Graded lines:
Using genetic variation in neuronal projections to understand functional organization

James W. Phillips

Magdalene College

Original Submission: September 2018

Corrections submitted: June 2019

This dissertation is submitted for the degree of

doctor of philosophy at

Department of Physiology, Development and Neuroscience

In the University of Cambridge

Declaration of originality

This dissertation is the result of my own work and includes nothing which is the outcome of work done by others or in collaboration, except where specifically indicated in the text. The thesis is not substantially the same as any that I have submitted, or will be submitting, for a degree, diploma or other qualification at this or any other university. This thesis does not exceed the limit of length (60,000 words) specified by the Degree Committee for the Faculty of Biology.

Name: James Phillips.....

Date: 10th July 2019

Abstract

James W. Phillips,

Graded lines: Using genetic variation in neuronal pathways to understand functional organization

In this thesis, I investigated the interactions between brain regions in mice, with an emphasis on the motor system. Whilst brain areas are typically studied in isolation, they exist to act upon one other, and we lack even relatively basic principles of how brain areas interact. Here, my colleagues and I used new genetic techniques to investigate this topic. There were two major focuses. Firstly, I investigated how thalamic pathways vary, using genetic sequencing to produce a high dimensional transcriptomic profile of forebrain communication pathways. This revealed a strikingly simple organisational system in rodent thalamus, with pathways serving diverse modalities varying in a common manner. Pathways serving systems as variable as vision, navigation, motor control and somatosensation showed a similar variation in gene expression, which was functionally relevant. This establishes a common reference frame for diverse modalities of cognition. Secondly, I began an ongoing investigation of the signals that motor cortex send to subcortical motor structures. This revealed a separation of signals that project to the basal ganglia and cerebellum, opposite to existing predictions. Basal Ganglia-projecting signals encoded motor execution, whilst cerebellar projecting pathways encoded preparation- and reward-related information. These experiments motivated further investigation of these pathways, which are described within. Together, these results

demonstrate the utility of focussing upon interactions between network nodes to reveal the contribution of individual nodes.

For my parents, for everything.

'...To explicate the uses of the Brain, seems as difficult a task as to paint the Soul, of which it is commonly said, That it understands all things but itself;'

Thomas Willis, Preface to Cerebri Anatome, Oxford, 1664

"My soul is a hidden orchestra; I know not what instruments, what fiddlestrings and harps, drums and tamboura I sound and clash inside myself. All I hear is the symphony."

Fernando Pessoa, The Book of Disquiet

Thesis chapters

Chapter 1 – *Introduction: Embodied signals*

Chapter 2 – *A common molecular and functional organization across modality in thalamus*

Chapter 3 – *Dissociable dynamics between genetically defined motor cortex outputs*

Chapter 4 – *Discussion: Graded lines: Examining continuous variability to reveal biological organization*

Thanks and acknowledgements

My PhD years have been complicated, and so I have many thanks to make.

I thank the entire thalamoSeq team, who tolerated my foibles and believed in the project from its time as an embryonic project. I especially mention Erina Hara and Anton Schulmann, whom I cannot thank enough, and Adam Hantman, whom encouraged and trusted such a juvenile scientist as myself to undertake this project . They know my faults better than any, and yet continued to believe in our team project. I also thank Nela Cicmil, John Stein, Richard Vaughan-Jones, for their exceptional mentoring. Finally, for the above and more, I thank Joshua Dudman and Steve Edgley.

I thank Sue Jones, Maryrose Franko, Ulrike Heberlein, Wyatt Korff, Chongxi Lai and Virginia Rutten for their friendship and advice throughout my PhD.

I thank the medical staff of the UK national health service who have cared for me since I was 5 weeks old, surgically operated on me, and supported me having the time needed to recover. I especially thank Dr Simon Curtis, Dr Khalid Rashid, Mr Andrew Stewart and Dr Saj Sengupta, and all of the Nurses in intensive care and my ward at Taunton hospital. I thank Jack Ellis, who made sure I could go back to being a scientist.

For Yves Weissenberger, I simply state that has been the very best of friends.

And finally, my parents, for everything.

Chapter 1 – Introduction: Embodied signals

This thesis uses recent advances in genetic tools to investigate the interactions between brain regions. The two experimental chapters differ substantially in their focus, and so have their own introductions. The purpose of this chapter is to introduce the common theme that unites them at a high level, and to explain the motivation for investigating the question. More specific detail is in each chapter. In the general thesis discussion, I seek to draw joint insights from the two chapters.

This thesis centers upon interactions between brain regions

A key goal of systems neuroscientists is to identify principles that explain how brains are organized and operate. These principles should ideally extend across both modalities (such as vision and motor control) and species, thus revealing general principles of nervous systems. Identifying such principles is of relevance not only for our understanding of brains, but also for the study and design of intelligent systems as a whole.

This search can be approached at multiple levels. At the smallest level, the features of individual neurons can be linked to their function: active dendrites, synaptic plasticity and sensitivity to neuromodulators have all been observed in multiple brain regions in different species, in systems processing quite different modalities of information, from the hippocampus to visual cortex to cerebellum (Citri & Malenka, 2007; Häusser, Spruston, & Stuart, 2000; Kandel, Jesssell, & Schwartz, 2012). At the local circuit level, feedforward inhibition, excitatory/inhibitory balance and lateral inhibition offer common

concepts/principles for wiring of different brain regions (Carandini, 2012; Denève & Machens, 2016; Koyama & Pujala, 2018; Pouille & Scanziani, 2001). At the more global level, ideas such as attention, reinforcement learning and the use of internal models offer broad concepts, again in a cross-species manner(Knudsen, 2007; Wolpert & Ghahramani, 2000). Further, these ideas are increasingly influencing the development of artificially intelligent systems, demonstrating their general use beyond the specific cases in which they were discovered(Hassabis, Kumaran, Summerfield, & Botvinick, 2017; Versace, Martinho-Truswell, Kacelnik, & Vallortigara, 2018). However, at present there is a significant gap in our understanding, and it involves how different brain regions interact to process information and produce useful output. Even the simplest of brains are composed of many different brain regions, each of which are strongly and typically reciprocally connected, often in loop-like arrangements connecting the regions(Alexander, DeLong, & Strick, 1986). It is at this level of analysis that we are relatively lacking in common principles. In short, we do not know if the same organizational principles are used to wire up the parts of the visual or somatosensory system are used to wire up parts of the motor or navigation-related systems.

This lack of knowledge may in part arise from the approach taken to understanding the disparate brain regions involved in a common operation. From the origins of neurology with Thomas Willis, scientists have sought to assign specific behavioural functions to specific brain regions(Molnár, 2004; Willis, 1664). This functional approach largely persists to this day. For example, in the visual circuitry of primates, the frontal eye fields are commonly ascribed the role of directing attention, the ventral visual stream to categorizing faces, and the dorsal visual stream to detecting motion(Goodale & Milner, 1992). Critically, these are behavioural functions.

A problem arises, however, because the link between behavior and neural activity is necessarily indirect, as most neurons are many synapses from behavioural effectors. Further, even in vision, signals relevant to a given computation or correlating with a

particular aspect of behavior are broadly distributed across the brain, disruption to a given behavior can occur from perturbation of multiple brain region, and these areas are often extensively interconnected(van Polanen & Davare, 2015; Zénon & Krauzlis, 2012). This suggests that an alternative viewpoint could be beneficial in producing a comprehensive understanding.

An alternative to investigating the *function* of a brain region is to investigate its *contribution*. This subtle change in wording moves us from seeing brain regions as soloists to viewing them as more as participants, team members and perhaps even adversaries in an active conversation. Critically, it leads us to instinctively focus not just upon behavior, but also upon how a brain region alters activity in other brain areas. Ie, what does brain region A do to brain region B in order to cause its apparent effects upon behavior? Here we have relatively few general principles. Are there equivalents of lateral inhibition, feedforward inhibition and excitatory/inhibitory balance at the brain-wide level? This, in broad terms, is the topic of this thesis.

Brain regions are engaged in a continual 'conversation'

Before outlining the approach to this problem, it is beneficial to review some of what is known about the relationship between brain regions. Lets consider two possibilities, lying at either end of a spectrum. One view, the 'switch board' view, is that brain areas engage in their own specific computations, for example selecting an action or identifying a locus for attention, and only provide an output to other areas once this computation is complete. Their contribution to the activity in other areas is therefore relatively minimal, akin to signaling a decision. Most of the time, brain areas would have minimal effect upon one another's activity. This view is analogous to that of a computer, where information is only sent once a local computation has been completed. Indeed, in the early conceptions of computing systems, this serial, step-by-step manner of computation was specifically argued by Von Neumann and others to be simpler, and thus easier to

engineer(Neumann, 1958; Waldrop, 2002). Until very recently, most computers have worked in a largely serial manner.

However, there is another possible extreme. This is that brain areas are engaged in a continual dialogue, with information being passed between them moment-by-moment. In such a view, each area is actively sculpting the activity in other brain areas, millisecond-by-millisecond. The activity in each brain region is intimately interconnected, and can only be fully understood in the context of each other branch area. Far from serial computation, such a brain would be not only parallel but indeed simultaneous and interactive: computation in any one brain area has marked effects elsewhere in the circuit.

Though a switch-board like metaphor for thought has been pervasive, the current evidence instead strongly supports the latter view. Far from a 'switch board' like arrangement, there is indeed abundant evidence for the view of an ongoing and behaviourally relevant interaction between brain regions, which we outline below. That is to say, regions continually, substantially and rapidly sculpt one another's activity. Thought is more like a conversation than a series of forcibly siloed soloists.

Network interactions in mammalian motor circuits

We will now focus upon the motor control circuit to expand upon, explain and illustrate these core points. The motor circuit provides a useful model for such studies of interactions, as it ultimately work produces a single, quantifiable output: the resulting behavior. Further, disorders of the motor system are both diverse and have been studied extensively, allowing scientists to leverage a wealth of clinical data to complement more discovery-focused, basic experiments.

Just as we outlined approximate discrete functions for different visual areas above, a similar discretization of function in the motor control realm has been attempted. We will focus upon three major nodes of the motor network: the basal ganglia, the cerebellum and the motor cortex. There are of course many other regions, but these are perhaps best studied.

Each of these three areas has been theorized to have specific functions in motor control. These proposed roles have been inferred based upon a combination of studying the effect of lesions and of examining the neural activity in each area.

For example, in motor control the Basal Ganglia have most commonly been argued to underlie action selection. This has been based upon the deficit in action initiation seen in disorders affecting basal ganglia, such as Parkinsonism and Huntington's disease. Further, Hikosaka's experiments in the 1980s show impairment of saccade selection following inactivation of the basal ganglia output substantia nigra reticulata (SNr) in primates, and a phasic reduction of output activity around saccade onset (Hikosaka and Wurtz 1983a). Since the SNr is inhibitory, this reduction would provide disinhibition of the downstream motor targets.

The cerebellum, by contrast, has been argued to contribute to internal models that compensate for internal delay lines in motor control systems (Miall et al 1993, 1993; Wolpert & Ghahramani, 2000). This has been based upon the observation that action selection *per se* is not impaired following lesions of this structure, but that the movements develop a disordered, clumsy manner. Further, when sensory feedback is delayed to the actor, movement develops deficits similar to those seen in cerebellar lesion patients (Miall et al., 1993).

The role of motor cortex in movement is perhaps the most enigmatic, and we review this specifically in chapter 3. Briefly, it has been linked to functions as diverse as action

selection, control of movement kinematics, feedback correction of errors and tutor-like roles for subcortical circuits(Crevecoeur, Thonnard, Lefevre, & Scott, 2016; Heindorf, Arber, & Keller, 2018; Shenoy, Sahani, & Churchland, 2013a). These are of course not mutually exclusive.

Whether these ascribed functions fully explain the functions of these areas is not the central point: the core point is that there is a proposed division of labour in the system, which allows us to study how the said division of labour is ultimately integrated to produce behaviour.

Despite this need for integration of activity, elements of the motor circuit are typically studied in isolation, with a corresponding deficit in understanding. That is to say, the relationship between activity in one area of the motor circuit and another area during behaviour is poorly understood. For example, the author is unaware of any study inactivating either the basal ganglia or cerebellar outputs and recording the resulting effects in motor circuits published prior to the past 2 years using electrophysiology, or indeed of recording from any two such sites simultaneously to examine activity relationships in normal behaving animals (but see (McCairn & Turner, 2015), with focus upon the diseased state). Inactivations without electrophysiological recordings are, however, frequent (eg, (Horak & Anderson, 1984)). There are some such studies in other area of motor control, but they are again few in number (such as thalamocortical processing, (K. C. Nakamura, Sharott, & Magill, 2012; Wurtz, McAlonan, Cavanaugh, & Berman, 2011)).

However, in the course of studying dysfunction in motor control, it has been apparent that abnormal activity is rarely consigned to a single brain region, suggesting network-level disruption. This has been done with sensors of metabolic markers, such as functional magnetic resonance imaging or 2-deoxyglucose uptake studies. For example, abnormal activity is seen in both the Basal Ganglia and Cerebellum in human dystonia,

in a correlated manner, suggesting an abnormal interaction between these two regions(Neychev, Fan, Mitev, Hess, & Jinnah, 2008). This has classically been interpreted as one diseased brain area driving dysfunction in other areas (Carrera & Tononi, 2014; Finger, Koehler, & Jagella, 2004; Otchy, Wolff, Rhee, Pehlevan, Kawai, Kempf, Gobes, & Ölveczky, 2015a; Schiff, 2008). Indeed, perturbative studies which artificially recreate motor deficits in animal models reproduce the network-wide abnormalities. Local perturbations to the basal ganglia's subnuclei lead to marked metabolic changes outside the basal ganglia in primates, in a manner similar to that seen in human patients with conditions such dystonia and Parkinsonism (Crossman, 2000; Crossman, Mitchell, & Sambrook, 1985; Eidelberg, 2009; Huang et al., 2007; Mitchell, Boyce, Sambrook, & Crossman, 1992; Mitchell, Jackson, Sambrook, & Crossman, 1985; Mitchell et al., 1990). Further, interventions and disease treatments that alter such network interactions also alter disease symptoms, and likely form the mechanism of symptom improvement in functional neurosurgery (Eidelberg, 2009; Stein & Aziz, 1999). Perhaps most intriguingly, lesions to the output of a diseased basal ganglia can yield marked improvements in motor symptoms(Hornyak, Rovit, Simon, & Couldwell, 2001). This is argued to be because the output is essentially flooding the brain with 'faulty' signals that the brain would be better off without, like switching off a faulty speaker (Stein and Aziz, 1999). Together, these results imply that what happens in one area of the motor network has profound effects upon other areas, and that this can be critical in dysfunction.

Direct experimental demonstration, through electrophysiology, of the pronounced effect of one region upon another can be seen in a pair of studies on cerebello-basal ganglia interaction in primates and rodents. In both lines of research, a basal ganglia-associated symptom is shown to be caused by a dysfunctional cerebellum driving dysfunction in a previously 'normal' basal ganglia. McCairn *et al* show that pharmacological disruption of the cerebellum can induce tic-like symptoms in primates, and this leads to distorted activity in the basal ganglia, the structure more typically linked to tics(McCairn,

Bronfeld, Belelovsky, & Bar-Gad, 2009; McCairn, Iriki, & Isoda, 2013). In rodents, the Khodakhah lab shows that a dystonia induced by manipulation of the cerebellum causes its effects via a cerebello-thalamo-basal ganglia projection, and that cutting this projection prevents the motor dysfunction(Calderon, Fremont, Kraenzlin, & Khodakhah, 2011; C. H. Chen, Fremont, Arteaga-Bracho, & Khodakhah, 2014). It is notable however, that though a role for this projection in dysfunction is now known, its contribution to normal function has only very recently been investigated(Le Xiao, Bornmann, Hatstatt-Burklé, & Scheiffele, 2018; Pidoux, Le Blanc, Levenes, & Leblois, 2018).

These examples all come from the dysfunctional brain, but signatures of ongoing interactivity are apparent in the normal brain also. The most direct evidence for moment-by-moment interaction between brain areas during behavior comes from oscillatory synchrony between regions, and its fluctuation with task features(Buzsaki, 2006). Rather than fluctuating independently, neural activity in different areas synchronizes at particular frequencies. Different frequencies are linked to different aspects of behavior, including relatively abstract components of it. For example, beta-oscillations in the cortico-basal ganglia network are correlated to movement preparation and post-movement feedback in a manner not entirely explained by the movement itself, though its precise role is not fully understood (H. Tan, Wade, & Brown, 2016; Torrecillos, Alayrangues, Kilavik, & Malfait, 2015; Zanos, Rembado, Chen, & Fetz, 2018). Fronto-visual cortex oscillations have been linked to attentional processes across multiple species(Thiele & Bellgrove, 2018). Importantly, such coherent oscillations occur not just during behavior, when task-related information could be being communicated, but also in the resting and sleeping state(Buzsaki, 2006). We thus have a view of a dynamic, interactive brain, with different frequencies of oscillatory activity between areas being reliable markers for different cognitive events.

Here, there is a gap: despite the pervasive nature of these oscillations, to date we lack a mechanistic understanding of how these oscillatory events occur, and how they are

regulated. What network events promote alpha-oscillatory activity, for example(Thiele & Bellgrove, 2018)? Why does a depletion of dopamine promote synchrony in the beta-frequency range(Little & Brown, 2014)? How is neural activity at gamma-frequency transformed differently through the same network node? These unanswered questions reflects a broader lack of knowledge regarding such brain-wide interactions, even at a relatively qualitative and phenomenological level(Buzsaki, 2006; Feingold, Lundqvist, Feingold, & Lundqvist, 2018). We now turn to outlining a potential path to understanding at least some of these processes at a mechanistic level.

Understanding embodied signals to understand the interaction between brain regions

What might 'understanding' approximately look like in such a system? We can envisage many ways that brain areas could alter the activity of one another, and how this could be supported mechanistically. An area could act to excite or inhibit the other, or perhaps some combination of these. This could obviously be achieved via differences in postsynaptic target and/or neurotransmitter choice. More subtly, a pathway could alter the gain of the activity in its target, as opposed to driving spikes directly. The former could be achieved by targeting the more distal regions of a pyramidal neuron dendrite, for instance, whilst the latter would be achieved via more proximal targeting. These areas have previously been shown to integrate information over different timescales, and to have different active dendritic conductances(Branco & Häusser, 2010; 2011; Branco, Clark, & Häusser, 2010). Other possibilities are that the input pathway could be specialized for carrying or even amplifying particular oscillatory frequencies, acting as a band-pass filter(Puil, Meiri, & Yarom, 1994). It could also act as a plasticity-inducing signal, for instance by differentially elevating postsynaptic calcium (linked to induction of synaptic weight changes)(Ellender, Harwood, Kosillo, Capogna, & Bolam, 2013). These examples are far from exhaustive.

The crucial point is that it is not just the message, but the messenger, that can be specialized for function. Notably, this is quite different to the viewpoint promoted by theory of communication by Shannon and others, which explicitly seeks to treat the message separately from the messenger(Shannon, 1997; Waldrop, 2002), and has been influential in seeking to understand neural computation(Rolls, Treves, Tovee, & Panzeri, 1997). Notably, Shannon himself warned of excessive, 'bandwagon'-like application of the tool he and others developed(Dimitrov, Lazar, & Victor, 2011). Indeed, this may be true in nervous systems, where the message alone is not all that matters. Rather, in nervous systems, messages are inherently embodied, and we must understand that body to understand the message. The properties of this body, the projection neurons, can vary as described above, greatly altering the effect of the message. It follows that an understanding of inter-area interaction would be facilitated by investigating the physical properties of the pathways involved. However, despite these being relatively basic properties of a given pathway, they are known for surprisingly few projections. There has been too much emphasis on the message, and not enough on the identity and properties of the messenger.

With this interaction-centered perspective, lets return to the motor control system, and the three elements that we spoke of earlier: the motor cortex, basal ganglia and cerebellum. What are the properties of these connections, what signals are exchanged, and how might these elements combine to produce the resulting network dynamics? We will explore the two sides of this in greater detail in later chapters: for now, a simple outline of the two sides of this problem that we shall investigate in this thesis.

Both the basal ganglia and cerebellum project to the motor cortex via the thalamus, targeting largely separate thalamic nuclei(Kuramoto et al., 2010; 2009). Might the pathways that carry these signals have specializations that facilitate computation, and might understanding these specializations help us understand the circuit better? For

example, the basal ganglia has been linked to reward based learning at relatively longer timescales, whilst the cerebellum has been linked to error-correction and learning at shorter timescales(Doya, 2000; Kunimatsu, Suzuki, Ohmae, & Tanaka, 2018; Lintz & Felsen, 2016). One might expect the two thalamic pathways carrying basal ganglia and cerebellar signals possesses differential sensitivity to reward-related dopamine, or target different parts of the dendrites of cortical neurons to be integrated over different timescales(see above for relevance of dendritic targeting location)? There is indeed some existing evidence for this, with differential layers of cortex being targeted by basal ganglia and cerebellar thalamus(Kuramoto et al., 2009). However, beyond this, relatively little is understood of the molecular differences, if any, between these pathways (existing studies do not well discriminate between subnuclei of the motor thalamus(Sawyer, Young, Groves, & Tepper, 1994)). Investigating these questions requires the ability to isolate the different pathways experimentally, an already challenging task given their small size and lack of established markers. Further, the range of potential variation in their properties is high-dimensional (gene expression alone: approximately 25,000), making case-by-case examination of them extremely burdensome. This to date has limited our understanding of how subcortical-cortical connectivity is specialized for contribution and function. Though motivated by motor control, these same difficulties apply across the investigation of all thalamic pathways. Addressing this will be the first half of the thesis.

In the second part of this thesis, we will investigate the other side of this connectivity: how the Basal ganglia and Cerebellum are influenced by the motor cortex. Both the basal ganglia and cerebellum receive relatively direct input from motor cortex, with some, though not total, separation of such projections (detailed in chapter 3). Though researchers often focus upon components directly related to movement, motor cortex itself exhibits a surprisingly diverse range of signals relating to behavior. Given the differing proposed roles of the basal ganglia and cerebellum in movement, might there be differences in the signals they receive from motor cortex? Stated another way, are a select

subset of the diversity of motor cortex signals routed preferentially to these two subcortical motor targets? For instance, could more context-related signals be sent to the basal ganglia to enable action selection, whilst signals more directly related to movement execution be sent to the cerebellum, to enable comparison of planned vs actual motor output and feedback(Reiner, 2010)? Achieving such targeted recordings has until recently required antidromic-collision testing to verify projection targeting of neurons, a technique that is both time intensive and greatly wasteful in terms of unit yield(Turner & DeLong, 2000). The topic of the second experimental chapter is addressing this question with recently developed, higher-throughput techniques.

Recent methodological advances facilitate mechanistic investigation of brain-wide circuits

Both experimental chapters rely heavily upon recent advances in technology, particularly genetics, to bypass these limitations that have slowed targeted investigation of the relevant projection neurons. Broadly, projection neurons that connect brain areas are interspersed with interneurons, which until recently made selective recording, perturbation or molecular analysis cumbersome. To record from multiple regions simultaneously at single unit resolution was more cumbersome still. However, this thesis was begun at a time when advances in genetics and neural recording technologies offered a new set of tools, allowing for mechanistic study of the interaction between brain regions at a very significantly improved level. Together, these advances facilitate qualitatively new kinds of experiments.

Two advances in genetics drive progress in two major ways. The first advance is the development of next-generation sequencing, RNAseq. RNAseq provides a transcriptome-wide read out of gene expression levels in a sample, giving indirect quantification of a wide range of cellular properties(Tasic et al., 2016). This provides a much more detailed and broad understanding of the cell types involved in said interactions than has been

possible with prior molecular profiling technologies. This is of clear benefit for investigating brain interactions. Not only can one provide a classification system for the projections between brain regions, one can use it to predict their contribution. One could, for example, identify candidate neuromodulator expression, or the presence of ion channels linked to particular firing rate profiles.

Further, it reveals the specific genes and gene classes that drive the observed differences. One can use this as a tool not only to produce classification of projection types, but also to hypothesize regarding the functions of said projections. Despite this, at the time this thesis was begun, nobody had done projection-specific RNAseq. This is the 1st genetics advance that we seek to leverage.

The second critical benefit of the genetics revolution in neuroscience is the burgeoning number of transgenic lines enabling selective targeting of neural populations with sensor and effector proteins. Once these lines are combined with other technologies, such as virus expression/ for example expressing cre-recombinase, allow experimentalists to record and perturb both more easily and more readily than before(Gerfen, Paletzki, & Heintz, 2013a; Tervo et al., 2016). This, combined with the RNAseq-facilitated increase in knowledge about cells to be targeted, allows an experimental precision not possible before(Gerfen, Paletzki, & Heintz, 2013b).

There is another, parallel advance in neuronal recording technologies. In optical imaging and electrophysiology, there have been orders of magnitude improvements in sensitivity and recording scale respectively(T.-W. Chen, Wardill, Sun, Pulver, Renninger, Baohan, Schreiter, Kerr, Orger, Jayaraman, Looger, Svoboda, & Kim, 2013b; Jun et al., 2017). It is now possible to record as many neurons with electrophysiology in a day as one could have recorded in a month less than a decade ago, and sensitive calcium sensors allow one to record with high fidelity from large numbers of genetically targeted neuronal populations.

This thesis uses these technologies to address the limitations addressed above, namely the challenge of investigating cortico-subcortical circuits, specifically thalamo-cortical and motor circuits. The first experimental chapter uses projection-specific RNAseq to investigate the organizational logic of the thalamus. Thalamus acts as the central hub for communication to and between forebrain areas, and so the organizational properties of its nuclei have a critical effect upon interactions between brain regions. By understanding more about the organization of the pathways themselves, we seek to provide a foundation for a functional anatomy of the brain's central communication hub. This in turn should provide insight into how information from subcortical areas reaches neocortex, as it will not only reveal molecular properties of those pathways but also facilitate their targeting with transgenic tools. The second chapter uses two-photon imaging to investigate the contribution of motor cortex during behavior. More specifically, it uses transgenic mice to target different output pathways from motor cortex to the basal ganglia and cerebellum for imaging during behavior.

Chapter 2 –
A common molecular and functional organization across modality in thalamus

Note on the candidate's contribution:

The research project described in this chapter began as a pilot project and became a Janelia small project team (ThalamoSeq project), both initiated and led by the myself. The resulting work is therefore highly collaborative. I contributed to all aspects of the project with the exception of carrying out the *in vitro* electrophysiological recordings, which were performed by Chenghao Liu in Sacha Nelson's lab at Brandeis.

This project has been posted as a preprint as below. I have only included aspects of the project to which I performed analysis. For this reason, I have not included the single cell anatomical reconstruction data present in later versions of the paper, as I did not do this analysis (though I had intellectual input to this).

Phillips, J.W., Schulmann, A., Hara, E., Liu, C., Wang, L., Shields, B.C., Korff, W., Lemire, A.L., Dudman, J., Nelson, S.B., Hantman A., 2018. A single spectrum of neuronal identities across thalamus, 241315. http://doi.org/10.1101/241315.

In which I was co-first author with AS and co-corresponding author with AH and SN.

2.1 – Introduction

Thalamus as the central communication hub of the forebrain

Brains process a wide range of information modalities, from vision to somatosensation, through to motor signals, through to higher cognitive signals involved in navigation and

planning. Uncovering common principles by which these diverse modalities of information are processed is a fundamental goal in neuroscience. However, we currently lack common, shared principles to how information is transformed across the brain's different diverse circuits. In this chapter, we study thalamus to gain insight into these questions.

The Thalamus, or *inner chamber*, sits at the center of the vertebrate brain(E. Jones, 2007). Present from fish through to humans, it is part of the diencephalon, lying between the telencephalon and the mesencephalon(Mueller, 2012). Matching its positional centrality, it is connectomically central to a diverse range of neural systems: It acts as a communication hub for almost every modality of information in the brain, and is thus an ideal place to examine shared, cross-modal organizational motifs. Further, and relatedly, it serves as an excellent model for the themes explored in the thesis introduction: the principles via which one area acts upon another, and whether conserved principles of this exist.

This role as a communication hub is demonstrated by its connectivity. Thalamus is almost exclusively a projection nucleus, lacking internal collaterals or local interneurons (the lack of interneurons in most nuclei is rodent-specific). The projections it makes beyond thalamus, however, are large, diverse and widespread. Almost all information to reach neocortex comes from these thalamic pathways, as does a significant fraction of input to subcortical structures, including the amygdala, striatum and hippocampus. Without thalamus, then, the cerebral cortex would be deprived of its knowledge of the world. It is thus a critical network node to understand.

Approaches and difficulties in categorizing thalamic nuclei

The diversity and centrality of thalamic pathways have made thalamus a center of investigation since the very first systematic studies of brains. In the 1660s, Thomas Willis and colleagues traced the optic nerve from the eye through to the thalamus, finding

that it terminated in a relatively small, circumscribed region, which Willis named the ‘Optic thalamus’(Molnár, 2004; Willis, 1664)(Figure 2.1A).The role of this optic thalamus was argued to be to carry visual signals, or ‘animal spirits’, from the eye through to the cerebral cortex, which Willis saw as the seat of higher cognitive function. Thalamus was thus seen as serving a relay function, acting as a gateway to cognition.

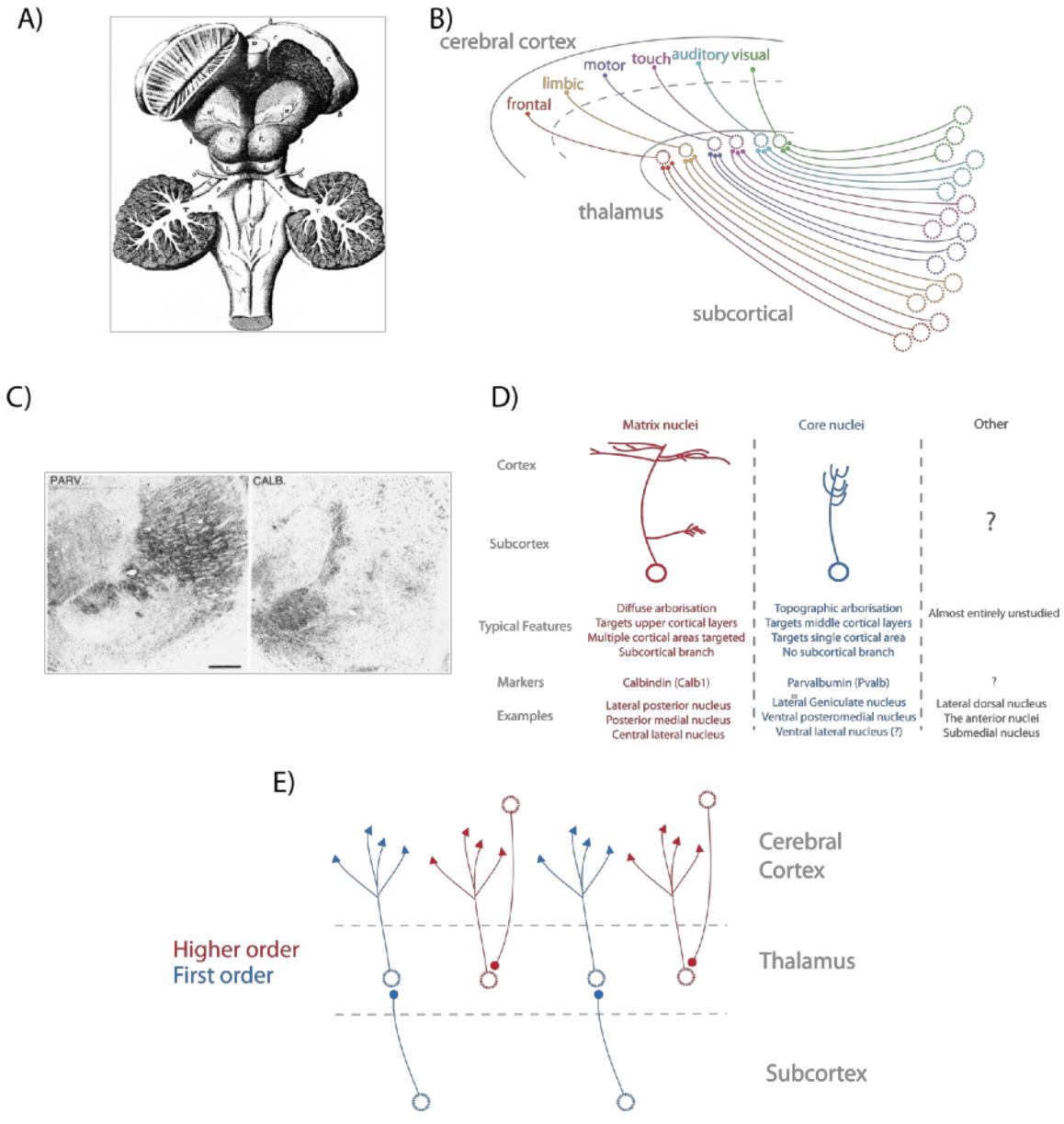


Figure 2.1 – Previous organizational schemes of thalamus divide it by modality or hierarchical level

a) Modality-based approaches to categorising thalamic nuclei. Image shows a depiction of the sheep brain, including the first depiction of the optic thalamus, from(Willis, 1664). Image taken from (Parent, 2012).
b) A schematic of thalamo-cortical pathways, with thalamic nuclei coloured by the modality of the subcortical system they receive input from.
c) Complementary expression of calcium binding proteins calbindin and parvalbum in thalamus. Expression of parvalbumin (left) and calbindin (right) in the primate thalamus. Image from(E. Jones, 1998). Note that several regions do not label strongly for either marker. Scale bar is 750um.
d) Features of the core/matrix nuclear classes. The core/matrix scheme divides thalamic nuclei into two major groups, based upon the expression of parvalbumin (core nuclei) or calbindin (matrix nuclei) and the corresponding differences in cortical arborisation pattern. However, some nuclei express neither marker.
e) The first/higher order nucleus scheme of Sherman and Guillery divides thalamic nuclei into two discrete groups, based upon whether they receive their driving input from cortex or subcortex.

This connectionist approach to describing and investigating thalamus remained the dominant approach until the turn of the 21st century. Via a combination of retrograde and anterograde tracing, thalamic nuclei projecting to and receiving input from a wide range

of modalities were identified, including motor-, limbic-, somatosensory-, auditory-, navigation-related regions(E. Jones, 2007; Paxinos & Franklin, 2012). For example, a particularly dense and large cluster of cells is found in the caudal lateral thalamus following injections to auditory cortex, whilst a different one, much more rostral, is identified following injections to motor cortex. Further, these nuclei receive input from corresponding subcortical structures, such as motor thalamus receiving input from the cerebellum (Figure 2.1B). These are thus referred to as auditory and motor thalamic nuclei respectively. This approach has an obvious assumption: this is that thalamic nuclei should be defined by the *modality* of information that they process. This we refer to as the modality-based taxonomy.

However, thalamic nuclei also differ in dimensions other than just the modality of information relayed. Firstly, there are multiple thalamic nuclei projecting to each region of neocortex. For example, injecting motor cortex with a retrograde tracer leads to labelling not only of classical lateral 'motor thalamus', but also of other smaller, more medial nuclei such as the rostral intralaminar nuclei. In vision, the spatially adjacent lateral posterior nucleus (LP) and dorsal lateral geniculate nucleus (LGd) both project to visual cortex, but only the LGd gets direct input from the retina. This led to the view of there being 'principal' thalamic neurons, dedicated to more direct sensory processing(E. Jones, 2007). Similar diversity has been observed in all thalamic modalities, leading to atlas' subdividing thalamus into approximately thirty nuclei, with the exact number varying by species and atlas (E. Jones, 2007; Paxinos & Franklin, 2012). Modality is therefore not the sole characteristic that varies between thalamus pathways.

This proliferation of identified nuclei serving each modality is only the first of several concerns that indicate a need for a revised taxonomy. Another is that perturbations of nuclei projecting to the same area of cortex can produce markedly different effects upon behavior. Perturbation of the motor thalamus leads to relatively modest effects upon the kinematics of movement, whilst inactivation or stimulation of the rostral intralaminar

nuclei can induce coma's or wakefulness respectively(Miall et al., 1998; Redinbaugh et al., 2017; van Donkelaar, Stein, Passingham, & Miall, 1999; 2000). This is particularly intriguing as the classic motor thalamus is, by volume and neuron count, a much larger nuclear group than the rostral intralaminar thalamic nuclei, yet has apparently much more subtle effects upon levels of behavioural activity when perturbed, at least in the dimensions typically quantified. This has led to additional divisions of nuclei based on functional grounds, with the medial nuclei being referred to in terms of arousal and alertness, whilst the more lateral groups are referred to more in terms of classical information relay functions(Smith et al., 2014). Different nuclei, serving the same modality, are therefore not simply duplicates of one another, but can have distinct functional roles.

These particular objections are still consistent with thalamus being a relay nucleus, albeit a diverse one. However, further investigation was suggestive of thalamus having direct roles in sculpting information flow, beyond simply conveying information. Two aspects of this are particularly critical going forward, and form the basis of two different approaches for viewing the organisation of thalamus. Both were proposed at the turn the 21st century, and have been highly influential in the field. We will review both of these now, as they are relevant both for this chapter and for the overall concluding thesis discussion.

