Evolutionary expansion of the corticospinal system is linked to dexterity in *Peromyscus* mice

Kelsey M. Tyssowski[1,2^], Jeremy D. Cohen[3], Jian-Zhong Guo[3], Phoebe R. Richardson[1,2], Karen E. Cortina[1,2], Isobel H. Smith[1,2], Deshawn C. Eijogu[1,2], Adam W. Hantman[3*^], Hopi E. Hoekstra[1,2*]

1. Department of Organismic and Evolutionary Biology, Harvard University
2. Department of Molecular and Cellular Biology, Harvard University
3. Department of Cell Biology and Physiology, University of North Carolina School of Medicine

*Senior author

^correspondence: tyssowski@g.harvard.edu, adam_hantman@med.unc.edu

Abstract

Animals have evolved behavioral variation to adapt to distinct environmental features. The expansion of neuron number is a potential neural mechanism underlying this behavioral adaptation. Corticospinal neurons (CSNs) are a classic example: an expansion in the corticospinal system in the primate lineage has been hypothesized to underlie their exceptional dexterity. However, the role of CSN number in behavior has been difficult to assess due to the large evolutionary distance between primates and less dexterous taxa with fewer CSNs. Here, we use deer mice (*Peromyscus maniculatus*) to overcome this challenge. We compared two closely related subspecies of deer mice: forest mice, which evolved dexterous climbing to support a semi-arboreal lifestyle, and prairie mice, which are less dexterous. We find that forest mice have about two-fold larger corticospinal tracts (CSTs) driven by an increase in CSN number in secondary motor and sensory cortical areas (M2 and S2). Furthermore, in a reach-to-grasp test of dexterity, forest mice display higher success and greater grasping flexibility, using multiple grasp types. In contrast, prairie mice use a stereotyped scooping motion, consistent with the idea that an increase in CSN number supports more dexterous movement. High-throughput neural recordings during this task revealed a difference in the timing of neural activity between forest and prairie mice in M2, but not in primary motor cortex (M1): in forest mice, the peak of activity was shifted towards the time of grasp. Forest mice also outperform their prairie counterparts on an ecologically relevant climbing task, where they spend more time upright crossing a thin rod, move faster, and right themselves more quickly when they fall, suggesting a general increase in motor dexterity not restricted to hand use. To assess whether the increase in CSN number contributes to observed behavioral adaptations, we generated forest-prairie F2 hybrid animals with shuffled genomes, neural features, and behavior. We find that the F2 hybrids with larger CSTs perform better on the rod crossing task, suggesting that expansion of the CS system likely supports the adaptive increase in climbing dexterity in forest mice. Together, our work establishes the forest-prairie deer mouse system as a novel model to investigate the role of neuron number expansion, and CSNs in particular, in dexterous movement.

Introduction

Evolution has tuned neural function and circuits to support behaviors that enable animals to adapt to their ecological niches. For example, genetic mutations impacting sensory receptor function impact behavior by altering animals' perception of the world (Baldwin and Ko, 2020), a mutation in an ion channel affects *Drosophila* mating calls (Ding et al., 2016), and evolved differential expression neurotransmitter and neuropeptide genes between subspecies impacts social and motor behaviors

(Bendesky et al., 2017; Tamvacakis et al., 2018). In contrast to these examples of molecular adaptations (for review, see Niepoth and Bendesky, 2020), we know less about how evolution alters circuitry to drive behavioral change.

Evolutionary circuit changes fall into two non-mutually exclusive broad categories. The first is changes in circuit connectivity (Isko et al., 2024; Roberts et al., 2022; Ye et al., 2024), such as in the pheromone-sensing circuits that govern mating behaviors in *D. melanogaster* versus *simulans* (Seeholzer et al., 2018), which can impact circuit activity (Baier et al., 2025). Additionally, neural circuits have also evolved changes in neuron number. Evolutionary changes in neuron number that have been explicitly linked to behavior have mostly been identified in sensory neurons (though see Ye et al., 2024), such as the olfactory circuits that alter feeding preference in *Drosophila* (Auer et al., 2020; Prieto-Godino et al., 2017; Takagi et al., 2024). Outside of sensory systems, there are reports of neuron number or neural projection expansion associated with behavioral adaptation (Chen et al., 2025; Isko et al., 2024), suggesting the importance of nervous system expansion as an evolutionary mechanism; however, most of these have not been directly linked to behavioral change. Over evolutionary time scales of tens of millions of years, neural expansion manifests as expansion of existing brain regions and generation of new ones (Striedter and Northcutt, 2019). A prime example of this brain region expansion is the expansion of cortex in the primate lineage, which is thought to confer increased dexterity and cognition (Geschwind and Rakic, 2013; Herculano-Houzel, 2011). While the correlation of cortex size with cognition and dexterity is well established, how the evolutionary expansion of specific cortical cell populations mediates this behavioral adaptation remains largely unknown.

One such cortical cell population that has undergone evolutionary expansion is corticospinal neurons (CSNs), which project from cortex to spinal cord. The expansion of the corticospinal system, both through changes in the number of CSNs and through changes in their synaptic partners in the spinal cord, has been hypothesized to underlie the evolution of dexterity (Heffner and Masterton, 1975, 1983; Iwaniuk et al., 1999; Lassek and Rasmussen, 1940; Lemon, 2008; Levi and Schwab, 2021; Nudo et al., 1995; Nyberg-Hansen and Rinvik, 1963; Strick et al., 2021; Welniarz et al., 2017). Consistent with this idea, experiments across taxa demonstrate that manipulating CSNs or the corticospinal tract (i.e., axons of CSNs; CST) impairs both fine-motor control of the hand or forepaw (Alstermark et al., 1989; Gu et al., 2017; Lawrence and Kuypers, 1968; Serradj et al., 2023; Wang et al., 2017; Whishaw et al., 1998), as well as complex locomotion, such as ladder-walking (Beier, 2019; Metz and Whishaw, 2002). Thus, CSNs have a clear role in dexterity, which can be defined as “finding a motor solution for any external situation, that is, to adequately solve any emerging motor problem correctly, quickly, rationally, resourcefully” (Bernstein, 1996). However, the role of the number of CSNs in mediating dexterity has been difficult to test directly. This difficulty is in part because animals with reported differences in CSN number or CST size (e.g., rodents and primates) diverged about 80 million years ago, over which time many other neural and morphological changes have accumulated, making it difficult to assess whether and how the expansion of CSNs specifically impacts dexterity.

To overcome this challenge, we take advantage of two closely related subspecies of deer mice (*Peromyscus maniculatus*) that have evolved differences in dexterous movement. Following the last glacial retreat (10k years ago), North American deer mice living in open or prairie habitats migrated into

forests where they evolved dexterous climbing behavior, presumably to support a semi-arboreal lifestyle, whereas populations living in prairies did not evolve this behavioral adaptation (Blair, 1950; Dice, 1944; Foster, 1959; Hibbard, 1968; King, JA, 1968; Osgood, Wilfred H., 1909). This evolution occurred multiple times across North America, resulting in at least four convergently evolved forest-mouse ecotypes (Figure 1A) (Kingsley et al., 2017). These subspecies are not only closely related but also interfertile, enabling the generation of hybrids to aid in identifying the genetic basis of evolved variation between subspecies (Hager et al., 2022; Kingsley et al., 2024). Compared to their prairie counterparts, forest mice have evolved shared morphological phenotypes, including longer tails and larger hind feet, that could impact their behavior (Dice, 1944; Hager et al., 2022; Hager and Hoekstra, 2021; Kingsley et al., 2017; Osgood, Wilfred H., 1909). However, these morphological changes are minimal compared to, for example, the difference between a rodent and primate hand, making it easier to account for them experimentally. In addition, behavioral studies of wild-caught and lab-raised forest and prairie mice reveal a dramatic behavioral difference: forest mice outperform prairie mice on tests of both vertical and horizontal climbing (Foster, 1959; Horner, B. Elizabeth, 1954; Kingsley et al., 2024; Thompson, 1990). Given the known role of the corticospinal system in dexterous behavior, together with the dramatic difference in dexterity between these subspecies, we hypothesized that the evolutionary expansion of the corticospinal system might mediate this behavioral adaptation.

Results

Forest mice have larger corticospinal tracts compared to prairie mice

We focus on one forest and one prairie mouse ecotype, which are recently diverged (by less than 1 million years; Bradley et al., 2019) and exist as laboratory colonies: *P. m. nubiterrae* (forest) and *P. m. bairdii* (prairie) (Osgood, Wilfred H., 1909) (Figure 1B-C). To determine whether CST size differs between ecotypes, we took advantage of CST-specific expression of protein kinase C gamma (PKCγ) (Miki, 1996; Mori et al., 1990) to label the CST in cleared, whole-mount spinal cords. Forest mouse spinal cords showed a greater PKCγ-positive CST volume compared to their prairie counterparts (Figures 2A-E). This difference was particularly striking in the cervical spinal cord, which controls the forelimbs, where forest mice showed approximately 2-fold greater CST volume than prairie mice (Figures 2A,C). However, even in the thoracic and lumbar spinal cord, which control the trunk and hindlimbs, forest mice displayed an approximately 1.5-fold larger CST volume (Figure 2D). While we found a 1.4-fold difference between subspecies in total spinal cord volume (Figure S1A), this difference was entirely due to differences in lumbar spinal cord volume, with forest mice having 1.8-fold larger lumbar spinal cords, but no difference in cervical spinal cord volume (Figure S1D-G). Nonetheless, we observe significant differences in CST size even when normalizing for total spinal cord volume (Figure S1B). The particularly large CST expansion in the cervical spinal cord of forest mice suggests that as axons travel down the spinal cord, an increased proportion of axons branch off and innervate the grey matter of the cervical spinal cord in forest mice compared to prairie mice. Consistent with this conclusion, we observed that a larger fraction of the total PKCγ volume was in cervical spinal cord in forest mice compared to prairie mice (Figure 2E). We observed no differences in CST length (Figure S1C), suggesting that the CST targets the same spinal segments in both subspecies. Together, these data support the conclusion that the CST is larger in forest mice compared to prairie mice, consistent with our hypothesis that CST expansion mediates an increase in dexterity in forest mice.

The difference we observe in CST volume could be due to an increase in the number of CST axons, or to differences in the morphology of individual axons, such as increased axon diameter or myelination. To distinguish between these possibilities, we counted axons within the PKCγ-positive CST region of the dorsal funiculus of the cervical spinal cord in coronal sections (Figure 2B). We identified axons using FluoroMyelin, a fluorescent myelin stain, since most CST axons are myelinated. We found that the number of axons in the CSTs of forest mice was about 2-fold greater than that of prairie mice, consistent with the increase in CST volume being due entirely to an increase in axon number (Figure 2H). To assess the ultrastructure of CST axons between subspecies, we performed electron microscopy on the ventral region of the dorsal funiculus, where the CST resides (Figure 2G). We found no difference between subspecies in axon density (Figure 2H), axon diameter (Figure 2I), or myelination, as measured by the g-ratio (i.e., the ratio of the total axon diameter to the axon diameter inside the myelin; Figure 2J). These data support the idea that the larger CST size in forest mice is entirely due to an increase in the number of axons, rather than differences in axon morphology. Although this does not entirely rule out the possibility that more axons branch before exiting the CST in forest mice, the most parsimonious explanation for this increase in axon number is a corresponding increase in CSN number.

Forest mice have more corticospinal neurons in secondary motor and sensory cortices

In rodents, CSNs reside in several cortical areas: primary motor cortex (M1, also called caudal forelimb area), secondary motor cortex (M2, also called rostral forelimb area), primary sensory cortex (S1), and secondary sensory cortex (S2) (Liang et al., 2011). The expansion of CSNs in forest mice could be uniform across these cortical areas or confined to a subset of areas. To determine whether the increase in CST size results from an expansion of a specific spatially defined subpopulation of CSNs, we labeled CSNs by injecting a retrogradely transported AAV serotype (Tervo et al., 2016) expressing nuclear-tagged GFP into the cervical spinal cord grey matter. We injected one spinal segment between C5-C8 to target the region of the spinal cord controlling the forelimbs (Figures 3A, S2A), where we expected, from our CST volume measurements, to observe a difference in CSN grey matter innervation between subspecies. We then quantified nuclei throughout the brain in light-sheet images of cleared brains and assigned them to cortical regions (or the rest of the brain) using a custom pipeline. To control for differences in viral labeling, we normalized our counts to the total number of labeled nuclei per brain, which did not differ between subspecies, nor did total brain volume (Figure S2B-C). We observed more CSNs in M2 and S2 in forest compared to prairie mice, but a similar number in M1 and S1 (Figure 3B-D). These data indicate that the increase in CST size is due to an increase in CSN number, specifically in secondary, but not primary, motor and sensory cortex.

CSNs are a subtype of neurons in layer V of cortex. We therefore considered the possibility that the expansion we observe in CSN number is due to a broader expansion of layer V or of sensorimotor cortex as a whole. We performed single-nucleus RNA sequencing (snRNA-seq) of frontal cortex, including all regions where CSNs reside, in five forest and five prairie mice, assessing a total of 107,000 cortical nuclei. We aligned reads to the *P. maniculatus* genome and integrated data from individual nuclei between datasets to compare between subspecies (see Methods). We clustered nuclei by gene expression and aligned our data to the BICNN atlas (Callaway et al., 2021) to identify

cell types. We observed no difference in the total number of neurons (or other cell types) between subspecies, indicating that forest mice do not have a general increase in neuron number (Figure S3A-D). When we looked for differences in neuron abundance at the level of neuron subtypes, we observed only a small but significant increase in the number of L6b neurons in prairie mice (Figure S3E-G). Notably, we were able to clearly distinguish layer V extratelencephalic (L5ET) neurons (a subclass that includes CSNs, as well as corticopontine, corticotectal, corticorubral neurons, and others). Still, we observed no difference in L5ET neuron abundance between subspecies. We note that the failure to observe a difference in abundance in L5ET neurons is not surprising given that CSNs are expected to make up only a small fraction of L5ET neurons, and suggests that the increase in CSN number is specific to CSNs rather than a side effect of increasing the size of layer V as a whole. Consistent with previous work suggesting that virally labeling CSNs is necessary to identify them in snRNA-seq data (Golan et al., 2023; Winter et al., 2023), we were unable to identify a clear CSN cluster among L5ET neurons (Figure S3H-K). However, we asked whether we could identify putative markers of the expanded CSN population through differential gene expression. Pseudo-bulk differential gene expression analysis between forest and prairie mice L5ET neurons revealed a significant enrichment in known CSN marker genes in the genes upregulated in forest compared to prairie mice (Figure S1L, but not vice versa: $p = 0.0068$ in forest-upregulated genes, $p = 0.558$ in prairie up-regulated genes, Fisher exact test) and revealed several known CSN markers that could mark the expanded CSN population. Therefore, CSN expansion in forest mice is not likley a side effect of a general increase in neuron number or in L5ET neurons but is specific to a subset of CSNs that may be distinguishable by specific CSN marker genes.

Forest mice display more dexterous hand use

Given the role of cortex, and CSNs more specifically, in dexterity, we next asked whether forest and prairie mice differ in dexterous movement. Because many studies across taxa have demonstrated the role of the CST in a reach-to-grasp task (e.g., Alstermark et al., 1989; Lawrence and Kuypers, 1968; Wang et al., 2017; Whishaw et al., 1998), we assessed forest and prairie mice in this behavioral paradigm, hypothesizing that forest mice would be more successful at this task given their expanded CSTs. We placed animals in a clear box with a slit and trained them to reach for a food pellet and measured their success over six days. We found that forest mice were more successful at this task both at the beginning and after six days of training compared to their prairie counterparts (Figure 4A), indicating that forest mice are more dexterous than prairie mice, consistent with their greater number of CSNs.

We next sought to understand the behavioral mechanisms underlying this difference in success. We reasoned that the freely moving version of the task requires high levels of dexterity: the animals must make many decisions about how to position their bodies and hands to grasp the pellet. We therefore asked whether we could train prairie mice to perform successfully in a more constrained setup with a consistent body and pellet position, reducing the animal's degrees of freedom and, thus, overall task complexity. Indeed, by reducing the complexity of the task, we were able to train prairie mice to achieve a success rate equivalent to that of forest mice (Figure 4B, Figure S4A). To more directly assess the relative flexibility of forest and prairie mice, we varied the position of the pellet. When we presented mice with pellets progressively further from their noses, we found that all prairie mice

eventually reached a point at which they were unable to perform the task. On the other hand, while the success rate of forest mice was lower at further pellet positions, they never reached a position at which they failed completely among the positions we tested (Figure 4B). Importantly, this difference in performance cannot be explained by differences in morphology. While the median arm plus hand of forest mice is about 3mm longer than that of prairie mice (Hager, Emily R., 2020; consistent with measurements in our sample, Figure S4B), forest mice can reach a pellet that is 15mm from their noses, whereas prairie mice can only reach a pellet 9mm from their noses. This large difference in distance at which mice can succeed indicates that differences in arm length cannot entirely account for differences in the ability to grasp more distant pellets, and that forest mice have greater flexibility in the pellet positions they can successfully reach than their prairie counterparts.

Previous reports in both rats and monkeys demonstrate that lesioning the CST alters the dexterous grasping part of the reach-to-grasp behavior. Specifically, animals transition from using dexterous grasping to using a scooping movement (Lawrence and Kuypers, 1968; Whishaw et al., 1998). We therefore asked whether forest animals with larger CSTs are more likely to use a dexterous, top-down grasping movement, whereas prairie mice use a less dexterous scooping movement. Indeed, we found that forest mice approach the pellet with their hands parallel to the platform and grasp it by surrounding it with all fingers from the top. In contrast, prairie mice approach the pellet with their hands at a more perpendicular angle to the platform to execute a scooping grasp (Figure 4C-D). We note that the grasping angle is not dependent on pellet position (Figure S4D-E). We next asked whether forest and prairie mice use qualitatively different grasping strategies. We divided grasps into four categories: top, describing an approach from the top that uses all fingers to grasp the pellet; side, describing a scooping approach from the side; pincer, which describes a grasp in which the pellet rests on the thumb-like digit one with another finger; and finger, which describes when the pellet is in between two fingers (Figure S4C). When we compared grasps at each mouse's most successful position, we found that forest mice use a wider variety of grasp types compared to prairie mice, which, with a small number of exceptions, almost exclusively use the side grasp (Figure 4E), suggesting that forest mice are more dexterous in the grasping part of the behavior, consistent with our hypothesis that corticospinal expansion supports greater dexterity.

We hypothesized that the more varied grasping strategy of forest mice may enable them to successfully grasp the pellet even if they inconsistently target their hands to the platform. In contrast, we hypothesized that because prairie mice can grasp the pellet in only one way, they need to be able to direct their hands to the optimal position to perform that grasp. To address this, we identified the ideal success position for each pellet position for each animal, that is, the median hand position of successful grasps. Then, to assess the variability in successful grasp positions, we calculated the median absolute deviation, i.e., the distance between the hand position of each successful grasp and this ideal hand position. We found that forest mice could successfully grasp the pellet at greater distances from their ideal hand position, indicating greater variability in successful grasp positions and suggesting that their increased dexterity in grasp strategy may enable greater dexterity in pellet targeting (Figure 4F-G). Consistent with this finding, we also observed that forest mice had greater variability not only in their targeting but also in the overall trajectory to the pellet (Figure S4I-J). We note that forest and prairie mice significantly differ in their pellet targeting and reach trajectory even when

accounting for pellet position (Figure S4G,J). We considered whether the difference in targeting might be due to differences in hand morphology, as forest mice have slightly longer fingers than prairie mice, which could allow them to grasp the pellet from a larger surrounding radius. We therefore measured how far each reach was from the pellet itself, but found this was not different between subspecies (Figure S4F) and a linear model accounting for this reach-pellet distance showed a significant impact of subspecies ($p = 0.001$), supporting our conclusion that forest mice use a wider range of grasp strategies that is not merely a result of their differing morphology.

