Cell-type landscape of the central amygdala: via single-cell RNA sequencing, multiplexed fluorescent in situ hybridization

New neuronal cell types were examined with scRNA-seq and multiplexed FISH

Our work revealed new inhibitory neuron cell types within specific spatial domains

Newly revealed subtypes had specialized proteins and long-range projections

Our work clarifies the spatial cell-type-specific landscape of the central amygdala

new transcriptomic *Isl1*-expressing and *Nr2f2*-expressing inhibitory neuron types, with specific proteins and projection targets
Neuronal cell types, projections, and spatial organization of the central amygdala

Timothy P. O’Leary, Rennie M. Kendrick, Brianna N. Bristow, Kaitlin E. Sullivan, Lihua Wang, Jody Clements, Andrew L. Lemire, and Mark S. Cembrowski

SUMMARY

The central amygdala (CEA) has been richly studied for interpreting function and behavior according to specific cell types and circuits. Such work has typically defined molecular cell types by classical inhibitory marker genes; consequently, whether marker-gene-defined cell types exhaustively cover the CEA and co-vary with connectivity remains unresolved. Here, we combined single-cell RNA sequencing, multiplexed fluorescent in situ hybridization, immunohistochemistry, and long-range projection mapping to derive a “bottom-up” understanding of CEA cell types. In doing so, we identify two major cell types, encompassing one-third of all CEA neurons, that have gone unresolved in previous studies. In spatially mapping these novel types, we identify a non-canonical CEA subdomain associated with Nr2f2 expression and uncover an Isl1-expressing medial cell type that accounts for many long-range CEA projections. Our results reveal new CEA organizational principles across cell types and spatial scales and provide a framework for future work examining cell-type-specific behavior and function.

INTRODUCTION

The central amygdala (CEA) plays a variety of functional roles. A significant body of work has focused on the involvement of the CEA in fear memory (for review, see 1,2), as well as its contributions to anxiety-like and defensive behaviors.3–7 Additional work has implicated the CEA in pain, motivation, reward, aggression, and satiety (for review, see 7). How a single brain region can support such a diverse repertoire of functions and behaviors remains unknown.8

One approach to mapping neural substrates with function relies on molecular “cell typing,” wherein marker genes are used to delineate, observe, and manipulate precise groups of neurons. As the CEA is largely comprised of inhibitory neurons, this approach has typically been performed by leveraging canonical inhibitory neuron marker genes (e.g., Oxtr, Crh, Nts, Tac2, Drd2, Prkcd, Sst).11–14 Importantly, such marker-gene-defined cell types within the CEA have been causally shown to have distinct contributions to a wide range of behaviors and functions.15–20

Although a variety of work has examined CEA cell types using these a priori marker genes or bulk transcriptional profiling,14,21–24 to date no published work has examined CEA organization from a whole-genome perspective at a single-cell resolution. Defining cell types at this fine level of granularity is likely to provide a more complete understanding of the cell-type organization of the CEA (for review, see 25,26); which may validate whether known marker genes capture all cell types, or conversely identify new cell types and organizational principles within the CEA.27,28 In either case, such results can provide a cell-type framework for integrating spatial organization with other structural and functional properties.29

Here, we sought to use a bottom-up approach to derive the inhibitory cell types that make up the CEA and register these cell types to spatial domains and long-range projection targets. To begin, we leveraged single-cell RNA sequencing (scRNA-seq) to identify cell types in the CEA, which recapitulated previously known cell types and revealed that a large number of CEA inhibitory neurons were not captured with previously known markers. Next, using multiplexed fluorescent in situ hybridization (mFISH) to map these known and novel cell types in space, we revealed new spatial organizational principles of the CEA that elaborate on current atlas parcellations. Finally, in combining mFISH with retrograde tracing, we found that many long-range CEA projections are associated with previously unresolved molecular cell types. In total,
our work reveals the principles of CEA local organization and provides insight into the molecular and spatial logic of CEA long-range projections.

RESULTS

Single-cell RNA sequencing identifies known and novel central amygdala inhibitory cell types

We began by performing scRNA-seq by microdissecting and dissociating the CEA, and manually capturing individual cells with neuron-like cell bodies.26,30 These captured cells underwent scRNA-seq (see STAR methods), with a total of 826 high-quality neurons retained for analysis (from n = 3 mice, with 4.5 ± 0.9 thousand expressed genes/cell and 141 ± 73 thousand transcripts/cell, mean ± SD). Visualizing transcriptomic data via UMAP dimensionality reduction31 in conjunction with Louvain-based clustering revealed 7 cell types in this scRNA-seq dataset (Figure 1A), which were reproducible across all mice (Figures S1A and S1B). All cell types exhibited transcriptomes consistent with inhibitory neurons (e.g., the expression of inhibitory neuron markers Gad1 and Gad2, Figure 1B). Consistent with the lack of excitatory neurons in the CEA,32 cells generally lacked expression of excitatory neuron markers (e.g. Slc17a7: Figure 1B; see also Figures S1C and S1D). Data for these neurons, as well as user-friendly analysis and visualization tools, are publicly available online (http://scrnaseq.janelia.org/cea).

We began by examining expression of inhibitory neuron marker genes that have been previously used to access CEA subpopulations (Oxtr, Crh, Nts, Tac2, Drd2, Prkcd, Sst, Drd1a, Calca, and Pvalb).11–14 Several of these genes (Oxtr, Crh, Nts, Tac2, Drd2) corresponded to individual scRNA-seq-derived clusters, although many were associated with the same underlying clusters (Figure 1C, left) (see also11). The previously used marker genes Prkcd and Sst each exhibited mixed expression across multiple scRNA-seq-clusters (Figure 1C, right), with the remaining previously used genes (Drd1a, Calca, Pvalb, Tac1) relatively rarely expressed (Figure S1E). The GABAergic intercalated neuron marker Tshz113 was not expressed in a cluster-specific way (Figure S1F), suggesting that all scRNA-seq-clusters were CEA-specific. As this analysis left two of the seven scRNA-seq clusters without corresponding markers from previous research, these findings suggest multiple new CEA inhibitory subtypes that may have gone unresolved in previous research.

Importantly, each scRNA-seq cluster could be uniquely identified by the expression of individual genes (Figure 1D), with the two previously unresolved clusters in particular associated with Nr2f2 and Isl1 expression. Of note, many other functionally relevant genes were differentially expressed within these clusters (Figure S1G): these genes spanned ligands and receptors typically used to differentiate inhibitory cell types, but also included a wide array of genes associated with other cellular and synaptic function. Consequently, these results argue that differences in gene expression between scRNA-seq-derived cell types likely drive markedly different functional properties. This finding, in conjunction with the large number of cells associated with previously unresolved Nr2f2 and Isl1-expressing cell types (~1/3 of all profiled cells: Figure 1E), suggests undiagnosed and important cell-type determinants of CEA organization and function.

Single-gene in situ hybridization suggests discrete spatial domains

Given that our scRNA-seq analysis suggested potentially new CEA inhibitory cell types and associated marker genes, we next sought to confirm and extend these results using in situ hybridization (ISH) across anterior, intermediate, and posterior sections of the CEA (Figure 2A). Using the Allen Mouse Brain atlas,35 we first identified the broad Gad1-expressing CEA population (Gad1 images: Figure 2B); segmented Gad1-expressing cells: Figure 2C). This expression was then used as a reference frame, wherein the scRNA-seq-derived marker gene expression from other animals was overlaid.

Each scRNA-seq-derived marker gene did indeed show expression within the CEA, with genes typically exhibiting expression in spatially restricted domains of the CEA rather than uniformly across the CEA (Figures 2B–2Q). Generally, marker genes associated with known inhibitory cell types tended to be expressed in lateral and/or posterior locations of the CEA (Figures 2D–2M), whereas markers for the previously unresolved cell types Nr2f2 and Isl1 were typically located in more medial and anterior sections. Overlapping and comparing gene expression of known and novel genes hinted at spatial boundaries discretely separating these cell types (Figures 2R and 2S). Quantitatively, and in agreement with our scRNA-seq work, Nr2f2 and Isl1-expressing cells encompassed approximately one-third of all cells labeled by expression of scRNA-seq-derived marker genes (Figure 2T).
Multiplexed in situ hybridization identifies discrete, abutting spatial domains

Although our single-gene ISH analysis was suggestive of a spatially abutting patterning of CEA cell types (Figures 2R and 2S), such an approach does not resolve relationships across genes within individual tissue sections, and thus only provides indirect evidence of such putatively discrete spatial domains. To circumvent this issue, we next employed multiplexed fluorescent ISH (mFISH), enabling spatially localized and
We next repeated this subdivision-based analysis by examining novel cell types (Figure 5). Notably, both Nr2f2- and Isl1-expressing clusters showed enrichment in the CEAm (cf. Figure 4, reproducible across n = 3 animals: Figure S4A). In particular, the Isl1-expressing cluster was enriched in anterior and intermediate CEAm subdivisions, whereas the Nr2f2-expressing cluster exhibited enrichment solely in the anterior CEAm subdivisions, whereas the Isl1-expressing cluster exhibited enrichment solely in the anterior division.
Figure 3. Multiplexed FISH analysis of the CEA

(A) mFISH overlays for anterior (left), intermediate (middle), and posterior (right) CEA sections from a representative animal. CEA outlined in white. Inset illustrates an expansion from the intermediate section. Probe targets are listed in accompanying box. Scale bar: 300 μm.

