Exploring Memory Representations with Activity-Based Genetics

Mark Mayford\textsuperscript{1} and Leon Reijmers\textsuperscript{2}

\textsuperscript{1}Molecular and Cellular Neurosciences Department, The Scripps Research Institute, La Jolla, California 92037
\textsuperscript{2}Department of Neuroscience, School of Medicine, Tufts University, Boston, Massachusetts 02111

Correspondence: mmayford@scripps.edu; leon.reijmers@tufts.edu

The brain is thought to represent specific memories through the activity of sparse and distributed neural ensembles. In this review, we examine the use of immediate early genes (IEGs), genes that are induced by neural activity, to specifically identify and genetically modify neurons activated naturally by environmental experience. Recent studies using this approach have identified cellular and molecular changes specific to neurons activated during learning relative to their inactive neighbors. By using opto- and chemogenetic regulators of neural activity, the neurons naturally recruited during learning can be artificially reactivated to directly test their role in coding external information. In contextual fear conditioning, artificial reactivation of learning-induced neural ensembles in the hippocampus or neocortex can substitute for the context itself. That is, artificial stimulation of these neurons can apparently cause the animals to “think” they are in the context. This represents a powerful approach to testing the principles by which the brain codes for the external world and how these circuits are modified with learning.

A central feature of nervous systems is that, to function properly, specific neurons must become active in response to specific stimuli. The nature of this selective activation and its modification with experience is the focus of much neuroscience research, ranging from studies of sensory processing in experimental animals to disorders of thought such as schizophrenia in humans. The central dogma of neuroscience is that perceptions, memories, thoughts, and higher mental functions arise from the pattern and timing of the activity in neural ensembles in specific parts of the brain at specific points in time. Until quite recently, the investigation of these “circuit”-based questions has primarily been limited to observational techniques, such as single unit recording, functional magnetic resonance imagery (fMRI), and calcium imaging, to document the patterns of neural activity evoked by sensory experience or even complex psychological contingencies in human fMRI studies. These techniques have been enormously successful and created a framework for understanding information processing in the brain. For example, recordings in the visual system have indicated that, in the primary visual cortex, neurons are tuned to the orientation of linear stimuli (Hubel and Wiesel 1962). In contrast, neurons in higher brain areas can respond to discrete items. The most striking example of
this specificity comes from in vivo recording in the human medial temporal lobe in which single units have been identified that respond to photos of the actress Halle Berry as well as her written name (Quiroga et al. 2005). This highly selective tuning of neural activity is suggestive of function, but how can this be directly tested? What would be the effect of stimulating just this rare population of neurons, a memory of the actress, a sensory illusion of her image? How does this type of specific firing arise? Do these neurons differ from their nonresponsive neighbors in terms of biochemistry, cell biology, or connectivity? Do they undergo molecular alterations when new information is learned about this individual and are these changes required for the learning? These types of questions have recently become accessible to study in mice through the use of activity-based genetic manipulation, in which neurons that are activated by a specific sensory stimulus can be altered to express any gene of experimental interest. These studies and approaches will be the focus of this work.

IMMEDIATE EARLY GENES FOR ACTIVITY MAPPING

The first observation that the expression of certain genes was responsive to neural activity was made almost 30 years ago with the identification of increased expression of *cfos* in the brain following seizure (Morgan et al. 1987). *cfos* is one of a class of genes known as immediate early genes (IEGs) that are defined by their rapid induction by pre-existing transcription factors without the need for de novo protein synthesis (Greenberg et al. 1986). This allows rapid transcription initiation, within 1 min of electrical stimulation, supported by RNA polymerase II that is bound to the promoter region under resting conditions (Saha et al. 2011). On neuronal activation, the stalled RNA polymerase is released, thereby enabling the extremely rapid induction of IEG transcription. In addition to rapid induction, the *cfos* protein and messenger RNA (mRNA) have a relatively short half-life, such that following the end of active transcription the levels of expressed protein rapidly return to baseline (Rahmsdorf et al. 1987; Wellington et al. 1993). In the induction of *cfos* with seizure, the expression peaks at 1 h and returns to baseline by 3 h, allowing the expression of this gene to provide a snapshot view of brain activity with this 3-h time window (Morgan et al. 1987). Since the initial identification of *cfos*, genetic screens have identified a wide array of immediate early genes that show similar responses to neural activity (Nedivi et al. 1993; Qian et al. 1993; Yamagata et al. 1993). The most notable are *arc* and *zif268/erg1*, which, along with *cfos*, have been used extensively as a surrogate measure for neural ensemble activity in experimental animals (Worley et al. 1991; Lyford et al. 1995).

There is an extensive literature on the regulation of IEGs in cultured neurons and non-neuronal cells in which transcription can be triggered in response to multiple classes of stimuli, including growth factors and cAMP as well as strong depolarization (Ghosh et al. 1994). The degree of synaptic or intrinsic neural activity required to induce IEG expression is not straightforward and varies between genes and brain regions. For example, in vivo stimulation of dentate gyrus (DG) granule cells with high-frequency bursts that are sufficient to produce long-term potentiation (LTP) causes strong induction of *zif268*, but *cfos* is only induced in these neurons with more prolonged stimulation protocols. In both cases, the induction requires N-methyl-D-aspartate (NMDA) receptor function (French et al. 2001). In hippocampal CA1 neurons, similar stimulation protocols failed to induce any of the IEGs. In hippocampal slice, the *cfos* gene is induced with high fidelity by 30 or more action potentials delivered at 10 Hz or greater frequency (Schoenenberger et al. 2009). Although the precise nature of activity required to induce IEG expression is not clear, the expression of IEGs has been used extensively as a measure of neural activation in response to environmental stimulation and has given results that often track with activity determined by other techniques, such as electrophysiological recording or with known anatomical function and connectivity. For example, *cfos* is activated in a tonotopic pattern in the auditory brainstem.
with auditory stimulation (Saint Marie et al. 1999) and in a somatotopic pattern in somatosensory cortex following whisker stimulation (Wagener et al. 2010). In learning and memory paradigms, such as fear conditioning, there is cfos activation in an array of areas, such as hippocampus, amygdala, prefrontal cortex, and anterior cingulate cortex, that are known to be active and functionally necessary for the learning or retrieval of contextual fear memories (Milanovic et al. 1998; Radulovic et al. 1998; Bontempi et al. 1999; Frankland et al. 2004; Knapska and Maren 2009).