To further investigate this idea, we assessed how forest and prairie mice fail. We reasoned that if forest mice indeed have a better ability to grasp the pellet than prairie mice, a larger fraction of prairie mice failures would be failures in grasping, whereas forest mouse failures would be failures in targeting. To address this, for each failed reach, we assessed the nearest point in the failed reach trajectory to the grasp position of a successful reach-to-grasp (fail-success distance). A reach that never comes close to a successful reach position is a failure in targeting, while a reach in which the mouse's hand arrives at a location where the animal has made a successful grasp is more likely to be an error in grasping itself. A comparison of the distribution of fail-success distances between forest and prairie mice revealed that while forest mice are more likely to fail by targeting poorly, prairie mice often fail despite targeting their hands to a position where they were previously successful (Figure 4F,H). In support of the idea that a difference in grasping dexterity might impact the ability of prairie mice to successfully grasp pellets far from their noses, we find that at far pellet distances, prairie mouse failures are more often targeting failures (Figure S4H). This difference between subspecies suggests that, at greater distances, where targeting becomes harder, grasping dexterity is more critical for success. Together, these data again support the idea that forest mice are more dexterous in their reach-to-grasp behavior. While prairie mice can learn to be successful at the task in a confined set of circumstances in which they can easily, consistently, and precisely target the ideal grasp location, forest mice, presumably due to their increased manual dexterity, do not need to achieve precise targeting to achieve similar levels of success.

Neural dynamics in forest mouse M2 differ from those of prairie mice

Given the role of cortex in dexterous behavior and the differences in cortical neuroanatomy we observe, we asked whether differences in CSN number between forest and prairie mice might impact neural dynamics. Using Neuropixel probes (Steinmetz et al., 2021) to achieve high-throughput electrophysiological recordings, we recorded from M1 (a.k.a. CFA, where we observed no CSN number differences between subspecies) and M2 (a.k.a. RFA, where we observed differences between subspecies) in trained, head-fixed animals during reaching behavior (Figure S5). We recorded from the same animals over 2-5 sessions and validated probe placement by comparing it to our retrogradely labeled brains. We observed robust neural signals aligned with the behavior, similar to that observed in lab mice in the same conditions (Sauerbrei et al., 2020). Because the main behavioral difference occurred during grasping in successful trials, we aligned neural signals during successful reaches to the time of the grasp. Interestingly, we observe a difference in the timing of the signal between forest and prairie mice in M2, but not M1 (Figure 5A-C). Whereas in prairie mice, the neural signal in M2 peaks before the time of the grasp and before the signal in M1, as expected from recordings in lab mice (Saiki-Ishikawa et al., 2025), in forest mice, the M2 signal is later, more aligned with the timing of

the grasp and more similar to the M1 signal. Notably, this is only true for positively modulated neurons (i.e., those that increase with behavior), but not for negatively modulated neurons, which show similar alignment in activity between forest and prairie animals. These findings indicate that neural dynamics in cortex differ between forest and prairie mice and are consistent with our behavioral results. While the most essential part of a successful reach for prairie mice appears to be the targeting phase before the grasp, for forest mice, it is the grasping itself. Furthermore, the fact that we observe a difference specifically in M2 suggests that the difference in CSN number, which we observe in M2 but not M1, may play a role in this difference in dynamics.

Forest mice display more dexterous climbing behavior

In addition to its role in hand use, the CST has also been implicated in complex locomotion tasks, such as ladder walking (Beier, 2019; Metz and Whishaw, 2002). Forest mice are thought to have evolved dexterous movement in response to the selective pressure to interact with a complex, three-dimensional environment, including climbing. Indeed, a rich natural history literature reports better performance in climbing tasks in *P. maniculatus* subspecies that evolved in forests compared to their closest non-forest-dwelling relatives (Foster, 1959; Horner, B. Elizabeth, 1954; King, JA, 1968; Kingsley et al., 2024; Thompson, 1990). We therefore tested animals on a rod-crossing task, where they crossed a thin rod between two platforms, similar to movement across a thin twig, which has been documented in wild deer mice (Graves et al., 1988), and similar to previous studies (Horner, B. Elizabeth, 1954; Kingsley et al., 2024). We made several adjustments to earlier versions of this task to enable us to measure animals' climbing dexterity, again, defined using Bernstein's definition: "to adequately solve any emerging motor problem correctly, quickly, rationally, resourcefully" (Bernstein, 1996). We altered our assay to: (1) train animals to cross to reduce the impact of the previously observed motivational difference between subspecies (Hager, Emily R., 2020; Kingsley et al., 2024); and (2) expose mice to progressively smaller and smaller rods to better identify the limits of the motor problems each subspecies could solve. We found that at the widest rod diameter, i.e., the easiest rod, both forest and prairie mice successfully cross the rod, although forest mice cross at a slightly higher velocity (Figure 6C). On narrower rods, we find that forest and prairie mouse performances diverge substantially (Figure 6A-F). Forest mice cross with higher velocities and spend most of their time moving and in an upright position, whereas prairie mice often stand still on the rod or hang upside down. We also find that when animals flip upside down on any rod size, forest mice right themselves more quickly. Finally, we observe that prairie mice more frequently fail to cross at all, with all forest mice succeeding in crossing the smallest rod without falling, but only 6 out of 11 prairie mice. Together, these data indicate that forest mice are better able to climb in a broader range of situations and are thus more dexterous in both climbing and reaching.

Climbing dexterity in forest mice is linked to corticospinal tract size

The evolutionary expansion of the corticospinal system, especially in the primate lineage, has been hypothesized to support dexterous movement, but testing this hypothesis has been difficult (Lemon, 2008; Levi and Schwab, 2021). While CST or CSN manipulations support the role of the corticospinal system in dexterity, how CST *size* or CSN *number* relates to dexterity remains unknown. To address this question, we took advantage of the interfertility of our forest and prairie deer mouse subspecies. We crossed forest and prairie deer mice and then crossed their progeny (F1s) to generate F2 hybrid

mice. F2 mice have recombinant, shuffled genomes, with both forest and prairie elements (Figure 7A). This shuffling means that each individual F2 hybrid animal inherits different combinations of forest and prairie genes and, therefore, different neural phenotypes and behaviors. We hypothesized that if CST size is important for dexterous motor behavior, we should observe larger CSTs in more dexterous F2 animals. We assessed 51 F2 mice on the rod crossing test at three rod sizes. As expected, F2 animals showed a wide distribution of rod crossing performance levels, as assessed by their velocity, time upright, time spent moving, and time to right themselves when they fall (Figure 7B-E). While we observe a correlation between F2 animals' performances on the wider two rods (Figure S6A-B), this is not correlated with their performance on the narrowest rod (Figure S6C). This discrepancy suggests that crossing easier rods successfully requires different genetic, neural, and behavioral mechanisms than crossing the more challenging, narrower rod. Therefore, we focused the rest of our analysis on performance on the narrowest rod, as we judged this to be most relevant to their dexterity and thus most likely to depend on CST size.

To determine whether CST size is linked to rod-crossing dexterity, we selected four "good" and four "bad" climbers based on their performance metrics. Interestingly, velocity and the fraction of time spent upright during crossing are not correlated (Figure S6D), indicating that these aspects of climbing dexterity are controlled by different neural and genetic mechanisms. Therefore, we identified "good climbers" as those in the top half of animals in both their velocity and time upright and "bad climbers" as those in the bottom half of animals in both measures. When we compared cervical CST volume between good and bad climbers, we found that good climbers have larger CSTs than bad climbers (Figure 7F). When we asked whether CST size correlates with rod crossing velocity in these eight mice, we also found a correlation (Figure 7G), but we did not find a correlation between CST size and fraction of time spent upright (Figure S6E). Importantly, we note that tail length, weight, and hind foot size, traits commonly altered in arboreal animals and hypothesized to be involved in climbing performance (Horner, B. Elizabeth, 1954; Kingsley et al., 2024, 2017; Young, 2023), are not correlated with velocity (Figure S6F-G). Furthermore, controlling for weight, tail length, and foot size, we still find a positive correlation between CST size and rod crossing velocity ($r^2 = 0.81$, $p = 0.014$, Pearson's correlation). While we cannot completely rule out the possibility that another neural or morphological feature that confers climbing dexterity is genetically linked to CST size, the most parsimonious explanation for these data is that a larger CST, and likely more CSNs, supports climbing dexterity.

Discussion

Here, we provide evidence that a forest-dwelling deer mouse subspecies evolved an expanded number of CSNs compared to a closely related prairie-dwelling mouse. We further find that forest mice are more dexterous in both a reach-to-grasp task and an ecologically relevant rod-crossing test of climbing. The difference between forest and prairie mouse behavior is specifically in dexterity, that is, the ability "to adequately solve any emerging motor problem correctly, quickly, rationally, and resourcefully", rather than in general motor coordination, that is, the ability to precisely control movements (Bernstein, 1996). In a reach-to-grasp task, forest mice are flexible in the conditions in which they can succeed whereas, prairie mice can learn to succeed only under specific circumstances and by using a stereotyped movement. Similarly, in a test of climbing performance, forest mice can dexterously climb narrow rods, whereas prairie mice can only successfully climb wider rods where

quick responses to momentary loss of balance are less important. This finding is consistent with the role of the corticospinal tract in dexterous movements in other taxa: monkeys, cats, rats, mice, and humans with damage to the corticospinal system have reduced hand dexterity (Alstermark et al., 1989; Gu et al., 2017; Lawrence and Kuypers, 1968; Serradj et al., 2023; Wang et al., 2017; Whishaw et al., 1998; Zhu et al., 2010), and CSN lesions also impact complex locomotion, such as walking on a ladder (Beier, 2019; Metz and Whishaw, 2002). These two behaviors may be related; that is, reduced hand dexterity may affect complex locomotor tasks such as climbing. On the other hand, the neural mechanisms that enhance hand dexterity could also improve animals' ability to solve motor problems beyond the hand. While we have not distinguished between these possibilities, our data suggest that expansion of the CS system likely confers dexterity in an ecologically relevant rod-crossing climbing task, supporting the idea that evolutionary expansion of the corticospinal system supports dexterous movement.

Variation in CSN number between forest and prairie mice likely influences behavior through altering neural circuit dynamics. Consistent with this, we observe that forest animals with more CSNs also show a shift in the timing of cortical activity during a reach-to-grasp task, with the peak of activity occurring later, closer to the timing of the grasp, compared to their prairie counterparts. While it is not clear to what extent this shift is due to CSN expansion, several lines of evidence suggest it may be influenced by CSN number. First, we observe the shift in activity only in M2, where we observe more CSNs, but not in M1, where we do not. In addition, previous work suggests that pyramidal tract neurons (a neural subtype that includes CSNs) that project to brainstem and potentially spinal targets exhibit less preparatory activity than other layer V extra-telencephalic neurons (Economo et al., 2018). This result raises the possibility that what we observe is a result of an increase in the proportion of these neurons in M2 driving the network to be active later. We also observe that while prairie mice are better at targeting the pellet's location during the task, forest mice are better at grasping the pellet and can flexibly use their hands to do so. This difference in targeting is consistent with the timing of their respective activity in M2, raising the possibility that the increase in CSN number enables better control of grasping, and the network activity shifts to support this behavior. Indeed, while M2 activity is often observed earlier than activity in M1, its timing relative to M1 can shift depending on behavioral context (Saiki-Ishikawa et al., 2025). Furthermore, a role for M2 CSNs in mediating specifically the dexterous grasping phase of the behavior is consistent with previous work (Serradj et al., 2023; Wang et al., 2018). However, future work that specifically records from CSNs or manipulates them in forest and prairie mice will be necessary to fully understand how CSN expansion impacts neural dynamics.

Across taxa—from birds to lizards to rodents to primates—arboreality has evolved repeatedly. Arboreal animals convergently evolved morphological characteristics such as long and sometimes prehensile tails, larger hands and feet, or longer whiskers, as well as behavioral factors such as differences in gait or hand use, often related to climbing dexterity (Grant and Goss, 2022; Hager et al., 2022; Kingsley et al., 2017; Young, 2023). Indeed, dexterity may be an arboreal trait, though other environmental pressures likely also influence hand use (Grillner and Wallén, 1985; Iwaniuk and Whishaw, 2000; Sustaita et al., 2013). However, less is known about neural adaptations to arboreality. We show that CSN number expansion may be a neural adaptation to life in a forested habitat. Consistent with this idea, primates, which also evolved arboreality, have evolved expanded corticospinal systems (Lemon,

2008; Levi and Schwab, 2021; Strick et al., 2021), although the length and depth of projection of CS fibers (but not CSN number) has been suggested not to be associated with evolution of arboreality (Iwaniuk et al., 1999). The repeated evolution of CS expansion across primates and forest deer mice makes deer mice a promising model system for understanding neural principles relevant to primate evolution and human disease.

Given that we observe a difference in CSN number in forest and prairie mice raised in the same laboratory conditions, this difference must be driven by genetic change rather than environmental factors. The genetic differences driving this difference in cell abundance are likely active during embryonic and early postnatal development, when CSN number is established (Greig et al., 2013). CSNs are born in embryonic development during the generation of layer V. After their birth, a defined gene expression program specifies their projection to the spinal cord, and also to specific spinal cord targets (Cederquist et al., 2013; Gu et al., 2019; Itoh et al., 2023; Sahni et al., 2021b, 2021a; Song et al., 2023). Finally, postnatally, most CSN axons are pruned from the spinal cord, and these neurons cease to be CSNs (Kamiyama et al., 2015, 2006; Low et al., 2008; Weimann et al., 1999). The difference we observe in CSN number is likely to occur at either the second or third step, as we do not observe a broad increase in the number of layer V or extra-telencephalic neurons. An intriguing possibility is that the expansion of CSNs we observe is due to the expansion of a specific subpopulation of CSNs with a particular developmental origin (Huilgol et al., 2023), synaptic partner in the spinal cord (Carmona et al., 2024; Fageiry et al., 2024; Ueno et al., 2018), or general projection target in the brain or spinal cord (Itoh et al., 2023; Nelson et al., 2021; Sahni et al., 2021b, 2021a; Song et al., 2023). Future work to understand the molecular and cellular basis of this evolved CSN difference will open the door to both evolutionary mechanisms research and a molecular handle for future investigation of the neural consequences of CSN expansion. The F2 hybrid deer mice provide a powerful system to investigate this question, given the ability to use quantitative trait loci mapping to reveal the genetic basis of morphological and behavioral variation (Bedford and Hoekstra, 2015).

In artificial neural networks, increasing network size, which is analogous to increasing neuron number, increases network performance via emergent properties, that is, new properties that cannot be predicted merely by understanding the properties of the smaller networks (Berti et al., 2025; Elman, 1990; Hopfield, 1982; Jozefowicz et al., 2016; Takiyama and Okada, 2012). Indeed, complex nervous system functions, such as cognition, including the motor cognition required for dexterity, can be understood as emergent properties of the brain (Chialvo, 2010; Shenoy et al., 2013; Thiebaut de Schotten and Forkel, 2022). While it is therefore likely that increasing the CSN number improves dexterity by altering emergent properties of the circuit, it has historically been challenging to identify those properties, given the need to compare relatively similar systems in which only neuron number has changed. Indeed, despite the ubiquity of variation in neuron number and its association with the evolution of cognition and dexterity, it is still largely unclear how neuron number alters circuit function to impact behavior (Arendt et al., 2019; Kverková et al., 2022; Lemon, 2008; Levi and Schwab, 2021; Williams and Herrup, 1988). Here, we identify deer mice as a system in which it is possible to address a fundamental question in nervous system function: how does changing neuron number alter neural dynamics to increase behavioral complexity?

Figures

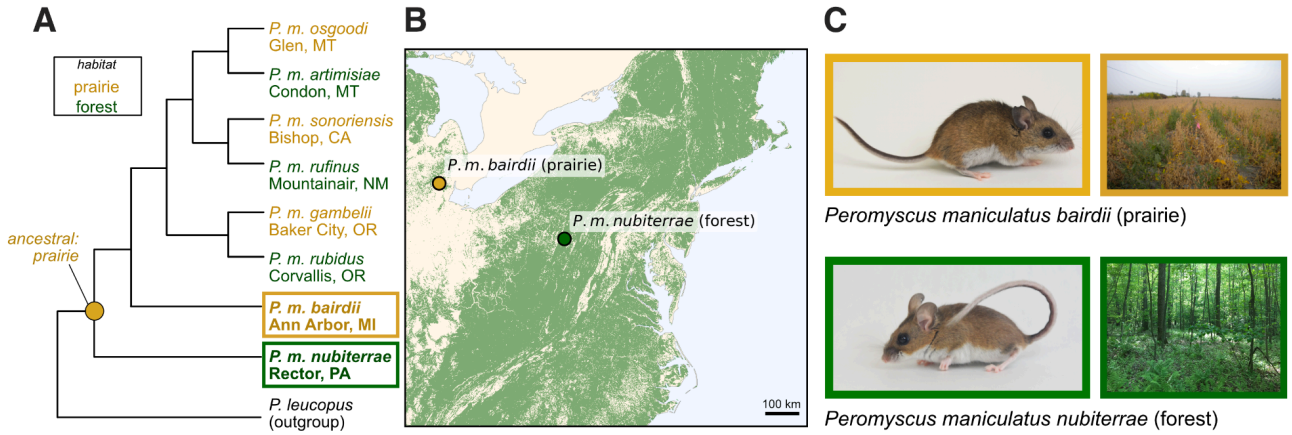


Figure 1: Forest and prairie deer mouse ecotypes

A. Cladogram showing repeated evolution of forest and prairie deer mouse ecotypes. The subspecies used in this study are outlined.

B. Terrain map showing the trapping locations of the mouse colonies used in this study. Green shading represents forests, specifically regions with greater than 30% tree canopy cover according to the United States Geological Survey National Land Cover Database.

C. Photos of representative prairie and forest mice from the laboratory colony, as well as their habitats at the trapping locations shown in B.

