imitation. *Curr Opin Neurobiol* **60**, 37-46, doi:10.1016/j.conb.2019.11.002 (2020).
19 Vallentin, D., Kosche, G., Lipkind, D. & Long, M. A. Neural circuits. Inhibition protects acquired song segments during vocal learning in zebra finches. *Science* **351**, 267-271, doi:10.1126/science.aad3023 (2016).
20 Roberts, T. F. *et al.* Identification of a motor-to-auditory pathway important for vocal learning. *Nat Neurosci*, doi:10.1038/nn.4563 (2017).
21 Scharff, C., Kirn, J. R., Grossman, M., Macklis, J. D. & Nottebohm, F. Targeted neuronal death affects neuronal replacement and vocal behavior in adult songbirds. *Neuron* **25**, 481-492 (2000).
22 Sanchez-Valpuesta, M. *et al.* Corticobasal ganglia projecting neurons are required for juvenile vocal learning but not for adult vocal plasticity in songbirds. *Proc Natl Acad Sci U S A* **116**, 22833-22843, doi:10.1073/pnas.1913575116 (2019).
23 Hahnloser, R., Kozhevnikov, A. & Fee, M. An ultra-sparse code underlies the generation of neural sequences in a songbird. *Nature* **419**, 65-70 (2002).
24 Brainard, M. S. & Doupe, A. J. Auditory feedback in learning and maintenance of vocal behaviour. *Nat Rev Neurosci* **1**, 31-40 (2000).
25 Mooney, R. Neurobiology of song learning. *Curr Opin Neurobiol* **19**, 654-660, doi:10.1016/j.conb.2009.10.004 (2009).
26 Fee, M. S. & Goldberg, J. H. A hypothesis for basal ganglia-dependent reinforcement learning in the songbird. *Neuroscience* **198**, 152-170, doi:10.1016/j.neuroscience.2011.09.069 (2011).
27 Mooney, R. The neurobiology of innate and learned vocalizations in rodents and songbirds. *Curr Opin Neurobiol* **64**, 24-31, doi:10.1016/j.conb.2020.01.004 (2020).
28 Tschida, K. A. & Mooney, R. Deafening drives cell-type-specific changes to dendritic spines in a sensorimotor nucleus important to learned vocalizations. *Neuron* **73**, 1028-1039, doi:10.1016/j.neuron.2011.12.038 (2012).
29 Shank, S. S. & Margoliash, D. Sleep and sensorimotor integration during early vocal learning in a songbird. *Nature* **458**, 73-77 (2009).
30 Wang, C. Z., Herbst, J. A., Keller, G. B. & Hahnloser, R. H. Rapid interhemispheric switching during vocal production in a songbird. *PLoS Biol* **6**, e250, doi:10.1371/journal.pbio.0060250 (2008).
31 Konopka, G. & Roberts, T. F. Animal Models of Speech and Vocal Communication Deficits Associated With Psychiatric Disorders. *Biological psychiatry* **79**, 53-61, doi:10.1016/j.biopsych.2015.07.001 (2016).
32 Araujo, D. J. *et al.* Foxp1 in Forebrain Pyramidal Neurons Controls Gene Expression Required for Spatial Learning and Synaptic Plasticity. *J Neurosci* **37**, 10917-10931, doi:10.1523/JNEUROSCI.1005-17.2017 (2017).
33 Bacon, C. *et al.* Brain-specific Foxp1 deletion impairs neuronal development and causes autistic-like behaviour. *Molecular psychiatry* **20**, 632-639, doi:10.1038/mp.2014.116 (2015).
34 Araujo, D. J. *et al.* FoxP1 orchestration of ASD-relevant signaling pathways in the striatum. *Genes & development* **29**, 2081-2096, doi:10.1101/gad.267989.115 (2015).
35 Prather, J. F., Peters, S., Nowicki, S. & Mooney, R. Precise auditory-vocal mirroring in neurons for learned vocal communication. *Nature* **451**, 305-310, doi:10.1038/nature06492 (2008).

36 Fujimoto, H., Hasegawa, T. & Watanabe, D. Neural coding of syntactic structure in learned vocalizations in the songbird. *J Neurosci* **31**, 10023-10033, doi:10.1523/JNEUROSCI.1606-11.2011 (2011).
37 Rizzolatti, G. & Arbib, M. A. Language within our grasp. *Trends Neurosci* **21**, 188-194 (1998).
38 Hamaguchi, K., Tschida, K. 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, doi:10.7554/eLife.01833 (2014).
39 Hickok, G. Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *J Cogn Neurosci* **21**, 1229-1243, doi:10.1162/jocn.2009.21189 (2009).
40 Buccino, G. & Amore, M. Mirror neurons and the understanding of behavioural symptoms in psychiatric disorders. *Current opinion in psychiatry* **21**, 281-285, doi:10.1097/YCO.0b013e3282fbcd32 (2008).
41 Iacoboni, M. & Dapretto, M. The mirror neuron system and the consequences of its dysfunction. *Nat Rev Neurosci* **7**, 942-951, doi:10.1038/nrn2024 (2006).
42 Rizzolatti, G. & Craighero, L. The mirror-neuron system. *Annu Rev Neurosci* **27**, 169-192, doi:10.1146/annurev.neuro.27.070203.144230 (2004).
43 Hamilton, A. F. Reflecting on the mirror neuron system in autism: a systematic review of current theories. *Developmental cognitive neuroscience* **3**, 91-105, doi:10.1016/j.dcn.2012.09.008 (2013).
44 Grabenhorst, F., Baez-Mendoza, R., Genest, W., Deco, G. & Schultz, W. Primate Amygdala Neurons Simulate Decision Processes of Social Partners. *Cell* **177**, 986-998 e915, doi:10.1016/j.cell.2019.02.042 (2019).

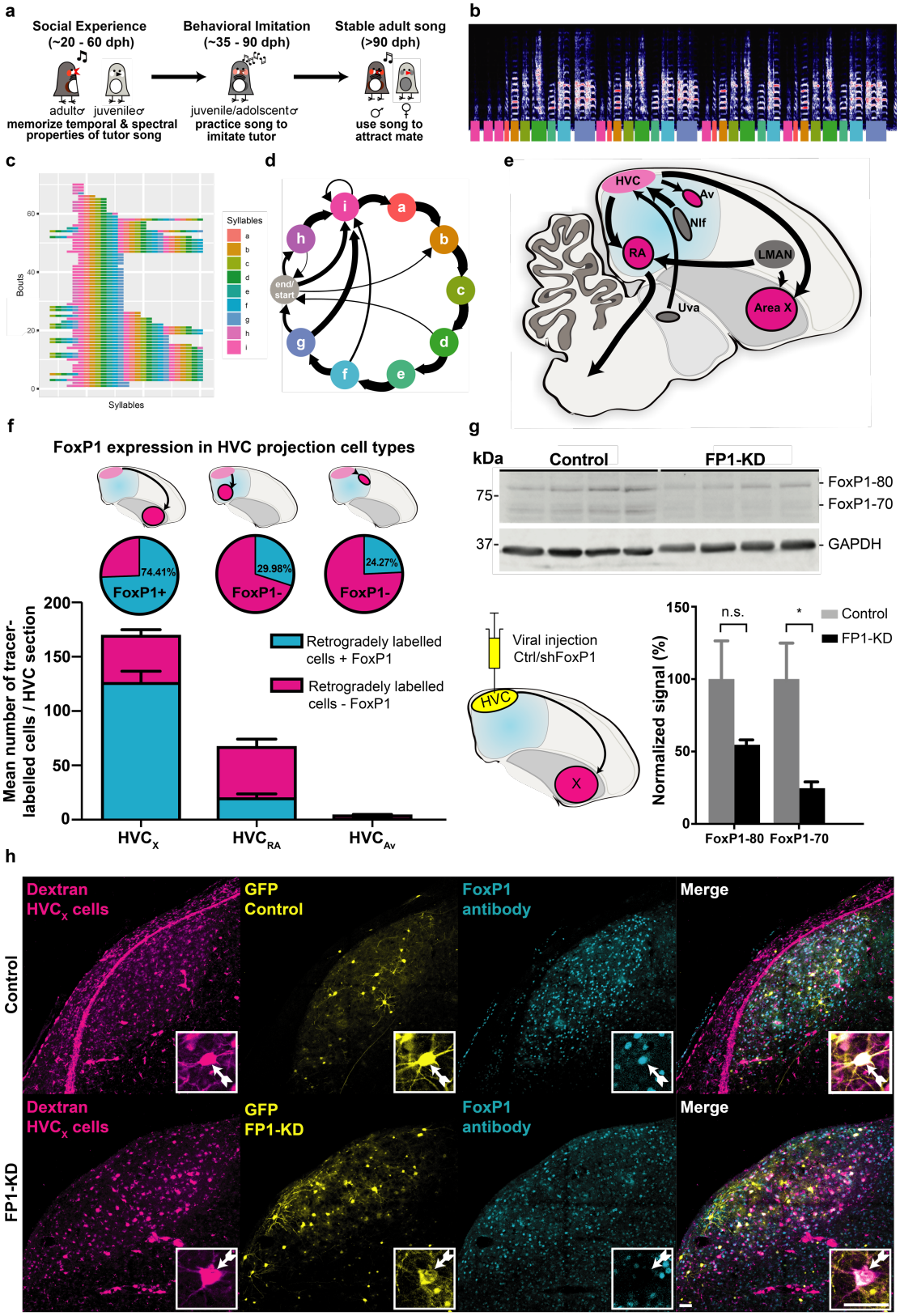


Figure 1. FoxP1 Expression in HVC of Control and Knock-down Zebra Finches.
a, Timeline of zebra finch song learning in juvenile males (dph = days post hatch). **b-d,** Example representations of an adult zebra finch song; each color represents a syllable or note in the song: **b,** Spectrogram of an adult male's song. Frequency is represented along the y-axis, time is represented along the x-axis, and color reflects intensity. Colored bars underneath indicate introductory notes (pink, i) and syllables (a-h). **c,** A syntax raster plot showing the syllables

