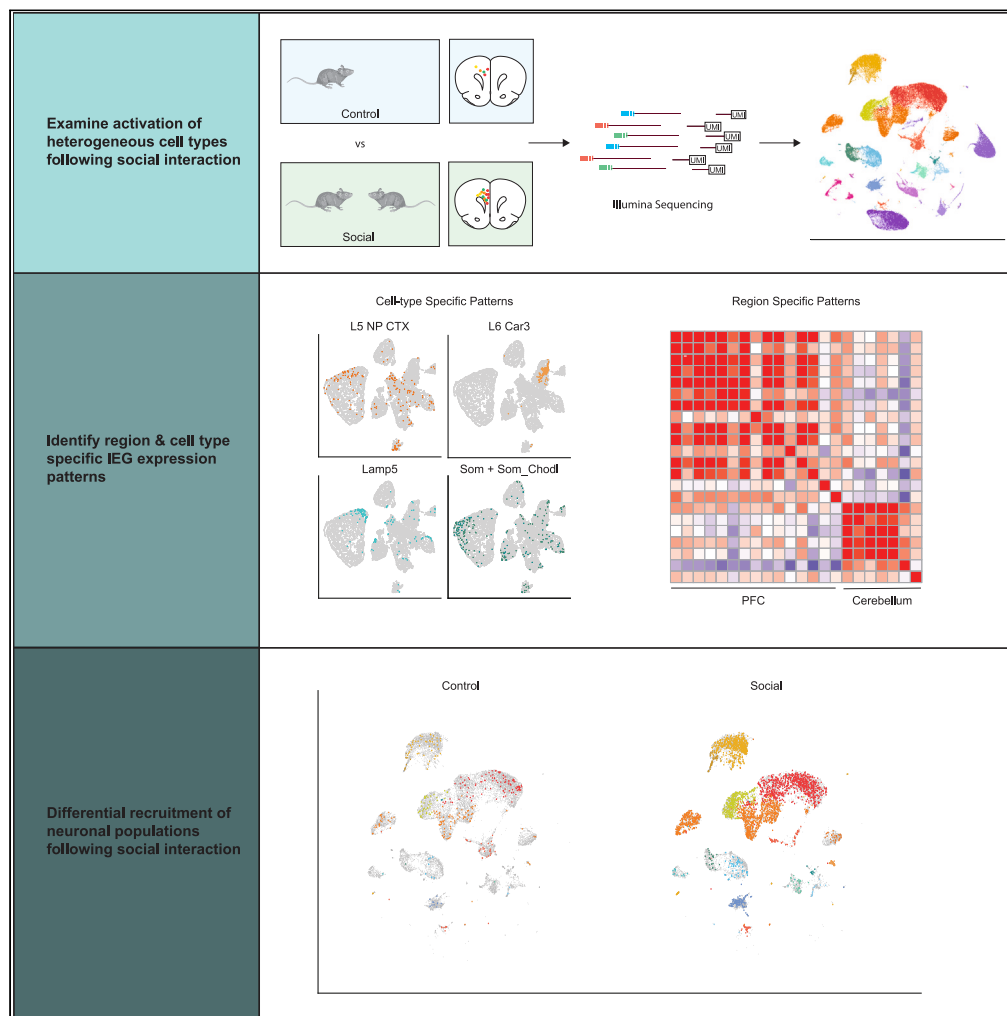## Article

Distinct transcriptional programs define a heterogeneous neuronal ensemble for social interaction

Hailee Walker,
Nicholas A. Frost

nick.frost@hsc.utah.edu

Highlights

Heterogeneous neuronal populations are recruited during social interaction

Excitatory and inhibitory populations are variably recruited in the mPFC

Active neuronal populations were identified by quantifying IEG expression

Cell-type and region-specific IEG transcriptional programs were observed

Article

Distinct transcriptional programs define a heterogeneous neuronal ensemble for social interaction

Hailee Walker[1] and Nicholas A. Frost[1,2,*]

SUMMARY

Social interactions are encoded by the coordinated activity of heterogeneous cell types within distributed brain regions including the medial prefrontal cortex (mPFC). However, our understanding of the cell types which comprise the social ensemble has been limited by available mouse lines and reliance on single marker genes. We identified differentially active neuronal populations during social interactions by quantifying immediate-early gene (IEG) expression using snRNA-sequencing. These studies revealed that distinct prefrontal neuron populations composed of heterogeneous cell types are activated by social interaction. Evaluation of IEG expression within these recruited neuronal populations revealed cell-type and region-specific programs, suggesting that reliance on a single molecular marker is insufficient to quantify activation across all cell types. Our findings provide a comprehensive description of cell-type specific transcriptional programs invoked by social interactions and reveal insights into the neuronal populations which compose the social ensemble.

INTRODUCTION

The medial prefrontal cortex (mPFC) plays a critical role in the regulation of normal social interactions.[1,2] Within the mPFC, it is the activity of heterogeneous neuronal sub-types which permits the encoding of diverse information relevant to social interactions,[2–4] anxiety-related behaviors,[5] and flexible representations of environment.[6]

Neurons may be segmented by layer, morphological appearance, connectivity, and specific patterns of gene regulation into heterogeneous populations with distinct functional properties within the cortical microcircuit. Normal social interactions require the activity of both excitatory and inhibitory neuron types.[7–9] The activity of both populations of neurons is critical for normal social behavior[8,9] indicating that social interactions likely utilize computations involving heterogeneous cell types. Single cell datasets have revealed that numerous transcriptionally defined cell types exist within the laminar cortex[10] and are non-randomly connected[11] suggesting distinct roles in microcircuit-level computations during behavior. Despite this there has not been a systematic description of the composition of a social ensemble by heterogeneous cell types.

Previous studies quantifying[4] or modulating[8,9] the activity of specified populations of neurons during social interaction have provided extraordinary insight into the role of specific molecularly or anatomically defined neurons in computations underlying normal social behavior. However, both electrophysiological and imaging recordings are limited by the number of neuron subtypes which can be assayed simultaneously, as well as the specificity by which neuronal populations can be defined. Similarly, 'tagging' approaches[12,13] may be limited in that they typically utilize a single molecular tag to define active populations of neurons which may introduce bias as not all neuron subtypes utilize the same transcriptional program.

We utilized a non-biased transcriptional approach[14] to identify heterogeneous neuronal populations within the mPFC and quantify their recruitment during social interactions. Prefrontal neurons isolated from wildtype mice following social interaction showed increased expression of a broad population of immediate-early genes (IEGs). Utilizing a panel of socially responsive IEGs we identified robust and reliable recruitment of activated cells within molecularly defined glutamatergic and GABAergic populations. Quantification revealed differential neuronal recruitment within distinct excitatory cell types, even within the same cortical layer, highlighting the precision with which neuronal ensembles are recruited during behavior.

RESULTS

Transcriptomic identification of heterogeneous cell types in the mPFC and cerebellum

Prefrontal microcircuits contain heterogeneous excitatory and inhibitory cell types which play functionally distinct roles in microcircuit function.[11] We utilized unbiased high-throughput single-nuclei RNA sequencing (snRNA-seq) to transcriptomically identify heterogeneous cell

populations within specific brain regions (Figures 1A–1C). We isolated nuclei from the mPFC or cerebellum and generated 10X sequencing libraries from which we identified 101,520 cells from the mPFC of 14 adult c57 bl6/J (WT) mice. Identified nuclei were clustered with Seurat and cell types were identified based on known cell-type specific marker genes[10,15] (Figures 1B and 1C). In total, twenty-two cell types were identified in the mPFC, each with unique transcriptional profiles (Figures 1C and 1D).

Neurons within the prefrontal cortex are organized into distinct layers (L1, 2/3, 5, and 6) and may be segmented by transcriptomic profile, morphology, and functional properties. We identified excitatory neurons and assigned them to cortical layers in the mPFC using previously identified marker gene profiles (Figure 1E). In total we identified 2 excitatory cell types found in layer 2/3 (L2/3 Exc and L2/3/5 Exc), 2 cell types found in layer 5 (L5 Exc and L5 NP Ctx), and 4 cell types which mapped to layer 6 (or L6 Car3, L6b CTX, L6 Exc Opkr3, and L6 Exc Syt6).

Inhibitory cell type markers Pvalb, Sst, and Vip mapped onto inhibitory clusters marked by expression of Gad2 (Figure 1E). Som Chodl neurons were defined by high expression of Chodl which has been associated with projecting interneurons.[15] Lamp5 and Meis2,Pbx3 inhibitory neurons express Gad2 and their respective marker genes. Meis2,Pbx3 neurons have previously been identified in the mPFC[10] though they are not well-characterized.

Roughly 20% of identified cells were non-neuronal and included astrocytes (Astro), oligodendrocytes (Oligo), oligodendrocyte progenitor cells (OPC), microglia, endothelial cells, vascular/leptomeningeal cells (VLMC), and pericytes (Figure 1E). Consistent with prior expectations,[10,16] excitatory neurons composed 53.24% of all identified cells in the mPFC, while inhibitory neurons made up a smaller (17.1%) proportion.

We identified an additional 90,719 cells from the cerebellum of the same 14 mice (Figures S1A–S1C). Cell clusters were identified based on previously described marker genes corresponding to known cerebellar cell types,[17] which mapped to non-overlapping populations of cerebellar cells (Figure S1D). We identified seventeen distinct cerebellar cell types in total (Figures S1A–S1C). We focused on excitatory and inhibitory neuronal populations in the remainder of this study.

Identification of region and cell type-specific patterns of IEG activation following social interaction

We reasoned that neuronal populations activated during social interaction would display increased IEG expression in activated cells as IEGs are quickly and transiently expressed in neurons following neuronal activity and are critical for experience-dependent plasticity.[18] We pooled cells across mice to examine the expression of a broad panel of IEGs[14] in each neuronal cluster in control or social animals and examined the proportion of neurons which expressed at least one copy of each candidate IEG. We considered an IEG to be socially responsive if the proportion of neurons expressing the gene increased by at least 2-fold in at least one cluster in either social condition (compared to control+1%). We identified a total of 51 socially responsive IEGs (34 in the mPFC only, 9 in the cerebellum only, 10 in both) (n = 86,559 neurons from 6 control mice, 39,233 neurons from 3 mice collected 10 min following social interaction, and 66,447 neurons from 5 mice collected 35 min following social interaction).