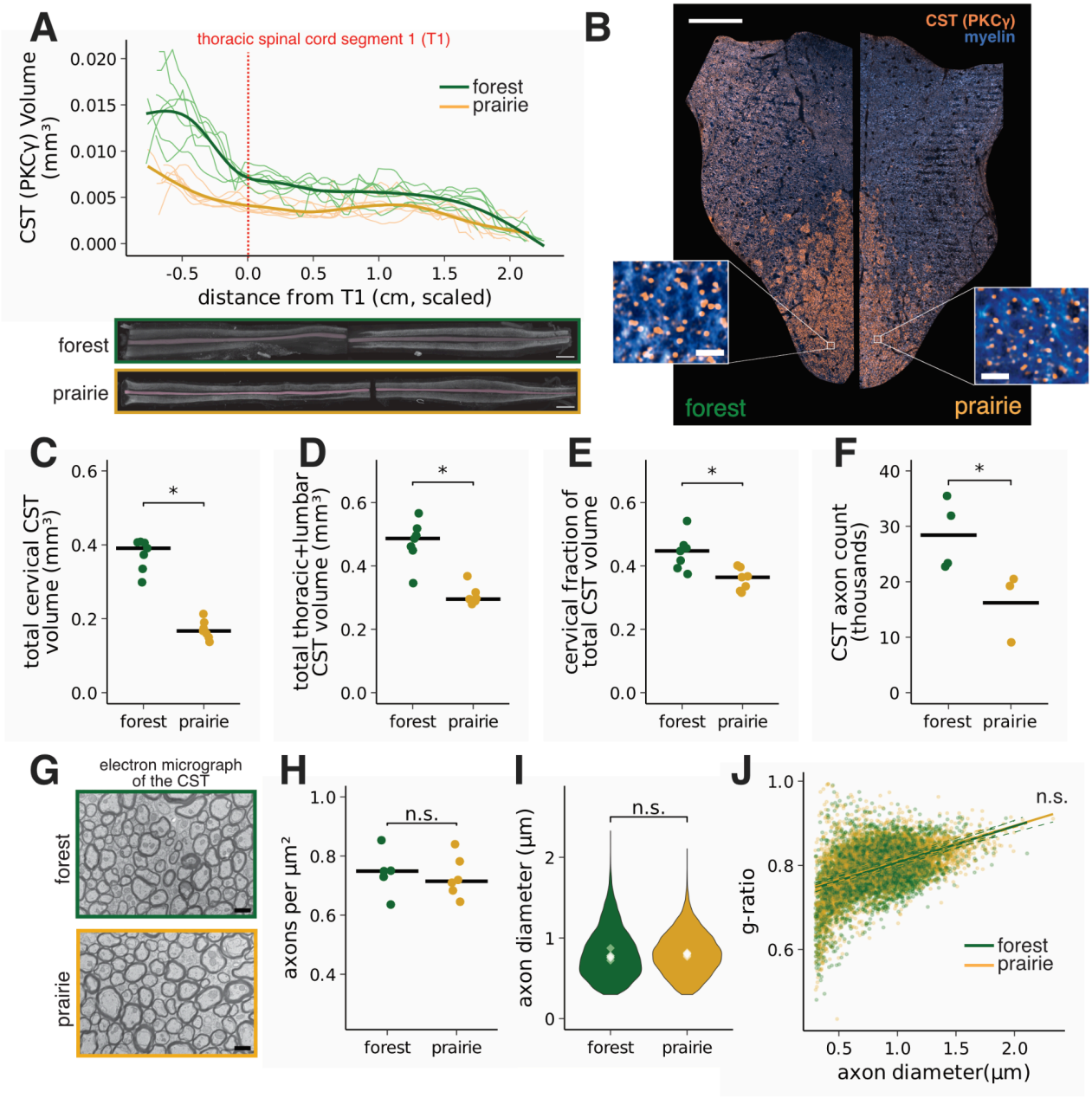


Figure 2: Forest mice have a larger corticospinal tract (CST) due to an increase in axon number A. Comparison of PKCγ+ corticospinal tract size between forest and prairie mice in whole-mount, cleared spinal cords. *Top:* Lines represent PKCγ+ corticospinal volume binned at 140µm and smoothed. Thin lines represent individual animals, and thick lines are a fitted regression for each subspecies. T1 marks the beginning of the thoracic spinal cord. n=7 animals of each subspecies. *Bottom:* Representative projections along the dorsal/ventral axis of 3D images of forest and prairie mouse spinal cords with PKCγ staining (cervical and lumbar segments imaged separately). The corticospinal tract is highlighted in pink. The other signal is PKCγ+ interneurons. Scale bar = 1mm.

B. Representative images of coronally sectioned hemisections of forest and prairie mouse cervical segment three dorsal funiculi stained with FluoroMyelin dye to label myelin and immunostained for PKC**γ**. Scale bar for full-sized image = 100μm. Scale bar for zoomed insets = 2μm.
C. Quantification of data shown in A. Points represent values for individual animals. Black lines represent the median of each subspecies. * = p = 0.0007, linear model comparing cervical CST volume between subspecies, accounting for sex, age, and total spinal cord volume.
D. Quantification of data shown in A. Points represent values for individual animals. Black lines represent the median of each subspecies. * = p = 0.020, linear model comparing thoracic+lumbar CST volume between subspecies, accounting for sex, age, and total spinal cord volume.
E. Quantification of data shown in A. Points represent values for individual animals. Black lines represent the median of each subspecies. * = p = 0.01, rank-sum test.
F. Quantification of the number of PKC**γ**+ axons in a hemisection of the dorsal funiculus (prepared as in B) from n=4 forest and n=3 prairie mice. Points represent values for individual animals. Black lines represent the mean of each subspecies. * = p 0.034, rank-sum test.
G. Representative electron micrographs of the ventral part of forest and prairie coronally sectioned dorsal funiculi. Scale bar = 1μm.
H. Quantification of axon density in electron micrographs from n = 6 animals of each subspecies. Points represent values for individual animals. Black lines represent the median of each subspecies. p = 0.66, rank-sum test.
I. Quantification of axon diameter in electron micrographs from n = 6 animals of each subspecies. Violin plots represent the distribution of axon diameters of all axons measured. White diamonds represent the medians of each animal. p = 0.24, rank-sum test.
J. Quantification of g-ratio (axon diameter/axon+myelin diameter) from n=6 animals of each subspecies. Points represent g-ratios of individual axons measured. Dashed lines are the trend lines for individual animals, and solid lines are the trend lines for each subspecies. p = 0.70, rank-sum test.

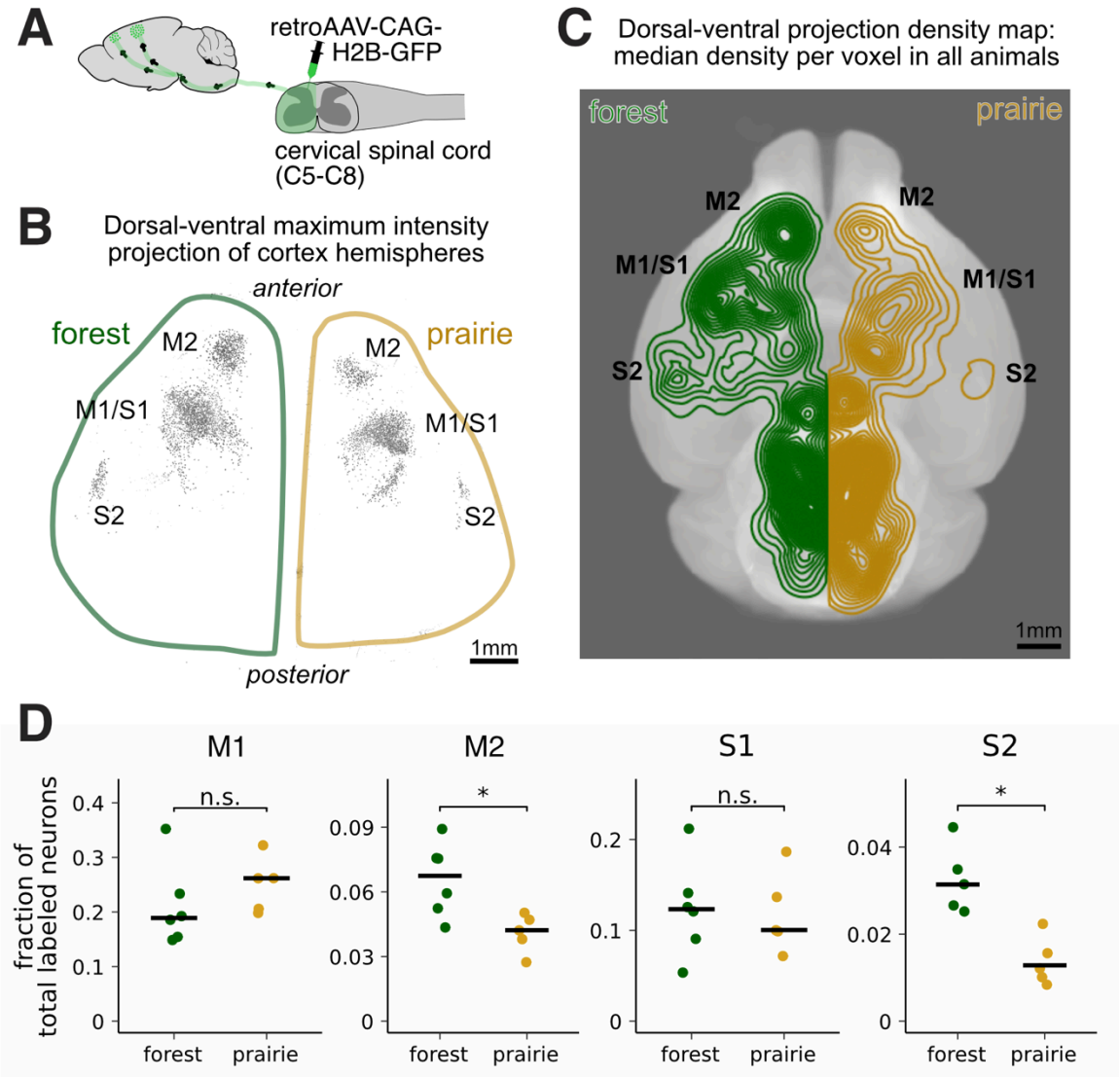


Figure 3: Forest mice have more corticospinal tract neurons specifically in secondary motor and sensory regions (M2 and S2)

A. Schematic of the injection. Retrograde AAV expressing nuclear-tagged-GFP was injected into segments C5-C8 of the spinal cord. We analyzed brains with tagged nuclei.

B. One representative brain from each subspecies showing a maximum intensity projection across the dorsal/ventral axis of the signal from labeled neuronal nuclei in contralateral cortex. Cortical areas labeled are primary motor (M1) and sensory (S1) cortex, and secondary motor (M2) and sensory (S2) cortex. The signal from M1 is in caudal forelimb area (CFA), and from M2 is in rostral forelimb area (RFA).

C. Density plot of the median fraction of nuclei labeled across the contralateral (to the injection) side of the brain. Data are shown as a projection along the dorsal/ventral axis, with cortical areas labeled (as in B). Data from n = 6 forest mouse brains, n = 5 prairie mouse brains.

D. Quantification of the fraction of total labeled neurons in each brain region labeled in B/C. * = FDR-adjusted $q < 0.05$, based on a rank-sum test.

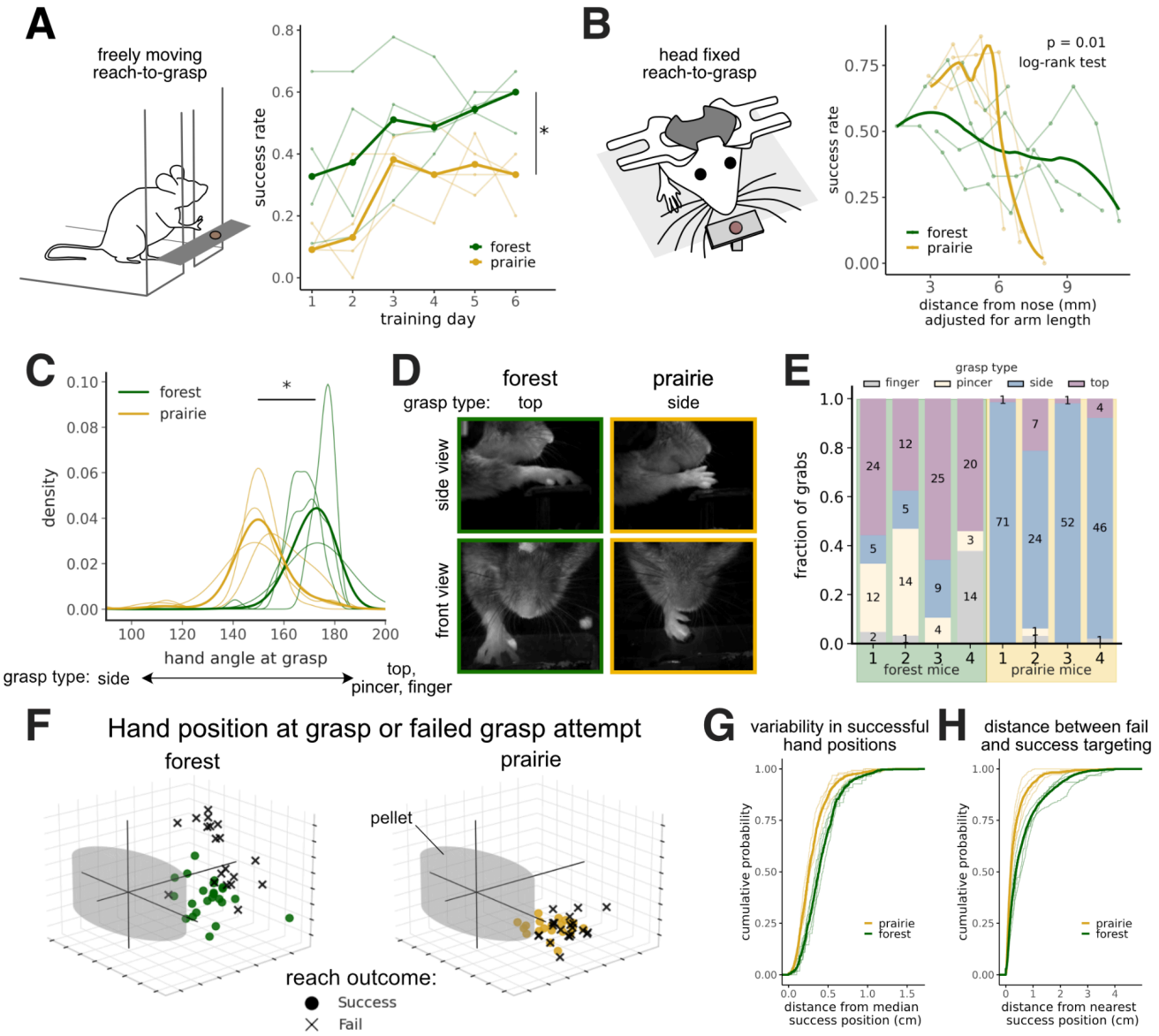


Figure 4: Forest mice show enhanced dexterity in a reach-to-grasp task compared to prairie mice

A. Comparison of forest and prairie mouse performance in a freely moving pellet-grasping task over 6 days of training. Thin lines represent individual animals (n=4 of each subspecies), thick lines represent the subspecies median. Forest mice successfully grasp the pellet at a higher rate, * p =0.018, repeated measures ANOVA on the difference between subspecies.

B. Comparison of forest and prairie mouse performance in a head-fixed pellet-grasping task. Pellet position is adjusted for arm length. Thin lines represent individual animals (n=4 of each subspecies), thick lines represent the subspecies median. Forest mice reach further than prairie mice before failing (p = 0.01, log-rank test).

C. Distribution of the angle of the hand relative to the platform in forest and prairie mice during successful reach attempts at their most successful pellet position, with 180 degrees being parallel to the platform (top grasp type). Lighter lines show distributions of individual animals (n=4 of each

subspecies), darker lines show the distributions of the subspecies. Forest mouse hand positions are closer to parallel to the platform (top-down reach) at the time of grasp, *p = 0.02, rank-sum test on median hand angle between subspecies.

D. Examples of images of forest and prairie mouse typical grasp types at the time of the start of the grasp.

E. Quantification of the frequency of four qualitative grasp types in forest and prairie mice. Data quantified are from each animal's preferred (i.e., most successful) pellet position. Forest mice have a greater diversity of grasp types. Grasp types are defined as follows: top = an approach from the top that uses all fingers to grasp the pellet; side = a scooping approach from the side; pincer = a grasp in which the pellet is grasped resting on the thumb-like digit one with another finger; and finger = when the pellet is grasped in between two fingers.

F. Example data plotting the 3D tracked hand position, i.e., center of the hand, at the time of grasp (or grasp attempt in the case of a fail trial). Each plot shows the data from one animal on one day at their preferred pellet position.

G. Cumulative distribution plot showing, for each successful reach attempt, the variability of target positions, i.e., the median absolute deviation or the distance of the hand at grasp from the median successful hand position. Light lines show individual animals, dark lines show subspecies distributions. Forest mice have a wider range of success positions than prairie mice (p = 0.03, rank-sum test comparing medians of each animal; significant (p = 0.0015) effect of subspecies on variability by ANOVA accounting for pellet position and distance of the hand from the pellet).

H. Cumulative distribution plot showing, for each failed reach attempt, the closest the hand comes to a successful grasp position. Light lines show individual animals, dark lines show subspecies distributions. Forest mice are further from a successful position when they fail (p = 0.03, rank-sum test comparing medians of each animal), suggesting forest mice also fail more often due to targeting their hand to the wrong position, whereas prairie mice fail at grasping even if their hand is targeted well.

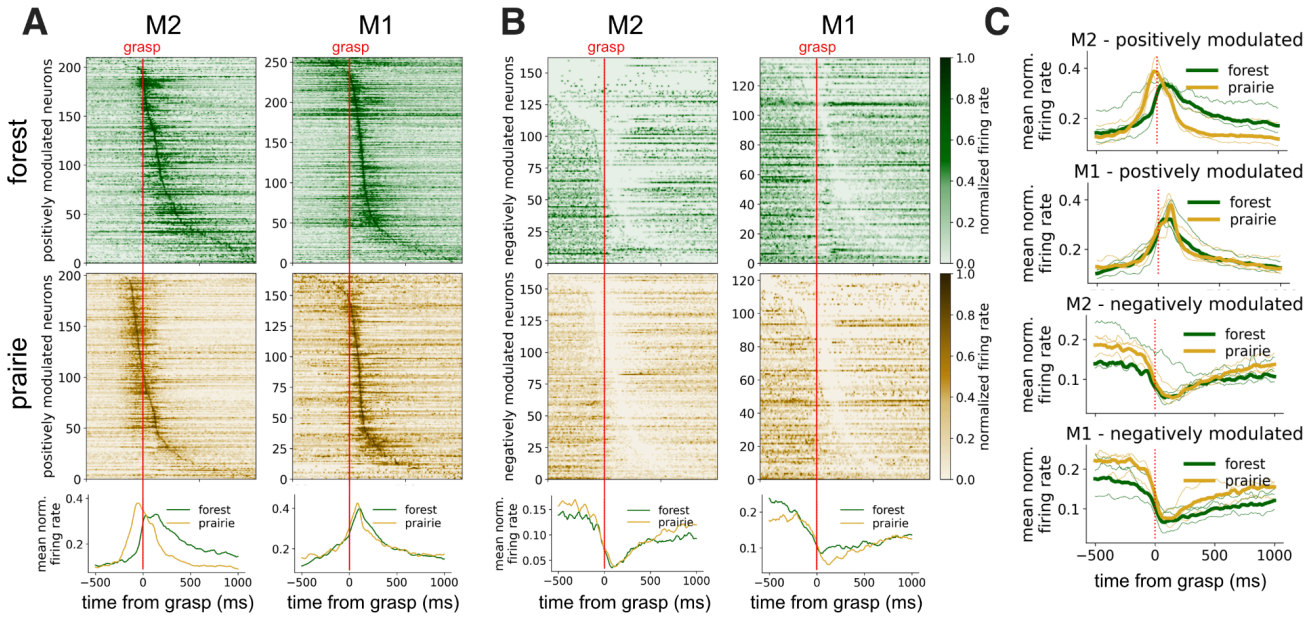


Figure 5: Neural activity in forest mouse M2 peaks closer to the grasp

A. (*top*) Each heatmap shows normalized firing rates for all positively modulated units recorded in a single region in a single example animal, averaged across successful trials (forest mouse sample recorded over three sessions, prairie mouse animal recorded over five sessions). Data is aligned to the time of the grasp. (*bottom*) Average firing rate across all units for the example units shown.