(B) Two-dimensional UMAP embedding of cells segmented from (A). Coloring indicates cluster membership according to a Leiden clustering algorithm (Levine et al., 2015; Traag et al., 2019).

(C) Expression of marker genes associated with each cluster, quantified as percent area covered.

Gene list and color conventions

| Gene   | Color |
|--------|-------|
| Gad1   |       |
| Prkcd  |       |
| Crcrl  |       |
| Isl1   |       |
| Nr2I2  |       |
| Cxcl14 |       |
| Fxyd6  |       |
| Nts    |       |
| Cck    |       |
| Carpt  |       |

Previously studied clusters

Novel clusters

Percent area covered

D

Anterior

Intermediate

Posterior

E

Number of cells

Previously studied
Novel

65%

35%
CEAm. Such work illustrates that these novel cell types account for spatial domains that have gone unresolved using previous marker genes. Moreover, these findings also suggest that there may be a CEAm subdivision in the posterior CEA, although no such anatomical subdivision is typically included in atlas parcellations. Specifically, \textit{Nr2f2} and \textit{Isl1} clusters were found in abutting posterior regions within the atlas-classified CEAl, exhibiting the same spatial patterning seen in more anterior subdivisions.

Strikingly, the \textit{Nr2f2}-expressing cluster was also the sole cluster that exhibited enriched expression across all CEA subdivisions (CEAm: enriched in anterior, CEAl: enriched in intermediate and posterior, CEAc: enriched in intermediate). Intriguingly, this \textit{Nr2f2}-expressing cell type did not conform to the standard CEA atlas parcellations, but rather formed an intermediate banded structure across the CEA that spanned and split conventional subdivisions. In collection, these findings suggest a cell-type-specific spatial organization that may elaborate upon more classical atlas definitions.

**Protein-level correlates of central amygdala marker genes and spatial domains**

To determine whether the novel cell-type organization revealed by our mFISH analysis could predict spatial domain organization at the protein level, we next performed immunohistochemistry. We identified antibodies that would allow for simultaneous labeling of Prkcd, \textit{Nr2f2}, and \textit{Isl1} proteins, and examined these three proteins in sections at comparable anterior-posterior locations to those examined in our mFISH analysis (Figure 6).

In agreement with previous mFISH organization, Prkcd-labeled cells were found laterally within the CEA, whereas \textit{Nr2f2}-and \textit{Isl1}-labeled cells occupied relatively medial neighboring domains (Figures 6A and 6B). Notably, although some coarse intermixing of \textit{Nr2f2}-and \textit{Isl1}-labeled cells was present in the medial CEA, cells typically were labeled by only one protein (expansion: Figure 6C, shuffled dataset: Figure 6D). Broadly, this spatial organization generalized across the anterior, intermediate, and posterior CEA sections (Figures 6E and 6F) and animals (Figure S4B). Thus, Prkcd, \textit{Nr2f2}, and \textit{Isl1} protein products demarcate abutting and distinct spatial CEA domains predicted by mFISH.

**Multiple long-range projections emerge from central amygdala**

As the function of the CEA is often interpreted according to long-range projection properties, we next examined if CEA projection neurons exhibited spatial patterning. Projection cells (“projectors”) were identified following the injection of a retrograde tracer into downstream regions of the CEA either rAAV2-retro-GFP or rAAV2-retro-tdt.\textsuperscript{38} We examined six well-studied long-range projection targets of the CEA: the substantia nigra, parabrachial nucleus, periaqueductal gray, parafascicular thalamus, lateral hypothalamus, and nucleus basalis stria terminalis (Figure 7; injection sites shown in Figure S5, and animal-to-animal reproducibility shown in Figures S6 and S7).

Across all examined downstream regions, projectors showed prominent spatial patterning within the CEA, manifesting as marked differences across CEA nuclei. In particular, four of these projection classes exhibited significant enrichment relative to chance for the CEAm exclusively (substantia nigra, parabrachial nucleus, periaqueductal gray, and parafascicular thalamus: Figures 7A-7L; all obeyed \( p < 0.001 \) for spatial enrichment relative to chance: Figure S7). Cells projecting to the remaining regions, the lateral hypothalamus and nucleus basalis stria terminalis, did not have this exclusive CEAm enrichment (lateral hypothalamus: CEAm and CEAc enrichment; nucleus basalis stria terminalis: CEAc enrichment; all \( p < 0.001 \), Figure S7). These results indicate that projectors from the CEA form spatially heterogeneous populations dependent on downstream projection site, similar to the spatial heterogeneity observed in gene expression, with the majority of projection classes examined being exclusively enriched in CEAm.

**Correlating long-range projections with marker gene expression**

The spatial domains associated with CEA long-range projections (Figure 7) appeared to coarsely agree with cluster-associated CEA boundaries and cell types (cf. Figure 3D). As the individual marker genes
Figure 4. Assessing spatial organization of previously studied CEA cell types

(A) Workflow for assessing the qualitative and quantitative spatial organization of cell types across CEA subdivisions, exemplified via \textit{Calcrl}-expressing cluster. Left: Spatial subdivisions of the anterior CEA (top) and cells overlaid on top of CEA subdivisions (bottom). Middle: Cell-density map of \textit{Calcrl}-expressing cell cluster for this individual section. Right: Enrichment analysis, across anterior sections from all animals, for \textit{Calcrl} expression by subdivision.

(B–D) As in (A), but for cells associated with all previously known cluster markers (i.e., \textit{Calcrl}, \textit{Prkcd}, \textit{Crh}, \textit{Sst}, \textit{Cxcl14}). Scale bar: 500 \textmu m.
Prkcd, Nr2f2, and Isl1 demarcated spatially abutting and cell-type-specific CEA domains (Figure 8A), we next sought to register long-range CEA projectors to the expression of these genes. To do this, we first imaged retrograde-virus-labeled CEA cells (as in Figure 7), and then quenched fluorophores and performed FISH for Prkcd, Nr2f2, and Isl1 in the same section (example projection: Figure 8B; see STAR methods). We focused on the substantia nigra, the parabrachial nucleus, and periaqueductal gray due to their dense spatial patterning of relatively medial projections, and included the lateral hypothalamus to contrast the relatively diffuse nature of these projectors.

As predicted from their medial location (Figures 7A–7I), CEA projectors to the substantia nigra, parabrachial nucleus, and periaqueductal gray showed the greatest Isl1 expression (Figure 8C, all obeyed $p_{ADJ} < 0.001$ for expression relative to non-projecting cells). Notably, each of these projections also showed enrichment of Nr2f2 expression ($p_{ADJ} < 0.001$ for all projections), but lacked enrichment for Prkcd expression ($p_{ADJ}$ not significant for all projections). Conversely, the relatively diffuse projections to the lateral hypothalamus (Figures 7M–7O) lacked enrichment for both Isl1 and Prkcd, but showed significant enrichment of Nr2f2 with a relatively modest effect size (Figure 8D). In total, these results illustrate that the previously unresolved Isl1-and Nr2f2-expressing populations comprise many different output pathways of the CEA.
Figure 6. Immunohistochemistry recapitulates spatial subdomains of the CEA at the protein level

(A) Overview of immunohistochemistry image showing Prkcd (blue), Nr2f2 (red), and Isl1 (green) labelling in a representative intermediate section, as well as a three-color overlay. Scale bar: 150 μm.

(B) Quantified relative protein abundance of Prkcd (blue), Nr2f2 (red), and Isl1 (green) for the intermediate section in (A).

(C) Expansion of dashed area from (B) illustrating that co-mingled medial cells typically are labeled by either Nr2f2 or Isl1 proteins.

