Article

Neuron

Spatial Gene-Expression Gradients Underlie Prominent Heterogeneity of CA1 Pyramidal Neurons

Highlights

- CA1 pyramidal cells (CA1 PCs) exhibit pronounced transcriptional heterogeneity
- Greatest heterogeneity is across long (dorsal-ventral; D-V) axis
- D-V differences in CA1 PCs are graded and continuous
- D-V differences within CA1 are comparable to differences across cell classes

Authors

Mark S. Cembrowski, Julia L. Bachman, Lihua Wang, Ken Sugino, Brenda C. Shields, Nelson Spruston

Correspondence

sprustonn@janelia.hhmi.org

In Brief

Using next-generation RNA sequencing, Cembrowski et al. show that CA1 pyramidal cells (CA1 PCs) exhibit marked variability across the long hippocampal axis. This heterogeneity emerges from a spectrum of gene-expression gradients, producing a highly variable continuum of CA1 PCs.

Accession Numbers

GSE67403

Cembrowski et al., 2016, Neuron *89*, 351–368
January 20, 2016
http://dx.doi.org/10.1016/j.neuron.2015.12.013

CellPress

Neuron
Article

Spatial Gene-Expression Gradients Underlie Prominent Heterogeneity of CA1 Pyramidal Neurons

Mark S. Cembrowski,[1] Julia L. Bachman,[1] Lihua Wang,[1] Ken Sugino,[1] Brenda C. Shields,[1] and Nelson Spruston[1,*]
[1]Janelia Research Campus, Howard Hughes Medical Institute, 19700 Helix Drive, Ashburn, VA 20147, USA
*Correspondence: sprustonn@janelia.hhmi.org
http://dx.doi.org/10.1016/j.neuron.2015.12.013

SUMMARY

Tissue and organ function has been conventionally understood in terms of the interactions among discrete and homogeneous cell types. This approach has proven difficult in neuroscience due to the marked diversity across different neuron classes, but it may be further hampered by prominent within-class variability. Here, we considered a well-defined canonical neuronal population—hippocampal CA1 pyramidal cells (CA1 PCs)—and systematically examined the extent and spatial rules of transcriptional heterogeneity. Using next-generation RNA sequencing, we identified striking variability in CA1 PCs, such that the differences within CA1 along the dorsal-ventral axis rivaled differences across distinct pyramidal neuron classes. This variability emerged from a spectrum of continuous gene-expression gradients, producing a transcriptional profile consistent with a multifarious continuum of cells. This work reveals an unexpected amount of variability within a canonical and narrowly defined neuronal population and suggests that continuous, within-class heterogeneity may be an important feature of neural circuits.

INTRODUCTION

Historically, a cell-type-based reductionist approach has been the dominant explanatory paradigm in tissue and organ physiology. In this framework, the system of interest is deconstructed into discrete cell types, such that a given type is structurally and functionally homogeneous when viewed relative to other types. Subsequently, the emergent processing of the system as a whole is understood through clarifying the contributions and interactions of different cell types. Although this approach has elucidated the operation of most tissues and organs in the body, insightful application to neuronal systems has thus far proved challenging. In part, this difficulty stems from a difference of scale, in that the cellular diversity in the brain is markedly greater than other systems; however, differences in organizational principles may also be present, wherein within-class variability may be an important aspect of neuronal systems (Soltesz, 2006).

Neurons in the brain are commonly divided into discrete cell types based upon morphology, marker-gene expression, electrophysiology, and location within their respective circuit motifs. Among the most prominent and highly studied motifs in the mammalian brain are those that exhibit a repeating and spatially extended geometry, for example, the hippocampus, the neocortex, and the cerebellum. Although both upstream inputs and downstream targets vary within any of these individual regions, it is commonly assumed that the stereotyped structure of these circuits lends itself to a conserved input-output transformation across the repeated motif.

In particular, the hippocampus is a model system to clarify the extent to which higher-level (e.g., functional and behavioral) and lower-level (e.g., cellular and circuit) properties covary across space. A wealth of evidence has illustrated that there is profound functional segregation across the long hippocampal axis (Moser and Moser, 1998; Strange et al., 2014), with both input to (Dolorfo and Amaral, 1998) and output from (Groenewegen et al., 1987; Kishi et al., 2006; Risold and Swanson, 1996) the hippocampus exhibiting graded topographical mapping across the long axis.

Do the intrinsic properties of the hippocampus also vary along this axis, or is hippocampal circuitry fixed and performing a conserved computation on its varied inputs? This question has been systemically investigated previously using gene-expression studies of CA3 and CA1 hippocampal pyramidal cells. From the Allen Brain Atlas (ABA) genome-wide in situ hybridization (ISH) database (Lein et al., 2007), nine discrete molecular subdomains have been identified along the dorsal-ventral axis in CA3, corresponding to hundreds of genes enriched in a subregion-specific manner (Thompson et al., 2008). In CA1, discrete cell types (Zeisel et al., 2015) and molecular domains (Dong et al., 2009) have also been proposed, although overall gene-expression differences have been suggested to be markedly smaller, with two independent population-level studies each identifying ~25 genes enriched at either of the poles (Dong et al., 2009; Leonardo et al., 2006).

The finding of minimal differences in CA1 pyramidal cells is perhaps surprising, as several lines of anatomical and physiological evidence have identified differences in CA1 in each of the dorsal-ventral (Amaral and Witter, 1989; Dougherty et al., 2012, 2013; Malik et al., 2015), proximal-distal (Graves et al., 2012; Igarashi et al., 2014; Jarsky et al., 2008), and superficial-deep (Lee et al., 2014; Mizuseki et al., 2011; Slomianka et al., 2011) axes of the hippocampus. Moreover, the presence of discrete subdomains in CA1 is also seemingly discordant with the graded topography of extrinsic hippocampal connectivity (Strange et al., 2014).

Motivated by this, we reexamined CA1 pyramidal cell variability in a systematic fashion by using next-generation RNA sequencing (RNA-seq) (Shin et al., 2014) in combination with in situ hybridization, immunohistochemistry, and electrophysiology. Using this approach, we show that marked differences are present in CA1 pyramidal cells at transcriptional, proteomic, and functional levels. This variability is much greater than previously thought, such that within-CA1 variability can rival the overall difference between CA1 and CA3 pyramidal cell populations. Strikingly, CA1 obeys a foundationally different organizational architecture than CA3: CA1 is best viewed as a population of neurons with large and continuous transcriptional variation along the dorsal-ventral axis, with no evidence from gene expression for tractable subdivision of this population into discrete populations with respect to anatomical position. This continuous but pronounced transformation of transcriptional identity across the long axis is consistent with the graded scheme of extrinsic connectivity and provides cellular insight into the functional heterogeneity at the hippocampal poles. This graded change of neuronal identity may be a general feature of supposedly stereotyped circuit motifs in the mammalian brain.

RESULTS

Strategy to Obtain the Transcriptome across the Three CA1 Axes

In the rodent, area CA1 of the hippocampus has a C shape that extends millimeters in the anterior-posterior, medial-lateral, and dorsal-ventral axes (Figure 1A; image from Brain Explorer 2, Allen Brain Atlas; Lau et al., (2008). As viewed in coronal sections, near its most rostral point, CA1 occupies approximately the top half of the continuous CA field (Figure 1B). More caudally, the full dorsal-ventral extent of the CA1 region can be visualized; here, the most lateral band of cells comprises CA1.

Conventionally, this three-dimensional structure is often described by three "natural" axes. The most prominent axis (dorsal-ventral) is typically referred to as the long axis and spans the length of CA1 (Figure 1C). Two-dimensional cross-sections perpendicular to the long axis reveal the proximal-distal axis (spanning from the CA2/CA1 "proximal" border to the CA1/subiculum "distal" border; Figure 1D) and the superficial-deep axis (spanning from "superficial" cell bodies located closest to stratum radiatum to "deep" cell bodies located closest to stratum oriens; Figure 1E).

To examine transcriptional variability in each of these three axes, we adopted a general approach to perform RNA-seq on labeled cells at the extreme locations of each axis (Figure 1F; see Experimental Procedures). First, for a given axis, we identified a labeling strategy that would allow fluorophore expression to be restricted to CA1 pyramidal cell subpopulations of interest. Second, we microdissected the associated regions and dissociated the tissue. Third, we manually purified for fluorescent cells (~100 cells/sample) (Hempel et al., 2007). Fourth, we constructed cDNA libraries that were then sequenced to produce raw RNA-seq datasets. For the ensuing analysis, we used the TuxedoSuite pipeline, which reports gene-expression values in FPKM (fragments per kilobase of exon per million fragments mapped) and uses a false discovery rate (FDR) <0.05 to assign genes as being differentially expressed in pairwise comparisons.

Pronounced Transcriptional Variability in the Dorsal-Ventral Axis

We began by investigating transcriptional differences in the dorsal-ventral axis. This axis has been previously investigated using both microarray (Leonardo et al., 2006) and ISH (Dong et al., 2009), providing a readily available comparison between RNA-seq and other techniques. To label cells across the long axis of CA1, we used the Vipr2-Cre driver line crossed to the Ai9 Cre-dependent RFP (tdTomato) reporter (hereafter, the Vipr2 line). In double-positive animals, RFP expression within CA1 was found to be restricted to cells in the pyramidal cell layer along the long axis (Figure 2A), with 100% of RFP-labeled cells being positive for the CA1 pyramidal cell marker Wfs1 (Hitti and Siegelbaum, 2014; Kohara et al., 2014) in dorsal CA1 (Figure 2B). Consistent with this, no overlap was seen with PV+ interneurons, which also have cell bodies in the pyramidal cell layer (Figure 2C). In addition, no RFP colabeling was seen with NPY+, SST+, or GABA+ interneurons (Figure S1A, available online).

Microdissecting the dorsal and ventral poles in Vipr2 mice, we obtained dorsal and ventral CA1 transcriptomes (Figure 1F). Transcriptional variability between different biological replicates (i.e., identical regions in different animals) was minimal ($r = 0.99 \pm 0.01$, Pearson correlation coefficient, mean ± SEM; Figures 2D and S2A). However, when comparing replicates across regions, the correlation coefficient dropped markedly (average replicate correlation $r = 0.93 \pm 0.01$, mean ± SEM; Figure S2A). Strikingly, comparing the averaged dorsal-to-ventral pyramidal cell transcriptome revealed hundreds of differentially expressed genes along the long axis (Figure 2E).

This result was surprising, given that two independent population-level studies (Dong et al., 2009; Leonardo et al., 2006) had found only tens of genes that that were regionally enriched (n = 26 in ISH; n = 23 in microarray with criterion of >3-fold enrichment at one pole; Figure 2F). Despite the small number of regionally restricted genes in either study, the overlap of genes between the two studies was small (n = 3), indicating that each study may have identified only a small subset of the total number of regionally enriched genes. Indeed, when compared with previous studies, we found a 10-fold greater number of regionally enriched genes (Figure 2F; see Experimental Procedures).

