

Integration of gene expression and brain-wide connectivity reveals the multiscale organization of mouse hippocampal networks

Michael S. Bienkowski^{1*}, Ian Bowman¹, Monica Y. Song¹, Lin Gou¹, Tyler Ard¹, Kaelan Cotter¹, Muye Zhu¹, Nora L. Benavidez¹, Seita Yamashita¹, Jaspar Abu-Jaber¹, Sana Azam¹, Darrick Lo¹, Nicholas N. Foster¹, Houri Hintiryan¹ and Hong-Wei Dong^{1,2,3*}

Understanding the organization of the hippocampus is fundamental to understanding brain function related to learning, memory, emotions, and diseases such as Alzheimer's disease. Physiological studies in humans and rodents have suggested that there is both structural and functional heterogeneity along the longitudinal axis of the hippocampus. However, the recent discovery of discrete gene expression domains in the mouse hippocampus has provided the opportunity to re-evaluate hippocampal connectivity. To integrate mouse hippocampal gene expression and connectivity, we mapped the distribution of distinct gene expression patterns in mouse hippocampus and subiculum to create the Hippocampus Gene Expression Atlas (HGEA). Notably, previously unknown subiculum gene expression patterns revealed a hidden laminar organization. Guided by the HGEA, we constructed the most detailed hippocampal connectome available using Mouse Connectome Project (<http://www.mouseconnectome.org>) tract tracing data. Our results define the hippocampus' multiscale network organization and elucidate each subnetwork's unique brain-wide connectivity patterns.

The hippocampus is a highly specialized cortical structure located in the mammalian medial temporal lobe. Numerous studies over the last several decades have suggested that there is structural and functional heterogeneity along the hippocampal longitudinal axis (dorsoventral axis in rodents, posterior/anterior axis in humans)^{1–6}. Behavioral rodent studies have suggested that dorsal parts of the hippocampus are cognitive regions that are responsible for the spatial navigation and cognitive functions, whereas ventral hippocampus parts are affective regions that are related to emotional and affective behaviors and are a core 'limbic system' component of the classic Papez circuit^{1,4}. However, gene expression and anatomical evidence suggests a more complicated organization in the hippocampus⁷. Decades of anatomical tract tracing studies in rats have suggested that there are topographically organized tripartite hippocampal projections from the entorhinal cortex (ENT)^{1,6}, whereas connectivity between different hippocampal components (dentate gyrus (DG), CA3, CA1, and the subiculum (SUB)) is organized as a continuous gradient with no definitive boundaries between the dorsal, intermediate, and ventral parts⁸.

Gene expression studies based on large-scale mRNA in situ hybridization data and modern RNAseq technologies have revealed a complex molecular heterogeneity of the mouse hippocampal formation^{1,9–14}, which suggests that the classic delineation of the hippocampus can be further refined. For example, recent mouse studies have revealed multiple discrete gene expression domains along the longitudinal axis with as many as nine CA3 subregions, whereas the classic CA1 has been divided into the dorsal, intermediate, and ventral domains^{1,9,10}. However, no consensus has been reached for how these hippocampal divisions should be refined, and many of these

studies focused only on 'representative levels' and did not investigate the entire hippocampus^{1,9,10,15–17}. Even less understood is how molecular heterogeneity of the hippocampus (transcriptomic data were mostly obtained in mouse) is correlated with its anatomical connectivity (which has been mostly conducted in rats). Although recent mouse connectomics approaches^{18,19} have collected large-scale connectivity data, mouse hippocampus connectivity and gene expression data have not been systematically analyzed together.

To understand the relationship between hippocampal gene expression and anatomical connectivity, we analyzed and annotated over 250 genes expressed throughout the entire hippocampus and subiculum, which are presented in the Allen Brain Atlas online gene expression database (<http://www.brain-map.org>). We mapped the full rostrocaudal extent of each gene expression domain onto all corresponding coronal and sagittal Allen Reference Atlas (ARA) levels²⁰ to create the HGEA. Guided by the HGEA, we examined inputs and outputs of each HGEA domain and found that the gene expression boundaries delineated by the HGEA strongly aligned with anatomical connectivity patterns. Accordingly, we created the most complete hippocampal network wiring diagram to date, which can serve as a foundation for the functional dissection of the hippocampus.

Results

General strategy for refining hippocampus parcellation on the basis of a combination of molecular characteristics and connectivity. To construct an accurate mouse hippocampus connectome, we refined the anatomical hippocampus delineation using a combination of large-scale gene expression and connectivity data

¹University of Southern California Stevens Neuroimaging and Informatics Institute, Center for Integrated Connectomics (CIC), Keck School of Medicine of University of Southern California, Los Angeles, CA, USA. ²Department of Neurology, Keck School of Medicine of University of Southern California, Los Angeles, CA, USA. ³Department of Physiology and Neuroscience, and Zilkha Neurogenetic Institute, Keck School of Medicine of University of Southern California, Los Angeles, CA, USA. *e-mail: michael.bienkowski@ini.usc.edu; hongwei.dong@ini.usc.edu

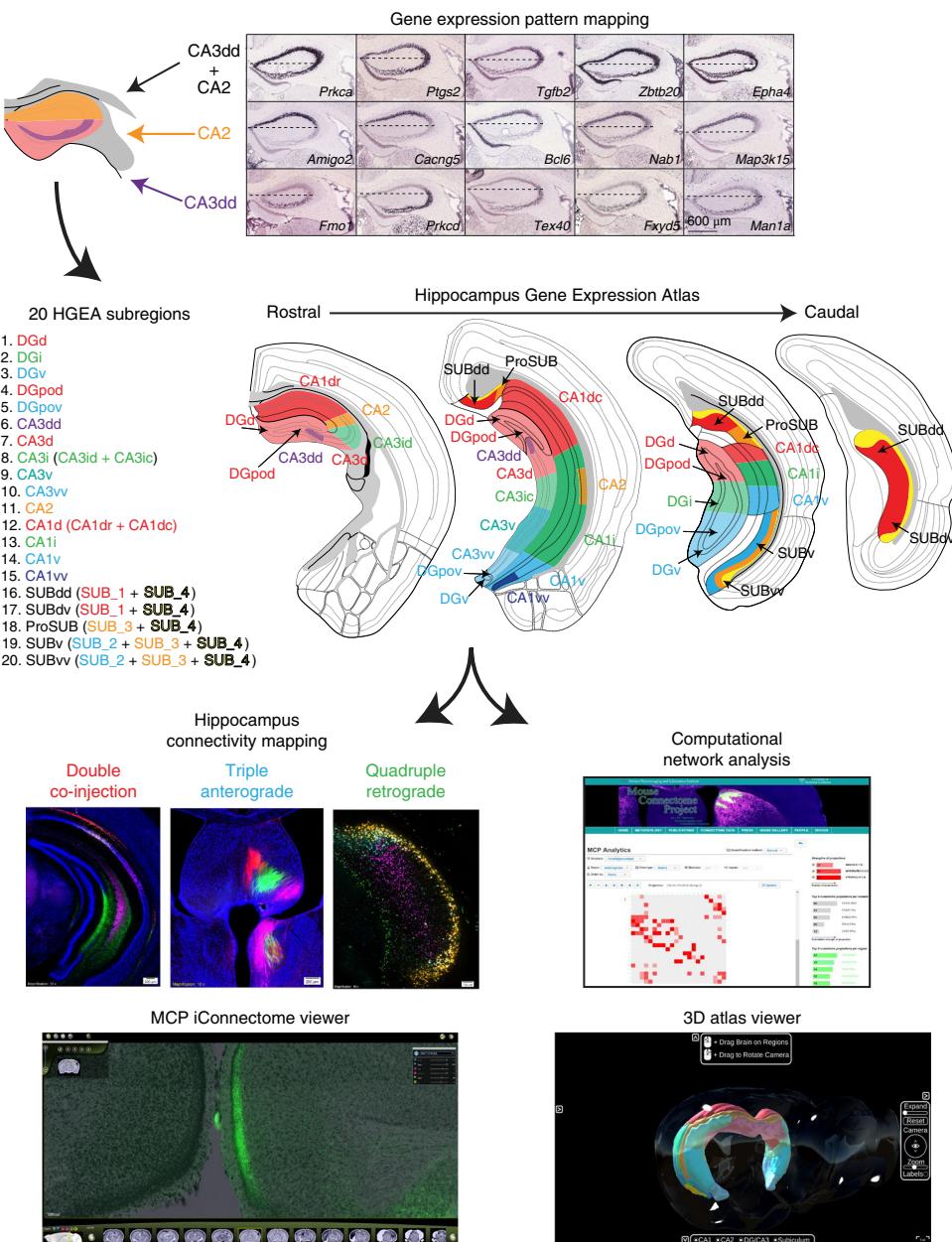


Fig. 1 | Experimental workflow. Top, HGEA subregions were defined and mapped by the consensus of multiple gene expression patterns. Scale bar represents 600 μ m. As an example, five gene expression distributions are shown that contrast CA3dd (bottom row) and CA2 (middle row), as well as five genes that were expressed in both CA3dd and CA2 (top row). Horizontal dashed line in each image corresponds to the division of CA2 genes versus CA3dd genes, which corresponds with the mapped boundary of the CA3dd/CA2 border at HGEA level 67. Consensus subregion boundaries were mapped in a similar fashion for DG, CA3, CA2, CA1, and SUB at all rostrocaudal levels of the hippocampus (four representative levels are shown in the middle panel). Following the creation of the atlas, we examined the connectivity of each HGEA subregion as part of the Mouse Connectome Project (MCP). We used combinations of multiple retrograde and anterograde tracers in a variety of experimental strategies (double co-injection, triple anterograde, quadruple retrograde). As a resource, we have made all data openly available online at our MCP website (<http://www.mouseconnectome.org>); including connectivity image data, HGEA atlas sections, annotation tables, network analyses, and three-dimensional (3D) visualization tools (*in situ* hybridization image data can also be found at Allen Institute's website, <http://www.brain-map.org>). To present the anatomical data in a clear and detailed manner, case number (white) and injection site (color-coded with fluorescent labeling color) are listed at the bottom of each image (all case numbers, injection sites, and tracers are listed in Supplementary Table 3).

(Fig. 1 and Supplementary Table 1). First, we systematically analyzed and annotated the spatial expression patterns of 258 hippocampal-enriched genes selected from thousands of genes in the Allen Brain Atlas database (<http://www.brain-map.org>; Supplementary Table 2). By analyzing multiple gene expression distributions rather than a single gene, we found that consensus patterns of similarities and differences became apparent (Fig. 1). Genes were sometimes expressed

uniquely in the major hippocampal regions (that is, CA3 versus CA1) or in distinct combinations (that is, DG and CA3, but not CA1). Other genes were more restricted in discrete parts of a region, thereby defining distinct domains that would otherwise have been indistinguishable. After determining a consensus gene-expression-based hippocampal parcellation, we mapped the boundaries of gene expression domains in the DG, CA3, CA2, CA1, and SUB for both

coronal and sagittal data to create the HGEA (Supplementary Figs. 1 and 2, and Supplementary Videos 1–5).

In creating the HGEA, we chose to manually analyze and delineate hippocampal boundaries resulting from inconsistent tissue sectioning across the coronal and sagittal Allen Brain Atlas database (see Methods). The Allen Brain Atlas database is filled with thousands of mouse brains, each identifying only one gene's expression pattern. Comparing different gene expression patterns across animals is complicated by differences in the oblique tissue sectioning of each brain (Supplementary Fig. 3). To account for this variability, we considered each part of each tissue section as being independent (based on local anatomical landmarks) when manually mapping boundaries to the atlas. By applying this approach, gene expression mapping across all atlas levels allows for a more complete understanding of the shifting gene expression boundaries across the rostrocaudal axis.

Accordingly, we identified more than 20 molecular domains that were mapped onto the HGEA (see Fig. 1): three in the DG granule cell layer (dorsal (DGd), intermediate (DGi), and ventral domains (DGv)), two in the DG polymorph layer (dorsal (DGpod) and ventral domains (DGpov)), five in the CA3 (a rostral-dorsal tip (CA3dd), dorsal (CA3d), intermediate (CA3i), ventral (CA3v), and a ventral tip (CA3vv)), four in the CA1 (dorsal (CA1d), intermediate (CA1i), ventral (CA1v), and ventral tip (CA1vv)), and five in the SUB (described below; sagittal gene expression data with sagittal HGEA atlas shown in Supplementary Fig. 4). Overall, these results consolidate and refine earlier gene expression domain descriptions in DG, CA3, and CA1 (refs. ^{1,9–14}). In addition, we compared the binary HGEA annotation to the more quantitative single-cell RNAseq database DropViz²¹ (<http://www.dropviz.org>) and found strong agreement between the two datasets. On average, $91.7 \pm 0.8\%$ of HGEA annotated genes were found to have positive expression in related DropViz 'global clusters' (analysis available at <http://www.mouseconnectome.org>). A principal component analysis (PCA) of 248 HGEA genes found in the DropViz database generated a highly similar clustering pattern of hippocampal cell types (meta cells) compared with PCA using all 32,307 DropViz genes (Supplementary Fig. 5). Finally, we calculated a dissimilarity matrix between the binary HGEA gene expression annotation and thresholded, binarized DropViz meta-cell data of the 248 HGEA genes found in both datasets. The dissimilarity matrix revealed strong similarities between specific HGEA regions and DropViz cell types, supporting the idea that each HGEA region is composed of distinct combinations of individual cell-types (Supplementary Table 4). After creating the HGEA, we systematically examined each domain's input and output connectivity, constructed the hippocampus wiring diagram, and analyzed hippocampal network organization. The entire tissue series for all tract tracing data are openly available by searching the case number through the iConnectome viewer at our website (<http://www.mouseconnectome.org>). Furthermore, we created several informatics and visualization tools to help navigating hippocampal networks (www.mouseconnectome.org/analytics/page/matrix; Fig. 1).

Delineating inputs/outputs of HGEA regions. DG gene expression subdomains and anatomical connectivity. The DG granule cell layer was divided into the dorsal, intermediate, and ventral domains (DGd, DGi, and DGv), whereas the DG polymorph layer contained two distinct dorsal and ventral cell groups (DGpod versus DGpov; Supplementary Fig. 6a). The DGpo is believed to contain a variety of neuronal cell types, including interneurons, which receive local inputs from the DG granule cells. The major DGpo projection neurons are the mossy cells, which are characterized by their longitudinally broad bilateral projections to the inner-third of the DG molecular layer²². Both gene expression and connectivity data determined at least two discrete populations of mossy cells in the DGpod

and DGpov, which displayed distinct projection patterns to innervate different segments of DG granular cells (Supplementary Fig. 6b). DGpod neurons generated dense bilateral axonal projections to the DGd and DGi (DGpod → DGi/DGd), whereas DGpov neurons generated axonal projections to the DGv and DGi (DGpov → DGv/DGi). This connectivity pattern suggests that DG information can be integrated only in a ventral-to-dorsal direction (DGv → DGpov → DGi → DGpod → DGd), but DGd information may not directly reach the DGv (Fig. 2e).

Our DG connectivity analysis revealed differences between the DGd versus DGi and DGv mossy fiber pathways. DGd mossy fibers followed a hairpin-shaped connectivity motif in which they first innervated the DGpo locally before traveling through the CA3 stratum lucidum and terminating on CA3 neurons along this path. DGd mossy fibers continued through CA3 up to 1 mm rostrally before turning back sharply to finally terminate in a topographically organized CA2 subregions (Supplementary Fig. 6c,e). In contrast, DGi and DGv mossy fibers traveled rostrally in a straight line to innervate CA3ic and CA3v/CA3vv, respectively (Fig. 2d and Supplementary Fig. 6d,f).

CA3, CA2, and CA1 gene expression subdomains and anatomical connectivity. The gene-expression-defined CA3dd, CA3d, and CA3id resembled the morphologically defined 'CA3c', 'CA3b', and 'CA3a' subregions first proposed by Lorente de Nò²³, but other CA3 and CA1 gene expression domains did not appear to match this classic delineation. In the CA1, gene expression patterns showed regional differences that corresponded to HGEA boundaries, although a more complex laminar pattern may exist (Fig. 2c). Finally, consistent with a previous study⁹, some genes uniquely identified the ventral tip of the CA1v (CA1vv) as a distinct single-layer domain (Supplementary Fig. 4b).

Our anterograde and retrograde co-injection data and multiple retrograde tracing data revealed that CA3, CA2, and CA1 domain connectivity patterns are highly consistent with other HGEA molecular domains (Fig. 2 and Supplementary Fig. 7). CA3 gene expression patterns revealed highly restricted subdomains along the dorsal/ventral axis that were used to map HGEA delineations (Fig. 2a,b). After connectivity mapping, we found anatomical evidence to distinguish the more rostrally located dorsal CA3i (CA3id) from the caudal CA3i (CA3ic; Fig. 2a,b), as well as the rostral and caudal CA1d (CA1dr versus CA1dc), but these subdivisions were not segregated in the gene-expression-based HGEA (Supplementary Fig. 7). Using multiple retrograde tracer injections, we found retrograde labeling patterns that appeared to be highly consistent with the boundaries outlined by the HGEA. Retrograde tracer injections into both CA3 and CA1 subregions produced labeling patterns that respected HGEA borders, suggesting that both CA3 associational connections and Schaffer collateral pathways are organized by CA3 gene expression domains (Fig. 2a,b). In addition, we found that anterograde and retrograde co-injections into individual CA3 subregions labeled discrete sets of bilateral CA3 and CA1 subregions (Fig. 2c and Supplementary Fig. 8). CA3 domains projected topographically and bilaterally to progressively more proximal CA1 domains (Fig. 2c,e). In general, CA3dd projected to distal CA1d; CA3d and CA3id projected to proximal CA1d, CA2, and CA1i; CA3ic projected to CA2, CA1i, and CA1v; CA3v projected to CA1v and SUBv; and CA3vv projected to CA1vv and SUBvv (note the dorsoventral limits of these terminal fields were highly restricted by HGEA boundaries rostrocaudally; Supplementary Fig. 8).

