

# Science



Supplementary Materials for

1139 **Vocal learning-associated convergent evolution in mammalian proteins and**  
1140 **regulatory elements**

1141

1142 Morgan E. Wirthlin, Tobias A. Schmid, Julie E. Elie, Xiaomeng Zhang, Amanda Kowalczyk,  
1143 Ruby Redlich, Varvara A. Shvareva, Ashley Rakuljic, Maria B. Ji, Ninad S. Bhat, Irene M.  
1144 Kaplow, Daniel E. Schäffer, Alyssa J. Lawler, Andrew Z. Wang, BaDoi N. Phan, Siddharth  
1145 Annaldasula, Ashley R. Brown, Tianyu Lu, Byung Kook Lim, Eiman Azim, Nathan L. Clark,  
1146 Wynn K. Meyer, Sergei L Kosakovsky Pond, Maria Chikina, Michael M. Yartsev, Andreas R.  
1147 Pfenning

1148

1149 Corresponding authors: [apfenning@cmu.edu](mailto:apfenning@cmu.edu), [myartsev@berkeley.edu](mailto:myartsev@berkeley.edu)

1150

1151 **This PDF file includes:**

1152 Materials and Methods

1153 Supplementary Text

1154 Figs. S1 to S7

1155 Tables S1 to S2

1156

1157 **Other Supplementary Materials for this manuscript include the following:**

1158 Movie S1.mp4

1159 DataS1.speciesAnnotations.4.1.7.1.csv

1160 DataS2.vocalLearningRERConvergeMaster.csv

1161 DataS3.HyPhyResults.1.xlsx

1162 DataS4.GeneOntology.1.xlsx

1163

1164 DataS5.BatM1OcrS.2.xlsx

1165 DataS6.BatM1OcrGo.1.xlsx

1166 DataS7.VocalLearningTACITResults.1.xlsx

1167 DataS8.TACIT\_ttest.zip

1168 DataS9.OCRMouseGeneAnnotation.1.xlsx

1169 DataS10.CellTypeEnrichments.1.xlsx

1170 **Materials and Methods**

1171 *Coding the vocal learning trait*

1172 The vocal learning trait was annotated for a set of 215 mammalian species that satisfied the  
1173 following conditions: (1) the species' genome was present in the Zoonomia whole-genome  
1174 Cactus alignment (241 assemblies total, (17, 99)), (2) the species' protein-coding gene sequences  
1175 were present in the TOGA gene alignment set described in (10) (427 species total), and (3) the  
1176 species was a member of the Boreoeutheria clade (Fig. 1A), given that our primary data were  
1177 restricted to species within this Magnorder.

1178 Although, in many studies, vocal production learning is treated as a binary ("presence /  
1179 absence") trait possessed by humans, bats, pinnipeds, and cetaceans alone among boreoeutherian  
1180 mammals, 'gold standard' tests for the trait have been performed for only a handful of species  
1181 within these clades, and recent reevaluations of the evidence suggest that a more nuanced coding  
1182 of this complex trait is needed, involving extensive behavioral testing of a greater taxonomic  
1183 diversity of species (2, 51, 100). To attempt to account for this, boreoeutherian species were  
1184 coded as vocal learners only if they belonged to an established vocal learning clade (i.e. bats,  
1185 pinnipeds, cetaceans, and humans) and presented evidence of song usage or a rich social acoustic  
1186 repertoire in the literature. Species potentially falling somewhere outside a simplified binary  
1187 coding were excluded from analyses based on the following criteria: (i) species with suggestive,  
1188 insufficient, or controversial evidence of vocal plasticity or learning (*Callithrix jacchus*,  
1189 *Heterocephalus glaber*, *Indri indri*, *Pan paniscus*, *Pan troglodytes*) or (ii) domesticated species  
1190 (or subspecies) given demonstrated connection between domestication and increased vocal  
1191 variability (101) (*Camelus bactrianus*, *Camelus dromedarius*, *Canis lupus familiaris*, *Capra*  
1192 *hircus*, *Equus asinus asinus*, *Equus caballus*, *Felis catus*, *Mustela putorius furo*, *Vicugna pacos*).  
1193 "In sum, this totaled a set of 215 species with genome assemblies, aligned protein predictions,  
1194 and vocal learning annotations (26 high-confidence vocal learners, 25 suspected vocal learners,  
1195 labeled as "unsure", and 164 high-confidence non-learners, Data S1)

1196

1197 *Evaluating the relationship between relative evolutionary rate of protein-coding genes and vocal*  
1198 *learning*

1199 RERconverge (11, 65, 102) was used to evaluate the relationship between relative evolutionary  
1200 rates (RERs) of protein-coding genes and the vocal learning trait across mammals. A set of  
1201 37,552 high-quality protein-coding amino acid alignments generated with TOGA using human  
1202 reference sequences mapped across a human-referenced MAF alignment of 427 species was  
1203 obtained (103). These alignments were subsequently filtered to remove duplicated species,  
1204 poorly represented proteins, and low-scoring alignments. Specifically, alignments with fewer  
1205 than 221 unique species (0.025 quantile of the distribution of unique species number for all  
1206 alignments), alignments with fewer than 189 total species with ungapped coverage of 50% of the  
1207 total alignment length (0.1 quantile), alignments with more than 97 duplicated species (0.95  
1208 quantile), and alignments with ungapped length <267 bp (0.01 quantile) were excluded. In total,  
1209 this resulted in excluding 4,723 transcripts representing 2,613 unique genes. Within the  
1210 remaining alignments, any sequences that did not cover 50% of the total alignment length were  
1211 excluded, and, when there were multiple sequences for a species within an alignment, the  
1212 sequence with the highest identity to the human reference sequence across the full alignment  
1213 length was retained. For genes with multiple transcripts, only the alignment with the longest

1214 median ungapped coverage was retained. From the remaining 16,209 protein-coding gene  
1215 alignments, branch lengths on the consensus species tree from (10) were estimated for each gene  
1216 using approximate maximum likelihood estimation with the WAG substitution model, as  
1217 implemented in the phangorn package in R (104, 105). RERs were generated from these  
1218 topology-constrained gene trees using RERconverge version 0.3.0 (11) with R version 4.1.0.  
1219 RER-to-phenotype correlations were generated using the vocal learning trait annotations  
1220 described above and RERconverge's correlateWithBinaryPhenotype tool, considering all  
1221 foreground branches and otherwise using default parameters. The p-values were corrected using  
1222 the Benjamini-Hochberg procedure (106) and protein-coding genes were considered to have  
1223 significant associations if their corrected p-values were less than 0.01. In addition, empirical p-  
1224 values were computed using phylogenetic permutations (24), which accounts for the structure of  
1225 the tree and the distribution of the trait across the tree, and only genes with permutation  $p < 0.01$   
1226 after Benjamini-Hochberg correction were considered for downstream analyses. To stratify  
1227 which clade or clades are driving changes in the RERconverge results, a Bayes Factor approach,  
1228 which measures the difference in evolutionary rates with that vocal learning clade relative to  
1229 other species (18), was employed. The results of these analyses are found in Data S2.

### 1230 *HyPhy Analysis*

1231 The RELAX analysis, implemented in the HyPhy package, was applied to the same set of  
1232 genomes for which the RERconverge analysis was run (10), with only the confident  
1233 vocal learning species included in the foreground set. Briefly, RELAX  
1234 infers and compares dN/dS distributions from test (vocal learners)  
1235 and reference (all other) lineages, and estimates the  
1236 intensification/relaxation parameter (K). When  $K \neq 1$  with sufficient  
1237 statistical support (LR test), test lineages were inferred to be  
1238 relaxed/intensified compared to the reference lineages. Multiple  
1239 ENSEMBL transcripts per gene were allowed to enable a more detailed  
1240 exploration of each gene. Given the computational constraints of the  
1241 software, HyPhy was run on the set of genes that showed greater  
1242 conservation or acceleration based on the RERconverge analysis. A  
1243 negative set of 100 transcripts randomly chosen from the set of  
1244 transcripts not associated with vocal learning ( $p > 0.5$ ) set, which were  
1245 matched to the RERconverge set to mirror the length distribution  
1246 (Data S3). To best align with the assumptions of RERconverge  
1247 analysis, transcripts with RELAX  $q < 0.01$  and the  
1248 relaxation/intensification parameter (K) estimate were used; values  
1249 of  $0 < K < 1$  indicate relaxed selection, and values  $K > 1$  – intensified  
1250 selection (12, 107). However, results from the BUSTED-PH analysis of associating  
1251 episodic positive selection with a binary phenotype of vocal learning were also provided (12).  
1252 For all analyses, synonymous substitution rates were allowed to vary from site to site because not  
1253 doing so risks subjecting selection analyses to strong confounding biases (108).

### 1254 *Gene Ontology Analysis*

1255 A set of genes were selected that were significantly accelerated or conserved in vocal learning  
1256 species based on a number of criteria (RERconverge adjusted  $P < 0.01$ ; Permulation Adjusted  $P$   
1257  $< 0.01$ ; HyPhy  $q < 0.01$ ). Those genes served as input to the EnrichR web interface with default  
1258 parameters (91). The dataset for “GO\_Biological Process” was used to find the overall function  
1259 of the genes in question. The “Human Phenotype” ontology was used to link the genes to  
1260 specific human phenotypes or disease states. Specific gene ontology categories were chosen for  
1261 visualization based on an adjusted  $p < 0.1$  (Data S4).

1262

### 1263 *Subjects for tracing and electrophysiological experiments*

1264 All experimental procedures were approved by the Institutional Animal Care and Use Committee  
1265 of the University of California, Berkeley. All animals were adult Egyptian fruit bats (weight  
1266 range 130-200g) maintained within the lab colony. For tracing experiments using  
1267 channelrhodopsin injected into ofM1, subjects were three females born in the lab. For tracing  
1268 experiments using AAV<sub>DJ</sub>-hSyn-Synaptophysin as an anterograde tracer injected into ofM1,  
1269 subjects were two males born in the lab. For the 3D mapping of ofM1 descending tracts using  
1270 dextran amine (Figure S4A-B, F and Movie S1), the subjects were two females born in the lab.  
1271 For electrophysiological recordings into ofM1, the four implanted subjects were wild-caught  
1272 bats. 10 other adult bats (seven males and three females) were used as companions during the  
1273 recording sessions to elicit vocal interactions. While the age of the four wild-caught bats could  
1274 not be precisely estimated, all were greater than four years old. Prior to the start of experiments,  
1275 bats were housed in a communal vivarium. After implantation for electrophysiology, the four  
1276 implanted bats were initially single housed and subsequently, following recovery from surgery,  
1277 were co-housed in pairs. All cages were kept in a humidity- and temperature-controlled room, on  
1278 a 12-hour reversed light-dark cycle. Bats were given *ad libitum* access to water and fed with fruit  
1279 mix supplemented with honey and vitamins.

1280

### 1281 *Anesthetic procedure for tracer injection and electrophysiological implant*

1282 The general anesthesia and surgical procedures used for Egyptian fruit bats were previously  
1283 described (109, 110). Anesthesia was induced using a cocktail of ketamine, demedotomidine,  
1284 and midazolam. The depth of anesthesia was evaluated by monitoring the breathing rate, body  
1285 temperature, and toe pinch reflex. The body temperature was monitored continuously using a  
1286 rectal probe and maintained around 34.5°C with a heating pad under the bat. If necessary to  
1287 maintain anesthesia longer than one hour, the bat was injected with a cocktail of  
1288 dexmedetomidine, midazolam, and fentanyl. Anesthesia was reversed with an injection of  
1289 atipamezole and the bat was hydrated subcutaneously with lactated ringer’s solution. For all  
1290 surgical procedures, antibiotics were given for one week post-surgery and analgesic pain  
1291 medication was given for three days post-surgery. For tracing experiments, the sutures were  
1292 removed within five days post-surgery.

1293

### 1294 *Anterograde tracer injection procedures*

1295 The procedures for the injection of the anterograde tracer into the brain and retrograde tracer into  
1296 the cricothyroid muscle were performed at separate times to allow for optimal expression of both  
1297 tracers (111). Two different tracers were used in order to validate the anatomical results across

1298 multiple techniques: AAV mediated channel-rhodopsin (ChR2) conjugated to GFP  
1299 (rAAV5/CamkII-hChR2(H134R)-EYFP, Lot#AV4316LM; UNC Vector Core, NC) (112) and a  
1300 custom synaptophysin/synapsin virus (SYN) from Byungkook Lim's lab that simultaneously  
1301 labeled fibers in eGFP and boutons in mRuby2 (AAV<sub>DJ</sub>-hsyn-mRuby2-T2S-Synap-eGFP; Lim  
1302 Lab, UCSD) (113, 114). Each anterograde tracer type was injected bilaterally into ofM1. The 3D  
1303 reconstruction of ofM1 projections and the localization of the decussation of the pyramidal tract  
1304 (Supp Fig 1 A-B and Supp Fig 1 F) were obtained in two females injected bilaterally in ofM1  
1305 with Dextran amine conjugated to Alexa Fluor 555 (ThermoFisher Scientific; D34679).