We next examined IEG expression by mouse. Because several clusters contained 0 counts of individual IEGs in one or more condition when examined by mouse we added 1% to both the numerator and denominator. This generated a somewhat larger list than the analysis of pooled data, and we observed an increase in the expression of many IEGs in the social condition compared to control mice (Figure 2B; mPFC: 10-min 50.3 $\pm$ 0.3 IEGs, 35-min 45.8 $\pm$ 1.5 IEGs; Cerebellum: 10-min 31.7 $\pm$ 1.2 IEGs, 35-min 28.7 $\pm$ 1.5 IEGs) which was significantly higher than that observed in shuffled data (Figure 2B; mPFC: 10-min 23.0 $\pm$ 2.1 IEGs, 35 min 23.4 $\pm$ 1.2 IEGs; Cerebellum: 23.3 $\pm$ 1.5 IEGs, 35 min 22.3 $\pm$ 2.0 IEGs, p < 0.0001 for shuffled v. real, two-way ANOVA). We observed cell-type specific differences in the number of IEGs which increased following social interaction (Figure 2A). Among excitatory neurons, we identified the largest number of socially-responsive IEGs in L6 Car3 and L2/3/5 pyramidal neurons (39 IEGs and 34 IEGs at 10 min post social interaction, respectively) whereas significantly fewer IEGs were identified in L5 NP CTX (20 IEGs at 10 min post social interaction).

Within inhibitory populations, LGE-derived Meis2, Pbx3-1 cells contained the most socially-responsive IEGs (33 IEGS increased > 2-fold at 10 min). Overall, a larger number of socially responsive IEGs were identified in the mPFC than in the cerebellum (Figure 2C; mPFC: 10-min 50.3 $\pm$ 0.3 IEGs, 35-min 45.8 $\pm$ 1.5 IEGs; Cerebellum: 10-min 31.7 $\pm$ 1.2 IEGs, 35-min 28.7 $\pm$ 1.5 IEGs, p < 0.0001 for region, two-way ANOVA). This increase in IEG expression was time-dependent such that a greater number of IEGs were positively modulated at 10 min post-interaction than at 35 min post-interaction, presumably due to egress from the nucleus (Figure 2C; p = 0.0273 for time point, two-way ANOVA). We restricted further analysis to the more conservative panel of IEGs identified in pooled data.

We next examined the degree of similarity between the population of IEGs which were socially-responsive across different cell types and conditions (Figures 2D and 2E). Pooled excitatory neurons in the mPFC displayed similar changes in IEG expression at 10 and 35 min which diverged from changes in expression observed in the cerebellum. (Excitatory mPFC neurons: 35-min mPFC vs. 10-min mPFC, correlation coefficient 0.29 $\pm$ 0.03, 35 min mPFC vs. 35 min mPFC correlation coefficient 0.22 $\pm$ 0.02, 35 min mPFC vs. 10 min cerebellum correlation coefficient 0.14 $\pm$ 0.01, 35 min mPFC vs. 35 min cerebellum correlation coefficient 0.10 $\pm$ 0.02, p < 0.0001, ANOVA). Similarly, we observed a greater similarity between IEG expression in pooled cerebellar excitatory neurons from the cerebellum at 10 min and cerebellar populations at 35 min than to mPFC expression at either timepoint (Excitatory cerebellar neurons: 35-min cerebellum vs. 10-min cerebellum correlation coefficient 0.36 $\pm$ 0.03, 35 min cerebellum vs. 35 min cerebellum correlation coefficient 0.31 $\pm$ 0.03, 35 min cerebellum vs. 10 min mPFC correlation coefficient 0.08 $\pm$ 0.01, 35 min cerebellum vs. 35 min mPFC correlation coefficient 0.10 $\pm$ 0.02, p < 0.0001, ANOVA). We observed similar findings when we examined inhibitory populations in the mPFC and the cerebellum at each timepoint (Figures 2D and 2E). This suggests that region-specific patterns of IEG expression are more distinct than cell-specific patterns of IEG expression.

To further examine the similarity in IEG expression between broad neuronal cell type (excitatory and inhibitory) in the mPFC and cerebellum, we performed principal component analysis (PCA) on the expression of 51 socially responsive IEGs. The first three principal
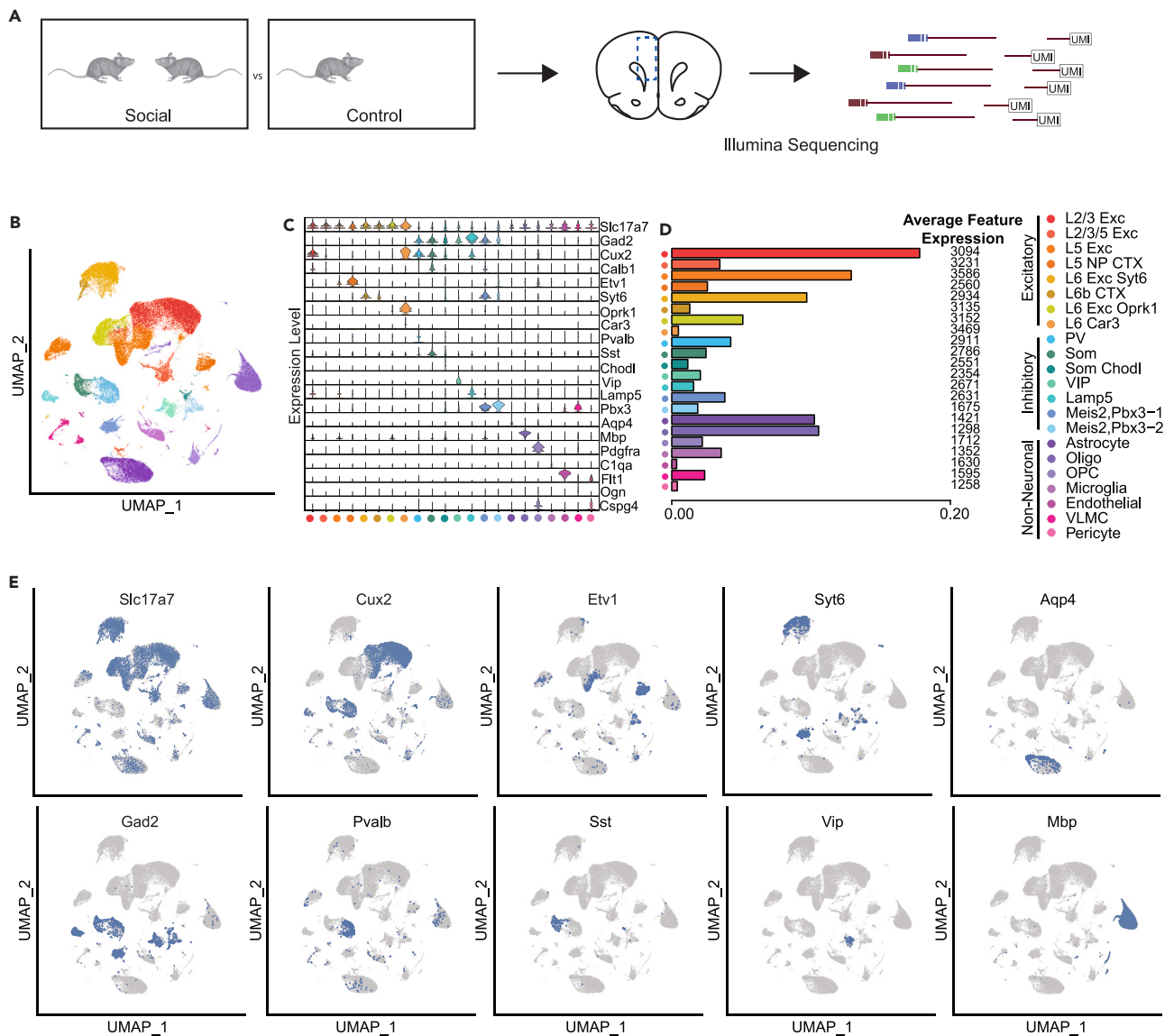
Figure 1. Identification of heterogeneous prefrontal cell types

(A) Overview of experimental workflow.

(B) UMAP visualization of clustered prefrontal cortex cells based on nearest neighbor calculations of 101,520 single-nuclei transcriptomes from WT mice (3 early social, 5 late social; 6 non-social controls). Individual clusters were identified based on gene expression profiles.

(C) Violin plot showing the expression level of known marker genes across cell types.[10,15]

(D) Bar plot representing the proportion of cells belonging to each cell type and average number of genes expressed by cells of that type (L2/3 Exc – 17.8%, L2/3/5 Exc – 3.5%, L5 Exc – 12.9%, 5 NP CTX – 2.6%, L6 Exc Syt6 – 9.7%, L6b CTX – 1.3%, L6 Exc Oprk1 – 5.1%, L6 Car3 – 0.5%, PV – 4.2%, Som – 2.5%, Som Chodl – 1.1%, VIP 2.1%, Lamp5 - 1.6%, Meis2,Pbx3-1 – 3.8%, Meis2, Pbx3-2 – 1.9%, Astrocytes – 10.2%, Oligodendrocytes – 10.5%, OPC – 2.2%, Microglia 3.6%, Endothelial 0.4%, VLMC – 2.4%, Pericyte – 0.4%).

(E) UMAP visualization as in B, with cells colored based on high level of expression (>~60% max) of given well known, cell-type specific marker genes.

components accounted for 80.0% of the variance, with differences between the cerebellum and mPFC, as well as between excitatory and inhibitory neuron types, accounting for the majority of IEG expression differences (Figure 3A). Examination of vectors drawn from the mean of the control condition to the mean of either social group suggested divergent patterns of IEG activation occurred in mPFC and cerebellar neurons (Figures 3B and 3C).

