AN INVESTIGATION OF INPUT INTEGRATION IN THE MOUSE BASAL GANGLIA

by

Barbara Jeanine Hunnicutt

A Dissertation

Presented to the Neuroscience Graduate Program
at the Oregon Health & Science University School of Medicine
in partial fulfillment of the requirements
for the degree of

Doctor of Philosophy

June 5, 2015

School of Medicine
Oregon Health & Science University

CERTIFICATE OF APPROVAL

This is to certify that the Ph.D. dissertation of
BARBARA JEANINE HUNNICUTT
has been approved on June 5, 2015

Advisor, Tianyi Mao, Ph.D.

Member and Chair, Mary Heinricher, Ph.D.

Member, Gary Westbrook, M.D.

Member, John Williams, Ph.D.

Member, Claudio Mello, Ph.D.

Table of Contents

List of Figures

List of Abbreviations

General

Abbreviation	Expanded name
AIBS	Allen Institute for Brain Science
PMBA	Paxinos Mouse Brain Atlas
AMBA	Allen Institute Mouse Brain Atlas
...	

Thalamic Nuclei

Abbreviation	Expanded name	PMBA Location	AIBA Location
AD	anterodorsal nucleus	AD	
AM	anteromedial nucleus	AM+AMV	
AV	anteroventral nucleus	AV+AVDM+AVVL	
CL	central lateral nucleus	CL	
CM	central medial nucleus	CM	
IAD	interanterodorsal nucleus	IAD	
IAM	interanteromedial nucleus	IAM	
IMD	intermediodorsal nucleus	IMD	
LD	laterodorsal nucleus	LD+LDVL+LDDM	
LG	lateral geniculate nucleus	VLG+DLG+VLG+IGL	
LP	lateral posterior nucleus	LP+LPLR+LPMP+LPMC	
MD	mediodorsal nucleus	MDC+MDL+MDM	
MG	medial geniculate nucleus	MGD+MGV+MGM	
PCN	paracentral nucleus	PC+OPC	
Pf	parafascicular nucleus	Pf	
Po	posterior nucleus	Po	
PR	perireuniens nucleus	vRe	
PT	parataenial nucleus	PT	
PVT	paraventricular nucleus	PVA+PV	
Re	reuniens nucleus	Re	
Rh	rhomboid nucleus	Rh	
SGN	supragenicualte nucleus	SG	
SMT	submedius nucleus	Sub	
SPFp	subprafascicular nucleus	SPFpc	

VAL	ventral anterior-lateral complex	VA+VL
VM	ventromedial nucleus	VM
VPL	ventral posterolateral nucleus	VPL+VPLpc
VPM	ventral posteromedial nucleus	VPM+VPMpc

Cortical subdivisions

Abbreviation	Expanded name	PMBA Location	AIBA Location
AI	anterior insular cortex	pregenual (AI+DI+GI)	
Aud	auditory cortex	Au1+AuD +AuV	
dACC	dorsal anterior cingulate cortex	Cg1	
FrA	frontal association area	FrA	
IL	infralimbic cortex	IL	
Ins	insular cortex	postgenual (AI+DI+GI)	
LO	lateral orbital cortex	LO+DLO	
M1	primary motor area	M1	
M2	secondary motor area	M2	
MO	medial orbital cortex	MO	
Piri	piriform cortex	Pir	
PrL	prelimbic cortex	PrL	
Pt	parietal association cortex	MPtA+LPtA+PtPR+PtPD	
Rhi	rhinal cortex	Ect+PRh+Lent	
Rsp	retrosplenial cortex	RSA+RSG	
S1/2	sensory cortex	S1 (all sub-regions)+S2	
Tem	temporal association cortex	TeA	
vACC	ventral anterior cingulate cortex	Cg2	
Vis	visual cortex	V1+V2 (all subregions)	
vM1	vibrissal motor cortex	M2	
VO	ventral orbital cortex	VO	

Acknowledgments

The work presented here would not have been possible without the help and support of many people. First, thank you to Tianyi Mao, my thesis advisor, I have learned a lot from you and I really appreciate your continual enthusiasm about science.

I am indebted to Deniz Kusefoglu, Brian Long, & Katrina Gertz who represent the three epochs of these projects, and my graduate school career. Deniz was the only person in the lab before me, she started all of the work presented here, and I am forever grateful to her for her hard work and wonderful spirit. Brian took up this project alongside me and was an irreplaceable partner for the first chapter of this thesis. I think we complemented each other perfectly, and I learned so much from him about science and life. Katrina was with me until the end, through countless hours? days? months? years?! of often grueling image analysis. I truly could not have survived the last two years without her.

Thank you to Haining for all of your technical help, and to the rest of the current and former Mao & Zhong lab members who contributed advice and criticism along the way: Shane, Danielle, Dale, Maozhen, Guang, Sheng, Clayton, Weihong, and Josh.

I would also like to thank the people who helped me through graduate school indirectly; Danielle Robinson who is there to talk me through anything large or small, Annie Logan for being by my side from classes to dissertation and everything in-between, and Frances Lefcort for giving me the voice in the back of my head telling me that *I can do this*. To all of my friends and family, I could not have done any of this without your love and support, especially Brent, who is the person I will always aspire to be. And finally, thank you Shane Tillo, for everything.

INTRODUCTION

Overview & Rationale

The basal ganglia are a group of forebrain nuclei integral to movement, decision making, and motivation. They consist of the striatum, the globus pallidus internal (GPi) and external segments (GPe), the subthalamic nucleus (STN), and the substantia nigra pars reticulata (SNr). The basal ganglia receive inputs from nearly all of the cerebral cortex, as well as the thalamus, the hippocampal formation, the amygdala, and several brainstem nuclei (Berendse et al. 1992; Berendse & Groenewegen 1990; Wall et al. 2013; McHaffie et al. 2005). The basal ganglia then integrate and relay this information to the thalamus, cortex, and brainstem to affect behavior (Steiner & Tseng 2010).

Dysfunctions associated with diseases of the basal ganglia have provided insight into the roles that the basal ganglia play in behavior (Yin 2014; DeLong & Wichmann 2007; Albin et al. 1989). Parkinson's and Huntington's diseases highlight the importance of the basal ganglia for both movement initiation and suppression (DeLong & Wichmann 2007; Plotkin & Surmeier 2015; Redgrave et al. 2010); neuropsychiatric disorders like Tourette's syndrome and obsessive-compulsive disorder point to a role in higher cognitive functions such as impulse control and decision making (DeLong & Wichmann 2007; Tritsch & Sabatini 2012); and the well-documented involvement of the basal ganglia in addiction emphasizes their importance in motivation and reward (Tomasi & Volkow 2013; Volkow et al. 2012). Although the wide array of inputs to the basal ganglia provides the substrate for these disparate functions, we know very little about the structural and computational organization necessary to produce a unified output.

Given the importance of input integration in the striatum, understanding the precise patterns of striatal input convergence is crucial to understand information processing in the basal ganglia. In the neocortex, the repeated laminar organization and the existence of clearly defined functional subregions has allowed for the study of circuit level integration in a systematic and reproducible manner (Katz & Callaway 1992; Mao et al. 2011). In contrast to the two-dimensionally organized cortex, the striatum is an oval mass of millions of

spherical neurons (Oorschot 1996; Wilson 1987), 96% of which are one of two nearly identical, homogeneously distributed cell types (Gertler et al. 2008; Steiner & Tseng 2010), and receives overlapping inputs from almost every other brain region (Veening et al. 1980; McGeorge & Faull 1989; Berendse & Groenewegen 1990; Wall et al. 2013; Guo et al. 2015). This structural homogeneity is likely the main reason why the details of striatal function have eluded us for decades, and why the striatum remains the largest part of the telencephalon without universally accepted subdivisions.

The work presented in this dissertation details the comprehensive and three-dimensional convergence patterns for all excitatory inputs to the mouse striatum (**Chapter 2**), as well as communication routes between input types (**Chapter 1 & 2**). To compare this information to the existing literature and place the analyses in historical context, the current understanding of excitatory input distributions in the striatum, as well as the information they are thought to convey are reviewed. Additionally, to explore the possible functional implications of this dataset, the known mechanisms of striatal input processing and current theories of striatal input integration are discussed.

Basal Ganglia Organization and Function

Striatal cell types

All excitatory inputs to the basal ganglia enter through either the striatum or STN. The striatum is the primary input nucleus, with the STN receiving a smaller fraction of the total inputs, mostly from motor related cortical areas (Nambu et al. 1996). In primates, the striatum is physically separated by the internal capsule into the caudate nucleus and the putamen, but in rodents the axons of the internal capsule traverse through the merged caudoputamen and create the bundles that give the striatum it's characteristic striated appearance (Steiner & Tseng 2010). Approximately 95% of neurons in the striatum are GABAergic medium spiny neurons (MSNs), which are also the only striatal output neurons. The remaining 5% of the neurons are either cholinergic (1-2%) or GABAergic (3-4%) interneurons (Steiner & Tseng 2010).

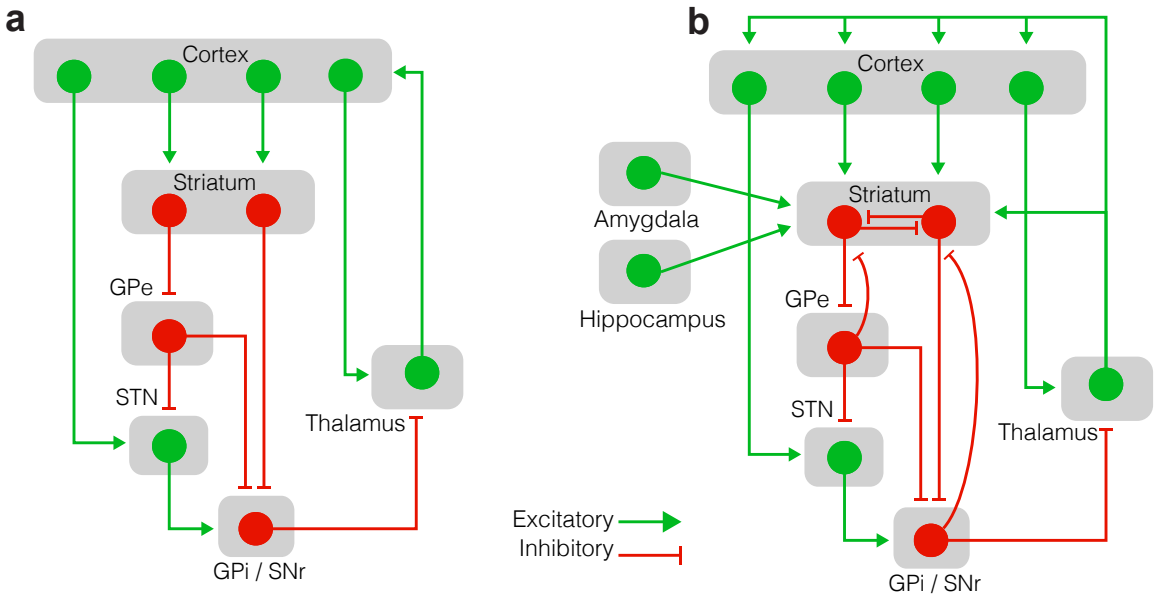


Figure 1 Overview of basal ganglia circuitry (**a**) Illustration of information flow through the basal ganglia (**b**) A more accurate illustration of the basal ganglia-thalamo-cortical loop, showing feed-back and feed-forward excitation and inhibition, as well as a wider array of striatal input convergence. Excitatory and inhibitory connections are shown in green and red, respectively.

MSNs can be separated into two subtypes, one expressing dopamine type-1 receptors (D1) and muscarinic M4 receptors, and the other expressing dopamine type-2 receptors (D2) and adenosine A2A receptors (Kreitzer 2009). These distinctions have both physiological and anatomical consequences. D1 receptors are g-protein coupled receptors that couple to $G\alpha s$ and activate adenylyl cyclase, and D2 receptors couple to $G\alpha i$ which inhibits adenylyl cyclase. Dopamine, a primary and ubiquitous neuromodulatory input to the striatum, has opposing effects on the excitability and synaptic function of the two MSN subtypes (Surmeier et al. 2007). Although the D1 and D2 receptor expressing MSNs are distributed homogeneously throughout the striatum, they have clearly segregated output pathways. D1-MSNs comprise the direct pathway, which send their axons directly to the output nuclei of the basal ganglia, the GPi and SNr, and D2-MSNs send projections to intermediate nuclei, the GPe and STN, creating the so-called indirect pathway (**Fig. 1a**). With the exception of the STN, which tonically excites the output nuclei, all basal ganglia projection neurons are inhibitory. The predominant theory of information transfer out of the basal ganglia, is the release of downstream targets from tonic inhibition (Steiner & Tseng 2010). Via this mechanism, an "excitatory" signal from the basal ganglia is a cessation of activity, releasing

the thalamus and brainstem from an inhibited state. The series of inhibitory connections between the striatum and the basal ganglia output nuclei result in the activation of the direct pathway having a net excitatory effect on the thalamus and cortex, and indirect pathway activation exerts a net inhibitory effect (**Fig. 1**).

Striatal output: mechanisms of behavioral control

The organization and basic physiology of the direct and indirect pathways has guided a long standing theory of basal ganglia function in which the balance of these two output pathways controls behavioral output (Albin et al. 1989). Biasing the direct pathway facilitates current behavior, and biasing the indirect pathway results in the inhibition of that behavioral pattern. This basic principle has been confirmed experimentally. Direct pathway activation in the striatum facilitates movement initiation and can enhance learning reinforcement, and indirect pathway activation leads to movement suppression and mimics punishment in learning paradigms (Kravitz et al. 2010; Kravitz et al. 2012). Although exogenous activation of these pathways is sufficient to alter behavior, it does not prove whether pathway biasing is a normal mechanism employed by the basal ganglia to control behavioral output (Nelson & Kreitzer 2014). Direct measurements of activity in the two pathways, recorded during behavioral tasks and in disease states, show that the balance of the two pathways does not seem to correlate with any aspect of behavior (Cui et al. 2013; Liang et al. 2008). Both direct and indirect pathway neurons simultaneously increase their activity levels during action initiation (Cui et al. 2013), they both positively modulate their responses during the cessation of movement, and both direct and indirect pathway neurons have the ability to alter their activity patterns in similar ways in response to reward expectation(Isomura et al. 2013). However, the direct and indirect pathways do exhibit differential changes in their activity *patterns* during action suppression and maintenance (Jin et al. 2014). These findings imply a much more complicated method of information processing in the basal ganglia than suggested by the initial hypotheses (Albin et al. 1989; Nelson & Kreitzer 2014). Since both striatal output pathways are activated during basal ganglia related behaviors, the mechanisms

that determine basal ganglia output are likely to be more dependent upon the precise connectivity patterns and timing of striatal inputs than on global activation imbalances across MSN subtypes.

Information processing through the basal ganglia

The proposed importance of striatal input patterning on the control of basal ganglia output is even more convincing when you look at the enormous convergence of information onto each MSN. The striatum is generally accepted as a site of input integration and not simply as a relay of inputs to downstream nuclei of the basal ganglia. This is due to the highly interconnected nature of MSNs (Kawaguchi et al. 1989; Plenz 2003), as well as the massive decrease in the number of total neurons as you move from the striatum to downstream nuclei. The striatum is the largest nucleus of the basal ganglia, containing approximately 5.5 million neurons bilaterally in the rat, all of the other nuclei combined consist of only ~200,000 neurons (Oorschot 1996). This difference is even more pronounced in humans with 100 million striatal neurons, and less than 500,000 neurons combined in the downstream nuclei of the basal ganglia (Steiner & Tseng 2010). The ratio of striatal input neurons to output neurons in the SNr and GPi is 30:1 in the rat, and 100:1 in the human. This lopsided distribution of input and output neurons accounts for an incredible amount information downsampling in the basal ganglia, but there is an even greater degree of input convergence at the cellular level. It has been estimated from electron microscopic data that within the neuropil occupied by an MSNs dendritic arbor, there are approximately 15 million corticostriatal synapses and 2845 overlapping MSNs, but a single cortical axon will make a maximum of 40 synapses in that space (Kincaid et al. 1998). Based on these estimates, each MSN receives inputs from roughly 5000 distinct cortical neurons, and this only accounts for half of the total excitatory input to a single MSN, since they also receive excitatory inputs from the thalamus, amygdala, and hippocampus (Huerta-Ocampo et al. 2013). When this information is combined with the convergence of MSNs on their

downstream nuclei, the total number of striatal inputs get consolidated by more than 4 orders of magnitude as the information is passed through the basal ganglia.

Striatal Input distribution at the cellular level

Conceptually, the distribution of inputs onto MSNs could be arranged in countless ways, but it's useful to consider two extreme scenarios. Inputs to MSNs could be arranged in a highly redundant fashion that allows for a given set of inputs to activate multiple MSNs simultaneously, which could then signal as functional subunits to downstream nuclei (**Fig. 2a**). Alternatively, the individual inputs to any given MSN could be unique, thereby allowing them to function as pattern detectors on an individual basis (**Fig. 2b**) (Steiner & Tseng 2010). Using the connectivity rates for corticostriatal neurons described above (Kincaid et al. 1998), the likelihood that any two MSNs will form synapses with the same corticostriatal axon can be determined. By expanding this to all possible connections, the cumulative probability distribution for co-innervation shows that there is almost no chance that more than 100 out of the 5000 cortical inputs received by a given MSN will be shared by any other MSN (Wilson 2000; Steiner & Tseng 2010). This supports the innervation pattern shown in **Figure 2b** as the estimation of input connectivity onto MSNs. Importantly, this arrangement allows for at least two distinct integration steps for information entering the basal ganglia before it reaches the output nuclei. The combination of inputs required to depolarize a single MSN to fire an action potential constitutes one integration step and is unique to each MSN (**Fig. 2b**), then the axons from many MSNs converge on cells in

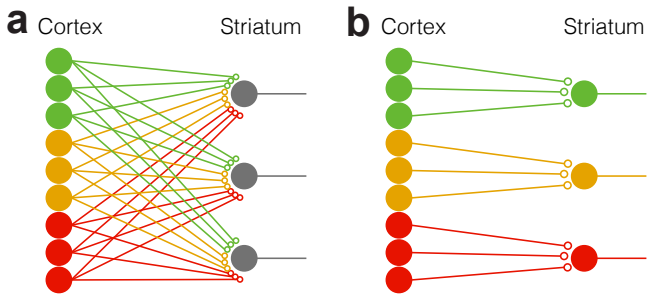


Figure 2 Hypothetical corticostriatal input distribution patterns at the cellular level. (**a**) Highly distributed input convergence across many striatal neurons. (**b**) Tarteted connectivity between cortical axons and their striatal targets. Adapted from: Steiner & Tseng, *Basal Ganglia Handbook of Structure and Function*, 2010.

downstream nuclei to set up the second integration step (Kawaguchi et al. 1990; Steiner & Tseng 2010) (**Fig. 1**).

Mesoscopic striatal input distribution overview

Mapping the unique innervation pattern of every MSN is not currently experimentally feasible. However, the total set of possible inputs to each MSN is not uniform across the striatum because no individual cortical or thalamic subregion projects to the entire striatal volume. Instead, projections from a single subregion of the cortex or thalamus will cover a portion of the striatum, overlapping with the projection fields of several other striatal inputs (McGeorge & Faull 1989; Berendse & Groenewegen 1990; Veening et al. 1980). This incomplete overlap of projection fields means that the possible combinations of inputs shifts almost continuously from one region of the striatum to the next, providing a vast opportunity for input integration.

Current maps of striatal input distributions are assimilated from a massive number of studies spanning over 50 years with techniques ranging from cellular ablation to cell-type specific viral infection (Webster 1961; Guo et al. 2015). Attempts to assimilate this data to create a comprehensive, three-dimensional picture of all of the excitatory inputs to the striatum have produced important insights into the broad architecture of the network, but have been fragmented and suffered from the inevitable loss of information that occurs when generalizations are made. Summary projection maps tend to describe unique projection fields for various inputs, even though there is almost always significant overlap in the few examples provided from the original data sets (Y. Smith et al. 2004; Berendse et al. 1992; Berendse & Groenewegen 1990; McGeorge & Faull 1989; Willuhn et al. 2003). More recent studies of striatal inputs have presented thorough and detailed information about input convergence from the entire brain, but only to a very limited portion of the anterior dorsal striatum (Wall et al. 2013; Guo et al. 2015). Although incomplete, these datasets are useful to understand mechanisms of input integration and basal ganglia function.

Organization and Function of Striatal Inputs

Corticostriatal inputs

The striatum receives excitatory glutamatergic inputs from all subregions of cerebral cortex except olfactory areas (McGeorge & Faull 1989). These inputs project to both the ipsilateral and contralateral striatum, synapsing primarily on the dendritic spines of MSNs, but also contact striatal interneurons (Doig et al. 2010; Guo et al. 2015; Reiner et al. 2010). Cortical inputs to striatum originate from at least two distinct cellular populations, the intratelencephalic (IT) and pyramidal tract (PT) type. IT-type neurons project only within the telencephalon and send axons either ipsilaterally or bilaterally in the striatum. These neurons originate primarily in cortical layers 3 and 5a. and are thought to carry sensory and motor planning information to the dorsolateral striatum where they preferentially innervate the direct pathway (Lei et al. 2004; Reiner et al. 2010). PT-type neurons originate in cortical layer 5b and send axons to the brainstem and spinal cord with axon collaterals projecting to the ipsilateral striatum where they preferentially innervate indirect pathway neurons (Lei et al. 2004), and are thought to convey an efference copy of motor commands (Wilson 1987; Reiner et al. 2010).

The axonal arbors of IT-type neurons can span up to 1 mm in the striatum, while PT-type axons have more focal projections of <500 μm (Cowan & Wilson 1994). These projections patterns have implications for striatal function. For instance, the large projection field of IT-type neurons puts them in a position to transmit information across a wider striatal volume, perhaps across modalities, and the unilateral, focal projection of a PT-type neuron could have a more directed role in striatal signaling. This balance between focal and distributed projections varies across cortical subregions, which may be important for the larger question of how input convergence and/or segregation from different subregions generates behavioral output (Reiner et al. 2010).

The most extensively studied excitatory inputs to the basal ganglia are those of the prefrontal cortex (PFC), which project to extensive, overlapping striatal volumes (Berendse et al. 1992; Haber 2006; Mailly et al. 2013). Subregions of PFC project in a generally

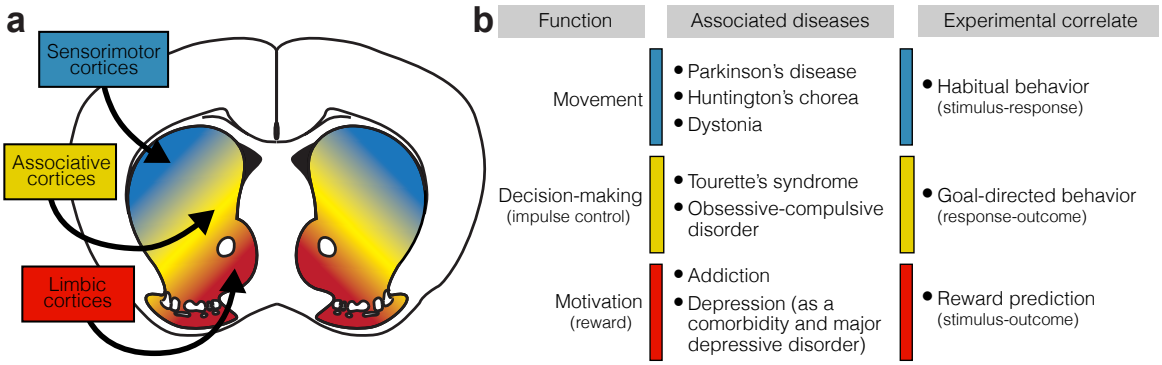


Figure 3 Overview of tripartite domain system in the striatum (**a**) Generalized distribution patterns for cortico-striatal projections. Projection fields broadly subdivide the striatum into the sensorimotor (blue), associative (yellow), and limbic (red) domains. (**b**) Associated functions (left), diseases (center), and behaviors (right) with each functional domain.

topographic manner so that ventral medial PFC (vmPFC), equivalent to infralimbic (IL) and ventral prelimbic (PrL) in rodents, projects to ventromedial striatum, and areas with increasing distance from vmPFC project in a dorsomedial to ventrolateral distribution ever more dorsolaterally (**Fig. 3a**), until the most dorsolateral portion of the striatum is dominated by inputs from motor cortices (Haber 2006; Euston et al. 2012; Voorn et al. 2004). There are known subregions with extensive overlap, such as those between dorsal anterior cingulate (dACC) and ventral/ lateral orbital (VLO) (Berendse et al. 1992), but no two areas have identical projection patterns (Steiner & Tseng 2010). Inputs from posterior cortical areas are not as well described, but connections have been demonstrated from all neocortical and mesocortical areas, as well as allocortical areas including the amygdala and subiculum (McGeorge & Faull 1989; Guo et al. 2015; Groenewegen et al. 1987; Novejarque et al. 2011).

Decades of anatomical data show that each striatal input converges with other cortical and subcortical projections (Webster 1961; Wall et al. 2013). The complete projection distribution of each input type reveals an organization principle in which functionally related inputs are more likely to have overlapping projection fields. When all cortical inputs to the striatum are taken into account, the projection overlap roughly divides the striatum into three functional domains, the limbic, associative, and sensorimotor domains, which are thought to constitute segregated circuits. The limbic domain integrates information on

affective state, motivation and reward, the associative domain takes in a variety of higher cognitive inputs, and the sensorimotor domain is the site of sensory and motor information convergence (**Fig. 3**) (Parent & Hazrati 1995; Joel & Weiner 1994; Gruber & McDonald 2012). The limbic striatum is located ventrally in both rodents and primates and encompasses the nucleus accumbens (NAc), the sensorimotor striatum is located dorsolaterally and falls primarily in the posterior putamen of primates, while the associative striatum occupies a poorly defined area between the limbic and sensorimotor domains but is generally considered to be the dorsomedial striatum in rodents and the caudate and anterior putamen of primates (**Fig. 3a**).

The Limbic Striatum

The cortical areas that send inputs to the limbic striatum, i.e. the NAc, are frontal regions known in the primate to mediate reward, motivation, and affect regulation (Haber & Knutson 2010; Euston et al. 2012). Animal studies of the NAc have shown it to be important for aspects of incentive based learning as well as motivated behavior, where it likely integrates information about context, motivation, and reward to influence behavioral output via other brain areas, functions consistent with its association to addictive behaviors (**Fig. 3b**) (Steiner & Tseng 2010; Yin & Knowlton 2006). The NAc is commonly separated into the core and shell, which have slightly different frontal input characteristics, but differences appear to fall along the ventromedial-dorsolateral gradient that spans the entire striatum. The core and shell mediate distinct functions, but this is likely due to unique NAc shell output targets relative to areas targeted by the rest of the striatum, and not due to differences between the core and shell in input distribution or the local circuit (Voorn et al. 2004). In addition to inputs from PFC, the limbic striatum receives extensive excitatory inputs from amygdala and hippocampus, as well as dopaminergic innervation from the ventral tagmental area (VTA). To understand the roles that hippocampal and amygdalar inputs play in striatal function, it is important to understand their local circuits.

The amygdala is highly interconnected with cortical and subcortical targets, including the basal ganglia. The amygdala receives inputs from every major sensory modality, making it a prime location for multisensory integration, and is also reciprocally connected to the hypothalamus, brainstem, and hippocampus. The autonomic connections allow the amygdala to both assess and influence systems that monitor and control affective states. This combination of inputs bestows the amygdala with the ability to directly associate incoming sensory information with affective states. Experimental evidence supports this hypothesis, showing that the amygdala is sufficient, and necessary, to make stimulus-outcome associations during instrumental learning tasks, i.e. Pavlovian associations (Stuber et al. 2011; Gruber & McDonald 2012). Specifically, as an input to the limbic striatum, activation of the basolateral amygdala (BLA) reinforces reward seeking behavior (Stuber et al. 2011), and neurons in the NAc core fire in response to reward predictive signals, which may be mediated by the Pavlovian associations relayed from the amygdala (Gruber & McDonald 2012). Damage to the NAc core decreases the overall response rate to a conditioned stimulus. This effect can be stimulus non-specific, so it is interpreted as a global decrease in motivation and mediated by affective information from the amygdala. Targeted dopamine depletion in the limbic striatum does not effect response rates, making it unlikely that this effect is mediated by dopaminergic neuromodulation (Gruber & McDonald 2012; Wise 2009).

Hippocampal structures are broadly thought to assess and store relational information about contextual, spatial and multimodal sensory aspects of episodic experiences (Gruber & McDonald 2012). The hippocampus receives inputs from neocortical areas through entorhinal and perirhinal cortices, including inputs from all sensory modalities. This robust sensory information is relayed to the dentate gyrus and then processed through the rest of the hippocampus to create a contextual and temporal record of the environment (Bird & Burgess 2008). Hippocampal outputs are sent to the subiculum where they are either relayed to subcortical areas or fed back to the entorhinal cortex. Behavioral studies have verified that the hippocampus is important for the formation and recall of associations, and specifically

with regard to the network formed with the frontal cortex and basal ganglia, associations between antecedents and outcomes. Lesion studies in the limbic striatum have identified strong contextual and spatial components to reward-seeking behavior, which is likely mediated by hippocampal inputs. Experimentally, the limbic striatum is necessary for animals to spatially orient their bodies toward and approach task-related stimuli, as well as form contextual associations between environments with positive or negative affect, i.e. conditioned place preference (Gruber & McDonald 2012).

The Sensorimotor Striatum

The primary function of the sensorimotor domain of the dorsolateral striatum is thought to be the acquisition and execution of stimulus-response relationships, i.e. habitual behavior (**Fig. 3b**). This aspect of instrumental learning is, by definition, uncoupled from reward value and outcome anticipation (Yin & Knowlton 2006). Anatomically, this behavioral distinction makes sense, as the dorsolateral striatum does not receive inputs from the amygdala or hippocampus, which carry affective and outcome associations respectively (see introduction to the *Limbic Striatum*), but does receive extensive and topographically overlapping sensory and motor input (Malach & Graybiel 1986; Nambu 2011). It is also interesting to note that dopaminergic inputs to the dorsolateral striatum are preferentially susceptible to degradation in the initial stages of Parkinson's disease, and early motor symptoms in patients are biased towards the automatic components of movement as well as an impaired ability to form habits (Redgrave et al. 2010). The outputs of this circuit eventually reach the motor cortices and brainstem motor networks to regulate movement (Yin 2014).

The Associative Striatum

The associative (i.e cognitive) domain of the dorsomedial striatum has been linked to a variety of functions including behavioral flexibility, allocentric navigation, and goal-directed forms of instrumental learning (**Fig. 3b**) (Gruber & McDonald 2012; Yin & Knowlton

2006). The navigation functions of the associative striatum are likely mediated by the wide range of hippocampal inputs that it receives (see introduction to the *Limbic Striatum*) . There are direct projections from the entorhinal cortex and subiculum, as well as indirect hippocampal information from two primary entorhinal targets, the mPFC and the retrosplenial cortex. Lesions in either the direct or indirect routes of information between the hippocampal formation and striatum impairs an animal's ability to adapt their behavior to changes in their spatial environment, but does not abrogate their ability to navigate completely, providing evidence for a role in the integration of spatial information for response flexibility (Gruber & McDonald 2012). Lesions in the associative striatum also impair non-spatial response flexibility when the modality of a cue is changed, possibly mediated by inputs from sensory modalities directly or indirectly via inputs from other multimodal areas (Khibnik et al. 2014; Gruber & McDonald 2012). The dorsomedial striatum is important for the outcome prediction required to make response-outcome relationships in goal-directed instrumental learning tasks, which is likely mediated by inputs from both the amygdala and hippocampus (Yin 2014). There are also inputs from many of the thalamic targets of basal ganglia output, providing it with direct feedback of internal behavioral states that could help in outcome prediction (Pan et al. 2010; Deniau & Chevalier 1992). Finally, the associative striatum is also strongly modulated by reward anticipation in all of these situations, making this domain generally important for monitoring recent actions and stimuli to anticipate their consequences (Yin & Knowlton 2006).

Thalamostriatal inputs

There has been a conspicuous lack of information regarding the role of the second largest source of excitatory input to the striatum, the thalamus (Huerta-Ocampo et al. 2013). Although the existence of thalamostriatal projections were definitively identified in the 50's (Cowan & Powell 1956), the importance of the thalamostriatal projections in basal ganglia driven behavior was largely ignored until recently. Previous theories of basal ganglia function defined the thalamus as a relay for information traveling from the output nuclei to the

cortex, and did not recognize it as an influential driver of striatal activity (Joel & Weiner 1994; Y. Smith et al. 2004; Y. Smith, Galvan, et al. 2014a) (**Fig. 1b**). Still, the function of thalamostriatal innervation is poorly understood. The proposed roles for each thalamic projection are generally only speculation based on the information content being transmitted (Jones 2007; Watson et al. 2012). For instance, thalamostriatal inputs that receive information from basal ganglia output nuclei may act as a positive feedback to facilitate the inputs they receive, or as a negative feedback loop to inhibit other pathways via D2-MSN activation. This would function to enable selected, and suppress unselected, behaviors.

Thalamic nuclei that relay sensory inputs may function to redirect attention based on salient sensory information. Similarly, the intralaminar nuclei (see below) receive information from a wide range of cortical and subcortical brain regions including a massive input from the medullary reticular formation, which regulates arousal and attention (Y. Smith, Galvan, et al. 2014a; Van der Werf et al. 2002). Based on this, as well as human studies where intralaminar activation results in "heightened attention", it is hypothesized that intralaminar thalamic inputs function in the striatum to shift attention towards relevant behaviors (Haber & McFarland 2001). The precise mechanism for these behavioral shifts is not clear, but thalamostriatal inputs preferentially activate indirect pathway neurons via cholinergic interneurons, which could allow for the selective suppression of non-relevant pathways (Ding et al. 2010).

The intralaminar and midline nuclei are the most thoroughly studied thalamic sources of striatal input, and together project to the majority of the striatal volume. The central medial (CM), paracentral (PCN), and central lateral (CL) nuclei make up the rostral intralaminar group, and the caudal intralaminar group is made up of the centromedian–parafascicular nuclear complex, which are separate nuclei in primates (CM/Pf) and fused in rodents (PF) (Y. Smith et al. 2004; Van der Werf et al. 2002). PF is unique in relation to other intralaminar and midline nuclei, in that it sends the majority of its afferents to the striatum instead of the cortex, in contrast to the other nuclei which have equal or cortically biased distributions (Y. Smith, Galvan, et al. 2014a). PF projects to the majority of the striatum, with lateral PF

innervating dorsolateral striatum and medial PF innervating ventral and medial striatum (Groenewegen & Berendse 1994). The midline nuclear group, made up of the paraventricular (PVT), parataenial (PT), interanterodorsal (IAD), interanteromedial (IAM), rhomboid (Rh) and reuniens (Re) nuclei, project primarily to the ventral striatum (Van der Werf et al. 2002; Li & Kirouac 2008). The midline and intralaminar nuclei are thought to project topographically to the striatum so that they overlap with inputs from the cortical area they are also directly connected to, maintaining the functional distinction of striatal regions defined by their cortical inputs (Haber & Calzavara 2009). These projections have been definitively characterized, however the full projection distribution for each of these nuclei is difficult to delineate, as the published data has been synthesized into summary figures that are inconsistent with examples of the raw data provided.

Although considerably less well characterized, many other thalamic nuclei send axons to the striatum. Several of the primarily sensory and motor related nuclei of the ventral and posterior groups; ventromedial (VM), ventral anterolateral (VAL), and posterior (Po) nuclei, send broad, topographic projections to the dorsal striatum (Erro et al. 2001; Veening et al. 1980), but there are inconsistent reports of thalamostriatal projections from the ventral posterolateral (VPL) and ventral posteromedial (VPM) nuclei (Erro et al. 2001; Pan et al. 2010). The medial nuclei, mediodorsal (MD) and intermediodorsal (IMD), form a primary thalamic projection to frontal cortex, which also send dense collaterals to the ventral and medial striatum (Groenewegen 1988; Veening et al. 1980). The anterior group, consisting of anteromedial (AM), anteroventral (AV), and anterodorsal (AD) nuclei, which primarily project to cingulate and retrosplenial cortices, have been passingly said to project to the striatum (Shibata 1993; van Groen & Wyss 1995), and have also been inconsistently found in more recent striatal projection characterizations (Guo et al. 2015; Wall et al. 2013; Pan et al. 2010). The lateral nuclei, lateral dorsal (LD) and lateral posterior (LP) nuclei, which have a wide range of posterior cortical afferents, have both been reported to project to the central striatum (Kamishina et al. 2009; Pan et al. 2010; Guo et al. 2015). Finally, the lateral and medial geniculate nuclei (LG and MG), which relay visual and auditory inputs respectively,

have controversial reports on producing striatal afferents, however it seems that the MG and not the LG projects to the caudal striatum (Ledoux et al. 1984; Veening et al. 1980). This list, while comprehensive, is very crude. The existence and distribution of many of these thalamostriatal projections are variably reported, and most have, at best, incomplete striatal projection maps.

A subgroup of these less studied nuclei, VAL, VM, and MD, are the thalamic targets of basal ganglia output, and important to examine further (Deniau & Chevalier 1992; Ilinsky et al. 1985; McFarland & Haber 2002; McHaffie et al. 2005). VAL, which is divided into VA and VL in primates, is reciprocally connected to motor, sensory, posterior cingulate, and retrosplenial cortices (Haber & Calzavara 2009; Mitchell & Cauller 2001). MD has reciprocal connections with all of prefrontal cortex (Ray & Price 1992). In fact, the cortical distribution of projections originating in MD is one characteristic used to define the boundaries of rodent prefrontal cortex (Mitchell & Cauller 2001). The final thalamic output target, VM, projects to layer 1 of almost the entire cortex, with some additional projections to the deeper layers of pregenual cortical areas (Herkenham 1979; Mitchell & Cauller 2001). Layer 1 contains very few cells, but has dense arborizations from the apical dendrites of cells residing in deeper layers, which makes the thalamocortical projections from VM particularly well positioned to alter the excitability of all layers in a given cortical area (Rubio-Garrido et al. 2009). The functional role of these projections is uncertain, but VM is poised to powerfully affect the excitability of the entire cortex. There are also inconsistent reports of PF being a target of basal ganglia output (Y. Smith, Wichmann, et al. 2014b; Haber & Calzavara 2009; Deniau & Chevalier 1992), but if it is, then its connectivity bias with the striatum over cortex and relationship to the reticular activating system would allow it to provide the striatum with integrated information on basal ganglia output and attention.

Mechanisms For Information Integration Across Parallel Circuits in the Basal Ganglia

Parallel vs integrative circuit overview

The topographic distribution of information that enters the basal ganglia via the striatum is thought to be grossly maintained as information travels through the basal ganglia, to the thalamus, and back to cortex (Steiner & Tseng 2010). The theory of a strict parallel organization of functionally related loops in the basal ganglia was pioneered several decades ago (Alexander et al. 1986; Parent & Hazrati 1995), but has since been criticized for it's inability to allow for the coordination of separate modalities during complex behavior (Joel & Weiner 1994; Steiner & Tseng 2010). For example, the limbic striatum is not directly connected to the motor cortex, but it is known that motivation and reward anticipation affect behavioral output (Yin & Knowlton 2006). To accommodate communication across the motor, associative, and limbic striatum, the circuit is now generally considered to be organized in parallel loops with mechanisms for cross-domain communication.