1306 For anterograde tracers injected intracranially into ofM1, bats were anesthetized and head fixed  
1307 in a stereotaxic device (Model 942; Kopf, CA). After opening the scalp with a scalpel, the tissue  
1308 was retracted to expose the skull. The center of three injection coordinates for ofM1 (AP:  
1309 +10.72mm, 10.22mm, 9.72mm; ML: +/- 3.2mm) were bilaterally measured from a common  
1310 reference point above the confluence of the sinus. A small craniotomy (1.2 mm long x 0.6mm  
1311 wide) was made above ofM1 to expose the surface of the brain while leaving the dura intact.  
1312 Bilateral injections were made along the anterior-posterior axis into each hemisphere using a  
1313 NanoFil syringe (36GA beveled needle; WPI, FL) attached to the stereotaxic device. The syringe  
1314 was slowly lowered to -1.2mm below the surface of the brain around layers V/VI of ofM1 and  
1315 allowed to rest for three minutes above the deep target. After pausing, 0.5 $\mu$ L of one of the two  
1316 anterograde tracers (ChR2, or SYN) were injected at a rate of 4nl/sec using a microinjection  
1317 pump (UMP3; WPI, FL). The needle was left in place for five minutes at each site. A total  
1318 volume of 1.5 $\mu$ L was delivered in each hemisphere. Upon completion of the six injections,  
1319 Kwik-Sil (WPI, FL) was used to fill the craniotomy and protect the brain and the tissue was  
1320 sutured.

1321

### 1322 *Retrograde injections into cricothyroid muscles*

1323 The retrograde tracer injection was performed approximately one month following the  
1324 anterograde tracer injection to optimize maximal expression of the virus and propagation of  
1325 cholera toxin B (CTB). Approximately one week before the planned perfusion time, the bats  
1326 were anesthetized according to the same procedures above. Once anesthetized, the neck was  
1327 shaved and the bat was placed on its back on a heating pad to facilitate access to the larynx. The  
1328 skin overlying the larynx was incised using a scalpel to reveal the larynx below the sternohyoid  
1329 and infrahyoid muscles. The tissue was retracted to expose the cricothyroid muscle caudal to the  
1330 inferior border of the thyroid cartilage and medial to the cricothyroid joint.

1331 Bilateral injections of cholera toxin B conjugated to fluorescent labels (ThermoFisher, C34778,  
1332 AlexaFluor 647) were made in the cricothyroid muscle at six different points, three on each side.  
1333 A NanoFil syringe (35GA beveled needle; WPI, FL) attached to the stereotaxic device was  
1334 slowly lowered approximately -0.4mm below the surface of the muscle. After waiting one  
1335 minute to allow the tissue to settle, 2 $\mu$ L of CTB was injected at a rate of 16nl/sec into each  
1336 injection site. The needle was left in place for one minute before moving to the next injection  
1337 site. Upon completion of the six injections of 2 $\mu$ L (12 $\mu$ L total) the tissue was sutured and two  
1338 surgical staples were placed over the sutures.

1339

### 1340 *Electrophysiological implant surgery*

1341 The surgical procedures for the implantation of four tetrode microdrives followed those  
1342 described previously in detail for Egyptian fruit bats (46, 110, 115). Each bat was implanted with  
1343 a lightweight microdrive (Harlan 4-Drive, Neuralynx; weight 2.1 g) loaded with four tetrodes  
1344 (~45 $\mu$ m diameter; four strands of platinum-iridium wire, 17.8 $\mu$ m in diameter, HML-insulated)  
1345 that could be moved independently. The tetrodes exited the microdrive through a guide cannula  
1346 with ~300 $\mu$ m horizontal spacing between tetrodes. On the day before surgery, each tetrode's tip  
1347 was cut flat using high-quality scissors (tungsten-carbide scissors ceramic coating, CeramaCut;  
1348 FST) and plated with Gold Plating Solution (Neuralynx) to reduce the impedance of individual  
1349 wires to 0.20-0.55M $\Omega$  (at 1kHz). The principal surgical steps to implant the microdrive were as  
1350 follows: after scoring the skull to improve adhesion and mechanical stability, a circular  
1351 craniotomy of 1.7 mm diameter was made in the skull over the left hemisphere 3.2mm lateral to  
1352 the midline and 10.7 mm anterior to the transverse sinus that runs between the posterior part of  
1353 the cortex and the cerebellum; after a durotomy, the microdrive was placed vertically such that  
1354 the tip of the microdrive's guide tube was flush with the brain's surface; the exposed craniotomy  
1355 was then filled with a biocompatible elastomer (Kwik-Sil, World Precision Instruments) to  
1356 protect the brain; a bone screw (FST) with a soldered stainless-steel wire was fixed to the skull in  
1357 the frontal plate contralateral to the microdrive to serve as a ground screw; an additional set of 3  
1358 bone screws were fixed to the skull to serve as anchor screws for maintaining mechanical  
1359 stability of the implant; finally the exposed skull and anchor screws were covered with a thin  
1360 layer of quick adhesive cement (C&B Metabond, Parkell) to provide a substrate for the adhesion  
1361 of dental acrylic that was added as final layer to secure the entire microdrive to the screws and to  
1362 the skull.

### 1363 *Electrophysiological and audio recording devices*

1364 Electrophysiological recordings were conducted using a wireless neural data logging device  
1365 ("neurologgers"; MouseLog-16 (vertical version), Deuteron Technologies) that interfaced with  
1366 the microdrive of each implanted animal. The neurologger amplified the voltage signals from the  
1367 16 channels of the four tetrodes referenced to the ground screw, performed analog-to-digital  
1368 conversion at a sampling rate of 31.25kHz, and stored the digitized data on a removable 32GB  
1369 SD card that can hold up to 9 hours of recording. The system has a bandwidth of 1Hz - 7kHz,  
1370 records voltage with a fine resolution of 3.3 $\mu$ v, and has a low level of noise generally close to the  
1371 limit of Johnson noise from the impedance of a given source. The neurologger and its lithium-  
1372 polymer battery were encapsulated in a house-made 3D-printed plastic casing to prevent damage  
1373 to the electronics, and weighed a total of 9.9g.

1374 The audio recordings of each individual bat vocalizations were performed using a call detector,  
1375 as previously described (46). In brief, a single-axis, low mass, piezo-ceramic accelerometers  
1376 (BU-27135, Knowles Electronics, sensitivity 0-10kHz) was mounted on a flexible rubber  
1377 necklace placed against the throat of the subject in a way that did not restrict normal behavior to  
1378 detect laryngeal vibrations produced during vocalizations. The signal of the accelerometer was  
1379 recorded, digitized at a sampling rate of 50Hz, and saved on removable SD cards with a wireless  
1380 audio data logging device ("audiologgers"; Audio Logger AL1, Deuteron Technologies)  
1381 mounted on the necklace on the back of the subject. The audiologger and its lithium-polymer  
1382 battery were encapsulated in a house-made 3D-printed plastic casing to prevent damage to the  
1383 electronics. All audiologgers and neurologgers were controlled and synchronized by a single

1384 transceiver. The Egyptian fruit bats in our experiment weighed more than 140g and could fly  
1385 with ease while equipped with both the neurologger and the audiologger.

### 1386 *Vocalizations and motor actions recording sessions*

1387 The four implanted bats were divided into two pairs that were independently recorded for 1-3  
1388 hours per day over multiple days (16 and 32 sessions) with two or three peers. These peers were  
1389 randomly chosen from a pool of ten bats (seven males, three females) and were used as  
1390 companions to increase the probability of vocal interactions implicating the subjects. During the  
1391 daily electrophysiological and audio recording sessions these groups of four to five bats (2  
1392 implanted + 2-3 companions) were housed in a rectangular prism (180 x 60 x 60cm) that had two  
1393 sides made of plexiglass, thereby permitting clear remote visual monitoring of bats behavior via  
1394 2 cameras (Flea 3 FLIR). The remaining sides of the enclosure were made of plastic mesh,  
1395 allowing bats to easily perch and crawl on the surface. The enclosure was placed in an  
1396 electromagnetically and acoustically shielded room and all recording sessions were conducted  
1397 during the dark cycle under red LED light. All bats were equipped with audiologgers such as to  
1398 record and identify their vocalizations. Water was given *ad libitum* and fruits were placed into  
1399 the cage such as to engage the animals into chewing and licking behavior. The experimenter was  
1400 monitoring the behavior of the animals in an ante-chamber via video cameras and an ambient  
1401 ultrasonic microphone (Earthworks, M50) centered 20cm above the cage ceiling and connected  
1402 to the main computer unit via an analog to digital convertor (MOTU, 896mk3). Using a house-  
1403 made keystroke annotation code written under Matlab, the experimenter was manually  
1404 annotating chewing, licking (self-grooming using licking movements), and quiet (staying still in  
1405 a relaxed position, wing closed) behaviors. The audio was recorded throughout the session  
1406 (sampling rate of 192kHz) from the ambient microphone using an in-house Matlab GUI  
1407 (VocOperant; [https://github.com/julieelie/operant\\_bats](https://github.com/julieelie/operant_bats)). The synchronization between the  
1408 microphone recording, the manual annotation of behaviors and the transceiver controlling the  
1409 audiologgers and neurologgers was achieved using transistor-transistor logic (TTL) pulses  
1410 generated by an UltraSoundGate Player 216H (Avisoft Bioacoustics) and sent via coaxial cables.  
1411 After each recording session, tetrodes were connected to a wired recording system (Digital Lynx,  
1412 Neuralynx) to monitor the neural signals and advance the tetrodes. Tetrodes were moved  
1413 downward once every one to two days (generally by 100 $\mu$ m) in order to record single units at  
1414 new sites.

### 1415 *Histology*

#### 1416 Perfusion

1417 Approximately one week following the injection of the retrograde CTB tracer, or after the last  
1418 day of electrophysiological recording, bats were administered an overdose of pentobarbital and  
1419 perfused transcardially with 250ml of pH 7.4 phosphate buffered saline (PBS) spiked with 0.5ml  
1420 heparin (1000 USP units/ml) ---followed by 250ml of fixative (3.7% formaldehyde in phosphate  
1421 buffered saline). When the brain was implanted with electrodes, tetrodes were left in the brain  
1422 for 30 minutes before extracting them. The brain was then carefully removed from the skull and  
1423 post-fixed overnight in the same fixative. To avoid over-fixation, the brain was removed from  
1424 fixative after 24 hours and switched into a 30% sucrose solution for cryoprotection. After  
1425 approximately two days or once the brain had sunk to the bottom, 40 $\mu$ m coronal sections were



1426 made using a microtome (HM450; ThermoFisher, MA) with a freezing stage.  
1427

1428 Staining and Immunocytochemistry

1429 The sections from the implanted bats were Nissl-stained with cresyl violet. Slides were imaged  
1430 using a light microscope to verify tetrode positions.  
1431

1432 The sections of the brains from bats injected with tracers were stained for VGLUT1 using a fresh  
1433 tissue floating immunohistochemistry protocol. Immunohistochemistry was conducted in 12-well  
1434 plates filled with 4ml of solution for washes and blockings and 48 well plates filled with 1 ml of  
1435 solution for primary and secondary incubation. Fifteen brainstem slices were selected from each  
1436 series centered around the NA region and three brainstem slices anterior to the target region were  
1437 selected for antibody control staining. Briefly, the tissue was placed in floating wells on a lab  
1438 rotator in a cold room at 4°C and washed in three separate PBS (0.025M, pH 7.4) solutions for  
1439 five minutes each wash. The tissue was then moved to a blocking solution containing 10% goat  
1440 serum (Sigma-Aldrich, G9023) in 0.3% triton-PBS (Triton X-100 - ACROS Organics, 21568-  
1441 2500 in 0.025M PBS) and rotated for 90 minutes at 4°C. The tissue was incubated overnight at  
1442 4°C in rabbit anti-VGLUT1 primary antibody (provided by Eiman Azim, Salk Institute for  
1443 Biological Studies and produced in Tom Jessell's lab at Columbia University) (116), which was  
1444 prepared in a 1:16,000 dilution in 5% goat serum and 0.3% triton-PBS. Control slices were  
1445 incubated in an antibody buffer without primary antibody. Approximately 16-24 hours after the  
1446 start of the primary antibody incubation, the tissue was moved into three separate 0.3% triton-  
1447 PBS washes for 10 minutes in each wash at 4°C. The secondary antibody was goat anti-rabbit  
1448 conjugated to a fluorescent protein (ThermoFisher A27012, A27018) that did not conflict with  
1449 the anterograde or retrograde tracers (selected wavelength 594 nm for bats injected with ChR2).  
1450 The tissue was incubated in the secondary solution diluted 1:500 in 5% goat serum and 0.3%  
1451 triton-PBS at room temperature for 90 minutes before three final washes in PBS (0.025M, pH  
1452 7.4) for 10 minutes each. DAPI was added to the secondary solution for the final 10 minutes of  
1453 incubation at 1:10,000 dilution (D1306; ThermoFisher, MA). The sections were then mounted on  
1454 glass slides and cover-slipped using ProLong Gold Antifade Mountant (P36934; ThermoFisher,  
1455 MA).

