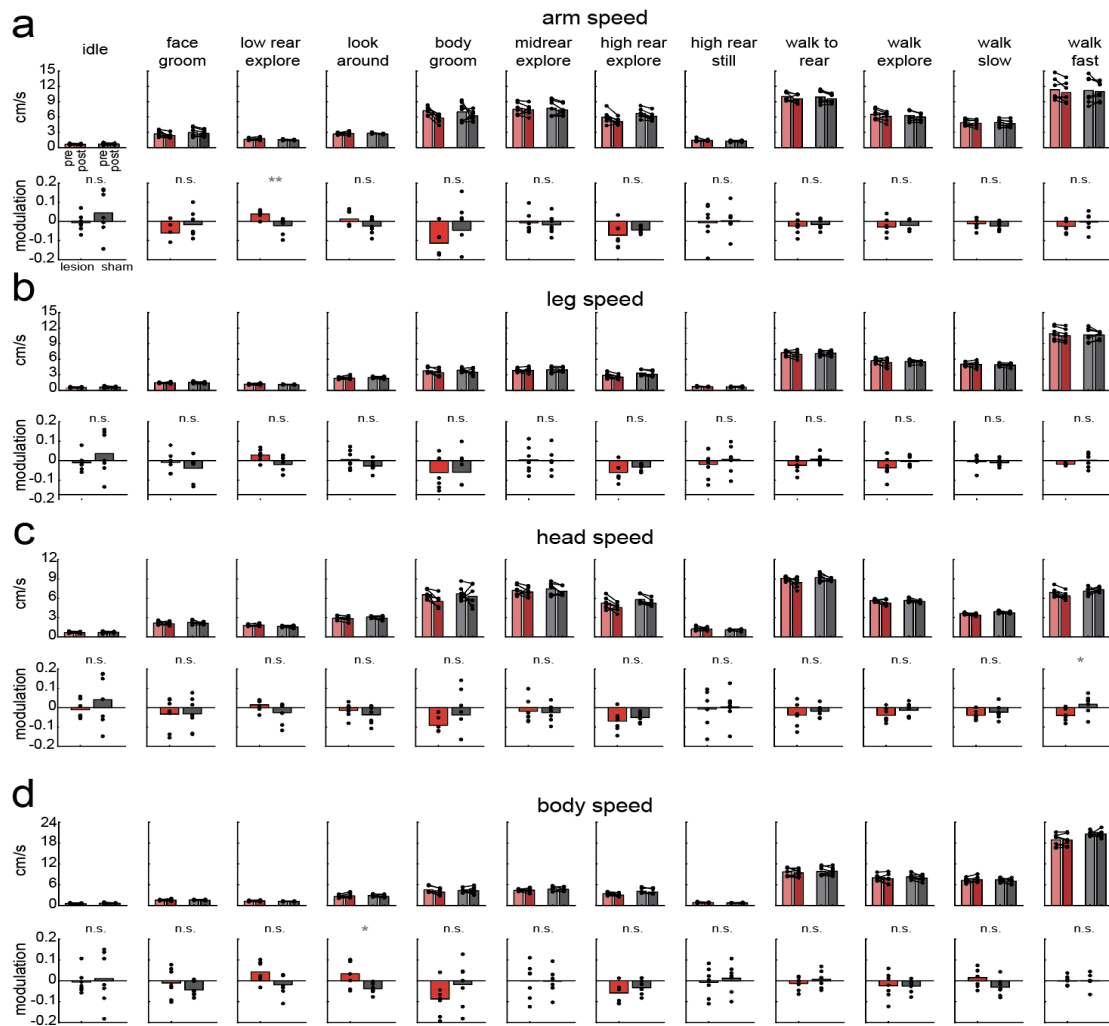
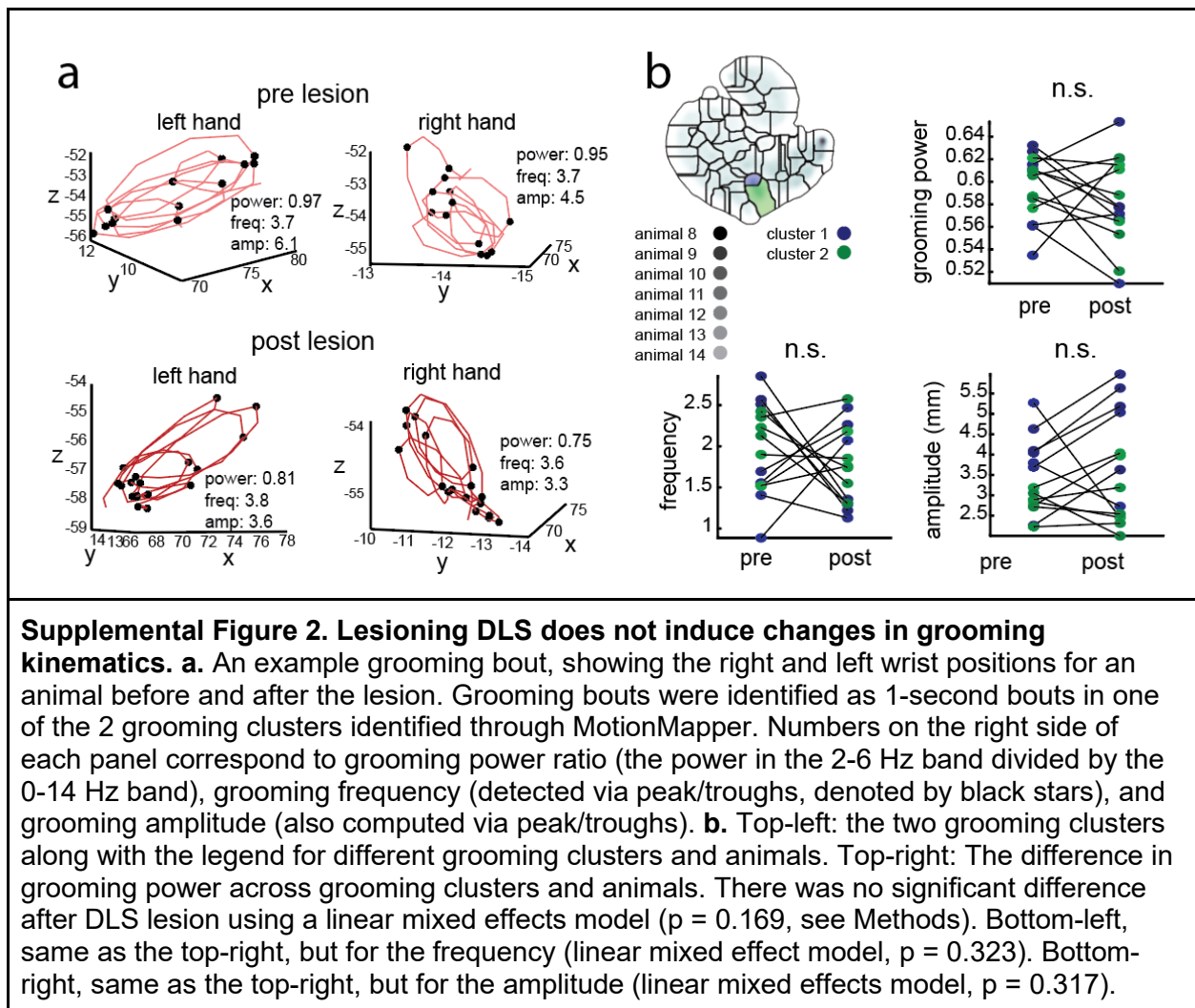


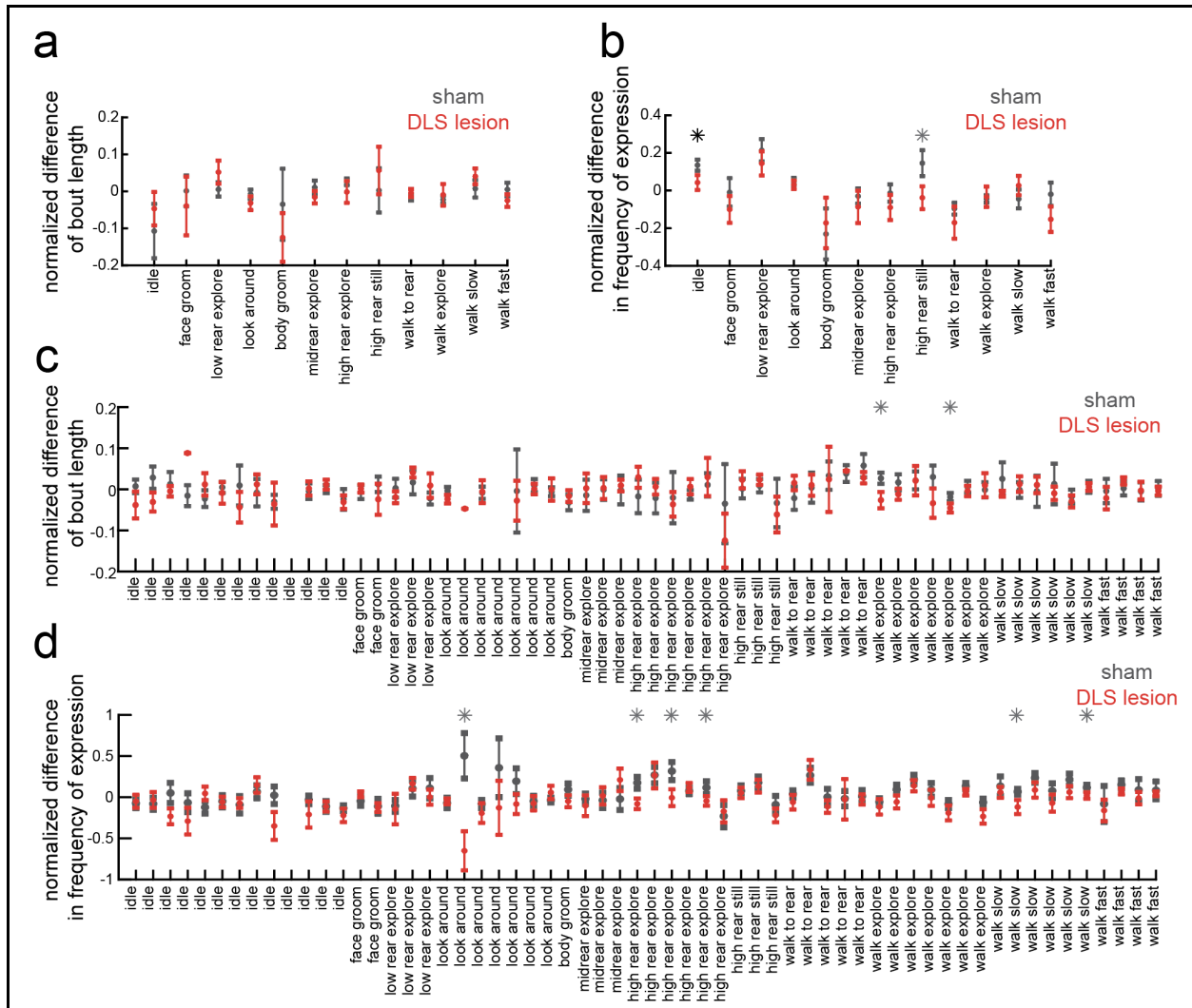
Differential kinematic coding in sensorimotor striatum across behavioral domains reflects different contributions to movement

In the format provided by the
authors and unedited

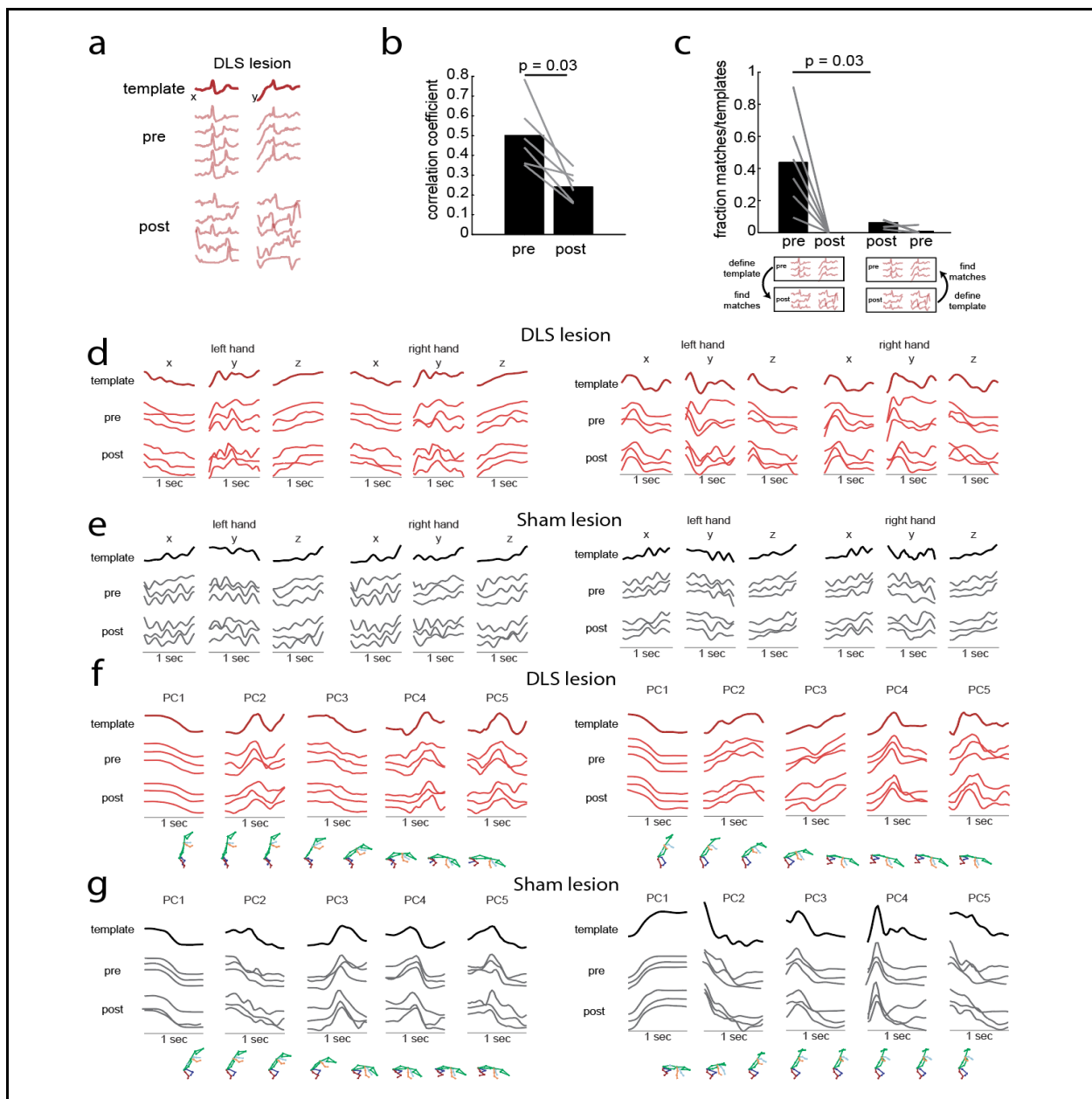


Supplemental Figure 1. Comparison of forelimb, hindlimb, head, and body speed across all coarsely-parsed behaviors. **a.