nature neuroscience

Supplementary information

https://doi.org/10.1038/s41593-025-01965-8

A common computational and neural anomaly across mouse models of autism

In the format provided by the authors and unedited

SUPPLEMENTARY MATERIALS FOR:

A common computational and neural anomaly across mouse models of autism

Jean-Paul Noel[1,2,3,4,*], Edoardo Balzani[5], Luigi Acerbi[6], Julius Benson[7], The International Brain Laboratory, Cristina Savin[7], & Dora E. Angelaki[7,8]

[1] Department of Neuroscience, University of Minnesota, Minnesota, USA
[2] Data Science Initiative, University of Minnesota, Minnesota, USA
[3] Minnesota Robotics Institute, College of Science and Engineering, University of Minnesota, Minnesota, USA
[4] Masonic Institute for the Developing Brain, University of Minnesota, Minnesota, USA
[5] Flatiron Institute, Simons Foundation, New York, USA
[6] Department of Computer Science, University of Helsinki, Helsinki, Finland
[7] Center for Neural Science, New York University, New York, USA
[8] Tandon School of Engineering, New York University, New York, USA

* Corresponding Author:
Noelx071@umn.edu (J.P. Noel)

Supplementary Text

Results: Blunted accumulation of sensory history in ASD mouse models

The “exp. bias model” may be more intuitively understood by example. Namely, if for sake of explanation we assume that the window over which animals count past sensory observations is a fixed uniform window of 10 trials (as opposed to an exponential of variable length) and over these trials, 8 were targets presented on the left, model 9 (“exp. no bias”) would suggest there is an 80% chance of the next target being presented on the left. Instead, in model 10 (“exp. bias”) the bins into which animals are counting “left” (or “right”) may be pre-filled – for example, with 5 counts each (i.e., α = 5 and β = 5). Thus, the model would suggest a 65% chance of the next target being presented on the left (8+5 vs. 2+5, or 13/20).

Results: Unit prior encoding shifts from visual and frontal cortices in ASD models

We briefly present the encoding of visual stimuli – as sanity check – before assessing the encoding of the subjective prior. In this regard, we find units that encode grating presentations throughout the brain (Steinmetz et al., 2019; IBL et al., 2023) with the primary visual cortex (VISp) showing the strongest difference (in mutual information) between contra-lateral responses (arguably driven by the stimulus itself) and ipsi-lateral responses (arguably further driven by recurrent neural dynamics; **Extended Data Fig. 8**). **Figure 5B** shows the PSTHs for a few example units, overlayed with the pGAM reconstruction and the estimated contribution of the visual stimulus kernel. We can observe that while for a few units (e.g., first two examples in **Fig. 5B**) the entirety of their response to grating presentation is (bottom-up) visually driven, for the majority of units (e.g., rest of first row and second row, **Fig. 5B**) their early response is sensory driven whereas their later latency responses are less so (putatively accounted for by the coupling filters). Some units showing increased firing after grating presentation were, in fact, according to the pGAM, not driven by visual stimulus at all (**Fig. 5B,** bottom row). Fittingly, these latter units showed a delayed evoked response relative to the “bottom-up” driven units. Overall, these results broadly align with classic characterizations of the visual pathways (Hubel, 1988; Felleman & Van Essen, 1991) and recent brain-wide surveys (Steinmetz et al., 2019; IBL et al., 2023), while also demonstrating the utility of including in the encoding model both externally-driven predictors and internal elements of neural dynamics.

Results: Lack of prediction errors in frontal cortex in ASD models

Most neurons recorded from across the brain (**Extended Data Fig. 10A**) and in frontal (**Extended Data Fig. 10B**) and visual (**Extended Data Fig. 10C**) cortices were from layers 2/3, 5, and 6 (no difference across genotypes, χ^2 test, all $p > 0.41$). The presence of sensory prediction errors in FC in wildtype animals, and the lack thereof in mouse models of ASD was true across L2/3, L5, and L6 layers (ANOVA interaction term, all $p < 0.001$; **Extended Data Fig. 10D**). The presence of prediction errors in visual cortices was driven by L2/3 across all genotypes, with L5/6 not demonstrating prediction errors (all $p > 0.71$; **Extended Data Fig. 10E**). Layers 1 and 4 did not have a sufficiently large number of units for this analysis.