(D) Shuffled analysis of the same intermediate section in (B) shows loss of selective markers and spatial domains.

(E) Composite immunohistochemistry image and quantified expression plot of a representative anterior CEA section. Scale bar: 150 μm.

(F) As in (E), but for a representative posterior CEA section. Scale bar: 150 μm.
Figure 7. CEA long-range projection neurons conform to restricted spatial domains

(A) Atlas schematic of the retrograde viral injection site for the substantia nigra modified from. 34
(B) Representative maximum intensity projection of labeled substantia nigra projecting cells (“projectors”) of the CEA. Scale bar: 250 μm. White outline indicates CEA boundaries.
(C) Locations of segmented cells (red), pooled across replicates, illustrating projector locations across anterior, intermediate, and posterior sections. For reference, substantia nigra projectors are superimposed over GAD1-expressing neurons (gray; as in Figure 2).
(D–R) As in (A–C), but for projectors to the parabrachial nucleus (D–F), periaqueductal gray (G–I), parafascicular thalamus (J–L), lateral hypothalamus (M–O), and nucleus basalis stria terminalis (P–R).
DISCUSSION

Here, we employed an unbiased scRNA-seq approach to identify cell types and associated marker genes in the CEA, and map these cell types onto spatial domains and long-range projections. In doing so, we recapitulated previous cell types and revealed two novel cell types with unique gene expression signatures (Figures 1–6), including a spatially banded intermediate Nr2f2-expressing and an abutting Isl1-expressing medial subtype. By simultaneously registering different molecularly defined cell types to long-range connectivity (Figures 7 and 8), we found that many long-range projectors of the CEA express these marker genes. As such, our findings reveal a novel CEA cell-type landscape and identify cells and molecules that can be accessed, perturbed, and interpreted in future experiments.

Transcriptomic approaches to resolving central amygdala cell types

“Cell typing” divides cells into groups based on variation in one or more properties,10 and has been used extensively to interpret the CEA by leveraging classical inhibitory neuron marker genes. Most studies have
focused on two marker genes, Prkcd and Sst, which are virtually non-overlapping in expression with associated cells differing in electrophysiology, spatially patterned inhibitory neuron types may drive the functional organization in intermediate sections that did not conform to any classical subdivisions, but rather seemed tectonically distinct from cells in both the medial and lateral subdivision. Importantly, this transition zone contains the highest density of CEA projections to lateral hypothalamus, consistent with our previous work has illustrated that different CEA subdivisions contain unique cell types and make distinct contributions to function, 55 suggesting that spatially patterned inhibitory neuron types may drive the functional organization in intermediate sections that did not conform to any classical subdivisions, but rather seemed tectonically distinct from cells in both the medial and lateral subdivision. Importantly, this transition zone contains the highest density of CEA projections to lateral hypothalamus, consistent with our

In our work here, we began with scRNA-seq to identify CEA cell types and marker genes de novo, rather than relying on previously known markers. In doing so, we identified new coherence and interrelationships in previously identified cell types, and moreover revealed new cell types that had been previously unresolved (Figure 1). To validate and spatially register our findings, scRNA-seq-derived marker genes were used to select mFISH probes for further analysis. Our mFISH work recapitulated previous spatial domains but also revealed more complex cell-type-specific organizational schemes (Figure 3).

This complement of scRNA-seq and mFISH provides a well-validated approach to understanding the CEA. From a conceptual standpoint, scRNA-seq cell-type assignments are derived in whole-genome space, whereas mFISH cell-type assignments are derived in a smaller (here, 12-dimensional) space. From a technical standpoint, mFISH also allows higher sensitivity of detection relative to scRNA-seq, circumventing issues of scRNA-seq drop-out. Ultimately, the general concordance across methodologies for our derived cell-type assignments illustrates that the CEA can be well-captured by a low-dimensional representation in gene-expression space, and that our derived cell type definitions encompass the spatial extent of the CEA.

As the spatial organization of cell types can be used as a framework to compare findings across studies, our use of spatially resolved mFISH provides an opportunity for external comparison of our findings. In comparing our mFISH results to two other FISH-based studies of the CEA, we recapitulated several key findings of previous work. Specifically, we confirmed the existence of several canonical CEA inhibitory neuron types, including Prkcd-, Sst-, Calcrl-, and Chh-expressing populations (Figures 1 and 3). Furthermore consistent with these prior studies, these cell types were largely relegated to the lateral regions of the CEA (Figure 3), resulting in a medial CEA region largely absent of classical inhibitory marker (Figure 4). Ultimately, our work builds upon previous studies by employing an unsupervised, data-driven approach to identify marker genes and cell types (Figure 1), and using mFISH to examine the combinatorial co-expression of twelve marker genes in individual brain sections (Figures 3 and S3). Our combination of data-driven cell types and higher multiplexing enabled us to fill previously noted gaps in CEA cell-type organization (Figures 1 and 3), in particular identifying Nr2f2- and Isl1-expressing novel CEA inhibitory neuron cell types biased toward the medial CEA (Figures 3, 5, and 6).

Spatial organization of central amygdala cell types and projections

Spatial domains within the CEA have frequently been used to understand CEA organization and interrelate CEA organization with function. Early work divided the CEA into gross medial, lateral, and capsular subdivisions based on anatomical features, morphology, and histochemical stains. Later complementary work has illustrated that different CEA subdivisions contain unique cell types and make distinct contributions to function, suggesting that spatially patterned inhibitory neuron types may drive the functional specialization of CEA subdivisions.

Here, using data-driven techniques to understand the cell-type composition of the CEA, we identified clear spatial patterning of many CEA cell types within these subdivisions (Figures 4 and 5). Importantly, although previous work has illustrated the expression of some classical inhibitory marker genes within the medial CEA, these genes only account for the minority of medial CEA neurons (>50% of cells lacked markers). Our findings here account for these missing neurons by uncovering previously unresolved medial-biased cell types within the CEA. These CEA cell types robustly express marker genes that have not yet been implicated in CEA cell-type identities, Nr2f2 and Isl1, and occupy restricted and largely reciprocal spatial domains across the anterior-posterior axis (Figures 5 and 6).

Of particular note, Nr2f2-expressing cells demonstrated a strikingly consistent and cohesive spatial organization in intermediate sections that did not conform to any classical subdivisions, but rather seemed to occupy a transition zone between medial and lateral subdivisions. Interestingly, early studies of the CEA identified a similar zone between medial and lateral subdivisions, which contained cells that were cytoarchitectonically distinct from cells in both the medial and lateral subdivision. Importantly, this transition zone contains the highest density of CEA projections to lateral hypothalamus, consistent with our
further suggests cell-type-specific contributions to CEA function. Specifically, more lateral cells transform and transmit local CEA signals, whereas the medial and molecularly distinct Isl1-expressing population conveys these processed signals to long-range targets. Thus, in addition to our work here harmonizing molecular, spatial, and connectivity properties of the CEA, it further suggests cell-type-specific computational roles emerging across the different CEA spatial domains.

**Hypothesized functional roles of novel cell types**

When considering potential functional roles for our newly revealed cell types, the multimodal nature of our cell typing enables us to derive predictions based on their molecular identity, spatial organization, and connectivity. For Isl1-expressing cells, the spatial, molecular, and connectivity profiles of this cell type all converge to suggest homology with a cell type of the striatum; namely, neurons that send projections as part of the “direct” striatal pathway. Spatially, the medial location of Isl1-expressing cells is consistent with previous work demonstrating that neurons in the medial CEA share a common structural and molecular profile to striatal neurons comprising the “direct” pathway.13,15,14 Molecularly, Isl1 is a transcription factor whose expression is required for striatal cells to assume direct cell-type identity,59,61 and fate-mapping of Isl1-expressing cells indicates that striatal direct long-range projectors arise from the same developmental pool as Isl1-expressing cells in the CEA.64 Finally, via connectivity, our work also illustrates that Isl1-expressing neurons also comprise a dominant long-range projecting population of CEA (Figures 7 and 8). In collection, our work synthesizes multiple lines of evidence suggesting structural, developmental, and molecular similarities between “direct” striatal long-range projectors and Isl1-expressing long-range projectors of the CEA.

