SPECIAL FEATURE

NEUROSCIENCE

Subregion-specific rules govern the distribution of neuronal immediate-early gene induction

Ben Jerry Gonzales[a,1], Diptendu Mukherjee[a,1], Reut Ashwal-Fluss[a,b], Yonatan Loewenstein[a,b,c], and Ami Citri[a,b,d,2]

[a]Institute of Life Sciences, The Hebrew University of Jerusalem, Edmond J. Safra Campus, Givat Ram, Jerusalem 91904, Israel; [b]The Edmond and Lily Safra Center for Brain Sciences, The Hebrew University of Jerusalem, Edmond J. Safra Campus, Givat Ram, Jerusalem 91904, Israel; [c]Federmann Center for the Study of Rationality, The Hebrew University of Jerusalem, Jerusalem 91904, Israel; and [d]Program in Child and Brain Development, Canadian Institute for Advanced Research, Toronto, ON M5G 1M1, Canada

Edited by Gene E. Robinson, University of Illinois at Urbana–Champaign, Urbana, IL, and approved October 1, 2019 (received for review August 18, 2019)

The induction of immediate-early gene (IEG) expression in brain nuclei in response to an experience is necessary for the formation of long-term memories. Additionally, the rapid dynamics of IEG induction and decay motivates the common use of IEG expression as markers for identification of neuronal assemblies ("ensembles") encoding recent experience. However, major gaps remain in understanding the rules governing the distribution of IEGs within neuronal assemblies. Thus, the extent of correlation between coexpressed IEGs, the cell specificity of IEG expression, and the spatial distribution of IEG expression have not been comprehensively studied. To address these gaps, we utilized quantitative multiplexed single-molecule fluorescence in situ hybridization (smFISH) and measured the expression of IEGs (*Arc*, *Egr2*, and *Nr4a1*) within spiny projection neurons (SPNs) in the dorsal striatum of mice following acute exposure to cocaine. Exploring the relevance of our observations to other brain structures and stimuli, we also analyzed data from a study of single-cell RNA sequencing of mouse cortical neurons. We found that while IEG expression is graded, the expression of multiple IEGs is tightly correlated at the level of individual neurons. Interestingly, we observed that region-specific rules govern the induction of IEGs in SPN subtypes within striatal subdomains. We further observed that IEG-expressing assemblies form spatially defined clusters within which the extent of IEG expression correlates with cluster size. Together, our results suggest the existence of IEG-expressing neuronal "superensembles," which are associated in spatial clusters and characterized by coherent and robust expression of multiple IEGs.

neurons | experience | inducible transcription | immediate-early genes | cell assemblies

The induction of activity-dependent transcription programs in the brain leads to long-lasting cellular adaptations necessary for the commitment of experiences to long-term memory (1–6). The primary response to neuronal stimulation involves the induction of a set of immediate-early genes (IEGs), which predominantly encode early-response transcription factors (7–10). This early wave of transcription has been shown to be responsible for the induction of a second wave of late-response genes, which include effector proteins, responsible for effecting changes in the neuronal circuits supporting learning and homeostasis (11, 12). IEGs are induced in the brain by a wide range of stimuli and have been implicated in many biological responses, including sensory and motor experience, spatial exploration, stress, and pharmacological intervention (13–22). The study of the IEG response has largely been performed in cultured neurons, bulk tissue, or genetically defined neuronal populations (23–28), leading to significant progress in understanding the regulation of IEG transcription (6) and its role in defining neural circuit function (12, 29). Progress has also been made in describing the segregation of IEG expression induced by discrete experiences (13, 28, 30–32), as well as defining the relationship between neuronal activity and transcription (25, 26, 33). Thus, the recent experiences of individual mice can be decoded from a minimal transcription signature comprising a handful of IEGs across multiple brain structures (25).

Despite this progress, major gaps remain in understanding how IEG expression is distributed within neuronal assemblies, i.e., populations of neurons in a brain region that are responsive to an experience (34). Uncovering the rules governing the distribution of IEG expression is expected to provide insight into the roles of IEG-expressing neuronal assemblies, as well as the mechanisms supporting their recruitment. Specific unanswered questions include the following: 1) Is IEG expression within neuronal populations coherent? i.e., are IEGs coexpressed across individual neurons, and is the expression of IEGs within individual neurons quantitatively correlated (15)? The role of IEGs will vary greatly depending on the coherence of their expression, since the function of individual genes may differ when coexpressed with different cohorts of other inducible genes. Gradation of IEG expression may also result in higher complexity of combinatorial action than could be expected from binary/all-or-none induction (1, 34, 35). 2) Does the IEG response exhibit cell type specificity? Are IEGs primarily expressed within one cell type? Are the rules governing cell specificity of IEG induction uniform within the tissue? The impact of IEG expression will vary depending on the type of neurons within which they are expressed (36, 37). Furthermore, differential

Significance

We are the cumulative product of past experiences, which define our interaction with the world. Experiences leave their mark, at least in part, through the induction of immediate-early genes (IEGs) within neuronal assemblies. However, the rules governing the distribution of IEG expression within neuronal assemblies have not been comprehensively studied. To this end, we applied single-molecule fluorescence in situ hybridization (smFISH) on striatal tissue following cocaine experience and analyzed a published dataset of single-cell RNA sequencing from cortical tissue following light exposure. We report principles governing IEG induction and identify region-specific rules defining cell-specific recruitment. Our results suggest the existence of "superensembles" of neurons found in spatially defined clusters and characterized by the coherent and robust expression of multiple IEGs.

Author contributions: B.J.G., D.M., R.A.-F., and A.C. designed research; B.J.G., D.M., and R.A.-F. performed research; R.A.-F. and Y.L. contributed new reagents/analytic tools; B.J.G., D.M., and R.A.-F. analyzed data; and B.J.G., D.M., and A.C. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]B.J.G. and D.M. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: ami.citri@mail.huji.ac.il.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1913658116/-/DCSupplemental.

expression within neuronal subtypes may be suggestive of circuit-level regulation of the inputs driving IEG induction. 3) Spatial aspects of IEG induction are also understudied—at the global scale of tissue organization, are macroscale "hot spots" of inducible transcription found, within which IEG expression is enriched? At the local, mesoscale level, are IEG-expressing neuronal assemblies clustered or uniformly distributed? This information is expected to be indicative of the existence of functional domains and provide insight regarding circuit mechanisms regulating the recruitment of neuronal assemblies.