Core/Matrix theory divides thalamus based on neurochemical and anatomical properties

The first is the Core/Matrix theory, proposed by Edward Jones(E. Jones, 1998; 2007; E. G. Jones, 2001; E. G. Jones & Hendry, 1989). It argues for a dichotomy of thalamic neurons, based upon a combination of neurochemical and anatomical information. It is based upon three observations.

Firstly, thalamic neurons vary not only in which cortical areas they target, but also in the layers and extent of their axonal terminations. This difference can be quite dramatic: some neurons send axons to single cortical areas, topographically, terminating predominantly in the middle cortical layers, whilst others can send axons across many cortical areas, with extensive axonal spread in cortical layer 1(Rubio-Garrido, Pérez-de-Manzo, & Clascá, 2007). This leads to a distinction between narrowly and diffusively arborizing thalamic axons.

The second observation is that thalamic neurons label with antibodies against either parvalbumin or calbindin (but see below), and this occurs across modalities (Figure 2.1C). For example, both visual and somatosensory thalamic nuclei have parvalbumin expressing neurons. This means that neurochemically, at least, thalamic nuclei of different modalities can be more similar than thalamic nuclei of the same modality. This separate expression pattern is shown in figure 2.1C.

Finally, these two observations are correlated: neurons with diffuse arborisations express calbindin, whilst neurons with narrow arborisations express parvalbumin (as examined below, the absoluteness/discreteness of this difference has, to the authors knowledge, not been quantified). Thus, the core/matrix theory proposes a dichotomy of thalamic neurons, based upon their axonal arborisation pattern, marked by calcium binding protein expression pattern. ‘Core’ parvalbumin-expressing neurons have a dense, local termination pattern, i.e. a ‘core’, whilst ‘matrix’ calbindin-expressing nuclei project broadly in layer 1, as well as relatively more focally in the deeper layers. These features are summarized in Figure 2.1D.

The difference in projection paths implies thalamic neurons do not simply relay information in a point-to-point manner: the ‘choice’ of thalamic neuron used to take information to a given cortical area will influence the postsynaptic effect that it will produce. Some paths will lead to synchronous excitation of a wide range of cortex, others

more locally and topographically. Jones speculates that the role of these matrix arborisations is to synchronously activate distributed neural representations, thus providing a potential solution to the binding problem: “Their superficial terminations can synchronize specific and nonspecific elements of the thalamocortical network in coherent activity that underlies cognitive events.”(E. G. Jones, 2001) Not only does this view diversify the roles thalamus can play, it also elegantly draws parallels to the organization of thalamic nuclei serving what would appear to be quite different functions. This is an example of a theory which addresses the theme of ‘embodied’ signals discussed in the introduction to this thesis.

There are, however, a number of unresolved issues with this theory that limit its acceptance. First, it was developed based almost entirely upon data in primates from a limited number of nuclei, and its validity in other species remains unproven. There are notable species differences in thalamus, such as the presence of interneurons and calcium binding protein distribution(E. Jones, 2007). Further, even within neurons matching the descriptions of these classes, there is very substantial heterogeneity, with neurons differing greatly in their extent of layer 1 arborisation(Kuramoto et al., 2009; H. Nakamura, Hioki, Furuta, & Kaneko, 2015). For example, neurons from the calbindin+ region of the mediodorsal nuclei in rodents do not all show obvious layer 1 arborisations(Kuramoto et al., 2016). Additional issues are raised below shortly.

First/Higher-order theory divides thalamus based upon input patterns

A second major theory also divides thalamus into two nuclear groups, this time called first- and higher-order nuclei. It was proposed by Murray Sherman and Ray Guillery, who were motivated by another discovery about thalamus that did not fit into its classic description as a relay nucleus(Sherman & Guillery, 2002). This is the observation that thalamic nuclei can be strongly excited by neocortex itself, and indeed this can be the dominant input to thalamic neurons. Studies had shown that some, but not all, thalamic

nuclei receive branching input from the neocortical layer 5 neurons that project outside the telencephalon, the pyramidal type (PT) neurons(Sherman & Guillery, 2002; Sumser, Mease, Sakmann, & Groh, 2017). These inputs had been shown *in vitro* to elicit strong excitatory post synaptic potentials (EPSPs), and so were called 'driving' inputs (Sherman, 2016). Much thalamic activity therefore reflects telencephalic, not pre-telencephalic, activity. This is confirmed *in* vivo, with whisker-stimulation responses in rodent posterior medial (POm) nucleus of thalamus being lost after inactivation of cerebral cortex(Mease, Sumser, Sakmann, & Groh, 2016).This is obviously strongly inconsistent with thalamus as a simple relaying for information travelling to the telencephalon.

This led Sherman/Guillery to propose that thalamic nuclei were divided into two groups based upon where they received their 'driving' input from(Sherman & Guillery, 2002). Those that received their driving input from subcortical structures, including the retina and sensory systems, were considered 'first-order'. This includes the lateral geniculate dorsal nucleus (LGNd) and ventrobasal (VB) nuclei, receiving retinal and trigeminal input respectively. Those that received neocortical layer 5 input were considered 'higher-order'. This includes lateral posterior (LP) and posterior medial (POm) nuclei. We note that there is an assumption that there is only one driver input (this, to our knowledge, has not been proven, and would require extensive examination of the properties of all inputs to a given thalamic nucleus).

Sherman/Guillery argue that these 'higher-order' nuclei act to facilitate cortico-cortical transmission between different cortical regions, via thalamus. The theory thus divides thalamic nuclei depending upon the dominant excitatory influence upon its neurons, with thalamus being argued to play a critical role in communication between forebrain areas, as well as in carrying information to neocortex.

As with the core/matrix theory, there are also limitations to this theory. These center around the issue that there is very limited data on which thalamic nuclei receive layer 5 input and which do not, and the theory is thus largely restricted to the sensory thalamic nuclei. In a 2017 abstract at the society for neuroscience, Yu and Sherman say: 'transthalamic pathways have been found in visual, somatosensory and auditory pathways, but it is unknown if they exist beyond purely sensory systems', and go onto demonstrate it in the sensorimotor circuitry(Mo & Sherman, 2017). For the large majority of thalamic nuclei, then, it is still not possible to list them as being 'higher' or 'first' order, as we lack sufficient anatomical information.

2.2.5 - Relationship and common limitations of these theories

These different approaches to categorizing thalamic nuclei are not necessarily in conflict. They could each be capturing different, but valid distinctions between groups of thalamic nuclei and their properties: i.e., modality, core/matrix and higher/first-order schemes could all reflect meaningful approaches to understanding and making predictions about thalamic nuclei in different contexts. There is also significant overlap in the core/matrix and first/higher-order theories. The 'core' and 'first-order' nuclei identified thus far are overlapping: both the LGd and VB are parvalbumin positive, have a restricted, topographic arborisation pattern and lack prominent input from cortical layer 5.

However, things become less clear when one moves to the matrix/higher-order nuclei. Calbindin is widely expressed across thalamus, and marks both the intralaminar and midline nuclei (ILM), as well as more lateral nuclei. As discussed above, lesions to these intralaminar nuclei have strikingly different effects than lesions to more lateral thalamic groups. Furthermore, the few single neuron reconstructions available of these neurons show a prominent lack of the layer 1 arborisation that is typical of 'matrix' cells classically defined, raising doubts about the accuracy of Calbindin as a marker of a uniform cell type (Kuramoto et al 2016). Finally, the status of ILM nuclei within the

higher/lower-order theory is relatively ambiguous, and to the author's knowledge has not been specifically addressed. These theories thus leave the status of a significant fraction of thalamic nuclei unresolved.

There are further, more fundamental issues with taking these theories as taxonomies for thalamus. All are based upon a very limited number of properties of a given nucleus, and these properties are either unmapped or absent in many nuclei. Relatedly, the criteria used were not identified in a statistically unsupervised manner, making the choice of markers somewhat arbitrary. As a simple example, why is the calcium binding protein calretinin, a marker of a subset of thalamic nuclei, not used as a marker in the core/matrix theory? Neurons can vary in an enormous parameter space, and the pattern of this variation in high dimensional space may differ substantially compared to that in the few chosen features. Further, such binary classifications may miss continuous variation of properties that could be functionally relevant.

2.2.6 - Existing characterization of molecular variation across adult thalamus is limited

So neurons may vary in ways not captured by the existing categories. Indeed, not only *can* neurons vary in this parameter space, but there is significant evidence that thalamic neurons do indeed vary so. Synaptic properties, responses to neuromodulators and action potential waveforms have all been shown to vary substantially within and between thalamic nuclei *in vitro*(Browne, Kang, Akk, Chiang, Schulman, Huguenard, & Prince, 2001a; Hu, Senatorov, & Mooney, 1994; J. Li, Bickford, & Guido, 2003; Puil et al., 1994; Steriade, Curró Dossi, & Contreras, 1993a). Furthermore, firing properties such as receptive field tuning, electrical coupling with neocortex and frequency of burst firing events have also been shown to vary *in vivo* also (K. C. Nakamura et al., 2012; Ramcharan, Gnadt, & Sherman, 2005; Roth et al., 2015; W. Sun, Tan, Mensh, & Ji, 2016). However, these properties have not been mapped for the large majority of

thalamus, and neither form part of the major organizational theories of thalamus nor are known to show systematic variation (with the exception of burst firing being more common in higher order primate sensory thalamus)(Ramcharan et al., 2005). Further, in most cases we also lack a molecular or circuit level understanding of why these properties vary across thalamic neurons, and how they affect the ultimate contribution of the pathway to changes in neuronal dynamics(Clascá, Rubio-Garrido, & Jabaudon, 2012).

There is, however, a deeper assumption that is in part fueled by, and that in part fuels, this lack of knowledge. This assumption is that thalamic neurons are largely similar, consisting of a single molecular 'cell type' with a few slight variations(E. Jones, 2007). A recent review highlighted a large number of morphological and developmental differences between thalamic neurons, but lacked a section on molecular diversity, and of electrophysiological properties the paper stated that "Studies in different mammal species have shown that thalamocortical neurons display roughly similar general membrane properties"(Clascá et al., 2012). The review went on to say, however, that this view could be a result of a lack of looking: it is highly labour intensive to evaluate these properties across many nuclei, and even identifying the cells recorded from is challenging in thalamus due to the unclear nuclear boundaries. Therefore, we may be missing important aspects of thalamic variance.

This work uses comprehensive RNAseq profiling of thalamus to identify principles of thalamic organization

Genetic approaches offer a powerful set of tools to rapidly produce high-dimensional portraits of neurons(Nelson, Sugino, & Hempel, 2006). Advances in genetic techniques have allowed scientists to bypass the labour intensive stage of making many individual, low dimensional recordings and instead make many thousands of measurements from a sample in a matter of hours. Research into many areas of the brain have greatly benefited

from such large scale molecular screens, first using microarrays and more recently using RNAseq. These approaches yield quantitative read outs of gene expression involved in many different cellular functions, and have led to a proliferation of identified cell types(Saunders et al., 2018; Tasic et al., 2016; Zeisel et al., 2018). It has also revealed a great deal of information that has fueled further experiments, by highlighting unusual or distinctive properties of individual cell types that would not have been predicted(Mandelbaum et al., 2018). This has predominantly focused upon neocortex, but hippocampus, striatum, amygdala and other areas have also been profiled in detail.

Thalamus, however, has been notably absent from detailed investigation using these methods (though, thalamic cells are included in the recent whole-brain mapping efforts, further analysis is extremely limited(Saunders et al., 2018; Zeisel et al., 2018)). There are only a handful of papers directly evaluating the genetic composition of adult thalamus, with these existing studies focusing upon either a very small number of nuclei or not using quantitative approaches. Further, none use RNAseq, which offers the advantage of greater number of detected genes, as well as improved intensity range and sensitivity. There is thus a relative poverty of knowledge about the molecular composition of thalamus.

In this study we produced a near-comprehensive genetic atlas of thalamic pathways using RNAseq, providing a high-dimensional portrait of thalamus to address these questions(Phillips, Schulmann, Hara, Liu, Wang, Shields, Korff, Lemire, Dudman, Nelson, & Hantman, 2018a). We did this by retrogradely labelling neurons in each thalamic nucleus to obtain samples for dissecting, allowing us to relate molecular properties directly to pathway identity. We use this to reveal a strikingly simple organizational logic for thalamic pathways.

2.2 – Methods

Animal care

Experimental procedures were approved by the Institutional Animal Care and Use Committee at the Janelia Research Campus. C57/Bl6 background strain mice were housed on a 12-hour light/dark cycle, with *ad libitum* food and water. The majority of mice were 8-12 weeks old. Transgenic mouse lines, where used, were chosen based on visual analysis of the GENSAT and Brandeis mice line collections, or their presence in existence RNAseq databases.

Acquisition of samples

Cells were fluorescently labelled to enable manual dissection. This was done through retrograde labelling via either viral or tracer injection into the major projection field of the nucleus of interest. For viral injections, rAAV2-retro expressing cre-dependent GFP or TdTomato under the CAG promoter were injected, with volumes of 50-100nL at 3 depths (see Supplementary Data Table 2)(Tervo et al., 2016). Minimum survival time was 3 weeks post-injection. Viruses were prepared by Janelia Virus Services. Non-viral retrograde tracer labelling used the lipophilic tracer DiI (2.5mg/mL in DMSO, injecting volumes of 50-200nL per site, from Molecular probes) or Lumafluor red retrobeads (diluted 3x in PBS, 50-200nL per site). For DiI injections, survival time was 3-5 days. Anterograde labelling of inputs to thalamus was also used in a small number of cases (see Supplementary Data Tables 1 and 2).

We referred to the Paxinos and Franklin mouse brain atlas to guide our dissections(Paxinos & Franklin, 2012). For the majority of thalamic regions, retrograde tracers labeled populations corresponding to identified thalamic nuclei (Supplementary Data Tables 1 and 2 for targeting details - all surgical injection coordinate are in these supplementary data tables.). However, the caudal intralaminar nuclei (Parafascicular complex) were less clearly delineated. This likely reflects additional heterogeneity within

this complex beyond that shown in atlases(T. Kita, Shigematsu, & Kita, 2016a; Mandelbaum et al., 2018).

At no stage were experimenters blinded to sample identity.

Manual cell sorting and RNAseq

Sorted pooled-cell RNAseq

Fluorescent cells were collected and sequenced as previously described(Hempel, Sugino, & Nelson, 2007; Sugino et al., 2017). Briefly, animals were deeply anaesthetized with isoflurane and euthanized. Coronal slices (200-300 μm) were cut by vibrotome and placed for 1 hour at room temperature with pronases and neural activity blockers in artificial cerebrospinal fluid (ACSF)(see Sugino et al 2017 for full information). This thinner thickness relative to earlier work was chosen as it was important to be accurate in the depth dimension when making dissections (thicker sections leading to more averaging across depth), though it also led to more curling and generally more difficult tissue handling.

Relevant regions were then microdissected with surgical scissors, and the tissue dissociated. For validation of accuracy of dissection, see results below. The resulting cell suspensions were diluted with filtered ACSF and washed at least 3 times by transferring them to clean dishes using suction pippettes. This process produces negligible contamination with non-fluorescent tissue (see main results section below)(Hempel et al., 2007). After the final wash, samples were aspirated in a small volume (3 μl) and lysed in 47μl XB lysis buffer using the Picopure kit (KIT0204, ThermoFisher) in a 200μL PCR tube (Axygen), incubated for 5 min at 42 ºC on a thermal cycler and stored at -80 ºC.

An important point of difference from earlier studies in our group (eg, Sugino et al 2017) is that for thalamus, we aimed to have 40 cells in each pooled cell sample, as opposed to a lower bound of 30 for earlier work (though this was not always reached). This was based upon early data showing that reaching threshold for quality metrics required more cells for thalamic neurons, potentially due to greater vulnerability to the stress of the dissection procedure.

Library preparation and sequencing was performed by the Janelia Quantitative Genomics core. RNA was isolated from each sample using the PicoPure RNA isolation kit (Life technologies) with on-column RNAase-free DNase I treatment (Qiagen). 1µL ERCC RNA spike-in mix at 10^{-5} dilution (Life technologies) was added to each sample. Amplification was then performed using the Ovation RNA-seq v2 kit (NuGEN), yielding 4-8µg of cDNA. cDNA was sheared using Covaris E220 sonicator according to the manufacturer's recommendations, and the Ovation rapid DR multiplexing kit (NuGEN) was used to make libraries for sequencing. Libraries were quantified by qPCR using Kapa Library Quantification for Illumina (Kapa Biosystems KK4854). Equimolar pools were made and sequenced on a HiSeq 2500 (Illumina).

Sorted single-cell RNAseq

Retrogradely labeled cells were isolated as described above, and collected into 8-well strips containing 3µL Smart-seq2 lysis buffer, flash-frozen on dry ice, and stored at -80ºC until further use(Picelli et al., 2013).

Upon thawing, cells were re-digested with Proteinase K, and 1µL ERCC RNA spike-in mix at 10^{-7} dilution (Life technologies) and barcoded RT primers were added. cDNA synthesis was done using the Maxima H Minus RT kit (Thermo Fisher) and E5V6NEXT template switch oligo, followed by heat inactivation of reverse transcriptase. PCR amplification using the HiFi PCR kit (Kapa Biosystems) and SINGV6 primer was performed with a modified thermocycling protocol (98ºC for 3 min, 20 cycles of 98ºC for

20s, 64ºC for 15s, 72ºC for 4 min, final extension at 72ºC for 5 min). Samples were then pooled across strips, purified with Ampure XP beads (Beckman Coulter), washed twice with 70% ethanol and eluted in water. These pooled strips were then combined to create the plate-level cDNA pool for tagmentation, and concentration was determined using Qubit High-Sensitivity DNA kit (Thermo Fisher).

Tagmentation and library preparation using 600 pg cDNA from each plate of cells was then performed with a modified Nextera XT (Illumina) protocol, but using the P5NEXTPT5 primer and tagmentation time extended to 15 minutes(Soumillon, Cacchiarelli, Semrau, van Oudenaarden, & Mikkelsen, 2014). The libraries were then purified following the Nextera XT protocol (at 0.6x ratio) and quantified by qPCR using Kapa Library Quantification (Kapa Biosystems). 6-10 plates were run on a NextSeq 550 high-output flow cell. Read 1 contained the cell barcode and unique molecular identifier (UMI). Read 2 contained a cDNA fragment from the 3' end of the transcript.

Multi-FISH

C57Bl/6J mice (~8 weeks old) were anesthetized with isoflurane then fixed via transcardial perfusion with PBS followed by 4% paraformaldehyde in PBS, pH 7.4. Brains were post-fixed at 4 ºC overnight, washed 3 times with PBS, and cryoprotected in a sucrose series of 10%, 20% then 30% in PBS at 4 ºC. All solutions were prepared RNase-free. Brains were sectioned (14µm) on a Leica CM3050S cryostat, mounted onto Fisher SuperFrost Plus slides, and stored at -80 ºC.

Multi-FISH was performed using the RNAscope Multiplex Fluorescent Assay platform from ACDBio, following the manufacturer's protocol. The probes used were: *Calb2* (ref 313641), *Necab1* (ref 428541), and *Tnnt1* (ref 466911), *Pvalb* (ref 421931), *Scn1a* (ref 434181), *Kcnab3* (ref 479081*), Kcnc3* (ref 445411). Fluorescent dyes were DAPI, Alexa

Fluor 488, Atto 550 and Atto 647. Images were acquired using a Zeiss LSM 880 confocal microscope, with an air 20x (0.8NA) objective unless otherwise specified.
For display purposes, images were brightened and background subtracted, independently for each channel using Zen software (version 2.3 SP1, Zeiss). Gamma correction was applied for the retrograde beads image in Fig. 1B. All changes were applied uniformly across the entire image.

Electrophysiology

Acute brain slices were prepared from P20-25 mice. Animals were deeply anesthetized with ketamine/xylazine/acepromizine and transcardially perfused with ice-cold oxygenated cutting solution containing (in mM): 74 NaCl, 3 KCl, 1 NaH_2PO_4, 25 $NaHCO_3$, 6 $MgCl_2$, 0.5 $CaCl_2$, 5 Sodium Ascorbate, 75 Sucrose, 10 Glucose. 300µm coronal slices containing the thalamus were cut on a vibratome (Leica), and then recovered for 15 min at 33 °C and for 15 min at room temperature in oxygenated cutting solution followed by at least another 1 hour at room temperature in oxygenated ACSF containing (in mM): 126 NaCl, 3 KCl, 1 NaH_2PO_4, 25 $NaHCO_3$, 2 $MgCl_2$, 2 $CaCl_2$, 10 Glucose. During recordings, slices were perfused with oxygenated 34-35 °C ASCF. For current clamp recordings of intrinsic properties, ACSF included 35µM d,l-2-amino-5-phosphonovaleric acid (APV), 20µM 6,7-dinitroquinoxaline-2,3-dione (DNQX) to block glutamatergic synaptic transmission and 50µM picrotoxin to block GABAergic synaptic transmission. For voltage clamp recordings of synaptic events, ACSF included TTX (0.5µM), APV as above, and DNQX (for mIPSCs) or picrotoxin (for mEPSCs). Target neurons in CM, VA, and VL were identified based on their distance to the mammillothalamic tract and nuclear borders were confirmed with calbindin immunostaining *post hoc*. For the intrinsic properties and mEPSC recordings whole-cell recording pipettes (6 – 8 MΩ) were filled with internal solution containing (in mM): 100 K-gluconate, 20 KCl, 10 HEPES, 4 Mg-ATP, 0.3 Na-GTP, 10 Na-phosphocreatine, and 0.1% biocytin and for mIPSCs containing: 120 KCl, 10 HEPES, 2 MgSO4, 4 Mg-ATP,

0.3 Na-GTP, 10 Na-phosphocreatine, and 0.1% biocytin. Recordings were obtained with Multiclamp 700B amplifiers (Molecular Devices) digitized at 10 kHz using IGOR Pro (WaveMetrics). Resting membrane potentials were adjusted to -70 mV and steady state series resistance was compensated. Neurons with high series resistance (> 25 MΩ current clamp; >20 MΩ Voltage clamp) or membrane potentials that changed by > 3 mV were excluded. Custom IGOR scripts and python programs were used to analyze the data. For each current clamp recorded neuron, threshold, amplitude, afterhyperpolarization, and half-width at half-height of the 16th-19th action potentials in trials with 20 to 40 Hz firing rate were averaged.

Immunohistochemistry

After recordings, slices were fixed with 4% paraformaldehyde and 2.5% sucrose in 5x phosphate-buffered saline (PBS) at 4 °C for 2-10 days. After washing with PBS, slices were blocked in PBS with 0.3% Triton and 3% BSA at 4 °C overnight and then incubated in PBS with 0.3% Triton and 3% BSA and rabbit anti-calbindin D-28k (Swant, 1:1000) at 4 °C overnight. After washing, they were incubated in PBS with 0.3% Triton, 3% BSA and 5% goat serum with fluorescent protein conjugated goat anti-rabbit IgG (Invitrogen, 1:1000) and streptavidin (Invitrogen, 1:1000) at 4 °C overnight to label calbindin-expressing neurons and biocytin-filled neurons.

ANALYSIS-SPECIFIC METHODS

Pooled-cell RNAseq analysis

Data processing and quality control

After removing Illumina adapter sequences using cutadapt, reads were mapped to the mouse reference genome (mm10) using STAR (version 2.5.3a) with ‘ENCODE settings’ for RNAseq(Dobin et al., 2012). Mean mapping rate was 82.29% with a standard

deviation of 2.25%. Unique unambiguous exon-mapping reads were summarized at the gene level with featureCounts (version 1.5.2) using Gencode version M13.

Contamination with common astrocytic, oligodendrocytic, erythrocytic, and microglial transcripts was low, consistent with a lack of substantial contamination by non-fluorescent cells (Figure 2.3A). To ensure the specificity of our dissections and to control for potential batch effects, we collected several nuclei through multiple independent labelling approaches, and showed that these samples cluster in a highly similar manner (Figure 2.3, see main text).

Differential gene expression

Differential expression was assessed using the Bioconductor package *edgeR*(Robinson, McCarthy, & Smyth, 2009). Low counts were removed by requiring a Transcripts per million (TPM) > 10 in at least 3 samples. This yielded a list of approximately 17,000 expressed genes. Counts were then fitted to a negative binomial generalized linear model, where each factor level represents a different thalamic nucleus, and a Likelihood Ratio Test was used to assess differential expression between groups. P-values were adjusted for multiple tests using the Benjamini-Hochberg method. Genes with false-discovery rate < 0.05 were considered differentially expressed. For selecting the most differentially expressed genes between any thalamic nuclei, we used an ANOVA-like test (ANODEV test for generalized linear models, as described in edgeR User manual 3.2.6), testing for differences between any of the 22 nuclei, and used the 500 genes with the lowest P-value.

To avoid bias due to differences in sample number when comparing numbers of differentially expressed genes between different profiles, the groups were subsampled (with replacement) to the size of the smallest group. Bootstrapped log2 fold changes were obtained over 100 iterations. For visualization, clustering, and machine learning of gene expression data, we used variance-stabilized counts produced by the variance-

stabilizing transformation in the *DESeq2* R package)(Anders & Huber, 2010; Love, Huber, & Anders, 2014).

Unsupervised clustering and functional enrichments

Hierarchical clustering was performed using 1 - Spearman's correlation as a distance metric and complete linkage for agglomeration. Groups were defined by splitting the tree at the level of 5 branches. We termed these profiles, not clusters, as we do not mean to imply discreteness between the classifications. PCA was done using the singular value decomposition based *prcomp* function in R. For functional enrichment of differentially expressed genes, we used the PANTHER Protein Class Ontology (http://data.pantherdb.org/PANTHER13/ontology/Protein_Class_13.0), which is a consolidated version of molecular function gene ontology. Over-representation in the top 100 genes with the highest PC1 loadings was assessed via hypergeometric test.

For defining voltage-gated ion channels and neurotransmitter receptors, we downloaded the IUPHAR/BPS database (http://www.guidetopharmacology.org/DATA/targets_and_families.csv). Voltage-gated ion channels were the genes defined in the database, while for neurotransmitter receptors we included ionotropic and metabotropic receptors for glutamate, GABA, glycine, acetylcholine, monoamines, neuropeptides, adenosine, and cannabinoids.

Single-cell RNAseq analysis

Data processing and quality control

Single-cell RNAseq data was trimmed for adapters using cutadapt and aligned to the mouse genome (mm10) using STAR (version 2.5.3a)(Dobin et al., 2012). To demultiplex cells, collapse UMIs and produce gene-wise counts for each cell, we used a modified version of the *Drop-seq_tools-1.13* pipeline (http://mccarrolllab.com/download/1276/). Briefly, read 1 was tagged based on the cell barcode and UMI, and this information was

added to read 2 by merging back the reads after mapping, followed by gene-wise tagging of reads that map onto exons and summarization of digital counts.

Single cells were required to have more than 20,000 UMIs and more than 2,500 genes detected per cell, which yielded a total of 1,972 cells. Of these, 20 cells were found to be significantly contaminated with oligodendroglial cell transcripts, leaving 1,952 cells for all downstream analyses. Quantification of ERCC spike-in control RNA indicated high accuracy and sensitivity of our single-cell profiling. Genes were considered expressed if their expression was detected in more than 10 cells.

Our single-cell sequencing was not comprehensive, and with improved sequencing approaches further genetic subdivisions may be identified. Single-cell and pooled-cell dissections were not precisely matched, for example motor-projecting midline nuclei were not dissected for single-cell RNAseq.

Single-cell clustering and marker genes

Single-cell clusters were defined using the *Seurat* R package (version 2.1)(Butler & Satija, 2017; Satija, Farrell, Gennert, Schier, & Regev, 2015). Data were log transformed and scaled. For identifying variable genes, genes were divided into 20 bins based on average expression, and genes that were more than 1 standard deviation away from average dispersion within a bin were used for downstream analysis. Single-cell clustering was performed separately for each projection system (Fig. 5A-C), and for all cells combined together using shared nearest neighbor (SNN) clustering and limiting the analysis to the top 10 principal components for distance calculation. Clusters were defined by the Louvain algorithm, and clustering resolution set to 0.6. Clusters of cells were visualized using t-distributed stochastic neighbor embedding (tSNE) using the top 10 principal components as input and perplexity set to 30. Marker genes for each cluster were required to be expressed in at least 80% of the cells in the cluster, to have a P-value $<10^{-5}$ (Likelihood Ratio Test), a log2 fold change > 0.5. Projection of single-cell data

onto pooled-cell principal components was obtained by multiplying (dot product) log-transformed and scaled single-cell data by the pooled-cell principal component loadings.

2.3 - Results

To understand the organization of thalamic pathways, we combined anatomical and genetic approaches to produce a near-comprehensive, projection-specific transcriptomic atlas of murine thalamus. This not only allowed us to examine the relationship between thalamic nuclei in higher dimensional space, but also examine the properties underlying these relationships and provide information to guide further mechanistic experiments. This large majority of the dataset consists of pooled-cell sequencing (see below), with a single cell component described later. For now, all data refers to sequencing groups of cells, as in (Hempel et al., 2007; Sugino et al., 2017).

Our first step was to achieve labelling of individual nuclei so that they could be individually microdissected, and this formed the large bulk of the experimental work. For the large majority of nuclei, retrograde tracers were injected into the primary projection field of that nucleus, leading to labelled cells in thalamus. After testing multiple tracers, the lipophilic dye DiI was chosen as the default tracer due to the reliability of labelling observed. After reviewing existing literature and the allen brain atlas, injections were made into 2-3 animals and the resulting labelling observed under the same dissecting microscope used during cell collection. Upon identifying a strategy that led to clear and consistent labelling, sample acquisition began. Nuclei were microdissected using surgical scissors and cells pooled together, and samples prepared for sequencing (see methods) (Figure 2.2A).

In two cases, additional anterograde labelling was used. This was used when the boundaries between nuclei were not clear based solely upon retrograde labelling, making microdissection challenging. For the visual thalamus, anterograde labelling from the

retina was used (Figure 2.2B), whilst for the motor thalamus, the outputs of the basal ganglia and cerebellum were labelled (ie, both outputs labelled in the same animals, Figure 2.2C). These label the previously identified subdvisions of these nuclei(E. Jones, 2007; Kuramoto et al., 2009; 2010). This also helped to minimize contamination of motor thalamic samples by anterior medial nucleus, which is spatially close to ventral anterior/lateral (VA/VL), labels from motor cortex, but does not receive prominent input from cerebellum or basal ganglia (though we cannot exclude some contamination).

The result of this was a near-comprehensive pooled-cell RNAseq atlas of thalamus (8 projection targets, 22 nuclei, 120 samples, full details of strategy for collecting each nuclei are available Tables S1 and S2)(Figure 2.2D).

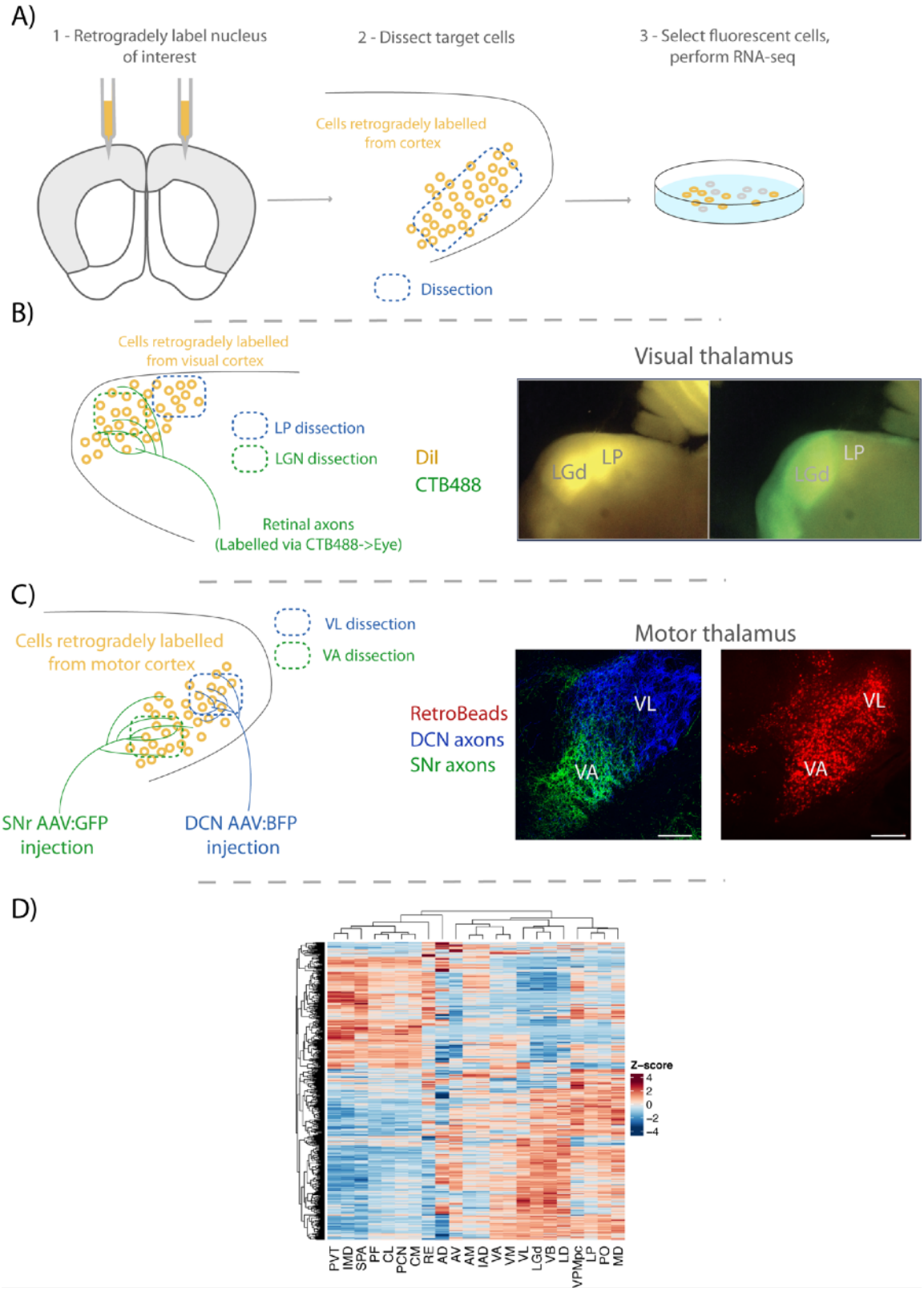

A)
1 - Retrogradely label nucleus of interest
2 - Dissect target cells
3 - Select fluorescent cells, perform RNA-seq
Cells retrogradely labelled from cortex
Dissection
B)
Cells retrogradely labelled from visual cortex
LP dissection
LGN dissection
Retinal axons (Labelled via CTB488->Eye)
DiI
CTB488
Visual thalamus
LGd
LP
LP
LGd
C)
Cells retrogradely labelled from motor cortex
VL dissection
VA dissection
SNr AAV:GFP injection
DCN AAV:BFP injection
RetroBeads
DCN axons
SNr axons
Motor thalamus
VL
VA
VL
VA
D)
Z-score
4
2
0
−2
−4
PVT
IMD
SPA
PF
CL
PCN
CM
RE
AD
AV
AM
IAD
VA
VM
VL
LGd
VB
LD
VPMpc
LP
PO
MD

Figure 2.2 – Experimental approach for obtaining near-comprehensive transcriptomic atlas of thalamus

a) Illustration of sample collection approach. Retrograde tracers were injected to the primary projection field of a nucleus, nuclei were manually dissected, and fluorescent cells picked.
b) Example combined anterograde/retrograde labelling approach for separating LP and LGN. CTB488 was injected to the retina, labelling retinal axons that terminate in the LGd, but do not enter LP. This was used to separate LGd and LP under the dissecting microscope (right panels). No scale bar available as image was taken during dissection.
c) As in b), but for VA/VL. Tracers labelling the output of cerebellum (deep cerebellar nuclei, DCN) and basal ganglia (substantia nigra reticulata, SNr) were used to mark the VL and VA subdivisions of motor thalamus, respectively. Scale bar = 200um.
d) Illustration of resulting dataset. Heatmap of the top 500 differentially expressed genes. Rows and columns are ordered by hierarchical clustering with Euclidean distance metric. Colors represent gene-wise Z-scores.

The accuracy and utility of our dataset relies on it lacking substantial contamination by non-neuronal tissue, and on the samples representing the underlying target tissue of interest. We assessed a range of markers of non-neuronal tissue, and found almost all samples to have 0 detected transcripts for these contamination markers(Figure 2.3A). Secondly, we acquired a number of samples in a different manner than in the main dataset, for example using transgenic lines to label the anterior (AV, AD, see supplementary data table 4 for details of the lines used), and with one exception, they clustered close to the replicate (Figure 2.3B). The similar localization of different samples from the same nucleus in the dimensionality reduction figures below is further evidence of consistent sampling (see below). We conclude that our sampling is reliable and robust to collection method.