We can leverage this multimodal homology between our novel Isl1-expressing cell type and richly functionally studied striatal “direct” neurons to make predictions about the functional roles Isl1-expressing neurons may play in the CEA. Broadly, direct neurons of the striatum are involved in promoting movement and action-oriented behavior through their projections to substantia nigra pars reticulata.83 Strikingly, the medial CEA, the subdivision in which Isl1-expressing neurons are contained, has been causally implicated in the expression of multiple action-oriented behaviors.13,35,64 Moreover, we also found that CEA Isl1-expressing cells project to the periaqueductal gray, which has been shown to be a central hub for motor function.64 Thus, Isl1-expressing cells may serve an analogous function to direct striatal neurons in CEA, conveying salient information into appropriate behavioral responses via extensive long-range projections. Future work could target Isl1-expressing cells based on their molecular identity and/or projection targets for causal interventions, aiming to causally map Isl1-associated function and behaviors.

**Limitations of the study**

The combination of previously known and novel cell types we identified here span the spatial extent of the CEA, and thus may reflect all of the inhibitory cell types of the CEA. Nonetheless, whether this collection of

*findings that lateral hypothalamus projectors expressed Nr2f2, but not Isl1 or Prkcd, at significantly higher levels than non-projectors (Figure 8). To our knowledge, our study is the first to find a unique molecular marker—Nr2f2—for this distinctive cell population (Figures 1, 3, and 6). Thus, our gene expression and connectivity findings provide a foothold by which this zone could be interrogated with high specificity, and further suggest that this zone plays a specialized role in CEA computation.*

A second cell type unresolved in previous work, adjacent Isl1-expressing cells, occupied the most medial component of the CEA. In addition to this cell type occupying a spatially distinct region of the CEA, by combining mFISH with long-range projection mapping, we revealed that this Isl1-expressing cell type predominantly encompasses many long-range projections of the CEA. Intriguingly, such a result illustrates that different CEA cell types co-vary in their intrinsic and connectivity properties, and likely thus are specialized for performing precise cell-type-specific contributions to CEA function. Specifically, more lateral cells transform and transmit local CEA signals, whereas the medial and molecularly distinct Isl1-expressing population conveys these processed signals to long-range targets. Thus, in addition to our work here harmonizing molecular, spatial, and connectivity properties of the CEA, it further suggests cell-type-specific computational roles emerging across the different CEA spatial domains.

**Hypothetical functional roles of novel cell types**

When considering potential functional roles for our newly revealed cell types, the multimodal nature of our cell typing enables us to derive predictions based on their molecular identity, spatial organization, and connectivity. For Isl1-expressing cells, the spatial, molecular, and connectivity profiles of this cell type all converge to suggest homology with a cell type of the striatum; namely, neurons that send projections as part of the “direct” striatal pathway. Spatially, the medial location of Isl1-expressing cells is consistent with previous work demonstrating that neurons in the medial CEA share a common structural and molecular profile to striatal neurons comprising the “direct” pathway.13,15,14 Molecularly, Isl1 is a transcription factor whose expression is required for striatal cells to assume direct cell-type identity,59,61 and fate-mapping of Isl1-expressing cells indicates that striatal direct long-range projectors arise from the same developmental pool as Isl1-expressing cells in the CEA.64 Finally, via connectivity, our work also illustrates that Isl1-expressing neurons also comprise a dominant long-range projecting population of CEA (Figures 7 and 8). In collection, our work synthesizes multiple lines of evidence suggesting structural, developmental, and molecular similarities between “direct” striatal long-range projectors and Isl1-expressing long-range projectors of the CEA.

We can leverage this multimodal homology between our novel Isl1-expressing cell type and richly functionally studied striatal “direct” neurons to make predictions about the functional roles Isl1-expressing neurons may play in the CEA. Broadly, direct neurons of the striatum are involved in promoting movement and action-oriented behavior through their projections to substantia nigra pars reticulata.83 Strikingly, the medial CEA, the subdivision in which Isl1-expressing neurons are contained, has been causally implicated in the expression of multiple action-oriented behaviors.13,35,64 Moreover, we also found that CEA Isl1-expressing cells project to the periaqueductal gray, which has been shown to be a central hub for motor function.64 Thus, Isl1-expressing cells may serve an analogous function to direct striatal neurons in CEA, conveying salient information into appropriate behavioral responses via extensive long-range projections. Future work could target Isl1-expressing cells based on their molecular identity and/or projection targets for causal interventions, aiming to causally map Isl1-associated function and behaviors.

**Limitations of the study**

The combination of previously known and novel cell types we identified here span the spatial extent of the CEA, and thus may reflect all of the inhibitory cell types of the CEA. Nonetheless, whether this collection of
cell types indeed comprehensively captures all inhibitory cell types of the CEA is unknown. Future work, potentially incorporating deeper sequencing and/or additional cells, will help to reveal whether there are additional rare types or finer subtypes within the CEA beyond those we have identified here.

Implications for future work
Our work here provides novel insight into CEA organization at both local and long-range scales and should help to guide and inform a variety of future experiments. By virtue of identifying cell types and respective marker genes (Figure 1), such cell types can be observed and manipulated in future experiments to derive the cell-type-specific underpinnings of CEA function. As such transcriptomic work also identifies genes that may play critical roles within specific cell types, our work here will similarly facilitate identifying molecular underpinnings of CEA structure and function. To facilitate this work in future research, we have hosted our scRNA-seq data along with analysis and visualization tools (http://scrnaseq.janelia.org/cea). As such cell types generally exhibited spatial organization across the CEA (Figures 2–6), such spatial patterning can act as a common framework for synthesizing results across studies. By revealing the long-range projections of specific marker-gene-defined cell types (Figures 7 and 8), our work will allow for conjunctive cell-type- and projection-specific labelling and manipulation, contributing to a more precise understanding of CEA long-range circuitry and function. As such, our work provides novel insight into CEA organization and will help to elucidate the cell-type-specific function of the CEA.

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 SUBJECT DETAILS
  - Single-cell RNA sequencing data acquisition and analysis
  - Chromogenic in situ hybridization
  - Tissue processing for histology
  - Multiplexed fluorescent in situ hybridization data acquisition
  - Multiplexed fluorescent in situ hybridization analysis
  - Immunohistochemical staining, imaging, and analysis
  - Viral tracing and analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION
Supplemental information can be found online at https://doi.org/10.1016/j.isci.2022.105497.

ACKNOWLEDGMENTS
MSC is supported by the University of British Columbia (Department of Cellular and Physiological Sciences, Djavad Mowafaghian Center for Brain Health, and the Faculty of Medicine Research Office), the Natural Sciences and Engineering Research Council of Canada (RGPIN-2019-04507), the Canadian Institutes of Health Research (PJT-419798), and the Canadian Foundation for Innovation (John R. Evans Leaders Fund 38369). RMK is supported by a Fulbright U.S. Student Program Award. This work was supported by resources made available through the NeuroImaging and NeuroComputation Center at the Djavad Mowafaghian Centre for Brain Health (RRID: SCR_019086). KES is supported by a Royal Canadian Legion Masters Scholarship in Veteran Health Research from the Canadian Institute for Military and Veteran Health Research. Collaboration between MSC and LW, JC, and ALL was supported by the Janelia Visiting Scientist Program. We thank members of the Cembrowski lab for helpful discussions and Jeffrey LeDue for insight and guidance in image acquisition.

AUTHOR CONTRIBUTIONS
MSC designed and supervised the project, LW and AL collected single-cell RNA-seq data, TPO performed retrograde tracing, JC constructed the web portal, RMK, KES and BNB performed in situ hybridization,
BNB performed immunohistochemistry, TPO, RMK, KES, BNB, and MSC analyzed data, TPO, RMK, KES, BNB, and MSC wrote the article.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

1. McCullough, K.M., Morrison, F.G., and Ressler, K.J. (2016). Bridging the Gap: towards a cell-type specific understanding of neural circuits underlying fear behaviors. Neuron. 135, 333-348. https://doi.org/10.1016/j.neuron.2016.07.025.

2. Fadok, J.P., Markovic, M., Toyove, P., and Luthi, A. (2018). New perspectives on central amygdala function. Curr. Opin. Neurobiol. 49, 141–147. https://doi.org/10.1016/j.conb.2018.02.009.

3. McCall, J.G., Al-Hasani, R., Siuda, E.R., Hong, D.Y., Norris, A.J., Ford, C.P., and Bruchas, M.R. (2015). CRH engagement of the locus coeruleus noradrenergic system mediates stress-induced anxiety. Neuron 87, 605–620. https://doi.org/10.1016/j.neuron.2015.07.002.

