NeuroResource

Neuron

High-Throughput Mapping of Single-Neuron Projections by Sequencing of Barcoded RNA

Highlights

- Single-neuron projections can provide insight into inter-areal information flow
- Exploiting the speed of DNA sequencing allows high-throughput single-neuron tracing
- Using MAPseq, thousands of single-neuron projections can be mapped in 1 week
- Individual locus coeruleus neurons have idiosyncratic projection patterns

Authors

Justus M. Kebschull, Pedro Garcia da Silva, Ashlan P. Reid, Ian D. Peikon, Dinu F. Albeanu, Anthony M. Zador

Correspondence

zador@cshl.edu

In Brief

Kebschull et al. present MAPseq, a method that exploits DNA sequencing to allow the rapid tracing of thousands of single-neuron projections. Using this method, they find that individual locus coeruleus neurons have idiosyncratic projection patterns.

Kebschull et al., 2016, Neuron *91*, 975–987
September 7, 2016
http://dx.doi.org/10.1016/j.neuron.2016.07.036

CellPress

Neuron
NeuroResource

High-Throughput Mapping of Single-Neuron Projections by Sequencing of Barcoded RNA

Justus M. Kebschull,[1,2] Pedro Garcia da Silva,[2,3] Ashlan P. Reid,[2] Ian D. Peikon,[1,2] Dinu F. Albeanu,[2] and Anthony M. Zador[2,*]
[1]Watson School of Biological Sciences, Cold Spring Harbor, NY 11724, USA
[2]Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA
[3]Champalimaud Centre for the Unknown, 1400-038 Lisbon, Portugal
*Correspondence: zador@cshl.edu
http://dx.doi.org/10.1016/j.neuron.2016.07.036

SUMMARY

Neurons transmit information to distant brain regions via long-range axonal projections. In the mouse, area-to-area connections have only been systematically mapped using bulk labeling techniques, which obscure the diverse projections of intermingled single neurons. Here we describe MAPseq (Multiplexed Analysis of Projections by Sequencing), a technique that can map the projections of thousands or even millions of single neurons by labeling large sets of neurons with random RNA sequences ("barcodes"). Axons are filled with barcode mRNA, each putative projection area is dissected, and the barcode mRNA is extracted and sequenced. Applying MAPseq to the locus coeruleus (LC), we find that individual LC neurons have preferred cortical targets. By recasting neuroanatomy, which is traditionally viewed as a problem of microscopy, as a problem of sequencing, MAPseq harnesses advances in sequencing technology to permit high-throughput interrogation of brain circuits.

INTRODUCTION

Neurons transmit information to distant brain regions via long-range axonal projections. In some cases, functionally distinct populations of neurons are intermingled within a small region. For example, nearby hypothalamic nuclei regulate basic drives including appetite, aggression, and sexual attraction (Kennedy et al., 2014; Sternson, 2013), and neurons from these nuclei project to distinct targets. In visual cortical area V1, responses to visual stimuli are matched to the properties of the higher visual areas to which the neurons project (Glickfeld et al., 2013; Movshon and Newsome, 1996). Findings such as these suggest that the information transmitted by individual neurons may be tailored to their targets. Such selective routing of information requires an anatomical substrate, but there is currently no high-throughput method for determining the diverse projection patterns of individual neurons.

At present, there is a steep tradeoff between throughput and resolution in anatomical approaches to mapping long-range connections. In conventional anterograde brain mapping studies, a fluorescent or enzymatic label is used to enable visualization of cell bodies and distal projections by light microscopy. Bulk techniques query the projections of many neurons labeled at a single injection site and thus sample the aggregate architecture of an entire neuronal population. There have been several large-scale efforts, including the Allen Brain Projection Atlas (Oh et al., 2014) and the iConnectome (Zingg et al., 2014), to systematically map mesoscopic connectivity. Although fast, such bulk methods obscure the diversity of the many projection neurons labeled in any one experiment. Consider, for example, a single source area that projects to three downstream areas (Figure 1A). This projection pattern enables neurons in the source area to send information to the three downstream areas. However, identical bulk projection patterns could arise in multiple ways: (1) from a one-to-one architecture, in which each neuron targets only a single downstream area (left); (2) from an all-to-all architecture, in which each neuron targets every downstream area (middle); or (3) from a host of more complicated architectures (right). With conventional bulk labeling, these three projection patterns, which have different functional implications, are indistinguishable without further experimentation.

Several alternative methods have been developed to complement conventional anterograde bulk labeling approaches. For example, genetically defined subpopulations of neurons within an area can be targeted by expressing a marker such as Cre recombinase (Gong et al., 2007; Harris et al., 2014; Huang, 2014). Similarly, subpopulations can be defined using retrograde (Lima et al., 2009; Oyibo et al., 2014; Wickersham et al., 2007a) or transsynaptic viruses (DeFalco et al., 2001; Wickersham et al., 2007b). However, because such approaches rely on positing defined subpopulations, they cannot easily be used to screen for the myriad possible complex projection patterns neurons might exhibit.

The most general and unbiased approach to distinguishing among the architectures in Figure 1A relies on single-neuron anterograde tracing. Current methods for achieving single-neuron resolution require individually labeling one or, at most, a few cells per brain (Economo et al., 2016), a labor-intensive approach that affords high resolution at the cost of low throughput. Although single-neuron tracing can be multiplexed by labeling individual neurons with different colored fluorophores (Ghosh et al., 2011; Livet et al., 2007), in practice the extent of

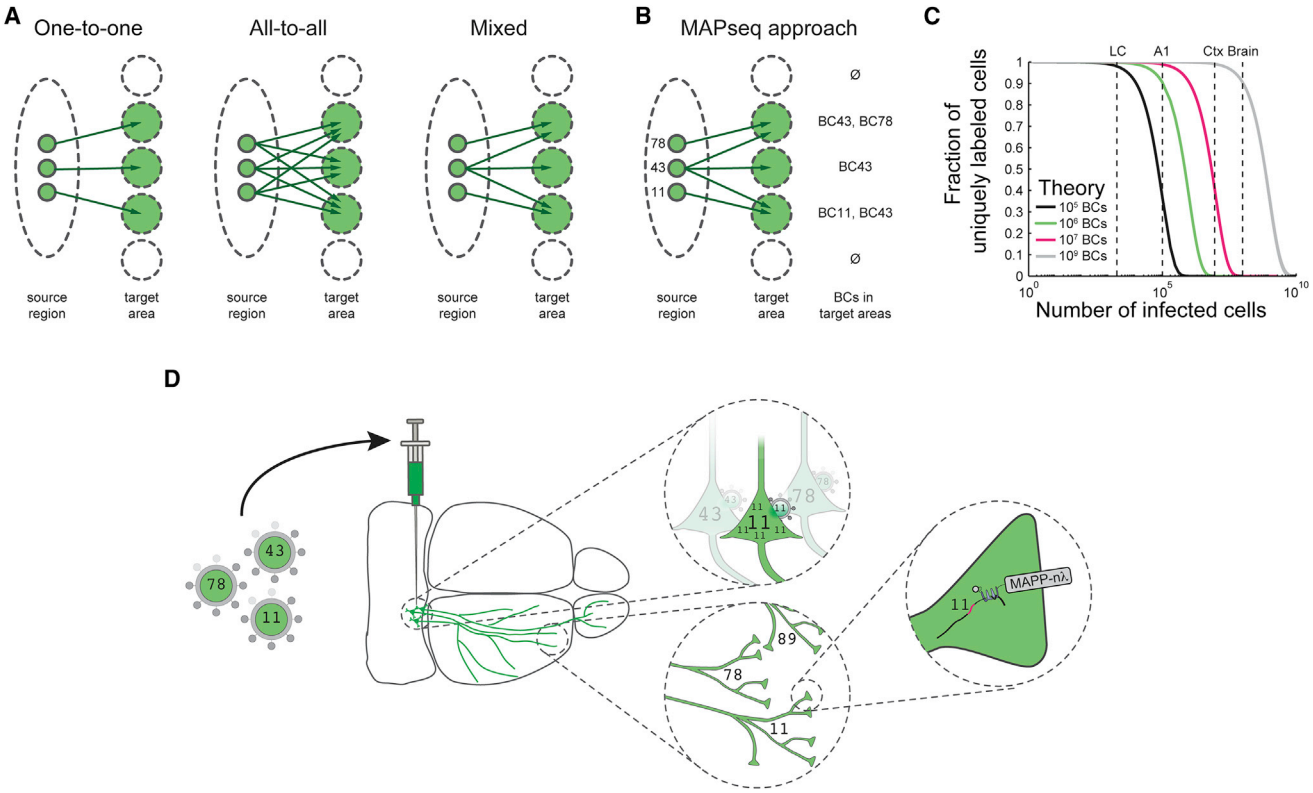

Figure 1. Barcoding Allows High-Throughput Single-Neuron Tracing

(A) Identical bulk mapping results can arise from different underlying projection patterns.

(B) Single-neuron resolution can be achieved by randomly labeling neurons with barcodes and reading out barcodes in target areas.

(C) The expected fraction of uniquely labeled cells is given by $F = (1\text{-}1/N)^{(k-1)}$, where N is the number of barcodes and k is the number of infected cells, assuming a uniform distribution of barcodes. The number of neurons for various mouse brain areas is indicated according to references (Herculano-Houzel et al., 2006; Schüz and Palm, 1989) (A1, primary auditory cortex; Ctx, neocortex).