To quantify these similar and divergent patterns of IEG expression we calculated the cosine similarity between vectors drawn between individual control samples, and social samples from either group (10-min or 35-min; Figures 3D–3G). As before this revealed a greater similarity between IEG expression in pooled excitatory neurons within the mPFC at 10 min and IEG expression within the mPFC at 35 min than
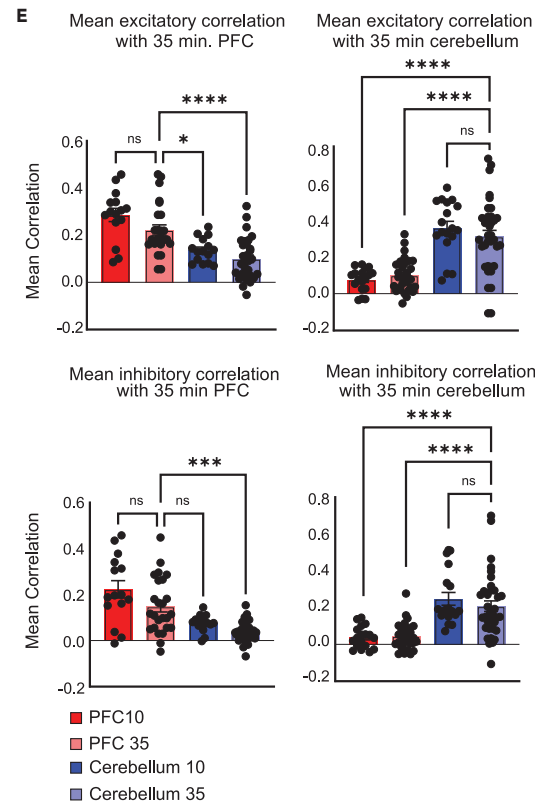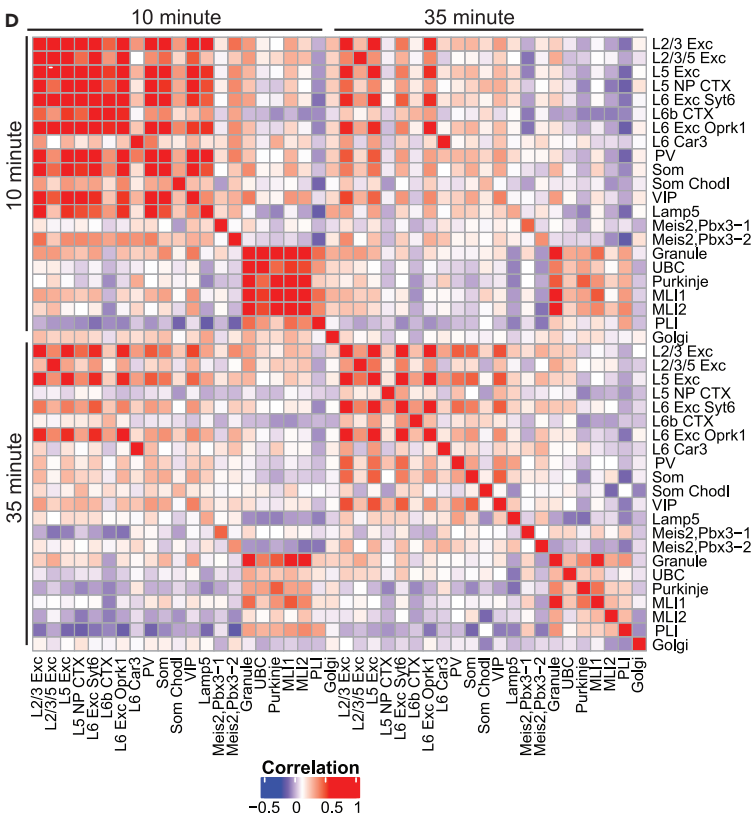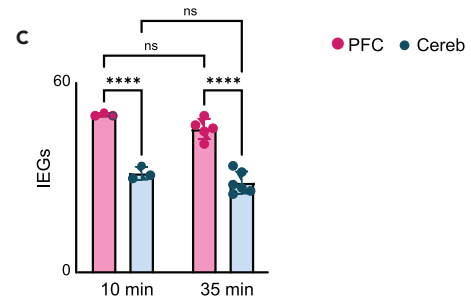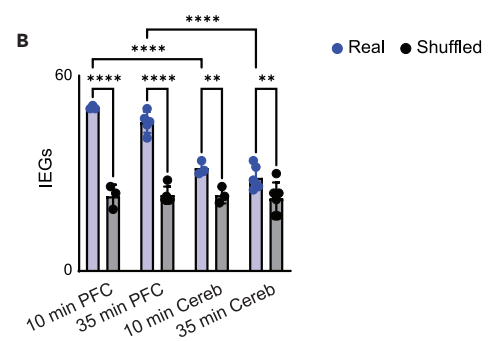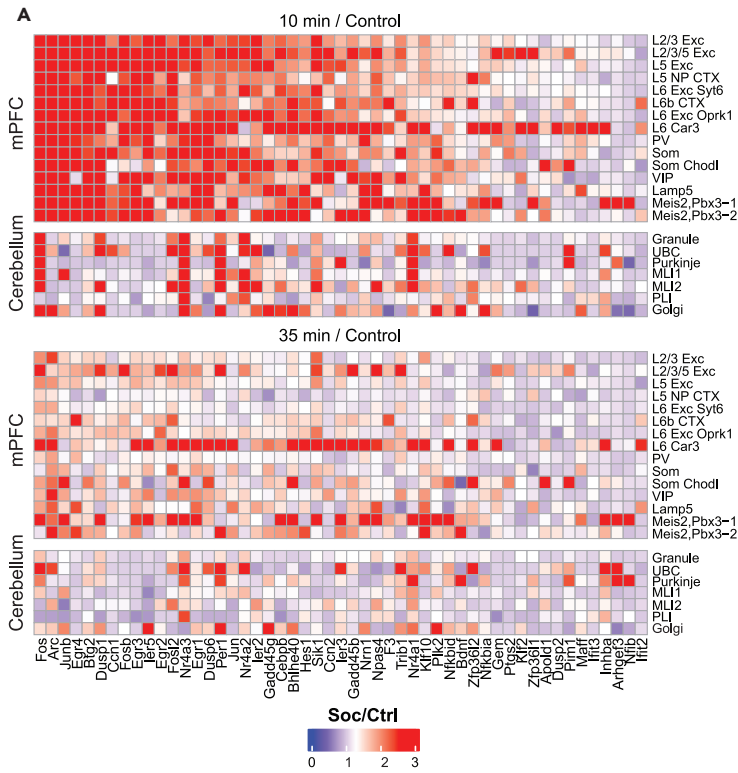
Figure 2. Region-specific patterns of IEG expression following social interaction

We identified 51 IEGs which exhibited greater than 2-fold increase in expression at either 10 or 35 min following social interaction (compared to cells from the same cluster in control mice).

(A) Color plot representing fold-change in the social condition (10 or 35 min) of each IEG (x axis) for each cell type within the mPFC or cerebellum. IEGs are sorted based on their fold-increase in mPFC cells. Because some clusters contained 0 counts of a few IEGs in the control condition we added 1% to the denominator for all IEGs.

(B) Bar graph depicting the number of IEGs per mouse which exhibit 2-fold or greater increase in at least one cell type compared to shuffled data (Real: mPFC 10-min 50.3 $\pm$ 0.3 IEGs, mPFC 35-min 45.8 $\pm$ 1.5 IEGs, cerebellum 10-min 31.7 $\pm$ 1.2 IEGs, cerebellum 35-min 28.7 $\pm$ 1.5 IEGs; Shuffled: mPFC 10-min 23.0 $\pm$ 2.1 IEGs, mPFC 35-min 23.4 $\pm$ 1.2 IEGs, cerebellum 10-min 23.3 $\pm$ 1.5 IEGs, cerebellum 35-min 22.3 $\pm$ 2.0 IEGs, $p < 0.0001$ for real vs. shuffled, $p < 0.0001$ for region and time point, two-way ANOVA with Sidak's multiple comparisons) with standard error.

(C) Bar graph demonstrating the number of IEGs which exhibit 2-fold or greater increase by region and time (mPFC 10-min 50.3 $\pm$ 0.3 IEGs, cerebellum 10-min 31.7 $\pm$ 1.2 IEGs, mPFC 35-min 45.8 $\pm$ 1.5 IEGs, cerebellum 35-min 28.7 $\pm$ 1.5 IEGs, $p < 0.0001$ for region, $p = 0.0273$ for time point, ANOVA with Uncorrected Fisher's LSD) with standard error.

(D) Correlation matrix showing similarity in IEG expressing proportions for each cell type, region and at each time point. Values were calculated by creating correlation matrix for each social sample/control sample comparison and then averaging by mouse and then across all mice.

(E) Bar graph of correlation coefficients between samples from each region and time point compared to 35 min excitatory (top left; 35-min excitatory mPFC vs.: 10-min mPFC, 0.29 $\pm$ 0.03 correlation coefficient, $p = 0.121$, 35-min mPFC, 0.22 $\pm$ 0.02 correlation coefficient, 10-min cerebellum, 0.14 $\pm$ 0.01 correlation coefficient, $p < 0.019$, 35-min cerebellum, 0.10 $\pm$ 0.02 correlation coefficient, $p < 0.0001$, one-way ANOVA with Sidak's multiple comparisons) and inhibitory mPFC neuron populations (bottom left; 35-min inhibitory mPFC vs.: 10-min mPFC, 0.22 $\pm$ 0.04 correlation coefficient, $p = 0.055$, 35-min mPFC, 0.15 $\pm$ 0.03 correlation coefficient, 10-min cerebellum, 0.08 $\pm$ 0.01 correlation coefficient, $p < 0.065$, 35-min cerebellum, 0.04 $\pm$ 0.01, $p < 0.0001$, one-way ANOVA with Sidak's multiple comparisons) or 35-min excitatory (top right; 35-min excitatory cerebellum vs.: 10-min mPFC, 0.08 $\pm$ 0.01 correlation coefficient, $p < 0.0001$, 35-min mPFC, 0.10 $\pm$ 0.02 correlation coefficient, $p < 0.0001$, 10-min cerebellum, 0.36 $\pm$ 0.03 correlation coefficient, $p = 0.586$, 35-min cerebellum, 0.31 $\pm$ 0.03 correlation coefficient, one-way ANOVA with Sidak's multiple comparisons) and inhibitory cerebellum neuron populations (bottom right; 35-min inhibitory cerebellum vs.: 10-min mPFC, 0.04 $\pm$ 0.01 correlation coefficient, $p < 0.0001$, 35-min mPFC, 0.04 $\pm$ 0.01 correlation coefficient, $p < 0.0001$, 10-min cerebellum, 0.25 $\pm$ 0.03 correlation coefficient, $p = 0.626$, 35-min cerebellum, 0.21 $\pm$ 0.03 correlation coefficient, one-way ANOVA with Sidak's multiple comparisons) with standard error. For all *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$.

between the mPFC and either cerebellar timepoint (Excitatory mPFC neurons: 35-min mPFC vs. 10-min mPFC, cosine similarity 0.51 $\pm$ 0.06, 35 min mPFC vs. 35 min mPFC cosine similarity 0.53 $\pm$ 0.11, 10 min cerebellum vs. 35 min mPFC cosine similarity 0.27 $\pm$ 0.03, 35 min cerebellum vs. 35 min mPFC cosine similarity 0.13 $\pm$ 0.03, $p < 0.0001$, ANOVA). Similarly, we observed a greater similarity between IEG expression observed in the cerebellum at 10 min and 35 min than between the cerebellum and the mPFC at either time point (Excitatory cerebellar neurons: 35-min cerebellum vs. 10-min cerebellum correlation coefficient 0.48 $\pm$ 0.03, 35 min cerebellum vs. 35 min cerebellum correlation coefficient 0.54 $\pm$ 0.08, 35 min cerebellum vs. 10 min mPFC correlation coefficient 0.18 $\pm$ 0.03, 35 min cerebellum vs. 35 min mPFC correlation coefficient 0.13 $\pm$ 0.03, $p < 0.0001$, ANOVA). We observed similar findings within inhibitory populations in the mPFC and the cerebellum across timepoints (Figures 3D–3G). Taken together these data suggest that regional differences in IEG-driven transcriptional programs exceed cell-type specific differences within a given region.