4. Botta, P., Demmou, L., Kasugai, Y., Markovic, M., Xu, C., Fadok, J.P., Lu, T., Poe, M.M., Xu, L., Cook, J.M., et al. (2015). Regulating anxiety with extrasynaptic inhibition. Nat. Neurosci. 18, 1493–1500. https://doi.org/10.1038/nn.4102.

5. Tye, K.M., Prakash, R., Kim, S.Y., Fenno, L.E., Grosenick, L., Zarabi, H., Thompson, K.R., Gradinani, V., Ramakrishnan, C., and Deisseroth, K. (2011). Amygdala circuitry mediating reversible and bidirectional control of anxiety. Nature 471, 356–362. https://doi.org/10.1038/nature09820.

6. Yu, K., Garcia da Silva, P., Albeau, D.F., and Li, B. (2016). Central amygdala somatostatin neurons gate passive and active defensive behaviors. J. Neurosci. 36, 6488–6496. https://doi.org/10.1523/JNEUROSCI.4419-15.2016.

7. Fadok, J.P., Krabbe, S., Markovic, M., Courtin, J., Xu, C., Massi, L., Botta, P., Bylund, K., Muller, C., Kovačević, A., et al. (2017). A competitive inhibitory circuit for selection of active and passive fear responses. Nature 542, 96–100. https://doi.org/10.1038/nature21007.

8. Janah, P.H., and Tye, K.M. (2015). From circuits to behaviour in the amygdala. Nature 517, 284–292. https://doi.org/10.1038/nature14188.

9. Beyeler, A., and Dabrowska, J. (2020). Neuronal diversity of the amygdala and the bed nucleus of the stria terminalis. Handb. Behav. Neurosci. 26, 63–100. https://doi.org/10.1016/b978-0-12-815134-1.00003-9.

10. Zeng, H., and Sanes, J.R. (2017). Neuronal cell-type and classification: challenges, opportunities and the path forward. Nat. Rev. Neurosci. 18, 530–546. https://doi.org/10.1038/nrn.2017.85.

11. McCullough, K.M., Morrison, F.G., Hartmann, J., Carlezon, W.A., Jr., and Ressler, K.J. (2018). Quantified coexpression analysis of central amygdala and extended amygdala subnuclei. eNeuro 5. ENEURO.0010-18.2018. https://doi.org/10.1523/ENEURO.0010-18.2018.

12. Wolfe, S.A., Sidhu, H., Patel, R.R., Kreifeldt, M., D’Ambrosio, S.R., Contet, C., and Roberto, M. (2019). Molecular, morphological, and functional characterization of corticotropin-releasing factor receptor 1-expressing neurons in the central nucleus of the amygdala. eNeuro 6. ENEURO.0087-19.2019. https://doi.org/10.1523/ENEURO.0087-19.2019.

13. Kim, J., Zhang, X., Muralidhar, S., LeBlanc, S.A., and Tonegawa, S. (2017). Basolateral to central amygdala CRF circuit facilitates learning and memory. Cell 162, 363–374. https://doi.org/10.1016/j.cell.2015.05.057.

14. Li, H., Penzo, M.A., Taniguchi, H., Kopec, C.D., Huang, J., and Zhang, X. (2021). Experience-dependent modification of a central amygdala fear circuit. Nat. Neurosci. 14, 332–339. https://doi.org/10.1038/nn.3322.

15. Yu, K., Ahrens, S., Zhang, X., Schiff, H., Ramakrishnan, C., Fenno, L., Deisseroth, K., Zhao, F., Luo, M.H., Gong, L., et al. (2017). The central amygdala controls learning in the lateral amygdala. Nat. Neurosci. 20, 1680–1685. https://doi.org/10.1038/s41593-017-0009-9.

16. Cai, H., Haubensak, W., Anthony, T.E., and Anderson, D.J. (2014). Central amygdala PKC-delta(+) neurons mediate the influence of multiple anorexigenic signals. Nat. Neurosci. 17, 1240–1248. https://doi.org/10.1038/nn.3767.

17. McCullough, K.M., Chatzinikos, C., Hartmann, J., Missig, G., Neve, R.L., Fenster, R.J., Carlezon, W.A., Jr., Daskalakis, N.P., and Ressler, K.J. (2003). Genome-wide translational profiling of amygdala Crh-expressing neurons reveals role for CREB in fear extinction learning. Nat. Commun. 11, 5180. https://doi.org/10.1038/s41467-020-18952-6.

18. Befort, K., Filliol, D., Ghate, A., Darcq, E., Matfias, A., Muller, J., Lardenois, A., Thibault, C., Dembele, D., Le Merrer, J., et al. (2008). Mu-opioid receptor activation induces transcriptional plasticity in the central extended amygdala. Eur. J. Neurosci. 27, 2973–2984. https://doi.org/10.1111/j.1460-9558.2008.06273.x.

19. Samineni, V.K., Grajailes-Reyes, J.G., Grajailes-Reyes, G.E., Tye, K.M., Copits, B.A., Pedersen, C., Ankudinov, E.S., Sackey, J.N., Sewell, S.B., Bruchas, M.R., and Gereau, R.W. (2021). Cellular, circuit and transcriptional framework for modulation of itch in the central amygdala. Elife 10, e68130. https://doi.org/10.7554/eLife.68130.

20. Zirlinger, M., Kreiman, G., and Anderson, D.J. (2001). Amygdala-enriched genes identified by microarray technology are restricted to specific amygdaloid subnuclei. Proc. Natl. Acad. Sci. USA 98, 5270–5275. https://doi.org/10.1073/pnas.091094698.
Cembrowski, M.S., Phillips, M.G., DiLisio, S.F., Shields, B.C., Winnubst, J., Chandrashekar, J., Bas, E., Winnubst, J., Menon, V., Graybuck, LT., Nguyen, T.N., Smith, K.A., Yao, Z., et al. (2018). Distinct descending motor cortex pathways and their roles in movement. Nature 563, 79–84. https://doi.org/10.1038/s41586-018-0642-9.

Economo, M.N., Viswanathan, S., Tasic, B., Bas, E., Winnubst, J., Menon, V., Graybuck, LT., Nguyen, T.N., Smith, K.A., Yao, Z., et al. (2018). Distinct descending motor cortex pathways and their roles in movement. Nature 563, 79–84. https://doi.org/10.1038/s41586-018-0642-9.

Cembrowski, M.S., and Spruston, N. (2019). Single-cell transcriptomics as a framework and roadmap for understanding the brain. J. Neurosci. Methods 326, 108353. https://doi.org/10.1016/j.neuroimage.2019.108353.

Economo, M.N., Viswanathan, S., Tasic, B., Bas, E., Winnubst, J., Menon, V., Graybuck, LT., Nguyen, T.N., Smith, K.A., Yao, Z., et al. (2018). Distinct descending motor cortex pathways and their roles in movement. Nature 563, 79–84. https://doi.org/10.1038/s41586-018-0642-9.

Kuerbitz, J., Arnett, M., Ehrman, S., Williams, M.T., Vorhees, C.V., Fisher, S.E., Winnubst, J., Menon, V., Graybuck, LT., Nguyen, T.N., Smith, K.A., Yao, Z., et al. (2018). Distinct descending motor cortex pathways and their roles in movement. Nature 563, 79–84. https://doi.org/10.1038/s41586-018-0642-9.

Kovalev, E., Shoham, S., and Soreq, H. (2013). Hippocampal microRNA-132 mediates stress-inducible cognitive deficits through its acetylcholinesterase target. Brain Str. Funct. 218, 59–72. https://doi.org/10.1007/s10049-011-0376-2.

Phipson, B., Zappia, L., and Oshlack, A. (2017). Gene length and detection bias in single cell RNA sequencing protocols. F1000Res. 6, 595. https://doi.org/10.12688/f1000research.11290.1.

McInnes, I., Healy, D.J., and Young, A.W. (2018). UMAP: uniform manifold approximation and projection for dimension reduction. Preprint at arXiv. https://doi.org/10.48550/arXiv.1802.03366.

Szpankowski, L., Fowler, B., Chen, P., et al. (2014). Low-coverage single-cell mRNA sequencing reveals cellular heterogeneity and activated signaling pathways in developing cerebral cortex. Nat. Biotechnol. 32, 1053–1058. https://doi.org/10.1038/nbt.2967.

Lein, E.S., Hawrylcz, M.J., Ao, N., Ayres, M., Bengis, A., Bernard, A., Boe, A.F., Boguski, M.S., Brookway, K.S., Byrnes, E.J., et al. (2007). Genome-wide atlas of gene expression in the adult mouse brain. Nature 445, 168–176. https://doi.org/10.1038/nature05433.