Until the development of Ca$^{2+}$ imaging approaches (Jercog et al. 2015), detection of IEG expression was one of the only ways of assessing large ensembles of neural activity with cellular resolution. A critical question in learning and memory (and neural processing more generally) is how similar is neural circuit activation with learning and with retrieval or during multiple retrieval trials. One approach to this question that takes advantage of the rapid and transient nature of IEG expression is known as catFISH (for cellular compartment analysis of temporal activity by fluorescence in situ hybridization) (Guzowski et al. 1999). Because mRNA is initially produced in the nucleus in an unspliced form and is later transported to the cytoplasm following splicing, the use of probes targeting unspliced nuclear RNA (currently active neurons) and processed cytoplasmic RNA (previously active neurons) can be used to determine the activity of neurons at two different but closely spaced time points. This approach was used to examine the activity of CA1 hippocampal neurons following the exploration of two identical environments relative to the exploration of two distinct environments. The results showed that there was a greater overlap in arc expression in CA1 neurons following the two epochs exploring identical environments. This is consistent with the role of the hippocampus in place recognition and with in vivo electrophysiological studies of place cell activity and its role in encoding location. It also supports the idea that IEG expression reliably reports behaviorally relevant neural activity.

### IEG-Based Transgensics

Although the IEGs offer a temporally crude measure of neural activation, they do parallel other methods of assessing activity in many brain areas. In addition, genetic studies have suggested a role for cfos, zif268, and arc in synaptic plasticity and memory, suggesting that their activation may be particularly important in identifying circuits undergoing plastic changes (Fleischmann et al. 2003; Plath et al. 2006; Baumgartel et al. 2008). One advantage of IEGs over more direct and temporally precise measures of neural activity is that by providing a transcriptional response they offer a molecular genetic conduit into environmentally activated neural ensembles. This was first shown in a simple transgenic mouse in which a promoter element from the cfos gene was used to drive expression of a lacZ marker protein (Smeyne et al. 1992). The expression of the marker was strongly induced throughout the brain following seizure. In addition, the lacZ expression was responsive to environmental stimuli showing strong induction in the superchiasmatic nucleus, an area involved in circadian rhythm regulation, following a light pulse during the animal’s dark cycle. The transgene incorporated just 600 bp of promoter region containing four different transcriptional response elements, all of which were required for normal stimulus-induced gene expression (Robertson et al. 1995).

More recently, a number of transgenic mouse lines have been developed that drive green fluorescent protein (GFP) in an activity-dependent manner (see Table 1 for a list of published IEG-based transgenic lines) (Barth et al. 2004; Wang et al. 2006; Reijmers et al. 2007). The use of a fluorescent marker compatible with live tissue imaging allows the recording of neurons in brain slices or in anesthetized animals specifically from recently active cfos-positive cells. This approach was used to examine spontaneously active neurons in the somatosensory cortex, where approximately 15% of neurons in layer 2/3 are cfos-GFP positive in the absence of any specific environmental stimuli (Yassin et al. 2010). These cfos-positive neurons showed an increase in excitatory drive, a reduc-
tion in inhibitory drive, and a greater degree of interconnectivity than their cfos-negative neighbors. These results indicate that ensembles of neurons with increased activity are maintained for several hours (the timing between cfos expression and electrophysiological recording) and that there may be a stable structure and connectivity in these spontaneously active ensembles. The presence of such spontaneously active ensembles is consistent with electrophysiological recordings that have suggested the importance of posttraining replay of neural activity patterns evoked during learning in the consolidation of memory (Wilson and McNaughton 1994; Foster and Wilson 2006; Karlsson and Frank 2009).

IEG-linked GFP reporter mice have also been used as a convenient means of repeatedly measuring neural activation patterns in intact animals by two-photon imaging through cranial windows. To examine the role of arc in the orientation selectivity responses of neurons in the primary visual cortex, Wang et al. (2006) generated a mouse in which GFP replaced the functional arc protein to produce a knock-out in which the activation of neural ensembles could be imaged through measurement of GFP expression levels. Heterozygous mice (carrying one functional arc allele and the activity-dependent arc-GFP) showed normal orientation tuning, whereas the homozygous, arc knockout mice showed impaired tuning of individual neurons measured electrophysiologically. Wang et al. then examined the activity of all neurons in the field using the GFP reporter and determined that the overall loss of tuning specificity was because of an increase in the number of weakly tuned cells rather than a loss of specificity in the highly tuned neurons. In another study, a transgenic mouse in which GFP was driven by the zif268 promoter was used to examine the formation of contextual representations in the neocortex following fear conditioning (Xie et al. 2014). These investigators found slowly developing and very sparse contextual representations that were stable for up to 2 months. Similar results were obtained using a cfos-GFP reporter to examine the formation of spatial representations in the retrosplenial cortex (Czajkowski et al. 2014).

Finally, IEG-linked GFP reporters have been used to detect neural activity on a brain-wide scale (Kim et al. 2015; Vousden et al. 2015). Serial two-photon tomography is a method for obtaining fluorescent images in an automated manner with the capacity to capture 260 50 µM serial sections throughout an entire mouse brain at 10× resolution within several hours (Ragan et al. 2012). This imaging approach was applied to cfos-GFP transgenic mice and combined with an analysis pipeline for automated cell counting and anatomical partitioning to obtain whole brain activity maps at cellular resolution (Kim et al. 2015). The results were generally consistent with previous studies showing activation in appropriate areas of the hypothalamus and amygdala with social interaction and served as a proof of principle for automated whole brain imaging of activity using an IEG-based reporter.