The CA2 has recently received a great deal of attention because of the discovery of many specific gene expression markers that extend its classic borders (Supplementary Fig. 3e,f)^{24–26}. Consistent with a recent monosynaptic rabies tracing study²⁶, our anterograde tracing data support the DGd input to CA2 along the rostrocaudal axis (Supplementary Fig. 6c,e). CA2 neurons received DGd input along

the rostrocaudal axis and then projected topographically to the entire CA1 axis (except CA1vv). CA2 fibers in CA1 were primarily targeted to stratum oriens, where a variety of inhibitory neuron cell types are distributed²⁷. More rostral neurons in CA2 projected to the CA1dr and CA1dc, whereas more caudal CA2 neurons projected to the CA1i and CA1v.

Subiculum gene expression subdomains and anatomical connectivity. In the SUB, we found that gene expression patterns subdivided the pyramidal layer into four layers, although the layer distribution was non-uniform (Fig. 3a–c and Supplementary Figs. 1, 2, and 4d). Similar to cerebral cortex parcellation, we defined five SUB subdivisions on the basis of the unique combination and distribution of their gene expression layers. First, a two-layered region, which is located in the ‘classic’ dorsal subiculum (SUBd), contained gene expression layers 1 and 4 (red and yellow in HGEA, respectively). At both the dorsal and ventral ends of this domain, layer 4 formed a thicker, bulb-like shape which distinguished the dorsal and ventral parts of the SUBd (SUBdd versus SUBdv; Fig. 3c). Next, a two-layered region containing gene expression layers 3 and 4 was located between the SUBdd and CA1d in a region that was first identified

previously as the prosubiculum (ProSUB), although its presence in the mouse has remained controversial^{23,28}. Furthermore, a three-layered subdivision comprising gene expression layers 2, 3, and 4 was comparable to the classic ‘ventral’ subiculum (SUBv; Fig. 3b). Finally, we also distinguished the ventral tip of SUBv (SUBvv), where layer 4 thickened and layer 2 thinned before curving to wrap around the other layers.

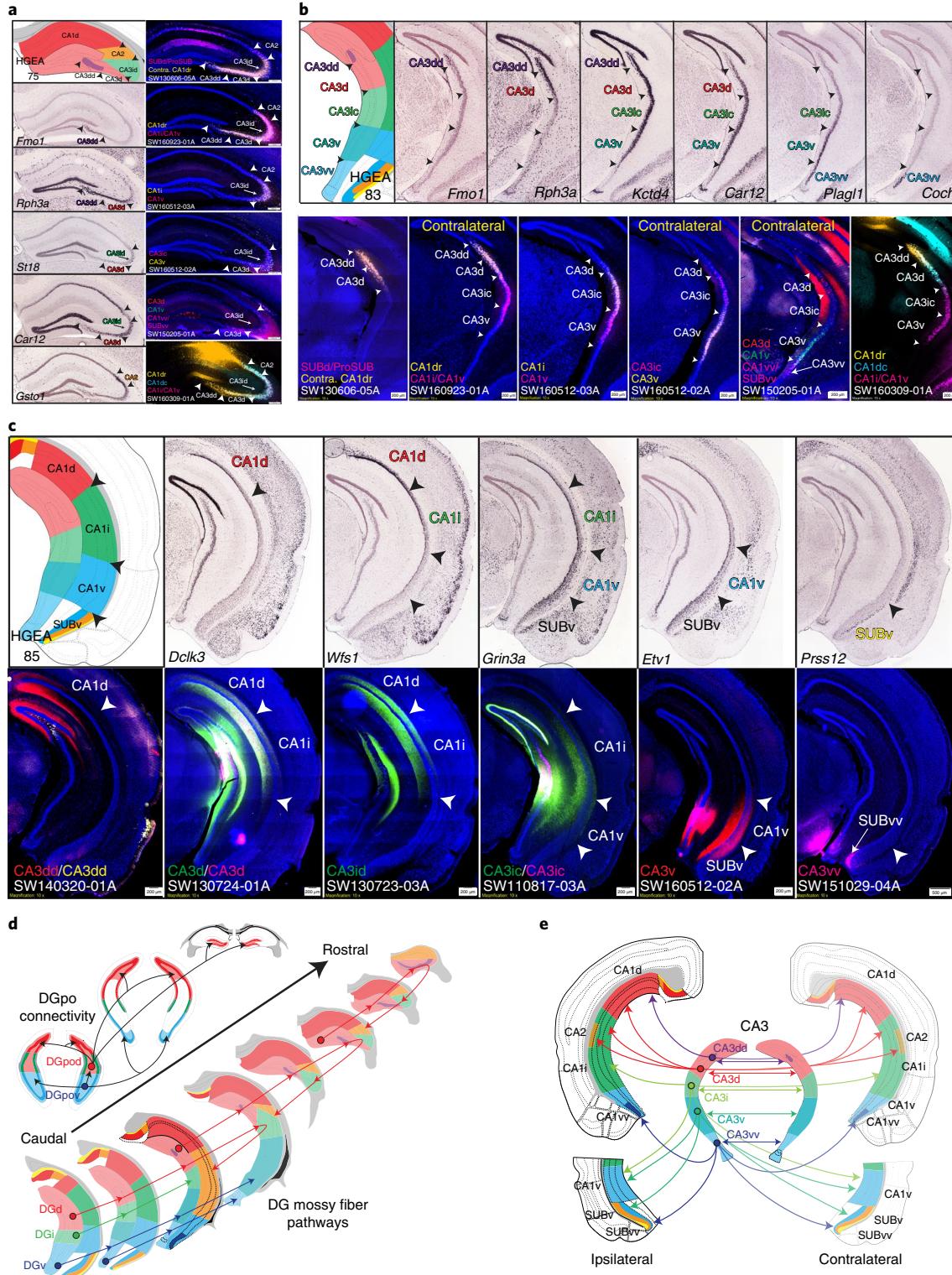
Our anatomical tracing data revealed that SUB connectivity was reflected by both the gene expression laminar organization and sub-regions outlined by the HGEA (Fig. 3a,b and Supplementary Video 6). In the SUBv and SUBvv, layers 2 and 4 mediated distinct connectivity pathways with CA1 (Fig. 3d), whereas layer 4 also contained thalamic-projecting neurons (Fig. 3e). Each CA1 domain targeted discrete combinations of SUB domains, although CA1d organization was particularly complex (Fig. 3f). CA1d output to the SUBdd and ProSUB was organized rostrocaudally, whereas CA3 input to CA1d was organized along the proximo-distal axis (Supplementary Fig. 7). Overall, SUB gene expression layer connectivity was organized similar to the isocortex: SUB layer 2 primarily mediated intrahippocampal connectivity similar to the cortico-cortical connectivity mediated by isocortex layers 2 and 3; SUB layers 1 and 3

Fig. 2 | CA3 and CA1 gene expression and anatomical labeling patterns match closely with HGEA delineation. For each image of tracer labeling, tracer injection sites and corresponding experimental case numbers (that is, SW130606-05A in a) are indicated in the bottom left corner (that is, ProSUB/SUBdd or contra CA1dr). **a**, Rostral hippocampus gene expression patterns and corresponding HGEA atlas section are shown on the left (*Fmo1*=CA3dd, *Rph3a*=CA3dd+CA3d, *St18*=CA3id, *Car12*=CA3d+CA3id, *Gst01*=CA2), and multiple retrograde labeling patterns from combinations of CA3 and CA1 subregions are shown on the right. Arrowheads demarcate HGEA subregion boundaries across the images. Retrograde injections into different combinations of CA3 and CA1 regions produced bilateral CA3 retrograde labeling patterns that appeared consistent with gene-expression-defined HGEA subregion boundaries. Retrograde tracer injection into the contralateral CA1dr (SW130606-05A) retrogradely labeled neurons in the CA3dd, CA3d, and CA3id (yellow) adjacent to CTB labeling (magenta) in the rostral CA2 (from ipsilateral ProSUB, some leakage into SUBdd). Retrograde injection into the CA1dr (SW160923-01A) revealed a distinct set of labeling in the CA3dd/CA3d (yellow), but overlapping labeling with CA1i/CA1v-projecting neurons in the CA2 and CA3id (magenta). Retrograde injections into the CA1i (SW160512-03A) retrogradely labeled neurons in the CA3id (yellow), but no labeling was observed from CA1v injection (magenta). Retrograde injections into the CA3ic (magenta) and CA3v (yellow; SW160512-02A) both produced restricted labeling in the ipsilateral CA3id, but in segregated superficial versus deep neuronal populations (compare with *St18* gene expression). Multiple retrograde injections in CA3d (red), CA1v (green), and CA1vv/SUBvv (magenta; SW150205-01A) revealed CA3d-projecting neurons in CA3d and CA3id (compare with *Car12* gene expression). Multiple retrograde tracer injections into the CA1dr (yellow), CA1dc (cyan), and CA1i/CA1v (magenta) produced overlapping labeling patterns in the CA3 and CA2, but with different labeling densities distinguishing each CA3 subregion. **b**, More caudal CA3 levels with gene expression patterns on top (*Fmo1*=CA3dd, *Rph3a*=CA3dd+CA3d, *Kctd4*=CA3dd+CA3d+CA3i+CA3v, *Car12*=CA3d+CA3id+CA3v, *Plagl1*=CA3v+CA3vv, *Coch*=CA3vv) and retrograde labeling patterns on bottom (same retrograde tracing experiments as described in a). In case SW130606-05A, contralateral CA1dr-projecting neurons (yellow), which were present in CA3dd, CA3d, and CA3id at more rostral levels (a), continued into caudal parts of CA3dd and CA3d, but not in CA3ic (compare with *Rph3a* gene expression at this level). Caudal levels of case SW160923-01A revealed that CA1dr-projecting neurons (yellow) in CA3dd and CA3d were adjacent to CA1i/CA1v-projecting neurons (magenta) in CA3ic, whereas these populations were overlapped in CA3id (a). In case SW160512-02A, CA1i-projecting neurons (yellow) were located in CA3d and CA3ic, whereas CA1v-projecting neurons (magenta) were located in CA3ic and CA3v. Together, cases SW130606-05A, SW160923-01A, and SW160512-02A indicate that CA3id and CA3ic both projected to CA1i, but could not be distinguished by projections to CA1d and CA1v, respectively (confirmed by anterograde labeling in c). Case SW160512-03A revealed associational connectivity of CA3ic (magenta) and CA3v (yellow) neurons in the contralateral CA3. Note the similarity in the CA3ic/CA3v boundary produced by CA3 associational connectivity in case SW160512-02A and Schaffer collateral CA1 connectivity in SW160512-03A. Case SW150205-01A revealed multiple retrograde labeling on the contralateral side after injections into CA3d (red), CA1v (green), and CA1vv/SUBvv (magenta). CA3d-projecting neurons were located in CA3d and CA3ic, CA1v-projecting neurons were located in CA3ic and CA3v, and CA1vv/SUBvv-projecting neurons were located adjacent in CA3vv. Finally, SW160309-01A showed different densities of CA1dr- (yellow) and CA1dc-projecting neurons (cyan) in CA3dd and CA3d adjacent to CA1i/CA1v-projecting neurons (magenta) in CA3ic and CA3v. **c**, Example gene expression patterns corresponded to HGEA CA1 subregions and defined CA1v/SUBv border (*Dclk3*=CA1d, *Wfs1*=CA1d+CA1i, *Grin3a*=CA1i+CA1v+SUBv, *Etv1*=CA1v+SUBv, *Prss12*=SUBv). Anterograde labeling of CA3 projections to CA1 revealed that partially overlapped topographic CA3 fiber distribution aligned well with gene-expression-defined HGEA CA1 subregion boundaries (additional rostrocaudal levels are shown in Supplementary Fig. 8). CA3dd projected to CA1d (more dense on medial side), CA3d and CA3id both projected to CA1d (more dense lateral side) and CA1i, CA3ic projected to CA1i and CA1v, CA3v projected to CA1v and SUBv, and CA3vv projected to CA1vv and SUBvv. **d**, Summary schematic of DGpo and DG mossy fiber connectivity organization (see Supplementary Fig. 5). DGpod neurons (red) projected bilaterally to the inner-third of the DG molecular layer across the rostrocaudal extent of the DGd and DGi, as well as directly targeting the contralateral DGpod. DGpov neurons (blue) produced a similar pattern, but instead targeted the DGv and DGi. Below, DGd granule cell mossy fiber pathways were unique compared with the DGi and DGv. DGd mossy fibers (red) innervated DGpo cells locally before extending rostrally through the CA3 up to 1mm before turning caudal to target a topographic part of CA2. The overall shape of this hairpin-like connectivity motif was different between neurons in the caudal versus rostral parts of DGd. In contrast, DGi (green) and DGv (blue) mossy fibers extended rostrally through DGpo and CA3 in a straight-line pattern. **e**, Summary schematic of CA3 subregion bilateral associational and Schaffer collateral projections to CA1. CA3dd (purple) projected to medial parts of CA1d, CA3d (red) projected to lateral parts of CA1d, CA2, and CA1i, CA3ic (green) projected to CA1i and CA1v, CA3v (teal) projected to CA1v and SUBv, and CA3vv (blue) projected to CA1vv and SUBvv. For the number of tracer experiments and cross-validated results, see Methods.

mediated different corticofugal pathways similar to isocortex layer 5; and SUB layer 4 contained thalamic-projecting neurons similar to isocortex layer 6 (Fig. 3g).

Multi-scale network organization of hippocampus connectivity. Using our anatomical tracer data, we manually annotated the inputs and outputs of all HGEA molecular domains to construct the hippocampus wiring diagram (Fig. 4a and Supplementary Table 5). To

further understand the network organization of hippocampus connectivity, we first annotated and weighted all of the intrahippocampal connections between the HGEA subregions (Supplementary Table 6) and constructed a weighted, directed intrahippocampal connectivity matrix (Fig. 4b). Using this data, we determined the intrahippocampal subnetworks by running 1,000 trials of a Louvain community detection algorithm at multiple gamma values (every 0.1 value between 0.1 and 20.0; 2,000,000 total iterations) and then



calculated a mean partition similarity (MPS) metric at each gamma value to determine the most representative consensus network partition. Gamma values with MPS peaks were observed at 0.15, 1.36, and 2.04 (Fig. 4b). Using these gamma values, we produced a schematic network graph and a reordered multiscale connectivity matrix featuring clustered network modules at gamma 2.04 as a subset of modules at gamma 1.36, which itself was a subset of modular networks at gamma 0.15 (Fig. 4b). We found that at gamma 0.15 the hippocampus was clearly divided into a ‘dorsal’ and ‘ventral’ hippocampus (Fig. 4b). At gamma 1.36, the dorsal hippocampus was divided into four modular subnetworks ([DGd, DGi, DGpod], [CA3dd, CA3d, CA3id, CA3ic], [CA1dr, SUBdd], and [CA1dc, CA2, ProSUB]) and the ventral hippocampus was divided into three modular subnetworks ([DGv, DGpov], [CA3v, CA1v, CA1i, SUBv, SUBdv], and [CA3vv, CA1vv, SUBvv]). Finally, at gamma 2.04, the only distinction from the network partition at gamma 1.36 was that the CA3dd and CA3d split from the CA3id and CA3ic. Overall, our intrahippocampal analysis indicates that the hippocampus is a multiscale structure that is organized as modular hierarchical subnetworks.

After defining the intrahippocampal subnetwork organization, we determined how each intrahippocampal subnetwork is connected to other brain regions as brain-wide neural networks. Using the annotated data shown in Supplementary Table 5, we ran the Louvain community detection algorithm using the same parameters as described for the intrahippocampal analysis. For this dataset, we found the highest MPS peak at gamma 9.74 with a corresponding consensus partition featuring 46 different communities. On the basis of these 46 different communities, we manually arranged each community into five subnetworks based on extrinsic connectivity similarity and constructed their network graphs (Fig. 4c,d). We found that the intrahippocampal subnetworks (from Fig. 4b) differentially innervated larger brain networks outside the hippocampus in a coordinated manner. The DG and CA3 regions contributed relatively little to extrahippocampal connectivity, but instead appeared to integrate information from multiple intrahippocampal subnetworks across the hippocampal axis (Fig. 4c). In contrast, the CA1/

CA2 and SUB regions were organized into three relatively distinct global networks (Fig. 4d).

Detailed connections of the three brain-wide extrahippocampal networks. *The CA1dr-SUBdd/SUBdv network.* Both SUBdd and SUBdv layer 1 neurons projected to complementary areas in the ventral retrosplenial cortex (RSPv), postsubiculum (POST), presubiculum (PRE), parasubiculum (PAR), and medial mammillary nucleus (MM), whereas SUBdd/SUBdv layer 4 neurons projected to the anterodorsal (AD), anteroventral thalamic nuclei (AV), and reunions thalamic nucleus (RE; Fig. 5a-d). The POST, PRE, and PAR generated dense projections through the fornix to the lateral mammillary nucleus (LM), which in turn sent massive bidirectional connections to the AD and AV. In parallel, both the SUBdd and SUBdv sent dense projections to the laterodorsal thalamic nucleus (LD), which had robust connectivity with the visual (VIS), retrosplenial (RSP), and POST/PRE/PAR cortical areas (Fig. 5b,e).