B. (*top*) Each heatmap shows normalized firing rates for all negatively modulated units recorded in a single region in a single example animal, averaged across successful trials(forest mouse sample recorded over three sessions, prairie mouse animal recorded over five sessions). Data is aligned to the time of the grasp. (*bottom*) Average firing rate across all units for the example units shown.

C. Average firing rate across all positively or negatively modulated units for n=3-4 forest and n=3 prairie mice. Thick lines represent the average of all units recorded in all animals, and thin lines represent the units from individual animals. Red lines mark the time of grasp.

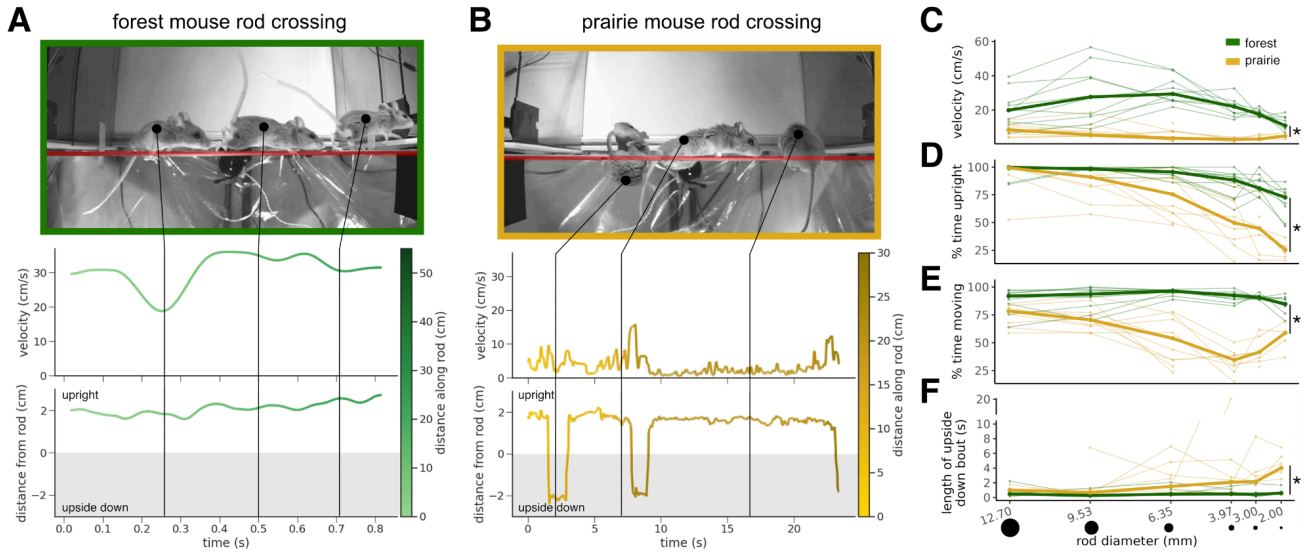


Figure 6: Forest mice show greater climbing dexterity than prairie mice

A/B. Example data from one forest (A) and one prairie (B) mouse crossing a 3.97mm rod. In this task, animals cross a rod between two platforms to receive a reward. Automated tracking of the back of the animal was used to determine the velocity and distance from the rod. The top image is made from video frames at the time points indicated by black lines. The rod is highlighted in red.

C-F. Measures of rod crossing performance across six different rod sizes. Animals crossed wider (easier) rods first and then were presented with progressively narrower rods. Thick lines show median performance, and thin lines show the performance of n=10 forest and n=11 prairie mice.

C. Forest mice move across the rod faster than prairie mice (*p=9.6e-07, repeated measures ANOVA).

D. Forest mice spend more time upright than prairie mice (*p=0.00019, repeated measures ANOVA).

E. Forest mice spend more time moving than prairie mice (*p=3.6e-08, repeated measures ANOVA).

F. Forest mice right themselves more quickly when they are upside down compared to prairie mice (*p= 0.0058, repeated measures ANOVA).

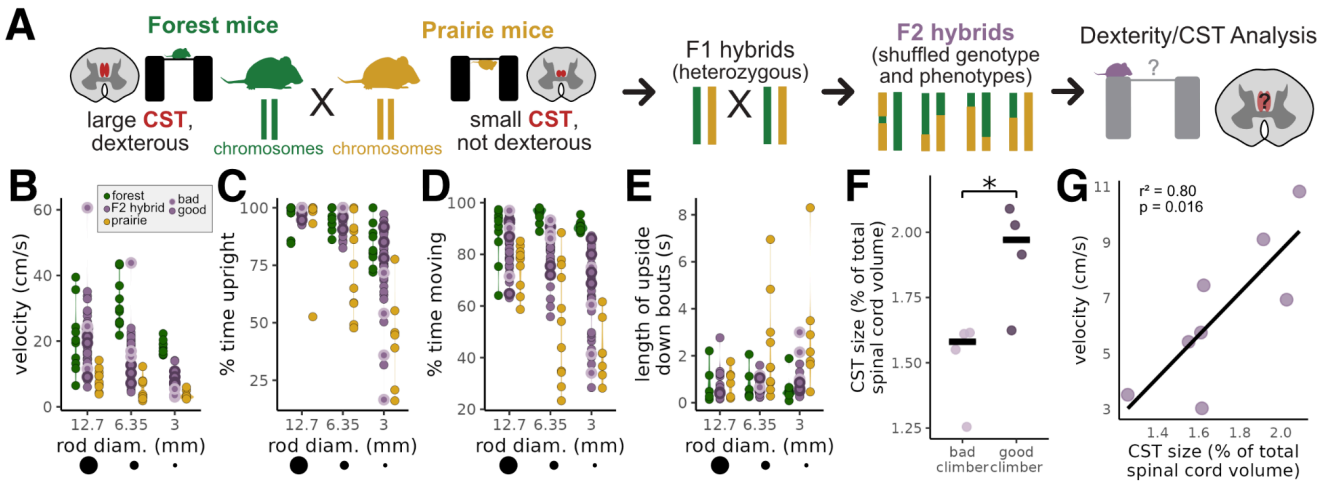


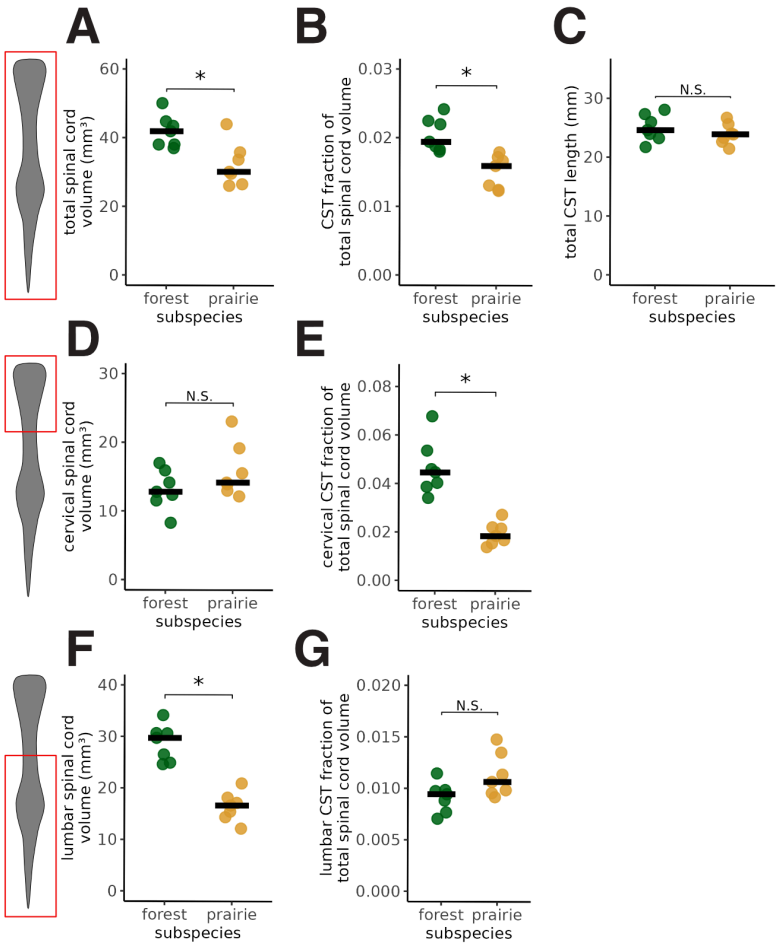
Figure 7: Corticospinal tract size is linked to climbing dexterity

A. Schematic of the F2 hybrid generation and testing. F2s were generated by crossing forest and prairie mice for two generations. F2 animals have shuffled genomes and phenotypes and can thus be used to separate subspecies' phenotypes.

B-E. Quantification of rod crossing performance in F2 hybrid animals (forest/prairie data from Figure 6 shown for comparison). Good and bad climbers chosen for CST analysis are indicated. Data from n=51 F2 animals. F2s show a broad distribution of climbing dexterity.

F. Good climber F2 animals have larger CSTs than bad climber F2 animals. Data is normalized by total spinal cord size. *p = 0.029, rank-sum test.

G. CST size is correlated with rod-crossing velocity in F2 animals. R^2 from Pearson correlation; r^2 = 0.81, p = 0.014 for a Pearson correlation accounting for morphological traits (weight, tail length, foot size).



Supplementary Figure 1: Corticospinal tract (CST) morphology

A. Forest mice have larger spinal cords. *p = 0.011, rank-sum test.

B. The CST comprises a larger fraction of total spinal cord volume in forest mice. *p = 0.00058, rank-sum test.

C. Total CST length (from cervical to lumbar spinal cord) is not different between forest and prairie mice. N.S.: p = 0.38, rank-sum test.

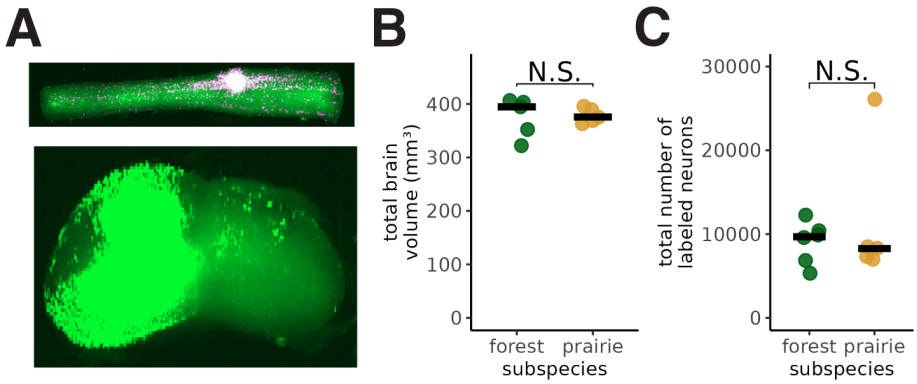
D. Cervical spinal cord volume is not different between forest and prairie mice. N.S.: p = 0.32, rank-sum test.

E. The CST comprises a larger fraction of cervical spinal cord volume in forest mice. *p = 0.00058, rank-sum test.

F. The lumbar spinal cord volume is larger in forest mice compared to prairie mice. *p = 0.00058, rank-sum test.

G. The CST comprises the same fraction of lumbar spinal cord volume in forest and prairie mice. N.S.: p = 0.073, rank-sum test.

All data is the same as in Figure 1 from n=7 animals of each subspecies.

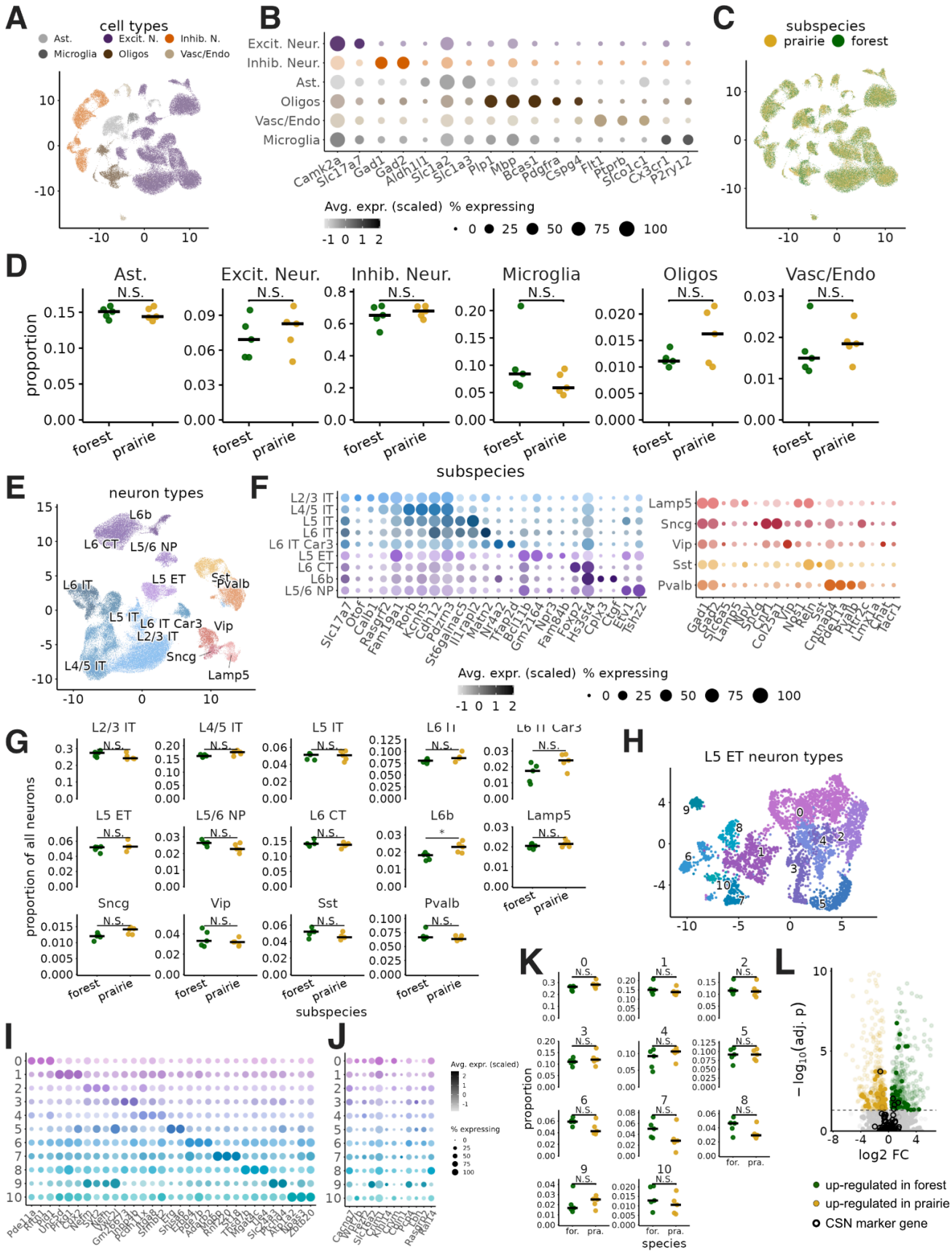


Supplementary Figure 2: Retrograde viral tracing

A. Example light-sheet images from a spinal cord injection. *Top*: maximum intensity projection along the dorsal-ventral axis of cervical/thoracic spinal cord, showing the injection site. Magenta = anti-GFP staining, green = GFP signal. *Bottom:* maximum intensity projection along the anterior-posterior axis of the injection site from the same animal. Only GFP signal is shown.

B. Forest and prairie mice do not differ in total brain volume, $p = 0.54$, rank-sum test. Data from n=6 foresta and n=5 prairie mice.

C. Forest and prairie mice did not differ in the total number of neurons labeled throughout the brain by retrograde viral injection, $p = 0.93$, rank-sum test. Data from n=6 foresta and n=5 prairie mice.


A
cell types
Ast.
Excit. N.
Inhib. N.
Microglia
Oligos
Vasc/Endo
B
Excit. Neur.
Inhib. Neur.
Ast.
Oligos
Vasc/Endo
Microglia
Camk2a
Slc17a7
Gad1
Gad2
Aldh1l1
Slc1a2
Slc1a3
Plp1
Mbp
Bcas1
Pdgfra
Cspg4
Flt1
Pteprb
Slco1c1
Cx3cr1
P2ry12
Avg. expr. (scaled)
% expressing
C
subspecies
prairie
forest
D
Ast.
Excit. Neur.
Inhib. Neur.
Microglia
Oligos
Vasc/Endo
N.S.
proportion
forest
prairie
subspecies
E
neuron types
L6b
L6 CT
L5/6 NP
L6 IT
L5 ET
Sst
Pvalb
L5 IT
L6 IT Car3
L2/3 IT
Vip
Sncg
L4/5 IT
Lamp5
F
L2/3 IT
L4/5 IT
L5 IT
L6 IT
L6 IT Car3
L5 ET
L6 CT
L6b
L5/6 NP
Lamp5
Sncg
Vip
Sst
Pvalb
Avg. expr. (scaled)
% expressing
G
L2/3 IT
L4/5 IT
L5 IT
L6 IT
L6 IT Car3
L5 ET
L5/6 NP
L6 CT
L6b
Lamp5
Sncg
Vip
Sst
Pvalb
N.S.
*
proportion of all neurons
forest
prairie
subspecies
H
L5 ET neuron types
I
J
Avg. expr. (scaled)
% expressing
K
proportion
for.
pra.
species
L
log2 FC
up-regulated in forest
up-regulated in prairie
CSN marker gene

Supplementary Figure 3: Single-nucleus RNA sequencing reveals no difference in L5ET or cortical neuron number between subspecies

A. UMAP plot of gene expression in all cortical nuclei isolated from forest and prairie mice, colored by their cell class, as determined by alignment to the BICNN database. Data is from 107,000 nuclei from n=5 animals from each subspecies. (Note that this is after integration – see Methods).

B. Expression of established marker genes for cell class of the neural populations shown in A.

C. UMAP plot of gene expression in cortical nuclei isolated from forest and prairie mice, colored by their subspecies.

D. Percentage of nuclei per sample assigned to each cell class. N.S. = $p > 0.05$, edgeR.

E. UMAP plot of gene expression in neuronal nuclei (as defined in A) of both forest and prairie mice. Coloring is by their subclass.