### Table 1. Currently available immediate early genes (IEGs)-based transgenic mouse lines

<table>
<thead>
<tr>
<th>Gene Type</th>
<th>Type</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fos-lacZ</td>
<td>Transgenic</td>
<td>Smeyne et al. 1992 -</td>
</tr>
<tr>
<td>Fos-taulacZ</td>
<td>Transgenic</td>
<td>Mehta et al. 2002 -</td>
</tr>
<tr>
<td>Fos-GFP</td>
<td>Transgenic</td>
<td>Barth et al. 2004 014135</td>
</tr>
<tr>
<td>Fos-GFP/tTA</td>
<td>Transgenic</td>
<td>Reijmers et al. 2007 018306</td>
</tr>
<tr>
<td>Fos-CreER&lt;sup&gt;T2&lt;/sup&gt;</td>
<td>Knockin</td>
<td>Guenthner et al. 2013 021882</td>
</tr>
<tr>
<td>Arc-CreER&lt;sup&gt;T2&lt;/sup&gt;</td>
<td>Knockin</td>
<td>Guenthner et al. 2013 021881</td>
</tr>
<tr>
<td>Arc-CreER&lt;sup&gt;T2&lt;/sup&gt;</td>
<td>BAC transgenic</td>
<td>Denny et al. 2014 022357</td>
</tr>
<tr>
<td>Arc-GFP</td>
<td>Knockin</td>
<td>Wang et al. 2006 007662</td>
</tr>
<tr>
<td>ERG1-GFP</td>
<td>BAC transgenic</td>
<td>Xie et al. 2014 014709-UCD</td>
</tr>
</tbody>
</table>

List of IEG-based transgenic and knockin mouse lines. Most are available from the Jackson Laboratory repository (www.jax.org).
The use of activity-dependent promoters to drive expression of fluorescent markers allows for targeted recording, repeated imaging for prolonged time periods, and whole brain activity mapping at a single time point. To probe the function of active neural ensembles, several groups have developed IEG-based transgenic lines that allow the expression of effector molecules in neurons that are active at a specific time point. By using the c-fos promoter in conjunction with the tetracycline system for control of gene expression, Reijmers et al. 2007 produced a transgenic mouse in which neurons that are naturally active in a given time window could be genetically modified to express essentially any gene, as shown in Figure 1A. The approach requires two transgenes introduced into the same animal. The first transgene consists of the c-fos promoter driving expression of the tetracycline transactivator (tTA), which is a transcription factor that can be regulated with the antibiotic doxycycline (Dox) (Gossen and Bujard 1992). The second transgene carries a tetracycline-responsive element promoter (TRE) to drive expression of any gene of interest. In the presence of Dox delivered in the animal’s diet, the c-fos promoter will drive expression of tTA in active neurons; however, transcription of the second TRE-linked gene will be blocked. When Dox is removed, a time window is open during which neurons that are sufficiently active to induce the c-fos-tTA will express the TRE-linked transgene via tTA-driven transcription at the TRE promoter. The investigators used these mice to examine the neural ensembles in the amygdala that are activated with learning and retrieval in contextual fear conditioning. They used fos-tTA to drive expression of a long-lasting lacZ reporter protein in neurons activated during learning and compared the degree to which these neurons were reactivated following memory retrieval. They found that the amygdala

---

**Figure 1.** Two systems for the genetic manipulation of active neural ensembles. (A) In this tetracycline (TET)-based system, two transgenes are required, a c-fos promoter-driven tetracycline transactivator (tTA) and a tetracycline-responsive element (TRE) promoter-driven gene of interest. In the presence of doxycycline (Dox) the tTA is expressed in electrically active (c-fos+) neurons but is prevented from activating expression of the gene of interest by the presence of Dox. In the absence of Dox, a window is opened during which active neurons that express tTA drive expression of the gene of interest from the TRE promoter. (B) This Cre-based system also uses two transgenes, a c-fos promoter-driven CreER2 and a gene of interest that is flanked by loxP sites and positioned in an inverted orientation to any neuronal promoter (Pr). The loxP sites are arranged such that Cre activity will lead to a single inversion event of the flanked DNA. In the absence of tamoxifen (TAM) the Cre recombinase is inactive so that no recombination takes place even in active neurons. On administration of TAM, any active (c-fos+) neurons will express the Cre, which is now active, and inverts the orientation of the gene of interest. This gene is then constitutively and permanently expressed from the neurons specific promoter.

---

Advanced Online Article. Cite this article as Cold Spring Harb Perspect Biol doi: 10.1101/cshperspect.a021832
neurons activated during learning are responsive primarily to shock, the unconditioned stimulus (US), and that after training these shock-US-responsive neurons were activated by the context, the conditioned stimulus (CS), alone and that the degree of reactivation was correlated with the strength of the expressed fear memory. These results suggest a model whereby pairing of CS and US during training produces plasticity within the circuit that allows the CS to now recruit US neurons, linking the two pathways during recall. These results are consistent with the literature on amygdala circuitry in fear conditioning (and also with fear conditioning in *Aplysia*; see Byrne and Hawkins 2015; Hawkins and Byrne 2015) and provide a validation of this genetic model, which has been used extensively in studies of the circuitry underlying memory as discussed in later sections.

More recently, several groups have developed mouse lines in which IEG-promoters are used to drive expression of a tamoxifen-regulated form of Cre recombinase (CreERT2) (Guenthner et al. 2013; Denny et al. 2014). As shown in Figure 1B, these mice can be used to activate or delete genes from subsets of neurons that are active at a particular point in time. In the absence of tamoxifen, the cfos-linked CreERT2 will be expressed in active neurons but fail to induce loxP-mediated recombination because the protein is restricted to the cytoplasm by the estrogen receptor (ER\(^{T2}\)) component of the molecule. Administration of tamoxifen opens a time window during which CreERT2 expressed in any cfos-positive neurons will be translocated to the nucleus to induce loxP-based recombination to activate or delete target genes. The use of Cre-based systems offers the advantage of permanent genetic tagging of active neural ensembles, the ability to delete endogenous genes, and the availability of a greater number of Cre-responsive genetic tools than are available for the tTA system.