1456

1457 *Imaging and anatomical quantification*

1458 Fluorescent imaging

1459 All imaging was conducted at the University of California, Berkeley Cancer Research  
1460 Laboratory Molecular Imaging Center and the Henry H. Wheeler Jr. Brain Imaging Center at UC  
1461 Berkeley. Preliminary imaging at magnification of 10x/20x using Plan-Apochromat 10x/20x  
1462 objective was conducted on Zeiss Axio Scan Z1 Slide Scanner. Slices were imaged in four  
1463 fluorescent channels – AF647 (Excitation: 653nm, Emission: 668nm, Gain: 0, Exposure time:  
1464 25ms, Filter Cube: 50 Cy5), AF488 (Excitation: 493nm, Emission: 517nm, Gain: 0, Exposure  
1465 time: 20ms, Filter Cube: 38 HE Green Fluorescent Protein), DAPI (Excitation: 353nm,  
1466 Emission: 465nm, Gain: 0, Exposure time: 5ms, Filter Cube: 49 DAPI), and AF598 (Excitation:  
1467 570nm, Emission: 618 nm, Gain: 0, Exposure time: 15ms, Filter Cube: 64 HE mPlum). All  
1468 images were taken with 100% fluorescent lamp strength on Hamamatsu Orca Flash Camera with  
1469 1x Camera Adapter.

1470 More specific imaging of nucleus ambiguus area and other target and control regions was  
1471 conducted at magnification of 63x with Plan Apochromat 63x/1.4 Oil DIC M27 Objective with  
1472 target area (135 $\mu$ m x 135 $\mu$ m) on Zeiss LSM 880 Confocal Microscope at the UC Berkeley  
1473 Molecular Imaging Center. The target region was localized by finding retrogradely labeled cells,  
1474 centering them, and taking a z-stack to encompass the entire cell volume. Each z-stack was taken  
1475 over 1.5 $\mu$ m depth with the above mentioned area and 1 $\mu$ m between each z-stack slice with 4  
1476 fluorescent channels – AF 647 (Laser wavelength: 633 nm, Excitation: 633nm, Emission:  
1477 697nm, Laser Power: 6.0%, Detector Gain: 650, Detector Digital Gain: 1, Detector Offset: 0),  
1478 AF 594 (Laser wavelength: 594nm, Excitation: 594nm, Emission: 659nm, Laser Power: 8.0%,  
1479 Detector Gain: 675, Detector Digital Gain: 1, Detector Offset: 0), AF 488 (Laser wavelength:  
1480 488nm, Excitation: 488nm, Emission: 552nm, Laser Power: 5.0%, Detector Gain: 625, Detector  
1481 Digital Gain: 1, Detector Offset: 0), and DAPI (Laser wavelength: 405nm, Excitation: 405nm,  
1482 Emission: 462nm, Laser Power: 6.0%, Detector Gain: 625, Detector Digital Gain: 1, Detector  
1483 Offset: 0). If relevant, AF 561 (Laser wavelength: 561nm, Excitation: 561nm, Emission: 621nm,  
1484 Laser Power: 8.0%, Detector Gain: 650, Detector Digital Gain: 1, Detector Offset: 0) was taken  
1485 on a separate z-stack due to maximum channels/image microscope limitations. All confocal  
1486 images were taken as 12-bit images with line-by-line imaging and 2x averaging.

1487 Deconvolution was conducted using HuygensPro software from the Biological Imaging Facility  
1488 at UC Berkeley. Theoretical PSF was generated using LSM 880 confocal microscope and 63x  
1489 objective parameters. Image histograms were created using a logarithmic mapping function and  
1490 background was generated using automatic estimation with an area radius of 0.7 $\mu$ m.

1491 Deconvolution was conducted for each channel with 200 maximum iterations, 30 signal to noise  
1492 ratio, 0.01 quality threshold, and with optimized iteration mode.

1493

#### 1494 3D model of ofM1 tracing and localization of decussation

1495 Series of coronal slices were manually stacked to build a 3D model of the fluorescent pathways  
1496 of fluorescently labeled dextran amine (DA) tracing to/from ofM1. Bats were injected with  
1497 fluorescent dextran amine in ofM1 and, following perfusion, coronal slices of the whole brain  
1498 separated by 200 $\mu$ m were stained for DAPI to create a uniform background marker. After slide  
1499 scanning the entire brain (see scanning settings above), the color levels for the DAPI and DA  
1500 channels were equalized for every coronal slice and exported into lossless tiffs from ZEISS ZEN  
1501 Microscope Software. The coronal slices were cut out from against the background and manually  
1502 aligned using Adobe Photoshop so that the edges lined up in a full stack from the rostral tip of  
1503 the olfactory bulbs to the caudal tip of the spinal cord. All images were then stacked in Imaris to  
1504 create a 3D model of the brain that can be rotated to observe the whole structure of the  
1505 corticobulbar pathway from ofM1 to NA (Fig. S4A-B and Movie S1).

1506

#### 1507 Image quantification

1508 Total numbers of cells, fibers, boutons, and DAPI-labeled cells of brains injected with SYN and  
1509 ChR2 were quantified using Imaris software (Version 9.2.1) at the UC Berkeley Cancer  
1510 Research Laboratory Molecular Imaging Center. Retrogradely labeled cells and boutons were  
1511 automatically counted using Imaris Spot Tracker. Initially, the diameter of retrogradely labeled  
1512 cells (15-30 $\mu$ m), boutons (1-3 $\mu$ m), and DAPI labeled nuclei (10-15 $\mu$ m) was measured using

1513 automatic measuring tools in Imaris in a 2D slice and imputed as a parameter for Spot Tracker.  
1514 After computing, x, y, and z, diameter and position was manually adjusted until the spot covered  
1515 the entire retrogradely labeled cell. Fibers were counted and traced using semi-automatic Imaris  
1516 AutoPath Tracer by manually tracing along the length of the fiber. Fiber diameter was set to  
1517 1.4 $\mu$ m across all z-stacks.

1518 Overlap was quantified manually by determining points of colocalization between fiber/cell and  
1519 fiber/cell/bouton. Z-stacks were exported as single multi-channel TIFF images and then opened  
1520 each Z-stack individually in Adobe Photoshop. Points of co-localization were individually and  
1521 manually marked. TIFF images were compared to the original z-stack to confirm the existence of  
1522 fiber/cell/bouton/DAPI-labeled cells over multiple slices and ensure real signal. The number of  
1523 retrogradely labeled cells with CTB that had at least one overlap with fibers and the number of  
1524 retrogradely labeled cells with CTB that had at least one overlap with fibers and boutons were  
1525 counted. If overlap on the same cell spanned multiple TIFF images, it was only counted in the  
1526 first slice in which it appeared.

1527

### 1528 *Acoustic data processing*

1529 Acoustic data logged as voltage traces on the SD cards of the audiloggers were extracted into  
1530 Matlab files and aligned across bats simultaneously recorded using a custom-made Matlab code  
1531 (<https://github.com/NeuroBatLab/LoggerDataProcessing/>). Potential vocalizations were then  
1532 detected and segmented from these piezo recordings using an in-house series of Matlab scripts  
1533 (<https://github.com/NeuroBatLab/SoundAnalysisBats>). The whole process consisted in three  
1534 major steps: 1) detection of sound events on call detector signals, 2) automatic classification  
1535 between vocalizations and noise, 3) manual curation of potential vocalizations.

- 1536 1. Detection: To focus on the detection of vocalizations emitted by the bat wearing the  
1537 collar, the signal of each call-detector was first band passed between 1 and 5kHz. As  
1538 previously described, this frequency range is not contaminated by airborne vocalizations  
1539 from other bats standing close to the collar wearer (46). After determining a noise  
1540 threshold from sections of silence during the recording session for that call-detector,  
1541 potential vocalizations were detected by threshold crossing on the amplitude envelope  
1542 (RMS, sampling frequency 1kHz) and any sound event above threshold for longer than  
1543 7ms was kept.
- 1544 2. Automatic step of data sorting between actual vocalizations and noise: Sound events  
1545 closer than 50ms were merged as a single sound sequence and a battery of 20 acoustic  
1546 measurements were applied on them (see (117) for mathematical definitions): RMS,  
1547 maximum amplitude, the five first momentum of the amplitude envelope taken as a  
1548 distribution (mean, standard deviation, kurtosis, skewness, and entropy), the five first  
1549 momentum of the frequency spectrum taken as a distribution, the three quartiles of the  
1550 frequency spectrum, the mean pitch saliency, and four parameters that pertain to the  
1551 sound as recorded from the ambient microphone (RMS, mean and maximum of the  
1552 amplitude envelope, and maximum value of cross-correlation between the microphone  
1553 and the call detector signals). Potential vocalizations among the detected sound events  
1554 were then identified using these 20 acoustic parameters as input to a support vector  
1555 machine trained on the data of two sessions manually sorted between vocalizations and  
1556 noise by an expert (JEE). This automatic sorting was set to be very conservative of  
1557 vocalizations by setting the threshold on the posterior probability of a vocalization at 0.2.

1558 3. Manual curation: to further eliminate noise and check the identity of the bat producing  
1559 vocalizations, each potential vocalization was aurally and visually scrutinized by an  
1560 expert (JEE) based on the inspection of the spectrograms of its signal as recorded from  
1561 the ambient microphone and from the call detectors of all the bats. After this step,  
1562 vocalizations further than 200ms apart were considered as distinct, while the others were  
1563 merged into a single sequence.

1564  
1565 *Neural data processing*

1566 Spike detection and sorting

1567 Neural data logged as voltage traces on the SD cards of the neurologgers were extracted into  
1568 Matlab files and aligned across simultaneously recorded bats and with the audio data, using  
1569 custom Matlab code (<https://github.com/NeuroBatLab/LoggerDataProcessing/>). Spike detection  
1570 and sorting was achieved using a generative algorithm (Kilosort 2.0,  
1571 <https://github.com/MouseLand/Kilosort/releases/tag/v2.0>, (118)) with the parameters set as  
1572 indicated under [https://github.com/julieelie/Kilosort2\\_Tetrode/configFiles/configFile16.m](https://github.com/julieelie/Kilosort2_Tetrode/configFiles/configFile16.m).  
1573 Clusters were further manually curated using Phy (<https://github.com/cortex-lab/phy>).

1574 Units obtained after manual curation were sorted between multi-units and single units by  
1575 applying thresholds on two metrics: 1) the consistency of the spike snippet (signal over noise  
1576 ratio of the spike snippet, SNR, estimated at the four contact points of the tetrode) and 2) the  
1577 respect of the refractory period (percent of violation of the refractory period, VRP). A unit had to  
1578 have at least one out of four SNR values above 5 and VRP<0.1% to be considered as a single  
1579 unit.

1580 For each unit, the spike snippet SNR was calculated for each electrode of the tetrode as:

1581 
$$SNR = \frac{\max(MeanSpike) - \min(MeanSpike)}{\sqrt{(StdSpike_{tMax}^2 - StdSpike_{tMin}^2) \div 2}}$$

1582 with *MeanSpike* the average spike snippet over all spikes, *StdSpike* the time-varying standard  
1583 deviation of the spike snippet over all spikes, and *tMax* and *tMin* the time point at which the  
1584 average spike snippet reaches maximum and minimum values, respectively.

1585 Applying the above mentioned thresholds on SNR and VRP of units manually curated with Phy  
1586 yielded 381 single units (SNR = 8.15 +/- 0.12; VRP = 0.0308 +/- 0.0015 %). 94.2% of single  
1587 units (359/381) had an index of contamination Q (118–120) below 0.2 (Q = 0.0531 +/- 0.0042,  
1588 N=381) and a significant test of refractoriness against a Poisson distribution at p<0.05.

1589  
1590 Firing rate calculations and analysis

1591 For this analysis, we selected the single units that were recorded during the sessions when the  
1592 subject had produced a minimum of ten vocalizations longer than 100 ms and had displayed  
1593 chewing, licking and quiet behaviors (n= 237 units). For each single unit, the time average firing  
1594 rate during the production of vocalizations longer than 100ms was estimated for the duration *D*  
1595 of each vocalization starting 10ms before vocalization onset as:

1596 
$$R = \frac{N_{AP}}{D + 10}$$
 with  $N_{AP}$  the number of action potentials occurring during the time window  $D +$   
1597  $10\text{ ms}$ .