The development of quantitative single-cell measurements of multiplexed gene expression with either high coverage (38) or high spatial resolution (39) support the investigation of these questions. Using data obtained from such quantitative approaches, we studied combinatorial IEG expression patterns at single-cell and single-copy resolution within the dorsal striatum following acute cocaine experience. Induction of IEG expression in the dorsal striatum has been previously shown to underlie the maladaptive plasticity induced by cocaine (36, 40–42), but the cellular distribution of IEG induction within the dorsal striatum has not been comprehensively studied. We observe that cocaine-induced IEG expression is graded and coherent. Addressing the relevance of our observations to other brain structures and stimuli, we analyzed an independent published dataset of single-cell RNA sequencing (scRNA-seq) of neurons from mouse visual cortex after exposure to light and observed graded coherent expression within this dataset as well (38).

The principal neuron type in the striatum is the spiny projection neuron (SPN), which is predominantly associated with 2 discrete pathways that are activated in a coordinated fashion during the execution of defined motor actions, with complementary, and largely opposing roles (43, 44). Neurons of the direct pathway (projecting directly to the midbrain) express the *Drd1* dopamine receptor (*Drd1*-SPNs), while *Drd2*-expressing neurons comprise the indirect pathway, projecting through the globus pallidus to the midbrain (43). A number of recent publications have alluded to a mesoscale local circuit organization of the striatum, whereby adjacent SPNs may form functional clusters (45–48). We addressed the pathway specificity of striatal IEG expression following cocaine, as well as macroscale and mesoscale organizational features. We observe region-specific rules governing the allocation of IEG expression among SPN subtypes, and clustering of IEG-expressing cell assemblies. Taken together, our observations suggest the existence of "superensembles" of neurons recruited by experience and found in spatially defined clusters, characterized by the coherent and robust expression of multiple IEGs.

Results

Transcriptionally Defined Cocaine Cell Assemblies Display Coherent IEG Expression. We exposed mice to acute cocaine (20 mg/kg; i.p.) following extended habituation, observing the expected reliable increase in locomotion (Fig. 1*A*) and IEG expression in the dorsal striatum (Fig. 1 *B–D* and *SI Appendix*, Fig. S1). Applying multicolored single-molecule fluorescence in situ hybridization (smFISH), we analyzed the expression of the IEGs *Arc*, *Egr2*, and *Nr4a1* in the striatum. These genes were selected as they had been found in a previous study to be coinduced following acute exposure to cocaine (25). Quantitative smFISH analysis demonstrated a clear trend of IEG induction by cocaine, both at the level of average puncta expression per cell (*SI Appendix*, Fig. S1*A*) as well as in the proportion of neurons displaying robust expression of each gene (*SI Appendix*, Fig. S1*B*; robust expression corresponds to the top one-third of expression under control conditions). Addressing the pairwise expression of the 3 IEGs, we observed positive baseline correlation of all 3 possible pairings, which increased following cocaine administration (Fig. 1*E*). This correlated expression is apparent when plotting the expression per cell of all 3 genes (Fig. 1*F*; control, $r^2 = 0.60$; cocaine, $r^2 = 0.68$). Addressing the gradation of IEG induction, we visualized the average expression of *Egr2* as a function of incremental rise of *Nr4a1* and *Arc* across expression bins (Fig. 1*G*). Cocaine drives a clear shift in the population coexpressing the 3 IEGs, from low levels of expression of *Nr4a1* and *Arc* at baseline (centered around 2 to 10 puncta per cell) to higher expression (centered around 5 to 20 puncta per cell), while maintaining a continuous gradation in expression (Fig. 1*G*). Thus, our data demonstrate coherent IEG induction over a wide range of expression levels following cocaine, rather than an all-or-none shift in expression, consistent with reports of graded IEG induction by experiences (49–52). However, while expression was largely correlated among the 3 IEGs, they differed in their characteristics of induction within cocaine-responsive cell assemblies: *Egr2* exhibited the lowest basal expression levels, and the highest fold-induction by cocaine, identifying it as a sensitive indicator of cocaine-engaged cell assemblies in the striatum (Fig. 1 and *SI Appendix*, Figs. S1 and S2).

Addressing macroscale spatial structure within the data, we mapped the cells displaying robust expression of each of the IEGs assayed within the striatum. Intriguingly, we observed that, following cocaine, these cells displayed clear patterns of spatial distribution, repeatedly localized to the medial and ventrolateral aspects of the striatum. In order to delineate the regions within which the densest populations of neurons robustly expressing each IEG are found, we performed a 2D kernel density estimate. The overlap of the identified high-density subregions revealed 2 "hot spots" of induced expression, corresponding to the medial (MS) and ventrolateral (VLS) aspects of the striatum (Fig. 1*H* and *SI Appendix*, Fig. S2). Thus, striatal cell assemblies encoding the experience of acute cocaine exhibit graded, coherent IEG expression, and are enriched within spatially discrete regions of the striatum.

Cortical Cell Assemblies Responding to Light Exhibit Coherent IEG Expression. In order to address whether graded coherence of IEG induction is a principle of broad relevance to IEG expression within diverse types of neuronal assemblies, we interrogated a comprehensive dataset of scRNA-seq of neurons in the visual cortex of mice stimulated by exposure to light after dark housing, which provided high-quality data for a large number of neurons (38). The published analysis of this dataset demonstrated graded and correlated induction of IEGs within cortical neurons, corresponding to our observations regarding cocaine-induced IEG expression within striatal neuronal assemblies. We analyzed this scRNA-seq dataset with the objective of addressing the coherence of IEG induction, and the correlation of IEG coherence with expression levels. We first identified 13 IEGs demonstrating significant induction across experimental replicates (Fig. 2*A*). Addressing the coexpression of IEGs, we observed the anticipated enrichment of coexpression of IEG pairs (expressing >1 read per IEG per cell) across single excitatory cortical neurons following stimulation (Fig. 2*B*). In order to estimate the coherence of IEG coexpression, neurons were grouped by the number of IEGs they coexpressed. Since the detection of IEG coexpression can be heavily impacted by sequencing depth, we randomly selected subpopulations of neurons from each of these groups exhibiting similar distributions of sequencing depth (*SI Appendix*, Fig. S3). This dataset was used to calculate the mean expression of each IEG within each group of neurons expressing different numbers of IEGs (1–13). Plotting the mean expression of each IEG (row) as a function of the number of coexpressed IEGs (column), demonstrated a strong correlation between the coherence of IEG expression and the expression levels of each IEG found within it (Fig. 2*C* and *SI Appendix*, Table S1). Minimal expression of each of the 13 IEGs was observed when they were expressed independently in single neurons, whereas maximal expression for each IEG was observed in neurons coexpressing