The CA2/CA1dc-ProSUB and CA1i/CA1v-SUBv networks. According to our intrahippocampal network analysis, the CA2/ProSUB and SUBv received distinct, but partially overlapping, intrahippocampal input: the ProSUB received primary input from the ‘dorsal’ hippocampus (primarily CA1dc), whereas the SUBv received input from the ‘ventral’ hippocampus (CA1i/CA1v). However, ProSUB and SUBv extrahippocampal connections were highly similar relative to the output of other CA1/SUB domains. Ultimately, our analysis grouped the CA1i/CA1v/SUBv and the CA2/CA1dc/ProSUB as one broad brain-wide network with the medial prefrontal cortex (MPF), amygdala, nucleus accumbens (ACB), and hypothalamus (Fig. 4d).

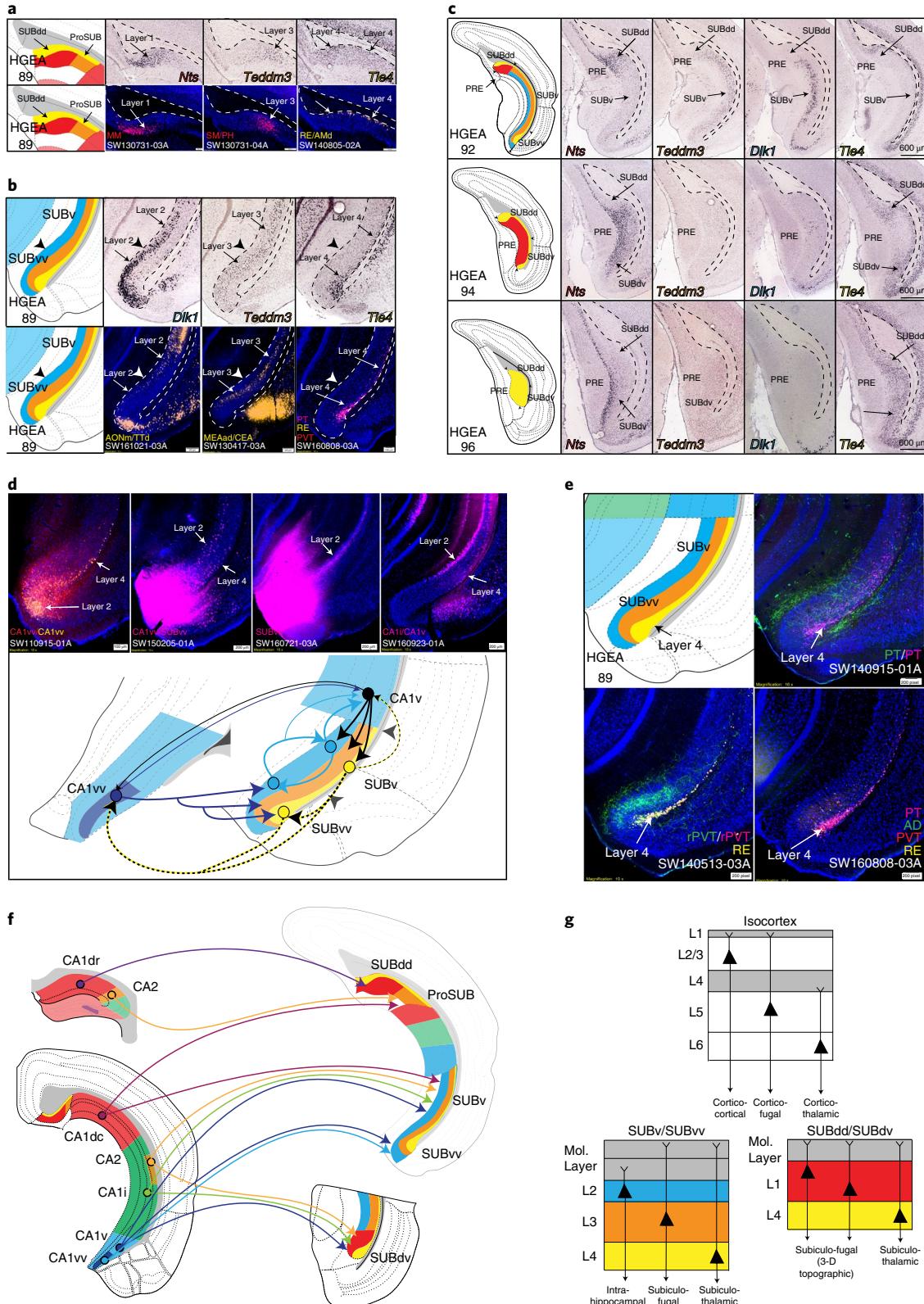
ProSUB and SUBv are the only SUB domains that target the MPF, which also received robust inputs from the CA1dc, CA1i, and CA1v. In addition, ProSUB and SUBv projections were complementary: the ProSUB primarily targeted the dorsal peduncular area (DP) and deep layer infralimbic cortex (ILA), whereas SUBv projected densely to superficial ILA layers and sparsely to ventral anterior cingulate cortex layer 6 (ACAv; Fig. 6a). The ProSUB and SUBv projected topographically to the amygdala, ACB, and hypothalamus.

Fig. 3 | SUB gene expression and anatomical labeling patterns produce similar laminar organization. **a**, Gene expression patterns of the classic dorsal subiculum define laminar and subregional organization that correspond to anatomical connectivity patterns. *Nts* and *Teddm3* expression demarcated HGEA layers 1 and 3, respectively, whereas *Tle4* expression identified the deeper HGEA layer 4 adjacent to the alveus. Below, example laminar retrograde labeling patterns of MM- (SW130731-03A), SM/PH- (SW130731-04A), and RE/AMd-projecting neurons (SW140805-02A) distinctly related to layers 1, 3, and 4, respectively. **b**, *Teddm3* and *Tle4* (layer 3 and 4) expression continued into the ‘classic’ ventral SUB, but both were located deeper in another layer identified by *Dlk1* expression (layer 2). Note that there was a thickening of layer 4 at the dorsal and ventral ends of subiculum that we refer to as the dorsal and ventral ‘bulbs’, respectively. Below, example laminar retrograde labeling patterns of AONm/TTd- (SW161021-03A), MEAad/CEA- (SW130417-03A), and multiple thalamic-projecting neuronal cell types (PT, RE, PVT; SW160808-03A) distinctly related to layers 2, 3, and 4, respectively. **c**, At caudal levels, the distribution of the gene expression layers changed following the disappearance of the CA1. Immediately caudal to CA1 at ARA level 92, layers 2 and 3 extended dorsally to border layer 1. At ARA level 93, a ventral region containing layer 1 appeared adjacent to the PAR (see Supplementary Fig. 1) and layer 1 became continuous at ARA level 94 following the emergence of the presubiculum (PRE). Finally, ARA level 96 represented the caudal end of the subiculum, where only gene expression layer 4 was present (*Nts* expression was in PRE layer 3). The combinations and distribution of these layers were used to define the five SUB subregions. Scale bars represent 600 μ m. **d**, Tracer co-injections in ventral CA1 and SUB regions revealed complex laminar-specific interconnectivity. CA1vv co-injection (SW110915-01A) produced anterograde labeling across all SUBvv layers and retrograde labeling in SUBvv layers 2 and 4. CTB injection into CA1vv/SUBvv (SW150205-01A) produced retrograde labeling in SUBv and SUBvv layer 4 and the deeper part of SUBv and SUBvv layer 2. PHAL injection into SUBvv (SW160721-03A) produced anterograde labeling that travels along SUBv and SUBvv layer 2. Retrograde injection into CA1v (SW160923-01A) produced labeling in SUBv layer 4 and superficial SUBv layer 2. Together, these data suggest that CA1v, CA1vv, SUBv, and SUBvv are bidirectionally connected through two different systems of connections mediated through layers 2 and 4 (see summary diagram below). **e**, Co-injection into the paratenial thalamus (PT, SW140915-01A) produced anterograde and retrograde labeling in SUBv and SUBvv layers 3 and 4, as well as anterograde labeling in the SUBvv deep molecular layer. Co-injection into the rostral paraventricular thalamus (rPVT, SW140513-03A) and Fluorogold injection into the reunions thalamus (RE) revealed strong retrograde labeling in SUBv and SUBvv layer 4 and anterograde labeling in SUBvv layer 3 and deep molecular layer. Multiple retrograde tracer injections within the PT, PVT, RE, and anterodorsal thalamus (AD, SW160808-03A) produced retrograde labeling that was highly restricted to layer 4. **f**, Summary schematic of CA1 and CA2 projections to the five SUB subregions. **g**, General laminar organization of pyramidal neurons in the isocortex (top) compared similarly to the gene expression laminar organization of the SUBv/SUBvv and SUBdd/SUBdv pyramidal neurons. Extension of dendrites into the molecular layer is based on interpretation of rabies-labeled morphology (for example, see Figs 5c and 6b). All *in situ* hybridization images shown in (a-c) are from the Allen Institute website (<http://www.brain-map.org>). For the number of tracer experiments and cross-validated results, see Methods.

Amygdala-projecting neurons are primarily distributed in SUBv/SUBvv layer 3, whereas amygdala inputs to SUB mostly terminated in the superficial or deep half of the molecular layer (Supplementary Fig. 9a,b; except for caudal part of the anterior basolateral amygdalar nucleus or BLAA, see below). Rabies viral tracing of the posterior basolateral amygdalar nucleus (BLAp)-projecting neurons revealed

that layer 3 SUB neurons extended thick dendritic shafts into the molecular layer that appeared to bifurcate after passing through the deep half of the molecular layer (as determined by PHAL-labeled BLAp fibers; Fig. 6b).

In the ACB, SUB projections were topographically distributed such that ProSUB and SUBv (as well as the SUBvv) pro-



jected to relatively distinct lateral, intermediate, and medial areas (Supplementary Fig. 9c). The CA1v contributed a broader terminal field that overlapped the ACB area, which was also innervated by both the SUBv and SUBvv. Axons arising from the ProSUB and SUBv traveled through the fornix and gave off collateral projections to innervate multiple hypothalamic nuclei around the fornix (perifornical band), in contrast with the SUBdd and SUBvv (medial band; Fig. 6c).

The CA1vv-SUBvv global neural networks. The CA1vv and SUBvv are part of a brain-wide network, with amygdala, hypothalamus, and basal forebrain regions that were distinct compared with the CA2-CA1dc-ProSUB and CA1i/CA1v-SUBv networks (Fig. 4d). First, the SUBvv was reciprocally connected with amygdala regions that received accessory olfactory bulb (AOB) input (the posteromedial cortical amygdala area (COApM), posterior amygdalar nucleus (PA), and the posterodorsal and posteroventral medial amygdalar nucleus (MEApd, MEApv; Supplementary Fig. 9b). Second, SUBvv axons projected to multiple hypothalamus regions along the ‘medial band’, including median preoptic nucleus (MPN), ventromedial hypothalamus (VMH), ventral premammillary nucleus (PMv), and multiple hypothalamic regions, such as the dorsomedial hypothalamic nucleus (DMH) and anterior preoptic nuclei that target neuroendocrine neurons (Fig. 6c). Third, SUBvv axon terminal fields were primarily located in the dorsomedial ACB shell, which received fewer inputs from the SUBv (Supplementary Fig. 9c). Finally, SUBvv fibers uniquely innervated the ventral lateral septum (LSv) and principal subnucleus of bed nuclei of stria terminalis (BSTpr), whereas SUBvv, SUBv, ProSUB, and CA1v fibers converged in the interfascicular and transverse BST (BSTif/tr; Supplementary Fig. 9d). Overall, the unique neuroanatomical connectivity of the SUBdd/SUBdv versus ProSUB/SUBv versus SUBvv supports our computational clustering results as parts of three separate brain-wide networks.

Global interactions of the multiple hippocampal subnetworks. Although our results indicate that hippocampal networks contribute to three relatively independent brain-wide networks, there are also examples in which hippocampal networks contribute as a whole. Previously, we described SUBdd’s relationship with the RSPv as part of the medial cortico-cortical networks¹⁸. Our new data indicate that all of the SUB subregions can influence the medial cortico-cortical network (that is, RSPv, ACAd, ACAv, and ILA) through relays in the anterior thalamus. First, although SUBdd and SUBdv projected directly to the RSPv, the AD/AV was also positioned to relay SUBdd/SUBdv information to the RSPv. In contrast, the ProSUB, SUBv, and SUBvv did not project directly to the medial cortical-cortical network, but instead projected to the dorsal and

ventral anteromedial thalamic nuclei (AMd/AMv), interanterodorsal/interanteromedial thalamic nuclei (IAD/IAM), and paraventricular (PVT) and parataenial (PT) thalamic nuclei (PVT/PT), which in turn projected to the dorsal anterior cingulate cortex (ACAd), ACAv, and ILA, respectively (Fig. 7a). Through these parallel, reciprocal thalamo-cortical and thalamo-subiculum connections, information from all of the hippocampal networks can be integrated at different nodes along the medial cortico-cortical network, which is hypothesized to transfer visual, auditory, and spatial information to the prefrontal cortex to influence attention, motor planning, and other aspect of goal-directed behavior, such as navigation and exploratory behavior^{1,18,29,30}.

Second, we found that a small cluster of neurons in the caudal BLAA (BLAA.c) were positioned as a unique hub between hippocampal subnetworks and medial prefrontal cortex, PAR, and amygdala (Fig. 7b). The distribution of CA3id/CA3ic retrograde labeling and anterogradely labeled PAR fibers uniquely identified this small cluster of BLAA.c neurons, but retrograde tracer injections in MPF, CA1, and SUB also label BLAA.c neurons (Fig. 7b). A co-injection that was targeted to the BLAA.c generated broad robust axon projections that innervated multiple regions across the hippocampal axis (Fig. 7b). Notably, BLAA.c neurons uniquely innervated the pyramidal cell layer of the CA1 and SUB directly, whereas all other amygdala regions primarily project to the molecular layer (compare with Supplementary Fig. 9a). Thus, BLAA.c neurons innervate multiple hippocampal networks and are positioned as a hub to integrate information between the medial prefrontal cortex, amygdala and hippocampus, whose interactions are critical to memory consolidation³¹.

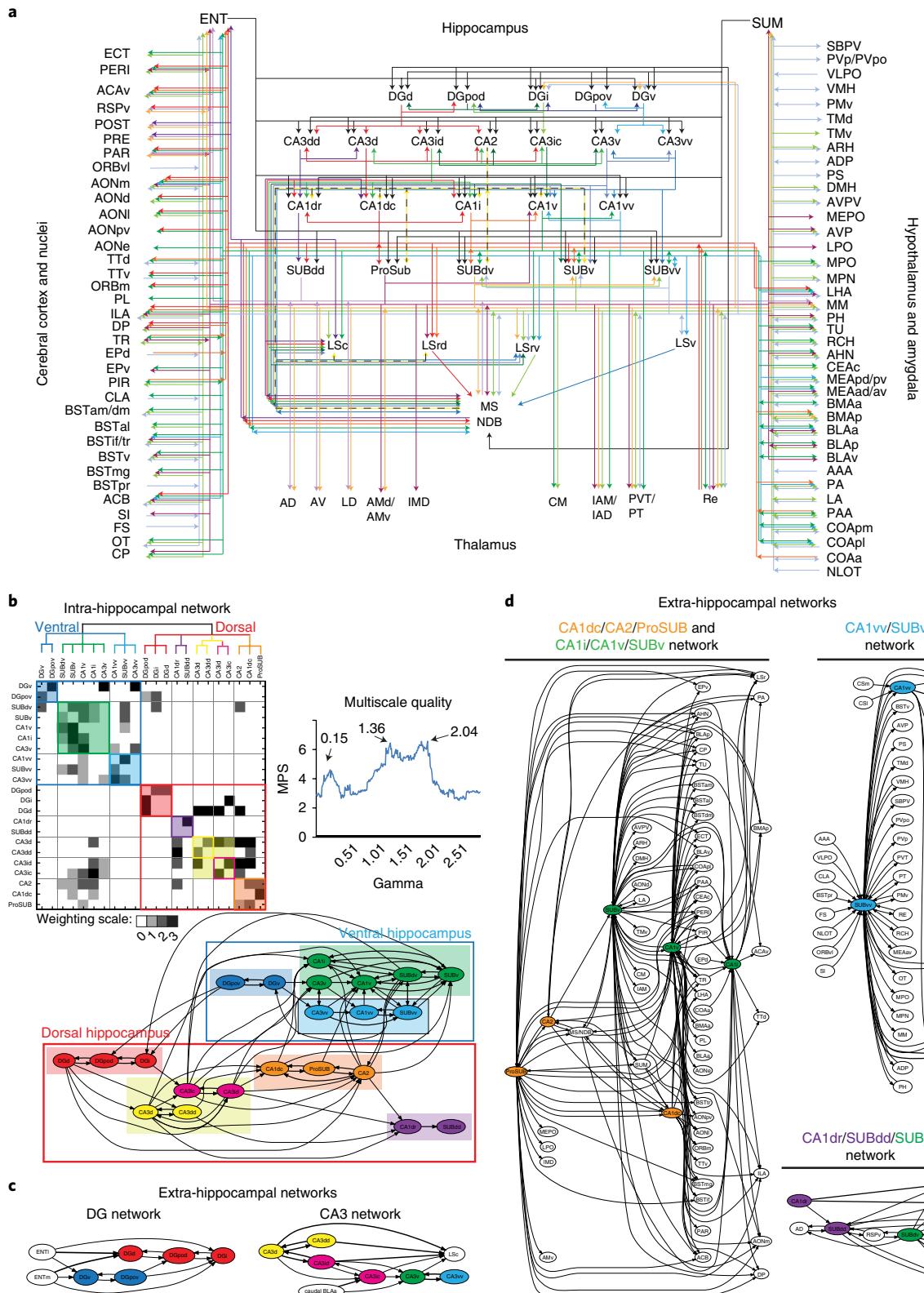
Third, the MM is the only hypothalamic region that received axonal inputs from the entire SUB in a three-dimensional topographic manner (Fig. 7c). Retrograde tracer injections into anterior versus posterior MM revealed regional and laminar specificities of MM-projecting SUB neurons. Retrograde injections into the anterior MM labeled neurons in the lateral/deep part of SUBdd layer 1 and deep layer 3 of the ProSUB, SUBv, and SUBvv (also some layer 2). In contrast, retrograde tracer injections into posterior MM labeled neurons in medial/superficial parts of SUBd layer 1 and layer 3 neurons in rostral SUBv. Injections that were limited to the ventral part of the posterior MM distinctly labeled medial SUBdv layer 1 neurons, but not SUBdd neurons. Together with the pattern of anterograde innervation, our data provide a comprehensive view of topographic SUB projection distribution across both the rostro-caudal and dorsoventral MM axis (Fig. 7c).