(D) In MAPseq, neurons are infected at low MOI with a barcoded virus library. Barcode mRNA is expressed, trafficked, and can be extracted from distal sites as a measure of single-neuron projections.

multiplexing is limited by the number of colors—at most 5–10—that can be resolved by microscopy.

Here we describe MAPseq (Multiplexed Analysis of Projections by Sequencing), a novel approach in which the speed and parallelization of high-throughput sequencing are exploited for brain mapping (Zador et al., 2012). MAPseq achieves multiplexing by using short, random RNA barcodes to uniquely label individual neurons (Mayer et al., 2015; Walsh and Cepko, 1992; Zador et al., 2012) (Figure 1B). The key advantage of using barcodes is that their diversity grows exponentially with the length of the sequence, overcoming the limited diversity of the resolvable color space. The pool of unique barcode identifiers is effectively infinite; even a 30-nt sequence has a potential diversity of $4^{30} \approx 10^{18}$ barcodes, far surpassing the $\sim 10^8$ neurons in the mouse brain (Herculano-Houzel et al., 2006). Because high-throughput sequencing can quickly and inexpensively distinguish these barcodes, with MAPseq we can potentially read out the projections of thousands or even millions of individual neurons in parallel within a single brain (Figure 1C).

In MAPseq, we uniquely label neurons in a source region by injecting a viral library encoding a diverse collection of barcode sequences. The barcode mRNA is expressed at high levels and transported into the axon terminals at distal target projection regions (Figure 1D). To read out single-neuron projection patterns, we then extract and sequence barcode mRNA from the injection site, as well as from each target region of interest. Spatial resolution of MAPseq is limited mainly by the precision of target dissection. Although MAPseq, like GFP tracing, does not distinguish fibers of passage, we minimize their contribution by avoiding large fiber bundles during the dissection of target areas. Using this procedure, the brain-wide map of projections from a given area can be determined in less than a week. By reformulating projection mapping as a problem of sequencing, MAPseq harnesses advances in high-throughput sequencing to permit efficient single-neuron circuit tracing.

RESULTS

As a proof of principle, we applied MAPseq to the locus coeruleus (LC), a small nucleus in the brainstem that is the sole source of noradrenaline to the neocortex (Foote and Morrison, 1987). Early bulk tracing experiments revealed that the LC projects broadly

throughout the ipsilateral hemisphere, leading to the view that the LC broadcasts a generalized signal that modulates overall behavioral state (Foote and Morrison, 1987; Foote et al., 1983; Loughlin et al., 1982; Waterhouse et al., 1983). This view has recently been supported by more sophisticated retrograde bulk tracing experiments, which reinforce the idea that LC neurons project largely indiscriminately throughout the entire ipsilateral hemisphere (Schwarz et al., 2015). However, other reports have challenged this model. Using double retrograde labeling methods, these experiments uncovered separate populations of LC neurons projecting to different areas of cortex (Chandler et al., 2014; Chandler and Waterhouse, 2012), raising the possibility that the LC exerts differential control over different cortical areas. To address this controversy, we applied MAPseq to LC to obtain a large number of projection patterns at single-neuron resolution.

In what follows, we first show that long-range projections of neurons can be efficiently and reliably determined using barcode mRNAs, the abundance of which we interpret, like GFP intensity, as a quantitative measure of projection strength. Next, we establish the theoretical and practical foundations of randomly labeling large numbers of neurons with a viral barcode library, critical for ensuring single-cell resolution for MAPseq. We then apply MAPseq to the LC and find that individual neurons have a variety of idiosyncratic projection patterns. Some neurons project almost exclusively to a single preferred target in the cortex or olfactory bulb, whereas others project more broadly. Our findings are consistent with, and reconcile, previous seemingly contradictory results about LC projections. Finally, we show that MAPseq can be multiplexed to two and potentially many injections in the same animal, which will allow the projection patterns from many brain areas to be determined efficiently and in the same brain without the need for registration across animals.

Using RNA to Trace Neurons

Conventional neuroanatomical tracing methods rely on filling neurons with dyes or proteins so that neural processes can be resolved by microscopy. An implicit assumption of these techniques—albeit one that has rarely been rigorously tested—is that the tracer fills the neuron abundantly and uniformly so that the strength of the signal corresponds to the quantity of labeled neural process, independent of distance from the soma. For barcode mRNAs to act as a comparable label in MAPseq, we sought to maximize the abundance and uniformity of barcode mRNA in distal processes. We used two strategies to achieve this goal.

First, we expressed a modified presynaptic protein that was designed to specifically bind to and transport barcode mRNA into axon terminals. We engineered this protein, which we denote MAPP-nλ, as part of a larger project aiming to read out synaptic connectivity using mRNA. To generate MAPP-nλ, we began with pre-mGRASP, a protein engineered to localize at the presynaptic terminal due to fusion with trafficking signals from the endogenous presynaptic protein NRXN1β (Kim et al., 2011). We then inserted four copies of the nλ RNA binding domain (Daigle and Ellenberg, 2007) into the cytoplasmic domain of the protein. Derived from the λ phage λ_N protein, the nλ domain is a 22-aa peptide that strongly and specifically binds to a 15-nt RNA hairpin, termed boxB. We added four copies of the boxB hairpin to the barcode mRNA, ensuring coupling of MAPP-nλ to the barcode mRNA and thus transport of barcode mRNA into axon terminals (Daigle and Ellenberg, 2007). Second, we delivered the barcode sequence using recombinant Sindbis virus, a virus that can rapidly achieve very high expression levels (Ehrengruber, 2002). We used a novel Sindbis packaging system that, unlike previous systems, is both neurotropic and propagation incompetent (Kebschull et al., 2016) (Figures S1A–S1G; Supplemental Information, note 1, available online). All components necessary for MAPseq are expressed from a dual promoter virus that generates two subgenomic RNAs (Figure 2A). The first encodes the MAPP-nλ protein. The second RNA encodes a random 30-nt barcode, as well as the boxB sequence, downstream of a GFP marker (Figures S1H and S1I). We reasoned that combining these two strategies would maximize our ability to reliably detect barcode mRNA in distal processes.

We injected barcoded virus into right LC (Figures S1J–S1L) and examined barcode localization by in situ RNA hybridization 44 hr after injection. We observed robust barcode mRNA localization in the soma and neuronal processes, in a pattern similar to that of co-expressed GFP (Figure 2B). This suggested that barcode mRNA could effectively fill local neuronal processes.

To determine whether the barcode mRNA fills distal neuronal processes uniformly, we exploited the particular anatomy of LC projection neurons. LC neurons that project to cortex send their processes all the way to the rostral end of the brain, before changing direction and moving caudally to innervate cortical areas (Figure 2C). Axons that project to visual cortex are therefore approximately twice as long as those that project to frontal cortices. From bulk tracing studies, it is known that LC innervation is homogeneous along the rostro-caudal axis (Schwarz et al., 2015; Waterhouse et al., 1983). Thus, if barcode mRNA were not efficiently transported to distal processes, we would expect to find more barcode mRNA in rostral regions of cortex. To assess this, we injected barcoded virus into LC, cut 300 μm coronal slices of the entire cortex (Figure 2D), and quantified the amount of barcode mRNA from each ipsilateral and contralateral slice. Consistent with previous results using GFP and other tracing methods (Schwarz et al., 2015; Waterhouse et al., 1983), we found approximately uniform projections throughout the ipsilateral cortex ($p = 0.972$ F-statistic versus constant model; Figure 2E); in particular, we found no evidence that distal processes were more weakly labeled than proximal processes. Also consistent with previous results, we observed much weaker (30.6-fold less; $p = 4 \times 10^{-31}$, paired Student's t test) projections to the contralateral cortex (Figure 2F). These results suggest that barcode mRNA fills distal and proximal processes with about equal efficacy so that the barcode mRNA can be interpreted in the same way as the fluorophores and dyes used in conventional tracing studies.

Unique Labeling of Neurons with Barcodes

In conventional single-neuron tracing, the main challenge to multiplexing is the low diversity of labels (fluorophores or enzymes) available to disambiguate individual neurons. To overcome this challenge, MAPseq labels neurons with short, random RNA barcodes delivered by infection with a diverse viral library. Ideally, each labeled neuron should have exactly one unique