292 sung over repeated song bouts, colors reflect the syllables produced. **d,** A representation of song syllable syntax as a
293 first-order Markov model, with thickness of arrows representing the probability of syllable transitions. **e,** Parasagittal
294 schematic of the song circuit, with relevant nuclei labelled: Area X, striato-pallidal basal ganglia nucleus; Av, nucleus
295 avalanche; HVC, premotor song nucleus; LMAN, lateral magnocellular nucleus of the anterior nidopallium; NIf, nucleus
296 interfacialis of the nidopallium; Uva, nucleus uvaeformis; RA, robust nucleus of the arcopallium. **f,** FoxP1 expression in
297 HVC projection neurons. Subtypes of HVC neurons, L-R: HVC_X neurons, HVC_{RA} neurons, HVC_{Av} neurons. (top) Schematic
298 of retrograde injections into areas downstream of HVC, (middle) pie charts showing the proportion of cells that express
299 FoxP1 for each cell-type, and (bottom) proportion of FoxP1 expressing neurons for each HVC subtype, per HVC section
300 (HVC_X: 74.4±2.2%, *n*=3 birds, 6 hemispheres; HVC_{RA}: 30.0±0.7%, *n*=2 birds, 4 hemispheres; HVC_{Av}: 24.3±2.5%, *n*=3 birds,
301 6 hemispheres). **g,** Western blot of lysates from HVC injected with control or shFoxP1 AAV. At least two isoforms of
302 FoxP1 were observed around 80 kDa and 70 kDa, with GAPDH as loading control. Schematic of injections for the tissues
303 collected after viral injections of rAAV9/ds-CBh-GFP (*n* = 4 birds) or pscAAV-GFP-shFoxP1 (*n* = 4 birds) into HVC, paired
304 with retrograde tracer in Area X. Graph shows quantification of FoxP1 protein. Signals were normalized to GAPDH,
305 averaged for each condition, and normalized to the Controls. Histograms represent average±SEM. (FoxP1-80: Control:
306 100±26.6% vs FP1-KD: 54.7±3.5%, Student's t-test, $P > 0.05$; FoxP1-70: Control: 100±25.0% vs FP1-KD: 24.7±4.4%,
307 Student's t-test, $P = 0.025$). **h,** Representative examples of HVC sections from Control (top) and FP1-KD (bottom) birds.
308 Injections were performed as in schematic in panel 'g'. Panels show HVC_X cells labelled with retrograde tracer in Area
309 X (magenta, left), GFP signal from AAV-Control/ AAV-shFoxP1 injection (yellow, middle left), FoxP1 staining with
310 antibody (cyan, middle right), and a merged composite (right). Inset boxes indicate one example cell per condition.
311 White arrowheads indicate the soma of the example neurons. The FP1-KD cell (bottom insets) does not show FoxP1
312 labelling, whereas it is present in the Control cell (top insets). Scale bars, 50 µm.

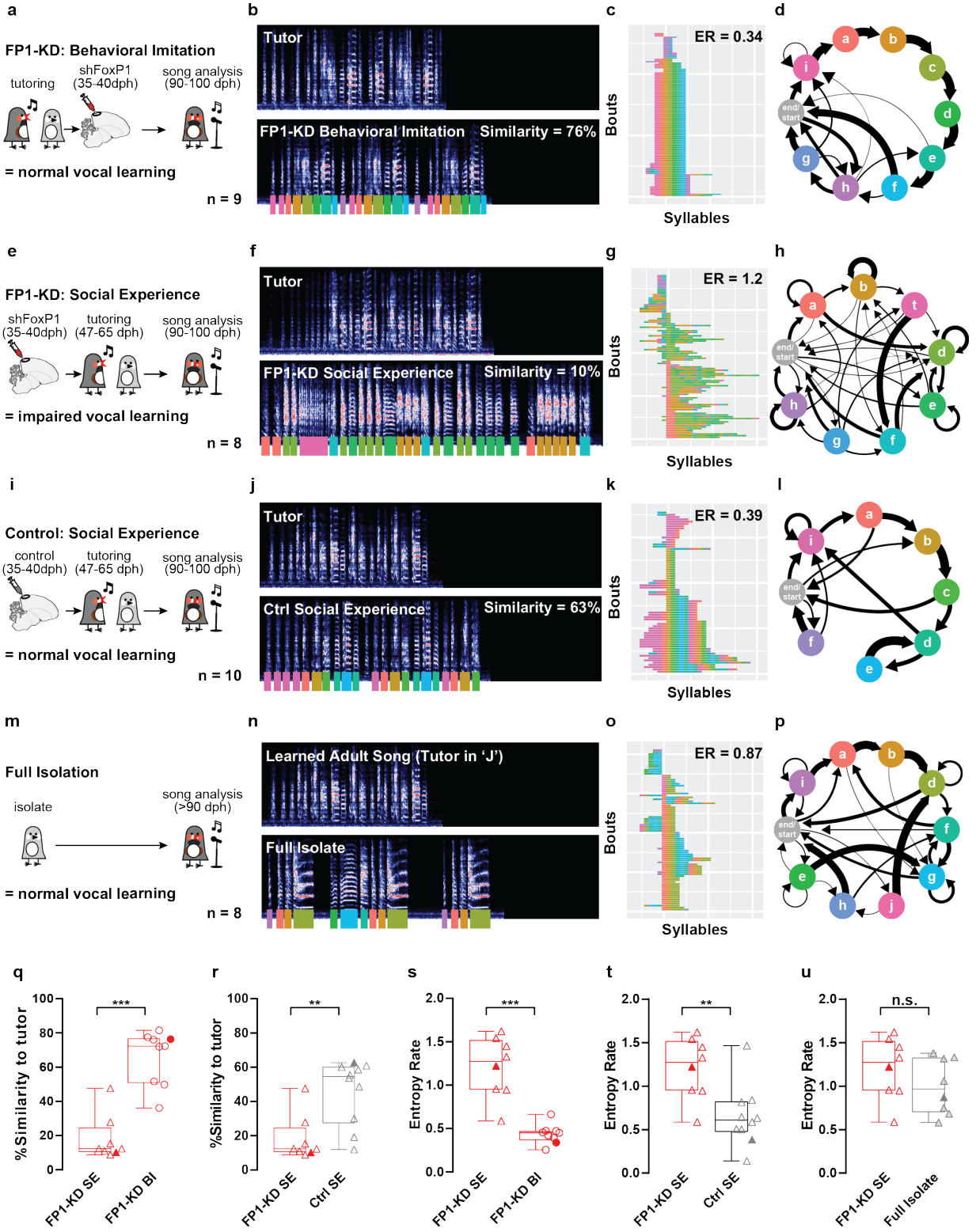


Figure 2. Song learning is impaired by FoxP1 KD

a, e, i & **m,** Timelines illustrating tutoring experience of FP1-KD behavioral imitation (**a**), FP1-KD social experience (**e**), Control social experience (**i**), and full isolate birds (**m**). **b, f, j** & **n,** Representative spectrograms from a single bird belonging to each experimental group and their tutor. Colored underlines reflect syllable labels used for subsequent syntax visualizations and analysis. Similarity refers to % similarity to tutor, also represented in **q** & **r**. The full isolate bird (**n**) had no tutor, and therefore has no similarity to tutor score. It is contrasted with the song of an adult zebra

finch with typical tutor exposure. **c, g, k** & **o,** Syntax raster plot illustrating the syntax stereotypy of an example bird for each condition. This is the same bird used for the spectrogram and syntax diagram. Each row reflects a single song bout, and each colored block reflects the syllable sung at that position in the bout. Rows are sorted according to syllable order. ER refers to entropy rate, also represented in **s-u**. **d, h, l** & **p,** Diagrams reflecting syllable transitions produced by an example bird for each condition. Line thickness is proportional to the transition probability from the originating syllable to the following. Transitions with a probability of less than 4% are omitted for clarity. **q.** FP1-KD social experience birds (n=8) have significantly lower song similarity to tutor than FP1-KD behavioral imitation birds (n=9, FP1-KD SE: 12.39 vs FP1-KD BE: 72.32, Mann-Whitney test, $P < 0.001$). Filled points correspond to the example birds shown in **f-h** and **b-d**. **r,** FP1-KD social experience birds (n=8) have significantly lower song similarity to tutor than Control social experience birds (n=10, FP1-KD SE: 12.39 vs Ctrl SE: 54.6, Mann-Whitney test, $P = 0.0031$). Filled points correspond to the example birds shown in **f-h** and **j-l**. **s,** FP1-KD social experience birds (n=8) have significantly higher song syntax entropy rates than FP1-KD behavioral imitation birds (n=9, FP1-KD SE: 1.274 vs FP1-BI: 0.4512, Mann-Whitney test by bird, $P < 0.001$). Filled points correspond to the example birds shown in **f-h** and **b-d**. **t,** FP1-KD social experience birds (n=8) have significantly higher song syntax entropy rates than Control social experience birds (n=10, FP1-KD SE: 1.274 vs Ctrl SE: 0.6113, Mann-Whitney test, $P = 0.0085$). Filled points correspond to the example birds shown in **f-h** and **j-l**. **u,** Song syntax entropy rates do not differ significantly between FP1-KD social imitation (n=8) and full isolate birds (n=8, FP1-KD SE: 1.274 vs Full Isolate: 0.9667, Mann-Whitney test, $P > 0.05$). Filled points correspond to example birds shown in **f-h** and **n-p**. For all box plots, median, 25th and 75th percentile, and minimum and maximum are reported. The single datapoints are overlaid on the side.

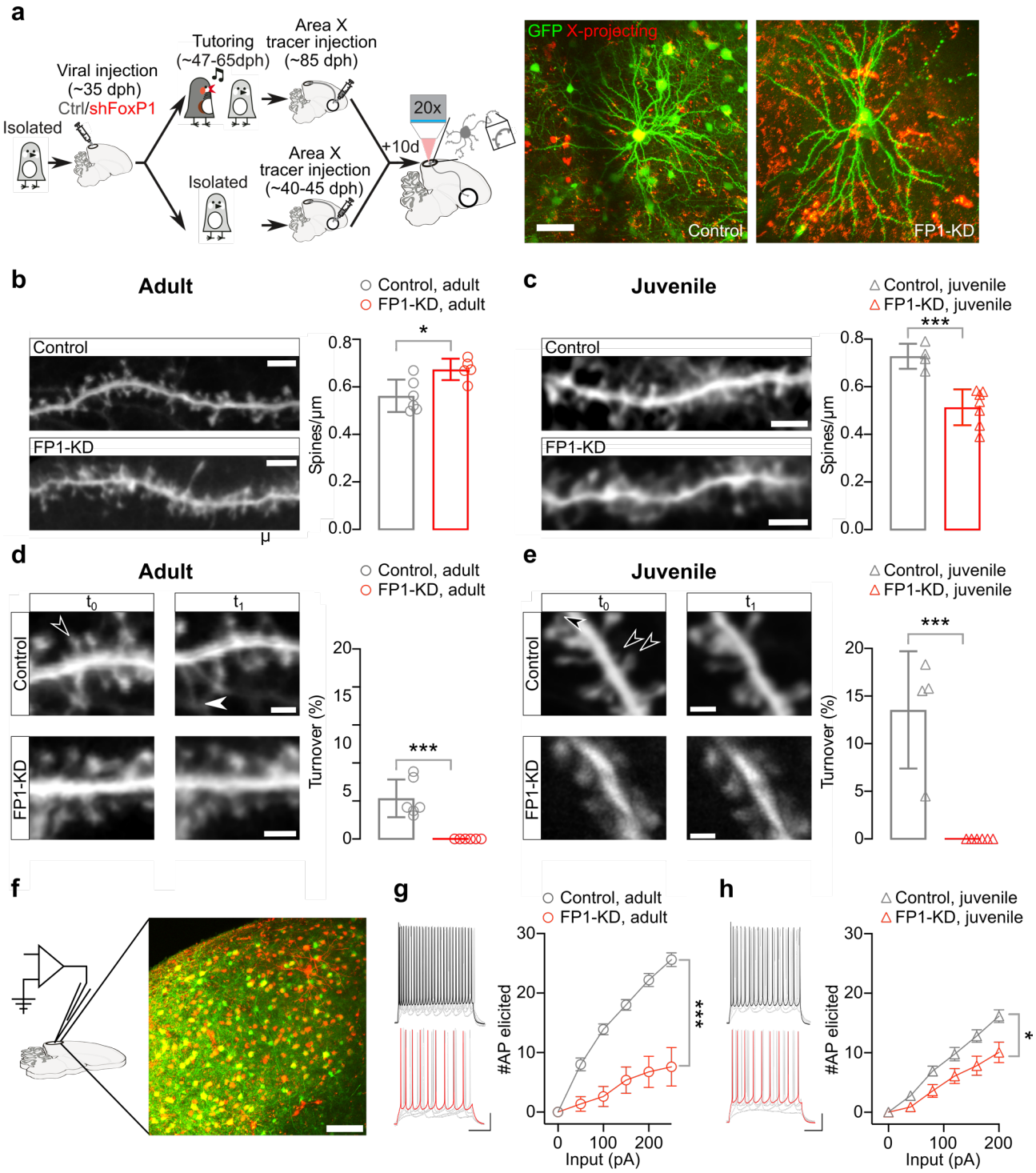


Figure 3. FP1-KD reduces structural plasticity and excitability of HVC$_X$ neurons