**IEG-Based Manipulation of Circuit Function**

As discussed in the introductory paragraphs, there is an extensive body of literature characterizing the patterns of neuronal firing in response to sensory stimuli. For example, neurons in the medial temporal lobe of human patients have been identified that seem to respond only to a specific individual (Quiroga et al. 2005). The IEG-based genetic tools provide a mechanism for directly probing the function of these sparse and dispersed ensembles in the representation of sensory information and memories. By driving the expression of genetic regulators of neural activity into the neural ensembles activated naturally by a specific environmental stimulus, the activity of these distributed ensembles can be experimentally controlled to test their behavioral relevance. In this section, we will discuss several recent studies using this approach to the study of the representation of context in fear conditioning.

Contextual fear conditioning is a hippocampal-dependent task in which animals learn to fear a place (the context) where they receive aversive foot shocks. How is the context (characterized by specific visual, tactile, and olfactory cues) represented by dispersed neural activity? Would artificially stimulating neurons naturally activated by the context be sufficient to substitute for the actual sensory experience of the context during fear conditioning? One attempt to address this question used the cfos-tTA system to drive expression of the excitatory hM3Dq receptor into neural ensembles that were active during contextual learning (Fig. 2A) (Garner et al. 2012). hM3Dq is a human Gq coupled muscarinic receptor that has been mutated so that it no longer responds to acetylcholine but responds instead to the synthetic ligand clozapine-N-oxide (CNO) (Alexander et al. 2009). In neurons expressing hM3Dq, application of CNO causes an 8–10-mV depolarization of the membrane potential and a subsequent increase in action potential firing. In this study, mice carrying both the cfos-tTA and TRE-hM3Dq transgenes were allowed to explore a distinct context (ctxA) causing the cfos ensembles activated by this environmental stimulation to express the hM3Dq transgene. The animals were then fear conditioned in a separate context (ctxB), and CNO was delivered to depolarize the hM3Dq-expressing ctxA ensemble of neurons. Under these conditions, the animals
fail to develop a fear response to either ctxA or ctxB, suggesting that the artificial activity produced in the ctxA neurons was not sufficient to substitute as a neural representation of ctxA but did interfere with the ability of the animals to represent the context in which they actually received foot shocks (ctxB). However, when the animals were placed in ctxB and received CNO to concurrently activate the ctxA neural ensemble, they showed a fear response that approached that of control animals. This suggests that the experimental animals formed an artificial contextual representation that incorporated both ctxA (artificial-internally generated) and ctxB (natural-sensory driven) neural ensembles.

Although these results failed to show that artificial activation of a distributed pattern of sensory evoked neurons could function to represent that sensory experience, it provided a number of important insights. First, the neurons that were activated artificially in this study were widely distributed throughout the neocortex and hippocampus. Previous studies had shown that local stimulation of a particular brain region or fiber bundle could serve as a CS in a conditioning paradigm (e.g., Huber et al. 2008), but the current results suggest that widely distributed, even global, patterns of activity can be meaningfully incorporated into neural representations. In fact, the brain is not silent in the absence of experimenter-provided stimuli yet this ongoing and dispersed “spontaneous” activity is rarely taken into account. Studies in the hippocampus, but also in the neocortex, indicate that spontaneous activity may not be random but instead reflect the replay of patterned activity associated with recent ex-

Figure 2. Manipulating contextual memory representations. This figure shows the arrangement of two sets of experiments to test the role of distributed neural ensembles in the coding of contextual memory. Both experiments use the cFos-tTA system discussed in Figure 1A to label neurons that are activated by exploration of a context (ctxA). In A, the active neurons express hM3Dq and in B, the neurons express ChR2. The animals are then fear conditioned in a distinct context (ctxB), whereas the ctxA neurons are activated with either clozapine N-oxide (CNO) or light, panels A and B, respectively. Memory retrieval is then tested in ctxA and ctxB. In panel A, the mice only show a fear response when the ctxA neurons are artificially activated while the animal is in ctxB, suggesting formation of a hybrid representation. In panel B, the animals show fear in ctxA even though they never received a shock in that context, suggesting that the artificial stimulation of the ChR2 positive neurons tagged in ctxA was able to substitute for (represent) that context. (Red circles) Neural ensembles expressing the genetic effector hM3Dq in A and ChR2 in B. (Blue circles) Neural ensembles naturally active by sensory input during training in ctxB.
periences (Kenet et al. 2003; Foster and Wilson 2006; Ji and Wilson 2007). The study by Garner et al. 2012 suggests that the integration of internally activated ensembles of neurons (ctxA in these experiments) at the time of new sensory-evoked learning (ctxB in this case) may represent a mechanism for integrating new information with previous experience, an important component of all higher forms of learning.

A more direct test of the functional relevance of distributed neural ensembles used the cfos-based genetic tagging approach in conjunction with channelrhodopsin (ChR2) (Boyden et al. 2005) and local stimulation in the hippocampus (Liu et al. 2012). In this study, the cfos	TA transgenic mouse was used to express a virally delivered TRE-ChR2 transgene in the DG region of the hippocampus (Fig. 2B). Animals received contextual fear conditioning to allow the expression of ChR2 in the learning-activated population of neurons. The animals were then moved to a neutral context that did not induce fear. Light-induced firing of the ChR2-labeled DG neurons in this neutral context was able to produce a fear response and this response was tightly linked to the light activation of the DG. Would stimulating any set of DG neurons produce a fear response, or do the neurons carry specific information about the context in which they were labeled? To address this question, active DG neurons were labeled with ChR2, while the animal explored one context (ctxA) but did not receive foot shocks. The animals were then fear conditioned in a distinct context (ctxB) while receiving Dox to prevent further labeling of active neurons. Stimulation of the ChR2-containing ctxA neurons did not produce a response in animals trained to fear only ctxB, indicating that the neurons carry context specific information. Taken together, the results suggest that the activation of a sparse (<5% of total neurons in the DG), distributed, population of neurons, originally activated during learning, was sufficient to produce an apparent context-specific memory retrieval event.