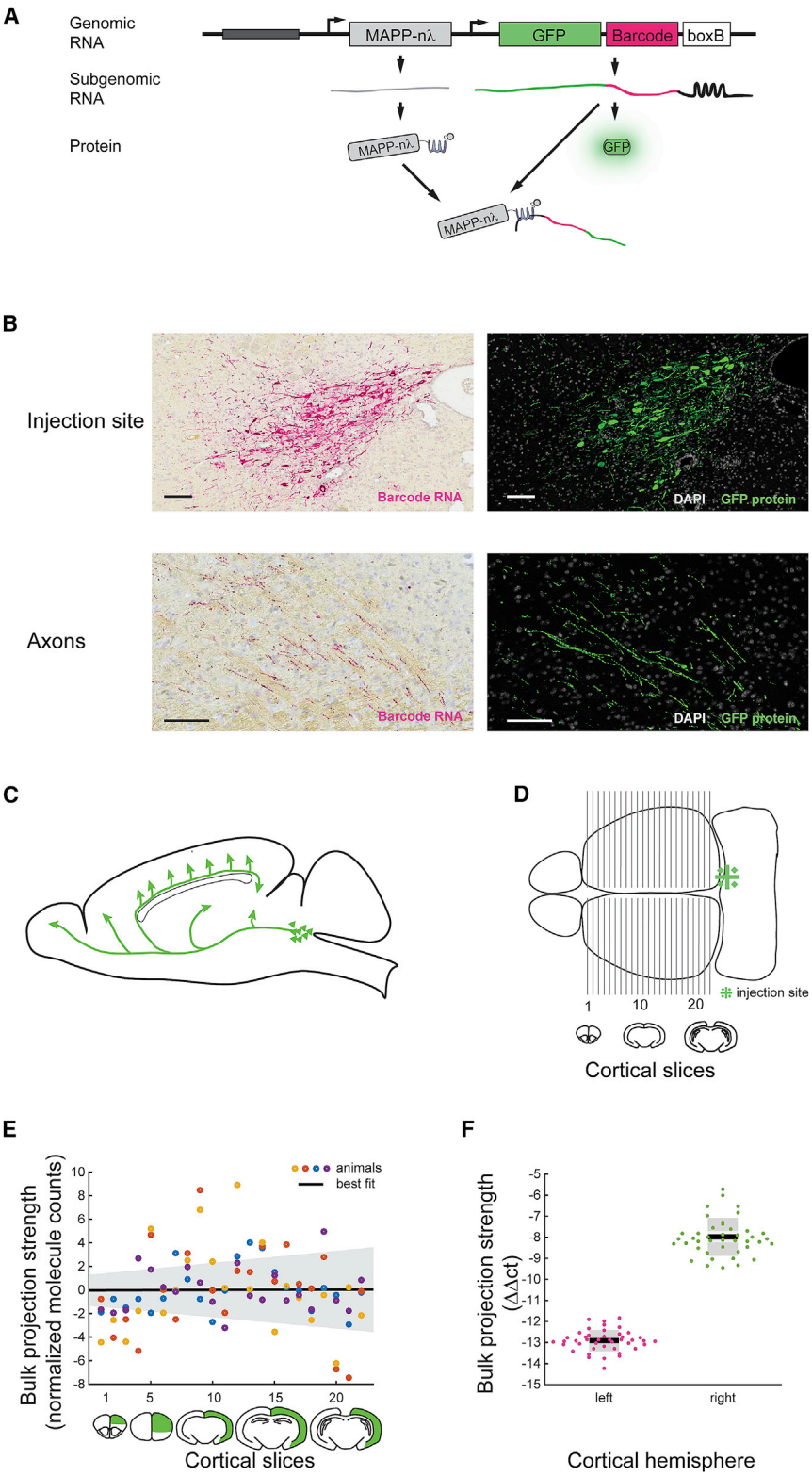

Figure 2. Barcoded Sindbis Virus Can Be Used for Projection Mapping

(A) A dual promoter Sindbis virus was used to deliver barcodes to neurons. The virus encoded GFP, barcodes, and MAPP-nλ.

(B) Barcode mRNA labeling of LC neurons is comparable to GFP labeling of these neurons in an adjacent 6 μm slice both at the injection site (top) and in the axon tract (bottom). Scale bar, 100 μm. Representative data from three animals.

(C) Axons from LC project rostrally from the cell body, before changing direction and innervating cortex. LC axons that project to frontal cortices have thus traveled only about half as long as axons innervating visual cortex.

(D) We injected right LC with MAPseq virus and dissected cortex along the anterior-posterior axis as shown.

(E) Bulk projection strength of LC to ipsilateral cortex as measured by barcode mRNA is independent of the anterior-posterior position of the cortical slice, suggesting a uniform RNA fill of LC axons. N = 4. The shaded area indicates the 95% confidence interval of the fit.

(F) qPCR for barcode mRNA shows approximately 30× stronger LC projections to ipsi- than to contralateral cortex. N = 2 animals and 21 cortical slices per animal. BC, barcodes. The y axis displays ΔΔct values, which are equivalent to the $\log_2$(foldchange of barcode mRNA per sample) normalized to β-actin levels in each sample and to the amount of barcode mRNA in the injection site of each animal (Livak and Schmittgen, 2001). Individual data points are plotted. The mean is indicated by a horizontal bar, SD error bars by a light gray area, and 95% confidence interval by a dark gray area.

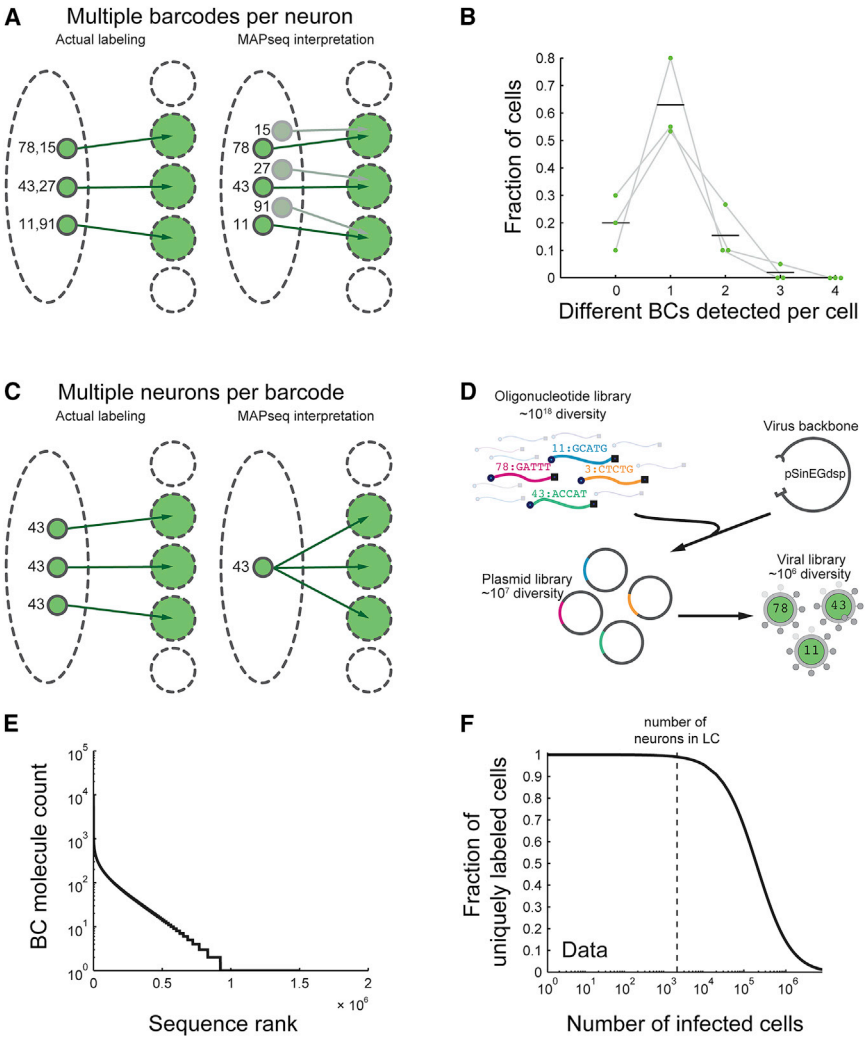

Figure 3. Random Labeling of Neurons with a Barcoded Virus Library Can Achieve Unique Labeling of Many Neurons

(A) When single neurons are labeled with several barcodes, MAPseq will overestimate of the number of neurons identified, but will not distort the projection patterns recorded for individual neurons.

(B) Single-cell isolation of GFP-positive, barcoded neurons, followed by sequencing of their barcode complement, reveals a low chance of double infection. We interpret neurons for which no barcodes were recovered as technical failures of cell isolation, rather than biological phenomena. N = 3 animals. Mean and individual data points are plotted.

(C) When several neurons share the same barcode, MAPseq misinterprets this as a single neuron whose projection pattern is given by the union of the projection patterns of the two infected neurons.

(D) High-diversity Sindbis virus libraries are produced by shotgun cloning random oligonucleotneuron_13280_gr1_4c.tif - ides into a plasmid followed by virus production.

(E) The virus library used in this work has a diversity of $\sim 10^6$ different barcodes (BC), but the distribution was non-uniform. The sequence rank is a number that ranges from 1 to the total number of barcodes, where 1 corresponds to the most abundant sequence, 2 to the second most abundant, and so on.

(F) Based on the empirically observed non-uniform barcode distribution, we determined that the virus library used is sufficiently diverse to uniquely label all of LC with low error rate.

barcode. Here we consider the factors that could lead to deviations from this ideal scenario: (1) more than one barcode per neuron (multiple labeling) and (2) more than one neuron per barcode (non-unique or degenerate labeling). As discussed below, deviations resulting from multiple labeling are much less of a concern than those resulting from degenerate labeling.

A neuron may express more than one barcode if it is infected by more than one viral particle. Such multiple labeling will lead to an overestimate of the number of neurons identified, but will not distort the projection patterns recorded for individual neurons (Figure 3A). Furthermore, even estimates of the relative abundances of different neuronal classes will, on average, be accurate. Assume, for example, that two neurons A and B are each labeled with ten different barcodes. In this scenario, MAPseq will discover ten instances of neuron A and ten of B, but even though the absolute number of neurons is incorrect, the fact that neurons A and B have distinct projection patterns, and that these patterns occur in a 1:1 ratio, is accurately inferred. Thus, multiple labeling will not, on average, lead to mischaracterization of neuronal classes or of their relative frequency in the population.