Fourth, it is well documented that the rat hippocampal Ammon’s horn (primarily field CA3 and CA1) generates massive descending projections to the lateral septum (LS), which in turn shares bidirectional connections with different hypothalamic neural networks

Fig. 4 | Multiscale neural network analysis of intra- and extra-hippocampal connections **a**, Unweighted connectome wiring diagram of all inputs and outputs of HGEA hippocampal subregions. Associational connections (for example, between CA3 subregions) are layered directly below each HGEA subregion group and extrinsic connections are layered further below. For annotated data, see Supplementary Table 5. Visit <http://www.mouseconnectome.org> for an interactive version of the wiring diagram. **b**, Reordered connectivity matrix and schematic diagram showing the modular hierarchical organization of intrahippocampal subnetworks as defined by current tracing data in HGEA subregions. MPS was calculated for multiple gamma values to determine which gamma values had the highest MPS peaks (0.15, 1.36, 2.04) to use for matrix reordering (multiscale quality graph). In the matrix, edges are shaded according to connectivity weight (0–3, see Supplementary Table 6) and colored boxes along the diagonal reflect modular communities at different scales. The large blue and red outlined boxes correspond to the two large dorsal and ventral hippocampus communities detected at 0.15 gamma, colored shaded boxes correspond to communities detected at 1.36 gamma, and smaller colored outlined boxes correspond to communities detected at 2.04 gamma. Matrix community color scheme corresponds to the organization of the schematic diagram below. In the schematic diagram, line weights refer to connectivity relationships at different scales. Thicker lines refer to modular connections at all scales (within community), whereas thinner lines show modular relationships only at larger scales. **c,d**, Five consensus brain-wide communities determined from multiscale community detection on annotated data in Supplementary Table 5 (**c**, DG and CA3 extrahippocampal networks; **d**, CA1/SUB extrahippocampal networks [CA1dc/CA2/ProSUB and CA1i/CA1v/SUBv network, CA1vv/SUBvv network, and CA1dr/SUBdd/SUBdv network]). Similar to the schematic shown in **b**, line weights refer to connectivity relationships at different scales. Node coloring is maintained from intrahippocampal subnetwork analysis in **b**. Different colored nodes in the extrahippocampal networks suggest that multiple intrahippocampal networks provide output to a larger brain-wide network.

mediating expression of motivated behavior³². We found that the hippocampal subnetworks had distinct hippocampo-septal, hippocampo-hypothalamic, and septo-hypothalamic projections that together formed parallel network systems (Fig. 8a). First, LSV and ventral LSr co-injections revealed distinct sets of input/output relationships between the CA1vv/SUBvv→LSv→medial hypothalamus

versus the CA1v/SUBv→LSrv→perifornical hypothalamus (as well as CA1dc/ProSUB→dorsal LSr→perifornical hypothalamus; Fig. 8a). Septo-hypothalamic projections were highly similar to the hippocampo-hypothalamic projection patterns (compare with Fig. 6c). Anterograde injections into each SUB region revealed a similar pattern: ProSUB and SUBv fibers differentially targeted a rostrocaudal



band of perifornical hypothalamic regions, whereas SUBvv fibers targeted the periventricular and medial hypothalamus rostrocaudally (note the SUBdd and SUBdv did not appreciably innervate the septum and completely avoided the hypothalamus except for MM). Multiple retrograde tracers targeted along the peri-fornical or medial hypothalamus further revealed that hypothalamus-projecting SUB neurons were richly heterogeneous across all of the gene expression layers with domain specificity (note that SUB-septal projections similarly arose from all layers; Fig. 8a). From a systems-level perspective, SUB was able to directly excite hypothalamus regions via direct glutamatergic projections or inhibit hypothalamus via recruitment of GABAergic septo-hypothalamic projections. LS projections to the MS/NDB provided a disinhibitory ‘positive’ feedback loop to all levels of the hippocampal network. In addition, SUM received robust septal input (and minor SUB input), providing an additional feedback loop to hippocampal networks downstream of hypothalamic innervation.

The SUB (or CA3 and CA1) generated sparse direct projections to the lateral hypothalamus (lateral hypothalamic area (LHA) lateral to the fornix). Instead, all of the CA3 domains (except CA3vv) and CA2 projected topographically to the caudal LSc, which in turn projected to the lateral band of hypothalamus, as revealed by multiple retrograde injections (Fig. 8a). In addition, hippocampal subnetworks innervated unique topographic ACB subregions (Supplementary Fig. 9c), which in turn generated projections to the lateral hypothalamus (Fig. 8a).

Finally, the ProSUB and SUBdv are anatomically positioned to mediate cognitive-limbic integration through their mixed input/output profiles (Fig. 8b). The ProSUB received input from CA1dc and CA2 and is part of the ‘cognitive’ dorsal hippocampal network, but provided output to ‘limbic’ brain regions (MPF, amygdala, and hypothalamus). The SUBdv appeared to have a reciprocal role in the ventral ‘limbic’ hippocampal network: SUBdv received limbic input from CA1i and CA1v, but sent its output to the visuospatial brain areas (RSPv, POST, PRE, PAR, AD, AV, and LD), which are implicated in spatial navigation. In this way, the SUBdv could influence spatial navigation based on emotionally salient information,

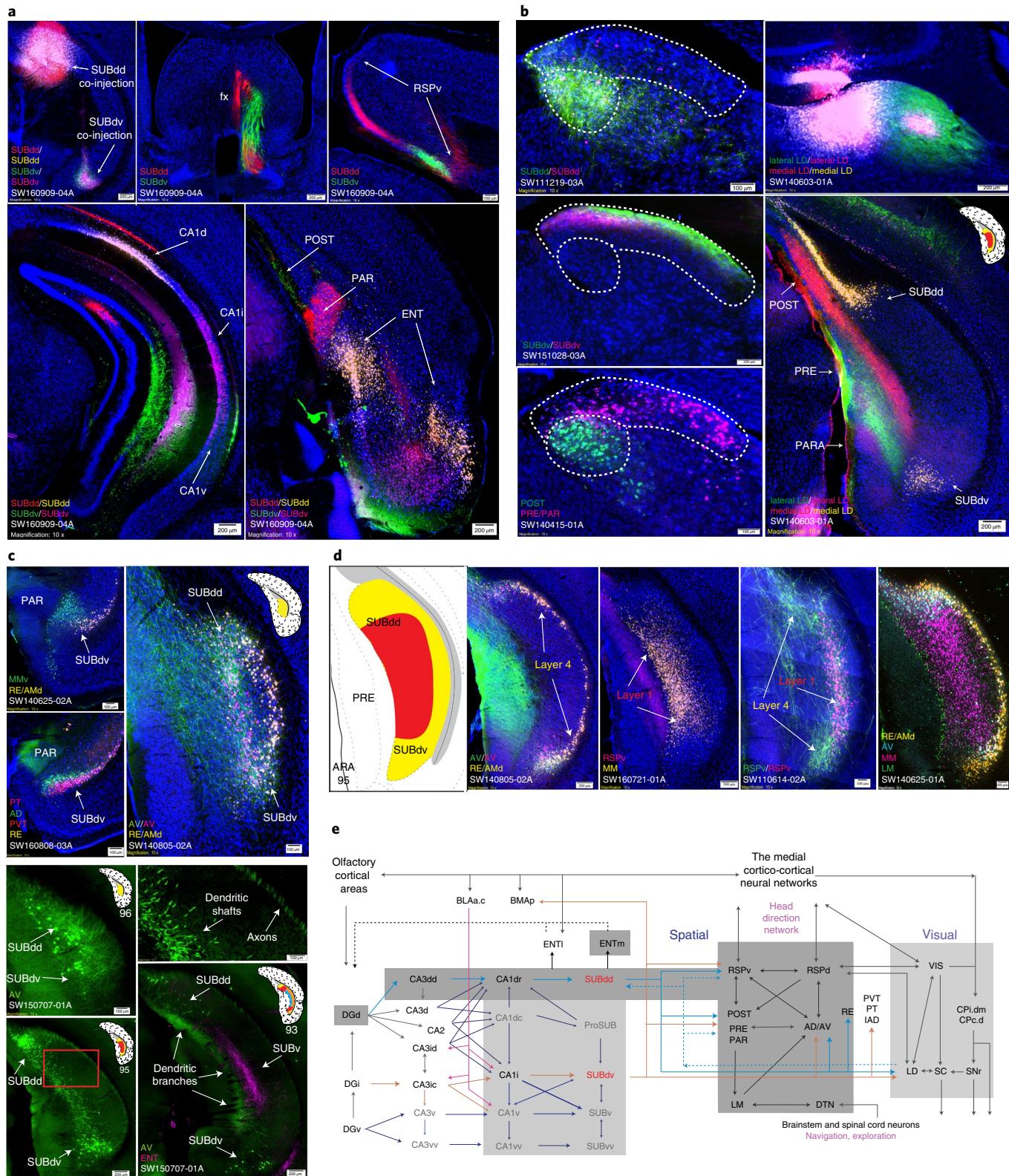
whereas the ProSUB could influence social behavior with regard to spatial location. In support of this hypothesis, two recent studies have shown that ProSUB neurons are involved in spatial working memory encoding (but not retrieval)³³, and both CA1dc and ProSUB contain neurons that encode reward location³⁴.

Discussion

The HGEA provides the first coherent anatomical and gene-expression-based framework for the rodent hippocampus (see Supplementary Table 7). Over the last decade, many studies have investigated gene expression in the hippocampus, but have yielded different interpretations of hippocampal organization. Two groups^{1,9,10} both analyzed Allen Brain Atlas data and concluded that the CA3 and CA1 can be subdivided into distinct gene expression domains. Recently, CA1 subdivisions have been further supported by multiplexed seqFISH mapping^{16,17}, but single-cell RNA approaches (that is, RNaseq and Div-Seq) have suggested a gradual transition of gene expression cell types, rather than discrete boundaries, along the CA1 axis^{12,15}. In relation to our findings, the HGEA appears markedly consistent with the CA1 subdivisions determined by seqFISH and these subdivisions are further supported by our anatomical connectivity data. An important distinction between seqFISH and single-cell RNA approaches is that seqFISH can view multiple gene expression patterns on the same histological section, whereas single-cell RNA approaches sample data from multiple sections and brains. Comparisons across multiple tissue sections introduce variations (see Supplementary Fig. 3) which could account for the gradient-like interpretation of expression. Studies which examine only one subregion (that is, CA1, but not CA3) and/or subsample from only one rostrocaudal level are limited toward understanding the bigger picture of hippocampal organization and how gene expression boundaries change across the rostrocaudal axis.

Our data are consistent with discrete HGEA gene expression domains with unique anatomical connectivity across the hippocampus. We found marked similarity between the HGEA subdivisions and anatomical labeling patterns (which were determined solely by tracer injection site size, shape, and location). Without the context

Fig. 5 | The role of the SUBdd and SUBdv in visuospatial integration and navigation **a**, Double co-injection experiments targeted to the SUBdd and SUBdv (SW160909-04A) revealed similar topographically organized anterogradely labeled projections, including, but not limited to, the fornix (middle top), RSPv (right top), and ENT (bottom right). In the hippocampus, SUBdd retrograde labeling was distributed in the CA1d, whereas SUBdv retrograde labeling was distributed adjacent within the CA1i (bottom left, highly comparable to HGEA CA1d/CA1i boundary). Both SUBdd and SUBdv projected to the POST, PRE, and PAR, but had reciprocal, complementary connectivity patterns. SUBdd projected to rostral POST and caudal PAR, whereas SUBdv projected to rostral PAR and caudal POST (caudal section shown in bottom right). In the ENT cortex, SUBdd and SUBdv received input from the lateral and intermediate bands, respectively. Anterograde and retrograde co-injection into the RSPv revealed retrograde labeling in SUBdd and SUBdv layer 1, whereas anterogradely labeled RSPv fibers primarily targeted the dorsal and ventral ends of layer 4. **b**, Anterograde and retrograde co-injections into the SUBdd (SW111219-03A) and SUBdv (SW151028-03A) produced labeling pattern distributions in the LD thalamus that were highly similar to retrograde labeling patterns from the POST versus PRE/PAR (SW140415-01A, outlined by dashed white lines). Double co-injections targeted to the medial versus lateral part of the LD thalamus produced retrograde labeling in SUBdd and SUBdv layer 4 with laminar specific anterogradely labeled projections to the POST/PRE/PAR. **c**, Organization of thalamic-projecting neurons. MM-projecting layer 1 neurons were distinct from RE/AMd-projecting layer 4 neurons in SUBdv (SW140625-02A). In the bulbs of layer 4, AD-projecting neurons (similar to LD- and AV-projecting neurons) were distributed more superficial than RE-, PVT-, and PT-projecting neurons (SW160808-03A). At the caudal end of the subiculum, the AV-projecting layer 4 neurons in the dorsal and ventral bulbs joined to become continuous along the medial part of the subiculum (SW140805-02A). Injection of G-deleted rabies virus into the AV thalamus (SW150707-01A) retrogradely labeled AV-projecting layer 4 neuron cell bodies, axons, and dendrites. AV-projecting neurons at ARA 96 sent their axons into the alveus, whereas their thick dendritic shafts extended rostrally through the SUB at ARA 95 to bifurcate into thin dendritic branches in the superficial molecular layer at ARA 93 and avoided PHAL-labeled ENT axon fibers in the deeper molecular layer. Top right, magnification of the red rectangle area in the bottom left panel. **d**, Comparison of multiple tracers revealed SUBdd and SUBdv projection cell types. In SW140805-02A, AV- and RE/AMd neurons were found to be relatively distinct projection neuron cell types in layer 4. Note that anterogradely labeled AV fibers also terminated specifically in the dorsal and ventral bulbs. In contrast, almost all of the RSPv-projecting neurons were a subset of MM-projecting neurons in layer 1 (SW160721-01A). Although RSPv-projecting SUB neurons were located in layer 1, anterogradely labeled fibers from the RSPv predominantly terminated in the layer 4 dorsal and ventral bulb regions (SW110614-02A). Finally, case SW140625-01A contained four retrograde tracer injections into the RE/AMd, AV, MM, and LM that provide a comprehensive picture of projection cell types in the SUBdd and SUBdv at ARA level 95 that is markedly similar to the HGEA. LM-projecting neurons were located primarily in PRE layer 3, MM-projecting neurons were located primarily in layer 1, and RE/AMd and AV-projecting neurons were located primarily in layer 4. **e**, Wiring schematic diagram of SUBdd and SUBdv network connections with brain regions that contribute to visuospatial behavior (see Discussion). For the number of tracer experiments and cross-validated results, see Methods.



of the gene expression boundaries to guide the injection site placement, it is unsurprising that previous anatomical studies concluded that hippocampal connectivity was organized as a topographic gradient. Our data indicate that injection sites restricted to HGEA boundaries produced labeling patterns (regional and laminar specific) that were similarly restricted to HGEA boundaries across the rostrocaudal hippocampus. This result was confirmed by antero-

grade and retrograde tracer injections into other brain regions outside of the hippocampus. In the case of the CA1 and SUB, the distribution of gene expression lamina defines subregions in the same way cortical layers define cortical subregions. In most cases, anatomical connectivity is defined by the gene expression lamina (that is, SUB layer 4 projects to thalamus); however, some connections are better described by subregion.

