1 **An adaptive behavioral control motif mediated by cortical axo-axonic inhibition**

2

3 Kanghoon Jung[1,2,5], Minhyeok Chang[1,6], André Steinecke[2,6], Benjamin Berke[1], Youngjin Choi[1],

4 Yasuhiro Oisi[2], David Fitzpatrick[2], Hiroki Taniguchi[2,3], and Hyung-Bae Kwon[1,2,4*]

5

6 [1]Department of Neuroscience, Johns Hopkins University School of Medicine, Baltimore,

7 Maryland 21205, USA

8 [2]Max Planck Florida Institute for Neuroscience, Jupiter, Florida 33458, USA

9 [3]Department of Pathology, Chronic Brain Injury program, Ohio State University, Columbus, Ohio

10 43210, USA

11 [4]Max Planck Institute of Neurobiology, Martinsried 82152, Germany

12 [5]Current address: Allen Institute for Neural Dynamics, Seattle, WA 98109, USA

13 [6]equally contributed

14

15 *Corresponding author: Hyung-Bae Kwon

16 The Solomon H Snyder Department of Neuroscience

17 Johns Hopkins University School of Medicine

18 733 N Broadway, Baltimore, MD 21205

19 hkwon29@jhmi.edu

20

21   **Abstract**

22   Neural circuits are reorganized with specificity during learning. Genetically-defined subgroups of

23   inhibitory interneurons are thought to play distinct roles in learning, but heterogeneity within

24   these subgroups has limited our understanding of the scope and nature of their specific

25   contributions to learning. Here we reveal that the chandelier cell (ChC), an interneuron type that

26   specializes in inhibiting the axon-initial segment (AIS) of pyramidal neurons, establishes cortical

27   microcircuits for organizing neural coding through selective axo-axonic synaptic plasticity. We

28   find that organized motor control is mediated by enhanced population coding of direction-tuned

29   premotor neurons, whose tuning is refined through suppression of irrelevant neuronal activity.

30   ChCs are required for learning-dependent refinements via providing selective inhibitory control

31   over pyramidal neurons rather than global suppression. Quantitative analysis on structural

32   plasticity of axo-axonic synapses revealed that ChCs redistributed inhibitory weights to

33   individual pyramidal neurons during learning. These results demonstrate an adaptive logic of the

34   inhibitory circuit motif responsible for organizing distributed neural representations. Thus, ChCs

35   permit efficient cortical computation in a target cell specific manner, which highlights the

36   significance of interneuron diversity.

37

38

39

Introduction

The reorganization of neural circuits during learning relies on a complex interplay between excitatory and inhibitory signals in the brain, accompanying highly specific synaptic modification. Various synaptic changes contributing to this interplay have been observed at the scale of individual cells and whole circuits. The observed formation and elimination of dendritic spines in cortical pyramidal neurons (PyNs) during learning have suggested that learning is associated with changes in neuronal connectivity [1–4]. At the circuit level, neuronal activity is synchronized in a subpopulation of neurons, resulting in a high correlation between neuronal activity patterns and learned behaviors [3,5,6]. It has been understood that GABAergic inhibition plays a critical role in shaping learning-dependent circuit changes. Interneurons are categorized by anatomical and electrophysiological features, and their distinct functions during learning have been shown in genetically-defined subtypes such as parvalbumin (PV)-, somatostatin (SOM)- and vasoactive intestinal polypeptide (VIP)-expressing neurons [7–11]. However, these subtypes still contain a remarkable degree of heterogeneity, such that the precise role of inhibition mediated by single-type interneurons remains an open question.

The ChC (i.e., 'axo-axonic cell') is a *bona fide* GABAergic interneuron subclass with respect to its distinct axonal geometry, subcellular synapse connectivity, and fast-spiking electrophysiological properties [12,13]. A single ChC exhibits a characteristic axonal geometry with many prominent vertical branches, which contain strings of synaptic boutons exclusively aligned along the AISs of neighboring PyNs. Because the AIS is the site of action potential initiation, ChCs can have decisive control over spike generation in an ensemble of PyNs, thereby regulating synchrony and oscillation of network activity. These functions are thought to be critical for higher-order cognitive processes like working memory [14–16]. In this study, using a transgenic mouse line specifically engineered to target ChCs, we examined the role of ChCs in sculpting cortical circuits involved in motor learning. We discovered that the ChC is essential for improving the direction-tuning of targeted cells by suppressing irrelevant activity at the population level. ChC shaped this sophisticated inhibitory motif among a specific set of neuron types. Thus, ChCs primarily take a "select and focus" strategy rather than serve as a uniform gain setter.

69

Results

Improved directional control of movement on multi-textured ball maze.

Previous studies have shown that ChCs increased activity during locomotion [17,18]. The premotor cortex (M2, anteromedial agranular cortex, homologous to supplementary motor areas in the primate brain) is involved in gating sensory inputs and motor outputs for movement planning, fine action control, and decision-making [19–21]. To investigate the role of ChCs in controlling the activity of M2 neurons during locomotion, we used a spatial navigation task that requires goal-directed motor control [22]. Mice were trained to traverse a multi-textured ball maze with four distinctive tactile surface cues and to reach a goal spot on the ball, where a water reward was given (**Fig. 1a**, See Methods). The task required directional motor control accurately aimed to the goal spot. While animals were navigating the ball maze, we recorded neuronal activity of the L2/3 premotor population using Ca^{2+} imaging via two-photon microscopy. Mice with water-restriction (experimental group) showed reward-seeking behavior on the ball maze, developed a preferred pathway, and generated goal-oriented locomotion with learning whereas mice *ad libitum* water (control group) did not (**Fig. 1b**). The number of successfully reaching the goal spot increased in the experimental group but not in the control group (**Fig. 1e**). Time spent in the goal (rewarded) quadrant (Q1) increased with training (**Fig. 1f**). We further examined the relationship between motor control and reward-seeking behavioral performance by analyzing the parameters of motor control. To examine whether general movement parameters such as speed and acceleration could be the main cause for these improvements, we compared the motor control between experimental and control conditions. The movement of the control group showed similar general motor parameters (e.g., speed and acceleration) to the experimental group (**Fig. 1c, d**). Overall latency to the reward decreased with training, indicating the increased effectiveness of the animals' movements at achieving rewards (**Fig. 1g**). We observed that mice showed increased proximity to the goal with learning, by tuning their movement to the goal (**Fig. 1h**). The movement accuracy as the percentage of goal-aimed movements increased with training (**Fig. 1i**), indicating that the enhanced movement direction (MD) controls toward the goal spot. Mice in the experimental group showed better control in

98 turning toward the goal during navigation than the control (**Fig. 1j, k**). These results indicate that

99 the directional control of movement is improved with learning.

100

101 **Motor learning refines direction coding through suppression of irrelevant activity.**

102 To determine how neurons mediate precise motor control with learning, we examined direction

103 coding of M2 neurons at the cellular and population levels. To test whether M2 is required for

104 neuronal adaptation to control task-related motor behavior, we used an adeno-associated virus

105 (AAV)-based approach to express tetanus toxin light chain (TeTxLC), which abolishes

106 presynaptic vesicle release [23,24], in M2 excitatory neurons (**Extended Data Fig. 1a**). Mice

107 expressing TeTxLC in M2 excitatory neurons (CaMKII-TeTxLC mice) showed deficits in learning

108 the task (**Extended Data Fig. 1e**). The reward-seeking behavioral parameters did not improve

109 with training (**Extended Data Fig. 1e-j**), suggesting that M2 neurons play an essential role in

110 mediating the motor learning required for the task.

111 We performed two-photon Ca^{2+} imaging in L2/3 of M2 with a genetically encoded

112 calcium indicator, GCaMP6s while mice navigated on the ball maze (**Extended Data Fig. 2a**).

113 Consistent with the established role of M2 in motor planning, overall population activity in L2/3

114 correlated with locomotion (**Extended Data Fig. 2b**). Individual premotor neuron activities bi-

115 directionally changed in a time-locked fashion at movement onset (**Extended Data Fig. 2c**).

116 The majority of neurons were movement-related (total 98.0%; positively correlated: 79.7 ± 3.3%;

117 negatively correlated: 18.4 ± 3.2%; and not correlated: 2.0 ± 0.4%; see Methods for

118 classification), and the proportion of movement-related neurons in the population remained

119 relatively constant throughout training sessions (**Extended Data Fig. 2d**).

120 We next measured direction selectivity of individual premotor neurons over training and

121 analyzed the changes. Distinct neuronal populations responded to different MDs (MDs) (**Fig. 2a,**

122 **Extended Data Fig. 2e-g**). A subset of L2/3 neurons exhibited tuning to the MD which was

123 defined as its preferred direction (PD). Its topological map was arranged in a salt-and-pepper

124 pattern in the M2 area (**Extended Data Fig. 2i, j**).

125 We estimated the likelihood distributions of neural tuning curves as a function of MD

126 (**Fig. 2b)**. Premotor neurons showed the highest likelihood to fire at their PD. We defined

127 population tuning precision as the reciprocal of the variance of total neural activity given MD.

128 Tuning of premotor neurons to MD became more precise over sessions. The precision of

129 population responses increased for all MDs in the experimental group, but not in the control

130 group (**Fig. 2c-d**). We next estimated the probability of the cell being active as a function of the

131 distance between MD and the cell's PD. When the MD was far from the cell's PD (distance from

132 PD > 90°), the probability of being active decreased with learning (**Fig. 2e**). When the MD was

133 opposite to the cell's PD, the probability of the cell being active was significantly lower in

134 Session 7 than in Session 1 for mice in the experimental condition, but not in the control

135 condition (**Fig. 2f**).

136   We took a Bayesian approach to decode the MD from neural activity (See Methods).

137 We computed the posterior probability density function to provide the most likely estimate of the

138 mouse's MD during navigation (**Fig. 2g**, **Extended Data Fig. 3a-c**). Moment-to-moment MDs

139 were decoded by the estimated MD corresponding to the maximum value of the posterior

140 probability density function. During behavior, posterior probability density functions of MD

141 sharpened over sessions in the experimental group, enhancing near the MD and waning at null

142 MDs (**Fig. 2h**). In contrast, posterior functions were flattened over sessions in the control group.

143 Consequently, we observed lower decoding errors in the experimental group (**Fig. 2i**). Similarly,

144 we tested neural coding of MD based on a population vector (PoV) as a vector sum of PDs of a

145 sparse population of active direction-tuned neurons at a given moment [25,26] (**Extended Data**

146 **Fig. 3d**). In this coding scheme, MDs were also flexibly encoded in a moment-to-moment

147 manner from the distinct sparsely distributed activity of neuronal ensembles (**Extended Data**

148 **Fig. 3e, f**). The fraction of decoded direction (PoV around 0°) increased and that of anti-

149 decoded direction (PoV around 180°) decreased in Session 7 in the experimental group, but not

150 in the control group, indicating an increased coding accuracy associated with learning

151 (**Extended Data Fig. 3g-i**). We further measured angular errors between the active neuron's

152 PD and MD to see how directional neuronal activity was accurately aligned with MD on a

153 moment-to-moment basis. Similar to the change in decoding error, the fraction of PD increased

154 and that of anti-PD decreased in Session 7 in the experimental mice, whereas the fraction of

155 anti-PD increased in Session 7 in the control mice, indicating that the fraction of irrelevant

156 activity reduced with learning (**Extended Data Fig. 3j, k**). To test whether the motor learning

157   mediates a direction-dependent change in network connectivity, we compared pairwise

158   correlations between neuronal pairs of diametrically opposed PDs. The pairwise correlations of

159   neuronal pairs decreased in the experimental group as a function of the difference in their PDs,

160   whereas they increased over sessions in the control group (**Fig. 2j**). These results suggest that

161   the observed reduction of incorrectly tuned responses shapes the direction coding of the M2

162   population.

163

164   **Manipulation of parvalbumin interneurons disrupts global neural activity.**

165   To determine how neuronal ensembles suppress irrelevant activity during learning, we asked

166   how interneurons might be uniquely involved. Several GABAergic interneuron subtypes have

167   been hypothesized to serve specific roles in regulating cortical functions by forming circuit motifs

168   with features like recurrence and feedforward inhibition [27]. We tested the regulative role of two

169   major interneuron groups that are known to form inhibitory synapses at specific compartments

170   of excitatory PyNs and comprise 70% of all cortical interneurons [28]; Specifically, PV-INs inhibit

171   perisomatic regions and SOM-INs target distal dendrites [29]. We bilaterally silenced activity of

172   PV-INs or SOM-INs during behavior by expressing an inhibitory opsin, eNpHR3.0, in the M2.

173   We found that inactivation of premotor PV-INs disrupted goal-directed motor performance, but

174   inactivation of SOM-INs did not, suggesting that perisomatic inhibition mediated by PV-INs may

175   be involved in circuit motifs that guide task-related directional control of movement (**Fig. 3a,**

176   **Extended Data Fig. 4**).

177   Perisomatic inhibition from PV-INs provides powerful inhibitory control over the principal

178   neuron population [29–32]. To further examine the role of PV-IN perisomatic inhibition in local

179   neuronal activity during motor control, we abolished presynaptic GABA release from PV-INs by

180   expressing Cre-dependent TeTxLC virus in PV-Cre mice (PV-TeTxLC mice) and monitored the

181   M2 neuron activity (**Fig. 3b**). Consistent with the notion that PV-INs are critical regulators of

182   cortical excitatory and inhibitory (E/I) balance [33,34], the abolishment of perisomatic inhibition in

183   PV-TeTxLC mice led to excessively increased overall activity of M2 neurons during movement

184   (**Fig. 3c-d**) [35] and hypersynchronization at movement initiation, which disrupted the propagation

185   of sequential network activity and decreased the sparseness of neural activity during

186 subsequent movements (**Fig. 3e, f**). Accordingly, the pairwise correlation and the proportion of

187 pairs with positive correlations were significantly higher in PV-TeTxLC mice than in PV control

188 mice (**Fig. 3g-h**). A significant proportion of PV-INs also showed increased activity for

189 movement initiation (**Fig. 3j**), suggesting that the excessive activity of M2 neurons in PV-

190 TeTxLC mice may result from unmasked responses of neurons that were previously inhibited by

191 PV-INs (**Fig. 3k-n**). These results suggest that PV-IN perisomatic inhibition is involved in task-

192 related motor control, but this manipulation also leads to global excitability changes that limits

193 the applicability of current approaches to specify mechanisms of directional motor control.

194

195 **Sparse, sequential activity is preserved in ChC manipulation.**

196 In our attempt to specify the underlying mechanisms of perisomatic inhibition, we targeted ChCs

197 by using the ChC-specific transgenic mouse line (Nkx2.1-2a-CreER::Ai14 as ChC control mice,

198 Nkx2.1-2a-CreER::Flex-FlpO mice as ChC-Flp mice, and Vipr2-Cre) [36]. This genetic strategy

199 allowed us to selectively label ChCs, including their distinctive axonal arborizations and

200 cartridge terminals (**Extended Data Fig. 5a-c**). A larger proportion of ChCs were found in the

201 upper L2/3 of M2, with a smaller number present in L5 (mean soma depth ± s.e.m = 178.9 ±

202 10.7 μm from pial surface) (**Extended Data Fig. 5d**). We confirmed that the axonal cartridges of

203 ChCs made synapses on the AISs of neighboring neurons (**Fig. 4a**). To find out if ChC

204 manipulation impacts global network excitability, we abolished presynaptic inhibitory GABAergic

205 inputs to the AISs of neighboring neurons by expressing FlpO-dependent TeTxLC (AAV9-CAG-

206 dfrt-TeTxLC-HA-WPRE) bilaterally in ChC-Flp mice and monitored activity of M2 neurons (**Fig.**

207 **4b**). In contrast to the hypersynchronous neuronal activity in PV-TeTxLC mice, we did not

208 observe a global change. Instead, we observed sparse, sequential activity propagation in both

209 ChC-TeTxLC and control mice during movement (**Fig. 4c, d**). Pairwise correlation indicated

210 local, not global, correlated activity between neuronal pairs in both groups (**Fig. 4e**); The

211 proportions of pairs with positive and negative correlations were similar between the two groups

212 (**Fig. 4f**). In contrast to PV-INs, a majority of ChCs showed increased activity during movement

213 rather than movement initiation (**Fig. 4g**), and ChC-TeTxLC manipulation did not evoke the

214 excessive activity of M2 neurons at movement onset shown in PV-TeTxLC mice (**Fig. 4h-k**).

215   These results demonstrate that ChCs have a distinct activity profile from PV-INs and that overall

216   network excitability is not globally altered by silencing the activity of ChCs.

217

218   **Chemogenetic inhibition of ChCs disrupts the direction selectivity of premotor neurons.**

219   The TeTxLC experiments suggest that the major role of ChCs is not control of general network

220   gain, but more target-specific control of neighboring neurons. We tested whether ChC

221   suppression would affect the direction tuning of M2 neurons during the behavioral task. We

222   bilaterally expressed FlpO-dependent chemogenetic silencer (AAV9-CAG-dfrt-hM4Di-mCherry)

223   in the M2 of ChC-Flp mice to suppress the activity of ChCs upon the application of clozapine N-

224   oxide (CNO), an hM4Di agonist (**Fig. 5a**). We first verified the effectiveness of CNO in brain

225   slices. Whole-cell electrophysiological recordings showed that CNO application reduced the

226   activity of ChCs expressing hM4Di (**Fig. 5b**). After 7 days of training, we administered either

227   saline or CNO. CNO administration in the ChC-Flp:hM4Di mice significantly impaired goal-

228   directed navigation performance and turning movement (**Fig. 5c-f**). CNO or saline injection

229   alone had no effect on behavioral performance (ChC-Flp:hM4Di mice with saline, ChC-Flp mice

230   with saline or CNO injection) (**Fig. 5d**).

231   To determine the role of ChC activity in the direction tuning of M2 neurons, we co-

232   expressed AAV9-CAG-dfrt-hM4Di-mCherry and AAV1-hSyn-GCaMP6s into the L2/3 M2. We

233   then used chemogenetic manipulations and performed Ca^{2+} imaging of ChCs and their

234   neighboring neurons during the behavioral task. In line with the *ex vivo* results, ChC activity was

235   suppressed upon CNO administration during locomotion (**Fig. 5g, h**). Consequently, we

236   observed that the coactivity of M2 neurons increased upon CNO application during locomotion

237   but not during rest (**Fig. 5i**), consistent with recent work demonstrating the inhibitory

238   postsynaptic effects of ChCs on M2 neurons *in vivo* [17,31]. We compared the likelihood

239   distributions of neural tuning curves as a function of MD with and without CNO administration.

240   We observed that neurons are more likely to be active at non-PDs in the CNO (ChC-silenced)

241   condition than in the saline condition (**Fig. 5j**), resulting in less precise population responses to

242   the MD (**Fig. 5k, l**). The probability of the cell being active at a certain MD increased upon CNO

243   administration (**Fig. 5m**). When the MD was at the opposite of the cell's PD, the probability of

244   the cell being active was higher after CNO administration than it had been in Session 7 or in the

245   saline condition (**Fig. 5n**). CNO application increased pairwise correlations between neuronal

246   pairs as opposed to the enhanced decorrelation between pairs seen with learning (**Fig. 5o**).

247   These results suggest that the suppression of ChC activity increased behaviorally irrelevant

248   neural activity, impairing the precise coding of movement.

249

250   **Variability of ChC activity increases during learning.**

251   To examine the activity dynamics of ChCs during learning, we expressed GCaMP6s in the L2/3

252   ChCs and recorded ChC activity while mice performed the behavioral task (**Fig. 6a**). Consistent

253   with previous reports [17,18], ChCs displayed increased activity during episodes of locomotion

254   (**Fig. 6b**), with a more synchronized pattern of activity in Session 1. In Session 7, activity

255   patterns became more diverse and desynchronous, which led to the formation of subclusters of

256   ChCs (**Fig. 6c**). To quantify this shift, we calculated the pairwise correlations between ChCs

257   during the behavioral task and found that the correlation decreased and ChCs formed

258   subclusters as learning progressed (**Fig. 6d, e**). We next examined how temporal relationships

259   between ChCs and M2 neurons change with learning. Various temporal relationships in activity

260   were shown between ChCs and neighboring M2 neurons (**Fig. 6f**). We calculated cell-to-cell

261   activity correlations (Pearsons' correlation) in pairs of ChCs and M2 neurons during locomotion

262   and rest epochs and compared their distributions of correlation coefficients. In Session 1, the

263   proportion of high correlation coefficient (corr(r) > 0.6) was decreased during locomotion

264   compared to rest, reflecting the increased decorrelation. In Session 7, the overall proportions of

265   positive correlation coefficient (corr(r) > 0) and of negative coefficient (corr(r) < 0) were

266   remarkably decreased during locomotion compared to rest, indicating the enhanced

267   decorrelation in gross pairs of ChC and M2 neuron (**Fig. 6g**). These results suggest that ChC

268   activity becomes more heterogeneous with learning and such decorrelated temporal relationship

269   between ChCs and neighboring M2 neurons implies an increased specificity of population

270   activity control.

271

272 **Abolition of ChC inhibition impairs refinement of direction coding and motor learning.**

273 We asked if the abolition of ChC-mediated inhibition would affect the refinement of direction

274 coding and learning ability. We abolished presynaptic ChC GABA release by expressing FlpO-

275 dependent TeTxLC (AAV9-CAG-dfrt-TeTxLC-HA-WPRE) in ChC-Flp mice, and performed two-

276 photon calcium imaging in M2 neurons across sessions (**Fig. 7a-c**). ChC-TeTxLC mice showed

277 a deficit in learning the task, indicating the necessary role of axo-axonic inhibition for learning

278 goal-directed motor control (**Fig. 7e-l**). Given the vast difference in the numbers of PV-INs and

279 ChCs [37], it is possible that the suppression of a similar number of PV-INs has similar effects to

280 the effects exerted by the selective suppression of ChCs. To test this possibility, we abolished

281 the inhibition presynaptic GABA release of a fraction of PV-INs that corresponds approximately

282 to the percentage of ChCs in the premotor network (**Extended Data Fig. 6a-c**). The sparse,

283 sequential activity was observed in sparse PV-TeTxLC mice similar to PV control (**Extended**

284 **Data Fig. 6d**). However, in contrast to ChC-TeTxLC mice, the sparse PV-TeTxLC mice showed

285 normal improvement of learned performance (**Extended Data Fig. 6e, f**).