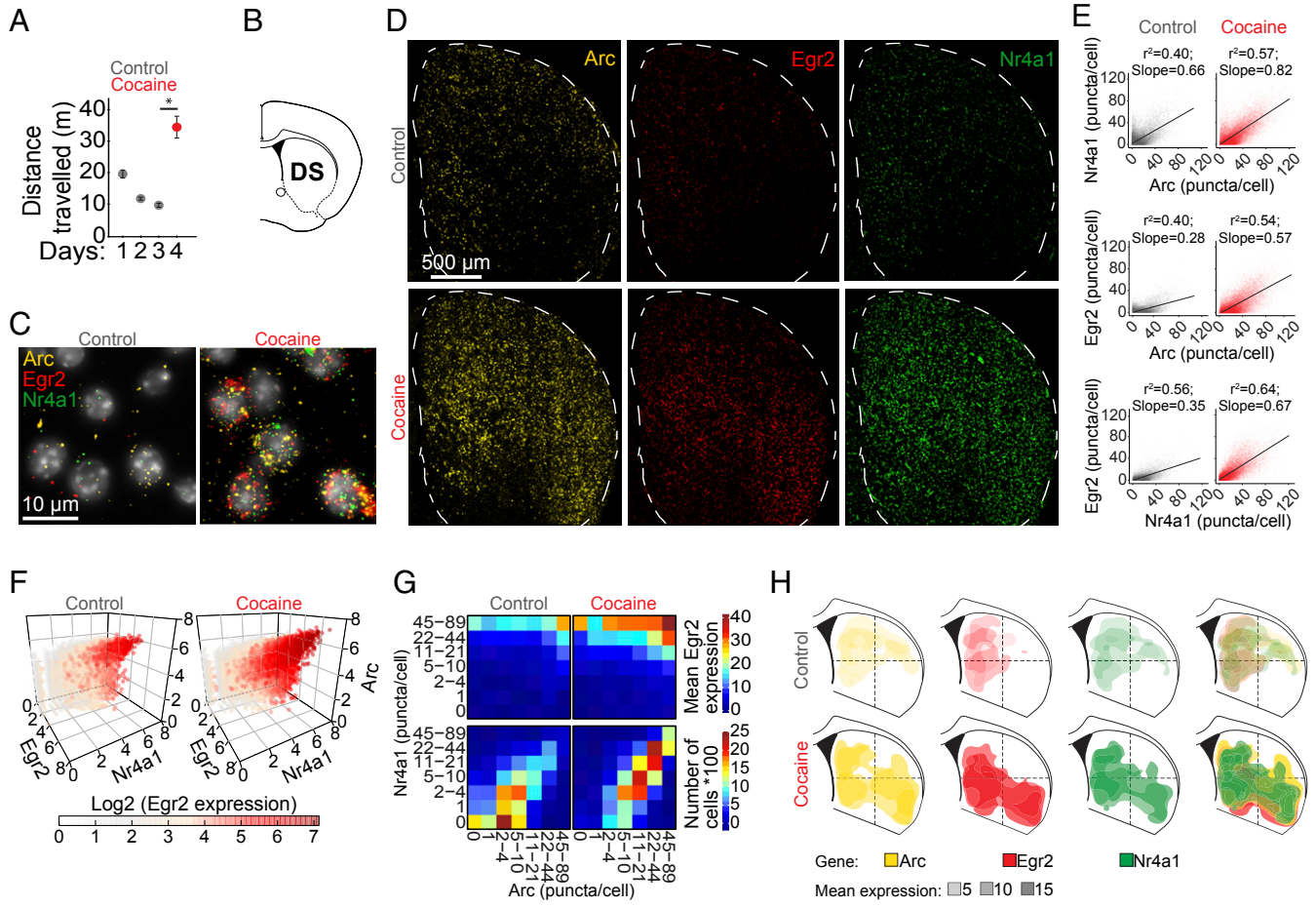

Fig. 1. Cocaine induces coherent IEG expression in subregions of the dorsal striatum. (*A*) Acute exposure to cocaine (i.p.; 20 mg/kg) drives increased locomotion ($n = 3$ mice cocaine/control; $^*P < 0.005$ Student's *t* test; error bars denote SEM). (*B*) Coronal sections through the dorsal striatum (DS) (~0.52 ± 0.1 mm from bregma) are assayed by multicolor smFISH for cocaine-induced IEG expression. (*C*) Cocaine induces expression of multiple IEGs in the dorsal striatum. High-magnification 40× images depicting expression of *Arc*, *Egr2*, and *Nr4a1* in control vs. 1 h following acute cocaine. (*D*) Representative images demonstrating cocaine induced expression of *Arc*, *Egr2*, and *Nr4a1* expression throughout the DS. (*E*) The expression of multiple IEGs in the dorsal striatum is positively correlated in control, and cocaine further increases this correlation (i.e., "coherence"). Individual dots depict pairwise correlations in the expression of *Arc*, *Egr2*, and *Nr4a1* within single cells. $n = 3$ sections from 3 mice (29,649 to 30,102 cells). $P < 0.0001$. (*F*) Expression of *Arc*, *Egr2*, and *Nr4a1* is induced coherently, such that high expression of each gene is correlated with high expression of the other 2 genes, within single cells. Each dot depicts expression of *Arc*, *Egr2*, and *Nr4a1* within a single cell [$\log_2$(puncta per cell + 1)]. The plot is color coded by *Egr2* expression from white (low) to red (high). $P < 0.0001$. (*G*) Heatmaps portraying the distribution of bins of cells expressing *Arc* and *Nr4a1* in control vs. cocaine (*Bottom*) as well as the extent of *Egr2* expression within these bins (*Top*) (control, 30,102 cells; cocaine, 29,649 cells). (*H*) Two-dimensional kernel density estimation was used to demarcate regions of maximal density of high expressing cells (for each IEG), demonstrating hot spots of high IEG expression in the medial and ventrolateral regions of the striatum. Yellow, *Arc*; red, *Egr2*; green, *Nr4a1*. The opacity of the demarcated areas corresponds to the mean per-cell expression.

all 13 IEGs. In contrast, the expression of housekeeping genes remained stable across the different groups (Fig. 2*C*). Taken together, these analyses demonstrate that IEG-expressing neuronal assemblies exhibit a gradation of IEG expression, and the expression levels of individual IEGs strongly correlate with the coherence of IEG coexpression.