a, (left) schematic of the experimental protocol and timeline of the experiments. (right) *In-vivo* two photon images of sample GFP-labelled (green) and retrogradely labelled (red) Control and FP1-KD HVC$_X$ neurons. Scale bar, 50 µm. **b,** (left) Representative *in-vivo* two photon images of GFP-expressing dendrite sections from Control (top) and FP1-KD (bottom) normally reared adult HVC$_X$ neurons. Scale bar, 5 µm. (right) Histograms reporting the average±SEM dendritic spine density (spines/µm) of Control and FP1-KD dendrites from adult HVC$_X$ neurons (Control adult: 0.56±0.03, *n*=821 spines, 6 cells, 2 animals; FP1-KD adult: 0.67±0.02, *n*=668 spines, 6 cells, 5 animals; Student's t-test, *P* = 0.01). **c,** (left) Representative *in-vivo* two photon images of GFP-expressing dendrites sections from Control (top) and FP1-KD (bottom) juvenile isolate HVC$_X$ neurons. Scale bar, 5 µm. (right) Histograms reporting the average±SEM

351 dendritic spine density (spines/µm) of Control and FP1-KD dendrites from juvenile HVC_X neurons (Control juvenile:
352 0.73±0.03, *n*=769 spines, 4 cells, 3 animals; FP1-KD juvenile: 0.51±0.03, *n*=745 spines, 6 cells, 5 animals; Student's t-
353 test, $P < 0.001$). **d,** (left) Representative images of Control and FP1-KD adult dendritic segments from HVC_X neurons
354 taken at two different times (t_0,t_1 across a 4-h imaging interval). Filled arrowhead indicates a representative gained
355 synapse. Empty arrowhead indicates a representative lost synapse. Scale bars, 2 µm. (right) Histograms representing
356 the average±SEM percent dendritic spine turnover (acquired + lost spines/total spines counted) from Control and
357 FP1-KD adults (Control adult: 4.3±0.8%, *n*=1126 spines, 6 cells, 2 animals; FP1-KD adult: 0.0±0.0%, *n*=1148 spines, 6
358 cells, 5 animals; Student's t-test, $P < 0.001$). **e,** (left) Representative images of Control and FP1-KD juvenile dendritic
359 segments from HVC_X neurons, taken at two different times (t_0,t_1 across a 2-h imaging interval). Empty arrowhead
360 indicates a representative lost synapse. Scale bars, 2 µm. (right) Histograms representing the average±SEM percent
361 dendritic spine turnover (acquired + lost spines/total spines counted) from Control and FP1-KD juveniles (Control
362 juvenile: 13.6±3.1%, *n*=650 spines, 4 cells, 3 animals; FP1-KD juvenile: 0.0±0.0%, *n*=735 spines, 6 cells, 5 animals;
363 Student's t-test, $P < 0.001$). **f,** Schematic of an *ex-vivo* slice and patch-clamp recording setup with high-resolution
364 image of HVC in a brain slice used for electrophysiology. Scale bar, 50 µm. **g,** Example traces (left, scalebars 20 mV,
365 100ms) and plot (right) reporting the number of action potentials elicited by somatic current injections in HVC_X
366 neurons from Control and FP1-KD adult brain slices. FoxP1 knockdown decreased the intrinsic excitability of HVC_X
367 neurons in adults (Controls, *n*= 10, 5 animals; FP1-KD *n*= 8, 3 animals; Two-Way Anova, Interaction $F_{10,160} = 30.87$,
368 Treatment $F_{1,16} = 34.56$, $P < 0.001$). **h,** Example traces (left, scalebars 20 mV, 100ms) and plot (right) reporting the
369 number of action potentials elicited by somatic current injections in HVC_X neurons from Control and FP1-KD juvenile
370 brain slices. FoxP1 knockdown decreased the intrinsic excitability of HVC_X neurons in isolate juveniles (Controls, *n*=
371 12, 3 animals; FP1-KD *n*= 6, 2 animals; Two-Way Anova, Interaction $F_{10,200} = 7.053$, Treatment $F_{1,20} = 7.627$, $P = 0.01$).
372 All data are reported as average±SEM. In the histograms, the single datapoints are overlaid on the side of each bar.

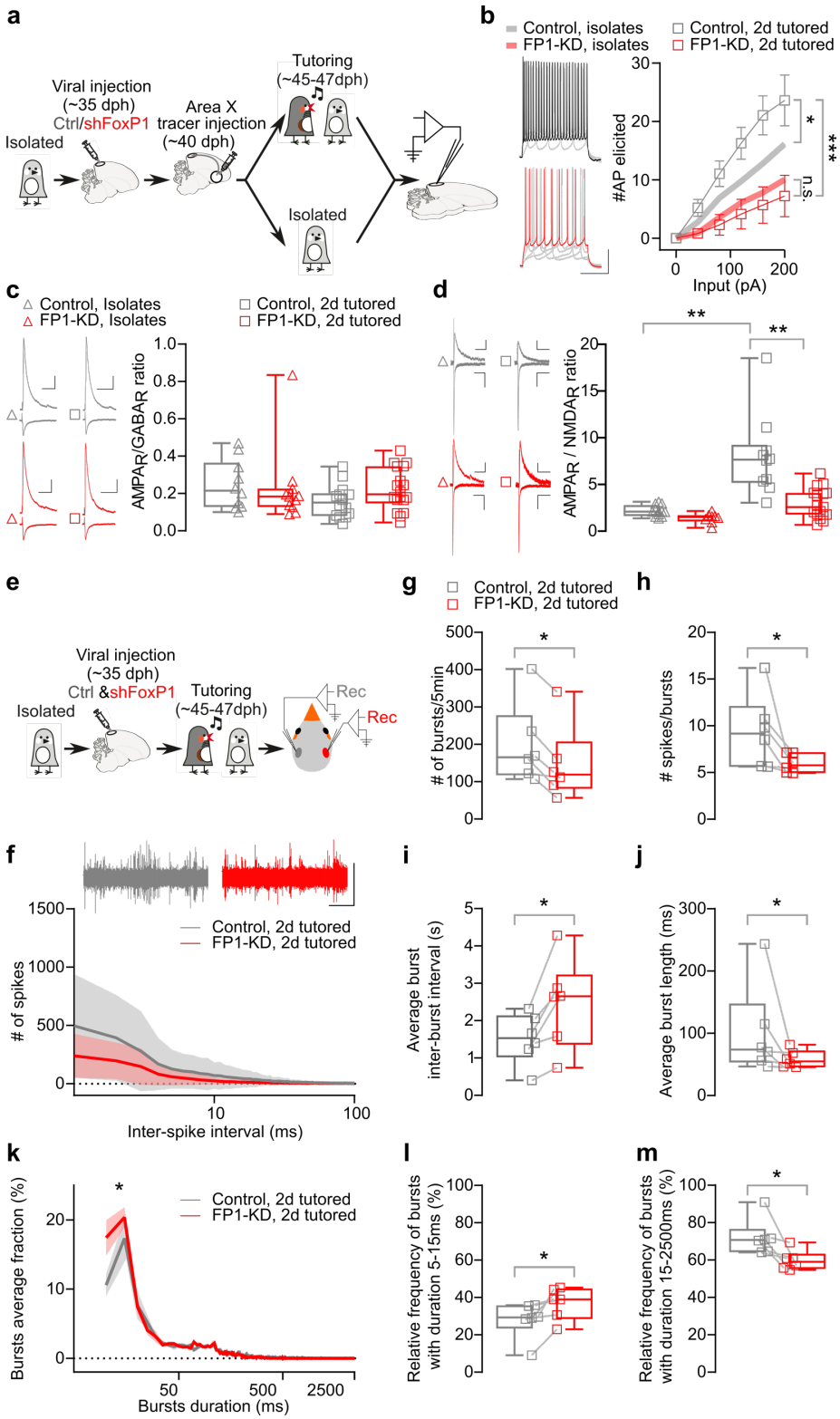


Figure 4. FP1-KD prevents experience-dependent synaptic strength modifications and reduces network-level bursting activity organization

a, Experimental timeline. **b,** Example traces (left) and plot (right) describing the number of spikes elicited by somatic depolarization steps in 2-day tutored Control and FP1-KD HVC_X neurons. Trendlines relative to the same experiments, but conducted in isolates, are reported here from Fig 3h. HVC_X neurons were more excitable in control birds subjected to the 2-day tutoring regime compared to isolates (Two-Way Anova, Interaction $F_{10,190} = 4.598$, Treatment $F_{1,19} = 5.376$, $P = 0.03$). This difference is not present in FP1-KD birds (Two-Way Anova, Interaction $F_{10,180} = 0.4578$, $P >$