In a follow-up study using the same techniques (Ramírez et al. 2013), the neurons activated by exploration of a specific context (ctxA), in the absence of shock, were labeled with ChR2. The artificial stimulation of these ChR2-expressing neurons with light was then paired with foot shock in a different context (ctxB). When animals were placed back into ctxA, where they had never received the foot-shock US, they showed a fear response. This result suggests that the artificial stimulation of the distributed cfos-positive population of neurons labeled during the natural exploration of ctxA was sufficient to produce a neural representation of that complex environment. In essence, the artificial stimulation and the natural sensory experience were interchangeable, at least for the purposes of fear conditioning. The arrangement of this experiment is similar to that discussed previously in Garner et al. (2012) with the training taking place in ctxB, while a competing neural ensemble representing ctxA was stimulated. The results were similar; the animal’s fear memory in the natural conditioning chamber (ctxB) was strongly reduced by the stimulation of ctxA neurons during training but recovered to normal levels during concurrent activation of the ctxA cells. Whether this reflects the integration of the A and B ensembles into a single complex representation, as suggested by Garner et al. (2012), or simply reflects the effects of the compound CS-A and CS-B each acquiring a lesser amount of associative strength is unclear.

Taken together these results indicate that the artificial activation of a sparse and distributed sensory-evoked population of neurons is sufficient to instantiate a neural representation of that sensory experience. The observation that artificial stimulation with ChR2, which will tend to activate all neurons simultaneously and was not synchronized to endogenous rhythms, is sufficient to produce this effect suggests that the temporal patterning seen in natural activity may not be critical, at least for this particular brain region and behavior. There are many remaining questions relating to this effect, such as in what other brain regions can it be produced, how many neurons need to be stimulated, what frequency parameters are required to produce these effects, what downstream circuits are required to produce the behavior, and how are they recruited? Perhaps the most inter-
Interesting but difficult question to address is what is the animal experiencing when the neurons for ctxA are activated. Is it experiencing a memory of that context, a hallucinatory sensory experience, or a poorly defined sense of fear or anxiety independent of coherent contextual features?

Some of these questions are examined in another recent paper studying contextual representations in the retrosplenial cortex (Cowansage et al. 2014). The retrosplenial cortex is one of the output areas of the hippocampus and projects to a wide variety of other cortical areas. It is required for the encoding and retrieval of both recent and remote contextual fear memories (Keene and Bucci 2008a,b). Similar to what was seen in the DG studies, Cowansage et al. (2014) showed that stimulation of the ChR2 labeled neural ensembles naturally activated in the retrosplenial cortex at the time of contextual fear learning was sufficient to produce a fear response. The hippocampus has been the focus of intense study regarding its role in spatial and contextual learning; however, the cortical contribution has been less well characterized. This study shows that at the time of learning, a neural ensemble representing the context forms in the neocortex as well as the hippocampus. To investigate the circuit relationship and relative contribution of these two neural representations of context, the investigators inactivated the hippocampus pharmacologically and found that, although this produced the well-characterized amnesia for natural context recall (Kim and Fanselow 1992), this effect could be overcome by direct stimulation of the retrosplenial neural ensembles. This shows that the retrosplenial contextual representation is downstream from and functions independent of the hippocampal representation.

Early studies by Penfield (1968) showed that, in a small number of cases, direct brain stimulation resulted in apparent perceptions or memories. Although it is not clear whether these results represent normal or pathological responses or actual past experience, they do suggest that direct stimulation of the brain can, in some cases, produce a coherent psychological representation. Because mice are unable to report subjective experience, it is difficult to know the psychological effect of artificial stimulation. To provide some insight into this question, Cowansage et al. (2014) compared the activation of downstream circuit elements with both natural recall (exposure to the context in which animals were shocked) and artificial ChR2 stimulation-induced recall. Using the catalase approach they found that stimulation of retrosplenial ctxA ensembles and natural memory retrieval in ctxA-activated overlapping subsets of neurons in the entorhinal cortex, as well as basolateral and central amygdala, suggesting that the two methods of producing memory retrieval are processed in a similar manner. This is perhaps the best insight that can be achieved in a nonverbal species; whatever is experienced with artificial stimulation in a mouse, we can say that it produces a state of brain activity that is similar to posttraining context exposure, which we define as natural memory retrieval.

The artificial stimulation of small ensembles of neurons in the DG region of the hippocampus allows the retrieval of contextual memories, presumably by activating more distributed ensembles of neurons in downstream hippocampal and cortical regions. Given the highly parallel nature of neural connectivity and processing, it is possible that although these ensembles may be sufficient for retrieval, they may not be necessary because other pathways could compensate. The question of necessity was addressed in two recent studies using a light-gated proton pump from archaeabacteria (ArchT) (Chow et al. 2010) to hyperpolarize and silence cflg-tagged neural ensembles in experiments analogous to those described above with ChR2 (Denny et al. 2014; Tanaka et al. 2014). Both papers also used contextual fear conditioning and examined the requirement for ensembles in three different hippocampal regions, the DG, CA3, and CA1. Expression of ArchT was driven into neurons that were active during learning in context fear conditioning and these ensembles were subsequently silenced in either the DG, CA3, or CA1 during memory retrieval. In each case, the silencing impaired the retrieval of the contextual fear memory. These results show that in each of the three major hippocam-
pal subregions, the reactivation of the ensemble of neurons active during initial learning is necessary for the subsequent retrieval of the memory. This is consistent with the view that retrieval involves the recruitment of patterns of brain activity produced during initial learning.