1598 For all other behaviors (chewing, licking, and quiet), the annotated period of time where the bat  
1599 was demonstrating that behavior was reduced by 1 second (true offset considered as 1s before

1600 keystroke) to conservatively accommodate for the annotator time response. Then the firing rate  
1601 was estimated in time segments of the same durations as those used to estimate firing rate during  
1602 vocalizations by randomly sampling without replacement into the period of time where the bat  
1603 was demonstrating the behavior of interest.

1604 Firing rate comparisons between pairs of behaviors were achieved by applying a test on the  
1605 deviance of the Poisson Generalized Linear Model (GLM) predicting the rate  $R$  as a function of  
1606 the category of behavior (function `fitglm` of Matlab). P-values were corrected for false detection  
1607 rate using the Benjamini-Hochberg procedure.

1608

#### 1609 Information on coherence during vocalization perception and production

1610 The relationships between each single unit activity and vocal production or vocalization  
1611 perception were quantified by calculating the coherence between the time varying amplitude of  
1612 vocalizations and the time varying firing rate of the unit during the vocalizations respectively  
1613 produced or heard. For this analysis, we selected the single units that were recorded during the  
1614 sessions when the subject had both produced and heard a minimum of twenty vocalizations  
1615 ( $n=219$  units). Vocal activity, as measured by the piezoelectric sensor of the call detector, and  
1616 neural activity, as represented by the arrival times of action potentials, were collected from  
1617 200ms prior to the onset to 200ms after the offset of each vocalization. The time-varying  
1618 amplitude of each sound extract was taken as the amplitude envelope calculated with a frequency  
1619 cut-off at 150Hz and sampled at 1000Hz (BioSound python package,  
1620 <https://github.com/theunissenlab/soundsig>, (117)). The spike pattern corresponding to each  
1621 sound extract was convolved with a Gaussian window of 1ms standard deviation to obtain a  
1622 time-varying firing rate sampled at 500 Hz. The coherence between the time-varying amplitude  
1623 and the time-varying firing rate was then calculated across all vocalizations that the animal had  
1624 produced/heard during the session to give the motor/auditory coherence. A multitaper approach  
1625 was implemented in the estimation of coherence to obtain an error measure (121). The  
1626 information on coherence was obtained by integrating all values of positive coherence up to the  
1627 Nyquist limit. Information was calculated on the estimate of coherence and on its lower and  
1628 upper bounds both for vocalizations produced and vocalizations heard, yielding a value of motor  
1629 information  $Info_{motor}$  and its corresponding lower and upper bounds  $Info_{motor}^{low}$  and  $Info_{motor}^{up}$   
1630 as well as a value of auditory information  $Info_{aud}$  and its lower and upper bounds  $Info_{aud}^{low}$  and  
1631  $Info_{aud}^{up}$ . For each unit, the information D-Prime was then calculated as:

$$1632 \quad DPrime = \frac{Info_{motor} - Info_{aud}}{\sqrt{E_{motor}^2 + E_{aud}^2}}, \text{ with } E_{motor} = Info_{motor}^{up} - Info_{motor}^{low} \text{ and } E_{aud} = Info_{aud}^{up} -$$

$$1633 \quad Info_{aud}^{low}.$$

1634

1635 The significance of the difference in information between produced (motor) and heard (auditory)  
1636 vocalization was assessed by a likelihood ratio test between two linear mixed effect (LME)  
1637 models predicting the information value with or without the type of information (motor vs  
1638 auditory) as a fixed effect and with the identity of the subject as a random variable.

1639

#### 1640 *Animals and sample collection for epigenomics*

1641 All animal procedures were in accordance with the National Institutes of Health Guide for the  
1642 Care and Use of Laboratory Animals and approved by the Institutional Animal Care and Use  
1643 Committees the University of California, Berkeley. Two adult (>1 year) male Egyptian fruit bats

1644 (*Rousettus aegyptiacus*), one male and one female, were housed socially in a large free-flight  
1645 vivarium. Bats were acoustically and socially isolated in a sound recording chamber (*external*  
1646 *dimensions*: 61 X 65 X 61cm; *internal dimensions*: 51 X 61 X 61cm) the night prior to  
1647 experiments. Bats were acoustically monitored to confirm non-vocalizing status prior to  
1648 experiments in order to control for the effects of activity-induced expression. To control for  
1649 circadian effects, all experiments were performed between 8 and 10am. Bats were administered  
1650 with an overdose of pentobarbital with an intraperitoneal injection. We then rapidly opened the  
1651 skull and removed the brain using round-tipped safety scissors. Brains were sliced coronally into  
1652 300  $\mu$ m sections in a vibrating microtome (Leica VT 1200) in ice-cold, oxygenated artificial  
1653 cerebrospinal fluid [119mM NaCl, 2.5mM KCl, 1mM NaH<sub>2</sub>PO<sub>4</sub> (monobasic), 26.2mM  
1654 NaHCO<sub>3</sub>, 11mM glucose] and regions of interest were excised under a dissection microscope.  
1655 Liver, muscle, and gonads were collected immediately. Tissues were preserved in a  
1656 cryoprotectant medium (CryoStor CS10, Biolife Solutions) in cryovials, which we placed in a  
1657 foam freezing container (CoolCell, Corning) and transferred to a -80°C freezer in order to ensure  
1658 a controlled freezing rate of -1°C per minute.

1659

#### 1660 *CryoATAC-seq protocol*

1661 Tissue samples were processed as described previously (61, 122, 123) with the following minor  
1662 differences in procedure and reagents. Cryopreserved samples were warmed in a 37°C water bath  
1663 for 2 minutes and then transferred into 12 mL PBS supplemented with a protease inhibitor  
1664 cocktail (Roche) and gently mixed by inversion. Samples were centrifuged at 300 rcf for 5  
1665 minutes at 4°C before aspirating all supernatant and resuspending the samples in ice cold lysis  
1666 buffer (61). Nuclei were isolated from dissected tissues using 30 strokes of homogenization with  
1667 the loose pestle (0.005 in. clearance) in 5mL of cold lysis buffer placed in a 15 mL glass Dounce  
1668 homogenizer (Pyrex #7722-15). Nuclei suspensions were filtered through a 70  $\mu$ m cell strainer,  
1669 pelleted by centrifugation at 2,000 x g for 10 minutes, resuspended in water, and filtered a final  
1670 time through a 40  $\mu$ m cell strainer. Sample aliquots were stained with DAPI (Invitrogen  
1671 #D1206), and nuclei concentrations were quantified using a manual hemocytometer under a  
1672 fluorescent microscope. Approximately 50,000 nuclei were input into a 50  $\mu$ L ATAC-seq  
1673 tagmentation reaction as described previously (61, 122). The resulting libraries were amplified to  
1674 1/3 qPCR saturation, and fragment length distributions estimated by the Agilent TapeStation  
1675 System showed high quality ATAC-seq fragment length periodicity. We shallowly sequenced  
1676 barcoded ATAC-seq libraries at 1-5 million reads per sample on an Illumina MiSeq and  
1677 processed individual samples through the ENCODE ATAC-seq pipeline (version 1.8.0, accessed  
1678 at <https://github.com/ENCODE-DCC/atac-seq-pipeline>) for initial quality control. We used the  
1679 QC measures from the pipeline (clear periodicity, library complexity, and minimal  
1680 bottlenecking) to filter out low-quality samples and re-pooled a balanced library for paired-end  
1681 deep sequencing on an Illumina NovaSeq 6000 System through Novogene services to target >30  
1682 million uniquely mapped fragments per sample after mitochondrial DNA and PCR duplicate  
1683 removal.

1684

#### 1685 *ATAC-seq data processing*

1686 Raw FASTQ files of ATAC-seq experiments were processed with the ENCODE ATAC-seq  
1687 pipeline (version 1.8.0, accessed at <https://github.com/ENCODE-DCC/atac-seq-pipeline>) to

1688 identify open chromatin region (OCR) peaks from sequenced samples. The ENCODE pipeline  
1689 was run using the mRouAeg1.p assembly (124). The pipeline was run with the default  
1690 parameters except for “atac.multimapping” : 0, “atac.cap\_num\_peak”: 300000,  
1691 “atac.smooth\_win”: 150, “atac.enable\_idr”: true, and “atac.idr\_thresh”: 0.1. Filtered bam files,  
1692 peak files, and signal tracks were generated for each biological replicate and the pool of  
1693 replicates for each tissue. To account for differences in sequencing depth between samples,  
1694 reproducible peak sets were identified, which were defined as peaks with an irreproducible  
1695 discovery rate (IDR, (125)) < 0.1 across pooled pseudo-replicates. All of our samples except for  
1696 the two liver samples were high-quality, displaying high fragment length periodicity, low PCR  
1697 duplicate rates, and concordance between biological replicates.

1698 In addition to identifying peak sets for individual tissues, consensus peak sets were identified to  
1699 serve as genome-wide background sets representing the intersection of the reproducible open  
1700 chromatin peaks identified from all processed tissues, from all cortical samples, or from all  
1701 motor cortex samples. These background sets were obtained using bedtools (version 2.25.0,  
1702 (126)) intersect with the -wa and -u options to combine reproducible peak sets. OCRs were  
1703 prepared for downstream analysis in the following way: peaks within 50 bp of one another were  
1704 combined using bedtools merge, preserving the summit location as the average of the summits of  
1705 all merged peaks. bedtools subtract with option -A was used to remove those peaks that were  
1706 within 1.5 kb from any annotated coding or noncoding exons, enabling us to exclude promoters,  
1707 coding sequences, and noncoding RNAs from our background set. Peaks greater than 1.5 kb in  
1708 width were also removed. In order to identify the complete set of exonic exclusion regions for  
1709 Egyptian fruit bat, the complete set of mRouAeg1.p annotations was used (124).

1710 To identify OCR peaks differentially active between tissues, the number of reads from each  
1711 tissue that aligned to the consensus peaksets described above was quantified using featureCounts  
1712 (127). The read counts at each peak between tissues were contrasted using the negative binomial  
1713 model in the DESeq2 R package (128). Differential peaks were identified based on their Wald  
1714 statistic value, with a statistical cutoff of unadjusted  $p < 0.05$  (Data S5).

1715 In order to identify OCR orthologs across species, the consensus peaksets were aligned as well as  
1716 their peak summits across all of the species present in the Zoonomia Cactus alignment (17, 99)  
1717 using halLiftover (93) with default parameters. In the case of the bat ofM1 and wM1 data, we  
1718 used consensus peaks that were common across subregions for both training and mapping. This  
1719 was done so that the bat M1 data would most closely mirror the bulk motor cortex tissue  
1720 collected in other species. Our conservative procedure led to strong training and evaluation of bat  
1721 motor cortex models. However, this also means that we have not directly traced the evolutionary  
1722 history of the ofM1 or wM1-specific open chromatin regions. The motor cortex was not  
1723 subdivided in any other species and so it would be impossible to validate the accuracy of these  
1724 models. The raw outputs of halLiftover were filtered and assembled into contiguous OCRs using  
1725 HALPER (94) with parameters -max\_frac 2.0, -min\_len 50, -protect\_dist 5, and -narrowPeak.  
1726 The sequences underlying these OCR summit orthologs' +/- 250bp were obtained using  
1727 fastaFromBed in bedtools (126).

1728 Single-nucleus ATAC-seq data for model training and evaluation was processed as described in  
1729 our previous work (23), and orthologs across species and their sequences were identified in the  
1730 same way that they were identified for OCRs from bulk ATAC-seq data. For evaluating house  
1731 mouse bulk motor cortex OCR ortholog overlap with OCRs from specific cell types, we re-  
1732 processed the data (downloaded from NEMO archive:

1733 [https://data.nemoarchive.org/biccn/grant/u19\\_cemba/cemba/epigenome/sncell/ATACseq/mouse/](https://data.nemoarchive.org/biccn/grant/u19_cemba/cemba/epigenome/sncell/ATACseq/mouse/)  
1734 ; metadata from Supplemental Table 1 of (71)) to include all cortical regions (instead of only  
1735 motor cortex) so that we could have enough reads to obtain sufficiently many peaks from each  
1736 cell type to have the power for overlap enrichment analysis.

1737 The raw sequencing reads were realigned with chromap with default parameters (129). The  
1738 fragment files from chromap were next processed with the ArchR comprehensive suite of  
1739 algorithms for processing snATAC data (130). Doublets and empty droplets were identified with  
1740 default recommended parameters for each biological replicate. The data were clustered together,  
1741 and cell clusters were identified, some of which contained predominantly low-quality cells. The  
1742 high-quality cell clusters contained cell types that were not previously annotated, cell types in  
1743 this data were re-annotated using the published mouse whole cortex snRNA-seq cell type  
1744 taxonomy defined by the Allen Brain Institute (131). Clusters were also visually inspected for  
1745 quality concerns, which led to one cluster being eliminated. The addGeneIntegrationMatrix()  
1746 function in ArchR was used to perform cell type label transfer from the mentioned snRNA-seq  
1747 dataset to the snATAC-seq dataset, followed by manual correction of minor mapping mistakes.  
1748 ArchR was then used to call peaks for distinct cell types across each biological replicate, and  
1749 reproducible peaks that were detected in more than half of the biological replicates were  
1750 identified. Finally, these peaks were filtered to retain enhancer regions, defined as regions >20k  
1751 base pairs away from transcription start sites (TSS). This was done using a reference mouse  
1752 genome annotation file containing TSS locations from Gencode version 15 (132) and applied to  
1753 the ArchR reproducible peaks.