Quality and External Validation of RNA-Seq Profiles

Our approach of sorting for labeled CA1 pyramidal cells should produce relatively pure CA1 transcriptomes. However, there remained the possibility that off-target cells may still be inadvertently included in the final pool of purified neurons and thus contaminate the resulting dataset. As such, we proceeded to characterize the purity of our transcriptomes. Examining cell-type-specific markers for off-target cells in our RNA-seq datasets, we found that each individual biological replicate showed no resolvable presence of gene cohorts associated with interneurons, CA2 pyramidal cells, subiculum pyramidal cells, or nonneuronal cells (Figure S2D). To quantitatively identify the extent that contamination should be resolvable in our

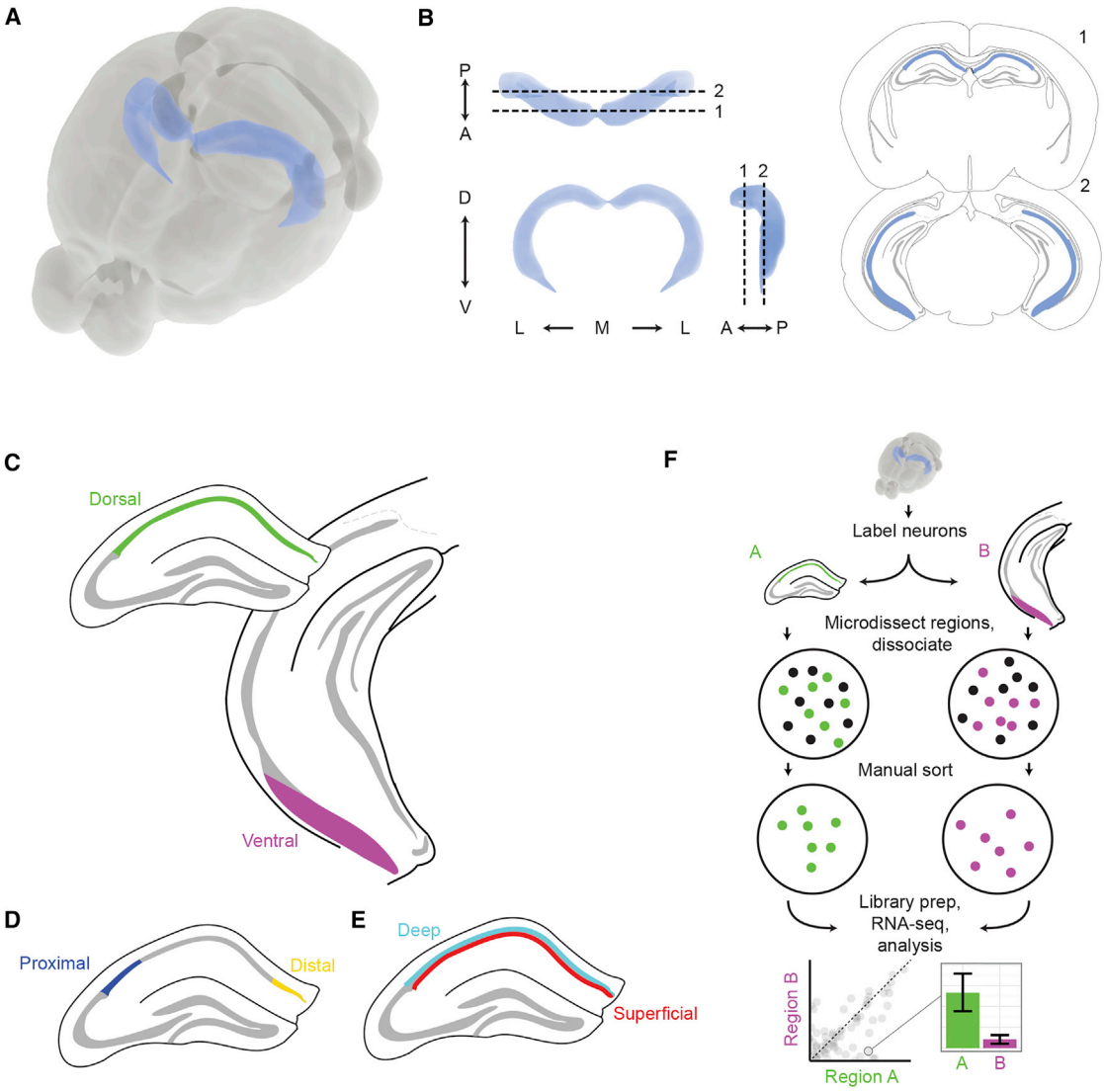


Figure 1. Interrogating the CA1 Transcriptome in Three Complementary Axes

(A) Area CA1 (blue) is highlighted in the mouse brain (gray). Image from Brain Explorer 2, Allen Brain Atlas.

(B) Left: CA1 in the orthogonal anterior-posterior (A-P)/medial-lateral (M-L)/dorsal-ventral (D-V) axes. Right: two coronal sections along the A-P axis illustrating dorsal and ventral CA1.

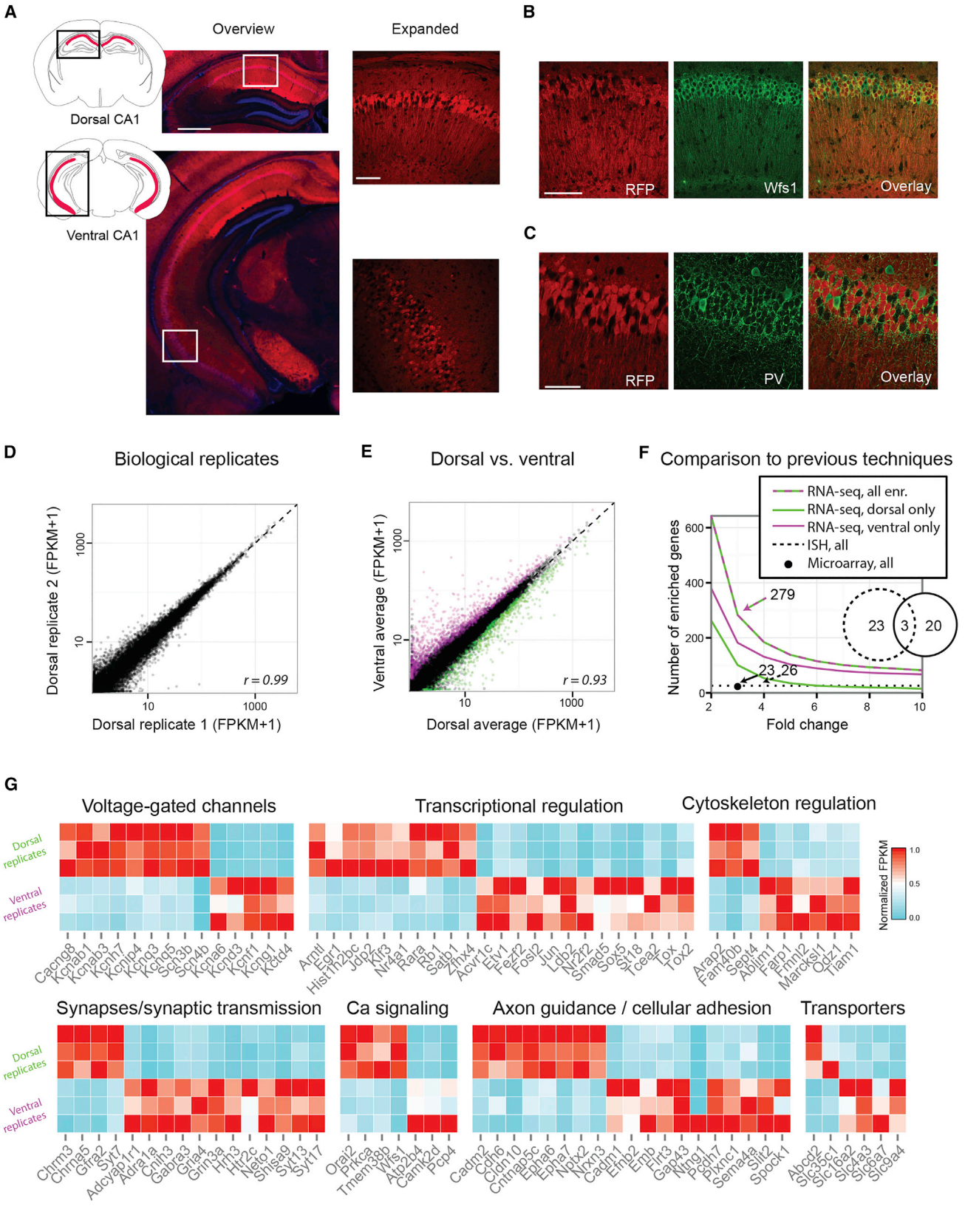
(C–E) The extremes of CA1 are highlighted in the "natural" dorsal-ventral (C), proximal-distal (D), and superficial-deep (E) axes.

(F) The general workflow to generate and analyze RNA-seq libraries; see Experimental Procedures. See also Figures S1–S3.

datasets, we profiled the transcriptomes for two canonical interneuron classes (PV+ interneurons and SST+ interneurons; see Experimental Procedures). Some marker genes for these cell types were present at >1,000 FPKM (Figure S2E), and thus the presence of even a single cell should be apparent in our pyramidal cell dataset (see Experimental Procedures). Given the absence of off-target gene cohorts, to the extent of our resolution we conclude that our transcriptomes reflect purely CA1 pyramidal cells.

As our RNA-seq data were internally reproducible and appeared to reflect pure CA1 populations, we next examined whether it was consistent with previous population-level gene-expression profiling of dorsal versus ventral CA1 (Dong et al., 2009; Leonardo et al., 2006).

First, we ensured consistency of our RNA-seq datasets with a prior microarray study that identified genes enriched at either the dorsal or ventral CA1 poles (Leonardo et al., 2006). We compared the extent of enrichment reported by microarray relative to enrichment obtained by RNA-seq (Figure S2F). Microarray enrichment was 4.3-fold on average, whereas RNA-seq found the same genes to be enriched on average 24.5-fold, demonstrating that RNA-seq accurately identified regionally enriched genes but with a markedly greater effect size.



(legend on next page)

Next, we examined whether our RNA-seq datasets could recapitulate and extend previous ISH work. Using marker genes previously identified as regionally enriched by ABA mining (Dong et al., 2009), we found our RNA-seq dataset gave 90% agreement with published results, with an average RNA-seq enrichment of 5.4-fold (Figure S2G; see Experimental Procedures). Second, using a strict definition to generate marker genes in our RNA-seq dataset (see Experimental Procedures), we identified 28 marker genes that we proceeded to examine in coronal ABA ISH images (Table S1). When inspecting the corresponding ISH profiles, we found consistent regional enrichment in ISH for 82% of genes (23/28; see Table S1). Similar results were obtained when validation was performed blind to the RNA-seq enrichment results by independent observers (75% and 71% of genes for two independent observers; see Experimental Procedures). Notably, 83% (19/23) of these validated hits were novel (i.e., not having been previously annotated as being regionally enriched in the ABA; Dong et al., 2009).

Finally, we ensured that our results were robust to the particular choice of computational method used to analyze the RNA-seq data. To this end, we crossvalidated our Cuffdiff-based pipeline by constructing a second, complementary pipeline comprised of count-based quantification with HTSeq (Anders et al., 2015) followed by differential expression computed by DESeq2 (Love et al., 2014; see Experimental Procedures). Applying this approach to analyze our aligned reads similarly returned hundreds of differentially expressed genes (Figure S2I). Of note, all genes identified in our Cuffdiff-based approach in the ensuing results were crossvalidated by the HTSeq-DESeq2 pipeline, in which every corresponding adjusted p-value returned by DESeq2 was found to be significant (geometric average $p_{adjusted}$ < 1e-28, range 3.4e-2 to 2.9e-120, n = 86; Figure S2K).

Thus, our findings were internally reproducible, validated by external datasets, and robust across different computational approaches. These findings are consistent with CA1 dorsal-ventral differences being much greater than previously thought.

Dorsal-Ventral Differences Incorporate Many Neuronally Relevant Genes

Previous transcriptional profiling of CA1 pyramidal cells revealed few region-specific genes with a clear neuronal significance, likely stemming from the modest overall number of region-specific genes. As our RNA-seq profiling revealed an abundance of transcriptional differences between the CA1 long-axis poles, we were next motivated to examine whether these differences included genes with neuronal relevance.

Regionally enriched genes were identified using strict criteria (>3-fold average dorsal-ventral difference, $FPKM_{MIN}$ = 10 in enriched population, and q < 0.05) and assigned to functional categories (see Experimental Procedures). For a host of neuronally relevant functions, we were able to identify multiple genes that were enriched between the two poles (Figure 2G; note this is a subset of the regionally enriched genes; see Table S2 for full list). Notably, these enriched categories were bidirectional (i.e., enriched genes found at both of the poles) and encompassed genes with both analog and binary enrichment (i.e., both graded and on-off differences between the dorsal-ventral poles).

Differences Are Present in Other Axes but Modest Relative to the Dorsal-Ventral Axis

We next examined transcriptional variability in the CA1 proximal-distal and superficial-deep axes. To our knowledge, no study has comprehensively examined transcriptional variability in the proximal-distal axis, and previous ABA mining efforts have identified few (<10) differentially expressed genes in the CA1 superficial-deep axis (Dong et al., 2009).