286 To determine the specific role of ChCs in suppressing irrelevant activity, we analyzed the

287 probabilistic distributions of neural tuning to MD as a function of learning. Compared to the

288 decrease in activity responding to non-PDs seen with learning, ChC-TeTxLC mice did not show

289 a reduced likelihood of non-PD activity over training (**Fig. 7m**). While the estimated probability of

290 the cell being active for non-PD movement (distance from PD > 90 degree) decreased in the

291 ChC control mice with learning, such reduction was not observed in ChC-TeTxLC mice (**Fig.**

292 **7n**). When the MD was opposite to the cell's PD, the probability of the cell being active

293 significantly decreased in the ChC-control group, but not in the ChC-TeTxLC group (**Fig. 7o**).

294 Consistently, we found that the decreases in pairwise correlation of neural responses with

295 learning were specific to ChC control but not ChC-TeTxLC mice (**Fig. 7p**). Additionally, posterior

296 probability density functions of MD were flattened over sessions in the ChC-TeTxLC group,

297 indicating impaired encoding of movement direction due to the abolition of ChC-mediated

298 inhibition (**Fig. 7q**). Given that intact ChC inhibition mediated the suppression of incorrectly

299 tuned responses with learning, these results collectively suggest a necessary role of ChCs in

300 refining distributed neural codes for active locomotion control.

301

Heterogeneous axo-axonic synaptic plasticity underlies organized motor control.

302

303 Our imaging data raised the possibility that ChC inhibitory synaptic strength may not be

304 uniformly altered by learning. To understand the cellular evidence in support of this claim, we

305 examined the structural rearrangement of axo-axonic synapses as proxies for the synaptic

306 strength changes in target neurons. We developed an automated detection algorithm to discern

307 the individual structures of ChC innervation on AISs of target PyNs, by using volumetric images

308 of M2 (**Extended Data Fig. 7**, See Methods). A large number of inhibitory presynaptic and

309 postsynaptic ChC-AIS structures, ChC axonal boutons, and gephyrin signals (GABAergic

310 inhibitory synapse scaffolding protein) along the AISs were systematically analyzed in both the

311 experimental group (n = 1634 AISs in 5 mice) and control group with *ad libitum* water (n = 3116

312 AISs in 6 mice) (**Fig. 8a-d**). We estimated structural synaptic strength by measuring presynaptic

313 structural efficacy (Pre-SSE) as a presynaptic feature and postsynaptic structural efficacy (Post-

314 SSE) as a postsynaptic feature based on ChC-AIS contacts and the corresponding gephyrin

315 intensity, respectively (**Extended Data Fig. 8a**). The Pre- and Post-SSE distributions showed

316 intrinsic heterogeneity in both control and experimental mice, indicating uneven inhibitory

317 weights for connections between ChCs and M2 neurons (**Fig. 8d-h**). These imaging results are

318 consistent with early Golgi staining and recent electron microscopy data that show non-random

319 distribution of ChC axonal terminations [18,38]. The strong correlation between Pre- and Post-SSE

320 showed the balance of presynaptic and postsynaptic strengths on AIS (**Fig. 8i**).

321 A subset of ChCs in the prelimbic cortex was reported to selectively control PyNs

322 located more superficially (< 50 μm from L1/2 border) for their specific projection to the

323 basolateral amygdala [31]. We tested whether the heterogeneity among axo-axonic synapses

324 depends on the cortical depth of AIS location. The overall synaptic strengths were not different

325 by the location of AISs in L2/3 of control mice and experimental mice (0 – 250 μm from L1/2

326 border, dotted and solid lines in **Fig. 8j**). However, we found that high and low portions of SSE

327 distribution increased in the experimental group, suggesting that the increased heterogeneity in

328 ChC-PyN connections with learning is not solely based on the cortical depth of AIS location

329 (**Fig. 8f, h**).

330   To decipher how learning reorganizes axo-axonic inhibition on AISs, the cumulative

331 density functions (CDFs) of SSEs in the experimental group of mice were compared to the

332 control group (**Fig. 8k, l**). If a subset of neurons is more tightly controlled by ChCs with learning

333 (reflecting a sparse and determinant inhibitory control), we expect to see an increased

334 percentage of AISs largely covered or scarcely covered by ChC boutons, respectively. Indeed,

335 we observed an increased portion of both weak and strong ends in the distributions of pre and

336 post-SSEs in the trained mice. Classifying AISs into three subgroups by their SSE strength

337 (high, mid, and low), 4-9% of ChC-AISs underwent structural changes from mid to high or low

338 (**Fig. 8m, n**). To validate whether this alteration was a result of training and not of high variance

339 among samples (**Extended Data Fig. 8b-e**), we performed a bootstrapping test (n = 3000

340 random sampling) on the data from each mouse (**Extended Data Fig. 8f-i**). Obtained from

341 multiple bootstrapping tests (n = 10), changes in the SSE distributions for trained mice showed

342 increases in high (> 1.5 for Pre-SSE, > 0.6 for Post-SSE) and low (< 0.5 for Pre-SSE, < 0.1 for

343 Post-SSE) value ranges (**Fig. 8o, p**). We found increased divergence only between the

344 experimental and control groups (**Extended Data Fig. 9f-i**, Exp. vs. Ctrl.), and not between

345 other condition comparisons (Exp. vs. Exp.; Control vs. Control; Unlabeled vs. Unlabeled). The

346 alterations in each subgroup were analogous and tested as significant (**Fig. 8q, r**). Thus, these

347 data suggest that ChCs differentially regulate the degree of inhibition applied to each axo-axonic

348 synapse during learning (**Fig. 8s**).

349

Discussion

351 Given the diversity of interneuron types and their specific connectivity in neural circuits, it has

352 been a longstanding belief that the impact of inhibition on neural computation depends on a

353 combination of interneuron cell-types and their connectivity patterns [30,39–41]. Genetically-defined

354 interneuron groups such as PV, SOM, and VIP, and their connectivity motifs have profoundly

355 advanced our understanding of cortical computation and psychiatric disorders. Different

356 interneuron subtypes display their own circuit motifs, but subtypes remain widely

357 heterogeneous. Therefore, we still do not fully understand the precise and unique roles of

358 inhibition mediated by most single interneuron subtypes.

359   Despite the discovery of ChCs over four decades ago [42] and their clinical relevance to

360   neuropsychiatric disorders such as schizophrenia characterized by GABAa receptor

361   dysregulation at AIS [43], the functional role of ChCs in cognition and behavior has been

362   enigmatic. Using a combination of chronic two-photon population imaging together with

363   analyses on the structural plasticity at axo-axonic synapses, we demonstrated a critical role of

364   the ChC in shaping the cortical coding of L2/3 M2 neurons during motor learning. ChC activity

365   was necessary for improving the precision and accuracy of directional movement control. ChCs

366   fine-tune the direction selectivity of individual premotor neurons and their collective tone.

367   Notably, global population activity was not perturbed when ChCs were blocked. The necessary

368   role of ChCs without perturbing overall network excitability implies that ChCs mediate their

369   regulatory effect by establishing a specific neuronal connectivity pattern for learning.

370   Quantification of a large number of ChC-AIS synapses uncovered revealed subsets of them

371   underwent modifications associated with learning. This analysis revealed a significant degree of

372   heterogeneity, which provides evidence of unequal inhibitory weighting among ChC-PyN

373   connections during learning.

374   Diverse interneurons have been shown to be involved in motor learning by shaping

375   neuronal population activity. Particularly, it has been shown that SOM-INs targeting the

376   dendrites and spines of PyNs in the primary motor cortex undergo plasticity during motor

377   learning and control learning-dependent sequential activation of PyNs in a manner comparable

378   to what is attributed to ChCs here [44]. In addition, inhibiting SOM-INs during new learning

379   destabilized previous learning-dependent sequential activity of PyNs, suggesting the importance

380   of SOM-IN inhibition for maintaining previously learned sequential activation patterns and

381   behavioral improvement when a new motor skill task is introduced. However, in our study, we

382   observed that acute optogenetic suppression of SOM-INs did not significantly disrupt navigation

383   task performance after learning had been established (**Fig. 3a, Extended Data Fig. 4**). Given

384   the reduced activity of SOM-IN activity during learning and partial restoration of naive-like

385   activity by post-learning SOM-IN reactivation [45], the acute post-learning inhibition of SOM-INs is

386   less likely to dominantly regulate neural representations formed with learning. We cannot rule

387   out the possibility that SOM-INs are involved in shaping learning by dendritic-specific

388   computation through synaptic plasticity on the dendritic spines of PyNs. It is possible that the

389   suppression of irrelevant activity that is attributed to ChCs in the present study can be partly

390   mediated by SOM-INs via their recurrent connectivity [46]. Subpopulations of SOM-INs increase

391   and decrease activities during motor learning in a task-specific manner [44]. It is likely that diverse

392   and task-specific activity of SOM-INs may mediate the temporal shift of motor learning-induced

393   sequential activity in PyNs, which can complement organized population coding in sequence

394   during goal-directed navigation.

395   The abolishment of perisomatic inhibition in PV-TeTxTC mice excessively increased the

396   overall activities of M2 neurons, consistent with the notion that PV-INs are critical regulators of

397   cortical excitatory and inhibitory (E/I) balance [33,34]. Perisomatic inhibition from PV-INs provides

398   powerful inhibitory control over the PyN population [29–32]. Fast-spiking PV-Basket cells (PV-BCs)

399   are well suited for regulating the balance, gain and network oscillation of relatively broad PyN

400   populations [47,48]. Balanced and delayed inhibition of fast-spiking INs with movement activity may

401   provide temporal sharpening of motor command and suppress irrelevant activities [35].

402   PV-BCs and ChCs have different dendritic arborizations and laminar locations of their

403   soma. They receive spatiotemporally distinct patterns of excitatory synaptic inputs from local

404   PyNs, long-range thalamocortical connections, and neuromodulatory inputs, and in turn form

405   different recurrent feedback excitation and inhibition patterns in a microcircuit [49,50]. The distinct

406   wiring features and output properties of PV-BCs and ChCs may give their differential control

407   over the spike-timing of PyNs [51]. Different cholinergic modulatory effects of synaptic inhibition

408   between ChCs and PV-BCs have been reported [52]. Notably, a previous *in vivo* study showed

409   that ChCs significantly increase their firing rate during arousal when brain states switched from

410   slow to theta oscillations in the hippocampus and showed a low-amplitude desynchronized field

411   potential in the prelimbic cortex, while the firing rates of PV-BCs and PyNs remained unchanged

412   [53]. Given that arousal is often associated with cholinergic signaling, differential modulation by

413   cholinergic receptor activation between ChCs and PV-BCs may result in their distinct

414   contributions to population activity control in a state-dependent manner. A future study is

415   needed to further clarify the functional difference between ChCs and other types of BCs in

416   behavior control, which is important to specify the role of perisomatic inhibitory circuit motifs.

417   Our findings support a model that ChCs exert differential strengths of control over

418   subsets of PyNs after learning. There exists substantial variability in the magnitude of ChC input

419   to PyN AISs [54,55]. This variability reflects the ability of ChCs to regulate their inhibitory strength

420   based on the characteristics of its target cell, and may have a role in shaping the functional

421   properties of PyNs. Variability in the number of axo-axonic ChC synapses correlated with

422   structural features of individual target PyNs including laminar depth of the soma, other sources

423   of perisomatic inhibition, and size of soma and AIS [18]. The innervated PyNs are not distributed

424   at random, but rather show a clustered distribution in a spatially heterogeneous fashion [56],

425   supporting the existence of target selectivity and/or avoidance by ChCs on local PyNs. Two

426   factors possibly select PyNs for stronger/weaker inhibition. First, ChCs differently inhibit PyNs

427   based on the projection targets of the PyNs. Recent studies showed that a subset of L2 ChCs

428   selectively innervates PyNs projecting to the basolateral amygdala over those projecting to the

429   contralateral cortex in the prelimbic cortex [31]. Second, ChCs differentially inhibit PyNs in an

430   activity-dependent manner during learning [57]. In this study, we analyzed a restricted subset of

431   ChCs in the L2 and sampled neighboring PyNs in the shallow part of L2/3. Deeper L2/3 PyNs

432   may have different projecting targets. Nevertheless, our findings on learning-dependent

433   changes of axo-axonic synapses in the shallow L2/3 suggest that the difference in ChC-

434   mediated inhibition is not solely resulted from the anatomical properties of target neurons. Given

435   experimental evidence for the activity-dependent structural plasticity of axo-axonic synapses [58–

436   60], it is conceivable that the plasticity of axo-axonic synapses provides an activity-dependent

437   regulatory mechanism in the cell-by-cell level to fine-tune neuronal excitability across diverse

438   cells involved in different functional networks. This can be mediated through precise inhibitory

439   synaptic plasticity, which has been theoretically suggested as essential for the formation and

440   maintenance of functional cortical circuitry [61,62]. Our findings suggest that the learning-

441   dependent synaptic plasticity of axo-axonic synapses provides individualized inhibition to each

442   PyN, which is adjusted to precisely balance excitation and inhibition based on the relevancy of

443   its activity to the task goal.

444   Our measures of direction selectivity and population vector coding of MD may not fully

445   capture high-order representations such as state-dependent action selection. It is presumed that

446 multiple other inputs, such as cholinergic, thalamic, and other cortico-cortical inputs, innervate

447 ChCs and PyNs. Premotor mechanisms of direction selectivity coding, movement direction, and

448 inhibitory synaptic plasticity coming from different inputs remain to be studied.

449   Here we revealed a specialized inhibitory role for ChCs in experience-dependent

450 plasticity of cortical coding, which allows for flexible and robust learned behavior. Given the

451 unknown functions of a number of heterogeneous interneuron types, our results suggest that

452 such heterogeneity may have evolved to serve specific functional needs. Unraveling the

453 multitude of inhibitory effects on neural representation and signal transmission across all cell

454 types is essential for understanding how cognitive functions are internally and externally

455 governed to give rise to meaningful behavior. Precise balance and target-specific inhibitory

456 synaptic plasticity by each inhibitory cell type are thought to contribute to the formation and

457 retrieval of distributed memory, providing fundamental neural circuit architecture for learning.

458 The heterogeneity of inhibitory cell types with their diverse receptors and innervation patterns

459 offer us a glimpse into how the intricate and versatile neural coding in the cortex can be adopted

460 to generate complex and flexible behavior. Our work sheds light on the importance of future

461 studies to investigate additional single-cell type specific plasticity rules and their impacts on

462 network activity and behavior.

463

464 **Methods**

465 **Animals.** C57BL/6J, PV-Cre (Cat.# 8069), SOM-Cre (Cat.# 13044) mice, Vipr2-Cre (Cat.#:

466 31332), Ai14 mice (Cat.# 7914) from Jackson laboratory (Bar Harbor, ME, USA) and Swiss

467 Webster (Cat.#24) mice from Charles River Laboratory (Boston, MA, USA) were used in this

468 study (4-9 weeks old, both sexes). All mice were maintained on a 12 hr light/ 12 hr dark cycle.

469 All experimental procedures were carried out in accordance with protocols approved by Johns

470 Hopkins University Animal Care and Use Committee, the Max Planck Florida Institute for

471 Neuroscience Institutional Animal Care and Use Committee, and National Institutes of Health

472 guidelines. For genetically targeting ChCs, Nkx2.1-2a-CreER and Vipr2-Cre mice were used.

473 Swiss Webster mice were used as the background strain for Nkx2.1-2a-CreER mice while

474 C57BL/6J mice were used as the background strain for Vipr2-Cre mice. Nkx2.1-2a-CreER and

ROSA-Flex-FlpO mouse lines were generated in the laboratory of Hiroki Taniguchi at the Max Planck Florida Institute for Neuroscience. For Nkx2.1-2A-CreER mice, a 2A-CreER cassette was inserted in the frame immediately after an open reading frame of an Nkx2.1 gene. The targeting vector containing 5' and 3' homology arms, a 2A-CreER cassette, an frt-Neo-frt cassette, and an HSV-TK gene was constructed. For FLEX-Flp mice, the targeting vector containing Rosa26 homology arms, a CAG promoter, and a FLEX-Flp cassette was constructed. Both targeting vectors were generated using a PCR-based cloning strategy. 129SVj/B6 F1 hybrid ES cells (V6.5) were electroporated with the targeting vectors and subjected to drug resistance tests. Neomycin-resistant ES clones for Nkx2.1-2A-CreER and FLEX-Flp were screened by mini-Southern blotting and PCR, respectively, for correct targeting. Positive ES clones were used for tetraploid complementation to obtain male heterozygous mice following standard procedures. Nkx2.1-2a-CreER::Ai14 and Nkx2.1-2a-CreER::Flex-FlpO mice were generated by crossing Nkx2.1-2a-CreER with Ai14 and ROSA-Flex-FlpO mouse lines, respectively. Both created ChC mice showed no visible behavioral phenotypes and were able to learn the behavior task used in this study.

Tamoxifen (Tmx) administration. Tmx was administered to timed pregnant Swiss Webster females that were bred to Nkx2.1-2a-CreER::Flex-FlpO males by oral gavage at E17 to induce CreER activity in the offspring. To achieve a high density of ChCs, the Tmx dose was adjusted to 3 mg/30 g of body weight. Tmx solution was prepared at a working concentration of 20 mg/ml in corn oil (SIGMA), kept protected from light, and kept refrigerated for no longer than one month.

Animal surgery and stereotactic viral injection. Surgeries were performed on 4–8 week-old mice. Mice were anesthetized with an intraperitoneal injection of an anesthetic cocktail containing ketamine (87.5 mg/kg) and xylazine (12.5 mg/kg) (Santa Cruz Animal Health). The animal's scalp was shaved and any remaining hair was removed with a hair remover lotion (Nair, Church & Dwight Co, Inc, Princeton, NJ, USA). Ophthalmic ointment (Puralube Vet Ophthalmic Ointment) was applied to prevent eyes from drying. Next, the animal was placed in

504   a stereotaxic device (Kopf, Model 900 Small Animal Stereotaxic Instrument). The surgical region

505   was scrubbed by 10% betadine solution (Purdue product LP, Stamford, CT, USA) and cleaned

506   with 70% alcohol solution. Body temperature (37–38 °C) was maintained by a thermostatically

507   controlled heating pad (Harvard Apparatus, Holliston, MA, USA). A small incision was made on

508   the scalp. A small craniotomy (~0.5 mm in diameter) was made above the injection site (the

509   right premotor cortex, A/P: +1.75 mm, M/L: +0.3 mm from the bregma, D/V: -0.25 mm from the

510   brain surface). AAV1-hSyn-GCaMP6s-WPRE-SV40 (400 nl, Penn Vector Core) for C57BL/6J

511   mice as the experimental group; a mixture of AAV1-CaMKII0.4-Cre-SV40 (Penn Vector core),

512   AAV1-Syn-Flex-GCaMP6s-WPRE-SV40 (Penn Vector core), and AAV9-CAG-Flex-tetanus toxin

513   light-chain (produced from ViGene Biosciences, MD) (1:1:1 ratio, total injection volume: 800 nl)

514   for C57BL/6J mice as CaMKII-TeTxLC; AAV9-EF1a-DIO-eNpHR3.0-eYFP-WPRE-hGH (400 nl,

515   Penn Vector Core) for PV-Cre as PV-eNpHR and SOM-Cre mice as SOM-eNpHR; a mixture of

516   AAV1-hSyn-GCaMP6s-WPRE-SV40 (400 nl) and AAV1-CAG-Flex-tdTomato-WPRE-bGH (400

517   nl, Penn Vector core) for PV-Cre mice as PV control; a mixture of AAV1-hSyn-GCaMP6s-

518   WPRE-SV40, AAV1-CAG-Flex-tdTomato-WPRE-bGH, and AAV9-CAG-Flex-tetanus toxin light-

519   chain (ViGene Biosciences, MD) (1:1:1 ratio, total injection volume: 800 nl) for PV-Cre mice as

520   PV-TeTxLC; AAV1-hSyn-GCaMP6s-WPRE-SV40 (400 nl) for Nkx2.1-2a-CreER::Ai14 as ChC

521   control; a mixture of AAV1-hSyn-GCaMP6s-WPRE-SV40 (300 nl) and AAV9-CAG-dfrt-hM4Di-

522   mCherry-WPRE (500 nl) for Nkx2.1-2a-CreER::Flex-FlpO mice as ChC-hM4Di; a mixture of

523   AAV1-hSyn-GCaMP6s-WPRE-SV40 (300 nl) and AAV9-CAG-dfrt-TeTxLC-HA-WPRE (500 nl)

524   for Nkx2.1-2a-CreER::Flex-FlpO mice as ChC-TeTxLC; a mixture of AAV1-hSyn-GCaMP6s-

525   WPRE-SV40 and AAVDJ-hSyn-FLEX-TeTxLC-P2A-dTom (1:2,000 dilution, a gift from Sandeep

526   Robert Datta Lab)(1:1 ratio, total injection volume: 800 nl) for PV-Cre as sparse PV-TeTxLC; a

527   mixture of AAV1-hSyn-GCaMP6s-WPRE-SV40 and AAVDJ-hSyn-FLEX-TeTxLC-P2A-dTom

528   (1:1 ratio, total injection volume: 800 nl) for Vipr2-Cre mice as ChC-TeTxLC; AAV1-Syn-Flex-

529   GCaMP6s-WPRE-SV40 (400nl, Penn Vector core) for Vipr2-Cre mice as ChC-GCaMP6; AAV1-

530   CAG-Flex-TdTomato (400nl, Addgene, Cat.#: 28306) for Vipr2-Cre mice were used for virus

531   injection. The viral constructs were injected via a beveled glass micropipette (tip size 10–20 µm

532   diameter, Braubrand) backfilled with mineral oil. Flow rate (100-150 nl/min) was regulated by a

533   syringe pump (World Precision Instruments). Following virus injection, skin adhesive (3M

534   vetbond) was applied to close the incision site. General analgesia (Buprenorphine SR, 0.6

535   mg/kg) was injected subcutaneously and mice were monitored until they recovered from

536   anesthesia. Around one week later, mice were anesthetized and hair was removed. The scalp

537   was removed in a circular shape, and the surface of the skull was cleaned. A craniotomy (3.2

538   mm in diameter) was made and a glass cranial window (3 mm diameter, Warner Instruments,

539   Hamden, CT, USA) was implanted on the virus injection site and a custom-made headplate was

540   attached to the exposed skull with dental adhesive (C&B-Metabond, Parkell inc, Edgewood, NY,

541   USA).

542

543   **Virus generation.** To create AAV9-CAG-dfrt-hM4Di-mCherry-WPRE and AAV9-CAG-dfrt-

544   TeTxLC-HA-WPRE, pAAV9-CAG-dfrt-hM4Di-mCherry-WPRE and pAAV9-CAG-dfrt-TeTxLC-

545   HA-WPRE vectors were created by infusion of dfrt-hM4Di-mCherry and dfrt-TeTxLC-HA `R-

546   products into pAAV9-CAG vectors. 293FT cells were transfected with pAAV-transfer vectors,

547   PAD-Helper vectors and pSerotypeSpecific (AAV9) plasmids using a standard calcium

548   phosphate method. After 70 hrs, cells were harvested in PBS and resuspended in 150 mM NaCl

549   and 50 mM Tris, pH 8.5. After 3 freeze-thaw cycles, the virus solution was spun down at 4000 x

550   g for 30 min and the supernatant was incubated in benzonase (100 U/ml) at 37 °C for 1 h.