1754

#### 1755 *Bat OCR Gene functional enrichment analyses*

1756 Gene ontology analyses was performed using GREAT version 4.0.4 (133) and g:Profiler (134).  
1757 GREAT was run on the genomic coordinates of ofM1 versus wM1 differential bat OCR peak  
1758 sets mapped to human (hg38) using halLiftover (93) and HALPER (94) with the same  
1759 parameters that were used for other analyses (Data S4). g:Profiler was also run to identify  
1760 functional enrichments in genes demonstrating relative evolutionary rate convergence in vocal  
1761 learners using default parameters (Data S6).

1762

1763

#### 1764 *Predicting OCR ortholog open chromatin activity across species and associating predictions* 1765 *with vocal learning*

1766 The Tissue-Aware Conservation Inference Toolkit (TACIT) was used to identify open chromatin  
1767 regions (OCRs) whose predicted open chromatin differences between species are associated with  
1768 differences in vocal learning (23). Specifically, the association between vocal learning and open  
1769 chromatin predictions across 222 boreoeutherian species from Zoonomia from models trained on  
1770 Egyptian fruit bat (this study), Brown Norway rat (20), C57BL/6J mouse (21), and Rhesus  
1771 macaque motor cortex was investigated. If open chromatin data was available from multiple  
1772 motor cortical regions, the intersection of OCRs was used, as described in our previous work  
1773 (23). In addition to using the intersection of OCRs across motor cortical regions, OCRs were  
1774 limited to regions < 1kb, > 20kb from the nearest protein-coding TSS, and not overlapping  
1775 protein-coding exons (20), and M1-PV+ OCRs were filtered in the same way neuron (22, 23,



1776 71). From the cross-species mapped OCR consensus peak sets, motor cortex OCRs that did not  
1777 have usable orthologs in at least half of the species with vocal learning annotations (provided in  
1778 Data S5) and at least three vocal learning species were removed, as these OCRs were unlikely to  
1779 have sufficient power to identify associations. Likewise, motor cortex OCRs that did not have a  
1780 usable ortholog in at least one species in Chiroptera (bats), the order with the largest number of  
1781 mammalian vocal learners, were removed. In addition, motor cortex OCRs that did not have at  
1782 least one usable ortholog in a non-chiropteran vocal learner were removed, as any association  
1783 found for such OCRs would likely be explainable by phylogeny (i.e., driven by bats) and  
1784 therefore not be able to be found through phyloglm (72, 73), the method that was used for  
1785 associating predicted open chromatin with phenotypes while correcting for phylogeny. After  
1786 applying these filtering steps, empirical p-values were calculated using phyloglm (72, 73) with  
1787 phylogenetic permutations (24) and the conditional p-value method (23, 135) and were corrected  
1788 using Benjamini-Hochberg (106) (Data S8).

1789 To determine if M1-PV+ OCRs with human orthologs near autism-associated genes tended to be  
1790 more significantly associated with vocal learning than other M1-PV+ OCRs with human  
1791 orthologs, the human orthologs of all M1-PV+ OCRs were filtered for those of OCRs tested by  
1792 TACIT. Then, autism-associated genes were obtained by downloading the complete human  
1793 autism gene list from the SFARI Gene database (<https://gene.sfari.org/database/human-gene/>) on  
1794 September 18, 2023 (82). These genes' transcription start sites were identified using human  
1795 GENCODE version 27 (132). OCRs with human orthologs within 1Mb of these genes  
1796 transcription start sites were found using window from bedtools version 2.29.2 (126). The  
1797 uncorrected empirical p-value distribution for the OCRs tested by TACIT with human orthologs  
1798 within 1Mb of the these genes was compared to that for the other OCRs tested by TACIT with  
1799 human orthologs using the Wilcoxon rank-sum test. This process was repeated for motor cortex  
1800 OCRs with human orthologs that were tested by TACIT, though the result was not statistically  
1801 significant. Both p-values were multiplied by two for a Bonferroni correction.

1802 To identify whether a specific clade was driving TACIT results, whether each of the vocal  
1803 learning clades showed differential predicted open chromatin relative to other closely related  
1804 species was tested. A t-test between each vocal learning clade within Laurasiatheria (bat,  
1805 cetacean, pinniped) was used to identify differences in the distribution relative to vocal non-  
1806 learning Laurasiatheria. For the human lineage, a t-test to identify whether that specific data point  
1807 was likely to be drawn from a normal distribution of the Euarchonta open chromatin predictions  
1808 was used (Data S10).

1809 From a broader set of motor cortex OCRs associated with vocal learning (permutations  
1810 unadjusted  $p < 0.05$ ), whether or not there was a bias to overlap open chromatin from specific  
1811 cortical cell types was tested. Cell type-specific open chromatin based on single-nucleus ATAC-  
1812 Seq experiments from the mouse motor cortex (74) was measured and processed as described in  
1813 our previous work (23). The bulk motor cortex data that the models are based on, seem to  
1814 contain a higher proportion of neurons relative to glial cells. To account for potential biases in  
1815 representation, an empirical p-value for the enrichment of each cell type comparing the number  
1816 of overlaps observed in the true data with the number of overlaps observed in 1,000 random  
1817 samples of the same number of OCRs that were not associated with vocal learning was  
1818 calculated.

1819

## 1820 **Supplementary Text**

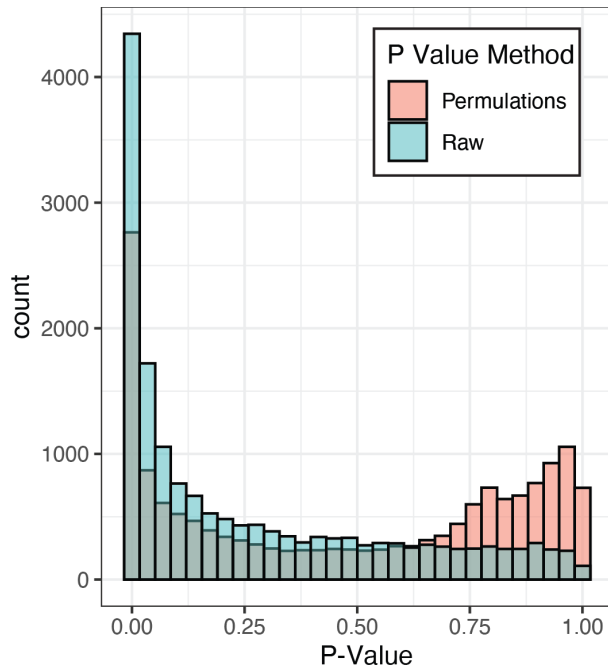
1821 Multiple lines of evidence suggest a model in which differences in gene regulatory networks of  
1822 transcription factors mediate behavioral differences in vocal learning across clades. At the  
1823 protein-coding level, we found evidence of vocal learning-associated convergent evolution in  
1824 200 genes. The 126 genes which are evolving more slowly in vocal learning species are enriched  
1825 for transcriptional regulatory function. At the noncoding level, TACIT, a machine learning  
1826 approach, identified 50 candidate enhancers whose predicted open chromatin states across  
1827 species is associated with vocal learning behavior, including OCRs near genes whose expression  
1828 was previously associated with vocal learning (4) (Table 1 and 2). The predicted and measured  
1829 differences in open chromatin are likely due to gains and losses of transcription factor binding  
1830 sites at those regions of open chromatin (23). These findings are consistent with previous work  
1831 linking vocal learning to transcription factors like *FOXP2*(62), *NEUROD6*, and the MEF2 family  
1832 (44). Although our *cis*-regulatory results clearly implicate motor cortex specializations, we  
1833 cannot rule out that these candidate enhancers and their nearby vocal learning-associated genes  
1834 impact a number of different brain regions and cell types.

1835 For both protein-coding genes and for open chromatin regions, we find greater concordance  
1836 between vocal learning species within Laurasiatheria (bats, cetaceans, and pinnipeds) than with  
1837 any of those clades and humans (Fig. 1B,C; Fig. 4B,C). These differences could arise because  
1838 humans are the most distantly related vocal learning species within our analysis. In addition, it  
1839 may be the case that our analysis is finding genetic signals of traits that are correlated with but  
1840 not sufficient for vocal learning (2, 100) Exciting extensions to this work would be to apply our  
1841 methodology to a continuous measure of vocal learning or to multiple vocal learning  
1842 components.

1843 Across our various approaches (protein-coding analysis, bat orofacial specialized, and OCR  
1844 analysis), the results broadly support the role of specific neurodevelopmental transcriptional  
1845 regulatory networks in the convergent evolution of vocal learning behavior. At the pathway  
1846 level, NFAT transcription factor signaling emerged as a common theme. The NFATC3 protein  
1847 showed increased conservation in vocal learning mammals (RERcConverge Tau p.adj. p =  
1848 0.0027; permutations p.adj. p = 0.019). An OCR near the *NFATC2IP* coding sequence showed  
1849 differential open chromatin between wingM1 and ofM1 in Egyptian fruit bat (Data S6). The  
1850 TACIT analysis also implicated NFAT signaling, with one of the candidate enhancers lying near  
1851 the *NFATC2* locus. Previous studies have demonstrated a potential role for NFAT signaling in  
1852 motor degeneration in Parkinson's disease, as well as in with gene expression specializations in  
1853 the songbird brain for vocal learning (136, 137). Further research is needed to identify the  
1854 specific role that NFAT signaling could play in specialization of motor control regions.

1855 These results from mammals show similarities to results from avian vocal learners at specific  
1856 genes. The behavior of vocal learning is associated with shared patterns of gene expression,  
1857 especially decreases in expression, in the vocal motor control region (4). Using TACIT, we were  
1858 able to find several candidate enhancers with lower predicted motor cortex open chromatin levels  
1859 in vocal learning mammals relative to vocal non-learning mammals near these convergently  
1860 evolved genes. These regulatory elements could contribute to the convergent evolution in the  
1861 expression of *DAAMI*, *VIP*, and *SCNA* genes that are shared between human and song-learning  
1862 birds. Out of the 50 candidate convergently evolved enhancers, several others showed  
1863 concordance with expression patterns currently observed only in songbirds, including OCRs  
1864 predicted to be more open in vocal learners near *KCNQ5* and *CNTNAP4* and OCRs predicted to

1865 be more closed in vocal learners near *PRKCE*, *MAGI2*, *CDH8*, and *ATXN1* (137). Finer- grained  
1866 single -cell profiling of gene expression between vocal motor and other subregions of M1 may  
1867 provide insight into whether any of these specializations are shared with human or other  
1868 mammalian vocal learners.



1869

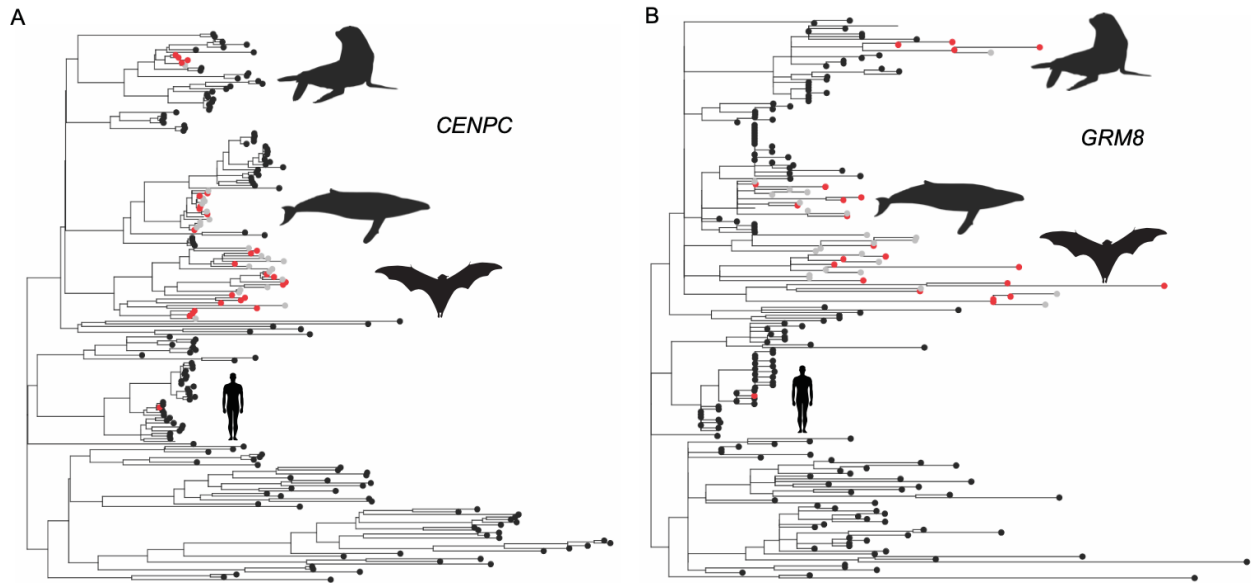
1870 **Fig. S1.**

1871 Histogram comparison of raw and permutation RERconverge P-values. RERconverge was run  
 1872 for aligned proteins across the Zoonomia mammal dataset. A histogram of RERconverge p-  
 1873 values is shown, which was obtained by applying to the Kendall rank correlation of the gene  
 1874 relative evolutionary rates to the trait of vocal learning (*blue*). After running RERconverge,  
 1875 phylogenetic permutations was applied to obtain a p-value adjusted for the distribution of the  
 1876 vocal learning trait across the phylogenetic tree (*red*). The application of permutations reduced  
 1877 the number of candidate vocal learning-associated genes, as indicated by the shift of p-values  
 1878 towards less significance.