To investigate the proximal-distal axis (Figure S3A), we identified two transgenic mouse lines that labeled predominantly opposite ends of the proximal-distal axis: Calb1-EGFP (hereafter, Calb1) was found to label proximal cells, and Sim1-Cre × Ai9 (hereafter, Sim1) was found to label distal cells (Figure S3B). In both lines, labeling was restricted to the pyramidal cell layer but did not overlap with PV+ interneurons (Figure S1B) or neurons positive for SST, NPY, or GABA (data not shown). Calb1 labeling was seen to abut but not overlap CA2 (Figure S1C), whereas Sim1 labeling was seen to terminate immediately before subiculum (Figure S1D).

For the superficial-deep axis (Figure S3C), we took advantage of the fact that CA1 pyramidal cells are born over a period of days and laminate the hippocampus in a deep-to-superficial fashion. In utero viral injections (see Experimental Procedures) were employed at E14 and E17 to label deep and superficial CA1 pyramidal cells, respectively. Labeling was seen to be restricted to the pyramidal cell layer, did not overlap with PV+ interneurons (Figure S1E), and did not colocalize with other canonical interneuron markers (SST, NPY, GABA; data not shown).

In all cases, dorsal CA1 was microdissected from tissue and processed according to our RNA-seq workflow (Figure 1F).

Figure 2. Dorsal-Ventral Differences Are Wide-Ranging and Prominent

(A) Vipr2-labeled cells in CA1. Scale bars, overview represents 500 μm, expanded represents 100 μm.

(B) Vipr2-labeled cells were immunopositive for Wfs1, a marker for CA1 pyramidal cells. Scale bar, 100 μm.

(C) Vipr2-labeled cells were immunonegative for PV, a marker for pyramidal cell layer interneurons. Scale bar, 50 μm.

(D) Representative scatterplot of gene expression across two biological replicates from dorsal CA1. Each data point corresponds to the measured expression of one gene.

(E) Scatterplot comparing averaged dorsal and ventral transcriptomes. Green and magenta points are differentially expressed genes enriched in dorsal and ventral populations, respectively.

(F) Comparison of enriched genes versus fold change from RNA-seq (individual poles shown as green- and magenta-filled lines, sum shown as green-magenta dashed line) versus previous microarray (filled black point; n = 23 for >3-fold enrichment) and ISH (black dashed line; n = 26; note ISH does not yield quantitative information on fold change). Inset indicates the enriched gene overlap between previous ISH and microarray studies.

(G) A subset of genes enriched in a pole-specific fashion with neuronal relevance. Top and bottom three rows depict dorsal and ventral replicates, respectively. Range is normalized to the highest replicate FPKM on a gene-by-gene basis. See also Figures S1 and S2.

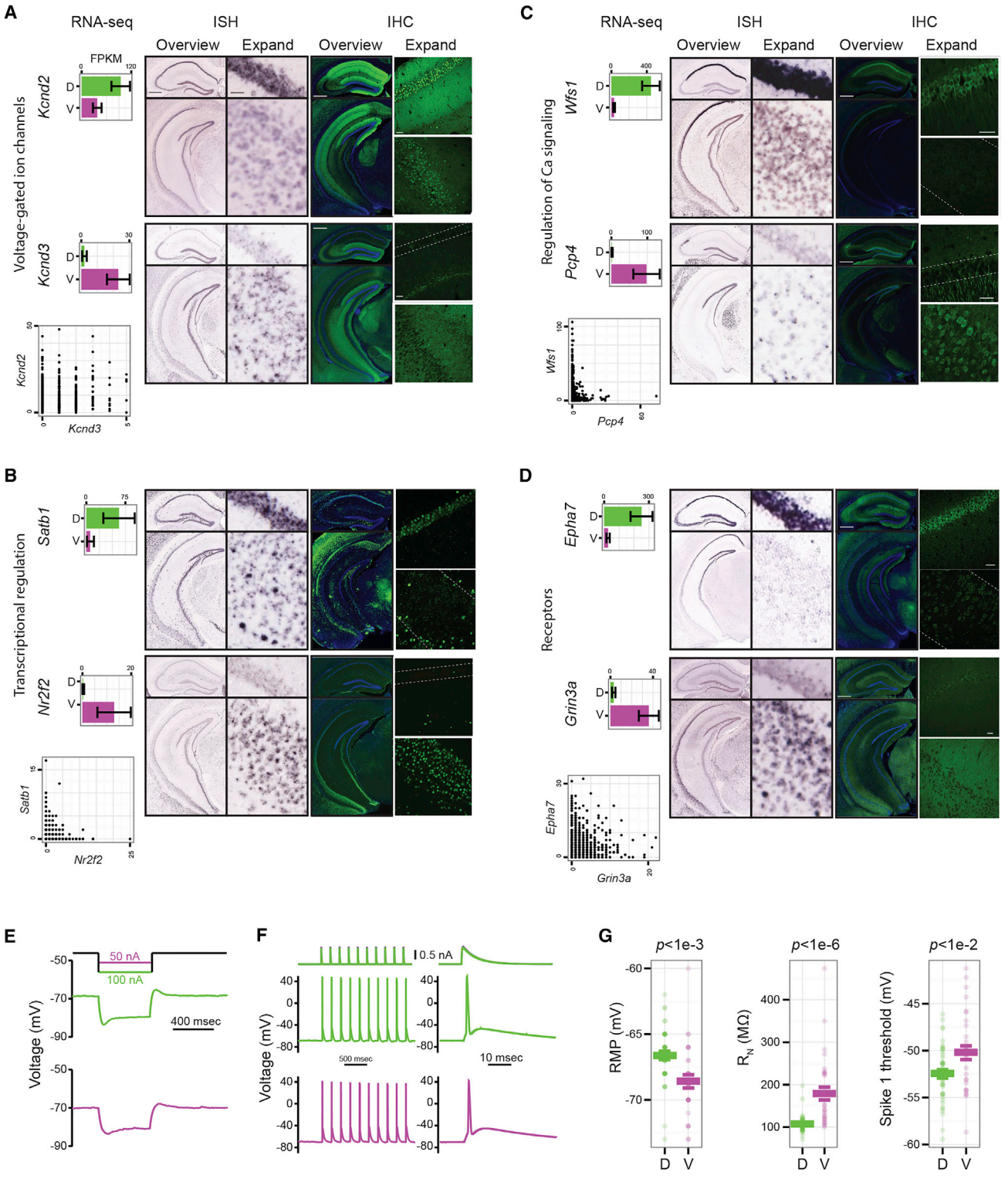


Figure 3. Protein Expression and Electrophysiology Vary along the CA1 Long Axis

(A) An example of two potassium-channel subunits identified by population-level RNA-seq (*Kcnd2* and *Kcnd3*) that were respectively enriched at the dorsal and ventral pole (column 1). For each subunit, ISH confirmed the expression pattern (columns 2 and 3), single-cell RNA-seq recapitulated the anticorrelated nature of gene expression (lower left), and IHC identified that the enrichment was present at the protein level (columns 4 and 5). Dashed line indicates pyramidal cell layer where not apparent. Error bars represent 95% confidence interval (CI).

(legend continued on next page)

High reproducibility was seen between biological replicates (0.98 ± 0.01 for within-proximal and within-distal comparison; 0.99 ± 0.01 for within-superficial and within-deep comparisons, mean ± SEM; Figures S2A and S2H). Here, however, the correlation across regions was almost as high (0.97 ± 0.01 for proximal-distal and superficial-deep comparisons; Figure S2A). This high correlation persisted despite the fact that several datasets appeared to have some contamination from interneurons (Figure S2D; see Experimental Procedures), suggesting that these high correlations are likely a lower bound of the true correlations between these CA1 subpopulations.

Using an approach to eliminate any confounds from interneurons (see Experimental Procedures), we searched for enriched genes across either the proximal-distal axis or the superficial-deep axis. Few genes were found to be enriched in these CA1 subpopulations relative to the dorsal-ventral dataset analyzed in the same manner (>2-fold different; result for proximal-distal was as follows: 33; result for superficial-deep was as follows: 71; and result for dorsal-ventral was as follows: 265; Figures S3E–S3G). Despite the relatively few genes found with this screen, our results both recapitulated previous findings and extended them to include a host of previously unannotated genes. For example, *Ndst4*, which was previously found to be enriched in the superficial-deep axis in ISH (Dong et al., 2009), was also identified as being enriched distally (Figure S3H). Other novel genes were identified as being differentially expressed in the proximal-distal axis (e.g., *Crtac1*; Figure S3I) and the superficial-deep axis (e.g., *Htr1a* and *Col11a1*; Figure S3J) and confirmed by the ABA.

Differentially Expressed Genes Produce Region-Specific Proteins

Having identified that the largest source of transcriptional variability in CA1 was present in the long axis, we investigated whether the observed transcriptional differences along this axis would also be apparent at the protein level. To this end, we identified hits from RNA-seq and crossvalidated our population-level transcriptional results histologically with single-cell RNA-seq, ABA in situ data, and immunohistochemistry (IHC) (see Experimental Procedures).

For a wide spectrum of proteins and associated ontologies, we were able to confirm clear labeling differences along the dorsal-ventral axis (Figures 3A–D; Figures S4A and S4B). Differences could be histologically validated at both the transcriptional and protein levels for voltage-gated channels (e.g., *Kcnd2*, *Kcnd3*, *Scn4b*, and *Kcnh7*, encoding Kv4.2, Kv4.3, Navβ4, and Kv11.3 subunits, respectively), transcription factors (*Satb1* and *Nr2f2*), regulators of calcium signaling (*Wfs1* and *Pcp4*), and receptors/auxiliary subunits (ephrin receptor *Epha7*, NMDA receptor subunit *Grin3a* encoding Nr3a, and the NMDAR/KAR auxiliary subunit *Neto1*).

Widespread Differences Are Present in CA1 Dorsal-Ventral Physiology

The pronounced differences in both the CA1 transcriptome and associated proteins suggest that physiological differences along the long axis should be present and readily apparent. To confirm this, we next performed patch-clamp experiments on CA1 pyramidal cells prepared from Vipr2 transverse slices, restricting our recording locations to dorsal and ventral locations along the long axis (i.e., ignoring intermediate CA1; see Experimental Procedures).

We drove CA1 pyramidal cells with two types of stimuli: hyperpolarizing DC current steps from rest were used to assess subthreshold properties (Figure 3E), and depolarizing EPSC-like waveforms were employed to assess suprathreshold properties (Figure 3F). From voltage responses to these stimuli, we extracted 16 electrophysiological parameters (Graves et al., 2012). In total, half of these parameters were significantly different between dorsal and ventral CA1 pyramidal cells ($p < 0.05$ for 8/16 properties, Table S4; see Experimental Procedures), including differences in both subthreshold properties (e.g., R_N, V_{rest}) and suprathreshold properties (e.g., spike threshold) (Figure 3G). Importantly, effect sizes were consistent across the dorsal-ventral axis whether considering either RFP-positive or RFP-negative cells solely (Figure S4C), indicating that these differences were a general feature of CA1 pyramidal cells.

Dorsal-Ventral Differences Do Not Emerge from Two Discrete Cell Types

Having verified that dorsal-ventral differences exist transcriptionally, proteomically, and physiologically, we next examined whether these transcriptional differences could be mapped onto discrete cell types. We first considered the most parsimonious explanation that could account for the transcriptional differences observed between the dorsal and ventral poles; namely, that CA1 is comprised of two distinct "dorsal" and "ventral" cell types. In this conceptual model (Figure 4A), at the extreme poles the relative cellular composition is biased toward pure populations. When moving from one pole to the opposite pole, the relative abundance of the two cell types changes.

This model predicts that the transcriptome sampled from intermediate CA1 cells should reflect a scaled average of the dorsal and ventral transcriptome weighted in proportion to the local relative abundance of the two cell types. For example, if the "dorsal" cell type made up 80% of cells midway along the long axis, then it would be expected that all dorsal marker-gene expression should be ~80% of that observed at the dorsal pole, and similarly all ventral marker-gene expression should be ~20% of the value observed at the ventral pole. Stated more formally, this conceptual model makes two predictions: first,

(B–D) As in (A), but with transcription factors (*Satb1* and *Nr2f2*) (B), calcium signaling regulators (*Wfs1* and *Pcp4*) (C), and receptors (*Epha7* and *Grin3a*) (D). Scale bars, overview represents 200 μm, expand represents 25 μm.