F. Expression of established marker genes for cell subclass of the neural populations shown in B.

G. Percentage of nuclei per sample assigned to each cell subclass. * = $p<0.05$, N.S. = $p > 0.05$, edgeR.

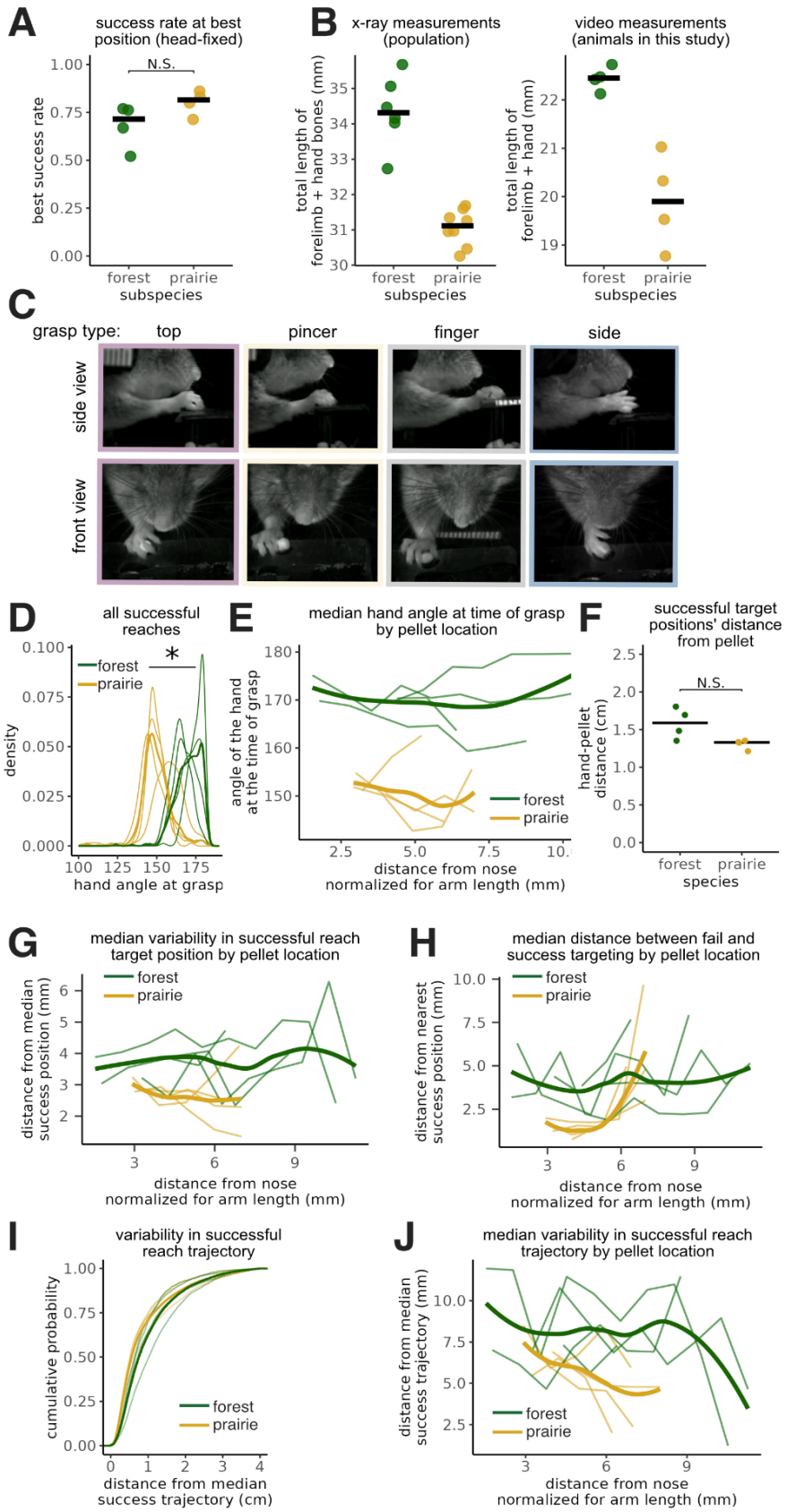
H. UMAP plot of gene expression in L5ET nuclei (as defined in E) of both forest and prairie mice. Coloring is by unbiased clustering results.

I. Expression of the top 3 marker genes for each cluster of L5ET neurons.

J. Expression of known CSN marker genes from Golan et al, 2023, and Arlotta et al., 2005 for each cluster of L5ET neurons.

K. Percentage of nuclei per sample in each cluster. N.S. = $p > 0.05$, edgeR.

L. Differential gene expression analysis of pseudo-bulk gene expression in L5ET neurons between forest and prairie mice. Dark circles mark CSN marker genes in Golan et al. Forest mouse differentially expressed genes are enriched for CSN marker genes ($p<0.05$, Fisher's exact test). Genes differentially expressed between subspecies in all neurons are shown in a lighter alpha shade.


A
success rate at best position (head-fixed)
N.S.
best success rate
1.00
0.75
0.50
0.25
0.00
forest
prairie
subspecies
B
x-ray measurements (population)
total length of forelimb + hand bones (mm)
35
34
33
32
31
30
forest
prairie
subspecies
video measurements (animals in this study)
total length of forelimb + hand (mm)
22
21
20
19
forest
prairie
subspecies
C
grasp type:
top
pincer
finger
side
side view
front view
D
all successful reaches
*
forest
prairie
density
0.100
0.075
0.050
0.025
0.000
100
125
150
175
hand angle at grasp
E
median hand angle at time of grasp by pellet location
angle of the hand at the time of grasp
180
170
160
150
forest
prairie
2.5
5.0
7.5
distance from nose normalized for arm length (mm)
F
successful target positions' distance from pellet
N.S.
hand-pellet distance (cm)
2.5
2.0
1.5
1.0
0.5
0.0
forest
prairie
species
G
median variability in successful reach target position by pellet location
forest
prairie
distance from median success position (mm)
6
5
4
3
2
3
6
9
distance from nose normalized for arm length (mm)
H
median distance between fail and success targeting by pellet location
forest
prairie
distance from nearest success position (mm)
10.0
7.5
5.0
2.5
3
6
9
distance from nose normalized for arm length (mm)
I
variability in successful reach trajectory
cumulative probability
1.00
0.75
0.50
0.25
0.00
forest
prairie
0
1
2
3
4
distance from median success trajectory (cm)
J
median variability in successful reach trajectory by pellet location
distance from median success trajectory (mm)
10.0
7.5
5.0
2.5
forest
prairie
3
6
9
distance from nose normalized for arm length (mm)

Supplementary Figure 4: Hand dexterity in a reach-to-grasp task

A. In a head-fixed reach-to-grasp task (Figure 4B), forest and prairie mice achieve equivalent dexterity. The success rate at each animal's most successful pellet position is not significantly different between subspecies (p=0.114, rank-sum test, n=4 of each subspecies).

B. Arm length measurements from x-ray data from Hager et al., 2020 (left) and those made from calibrated video images used in this study (right). We find similar differences between populations (x-ray: mean 3.2mm +/- 1.06 95% CI, video: mean 2.5mm +/- 1.47 95% CI), indicating that the values we used to normalize our data are reasonable. The difference in absolute lengths between x-ray and video measurements come from: (1) x-ray measurements sum bones that overlap and do not contribute to overall bone length, (2) x-ray measurements start at the shoulder socket, which is difficult to determine in video images. Video measurements were made blinded to the animal's identity.

C. Examples of the qualitative grasp types (Figure 4E). Images taken just as the animal begins to move the pellet.

D. Forest mice use a more parallel, top-down grasp compared to their prairie counterparts. This is the same as Figure 4B, except using all successful reaches, not just those at the animals' most successful positions. Lighter lines show distributions of individual animals (n=4 of each subspecies), darker lines show the distribution of the subspecies. Forest mouse hand positions are closer to parallel to the platform (top-down reach) at the time of grasp, *p = 0.03, rank-sum test on median hand angle between subspecies.

E. Forest mice use a top-down grasping strategy, and prairie mice use a side grasping strategy regardless of pellet position. Lighter lines show distributions of individual animals (n=4 of each subspecies), darker lines show the distribution of the subspecies. There is a significant impact of subspecies on hand angle (p=0.013) in a linear model accounting for pellet position.

F. Despite differences in hand morphology, forest mice do not show significant differences in the distance of their hand from the pellet during targeting (p = 0.057, rank-sum test). Data from n=4 of each subspecies, median distance between hand target position and pellet position plotted.

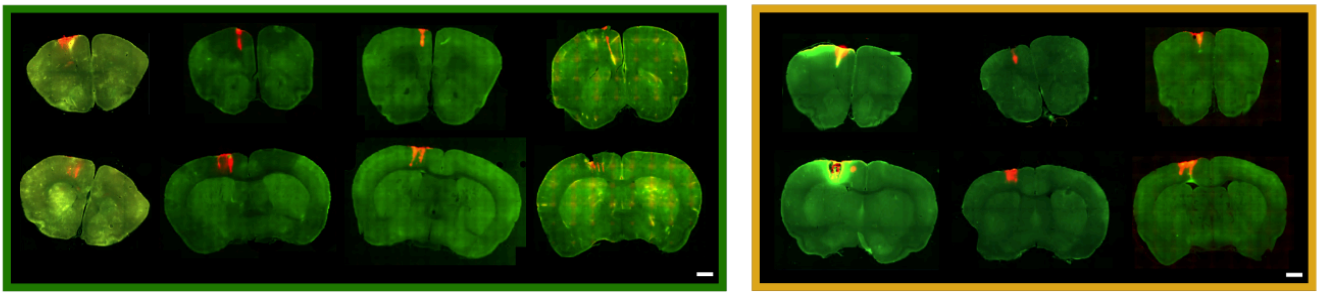
G. Forest mice have greater variability in targeting the pellet at all pellet positions. This is the same data as in Figure 4G, divided by pellet position. Lighter lines show the median MAD (median absolute difference) for each pellet position; darker lines show the subspecies median. A linear model accounting for pellet position and hand distance from the pellet for each reach finds a significant contribution of subspecies to targeting variability (p = 0.002).

H. Forest mice have greater variability in targeting the pellet even when accounting for pellet position. This is the same data as in Figure 4H, divided by pellet position. Lighter lines show the median distance between failure and successful reaches for each pellet position; darker lines show the subspecies median. A linear model accounting for pellet position for each reach finds a significant contribution of subspecies to fail-success distance (p = 5.438096e-07), as well as a significant effect of pellet position (p = 1.69e-13).

I. Cumulative distribution plot showing, for each successful reach attempt, the variability of 3D reach trajectory, i.e., the median absolute deviation or the distance of each trajectory from the median successful 3D trajectory. Light lines show individual animals, dark lines show subspecies distributions.

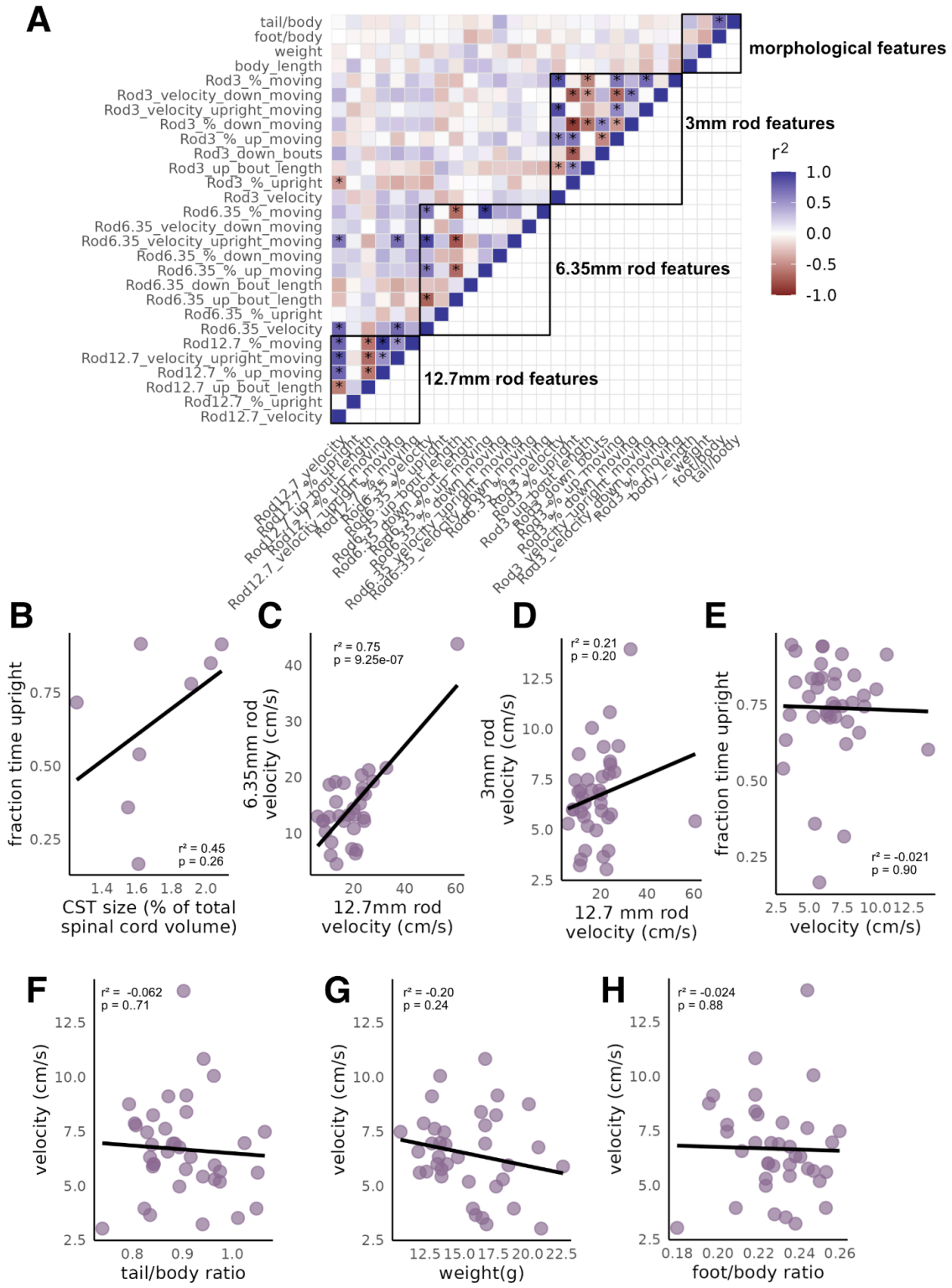
Forest mice have a wider range of trajectories than prairie mice (p = 0.01, effect of subspecies in a linear model accounting for distance from pellet.

J. Forest mice have greater variability in their trajectories regardless of pellet position. This is the same data as (I) divided by pellet position. Lighter lines show the median distance from the median trajectory for each pellet position; darker lines show the subspecies median. A linear model accounting for pellet position for each reach finds a significant contribution of subspecies to trajectory variability (p = 0.01).



Supplementary Figure 5: Tract traces from neural recordings.

Sections from forest (left) and prairie (right) deer mouse brains showing the location of the neural recording probe in each animal from DiI staining. Each column is one animal. The top row is RFA, and the bottom row is CFA. Scale bar is 1mm.



Supplementary Figure 6. F2 hybrid mice

A. Correlation matrix of behavioral and morphological features in F2 animals. * = q <0.05, FDR-corrected p-values on r^2 from a Perason correlation. Data is from n=51 F2 animals.

B. Fraction of time upright on the rod does not significantly correlate with CST size. r^2 from a Pearson correlation. Data is from n=8 F2 animals in which we measured CST size.
C. Performance on the larger two rods is significantly correlated (data same as in A). r^2 from a Pearson correlation.
D. Performance on the largest and smallest rods was not significantly correlated (data same as in A). r^2 from a Pearson correlation.
E. On the smallest rod, velocity and the percentage of time spent upright are not significantly correlated (data same as in A). r^2 from a Pearson correlation.
F. Velocity is not correlated with tail length (normalized to body size, data same as in A). r^2 from a Pearson correlation.
G. Velocity is not correlated with weight (data same as in A). r^2 from a Pearson correlation.
H. Velocity is not correlated with hind foot length (normalized to body size, data same as in A). r^2 from a Pearson correlation.

Methods

Animals

We used two subspecies of deer mice: *P. m. maniculatus* and *P. m. nubiterrae*, which are representatives of prairie and forest mice, respectively. Forest mice were descendants of a colony started from wild-caught deer mice trapped in Rector, PA, in 2010 (Kingsley et al., 2017). Prairie mice were descendants of the colony at the Peromyscus Genetic Stock Center (University of South Carolina), which are descended from wild-caught mice originally captured in Ann Arbor, MI, in 1948. Colonies are maintained using an outbred breeding scheme to maintain maximal genetic diversity. Animals used in this study were between 3 months and 1 year of age. We used animals from different families (i.e., did not use only siblings/littermates) in all experiments to prevent confounding differences between subspecies with differences between families.

At Harvard, mice were housed at 23°C on a 16:8 h light:dark cycle in standard mouse cages with corncob bedding, cotton nestlet, Enviro-Dri, and a red tube. Mice were fed standard lab mouse chow (LabDiet Prolab Isopro RMH 3000 5P75) ad libitum. At UNC, mice were housed at 23°C on a 12:12 h light:dark cycle in standard mouse cages with corncob bedding, cotton nestlet, and a red hut. Mice were fed standard lab mouse chow (LabDiet 5V5R - PicoLab® Select Rodent 50 IF/6F) ad libitum. All experiments were conducted in accordance with Harvard IACUC protocol 11-05 or UNC School of Medicine IACUC protocol 23-198.

Spinal cord/brain clearing and antibody staining

For CST staining in forest and prairie subspecies and retrograde tracing experiments, animals were transcardially perfused with PBS with heparin (20U/mL) followed by 4% paraformaldehyde in PBS. Brains were dissected out of the skull, and spinal cords were dissected using hydraulic extrusion (Richner et al., 2017) and then post-fixed overnight in 4% paraformaldehyde (PFA). Samples were cleared using a solvent-based method (adapted from Chi et al., 2018). Briefly, they were dehydrated at 4°C using a methanol gradient with a minimum of 30 minutes at each step (20%, 40%, 60%, 80%, 100%) in B1n buffer (0.3 M glycine, 0.1% Triton X-100, 0.01% sodium azide, pH 7). They were then placed in dichloromethane overnight at 4°C and then rehydrated using the reverse methanol gradient. The spinal cords were then placed in B1n buffer overnight. Brains were further cleared using the uniclear method (Chen and Svoboda, 2020). They were incubated shaking at 37°C for 1 week in SiBP buffer (0.2mM, 0.08% SDS, 16% 2-methyl-2-butanol, 8% 2-propanol in water), changing the buffer after 3 hours, 6 hours, and then daily for the next week. They were then placed in B1n buffer overnight.

To stain, samples were incubated with antibody in PTxw buffer (PBS, 0.1% Triton X-100, 0.05% Tween20, 0.01% sodium azide) for 3-7 days (spinal cords) or 10-14 days, with a refresh in antibody after 5-7 days (brains), rotating at room temperature. They were then washed in PTxw and incubated with 1:200 secondary donkey-anti-mouse/rabbit-AlexaFluor647 antibody for 3 days (spinal cords) or 10 days (brains). Antibodies used were: A-7 mouse-anti-PKCy (sc-166451) at a concentration of 1:200, rabbit-anti-GFP (Invitrogen, A-11122) at 1:1000. They were then washed again with PTxw.