The heterogeneous social ensemble

The broad similarities in IEG expression in mPFC excitatory and inhibitory neurons following social behavior could be driven by a single transcriptional program shared by multiple cell types, or distinct transcriptional programs utilized by one or more cell types. To examine this question, we identified IEG+ mPFC neurons which expressed 4 or more IEGs at high levels, then clustered these IEG+ neurons based on the expression of IEGs from the original 134 gene panel (Figures 4A and 4B). This generated 12 total clusters. We first examined whether individual IEGs were preferentially utilized by individual or distributed clusters. We found that some IEGs including Nr4a3 and Npas4 belonged primarily to a handful of clusters, whereas others including Fos and Arc were broadly utilized (Figure 4C). We next examined the corollary question of whether individual neuron subtypes predominantly used one or more transcriptional programs. Again, we found evidence for both (Figures 4B–4E; Table S2) with certain cell types utilizing a restricted subset of transcriptional programs and others being more uniformly distributed.

To further examine this question, we examined the expression of individual IEGs in different cell types following social interaction (Figures 5 and S2). This investigation revealed two primary findings. First, individual IEGs were generally expressed only in a relatively small proportion of cells (roughly 10%) following social interaction. Second, IEG expression was not uniform across cell types. For instance, whereas increased Arc expression was observed in the majority of mPFC excitatory neurons, its expression only increased in 2 cerebellar cell types (Figures 5 and S2). Therefore, we did not identify a single IEG which could accurately capture all activated neurons of each subtype. This suggests that using only a single marker gene may be insufficient to compare activation between different types of neurons.

We therefore reasoned that activated cells (IEG+ cells) within different clusters could be more reliably detected using a region-specific panel of IEGs than choosing a single gene. To identify this panel, we utilized all IEGs identified as socially responsive within a given region (defined as > 2-fold change in number of cells expressing). This corresponded to a total of 43 IEGs in the mPFC and 18 IEGs in the cerebellum (Figures 6A and S3A). Some socially responsive IEGs were identified in both cerebellar and mPFC cell-types (Arc, Dusp1, Fos, Ier2, Junb, Nr4a1-3, Per1), others (Klf2, Nfknbid, Pim1, Plk2, Sik2, Arhgef3, Inhba, Nfib, Figure S3) increased significantly only in
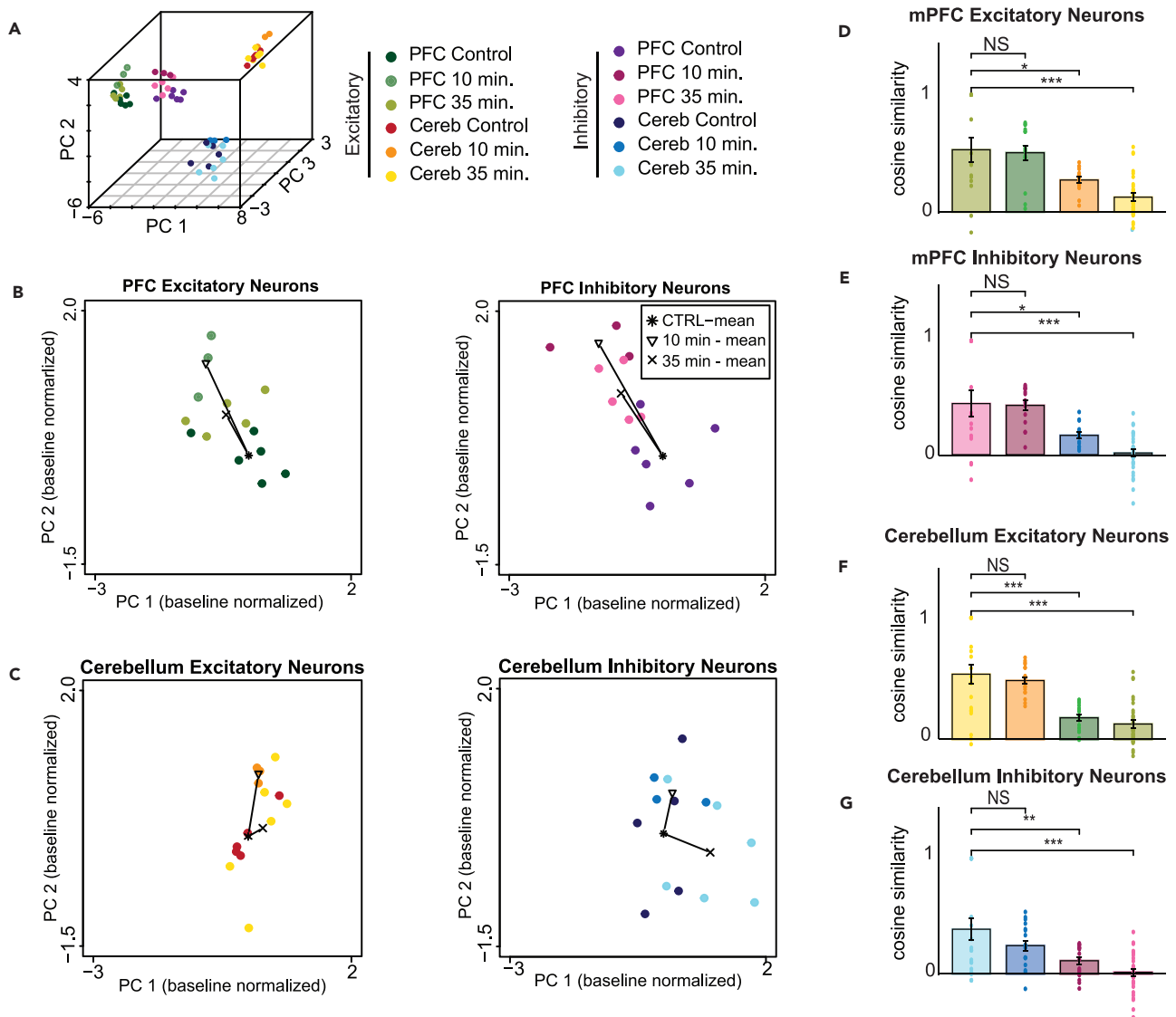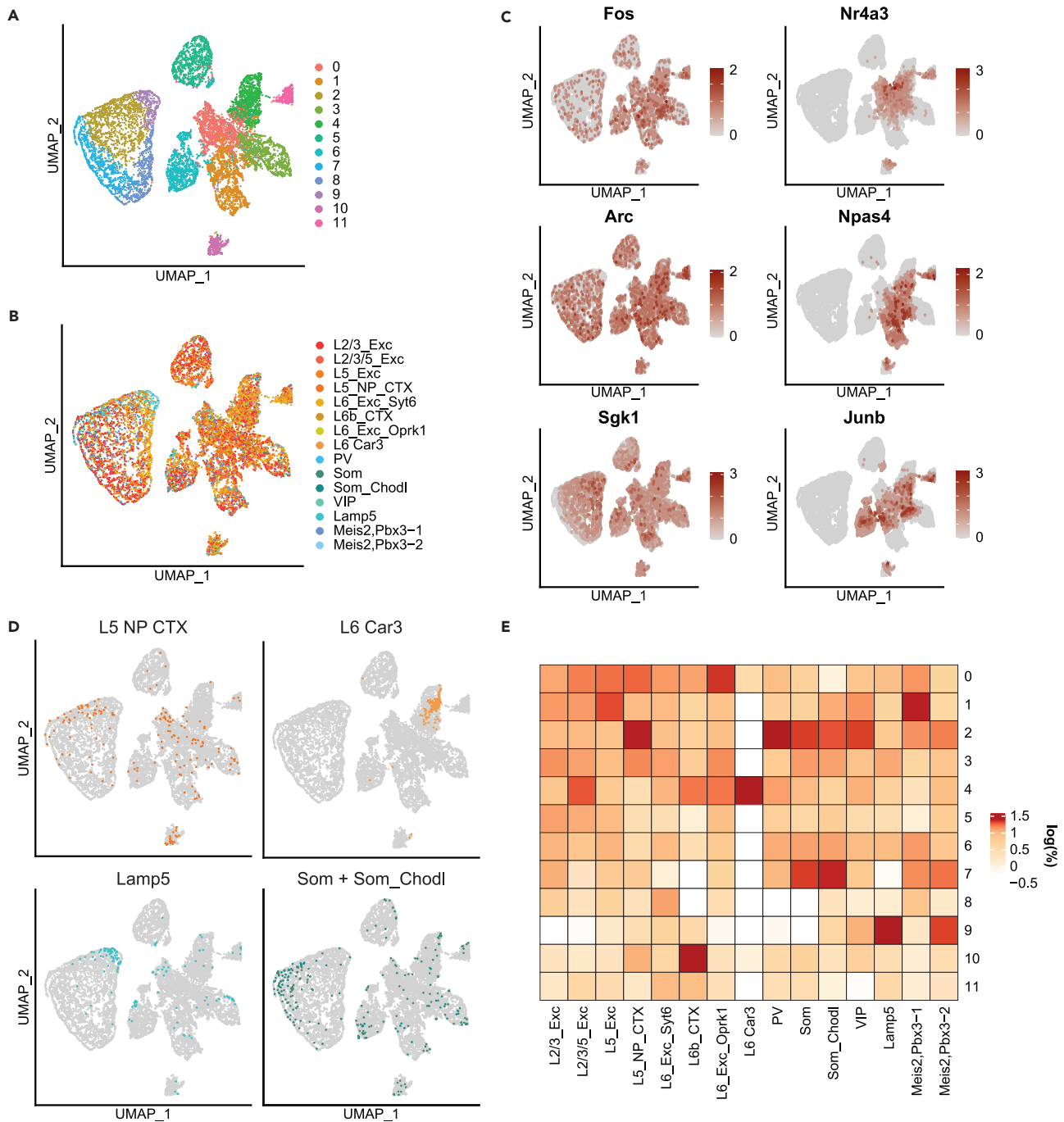
Figure 3. Comparison of region-specific and cell-type specific patterns of IEG expression following social interaction

(A) IEG expression averaged across excitatory or inhibitory neuron types for each region and time point plotted following principal components analysis. Each point represents a sample from an individual mouse plotted in principal component space revealing regional and cell type (excitatory vs. inhibitory) specific differences in baseline IEG expression.