0.05). Among 2 day-tutored birds, FoxP1 knockdown resulted in a decreased intrinsic excitability of HVC_X neurons (Controls, *n*= 9, 4 animals; FP1-KD *n*= 10, 3 animals; Two-Way Anova, Interaction $F_{10,170}$ = 9.308, Treatment $F_{1,17}$ = 10.73, *P* = 0.005). Data are reported as average±SEM. **c,** Representative traces (left), box and scatter plots (right) of evoked AMPAR and GABAR mediated currents recorded at -60 mV or 10 mV respectively (scale bars:100 pA; 100 ms) in Control or FP1-KD HVC_X neurons from isolate and 2-day tutored birds. Box plots represent the AMPAR/GABAR currents amplitude ratios from isolates (open triangles) and 2-day tutored birds (open squares). FoxP1 knockdown has no significant effect on the AMPAR/GABAR ratios for both bird groups (Isolates Control: 0.25±0.04, *n*=10, 5 animals; isolated FP1-KD: 0.23±0.06, *n*=11, 5 animals; 2d-tutored Control: 0.16±0.02, *n*=14, 5 animals; 2d-tutored FP1-KD: 0.22±0.03, *n*=18, 6 animals; Kruskal-Wallis test, *P* > 0.05). **d,** Representative traces (left), box and scatter plot (right) of evoked AMPAR and NMDAR mediated currents recorded at -70 mV or +40 mV respectively (scale bars:50 pA; 100 ms) in Control or FP1-KD HVC_X neurons in isolates and 2-day tutored birds. Box plots represent the AMPAR/ NMDAR current ratios recorded from isolates (open triangles) and 2-day tutored birds (open squares). (isolate Control: 2.2±0.2, *n*=12, 6 animals; isolate FP1-KD: 1.4±0.2, *n*=8, 5 animals; 2d-tutored Control: 8.0±1.4, *n*=10, 6 animals; 2d-tutored FP1-KD: 2.9±0.4, *n*=17, 6 animals; Kruskal-Wallis test, *P* < 0.001; Dunn's multiple comparisons, isolate Control vs 2d-tutored Controls, *P* = 0.002, isolate FP1-KD vs 2d-tutored Controls, *P* < 0.001, 2d-tutored Controls vs 2d-tutored FP1-KD, *P* = 0.008). **e,** Schematic representing the experimental timeline. Birds received AAV injections in HVC to knock-down FoxP1 in one hemisphere, and control virus in the other hemisphere (pseudo-randomized). The birds were kept in isolation from male tutors, and ~10 days later they received 48h tutoring, after which HVC extracellular activity was recorded in both hemispheres (*n* = 6 birds, 3-5 recordings/hemisphere). **f,** (top) Sample traces (scalebar 0.5V, 1s) and (bottom) average inter-spike interval distribution (bin 1ms, 1-100ms, logarithmic scale, 300s/recording) (Control hemispheres vs FP1-KD hemispheres, Two-Way Anova, Interaction $F_{99,990}$ = 1.222, *P* > 0.05). Data are reported as average (thick line)±SEM (semitransparent contour) **g,** Box and scatter plot of the total number of bursts (Control hemispheres: 199.4 ± 44.4 vs FP1-KD hemispheres: 148.0 ± 41.2, Wilcoxon matched-pairs signed rank test, *P* = 0.03). **h,** Box and scatter plot of the average number of spikes in a burst (Control hemispheres: 9.4 ± 1.6 vs FP1-KD hemispheres: 6.0 ± 0.4, Wilcoxon matched-pairs signed rank test, *P* = 0.03). **i,** Box and scatter plot of the average inter-burst interval (Control hemispheres: 1.5 ± 0.3 vs FP1-KD hemispheres: 2.4 ± 0.5, Wilcoxon matched-pairs signed rank test, *P* = 0.03). **j,** Box and scatter plot of the average burst length (Control hemispheres: 101.6 ± 30 vs FP1-KD hemispheres: 58.6 ± 5.8, Wilcoxon matched-pairs signed rank test, *P* = 0.03). **k,** Plot representing the relative distribution of burst duration, normalized for each recording (5ms duration bins, 5-2500ms, logarithmic scale; Control hemispheres vs FP1-KD hemispheres, Two-Way Anova, interaction $F_{498,4980}$ = 1.972, *P* < 0.001, Control vs FP1-KD $F_{1,10}$ = 5.570, *P* = 0.04). Data are reported as average (thick line)±SEM (semitransparent contour). **l,** Box and scatter plot of the average relative prevalence of bursts with durations between 5 and 15ms (Control hemispheres: 27.9 ± 4.0 vs FP1-KD hemispheres: 36.8 ± 3.5, Wilcoxon matched-pairs signed rank test, *P* = 0.03). **m,** Box and scatter plot of the average relative prevalence of bursts with durations between 15 and 2500ms (Control hemispheres: 72.1 ± 4.0 vs FP1-KD hemispheres: 59.8 ± 2.2, Wilcoxon matched-pairs signed rank test, *P* = 0.03). For all box plots, median, 25th and 75th percentile, and minimum and maximum are reported. The single datapoints are overlaid on the side.

Supplementary Materials: Autism-linked gene FoxP1 selectively regulates the cultural transmission of learned vocalizations

Materials and Methods

Figures ED1 to ED5

Materials and Methods

Animals

Experiments described in this study were conducted using juvenile and adult male zebra finches (30-110 days post hatch (dph)). We raised juvenile male zebra finches either in isolation from an adult song model (isolates), with normal access to an adult song model (non-isolates), or isolates that were exposed to 2 days of tutoring by an adult song model (2-day tutored). All procedures were performed in accordance with protocols approved by Animal Care and Use Committee at UT Southwestern Medical Center.

Viral Vectors

The following adeno-associated viral vectors were used in these experiments: rAAV2/9/ds-CBh-GFP (The University of North Carolina at Chapel Hill Gene Therapy Center Vector Core), and pscAAV-GFP-shFoxp1 (IDDRC Neuroconnectivity Core at Baylor College of Medicine). All viral vectors were aliquoted and stored at -80°C until use.

Constructs

pscAAV-GFP-shFoxP1 was generated by PCR-amplification of U6-shFoxp1 from pLKO.1 (TRCN0000072005, Broad Institute) while adding NotI and BamHI sites, then ligating into pscAAV9-CBh-GFP (Xiao et al., 2018) digested with these enzymes. The PCR program used was: 98°C 2 min, 35 x (98°C 10 s, 55°C 15 s, 72°C 5 s), 72°C 7 min. The primers used were forward (5'-3'): ATAAGAATGCGGCCGCTTTCCCATGATTCCTTC, and reverse (3'-5'): CGCGGATCCAAAAAGCTAACACT AAACG. The scrambled control (shCtrl) was designed by uploading shFoxp1 into https://www.genscript.com/ssl-bin/app/scramble and selecting the chicken mRNA pool. Ensembl BLAST found only 3 hits of 15/21 bases in the zebra finch genome. pCMV- FoxP1^{WT} was generated by subcloning zebra-finch specific FoxP1 cDNA from pGEMTeasy (Haesler et al., 2004) into pCMV-Tag4a using EcoRI.

Stereotaxic Surgery

All surgical procedures were performed under aseptic conditions. Birds were anaesthetized using isoflurane inhalation (0.8-1.5%) and placed in a stereotaxic surgical apparatus. The centers of HVC and RA were identified with electrophysiological recordings, and Area X and Av were identified using stereotaxic coordinates.

35 Viral injections to HVC were performed using previously described procedures (Roberts et al., 2012; Roberts et al.,
36 2017). Briefly, adeno-associated viral vectors (pscAAV-GFP-shFoxP1 or rAAV9/ds-CBh-GFP) were injected into HVC (50
37 nl per injection and ~60 injections, for a total of ~3 µl) at ~35 dph, and the transgenes were allowed to express for a
38 minimum of 10 days. We also injected 500-950nl of differently conjugated tracers (Dextran, AlexaFluor 488 or 594,
39 10,000MW, Invitrogen) bilaterally into birds' Area X, Av, and RA respectively. Tracer injections were performed using
40 previously described procedures (Roberts et al., 2012; Roberts et al., 2017; Xiao et al., 2018) at the following
41 approximate stereotaxic coordinates relative to interaural zero and the brain surface (rostral-caudal, medial-lateral,
42 dorsal-ventral, in mm, head angle): HVC (0, ±2.4, 0.1-0.6, with 30° head angle); RA (-1.0, ±2.4, 1.7-2.4, with 30° head
43 angle); Area X (5.1, ±1.6, 3.3, with 45° head angle), and Av (1.75, ±2.0, 1.0, with 45° head angle).
44
45 **Tutoring Conditions**
46 Social Experience (SE)**:** juvenile male zebra finches raised in isolation from an adult song model were injected with
47 viruses into HVC at 35-40 dph. After optimal viral expression, these juveniles were then housed with a song tutor
48 between days ~47 and 65 dph. All birds were separated from their tutors at 65 dph and raised to adulthood.
49 Behavioral Imitation (BI): juvenile male zebra finches were reared with a song tutor and injected with viruses into HVC
50 at 35-40 dph. Normally reared juveniles were raised with an adult song model before and after viral injections,
51 separated from their tutors at 60 days of age, and raised to adulthood.
52 2-day Tutored: juvenile male zebra finches that were raised as isolates were injected with viruses into HVC at ~35 dph.
53 After allowing time for viral expression, a song tutor was placed into the isolate's cage for two days of tutoring. Birds
54 were then separated from their tutors and raised to adulthood.
55 Isolates: juvenile male zebra finches raised in isolation from an adult song model were injected with viruses into HVC
56 at ~ 35 dph, and raised to adulthood without exposure to an adult song model.
57
58 ***In vivo* two-photon imaging**
59 We conducted longitudinal dendritic spine imaging in male juvenile zebra finches raised in isolation from an adult song
60 model (isolates) aged 45–55d or adult zebra finches 90-100d (table details). Viruses (pscAAV-GFP-shFoxP1 or rAAV9/ds-
61 CBh-GFP) were allowed to express for a minimum of 14 days before a cranial window over HVC was made. Birds were
62 anaesthetized by isoflurane inhalation (2%) and positioned in a stereotaxic apparatus for cranial windowing. The scalp
63 overlying HVC was removed and the scalp margins were sealed to the surface of the skull using Vetbond (*n*-butyl
64 cyanoacrylate). Bilateral craniotomies (~1–1.5 mm^2) were made in the skull overlying HVC. The dura mater was
65 excised, leaving the pia mater, the 60–150-µm-thick layer of neural tissue, and the lateral telencephalic ventricle
66 overlying HVC intact. A custom-cut coverslip (no.-1 thickness) was placed directly on the pial surface, then sealed to
67 the skull with dental acrylic. A head post was also affixed to the skull with dental acrylic as per previously described