Supplementary References

Hubel, D. H. (1988). Eye, Brain, and Vision. Scientific American Library.

International Brain Lab, Benson, B., Benson, J., Birman, D., Bonacchi, N., Carandini, M., Catarino, J.A., Chapuis, G.A., Churchland, A.K., Dan, Y., Dayan, P., DeWitt, E.E.J., Engel, T.A., Fabbri, M., Faulkner, M., Fiete, I.R., Findling, C., Freitas-Silva, L., Gercek, B., Harris, K.D., Hausser, M., Hofer, S.B., Hu, F., Hubert, F., Huntenburg, J.M., Khanal, A., Krasniak, C., Langdon, C., Lau, PYP., Mainen, Z.F., Meijer, G.T., Miska, N.J., Mrsic-Flogel, T.D., Noel, J.P., Nylund, K., Pan-Vazquez, A., Pouget, A., Rossant, C., Roth, N., Schaeffer, R., Schartner, M., Shi, Y., Socha, K.Z., Steinmetz, N.A., Svoboda, K., Urai, A.E., Wells, M.J., West, S.J., Whiteway, M.R., Winter, O., Witten, I.B. (2023). A brain-wide map of neural activity during complex behavior. bioRxiv, 2023.07. 04.547681

Steinmetz NA, Zatka-Haas P, Carandini M, Harris KD. Distributed coding of choice, action and engagement across the mouse brain. Nature. 2019 Dec;576(7786):266-273. doi: 10.1038/s41586-019-1787-x. Epub 2019 Nov 27. PMID: 31776518; PMCID: PMC6913580.

Figure S1. Model and parameter recovery.