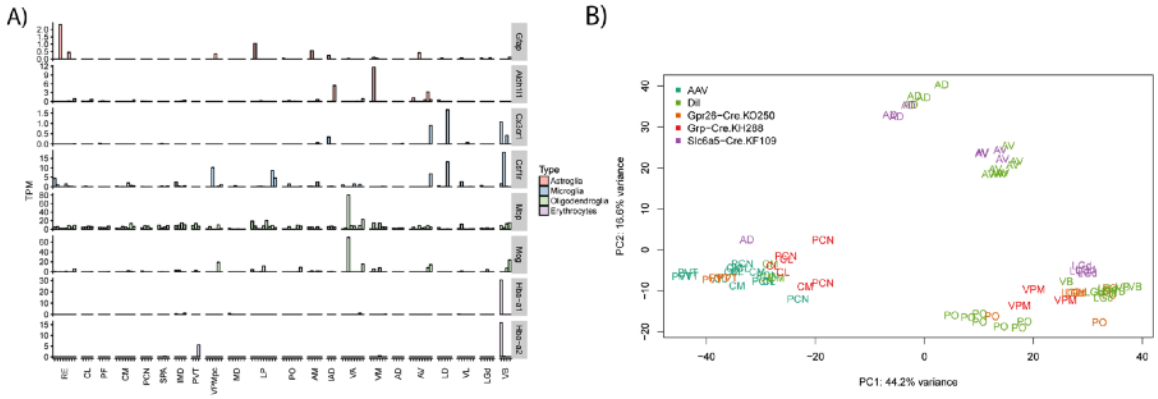


Figure 2.3 - Pooled-cell RNAseq quality controls

a) Markers of non-neuronal sample contamination are low across our dataset. Expression (transcripts per million, TPM) in pooled-cell samples shown for 8 genes marking astrocytes, microglia, oligodendrocytes and erythrocytes. Only a small number of samples showed expression of contamination markers.

b) Samples of the same nucleus obtained via different labelling methods cluster similarly. Principal components analysis of those samples, for which multiple collection methods were used (i.e. GENSAT lines in addition to retrograde labeling) using the top 500 genes with highest variance. Samples are colored by collection approach or transgenic line used.

Our first goal was to understand the relationship between thalamic nuclei. The dataset offers the opportunity to do so in a relatively unsupervised manner, by evaluating the genetic distance between different nuclei. We therefore first used hierarchical clustering to explore the relationship between the thalamic nuclei (on the 500 most differentially expressed genes via an ANOVA-like test). This allowed us to identify different groups of thalamic nuclei based upon transcriptomic similarity (figure 2.4).

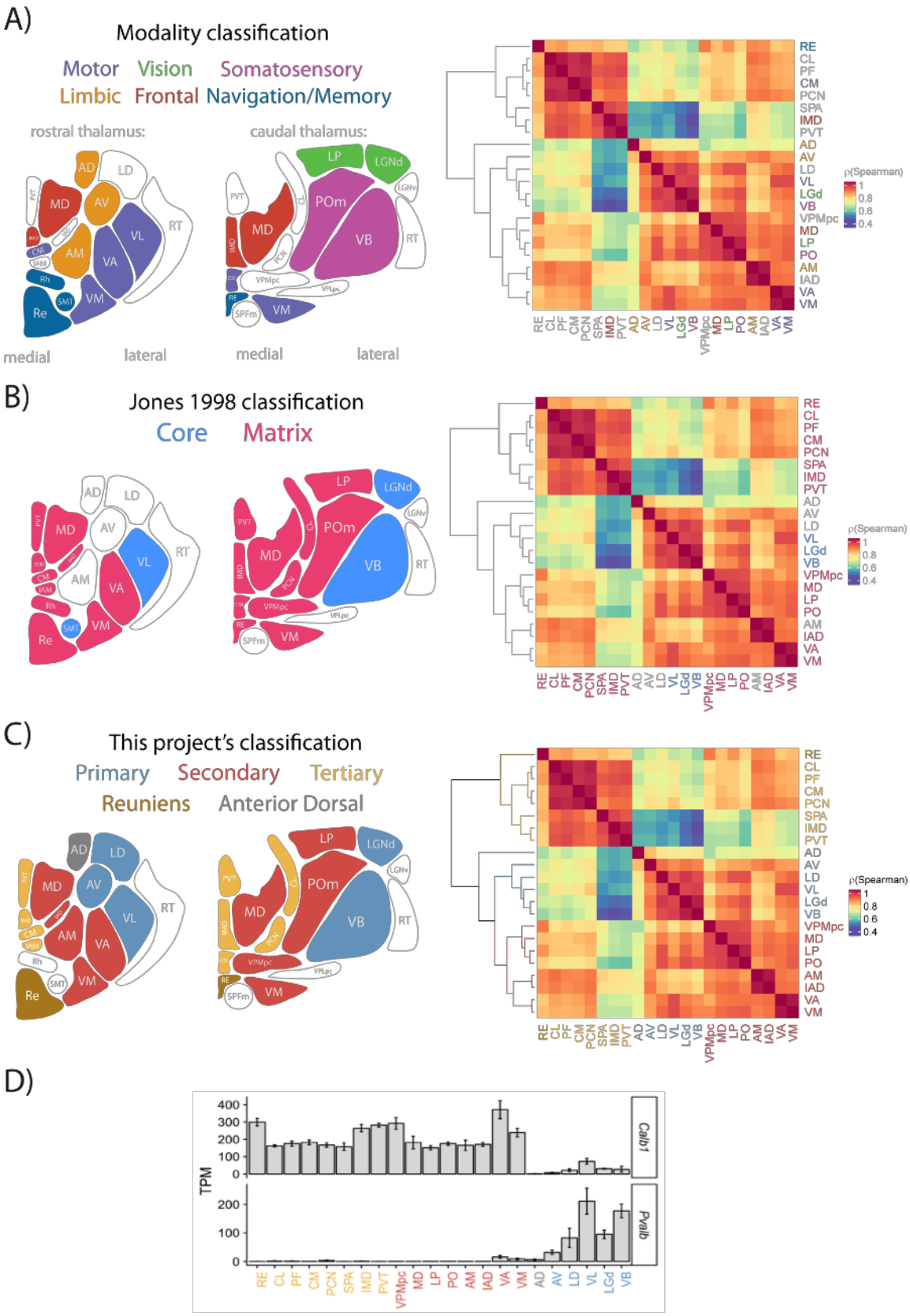

A)
Modality classification
Motor Vision Somatosensory
Limbic Frontal Navigation/Memory
rostral thalamus:
caudal thalamus:
medial
lateral
medial
lateral
B)
Jones 1998 classification
Core Matrix
C)
This project's classification
Primary Secondary Tertiary
Reuniens Anterior Dorsal
ρ(Spearman)
1
0.8
0.6
0.4
RE CL PF CM PCN SPA IMD PVT AD AV LD VL LGd VB VPMpc MD LP PO AM IAD VA VM
D)
TPM
Calb1
Pvalb

Figure 2.4 - Clustering of thalamic gene expression profiles identifies topographic genoarchitecture

A. Modality-based approach does not explain results of hierarchical clustering. Left) Nuclei coloured by modality of information conveyed. Right) Hierarchical clustering of thalamic nuclei using Spearman's correlation of top 500 most differentially expressed genes across all 22 nuclei. Nuclei coloured as in left panel.
B. Core/matrix dichotomy partly explains the groupings found via hierarchical clustering. Left) Nuclei coloured by classification within core/matrix theory, based upon calcium binding protein expression. Right) As in a), but coloured by core/matrix identity.
C. Nuclei belonging to each of the three major classes. Reuniens and AD form their own clusters. Major clusters defined as the top 5 branches of cluster dendrogram.
D. Comparison of core/matrix theory (in b) with the clusters from (in c). Expression of *Calb1* and *Pvalb* (in transcripts per million, TPM, mean ± SEM) for each nucleus, with nuclei on the x-axis colored by their profile from Fig. 1. The core/ matrix organizational theory proposes that thalamus is divided into two discrete groups, expressing calbindin or parvalbumin. However, the first major branch splits the secondary and tertiary groups, both of which are marked by *Calb1* and would thus both be considered 'matrix' nuclei in this theory. Thus existing markers for thalamic nuclei subgroups do not adequately reflect thalamic organizational structure.

What accounts for the resulting dendrograms? Modality, such as vision, motor or somatosensory, did not explain the resulting groupings. To illustrate this, figure 2.4A shows coronal sections of the thalamus on the left, with nuclei coloured by modality, and on the right the dendrogram with samples coloured by the modality of information they process. Individual nuclei of the motor, limbic, visual, frontal and somatosensory

thalamic nuclear groups all come out in separate parts of the dendrogram. If modality is not the explanation, perhaps the cross-modal core and matrix theory of Jones could explain the resulting clustering(E. G. Jones & Hendry, 1989)? Figure 2.4B shows nuclei coloured by whether they predominantly express parvalbumin (core) or calbindin (matrix). Here we see that there is indeed some grouping, but that the major splits of the dendrogram are not explained by calcium buffer expression (see also Figure 2.4D). Core/ matrix theory thus appears to be a smaller part of a larger cross-modal organizational scheme.

To examine what this more fundamental organizational scheme was, we cut the dendrogram to form 5 major groups, which we refer to as profiles (so as not to use the term clusters, with its implication of discreteness. See below). This led to three major groups containing more than a single nucleus. These three main groups were coloured blue, red and yellow, and were termed primary, secondary and tertiary respectively, for reasons that will be apparent shortly. Rather than being explained by modality or by expression of individual calcium binding proteins, our three major profiles were best distinguished by anatomy, with nuclei of each profile occupying a characteristic position along the mediolateral axis of the thalamus (Figure 2.4C). We thus find that the architecture of thalamus is dominated by genetic differences that are organized topographically.

Two of the five profiles contained just a single nucleus. These were the anterior dorsal nucleus (AD) and nucleus reuniens (RE). We did not explore in detail the reason for the two minor clusters existing. These minor profiles contained just a single nucleus each, the anterior dorsal and reuniens nuclei. It is noted though that during examination of the allen brain atlas, we found that many markers of brain-stem nuclei were also labelled by markers of the anterior dorsal nucleus, and the hypothalamus was commonly labelled by markers of the nucleus reuniens. As we lacked RNAseq data from these other areas, we did not analyse this intriguing finding further.

We note that the number of profiles chosen is somewhat arbitrary. As we shall we see, these groups are used further only as description for narrative simplicity: the core conclusions of the dataset come from fully unsupervised dimensionality reduction methods without the need to define cluster numbers.

To further understand the pattern of gene expression differences between the thalamic projection profiles, we performed principal component analysis (PCA) on our data. This identifies new axis of variance, ranked by total explained variance, in a deterministic and unsupervised manner. We found that the principal component 1 (PC1, 38% explained variance) separated nuclei into the same major profiles identified via hierarchical clustering (Figure 2.5A), closely matching the hierarchical clustering. Again, position on this first dimension strongly correlated with mediolateral position, demonstrating topogenetic architecture in thalamus (Figure 2.5B). We can now see why we named the groups primary, secondary and tertiary: there is a progressive genetic difference from one group to another. The progressive difference from primary to tertiary nuclei was evident in the number of genes differentially expressed between the groups, with the primary and tertiary nuclei being most distinct, and the other two comparisons being less so (Figure 2.5C). Thus, nuclei of the major profiles sit continuously along a single axis of genetic variance which is aligned with the mediolateral spatial axis of thalamus.

There are two further points to make here. The first is that the major axis of variance was prominently enriched in genes encoding neurotransmitter receptors, ion channels, and signaling molecules (Figure 2.5D and 2.5E). The second is that this pattern produced by PCA is also identified via other dimensionality reduction techniques: we show the robustness of this result by using a different dimensionality reduction technique, MDS, using a different distance metric (figure 2.5F, see legend). There is thus a robust, single axis of variance separating major thalamic profiles, with a diverse but functionally enriched gene set underlying this.

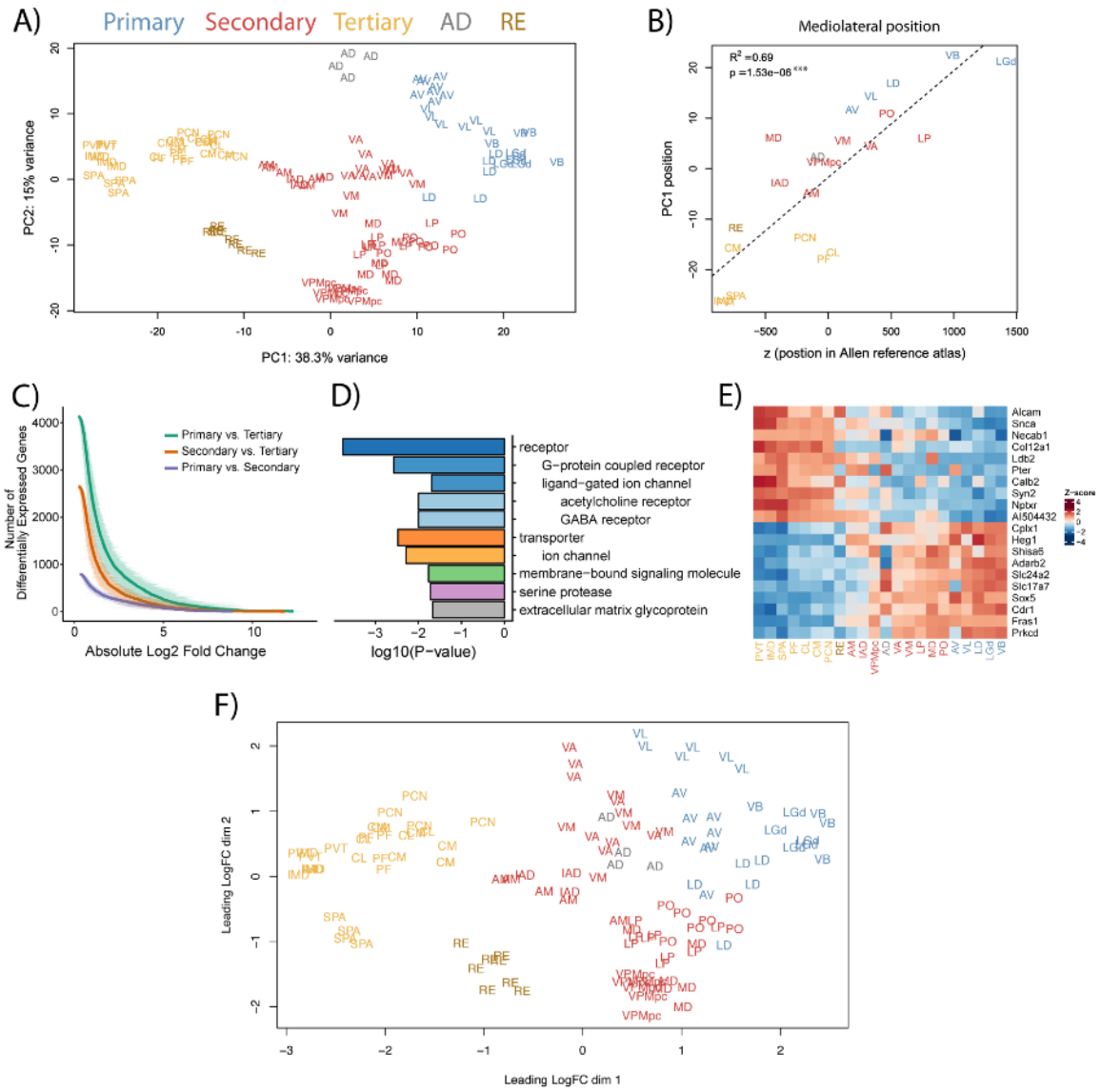


Figure 2.5 - A genetically diverse and topographic axis of variance separates major thalamic profiles

A. PCA showing separation of functional nuclear profiles in the first two principal components. The underlying gene set and color scheme are the same as in 1C.

B. PC1 position is highly correlated with mediolateral position of the nuclei. Mediolateral positions are z voxel coordinates of nuclei centers in the Common Coordinate Framework (CCF) in the Allen Mouse Brain Atlas. 1 voxel = 10 μm.

C. Primary nuclei are farthest from tertiary nuclei, with secondary nuclei being intermediate. Plot shows the number of differentially expressed genes at each log fold change level (shown as mean ± standard deviation) for the three comparisons.

D. Genes relevant to neurotransmission are overrepresented amongst the top 100 genes with the highest absolute PC1 loadings in our dataset. The ten most highly overrepresented PANTHER protein class terms are shown. P-values based on hypergeometric test. Indentation indicates gene subfamily.

E. Heatmap of genes with strongest positive and negative loadings on PC1. Nuclei are ordered by their mean position on PC1 of Fig. 2A. Colors represent gene-wise Z-scores.

F. Multidimensional scaling using an alternative distance metric also identifies the same leading axis of variance. Distance was defined as the quadratic mean of the log2 fold changes of the top 500 differentially expressed genes between any two samples (meaning that the gene set used for the distance comparison varies between each sample comparison made).

These results show that there is a cross-modal architecture in thalamus, and that it has a surprisingly rich genetic basis. Given the prominent differences in receptor and ion channel expression between thalamic profiles, we asked whether these profiles correspond to functionally different classes of neurons. We first did this based upon further examination of the RNAseq dataset. We performed PCA on the expression profiles of voltage-gated ion channel or neurotransmitter/modulator receptor encoding genes (Fig. 2.6A, left and right respectively). Analysis with these limited gene sets reproduced the separation of profiles in PC1 (Fig. 2.6A), confirming that ion channel and

receptor profiles are organized along the same axis identified in Fig. 2. This is suggestive of a progressive shift in the electrophysiological properties along the first axis of genetic variance.

To assess this, we first examined the literature to see if channels and receptors at one end of the axis were consistently linked to particular firing properties in neurons. We found that genes linked to high firing rates via faster channel kinetics, such as Kv3 channels (*Kcnc1*, *Kcnc3*), the *Scn8a* channel, and the *Kcnab3* subunit, tended to be progressively elevated toward primary profile nuclei(Figure 2.6B) (Freeman, Desmazières, Fricker, Lubetzki, & Sol-Foulon, 2015; Leicher, Bähring, Isbrandt, & Pongs, 1998; Rudy & McBain, 2001). Though inferring *in* vitro spiking properties solely from ion channel expression is fraught with difficulties, this raised the possibility that action potential width may progressively narrow from tertiary to primary nuclei.

We assessed the action potential width experimentally (Figure 2.6C). Whole-cell recordings were made from the motor-related nuclei CM, VA and VL (representing the three main nuclear profiles, Figure 2.6C top row) and confirmed this prediction. Neurons recorded within VL have the narrowest action potential width and those in CM the widest. In addition, other electrophysiological properties showed a systematic gradient ranging from VL through VA to CM (Figure 2.6C, bottom row). Prior work has shown substantial differences in electrophysiological properties between different thalamic nuclei, but to date this has not been incorporated into thalamic organizational schemes (Fogerson & Huguenard, 2016; K. C. Nakamura et al., 2012; Puil et al., 1994). Our electrophysiological recordings thus show a systematic variation of neuronal firing properties from primary to tertiary profiles.

Though we did not assess it experimentally, the strong enrichment of neuromodulatory genes, especially among nuclei of the secondary and tertiary type, suggests further differential modulation of inputs across these profiles (Figure 2.6B, right). These results converge to show that the leading axis of genetic variance in thalamus is marked by a

rich complement of functionally relevant genes, with implications for neural function, and are indicative of a spectrum of transformations across thalamic profiles and the underlying axis of genetic variance.

Notably, transcription factors were relatively less closely matched to the single axis identified above, and the leading axis of variance in this data explained less variance than did the ion channel/receptors axes (Figure 2.6D). This is somewhat surprising, as transcription factors are typically invoked as the driving force behind gene expression differences. This is a point that we explore in the thesis discussion.

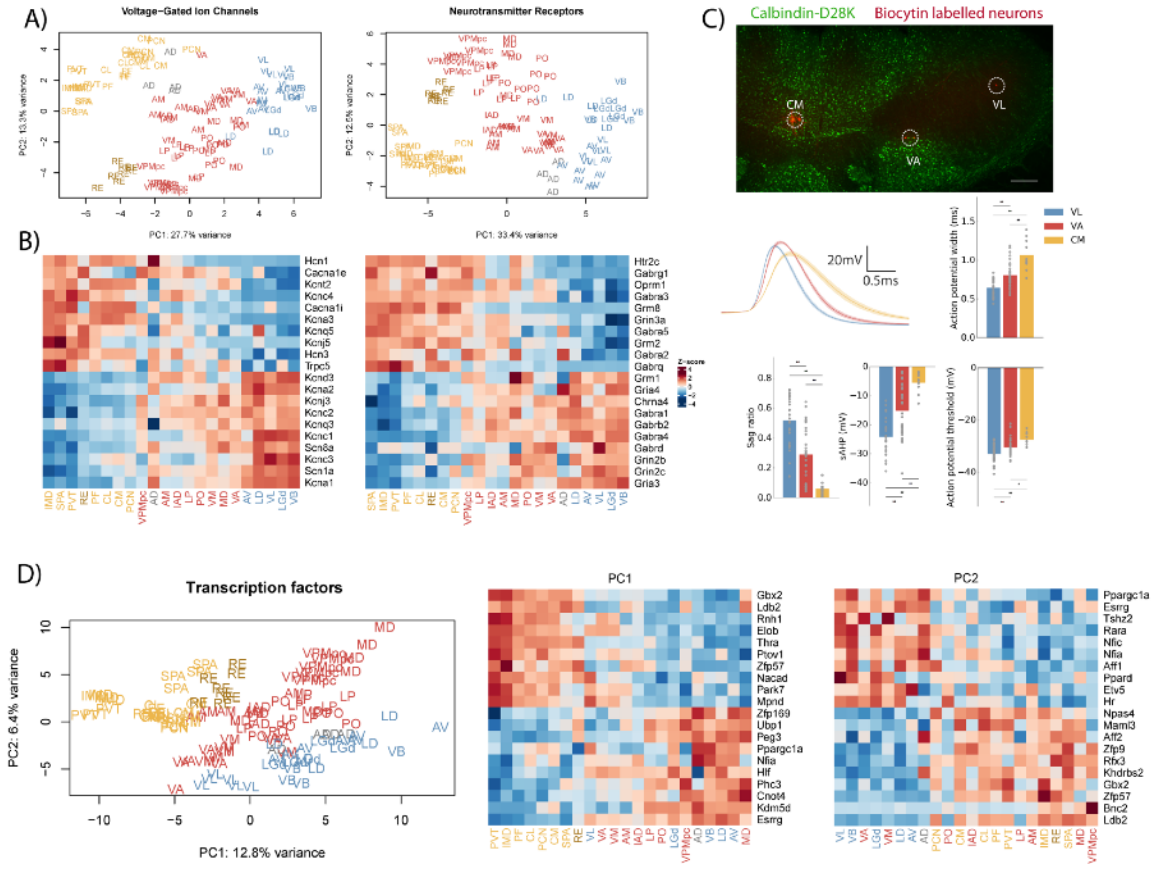


Figure 2.6 – Functionally relevant genes and electrophysiological properties vary systematically across nuclear profiles

A. PCA including only genes encoding voltage-gated ion channels and neurotransmitter/neuromodulator receptors. Colors as in Fig. 1C.
B. Heatmap for genes with the highest gene loadings in PC1 from Fig. 3A. Voltage-gated ion-channels on the left and neurotransmitter receptors on the right. Colors represent gene-wise Z-scores.
C. Systematic variation of electrophysiological properties across profiles. Whole-cell patch-clamp recordings from VL (primary thalamus), VA (secondary thalamus) and CM (tertiary thalamus). Top row: neurons were labelled with biocytin (red) and localized to individual nuclei with the aid of Calbindin-D28K immunolabelling (green). Scale bar = 100µm. Middle row, left shows average action potential shape for VL, VA and CM neurons (mean±SEM). Remaining panels show comparisons for four physiological measurements across these nuclei (One-way ANOVA with *post-hoc* Tukey HSD test, all comparisons $P<0.05$). Sample contained 29 VL neurons, 34 VA neurons and 10 CM neurons.
D. As with A and B, but for transcription factor-encoding genes.

Thus far we analyzed thalamic nuclei by pooling projection neurons from specific anatomical positions. By averaging across many cells, the resulting transcriptional profiles from this approach could be composed of homogenous populations, an average of multiple discrete subtypes, or of cells with graded differences. To distinguish between these possibilities, we profiled the transcriptomes of individual neurons from motor, somatosensory, visual, auditory, and prefrontal projection classes using single-cell RNAseq. Cells were retrogradely labelled as for pooled-cell collection (see supplementary table 3). However, we did not collect midline-localised intralaminar nuclei cells projecting to vision, somatosensory, auditory or motor cortices. This was because they were sparse and diffuse, meaning a large amount of white matter needed to be sifted through the filter the cells, which would have been restrictive for fast processing of healthy cells.

We quality controlled our single-cell data in several ways. Firstly, we filtered individual cells by total number of detected transcripts per cell (Figure 2.7A, left) and by total number of detected genes (Figure 2.7A, right). Secondly, we evaluated the principal components of our data for components with strong weightings for contamination-related genes, finding that PC3 and PC7 have genes marking oligodendrocytes and epithelial cells (Figure 2.7B) respectively. Cells were removed based upon their position in these axes (see figure 2.7B legend). Thirdly and particularly notably, we find that the ranking of genes on the first principal component were highly similar between single-cell and pooled-cell data (this was done using the limited gene set as in figure 3.3A).

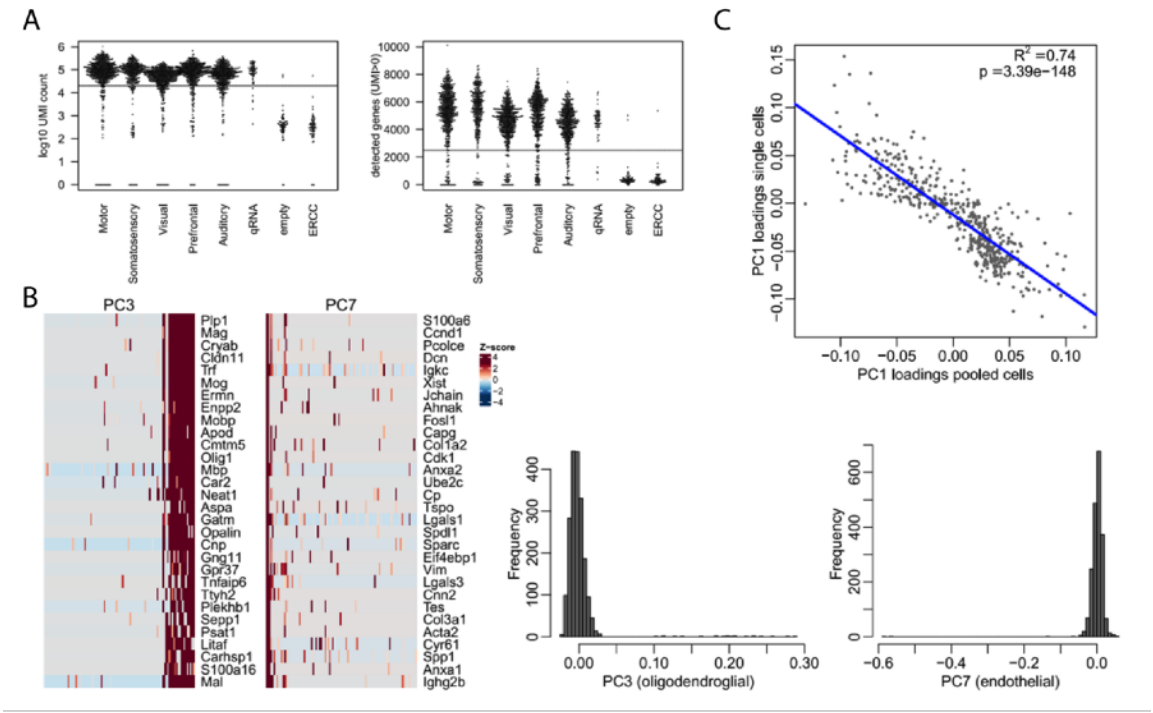


Figure 2.7 – Quality control for single-cell RNAseq data

A) Unique molecular identifier count (left) and gene detection rate (right) for all collected single cells. Cutoffs for downstream use are indicated by dashed lines.

B) PCA on the single-cell RNAseq data revealed that principal components 3 and 7 represented non-neuronal contamination from oligodendrocytes and endothelial cells. Cells with PC3 position >0.05 and PC7 position <0.1 were removed.

C) PC1 loadings for the most differentially expressed genes between nuclei (gene set as in Fig.1) are highly similar in pooled-cell and single-cell RNAseq data.

We used this resulting dataset to evaluate heterogeneity in thalamic cells projecting to the same cortical area. Analysis of this single-cell RNAseq dataset resulted in multiple clusters for each projection, and cluster markers included many genes that also distinguished nuclei from pooled data (Figure 2.8A,B). Single-cell clusters also separated along the first principal component derived from our pooled-cell RNAseq dataset (Figure

2.8C). Further, these groups did not appear clearly discrete, but rather These results indicate that the genetic axis identified in the bulk data also applies to the single cell data.

Our single cell RNAseq could give the appearance of a spectrum due to noise. We therefore next sought to confirm our results with a technique with higher spatial resolution and greater sensitivity. Using multiplex-fluorescent in situ hybridization (multiFISH), we found that markers for the single-cell clusters were spatially separated at the single-cell level, but with intermediate cells especially prominent near the anatomical boundaries between thalamic nuclei (Fig. 2.8D, Figure 2.9). This co-expression of profile marker genes was therefore confirmed by multi-FISH (Figure 2.10). Together, these results are consistent consistent with spatially organized heterogeneity or a gradient-like organization rather than intermingled, discrete populations(Cembrowski, Bachman, Wang, Sugino, Shields, & Spruston, 2016a; Gokce, Stanley, Treutlein, Neff, Camp, Malenka, Rothwell, Fuccillo, Südhof, & Quake, 2016a). Though we examined only 3 genes *in situ* due to the prohibitive time and cost of multi-FISH, emerging technologies should enable examination these spatial gradients across hundreds of genes(Xiao Wang et al., 2018).

Given the strong relationship between PC1 and anatomical nuclei position (Figure 2.4B), the presence of single-cell clusters occupying similar PC1 positions (e.g. clusters 1 and 2 of the motor projections neurons, or clusters 1 and 2 of the somatosensory nuclei, suggested that distinct neuron types could also coexist within anatomical boundaries of nuclei. We examined this possibility by performing multi-color fluorescent *in situ* hybridization (multi-FISH) for genes which distinguished amongst the major profiles (*e.g. Pvalb* and *Tnnt1*). Taking motor thalamus as an example, *Pvalb* and *Tnnt1* expressing cells were found within the anatomical boundaries of a single thalamic nucleus (VL; Fig. 2.8E). Some individual VL cells expressed both *Pvalb* and *Tnnt1; Pvalb*-selective, intermediate, and *Tnnt1*-selective cells were distributed along the

mediolateral axis (Fig. 2.8E). Therefore, spatially organized transcriptomic differences can exist even within individual thalamic nuclei.

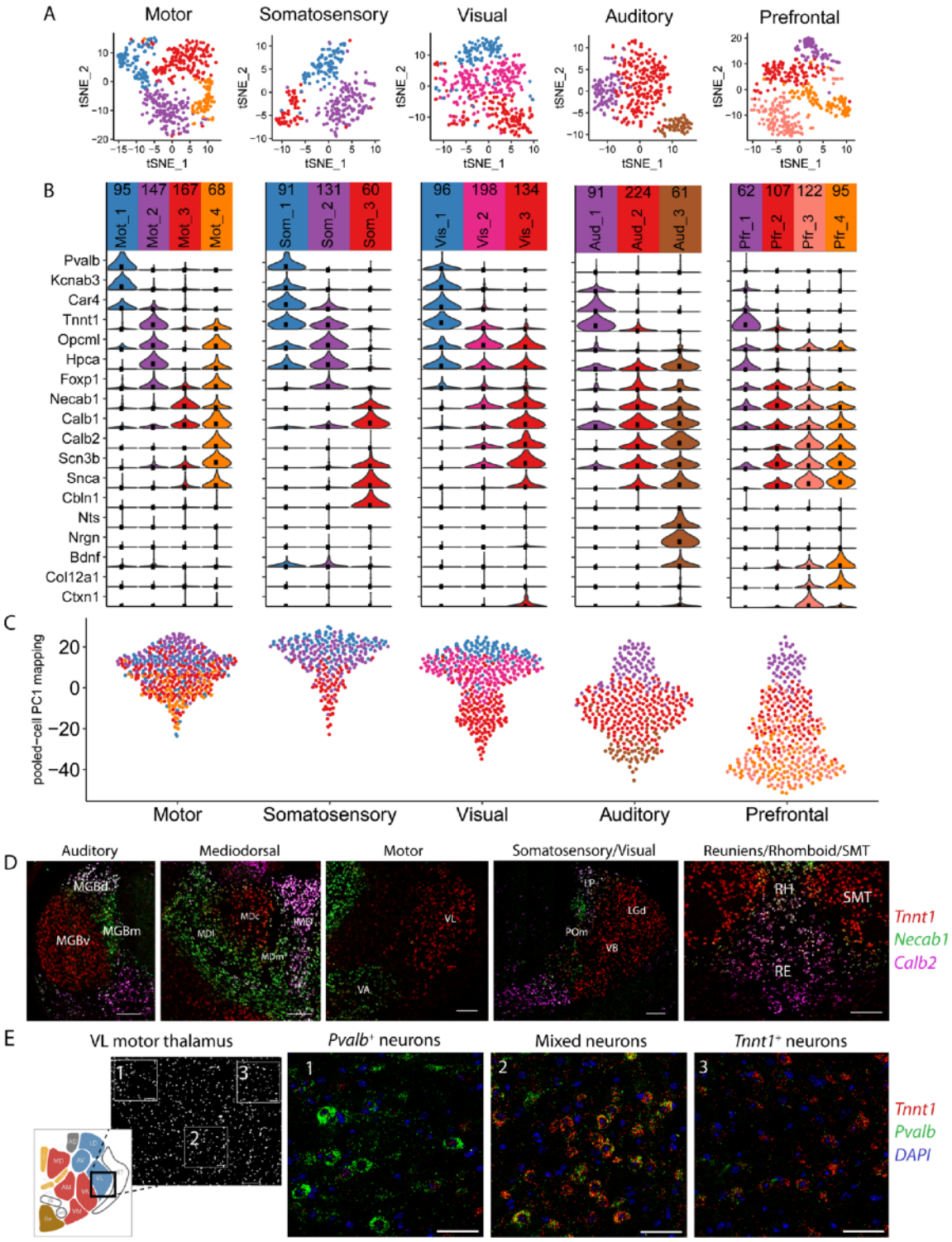


Figure 2.8 – A topographic spectrum of thalamic cell identities within and between thalamic nuclei

A. Overview of single-cell clusters within each projection system visualized via tSNE. Cells colored by cluster identity.
B. Violin plots for cluster marker genes (inclusion criteria were Likelihood Ratio Test P-value < 10^{-5} and log2 fold change > 0.5 for each cluster).
C. Projection of single-cell data onto pooled-cell PC1 from Fig. 2A. Each dot is a single cell colored as in A.
D. Topographic distribution of marker genes within 6 major thalamic modalities. Multi-FISH with probes for *Calb2* (pink), *Tnnt1* (red) and *Necab1* (green). See Fig. S7 and S8 for expanded views and quantification. Scale bars = 200 μm.
E. Multi-FISH within VL thalamus comparing genes marking clusters in single-cell data. Left panel shows field of view (coronal section, scale bar = 100 μm). Right three panels (scale bars = 50μm) show expansion of the three boxed areas, moving left to right. Middle box shows intermediate cells expressing both single cell cluster markers. (red = *Tnnt1,* green = *Pvalb*, blue = DAPI).

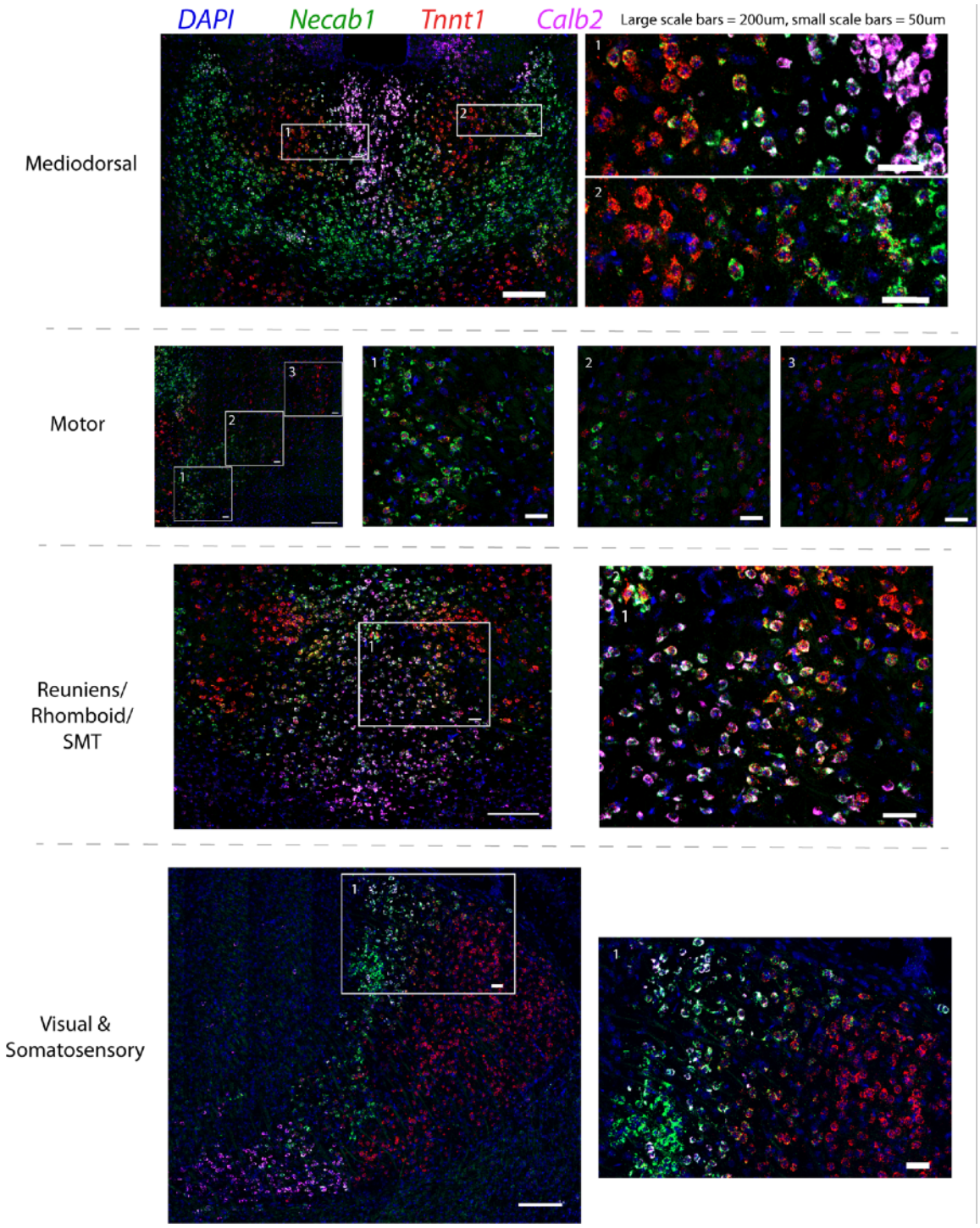


Figure 2.9 - Multi-FISH show cells with mixed expression of profile marker genes

Expanded views of example regions showing intermediate cells expressing combinations of *Tnnt1*, *Necab1* and *Calb2*, which are preferentially expressed in primary, secondary and tertiary nuclear profiles respectively.