McInnes, L., Healy, J., and Melville, J. (2018). UMAP: uniform manifold approximation and projection for dimension reduction. Preprint at arXiv. https://doi.org/10.48550/arXiv.1802.03366.

Szpankowski, L., Fowler, B., Chen, P., et al. (2014). Low-coverage single-cell mRNA sequencing reveals cellular heterogeneity and activated signaling pathways in developing cerebral cortex. Nat. Biotechnol. 32, 1053–1058. https://doi.org/10.1038/nbt.2967.

Lein, E.S., Hawrylcz, M.J., Ao, N., Ayres, M., Bengis, A., Bernard, A., Boe, A.F., Boguski, M.S., Brookway, K.S., Byrnes, E.J., et al. (2007). Genome-wide atlas of gene expression in the adult mouse brain. Nature 445, 168–176. https://doi.org/10.1038/nature05433.
59. Lu, K.M., Evans, S.M., Hirano, S., and Liu, F.C. (2014). Dual role for Islet-1 in promoting striatogniral and repressing striatopallidal genetic programs to specify striatogniral cell identity. Proc. Natl. Acad. Sci. USA 111, E168–E177. https://doi.org/10.1073/pnas.1319138111.

60. Ehrman, L.A., Mu, X., Waclaw, R.R., Yoshida, Y., Vorhees, C.V., Klein, W.H., and Campbell, K. (2013). The LIM homeobox gene Isl1 is required for the correct development of the striatogniral pathway in the mouse. Proc. Natl. Acad. Sci. USA 110, E4026–E4035. https://doi.org/10.1073/pnas.1308275110.

61. Stanley, G., Gokce, O., Malenka, R.C., Sudhof, T.C., and Quake, S.R. (2020). Continuous and discrete neuron types of the mammalian brain. Nat. Protoc. https://doi.org/10.1038/nprot.2020.659.

64. Han, W., Tellez, L.A., Rangel, M.J., Jr., Motta, E., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21. https://doi.org/10.1093/bioinformatics/bts635.

69. R Development Core Team. (2008). R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing).

70. Satija, R., Farrell, J.A., Gennert, D., Schier, A.F., and Regev, A. (2015). Spatial reconstruction of single-cell gene expression data. Nat. Biotechnol. 33, 495–502. https://doi.org/10.1038/nbt.3192.

71. Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., 3rd, Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive integration of single-cell data. Cell 177, 1886–1902.e21. https://doi.org/10.1016/j.cell.2019.05.031.

72. Beyeler, A., Chang, C.J., Silvestre, M., Léveque, C., Namburi, P., Wildes, C.P., and Tye, K.M. (2018). Organization of valence-encoding and projection-defined neurons in the basolateral amygdala. Cell Rep. 22, 905–918. https://doi.org/10.1016/j.celrep.2017.12.097.

73. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat. Methods 9, 676–682. https://doi.org/10.1038/nmeth.2019.

74. Arganda-Carreras, I., Sorzano, C.O.S., Thévenaz, P., Muñoz-Barrutia, A., Kybic, J., Marabini, R., Carazo, J.M., and Ortíz-de Solorzano, C. (2010). Non-rigid consistent registration of 2D image sequences. Phys. Med. Biol. 55, 6215–6242. https://doi.org/10.1088/0031-9155/55/20/012.

75. Levine, J.H., Simonds, E.F., Bendall, S.C., Davis, K.L., Amir, E.A.D., Tadmor, M.D., Lavín, O., Fienberg, H.G., Jager, A., Zunder, E.R., et al. (2015). Data-driven phenotypic dissection of AML reveals progenitor-like cells that correlate with prognosis. Cell 162, 184–197. https://doi.org/10.1016/j.cell.2015.05.047.

76. Cao, J., Spielmann, M., Qiu, X., Huang, X., Ibrahim, D.M., Hill, A.J., Zhang, F., Mundlos, S., Christiansen, L., Steemers, F.J., et al. (2019). The single-cell transcriptional landscape of mammalian organogenesis. Nature 566, 496–502. https://doi.org/10.1038/s41586-019-0969-x.

77. Levine, J.H., Simonds, E.F., Bendall, S.C., Davis, K.L., Amir, E.A.D., Tadmor, M.D., Lavín, O., Fienberg, H.G., Jager, A., Zunder, E.R., et al. (2015). Data-driven phenotypic dissection of AML reveals progenitor-like cells that correlate with prognosis. Cell 162, 184–197. https://doi.org/10.1016/j.cell.2015.05.047.

78. Cao, J., Spielmann, M., Qiu, X., Huang, X., Ibrahim, D.M., Hill, A.J., Zhang, F., Mundlos, S., Christiansen, L., Steemers, F.J., et al. (2019). The single-cell transcriptional landscape of mammalian organogenesis. Nature 566, 496–502. https://doi.org/10.1038/s41586-019-0969-x.

79. Qiu, X., Hill, A., Packer, J., Lin, D., Ma, Y.A., and Trapnell, C. (2017). Single-cell mRNA quantification and differential analysis with Censor. Nat. Methods 14, 309–315. https://doi.org/10.1038/nmeth.4150.

80. Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M., Lennon, N.J., Livak, K.J., Mikkelson, T.S., and Rinn, J.L. (2014). The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nat. Biotechnol. 32, 381–386. https://doi.org/10.1038/nbt.2859.
STAR METHODS

KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---------------------|--------|------------|
| **Antibodies**      |        |            |
| Anti-Nr2f2 mouse IgG2a monoclonal antibody | Abcam | RRID: AB_742211 |
| Anti-PKC delta rabbit monoclonal antibody | Thermo Fisher | RRID: AB_2809759 |
| Anti-Islet 1 mouse IgG1 monoclonal antibody | Novus Biologicals | NBP2-45452 |
| **Critical commercial assays** |        |            |
| Nr2f2 ISH probe | Advanced Cell Diagnostics | 480301-T1 |
| Isl1 ISH probe | Advanced Cell Diagnostics | 451931-T2 |
| Certf ISH probe | Advanced Cell Diagnostics | 452281-T3 |
| Gad1 ISH probe | Advanced Cell Diagnostics | 400951-T4 |
| Cxcl14 ISH probe | Advanced Cell Diagnostics | 459741-T5 |
| Cck ISH probe | Advanced Cell Diagnostics | 402271-T6 |
| Prkcd ISH probe | Advanced Cell Diagnostics | 441791-T7 |
| Fxyd6 ISH probe | Advanced Cell Diagnostics | 430971-T8 |
| Sst ISH probe | Advanced Cell Diagnostics | 404631-T9 |
| Crh ISH probe | Advanced Cell Diagnostics | 316091-T10 |
| Nts ISH probe | Advanced Cell Diagnostics | 420441-T11 |
| Cartpt ISH probe | Advanced Cell Diagnostics | 432001-T12 |
| **Deposited data** |        |            |
| scRNA-seq data | This paper | GEO: GSE184666 |
| **Experimental models: Organisms/strains** |        |            |
| C57BL/6 mice | Jackson Laboratory | RRID: IMSR_JAX:000664 |
| **Software and algorithms** |        |            |
| R v4 | R Foundation for Statistical Computing, Vienna, Austria, 2008 | RRID:SCR_001905 |
| Seurat v4 | Satija et al., 2015 | RRID:SCR_007322 |
| Fiji v2 | Schindelin et al., 2012 | RRID:SCR_002285 |
| Prism v4 | https://www.graphpad.com/scientific-software/prism/ | RRID:SCR_002798 |
| **Code** | This paper | https://doi.org/10.6084/m9.figshare.17031605 |
| **Other** |        |            |
| Retrograde AAV | rAAV2-retro-CAG-tdT | Janelia Viral Core |
| Retrograde AAV | rAAV2-retro-CAG-GFP | Janelia Viral Core |

RESOURCE AVAILABILITY

**Lead contact**
Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Mark Cembrowski (mark.cembrowski@ubc.ca).

**Materials availability**
This study did not generate new unique reagents.
Data and code availability

- Single-cell RNA-seq data have been deposited at GEO and are publicly available as of the date of publication. Accession numbers are listed in the key resources table.
- All original code has been deposited and is publicly available as of the date of publication. DOIs are listed in the key resources table.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Unless otherwise noted, experiments used male wild-type C57BL/6 mice between 8–20 weeks of age, with mice maintained on a 12-h light-dark cycle with ad libitum access to food and water. All experimental procedures were approved by the University of British Columbia Animal Care Committee and the Janelia Research Campus Institutional Animal Care and Use Committee.