551   Another centrifugation step of 4000 x g for 25 minutes was followed by filtration of the

552   supernatant through 0.45 μm and 0.22 μm. Next, the virus solution is put on top of a sterile

553   Iodixanol gradient in an ultracentrifugation tube (15%, 25%, 40%, and 60% Iodixanol in 1 M

554    $MgCl_2$, 2.5 M KCl, 5 M NaCl in PBS). After centrifugation at 31,000 RPM for 6 hrs, the 40%

555   Iodixanol band was harvested and stored in DPBS and concentrated using 100,000

556   concentrator vials (Amicon). The virus solution was aliquoted and stored at -80 °C until use.

557

558   **Behavior training in a floating ball maze.** More than 2 weeks (19~25 d) after virus injection,

559   mice were trained to perform a tactile-based spatial navigation task while head-restrained on a

560   custom-made spherical floating ball maze [22]. Mice in the water-restricted group were water-

561   restricted at 1 ml per day for a week before training starts. Mice in the control group were given

562   *ad libitum* water access and, in addition, 2 ml of 10% sucrose were given 30 minutes before

563   training. The floating ball maze was made of an air-supported 8" diameter spherical styrofoam

564   ball with quadrants that have different tactile surface textures (ex. plain, grooved, striped, and

565   grid). A simple auditory tone (8 kHz for 2 s) was delivered at the beginning and end of training.

566   During navigation, mouse movement was monitored by a CMOS camera (Thorlabs) and ball

567   movement was continuously monitored and recorded by a Bluetooth motion sensor (LPMS-B,

568   LP research) placed at the center of the ball. Mice were trained to navigate to find a hidden goal

569   location (surface area of 30 ° solid angle) in which a 10% sucrose reward (~10 µl per a reward)

570   was delivered if they were on the location. To facilitate animals' exploration, a brief air puff (60

571   psi, 200 ms) was applied to their hind limbs if they did not move for 30 seconds. The reward

572   was delivered via a lick port positioned in front of the mouse's mouth and its timing was

573   controlled by a solenoid valve (NResearch, West Caldwell, NJ, USA). To prevent excessive stay

574   within the goal location and excessive access to the goal, a 3-second refractory period was

575   assigned for the reward; Either a stay in the goal location or a re-visit to the goal location within

576   3 seconds from the last reward did not provide an additional reward. The behavioral setup was

577   controlled by a custom-written code in MATLAB (Mathworks, Natick, MA, USA). Daily training

578   sessions lasted 800 seconds and occurred for 7 days.

579

580   **Optogenetic inhibition of parvalbumin (PV) and somatostatin (SOM)-positive**

581   **interneurons.** We bilaterally injected AAV9-EF1a-DIO-eNpHR3.0-eYFP-WPRE-hGH (400 nl,

582   Penn Vector Core) into the L2/3 M2 of PV-Cre mice or SOM-Cre mice and implanted fiber-optic

583   cannulae (200 µm core, 0.37 NA, BFL37-2000, Thorlabs) with head plates. After two weeks of

584   recovery from surgery and one week of water restriction, mice were trained on a floating ball

585   maze for a session (800 s) per day for 7 days. After 7 days of training, mice were trained for a

586   session and simultaneously received bilateral optogenetic inhibition using 589nm-light

587   administration (continuous 800 s during the session, ~10 mW power at the optic-fiber tip, MBL-

588   FN-589, Changchun New Industries Optoelectronic Technology) through the fiber-optic

589   cannulae. On the following day, mice were trained without optogenetic inhibition as a control.

590

591 **Chemogenetic inhibition of ChCs.** We bilaterally injected a mixture of AAV1-hSyn-GCaMP6s-

592 WPRE-SV40 (300 nl) and AAV9-CAG-dfrt-hM4Di-mCherry-WPRE (500 nl) into the L2/3 M2 of

593 ChC-Flp mice. After 7 days of training, to suppress the activity of ChCs expressing AAV9-CAG-

594 dfrt-hM4Di-mCherry in a FlpO-dependent manner, mice were administered i.p. with clozapine N-

595 oxide (CNO, Tocris, Cat.#: 4936) 45 mins before a behavioral session and underwent the

596 session. On the following day, mice were administered i.p. with saline as control and underwent

597 the session.

598

599 *In vivo* **two-photon imaging.** Imaging was conducted with a two-photon microscope (Bruker)

600 using two pulsed Ti:sapphire lasers tuned to a wavelength of 920 nm (MaiTai HP DeepSee,

601 Newport Spectra-Physics) for GCaMP6 signals or a wavelength of 1,045 nm (HighQ-2, Newport

602 Spectra-Physics) for tdTomato and mCherry signals. The calcium imaging dataset was motion-

603 corrected using a custom-written MATLAB code based on a full-frame cross-correlation image

604 alignment algorithm. Regions of interest (ROIs) were semi-automatically drawn using a custom

605 algorithm based on fluorescence intensity, cell size, and cell shape by visually inspecting

606 movies, average of movies, and standard-deviation of movies, and then selecting neurons that

607 showed fluorescence transients at least once during a session. The same area was imaged

608 across training days to trace changes from same ensembles, but due to intrinsic tissue

609 movement, neuronal activity was realigned when calcium signals were analyzed. All pixels

610 within each ROI were averaged to create a fluorescence time series, *F*. Surrounding 'neuropil

611 ROIs' were drawn from 0.3 μm outside of the border of each neuronal ROI to 10.0 μm outside.

612 Neuropil ROIs excluded adjacent neuronal ROIs. Pixels within neuropil ROIs that showed

613 apparent calcium transients exceeding 3 standard deviations of the difference in fluorescence

614 between each neuropil ROI pixel time series and the neuronal ROI time series were excluded

615 from further analysis. The remaining neuropil ROI pixels were considered as background

616 fluorescence and averaged to create a time-varying background fluorescence trace. The time-

617 varying baseline (*F0*) of a fluorescence trace was estimated by the following procedure: A

618 preliminary baseline fluorescence time series, *Pre F0*, was loess smoothed with 120 frames.

619 Preliminary *ΔF*, *Pre ΔF*, was obtained by subtracting *Pre F0* from *F*. Noise of Pre *ΔF* was

620   estimated by subtracting loess smoothed *Pre ΔF* from the standard deviation of *Pre ΔF*. The

621   offset of *ΔF* was determined by the mean of the distribution of *Pre ΔF* that did not exceed two

622   times the noise. The baseline fluorescence trace, *F0*, was estimated by adding the offset to *Pre*

623   *F0*. *ΔF/F* for neuronal ROIs was obtained by subtracting *F0* from *F* and dividing it by *F0*. The

624   background fluorescence trace *ΔF/F* for neuropil ROIs was estimated in the same manner, i.e.,

625   as the *ΔF/F* for neuronal ROIs. The background fluorescence trace was subtracted with a 0.7

626   weight from the neuronal ROI fluorescence trace for each neuronal ROI. For pixel-based

627   imaging analysis of MD, fluorescence change maps of images were color-coded by MD and

628   averaged. For pixel-based analysis of different movement paths, fluorescence change maps of

629   images for each movement path were differently color-coded and overlaid.

630

631   **Tissue fixation, immunohistochemistry, and acquisition of confocal and light-sheet**

632   **microscope images.** Animals were deeply anesthetized with a mixture of ketamine and

633   xylazine and then perfused transcardially, first with phosphate buffered saline (PBS, pH 7.4) and

634   next with 4% paraformaldehyde (PFA) dissolved in PBS. The brains were removed and

635   postfixed in 4% PFA overnight at 4 °C. For confocal microscope images, the brains were

636   embedded in 10% melted gelatin solution for 50 minutes at 50°C. Melted gelatin solution was

637   then refreshed and the gelatin solution with the embedded brains was solidified at 4 °C for about

638   30 minutes. Then, the solidified gel was trimmed around the embedded brain to a small cube

639   shape and the cube was stored into 4% PFA overnight. The gelatin cube brain was coronally

640   sectioned (100 μm thick) using a vibratome (Leica Biosystems, Buffalo Grove, IL, USA). For

641   parvalbumin immunohistochemistry, after blocking with a solution (1 % Normal Donkey

642   Serum/0.1 Triton-X in PBS), brain slices were incubated in primary antibodies (1:500 mouse

643   anti-parvalbumin, #P3088, Sigma) for 48 hours at 4 °C. After washing 3 times in PBS, the brain

644   slices are then incubated in a secondary antibody (Cy5-conjugated donkey anti-mouse IgG,

645   Cat.#: 715-585-150, diluted 1:1,000 in PBS, Jackson ImmunoResearch Laboratories) for 2

646   hours at room temperature, followed by 3 washes in PBS. The brain slices were mounted on the

647   slide glass with a DAPI-containing mounting solution (DAPI Fluoromount-G, SouthernBiotech,

648   USA). For Nkx2.1-2a-CreER::Ai14 and Nkx2.1-2a-CreER::Flex-FlpO mice

649   immunohistochemistry, after intraperitoneal administration of Ketamine/Xylazine mixture (50

650   mg/kg Ketamine, VETCO, 5 mg/kg Xylazine, AKORN) and a foot pinch to check adequate

651   sedation, the mice were trans-cardially perfused with 15 ml cold saline solution followed by 20

652   ml of cold 4% PFA solution in PBS. After post-fixation in 2% PFA in PBS overnight, 60 μm brain

653   sections were prepared using a Leica vibratome and collected in PBS, permeabilized in 0.5%

654   TritonX-100 in PBS followed by washing for 1 h in blocking reagent (BR: 0.1% TritonX-100, 10%

655   donkey serum in PBS). The sections were treated in primary antibody solution in BR overnight

656   at 4°C. The following primary antibodies were used: rat anti HA (1:500, Roche, #11-867-423-

657   001), chicken anti GFP (1:800, ABCAM, Cat.#: ab13970), and mouse anti AnkyrinG (1:500, UC-

658   Davis/NIH NEUROMAB, Cat.#: clone N106/36 75-146). After three washes (10 min) in PBS, the

659   slices were incubated in BR with secondary antibodies (Jackson ImmunoResearch) for 2 h at

660   room temperature containing appropriate fluorophores (CyX or Alexa) at 1:1000. The slices

661   were embedded in FluoromountG (Thermo Fisher) [K1] after two final washing steps in PBS.

662   The fixed and immuno-stained brain slices were imaged using an upright confocal microscope

663   (Zeiss LSM880, Germany). For light-sheet microscope images, brains were perfused as

664   described above and were coronally sectioned (2 mm thick) using a vibratome (Leica

665   Biosystems, Buffalo Grove, IL, USA). The brain slices were cleared using the Passive Clarity

666   protocol (Yang et al., 2014). Images of the cortex were collected with a Zeiss Light-sheet

667   microscope at 25x magnification. Light-sheet microscope image processing and 3D rendering

668   were performed with arivis Vision4D.

669

670   **Preparation for acute cortical slices.** Mice were anesthetized with isoflurane and decapitated.

671   The brain was removed and rapidly placed into ice-cold cutting solution containing (in mM): 215

672   Sucrose, 26 $NaHCO_3$, 20 Glucose, 4 $MgCl_2$, 4 $MgSO_4$, 2.5 KCl, 1 $CaCl_2$, and 1.6 NaH_2PO_4.

673   Cortical slices (300 μm thick) were prepared using a VT1000S vibratome (Leica). Slices were

674   incubated at 32 °C for 30 minutes in a holding chamber filled with artificial cerebrospinal fluid

675   (ACSF) containing (in mM): 124 NaCl, 26 $NaHCO_3$, 10 glucose, 3 KCl, 1.25 NaH_2PO_4, 2.5

676     $CaCl_2$, and 1.3 $MgSO_4$. Slices were then placed at room temperature and allowed to recover

677   for 1 hour before recording. All solutions were equilibrated with carbogen (95% O_2/5% CO_2).

678

679 **Electrophysiology.** Whole-cell patch-clamp recordings (electrode resistance 5-9 MΩ) were

680 performed at room temperature using a Multiclamp 700B amplifier (Molecular Devices). Layer

681 2/3 chandelier cells were patched in current-clamp configuration using potassium-based internal

682 solution (in mM: 10 NaCl, 1 MgCl2, 4 NA2-ATP, 124 K-gluconate, 10 Na2-phosphocreatine, 16

683 KCl, 0.4 Na-GTP, 10 HEPES) in ACSF containing (in mM: 124 NaCl, 26 NaHCO3, 10 glucose,

684 3 KCl, 1.25 NaH2PO4, 2.5 CaCl2, and 1.3 MgSO4). To examine the effect of DREADD virus

685 hM4Di, we recorded the firing properties of hM4Di-expressing neurons before and after applying

686 10 μM clozapine N-oxide (CNO, Tocris, Cat.#: 4936) (in ACSF).

687

688 **Data acquisition and analysis.** All data analysis was performed using Excel (Microsoft), and

689 custom codes and toolboxes in MATLAB. Behavioral sessions in which the mouse made more

690 than three successful accesses to the goal were analyzed. The movement signals of the ball

691 maze including quaternion and angular velocity were sampled at 100 Hz and smoothed using a

692 lowess regression of 500 ms width. We extracted animals' virtual coordinates on the ball relative

693 to the center of the hidden goal spot. We acquired mouse movement-related information based

694 on translational movement (forward and backward; right and left) by using a precise Bluetooth

695 motion sensor (LPMS-B, LP research, Tokyo, Japan).

696

697 **Identification of movement bouts.** We classified each behavior frame (sampled at 100 Hz, 1

698 ms bin size) as occurring during either rest or movement. We defined rest frames as those in

699 which the movement speed is less than 0.02 m/sec and acceleration is less than 0.2 m/sec. We

700 defined movement frames as those in which a movement speed is greater than 0.03 m/sec and

701 not classified as rest frames. Parameter values were determined by visually comparing videos

702 with movement speed and acceleration traces. The uncategorized frames which do not meet

703 the above criteria were further categorized as following: If both ends of the bout of

704 uncategorized frames were connected to rest frames and the duration of the bout of

705 uncategorized frames was less than 1 second, the uncategorized frames were identified as rest

706 frames. If either both ends of the bout of uncategorized frames were connected to movement

707 frames or the state of connected frames of one end of the bout was different from that of the

708 other end, the uncategorized frames were identified as movement frames.

709

710 **Movement direction.** Movement direction was determined by the inverse tangent of forward-

711 backward speed over right-left speed. Goal-angle was defined as the angle passing through the

712 goal spot from the mouse's two-dimensional virtual position. Goal-directed movement epochs

713 were determined if the mouse's movement direction was within the goal-angle. Percentage of

714 goal-directed movement was determined by the ratio of goal-directed movement epochs to

715 overall movement epochs. Goal-directed movement bout was identified as a bout of movement

716 in which mice pursued the hidden goal location and received the reward. Non-goal-directed

717 movement bouts were defined as bouts wherein the animal does not reach the goal.

718

719 **Classification of movement-related neurons.** Modulation of neural activity was frequently

720 observed between movement and rest epochs. We classified individual neurons into three

721 categories, movement-positive, movement-negative, and non-movement. Movement positive or

722 negative neurons were identified if their mean activity during movement epochs was significantly

723 higher or lower than during rest epochs, respectively. If there was no significant difference in

724 mean neuronal activity between movement and rest epochs, neurons were classified as non-

725 movement cells.

726

727 **Estimation of probabilistic neural tuning curves.** To estimate calcium events, calcium

728 signals were deconvolved using an Online Active Set method to Infer Spikes (OASIS) with

729 FOOPSI approach and auto-regressive models with order 2. The values of calcium event

730 estimate s were threshold to 0, producing binary events that represent estimated calcium

731 events. To focus on the tuning curves of premotor neurons to movement direction, we therefore

732 excluded periods of immobility. Using this event estimate, we computed the probability of a

733 neuron to be active $P(active)$ as the period of activity of a neuron over the total period of a

734 recording session, which corresponds to the marginal likelihood probability distribution in a

735 Bayesian approach. We divided the movement direction in 5º bins. Each bin represents a

736   discrete state of movement direction. We then computed the likelihood probability distribution

737   that a cell is active given movement direction, $P(active|MD_i)$, as the period of activity in

738   direction state MD_i over the total period in state MD_i. The distribution was smoothed with a

739   Savitzky-Golay filter. To test statistical deviation from random firing patterns, we carried out a

740   bootstrapping procedure as follows. For each cell, deconvolved data were circularly shifted by

741   different random values of shifts and the likelihood probability distribution of the randomized

742   data was estimated. This procedure was repeated 100 times with different random values of

743   shifts. The mean value and standard deviation of the randomized dataset were used to compute

744   the z-scored likelihood distribution of actual data. A cell is defined as an active cell if the

745   likelihood distribution of actual data is significantly deviated from that of the randomized dataset

746   ($P < 0.05$).

747   Neurons were considered to be directionally tuned if they showed significant tuning with

748   movement direction (ANOVA, F test, $P < 0.05$) and they showed a good fit of tuning function. To

749   obtain the tuning curve, the z-scored likelihood distribution of actual data was fitted with the sum

750   of two von Mises distributions [63]:

751   $$R(\theta) = a_0 + a_1 e^{k\,cos(\theta - \theta_0)} + a_2 e^{k\,cos(\theta - \theta_0 + 180°)}$$

752   , where R is the response to movement in a direction θ, θ_0 is the preferred direction, and a0, a1,

753   a2, and k are fitting parameters. We estimated fitting parameters by using lsqcurvefit in the

754   Matlab function. The preferred direction PD_i for neuron *i* was defined as the movement direction

755   corresponding to a peak in the tuning curve response. Neurons were considered to be direction-

756   selective if they fit the von Mises distribution well ($P < 0.05$) and if their direction-selective index

757   (DSI) defined as $DSI = \frac{R_{pref} - R_{oppo}}{R_{pref} + R_{oppo}}$ was greater than 0.4, where R_{pref} and R_{oppo} are the mean of

758   the normalized response to the preferred direction and that of the opposite direction (the

759   preferred direction + π), respectively.

760

761   **Bayesian decoding analysis.** In addition to the estimation of the likelihood distribution,

762   $P(active|MD_i)$, we computed the probability of being in a given state of movement direction

763   MD_i, $P(MD_i)$, as the period in the direction MD_i over the total period of a recording session,

764   which corresponds to the prior probability distribution. Using Bayesian approach [64], we inferred

765   the posterior probability distribution function that the animal is in a movement direction given

766   neuronal activity at a time *t*.

767
$$P(MD|active) \ = \ \frac{P(active|MD)P(MD)}{P(active)} \ = \ \prod_{i=1}^{N} \blacksquare \frac{P(active_i|MD)P(MD)}{P(active_i)}.$$

768   where $P(MD|active)$ is a vector of a posterior probability distribution of behavioral states and *N*

769   corresponds to the number of neurons used. To decode the direction of movement, we

770   considered the direction associated with the maximum *a posteriori* (MAP).

771
$$\hat{S} \ = \ arg \ max \ P(MD|active) \ = \ arg \ max \ \frac{P(active|MD)P(MD)}{P(active)}$$

772   where Ŝ is the decoded movement direction. The decoding error was computed as the angular

773   difference between the decoded and actual movement direction at the observed time.

774

775   **Pairwise correlations.** We calculated pairwise correlations between neuronal pairs by

776   measuring coincident events [65]. For neuron i, the events E_i were vectorized in a binary manner.

777   The pairwise correlation for the pair i and j is given by

778
$$CI_{ij} = \frac{E_i \cdot E_j - N\langle E_i \rangle \langle E_j \rangle}{N\sqrt{\langle E_i \rangle \langle E_j \rangle}}$$

779   , where *N* is the number of image frames, and $\langle E_i \rangle$ is the expected number of events for neuron

780   *i*. We categorized positively and negatively correlated pairs if the CI is greater than 0 and less

781   than 0, respectively.

782

783   **Automated detection of ChC-AIS.** To label ChCs in M2, we injected AAV1-CAG-Flex-

784   TdTomato (400nl, Addgene, Cat.#: 28306) into L2 of M2 in Vipr2-Cre mice. Once the brain

785   slices were stained with immunohistochemistry and imaged through the confocal microscope

786   described in the tissue fixation, immunohistochemistry, and acquisition of confocal and light-

787   sheet microscope images section, each stack of tiled confocal images (generally 3800 X 1950 X

788   10 in pixel or 400 X 206 X 10 in μm) from three different fluorescent channels (green for

789   gephyrin puncta antibody, red for TdTomato expressed in ChC projection, and far-red for AnkG

790   antibody targeting AIS) was processed simultaneously to detect AISs innervated by axonal

791   cartridges of ChC and colocalized gephyrin puncta using a custom-written program in Matlab.