68 procedures (Daliparthi et al., 2019). Birds were placed onto a custom stage under an Ultima IV Bruker laser scanning
69 microscope running Prairie View Software. Only HVC neurons that expressed both the retrograde tracer Dextran Alexa
70 Fluor 594 from Area X (HVC_X neurons) and GFP from viruses were chosen for spine imaging. Dendritic segments of
71 these neurons were imaged at high resolution during the bird's subjective night-time (1024 × 1024 pixels, 76 × 76 μm^2,
72 3.2 µs per pixel, averaging two samples per pixel with 1-µm *z* steps, focused through a ×20, NA 1.0 Zeiss IR-Achroplan
73 immersion objective). Birds were then returned to a darkened holding cage and allowed to sleep, and were re-imaged
74 2-4 hrs later.
75
76 **Spine Image Analysis**
77 Dendritic spine images were analyzed as reported previously (Roberts et al., 2010). Briefly, three-dimensional image
78 stacks were auto-aligned and smoothed using a Gaussian filter (ImageJ; http://rsbweb.nih.gov/ij/) and the same
79 dendritic segment, imaged twice with a 2 hr interval, was selected. Images exhibiting changes in fluorescence or
80 rotational artefacts were excluded from further analysis. Sets of selected three-dimensional image stacks were scored
81 by 4 researchers blind to the experimental condition. To assess spine growth and retraction, we compared individual
82 dendritic spines across 2-4 hr time intervals and calculated spine stability (N_{stable}/N_{total}), spine elimination (N_{lost}/N_{total}),
83 spine addition (N_{gained}/N_{total}) and spine turnover ($(N_{gained} + N_{lost})/2N_{total}$), where N_{stable} is the number of spines that were
84 stable over the time interval, N_{lost} is the number of spines lost over the time interval, N_{gained} is the number of spines
85 gained over the time interval and N_{total} is the total number of spines from the first imaging time point. Changes in spine
86 density (N_{total} divided by dendritic length in µm) were measured from the same dendritic segments used to assess spine
87 turnover.
88
89 ***In-vivo* extracellular recordings**
90 Isolated, juvenile (dph 30-35) birds received a unilateral (pseudo-randomized) injection of pscAAV-GFP-shFoxp1 in HVC.
91 The other hemisphere received the control virus (rAAV9/ds-CBh-GFP). After 10-12 days, we recorded HVC
92 electrophysiological activity from both hemispheres (3-5 recordings/hemisphere, minimum 5 minutes/recording after
93 an initial 5 minutes to allow the activity to stabilize after the electrode dip). Recordings from both HVCs were sampled
94 in a pseudorandomized order. Signals were acquired at 10KHz, and filtered (high-pass 300Hz, low-pass 20KHz). We
95 used Spike2 to analyze the spikes whose amplitude reached a threshold of 0.3V (determined based on the average
96 noise level among all the recordings). A Spike2 script was used to analyze the characteristics of spikes in bursting
97 patterns. Bursts were defined as a minimum of 2 spikes separated by 10ms or less, and the burst epoch was considered
98 terminated if no spike was detected for 100ms after the last spike of the burst. Values for each recording were averaged
99 per hemisphere, and statistically compared pairwise where indicated to reveal differences. For the comparison of
100 bursts length, the data from each hemisphere was normalized to produce a frequency distribution of the burst lengths

101 for that hemisphere (bin size 1ms). The data were then compared across treatments, and further subdivided into two
102 burst duration categories: 5-15ms, and 15-2500ms.
103
104 ***Ex vivo* physiology**
105 **Slice preparation.** All extracellular solutions were adjusted to 310 mOsm, pH 7.3–7.4, and aerated with a 95% O_2/5%
106 CO_2 mix. Zebra finches were first deeply anesthetized with isoflurane. Once the bird was no longer responsive to a toe
107 pinch, it was quickly decapitated. The brain was removed from the skull and submerged in cold (1– 4°C) oxygenated
108 dissection buffer. Acute sagittal 230-250 µm brain slices were cut in dissection buffer at 4°C containing (in mM): 225
109 sucrose, 3 KCl, 1.25 NaH_2PO_4, 26 $NaHCO_3$, 10 D-(+)-glucose, 2 $MgSO_4$, .5 $CaCl_2$, 2 Kynurenic acid. Individual slices were
110 incubated in a custom-made holding chamber saturated with 95% O_2/5% CO_2 at 34 °C for 20 minutes, and then kept at
111 30 °C for a minimum of 45 minutes in artificial cerebrospinal fluid (aCSF), containing (in mM): 126 NaCl, 3 KCl, 1.25
112 NaH_2PO_4, 26 $NaHCO_3$, 10 D-(+)-glucose, 2 $MgSO_4$, 2 $CaCl_2$.
113
114 **Slice electrophysiological recording.** Slices were constantly perfused in a submersion chamber with 32°C oxygenated
115 normal aCSF. Patch pipettes were pulled to a final resistance of 3-5 MΩ from filamented borosilicate glass on a Sutter
116 P-1000 horizontal puller. HVC_X cell bodies double-labelled with GFP and Dextran AlexaFluor594 were visualized by
117 epifluorescence imaging using a water immersion objective (×40, 0.8 numerical aperture) on an upright Olympus BX51
118 WI microscope, with video-assisted infrared CCD camera (Q-Imaging Rolera). Data were low-pass filtered at 10 kHz and
119 acquired at 2 kHz with an Axon MultiClamp 700B amplifier and an Axon Digidata 1550B Data Acquisition system under
120 the control of Clampex 10.6 (Molecular Devices).
121 For **voltage clamp** whole-cell recordings of HVC_X projecting neurons, the internal solution contained (in mM): 120 Cs-
122 gluconate, 10 HEPES, 5 tetraethylammonium-Cl, 2.8 NaCl, 0.6 EGTA, 4 MgATP, 0.3 NaGTP, 5 BAPPTA, 7 QX314 chloride
123 (adjusted to pH 7.3-7.4 with CsOH, 297 mOsm). For **current clamp** recordings, the internal solution contained: 116 K-
124 gluconate, 20 HEPES, 6 KCl, 2 NaCl, 0.5 EGTA, 4 MgATP, 0.3 NaGTP, 10 Na-phosphocreatine (adjusted to pH 7.3-7.4
125 with KOH, 299 mOsm).
126
127 Electrically evoked synaptic currents were measured by delivering one electric stimulus (1 ms, 10–30 µA) every 12 s,
128 with an isolation unit, through a glass stimulation monopolar electrode filled with aCSF, placed at about 50–100 µm
129 from the recorded HVC_X neuron. Synaptic responses were monitored at different stimulation intensities prior to
130 baseline recording. "Normal" stimulation was defined as a stimulation reliably evoking a synaptic current in the range
131 100 pA to 1 nA.
132
133 Excitatory and inhibitory **synaptic** currents were recorded in the whole-cell voltage clamp mode with the Cs-based
134 patch pipette solution in order to measure the E/I ratio. Only recordings with series resistance below 20 MΩ were

135 included. Excitatory post synaptic currents (EPSCs) and inhibitory post synaptic currents (IPSCs) were recorded at the
136 reversal potential for IPSCs (+10 mV) and EPSCs (-60 mV) in the presence of the NMDA receptor antagonist APV (100
137 µM), respectively. We also used the Cs-based pipette solution to measure the ratio between N-methyl-D-aspartate-
138 receptor-mediated currents (I_{NMDA}) and alpha-amino-3-hydroxy-5-methyl-4-isoxazolepropionic-acid-receptor-
139 mediated currents (I_{AMPA}) in HVC_X neurons. We added the $GABA_A$ receptor antagonist Picrotoxin (10 µM) to the aCSF
140 for these recordings. I_{AMPA} was recorded at a holding potential of $V_h = -60$ mV and measured at their peak. I_{NMDA} were
141 recorded in the same cell at $V_h = +40$ mV. I_{NMDA} amplitude was calculated as the mean between 95 and 105 ms after
142 the electric stimulation artifact, to minimize the possible contamination by I_{AMPA}.
143 Access resistance (10–20 MΩ) was monitored throughout the experiment.
144
145 **Intrinsic excitability**: Neuronal intrinsic excitability was examined with the potassium gluconate-based pipette
146 solution. After whole-cell current clamp mode was achieved, cells were maintained at -80 mV. Input resistances were
147 monitored by injecting a 150-ms hyperpolarizing current (40pA) to generate a small membrane potential
148 hyperpolarization from the resting membrane potentials. Firing rate represents the average value measured from 1 to
149 3 cycles (700ms duration at 0.1 Hz, –200 to +250pA range with 50pA or 40pA steps increment, every 12 s).
150
151 **Western Blotting in HVC**
152 Samples for western blotting were collected over 2 days between 9am and 12pm from 90-92 dph birds which received
153 either a control (rAAV9/ds-CBh-GFP) or FoxP1 knock-down (pscAAV-GFP-shFoxP1) injection into HVC. Briefly, birds
154 were anesthetized with isoflurane, and their brains were dissected. 230 µm thick sagittal sections were collected in
155 aCSF using a Vibratome VT1200. After confirming the GFP fluorescence under a microscope, HVC was dissected out
156 from these sections. PBS was pipetted out, and the tissue was lysed in RIPA buffer containing protease and phosphatase
157 inhibitors. Protein quantification was done using a Bradford assay, and 50 µg protein/well was used for Western blots.
158 Proteins were run on a 10% SDS-PAGE resolving gel, with 5% stacking gel at 80V until the loading dye front ran off the
159 gel, and were then transferred to an Immuno-Blot PVDF Membrane (Bio-Rad Laboratories) at 250 mA for 2 hr at 4°C.
160 The membrane was dried at room temperature for 1 h, reactivated in methanol, and blocked for 1 hr in 5% milk in TBS.
161 The membrane was cut above 50 kDa, and incubated with primary antibodies in 5% milk in TBS with 0.1% Tween-20
162 (TBS-T) overnight at 4°C. The following day, it was washed in TBS-T, incubated with secondary antibodies in 5% milk in
163 TBS-T for 1 hr at room temperature, washed in TBS-T again, and imaged with TBS on an Odyssey Infrared Imaging
164 System (LI-COR Biosciences). The following antibodies were used: rabbit α- FOXP1 (Spiteri et al., 2007) (1:5000), mouse
165 α-GAPDH (#MAB374, Millipore, 1:10,000), donkey α-rabbit IgG IRDye 800 (#926-32213, LI-COR Biosciences, 1:20,000),
166 donkey α-mouse IgG IRDye 680 (#926-68072, LI-COR Biosciences, 1:20,000). The images were quantified using the
167 Odyssey Imaging System (LI-COR Biosciences).
168

169 **Western Blotting in 293T cells**

170 293T cells were cultured in high-glucose DMEM containing 4 mM L-glutamine, 10% fetal bovine serum (#10437028,
171 Invitrogen), and 1% antibiotic-antimycotic (#15240- 062, Invitrogen) at 37°C under 5% CO_2. Cells were plated in 6-well
172 plates and transfected with 0.1 pmol/well of overexpression construct (pCMV-FoxP1) and 0.4 pmol/well of hairpin
173 constructs (pscAAV-GFP-shFoxP1 or shCtrl) using FuGENE 6 Transfection Reagent (#E2691, Promega) according to the
174 manufacturer's instructions. After 24-48h, cells were washed twice with PBS and collected into tubes by scraping in
175 Lysis Buffer (50 mM Tris pH 8.0, 0.1 mM EDTA, 250 mM NaCl, 50 mM NaF, 0.1 mM $NaVO_4$, 10% glycerol, 0.5% IGEPAL
176 CA-630, 1 mM DTT, 0.7 mM PMSF). Samples were incubated for 10 min at 4°C with rotation, and centrifuged at 15,000
177 rpm for 10 min at 4°C. The supernatant was quantified using a Bradford assay, mixed with loading buffer, incubated at
178 95°C for 5 min, and stored at -20°C. Western blotting was performed as described for HVC tissue, and imaged on an
179 Odyssey Infrared Imaging System (LI-COR Biosciences).