(B and C) All points were normalized to the mean of the control condition. B and C demonstrate normalized values for each mouse colored by control and social conditions. Both the mean position of each condition, and the trajectory between the mean control and social conditions are overlaid, demonstrating similar trajectories for each mPFC timepoint and cell-type, which are distinct from trajectories observed in the cerebellum for excitatory or inhibitory neurons.

(D–G) Bar plots depicting cosine similarity for each social condition compared to all other conditions within the same neuronal subtype for excitatory mPFC (D; 35 min excitatory mPFC, similarity with: 35-min mPFC, 0.53 ± 0.11, 10-min mPFC, 0.51 ± 0.06, 10-min cerebellum, 0.27 ± 0.03, 35-min cerebellum, 0.13 ± 0.03, $p < 0.0001$, ANOVA with Dunnett's multiple comparisons test), inhibitory mPFC (E; 35 min inhibitory mPFC similarity with: 35-min mPFC, 0.45 ± 0.12, 10-min mPFC, 0.43 ± 0.05, 10-min cerebellum, 0.17 ± 0.05, 35-min cerebellum, 0.02 ± 0.03, $p < 0.0001$, ANOVA with Dunnett's multiple comparisons test), excitatory cerebellum (F; 35 min excitatory cerebellum similarity with: 35-min cerebellum, 0.54 ± 0.08, 10-min cerebellum, 0.48 ± 0.03, 10-min mPFC, 0.18 ± 0.03, 35-min mPFC, 0.13 ± 0.03, $p < 0.0001$, ANOVA with Dunnett's multiple comparisons test), or inhibitory cerebellum (G; 35 min inhibitory cerebellum similarity with: 35-min cerebellum, 0.39 ± 0.09, 10-min cerebellum, 0.25 ± 0.04, 10-min mPFC, 0.11 ± 0.03, 35-min cerebellum, 0.02 ± 0.03, $p < 0.0001$, ANOVA with Dunnett's multiple comparisons test). For all *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

cerebellar neurons, and still more were only identified in mPFC datasets (Apold, Bhlhe40, Btg2, Ccn1, Ccn2, Cebpb, Dusp2, Dusp6, Egr1-4, F3, Fosb, Fosl2, Gadd45b, Gadd45g, Gem, Hes1, Ier3, Ier5, Ifit2, Ifit3, Jun, Klf10, Maff, Nfkbia, Npas4, Nrn1, Ptgs2, Sik1, Trib1, Zfp36l1&2; Figure 6).

Figure 4. Cell-type specific transcriptional programs following social interaction

(A) UMAP clustering of IEG+ neurons based on IEG expression.

(B) Colored dots correspond to cell type.

(C) Expression of indicated IEGs within clustered neurons.

(D) UMAPS with colored dots corresponding to indicated cell type.

(E) Heatmap indicates the percentage of each cell type (X axis) belonging to each cluster (Y axis) on a log scale. True percentages are shown in Table S2.

IEG+ cells in the mPFC were defined by high expression of 4 or more socially-responsive IEGs from the mPFC panel (>95th percentile of control expression; Figures 6B and 6C). Note IEG+ mPFC neurons are the same neurons described in Figure 4. We then calculated the proportion of IEG+ cells in each cluster to quantify heterogeneous cell populations activated by social interaction. We observed a significant
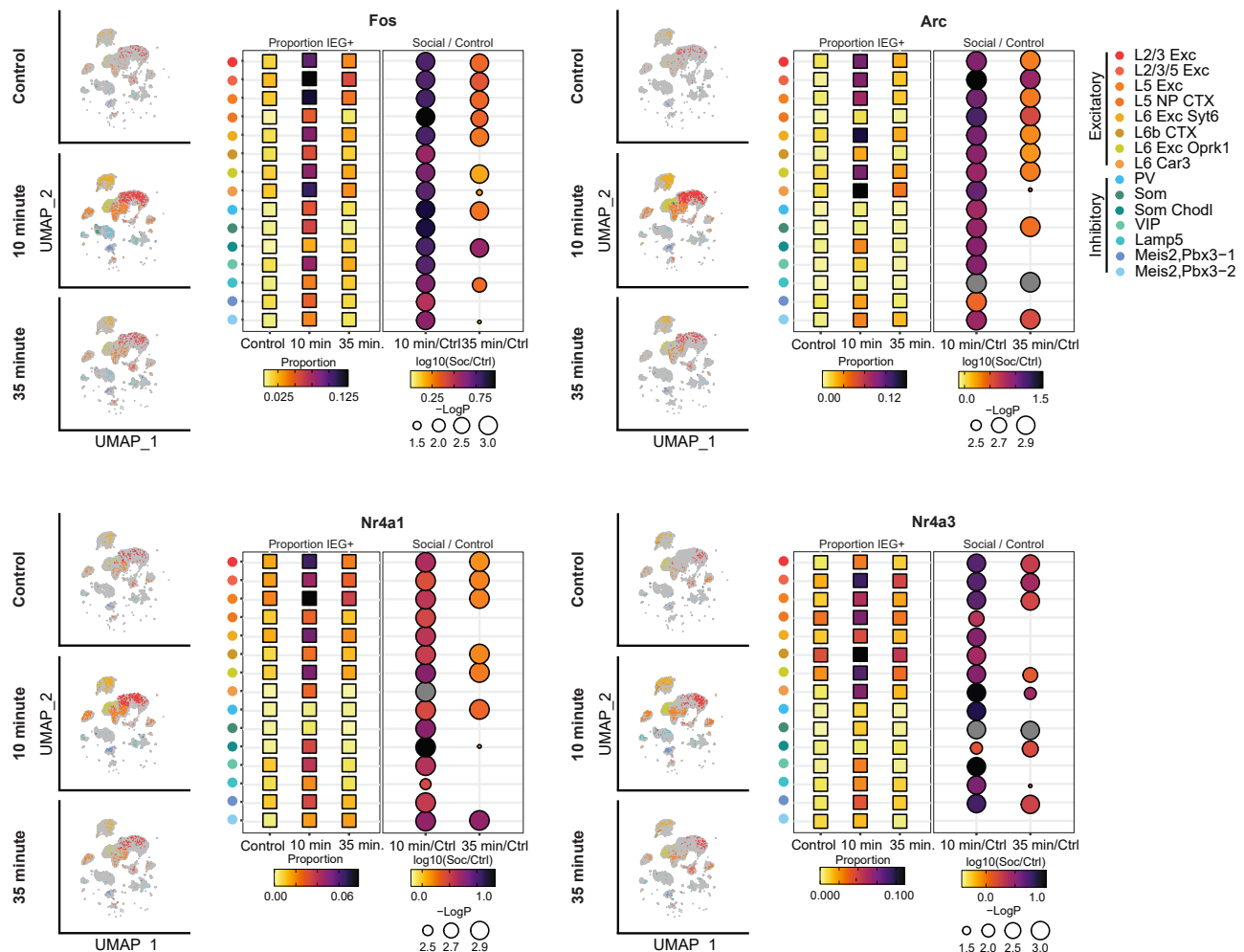
Figure 5. IEG expression in heterogeneous prefrontal neuron populations following social interaction

Cells expressing Fos, Arc, Nr4a1, and Nr4a3 (>95th percentile of control) are plotted for each timepoint and quantified. Individual colored dots represent individual neurons expressing corresponding IEG, overlaid on clustered dataset randomly down sampled to show equivalent cell counts in each condition. Corresponding dot plots represent the proportion of cells expressing the corresponding IEG at 10 min and 35 min post social interaction, and the ratio of the proportions of positive cells in the post-social vs. control conditions (also see Table S3; *p*-values calculated based on shuffled data).

increase in the proportion of IEG+ neurons in excitatory and inhibitory prefrontal cell clusters at both 10 and 35 min following social interaction (compared both to control and shuffled data; Figure 6D).

We defined IEG+ neurons in the cerebellum as those having high expression of 2 or more socially-responsive IEGs (>95th percentile of control expression; Figures S3B and S3C). The proportion of IEG+ cells increased in both excitatory and inhibitory neurons at 10 and 35 min following social interaction (compared to control or shuffled data), but the increase was blunted compared to that observed in the mPFC (10 min mPFC, 12.6 ± 2.8 fold-increase, 10 min Cerebellum, 3.6 ± 0.5 fold-increase, 35-min mPFC 2.5 ± 0.5 fold-increase, 35-min Cerebellum 1.6 ± 0.3 fold-increase, *p* = 0.04 for region & *p* = 0.013 for time point, two-way ANOVA), potentially reflecting restricted or regional recruitment of cerebellar activity during social interaction.

The relative abundance and the proportional increase in IEG+ neurons varied by cell type within both the mPFC and cerebellum. Even within a given layer, we observed unequal recruitment of different cell types (for instance L5 Exc vs. L5 NP CTX, or L6 Car3 vs. L6b CTX; Figure 6D). Examination of the total number of cells of each type which composed the social ensemble also revealed the importance of cell abundance; cerebellar granule cells - due to their abundance - were the largest component of the social ensemble (19.3% of significantly active cells 10-min post social interaction, and 42.1% of significantly active cells 35-min post social interaction; Figure S4).