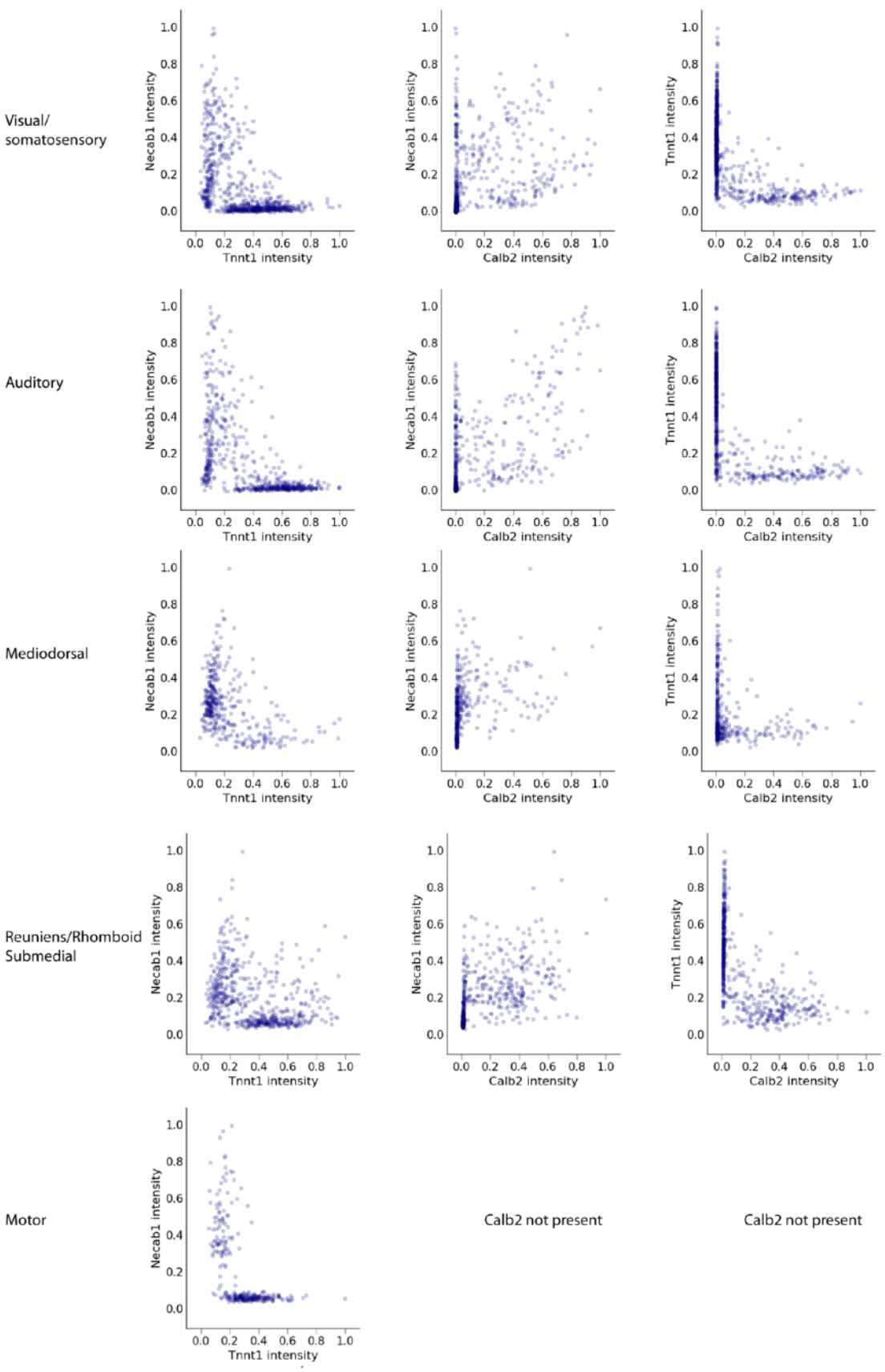

Visual/
somatosensory
Auditory
Mediodorsal
Reuniens/Rhomboid
Submedial
Motor
Necab1 intensity
Tnnt1 intensity
Calb2 intensity
Calb2 not present
Calb2 not present

Figure 2.10 - Quantification of multi-FISH images shows intermediate cells
Quantification of multi-FISH gene expression images confirms the co-expression of markers within the same neurons, supporting a gradient-like organisation. Regions of interest (ROIs) were drawn in ImageJ on images like those in figure 2.9. Intensity was normalized first to the ROI size, then divided by the maximum for that channel. Only cells that express at least one of the marker genes were included.

3.4 Discussion

Contributions of this study

In this study we produced a near-comprehensive transcriptomic atlas of mouse thalamus, and used it to understand the organization of thalamic pathways(Phillips, Schulmann, Hara, Liu, Wang, Shields, Korff, Lemire, Dudman, Nelson, & Hantman, 2018a). In contrast to previous studies that were limited to the evaluation of a small number of markers or nuclei, the high-dimensional nature of our dataset allowed us to address the organization of thalamus in an unsupervised, thalamus-wide manner.

The critical experimental effort and advance in this study was to use anatomical information to guide collection of samples for RNAseq, thus allowing us to identify the genetic profiles of specific thalamic pathways. Simply gathering all of the thalamic neurons and performing single cell sequencing would not have allowed us to infer thalamus-wide information, as it was imperative we knew *where* the genetic information was from (eg, (Saunders et al., 2018; Zeisel et al., 2018)). To the authors knowledge this presents the first coupling of projection type with transcriptomic identity at this scale (near-comprehensive coverage of a major brain region).

We were also able to use this dataset revise the organization of thalamus significantly. Firstly, our analyses reveal that, regardless of modality of information conveyed, thalamic pathways vary in a strikingly similar way. Stated in opposite form, almost all regions of cortex receive input from a shared set of thalamic pathway gene expression profiles, at least in terms of gene expression. We term these primary, secondary and tertiary profiles. There is thus a cross-modal architecture in thalamus, such that the same basic pattern of variation occurs across thalamic modalities. This is summarised in figure 2.11.

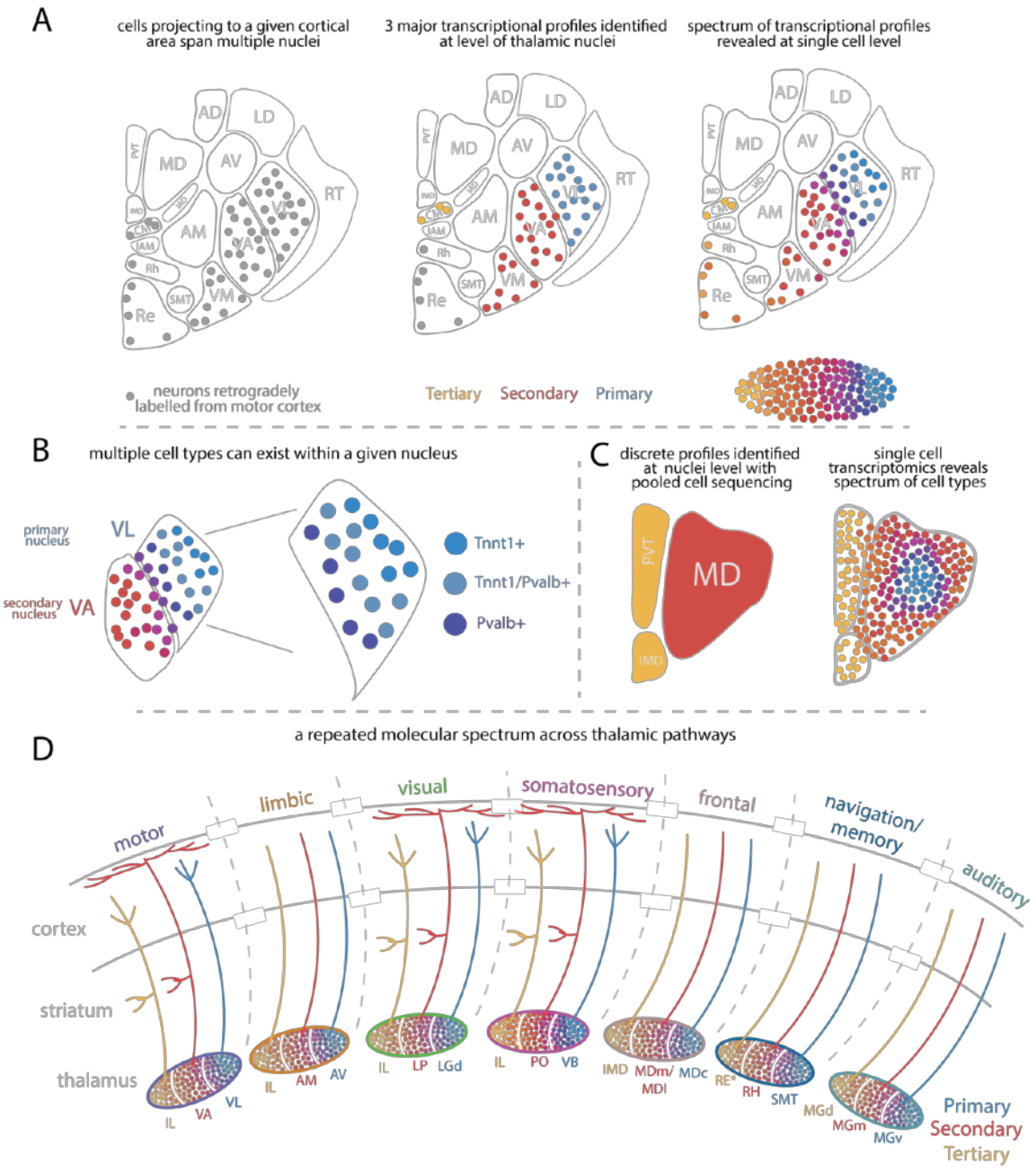


Figure 2.15 – Summary: a conserved molecular spectrum across thalamic projection systems

A. Thalamic neurons have been divided into discrete nuclei. Neurons going to a given cortical area are distributed across multiple nuclei (left), in this example, motor thalamus. We applied pooled cell RNAseq to each nucleus, and found three major classes of thalamic nuclei (middle for example of the three main cellular profiles).

Single cell sequencing revealed that these profiles lie along a continuous spectrum (right).

B. Heterogeneity was present within single nuclei, with multiple cellular classes within VL. Again, intermediate cells were present.

C. This spectrum extended across nuclei serving different projection systems, and again aligned with heterogeneity within nuclei boundaries. For example, a similar spectrum of molecular variation was present in the prefrontal-projecting mediodorsal nucleus.

D. In summary, three major thalamic gene expression profiles are repeated with variation across each projection system and there is continuous cell-type variation between them. Electrophysiology-related genes are the most differentially expressed along the axis, in a manner predictive of action potential waveform (tested in motor thalamus). These gene expression profiles correlate with differing morphological projections, as shown for the motor, visual and somatosensory systems, but morphological features also likely show projection-system specific differences. Nuclear acronyms are as in Supplementary Data Table 1. IL = intralaminar nuclei. * by Reuniens indicates that tertiary status is inferred by expressions of tertiary markers, though it contains additional genetic differences that make it cluster separately in hierarchical clustering.

This is not a wholly new proposal, and our nomenclature has commonalities with existing theories. For example, ‘core’ nuclei from the theory of Jones are encompassed within the ‘primary’ profile identified here, whilst the intralaminar and midline groups are encompassed within the ‘tertiary’ profile(E. G. Jones, 2001; Sherman & Guillery, 2002; Smith et al., 2014). However, there are also important differences. We show that calbindin, the most widely used marker for thalamic cell types, in fact labels a very broad range of neuronal identities, and that nuclei lacking classic thalamic ‘markers’ nonetheless group reliably into the profiles identified here. Indeed, the separation of subsets within the calbindin group is much larger than between the calbindin and parvalbumin groups. We note that despite extensive searching, we never found genes that

reliably and uniquely marked only the secondary nuclear profile in thalamus. This speaks to the limitations of defining statistical groups by individual markers, as has been the practice previously.

Perhaps the most important difference between this classification and others, however, is that it is based upon unsupervised analysis of data from all modalities of thalamus. This is in contrast to earlier theories which have been based predominantly on data from sensory nuclei, and have not been tested beyond it. Whilst one could intuitively expect a similarity between touch and visual processing, it is, at least to this author, a much more intriguing finding and question as to why sensory and non-sensory pathways would be organized similarly. We will explore this theme in the overall thesis discussion.

As well as defining a thalamic organizational scheme that will allow cross-modal comparison of thalamic circuits, our work suggests much greater functional diversity within thalamus than had previously been appreciated. We revealed a striking genetic diversity between thalamic nuclear profiles, including those genes encoding kinases, ion channels and receptors, and used this to make correct predictions about electrophysiological properties (namely, action potential width). This is in substantial contrast to the existing consensus view in thalamic research, that thalamic nuclei are largely genetically and electrophysiologically identical except for developmental genes(Clascá et al., 2012). Why such diversity exists is, at present, largely unexplained. However, to take just one example, different kinases have been linked to effects on wide ranging cell functions, including plasticity, a topic scarcely explored in thalamus (Kandel et al., 2012). Our dataset thus offers an exciting motivation and resource for future exploratory and hypothesis-based research seeking to relate neuronal properties to circuit function.

Why was this diversity, in particular the striking pattern of electrophysiological variation across thalamus, not observed previously? Earlier genetic studies have typically focused

upon development of thalamic circuits, and thus primarily evaluated transcription factor differences rather than other gene classes (E. G. Jones & Rubenstein, 2004; Murray, Choudary, & Jones, 2007; Nagalski et al., 2015). Further, these prior genetic studies of thalamus have not used RNAseq, and therefore were limited in sensitivity and/or range of genes detected(Yuge et al., 2010). Ion channels and receptors are relatively low in expression levels, meaning they would be biased against in these screens. Finally, whilst areas such as neocortex and hippocampus have been used as a model for topics such as synaptic plasticity, thalamus has less often been studied for at the single cell level. There has therefore been little motivation to search for protein expression patterns that could explain thalamic functions. Together, these reasons likely explain the lack of prior recognition of thalamus' molecular diversity.

As for the systematic pattern of electrophysiological diversity being missed, this likely reflects the manner of investigation. Studies of thalamus electrophysiologically have typically focused upon only a single, or at most two to three, nuclei, investigating within rather than between-nuclei variation(Browne, Kang, Akk, Chiang, Schulman, Huguenard, & Prince, 2001b; Jhangiani-Jashanmal, Yamamoto, Gungor, & Paré, 2016; J. Li, 2003; Steriade, Curró Dossi, & Contreras, 1993b). To the authors knowledge, no study has compared nuclei representing the three profiles identified above before. This means that no study has evaluated electrophysiological properties across the leading axes of variance that we reveal here, explaining why the principles identified in this paper have been previously been missed. There have, however, been many studies identifying electrophysiological and pharmacological differences between thalamic nuclei, and the dataset gathered here should provide a valuable resource for identifying the genetic and molecular specializations underlying these differences(Clascá et al., 2012; E. Jones, 2007).

Finally, we show that rather than multiple, unrelated nuclei projecting to a single cortical region, thalamic pathways lie on a continuous spectrum of neuronal identities. That is,

there is a progressive, continuous difference between the genetic profiles that project to a given cortical region, without clear discreteness along this axis. This is contrast to the current view of there being clearly separate thalamic nuclei(E. Jones, 2007; Paxinos & Franklin, 2012). The key evidence for this is the co-localisation of markers of the different nuclear profiles, shown via both single cell RNAseq and multi-FISH. These intermediate cells were topographically localized to the boundaries of the nuclei. To our knowledge, there is only one prior study reporting the co-localization of different calcium binding proteins within the same thalamic neurons, namely a study from Sherman lab in auditory thalamus (Lu, Llano, & Sherman, 2009). However, again this likely reflects the lack of prior research on this. It will be intriguing to investigate whether other neuronal properties, such as anatomical projection pattern, also vary continuously along this axis. Such a continuous, rather than discrete, variation provides an obvious increase in possible functional contributions, and has recently been observed in other brain regions including hippocampus and striatum(Cembrowski & Menon, 2018; Cembrowski, Bachman, Wang, Sugino, Shields, & Spruston, 2016a; Gokce, Stanley, Treutlein, Neff, Camp, Malenka, Rothwell, Fuccillo, Südhof, & Quake, 2016b; Tasic et al., 2016).

Despite arguing for a continuous variation of cell types in thalamus, it is important to highlight that we are not advocating an end to the use of nuclear names in thalamus – they are an important part of language used by neuroscientist. Rather, we argue simply that they should be viewed analogously to names of colours: ie, the use of ‘red’, ‘green’ and ‘blue’ to discretely label positions along a continuous spectrum, and ‘CM’, ‘VA’, ‘VL’ to label positions along the spectrum of cells projecting to motor cortex. Further, our decision to have three labels (primary, secondary and tertiary), as opposed to four, is also arbitrary. Indeed, the first branch within the tertiary group divides calretinin+ from calretinin- nuclei (one could therefore speak of tertiary-a and tertiary-b). We chose three largely for simplicity and for its ease of relation to existing theories of thalamic organization (with loose corresponding of tertiary and intralaminar/midline, for instance).

Limitations of this study

There are at least three key limitations to this study that further work could improve upon.

The first is that our dataset, with the exception of electrophysiology data from motor-projecting thalamus, is solely genetic. The link between gene expression and resulting cellular properties is both difficult to predict, indirect and indeterminate: other factors relating to the animal's experience play a critical role. There could other axes of variance, for example anatomical projection pattern, that differ strikingly along an axis with minimal genetic variance. It will thus be important to compare classifications developed here with those that result from examining other information types, such as single neuron reconstructions (currently not comprehensive enough to detect systematic patterns).

The second limitation is that RNAseq remains imperfectly sensitive, especially at the single cell level, likely detecting at most a third of the transcripts from the cell body(Tasic et al., 2016). This is partly due to inefficiencies in amplification of the mRNA. This makes it likely that many low expressed genes are being missed in each cell, potentially obscuring detection of particular neuronal subtypes. However, we view this as unlikely to fundamentally change our core results, as the pooled cell data is in strong concordance with the single cell data, and has much higher sensitivity(Sugino et al., 2017). Further, the central results from our dimensionality reduction analysis are robust to gene selection method, suggesting our results do not strongly depend upon specific gene sets (see Figure 2.5, for PCA vs MDS with different gene selection metrics).

Thirdly, there likely exists further diversity not captured here due to our single cell sequencing not being comprehensive across thalamus. During manual examination of the

Allen atlas, we found homogeneity within the anterior nuclei, parafascicular and lateral dorsal nuclei. This diversity could reflect genetic correlates of known anatomical diversity within some of these nuclei(E. Jones, 2007; T. Kita, Shigematsu, & Kita, 2016b). Further single cell efforts should be preferentially targeted to those nuclei.

There are some additional caveats that should be kept in mind. Firstly, we did not verify that the spectrum of transcriptional types is reflected in protein expression. It is possible that a nonlinearity in the translational process exists that creates clearly discrete cell types from a gradient of underlying transcriptional states. Secondly, our results may only apply in the adult: in development transcription factors could be far more prominent sources of heterogeneity/cell type identity. Ultimately, integrated study of multiple modalities of information in the same cell will be needed to definitively address cellular identity, and to clarify the best way to define a 'cell type' (my suspicion is that the notion of 'cell type' will become more fluid, and the never ending quest to find more cell types will not be seen as profitable to producing scientific understanding. In essence, what constitutes a cell type will depend on what one is seeking to understand, a point explored further in the discussion to the overall thesis).

Developmental sequence as a source of the topographic spectrum

The dominant approach to describing the diversity of cells in the body has been to categorise them into discrete classes(Sugino et al., 2017). In such a view, distinct transcriptional cascades act to funnel neuronal precursors to discrete end points, consisting of relatively homogenous 'cell types'. Indeed, the mechanism of mutually inhibitory transcription factors would appear to enforce such discreteness.

The possibility of continuous cell feature variation has, perhaps consequently, only recently attracted substantial attention(Cembrowski & Menon, 2018; Huntenburg, Bazin, & Margulies, 2018). In such continuous variation, cells within a cell type differ to a

similar extent previously used to define discrete cell types, but instead have a continuum of intermediate cells between them. Thanks to the advent of single cell RNAseq and multi-FISH, such graded variation has now been observed in neocortex, striatum and hippocampus (Cembrowski, Bachman, Wang, Sugino, Shields, & Spruston, 2016b; Gokce, Stanley, Treutlein, Neff, Camp, Malenka, Rothwell, Fuccillo, Südhof, & Quake, 2016a; Tasic et al., 2016). It has also been seen outside of the neurons: a gradient of cell types exists in vascular tissue, also(Vanlandewijck et al., 2018). It may therefore reflect a broader principle of cellular diversity. However, the circuit functions and mechanisms underlying such gradients are not yet understood.

We observed such graded variation in thalamus, and argue here that thalamus offers several advantages for further investigation of these topics. Though a topography of such continuous cell types has also been observed in hippocampus, thalamus has far greater genetic simplicity than hippocampus(Cembrowski, Bachman, Wang, Sugino, Shields, & Spruston, 2016b). For instance, hippocampus has many interneurons, whilst rodent thalamus (LGd apart) entirely lacks them. This means that a given spatial position in thalamus is largely genetic homogenous, greatly simplifying the investigators task: spatial location and genetic identity are closely coupled. This could make ‘projecting graded transcriptomic identities onto higher-order features’ (such as anatomical projection pattern), as argued for by Cembrowski and Menon, a much easier task(Cembrowski & Menon, 2018).

We can exploit this simplicity to ask about the mechanisms underlying this cell type spectrum. Though relatively closely linked to the mediolateral position of a cell, it is not purely a mediolateral spectrum. For example, the gradient from Reuniens to Rhomboid, visible in Figure 2.9, runs ventral-dorsal. Still, however, the cell type variation is topographic.

What underlies the topographic spectrum? Rather than simply reflecting the final position of a cell, we view genetic identity in thalamus as likely being coupled to the birth date of the neuron. The concept of a 'neurogenetic spectrum' was introduced in thalamic development in the 1980s (and been essentially forgotten), when it was observed in multiple thalamic nuclei that neurons in different locations of a nucleus were born in different times(Altman & Bayer, 1988; 1989). Neuronal birth date and final location were correlated. More recently, clonal analysis has shown close coupling of neuronal birthdate of a neuron at the thalamus-wide level(Shi et al., 2017).

Studies examining calcium binding protein expression across thalamic development offer a more direct coupling of gene expression profile to developmental stage. They show that calcium binding protein labelling in thalamus indeed appears as one would predict, with calretinin, a marker for the most tertiary (and medially-located) neurons, appearing first during development, then calbindin, then finally parvalbumin(Frassoni, Arcelli, Selvaggio, & Spreafico, 1998; Puelles, Sánchez, Spreafico, & Fairén, 1992; J.-Y. Zhang et al., 2015). Notably, whether the neurons are *born* earlier, or simply show labelling sooner, is not clearly separated in all of these studies.

Testing this coupling of birth date and gene expression profile directly would require sequencing of neurons by birthdate. Unfortunately, existing methods for fluorescently labelling neurons at a particular birthdate involve potentially mutagenic reactions, which could distort RNAseq results (though whether this would alter the underlying conclusions substantially is doubtful) (Habib et al., 2016; Salic & Mitchison, 2008). Recent advances in lineage mapping in 'simpler' model organisms could provide an alternative approach (Briggs et al., 2018; Farrell et al., 2018; Wagner et al., 2018).

The mechanism underlying this gradient of cellular diversity remains unresolved, and, as noted above, is difficult to explain using the classic transcription factor code approach. Another possibility is that dynamic factors extrinsic to the neuron during its

developmental period could drive the resulting specialization. This has already been seen in thalamus: refinement of the retinogeniculate projection has been shown to rely on cortical activity, and thalamic transcriptional state is influenced by peripeheral input also(Frangeul et al., 2016; Thompson, Picard, Min, Fagiolini, & Chen, 2016).

One possibility is that such gradients of cellular properties could arise from genome:neuronal activity interactions(Wong et al., 2018; L. I. Zhang & Poo, 2001). The oscillatory pattern of thalamocortical interactions changes markedly over development, and this could provide a drive for different gene expression patterns. For example, neurons born earlier could receive input from slower oscillations. In such a scheme, thalamic neurons are not born with a particular gene expression profile predetermined, but rather specialize based upon the neural activity patterns currently propagating through forebrain circuits. Notably, calcium levels, controllers of which are such a prominent marker of thalamic cell types, have been linked to cell type(O'Leary, Williams, Franci, & Marder, 2014; Spitzer, Root, & Borodinsky, 2004). One could test this by perturbing the patterns of input to, or firing patterns of, developing thalamic neurons, and seeing how this alters their transcriptional fate.

Conclusion

The transcriptomic dataset obtained in this project provides information critical for more detailed investigation of individual nuclei and thalamus-wide organization. Further, our analysis shows striking covariation in the organization of thalamic pathways serving all modalities, establishing a simple and comprehensive thalamic functional architecture derived from comprehensive examination of thalamic transcriptomes. Previously defined functional groups in thalamus map onto profiles along the single axis identified here. However, by identifying an organization in a statistically unsupervised manner based upon a broad set of genes, we provide a more fundamental spectrum on which an increasingly detailed understanding of thalamus can be built (Figure 2.11). The genetic

axis defining thalamic cell types identified here is dominated by genes that can directly shape neuronal properties. Together, our results reveal deep similarities in the properties of thalamic pathways serving diverse modalities of cognition, establishing a common framework for understanding the central communication hub of thought.

SUPPLEMENTARY TABLES

Supplementary Data Table 1 – Sample collection approaches (pooled-cell RNAseq) [Validation] refers to samples in Figure 2.3. For transgenic lines used, see Supplementary Data Table 4.

Region name:	**Labelling methods:**	**Replicates:**
Anterior dorsal (AD)	Retrograde DiI from retrosplenial cortex. [Validation: Slc6a5-Cre.KF109 transgenic line]	4,[3]

Anterior medial (AM)	Retrograde DiI from cingulate cortex. Retrograde AAV from cingulate cortex.	4,1
Anterior ventral (AV)	Retrograde DiI from retrosplenial cortex. [Validation: Slc6a5-Cre.KF109 transgenic line]	8,[4]
Central lateral (CL)	rAAV2-retro from striatum. [Validation: Grp-Cre.KH288 transgenic line]	4
Central medial (CM)	rAAV2-retro from striatum. Retrograde DiI to striatum. [Validation: Grp-Cre.KH288 transgenic line]	4,3,[1]
Interanterodorsal nucleus (IAD)	Retrograde DiI from cingulate cortex	3
Intermediodorsal nucleus (IMD)	rAAV2-retro from striatum.	4
Laterodorsal nucleus (LD)	Retrograde DiI from retrosplenial cortex.	5
Lateral Geniculate (dorsal) nucleus (LGd)	Retrograde DiI from visual cortex with anterograde CTB-488 tracer injected to retina (Cholera Toxin Subunit B, Alexa Fluor 488™, Molecular Probes). [Validation: Slc6a5-Cre.KF109 and Gpr26-Cre.KO250 transgenic lines]	5,[3],[4]

Lateral posterior nucleus (LP)	Retrograde DiI from visual cortex with anterograde CTB-488 tracer injected to retina (Cholera Toxin Subunit B, Alexa Fluor 488™, Molecular Probes).	9
Medial geniculate body (MGB)	Not collected for pooled-cell RNAseq in this study. Consists of ventral (MGBv), dorsal (MGBd) and medial (MGBm) subdivisions.	na
Mediodorsal nucleus (MD)	Retrograde DiI from frontal cortex. Collected without subdividing the three subdivisions (central: MDc, medial: MDm, lateral: MDl).	6
Paracentral nucleus (PCN)	rAAV2-retro from striatum. [Validation: Grp-Cre.KH288 transgenic line]	4,[3]
Parafascicular nucleus (PF)	rAAV2-retro from striatum. This did not label the entire PF complex, and likely reflects a genetic subset.	4
Posterior medial nucleus (PO)	Retrograde DiI from somatosensory cortex. [Validation: Gpr26-Cre.KO250 transgenic line]	7,[4]
Paraventricular thalamus (PVT)	rAAV2-retro from striatum [Validation: Gpr26-Cre.KO250 transgenic line]	3,[3]
Reuniens nucleus (RE)	Retrograde DiI from cingulate cortex	8

Rhomboid nucleus (RH)	Not collected for pooled-cell sequencing in this study.	na
Subparafascicular nucleus (SPA)	rAAV2-retro from striatum. This did not label the entire parafascicular complex, and likely reflects a genetic subset (medially biased).	4
Submedial thalamus (SMT)	Not collected for pooled-cell RNAseq in this study.	na
Ventral anterior nucleus (VA)	Retrograde DiI from motor cortex. Anterograde labelling of inputs from SNr and DCN with viral tracers (AAV2/1-CAG-GFP and/or AAV2/1-CAG-BFP).	8
Ventral medial nucleus (VM)	Retrograde DiI from motor cortex. Anterograde labelling of inputs from SNr and DCN with viral tracers (AAV2/1-CAG-GFP and/or AAV2/1-CAG-BFP).	6
Ventral lateral nucleus (VL)	Retrograde DiI from motor cortex. Anterograde labelling of inputs from from SNr and DCN with viral tracers (AAV2/1-CAG-GFP and/or AAV2/1-CAG-BFP).	6
Ventrobasal nucleus (VB)	Retrograde DiI from somatosensory cortex [Validation: Grp-Cre.KH288 transgenic line for VPM]	4,[3]
Ventroposteromedial parvocellular (VPMpc)	Retrograde DiI from insular cortex.	6

Supplementary Data Table 2 – Coordinates for retrograde labelling/trace injections (pooled-cell RNAseq)

All depths relative to brain surface. If depth not stated, injections were made at 300µm and 600µm deep.

Sample area:	**Injection coordinates (in millimeters, from bregma, depth from brain surface)**	***Additional comments***
AD	1.7 caudal, 0.25 lateral, 0.4, 1 deep.	
AM	1.35 rostral, 0.2 lateral, 1.25, 1.8 deep.	
AV	1.7 caudal, 0.25 lateral, 0.4, 1 deep.	
CL	0.7 rostral, 1.9 lateral, 2.75, 3.00, 3.25 deep.	For CL, the top 1/3 of the rostral intralaminar nuclei was taken.
CM	0.7 rostral, 1.9 lateral, 2.75, 3.00, 3.25 deep.	For CM, the bottom 1/3 of the rostral intralaminar nuclei was taken.
IAD	1.35 rostral, 0.2 lateral, 1.25, 1.8 deep.	

LD	1.7 caudal, 0.25 lateral, 0.4, 1 deep.	
LGd	2.9 caudal, 2.4 lateral. 2.9 caudal, 1.7 lateral. 3.8 caudal, 3.0 lateral 3.8 caudal, 2.0 lateral	Retina was injected with anterograde tracer (CTB) to distinguish LGd from LP
LP	2.9 caudal, 2.4 lateral. 2.9 caudal, 1.7 lateral. 3.8 caudal, 3.0 lateral. 3.8 caudal, 2.0 lateral.	Retina was injected with anterograde tracer (CTB) to distinguish LGd from LP
MD	2 rostral, 1 lateral, 3.0, 2.5, 2.0 deep. 2 rostral, 1.6 lateral, 3.0, 2.4,1.5 deep. 1.7 rostral, 0.4 lateral, 3.0, 2.4, 1.5 deep.	
PCN	0.7 rostral, 1.9 lateral, 2.75, 3.00, 3.25 deep.	For PCN, the middle 1/3 of the rostral intralaminar nuclei was taken.

PF	0.7 rostral, 1.9 lateral, 2.75, 3.00, 3.25 deep.	
PO	0.85, 1.5 caudal,2.75 lateral, 0.4 deep. 1.2 caudal, 2.25 lateral, 0.4 deep.	
RE	1.35 rostral, 0.2 lateral, 1.25+1.8mm deep.	
SPA	0.7 rostral, 1.9 lateral, 2.75, 3.00, 3.25 deep.	
VA	0.3 rostral, 1.6 lateral. 0.8 rostral, 1.6 lateral. 0.5 rostral, 1.4 lateral.	
VM	0.3 rostral, 1.6 lateral 0.8 rostral, 1.6 lateral. 0.5 rostral, 1.4 lateral.	
VL	0.3 rostral, 1.6 lateral. 0.8 rostral, 1.6 lateral. 0.5 rostral, 1.4 lateral.	
VB	0.85, 1.5 caudal, 2.75 lateral, 0.4 deep. 1.2 caudal, 2.25 lateral, 0.4 deep.	
VPMpc	0.25 caudal, 4 lateral, 2.4,2.7,2.9 deep.	

Supplementary Data Table 3 – Injection coordinates for retrograde labelling (single-cell RNAseq)

Sample area:	Injection coordinates (in millimeters, from bregma, depth from brain surface)
Auditory Thalamus	3.5 caudal, 4.5 lateral, 2.25 deep. 3.16 caudal, 4.5 lateral, 2.0 deep.
Mediodorsal (Prefrontal) Thalamus	2 rostral, 1 lateral, 3.0, 2.5,2.0 deep. 2 rostral, 1.6 lateral, 3.0,2.4,1.5 deep. 1.7 rostral, 0.4 lateral, 3.0, 2.4, 1.5 deep.
Motor Thalamus	0.3 rostral, 1.6 lateral. 0.8 rostral, 1.6 lateral. 0.5 rostral, 1.4 lateral.
Somatosensory Thalamus	0.85 and 1.5 caudal, 2.75 lateral, 0.4 deep. 1.2 caudal, 2.25 lateral, 0.4 deep.
Visual thalamus	2.9 caudal, 2.4 lateral. 2.9 caudal, 1.7 lateral. 3.8 caudal, 3.0 lateral. 3.8 caudal, 2.0 lateral.

Supplementary Data Table 4 – Transgenic mice used in this study

Transgenic line:	Areas collected
Slc6a5-Cre.KF109 (GENSAT)	Anterior dorsal, Anterior ventral, Dorsal lateral geniculate (LGd).
Gpr26-Cre.KO250. (GENSAT)	Paraventricular nucleus (PVT), Posterior medial (PO)
Grp-Cre.KH288 (GENSAT)	Rostral intralaminar nuclei (Central lateral, paracentral), ventral posteromedial nucleus (VPM).

Chapter 3

Dissociable dynamics between genetically defined motor cortex outputs

Introduction:

The previous chapter investigated how signals are routed to neocortex via thalamus. The original motivation for that study was to understand the organisation of inputs from subcortical motor structures to motor cortex. We now return to this theme from the other side, investigating motor cortex itself, and how it interacts with subcortical motor structures. The study of motor cortex has a long history, dating at least to studies by Ferrier, Hitzig and Fritsch from the 19th century that elicited movement from cortex via electrical stimulation(Hatsopoulos & Suminski, 2011). However, despite intensive study, to date a clear understanding of its actual role in movement generation remains elusive.

Motor cortex in rodents

Due to the advent of genetic tools and ease of experimentation compared to primates, motor control is increasingly studied in rodents. This led to a number of studies seeking to map which areas of rodent cerebral cortex can elicit movement, and at which stimulation strength (with lower stimulation strength required taken to imply a closer link to movement) (Harrison, Ayling, & Murphy, 2012; Ramanathan, Conner, & Tuszynski, 2006; Tennant et al., 2010).These studies identify an area 0-1mm rostral and 1.5mm lateral to bregma in mice, and in a corresponding position in rats, from which forelimb movements can be evoked(Paxinos & Franklin, 2012). The area is relatively large and imprecise in its boundaries, so we take the center of this area as forelimb M1 for this study. Of course, just because an area being stimulated does not elicit movement directly

does not mean it is not involved in the computations involved in motor control: however, this area makes a sensible place to begin investigation, and is our focus in this chapter.