Cross-domain integration during cortical input integration

At the level of inputs there are two general categories of input integration considered to account for integration across functional domains. First, as has been thoroughly discussed, striatal inputs do not obey any hard boundaries in regard to the sensorimotor, associative, and limbic striatum. Instead, there is a distribution of slightly overlapping projections that slowly biases the combination of inputs attributed to each domain (Groenewegen et al. 1990; Steiner & Tseng 2010). If these functional domains are the final word in striatal organization, then the overlap at boundaries could be a mechanism of cross-domain integration. Second, corticostriatal inputs have a characteristic projection pattern, with a dense innervation zone that follows the topographic organization of the domain system, but they also have a diffuse projection that spans large striatal volumes and does not obey the boundaries of domains (Mailly et al. 2013). The dense projections form complex convergence patterns that have only really been investigated crudely at the level of the tripartate domain, and in isolated studies (Reig & Silberberg 2014). Given the relatively sparse innervation from single corticostriatal axons (see Introduction to: *Information processing*

through the basal ganglia), it may be difficult for the diffuse corticostriatal inputs to drive MSN activity on their own, and convergence from other inputs may be obligatory for their function. This projection organization, consisting of targeted dense projections and convergent diffuse projections, would be a straight forward set up for parallel loops that communicate across domains, with the dense projections being poised to drive MSN activity in a region specific manner, and the diffuse projections acting as the integrators.

Cross-domain integration in the globus pallidus

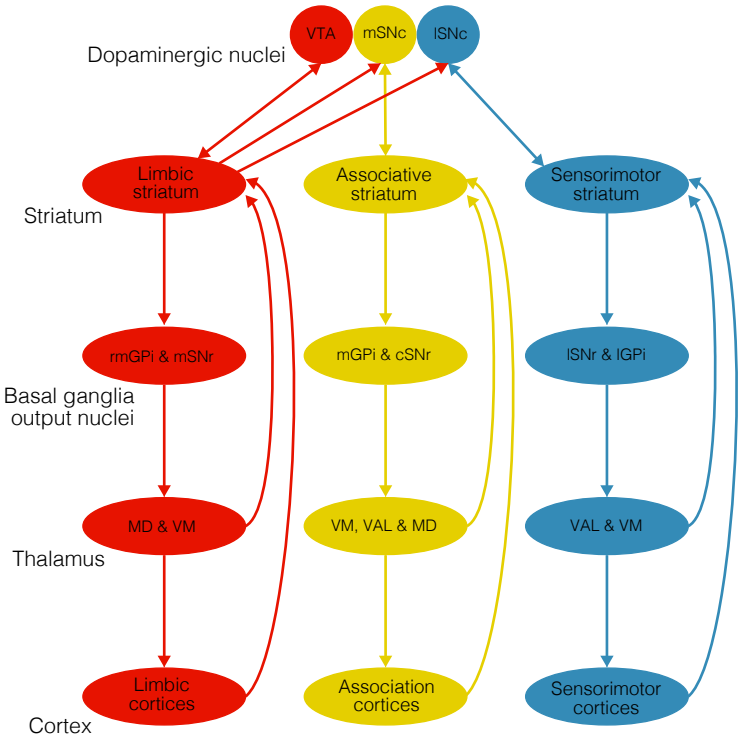


Figure 4 Segregated information flow from the limbic, associative, and sensorimotor domains of the striatum to the basal ganglia output nuclei, thalamus, and cortex. Proposed cross-domain communication via the dopaminergic nuclei is also illustrated.

Striatal projections to the two segments of the globus pallidus, the GPe (the primary target of the indirect pathway) and the GPi (a target of the direct pathway and one of the basal ganglia output nuclei), are topographically organized, preserving the three-dimensional organization of inputs to the striatum, and therefore the functional domains described previously in both the indirect and direct pathways (**Fig. 4**). The lateral GPe and GPi receive

inputs from the sensorimotor striatum, the medial GPe and GPi receive inputs from the associative striatum, and the ventromedial edges of the GPe and GPi are innervated by the limbic striatum (Parent & Hazrati 1995; Joel & Weiner 1994). Palladial neurons have long dendrites that are not confined to the putative functional domains created by the topographic distribution of striatal inputs. The proximal dendrites of palladial neurons receive topographic input from the striatum, but the distal dendrites can receive input from other functional regions. Interestingly, these palladial neurons send projections back to the striatum that are both reciprocal and non-reciprocal (Steiner & Tseng 2010). This cross-domain integration and non-reciprocal connectivity points to the palladium as a likely site of input integration, both within and across functional domains.

Cross-domain integration via the substantia nigra

The striatal projections to both parts of the substantia nigra are also topographic. The projections to the SNr, the other basal ganglia output nucleus, are arranged so that the sensorimotor striatum projects to the lateral SNr, the limbic striatum projects to the medial SNr, and the associative striatum innervates a central area in the SNr (Joel & Weiner 2000) (**Fig. 4**). The substantia nigra pars compacta (SNc), along with the VTA, provide dense dopaminergic innervation to the striatum, but also receive striatal input. In rodents the limbic striatum projects to the VTA and both the medial and lateral SNc, the associative striatum only projects to the medial SNc, and the sensorimotor striatum projects to the lateral SNc (Joel & Weiner 2000). In primates, the nigral projections from the associative striatum overlap with those from the sensorimotor striatum, which may be a functional difference across species, or could be due to differences in the demarcation of domain boundaries (Haber et al. 2000). The organization of this pathway is proposed to be a major source of cross-talk between striatal subregions. Haber and colleagues found that the striato-nigro-striatal subcircuit, i.e. a projection from the striatum to SNc/VTA that then projects back to the striatum, is arranged in an ascending loop. MSN input from the NAc shell causes an indirect activation of dopaminergic neurons that project to the NAc core, then the NAc

core influences the central striatum, and the central striatum influences the dorsolateral striatum (Haber et al. 2000). The loop organization is not as robust in rodents, but the gradual dorsolateral shift in dopaminergic feedback is consistent. In both primates and rodents, the ventral striatum influences a wide range of dopamine neurons, but receives inputs from fewer dopamine cells than it contacts, and the dorsolateral striatum influences a small range of dopamine neurons and receives dopaminergic innervation from a relatively large swath of the SNc/VTA (Steiner & Tseng 2010). This organization allows information to propagate across striatal domains so that inputs related to motivation and cognition can influence motor decision making.

Since the feed-forward dopaminergic innervation from the limbic striatum reaches almost the entire striatum in the rodent, a mechanism has been proposed by which the limbic striatum essentially acts as a behavioral filter for the rest of the striatum (Joel & Weiner 2000). The limbic striatum has inputs capable of assigning affective value to stimuli in a context dependent manner (see Introduction to: *The Limbic Striatum*). These value judgements could then be passed on to the associative and sensorimotor domains of the striatum through a combination of lateral inhibition within the striatum and via feed-forward neuromodulation from the dopamine system to bias competition between either goal-directed or habitual responses (Gruber & McDonald 2012).

Cross-domain integration in the thalamus

The output from the basal ganglia to the thalamus maintains the general topographic distribution set up by cortical inputs to the striatum (Haber & McFarland 2001). The ventral and medial SNr and GPi project to MD and VM, and the lateral portions project primarily to VAL and VM (Herkenham 1979; Haber & Calzavara 2009; Deniau & Chevalier 1992; Y. Smith, Galvan, et al. 2014a). Since VAL innervates sensorimotor areas and MD targets PFC (see Introduction to: *Thalamostriatal inputs*), this constitutes feedback to the same cortical areas that originally provided input to each information stream. The projections from VAL and MD have also been shown to project to areas in the striatum that receive convergent

input from the same cortical subregions that VAL and MD innervate directly (McFarland & Haber 2000). This evidence supports a maintenance of parallel information pathways in thalamocortical and thalamostriatal projections. There is, however, evidence in primates that these nuclei receive non-reciprocal corticothalamic inputs, which primarily originate in prefrontal areas (McFarland & Haber 2002). In summary, while the broad topographic arrangement of information appears to remain segregated through the thalamus as it is relayed to the cortex and fed back to the striatum, there may be integration of higher cognitive information during these steps.

Summary & Motivation

The two main goals of this introduction were to provide the anatomical and behavioral background necessary to hypothesize how patterns of input integration in the striatum could account for the known behavioral functions of the basal ganglia, and highlight the incomplete state of our current knowledge of striatal anatomy. While we have extensive data on the existence of excitatory inputs from nearly the entire cortex, thalamus, amygdala, and hippocampal formation to the striatum, we know very little about full input distributions and input convergence patterns.

Broadly, information flow through the basal ganglia begins with excitatory inputs entering the striatum from the cortex and thalamus, this information is passed to downstream nuclei within the basal ganglia, then to the thalamus, and finally back to the cortex (**Fig. 4**) (Steiner & Tseng 2010). Although it is centrally positioned as both a primary input and output target, the thalamus is the least characterized part of this circuit in the current literature. For this reason, we first focused on creating a map of the thalamic connections with the cortex and striatum (**Chapters 1 & 2**). **Chapter 1** presents a comprehensive map of the projections from the thalamus to the cortex. We developed a novel localization method which allowed us to computationally align 254 separate experiments and generate the most complete map of thalamocortical projections ever created. **Chapter 2** describes the complete distribution of cortical inputs to the mouse

striatum, and identifies the thalamostriatal convergence patterns with each cortical subregion. This information subsequently allowed us to objectively segment the striatum into functionally distinct subregions. The work presented in this dissertation details the comprehensive and three-dimensional convergence patterns for all excitatory inputs to the mouse striatum (**Chapter 2**), as well as communication routes between input modalities (**Chapters 1 & 2**), shedding light on mechanisms of input convergence and information flow through the cortico-thalamo-basal ganglia circuit with unprecedented breadth and detail.

CHAPTER 1:
A COMPREHENSIVE THALAMOCORTICAL MAP AT THE MESOSCOPIC LEVEL

Barbara J. Hunnicutt[1-3], Brian R. Long[1-2,4], Deniz Kusefoglu[2-3], Katrina J. Gertz[2], Haining Zhong[2], and Tianyi Mao[2]

[1] These authors contributed equally to this work

[2] Vollum Institute, Oregon Health and Science University, Portland, Oregon.

[3] Neuroscience Graduate Program, Oregon Health and Science University, Portland, Oregon.

[4] Present address: Allen Institute for Brain Science, Seattle, Washington.

Presented as published: *Nature Neuroscience* 17, 1276–1285 (2014) doi:10.1038/nn.3780

Forward

All of the authors participated in designing the experiments. The initial third of the viral tracer injections, imaging, and some exploratory analyses were performed solely by Deniz Kusefoglu, the next third were done by myself and Deniz, and the remaining injections and imaging were performed by myself and Katrina Gertz. I did all of the manual image segmentation, manual analysis and alignment steps. Tianyi Mao, Katrina Gertz, and I performed blind replications of the manual scoring of thalamocortical projection distributions. Katrina Gertz did the cell counting shown in Supplementary Figure 1, and Katrina and I performed the retrograde analyses shown in Supplementary Figure 10. Deniz Kusefoglu wrote the initial code to import and segment images for analysis in MATLAB, which was modified by both myself and Brian Long. For all computational alignment and analysis Brian Long and I designed each analysis step, Brian wrote the MATLAB code generate and analyze the data, and then I checked the output for accuracy and biological relevance. I made all of the final figures except Supplementary Figure 5, which was made by Brian, and I performed the literature review presented in Supplementary Figure 13. Tianyi Mao provided input at all steps. Haining Zhong helped with replication and modification of Brian's analysis code upon resubmission of the manuscript for publication. Brian, Tianyi, Haining, and I wrote the manuscript.

Abstract

The thalamus relays sensory and motor information to the cortex and is an integral part of cortical executive functions. However, the precise distribution of thalamic projections to the cortex is poorly characterized, particularly in mouse. We employed a systematic, high-throughput viral approach to visualize thalamocortical axons with high sensitivity. We then developed algorithms to directly compare injection and projection information across animals. By tiling the mouse thalamus with overlapping 254 injections, we constructed the comprehensive map of thalamocortical projections. We determined the projection origins of specific cortical sub-regions, and verified that the characterized projections formed functional synapses using optogenetic approaches. As an important application, we determined the optimal stereotaxic coordinates for targeting specific cortical sub-regions and expanded these analyses to localize layer-preferential projections. This dataset will serve as a foundation for functional investigations of thalamocortical circuits. Our approach and algorithms will also provide an example for analyzing the projection patterns of other brain regions.

Introduction

Anatomical connections provide structural substrates for information processing in the brain, yet neuroanatomical maps in most model organisms are incomplete(Bohland et al. 2009). This is especially true in mouse, where there are few comprehensive characterizations of anatomical connectivity despite being a primary model for studying neural function(Bohland et al. 2009). Anatomical connectivity at the mesoscopic level is critical for understanding of how circuits subserve behaviors and is necessary for investigation of circuit function using genetic manipulation(Bohland et al. 2009; Luo et al. 2008; Lichtman & Denk 2011).

The thalamus is integral to the flow of information into and within the brain via its extensive interconnection with the peripheral and central nervous systems(Jones 2007; Sherman & Guillery 2009; Groenewegen et al. 1990; Evarts & Thach 1969; Steriade & Llinás 1988; Asanuma & Fernandez 1974). Thalamocortical projections are the primary drivers of cortical activity in sensory areas(Sherman & Guillery 2009) and associative brain regions, such as the frontal cortex(Nauta 1971; Kolb 1977; Weinberger 1993). The thalamus contains ca. 40 nuclei(Jones 2007; Berman & Jones 1982; Walker 1938), each innervating a different combination of cortical areas. Thalamic inputs to the frontal cortex are poorly characterized compared to thalamic inputs to primary sensory cortices, and our knowledge of the thalamo-frontal pathway is based on an amalgam of tracing studies from primates, cats, and rats spanning several decades(Jones 2007). Gaining a complete representation of each thalamo-frontal projection pathway from these studies has been difficult, due to variability between techniques and inconsistencies in anatomical boundary definitions(Jones 2007). A systematic characterization of thalamo-frontal pathways is necessary for investigating the function of frontal sub-regions.

It remains challenging to create a comprehensive thalamocortical projection map from individual thalamic subdivisions in mouse. First, the potential target area spans the entire cortex, necessitating a high-throughput microscopic method that can image the projections throughout the cortex at sufficiently high resolution and sensitivity(Bohland et al. 2009).

Next, demarcating the cytoarchitectural boundaries for mouse thalamic nuclei is difficult because they are less distinct than the boundaries in other mammalian brains(Jones 2007). Furthermore, a comprehensive neuroanatomical dataset requires robust analysis methods to combine anatomical data across experimental animals(Bota et al. 2012). Finally, it remains a major challenge to process, analyze, summarize, and present large anatomical datasets.

To overcome these challenges, we have developed a high-throughput approach using bilateral, two-color, anterograde, focal viral injections into mouse thalami. We then imaged injected brains at sub-micrometer resolution, providing single axon sensitivity. We developed algorithms to localize injections within a model thalamus, allowing us to compare injection and projection information across animals. We identified the origins of thalamic inputs to 19 cortical sub-regions in mouse, focusing on poorly understood thalamo-frontal pathways. We further localized the origins of layer-specific cortical projections to vibrissal motor cortex (vM1). Based on coordinates extracted from our analyses, we performed viral injections encoding channelrhodopsin, and optogenetically confirmed that the anatomically characterized projections form functional synapses. Our data provide a practical guide for viral injection, imaging, and manipulation of thalamocortical circuits in mice. This method and associated analyses can be adapted to develop comprehensive neuroanatomical connectivity maps in other brain regions.

Materials & Methods

All animal experiments were conducted according to National Institutes of Health guidelines for animal research and were approved by the Institutional Animal Care and Use Committee. All measurements are listed as mean ± standard deviation, unless otherwise indicated. All calculations were performed in MATLAB (MathWorks). The raw data and the analyzed data are publicly available at http://digitalcollections.ohsu.edu/projectionmap. The original resolution images are available upon request as hard drive format.

Stereotaxic viral injections

Injections were performed as described(Mao et al. 2011) with optimizations/ modifications. Briefly, C57BL/6J male and female mice were anesthetized (1–2% isoflurane) at P14–18 and stabilized in a custom stereotaxic apparatus (modified from a David Kopf system). A dental drill (Henry-Schein) was used to drill holes through the skull. A pulled glass micropipette (Drummond; tip diameter: 10–15 μm), beveled sharp, was backfilled with AAV (serotype 2/1) that expresses either eGFP (Addgene 28014) or tdTomato (A gift from J. Magee). AAV2/1 is a hybrid serotype that has AAV2 inverted terminal repeats, AAV1 capsid proteins, and widespread neuronal tropism(Harris et al. 2012). The transgenes were driven by CAG promoter and included a WPRE element to enhance the expression. The viruses were prepared by the University of Pennsylvania vector core and viral titers >5.0 x 10^{12} GC/mL were used. Unless noted otherwise, a 10 nL volume of virus was dispensed at a speed of 5 nL/s using a hydraulic injector (Narishige), followed by a 5–10 minute waiting period. The pipette was retracted 0.3 mm at 0.008 mm/s, paused for 3 minutes, and then retracted at a rate of 0.008 mm/s. This process minimized the undesirable infection of cells along the injection path. Up to four injections were performed in each animal (two colors and two hemispheres). Coordinates for injections ranged from: 0.5 – –1.6 anterior to posterior, 0 – 1.6 lateral, and 2.8 – 4.2 deep from the pia (in mm from bregma). Total 254 highly-overlapped injections were used to insure the final coverage of the thalamus was > 90% volume.

Sectioning and imaging

14 days after viral infection, mice were perfused transcardially with 25 mL phosphate buffered saline (PBS) followed by 50 mL of 4% paraformaldehyde (PFA). The brain was post-fixed in 4% PFA at 4°C overnight and then placed in 30% sucrose in PBS at 4°C overnight. The brain was centered and aligned in a rectangular mold, embedded in Optimal Cutting Temperature medium (TISSUE-TEK), and sectioned coronally on a cryostat (Thermo Scientific) at 50 μm thickness. The sections from the most anterior section of the cortex to the most posterior section of the thalamus were floated in PBS and then collected onto Superfrost-Plus microscope slides (FisherBrand). Slides were mounted using Fluoromount (Sigma) and covered with number 1.5 cover glass (Gold Seal, Fisher).

All sections on the slides were imaged with a 20X objective (0.5 μm/pixel) on the Nanozoomer slide scanner (Hamamatsu), at a fixed exposure time. Because injection sites were often overexposed under these settings, they were re-imaged at a lower exposure with either a 5X objective on a Zeiss Axio Imager or using shorter exposure times on the Nanozoomer. Axio images were matched to their corresponding Nanozoomer section images through rigid translation and rotation using manually selected anatomical landmarks visible in both images. After imaging, injections that extended beyond the lateral or ventral borders of the thalamus were excluded.

Each brain was processed and imaged equally and randomly without any knowledge of the injection locations. Data distribution was assumed to be normal but this was not formally tested.

Cell counting

Confocal images were collected (Zeiss, LSM780) for DAPI (Vector Labs) stained sections across the center of an AAV2/1-eGFP thalamic injection site from 17 mice. The fraction of cells found to be both DAPI- and eGFP-positive indicated the percentage of DAPI-positive cells infected (**Supplementary Fig. 1d**). To calculate the percentage of

neurons infected, thalamus sections across the center of AAV2/1-eGFP injections from 5 mice were incubated with mouse anti-NeuN (Millipore), followed by Alexa-594 goat anti-mouse secondary antibody (Life Technologies) and DAPI. The fraction of DAPI-positive cells that were also NeuN- and eGFP-positive indicated the percentage of infected neurons at the injection site (**Supplementary Fig. 1b–e**). To confirm these results, the thalamus sections from 3 mice injected with eGFP-expressing AAV2/1 were stained with NeuroTrace (Life Technologies) and DAPI and were analyzed in the same fashion (data not shown). To evaluate the viral tropism, eGFP and tdTomato expressing AAV were mixed (1:1) and co-injected into the thalamus in 4 mice. The same imaging process was used as with single viral injections (**Supplementary Fig. 1f–h)**.

Thalamus and injection site segmentation

Individual sections were isolated from the full slide images by determining an intensity threshold that would distinguish tissue from background pixels. The outline of the thalamus was manually traced to generate a thalamus mask (**Fig. 1c**). The front of the thalamus was defined as the first slice posterior to the anterior commissure (AC) crossing the midline, and the back of the thalamus was defined as one slice posterior to the end of the lateral geniculate nucleus (LGN)(Paxinos 2004; Lein et al. 2007). In addition, the medial and lateral geniculate nuclei were not included due to their already well-characterized anatomy in the auditory and visual systems, respectively. Finally, the posterior portion of the reticular thalamic nucleus (RT), which does not produce cortical projections(Kolmac & Mitrofanis 1997) and the posterior portion of the ventral medial nucleus (VM) were excluded from the traced masks due to technical difficulties in visualizing their borders. We segmented each injection site into a binary mask by applying independent intensity thresholds in green and red channels, utilizing a supervised MATLAB routine based on Otsu's method(Otsu 1975). Traveling axon bundles that were above threshold in the thalamus were manually excluded from the associated injection site.

Thalamus registration and alignment

The model thalamus and registered injections were created as described in **Supplemental Table 1**. 1) Two manually selected midline points were used to rotate and align the thalamus masks. 2) To align the masks in the correct y position and to correct for the cutting angle tilt about the x-axis (i.e., rotation around the x-axis), we used anatomical landmarks to estimate the tilt angle (**Supplementary Fig. 5)**. A separately traced ABA thalamus mask was rotated to the same tilt angle and the mask stack was resampled as 50 μm slices. 3) The centers of mass of these slices were used to direct the position of experimental thalamus masks in y. The center of mass is defined as the unique point where the weighted relative position of the distributed mass sums to zero and was calculated as:

$$R = \frac{1}{M}\int_{v} \rho(\boldsymbol{r})\boldsymbol{r}dV$$

where M is the sum of the masses of each point $\mathbf{r}$ in a volume V with constant density $\varrho(\mathbf{r})$.

The aligned thalamus masks were then rotated to a tilt angle of 0 degrees and re-sampled as 50 μm slices (**Supplementary Fig. 5**). The thalamus masks were down-sampled to a 36.4 x 36.4 x 50 μm voxel size. 4) To control for the cutting angle tilt about the y-axis, the aligned 3D mask was sheared to maximize left and right asymmetry. 5) The overall size of the thalamus was scaled: i) in z so that the midline distance from the beginning to end of the thalamus matched the ABA thalamus, ii) in x-y isotropically to match the total area of the central slices with that of the corresponding ABA thalamus, and iii) the 3D thalamus masks were scaled in x to match the average width of all thalamus masks. 6) The masks were visually inspected and 18/75 brains underwent minor scaling or position adjustments in the z dimension. All brains were further aligned with each other in y based on their center of mass. All experimental masks were summed and then segmented according to a threshold that retains the volume of averaged thalamus mask volumes, producing the model thalamus.

We used Dice's coefficient to assess the similarity between two thalamic structures (**Fig. 2c** and **Supplementary Fig. 6d**). Dice's coefficient is defined as:

$$D(A,B) = {}^{2\sum(A\ \&\ B)}\Big/_{(\sum A + \sum B)}$$

where A and B are two binary volumes, '&' is the logical AND operator and indicates the sum of all elements in X(Dice 1945). To further quantify the variability of thalamus masks (**Fig. 2a** and **Supplementary Fig. 6**), we overlaid the borders of each thalamus mask and measured the distribution of boundary points at 18 locations (6 locations per slice for 3 z slices).

Injection site masks were processed identically to their corresponding thalamus masks so that they are registered to the model thalamus. All injection site masks were summed to quantify the injection coverage at individual voxels (**Fig. 2d–f**).

Atlas Alignment

To register known thalamic nuclei within our model thalamus, we traced, scaled, and re-sampled 25 nuclei from both the ABA and PMBA as 3D volumes that are aligned with our model thalamus. Differences in animal age and tissue preservation techniques resulted in size and shape differences between the two atlases, so each atlas was scaled separately to best fit our model thalamus. The correspondence between the nuclei of individual experimental thalami and the atlas nuclei were assessed by manually tracing four cytoarchitecturally identifiable thalamic structures (nuclei AD, AV and PT and fiber tract fr) from five randomly selected brains (**Fig. 2b**). The similarity between the atlas and experimental nuclei was assessed using Dice's coefficient (**Fig. 2c** and **Supplementary Fig. 6e**). Notably, the values of Dice's coefficients for comparing nuclei in are lower than those for comparing the thalami of all brains (**Supplementary Fig. 6d–e**) because this coefficient is inversely dependent on volume. For example, the average volume of the traced nuclei is 0.24 mm^3 (1.3% of the model thalamus volume), and our position variability (~100 μm) affects D for nuclei more than for the larger thalamus masks. The similarity matrix shows that (1) each traced nucleus is more similar to a corresponding nucleus in another brain or atlas, than to other nuclei, (average D = 0.53 for comparing the same nuclei and D = 0.02 for comparing

different nuclei) and (2) similarities between traced nuclei and atlas nuclei ($D = 0.53 \pm 0.10$) are comparable to that of the atlases to each other ($D = 0.60 \pm 0.11$) (**Fig. 2c** and **Supplementary Fig. 6e**).

Confidence maps and thalamic origins of projections to the cortical sub-regions

For each thalamic injection, projection distributions were blindly scored by three independent experts. The presence/absence, strength (dense or sparse), coverage (full or partial ROI coverage), and specificity (whether the projection also goes to an adjacent ROI) were determined (**Supplementary Fig. 8b**). All final scoring decisions were reached by consensus. The cortical area boundaries were based on the PMBA. Injections are referred to as being "positive" or "negative" for a given cortical ROI, where "positive" indicates the presence of a projection and "negative" indicates the absence of a projection for this particular ROI. To control for our alignment variability (~100 μm) across thalamus masks (**Fig. 2a** and **Fig. 3b**), an injection core was produced by eroding each 3D injection mask by 100 μm.

A confidence map, which defines the thalamic origin of cortical projections, was created for each target projection region. As shown in **Supplementary Figure 8**, a confidence map was developed by grouping injections (**Fig. 3b**) that met each of eight independent criteria. Meeting each criterion would give a thalamic voxel a score of 1 and meeting all criteria would result in a maximal confidence level of 8. For example, criteria (A) requires a voxel to be included in the core of an injection producing specific projections, but may not be in any negative injections, (see **Supplementary Fig. 8c** for the remaining seven criteria descriptions). The binary masks produced by each grouping criteria were summed to create the confidence map (**Supplementary Fig. 8d–g**).

Overall, the confidence maps incorporate information about the intensity and specificity of projections, as well as the variability in thalamus transformation and alignment. The confidence map therefore represents the likelihood of a thalamic voxel projecting to a particular target.

Voxel clustering based on projection confidence maps

The model thalamus and individual confidence maps were down sampled to 150 x 150 x 150 μm voxels. The thalamic voxels were then subjected to agglomerative hierarchical clustering (MATLAB) based on their confidence map values across the 19 target regions in a 19-D space using the city-block metric and average linkage with a set maximum of 11 clusters.

Quantifying the nuclear origins of thalamocortical projections

Atlas nuclei previously aligned to the model thalamus were overlaid onto our confidence maps. We calculated the fraction of injection-covered nucleus volume occupied by the confidence map at three confidence levels ($C \geq 3$, $C \geq 5$ and $C \geq 7$; **Supplementary Fig. 8d**). These values were averaged across the ABA and PMBA atlases to create the confidence threshold data (**Fig. 6a, d** and **Supplementary Fig. 12a**). The average nuclear fraction from these three thresholds (**Fig. 6a, d and Supplementary Fig. 12a**) gives the final nucleus projection data that forms the basis of our visualization and clustering results (**Fig. 6e**).

Clustering nuclei and projection regions

Each nucleus was assigned a point in a 19-dimensional space corresponding to the fraction of the thalamic nucleus volume occupied by projections to each cortical area (**Fig. 6e**). We performed a cluster analysis on the nuclei using a Euclidean distance metric and minimum linkage. The projection regions were similarly assigned a point in 25-dimensional space corresponding to the 25 nuclei, and clustered using the same method.

vM1 injection and projection analysis

The boundaries of vM1 were based on previous characterizations(Mao et al. 2011) and were defined as follows: dorsally by the pial surface, medially by a line that connects the top arc of the cingulum to the point that the pia folds towards the midline, and laterally by a line

from the cingulum to the pia that is parallel to the midline (**Fig. 7a**). vM1 was delineated independently for each hemisphere from three sections: the section where the corpus callosum merges plus one section anterior and one posterior.

Because layer depth and thickness varies depending on the position within vM1, we normalized all depths to that at the medial edge of vM1 (**Supplementary Fig. 14)**. Specifically, for pixels at angle θ, their depths are linearly transformed to the medial depth based on the layer boundaries at θ and at the medial boundary (**Supplementary Fig. 14a**). The normalized depths were separated into 100 bins, and florescence intensity values within each bin were averaged and normalized to the background fluorescence estimated from an unlabeled cortical region on the same section. These normalized fluorescence intensity traces were further background subtracted using the minimum values of the respective traces lying in the vicinity of cell body layer at the L1–L2/3 boundary.

A thalamic injection was considered to produce layer-preferential projections if it met two criteria. First, the average fluorescence intensity within either L2/3–5a or L5b had to be significantly greater than background fluorescence measured from the same depths in vM1 brains that did not contain projections to vM1. The threshold for each depth and each color was the median plus interquartile range of the background fluorescence levels in brain sections not containing vM1 projections. If ≥25% of the bins in either the L5b or L2/3–5a region were considered above threshold, then the second criteria would be evaluated. Second, after subtracting the layer specific background fluorescence, the intensity was averaged within L2/3–5a and L5b and a layer preference index, α, was computed from these average intensities in L2/3–5a and L5b ($I_{2/3–5a}$ and I_{5b} respectively):

$$\alpha = \log_2\left(I_{2/3-5a} \middle/ I_{5b}\right)$$

$\alpha = 0$ indicates equal intensities in the two regions while $\alpha > 0$ indicates higher fluorescence intensity in L2/3–5a and $\alpha < 0$ indicates higher fluorescence intensity in L5b. A threshold was set at 1.1: if α was greater than 1.1, the injection was classified as strongly L2/3–5a

preferential, and if α was less than –1.1, the injection was classified as L5b preferential. Based on this classification, we created vM1 layer-preferential thalamus confidence maps by scoring each voxel against the following 4 criteria (each criterion gives a score of 1 and meeting all criteria gives a maximal confidence level of 4): 1) the voxel is in a layer preferential injection, 2) the voxel is in the core of a layer preferential injection, 3) the voxel is in the core of a layer preferential injection with strong intensity, and 4) the voxel is not in the core of an injection with opposite layer preference (**Fig. 7**).

Statistics

Statistics comparisons were performed using a t test. N indicates the number of independent brains. The significance level was set at 0.05.

Photostimulation and electrophysiology

Mice were injected at P14–16 with 10–20 nL of an AAV2/1 virus encoding ChR2-H134R-TdTomato (Addgene: 28017). Cortical brain slices were prepared 14 days later from mice anesthetized with an intraperitoneal injection of ketamine/xylazine (0.13 mg ketamine/ 0.01 mg xylazine/g body weight) and perfused transcardially with ice cold ACSF containing (in mM): 127 NaCl, 25 $NaHCO_3$, 25 D-glucose, 2.5 KCl, 1 $MgCl_2$, 2 $CaCl_2$, and 1.25 NaH_2PO_4, pH 7.25–7.35, ~310 mOsm, and bubbled with 95% O_2 / 5% CO_2. The brain was removed and placed into ice-cold cutting solution containing (in mM): 110 choline chloride, 25 $NaHCO_3$, 25 D-glucose, 11.5 sodium ascorbate, 7 $MgCl_2$, 3 sodium pyruvate, 2.5 KCl, 1.25 NaH_2PO_4, and 0.5 $CaCl_2$. 300 μm thick modified coronal slices were vibratome sectioned (Leica 1200S) at an angle to achieve a cut perpendicular to the pial surface for each recorded brain area. Slices were incubated in oxygenated ACSF for 45 min at 34^{o}C, and then maintained in an oxygenated holding chamber at room temperature.

Subcellular channelrhodopsin-assisted circuit mapping (sCRACM) and electrophysiology were performed as previously described(Mao et al. 2011; Petreanu et al. 2009). The excitatory postsynaptic currents ($EPSC_{sCRACM}$) were recorded in voltage clamp (holding potentials were –70 mV or –75 mV) while blue light was stimulated the thalamic axons

transfected with Channelrhodopsin. Each map was repeated 2–4 times. After sCRACM maps were obtained, a cell was counted as a positive responder if there was any excitatory postsynaptic current amplitude >6x the standard deviation of the baseline.

Retrograde bead injections, imaging, and analysis

Retrograde tracing was performed using fluorescent latex microspheres (LumaFluor: Red Microbeads IX and Green Microbeads IX) at a 1:2 dilution in PBS. Injections were performed similarly to the viral injections with P27 mice (tip diameter: 40–60 μm). 3 days later, mice were perfused as described above, with the exception that brains were not post-fixed following perfusion. Brains were sectioned coronally on a vibratome (Leica BT1200S) at 100 μm thickness. Sections were floated, collected, mounted, and covered as described above. All sections on the slides were imaged (Olympus MVX10), at a fixed exposure time, using a Retiga 2000R camera. From these images, the cortical injection sites as well as the approximate distribution of fluorescent thalamic somas were manually mapped onto thalamic sections (**Supplementary Fig. 10**).

Results

Labeling and imaging thalamocortical projections

To visualize thalamic projections, we stereotaxically injected two recombinant adeno-associated viruses (serotype 2/1; AAV2/1)(Mao et al. 2011; Harris et al. 2012; McFarland et al. 2009; Aschauer et al. 2013) encoding eGFP and tdTomato respectively, bilaterally into the mouse thalamus (**Fig. 1a–c** and **Supplementary Fig. 1**) Thalamic projections do not cross the midline in mouse(Hoover & Vertes 2007) (**Supplementary Fig. 1a**), which allowed us to inject, image and analyze each hemisphere independently. Bilateral, two-color viral injections quadrupled the throughput of subsequent data collection, consolidated the total amount of data (~0.5 TB/animal), and eased computational demands for data processing. In addition, two-color labeling highlights topographic projection patterns from adjacent thalamic volumes(Wang & Burkhalter 2007; Thompson & Swanson 2010) (**Supplementary Fig. 2**). By using a hydraulic apparatus to deliver ~10 nL of AAV, a consistently small infection volume was achieved; measuring 0.30 ± 0.23 mm^3, corresponding to ~1.6% of the total thalamic volume, and of 630 ± 160 μm (n = 188 injections) wide in the medial-lateral axis (**Fig. 1c**). 67.4 ± 10.3% of cells expressed detectable levels of fluorescent protein at the injection center, with an 88.8 ± 4.4% infection rate for neurons (**Supplementary Fig. 1b–e**).

Brains were paraformaldehyde fixed and cryostat sectioned coronally at 50 μm (**Fig. 1a–c**). All sections of each brain, from the start of the frontal cortex through the end of the thalamus, were fluorescently imaged in their entirety under identical conditions using a Hamamatsu Nanozoomer imaging system (0.5 μm/pixel) providing sufficient resolution to detect single axons (**Fig. 1b**). The thalami were re-imaged to avoid saturation of the injection sites (**Fig. 1c**). We successfully imaged 75 mouse brains containing a total of 254 injections, resulting in ~40 TB of imaging data.

An overview of data analysis

To analyze and compare thalamic injections across animals, we developed a suite of custom algorithms using MATLAB (MathWorks). The goal of these algorithms is to align

individual injection sites onto a model thalamus (**Fig. 1d**), such that injection and projection information can be compared across brains. **Supplementary Figure 3** schematically illustrates our approach. We manually traced each thalamus from the section images to generate a binary thalamus mask. Injection sites were masked by applying an intensity threshold to the images using a threshold determined by the Otsu's method(Otsu 1975) (**Fig. 1c** and **Supplementary Fig. 3b**). We then aligned and stacked each brain's thalamus mask sections to create a 3D volume mask (**Fig. 1d, Supplementary Fig. 4,** and see **Methods**). We normalized the 3D masks and their corresponding injection site masks, corrected them for variability in cutting angle, and aligned them using anatomical landmarks (**Supplementary Fig. 4, Supplementary Fig. 5,** and see **Methods**). The aligned 3D thalamus masks were then averaged to produce a model thalamus (**Supplementary Fig. 3c** and **Supplementary Fig. 6a**)**,** and each injection site was mapped onto the model (**Fig. 1d**).

We then determined the cortical projection targets for each injection, and combined the injection and target information for all 254 injections to localize the precise thalamic origin of the cortical projections (**Supplementary Fig. 3d**). We aligned two widely used atlases to the model for nucleus-specific analysis (**Supplementary Fig. 3e**). Notably, we also used our comprehensive dataset to create a nucleus-independent assessment of subdivisions within the thalamus (**Supplementary Fig. 3e**).

Assessment of thalamus alignment and injection coverage

After normalization and alignment (see above and **Methods**), individual thalami were highly similar to each other, with only 3.7% variability in the thalamic volume (percent s.d.), and 102 ± 51 μm (mean ± s.d.) variability in the thalamic border location (**Fig 2a** and **Supplementary Fig. 6b–c**). This variability is nearly identical to that measured with alternative data collection methods such as serial block-face imaging (102.5 μm ± 45 μm) (Ragan et al. 2012). The high degree of similarity between the individual masks and the model thalamus, was confirmed using Dice's coefficient (D = 0.94 ± 0.01; **Supplementary Fig. 6d**). To facilitate subsequent data analysis, thalamus masks were down-sampled to a

voxel size of 36.4 X 36.4 X 50 μm (x, y, and z, respectively), which is more than 2 fold smaller than the variability across individual thalami.

We then aligned our model thalamus to two atlases: the Allen Brain Atlas (ABA, http://mouse.brain-map.org) and the Paxinos Mouse Brain Atlas (PMBA)(Paxinos 2004) (**Supplementary Fig. 3e**, **Fig. 2b**, and refer to **Table 1** for all anatomical structure abbreviations). To verify this alignment, four cytoarchitecturally identifiable structures (AD, AV, PT and fr) were traced from 5 randomly selected experimental brains and compared to their corresponding atlas structures (**Fig. 2b–c**). The overall shape, orientation, and location of the thalamic structures were highly similar among the brains and atlases as quantified using Dice's coefficient (**Fig. 2c** and **Supplementary Fig. 6e**). While variability across brains remained, the structures from experimental brains were as similar to the atlases ($D = 0.53 \pm 0.10$) as the atlases were to one another ($D = 0.60 \pm 0.11$; $p = 0.35$, t-test). We concluded that the alignment of individual nuclei to our model was accurate.

We distributed the injections throughout the thalamus (**Supplementary Fig. 7a**) such that 93.4% of the thalamus was covered by at least 1 injection, and 85.3% was covered by at least 2 injections (**Fig. 2d–f** and **Supplementary Fig. 7c**). The majority of thalamic nuclei are fully covered (**Fig. 2f** and **Supplementary Fig. 7b**); however, we excluded the geniculate nuclei from the dataset. The center of the thalamus was more highly sampled because injections that extended beyond the lateral or ventral borders of the thalamus were excluded (**Fig. 2d**, and **Supplementary Fig. 7a, c**).

Mapping the thalamic origins to cortical targets

Using this dataset, we sought to identify the thalamic sources of projections to each of 19 cortical sub-regions of interest (ROI's), which were defined by their boundaries in the PMBA (**Fig. 3a**). We noted the strength and specificity of projections from each of our injections to all cortical areas using a manual scoring system, as detailed in **Supplementary Figure 8**. Independently, three experts blindly performed this analysis.

We then used the projection scores for each injection to perform a simple injection site grouping method, which allowed us to localize the thalamic sub-volumes projecting to each cortical ROI. First, to account for the alignment variability between thalami (102 ± 51 μm), we eroded each aligned injection site by 100 μm to produce the injection "core" (**Fig. 1c**, and **Fig. 3b**). The core, as compared to the periphery, represents the volume of an injection that we are more confident is accurately localized within the model thalamus.