Region-Specific Rules Govern Allocation of IEG Expression to SPN Subtypes. While scRNA-seq data, such as that analyzed in Fig. 2, can provide powerful insight regarding coherent coexpression of multiple IEGs, it lacks spatial resolution, which is a strong suit of multiplexed smFISH. Returning to the striatal smFISH data following cocaine exposure, we focused our analysis on the VLS and MS aspects of the striatum and addressed the cell type specificity of IEG induction by studying the correlation of IEG expression with that of markers of SPN subtypes (*Drd1*$^+$ [*Drd1a* paralog] and *Drd2*$^+$; Fig. 3*A*); the threshold to define *Drd1*$^+$ and *Drd2*$^+$ subpopulations was set at ≥8 puncta, with each population corresponding to ~45% of the total population of striatal neurons, on par with the majority of the literature describing the distribution of SPN subtypes in the striatum (53–55). We observed an interesting contrast in the identity of cell assemblies engaged within these 2 substructures. In the VLS, following cocaine, *Egr2* expression was strongly correlated with *Drd1* expression, and not correlated with *Drd2* (Fig. 3 *B* and *D* and *SI Appendix*, Fig. S4*A*). In contrast, in the MS, Egr2 induction by cocaine was correlated with both *Drd1* and *Drd2* expression, albeit more strongly with *Drd1* (Fig. 3 *C* and *E* and *SI Appendix*, Fig. S4*B*). Notably, we observed no cocaine-driven changes in the fraction of *Drd1*- or *Drd2*-expressing striatal cells, their *Drd1* or *Drd2* expression levels, or the distribution of per cell expression (*SI Appendix*, Fig. S5 *A* and *B*). Also, in accordance with previous studies (56), we observed *Drd1* expression to be largely anticorrelated to *Drd2* expression in both the MS and VLS (*SI Appendix*, Figs. S4 and S5 *C* and *D*). These results demonstrate that, within the striatum, whose cellular composition is relatively homogeneous,

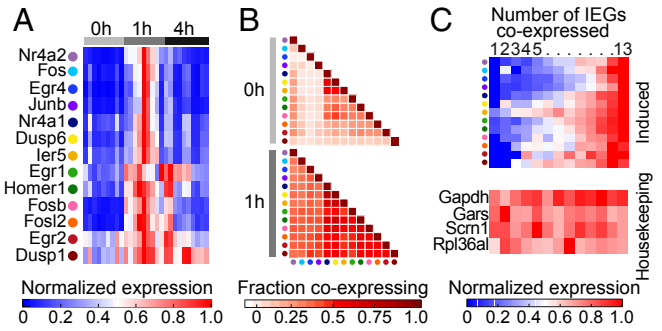

Fig. 2. Coherent and graded IEG induction in excitatory neurons of mouse visual cortex. (*A*) Selection of IEGs displaying maximal induction at 1 h. The heatmap portrays the induction dynamics of 13 significantly induced IEGs (color coded; FDR-corrected $P < 0.05$) in visual cortex following light exposure. Multiple experimental replicates are included for each condition (0, 1 h, and 4 h following light exposure). Expression is normalized to the peak expression for each gene, scaled from low (blue) to maximal (red). (*B*) Expression matrix illustrating the fraction of excitatory neurons coexpressing each pair of IEGs, graded from white (low) to red (high). (*C*) Expression of individual IEGs scales with the number of IEGs coexpressed, while expression of housekeeping genes remains stable regardless of IEG coexpression. The heatmaps depict the mean expression of each gene (row) as a function of the number of IEGs coexpressed (column). Expression is normalized to the peak expression for each gene, scaled from low (blue) to maximal (red).

subregion-specific rules govern the identity of neuronal subtypes recruited to express IEGs following experience.

Mesoscale Spatial Organization of Homotypic SPNs and IEG-Defined Assemblies. Recent studies have alluded to possible mesoscale local circuit organization of the striatum, whereby adjacent SPNs may form functional clusters supporting specific motor actions (45–48). We queried our dataset of high-resolution smFISH for the existence of mesoscale structure. In order to address local spatial clustering of cells of homogeneous character, we focused on "robust expressors," i.e., cells expressing IEG markers *Egr2*, *Arc*, or *Nr4a1* to levels defined by the top one-third of expression under control conditions. For these robust expressors, we calculated the fraction of neighboring cells expressing the same gene ("homotypic"; Fig. 4*A*). Comparing fractions across increasing distances, we observed that neurons expressing high levels of the genes studied herein display a higher probability to be found adjacent to homotypic neurons (Fig. 4*B* and *SI Appendix*, Fig. S6*A*). Notably, a similar organization was observed when separately assessing the MS and VLS (*SI Appendix*, Fig. S7*A*). Shuffling the expression data between striatal cells significantly decreased the frequency of positive neighbors, demonstrating that mesoscale organization is a feature of specific structure within the data (Fig. 4*B* and *SI Appendix*, Figs. S6 and S7 *A* and *B*). In order to estimate the size of these clusters, we selected cells within homogeneous neighborhoods (homotypic cells) and associated them in clusters if they were found within 20 μm of each other. For SPN markers *Drd1* and *Drd2*, we observed a high frequency of relatively small clusters (<10 cells/cluster). In contrast, large *Egr2* clusters were more frequent, reaching up to 40 to 60 cells in the largest clusters (Fig. 4 *C* and *D*). Furthermore, for *Egr2*, cluster size correlated with mean *Egr2* expression level, while the expression of *Drd1* and *Drd2* did not relate to cluster size (Fig. 4*E*). *Nr4a1*$^+$ neurons and *Arc*$^+$ neurons exhibited mesoscale clustering similar to that observed for *Egr2*, as well as a trend toward correlation of IEG expression with cluster size (*SI Appendix*, Figs. S6 and S7). Our results imply a mesoscale organization of striatal neurons, such that SPNs are localized in compact homogeneous clusters, while IEG-defined cell assemblies associate in larger clusters, within which IEG expression correlates with cluster density.