DISCUSSION

Normal social interactions depend on reliable encoding of social information by the prefrontal cortex. During social interactions, information relevant to these interactions must be received, decoded, and routed through the mPFC to distributed brain regions. Thus, the neuronal
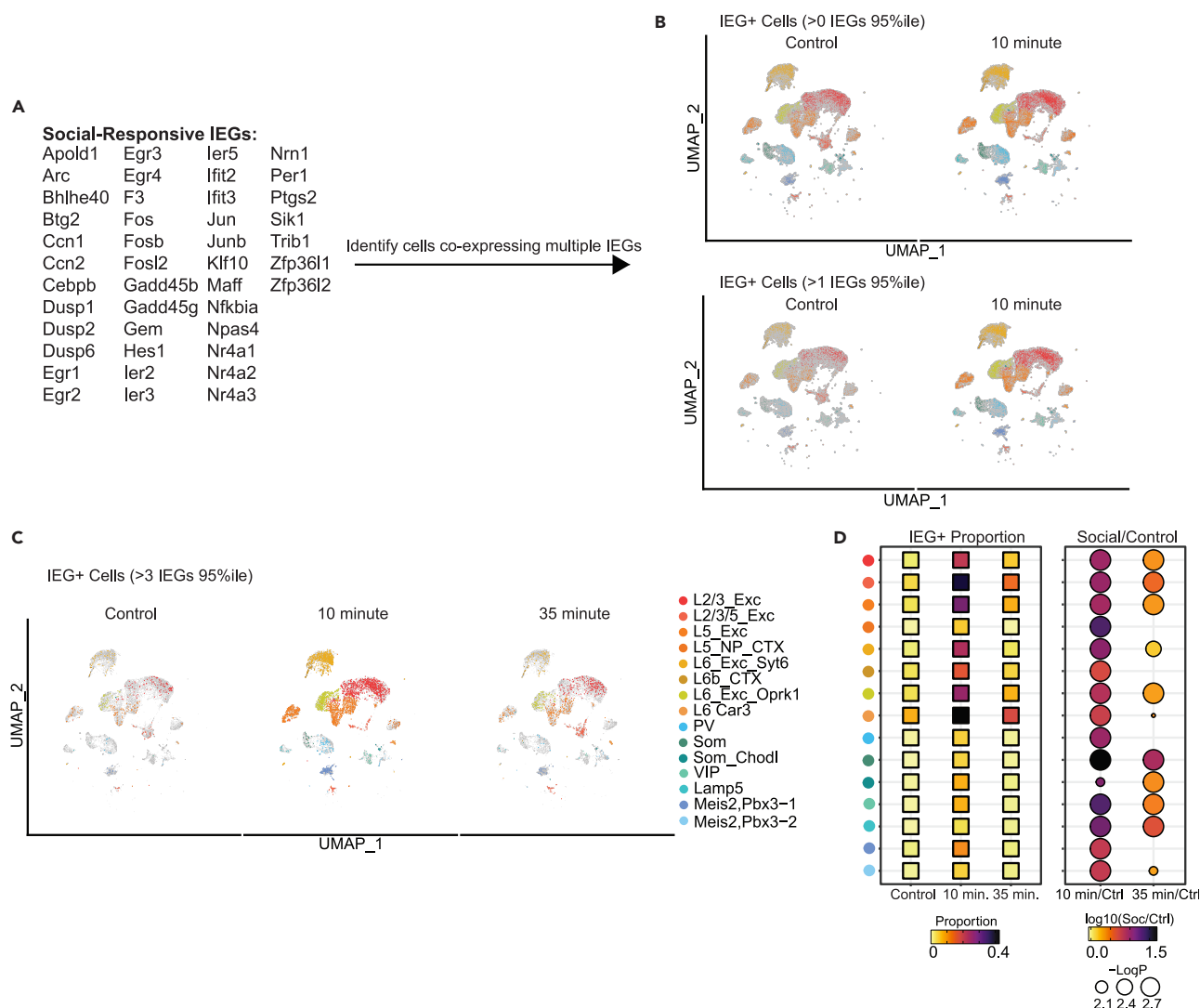
Figure 6. IEG expression defines a heterogeneous social ensemble in the prefrontal cortex

(A) We utilized a panel of 43 socially responsive IEGs in the mPFC (>2-fold expression in social condition compared to control).

(B) UMAPs showing the proportion of cells of each type which express at least a single IEG (top) or 2 or more IEGs (bottom) in the control and 10-min post-social condition.

(C) Proportion of cells expressing 4 or more IEGs above the 95th percentile of expression in the control group. UMAPs are randomly down sampled by condition to show the same number of cells for each condition (B&C).

(D) Boxes on left are colored to represent the proportion of IEG+ neurons in each cell type in control or social conditions (n = 31,290 neurons from 6 control mice, 24,036 neurons from 5 35-min post social mice and 16,121 neurons from 3 10-min post social mice). We then calculated the proportion of IEG+ neurons at each post-social timepoint. Circles are colored to represent the ratio of IEG+ cells in each social condition relative to the control condition. The diameter of each circle represents Log10 p-value (also see Table S4; p-values calculated based on shuffled data).

populations which we define as neuronal ensembles consist of heterogeneous cell types which play distinct roles in the processing and communication[4,19] of information within brain circuits. Cells within the prefrontal microcircuit are nonrandomly connected,[11] and perturbing the recruitment of one cell type may have unexpected consequences on network function. Indeed, altered recruitment of even a single cell type has been demonstrated to alter social behavior,[8,9] presumably by disrupting computations within the prefrontal microcircuit.

This study builds on prior studies which have collectively demonstrated critical features of ensemble representations. First, ensembles are composed of neurons which may be reliably activated or repressed during social interaction,[2,3,6] permitting the generation of social ensembles which are orthogonal to representations of other types of information despite utilizing overlapping cells. This oppositional activation and repression of activity within the mPFC is most likely driven at least in part by recurrent inhibition. While prior studies have implicated specific inhibitory neuronal populations in the regulation of social interactions[7,8,20,21] our data supports a model in which heterogeneous inhibitory neuron populations are recruited alongside excitatory neurons during social interaction.

Second, selective routing of activity by defined projections during complex behaviors including social interactions provides an anatomic substrate for distributed processing of social information as well as the top-down control of social interactions by prefrontal circuits.[22,23] We observe activation of pyramidal neurons across multiple cortical layers, arguing that the anatomic segregation of projections with distinct cortical and subcortical targets is insufficient to achieve specific temporal patterns of recruitment[19] or behavioral encoding[4] observed during social interaction. This selective recruitment may instead be accomplished through distinct patterns of local or distributed inhibitory[11,24–26] and excitatory input.[11] Indeed, the mPFC does not function in isolation, and receives input from various regions including the ventral hippocampus (vHPC). The vHPC encodes information relevant to social interactions[27–29] and anxiety,[5,30] with recent work demonstrating that vHPC projections are composed of at least two populations. These populations convey distinct types of information relevant to anxiety-related behaviors, and synapse onto distinct populations of neurons within the mPFC.[31] The role of these neurons in conveying information about social interactions is not described, but the implication is that even within region-defined inputs, subsets of projecting neurons may differentially drive activation of different populations within the mPFC microcircuit.

Critically, our data emphasize that complex and heterogeneous transcriptional programs are evoked within single neurons. In our dataset, we observed recruitment of multiple IEGs within individual neurons following social interaction, and the population of socially responsive IEGs varied markedly between the prefrontal cortex and cerebellum. Even within regions, individual IEGs varied in their ability to identify heterogeneous populations of neurons. For example, whereas Fos expression increased across both excitatory and inhibitory neurons, Arc expression increased less reliably in inhibitory neuron populations. Notably, even within excitatory or inhibitory populations, multiple transcriptional programs were observed and utilized by different cell types (Table S7). Previous studies have utilized cell-type specific IEG panels to define activated neurons.[14] In this work, we utilized a single broad panel of IEGs to define IEG positivity within each region as we observed that while some mPFC cell types utilized only a handful of transcriptional programs (In particular L6 Car3; see Figure 4), the majority of cell types utilized a broad range of these programs. To a certain degree some of this heterogeneity may be driven by our behavioral paradigm as well as efflux of transcribed IEGs from the nucleus. This rapid efflux unfortunately hinders our ability to link rapid and delayed transcriptional programs which might be triggered by social interactions.

The source and types of information being processed within the mPFC during social interactions are critical but not fully understood, and includes information related to sensory stimuli[32] and internally generated states. These latter states may include information related to recollection or prior experience,[28] social hierarchy[33] or competing motivations.[34] The mPFC also represents information related to novelty and anxiety which we do not address here, however, is likely important in the context of social behavior alterations in disease states such as neurodegeneration and autism spectrum disorder. These states may also result in recruitment of differing subpopulations within the mPFC.

Finally, different inputs may differentially activate distinct populations of prefrontal neurons resulting in the recruitment of distinct computational units or cell populations.[31] Our data suggest that these computations utilize distinct excitatory and inhibitory populations to different degrees, which may be critical for generating distinct representations of distinct types of behavioral information.

Limitations of the study

Social behavior consists of a complex repertoire of social interactions which may be applied in a flexible and context-dependent manner. The behavioral paradigm utilized in this study assesses a variety of social interactions which occur in a single environment following social isolation. We utilized this approach to elicit a strong neuronal response; future studies could utilize similar strategies to examine populations recruited by different sub-behaviors or examine how disease states or circuit perturbation alters this recruitment.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Animals
- METHOD DETAILS
  - Social interaction assay
  - Single-nuclei RNA sequencing
  - Cell type clustering
  - Identification of region-specific panels of immediate-early genes
  - Comparison of IEG expression patterns by region and cell type
  - Quantification of a distributed social ensemble
- QUANTIFICATION AND STATISTICAL ANALYSIS

AUTHOR CONTRIBUTIONS

N.A.F. and H.W. designed the experiments and analysis. H.W. performed all experiments and data analysis. N.A.F. and H.W. wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

1. Gutzeit, V.A., Ahuna, K., Santos, T.L., Cunningham, A.M., Sadsad Rooney, M., Muñoz Zamora, A., Denny, C.A., and Donaldson, Z.R. (2020). Optogenetic reactivation of prefrontal social neural ensembles mimics social buffering of fear. Neuropsychopharmacology 45, 1068–1077. https://doi.org/10.1038/s41386-020-0631-1.

2. Liang, B., Zhang, L., Barbera, G., Fang, W., Zhang, J., Chen, X., Chen, R., Li, Y., and Lin, D.T. (2018). Distinct and Dynamic ON and OFF Neural Ensembles in the Prefrontal Cortex Code Social Exploration. Neuron 100, 700–714.e9. https://doi.org/10.1016/j.neuron.2018.08.043.

3. Frost, N.A., Haggart, A., and Sohal, V.S. (2021). Dynamic patterns of correlated activity in the prefrontal cortex encode information about social behavior. PLoS Biol. 19, e3001235. https://doi.org/10.1371/journal.pbio.3001235.

4. Murugan, M., Jang, H.J., Park, M., Miller, E.M., Cox, J., Taliaferro, J.P., Parker, N.F., Bhave, V., Hur, H., Liang, Y., et al. (2017). Combined Social and Spatial Coding in a Descending Projection from the Prefrontal Cortex. Cell 171, 1663–1677.e16. https://doi.org/10.1016/j.cell.2017.11.002.

5. Adhikari, A., Topiwala, M.A., and Gordon, J.A. (2010). Synchronized activity between the ventral hippocampus and the medial prefrontal cortex during anxiety. Neuron 65, 257–269. https://doi.org/10.1016/j.neuron.2009.12.002.

6. Frost, N.A., Donohue, K.C., and Sohal, V. (2023). Context-invariant socioemotional encoding by prefrontal ensembles. Preprint at bioRxiv. https://doi.org/10.1101/2023.10.19.563015.

7. Bicks, L.K., Yamamuro, K., Flanigan, M.E., Kim, J.M., Kato, D., Lucas, E.K., Koike, H., Peng, M.S., Brady, D.M., Chandrasekaran, S., et al. (2020). Prefrontal parvalbumin interneurons require juvenile social experience to establish adult social behavior. Nat. Commun. 11, 1003. https://doi.org/10.1038/s41467-020-14740-z.