Next, we combined the volumes of all injections that projected to a given ROI ("positive" injections), and then the volumes of the injections that did not project to that region ("negative" injections) were subtracted from the combined total. This process resulted in a better localization of the thalamic volume projecting to each ROI than if the summed positive injection volume was used alone (**Fig. 3b–c** and **Supplementary Fig. 8e–g**). By employing this method, we localized volumes at a finer resolution than at the size of a single injection. The grouping method described above was expanded to assign higher confidence to injection site cores, as well as account for different projection properties, such and strength and specificity (see **Supplementary Fig. 8** and **Methods** for detailed grouping). This analysis resulted in confidence maps in which the value of each thalamic voxel, volumetric pixel, indicates our certainty that the thalamic voxel projects to a particular ROI (**Fig. 3d–e** and **Supplementary Fig. 8d–g**), where a confidence value of 8 is highest, and a confidence value of 0 means that no projections originated from that voxel. We have provided confidence map summaries for the nine sub-regions of the frontal cortex: FrA, dACC, vACC, PrL, IL, MO, VO, LO, AI (**Fig. 3d–e, Supplementary Movies 1–9,** and **Supplementary Fig. 9** for full confidence maps to all sub-regions). Each confidence map contains a continuous positive volume, which is unique for each target region. To validate the projection sources predicted by our confidence maps, we performed injections of fluorescent retrograde beads in a subset of our characterized areas (**Supplementary Fig. 10**). We observed all retrogradely transported beads were localized within the predicted confidence map.

One advantage of having confidence maps across many cortical sub-regions is that we could directly compare the thalamic origins of functionally related cortical sub-regions. For example, PrL and IL are both crucial in fear learning, but PrL is associated with the 'high fear' behavior state, and IL is associated with the 'low fear' state(Knapska et al. 2012). By comparing the confidence maps for PrL and IL, we localized the shared and unique thalamic origins to PrL and IL (**Fig. 3f**), suggesting that differential thalamic inputs may contribute to their functional differences. Such comparisons allow for the selective targeting of thalamic projections to PrL and IL for future functional studies.

Defining thalamic subdivisions based on cortical targets

The thalamus is commonly subdivided into anatomically and functionally similar nuclear groups(Jones 2007). While useful, these divisions ignore ambiguity in nuclear borders, differences in projection patterns within a single nucleus, and the possibility that cytoarchitecturally defined nuclei may not always be the relevant functional unit within the thalamus(Sherman & Guillery 2009). Since our confidence maps provide distinct topographic information (**Fig. 4a**), we determined whether the thalamus could be instead sub-divided based on cortical projection patterns alone.

The thalamus was divided into 150 x 150 x 150 μm voxels (**Fig. 4b**), which were then clustered (agglomerative hierarchical clustering, MATLAB) based on their confidence values for all 19 cortical sub-regions (**Fig. 4b–c**). We then applied a threshold to identify the 11 largest thalamic voxel clusters (**Fig. 4c–e**). Notably, the thalamic voxels comprising each cluster were spatially grouped and largely continuous, and similar to the thalamic nuclear groups (**Fig. 4d)**. However, the voxel clusters and nuclear groups were not identical. While several nuclear groups were comprised of one or two closely related clusters (**Fig. 4f–g**, anterior and intralaminar nuclear groups), other nuclear groups contained several largely divergent clusters (e.g. the medial and ventral groups, **Fig. 4f–g**), suggesting functional homogeneity in some nuclear groups, but significant heterogeneity in others.

Optimal injection sites and functional confirmation

Stereotaxic viral delivery of optogenetic and pharmacogenetic reagents to manipulate neuronal activities has become an important method to dissect functional circuitry. Currently, studies involving the mouse thalamus that employ these methods primarily rely on the empirical determination of the injection coordinates based on a small number of trials. Using the confidence maps developed here, we have simulated injections throughout our model thalamus and determined the optimal injection coordinates for targeting projections to a specific ROI (**Fig. 5a** and **Supplementary Fig. 11** for all optimal injection coordinates).

Since anatomical projections do not always guarantee functional connectivity(Mao et al. 2011; Dantzker & Callaway 2000; Shepherd et al. 2005), we sought to verify that the observed anatomical axonal projections form functional connections at each target region, which also allowed us to verify the validity of the optimal injection coordinates. We injected AAV2/1 expressing channelrhodopsin2 (ChR2) using our optimal injection coordinates to target thalamic projections to eight frontal sub-regions. Whole cell recordings were made in each projection target area, shown here for dACC (**Fig. 5b**), and postsynaptic responses were observed upon activation of the ChR2+ thalamic axons with blue light stimulation (**Fig. 5b–d** and **Methods**). 48 out of 50 cells recorded showed excitatory responses; specifically, we recorded responses from 4/4 cells in AI, 13/13 in VO/LO, 4/4 in MO, 3/5 in IL, 5/5 in PrL, 19/19 in dACC/vACC (cells with responses/total cells recorded; **Fig. 5d**), indicating that the anatomically defined projections corresponded to functional thalamocortical synaptic connections.

Grouping thalamic nuclei based on cortical targets

As described earlier, nucleus locations from both the ABA and PMBA were aligned to our model thalamus (**Fig. 2b–c**), allowing us to localize the origins of cortical projections to individual nuclei (**Fig. 6a–c**). To compute the fraction of each nucleus that projects to a given ROI, the nuclear boundaries aligned from the atlases were overlaid onto the confidence maps (**Fig. 6b**), and the coverage of a given nucleus was averaged between the

two atlases. The coverage distribution across nuclei is shown for projections to select frontal sub-regions (**Fig. 6a, d** and see **Supplementary Fig. 12** for all areas). We performed a cluster analysis using the nuclear localization of the confidence data for all 19 cortical sub-regions to identify projection patterns across thalamic nuclei (**Fig. 6e**). Functionally related cortical sub-regions formed tight clusters when grouped according to the origin of their thalamic inputs, suggesting that our comprehensive anatomical dataset can be predictive of functional relationships, which validates our approach. It is important to note that there are limits to the resolution of this method: small (<300 μm wide) and intricately shaped nuclei will be difficult to separate from their neighbors.

We compared our nuclear localized thalamocortical projection data to literature data for rat (**Fig. 6a, d**, **Supplementary Fig. 12** and **Supplementary Fig. 13**), because primary anatomical data for mouse is sparse(Jones 2007). Overall, our nucleus projection data are largely consistent with the cumulative rat anatomical data, but we have indicated discrepancies between our findings and the rat literature with asterisks (**Fig. 6a, d** and **Supplementary Fig. 13a** for full literature list). Several factors may contribute to these discrepancies. First, the boundary definitions between cortical sub-regions vary across atlases, so the atlas used in each study will impact their findings (**Supplementary Fig. 13b**) (Paxinos 2004; Groenewegen 1988; Ray & Price 1992; Lein et al. 2007), as exemplified by FrA(Paxinos 2004; Lein et al. 2007; Haque et al. 2010; Desbois & Villanueva 2001). Second, localization of projection origins within specific thalamic nuclei can vary both due to the atlas used and the ability to precisely target individual nuclei, as demonstrated by the discrepancies in projections from CM reported in the literature(Hoover & Vertes 2007; Haque et al. 2010; Berendse & Groenewegen 1991; Vertes et al. 2012; Van der Werf et al. 2002) (**Supplementary Fig. 13a**). To avoid anatomical bias, we averaged nucleus localization data between two atlases (ABA and PMBA) (**Supplementary Fig. 13c**), and created our confidence maps independent of nuclear boundaries (**Fig. 3e**). In addition, most studies cannot identify the regions of the thalamus that do not project to a given ROI because they lack the comprehensive dataset necessary to do so. Using our approach, we are able to

present this underreported feature of the thalamocortical connectome (**Fig. 6** and **Supplementary Fig. 12**).

Thalamic origins of layer-preferential projections in vM1

Different layers of the same cortical area play distinct roles in information integration. We analyzed the primary vibrissal motor cortex (vM1) to test whether our dataset could be used to identify thalamic volumes preferentially innervating specific cortical layers. We previously showed that the posterior "sensory" thalamus is more likely to project to layers 2/3 and 5a (L2/3–5a) in vM1, whereas the anterior "motor" thalamus projects to layer 5b (L5b) as well as L2/3–5a(Hooks et al. 2013). However, based on subjective assessments of a small number of injections, we had estimated the thalamic volumes responsible for these projections. To accurately localize the thalamic origin of layer-specific projections to vM1, we quantified the fluorescence intensity of thalamic projections to L2/3–5a and L5b for each injection (**Fig. 7a–c** and **Supplementary Fig. 14**), and created modified confidence maps to characterize the thalamic volumes associated with layer preferential projections (**Fig. 7d–e** and **Methods**). Several nuclei, including PCN, AM, LD, and VAL, contained volumes preferentially innervating L5b (**Fig. 7d–g**). While previous research shows that these nuclei send vM1 projections broadly to both L5b and L2/3–5a(Hooks et al. 2011), we found the first evidence of preferential projections to L5b in vM1. Since L5b neurons provide the only direct motor output from vM1, these projections may play a direct role in motor control. The thalamic projections that preferentially target L2/3–5a arose from a more posterior-central thalamic volume, identified here as Po, LP, Pf and SPFp (**Fig. 7d–g**). This confirmed previous results, which suggest preferential projections from a region containing Po to L2/3–5a in vM1(Hooks et al. 2011). Furthermore, when we compared each layer-preferential thalamic volume to the thalamic voxel clusters identified in **Figure 4c**, we found that several clusters displayed strong preference to specific vM1 layers. For example, 81% of cluster 11 preferentially projected to L2/3–5a of vM1, while only 0.3% projected

preferentially to L5b (**Supplementary Fig. 14d**). We concluded that future studies can use our method to identify thalamic volumes targeting detailed anatomical features.

Discussion

Mesoscopic connectivity maps are crucial for studying interactions among multiple brain regions and for linking cellular circuit mechanisms to behaviors. In this study, by using anterograde viral tracing, high-throughput whole brain imaging, and custom development of alignment and analysis software, we established a mesoscopic thalamo-centric projection map to the cortex in mouse, identified unique thalamic sub-volumes projecting to each cortical sub-region, and determined the optimal injection coordinates for optogenetically or anatomically targeting specific cortical regions. Our maps also permitted the identification of shared and unique thalamic sources to different cortical regions, such as PrL and IL (**Fig. 3f**), providing an entry point for teasing out their common and distinct functions. Additionally, our systematic approach allowed us to functionally subdivide the thalamus based solely on cortical projection patterns (**Fig. 4**). We further identified the thalamic volumes that give rise to layer-preferential projections to vM1 (**Fig. 7**). Our results provide a foundation for understanding the function of the thalamus and frontal cortex, as well as for investigating and manipulating the microcircuits within and between thalamic and cortical sub-regions.

Historically, the extensive time and labor required to image and map long-range projections limited the number of tracer injections used in anatomical studies, and necessitated the reliance on subjective assessments to compare across experiments. Recent advances in high-throughput fluorescent imaging facilitate the generation of large anatomical image datasets(Thompson & Swanson 2010; Hintiryan et al. 2012; Oh et al. 2014; Zingg et al. 2014), allowing researchers to access vast amounts of anatomical information. However, extracting relevant biological information from these data remains a major challenge for several reasons: variability across experiments, both due to intrinsic size differences and experimental manipulation, makes it difficult to compare across experiments directly; the resolution is limited to the size and shape of the tracer injection site; and the tools needed for data analysis have not kept up with technological advances in data collection, impeding efforts to turn images into knowledge.

We addressed variability issues by tightly controlling the animal age (P30 ± 2), computationally correcting for angled sectioning, and normalizing individual thalami to a standard volume. The variability among our thalamic mask boundaries is 102 ± 51 μm, comparable to that observed in the absence of mechanical sectioning(Ragan et al. 2012). By creating a comprehensive, age-matched thalamocortical projection map, we have provided a framework that others can build upon to understand differences across age groups, cell types, and species. These variances may explain some of the discrepancies seen across the 43 anatomical studies we evaluated in rat and our data in mouse (**Supplementary Fig. 13**).

Another major limitation of mesoscopic mapping is that the size of the tracer injection limits the resolution. To reliably identify the origin of the each mapped projection, tracer injections must target a single defined brain region. This is straightforward in cortical areas with large, superficial sub-regions(Oh et al. 2014; Zingg et al. 2014); however, this task is difficult, if not impossible, in the thalamus due to the complex shape and small size of many thalamic nuclei. We overcame this limitation by analyzing the intersectional areas of overlapping injections (**Fig. 3b–c**), which allowed us to localize volumes smaller than a single injection. We could only have obtained our confidence maps and optimal injection coordinates by integrating information from a large number of highly overlapping injections. Furthermore, to maximize our resolution, we used the smallest replicable viral infection volume (~0.3 mm^2, and laterally ~600 μm). We estimated our resolution to be larger than our variability (~100 μm) and smaller than our injection size (~600 μm), which is sufficient for most thalamic targeting, but small thalamic nuclei may require smaller injection volumes or more closely spaced injections to precisely discriminate their boundaries (e.g. **Supplementary Fig. 10**). Because of the heavy dependence on viral infection to deliver molecular reagents in systems neuroscience, our map, which is at the equivalent 'operational scale', can serve as a guide for targeting these tools.

By exploiting injections that both do and do not project to each cortical ROI, we were able to identify the entire thalamic volume that does not project to each cortical ROI. From an anatomical point of view, characterization of non-projecting regions is particularly

important because it has been estimated that only ~10% of all possible connections within the rodent brain are fully characterized at the mesoscopic level, largely due to a lack of definitive information on non-existent projections(Bohland et al. 2009).

As stated by Sherman and Guillery, 'The concept of the thalamic nucleus as a single structural, functional, and connectional entity has barely survived advancing techniques and new information. We stay with the thalamic nuclei as one of our prime analytical tools because, as yet, we have little to use in its place'. Here, our comprehensive projection map provided us with a unique opportunity to establish a nucleus-independent map of thalamic projections that transcends what we have learned from a nucleus-based framework (**Fig. 3–5**). Although we related our results to thalamic nuclei, we created our confidence maps independent of nuclear boundaries. This enabled us to unbiasedly identify the precise thalamic volumes responsible for projections to specific cortical sub-regions and cortical layers.

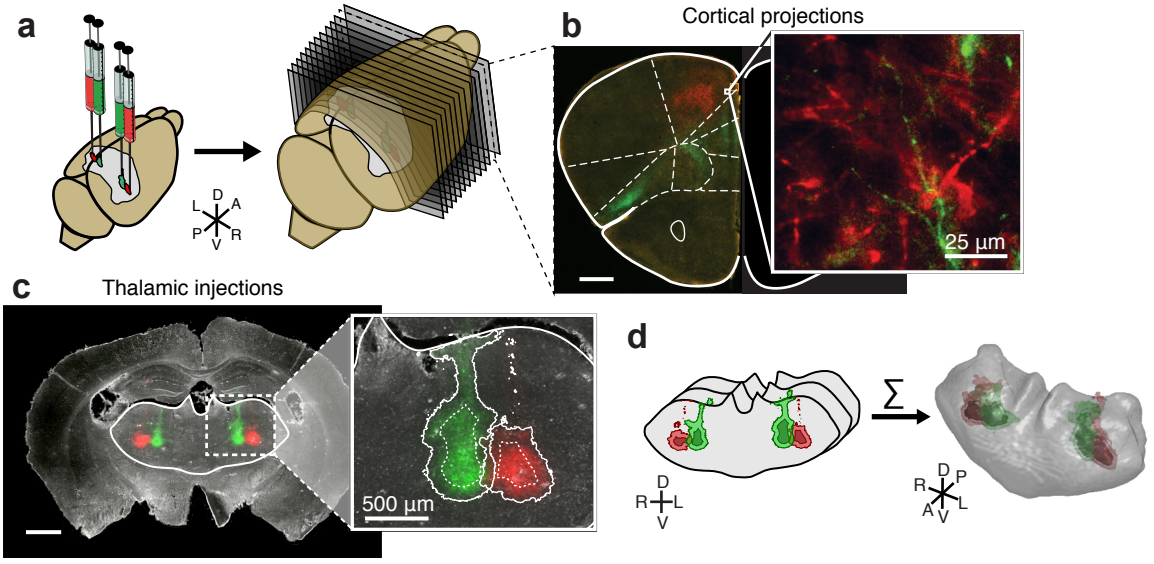
Our maps were obtained in adolescent mice, which is a dynamic period for prefrontal cortex (PFC) associated behaviors(Spear 2000; Van Eden & Uylings 1985; CASEY et al. 2008). We found that thalamocortical projections from at least 25 nuclei have reached PFC and form functional synapses by P30 (**Fig. 5** and **6**). Since the frontal sub-regions innervated by each nucleus are comparable to those seen in the adult rat (**Supplementary Fig. 13**), our data suggest that thalamocortical projections to PFC have reached their final targets by P30 in mouse. We therefore propose that the behavioral changes that occur during adolescence are more likely due to local refinements and synaptic pruning than larger rearrangements in thalamocortical projection distributions to PFC sub-regions.

In light of novel tools for imaging, physiology, and cell-type specific manipulations in mouse(Luo et al. 2008), the mesoscopic data provided here will serve as a critical reference for applying these tools to study circuit function. The results from over 43 disparate studies were necessary to summarize only a fraction of the thalamocortical projections in rat that are described here in mouse (**Supplementary Fig. 13**), which is a testament to the power of the high-throughput imaging and computational analyses used in this study. The ability to

directly compare across animals and experiments is a crucial step for extracting useful biological information from large anatomical datasets. Our results present an example for large-scale data integration and analysis, and will inform future studies in systems neurobiology.

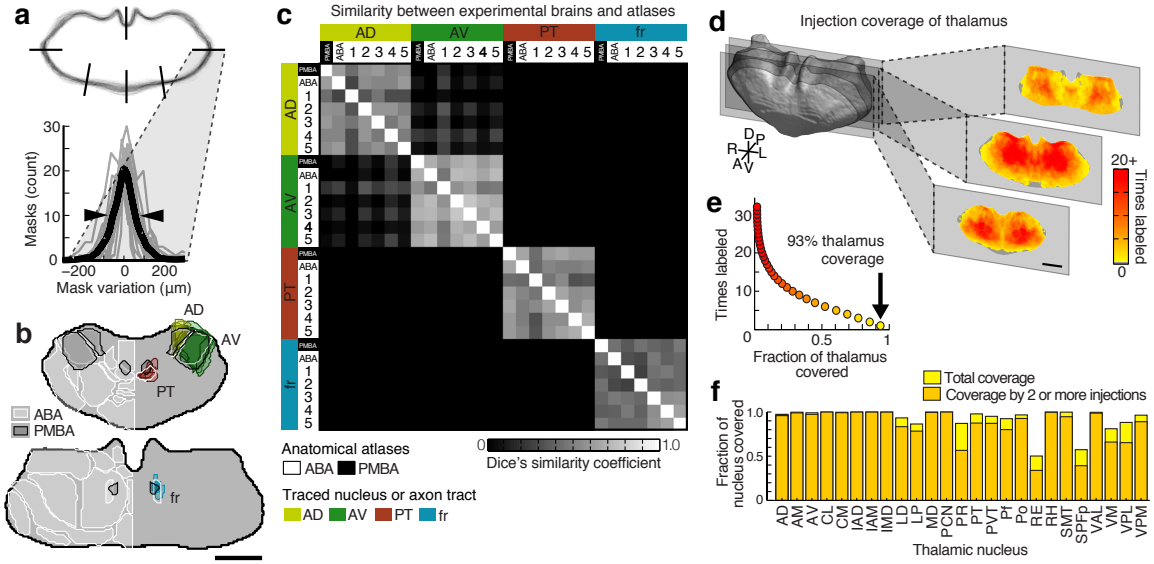
Figures

Figure 1. Systematic mapping of fluorescently labeled thalamocortical projections using high-throughput, high-resolution imaging.



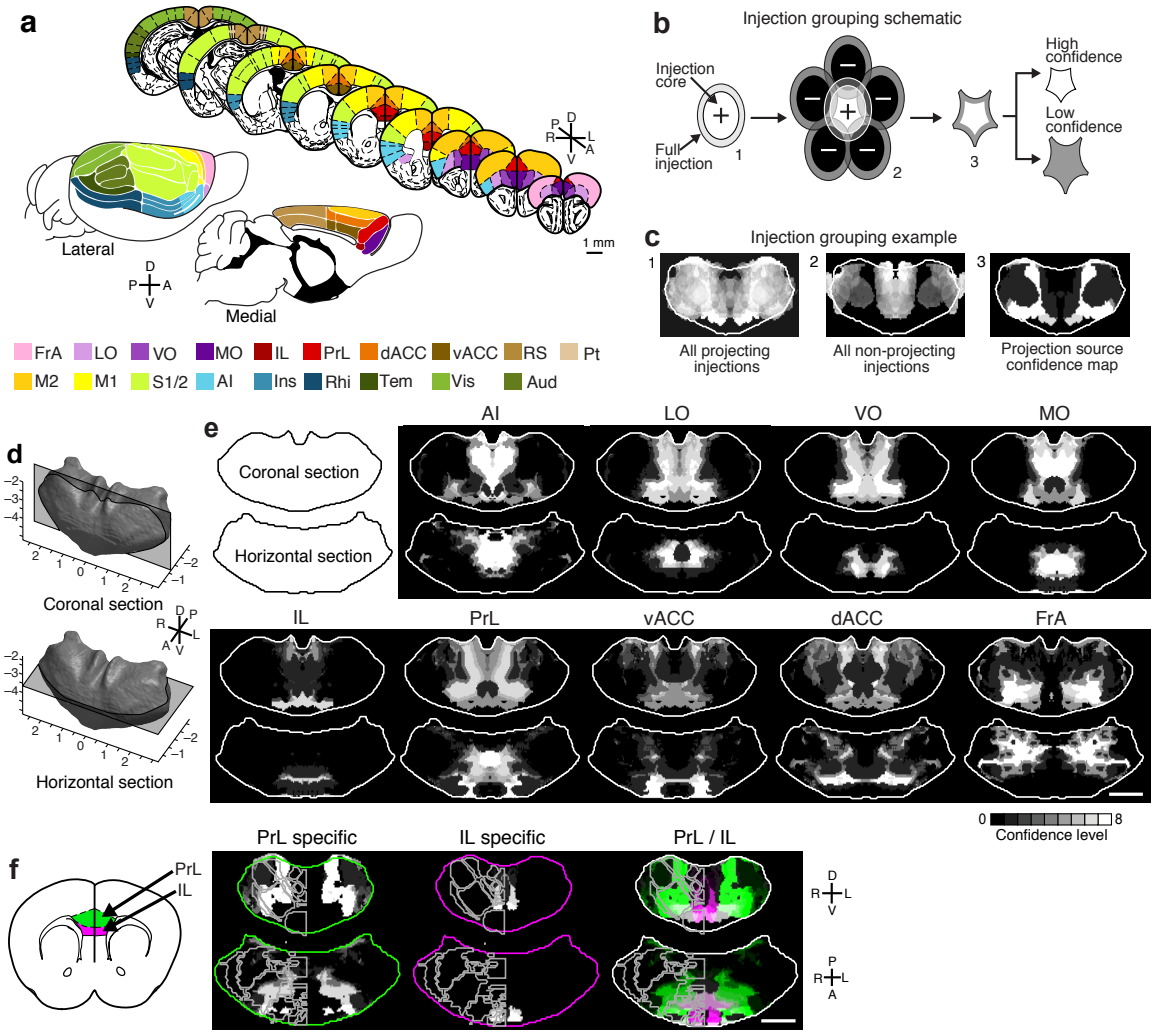
(**a**) Illustration showing bilateral viral injections driving the expression of tdTomato (red) and eGFP (green) in the mouse thalamus (left), followed by sectioning (50 µm/section, right) and high resolution imaging under identical conditions (right). (**b**) Representative coronal section showing thalamocortical projections to specific frontal sub-regions, with a zoomed-in image showing that full-resolution images allow the identification of single axons (inset). (**c**) Example fluorescent image showing viral injection sites on the dark-field image of the brain section (left). Solid white line: the thalamus mask. Zoom-in of the injection sites (right) shows the injection site masks created by intensity thresholding (solid line), as well as the injection site "cores" created by eroding the injection by 100 µm (dashed line) (see Methods). (**d**) The outline of the thalamus is manually traced in each coronal section and combined with the injection site masks created in panel **c**. These masks are then stacked to create a 3D representation of each thalamus. All scale bars are 1 mm, unless otherwise indicated. A (anterior), P (posterior), L (left), R (right), D (dorsal), V (ventral) throughout all figures.

Figure 2. Assessment of variability across brains, atlas alignment, and injection coverage of the thalamus.



(**a**) Top: aligned coronal thalamus sections from 75 brains (gray outlines). Black lines indicate 6 of 18 line profiles used to calculate thalamus edge variability. Bottom: thalamus edge variability after normalization at 18 locations (gray traces) and their average (black trace, full-width half-maximum = 102 ± 51 μm, arrowheads). (**b**) Two representative coronal sections through the averaged model thalamus (gray), overlaid with three thalamic nuclei (AD, yellow; AV, green; PT, red) and one axon tract (fr, blue) traced from 5 experimental brains. These atlas structures are also shown for the Paxinos Mouse Brain Atlas (PMBA, black) and the Allen Brain Atlas (ABA, white). (**c**) Dice's similarity coefficient across the traced nuclei and axon bundle in 5 experimental animals, the ABA, and the PMBA, showing that each traced structure is well aligned to that same structure in other experimental brains and in each atlas. Data are symmetric across the diagonal. (**d**) The model thalamus (left) with coronal sections through the model thalamus showing injection coverage within the thalamus (i.e. how many times a voxel is hit by independent viral injections). See **Supplementary Fig. 7** for full coverage maps. (**e**) The fraction of the thalamic volume covered by a given number of injections, with 93.4% of the thalamus covered by at least 1 injection (arrow). (**f**) The fraction of each thalamic nucleus covered by at least one and at least two injections. All scale bars are 1 mm.

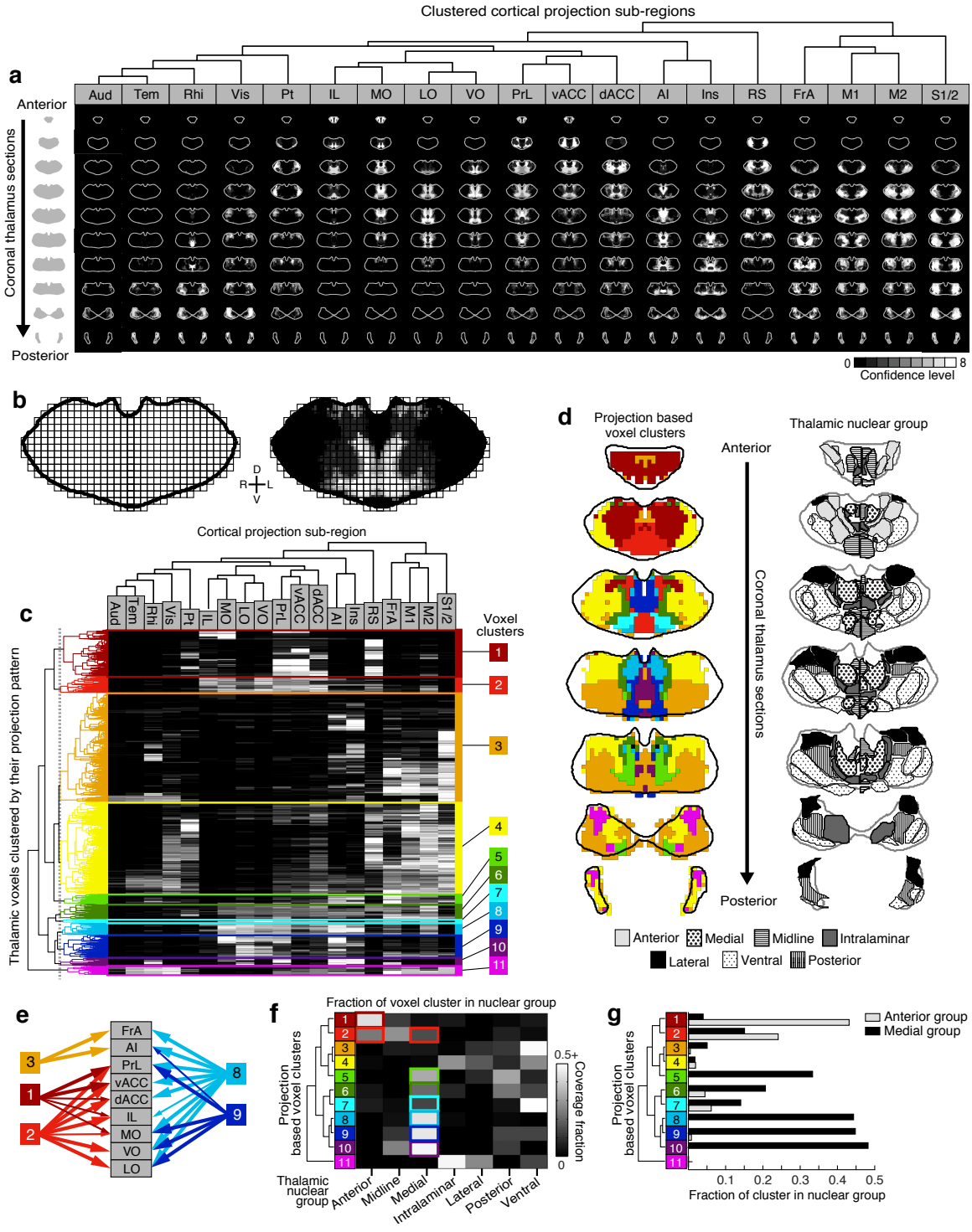
Figure 3. Localization of the thalamic origins of cortical projections.



(**a**) Illustration showing the 19 cortical areas examined (modified from PMBA (Paxinos 2004; Watson et al. 2012)): whole brain views (lower) and coronal sections (upper). (**b**) Simplified schematic of thalamic localization method: 1. The volume of each injection is eroded to generate an injection core (Fig. 1 and Methods); 2. Injections that send projections to a region of interest (ROI) (positive injections) are summed, and those that do not (negative injections) are subtracted; 3. Resulting in the precise thalamic volume projecting to each ROI, which is separated into eight confidence levels (**Supplementary Fig. 8**). (**c**) Example thalamus section illustrating the injection grouping method summarized in panel b. Sections 1 and 2: injections that do project and do not project to a target, respectively. Section 3: the refined thalamic volume projecting to that target (i.e. the confidence map). (**d**) Diagram of the thalamus sections shown in **e**. Coronal section –1.16 mm posterior to bregma (top), horizontal section –3.52 mm ventral to bregma (bottom). (**e**) Confidence maps (gray scale) show the thalamic origin of projections to nine frontal brain areas. See **Supplementary**

Figure 9 for full confidence maps to all cortical areas. (**f**) Localization of exclusive and shared thalamocortical projections to PrL (green) and IL (magenta) through direct comparison of their confidence maps (right). Nuclear boundaries shown on left half of each thalamus section (PMBA). Coronal section –0.61 mm posterior to bregma (top), horizontal section –4.06 mm ventral to bregma (bottom). All scale bars are 1 mm.

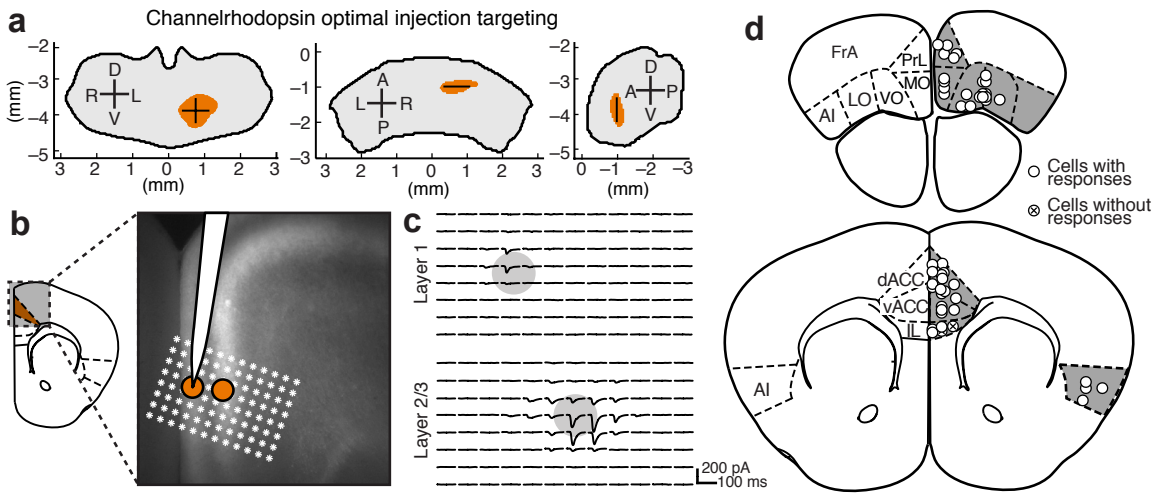
Figure 4. Localizing thalamic subdivisions based on cortical projection patterns.



(**a**) Summary of confidence maps to all cortical sub-regions, clustered based on confidence map similarity (determined in panel **c**). See **Supplementary Fig. 9** for large confidence maps. (**b**) The thalamus is down-sampled into 150 x 150 x 150 μm voxels (left) and the

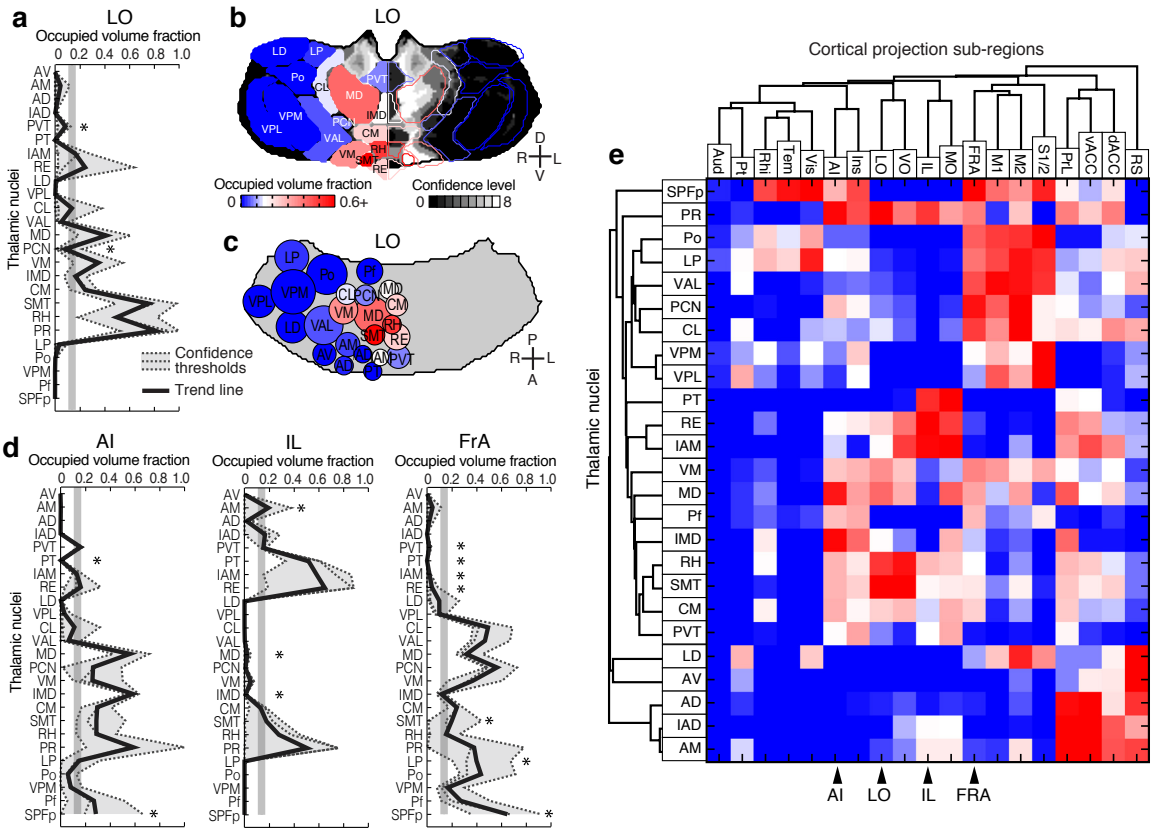
average confidence level within each voxel is determined for each cortical projection (example, right). (**c**) All thalamic voxels (rows) are hierarchically clustered based on their cortical projection patterns, and cortical sub-regions (columns) are clustered based on which thalamic voxels innervate them. The average confidence level is indicated in gray scale, as in **a**. A threshold (gray dashed line, left) was applied to identify 11 distinct clusters. (**d**) Coronal thalamus sections showing the spatial location of clusters from **c** (left), with the corresponding atlas sections (PMBA, left half & ABA, right half) showing thalamic nuclear groups for comparison (right). (**e**) Schematic showing the convergence and divergence of projections for several clusters. (**f**) Overlap between voxel clusters (rows) and atlas-defined nuclear groups (columns). Colored boxes highlight the clusters that are dominant in (compose >10% of) the anterior and medial thalamic groups. Some nuclear groups are covered by relatively few clusters that have closely related projection patterns (e.g., the anterior group mainly contains clusters 1 and 2), while other groups contain clusters with disparate projection patterns (e.g., the medial group contains clusters 2 & 5–10). (**g**) Coverage of the anterior and medial nuclear groups by each voxel cluster.

Figure 5. Targeting anatomically defined thalamocortical projections to verify that they form functional synapses.



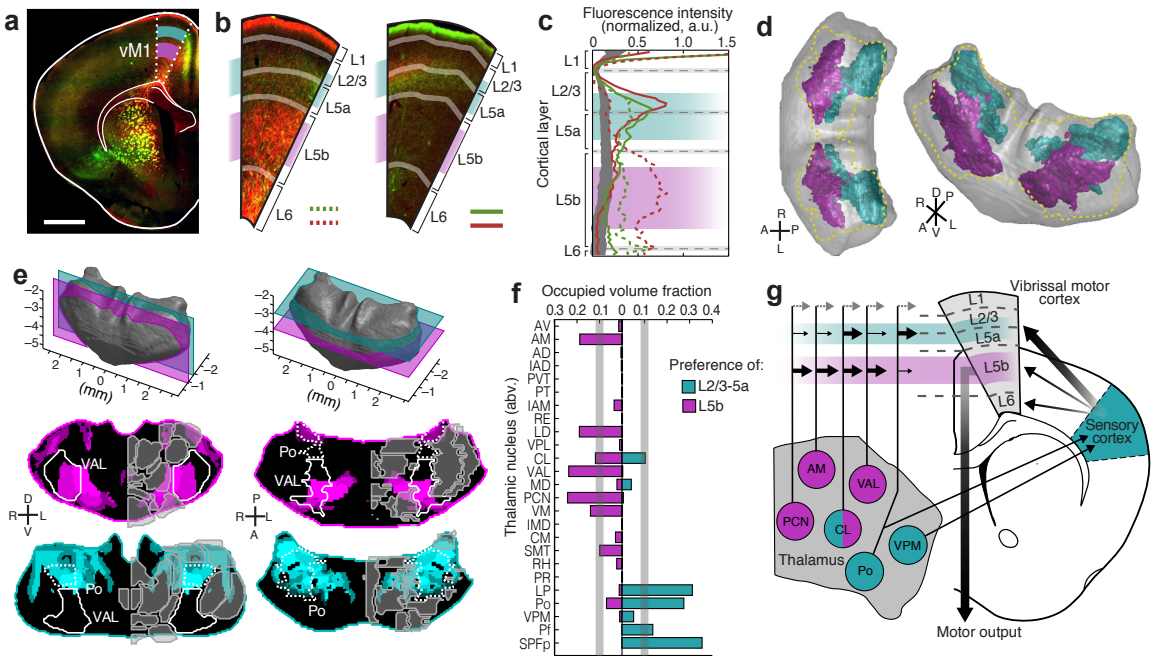
(**a**) Optimal injection coordinates for dACC, i.e. the most probable location to inject in order to target thalamic projections to dACC, determined from the confidence maps in Fig. 3. Anterior (left), dorsal (middle), and lateral (right) views of the target thalamic volume. (See **Supplementary Fig. 11** for details). Shown in millimeters relative to bregma. (**b–c**) Optimal injection coordinates were used to target thalamic projections to dACC. (**b**) Image of thalamic axons expressing fluorescently-tagged channelrhodopsin 2 (ChR2) in dACC. Orange circles indicate the location of two neurons recorded sequentially in layers 1 and 2/3 of dACC during optogenetic activation of the ChR2 expressing thalamic axons. White stars indicate the location of ChR2 stimulation by blue light (8 x 12 grid, 50 μm spacing). (**c**) Current recordings of the two neurons shown in panel b, showing synaptic currents elicited by light stimulation of thalamic axons. Each current trace corresponds to the white star grid in panel b rotated 30° counter-clockwise. The center of each circle indicates the location of the cell body. (**d**) The approximate locations of all neurons recorded (white circles) are shown, and crossed circles indicate no postsynaptic response. The approximate position shows the cortical layer (superficial: layer 1, middle: layers 2–4, and deep: layers 5–6), and the anterior-posterior extent of each area is collapsed into a single section (schematic modified from PMBA).