Discussion

In this study, we address the cellular distribution of IEG expression within striatal neurons following acute exposure to cocaine, investigating the coherence of IEG expression, and the distribution of IEG expression within SPN subtypes, as well as the macroscale and mesoscale spatial organization of IEG expression. The broad relevance of observations regarding coherence of IEG expression was further investigated utilizing a dataset of scRNA-seq interrogation of mouse visual cortex following exposure to light. Based on our results, we propose several principles defining the distribution of IEG expression within neuronal assemblies. Consistent with previous observations, we find that the induction of IEGs within neuronal assemblies is graded (1, 4), with substantial variation between

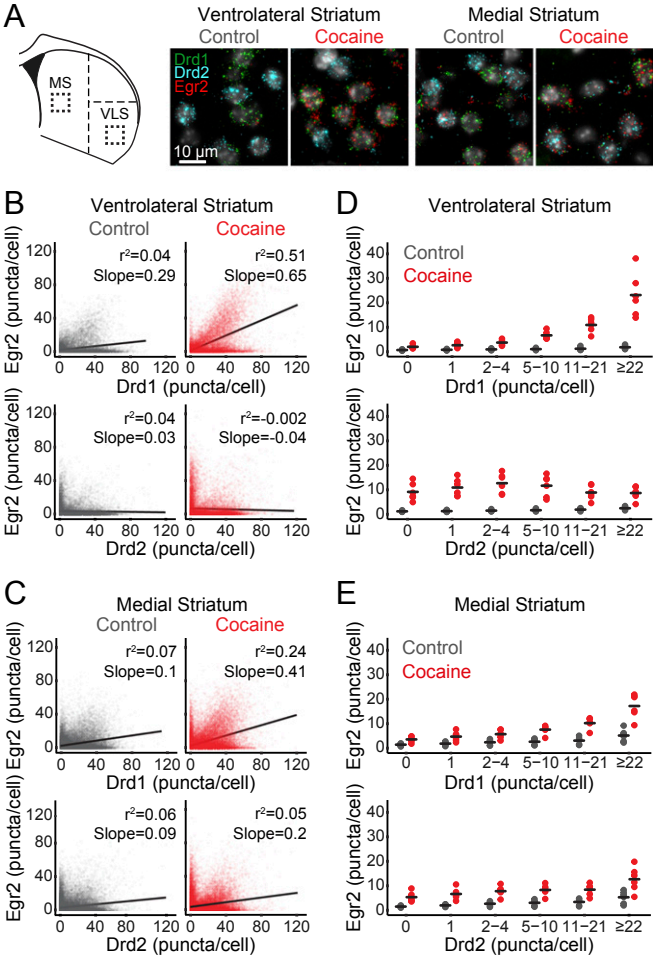

Fig. 3. Differential recruitment of *Drd1* and *Drd2* SPNs in ventrolateral vs. medial striatum. (*A*) Cartoon depiction of the location of ventrolateral (VLS) and medial (MS) striatal regions and representative high-magnification 40× images showing *Egr2* expression within *Drd1*- or *Drd2*-expressing SPNs in VLS (*Left*) and MS (*Right*) in control vs. cocaine. (*B* and *C*) Scatter plots of *Egr2* expression as a function of *Drd1* or *Drd2* expression within single cells in the VLS (*B*) and MS (*C*); $n = 6$ sections from 3 mice in each condition; VLS, 18,165 to 18,622 cells; MS, 31,008 to 31,920 cells. $P < 0.0001$. (*D* and *E*) Cocaine-induced *Egr2* expression correlates with *Drd1* expression in the VLS (*D*), and *Drd1*, as well as *Drd2* expression in the MS (*E*). Each dot represents the mean gene expression for each replicate bin, while the line represents the mean. $n = 6$ sections from 3 mice for each condition; VLS, 18,165 to 18,622 cells; MS, 31,008 to 31,920 cells ($P < 0.005$ for all conditions except VLS *Drd2*, which was nonsignificant; ANOVA).