Nevertheless, to simplify the interpretation of MAPseq results, we sought to minimize the MOI by titrating the concentration and volume of virus injected. To estimate the MOI, we isolated a total of 45 individual neurons from three animals injected with MAPseq virus into the right LC and sequenced the barcodes within each neuron. On average, infected LC neurons contained 1.2 ± 0.1 barcodes each, implying MOI of 0.43 (Figure 3B). Only 21% ± 11% of neurons contained more than one barcode, with most of these neurons carrying two barcode sequences and only 1.7% ± 2.9% of neurons containing three barcode sequences.

The second deviation from the ideal scenario is non-unique labeling. If two neurons share the same barcode, then MAPseq misinterprets this as a single neuron whose projection pattern is given by the union of the projection patterns of the two infected neurons (Figure 3C). The probability that two neurons are infected by the same barcode depends on the number of infected cells relative to the number of available barcodes. Trivially, if the number of infected cells is larger than the number of available barcodes, unique labeling of all neurons cannot be achieved. Conversely, if the number of available barcodes is much higher than the number of infected cells, every neuron will be labeled with a different barcode purely by chance.

To determine whether our barcode diversity was sufficient to ensure unique labeling, we first quantified the number of neurons in the LC. We counted 1,985 ± 132 (N = 6 animals) neurons expressing tyrosine hydroxylase, a noradrenergic marker. The size of this neuronal population is approximately 15 orders of magnitude smaller than the theoretical diversity ($4^{30} \approx 10^{18}$) of a library of 30-nt barcodes, so in theory, unique labeling would be virtually certain. In practice, however, the actual diversity of a virus library is limited by bottlenecks in plasmid and virus generation (Figure 3D), so further analysis was required to determine whether the viral library was sufficiently diverse.

We therefore computed how the fraction of uniquely labeled neurons scaled with the diversity of the viral library. This problem is formally equivalent to a generalization of the classic "birthday problem," which concerns the probability that in a group of k people (neurons), some pair will have the same birthday (barcode). Assuming that all barcodes are equally abundant in the library, we can express the expected fraction F of uniquely labeled neurons as

$$F = \left(1 - \frac{1}{N}\right)^{k-1},$$

where k is the number of infected neurons (assuming one barcode per cell), and N is the barcode diversity (see Supplemental Information, note 2). Thus, if k = 1,000 LC neurons were infected with a library of diversity of $N = 10^6$, on average 99.9% of all neurons would be labeled uniquely. This expression holds only for a library of equally abundant barcodes; if some barcodes are overrepresented, the fraction F of uniquely labeled neurons decreases (in the same way that if birthdays tend to fall on a particular day, the probability of finding a shared birthday in a group increases; Munford, 1977). However, analysis of the actual viral barcode library determined by sequencing (Figure 3E) revealed that in practice, these deviations from uniformity had only a minor effect (Figure S2A). Thus, under our conditions, the vast majority (>99%) of neurons will be uniquely labeled, even taking into account the uneven barcode distribution in the viral library (Figure 3F; Supplemental Information, note 2).

We also used a second, more empirical approach to estimate the extent of degenerate labeling. Since we used the same viral library to infect neurons in different animals, barcode sequences found in more than one animal represent degeneracy. We therefore looked for overlap in the recovered barcodes from four independent injections of the same virus library. Out of the 992 unique barcodes that we recovered from traced neurons, only three barcodes were present in more than one animal, and no barcode was present in more than two animals. This empirically measured rate of degenerate labeling is in close agreement with our expectations based on the theoretical considerations above. Moreover, two of the three repeated barcodes were among the most abundant barcodes in the virus library (Figure S2B), and would thus be expected a priori to have the highest probability of double labeling. This analysis provides an independent confirmation that the error rate due to non-unique labeling by the barcode library is very low in our experiments.

In addition to non-unique barcode labeling, MAPseq is subject to other errors that differ from those associated with conventional tracing approaches. We used several approaches to quantify these errors and find that the overall MAPseq error rate was low both for false positives (1.4% ± 0.8%, mean ± SE) and for false negatives (8.6% ± 6%, mean ± SE) (see Supplemental Information, note 3). MAPseq thus provides a reliable measure of axon projections.

Sequencing of Barcode mRNAs Reveals Diverse Single-Neuron Projection Patterns

The goal of MAPseq is to quantify the projection patterns of large populations of neurons in parallel. We therefore developed a method to determine the amount of each barcode in each dissected target (Figure 4A). Forty-four hours after injection of barcoded virus into right LC, we performed reverse transcription on barcode mRNA extracted from dissected target regions. To overcome distortions introduced during amplification (Kebschull and Zador, 2015), and to allow a precise count of barcode cDNA molecules, we designed reverse transcription primers to tag each individual barcode mRNA molecule with a random 12-nt unique molecular identifier (UMI). We also added a 6-nt slice-specific identifier (SSI) to allow multiplexing of samples within a single high-throughput sequencing flow cell. We then amplified, pooled, and sequenced these SSI-UMI-barcode cDNAs (Figure S3). We developed a conservative computational pipeline to minimize noise due to RNA contamination and to correct for sequencing and other errors (Supplemental Information, note 4). Finally, we converted barcode abundance in the target areas to a matrix of single-neuron projection patterns.

We used MAPseq to determine the projection patterns of a total of 995 barcodes labeled in four animals (249 ± 103 barcodes per animal), roughly corresponding to an equal number of LC neurons. For each animal, we analyzed the barcode mRNAs extracted and amplified from the olfactory bulb and from 22 coronal slices (300 μm) taken from the cortex ipsilateral to the LC injection (Figures 4B and 4C). Although, like conventional GFP tracing, MAPseq does not distinguish between synaptic connections and fibers of passage, we minimized the contribution of large fiber tracts in the white matter during dissection, so most of the mRNA barcode signal was likely generated from axons terminating in the regions of interest. Dissection of coronal slices allowed us to probe the organization of projections along the rostral-caudal axis, but we were insensitive to any additional structure along the medio-lateral axis. Because individual barcode cDNA molecules are tagged with a UMI before amplification, we obtained a precise quantification (subject to Poisson counting statistics; see Supplemental Information, note 5) of the amount of each barcode sequence in each dissected target. In this way, we could infer the projection strength—the density of axon per tissue volume—of each neuron to each coronal target area. For example, we recovered 223 copies of BC28 in slice 5, but none in slice 20, indicating that the projection strength to slice 5 is at least a factor of 200 higher than our detection floor (Figure 4D).

Inspection of the projection patterns revealed that in contrast to the simplest prediction from conventional bulk tracing, single neurons did not project uniformly throughout the ipsilateral hemisphere. Instead, neurons projected in diverse and idiosyncratic ways to specific targets, innervating some areas hundreds of times more strongly than others (Figure 4D). Some neurons