Figure 6. Nuclear localization of the thalamic origins of frontal projections.



(**a–c**) Three representations of the nucleus origin data for LO (See **Supplementary Fig. 12** for remaining cortical sub-regions). (**a**) The fractions of each thalamic nucleus projecting to LO are shown in three confidence levels (dashed lines) with their average (black line). Vertical gray line: the inflection point in the color scale used in panels b, c, and e. Asterisks indicate potential differences between localized thalamocortical projection origins and literature data in rat (see **Supplementary Fig. 13a** for details). (**b**) Single coronal section through the confidence map for LO (gray scale) overlaid with nuclear subdivisions from the ABA. The atlas is colored on the left to indicate the fraction of each nucleus covered by the average confidence trace (black line in panel a), with the inflection point (white) at 15%. (**c**) Spatial representation of all nuclei projecting to the LO. Circle diameters correspond to the relative size of each nucleus and their positions correspond to their relative center-of-mass location within the thalamus in the anterior-posterior and medial-lateral axes. Color scale is the same as in panel **b**. (**d**) The fractions of each thalamic nucleus projecting to AI, IL, and FrA, shown in three confidence levels (dashed lines) with their average (black line), as described in panel **a**. (**e**) Aggregate nucleus coverage map for all cortical areas. Nuclei (rows) and cortical sub-regions (columns) are hierarchically clustered based on output and input similarity, respectively. Color scale is the same as panel **b**.

Figure 7. Cortical layer preferences of thalamic projections to primary vibrissal motor cortex (vM1).



(**a**) Coronal brain section showing layers 2/3 and 5a (L2/3–5a, cyan) and layer 5b (L5b, magenta) in vM1 (white dashed outline). (**b**) Example coronal sections of vM1 (left image magnified from panel **a**), showing thalamocortical projections with preference to L5b (red projections, left), and L2/3–5a (red and green projections, right). (**c**) Normalized fluorescence intensity plots for red and green projections in panel b (left image: dashed lines, right image: solid lines). Fluorescence is averaged radially across vM1 to determine layer preference (see **Methods** and **Supplementary Fig. 14**). Background fluorescence is calculated from brains without vM1 projections (gray line). (**d**) Dorsal (left) and oblique (right) views of a 3D thalamus rendering, showing the volumes that are associated with preferential axonal projections to L2/3–5a (cyan) and L5b (magenta). The total volume projecting to all layers of vM1 is shown (gold dashed line). (**e**) Representative coronal (left) and horizontal (right) sections of modified confidence maps for L2/3–5a (cyan) and L5b (magenta) preferential vM1 projections. Outlines of thalamic nuclei are overlaid on each section image. (**f**) The occupied fraction of each thalamic nucleus containing layer-preferential projections to L2/3–5a and 5b of vM1. An arbitrary 10% threshold is indicated (gray line). (**g**) Schematic showing layer preferential input from thalamus to vM1 in the context of a motor-sensory circuit diagram. All scale bars are 1 mm.

Supplementary Information

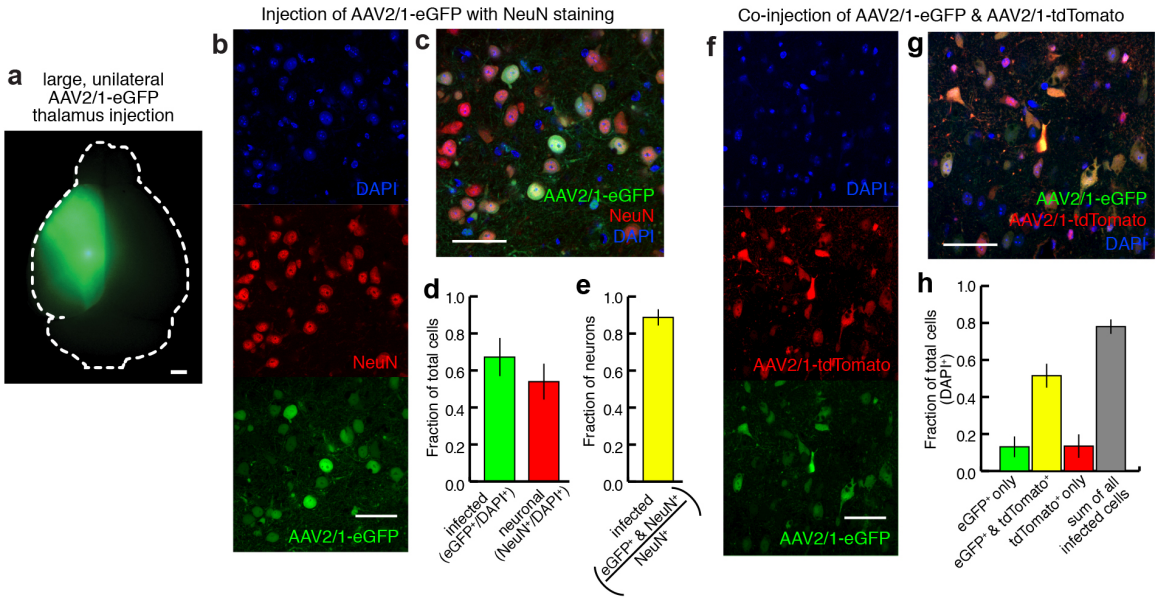
Supplementary movies 1–9. Full confidence maps as 3D stacks for all analyzed cortical subregions.

Distance from bregma for each section is indicated. All scale bars are 1 mm. These files can be accessed through the OHSU Library at http://digitalcollections.ohsu.edu/confidencemaps.

Supplementary Table 1 Correction methods for brain-to-brain variability.

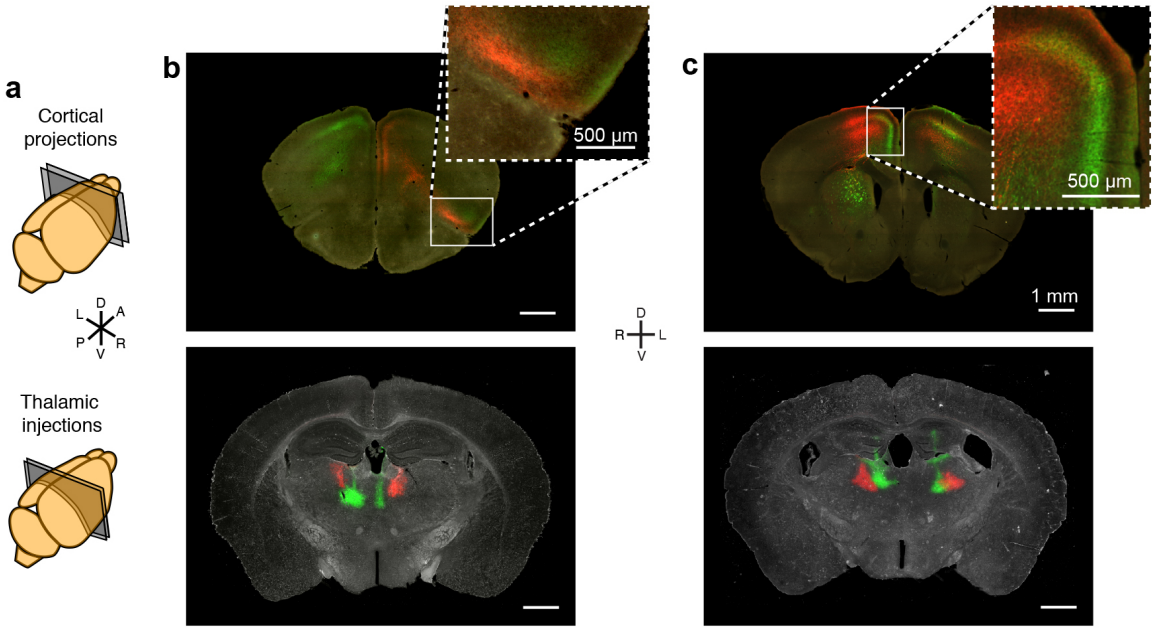
Source of variability	Correction Method
1. Section rotation and misalignment in the x (medial-lateral) axis	Rotate and align individual thalamic mask sections using two hand-selected midline points along the center y (dorsal-ventral) axis.
2. Cutting angle tilt about the x axis (i.e. rotation around the x axis)	Cutting angle tilt relative to the ABA thalamus is estimated from the z (anterior-posterior) coordinates of: the anterior and posterior points of the corpus callosum along the midline, the medial posterior point of the anterior commissure, and the anterior dentate gyrus, as illustrated in **Supplementary Figure 5**. The tilt is corrected at step 3.
3. Misalignment along the y axis	Multiple steps were involved. A) To control for the tilt about the x axis, the ABA thalamic mask is rotated to the tilt angle determined in step 2 and re-sampled as 50 μm slices. B) The centers of mass of individual thalamic mask sections were aligned to the corresponding adjusted atlas coordinates. C.) The aligned thalamus is rotated back to correct for its tilt angle.
4. Cutting angle tilt about the y axis (i.e. rotation around the y axis)	The 3D thalamus mask is sheared to minimize left-right asymmetry. Unlike a pure rotation, this transformation preserved the midline locations in our coordinate system, which were determined by hand-selected points. The degree of shear correction is determined by measuring the left-right asymmetry between 0.7 mm-wide sagittal bands located 0.4 mm lateral to the midline.
5. Overall size of the thalamus	A) Each 3D thalamic mask was scaled in z so that the distance along the midline between 1% (anterior) and 99% (posterior) of the thalamus voxels in a 110 μm sagittal band matched that of the ABA thalamus. B) The thalamus is scaled isotropically in x-y based on matching the central volume (0.8 mm thick in the A-P axis) to the corresponding ABA thalamus volume. C) The thalamus is scaled in x (medial-lateral) to the widest point of the ABA thalamus.

Supplementary Figure 1. Characterization of viral injections.



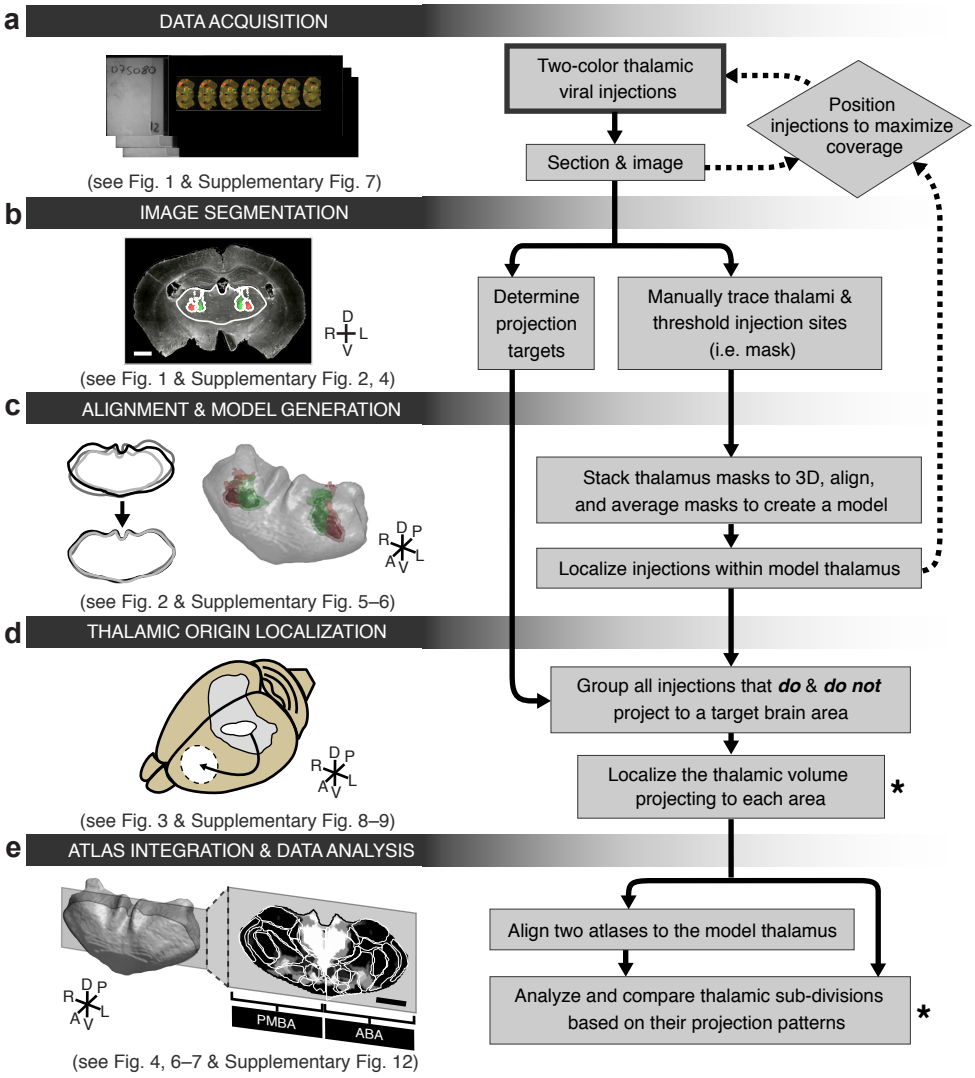
(**a**) Dorsal view of a mouse brain (dashed white outline) after receiving a large, unilateral thalamic injection (~100 nL); demonstrating that thalamic projections do not cross the midline in mouse. (**b–c**) Representative confocal image at the center of a thalamic injection site showing all cell nuclei (DAPI, blue), neurons (NeuN, red), and virally labeled cells (eGFP, green), shown individually in **b**, and overlaid in **c**. (**d**) The fraction of total cells infected at the center of an injection site (n = 17 green bar, mean ± s.d.). The fraction of total cells that are neurons at the center of an injection site (n = 5: red bar, mean ± s.d.). (**e**) The fraction of neurons infected at the center of an injection site (n = 5: yellow bar, mean ± s.d.). (**f–g**) Representative confocal image at the center of a thalamic co-injection of two viruses showing all cell nuclei (DAPI, blue), cells infected by virus expressing the fluorophore tdTomato (red), and cells infected by virus expressing the fluorophore eGFP (green), shown individually in **f**, and overlaid in **g**. (**h**) Quantification of the fraction of total cells expressing one or both co-injected viruses, showing that total infectable cell population is larger than a single injection will infect on average (n = 4, mean ± s. d.). All scale bars are 50 μm.

Supplementary Figure 2. Two-color injections reveal topographic projection information.



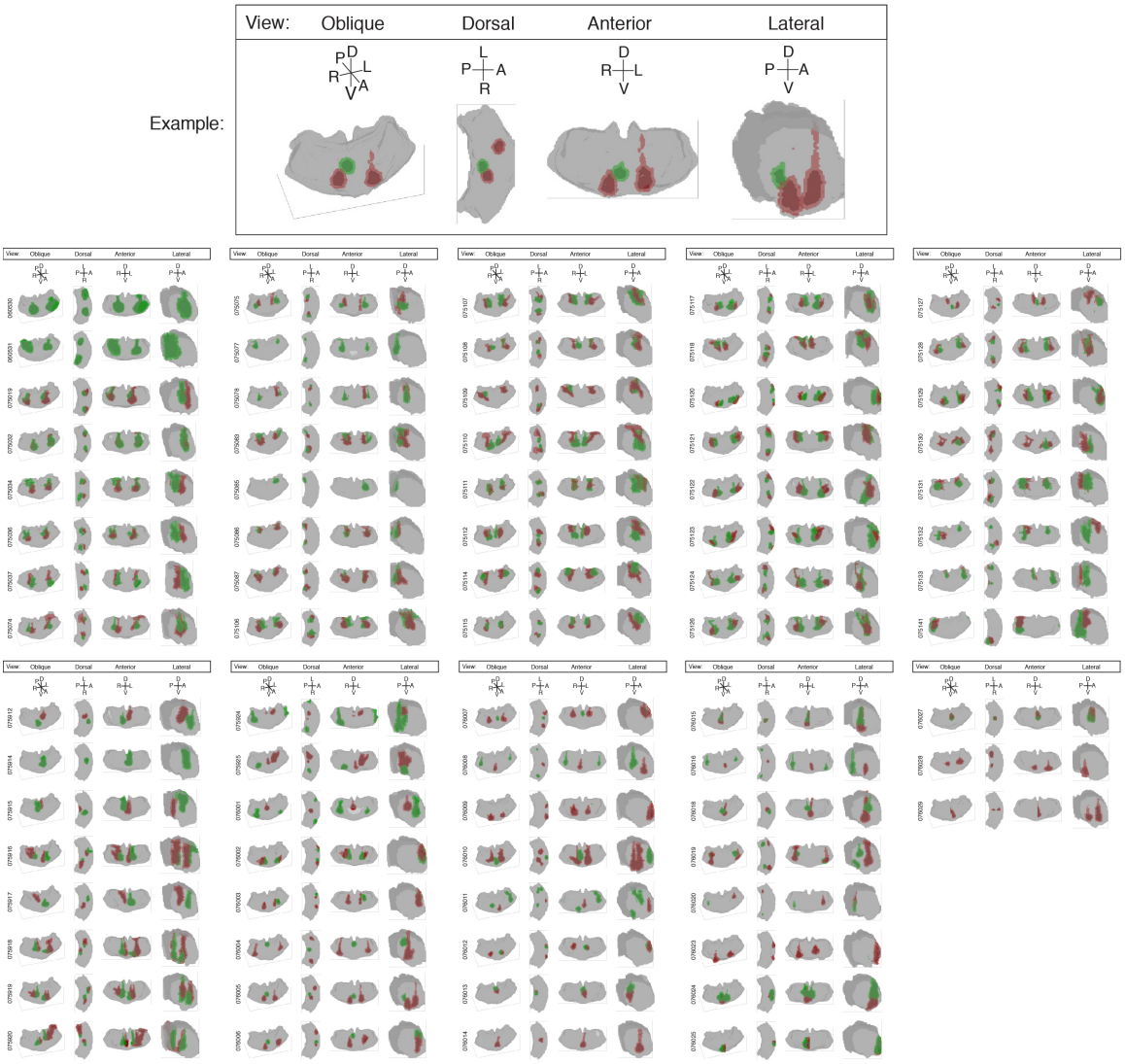
(**a**) Illustration showing the section locations for images in **b** and **c**. (**b–c**) Example thalamus sections showing injection sites (bottom), and example cortical sections showing thalamocortical projections (top). The insets for **b** and **c** show a zoomed image of juxtaposed red and green projections. (**b**) The green injection in the right hemisphere projects to AI, while the red injection, which originated lateral to the green injection in the thalamus, sends axonal projections to more medial brain structures, including LO, which is immediately adjacent to AI. (**c**) The green injection in the left hemisphere projects primarily to cortical layers 1 and 2/3 in the cingulate cortex, while the red injection projects to all cortical layers of motor cortex immediately adjacent to the green projections, largely maintaining separation of their respective projection fields. All scale bars are 1 mm unless otherwise specified.

Supplementary Figure 3. Schematic flow chart of anatomical tracing and analysis methods.



(**a**) Data acquisition: the entire thalamus is labeled with small viral injections, sectioned, and imaged at a 0.5 μm/pixel resolution. (**b**) Image segmentation and projection identification: binary masks of the thalamus and injection sites are created and cortical sub-regions are scored for the presence or absence of projections. (**c**) Data alignment and model generation: thalamus masks are rotated, sheared, scaled and aligned. The aligned masks are then averaged to create a model thalamus. All injection sites are subsequently mapped onto the model thalamus. (**d**) Thalamic origin localization: injection site location and projection information are combined to define the precise thalamic volume projecting to a region of interest. (**e**) Atlas integration and data analysis: thalamic output patterns are determined using either the classic nuclear subdivision of the Paxinos Mouse Brain Atlas (PMBA) and the Allen Brain Atlas (ABA), or through atlas independent projection analyses. The thick-lined box indicates the starting point of the experiment while asterisks indicate the major output of the data analysis. All scale bars are 1 mm.

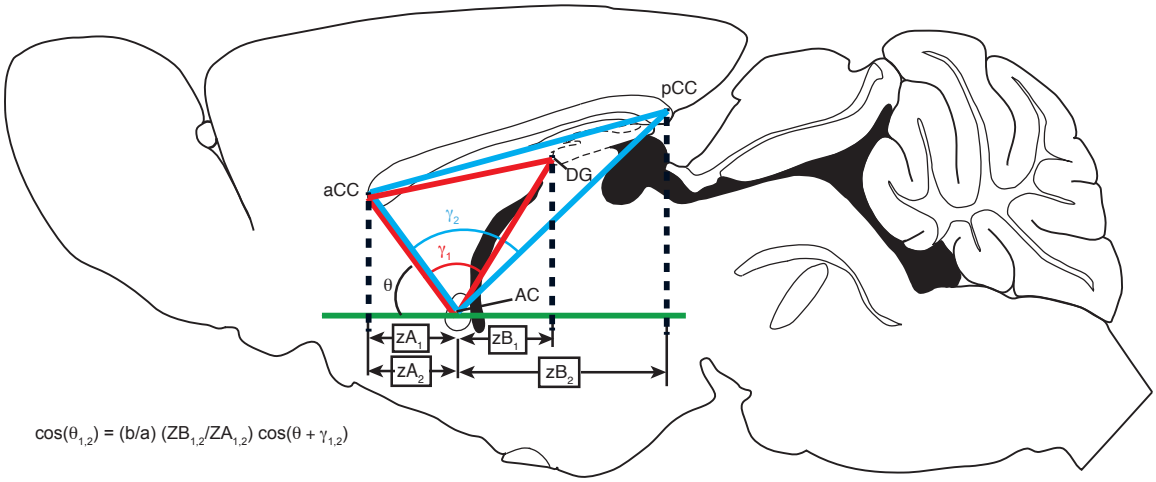
Supplementary Figure 4. 3D rendering of the thalami and corresponding viral injections of all 75 experimental brains.



Each row shows an oblique, dorsal, anterior, and lateral view of an individual experimental thalamus. Axis descriptions are shown for the first thalamus only. Viral injection sites (red/green fluorescence) are shown within the traced thalamus from each brain (gray). A darker center of each injection site represents the eroded 'core' of the injection as described in Figure 1.

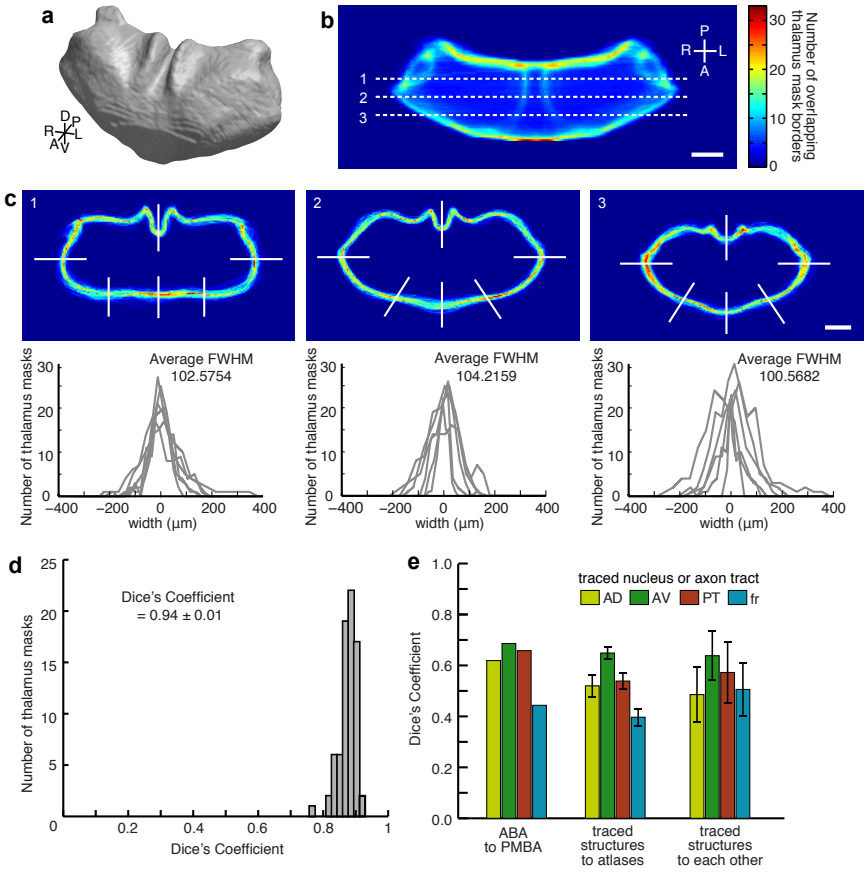
* This figure can be accessed at higher resolution through the OHSU Library at http://digitalcollections.ohsu.edu/items/show/26984.

Supplementary Figure 5. Slice angle estimation from anatomical landmark positions.



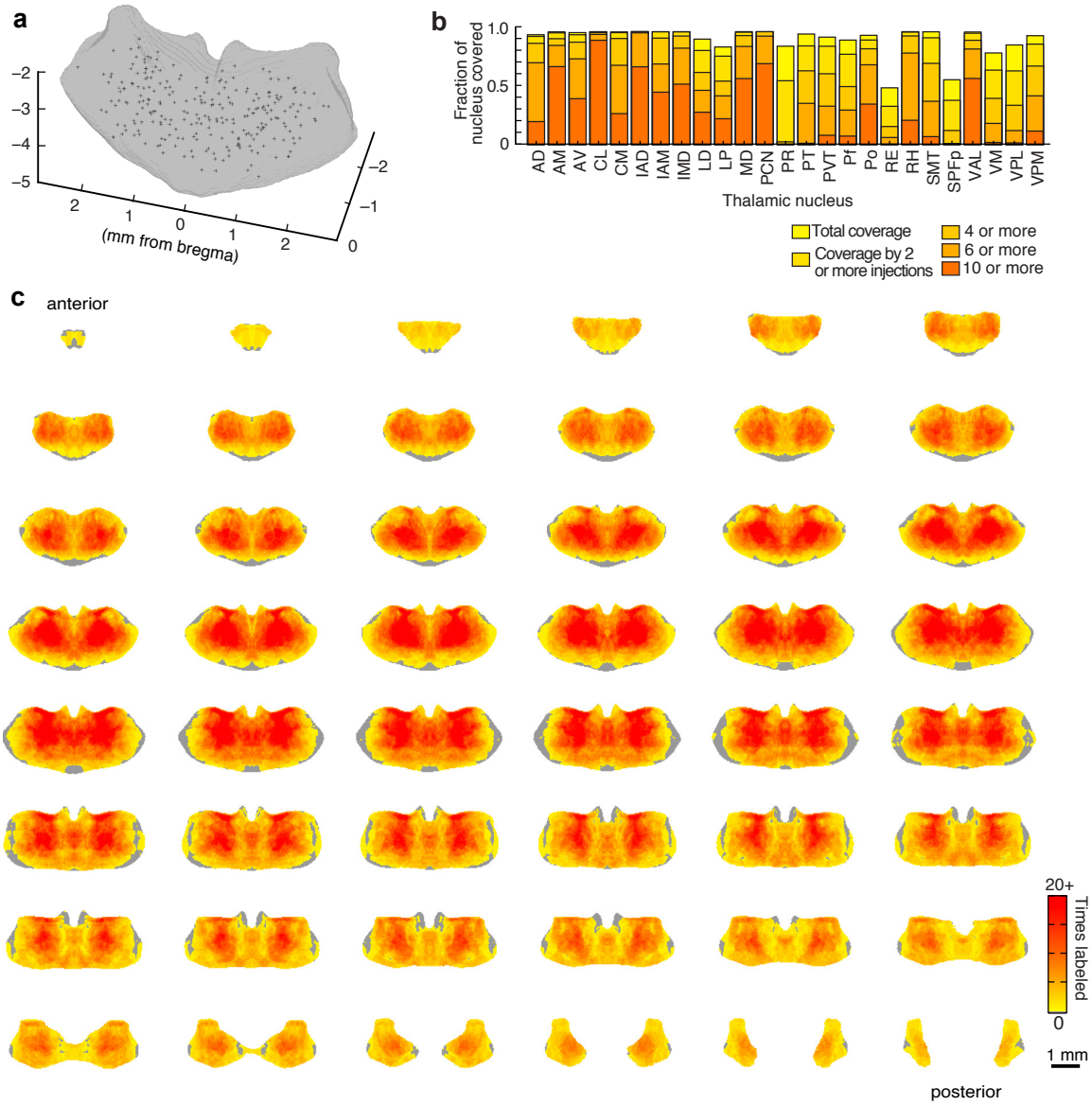
Slice angle is estimated using the slice locations of several landmarks. In each case, the landmarks form a triangle, either between the medial anterior corpus callosum (aCC), the anterior commissure (AC) and the anterior dentate gyrus (DG) (i = 1, red) or between the aCC, the AC, and the medial posterior corpus callosum (pCC) (i = 2, blue). Assuming the sides of the triangle are fixed and proportional to the atlas dimensions (ABA), we can define the angles γ_1 and γ_2, using the red and blue coordinates, respectively. The fixed sides of the red triangle constrain the relationship between zA_1 and zB_1 so that θ_1 (estimated cutting angle) can be found from slice locations zA_1 and zB_1 only (and similarly for θ_2). θ_1 and θ_2 were determined from the equation numerically, and the final cutting angle tilt estimation was the average of θ_1 and θ_2. Baseline θ_{ABA} values for ABA sections were measured from ABA landmark locations, and brains with $|\theta_C - \theta_{ABA}| > 2°$ were subjected to rotation correction.

Supplementary Figure 6. Variability across thalamus masks and cytoarchitecturally identifiable thalamic structures.



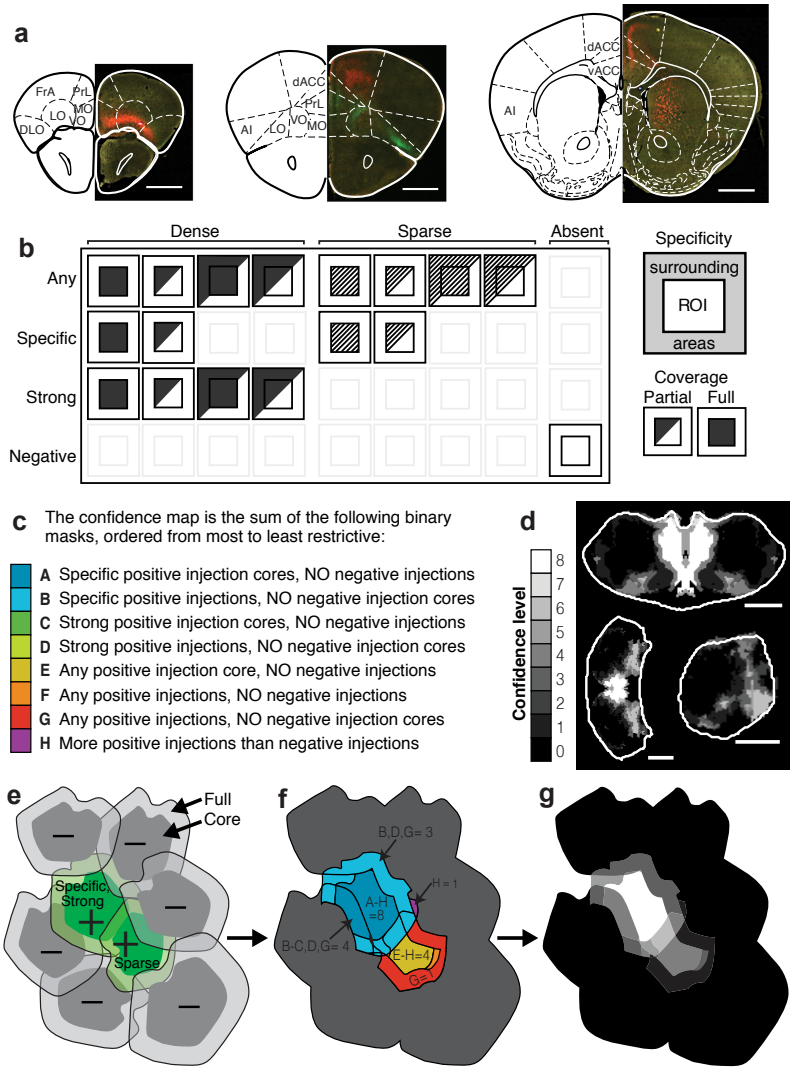
(**a**) 3D rendering of the model thalamus created by averaging individual thalamus masks (oblique view). (**b**) To assess thalamus mask variability, we created the 3D perimeter matrix *P*, which contains the sum of the perimeter voxels of all 75 experimental thalamus masks. Color map indicates the summed projection of *P* along dorsal-ventral axis. (**c**) The top panels show x-y cross sections through *P* at the z locations marked in panel **b** as 1, 2, and 3. The bottom panels show profiles at the locations indicated in top of each panel (white lines). Full width half maximum (FWHM) across all profiles is 102 ± 51 μm (mean ± s.d.). (**d**) Histogram of Dice's Similarity Coefficient for individual thalamus masks. Individual thalamus masks are highly similar to the model thalamus, with Dice's coefficient between individual masks and the model $D = 0.94 \pm 0.01$ (mean ± s.d.). (**e**) Four cytoarchitecturally identifiable thalamic structures (AD, AV, PT, and fr) were manually traced in five experimental brains and two atlases (PMBA and ABA), as described in **Figure 2b–c**. Dice's Similarity Coefficient is plotted to compare the identified structures in two separate atlases (left, n = 2), to compare the atlases to the experimentally traced structures (middle, n = 5, mean ± s.d.), and to compare the experimentally traced structures to each other (right, n = 10, mean ± s.d.). All scale bars are 1 mm.

Supplementary Figure 7. Full characterization of injection coverage within the model thalamus.



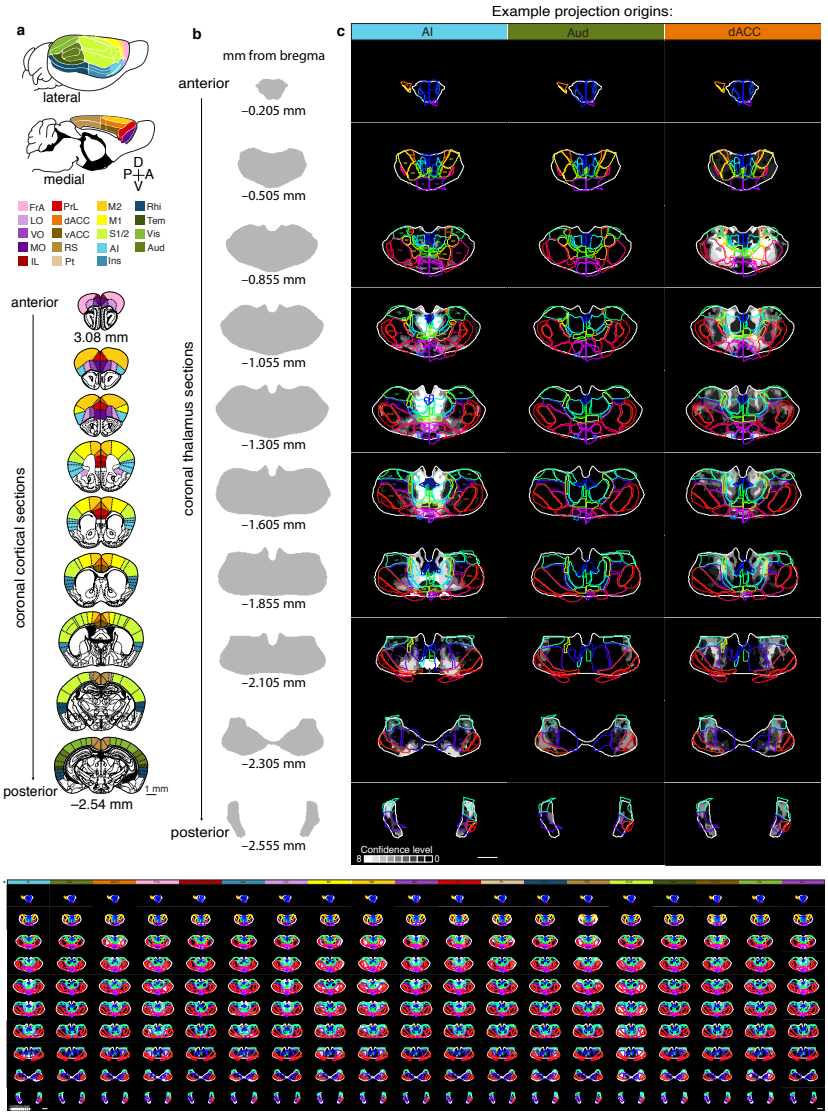
(**a**) Oblique view of the model thalamus (gray) with the center of mass of all 254 injection sites (black crosses). (**b**) Coverage of the thalamus broken down by nucleus. Fraction of each nucleus covered by at least 1, and at least 2, at least 4, at least 6 and at least 10 injections, showing that the entire volume of most thalamic nuclei is sampled at least 4 times. (**c**) Full coverage montage of the thalamus. Coronal slices through the thalamus starting at –0.155 mm posterior to bregma, and continuing in 50 μm increments to the posterior end of the thalamus. The color scale indicates the number of times an independent injection labeled a given location of the thalamus.

Supplementary Figure 8. Projection scoring criteria and injection grouping method.



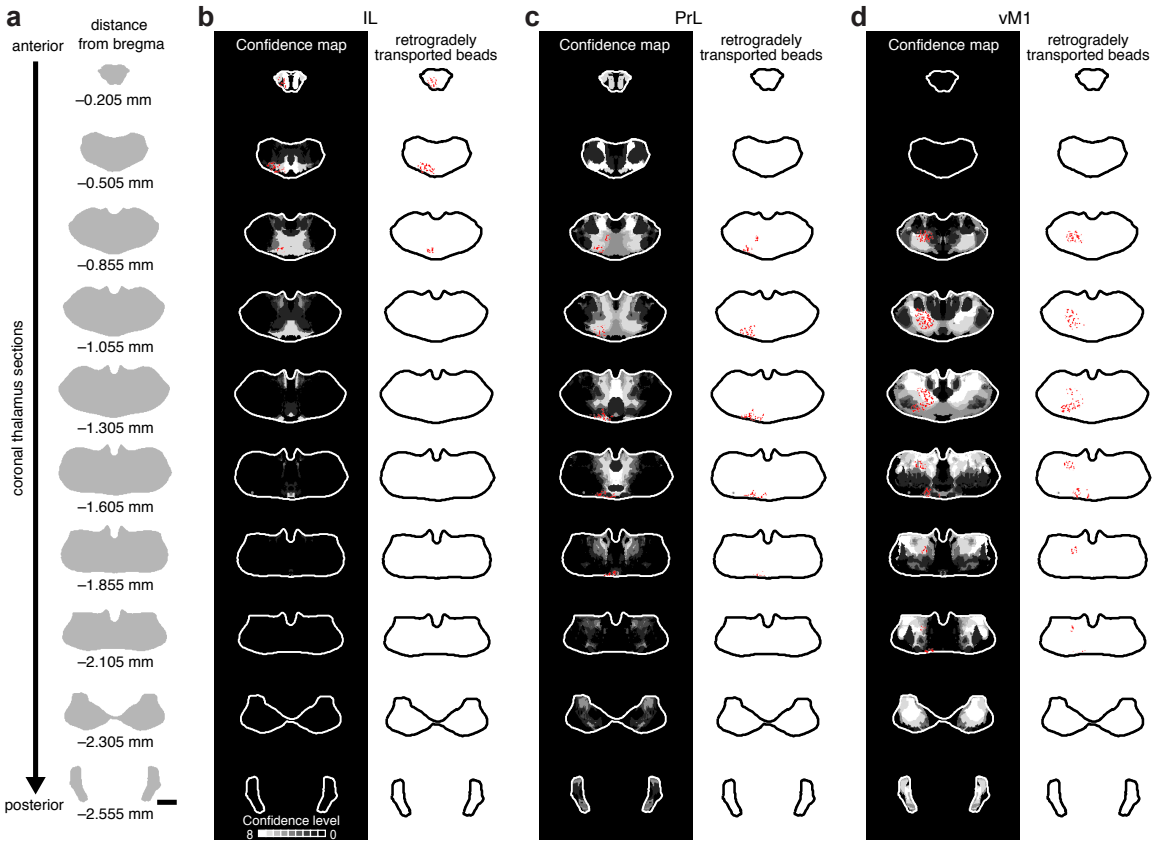
(**a**) PMBA sections with corresponding mouse brain images on the right, showing example projection distributions within several regions of interest (ROIs) in the frontal cortex. (**b**) Diagram of projection scoring criteria: strength (either dense or sparse labeling within the ROI) and specificity (either projections locally confined to the ROI or projections that also go to surrounding areas). Partial or full coverage within the ROI was also scored, but differences in thalamic origin profiles were not found (data not shown). (**c**) Description of each binary mask created by grouping the scored injections as described. (**d**) Example confidence map created by summing the binary masks created in **c**, where each mask adds a value of 1, and all masks (A–H) must be true at a given point to reach a confidence level of 8 (white). (**e–g**) Illustration of how a confidence map is created. (**e**) Illustration of eight injections and their corresponding score, showing the full injection as well as the eroded core. (**f**) Sample of the binary masks created from panel **e** as described in panel **c**. (**g**) Confidence map created from panel **f** as described in panel **d**. All scale bars are 1 mm.