This region from which movements can be elicited has now been studied in some detail by anatomists. It has direct corticospinal projections, as well as projections to collicular, basal ganglia and brain stem regions including the pons(Hooks, Papale, Paletzki, Feroze, Eastwood, Couey, Winnubst, Chandrashekar, & Gerfen, 2018a; Xuhua Wang et al., 2017). This closely resembles the primate brain, though direct contacts onto spinal motor neurons have not been shown(Kandel et al., 2012). It also projects within cerebral cortex, including to sensory cortex, where it contacts a diverse range of both excitatory and inhibitory neurons(Kinnischtzke, Simons, & Fanselow, 2014). This projection profile puts it in an ideal position to influence a wide range of targets relevant for motor control.

Further, though often viewed as not exhibiting dexterous behaviour, rodents can exhibit complex, often bimanual, motor control, including manipulation of small pellets for eating(J.-Z. Guo et al., 2015). Rodents, and their motor cortex, therefore offer a model for studying non-reflexive, purposeful motor control.

The role of motor cortex in rodent motor control

What is the function of motor cortex in movement? Many roles have been proposed, and none are universally accepted. These proposed functions vary from control of basic kinetics (the forces a muscle produces), to kinematics (such as direction, velocity of a controlled element), through to encoding of corrective motor commands, tutorial signals for subcortical areas, and even *inhibition* of movements (Ebbesen, Doron, Lenschow, & Brecht, 2016; Hatsopoulos, 2005; Heindorf et al., 2018; Otchy, Wolff, Rhee, Pehlevan, Kawai, Kempf, Gobes, & Ölveczky, 2015b). With the exception of the latter two, these theories make a prediction that motor cortex outputs should be active primarily *during*

movement, potentially with internal computations preceding this. Further, perturbation of motor cortex should directly alter movement.

What do experiments show? Transient experimental perturbation of motor cortex can elicit complex motor behaviours in both primates and rodents(Graziano, Taylor, & Moore, 2002; J.-Z. Guo et al., 2015).This, combined with close coupling of motor cortical activity with movement kinematics in primates, has supported the view that motor cortex can command precise movement kinematics/kinetics. As in primate motor cortex, most recording studies in rodent motor cortex show prominent peri-movement modulation, including both increases and decreases in activity (Estebanez, Hoffmann, Voigt, & Poulet, 2017; Hasegawa et al., 2017; D. Huber et al., 2012; Peters, Lee, Hedrick, O'Neil, & Komiyama, 2017). However, the closeness of the coupling between neuronal activity and kinematics is not so well explored in rodents.

In both primates and rodents, whilst most studies show some form of deficit in motor control following motor cortical inactivation, typically the basic movement itself is intact(Heindorf et al., 2018). Often, for example, only the timing may be disrupted, leaving the approximate kinematics relatively unchanged. This is especially true in rodents, and has led to the suggestion that only highly dexterous movement is directly controlled by cortex, or that motor cortex acts to 'fine tune' the more primitive motor commands that subcortical areas can provide(J.-Z. Guo et al., 2015). In summary, though it remains unclear what the precise 'function' of motor cortex is, it is clear that it is possible to produce movement without it, but that this movement is not identical to that produced in the unperturbed state. A clear understanding of what this difference constitutes is elusive.

Limitations in existing research on motor cortex

There are a number of limitations with existing studies of motor cortex. The first is the difficulty of interpreting the results of inactivation experiments. Inactivating an area of the brain can have additional, off-target effects upon downstream neural circuits, for example producing a sudden and dramatic change in excitatory/inhibitory balance(Otchy, Wolff, Rhee, Pehlevan, Kawai, Kempf, Gobes, & Ölveczky, 2015b). For example, if a brain area outputs a tonic resting firing rate, inactivating it will not only remove task-related activity (on-target), but also this tonic activity (off-target), the reduction of which would have effects by itself. This leads to difficulties with the interpretation of inactivation studies, as one cannot know whether the result is due to on-target or off-target effects. This could potentially be ameliorated by having clear predictions to make ahead of said experiments, yet to date experimental perturbation in motor cortex tend not to be designed as direct tests of hypothesis inspired by data, due to the relatively imprecise nature of current models.

Another significant issue is that the large majority of studies cue the movement initiation and immediately reward movement completion, such that behavioural output itself is temporally confounded by both the initiation and reward cues(D. Huber et al., 2012; Peters et al., 2017). This cueing is often done to reduce the additional training burden of adding delayed periods, which prolongs training. However, it also means that one struggles to cleanly separate movement from cue-responsive signals during analysis. Studying internally generated movements (ie, uncued) could help to avoid this difficulty, but would bear the cost of less stereotyped behaviour. Further, only a minority of behaviour is due to an externally triggered cue. In watching my mice in their cages (as well as my beautiful pet hamster called Cosmo), it is apparent that the large majority of their behaviour is either entirely internally generated, or in response to a stimulus they themselves have created by exploring their environment (such as falling off their cage bars whilst practicing their climbing skills). Natural movement is thus quite different to that typically studied in the lab.

A final issue is that recordings from motor cortex typically lack cell type identity, due to the technical difficulty of achieving this. We consider this final issue in the following section.

The contribution of motor cortex to subcortical areas - L5PT and L5IT neurons

To use the language of the introduction to this thesis, it may be challenging at present, or perhaps even an incorrect approach, to try to identify the *function* of motor cortex. An alternative is to identify the *contribution* that motor cortex makes to brain dynamics during movement. This is, of course, a large topic all by itself. An intermediate step to understanding this would be to identify the information that motor cortex sends to its different targets during behaviour, i.e., what are the different routes through which it can influence its major targets, such as the basal ganglia, striatum and superior colliculus, as well as other cortical areas? Once this is identified, one can design hypothesis-driven experiments to investigate what effects those signals have once they reach their target.

Here we shall focus upon the main output layer of neocortex, layer 5. This layer has been divided into two major classes of projection neuron(Reiner, 2010). The first is the pyramidal tract (herein, PT) neuron, which projects to targets including brain stem nuclei (such as the cerebellar relays), thalamus and striatum. This is the only output neuron of neocortex to leave the Telecenphalon, but remains exclusively ipsilateral to its cell body. It is also notable that there is substantial anatomical and genetic diversity amongst these neurons, such that individual neurons do not project to all of the PT targets, but rather a subset(Hattox & Nelson, 2007). Morphologically, they have elaborate dendrites in layer 1, and electrophysiologically are characterised by a lack of spike frequency adaptation upon sustained current injection(Hattox & Nelson, 2007).

The second type of layer 5 output is the layer 5 intratelencephalic neuron (herein, IT). These neurons project exclusively within the forebrain to the striatum and within cerebral

cortex itself. Unlike PT neurons, IT neurons also project contralaterally, thus providing a link between the two cortical hemispheres and between cortex and contralateral striatum. Though their dendrites reach layer 1, their ramifications are notably sparser than those of PT neurons. Electrophysiologically, their spike trains show prominent spike frequency adaptation, in contrast to PT neurons(Hattox & Nelson, 2007). Further, IT neuron's apical dendrites integrate over a longer temporal window than PT neurons(Dembrow, Zemelman, & Johnston, 2015). The functional relevance of these observations remains unclear.

The relationship between PT and IT neurons is also of note. The mean depth of PT neurons is slightly deeper than IT neurons, with PT neurons residing predominantly in L5b, and IT neurons in L5a. IT neurons substantially monosynaptically contact PT neurons, whilst monosynaptic inputs from PT to IT neurons are extremely sparse(Harris & Shepherd, 2015; Reiner, 2010). Though of course multi-synaptic PT to IT interactions will occur, this suggests a relatively feedforward interaction monosynaptically, with IT neurons driving PT neurons, but not vice versa.

In summary, PT neurons have extensive distal dendrites and project to a diverse range of ipsilateral subcortical targets, whilst IT neurons have relatively smaller dendrites and project bilaterally, and to forebrain targets only. IT neurons can drive PT neurons monosynaptically, but not vice versa. This is summarised in figure 3.1. Though these individual features have not been examined in all major , they are assumed to be relatively general features of PT and IT classes across cerebral cortex(Harris & Shepherd, 2015). Together, this suggests two types of layer 5 outputs to ipsilateral striatum, but separate information streams to the contralateral forebrain as opposed to the ipsilateral brainstem. Our next question is to understand the different information that reaches these two parts.

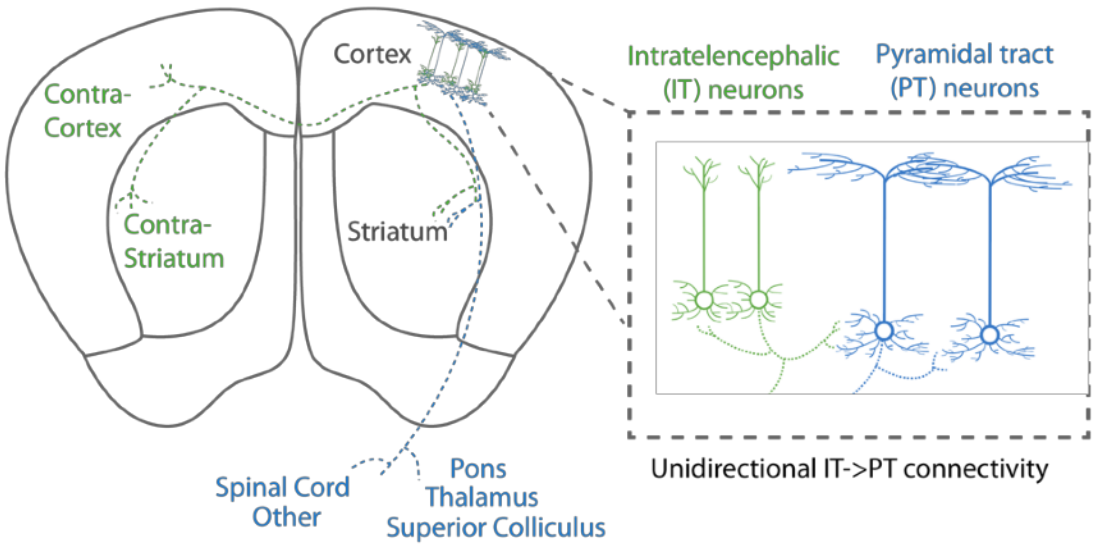


Figure 3.1. PT and IT neurons comprise the two major types of layer 5 output neuron

Left) Pyramidal tract (PT, blue) and Intratelencephalic (IT, green)-type layer 5 projection neurons have different brain-wide targets. Their defining characteristics are that PT neurons project outside of telencephalon, whilst IT neurons cross to the other telencephalic hemisphere. Right) Illustration of IT and PT type neurons. PT neurons lie deeper within layer 5 (L5) than do IT neurons (Broadly, L5a for IT, L5b for PT,), and their distal dendrites cover a broader area in cortical layer 1. (Harris & Shepherd, 2015; Hattox & Nelson, 2007)

Existing output-type-specific studies of motor cortex

Identified recordings from these two subtypes are few in number, as most studies know neither if their neurons are in layer 5 nor which class the neuron recorded from falls into. This is largely because it is a laborious experiment to identify the projection target of a recorded neuron: not only must one antidromically confirm the neuronal identity, but this assumes the neuron is even of the IT or PT class. Nonetheless, a small number of recordings have identified PT vs IT information during recording(Bauswein, Fromm, &

Preuss, 1989; N. Li, Chen, Guo, Gerfen, & Svoboda, 2015a; Soma et al., 2017; Turner & DeLong, 2000).

This has been examined in both rodents and primates. Two rodent studies show that in cued motor tasks, PT type neurons are more directionally tuned, with IT neurons responding without relative bias (on average) for ipsi vs contralateral movements. This has been interpreted as arguing for a 'feedforward' basis for 'planning and movement(N. Li, Chen, Guo, Gerfen, & Svoboda, 2015b; Soma et al., 2017). However, these studies do not find other major differences between the cell types. In primates, two studies report during unimanual reaching tasks that PT neurons tend to exhibit tuning to multiple task stages, such as peaks at initiation and end of movement, whilst IT neurons tend to be tuned for single task events(Bauswein et al., 1989; Turner & DeLong, 2000). These studies again have the limitations outlined above: it is challenging to separate different components of the task cleanly due to them not being cleanly separated in time.

The approach of this thesis chapter

In this study, we sought to assess how different, genetically defined outputs of motor cortex relate to behaviour. The use of genetic approaches allowed us to record from large numbers of identified PT and IT neurons, meaning we could record neurons regardless of whether they were significantly task modulated or not. Further, we used a relatively distinctive motor task. As noted above, any previous studies of cortex's role in motor control have relied upon cue-initiated motor behaviours, and typically the outcome of the motor task (delivery of reward/no-reward) has been tightly locked to the end of movement. The task used in this study was almost entirely lacking in external cues, and temporally separated movement planning, from movement, from reward(Panigrahi et al., 2015). We used this to assess the signals sent to different subcortical targets by motor cortex.

Methods

Experimental methods

Animal Care

Male and female mice, typically aged 8-16 weeks at time of surgery, were used in this study. All procedures were approved by the Janelia Research Campus Institutional Animal Care and Use Committee (IACUC) and were consistent with the standards of the Association for Assessment and Accreditation of Laboratory Animal Care. Mice were water restricted (typically to 1ml water/day), and their weight and signs of health were monitored daily(Z. V. Guo et al., 2014).

Transgenic lines for selective layer 5 subset labelling

Tlx3-cre (PL56) and Sim1-Cre (KJ18) mice of C57/Bl6 background were used to label the two Layer 5 projection neuron populations (IT and PT neurons respectively)(Gerfen, Paletzki, & Heintz, 2013a). These lines have previously been used to study IT and PT populations independently (Hooks, Papale, Paletzki, Feroze, Eastwood, Couey, Winnubst, Chandrashekar, & Gerfen, 2018b; N. Li, Chen, Guo, Gerfen, & Svoboda, 2015c). As previously shown, we confirmed that in the IT line there was minimal/absent labelling in the pyramidal tract, and similarly in the corpus callosum for the PT line. We note that this does not mean the lines label all of the respective populations, they may be capturing subsets of the IT and PT populations.

Behavioural setup

Behavioural code was adapted from Panigrahi *et al* and run from an Arduino microcontroller (Panigrahi et al., 2015). Behavioural signals were recorded by a Blackrock Microsystems data acquisition system, sampled at 1kHz.

After surgery (see below), mice were given 5 days recovery prior to beginning water restriction (1ml water/day). Following 3-5 days of initial water restriction, they underwent 10-15 days training, which simply consisted of them attempting to perform the task without outside input. No metrics other than ability to complete a session of 120 trials in under 20 minutes were used. The priority was on imaging relatively quickly post-surgery, due to the deterioration of optical imaging quality that occurs over time due to dural thickening(Goldey et al., 2014).

Mice were head-fixed in a custom made head restraint box using the RIVETS head-fixation apparatus, which is a 3D plastic-polymer that 'clips' from both sides, using pressure to stabilize the 3D printed element (see Osborne & Dudman, 2014 for methods). The mouse's front paws rested on a metal bar attached to a spring-loaded joystick, which had unconstrained maneuverability in the horizontal plane. There were up to 120 trials per session (some sessions were incomplete), with one reward being available per trial. Mice were rewarded for pushing the joystick a specific distance in either in either the x (front/back herein) or y(left right herein) axis, such that the reach boundaries formed a square, not a circle. This threshold changed over the course of a session, as in figure 3.2C. This threshold was not cued to the mouse in any way, but rather had to be identified through experience.

Imaging experiments

3mm-wide circular imaging windows were made over the left cortical hemisphere in all animals. Window implants were centered on the virus injection center, and fixed in place using cyanoacrylate glue and dental acrylic. Windows (custom ordered from Potomac

photonics) were made by placing three windows together, with the top window being 3.5mm, the bottom two being 3mm, such that the top window rested on thinned skull area, as in the protocol of Goldey et al 2014. This triple window arrangement was used to increase downward pressure on the brain and stabilize the brain motion, after use of double windows was found to lead to excessive movement. We note that the effect of such additional pressure on the brain is not well known, but appears necessary for motor cortex which anecdotally is less stable than barrel cortex.

Viruses were AAV 2/1-Flex-GCaMP6f (T.-W. Chen, Wardill, Sun, Pulver, Renninger, Baohan, Schreiter, Kerr, Orger, Jayaraman, Looger, Svoboda, & Kim, 2013a), diluted to $2*10^{12}$gc/ml. 5 injections performed in a cross-shape, centered on 1.6 lateral, 0.6 rostral. 20nL was ejected at 600um depth. See the introduction for a discussion of why this site was chosen for imaging.

Imaging was performed with a custom built two photon laser scanning microscope running scanImage software (latest versions, from 2013-2017). GCaMP6f was excited with a ti:sapphire laser, tuned to 920nm. Imaging was typically performed at 33Hz via bidirectional scanning with a resonant galvo. Power at sample did not exceed 150mW. In poorer quality windows, frame rate was halved to allow an increase in peak pulse power. This was done to minimise photodamage, as thermal effects are thought to be the major cause of photodamage(Podgorski & Ranganathan, 2016). All mice in which imaging at >350um depth was achievable were included in the analysis, with depth of recording ranging from 350um-450um, depending upon imaging clarity. This corresponds to the proximal dendritic region of the apical dendrite.

Data analysis methods

All data analysis was performed in Python using custom-written scripts unless otherwise stated.

Behavioural data analysis

From the joystick position, outward (against the spring loading) velocity was extracted. Reaches were then extracted from this as follows. Velocity signals were extracted, then z-scored and smoothed (smoothing with savgolay filter, window length 90ms, polyorder of 2). Candidate reach start and stop times were identified via thresholding of this trace, combined with the joystick's displacement, such that a reach is begun when the velocity, and ended when the joystick has returned close to the initial position and joystick velocity has shifted near zero (code for this algorithm is available upon request). The resulting candidate reaches were then filtered: reaches that failed to achieve a specified maximum outward velocity, displacement, and reaches exceeding 2 seconds in duration were discarded (shorter/longer durations did not qualitatively change conclusions).

Imaging data analysis

Imaging data was motion corrected in two stages. Firstly, an image average was taken for a session across all frames. Secondly, each frame was then then motion registered to that image, based upon a Fourier-based cross-correlation approach to detect the optimal displacement. The average was then re-taken, and the process repeated 3 times. The result of this image registration process was examined by eye for each session to check for errors.

Region of interest (ROI) extraction was done manually in imageJ software. ROIs with high baseline fluorescence, a putative marker for unhealthy cells, were not used(Packer, Russell, Dalgleish, & Häusser, 2014). Code from Bethge lab was used to deconvolve to inferred rates from the fluorescence traces (Theis et al., 2015). We note that this is not an attempt to claim specific firing rates of neurons, but rather to reduce the distorting effect of the calcium sensors' slow kinetics on the inferred activity. We did not attempt to

calibrate these inferred spike rates with real rates (in the same neuron), preferring instead to verify our core conclusions with electrophysiology.

All statistical analysis was performed in Python 2.7, using numpy and scikitlearn, with methods as described below.

Results

Mice perform a self-initiated forelimb motor task for reward

In this study, we used a behavioural task specifically designed to separate different behavioural components in time. This was an uncued joystick movement task, developed in a previous project in the Dudman lab (Panigrahi et al., 2015; Yttri & Dudman, 2016). As in other rodent motor control tasks, head-fixed mice must move a joystick to obtain a water reward, delivered through a lickport (Figure 2A,B). In this task, mice made regular, discrete joystick displacements, successfully obtaining reward (Figure 2D).

There are several features of the task that differ from prior work and place significantly different requirements upon the animal. Firstly, the reach magnitude required to obtain reward varies in a block-wise manner, as in figure 3.2C. This has the effect of producing variable velocity reaches. However, the animal is not directly cued to this threshold in any way, and only receives a reward one second *after* successfully passing the threshold. This reward delivery is marked by an auditory ‘click’ sound. The mouse therefore only finds out if its reach is successful after the reach has been completed for almost a second. This separates movement execution and reward signals. Secondly, following reward delivery, a three and a half second time out period occurs in which no reward can be obtained (Figure 3.2B). In order to not waste energy, the mice should therefore learn to initiate reaches at the correct time. Critically, the start of a new trial after this time out is not cued in any way either, and so motor planning signals can be separated from responses to external cues. In summary, the only cue the animal gets regarding the

success of its outcome, or availability of reward, is a click sound given when reward arrives. The timecourse of a trial is shown in figure 3.2B.

We note that this task does not make precise demands on the movement kinematics, which is a significant point of departure from classic approaches to studying motor control (such as those used in primates.). However, motor control in the real world rarely has a specific kinematics requirement.

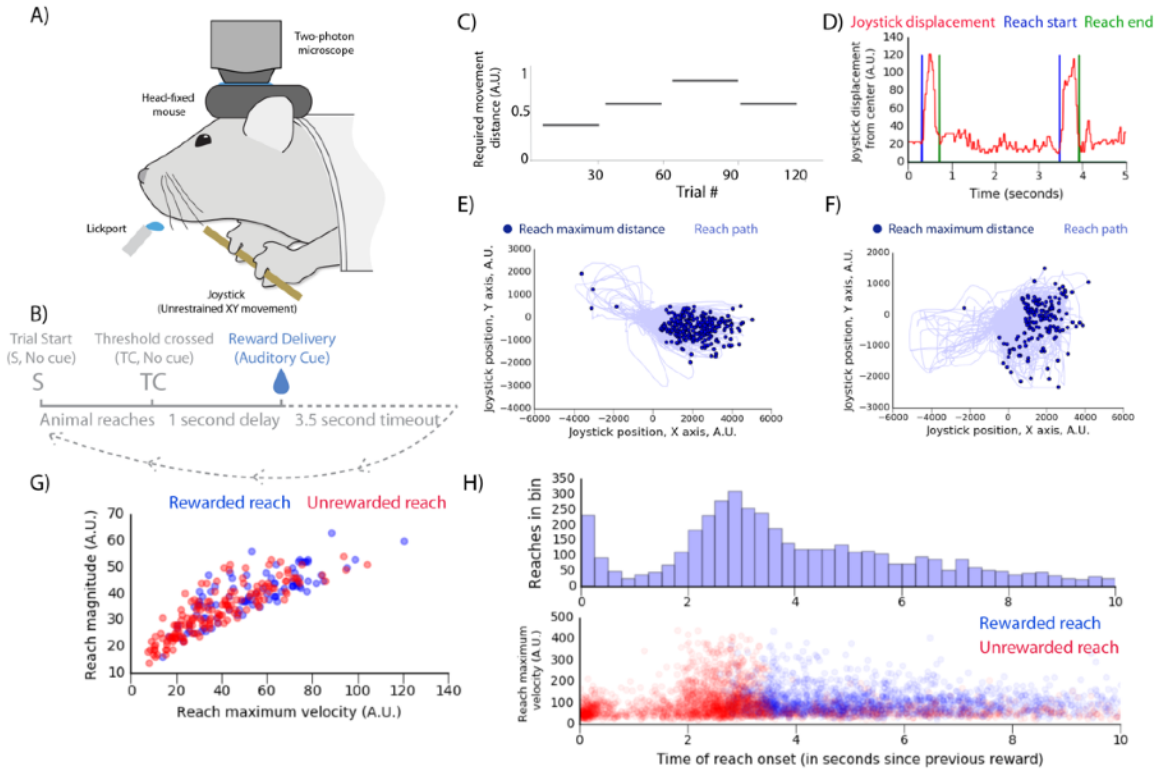


Figure 3.2. Mice perform a self-initiated forelimb motor task for reward

a) Head-fixed mice perform a joystick movement task to obtain lickport reward (image adapted from version by Josh Dudman).

b) Trial structure for the task. Mice must reach to obtain reward, with reward given 1 second after crossing the reach threshold. There is then a 3.5 second timeout until reward is available again.

c) The task consisted of 4 blocks of 30 trials, with the reach amplitude required for reward differing between blocks. The precise value of thresholds was varied over the course of the study according to individual mouse performance (for example, if mice were slow in completing training, the thresholds were lowered).
d) Example trace showing joystick position (red), and the detection of start (blue) and stop (green) of reaches by the reach detection algorithm.
e) Shows the reach end points for a single behavioural session, showing the unidirectional movement direction for this mouse.
f) Same as E., different session, showing a more scattered each end point.
g) Example session showing close coupling of reach velocity and reach magnitude. Each circle is one reach.
h) Plots showing reach start times relative to previous reward. Includes all detected reaches in dataset. Top) Histogram of reach times (250ms bins). Mice show a reduction in reach frequency following reward, but show a trend to persistently reach early. Bottom row: as in top row, but showing individual reaches, coloured by success or failure.

Mice were trained for 12-15 days, after which time discrete reaches were apparent (figure 3.2D for example trace). Though we did not reinforce any particular direction of movement, mice typically reached into a favoured region of space (example in Figure 3.2E), though some mice had more scattered movement trajectories (example in Figure 3.2F). Reach velocity and reach magnitude were closely coupled (example session in figure 3.2G), as previously reported(Panigrahi et al., 2015). Due to the varying reach threshold and the reach time out period, reaches with similar kinematic properties could be rewarded or unrewarded, allowing independent analysis of movement properties and movement outcome (figure 3.2G, 3.2H).

When examining reach timing, two notable features stood out. Firstly, some mice had a tendency to make small, low-velocity reaches coincident with reward delivery. These reaches were excluded from further analysis (Figure 2H, reaches beginning less than 1 second after previous reward were excluded). Secondly, mice frequently reached slightly prematurely, thus failing to obtain reward. This may reflect the relatively short training duration used in this study. Similar observations have been made in rats, in which the rats persistently slightly undershoot the optimal time to obtain reward (see figure 1D, (Rueda-Orozco & Robbe, 2015)). This may be an attempt of animals to increase their reward rate. We note that there is a cost to early reaches, as the joystick required substantial effort to move (and is spring loaded).

Two-photon calcium imaging reveals mixed task tuning in motor cortex outputs

To investigate cortical dynamics in this task, we used two-photon calcium imaging to record the activity of layer 5 output neurons whilst mice performed the behavior. Full details are in methods. Briefly, the calcium sensor GCaMP6f was expressed in motor cortex via injection of AAV-GCaMP6f-syn-flex in either Sim1-cre (10 mice, 19 imaging sessions) or Tlx3-cre mice (7 mice, 14 imaging sessions) (T.-W. Chen, Wardill, Sun, Pulver, Renninger, Baohan, Schreiter, Kerr, Orger, Jayaraman, Looger, Svoboda, & Kim, 2013a). These lines selectively label the layer 5 PT and IT populations respectively(Gerfen, Paletzki, & Heintz, 2013a; N. Li, Chen, Guo, Gerfen, & Svoboda, 2015c).

We will first analyze data from these two mouse lines together to identify the components of motor cortex output, and then use this genetic specificity to explore these two projections separately.

Proximal dendrites close to the cell body were imaged as mice performed the task (an example field of view is shown in figure 3.3A). Imaging of proximal dendrites was done

due to the cell bodies of the imaged neurons typically being beyond the depth of imaging for green-wavelength two (>500μm deep). Due to the slow dynamics of the calcium sensor, we deconvolved to extract an inferred spike rate (see methods and figure 3.3B) (Theis et al., 2015). See discussion for limitations of calcium imaging as a proxy for neural activity.

This experimental approach allowed comparison of neural activity with behavioural variables. The question of which recorded regions of interest (herein, units) to include in analysis was complicated by two considerations. In classic single unit electrophysiology studies, units are usually selected for modulation by behavioural parameters prior to the full behavioural session being done. This is typically done relative to an estimate of the baseline state. However, this makes the assumption that there is a discrete difference between task-active and task-inactive units, whereas in fact task tuning could be continuous. In addition, in the behavioural task used here there is no obvious baseline state, as the mice must be estimating the remaining waiting time until the next available reward period. They are thus continually engaged in a behavioural task, complicating the question of what it means to be task-modulated.

Using the entire dataset, we first aligned the mean neural population response to the two key behavioral metrics – joystick velocity and lick rate. As expected, the joystick movement occurred 1 second before reward, corresponding to the threshold crossing (Figure 3.3C, same x-axis as 3.3D), with lick rate peaking in the second following reward (note some anticipatory licking). In comparing this to neural activity, we observed a diverse range of signals in the motor cortex output, with the large majority of activity in the mean trace *following* movement (Figure 3.3D). This was apparent both in the mean trace of units, and in the time of peak activity of individual neurons (Figures 3.3D,E). There were two major peaks, one in the period between reach and reward, and another peaking a second after reward delivery. The neural activity underlying this could reflect several different processes. For example, some units peak during licking, others during

periods of little movement, prior to the next trial start. We shall explore this further below. We conclude from this that in a self-paced, un-cued behavioural task, forelimb motor cortex contains a diverse range of output signals that predominantly follow forelimb movements.

Notably, there was no clear evidence of bimodality in the degree to which a neuron was significantly modulated by the task (see figure 3.2F for example of unit significance scores). In exploratory analysis, we tried three different approaches to unit inclusion/ exclusion, and found they led to qualitatively identical results (see supplementary figure 3.1, using a PCA analysis described below). Importantly, the use of a significance threshold led to the artificial appearance of bimodality in the distribution of units in the leading principal components. Going forward, we will analyse all units in the dataset unless stated otherwise. This is a notable difference from most studies of motor cortex, which often evaluate modulation *prior* to continuing a recording.

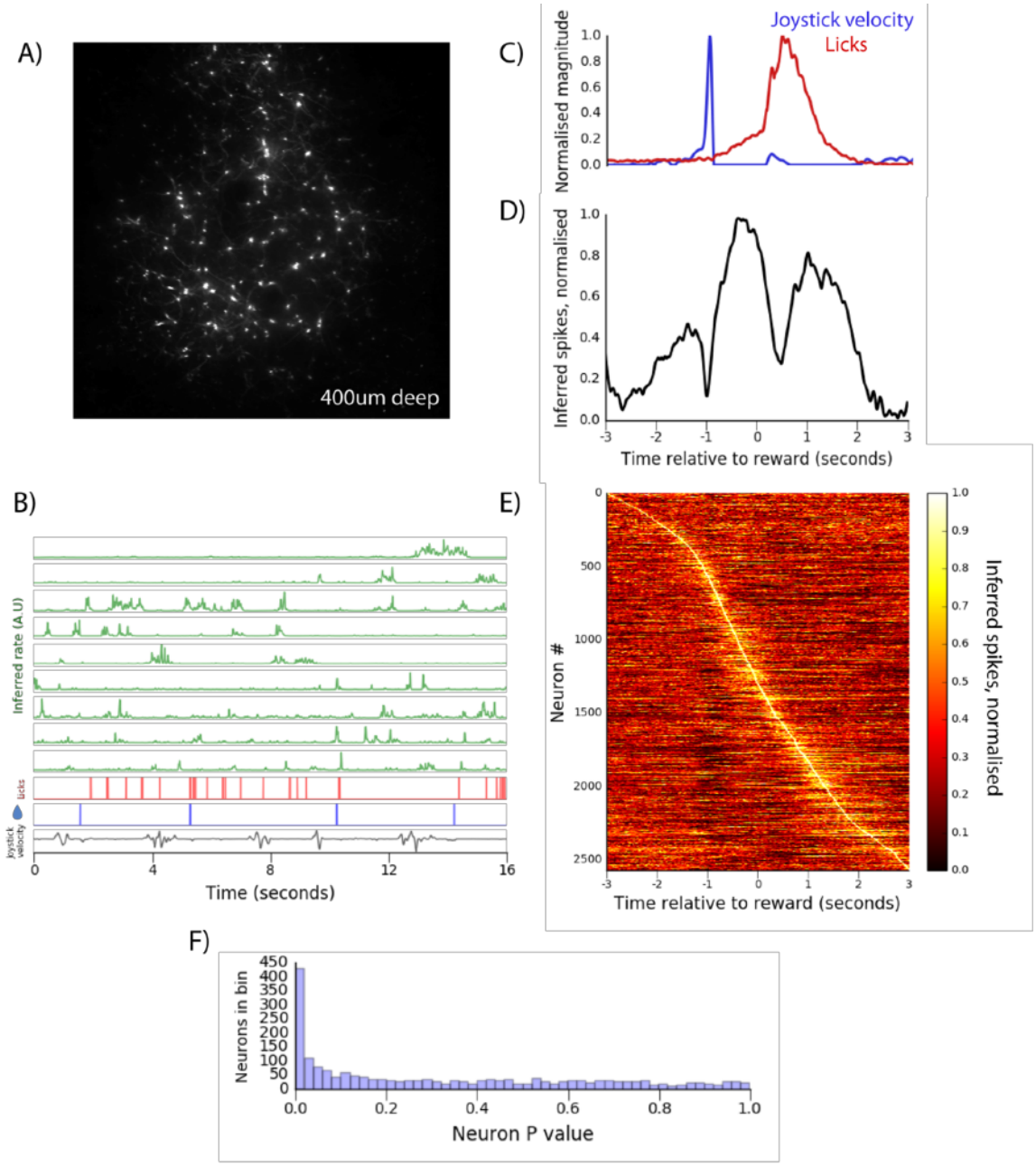


Figure 3.3. ***Two-photon calcium imaging reveals mixed task tuning in motor cortex outputs***

a) Example field of view during two-photon imaging. Image shows the mean image intensity across an entire recording session. Each fluorescent circle is a proximal dendrite.

b) Example of the data resulting from the recordings. Green rows shows inferred spike for 9 region of interest (R.O.Is). Bottom three rows show licks (red), water delivery (blue) and joystick velocity (black).
c) Average (mean) joystick velocity (thresholded at 0 to show outward components only) and lick rate across all mice across all trials, aligned to reward onset.
d) Average (mean) inferred spike rate across all units in the dataset, aligned as in C. Large majority of neural activity follows the reach and reward, rather than preceeding. The mean trace was normalized in the range 0-1 *after* averaging, so more active units contribute more strongly to the mean.
e) Heatmap showing inferred spike rate for all units in the dataset (see methods), aligned as in C,D. Each row was normalized in row 0-1.
f) Units binned by significance value (Rank sum significance test) shows lack of bimodality. Bin size = 0.02 (50 bins total).

Cell-type specific imaging reveals PT vs IT specialization

As introduced above, cortical layer 5 has two different classes of output neuron, the PT and IT classes. The bimodal nature of the mean activity and the diversity of the timing of the neuronal activity (figure 3.3D,E) raises the possibility that different classes of output neurons could exhibit different dynamics. One might expect the brain-stem projecting cell type, the PT neurons, to show the clearer peri-movement modulation, whilst those projecting exclusively within cortex and to striatum, the IT neurons, could show greater reward-related activity.

We directly assessed this using our separate PT and IT recordings, surprisingly finding the opposite result. The activity of PT and IT populations were substantially different, a difference apparent even in the mean activity trace for each population (Figure 3.4A). Peak IT neuron activity occurred significantly earlier than PT activity ($P<4.07*10^{-15}$, students independent t-test) and lacked a later component during the inter-trial interval

(Figure 3.4A,B). IT activity was closest to the reach period, with the peak in PT activity coming two seconds *after* the reach. We show the normalized traces for all units in Figures 3.4C,D. We note that a peri-movement suppression was present in the PT population closely matching the time of the excitation in the IT population (Figures 3.4 A,C,D). These results are inconsistent with an elevation in PT neuron activity being the command trigger for movement in this behavioral task.

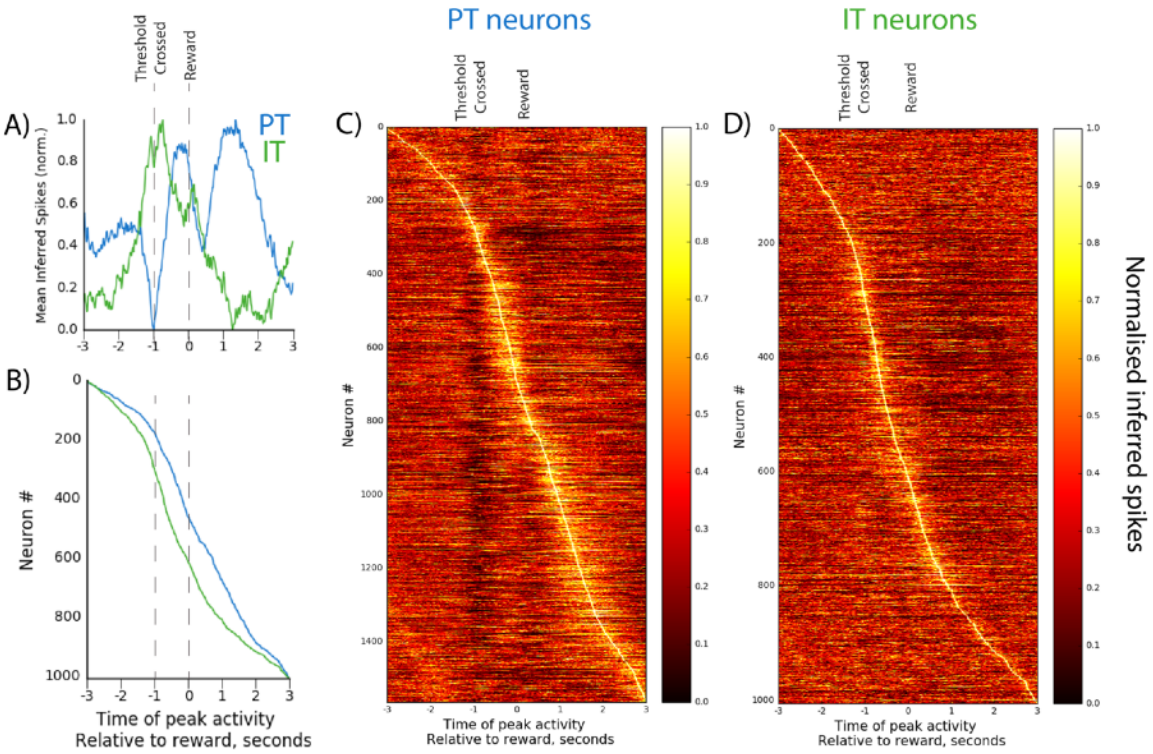


Figure 3.4 Cell-type specific imaging reveals PT vs IT specialization

a) Mean inferred spikes of PT and IT neurons, aligned to reward. Y axis normalized to 0-1 range *after* averaging, such that more active neurons influence the mean more strongly.

b) Time of peak activity for PT and IT neurons (PT neurons subsampled as in Fig3.5).

c) Heatmap showing all PT neurons in the dataset, ranked by time of peak activity. Each row shows the normalized activity (scaled 0 to 1) for each unit, aligned to reward.
d) As in d, but for IT neurons.