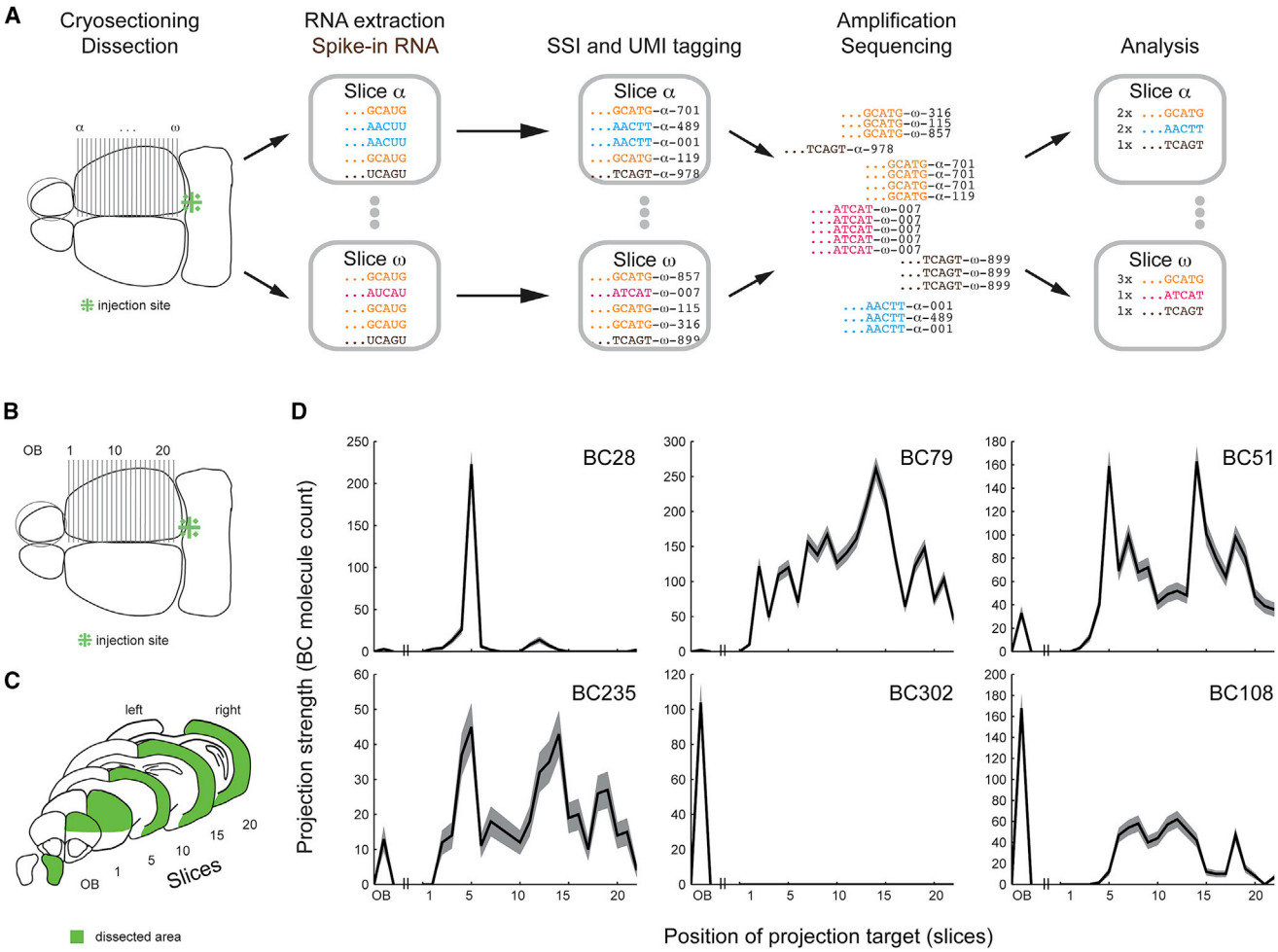

Figure 4. MAPseq Reveals Large Diversity of Projections from LC

(A) Barcode mRNAs from target areas are sequenced as described (SSI, slice specific identifier; UMI, unique molecular identifier).

(B–D) Barcodes from ipsilateral olfactory bulb and cortex (B and C) show projection patterns (D) with single or multiple peaks in cortex and/or olfactory bulb. The shaded area indicates Poisson error bars given by the square root of barcode (BC) counts per slice.

(e.g., BC28) showed specific projections to only a small part of cortex, whereas others (e.g., BC79) projected more broadly, or projected to multiple areas (e.g., BC51 or BC235). Projections to the olfactory bulb appeared independent of projections to cortex, with some neurons projecting exclusively to the olfactory bulb (e.g., BC302), some projecting exclusively to cortex (e.g., BC79), and others projecting to both (e.g., BC108).

The small fraction of multiply infected neurons revealed by single-cell sequencing (21% ± 11%; Figure 3B) provided a convenient internal measure of the reliability of MAPseq. Within each animal, we found pairs of very similar projection patterns, as would be expected if they arose from double labeling of the same neuron (Figures 5A and S4A). By comparing the similarity of projection patterns within and across animals, we estimated the total number of barcode pairs arising from doubly labeled neurons (18.6% ± 2%; Figures 5B and S4B). The fact that this estimate is in such close agreement with the independent estimate of the number of doubly labeled neurons from single-cell sequencing supports the view that MAPseq provides a robust measure of single-neuron projection profiles.

To assess the structure of the LC projection to cortex and olfactory bulb, we sorted all traced neurons by their maximum projection (Figure 6A). The maximum projections of individual LC neurons tile the entire cortex. To compare across the population, we normalized the total barcode count of the projection of each neuron to one, although interestingly there was no correlation between expression level in the LC and the maximal projection strength to cortex, as would have been expected if differences across neurons were dominated by expression level ($R = -0.06$; $p = 0.09$; Figure S5A). Only in the aggregate do these projections re-create the apparently homogeneous LC innervation of cortex previously described by bulk methods (Figure 2E). Consistent with previous results (Shipley et al., 1985), a considerable fraction (23% ± 4.7%) of all mapped neurons projected to the olfactory bulb.

We then asked if we could find structure, or even stereotypic projection cell types, in the single-cell dataset. We investigated

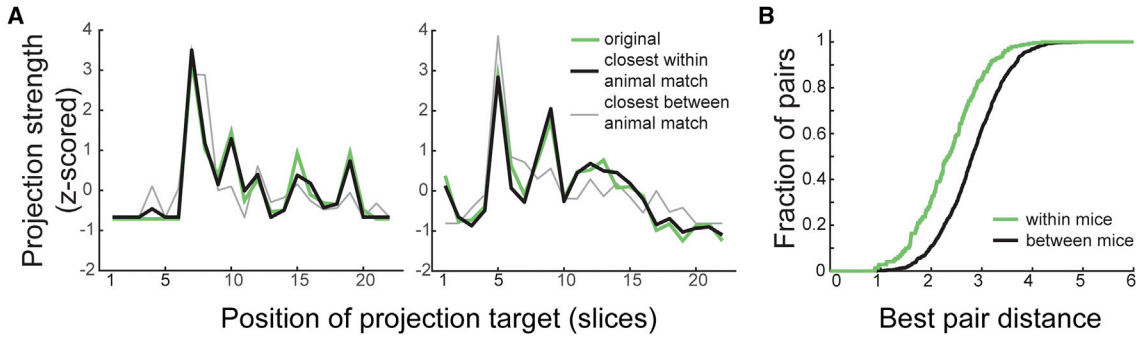

Figure 5. MAPseq Provides a Robust Readout of Single-Neuron Projection Patterns

(A) Two representative pairs of barcodes with projection patterns more similar than expected by chance for two distinct neurons, likely the result of double infection of a single neuron. The close agreement between the two barcode profiles indicates that MAPseq provides a reliable measure of projection patterns. The closest match across animals is indicated in gray for comparison.

(B) Cumulative distribution of distances between the best barcode pairs within one animal and across animals. The shift in the within-animal distribution reflects the higher fraction of closely matched projection profiles, consistent with double infection. Representative data from one animal.

the cortex-wide projection patterns of LC by reducing the dimensionality of the projection dataset using Euclidean distance-based t-SNE (t-distributed stochastic neighbor embedding) (Van der Maaten and Hinton, 2008). Neurons with maximum projections close to each other in physical space along the rostro-caudal axis also fall closely together in t-SNE space (Figure S5B), indicating that the location of the maximum projection target at least partially describes the individual neuron projection patterns. However, hierarchical clustering of the projection profiles of neurons that project to cortex did not uncover distinct cell classes (Figure S5C). Although we cannot rule out the possibility that there is further structure in the projection patterns that would be revealed by higher-resolution dissection, our data suggest the intriguing possibility that LC projections are equipotential for projecting to all targets, and the choice is arbitrary for each neuron. How the circuit might exploit such random connectivity raises interesting computational challenges.

Although many LC axons projected very strongly to a narrow target, these axons often sent minor collaterals to a much broader area, like a plant with a single main stalk and many minor growths. The average number of projection peaks per LC neuron was 1.6 ± 0.8 (Figure S5D), and the fall-off to half the maximum projection strength of individual neurons occurred on average in <300 μm (Figure 6B). Nonetheless, every cortically projecting neuron innervated on average 65% ± 23% of cortex at a detectable level (see cumulative distribution of projection widths in Figure 6C). Importantly, this broad, weak innervation of cortex cannot be explained simply by contamination of our dataset by fibers of passage. LC axons innervate cortex starting in the rostral end and moving caudally (Figure 3C). We would therefore expect signals from accidentally dissected fibers of passage to only precede strong projection targets of individual neurons along the rostro-caudal axis. However, we find that strong projection targets are both preceded and followed by low-level projections (e.g., see Figure 4D; BC79 and BC51). We therefore conclude that the observed weaker signals are not the result of fibers of passage but represent real, but weak, projections.

The fact that many neurons had a strong preferred target in cortex or olfactory bulb, but also projected weakly to a much broader area, provides a way to reconcile apparently conflicting results about the specificity of LC projections. Recent experiments in which retrograde viral labeling was combined with anterograde tracing of axons concluded that as a population, LC neurons project largely indiscriminately throughout cortex and the rest of the brain examined (Schwarz et al., 2015). However, using this approach, a neuron labeled retrogradely from a weak projection is indistinguishable from one labeled from a strong projection, so at the level of the population (i.e., after summing the projection patterns of strongly and weakly projecting neurons), it may appear that projections are nonspecific. Thus, although the results of this study may appear to contradict those obtained by MAPseq at single-neuron resolution, simulations of retrograde labeling in combination with anterograde bulk tracing based on our MAPseq dataset demonstrate that there is no contradiction (Figures S5E and S5F) at the bulk level.

MAPseq Scales to Several Injection Sites

MAPseq can readily be extended to determine the projections of two or more regions in a single animal. As a proof of principle, we injected the same library of MAPseq virus bilaterally into left and right LC, and dissected coronal slices of left and right cortex and the olfactory bulbs, as well as both injection sites (Figure 7A). Each barcode was expressed predominantly in either the left or right LC (Figure 7B); barcode expression at the site contralateral to the injection, due to contralaterally projecting fibers and/or contamination, is much lower. Thus, each barcode can be reliably assigned to the appropriate injection site. As expected, parallel injections recapitulated the projection pattern observed with single injections (Figure 7C). Multiplexing MAPseq to dozens of injections per animal may be feasible, reducing the labor and cost of brain-wide projection mapping efforts, and eliminating the need to map data from multiple animals to an average reference brain (Oh et al., 2014; Zingg et al., 2014).

