Current Biology, Volume *31*

Supplemental Information

Reinforcement learning links spontaneous cortical dopamine impulses to reward

Conrad Foo, Adrian Lozada, Johnatan Aljadeff, Yulong Li, Jing W. Wang, Paul A. Slesinger, and David Kleinfeld

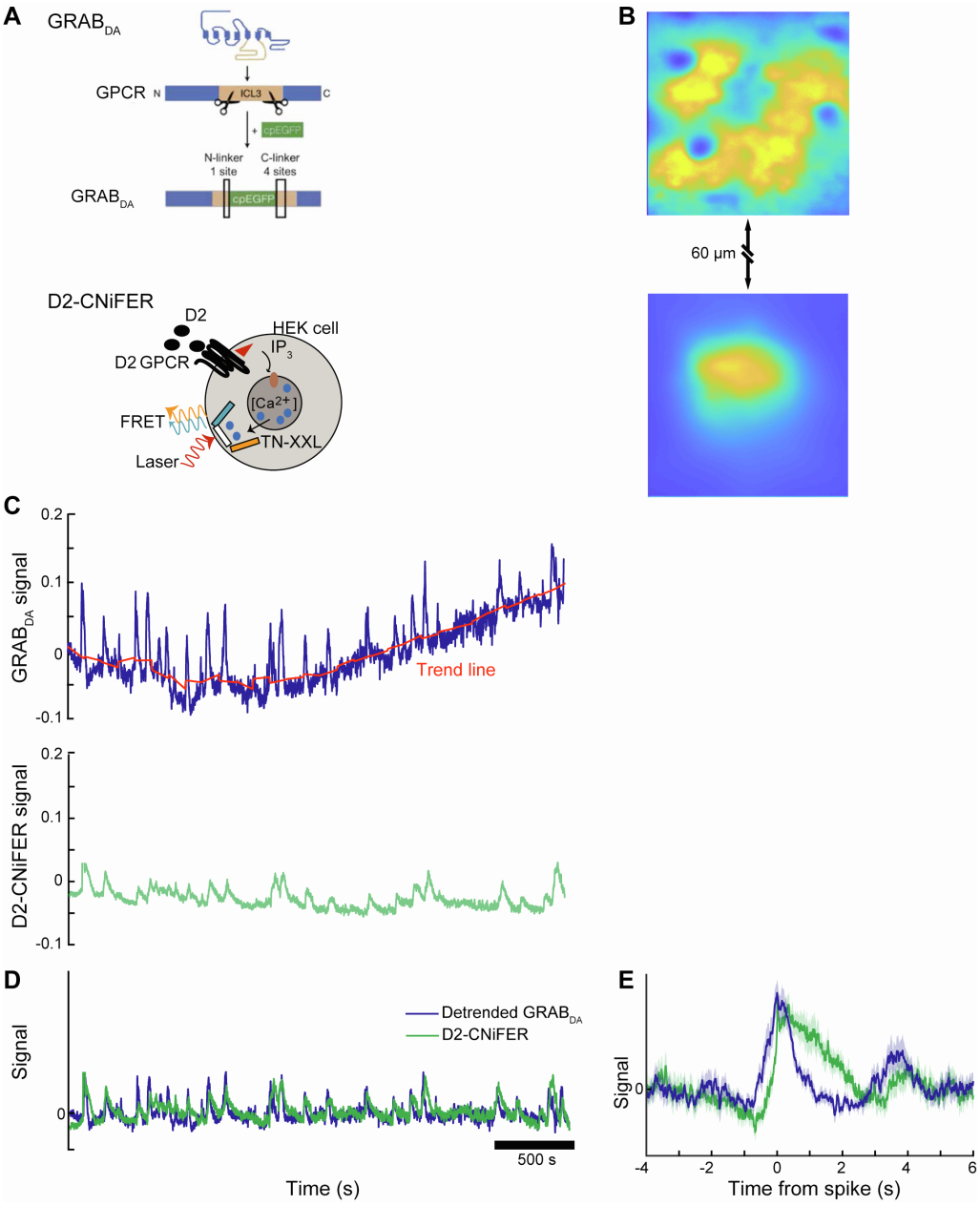

Figure S1. Comparison of signals from genetically expressed $GRAB_{DA}$ and implanted D2-CNiFER cells, related to Figure 1.