8. Selimbeyoglu, A., Kim, C.K., Inoue, M., Lee, S.Y., Hong, A.S.O., Kauvar, I., Ramakrishnan, C., Fenno, L.E., Davidson, T.J., Wright, M., and Deisseroth, K. (2017). Modulation of prefrontal cortex excitation/inhibition balance rescues social behavior in CNTNAP2-deficient mice. Sci. Transl. Med. 9, eaah6733. https://doi.org/10.1126/scitranslmed.aah6733.

9. Yizhar, O., Fenno, L.E., Prigge, M., Schneider, F., Davidson, T.J., O'Shea, D.J., Sohal, V.S., Goshen, I., Finkelstein, J., Paz, J.T., et al. (2011). Neocortical excitation/inhibition balance in information processing and social dysfunction. Nature 477, 171–178. https://doi.org/10.1038/nature10360.

10. Bhattacherjee, A., Djekidel, M.N., Chen, R., Chen, W., Tuesta, L.M., and Zhang, Y. (2019). Cell type-specific transcriptional programs in mouse prefrontal cortex during adolescence and addiction. Nat. Commun. 10, 4169. https://doi.org/10.1038/s41467-019-12054-3.

11. Jiang, X., Shen, S., Cadwell, C.R., Berens, P., Sinz, F., Ecker, A.S., Patel, S., and Tolias, A.S. (2015). Principles of connectivity among morphologically defined cell types in adult neocortex. Science 350, aac9462. https://doi.org/10.1126/science.aac9462.

12. Hughes, B.W., Siemsen, B.M., Tsvetkov, E., Berto, S., Kumar, J., Cornbrooks, R.G., Akiki, R.M., Cho, J.Y., Carter, J.S., Snyder, K.K., et al. (2023). NPAS4 in the medial prefrontal cortex mediates chronic social defeat stress-induced anhedonia-like behavior and reductions in excitatory synapses. Elife 12, e75631. https://doi.org/10.7554/eLife.75631.

13. Xing, B., Mack, N.R., Guo, K.M., Zhang, Y.X., Ramirez, B., Yang, S.S., Lin, L., Wang, D.V., Li, Y.C., and Gao, W.J. (2021). A Subpopulation of Prefrontal Cortical Neurons Is Required for Social Memory. Biol. Psychiatry 89, 521–531. https://doi.org/10.1016/j.biopsych.2020.08.023.

14. Wu, Y.E., Pan, L., Zuo, Y., Li, X., and Hong, W. (2017). Detecting Activated Cell Populations Using Single-Cell RNA-Seq. Neuron 96, 313–329.e6. https://doi.org/10.1016/j.neuron.2017.09.026.

15. Yao, Z., van Velthoven, C.T.J., Nguyen, T.N., Goldy, J., Sedeno-Cortes, A.E., Baftizadeh, F., Bertagnolli, D., Casper, T., Chiang, M., Crichton, K., et al. (2021). A taxonomy of transcriptomic cell types across the isocortex and hippocampal formation. Cell 184, 3222–3241.e26. https://doi.org/10.1016/j.cell.2021.04.021.

16. Xu, P., Chen, A., Li, Y., Xing, X., and Lu, H. (2019). Medial prefrontal cortex in neurological diseases. Physiol. Genomics 51, 432–442. https://doi.org/10.1152/physiolgenomics.00006.2019.

17. Kozareva, V., Martin, C., Osorno, T., Rudolph, S., Guo, C., Vanderburg, C., Nadaf, N., Regev, A., Regehr, W.G., and Macosko, E. (2021). A transcriptomic atlas of mouse cerebellar cortex comprehensively defines cell types. Nature 598, 214–219. https://doi.org/10.1038/s41586-021-03220-z.

18. Okuno, H. (2011). Regulation and function of immediate-early genes in the brain: Beyond neuronal activity markers. Neurosci. Res. 69, 175–186. https://doi.org/10.1016/j.neures.2010.12.007.

19. Brumback, A.C., Ellwood, I.T., Kjaerby, C., Iafrati, J., Robinson, S., Lee, A.T., Patel, T., Nagaraj, S., Davatolhagh, F., and Sohal, V.S. (2018). Identifying specific prefrontal neurons that contribute to autism-associated abnormalities in physiology and social behavior. Mol. Psychiatry 23, 2078–2089. https://doi.org/10.1038/mp.2017.213.

20. Scheggia, D., Managò, F., Maltese, F., Bruni, S., Nigro, M., Dautan, D., Latuske, P., Contarini, G., Gomez-Gonzalo, M., Requie, L.M., et al. (2020). Somatostatin interneurons in the prefrontal cortex control affective state discrimination in mice. Nat. Neurosci. 23, 47–60. https://doi.org/10.1038/s41593-019-0551-8.

21. Chen, Q., Deister, C.A., Gao, X., Guo, B., Lynn-Jones, T., Chen, N., Wells, M.F., Liu, R., Goard, M.J., Dimidschstein, J., et al. (2020). Dysfunction of cortical GABAergic neurons

leads to sensory hyper-reactivity in a Shank3 mouse model of ASD. Nat. Neurosci. *23*, 520–532. https://doi.org/10.1038/s41593-020-0598-6.

22. Dulka, B.N., Bress, K.S., Grizzell, J.A., and Cooper, M.A. (2018). Social Dominance Modulates Stress-induced Neural Activity in Medial Prefrontal Cortex Projections to the Basolateral Amygdala. Neuroscience *388*, 274–283. https://doi.org/10.1016/j.neuroscience.2018.07.042.

23. Li, A., Jing, D., Dellarco, D.V., Hall, B.S., Yang, R., Heilberg, R.T., Huang, C., Liston, C., Casey, B.J., and Lee, F.S. (2021). Role of BDNF in the development of an OFC-amygdala circuit regulating sociability in mouse and human. Mol. Psychiatry *26*, 955–973. https://doi.org/10.1038/s41380-019-0422-4.

24. Lee, A.T., Gee, S.M., Vogt, D., Patel, T., Rubenstein, J.L., and Sohal, V.S. (2014). Pyramidal neurons in prefrontal cortex receive subtype-specific forms of excitation and inhibition. Neuron *81*, 61–68. https://doi.org/10.1016/j.neuron.2013.10.031.

25. Morishima, M., and Kawaguchi, Y. (2006). Recurrent connection patterns of corticostriatal pyramidal cells in frontal cortex. J. Neurosci. *26*, 4394–4405. https://doi.org/10.1523/JNEUROSCI.0252-06.2006.

26. Liu, L., Xu, H., Wang, J., Li, J., Tian, Y., Zheng, J., He, M., Xu, T.L., Wu, Z.Y., Li, X.M., et al. (2020). Cell type-differential modulation of prefrontal cortical GABAergic interneurons on low gamma rhythm and social interaction. Sci. Adv. *6*, eaay4073. https://doi.org/10.1126/sciadv.aay4073.

27. Okuyama, T., Kitamura, T., Roy, D.S., Itohara, S., and Tonegawa, S. (2016). Ventral CA1 neurons store social memory. Science *353*, 1536–1541. https://doi.org/10.1126/science.aaf7003.

28. Phillips, M.L., Robinson, H.A., and Pozzo-Miller, L. (2019). Ventral hippocampal projections to the medial prefrontal cortex regulate social memory. Elife *8*, e44182. https://doi.org/10.7554/eLife.44182.

29. Tao, K., Chung, M., Watarai, A., Huang, Z., Wang, M.Y., and Okuyama, T. (2022). Disrupted social memory ensembles in the ventral hippocampus underlie social amnesia in autism-associated Shank3 mutant mice. Mol. Psychiatry *27*, 2095–2105. https://doi.org/10.1038/s41380-021-01430-5.

30. Adhikari, A., Topiwala, M.A., and Gordon, J.A. (2011). Single units in the medial prefrontal cortex with anxiety-related firing patterns are preferentially influenced by ventral hippocampal activity. Neuron *71*, 898–910. https://doi.org/10.1016/j.neuron.2011.07.027.

31. Sanchez-Bellot, C., AlSubaie, R., Mishchanchuk, K., Wee, R.W.S., and MacAskill, A.F. (2022). Two opposing hippocampus to prefrontal cortex pathways for the control of approach and avoidance behaviour. Nat. Commun. *13*, 339. https://doi.org/10.1038/s41467-022-27977-7.

32. Levy, D.R., Tamir, T., Kaufman, M., Parabucki, A., Weissbrod, A., Schneidman, E., and Yizhar, O. (2019). Dynamics of social representation in the mouse prefrontal cortex. Nat. Neurosci. *22*, 2013–2022. https://doi.org/10.1038/s41593-019-0531-z.

33. Zhou, T., Zhu, H., Fan, Z., Wang, F., Chen, Y., Liang, H., Yang, Z., Zhang, L., Lin, L., Zhan, Y., et al. (2017). History of winning remodels thalamo-PFC circuit to reinforce social dominance. Science *357*, 162–168. https://doi.org/10.1126/science.aak9726.

34. Jennings, J.H., Kim, C.K., Marshel, J.H., Raffiee, M., Ye, L., Quirin, S., Pak, S., Ramakrishnan, C., and Deisseroth, K. (2019). Interacting neural ensembles in orbitofrontal cortex for social and feeding behaviour. Nature *565*, 645–649. https://doi.org/10.1038/s41586-018-0866-8.

35. Edgar, R., Domrachev, M., and Lash, A.E. (2002). Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res. *30*, 207–210. https://doi.org/10.1093/nar/30.1.207.

36. Ayhan, F., Douglas, C., Lega, B.C., and Konopka, G. (2021). Nuclei isolation from surgically resected human hippocampus. STAR Protoc. *2*, 100844. https://doi.org/10.1016/j.xpro.2021.100844.

37. Lebrigand, K., Magnone, V., Barbry, P., and Waldmann, R. (2020). High throughput error corrected Nanopore single cell transcriptome sequencing. Nat. Commun. *11*, 4025. https://doi.org/10.1038/s41467-020-17800-6.

38. Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W.M., 3rd, Zheng, S., Butler, A., Lee, M.J., Wilk, A.J., Darby, C., Zager, M., et al. (2021). Integrated analysis of multimodal single-cell data. Cell *184*, 3573–3587.e29. https://doi.org/10.1016/j.cell.2021.04.048.