DISCUSSION

We have described MAPseq, the first application of sequencing to neuroanatomical tracing. MAPseq is a simple, rapid, and

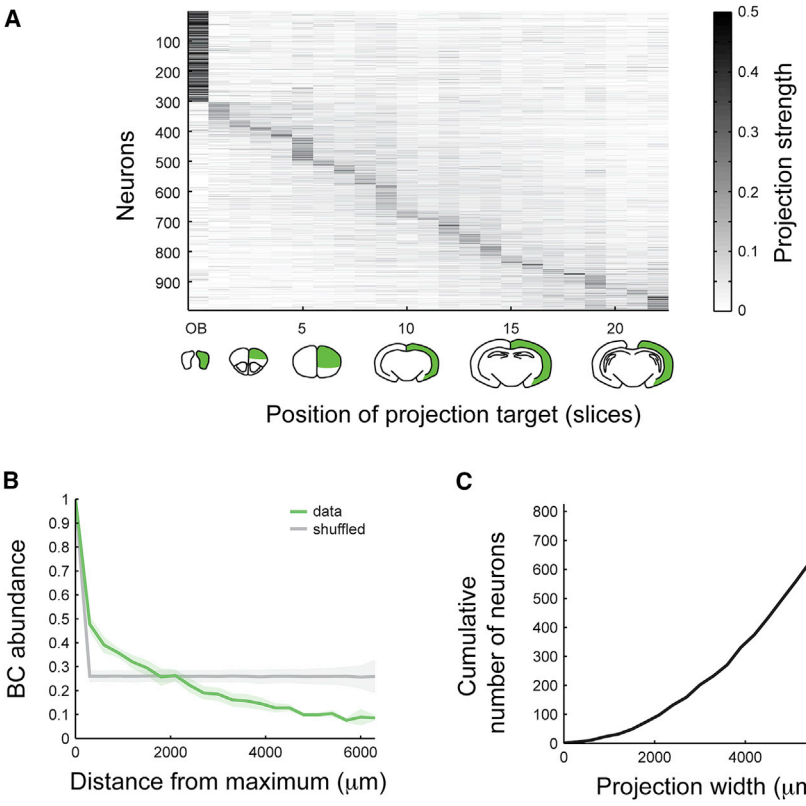

Figure 6. LC Neurons Tile Cortex with Their Maximum Projections, but Innervate Large Areas of Cortex at a Low Level

(A) A heatmap of all 995 projection patterns from four animals shows a strong diagonal component after sorting by maximum projection site. Barcode abundances are normalized to sum to one across target areas and are color coded as indicated.

(B) Average cortical drop-off rate from maximum for all barcodes shows a rapid drop-off and a structure that is different from the drop-off after randomly shuffling slices for all neurons. N = 4. The shaded area around the curve indicates the SD across animals.

(C) Cumulative distribution of cortical projection widths indicates a broad, low-intensity innervation of cortex by individual LC neurons. BC, barcode.

inexpensive approach to determining the projection patterns of myriad single neurons at one or more injection sites in a single animal. As a proof of principle, we applied MAPseq to the LC. In contrast to previous bulk labeling studies that reported diffuse and non-specific projections from the LC, this single-neuron resolution analysis reveals that individual LC neurons have idiosyncratic projection patterns with preferred cortical targets, and reconciles a controversy about the specificity of LC projection patterns. MAPseq, which complements rather than replaces conventional approaches, can readily be applied to other brain regions and organisms, and with further development may be combined with information about gene expression and neural activity.

The cost of sequencing the human genome was several billion dollars in 2003, but today it is less than one thousand dollars—a decrease of over six orders of magnitude in just over a dozen years (Hayden, 2014; Sheridan, 2014). This precipitous drop in sequencing costs continues unabated, at a rate faster even than Moore's law (the rate at which computers improve). At the same time, DNA sequencing has evolved from a specialized tool for determining the sequences of genomes into a versatile technology for determining gene expression levels, discovering new species, tracking cell fates, and understanding cancer growth, among many other applications (Reuter et al., 2015). With advances of technology and novel assays, DNA sequencing has revolutionized disparate areas of biology. By harnessing sequencing for neuroanatomical tracing, MAPseq may accelerate our understanding of neural circuits.

High-Throughput Sequencing and Neuroanatomy

In the present work, we applied the simplest form of MAPseq to LC. We expressed the virus ubiquitously, without any specific targeting to a cell type using, e.g., Cre recombinase, and used gross dissection to probe projections at a spatial resolution of 300 μm coronal sections. Because the LC projects throughout the entire cortex, this relatively low spatial resolution was well suited to address previously unresolved questions about LC anatomy. Gross dissection can achieve somewhat higher spatial resolution (~1–2 mm × 300 μm × cortical thickness; Figures S7A and S7B), sufficient to distinguish among cortical areas in the mouse.

The application of MAPseq to the LC highlights both the advantages and the tradeoffs of MAPseq compared with traditional methods. The main advantages of MAPseq are higher throughput and that it is less labor intensive. The total amount of time required to obtain the projection patterns from all labeled LC neurons in a single mouse was under 1 week (dissection 2 days post-injection, followed by 1–2 days of tissue processing to generate sequencing libraries, followed by 1–2 days of sequencing), of which only a relatively small fraction is spent on "hands-on" labor (as opposed to waiting). By comparison, tracing a single neuron's projections using traditional methods requires a week or more of hands-on labor and even with automation requires days. Processing time with MAPseq does not depend on the number of source (barcoded) neurons—as many as 500,000 neurons can be labeled in a single subject (unpublished data)—and MAPseq can be multiplexed to several injection sites (Figure 7), so the throughput of MAPseq could be five or more orders of magnitude higher than conventional single-neuron tracing.

Therefore, MAPseq is, without any further technical development, well positioned to address questions that depend on single-neuron resolution at the source but do not require very high spatial resolution at the target site. For example, MAPseq could readily be used to systematically and efficiently compare

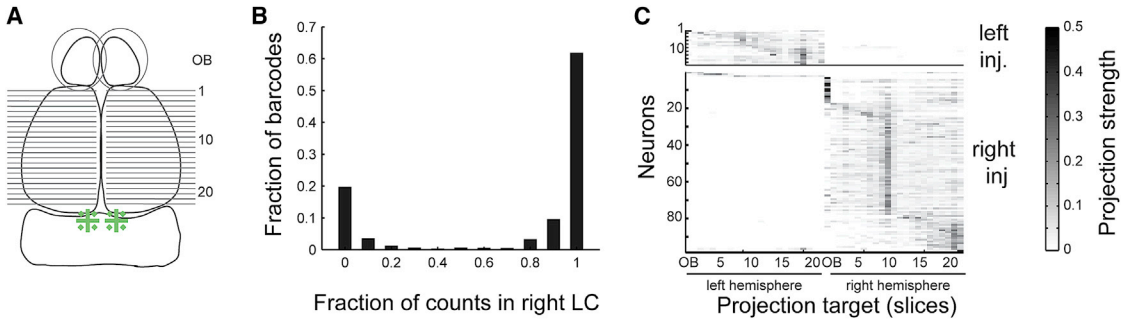

Figure 7. MAPseq Can Be Multiplexed to Several Injection Sites

(A) Following bilateral injection of barcoded Sindbis virus into LC, left and right olfactory bulb and cortex were dissected as before.

(B) Histogram of the fraction of barcode counts in the right versus left injection site across barcodes. Barcodes show strong abundance differences in the left and right injection sites, allowing them to be assigned to one of the two sites.

(C) Bilateral injections produce the projection pattern expected from unilateral injections. Differences in the number of neurons traced from the left and right LC arise from injection variability.

the projection patterns of other neuromodulatory systems, such as cholinergic or serotonergic, to the LC noradrenergic system considered here. MAPseq could also be used to test whether individual neurons in the primary auditory (or somatosensory) cortex project to specific subsets of higher auditory (or somatosensory) areas, analogous to the "ventral" and "dorsal" processing streams in the visual system. Furthermore, the Sindbis virus used to deliver barcodes has a very wide host range (Xiong et al., 1989), from arthropods to mammals (including primates). A particularly appealing application of MAPseq is thus single-neuron tracing in nonhuman primates and in less common model systems, where the cost per specimen renders traditional single-neuron tracing approaches prohibitive, and which benefit from the massive multiplexing achievable in MAPseq. It is easy to envision myriad other applications of MAPseq in its current form.

The main tradeoff of MAPseq compared with conventional methods is spatial resolution. There are several potential paths to achieving higher spatial resolution. One simple extension of the protocol presented here is to use laser-capture microdissection, which can readily resolve MAPseq target areas to several hundred microns (unpublished data), and which in principle can achieve single-neuron resolution (Espina et al., 2006). Using fluorescent in situ sequencing (FISSEQ) (Lee et al., 2014), even higher resolution can be achieved (unpublished data); barcodes can be localized with subcellular resolution, allowing for a kind of "infinite color Brainbow" (Livet et al., 2007). At present, however, FISSEQ is not fully automated and has considerably lower throughput than conventional sequencing.

What are the ultimate limits of MAPseq spatial resolution? The factors limiting MAPseq can be understood by analogy with those limiting optical microscopy. Just as spatial resolution—effective pixel size—in microscopy is limited by the optics and detectors, so is the spatial resolution of MAPseq limited by the spatial precision with which brain regions are sampled—potentially subcellular with FISSEQ. The effective spatial resolution of MAPseq may also be limited by the amount of barcode mRNA in each spatially defined region, analogous to the low-light photon-counting limit in microscopy. The precise spatial scale at which the barcode mRNA "shot noise" limit is reached in MAPseq is determined by the interplay of several factors, including the diameter of the axonal projections, expression level of the barcode, and the efficiency with which barcodes can be recovered and amplified. Because mRNA shot noise did not appear limiting in the present experiments, we did not invest significant effort in optimizing these parameters. However, the fact that barcodes can be detected even in fine LC axons suggests that MAPseq may achieve relatively high spatial resolution.