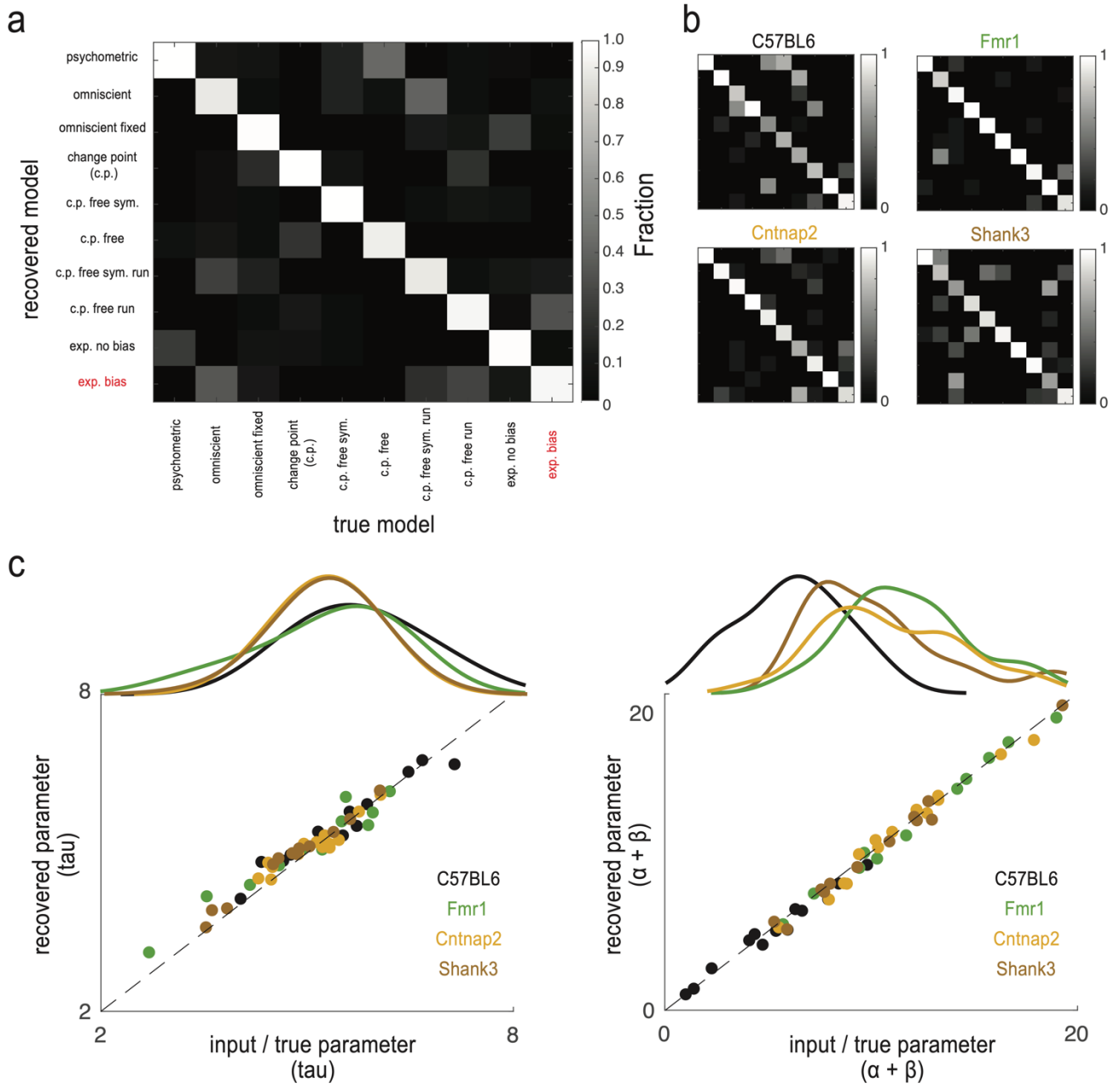

Figure S1. Model and parameter recovery. A. Model recovery across all animals and genotypes. We generated 750 artificial datasets by picking at random parameters for each of the 10 models presented in the main paper, and using the stimuli presented to the actual animals (n = 75, thus 10 * 75 = 750 datasets) to generate simulated choices. The parameters chosen were bounded by the minimum and maximum parameters observed in real data. Each of these datasets were then fit with the 10 models in the same manner we fit real datasets (20 optimizations with BADS[75] and cross-validating log likelihoods with odd vs. even trials; total of 150k fits). This resulted in a best-fit model for each artificial dataset, from which we created a confusion matrix. As it can be observed, the recovered model almost always matched the true model (92.6% of cases), with the "*exp. bias*" model (overall winning model) being seldomly confused with the "*c.p. free run*" model (6.1% of cases), and almost never with any other (see **Figure S6** and main text for further discussion). **B.