CellPress
OPEN ACCESS

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
Adult C57BL/6J mice	The Jackson Laboratory	Strain #000664
Chemicals, peptides, and recombinant proteins		
Actinomycin-D	Sigma-Aldrich	Cat #A1410-10MG
Pure Prep Nuclei Isolation Kit	Sigma-Aldrich	Cat #NUC201
RNAse inhibitor	New England Biolabs	Cat #M0314S
Ultrapure Bovine Serum Albumin (BSA)	Fisher Scientific	Cat #AM2616
Phosphate Buffered Saline (10X) pH 7.4, DNAse and RNAse free	Fisher bioreagents	BP399500
Critical commercial assays		
Chromium Next GEM single Cell 3′ v3.1	10X Genomics	https://www.10xgenomics.com
Deposited data		
snRNA-seq FASTQ files	This Paper	https://www.ncbi.nlm.nih.gov/geo/query.cgi?acc=GSE269499
Original code	This Paper	https://github.com/nickfrostneuro/single-cell
Software and algorithms		
CellRanger v7.0.1	10x Genomics	https://support.10xgenomics.com/single-cell-gene-expression/software/overview/welcome
R v4.3.0	https://cran.r-project.org/	https://cran.r-project.org
Seurat v4.2.1	Hao et al.[38]	https://github.com/satijalab/seurat/tree/develop
Prism v10.1.1	GraphPad	https://www.graphpad.com
Other		
Flowmi Cell strainers, 40 uM	Fisher Scientific	Cat #14100150
1.5 mL Pestle	Kimble	Cat #749521-1500
Eppendorf Centrifuge 5425R	Eppendorf	https://www.eppendorf.com/us-en/
1.5 mL DNA LoBind Tubes	Eppendorf	Cat #022431021
2 mL Protein LoBind Tubes	Thermo Scientific	Ref #88379
Countess 3 Cell Counter	Invitrogen	https://www.thermofisher.com/us/en/home.html
Countess3 slides	Fisher Scientific	Cat #C10228

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by Nicholas Frost (Nick.Frost@hsc.utah.edu).

Materials availability

This study did not generate new unique materials.

Data and code availability

- Data: All single-nuclei RNA-seq data has been deposited on NCBI's Gene Expression Omnibus[35] and is accessible through accession number GSE269499 (https://www.ncbi.nlm.nih.gov/geo/query.cgi?acc=GSE269499).

- Code: All original code necessary to perform analysis included in this manuscript has been deposited on GitHub (https://github.com/nickfrostneuro/single-cell) and is publicly available.
- Additional Information: Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Animals

Nuclei suspensions from mouse prefrontal cortex (mPFC) and cerebellum were collected from 14 C57BL/6J male and female mice aged 4–5 months. Mice were group housed in our vivarium on a 12-h light dark cycle with *ad libitum* food and water until isolation 3 days prior to tissue collection. Novel juveniles (C57BL/6J) used in social interaction assay were sex-matched and aged 4–6 weeks. All animal use was approved by the University of Utah Institutional Animal Care and Use Committee (IACUC Protocol 21–07014).

METHOD DETAILS

Social interaction assay

Three days prior to the day of sequencing mice were singly housed to reduce social-interaction related neural activity prior to the behavioral assay. In addition, one day prior all mice were habituated to the opening of their cage lid by lifting each lid for about 10 s and then replacing it every 5 min for 2 h. One hour prior to euthanasia, 5 novel sex-matched juvenile conspecifics were sequentially introduced into the home cage of each mouse in the social cohort for a total of 25 min of social interaction. This sequential interaction maximized the amount of time that mice spent investigating the novel juvenile. Aggressive behavior was rarely observed in this paradigm. Control mice were subjected to sequential cage lid removal to match lid openings experienced by the social cohort. Social mice engaged in sequential 5-min blocks of social interactions with 5 novel sex-matched juveniles. At a specified time point following removal of the final mouse (or final lid opening) mice were euthanized by isoflurane overdose immediately followed by decapitation and tissue dissection. To prevent light-cycle induced changes in activity or IEG expression all behavioral experiments were performed between 6:00–7:00 AM. Tissue was therefore harvested between 7:00–8:00 AM for all animals.

Single-nuclei RNA sequencing

Mice were deeply anesthetized by administration of isoflurane in a gas chamber flowing at 5% isoflurane. Anesthesia was confirmed by negative toe pinch response. Following decapitation whole brains were removed and quickly submerged in ice-cold 1X phosphate buffered saline with 15 μM actinomycin-D to prevent artificially induced transcriptional perturbation.[14] Nuclei were isolated using a protocol developed based on previously described methods.[36] Briefly, regions of interest were micro-dissected and transferred into 15 μM actinomycin-D (Sigma-Aldrich #A1410-10MG) containing lysis buffer and homogenized with a pestle. Samples were spun down and supernatant was removed and replaced with new lysis buffer with Triton-X (MilliporeSigma #NUC201) two additional times. Samples were filtered through a 40 μM Flowmi cell strainer (Fisher Scientific #14100150). Nuclei are isolated and myelin was removed using a sucrose gradient (MilliporeSigma #NUC201). Nuclei were then resuspended in a resuspension buffer containing 0.5% Ultrapure BSA (Fisher Scientific #AM2616) to prevent clumping of nuclei and 0.2 U/ul RNAse inhibitor (New England Biolabs #M0314S) to prevent RNA degradation. Samples were spun down, supernatant removed and resuspended in 100 μL resuspension buffer, and filtered a final time through a 40 μM cell strainer. Nuclei counts were calculated on a Countess3 and Nexcellom cell counter and samples were diluted to around 700-1,000 nuclei/uL. All samples remained on ice throughout this process when outside of the centrifuge. Samples were delivered to the University of Utah High-Throughput sequencing core immediately following nuclei isolation for library preparation.

Barcoded snRNA cDNA libraries were generated for each sample with the 10x Genomics Chromium Next GEM Single Cell 3' Gene Expression Library prep v3.1, targeting 7,000–10,000 cells per sample. Libraries were generated following the 10X protocol, apart from extending the PCR elongation step from 1 to 3 min.[37] The prepared cDNA library was then sequenced on the Illumina NovaSeq 6000. We generated 200–300 M reads per sample and mapped them to the GRCm39 with the 10x Genomics Cell Ranger software (v4.0.8).

Cell type clustering

Filtered gene expression matrices generated by Cell Ranger were clustered using Seurat (v4.2.1) in R (v 4.3.0).[38] Nuclei with fewer than 800 features or greater than 6000 features were removed to eliminate low quality nuclei and multiplets. In addition, nuclei which had greater than 5% mitochondrial genes were removed. Next, we performed dimensionality reduction by principal component analysis. We utilized the top 30 principal components to calculate the shared nearest neighbor at 0.08 resolution, followed by UMAP dimensionality reduction and visualization. Samples from the mPFC and cerebellum were clustered independently. Each sample retained its sample identity and social status. Any clusters that did not appear in a minimum of 75% of samples or in which 75% or more of the cells belonged to a single animal were removed. Clusters were assigned to known cell types using known cell type specific marker genes.

Identification of region-specific panels of immediate-early genes

We started by examining 134 genes previously identified as immediate-early genes.[14] We identified cells which contained at least one copy of a given gene and considered a given IEG socially-responsive if the number of identified neurons increased within any cluster by at least 2-fold

(with 1% added to denominator) following social interaction (relative to control mice) within pooled experimental groups. Most clusters of cells demonstrated a social interaction-dependent increase of some, but not all IEGs. Because we observed marked region-specific differences in socially responsive IEGs we utilized region-specific panels which included IEGs which displayed a greater than 2-fold increase in expression following social interaction. IEG+ neurons in the mPFC were defined as cells expressing four or more socially-responsive IEGs (43 genes) at above the 95th percentile control level, and IEG+ neurons in the cerebellum were defined as those neurons which expressed two or more socially-responsive IEGs (18 genes) at above the 95th percentile.

Comparison of IEG expression patterns by region and cell type

To compare the pattern of IEG expression changes driven by social interaction we then determined the fold-increase in the proportion of IEG+ neurons within each identified population of neurons for socially-responsive IEG (using 51 IEGs found to be socially-responsive in either the mPFC or cerebellum). As some smaller clusters had 0 IEG+ cells in the control condition we conservatively added 1 percent to the numerator and denominator. We averaged the fold-difference calculated for each social mouse compared to every control mouse, then calculated and plotted the average across social mice for each condition as a heatmap of neuron types (y) and IEGs (x) sorted by average fold-increase observed in the mPFC. We generated similarity matrices corresponding to Pearson's pairwise correlations calculated between vectors containing fold-increases in expression (social/control) of each IEG for each mPFC and cerebellar cell type at 10 and 35 min which was then averaged across samples.

To further delineate cell-type or regional differences in socially-responsive IEGs we generated counts matrices of socially-responsive IEGs pooled by excitatory or inhibitory neurons from each animal, to generate a matrix corresponding to the expression of all socially-responsive IEGs (51 in total) across all control, 10 min post-social, and 35 min post-social animals. We next performed dimensionality reduction via PCA analysis. We found that the first three principal components explained over 80% of the variance. Visualization of the position of each sample in 3-dimensional component space revealed that samples clustered by region and cell-type. To examine how social interaction alters IEG expression we separated each group of samples by region and cell-type (excitatory or inhibitory) and normalized their position by subtracting the mean control position from each sample to visualize the trajectory of social interaction-driven IEG expression. We then compared the trajectories between mice both within and across condition (cell-type and region) by calculating the cosine similarity between each pair of vectors.

$$\cos \theta = \frac{Vector\ 1 \cdot Vector\ 2}{||Vector\ 1||\ ||Vector\ 2||}$$

To examine IEG+ neurons for distinct transcriptional patterns, IEG+ neurons (i.e., those which expressed 4 or more IEGs at >95%ile) were clustered based on the 134 IEG panel. This resulted in 12 distinct clusters (0–11) which differed both in IEG expression and member neuron populations.

Quantification of a distributed social ensemble

We quantified the composition of social ensembles at 10 and 35 min by determining the excess proportion of IEG+ neurons in each cell type by dividing the proportion of IEG+ neurons in the control condition by the proportion of neurons identified in the same cluster following social interaction, then normalizing to the number of cells identified assigned to the cluster in the social condition.

QUANTIFICATION AND STATISTICAL ANALYSIS

Differences between social and control group IEG positive proportions were determined by comparing the real proportional increase between experimental and control groups to a distribution of shuffled values for each cell type. For shuffling, the number of real cells belonging to each experimental group were selected from within each cell type and the proportion of IEG+ cells in each randomly selected cell population were calculated. The same comparisons were then calculated for the shuffled data as was done in real data. This was repeated 1000 times to create a distribution of 1000 shuffled values for each cell type. To compare patterns of IEG expression across cell types, we calculated Pearson's pairwise correlation coefficients and the cosine similarity as described, then performed one-way ANOVA with multiple comparisons to calculate differences between groups using GraphPad Prism (v 9.4.1). Two-way ANOVAs were also calculated in GraphPad Prism.