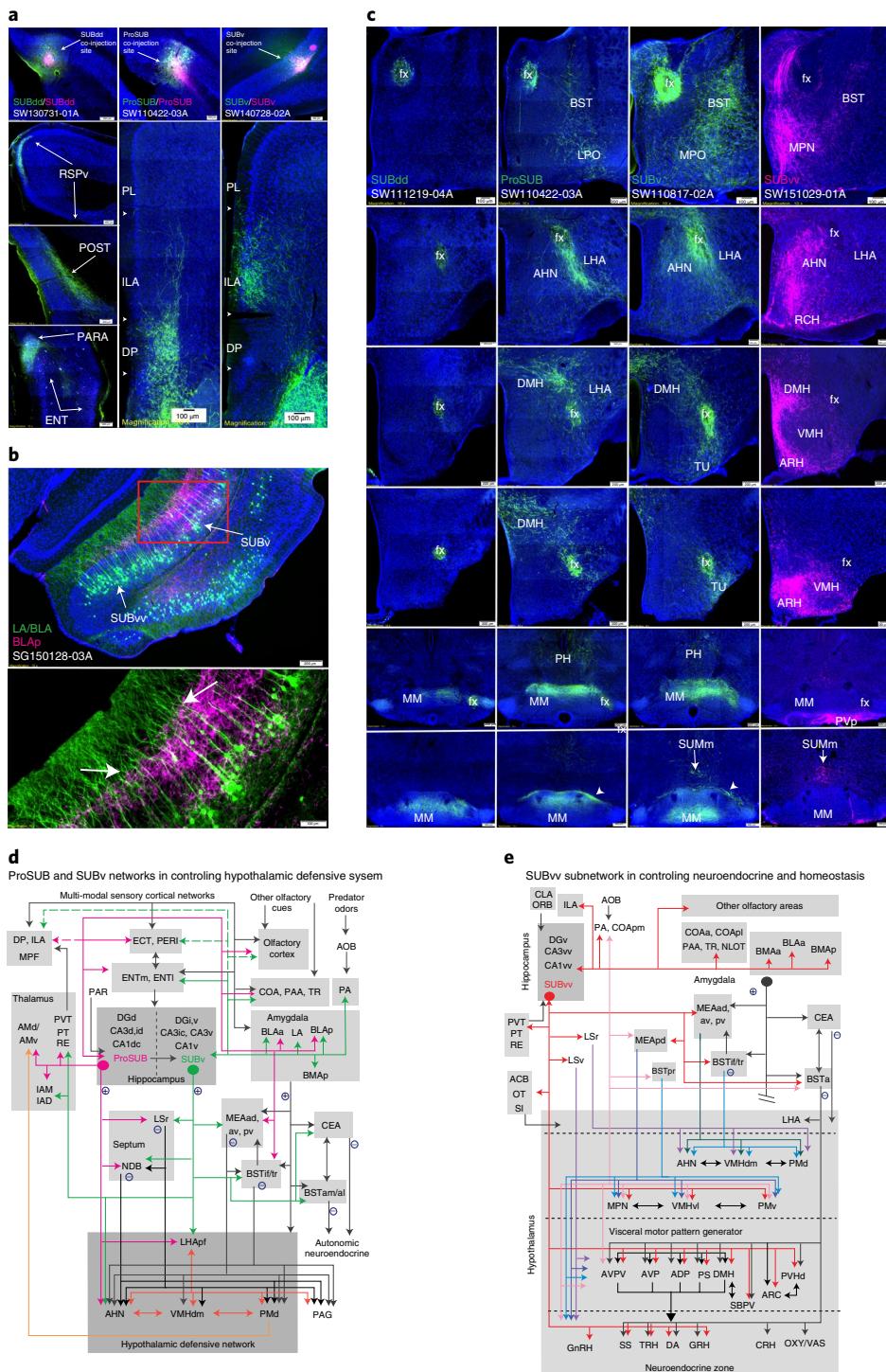


Fig. 6 | Distinct connectivity patterns of the ProSUB, SUBv, and SUBvv. **a**, SUBdd, ProSUB, and SUBv co-injections revealed that, although the SUBdd targeted the RSPv, POST/PRE/PAR, and ENT lateral band, ProSUB and SUBv neurons targeted complementary parts of the MPF. Anterogradely labeled ProSUB fibers robustly innervated the DP, with lighter input to ILA deep layers (SW110422-03A), whereas SUBv fibers densely targeted ILA superficial layers (SW140728-02A). **b**, Subiculum projections to the amygdala arose primarily from SUBv and SUBvv layer 3 (also ProSUB), whereas amygdala fibers primarily terminated in different parts of the molecular layer (Supplementary Fig. 8a,b). Injection of G-deleted rabies virus into the BLAp (SG150128-03A) revealed the dendritic branches of BLAp-projecting neurons (green) among PHA-L labeled BLAp fibers (pink). Closer examination of the amygdala-projecting dendrites revealed a thick shaft that bifurcated after passing through the BLAp terminal field (arrow). **c**, Subicular projections to the hypothalamus were organized along rostrocaudal bands (all image columns are organized rostral (top) to caudal (bottom)). SUBdd (and SUBdv) fibers bypassed most of the hypothalamus via the fornix to innervate the MM (left column). The ProSUB and SUBv (middle columns) differentially innervated multiple hypothalamic nuclei around the fornix (perifornical band), such as the anterior hypothalamic nucleus (AHN). In contrast, SUBvv fibers (right column) targeted multiple hypothalamic nuclei along the periventricular and medial hypothalamus (medial band). **d,e**, Schematic models showing ProSUB and SUBv connectivity with the hypothalamic defensive behavior network (**d**) and SUBvv connectivity with brain networks controlling metabolism, sexual behavior, and neuroendocrine function (**e**). For the number of tracer experiments and cross-validated results, see Methods.

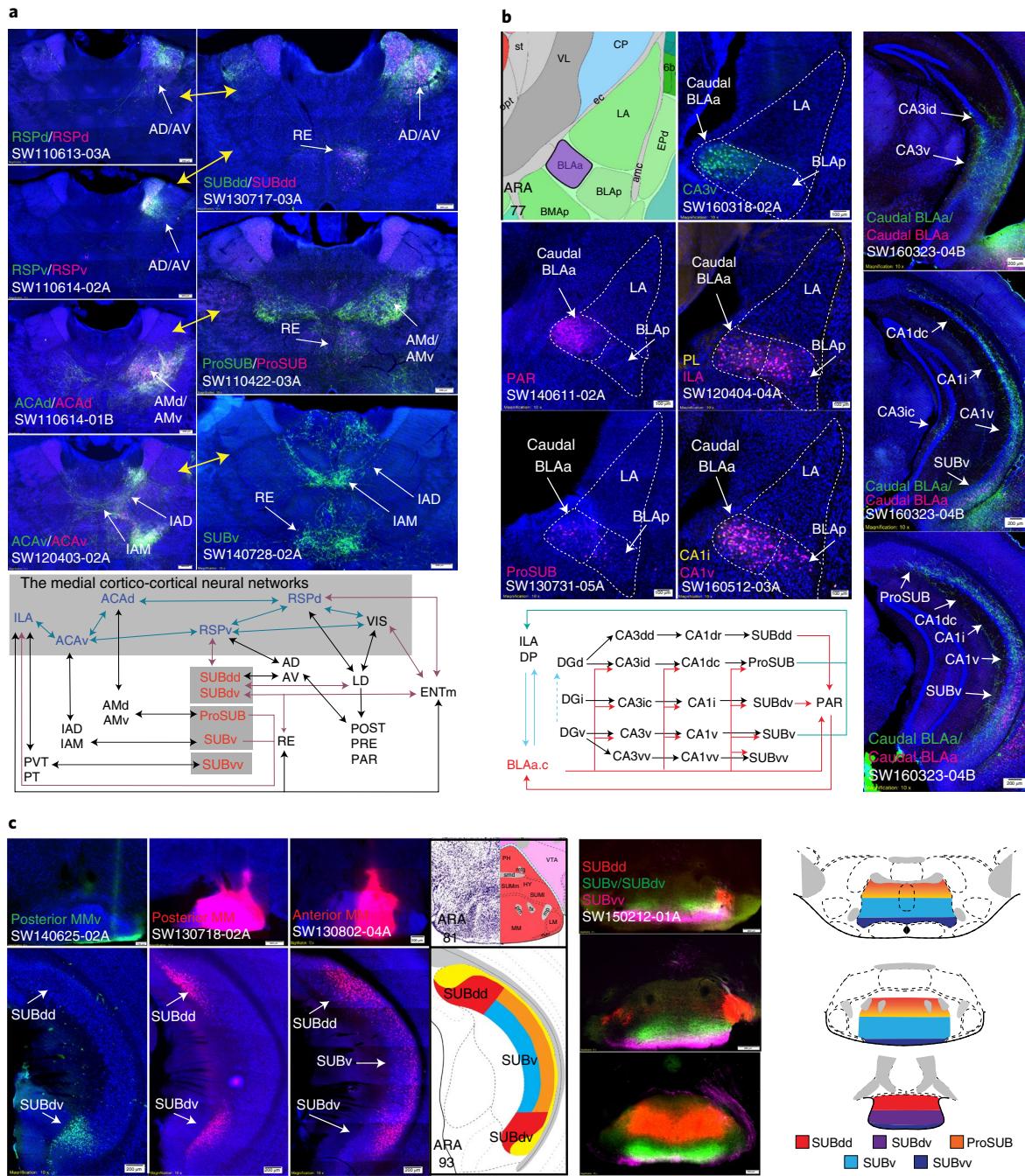


Fig. 7 | Brain systems with multiple interactions among hippocampal neural networks **a**, Subiculum regions were topographically, bidirectionally connected with the anterior thalamic nuclei (right) in a way that was highly similar to the projection patterns observed following injections in RSP and ACA cortices (left, yellow arrows refer to similar connectivity patterns). Anterior thalamic regions served as discrete relays for SUB information to modulate different levels of the medial cortico-cortical network (schematic below). **b**, Retrograde tracer injections into the CA3v specifically labeled a small dense cluster of neurons in the caudal BLAA (SW160318-02A) that was also specifically innervated by the PAR (SW140611-02A). Dashed white lines outline cytoarchitectural boundaries of the LA, caudal BLAA, and BLAp. In addition, the caudal BLAA also contained retrogradely labeled neurons following injections into many brain regions, including the ProSUB, CA1i, CA1v, PL, and ILA. Double co-injection into the caudal BLAA (SW160323-04B) revealed a broad distribution of anterogradely labeled fibers that, in contrast with other amygdala nuclei (Fig. 6b), directly targeted hippocampal neuron cell bodies in the CA3id, CA3ic, CA3v, ProSUB, CA1dc, CA1i, CA1v, and SUBv. Taken together, our data indicate that the caudal BLAA is a unique amygdalar region positioned as a hub between multiple hippocampal networks and MPF. **c**, Organization of the subiculum projections to MM. An injection isolated specifically to the ventral part of posterior MM (MMv, SW140625-02A) only labeled neurons in the SUBdv, whereas a larger injection into the posterior MM (SW130718-02A) specifically labeled neurons in both the medial/superficial part of SUBdd and SUBdv layer 1. Although posterior MM injections labeled superficial neurons in layer 1, injection into the anterior MM (SW130802-04A) retrogradely labeled neurons deeper in SUBdd/SUBdv layer 1 as well as layer 3 neurons in SUBv. Triple anterograde tracer injections into the SUBdd, SUBdv/SUBv, and SUBvv (SW150212-01A) revealed bilateral fiber terminals across the rostrocaudal MM. Overall, the entire SUB topographically projected in unique termination zones across the rostrocaudal axis of the MM (summarized in the diagram on the right). For the number of tracer experiments and cross-validated results, see Methods.

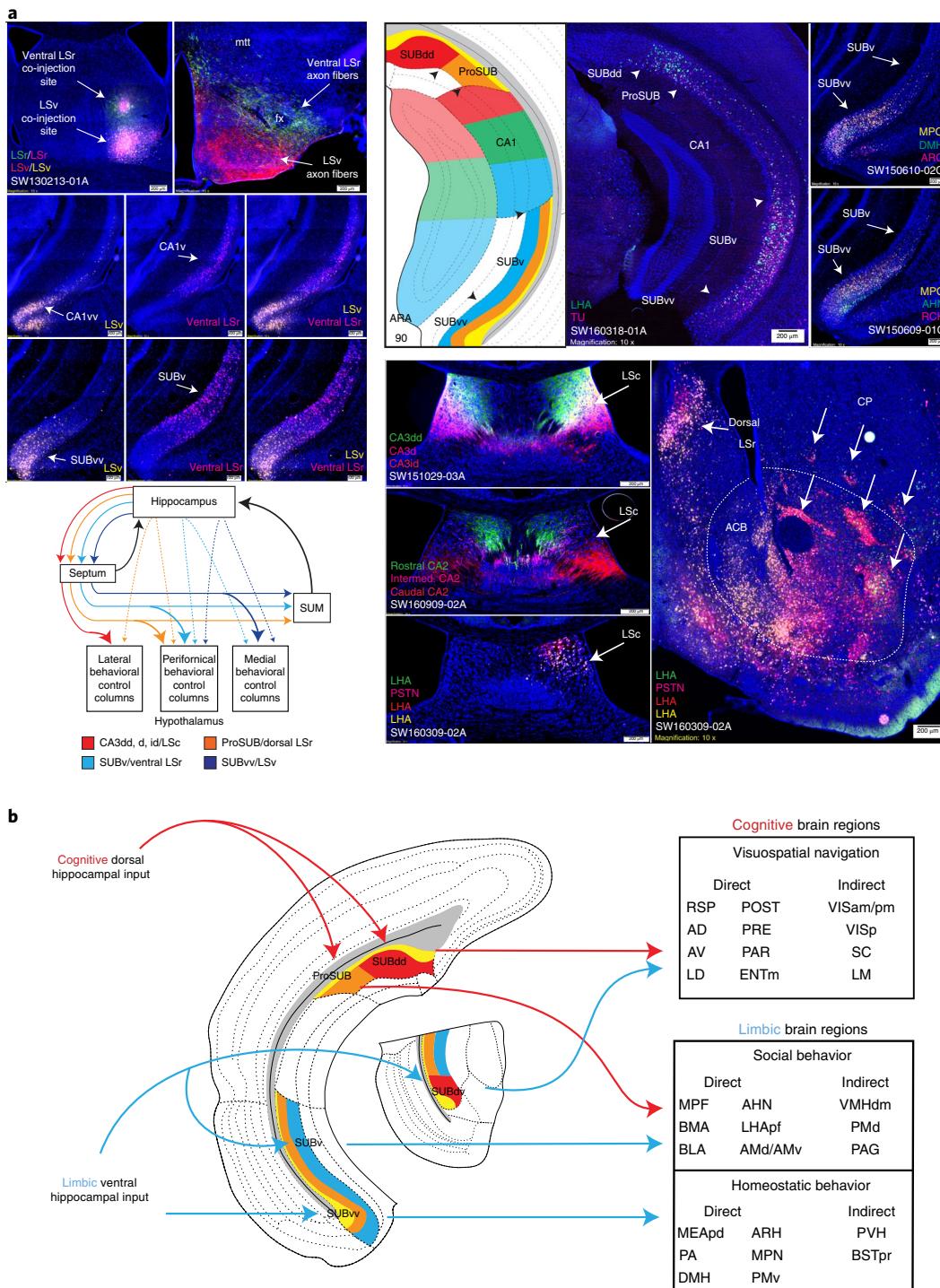


Fig. 8 | Hippocampal-septo-hypothalamic networks and cognitive-limbic integration through subiculum output pathways **a**, Hippocampal network output contributes to parallel septo-hypothalamic pathways. Double co-injection experiment into the ventral LSr and LSV produced retrograde labeling distinctly in all layers of the CA1v/SUBv and CA1vv/SUBvv, respectively, and anterograde labeling in the perifornical and medial bands of the hypothalamus (SW130213-01A; highly similar to SUB hypotalamic projection shown in Fig. 6c). Multiple retrograde tracers injected along the perifornical and medial bands produced a heterogeneous mixture of retrograde labeling across all layers of ProSUB/SUBv (SW160318-01A) and SUBvv (SW150610-02C and SW150609-01C). Multiple dorsal CA3 and CA2 subregions topographically innervated the LSc (SW151029-03A and SW160909-02A), which overlapped retrograde labeling from multiple retrograde tracer injections into the lateral hypotalamic band (lateral to perifornical band, SW160309-02A). Lateral hypotalamic band (also SI) received input from the nucleus accumbens (arrows identify striosome-like multi-labeled structures). Overall, hippocampal networks formed four parallel networks that directly and indirectly (via septum) innervated broad regions of the hypotalamus for motivated behavior. **b**, From a systems perspective, the ‘cold cognitive’ dorsal hippocampus and ‘hot affective’ ventral hippocampus ultimately provided input to distinct subiculum regions. However, the outputs of the ProSUB and SUBvv suggest pathways for cognitive-limbic crossover function in two ways. The ProSUB received cognitive visuospatial information from dorsal hippocampus and then provided output to limbic regions that are similar to the output of the SUBv. In reciprocal fashion, the SUBvv received emotionally salient limbic information from ventral hippocampus and projected to cognitive spatial navigation brain regions similar to the SUBdd. For the number of tracer experiments and cross-validated results, see Methods.

Our analysis of hippocampal connectivity revealed new insights into the hippocampus as a multiscale hierarchical network. Each HGEA subregion is part of small hippocampal subnetworks that, in turn, are constituent members of larger ‘dorsal’ and ‘ventral’ subnetworks. In this way, we can better understand how the broadly defined dorsal/ventral hippocampal dichotomy (suggested by behavioral studies) could arise from the network connectivity of a multitude of smaller, more discrete gene expression-specific and anatomically-specific subregions.

The HGEA provides a roadmap: a structural foundation for hippocampal network connectivity that implies unique functional roles for each molecular domain. Guided by the structural connectivity, functional hypotheses can be investigated by targeting and exploiting each region’s distinct gene expression. In the future, functional information can be added directly to the HGEA to provide a fully comprehensive ‘dynamic’ hippocampal connectome. On the basis of our extrahippocampal analysis and evidence from the literature, we propose testable functional hypotheses for the three CA1/SUB brain-wide networks.

First, the CA1dr/SUBdd/SUBdv network has an essential role in regulating spatial orientation and navigation. The CA1dr and SUBdd has neuronal cell types that are consistent with a role in processing and encoding spatial orientation. The CA1dr contains the highest density of place cells³⁵, whereas the SUBdd contains boundary vector cells³⁶. Furthermore, we found the SUBdd was bidirectionally connected with the caudodorsal medial entorhinal cortex (ENTm), a region that contains the highest density of the most spatially sensitive grid cells³⁷ as well as a variety of spatially modulated neuronal cell types³⁸. As we described, the SUBdd and SUBdv projected to complementary areas in the RSPv, POST, PRE, PAR, AD, AV, and RE. All of these brain structures (except for the SUBdv) have been shown to contain head-direction cells³⁹. Head-direction signals arise from vestibular input to the dorsal tegmental nucleus (DTN) and are relayed to the anterior thalamus via the LM. Together, these regions form a highly interconnected head-direction-sensitive neural network that is robustly interconnected with the CA1dr-SUBdd/SUBdv network (Fig. 5e). In addition, the RSPd and LD are hubs that suggest integration between the spatial network and the visual system (VIS cortices and SC). To produce locomotor behavior for spatial navigation, the visual and spatial networks are connected with both the dorsomedial caudoputamen (CP.i.dm)³⁸, SNr (goal-directed behavior), and SC (orientation and eye movement).