PT and IT differences represent the leading principal component of motor cortex output

Our results so far suggest that cell-type related differences are leading contributors to the variance in the output of layer 5. To confirm this we used the dimensionality reduction method principal components analysis (PCA). Briefly, PCA identifies new axes that explain the most linear variance. If the axis identified via unsupervised analysis are discriminative between our cell types, it implies that our cell type differences are indeed leading contributors to cortical variance. We applied PCA to a matrix containing the mean activity of each R.O.I. as rows, with time as columns (matching the sizes of our PT and IT samples, see methods for justification and approach). Prior to PCA, activity of each unit was scaled in the range 0-1 (min-max)(to show conclusions are robust to preprocessing approach, see supplementary figure 3.3, 3.4 for normalisation with z-score and no normalisation.), and data was binned into 200ms bins.

The two leading components identified are shown in figure 3.5A, and the top five components in figure 3.5B along with the percentage explained variance (we note that the fraction explained variance would be higher were we to exclude non-significantly modulated units, additional principal components shown in supplementary figure 3.2). The first component showed a peak around threshold crossing, with a suppression beginning following reward. The second principal component consisted of a prolonged suppression following movement, gradually recovering beginning around reward. Notably, different unit inclusion/exclusion criteria led to qualitatively similar results

(ssupplementary figure 3.1, 3.2 as mentioned above). Further, these conclusions were qualitatively robust to the method of normalization used prior to PCA (as stated in paragraph above).

Do PT or IT neurons contribute differently to each component? We can address this directly by examining the weightings of neurons onto the principal components. This was indeed true in the very first principal component. IT neurons had significantly more positive weightings onto this did PT neurons (Figure 3.5C, D Top row, $P<2.04*10^{-44}$, independent students t-test with equal variance.). This is consistent with the mean activity traces showing IT excitation around reach time. This was also true for the third principal component, which also had a peak at movement onset (see supplementary figure Y). (Figure 3.5 is repeated for PC3 and PC4 in supplementary figure 3.2). The first principal component of variance in layer 5 motor cortex output is thus driven at least in part by differences between genetically defined cell types, and this difference is opposite to what would have predicted based upon a 'motor command' model of PT-type neurons.

This does not necessarily identify the most discriminative dimension: IT and PT neurons could be more different in a different dimension. We examined how these differences between IT and PT are distributed across the population. Figure 3.4A shows very substantial difference between IT and PT populations. This could arise from a spectrum of possibilities. At one extreme, a small number of highly active units could differ by a large amount, at the other extreme a much smaller difference could be present across all units in the dataset. If the former, then a supervised classifier should perform poorly on classifying most cells. If the latter, it should perform well. To test this we applied the supervised learning method linear discriminant analysis (LDA), which identifies new axes designed to maximize separation between groups (As it makes little sense to classify units without task modulation, we applied a significance threshold, see figure 3.5E legend). This classified 194/256 (75.8%) of PT neurons correctly, and 200/256 (78.1%) of IT neurons correctly. This implies that the differences in tuning are relatively

distributed across active units in the population, and observed differences are not simply the product of a few strongly modulated neurons. PT and IT neurons are thus substantially different across the population.

Both for illustrative purposes and to confirm that the principal component patterns exist in individual units, we plotted the units with most positive and negative weightings onto PC1 and PC2 (Figure 3.6).

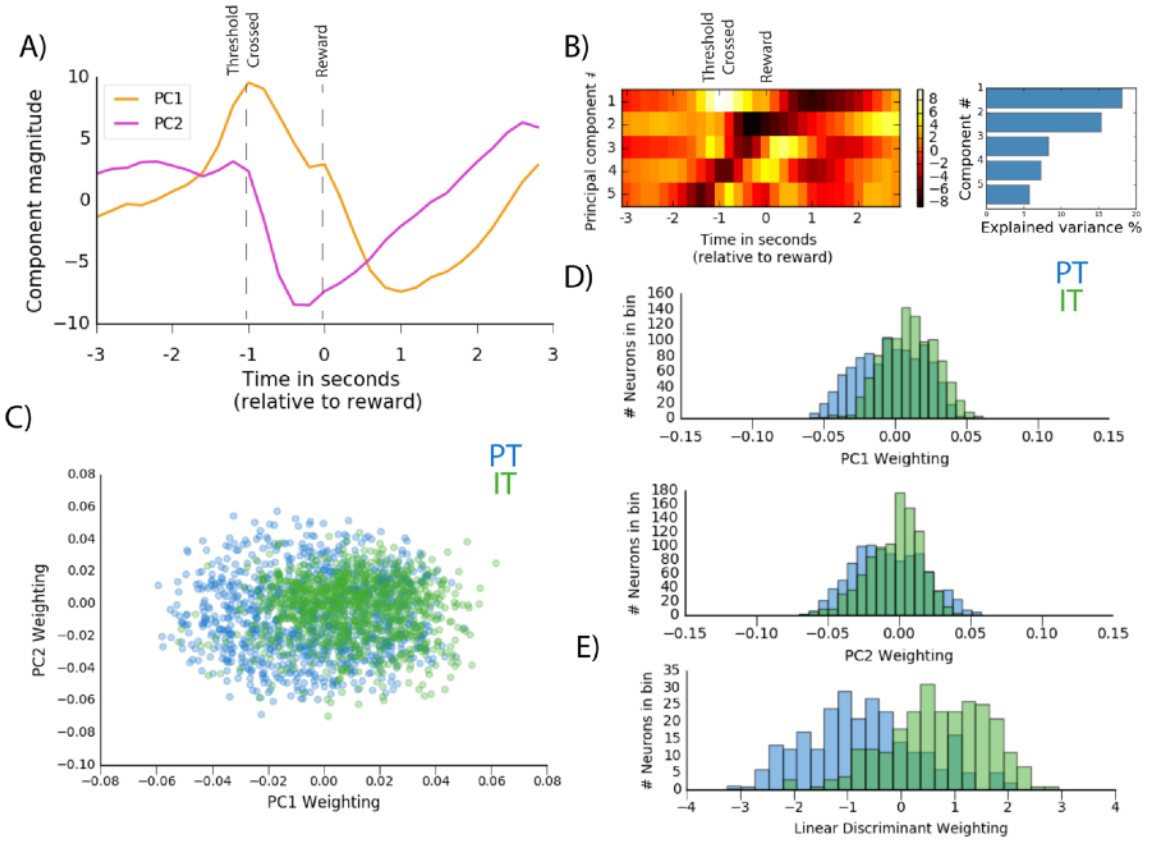


Figure 3.5 – Cell-type-related differences are the dominant components of cortical output variance

a) Time course of the first two principal components of neuronal activity, aligned to reward. Note striking similarity to 3.4A. This and all other subfigures include all units in the dataset.
b) Time course of the top 5 principal components (left), and the explained variance of each (right). Color bar in A.U.
c) Loadings of individual units onto each principal component, colored by unit subtype.
d) Histograms showing the distribution of units across the two major principal components (PC1, top, and PC2, bottom). PC1: $P<2.04*10^{-44}$, PC2: $P<2.16*10^{-5}$, independent students t-test with equal variance..
e) Histograms showing the distribution of unit weightings following linear discriminant analysis. This was done using units with a significance threshold (< 0.05 as in supplementary figure 3.1, middle row), as it made little sense to classify units with little pattern of activity. This classified 194/256 (75.8%) of PT neurons correctly, and 200/256 (78.1%) of IT neurons correctly.

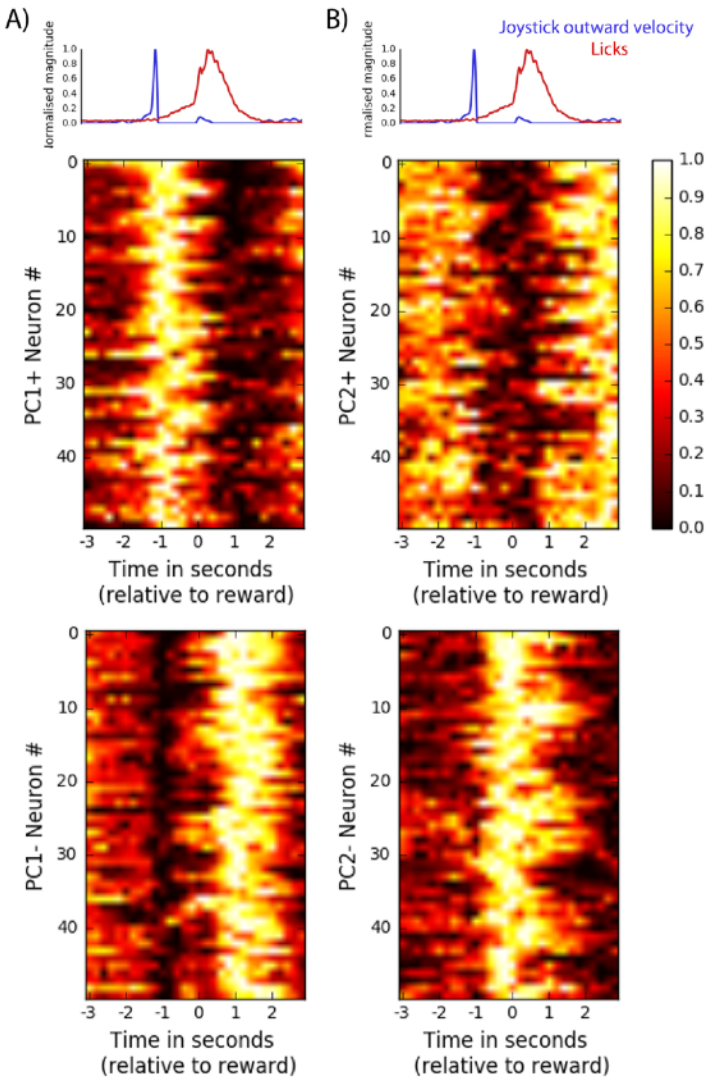


Figure 3.6 – Heatmaps showing units with most positive/negative weightings on PC1 and PC2

a) Plot showing the units with strongest positive (top) and negative (bottom) weights onto principal component 1. Shows inferred spike rate, normalized to range 0-1, as reflected in the color bar.

b) Same, but for principal component 2.

Preliminary analysis shows cell-type specific correlation of neural responses with reach velocity

Prior work in our group has shown prominent tuning of neurons in striatum to reach velocity in this task, and shown that targeted modulation of basal ganglia circuits can alter velocity of reaches in a dopamine-dependent manner(Panigrahi et al., 2015; Yttri & Dudman, 2016). Further, mice appear to modulate reach displacement by modulating reach velocity, suggesting it could be a property that their motor circuit learns to control to perform the task. Based upon this, we asked whether units in our dataset show tuning to velocity.

We found an enrichment of units significantly modulated by velocity relative to a flat distribution in both PT and IT populations(using spearmann's rank sum correlation test, comparing the sum response of the unit in the 250ms window either side of the timepoint of maximum reach velocity for a given reach)(figure 3.7). For this analysis, we only included units modulated by reaches to begin with: it makes little sense to look for velocity tuning in neurons not modulated by the behavior (done as in supplementary figure 3.1, middle row).

Further, this was significantly more prominent in the PT population than the IT population (Chi Square test, contingency value 4.13, P value<0.042). We found that PT neurons had a significantly different mean correlation value to velocity than IT neurons, with PT neurons tending to be *negatively* modulated by reach velocity (Figure 3.7D, Student's t test for independent samples, $P<1.60*10^{-5}$). This is consistent with the earlier results using principal components analysis across the entire dataset. PT neurons thus appear to be negatively modulated by increasing velocity of movement on average.

A potential concern is that the correlation is due to a different between responses to very small, potentially spurious, reaches detected by the algorithm. To check for this we more than doubled the reach exclusion criteria to select for higher velocity reaches. Using this

increased threshold, these results remained statistically significant, with the exception of IT neurons, where significantly correlated units were detected at chance level (supplementary figure 3.5). We discuss this further below.

We note that this enrichment of significantly modulated units is relatively modest. for reasons stated in the discussion, we are pursuing these observations and analysis in more detail with an electrophysiology dataset now collected but yet to be analysed in detail, as imaging is not the optimal technique for detecting such correlations. This is thus a preliminary finding.

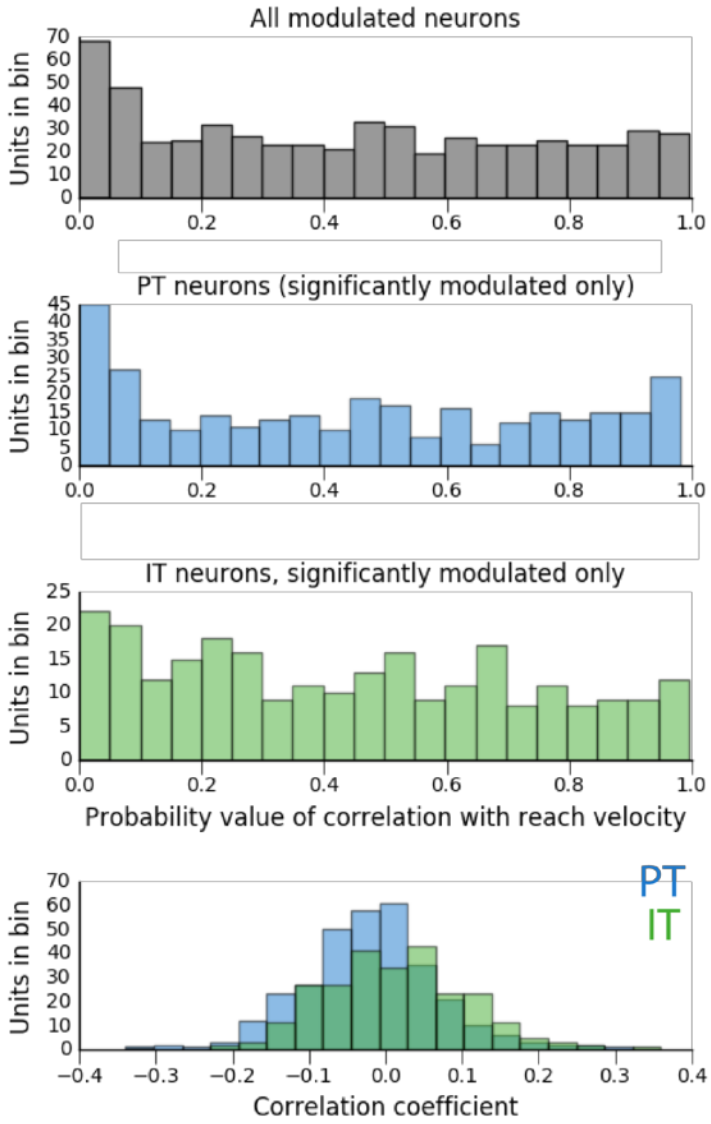


Figure 3.7 – Preliminary analysis shows cell type-specific correlation of responses with reach velocity

a) Spearmann correlation probability value for all significantly modulated units in the dataset. This is done on all reaches exceeding amplitude and velocity of 15 (A.U.). 20 bins spanning the probability range 0-1.

b) As in a), but for PT neurons only. 46/318 PT neurons were significantly correlated with velocity ($P<0.05$), with 15.9 expected by chance (using p value of 0.05).

c) As in a), but for IT neurons only. 22/256 IT neurons were significantly correlated with velocity ($P<0.05$), with 12.8 expected by chance (using p value of 0.05).
d) Comparison of the spearman correlation coefficients for each unit class. 20 bins spanning -0.28 to 0.41 (min and max values in dataset). PT neurons had a significantly more negative correlation than IT neurons (see text).

Discussion:

In this chapter we sought to advance our understanding of how motor cortex contributes to movement by recording its major output classes separately, using a behavioural task designed in part to separate key task variables in time. There are three central results. The first is that in this atypical motor control task, motor cortex activity exhibits dominant components not directly related to movement execution. The second is that this occurs in a cell-type-biased manner, opposite to that which had been previously posited. Thirdly, motor cortex output activity correlates with movement velocity in a cell-type specific manner. We now assess these points individually, before discussing future work.

Motor cortex exhibits dominant components not directly related to movement execution

Motor cortex is most commonly linked to the direct representation or generation of motor commands(Shenoy, Sahani, & Churchland, 2013b). However, studies examining this typically use stereotyped, cued movements, and preselect for neurons modulated by movement. Examining all neurons across a large dataset, we found that the first principal component showed prominent modulation throughout movement, including hundreds of milliseconds after movement. This is evidently not consistent with motor cortex existing solely to generate motor commands, as activity was present during periods of motor inactivity.

The post-movement activity is particularly intriguing. PC1 shows a prominent trough around the timing of reward and after it. This means that neurons with a negative weighting on PC1 show a peak in activity following reward, whilst being suppressed around movement execution. This is evidently not driving reach movement *per se*, but could reflect feedback activity evaluating the success or failure of a reach. Such outcome evaluation-related components of activity have been detected in human experiments, where it is linked to beta-frequency activity(H. Tan, Jenkinson, & Brown, 2014; H. Tan, Wade, & Brown, 2016). Another possibility is that the activity is involved in actively suppressing any subsequent reaches until the time at which reward becomes available again. If so, experimental inhibition of motor cortex during this time-window would be expected to trigger an early movement. This would support the view of motor cortex acting to suppress movement (Ebbesen et al., 2016; Ebbesen & Brecht, 2017).

There are two major limitations to this observation of post-movement activity. The first is that we could not adequately separate licking from reward consumption. This is because mice failed to lick when reward was not given, and when we introduced 'dummy' reward click sounds, mice quickly learnt to withhold vigorous licking to the false reward (pilot experiment, data not shown). Further, spontaneous licking was much less vigorous than that following reward, preventing clean comparison of rewarded vs unrewarded licks. The other, related issue is whether these neurons were perhaps in the wrong area: are these simply lick-related, non-reach neurons? This is possible, as reach and lick-related regions are within 1mm of each other. However, we feel this is unlikely to fully explain the observation, as this delayed-post-movement response was present to a lesser extent in the IT population recorded from the same location, and the PT post-reward activity significantly outlasted the duration of licking. Dissociating these factors will be much easier with electrophysiology, where responses to individual licks can be analysed due to the much greater temporal resolution.

The leading component of variance in output activity is aligned to genetic identity

The intent of this project was to use cell-type specific imaging to find different components of motor cortex activity in the separate output pathways of motor cortex that go to the basal ganglia and cerebellum. Prior work has proposed that different information should go to these two structures, with context-related information going to the basal ganglia to facilitate action selection, and execution-related signals going to the cerebellum to facilitate movement execution(Reiner, 2010)[1]. Such links between linking different components of excitatory neuron activity and genetic identity have been speculated before in rodent motor cortex(Zagha, Ge, & McCormick, 2015).

With cell-type specific imaging we found that that, in contrast to expectations, the neurons with strongest movement-locked activity were actually in the IT neurons that provide strongest input to striatum, whilst the pathways projecting outside the forebrain (PT) were strongly modulated at pre and post-movement timepoints. Further, this difference is reflected in the leading principal component of neural activity in motor cortex output. This raises the possibility that IT, not PT neurons could be the major driving of movement from motor cortex. This would be consistent with the surprising observation that IT-neurons more strongly innervate and drive the dopamine-1 receptor expressing medium spiny neurons in striatum, linked to the invigoration of movement(Ballion, Mallet, Bézard, Lanciego, & Gonon, 2008; Deng et al., 2015; Lei, 2004; Panigrahi et al., 2015). Importantly, these results are also not consistent with a simple 'feedforward' circuit in which PT neurons gain their response properties by being driven by IT neurons(N. Li, Chen, Guo, Gerfen, & Svoboda, 2015b). We further note that this relative lack of positive modulation of PT neurons has been observed in a recent

[1] "These two neuron types, however, differ genetically, morphologically, and functionally, with IT-type neurons conveying sensory and motor planning information to striatum and PT-type neurons conveying an efference copy of motor commands (for motor cortex at least)."(Reiner, 2010)

study using similar methodology to ours, though that study did not image IT neurons also(Peters et al., 2017).

Why do other studies not show such clear separation of IT and PT neuron profiles, showing instead only a directional tuning bias between the two groups(N. Li, Chen, Guo, Gerfen, & Svoboda, 2015b; Soma et al., 2017)? The most likely difference would appear to be the tasks since, as mentioned above, the task we use is almost entirely lacking in external cues, relying instead on inference of the optimal timing distribution of when to make reaches. This task also does not make precise requirements of the animal with respect to the movement executed: it is not a classic motor control task in this sense. It is possible that a task requiring online control more strongly might elicit a greater PT contribution. Another potential difference is that in this study we solely used genetic labelling, and it may be that the Sim1 and Tlx3 lines are biased toward particular subsets of the IT and PT neurons (few, if any transgenic lines are fully penetrant due to variations in cre recombinase expression levels). The other studies used antidromic collision to identify their neurons. However, this is unlikely to explain the difference, as the Li *et al* study also uses the transgenic lines we used here for imaging, and they qualitatively match their electrophysiology result. It is however an important limitation to note that we are not necessarily sampling all PT and IT neurons equally. Finally, we note that though this site in motor cortex was chosen based upon several prior studies (see introduction + methods), it is possible that in other areas in motor cortex may show quite different results, and there may be a 'PT' hotspot in which PT neurons are much more strongly locked to motor execution.

Cell-type specific reach velocity tuning in motor cortex outputs

Here we showed a cell-type specific correlation with velocity, with PT neurons preferentially correlating with velocity. However, notably this was predominantly a negative correlation: the neurons projecting out of the forebrain seem to be relatively *less*

active on larger movements. These correlations were relatively sparse in the population, only modestly higher than the expected false discovery rate. This is notably lower than in the prior striatal recordings. This may reflect genuine physiology, but there are reasons to suspect that this is a significant underestimate of the true population coding.

We used imaging rather than electrophysiology, due to the higher unit yield and relative ease of performing cell-type specific recordings. This is a less direct detector of neuronal activity. Calcium imaging using GCaMP is a nonlinear read out of a nonlinear signal, and existing deconvolution algorithms still perform relatively poorly. This is true even with imaging in superficial neurons, whilst our imaging was near the depth limits of two photon microscopy. Another possibility is that proximal dendrites do not accurately reflect the activity in the cell body. However, a prior study comparing event detection in soma versus proximal dendrites substantially more superficial than ours found very strong correlation between the two signals, suggesting proximal dendritic imaging is indeed a good proxy for cell bodies(Peters et al., 2017). The detection of a cell-type specific enrichment in velocity tuning is a promising result, but is, for now, preliminary. A further concern is that the reward-related activity increase could be related to calcium increases driven by reward-related neuromodulators: ie, reward could alter the relationship between calcium and neural activity.

Future work

The goal of this study was not to identify the correlates of neurons with specific movement parameters, but rather identify dominant components of the cortex outputs across the whole task to allow us to later investigate the mechanisms of such components. Prior work in our group using this task has shown correlation of neural responses in striatum with velocity(Panigrahi et al., 2015; Yttri & Dudman, 2016), and begun unravelling the mechanisms at the mechanistic level in the striatal circuitry. This chapter, combined with the previous one, will allow us to take this mechanistic

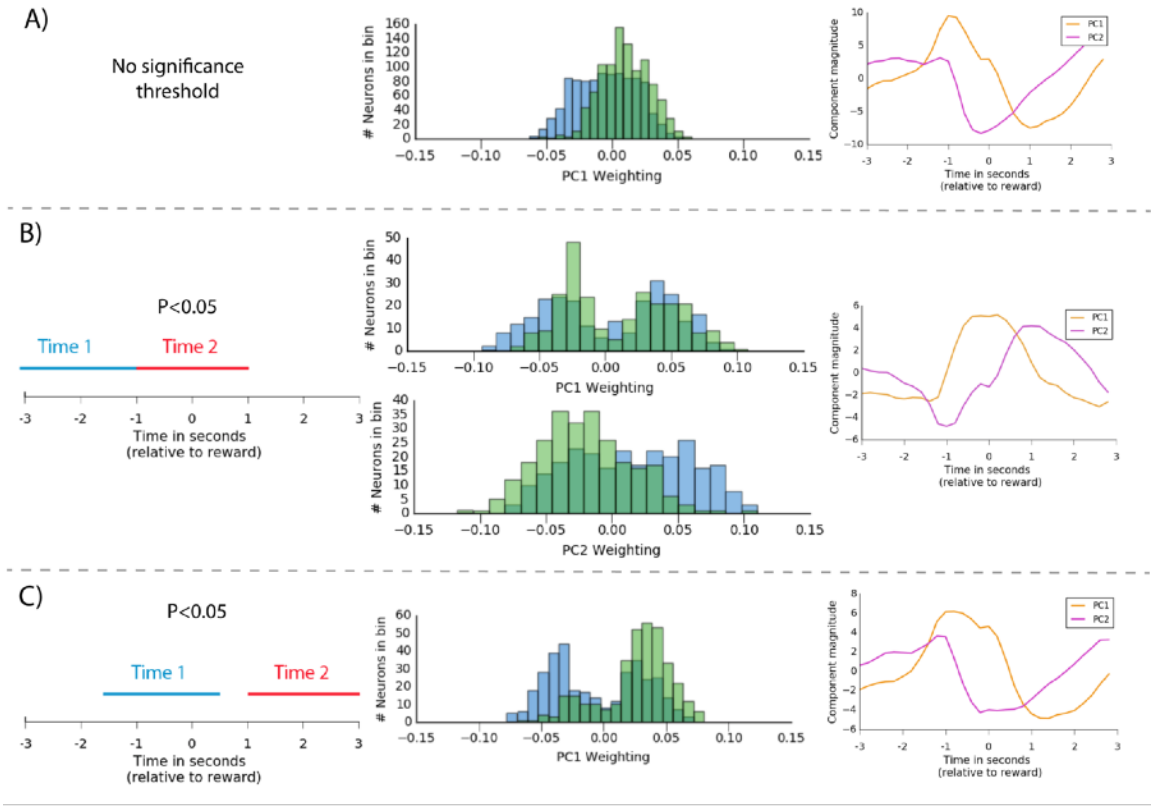
investigation to the circuits level, investigating how the different elements of the motor system interact to alter velocity signals.

There are two lines of experiment to this. The first is to perform electrophysiological recordings from identified PT and IT neurons. These experiments are complete, but await in depth analysis. In the course of these experiments we also inactivated PT and IT neurons at specific phases of the behaviour to explore the cell type's contribution to behaviour. The second is to use transgenic lines developed from chapter 2's data to target the different thalamic inputs to cortex, and identify whether, for example, basal ganglia recipient thalamus selectively modulates velocity signals in cortex. We are currently having difficulties reaching high viral expression in the cerebellar thalamus, potentially due to viral tropism issues.

Conclusions

Together, the data suggest that motor cortex contributes diverse signals to subcortical motor signals, in a pathway specific manner. We show, using cell-type specific imaging, that motor cortex sends diverse signals to subcortical circuits, including both during and after reaches. This is cell-type specific, with PT neurons predominantly suppressed by movement and IT neurons predominantly excited, in contrast to expectations. Finally, we show tentative evidence of a specificity of velocity tuning across these two cell classes. Current work is combining electrophysiology with targeted perturbations of the cortico-basal-ganglia-thalamic loop to reveal the circuit mechanisms of such cell-type specific dynamics.

Chapter 3 Supplementary Materials



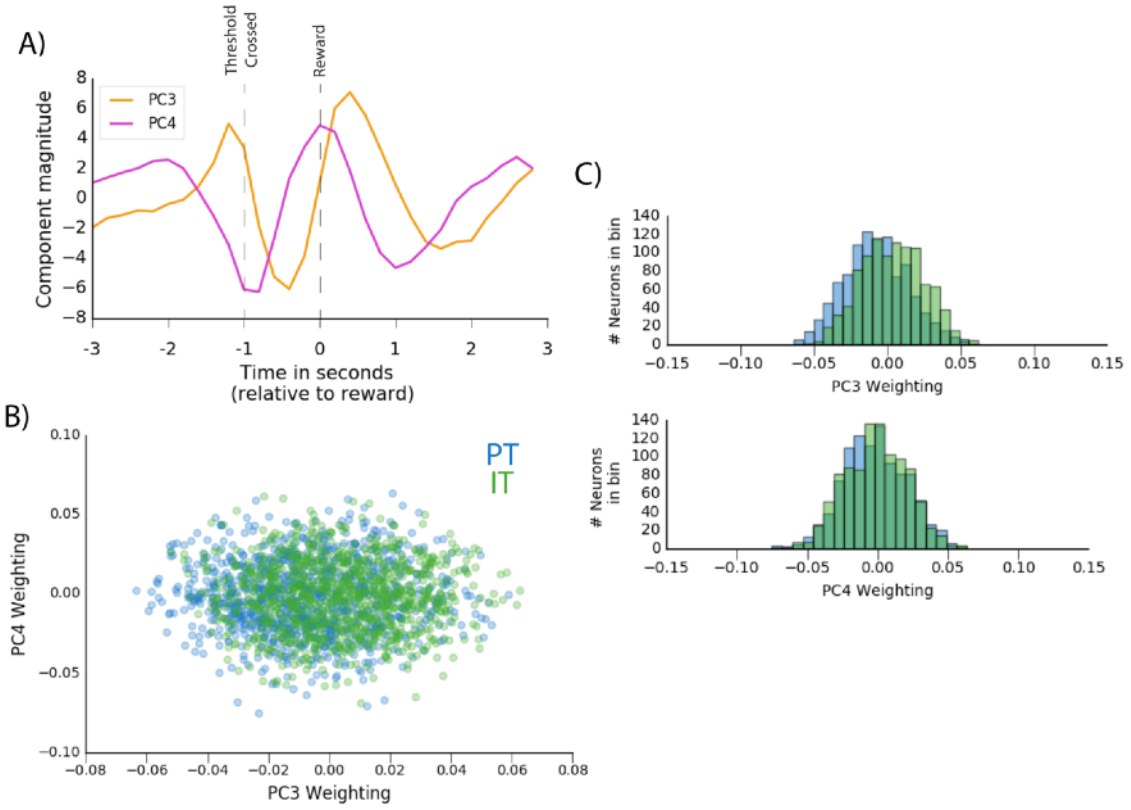
Supplementary Figure 3.1 – Composition and shape of leading components is robust to unit selection approach

Significance thresholding leads to artificial appearance of bimodality in leading principal component.

All analysis here was done normalizing individual traces in the range 0-1, similar results obtained with other preprocessing steps.

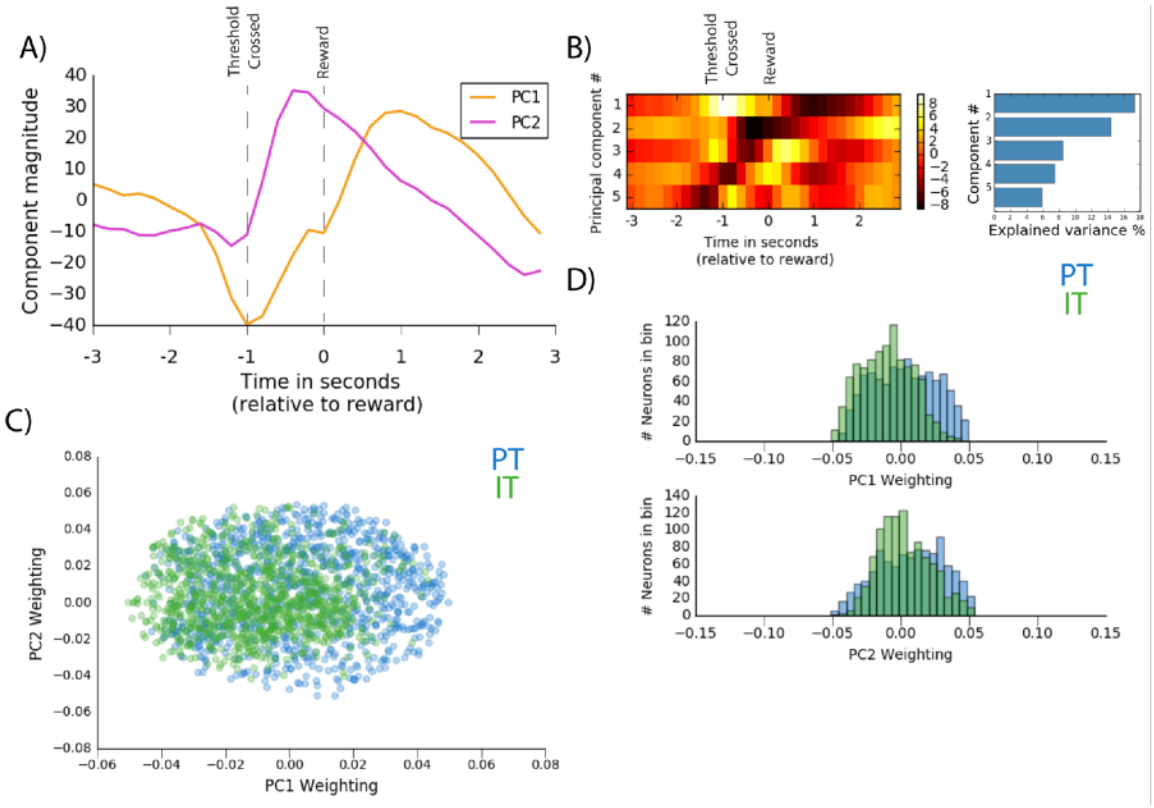
For all neural activity analysis following, units were filtered by a significance test, with units included if P<0.05, wilcoxon rank sum test, no correction for multiple comparisons. As there was no clear baseline in this task, we simply compared the mean activity in a period at reach onset up until 2 seconds post-reward to mean activity in a period 2 prior

to reach onset. Results were not sensitive to the precise metric used, nor to the significance threshold used. Using this metric, 36.6% of recorded units were significantly modulated during the task (637/1566 PT neurons, 304/1006 IT neurons). Note the flipping of the sign simply reflects flipping of unit weights (ie, PC2+PC1).



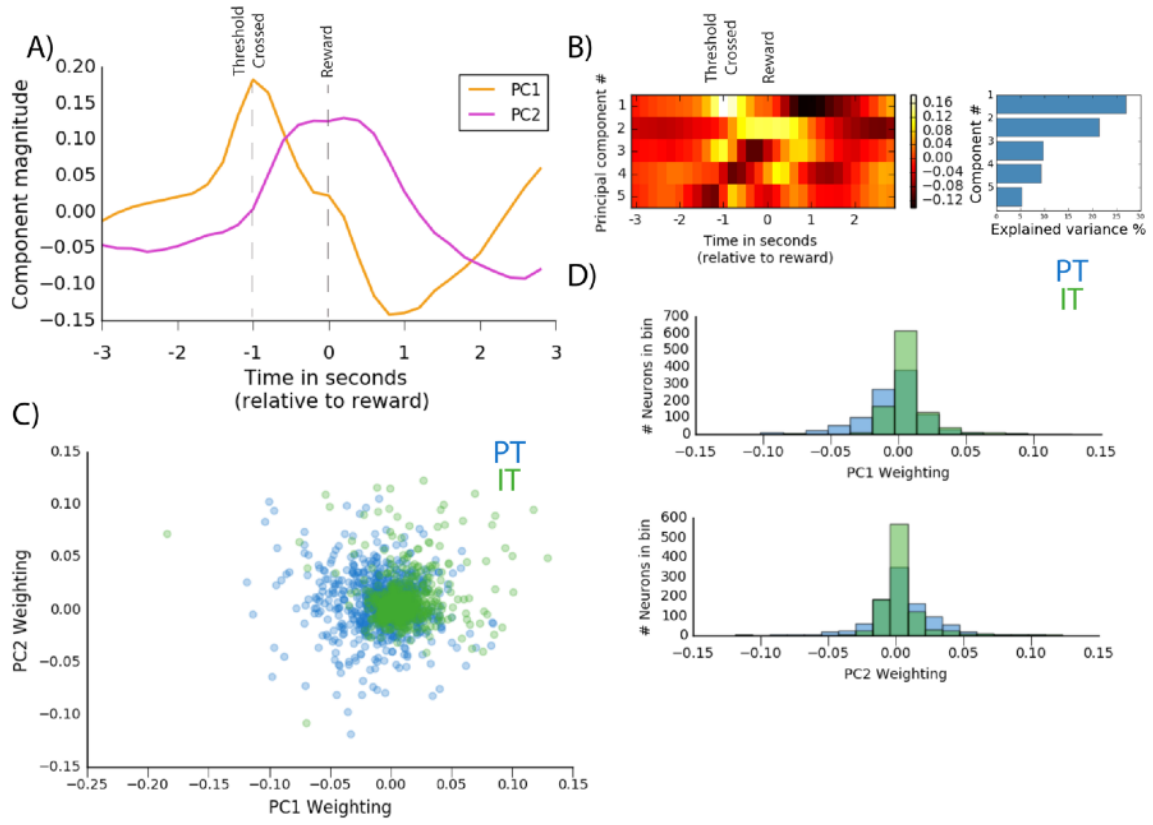
Supplementary Figure 3.2 – Additional principal components

As in figure 3.5, but for the third and fourth principal component. Loadings on PC3 were significantly different between IT and PT (student's independent t-test, $P<1.28*10^{-29}$). Non significant for PC4 (Student's independent t-test, $P>0.5$).