(E) Voltage responses to subthreshold DC currents for patched dorsal and ventral CA1 cells (100 and 50 nA, respectively).

(F) Representative responses to an EPSC-like stimulus train for dorsal and ventral cells (left: full 10 pulse train; right: expansion of single pulse).

(G) A subset of the electrophysiological parameters found to be significantly different between dorsal and ventral CA1 PCs. See also Figure S4 and Table S4. Error bars represent SEM.

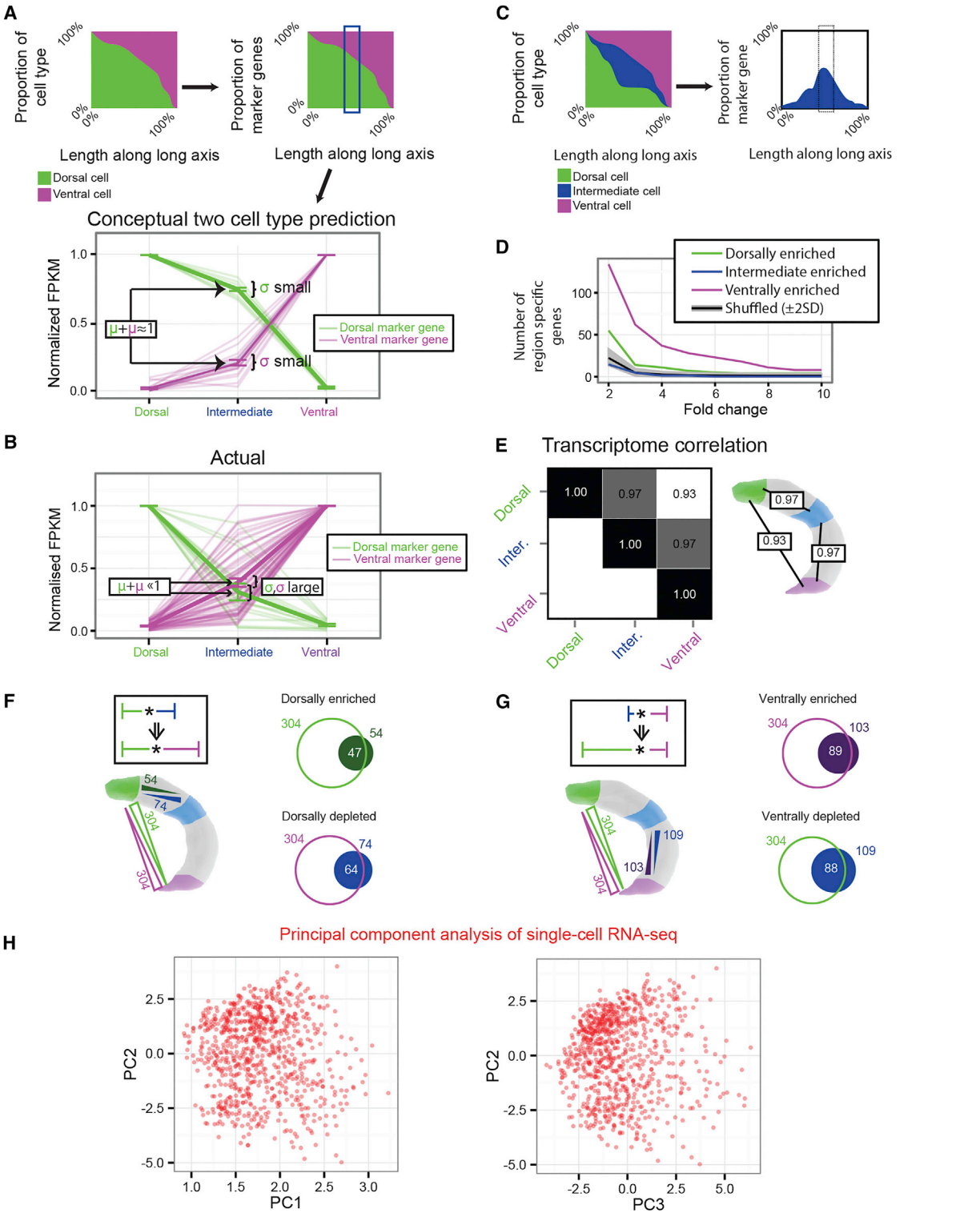


Figure 4. RNA-Seq Demonstrates Gene Expression Is Graded and Continuous across the Long Axis

(A) Conceptual model for two cell types. If a gradient of two cell types exists along the long axis (top left), then an identical expression gradient should be observed by all genes exclusive to one cell type (top right). This model predicts all marker genes of a given pole should exhibit a quantitatively similar profile along the long axis (bottom: μ, mean; σ, standard deviation). Individual marker genes for dorsal and ventral CA1 are represented by thin green and magenta lines, respectively, with the maximum average FPKM value normalized to one for each gene.

(legend continued on next page)

the sum of the mean values of dorsal and ventral marker genes should equal one, reflecting the ratio of the two cell types; second, variance about each mean value should be small, as all marker-gene expression from a given pole should be scaled by a similar amount (i.e., by the cell-type fraction).

To directly test these predictions, we returned to the Vipr2 transgenic animal used to profile dorsal and ventral CA1 transcriptomes. Exploiting the fact that RFP labeling terminated at the CA1-subiculum border (defined by either Wfs1 immunolabeling or a widening of the pyramidal cell layer; Figure S1F), we microdissected tissue intermediate along the long axis and purified fluorescent CA1 pyramidal cells for transcriptional profiling. The resulting transcriptome was reproducible (r = 0.99 ± 0.00 across biological replicates, mean ± SEM; Figure S2H) and was devoid of any resolvable contamination (Figure S2D). From this transcriptome, it was apparent that both predictions arising from the two-cell-types hypothesis failed: the mean value for both dorsal and ventral marker genes was less than 0.5, and a large spread about the mean was observed for both sets of marker genes (Figure 4B).

Intermediate Cells Exhibit Intermediate Transcriptomes

One explanation that could account for dorsal and ventral marker-gene mean values each being less than 0.5 would be the emergence of a third cell type intermediate along the long axis. The presence of an intermediate cell type would be reflected in the emergence of genes enriched intermediate along the long axis (Figure 4C). However, the intermediate CA1 transcriptome was effectively devoid of locally enriched genes: this region exhibited considerably fewer marker genes than at the dorsal and ventral poles and was at or below the number expected by chance (using replicate permutations, see Experimental Procedures) (Figure 4D).

Given that intermediate CA1 did not contain locally enriched genes, it is possible that CA1 pyramidal cells spatially intermediate along the long axis were also transcriptionally intermediate. We investigated this in two complementary ways, considering the gross relationships between whole transcriptomes as well as the finer relationships of specific genes. First, we compared correlation coefficients across the entire transcriptome. We found that the intermediate CA1 pyramidal cells were approximately equally correlated with the dorsal and ventral cells, and both of these correlations were markedly higher than the dorsal-ventral correlation (Figure 4E). Second, we compared relationships on a gene-by-gene basis. A transcriptionally monotonic CA1 would suggest that a differentially expressed gene found at a short spatial scale should be predictive of the same differentially expressed gene on a longer spatial scale. For example, if a gene is differentially expressed (e.g., enriched) dorsally relative to intermediate CA1, it should follow that the same gene is differentially expressed and enriched dorsally relative to ventral CA1. To examine this, we identified all genes that were significantly enriched or depleted intermediate relative to either pole (short spatial-scale relationship), and found that the corresponding long spatial-scale relationship was correctly predicted ~80% of the time (Figures 4F and 4G).

CA1 Single-Cell RNA-Seq Data Exhibit a Continuum

If indeed CA1 cells exist in a continuum rather than in discrete cell classes, it would be expected that single-cell transcriptomes should not exhibit clear subgroups. As our population-level RNA-seq data cannot resolve this, we next analyzed recent work from single-cell RNA-seq (n = 827 CA1 pyramidal cells) (Zeisel et al., 2015). Using principal components analysis (PCA) across all genes, we found that CA1 cells conformed to a wide continuum that did not exhibit obvious discrete subgroups (Figures 4H and S5C). As single-cell RNA-seq can exhibit substantial noise, we ensured that this finding did not simply reflect noisy expression dominating the overall classification. Notably, a continuum of CA1 cells was observed when including only genes corresponding to two previously identified putative cell types (Zeisel et al., 2015; Figures S5A, S5B, and S5D) as well as when including genes enriched at the two poles of the hippocampus (Figure S5E).

CA1 Transcriptional Identity Is Largely Established by Spatial-Expression Gradients

The previous RNA-seq results suggested that gene expression along the long axis was monotonic but lacked global cell-type rules that would produce consistent scaling across genes. Histologically, this predicts that marker-gene expression should be maximal at one pole and decay along the long axis, with expression profiles exhibiting pronounced gene-to-gene variability (Figure 5A). In agreement with this, ABA ISH profiles of dorsal and ventral marker genes showed marked variability in expression across multiple axes. At the dorsal pole (Figures 5B and S6A), where we obtained transcriptomes in all three axes, individual dorsally enriched genes could be seen to label CA1 cells ubiquitously (e.g., *Wfs1*), or to also be polarized in other axes in dorsal CA1 (e.g., *Enpp2*, enriched proximal and superficial, or *Car2*, enriched distal and deep). Variability in decay was also observable across the long axis (Figure S6A). For ventral CA1 (Figures 5C and S6B), marked expression variability was also observed.

(B) Transcriptional profile of dorsal and ventral marker genes across the long axis. Error bars represent SD.

(C) Conceptual model including a third, intermediate cell type. If an intermediate cell type exists (left), it should give rise to genes enriched intermediate relative to the two poles of CA1 (right).

(D) The number of region-specific genes as a function of fold change for dorsal, intermediate, and ventral locations, as well as replicate-shuffled data (see Experimental Procedures).

(E) Pearson correlation coefficients between averaged dorsal, intermediate, and ventral replicates.

(F and G) The number and directionality of differentially expressed genes for all pairwise comparisons between dorsal, intermediate, and ventral transcriptomes. Statistically significant pairwise comparisons between the intermediate location and either the dorsal (F) or ventral (G) pole were seen to be highly predictive of the same comparison across the full length of the hippocampus, independent of both the specific pole and the directionality of enrichment.

(H) Principal component analysis of 827 individual CA1 transcriptomes from single-cell RNA-seq. See also Figures S1, S2, and S5.