180

181 **Immunohistochemistry**

182 Immunohistochemistry experiments were performed following standard procedures. Briefly, birds were anesthetized
183 with Euthasol (Virbac, TX, USA) and transcardially perfused with PBS, followed by 4% paraformaldehyde in PBS. Free-
184 floating sagittal sections (40 µm) were cut using a cryostat (Leica CM1950, Leica). Sections were first washed in PBS.
185 The tissues were then blocked in 5% normal donkey serum in PBST (0.3% Triton-X in PBS) for 1 hr at RT, incubated with
186 primary antibodies diluted in the blocking buffer (5% donkey serum in PBST), first for 1 hr at RT, and then at 4°C for 48
187 hrs. Slices were then washed with PBS, and incubated with fluorescent secondary antibodies (diluted in blocking buffer)
188 at RT for 2 hours. After a final PBS wash, sections were mounted onto slides with Fluoromount-G (eBioscience, CA,
189 USA). Composite images were acquired and stitched using an LSM 880 laser-scanning confocal microscope (Carl Zeiss,
190 Germany). Primary antibody: mouse anti-FoxP1 (ab32010, Abcam), 1:1000 concentration. Secondary antibody: donkey
191 anti-mouse, conjugated to AlexaFluor405 (ab175658, Abcam). All image analysis was done using ImageJ, and graphs
192 were prepared in GraphPad Prism7.

193

194 **Song analysis**

195 A single day of songs from each FP1-KD behavioural imitation (n=9), FP1-KD social experience (n=8) and control social
196 experience bird (n=10) was selected for analysis when the birds were between 90-100dph. Adult isolate songs were
197 obtained from 8 birds ranging in age from 99-463 dph. Birds were recorded using Sound Analysis Pro 2011
198 (Tchernichovski et al., 2000).

199

200 **Syntax Analysis**

From the single day of song for each bird, a random subset of 30 song files were selected for syllable labeling. Syllables were labeled by hand by an expert based on the song spectrograms using a custom MATLAB program. Syllable labels, onset times, and offset times were exported to R for further analysis. Entropy rate was calculated using the *ccber* package in R (Davis et al., 2017) for each bird based on the syllable labels, including a label reflecting the start or end of a song file. Entropy rate reflects an overall measure of the predictability of syllable transitions as a first order Markov chain, calculated according to the formula:

$$\text{Entropy Rate} = -\sum_{ij} \pi_i P_{ij} \log_2 P_{ij}$$

Where P_{ij} is the probability of transitioning from state *i* to state *j*, and π_i is the stationary distribution of the model for state *i*. A perfectly predictable sequence, where every syllable is always followed by the same syllable will have an entropy score of 0. A maximally entropic syllable sequence, where there is an equal probability that a given syllable transitions to any of the possible syllables (or a file end), will have an entropy rate of $\log_2 K$, where *K* is the number of different states, or in this case syllables plus an additional state representing the end of a file.

Syllable transition probabilities used to determine arrow thickness in Figure 2d, h, l & p were calculated using the *markovchain* package in R (Spedicato & Signorelli, 2014). Each syllable, as well as file boundaries and gaps lasting longer than 100ms were considered 'inter-bout gap' states in the first order Markov chain. Each pairwise transition between states (syllables or inter-bout gaps) is tallied to determine the probability of each of the possible states occurring following a given current state in order to generate a two-dimensional transition probability matrix. Arrow thicknesses in Figure 2d, h, l & p are directly proportional to the probability that the origin syllable would transition to the syllable designated by the arrow.

Syntax raster plots shown in Figure 2 c, g, k & o were created using a custom R code. Syllable labels are first arranged into bouts. Bouts are considered strings of at least two syllables which are not separated by a gap longer than 100ms or a file boundary. All bouts are then arranged according to user-specified primary and secondary alignment syllables. These alignment syllables are chosen to maximize the overall alignment of all bouts in the final raster plot. Each bout is shifted along the x axis such that the first occurrence of the primary alignment syllable occupies the 0 position across all bouts, with the order of syllables maintained within each bout. If a bout does not contain the primary alignment syllable, it is shifted such that the first secondary alignment syllable occupies the 0 position, and the order of syllables within the bout is maintained. Bouts which contain neither the primary nor secondary alignment syllables are plotted above the others at an offset, such that the last syllable in that bout occupies the x position -1. Once aligned, the bouts are ordered along the y axis alphabetically based on the syllable labels following the alignment syllable. The result is a

representation of the sequence of syllables across all labeled bouts from a single bird arranged to maximize the emergence of patterns and dominant sequences.

Similarity Scoring

The lack of stereotyped motif structure in the isolate tutored shFoxP1 birds made it difficult to use standard automated song similarity scoring programs. Instead, similarity between spectrograms was evaluated by 11 human experts. Each was presented with a total of 126 spectrogram pairs, and were instructed to rate the similarity of the two spectrograms on a scale of 1 (not similar) to 10 (very similar), following a 5 comparison training set. The spectrograms used were all generated using Sound Analysis Pro 2011 (Tchernichovski et al., 2000) and all reflected a duration of approximately 2.5 seconds. All spectrograms began at the beginning of a song bout, and with the exception of songs which lacked a typical motif structure, included at least one full song motif. The order of comparisons within the test set was randomized for each participant, and included 4 comparisons between tutee and tutor for each tutee in the FP1-KD behavioural imitation, FP1-KD social experience and control social experience groups (4 x 27 birds), 8 comparisons between a tutor and itself to ensure that scorers were using the full range of the scale, and 10 duplicated tutor-tutee pairs to ensure that the scorers were internally consistent.

No individual scorer differed from the mean score by more than an average of 2 standard deviations, none differed by more than an average of 2 points on the duplicated comparisons, and all but one made use of the full 10 point scale (they gave at least one comparison a score of 0 and at least one a score of 10). One scorer never gave any comparison a perfect 10/10 score, so their scores were rescaled such that they spanned the full 1-10 range.

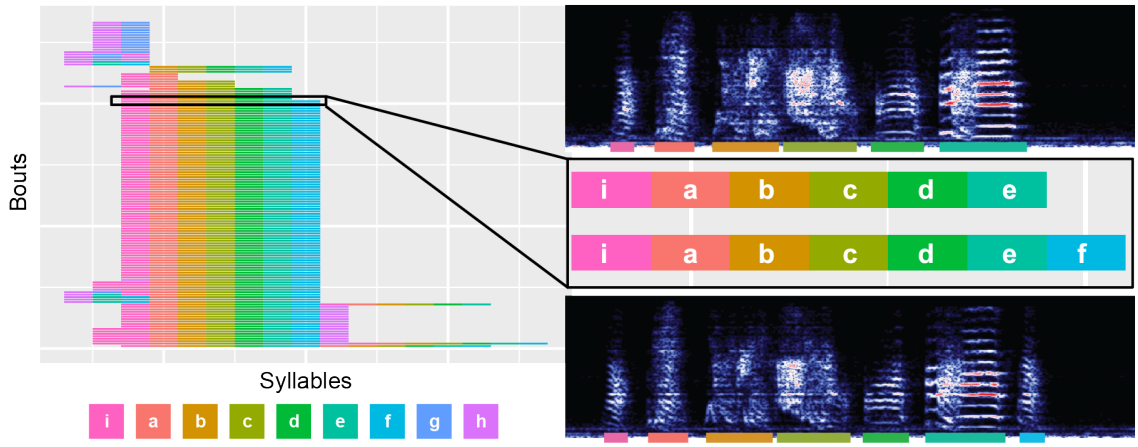
A % similarity score was calculated for each bird by taking the mean similarity rating for all 4 comparisons to the tutor across all scorers.

The full scoring set is available at https://forms.gle/9TDu1fwGGYXWKhgB6.

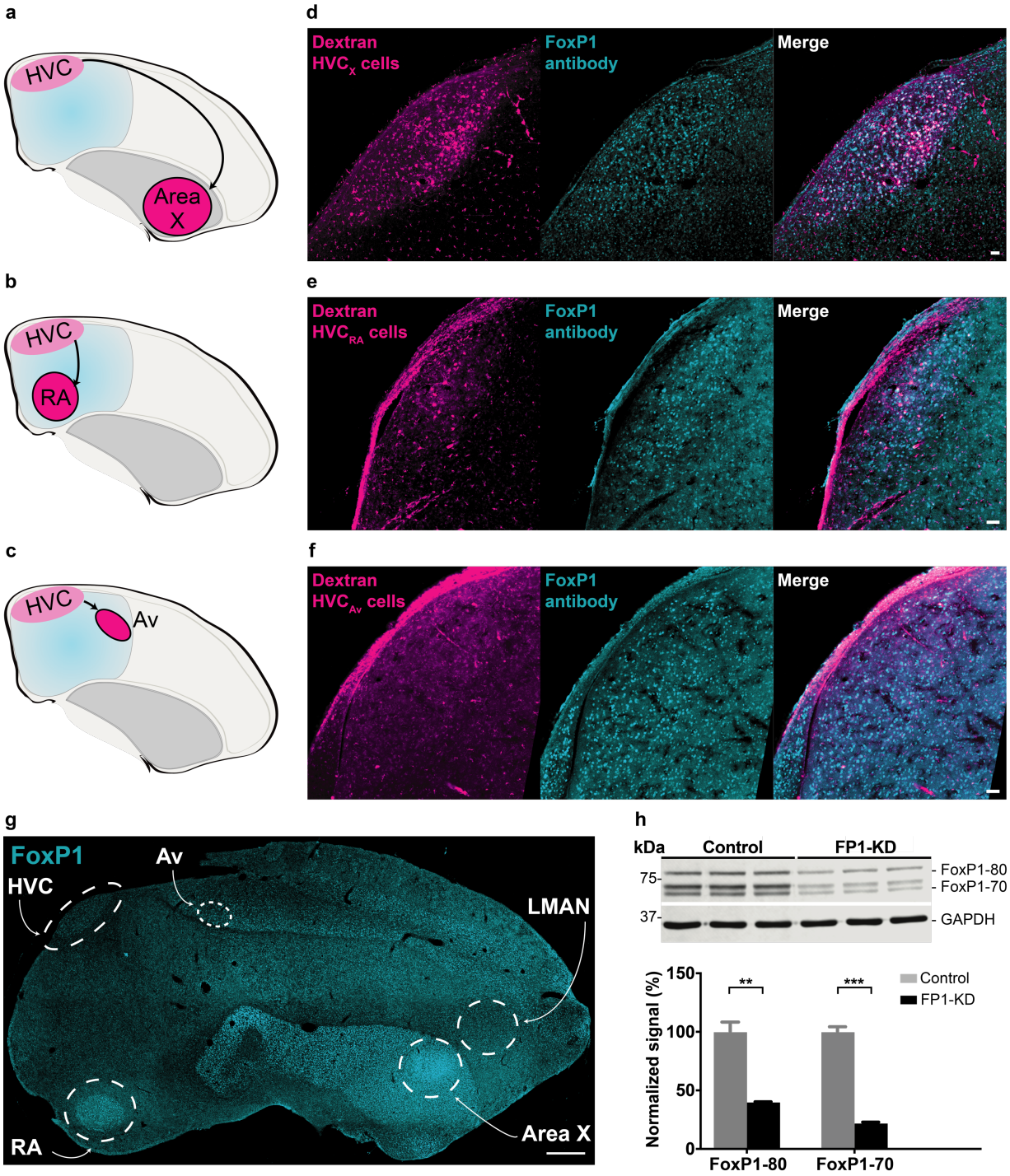
Statistical Analysis

All data were tested for normality using the Shapiro-Wilk Test. Consequently, parametrical or non-parametric statistical tests were used as appropriate: t-test (unpaired or paired, as indicated), Mann-Whitney (unpaired) or Wilcoxon matched-pairs signed rank test (paired) tests were used where appropriate. One-way Anova or Kruskal-Wallis test was performed when comparisons were made across more than two conditions. Two-way Anova (post-hoc Sidak) was used to test differences between two or more groups across different conditions. Statistical significance refers to $*P < 0.05$, $** P < 0.01$, $*** P < 0.001$. Statistical details for all experiments are included in the corresponding figure legends.