** Same as **A,** but restricted to each genotype (including parameter ranges, number of sessions, trials, etc.). Same conclusion. **C.** Parameter recovery for the "*exp. bias*": x-axis shows the true parameters used (those estimated from the real dataset) and y-axis shows the recovered parameters in an independent run. Both for the tau parameter (dictating the space of the exponential weighting; left), and the sum of $\alpha + \beta$ of the Beta distribution (dictating the influence of the hyper-prior; right), true and recovered parameters span the diagonal, indicating proper parameter recovery. The scatter plots also show individual data and marginal distributions are shown on top of the scatter plots. Data shows a larger $\alpha + \beta$ in mouse models of ASD relative to the wildtype.

Figure S2. Quality of model fittings.

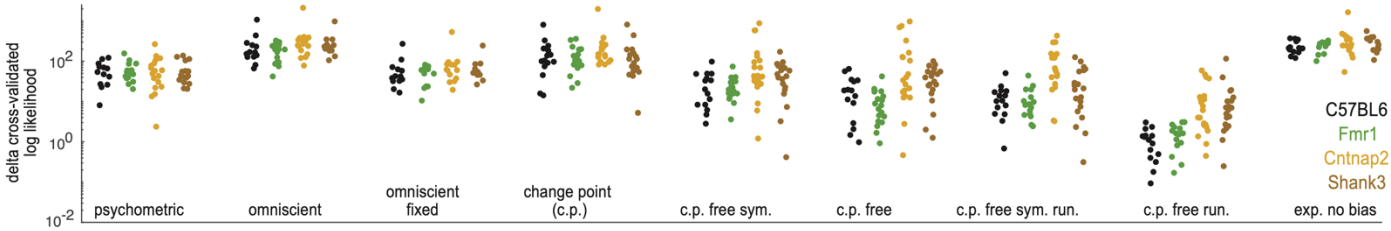

Figure S2. Quality of model fittings. Difference in cross-validated log likelihood to the best-fitting model ("exp bias", not shown here, as x = 0). All subjects are shown individually, and of note, the y-axis is on a log-scale. Data is the same as in **Figure 2A.**

Figure S3. Change-point detection model parameters.