1879

1880

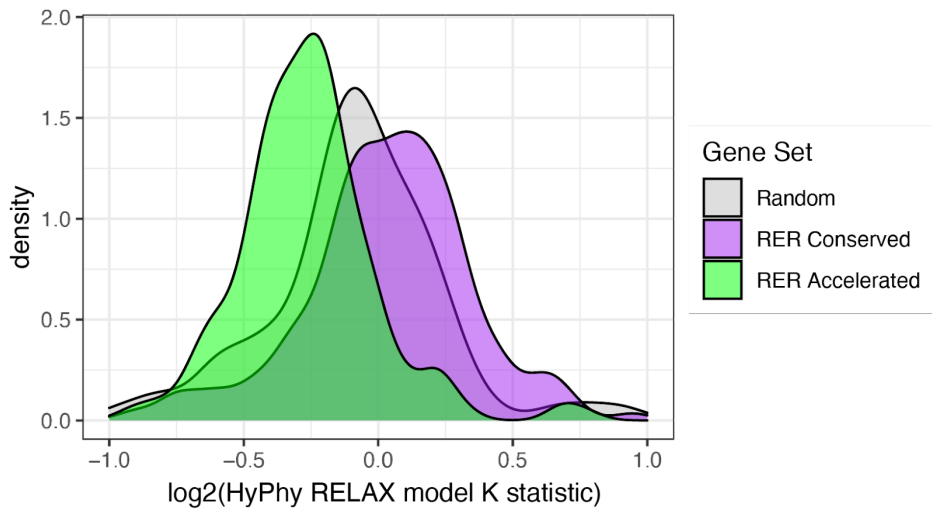
1881



1882  
1883

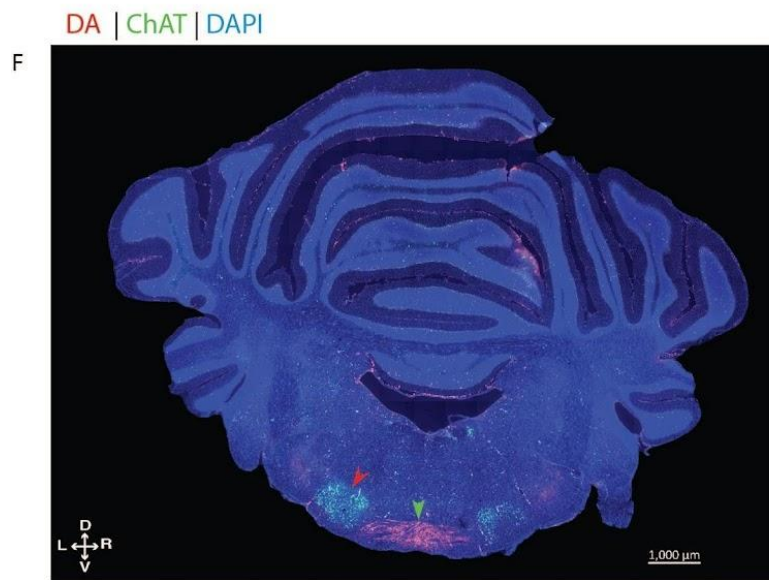
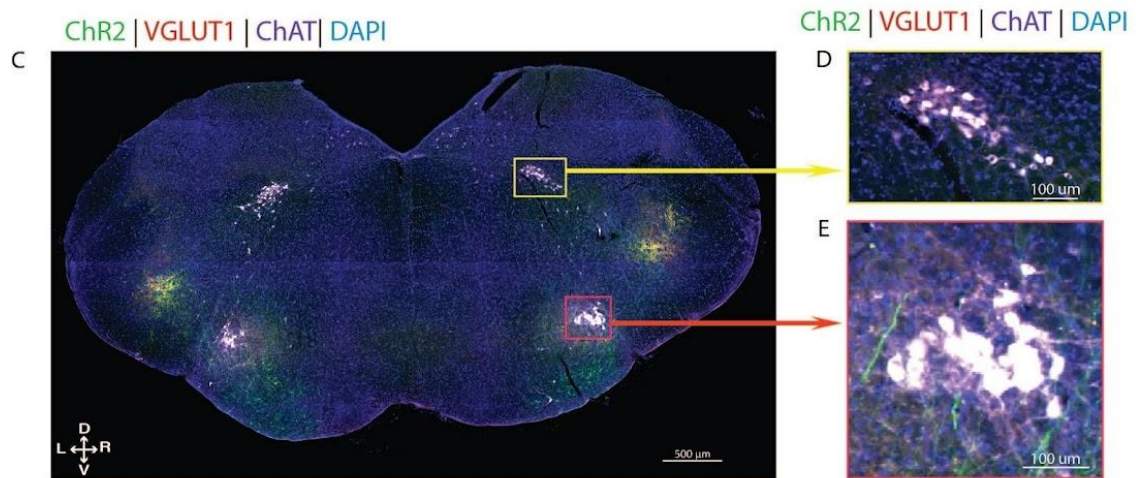
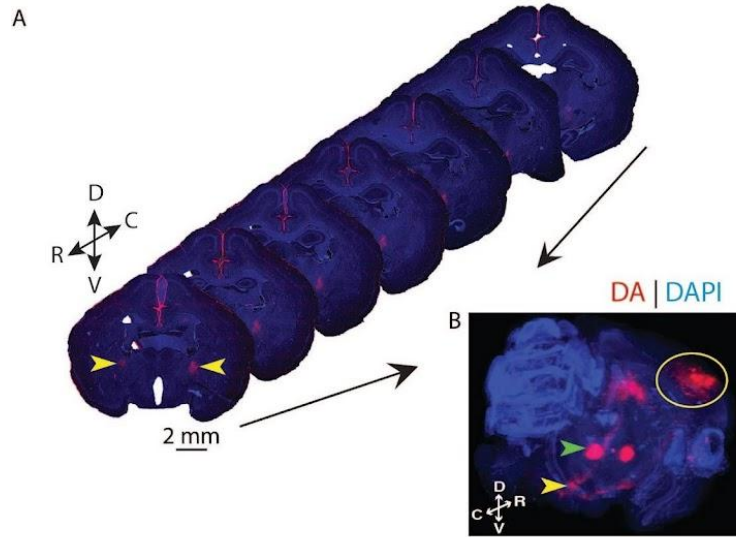
1884 Fig. S2. Example genes implicated by RERConverge include both CENPC (Conserved) (A) and GRM8  
 1885 (Accelerated) (B), which exhibit relative evolutionary rate (RER) shifts in vocal learning mammals (red)  
 1886 relative to vocal non-learners (black). Branch lengths illustrate the evolutionary rate of the protein-coding  
 1887 genes on the branches. Species with non-confident annotations of behavior are shown in gray.

1888  
1889  
1890  
1891  
1892  
1893  
1894



1895

1896 **Fig. S3.**  
1897 Comparison of genes implicated by RERconverge and HyPhy RELAX.  
1898 RERconverge was run for aligned proteins across the Zoonomia mammal dataset (see methods).  
1899 Genes were split between conserved (*purple*, RERconverge  $p < 0.01$ ; permutations  $p < 0.01$ ,  $\tau < 0$ ) and accelerated (*green*, RERconverge  $p < 0.01$ ; permutations  $p < 0.01$ ,  $\tau > 0$ ).  
1901 Additionally, a set of genes outside those groups with the same length distribution as genes from  
1902 vocal learners and non-learners was chosen as a control (*gray*). For each group, the density plot  
1903 shows the distribution of the HyPhy RELAX model K statistic, which will be positive if there is  
1904 more constraint within vocal learning clades and negative if there is more constraint outside  
1905 vocal learning clades.  
1906  
1907



1909 **Fig. S4.**  
1910 Reconstruction of anterograde and retrograde tracing from orofacial motor cortex (ofM1). **(A)**  
1911 Example images from one bat injected bilaterally with dextran amine tracer in ofM1 showing  
1912 seven sequential coronal planes separated by 240 $\mu\text{m}$ , aligned by hand and stacked into a 3D  
1913 volume shown in **(B)**. The anterograde propagation of the tracer (*red*) is visible in the pyramidal  
1914 tract (*yellow arrows*) against the background DAPI stain (*blue*). **(B)** 3D side view of the stacked  
1915 images showing the bilateral anterograde propagation of the tracer (*red*) projecting from ofM1  
1916 (*yellow circle*) down the pyramidal tract (*yellow arrow*) to the brainstem. Retrograde propagation  
1917 of the tracer shows cell bodies in the thalamus (*green arrow*) that send afferents into ofM1. See  
1918 Movie S1 for a 3D rotational view of the image stack. **(C-E)** Anterograde tracers are not found in  
1919 any other brainstem motor nuclei. **(C)** Coronal slice in brainstem following bilateral injections in  
1920 ofM1 with rAAV5/CamkII-hChR2(H134R)-EYFP (ChR2) labeling fibers in *green*. Synaptic  
1921 boutons were histologically labeled with VGLUT1 in *red*. Bilateral retrograde injections of CTB  
1922 (*white*) were delivered within the same bat into the cricothyroid muscles to label laryngeal  
1923 motoneurons in NA, and in nearby neck and tongue muscles, which enabled the labeling of cells  
1924 in the hypoglossal nucleus. **(D)** Magnification of hypoglossal nucleus with motoneurons  
1925 controlling the neck and tongue muscles (*white*). **(E)** Magnification of NA with laryngeal  
1926 motoneurons (*white*). Note the absence of cortical axons labeled with the anterograde tracer  
1927 within the hypoglossal nucleus or any other motor nuclei within the medulla except for NA. **(F)**  
1928 Example coronal section from Egyptian fruit bat medulla showing the decussation of the  
1929 pyramidal tract (*green arrowhead*) labeled by anterograde injections of fluorescent dextran  
1930 amine into ofM1. Note the decussation occurs at the level of the facial nucleus (*red arrowhead*),  
1931 rostral to the location of nucleus ambiguus. Motoneurons are labeled with choline  
1932 acetyltransferase (ChAT) in *green*.

