A Genetic Locus Mediating Attentional Processing

Gershon Z[1], Bonito-Oliva A[1], Kanke M[2], Terceros A[1], Fak J, Ianone A[3], Gebrehemedin M[1], De Marco Garcia NV[3], Sethupathy P[2], Rajasethupathy P[1]*.

[1]Laboratory of Neural Dynamics & Cognition, The Rockefeller University, New York, NY 10065 USA, [2]Department of Biomedical Sciences, Cornell University, Ithaca, NY 14853 USA, [3]Center for Neurogenetics, Brain and Mind Research Institute, Weill Cornell Medicine, New York, NY 10021, USA

** Correspondence: priya@rockefeller.edu*

Summary

Attention is required for most higher-order cognitive functions, but despite extensive and careful study, central unifying principles have been challenging to elicit. To provide a new perspective, we took a forward genetics approach to identify genes with large contributions to attentional performance. We studied 200 genetically diverse mice on measures of pre-attentive processing and through genetic mapping identified a small locus on chromosome 13 (95%CI: 92.22- 94.09 Mb), driving substantial variation (19%) in this trait. Further characterization of the locus revealed a causative gene, *Homer1a*, a synaptic protein, whose down-regulation specifically in prefrontal excitatory cells during a developmental critical period (<p14) led to significant improvements in multiple measures of attentional performance in the adult. Subsequent molecular and physiological studies revealed that prefrontal *Homer1* down-regulation is associated with GABAergic receptor up-regulation in those same cells and an overall inhibitory tone in prefrontal cortex. This inhibitory tone was relieved during task performance, where large increases in locus-coeruleus (LC) to prefrontal cortex (PFC) coupling led to sustained elevations in PFC activity, specifically prior to cue-onset, predicting short-latency correct responses. Notably high-*Homer1a*, low-attentional performers, exhibited constantly elevated LC-PFC correlations and PFC response magnitudes both at baseline and during task. Thus, rather than overall increases in neural activity, a scalable dynamic range of LC-PFC coupling and of pre-cue PFC responses supported attentional performance. We thus identify a gene with outsized contributions to attentional performance - *Homer1* - and link this with prefrontal inhibitory tone as an important component of dynamic task-dependent neuromodulation during attention.

Attention is the process of directing cognitive resources to particular stimuli and is a pre-requisite for higher-order cognition, such as short-term memory. It is a complex cognitive process that can be constant or fluctuating, occurring on slow or rapid time-scales, and can be broadly distributed in the brain but also highly specific to particular stimuli[1-5]. Years of foundational research have highlighted the importance of enhanced prefrontal activity in mediating attention. It is also well appreciated that the PFC interfaces with a broader neuroanatomically distributed network to enable attention. For instance, long-range recurrence with PFC, i.e., via thalamus[6,7], and neuromodulation[8-12], i.e., via adrenergic, dopaminergic and other systems, are thought to be key mediators of attention. Indeed, many ADHD medicines target these circuits[13]. However, there exist important limitations in our understanding of the specific circuits, cell-types, and underlying molecular pathways involved in this cognitive process. More importantly, we lack an understanding of which nodes in these complex pathways are most critical, which if identified, can inform more unifying models and therapeutic strategies for attentional processing.

In the past, unbiased genetic mapping approaches enabled the identification of genes with outsized contributions to a behavioral trait[14-17]. Further investigations of these genes identified critical cell types and circuits that led to unifying cellular models of behavior. Toward this goal, we recently performed genetic mapping in mice and identified a single gene with strong outsized contributions to short-term memory[18]. Building on the fruition of this previous work, we studied a cohort of genetically diverse mice on a pre-attentive processing task and identified a significant genetic locus on chromosome 13 linked to variation in this trait. Further characterization of the locus revealed a causative gene, *Homer1a*, a synaptic protein, whose chronic down-regulation in PFC during development, enhances attentional performance in adults. Notably, low-Homer, high-attention mice were associated with increased baseline GABAergic tone in PFC, but rapid adrenergic facilitation of prefrontal activity during task performance. We thus identify a single gene with outsized contributions to attentional performance – *Homer1a* – and link this with increased dynamic range, rather than overall magnitude, of adrenergic coupling to PFC as an important component of attention.

Identification of a QTL linked to pre-attentive processing.

The Diversity Outbred (DO) resource is a mouse population that is derived from eight founder strains, whose genetic diversity, i.e., SNP density and allelic heterozygosity, is comparable to that of the human population, providing an ideal platform for high-resolution genetic mapping (Fig. 1a). Based on our previous work characterizing the DO founders, 19th and 25th generations, we determined that successful mapping would require 1) An automated and robust behavioral assay with minimal training, thus narrowing the observed variance to genetic and task-associated features, and 2) Approximately 200 mice to detect a quantitative trait locus (QTL) that shifts the trait mean by 1 standard deviation at 95% confidence. Since traditional tasks of attention require extensive training, reward-associations, and other potential confounds for genetic mapping, we selected and optimized a simple, robust assay for pre-attentive process, i.e., pre-pulse inhibition of startle response (PPI).

Briefly, we measured the acoustic startle response of each DO mouse after a 120 dB tone (the startle stimulus) as well as when the startle stimulus is preceded by a weaker tone (the pre-pulse, 3, 6, or 12 dB), and then calculated the ratio as the measure of pre-attentive processing (PPI). Notably, while the startle to pulse is considered a "bottom-up" process[19], the inhibition of this startle is a "top-down" process[20] that has been linked to downstream measures of cognition, including selective[21] and sustained[22] attention.

We tested 191 mice (Methods) for performance on this PPI task and first confirmed that the phenotypic variability of the DO greatly surpassed that of the C57BL/6J (B6) classical inbred line, as would be expected from the underlying genetic variation (Fig. 1b). To minimize potential confounds in PPI performance that may not be due to pre-attentive processing, i.e., deficits in sensory or motor impairments, we assessed and excluded 15 mice that exhibited greater PPI3 than PPI12, suggesting potential hearing impairment. With the remaining mice, we assessed and found no significant correlations between PPI and startle response, gross motor activity, or body weight (Extended Data Fig.1b-e).

We next genotyped the 176 DO mice using the GigaMUGA platform (114,184 loci had variability in our cohort) and performed QTL mapping for pre-attentive processing based on pre-pulse inhibition (PPI) scores using R/qtl2[23] (Extended Data Fig. 1f-g). As in our previous DO study[18], founder haplotype reconstructions showed extensive allelic heterozygosity genome-wide (Extended Data Fig. 1a) and we observed approximately equal founder contributions across our cohort. The QTL analysis revealed a statistically significant locus for PPI6 on chromosome 13 with a genome-wide p value of ≤0.01 (LOD score = 8.22, 95%CI: 92.22- 94.09 Mb). This locus was also confirmed to be statistically significant using a second mapping approach, miQTL (Fig. 1c). QTL mapping of PPI3 and PPI12 did not reveal any loci that surpassed significance thresholds, but a suggested peak for PPI3 indeed mapped to the same Chr 13 QTL (Extended Data Fig. 1g, left; LOD score = 5.88, 95%CI: 90.51-94.09 Mb), further adding confidence to the functional significance of this locus.

Next, to increase confidence in this locus we performed an allele effects analysis (Methods) and found that the B6 haplotype (henceforth referred to as Chr13QTL^{B6}) was associated with high performance while the WSB/EiJ haplotype (henceforth referred to as Chr13QTL^{WSB}) was associated with low performance (Fig. 1d-e). We then asked whether new strains of recombinant inbred Collaborative Cross (CC) lines, which have the same multi-parent origins as the DO (Fig. 1f), that possess either Chr13QTL^{B6} or Chr13QTL^{WSB} would separate into high and low performers, respectively. We analyzed the genomes of existing CC lines and selected two that were homozygous for our desired Chr13QTL^{B6} (CC083) or Chr13QTL^{WSB} (CC025) haplotypes, while maintaining distinctive mosaic representations of the founder genomes at other loci. We compared PPI performance between CC083 and CC025 and found that CC083 have significantly greater PPI than CC025 (Fig. 1g, n(CC083)=27, n(CC025)=24, p(PPI6)=0.05, two-way ANOVA with Sidak's test for multiple comparison; Extended data Fig. 2a, Welch-corrected t-test, $p < 0.001$) which, similarly to the DO, was not due to differences in peak startle, body weight, or gross motor behavior (Extended Data Fig. 2b-d). These data further

support that the genotypic differences at the Chr13 locus, specifically WSB vs B6 genotype, drive significant variation in pre-attentive processing.

Since pre-attentive processing is linked to downstream measures of attention, we next asked whether the Chr 13 QTL drives differences in attentional performance, as measured by an operant signal detection task (SDT). Here, mice are trained to nose-poke in response to a 5s auditory cue within 10s of cue onset to receive a food reward. Once the mice have sufficiently learned the task (Methods), their attentional load is then challenged by decreasing the length of the cue to 1s and reducing the response window to 6s from cue onset (Fig. 1i). Similar signal detection tasks have been widely used to assay attention[24]. They provide multiple metrics to track attention including accuracy, latency of response, and trial omissions. During the initial 5 s cue training, there were no significant differences in learning the task, but CC083's were already exhibiting fast latency responses, and after increased attentional load during the 1s trials, the CC083s significantly outperformed the CC025s in all of measures of attention including accuracy, latency to all responses, proportion of delayed responses, latency to correct responses (Fig. 1j-m). Notably, the strains did not differ in other cognitive measures such as short or long term memory, nor in measures of motor, motivational, or social behavior (Extended Data Fig. 2e-h). We did observe differences in measures of anxiety (Extended Data Fig. 2i), which requires further consideration given important dependencies between anxiety and attention. In sum, these data together suggest that genetic variation at the Chr 13 locus drives differences in attentional performance.

Chr 13 QTL effects on attention are driven by *Homer1*

With a QTL in hand, we next wanted to understand which gene within the locus was driving the changes in attentional performance. Thus, we performed bulk RNA sequencing from DO high and low performers, focusing on the prefrontal cortex (PFC) for its central role in attentional processing, but also including related brain areas such as mediodorsal thalamus (MD), and the ventral tegmental area (VTA). We found that samples stratified by performance in PFC and MD, but not in VTA (Extended Data Fig. 3a), leading us to then ask which locus genes (Extended Data Table 1) were differentially expressed in MD or PFC between high and low performers. Of all locus genes, only *Homer1* was significantly differentially expressed, with substantial downregulation in PFC in high-performers (Fig. 2a, n=3 each, adjusted $p<0.001$). *Homer1* has several transcript variants due to alternative splicing[25] (Extended Data Fig. 3b), and thus we assessed whether differential expression was uniform across splice isoforms. Strikingly, only the short, activity-dependent isoforms, *Homer1a*[26,27] and *Ania3*[25], were differentially expressed between DO high- and low-performers (Fig. 2b, p(Homer1a)=0.003, p(Ania3)=0.007, two-way ANOVA with post hoc Holm-Sidak test for multiple comparisons). Furthermore, bulk RNAseq from high- (CC083) and low- (CC025) performing CC strains also confirmed significant differences in *Homer1*, as well as gene ontologies (GOs) relating to *Homer1*'s function in excitatory neurotransmission and activity (Fig. 2c). Similar to the DO mice, the differential Homer1 expression in CC's was driven by downregulation of *Homer1a* and *Ania3* isoforms in the high performing CC083s (Fig. 2d, 2-way ANOVA $p<0.001$, Holm-Sidak test for multiple comparisons p(Homer1a)<0.001, p(Ania3)<0.001).

Since Homer1a is better characterized and conserved than *Ania3*[25,28], we next assessed whether *Homer1a* manipulations could drive behavioral changes in attentional performance (Fig. 2e; Extended Data Fig. 4). To knock down *Homer1a*, we designed and tested AAV-based short-hairpin RNAs (shRNAs) to target the *Homer1a* isoform *in vitro* (Extended Data Fig. 4a) and selected the most effective shRNA for bilateral PFC injections *in vivo* and behavioral testing (Fig. 2e; Extended Data Fig. 4). To over-express *Homer1a,* which has endogenous expression primarily in excitatory pyramidal neurons, we cloned the *Homer1a* coding sequence into an AAV-based CamKII(1.3)-eYFP vector (Extended Data Fig. 5a) for bilateral PFC injection and behavioral testing (Fig. 2f; Extended Data Fig. 5b-d). To our surprise, we did not see any significant behavioral effect for either the knockdown or overexpression experiments (Fig. 4e & 5f; Extended Data Fig. 4f & 5e).

To assess whether the effects of *Homer1a* may be developmental in origin, which was supported by prior work on germline knockouts[29-32], we profiled the expression of *Homer1a*, *Ania3*, and *Homer1b/c* in CC083 and CC025 mice across postnatal development. We found that, the expression of *Homer1a* and *Ania3,* but not that of Homer1b/c, diverged between strains as early as p14 (Fig. 2g; two-way ANOVA p=0.02), suggesting possible developmental roles in regulating attentional processing. To test this hypothesis, we knocked down *Homer1a* and *Ania3* isoforms during early developmental stages (<p14) to evaluate the effect on adult behavior. We injected the AAV shRNA targeting *Homer1a* pooled together with an AAV-based shRNA for *Ania3*, which we validated *in vitro* (Extended Data Fig. 6) and injected bilaterally into the PFC of B6 pups between p0-p2 (Fig. 2h; Extended Data 7a; referred to as KD_{dev}). Despite the developmental *Homer1a* knockdown being less effective than the adult manipulation (~80% in adults and ~60% in pups; Extended Data Fig. 4e & Extended Data Fig. 7b), we observed a large and significant improvement in measures of pre-attentive processing (PPI, Fig. 2h, Extended Data Fig. 7c-f) and attentional performance (SDT, Fig. 2i-l) in KD_{dev} mice compared with scramble controls. For instance, when comparing response latencies on correct trials, the most sensitive measure of attentional performance, KD_{dev} mice exhibited significantly faster response latencies than controls, that persisted throughout the extent of both cue length phases (Fig. 2k-l, repeated-measures two-way ANOVA p(5s cue)=0.04, p(1s cue)<0.001). As with CC mice, KD_{dev} and controls displayed no significant differences in measures of short-term memory or motor behavior, but notably in contrast to CC mice, they also displayed no significant differences in measures of anxiety (Extended Data Fig. 7g-i). Altogether, these results led us to conclude that reducing the expression of *Homer1a* and *Ania3* in PFC during early postnatal development (<p14) is sufficient to improve pre-attentive processing and attentional performance in adult. This raises two important questions: 1) how does endogenous differential expression of short *Homer1* isoforms throughout development affect cellular functions underlying attention in the adult, and 2) how do these cellular and molecular changes influence neural dynamics supporting attention?

Low Homer1a expressing neurons up-regulate GABA-receptors.

To better understand the differences in transcriptional programs associated with *Homer1* we performed single cell RNAseq (scRNAseq) from PFC of adult CC083 and CC025 mice (Fig. 3a; Methods). After removing low quality cells (Methods) we obtained 70,920 total cells (Extended Data Fig. 8a; 40,897 from CC083 and 30,023 from CC025, n=2 biological replicates of 3 mice pooled per strain per replicate). We performed graph-based weighted nearest neighbors clustering analysis and identified the major cell types based on cluster-wide expression of several canonical marker genes[33] (Fig 3b; Extended Data Fig. 8b-c; Methods).

Since *Homer1* is primarily expressed in neurons[34], we sub-clustered the neurons (4,633 cells) and re-clustered them based on the first 50 principal components, identifying 15 distinct neuronal clusters (Fig. 3c; Methods). We determined that 14 of the clusters were glutamatergic and 1 was GABAergic based on expression of the marker genes *Slc17a6*, *Slc17a7*, *Slc32a1*, and *Gad1* (Fig. 3d); and consistent with previous studies[35,36], *Homer1* expression was primarily restricted to glutamatergic neurons (Fig. 3e). Of the 14 glutamatergic clusters, 5 showed substantial downregulation of *Homer1* in CC083 cells compared to CC025 cells (Fig. 3f, referred to as *Homer1* DE clusters; t-tests, * indicates $p < 0.1$). To ask how differing levels of *Homer1* expression affects transcriptional programs and cellular function, we performed differential expression analysis on the *Homer1* DE clusters between strains using MAST[37] (Methods). GO analysis of molecular function for genes upregulated in the CC083 cells from the *Homer1* DE clusters showed an overrepresentation of genes associated with inhibitory GABA signaling, while CC025 terms overrepresented glutamatergic signaling (Fig. 3g-h; Extended Data 8e). In fact, in the *Homer1* DE clusters, the CC083 cells uniformly upregulate several GABA receptor sub-types while downregulating several glutamatergic receptor subtypes, with almost no differential expression of other neurotransmitter receptors or transporters (Fig. 3i). These data indicate that lower neuronal expression of *Homer1* in a subset of prefrontal excitatory neurons yields enhanced GABAergic to glutamatergic receptor balance within those same neurons.

We thus turned to assessing the transcriptional programs of upstream GABAergic neurons. To do so, we performed differential expression analysis using MAST on the GABAergic cluster, in which, interestingly, *Homer1* is significantly upregulated in the CC083s (Fig. 2f, $p = 0.021$). Due to the well-studied contributions of neuromodulation in attentional processing[12], we assessed expression differences of markers for the most common neuromodulatory systems and found that CC083 GABAergic neurons had higher expression of genes associated with adrenergic and cholinergic signaling than the CC025s (Fig. 3j). Furthermore, pathway enrichment analysis[38,39] indicated significant overrepresentation of genes related to noradrenergic signaling in CC083s (Extended Data Fig. 8f). Given its historical significance in attentional regulation[4,9,40,41], as well as its role as a target of medications to treat ADHD[42,43], we further analyzed the adrenergic receptors. We found that the higher expression of adrenergic marker genes in CC083 GABAergic cells is driven primarily by the adrenergic receptor *Adra1b* (Fig. 3k), which, along with *Adra1a*, appear to be preferentially expressed in the GABAergic cluster (Extended Data Fig. 8d). Taken together, these data suggest that prefrontal cells with

lower *Homer1* expression are associated with increased GABAergic signaling onto those neurons, likely supported by adrenergic neuromodulation. With this is mind, we were keen to explore the neural dynamics of animals naturally varying in PFC *Homer1* expression during an attention task.