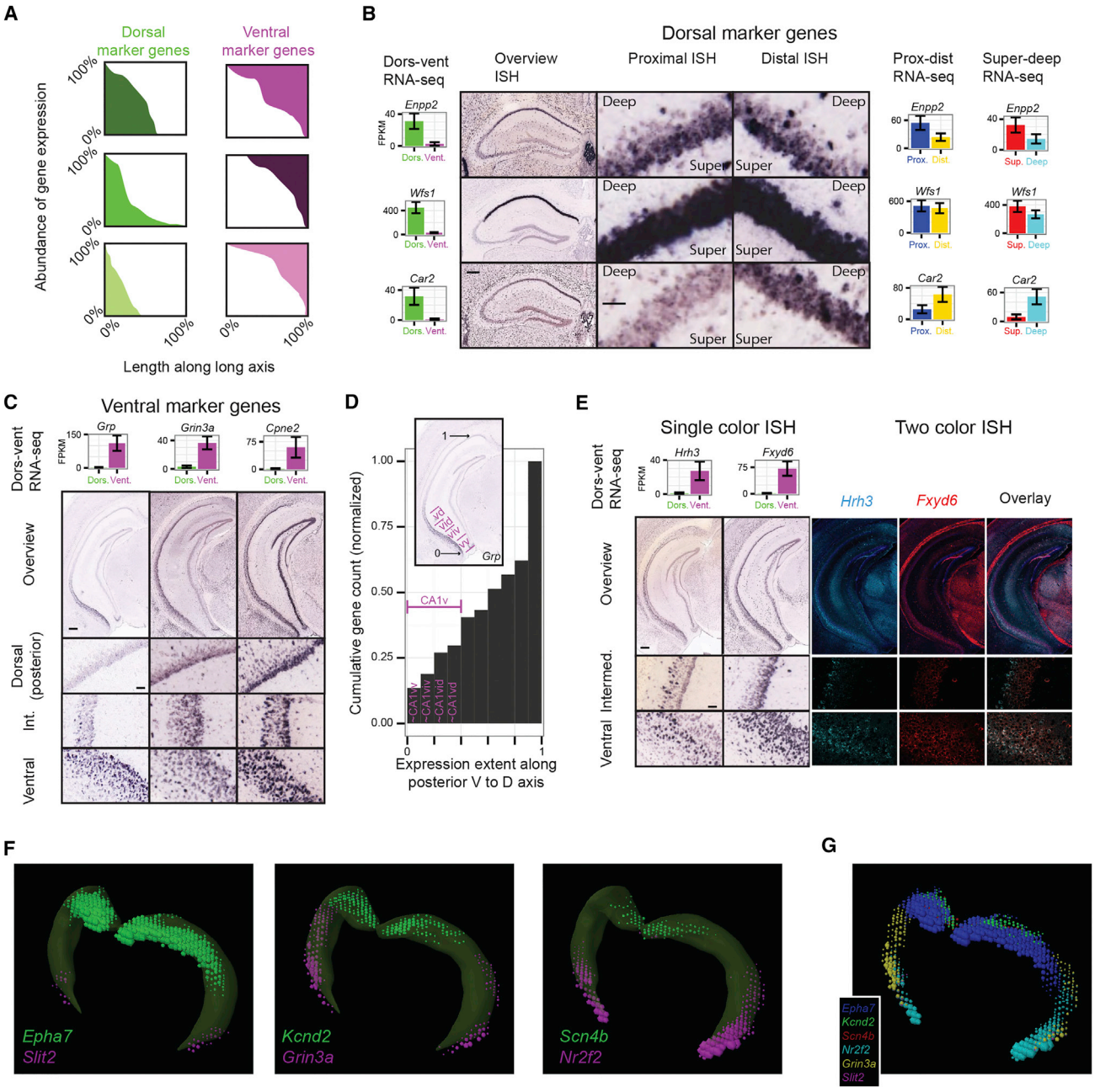


Figure 5. CA1 Transcriptional Identity Emerges from a Heterogeneous Collection of Gene-Expression Gradients

(A) Schematic of the spatial profile of marker genes consistent with RNA-seq. Marker-gene expression is maximal at one pole, decays along the long axis, and exhibits variability in spatial profile from gene to gene.

(B) Dorsal marker genes *Enpp2*, *Wfs1*, and *Car2* identified by RNA-seq (left) showed different patterns of enrichment in the other axes of CA1, apparent in both ISH (middle) and RNA-seq (right). Error bars represent 95% CI.

(C) Ventral marker genes *Grp*, *Grin3a*, and *Cpne2* identified by RNA-seq (top) showed different patterns of enrichment, as seen in ISH (bottom). Error bars represent 95% CI.

(D) Cumulative distribution of the dorsal-ventral extent of ventral marker-gene expression in posterior CA1 (n = 38 genes total).

(E) Ventral marker genes *Hrh3* and *Fxyd6* (top left) appeared to label district groups of cells in intermediate CA1 (middle) but overlapped at the ventral pole of CA1 (bottom). Scale bars for ISH, overview represents 250 μm, expanded represents 25 μm. Error bars represent 95% CI.

(F) Gene expression from the ABA Brain Explorer for representative novel regionally restricted genes. A range of spatial expressions was found for dorsal (*Epha7*, *Kcnd2*, *Scn4b*) and ventral (*Slit2*, *Grin3a*, *Nr2f2*) genes. CA1 is shown in green. Genes depicted were involved in neurotransmission (*Grin3a*), transcriptional regulation (*Nr2f2*), intrinsic excitability (*Kcnd2*, *Scn4b*), and axon guidance (*Epha7*, *Slit2*).

(G) Superposition of these representative genes produced a spatially heterogeneous CA1 across the long axis. See also Figure S6.

We found genes exhibiting previously described sharp boundaries (Dong et al., 2009; e.g., *Grp*; Figure 5C), but also identified novel genes with a more gradual taper across the long axis, either appearing restricted to specific laminae (e.g., *Cpne2*) or graded across multiple laminae (e.g., *Grin3a*; Figures 5C and S6B). Ventral CA1 has been previously partitioned into four discrete subdomains (VD, VID, VIV, VV; Figure 5D inset; Dong et al., 2009). To compare to this work and summarize the spatial extent of ventral marker-gene expression along the long axis, we calculated the normalized length of expression for each marker gene along the long axis in posterior sections (see Experimental Procedures). Notably, few genes (<30%) terminated within the previously described four subdomains (Figure 5D); indeed, ventral markers exhibited a wide range of termination points, with more than a third extending across the full dorsal-ventral axis in posterior CA1 (i.e., approximately midway across the long axis).

Previous characterizations of hippocampal pyramidal cell gene expression have identified reciprocal boundaries between marker genes (Dong et al., 2009; Thompson et al., 2008). Examining regionally restricted transcripts suggests that this is not an organizational principle in CA1, as the vast majority of marker genes show clear overlap in histological profiles at their respective poles (Figure S6). Notably, although some genes could be seen to exhibit apparent reciprocal boundaries at intermediate locations along the long axis (e.g., the ventral marker genes *Hrh3* and *Fxyd6*, which appeared to be polarized to opposite laminae in intermediate CA1; Figure 5E), two-color ISH revealed that the apparent reciprocal boundaries broke down at more ventral CA1 locations.

To conclude, ABA histology confirmed our RNA-seq prediction (Figure 5A) that gene expression along the long axis was monotonic and exhibited marked gene-to-gene variability and thus did not conform to discrete transcriptional cell types enriched at different positions along the dorsal-ventral axis. Combining these findings, CA1 pyramidal cells exhibit pronounced transformations in their transcriptional profile across space, which encompasses regionally restricted genes involved in a host of neuronally relevant ontologies (Figures 5F and 5G).

Variability within CA1 Is Comparable to Variability across Pyramidal Cell Types

From the pronounced variability seen between the two poles of CA1, we were next motivated to compare the variability found within CA1 at these extremes to the variability between CA1 and CA3 pyramidal cells from the same region (Figure 6A). As a host of morphological, physiological, anatomical, and functional properties vary between CA3 and CA1 pyramidal cells, quantifying the transcriptional difference between CA3 and CA1 pyramidal cells provides a quantitative framework for evaluating the extent of transcriptional variability within CA1.

To label CA3 pyramidal cells, we identified a Cre line (Mpp3-Cre) that, when crossed to a Cre-dependent reporter (see Experimental Procedures), exhibited expression in the CA3 pyramidal cell layer (Figure 6B). Labeling did not overlap with PV+ cells (Figure S1G) or with SST+, NPY+, and GABA+ cells (data not shown). From dorsal and ventral CA3 transcriptomes, we found that the resulting RNA-seq datasets were highly reproducible (r = 0.99 ± 0.01, mean ± SEM; Figure S2H) and devoid of contamination from other cell types (Figure S2D). Comparisons were then made between CA3 dorsal and ventral RNA-seq datasets and the corresponding dorsal and ventral CA1 datasets (Figure 6C).

Interestingly, by multiple metrics, the transcriptional variability within the CA1 pyramidal cell population (“within class”) approached or matched the magnitude of variability between CA3 and CA1 cell types (“across class”). First, the correlation was similar in magnitude for the within-class versus across-class comparison (Figure 6D; Pearson’s correlation coefficient; “within” results were as follows: 0.926 ± 0.007; “across” results were as follows: 0.933 ± 0.006, p = 0.51, mean ± SEM; see Experimental Procedures). This relationship was robust: it was recapitulated for nonparametric quantification of correlation (Spearman correlation coefficient; “within” results were as follows: 0.903 ± 0.002, “across” were as follows: 0.907 ± 0.001, p = 0.24, mean ± SEM) as well as correlation of log-transformed FPKM values (Pearson correlation coefficient; “within” results were as follows: 0.960 ± 0.003; “across” results were as follows: 0.959 ± 0.001, p = 0.81, mean ± SEM). Qualitatively identical results were found for correlation of count-based quantification (Figure S7A). Second, the number of regionally enriched genes was similar across and within class, such that the number of within-class differences was ~60%–70% of that seen across class (Figures 6E and S7B). Third, using the Allen Gene Expression Atlas (AGEA) (Ng et al., 2009) to compare whole-genome ISH correlations, we found that lowest voxel correlation was the within-class CA1 comparison, rather than across classes at either pole (Figure S7D).

Next, we investigated enriched genes in terms of functional consequences (see Experimental Procedures). Analysis of gene ontologies, pathways, and potential upstream regulators all revealed many significant features for the within-CA1 comparison, many of which exceeded the significance for across-class comparisons (Figures S7F–S7H). These results attest to the differences within CA1 not only being numerically similar to those across classes but also mapping onto features with potential functional relevance.

Finally, we investigated whether the similar across- versus within-class differences would be found when comparing dentate gyrus granule cells to CA1 and CA3 pyramidal cells. Obtaining dorsal and ventral granule-cell RNA-seq datasets and using hierarchical clustering to identify distances between datasets (see Experimental Procedures), we found that the first bifurcation of the clustering corresponded to pyramidal versus granule cells (Figure S7E), indicating that across- versus within-class differences were similar only within different classes of pyramidal cells. Notably, this analysis using agglomerative clustering again illustrated that the distance between dorsal and ventral CA1 was almost identical to the distance between CA1 and CA3.

Projection-Specific Transcriptomes Exhibit Further CA1 PC Variability Emerging from Gene-Expression Gradients

The previous anatomical dorsal and ventral regions used to characterize the transcriptional differences may miss finer details corresponding to functional subpopulations of CA1 pyramidal

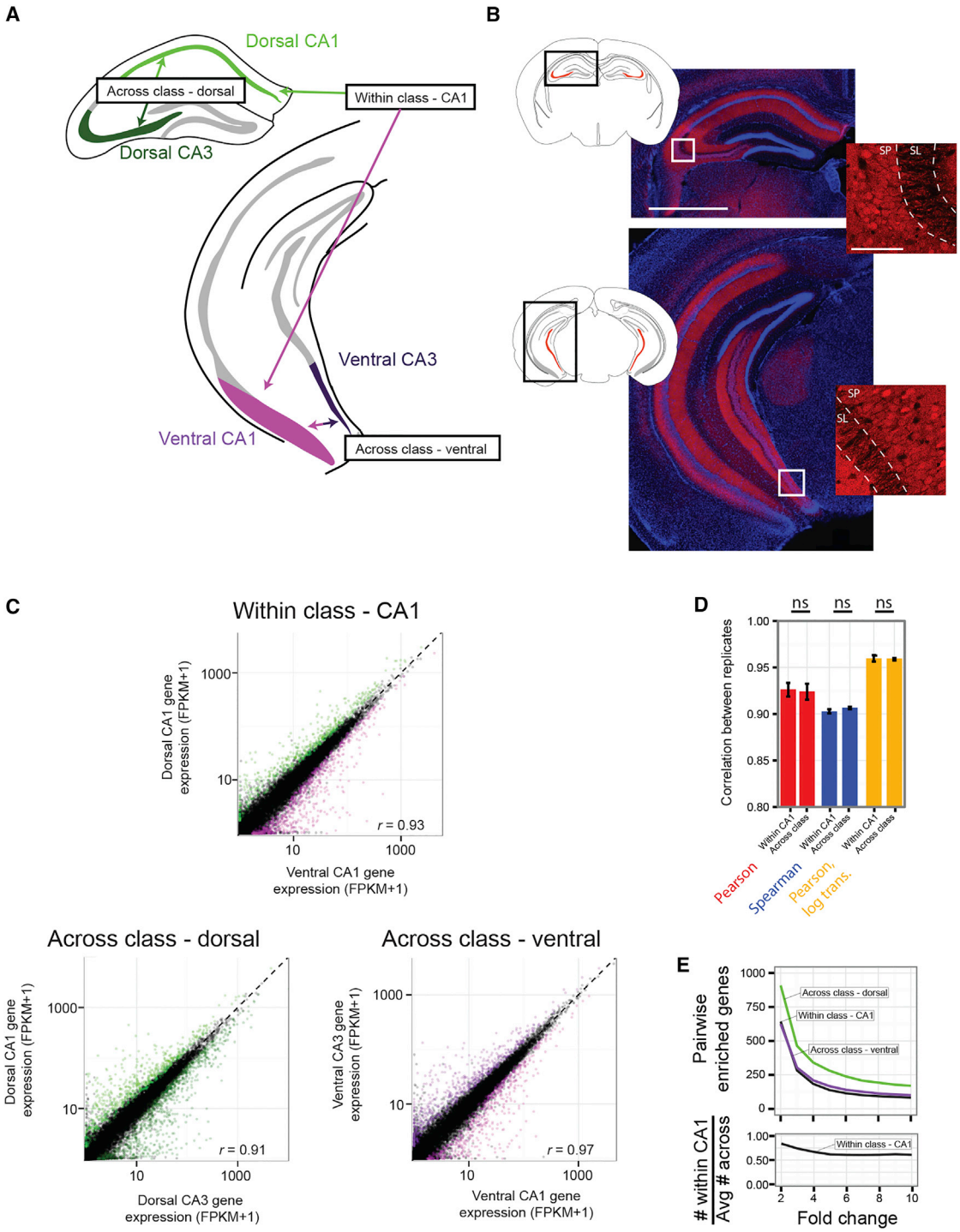


Figure 6. Variability within CA1 Is Comparable to Variability across Different Canonical Cell Populations