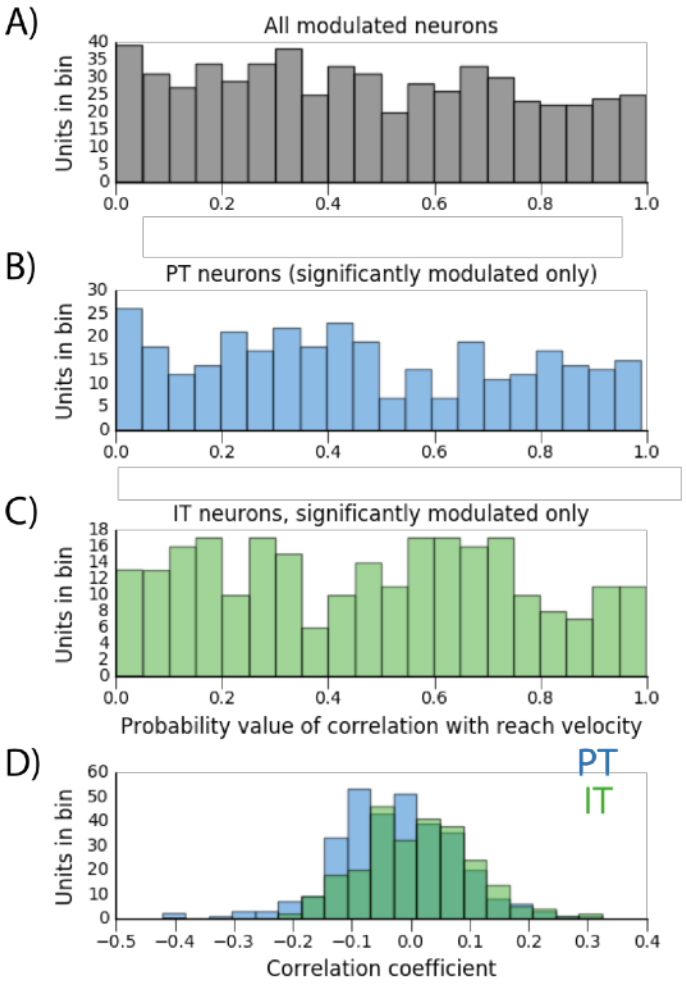
Supplementary Figure 3.3 – Preprocessing methods and their effect upon resulting principal components (Z-score)

This figure repeats the analysis of figure 3.5, but using a z-score normalisation of unit activity instead of scaling 0-1.



Supplementary Figure 3.4– Preprocessing methods and their effect upon resulting principal components (No normalization)

This figure repeats the analysis of figure 3.5, without using normalisation of unit activity.



Supplementary Figure 3.5– Reach correlations with more stringent reach exclusion criteria

a) Spearmann correlation probability value for all significantly modulated units in the dataset. This is done on all reaches exceeding amplitude and velocity of 40 (A.U.). 20 bins spanning the probability range 0-1.

b) As in a), but for PT neurons only. 26/318 PT neurons were significantly correlated with velocity ($P<0.05$), with 15.9 expected by chance.

c) As in a), but for IT neurons only. 13/256 IT neurons were significantly correlated with velocity ($P<0.05$), with 12.8 expected by chance.

d) Comparison of the spearman correlation coefficients for each unit class. 20 bins spanning -0.28 to 0.41 (min and max values in dataset). PT neurons had a

significantly more negative correlation than IT neurons (see text). (Independent t-test, $P<9.9*10^{-9}$).

Chapter 4 – Discussion

Graded lines: Examining continuous variability to reveal biological organization

In this thesis I used genetic tools to investigate the interactions between different brain areas, with a particular focus upon the mammalian motor control system. The first experimental chapter used genetics to examine the static structure of neural circuits: how the mammalian thalamus is organized. The second experimental chapter used genetic techniques to isolate the different components of the dynamic organization: the activity of the layer 5 outputs of motor cortex during behaviour. I have dealt with chapter specific comments and issues in discussion sections of each chapter.

Though the two chapters have significantly different focuses, there is a common theme that emerges and that changed my view of how gene expression can be used to understand the brains organization and operation. The goal of this discussion is to highlight this. This theme relates to the notion of a cell type, what we mean by a cell type, and how it is and isn't useful for understanding nervous system function. Though much of this chapter may sound more fitting for an introductory chapter, I only began seriously thinking about these questions once I had already obtained results. Further, this discussion is quite orthogonal to the original intent of the thesis. They are therefore more fitting for discussion, as a reflection on how this thesis has changed my view of the nervous system.

Lumpers, Splitters….. or connoisseurs of variance?

When I began this thesis, I approached the problem of understanding long-range motor circuits from the dominant perspective of genetically defined, discrete cell types, each with their own marker genes and their own unique connectivity profile. One can view this as a kind of 'labelled line' approach, with distinct types of information being carried by distinct pathways. However, in the process of performing the project, I have come to question the utility of this, and have become frustrated in the way in which genetic analysis is moving. I am going to explore these points now using the results within this thesis.

The notion of a cell type dates back to the origin of cell theory: the theory that all living organisms are composed of cells(turner, 1890). Cell theory has undergone revisions and refinements since Hooke first used the term 'Cell', but broadly within biology there is an attempt to link specializations in specific, discrete cell types to specific, discrete functions. A rigorous, exact definition of a cell type in neuroscience is lacking (and perhaps not as useful as often supposed), but, again, broadly it is meant to describe a group of cells with discrete, relatively homogenous features in higher dimensional space(Nelson et al., 2006). In this way, there have been attempts to link discrete cell types to specific circuit functions.

Cajal famously used diversity in neural features (in his case, morphology) to identify discrete cell classes, but the approach both pre-dates him and continues to this day(Kandel et al., 2012). The logic is simple: one searches for features present in one cell group but not another, and uses these to define discrete categories of cells. These, in turn, are investigated and linked to functions. In short, discrete categorization allows a taxonomy of cell types to be formed, and these cell types can then be allocated their own role.

Cell typists and geneticists in general have recently been given powerful new tools, in particular next generation RNAseq(Poulin, Tasic, Hjerling-Leffler, Trimarchi, &

Awatramani, 2016). This has revealed additional heterogeneity in neuronal systems, including identification of novel marker genes and indeed novel cell types(Tasic et al., 2016; Zeisel et al., 2018). The technology is receiving intense attention from disparate fields of neuroscience(Cembrowski, Bachman, Wang, Sugino, Shields, & Spruston, 2016b; Hrvatin et al., 2017; Phillips, Schulmann, Hara, Liu, Wang, Shields, Korff, Lemire, Dudman, Nelson, & Hantman, 2018b).

However, there is a debate in biology between the lumpers and the splitters, relating to how to interpret and group observations, that is important to consider. The debate concerns at which to divide or unify observations: Splitters seek to find ever finer categories, arguing that the devil is in the detail. They, relatively speaking, prefer to divide observations into further categories. Lumpers, contrastingly, do the opposite, preferring to keep categories larger and more heterogenous. At the moment, the splitters appear to be winning in the RNAseq world, as we identify ever finer types of cell type. This in in part inevitable: the addition of new tools allows people to make new discoveries, the most straightforward of which is to define a new cell type. However, categories are ultimately somewhat arbitrary: though criteria exist for determining the number of clusters present in data, the decision at which point to define two separate clusters must be chosen by the analyst, or at least the parameter tweaked.

This discussion is intended to highlight the benefit of, and argue for, adopting a complementary approach, one that seeks to instead understand patterns of variance to reveal biological structure. In short, we should be neither lumpers nor splitters, but rather, connoisseurs of variance. In this way we are consciously trying to move away from discrete categorizing, and recognize that categories are only useful insofar as they produce understanding: we should not define cell types for cell type's sake. After all, in reality nature (and evolution) does not care about words such as 'cell type'. It instead cares about how things vary, and about how to make them vary usefully. In trying to understand nature, then, we should do the same.

This discussion seeks to highlight the benefits of such energetic connoisseurship of variance. There are several steps to this.

Within-'cell type' variability is high

The first relevant point arising from this thesis is that the variance of a given feature within a 'cell-type', with 'cell-type' defined either by transgenic line or by marker gene expression, is remarkably high in this study. This was true both in the RNAseq-centered chapter 2 and in the functional imaging-centered chapter 3. In both cases, the value of a feature that is typical of a cell type extends essentially continuously from a maximum value to essentially zero. The features were clearly non-binary. This makes the notion of a discrete cell type somewhat fuzzy already, since the cells within a cell type vary a great deal. Lets consider this first in the instance of genetics, and then in the case of understanding neuronal 'tuning' to task features.

In the context of genetics, most studies are focussed upon using clustering algorithms to identify discrete cell types: a general pipeline of clustering neurons into discrete types, then identifying unique markers for each cluster, is performed. Indeed, I am not aware of any RNAseq study emphasizing quantification of the difference in magnitude between *within* class variance and *between* class variance(though see (Cembrowski & Menon, 2018), and below, for discussion). The emphasis is entirely focused upon quantifying the 'number' of cell types, as though this were an end in itself.

In the case of the RNAseq chapter of this thesis, quantification of the 'marker' genes for primary, secondary and tertiary profiles reveals expression ranging from high expression all the way to zero, without an obvious discrete cut off (figure 2.10). For example, two thalamic neurons of 'primary' type could vary in their expression of 'marker gene' Tnnt1 by an order of magnitude. Indeed, for 'secondary' nuclei there is no identified perfect

marker: Necab1 is also expressed in tertiary nuclei substantially. Further, two cells of *different* types could actually be closer to one another in high dimensional feature space than two cells of the same 'type', a point that we explore below (see figure 2.8C). Not only does this obscure the utility of using such features as category markers, it means that a large fraction of variance is unexplained simply by devising a taxonomy of discrete cell types. An explanation of neuronal diversity based solely upon discrete cell types will therefore be necessarily limited.

A similar attempt to discretize and categorize neural variation exists in the realm of neural activity analysis. Studies typically class cells by which stimuli or actions they respond to, based upon a statistical significance cut off[2]. For example, in examining motor thalamus neurons Gaidica *et al* find 'functionally distinct neuronal populations in Mthal whose activity changes briefly around movement onset'(Gaidica, Hurst, Cyr, & Leventhal, 2018). However, in this case, and typically in others, there is no test of distribution to be made, to check for genuine discreteness (For example, Hattox & Nelson 2007, the title of which states discrete cell features without plotting individual data points, only averages).

Now, with the collection of ever larger neural populations, one is able to estimate the underlying population more fully and more accurately, and this has led to claims that there is actually 'non-categorical' encoding of variables. For example, Raposo *et al* claim in parietal cortex of rodents state that '*Newly designed tests revealed that task parameters and temporal response features were distributed randomly across neurons, without evidence of categories'*(Raposo, Kaufman, & Churchland, 2014). This finding was deemed surprising enough at the time to form the title of their paper: "A category-free neural population supports evolving demands during decision-making'.

Our data in large part agree with this claim. In chapter 3 we found that the contributions of neurons to each principal component varied essentially continuously. In the case of the functional imaging, the principal component weightings are both overlapping for PT and IT neurons, and extend from maximum to minimum. In none of the analysis in chapter 3 does a clear 'discrete' grouping of tuned cells emerge. Rather, a continuum is present. Clearly, then, 'cell type' alone is not enough to explain why some neurons end up with a given response profile. This strikes at the limitations of the genetics approach, which often implicitly seems to map each individual cell type onto individual components in the data.

There is, however, an intriguing disagreement between our data and the conclusions of Raposo *et al*. Raposo sate that "*temporal response features were distributed randomly across neurons".* In this thesis, with the addition of genetic specificity, ie, cre-targeted reporters, shows that different aspects of the neural variation are accounted for by different parts of the population (Figure 3.5,3.7). So, we can agree with the non-categorical part, but not with the random distribution. This presents an exciting motivation to revisit an existing result with genetic technology, as well as others that find intriguingly different components of neural activity without investigating their basis in cell types(Zagha et al., 2015).

Together, the results of the two experimental chapters illustrate that neural features vary substantially within classic 'cell types'. Further, this variation is often not obviously discrete, at least in the leading dimensions of variance. This is more consistent with a gradation of labelled lines, rather than discrete, cleanly separation contributions for each pathway.

Spectrum-like variation of cell types*: *categorization may miss the most relevant variance

Viewing neuronal cell types not as point-like, homogenous entities, but rather as spectrum-like in their variance, opens several topics for inquiry. As it is most typically used to define cell types, we will now focus primarily upon genetic variance.

The first is to consider whether this 'within class' genetic variability is functionally relevant, and is ignoring it therefore impeding understanding? This is difficult to directly answer. However, if within class variability *does* exceed between class variability, it follows that it will likely be functionally relevant if the between class variability itself is assumed to be of functional relevance[3]. Since even between class variability of features is difficult to link conclusively to differences in function with current techniques, this point will await further experimental investigation.

However, there are intriguing possibilities here for further investigation. One could imagine a coupling of within-class gene expression variability with within-class axonal morphology, for instance, or that within-class variation of gene expression could be linked to variation in the electrophysiological properties. The resulting heterogeneity could provide a spectrum of electrophysiological properties, allowing generation of different resonant frequencies in neuronal populations. Such relatively continuous distribution of neural activity frequencies is indeed seen in nervous systems, with different peaks becoming more prominent at different times(Buzsaki, 2006). Notably, most theoretical models lack this heterogeneity, though, some theorists are beginning to explore the potential benefits of such heterogeneity on population dynamics (Landau,

[3] Another possibility is that within-class variability occurs via different classes of genes, which may be of less functional relevance. This will likely at least in part be true, as within-class variability will in part reflect variability in cell-health related genes, for example.

Egger, Dercksen, Oberlaender, & Sompolinsky, 2016; Tripathy, Padmanabhan, Gerkin, & Urban, 2013).

New techniques for combining observations from multiple experimental modalities, such as Patch-Seq, could investigate this(Muñoz-Manchado, Gonzales, Zeisel, Munguba, Bekkouche, Skene, Lönnerberg, Ryge, Harris, Linnarsson, & Hjerling-Leffler, 2018a). Currently, however, such techniques are often more noisy than when measurements are made in isolation, which could obscure correlation. Improved *in situ* sequencing could help address this(Xiao Wang et al., 2018). In addition, improved tools for predicting neural features from gene expression data could facilitate investigations such as this by making hypothesis more precise ahead of testing.

In short, within class variability in gene expression could be analysed to reveal novel variability in properties directly relevant for neural function, such as electrophysiological or morphological properties. Therefore, it should not be so readily discarded as noise.

There is a second way in which this spectrum-like conceptualization of cell classes leads us to reconsider our understanding of neuronal variability. Had we simply defined three major cell groups and identified marker genes in chapter 2, we would have missed a central result: that there is a sequential shift from primary, to secondary, to tertiary classes in genetic distance and in neuronal features across thalamus. We revealed this by using principal components analysis in addition to more typical clustering methods used in RNAseq analysis. Analysis of this kind is not often performed for investigating neuronal variability. More often, hierarchical clustering or package-specific algorithms such as those use in chapter 2 are used to define cell types(Tasic et al., 2016). However, such a sequential relationship is only indirectly observable in the hierarchical clustering analysis typically used to group cells, and non-linear methods such as tSNE will not reveal it either. Rather, one must use linear analysis techniques to extract such variance patterns. By examining patterns of linear variance, we revealed a biologically relevant

pattern of variability with close links to neural function. Further, there were intermediate neurons along this axis: rather than three distinct groups of cells, we revealed a spectrum of neural variability, with one of the groups intermediate between the other type. Similar analyses have recently revealed similar such arrangements in other groups of neurons(Muñoz-Manchado, Gonzales, Zeisel, Munguba, Bekkouche, Skene, Lönnerberg, Ryge, Harris, Linnarsson, & Hjerling-Leffler, 2018b). In short, exploring the relationsship between groups of cells, not simply identifying the number of such groups, should be prioritised.

Approaches for further investigating patterns of variance in genetic data

The focus of this discussion has, thus far, been in highlighting the within-category variance in data, primarily in genetics data, and how this has been used in this thesis to reveal organizational principles in thalamus. This all centers around the notion of continuous variation of cell types. This emerging topic has recently been excellently reviewed by my Janelia colleagues Vilas Menon and Mark Cembrowski(Cembrowski & Menon, 2018). Here I want to briefly outline topics of further research that are motivated by this view of gene expression space and that are not examined in detail by their review.

The first is that a continua of cell types is not readily explained by the 'standard model' of how cell types are fixed: by attractor-like, competitive interaction between transcription factors (such as Hox genes defining discrete segments during development). How a non-binary gene expression difference is setup is not well explained by existing models of gene expression, revealing a notable gap in our understanding of the genetic basis of cell types. Further, such continuous variability of cell types raises the possibility of dynamic cell types, in which the specific position along the continua could vary over time in an adult animal. For example, even 'marker' genes could come on and off in accordance with changing activity patterns in the brain over time. Indeed, major differences in gene expression in response to visual stimulus exposure have recently been

shown(Hrvatin et al., 2017), but this again is a relatively nascent area of investigation. We note that it is of course unlikely that a neuron's axonal morphology varies dramatically in the adult animal: existing axonal imaging studies show relative stabilization after the early postnatal period(Portera-Cailliau, Weimer, De Paola, Caroni, & Svoboda, 2005). However, less dramatic shifts in neuronal features could potentially occur. In this way, the relentless march from pluripotency to a fixed, lifeless state currently assumed may not be fully accurate.

Further, a continua of cell types of course raises the question of whether 'type' is the right term to be using in the first place. In chapter 2 we veered away from using the word 'type' and 'cluster', instead employing the less discrete term 'profile'. This change in language, adopted more broadly, might help to instill a different of attitude in mind. It would lead people to ask not how many 'types' there are, but rather what factors make up the genetic composition of a cell. We could term these 'cassettes', different aspects of the gene expression profile that are on or off to define a given gene profile.

What might such an understanding this look like? There is an immediate, obvious possibility: that the variance pattern within cell types is the same for different cell types. Ie, PT and IT neurons could have the same spectrum of ion channel features within them, for instance. A cell would therefore have a 'cassette' defining its IT or PT nature, and another defining where along this ion channel spectrum a cell should lie.

In the context of thalamus, a cell in VA might have a 'cassette' defining it as a thalamic neuron, a 'cassette' defining it as a motor-modality neuron, and another defining it as being a 'secondary' type neuron. The goal of genetics analysis, then, would be to identify the number of 'cassettes' (analogous to 'factor') driving a given cell profile, and what influences them. This, again, to me would represent an understanding of gene expression that is much more fundamental than simple identification of marker genes.

Further, one could envisage these ‘cassettes’ might be shared across brain regions, such that there could be equivalents of the spectrum of neurons identified in thalamus elsewhere in the brain. This would explain why markers of cell types in one area so often also define cell types in other areas (for example, Rorb, or Cblln2). We lack direct evidence of this at present, but there is an intriguing clue that suggests such a possibility.

I have long been intrigued as to why calcium binding proteins (CBPs) are such excellent markers of cell types (I use this word begrudgingly) around the nervous system. From cerebral cortex down to the brainstem, from mammals to reptiles, a relatively small number of calcium buffers (typically parvalbumin, calbindin1 and calretinin) divide neural groups in disparate brain areas and species. Though calcium is linked to diverse cellular processes, why a cell would express one CBP instead of another is not understood.

What, then, does this CBP expression variation correspond to and correlate with? Is there a gene expression ‘cassette’ corresponding to parvalbumin expression, for example, and if so, what other features of a neuron does it correlate with? Parvalbumin interneurons in cerebral cortex, hippocampus and striatum all have relatively local axonal arborizations and specializations for fast firing rates: this is intriguing similar to the axonal and electrophysiological features in the ‘primary’ group in our data. Examining this across other areas of the nervous system should be technically straightforward, though laborious. A simpler test would be to look for genetic covariates of these calcium binding proteins brain-wide, and then seek to understand what they are driving. This is on my to-do list (and has been for far too long).

Summary

In summary, this thesis sought to use advances in genetics to investigate the pathways connecting different brain regions. In doing this, a common theme has been the relatively continuous variation of features in neural populations investigated. This was true in both chapters, though especially prominent and consequential in the RNAseq chapter.

In this discussion I have argued that an excessive focus on finding unique 'marker genes' is limiting the utility of RNAseq for revealing biological structure and function. Not only does this lead to one ignoring the substantial variance within a 'type', but also to not studying how different cell types are organized in higher dimensional space: for example, is there a sequential change across multiple cell types, as we identified in chapter two? We sought to understand how gene expression profiles relate to one another, revealing a strikingly simple logic at the heart of the inner chamber of the forebrain. This, at least to me, represents moving beyond categorizing and toward understanding, making testable predictions that extend across different modalities: in short, it allows the identification of general principles.

Thesis references

Alexander, G. E., DeLong, M. R., & Strick, P. L. (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annual Review of Neuroscience*, *9*(1), 357–381. http://doi.org/10.1146/annurev.ne.09.030186.002041

Altman, J., & Bayer, S. A. (1988). Development of the rat thalamus: II. Time and site of origin and settling pattern of neurons derived from the anterior lobule of the thalamic neuroepithelium. *Journal of Comparative Neurology*, *275*(3), 378–405. http://doi.org/10.1002/cne.902750305

Altman, J., & Bayer, S. A. (1989). Development of the rat thalamus: VI. The posterior lobule of the thalamic neuroepithelium and the time and site of origin and settling pattern of neurons of the lateral geniculate and lateral posterior nuclei. *Journal of Comparative Neurology*, *284*(4), 581–601. http://doi.org/10.1002/cne.902840407

Anders, S., & Huber, W. (2010). Differential expression analysis for sequence count data. *Genome Biology*, *11*(10), R106. http://doi.org/10.1186/gb-2010-11-10-r106

Ballion, B., Mallet, N., Bézard, E., Lanciego, J. L., & Gonon, F. (2008). Intratelencephalic corticostriatal neurons equally excite striatonigral and striatopallidal neurons and their discharge activity is selectively reduced in experimental parkinsonism. *European Journal of Neuroscience*, *27*(9), 2313–2321. http://doi.org/10.1111/j.1460-9568.2008.06192.x

Bauswein, E., Fromm, C., & Preuss, A. (1989). Corticostriatal cells in comparison with pyramidal tract neurons: contrasting properties in the behaving monkey. *Brain Research*, *493*(1), 198–203.

Branco, T., & Häusser, M. (2010). The single dendritic branch as a fundamental functional unit in the nervous system. *Current Opinion in Neurobiology*, *20*(4), 494–502. http://doi.org/10.1016/j.conb.2010.07.009

Branco, T., & Häusser, M. (2011). Synaptic Integration Gradients in Single Cortical Pyramidal Cell Dendrites. *Neuron*, *69*(5), 885–892. http://doi.org/10.1016/j.neuron.2011.02.006

Branco, T., Clark, B. A., & Häusser, M. (2010). Dendritic discrimination of temporal input sequences in cortical neurons. *Science*, *329*(5999), 1671–1675. http://doi.org/10.1126/science.1189664

Briggs, J. A., Weinreb, C., Wagner, D. E., Megason, S., Peshkin, L., Kirschner, M. W., & Klein, A. M. (2018). The dynamics of gene expression in vertebrate embryogenesis at single-cell resolution. *Science*, *360*(6392), eaar5780–17. http://doi.org/10.1126/science.aar5780

Browne, S. H., Kang, J., Akk, G., Chiang, L. W., Schulman, H., Huguenard, J. R., & Prince, D. A. (2001a). Kinetic and pharmacological properties of GABA(A) receptors in single thalamic neurons and GABA(A) subunit expression. *Journal of Neurophysiology*, *86*(5), 2312–2322. http://doi.org/10.1152/jn.2001.86.5.2312

Butler, A., & Satija, R. (2017). Integrated analysis of single cell transcriptomic data across conditions, technologies, and species. *Book*, 1–18. http://doi.org/10.1101/164889

Buzsaki, G. (2006). Rhythms of the Brain. Oxford University Press.

Calderon, D. P., Fremont, R., Kraenzlin, F., & Khodakhah, K. (2011). The neural substrates of rapid-onset Dystonia- Parkinsonism. *Nature Publishing Group*, *14*(3), 357–365. http://doi.org/10.1038/nn.2753

Carandini, M. (2012). From circuits to behavior: a bridge too far? *Nature Publishing Group*, *15*(4), 507–509. http://doi.org/10.1038/nn.3043

Carrera, E., & Tononi, G. (2014). Diaschisis: past, present, future. *Brain : a Journal of Neurology*, *137*(Pt 9), 2408–2422. http://doi.org/10.1093/brain/awu101

Cembrowski, M. S., & Menon, V. (2018). Continuous Variation within Cell Types of the Nervous System. *Trends in Neurosciences*, *41*(6), 337–348. http://doi.org/10.1016/j.tins.2018.02.010

Cembrowski, M. S., Bachman, J. L., Wang, L., Sugino, K., Shields, B. C., & Spruston, N. (2016a). Spatial Gene-Expression Gradients Underlie Prominent Heterogeneity of CA1 Pyramidal Neurons. *Neuron*, *89*(2), 351–368. http://doi.org/10.1016/j.neuron.2015.12.013

Cembrowski, M. S., Bachman, J. L., Wang, L., Sugino, K., Shields, B. C., & Spruston, N. (2016b). Spatial Gene-Expression Gradients Underlie Prominent Heterogeneity of CA1 Pyramidal Neurons. *Neuron*, *89*(2), 351–368. http://doi.org/10.1016/j.neuron.2015.12.013

Chen, C. H., Fremont, R., Arteaga-Bracho, E. E., & Khodakhah, K. (2014). Short latency cerebellar modulation of the basal ganglia. *Nature Publishing Group*, *17*(12), 1767–1775. http://doi.org/10.1038/nn.3868

Chen, T.-W., Wardill, T. J., Sun, Y., Pulver, S. R., Renninger, S. L., Baohan, A., Schreiter, E. R., Kerr, R. A., Orger, M. B., Jayaraman, V., Looger, L. L., Svoboda, K., & Kim, D. S. (2013a). Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature*, *499*(7458), 295–300. http://doi.org/10.1038/nature12354

Chen, T.-W., Wardill, T. J., Sun, Y., Pulver, S. R., Renninger, S. L., Baohan, A., Schreiter, E. R., Kerr, R. A., Orger, M. B., Jayaraman, V., Looger, L. L., Svoboda, K., & Kim, D. S. (2013b). Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature*, *499*(7458), 295–300. http://doi.org/10.1038/nature12354

Citri, A., & Malenka, R. C. (2007). Synaptic Plasticity: Multiple Forms, Functions and Mechanisms. *Neuropsychopharmacology*, *33*(1), 18–41. http://doi.org/10.1038/sj.npp.1301559

Clascá, F., Rubio-Garrido, P., & Jabaudon, D. (2012). Unveiling the diversity of thalamocortical neuron subtypes. *European Journal of Neuroscience*, *35*(10), 1524–1532. http://doi.org/10.1111/j.1460-9568.2012.08033.x

Crevecoeur, F., Thonnard, J. L., Lefevre, P., & Scott, S. H. (2016). Long-Latency Feedback Coordinates Upper-Limb and Hand Muscles during Object Manipulation Tasks. *eNeuro*, *3*(1), 1–12. http://doi.org/10.1523/ENEURO.0129-15.2016

Crossman, A. R. (2000). Functional anatomy of movement disorders. *Journal of Anatomy*, *196 (Pt 4)*(Pt 4), 519–525. http://doi.org/10.1046/j.1469-7580.2000.19640519.x

Crossman, A. R., Mitchell, I. J., & Sambrook, M. A. (1985). Regional brain uptake of 2-deoxyglucose in N-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP)-induced parkinsonism in the macaque monkey. *Neuropharmacology*, *24*(6), 587–591.

Dembrow, N. C., Zemelman, B. V., & Johnston, D. (2015). Temporal Dynamics of L5 Dendrites in Medial Prefrontal Cortex Regulate Integration Versus Coincidence Detection of Afferent Inputs. *Journal of Neuroscience*, *35*(11), 4501–4514. http://doi.org/10.1523/JNEUROSCI.4673-14.2015

Denève, S., & Machens, C. K. (2016). Efficient codes and balanced networks. *Nature Neuroscience*, *19*(3), 375–382. http://doi.org/10.1038/nn.4243

Deng, Y., Lanciego, J., Goff, L. K.-L., Coulon, P., Salin, P., Kachidian, P., et al. (2015). Differential organization of cortical inputs to striatal projection neurons of the matrix compartment in rats. *Frontiers in Systems Neuroscience*, *9*, 1–14. http://doi.org/10.3389/fnsys.2015.00051

Dimitrov, A. G., Lazar, A. A., & Victor, J. D. (2011). Information theory in neuroscience. *Journal of Computational Neuroscience*, *30*(1), 1–5. http://doi.org/10.1007/s10827-011-0314-3

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., et al. (2012). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, *29*(1), 15–21. http://doi.org/10.1093/bioinformatics/bts635

Doya, K. (2000). Complementary roles of basal ganglia and cerebellum in learning and motor control. *Current Opinion in Neurobiology*, *10*(6), 732–739.

Ebbesen, C. L., & Brecht, M. (2017). Motor cortex — to act or not to act? *Nature Publishing Group*, *18*(11), 694–705. http://doi.org/10.1038/nrn.2017.119

Ebbesen, C. L., Doron, G., Lenschow, C., & Brecht, M. (2016). Vibrissa motor cortex activity suppresses contralateral whisking behavior. *Nature Neuroscience*, *20*(1), 82–89. http://doi.org/10.1038/nn.4437

Eidelberg, D. (2009). Metabolic brain networks in neurodegenerative disorders: a functional imaging approach. *Trends in Neurosciences*, *32*(10), 548–557. http://doi.org/10.1016/j.tins.2009.06.003

Ellender, T. J., Harwood, J., Kosillo, P., Capogna, M., & Bolam, J. P. (2013). Heterogeneous properties of central lateral and parafascicular thalamic synapses in the striatum. *The Journal of Physiology*, *591*(1), 257–272. http://doi.org/10.1113/jphysiol.2012.245233

Estebanez, L., Hoffmann, D., Voigt, B. C., & Poulet, J. F. A. (2017). Parvalbumin-Expressing GABAergic Neurons in Primary Motor Cortex Signal Reaching. *CellReports*, *20*(2), 308–318. http://doi.org/10.1016/j.celrep.2017.06.044

Farrell, J. A., Wang, Y., Riesenfeld, S. J., Shekhar, K., Regev, A., & Schier, A. F. (2018). Single-cell reconstruction of developmental trajectories during zebrafish embryogenesis. *Science*, *360*(6392), eaar3131–15. http://doi.org/10.1126/science.aar3131

Feingold, J., Lundqvist, M., Feingold, J., & Lundqvist, M. (2018). Neural Oscillations: Sustained Rhythms or Transient Burst- Events? *Trends in Neurosciences*, *41*(7), 415–417. http://doi.org/10.1016/j.tins.2018.04.004

Finger, S., Koehler, P. J., & Jagella, C. (2004). The Monakow concept of diaschisis: origins and perspectives. *Archives of neurology* (Vol. 61, pp. 283–288). American Medical Association. http://doi.org/10.1001/archneur.61.2.283

Fogerson, P. M., & Huguenard, J. R. (2016). Tapping the Brakes: Cellular and Synaptic Mechanisms that Regulate Thalamic Oscillations. *Neuron*, *92*(4), 687–704. http://doi.org/10.1016/j.neuron.2016.10.024

Frangeul, L., Pouchelon, G., Telley, L., Lefort, S., Luscher, C., & Jabaudon, D. (2016). A cross-modal genetic framework for the development and plasticity of sensory pathways. *Nature*, *538*(7623), 96–98. http://doi.org/10.1038/nature19770

Frassoni, C., Arcelli, P., Selvaggio, M., & Spreafico, R. (1998). Calretinin immunoreactivity in the developing thalamus of the rat: a marker of early generated thalamic cells. *Nsc*, *83*(4), 1203–1214.

Freeman, S. A., Desmazières, A., Fricker, D., Lubetzki, C., & Sol-Foulon, N. (2015). Mechanisms of sodium channel clustering and its influence on axonal impulse conduction. *Cellular and Molecular Life Sciences*, *73*(4), 723–735. http://doi.org/10.1007/s00018-015-2081-1

Gaidica, M., Hurst, A., Cyr, C., & Leventhal, D. K. (2018). Distinct Populations of Motor Thalamic Neurons Encode Action Initiation, Action Selection, and Movement Vigor. *Journal of Neuroscience*, *38*(29), 6563–6573. http://doi.org/10.1523/JNEUROSCI.0463-18.2018

Gerfen, C. R., Paletzki, R., & Heintz, N. (2013a). GENSAT BAC Cre-Recombinase Driver Lines to Study the Functional Organization of Cerebral Cortical and Basal Ganglia Circuits. *Neuron*, *80*(6), 1368–1383. http://doi.org/10.1016/j.neuron.2013.10.016

Gerfen, C. R., Paletzki, R., & Heintz, N. (2013b). GENSAT BAC Cre-Recombinase Driver Lines to Study the Functional Organization of Cerebral Cortical and Basal Ganglia Circuits. *Neuron*, *80*(6), 1368–1383. http://doi.org/10.1016/j.neuron.2013.10.016

Gokce, O., Stanley, G. M., Treutlein, B., Neff, N. F., Camp, J. G., Malenka, R. C., Rothwell, P. E., Fuccillo, M. V., Südhof, T. C., & Quake, S. R. (2016a). Cellular Taxonomy of the Mouse Striatum as Revealed by Single-Cell RNA-Seq. *CellReports*, *16*(4), 1126–1137. http://doi.org/10.1016/j.celrep.2016.06.059

Gokce, O., Stanley, G. M., Treutlein, B., Neff, N. F., Camp, J. G., Malenka, R. C., Rothwell, P. E., Fuccillo, M. V., Südhof, T. C., & Quake, S. R. (2016b). Cellular Taxonomy of the Mouse Striatum as Revealed by Single-Cell RNA-Seq. *CellReports*, *16*(4), 1126–1137. http://doi.org/10.1016/j.celrep.2016.06.059

Goldey, G. J., Roumis, D. K., Glickfeld, L. L., Kerlin, A. M., Reid, R. C., Bonin, V., et al. (2014). Removable cranial windows for long-term imaging in awake mice. *Nature Protocols*, *9*(11), 2515–2538. http://doi.org/10.1038/nprot.2014.165

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, *15*(1), 20–25. http://doi.org/10.1038/nature16442

Graziano, M. S. A., Taylor, C. S. R., & Moore, T. (2002). Complex movements evoked by microstimulation of precentral cortex. *Neuron*, *34*(5), 841–851.

Guo, J.-Z., Graves, A. R., Guo, W. W., Zheng, J., Lee, A., Rodríguez-González, J., et al. (2015). Cortex commands the performance of skilled movement. *eLife*, *4*, e10774. http://doi.org/10.7554/eLife.10774

Guo, Z. V., Hires, S. A., Li, N., O'Connor, D. H., Komiyama, T., Ophir, E., et al. (2014). Procedures for Behavioral Experiments in Head-Fixed Mice. *PLoS ONE*, *9*(2), e88678–16. http://doi.org/10.1371/journal.pone.0088678

Habib, N., Li, Y., Heidenreich, M., Swiech, L., Avraham-Davidi, I., Trombetta, J. J., et al. (2016). Div-Seq: Single-nucleus RNA-Seq reveals dynamics of rare adult newborn neurons. *Science*, *353*(6302), 925–928. http://doi.org/10.1126/science.aad7038

Harris, K. D., & Shepherd, G. M. G. (2015). The neocortical circuit: themes and variations. *Nature Neuroscience*, *18*(2), 170–181. http://doi.org/10.1038/nn.3917

Harrison, T. C., Ayling, O. G. S., & Murphy, T. H. (2012). Distinct Cortical Circuit Mechanisms for Complex Forelimb Movement and Motor Map Topography. *Neuron*, *74*(2), 397–409. http://doi.org/10.1016/j.neuron.2012.02.028

Hasegawa, M., Majima, K., Itokazu, T., Maki, T., Albrecht, U.-R., Castner, N., et al. (2017). Selective Suppression of Local Circuits during Movement Preparation in the Mouse Motor Cortex. *CellReports*, *18*(11), 2676–2686. http://doi.org/10.1016/j.celrep.2017.02.043

Hassabis, D., Kumaran, D., Summerfield, C., & Botvinick, M. (2017). Neuroscience-Inspired Artificial Intelligence. *Neuron*, *95*(2), 245–258. http://doi.org/10.1016/j.neuron.2017.06.011

Hatsopoulos, N. G. (2005). Encoding in the Motor Cortex: Was Evarts Right After All? Focus on “Motor Cortex Neural Correlates of Output Kinematics and Kinetics During Isometric-Force and Arm-Reaching Tasks.” *Journal of Neurophysiology*, *94*(4), 2261–2262. http://doi.org/10.1152/jn.00533.2005

Hatsopoulos, N. G., & Suminski, A. J. (2011). Sensing with the Motor Cortex. *Neuron*, *72*(3), 477–487. http://doi.org/10.1016/j.neuron.2011.10.020

Hattox, A. M., & Nelson, S. B. (2007). Layer V Neurons in Mouse Cortex Projecting to Different Targets Have Distinct Physiological Properties. *Journal of Neurophysiology*, *98*(6), 3330–3340. http://doi.org/10.1152/jn.00397.2007

Häusser, M., Spruston, N., & Stuart, G. J. (2000). Diversity and dynamics of dendritic signaling. *Science*, *290*(5492), 739–744.