** Top row: The mean speed (computed in 3D in cm/s) for both forelimbs across all 7 sham (gray bars) and DLS-lesioned animals (red bars) before (lighter bars on the left) and after (darker bars on the right) for each coarsely-parsed behavior. The behavior is listed above each plot. Each black dot corresponds to an animal, with lines connecting data from the same animal before/after surgery. Bottom row: The modulation $[(\text{post-pre})/(\text{post+pre})]$ for the arm speed. For all plots, a single star is given for $p < 0.05$, and two stars are given for $p < 0.01$. P-values are computed via a ranksum test across both cohorts ($n = 7$ in each cohort). Stars are gray if they do not reach significance according to the Benjamini-Hochberg procedure, where $Q = 0.1$. Only one test was significant with ranksum $p < 0.05$ (the change in speed for low-rear explore), but this was not significant after the B-H procedure. **b.** Same as **a**, but for the average speed of each hindlimb. There were no significant differences in changes in hindlimb speed across cohorts. **c.** Same as **a**, but for the speed of the head (averaged across both ears). Here there was significant p-value ($p < 0.05$), for walking fast, but this was not significant after accounting for multiple comparisons using the B-H procedure (as in **a**). **d.** Same as **a**, but for the trunk (the keypoint on the mid-back). There were no significant differences in changes in body speed across cohorts.

