Journal of Chemical Neuroanatomy 100 (2019) 101653

Contents lists available at ScienceDirect

Journal of Chemical Neuroanatomy

journal homepage: www.elsevier.com/locate/jchemneu

Review

DNA sequencing in high-throughput neuroanatomy

Justus M Kebschull

Stanford University, Stanford, CA, United States

ARTICLE INFO

Keywords:
projection mapping
connectivity
tracing
MAPseq
BARseq
DNA sequencing
mesoscale connectivity
barcodes
in situ sequencing
Rosetta brain

ABSTRACT

Mapping brain connectivity at single cell resolution is critical for understanding brain structure. For decades, such mapping has been principally approached with microscopy techniques, aiming to visualize neurons and their connections. However, these techniques are often very labor intensive and do not scale well to the complexity of mammalian brains. We recently leveraged the speed and parallelization of DNA sequencing to map the projections of thousands of single neurons in single experiments, and to map cortical mesoscale connectivity in single mice. Here, I review the state of sequencing-based neuroanatomy, and discuss future directions in synaptic connectivity mapping and comparative connectomics.

1. Introduction

Brains are complex arrangements of millions of neurons that vary in their physiological responses to stimuli, gene expression patterns, developmental origin and connectivity patterns. Averaging over neural diversity can produce misleading results and leaves us with an incomplete picture of brain function. Recently, great strides have been made in measuring neuronal activity (e.g. refs Jun et al., 2017; Stringer et al., 2018; Ahrens et al., 2013), gene expression (e.g. refs Zeisel et al., 2018; Saunders et al., 2018; Tasic et al., 2017) and even the developmental history (e.g. refs Kalhor et al., 2018; Kalhor et al., 2016; McKenna et al., 2016; Frieda et al., 2016; Raj et al., 2018; Alemany et al., 2018; Spanjaard et al., 2018) of large numbers of individual neurons. In contrast, exploring the heterogeneity of neuronal connectivity patterns remains difficult, as traditional neuroanatomical approaches face a steep trade-off between throughput and resolution. To overcome this problem, we introduced the use of DNA sequencing and cellular barcoding to develop high-throughput, high-resolution neuroanatomical tracing techniques (Zador et al., 2012; Kebschull et al., 2016a).

In this review, I will elucidate the benefits of sequencing-based neuroanatomy over traditional approaches, and will detail the current state of available methods. Finally, I will discuss future applications of DNA sequencing in producing fully annotated “Rosetta brains,” that integrate connectional, functional and transcriptomic measures of the same cells (Marblestone et al., 2014), and in comparative connectomics (van den Heuvel et al., 2016).

2. Why sequencing?

Brain mapping by microscopy relies on the direct visualization of neurons and their connections. The conceptually simplest approach is the imaging of all cells in a tissue volume at nanometer resolution, followed by reconstruction of the imaged cells. This approach is being used with great success in electron microscopy (EM) based mapping of local circuits (see e.g. refs (Zheng et al., 2018; Kasthuri et al., 2015; White et al., 1986)) and is bolstered by significant advances in automatic circuit reconstruction. Expanding such a label free approach beyond local circuits however faces significant challenges in sample handling, as hundreds of thousands of EM images need to be taking in succession, with very low tolerance for error. Dedicated and highly specialized efforts are underway that address these challenges (see for example IARPA’s MICrONS project, aiming to reconstruct 1mm^3 of mouse cortex). However, EM connectomics at the scale of mammalian circuits remain out of reach for the average researcher.

As a result, long-range connectivity mapping today generally relies on the use of labels, such as a fluorescent dye or protein. These labels make specific neurons and their processes stand out from the rest of the brain (see for example refs (Kita and Kita, 2012; Ghosh et al., 2012; Economo et al., 2016; Gong et al., 2016a)) and thus allow imaging at lower resolution.

Label-based approaches, however, face a steep tradeoff between mapping resolution and throughput. Currently, only a small number of spectrally distinguishable labels are available, limiting the number of independent labels per brain to approximately four. Combinatorial use

E-mail address: justus@kebschull.me.

https://doi.org/10.1016/j.jchemneu.2019.101653
Received 26 January 2019; Received in revised form 30 March 2019; Accepted 24 May 2019
Available online 04 June 2019

of a small number of different fluorophores has been proposed as a way to greatly expand the number of available labels in *brainbow* and related approaches (Livet et al., 2007; Cai et al., 2013). Here cells are labeled with composite 'colors' by expressing different ratios of three fluorophores in a process akin to how colors are produced by a rgb computer display. In practice, however, it has proven challenging to reliably read out different color combinations in axons – where individual fluorophores are sparse – limiting the utility of this approach for brain mapping. As a result, currently only approximately ten, and often closer to three colors can be used as independent labels in the same brain. A mouse brain, in contrast, contains 100 million neurons (Herculano-Houzel et al., 2006) and even single brain areas, such as auditory cortex, still contain 300,000 neurons (Herculano-Houzel et al., 2013).

The gold standard in single-neuron resolution brain mapping is to label a single neuron per brain, either by viral infection (e.g. ref (Ghosh et al., 2012)) or single neuron electroporation (e.g. ref (Han et al., 2018)), and to image its entire axonal tree (Kita and Kita, 2012; Ghosh et al., 2012). This approach provides definitive single cell resolution and high spatial resolution. As only a single neuron is traced per animal, throughput however is low, and collecting sufficient data to comprehensively describe the diversity of projection patterns in a brain area is often prohibitively labor intensive (Han et al., 2018). Recent technological advances now allow volumetric imaging of entire mammalian brains at high resolution, and make it possible to trace dozens of individual, identically labeled single neurons per brain (Economo et al., 2016; Gong et al., 2016b; Winnubst et al., 2019). In particular, the MouseLight project at Janelia has recently released an impressive dataset of 1000 reconstructed mouse neurons scattered across the brain (Winnubst et al., 2019). Collecting large numbers of single neuron projection patterns from individual areas of interest however remains challenging, requiring many experiments, and specialized, expensive equipment.