A. Schematic of design of the genetically expressed dopamine sensor, $GRAB_{DA}$, compared to D2-CNiFERs. $GRAB_{DA}$ is constructed by inserting the dopamine binding site on the D2-GPCR into cpGFP. Binding of DA causes conformational changes in the binding site that effect the efficiency of fluorescence.

B. Averaged image showing region of $GRAB_{DA}$ expression and a D2-CNiFER implant. $GRAB_{DA}$ expression was induced using a viral vector.

C. Simultaneous measurement of genetically expressed $GRAB_{DA}$ (blue, top) and implanted D2-CNiFER cells (green, middle). A small region of interest near the center of the region of $GRAB_{DA}$ expression was averaged and a fluorescence trace was calculated. The $GRAB_{DA}$ signal had significant drift in baseline on the scale of tens of minutes, but had better signal-to-noise and temporal resolution than the D2-CNiFER signal.

D. Comparison of normalized detrended $GRAB_{DA}$ signal and normalized D2-CNiFER signal. The $GRAB_{DA}$ signal did not exhibit the decay tail that the D2-CNiFER signal had, and was about twice as bright; small transients that were detected by $GRAB_{DA}$ were not always detected by the D2-CNiFERs. Transients occurring in quick succession as observed by $GRAB_{DA}$ appeared as a single, longer transient when observed by D2-CNiFERs.

E. Normalized average transient triggered response of $GRAB_{DA}$ (blue) and D2-CNiFER (green) signals. The $GRAB_{DA}$ signal both rose and decayed more rapidly than the D2-CNiFER signal, as one would expect; the change in fluorescence in the D2-CNiFER signal requires activation of a second messenger pathway that is not necessary in $GRAB_{DA}$.