Extended Data Figures 1-5



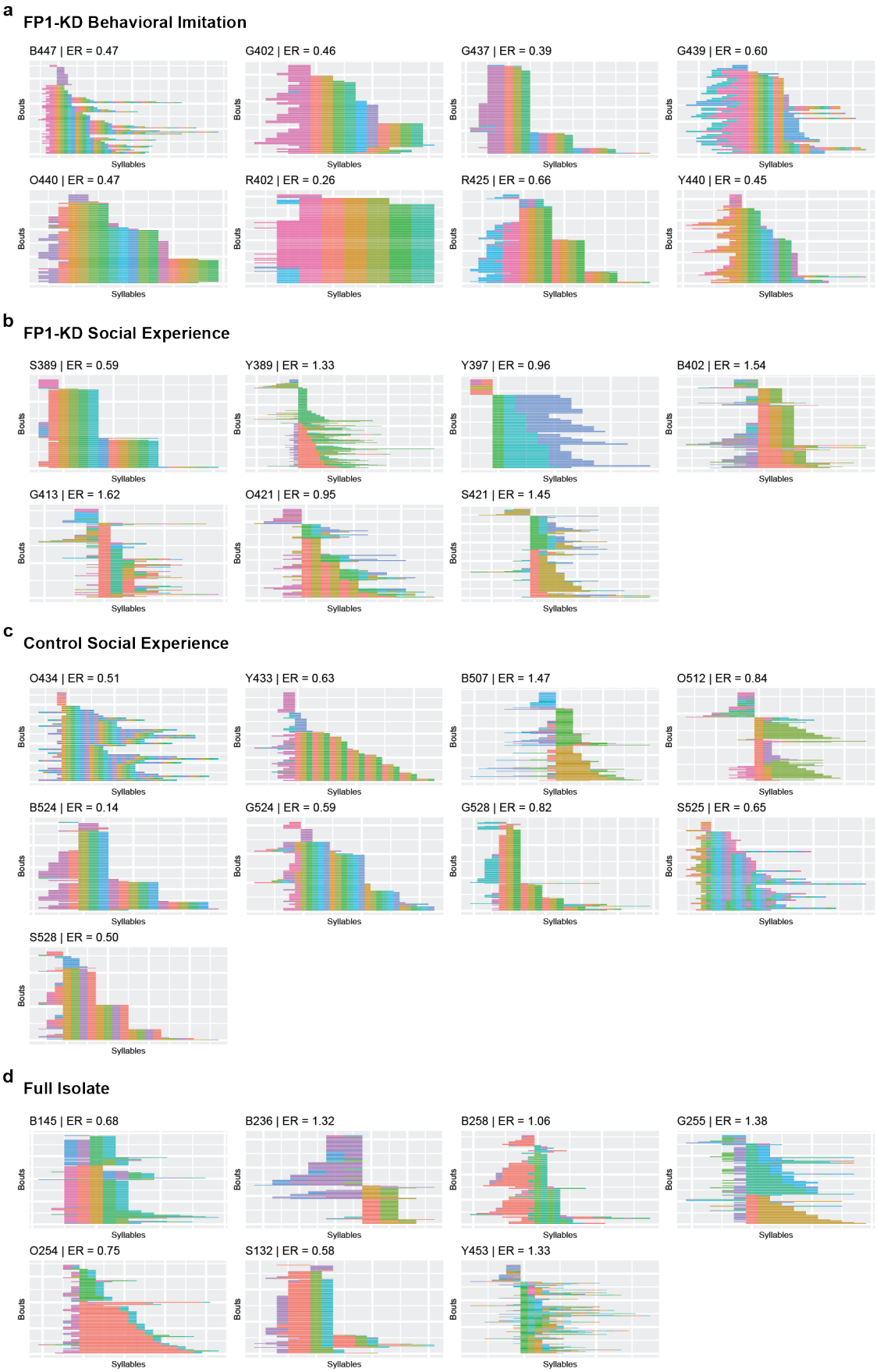
Extended Data Figure 1: Interpreting a syntax raster plot.
Schematic, featuring song data from the non-isolate shFoxP1 bird featured in figure 2**b-d**, which illustrates the relationship between raw spectrograms and syntax raster plot visualizations used in figures 1 and 2. The syntax raster plot represents 30 randomly selected song files from a single bird. These files are segmented into syllables by amplitude thresholding, then syllables are labelled by hand. The song files are then separated into song bouts, defined as more than two syllables produced consecutively with a gap shorter than 100ms between them. These bouts are then sorted according to the syllables they contain so that identical bouts appear together. Each song bout is then represented as a row of blocks whose colour reflects the syllable label at each position within the bout. Birds with very stereotyped syllable order, such as the one in this figure, will have raster plots with a consistent vertical stripe pattern, indicating that the same syllable is consistently produced in the same position.



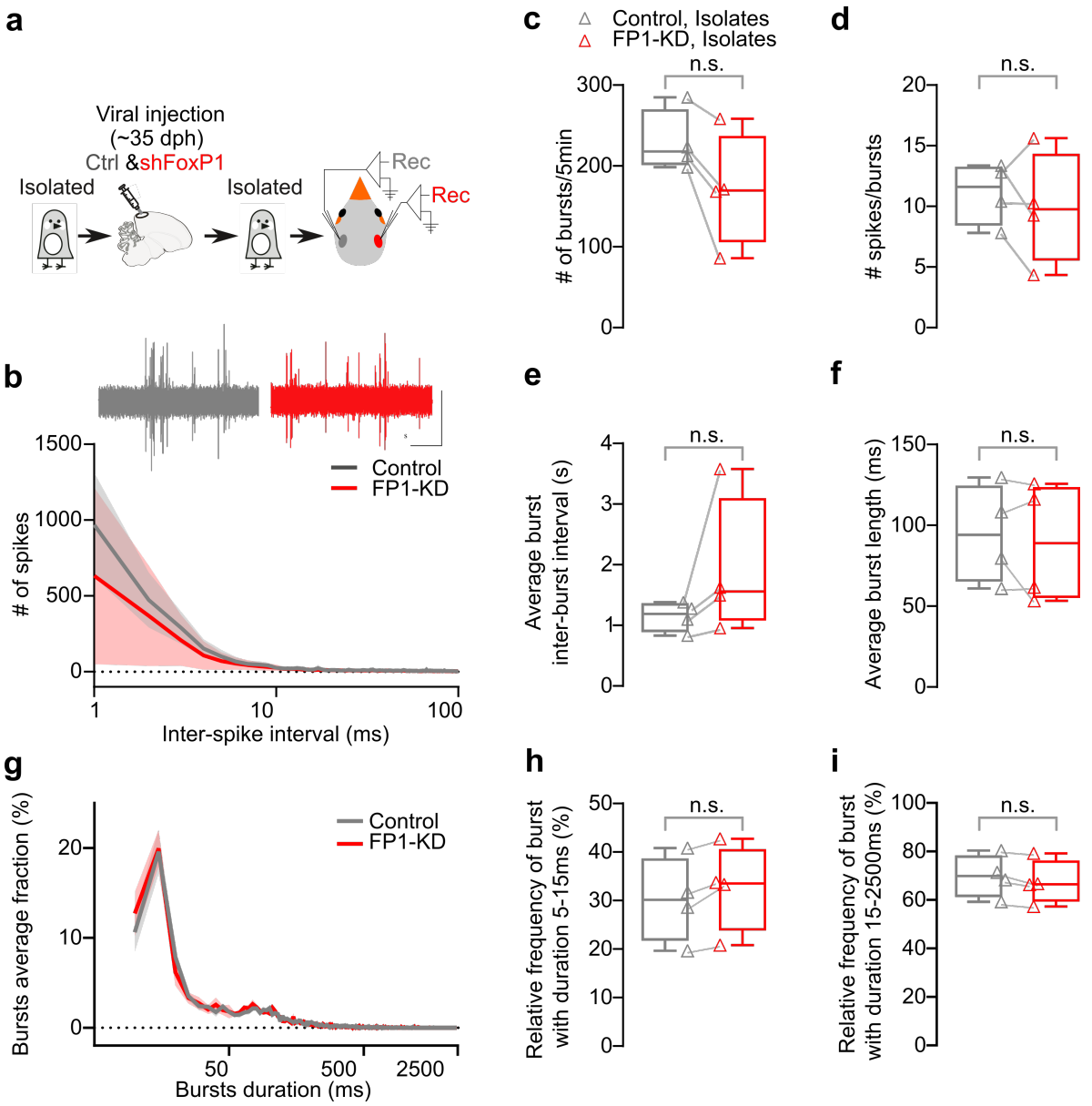
Extended Data Figure 2: FoxP1 expression differs between HVC subtypes, various song nuclei, and in cells overexpressing zfFoxP1.

a-f, FoxP1 labelling in various subtypes of HVC projection neurons: HVC_X (**a**,**d**), HVC_{RA} (**b**,**e**), and HVC_{Av} (**c**,**f**). **a-c,** Schematic of Dextran injections for retrograde labelling of the various HVC neuronal subtypes; **d-f,** Examples of sagittal sections through HVC showing retrograde labelling with Dextran-594 (magenta, left), immunolabelling with FoxP1 (cyan, middle), and a composite image showing coexpression of FoxP1 in the respective HVC neurons (merge, right). Scale bars, 50 µm. **g**, FoxP1 staining in a sagittal brain slice: FoxP1 is prominently present in Area X, HVC, RA and the mesopallium (including Av), and absent in LMAN (as in Mendoza et al., 2015). Scale bar, 500 µm. **h,** (top) Western blot of lysates from 293T cells overexpressing FoxP1, and either control (n = 3) or shFoxP1 (n = 3) AAV constructs. At least