Enhanced range of LC-PFC coupling and pre-cue PFC responses support attention

How do the Homer1a-associated molecular and cellular changes in CC mice contribute to changes in neural dynamics underlying attentional processing? Are there roles for long-range inhibitory recurrence via thalamus, or neuromodulation via locus coeruleus, or both, linked to attentional performance? And are there contributions of non-neuronal cells (i.e., oligodendrocyte precursor cells (OPCs) or mature oligodendrocytes (OLs) (Extended Data Fig. 9a-d) in regulating prefrontal dynamics and attentional processing? To address these and other questions, we moved to an *in vivo* preparation to record multi-area brain activity as mice performed the operant signal detection task. We injected AAV1/9-GCaMP or JRGECO1a into locus coeruleus (LC, GCaMP), mediodorsal thalamus (MD, GCaMP), PFC oligodendrocytes (OL, GCaMP; creation and validation of OL-specific AAV-MAG-GCaMP in Extended Data Fig. 9e-g) and PFC neurons (PFC, JRGECO1a), implanted optical fibers above each region, and used a custom dual-color fiber photometry system to record bulk calcium signals from these regions simultaneously in behaving mice (Fig. 4a). Because CC025 mice did not tolerate intracranial implants we used B6 mice as “low performers” since they have comparable *Homer1a* expression and behavioral performance to CC025s. Multi-area neural activity recordings from a given animal were frame projected onto a camera sensor, and custom Matlab scripts (Methods) were used to extract time-series data, regress out motion-related artifacts, and align to behavioral data (Example alignment from one trial of a CC083 mouse shown in Fig. 4b).

We first analyzed baseline activity patterns in CC083 and B6 mice and noticed substantially elevated LC activity ($p<0.01$, unpaired t-test), and depressed PFC activity ($p<0.001$, unpaired t-test), in CC083 mice (Fig. 4c). This, together with the scSeq observation of increased adrenergic Adra1b reception onto PFC GABAergic cells (Fig. 3k), suggested that LC may contribute to strong baseline inhibitory tone in PFC through feed-forward inhibition. Indeed, we found that LC-PFC functional correlations at baseline were close to Pearson’s $r=0$ in CC083s compared to ~0.7 in B6 (Fig. 4d, $p<0.001$, unpaired t-test). However, as mice started training on the signal detection task, we noticed a steady increase in LC-PFC functional correlations in CC083 mice that mirrored their steady improvement in task performance and measures of attention, and which was not observed in B6 mice, presumably due to already high baseline correlations precluding further enhancements during task (Fig. 4e, example raw traces from day 2 and day 11 are shown). Notably, this improved LC-PFC coupling in CC083 mice was strongest in the first four minutes (block 1) of each day’s session (Fig. 4e, left panel), after which these correlations reduced back toward baseline (blocks 2-5; Fig. 4e, right panel). Meanwhile, block 1 elevations in LC-PFC coupling in CC083s were associated with large elevations in PFC pre-cue responses, which strikingly persisted throughout the remaining blocks of each day’s session (Fig. 4f). Furthermore, such pre-cue activity was significantly elevated on short-latency correct trials compared with long-latency incorrect trials, and specifically in CC083 mice (Fig. 4g, example raw traces from correct and incorrect trials shown).

Again, these dynamic task- and cue- dependent PFC activity changes were not observed in high-*Homer1a* B6 mice, which exhibited relatively high prefrontal responses throughout baseline, early, and late days of the task (Fig. 4c,f,g).

Taken together, these results suggest a model in which low *Homer1a* mediates high baseline inhibitory tone, large dynamic range in LC-PFC coupling, targeted elevations in pre-cue activity, for short-latency accurate responses and high attentional performance. These results therefore underscore that the critical component of neuromodulation during attention may not be overall increases in adrenergic activity, but rather frequent LC-PFC resets that enable increased signal to noise and targeted behavioral responses.

Discussion

Here we performed unbiased genetic mapping in outbred mice and identified a short segment on chromosome 13 that is critical for explaining variation in attentional performance. Within this locus, we identified the causative gene to be *Homer1a*, a scaffolding protein with known roles in regulating excitatory glutamatergic synaptic transmission. We found that while manipulation of *Homer1a* in the adult led to no significant changes in attentional performance, knockdown specifically in prefrontal excitatory cells during a developmental critical period (<p14) led to significant improvements in pre-attentive processing and attentional performance in the adult.

A rich history of work on *Homer1* and its isoforms have revealed important roles in multiple cognitive domains[29-32], but little is known about its role in attention. Given strong *Homer1a* association with schizophrenia[44,45], autism[46,47] and ADHD[48,49], it's possible that *Homer1a*-related early dysfunctions in sensory gating and attention provide common etiology driving diverse downstream neuropsychiatric symptoms. Thus, prefrontal *Homer1a* may be a critical hub for deeper mechanistic understanding of attention, indeed it has outsized contribution to the trait (19%) and may therefore point us toward unifying cellular models.

In our initial attempts to understand how *Homer1a* might shape behavioral improvements in attentional performance, we explored the molecular programs associated with *Homer1a* expressing neurons as well as their associated circuit physiology in the context of their inputs and outputs. Through cellular resolution RNA sequencing analysis, we found that low-*Homer1a,* high attention mice (CC083) only down-regulate *Homer1a* expression in a subset of prefrontal excitatory neurons, which in turn is associated with significant upregulation of GABAergic receptors in these same cells. Assuming these receptors receive inputs from local GABAergic neurons, we further analyzed these inhibitory neurons and found enrichment of a specific adrenergic receptor gene, *Adra1b*, in these cells. These data together suggested that in mice with high attentional performance, chronic downregulation of *Homer1a* may drive homeostatic scaling, favoring inhibitory inputs and overall inhibitory tone in prefrontal cortex. Indeed *in vivo* simultaneous neural activity recordings of LC, MD, and PFC revealed strongly reduced baseline prefrontal activity in high attention mice. However, during task performance, this inhibitory tone was relieved and large increase in the gain of LC-to-PFC functional correlations at the start of a task led to sustained elevations in prefrontal activity, specifically prior to cue onset, for the remainder of the task session. Notably high-

Homer1a, low-attentional performers, exhibited constantly elevated LC-PFC correlations and PFC response magnitudes both at baseline and during task. Thus, rather than overall increases in neural activity, a scalable dynamic range of LC-PFC coupling and targeted pre-cue PFC responses enabled attentional performance.

While the current investigations highlight new mechanisms of attention related to the subtle interplay of prefrontal inhibitory tone with dynamic neuromodulation of PFC, an understanding of the more complete effects of *Homer1a* require deeper investigation. For instance, GABAergic cells from low-*Homer1a* mice also upregulate cholinergic signaling in PFC, suggesting cross-talk between neuromodulatory systems, likely over diverse time-scales. We also note that throughout the study there were interesting links between oligodendrocyte function and attentional performance that may be important to investigate further. For instance, the top 20 genes upregulated in PFC of the DO high performers mapped almost exclusively to expression in the oligodendrocyte lineage (Extended Data Fig. 9a-b), and a similar shift in oligodendrocyte transcriptional programs was also noticeable between the CC high and low performers (Extended Data Fig. 9c-d). Furthermore, preliminary neural activity recordings in prefrontal oligodendrocytes revealed dynamic changes in calcium responses during task that were correlated with LC responses. As such calcium events have been linked with myelination[50], this supports roles for activity-dependent myelination in attentional processing[51-54]. Overall, the development and characterization of mouse lines that model high and low attention, together with the identification of a single gene with outsized contributions to the trait, provides tractable inroads into more comprehensive models of attentional processing.

Acknowledgements
We are grateful to Mary E. Hatten, Cori Bargmann, Bob Darnell, and Rajasethupathy lab members for helpful discussions. We thank Josue Regalado for help with photometry experiments and analysis, Vivian Li for help with surgeries, and Brian Fabella and James Hudspeth for help with ABR testing. We thank Paul Greengard's laboratory for sharing behavioral instrumentations. We are thankful to core facilities at Rockefeller (Precision Instrumentation, FACS), MSKCC (Genomics & Sequencing), and Univ of Arizona (Viral Core). This work is supported by a Kavli Neuroscience Institute pilot grant from the Rockefeller University (to A.T), and grants from the Harold & Leila Mathers Foundation and the National Institutes of Health under award number DP2AG058487 (to P.R).

Author Contributions
ZG, PR and PS conceived the study. ZG, ABO, and PR designed the experiments. ZG selected and optimized the behaviors used in this study, performed mouse behaviors (together with ABO), molecular studies using RNAseq, scRNAseq (together with AT), cloning and cell-based assays, and in vivo neural activity recordings and analysis, supervised by PR. MK performed QTL analysis as well as RNAseq and scRNAseq analysis, supervised by PS. AT assisted with designing and performing the scRNAseq experiments and analysis. JF and MG assisted with surgeries and histology. AI helped design experiments and perform surgeries for the developmental study, supervised by ND. ZG and PR wrote the paper with input from all authors.

Declaration of Interests: The authors declare no competing interests.

References

1 Robbins, T. W. Arousal systems and attentional processes. *Biol Psychol* **45**, 57-71 (1997). https://doi.org:10.1016/s0301-0511(96)05222-2

2 Petersen, S. E. & Posner, M. I. The attention system of the human brain: 20 years after. *Annu Rev Neurosci* **35**, 73-89 (2012). https://doi.org:10.1146/annurev-neuro-062111-150525

3 Harris, K. D. & Thiele, A. Cortical state and attention. *Nat Rev Neurosci* **12**, 509-523 (2011). https://doi.org:10.1038/nrn3084

4 Arnsten, A. F., Wang, M. J. & Paspalas, C. D. Neuromodulation of thought: flexibilities and vulnerabilities in prefrontal cortical network synapses. *Neuron* **76**, 223-239 (2012). https://doi.org:10.1016/j.neuron.2012.08.038

5 Buschman, T. J. & Kastner, S. From Behavior to Neural Dynamics: An Integrated Theory of Attention. *Neuron* **88**, 127-144 (2015). https://doi.org:10.1016/j.neuron.2015.09.017

6 Halassa, M. M. & Kastner, S. Thalamic functions in distributed cognitive control. *Nat Neurosci* **20**, 1669-1679 (2017). https://doi.org:10.1038/s41593-017-0020-1

7 Varela, C. Thalamic neuromodulation and its implications for executive networks. *Front Neural Circuits* **8**, 69 (2014). https://doi.org:10.3389/fncir.2014.00069

8 Lee, S. H. & Dan, Y. Neuromodulation of brain states. *Neuron* **76**, 209-222 (2012). https://doi.org:10.1016/j.neuron.2012.09.012

9 Aston-Jones, G. & Cohen, J. D. An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. *Annu Rev Neurosci* **28**, 403-450 (2005). https://doi.org:10.1146/annurev.neuro.28.061604.135709

10 Robbins, T. W. & Arnsten, A. F. The neuropsychopharmacology of fronto-executive function: monoaminergic modulation. *Annu Rev Neurosci* **32**, 267-287 (2009). https://doi.org:10.1146/annurev.neuro.051508.135535

11 Bargmann, C. I. & Marder, E. From the connectome to brain function. *Nat Methods* **10**, 483-490 (2013). https://doi.org:10.1038/nmeth.2451

12 Thiele, A. & Bellgrove, M. A. Neuromodulation of Attention. *Neuron* **97**, 769-785 (2018). https://doi.org:10.1016/j.neuron.2018.01.008

13 Chappell, P. B. *et al.* Guanfacine treatment of comorbid attention-deficit hyperactivity disorder and Tourette's syndrome: preliminary clinical experience. *J Am Acad Child Adolesc Psychiatry* **34**, 1140-1146 (1995). https://doi.org:10.1097/00004583-199509000-00010

14 Dudai, Y., Jan, Y. N., Byers, D., Quinn, W. G. & Benzer, S. dunce, a mutant of Drosophila deficient in learning. *Proc Natl Acad Sci U S A* **73**, 1684-1688 (1976). https://doi.org:10.1073/pnas.73.5.1684

15 Bargiello, T. A., Jackson, F. R. & Young, M. W. Restoration of circadian behavioural rhythms by gene transfer in Drosophila. *Nature* **312**, 752-754 (1984). https://doi.org:10.1038/312752a0

16 de Bono, M. & Bargmann, C. I. Natural variation in a neuropeptide Y receptor homolog modifies social behavior and food response in C. elegans. *Cell* **94**, 679-689 (1998). https://doi.org:10.1016/s0092-8674(00)81609-8

17 Bendesky, A. *et al.* The genetic basis of parental care evolution in monogamous mice. *Nature* **544**, 434-439 (2017). https://doi.org:10.1038/nature22074

18 Hsiao, K. *et al.* A Thalamic Orphan Receptor Drives Variability in Short-Term Memory. *Cell* **183**, 522-536.e519 (2020). https://doi.org:10.1016/j.cell.2020.09.011

19 Koch, M. & Schnitzler, H. U. The acoustic startle response in rats--circuits mediating evocation, inhibition and potentiation. *Behav Brain Res* **89**, 35-49 (1997). https://doi.org:10.1016/s0166-4328(97)02296-1

20 Li, L., Du, Y., Li, N., Wu, X. & Wu, Y. Top–down modulation of prepulse inhibition of the startle reflex in humans and rats. *Neuroscience & Biobehavioral Reviews* **33**, 1157-1167 (2009). https://doi.org:10.1016/j.neubiorev.2009.02.001

21 Scholes, K. E. & Martin-Iverson, M. T. Disturbed prepulse inhibition in patients with schizophrenia is consequential to dysfunction of selective attention. *Psychophysiology* **47**, 223-235 (2010). https://doi.org:10.1111/j.1469-8986.2009.00927.x

22 Scholes, K. E. & Martin-Iverson, M. T. Relationships between prepulse inhibition and cognition are mediated by attentional processes. *Behav Brain Res* **205**, 456-467 (2009). https://doi.org:10.1016/j.bbr.2009.07.031

23 Broman, K. W. *et al.* R/qtl2: Software for Mapping Quantitative Trait Loci with High-Dimensional Data and Multiparent Populations. *Genetics* **211**, 495-502 (2019). https://doi.org:10.1534/genetics.118.301595

24 Turner, K. M., Peak, J. & Burne, T. H. J. Measuring Attention in Rodents: Comparison of a Modified Signal Detection Task and the 5-Choice Serial Reaction Time Task. *Frontiers in Behavioral Neuroscience* **9**, 370 (2016). https://doi.org:10.3389/fnbeh.2015.00370

25 Bottai, D. *et al.* Synaptic Activity-Induced Conversion of Intronic to Exonic Sequence in Homer 1 Immediate Early Gene Expression. *The Journal of Neuroscience* **22**, 167 (2002). https://doi.org:10.1523/JNEUROSCI.22-01-00167.2002

26 Brakeman, P. R. *et al.* Homer: a protein that selectively binds metabotropic glutamate receptors. *Nature* **386**, 284-288 (1997). https://doi.org:10.1038/386284a0

27 Kato, A., Ozawa, F., Saitoh, Y., Hirai, K. & Inokuchi, K. vesl, a gene encoding VASP/Ena family related protein, is upregulated during seizure, long-term potentiation and synaptogenesis. *FEBS Lett* **412**, 183-189 (1997). https://doi.org:10.1016/s0014-5793(97)00775-8

28 Xiao, B. *et al.* Homer Regulates the Association of Group 1 Metabotropic Glutamate Receptors with Multivalent Complexes of Homer-Related, Synaptic Proteins. *Neuron* **21**, 707-716 (1998). https://doi.org:10.1016/S0896-6273(00)80588-7

29 Szumlinski, K. K. *et al.* Behavioral and neurochemical phenotyping of Homer1 mutant mice: possible relevance to schizophrenia. *Genes Brain Behav* **4**, 273-288 (2005). https://doi.org:10.1111/j.1601-183X.2005.00120.x

30 Lominac, K. D. *et al.* Distinct Roles for Different Homer1 Isoforms in Behaviors and Associated Prefrontal Cortex Function. *The Journal of Neuroscience* **25**, 11586 (2005). https://doi.org:10.1523/JNEUROSCI.3764-05.2005

31 Jaubert, P. J. *et al.* Complex, multimodal behavioral profile of the Homer1 knockout mouse. *Genes, Brain and Behavior* **6**, 141-154 (2006). https://doi.org:10.1111/j.1601-183X.2006.00240.x

32 Datko, M. C. *et al.* Behavioral and Neurochemical Phenotyping of Mice Incapable of Homer1a Induction. *Frontiers in Behavioral Neuroscience* **11**, 208 (2017). https://doi.org:10.3389/fnbeh.2017.00208

33 Zeisel, A. *et al.* Molecular Architecture of the Mouse Nervous System. *Cell* **174**, 999-1014.e1022 (2018). https://doi.org:10.1016/j.cell.2018.06.021

34 Shiraishi-Yamaguchi, Y. & Furuichi, T. The Homer family proteins. *Genome Biology* **8**, 206 (2007). https://doi.org:10.1186/gb-2007-8-2-206

35 Ango, F. *et al.* Dendritic and axonal targeting of type 5 metabotropic glutamate receptor is regulated by homer1 proteins and neuronal excitation. *J Neurosci* **20**, 8710-8716 (2000). https://doi.org:10.1523/JNEUROSCI.20-23-08710.2000

36 Petralia, R. S. *et al.* Glutamate receptor targeting in the postsynaptic spine involves mechanisms that are independent of myosin Va. *Eur J Neurosci* **13**, 1722-1732 (2001). https://doi.org:10.1046/j.0953-816x.2001.01553.x

37 Finak, G. *et al.* MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biology* **16**, 278 (2015). https://doi.org:10.1186/s13059-015-0844-5

38 Chen, E. Y. *et al.* Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* **14**, 128 (2013). https://doi.org:10.1186/1471-2105-14-128

39 Kuleshov, M. V. *et al.* Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Research* **44**, W90-W97 (2016). https://doi.org:10.1093/nar/gkw377

40 de Lecea, L., Carter, M. E. & Adamantidis, A. Shining light on wakefulness and arousal. *Biol Psychiatry* **71**, 1046-1052 (2012). https://doi.org:10.1016/j.biopsych.2012.01.032

41 Lovett-Barron, M. *et al.* Ancestral Circuits for the Coordinated Modulation of Brain State. *Cell* **171**, 1411-1423 e1417 (2017). https://doi.org:10.1016/j.cell.2017.10.021