Faced with the low throughput of single neuron tracing, the majority of brain mapping today is done at the coarser resolution of brain areas. Here, the number of available colors is a good match to the often small number of brain areas of interest. Even comprehensive area-to-area maps can be produced in concerted efforts, tracing the projections of one or a few brain areas per animal (Oh et al., 2014; Zingg et al., 2014; Harris et al., 2018). Such bulk projection tracing, however, obscures the underlying and often diverse projection patterns of single neurons. Insight into information routing and cross-area computations is limited. This problem can be partially addressed by probing the collateralization patterns of individual neurons using several retrograde tracers of different colors, or restricting bulk tracing to projection defined subpopulations of neurons (see Box 1; Table 1). Again, however, the resolution is limited by the low number of labels, as well as the efficiencies of each retrograde labeling step.

To transform single-neuron resolution brain mapping into a high-throughput method broadly accessible to the research community, we recently proposed to draw on cellular barcoding strategies (Kebschull and Zador, 2018) to multiplex single neuron tracing. Cellular barcoding uses short nucleic acid sequences, "barcodes," rather than colors to uniquely label cells (Zador et al., 2012; Kebschull et al., 2016a). The key advantage of barcode labeling is that even short sequences have extremely high diversity. A stretch of 30 random nucleotides, for example, can take any of $4^{30} \approx 10^{18}$ different sequences. In principle, a 30-nt barcode therefore provides enough labels to uniquely label every one of the 100 million neurons in the mouse brain 10 billion times over (Fig. 1). Importantly, each of these sequences can be reliably, rapidly and inexpensively read out and quantified by readily accessible high-throughput sequencing technologies (Hayden, 2014).

3. MAPseq: a sequencing-based method for high-throughput single neuron tracing

Multiplexed Analysis of Projections by sequencing (MAPseq) uses cellular barcoding and a DNA sequencing-based readout to map the area-to-area projections of thousands of individual neurons in parallel in a single experiment (Kebschull et al., 2016a). For MAPseq, we first produce a complex mixture of barcoded Sindbis viruses (Kebschull et al., 2016b), in which each virus particle encodes one of millions of different barcode sequences (Fig. 2a). We then inject this mixture into a brain area, at concentrations such that each infected neuron takes up roughly one virus particle (Fig. 2b). Infected neurons are thus labeled with the barcode of the infecting virus particle. Within the cell, barcodes are amplified through transcription. Alongside the barcode mRNA, the virus expresses an engineered presynaptic protein, MAPP-nλ that is based on pre-mGRASP (Kim et al., 2011). MAPP-nλ contains the viral RNA binding domain nλ in its intracellular tail that specifically binds to a series of four boxB hairpins included in the barcode mRNA (Daigle and Ellenberg, 2007). When MAPP-nλ is trafficked into the presynaptic terminals of the labeled neurons, it drags the barcode mRNA along and thus moves barcode mRNA from the cell body into the axonal processes of each neuron (Fig. 2c,d). To map the projections of such barcode-filled neurons, we then simply dissect out potential target areas as well as tissue at the injection site, and sequence all barcodes contained in them (Fig. 2e). Matching the barcodes found in each target area to barcodes in the injection site then constitutes projection mapping, as barcodes detected far away from the cell body must have traveled there in the axonal processes of the neuron. MAPseq thus avoids all manual tracing, and can map the projections of thousands and potentially millions of neurons in parallel at dissection-limited spatial resolution.

Consequences of viral barcode delivery

For MAPseq to report single neuron projection patterns, every neuron must be uniquely labeled by a barcode sequence. As viral delivery of barcodes to neurons is random, two potential sources of error must be considered and controlled for. On the one hand, one neuron might be infected by more than one virus particle and would therefore be labeled with multiple barcodes. On the other hand, two or more neurons might get infected by virus particles containing the same barcode sequence and would thus be labeled degenerately.

Multiple labeling, where a neuron carries two distinct barcodes, will result in this neuron effectively being traced twice. Importantly, however, the tracing information for each of the two barcodes is accurate. As viral infection, to a first approximation, is a random Poisson process, multiple labeling will cause an overestimation of the number of traced neurons in a MAPseq experiment, but will leave the relative abundances of different projection patterns unchanged. Multiple labeling therefore slightly alters the interpretation of each recovered barcode, but is largely inconsequential for MAPseq tracing. Nevertheless, the frequency of multiple labeling can be controlled by adjusting viral titer and injection volume, and can be measured by single cell barcode sequencing (Kebschull et al., 2016a).

Degenerate labeling, where several neurons share the same barcode, results in loss of single cell resolution. MAPseq cannot distinguish identically labeled cells, and the aggregate (bulk) projection pattern of these cells is reported. This is a severe error and must be minimized for a successful experiment. The rate of degenerate labeling is completely determined by the number of infections relative to the number of available barcodes. In the extreme, if more cells are labeled than there are different barcode sequences in the virus library, degenerate labeling is certain. As a rule of thumb, labeling a given number of cells with a virus library 100x larger than this number will result in a degenerate labeling rate of 1% (for a more quantitative discussion of this issue see ref (Kebschull and Zador, 2018)). Acceptable error rates vary by the particular experiment, and can be controlled for by adjusting the size of the virus library and the number of infected cells per mouse.

Box 1
Approaches to mapping axonal collateralization.

Individual neurons often project to more than one target area in the brain. They thus simultaneously distribute identical sets of information to different postsynaptic targets, suggesting different underlying computations than if each neuron projected to only one target area. A variety of methods exist to map such collateralizations (Table 1).