Supplementary Figure 9. Full confidence maps for the thalamic origin of all cortical projections.



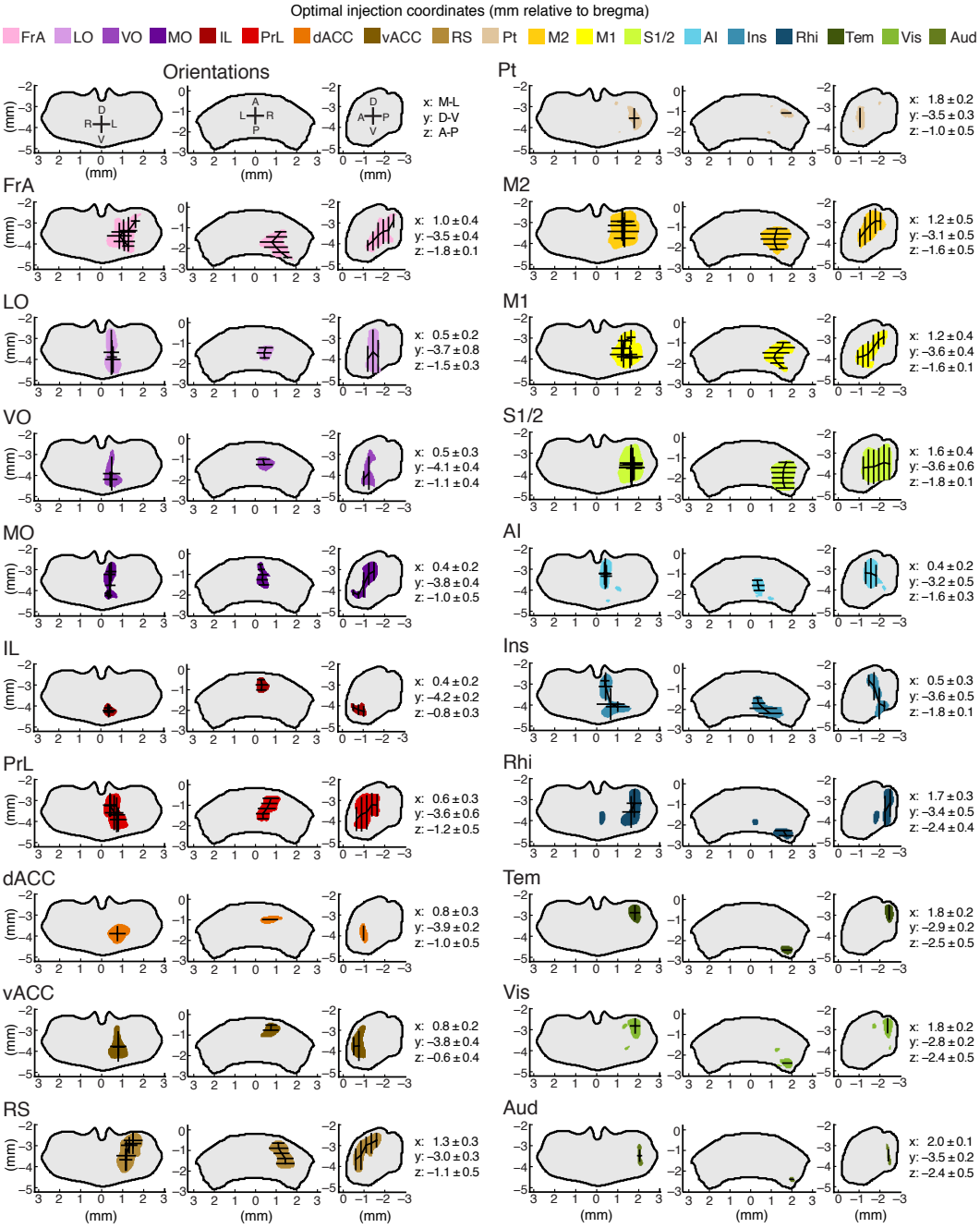
(**a**) The cortical areas examined (modified from PMBA[25] and reference 47) are shown. Lateral and medial whole brain views (top) and coronal sections from anterior to posterior (bottom). (**b**) Reference coronal sections for the confidence maps shown in panel **c**, with the bregma coordinates noted in the anterior-posterior axis. (**c**) Each column contains the confidence map for a given cortical sub-region, arranged alphabetically, with coronal sections corresponding to those in panel **b**. The confidence maps (gray scale) show the thalamic origin of projections to nineteen cortical sub-regions: FrA, LO, VO, MO, IL, PrL, vACC, dACC, RS, Pt, M2, M1, S1/2, AI, Ins, Rhi, Tem, Vis, and Aud. The outline of the model thalamus is shown in white for each section, and the aligned nuclei from PMBA (left) and the ABA (right), are overlaid as colored outlines (nuclei labeled for AI only). All scale bars are 1 mm. *This figure can be accessed at higher resolution through the OHSU Library at http://digitalcollections.ohsu.edu/items/show/26983.

Supplementary Figure 10. Retrograde bead injections fall within the thalamic volume predicted by the confidence maps.



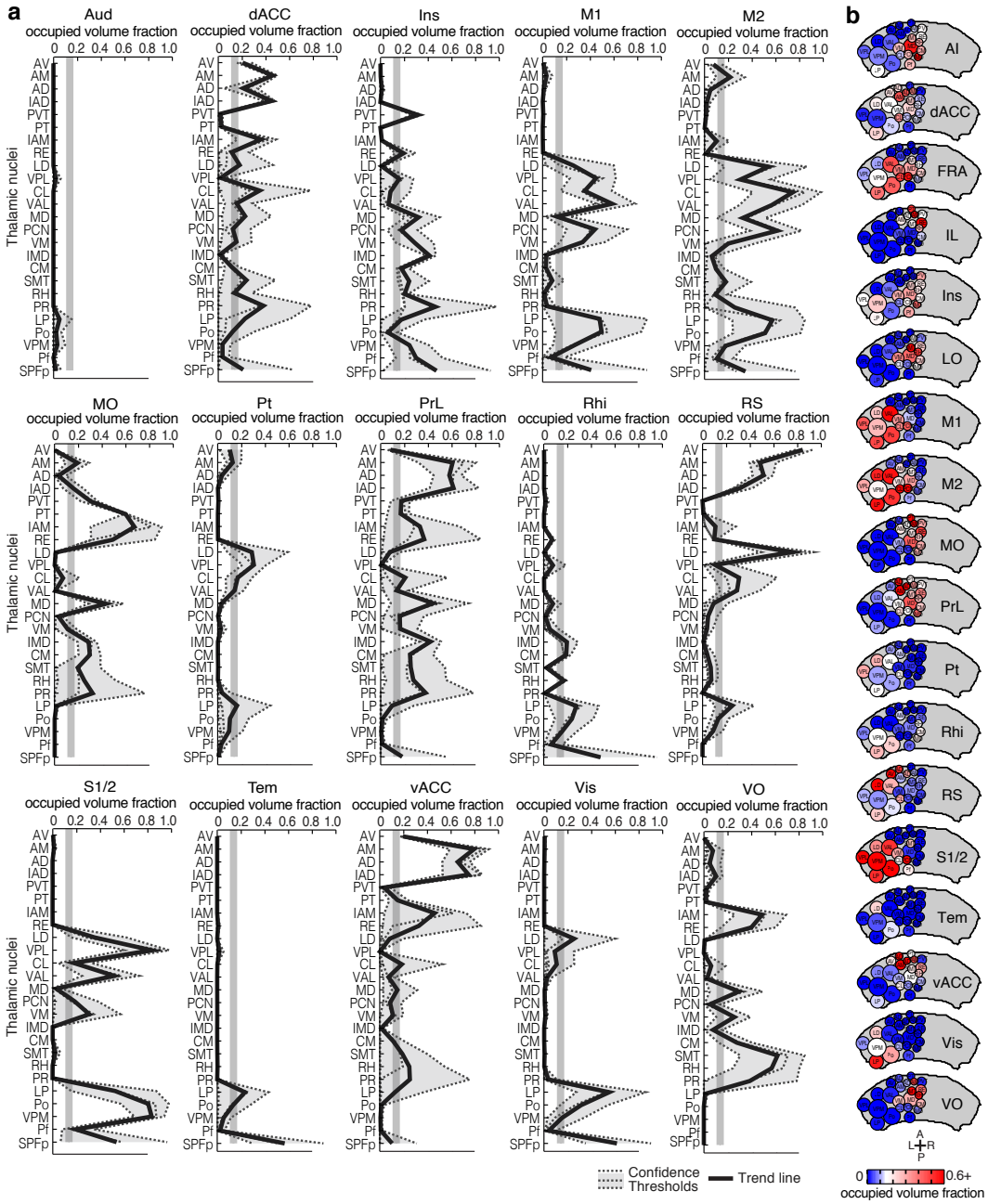
(**a**) Reference coronal sections for the confidence maps shown in panel **b–d**, with the bregma coordinates noted in the anterior-posterior axis. (**b–d**) Thalamic location of retrogradely transported beads 3 days after injection in the indicated cortical sub-region (total of 10 mice injected and 3 analyzed in detail). The confidence maps (gray scale, left: see **Fig. 3** for detail) show the predicted thalamic origin of projections to three cortical sub-regions; (**b**) IL, (**c**) PrL, and (**d**) vM1. The locations of beads that were retrogradely transported from the indicated cortical region are overlaid on the left of each confidence map section (red dots), and shown independently on the right. The outline of the model thalamus is shown for each section (left: white, right: black). All scale bars are 1 mm.

Supplementary Figure 11. Optimal injection coordinates used to target specific thalamocortical projections.



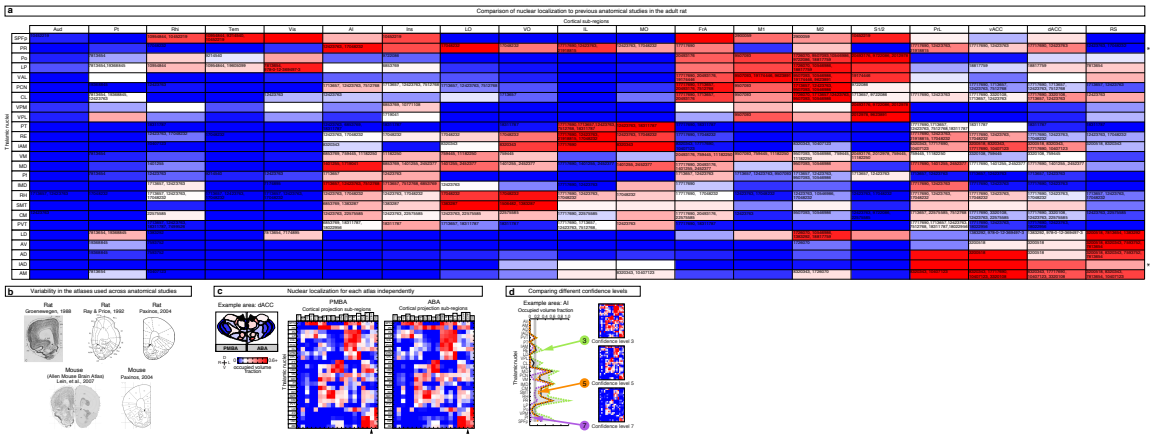
Optimal injection coordinates, i.e. the most probable location to inject in order to target a specific thalamocortical pathway, were determined for each of the 19 cortical sub-regions: FrA, LO, VO, MO, IL, PrL, vACC, dACC, RS, Pt, M2, M1, S1/2, AI, Ins, Rhi, Tem, Vis, and Aud. The orientation of each thalamus view is indicated in the top left column. The colored areas for each cortical projection sub-region denote the location with the highest probability to label those projections. A single set of coordinates is listed to the right of each cortical sub-region. All coordinates are shown in millimeters relative to bregma.

Supplementary Figure 12. Nuclear localization of the thalamic origins of cortical projections.



(**a**) The fractions of each thalamic nucleus projecting to Aud, dACC, Ins, M1, M2, MO, Pt, PrL, Rhi, RS, S1/2, Tem, vACC, Vis, and VO are shown in three confidence levels (dashed lines) with their average (black line). Vertical gray line: the inflection point in the color scale used in panel **b**. (**b**) Spatial representations of all nuclei projecting to the each cortical sub-region, arranged alphabetically. Circle diameters correspond to the relative size of each nucleus and their positions correspond to their relative center-of-mass location within the thalamus in the anterior-posterior and medial-lateral axes. Auditory cortex is not shown because we did not identify any projections from our injections.

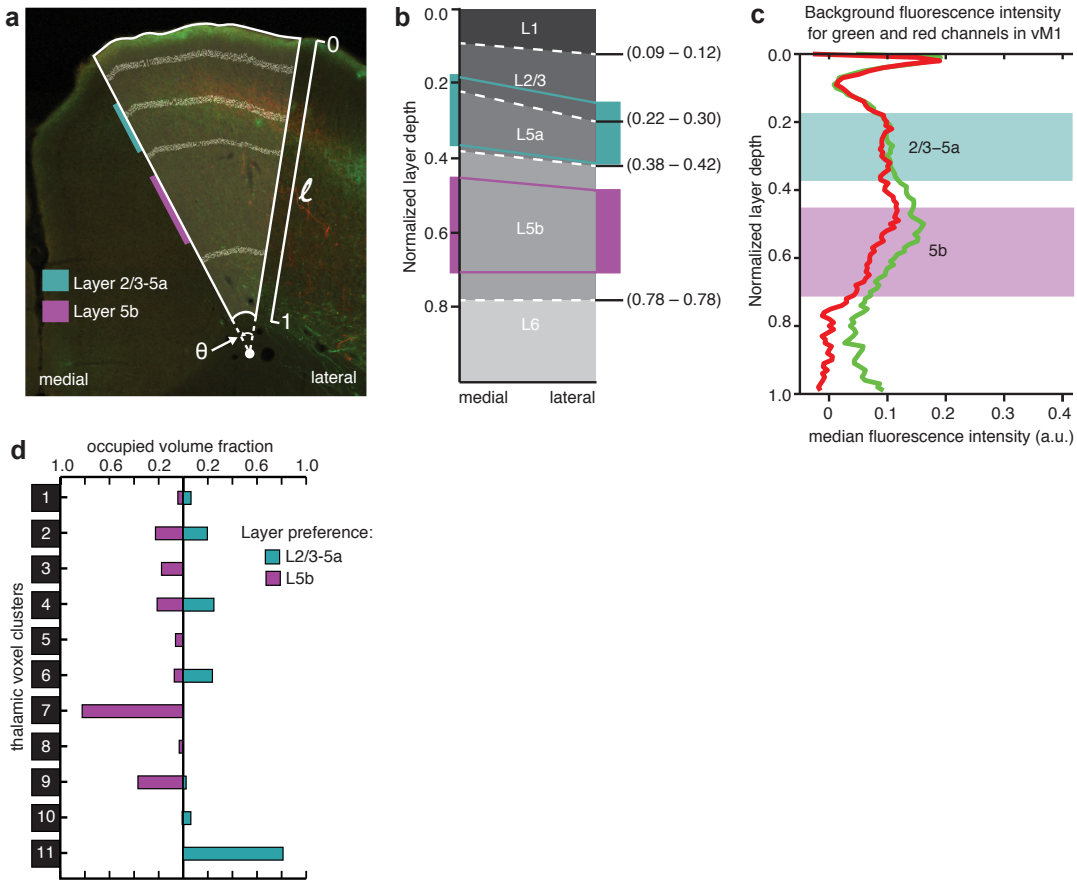
Supplementary Figure 13. Localizing and verifying the thalamic nuclear origins of cortical projections.



(**a**) Comparison of our thalamic nuclear localization data in mouse to previous studies of thalamocortical projections in the adult rat. Nuclear localization map as described in **Figure 6e**, overlaid with the PubMed ID #'s for publications with evidence for the corresponding thalamocortical projection. Some specific projections are inferred from data presented in the publications. Asterisks indicate thalamic nuclei with little or no anatomical data in the rat literature. (**b**) Examples of several atlases used in literature to classify cortical sub-regions in the literature. (**c**) Comparison of nuclear localization maps for two atlases separately (ABA and PMBA). Left: Coronal section showing the nuclear subdivisions from the PMBA (left) and ABA (right). Nucleus colors indicate the fraction of each nucleus covered by the dACC confidence map described in **Figure 6**. Right: Aggregate nucleus coverage map of all thalamocortical projections using the PMBA and ABA separately. Nuclei (rows) and cortical sub-regions (columns) are ordered as in **Figure 6e**. The ABA and PMBA specific maps were averaged to produce **Figure 6e**. (**d**) Comparison of nuclear localization maps for separate confidence levels (using averaged atlases). Left: The fraction of each nucleus projecting to AI, shown for three confidence levels (dashed lines, 3, 5 and 7) and their average (black line) as described in **Figure 6**. Right: Nucleus localization maps for the three confidence levels. Nuclei (rows) and cortical sub-regions (columns) are ordered as in **Figure 6e**. All three maps are averaged to produce **Figure 6e**.

* This figure can be accessed at higher resolution through the OHSU Library at http://digitalcollections.ohsu.edu/items/show/26985.

Supplementary Figure 14. vM1 layer boundaries and background fluorescence.



(**a**) Representative vM1 image showing the layer boundaries. Relative thicknesses of cortical layers in vM1 vary from the lateral edge to the medial edge. The parameters used in calculations are indicated: *l*: relative cortical depth (0 at the pia and 1 at white matter); θ: the angel of the ROI. Horizontal lines: layer boundaries. The depth regions analyzed for layer preferential projections are indicated (color bars). (**b**) Schematic of layer depth variation from medial to lateral edge. The boundaries at the medial and lateral edges are based on previous characterizations. At other radial positions, the layer boundaries were linearly interpolated from the lateral to medial edges. The boundaries were normalized to the cortical layer depth. Colored regions and lines indicate the regions analyzed for layer preferential projections. (**c**) Background florescence along the normalized cortical depth. All fluorescence was first mapped to the corresponding layer depth at the medial edge before being binned and averaged. Background fluorescence was calculated as the median fluorescence at a depth bin across brains manually determined not to have projections from the thalamus to vM1. (**d**) Assessment of 11 thalamic voxel clusters' layer-preferential projections to vM1. The clusters were determined as described in **Figure 4**. The occupied fraction of each thalamic voxel cluster containing layer-preferential projections to L2/3–5a (cyan) or 5b (magenta) of vM1 is plotted as a fraction of total cluster volume.

Acknowledgments & Author Contributions

We thank Dr. Karel Svoboda and HHMI, Janelia Farm Research Campus for their generous support in the initiation and data collection phase of this project. All Nanozoomer imaging were collected in Janelia Farm Research Campus. We thank Dr. Claudio Mello, Maozhen Qin, Dr. Jian Qiu, Jessica Li, Dr. Stefanie Kaech Petrie, Anthony Gilmore, Dr. Yi Zuo, and Dr. Daniel Lioy for technical support. We thank Drs. Karel Svoboda, Jeff Lichtman, Kun Huang and Bo Li for their discussions and Drs. Karel Svoboda, John Williams, and John Adelman for comments on the manuscript. We thank Robin Champieux and Kyle Banerjee at the OHSU Library for establishing a web data depository. We also thank the anonymous reviewers who gave critical suggestions that helped us to improve the data analysis, presentation and depository. This work is supported by an ARRA P30 grant (PI: Dr. Richard Goodman), an NSF-GRF fellowship (BJH), ARCS Foundation Portland Chapter (BJH), an NIH Director's Innovator Award (HZ), an NIH R01 grant (NS081071; TM), and the Medical Research Foundation (TM).

All authors participated in designing the experiments. BJH, DK, KG, and TM performed the experiments, BJH, BRL, KG, DK, HZ and TM analyzed the data. BJH, BRL, KG, HZ and TM wrote the manuscript.

CHAPTER 2:
A COMPREHENSIVE MAP OF EXCITATORY INPUT CONVERGENCE IN THE MOUSE STRIATUM

Barbara J. Hunnicutt[1-2], Katrina J. Gertz[1], Haining Zhong[1], and Tianyi Mao[1]

[1] Vollum Institute, Oregon Health and Science University, Portland, Oregon.

[2] Neuroscience Graduate Program, Oregon Health and Science University, Portland, Oregon.

Presented prior to publication

Forward

Two datasets were used in this chapter, the thalamic projection dataset was generated in **Chapter 1** of this dissertation (Hunnicutt et al. 2014), and cortical projection dataset was generated by the Allen Institute for Brain Science (AIBS) (Oh et al. 2014). The raw data from the AIBS was acquired at the 2014 Summer Workshop on the Dynamic Brain, co-hosted by the AIBS and the Computational Neuroscience Program at the University of Washington. Parts of this project evolved from group projects at this course with Peter Keating, Elle O'Brien, Agata Budzillo, and Nick Hollon. I would like to thank Peter, Elle, Agata and Nick, as well as the TAs from the Allen Institute for helping me formulate and test ideas in the initial stages of this project.

All analyses for this chapter were performed by myself and Katrina Gertz. I wrote all analysis code for the cortical projection data, as well as the final thalamostriatal analyses, but some parts of the thalamus analysis code were modified from code written by Brian Long for **Chapter 1** of this dissertation. I manually segmented each striatum, Haining Zhong and I designed the striatum alignment method, Haining wrote the striatum alignment program, and Haining and I all manually checked the final alignment for each brain. The manual training for the machine learning steps of the thalamostriatal projection data were performed by Katrina Gertz. Katrina and I manually checked and modified the machine learning output from both the thalamocortical and thalamostriatal datasets. I performed all remaining data analysis, generated all of the figures, and wrote the manuscript. Tianyi Mao provided input at all steps.

Abstract

The basal ganglia receive excitatory inputs from nearly the entire cerebral cortex, thalamus, amygdala and hippocampus. This confluence of multi-modal information allows the basal ganglia to affect behaviors ranging from movement to motivation, yet the precise structural organization necessary to produce these unified outputs is poorly understood. To investigate the extent of input integration in the basal ganglia, we mapped all excitatory projections to the basal ganglia input nucleus, the striatum. These projection maps allowed us to determine the comprehensive, three-dimensional distributions of excitatory striatal input convergence. We were also able to create the first objective division system for the striatum, and identify several putative functional subdomains. Finally, since the thalamus is both a primary input and output target of the basal ganglia, we detailed the complete organization of the thalamus in the cortico-thalamo-basal ganglia circuit. Together, these findings describe the basal ganglia circuit with both a breadth and detail that has never before existed, and will serve as a guide for future studies of basal ganglia function.

Introduction

The basal ganglia are a group of forebrain nuclei integral to movement, decision making, and motivation (Yin 2014; DeLong & Wichmann 2007; Albin et al. 1989). Diseases of the basal ganglia can affect each of these behaviors independently, supporting a view that these functions are governed by distinct processes within the basal ganglia circuitry (DeLong & Wichmann 2007; Plotkin & Surmeier 2015; Redgrave et al. 2010; Volkow et al. 2012).

The primary input nucleus of the basal ganglia, the striatum, receives excitatory projections from nearly all of the cerebral cortex, as well as much of the thalamus, hippocampus, and amygdala (Berendse et al. 1992; Berendse & Groenewegen 1990; Wall et al. 2013; McHaffie et al. 2005). This wide array of inputs provides the substrate necessary to accomplish the disparate functions performed by the basal ganglia. However, very little is known about the structural organization necessary to produce a unified output in each case.

Structurally, the striatum is made up of millions of spherical neurons (Oorschot 1996; Wilson 1987), 96% of the which are one of two nearly identical and homogeneously distributed cell types (Gertler et al. 2008; Steiner & Tseng 2010). Considering this cellular homogeneity, along with the fact that the striatum receives overlapping inputs from almost every other brain region (Veening et al. 1980; McGeorge & Faull 1989; Berendse & Groenewegen 1990; Wall et al. 2013; Guo et al. 2015), it is not surprising that the details of striatal function have eluded researchers for decades, or that the striatum remains the largest part of the telencephalon without universally accepted subdivisions.

The distribution of striatal inputs from the three subdivisions of telencephalic cortex, the neocortex, mesocortex and allocortex, can be used to roughly divide the striatum into three domains, the sensorimotor, associative, and limbic (McGeorge & Faull 1989; Parent & Hazrati 1995). These areas are thought to constitute functionally segregated circuits. The limbic domain receives information from the hippocampus, amygdala, and other limbic cortices on affective state, motivation and reward; the associative domain integrates a variety of higher cognitive inputs as well as overlapping projections from both neocortical and allocortical areas; and the sensorimotor domain has a dense convergence of sensory and

motor information (Joel & Weiner 1994; Gruber & McDonald 2012). The boundaries of these domains are ill-defined, but they generally occupy the ventral, dorsomedial, and dorsolateral striatum, respectively (Draganski et al. 2008).

The remaining excitatory inputs to the striatum come from the thalamus, which accounts for ~1/3 of all glutamatergic striatal synapses (Huerta-Ocampo et al. 2013). The thalamus can be broken into as many as 40 nuclei, most of which have been shown to send projections to the striatum (Berendse & Groenewegen 1990; Veening et al. 1980; Wall et al. 2013). The midline nuclear group, made up of the paraventricular (PVT), parataenial (PT), interanterodorsal (IAD), interanteromedial (IAM), rhomboid (Rh) and reuniens (Re) nuclei, project primarily to the ventral striatum (Van der Werf et al. 2002; Li & Kirouac 2008). The intralaminar nuclear group, made up of the central medial (CM), paracentral (PCN), central lateral (CL), and centromedian–parafascicular (Pf) nuclei, project to the majority of the striatal volume (Y. Smith et al. 2004; Van der Werf et al. 2002), and several of the ventral and posterior nuclei, including the ventromedial (VM), ventral anterolateral (VAL), and posterior (Po) nuclei, send broad topographic projections to the dorsal striatum (Erro et al. 2001; Veening et al. 1980). These projection distributions are reminiscent of the striatal areas encompassed by the limbic, associative, and sensorimotor domains, respectively.

To identify the full range of input convergence for all excitatory inputs to the mouse striatum, we aligned and analyzed two large anatomical datasets consisting of 335 anterograde tracer injections. The first dataset contains highly redundant injections in the thalamus, and was originally generated in our lab to construct a comprehensive map of thalamocortical projections (Hunnicutt et al. 2014). The second dataset is a subset of the Allen Mouse Brain Connectivity Atlas, which was created by the Allen Institute for Brain Science (AIBS) as a freely available online resource (Oh et al. 2014), (http://connectivity.brain-map.org/). Here, we have utilized the high-resolution and complementary advantages of the two datasets to map the axonal projection distributions for all of the excitatory inputs to the mouse striatum. The result is the most complete three-dimensional

map of input convergence ever created for the striatum, and has allowed for the objective parcellation of the striatum into anatomically defined sub-regions.

Materials & Methods

Thalamus data generation overview

Thalamic injection data was generated as described previously (Hunnicutt et al. 2014), but will be explained briefly to highlight the similarities and differences of the two datasets used in this study. Injections were performed in male and female wild-type C57BL/6J mice using a hydraulic apparatus to stereotaxically pressure inject ~10 nL of recombinant adeno-associated virus (serotype 2/1; AAV2/1) encoding either eGFP (Addgene 28014) or tdTomato (A gift from J. Magee) (Harris et al. 2012). Two weeks post injection, at postnatal day 30±2, each brain was fixed in 4% paraformaldehyde, cryostat (Thermo Scientific) sectioned at 50 μm, and imaged using a Hammamatsu Nanozoomer imaging system, resulting in 0.5 μm/pixel lateral resolution for the full-brain fluorescence image sets of all injections. Injection sites were imaged at a lower exposure on either the Nanozoomer or a Zeiss Axio Imager to avoid overexposure. Injection site images were matched to their corresponding full brain Nanozoomer section images through rigid translation and rotation using manually selected anatomical landmarks visible in both images. The thalamus was manually traced from the full brain images, and injection sites were segmented from background fluorescence in the green and red channels using a supervised MATLAB routine (Otsu 1975). The injection sites and thalami were aligned as described previously (Hunnicutt et al. 2014).

Thalamostriatal projection segmentation

The outline of the striatum was manually traced in each image set to generate a striatum mask. The front of the striatum was defined as the first slice containing the nucleus accumbens (NAc), where the anterior commissure (ac) separates from the rostral migratory stream. The border of the dorsal striatum was determined by the lateral ventricle (VL) and corpus callosum (cc). The ac was included in the striatum mask until it became medial of the VL, and, once it fuses across the midline, formed the ventral border of the striatum. The

globus pallidus (GPe) and internal capsule formed the medial border of the striatum in more posterior sections.

To localize thalamostriatal projections, and distinguish them from passing thalamocortical axons, the machine-learning plugin for ImageJ, Trainable WEKA Segmentation (http://fiji.sc/Trainable_Weka_Segmentation) was used. To prepare the image sets for training, each slice containing striatum was background subtracted, a 12-pixel gaussian filter was applied, and the striatum mask was used to limit the image to only the striatal volume. The images were then split into single channels (red or green) and converted to an 8-bit grayscale format.

The WEKA Segmentation program was manually trained to distinguish between diffuse (thalamostriatal), bundled (thalamocortical), or residual background fluorescence. Visually, thalamocortical axons can be identified as being highly directionally oriented and generally brighter than the unbundled thalamostriatal projections, which have a diffuse, spidery appearance. However, these distinctions are slightly different for projections from each thalamic nucleus, so separate training was required for each brain. For each channel, 3-6 sections were trained and this training was then applied to the remaining ~70 sections containing striatum.

The Trainable Weka Segmentation parameters were as follows; six image filters were selected, the Entropy, Membrane Projections, Neighbors, Structure, and Variance. Classes were homogenized, and the other settings were left on their default values (membrane thickness: 1, membrane patch size: 19, minimum sigma: 1.0, maximum sigma: 16.0, classifier: fast random forest of 200 trees with 2 features per tree). Once training was complete, the classifier was applied to the complete stack of each channel's striatum mask, generating a probability map for each of the 3-levels, which conveys the certainty that a given pixel belongs in each of the 3-levels.

A threshold was manually determined for the diffuse projection probability map and applied to the full stack. This single-level threshold was chosen to encompass the largest possible region of correctly trained diffuse projections throughout the striatum. Any

mistakes not corrected for in the Trainable Weka Segmentation process were then manually edited in MATLAB using custom programs. The final output was a binary projection mask encompassing the full thalamostriatal projection range of each injection.

Striatum registration and alignment

To facilitate comparison across experiments and datasets, each experimental striatum mask was aligned to the striatum of the AIBS average template brain (Kuan et al. 2015). First, each section image was rotated vertically based on manually selected midpoints and down-sampled to 25 μm per pixel. The rotation due to aberrant slicing angle in the dorsal-ventral (D-V) axis was estimated using manually selected landmarks, and the rotation in the medial-lateral (M-L) axis was estimated using the center of mass of each hemisphere. The average template brain was rotated using these estimated angles to mimic the aberrant slicing angle of the experimental brain.

A center of mass curve is then generated from the segmented striatum of the rotated and averaged template brain, and the experimental brain sections are aligned in M-L and D-V axes using this information. In the M-L axis, only the top half of the striatum was used to calculate the center of mass due to variability in the ventral striatum masks. Additionally, the first several sections of the striatum (variable depending on D-V rotation angle) were aligned using the center of mass of the anterior commissure because the range of D-V slicing angles made the shape of these sections inconsistent. Isolated sections with tissue damage were skipped, and the striatum mask, as well as the projection masks, were averaged from the sections before and after the damage.

The experimental striatum was scaled in the anterior-posterior (A-P) axis to fit the rotated average template brain based on the first and last section with the corpus callosum crossing the midline. A linear scaling in the D-V axis was applied based on the average distance from the top of the striatum to the center of mass of the anterior commissure in the front, and the average distance from the top to bottom of the striatum in sections posterior to the anterior commissure crossing the midline. Scaling in the M-L axis was

determined by an average with of the dorsal striatum above the center of mass. The section images are again aligned to the average template brain in the D-V axis using anterior commissure for the first several sections, and the dorsal border of the striatum for posterior sections, and realigned in the M-L axis based on the center of mass of the top half of the striatum.

The experimental brains were then rotated in all axes to align with the original coordinates of the average template brain, and then subjected to one more round of iterative alignment in each axis as described above. Finally, after visual inspection, if manual adjustments to the alignment were necessary, they were fed back to a point just before the average template brain is rotated to the aberrant slicing angle of the experimental brain, and the process is repeated. The corresponding thalamic projection masks were aligned concurrently with the striatum masks.

Cortex data generation overview

The data generation pipeline for the AIBS dataset is analogous, but slightly different, to that used in the thalamocortical dataset (Oh et al. 2014). Briefly, since cortical projections do cross the midline, one stereotaxic, iontophoretic injection of AAV2/1 encoding eGFP was performed per animal (Wang et al. 2014). Both male and female wild-type and Cre expressing (see below) C57BL/6J mice were used. Two weeks post infection, at postnatal day 56±2, the animals were paraformaldehyde fixed and imaged using a TissueCyte 1000 serial two-photon tomography system. By imaging prior to sectioning, this dataset avoids the tissue distortion created in the thalamocortical dataset during sectioning, allowing for full-brain registration across experiments. The images have a z-resolution of 100 μm and a lateral resolution of 0.35 μm/pixel. The background fluorescence from 1231 of these image sets was iteratively registering and aligned to generate an anatomically salient average template mouse brain, to which all of the AIBS Mouse Connectivity Atlas data are aligned (Kuan et al. 2015). Our thalamostriatal projections were also mapped onto this average template brain to facilitate comparisons across datasets (**Fig. 1b**).

Projections were identified in the AIBS image sets using an AIBS custom image segmentation algorithm that identifies fluorescent pixels and produced a full-resolution (0.35 μm/pixels) binary mask of positive pixels (**Fig. 1d**) (Kuan et al. 2015). This data was then binned into 100x100x100 μm (100^3 μm) voxels, where the value in each voxel represents the fraction of the original pixels within that voxel that contained fluorescence. The voxelized data identified fluorescence in the striatum that resulted from bundled axons traveling to subcortical structures as well as corticostriatal axons. To remove the fluorescence resulting from bundled axons, as well as contamination from axons traveling adjacent to the striatum in the corpus callosum, the traveling axons were manually subtracted using custom MATLAB software, and the full resolution image sets as a guide (http://connectivity.brain-map.org).

The volume and density of projections to all brain regions defined in the AMBA ontology was determined for all injections in the AIBS Mouse Brain Connectivity Atlas. Since the striatum does not have accepted subdivisions this data was not useful for the analysis of corticostriatal projections. However, this data was utilized to identify nucleus specific corticothalamic projection densities for the injections analyzed in this paper. This data was downloaded directly from the AIBS API (http://www.brain-map.org/api/index.html)(Oh et al. 2014).

Selection of cortical injections for projection analysis

The AIBS Mouse Brain Connectivity Atlas contains 1029 cerebral cortex injections, which cover all of neocortex, mesocortex, and allocortex (Oh et al. 2014). It is known that olfactory information does not project directly to the striatum so olfactory areas and the piriform cortex were excluded (McGeorge & Faull 1989), leaving 957 injections. This was also checked through a search for olfactory to striatal projections in the AIBS Mouse Brain Connectivity Atlas (data not shown). These 957 injections include both wildtype and cell-type specific cre lines, 177 of these injections are in wildtype C57BL/6J animals. However, many of the wildtype injections had insufficient subregion specificity to map projections in a

subregion specific manner, so three sets of cre lines were also included, A930038C07Rik-Tg1-Cre, Rbp4-Cre_KL100, and Cux2-IRES-Cre. The cre lines were chosen to span cortical layers 2/3 (L2/3) and 5 (L5), so as not to bias the dataset towards IT or PT-type corticostriatal projections (Kress et al. 2013), and contain injections in all of the cortical subregions analyzed. This added another 177 injections to the 177 wildtype injections, totaling 354 to choose from. No cortical layer 4 (L4) or layer 6 (L6) lines were chosen because they do not project to striatum (Briggs 2010). One injection each of Etv1-CreERT2, Gpr26-Cre_KO250, and Grp-Cre_KH288 were used to supplement the lack of a specific L5 or L2/3 injection using the 3 primary cre lines in auditory and insular cortices, respectively. The amygdala and hippocampus were primarily targeted by wildtype injections, but required different cre lines since they have different gene expression patterns from neocortex and mesocortex. The meta data for each injection identifies the primary and secondary brain areas infected, which was used as final screening process for subregion specificity before each brain was manually checked for targeting accuracy. Some small subregions were grouped with functionally similar areas if few or no specific injections could be found. This was necessary for LO/VO, dACC/vACC, Ect/Peri/Tem, and AI/GI/DI. Injections specific to subregions of large areas, such as visual and sensory cortices were selected to insure full coverage of the entire volume, but not analyzed independently. In the end 127 injections were found to specifically target 15 subregions. All areas contain at least one wildtype, one L2/3, and one L5 injection and contain 8 injections on average, with considerable variability depending on the size of the subregion, with the fewest being insular cortex with 3 injections and most being sensory cortex with 21 injections (see **Table 1**).

For hippocampal areas, while some injections included in this dataset had CA1 or CA3 as a primary target, only injections that at least partially covered the subiculum sent projections to the striatum (data not shown). For amygdalar areas, the primary volumes of the amygdala injections in this dataset are in the BLA, and BMA, but they also cover parts of the CEA, PA, MEA, and PAA, spanning both pallial and subpallial parts of the amygdaloid complex (see **Table 1**).

Subregions of the isocortex, hippocampus, and amygdala can all be broadly defined as telencephalic cortical areas, and separated into neocortex, mesocortex and allocortex. Neocortex is primarily six-layered and comprised of the primary sensory and motor cortices. Mesocortex, also called the paralimbic cortex, is generally three-layered and made up of associative subregions in frontal cortex as well as subregions at the interface between allocortex and neocortex, such as insular and perirhinal cortices. Allocortex is the evolutionarily oldest part of cortex, and comprised of piriform cortex, hippocampus and the subiculum (McGeorge & Faull 1989). Although the amygdaloid complex has both telencephalic (pallial) and subpallial origins, it is situated within allocortex, between piriform cortex and the subiculum (Pabba 2013). Being functionally related to both the hippocampus through the limbic system and the piriform cortex with olfactory processing (Novejarque et al. 2011), it will be grouped here with allocortex.