MAPseq Extensions

Information about single-neuron projection patterns obtained by MAPseq may be combined with information about other dimensions of neuronal function. For example, single-neuron projection patterns obtained by MAPseq could be associated with information about gene expression. One approach exploits transgenic mice that mark defined neuron classes with Cre recombinase. Although expression of barcodes delivered using the RNA virus Sindbis cannot readily be controlled with Cre, barcodes could be delivered using a DNA virus like AAV or a retrovirus like lentivirus. More general approaches, such as single-cell isolation (Figure S6), might associate several genes or even a whole transcriptome with projection patterns. MAPseq data could also be combined with recordings of neural activity obtained by calcium imaging. Taken together, the combination of connectivity, gene expression, and activity could provide a richer picture of neuronal function than any of these alone (Marblestone et al., 2014).

In its current form, MAPseq provides information about neuronal projections, but not about synaptic connections. MAPseq can tell us, for example, that a given neuron in the thalamus projects to both the amygdala and the cortex, but it cannot resolve whether the cortical target is an excitatory or an inhibitory neuron, or both. In this respect, it is similar to most standard tracing methods, such as GFP, which do not have synaptic specificity. However, the MAPP-nλ protein we engineered includes trafficking signals from the presynaptic protein *NRXN1β*, and was originally designed to allow us to join pre- and postsynaptic barcodes to detect synaptic connections. The joined barcode

pairs can be sequenced using high-throughput methods and provide an efficient and scalable method for determining the full connectome of a neural circuit (Zador et al., 2012).

Uniformity versus Specificity of LC Projections

The LC sends projections to most ipsilateral brain areas, with the notable exception of the striatum. However, how broadly individual neurons innervate those target areas is subject to debate. Classical retrograde tracing studies suggested a topographic organization of neocortical (Waterhouse et al., 1983) and brain-wide (Loughlin et al., 1986) projection neurons in the LC. Consistent with this, double retrograde labeling studies reported that the LC projections to frontal and motor cortices (Chandler et al., 2014; Chandler and Waterhouse, 2012) overlap minimally. In contrast, other double retrograde studies found overlap between neurons projecting to separate structures along the same processing stream (Simpson et al., 1997), or structures as different as forebrain and cerebellum (Steindler, 1981). More recent work using retrograde viral labeling combined with anterograde tracing concluded that LC neurons project largely indiscriminately throughout both cortex and the rest of the brain (Schwarz et al., 2015).

Our single-cell resolution data reconcile these conflicting datasets. We find that individual LC neurons have very specific projection targets in cortex and olfactory bulb, but are not limited to a single target. We further observe that many LC neurons that project to cortex innervate a large fraction of the cortex at least weakly, in addition to having preferred projection targets. To the extent that retrograde viral tracers may not distinguish between strong and weak projections, it may be this feature of single-neuron projections, in combination with bulk tracing, that leads Schwarz and colleagues (Schwarz et al., 2015) to conclude that LC neurons largely indiscriminately project throughout cortex and the bulb, and indeed we can replicate their results by simulating retrograde tracing on our single-cell dataset.

The LC is the sole source of noradrenaline to the cortex. Noradrenaline exerts a powerful influence on an animal's behavioral state. Noradrenaline levels control the overall level of vigilance; they are lowest during sleep and are increased in response to stimuli such as pain. Noradrenaline gates attention, enhances formation of long-term memory, and is thought to regulate the exploration-exploitation balance (Aston-Jones and Cohen, 2005; Sara, 2009). Traditionally, it has been assumed that the levels of neuromodulators such as noradrenaline represent a global signal, broadcast indiscriminately throughout the cortex. However, the specificity of the single-neuron projections patterns uncovered by MAPseq suggests that different brain regions could be subject to differential control. Whether this potential for differential control is actually realized, and what functional role it plays, remain to be determined.

Conclusion

Applying MAPseq to the LC, we discovered unexpected structure that could not have been resolved by previous methods lacking single-neuron resolution. MAPseq also lays the foundation for using sequencing to decipher local neuron-to-neuron connectivity (Zador et al., 2012). Using DNA sequencing technology, experimenters have gained unprecedented insight into the heterogeneity of cell populations at the single-cell level (Navin et al., 2011). By leveraging this sequencing technology, MAPseq empowers neuroscience researchers to efficiently do the same for populations of projection neurons examined at the single-neuron level.

EXPERIMENTAL PROCEDURES

Animal procedures were approved by the Cold Spring Harbor Laboratory Animal Care and Use Committee and carried out in accordance with NIH standards.

MAPseq

Forty-four hours after injection of MAPseq virus into LC, we flash froze the brain and cut it into 300 μm coronal sections using a cryostat. We dissected the cortical regions on dry ice and extracted total RNA from each sample. We then performed gene-specific reverse transcription for the barcode mRNA, produced double-stranded cDNA, and PCR amplified it to produce an Illumina sequencing library, which we sequenced at paired end 36 on an Illumina NextSeq sequencing machine.

Data Analysis

We processed the sequencing data to determine the exact copy number of each barcode sequence in each target area and in the injection site, and produced a barcode matrix where each row corresponds to one specific barcode sequence, each column corresponds to a target area or the injection site, and each entry is the copy number of that barcode mRNA in the respective area. All data are expressed as mean ± SD unless otherwise stated.

For full details on the experimental procedures, please refer to the Supplemental Experimental Procedures.

ACCESSION NUMBERS

All high-throughput sequencing datasets are publicly available under SRA accession codes SRA: SRS1204613 (library ZL067; virus library), SRS1204589 (libraries ZL068, ZL070, ZL071, ZL072; unilateral MAPseq datasets), SRS1204614 (libraries ZL073, ZL074; bilateral MAPseq datasets), and SRS1204626 (libraries ZL075 and ZL078; single-cell dataset).

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures seven figures, five notes, and two data files and can be found with this article online at http://dx.doi.org/10.1016/j.neuron.2016.07.036.

AUTHOR CONTRIBUTIONS

J.M.K., I.D.P., and A.M.Z. conceived the study. J.M.K. and P.G.S. performed the experiments. J.M.K. and A.P.R. performed the single-cell isolation. J.M.K. and A.M.Z. analyzed the data. J.M.K. and A.M.Z. wrote the paper. A.M.Z. and D.F.A. supervised the project.

ACKNOWLEDGMENTS

The authors would like to acknowledge Vasily Vagin, Fred Marbach, Alex Koulakov, Brittany Cazakoff, and Alexander Vaughan for useful discussions and assistance, and Barry Burbach and Diana Gizatulliana for technical support. We would further like to thank Anirban Paul for help with single-cell isolations. This work was supported by the following funding sources: NIH (5RO1NS073129 and 5RO1DA036913 to A.M.Z.); Brain Research Foundation (BRF-SIA-2014-03 to A.M.Z.); IARPA (MICrONS to A.M.Z.); Simons Foundation (382793/SIMONS to A.M.Z.); Paul Allen Distinguished Investigator Award (to A.M.Z.); PhD fellowship from the Boehringer Ingelheim Fonds to J.M.K.; PhD fellowship from the Genentech Foundation to J.M.K.; PhD fellowship from the Fundação para a Ciência e Tecnologia, Portugal, to P.G.S.; and

Pew Scholarship and CSHL startup funds to D.F.A. This work was performed with assistance from CSHL Shared Resources, which are funded, in part, by the Cancer Center Support Grant 5P30CA045508.

Received: April 29, 2016
Revised: June 18, 2016
Accepted: July 9, 2016
Published: August 18, 2016

REFERENCES

Aston-Jones, G., and Cohen, J.D. (2005). An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. Annu. Rev. Neurosci. *28*, 403–450.

Chandler, D., and Waterhouse, B.D. (2012). Evidence for broad versus segregated projections from cholinergic and noradrenergic nuclei to functionally and anatomically discrete subregions of prefrontal cortex. Front. Behav. Neurosci. *6*, 20.

Chandler, D.J., Gao, W.J., and Waterhouse, B.D. (2014). Heterogeneous organization of the locus coeruleus projections to prefrontal and motor cortices. Proc. Natl. Acad. Sci. USA *111*, 6816–6821.

Daigle, N., and Ellenberg, J. (2007). LambdaN-GFP: an RNA reporter system for live-cell imaging. Nat. Methods *4*, 633–636.

DeFalco, J., Tomishima, M., Liu, H., Zhao, C., Cai, X., Marth, J.D., Enquist, L., and Friedman, J.M. (2001). Virus-assisted mapping of neural inputs to a feeding center in the hypothalamus. Science *291*, 2608–2613.

Economo, M.N., Clack, N.G., Lavis, L.D., Gerfen, C.R., Svoboda, K., Myers, E.W., and Chandrashekar, J. (2016). A platform for brain-wide imaging and reconstruction of individual neurons. eLife *5*, e10566.

Ehrengruber, M.U. (2002). Alphaviral gene transfer in neurobiology. Brain Res. Bull. *59*, 13–22.

Espina, V., Wulfkuhle, J.D., Calvert, V.S., VanMeter, A., Zhou, W., Coukos, G., Geho, D.H., Petricoin, E.F., 3rd, and Liotta, L.A. (2006). Laser-capture microdissection. Nat. Protoc. *1*, 586–603.