The most common method for mapping the collateralization patterns of neurons in a brain area of interest is multi-color retrograde tracing (Kuypers et al., 1980; Cavada et al., 1984). Each target area is injected with a retrograde tracer (e.g. CTb (Trojanowski et al., 1982; Wan et al., 1982), retrobeads (Katz et al., 1984), CAV (Soudais et al., 2001), HSV (Ugolini et al., 1987), AAVretro (Tervo et al., 2016)) that is labeled with a different color. The area of interest is then imaged, and the number of single, double, etc. labeled neurons quantified. While this approach is simple and provides cellular spatial resolution in the area of interest and high, injection-limited resolution at the target sites, retrograde tracing scales poorly with the number of target sites tested. Each retrograde tracer has a limited probability of infecting any given neuron that projects to the injection site. As individual tracer injections are independent of each other, the probability of observing a neuron labeled by more than one tracer decreases exponentially, even if all neurons project to all target areas. Multicolor retrograde tracing is therefore usually limited to testing two, or three target sites and requires careful controls to measure the efficiencies of retrograde labeling. Moreover, no information about projection strength can be obtained, as retrograde labeling is largely binary.

Collateralization maps in part overcome the limited throughput and lack of projection strength quantification of multi-color retrograde tracing by pairing retrograde labeling with anterograde tracing (Schwarz et al., 2015; Beier et al., 2015). In this method, a retrograde virus expressing a recombinase (e.g. CAV-cre; AAVretro-cre) is injected into a single target area. At the same time, a recombinase reporter virus (e.g. AAV expressing cre-dependent GFP) is injected into the area of interest. All reporter labeled neurons in the brain are therefore known to project to the injection site of the first virus, and to have their cell bodies in the injection site of the second. Imaging the axonal arbors of the labeled neurons then allows to efficiently determine the projection targets of neurons that also project to the injection site of the retrograde virus. Moreover, imaging of axonal arbors allows the discovery of unexpected projection patterns within brain areas. However, like multicolor retrograde tracing, collateralization mapping is a binary readout of projection to the area injected with retrograde virus, and is limited to testing pairwise projection pairings.

Single-neuron anterograde tracing avoids the drawbacks of retrograde tracing altogether by labeling only a single neuron per brain with a cell-filling dye or protein (delivered by electroporation, or viral or dye injection), such that all processes of that neuron can be traced across the brain, without confusing them for another's (Kita and Kita, 2012; Ghosh et al., 2012; Economo et al., 2016; Han et al., 2018). The number of potential target areas does not influence the efficiency of cell filling, and imagining individual axonal arbors provides unprecedented structural information, allowing the quantification of branch points, angles and geometries. Single-cell tracing, however, is extremely laborious, often making it difficult to obtain a sufficient number of reconstructed neurons that would reflect the diversity of projection patterns in an area. Recent technical improvements that allow imaging entire brain volumes at very high resolution now allow labeling dozens of cells in the same brain, and tracing their processes without confusing one for another (Economo et al., 2016; Gong et al., 2016a; Winnubst et al., 2019). However, throughput is still limited to low tens of neurons per brain area, and relies on specialized equipment.

MAPseq overcomes the limited throughput of fluorophore-based single neuron tracing, mapping the projections of thousands of neurons per area (Kebschull et al., 2016a; Han et al., 2018). The increased throughput, however, comes at the cost of lower spatial resolution that is completely determined by dissection, and cannot easily approach that of imaging. Combining MAPseq with *in situ* barcode readout in BARseq addresses the reduced resolution at the injection site, and potentially in target sites.

Each approach comes with certain advantages and disadvantages. When choosing the appropriate tool to map axonal collateralization patterns, one therefore needs to tradeoff the scale of the mapping (a few select target sites or brain-wide?), the resolution required at each of the target sites (imaging resolution or injection/dissection?), and the number of cells that need to be examined, based on the biological question that is to be answered.

4. Demonstrations of MAPseq to delineate projection architecture

We first applied MAPseq to tracing the projections of locus coeruleus (LC) (Kebschull et al., 2016a), a major noradrenergic nucleus controlling arousal. In contrast the homogeneous LC innervation of cortex shown by bulk tracing (Schwarz et al., 2015; Waterhouse et al., 1983), MAPseq revealed idiosyncratic projection patterns of individual LC neurons, with preferred projection patterns to specific positions along the anterior-posterior axis of cortex. As a population, LC neurons tiled cortex with these preferred projection targets. The specificity of LC projections at single-cell resolution opens the possibility that arousal signals carried by different LC neurons could differentially act on individual brain areas (see also refs (Chandler and Waterhouse, 2012; Chandler et al., 2014)).

Though elucidating the LC output architecture served as a perfect test case, MAPseq has the potential to provide clarity to even more complicated questions. We thus turned our attention to the rules underlying cortical wiring. Structurally, cortex is formed from repeated modules, all tightly interconnected, and each specialized to process specific information. Still we know very little about how these areas communicate. According to a popular "one-neuron—one-target" model, individual neurons project to only a single target area and provide that area with perfectly tailored information (Nakamura et al., 1993; Segraves and Innocenti, 1985; Rockland, 2013; Sincich and Horton, 2003; Yamashita et al., 2013). This model suggests that an area like primary visual cortex (V1) separates the information it receives from the eyes into independent streams, each processing specific visual features (Glickfeld et al., 2013).

Limited by the throughput of classical single neuron tracing, however, proof for the "one-neuron—one-target" model was lacking. We used MAPseq to map the projections of hundreds of individual V1 neurons (Han et al., 2018). Contrary to expectation, most neurons projected to more than one area and broadcasted information to multiple areas simultaneously. We then investigated the higher-order structure of the broadcasting neuron population – an analysis possible only due to the large number of neurons traced. We found that the set of projection targets of individual neurons was non-random, with some combinations occurring more often and others less often than expected by chance. Our results demonstrated the existence of specialized classes of broadcasting neurons in V1. This outcome challenges the model of one-to-one lines of information flow in cortex and requires the development of more complex functional models of intra-cortical communication.