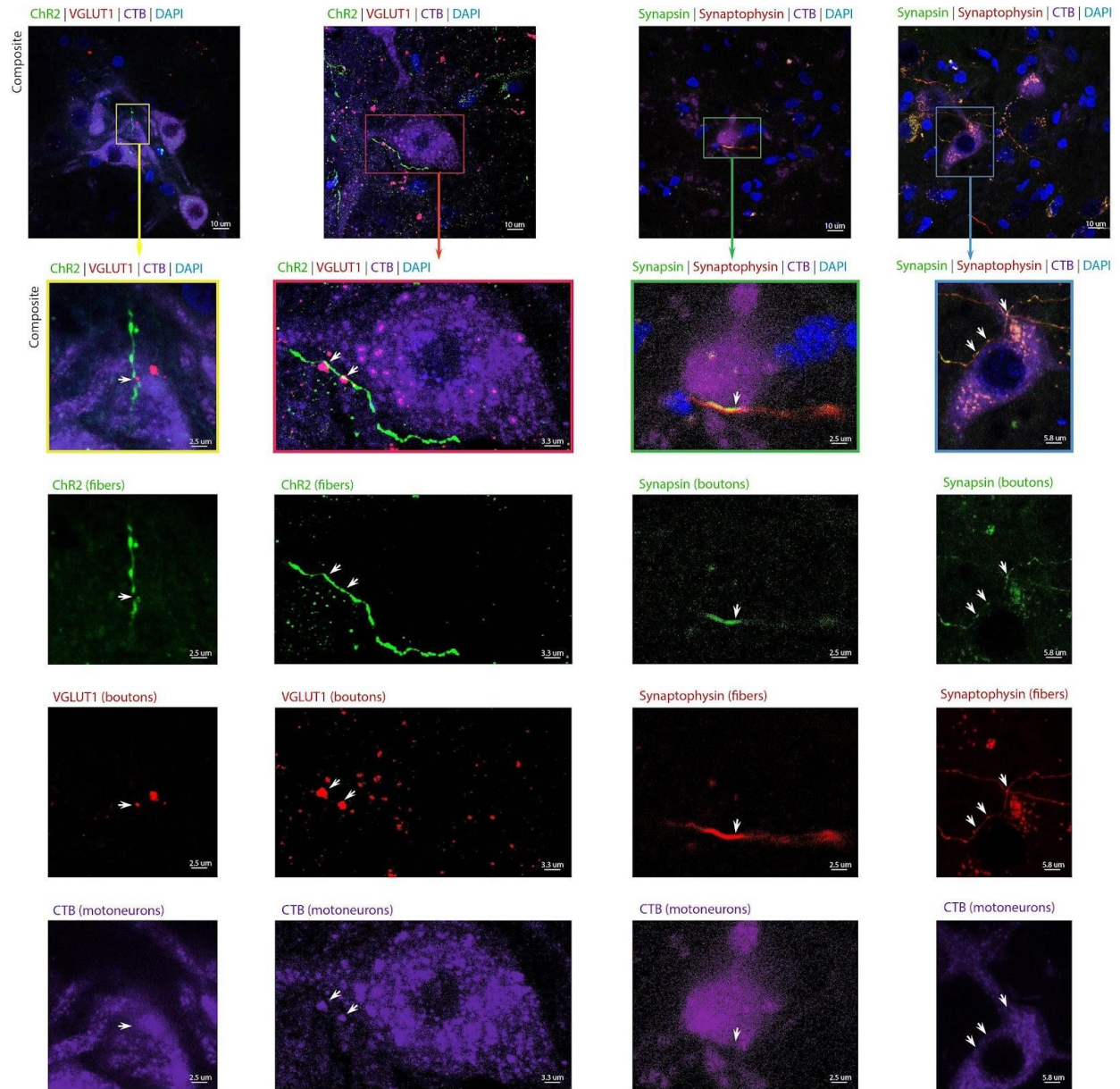
1933

1934

1935

1936





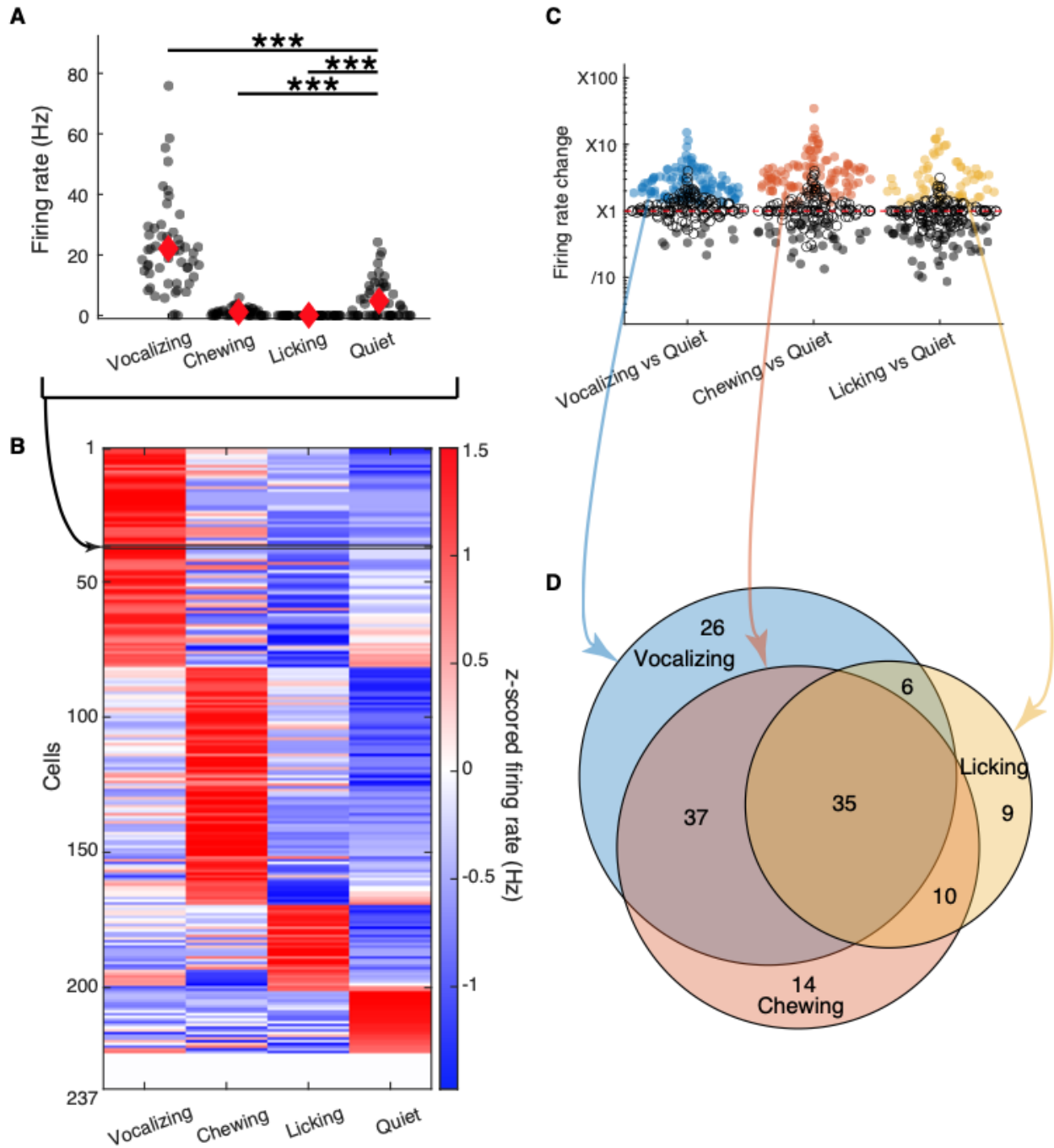
1937

1938

1939 **Fig. S5.** Examples of triple colocalization between corticobulbar fibers, synaptic boutons, and laryngeal  
 1940 motoneurons.

1941 (First 2 columns) Example confocal images of nucleus ambiguus (NA) showing triple  
 1942 colocalization (*white arrows*) between corticobulbar axons from ofM1 labeled with  
 1943 rAAV5/CamkII-hChR2(H134R)-EYFP (ChR2, *green*), CTB-labeled cricothyroid motoneurons  
 1944 (*purple*), and VGLUT1-labeled presynaptic boutons (*red*) with a background DAPI stain (*blue*).

1945 (Last 2 columns) Example confocal image of NA showing triple colocalization (*white arrows*)  
 1946 between corticobulbar axons from ofM1 labeled with AAV<sub>DJ</sub>-hsyn-mRuby2-T2S-Synap-eGFP  
 1947 (*fibers in red, boutons in green*), and CTB-labeled cricothyroid motoneurons (*purple*), with a  
 1948 background DAPI stain (*blue*). Across all columns, the first two rows are composite images, with  
 1949 the second row being insets of the first row. The three last rows depict three of the four  
 1950 individual channels merged to obtain the insets of the second row.



1952

1953

1954

1955

1956

1957

1958

1959

1960

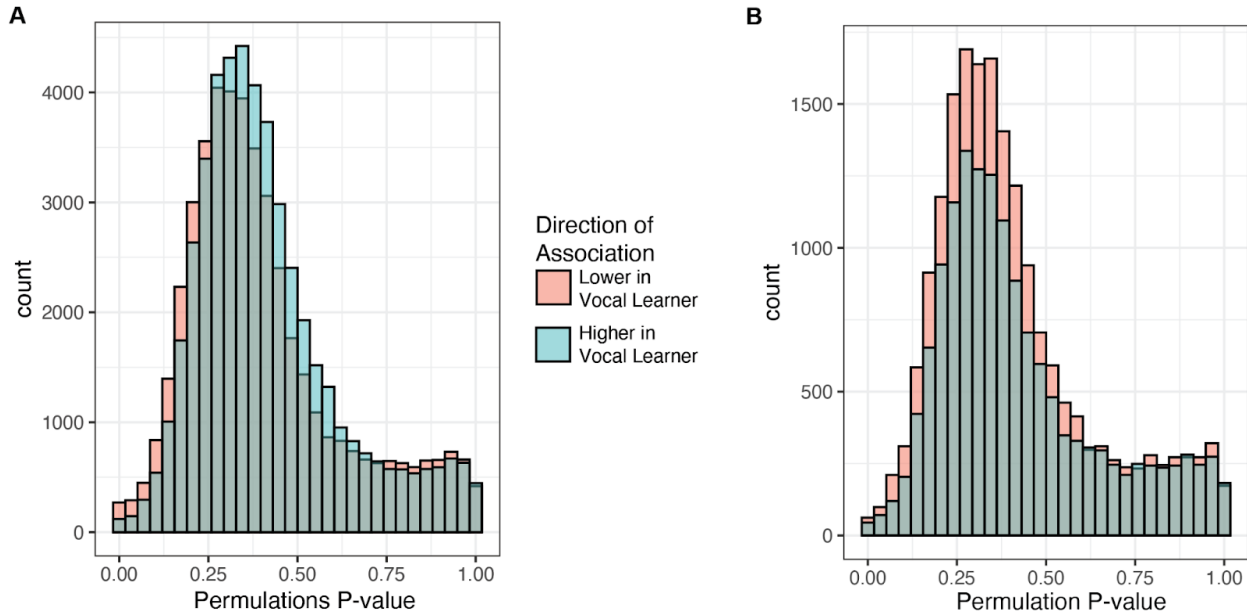
**Fig. S6.**

ofM1 neurons activity is modulated by vocalizing, chewing and licking.

(A) Time average firing rate for an example ofM1 neuron during vocalization production (n=53 vocalizations) and duration matched period of time when the bat was chewing, licking, or quiet and immobile. Black dots are individual events, red diamonds are mean +/- SEM. The neuron was significantly excited during vocal production as compared to quiet and significantly inhibited during chewing and licking as compared to quiet (Anova on Poisson GLM, all  $p < 0.001$ ). (B) Z-scored firing rate of ofM1 neurons, the activity of which could be evaluated

1961 during vocalizing, chewing, licking, and quiet ( $n=237$ ). **(C)** Change in firing rate during orofacial  
1962 motor actions as compared to quiet for ofM1 neurons. Each circle represents a neuron. The  
1963 significance of the Anova on Poisson GLM is signified by filled circles ( $p<0.001$  after FDR  
1964 correction;  $n=237$ ). As expected, neurons in the orofacial motor cortex are modulated during  
1965 orofacial motor actions as compared to quiet periods **(D)** Venn Diagram displaying the number  
1966 of neurons that were excited (significant increase in firing rate as compared to quiet; Anova on  
1967 Poisson GLM and FDR corrected  $p$ -value  $< 0.001$ ) by vocalizing, chewing, or licking ( $n=137$ ).  
1968 26 neurons were excited by vocal production only and not by chewing or licking.

1969



1970

1971

**Fig. S7.**

1972

TACIT identified stronger associations between regions that are predicted to have lower

1973

chromatin activity and the vocal learning ability of species. Histograms of TACIT permutation p-

1974

values are plotted separately for OCRs whose chromatin is predicted to be less open (*red*) vs.

1975

more open (*blue*) in vocal learners as opposed to vocal non-learners. Results are plotted

1976

separately for motor cortex (**A**) and parvalbumin positive inhibitory interneuron (**B**).