42 Banaschewski, T., Roessner, V., Dittmann, R. W., Santosh, P. J. & Rothenberger, A. Non-stimulant medications in the treatment of ADHD. *Eur Child Adolesc Psychiatry* **13 Suppl 1**, I102-116 (2004). https://doi.org:10.1007/s00787-004-1010-x

43 Cinnamon Bidwell, L., Dew, R. E. & Kollins, S. H. Alpha-2 adrenergic receptors and attention-deficit/hyperactivity disorder. *Curr Psychiatry Rep* **12**, 366-373 (2010). https://doi.org:10.1007/s11920-010-0136-4

44 Norton, N. *et al.* Mutation screening of the Homer gene family and association analysis in schizophrenia. *Am J Med Genet B Neuropsychiatr Genet* **120B**, 18-21 (2003). https://doi.org:10.1002/ajmg.b.20032

45 Spellmann, I. *et al.* Homer-1 polymorphisms are associated with psychopathology and response to treatment in schizophrenic patients. *J Psychiatr Res* **45**, 234-241 (2011). https://doi.org:10.1016/j.jpsychires.2010.06.004

46 Kelleher, R. J., 3rd *et al.* High-throughput sequencing of mGluR signaling pathway genes reveals enrichment of rare variants in autism. *PLOS ONE* **7**, e35003 (2012). https://doi.org:10.1371/journal.pone.0035003

47 Gai, X. *et al.* Rare structural variation of synapse and neurotransmission genes in autism. *Mol Psychiatry* **17**, 402-411 (2012). https://doi.org:10.1038/mp.2011.10

48 Hong, Q. *et al.* Prefrontal cortex Homer expression in an animal model of attention-deficit/hyperactivity disorder. *J Neurol Sci* **287**, 205-211 (2009). https://doi.org:10.1016/j.jns.2009.07.024

49 Naaijen, J. *et al.* Glutamatergic and GABAergic gene sets in attention-deficit/hyperactivity disorder: association to overlapping traits in ADHD and autism. *Translational Psychiatry* **7**, e999-e999 (2017). https://doi.org:10.1038/tp.2016.273

50 Baraban, M., Koudelka, S. & Lyons, D. A. Ca2+ activity signatures of myelin sheath formation and growth in vivo. *Nature Neuroscience* **21**, 19-23 (2018). https://doi.org:10.1038/s41593-017-0040-x

51 Gibson, E. M. *et al.* Neuronal Activity Promotes Oligodendrogenesis and Adaptive Myelination in the Mammalian Brain. *Science* **344**, 1252304 (2014). https://doi.org:10.1126/science.1252304

52 Hughes, E. G., Orthmann-Murphy, J. L., Langseth, A. J. & Bergles, D. E. Myelin remodeling through experience-dependent oligodendrogenesis in the adult somatosensory cortex. *Nature Neuroscience* **21**, 696-706 (2018). https://doi.org:10.1038/s41593-018-0121-5

53 Geraghty, A. C. *et al.* Loss of Adaptive Myelination Contributes to Methotrexate Chemotherapy-Related Cognitive Impairment. *Neuron* **103**, 250-265.e258 (2019). https://doi.org:10.1016/j.neuron.2019.04.032

54 Noori, R. *et al.* Activity-dependent myelination: A glial mechanism of oscillatory self-organization in large-scale brain networks. *Proceedings of the National Academy of Sciences* **117**, 13227-13237 (2020). https://doi.org:10.1073/pnas.1916646117

Methods

Animals

C57Bl6/J (B6) and Diversity Outbred (DO, 25th generation) male mice were purchased from The Jackson Laboratory. Collaborative Cross (CC) male mice from the CC083 and CC025 lines were purchased from the University of North Carolina at Chapel Hill. All mice were bought at six to eight weeks old, group housed three to five per cage and kept under a 12 h light–dark cycle in a temperature-controlled environment with ad libitum food and water, unless mice were food restricted for behavioral assays. All procedures were done in accordance with guidelines approved by the Institutional Animal Care and Use Committees (protocol #22087-H) at The Rockefeller University. Number of mice used for each experiment was determined based on expected variance and effect size from previous studies and no statistical method was used to predetermine sample size. DO phenotyping was performed with all males to sufficiently power the study at affordable cost, but future studies will use female-only or mixed cohorts.

Surgeries

Surgical procedures and viral injections were carried out under protocols approved by Rockefeller University IACUC and were performed in mice anesthetized with 1%–2% isoflurane using a stereotactic apparatus (Kopf) under a heating pad. Paralube vet ointment was applied on the eyes to prevent drying. Virus was injected using a 34–35 G beveled needle in a 10ul NanoFil Sub-Microliter Injection syringe (World Precision Instruments) controlled by an injection pump (Harvard Apparatus). All viruses were injected at a volume of 1 µL and a rate of 100nL/min (unless mentioned otherwise), and the needle was removed 10 min after the injection was done to prevent backflow of the virus. All injection coordinates were relative to bregma.

For adult knockdown manipulations: B6 mice were bilaterally injected at the age of 8 weeks in the PFC (A/P: 1.8 mm, M/L: ±0.3 mm, D/V: -1.75 mm) with an AAV9 expressing either a U6-scramble (non-targeting) shRNA-CMV-mCherry (titer: 9.87. x 10^{12} GC/mL) or U6-Homer1a-targeted shRNA-CMV-mCherry (titer: 4.8 x 10^{12} GC/mL) construct.

For adult overexpression manipulations: B6 mice were bilaterally injected (2 injections/hemisphere) at the age of 8 weeks in the PFC (A/P: 1.8 mm, M/L1: ±0.3 mm, M/L2: ±0.45 mm, D/V: -1.75 mm) with an AAV9 expressing either CaMKII(1.3)-eYFP (titer: 1.0 x 10^{13} GC/mL) or CaMKII(1.3)-Homer1a-eYFP (titer: 1.0 x 10^{13} GC/mL) construct at a volume 0.5 µL for each injection. pAAV.CamKII(1.3).eYFP.WPRE.hGH was a gift from Karl Deisseroth (Addgene plasmid # 105622 ; http://n2t.net/addgene:105622 ; RRID:Addgene_105622).

For developmental knockdown experiments: B6 mice were bilaterally injected in PFC at p0 pups (A/P: ~0.3 mm, M/L: ~±0.15-0.2 mm, ~-0.7-0.8 mm) and again at p11 (A/P: 0.51, M/L: ±0.17, D/V: -1.5 mm) with an AAV9 expressing either a U6-scramble (non-targeting) shRNA-EF1a-mCherry (titer: 4.8. x 10^{12} GC/mL) or pooled U6-Homer1a-targeted shRNA-EF1a-mCherry and U6-Ania3-targeted shRNA-EF1a-mCherry (titer: 2.8 x 10^{12} GC/mL) construct at a volume of 0.1 µL both times.

For multi-fiber photometry experiments: A mixture of AAV9-CaMKII(0.4)-Cre (titer: 1.0 x 10^{13}) and AAV1-Cag-Flex-JRGECO1a (titer: 1.0 x 10^{13}) was injected into PFC (A/P: 1.85 mm, M/L: 0.35 mm, D/V: -2.55 mm) at a combined volume of 1 µL. AAV9-Syn-GCaMP6f (titer: 1.4 x 10^{13} GC/mL) was injected ipsilaterally into MD (A/P: -1.6 mm, M/L: 0.45 mm, D/V: -3.2 mm). AAV(Olig001)-MAG-GCaMP6f (titer: 1 x 10^{13} GC/mL) was injected into PFC contralaterally (coordinates: A/P: 1.85 mm, M/L: -0.35 mm, D/V: -2.55 mm). AAV1-Cag-GCaMP6f (titer: 2.6 x 10^{12}) was also injected contralaterally to the initial injection (Cag-Flex-JRGECO1a) into LC (A/P: -5.4 mm, M/L: -0.85 mm, D/V: -3.6 mm). Following the completion of all injections, mice were implanted with 1.25 mm ferrule-coupled optical fibers (0.48 NA, 400 mm diameter) dorsally to the targeted structure (PFC, MD, LC). PFC and LC cannulas were implanted at a 10º angle laterally to their respective side. Cannulas were fixed to the skull with dental cement, as was a light weight titanium head plate implant, which was used to enable stable head fixation during tethering of the animal to optical patchcords each day.

Mice recovered for 5 weeks after Homer1 manipulations and 3 weeks after photometry implants before experiments began.

Behavioral protocols

Acoustic Startle Response and Prepulse inhibition

ASR/PPI testing was performed as described previously (Jin et al. 2019, Cholinergic Neurons of the Medial Septum Are Crucial for Sensorimotor Gating) Startle was measured using a San Diego Instruments SR-Lab Startle Response System. Mice were placed into Plexiglas cylinders resting on a Plexiglas platform with the chamber light on for the entire duration of the experiment. Acoustic stimuli were produced by speakers placed 33 cm above the cylinders. Piezoelectric accelerometers mounted under the cylinders transduced movements of the mice, which were digitized and stored by an interface and computer assembly. Beginning at startle stimulus onset, 65 consecutive 1 ms readings were recorded to obtain the amplitude of the mouse's startle response. For the acoustic startle sessions, the intertrial interval between stimulus presentations averaged on 15 s (range: 7–23 s). A 65 dB background was presented continuously throughout the session. Startle pulses were 40 ms in duration, prepulses were 20 ms in duration, and prepulses preceded the pulse by 100 ms (onset–onset). The Plexiglas holders were wiped clean and allowed to dry between runs. The acoustic startle sessions consisted of three blocks. Sessions began with a 5 min acclimation period followed by delivery of five startle pulses (120 dB). This block allowed startle to reach a stable level before specific testing blocks. The next block tested response threshold and included four each of five different acoustic stimulus intensities: 80, 90, 100, 110, and 120 dB (data not shown) presented in a pseudorandom order. The third block consisted of 42 trials including 12 startle pulses (120 dB) and 10 each of 3 different prepulse trials (i.e., 68, 71 and 77 dB preceding a 120 dB pulse). PPI was calculated as follows using the trials in the third block: 100 − ([(average startle of the prepulse + pulse trials)/average startle in the pulse alone trial] × 100). In all experiments, the average startle magnitude over the record window (i.e., 65 ms) was used for all data analysis.

Signal Detection Task (SDT)

Experiments were performed within a Habitest Modular Arena and controlled/recorded by Graphic State 4 software (Colbourn).

3 days prior to the experiment, mice were gradually food restricted to 85% of their body weight by providing ~2g of food per mouse per day and habituated to chocolate pellets by providing 2/3 pellets per mouse per day in their home cage. From the start of food deprivation and for the entire duration of the experiment, body weight and overall well-being were monitored by daily observation and weighting. All training and testing occur immediately before daily feeding.

The protocol is divided into multiple phases:

- **Magazine Shaping.** The box is configured to have the chocolate pellet magazine and dispenser, the white LED chamber light, speaker. The mouse enters the box with chamber light off. A reward pellet is dispensed into the magazine on a variable 7-13 sec (VI. 10) schedule and at the same time the magazine light goes on. If the mouse retrieves the pellet, the program returns to VI 10 schedule of reward delivery. Alternatively, if the mouse does not retrieve the pellet within a variable 1-5 min period, the program returns to VI 10 schedule of reward. The session ends after 20 min. *When 75% of the cohort are retrieving ≥15 pellets during the magazine shaping phase, the experiment moves to the next phase (usually 1 or 2 days).*
- **Nose Poke Shaping.** The box configuration is enriched by the nosepoke port and will stay unchanged until the end of the experiment. The mouse enters the box with chamber light off and is left to explore the box with the new element. Whenever the mouse pokes in the nosepoke port, a reward is dispensed. The session ends when the mouse receives 80 rewards or 20 min has elapsed. *When 75% of the cohort is retrieving ≥15 pellets during nosepoke shaping phase, the experiment moves to the next phase (usually ~3 days).*
- **SDT– 5 sec Cue Training.** The mouse enters the box with chamber light off. The session begins with an initial pre-cue delay period of the variable duration of 3-5 sec. If the mouse pokes during this time, the program moves to anticipatory response contingency (see below). Otherwise, it is followed by a 8 kHz pure tone auditory cue (~71 dB) that lasts for up to 5 sec. If the mouse pokes during the cue, the magazine lights up, a chocolate pellet is dispensed and the program moves to ITI contingency (see below). If, on the other hand, the mouse doesn't poke during the 5 sec cue, the cue turns off and the program moves to post-cue response period that lasts up to 5 sec. If the mouse pokes during this phase, the magazine lights up, a chocolate pellet is dispensed and the program moves to ITI contingency. If, on the other hand, the mouse doesn't poke during the post-cue response period, the program moves to time out contingency (see below).

ITI contingency: the magazine light turns off, after a 10 schedule delay, the program returns to pre-cue delay period. If, on the other hand, the mouse pokes during ITI contingency, the program goes to anticipatory response contingency.

Anticipatory response contingency: the chamber light turns on for 10 sec. If the mouse pokes during this time, the program restarts anticipatory response contingency. If, on the

other hand, it doesn't poke, the chamber light turns off and the program moves to pre-cue delay period.
Time out contingency: the chamber light turns on for 10 sec. If the mouse pokes during this time, the program moves to delayed response contingency. If it doesn't, the chamber light turns off, the trial is considered omitted and the program moves to pre-cue delay period.
Delayed response contingency: the chamber light turns on for 10 sec. If the mouse pokes during this time, the program restarts delayed response contingency. If it doesn't, the chamber light turns off, the trial is considered omitted and the program moves to pre-cue delay period. Session ends when either the mouse has reached 100 correct responses or 20 min elapses. *When 75% of the cohort is getting ≥70% trials rewarded for 2 consecutive days in SDT Training 1, the experiment moves to the next phase.*

- **SDT– 1 sec Cue Training.** This phase is exactly as the 5 sec cue training, with the only exception that the tone (cue) stays on for up to 1 sec vs 5 sec. The session ends when either the mouse has reached 100 correct responses or 20 min elapses.

Y-maze
Tests consisted of a single 5 min trial, in which the mouse was allowed to explore all three arms of a Y maze (arm dimensions: 12" x 3" x 5" in (L x W x H) while being recorded using a ceiling mounted camera under red light illumination. Mice were acclimated to the experimental site for 1 hr before all experiments. Whenever possible, the experimenter was blind to the viral condition of all mice during behavioral testing, with the exception of CC083 vs CC025 tests due to the difference in their coat color. The animal behavior was automatically tracked and analyzed by the EthoVision XT (Noldus) software for 1) total number of entries into each arm, 2) sequences of arm entries, 3) and distance moved (inch). Correct alternation (% of total number of arm entries) was defined as consecutive entries in 3 different arms. Total number of entries into each arm as well as total distance moved in the apparatus served as controls to exclude confounding factors to the memory performance, such as arm bias and/or differences in gross motor activity.

Novel object recognition task (NORT)
This test began with 2 days of habituation where the mice are allowed to explore an empty square arena (16'' x 16'' x 14'' (L x W x H) for 15 minutes. During training (day 3), mice are re-introduced in the arena where are now present two identical objects positioned in the back left and right corners of the cage. Each animal was placed in the middle point of the wall opposite to the objects and allowed to explore them for 15 min. At the end of the training phase, mice return to their home cage for 15 min, while the box and the objects were cleaned with 10% Ethanol and then water. Successively, the mice were re-entered into the arena for the test, during which one of the two (familiar) objects was replaced with a new one (novel), totally different in color, texture and shape. Each mouse was left free to explore the objects for 5 min. The entire experiment was recorded using a ceiling mounted camera and the animal behavior was automatically tracked and analyzed by the EthoVision XT (Noldus) software. Two measures were considered: 1) the total exploration time (sec) spent by the animal interacting with the two familiar objects during the training,

in order to evaluate objects bias and 2) the exploration time spent by the animal interacting with the novel object over the total exploration time (e.g., [novel/(familiar + novel)] × 100) during the test. Object exploration time is defined as time during which the mouse nose was in contact with the object or directed toward it at a distance ≤ 2 cm.

Elevated plus maze (EPM)

This test is commonly used to evaluate anxiety-like behavior in rodents (Lister, 1987). The apparatus was composed of four black plastic arms, arranged as a cross, located 55 cm above the plane of a laboratory bench and illuminated by a 60 w lamp located above the apparatus. Two close arms, opposite to each other were enclosed by lateral walls (50 × 6 × 40 cm), whereas the other two open arms were without walls (50 × 6 × 0.75 cm); the close and open arms delimited a small squared area (6 × 6 cm) called center. Each mouse was placed into the center of the maze, facing one of the two open arms, and its behavior was video-recorded for 5 min and automatically analyzed by the EthoVision XT software (Noldus) for the time spent by the mice in each of the three compartments (open, close, center), which was measured by an observer blind to the experimental groups.

3-Chamber Social Interaction

Tests used a (18'' x 18'' x 12'' (L x W x H) clear acrylic arena, which was divided into 3 chambers of equal area (18" x 6" x 12" (L x W x H) that were separated by walls 6" in length on each side, so that there was a 6" long separation in each wall that a mouse could pass through. Mice were habituated to the testing area for 1 hr prior to the start for the experiment. The test began with a 5 min habituation phase to the center chamber, in which the openings in the walls were obstructed so that the mice could not see or enter either opening. Mice were then put in a transfer cage for 1 minute as the center walls were opened, after which the mice were returned to the center chamber for a 5 min habituation phase to all 3 chambers of the arena. Mice were then returned to the transfer cage for 5 min and the arena as wiped down with 10% ethanol, and wire cups were placed upside down in the center of the outer 2 chambers either with a non-social stimulus (foam figurine) or a novel, age- and strain-matched mouse underneath. Mice were then placed back in the center chamber and allowed to explore for 15 min. Behavior was video-recorded and automatically analyzed by the EthoVision XT Software (Noldus) for time spent in each chamber, time spent in a 3 cm proximity to the social or the non-social stimulus, which was measured by an observer blind to the experimental groups.