5. Scaling up: mapping connectivity networks using sequencing

In MAPseq and traditional bulk tracing, brain mapping is limited to a single source area, in which the cell bodies of all traced neurons reside. Obtaining a global understanding of brain connectivity therefore requires data collation from many individual experiments, each

Table 1
Characteristics of various methods to map the collateralization patterns of individual neurons.

Method	# of target sites	Single neuron resolution	Projection strength readout	Number of cells traced per brain	Resolution at target	Resolution at source
Retrograde tracing	≤ 4	Yes	No	1000s	Injection limited	Cellular
Collateralization maps	1 retrograde site + brain-wide collaterals	No	No/yes	n/a	Injection limited/imaging resolution	Cellular
Fluorophore-based single cell tracing	Brain-wide	Yes	Yes	1 to tens	Imaging resolution	Cellular
MAPseq	Brain-wide	Yes	Yes	1000s	Dissection limited	Dissection limited
BARseq	Brain-wide	Yes	Yes	1000s	Dissection limited	Cellular

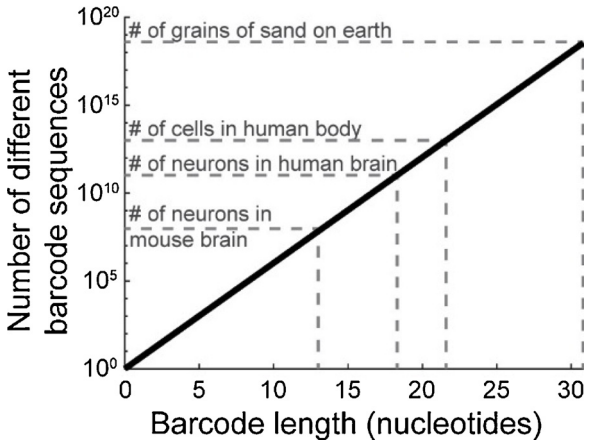

Fig. 1. The number of distinct barcode sequences grows exponentially with increasing sequence length. As a result, an 18 nucleotide barcode can form more sequences than there are neurons in the human brain, and a 22 nucleotide barcode more than there are cells in a human body.

interrogating the projections from a different source area (Oh et al., 2014; Zingg et al., 2014; Harris et al., 2018). Such an approach, however, is labor intensive, costly, and suffers from complications in cross-brain integration and animal-to-animal variability. As a result, global connectivity atlases exist only for a small set of model species, and even there are generally restricted to a certain age and sex (Oh et al., 2014; Zingg et al., 2014; Harris et al., 2018,Markov et al., 2014; Felleman and Van Essen, 1991). A systematic understanding of the effects of age, sex, strain or disease mutations on global connectivity therefore lacking.

To overcome these limitations we developed muMAPseq (multi-source MAPseq), a method that allows brain-wide projection tracing from hundreds of source areas in the same animal and experiment (Huang et al., 2018). MuMAPseq draws on the large number of labels provided by barcodes to uniquely label tens of thousands of cells across all source areas. This is achieved by injecting the same large pool of barcoded Sindbis virus into multiple source areas, without regard for which barcodes are used in which area. After viral expression, all potential target areas and all source areas (note that source areas can be target areas) are dissected and sequenced. Every barcode is then assigned a source area *post hoc* by abundance, exploiting the observation that barcodes are much more abundant in cell bodies than in axons.

We applied muMAPseq to mapping the cortico-cortical mesoscale connectome of two adult male C57BL/6 J animals, recovering approximately 70,000 uniquely labeled projection neurons across ~120 source areas and a total of ~280 target areas in each experiment. Constrained by low sequencing depth and other technical limitations, we analyzed area-to-area bulk connectivity data, collapsing projection patterns across neurons in individually dissected samples. We found muMAPseq derived connectivity data to be highly reproducible and to closely match the Allen Connectivity Atlas reference dataset (Oh et al., 2014). We then proceeded to compare the C57BL/6 J datasets to those obtained from mapping two male BTBR mice, an inbred mouse strain notable for the absence of a corpus callosum (McFarlane et al., 2008; Wahlsten et al., 2003). As expected muMAPseq revealed a complete absence of contralateral connections in BTBR mice. Despite this large scale disruption of brain-wide connectivity, however, we found that the ipsilateral cortical connectome structure was largely maintained of BTBR mice. Our study for the first systematically mapped the mesoscale connectomes of different mouse strains, allowing network-level comparison. MuMAPseq allows the production of tailor made connectivity atlases for non-standard model systems in any lab, for a few thousands of dollars, again at dissection-limited spatial resolution.