The CA2/CA1dc-ProSUB and CA1i/CA1v-SUBv intrahippocampal subnetworks are constituent parts of a much larger and complex brain-wide network that guides social behavior (particularly aggression/defensive behavior; Fig. 6d). For example, the CA2 and CA1v have both been shown to be critical for social memory and are a putative site for social memory engrams^{40,41}. In addition, the ProSUB and SUBv target complementary parts of the amygdala (that is, the BLAA and BLAp) and the perifornical hypothalamic nuclei (that is, AHN and PMd; Fig. 6b,c and Supplementary Fig. 9a,b), which are involved in fear memory and defensive behaviors^{42–45}. The PMd generates dense projections back to the AMy, which is involved with processing contextual fear memories to predatory threats^{46,47} and is bidirectionally connected with the ProSUB.

Finally, the most distinguishing features of the CA1vv-SUBvv network are robust projections that directly innervate neuroendocrine-related medial hypothalamic areas and other forebrain structures to regulate homeostasis, metabolism, and sexual behavior. In contrast with the ProSUB and SUBv, SUBvv axons robustly innervated hypothalamic nuclei along a medial rostrocaudal band, including the periventricular zone, which contains multiple neuroendocrine cell types (Fig. 6c)⁴⁸. Furthermore, the SUBvv innervated the anteromedial BST nucleus (BSTam; Supplementary Fig. 9d) and several densely interconnected hypothalamic nuclei that target the

PVH, namely the anterodorsal preoptic (ADP), anteroventral preoptic (AVP), anteroventral periventricular (AVPV), parastrial (PS) and DMH (Fig. 6c). These interconnected hypothalamic structures have been hypothesized to be a hypothalamic visceral motor pattern generator network that coordinates neuroendocrine and autonomic activities in maintaining homeostasis⁴⁹. In addition, the SUBvv was robustly connected to the paraventricular (PVT) thalamic nucleus, which receives hypothalamic inputs and viscerosensory and pain-related information from brainstem structures (that is, nucleus solitarius tract and parabrachial nucleus)⁵⁰.

Hypothesis for cognitive-limbic integration. Although the SUBdd and SUBv/SUBvv generally fit the classic view of a distinct dorsal/ventral functional dichotomy, the ProSUB and SUBdv provide anatomical substrates for cognitive and limbic crossover function through their extrahippocampal output (Fig. 8b). The ProSUB is a region that receives input from cognitive-related dorsal hippocampal circuits, but has extrinsic projections to limbic areas (similar to SUBv). Conversely, the SUBdv receives input from ventral hippocampal circuits, but extrinsically projects to cognitive areas in the visuospatial network (similar to SUBdd). This suggests the ProSUB may use spatial information to influence hypothalamic-mediated social behaviors, whereas the SUBdv may use emotionally salient information for environmental navigation via its projections to RSPv, POST/PRE/PAR, AD, AV, etc.

Conclusion

In conclusion, we have shown that the hippocampus can be divided into multiple subregions with distinct combinatorial gene expression patterns and connectivity. The multiple hippocampal subregions are wired together as a multiscale hierarchical network that provides output to three brain-wide networks related to spatial navigation, social, and homeostatic and reproductive behaviors. Ultimately, the HGEA provides an anatomical and gene expression roadmap for future functional dissection of the hippocampus.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at <https://doi.org/10.1038/s41593-018-0241-y>.

Received: 20 September 2017; Accepted: 17 July 2018;

Published online: 8 October 2018

References

1. Fanselow, M. S. & Dong, H.-W. Are the dorsal and ventral hippocampus functionally distinct structures? *Neuron* **65**, 7–19 (2010).
2. Maruszak, A. & Thuret, S. Why looking at the whole hippocampus is not enough—a critical role for anteroposterior axis, subfield and activation analyses to enhance predictive value of hippocampal changes for Alzheimer’s disease diagnosis. *Front. Cell. Neurosci.* **8**, 95 (2014).
3. Moser, M. B. & Moser, E. I. Functional differentiation in the hippocampus. *Hippocampus* **8**, 608–619 (1998).
4. Poppenk, J., Evensmoen, H. R., Moscovitch, M. & Nadel, L. Long-axis specialization of the human hippocampus. *Trends Cogn. Sci.* **17**, 230–240 (2013).
5. Risold, P. Y. & Swanson, L. W. Structural evidence for functional domains in the rat hippocampus. *Science* **272**, 1484–1486 (1996).
6. Swanson, L. W. & Cowan, W. M. An autoradiographic study of the organization of the efferent connections of the hippocampal formation in the rat. *J. Comp. Neurol.* **172**, 49–84 (1977).
7. Strange, B. A., Witter, M. P., Lein, E. S. & Moser, E. I. Functional organization of the hippocampal longitudinal axis. *Nat. Rev. Neurosci.* **15**, 655–669 (2014).
8. Amaral, D. G. Emerging principles of intrinsic hippocampal organization. *Curr. Opin. Neurobiol.* **3**, 225–229 (1993).
9. Dong, H.-W., Swanson, L. W., Chen, L., Fanselow, M. S. & Toga, A. W. Genomic-anatomic evidence for distinct functional domains in hippocampal field CA1. *Proc. Natl. Acad. Sci. USA* **106**, 11794–11799 (2009).

10. Thompson, C. L. et al. Genomic anatomy of the hippocampus. *Neuron* **60**, 1010–1021 (2008).
11. Zeisel, A. et al. Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science* **347**, 1138–1142 (2015).
12. Cembrowski, M. S., Wang, L., Sugino, K., Shields, B. C. & Spruston, N. HippoSeq: a comprehensive RNA-seq database of gene expression in hippocampal principal neurons. *eLife* **5**, e14997 (2016).
13. Habib, N. et al. Div-Seq: Single-nucleus RNA-Seq reveals dynamics of rare adult newborn neurons. *Science* **353**, 925–928 (2016).
14. Lein, E. S., Callaway, E. M., Albright, T. D. & Gage, F. H. Redefining the boundaries of the hippocampal CA2 subfield in the mouse using gene expression and 3-dimensional reconstruction. *J. Comp. Neurol.* **485**, 1–10 (2005).
15. Cembrowski, M. S. & Spruston, N. Integrating results across methodologies is essential for producing robust neuronal taxonomies. *Neuron* **94**, 747–751.e1 (2017).
16. Shah, S., Lubeck, E., Zhou, W. & Cai, L. In situ transcription profiling of single cells reveals spatial organization of cells in the mouse hippocampus. *Neuron* **92**, 342–357 (2016).
17. Shah, S., Lubeck, E., Zhou, W. & Cai, L. seqFISH accurately detects transcripts in single cells and reveals robust spatial organization in the hippocampus. *Neuron* **94**, 752–758.e1 (2017).
18. Zingg, B. et al. Neural networks of the mouse neocortex. *Cell* **156**, 1096–1111 (2014).
19. Oh, S. W. et al. A mesoscale connectome of the mouse brain. *Nature* **508**, 207–214 (2014).
20. Dong, H. W. *The Allen Reference Atlas: A Digital Color Brain Atlas of the C57BL/6J Male Mouse*. (John Wiley & Sons Inc, Hoboken, NJ, 2008).
21. Saunders, A. et al. Molecular diversity and specializations among the cells of the adult mouse brain. *Cell* **174**, 1015–1030 (2018).
22. Scharfman, H. E. The enigmatic mossy cell of the dentate gyrus. *Nat. Rev. Neurosci.* **17**, 562–575 (2016).
23. Lorente de Nó, R. Studies on the structure of the cerebral cortex. II. Continuation of the study of the ammonic system. *J. Psychol. Neurol.* **46**, 113–177 (1934).
24. Cui, Z., Gerfen, C. R. & Young, W. S. III Hypothalamic and other connections with dorsal CA2 area of the mouse hippocampus. *J. Comp. Neurol.* **521**, 1844–1866 (2013).
25. Dudek, S. M., Alexander, G. M. & Farris, S. Rediscovering area CA2: unique properties and functions. *Nat. Rev. Neurosci.* **17**, 89–102 (2016).
26. Kohara, K. et al. Cell type-specific genetic and optogenetic tools reveal hippocampal CA2 circuits. *Nat. Neurosci.* **17**, 269–279 (2014).
27. Sik, A., Penittonen, M., Ylinen, A. & Buzsáki, G. Hippocampal CA1 interneurons: an *in vivo* intracellular labeling study. *J. Neurosci.* **15**, 6651–6665 (1995).
28. Ding, S. L. Comparative anatomy of the prosubiculum, subiculum, presubiculum, postsubiculum, and parasubiculum in human, monkey, and rodent. *J. Comp. Neurol.* **521**, 4145–4162 (2013).
29. Vann, S. D., Aggleton, J. P. & Maguire, E. A. What does the retrosplenial cortex do? *Nat. Rev. Neurosci.* **10**, 792–802 (2009).
30. Weible, A. P. Remembering to attend: the anterior cingulate cortex and remote memory. *Behav. Brain Res.* **245**, 63–75 (2013).
31. Kitamura, T. et al. Engrams and circuits crucial for systems consolidation of a memory. *Science* **356**, 73–78 (2017).
32. Risold, P. Y. & Swanson, L. W. Connections of the rat lateral septal complex. *Brain Res. Brain Res. Rev.* **24**, 115–195 (1997).
33. Cembrowski, M. S. et al. Dissociable structural and functional hippocampal outputs via distinct subiculum cell classes. *Cell* **173**, 1280–1292.e18 (2018).
34. Gauthier, J. L. & Tank, D. W. A dedicated population for reward coding in the hippocampus. *Neuron* **99**, 179–193.e7 (2018).
35. O'Keefe, J. & Dostrovsky, J. The hippocampus as a spatial map: preliminary evidence from unit activity in the freely-moving rat. *Brain Res.* **34**, 171–175 (1971).
36. Lever, C., Burton, S., Jeewajee, A., O'Keefe, J. & Burgess, N. Boundary vector cells in the subiculum of the hippocampal formation. *J. Neurosci.* **29**, 9771–9777 (2009).
37. Hafting, T., Fyhn, M., Molden, S., Moser, M.-B. & Moser, E. I. Microstructure of a spatial map in the entorhinal cortex. *Nature* **436**, 801–806 (2005).
38. Sargolini, F. et al. Conjunctive representation of position, direction, and velocity in entorhinal cortex. *Science* **312**, 758–762 (2006).
39. Taube, J. S. The head direction signal: origins and sensory-motor integration. *Annu. Rev. Neurosci.* **30**, 181–207 (2007).
40. Okuyama, T., Kitamura, T., Roy, D. S., Itohara, S. & Tonegawa, S. Ventral CA1 neurons store social memory. *Science* **353**, 1536–1541 (2016).
41. Hitti, F. L. & Siegelbaum, S. A. The hippocampal CA2 region is essential for social memory. *Nature* **508**, 88–92 (2014).
42. Fanselow, M. S. & Poulos, A. M. The neuroscience of mammalian associative learning. *Annu. Rev. Psychol.* **56**, 207–234 (2005).
43. Totove, P., Fadok, J. P. & Lüthi, A. Neuronal circuits for fear and anxiety. *Nat. Rev. Neurosci.* **16**, 317–331 (2015).
44. Swanson, L. W. Cerebral hemisphere regulation of motivated behavior. *Brain Res.* **886**, 113–164 (2000).
45. Canteras, N. S., Pavesi, E. & Carobrez, A. P. Olfactory instruction for fear: neural system analysis. *Front. Neurosci.* **9**, 276 (2015).
46. Gross, C. T. & Canteras, N. S. The many paths to fear. *Nat. Rev. Neurosci.* **13**, 651–658 (2012).
47. de Lima, M. A. X., Baldo, M. V. C. & Canteras, N. S. A role for the anteromedial thalamic nucleus in the acquisition of contextual fear memory to predatory threats. *Brain Struct. Funct.* **222**, 113–129 (2017).
48. Biag, J. et al. Cyto- and chemoarchitecture of the hypothalamic paraventricular nucleus in the C57BL/6J male mouse: study of immunostaining and multiple fluorescent tract tracing. *J. Comp. Neurol.* **520**, 6–33 (2012).
49. Thompson, R. H. & Swanson, L. W. Structural characterization of a hypothalamic visceromotor pattern generator network. *Brain Res. Brain Res. Rev.* **41**, 153–202 (2003).
50. Li, S. & Kirovac, G. J. Sources of inputs to the anterior and posterior aspects of the paraventricular nucleus of the thalamus. *Brain Struct. Funct.* **217**, 257–273 (2012).

Acknowledgements

This manuscript is dedicated to the memory of Sigmund J. Bienkowski. The authors would like to thank A.W. Toga, L.W. Swanson, and H. Karten for advising the Mouse Connectome Project. This work was supported by NIH/NIMH R01 MH094360-01A1 (H.-W.D), NIH/NCI U01 CA198932-01 (H.-W.D), NIH/NIMH RF1 MH114112-01 (H.-W.D), NIH/NIMH U01 MH114829-01 (H.-W.D), NIH/NIMH F32 MH107071-01-A1 (M.S.B), and a Pilot fund (H.-W.D.) of ADRC center grant P50-AG05142.

Author contributions

M.S.B. and H.-W.D. conceived, designed and managed the project. M.S.B. and H.-W.D. wrote the manuscript. M.S.B. performed manual analysis of all raw image data, including the gene expression and connectivity annotation, created the HGEA, and prepared figures for publication. I.B. led the informatics team and wrote the code for computational network analysis. M.S.B., I.B., and H.-W.D. constructed the neural networks diagrams. M.S.B., M.Y.S., L.G., M.Z., and N.L.B. performed stereotaxic surgeries to generate anatomical connectivity data. K.C. and M.S.B. performed integrative analysis of HGEA annotation with DropViz data. T.A. created the three-dimensional HGEA atlas rendering along with online three-dimensional atlas viewer and supplementary videos. S.Y. managed the iConnectome website and created online informatics and visualization tools. J.A.-J., S.A., and D.L. performed image processing for image data upload to the iConnectome viewer. N.L.B., N.N.F., and H.H. offered constructive guidance for the manuscript edits.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41593-018-0241-y>.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to M.S.B. or H.-W.D.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2018

Methods

Allen Gene Expression Atlas annotation, analysis, and mapping. The Allen Institute provides an online open-access resource of thousands of mouse *in situ* hybridization experiments each visualizing gene expression patterns of a single gene (<http://www.brain-map.org>), methodology can be found online in the Documentation section¹. The archive of mouse gene expression patterns can be searched either directly by name or by gross anatomical structure as listed in the ARA²⁰. The Allen Gene Expression Atlas (AGEA) allows for a point-to-point correlation comparison of two anatomical locations in the mouse brain. To discover gene expression patterns within the hippocampal formation, we applied all of these methods to manually search for genes that were located within the major cell layers of the mouse hippocampus and subiculum. As described previously¹⁰, gene expression patterns were observed to be a nested mosaic, such that one gene could be expressed in an entire structure while another gene's expression could be limited to various parts of that structure¹⁰. To identify hippocampal subregions, we examined combinatorial gene expression patterns of hundreds of hippocampal-expressed genes that displayed restricted expression patterns until we could determine the least common pattern that defines a unique subregion. For some areas, a single gene clearly demarcates a subregion whereas other subregions are interpreted based on the presence or absence of multiple genes. Overall, we manually annotated the expression pattern of over 250 genes across the entire hippocampal formation (Supplementary Table 2).

A major consideration when observing and annotating Allen Brain Atlas gene expression data is the highly varied histological tissue sectioning. Different sectioning angles create inaccuracies when comparing individual sections across animals and matching and registering whole tissue sections directly on to ARA levels. Point-to-point comparisons across different brains can be inaccurate as one part of the tissue can align well while another part is mismatched. For the hippocampus, this problem is particularly relevant to the dorsal/ventral axis and may be the reason for previous observations of many small subdomains in the ventral CA3 (ref. ¹⁰) and ventral CA1 (ref. ⁹) as well as the interpretation of gene expression gradients. To accurately map rostrocaudal HGEA boundaries, we used local tissue landmarks at multiple points throughout the tissue section to estimate the sectioning angle, treat each section as a composite gradient of multiple rostrocaudal ARA levels, and appropriately match the data to the HGEA (Supplementary Fig. 3a). Notably, tissue sections from the same brain are cut at the same angle while tissue sections from other brains are often cut at different angles (Supplementary Fig. 3b,c). Oblique sectioning angles can also create both medial/lateral and dorsal/ventral gradients.

Final determinations of HGEA boundaries are accomplished by examining the positioning of gene expression across multiple tissue sections and interpreting the changes that occur between. For example, the caudal boundary of the CA2 with the CA1i has been of recent debate¹⁵. The rostrocaudal progression of the CA2 'marker gene' *Amigo2* can be observed in both the coronal and sagittal sections (Supplementary Fig. 3e,f). At level HGEA 80, dense *Amigo2* expression is located within the CA2, with lighter expression within the CA3v and CA3vv (Supplementary Fig. 3e). At HGEA 81, the dense *Amigo2* CA2 expression is now present in a large continuous area between the CA1d and CA1v (caudal to the CA3 in HGEA 80). Finally, at HGEA 82, *Amigo2* CA2 expression is limited to a sparse lamina of neurons deep to the CA1i. A similar progression from medial to lateral can be observed in the sagittal tissue sections (Supplementary Fig. 3f). In sagittal HGEA 4, *Amigo2* expression is present in dorsal and ventral parts of the hippocampus separated by CA3. At sagittal HGEA 2, *Amigo2* CA2 expression becomes continuous between the CA1d and CA1v. Finally, at sagittal HGEA 1, *Amigo2* expression is limited to a superficial lamina with CA1i located more deeply (CA2 rostral to CA1i). Considered together, the coronal and sagittal *Amigo2* expression show that the caudal extension of the CA2 is interposed between the CA1d and CA1v in the dorsoventral direction and the CA3 and CA1i in the rostrocaudal direction. The boundary between the CA2 and CA1i is non-parallel to the coronal sectioning plane resulting in CA2 neurons located deep to CA1i neurons at HGEA 81. This progression is similar across all CA2-expressed genes.