QTL mapping in Diversity Outbred Mice

Genotype Identification & Haplotype Reconstruction

SNP locations and genotypes for the eight founder strains were acquired from ftp.jax.org/MUGA and the consensus genotype for each founder strain and each SNP was determined from the multiple individuals that were genotyped. SNP genotypes for the 182 DO mice were determined using a high density mouse universal genotyping array, GigaMUGA (geneSeek). A total of 114,184 SNPs were detected on the 19 autosomes and X chromosomes. Using R/qtl2 (Broman et al., 2019), founder haplotype probabilities were reconstructed for all samples and then converted to additive allelic dosages and scaled to 1. Realized genetic relationship matrices, often referred to as kinship matrices, were estimated using the leave one chromosome out (LOCO) method,

so that the kinship term does not absorb variation explained by the putative QTL. Another QTL mapping software package for multi-parental populations (MPP), miQTL, was used to confirm findings from R/qtl2, and to visualize and assess the level of heterozygosity at the locus of interest.

QTL Mapping

Phenotype values from the prepulse inhibition performance were subject to Box-Cox transformation. Then, using R/qtl2, an additive single locus linear mixed model was fit at positions across the genome, producing a genome scan. Potential population. structure was controlled for through the inclusion of a random effect to account for correlation structure measured by the kinship matrix. This was performed in R/qtl2 using the leave one chromosome out (LOCO) method (Kang et al., 2010). For confirmation of the QTL results, we performed a multiple imputation genome scan (11 imputations) using miQTL, to assess whether uncertainty in founder haplotype reconstruction was strongly influencing the results. Genome-wide significance thresholds (alpha = 0.05) for the genome scans were determined through 1000 permutations of the diplotype.

Analysis of Founder Contributions

To determine the founder haplotype effects driving the Chr13QTL, we first estimated best linear unbiased predictors (BLUPs), which constrain potentially unstable effects by fitting the QTL term as a random effect. To further confirm these results, we used Diploffect, to estimate posterior credible intervals for the haplotype effects as well as the proportion of variance explained by the QTL (sometimes referred to as the locus heritability).

RNA Expression Analysis

RNA extraction from brain tissues

For tissue extraction, mice were sacrificed by cervical dislocation and immediately decapitated in compliance with IACUC protocol # 22087-H. The targeted brain regions were harvested from 1 mm brain slices, obtained by brain matrices (ZIVIC) using 1.0 mm tissue punches and transferred to a tube containing 300 mL of ice-cold lysis buffer and 3 mL b-mercaptoethanol (Total RNA Purification kit, NORGEN; following the manufacturer's protocol). Samples were then homogenized by passing a 25G insulin syringe six times and left on ice. For RNA extraction, the total RNA Purification kit was used according to the manufacturer's instructions (NORGEN). RNA quality was evaluated by Bioanalyzer 2100 (Eukaryote Total RNA Nano chip, Agilent) at the Rockefeller University Genomic Resource Center (RIN > = 7.50 and free of genomic DNA contamination). RNA samples were then aliquoted and stored at 80ºC.

RNA sequencing (RNA-seq) analysis

For RNAseq, RNA libraries were prepared from 100ng of total RNA per sample for 6 DO mice, 3 brain regions per mouse using the TruSeq stranded mRNA LT kit (Cat# RS-122-2101, Illumina). These synthetic RNAs cover a range of concentrations, length, and GC content for validation of the fidelity and dose-response of the library prior to downstream procedures. Libraries prepared with unique barcodes were pooled at equal molar ratios following manufacturer's protocol (Cat# 15035786 v02, Illumina). The pool was denatured and subject to paired-end 50x sequencing on the Hi-Seq 2500 platform. An average of

67 million reads per sample were obtained. Sequencing reads were aligned to the mouse genome (mm10) using STAR (v2.4.2a) and aligned reads were quantified using Salmon (v0.8.2). Approximately 90% of the reads mapped uniquely. Hierarchical clustering and Principal Components Analysis were performed following Variance Stabilizing Transformation (VST) from DESeq2, which is on the log2 scale and accounts for library size differences. The hierarchical clustering heatmap shows the Euclidean distances of VST of the counts between samples.

Quantitative PCR (qPCR)

For quantitative PCR, each reverse transcription was performed with 0.2 mg RNA using the High-Capacity RNA-to-cDNA kit (Applied Biosystems #4387406), in a final volume of 20 µL. Primers for reverse transcription were equal mixtures of poly-T nucleotides and random hexamers. Negative controls (omitting reverse transcriptase enzyme) were performed for each sample. The cDNA products were diluted 1:1 and 2 µL was analyzed by qPCR using custom primer sets and PowerUp SYBR Green Master Mix (10 µL total reaction, Applied Biosystems. #A25742). RT-qPCRs were performed on a Quantstudio3 from Applied. Biosystems. Every reaction was systematically run in triplicate. Conditions were the following: 50ºC 2 min, 95ºC 10 min, 40 x (95ºC 15 s, 60ºC 1 min). qPCR Ct values were analyzed using the LightCycler software. Detection threshold was set at DRn = 0.3, with this limit always within the 2n exponential amplification phase of genes. Mean of technical triplicate values were reported. All mice gene expression Ct values were normalized with the reference gene Ube2d2a using dCt method to determine the relative mRNA expression of each gene.

Single-Cell Sequencing

Single-cell dissociation and Single-Cell RNA-sequencing

Single cell suspensions of prefrontal cortex were prepared as described previously (Zeissel et al 2018). Briefly, mice were sacrificed with an overdose of isoflurane, followed by transcardial perfusion with carbogenated (95% O_2, 5% CO_2) Hanks' Balanced Salt Solution (HBSS). Brains were removed, 500µm sections were collected, and the prefrontal cortex region was isolated. The tissue was dissociated using papain (LS003124, Worthington) dissolved in Hibernate A buffer (NC1787837, Fisher Scientific) and incubated for 25-30 min at 37º C, followed by manual trituration using fire polished Pasteur pipettes and filtering through a 40µm cell strainer (BAH136800040, Millipore Sigma). Cells were washed with wash buffer (PBS + 1% BSA) and centrifuged at 200 g for 5 min, the supernatant was carefully removed, and cells were resuspended in ~500ul wash buffer and 10% DAPI. Flow cytometry was done using a BD FACS Aria III Cell Sorter (BD FACSDiva Software, version 8.0.1) with a 100-µm nozzle. The cell suspensions were first gated on forward scatter, then within this population based on lack of DAPI staining. Cells were collected in wash buffer, manually counted using a Burker chamber, and suspension volumes were adjusted to a target concentration of 700 -1000 cells/µl. Single cell RNA-sequencing was carried out with the 10X Genomics Chromium Single Cell Kit Version 3. Manufacturer's instructions were followed for downstream cDNA synthesis (12-14 PCR cycles), library preparation, and sequencing.

Single-Cell RNA sequencing data analysis

RAW sequencing reads were aligned to the GRCm38/mm10 mouse reference genome and a custom SEQC code was used to generate a per-cell count matrix. The Python Scanpy package (version 1.9.3) was used to further analyze the data. Replicates were merged and doublets were removed using Scrublet (Wolock et al. 2019). Cells with <2,500 UMIs per cell and <1,000 genes per cell, and genes detected in <3 cells were removed. Per-cell counts were normalized to equal the median of total counts per cell, and log-transformed. Principal component analysis was used to reduce the dimensionality to 50 principal components. A nearest neighbor graph was computed between cells using these principal components, and Leiden clustering was applied to separate the cells into distinct clusters of major cell types. Known gene markers were used (Zeissel et al. 2018) to assign cell types. Once the neuronal cluster was identified, it was subsetted and re-clustered using the first 50 principal components to identify inhibitory and excitatory neurons. Clusters with differential *Homer1* expression between cc083 and cc025 strains were identified using t-tests. Clusters with significantly different *Homer1* expression between strains were merged, and the "MAST" R package (Finak et al. 2015) was used to identify differentially expressed genes between strains for the merged cluster as well as all individual clusters. Gene set enrichment analysis was performed using the fast gene set enrichment analysis (GSEA) package (fgsea v1.18.0), the GO_Molecular_Function_2021 gene set, and the Elsevier_Pathway_Collection gene-set libraries using Enrichr (Chen et al. 2013; Kuleshov et al. 2016; Xie et al. 2021).

shRNA Experiments In Vitro & In Vivo

We used the following shRNAs for gene knockdown (which were then subcloned into a pscAV-U6-mCherry construct, VectorBuilder/VigeneBiosciences):
Homer1a (GenBank: NM_011982.4), Targeting sequence: GGTTTCAGAAACTCTTGAA; Ania3 (GenBank: NM_001347598.1),Targeting sequences: GGAGACATAGTTCTTCTTA, GCTAAGCTAGAGCCATCTA.

For gene expression, coding sequences of Homer1a and Ania3 was cloned from mouse cortical cDNA and subsequently subcloned into a pAAV.CamKII(1.3).eYFP.WPRE.hGH expression vector using standard molecular cloning techniques. Constructs were verified first by Sanger sequencing, and then diagnostics for ITR integrity, by digestion with SmaI, before AAV production.

Generation of AAV-MAG-GCaMP6f

We identified the mouse *MAG* gene locus (including introns and a 3 kb upstream potential promoter region) using the UCSC genome browser on the reverse strand of Chr 7qB1: 30,899,176-30,917,832 in the July 2007 mm9 alignment (Chromosome 7: 30,598,601-30,617,298 in the GRC38/mm10 alignment). Sequence conservation was assessed using the VISTA genome browser and the putative MAG promoter was screened for regions of >50% interspecies sequence similarity, which were then evaluated for transcription factor binding sites, especially OL-lineage specific Olig1 and Olig2, using the Wilmer Bioinformatics Resource and the P-match 1.0 program (http://gene-regulation.com/pub/programs.html#pmatch). This method yielded a 2.5 Kb putative MAG promoter region. The putative MAG promoter was cloned from mouse cortical cDNA

using standard molecular cloning techniques and replaced the Syn promoter from the pAAV.Syn.GCaMP6f.WPRE.SV40 plasmid. The pAAV-MAG-GCaMP6f construct was packaged using the Olig001 capsid, which has high oligodendrocyte tropism. (University of Arizona Viral Production Core).

Histology & Immunohistochemistry

Mice were transcardially perfused with PBS and 4% paraformaldehyde in 0.1M PB, then brains were post-fixed by immersion for 24 h in the perfusate solution followed by 30% sucrose in 0.1M PB at 4ºC. The fixed tissue was cut into 40 mm coronal sections using a freezing microtome (Leica SM2010R), stained with DAPI (1:1000 in PBST), and mounted on slides with ProLong Diamond Antifade Mountant (Invitrogen). For immunostaining, the fixed sections were permeabilized with 70% methanol for 15 min before blocking with 5%normal donkey serum in PBS for 1 h and incubated with primary antibodies overnight at 4º C. Sections were washed three times in PBS and incubated with appropriate secondary antibodies overnight at 4º C. Afterward, coverslips were mounted using ProLong Diamond Antifade mounting medium for image collection. Primary and secondary antibodies include rabbit polyclonal anti-NeuN (Millipore AN78), rabbit anti-Iba1 (Wako, 019-19741), rabbit polyclonal anti-Olig2 (Millipore, AB9610), and mouse monoclonal anti-GFAP (Millipore MAB360), Alexa Fluor 647 donkey anti-rabbit IgG (Jackson ImmunoResearch, Cat# 711-606-152), Alexa Fluor 647 donkey anti-mouse IgG (Jackson ImmunoResearch, Cat# 715-606-151), and DAPI (Cayman Chemical, Cat#28718-90-3). For immunohistochemistry staining, epifluorescent images were obtained at room temperature on a Nikon Eclipse T*i* microscope using a Nikon 4x (NA 0.13, dry), 10x (NA 0.30, dry), or 20x (NA 0.45, dry), objectives with the same settings and configurations for each objective across all samples within each experiment.

In Vivo Multi Site Photometry Recordings

Photometry Setup

A custom dual-color, multi-fiber photometry setup was built. For GCaMP6f imaging, excitation of the 470 nm (imaging) and 405 nm (isosbestic control) wavelengths were provided by LEDs (Thorlabs M470F3, M405FP1), which were collimated into a dichroic mirror holder with a 425 nm long pass filter (Thorlabs DMLP425R). This was coupled to another dichroic mirror holder with a 495 nm long pass dichroic (Semrock FF495-Di02-25x36) which redirected the excitation light on to a custom branching low-autofluorescence fiberoptic patchcord of three bundled 400 mm diameter 0.57NA fibers (BFP(3)_400/440/PKMJ-0.57_1m_SMA-3xFC_LAF, Doric Lenses) using a 20x/0.5NA Objective lens (Nikon CFI SFluor 20X, Product No. MRF00100). GCaMP6f fluorescence from neurons below the fiber tip in the brain was transmitted via this same cable back to the mini-cube, where it was passed through a GFP emission filter (Semrock FF01-520/35-25), amplified, and focused onto a high sensitivity sCMOS camera (Prime 95b, Photometrics). For JRGECO1a imaging, a second light path was built so that excitation of the 565 nm (imaging) and 470 nm (isosbestic control) wavelengths were provided by LEDs (Thorlabs M565F3^{h}, M470F3), which were collimated into a dichroic mirror holder with a 505 nm long pass dichroic (Thorlabs DMLP505R). This was coupled to another

dichroic mirror holder with a 573 nm long pass dichroic (Semrock Di02-R561-25x36), which redirected the excitation light on to a low-autofluorescence monofiberoptic patchcord with a 400 mm diameter 0.57NA fiber (MFP_400/440/PKMJ-0.57_1m_SMA-FC_LAF, Doric Lenses) using a 20x/0.5NA Objective lens (Nikon CFI SFluor 20X, Product No. MRF00100). JRGECO1a fluorescence from neurons below the fiber tip in the brain was transmitted via this same cable back to the mini-cube, where it was passed through a RFP emission filter (Semrock FF01-607/36-25), amplified, and focused onto a high sensitivity CMOS camera (BFS-PGE-50S5M-C, Teledyne FLIR).
Each of the multiple branch ends of the branching fiberoptic patchcord as well as the monofiberoptic patchchord were coupled to four 2 m low-autofluorescence patchcords (MFP_400/430/1100-0.57_2m_FCZF1.25_LAF, Doric Lenses) which is used to collect emission fluorescence from 1.25mm diameter light weight ferrules (MFC_400/430-0.48_ZF1.25, Doric Lenses) using a mating sleeve (SLEEVE_BR_1.25, Doric Lenses). An microcontroller (Arduino Uno) was programmed to take trigger inputs from the Operant Behavior Setup or MATLAB and synchronize the camera shutters and alternate triggering of the 405 nm and 565 nm LEDs together and both 470 nm LEDs together. Custom TTL converters were used to read in frame acquisition times to the Habitest Modular system (above), which were integrated with events from the behavior in Graphic State 4. Bulk activity signals were collected using the PVCAM (GCaMP) and Spinnaker (JRGECO) software, and data were further post-processed and analyzed using custom MATLAB scripts.

QUANTIFICATION AND STATISTICAL ANALYSIS

Behavior Statistical Reporting

Sample sizes were selected based on expected variance and effect sizes from the existing literature, and no statistical methods were used to determine sample size a priori. Prior to experiments being performed, mice were randomly assigned to experimental or control groups. The investigator was blinded to all behavioral studies (except for CC083 versus CC025 cohorts, where coat color differences prevent blinding during experimentation). Data analyses for calcium imaging were automated using MATLAB scripts. Statistical tests were performed in MATLAB 2022b or Graphpad Prism 9.

Gene Expression Statistics

Differential gene expression between high and low performing DO mice was determined in R (3.5.0) using the DESeq2. P values were determined using a Wald test and p values were corrected using the Benjamini-Hochberg (BH) method.

Multi-Fiber Photometry Data Processing

For analysis, the images captured by the sCMOS camera were post-processed using custom MATLAB scripts. Regions of interest were manually drawn for each fiber to extract fluorescence values throughout the experiment. The 405 nm (GCaMP) or 470 nm (JRGECO) reference traces were scaled to best fit the 470 nm (GCaMP) or 565. Nm. (JRGECO) signal using least-squares regression. The normalized change in fluorescence (dF/F) was calculated by subtracting the scaled 405 nm or 470. nm reference traces from the 470 nm or 565 nm signals, respectively, and dividing those

values by the scaled reference traces. The true baseline of each dF/F trace was determined and corrected by using the MATLAB function ‘‘msbackadj’’ estimating the baseline over a 200 frame sliding window, regressing varying baseline values to the window’s data points using a spline approximation, then adjusting the baseline in the peak range of the dF/F signal. Task events (for example, cue on/offsets, and nosepokes), were time stamped via the Graphic State 4 software,

Multi-Fiber Photometry Data Analysis

Total mean activity, for different task phases, and different strains, were quantified as area under the curve (AUC) of dF/F responses. To facilitate comparison across mice, dF/F responses were z-scored and shifted above 0. AUC was calculated using MATLAB “trapz” function and normalized with the recorded time. Local maximum was calculated using the “max” (MATLAB) function. Pearson Correlation of the dF/F responses was performed between different regions using the ‘‘corr’’ (MATLAB) function. To ensure that correlation values were significantly more than chance, each timeseries was scrambled 10,000 times randomly, for each trial across all mice. All such chance correlation coefficients were pooled to calculate mean and standard deviation of chance correlations.