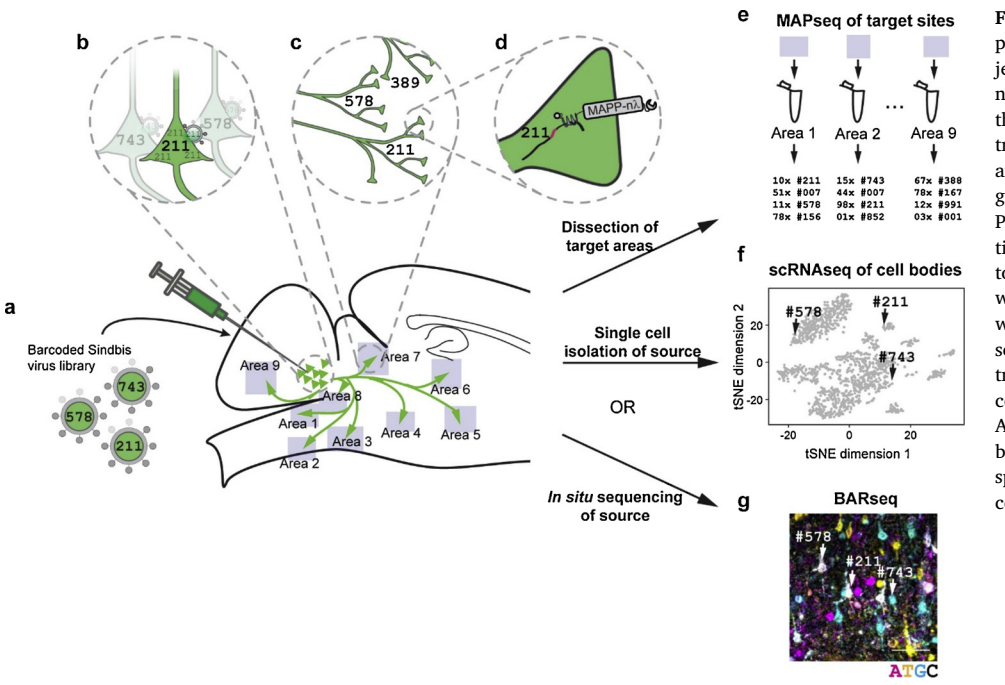

Fig. 2. MAPseq and its extensions. (a) A large pool of barcoded Sindbis virus particles is injected into a source area. (b) Every infected neuron is labeled with the barcode carried by the infecting virus particle. (c) The barcode is transcribed into mRNA, which is trafficked into axonal processes by (d) the co-expressed engineered presynpatic protein MAPP-nλ. (e) Projection mapping then requires the dissection and sequencing of potential target areas, to obtain a list of which barcode is found where, and thus of which neuron projects where. (f) MAPseq can be combined with scRNAseq of the barcoded cell bodies to obtain transcriptional information in addition to the connectivity info for every cell. (g) Alternatively, barcodes at the injection site can be sequenced *in situ* by BARseq to provide spatial information to complement MAPseq connectivity information.

6. Integration of connectivity data with other modalities

Annotation of connectivity data with each cells' morphology, transcriptome and activity patterns, i.e. the production of a 'Rosetta brain' (Marblestone et al., 2014), is a long standing goal in neuroscience. Sequencing-based neuroanatomical methods in particular lend themselves to such integration, as the barcodes used for connectivity mapping are mRNAs, and are thus already in the same modality as the cellular transcriptome. As a first step in this direction, we recently combined MAPseq tracing at target sites with single cell RNA sequencing of barcoded cell bodies (Fig. 2f). This study simultaneously revealed the projection patterns and transcriptional state of cells in the developing somatosensory cortex and thus allowed the identification of transcriptional modules that correlate with different projections (Klingler et al., 2018).

While powerful in its own right, it is challenging to expand the single cell RNA sequencing approach to the Rosetta brain beyond the integration of transcriptomic and projectional data: Using single cell RNA sequencing to read out a cell's transcriptome requires the dissociation of tissue into a suspension of single neurons. All spatial information is lost, precluding the analysis of spatial domains or gradients within the sequenced area, as well as investigation of neuronal morphology.

Recent breakthroughs in single molecule *in situ* RNA detection methods now provide an attractive alternative pathway towards a 'Rosetta brain' (Wang et al., 2018; Lee et al., 2015; Shah et al., 2018). In these methods, mRNAs are detected *in situ* in tissue slices, combining the benefits of sequencing with the utility of spatial information. To allow the integration of barcode based connectivity information with other modalities, we recently developed one such method, BARseq, to determine specifically the sequence and position of barcode mRNAs in tissue. This achieved by amplifying each barcode *in situ* and then reading its sequence out using Illumina sequencing-by-synthesis chemistry in the tissue slice. In a proof-of-principle experiment, we combined BARseq of barcoded cell bodies, with MAPseq projection mapping of the same cells in the primary auditory cortex (Chen et al., 2018) (Fig. 2g). We were thus able to precisely determine the cell body location of each barcoded neuron within the auditory cortex, as well as each neuron's long-range projection targets. The MAPseq data revealed a large diversity of projection neuron classes. Intriguingly, when we related single cell projection patterns of neurons to the cortical layer they originate from, we found that that only the broad divisions of projection patterns, but not more subtle differences, were reflected in clear differences in layer distributions.

BARseq is compatible with other tools for the highly multiplexed detection of gene expression *in situ* (Wang et al., 2018; Lee et al., 2015; Shah et al., 2018) and thus lays the foundation for a Rosetta brain, where barcode based connectomics readouts are combined with morphological and spatial information, as well as information about transcriptional cell types. Moreover, we envision that in the future it should be possible to register BARseq'd cells to 2-photon functional imaging data of the same cells by spatial alignment (compare refs (Ko et al., 2011; Ko et al., 2013)), adding functional information to connectional and transcriptomic data. Finally, this approach could be combined with CRISPR Cas9 generated evolving barcodes (Kalhor et al., 2018; Kalhor et al., 2016; McKenna et al., 2016; Frieda et al., 2016; Raj et al., 2018; Alemany et al., 2018; Spanjaard et al., 2018), that record lineage information, to combine connectional, transcriptomic and functional data, with developmental information for an integrated view of brain circuits.

7. Future directions

Barcode based connectomics is a very young field with many future applications and directions, now that the feasibility of brain mapping by sequencing is established. Here I will highlight two of my personal favorites, namely large-scale synaptic connectivity mapping, and comparative connectomics (van den Heuvel et al., 2016) of multi-modal connectomes.

7.1. Synaptic connectivity mapping

While initially proposed as an approach for synaptic connectivity mapping (Zador et al., 2012), sequencing has so far mainly been applied to projection mapping. Nevertheless, we established the feasibility of purely sequencing based connectome mapping at the synaptic level by SYNseq (Peikon et al., 2014). In this method barcode mRNAs from pre- and postsynaptic cells are trafficked to the synapse, where they are bound into a transsynaptic protein-barcode complex. This complex is then extracted from the brain, and the pre- and postsynaptic barcode

mRNAs in each complex are joined into barcode-pairs *in vitro* by droplet overlap RT-PCR. Each pre-post pairing is then read out by Illumina sequencing of the barcode-pairs resulting in a connectivity matrix. Unfortunately, inefficient barcode joining in droplets so far prevented the application of SYNseq to any biological question. Since the development of SYNseq, technological breakthroughs in the field of single cell sequencing (10X genomics; (Zeisel et al., 2018; Saunders et al., 2018; Macosko, 2015); split-pool combinatorial barcoding (Rosenberg et al., 2019; Cao et al., 2017)) have greatly improved our ability to handle large numbers of individual synaptic complexes with or without droplets and to manipulate nucleic acids at this scale. Applying these lessons to SYNseq, or similar approaches based on barcoded synaptosome preparations, should improve the efficiency of barcode-pair generation, and allow the sequencing of a connectome.