Single-cell RNA sequencing data acquisition and analysis

To isolate neurons from the CEA, brain sections were obtained from 3 male C57BL/6 mice (≥8 weeks of age). The CEA was microdissected and dissociated, and manual purification was used to capture cells with neuron-like cell bodies in capillary needles in approximately 0.1–0.5 mL ACSF cocktail. Cells were placed into 8-well strips containing 3 μL of cell collection buffer (0.1% Triton X-100, 0.2 U/μL RNase inhibitor (Lucigen, Middleton, WI), and generally processed according to published methodology. Specifically, each strip of cells was flash frozen on dry ice, then stored at −80°C until cDNA synthesis. Cells were lysed by adding 1 μL lysis mix (50 mM Tris pH 8.0, 5 mM EDTA pH 8.0, 10 mM DTT, 1% Tween 20, 1% Triton X-100, 0.1 g/L Proteinase K (Roche), 2.5 mM dNTPs (Takara), and ERCC Mix 1 (Thermo Fisher) diluted to 1 x 10^-6) and 1 μL 10 μM barcoded RT primer (E3V6NEXT primer from modified to add a 1 bp spacer before the barcode, extending the barcode length from 6 bp to 8 bp, and designing the 384 barcodes to tolerate 1 mismatch error correction). The samples were incubated for 5 min at 50°C to lyse the cells, followed by 20 min at 80°C to inactivate the Proteinase K. Reverse transcription master mix (2 μL SX buffer (Thermo Fisher Scientific), 2 μL 5M Betaine (Sigma-Aldrich, St. Louis, MO), 0.2 μL 50 μM E5V6NEXT template switch oligo (Integrated DNA Technologies, Coralville, IA), 0.1 μL 200 U/μL Maxima H-RT (Thermo Fisher Scientific), 0.1 μL 40 U/μL RNasin (Lucigen), and 0.6 μL nuclease-free water (Thermo Fisher Scientific) was added to the approximately 5.5 μL lysis reaction and incubated at 42°C for 1.5 h, followed by 10 min at 75°C to inactivate reverse transcriptase. PCR was performed by adding 10 μL 2X HiFi PCR mix (Kapa Biosystems) and 0.5 μL 60 μM SingV6 primer with the following conditions: 98°C for 3 min, 20 cycles of 98°C for 20 s, 64°C for 15 s, for 4 min, with a final extension step of 5 min at 72°C. Samples were pooled across the plate to yield approximately 2 mL pooled PCR reaction. From this, 500 μL was purified with 300 μL AmPure XP beads (0.6x ratio; Beckman Coulter, Brea, CA), washed twice with 75% ethanol, and eluted in 20 μL nuclease-free water. The cDNA concentration was determined using Qubit High-Sensitivity DNA kit (Thermo Fisher Scientific).

Nine plates were analyzed in total, with six hundred pg cDNA from each plate of cells used in a modified Nextera XT (Illumina, San Diego, CA) library preparation with 5 μM PSNEXTPT5 primer. The resulting libraries were purified according to the Nextera XT protocol (0.6x ratio) and quantified by qPCR using Kapa Library Quantification (Kapa Biosystems). Plates were pooled on a NextSeq 550 high-output flowcell with 26 bases in read 1, 8 bases for the i7 index, and 50 bases for read 2. Alignment and count-based quantification of single-cell data was performed by removing adapters, tagging transcript reads to barcodes and UMIs, and aligned to the mouse genome via STAR.

A total of 829 cells were captured, with 826 retained throughout our analysis (3 cells eliminated due to Seurat thresholds; see below). These 826 cells retained for analysis exhibited 4.5 ± 0.9 thousand expressed genes/cell from 141 ± 73 thousand transcripts/cell (mean ± SD). No blinding or randomization was used for the construction or analysis of this dataset. No a priori sample size was determined for the number of animals or cells to use; note that previous methods have indicated that several hundred cells from a single animal is sufficient to resolve heterogeneity within excitatory neuronal cell types.
Computational analysis was performed in R (RRID:SCR_001905) using a combination of Seurat v4 (RRID:SCR_007322) and custom scripts. To analyze our data, a Seurat object was created via CreateSeuratObject(min.cells = 3, min.features = 200) (minimum thresholding resulting in the exclusion of 3 cells), normalized via NormalizeData(), variable features identified via FindVariableFeatures(selection.method='vst', n.features=2000) and scaled via ScaleData(). Data was processed via RunPCA(), JackStraw(num.replicate=100), RunTSNE(), FindNeighbors(), FindClusters(resolution=0.6), and RunUMAP(reduction='pca'), with 30 dimensions used throughout the analysis. This processed Seurat object was then used for downstream analysis. Subpopulation-specific enriched genes obeying $p_{\text{ADJ}} < 0.05$ were obtained with Seurat via FindMarkers(), where $p_{\text{ADJ}}$ is the adjusted p value from Seurat based on Bonferroni correction. Functionally relevant differentially expressed genes were obtained using FindMarkers(), allowing for both cluster-specific enriched and depleted genes obeying $p_{\text{ADJ}} < 0.05$, and manually identified for functional relevance. Raw and processed scRNA-seq datasets have been deposited in the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus under GEO: GSE184666.

**Chromogenic in situ hybridization**

Analysis of chromogenic in situ hybridization was completed with Fiji (RRID:SCR_002285) using images from the Allen Mouse Brain Atlas, $^{35}$ with the following series: GAD1: 79556706; Calcif: 74988670; Chh: 292; Cck: 200; Cartpt: 72077479; Cxcl14: 74272041; Nr2f2: 112646890; Isl1: 596. Images of the CEA were obtained at three points, corresponding to the anterior ($-0.82$ to $-0.94$ from bregma), intermediate ($-1.22$ to $-1.34$ from bregma) and posterior ($-1.70$ to $-1.82$ from bregma) regions. For each image, the boundary of the CEA was traced using eight points. Each image was inverted and processed with a Gaussian blur filter before a manual threshold was applied to binarize images and separate signal from background. The Watershed function was then applied to assist in segmentation of adjacent cells. The locations of cells between 75 and 400 $\mu$m$^2$ were measured within the CEA using Analyze Particles. To assist in the comparison of the spatial expression of multiple marker genes, the location of segmented cells was translated onto a common CEA template using a Procrutes transformation in R. $^{72}$ The locations of cells were plotted relative to the dorsal-most point of the CEA adjacent to the medial wall of the basolateral amygdala.

**Tissue processing for histology**

For mFISH and viral tracing, brains were extracted from mice immediately after trans-cardiac perfusion (10 mL of PBS followed by 50 mL of 4% PFA, pH 7.4). Brains were then post-fixed in 4% PFA for 24 h, and subsequently cryoprotected in 30% sucrose for at least 48 h. Tissue was sectioned in the coronal plane using a cryostat, at a thickness of either 20 $\mu$m (mFISH) or 100 $\mu$m (projection mapping). Tissue sections were then mounted onto glass slides and stored at $-80^\circ$C until use (20 $\mu$m sections), or were processed the following day (100 $\mu$m sections). For retrograde viral tracing, tissue was counterstained with DAPI (1:1000) for 10 min and a coverslip added with PVA-Dabco mounting medium.

For immunohistochemistry, mice were deeply anesthetized with isoflurane and underwent transcardial perfusion (10 mL of 1x PBS followed by 50 mL of 4% PFA in 1x PBS, pH 7.4). Brains were extracted and post-fixed in 4% PFA overnight. Brains were then sectioned in the coronal plane using a vibrating blade microtome (Leica VT 1200S, Leica Microsystems, Wetzlar, Germany) at 40 $\mu$m thickness.