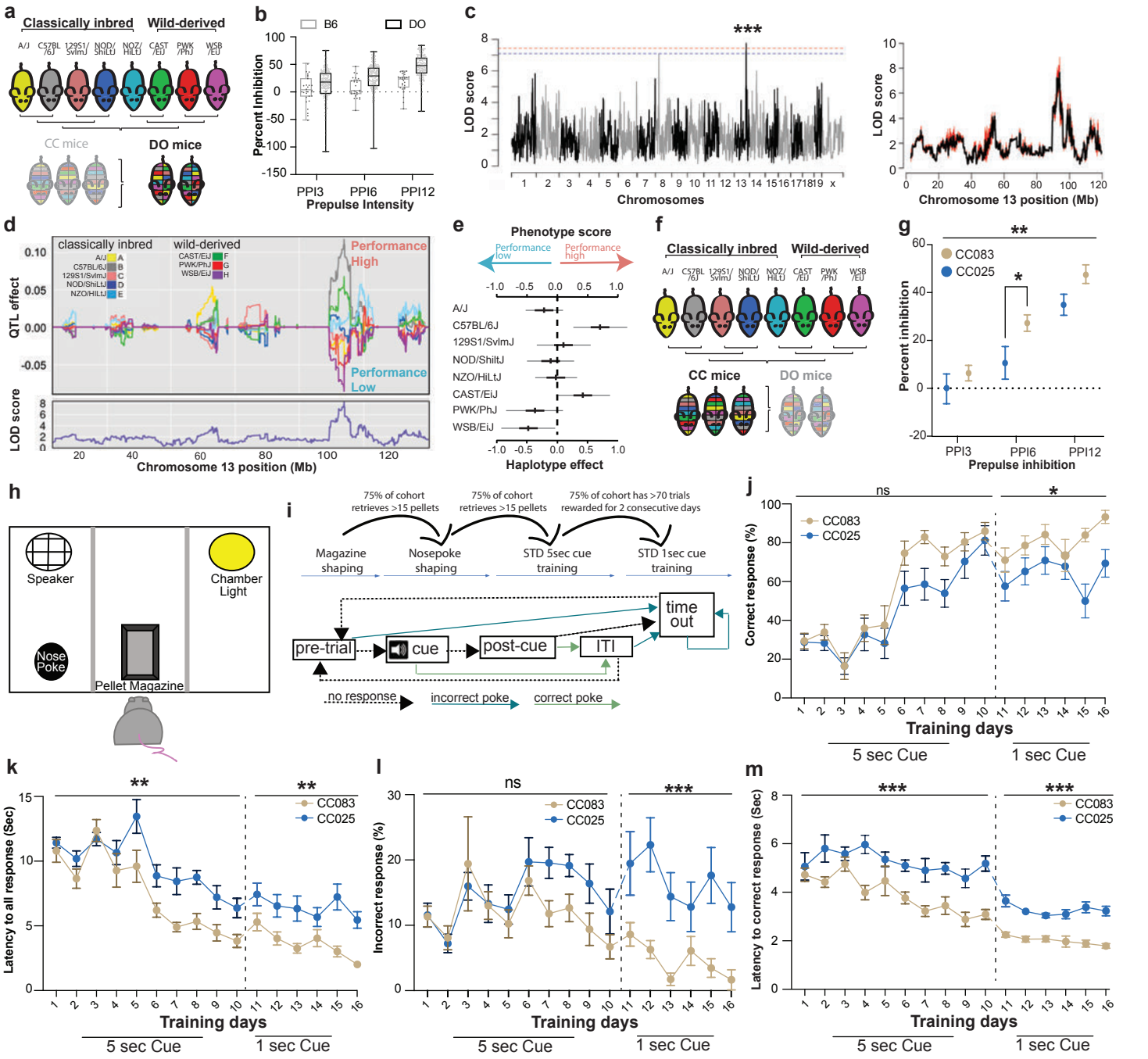


Fig 1. Identification of a QTL Associated with Pre-Attentive Processing

a) Outbreeding scheme to generate the DO mice. **b)** Sensory gating performance (as measured by PPI) in B6 (n=27) and DO (n=176) mice measured as percent inhibition at 3 different prepulse intensities: 3, 6, and 12 dB above background (PPI3, 6, and 12, respectively; Methods). Boxes indicate 2nd and 3rd quartiles with median and range. **c)** Left: QTL analysis (by miQTL) for Prepulse Inhibition. Significance thresholds after 50 imputations of genotype, blue: 90%, red: 95%. Right: Mapping analyses performed using both R/qtl2 (black) and miQTL (red) revealing minimal fluctuation in LOD score across imputations (overlapping bands). **d)** Effect of each founder allele on PPI performance along chromosome 13 (Chr 13) (x axis), as measured by the founder coefficients from the linkage model (y axis). Coefficients diverge substantially at the peak QTL. Logarithm of odds (LOD) score at each chromosomal position shown. **e)** Haplotype representation at the Chr 13 locus and corresponding z-scored phenotypes of each founder strain, quantified as mean ± SEM. **f)** Outbreeding scheme to generate the CC mice. **g)** PPI3, 6, and 12 values for CC083 (n=27) and CC025 (n=24) mice. Data are mean ± SEM. Two-way ANOVA $F(1,147)=9.40$ $p=0.003$ for strain main effect followed by Sidak's test for multiple comparison $p(PPI6)=0.05$. **h)** Cartoon of the operant wall of the arena used for the auditory signal detection task (aSDT), showing the position of the speaker, chamber light, nose poke and pellet magazine. **i)** Schematic of the aSDT protocol. **j-m)** Performance of CC083 (n=10 for 5 sec cue and 1 sec cue) and CC025 (n=10 for 5 sec cue and n=9 for 1 sec cue) mice during aSDT across days, showing **(j)** percent correct, **(k)** mean latency of response, **(l)** percent of incorrect responses (not including omissions), and **(m)** mean latency to correct response. Significant differences between strains reported as *p<0.05, **p<0.01, and ***p<0.001 by two-way ANOVA with repeated measures and post-hoc Holm-Sidak test for multiple comparisons.

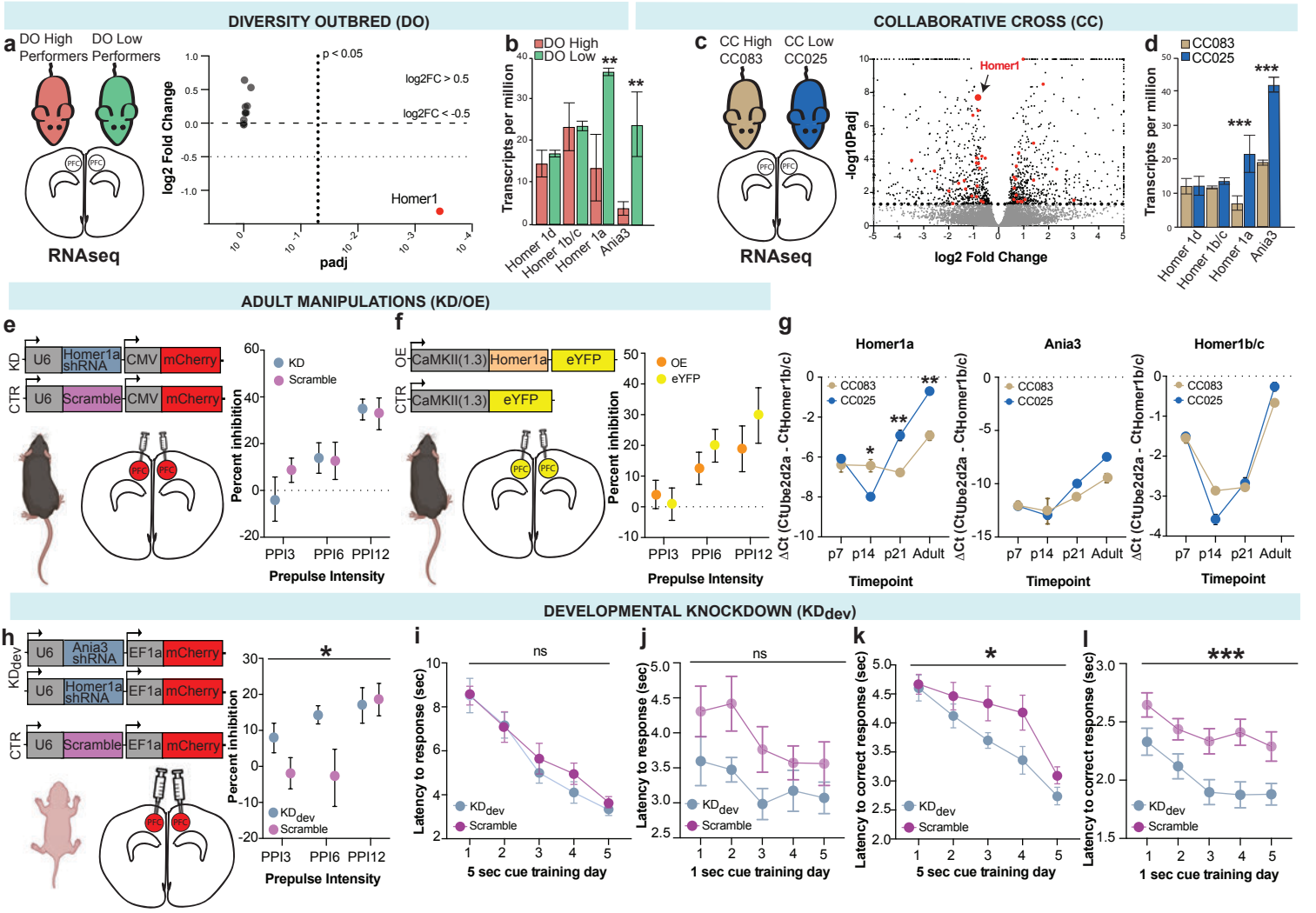


Fig 2. Chr 13 QTL effects map to Homer1a as a developmental modifier of attentional processing

a) Volcano plots of differential expression between DO high and low performers for all locus genes (n=3 biological replicates per strain). Dashed lines indicate significance thresholds (adjusted p<0.05 and log2FC>0.5 or <-0.5). Only Homer1 crosses both thresholds (red). **b)** Expression levels of Homer1 isoforms in PFC in DO high- (pink) and low- (green) performers (n=3 biologically independent samples per performance group, data are mean ± SEM, significant differential expression of Homer1a and Ania3, p<0.01 for Homer1a and Ania3, by two-way ANOVA with post hoc Holm-Sidak test. **c)** PFC RNAseq from CC high (CC083, gold) and low (CC025, blue) performers visualized as volcano plots showing the significance and -log-transformed adjusted p-value distribution after DESeq2 differential gene expression analysis. Red dots indicate genes with gene ontologies relating to Homer1's function in excitatory neurotransmission as well as postsynaptic structure and activity. **d)** PFC qPCR from CC high and low performers assessing differential expression of Homer1 isoforms (n=3, data are mean ± SEM, p<0.001 by two-way ANOVA with post hoc Holm-Sidak test for multiple comparisons. **e)** Left panel: Schematic of constructs and injection location (PFC) for the knockdown in adult B6 mice. Right panel: Prepulse Inhibition (PPI) in knock down (KD, n=15) and control (Scramble, n=15) mice measured as percent inhibition at 3 different prepulse intensities: 3, 6, and 12 dB above background (PPI3, 6, and 12, respectively). Data are mean ± SEM. **f)** Left panel: Schematic of constructs and injection location (PFC) for overexpression in adult B6 mice. Right panel: Prepulse Inhibition in overexpressed (OE, n=11) and control (eYFP, n=10) mice measured for PPI3, 6, and 12. Data are mean ± SEM. **g)** PFC expression of Homer1a, Ania3, and Homer1b/c in CC083 and CC025 mice at p7, p14, p21, and in adulthood by qPCR (n = 3 biological replicates per strain per timepoint, significant differences between strains for Homer1a by two-way ANOVA with post hoc Holm-Sidak test, p = 0.02 at p14, p=0.002 at p21, and p<0.001 at adult; and for Ania3 p=0.002 at adult. **h)** Left panel: Schematic of constructs and injection location (PFC) for knockdown in neonatal B6 mice (p0-2). Right panel: PPI in developmental knock down (KDdev, n=15) and controls (Scramble, n=13) Data are mean ± SEM, significant differences between injection groups, p<0.01, by two-way ANOVA. **i-l)** Performance of KDdev and Scamble mice during the aSDT across 5 sec cue (n=13 per group) and 1 sec cue (n(KDdev)=11, n(Scramble)=10). Performance shown as latency to all responses during **i)** 5 sec cue training, **j)** 1 sec cue training and **k)** latency to correct response during 5 sec cue and **l)** 1 sec cue training. Significant differences between strains measure by repeated-measures two-way ANOVA, *p<0.05 and ***p<0.001.

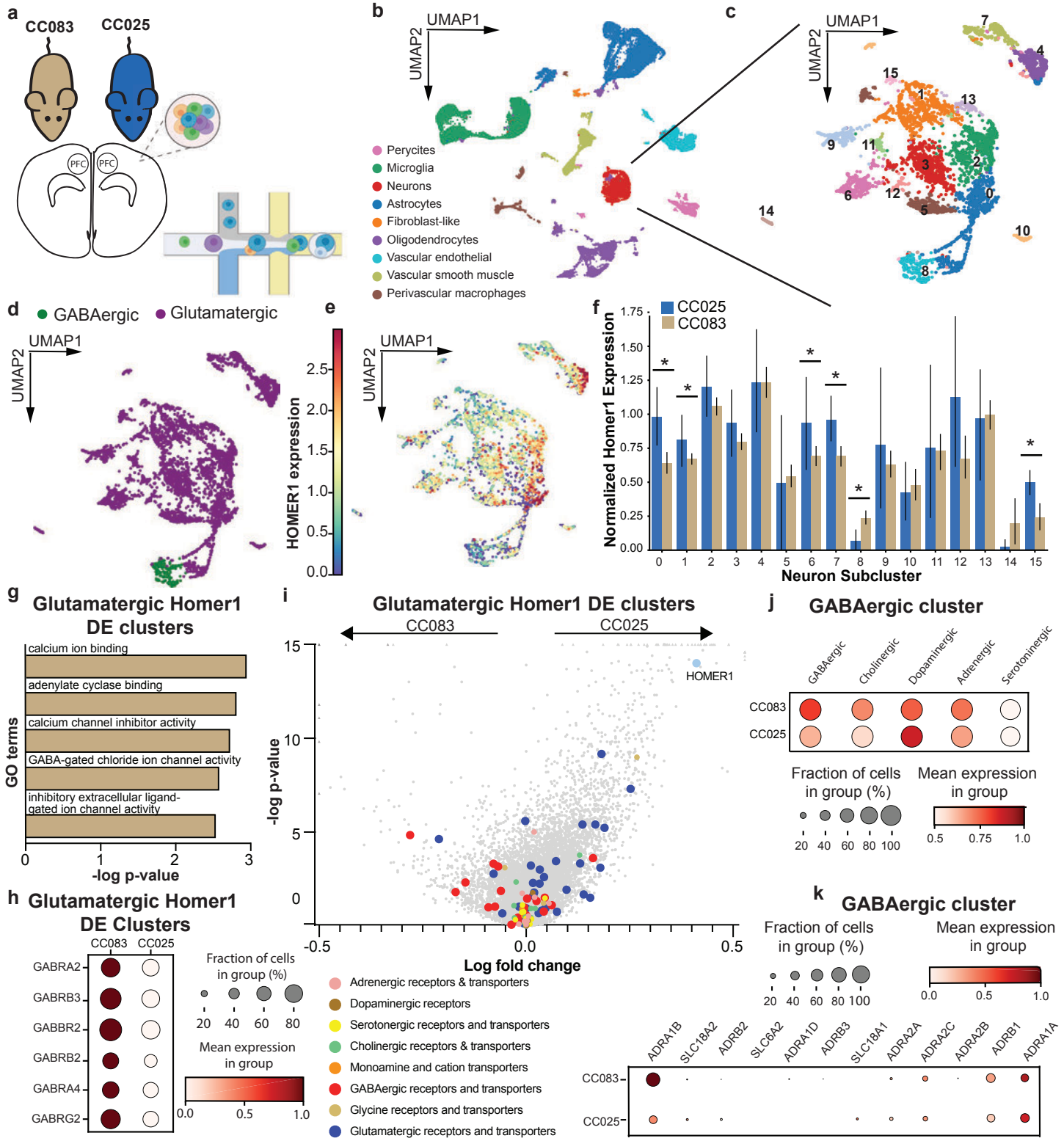


Fig. 3 Low Homer1-Expressing Glutamatergic Neurons Upregulate GABA Receptors

a) Schematic representation of scRNAseq workflow. **b)** UMAP visualization of all cells collected (n = 70,90 cells) collected from CC025 (n=6) and CC083 (n=6) mice clustered based on transcriptional profile. **c)** UMAP visualization sub-clustering all cells identified as neurons (n = 4633 cells). **d)** Distribution of excitatory (glutamatergic) and inhibitory (GABAergic) neurons identified by expression of canonical marker genes. **e)** Scaled Homer1 expression (by MAST) in neuronal clusters in UMAP plot (from c). **f)** Differential Homer1 expression between CC083 and CC025 neurons by cluster by MAST analysis. Homer1 was differentially expressed between CC025 and CC083 in glutamatergic clusters 0, 1, 6, 7, and 15 as well as GABAergic cluster 8 ($p < 0.1$). Data is shown as mean ± SD. **g)** Gene ontology (GO) analysis of molecular function for genes upregulated in glutamatergic neurons with differential Homer1 expression showed that CC083 excitatory neurons have increased transcription of genes associated with inhibitory GABA ion channel activity. **h)** Dot plot showing scaled expression of GABAergic receptors driving GO analysis (from g). The size of each dot corresponds to the percentage of cells in the merged cluster expressing each gene and color intensity indicates relative, scaled expression of that gene. **i)** Volcano plot depicting differential gene expression in the merged Homer1 DE cluster between CC025 and CC083. Positive change indicates higher expression in CC025 excitatory neurons relative to CC083, cells. Colored dots indicate genes encoding receptors and transporters of the most common neurotransmitter systems. **j)** Dot plot of markers for common neuromodulatory systems in GABAergic cluster 8 for cells from CC025 and CC083 mice. The size of each dot corresponds to the percentage of each strain's cluster 8 cells expressing marker gene sets from each neuromodulatory system and the color intensity indicates the relative, scaled expression of the marker gene set. **k)** Dot plot of adrenergic receptors and transporters between CC025 and CC083 cells in GABAergic cluster 8. The size of each dot corresponds to the percentage of each strain's cluster 8 cells expressing each gene and the color intensity indicates the relative, scaled expression of the gene.