Foote, S.L., and Morrison, J.H. (1987). Extrathalamic modulation of cortical function. Annu. Rev. Neurosci. *10*, 67–95.

Foote, S.L., Bloom, F.E., and Aston-Jones, G. (1983). Nucleus locus ceruleus: new evidence of anatomical and physiological specificity. Physiol. Rev. *63*, 844–914.

Ghosh, S., Larson, S.D., Hefzi, H., Marnoy, Z., Cutforth, T., Dokka, K., and Baldwin, K.K. (2011). Sensory maps in the olfactory cortex defined by long-range viral tracing of single neurons. Nature *472*, 217–220.

Glickfeld, L.L., Andermann, M.L., Bonin, V., and Reid, R.C. (2013). Cortico-cortical projections in mouse visual cortex are functionally target specific. Nat. Neurosci. *16*, 219–226.

Gong, S., Doughty, M., Harbaugh, C.R., Cummins, A., Hatten, M.E., Heintz, N., and Gerfen, C.R. (2007). Targeting Cre recombinase to specific neuron populations with bacterial artificial chromosome constructs. J. Neurosci. *27*, 9817–9823.

Harris, J.A., Hirokawa, K.E., Sorensen, S.A., Gu, H., Mills, M., Ng, L.L., Bohn, P., Mortrud, M., Ouellette, B., Kidney, J., et al. (2014). Anatomical characterization of Cre driver mice for neural circuit mapping and manipulation. Front. Neural Circuits *8*, 76.

Hayden, E.C. (2014). Technology: The $1,000 genome. Nature *507*, 294–295.

Herculano-Houzel, S., Mota, B., and Lent, R. (2006). Cellular scaling rules for rodent brains. Proc. Natl. Acad. Sci. USA *103*, 12138–12143.

Huang, Z.J. (2014). Toward a genetic dissection of cortical circuits in the mouse. Neuron *83*, 1284–1302.

Kebschull, J.M., and Zador, A.M. (2015). Sources of PCR-induced distortions in high-throughput sequencing data sets. Nucleic Acids Res. *43*, e143.

Kebschull, J.M., Garcia da Silva, P., and Zador, A.M. (2016). A new defective helper RNA to produce Sindbis virus that infects neurons but does not propagate. Front. Neuroanat. *10*, 6439.

Kennedy, A., Asahina, K., Hoopfer, E., Inagaki, H., Jung, Y., Lee, H., Remedios, R., and Anderson, D.J. (2014). Internal states and behavioral decision-making: toward an integration of emotion and cognition. Cold Spring Harb. Symp. Quant. Biol. *79*, 199–210.

Kim, J., Zhao, T., Petralia, R.S., Yu, Y., Peng, H., Myers, E., and Magee, J.C. (2011). mGRASP enables mapping mammalian synaptic connectivity with light microscopy. Nat. Methods *9*, 96–102.

Lee, J.H., Daugharthy, E.R., Scheiman, J., Kalhor, R., Yang, J.L., Ferrante, T.C., Terry, R., Jeanty, S.S.F., Li, C., Amamoto, R., et al. (2014). Highly multiplexed subcellular RNA sequencing in situ. Science *343*, 1360–1363.

Lima, S.Q., Hromádka, T., Znamenskiy, P., and Zador, A.M. (2009). PINP: a new method of tagging neuronal populations for identification during in vivo electrophysiological recording. PLoS ONE *4*, e6099.

Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Δ Δ C(T)) Method. Methods *25*, 402–408.

Livet, J., Weissman, T.A., Kang, H., Draft, R.W., Lu, J., Bennis, R.A., Sanes, J.R., and Lichtman, J.W. (2007). Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. Nature *450*, 56–62.

Loughlin, S.E., Foote, S.L., and Fallon, J.H. (1982). Locus coeruleus projections to cortex: topography, morphology and collateralization. Brain Res. Bull. *9*, 287–294.

Loughlin, S.E., Foote, S.L., and Bloom, F.E. (1986). Efferent projections of nucleus locus coeruleus: topographic organization of cells of origin demonstrated by three-dimensional reconstruction. Neuroscience *18*, 291–306.

Marblestone, A.H., Daugharthy, E.R., Kalhor, R., Peikon, I.D., Kebschull, J.M., Shipman, S.L., Mishchenko, Y., Lee, J.H., Kording, K.P., Boyden, E.S., et al. (2014). Rosetta brains: a strategy for molecularly-annotated connectomics. arXiv *q-bio*.NC. arXiv:1404.5103.

Mayer, C., Jaglin, X.H., Cobbs, L.V., Bandler, R.C., Streicher, C., Cepko, C.L., Hippenmeyer, S., and Fishell, G. (2015). Clonally related forebrain interneurons disperse broadly across both functional areas and structural boundaries. Neuron *87*, 989–998.

Movshon, J.A., and Newsome, W.T. (1996). Visual response properties of striate cortical neurons projecting to area MT in macaque monkeys. J. Neurosci. *16*, 7733–7741.

Munford, A.G. (1977). A note on the uniformity assumption in the birthday problem. Am. Stat. *31*, 119.

Navin, N., Kendall, J., Troge, J., Andrews, P., Rodgers, L., McIndoo, J., Cook, K., Stepansky, A., Levy, D., Esposito, D., et al. (2011). Tumour evolution inferred by single-cell sequencing. Nature *472*, 90–94.

Oh, S.W., Harris, J.A., Ng, L., Winslow, B., Cain, N., Mihalas, S., Wang, Q., Lau, C., Kuan, L., Henry, A.M., et al. (2014). A mesoscale connectome of the mouse brain. Nature *508*, 207–214.

Oyibo, H.K., Znamenskiy, P., Oviedo, H.V., Enquist, L.W., and Zador, A.M. (2014). Long-term Cre-mediated retrograde tagging of neurons using a novel recombinant pseudorabies virus. Front. Neuroanat. *8*, 86.

Reuter, J.A., Spacek, D.V., and Snyder, M.P. (2015). High-throughput sequencing technologies. Mol. Cell *58*, 586–597.

Sara, S.J. (2009). The locus coeruleus and noradrenergic modulation of cognition. Nat. Rev. Neurosci. *10*, 211–223.

Schüz, A., and Palm, G. (1989). Density of neurons and synapses in the cerebral cortex of the mouse. J. Comp. Neurol. *286*, 442–455.

Schwarz, L.A., Miyamichi, K., Gao, X.J., Beier, K.T., Weissbourd, B., DeLoach, K.E., Ren, J., Ibanes, S., Malenka, R.C., Kremer, E.J., and Luo, L. (2015). Viral-genetic tracing of the input-output organization of a central noradrenaline circuit. Nature *524*, 88–92.

Sheridan, C. (2014). Illumina claims $1,000 genome win. Nat. Biotechnol. *32*, 115.

Shipley, M.T., Halloran, F.J., and de la Torre, J. (1985). Surprisingly rich projection from locus coeruleus to the olfactory bulb in the rat. Brain Res. *329*, 294–299.

Simpson, K.L., Altman, D.W., Wang, L., Kirifides, M.L., Lin, R.C., and Waterhouse, B.D. (1997). Lateralization and functional organization of the locus coeruleus projection to the trigeminal somatosensory pathway in rat. J. Comp. Neurol. *385*, 135–147.

Steindler, D.A. (1981). Locus coeruleus neurons have axons that branch to the forebrain and cerebellum. Brain Res. *223*, 367–373.

Sternson, S.M. (2013). Hypothalamic survival circuits: blueprints for purposive behaviors. Neuron *77*, 810–824.

Van der Maaten, L., and Hinton, G. (2008). Visualizing data using t-SNE. J. Mach. Learn. Res. *9*, 2579–2605.

Walsh, C., and Cepko, C.L. (1992). Widespread dispersion of neuronal clones across functional regions of the cerebral cortex. Science *255*, 434–440.

Waterhouse, B.D., Lin, C.S., Burne, R.A., and Woodward, D.J. (1983). The distribution of neocortical projection neurons in the locus coeruleus. J. Comp. Neurol. *217*, 418–431.

Wickersham, I.R., Finke, S., Conzelmann, K.-K., and Callaway, E.M. (2007a). Retrograde neuronal tracing with a deletion-mutant rabies virus. Nat. Methods *4*, 47–49.

Wickersham, I.R., Lyon, D.C., Barnard, R.J.O., Mori, T., Finke, S., Conzelmann, K.-K., Young, J.A.T., and Callaway, E.M. (2007b). Monosynaptic restriction of transsynaptic tracing from single, genetically targeted neurons. Neuron *53*, 639–647.

Xiong, C., Levis, R., Shen, P., Schlesinger, S., Rice, C.M., and Huang, H.V. (1989). Sindbis virus: an efficient, broad host range vector for gene expression in animal cells. Science *243*, 1188–1191.

Zador, A.M., Dubnau, J., Oyibo, H.K., Zhan, H., Cao, G., and Peikon, I.D. (2012). Sequencing the connectome. PLoS Biol. *10*, e1001411.

Zingg, B., Hintiryan, H., Gou, L., Song, M.Y., Bay, M., Bienkowski, M.S., Foster, N.N., Yamashita, S., Bowman, I., Toga, A.W., and Dong, H.W. (2014). Neural networks of the mouse neocortex. Cell *156*, 1096–1111.