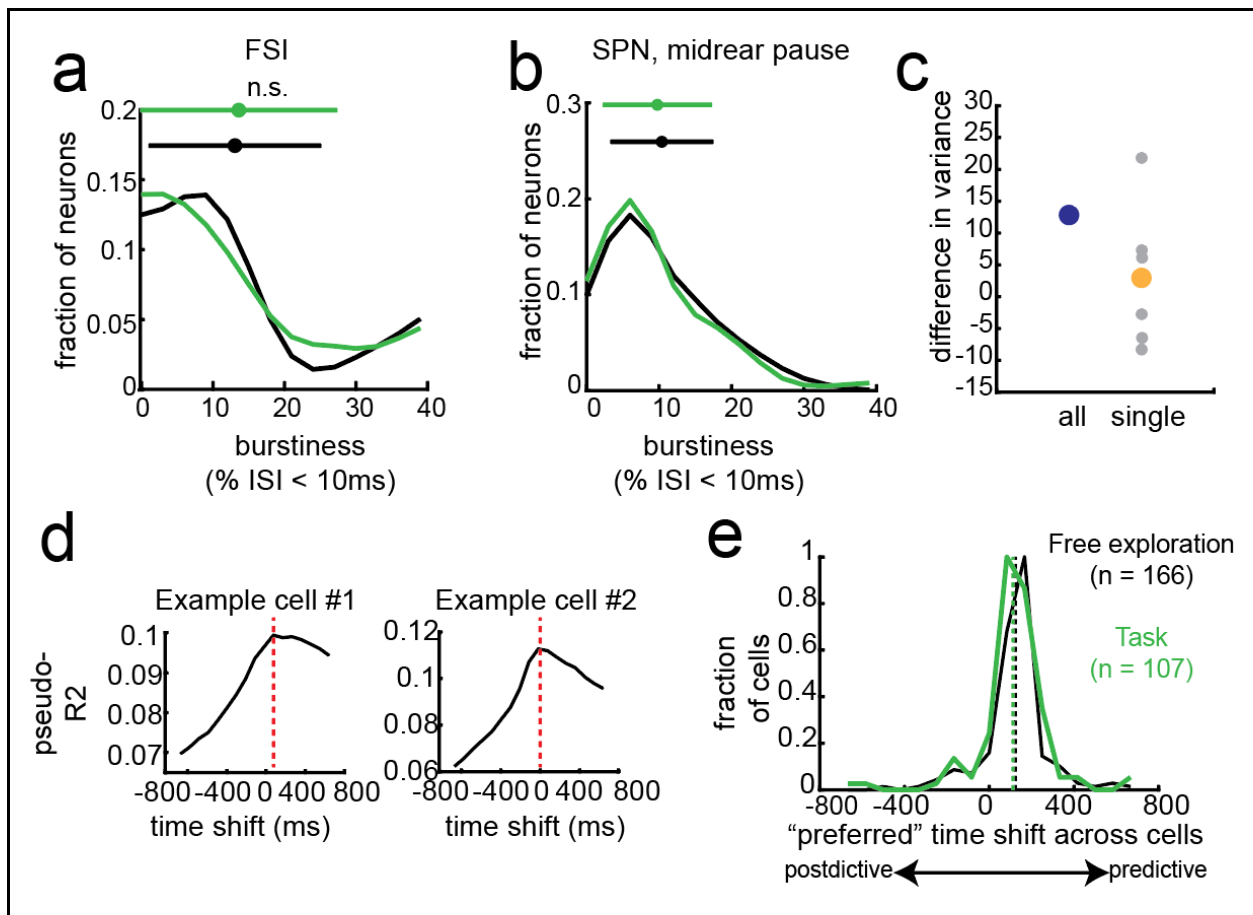




Supplemental Figure 3. Difference in the frequency of expression and bout length for coarse and fine behaviors after sham or DLS lesion. **a.** The normalized change in the bout length of behavioral expression for coarse behaviors before and after surgery for sham (black) and DLS-lesioned (red) animals. The normalized change is computed by quantifying $[(\text{post} - \text{pre})/(\text{post} + \text{pre})]$. Stars above the data points indicate $p < 0.05$ when computed through a ranksum test. Black stars indicate cases where the p-value was significant when accounting for multiple comparisons through the Benjamini-Hochberg procedure (using $Q = 0.1$). Here, there was no coarse behavior that had a significant difference. **b.** Same as **a**, but for the frequency of behavior expression. Here, there were two behaviors with ranksum $p < 0.05$, and one was significant after the B-H procedure (idle). However, note that the lesioned animals exhibit a smaller change after lesion (meaning their behavior is more similar to their behavior pre-surgery) compared to control animals. **c.** Similar to **a**, but for all finely-parsed behaviors. Here, there were no behaviors that had a significant difference. **d.** Similar to **b**, but for all finely parsed behaviors. Here, there were no behaviors that had a significant difference.