**Multiplexed fluorescent in situ hybridization data acquisition**

All probes were purchased from Advanced Cell Diagnostics (Hayward, CA), and were as follows: Nr2f2 (480301-T1), Isl1 (451931-T2), Calcif (452281-T3), Gad1 (400951-T4), Cxcl14 (459741-T5), Cck (402271-T6), Prkcd (441791-T7), Fxyd6 (430971-T8), Sst (404631-T9), Crh (316091-T10), Nts (420441-T11), Cartpt (432001-T12). Generally, our mFISH approach proceeded according to published work. $^{73,74}$ Briefly, mice were selected for mFISH, and tissue was processed for histology as above. For use, the tissue underwent pretreatment and antigen retrieval per the User Manual for Fixed Frozen Tissue (Advanced Cell Diagnostics). All 12 probes with unique tails (T1-T12) were hybridized to the tissue and amplified. To visualize four targets within one round of imaging, cleavable fluorophores with matching tails (T1-T12) were added four at a time. In the first round, the section was counterstained with DAPI mounted in ProLongGold antifade mounting medium, and coverslipped for imaging the subsequent day. In the other two rounds (T5-T8 and T9-T12), the sections were mounted in ProLongGold antifade mounting medium before coverslipping. At the end of each round, the tissue was decoverslipped by soaking undisturbed in 4X SSC until the coverslip fell off the slide.
mFISH images were acquired with a 63x objective on a SP8 Leica white light laser confocal microscope (Leica Microsystems, Concord, Ontario, Canada). Z-stacks were obtained with a step size of 0.45 μm for each imaging round. Final composite images represent pseudocolored maximum intensity projections, with channelsopaquely overlaying one another ordered from highest to lowest expression. Processing of multiplexed FISH images generally followed our previously published analysis pipeline for FIJI. In brief, the DAPI signal from each round was used to rigidly register images across rounds, followed by nonlinear elastic registration via bUnwarpJ to correct any nonlinear tissue warping. Individual nuclei were segmented from DAPI images and dilated by a factor of 5 μm to account for the surrounding cytosol. Signal from each probe was binarized, with a manual threshold chosen for each probe image. The number of pixels contained within each segmented ROI (i.e., cell) was summed and normalized to the total pixel area of the cell to give counts-per-area, which were converted into percent area covered (PAC) values by multiplying by 100.

**Multiplexed fluorescent in situ hybridization analysis**

A total of 3 mature male C57BL/6 mice, comprising 8 total brain sections (see Figure S4), were used for mFISH. Cells were required to express GAD1 to be included, and to have at least one marker gene expression to an appreciable (85th percentile) level, resulting in 3,959 cells retained for further analysis. UMAP dimensionality reduction was performed on within-gene z-scored PAC values using the umap package (15 nearest neighbors, all other parameters default), and cells were clustered on the pooled dataset (i.e. across anterior-posterior sections and animals) using a Leiden community detection algorithm via the Monocle package in R. A resolution of $4 \times 10^{-4}$ was used, producing 7 clusters with strong agreement between relative locations of cell groups in UMAP embedding and their cluster assignment. Statistical tests for marker genes for each cluster are shown for the genes with greatest fold enrichment, based on the ratio of median gene expression of that cluster relative to the pooled remaining data.

To gain insight into the spatial organization of cell types identified by clustering in our dataset, we mapped spatial locations of cell types onto existing lateral, medial, and capsular subdivisions of the CEA in both qualitative and quantitative fashion. To do this, DAPI images of anterior, intermediate, and posterior CEA sections were subdivided based on previously-characterized anatomical landmarks, conducted blind to gene expression. To qualitatively compare the spatial organization of cluster-defined cell types to anatomical subdivisions, two-dimensional kernel density estimation was applied to a dataset of the spatial locations of individual cells via the stat_density2d function in R. The resultant density plots reflected a transformation from the space of single cells to one of density contours, colored according to the cell type represented in an area of high cell density. Subdivisions were overlayed onto these plots, and the degree to which cell type density conformed to subdivision boundaries was examined. To quantitatively compare anatomical subdivisions and the spatial distribution of cell types, Monte-Carlo shuffle analysis with $N = 1000$ simulations was used to identify cell types overrepresented in each subdivision at a given anterior-posterior position.

**Immunohistochemical staining, imaging, and analysis**

Immunohistochemical staining was performed on free-floating tissue sections. Sections were washed in 1xPBS (3 × 5 minutes each), then incubated in blocking buffer (5% NGS in 0.3% Triton-PBS) for 1 h at room temperature on a shaking platform. Subsequently, tissue was incubated at 4°C overnight in the following primary antibody solution: anti-Nr2f2 mouse IgG2a monoclonal antibody (1:50, cat# ab41859, Abcam, RRID: AB_742211), anti-PKC delta rabbit monoclonal antibody (1:50, cat# MA5-32482, Thermo Fisher, RRID: AB_2809759), anti-Islet 1 mouse IgG1 monoclonal antibody (1:50, cat# NBP2-45452, Novus Biologicals, RRID: N/A) diluted in 0.3% Triton-PBS. Tissue was washed in 0.3% Triton-PBS (3 × 5 minutes each), followed by a 2–3 h room temperature incubation in Alexa Fluor secondary antibodies (Thermo Fisher Scientific Inc., Waltham, MA). Sections were subsequently washed in 1xPBS (3 × 5 minutes each), mounted to slides and coverslipped with mounting media containing DAPI (H-1200, Vector Laboratories, Burlingame, CA).

Imaging was performed on a LSM 880 confocal microscope (Carl Zeiss Microscopy, Jena, Germany) with a 20x objective at 2x zoom. Images are shown as maximum intensity projections, tiled in X and Y to capture the extent of the CEA. Post-processing and analyses of images were performed using FIJI, including brightness adjustments and pseudocolouring to facilitate visual comparisons across channels.
Individual cells were segmented from images based upon DAPI nuclei, as performed in mFISH analysis. To facilitate comparisons to mFISH analysis, the immunohistochemical signals underwent thresholding, and the normalized optical areas covered by each of Prkcd, Nr2f2, and Isl1 in DAPI-segmented nuclei were calculated. The relative protein abundance was then computed, on a cell-by-cell basis, by normalizing these values by the total abundance of all three proteins. Visualized signal intensity for each of these abundances is shown in colored range values, with cells lacking all three proteins shown black. For shuffled analysis, values of each protein were randomly shuffled within a given section, and total abundance values recomputed and visualized.

**Viral tracing and analysis**

Stereotaxic surgery was performed per published protocols, where either rAAV2-retro-CAG-GFP or rAAV2-retro-CAG-tdT retrograde viruses were injected into regions that received projections from the CEA. These regions included the nucleus basalis stria terminalis (0.25 A/P, 0.9 M/L, −4.5 D/V), lateral hypothalamus (−1.0 A/P, 1.3 M/L, −5.75 D/V), periaqueductal gray (−4.5 A/P, 0.5 M/L, −2.83D/V), parabrachial nucleus (−5.0 A/P, 1.35 M/L, −4.0 D/V), substantia nigra (−3.0 A/P, 1.5 M/L, −4.8 D/V) and parafascicular thalamus (−2.20 A/P, 0.70 M/L, −3.60 D/V). Injections were completed using pulled glass pipettes, with 100 nL of virus injected at a given site. Following injections, pipettes remained at the injection site for 3 min to permit diffusion of the virus. Mice were sacrificed 6–7 days following surgery.

Imaging for retrograde viral labeling was completed using an LSM 880 confocal microscope (Carl Zeiss Microscopy, Jena, Germany) with a 20x objective. Images are shown as either single optical sections or maximum intensity projections, tiled in X and Y to encompass the CEA extent. Post-processing and analysis of images was performed in FIJI, and included uniform brightness adjustments and pseudocoloring.

For correlating projection targets with marker gene expression, viral-labeled cells were first imaged with a 63x objective on a SP8 Leica white light laser confocal microscope. Projections to the substantia nigra (n = 17 cells), parabrachial nucleus (n = 28 cells), periaqueductal gray (n = 27 cells), and lateral hypothalamus (n = 39 cells) were examined, with both anterior and intermediate sections from n = 2 animals used in all projections (in total, comprising 111 projection neurons from 9,981 segmented CEA cells). Posterior sections were not used due to the small number of CEA projectors in these sections (Figure 7). Afterward, fluorescent tracers were cleaved as in, and tissue was hybridized with the following probes from Advanced Cell Diagnostics: Nr2f2 (480301-T1), Isl1 (451931-T2), Prkcd (441791-T3), and Sst (404631-T4) (note that Sst expression was not examined further). Expression was detected and quantified as in Multiplexed fluorescent in situ hybridization data acquisition.

**QUANTIFICATION AND STATISTICAL ANALYSIS**

In general, box-and-whisker plots show distribution of gene expression within cell populations according to the following conventions: midline denotes median, boxes denote first and third quartiles, whiskers denote remaining data points up to at most 1.5 * interquartile range, with outlier values beyond whiskers not shown unless otherwise specified. Unless otherwise noted, pADJ-values are computed via a Mann-Whitney U test with FDR correction for multiple comparisons. Statistical significance for adjusted pADJ-values is denoted as: ns: pADJ ≥ 0.05; *: pADJ < 0.05, **: pADJ < 0.01, ***: pADJ < 0.001.