(A) Schematic illustrating within-class and across-class comparisons.

(B) Mpp3 labeling near the dorsal and ventral poles of CA3. Scale bars, overview represents 1 mm, expanded represents 100 μm.

(C) Scatterplots comparing averaged transcriptomes for dorsal CA1 versus dorsal CA3 (lower left), dorsal CA1 versus ventral CA1 (top), and ventral CA1 versus ventral CA3 (lower right). Colored points indicate differentially expressed genes.

(D) Correlation of replicates, within CA1 and across classes. Error bars represent SEM.

(E) Top: the number of genes found enriched in pairwise comparisons as a function of fold change for within- and across-class comparisons. Bottom: the number of within-class enriched genes for CA1, normalized by the average number of genes enriched in across-class comparisons. See also Figures S1, S2, and S7.

cells within these regions. To examine this possibility, we labeled projection-defined subpopulations of CA1 pyramidal cells via retrograde bead injections (see Experimental Procedures), microdissected out the same region as the previous datasets, and compared the resulting projection- and location-based transcriptome to the corresponding location-based transcriptome.

We examined projections from ventral CA1, which has multiple extrahippocampal projections and exhibited marked variability in gene expression (Figures 4, 5, and S6). We first considered CA1 neurons that projected to the amygdala (Figure 7A). Gene expression in amygdala-projecting neurons was found to vary from the more general location transcriptome (Figures 7A and 7E). Encouraged by this finding, we next examined neurons that projected to the nucleus accumbens (NAc) (Figure 7B); here again, we identified a large degree of variability between NAc-projecting neurons of ventral CA1 and the ventral transcriptome (Figure 7E). Importantly, comparisons between amygdala- and NAc-projecting neurons also exhibited a preponderance of marker genes (Figure 7E), indicating that the two projection transcriptomes were not redundant.

We examined ISH profiles for genes that were either enriched or depleted in projection transcriptomes relative to the location transcriptome. We found that these genes were again expressed in gradients across the long axis with a variety of profiles (Figure 7C), varying in both the extent across the dorsal-ventral axis and the laminae of the superficial-deep axis. This recapitulated the same organizational scheme when considering geography alone (Figures 4 and 5), showing that functional cohorts of ventral CA1 cells with distinct projections obeyed the same rules as those identified based on ventral position alone.

For comparison, we examined the transcriptome of dorsal CA1 pyramidal cells that projected to the dorsal postsubiculum, found to be a dense group of cells that were observed predominantly in the proximal CA1 (Figure 7D). Differentially expressed genes were identified between postsubiculum-projecting neurons and dorsal CA1 neurons; however, the differences were largely analog in nature (Figure 7E), reminiscent of the small number of marker genes present in the proximal-distal axis of dorsal CA1 (Figure S3).

Notably, when comparing transcriptomes at opposite ends of the long axis, the differences between projection-defined neurons exceeded those defined by location alone (Figure 7F). Visualizing all datasets simultaneously (Figure 7G; see Experimental Procedures), projection-defined populations separated more than solely location-defined populations, further underscoring the profound variability of CA1 gene expression.

DISCUSSION

Here, we used RNA-seq to transcriptionally profile subpopulations of CA1 pyramidal cells. RNA-seq is a relatively new but promising technology for neuroscience (Shin et al., 2014) that has already provided insight into heterogeneity in a variety of principal neuron classes (Belgard et al., 2011; Pollen et al., 2014; Usoskin et al., 2015; Zeisel et al., 2015; Zhang et al., 2014). Our CA1 pyramidal cell RNA-seq revealed a wealth of subpopulation-enriched genes that have gone undetected by previous techniques. Through identifying, quantifying, and contextualizing this heterogeneity, we demonstrated that CA1 pyramidal cell identity emerges from a variety of gene-expression gradients, producing transcriptional variability between the two poles of CA1 that was quantitatively similar to that observed when comparing CA1 and CA3. This work illustrates a surprising amount of variability within a canonical neuronal population and explicitly demonstrates that, at the level of gene expression, graded changes in neuronal identity can be an important organizational principle of neuronal populations.

Continuous versus Discrete Cell Populations in the Hippocampus

Previous work has illustrated that pyramidal cells in the major CA regions can exhibit multiple subdomains with discrete, reciprocal boundaries (Dong et al., 2009; Thompson et al., 2008). Here, however, we primarily found marker genes that obeyed fundamentally different spatial patterns when considering both the profiles of individual genes and the interrelationships of boundaries across expression profiles.

Expression of individual genes was found to be primarily monotonic along the long axis; that is, maximal expression occurred at one pole and decayed toward the opposite pole. The spatial profile of this decay varied on a gene-by-gene basis when considered either quantitatively with RNA-seq or histologically with ISH. This gene-to-gene variability contrasts with CA3, where cohorts of regionally enriched genes recapitulate similar spatial patterns (Thompson et al., 2008). Additionally, no marker genes were found to exhibit peak expression at an intermediate point along the long axis. This lack of intermediate marker genes in CA1 is consistent with previous work identifying intermediate CA1 largely by the decreased expression of marker genes (Dong et al., 2009) but again differing from CA3, which has multiple subdomains along the long axis that can be viewed directly by local marker-gene expression (Thompson et al., 2008).

The boundaries between marker genes also obey a fundamentally different relationship between CA1 and CA3 pyramidal cells. The vast majority of marker genes found here exhibited near-ubiquitous expression at their respective poles, guaranteeing overlap between marker genes of the same pole; indeed, even transcripts that seemed to be nonoverlapping at intermediate locations exhibited a high degree of overlap near their respective poles (Figure 5E). The lack of clear, abutting boundaries was also found when considering boundaries for marker genes of opposite poles: very few marker genes were seen to be expressed at intermediate locations, again guaranteeing that no reciprocal boundaries existed for dorsal versus ventral marker genes. In contrast, regionally enriched marker genes in CA3 were frequently found to have reciprocal boundaries (Thompson et al., 2008).

In principle, the discrepancy in organizational schemes found between CA1 and CA3 could arise due to differences in methodology (RNA-seq versus ISH). We controlled for this by applying the same analysis to examine dorsal-ventral differences in CA3 as was used for CA1 (see Experimental Procedures). Importantly, this recapitulated previously identified regionally enriched genes and corresponding boundaries (Figure S7I), and only ~15% of genes recovered obeyed a gradient pattern in CA3. Therefore, these CA3-CA1 differences do not arise

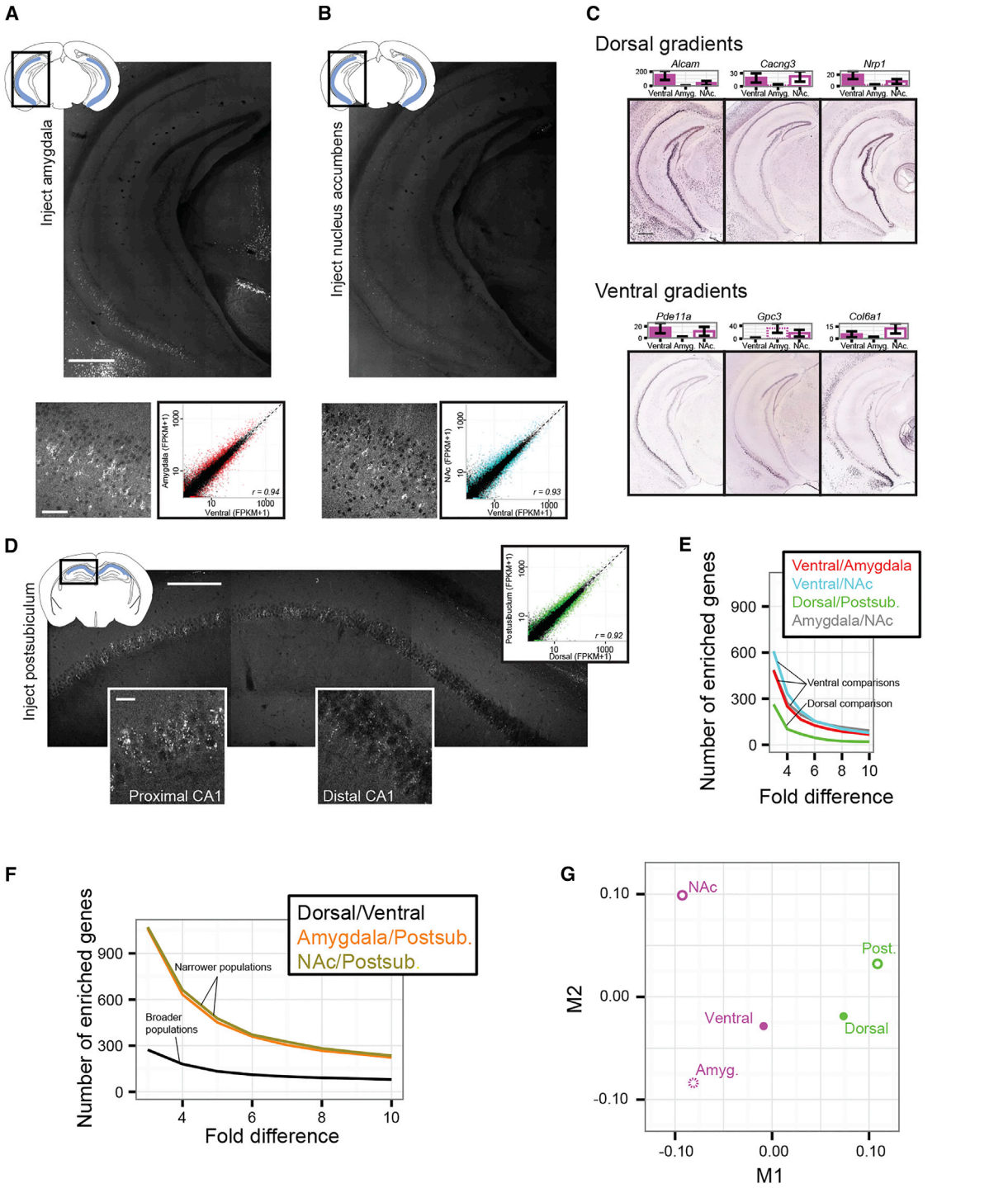


Figure 7. Projection-Specific Transcriptomes Exhibit Greater Variability Than Location-Defined Transcriptomes

(A) Top: retrograde bead labeling in ventral CA1 following amygdala injection. Scale bar, 500 μm. Bottom left: expansion of label in ventral CA1. Scale bar, 100 μm. Bottom right: scatterplot of ventral versus amygdala-specific transcriptomes.