Combination of synaptic barcode localization with *in situ* barcode sequencing (Chen et al., 2018), and expansion microscopy (Chozinski et al., 2016; Chang et al., 2017; Chen et al., 2015; Wassie et al., 2019) (ExM) provides an alternative avenue for high-throughput synaptic connectomics. Here, pre- and post-synaptically localized barcodes would be detected and read out by *in situ* sequencing in expanded tissue. The high effective resolution of ExM in combination with synaptic marker staining, or the restriction of *in situ* sequencing to synapses, would then allow to map synaptic connectivity by simply associating proximal barcodes with each other (compare also refs (Marblestone et al., 2014; Yoon et al., 2017; Underwood, 2016; Mishchenko, 2010)). Such an approach would combine the throughput of sequencing based neuroanatomy with the spatial resolution afforded by ExM. Moreover, this approach is particularly well suited to produce a synaptic resolution “Rosetta brain”. Connectivity readout can be readily combined with antibodies that distinguish various kinds of synapses. FISH or *in situ* sequencing can be added for reading out endogenous gene expression, and the entire ExM volume can be registered to previously acquired 2-photon Ca^{2+} imaging datasets, to add activity information.

7.2. Comparative connectomics

Mapping brain connectivity in individual model systems will undoubtedly provide a valuable foundation for all future studies of the mapped system. Modern brains and their circuits, however, where shaped by millions of years of evolution. Understanding not only how brains work, but also how they happened to work in a particular way, as well as what principles govern brain function, therefore require the comparison of brain circuits across species. So far such comparisons were hindered by a lack of high-throughput tools to map circuits. Sequencing based connectivity mapping now provides an opportunity to overcome this bottleneck and to provide the data needed in the emerging field of comparative connectomics (van den Heuvel et al., 2016). Using MAPseq and related technologies large scale circuits can be mapped at single cell resolution in various species. Integration of connectivity data with other modalities such as transcriptomic and functional data will then facilitate the alignment (Stuart et al., 2018; Welch et al., 2019) of different circuit elements across species, such that we can arrive at an understanding how modern circuits evolved from their evolutionary precursors. Areas of particular interest for comparative connectomics include the elaborating visual system in the mammalian lineage, the evolution of the mammalian cortex and avian pallium from a common ancestor, and – on shorter evolutionary timescales – the circuit changes that might underlie different parenting behaviors in monogamous versus polygamous mice (Bendesky et al., 2017).

8. Conclusion

By leveraging the vast combinatorial space of nucleic acid sequences and the power of modern DNA sequencing approaches, sequencing based high-throughput neuroanatomy is already providing unprecedented insights into the structure of the nervous system. A recent confluence of technological breakthroughs in *in situ* sequencing, single cell RNA sequencing and expansion microscopy now place neuroanatomical barcoding in a unique position to provide an integrated view of brain structure and function.

Competing interests

The author declares no competing interests.

Acknowledgements

I would like to thank Kevin Beier for feedback on this manuscript, and Tony Zador, Longwen Huang, Ian Peikon, Xiaoyin Chen, Vas Vagin, and Diana Ravens for many discussions over the years. This work was supported by a Coffin Childs Postdoctoral Fellowship to J.M.K.