792   Features of processed images were segmented automatically by boundary detection

793   (bwboundaries function, Matlab) in each planar image of red and far-red channels after

794   binarization by adaptive threshold (adaptthresh function, Matlab, sensitivity = 0.08 and

795   neighborhood size = 19 for AIS, sensitivity = 0 and neighborhood size = 15 for ChC). Segments

796   in different axial planes were then filtered through several criteria including intensity, size, and

797   width and registered to recover the whole AISs or ChC cartridges. Only AISs longer than 10 μm

798   were selected for the further analysis. For each registered AIS, we overlaid its segment on

799   detected ChC cartridges in each axial plane and classified it as a ChC-AIS when any

800   colocalized area existed. For demonstration purposes, fluorescence images shown in figures

801   are maximum intensity projections of corresponding volumetric stack with pseudo-colors.

802

803   **Quantification of characteristic parameters.** For each detected ChC-AIS segment, we

804   calculated the lengthwise axis. Projected intensity profiles of ChC and AIS on the axis were then

805   applied to measure the total AIS length (L_{AIS}), the ChC cartridge length (L_{ChC}), and the AIS

806   length covered by ChC (L_{ChC_AIS}). To ensure the exact L_{ChC} measurement, we extended the ROI

807   5 μm longer than the detected AIS segments we identified. Half of the maximum intensity for

808   each profile was used as a threshold for measurement. The area of ChC on AIS ($A_{ChC-AIS}$) was

809   calculated as the mask size for segmentation. We computed the normalized ChC-AIS size, the

810   presynaptic parameter, as

811   $$PRE = \frac{A_{ChC-AIS}}{L_{AIS}} \times \frac{L_{ChC-AIS}}{L_{ChC}}$$

812   from these morphological quantities.

813   We used the particle detecting algorithm u-track [66] to identify gephyrin puncta in each plane of

814   the green channel. This algorithm fitted Gaussian kernels approximating the two-dimensional

815   point spread function of the microscope (σ = 1.5 pixel or 234 nm) around local intensity maxima,

816   where both position and amplitude were free parameters in the fit. The intensity of detected

817   gephyrin punctum is represented by a modified z-score calculated from the median and

818   standard deviation of all the detected gephyrin puncta from the tiled image.

819   To evaluate the postsynaptic strength, a detected gephyrin punctum was classified as

820   AIS-associated when its detected position resides in the boundary of a registered AIS. In the

821   same manner, an AIS-associated gephyrin punctum was labeled ChC-AIS associated when its

822   position colocalizes in the boundary of ChC on AIS simultaneously. Summation of the z-score of

823   gephyrin puncta associated with the total AIS (Z_{AIS}) and only a portion of ChC-AIS ($Z_{ChC-AIS}$) was

824   used for calculation of the corresponding gephyrin intensity, the postsynaptic parameter, as

825   follows.

826    $$POST = \frac{Z_{ChC-AIS}}{Z_{AIS}}$$

827   The corresponding gephyrin intensity equals 1 when all the gephyrin puncta on the AIS are

828   located on the corresponding ChC. When the ChC does not have corresponding gephyrin

829   puncta on the bound AIS, its value is 0.

830

831   **Statistics.**

832   Statistical analyses were performed using either MATLAB (Mathworks) or OriginPro (OriginLab).

833   All statistical tests are described in the corresponding figure legends. Normality of distributions

834   was tested for each dataset using Lilliefors test to decide whether to use parametric or non-

835   parametric tests. For parametric tests, two-tail Student's t-test, two-sample t-test, two-tailed

836   paired t-test, one-way ANOVA, one-way repeated measures ANOVA, and two-way repeated

837   measures ANOVA were performed to compare population means across animals and

838   conditions except where noted. For ANOVAs, Fisher post hoc tests were used for multiple

839   comparisons, When sphericity is violated, we used Greenhouse-Geisser correction. Chi-square

840   analyses were used to compare proportions. Pearson's linear correlation coefficient was used to

841   measure the correlation between two variables. For non-parametric tests, two-tailed Wilcoxon-

842   Mann-Whitney test and Friedman test with Dunn post hoc tests were performed to compare the

843   population median and differences with dependent variables across animals and conditions,

844   respectively. Two-sample Kolmogorov-Smirnov test was performed to compare probability

845   distributions between groups. The number of animals and the number of cells used for analysis

846   are specified in the figures, the figure legends, and the text.

847

848   **Data availability**

849   Additional data that supports the findings of this study are available from the corresponding

850   author upon reasonable request.

851

Code availability

The code for semi-automated ROI selection can be found at

https://github.com/fitzlab/CellMagicWand. Custom MATLAB scripts used in this study are

available from the corresponding author upon reasonable request.

856

867

Author contributions

K.J. and H-B.K. conceived and designed the study. K.J. designed and optimized head-fixed ball

maze system. K.J performed most of the experiments including viral injection, electrophysiology,

optogenetic manipulation, two-photon imaging and behavioral training. K.J. wrote behavior and

calcium imaging data analysis programs and analyzed all data. M.C. and D.C performed

immunostaining and data analysis. M.C. developed an algorithm detecting ChC-AIS synapses.

B.B. helped with virus injection surgery. A.S. and H.T. made Flex-FlpO mice and provided

Nkx2.1-2a-CreER::Ai14 and Nkx2.1-2a-CreER::Flex-FlpO mice. Y.O. and H.T. generated

Nkx2.1-2a-CreER mice. D.F. and H.T. provided advice on the manuscript organization and data

interpretation. K.J. and H-B.K. wrote the manuscript.

878

879 **Competing interests**

880 The authors declare no competing interests.

881

882

References

1. Xu, T. *et al.* Rapid formation and selective stabilization of synapses for enduring motor memories. *Nature* **462**, 915 919 (2009).

2. Lai, C. S. W., Franke, T. F. & Gan, W.-B. Opposite effects of fear conditioning and extinction on dendritic spine remodelling. Nature **483**, 87 91 (2012).

3. Peters, A. J., Chen, S. X. & Komiyama, T. Emergence of reproducible spatiotemporal activity during motor learning. Nature **510**, 263 267 (2014).

4. Hofer, S. B., Mrsic-Flogel, T. D., Bonhoeffer, T. & Hubener, M. Experience leaves a lasting structural trace in cortical circuits. Nature **457**, 313 317 (2009).

5. Petreanu, L. et al. Activity in motor-sensory projections reveals distributed coding in somatosensation. Nature **489**, 299 303 (2012).

6. Huber, D. et al. Multiple dynamic representations in the motor cortex during sensorimotor learning. Nature **484**, 473 478 (2012).

7. Taniguchi, H. et al. A resource of Cre driver lines for genetic targeting of GABAergic neurons in cerebral cortex. *Neuron* **71**, 995 1013 (2011).

8. Adesnik, H., Bruns, W., Taniguchi, H., Huang, Z. J. & Scanziani, M. A neural circuit for spatial summation in visual cortex. Nature **490**, 226 231 (2012).

9. Lee, S.-H., Kwan, A. C. & Dan, Y. Interneuron subtypes and orientation tuning. Nature **508**, E1 2 (2014).

10. Fu, Y. et al. A cortical circuit for gain control by behavioral state. *Cell* **156**, 1139 1152 (2014).

11. Kvitsiani, D. et al. Distinct behavioural and network correlates of two interneuron types in prefrontal cortex. Nature **498**, 363 366 (2013).

12. Somogyi, P. A specific 'axo-axonal' interneuron in the visual cortex of the rat. *Brain Res* **136**, 345–350 (1977).

13. Taniguchi, H., Lu, J. & Huang, Z. J. The spatial and temporal origin of chandelier cells in mouse neocortex. *Science* **339**, 70 74 (2013).

14. Yang, J.-M. et al. erbb4 Deficits in Chandelier Cells of the Medial Prefrontal Cortex Confer Cognitive Dysfunctions: Implications for Schizophrenia. *Cereb Cortex* **29**, 4**334**–4346 (2018).

15. Fazzari, P. et al. Control of cortical GABA circuitry development by Nrg1 and ErbB4 signalling. Nature **464**, 1376–1380 (2010).

16. Volk, D. & Lewis, D. Impaired prefrontal inhibition in schizophrenia: relevance for cognitive dysfunction. *Physiol Behav* **77**, 501–**505** (2002).

17. Dudok, B. et al. Recruitment and inhibitory action of hippocampal axo-axonic cells during behavior. Neuron **109**, 3838-3850.e8 (2021).

18. Schneider-Mizell, C. M. et al. Structure and function of axo-axonic inhibition. *Elife* **10**, 2020.03.31.018952 (2021).

19. Gremel, C. M. & Costa, R. M. Premotor cortex is critical for goal-directed actions. *Front Comput Neurosc* **7**, 110 (2013).

20. Li, N., Daie, K., Svoboda, K. & Druckmann, S. Robust neuronal dynamics in premotor cortex during motor planning. Nature **532**, 459 464 (2016).

21. Siniscalchi, M. J., Phoumthipphavong, V., Ali, F., Lozano, M. & Kwan, A. C. Fast and slow transitions in frontal ensemble activity during flexible sensorimotor behavior. *Nat Neurosci* **19**,

1234 1242 (**20**16).

22. Lee, D. et al. Temporally precise labeling and control of neuromodulatory circuits in the mammalian brain. *Nat Methods* **14**, 49**5**–503 (2017).

23. Oh, W. C., Lutzu, S., Castillo, P. E. & Kwon, H. B. De novo synaptogenesis induced by GABA in the developing mouse cortex. Science **353**, 1037 1040 (2016).

24. Zaccolo, M. & Pozzan, T. Discrete microdomains with high concentration of cAMP in stimulated rat neonatal cardiac myocytes. Science **295**, 1711 1715 (2002).

25. Seung, H. S. & Sompolinsky, H. Simple models for reading neuronal population codes. *Proceedings of the National Academy of Sciences of the United States of America* **90**, 10749 10753 (1993).

26. Georgopoulos, A. P., Schwartz, A. B. & Kettner, R. E. Neuronal Population Coding of Movement Direction. Science **233**, 1416–1419 (19**8**6).

27. Isaacson, J. S. & Scanziani, M. How Inhibition Shapes Cortical Activity. Neuron **72**, 231 243 (2011).

28. Rudy, B., Fishell, G., Lee, S. & Leffler, J. H. Three groups of interneurons account for nearly 100% of neocortical GABAergic neurons. *Dev Neurobiol* **71**, 45 61 (2011).

29. Markram, H. et al. Interneurons of the neocortical inhibitory system. *Nat Rev Neurosci* **5**, 793 807 (2004).

30. Kepecs, A. & Fishell, G. Interneuron cell types are fit to function. Nature 505, 3**18** 326 (2014).

31. Lu, J. et al. Selective inhibitory control of pyramidal neuron ensembles and cortical subnetworks by chandelier cells. Nat Neurosci 20, 1377 1383 (2017).

32. Buhl, E. H., Halasy, K. & Somogyi, P. Diverse sources of hippocampal unitary inhibitory postsynaptic potentials and the number of synaptic release sites. Nature **368**, 823 828 (1994).

33. Ferguson, W.-J. G. B. R. PV Interneurons: Critical Regulators of E/I Balance for Prefrontal Cortex-Dependent Behavior and Psychiatric Disorders. *Front Neural Circuit* **12**, 479 (2018).

34. Vogels, T. P. & Abbott, L. F. Gating multiple signals through detailed balance of excitation and inhibition in spiking networks. Nat Neurosci 12, 483 4**91** (2009).

35. Isomura, Y., Harukuni, R., Takekawa, T., Aizawa, H. & Fukai, T. Microcircuitry coordination of cortical motor information in self-initiation of voluntary movements. Nat Neurosci 12, 1586 1593 (2009).

36. Hayano, Y. et al. IgSF11 homophilic adhesion proteins promote layer-specific synaptic assembly of the cortical interneuron subtype. *Sci Adv* **7**, eabf1600 (2021).

37. Tasic, B. et al. Shared and distinct transcriptomic cell types across neocortical areas. Nature **563**, 72 78 (2018).

38. Defelipe, J., Hendry, S. H. C., Jones, E. G. & Schmechel, D. Variability in the terminations of GABAergic chandelier cell axons on initial segments of pyramidal cell axons in the monkey sensory-motor cortex. *J Comp Neurol* **231**, 364–384 (1985).

39. Tremblay, R., Lee, S. & Rudy, B. GABAergic Interneurons in the Neocortex: From Cellular Properties to Circuits. Neuron 91, 260–92 (2016).

40. Wamsley, B. & Fishell, G. Genetic and activity-dependent mechanisms underlying interneuron diversity. Nat Rev Neurosci 18, 299–309 (2017).

41. Cossart, R. The maturation of cortical interneuron diversity: how multiple developmental journeys shape the emergence of proper network function. *Curr Opin Neurobiol* **21**, 160–168

(2011).

42. Szentágothai, J. & Arbib, M. A. *Conceptual Models of Neural Organization*. (1974).

43. Lewis, D. A., Curley, A. A., Glausier, J. R. & Volk, D. W. Cortical parvalbumin interneurons and cognitive dysfunction in schizophrenia. *Trends Neurosci* **35**, 57 67 (2012).

44. Adler, A., Zhao, R., Shin, M. E., Yasuda, R. & Gan, W.-B. Somatostatin-Expressing Interneurons Enable and Maintain Learning-Dependent Sequential Activation of Pyramidal Neurons. Neuron **102**, 202-216.e7 (2019).

45. Makino, H. & Komiyama, T. Learning enhances the relative impact of top-down processing in the visual cortex. Nat Neurosci 18, 1116 1122 (2015).

46. Jiang, X. et al. Principles of connectivity among morphologically defined cell types in adult neocortex. Science **350**, aac9462–aac9462 (2015).

47. Hu, H., Gan, J. & Jonas, P. Fast-spiking, parvalbumin+ GABAergic interneurons: From cellular design to microcircuit function. Science **345**, 1255263 (20**14**).

48. Armstrong, C. & Soltesz, I. Basket cell dichotomy in microcircuit function. *J Physiology* **590**, 683–694 (2012).

49. Andrási, T. et al. Differential excitatory control of 2 parallel basket cell networks in amygdala microcircuits. *Plos Biol* **15**, e2001421 (2017).

50. Gonzalez-Burgos, G. & Lewis, D. A. GABA Neurons and the Mechanisms of Network Oscillations: Implications for Understanding Cortical Dysfunction in Schizophrenia. *Schizophrenia Bull* **34**, 944–961 (2008).

51. Vereczki, V. K. et al. Synaptic Organization of Perisomatic GABAergic Inputs onto the Principal Cells of the Mouse Basolateral Amygdala. *Front Neuroanat* **10**, 20 (2016).

52. Szabó, G. G., Holderith, N., Gulyás, A. I., Freund, T. F. & Hájos, N. Distinct synaptic properties of perisomatic inhibitory cell types and their different modulation by cholinergic receptor activation in the CA3 region of the mouse hippocampus. *European J Neurosci* **31**, 2234–2246 (2010).

53. Massi, L. et al. Temporal Dynamics of Parvalbumin-Expressing Axo-axonic and Basket Cells in the Rat Medial Prefrontal Cortex In Vivo. J Neurosci **32**, 16496–16502 (2012).

54. Schneider-Mizell, C. M. et al. Structure and function of axo-axonic inhibition. Elife 10, e73783 (2021).

55. DeFelipe, J., Hendry, S. H. C., Jones, E. G. & Schmechel, D. Variability in the Terminations of GABAergic Chandelier Cell Axons on Initial Segments of Pyramidal Cell Axons in the Monkey Sensory-Motor Cortex. J Comp Neurol 231, 364–384 (1985).

56. Blazquez-Llorca, L. et al. Spatial distribution of neurons innervated by chandelier cells. *Brain Struct Funct* **220**, 2817 2834 (2015).

57. Pan-Vazquez, A., Wefelmeyer, W., Sabater, V. G., Neves, G. & Burrone, J. Activity-Dependent Plasticity of Axo-axonic Synapses at the Axon Initial Segment. Neuron **106**, 265-276.e6 (2020).

58. Grubb, M. S. & Burrone, J. Activity-dependent relocation of the axon initial segment fine-tunes neuronal excitability. Nature **465**, 1070–1074 (2010).

59. Kuba, H., Oichi, Y. & Ohmori, H. Presynaptic activity regulates Na+ channel distribution at the axon initial segment. Nature 465, 1075–1078 (2010).

60. Kuba, H. Structural tuning and plasticity of the axon initial segment in auditory neurons. J Physiology 590, 5571–5579 (2012).