1977

OCR Ortholog Coordinates (mm10)	TACIT P-Value	TACIT Coef.	Permu- lations P-Value	Permu- lations BH Adj. P-Value	Gene Name	RER Conv. Result	Additional Genes
chr15:16107524-16107960	1.71E-02	-10.32	1.40E-05	0.05	Cdh9 (138)	No	-
chr18:65510387-65511111	3.99E-02	-9.46	3.70E-05	0.09	Zfp532	No	Malt1
chr2:131650717-131651399	1.44E-02	-10.89	3.70E-05	0.09	Adra1d	No	Smox
chr9:8176928-8177666	1.56E-02	-9.66	4.20E-05	0.09	Cep126	No	
chr18:68703945-68704585	1.46E-02	-13.68	3.50E-05	0.09	4930546 C10Rik	No	Tcf4 (139),
chr13:48323987-48324750	2.84E-02	15.06	2.90E-05	0.08	Id4	No	
chr10:5735065-5735721	2.73E-02	-10.25	2.50E-05	0.07	Fbxo5	No	<b>VIP</b>
chr19:56747655-56748356	1.70E-02	-9.31	1.10E-05	0.05	Adrb1	No	
chr18:15238033-15239015	2.36E-02	-9.96	6.00E-06	0.04	Kctd1	No	AQP4
chr16:65629047-65629933	2.51E-02	-10.04	9.00E-06	0.04	Chmp2b (140)	No	VGLL3
chr17:94243749-94244469	1.35E-02	-7.06	2.00E-05	0.07	Mettl4	No	
chr8:60393577-60394004	2.54E-02	-9.83	2.60E-05	0.07	Gm10283	No	
chr2:168519584-168519969	3.82E-02	-7.50	3.00E-06	0.04	Nfatc2	No	
chr17:72819735-72820272	6.78E-06	-5.90	4.40E-05	0.10	Ypel5	No	ALK
chr7:37573064-37573594	4.61E-03	-11.53	1.00E-06	0.02	Zfp536	No	Tshz3 (141-143)
	1.33E-02	-12.10	9.00E-06	0.04	NA	No	ZNHIT6, CCN1

	1.10E-02	10.44	2.80E-05	0.08	NA	No	
chr12:75463900-75464505	3.10E-02	12.42	9.00E-06	0.04	Gphb5	No	<u>PPP2R5E</u> (144)
chr2:145057020-145057789	2.14E-02	-5.83	2.00E-06	0.03	Slc24a3 (145)	No	Dtd1
chr15:61299599-61300306	1.91E-02	-10.06	4.50E-05	0.10	A1bg	No	Myc
chr18:25702100-25702785	1.86E-03	-6.47	1.00E-06	0.02	Celf4 (76)	p<0.01	
chr8:100212225-100213081	1.44E-02	-14.07	1.50E-05	0.05	<b>Cdh8</b> (146)	p<0.1	
chr13:26914286-26915174	1.77E-02	-11.20	1.50E-05	0.05	Prl	No	HDGFL1
chr7:96710377-96711352	2.92E-02	-9.48	3.70E-05	0.09	Tenm4	No	Nars
	3.39E-02	-6.45	1.00E-06	0.02	NA	No	ZNF704
chr14:28101682-28102178	2.39E-02	-10.64	3.20E-05	0.08	Erc2	No	Wnt5a
chr2:61494077-61494853	2.69E-03	15.75	1.00E-06	0.02	Tank (147)	No	Tbr1 (148, 149)
chr17:63914921-63915759	2.01E-02	-11.71	1.40E-05	0.05	Fer	No	
chr12:97833162-97833781	5.23E-03	-14.31	1.50E-05	0.05	Galc (150)	p<0.01	
chr2:33543268-33544008	2.30E-02	13.70	1.00E-05	0.05	Zbtb43	No	<u>Lmx1b</u> (151)
chr8:18790099-18790817	4.64E-02	-10.55	2.50E-05	0.07	Angpt2	No	
	3.85E-02	-9.41	4.00E-06	0.04	NA	No	SORCS3
chr5:20750587-20751246	1.33E-02	-10.20	5.00E-06	0.04	Phtf2	No	<b>Magi2</b> (152)
	5.68E-03	10.75	8.00E-06	0.04	NA	No	DPPA4
chr12:86795973-86796457	3.78E-07	5.92	5.00E-06	0.04	Lrrc74a	No	
	2.48E-03	6.51	8.00E-06	0.04	NA	p<0.1	Galc (150)

chr17:86429786-86430568	6.81E-03	-11.21	2.50E-05	0.07	<b>Prkce</b> (153)	No	Epas1
chr1:21773189-21773501	1.27E-03	6.73	2.00E-05	0.07	<b>Kcnq5</b> (154)	No	
chrX:87178627-87179266	3.10E-02	5.49	2.00E-06	0.03	Il1rapl1 (155)	No	
	5.14E-05	7.45	2.00E-06	0.03	NA	p<0.01	<b>DAAM1</b> (156), <b>DACT1</b> (157)
chr12:67321023-67321586	5.17E-02	-15.62	1.70E-05	0.06	Mdga2 (158)	No	
chr17:86352256-86352626	7.58E-02	-6.33	5.00E-06	0.04	<b>Prkce</b> (153)	No	Epas1
chr5:18822320-18822911	1.62E-01	-3.66	7.00E-06	0.04	<b>Magi2</b> (152)	No	
chr8:35222316-35223646	1.23E-01	-10.39	3.40E-05	0.09	Ppp1r3b	No	

1979 **Table S1. M1 OCRs whose predicted open chromatin in boreoeutherian mammals is**  
1980 **significantly associated with vocal learning.**

1981 Coordinates are reported for the mouse genome (mm10). The coefficients shown are those  
1982 reported by phyloglm for the association between the vocal learning trait and the OCR orthologs'  
1983 open chromatin prediction. The permutations p-values represent the adjusted p-values after  
1984 Benjamini-Hochberg correction. Nearest genes are indicated, with genes previously associated  
1985 (in some cases specifically, in other cases as part of larger deletions) with speech delay or  
1986 disability underlined (with associated references) and genes previously shown to be convergently  
1987 regulated in humans and song-learning birds (4) in bold red. Genes that have vocal learning  
1988 associated specialization in the songbird brain are in bold black (137). A summary of the  
1989 RERConvege unadjusted p-value for the nearest annotated mouse gene is provided. Human  
1990 coordinates in the hg38 assembly are used when a mouse ortholog does not exist.

1991

OCR Ortholog Coordinates (mm10)	TACIT P-Value	TACIT Coef.	Permutations P-Value	Perm. BH Adj. P-Value	Gene Name	RER Converge Result	Additional Genes
chr13:45770571.45771071	2.32E-02	-8.44	1.40E-05	0.08	<b>Atxn1</b> (159, 160)	No	GMPR
chr2:42511067.42511567	1.26E-03	-14.22	6.00E-06	0.05	Lrp1b (161)	No	
chr2:56025027.56025527	3.21E-02	-7.40	3.00E-06	0.03	Kcnj3 (162)	No	
chr2:148111903.148112403	5.08E-04	-14.11	1.00E-06	0.03	Foxa2	No	
chr6:61939127.61939627	5.49E-04	-14.67	2.00E-06	0.03	Ccser1	p<0.1	<b>Snca</b> (163, 164)
chr8:112642458.112642958	2.45E-02	7.48	1.20E-05	0.08	<b>Cntnap4</b> (81)	No	Mon1b

1992 **Table S2. PV OCRs whose predicted open chromatin state in boreoeutherian mammals is**  
 1993 **significantly associated with vocal learning.**

1994 Coordinates are provided for the mouse (mm10) OCR orthologs; otherwise this table is in the same  
 1995 format as Table S1.



1996 **Movie S1.**  
1997 **3D rotational view of fluorescent dextran amine bilaterally injected in ofM1.** See legend of  
1998 Figure S4A-B for details. The anterograde propagation of the fluorescent tracer dextran amin  
1999 (*red*) is visible from the injection point to the pyramidal tract and down to the medulla. A  
2000 background DAPI stain (blue) is used to better visualize the brain. Note, that the decussation of  
2001 the pyramidal tract, as further depicted in Fig. S4F, is rostral to the location of NA.  
2002  
2003 **DataS1. (separate file: speciesAnnotations.4.1.7.1.csv)**  
2004 Table of vocal learning annotations across mammals considered in this study. Species are  
2005 referenced based on their ID in the TOGA alignment, propername, and NCBI accession ID.  
2006 **DataS2. (separate file: vocalLearningRERConvergeMaster.csv)**  
2007 This file contains the output from running RERconverge for the vocal learning trait on the  
2008 Zoonomia genomes. The “Rho” column corresponds to a rank sum test between vocal learners  
2009 and non-learners. Positive values have higher evolutionary rates in vocal learning species. The  
2010 “permP” column includes p-values after conducting permutations. Both the RERconverge and  
2011 permutations p-values were adjusted using Benjamini Hochberg. The final four columns  
2012 correspond to the Bayes Factor calculated for that group of species. Bayes Factors > 5 are  
2013 considered a strong signal of selection.  
2014 **DataS3. (separate file: HyPhyResults.1.xlsx)**  
2015 This file contains the output from running HyPhy on selected set of vocal learning genes.  
2016 1 ID: gene  
2017 2 Sequences: number of sequences in the alignment  
2018 3 Sites: length of the alignment  
2019 4 FG q-value: q-value (Benjamini-Hochberg) of the BUSTED test on foreground branches  
2020 5 BG q-value: q-value (Benjamini-Hochberg) of the BUSTED test on background branches  
2021 6 DIFF q-value: q-value (Benjamini-Hochberg) for the difference between FG and BG  
2022 7 S.Sites (FG): Sites that are showing support for positive selection in the Foreground branches  
2023 with evidence ratio of 100 or higher  
2024 8 S.Sites (BG): Sites that are showing support for positive selection in the Background branches  
2025 with evidence ratio of 100 or higher  
2026 9 L (FG): tree length foreground  
2027 10 L (BG): tree length background  
2028 11 RELAX q-value  
2029 12 RELAX K: the foreground is modeled as  $(\omega)^K$  so it matters if  $K < 1$  or  $K > 1$   
2030 13 -18 Foreground parameters for BUSTED-PH  
2031 19 PH: Selection associated with phenotype  
2032 20 TYPE: Binary summary of 4,5,6.

2033

2034 **DataS4. (separate file: GeneOntology.1.xlsx)**

2035 This file contains the result of running gene set enrichment analysis using ENRICHR. Different  
2036 tabs correspond to different categories (GO Biological Process; Human Phenotype) and different  
2037 gene sets (Conserved and Accelerated). The input genes are the intersection of significant genes  
2038 from RERconverge and HyPhy.

2039

2040

2041 **DataS5. (separate file: BatM1OcrS.2.xlsx)**

2042 The file contains a table of 348 open chromatin regions (OCRs) discovered from ATAC-seq  
2043 analyses with differential activity between the bat orofacial and wing subregions of primary  
2044 motor cortex (ofM1 and wM1, respectively). The first four columns provide locations in the  
2045 *Rousettus aegyptiacus* genome assembly (mRouAeg1.p), with the fourth column presenting the  
2046 relative position of the ATAC peak summit in the OCR. Columns 5-9 provide output of DEseq2  
2047 analysis (128). Proximal genes were identified from the mRouAeg1.p assembly gene annotations  
2048 (124). The full list of TACIT vocal learning associated peaks (intersections from which are  
2049 provided in column 10) can be found in Tables S1 and S2. The full list of genes under RER  
2050 convergence in vocal learning mammals (intersections from which are provided in column 11)  
2051 can be found in Data S1.

2052

2053 **DataS6. (separate file: BatM1OcrGo.1.xlsx)**

2054 This file contains the results of the gene functional enrichment analyses performed using  
2055 GREAT on the OCRs differential between ofM1 and wM1 (provided in Data S3) mapped to the  
2056 human genome assembly (hg38), prepared as described in the Materials and Methods section  
2057 “Gene functional enrichment analyses.” The file contains two tabs, one for the enriched terms  
2058 from GO Biological Processes (Data\_S4A\_GOBiologicalProcess) and the other for those from  
2059 the set of GO Molecular Functions (Data\_S4B\_GOMolecularFunction).

2060

2061 **DataS7. (separate file: VocalLearningTACITResults.1.xlsx)**

2062 The results of running TACIT for vocal learning on motor cortex and PV+ inhibitory  
2063 interneurons. The “Exp\_Pvalue” is calculated using permutations and then adjusted using  
2064 Benjamini Hochberg (bh column). The number of permutations is the “Trials” column.

2065

2066 **DataS8. (separate file: TACIT\_ttest.zip)**

2067 This archive contains a series of .csv files that hold the results of t-test applied across motor  
2068 cortex (mcx) and PV+ interneuron (pv) open chromatin. Each vocal learning clade is calculated  
2069 separately.

2070

2071 **DataS9. (separate file: OCRMouseGeneAnnotation.1.xlsx)**

2072 This file contains a subset of results from Data S7 that pass the significance threshold. The  
2073 HOMER pipeline was used to annotate the nearest mouse gene. Coordinates are reported for the  
2074 mouse genome (mm10). The coefficients shown are those reported by phyloglm for the  
2075 association between the vocal learning trait and the OCR orthologs' open chromatin prediction.  
2076 The permutations p-values represent the adjusted p-values after Benjamini-Hochberg correction.  
2077 Nearest genes are indicated, with genes previously associated (in some cases specifically, in  
2078 other cases as part of larger deletions) with speech delay or disability underlined (with associated  
2079 references) and genes previously shown to be convergently regulated in humans and song-  
2080 learning birds (4) in bold red. Genes that have vocal learning associated specialization in the  
2081 songbird brain are in bold (137). A summary of the RERConvege pvalue for the nearest  
2082 annotated mouse gene is provided. Human coordinates are used when a mouse ortholog does not  
2083 exist.

2084

2085 **DataS10. (separate file: CellTypeEnrichments.1.xlsx)**

2086 This file contains a subset of results of exploring cell type-specific open chromatin intersecting  
2087 peaks of interest from the bulk result. The number of overlapping peaks is given separately for  
2088 the sets that have higher or lower predicted open chromatin in vocal learners (vl\_down\_num,  
2089 vl\_up\_num). A permutation test is used to identify the significance of the overlap. The  
2090 permutation p-value is provided along with the log2 of the enrichment over the expected value  
2091 from the permutation test.

2092

2093 **References and Notes:**

2094

2095

2096

2097

2098