Comparison of HGEA gene expression annotation with single-cell RNAseq database. To relate our qualitative gene expression annotation approach to more quantitative single-cell RNAseq methods, we compared the binary HGEA gene expression annotation to the Dropviz drop-seq database²¹ (<http://www.dropviz.org>). The Dropviz database contains single-cell gene expression from 113,000 hippocampal cells and over 32,000 genes and clustering analysis reveals that the hippocampal neurons can be divided into seven global clusters (Interneuron_Gad2 versus Cajal-Retzius versus Dentate_C1ql2 versus CA3CA3_Pvrl3-Rgs15-Calb2 versus CA1_Subiculum_Postsubiculum_Entorhinal_Fibcd1-Dcn-Cbln1-Ptgfr-Fezf2 versus Subiculum_Entorhinal_Nxph3 versus Subiculum_Slc17a6) that contains multiple 'subcluster' cell-types (27 interneurons and 31 principal neurons, note the subiculum neurons are divided among three different clusters). DropViz data distribution was assumed to be normal, but this was not formally tested.

First, to compare the relative correspondence of our HGEA gene expression annotation to the presence of positive gene expression within a hippocampal region, we used the DropViz query feature to input HGEA annotated genes for

each region and compared their expression within a corresponding DropViz global cluster (that is, DGd-expressed genes within Dentate_C1ql2, CA3dd-expressed genes within CA3CA3_Pvrl3-Rgs15-Calb2; see the "DropViz Comparative Analysis" Excel spreadsheet hosted at <http://www.mouseconnectome.org/MCP/page/papers>). Comparisons for each HGEA SUB layer were performed for each of the 3 global SUB-related global clusters. In addition, we compared the significance of this expression within a global cluster versus the rest of the hippocampus, although only the 'marker genes' we identified would be expected to be different (all data and comparisons exported from DropViz website). For each set of annotated genes within a HGEA region, most of the genes were found to have positive expression within the DropViz global cluster. Notably, many of the same 'marker genes' identified in our annotation as unique to a single HGEA region had significant positive expression within the global cluster and 0 expression in the rest of the hippocampus (highlighted green in online Excel spreadsheet). HGEA genes that were annotated as having positive expression, but showed 0 expression within the DropViz dataset, were queried for their expression within the Interneuron_Gad2 global cluster. Genes which were found to be expressed within other relevant global clusters that could account for positive expression within an HGEA region (Interneuron_Gad2 or one of the other SUB-containing global clusters) were highlighted orange. All other genes which were annotated as having expression in the HGEA annotation, but 0 expression in any relevant global clusters were highlighted red (note, much of the discrepancy comes from repetition of a few genes that DropViz reported as not expressed anywhere in the hippocampus, but were annotated as expressed in multiple HGEA regions). In total, the number and proportion of genes with positive expression (non-red highlighted) and 0 expression (red highlighted) were calculated for each comparison to evaluate the relative similarity at the global cluster scale (for each SUB layer, the three comparisons were averaged to create one percentage value). Overall, the percentage of similar positive expression between each HGEA region versus DropViz global cluster averaged $91.7 \pm 0.8\%$.

To further analyze similarities between the HGEA annotation and DropViz data, we performed a PCA using DropViz's meta-cell data. According to DropViz's website, a 'meta-cell' contains the aggregate UMI counts for all single cells that belong to a subcluster so that there is one meta-cell per subcluster. Using the UMI count data from DropViz, we performed two separate PCA clustering analyses on the meta-cells using all 32,307 DropViz gene set and compared it to the more limited set of 248 HGEA annotated genes. We found that PCA clustering from the 248 HGEA genes performed remarkably similar to the clustering with all 32,307 DropViz genes, suggesting that the more limited HGEA annotated gene set can still recapitulate the overall differences in hippocampal cell type gene expression (see Supplementary Fig. 5).

Finally, we reasoned that the annotated gene expression patterns that define HGEA subregions may reflect the distinct gene expression profiles of the individual meta-cells if each HGEA region is also composed of different cell-types. To compare the Dropviz meta-cell data to the binary HGEA gene annotation, we first binarized the UMI count data for the HGEA genes within the DropViz meta-cells after thresholding out UMI counts less than 30. Then, we calculated a Rogers-Tanimoto dissimilarity matrix (SciPy Python library) for the HGEA annotated genes between where they were expressed (HGEA subregion) and which meta-cell they were expressed in (Supplementary Table 4; all coefficient values between 0 and 1, lower value means more similar and vice versa). As expected, the dissimilarity matrix showed that meta-cells were more similar to corresponding HGEA subregion and more dissimilar to other HGEA regions (for example, CA1 meta-cells had lower dissimilarity values in HGEA CA1 regions, etc.). Notably, our annotated 'Putative Interneurons' were highly similar to the DropViz interneuron meta-cells. Some HGEA subregions were notably similar to a specific meta-cell. For example, the Neuron.Slc17a7.Calb2-Vgll3 meta-cell had a low dissimilarity value for the CA3vv region. Consistent with this similarity, DropViz query feature reports the CA3vv 'marker gene' *Coch* as being significantly expressed in the Neuron.Slc17a7.Calb2-Vgll3 subcluster ($2.77, P = 7.44 \times 10^{-192}$) and almost 0 expression in all other CA3 subclusters.

Mouse Connectome Project methodology. Anatomical tracer data was generated as part of the Mouse Connectome Project (MCP) within the Center for Integrative Connectomics (CIC) at the University of Southern California Mark and Mary Stevens Neuroimaging and Informatics Institute (formerly the Laboratory of Neuro Imaging at the University of California, Los Angeles). Some MCP experimental procedures for data generation and online publication have been described previously^{18,52,53}.

We systematically and carefully mapped neuronal connectivity of every molecular domain of the hippocampus and SUB to determine their connectivity (for injection site list, see Supplementary Table 3). We used multiple fluorescent tracing strategies with a combination of classic tract-tracing and viral tracing methods. First, we used a double co-injection approach that injects two different tracer cocktails each containing one anterograde and one retrograde tracer to simultaneously visualize two sets of input/output connectivity¹⁸. To investigate the convergence or divergence of axonal fiber pathways either into or out of the hippocampus, we used a triple anterograde tracing approach with individual injections of PHAL and EGFP- and tdTomato- expressing adeno-associated viruses

(AAV). Finally, to compare neuronal projection cell types and fiber pathway origin, we used a quadruple retrograde tracing method with individual injections of four different retrograde tracers (cholera toxin subunit B (CTb) conjugated with 488, 555, or 647, as well as FG) into four different hippocampal projection targets.

Subjects. All tracer experimental data was generated using 2–6-month-old male C57BL/6J mice (Jackson Laboratories). Mice were pair-housed within a room that was controlled for temperature (21–22 °C), humidity (51%), and light (12 h light:12 h dark cycle with lights on at 6:00 a.m. and off at 6:00 p.m.). Subjects had ad libitum access to tap water and mouse chow throughout the experiments. Rabies injection surgeries were performed in a BSL-2 level environment and performed by individuals who had been rabies-vaccinated. Following surgery, rabies-infected animals were individually housed in a separate BSL-2 level facility. All experiments were conducted according to the regulatory standards set by the National Institutes of Health Guide for the Care and Use of Laboratory Animals and by the institutional guidelines set by the Institutional Animal Care and Use Committee at USC and the Animal Research Committee at UCLA. This study's protocol was approved by the Institutional Animal Care and Use Committee at USC and the Animal Research Committee at UCLA.

Tracer injection experiments. The Mouse Connectome Project's standard experimental approach is dual co-injections of anterograde and retrograde tracers into different brain areas within the same mouse. Each co-injection contained an anterograde (*phaseolus vulgaris* leucoagglutinin [PHAL] or biotinylated dextran amine [BDA]) and a retrograde (cholera toxin subunit b [CTb] or Fluorogold [FG]) tracer. PHAL (2.5%; Vector Laboratories) and CTb conjugated to Alexa Fluor 647 (CTB-647, 0.25%; Invitrogen) were co-injected, whereas BDA (biotinylated dextran amine, 5%; Invitrogen) or AAV1.CAG.RFP (Penn Vector Core, originally created at Allen Institute for Brain Sciences) was injected in combination with FG (1%; Fluorochrome, LLC).

To provide further details on specific connectivity patterns, we also performed quadruple retrograde tracer, triple anterograde tracer, and rabies/PHAL experiments. Quadruple retrograde tracer experiments involved four different injection sites receiving a unique injection of either 0.25% CTB-647, CTB conjugated to Alexa Fluor 555 (CTB-555, 0.25%; Invitrogen), CTB conjugated to AlexaFluor 488 (CTB-488, 0.25%; Invitrogen), or 1% FG. Triple anterograde tracing experiments involved three separate injections of 2.5% PHAL, AAV1.CAG.RFP, and AAV1.hSyn.GFP (Penn Vector Core, originally created at Allen Institute for Brain Sciences). For rabies/PHAL experiments, we used G-deleted rabies-GFP (Salk Institute vector core) at the same or separate injection site as PHAL. G-deleted rabies virion constructs are pseudotyped with the normal rabies glycoprotein (G)-containing viral envelope but are unable to incorporate rabies glycoprotein in new virions, working similar to classic retrograde tracers. However, although most retrograde tracers only label cell bodies, G-deleted rabies-GFP functions as a retrograde tracer that produces bright fluorescent labeling of both the cell body and dendrites. Together with PHAL axonal labeling, G-deleted rabies is a useful tool for examining anterogradely-labeled axons and their approximate location adjacent to retrogradely-labeled dendrites and cell bodies. All cases used in this study are listed in Supplementary Table 3 and are available online at www.MouseConnectome.org. No statistical methods were used to pre-determine sample sizes but our sample sizes are similar to those reported in previous publications^{18,52}. In most cases, anterograde tracing results are cross-validated by retrograde labeling injections at anterograde fiber terminal fields and vice versa. No data has been excluded from this study and all image data generated is published online as part of the Mouse Connectome Project (<http://www.mouseconnectome.org>).

Stereotaxic surgeries. Mice were anesthetized in an induction chamber primed with isoflurane (Hospira) and subsequently mounted to a Kopf stereotaxic apparatus where they were maintained under anesthetic state via a vaporizer (Datex-Ohmeda). For dual co-injections, tracer cocktails were delivered iontophoretically via glass micropipettes (outer tip diameter of 15–20 µm) using alternating 7-s pulsed positive 5-µA current for 5 (BDA or AAV/FG) or 10 min (PHAL-CTB-647) generated by a current source (Stoelting). Triple anterograde tracing experiments were performed using similar iontophoretic parameters although the duration was different (5 min for PHAL, 1.5 min for AAVs). In quadruple retrograde tracing experiments, retrograde tracers were loaded into glass micropipettes that were connected to a picoPump pressure injector. At each injection site, 50 nl of retrograde tracer was pressure injected at a rate of 10 nl/min. For PHAL/rabies experiments, PHAL was delivered iontophoretically and 50 nl of G-deleted Rabies-GFP was pressure injected following the same parameters described above. In all experiments, pipettes were left in place for an additional 5 min following delivery to avoid diffusion of tracers along the needle track. Animals survived for 7 d before being sacrificed except animals injected with AAV1.hSyn.GFP or AAV1.CAG.RFP were sacrificed 3 weeks following surgeries to ensure fluorescent labeling had sufficiently labeled distant axon terminals.

Histology and immunohistochemical processing. Each animal was deeply anesthetized with an overdose injection of sodium pentobarbital and transcardially perfused with approximately 50 ml of 0.9% saline solution followed by

50 ml of 4% paraformaldehyde (PFA; pH 9.5). The brains were post-fixed in 4% PFA for 24–48 h at 4 °C.

For tissue sectioning, fixed brains were embedded in 3% Type I-B agarose (Sigma-Aldrich) and sectioned into four series of 50-µm-thick coronal sections with a Compresstome (VF-700, Precisionary Instruments). For all experiments, one series of sections was stained for NeuroTrace 435/455 (NT; 1:1000; Invitrogen, #N21479) and PHAL (if necessary) immunofluorescence using the free-floating method and the other three series were stored in cryopreservant under –20 °C if additional staining was needed. For PHAL immunostaining, sections were placed in a blocking solution containing normal donkey serum (Vector Laboratories) and Triton X (VWR) for 1 h. Following 3–5-min rinses, sections were incubated in PHAL primary antiserum (KPBS solution comprised of donkey serum, Triton, and 1:1,000 rabbit anti-PHAL antibody (Vector Laboratories, #AS-2300; see our previous study for validation of this antibody^{18,52}) for 48–72 h at 4 °C. Sections were rinsed three times in KPBS and then soaked for 3 h in the secondary antibody solution (donkey serum, Triton and either 1:500 anti-rabbit IgG conjugated with Alexa Fluor 488 for dual co-injection experiments or 1:500 anti-rabbit IgG conjugated with Alexa Fluor 647 for triple anterograde and rabies tracing experiments (Invitrogen, 488:#A-21206, 647: #A-31573). After processing, sections were mounted and coverslipped using 65% glycerol.

Imaging and post-acquisition processing. All tissue sections were scanned as high-resolution virtual slide image (VSI) files using an Olympus VS110 high-throughput microscope fitted with a 10× objective lens. Images were captured tile by tile using appropriately-matched fluorescent filters and then assembled together as whole-brain images. The number of color channels in each image depends on the number of tracers used in each experiment. For dual co-injection experiments, each image contains five color channels (green = PHAL, magenta = CTB-647, red = BDA, yellow = FG, and blue = NT). Images from quadruple retrograde tracer experiments contain five channels (green = CTB-488, magenta = CTB-647, red = CTB-555, yellow = FG, blue = NT), images from triple anterograde tracing data contain four channels (green = AAV1.hSyn.GFP, red = AAV1.CAG.RFP, magenta = PHAL), and images from rabies experiments contain three channels (green = G-deleted Rabies-GFP, magenta = PHAL, blue = NT). Before online publication, all images are aligned to the correct left-right orientation, matched to the nearest ARA atlas level, and converted to tiff format before being registered (detailed below). Following registration and registration refinement, the NT blue channel was converted to a bright-field image. Next, each channel for every image was adjusted for brightness and contrast to maximize labeling visibility and quality in iConnectome. Following final modifications (that is, skewness, angles) and JPEG2000 file format conversions, all images were published to iConnectome (<http://www.mouseconnectome.org>). For publication, some tracer colors have been changed from the raw images for better multicolor visualization.

Tracer experiment reproducibility. All injection experiments shown in this manuscript are listed in Supplementary Table 3, although additional supporting data and relevant injection sites can be found at <http://www.mouseconnectome.org> (currently, 553 total experiments, 1,194 injection sites labeling 1,872 pathways across the mouse brain). For each figure panel with tracer experiments, we list the number of times the experiment has been repeated and the number of cases which cross-validate the result using the data in Supplementary Table 3:

- Figure 2a,b. Retrograde injections were repeated with similar results [CA3d (2), CA3ic (4), CA3v (2), CA1dr (4), CA1i (5), CA1v (4), CA1vv (3), CA1dc (4), SUBdd (8), ProSUB (3)] and are cross-validated by anterograde injections [CA3dd (2), CA3d (2), CA3id (2), CA2 (3)].
- Figure 2c. Some CA3 anterograde injections were repeated with similar results [CA3dd (2), CA3d (2), CA3id (2), CA3ic (2), CA3v (1), CA3vv (1)] and are cross-validated by retrograde CA1 injections (CA1dr (4), CA1dc (4), CA1i (5), CA1v (4), CA1vv (3)).
- Figure 3a. Retrograde injections were repeated with similar results [MM (9), SM/PH (3), RE/AMD (5)] and cross-validated by SUB anterograde injections (SUBdd (8) and ProSUB (2)).
- Figure 3b. Some retrograde injections were repeated with similar results [AONM/TD (1), MEAad/CEA (3), RE (5), PVT (2), PT (2)] and cross-validated by SUB anterograde injections [SUBv (2) and SUBvv (3)].
- Figure 3d. Retrograde injections were repeated with similar results [CA1i (5), CA1v (4), CA1vv (3), SUBvv (1)] and cross-validated by SUB anterograde injections [SUBv (2) and SUBvv (3)]. Anterograde injections were repeated with similar results [CA1vv (1), SUBv (2)] and cross-validated by retrograde injections (CA1vv (3) and SUBv (3)).
- Figure 3e. Retrograde injections were repeated with similar results [AD (1), RE (5), PVT (2), PT (2)] and cross-validated by SUB anterograde injections [SUBv (2) and SUBvv (3)]. Anterograde injections were cross-validated by retrograde injections [SUBv (2) and SUBvv (1)].
- Figure 5a. Anterograde injections were repeated with similar results [SUBdd (9), SUBdv (4)] and cross-validated by retrograde injections [RSpv (3), POST (2), PRE (2), PAR(2)]. Retrograde injections were repeated with similar results [SUBdd (8), SUBdv (3)] and cross-validated by anterograde injections [CA1dc (2) and CA1i (1)].