1014   61. Vogels, T. P., Sprekeler, H., Zenke, F., Clopath, C. & Gerstner, W. Inhibitory Plasticity
1015   Balances Excitation and Inhibition in Sensory Pathways and Memory Networks. Science 334,
1016   1569 1573 (2011).
1017   62. Barron, H. C., Vogels, T. P., Behrens, T. E. & Ramaswami, M. Inhibitory engrams in
1018   perception and memory. *Proc National Acad Sci* **114**, 6666 6674 (2017).
1019   63. Montijn, J. S., Vinck, M. & Pennartz, C. M. Population coding in mouse visual cortex:
1020   response reliability and dissociability of stimulus tuning and noise correlation. Front Comput
1021   Neurosc 8, 58 (2014).
1022   64. Etter, G., Manseau, F. & Williams, S. A Probabilistic Framework for Decoding Behavior
1023   From in vivo Calcium Imaging Data. Front Neural Circuit 14, 19 (2020).
1024   65. Komiyama, T. et al. Learning-related fine-scale specificity imaged in motor cortex circuits of
1025   behaving mice. Nature 464, 1182 1186 (2010).
1026   66. Jaqaman, K. et al. Robust single particle tracking in live cell time-lapse sequences. Nat
1027   Methods 5, 695–702 (2008).
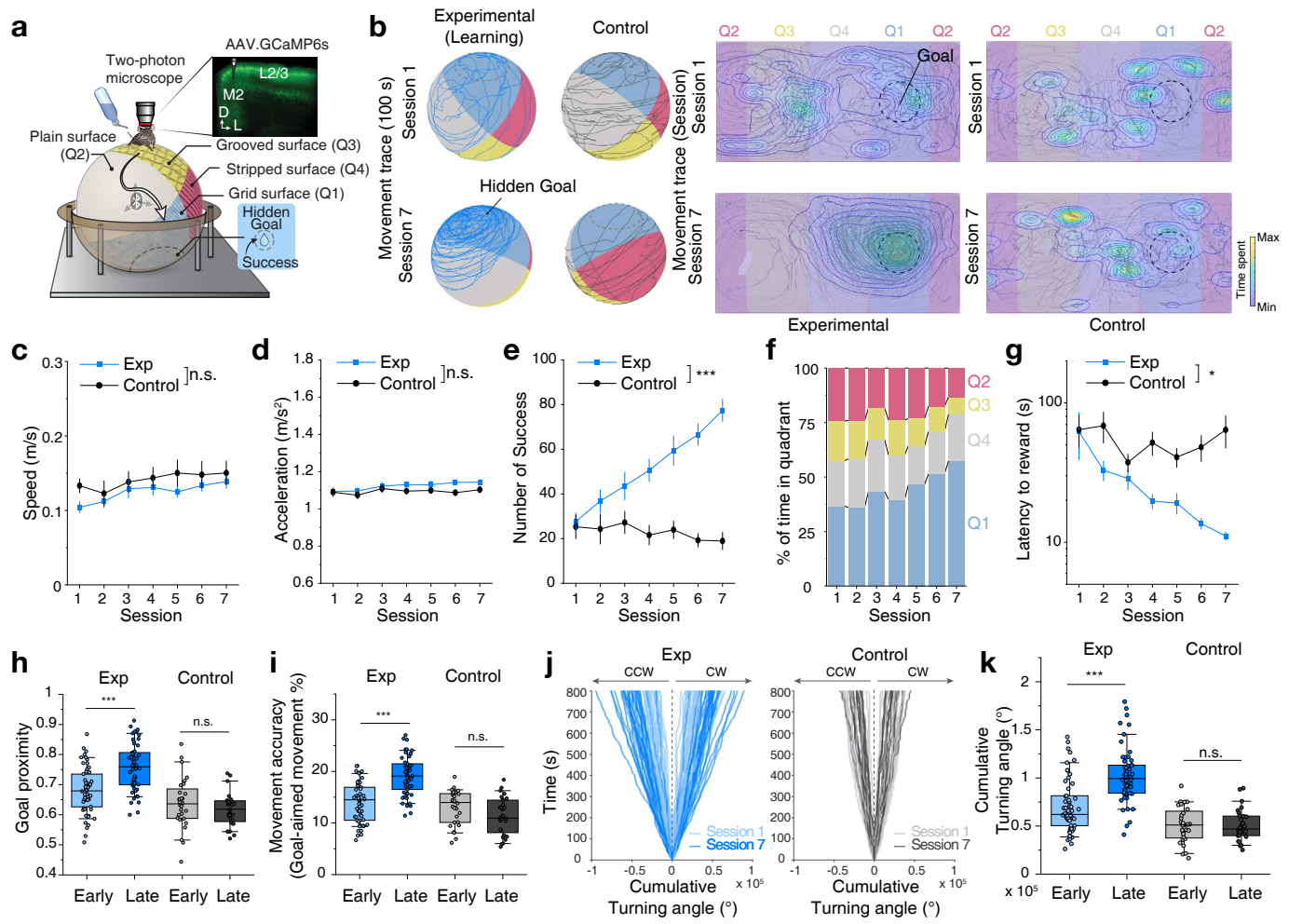
Figure 1 - Jung et al.
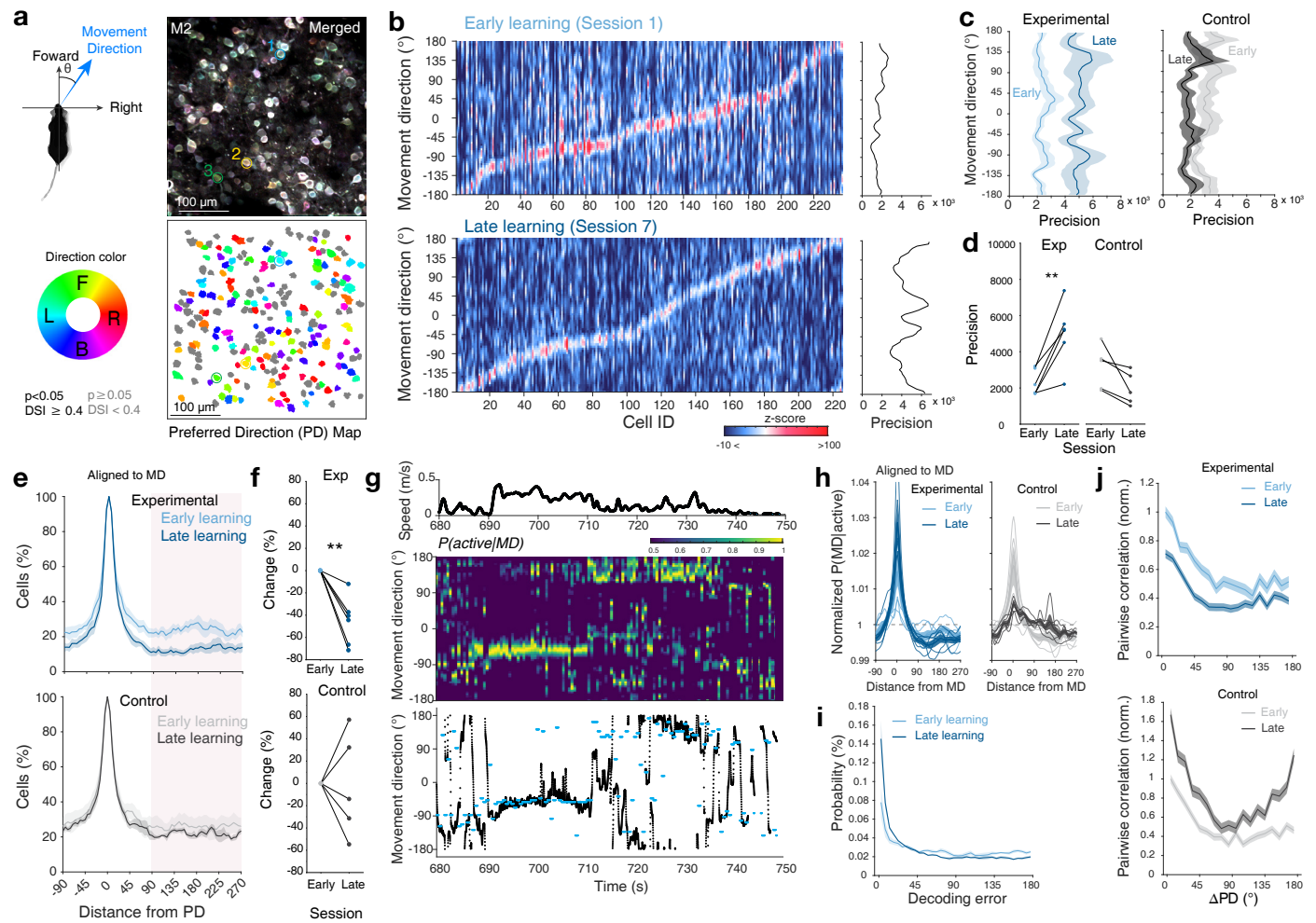
Figure 2 - Jung et al.
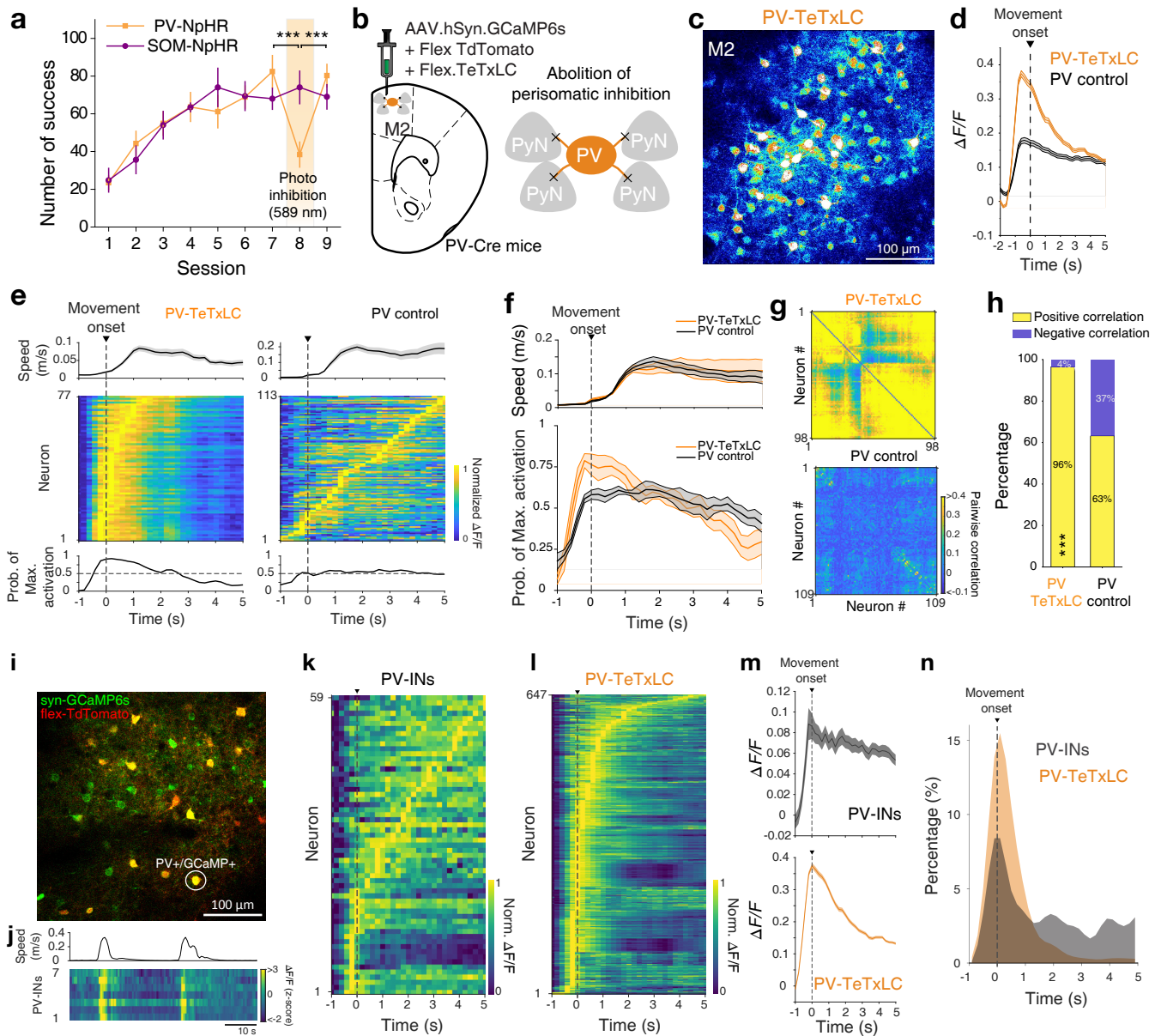
Figure 3 - Jung et al.
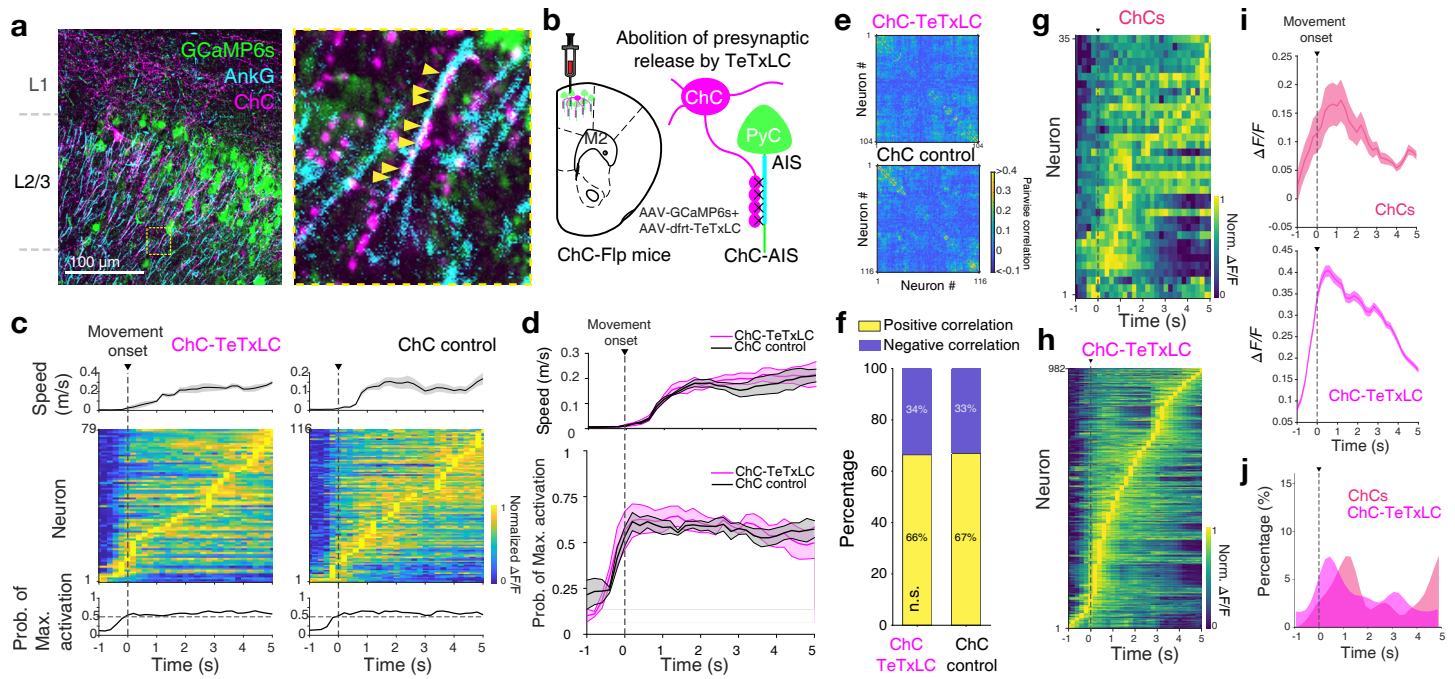
Figure 4 - Jung et al.
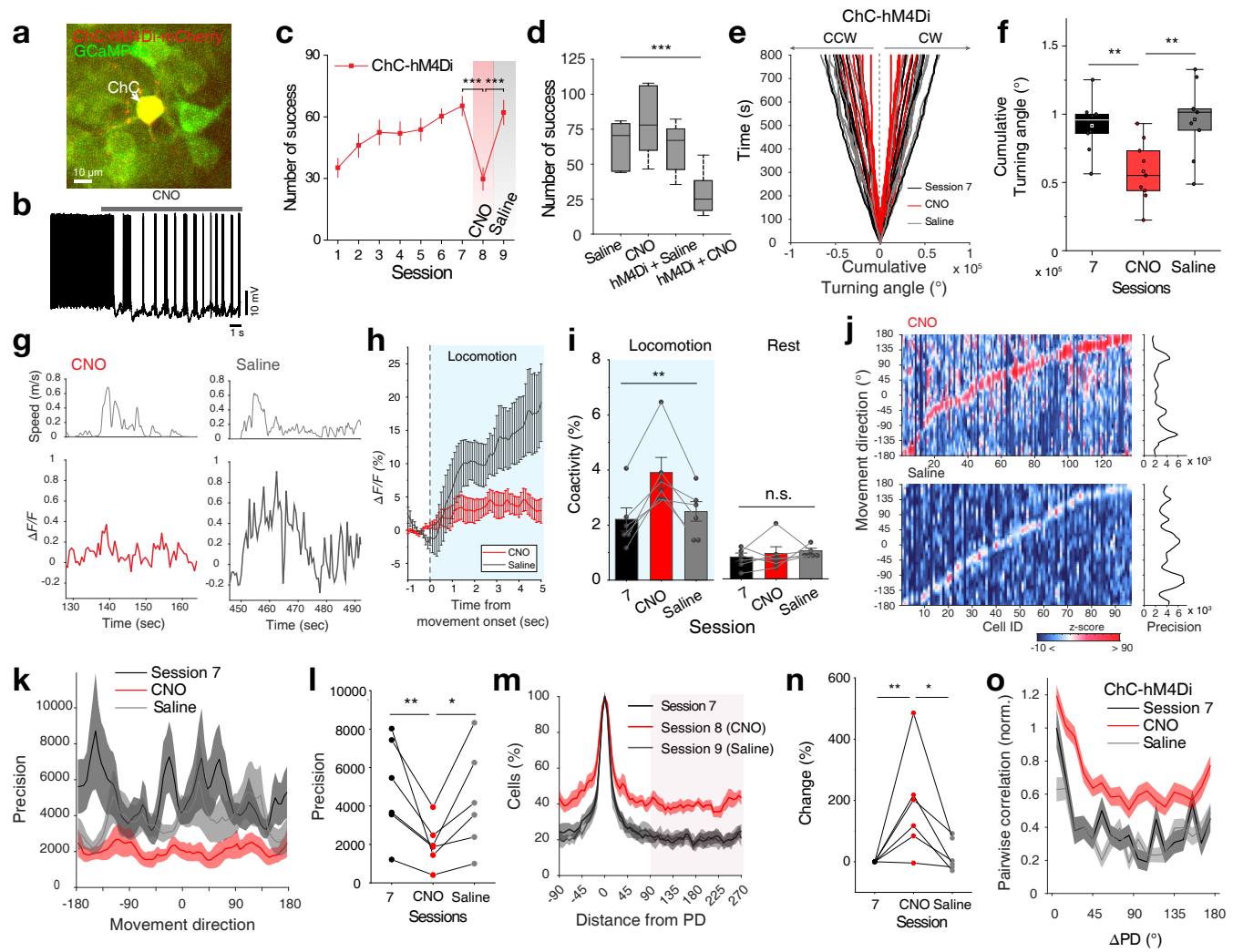
Figure 5 - Jung et al.
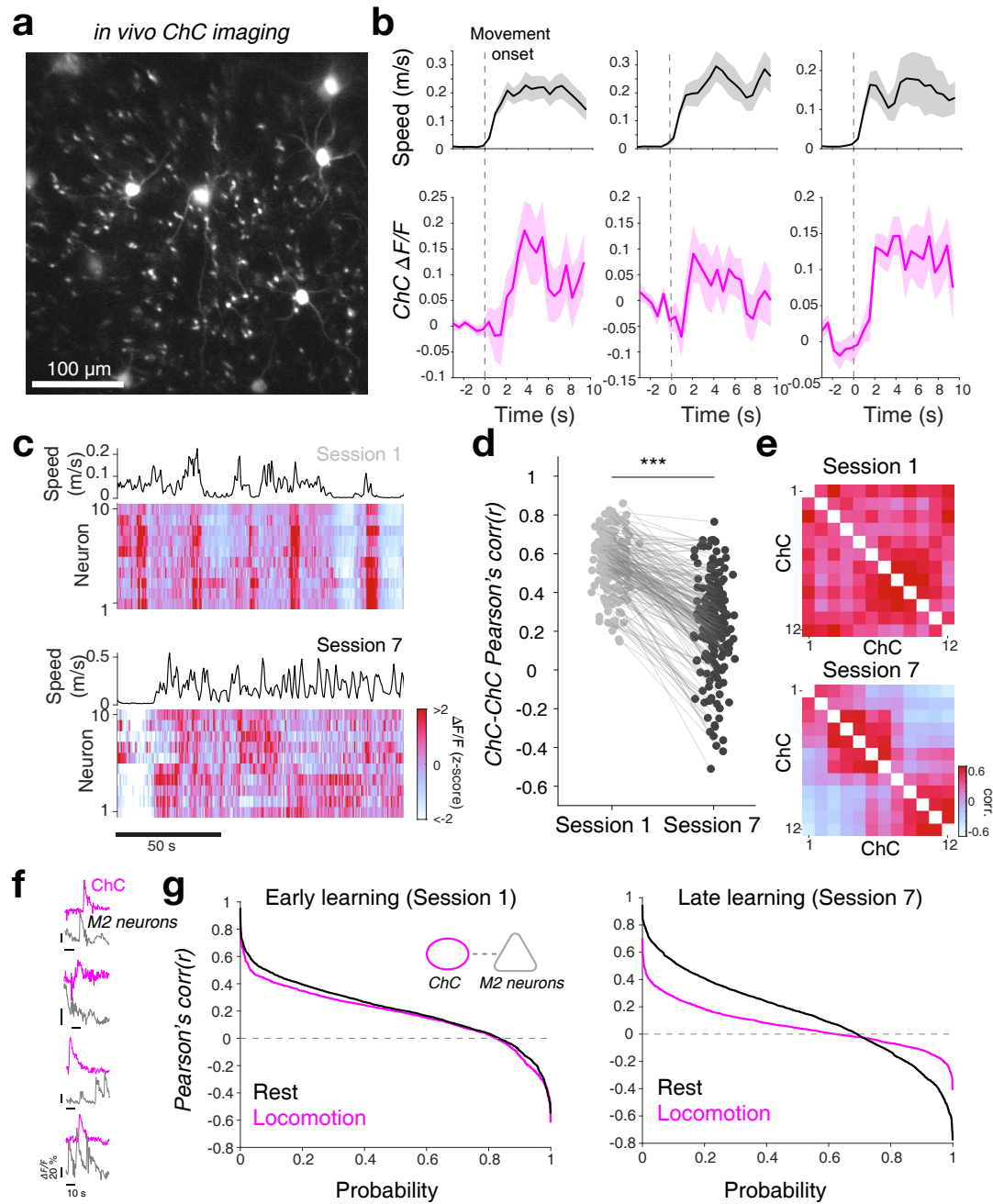
Figure 6 - Jung et al.
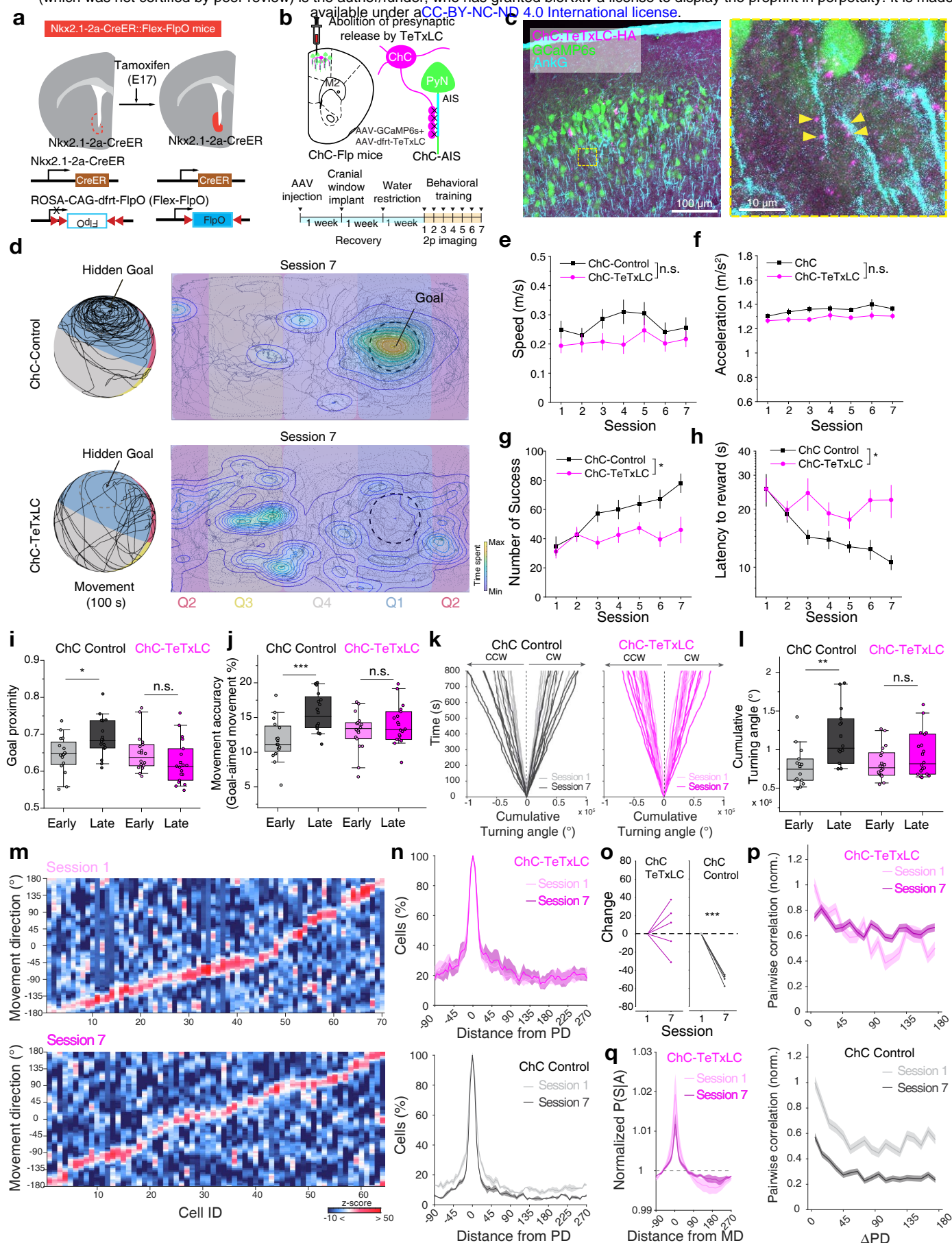
Figure 7 - Jung et al.
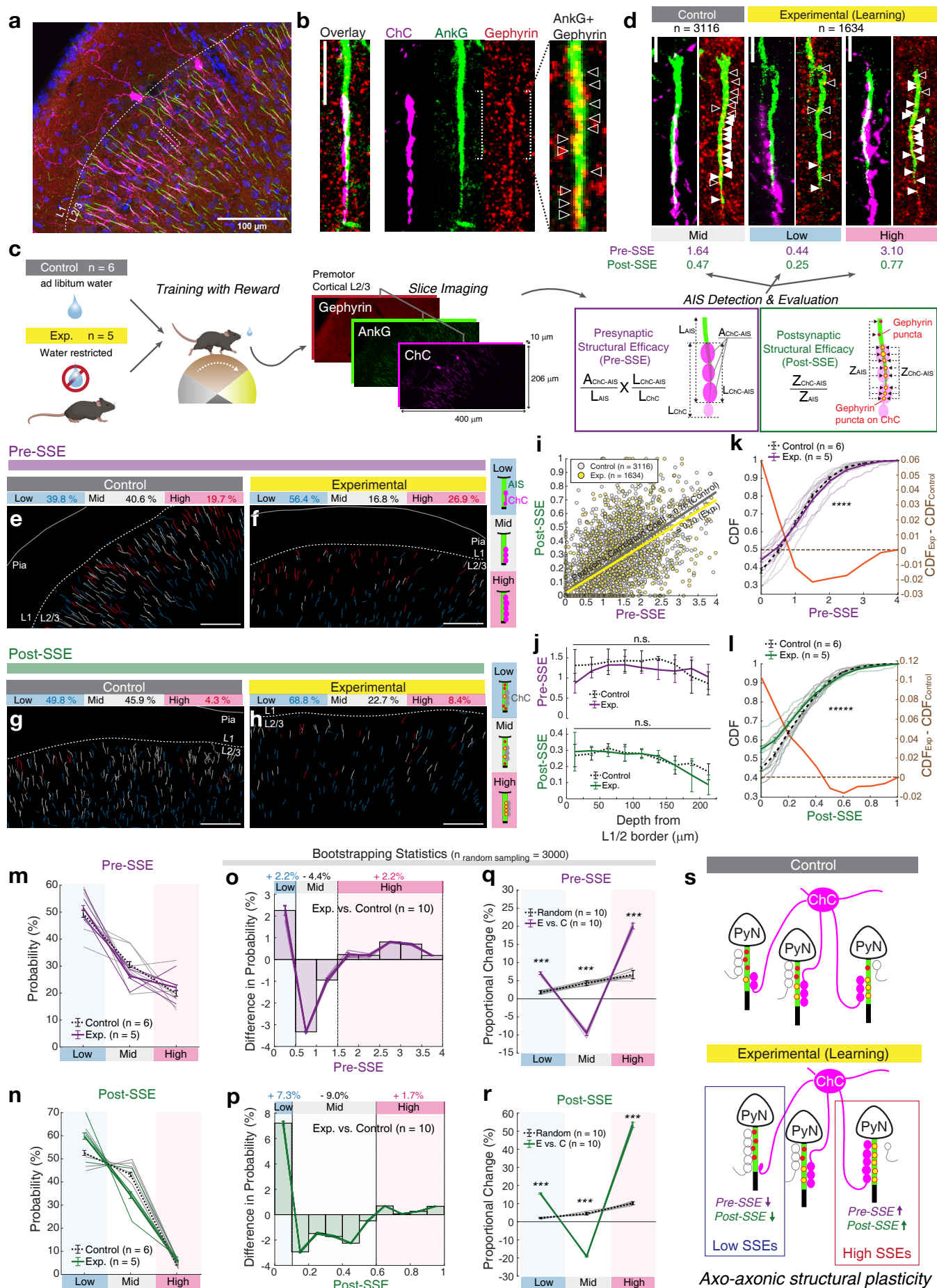
Figure 8 - Jung et al.