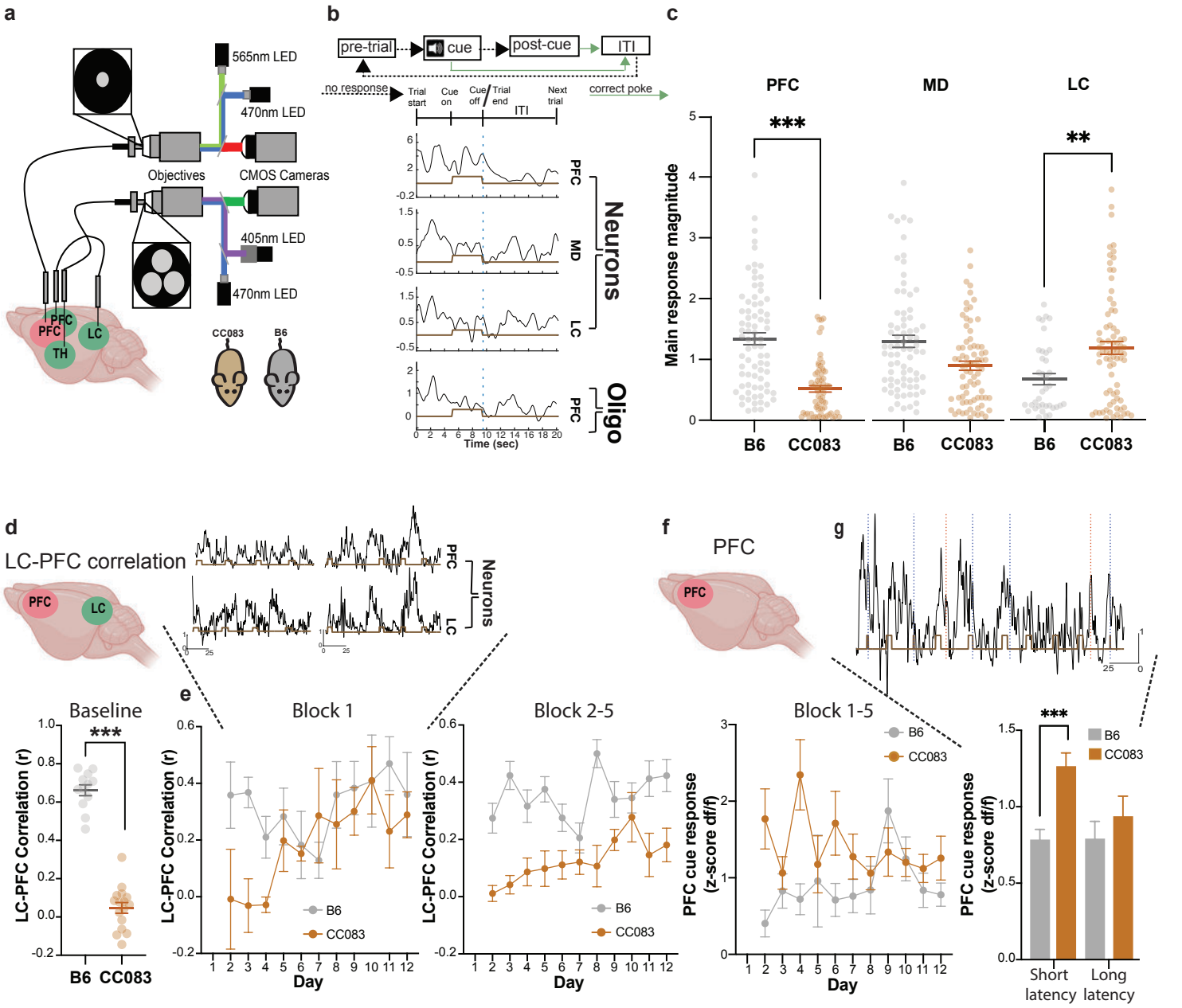
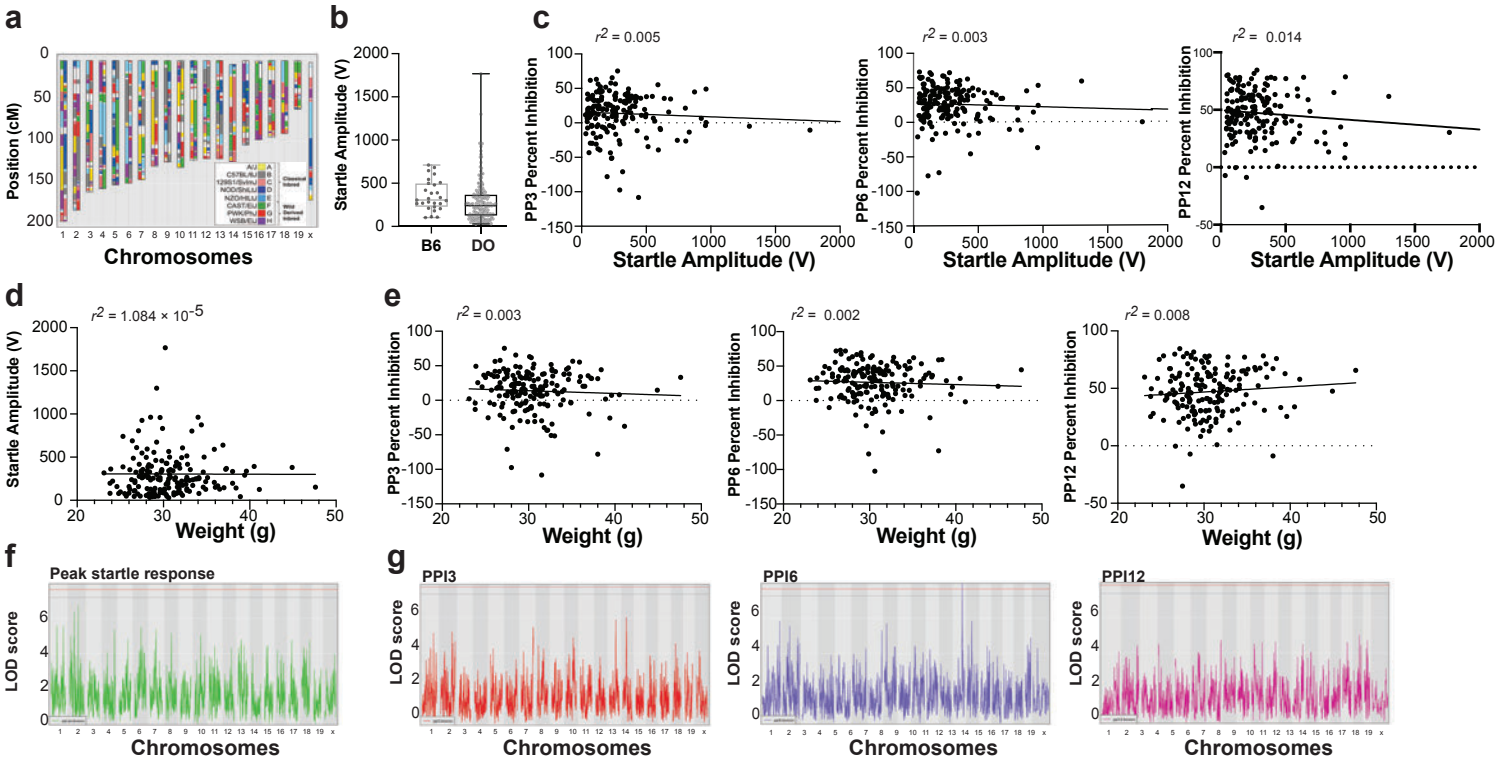
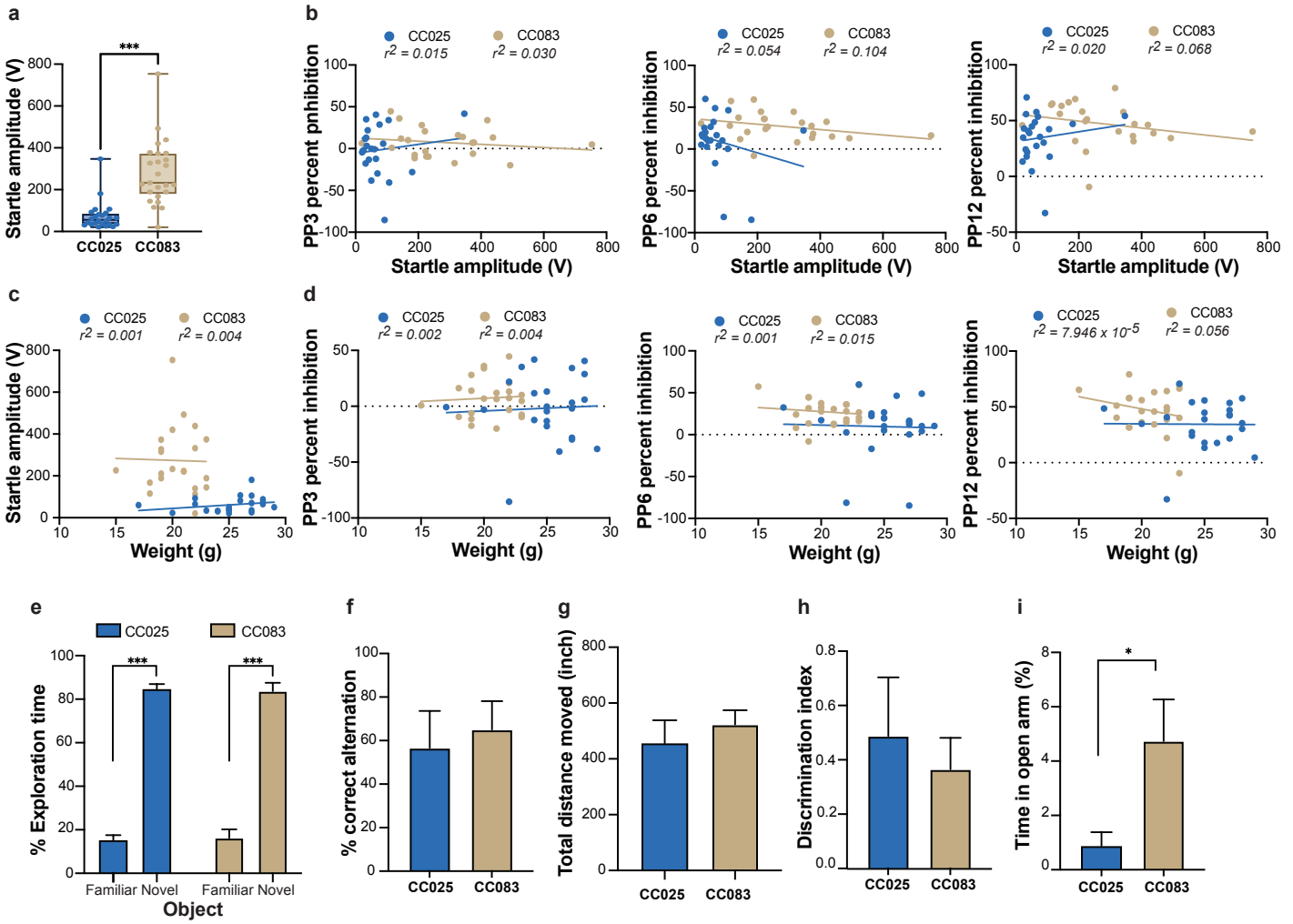


Fig 4. Enhanced dynamic range of LC-PFC coupling and of pre-cue PFC responses support attention

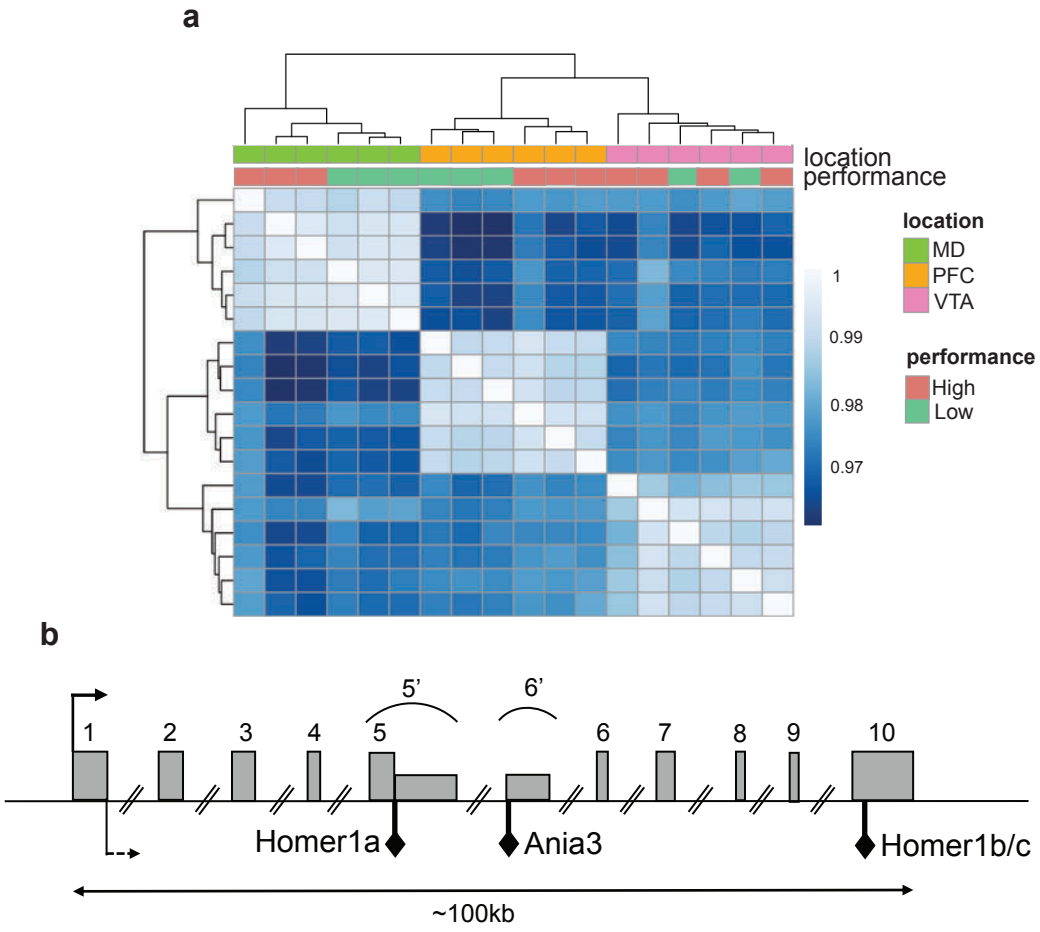
a) Schematic of dual-color, 4-region photometry system. Simultaneous 565 nm, 470 nm, and 405 nm recordings were taken from PFC neurons (red), MD neurons (green), LC neurons (green), and PFC oligodendrocytes (green) in B6 or CC083 strains. **b)** Top: Schematic representation of aSDT trial structure. Bottom: example jRGECO1a (PFC neurons, row 1) and GcaMP6f (MD neurons, LC neurons, and PFC oligodendrocytes (rows 2-4, respectively) traces aligned to an aSDT trial. **c)** Average activity (area under responses) in home cage for B6 (n = 5) versus CC083 (n = 4) during 5 minute recordings. **d)** Pairwise Pearson's correlations between LC and PFC neuronal activity in home cage. Welch-corrected t-test between (B6 versus CC083, n = 4 ea., 5 minute recordings, p<0.001). **e)** Top: representative traces from PFC neurons (top) and LC neurons (bottom) from day 2 (left) and day 11 (right), where the Y-axis is z-scored df/f and the X-axis is time measured in seconds. Yellow rectangles indicate cues. Bottom: Pairwise Pearson's correlations between LC and PFC neuronal during the aSDT session. Each 20 minute session was split into 5, 4-minute blocks. Data is shown from the first 4-minute block (left) and blocks 2-5 (right) as mean ± SEM. **f)** Maximum PFC magnitudes 5 seconds before cue onset throughout the entire aSDT session across days in B6 (n = 5) and CC083 (n = 4) mice. Data is shown as mean ± SEM. **g)** Top: Representative trace from PFC where the Y-axis is z-scored df/f and the X-axis is time measured in seconds. Yellow rectangles indicate cues, orange dotted lines indicate delayed responses, and blue dotted lines indicate correct responses. Bottom: Maximum PFC magnitudes for B6 (n = 5) and CC083 (n = 4) mice on trials separated by short and long response latencies. Two-way ANOVA, followed by Holm-Sidak test for multiple comparisons (p<0.001).



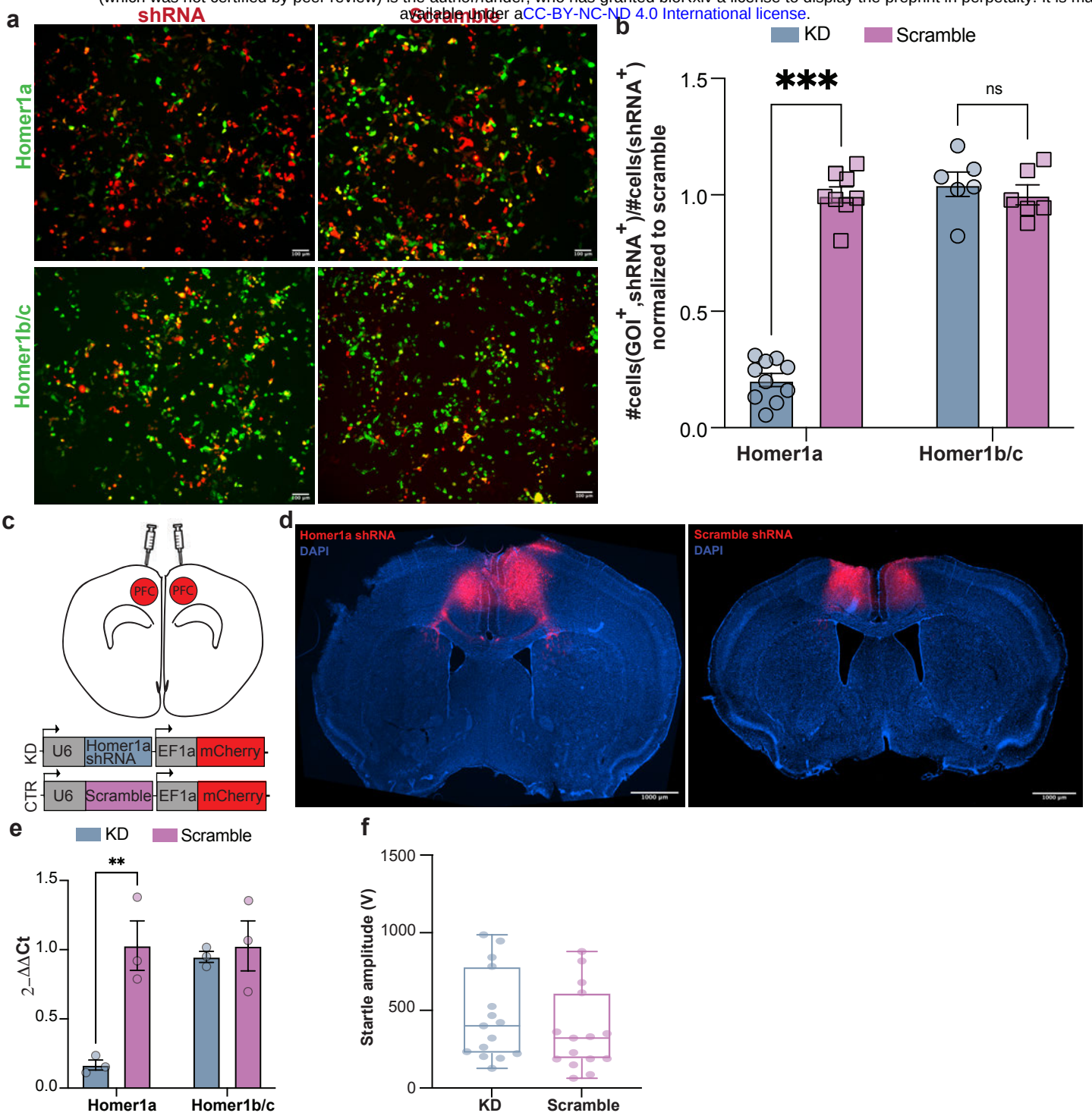
Extended data figure 1. Additional genetic and behavioral characteria tion of DO mice. Related to Fig. 1
a) Haplotype reconstruction of a representative DO mouse from the 25th generation of the population. Colors correspond to the founder lines (shown in legend) for which the genomic contribution is attributed at each depicted locus. **b)** Startle response assessed during the PPI experiment in B6 (black, n=27) and DO (grey, n=176) mice measured as startle amplitude (V). Data are expressed as mean ± SEM. **c-e)** Correlations in DO mice (n = 176) between **c)** ASR, measured as the magnitude of startle amplitude, and PPI, measured as percent inhibition, at 3 (PP3, r2=0.005), 6 (PP6, r2=0.003), and 12 (PP12, , r2=0.014) dB above background (Methods), **d)** weight and ASR (r2 = 1.084 x 10-5), and **e)** weight and PPI (PP3, r2=0.003; PP6, r2=0.002; PP12, r2=0.008) dB above background. **f,g)** QTL mapping analysis of ASR/PPI behavioral task (by R/qtl2) shown as Manhattan plots of **f)** ASR, and **g)** PPI at 3 (PPI3, red), 6 (PPI6, blue), and 12 dB (PPI12, magenta) above background (Methods). (n=176; blue lines indicate 90% confidence threshold and red lines indicate 95% confidence threshold)