289 two isoforms of FoxP1 were observed around 80 kDa and 70 kDa, with GAPDH as loading control. (bottom)
290 Quantification of top blot. FoxP1 protein signals were normalized to GAPDH, averaged for each condition, and
291 normalized to Controls. Histograms represent average±SEM% of the normalized signal. Cells transduced with pscAAV-
292 GFP-shFoxP1 (FP1-KD) showed significantly reduced FoxP1 expression than those with the control virus. (FoxP1-80:
293 Control: 100±8.9%, n=3, vs FP1-KD: 40.0±0.7%, n=3, Student's t-test, $P = 0.0025$; FoxP1-70: Control: 100±4.8%, n=3,
294 FP1-KD: 22.0±1.3%, n=3, Student's t-test, $P < 0.001$).


a
FP1-KD Behavioral Imitation
B447 | ER = 0.47
G402 | ER = 0.46
G437 | ER = 0.39
G439 | ER = 0.60
O440 | ER = 0.47
R402 | ER = 0.26
R425 | ER = 0.66
Y440 | ER = 0.45
b
FP1-KD Social Experience
S389 | ER = 0.59
Y389 | ER = 1.33
Y397 | ER = 0.96
B402 | ER = 1.54
G413 | ER = 1.62
O421 | ER = 0.95
S421 | ER = 1.45
c
Control Social Experience
O434 | ER = 0.51
Y433 | ER = 0.63
B507 | ER = 1.47
O512 | ER = 0.84
B524 | ER = 0.14
G524 | ER = 0.59
G528 | ER = 0.82
S525 | ER = 0.65
S528 | ER = 0.50
d
Full Isolate
B145 | ER = 0.68
B236 | ER = 1.32
B258 | ER = 1.06
G255 | ER = 1.38
O254 | ER = 0.75
S132 | ER = 0.58
Y453 | ER = 1.33
Bouts
Syllables

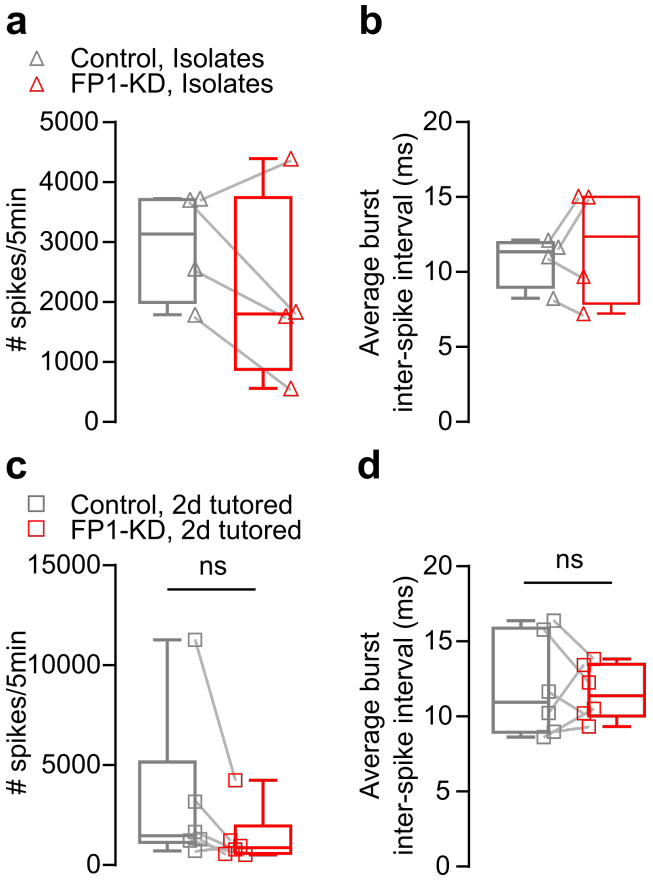
297 **Extended Data Figure 3: HVC FoxP1 knock-down in isolate birds impairs subsequent song sequence stereotypy.**
298 Syntax raster plots from each bird included in song syntax analysis. ER refers to the first order Markov entropy rate
299 score for that individual bird.



Extended Data Figure 4: Unilateral HVC FoxP1 knock-down does not modify in-vivo extracellular recorded spiking and bursting pattern in isolated birds.

a, Schematic representing the experimental timeline. Birds received AAV injections in HVC to knock-down FoxP1 in one hemisphere, and control virus in the other hemisphere (pseudo-randomized). The birds were kept in isolation from male tutors, and ~10 days later, HVC extracellular activity in both hemispheres (n = 4 birds, 3-5 recordings/hemisphere) was recorded. **b,** (top) Sample traces (scalebar 0.5V, 1s) and (bottom) average inter-spike interval distribution (bin 1ms, 1-100ms, logarithmic scale, 300s/recording) (Control vs FP1-KD hemispheres, Two-Way Anova $F_{1,6} = 0.7424$, $P > 0.05$). **c,** Box- and scatter plots of the total number of bursts (Control hemispheres: 229.6±19.1 vs FP1-KD hemispheres: 170.8±35.2, Wilcoxon matched-pairs signed rank test, $P > 0.05$). **d,** Box- and scatter plots of the average number of spikes in a burst (Control hemispheres: 11.1±1.3 vs FP1-KD hemispheres: 9.9±2.3, Wilcoxon matched-pairs signed rank test, $P > 0.05$). **e,** Box- and scatter plots of the average inter-burst interval (Control hemispheres: 1.1±0.1 vs FP1-KD hemispheres: 1.9±0.6, Wilcoxon matched-pairs signed rank test, $P > 0.05$). **f,** Box- and scatter plots of the average burst length (Control hemispheres: 94.6±15.2 vs FP1-KD hemispheres: 89.2±18.5, Wilcoxon matched-pairs signed rank test, $P > 0.05$). **g,** Plot representing the relative distribution of bursts duration, normalized for each recording (5ms duration bins, 5-2500ms, logarithmic scale; Control hemispheres vs FP1-KD hemispheres: Two-Way Anova, Interaction $F_{498,2988}$

317 = 0.05189, $P > 0.05$). **h,** Box- and scatter plots of the average relative prevalence of bursts with duration comprised
318 between 5 and 15ms (Control hemispheres: 30.2±4.4 vs FP1-KD hemispheres: 32.6±4.5, Wilcoxon matched-pairs signed
319 rank test, $P > 0.05$). **i,** Box- and scatter plots of the average relative prevalence of bursts with duration comprised
320 between 15 and 2500ms (Control hemispheres: 69.8±4.4 vs FP1-KD hemispheres: 67.4±4.5, Wilcoxon matched-pairs
321 signed rank test, $P > 0.05$).



Extended Data Figure 5: Spiking properties in unilaterally-injected FP1-KD birds remain unchanged, irrespective of previous social experience.

a, Box- and scatter plots of the average number of spikes recorded in isolates FP1-KDshFoxP1 and Control hemispheres (Control hemispheres: 2946±472 vs FP1-KD hemispheres: 2141±806, Wilcoxon matched-pairs signed rank test, $P > 0.05$). **b,** Box- and scatter plots of the average inter-spike interval in bursting events recorded in isolates FP1-KD and Control hemispheres (Control hemispheres: 10.8±0.9 vs FP1-KD hemispheres: 11.8±2.0, Wilcoxon matched-pairs signed rank test, $P > 0.05$). **c,** Box- and scatter plots of the average number of spikes recorded in 2d-tutored FP1-KD and Control hemispheres (Control hemispheres: 3222±1646 vs FP1-KD hemispheres: 1379±584, Wilcoxon matched-pairs signed rank test, $P > 0.05$). **d,** Box- and scatter plots of the average inter-spike interval in bursting events recorded in 2d-tutored FP1-KD and Control hemispheres (Control hemispheres: 11.9±1.4 vs FP1-KD hemispheres: 11.6±0.8, Wilcoxon matched-pairs signed rank test, $P > 0.05$).

References

Daliparthi, V. K., Tachibana, R. O., Cooper, B. G., Hahnloser, R. H., Kojima, S., Sober, S. J., & Roberts, T. F. (2019). Transitioning between preparatory and precisely sequenced neuronal activity in production of a skilled behavior. *ELife*, *8*, 481960. https://doi.org/10.7554/eLife.43732

Davis, E. P., Stout, S. A., Molet, J., Vegetabile, B., Glynn, L. M., Sandman, C. A., … Baram, T. Z. (2017). Exposure to unpredictable maternal sensory signals influences cognitive development across species. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(39), 10390–10395. https://doi.org/10.1073/pnas.1703444114

Haesler, S., Wada, K., Nshdejan, A., Morrisey, E. E., Lints, T., Jarvis, E. D., & Scharff, C. (2004). FoxP2 Expression in Avian Vocal Learners and Non-Learners. *Journal of Neuroscience*, *24*(13), 3164–3175. https://doi.org/10.1523/JNEUROSCI.4369-03.2004

Mendoza, E., Tokarev, K., Düring, D. N., Retamosa, E. C., Weiss, M., Arpenik, N., & Scharff, C. (2015). Differential coexpression of FoxP1, FoxP2, and FoxP4 in the Zebra Finch (Taeniopygia guttata) song system. *The Journal of Comparative Neurology*, *523*(9), 1318–1340. https://doi.org/10.1002/cne.23731

Roberts, T. F., Gobes, S. M. H., Murugan, M., Ölveczky, B. P., & Mooney, R. (2012). Motor circuits are required to encode a sensory model for imitative learning. *Nature Neuroscience*, *15*(10), 1454–1459. https://doi.org/10.1038/nn.3206

Roberts, T. F., Hisey, E., Tanaka, M., Kearney, M. G., Chattree, G., Yang, C. F., … Mooney, R. (2017). Identification of a motor-to-auditory pathway important for vocal learning, *20*(7). https://doi.org/10.1038/nn.4563

Roberts, T. F., Tschida, K. A., Klein, M. E., & Mooney, R. (2010). Rapid spine stabilization and synaptic enhancement at the onset of behavioural learning. *Nature*, *463*(7283), 948–952. https://doi.org/10.1038/nature08759

Spedicato, G. A., & Signorelli, M. (2014). 4e R Package “markovchain”: Easily Handling Discrete Markov Chains in R.

Spiteri, E., Konopka, G., Coppola, G., Bomar, J., Oldham, M., Ou, J., … Geschwind, D. H. (2007). Identification of the transcriptional targets of FOXP2, a gene linked to speech and language, in developing human brain. *American Journal of Human Genetics*, *81*(6), 1144–1157. https://doi.org/10.1086/522237

Tchernichovski, O., Nottebohm, F., Ho, C. E., Pesaran, B., & Mitra, P. P. (2000). A procedure for an automated measurement of song similarity. *Animal Behaviour*, *59*(6), 1167–1176. https://doi.org/10.1006/anbe.1999.1416

Xiao, L., Chattree, G., Oscos, F. G., Cao, M., Wanat, M. J., & Roberts, T. F. (2018). A Basal Ganglia Circuit Sufficient to Guide Birdsong Learning. *Neuron*, *98*(1), 208-221.e5. https://doi.org/10.1016/j.neuron.2018.02.020