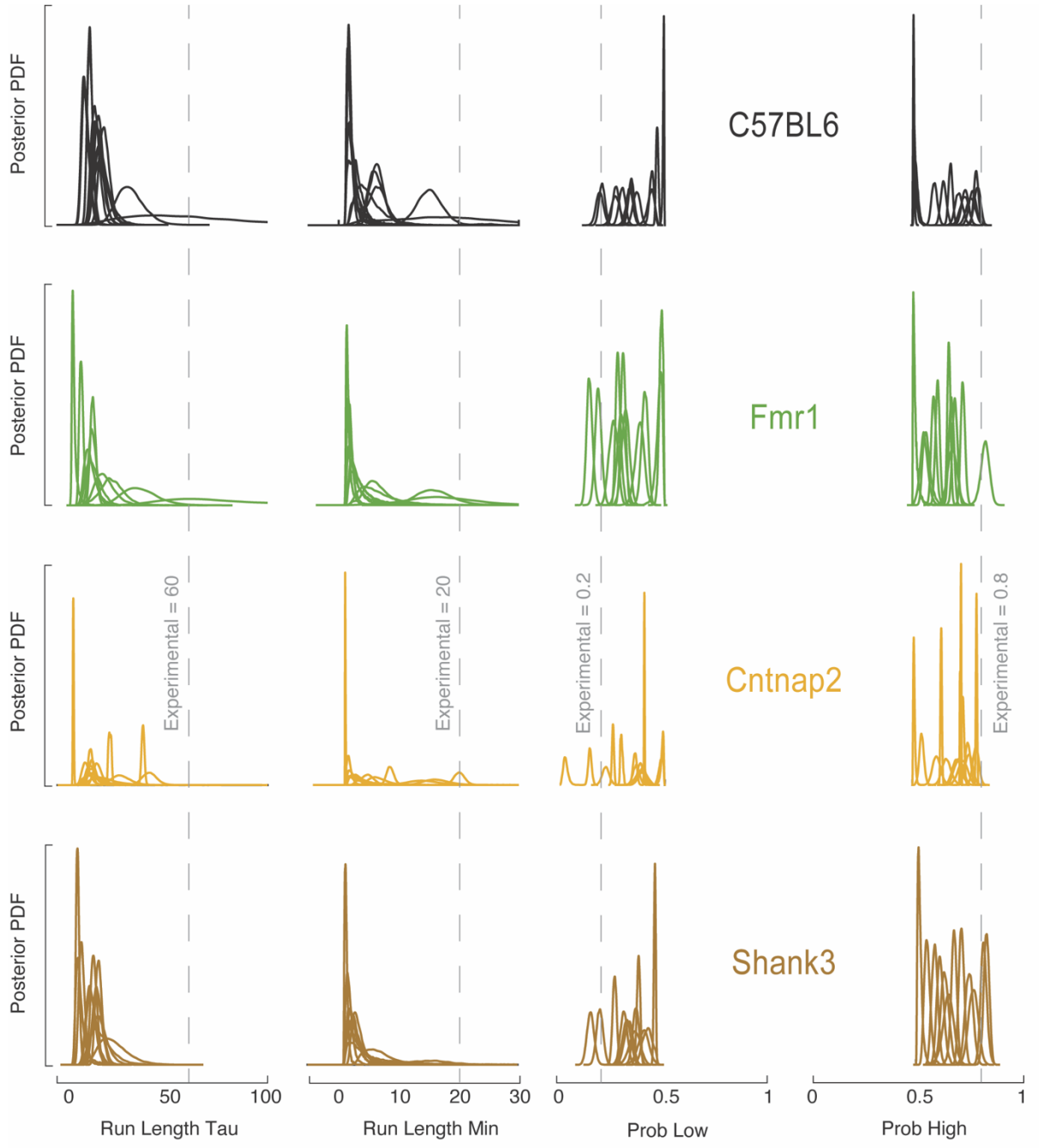

Figure S3. Change-point detection model parameters. We fit the Bayesian *change-point free run* model – the most flexible of the change-point models – to the animals behavioral responses. This model performs worse than the biased weighted average in general (i.e., when considering all animals), and statistically so for the Cntnap2 and Shank3 animals. It performs equally to the biased weighted average model in C57BL6 and Fmr1 animals. Here, we estimate full posteriors of the *c.p. free run* model via Variational Bayesian Monte Carlo (VBMC[103]). These, demonstrate that the posteriors were generally well-behaved (i.e., sufficiently narrow not to encompass the full parameter-space), yet far from the true experimental values. In particular, it is worth noting that the time constant and minimum run lengths estimated by animals were very concentrated near 0. In other words, animals did not use a generative model wherein there was a representation of blocks. Instead, the model parameters attempted to exclude this notion by making "blocks" as short as possible, effectively rendering this *c.p. free run* model akin to the heuristic models detailed in the main text.

Figure S4. Electrode track histology with CCF-atlas alignment overlaid.