Heindorf, M., Arber, S., & Keller, G. B. (2018). Mouse Motor Cortex Coordinates the Behavioral Response to Unpredicted Sensory Feedback. *Neuron*, 1–21. http://doi.org/10.1016/j.neuron.2018.07.046

Hempel, C. M., Sugino, K., & Nelson, S. B. (2007). A manual method for the purification of fluorescently labeled neurons from the mammalian brain. *Nature Protocols*, *2*(11), 2924–2929. http://doi.org/10.1038/nprot.2007.416

Hooks, B. M., Papale, A. E., Paletzki, R. F., Feroze, M. W., Eastwood, B. S., Couey, J. J., Winnubst, J., Chandrashekar, J., & Gerfen, C. R. (2018a). Topographic precision in sensory and motor corticostriatal projections varies across cell type and cortical area. *Nature Communications*, 1–16. http://doi.org/10.1038/s41467-018-05780-7

Hooks, B. M., Papale, A. E., Paletzki, R., Feroze, M., Eastwood, B. S., Couey, J. J., Winnubst, J., Chandrashekar, J., & Gerfen, C. R. (2018b). Cell type-specific variation of somatotopic precision across corticostriatal projections. *Book*, 1–30. http://doi.org/10.1101/261446

Horak, F. B., & Anderson, M. E. (1984). Influence of globus pallidus on arm movements in monkeys. I. Effects of kainic acid-induced lesions. *Journal of Neurophysiology*, *52*(2), 290–304. http://doi.org/10.1152/jn.1984.52.2.290

Hornyak, M., Rovit, R. L., Simon, A. S., & Couldwell, W. T. (2001). Irving S. Cooper and the early surgical management of movement disorders. Video history. *Neurosurgical focus* (Vol. 11).

Hrvatin, S., Hochbaum, D. R., Nagy, M. A., Cicconet, M., Robertson, K., Cheadle, L., et al. (2017). Single-cell analysis of experience-dependent transcriptomic states in the mouse visual cortex. *Nature Neuroscience*, *21*(1), 120–129. http://doi.org/10.1038/s41593-017-0029-5

Hu, B., Senatorov, V., & Mooney, D. (1994). Lemniscal and non-lemniscal synaptic transmission in rat auditory thalamus. *The Journal of Physiology*, *479 (Pt 2)*(Pt 2), 217–231. http://doi.org/10.1111/(ISSN)1469-7793

Huang, C., Tang, C., Feigin, A., Lesser, M., Ma, Y., Pourfar, M., et al. (2007). Changes in network activity with the progression of Parkinson's disease. *Brain : a Journal of Neurology*, *130*(Pt 7), 1834–1846. http://doi.org/10.1093/brain/awm086

Huber, D., Gutnisky, D. A., Peron, S., O'Connor, D. H., Wiegert, J. S., Tian, L., et al. (2012). Multiple dynamic representations in the motor cortex during sensorimotor learning. *Nature*, *484*(7395), 473–478. http://doi.org/10.1038/nature11039

Huntenburg, J. M., Bazin, P.-L., & Margulies, D. S. (2018). Large-Scale Gradients in Human Cortical Organization. *Trends in Cognitive Sciences*, *22*(1), 21–31. http://doi.org/10.1016/j.tics.2017.11.002

Jhangiani-Jashanmal, I. T., Yamamoto, R., Gungor, N. Z., & Paré, D. (2016). Electroresponsive properties of rat central medial thalamic neurons. *Journal of Neurophysiology*, *115*(3), 1533–1541. http://doi.org/10.1152/jn.00982.2015

Jones, E. (1998). Viewpoint: The core and matrix of thalamic organization. *Neuroscience*.

Jones, E. (2007). The Thalamus. *Cambridge University Press*.

Jones, E. G. (2001). The thalamic matrix and thalamocortical synchrony. *Trends in Neurosciences*, *24*(10), 595–601. http://doi.org/10.1016/S0166-2236(00)01922-6

Jones, E. G., & Hendry, S. H. C. (1989). Differential Calcium Binding Protein Immunoreactivity Distinguishes Classes of Relay Neurons in Monkey Thalamic Nuclei. *The European Journal of Neuroscience*, *1*(3), 222–246.

Jones, E. G., & Rubenstein, J. L. R. (2004). Expression of regulatory genes during differentiation of thalamic nuclei in mouse and monkey. *Journal of Comparative Neurology*, *477*(1), 55–80. http://doi.org/10.1002/cne.20234

Jun, J. J., Steinmetz, N. A., Siegle, J. H., Denman, D. J., Bauza, M., Barbarits, B., et al. (2017). Fully integrated silicon probes for high-density recording of neural activity. *Nature*, *551*(7679), 232–236. http://doi.org/10.1038/nature24636

Kandel, E., Jesssell, T., & Schwartz, J. (2012). The Principles of Neural Science. http://doi.org/1

Kinnischtzke, A. K., Simons, D. J., & Fanselow, E. E. (2014). Motor Cortex Broadly Engages Excitatory and Inhibitory Neurons in Somatosensory Barrel Cortex. *Cerebral Cortex*, *24*(8), 2237–2248. http://doi.org/10.1093/cercor/bht085

Kita, T., Shigematsu, N., & Kita, H. (2016a). Intralaminar and tectal projections to the subthalamus in the rat. *European Journal of Neuroscience*, *44*(11), 2899–2908. http://doi.org/10.1111/ejn.13413

Kita, T., Shigematsu, N., & Kita, H. (2016b). Intralaminar and tectal projections to the subthalamus in the rat. *European Journal of Neuroscience*, *44*(11), 2899–2908. http://doi.org/10.1111/ejn.13413

Knudsen, E. I. (2007). Fundamental components of attention. *Annual Review of Neuroscience*, *30*(1), 57–78. http://doi.org/10.1146/annurev.neuro.30.051606.094256

Koyama, M., & Pujala, A. (2018). Mutual inhibition of lateral inhibition: a network motif for an elementary computation in the brain. *Current Opinion in Neurobiology*, *49*, 69–74. http://doi.org/10.1016/j.conb.2017.12.019

Kunimatsu, J., Suzuki, T. W., Ohmae, S., & Tanaka, M. (2018). Different contributions of preparatory activity in the basal ganglia and cerebellum for self-timing. *eLife*, *7*, 743. http://doi.org/10.7554/eLife.35676

Kuramoto, E., Fujiyama, F., Nakamura, K. C., Tanaka, Y., Hioki, H., & Kaneko, T. (2010). Complementary distribution of glutamatergic cerebellar and GABAergic basal ganglia afferents to the rat motor thalamic nuclei. *European Journal of Neuroscience*, *33*(1), 95–109. http://doi.org/10.1111/j.1460-9568.2010.07481.x

Kuramoto, E., Furuta, T., Nakamura, K. C., Unzai, T., Hioki, H., & Kaneko, T. (2009). Two Types of Thalamocortical Projections from the Motor Thalamic Nuclei of the Rat: A Single Neuron-Tracing Study Using Viral Vectors. *Cerebral Cortex*, *19*(9), 2065–2077. http://doi.org/10.1093/cercor/bhn231

Kuramoto, E., Pan, S., Furuta, T., Tanaka, Y. R., Iwai, H., Yamanaka, A., et al. (2016). Individual mediodorsal thalamic neurons project to multiple areas of the rat

prefrontal cortex: A single neuron-tracing study using virus vectors. *Journal of Comparative Neurology*, *525*(1), 166–185. http://doi.org/10.1002/cne.24054

Landau, I. D., Egger, R., Dercksen, V. J., Oberlaender, M., & Sompolinsky, H. (2016). The Impact of Structural Heterogeneity on Excitation-Inhibition Balance in Cortical Networks. *Neuron*, *92*(5), 1106–1121. http://doi.org/10.1016/j.neuron.2016.10.027

Le Xiao, Bornmann, C., Hatstatt-Burklé, L., & Scheiffele, P. (2018). Regulation of striatal cells and goal-directed behavior by cerebellar outputs. *Nature Communications*, 1–14. http://doi.org/10.1038/s41467-018-05565-y

Lei, W. (2004). Evidence for Differential Cortical Input to Direct Pathway versus Indirect Pathway Striatal Projection Neurons in Rats. *Journal of Neuroscience*, *24*(38), 8289–8299. http://doi.org/10.1523/JNEUROSCI.1990-04.2004

Leicher, T., Bähring, R., Isbrandt, D., & Pongs, O. (1998). Coexpression of the KCNA3B gene product with Kv1.5 leads to a novel A-type potassium channel. *The Journal of Biological Chemistry*, *273*(52), 35095–35101.

Li, J. (2003). Distinct Firing Properties of Higher Order Thalamic Relay Neurons. *Journal of Neurophysiology*, *90*(1), 291–299. http://doi.org/10.1152/jn.01163.2002

Li, J., Bickford, M. E., & Guido, W. (2003). Distinct Firing Properties of Higher Order Thalamic Relay Neurons. *Journal of Neurophysiology*, *90*(1), 291–299. http://doi.org/10.1152/jn.01163.2002

Li, N., Chen, T.-W., Guo, Z. V., Gerfen, C. R., & Svoboda, K. (2015a). A motor cortex circuit for motor planning and movement. *Nature*, *519*(7541), 51–56. http://doi.org/10.1038/nature14178

Lintz, M. J., & Felsen, G. (2016). Basal ganglia output reflects internally-specified movements. *eLife*, *5*, 111. http://doi.org/10.7554/eLife.13833

Little, S., & Brown, P. (2014). The functional role of beta oscillations in Parkinson's disease. *Parkinsonism & Related Disorders*, *20 Suppl 1*, S44–8. http://doi.org/10.1016/S1353-8020(13)70013-0

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12), 31–21. http://doi.org/10.1186/s13059-014-0550-8

Lu, E., Llano, D. A., & Sherman, S. M. (2009). Different distributions of calbindin and calretinin immunostaining across the medial and dorsal divisions of the mouse medial geniculate body. *Hearing Research*, *257*(1-2), 16–23. http://doi.org/10.1016/j.heares.2009.07.009

Mandelbaum, G., Taranda, J., Haynes, T., Hochbaum, D., Huang, K. W., Hyun, M., et al. (2018). Distinct Cortical-Thalamic-Striatal Circuits Through the Parafascicular Nucleus. *Book*, 1–67. http://doi.org/10.1101/370734

McCairn, K. W., & Turner, R. S. (2015). Pallidal stimulation suppresses pathological dysrhythmia in the parkinsonian motor cortex. *Journal of Neurophysiology*, *113*(7), 2537–2548. http://doi.org/10.1152/jn.00701.2014

McCairn, K. W., Bronfeld, M., Belelovsky, K., & Bar-Gad, I. (2009). The neurophysiological correlates of motor tics following focal striatal disinhibition. *Brain : a Journal of Neurology*, *132*(8), 2125–2138. http://doi.org/10.1093/brain/awp142

McCairn, K. W., Iriki, A., & Isoda, M. (2013). Global Dysrhythmia of Cerebro-Basal Ganglia-Cerebellar Networks Underlies Motor Tics following Striatal Disinhibition. *Journal of Neuroscience*, *33*(2), 697–708. http://doi.org/10.1523/JNEUROSCI.4018-12.2013

Mease, R. A., Sumser, A., Sakmann, B., & Groh, A. (2016). Cortical Dependence of Whisker Responses in Posterior Medial Thalamus In Vivo. *Cerebral Cortex*, *26*(8), 3534–3543. http://doi.org/10.1093/cercor/bhw144

Miall, R. C., Price, S., Mason, R., Passingham, R. E., Winter, J. L., & Stein, J. F. (1998). Microstimulation of movements from cerebellar-receiving, but not pallidal-receiving areas of the macaque thalamus under ketamine anaesthesia. *Experimental Brain Research*, *123*(4), 387–396.

Miall, R. C., Weir, D. J., behavior, D. W. J. O. M., 1993. (1993). Is the cerebellum a smith predictor? *Taylor & Francis*

, *25*(3), 203–216. http://doi.org/10.1080/00222895.1993.9942050

Mitchell, I. J., Boyce, S., Sambrook, M. A., & Crossman, A. R. (1992). A 2-deoxyglucose study of the effects of dopamine agonists on the parkinsonian primate brain. Implications for the neural mechanisms that mediate dopamine agonist-induced dyskinesia. *Brain : a Journal of Neurology*, *115 (Pt 3)*, 809–824.

Mitchell, I. J., Jackson, A., Sambrook, M. A., & Crossman, A. R. (1985). Common neural mechanisms in experimental chorea and hemiballismus in the monkey. Evidence from 2-deoxyglucose autoradiography. *Brain Research*, *339*(2), 346–350.

Mitchell, I. J., Luquin, R., Boyce, S., Clarke, C. E., Robertson, R. G., Sambrook, M. A., & Crossman, A. R. (1990). Neural mechanisms of dystonia: evidence from a 2-deoxyglucose uptake study in a primate model of dopamine agonist-induced dystonia. *Movement Disorders : Official Journal of the Movement Disorder Society*, *5*(1), 49–54. http://doi.org/10.1002/mds.870050113

Mo, C., & Sherman, S. M. (2017). A sensorimotor pathway via higher order thalamus. Presented at the Society For Neuroscience.

Molnár, Z. (2004). Thomas Willis (1621-1675), the founder of clinical neuroscience. *Nature reviews. Neuroscience* (Vol. 5, pp. 329–335). Nature Publishing Group. http://doi.org/10.1038/nrn1369

Mueller, T. (2012). What is the Thalamus in Zebrafish? *Frontiers in Neuroscience*, *6*, 64. http://doi.org/10.3389/fnins.2012.00064

Muñoz-Manchado, A. B., Gonzales, C. B., Zeisel, A., Munguba, H., Bekkouche, B., Skene, N. G., Lönnerberg, P., Ryge, J., Harris, K. D., Linnarsson, S., & Hjerling-Leffler, J. (2018a). Diversity of Interneurons in the Dorsal Striatum Revealed by Single-Cell RNA Sequencing and PatchSeq. *CellReports*, *24*(8), 2179–2190.e7. http://doi.org/10.1016/j.celrep.2018.07.053

Muñoz-Manchado, A. B., Gonzales, C. B., Zeisel, A., Munguba, H., Bekkouche, B., Skene, N. G., Lönnerberg, P., Ryge, J., Harris, K. D., Linnarsson, S., & Hjerling-

Leffler, J. (2018b). Diversity of Interneurons in the Dorsal Striatum Revealed by Single-Cell RNA Sequencing and PatchSeq. *CellReports*, *24*(8), 2179–2190.e7. http://doi.org/10.1016/j.celrep.2018.07.053

Murray, K. D., Choudary, P. V., & Jones, E. G. (2007). Nucleus- and cell-specific gene expression in monkey thalamus. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(6), 1989–1994. http://doi.org/10.1073/pnas.0610742104

Nagalski, A., Puelles, L., Dabrowski, M., Wegierski, T., Kuznicki, J., & Wisniewska, M. B. (2015). Molecular anatomy of the thalamic complex and the underlying transcription factors. *Brain Structure and Function*, *221*(5), 2493–2510. http://doi.org/10.1007/s00429-015-1052-5

Nakamura, H., Hioki, H., Furuta, T., & Kaneko, T. (2015). Different cortical projections from three subdivisions of the rat lateral posterior thalamic nucleus: a single-neuron tracing study with viral vectors. *European Journal of Neuroscience*, *41*(10), 1294–1310. http://doi.org/10.1111/ejn.12882

Nakamura, K. C., Sharott, A., & Magill, P. J. (2012). Temporal Coupling with Cortex Distinguishes Spontaneous Neuronal Activities in Identified Basal Ganglia-Recipient and Cerebellar-Recipient Zones of the Motor Thalamus. *Cerebral Cortex*, *24*(1), 81–97. http://doi.org/10.1093/cercor/bhs287

Nelson, S. B., Sugino, K., & Hempel, C. M. (2006). The problem of neuronal cell types: a physiological genomics approach. *Trends in Neurosciences*, *29*(6), 339–345. http://doi.org/10.1016/j.tins.2006.05.004

Neumann, von, J. (1958). The Computer and the Brain.

Neychev, V. K., Fan, X., Mitev, V. I., Hess, E. J., & Jinnah, H. A. (2008). The basal ganglia and cerebellum interact in the expression of dystonic movement. *Brain : a Journal of Neurology*, *131*(9), 2499–2509. http://doi.org/10.1093/brain/awn168

Osborne, J. E., & Dudman, J. T. (2014). RIVETS: a mechanical system for in vivo and in vitro electrophysiology and imaging. *PLoS ONE*, *9*(2), e89007. http://doi.org/10.1371/journal.pone.0089007

Otchy, T. M., Wolff, S. B. E., Rhee, J. Y., Pehlevan, C., Kawai, R., Kempf, A., Gobes, S. M. H., & Ölveczky, B. P. (2015a). Acute off-target effects of neural circuit manipulations. *Nature*, *528*(7582), 358–363. http://doi.org/10.1038/nature16442

O'Leary, T., Williams, A. H., Franci, A., & Marder, E. (2014). Cell Types, Network Homeostasis, and Pathological Compensation from a Biologically Plausible Ion Channel Expression Model. *Neuron*, *82*(4), 809–821. http://doi.org/10.1016/j.neuron.2014.04.002

Packer, A. M., Russell, L. E., Dalgleish, H. W. P., & Häusser, M. (2014). Simultaneous all-optical manipulation and recording of neural circuit activity with cellular resolution in vivo. *Nature Methods*, *12*(2), 140–146. http://doi.org/10.1038/nmeth.3217

Panigrahi, B., Martin, K. A., Li, Y., Graves, A. R., Vollmer, A., Olson, L., et al. (2015). Dopamine Is Required for the Neural Representation and Control of Movement Vigor. *Cell*, *162*(6), 1418–1430. http://doi.org/10.1016/j.cell.2015.08.014

Parent, A. (2012). The History of the Basal Ganglia: The Contribution of Karl Friedrich Burdach. *Neuroscience and Medicine*, *03*(04), 374–379. http://doi.org/10.4236/nm.2012.34046

Paxinos, G., & Franklin, K. (2012). The Mouse Brain in Stereotaxic Coordinates, 4th Edition. *Academic Press*.

Peters, A. J., Lee, J., Hedrick, N. G., O'Neil, K., & Komiyama, T. (2017). Reorganization of corticospinal output during motor learning. *Nature Publishing Group*, 1–22. http://doi.org/10.1038/nn.4596

Phillips, J. W., Schulmann, A., Hara, E., Liu, C., Wang, L., Shields, B., Korff, W., Lemire, A., Dudman, J., Nelson, S. B., & Hantman, A. (2018a). A single spectrum of neuronal identities across thalamus. *Bioxriv*, 1–31. http://doi.org/10.1101/241315

Phillips, J. W., Schulmann, A., Hara, E., Liu, C., Wang, L., Shields, B., Korff, W., Lemire, A., Dudman, J., Nelson, S. B., & Hantman, A. (2018b). A single spectrum of neuronal identities across thalamus, 1–31. http://doi.org/10.1101/241315

Picelli, S., Björklund, Å. K., Faridani, O. R., Sagasser, S., Winberg, G., & Sandberg, R. (2013). Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nature Methods*, *10*(11), 1096–1098. http://doi.org/10.1038/nmeth.2639

Pidoux, L., Le Blanc, P., Levenes, C., & Leblois, A. (2018). A subcortical circuit linking the cerebellum to the basal ganglia engaged in vocal learning. *eLife*, *7*. http://doi.org/10.7554/eLife.32167

Podgorski, K., & Ranganathan, G. (2016). Brain heating induced by near-infrared lasers during multiphoton microscopy. *Journal of Neurophysiology*, *116*(3), 1012–1023. http://doi.org/10.1152/jn.00275.2016

Portera-Cailliau, C., Weimer, R. M., De Paola, V., Caroni, P., & Svoboda, K. (2005). Diverse Modes of Axon Elaboration in the Developing Neocortex. *PLOS Biology*, *3*(8), e272–15. http://doi.org/10.1371/journal.pbio.0030272

Pouille, F., & Scanziani, M. (2001). Enforcement of temporal fidelity in pyramidal cells by somatic feed-forward inhibition. *Science*, *293*(5532), 1159–1163. http://doi.org/10.1126/science.1060342

Poulin, J.-F., Tasic, B., Hjerling-Leffler, J., Trimarchi, J. M., & Awatramani, R. (2016). Disentangling neural cell diversity using single-cell transcriptomics. *Nature Neuroscience*, *19*(9), 1131–1141. http://doi.org/10.1038/nn.4366

Puelles, L., Sánchez, M. P., Spreafico, R., & Fairén, A. (1992). Prenatal development of calbindin immunoreactivity in the dorsal thalamus of the rat. *Nsc*, *46*(1), 135–147.

Puil, E., Meiri, H., & Yarom, Y. (1994). Resonant behavior and frequency preferences of thalamic neurons. *Journal of Neurophysiology*, *71*(2), 575–582.

Ramanathan, D., Conner, J. M., & Tuszynski, M. H. (2006). A form of motor cortical plasticity that correlates with recovery of function after brain injury. *Proceedings of the National Academy of Sciences*, *103*(30), 11370–11375. http://doi.org/10.1073/pnas.0601065103

Ramcharan, E. J., Gnadt, J. W., & Sherman, S. M. (2005). Higher-order thalamic relays burst more than first-order relays. *Proceedings of the National Academy of Sciences*

of the United States of America, *102*(34), 12236–12241. http://doi.org/10.1073/pnas.0502843102

Raposo, D., Kaufman, M. T., & Churchland, A. K. (2014). A category-free neural population supports evolving demands during decision-making. *Nature Publishing Group*, *17*(12), 1784–1792. http://doi.org/10.1038/nn.3865

Redinbaugh, M. J., Phillips, J. M., Kambi, N. A., Mohanta, S., Raz, A., & Saalmann, Y. B. (2017). Stimulation of central lateral thalamus restores conscious processing under propofol and isoflurane anaesthesia. *Society for Neuroscience Abstract*, 1–12.

Reiner, A. (2010). Corticostriatal projection neurons – dichotomous types and dichotomous functions. *Frontiers in Neuroanatomy*, *4*, 1–15. http://doi.org/10.3389/fnana.2010.00142

Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2009). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, *26*(1), 139–140. http://doi.org/10.1093/bioinformatics/btp616

Rolls, E. T., Treves, A., Tovee, M. J., & Panzeri, S. (1997). Information in the neuronal representation of individual stimuli in the primate temporal visual cortex. *Journal of Computational Neuroscience*, *4*(4), 309–333. http://doi.org/10.1038/sj.npp.1301559

Roth, M. M., Dahmen, J. C., Muir, D. R., Imhof, F., Martini, F. J., & Hofer, S. B. (2015). Thalamic nuclei convey diverse contextual information to layer 1 of visual cortex. *Nature Neuroscience*, *19*(2), 299–307. http://doi.org/10.1038/nn.4197

Rubio-Garrido, P., Pérez-de-Manzo, F., & Clascá, F. (2007). Calcium-binding proteins as markers of layer-I projecting vs. deep layer-projecting thalamocortical neurons: A double-labeling analysis in the rat. *Neuroscience*, *149*(1), 242–250. http://doi.org/10.1016/j.neuroscience.2007.07.036

Rudy, B., & McBain, C. J. (2001). Kv3 channels: voltage-gated K+ channels designed for high-frequency repetitive firing. *Trends in Neurosciences*, *24*(9), 517–526.

Rueda-Orozco, P. E., & Robbe, D. (2015). The striatum multiplexes contextual and kinematic information to constrain motor habits execution. *Nature Neuroscience*, *18*(3), 453–460. http://doi.org/10.1038/nn.3924

Salic, A., & Mitchison, T. J. (2008). A chemical method for fast and sensitive detection of DNA synthesis in vivo. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(7), 2415–2420. http://doi.org/10.1073/pnas.0712168105

Satija, R., Farrell, J. A., Gennert, D., Schier, A. F., & Regev, A. (2015). Spatial reconstruction of single-cell gene expression data. *Nature Biotechnology*, *33*(5), 495–502. http://doi.org/10.1038/nbt.3192

Saunders, A., Macosko, E. Z., Wysoker, A., Goldman, M., Krienen, F. M., de Rivera, H., et al. (2018). Molecular Diversity and Specializations among the Cells of the Adult Mouse Brain. *Cell*, *174*(4), 1015–1030.e16. http://doi.org/10.1016/j.cell.2018.07.028

Sawyer, S. F., Young, S. J., Groves, P. M., & Tepper, J. M. (1994). Cerebellar-responsive neurons in the thalamic ventroanterior-ventrolateral complex of rats: in vivo electrophysiology. *Nsc*, *63*(3), 711–724.

Schiff, N. D. (2008). Central Thalamic Contributions to Arousal Regulation and Neurological Disorders of Consciousness. *Annals of the New York Academy of Sciences*, *1129*(1), 105–118. http://doi.org/10.1196/annals.1417.029

Shannon, C. E. (1997). The mathematical theory of communication. *M.D. computing : computers in medical practice* (Vol. 14, pp. 306–317).

Shenoy, K. V., Sahani, M., & Churchland, M. M. (2013a). Cortical Control of Arm Movements: A Dynamical Systems Perspective. *Annual Review of Neuroscience*, *36*(1), 337–359. http://doi.org/10.1146/annurev-neuro-062111-150509

Shenoy, K. V., Sahani, M., & Churchland, M. M. (2013b). Cortical Control of Arm Movements: A Dynamical Systems Perspective. *Annual Review of Neuroscience*, *36*(1), 337–359. http://doi.org/10.1146/annurev-neuro-062111-150509

Sherman, S. M. (2016). Thalamus plays a central role in ongoing cortical functioning. *Nature Neuroscience*, *19*(4), 533–541. http://doi.org/10.1038/nn.4269

Sherman, S. M., & Guillery, R. W. (2002). The role of the thalamus in the flow of information to the cortex. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *357*(1428), 1695–1708. http://doi.org/10.1098/rstb.2002.1161

Shi, W., Xianyu, A., Han, Z., Tang, X., Li, Z., Zhong, H., et al. (2017). Ontogenetic establishment of order-specific nuclear organization in the mammalian thalamus. *Nature Neuroscience*, *20*(4), 516–528. http://doi.org/10.1038/nn.4519

Smith, Y., Galvan, A., Ellender, T. J., Doig, N., Villalba, R. M., Huerta-Ocampo, I., et al. (2014). The thalamostriatal system in normal and diseased states. *Frontiers in Systems Neuroscience*, *8*, 5. http://doi.org/10.3389/fnsys.2014.00005

Soma, S., Saiki, A., Yoshida, J., Ríos, A., Kawabata, M., Sakai, Y., & Isomura, Y. (2017). Distinct Laterality in Forelimb-Movement Representations of Rat Primary and Secondary Motor Cortical Neurons with Intratelencephalic and Pyramidal Tract Projections. *Journal of Neuroscience*, *37*(45), 10904–10916. http://doi.org/10.1523/JNEUROSCI.1188-17.2017

Soumillon, M., Cacchiarelli, D., Semrau, S., van Oudenaarden, A., & Mikkelsen, T. S. (2014). Characterization of directed differentiation by high-throughput single-cell RNA-Seq. *Book*, 1–13. http://doi.org/10.1101/003236

Spitzer, N. C., Root, C. M., & Borodinsky, L. N. (2004). Orchestrating neuronal differentiation: patterns of Ca2+ spikes specify transmitter choice. *Trends in Neurosciences*, *27*(7), 415–421. http://doi.org/10.1016/j.tins.2004.05.003

Stein, J. F., & Aziz, T. Z. (1999). Does imbalance between basal ganglia and cerebellar outputs cause movement disorders? *Current Opinion in Neurology*, *12*(6), 667–669.

Steriade, M., Curró Dossi, R., & Contreras, D. (1993a). Electrophysiological properties of intralaminar thalamocortical cells discharging rhythmic (approximately 40 HZ) spike-bursts at approximately 1000 HZ during waking and rapid eye movement sleep. *Nsc*, *56*(1), 1–9.

Steriade, M., Curró Dossi, R., & Contreras, D. (1993b). Electrophysiological properties of intralaminar thalamocortical cells discharging rhythmic (approximately 40 HZ) spike-bursts at approximately 1000 HZ during waking and rapid eye movement sleep. *Nsc*, *56*(1), 1–9.

Sugino, K., Clark, E., Schulmann, A., Shima, Y., Wang, L., Hunt, D. L., et al. (2017). The Transcriptional Logic of Mammalian Neuronal Diversity. *Book*, 1–45. http://doi.org/10.1101/208355

Sumser, A., Mease, R. A., Sakmann, B., & Groh, A. (2017). Organization and somatotopy of corticothalamic projections from L5B in mouse barrel cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(33), 8853–8858. http://doi.org/10.1073/pnas.1704302114

Sun, W., Tan, Z., Mensh, B. D., & Ji, N. (2016). Thalamus provides layer 4 of primary visual cortex with orientation- and direction-tuned inputs. *Nature Publishing Group*, *19*(2), 308–315. http://doi.org/10.1038/nn.4196

Tan, H., Jenkinson, N., & Brown, P. (2014). Dynamic Neural Correlates of Motor Error Monitoring and Adaptation during Trial-to-Trial Learning. *Journal of Neuroscience*, *34*(16), 5678–5688. http://doi.org/10.1523/JNEUROSCI.4739-13.2014

Tan, H., Wade, C., & Brown, P. (2016). Post-Movement Beta Activity in Sensorimotor Cortex Indexes Confidence in the Estimations from Internal Models. *Journal of Neuroscience*, *36*(5), 1516–1528. http://doi.org/10.1523/JNEUROSCI.3204-15.2016

Tasic, B., Menon, V., Nguyen, T. N., Kim, T. K., Jarsky, T., Yao, Z., et al. (2016). Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. *Nature Neuroscience*, *19*(2), 335–346. http://doi.org/10.1038/nn.4216

Tennant, K. A., Adkins, D. L., Donlan, N. A., Asay, A. L., Thomas, N., Kleim, J. A., & Jones, T. A. (2010). The Organization of the Forelimb Representation of the C57BL/6 Mouse Motor Cortex as Defined by Intracortical Microstimulation and Cytoarchitecture. *Cerebral Cortex*, *21*(4), 865–876. http://doi.org/10.1093/cercor/bhq159

Tervo, D. G. R., Hwang, B.-Y., Viswanathan, S., Gaj, T., Lavzin, M., Ritola, K. D., et al. (2016). A Designer AAV Variant Permits Efficient Retrograde Access to Projection Neurons. *Neuron*, *92*(2), 372–382. http://doi.org/10.1016/j.neuron.2016.09.021

Theis, L., Berens, P., Froudarakis, E., Reimer, J., Rosón, M. R., Baden, T., et al. (2015, February 28). Supervised learning sets benchmark for robust spike detection from calcium imaging signals. *arXiv.org*.

Thiele, A., & Bellgrove, M. A. (2018). Neuromodulation of Attention. *Neuron*, *97*(4), 769–785. http://doi.org/10.1016/j.neuron.2018.01.008

Thompson, A. D., Picard, N., Min, L., Fagiolini, M., & Chen, C. (2016). Cortical Feedback Regulates Feedforward Retinogeniculate Refinement. *Neuron*, *91*(5), 1021–1033. http://doi.org/10.1016/j.neuron.2016.07.040

Tripathy, S. J., Padmanabhan, K., Gerkin, R. C., & Urban, N. N. (2013). Intermediate intrinsic diversity enhances neural population coding. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(20), 8248–8253. http://doi.org/10.1073/pnas.1221214110

Turner, R. S., & DeLong, M. R. (2000). Corticostriatal activity in primary motor cortex of the macaque. *Journal of Neuroscience*, *20*(18), 7096–7108.

van Donkelaar, P., Stein, J. F., Passingham, R. E., & Miall, R. C. (1999). Neuronal activity in the primate motor thalamus during visually triggered and internally generated limb movements. *Journal of Neurophysiology*, *82*(2), 934–945. http://doi.org/10.1152/jn.1999.82.2.934

van Donkelaar, P., Stein, J. F., Passingham, R. E., & Miall, R. C. (2000). Temporary inactivation in the primate motor thalamus during visually triggered and internally generated limb movements. *Journal of Neurophysiology*, *83*(5), 2780–2790. http://doi.org/10.1152/jn.2000.83.5.2780

van Polanen, V., & Davare, M. (2015). Interactions between dorsal and ventral streams for controlling skilled grasp. *Neuropsychologia*, *79*(Part B), 186–191. http://doi.org/10.1016/j.neuropsychologia.2015.07.010

Vanlandewijck, M., He, L., Mäe, M. A., Andrae, J., Ando, K., Del Gaudio, F., et al. (2018). A molecular atlas of cell types and zonation in the brain vasculature. *Nature*, *554*(7693), 475–480. http://doi.org/10.1038/nature25739

Versace, E., Martinho-Truswell, A., Kacelnik, A., & Vallortigara, G. (2018). Priors in Animal and Artificial Intelligence: Where Does Learning Begin? *Trends in Cognitive Sciences*, 1–3. http://doi.org/10.1016/j.tics.2018.07.005

Wagner, D. E., Weinreb, C., Collins, Z. M., Briggs, J. A., Megason, S. G., & Klein, A. M. (2018). Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo. *Science*, *360*(6392), 981–987. http://doi.org/10.1126/science.aar4362

Waldrop, M. M. (2002). The Dream Machine. Penguin Books.

Wang, Xiao, Allen, W. E., Wright, M. A., Sylwestrak, E. L., Samusik, N., Vesuna, S., et al. (2018). Three-dimensional intact-tissue sequencing of single-cell transcriptional states. *Science*, *361*(6400), eaat5691–18. http://doi.org/10.1126/science.aat5691

Wang, Xuhua, Liu, Y., Li, X., Zhang, Z., Yang, H., Zhang, Y., et al. (2017). Deconstruction of Corticospinal Circuits for Goal- Directed Motor Skills. *Cell*, *171*(2), 440–448.e14. http://doi.org/10.1016/j.cell.2017.08.014

Willis, T. (1664). *Cerebri Anatome*.

Wolpert, D. M., & Ghahramani, Z. (2000). Computational principles of movement neuroscience. *Nature Neuroscience*, *3 Suppl*(Supp), 1212–1217. http://doi.org/10.1038/81497

Wong, F. K., Bercsenyi, K., Sreenivasan, V., Portalés, A., Fernández-Otero, M., & Marín, O. (2018). Pyramidal cell regulation of interneuron survival sculpts cortical networks. *Nature*, *557*(7707), 668–673. http://doi.org/10.1038/s41586-018-0139-6

Wurtz, R. H., McAlonan, K., Cavanaugh, J., & Berman, R. A. (2011). Thalamic pathways for active vision. *Trends in Cognitive Sciences*, *15*(4), 177–184. http://doi.org/10.1016/j.tics.2011.02.004

Yttri, E. A., & Dudman, J. T. (2016). Opponent and bidirectional control of movement velocity in the basal ganglia. *Nature*, *533*(7603), 402–406. http://doi.org/10.1038/nature17639

Yuge, K., Kataoka, A., Yoshida, A. C., Itoh, D., Aggarwal, M., Mori, S., et al. (2010). Region-specific gene expression in early postnatal mouse thalamus. *Journal of Comparative Neurology*, *519*(3), 544–561. http://doi.org/10.1002/cne.22532

Zagha, E., Ge, X., & McCormick, D. A. (2015). Competing Neural Ensembles in Motor Cortex Gate Goal-Directed Motor Output. *Neuron*, *88*(3), 565–577. http://doi.org/10.1016/j.neuron.2015.09.044

Zeisel, A., Hochgerner, H., Lönnerberg, P., Johnsson, A., Memic, F., van der Zwan, J., et al. (2018). Molecular Architecture of the Mouse Nervous System. *Cell*, *174*(4), 999–1014.e22. http://doi.org/10.1016/j.cell.2018.06.021

Zénon, A., & Krauzlis, R. J. (2012). Attention deficits without cortical neuronal deficits. *Nature*, *489*(7416), 434–437. http://doi.org/10.1038/nature11497

Zhang, J.-Y., Lin, Y.-T., Gao, Y.-Y., Chao-Xi, Zhang, X.-B., Zhang, X.-W., & Zeng, S.-J. (2015). Distinction in the immunoreactivities of two calcium-binding proteins and neuronal birthdates in the first and higher-order somatosensory thalamic nuclei of mice: Evolutionary implications. *Journal of Comparative Neurology*, *523*(18), 2738–2751. http://doi.org/10.1002/cne.23813

Zhang, L. I., & Poo, M.-M. (2001). Electrical activity and development of neural circuits. *Nature Neuroscience*, *4*(Supp), 1207–1214. http://doi.org/10.1038/nn753