(B) As in (A), but following NAc injection.

(C) ISH for projection-enriched or projection-depleted genes. Scale bar, 200 μm. Error bars represent 95% CI.

(D) As in (A), but following postsubiculum injection and comparing to dorsal transcriptome. Scale bar, overview represents 200 μm, expanded represents 20 μm.

(E) Enriched genes versus fold change for projection-specific populations relative to corresponding location-defined populations.

(F) As in (D), but comparing transcriptomes at opposite poles.

(G) Relationships between projection-defined and geographically defined transcriptomes visualized by multidimensional scaling.

from methodological differences but rather bona fide differences in organizational principles.

Ultimately, the large variability across marker-gene expression, combined with the lack of reciprocal boundaries for this expression, suggests that there is a lack of global rules that determine CA1 transcriptional identity along the long axis, and thus CA1 does not adhere to a small number of discrete dorsal-ventral cell types. This gene expression is not consistent with a strict "dichotomy" interpretation of hippocampal processing or with a more generalized tripartite model of three gross dorsal, intermediate, and ventral compartments (Fanselow and Dong, 2010; Strange et al., 2014). Here, we found that CA1 transcriptional identity cannot be easily distilled into distinct compartments along the long axis; rather, the emergent scheme is one of multiple, relatively independent gene-expression gradients that superpose to produce a highly heterogeneous and continuously variable population of pyramidal cells (Figures 5F and 5G).

Anatomical Consequences of CA1 Gene-Expression Gradients

In identifying that CA1 transcriptional identity along the long axis emerges from gene-expression gradients, our work qualitatively agrees with the graded spatial topography of CA1 inputs and outputs. Consistent with this, many novel differentially expressed genes were found to be involved in the establishment and maintenance of neuronal networks (Figure 2G), and for several marker (on-off) genes we were able to directly demonstrate clear expression gradients across the long axis (e.g., *Epha7*, *Slit2*, *Ntng1*; Figures 5F, 5G, and S6B). These results represent a departure from the previous model of graded afferent and efferent hippocampal connections being superimposed on discrete subpopulations of cells (Strange et al., 2014).

Because CA1 cell transcriptional identity emerges from spatial gradients, whereas CA3 identity emerges from discrete microdomains (Thompson et al., 2008), the spatial arrangements of CA3 and CA1 cells might initially seem to be mismatched. However, it is well known that individual CA3 cells can project broadly across the dorsal-ventral axis (Amaral and Witter, 1989; Ishizuka et al., 1990; Ropireddy and Ascoli, 2011). Therefore, it may be that individual genetically defined CA3 microdomains disperse information in a graded fashion, such that a given CA1 pyramidal cell receives and combines information from multiple CA3 subtypes.

Pronounced Transcriptional Variability within the CA1 Pyramidal Cell Population

Our findings of within-class variability being comparable to across-class variability at the two poles of CA1 suggest that the identity of CA1 pyramidal cells changes dramatically along the dorsal-ventral axis. Moreover, even transcriptomes from the same pole but associated with different projections can exhibit pronounced differences (Figure 7). This has important ramifications for how processing in the trisynaptic loop is viewed. Although this loop is repeated along the dorsal-ventral length of the hippocampus, our work suggests that in addition to being a stereotyped circuit that subserves a general function, CA1 cell identity and function evolve along the long axis as well as locally within a given location in the dorsal-ventral axis. Consequently, these gene-expression differences might suggest that the trisynaptic loop performs specific region- and projection-dependent variants of its core computation. Given the concomitant change in location and projection targets of dorsal and ventral CA1, this disparity of CA1 pyramidal cell identity at the two poles may be an important feature underlying the functional segregation of the dorsal and ventral regions of the hippocampus (Fanselow and Dong, 2010; Moser and Moser, 1998), allowing the two subregions to extract and impart foundationally different types of information.

In principal, similar higher-order features can emerge from disparate gene-expression profiles (Goaillard et al., 2009; Prinz et al., 2004). However, a collection of work demonstrates graded CA1 electrophysiological (Malik et al., 2015), morphological (Malik et al., 2015), and connectivity (Amaral and Witter, 1989) properties, suggesting the graded transcriptional scheme found here grossly maps onto similar higher-order organizational features. As such, understanding the precise interrelationships between these levels of granularity will yield key insight into hippocampal processing. Our work here begins to examine interrelationship between gene expression and projection target (Figure 7), an important step toward creating a coherent, unified description of covariance of CA1 gene expression, intrinsic properties, and network connectivity.

Continuously Variable Cell Properties in the Brain

To what extent do the gradual, continuous changes found across the spatial extent of CA1 generalize to other regions of the brain? Within the hippocampus, previous transcriptional analysis has shown that CA3 is best described by discrete cell types with sharp boundaries; however, some individual gene-expression profiles were identified with more gradual changes that did not conform to this discrete framework (Thompson et al., 2008). Similarly, in the dentate gyrus, graded gene-expression differences have also been reported along the dorsal-ventral axis (Fanselow and Dong, 2010; Lein et al., 2007). These findings suggest that graded gene-expression differences are present, to variable extents, in multiple principal cell types in the hippocampus. Recent work has also illustrated that memory representations in humans vary in a graded fashion across the hippocampal long axis (Collin et al., 2015).

In the neocortex, the molecular specification cues present during development exhibit smooth, graded profiles. These profiles span multiple cortical areas, potentially suggesting that graded cell properties could be found across different neocortical regions (Sansom and Livesey, 2009). Consistent with this, studies in the primate (Bernard et al., 2012) and human (Hawrylycz et al., 2012) have shown that spatial proximity correlates with gene-expression similarity for regions across the neocortex. Moreover, continuously variable immunohistochemical (Kondo et al., 1994; Xu et al., 2003), morphological (Elston, 2002), and anatomical (Freese and Amaral, 2005) neocortical properties have been found, and areal boundaries between higher-order association areas can be graded rather than sharp (Komatsu et al., 2005). These neocortical properties suggest that continual variation may be a general feature of repeated, spatially extended circuit motifs in the brain.

We conclude by noting that although gradients of cellular properties are found throughout many regions of the brain, our work here is the first to explicitly and quantitatively show that this can be an organizational principle in a supposedly stereotyped, mature neuronal population. The high degree of variability in gene expression found here correlates with differences in protein products (Figure 3), electrophysiological properties (Figure 3), and network connectivity (Figure 7), suggesting that these transcriptional differences underlie a host of higher-level variability in the CA1 pyramidal cell population. The fact that the differential expression of a single gene can have significant consequences on cellular function, combined with our findings of hundreds of differentially enriched genes within CA1, suggests that these functional differences identified here are likely a small subset of the total functional differences in a markedly variable CA1 population.

EXPERIMENTAL PROCEDURES

Mice were housed on a 12 hr light/dark cycle with ad libitum food and water access. Experimental procedures were approved by the Institutional Animal Care and Use Committee at the Janelia Research Campus.

Manual Sorting, Library Preparation, and Sequencing

A combination of transgenic mice, viral injections in utero, and retrograde bead injections were used to fluorescently label neurons in the hippocampus. Manual sorting was used to purify fluorescent neurons from microdissected slices according to previous methods (Hempel et al., 2007). Total RNA was isolated from each sample, ERCC spike-in controls were added, and cDNA libraries were amplified from this material. Libraries were sequenced on a HiSeq 2500 (Illumina, San Diego, CA), and single-end 100 bp reads were generated.

RNA-Seq Read Alignment, Quantification, Differential Expression, and Analysis

RNA-seq reads for each library were aligned with TopHat v2.0.6 (http://ccb.jhu.edu/software/tophat/index.shtml) (Trapnell et al., 2009), with quantification and differential expression performed by Cuffdiff v2.1.1 (http://cole-trapnell-lab.github.io/cufflinks/) (Trapnell et al., 2010). The processed data were analyzed in the R environment using a combination of cummeRbund v3.0 (http://compbio.mit.edu/cummeRbund/) (Goff et al., 2013) and custom scripts. Analysis conventions were as follows: FDR < 0.05 was used for differential expression; a gene was considered X-fold enriched in a given region, relative to other regions, when the mean FPKM value was at least X-fold greater for all corresponding pairwise comparisons (e.g., for gene A to be X-fold enriched dorsally relative to intermediate and ventral CA1, $FPKM_{A,dorsal} > X \cdot FPKM_{A,intermediate}$ and $FPKM_{A,dorsal} > X \cdot FPKM_{A,ventral}$); Pearson correlation coefficients were used to compare across datasets (except Figures 6D and S7A, where correlation coefficients were as specified); and error bars for FPKM values were taken from Cuffdiff's 95% CI model. Gene expression was required to obey FPKM > 10 in at least one population to be included in differential expression or enriched population analyses. Results from the TuxedoSuite pipeline were cross-validated by using count-based quantification (HTSeq) (Anders et al., 2015) and differential expression (DESeq2) (Love et al., 2014).

Histology

Mice were deeply anesthetized with isoflurane and perfused with phosphate-buffered saline (PBS) followed by 4% paraformaldehyde (PFA) in PBS. Brains were dissected, postfixed in 4% PFA, and sectioned. Immunohistochemistry was performed according to previously established protocols, with antigen retrieval used for Nr3a, Satb1, and Epha7; in situ hybridization was performed with antigen retrieval, pretreatment, hybridization, amplification, and detection according to User Manual for Fixed Frozen Tissue (ACD).

Electrophysiology

Slices and recordings from 3- to 4-week-old Vipr2-Cre(KE2) × Ai9 mice of either sex were made according to previous approaches (Graves et al., 2012). Analysis was performed in Igor Pro and R with custom scripts with previous conventions for quantifying electrophysiological parameters (Graves et al., 2012). Cells that exhibited a burst-firing phenotype (n = 8/85), corresponding to a small population of CA1 cells that are a different class than regular spiking cells (Graves et al., 2012), were excluded from analysis.

Analysis of ABA ISH Database and Brain Explorer

When validating the results of RNA-seq with the ABA, we examined coronal ABA images at A-P locations previously used to examine CA1 gene expression (Dong et al., 2009), which also corresponded to the microdissected areas used for RNA-seq (Figure 1). For visualizing expression profiles in Brain Explorer, genes were chosen that were selectively expressed in pyramidal cells within CA1 (e.g., excluding genes that were present in interneurons and glial cells, which confounded visualization). Intensity ranged between 123 and >260, and density ranged between 0.0248 and >0.1.

Fluorescence Imaging

Images of large regions of tissue (i.e., complete dorsal and ventral CA1) were acquired on a whole-slide digital scanner (Pannoramic 250 Flash, Perkin Elmer, Waltham, PA) using a 20× objective. Cellular resolution images were acquired with a confocal microscope (LSM 710 Carl Zeiss Microscopy, Jena, Germany) using a 20× objective. Some images were postprocessed in Fiji, including pseudocoloring to facilitate visual comparisons across channels and/or to adhere to the coloring conventions of the dorsal-ventral, proximal-distal, and superficial-deep axes.

Publicly Available Resources

Analysis scripts (http://dx.doi.org/10.6084/m9.figshare.2013267), electrophysiological data (http://dx.doi.org/10.6084/m9.figshare.2009718), and high-resolution histological images (http://dx.doi.org/10.6084/m9.figshare.2016057) are available through Figshare.

ACCESSION NUMBERS

Raw and processed RNA-seq datasets were deposited in the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus under GEO: GSE67403.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, seven figures, and four tables and can be found with this article online at http://dx.doi.org/10.1016/j.neuron.2015.12.013.

AUTHOR CONTRIBUTIONS

M.S.C., J.L.B., K.S., and N.S. designed the experiments; L.W. and K.S. collected the RNA-seq data; M.S.C. analyzed the RNA-seq data; M.S.C and J.L.B. collected and analyzed the electrophysiological data; and M.S.C. and B.C.S. collected and analyzed the histological data. M.S.C. and N.S. wrote the manuscript with input from the other authors.