To image, spinal cords were refractive index matched in RI solution (a 60:40 mix of OptiPrep and 2,2-thiodiethanol with 0.6g/mL Histodenz) overnight. To fit into the microscope, spinal cords were cut

in half at the thoracic segments before imaging. Each half was mounted in a glass capillary tube. They were then imaged in RI solution in a Zeiss lightsheet microscope with a 5x objective, and a 1.525 refractive index. Brains were refractive index-matched and imaged in EasyIndex from LifeCanvasTechnologies (EI-500-1.52), also in a Zeiss lightsheet microscope with a 5x objective, and a 1.52 refractive index.

For F2 experiments, we modified our clearing protocol due to changes in university regulations on the use of dichloromethane. We cleared samples using the Life Canvas Technology passive clearing kit (C-PCK-500-1.52) following the instructions of the manufacturer. We note that our previous method altered spinal cord volume more than this method, and absolute volumes are not comparable between methods. Antibody staining was performed as described above. Spinal cords were imaged in EasyIndex on a Zeiss lightsheet microscope with a 5x objective and a 1.52 refractive index.

CST volume quantification

To quantify the corticospinal tract (CST) size, images from the microscope were first stitched together in Arivis software. Then, using Arivis, we developed a pipeline that identifies the CST using a machine learning model. We then divided the CST into segments along the anterior-posterior axis of the spinal cord and quantified volume in those segments. Total spinal cord volume was quantified by thresholding the green autofluorescence signal. To compare between animals, the distance along the spinal cord was scaled using the length of the thoracic segment (we did not use total length due to the potential for differences in where the cut between the brain and spinal cord was made during dissection). For plotting, data were aligned such that T1 is the 0 point in all scaled data. Data were smoothed using the lowess function in R with f set to 0.5. Statistical testing for volume differences was performed in R using a linear model: CST volume ~ subspecies + age + total spinal cord volume + sex. We observed significant differences only for subspecies. Rank-sum tests were performed using the wilcox.text function in R.

Dorsal funiculus axon immunostaining, confocal imaging, and quantification

Animals were perfused and spinal cords fixed as described above. Spinal cords were then mounted in agarose and sectioned on a vibratome at 75 μm thickness. We note that we found it is imperative not to freeze the sample to enable observation of axons. After sectioning, spinal cords were blocked in normal donkey serum, immunostained as free-floating sections with 1:200 A-7 PKCy antibody, and gently shaken at room temperature overnight. They were then washed with PBS and stained with AlexaFluor 55 or 647 secondary antibody and Green FluoroMyelin (F34651) at 1:300 for 3 hours at room temperature. The dorsal funiculus of each section was imaged on a Zeiss LSM-900 confocal microscope with a 60x oil objective with the following settings: 488 laser at 0.10% and 0.88 AU pinhole, 640 laser at 0.6% and pinhole of 0.78 AU or 561 laser at 0.2% power with 0.87 AU pinhole, scan speed of 4, gain of 650V, and averaging of 2.

Image tiles were stitched in Zen software and then exported as TIFF files. The CST area was manually segmented out using thresholding of the PKCy signal, with the experimenter blinded to the subspecies of the animal. Then, using custom Python code, 10-20 random ROIs, either 500x500 pixels or 250x250 pixels within the CST were selected for axon counting. Due to injury to the tissue during sectioning and

staining, not all ROIs had quantifiable axon signal, and ROIs without such signal were manually filtered out with the experimenter blinded to the subspecies of the animal, ensuring that at least 5 500x500 ROIs or 10 250x250 ROIs remained. The FluoroMyelin signal on good squares was inverted, and 6 ROIs were used to train a custom Cellpose (Stringer et al., 2021) model. The Cellpose model was then run on all good ROIs and manually checked to confirm good axon detection. We then filtered out any ROIs from Cellpose smaller than 0.1 um^2 as likely noise. We chose this value as the division between two peaks in a histogram of ROI size. Total axon number for each animal was calculated by averaging the number of axons per micron in all of the ROIs counted, and then multiplying by the area of the CST. Statistical analysis was performed in python using the scipy.stats package.

Electron Microscopy preparation, imaging, and quantification

Animals were transcardially perfused with cold Ames buffer followed by cold 2% paraformaldehyde/2.5% glutaraldehyde in 0.1M cacodylate pH 7.4 with 0.04% calcium chloride. Spinal cords were dissected out and then incubated overnight in the same solution. After washing, spinal cords were coronally sectioned on a vibratome, followed by fixing of the sections in 3% glutaraldehyde in 0.1M cacodylate pH 7.4 with 0.04% calcium chloride. The dorsal funiculus was then dissected out and used in the following steps. After washing in water and maleate buffer pH 5.2, sections were incubated up to 24h in 1% uranyl acetate. They were then washed again in a maleate buffer, followed by water. They were then incubated in lead aspartate for 3h at 60°C. After washing with water, samples were dehydrated in an ethanol gradient at 4°C. They were then incubated in 100% ethanol at room temperature, then in propylene oxide. They were then embedded in LX112 resin. Sections were then sectioned on a microtome to create ultrathin sections. Immediately before imaging, they were treated with 0.2% lead citrate. Sections were imaged on a transmission electron microscope at 4800x. We imaged the ventral-most region of the dorsal funiculus and imaged at least 12 regions per dorsal funiculus.

To quantify the images, we used AxonDeepSeg (Zaimi et al., 2018) to segment myelin and axons, following instructions for the published version with an additional myelin cutter model (https://github.com/axondeepseg/nn-axondeepseg) to better separate individual axons. We manually inspected and adjusted a subset of the images (10%) drawn from all animals. We compared the results from our manually refined dataset to the whole dataset and observed no differences, so we did not perform the manual refinement for the final data analysis (which enabled us to quantify all 12 images per animal). We used AxonDeepSeg to generate tables of axon size, myelin thickness, and g-ratio that we then analyzed using custom R scripts.

Retrograde labeling surgeries

Animals were anesthetized with isoflurane for the duration of the surgery. After shaving the back, an incision was made in the skin. After tearing the muscle with forceps, the spinal column was clamped and stabilized using the Narishige spinal cord clamp (STS-A). The T1 vertebrae was located, and cervical injections were made 2-3 vertebral segments anterior to T1. The dura was removed from the spinal cord using a needle. Using a stereotaxic manipulator holding a pulled glass pipette, 25nL of virus was injected 1mm lateral from the midline (as defined by the central artery) at three depths: 500nm, 300nm, 200nm, for a total of 75nL injected. Each injection was done over 3 minutes with 3

minutes following the injection without moving the needle. Gelfoam was added to the spinal column to aid in healing, and the wound was closed with wound clips. Animals were perfused (see below) 3 weeks following the surgery. Animals were injected with SL1-AAV (Tervo et al., 2016) with H2B-linked green fluorescent protein (GFP) under the CAG promoter.

Brain-wide retrograde labeling counting

Image tiles were stitched in Arivis software and then exported as TIFF files. Tiff files were stitched together into full 3D brain images and downsampled 3x using ANTS (Avants et al., 2009). Images were then subdivided into smaller, overlapping 3D cubes (512x512x512 voxels) for counting. An ilastik (Berg et al., 2019) model using a subset of the data taken from all brains was trained on the far red channel to identify GFP+ nuclei, background, and noise signal. We eliminated two samples at this point (one of each subspecies) due to poor perfusion, resulting in pervasive signal from blood vessels. Using custom Python scripts, the GFP+ nuclei masks returned from ilastik were stitched together and assigned labels (where each nucleus has a unique label). To compare the same brain regions between animals, the green channel from the brain imaging (mostly autofluorescence because the clearing process quenches most of the GFP signal) was registered to a template brain generated by averaging the green channel autofluorescence signal of 40 non-injected brains imaged with a lightsheet microscope. Label files were then further downsampled, again using custom Python scripts, preserving the total count, and this downsampled image was transformed into template brain space using the registration from the green channel signal. Cortical regions were annotated based on manual adjustment of the Allen brain annotations registered to our template brain. To normalize for viral injection variation, we normalized nuclei counts by the total number of nuclei in the brain. To ensure that this baseline count was not excessively affected by noise, we masked ventricles and other locations where there was noise common across samples. Furthermore, we created an average of the signal across samples (which was made up of samples that had injections on both sides of the spinal cord), and we masked areas where signal was either non-symmetrical or did not pass a threshold. We then used these regions to quantify the total number of neurons in each brain. We eliminated S2 from the analysis of one animal because of tissue damage in this region, resulting in no identifiable GFP signal. Downstream data analysis and plotting were performed using custom Python and R scripts.

Single nucleus RNA-sequencing

Five animals of each subspecies were sacrificed using carbon dioxide, followed by cervical dislocation. Brains were immediately removed, and the cortex was dissected from the rest of the brain. The anterior half of the cortex was extracted from both hemispheres for each animal (this includes all of sensorimotor cortex, as well as other areas) and flash-frozen in liquid nitrogen. Samples were stored at -70°C until further processing. Nuclei were extracted as described (Matson et al., 2018) except that the high sucrose buffer had 1.2M sucrose. Sequencing libraries were prepared using the Parse Evercode WT Mega kit (along with other samples for different projects) according to the instructions of the manufacturer. Samples were sequenced on a NovaSeq S4. Data were demultiplexed using python-based pipelines from Parse and aligned to the *P. maniculatus bairdii* genome (Peromyscus maniculatus (North American deer mouse) Genome Assembly). To avoid confounds from mapping to the genome of one of the subspecies, before alignment, genomes were “personalized” to each subspecies, using a method similar to the generation of the *P. polionotus* genome in (Bendesky et al.,

2017). Briefly, we aligned our data to the *P. maniculatus bairdii* genome using STAR aligner (Dobin et al., 2013). Then we performed variant calling for the individuals we sequenced for each subspecies using GATK. We then incorporated variants back into the final genomes to which we mapped reads. We did the final alignment using STAR aligner as a part of the Parse pipeline. Cells that passed a UMI threshold determined automatically by the Parse preprocessing script were carried on to the next steps.

Data were analyzed as a cell x gene count matrix in Seurat version 4.3.0. Data were normalized with the SCTransform method and integrated across samples and subspecies using harmony (Korsunsky et al., 2019). Data were then clustered using the default Seurat settings with the FindClusters and FindNeighbors functions. Data were then mapped to the BICNN mouse motor cortex dataset (Callaway et al., 2021) using Azimuth to identify cell types. Clusters with low prediction scores for the mapping (median prediction score < 0.75 out of 1.0) were further assessed by finding marker genes. If clusters showed markers of multiple larger cell types (e.g., neurons and oligodendrocytes) or if they appeared to be non-cortical cell types (e.g., from dissection error), they were removed from the dataset. After removing this contaminating signal, this alignment resulted in 96% of cells with a prediction score greater than 0.75 when mapping to gene classes. The remaining clusters were then re-normalized, re-integrated, and re-clustered. We assessed differential abundance between subspecies for each cell type using edgeR (Robinson et al., 2010), as described in (Chen et al., 2025). We then subset neurons, re-normalized, re-integrated, and re-clustered, and used a combination of the Azimuth mapping and manual marker assessment to assign clusters to the BICNN cell types. At this stage, we eliminated two small clusters that had both strong excitatory neuron and strong oligodendrocyte markers. We again assessed differential abundance using edgeR. For alignment to subclasses, we found that 78% of neurons had an alignment prediction score greater than 0.75. When we took the subset of only L5ET neurons, we again renormalized, reintegrated, and reclustered (resolution = 0.3), and we performed differential abundance analysis as before. For pseudo-bulk differential gene expression analysis, we combined RNAseq reads for all L5ET neurons (or all neurons) by subspecies and sample. We then used DESeq2 (Love et al., 2014) with default settings to calculate differential expression (DE). We eliminated genes that were DE in all neurons to focus specifically on genes DE in L5ET neurons. We compared those with marker genes for CSNs from the list of L5vsCSN genes in (Golan et al., 2023).

Reach-to-grasp behavior experiments and analysis

Freely moving reach-to-grasp experiments: Freely moving pellet reaching experiments were adapted from the protocol described in Xu et al., 2009. Briefly, food-restricted animals were placed in a clear acrylic box with a slit and presented with a pellet in a small divot on a rotating platform 1.25cm from the front of the box. In the shaping phase, the slit was in the middle of the box. The shaping phase lasted until animals reached for 30 pellets 70% with a single hand. Animals were then moved to the training phase, in which they were placed into a box with the slit on the side of their preferred hand from the shaping phase. They were then presented with 30 pellets in each session and scored on how many of the pellets they successfully grabbed and ate. Animals were trained for 6 days. Animals were videotaped with FLIR Blackfly3 cameras at 500fps. Data was scored manually. Analysis and statistics were done using custom R scripts.

Head-fixed reach-to-grasp experiments: Head-fixed pellet reaching experiments were adapted from previous established protocols in *Mus musculus* (Guo et al., 2021, 2015; Sauerbrei et al., 2020). Mice of either sex aged 16–50 weeks were fitted with head posts as described previously (Guo et al., 2015), acclimated to head restraint, and water restricted (0.04-0.05 ml/g body weight) to 75-80% of their original body weight to encourage task participation. In each session, mice first learned to retrieve food pellets resting on a rotating table (Guo et al., 2015; Sauerbrei et al., 2020), followed by a moveable vertical pellet dispenser that presents one pellet per trial for a limited period of time (3-5s; Sauerbrei et al., 2021). Each trial is video recorded using two high-speed cameras (FLIR, Flea 3, 500 fps) with manual iris and focus lenses, (Tamron 13VM1040ASIR 10–40 mm, f/1.4) one placed in front and one to the right side of the animal, illuminated with custom IR LED light sources, and calibrated for 3D tracking using DeepLabCut version 3.0.0 (Mathis et al., 2018; Nath et al., 2019) with default settings. We tracked forelimb and digit movements first using a 2D DeepLabCut model trained and run using default settings, and then filtered using the ARIMA filter with ARdegree and MAdegree set to 1 and p_bound set to 0.2. Triangulation was then run on the filtered data using default settings to create the 3D output. The cameras, acoustic cue, and pellet dispenser were controlled using Wavesurfer software (Adam Taylor, Janelia Scientific Computing; http://wavesurfer.janelia.org/) and a custom Arduino controller (Peter Polidoro, Janelia Experimental Technology).

Arm measurements: Arm measurements were taken by an experimenter blinded to the animal subspecies from video frames where the arm and fingers were fully extended. Lines were drawn from just before the shoulder (judging where the video showed movement), roughly below the posterior edge of the eye, to the tip of the longest finger, tracing the length of the arm. This was used to find the length of the arm in pixels. Pixels were converted into mm using a ruler roughly in the same plane of the frame as the arm (attached to the pellet platform). A similar procedure was used to measure the distance between the animals' noses and the pellet position. This information was used to calculate the adjusted arm length, which was (distance from nose) - (arm length - minimum arm length of all animals), which essentially treats the distance for animals with longer arms as shorter by a value proportional to the length of their arms. For x-ray measurements, we used the length of the humerus + ulna + metacarpals + proximal and distal phalanges to get the total hand and arm length. We note that in our video data, we did not measure from the shoulder socket, as we did in x-ray data, and that x-ray data also is an overestimation of arm length due to overlap between the individual bones.

3D video analysis: Success or failure was called manually. Similarly, grasp time was determined as the time when the animal first started to move the pellet, as notated by manual observation. Hand angle was calculated from the DeepLabCut tracking data, taking a line between digit two and digit four and computing the angle between that line and the pellet platform. To compare similar parts of the grasp between animals and reduce noise in the angle measurement, the hand angle used in analysis was the angle at the time within 50ms before the grasp, where digits 2 and 4 were most spread apart (though we observed similar differences when using the grasp time exactly). To determine the frequency of each qualitative grasp type (top, side, finger, pincer), an observer blinded to the identity of the mouse classified each grasp (all grasps were shown in a scrambled order). To calculate hand position at the time of grasps, the 3D position from DeepLabCut tracking was used. The median success position

used to calculate the median absolute difference (MAD) was the median x, median y, and median z from all successes for each animal for each behavioral session. Sessions with fewer than five successful reaches were discarded. To calculate pellet position, we tracked the pellet in DeepLabCut and took the median 3D pellet position before the animal reached for each day. To calculate the lift time (i.e., the start of the behavioral sequence), we first determined the times when the position of the wrist peaked in the y plane (up-down) and then defined the lift as the time before the peak where the acceleration started to increase. To calculate the fail position, we considered all positions between the lift time and 250ms after the lift and calculated the position at which the hand was closest to a position at which the animal successfully grasped the pellet in other trials. We again analyzed each behavioral session separately and eliminated sessions with fewer than five successful reaches. We further eliminated any reaches in which the hand was not within 4.2 cm from the pellet position, observing that those were mostly instances where the animal did not fully attempt to reach, as indicated by not fully raising the arm to the table. We determine this threshold value by taking the 99th percentile of the hand-pellet distances from successful reaches. The reach trajectory was calculated by taking the time between lift and grasp for each successful reach, interpolating the 3D location values so that each reach was on the same timeline, starting with lift and ending with reach. Then the Euclidean distance between each trajectory and the median trajectory for each behavioral session was calculated.

Statistical analysis: Data analysis was done using custom R and Python scripts. Rank-sum tests were performed using default settings of R's wilcox.test() function. Linear regression was performed using the lmerTest package, fitting a linear mixed effects model with random intercepts for each mouse, with the following tests:

(1) Hand angle: angle ~ subspecies*pellet_position + (1|mouse)
(2) Success position MAD: succ_position ~ subspecies*pellet_position*pellet_hand_distance + (1|mouse)
(3) Fail-success distance: fail_succ_dist ~ subspecies*pellet_position + (1|mouse)
(4) Distance from median trajectory: traj_dist ~ subspecies*pellet_position + (1|mouse)

Neural recording experiments and analysis

At least one day prior to recording, two craniotomies were made over M1 and M2 within the following coordinates: M1, A/P: 0.0-0.5 mm, M/L: 1.0-2.0 mm; M2, AP: 1.0-2.0 mm, M/L: 0.8-1.3 mm. Depth targeting layer V: ~0.4-1.2 mm from surface. Coordinates for recordings were determined based on the center of the CSN population in the retrograde tracing experiments. Electrophysiological recordings were performed similarly to prior work in *Mus musculus* (Guo et al., 2021; Sauerbrei et al., 2020), using one 4-shank 384-channel Neuropixel 2.0 probe (IMEC, https://www.neuropixels.org/; Steinmetz et al., 2021) per area (M1/M2). Data were spike-sorted using kilosort4 (Pachitariu et al., 2024) and manually refined using the phy interface. Multi-units were removed from further analysis. Recordings were synced to the cameras, enabling us to match tracked behavior (see above) with neural activity. Neural data were binned (1 ms) and smoothed using a 20-ms Gaussian kernel. Data were then normalized at the level of the unit using the following formula: (activity – minimum activity)/(maximum activity – minimum activity +2). This normalization was applied to ensure that each neuron contributed on a comparable scale, independent of absolute firing rate. We adopted min–max normalization for three

primary reasons. First, it standardized activity distributions across sessions. Second, it allowed fair comparison across subspecies by placing all neurons on the same bounded scale. Third, given our relatively low sample sizes, min–max scaling stabilized variance estimates, reduced the influence of outlier neurons, and ensured that population-level analyses reflected consistent task-related dynamics. We isolated modulated neurons as those that had a significant (FDR-adjusted $q < 0.05$) difference in spike events in a 500 ms window centered 750 ms before the beginning of behavior, compared to a 500 ms window centered 250 ms after the beginning of the behavior. We further analyzed only modulated neurons. We identified the start of behavior as the lift time, determined from behavior tracking data. We first determined the times when the position of the wrist peaked in the y plane (up-down) and then defined 'lift' as the time before the peak where acceleration started to increase.