References

Ahrens, M.B., Orger, M.B., Robson, D.N., Li, J.M., Keller, P.J., 2013. Whole-brain functional imaging at cellular resolution using light-sheet microscopy. Nat. Methods 10, 413–420.
Alemany, A., Florescu, M., Baron, C.S., Peterson-Maduro, J., Van Oudenaarden, A., 2018. Whole-organism clone tracing using single-cell sequencing. Nature 556, 108–112.
Beier, K.T., et al., 2015. Circuit architecture of VTA dopamine neurons revealed by systematic input-output mapping. Cell 162, 622–634.
Bendesky, A., et al., 2017. The genetic basis of parental care evolution in monogamous mice. Nature 544, 434–439.
Cai, D., Cohen, K.B., Luo, T., Lichtman, J.W., Sanes, J.R., 2013. Improved tools for the Brainbow toolbox. Nat. Publ. Gr. 1–10.
Cao, J., et al., 2017. Comprehensive Single Cell Transcriptional Profiling of a Multicellular Organism by Combinatorial Indexing.
Cavada, C., Huisman, A.M., Kuypers, H.G.J.M., 1984. Retrograde double labeling of neurons: the combined use of horseradish peroxidase and diamidino yellow dihydrochloride (DY·2HCl) compared with true blue and DY·2HCl in rat descending brainstem pathways. Brain Res. 308, 123–136.
Chandler, D., 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.
Chandler, D.J., Gao, W.J., Waterhouse, B.D., 2014. Heterogeneous organization of the locus coeruleus projections to prefrontal and motor cortices. Proc. Natl. Acad. Sci. 111, 6816–6821.
Chang, J.-B., et al., 2017. Iterative expansion microscopy. Nat. Methods. https://doi.org/10.1038/nmeth.4261.
Chen, X., Kebschull, J.M., Zhan, H., Sun, Y.-C., Zador, A.M., 2018. Spatial organization of projection neurons in the mouse auditory cortex identified by in situ barcode sequencing. bioRxiv 294637. https://doi.org/10.1101/294637.
Chen, F., Tillberg, P.W., Boyden, E.S., 2015. Expansion microscopy. Science (80-) 347.
Chozinski, T.J., et al., 2016. Expansion microscopy with conventional antibodies and fluorescent proteins. Nat. Methods 13, 485–488.
Daigle, N., Ellenberg, J., 2007. λN-GFP: an RNA reporter system for live-cell imaging. Nat. Methods 4, 633–636.
Economo, M.N., et al., 2016. A platform for brain-wide imaging and reconstruction of individual neurons. Elife 5, 13.
Felleman, D.J., Van Essen, D.C., 1991. Distributed hierarchical processing in the primate cerebral cortex. Cereb. Cortex 1, 1–47.
Frieda, K.L., et al., 2016. Synthetic recording and in situ readout of lineage information in single cells. Nature 541, 107–111.
Ghosh, S., et al., 2012. 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., Reid, R.C., 2013. Cortico-cortical projections in mouse visual cortex are functionally target specific. Nat. Neurosci. 16, 219–226.
Gong, H., et al., 2016a. High-throughput dual-colour precision imaging for brain-wide connectome with cytoarchitectonic landmarks at the cellular level. Nat. Commun. 7, 12142.
Gong, H., et al., 2016b. High-throughput dual-colour precision imaging for brain-wide connectome with cytoarchitectonic landmarks at the cellular level. Nat. Commun. 7, 12142.
Han, Y., et al., 2018. The logic of single-cell projections from visual cortex. Nature 556, 51–56.
Harris, J.A., et al., 2018. The organization of intracortical connections by layer and cell class in the mouse brain. bioRxiv 292961. https://doi.org/10.1101/292961.
Hayden, E.C., 2014. Technology: the $1,000 genome. Nature 507, 294–295.
Herculano-Houzel, S., Mota, B., Lent, R., 2006. Cellular scaling rules for rodent brains. Proc. Natl. Acad. Sci. 103, 12138–12143.
Herculano-Houzel, S., Watson, C., Paxinos, G., 2013. Distribution of neurons in functional areas of the mouse cerebral cortex reveals quantitatively different cortical zones. Front. Neuroanat. 7, 35.
Huang, L., et al., 2018. High-throughput mapping of mesoscale connectomes in individual