Confidence map generation

Confidence maps, which define the thalamic origin of projections, were created to determine which regions of the thalamus sent projections to: (1) striatal volumes with subregion specific corticostriatal projections (**Fig. 4c**), (2) striatal volumes with dense projections from cortical subtypes (**Fig. 5f**), (3) striatal volumes with high or low diffuse cortical projection convergence (**Fig. 5g**), and (4) striatal subdivision created by clustering voxels with common cortical input patterns (**Fig. 7a**). To control for alignment variability (~100 μm) across thalamus masks (Hunnicutt et al. 2014), an injection "core" was produced by eroding the injection "shell" of each three-dimensional injection mask by 100 μm. For each injection, a positive core adds 2 to the confidence map and a positive shell adds 1. Similarly, negative injection cores subtract 2, and a negative shell subtracts 1 (Exception: shells are only subtracted for the two easiest to meet criteria in each grouping method, and subsequent criteria only subtract negative cores as 1). (see **Figure 4a** for a simplified schematic of this process). For case (1) above, a six level confidence map was created by determining the inclusion of each injection in the following three groups; 10% of the diffuse

target volume covered by the projection, 5% of the dense target volume covered by the projection, and 50% of the dense target volume covered by the projection. Thalamic volumes occupied by the cores of injections that did not meet any of these criteria were set to zero. For cases(2),(3), and (4) there was not projection density data, but instead only binary volumes, so the injection grouping had to be adapted. For these groups 8 level confidence maps were created by determining the inclusion of each injection in the following four groups; 10% of the target volume covered by the projection, 10% of the projection volume within the target, 25% of the target volume covered by the projection, and 25% of the projection volume within the target. Thalamic volumes occupied by the cores of injections that did not meet any of these criteria were set to zero.

The nuclear coverage of the confidence maps in all groups was calculated for both the aligned PMBA and AMBA, and then an average was taken between them (Hunnicutt et al. 2014). For case (1) the mean nuclear coverage was determined was by averaging 6 coverage values, confidence levels 1, 3, and 5 in both atlases (**Fig. 4d**). For cases(2),(3), and (4) the mean nuclear coverage was determined was by averaging confidence levels 1, 4, and 7 in both atlases (**Fig. 5f-g** and **Fig. 7b**). All summary data is shown as an average across atlases (**Fig. 4b, d-e, 5f-g, 7b, 8c-d**).

Overall, the confidence maps incorporate information about the breadth and specificity of projections, as well as the variability in thalamus alignment. The confidence map therefore represents the likelihood of a thalamic voxel projecting to a particular target.

Voxel clustering for striatal segmentation

Each voxel was assigned a point in a 15-dimensional space corresponding to the density of projections from each cortical subregion. The optimum distance metric was determined by comparing the cophenetic correlation coefficient across methods, and Spearman's rank correlation metric was selected with a cophenetic coefficient of 0.78. This distance metric and an average linkage was used to perform cluster analysis on the striatal voxels. The maximum number of voxel clusters was determined by manually thresholding the resulting

dendrogram. The projection regions were similarly assigned a point in 25-dimensional space corresponding to the 25 nuclei, and clustered using the same method.

Results

Localizing the striatal projection fields for anterograde thalamic injections

To map projections from the entire thalamus to the striatum, we utilized 218 anterograde tracer injections generated previously (see **Methods** for details)(Hunnicutt et al. 2014). Since thalamic projections do not cross the midline in mouse, up to 4 anterograde viral tracer injections could be analyzed independently in each animal, one red and one green in each hemisphere (Hoover & Vertes 2007) (**Fig. 1a**). High-resolution, full-brain image sets were acquired for each experiment, which were previously used to localize the injections within the thalamus (Hunnicutt et al. 2014). The thalamic injections are ~630 μm in diameter in the medial-lateral axis, and they tile the thalamus so that each pixel is sampled by 4 injections on average (Hunnicutt et al. 2014). This high sampling density allows us to localize the origin of thalamic projections to a volume smaller than any individual injection, which was necessary since most thalamic nuclei are smaller in at least one dimension than the mean injection diameter. Using the full-brain image sets, the striatal projection distribution was then localized for each injection.

Thalamic axons on their way to cortex travel through the striatum in axon bundles. Since passing thalamocortical axons are oftentimes intermingled with thalamostriatal projections, automated localization of thalamostriatal projections is difficult. To localize striatal projection distributions in a way that distinguished between thalamocortical and thalamostriatal axons, we employed a semi-automated, machine-learning image segmentation plugin for ImageJ, called Trainable WEKA Segmentation (http://fiji.sc/Trainable_Weka_Segmentation). The segmentation program was manually trained on a subset of striatum images to distinguish between the highly directional axon bundles and the diffuse axon terminal. The program then used this information to identify diffuse projections in the remaining ~70 images containing striatum. The output of the WEKA image segmentation was manually checked to ensure accurate projection localization, and the final result was a binary mask of the distribution of thalamostriatal projections for each injection (**Fig. 1b** and **Methods** for details).

To compare striatal projection distributions across both animals and datasets, each striatum was aligned to an average template brain created by the AIBS as a common coordinate framework for the C57BL/6J mouse brain (**Fig. 1b**) (Kuan et al. 2015). The striatum was manually traced in each image. Using the shape of the striatum, as well as fiducial landmarks, each section image was rotated and scaled to align with the average template brain (see **Methods**). The projection masks were aligned using the same transformations, resulting in the localization of the striatal projection distributions from each of the 218 thalamic injections within the common coordinate system of the average template brain.

Localizing the striatal projection fields for anterograde cortical injections

The remaining excitatory inputs to the striatum come from all neocortical and mesocortical subregions, the hippocampal formation and the amygdaloid complex of allocortex (**Fig. 1e**). In contrast to the relatively small nuclei of the thalamus, these subregions are large enough to be labeled specifically by single viral injections, and could therefor be mapped using a smaller number of targeted injections. To do this, we analyzed the striatal projection distributions for a subset of the 1029 cerebral cortical injections included in the AIBS Mouse Brain Connectivity Atlas(Oh et al. 2014), (http://connectivity.brain-map.org/). To get a sufficient number of subregion specific injections, both Cre and wildtype lines were included in our dataset (see **Methods** and **Table 1**). All subregions contain at least one wildtype, one L2/3 specific, and one L5 specific injection and were targeted by 8 injections on average. A total of 127 anterograde viral injections targeting 15 subregions were used (**Fig. 1e**).

While the precise details of the thalamic and cortical dataset generation are different, the general data pipeline and output are very similar (see **Methods** for details). Like the thalamic data, the cortical dataset consists of injections of AAV2/1 driving the anterograde expression of eGFP. After two weeks of viral expression, the brains are fixed and full-brain, high-resolution image sets are generated, from which fluorescent projections are localized

and converted to a binary mask (Oh et al. 2014). The full-resolution projection data was down-sampled to 100x100x100 μm voxels, where the value in each voxel represents the fraction of the original pixels within that voxel that contained fluorescence. This voxelized projection information is the dataset used in the present study.

Since the AIBS image segmentation identifies all fluorescence, there is no distinction between striatal fluorescence originating from traveling axon bundles and diffuse axon terminals. To make sure the analyzed data only contained corticostriatal projections, the voxelized data was compared to the original high-resolution image sets and manually corrected to remove fluorescence resulting from traveling axon bundles. (see **Methods**).

In both primates and rodents, corticostriatal projections have two distinct innervation patterns, a core projection of densely packed terminals that constitute the canonical projections for a given cortical area, i.e. sensory cortex projections to dorsolateral striatum, and a diffuse projection field that spans a less defined volume (Haber 2006; Mailly et al. 2013). To investigate the distributions of both of these projection types, 3 levels of projection densities were mapped for all cortical injections. The most inclusive level, a threshold of 0.5% fluorescence per voxel, was chosen to encompass all diffusely projecting axon terminals, the least inclusive threshold of 20% selects only the dense core of the projection fields, and an arbitrary middle threshold of 5% selects an intermediate volume between dense and diffuse (**Fig. 1d**).

Corticostriatal input distribution patterning

To examine the complete distribution of subregion specific projections from the cortex to the striatum, the projections distributions were summed for all injections within each area (**Fig. 2**). This process was performed separately for the dense, moderate, and diffuse projection densities, allowing for the visual comparison of core dense projections and the distributed diffuse projections for each input type (**Fig. 2a**). The projection distributions, while similar for some areas, were never identical across any two input types, and formed either one or two continuous projection volumes in all cases. As expected, the dense

projection distributions follow the projection patterns described previously for each striatal input (McGeorge & Faull 1989; Berendse et al. 1992; Groenewegen et al. 1987). The distribution of each projection volume in the dorsal-ventral (D-V), medial-lateral (M-L), and anterior-posterior (A-P) axes are quantified in **Figure 2b** as the fraction of the striatum containing either dense or diffuse projections in each axis. Interestingly, subregions that targeted two distinct volumes of the striatum, such as IL and Sub, always segregate between the anterior and posterior striatum, and largely avoided any of the central volume (**Fig. 2b**). For the sensorimotor subregions FrA, M1/2, and S1/2, both the dense and diffuse projections target a central volume in the D-V and A-P axes but are all biased laterally. In contrast, LO/VO has very defined dense projection in the anterior, medial striatum, but covers almost the entire striatal volume with diffuse projections, potentially allowing LO/VO to influence activity throughout the striatum (**Fig. 2b**).

To explore whether the defined projection patterns seen in the subregion specific injection groupings were due to subregion specificity, or to topographic proximity of the injections in each group, we grouped injections according to their distributions in the A-P and M-L axes (**Fig. 3a, d**). The striatal projections from the most medial cortical group in the M-L axis have a strong anterior and medial projection bias (**Fig. 3a-c**, dark green). The distribution of dense striatal projections moves generally posterior and lateral as the cortical groups become more lateral, but the groups are significantly more heterogeneous in their projection distributions than the cortical subregions, covering most of the striatal volume with diffuse projections in all but the most medial cortical group (**Fig. 3c**). The cortical groups in the A-P axis have a similar specificity to the M-L groups, where one group has a spatially defined projection distribution (**Fig. 3e**, light orange), but the rest of the groups have only a general trend in their dense projection distributions moving anterior to posterior as the cortical groups move posterior and project diffusely to the entire striatal volume (**Fig. 3d-f**). Overall, the injection groupings resulted in projection distributions that were both more regionally specific and distinct from other groups when compared to projections from A-P or M-L groupings.

Convergence of thalamostriatal inputs with sub-region specific corticostriatal inputs

One putative role of the striatum in basal ganglia processing is in the integration of information from multiple modalities to form a unified output (Steiner & Tseng 2010; Seger & Miller 2010). Cortical inputs are generally thought of as the drivers of striatal activity, and the commonly accepted subdivisions of striatum, i.e. sensorimotor, limbic, and associative, are defined by their most predominant cortical input source (Draganski et al. 2008). However, the thalamus is the second largest source of striatal input, contributing nearly one third of all excitatory synapses (Huerta-Ocampo et al. 2013). Despite this significant contribution, the thalamus if often ignored as a driver of striatal activity.

To understand the full range of cortical and thalamic input integration in the striatum, we identified all thalamic volumes whose projections converge in the striatum with projections from each cortical subregion (**Fig. 4**). Since many thalamic nuclei are smaller than the average injection diameter, we used a grouping method that allowed us to localize thalamic volumes smaller than any given injection (**Fig. 4a**). Injections that converge in the striatum with a given cortical subregion are summed, while the injections that do not converge are subtracted (see **Methods**). To account for the alignment variability across experiments, the injection sites were eroded by 100 μm to produce an injection "core" (Hunnicutt et al. 2014). Other factors, such as the extent of projection overlap and whether convergence occurs with dense or diffuse cortical projections, are also taken into account, and the result is a confidence map of the entire thalamic volume where the white identifies the origin of thalamic projections that we are most confident is accurately localized within the thalamus (see **Methods** for details). The result provides a complete map of the origins of thalamic projections convergent with subregion specific cortical projections (**Fig. 4c**).

The thalamus is made up of over 25 distinct nuclei which provide the striatum with information on everything from incoming sensory stimuli to attentional states (Sherman & Guillery 2001; Haber & McFarland 2001; Y. Smith, Galvan, et al. 2014a). Knowledge of

which nuclei contribute projections in our thalamic confidence maps is important for understanding the functional consequences of each convergent projection. To gain insight into this, we determined which thalamic nuclei encompass the thalamostriatal projection origins (**Fig. 4b**). The thalamic injections were originally aligned to two atlases, the Paxinos Mouse Brain Atlas (PMBA) and the AIBS Mouse Brain Atlas (AMBA) (Hunnicutt et al. 2014). We overlaid these atlases onto the thalamostriatal confidence maps, calculated the fraction of each nucleus covered by each confidence level, and averaged those values across the two atlases to account for differences between them (**Fig. 4b**). The fractions of each nucleus covered by projections convergent with each cortical subregion are summarized for confidence levels 1, 3, 5, and their average in **Figure 4d**.

There are several thalamic nuclei that do not project to any striatal volumes, VPM, VPL, RT, and AV, which is consistent with previous literature. RT is a thalamic feedback nucleus(Kolmac & Mitrofanis 1997), and AV sends strong projections to RSP and Sub, but has not been shown to synapse in the striatum (van Groen & Wyss 1995; Pan et al. 2010). VPM and VPL are the primary somatosensory relay nuclei, and have been shown to send axons through the striatum, but do not synapse within the striatal volume (Zhang & Deschênes 1998; Pan et al. 2010; J. B. Smith et al. 2012). These results validate our image segmentation methods, which excluded fluorescence resulting from traveling axons (see **Methods**).

Corticostriatal input convergence patterns

Given the obvious abundance of convergent cortical and thalamic projections throughout the striatum (**Fig. 2, 4**), we wanted to determine if convergence was ubiquitous across the striatum, or if there were regions with more access to information integration than others. To do this, the bilateral distributions of dense and diffuse projections were summed for all cortical subregions. The resulting heat maps show that the distribution of input convergence is not uniform across the striatum, and it is dramatically different when you look at the diffuse projections versus the dense projections (**Fig. 5a**). For the diffuse

projections, nearly all pixels have 5 or more convergent inputs and the convergence value per pixel rises steeply around 9 (**Fig. 5a**, dashed black line). If this value is used to divide the striatum, ~37% of ipsilateral and 62% of bilateral pixels have 8 or less convergent inputs, with an average bilateral convergence of 6.6 ± 0.84 (mean ± standard deviation) inputs, leaving 63% of ipsilateral and 38% of bilateral pixels with 9 or more convergent inputs, and an average bilateral convergence of 10.7 ± 1.1 inputs. Conversely, for dense projections, 90% of bilateral pixels have 5 or fewer convergent inputs, with a maximum value of 7, and an average convergence of 3.4 ± 0.4 inputs ipsilaterally and 2.7 ± 0.4 inputs bilaterally.

The dense projections are not only less likely to converge, which is expected given that they each cover a smaller striatal volume than the diffuse projections, their distribution of input convergence is very different than that of the diffuse projections. The diffuse projections converge primarily in the dorsal striatum, whereas dense projection convergence is biased towards the ventral striatum (**Fig. 5b**). These areas of high and low input convergence in the striatum could constitute functionally distinct subdivisions, and would be good targets to compare forms of input integration in the striatum.

To determine which cortical subregions converge with one another, the fraction of each striatal projection field convergent with all other cortical subregions was calculated (**Fig. 5c**). As expected, the diffuse projections have a much higher proportion of convergence than the dense projections. However, there are areas that have universally very little crosstalk, such as Ptl, RS, IL and Sub with the motor areas M1/2 or FrA. There are also subregions of consistently high input convergence, such as the motor (FrA and M1/2) and sensory (S1/2 and AI/GI/DI, which includes areas processing gustatory sensory information) subregions, that have a high proportion of convergence for both their diffuse and dense projections (**Fig. 5c**).

To determine if either the high or low convergence areas identified in **Figure 5a** could be attributed to a type of cortical input, the projection distributions of cortical subtypes were examined (**Fig. 5d-e**). Neocortical subregions are entirely sensory and motor, mesocortical subregions are broadly considered associative, and the allocortical subregions

examined here carry limbic information from hippocampal and amygdalar areas (**Fig.1e**). No single cortical type accounted for the full distribution of high or low input convergence for either dense or diffuse projections. The dense projection distributions for cortical subtypes were consistent with previous results (McGeorge & Faull 1989), and the diffuse projections of each group contacted almost 100% of the ipsilateral striatum (**Fig. 5e**).

We then examined thalamic convergence with the ipsilateral striatal volumes containing high/low cortical convergence, and projections from cortical subtypes (**Fig. 5f-g**). The target data, i.e. the projection distributions in the striatum, for these datasets are fundamentally different than those for the cortical projection fields localized in **Figure 4**. Instead of having graded projection densities, i.e. dense, moderate, and diffuse, there is only one projection level, present or absent. This fact required the thalamic injection grouping method to be adapted. It still uses the eroded core to account for alignment variability, but also takes into account the fraction of the target field covered and the fraction of each projection within the target field (see **Methods** for details).

To highlight the differences across the largely complementary allocortical and neocortical projection fields, thalamic convergence with only the non-overlapping dense projections are shown here, but the thalamic convergence with mesocortical projections are shown for the full distribution of dense projections (**Fig. 5f**). The most striking differences are the lack of convergent neocortical projections for the midline nuclei, PT, PVT, and RE, as well as very low confidence for projections from the anterior nuclei, AM, AD, IAD, and IAM, yet strong convergence for these areas by allocortical projections. Conversely, the neocortical projections receive convergent input from the ventral nuclei, VM and VAL, which do not contact allocortical projections at all (**Fig. 5f**).

The volume of the striatum with low (<9) convergent diffuse cortical projection (**Fig. 5a**), did not receive any input from the anterior nuclei. Interestingly, although it only accounts for 63% of the ipsilateral striatal volume, the striatal area with high diffuse cortical projection convergence receives inputs from every thalamic nucleus that projects to the striatum (**Fig. 5g**).

Subdividing the striatum based on common corticostriatal input patterns

We then sought to subdivide the striatum in an objective and unbiased manner using the distributions of all corticostriatal projections. To do this, the striatum was down-sampled to 150x150x150 μm voxels, and the density (diffuse, moderate, or dense) of projections within each voxel was determined for all cortical subregions (**Fig. 6a**). The voxels were then clustered (agglomerative hierarchical clustering, MATLAB) based on their input density from each of the 15 cortical subregions (**Fig. 6b**). Although the clustering contains no spatial information, the clustered voxels could be mapped back onto the striatum to form largely continuous volumes. The dendrogram threshold was gradually decreased until the consistency across linkage points was too high to create coherent subdivisions (**Fig. 6c**). This gradual increase in the maximum number of striatal clusters allowed for the examination of input relationships throughout the striatum.

When the striatum is divided into two clusters, a relatively small dorsomedial subdivision and a large region that encompasses both the ventral and the dorsolateral striatum are formed. This division highlights the relatively specific and convergent inputs from Ptl, RS, Aud, and Vis to the dorsomedial striatum (**Fig. 6b-d**, red). When the threshold is lowered slightly more, the 3 clusters produced are highly reminiscent of the functional subdivisions that the striatum is generally broken into, the ventral limbic region, the dorsomedial associative region, and dorsolateral sensorimotor region (Gruber & McDonald 2012), which constitutes the first three dimensional definition of the precise boundaries of these subdivisions. Interestingly, at this level the ventral striatum is also clustered with the most posterior striatum, and this region is separated into its own cluster as the threshold is lowered to produce four clusters. The boundaries of these four clusters are carried across the voxel image to highlight their input distributions (**Fig 6b-c**). The posterior 4th cluster has an input distribution that falls somewhere between the dorsomedial and ventral clusters (**Fig. 6b-d**, green).

As the threshold was gradually lowered to 14 clusters, all but one new cluster broke down the dorsomedial striatum into ever smaller subdivisions, further highlighting the input heterogeneity in this area (**Fig. 6c-d**, pink and red-yellow). The new cluster separated the posterior group into motor and non-motor subdivisions (**Fig. 6c-d**, dark-light green). Finally, at the 15th cluster, the ventral striatum was separated into two large subdivisions, one with and one without primary sensory and motor inputs (**Fig. 6c-d**, mid-dark blue).

The contributions of thalamic inputs to each of these cortically defined clusters are shown in **Figure 7**. Since the cluster distributions are binary values, the same injection grouping method used for cortical subtypes was used here (**Fig. 5f**, and **see Methods**).

Organization of the thalamus in cortico-thalamo-basal ganglia loops

The basal ganglia receive excitatory input from the cortex and thalamus via the striatum. This information is processed and eventually passed to the basal ganglia output nuclei, which transmit it to the thalamus. The thalamus will either pass the information to the cortex or act as a feedback to the striatum, thus completing the cortico-thalamo-basal ganglia loop (**Fig. 8a**). To get a comprehensive picture of the role of the thalamus in this circuit, the complete distributions of thalamostriatal, thalamocortical, and corticothalamic projections were directly compared to one another (**Fig. 8b-c**). First, the thalamostriatal projection origins described in **Figure 4** were overlaid with subregion specific thalamocortical projection origins previously published using this dataset (Hunnicutt et al. 2014). This analysis identified three thalamic projection types. Those that projected only to a given cortical subregion without any corticostriatal convergence (cyan), thalamostriatal projections that converge with corticostriatal inputs from areas they do not project to themselves (magenta), and thalamostriatal projections that also sent feed forward thalamocortical projections to the same subregions that they converge with in the striatum (**Fig. 8b**, white). The average nuclear localization of these distinct and feed forward projections are summarized in **Figure 8c**. For all 127 cortical injections, nucleus specific corticothalamic projection densities were accessed directly from the AIBS Mouse

Connectivity Atlas API (see **Methods**). This data is displayed alongside the thalamostriatal and thalamocortical nuclear coverage data as the maximum projection density in each nucleus (**Fig. 8c**, green).

The thalamocortical, thalamostriatal, and corticothalamic data were then rearranged to investigate the relationship between the targets of the thalamic basal ganglia output, MD, Pf, VAL, and VM, and specific cortical subregions (**Fig. 8d**). This direct comparison highlights the interconnectedness of all of the targets of basal ganglia out with motor areas, particularly M1/2 which receives thalamocortical inputs from all of the output nuclei, converges in the striatum with projections from all of the output nuclei, and sends corticothalamic projections to all of the output nuclei (**Fig. 8d**, white and green). This is in contrast to S1/2, which is interconnected with Pf, VM, and VAL via thalamocortical and thalamostriatal projections, but only converges with thalamostriatal MD and does not send or receive MD projections directly. Finally, Pf and MD have a higher proportion of thalamostriatal projections that are not convergent with their direct cortical targets (magenta), than VAL and VM, which tend to have a fully interconnected organization (white). Together, these data provide a comprehensive picture of cortical and thalamic input integration in cortico-thalamo-basal ganglia loops.

Discussion

The preponderance of input integration in the striatum makes understanding the precise striatal convergence patterns crucial before any system of information processing in the basal ganglia can be elucidated. Here, we generated a comprehensive, three-dimensional map of excitatory input convergence in the mouse striatum. This map provides the first objective way to identify functionally relevant subregions in the cellularly homogenous striatum. In the future, this dataset will serve as a guide for functional studies of striatal input physiology, as well as for modeling the mechanisms of information integration in the basal ganglia.

Utilizing comprehensive cortical input convergence data, segmentation of the striatum could be performed in an unbiased, data-driven manner, and allowed for the anatomical significance of a variable number of clustered subdivisions to be assessed. Since the limbic, associative, and sensorimotor domains were originally defined by their unique inputs (Parent & Hazrati 1995), it is not surprising that our analyses isolated clusters reminiscent of these domains. However, the three dimensional distribution of these areas has never before been described. When the data is used to divide the striatum into four clusters, the newly created posterior subregion has interesting input distribution characteristics. It's coverage by auditory and visual projections makes it more similar to the dorsomedial striatum, while the lack of parietal and retrosplenial inputs gives it ventral striatal characteristics. Perhaps most interestingly, this region receives less input from the lateral and ventral orbital cortices than any other clustered portion of the striatum (**Fig. 6c**). Since the input distribution of this area is similar to both the dorsomedial and ventral clusters, it could function as a transitional region between them. Alternatively, it could constitute a 4th, previously unidentified, functional subdivision of the striatum with unique behavioral roles (**Fig 6c-d**, green). When the striatum was segmented into 15 clusters the ventral striatum was separated into two subdivisions (**Fig. 6c-d**, mid-dark blue). The anatomical reason appears to be the presence of motor and sensory projections in the more lateral subdivision. This is interesting for several reasons, the first being that the ventral striatum is generally thought to lack sensorimotor input (Gruber & McDonald 2012; Draganski et al. 2008). When the inputs

from primary sensory and motor cortices to this cluster are examined closer, it is notable that they predominantly contain moderate or diffuse projections, instead of the dense projections for which the canonical projection distributions are described, which may explain this omission in the literature. This putative subregion defines an area of the striatum fully contacted by both limbic and motor cortices.

Upon first inspection, some aspects of this dataset may be misleading. First, the thalamic dataset does not include projection density information for any of the analyses. Therefore none of the data presented reflects the strength of thalamostriatal projections. Instead, the brighter values in the confidence maps represent the certainty with which the origin of a projection is localized (e.g. **Fig. 4c**), and the larger values in the nucleus specific data represent the fraction of the nucleus from which striatal projections originated (e.g. **Fig. 4d**). It is entirely possible, that small portions of nuclei produce widespread, dense thalamostriatal projections. For instance, the thalamic nucleus MD is known to send targeted projections from fractions of its total volume (Groenewegen 1988). This likely accounts for the strong coverage of only medial MD by projections convergent with the subiculum, which both target the ventral striatum (**Fig. 4b**).

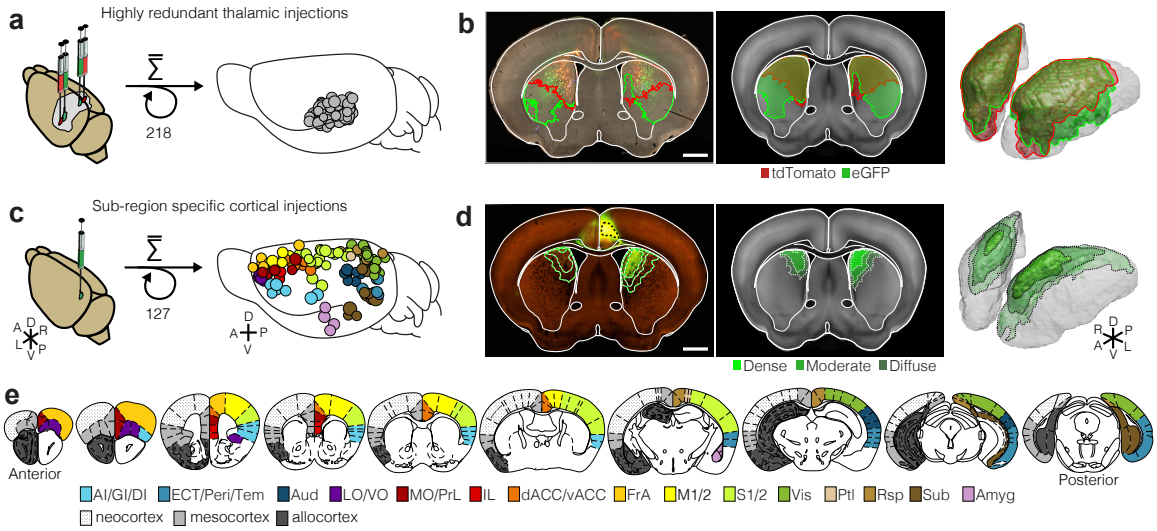
Additionally, while the alignment variability for the thalamostriatal data is ~100 μm, this number does not reflect the certainty in the nuclear localizations. The shape of nuclei can change dramatically in the A-P axis, and be drastically different across atlases (e.g. VAL). The thalamostriatal dataset is aligned to both the AMBA and PMBA and the average coverage value across them is shown for all data, so the alignment of nucleus specific data may vary across both nuclei and experiments. There are also nuclei that are as small or smaller in a given dimension than the alignment variability (e.g. CL), which makes the fraction covered easily altered be a single experiment. This caveat also applies to the corticothalamic dataset obtained from the AIBS Mouse Connectivity Atlas (**Fig. 8c-d**). These nuclear localization issues can be circumvented for thalamostriatal projections by examining the confidence map data.

Through the integration of two large-scale anatomical datasets (Oh et al. 2014;

Hunnicutt et al. 2014), the organization of the cortico-thalamo-basal ganglia circuit has been described with a breadth that has never before existed, and could not have been accomplished by either dataset independently. This effort speaks to both the power and necessity of data sharing and data mining. Taken together, the data presented here describe the complete distribution of excitatory inputs to the striatum, delineate boundaries for the three known functional domains within the striatum, identify novel areas which may constitute functional subdivisions, and detail the complete organization of the thalamus in the cortico-thalamo-basal ganglia circuit.

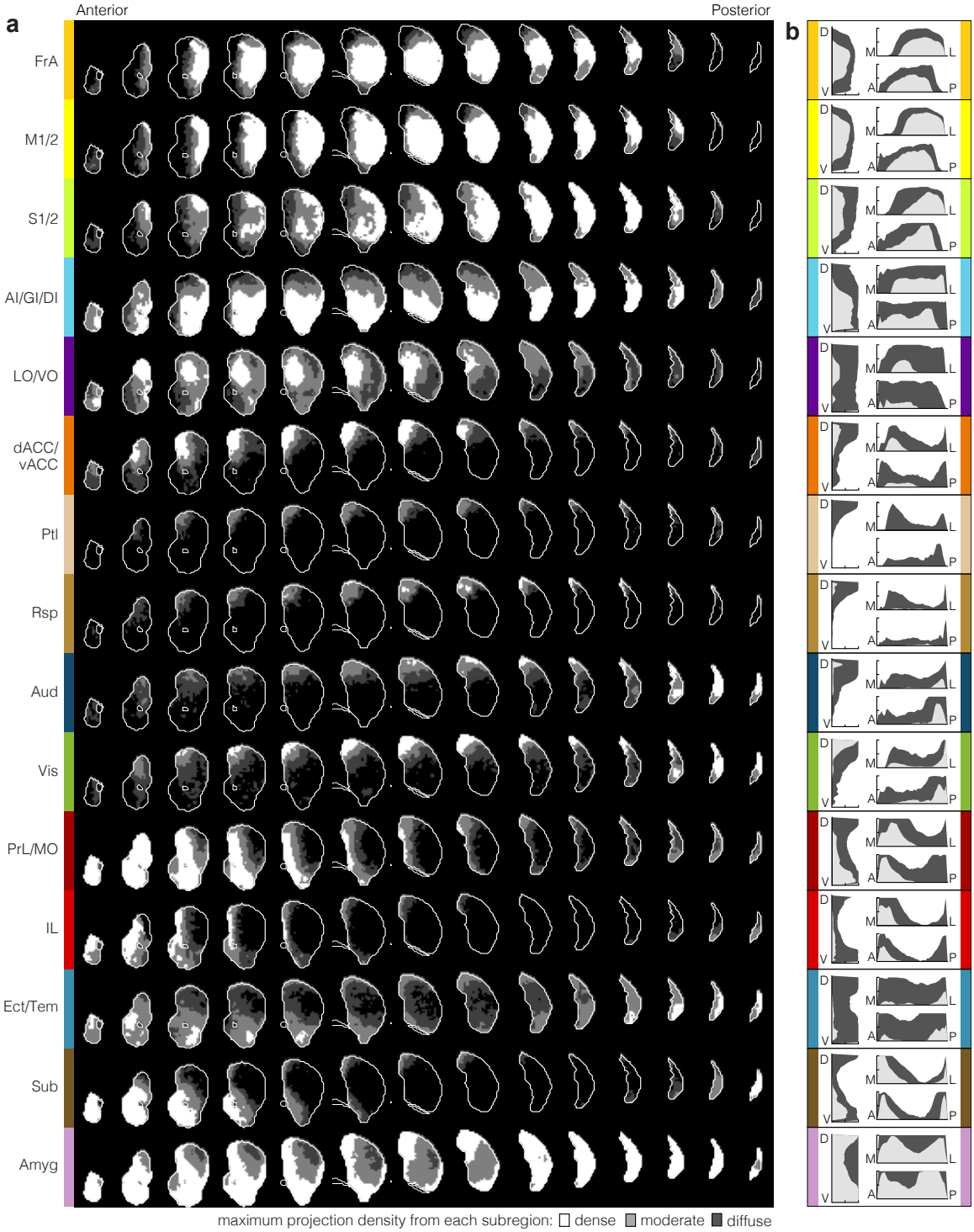
Figures

Figure 1. Integrating large-scale anatomical datasets to investigate whole-brain striatal input convergence.



(**a**) Illustration of thalamostriatal data generation. Bilateral injections of virus driving the expression of tdTomato (red) and eGFP (green) in the mouse thalamus (left), followed by sectioning (50 μm per section) and high-resolution imaging. Each injection is then localized and aligned within an average model thalamus (right). (**b**) Representative coronal section showing thalamostriatal projections (left) and the segmented projection masks (red and green areas) aligned to the AIBS average template brain (center), with 3D view (right) (**c**) Illustration of corticostriatal data generation. Unilateral injection of virus driving the expression eGFP (green) in the mouse cortex (left), followed by 2-photon imaging and then serial sectioning (100 μm per section). Each injection is then localized and the full brain is aligned to the AIBS average template brain (right). (**d**) Representative coronal section showing a cortical injection and its striatal projections (left), and the corresponding segmented projection mask (center) with three projection density thresholds (green lines) and the segmented injection site (dashed black line), with 3D view (right). (**e**) Coronal sections showing the 15 subregions targeted by cortical injections (right of each section) and the cortical subtypes they occupy (left of each section, modified from the PMBA (Paxinos 2004)). (**f**) Collapsed sagittal, coronal, and horizontal view showing the widest points of the average template brain (light gray) and striatum (dark gray) in each axis with anatomical directions used throughout the paper shown for reference.

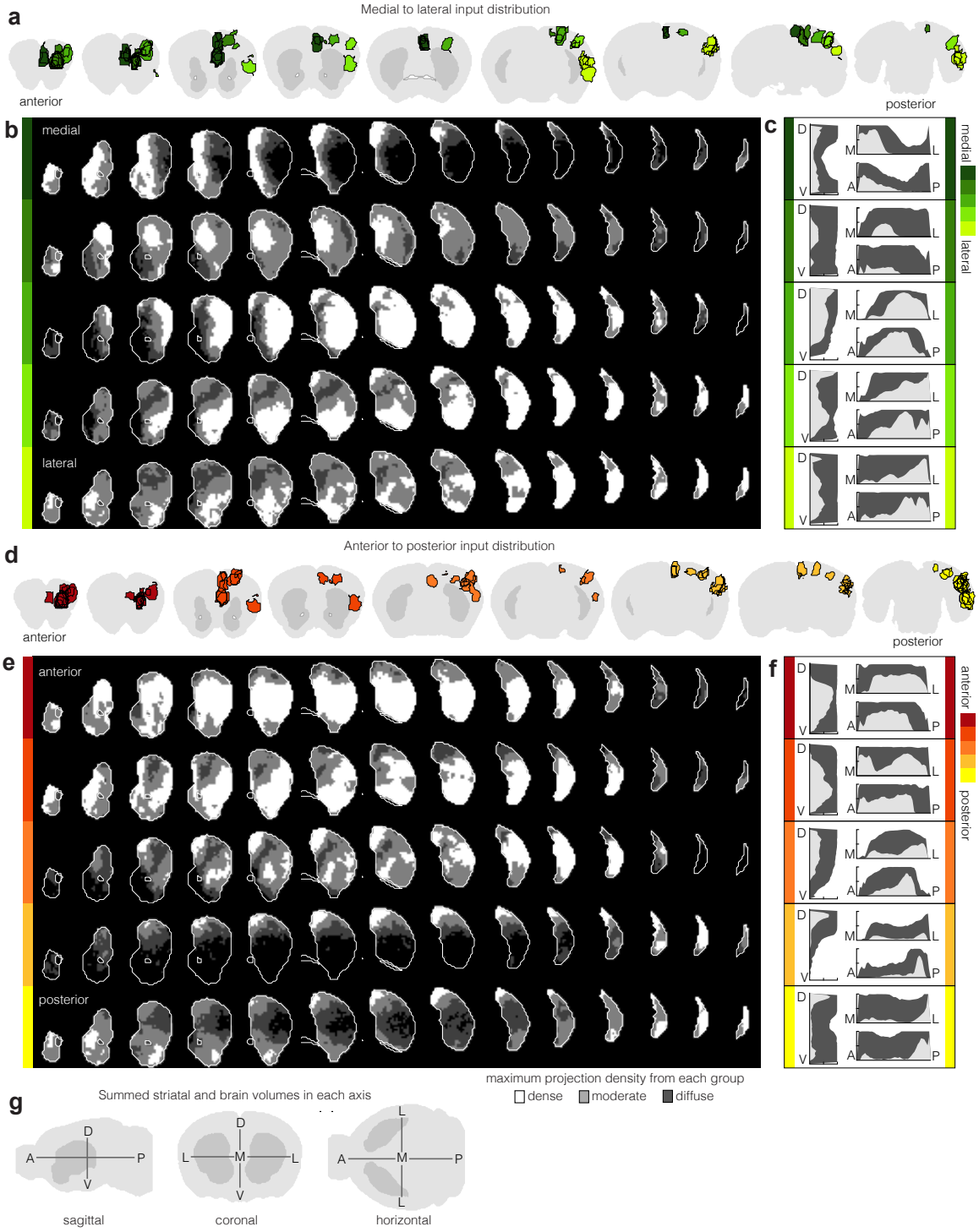
Figure 2. Comprehensive mapping of cortical inputs to the striatum.



(**a**) Coronal sections through the ipsilateral striatum (starting 1.6 mm anterior to bregma and continuing posterior in 300 μm steps) showing the striatal projection distributions for all cortical subregions (rows). The maximum projection densities (dense, moderate, or diffuse) are indicated for the sum of all injections within each cortical subregion. (**b**) Projection distribution plots in the dorsal-ventral (D-V), medial-lateral (M-L), and anterior-posterior (A-

P) axes for each cortical subregion shown in **a**. Coverage of the striatum by dense (light gray) and diffuse (dark gray) projections were calculated at 100μm intervals in each axis as the normalized fraction of the striatum covered at each point.