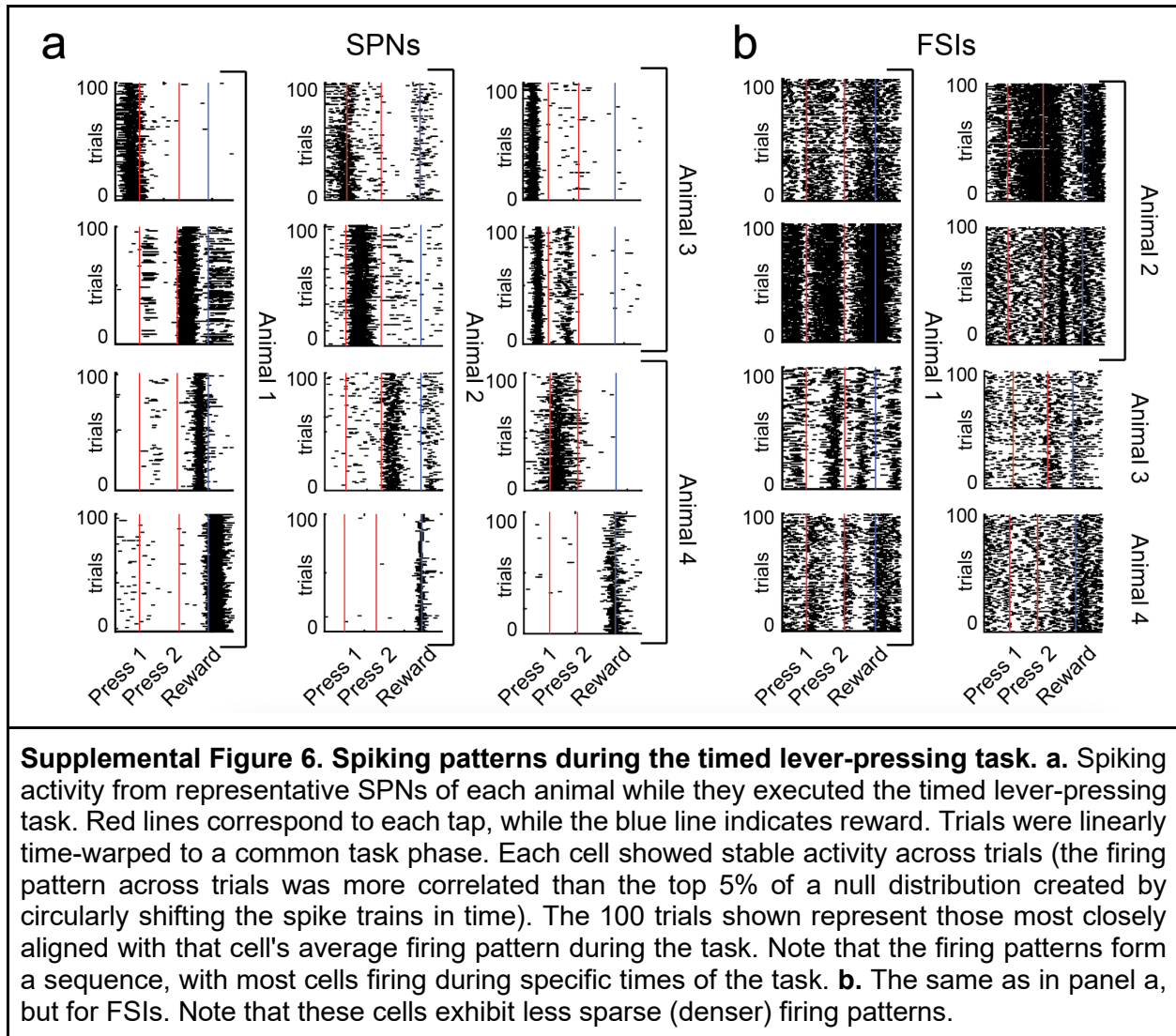


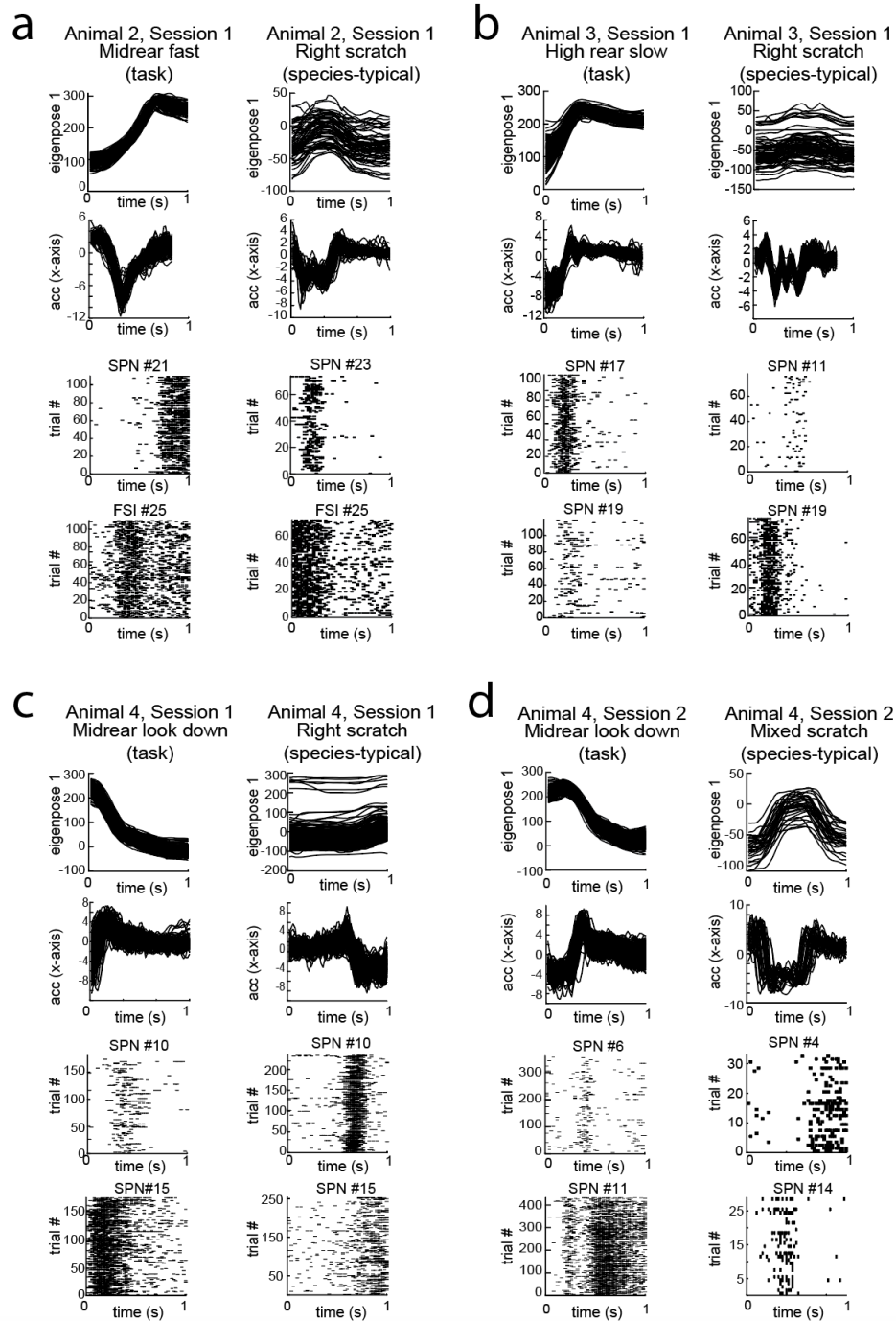
Supplemental Figure 4. Comparison of templates for task, grooming, and whole-body behaviors. **a.** Five randomly-chosen matches to a task-specific template before (top) and after (bottom) lesion. **b.** The mean correlation in hand position (averaged across the x- and y-value for both hands) was significantly less correlated across the 5000 trials before ('pre') and compared to after lesion ('post', signrank $p = 0.03$, $n = 6$ animals). **c.** The fraction of matches that match the template as defined before (left) or after (right) surgery. The left two bars correspond to data in which the template is defined from pre-lesion data (as in Figure 3), while on the right, the template is defined from post-lesion data. There are fewer matches to the template in the post-lesion data (signrank $p = 0.03$, $n = 6$ animals). **d.** The x, y, and z positions of both hands for two example grooming templates (dark red) and three example matches (lighter red) per template. **e.** Same as **d**, but for sham animals. Templates are in black; matches in gray. **f.** Same as **d**, but for the first five eigenposes (denoted by PCs) for whole-body templates. The skeletons indicate the template. **g.** Similar to **f**, but for sham animals.