Extended data figure 2. Behavioral phenotype and covariate characteria tion of CC083 and CC025 mice, related to Fig. 1
a) ASR, measured as the magnitude of the startle amplitude in CC025 (blue, n=24) and CC083 (gold, n=27) and mice. **B-d**) Correlations in CC025 (blue, n=24) and CC083 (gold, n=27) between **b)** ASR and PPI, measured as percent inhibition, at 3 dB (PP3, CC025 r2=0.015; CC083 r2=0.030), 6 dB (PP6, CC025 r2=0.054; CC083 r2=0.104), and 12 dB (PP12, CC025 r2=0.020; CC083 r2=0.068) above background (Methods), **c)** weight and ASR (CC025 r2=0.001; CC083 r2=0.004), and **d)** weight and PPI (PP3, CC025 r2=0.002, CC083 r2=0.004; PP6, CC083, r2=0.015; CC025, r2=0.001; PP12, CC025, r2=7.946 x 10-5; CC083 r2= 0.056). **e)** Short-term memory measured in a novel object recognition test for CC025 (blue, n=9) and CC083 (gold, n=10) mice, measured as time spent exploring the novel object vs the familiar one and expressed as percentage of total exploration time during a 10 min test. 2-way ANOVA showed significant main effect for novelty ($F(1,34)=440.9$, $p<0.001$), but not for strain ($F(1,34)= 1.540e-019$, $p>0.99$). **f)** Working memory performance assessed in a Y-maze apparatus for CC025 (blue, n=9) and CC083 (n=13) mice, measured as correct alternations performed and expressed as % over total alternations. No significant effect found by t-test. **g)** Gross motor activity evaluated by distance moved (in) by CC025 (n=11) and CC083 (gold, n=11) mice during 10 min in a square open field arena. **h)** Social behavior for CC025 (blue, n=9) and CC083 (gold, n=8) mice, expressed as discrimination index determined by exploration time (Methods) in a 3-chamber social interaction test. Data for e-i are expressed as mean ± SEM. i**)** Anxiety-like behavior measured as time spent in the open arm (%) during a 5 min elevated plus maze test for CC025 (blue, n=11) and CC083 (gold, n=10) mice. Welch-corrected t-test showed a significant difference between strains ($p=0.03$).

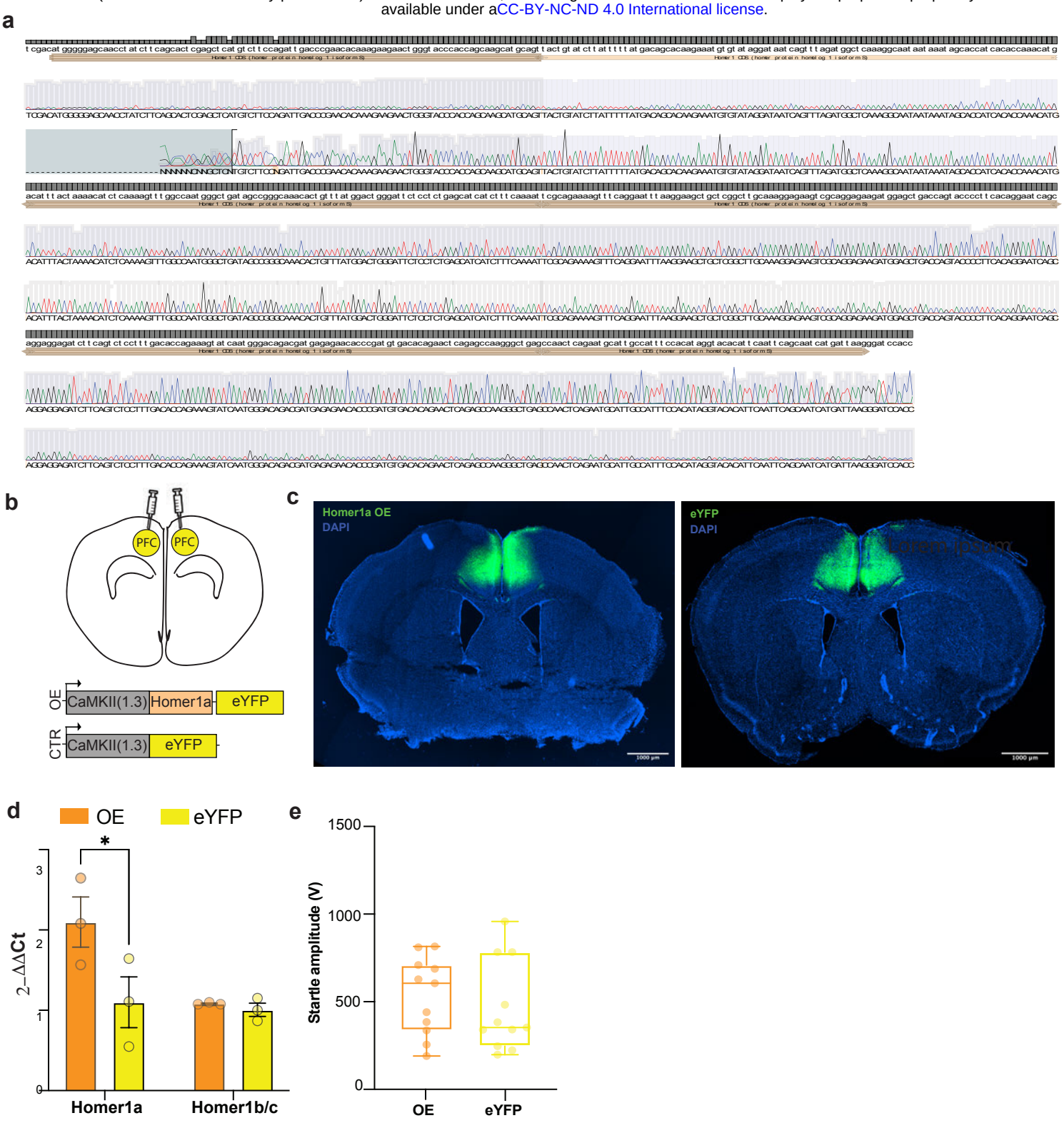


Extended data figure 3. DO RNAseq Clustering Information and Homer Exon Structure, related to Fig.2
a) Heatmap of hierarchical clustering by Euclidean distance among gene expression profiles in DO high (pink, n = 3) and low performers (blue, n = 3) as highlighted in Figure 1b and from three brain regions per mouse: MD (mediodorsal thalamus, green), PFC (prefrontal cortex, orange) and Ventral tegmental area (VTA, pink). Clustering is visible by brain region and performance in MD and PFC. **b)** Schematic representation of the Homer1 genomic exon structure. The bent arrow at the 5' end of exon 1 (solid line, above) indicates the putative transcription start site, while the bent arrow at the 3' end of exon 1 (dashed line, below) represents the translation start site. Black diamonds (below) indicate the translation stop sites of Homer1a, Ania3, and Homer1b/c, respectively. To create Homer1a, exon 5 extends into intron 5 to create the Homer1a-specific exon (5') through alternative splicing. Ania3 is generated by alternative splice usage of intron 5 sequence downstream of exon 5' as the Ania3-specific exon 6'. (Adapted from Bottai et al. 2002).

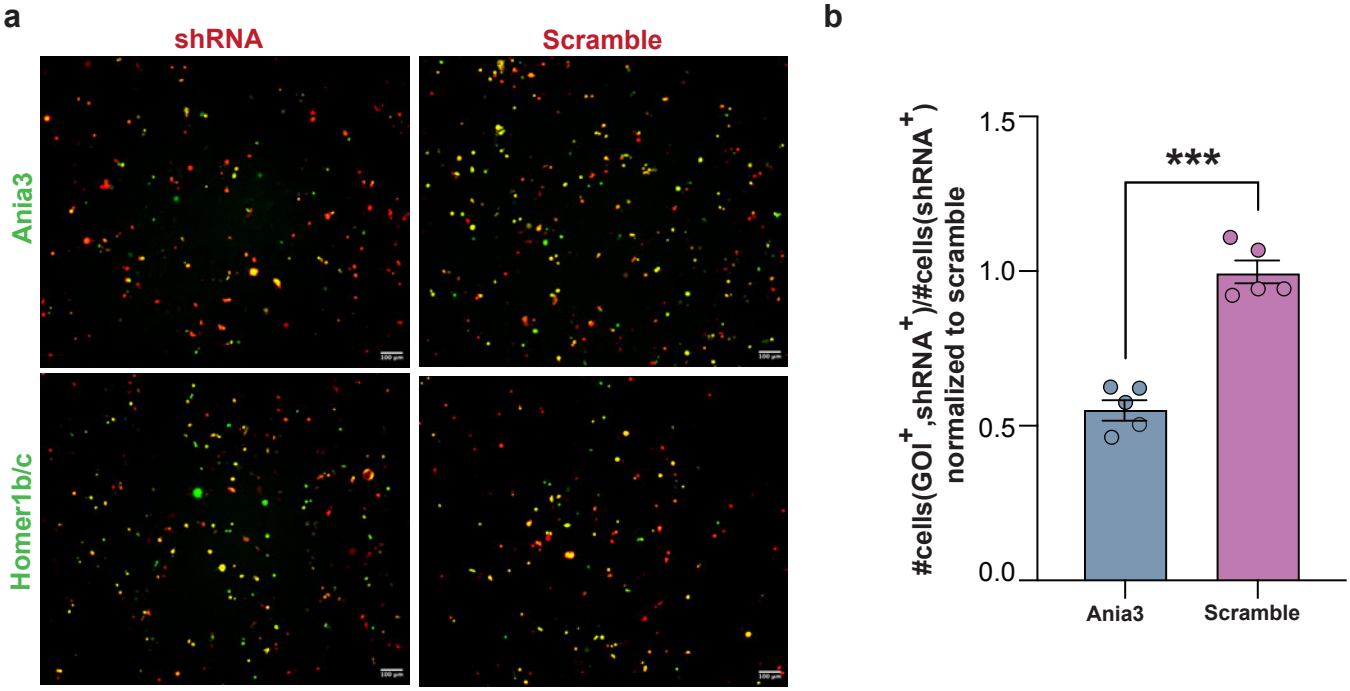


Extended data figure 4. Validation of Homer1a shRNA in vitro and in vivo and ASR for Homer1a knockdown, related to Fig. 2

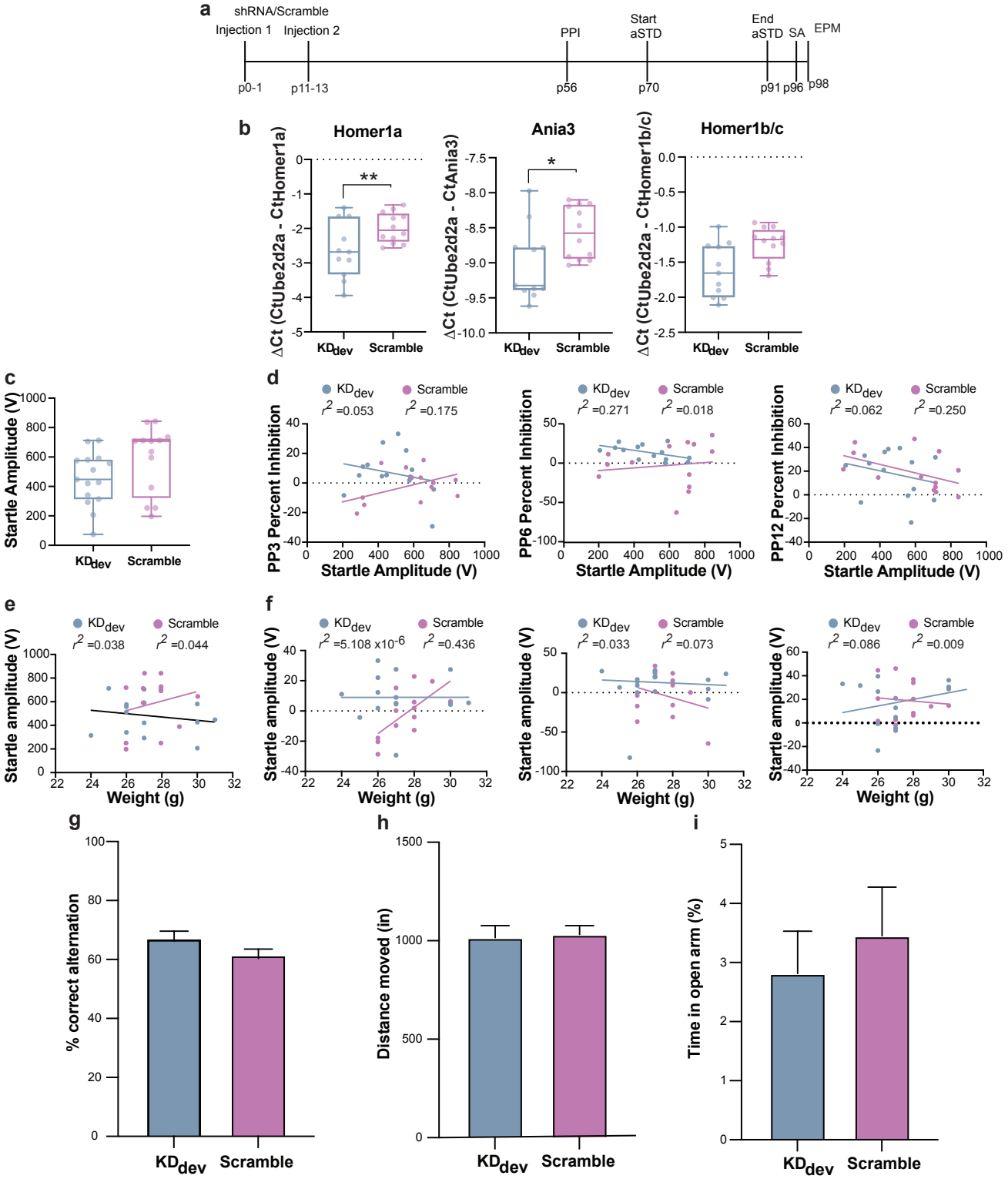
a) in vitro validation of Homer1a gene knockdown construct. Representative images of HEK cells co-transfected with Homer1a (left column) or Scramble (right column) shRNA (red) and Homer1a (top row) or Homer1b/c (bottom row) expression constructs (green). Images were collected with x10 objective and processed using the NIS-Elements (Nikon). Scale bar: 100 µm. **b)** Quantification of shRNA-mediated gene knockdown, expressed as the fraction of cells co-expressing a Homer1 expression construct and shRNA construct relative to the total number cells expressing the shRNA construct, normalized to the respective scramble control experiments, given as mean ± SD. 2-way ANOVA showed significant main effects for Homer1 isoform expression ($F(1,26)= 118.8$, $p<0.001$) and shRNA construct ($F(1,26)=94.12$, $p<0.001$), as well as a significant interaction between those variables ($F(1,26)= 18.8$, $p<0.001$). Holm-Sidak test for multiple comparisons, which, in cells co-transfected with the Homer1a expression construct, showed a significant difference in Homer1a expression between the shRNA (n=10 fields of view across 2 independent experiments) and scramble (n=8 fields of view across 2 independent experiments) constructs ($p < 0.001$), while, in cells transfected with the Homer1c expression construct, there was no significant difference between the shRNA (n=6 fields of view across 2 independent experiments) and scramble (n=6 fields of view across 2 independent experiments). **c)** Schematic of constructs and injection location (PFC) for the Homer1a knockdown manipulation in adult B6 mice. **d)** Histology performed 8 weeks after bilateral injection of viral gene knockdown of AAV-U6-Homer1a shRNA-CMV-mCherry (left; KD) and AAV-U6-Scramble-CMV-mCherry (right; control) constructs into PFC (A/P=+1.8, M/L=± 0.3, D/V=-1.75, relative to bregma) showing viral transduction in the target area. DAPI (blue), mCherry fluorescent marker (red). 4x images were collected and tiled together to generate high-resolution images of brain sections using the NIS-Elements (Nikon). Scale bars: 1000µm; **e)** Homer1a and Homer1b/c expression levels (relative to controls) in prefrontal cortex samples dissected from KD (blue, n=3) and control (purple, n=3) mice measured by qPCR. 2-way ANOVA showed significant main effects for shRNA construct ($F (1, 8) = 13.11$, $p=0.007$) and Homer1 isoform expression ($F (1, 8) = 8.999$, $p=0.02$), as well as a significant interaction between those variables ($F (1, 8) = 9.073$, $p=0.02$). Post hoc Holm-Sidak test for multiple comparisons shows a significant difference in Homer1a expression ($p=0.003$) but not in Homer1b/c expression ($p=0.68$) between mice injected with the Homer1a shRNA and controls. Data are expressed as mean ± SEM. **f)** ASR in KD (blue, n=15) and control (purple, n=15). No significant difference was found by t-test.