mice. bioRxiv 422477. https://doi.org/10.1101/422477.
Jun, J.J., et al., 2017. Fully integrated silicon probes for high-density recording of neural activity. Nature 551, 232–236.
Kalhor, R., et al., 2018. Developmental barcoding of whole mouse via homing CRISPR. Science 361, eaat9804.
Kalhor, R., Mali, P., Church, G.M., 2016. Rapidly evolving homing CRISPR barcodes. Nat. Methods 14, 195–200.
Kasthuri, N., et al., 2015. Saturated reconstruction of a volume of neocortex. Cell 162, 648–661.
Katz, L.C., Burkhalter, A., Dreyer, W.J., 1984. Fluorescent latex microspheres as a retrograde neuronal marker for in vivo and in vitro studies of visual cortex. Nature 310, 498–500.
Kebschull, J.M., et al., 2016a. High-throughput mapping of single-neuron projections by sequencing of barcoded RNA. Neuron 91, 975–987.
Kebschull, J.M., da Silva, P., Zador, A.M., 2016b. A new defective helper RNA to produce recombinant sindbis virus that infects neurons but does not propagate. Front. Neuroanat. 10, 6439.
Kebschull, J.M., Zador, A.M., 2018. Cellular barcoding: lineage tracing, screening and beyond. Nat. Methods 15, 871–879.
Kim, J., et al., 2011. mGRASP enables mapping mammalian synaptic connectivity with light microscopy. Nat. Methods 9, 96–102.
Kita, T., Kita, H., 2012. The subthalamic nucleus is one of multiple innervation sites for long-range corticofugal axons: a single-axon tracing study in the rat. J. Neurosci. 32, 5990–5999.
Klingler, E., et al., 2018. Single-cell molecular connectomics of intracortically-projecting neurons. bioRxiv 378760. https://doi.org/10.1101/378760.
Ko, H., et al., 2011. Functional specificity of local synaptic connections in neocortical networks. Nature 473.
Ko, H., et al., 2013. The Emergence of Functional Microcircuits in Visual Cortex. https://doi.org/10.1038/nature12015.
Kuypers, H.G., Bentivoglio, M., Catsman-Berrevoets, C.E., Bharos, A.T., 1980. Double retrograde neuronal labeling through divergent axon collaterals, using two fluorescent tracers with the same excitation wavelength which label different features of the cell. Exp. Brain Res. 40, 383–392.
Lee, J.H., et al., 2015. Fluorescent in situ sequencing (FISSEQ) of RNA for gene expression profiling in intact cells and tissues. Nat. Protoc. 10, 442–458.
Livet, J., et al., 2007. Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. Nature 450, 56–62.
Macosko, E.Z., et al., 2015. Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. Cell 161, 1202–1214.
Marblestone, A.H., et al., 2014. Rosetta brains: a strategy for molecularly-annotated connectomics. arXiv.org.
Markov, N.T., et al., 2014. A weighted and directed interareal connectivity matrix for macaque cerebral cortex. Cereb. Cortex 24, 17–36.
McFarlane, H.G., et al., 2008. Autism-like behavioral phenotypes in BTBR T+tf/J mice. Genes Brain Behav. 7, 152–163.
McKenna, A., et al., 2016. Whole-organism lineage tracing by combinatorial and cumulative genome editing. Science (80-) 353 aaf7907–aaf7907.
Mishchenko, Y., 2010. On Optical Detection of Densely Labeled Synapses in Neuropil and Mapping Connectivity with Combinatorially Multiplexed Fluorescent Synaptic Markers. PLoS One 5, e8853.
Nakamura, H., Gattass, R., Desimone, R., Ungerleider, L., 1993. The modular organization of projections from areas V1 and V2 to areas V4 and TEO in macaques. J. Neurosci. 13.
Oh, S.W., et al., 2014. A mesoscale connectome of the mouse brain. Nature 508, 207–214.
Peikon, I.D., Gizatullina, D.I., Zador, A.M., 2014. In vivo generation of DNA sequence diversity for cellular barcoding. Nucleic Acids Res. 42 e127–e127.
Raj, B., et al., 2018. Simultaneous single-cell profiling of lineages and cell types in the vertebrate brain. Nat. Biotechnol. https://doi.org/10.1038/nbt.4103.
Rockland, K.S., 2013. Collateral branching of long-distance cortical projections in monkey. J. Comp. Neurol. 521, 4112–4123.
Rosenberg, A.B. et al. Scaling single cell transcriptomics through split pool barcoding doi: https://doi.org/10.1101/105163.
Saunders, A., et al., 2018. Molecular diversity and specializations among the cells of the adult mouse brain. Cell 174, 1015–1030 e16.
Schwarz, L.A., et al., 2015. Viral-genetic tracing of the input-output organization of a central noradrenaline circuit. Nature 524, 88–92.
Segraves, M., Innocenti, G., 1985. Comparison of the distributions of ipsilaterally and contralaterally projecting corticocortical neurons in cat visual cortex using two fluorescent tracers. J. Neurosci. 5.
Shah, S., et al., 2018. Dynamics and Spatial Genomics of the Nascent Transcriptome by Intron seqFISH. Cell 174, 363–376 e16.
Sincich, L.C., Horton, J.C., 2003. Independent projection streams from macaque striate cortex to the second visual area and middle temporal area. J. Neurosci. 23, 5684–5692.
Soudais, C., Laplace-Builhe, C., Kissa, K., Kremer, E.J., 2001. Preferential transduction of neurons by canine adenovirus vectors and their efficient retrograde transport in vivo. FASEB J. 15, 2283–2285.
Spanjaard, B., et al., 2018. Simultaneous lineage tracing and cell-type identification using CRISPR–Cas9-induced genetic scars. Nat. Biotechnol. https://doi.org/10.1038/nbt.4124.
Stringer, C., et al., 2018. Spontaneous behaviors drive multidimensional, brain-wide population activity. bioRxiv 306019. https://doi.org/10.1101/306019.
Stuart, T., et al., 2018. Comprehensive integration of single cell data. bioRxiv 460147. https://doi.org/10.1101/460147.
Tasic, B., et al., 2017. Shared and distinct transcriptomic cell types across neocortical areas. bioRxiv 229542. https://doi.org/10.1101/229542.
Tervo, D.G.R., et al., 2016. A designer AAV variant permits efficient retrograde access to projection neurons. Neuron 92, 372–382.
Trojanowski, J.Q., Gonatas, J.O., Gonatas, N.K., 1982. Horseradish peroxidase (HRP) conjugates of cholera toxin and lectins are more sensitive retrogradely transported markers than free HRP. Brain Res. 231, 33–50.
Ugolini, G., Kuypers, H.G.J.M., Simmons, A., 1987. Retrograde transneuronal transfer of Herpes simplex virus type 1 (HSV 1) from motoneurones. Brain Res. 422, 242–256.
Underwood, E., 2016. Neuroscience. Barcoding the brain. Science 351, 799–800.
van den Heuvel, M.P., Bullmore, E.T., Sporns, O., 2016. Comparative Connectomics. Trends Cogn. Sci. 20, 345–361.
Wahlsten, D., Metten, P., Crabbe, J.C., 2003. Survey of 21 inbred mouse strains in two laboratories reveals that BTBR T/+ tf/tf has severely reduced hippocampal commissure and absent corpus callosum. Brain Res. 971, 47–54.
Wan, X.C., Trojanowski, J.Q., Gonatas, J.O., 1982. Cholera toxin and wheat germ agglutinin conjugates as neuroanatomical probes: their uptake and clearance, transganglionic and retrograde transport and sensitivity. Brain Res. 243, 215–224.
Wang, X., et al., 2018. Three-dimensional intact-tissue sequencing of single-cell transcriptional states. Science 361, eaat5691.
Wassie, A.T., Zhao, Y., Boyden, E.S., 2019. Expansion microscopy: principles and uses in biological research. Nat. Methods 16, 33–41.
Waterhouse, B.D., Lin, C.S., Burne, R.A., Woodward, D.J., 1983. The distribution of neocortical projection neurons in the locus coeruleus. J. Comp. Neurol. 217, 418–431.
Welch, J., et al., 2019. Integrative Inference of Brain Cell Similarities and Differences from Single-cell Genomics. https://doi.org/10.1101/459891.
White, J.G., Southgate, E., Thomson, J.N., Brenner, S., 1986. The structure of the nervous system of the nematode Caenorhabditis elegans. Philos. Trans. R. Soc. Lond., B, Biol. Sci. 314, 1–340.
Winnubst, J., et al., 2019. Reconstruction of 1,000 projection neurons reveals new cell types and organization of long-range connectivity in the mouse brain. bioRxiv 537233. https://doi.org/10.1101/537233.
Yamashita, T., et al., 2013. Membrane potential dynamics of neocortical projection neurons driving target-specific signals. Neuron 80, 1477–1490.
Yoon, Y.-G., et al., 2017. Feasibility of 3D reconstruction of neural morphology using expansion microscopy and barcode-guided agglomeration. Front. Comput. Neurosci. 11, 97.
Zador, A.M., et al., 2012. Sequencing the connectome. PLoS Biol. 10, e1001411.
Zeisel, A., et al., 2018. Molecular architecture of the mouse nervous system. Cell 174, 999–1014 e22.
Zheng, Z., et al., 2018. A complete electron microscopy volume of the brain of adult Drosophila melanogaster. Cell 174, 730–743 e22.
Zingg, B., et al., 2014. Neural networks of the mouse neocortex. Cell 156, 1096–1111.