- Figure 5b. Retrograde injections were repeated with similar results [SUBdd (8), SUBdv (3), POST (2), PRE (2), PAR (2)] and cross-validated by anterograde injections [LD (2)]. Anterograde injections were repeated with similar results [SUBdd (9), SUBdv (4), LD (2)] and cross-validated by retrograde injections [POST (2), PRE (2), PAR (2)].
- Figure 5c. Some retrograde injections were repeated with similar results [MM (9), RE/AMd (5), PT (3), AD (1), PVT (2), AV (4)] and cross-validated by anterograde injections [SUBdd (9) and SUBdv (4)].
- Figure 5d. Some retrograde injections were repeated with similar results [MM (9), RE/AMd (5), AV (4), RSPv (3), LM (1)] and cross-validated by anterograde injections [SUBdd (9) and SUBdv (4)].
- Figure 6a. Anterograde injections were repeated with similar results (SUBdd (9), ProSUB (2), SUBv (2)) and cross-validated by retrograde injections [RSPv (3), POST (2), PAR (2), ILA (1)].
- Figure 6c. Anterograde injections were repeated with similar results [SUBdd (9), ProSUB (2), SUBv (2), SUBvv (3)] and cross-validated by 21 hypothalamus retrograde injections.
- Figure 7a. Some retrograde injections were repeated with similar results [RSPv (1), RSPv (3), ACAd (1), ACAv (1), SUBdd (8), ProSUB (3)]. Anterograde injections were repeated with similar results [RSPd (1), RSPv (1), ACAd (1), ACAv (1), SUBdd (9), ProSUB (2), SUBv (2)].
- Figure 7b. Some retrograde injections were repeated with similar results [CA3v (2), PL (1), ILA (1), ProSUB (3), CA1i (5), CA1v (4)].
- Figure 7c. Retrograde injections were repeated with similar results [MM (9) and are cross-validated by anterograde injections (SUBdd (9), SUBdv (4), ProSUB (2), SUBv (2), SUBvv (3))].
- Figure 8a. These experiments were not repeated but are cross-validated by anterograde injections [CA1v (2), CA1vv (1), SUBv (2), and SUBvv (3)].
- Supplementary Fig. 6c. DGd anterograde injections were repeated with similar results [DGd(5)] and cross-validated by CA3 retrograde injections [CA3dd(1), CA3d (2), and CA3id (1)].
- Supplementary Fig. 6b,d,e,f were not repeated.
- Supplementary Fig. 7a. Anterograde injections were repeated with similar results [CAAdd (2) and CA3d (2)] and cross-validated by CA1 retrograde injections [CA1dr(4), CA1dc (4), and CA1i (5)].
- Supplementary Fig. 7b. Anterograde injections were repeated with similar results [CA3dd (2) and CA3d (2)] and cross-validated by CA1 retrograde injections [CA1dr(4), CA1dc (4), CA1i (5)].
- Supplementary Fig. 7c. Retrograde injections were repeated with similar results [SUBdd (8) and ProSUB (3)] and cross-validated by CA1 anterograde injections [CA1dr(1), CA1dc (2), and CA2 (3)].
- Supplementary Fig. 7d. Anterograde injections were repeated with similar results [CA1dr (1) and CA1dc (2)] and cross-validated by SUB retrograde injections [SUBdd (8) and ProSUB (3)].
- Supplementary Fig. 8. Anterograde injections were repeated with similar results [CA3dd (2), CA3d (2), CA3id(2), CA3ic (2), CA3v(1), and CA3vv(1)] and cross-validated by CA1 retrograde injections [CA1dr(4), CA1dc (4), CA1i (5), CA1v(4), CA1vv(3), SUBv(2), and SUBvv(1)].
- Supplementary Fig. 9a. Anterograde injections were not repeated but are cross-validated by SUB retrograde injections [SUBv (2) and SUBvv (1)]. Retrograde injections were repeated with similar results [CEA/MEA (3), LA (2), and mBLAA (2)] and are cross-validated by SUB anterograde injections [SUBv (2) and SUBvv (3)].
- Supplementary Fig. 9b. Anterograde injections were repeated with similar results [SUBv (2) and SUBvv (3)] and are cross-validated by 6 amygdala retrograde injections.
- Supplementary Fig. 9c. Anterograde injections were repeated with similar results [CA1v (2), ProSUB (2), SUBv (2) and SUBvv (3)].
- Supplementary Fig. 9d. Anterograde injections were repeated with similar results [CA1v (2), ProSUB (2), SUBv (2) and SUBvv (3)].

Construction and analysis of intrahippocampal connectivity matrix. Data collection was not randomized and analyses were not performed blind to the conditions of the experiments, as injection site location is apparent in tissue sections when annotating connectivity data. For intrahippocampal connectivity, the presence or absence of anterograde or retrograde labeling within all hippocampal regions was manually annotated with both directionality and labeling density weight (Supplementary Table 6; weighting scale = 0–3; 0 = none to sparse labeling, 1 = minor labeling, 2 = moderate labeling, 3 = robust labeling). The annotated data was used to construct a directed, weighted connectivity matrix using the gene-expression-defined and additional anatomically defined hippocampus regions as the major network nodes (Fig. 4b). The connectivity matrix can be read in two ways. For each node listed on the vertical axis, the output connections are listed across the row. Alternatively, the inputs to each node can be interpreted by reading down the columns on the horizontal axis.

Separately, all hippocampal region inputs and outputs were manually annotated from exemplar injection cases and listed in Supplementary Table 5 (non-weighted). Data from Supplementary Table 5 was used for extrahippocampal analysis and wiring diagram construction (Fig. 4c,d).

Hippocampal network analysis and visualization. We applied a multi-scale strategy³⁴ to explore the network organization of the intrahippocampal anatomical data. First, we executed a parallel Louvain modularity maximization analysis using the Brain Connectivity Toolbox for Python (bctpy; <https://github.com/aestrivex/bctpy>, <https://sites.google.com/site/bctnet/>) on USC Stevens Neuroimaging and Informatics Institute computational resources⁵⁵. The algorithm was executed from gamma values of 0.01 to 20.00, at increments of 0.01. To account for variability in the Louvain results, the algorithm maximized modularity over 1000 runs per iteration (2,000,000 total runs across all gamma values). We calculated a MPS metric to identify the most relevant of the 2000 scales³⁶. Finally, we computed a consensus partition from the maximization runs calculated at each scale³⁷.

Noteworthy high MPS value peaks were recorded at gamma 0.15, 1.36, and 2.04. Gamma 0.15 is the lowest gamma value featuring an MPS peak (4.23). The corresponding consensus partition exhibits a distinct bifurcation between dorsal hippocampus and ventral hippocampus regions. Gamma 2.04 exhibits the highest MPS value (6.57) of all scales, whereas gamma 1.36 contained the second highest MPS value (6.44). For a multi-scale representation of the intrahippocampal data, we therefore assigned each community in the consensus partition of gamma 2.04 as a subset of the larger communities (subnetwork modules) defined by the consensus partition at 1.36 and 0.15 (connectivity matrix in Fig. 4b). In addition, we constructed a representative network connectivity graph using pydot (<https://github.com/erocarrera/pydot>) and matplotlib Python libraries³⁸. The network graph contains three levels of decreasing edge weights corresponding to intra-community, inter-community, and inter-module connectivity with the densest weights referring to intra-community connections at gammas 0.15, 1.36, and 2.04. All related code and data are available at <https://git.ini.usc.edu/ibowman/HGEA>. In addition, we have created several informatics and visualization tools to view the data that are openly available on our Mouse Connectome Project website (<http://www.mouseconnectome.org/Analytics/page/matrix>).

For extrahippocampal analysis, we employed a similar multiscale approach using the non-weighted annotated connectivity data in Supplementary Table 5. For this dataset, we found the highest MPS peak at gamma 9.74 with a corresponding consensus partition featuring 46 different communities:

```
[['AAA'], ['ACAv', 'AOND', 'COAp', 'SUBvv', 'TMD', 'TMv'], ['ACB', 'AONI', 'ARH', 'BMAp', 'BSTTr', 'CA2'], ['AD'], ['ADP', 'CLA', 'CM'], ['AHN', 'BMAa', 'BSTdm'], ['AMD', 'IAD', 'PH', 'SUBdv'], ['AMv', 'BSTif', 'IMD', 'LPO', 'MEPO', 'ProSUB'], ['AO', 'BLAA', 'BSTm', 'CA1v', 'PL'], ['AONm', 'BStal'], ['AONpv', 'BSTM', 'CA1dc', 'ORBm', 'SUM'], ['AV'], ['AVP', 'BStv', 'DGD', 'ENT', 'ENTr'], ['AVPV'], ['BLAp', 'BLAv', 'EPD'], ['BStTr', 'CSI', 'CSm'], ['CA1dr', 'LA', 'POST', 'PRE'], ['CA1i', 'DP', 'SUBv'], ['CA1vv', 'LSv', 'MEApv', 'MM'], ['CA3d', 'CA3dd', 'CA3id', 'LSc'], ['CA3ic', 'caudal BLAa'], ['CA3v', 'CA3vv'], ['CEAc', 'TTd', 'TTv', 'TU', 'VLPO'], ['COAa', 'PAA', 'PIR', 'TR'], ['COAp', 'CP', 'ECT', 'PA'], ['DGd'], ['DGpod'], ['DGpov'], ['DGv'], ['DMH'], ['Epv', 'PERI'], ['FS', 'MPN', 'MPO', 'OT'], ['IAM', 'NLot', 'ORBv'], ['ILA', 'LHA'], ['LC'], ['LD'], ['Lsr', 'PS'], ['MEAad', 'MEAav', 'MEApd', 'PAR', 'RC', 'RE'], ['MS/NDB'], ['PMv', 'PT', 'PVT', 'PVp'], ['PVpo'], ['RSPV'], ['SBPV'], ['SI'], ['SUBdd'], ['VMH']]], [[['AAA', 'BStTr', 'CLA', 'EPD', 'FS', 'LA', 'NLot', 'ORBv', 'PAA', 'SI', 'TTv', 'VLPO'], ['ACAv', 'AONi', 'BLAA', 'CA1dr', 'CA1v', 'DGv', 'PL', 'POST', 'PRE', 'TMv'], ['ACB', 'ARH', 'BStv', 'CA3ic', 'TU', 'caudal BLAa'], ['AD', 'AV', 'AVP', 'BMAp', 'CA2', 'ENT', 'LD', 'RSPv', 'SUBdd'], ['ADP', 'DGd', 'DGpod', 'DGpov'], ['LC', 'MS/NDB'], 'CA2', 'PMv', 'PS', 'PVp', 'PVpo', 'SBPV', 'SUBv', 'TMD', 'VMH'], ['AHN', 'CA3id', 'COAp', 'DGd', 'DMH', 'PH'], ['AMD', 'COAa', 'IAD', 'PERI', 'PIR', 'PT', 'RE', 'SUBdv'], ['AMv', 'AONi', 'BMAa', 'CA3d', 'ENTm', 'IMD', 'LPO', 'MEPO', 'PAR', 'ProSUB'], ['AOND', 'BStdm'], ['AONm', 'SUBv'], ['AONpv', 'CA1dc', 'ORBm', 'SUM'], ['AVP', 'BStTr', 'CA1i', 'DP'], ['BLAp'], ['BLAv', 'CA3dd', 'LSc'], ['BStal'], 'CEAc', 'LSv', 'MEApd'], ['BSTM', 'CA1v', 'ENTr'], ['BStif', 'TTd'], ['BStmg', 'IAM'], ['CA1vv', 'LSv', 'MEApv'], ['CA3v', 'CA3vv', 'MEAad', 'MEAav', 'RC'], ['CM', 'TR'], ['COAp'], ['CP'], ['CSI', 'CSm', 'MM'], ['Epv', 'MPN'], ['ILA', 'LHA'], ['MPO'], ['OT']]]]
```

We manually arranged each of these 46 different communities into five subnetworks based on the similarity of their extrinsic connectivity with the intrahippocampal communities found in the previous analysis. The resulting community structure was graphed using pydot (<https://github.com/erocarrera/pydot>) and matplotlib Python libraries³⁸ with three levels of decreasing edge weights corresponding to intra-community, inter-community, and inter-module connectivity (employing a similar process as the intrahippocampal analysis). All related code and data are available at <https://git.ini.usc.edu/ibowman/HGEA>.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Code availability. All related code and data are available at <https://git.ini.usc.edu/ibowman/HGEA> and are included in the Supplementary Software file.

Data availability

All images of mouse *in situ* hybridization gene expression data are available at the Allen Brain Atlas website (<http://www.mouse.brain-map.org>). Gene accession codes for all genes and the *in situ* hybridization probe sequences can be found by querying the gene name through the Allen Brain Atlas website (<http://www.mouse.brain-map.org>). All anatomical tracer image data is available through our

iConnectome viewer as part of the Mouse Connectome Project at USC (<http://www.mouseconnectome.org>). The HGEA stereotaxic coordinate atlas, 3-D atlas viewer, and annotation data are available for download at <http://www.mouseconnectome.org/MCP/page/papers>. Additional informatics and visualization tools are available online at <http://www.mouseconnectome.org/Analytics/page/matrix>.

References

51. Lein, E. S. et al. Genome-wide atlas of gene expression in the adult mouse brain. *Nature* **445**, 168–176 (2007).
52. Hintiryan, H. et al. The mouse cortico-striatal projectome. *Nat. Neurosci.* **19**, 1100–1114 (2016).
53. Hintiryan, H. et al. Comprehensive connectivity of the mouse main olfactory bulb: analysis and online digital atlas. *Front. Neuroanat.* **6**, 30 (2012).
54. Betzel, R. F. & Bassett, D. S. Multi-scale brain networks. *Neuroimage* **160**, 73–83 (2017).
55. Blondel, V. D., Guillaume, J.-L., Lambiotte, R. & Lefebvre, E. Fast unfolding of communities in large networks. *J. Stat. Mech.* **2008**, P10008 (2008).
56. Gu, S. et al. Controllability of structural brain networks. *Nat. Commun.* **6**, 8414 (2015).
57. Lancichinetti, A. & Fortunato, S. Consensus clustering in complex networks. *Sci. Rep.* **2**, 336 (2012).
58. Droettboom, M. et al. Matplotlib/Matplotlib: V2. 0. 0. <https://doi.org/10.5281/zenodo.248351> (2017).

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated
- Clearly defined error bars
State explicitly what error bars represent (e.g. SD, SE, CI)

Our web collection on [statistics for biologists](#) may be useful.

Software and code

Policy information about [availability of computer code](#)

Data collection

No software was used for data collection.

Data analysis

All custom code that we developed for data analysis is available at <https://git.ini.usc.edu/ibowman/HGEA> and is included in the Supplementary Software file. In addition, we have used several open-source Python packages: scipy (v0.17.0), matplotlib (v1.5.1), ipython (v2.4.1), ipython-genutils (v0.2.0), jupyter (v1.0.0), jupyter-client (v5.2.3), jupyter-console (v5.2.0), jupyter-core (v4.4.0), pandas (v0.23.3), sympy (v1.2), nose (v1.3.7), pydot (v1.2.3), bctpy (v0.5.0), pyflakes (v1.1.0), pandas (v0.23.3), cycler (v0.9.0), plotly (v3.0.0).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All images of mouse *in situ* hybridization gene expression data are available at the Allen Brain Atlas website (www.mouse.brain-map.org). All anatomical tracer image data is available through our iConnectome viewer as part of the Mouse Connectome Project at USC (www.MouseConnectome.org). The HGEA stereotaxic coordinate atlas, 3-D atlas viewer, and annotation data are available for download at www.MouseConnectome.org/MCP/papers. Additional informatics and visualization tools are available online at www.MouseConnectome.org/Analytics/page/matrix.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size

No sample size calculation was performed. For gene expression analysis, 255 genes were chosen for manual annotation by examining the ~top 50 highly-expressed genes in the Allen gene expression database for the dentate gyrus, CA3, CA1, and subiculum. In addition, we examined genes that were included in Thompson et al. 2008 as well as other notable genes we found by browsing through thousands of gene expression profiles in the Allen database. As we have examined many more genes that were not annotated, we do not believe that additional annotation will alter the interpretations or results presented in this manuscript. For the anatomical tracing, we systematically injected multiple tracers to target each genetic HGAA subregion and determine their inputs and outputs. In addition, we have made multiple injections outside of the hippocampus to confirm the specificity. In many cases, we have both anterograde and retrograde data to describe an anatomical pathway and the combination of the two tracer types provides cross-validation. In total, we have listed 79 total mice that contain 177 different tracers that we believe sufficiently describe the anatomical network connectivity of the hippocampus.

Data exclusions

No analyzed data was excluded

Replication

Anatomical connections were often cross-validated by the use of both anterograde and retrograde tracers. For example, if area A->B, this same axonal projection can be labeled by an anterograde injection in area A or a retrograde injection in area B. In this way, many of the connections we present are replicated both by multiple experiments and different tracers. Statements as to the number of repeated experiments and number of experiments which cross-validate the result are included in the figure legend.

Randomization

We do not believe randomization of the mice is relevant as we do not make statistical comparisons.

Blinding

The investigators were not blinded as the injection site location is apparent when analyzing the distribution of anatomical labeling throughout the brain.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	Unique biological materials
<input type="checkbox"/>	Antibodies
<input checked="" type="checkbox"/>	Eukaryotic cell lines
<input checked="" type="checkbox"/>	Palaeontology
<input type="checkbox"/>	Animals and other organisms
<input checked="" type="checkbox"/>	Human research participants

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	ChIP-seq
<input checked="" type="checkbox"/>	Flow cytometry
<input checked="" type="checkbox"/>	MRI-based neuroimaging

Antibodies

Antibodies used

A donkey anti-rabbit PHAL primary antibody (Vector Laboratories, #AS-2300) and rabbit IgG secondary antibodies conjugated to either AlexaFluor 488 or AlexaFluor 647 (Invitrogen, 488: #A-21206, 647: #A-31573) were used to localize PHAL tracer within brain tissue sections.

Validation

Rabbit anti-PHAL antibody (Vector Laboratories, #AS-2300) has been used and validated in our previous studies (Biag et al, 2012; Hintiryan et al, 2016)

Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals

All tracer experimental data was generated using 2-6 month-old male C57BL/6J mice (Jackson Laboratories)

Wild animals

Study did not involve wild animals

Field-collected samples

Study did not involve field-collected samples.