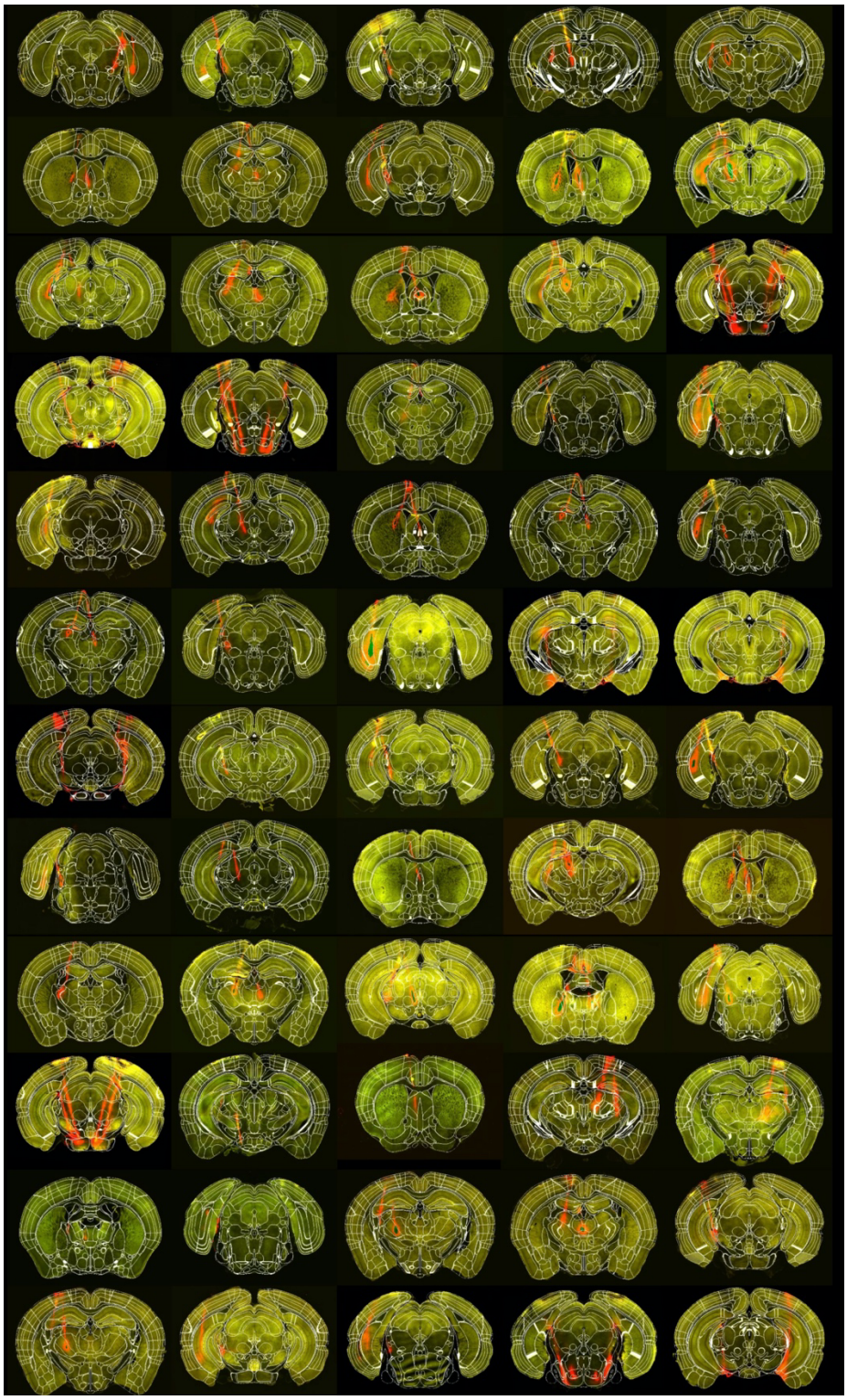

Figure S4. Electrode track histology with CCF-atlas alignment overlaid. Sixty coronal histology slices are shown. The dye is shown in red, tracking the electrode tracks. Allen Common Coordinate Framework (CCF) atlas is shown overlayed in white. Tracing these tracks rendered the problem of localizing units from 3-dimensional (in the brain), to 1-dimensional (along the track). Histology and physiology were then aligned (Falkner, 2020) the track to solve the 1-dimensional problem.

Figure S5. Demixed PCA variance explained in the subspace encoding the subjective prior.

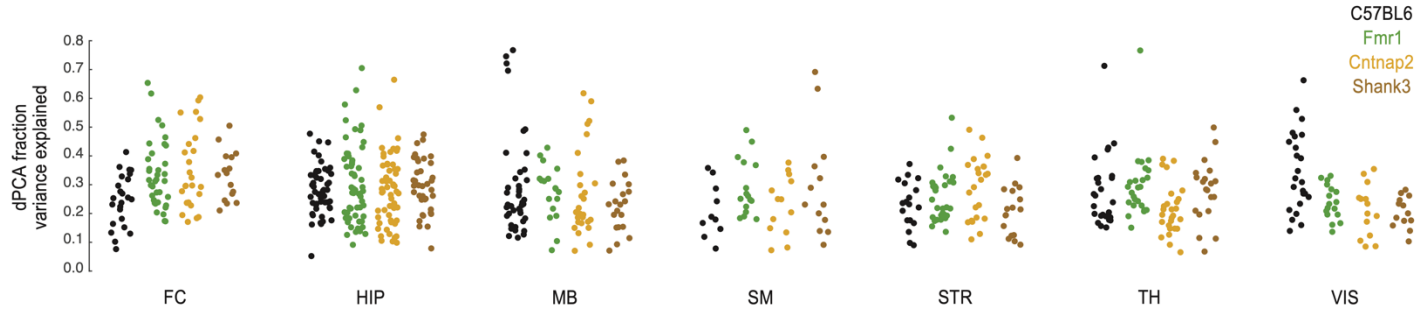

Figure S5. Demixed PCA variance explained in the subspace encoding the subjective prior. dPCA variance explained is expressed as a fraction of variance explained by PCA to normalize across brain areas. Each session is shown as an individual dot.

Figure S6. Informativeness of tuning functions regarding the subjective prior.