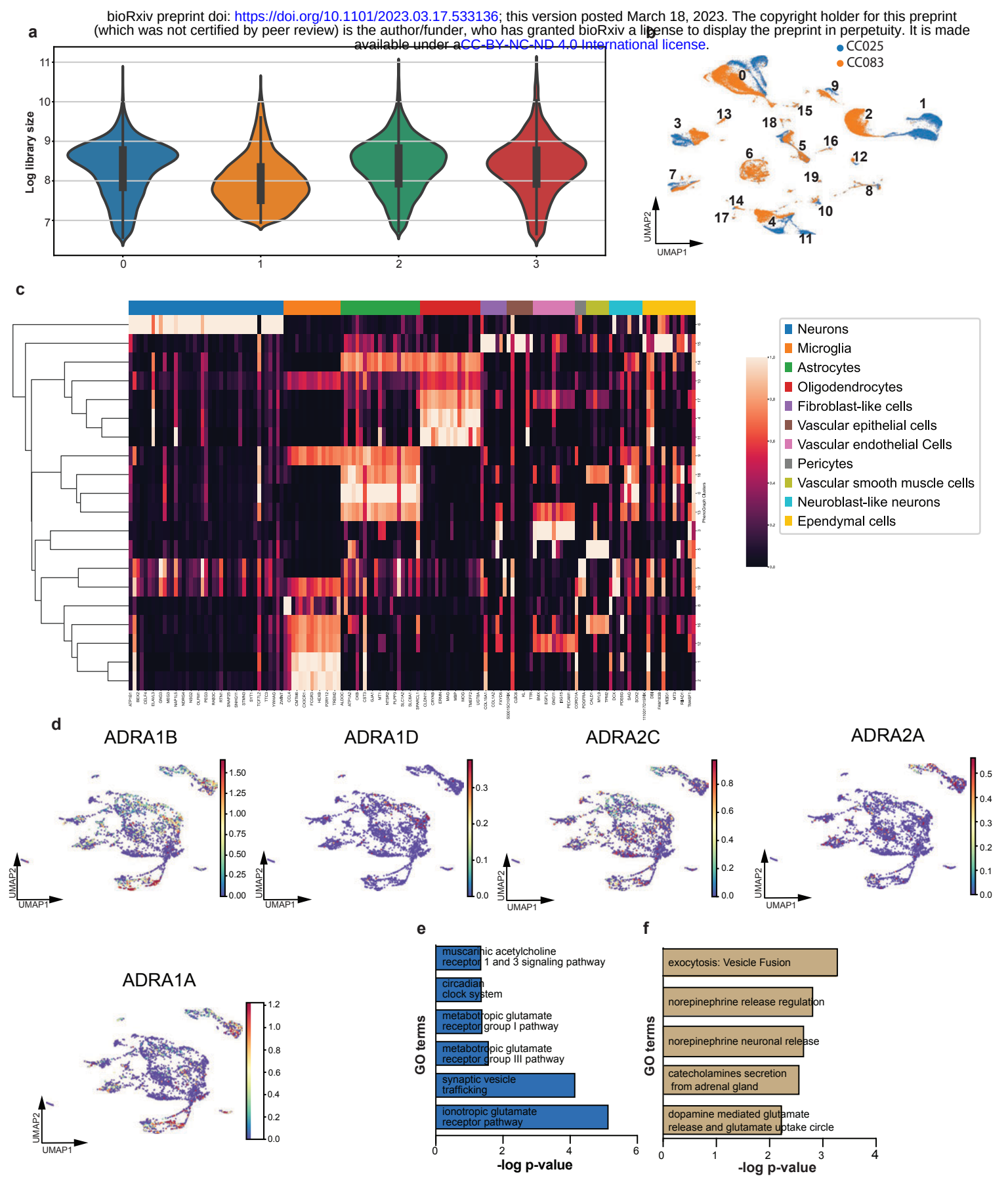
Extended data figure 5. Characteria tion of Homer1a overexpression manipulation, related to Fig. 2
a) Electropherogram of AAV-CaMKII(1.3)-Homer1a¬-eYFP overexpression construct aligned to the Homer1a coding sequence (tan bar near the top of each line). The shaded area indicates regions covered by the Sanger sequencing, height of grey boxes at the top of each line is proportional to the number of sequencing runs aligned to the reference sequence (maximum # of sequencing runs in image = 2). **b)** Schematic of constructs and injection location (PFC) for the Homer1a overexpression manipulation in adult B6 mice. **c)** Histology performed 8 weeks after bilateral injection of viral gene knockdown of AAV-CaMKII(1.3)-eYFP (left; OE) and AAV-CaMKII(1.3)-eYFP (right; control) constructs into PFC (A/P=+1.8, M/L=±0.3, D/V=-1.75, relative to bregma) showing viral transduction in the target area. DAPI (blue), eYFP fluorescent marker (green). 4x images were collected and tiled together to generate high-resolution images of brain sections using the NIS-Elements (Nikon). Scale bars: 1000µm. **d)** Homer1a and Homer1b/c expression levels (relative to controls) in prefrontal cortex samples dissected from OE (orange, n=3) and control (yellow, n=3) mice measured by qPCR. 2-way ANOVA showed significant main effects for expression construct ($F(1,8)=5.652$, $p=0.04$) and Homer1 isoform expression ($F(1,8)=6.010$,$p=0.04$. Post hoc Holm-Sidak test for multiple comparisons shows a significant difference in Homer1a expression ($p=0.03$) but not in Homer1b/c expression ($p=0.81$) between mice injected with the Homer1a overexpression and eYFP control constructs. Data are expressed as mean ± SEM. **e)** ASR in OE (orange, n=15) and control (yellow, n=15). No significant difference was found by t-test.



Extended data figure 6. In vitro characteria tion of Ania3 shRNA construct, related to Fig. 2
a) in vitro validation of gene knockdown constructs. Representative images of HEK cells co-transfected with Ania3 (left column) or Scramble (right column) shRNA (red) and Ania3 (top row) or Homer1b/c (bottom row) expression constructs (green). Images were collected with x10 objective and processed using the NIS-Elements (Nikon). Scale bar: 100 µm. **b)** Quantification of shRNA-mediated gene knockdown, expressed as the fraction of cells co-expressing the Ania3 expression construct and shRNA or scramble construct relative to the total number cells expressing the shRNA or scramble, normalized to the scramble control experiments, given as mean ± SD. In cells transfected with the Ania3 expression construct, there was a significant difference in Ania3 expression between the cells co-transfected with the shRNA (n=5 fields of view), and scramble (n=5 fields of view) constructs ($p<0.001$). Pairwise comparisons were made using a student's t test.

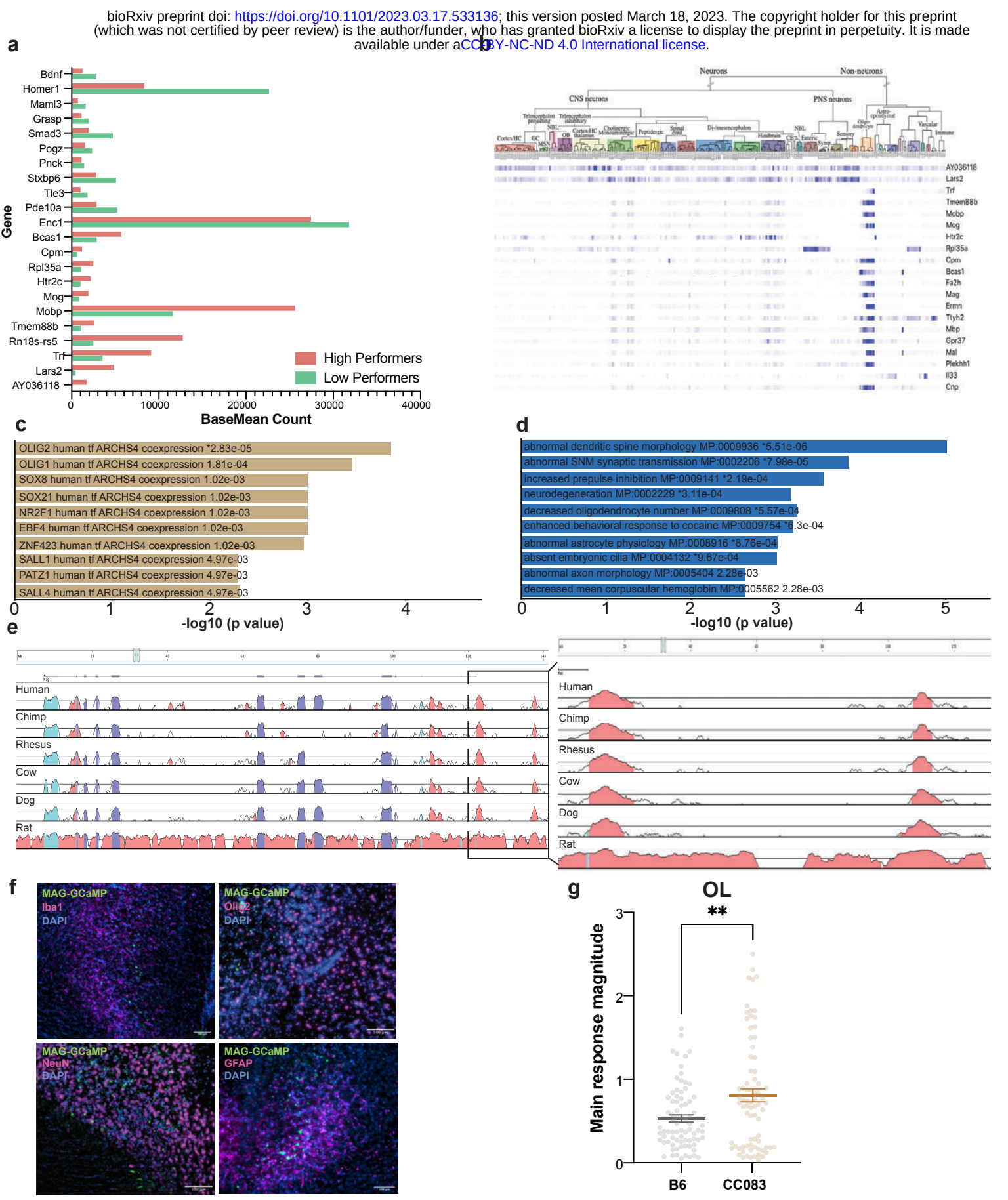


Extended data figure 7. In vivo characteria tion of developmental knockdown (KDdev) manipulation, related to Fig. 2
a) Experimental timeline, indicating virus injection time (p0-1 and p11-13), and experiments sequence: Prepulse inhibition (PPI) at p56; acoustic signal detection task (aSDT) from p70 to p91; Spontaneous alternation (SA) at p96, and Elevated plus maze (EPM) at p97. **b)** ex vivo validation of knockdown manipulation assessed by quantification of Homer1a (left), Ania3 (center) and Homer1b/c (right) levels measured by qPCR in prefrontal cortex samples dissected from KDdev (blue, n=15) and Scramble (purple, n=11). 2-way ANOVA showed significant main effects for shRNA construct ($F(1,72)= 15.07$, $p<0.001$) and Homer1 isoform expression ($F(1,72)=15.07$, $p<0.001$). Post hoc Holm-Sidak test for multiple comparisons shows a significant difference in Homer1a ($p=0.009$) and Ania3 ($p=0.05$) expression, but not in Homer1b/c expression ($p=0.18$) between mice injected with the pooled shRNA and controls. **c)** ASR in KDdev (blue, n=15) and Scramble (purple, n=13) mice, measured as the magnitude of startle amplitude. **d-f)** Correlations between **d)** ASR and PPI, measured as percent inhibition, at 3 dB (PP3, KDdev, r2 = 0.053; Scramble, r2 = 0.175), 6 dB (KDdev, r2 = 0.271; Scramble, r2 = 0.018), and 12 dB (KDdev, r2 = 0.062; Scramble, r2 = 0.250) above background (Methods), **e)** weight (g) and ASR (KDdev, r2 = 0.038; Scramble, r2 = 0.044), and **f)** weight and PPI (PP3: KDdev r2 = 5.108 x10-6, Scramble r2 = 0.436; PP6: KDdev r2 = 0.033, Scramble r2 = 0.073; PP12: KDdev r2 = 0.086, Scramble r2 = 0.009). **g)** Working memory performance assessed in a Y-maze, measured as correct alternations performed and expressed as % over total alternations. **h)** Gross motor activity measured as distance moved (inch) by KDdev (blue, n=13) and Scramble (purple, n=12) during a 6 min spontaneous alternation test in a Y-maze. **i)** Anxiety-like behavior measured as time spent in the open arm of the elevated plus maze (%) during a 5 min test.



Extended Data Fig. 8 Additional Information for scRNAseq Experiments, Related to Fig. 3

a) Violin Plots of library size for each biological replicate. **b)** UMAP visualization of initial clusters colored by strain. **c)** Heatmap of marker genes used for cell type identification of clusters (from b). Initial clusters were merged by cell type to form clusters in Fig. 3b. **d)** Scaled expression of adrenergic receptor genes in neuronal clusters in UMAP plot (Fig. 3c). **e)** Functional pathway enrichment analysis for CC025 cells in the Homer1 DE clusters using the Reactome_2022 gene set library in Enrichr. **f)** Functional pathway enrichment analysis for CC083 cells in the Homer1 DE clusters using the Elsevier_Pathway_Collection gene set library in Enrichr.



Extended data figure 9 Generation and validation of MAG-GCaMP6f AAV construct, related to Fig. 4

a) The most differentially expressed genes between DO high- (n=3) and DO low (n=3) from bulk PFC RNAseq (Fig. 2b). Genes are ordered by Log2FC relative to high-performers (11 upregulated and 11 downregulated). **b)** Typical cell-type expression patterns for genes most upregulated in PFC of DO mice with high PPI performance. Differential gene expression between performance groups was determined by DESeq2, n = 6, biologically independent samples. Cell-type expression patterns were determined by the Linnarsson Lab adult mouse single cell gene expression database(http://www.mousebrain.org). Genes selected for each group had log2 fold change ≥ 0.7 and adjusted p value ≤ 0.05. **c)** Transcription factor co-expression enrichment analysis (from ARCHS4 database) of genes significantly upregulated in the PFC of CC083 mice (n=3) relative to CC025 mice (n=3) from bulk RNAseq data (Fig. 2e). **d)** Mammalian phenotype ontology enrichment analysis (from MGI Mammalian Phenotype database with Level 4 cutoff) of genes upregulated in CC025 cells within neuron subcluster 0 from scRNAseq (Fig. 3). **e)** VISTA plots of mouse sequence at the MAG locus and its genomic sequence conservation in humans, chimpanzees, rhesus macaques, cows, dogs, and rats. The box indicates the 2.3 kb upstream promoter region zoomed-into on the right with 2 regions of high conservation, which suggests they are likely functionally important. Purple shading indicates coding sequence conservation, blue shading indicates UTR sequence conservation, and pink shading indicates non-coding sequence conservation. **f)** Immunohistochemistry performed 4 weeks after PFC injections (A/P=+1.85, M/L=-0.35, D/V=-2.55, relative to bregma) of AAV-MAG-GCaMP6f (green) showing viral transduction in the target area. 20x images were collected of sections stained with antibodies (magenta) raised against the microglial marker Iba1 (top left), neuronal marker NeuN (bottom left), oligodendroglial lineage marker Olig2 (top right), and astrocytic marker GFAP (bottom right), as well as DAPI (blue). **g)** Average activity (area under responses) in home cage for B6 (n = 5) versus CC083 (n = 4) during 5 minute recordings from PFC oligodendrocytes.

a

ENSEMBLE ID	Gene Symbol	Gene Name	Start (bp)	Stop (bp)	Strand
ENSMUSG00000014850	Msh3	mutS homolog 3 (E. coli)	92211872	92355003	-
ENSMUSG00000021707	Dhfr	dihydrofolate reductase	92354783	92389053	+
ENSMUSG00000045034	Ankrd34b	ankyrin repeat domain 34B	92425896	92441658	+
ENSMUSG00000034334	Fam151b	family with sequence similarity 151, member B	92449622	92484168	-
ENSMUSG00000021706	Zfyve16	zinc finger, FYVE domain containing 16	92487108	92530868	-
ENSMUSG00000046957	Spz1	spermatogenic leucine zipper 1	92574631	92576232	-
ENSMUSG00000021703	Serinc5	serine incorporator 5	92611138	92711947	+
ENSMUSG00000021702	Thbs4	thrombospondin 4	92751586	92794818	-
ENSMUSG00000021704	Mtx3	metaxin 3	92844787	92858230	+
ENSMUSG00000047419	Cmya5	cardiomyopathy associated 5	93040713	93144724	-
ENSMUSG00000042167	Papd4	PAP associated domain containing 4	93147400	93192283	-
ENSMUSG00000007617	Homer1	homer scaffolding protein 1	93303757	93404129	+
ENSMUSG00000021690	Jmy	junction-mediating and regulatory protein	93430097	93499808	-
ENSMUSG00000074768	Bhmt	betaine-homocysteine methyltransferase	93616891	93637961	-
ENSMUSG00000042118	Bhmt2	betaine-homocysteine methyltransferase 2	93655720	93674302	-
ENSMUSG00000042102	Dmgdh	dimethylglycine dehydrogenase precursor	93674433	93752831	+
ENSMUSG00000042082	Arsb	arylsulfatase B	93771679	93943016	+
ENSMUSG00000045312	Lhfpl2	lipoma HMGIC fusion partner-like 2	94057796	94195409	+

Extended Data Table 1
Table 1: Protein-coding genes within the 95% CI surrounding the Chr 13 QTL identified by rQTL2.