Climbing behavior experiments and analysis

Animals were water-restricted to motivate participation in the task. Water was removed from home cages 1 day before the task, and animals were given at least 40 µL of water per gram of body weight per day (mostly during the task, but they were supplemented if they did not drink enough). They were placed in an arena with two platforms separated by a 30cm rod with a plastic net underneath to catch them if they fell. Rods were stainless steel rods from McMaster Carr of the following diameters: 0.5 inch (½", 12.7mm), 0.375 inch (⅜", 9.53mm), 0.25 inch (¼", 6.35mm), 0.15625 inch (5/32", 3.97mm), 3mm, 2mm. Animals were trained over a ~3-day period. First, they were trained to use the water-dispensing nose-poke on an enclosed platform. Next, they were exposed to the opportunity to cross to the other platform (forest and prairie mice) or the thickest rod (F2s) and were able to get water at either platform. Finally, they were exposed to a scenario, programmed using Bonsai (Lopes et al., 2015), in which they only received water when they successfully moved from one platform to another: water dispensers only turned on when the animal crossed an infrared beam at the entrance to each platform consecutively. Animals were given 1 day on each rod size. They failed if they fell off the rod more than 3 times in a 15-minute period or if they refused to cross the rod in that period. Animals who failed were given one more day to attempt that rod, but if they failed again, they were not given the next rod. We attempted to stop animals from crossing after 18 crosses per day (when they should have received their target water allocation for the day), but some animals performed slightly more crosses due to difficulty in removing them from the behavioral assay. Animals were removed from the behavior setup after 15 minutes, regardless of the number of crosses they achieved. Animals were videotaped using three synced Blackfly3 cameras recording at 120 fps: one from the top and two from the sides, angled up from the bottom. Video collection was done using campy (https://github.com/ksseverson57/campy). A side camera view was used for all analyses in this study. Tests of climbing behavior were performed during the animal's dark cycle, illuminated by red and infrared lights.

Videos of climbing behavior were cropped to only the time when the animal is actually on the rod using either DeepEthogram (Bohnslav et al., 2021) or DeepLabCut (Mathis et al., 2018). Data was tracked in two dimensions using a DeepLabCut. All downstream analysis was performed using data that had a >0.65 likelihood of detection of the back, as reported by DeepLabCut. Data was smoothed using a 0.02-second Gaussian kernel. Time upright was calculated by calculating the difference between the minimum and maximum positions of the back. If it was greater than 100 pixels, the animal was

determined to have flipped upside down, and the midpoint between the minimum and maximum value was considered to be the cutoff for calling upright or upside down. Most animals spent very little time near that midpoint value. Whether or not a mouse was moving was determined by a velocity threshold of greater than 5 pixels/frame.

Generation and analysis of F2 hybrids

Four pairs of forest and prairie cross-subspecies mating pairs were set up in both directions (i.e., with the female being forest or prairie). The progeny of those crosses (F1s) were then crossed to animals that were not their siblings to generate F2 animals. Animals were run on the behavior assay. To facilitate performing the experiment on all animals, we ran them on three of the rod sizes: 12.7mm, 6.35mm, and 3mm. A subset of animals (5) was not run on the 6.35mm rod. Video data was missing for one day for two animals due to errors in video recording. Weights were measured at the time of the behavior. Foot, tail, and body length were measured at the time of sacrifice. To choose animals to use for the CST size analysis, we set thresholds for the fraction of time they spent upright and their velocities, which are not correlated with one another. We determined “good” climbers to be in the better half of all mice on both metrics and “bad” climbers to be in the worse half. Eleven mice were classified as “good” climbers, and eleven as “bad” climbers, and we picked randomly among them. We then watched videos to qualitatively confirm good and bad climbers. These animals were transcardially perfused, and spinal cords were stained as described above.

Acknowledgements

K.M.T. was funded by a Life Sciences Research Foundation Postdoctoral Fellowship sponsored by the Howard Hughes Medical Institute (HHMI) and a NIH K99 BRAIN Initiative Advanced Postdoctoral Career Transition Award (K99NS133031). H.E.H.’s lab was supported by HHMI. K.E.C. and D.C.E. were supported by the Harvard College Research Program and the Program for Research in Science and Engineering. Funding was provided by the A.W.H. laboratory startup funds. The Harvard Center for Biological Imaging assisted with light microscopy. The Harvard Center for Brain Sciences Neurotechnology core assisted with behavioral assay set-up and engineering. The Harvard Medical School Electron Microscopy Core assisted with electron microscopy. The Bauer Core at Harvard performed sequencing. The Image Analysis Collaborative at Harvard Medical School provided advice on whole-brain retrograde tracing analysis. We thank Wei-Chung Lee for advice on electron microscopy protocols and Caroline Hu for early work on viral tracing in deer mice and advice on surgeries. We thank Chris Kirby and the Harvard Office of Animal Resources for animal husbandry. We thank Jenny Chen for advice on single cell data analysis and assistance with single cell library generation. We thank Kevin Cross and Stefen Lemke for advice on neural activity data analysis, Nick Tustison for advice on whole brain registration, and members of the Hoekstra lab for helpful discussion throughout the project. We thank Evan Kingsley, Andi Kautt, Nacho Sanguinetti, and Jenny Chen for feedback on the manuscript.

Author Contributions

K.M.T.: Conceptualization, Data Curation, Formal Analysis, Funding Acquisition, Investigation, Methodology, Project Administration, Software, Supervision, Visualization, Writing - Original Draft Preparation, Writing - Review & Editing; J.D.C.: Data Curation, Investigation, Methodology, Writing -

Review & Editing; J-Z.G.: Data Curation, Investigation, Methodology; P.R.R.: Data Curation, Investigation, Methodology; K.E.C.: Data Curation, Investigation, Methodology; I.H.S: Data Curation, Investigation; D.C.E.: Data Curation, Investigation; A.W.H.: Conceptualization, Funding Acquisition, Supervision, Writing - Review & Editing; H.E.H.: Conceptualization, Funding Acquisition, Supervision, Writing - Review & Editing

References

Alstermark, B., Isa, T., Lundberg, A., Pettersson, L.-G., Tantisira, B., 1989. The effect of low pyramidal lesions on forelimb movements in the cat. Neuroscience Research 7, 71–75. https://doi.org/10.1016/0168-0102(89)90038-2

Arendt, D., Bertucci, P.Y., Achim, K., Musser, J.M., 2019. Evolution of neuronal types and families. Current Opinion in Neurobiology, Neuronal Identity 56, 144–152. https://doi.org/10.1016/j.conb.2019.01.022

Auer, T.O., Khallaf, M.A., Silbering, A.F., Zappia, G., Ellis, K., Álvarez-Ocaña, R., Arguello, J.R., Hansson, B.S., Jefferis, G.S.X.E., Caron, S.J.C., Knaden, M., Benton, R., 2020. Olfactory receptor and circuit evolution promote host specialization. Nature 579, 402–408. https://doi.org/10.1038/s41586-020-2073-7

Avants, B.B., Tustison, N., Johnson, H., 2009. Advanced Normalization Tools (ANTS). Insight j 2, 1–35.

Baier, F., Reinhard, K., Nuttin, B., Sans-Dublanc, A., Liu, C., Tong, V., Murmann, J.S., Wierda, K., Farrow, K., Hoekstra, H.E., 2025. The neural basis of species-specific defensive behaviour in Peromyscus mice. Nature 645, 439–447. https://doi.org/10.1038/s41586-025-09241-2

Baldwin, M.W., Ko, M.-C., 2020. Functional evolution of vertebrate sensory receptors. Hormones and Behavior 124, 104771. https://doi.org/10.1016/j.yhbeh.2020.104771

Bedford, N.L., Hoekstra, H.E., 2015. Peromyscus mice as a model for studying natural variation. eLife 4, e06813. https://doi.org/10.7554/eLife.06813

Beier, K.T., 2019. Hitchhiking on the neuronal highway: Mechanisms of transsynaptic specificity. Journal of Chemical Neuroanatomy 99, 9–17. https://doi.org/10.1016/j.jchemneu.2019.05.001

Bendesky, A., Kwon, Y.-M., Lassance, J.-M., Lewarch, C.L., Yao, S., Peterson, B.K., He, M.X., Dulac, C., Hoekstra, H.E., 2017. The genetic basis of parental care evolution in monogamous mice. Nature 544, 434–439. https://doi.org/10.1038/nature22074

Berg, S., Kutra, D., Kroeger, T., Straehle, C.N., Kausler, B.X., Haubold, C., Schiegg, M., Ales, J., Beier, T., Rudy, M., Eren, K., Cervantes, J.I., Xu, B., Beuttenmueller, F., Wolny, A., Zhang, C., Koethe, U., Hamprecht, F.A., Kreshuk, A., 2019. ilastik: interactive machine learning for (bio)image analysis. Nat Methods 16, 1226–1232. https://doi.org/10.1038/s41592-019-0582-9

Bernstein, N.A., 1996. Dexterity and its development, Dexterity and its development. Lawrence Erlbaum Associates, Inc, Hillsdale, NJ, US.

Berti, L., Giorgi, F., Kasneci, G., 2025. Emergent abilities in large language models: A survey. https://doi.org/10.48550/arXiv.2503.05788

Blair, W.F., 1950. Ecological factors in speciation of Peromyscus. Evolution 4, 253–275. https://doi.org/10.1111/j.1558-5646.1950.tb01396.x

Bohnslav, J.P., Wimalasena, N.K., Clausing, K.J., Dai, Y.Y., Yarmolinsky, D.A., Cruz, T., Kashlan, A.D., Chiappe, M.E., Orefice, L.L., Woolf, C.J., Harvey, C.D., 2021. DeepEthogram, a machine learning pipeline for supervised behavior classification from raw pixels. eLife 10, e63377. https://doi.org/10.7554/eLife.63377

Bradley, R.D., Francis, J.Q., Platt, R.N., Soniat, T.J., Alvarez, D., Lindsey, L.L., 2019. Mitochondrial DNA sequence data indicate evidence for multiple species within Peromyscus maniculatus. Museum of Texas Tech University.

Callaway, E.M., et al., 2021. A multimodal cell census and atlas of the mammalian primary motor cortex. Nature 598, 86–102. https://doi.org/10.1038/s41586-021-03950-0

Carmona, L.M., Thomas, E.D., Smith, K., Tasic, B., Costa, R.M., Nelson, A., 2024. Topographical and cell type-specific connectivity of rostral and caudal forelimb corticospinal neuron populations. Cell Reports 43. https://doi.org/10.1016/j.celrep.2024.113993

Cederquist, G.Y., Azim, E., Shnider, S.J., Padmanabhan, H., Macklis, J.D., 2013. Lmo4 establishes rostral motor cortex projection neuron subtype diversity. J. Neurosci. 33, 6321–6332. https://doi.org/10.1523/JNEUROSCI.5140-12.2013

Chen, J., Richardson, P.R., Kirby, C., Eddy, S.R., Hoekstra, H.E., 2025. Cellular evolution of the hypothalamic preoptic area of behaviorally divergent deer mice. eLife 13, RP103109. https://doi.org/10.7554/eLife.103109

Chen, S., Svoboda, K., 2020. "Uniclear" water-based brain clearing for light sheet imaging of electrode tracks.

Chi, J., Crane, A., Wu, Z., Cohen, P., 2018. Adipo-Clear: A Tissue Clearing Method for Three-Dimensional Imaging of Adipose Tissue. J Vis Exp 58271. https://doi.org/10.3791/58271

Chialvo, D.R., 2010. Emergent complex neural dynamics. Nature Phys 6, 744–750. https://doi.org/10.1038/nphys1803

Dice, L.R., 1944. Variation of the deer mouse (Peromyscus maniculatus) in parts of Idaho, Montana, Utah, and Wyoming. Contributions from the Laboratory of Vertebrate Biology 28, 1–13.

Ding, Y., Berrocal, A., Morita, T., Longden, K.D., Stern, D.L., 2016. Natural courtship song variation caused by an intronic retroelement in an ion channel gene. Nature 536, 329–332. https://doi.org/10.1038/nature19093

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., Gingeras, T.R., 2013. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21. https://doi.org/10.1093/bioinformatics/bts635

Economo, M.N., Viswanathan, S., Tasic, B., Bas, E., Winnubst, J., Menon, V., Graybuck, L.T., Nguyen, T.N., Smith, K.A., Yao, Z., Wang, L., Gerfen, C.R., Chandrashekar, J., Zeng, H., Looger, L.L., Svoboda, K., 2018. Distinct descending motor cortex pathways and their roles in movement. Nature 563, 79–84. https://doi.org/10.1038/s41586-018-0642-9

Elman, J.L., 1990. Finding structure in time. Cognitive Science 14, 179–211. https://doi.org/10.1207/s15516709cog1402_1

Fageiry, S., Warriner, C.L., Loper, J., Pianski, L., Reardon, T., Jessell, T.M., Costa, R.M., Miri, A., 2024. Distinct anatomical and functional corticospinal inputs innervate different spinal neuron types. https://doi.org/10.1101/2024.04.29.591683

Foster, D.D., 1959. Differences in behavior and temperament between two races of the deer mouse. Journal of Mammalogy 40, 496–513. https://doi.org/10.2307/1376266

Geschwind, D.H., Rakic, P., 2013. Cortical evolution: Judge the brain by its cover. Neuron 80, 633–647. https://doi.org/10.1016/j.neuron.2013.10.045

Golan, N., Ehrlich, D.B., Bonanno, J., O'Brien, R.F., Murillo, M., Kauer, S., Ravindra, N., Dijk, D. van, Cafferty, W.B., 2023. Anatomical diversity of the adult corticospinal tract revealed by single cell transcriptional profiling. J. Neurosci. https://doi.org/10.1523/JNEUROSCI.0811-22.2023

Grant, R.A., Goss, V.G.A., 2022. What can whiskers tell us about mammalian evolution, behaviour, and ecology? Mammal Review 52, 148–163. https://doi.org/10.1111/mam.12253

Graves, S., Maldonado, J., Wolff, J.O., 1988. Use of ground and arboreal microhabitats by *Peromyscus leucopus* and *Peromyscus maniculatus*. Can. J. Zool. 66, 277–278. https://doi.org/10.1139/z88-040

Greig, L.C., Woodworth, M.B., Galazo, M.J., Padmanabhan, H., Macklis, J.D., 2013. Molecular logic of neocortical projection neuron specification, development and diversity. Nat Rev Neurosci 14, 755–769. https://doi.org/10.1038/nrn3586

Grillner, S., Wallén, P., 1985. Central pattern generators for locomotion, with special reference to vertebrates. Annual Review of Neuroscience 8, 233–261. https://doi.org/10.1146/annurev.ne.08.030185.001313

Gu, Z., Kalambogias, J., Yoshioka, S., Han, W., Li, Zhuo, Kawasawa, Y.I., Pochareddy, S., Li, Zhen, Liu, F., Xu, X., Wijeratne, H.R.S., Ueno, M., Blatz, E., Salomone, J., Kumanogoh, A., Rasin, M.-R., Gebelein, B., Weirauch, M.T., Sestan, N., Martin, J.H., Yoshida, Y., 2017. Control of species-dependent cortico-motoneuronal connections underlying manual dexterity. Science 357, 400–404. https://doi.org/10.1126/science.aan3721

Gu, Z., Ueno, M., Klinefelter, K., Mamidi, M., Yagi, T., Yoshida, Y., 2019. Skilled movements in mice require inhibition of corticospinal axon collateral formation in the spinal cord by semaphorin signaling. J. Neurosci. 2832–18. https://doi.org/10.1523/JNEUROSCI.2832-18.2019

Guo, J.-Z., Graves, A.R., Guo, W.W., Zheng, J., Lee, A., Rodríguez-González, J., Li, N., Macklin, J.J., Phillips, J.W., Mensh, B.D., Branson, K., Hantman, A.W., 2015. Cortex commands the performance of skilled movement. eLife 4, e10774. https://doi.org/10.7554/eLife.10774

Guo, J.-Z., Sauerbrei, B.A., Cohen, J.D., Mischiati, M., Graves, A.R., Pisanello, F., Branson, K.M., Hantman, A.W., 2021. Disrupting cortico-cerebellar communication impairs dexterity. eLife 10, e65906. https://doi.org/10.7554/eLife.65906

Hager, Emily R., 2020. Morphological and behavioral evolution in forest deer mice. Harvard University.

Hager, E.R., Harringmeyer, O.S., Wooldridge, T.B., Theingi, S., Gable, J.T., McFadden, S., Neugeboren, B., Turner, K.M., Jensen, J.D., Hoekstra, H.E., 2022. A chromosomal inversion contributes to divergence in multiple traits between deer mouse ecotypes. Science 377, 399–405. https://doi.org/10.1126/science.abg0718

Hager, E.R., Hoekstra, H.E., 2021. Tail length evolution in deer mice: Linking morphology, behavior, and function. Integrative and Comparative Biology 61, 385–397. https://doi.org/10.1093/icb/icab030

Heffner, R., Masterton, B., 1975. Variation in form of the pyramidal tract and its relationship to digital dexterity. BBE 12, 161–174. https://doi.org/10.1159/000124401

Heffner, R.S., Masterton, R.B., 1983. The role of the corticospinal tract in the evolution of human digital dexterity. Brain Behav Evol 23, 165–183. https://doi.org/10.1159/000121494

Herculano-Houzel, S., 2011. Not all brains are made the same: new views on brain scaling in evolution. Brain Behav Evol 78, 22–36. https://doi.org/10.1159/000327318

Hibbard, C., 1968. Palentology, in: King, J. (Ed.), Biology of Peromyscus (Rodentia). American Society of Mammalogists, pp. 6–24.

Hopfield, J.J., 1982. Neural networks and physical systems with emergent collective computational abilities. Proceedings of the National Academy of Sciences 79, 2554–2558. https://doi.org/10.1073/pnas.79.8.2554

Horner, B. Elizabeth, 1954. Arboreal adaptations of Peromyscus with special reference to use of the tail. University of Michigan, Ann Arbor, Michigan.