ACKNOWLEDGMENTS

The authors would like to thank Brett Mensh, Erik Bloss, William Kath, and other members of the Spruston laboratory for helpful discussions; Jim Trimmer for the gift of the Nr3a antibody; Masahiko Watanabe for the gift of the Neto1 antibody; Deanna Otstot, Kendra Morris, Cassie Ruiz, Jared Rouchard, and Salvatore DiLisio for help with animal care and injections; and Andrew Lemire and Serge Picard for helping to coordinate and generate RNA-seq data. We are also grateful to the Allen Institute for Brain Science for providing the public Allen Mouse Brain Atlas (http://mouse.brain-map.org/) and to the creators of TuxedoSuite, HTSeq, and DESeq2 for providing essential, publicly available

RNA-seq analysis and visualization software. Funding was provided by the Howard Hughes Medical Institute.

Received: March 31, 2015
Revised: November 30, 2015
Accepted: December 7, 2015
Published: January 14, 2016

REFERENCES

Amaral, D.G., and Witter, M.P. (1989). The three-dimensional organization of the hippocampal formation: a review of anatomical data. Neuroscience *31*, 571–591.

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq–a Python framework to work with high-throughput sequencing data. Bioinformatics *31*, 166–169.

Belgard, T.G., Marques, A.C., Oliver, P.L., Abaan, H.O., Sirey, T.M., Hoerder-Suabedissen, A., García-Moreno, F., Molnár, Z., Margulies, E.H., and Ponting, C.P. (2011). A transcriptomic atlas of mouse neocortical layers. Neuron *71*, 605–616.

Bernard, A., Lubbers, L.S., Tanis, K.Q., Luo, R., Podtelezhnikov, A.A., Finney, E.M., McWhorter, M.M., Serikawa, K., Lemon, T., Morgan, R., et al. (2012). Transcriptional architecture of the primate neocortex. Neuron *73*, 1083–1099.

Collin, S.H., Milivojevic, B., and Doeller, C.F. (2015). Memory hierarchies map onto the hippocampal long axis in humans. Nat. Neurosci. *18*, 1562–1564.

Dolorfo, C.L., and Amaral, D.G. (1998). Entorhinal cortex of the rat: topographic organization of the cells of origin of the perforant path projection to the dentate gyrus. J. Comp. Neurol. *398*, 25–48.

Dong, H.W., Swanson, L.W., Chen, L., Fanselow, M.S., and Toga, A.W. (2009). Genomic-anatomic evidence for distinct functional domains in hippocampal field CA1. Proc. Natl. Acad. Sci. USA *106*, 11794–11799.

Dougherty, K.A., Islam, T., and Johnston, D. (2012). Intrinsic excitability of CA1 pyramidal neurones from the rat dorsal and ventral hippocampus. J. Physiol. *590*, 5707–5722.

Dougherty, K.A., Nicholson, D.A., Diaz, L., Buss, E.W., Neuman, K.M., Chetkovich, D.M., and Johnston, D. (2013). Differential expression of HCN subunits alters voltage-dependent gating of h-channels in CA1 pyramidal neurons from dorsal and ventral hippocampus. J. Neurophysiol. *109*, 1940–1953.

Elston, G.N. (2002). Cortical heterogeneity: implications for visual processing and polysensory integration. J. Neurocytol. *31*, 317–335.

Fanselow, M.S., and Dong, H.W. (2010). Are the dorsal and ventral hippocampus functionally distinct structures? Neuron *65*, 7–19.

Freese, J.L., and Amaral, D.G. (2005). The organization of projections from the amygdala to visual cortical areas TE and V1 in the macaque monkey. J. Comp. Neurol. *486*, 295–317.

Goaillard, J.M., Taylor, A.L., Schulz, D.J., and Marder, E. (2009). Functional consequences of animal-to-animal variation in circuit parameters. Nat. Neurosci. *12*, 1424–1430.

Goff, L., Trapnell, C., and Kelley, D. (2013). cummeRbund: analysis, exploration, manipulation, and visualization of Cufflinks high-throughput sequencing data, R package version 2.8.0.

Graves, A.R., Moore, S.J., Bloss, E.B., Mensh, B.D., Kath, W.L., and Spruston, N. (2012). Hippocampal pyramidal neurons comprise two distinct cell types that are countermodulated by metabotropic receptors. Neuron *76*, 776–789.

Groenewegen, H.J., Vermeulen-Van der Zee, E., te Kortschot, A., and Witter, M.P. (1987). Organization of the projections from the subiculum to the ventral striatum in the rat. A study using anterograde transport of Phaseolus vulgaris leucoagglutinin. Neuroscience *23*, 103–120.

Hawrylycz, M.J., Lein, E.S., Guillozet-Bongaarts, A.L., Shen, E.H., Ng, L., Miller, J.A., van de Lagemaat, L.N., Smith, K.A., Ebbert, A., Riley, Z.L., et al. (2012). An anatomically comprehensive atlas of the adult human brain transcriptome. Nature *489*, 391–399.

Hempel, C.M., Sugino, K., and Nelson, S.B. (2007). A manual method for the purification of fluorescently labeled neurons from the mammalian brain. Nat. Protoc. *2*, 2924–2929.

Hitti, F.L., and Siegelbaum, S.A. (2014). The hippocampal CA2 region is essential for social memory. Nature *508*, 88–92.

Igarashi, K.M., Ito, H.T., Moser, E.I., and Moser, M.B. (2014). Functional diversity along the transverse axis of hippocampal area CA1. FEBS Lett. *588*, 2470–2476.

Ishizuka, N., Weber, J., and Amaral, D.G. (1990). Organization of intrahippocampal projections originating from CA3 pyramidal cells in the rat. J. Comp. Neurol. *295*, 580–623.

Jarsky, T., Mady, R., Kennedy, B., and Spruston, N. (2008). Distribution of bursting neurons in the CA1 region and the subiculum of the rat hippocampus. J. Comp. Neurol. *506*, 535–547.

Kishi, T., Tsumori, T., Yokota, S., and Yasui, Y. (2006). Topographical projection from the hippocampal formation to the amygdala: a combined anterograde and retrograde tracing study in the rat. J. Comp. Neurol. *496*, 349–368.

Kohara, K., Pignatelli, M., Rivest, A.J., Jung, H.Y., Kitamura, T., Suh, J., Frank, D., Kajikawa, K., Mise, N., Obata, Y., et al. (2014). Cell type-specific genetic and optogenetic tools reveal hippocampal CA2 circuits. Nat. Neurosci. *17*, 269–279.

Komatsu, Y., Watakabe, A., Hashikawa, T., Tochitani, S., and Yamamori, T. (2005). Retinol-binding protein gene is highly expressed in higher-order association areas of the primate neocortex. Cereb. Cortex *15*, 96–108.

Kondo, H., Hashikawa, T., Tanaka, K., and Jones, E.G. (1994). Neurochemical gradient along the monkey occipito-temporal cortical pathway. Neuroreport *5*, 613–616.

Lau, C., Ng, L., Thompson, C., Pathak, S., Kuan, L., Jones, A., and Hawrylycz, M. (2008). Exploration and visualization of gene expression with neuroanatomy in the adult mouse brain. BMC Bioinformatics *9*, 153.

Lee, S.H., Marchionni, I., Bezaire, M., Varga, C., Danielson, N., Lovett-Barron, M., Losonczy, A., and Soltesz, I. (2014). Parvalbumin-positive basket cells differentiate among hippocampal pyramidal cells. Neuron *82*, 1129–1144.

Lein, E.S., Hawrylycz, M.J., Ao, N., Ayres, M., Bensinger, A., Bernard, A., Boe, A.F., Boguski, M.S., Brockway, K.S., Byrnes, E.J., et al. (2007). Genome-wide atlas of gene expression in the adult mouse brain. Nature *445*, 168–176.

Leonardo, E.D., Richardson-Jones, J.W., Sibille, E., Kottman, A., and Hen, R. (2006). Molecular heterogeneity along the dorsal-ventral axis of the murine hippocampal CA1 field: a microarray analysis of gene expression. Neuroscience *137*, 177–186.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550.

Malik, R., Dougherty, K.A., Parikh, K., Byrne, C., and Johnston, D. (2015). Mapping the electrophysiological and morphological properties of CA1 pyramidal neurons along the longitudinal hippocampal axis. Hippocampus. Published online October 10, 2015. http://dx.doi.org/10.1002/hipo.22526.

Mizuseki, K., Diba, K., Pastalkova, E., and Buzsáki, G. (2011). Hippocampal CA1 pyramidal cells form functionally distinct sublayers. Nat. Neurosci. *14*, 1174–1181.

Moser, M.B., and Moser, E.I. (1998). Functional differentiation in the hippocampus. Hippocampus *8*, 608–619.

Ng, L., Bernard, A., Lau, C., Overly, C.C., Dong, H.W., Kuan, C., Pathak, S., Sunkin, S.M., Dang, C., Bohland, J.W., et al. (2009). An anatomic gene expression atlas of the adult mouse brain. Nat. Neurosci. *12*, 356–362.

Pollen, A.A., Nowakowski, T.J., Shuga, J., Wang, X., Leyrat, A.A., Lui, J.H., Li, N., 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.

Prinz, A.A., Bucher, D., and Marder, E. (2004). Similar network activity from disparate circuit parameters. Nat. Neurosci. *7*, 1345–1352.

Risold, P.Y., and Swanson, L.W. (1996). Structural evidence for functional domains in the rat hippocampus. Science *272*, 1484–1486.

Ropireddy, D., and Ascoli, G.A. (2011). Potential synaptic connectivity of different neurons onto pyramidal cells in a 3D reconstruction of the rat hippocampus. Front. Neuroinform. *5*, 5.

Sansom, S.N., and Livesey, F.J. (2009). Gradients in the brain: the control of the development of form and function in the cerebral cortex. Cold Spring Harb. Perspect. Biol. *1*, a002519.

Shin, J., Ming, G.L., and Song, H. (2014). Decoding neural transcriptomes and epigenomes via high-throughput sequencing. Nat. Neurosci. *17*, 1463–1475.

Slomianka, L., Amrein, I., Knuesel, I., Sørensen, J.C., and Wolfer, D.P. (2011). Hippocampal pyramidal cells: the reemergence of cortical lamination. Brain Struct. Funct. *216*, 301–317.

Soltesz, I. (2006). Diversity in the Neuronal Machine: Order and Variability in Interneuronal Microcircuits (Oxford University Press).

Strange, B.A., Witter, M.P., Lein, E.S., and Moser, E.I. (2014). Functional organization of the hippocampal longitudinal axis. Nat. Rev. Neurosci. *15*, 655–669.

Thompson, C.L., Pathak, S.D., Jeromin, A., Ng, L.L., MacPherson, C.R., Mortrud, M.T., Cusick, A., Riley, Z.L., Sunkin, S.M., Bernard, A., et al. (2008). Genomic anatomy of the hippocampus. Neuron *60*, 1010–1021.

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics *25*, 1105–1111.

Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., and Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat. Biotechnol. *28*, 511–515.

Usoskin, D., Furlan, A., Islam, S., Abdo, H., Lönnerberg, P., Lou, D., Hjerling-Leffler, J., Haeggström, J., Kharchenko, O., Kharchenko, P.V., et al. (2015). Unbiased classification of sensory neuron types by large-scale single-cell RNA sequencing. Nat. Neurosci. *18*, 145–153.

Xu, L., Tanigawa, H., and Fujita, I. (2003). Distribution of alpha-amino-3-hydroxy-5-methyl-4-isoxazolepropionate-type glutamate receptor subunits (GluR2/3) along the ventral visual pathway in the monkey. J. Comp. Neurol. *456*, 396–407.

Zeisel, A., Muñoz-Manchado, A.B., Codeluppi, S., Lönnerberg, P., La Manno, G., Juréus, A., Marques, S., Munguba, H., He, L., Betsholtz, C., et al. (2015). Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. Science *347*, 1138–1142.

Zhang, Y., Chen, K., Sloan, S.A., Bennett, M.L., Scholze, A.R., O'Keeffe, S., Phatnani, H.P., Guarnieri, P., Caneda, C., Ruderisch, N., et al. (2014). An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. J. Neurosci. *34*, 11929–11947.