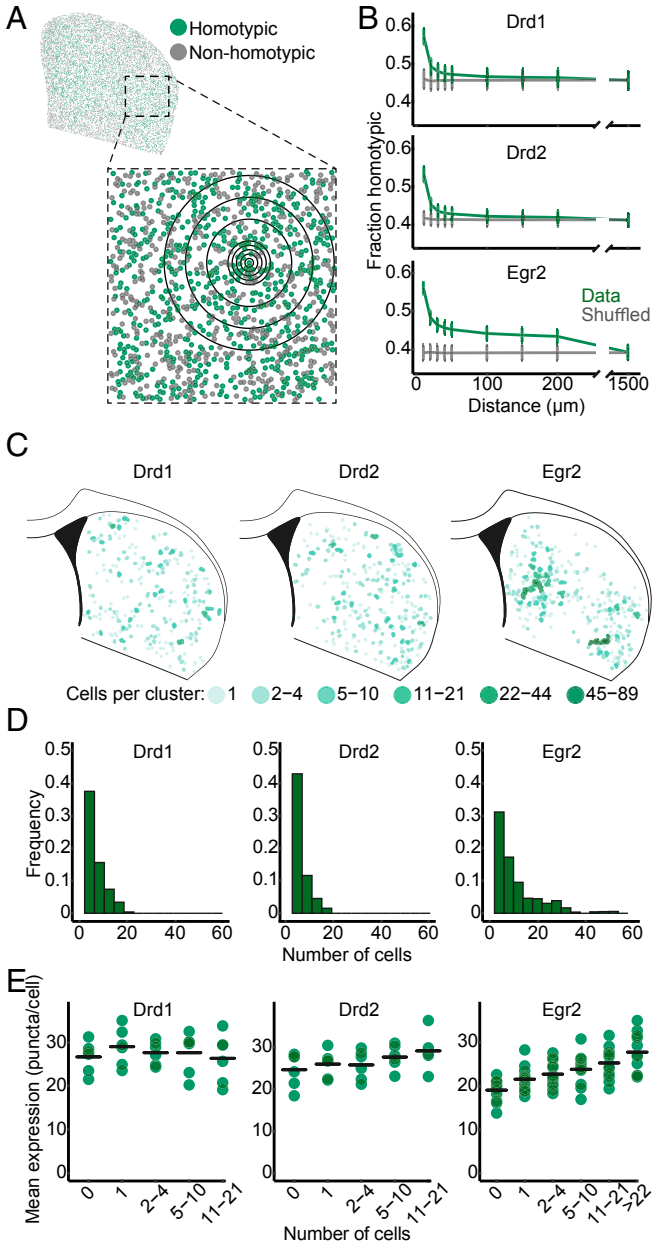

Fig. 4. Neurons in striatal ensembles associate in clusters, within which IEG expression correlates to cluster size. (*A*) Illustration of strategy used to identify the fraction of homotypic neighboring cells. For each cell exhibiting robust expression of a given gene (*Drd1*/*Drd2* ≥ 8, *Egr2* ≥ 6 puncta per cell), the fraction of homotypic neighboring cells is calculated at increasing radii. (*B*) Plots of the fraction of homotypic cells as a function of distance, compared to data shuffled for expression levels between cells, while maintaining their spatial organization. Data represented as mean ± SEM; n = 6 to 12 sections from 3 mice (61,491 to 122,195 cells). (*C*) Representative depiction of *Drd1*, *Drd2*, and *Egr2* neuronal clusters. Neighboring homotypic cells were clustered if their distance was below 20 μm. (*D*) Quantification of the frequency of clusters of different size for *Drd1*/*Drd2*/*Egr2*. (*E*) Mean *Egr2* (but not *Drd1* or *Drd2*) expression correlates with increased cluster size (*Egr2*, $P < 0.0001$; *Drd1*, $P = 0.8$; *Drd2*, $P = 0.3$; ANOVA). Each dot represents the mean expression of each gene within a replicate section. n = 6 to 12 sections from 3 mice; 61,491 to 122,195 cells.

cells (49–52). It should be noted that IEGs vary in their basal expression levels (i.e., background levels), in their specificity of induction by experience (25), and in the level of induction they achieve (26, 28, 33, 38). Still, we find that the absolute levels of expression of each IEG, within individual cells coexpressing multiple IEGs, are largely correlated. Two different hypotheses provide simple explanations for this coherence of IEG coexpression within a broad distribution: 1) Neurons in the same brain region and of the same type are engaged by experience, to drive the induction of IEGs, to a similar extent (transitioning from "off" to "on"), but differ in their dynamics of engagement, such that the diversity of expression observed at any one time point corresponds to different phases in the development of the full response and its decay back to control levels. Alternatively, 2) neurons in the same brain region and of the same type are engaged to different extents by a given experience. In the case of cocaine, striatal IEG expression has been clearly shown to depend upon the expression of dopamine receptors in striatal neurons (42, 53, 57, 58). As we observe a strong correlation between *Egr2* expression and the expression of *Drd1* and *Drd2* dopamine receptors, we favor the second hypothesis, whereby neurons within an assembly are recruited differentially, according to their sensitivity to the inducing stimulus. This would imply that the synaptic expression of dopamine receptors conferring sensitivity to cocaine corresponds to the cellular expression levels of dopamine receptor genes, a point for future investigation.

Addressing the identity of neurons recruited by experience, our smFISH analysis uncovered surprising region-specific rules. While in VLS neurons, *Egr2* was induced selectively within the *Drd1*$^+$ population, in the MS *Egr2* induction was observed both within *Drd1*- and *Drd2*-expressing neurons (albeit with a preference for *Drd1*$^+$). Thus, macroscale spatial rules contribute to defining the identity of neurons recruited to assemblies. These rules may theoretically relate to engagement of inhibitory neuronal populations specifically gating the recruitment of a neuronal subtype within a defined spatial domain, or to differential engagement of neuronal subtypes by long-range inputs. Future research is required in order to define the local and long-range circuit bases of these macroscale, subregion-specific rules.

Recent measurements of striatal activity have identified clusters of neuronal coactivation that were correlated to specific motor sequences (45–47). In addition, astrocyte-mediated homotypic regulation of neighboring striatal neurons has been described (48). Thus, mesoscale structure is evident in the activity of striatal neurons, potentially subject to local regulation, and functionally associated to encode specific actions. Addressing mesoscale organization within the smFISH data, we find a number of interesting features. Both *Drd1*$^+$ and *Drd2*$^+$ neurons are enriched within small homotypic clusters throughout the striatum. Furthermore, IEG-expressing neuronal assemblies (as defined by cocaine-induced *Egr2*, *Arc*, or *Nr4a1* expression) associate in large clusters, and the size of the cluster tends to correlate with the cellular expression of these IEGs. This correlation of IEG expression with cluster size suggests that IEG induction within spatially clustered neurons may be subject to local amplification, possibly through common anatomical connections, a topic for further investigation. It should be noted that our smFISH investigation utilized thin (14-μm) sections through the striatum. Thus, our results likely underrepresent the spatial architecture of constitutive and induced gene expression in the striatum, as mesoscale organization likely extends in 3 dimensions within the striatum. Future research could address the relationship between coactivation of striatal neurons and the distribution of IEG expression within the striatum, as well as a potential role for astrocytic domains in the definition of the boundaries of IEG-expressing clusters.

Both smFISH and scRNA-seq datasets demonstrate a clear gradation of IEG expression, whereby high levels of IEG expression are observed to correspond to coinduction of multiple IEGs. Our smFISH analysis further associates high IEG expression with mesoscale clustering. Thus, at the extreme end of the spectrum exists a population of neurons within which multiple

IEGs are coherently coexpressed at robust levels, and which preferentially associates within relatively tight clusters, defining a superensemble of transcriptionally coherent, spatially packed neurons. It should be noted that, due to the gradation of expression, coherence, and spatial clustering, specifying a threshold for inclusion within the superensemble is nontrivial. Whether superensembles bear greater responsibility for information coding than adjacent neurons with reduced coherence or spatial packing, is a point worthy of future investigation. Furthermore, future studies will address the relationship between the IEG representation of multiple experiences in striatal neurons—addressing overlap between neuronal assemblies encoding different experiences, as well as the broader relevance of principles defined here of coherence, subtype specificity, and macroscale and mesoscale spatial organization.