In associative conditioning tasks, the unconditioned stimulus (US) generally has an emotional valence; for example, food is generally positive or rewarding, whereas shock, the US in fear conditioning, is aversive. Two recent studies examined the circuits that encode emotional valence in the hippocampus and amygdala using the cFos-tagging approach. The first study extended on the finding of a neural representation for context fear in the DG (Redondo et al. 2014). The investigators genetically tagged DG neurons active during fear conditioning with channelrhodopsin (ChR2), and when stimulated with light these neurons produced a fear response. They next asked whether the emotional valence of this fear memory could be reversed by pairing light stimulation of the DG ensembles with a positive valence US. Following the pairing, animals that previously avoided the context in which they received a shock now sought out the context, suggesting that there had been a switch in emotional valence associated with the context. This result is consistent with the view that the DG neural ensemble encodes a contextual representation (i.e., substitutes for the conditioning box), so that when the artificial activation of these ensembles is paired with a positive or negative stimulus, the box itself acquires this emotional valence as occurs in natural associative conditioning. Interestingly, when the same experiment was conducted using stimulation of cFos-labeled amygdala neural ensembles, the switch in valence was not observed. This suggests that the amygdala ensembles activated during learning are hard wired for a particular emotional valence.

The amygdala circuitry for emotional valence was explored in another recent paper (Gore et al. 2015). The study genetically introduced ChR2 into neurons activated by either a positive valence US (nicotine), or a negative valence US (shock). The investigators found two sparse (approximately 3%) and distinct populations of basolateral amygdala neurons were labeled by the stimuli, with little overlap in the cells activated by positive and negative valence stimuli. They went on to show that the artificial stimulation of these neurons could serve as a US (positive or negative valence depending on the cell population stimulated) in an array of associative and instrumental conditioning tasks. To examine the effects of conditioning on the US neurons, animals were conditioned by pairing a CS with artificial stimulation of the US neurons to produce an associative memory. After this training, the CS alone was now able to activate the US neurons of the paired emotional valence, similar to what was seen in previous studies with fear conditioning (Reijmers et al. 2007). Taken together, this study presents a view in which the amygdala contains distinct populations of neurons that confer positive or negative valence to a stimulus, presumably because of their output targets. During associative conditioning, the CS develops the ability to activate the subpopulation of neurons that were activated by the US to which it was paired and those neurons output the appropriate emotional valence to guide behavior.

Biochemical and Cellular Studies in Active Circuits

The formation of memories presumably involves cellular and molecular changes at specific points within the processing stream of incoming sensory information such that it is altered to produce new behavioral outputs following learning. For example, in our fear-conditioning model, when exposed to a new environment (context), mice show exploratory behavior. However, if the animal received an aversive foot shock while in this context, their subsequent behavior to the same sensory cues of the context will be immobility or freezing. The physical instantiation of this context memory, the engram, is presumably reflected in cellular and molecular changes at specific points in processing of the contextual cues, which diverts or enhances activity to the fear-output circuits in the amygdala. Where these changes occur,
whether they are distributed or localized, their cellular and molecular nature and how they alter the processing of information from sensory input to motor output are unknown but required for a deep understanding of the engram in this behavioral model. The IEG-based genetic technology enables the targeting of cellular and biochemical studies specifically to active neurons as opposed to their inactive neighbors and has been used in several recent studies to look for specific changes that may be relevant to the underlying cellular mechanisms of memory.

Forms of activity-dependent synaptic plasticity, such as LTP and long-term depression (LTD), represent attractive candidates for a cellular mechanism of learning. The idea remains controversial and a detailed discussion of the issue can be found in Angelakos and Abel (2015) and Basu and Siegelbaum (2015) and in several recent reviews (Morris 2013; Kandel et al. 2014; Mayford 2014). Perhaps the best support for LTP and LTD as a synaptic mechanism of memory comes from a recent study in which pairing optogenetic stimulation of inputs to the amygdala with foot shock produced a fear memory in which the optogenetic stimulation served as the CS. When LTD and LTP were subsequently produced by the appropriate stimulation of the ChR2 positive fibers carrying the CS, the behavioral fear response showed a corresponding decrease and increase, respectively, providing a direct link between these forms of plasticity and behavior, albeit in a nonnatural memory. The molecular signaling involved in the production of LTP is complex but may converge on at least one common mechanism, the insertion of new 2-amino-3-5-methyl-3-oxo-1,2-oxazol-4-yl propanoic acid (AMPA)-type glutamate receptors into dendritic spines that have undergone plasticity leading to an increased postsynaptic response (Shi et al. 1999; Rumpel et al. 2005; Kessels and Malinow 2009).

To examine glutamate receptor trafficking specifically in circuits activated during learning, Matsuo et al. (2008) used the cfas-tTA-based genetic tagging approach to drive expression of a GFP-linked GluR1 receptor subunit. Animals received either standard contextual fear conditioning or an unpaired protocol in which they were exposed to the conditioning chamber and later received immediate foot shocks in a separate context. Novel context exposure induced GFP-GluR1 expression in approximately 25% of hippocampal CA1 neurons and this was not altered by the shock US, consistent with a role for the hippocampus in encoding contextual information. Twenty-four hours after conditioning, the investigators examined the distribution of GFP-tagged glutamate receptors on dendritic spines and found that the paired training caused an increase in receptor trafficking specifically to mushroom-type spines. Because the expression of the GFP-GluR1 transgene was induced by the training itself and the protein did not reach the dendrites for several hours, the results were consistent with the idea that the training produced a molecular change or tag at some synapses allowing them to capture the newly synthesized receptors when they arrived. The notion of synaptic tagging was developed in studies of LTP to explain how newly expressed genes, required for long-term maintenance of the plasticity and of memories, could influence only those synapses that underwent the initial short-term plasticity (Frey and Morris 1997, 1998; see also Martin et al. 1997 in Aplysia and Si and Kandel 2015). Although these results are purely correlative, they are consistent with a subtle associative change in hippocampal synaptic function following fear conditioning. They also serve to indicate some of the difficulties likely to be encountered in trying to identify and functionally test the role of molecular changes with learning in mammalian systems that are likely to occur in a limited group of neurons and on a limited number or type of synapse on those neurons.