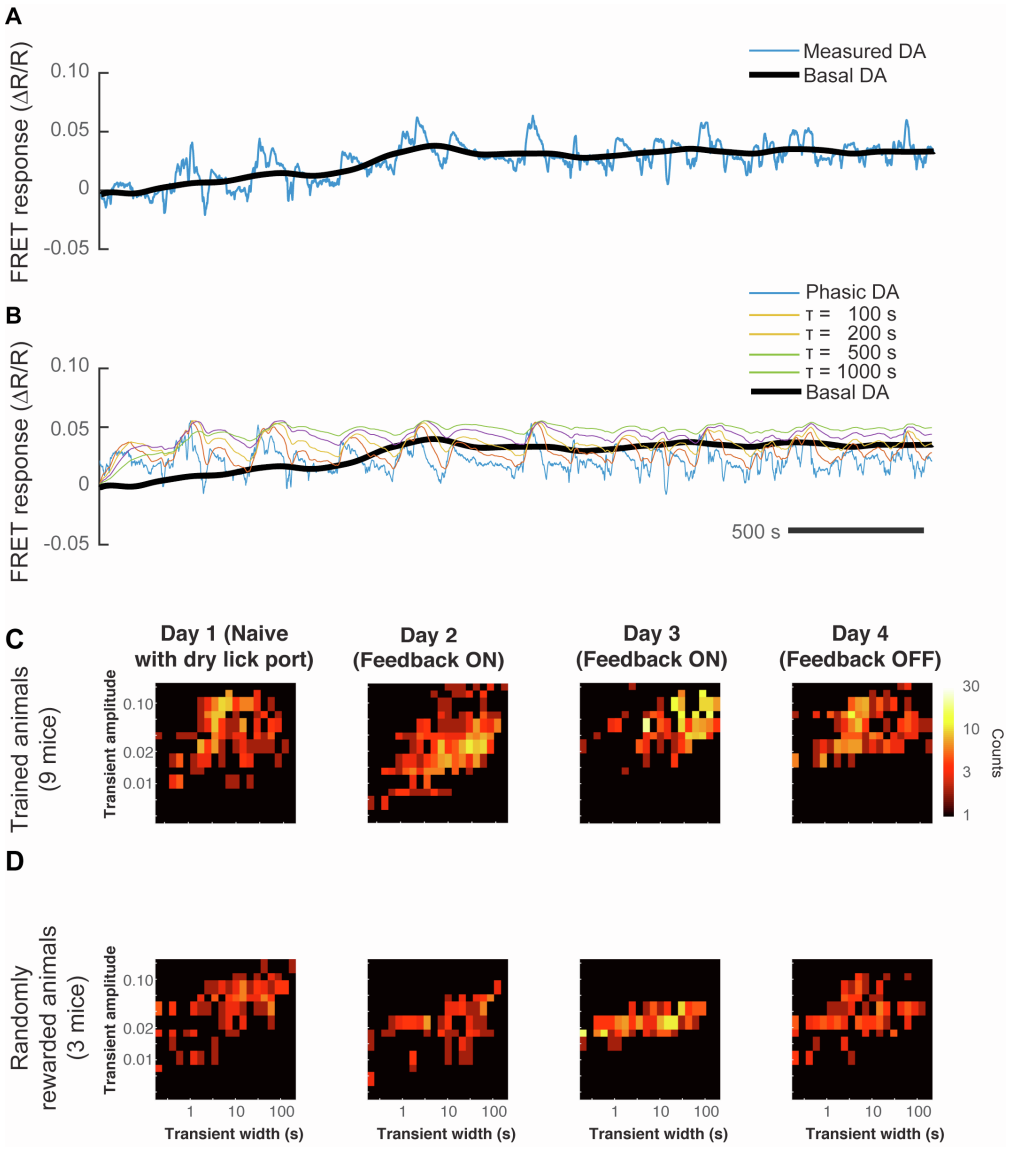

Figure S2. Additional analysis of dopamine levels, related to Figure 3.

A. Extraction of basal DA from the measured $[DA]_{ex}$. A LOESS fit (tricubic weighting function, linear fit) was applied to the measured signal (blue) to extract DA transients. The window size was 940 s and the step size was 11 s. The transients were subtracted from the total DA signal to get the basal DA (black).

B. Leaky integration of the phasic DA signal does not reproduce the ramping basal DA signal. The phasic DA signal (blue), was extracted from the measured $[DA]_{ex}$ using a LOESS fit. A leaky integrator, with exponential decay time τ, was applied to this signal (yellow, green) for values of τ = 100 s, 200 s, 500 s, and 1000 s. The leaky integrator fails to reproduce the shift in basal DA that we observe. Although the integrator with a half-decay time of 500 s shows a similar shift upwards in the DA response, it ramps up to this level much quicker, i.e., around 100 s, than the 1000 s that we observe (black line).

C. Two-dimensional histograms showing the change in dopamine transient properties over training. Transient amplitudes during Day 3 of feedback training were significantly larger than those in the naïve animal. Transient widths during both Days 2 and 3 of feedback training were significantly longer than those in the naïve animal. Transient properties when feedback was turned off on Day 4 did not significantly differ from the naïve animal. The average widths of transients were 15.1 ± 1.3 s, 25.4 ± 2.1 s, 43.1 ± 2.5 s, and 18.5 ± 1.7 s; the corresponding average amplitudes were 0.056 ± 0.002, 0.045 ± 0.003, 0.081 ± 0.004, and 0.056 ± 0.002 for Days 1, 2, 3, and 4, respectively.

D. Two-dimensional histograms showing dopamine transient properties when animals were randomly rewarded. Transient width was significantly shorter on Days 3 and 4 compared to Days 1 and 2. Transient amplitudes were lower when animals rewarded compared to when they were not. The average widths of transients were 22.4 ± 2.1 s, 24.6 ± 2.3 s, 13.1 ± 1.3 s, and 14.7 ± 1.6 s; the corresponding average amplitudes were 0.052 ± 0.003, 0.029 ± 0.002, 0.027 ± 0.0005, and 0.039 ± 0.002 for Days 1, 2, 3, and 4 respectively.