Figure 1. Improved directional control of movement with learning.

(a) Schematic of the experimental setup for the spatial navigation task on a multi-textured floating ball maze, in which head-fixed mice freely navigate toward a goal spot in the target quadrant (Q1) based on tactile cues. Two-photon calcium imaging was simultaneously performed to monitor neuronal activity in layer 2/3 of premotor cortex (M2).

(b) Representative movement trace (100 s) of mice in the experimental (water restricted) and control (water *ad libitum*) groups navigating on the multi-textured floating ball maze from training sessions 1 to 7 (left) and their 2-dimensional projection (right). A circle with a dashed line indicates the goal spot on the ball. Heat map and contour lines of the times of mice spent on the location are presented.

(c) Average movement speed (n = 27 for the experimental group; n = 14 for the control group; Two-way repeated measures ANOVA, F_{group} = 3.85, P = 0.072).

(d) Average movement acceleration (two-way repeated measures ANOVA, F_{group} = 154, P = 0.24).

(e) Average number of successes obtained in training sessions (two-way repeated measures ANOVA, F_{group} = 26.97, P = 1.73 × 10^{-4}).

(f) Percentage of time spent in each quadrant during navigation in the experimental group.

(g) Average latency to reward (two-way repeated measures ANOVA, F_{group} = 7.43, P = 0.017).

(h) Average goal proximity (two-tailed *t*-test, t = -5.19, P = 1.03 × 10^{-6} for the experimental group; t = 1.15, P = 0.26 for the control group).

(i) Movement accuracy (two-tailed *t*-test, t = -7.04, P = 2.01 × 10^{-10} for the experimental group; t = 1.74, P = 0.09 for the control group).

(j) Cumulative turning angle over time in Sessions 1 and 7.

(k) Comparison of cumulative turning angle between the experimental and control groups (two-tailed *t*-test, t = -5.84, P = 5.76 × 10^{-8} for the experimental group; t = 0.019, P = 0.98 for the control group). n.s., not significant; **P < 0.01; ***P < 0.001; Error bars indicate s.e.m. In the box plot, the midline, box size, and whisker indicate median, 25-75th percentile, and 10-90th percentile, respectively.

Figure 2. Motor learning refines direction coding through the suppression of irrelevant activity.

(a) Schematic of movement direction based on forward-backward and right-left speeds. An average pixel-based activity map of movement direction (top). Corresponding merged activity map (top). Color-coded map of preferred directions of direction-tuned neurons (von Mises fitting $P < 0.05$ and Direction Selectivity Index, DSI ≥ 0.4), colored according to their preferred direction. Non direction-tuned neurons (DSI < 0.4) are shown in gray.

(b) Example heatmap of premotor neurons' tuning for movement direction in early learning (session 1, top) and late learning (session 7, bottom) from a mouse in the experimental group. Data is sorted from the location of peak likelihood probability P (active|MD). Corresponding precisions of population responses for each movement direction (right).

(c) Changes of average precision curves of population responses in the experimental and control groups.

(d) Comparison of population response precisions between early and late sessions (n = 6 mice for the experimental group, two-tailed paired t-test, $t = -4.98$, $P = 0.0042$; n = 5 mice for the control group, $t = 2.56$, $P = 0.063$).

(e) Normalized percentage of active cells in the population as a function of distance from the preferred direction (n = 830 cells for early, n = 937 cells for late in the experimental group; n = 705 cells for early, n = 671 cells for late in the control group).

(f) Changes in the percentage of active cells from the early session and late session (two-tailed Student's t-test. $t = 5.149$, $P = 0.0036$ for the experimental group; $t = 0.106$, $P = 0.920$ for the control group).

(g) Movement speed (top), posterior probabilities (middle), and corresponding actual and decoded movement direction estimated with maximum a posteriori (MAP)(bottom).

(h) Changes in posterior probabilities normalized by chance level as a function of distance from movement direction with learning for experimental and control groups.

(i) Probability distributions of decoding error.

(j) Pairwise correlations with respect to ΔPD normalized by early session Session 1 (angular difference in preferred direction between neuronal pairs, n = 58456 pairs for Session 1, n = 67143 pairs for Session 7 for the experimental group; n = 51285 pairs for Session 1, n = 45156 pairs for Session 7 for the control group). **$P < 0.01$; Error bars and shading indicate s.e.m.

Figure 3. Manipulation of parvalbumin interneurons alters global neural activity.

(a) Behavioral performance of the number of successes with training and photo-inhibition. Inactivation of PV-INs, but not SOM-INs, resulted in a deficit in finding a hidden goal (n = 8 mice for PV-NpHR, one-way repeated measures ANOVA, $F_{session}$ = 30.0, P = 8.69 × 10^{-6}, Fisher multiple comparisons tests, ***P <0.001; n = 5 mice for SOM-NpHR, $F_{session}$ = 0.42, P = 0.67).

(b) Schematic of blockage of GABA release from PV-INs by expressing tetanus toxin light chain (TeTxLC) in PV-Cre mice.

(c) Representative field of view showing increased excitability of premotor neurons in PV-TeTxLC mice.

(d) Average fluorescence transient traces of premotor neurons aligned to movement onset in PV-TeTxLC and PV control mice (n = 647 cells from 6 mice for PV-TeTxLC, n = 480 cells from 7 mice for PV control).

(e) An example of movement speed (top), the normalized activity of neurons (middle), and probability of maximum neuron activation (bottom) aligned to movement onsets in PV-TeTxLC and PV control.

(f) Average movement speed (top) and corresponding changes of the probability of maximum neuron activation (bottom) aligned to movement onsets.

(g) Example pairwise correlation matrices of PV-TeTxLC and PV control mice.

(h) Population fraction of neuronal pairs with positive and negative correlation (n = 31660 pairs for PV-TeTxLC, n = 24061 pairs for control; chi-square test, χ^2 = 1.03 x 10^4, P = 0)

(i-n) The activity of PV-INs and ChCs related to the movement initiation.

(i) An example field of view showing premotor neurons expressing GCaMP6s (green) and FLEX-tdTomato (red) in PV-Cre mice.

(j) An example trace of movement speed (top) and corresponding heatmap of fluorescent transient signals of PV-INs (bottom).

(k) The normalized activity of PV-INs aligned to movement onset (n = 59 cells from 7 mice in PV control).

(l) The normalized activity of neurons aligned to movement onset in PV-TeTxLC group (n = 647 cells from 6 mice in PV-TeTxLC).

(m) Average ΔF/F traces of PV-INs (top) and neurons in PV-TeTxLC group (bottom) aligned to movement onset.

(n) Probability distributions of peak activity timing of PV-INs and neurons in PV-TeTxLC group

aligned to movement onset. ***$P < 0.001$; Error bars and shadings indicate s.e.m.

Figure 4. Sparse, sequential activity is preserved in ChC manipulation.

(a) Example image of the axonal projection of ChC (expressing mCherry, in magenta) to AISs (post-hoc AnkG staining, in cyan) of neighboring L2/3 neurons (expressing GCaMP6s, in green). Axonal boutons of ChC innervating an AIS (right).

(b) Schematic of selective abolition of GABA release from ChCs in M2 by expressing tetanus toxin light chain (TeTxLC) in ChC-Flp mice (Nkx2.1-2a-CreER::Flex-FlpO mice).

(c) An example of movement speed (top), the normalized activity of neurons (middle), and probability of maximum neuron activation (bottom) aligned to movement onsets in ChC-TeTxLC and ChC control.

(d) Average movement speed (top) and corresponding changes in the probability of maximum neuron activation (bottom) aligned to movement onsets.

(e) Example pairwise correlation matrices of ChC-TeTxLC and ChC control mice.

(f) Population fraction of neuronal pairs with positive and negative correlation (n = 14082 pairs for ChC-TeTxLC, n = 57956 pairs for ChC control; chi-square test, χ^2 = 1.21, P = 0.27).

(g-j) The activity of ChCs related to the movement initiation.

(g) The normalized activity of ChCs aligned to movement onset (35 cells from 3 mice).

(h) The normalized activity of neurons aligned to movement onset in ChC-TeTxLC group (n = 982 cells from 9 mice).

(i) Average ΔF/F traces of ChCs (top) and neurons in ChC-TeTxLC group (bottom) aligned to movement onset.

(j) Probability distributions of peak activity timing of ChCs and neurons in ChC-TeTxLC group aligned to movement onset. Error bars and shading indicate s.e.m.

Figure 5. Chemogenetic inhibition of ChCs disrupts the direction selectivity of premotor neurons.

(a) *In vivo* two-photon Ca^{2+} imaging of ChC and neighboring neurons expressing GCaMP6s and selectively coexpressing the chemogenetic silencer, hM4Di-mCherry in ChCs of M2, in a FlpO-dependent manner, using Nkx2.1-2a-CreER::Flex-FlpO mice (ChC-Flp mice).

(b) Bath-application of CNO (10 µM) reduced the firing rate in ChCs expressing hM4Di. Example whole-cell current-clamp recording from a ChC.

(c) The average number of successes of ChC-hM4Di mice increased with training from sessions 1 to 7. CNO or saline injection was followed on session 8 or 9 (n = 9 mice, one-way repeated measures ANOVA, $F_{session}$ = 31.1, $P = 3.09 \times 10^{-6}$; Fisher multiple comparisons tests, Session 7 vs. CNO, $P = 2.40 \times 10^{-6}$; Session CNO vs. Saline, $P = 7.58 \times 10^{-6}$; Session 7 vs. Saline, $P = 0.53$).

(d) Cumulative turning angle of ChC-hM4Di mice across conditions (Session 7, CNO, and Saline).

(e) Comparison of cumulative turning angle across conditions (one-way repeated measures ANOVA, $F_{session}$ = 8.99, $P = 0.0024$; Fisher multiple comparisons tests, Session 7 vs. CNO, $P = 0.0035$; Session 7 vs. Saline, $P = 0.45$; CNO vs. Saline, $P = 0.0013$).

(f) Comparison of performance between conditions (n = 6 mice for ChC control with saline or CNO; n = 9 mice for ChC-hM4Di with saline or CNO, one-way ANOVA F_{sesion} = 9.50, $P = 2.06 \times 10^{-4}$).

(g) Example ΔF/F traces of a ChC during locomotion in CNO and saline conditions.

(h) Average ΔF/F traces of ChCs aligned to movement onset in CNO and saline conditions (6 ChC-hM4Di mice, n = 497 cells for CNO; n = 475 cells for Saline).

(i) Coactivity percentage of neurons during periods of locomotion and rest. Friedman test, χ^2 = 9.33, $P = 0.0094$ for locomotion; χ^2 = 0.33, $P = 0.85$ for rest).

(j) Example tuning curves of each individual premotor neuron for movement direction in CNO (top) and saline (bottom) conditions. Data is sorted from the location of peak likelihood probability P (active|MD). Corresponding precisions of population responses for each movement direction (right).

(k) Average precision curves of population responses across later learning, CNO, and saline conditions.

(l) Comparison of population response precisions (one-way repeated measures ANOVA, $F_{session}$ = 6.39, $P = 0.016$; Fisher multiple comparisons tests, Session 7 vs. CNO, $P = 0.0068$; Session CNO vs. Saline, $P = 0.023$; Session 7 vs. Saline, $P = 0.486$).

(m) Normalized percentage of active cells in the population as a function of distance from the preferred direction (n = 484 cells for Session 7, n = 407 cells for CNO, n = 582 cells for Saline).

(n) Changes in the percentage of active cells across conditions. (one-way repeated measures ANOVA with Greenhouse-Geisser correction, $F_{session}$ = 7.26, P = 0.037; Fisher multiple comparisons tests, Session 7 vs. CNO, P = 0.006; Session CNO vs. Saline, P = 0.011; Session 7 vs. Saline, P = 0.372).

(o) Pairwise correlations with respect to ΔPD normalized by Session 7 (angular difference in preferred direction between neuronal pairs, n = 17933 pairs for Session 7; n = 24205 pairs for CNO; n = 22735 pairs for Saline). *P < 0.05, **P < 0.01, ***P < 0.001; n.s., not significant. Error bars and shading indicate s.e.m.

Figure 6. Variability of ChC activity increases during learning

(a) Representative field of view showing L2 ChCs expressing GCaMP6.

(b) Example fluorescence traces from ChCs that show increased activity during episodes of locomotion.

(c) Movement speed and the heatmap of ChC activity during navigation in Sessions 1 and 7. The synchronous activation of ChCs in Session 1 during locomotion becomes asynchronous in Session 7.

(d) Changes in the correlation between a pair of ChCs (n = 189 pairs in 35 cells from 3 mice; two-tailed Wilcoxon signed-rank test, $Z = 11.92$, $P = 0$).

(e) Example correlation matrix of ChCs in Sessions 1 and 7.

(f) Various temporal relationships between activity of ChC (magenta) and neighboring M2 neurons (gray) during episodes of locomotion in ChC-hM4Di mice.

(g) Cell-to-cell Pearson's correlation of ChC-M2 neuron pairs in Sessions 1 and 7 during locomotion and rest epochs (4 ChC-hM4Di mice, n = 1957 pairs for Session 1, two-sample Kolmogorov-Smirnov test, locomotion vs. rest, $D = 0.076$, $P = 2.12 \times 10^{-5}$; n = 1723 pairs for Session 7, locomotion vs. rest, $D = 0.28$, $P = 4.049 \times 10^{-61}$). ***$P < 0.001$; Shadings indicate s.e.m.

Figure 7. Abolition of ChC inhibition impairs the refinement of direction coding and motor learning.

(a) Generation of Nkx2.1-2a-CreER::Flex-FlpO mice. Nkx2.1-2a-CreER::Flex-FlpO mice were generated by crossing Nkx2.1-2a-CreER with ROSA-Flex-FlpO mouse lines. The targeting vector containing Rosa26 homology arms, a CAG promoter, and a FLEX-Flp cassette was constructed. Similar to Nkx2.1-2a-CreER::Ai14, tamoxifen was administered to timed pregnant SW females by oral gavage at E17.

(b) Schematic for selective abolition of GABA release from ChCs in M2 by expressing tetanus toxin light chain (TeTxLC) in ChC-Flp mice (Nkx2.1-2a-CreER::Flex-FlpO mice).

(c) A representative image of ChC expressing TeTxLC-HA and neighboring neurons expressing GCaMP6s in Nkx2.1-2a-CreER::Flex-FlpO mice (left) and post-hoc validation of ChCs' axonal projection to the axon-initial segment of neighboring PyNs by Ankyrin G (AnKG) staining (right).

(d-l) Behavioral impact of ChC manipulation.

(d) Representative movement traces of mice in ChC-TeTxLC (top) and ChC control (bottom) navigating on the multi-textured floating ball maze in Session 7 (100 s). A circle with a dashed line indicates the goal spot on the ball. Heat map and contour lines of the times of mice spent on the location are presented.

(e) Average movement speed (n = 9 for ChC-TeTxLC group; n = 8 for ChC control group; two-way repeated measures ANOVA, F_{group} = 1.16, P = 0.32).

(f) Average movement acceleration (two-way repeated measures ANOVA, F_{group} = 3.22, P = 0.12).

(g) Average number of successes obtained in training sessions (two-way repeated measures ANOVA, F_{group} = 8.54, P = 0.011).

(h) Average latency to reward (two-way repeated measures ANOVA, F_{group} = 7.81, P = 0.0136).

(i) Average goal proximity (two-tailed t-test, t = -2.52, P = 0.017 for ChC control; t = 1.38, P = 0.177 for ChC-TeTxLC).

(j) Movement accuracy (two-tailed t-test, t = -3.80, P = 6.67×10^{-4} for ChC control; t = -1.05, P = 0.30 for ChC-TeTxLC).

(k) Cumulative turning angle over time in Sessions 1 and 7.

(l) Comparison of cumulative turning angle between ChC-TeTxLC and ChC control (two-tailed t-test, t = -3.21, P = 3.51×10^{-3} for ChC control; t = -1.39, P = 0.175 for ChC-TeTxLC).

(m) Example tuning curves of individual premotor neurons for movement direction in sessions 1 (top) and 7 (bottom) in ChC-TeTxLC condition. Data is sorted from the location of peak likelihood probability P(active|MD).

(n) Normalized percentage of active cells in the population as a function of distance from the preferred direction (n = 378 cells for Session 1, n = 416 cells for Session 7 in ChC-TeTxLC; n = 665 cells for Session 1, n = 689 cells for Session 7 in ChC control).

(o) Changes of the percentage of active cells from Session 1 to Session 7 (n = 5 mice for ChC-TeTxLC, two-tailed Student's t-test. $t = 0.56$, $P = 0.606$; n = 4 mice for ChC control. $t = -17.97$, $P = 3.76 \times 10^{-4}$).

(p) Pairwise correlations with respect to ΔPD normalized by Session 1 (angular difference in preferred direction between neuronal pairs, n = 15028 pairs for Session 1, n = 22589 pairs for Session 7 for ChC-TeTxLC; n = 57956 pairs for Session 1, n = 62104 pairs for Session 7 for ChC control).

(q) Changes in posterior probabilities normalized by chance level as a function of distance from MD with learning. n.s., not significant; *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; Error bars and shading indicate s.e.m. In the box plot, the midline, box size, and whisker indicate median, 25-75th percentile, and 10-90th percentile, respectively.

Figure 8. Heterogeneous axo-axonic synaptic plasticity underlies organized motor control.

(a, b) Immunostaining of the axon initial segment (AIS) by AnkG (green), and inhibitory postsynaptic gephyrin (red) visualizes the corresponding synaptic composition (yellow) of ChC (magenta) to AIS contacts. DAPI stains nuclei (blue).

(c) Schematics of automated detection of ChC-innervated AISs (ChC-AISs).

(d) Representative examples of ChC-AIS evaluation of presynaptic and postsynaptic structural efficacy (Pre-SSE and Post-SSE). Filled and empty arrows indicate ChC-AIS and non-ChC-AIS synapses respectively. SSE values were classified into three subgroups (high: > 1.5 for Pre-SSE, > 0.6 for Post-SSE; low: < 0.5 for Pre-SSE, < 0.1 for Post-SSE; mid between high and low).

(e-h) Examples of Pre-SSE and Post-SSE distribution in each slice for the experimental (learning) group and control group.

(i) Scatter plot of Pre-SSEs and Post-SSEs for each group.

(j) SSEs by the position of AISs within the cortical layer 2 for each group, which is measured by the depth from the border of layers 1 and 2.