Methods

Detailed methods can be found in *SI Appendix* and raw data are found in Dataset S1. Five- to 6-wk-old male C57BL/6 mice weighing 19 to 21.5 g were used in this study. Behavioral analysis was performed as previously described (25). The smFISH protocol was implemented according to manufacturer guidelines (ACD RNAscope fresh frozen tissue pretreatment and fluorescent multiplex assay manuals; catalog #320513 and #320293). R, version 3.4.4, was used for all statistical analysis and graphical representations.

ACKNOWLEDGMENTS. We are grateful to Dr. Naomi Melamed-Book for support in establishing the imaging system required for efficient acquisition of multiplexed smFISH signals. We are also grateful to Ms. Adi Doron for insight regarding mesoscale spatial clustering analysis. We further appreciate members of the A.C. laboratory (Noa Rivlin, Gal Atlan, Anna Terem, and Dr. David Lipton) for helpful comments on the manuscript; Dr. Inbal Goshen for critical comments on data, writing, and presentation; Prof. David Clayton for insightful discussions and constructive critique of the manuscript; and our anonymous reviewers for their helpful commentary and suggestions. Work in the A.C. laboratory is funded by the European Research Council (Grant ERC 770951), The Israel Science Foundation (Grants 1062/18, 393/12, 1796/12, and 2341/15), EU Marie Curie (Grant PCIG13-GA-2013-618201), Brain and Behavior Foundation (Grant NARSAD 18795), German–Israel Foundation (Grant 2299-2291.1/2011), and Binational Israel–United States Foundation (Grant 2011266). Y.L. is funded by the Israel Science Foundation (Grant 757/16) and the Gatsby Charitable Foundation.