One of the weakest links in the LTP-memory connection is the paucity of studies that directly observe synaptic potentiation with learning. This may be attributable to the presumed paucity of neurons and synapses that contribute to any single learning event. To overcome this problem, Whitlock et al. (2006) used an array of recording electrodes implanted in the apical dendrites of the hippocampal CA1 region to measure synaptic responses before and after
learning in an inhibitory avoidance task. They found a modest increase in synaptic response on 14 of 44 electrodes with the conditioning, indicating a sparse but localized potentiation. Several studies have reported LTP-like synaptic potentiation in the lateral amygdala with auditory fear conditioning (McKernan and Shin-nick-Gallagher 1997; Rogan et al. 1997; Rumpel et al. 2005). The ability to detect plasticity in this region may reflect the recruitment of a large number of neurons in this particular behavioral task. In one study, it was estimated that one-third of the lateral amygdala neurons showed plasticity following conditioning (Rumpel et al. 2005). The identification of plastic changes in more sparsely recruited networks may be more difficult.

This issue was recently addressed in a study using the c-fos-tagging approach to look for physiological changes in DG neurons activated with contextual fear conditioning, the same neurons that can produce an apparent contextual representation when artificially stimulated (Ryan et al. 2015). Whole-cell recording of the c-fos-positive DG neurons from fear-conditioned animals showed an increase in excitatory postsynaptic current (EPSC) amplitude and in the ratio of AMPA/NMDA current relative to their negative neighbors. The c-fos-positive neurons also showed an increase in dendritic spine density. Both of these changes were blocked by immediate posttraining administration of the protein synthesis inhibitor anisomycin, which also blocked the contextual fear memory. Taken together, these results suggest that the active ensemble of neurons undergo a learning-related increase in synaptic strength and connectivity. However, a similar change was also found in neurons from the animals that received only context exposure. This could reflect plasticity resulting from the encoding of a novel context, which might be the expected function of hippocampal ensembles. Alternately, it could reflect differences in a constitutively active subpopulation of neurons as the results are quite similar to what has been described previously in c-fos-positive cortical neurons from animals not receiving any environmental stimuli (Yassin et al. 2010).

Memory Allocation

Are the neurons that are recruited to encode a specific representation or memory determined exclusively by their preexisting connectivity and synaptic responses within the circuit processing the relevant sensory cues or is there a greater flexibility in the precise set of neurons recruited in any particular circumstance? A number of recent studies have been put forward to support a memory allocation model, which posits that in a given brain area many neurons have the potential to be recruited for the encoding of a particular stimulus and that the specific neurons are chosen for this task based on their preexisting state of excitability.

The initial idea for this model came from a study attempting to rescue memory deficits in a cyclic AMP response element-binding protein (CREB)-deficient mouse line (Han et al. 2007). Viral vectors carrying a wild-type copy of CREB were injected into the lateral amygdala of CREB-deficient animals and resulted in a complete rescue of the impairments in auditory fear conditioning, even though the virus infected only about 20% of the neurons. How could such a small number of restored cells support normal memory encoding? The viral construct carried a GFP marker to allow identification of the infected cells. When the investigators examined the neurons activated during a memory retrieval trial, using staining for the IEG arc, they found that almost all of the arc was expressed in GFP-positive cells, indicating that the infected neurons were selectively recruited to the fear memory representation. This was not an effect that was specific to the CREB mutant mouse line as viral elevation CREB levels in wild-type animals produce the same effect. In subsequent studies, it was shown that the specific silencing or lesioning of the CREB-elevated neurons impaired the fear memory showing that these specific neural ensembles contributed to expression of the memory (Han et al. 2009; Zhou et al. 2009).

How could CREB, a transcription factor implicated in the consolidation of memory, alter the recruitment of neurons during learning? The first clue came from the observation that CREB-transfected neurons showed a significant
increase in excitability (Zhou et al. 2009). This suggests that it was not necessarily CREB itself that was responsible but simply that neurons with elevated excitability at the time of learning are those more likely to be recruited to the representation of that event. This idea was tested by increasing neuronal excitability directly using expression of a dominant negative mutant of the potassium channel KCNQ2 or the hM3Dq receptor in viral transfection experiments analogous to those described above for CREB (Yiu et al. 2014). In both cases, increased excitability was sufficient to recruit the transfected neurons to the fear memory representation. Using the regulated hM3Dq system, it was further shown that the increase in excitability was effective only immediately before training but not after learning had occurred. Although these studies were restricted to the lateral amygdala and fear conditioning, similar effects have recently been described in the insular cortex with conditioned taste aversion, suggesting that memory allocation may be a more general phenomena (Sano et al. 2014). The memory allocation model is attractive as a mechanism for linking memories for events that are closely spaced in time (Silva et al. 2009). According to this view, neural activity and the formation of long-term memory evokes CREB-dependent transcription, which in turn opens a time window of increased excitability in those active neurons. When new information is acquired during this time window, it will be more likely to recruit neurons that participated in the previous learning event creating a link at the circuit level between events that are closely spaced in time.