(k, l) CDFs from individual mice ($P = 5 \times 10^{-5}$ for **k** and $P = 2 \times 10^{-11}$ for **l**, error bars = s.e.m., left axis) and the difference between the averages for the experimental and control condition (Exp. vs. control, right axis).

(m, n) Probabilities by Pre-SSE and Post-SSE strength subgroups (high, mid, and low) from individual mice. Error bars indicate s.e.m.

(o, p) Differences of probability distributions for Pre-SSE and Post-SSE (Exp. vs. control) and

(q, r) proportional changes (Exp. vs. Control) in the probability of each subgroup which is confirmed by 10 independent robust random samplings (n = 3000 each). In both SSEs, high and low subgroups were increased. The results were compared to proportional changes between unlabeled random pairs. For bootstrapping statistics (**o-r**), the mean and error of each bin were calculated from ten distributions generated by independent robust random sampling (n = 3000). For each random sampling, the probability distribution of a mouse was randomly selected from each condition (Exp. and Control) and used to calculate differences. Error bars indicate S.D., $P = 2.0 \times 10^{-4}$ for all bins.

(s) Model for axo-axonic structural plasticity by heterosynaptic competition. The CDFs were tested

by two-sample Kolmogorov-Smirnov test and the other distributions were tested by the two-tailed

Wilcoxon-Mann-Whitney test. Scale bar = 100 µm (**a, e-h**), 5 µm (**b, d**). Every fluorescent image is

presented by maximum intensity projection of the corresponding volumetric stack with

pseudo-colors. ****P* < 0.001; *****P* < 0.0001; ******P* < 0.00001.
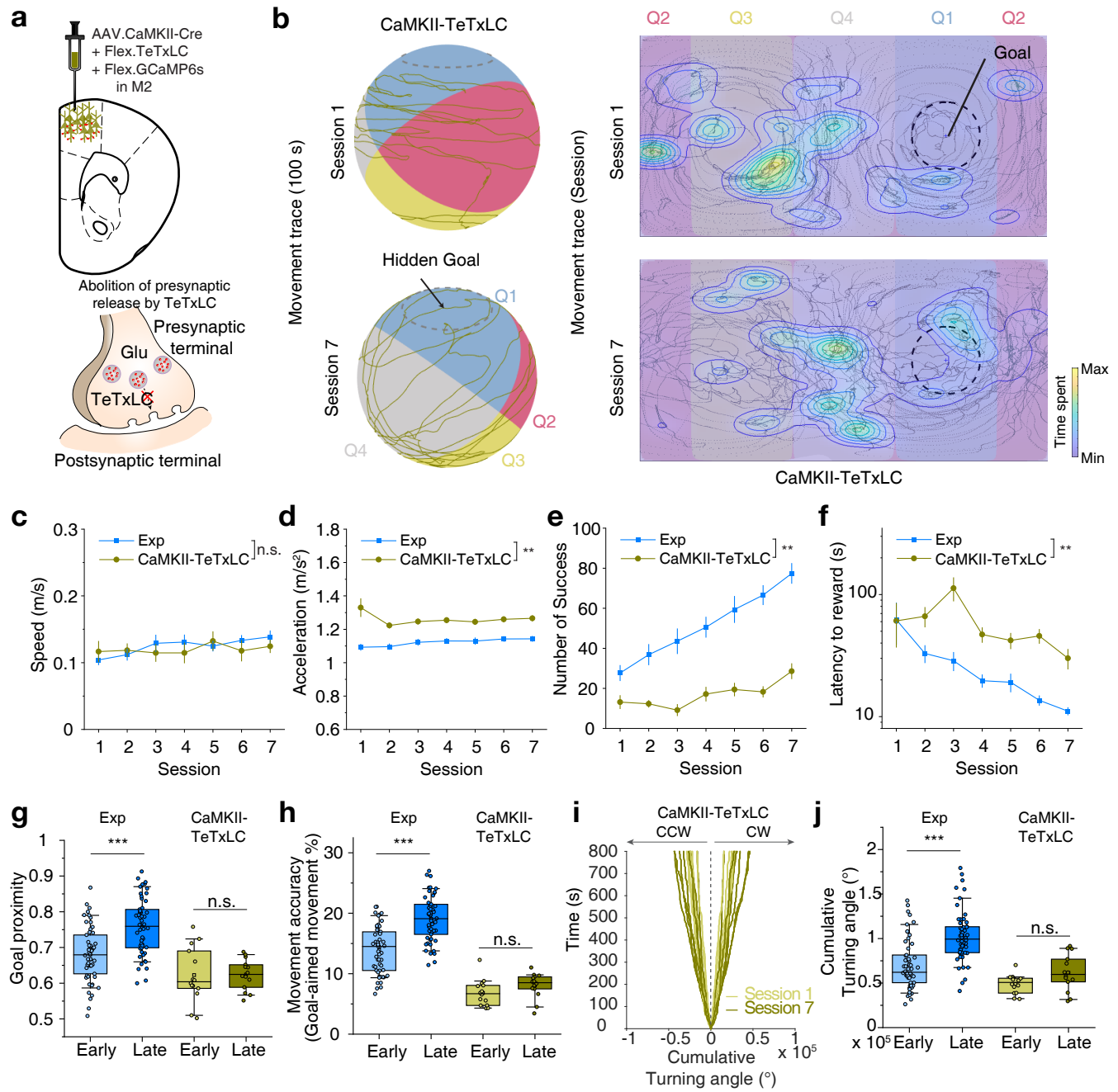
EDF 1 - Jung et al.

Extended Data Figure 1. The premotor cortex is required for organized purposive motor control.

(a-j) Blockage of local glutamate release in the premotor cortex impaired organized motor control.

(a) Schematic of virus injection of tetanus toxin light-chain, which blocks local glutamate transmission from excitatory neurons in the premotor cortex by co-expressing AAV1.CaMKII-Cre and AAV1.Flex-tetanus toxin light-chain for CaMKII-TeTxLC group.

(b) Representative movement trace (100 s) of a mouse in CaMKII-TeTxLC group on the ball (left) from training sessions 1 to 7 and its 2-dimensional projection (right).

(c) Average movement speed of CaMKII-TeTxLC (n = 27 for the experimental group, Exp; n = 7 for CaMII-TeTxLC; two-way repeated measures ANOVA, $F_{group} = 0.45$, $P = 0.53$).

(d) Average movement acceleration (two-way repeated measures ANOVA, $F_{group} = 19.59$, $P = 0.0044$).

(e) Average number of successes obtained in training sessions (two-way repeated measures ANOVA, $F_{group} = 24.11$, $P = 0.0027$).

(f) Average latency to reward of CaMKII-TeTxLC (two-way repeated measures ANOVA, $F_{group} = 24.11$, $P = 3.50 \times 10^{-4}$).

(g) Average goal proximity of CaMKII-TeTxLC (two-tailed t-test, $t = -5.19$, $P = 1.03 \times 10^{-6}$ for Exp; $t = -0.06$, $P = 0.95$ for CaMKII-TeTxLC).

(h) Movement accuracy of CaMKII-TeTxLC (two-tailed t-test, $t = -7.04$, $P = 2.01 \times 10^{-10}$ for Exp; $t = -1.54$, $P = 0.135$ for CaMKII-TeTxLC).

(i) Cumulative turning angle of CaMKII-TeTxLC over time in Sessions 1 and 7.

(j) Comparison of cumulative turning angle between Exp and CaMKII-TeTxLC (two-tailed t-test, $t = -5.84$, $P = 5.76 \times 10^{-8}$ for Exp; $t = -2.05$, $P = 0.050$ for CaMKII-TeTxLC). n.s., not significant; **$P < 0.01$; ***$P < 0.001$; Error bars indicate s.e.m. In the box plot, the midline, box size, and whisker indicate median, 25-75th percentile, and 10-90th percentile, respectively.

EDF 2 - Jung et al.
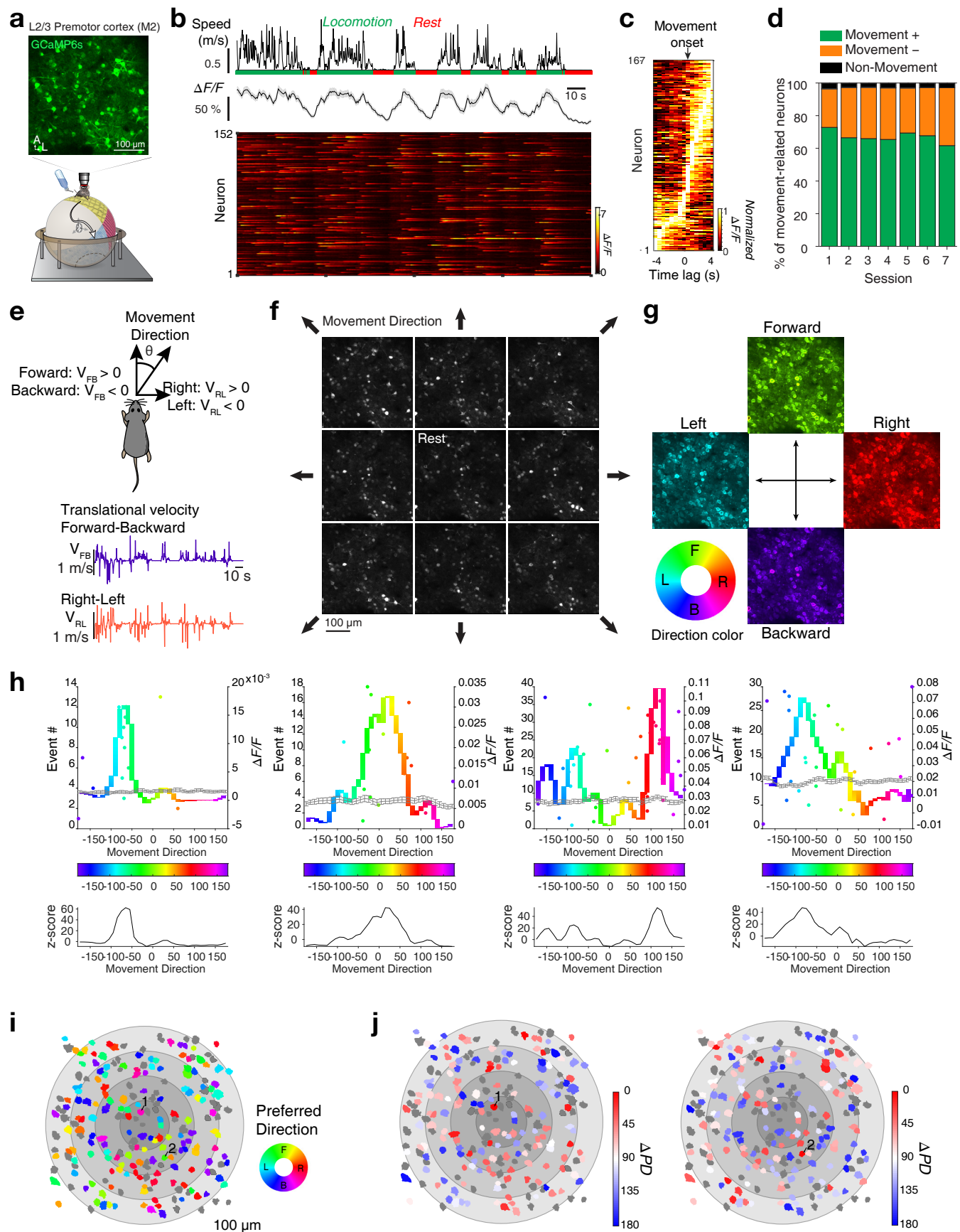
Extended Data Figure 2. Movement direction selective responses of premotor neurons.

(a) Schematic of the navigation task on ball maze with 2-photon calcium imaging and a representative field of view of layer 2/3 premotor neurons expressing GCaMP6.

(b) An example trace of mouse movement speed (top), aligned averaged fluorescence transients (middle), and a corresponding heat-map raster plot (bottom).

(c) Normalized fluorescence transients of premotor neurons aligned to movement onsets.

(d) Summary graph of movement-related neurons throughout training.

(e) Schematic that depicts the estimation of movement direction based on forward-backward and right-left speeds.

(f) Average fluorescence imaging frames during responses to varying movement directions.

(g) A color-coded, pixel-based map of neuron activity with respect to movement direction tuning.

(h) Top, four example calcium transient events by movement direction (dots) and direction tuning curves of premotor neurons (top, colored-curve). Movement direction is color-coded. The gray line indicates the average tuning curve from shuffled data. Bottom, corresponding z-scores of the actual tuning curves normalized by the tuning curves of the shuffled data.

(i) An exemplary preferred direction map. The color indicates the preferred direction of individual cells. Black and gray color indicate non-direction-tuned neurons. The salt-and-pepper layout of color indicates functional heterogeneity in direction-selectivity in the premotor cortex.

(j) Two examples of the heterogeneous spatial distribution of ΔPD (angular difference in preferred direction) between neuronal pairs. The same neuronal population as in **i,** with color-coding by ΔPD (angular difference in preferred direction between reference neuron and another direction-tuned neuron). Numbered neurons indicate reference neurons.
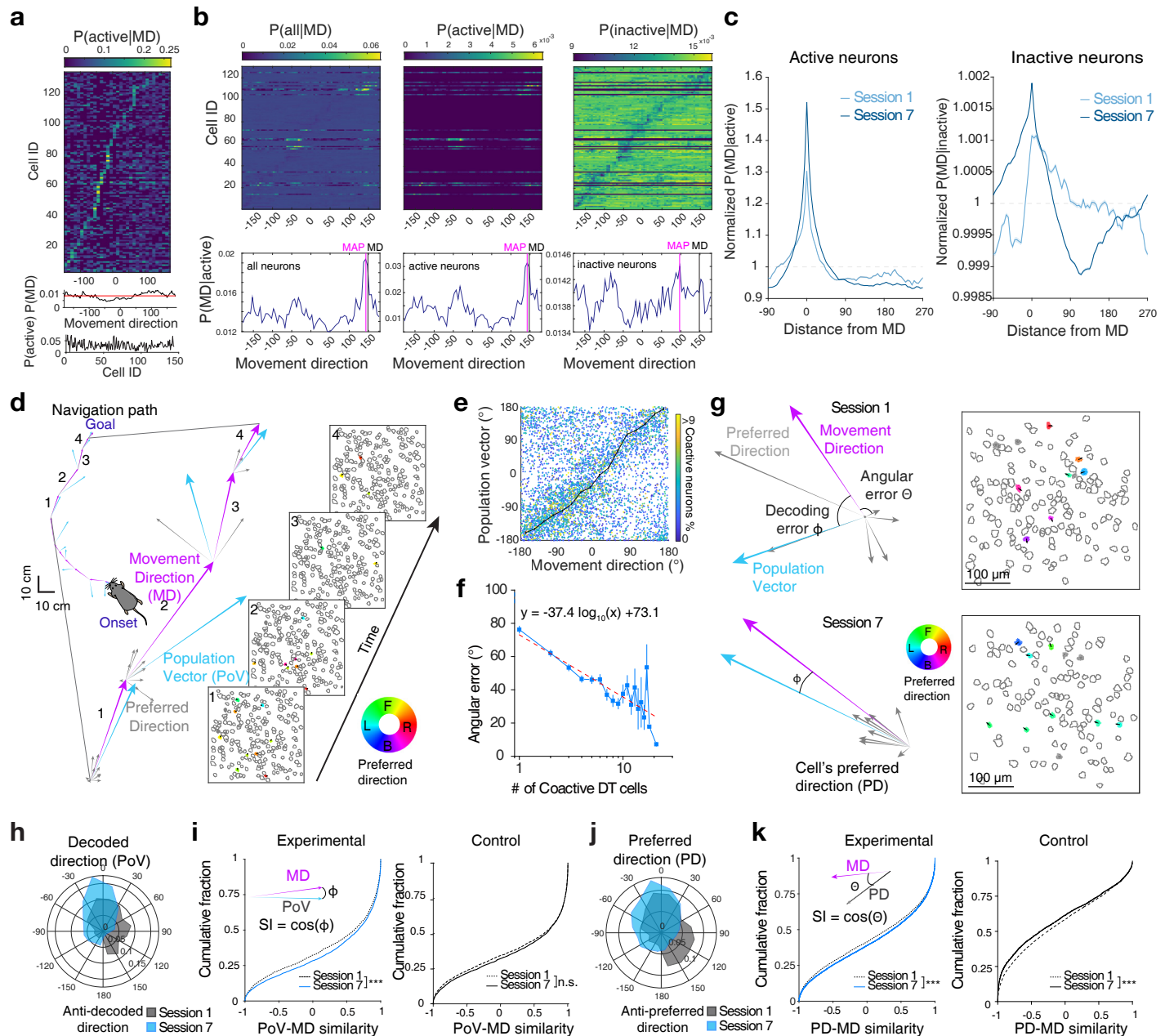
EDF 3 - Jung et al.

Extended Data Figure 3. Population coding of movement direction in premotor neurons.

(a) Movement direction tuning curves for each individual premotor neuron sorted from peak probability location (top). The prior probability of movement direction (middle). The marginal likelihood of being active (bottom).

(b) Tuning curves of neurons corresponding to movement direction at a given moment (top) and posterior probability of movement direction (bottom) given activity from all neurons (left) from active neurons (center), and from inactive neurons (left). Actual movement direction (MD) and decoded movement direction estimated with maximum a posteriori (MAP) are shown.

(c) An example of changes in posterior probabilities, *P(MD|A),* normalized by chance level (dotted line) from a mouse of the experiential group, for active (left) and inactive neurons (right) as a function of distance from movement direction with learning.

(d) Sparse population coding of movement direction during navigation; animal's movement direction (MD, magenta), preferred directions of individual active direction-tuned neurons (PD, gray), and population vector (PoV, the vector sum of the preferred directions, blue) are superimposed. Corresponding maps of active neurons during movement.

(e) Comparison between movement direction and sparse population vector. Each dot is color-coded by the percentage of coactive neurons in the population.

(f) The angular error between the population vector and actual movement direction as a function of the number of coactive direction-tuned (DT) cells. The dotted line with a negative slope coefficient indicates a linear fit of the data.

(g) An example of population coding of movement direction in session 1 (top) and session 7 (bottom). Magenta, cyan, and gray arrow lines indicate an animal's movement direction, population vector direction, and active neurons' preferred direction, respectively (left). Preferred direction map of the active population (right, filled ROIs). The filled color and ROI arrow indicate the corresponding preferred direction.

(h) Polar distribution of angular errors between PoV and MD in the experimental group (session 1, gray; session 7, light blue).

(i) Changes of the cumulative distribution of similarity index between PoV and MD with learning (Two-sample Kolmogorov-Smirnov test, D = 0.0495, *P* = 0.0003 for the experimental group, n = 6 mice; D = 0.027, *P* = 0.12 for the control group, n = 5 mice).

(j) Polar distribution of angular errors between PD and MD in the experimental group (session 1, gray; session 7, light blue).

(k) Changes of the cumulative distribution of similarity index between active neurons' PD and MD with learning (Two-sample Kolmogorov-Smirnov test, D = 0.0533, $P < 2.68 \times 10^{-10}$ for the experimental group, n = 6 mice; D = 0.0584, $P < 2.61 \times 10^{-10}$ for the control group, n = 5 mice). ***P < 0.001; n.s., not significant. Error bars indicate s.e.m.
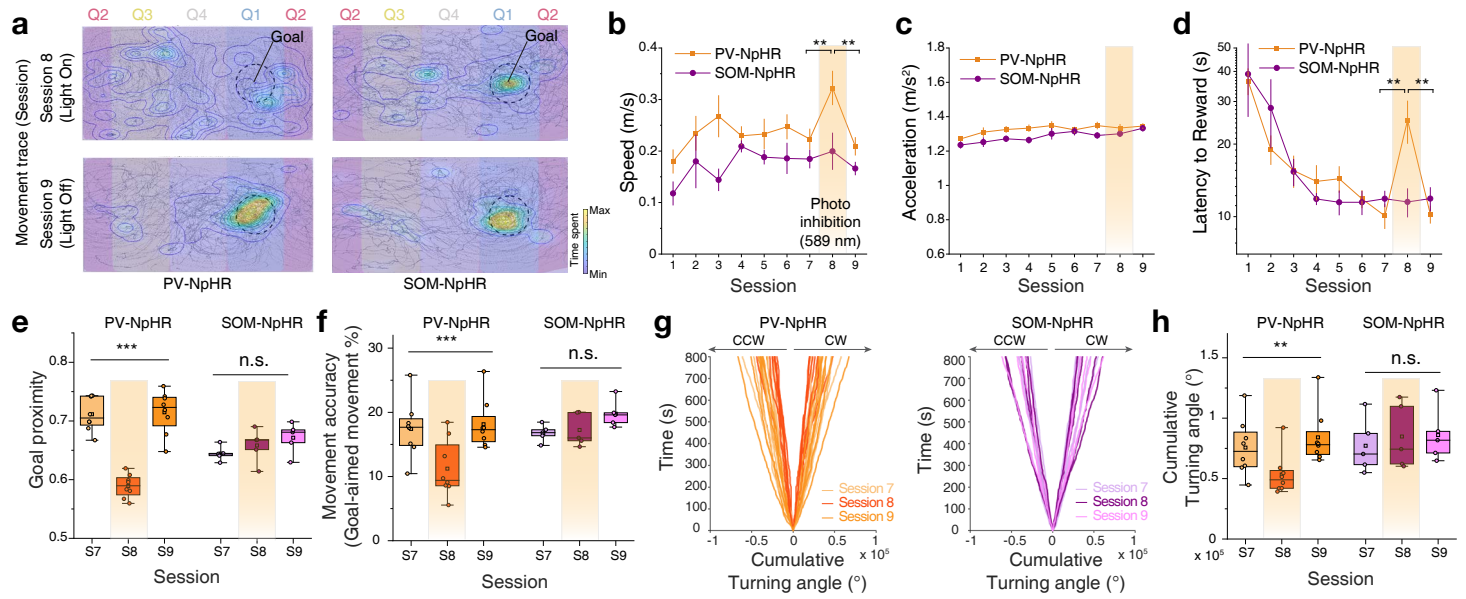
EDF 4 - Jung et al.