Supplemental Figure 5. Firing statistics for FSIs and SPNs across behavioral domains.

a. The distribution of bursting during free exploration (black) or the task (green), with the mean $\pm$ standard deviation plotted above. The variance during the task was not significantly different from that during free exploration ($n = 33$ for both groups, F-test $F = 1.30$, $p = .462$). **b.** Similar to Fig 5e, but only considering a single behavior, midrear pause. There was no difference in the mean burstiness (signrank $p = 0.13$, $n = 146$ cells that were sufficiently active during this behavior in either domain) or the variance across bursting distributions ($p = 0.53$). **c.** The difference in the variances of the bursting distributions for all behaviors (blue) and for each rearing behavior (gray, with average in yellow). This indicates that when considering a more restricted set of behaviors, as we might expect to see in the task, the difference in variance decreases. **d.** The GLM fit (y-axis, quantified by the pseudo- R^2) for two example cells during species-typical behaviors expressed during free exploration. The spike train was shifted relative to the behavior according to the time-shift on the x-axis; positive numbers correspond to spiking being better correlated with future movements. Note that for both cells, spikes tend to be better aligned with movements in the future. The red dashed line indicates the time shift that returns the highest model fit. **e.** Comparison of the time shift for species-typical-active and task-specific active neurons. Here we only included neurons for which the GLM fits across time shifts reported a reliable peak at a single time value (specifically, that the curve was well-fit, with $RMSE > 0.1$, by a 2nd-order polynomial whose first coefficient was < -0.01 , indicating an upside-down U-shaped curve.) The dashed lines indicate the average time shift for cells in each domain; there is no significant difference in the time-shift (ranksum $p = 0.24$).





Supplemental Figure 7. Comparison of neural activity during movement templates identified during the task or during free exploration. a.-d. Neural activity from SPNs and FSIs during species-typical and task-derived templates for three different animals and sessions. The top two rows of each plot correspond to the eigenpose and accelerometer of the template matches. The bottom two rows correspond to the raster plots for each cell during template matches. Cell numbers (specific to each session) are listed above each subplot. Note in some cases, the same cells are plotted for task and species-typical templates.