Huilgol, D., Levine, J.M., Galbavy, W., Wang, B.-S., He, M., Suryanarayana, S.M., Huang, Z.J., 2023. Direct and indirect neurogenesis generate a mosaic of distinct glutamatergic projection neuron types in cerebral cortex. Neuron 0. https://doi.org/10.1016/j.neuron.2023.05.021

Isko, E.C., Harpole, C.E., Zheng, X.M., Zhan, H., Davis, M.B., Zador, A.M., Banerjee, A., 2024. Selective expansion of motor cortical projections in the evolution of vocal novelty. https://doi.org/10.1101/2024.09.13.612752

Itoh, Y., Sahni, V., Shnider, S.J., McKee, H., Macklis, J.D., 2023. Inter-axonal molecular crosstalk via Lumican proteoglycan sculpts murine cervical corticospinal innervation by distinct subpopulations. Cell Reports 42, 112182. https://doi.org/10.1016/j.celrep.2023.112182

Iwaniuk, A.N., Pellis, S.M., Whishaw, I.Q., 1999. Is digital dexterity really related to corticospinal projections?: a re-analysis of the Heffner and Masterton data set using modern comparative statistics. Behavioural Brain Research 101, 173–187. https://doi.org/10.1016/S0166-4328(98)00151-X

Iwaniuk, A.N., Whishaw, I.Q., 2000. On the origin of skilled forelimb movements. Trends in Neurosciences 23, 372–376. https://doi.org/10.1016/S0166-2236(00)01618-0

Jozefowicz, R., Vinyals, O., Schuster, M., Shazeer, N., Wu, Y., 2016. Exploring the limits of language modeling. https://doi.org/10.48550/arXiv.1602.02410

Kamiyama, T., Kameda, H., Murabe, N., Fukuda, S., Yoshioka, N., Mizukami, H., Ozawa, K., Sakurai, M., 2015. Corticospinal tract development and spinal cord innervation differ between cervical and lumbar targets. J. Neurosci. 35, 1181–1191. https://doi.org/10.1523/JNEUROSCI.2842-13.2015

Kamiyama, T., Yoshioka, N., Sakurai, M., 2006. Synapse elimination in the corticospinal projection during the early postnatal period. Journal of Neurophysiology 95, 2304–2313. https://doi.org/10.1152/jn.00295.2005

King, JA (Ed.), 1968. Biology of Peromyscus (Rodentia).

Kingsley, E.P., Hager, E.R., Lassance, J.-M., Turner, K.M., Harringmeyer, O.S., Kirby, C., Neugeboren, B.I., Hoekstra, H.E., 2024. Adaptive tail-length evolution in deer mice is associated with differential Hoxd13 expression in early development. Nat Ecol Evol 8, 791–805. https://doi.org/10.1038/s41559-024-02346-3

Kingsley, E.P., Kozak, K.M., Pfeifer, S.P., Yang, D.-S., Hoekstra, H.E., 2017. The ultimate and proximate mechanisms driving the evolution of long tails in forest deer mice: Mechanisms driving evolution of long-tailed deer mice. Evolution 71, 261–273. https://doi.org/10.1111/evo.13150

Korsunsky, I., Millard, N., Fan, J., Slowikowski, K., Zhang, F., Wei, K., Baglaenko, Y., Brenner, M., Loh, P., Raychaudhuri, S., 2019. Fast, sensitive and accurate integration of single-cell data with Harmony. Nat Methods 16, 1289–1296. https://doi.org/10.1038/s41592-019-0619-0

Kverková, K., Marhounová, L., Polonyiová, A., Kocourek, M., Zhang, Y., Olkowicz, S., Straková, B., Pavelková, Z., Vodička, R., Frynta, D., Němec, P., 2022. The evolution of brain neuron numbers in amniotes. Proceedings of the National Academy of Sciences 119, e2121624119. https://doi.org/10.1073/pnas.2121624119

Lassek, A.M., Rasmussen, G.L., 1940. A comparative fiber and numerical analysis of the pyramidal tract. Journal of Comparative Neurology 72, 417–428. https://doi.org/10.1002/cne.900720209

Lawrence, D.G., Kuypers, H.G.J.M., 1968. The functional organization of the motor system in the monkey. Brain 91, 1–13.

Lemon, R.N., 2008. Descending pathways in motor control. Annu. Rev. Neurosci. 31, 195–218. https://doi.org/10.1146/annurev.neuro.31.060407.125547

Levi, A.D., Schwab, J.M., 2021. A critical reappraisal of corticospinal tract somatotopy and its role in traumatic cervical spinal cord syndromes. Journal of Neurosurgery: Spine 36, 653–659. https://doi.org/10.3171/2021.7.SPINE21546

Liang, H., Paxinos, G., Watson, C., 2011. Projections from the brain to the spinal cord in the mouse. Brain Struct Funct 215, 159–186. https://doi.org/10.1007/s00429-010-0281-x

Lopes, G., Bonacchi, N., Frazão, J., Neto, J.P., Atallah, B.V., Soares, S., Moreira, L., Matias, S., Itskov, P.M., Correia, P.A., Medina, R.E., Calcaterra, L., Dreosti, E., Paton, J.J., Kampff, A.R., 2015. Bonsai: an event-based framework for processing and controlling data streams. Front Neuroinform 9, 7. https://doi.org/10.3389/fninf.2015.00007

Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology 15, 550. https://doi.org/10.1186/s13059-014-0550-8

Low, L.K., Liu, X.-B., Faulkner, R.L., Coble, J., Cheng, H.-J., 2008. Plexin signaling selectively regulates the stereotyped pruning of corticospinal axons from visual cortex. Proceedings of the National Academy of Sciences 105, 8136–8141. https://doi.org/10.1073/pnas.0803849105

Mathis, A., Mamidanna, P., Cury, K.M., Abe, T., Murthy, V.N., Mathis, M.W., Bethge, M., 2018. DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. Nat Neurosci 21, 1281–1289. https://doi.org/10.1038/s41593-018-0209-y

Matson, K.J.E., Sathyamurthy, A., Johnson, K.R., Kelly, M.C., Kelley, M.W., Levine, A.J., 2018. Isolation of adult spinal cord nuclei for massively parallel single-nucleus RNA sequencing. J Vis Exp 58413. https://doi.org/10.3791/58413

Metz, G.A., Whishaw, I.Q., 2002. Cortical and subcortical lesions impair skilled walking in the ladder rung walking test: a new task to evaluate fore- and hindlimb stepping, placing, and co-ordination. Journal of Neuroscience Methods 115, 169–179. https://doi.org/10.1016/S0165-0270(02)00012-2

Miki, A., 1996. Developmental expression of A, B and Y-subspecies of protein kinase C in the dorsal corticospinal tract in the rat spinal cord. Neuroscience 75, 939–948. https://doi.org/10.1016/0306-4522(96)00365-x

Mori, M., Kose, A., Tsujino, T., Tanaka, C., 1990. Immunocytochemical localization of protein kinase C subspecies in the rat spinal cord: Light and electron microscopic study. J. Comp. Neurol. 299, 167–177. https://doi.org/10.1002/cne.902990204

Nath, T., Mathis, A., Chen, A.C., Patel, A., Bethge, M., Mathis, M.W., 2019. Using DeepLabCut for 3D markerless pose estimation across species and behaviors. Nat Protoc 14, 2152–2176. https://doi.org/10.1038/s41596-019-0176-0

Nelson, A., Abdelmesih, B., Costa, R.M., 2021. Corticospinal populations broadcast complex motor signals to coordinated spinal and striatal circuits. Nat Neurosci 24, 1721–1732. https://doi.org/10.1038/s41593-021-00939-w

Niepoth, N., Bendesky, A., 2020. How natural genetic variation shapes behavior. Annual Review of Genomics and Human Genetics 21, 437–463. https://doi.org/10.1146/annurev-genom-111219-080427

Nudo, R.J., Sutherland, D.P., Masterton, R.B., 1995. Variation and evolution of mammalian corticospinal somata with special reference to primates. Journal of Comparative Neurology 358, 181–205. https://doi.org/10.1002/cne.903580203

Nyberg-Hansen, R., Rinvik, E., 1963. Some comments on the pyramidal tract, with special reference to its individual variations in man. Acta Neurologica Scandinavica 39, 1–30. https://doi.org/10.1111/j.1600-0404.1963.tb05384.x

Osgood, Wilfred H., 1909. Revision of the mice of the American genus Peromyscus. US Department of Agriculture.

Pachitariu, M., Sridhar, S., Pennington, J., Stringer, C., 2024. Spike sorting with Kilosort4. Nat Methods 21, 914–921. https://doi.org/10.1038/s41592-024-02232-7

Peromyscus maniculatus (North American deer mouse) Genome Assembly, n.d.

Prieto-Godino, L.L., Rytz, R., Cruchet, S., Bargeton, B., Abuin, L., Silbering, A.F., Ruta, V., Dal Peraro, M., Benton, R., 2017. Evolution of acid-sensing olfactory circuits in Drosophilids. Neuron 93, 661-676.e6. https://doi.org/10.1016/j.neuron.2016.12.024

Richner, M., Jager, S.B., Siupka, P., Vaegter, C.B., 2017. Hydraulic extrusion of the spinal cord and isolation of dorsal root ganglia in rodents. JoVE 55226. https://doi.org/10.3791/55226-v

Roberts, R.J.V., Pop, S., Prieto-Godino, L.L., 2022. Evolution of central neural circuits: state of the art and perspectives. Nat Rev Neurosci 1–19. https://doi.org/10.1038/s41583-022-00644-y

Robinson, M.D., McCarthy, D.J., Smyth, G.K., 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139–140. https://doi.org/10.1093/bioinformatics/btp616

Sahni, V., Itoh, Y., Shnider, S.J., Macklis, J.D., 2021a. Crim1 and Kelch-like 14 exert complementary dual-directional developmental control over segmentally specific corticospinal axon projection targeting. Cell Reports 37. https://doi.org/10.1016/j.celrep.2021.109842

Sahni, V., Shnider, S.J., Jabaudon, D., Song, J.H.T., Itoh, Y., Greig, L.C., Macklis, J.D., 2021b. Corticospinal neuron subpopulation-specific developmental genes prospectively indicate mature segmentally specific axon projection targeting. Cell Reports 37. https://doi.org/10.1016/j.celrep.2021.109843

Saiki-Ishikawa, A., Agrios, M., Savya, S., Forrest, A., Sroussi, H., Hsu, S., Basrai, D., Xu, F., Miri, A., 2025. Hierarchy between forelimb premotor and primary motor cortices and its manifestation in their firing patterns. eLife 13, RP103069. https://doi.org/10.7554/eLife.103069

Sauerbrei, B.A., Guo, J.-Z., Cohen, J.D., Mischiati, M., Guo, W., Kabra, M., Verma, N., Mensh, B., Branson, K., Hantman, A.W., 2020. Cortical pattern generation during dexterous movement is input-driven. Nature 577, 386–391. https://doi.org/10.1038/s41586-019-1869-9

Seeholzer, L.F., Seppo, M., Stern, D.L., Ruta, V., 2018. Evolution of a central neural circuit underlies Drosophila mate preferences. Nature 559, 564–569. https://doi.org/10.1038/s41586-018-0322-9

Serradj, N., Marino, F., Moreno-López, Y., Bernstein, A., Agger, S., Soliman, M., Sloan, A., Hollis, E., 2023. Task-specific modulation of corticospinal neuron activity during motor learning in mice. Nat Commun 14, 2708. https://doi.org/10.1038/s41467-023-38418-4

Shenoy, K.V., Sahani, M., Churchland, M.M., 2013. Cortical control of arm movements: A dynamical systems perspective. Annu. Rev. Neurosci. 36, 337–359. https://doi.org/10.1146/annurev-neuro-062111-150509

Song, J.H.T., Ruven, C., Patel, P., Ding, F., Macklis, J.D., Sahni, V., 2023. Cbln1 Directs Axon Targeting by Corticospinal Neurons Specifically toward Thoraco-Lumbar Spinal Cord. J. Neurosci. 43, 1871–1887. https://doi.org/10.1523/JNEUROSCI.0710-22.2023

Steinmetz, N.A., Aydin, C., Lebedeva, A., Okun, M., Pachitariu, M., Bauza, M., Beau, M., Bhagat, J., Böhm, C., Broux, M., Chen, S., Colonell, J., Gardner, R.J., Karsh, B., Kloosterman, F., Kostadinov, D., Mora-Lopez, C., O'Callaghan, J., Park, J., Putzeys, J., Sauerbrei, B., van Daal, R.J.J., Vollan, A.Z., Wang, S., Welkenhuysen, M., Ye, Z., Dudman, J.T., Dutta, B., Hantman, A.W., Harris, K.D., Lee, A.K., Moser, E.I., O'Keefe, J., Renart, A., Svoboda, K., Häusser, M., Haesler, S., Carandini, M., Harris, T.D., 2021. Neuropixels 2.0: A miniaturized high-density probe for stable, long-term brain recordings. Science 372, eabf4588. https://doi.org/10.1126/science.abf4588

Strick, P.L., Dum, R.P., Rathelot, J.-A., 2021. The cortical motor areas and the emergence of motor skills: A neuroanatomical perspective. Annu. Rev. Neurosci. 44, 425–447. https://doi.org/10.1146/annurev-neuro-070918-050216

Striedter, G.F., Northcutt, R.G., 2019. Brains Through Time: A Natural History of Vertebrates. Oxford University Press.

Stringer, C., Wang, T., Michaelos, M., Pachitariu, M., 2021. Cellpose: a generalist algorithm for cellular segmentation. Nat Methods 18, 100–106. https://doi.org/10.1038/s41592-020-01018-x

Sustaita, D., Pouydebat, E., Manzano, A., Abdala, V., Hertel, F., Herrel, A., 2013. Getting a grip on tetrapod grasping: form, function, and evolution. Biological Reviews 88, 380–405. https://doi.org/10.1111/brv.12010

Takagi, S., Sancer, G., Abuin, L., Stupski, S.D., Roman Arguello, J., Prieto-Godino, L.L., Stern, D.L., Cruchet, S., Álvarez-Ocaña, R., Wienecke, C.F.R., van Breugel, F., Jeanne, J.M., Auer, T.O., Benton, R., 2024. Olfactory sensory neuron population expansions influence projection neuron adaptation and enhance odour tracking. Nat Commun 15, 7041. https://doi.org/10.1038/s41467-024-50808-w

Takiyama, K., Okada, M., 2012. Maximization of learning speed in the motor cortex due to neuronal redundancy. PLoS Comput Biol 8, e1002348. https://doi.org/10.1371/journal.pcbi.1002348

Tamvacakis, A.N., Senatore, A., Katz, P.S., 2018. Single neuron serotonin receptor subtype gene expression correlates with behaviour within and across three molluscan species. Proceedings of the Royal Society B: Biological Sciences 285, 20180791. https://doi.org/10.1098/rspb.2018.0791

Tervo, D.G.R., Hwang, B.-Y., Viswanathan, S., Gaj, T., Lavzin, M., Ritola, K.D., Lindo, S., Michael, S., Kuleshova, E., Ojala, D., Huang, C.-C., Gerfen, C.R., Schiller, J., Dudman, J.T., Hantman, A.W., Looger, L.L., Schaffer, D.V., Karpova, A.Y., 2016. A designer AAV variant permits efficient retrograde access to projection neurons. Neuron 92, 372–382. https://doi.org/10.1016/j.neuron.2016.09.021

Thiebaut de Schotten, M., Forkel, S.J., 2022. The emergent properties of the connected brain. Science 378, 505–510. https://doi.org/10.1126/science.abq2591

Thompson, D.B., 1990. Different spatial scales of adaptation in the climbing behavior of Peromyscus maniculatus: geographic variation, natural selection, and gene flow. Evolution 44, 952–965. https://doi.org/10.2307/2409558

Ueno, M., Nakamura, Y., Li, J., Gu, Z., Niehaus, J., Maezawa, M., Crone, S.A., Goulding, M., Baccei, M.L., Yoshida, Y., 2018. Corticospinal circuits from the sensory and motor cortices differentially regulate skilled movements through distinct spinal interneurons. Cell Reports 23, 1286-1300.e7. https://doi.org/10.1016/j.celrep.2018.03.137

Wang, X., Liu, Y., Li, X., Zhang, Z., Yang, H., Zhang, Yu, Williams, P.R., Alwahab, N.S.A., Kapur, K., Yu, B., Zhang, Yiming, Chen, M., Ding, H., Gerfen, C.R., Wang, K.H., He, Z., 2017. Deconstruction of corticospinal circuits for goal-directed motor skills. Cell 171, 440-455.e14. https://doi.org/10.1016/j.cell.2017.08.014

Wang, Z., Maunze, B., Wang, Y., Tsoulfas, P., Blackmore, M.G., 2018. Global connectivity and function of descending spinal Input revealed by 3D microscopy and retrograde transduction. J. Neurosci. 38, 10566–10581. https://doi.org/10.1523/JNEUROSCI.1196-18.2018

Weimann, J.M., Zhang, Y.A., Levin, M.E., Devine, W.P., Brulet, P., McConnell, S.K., 1999. Cortical neurons require Otx1 for the refinement of exuberant axonal projections to subcortical targets. Neuron 24, 819–831. https://doi.org/10.1016/S0896-6273(00)81030-2

Welniarz, Q., Dusart, I., Roze, E., 2017. The corticospinal tract: Evolution, development, and human disorders. Developmental Neurobiology 77, 810–829. https://doi.org/10.1002/dneu.22455

Whishaw, I.Q., Gorny, B., Sarna, J., 1998. Paw and limb use in skilled and spontaneous reaching after pyramidal tract, red nucleus and combined lesions in the rat: behavioral and anatomical dissociations. Behavioural Brain Research 93, 167–183. https://doi.org/10.1016/S0166-4328(97)00152-6

Williams, R.W., Herrup, K., 1988. The control of neuron number. Annual Review of Neuroscience 11, 423–453.

Winter, C.C., Jacobi, A., Su, J., Chung, L., van Velthoven, C.T.J., Yao, Z., Lee, C., Zhang, Z., Yu, S., Gao, K., Duque Salazar, G., Kegeles, E., Zhang, Yu, Tomihiro, M.C., Zhang, Yiming, Yang, Z., Zhu, J., Tang, J., Song, X., Donahue, R.J., Wang, Q., McMillen, D., Kunst, M., Wang, N., Smith, K.A., Romero, G.E., Frank, M.M., Krol, A., Kawaguchi, R., Geschwind, D.H., Feng, G., Goodrich, L.V., Liu, Y., Tasic, B., Zeng, H., He, Z., 2023. A transcriptomic taxonomy of mouse brain-wide spinal projecting neurons. Nature 624, 403–414. https://doi.org/10.1038/s41586-023-06817-8

Xu, T., Yu, X., Perlik, A.J., Tobin, W.F., Zweig, J.A., Tennant, K., Jones, T., Zuo, Y., 2009. Rapid formation and selective stabilization of synapses for enduring motor memories. Nature 462, 915–919. https://doi.org/10.1038/nature08389

Ye, D., Walsh, J.T., Junker, I.P., Ding, Y., 2024. Changes in the cellular makeup of motor patterning circuits drive courtship song evolution in Drosophila. Current Biology 34, 2319-2329.e6. https://doi.org/10.1016/j.cub.2024.04.020

Young, J.W., 2023. Convergence of arboreal locomotor specialization: morphological and behavioral solutions for movement on narrow and compliant supports, in: Bels, V.L., Russell, A.P. (Eds.), Convergent Evolution, Fascinating Life Sciences. Springer International Publishing, Cham, pp. 289–322. https://doi.org/10.1007/978-3-031-11441-0_11

Zaimi, A., Wabartha, M., Herman, V., Antonsanti, P.-L., Perone, C.S., Cohen-Adad, J., 2018. AxonDeepSeg: automatic axon and myelin segmentation from microscopy data using convolutional neural networks. Sci Rep 8, 3816. https://doi.org/10.1038/s41598-018-22181-4

Zhu, L.L., Lindenberg, R., Alexander, M.P., Schlaug, G., 2010. Lesion load of the corticospinal tract predicts motor impairment in chronic stroke. Stroke 41, 910–915. https://doi.org/10.1161/STROKEAHA.109.577023