Extended Data Figure 4. Silencing perisomatic inhibition disrupts organized motor control.

(a) 2-dimensional projections of representative movement traces of mice on the ball in PV-NpHR and SOM-NpHR groups with photoinhibition (Session 8) and without photoinhibition (Session 9). Blockage of local glutamate release in the premotor cortex impaired organized motor control.

(b) Average movement speed of PV-NpHR and SOM-NpHR (n = 8 mice for PV-NpHR, one-way repeated measures ANOVA with Greenhouse-Geisser correction for Sessions 7, 8, and 9, $F_{session}$ = 10.16, P = 0.0019, Fisher multiple comparisons tests, Session 7 vs. 8, P = 0.003; Session 7 vs. 9, P = 0.61, Session 8 vs. 9, P = 0.001; n = 5 mice for SOM-NpHR, $F_{session}$ = 1.08, P = 0.38).

(c) Average movement acceleration (PV-NpHR, one-way repeated measures ANOVA for sessions 7, 8, and 9, $F_{session}$ = 0.45, P = 0.65; SOM-NpHR, $F_{session}$ = 3.69, P = 0.07).

(d) Average latency to reward (PV-NpHR, one-way repeated measures ANOVA with Greenhouse-Geisser correction for sessions 7, 8, and 9, $F_{session}$ = 9.55, P = 0.014, Fisher multiple comparisons tests, Session 7 vs. 8, P = 0.002; Session 7 vs. 9, P = 0.97, Session 8 vs. 9, P = 0.0021; SOM-NpHR, $F_{session}$ = 0.094, P = 0.84).

(e) Average goal proximity of PV-NpHR and SOM-NpHR in Sessions 7, 8, and 9 (PV-NpHR, one-way repeated measures ANOVA with Greenhouse-Geisser correction, $F_{session}$ = 40.67, P = 3.05 × 10^{-5}, Fisher multiple comparisons tests, Session 7 vs. 8, P = 2.12 × 10^{-6}; Session 7 vs. 9, P = 0.84, Session 8 vs. 9, P = 1.56 × 10^{-6}; SOM-NpHR, $F_{session}$ = 1.57, P = 0.28).

(f) Movement accuracy of PV-NpHR and SOM-NpHR in Sessions 7, 8, and 9 (PV-NpHR, one-way repeated measures ANOVA with Greenhouse-Geisser correction, $F_{session}$ = 13.9, P = 4.72 × 10^{-4} Fisher multiple comparisons tests, Session 7 vs. 8, P = 7.39 × 10^{-4}; Session 7 vs. 9, P = 0.62, Session 8 vs. 9, P = 2.84 × 10^{-4}; SOM-NpHR, $F_{session}$ = 3.06, P = 0.14).

**(g)** Cumulative turning angle of PV-NpHR and SOM-NpHR over time.

(h) Comparison of cumulative turning angle between PV-NpHR and SOM-NpHR (PV-NpHR, Friedman test for Sessions 7, 8, and 9, $\chi^2(2)$ = 13, P = 0.0015, Dunn's multiple comparisons tests, Session 7 vs. 8, P = 0.037; Session 7 vs. 9, P = 0.95, Session 8 vs. 9, P = 0.0014; SOM-NpHR, $\chi^2(2)$ = 2.8, P = 0.25). n.s., not significant; **P < 0.01; ***P < 0.001; Error bars indicate s.e.m. In the box plot, the midline, box size, and whisker indicate median, 25-75th percentile, and 10-90th percentile, respectively.
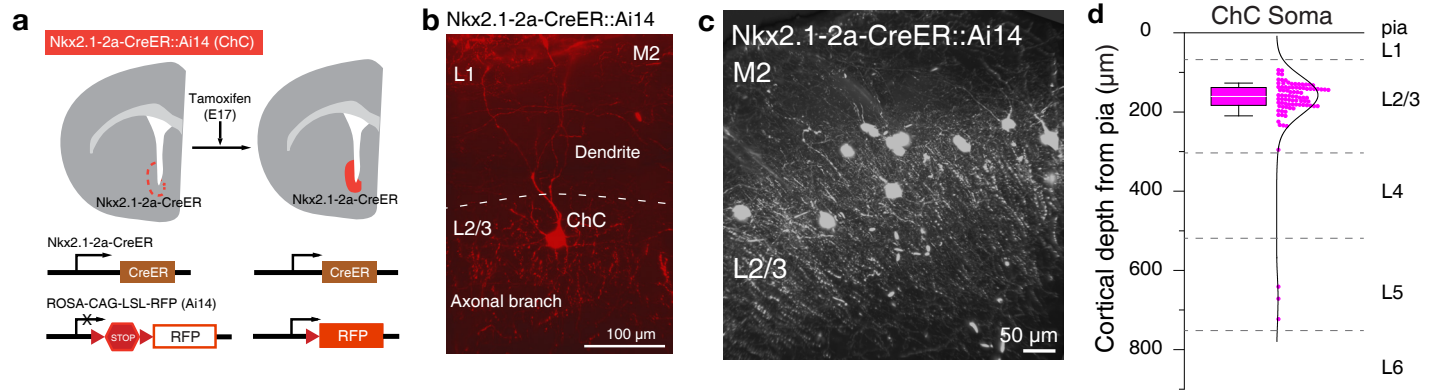
EDF 5 - Jung et al.

**Extended Data Figure 5. Genetical labeling of cortical chandelier cells in L2/3 premotor cortex.**

(a) Generation of Nkx2.1-2a-CreER::Ai14 mice. Nkx2.1-2a-CreER::Ai14 were generated by crossing Nkx2.1-2a-CreER with Ai14 mouse lines. A 2A-CreER cassette was inserted into the frame immediately after an open reading frame of an Nkx2.1 gene. To induce CreER activity in the offspring, tamoxifen was administered to timed pregnant SW females by oral gavage at E17.

(b) A representative image of ChC located in layer 2/3 premotor cortex.

(c) Light-sheet microscope image of ChCs' densely branching axonal cartridges in layer 2/3 of the premotor cortex.

(d) Cortical depth of ChC soma location from pia (n = 90 ChCs from 5 mice). Most of the ChCs marked by viral expression were located in the upper L2/3 (87 out of 90 ChCs) and 3 % of the ChCs were in the L5 (3 out of 90 ChCs). In the box plot, the white line, box size, and whisker indicate median, 25-75th percentile, and 10-90th percentile, respectively.
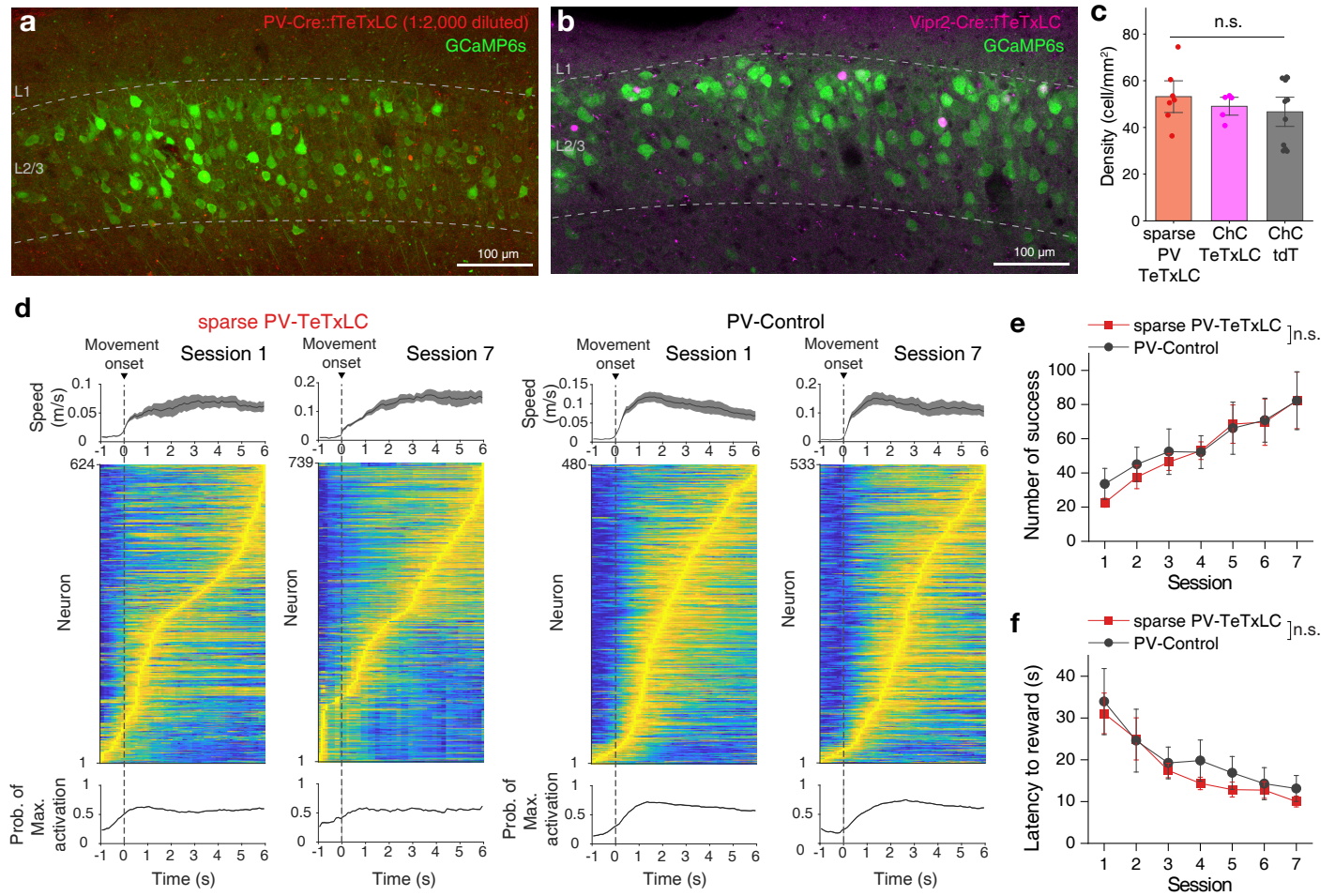
EDF 6 - Jung et al.

Extended Data Figure 6. Sparse expression of tetanus toxin light chain in parvalbumin interneurons.

(a) Example image of the expressions of AAV-hSyn-GCaMP6s (green) and AAV-hSyn-FLEX-TeTxLC-P2A-NLS-dTomato (red, 1:2,000 diluted) in L2/3 premotor cortex of PV-Cre mice (sparse PV-TeTxLC).

(b) Example image of the expressions of AAV-hSyn-GCaMP6s (green) and AAV-hSyn-FLEX-TeTxLC-P2A-NLS-dTomato (magenta) in L2/3 premotor cortex of Vipr2-Cre mice (ChC-TeTxLC).

(c) The density of cells expressing TeTxLC (n = 7 mice for sparse PV-TeTxLC; n= 5 mice for expression of ChC-TeTxLC; n= 11 mice for expression of FLEX-tdTomato in Vipr2-Cre mice, ChC-tdT; one-way ANOVA, $F_{group} = 0.62$, $P = 0.547$).

(d) Average movement speed (top), the normalized activity of premotor neurons (middle), and probability of maximum neuron activation (bottom) aligned to movement onset for sessions 1 and 7 in sparse PV-TeTxLC (n = 624 cells for session 1 and 739 cells for session 7, 7 mice) and PV control (n = 480 cells for session 1 and 533 cells for session 7, 7 mice).

(e) The average number of successes for sparse PV-TeTxLC and PV control mice increased with learning (n = 7 mice for sparse PV-TeTxLC; n = 7 mice for PV control; two-way repeated measures ANOVA, $F_{group} = 0.050$, $P = 0.831$).

(f) The average latency to reward for sparse PV-TeTxLC and PV control mice deceased with learning (n = 7 mice for sparse PV-TeTxLC; n = 7 mice for PV control; two-way repeated measures ANOVA, $F_{group} = 0.346$, $P = 0.578$). n.s., not significant; Error bars indicate s.e.m.

EDF 7 - Jung et al.
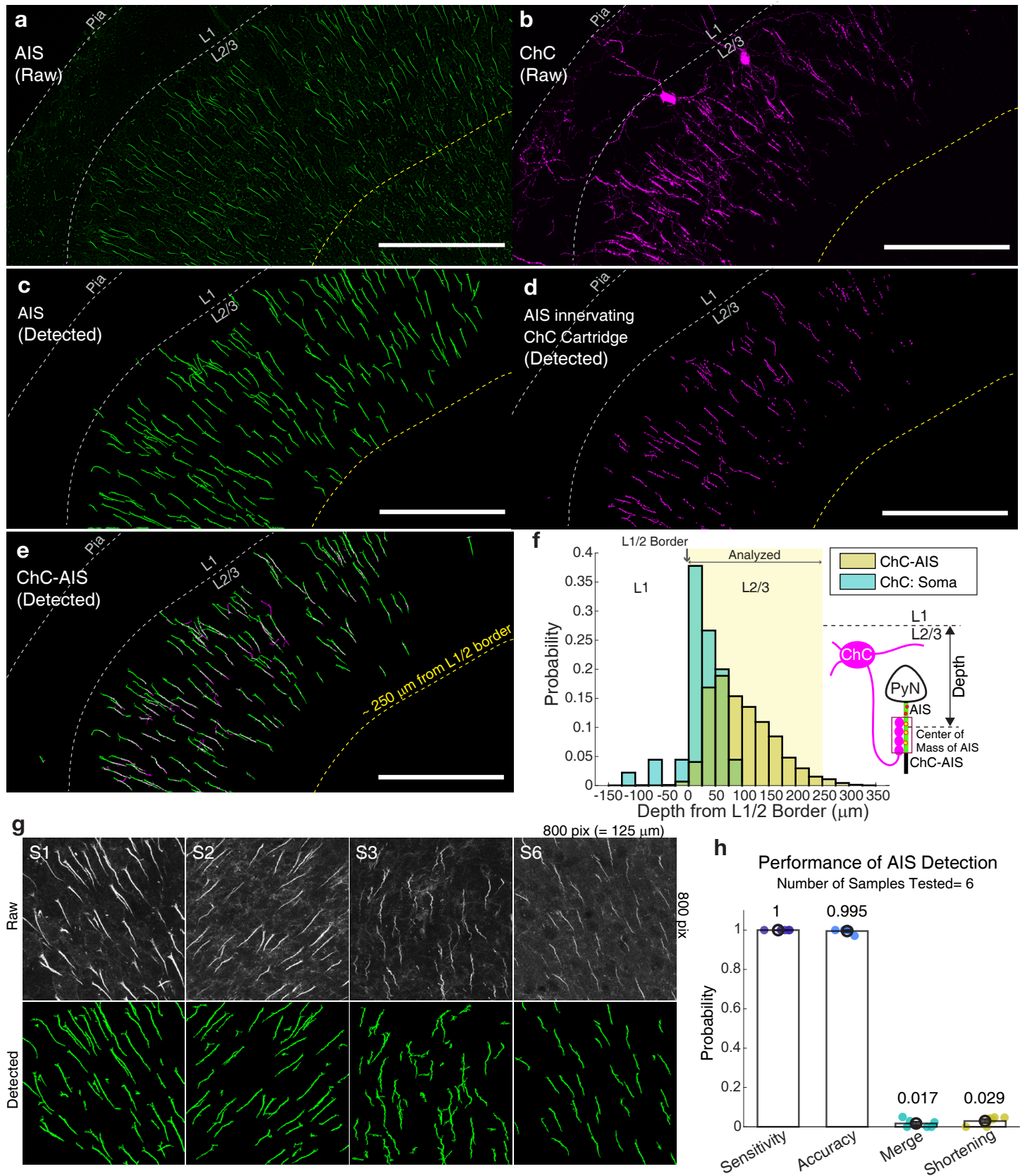
Extended Data Figure 7. Automated detection of ChC-AIS interaction.

(a-b) Example images of AIS and ChC in the M2 of Vipr2-Cre mice expressing AAV1-CAG-Flex-tdTomato.

(c-d) Detected AISs and ChC axonal bouton cartridges by an automated detection tool.

(e) Detected ChC-innervated AISs (ChC-AISs) with the white colocalized area. Representative images are maximum intensity projections of 400 X 206 X 10 µm^3. Scale bar = 100 µm for **a-e**. The dotted lines represent the layer 1/2 border (white) and the 250 µm-deep from the border approximately (yellow).

(f) Probability distribution of ChC somas and ChC-AISs by their position in layer 2. Only ChC-AISs within the range from 0 to 250 µm from layer 1/2 borders were considered as in layer 2 and analyzed further.

(g) Examples of samples (125 X 125 X 10 µm^3) used to evaluate the performance of automated AIS detection.

(h) Performance of AIS detection evaluated by comparing the results to manual detection in 4 different categories (Sensitivity = probability that a manually detected AIS is detected automatically. Accuracy = probability that an automatically detected AIS is detected manually. Merge = probability that an automatically detected AIS corresponds to multiple AISs in manual detection. Shortening = probability that an automatically detected AIS length is significantly shorter than manual detection).

EDF 8 - Jung et al.
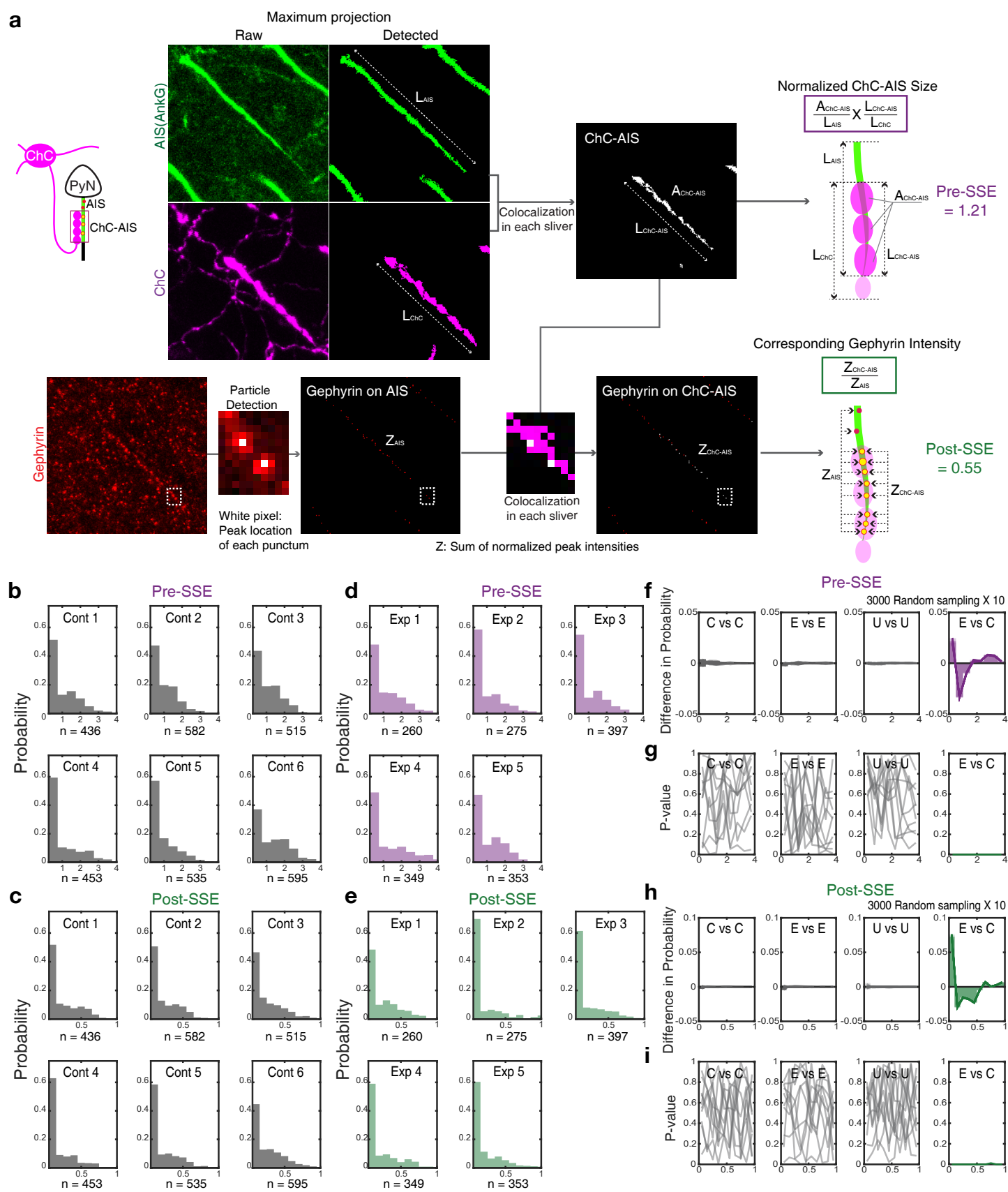
Extended Data Figure 8. Evaluation of ChC-AIS synaptic structural efficacy.

(a) Automatically detected segments were processed to extract quantities as the total AIS length (L_{AIS}), the ChC cartridge length (L_{ChC}), the AIS length covered by ChC (L_{ChC_AIS}), the area of ChC on AIS ($A_{ChC-AIS}$), and z-scored intensity of gephyrin puncta on AIS and ChC-AIS (Z_{AIS} and $Z_{ChC-AIS}$) for characterization. Those quantities were used to compute the defined characteristic value of presynaptic structural efficacy (Pre-SSE) and postsynaptic structural efficacy (Post-SSE). Every image represents the maximum intensity projection of the corresponding volumetric stack (Details in Methods).

(b-e) Normalized histograms of Pre-SSE and Post-SSE from mice in the control group (n = 6 mice) and the experimental (learning) group (n = 5 mice).

(f-i) We performed a random sampling test (n = 3000) repetitively (10 times) to compare individual distributions after training. The difference in probability was significant between the Experimental and Control condition (E vs C). We did not observe any significant difference between any other pair (random sampling from the Control and Control (C vs C), Exp. vs. Exp. (E vs E), or Unlabeled vs. Unlabeled (U vs U)). The distributions were compared by a two-tailed Wilcoxon-Mann-Whitney test.