CONCLUSIONS

One of the critical difficulties in studying information processing and its modification with learning in the mammalian brain is identifying and manipulating the neurons involved in representing any specific environmental stimulus or memory. A particular complex item, a place, or an individual, is likely represented in higher brain areas in the activity pattern of a relatively small and anatomically dispersed group of neurons. In this review, we have discussed the use of genetic tools in the mouse that allow the molecular identification and genetic modification of neurons that are electrically active (because of natural environmental stimuli) at a specific point in time. This allows experiments to focus on just those neurons that were active in response to a given stimulus. This has been used to identify cellular and molecular changes that are specific to active neurons relative to their inactive neighbors. Active neural ensembles in the DG and somatosensory cortex have been shown to have increased excitability, interconnectivity, and dendritic spine density. In the hippocampus, learning induced an increase in glutamate receptor trafficking specifically in neurons activated during context learning. Although the functional significance of these cellular changes is still unclear, the approach is revealing that active neural ensembles show underlying molecular differences compared to apparently identical neurons in the same brain region that are not activated by behaviorally relevant stimuli.

In a second line of research, effectors of neural activity were introduced into naturally active ensembles to explore the nature of neural representations. Previous experiments were limited to recording the sensory- or learning-evoked activity of neurons and inferring a role in producing the psychological or perceptual manifestation of that sensory experience or memory. By introducing ChR2 into cfos-active neural ensembles during learning of a specific location in context fear conditioning, the activity of naturally activated neurons could actually be replayed artificially to directly test the relevance of that specific pattern of neural activity in encoding that specific context. The results suggest that even with the relatively crude IEG-based labeling of active ensembles, their reactivation via CHR2 produced a coherent representation/perception of the context. This result has been found in both the hippocampus and retrosplenial cortex with complementary studies in the hippocampus showing that suppression of these ensembles is sufficient to impair context recall. This is the first direct link between dispersed patterns of sensory-evoked neural activity and a coherent perceptual representation. In relation to learning and memory, these results should be
viewed as an important first step in identifying the relevant principles of information processing rather than identification of the elusive “engram.” The studies show that the stimulation of the correct pattern of neurons in one brain area is able to substitute for, or represent, a complex environmental stimulus (the context box), not that these neurons undergo the physical changes necessary to link that context to the fear-evoking shock. For example, if one could stimulate retinal neurons in precisely the same pattern as when an animal explored a specific context, then the animal would presumably perceive that they were in the box, and freeze if fear conditioned, but we would not suggest that the engram for that memory lies in the retina.

The mammalian brain is an immensely complex information-processing system that transforms an ongoing stream of sensory information into evolutionarily advantageous behavioral outputs. Learning is the process by which specific environmental contingencies (e.g., paring a specific location with an aversive stimulus) alter this processing to produce altered behavioral outputs. The IEG-based genetic tagging approach provides the ability to focus studies specifically on those neurons that are activated by any given set of environmental contingencies. It should prove useful in future studies to identify and genetically test specific models for molecular changes that may underlie learning, such as Hebbian forms of plasticity. A deep understanding of learning and memory will require not just identifying these molecular mechanisms and their location(s) in the brain, but also identifying how they alter the processing of information to evoke the learned responses. The new approaches for recording and manipulating active neural ensembles described in this and previous work offers an experimental framework in which this type of question can begin to be addressed.

REFERENCES


M. Mayford and L. Reijmers


Qian Z, Gilbert ME, Colicos MA, Kandel ER, Kuhl D. 1993. Tissue-plasminogen activator is induced as an immedi-


Radulovic J, Kammermeier J, Spiess J. 1998. Relationship between fos production and classical fear condition-


Saint Marie RL, Luo L, Ryan AF. 1999. Spatial representation of frequency in the rat dorsal nucleus of the lateral lem-


# Exploring Memory Representations with Activity-Based Genetics

Mark Mayford and Leon Reijmers

*Cold Spring Harb Perspect Biol* published online December 18, 2015

## Subject Collection: Learning and Memory

<table>
<thead>
<tr>
<th>Large-Scale Fluorescence Calcium-Imaging Methods for Studies of Long-Term Memory in Behaving Mammals</th>
<th>The Role of Functional Prion-Like Proteins in the Persistence of Memory</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pablo Jercog, Thomas Rogerson and Mark J. Schnitzer</td>
<td>Kausik Si and Eric R. Kandel</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Exploring Memory Representations with Activity-Based Genetics</th>
<th>Working Memory: Maintenance, Updating, and the Realization of Intentions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mark Mayford and Leon Reijmers</td>
<td>Lars Nyberg and Johan Eriksson</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>The Origins and Organization of Vertebrate Pavlovian Conditioning</th>
<th>Memory Retrieval in Mice and Men</th>
</tr>
</thead>
<tbody>
<tr>
<td>Michael S. Fanselow and Kate M. Wassum</td>
<td>Aya Ben-Yakov, Yadin Dudai and Mark R. Mayford</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>The Corticohippocampal Circuit, Synaptic Plasticity, and Memory</th>
<th>Reconsolidation and the Dynamic Nature of Memory</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jayeeta Basu and Steven A. Siegelbaum</td>
<td>Karim Nader</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Motor Learning and the Cerebellum</th>
<th>Memory Consolidation</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>The Striatum: Where Skills and Habits Meet</th>
<th>Structural Components of Synaptic Plasticity and Memory Consolidation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ann M. Graybiel and Scott T. Grafton</td>
<td>Craig H. Bailey, Eric R. Kandel and Kristen M. Harris</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Molecular Genetic Strategies in the Study of Corticohippocampal Circuits</th>
<th>Associative Learning in Invertebrates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Christopher C. Angelakos and Ted Abel</td>
<td>Robert D. Hawkins and John H. Byrne</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Nonassociative Learning in Invertebrates</th>
<th>The Regulation of Transcription in Memory Consolidation</th>
</tr>
</thead>
<tbody>
<tr>
<td>John H. Byrne and Robert D. Hawkins</td>
<td>Cristina M. Alberini and Eric R. Kandel</td>
</tr>
</tbody>
</table>

For additional articles in this collection, see [http://cshperspectives.cshlp.org/cgi/collection/](http://cshperspectives.cshlp.org/cgi/collection/)

---

Copyrigh © 2015 Cold Spring Harbor Laboratory Press; all rights reserved