1. D. F. Clayton, The genomic action potential. *Neurobiol. Learn. Mem.* **74**, 185–216 (2000).
2. C. M. Alberini, Transcription factors in long-term memory and synaptic plasticity. *Physiol. Rev.* **89**, 121–145 (2009).
3. A. Holtmaat, P. Caroni, Functional and structural underpinnings of neuronal assembly formation in learning. *Nat. Neurosci.* **19**, 1553–1562 (2016).
4. D. F. Clayton *et al*., The role of the genome in experience-dependent plasticity: Extending the analogy of the genomic action potential. *Proc. Natl. Acad. Sci. U.S.A.*, 10.1073/pnas.1820837116 (2019).
5. A. E. West, M. E. Greenberg, Neuronal activity-regulated gene transcription in synapse development and cognitive function. *Cold Spring Harb. Perspect. Biol.* **3**, a005744 (2011).
6. H. Bading, Nuclear calcium signalling in the regulation of brain function. *Nat. Rev. Neurosci.* **14**, 593–608 (2013).
7. A. Lanahan, P. Worley, Immediate-early genes and synaptic function. *Neurobiol. Learn. Mem.* **70**, 37–43 (1998).
8. T. A. Terleph, L. A. Tremere, "The use of immediate early genes as mapping tools for neuronal activation: Concepts and methods" in *Immediate Early Genes in Sensory Processing, Cognitive Performance and Neurological Disorders*, R. Pinaud, L. A. Tremere, Eds. (Springer, Boston, MA, 2006), pp. 1–10.
9. S. Loebrich, E. Nedivi, The function of activity-regulated genes in the nervous system. *Physiol. Rev.* **89**, 1079–1103 (2009).
10. B. Pérez-Cadahía, B. Drobic, J. R. Davie, Activation and function of immediate-early genes in the nervous system. *Biochem. Cell Biol.* **89**, 61–73 (2011).
11. J. M. Gray, I. Spiegel, Cell-type-specific programs for activity-regulated gene expression. *Curr. Opin. Neurobiol.* **56**, 33–39 (2019).
12. E.-L. Yap, M. E. Greenberg, Activity-Regulated transcription: Bridging the gap between neural activity and behavior. *Neuron* **100**, 330–348 (2018).
13. S. Ribeiro, C. V. Mello, Gene expression and synaptic plasticity in the auditory forebrain of songbirds. *Learn. Mem.* **7**, 235–243 (2000).
14. G. Feenders *et al*., Molecular mapping of movement-associated areas in the avian brain: A motor theory for vocal learning origin. *PLoS One* **3**, e1768 (2008).
15. K. Minatohara, M. Akiyoshi, H. Okuno, Role of immediate-early genes in synaptic plasticity and neuronal ensembles underlying the memory trace. *Front. Mol. Neurosci.* **8**, 78 (2016).
16. M. C. Saul *et al*., Transcriptional regulatory dynamics drive coordinated metabolic and neural response to social challenge in mice. *Genome Res.* **27**, 959–972 (2017).
17. J. Zhang *et al*., c-*fos* regulates neuronal excitability and survival. *Nat. Genet.* **30**, 416–420 (2002).
18. M.-H. Han, S. J. Russo, E. J. Nestler, "Molecular, cellular, and circuit basis of depression susceptibility and resilience" in *Neurobiology of Depression*, J. Quevedo, A. Carvalho, C. Zarate, Eds. (Academic, London, 2019), pp. 123–136.
19. D. M. Walker *et al*., Cocaine self-administration alters transcriptome-wide responses in the brain's reward circuitry. *Biol. Psychiatry* **84**, 867–880 (2018).
20. D. L. Marcus *et al*., Quantitative neuronal c-*fos* and c-*jun* expression in Alzheimer's disease. *Neurobiol. Aging* **19**, 393–400 (1998).
21. H. Jin, D. F. Clayton, Localized changes in immediate-early gene regulation during sensory and motor learning in zebra finches. *Neuron* **19**, 1049–1059 (1997).
22. F. C. Cruz *et al*., New technologies for examining the role of neuronal ensembles in drug addiction and fear. *Nat. Rev. Neurosci.* **14**, 743–754 (2013).
23. F. E. Henry, K. Sugino, A. Tozer, T. Branco, S. M. Sternson, Cell type-specific transcriptomics of hypothalamic energy-sensing neuron responses to weight-loss. *eLife* **4**, e09800 (2015).
24. Y. Lin *et al*., Activity-dependent regulation of inhibitory synapse development by Npas4. *Nature* **455**, 1198–1204 (2008).
25. D. Mukherjee *et al*., Salient experiences are represented by unique transcriptional signatures in the mouse brain. *eLife* **7**, e31220 (2018).
26. K. M. Tyssowski *et al*., Different neuronal activity patterns induce different gene expression programs. *Neuron* **98**, 530–546.e11 (2018).
27. A. R. Mardinly *et al*., Sensory experience regulates cortical inhibition by inducing IGF1 in VIP neurons. *Nature* **531**, 371–375 (2016).
28. O. Whitney *et al*., Core and region-enriched networks of behaviorally regulated genes and the singing genome. *Science* **346**, 1256780 (2014).
29. L. Luo, E. M. Callaway, K. Svoboda, Genetic dissection of neural circuits: A decade of progress. *Neuron* **98**, 256–281 (2018).
30. J. F. Guzowski, B. L. McNaughton, C. A. Barnes, P. F. Worley, Environment-specific expression of the immediate-early gene Arc in hippocampal neuronal ensembles. *Nat. Neurosci.* **2**, 1120–1124 (1999).
31. J. Xiu *et al*., Visualizing an emotional valence map in the limbic forebrain by TAI-FISH. *Nat. Neurosci.* **17**, 1552–1559 (2014).
32. Q. Zhang, Q. He, J. Wang, C. Fu, H. Hu, Use of TAI-FISH to visualize neural ensembles activated by multiple stimuli. *Nat. Protoc.* **13**, 118–133 (2018).
33. K. M. Tyssowski, J. M. Gray, The neuronal stimulation-transcription coupling map. *Curr. Opin. Neurobiol.* **59**, 87–94 (2019).
34. M. Mayford, L. Reijmers, Exploring memory representations with activity-based genetics. *Cold Spring Harb. Perspect. Biol.* **8**, a021832 (2015).
35. M. Stevense, T. Muramoto, I. Müller, J. R. Chubb, Digital nature of the immediate-early transcriptional response. *Development* **137**, 579–584 (2010).
36. R. Chandra, M. K. Lobo, Beyond neuronal activity markers: Select immediate early genes in striatal neuron subtypes functionally mediate psychostimulant addiction. *Front. Behav. Neurosci.* **11**, 112 (2017).
37. I. Spiegel *et al*., Npas4 regulates excitatory-inhibitory balance within neural circuits through cell-type-specific gene programs. *Cell* **157**, 1216–1229 (2014).
38. S. Hrvatin *et al*., Single-cell analysis of experience-dependent transcriptomic states in the mouse visual cortex. *Nat. Neurosci.* **21**, 120–129 (2018).
39. X. Pichon, M. Lagha, F. Mueller, E. Bertrand, A growing toolbox to image gene expression in single cells: Sensitive approaches for demanding challenges. *Mol. Cell* **71**, 468–480 (2018).
40. P. Gao, J. C. de Munck, J. H. W. Limpens, L. J. M. J. Vanderschuren, P. Voorn, A neuronal activation correlate in striatum and prefrontal cortex of prolonged cocaine intake. *Brain Struct. Funct.* **222**, 3453–3475 (2017).
41. J. M. Caster, C. M. Kuhn, Maturation of coordinated immediate early gene expression by cocaine during adolescence. *Neuroscience* **160**, 13–31 (2009).
42. R. Moratalla, M. Xu, S. Tonegawa, A. M. Graybiel, Cellular responses to psychomotor stimulant and neuroleptic drugs are abnormal in mice lacking the D1 dopamine receptor. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 14928–14933 (1996).
43. A. C. Kreitzer, R. C. Malenka, Striatal plasticity and basal ganglia circuit function. *Neuron* **60**, 543–554 (2008).
44. A. V. Kravitz, L. D. Tye, A. C. Kreitzer, Distinct roles for direct and indirect pathway striatal neurons in reinforcement. *Nat. Neurosci.* **15**, 816–818 (2012).
45. A. Klaus *et al*., The spatiotemporal organization of the striatum encodes action space. *Neuron* **95**, 1171–1180.e7 (2017). Erratum in: *Neuron* **96**, 949 (2017).
46. G. Barbera *et al*., Spatially compact neural clusters in the dorsal striatum encode locomotion relevant information. *Neuron* **92**, 202–213 (2016).
47. J. E. Markowitz *et al*., The striatum organizes 3D behavior via moment-to-moment action selection. *Cell* **174**, 44–58.e17 (2018).
48. R. Martin, R. Bajo-Grañeras, R. Moratalla, G. Perea, A. Araque, Circuit-specific signaling in astrocyte-neuron networks in basal ganglia pathways. *Science* **349**, 730–734 (2015).
49. C. V. Mello, D. S. Vicario, D. F. Clayton, Song presentation induces gene expression in the songbird forebrain. *Proc. Natl. Acad. Sci. U.S.A.* **89**, 6818–6822 (1992).
50. C. V. Mello, D. F. Clayton, Song-induced ZENK gene expression in auditory pathways of songbird brain and its relation to the song control system. *J. Neurosci.* **14**, 6652–6666 (1994).
51. E. D. Jarvis, F. Nottebohm, Motor-driven gene expression. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 4097–4102 (1997).
52. E. D. Jarvis, C. Scharff, M. R. Grossman, J. A. Ramos, F. Nottebohm, For whom the bird sings: Context-dependent gene expression. *Neuron* **21**, 775–788 (1998).

53. C. R. Gerfen *et al.*, D1 and D2 dopamine receptor-regulated gene expression of striatonigral and striatopallidal neurons. *Science* **250**, 1429–1432 (1990).
54. D. A. Burke, H. G. Rotstein, V. A. Alvarez, Striatal local circuitry: A new framework for lateral inhibition. *Neuron* **96**, 267–284 (2017).
55. C. R. Gerfen, D. J. Surmeier, Modulation of striatal projection systems by dopamine. *Annu. Rev. Neurosci.* **34**, 441–466 (2011).
56. O. Gokce *et al.*, Cellular taxonomy of the mouse striatum as revealed by single-cell RNA-seq. *Cell Rep.* **16**, 1126–1137 (2016).
57. H. Steiner, C. R. Gerfen, Cocaine-induced c-fos messenger RNA is inversely related to dynorphin expression in striatum. *J. Neurosci.* **13**, 5066–5081 (1993).
58. A. J. Cole, R. V. Bhat, C. Patt, P. F. Worley, J. M. Baraban, D1 dopamine receptor activation of multiple transcription factor genes in rat striatum. *J. Neurochem.* **58**, 1420–1426 (1992).