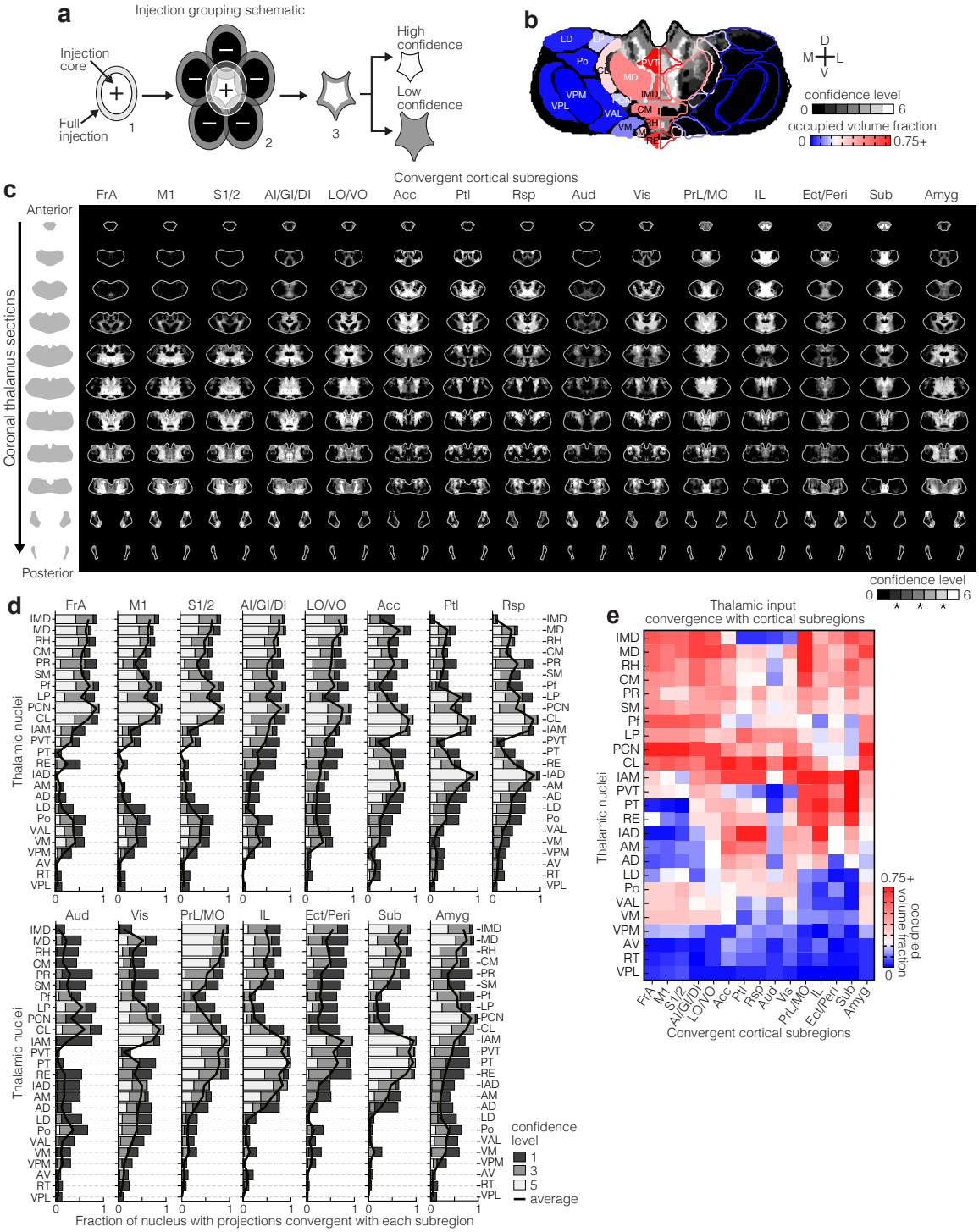
Figure 3. Anterior-posterior and medial-lateral organization of corticostriatal inputs.



(**a**) Coronal sections through the average template brain showing the relative location of injections in each of five medial to lateral (M-L) groups (dark to light green). (**b**) Coronal sections (starting 1.6 mm anterior to bregma and continuing posterior in 300 μm steps) through the ipsilateral striatum showing the striatal projection distributions for the M-L injection groups shown in panel **a** (rows). (**c**) Distribution plots for the dense (light gray) and

diffuse (dark gray) projections of each M-L cortical group shown in panel **b** (calculated as described in **Figure 2**). (**d**) Coronal sections through the average template brain showing the relative location of injections in each of five anterior to posterior (A-P) groups (red to yellow). (**e**) Coronal sections (at 300 μm intervals) from anterior to posterior through the ipsilateral striatum showing the striatal projection distributions for the A-P injection groups shown in panel **d** (rows). (**f**) Distribution plots for the dense (light gray) and diffuse (dark gray) projections of each A-P cortical group shown in panel **e** (calculated as described in **Figure 2**). (**g**) Example sagittal, coronal and horizontal sections through the mouse brain showing the location of the striatum in each dimension.

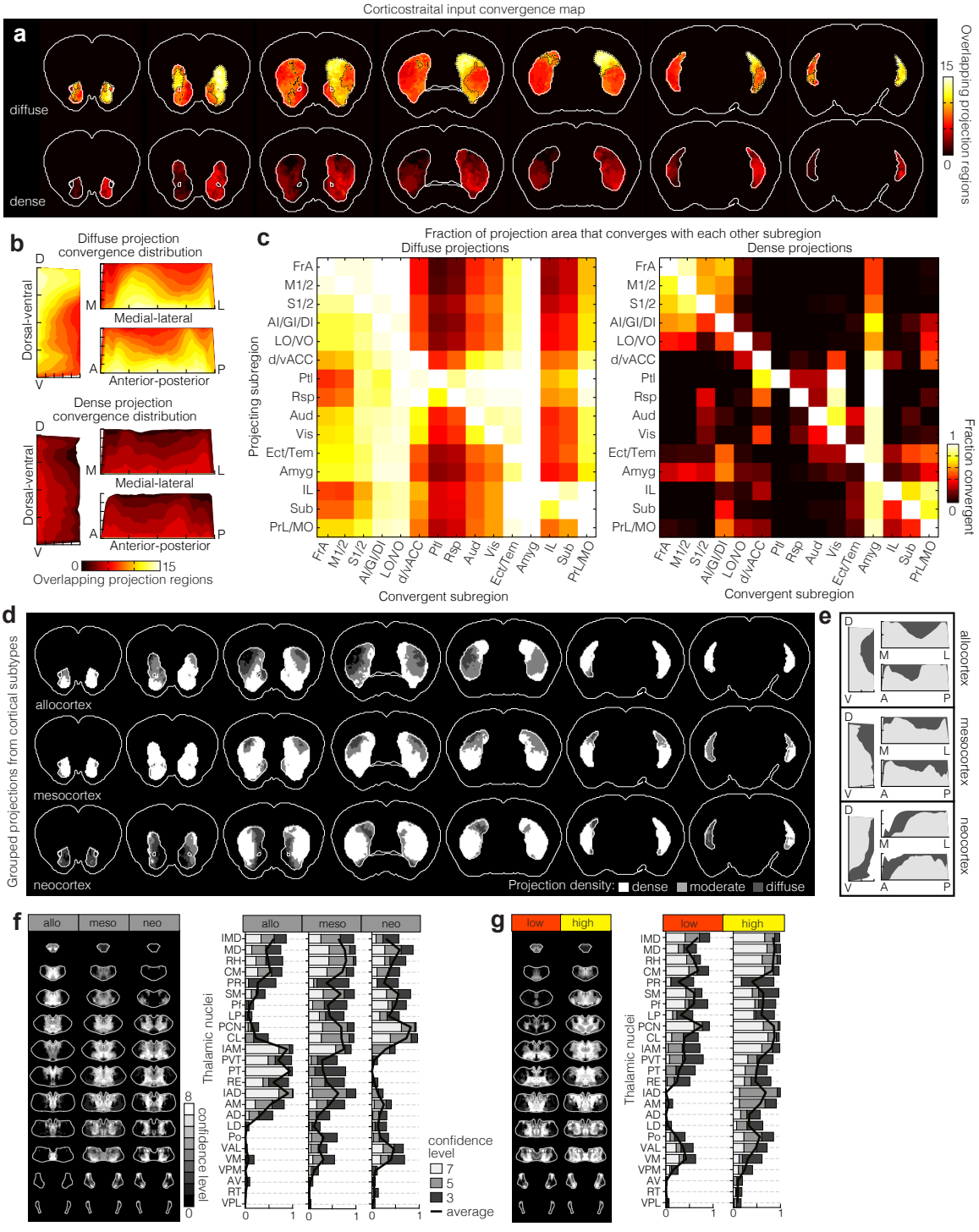
Figure 4. Thalamostriatal projections that converge with subregion specific corticostriatal projections.



(**a**) Simplified schematic of thalamic localization method. (1) The volume of each injection is eroded to generate an injection core. (2) Injections that send convergent projections to the striatum are summed and those that do not (negative injections) are subtracted. (3) This results is the precise thalamic volume sending convergent projections to each cortical

subregion, which is separated into six confidence levels (see **Methods**, modified from (Hunnicutt et al. 2014)). (**b**) Single coronal section through the confidence map for Sub (gray scale) overlaid with nuclear subdivisions from the AMBA. The atlas is colored on the left to indicate the fraction of each nucleus covered by the average of confidence levels 1, 3, and 5 also averaged across the PMBA and AMBA. The color scale minimum (blue) is 0%, inflection point (white) is 25%, and the peak coverage (red) is 75%. (**c**) Summary of thalamic confidence maps describing the origin of thalamostriatal projections convergent with each cortical subregion. For each convergent subregion, coronal thalamus sections are shown from -0.155 mm posterior to bregma, and continuing in 250 μm increments posterior (columns). (**d**) The fractions of each thalamic nucleus covered by confidence levels 1, 3 and 5 (dark, mid and light gray bars, respectively), with their average (black line). (**e**) Aggregate nucleus coverage map for the thalamostriatal projections convergent with each cortical subregion. Nuclei (rows) and cortical subregions (columns) are hierarchically clustered on the basis of output and input similarity, respectively. Color scale is the same as **b**.

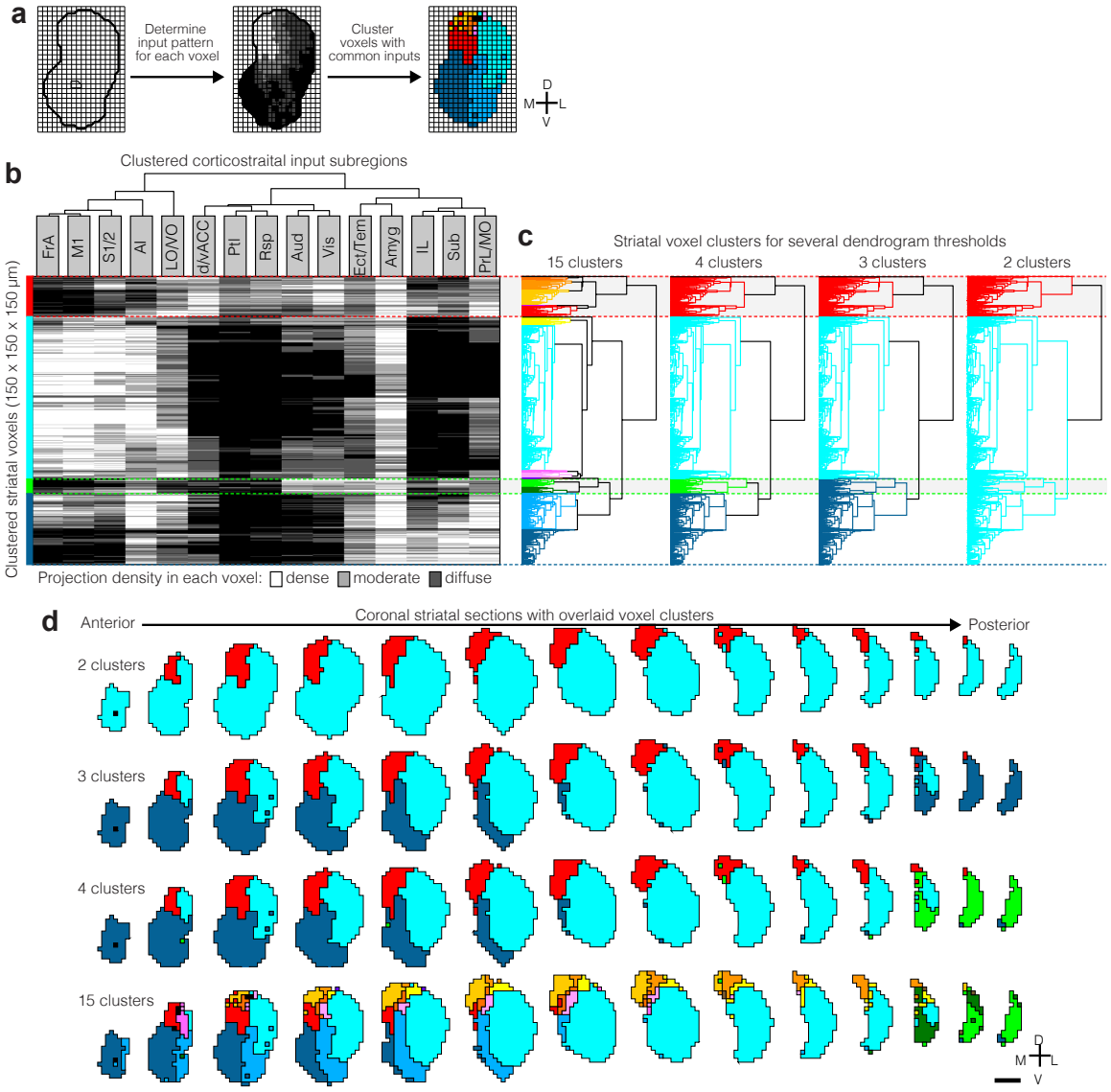
Figure 5. Input convergence and cortical subtype distributions.



(**a**) Coronal sections through the average template brain showing the cumulative bilateral convergence of diffuse (top) and dense (bottom) projections from all cortical subregions (heat map). Areas above and below 9 convergent subregions are indicated (black dashed line). Sections start 1.6 mm anterior to bregma, the second slice is 300 μm posterior, and the rest continue in 600 μm steps. (**b**) Distribution plots for striatal convergence of diffuse (top)

and dense (bottom) projections, shown for the cumulative number of converging projections in each axis (calculated as described in **Figure 2**). (**c**) Subregion specific convergence plots for diffuse (left) and dense (right) projections. The color scale indicates the fraction of the projection field from the subregion in each row covered by the projection field from the subregion in each column. (**d**) Coronal sections through the average template brain showing the bilateral distribution of dense, moderate and diffuse projections from all allocortical, mesocortical, and neocortical subregions. Section positions are the same as in panel **a**. (**e**) Distribution plots for the dense (light gray) and diffuse (dark gray) projections of each cortical subtype shown in panel **d** (calculated as described in **Figure 2**). (**f**) Thalamic confidence maps for the origins of thalamostriatal projections convergent with each cortical subtype (left), allocortical and neocortical is calculated for the non-overlapping volumes of their dense projections, and mesocortical is calculated for the full distribution of its dense projections (section locations as in **Figure 4c**). The fractions of each thalamic nucleus covered by confidence levels 3, 5 and 7 (dark, mid and light gray bars, respectively), with their average (black line) are shown on the left (see **Figure 4b** and **Methods** for details). (**g**) Summary of thalamic confidence maps for the origins of thalamostriatal projections to striatal volumes of high and low diffuse projection convergence determined in panel **a** (section positions are the same as in **Figure 4c**). The fractions of each thalamic nucleus covered by confidence levels 3, 5 and 7 (dark, mid and light gray bars, respectively), with their average (black line) are shown on the left (see **Methods** for details).

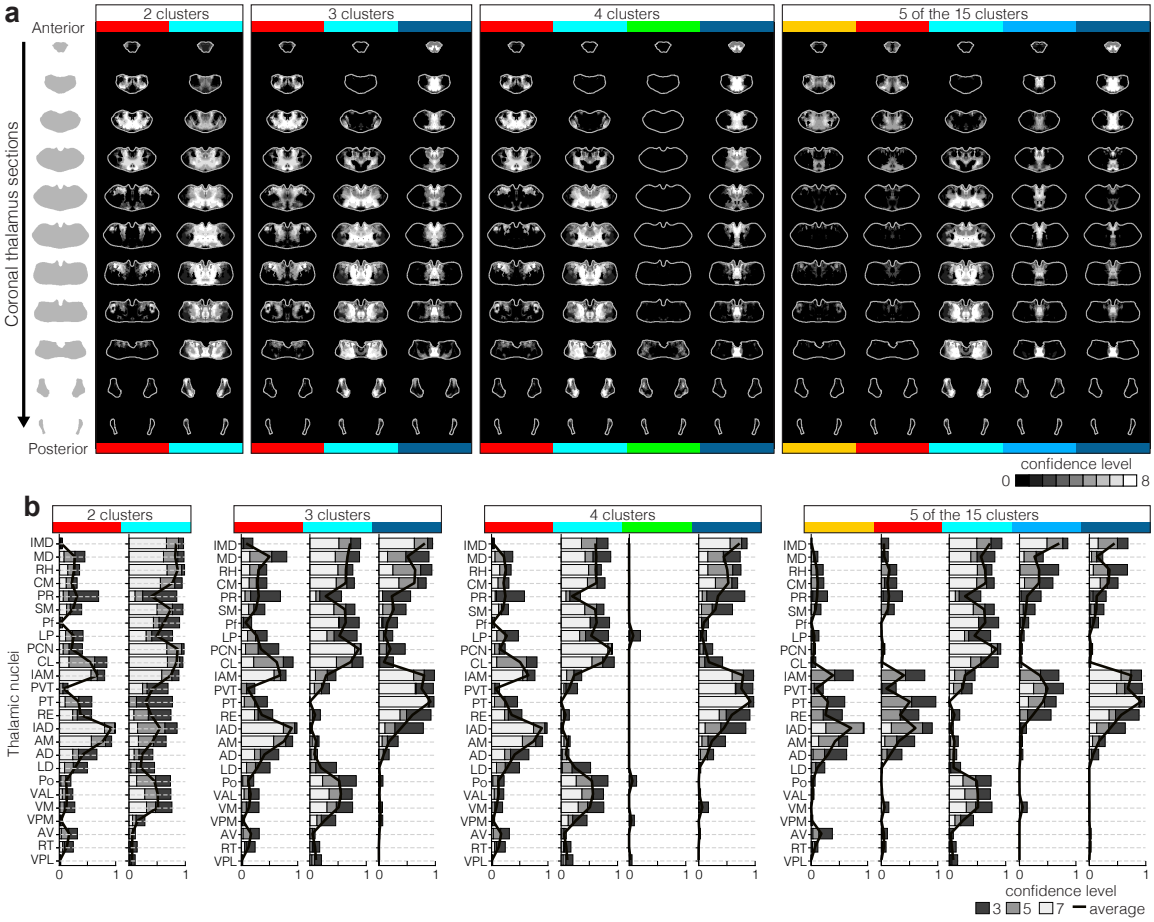
Figure 6. Striatal segmentation based on cortical input convergence.



(**a**) Schematic of voxel clustering. The ipsilateral striatum was down-sampled into 150 × 150 × 150 µm voxels (left), the projection density (dense, moderate, or diffuse) to each voxel was determined for each cortical subregion (example, center), and the sum of this information was used to cluster voxels with common inputs (example, right). (**b**) All striatal voxels (rows) were hierarchically clustered based on their cortical input patterns, and cortical subregions (columns) were clustered based on which striatal voxels they innervated. The projection densities in each voxel are indicated in gray scale, as in determined in **Figure 2a**. (**c**) Four separate thresholds were applied to the voxel dendrogram to produce 2, 3, 4, and 15 clusters (right to left). The leaf borders for the threshold producing 4 clusters are carried across the clustered voxels panel **b** for comparison. Clusters containing only one voxel were ignored in our analyses (top linkage node for 3, 4 and 15 clusters). (**d**) Coronal sections through the

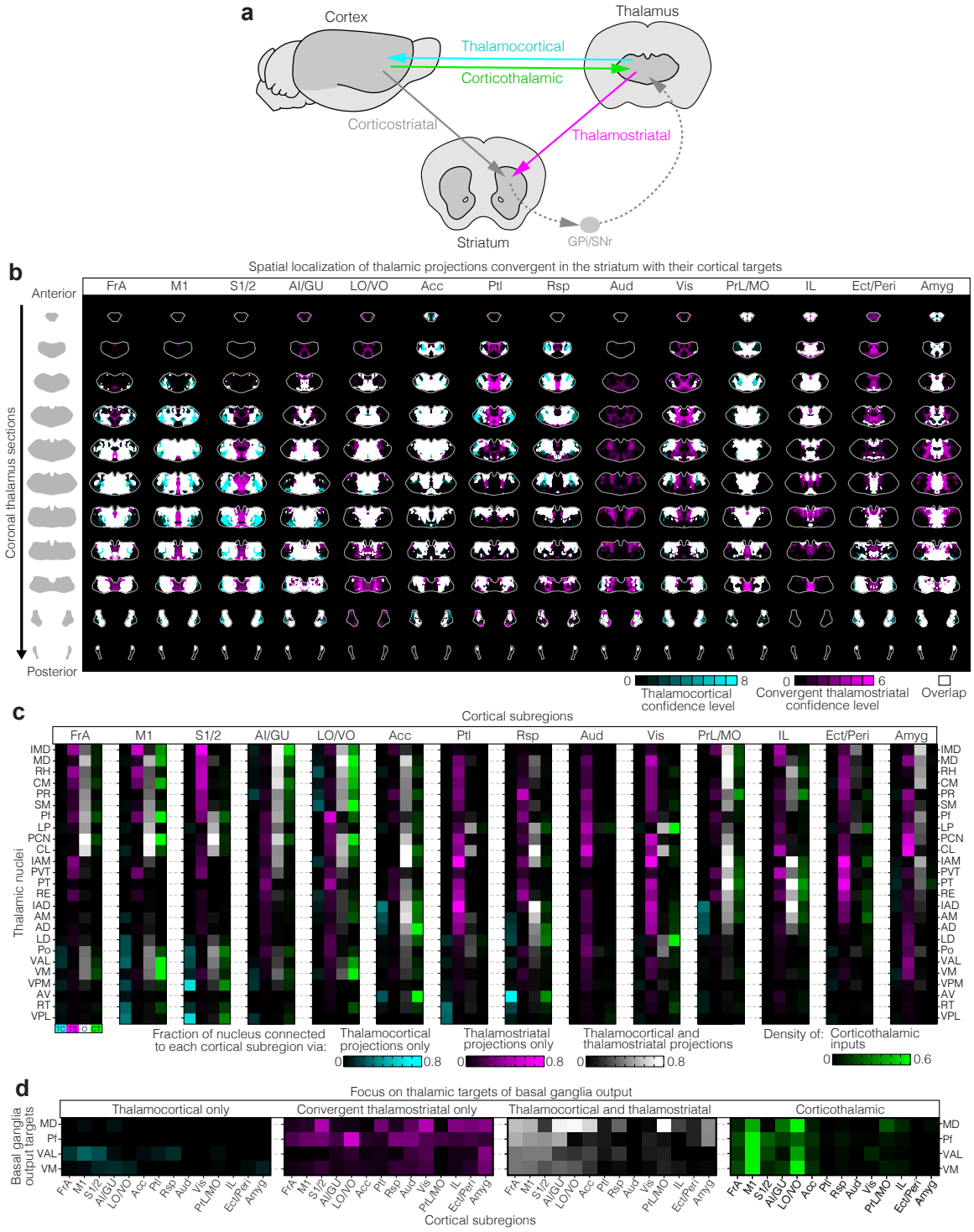
ipsilateral striatum (starting 1.6 mm anterior to bregma and continuing posterior in 300 μm steps) showing the spatial location of the clusters determined in panel **c** (rows).

Figure 7. Thalamic origins of inputs to striatal clusters.



(**a**) Summary of thalamic confidence maps for the origins of thalamostriatal projections to the striatal voxel clusters found in **Figure 6**. Confidence maps are shown for all subdivisions of 2, 3, and 4 clusters, as well as 5 of the 15 clusters. The confidence levels are shown in grayscale (section positions are the same as in **Figure 4c**). (**b**) The fractions of each thalamic nucleus covered by confidence levels 3, 5 and 7 (dark, mid and light gray bars, respectively), with their average (black line) are shown for the confidence maps in panel **a** (see **Methods** for details).

Figure 8. Complete organization of the thalamus in cortico-thalamo-basal ganglia loops.



(**a**) Diagram showing of the connection between the cortex, thalamus, striatum, and the output nuclei of the basal ganglia, the GPi and SNr, which collectively make up the cortico-thalamo-basal ganglia loop. (**b**) Overlaid confidence maps for the origins of thalamostriatal

projections that converge with subregion specific corticostriatal projections (magenta, same as **Figure 4c**), and confidence maps for the origins of subregion specific thalamocortical projections (cyan, previously published data (Hunnicutt et al. 2014)). Overlapping volumes are shown in white. (see **Figure 4** and **Methods** for details). (**c**) Nuclear localization for the confidence maps shown in panel **b**. Values are represented as the fraction of each thalamic nucleus covered by the average of confidence levels 1, 3, and 5 for thalamostriatal projections (magenta), the average of confidence levels 1, 4, and 7 for thalamocortical projections (cyan) and the average of confidence levels 1, 3, and 5 for thalamostriatal projections that lie within the white overlapping volume shown in panel **b**. The density of subregion specific corticostriatal projections to each nucleus is shown in green. (**d**) The nuclear localization data from panel **c** are grouped by projection type and only shown for the thalamic targets of basal ganglia output, MD, Pf, VAL, and VM.

Supplementary Information

Table 1. AIBS Mouse Connectivity Atlas injection metadata.

Experiment ID	Primary target structure	Cortical group	Cortical layer	Transgenic-line	Strain	Sex	Injection coordinates (mm from bregma)	Injection volume (mm^3)
100140756	MOs	FrA	All	WT	C57BL/6J	M	[2400, 2600, 7100]	0.124482
100140949	RSPv	Rsp	All	WT	C57BL/6J	M	[7700, 1200, 6000]	0.117875
100141219	VISp	Vis	All	WT	C57BL/6J	M	[9100, 1300, 8000]	0.173612
100141599	VISam	Vis	All	WT	C57BL/6J	M	[8100, 1000, 7300]	0.119866
100141796	VISl	Vis	All	WT	C57BL/6J	M	[9400, 2400, 9300]	0.158867
100142655	SSp-tr	S1/2	All	WT	C57BL/6J	M	[7000, 1600, 8200]	0.11662
100148142	RSP	Rsp	All	WT	C57BL/6J	M	[7200, 1200, 6000]	0.07169
100149109	AUDp	Aud	All	WT	C57BL/6J	M	[7800, 2800, 9800]	0.049439
112162251	SSp-bfd	S1/2	All	WT	C57BL/6J	M	[5700, 2100, 9300]	0.117388
112229103	PTLp	Ptl	All	WT	C57BL/6J	M	[7700, 1000, 6800]	0.03226
112306316	ORBl	LO/VO	All	WT	C57BL/6J	M	[3100, 3500, 7000]	0.324949
112595376	RSPv	Rsp	All	WT	C57BL/6J	M	[8600, 1600, 6800]	0.066831
112596790	AI	AI/GI/DI	All	WT	C57BL/6J	M	[3800, 4300, 8600]	0.294591
112670853	MOp	M1/2	All	WT	C57BL/6J	M	[3100, 2600, 8000]	0.214222
112881858	AUDv	Aud	All	WT	C57BL/6J	M	[8300, 3400,10000]	0.086969
112936582	SSp-m	S1/2	All	WT	C57BL/6J	M	[4600, 2700, 8900]	0.175167
113144533	BLAa	Amyg	All	WT	C57BL/6J	M	[6700, 5800, 8900]	0.310123
113887162	VISp	Vis	All	WT	C57BL/6J	M	[8300, 1300, 8100]	0.35562
116903968	VISal	Vis	All	WT	C57BL/6J	M	[8900, 1800, 9000]	0.156213
117298988	SSs	S1/2	All	WT	C57BL/6J	M	[6400, 3700, 9700]	0.282782
120491896	AUDp	Aud	All	WT	C57BL/6J	M	[7900, 3000, 9800]	0.064253
120814821	MOp	M1/2	L5	Rbp4-Cre_KL100	unknown	F	[4100, 2200, 7300]	0.156917
120875816	SSp-bfd	S1/2	L5	Rbp4-Cre_KL100	unknown	F	[5900, 2100, 8800]	0.079808
121510421	VISp	Vis	L5	Rbp4-Cre_KL100	unknown	F	[9400, 1400, 8000]	0.035524
122641784	SUBv	Sub	All	WT	C57BL/6J	M	[8100, 5400, 8900]	0.056891
125832322	BMAp	Amyg	All	WT	C57BL/6J	M	[7300, 6200, 8500]	0.253454
126117554	ACAv	d/vACC	L5	Rbp4-Cre_KL100	unknown	M	[5200, 2400, 6100]	0.142052
126908007	SSp-bfd	S1/2	All	WT	C57BL/6J	M	[5800, 2000, 9200]	0.325616
127222723	SUBv	Sub	All	WT	C57BL/6J	M	[9100, 4200, 9300]	0.50004
127649005	CA3	Sub	All	WT	C57BL/6J	M	[8100, 4900, 8500]	0.540497
127795906	SUBd	Sub	All	WT	C57BL/6J	M	[8900, 2300, 8100]	0.28667
139426984	ACAd	d/vACC	All	WT	C57BL/6J	M	[5300, 1900, 5900]	0.086324
139520203	ACAv	d/vACC	All	WT	C57BL/6J	M	[5400, 1900, 6000]	0.116901
141602484	MOp	M1/2	All	WT	C57BL/6J	M	[4200, 1900, 7300]	0.180769
142656218	PERI	Ect/Peri/Tem	All	WT	C57BL/6J	M	[9400, 3800, 9800]	0.123722
146077302	VISam	Vis	All	WT	C57BL/6J	M	[8600, 900, 7200]	0.052389
146858006	AUDd	Aud	All	WT	C57BL/6J	M	[8600, 2500, 9800]	0.084186
152994878	SUBv	Sub	All	WT	C57BL/6J	M	[9300, 4100, 9300]	0.116975
156394513	MOp	M1/2	L5	A930038C07Rik-Tg1-Cre	B6.Cg	F	[3800, 2100, 7100]	0.104113
156493815	AUDp	Aud	L5	Etv1-CreERT2	B6.Cg	F	[7800, 2600, 9700]	0.2045
156741826	ORBl	LO/VO	L5	Rbp4-Cre_KL100	unknown	M	[2600, 3700, 7100]	0.397148
157062358	TEa	Ect/Peri/Tem	All	WT	C57BL/6J	M	[9500, 2800, 9600]	0.060422
157063781	SUBd	Sub	All	WT	C57BL/6J	M	[8900, 2200, 8100]	0.113087
157556400	ILA	IL	All	WT	C57BL/6J	M	[4000, 3700, 6200]	0.105299
157654817	SSp-m	S1/2	All	WT	C57BL/6J	M	[4200, 3400, 9200]	0.095364
157710335	MOs	FrA	All	WT	C57BL/6J	M	[2600, 2700, 7400]	0.087187
157711748	PL	PrL/MO	All	WT	C57BL/6J	M	[3200, 3300, 6100]	0.108321
158255941	VISp	Vis	L5	A930038C07Rik-Tg1-Cre	B6.Cg	F	[9400, 1400, 8300]	0.039007
158314278	AUDd	Aud	All	WT	C57BL/6J	M	[7900, 2300, 9700]	0.243249
158435116	ORBvl	LO/VO	All	WT	C57BL/6J	M	[2900, 3800, 6700]	0.125197
159319654	ACAv	d/vACC	L5	A930038C07Rik-Tg1-Cre	B6.Cg	F	[5200, 2100, 6000]	0.045202

159832064	RSPv	Rsp	L5	Rbp4-Cre_KL100	unknown	F	[7800, 1400, 6000]	0.117248
161458737	ACAd	d/vACC	L5	Rbp4-Cre_KL100	unknown	F	[4600, 2400, 6000]	0.35836
166054929	RSPd	Rsp	L5	Rbp4-Cre_KL100	unknown	M	[8500, 1000, 6100]	0.250529
166082128	MOp	M1/2	L5	Rbp4-Cre_KL100	unknown	M	[3300, 2400, 7800]	0.322731
166083557	TEa	Ect/Peri/Tem	L5	Rbp4-Cre_KL100	unknown	F	[9600, 2800, 9500]	0.143854
166153483	AId	AI/GI/DI	L5	Rbp4-Cre_KL100	unknown	M	[3300, 4300, 8100]	0.345466
166271142	RSP	Rsp	L5	A930038C07Rik-Tg1-Cre	B6.Cg	F	[7700, 1200, 6000]	0.094002
166323186	SSp-ll	S1/2	L5	A930038C07Rik-Tg1-Cre	B6.Cg	M	[6200, 1500, 7700]	0.035333
166323896	ACAd	d/vACC	L5	A930038C07Rik-Tg1-Cre	B6.Cg	M	[4200, 2000, 6100]	0.163949
166324604	VISp	Vis	L5	A930038C07Rik-Tg1-Cre	B6.Cg	M	[8600, 1300, 8100]	0.077434
166461899	VISl	Vis	L5	A930038C07Rik-Tg1-Cre	B6.Cg	F	[9500, 2300, 9300]	0.032779
167794131	VISal	Vis	L5	Rbp4-Cre_KL100	unknown	F	[8600, 1700, 8800]	0.115341
168002073	MOp	M1/2	L234	Cux2-IRES-Cre	unknown	M	[3800, 1900, 7100]	0.202714
168003640	SSp-bfd	S1/2	L234	Cux2-IRES-Cre	unknown	M	[6100, 2100, 8900]	0.231043
168163498	SSs	S1/2	L5	Rbp4-Cre_KL100	unknown	M	[5800, 3600, 9700]	0.242035
168164972	ORBl	LO/VO	L5	A930038C07Rik-Tg1-Cre	B6.Cg	M	[2700, 3700, 7000]	0.288391
168165712	PTLp	Ptl	L5	A930038C07Rik-Tg1-Cre	B6.Cg	F	[7900, 1100, 7400]	0.015441
171276330	SSp-bfd	S1/2	L5	A930038C07Rik-Tg1-Cre	B6.Cg	M	[5900, 2100, 8700]	0.012751
174361746	AIp	AI/GI/DI	All	WT	C57BL/6J	M	[6300, 4700, 9900]	0.363888
176430283	AUDd	Aud	L234	Cux2-IRES-Cre	unknown	M	[7800, 2400, 9500]	0.293934
178488859	PRE	Sub	NA	Satb2-Cre_MO23	unknown	F	[9600, 4100, 8600]	0.737451
180719293	MOp	M1/2	All	WT	C57BL/6J	M	[2400, 3000, 7700]	0.769831
180917660	VISC	AI/GI/DI	All	WT	C57BL/6J	M	[5600, 4400, 9800]	1.16191
181600380	AUDp	Aud	L5	Gpr26-Cre_KO250	unknown	M	[7900, 2700, 9800]	0.041505
182226839	CA1	Sub	NA	Vipr2-Cre_KE2	unknown	M	[9200, 3700, 9500]	0.175034
182294687	BMAp	Amyg	NA	Rbp4-Cre_KL100	unknown	F	[7100, 6500, 8500]	0.217468
182467026	SSp-bfd	S1/2	L5	A930038C07Rik-Tg1-Cre	B6.Cg	F	[6400, 2000, 9000]	0.01632
182616478	MOp	M1/2	L234	Cux2-IRES-Cre	unknown	F	[6200, 1200, 6900]	0.120931
182794184	ECT	Ect/Peri/Tem	L5	Rbp4-Cre_KL100	B6.Cg	F	[9700, 4000, 9800]	0.066608
183461297	ACAd	d/vACC	L234	Cux2-IRES-Cre	unknown	F	[5300, 2100, 5900]	0.086148
183470468	ACAv	d/vACC	L234	Cux2-IRES-Cre	unknown	F	[5500, 2100, 5900]	0.182905
183471174	ORBl	LO/VO	L234	Cux2-IRES-Cre	unknown	F	[2500, 3600, 7100]	0.204871
184167484	PTLp	Ptl	L234	Cux2-IRES-Cre	B6.Cg	F	[8000, 800, 7300]	0.12762
184168899	AId	AI/GI/DI	L234	Grp-Cre_KH288	B6;FVB;ICR	M	[3600, 4700, 8800]	0.094153
263106036	PL	PrL/MO	L5	Rbp4-Cre_KL100	unknown	M	[3700, 3500, 6000]	0.321523
263242463	MOs	FrA	L5	Rbp4-Cre_KL100	unknown	F	[2600, 2700, 7200]	0.169063
263780729	VISam	Vis	L234	Cux2-IRES-Cre	unknown	M	[9000, 1400, 7900]	0.171675
264629246	SSp-ll	S1/2	L5	A930038C07Rik-Tg1-Cre	B6.Cg	F	[6400, 1400, 7400]	0.011544
264630019	VISl	Vis	L5	A930038C07Rik-Tg1-Cre	B6.Cg	F	[9400, 2600, 9600]	0.029828
266486371	SSp-bfd	S1/2	L234	Cux2-IRES-Cre	B6.Cg	F	[7200, 1300, 8700]	0.105705
266487079	VISal	Vis	L234	Cux2-IRES-Cre	B6.Cg	M	[8700, 1700, 8900]	0.108468
266644610	SSp-ll	S1/2	L234	Cux2-IRES-Cre	unknown	F	[6100, 1400, 7600]	0.136925
272414403	ECT	Ect/Peri/Tem	All	WT	C57BL/6J	M	[9500, 3600, 9800]	0.258675
272735030	SSp-bfd	S1/2	L5	Rbp4-Cre_KL100	B6.Cg	M	[7400, 1600, 8500]	0.081838
272735744	RSP	Rsp	L234	Cux2-IRES-Cre	B6.Cg	M	[8200, 800, 6000]	0.091458
272737914	GU	AI/GI/DI	All	WT	C57BL/6J	M	[4200, 4700, 9100]	0.195609
272821309	VISp	Vis	L234	Cux2-IRES-Cre	B6.Cg	F	[8800, 1000, 8100]	0.126388
277710753	BLAa	Amyg	All	WT	C57BL/6J	M	[6300, 5900, 8800]	0.211792
278317239	ECT	Ect/Peri/Tem	L234	Cux2-IRES-Cre	B6.129	M	[9900, 4100, 10000]	0.287965
278317945	SSp-m	S1/2	L234	Cux2-IRES-Cre	B6.129	M	[5100, 2200, 9000]	0.143174
283019341	ORBm	PrL/MO	L5	A930038C07Rik-Tg1-Cre	B6.Cg	M	[3500, 3400, 6100]	0.386955
283020912	VISp	Vis	L5	A930038C07Rik-Tg1-Cre	B6.Cg	M	[8600, 1200, 8000]	0.043266
286299886	MOs	FrA	L234	Cux2-IRES-Cre	B6.Cg	F	[2200, 3000, 7300]	0.062127
286300594	SSp-ll	S1/2	L234	Cux2-IRES-Cre	B6.Cg	M	[5300, 1700, 8100]	0.162377
286312782	SSp-ul	S1/2	L5	Rbp4-Cre_KL100	B6.Cg	F	[5200, 2200, 8000]	0.080407
286313491	ILA	IL	L5	Rbp4-Cre_KL100	B6.Cg	F	[3900, 3500, 6100]	0.16043
286610216	SUBv	Sub	NA	Syt17-Cre_NO14	B6.Cg	F	[8600, 5300, 9300]	0.082374
286610923	CA1	Sub	NA	Syt17-Cre_NO14	B6.Cg	M	[8900, 3700, 9600]	0.188479

287494320	ILA	IL	L234	Cux2-IRES-Cre	B6.Cg	F	[3900, 3500, 6100]	0.172357
287495026	VISal	Vis	L234	Cux2-IRES-Cre	B6.Cg	F	[9200, 2100, 9300]	0.111757
287769286	ORBvl	LO/VO	L5	Rbp4-Cre_KL100	B6.Cg	F	[2900, 3700, 6900]	0.274778
292373346	RSPv	Rsp	L5	A930038C07Rik-Tg1-Cre	unknown	F	[8700, 1500, 6700]	0.091962
292374068	MOp	M1/2	L5	A930038C07Rik-Tg1-Cre	B6.Cg	M	[3300, 2600, 7600]	0.119589
292374777	PL	PrL/MO	L5	A930038C07Rik-Tg1-Cre	unknown	F	[3600, 3300, 6100]	0.1849
292476595	MOs	FrA	L5	A930038C07Rik-Tg1-Cre	unknown	F	[2800, 2800, 7700]	0.072795
292532065	CA1	Sub	NA	Ppp1r17-Cre_NL146	FVB/N-Crl:CD1(ICR)	M	[8600, 2100, 8000]	0.212305
292792016	SSp-ul	S1/2	L5	A930038C07Rik-Tg1-Cre	B6.Cg	M	[5200, 2000, 8000]	0.010348
292792724	PTLp	Ptl	L5	A930038C07Rik-Tg1-Cre	B6.Cg	M	[7900, 1300, 8200]	0.032602
293471629	ORBm	PrL/MO	L234	Cux2-IRES-Cre	B6.Cg	M	[3000, 3300, 6000]	0.168764
294396492	PL	PrL/MO	L234	Cux2-IRES-Cre	B6.Cg	F	[3400, 3200, 6100]	0.05761
294481346	VISam	Vis	L5	Rbp4-Cre_KL100	unknown	F	[8700, 1100, 7300]	0.127075
294482052	VISp	Vis	L5	Rbp4-Cre_KL100	B6.Cg	M	[8800, 1300, 8200]	0.05166
294484177	RSPagl	Rsp	L234	Cux2-IRES-Cre	B6.Cg	F	[9800, 900, 7100]	0.066937
296048512	AI	AI/GI/DI	L5	Rbp4-Cre_KL100	B6.Cg	M	[3600, 4500, 8200]	0.367829
297652799	SSp-tr	S1/2	L5	Rbp4-Cre_KL100	B6.Cg	F	[6400, 1400, 7600]	0.308236
298324391	VISam	Vis	L5	A930038C07Rik-Tg1-Cre	unknown	M	[8900, 1000, 7400]	0.01477

CHAPTER 3:
DISCUSSION & FUTURE DIRECTIONS

Advances in anatomical mapping and data sharing

The collective effort of the scientific community over the last century to map the neuronal connections within brain has produced a wealth of anatomical data on the circuitry of the basal ganglia. From the detailed drawings of the structure of striatal input created by the Spanish neuroanatomist Santiago Ramon y Cajal in 1911 (Haycock & Bro 1975), to modern cell-type specific retrograde studies of striatal input convergence (Wall et al. 2013), and all of the careful and laborious work in between (Cowan & Powell 1956; McGeorge & Faull 1989; Groenewegen et al. 1999; McFarland & Haber 2000; Pan et al. 2010), there is a wealth of information on basal ganglia neuroanatomy. The dataset presented here is not novel in its content, but rather in its breadth. By localizing all anatomical data in a common coordinate system, i.e. in the AIBS average template brain and our model thalamus, we could directly compare the projection patterns across all input modalities.