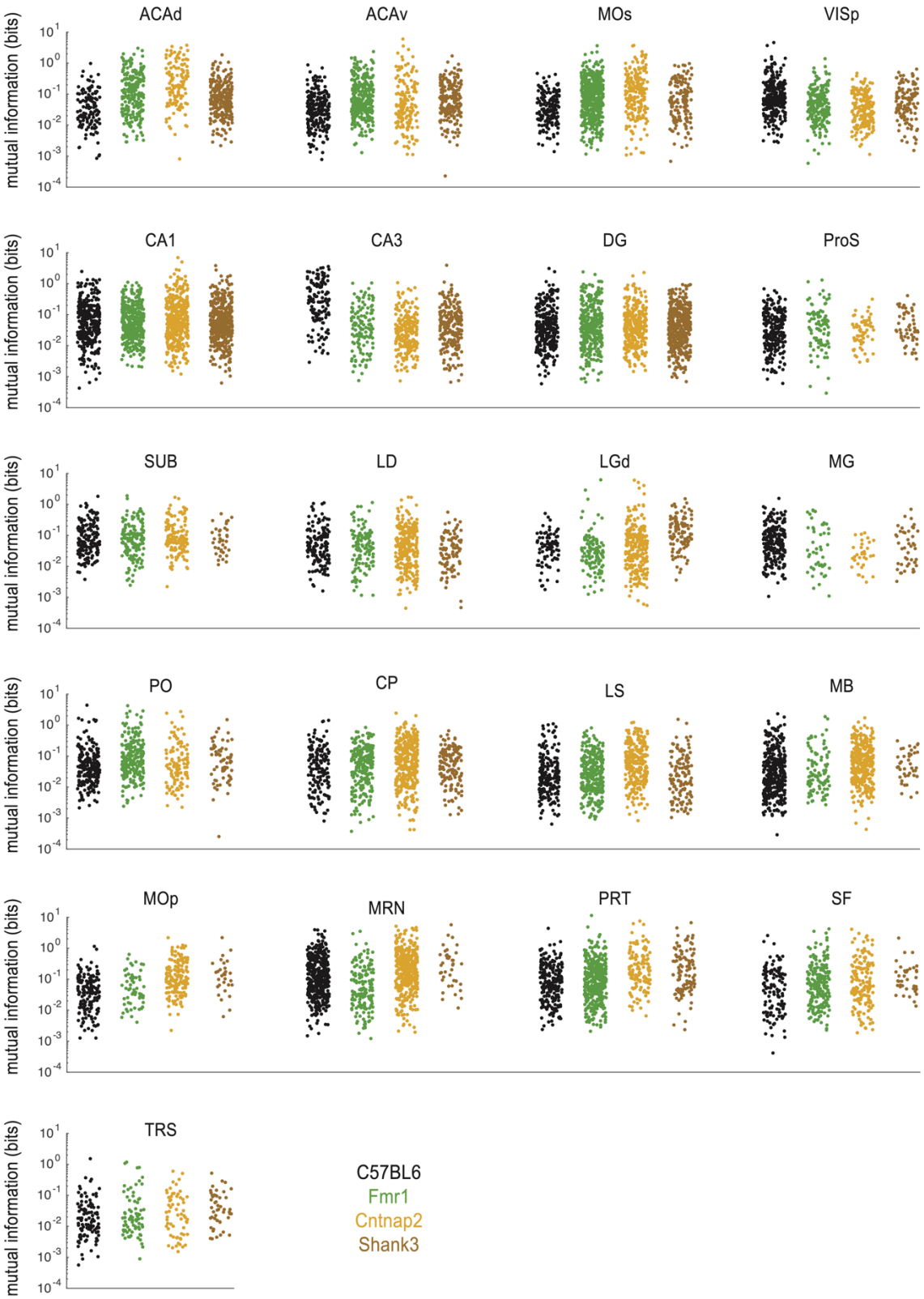

Figure S6. Informativeness of tuning functions (measured in mutual information) regarding the subjective prior value as a function of brain area (panels) and genotype (colors, follow convention from the rest of plots). Each unit is shown as an individual dot, y-axis is in a log-scale.

Supplementary Materials

Table S1. Categorization of areas and their acronyms.

HIP	CA1	Field CA1
	CA3	Field CA3
	DG	Dentate gyrus
	POST	Postsubiculum
	ProS	Prosubiculum
	SUB	Subiculum
MB	MB	Midbrain
	MRN	Midbrain reticular nucleus
	SC	Superior colliculus
	SNr	Substantia nigra reticular part
	PRT	Pretectal region
FC	ACAd	Anterior cingulate area dorsal part
	ACAv	Anterior cingulate area ventral part
	MOs	Secondary motor area
SM	MOp	Primary motor area
	RSPd	Retrosplenial area dorsal part
	RSPv	Retrosplenial area ventral part
STR	CP	Caudoputamen
	LS	Lateral septal complex
	TRS	Triangular nucleus of septum
	SF	Septofimbrial nucleus
TH	LD	Lateral dorsal nucleus of thalamus
	LGd	Dorsal part of the lateral geniculate complex
	LP	Lateral posterior nucleus of the thalamus
	MG	Medial geniculate complex
	PO	Posterior complex of the thalamus
	PVT	Paraventricular nucleus of the thalamus
	TH	Thalamus
	SGN	Suprageniculate nucleus
VIS	VISa	Anterior visual area
	VISam	Anteromedial visual area
	VISp	Primary visual area
Other	P	Pons
	PRNr	Pontine reticular nucleus