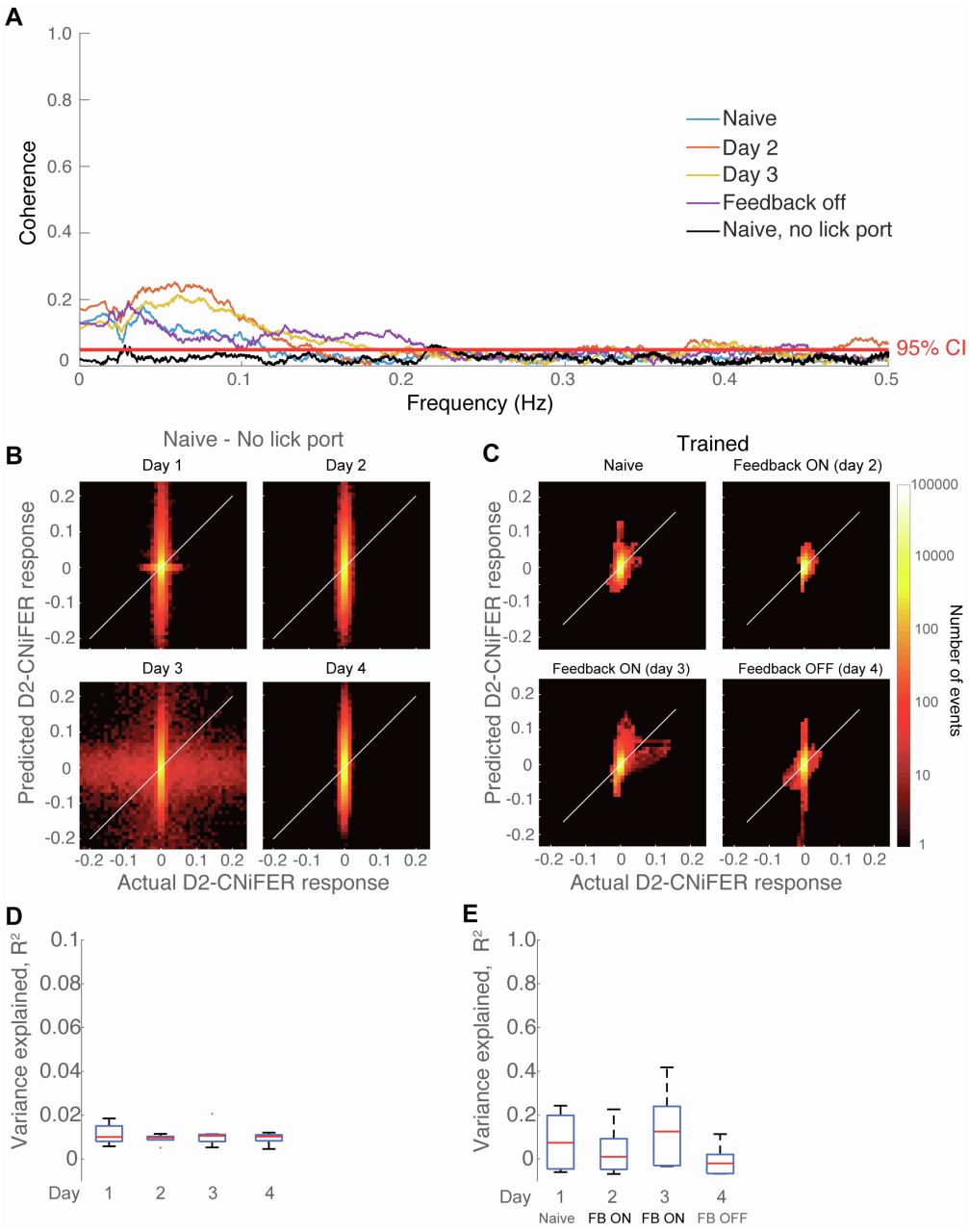

Figure S3. Introduction of a dry lick port introduces a small correlation between running and dopamine release, related to Figure 3.

A. Average spectral coherence between running and phasic dopamine release across animals (9 mice with lick port, 7 mice without lick port). In the presence of a lick port, coherence was significant at frequencies below 0.2 Hz. In the absence of a lick port, coherence was not significant. The coherence was calculated using the multi-taper method; the bandwidth was 0.02 Hz from averaging with 143 tapers.

B. Histograms of the predictions of a linear model of $[DA]_{ex}$ as a function of running speed versus the measured $[DA]_{ex}$ in the absence of a lick port during four consecutive days of experiments. White line shows the expected distribution of a perfectly predictive model. The model was fit to the data in the frequency domain, making use of the convolution theorem. Cross-spectral power was calculated with a multitaper estimate. A new model was fit for each trial; each histogram uses data from all trials within a given day of the experiment.

C. Same as panel B, but in the presence of a lick port. Animals were trained to increase $[DA]_{ex}$ for these data.

D. Variance explained by linear model of $[DA]_{ex}$ as a function of running speed in the absence of a lick port for different days of the experiment. This was calculated directly from the data shown in panel B. Each trial was a separate data point. R^2 was 0.011 ± 0.005, 0.009 ± 0.002, 0.010 ± 0.005, and 0.009 ± 0.003 for Days 1, 2, 3, and 4 respectively.

E. Same as panel D, but in the presence of a lick port. R^2 was 0.1 ± 0.1, 0.08 ± 0.1, 0.2 ± 0.2, and 0.04 ± 0.07 for naïve, Day 2 of training, Day 3 of training, and feedback OFF days respectively.