Historically, to investigate whole brain connectivity, subjective assessments of projection trends were used to compare across experiments. This method, although useful, is laborious and inevitably results in experimental biases when data is subjectively consolidated (Y. Smith et al. 2004; Berendse et al. 1992; Berendse & Groenewegen 1990; McGeorge & Faull 1989; Willuhn et al. 2003). This method was manageable when these datasets were small, usually consisting of less than 20 tracing experiments at a time, but recent advances in high-throughput fluorescent imaging have facilitated the generation of large anatomical datasets with hundreds of tracer injections (Thompson & Swanson 2010; Hintiryan et al. 2012; Oh et al. 2014; Zingg et al. 2014). Extracting relevant biological information from these data is nearly impossible when the experiments need to be compared directly, necessitating the adoption of computational methods to analyze these large anatomical studies.

To combine data from independent experiments, each dataset must be aligned well enough to compare them directly. Intrinsic size differences between brains, as well as distortions caused by experimental manipulation, provide a lower limit to how well any two

experiments can be aligned. The estimated maximum computational resolution in the mouse brain is ~100 μm(Ragan et al. 2012), and falls into the mesoscale, a mapping resolution that balances the caveats associated with microscopic and macroscopic mapping techniques. Combining datasets with higher resolution, such as single cell tracing, introduces relatively more variability and increased labor, and macroscopic mapping techniques, such as resting state functional magnetic resonance imaging (rs-fMRI), do not provide the necessary detail to localize functional brain regions (Bohland et al. 2009; Mitra 2014). The datasets described in this dissertation are at the mesoscopic level (see **Chapters 1 & 2**), a resolution which sacrifices cellular specificity, but provides both a reproducible and detailed description of the connectivity patters across the cortex and striatum.

Many of the functional subregions of the thalamus, i.e. thalamic nuclei, are smaller than the resolution obtained with mesoscopic mapping techniques and incompatible with the shape of tracer injections (Jones 2007; Oh et al. 2014), making nuclear localization difficult. We developed a novel method to localize the origin of anterograde projections. By supersampling the thalamus, we can localize thalamic volumes smaller than individual injection sites, allowing us to isolate regions of the thalamus as small as most thalamic nuclei (see **Chapter 1: Methods & Supplementary Fig. 8**). Using this method we produced the most comprehensive maps of thalamocortical and thalamostriatal connectivity ever created.

A related challenge associated with analyzing these large anatomical datasets is the ability to compare across studies. It is difficult enough to computationally align several experiments generated with identical methods, but aligning analogous datasets generated by separate groups poses an even greater challenge. To utilize the abundance of data currently being generated by the scientific community, it is crucial that the data is generated and documented in a form that allows for lateral comparisons. For our analysis of the cortico-thalamo-basal ganglia circuit, we integrated an anterograde cortical dataset generated by the Allen Institute for Brain Science (AIBS) with our anterograde thalamic dataset (Oh et al. 2014; Hunnicutt et al. 2014). This was possible thanks to the commitment made by the AIBS to data sharing

and open science. Combining the AIBS corticostriatal dataset with our thalamostriatal and thalamocortical datasets allowed us to investigate excitatory input convergence in the basal ganglia with unprecedented breadth, gaining insight into a brain region that has remained one of the most poorly understood for decades (see **Chapter 2**).

Age differences across studies

Our thalamocortical and thalamostriatal projection maps were generated in adolescent mice (P30), but the AIBS corticostriatal data was generated in adult mice (P56) (Hunnicutt et al. 2014; Oh et al. 2014). The overall structure of the mouse neo-cortex, including relative cortical layer width, has just reached steady state by P30 (Van Eden & Uylings 1985; Spear 2000), and sensorimotor cortex is considered functionally mature using behavioral and functional measurements by P28 (Yang et al. 2012; Weiler et al. 2008; Mao et al. 2011), but subregions of prefrontal cortex are commonly accepted as functionally immature during adolescence, which occurs between P25 and P50 in rodents (Sturman & Moghaddam 2011; Spear 2000). The functional maturation of prefrontal cortex likely involves some structural changes, so we wanted to ensure that the gross anatomical connectivity of adolescent and adult mice were the same. We carefully compared the thalamo-prefrontal projection patterns seen in our data to the existing literature in adult rodents (see **Chapter 1: Supplementary Fig. 13**). We found that these projection distributions are comparable across ages, suggesting that the dramatic behavioral changes occurring during this epoch (at least after P30) are more likely due to local refinements and synaptic pruning, rather than larger-scale rearrangements in the complement of thalamic nuclei projecting to prefrontal cortex. Although we cannot definitively resolve the maturation time point, our data suggest that thalamic projections to prefrontal cortex have reached their final targets by P30 in mouse. Prefrontal cortex is considered to be the brain region that becomes functionally mature last, making us confident that the broad connectivity patterns for all mouse brain regions are comparable at P30 and P56.

Striatal segmentation

There was a time when treating the striatum as a single functional unit was sufficient to investigate the basic circuit of the basal ganglia. As with many categorization systems, we eventually reach a point where we can no longer assimilate information into the current model and must update our framework to accommodate new data. As researchers identified cell density differences, and differential sensitivity of the ventral striatum to psychostimulant drugs, the striatum was divided into the ventral nucleus accumbens (NAc) and the dorsal striatum (Steiner & Tseng 2010). More recently, behavioral data has made it clear that this dorsal-ventral divide was insufficient to explain the heterogeneity of striatal function, and a third functional subdivision was created. Guided by both anatomical input distributions and the identification of differences in behavioral control, the striatum is currently thought to consist of three functional domains, the limbic, associative and sensorimotor domains (Parent & Hazrati 1995; Voorn et al. 2004; Haber & McFarland 1999) (see the **Introduction**, **Fig. 3**).

The demarcation of these three functional domains in the striatum has been useful for forming and testing hypotheses about basal ganglia function. However, the input distributions across these subdivisions are not entirely uniform (see the **Introduction** and **Chapter 2**, **Fig. 2,6**). Although it is possible that the limbic, associative and sensorimotor domains are the final organizational step in striatal input processing, given the heterogeneity of inputs across them, it seems unlikely. It is clear that the anatomical organization of the basal ganglia is set up to maintain the gross topography of striatal inputs as they are processed and relayed to the output nuclei and subsequently to the thalamus and cortex (Draganski et al. 2008). This organization, although amenable to a tripartite domain system in the striatum, does not rely on one. Any division system, or conversely, no division system and instead a gradient of functionality, could be consistent with this anatomical set-up. The only clear division is imposed on projections from the output nuclei as they diverge between VAL and MD (Haber & Calzavara 2009), but it is already assumed that there are more than two striatal subdivisions endowed with unique behavioral functionality (Gruber &

McDonald 2012; Yin 2014). Using the comprehensive dataset created here, the striatum was objectively parcellated into a variable number of subdivisions. When the striatum is segmented into three anatomically defined subdivisions, the resulting volumes that are demarcated are largely homologous to the three behaviorally defined domains described above, effectively circumventing decades of research. Additionally, at least two additional putative functional domains have been identified. A posterior area was identified with a pattern of cortical and thalamic inputs reminiscent of both the limbic and associative domains, but fundamentally distinct in several ways, and the limbic domain was subdivided into lateral and medial aspects that have differential access to sensorimotor information (see **Chapter 2**, **Fig. 6**).

The three striatal subdivisions identified here are located in the same relative locations as the limbic, associative, and sensorimotor domains, with the main difference being that their precise boundaries are demarcated throughout the entire striatum. The first subdivision and the sensorimotor striatum are located dorsolaterally, the second subdivision and the limbic striatum are located ventrally and encompasses the nucleus accumbens (NAc), while the third subdivision and the associative striatum occupy a poorly defined area between the limbic and sensorimotor domains but are generally located in the dorsomedial striatum (see **Introduction: Fig. 3 & Chapter 2: Fig. 6**).

The cortical input patterns to these three subdivisions are also largely consistent with the known inputs to the sensorimotor, associative, and limbic domains in the striatum. As described previously, the sensorimotor domain receives extensive and topographically overlapping sensory and motor inputs (Malach & Graybiel 1986; Nambu 2011) (see **Introduction: Fig. 3**). Behaviorally, this area is necessary for the acquisition and execution of stimulus-response relationships, i.e. habitual behavior. This form of learning is, by definition, uncoupled from reward value and outcome anticipation (Yin & Knowlton 2006). Anatomically, this has been explained by a lack of input from the amygdala and hippocampus, which carry affective and outcome associations respectively (see **Introduction**). When we compare the sensorimotor domain to the analogous subdivision

created by our analyses (cyan, **Chapter 2: Fig. 6**), we do see strong overlapping input from the primary sensory and motor cortices, frontal associative motor cortex, and insular areas conveying gustatory sensory information (M1, S1/2, FrA, and AI). Interestingly, we also see broad innervation from the amygdala and weak innervation by prefrontal areas such as lateral & ventral orbital cortices (see **Chapter 2: Fig. 6b**). These inputs may have been overlooked previously due to their relatively low density compared to the sensory and motor inputs in this region.

The associative domain of the striatum is important for behavioral flexibility, allocentric navigation, and goal-directed forms of instrumental learning (Gruber & McDonald 2012; Yin & Knowlton 2006). The navigational functions of the associative striatum are thought to be mediated by the hippocampal information it receives via direct projections from the entorhinal cortex and subiculum, as well as indirect hippocampal information from two primary entorhinal targets, the medial prefrontal cortex and retrosplenial cortex. The non-spatial response flexibility is thought to be mediated either by inputs from sensory modalities directly or indirectly via inputs from other multimodal areas (Khibnik et al. 2014; Gruber & McDonald 2012). Finally, the associative striatum is strongly modulated by reward anticipation, which is attributed to its amygdalar inputs (Yin & Knowlton 2006). In the analogous subdivision created by our analysis (red, **Chapter 2: Fig. 6**), all of these inputs are present. In fact, this subdivision receives input from all subregions examined except motor cortices, making it the predominant site of input convergence within the striatum (see **Chapter 2: Fig. 5 & 6**), which is congruent with its role as an associative area.

The limbic domain integrates information on affective state, motivation and reward (Haber & Knutson 2010; Euston et al. 2012), and receives extensive excitatory inputs from the prefrontal cortex, the amygdala, and the hippocampal formation (Gruber & McDonald 2012). The analagous subdivision created in our analysis (dark blue, **Chapter 2: Fig. 6**) also receives strong projections from the amygdala, hippocampal structures (the subiculum and entorhinal cortex), as well as from several prefrontal cortical areas including infralimbic, prelimbic, and orbital cortices. In addition to the expected inputs, this subdivision also

received broad projections from insular cortex and partial innervation by sensory, motor, auditory, and visual areas.

These inputs that only partially cover the subdivision analagous to the limbic domain may define functionally significant subregions. When the clusters are expanded from 3 to 4, the posterior aspect of this limbic-like cluster becomes the next subregion. The primary difference between this 4th subregion and the canonical limbic domain is the presence of strong visual and auditory inputs, possibly endowing this area with the ability to directly integrate these two sensory modalities with motivation and reward. As the striatum is segmented into more and more clustered subdivisions, the next 10 clusters break the dorsomedial, associative-like subregion into ever smaller volumes, which adds to the evidence for this subregion being important for a variety of associative tasks. Finally, at the 15th cluster, the ventral striatum was separated into two proportional subdivisions (mid-dark blue, **Chapter 2: Fig. 6c-d**). The anatomical distinction between these two subdivisions is the presence of motor and sensory inputs to the lateral aspect of the limbic-like cluster. The limbic striatum is generally thought to lack sensorimotor input (Gruber & McDonald 2012; Draganski et al. 2008), but this may have been overlooked in previous studies due to their significantly weaker projection density when compared to the canonical dorsolateral sensorimotor projections. This putative subregion defines an area of the striatum capable of directly integrating limbic and sensorimotor information. The novel limbic subdivisions described above each possess interesting and experimentally tractable characteristics to be tested in future studies of striatal function.

Striatal input integration

The differences observed between dense and diffuse corticostriatal projection distributions, and more specifically, differences in their convergence patterns (see **Chapter 2: Fig. 5a-c**) could represent an important mechanism for integrating multimodal information in the basal ganglia. There is mounting behavioral and anatomical evidence that the primary information streams remain segregated as they pass from the striatum, through the output

nuclei, and back to the thalamus (Draganski et al. 2008; Haber & Calzavara 2009; Parent & Hazrati 1995). However, this strict parallel organization does not allow for the coordination of separate modalities during complex behavior (Joel & Weiner 1994; Steiner & Tseng 2010). Mechanisms for cross-talk between adjacent functional domains have been proposed to occur through dopaminergic feedback (Joel & Weiner 2000; Haber et al. 2000), lateral inhibition in the striatum(Gruber & McDonald 2012), and cortical feedback to the thalamic targets of basal ganglia output (Haber & Calzavara 2009), but none of these theories posits direct communication across modalities in the striatum. A projection organization consisting of targeted dense projections and convergent diffuse projections would be a simple mechanism for parallel loops that communicate across domains, with the dense projections being poised to drive striatal activity in a region specific manner, and the diffuse projections acting as the integrators. Furthermore, the stark distinction between the striatal volume with high diffuse corticostriatal projection convergence and low convergence (Fig. 5a), could constitute a functionally relevant subdivision, worthy of further investigation.

Future directions

Paradoxically, the simplicity of this evolutionarily ancient brain region has made it complicated to study. The striatum receives inputs from nearly every other brain region, but consists almost entirely of two identical, spherical cell types and does not have a laminar organization or any other cytoarchitectural demarcations to identify distinct subregions (Steiner & Tseng 2010). This lack of functional subdivisions has made it nearly impossibly to systematically perturb the system in order to investigate how striatal inputs mediate behavior. Most of our current knowledge about the role of the basal ganglia in behavior comes from dysfunctions associated with diseases of the basal ganglia (see **Introduction: Fig. 3b**) (Yin 2014; DeLong & Wichmann 2007; Albin et al. 1989). Although the involvement of the basal ganglia in Parkinson's disease, Tourette's syndrome, and addiction, point to a broad role in decision making (DeLong & Wichmann 2007; Plotkin & Surmeier 2015; Redgrave et al. 2010; Tomasi & Volkow 2013; Volkow et al. 2012), understanding these dysfunctions does

not necessarily shed light on the structural and computational organization required to produce these decisions normally. We must understand how the system functions in healthy individuals if we intend to restore function in the diseased brain.

The fundamental motivation for generating this dataset was to create a comprehensive and objective map of all excitatory inputs to the mouse striatum, which could act as a guide for the investigations of basal ganglia function. With this goal in mind, the data generation and data analyses were designed to facilitate follow-up physiological and behavioral studies. Injection coordinates required to target these areas are provided for the thalamocortical and corticostriatal datasets. Using this information, analogous injections can be performed with virus that instead expresses optogenetic activators or inhibitors (Boyden et al. 2005), allowing for the manipulation of targeted inputs. These studies will be invaluable for investigations of input integration in the striatum, both within and across modalities. The ability to target specific inputs, particularly in the newly identified functional domains, will also facilitate future studies into the behavioral functions of these circuits.

References

Albin, R.L., Young, A.B. & Penney, J.B., 1989. The functional anatomy of basal ganglia disorders. *Trends in neurosciences*, 12(10), pp.366–375.

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, pp.357–381.

Asanuma, H. & Fernandez, J.J., 1974. Organization of projection from the thalamic relay nuclei to the motor cortex in the cat. *Brain research*, 71(2-3), pp.515–522.

Aschauer, D.F., Kreuz, S. & Rumpel, S., 2013. Analysis of transduction efficiency, tropism and axonal transport of AAV serotypes 1, 2, 5, 6, 8 and 9 in the mouse brain. *PloS one*, 8(9), p.e76310.

Berendse, H.W. & Groenewegen, H.J., 1990. Organization of the thalamostriatal projections in the rat, with special emphasis on the ventral striatum. *The Journal of comparative neurology*, 299(2), pp.187–228.

Berendse, H.W. & Groenewegen, H.J., 1991. Restricted cortical termination fields of the midline and intralaminar thalamic nuclei in the rat. *NSC*, 42(1), pp.73–102.

Berendse, H.W., Galis-de Graaf, Y. & Groenewegen, H.J., 1992. Topographical organization and relationship with ventral striatal compartments of prefrontal corticostriatal projections in the rat. *The Journal of comparative neurology*, 316(3), pp.314–347.

Berman, A.L. & Jones, E.G., 1982. *The Thalamus and Basal Telencephalon of the Cat*,

Bird, C.M. & Burgess, N., 2008. The hippocampus and memory: insights from spatial processing. *Nature reviews. Neuroscience*, 9(3), pp.182–194.

Bohland, J.W. et al., 2009. A Proposal for a Coordinated Effort for the Determination of Brainwide Neuroanatomical Connectivity in Model Organisms at a Mesoscopic Scale O. Sporns, ed. *PLoS Computational Biology*, 5(3), p.e1000334.

Bota, M., Dong, H.W. & Swanson, L.W., 2012. Combining collation and annotation efforts toward completion of the rat and mouse connectomes in BAMS. *Frontiers in neuroinformatics*.

Boyden, E.S. et al., 2005. Millisecond-timescale, genetically targeted optical control of neural activity. *Nature Neuroscience*, 8(9), pp.1263–1268.

Briggs, F., 2010. Organizing principles of cortical layer 6. *Frontiers in neural circuits*, 4, p.3.

CASEY, B.J., Getz, S. & Galvan, A., 2008. The adolescent brain. *Developmental Review*, 28(1), pp.62–77.

Cowan, R.L. & Powell, E.W., 1956. A study of thalamo-striate relations in the monkey. *Brain : a journal of neurology*, 79(2), pp.364–390.

Cowan, R.L. & Wilson, C.J., 1994. Spontaneous firing patterns and axonal projections of single corticostriatal neurons in the rat medial agranular cortex. *Journal of Neurophysiology*, 71(1), pp.17–32.

Cui, G. et al., 2013. Concurrent activation of striatal direct and indirect pathways during action initiation. *Nature*.

Dantzker, J.L. & Callaway, E.M., 2000. Laminar sources of synaptic input to cortical inhibitory interneurons and pyramidal neurons. *Nature Neuroscience*, 3(7), pp.701–707.

DeLong, M.R. & Wichmann, T., 2007. Circuits and circuit disorders of the basal ganglia. *Archives of neurology*, 64(1), pp.20–24.

Deniau, J.M. & Chevalier, G., 1992. The lamellar organization of the rat substantia nigra pars reticulata: distribution of projection neurons. *NSC*, 46(2), pp.361–377.

Desbois, C. & Villanueva, L., 2001. The organization of lateral ventromedial thalamic connections in the rat: a link for the distribution of nociceptive signals to widespread cortical regions. *NSC*, 102(4), pp.885–898.

Dice, L.R., 1945. Convergent inputs from electrically and topographically distinct orexin cells to locus coeruleus and ventral tegmental area. *Ecology*.

Ding, J.B. et al., 2010. Thalamic gating of corticostriatal signaling by cholinergic interneurons. *Neuron*, 67(2), pp.294–307.

Doig, N.M., Moss, J. & Bolam, J.P., 2010. Cortical and thalamic innervation of direct and indirect pathway medium-sized spiny neurons in mouse striatum. *Journal of Neuroscience*, 30(44), pp.14610–14618.

Draganski, B. et al., 2008. Evidence for segregated and integrative connectivity patterns in the human Basal Ganglia. *Journal of Neuroscience*, 28(28), pp.7143–7152.

Erro, M. et al., 2001. Striatal input from the ventrobasal complex of the rat thalamus. *Histochemistry and cell biology*, 115(6), pp.447–454.

Euston, D.R., Gruber, A.J. & McNaughton, B.L., 2012. The role of medial prefrontal cortex in memory and decision making. *Neuron*, 76(6), pp.1057–1070.

Evarts, E.V. & Thach, W.T., 1969. Motor mechanisms of the CNS: cerebrocerebellar interrelations. *Annual review of physiology*, 31, pp.451–498.

Gertler, T.S., Chan, C.S. & Surmeier, D.J., 2008. Dichotomous anatomical properties of adult striatal medium spiny neurons. *Journal of Neuroscience*, 28(43), pp.10814–10824.

Groenewegen, H.J., 1988. Organization of the afferent connections of the mediodorsal thalamic nucleus in the rat, related to the mediodorsal-prefrontal topography. *NSC*, 24(2), pp.379–431.

Groenewegen, H.J. & Berendse, H.W., 1994. The specificity of the "nonspecific" midline and intralaminar thalamic nuclei. *Trends in neurosciences*, 17(2), pp.52–57.

Groenewegen, H.J. et al., 1999. Convergence and segregation of ventral striatal inputs and outputs. *Annals of the New York Academy of Sciences*, 877, pp.49–63.

Groenewegen, H.J. et al., 1987. Organization of the projections from the subiculum to the ventral striatum in the rat. A study using anterograde transport of Phaseolus vulgaris leucoagglutinin. *NSC*, 23(1), pp.103–120.

Groenewegen, H.J. et al., 1990. The anatomical relationship of the prefrontal cortex with the striatopallidal system, the thalamus and the amygdala: evidence for a parallel organization. *Progress in brain research*, 85, pp.95–116– discussion 116–8.

Gruber, A.J. & McDonald, R.J., 2012. Context, emotion, and the strategic pursuit of goals: interactions among multiple brain systems controlling motivated behavior. *Frontiers in behavioral neuroscience*, 6, pp.1–26.

Guo, Q. et al., 2015. Whole-brain mapping of inputs to projection neurons and cholinergic interneurons in the dorsal striatum. *PloS one*, 10(4), p.e0123381.

Haber, S. & McFarland, N.R., 2001. The place of the thalamus in frontal cortical-basal ganglia circuits. *The Neuroscientist : a review journal bringing neurobiology, neurology and psychiatry*, 7(4), pp.315–324.

Haber, S.N., 2006. Reward-Related Cortical Inputs Define a Large Striatal Region in Primates That Interface with Associative Cortical Connections, Providing a Substrate for Incentive-Based Learning. *Journal of Neuroscience*, 26(32), pp.8368–8376.

Haber, S.N. & Calzavara, R., 2009. The cortico-basal ganglia integrative network: the role of the thalamus. *Brain research bulletin*, 78(2-3), pp.69–74.

Haber, S.N. & Knutson, B., 2010. The reward circuit: linking primate anatomy and human imaging. *Neuropsychopharmacology*, 35(1), pp.4–26.

Haber, S.N. & McFarland, N.R., 1999. The concept of the ventral striatum in nonhuman primates. *Annals of the New York Academy of Sciences*, 877, pp.33–48.

Haber, S.N., Fudge, J.L. & McFarland, N.R., 2000. Striatonigrostriatal pathways in primates form an ascending spiral from the shell to the dorsolateral striatum. *Journal of Neuroscience*, 20(6), pp.2369–2382.

Haque, T. et al., 2010. Thalamic afferent and efferent connectivity to cerebral cortical areas with direct projections to identified subgroups of trigeminal premotoneurons in the rat. *Brain research*, 1346(C), pp.69–82.

Harris, J.A., Oh, S.W. & Zeng, H., 2012. Adeno−Associated Viral Vectors for Anterograde Axonal Tracing with Fluorescent Proteins in Nontransgenic and Cre Driver Mice. *Current protocols in* ….

Haycock, J.W. & Bro, S., 1975. Corpus striatum (Translation of S. Ramón y Cajal). translated from Corps Strié, chapter 23, in "Histologie du systéme nerveux de l'homme et des vertébrés" 1911. *Behavioral biology*, 14(3), pp.387–402.

Herkenham, M., 1979. The afferent and efferent connections of the ventromedial thalamic nucleus in the rat. *The Journal of comparative neurology*, 183(3), pp.487–517.

Hintiryan, H. et al., 2012. Comprehensive connectivity of the mouse main olfactory bulb: analysis and online digital atlas. *Frontiers in Neuroanatomy*, 6, pp.1–16.

Hooks, B.M. et al., 2011. Laminar analysis of excitatory local circuits in vibrissal motor and sensory cortical areas. *PLoS biology*, 9(1), p.e1000572.

Hooks, B.M. et al., 2013. Organization of cortical and thalamic input to pyramidal neurons in mouse motor cortex. *Journal of Neuroscience*, 33(2), pp.748–760.

Hoover, W.B. & Vertes, R.P., 2007. Anatomical analysis of afferent projections to the medial prefrontal cortex in the rat. *Brain structure & function*, 212(2), pp.149–179.

Huerta-Ocampo, I., Mena-Segovia, J. & Bolam, J.P., 2013. Convergence of cortical and thalamic input to direct and indirect pathway medium spiny neurons in the striatum. *Brain structure & function*.

Hunnicutt, B.J. et al., 2014. A comprehensive thalamocortical projection map at the mesoscopic level. *Nature Neuroscience*, 17(9), pp.1276–1285.

Ilinsky, I.A., Jouandet, M.L. & Goldman-Rakic, P.S., 1985. Organization of the nigrothalamocortical system in the rhesus monkey. *The Journal of comparative neurology*, 236(3), pp.315–330.

Isomura, Y. et al., 2013. Reward-modulated motor information in identified striatum neurons. *Journal of Neuroscience*, 33(25), pp.10209–10220.

Jin, X., Tecuapetla, F. & Costa, R.M., 2014. Basal ganglia subcircuits distinctively encode the parsing and concatenation of action sequences. *Nature Neuroscience*, 17(3), pp.423–430.

Joel, D. & Weiner, I., 2000. The connections of the dopaminergic system with the striatum in rats and primates: an analysis with respect to the functional and compartmental organization of the striatum. *NSC*, 96(3), pp.451–474.

Joel, D. & Weiner, I., 1994. The organization of the basal ganglia-thalamocortical circuits: open interconnected rather than closed segregated. *NSC*, 63(2), pp.363–379.

Jones, E.G., 2007. *The Thalamus 2 Volume Set*, Cambridge University Press.

Kamishina, H. et al., 2009. Cortical connections of the rat lateral posterior thalamic nucleus. *Brain research*, 1264, pp.39–56.

Katz, L.C. & Callaway, E.M., 1992. Development of local circuits in mammalian visual cortex. *Annual Review of Neuroscience*, 15, pp.31–56.

Kawaguchi, Y., Wilson, C.J. & Emson, P.C., 1989. Intracellular recording of identified neostriatal patch and matrix spiny cells in a slice preparation preserving cortical inputs. *Journal of Neurophysiology*, 62(5), pp.1052–1068.

Kawaguchi, Y., Wilson, C.J. & Emson, P.C., 1990. Projection subtypes of rat neostriatal matrix cells revealed by intracellular injection of biocytin. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 10(10), pp.3421–3438.

Khibnik, L.A., Tritsch, N.X. & Sabatini, B.L., 2014. A Direct Projection from Mouse Primary Visual Cortex to Dorsomedial Striatum. *PloS one*, 9(8), p.e104501.

Kincaid, A.E., Zheng, T. & Wilson, C.J., 1998. Connectivity and convergence of single corticostriatal axons. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 18(12), pp.4722–4731.

Knapska, E., Marcias, M. & Mikosz, M., 2012. Functional anatomy of neural circuits regulating fear and extinction

. *Proceedings of the National Academy of Sciences*, 42, pp.1–6. Available at: http://www.pnas.org/content/109/42/17093.full.pdf+html?sid=da60374a-6818-46aa-8585-6003f58e54c8.

Kolb, B., 1977. Studies on the caudate-putamen and the dorsomedial thalamic nucleus of the rat: implications for mammalian frontal-lobe functions. *Physiology & Behavior*, 18(2), pp.237–244.

Kolmac, C.I. & Mitrofanis, J., 1997. Organisation of the reticular thalamic projection to the intralaminar and midline nuclei in rats. *The Journal of comparative neurology*, 377(2), pp. 165–178.

Kravitz, A.V. et al., 2010. Regulation of parkinsonian motor behaviours by optogenetic control of basal ganglia circuitry. *Nature*, 466(7306), pp.622–626.

Kravitz, A.V., Tye, L.D. & Kreitzer, A.C., 2012. Distinct roles for direct and indirect pathway striatal neurons in reinforcement. *Nature Neuroscience*.

Kreitzer, A.C., 2009. Physiology and pharmacology of striatal neurons. *Annual Review of Neuroscience*, 32, pp.127–147.

Kress, G.J. et al., 2013. Convergent cortical innervation of striatal projection neurons. *Nature Neuroscience*, 16(6), pp.665–667.

Kuan, L. et al., 2015. Neuroinformatics of the Allen Mouse Brain Connectivity Atlas. *METHODS*, pp.1–14.

Ledoux, J.E., Sakaguchi, A. & Reis, D.J., 1984. Subcortical efferent projections of the medial geniculate nucleus mediate emotional responses conditioned to acoustic stimuli. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 4(3), pp.683–698.

Lei, W. et al., 2004. Evidence for differential cortical input to direct pathway versus indirect pathway striatal projection neurons in rats. *Journal of Neuroscience*, 24(38), pp.8289–8299.

Lein, E.S. et al., 2007. Genome-wide atlas of gene expression in the adult mouse brain. *Nature*, 445(7124), pp.168–176.

Li, S. & Kirouac, G.J., 2008. Projections from the paraventricular nucleus of the thalamus to the forebrain, with special emphasis on the extended amygdala. *The Journal of comparative neurology*, 506(2), pp.263–287.

Liang, L., DeLong, M.R. & Papa, S.M., 2008. Inversion of dopamine responses in striatal medium spiny neurons and involuntary movements. *Journal of Neuroscience*, 28(30), pp. 7537–7547.

Lichtman, J.W. & Denk, W., 2011. The big and the small: challenges of imaging the brain's circuits. *Science (New York, NY)*, 334(6056), pp.618–623.

Luo, L., Callaway, E.M. & Svoboda, K., 2008. Genetic dissection of neural circuits. *Neuron*, 57(5), pp.634–660.

Mailly, P. et al., 2013. The rat prefrontostriatal system analyzed in 3D: evidence for multiple interacting functional units. *Journal of Neuroscience*, 33(13), pp.5718–5727.

Malach, R. & Graybiel, A.M., 1986. Mosaic architecture of the somatic sensory-recipient sector of the cat's striatum. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 6(12), pp.3436–3458.

Mao, T. et al., 2011. Long-range neuronal circuits underlying the interaction between sensory and motor cortex. *Neuron*, 72(1), pp.111–123.

McFarland, N.R. & Haber, S.N., 2000. Convergent inputs from thalamic motor nuclei and frontal cortical areas to the dorsal striatum in the primate. *Journal of Neuroscience*, 20(10), pp.3798–3813.

McFarland, N.R. & Haber, S.N., 2002. Thalamic relay nuclei of the basal ganglia form both reciprocal and nonreciprocal cortical connections, linking multiple frontal cortical areas. *Journal of Neuroscience*, 22(18), pp.8117–8132.

McFarland, N.R. et al., 2009. Comparison of transduction efficiency of recombinant AAV serotypes 1, 2, 5, and 8 in the rat nigrostriatal system. *Journal of Neurochemistry*, 109(3), pp.838–845.

McGeorge, A.J. & Faull, R.L., 1989. The organization of the projection from the cerebral cortex to the striatum in the rat. *NSC*, 29(3), pp.503–537.

McHaffie, J.G. et al., 2005. Subcortical loops through the basal ganglia. *Trends in neurosciences*, 28(8), pp.401–407.

Mitchell, B.D. & Cauller, L.J., 2001. Corticocortical and thalamocortical projections to layer I of the frontal neocortex in rats. *Brain research*, 921(1-2), pp.68–77.

Mitra, P.P., 2014. The circuit architecture of whole brains at the mesoscopic scale. *Neuron*, 83(6), pp.1273–1283.

Nambu, A., 2011. Somatotopic organization of the primate Basal Ganglia. *Frontiers in Neuroanatomy*, 5, p.26.

Nambu, A. et al., 1996. Dual somatotopical representations in the primate subthalamic nucleus: evidence for ordered but reversed body-map transformations from the primary motor cortex and the supplementary motor area. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 16(8), pp.2671–2683.

Nauta, W.J., 1971. The problem of the frontal lobe: a reinterpretation. *Journal of psychiatric research*, 8(3), pp.167–187.

Nelson, A.B. & Kreitzer, A.C., 2014. Reassessing models of Basal Ganglia function and dysfunction. *Annual Review of Neuroscience*, 37, pp.117–135.

Novejarque, A. et al., 2011. Amygdaloid projections to the ventral striatum in mice: direct and indirect chemosensory inputs to the brain reward system. *Frontiers in Neuroanatomy*, 5, p.54.

Oh, S.W. et al., 2014. A mesoscale connectome of the mouse brain. *Nature*, 508(7495), pp. 207–214.

Oorschot, D.E., 1996. Total number of neurons in the neostriatal, pallidal, subthalamic, and substantia nigral nuclei of the rat basal ganglia: a stereological study using the cavalieri and optical disector methods. *The Journal of comparative neurology*, 366(4), pp. 580–599.

Otsu, N., 1975. A threshold selection method from gray-level histograms. *Automatica*.

Pabba, M., 2013. Evolutionary development of the amygdaloid complex. *Frontiers in Neuroanatomy*, 7, p.27.

Pan, W.X., Mao, T. & Dudman, J.T., 2010. Inputs to the dorsal striatum of the mouse reflect the parallel circuit architecture of the forebrain. *Frontiers in Neuroanatomy*, 4, p.147.

Parent, A. & Hazrati, L.N., 1995. Functional anatomy of the basal ganglia. I. The cortico-basal ganglia-thalamo-cortical loop. *Brain research. Brain research reviews*, 20(1), pp.91–127.

Paxinos, G., 2004. The mouse brain in stereotaxic coordinates.

Petreanu, L. et al., 2009. The subcellular organization of neocortical excitatory connections. *Nature*, 457(7233), pp.1142–1145. Available at: http://www.nature.com/nature/journal/vaop/ncurrent/full/nature07709.html.

Plenz, D., 2003. When inhibition goes incognito: feedback interaction between spiny projection neurons in striatal function. *Trends in neurosciences*, 26(8), pp.436–443.

Plotkin, J.L. & Surmeier, D.J., 2015. Corticostriatal synaptic adaptations in Huntington's disease. *Current Opinion in Neurobiology*, 33C, pp.53–62.

Ragan, T. et al., 2012. Serial two-photon tomography for automated ex vivo mouse brain imaging. *Nature Methods*, 9(3), pp.255–258.

Ray, J.P. & Price, J.L., 1992. The organization of the thalamocortical connections of the mediodorsal thalamic nucleus in the rat, related to the ventral forebrain-prefrontal cortex topography. *The Journal of comparative neurology*, 323(2), pp.167–197.

Redgrave, P. et al., 2010. Goal-directed and habitual control in the basal ganglia: implications for Parkinson's disease. pp.1–13.

Reig, R. & Silberberg, G., 2014. Multisensory integration in the mouse striatum. *Neuron*, 83(5), pp.1200–1212.

Reiner, A. et al., 2010. Corticostriatal projection neurons - dichotomous types and dichotomous functions. *Frontiers in Neuroanatomy*, 4, p.142.

Rubio-Garrido, P. et al., 2009. Thalamic Input to Distal Apical Dendrites in Neocortical Layer 1 Is Massive and Highly Convergent. *Cerebral Cortex*, 19(10), pp.2380–2395.

Seger, C.A. & Miller, E.K., 2010. Category learning in the brain. *Annual Review of Neuroscience*, 33, pp.203–219.

Shepherd, G.M.G. et al., 2005. Geometric and functional organization of cortical circuits. *Nature Neuroscience*, 8(6), pp.782–790.

Sherman, S.M. & Guillery, R.W., 2009. *Exploring the Thalamus and Its Role in Cortical Function*, MIT Press (MA).

Sherman, S.M. & Guillery, R.W., 2001. Exploring the thalamus.

Shibata, H., 1993. Efferent projections from the anterior thalamic nuclei to the cingulate cortex in the rat. *The Journal of comparative neurology*, 330(4), pp.533–542.

Smith, J.B., Mowery, T.M. & Alloway, K.D., 2012. Thalamic POm projections to the dorsolateral striatum of rats: potential pathway for mediating stimulus-response associations for sensorimotor habits. *Journal of Neurophysiology*, 108(1), pp.160–174.

Smith, Y. et al., 2004. The thalamostriatal system: a highly specific network of the basal ganglia circuitry. *Trends in neurosciences*, 27(9), pp.520–527.

Smith, Y., Galvan, A., et al., 2014a. The thalamostriatal system in normal and diseased states. *Frontiers in systems neuroscience*, 8, p.5.

Smith, Y., Wichmann, T. & DeLong, M.R., 2014b. Corticostriatal and mesocortical dopamine systems: do species differences matter? *Nature reviews. Neuroscience*, 15(1), p.63.

Spear, L.P., 2000. The adolescent brain and age-related behavioral manifestations. *Neuroscience and biobehavioral reviews*, 24(4), pp.417–463.

Steiner, H. & Tseng, K.Y., 2010. *Handbook of Basal Ganglia Structure and Function*, Academic Press.

Steriade, M. & Llinás, R.R., 1988. The functional states of the thalamus and the associated neuronal interplay. *Physiological Reviews*, 68(3), pp.649–742.

Stuber, G.D. et al., 2011. Excitatory transmission from the amygdala to nucleus accumbens facilitates reward seeking. *Nature*, 475(7356), pp.377–380.

Sturman, D.A. & Moghaddam, B., 2011. Reduced neuronal inhibition and coordination of adolescent prefrontal cortex during motivated behavior. *Journal of Neuroscience*, 31(4), pp.1471–1478.

Surmeier, D.J. et al., 2007. D1 and D2 dopamine-receptor modulation of striatal glutamatergic signaling in striatal medium spiny neurons. *Trends in neurosciences*, 30(5), pp.228–235.

Thompson, R.H. & Swanson, L.W., 2010. Hypothesis-driven structural connectivity analysis supports network over hierarchical model of brain architecture. *Proceedings of the National Academy of Sciences*, 107(34), pp.15235–15239.

Tomasi, D. & Volkow, N.D., 2013. Striatocortical pathway dysfunction in addiction and obesity: differences and similarities. *Critical Reviews in Biochemistry and Molecular Biology*, 48(1), pp.1–19.

Tritsch, N.X. & Sabatini, B.L., 2012. Dopaminergic modulation of synaptic transmission in cortex and striatum. *Neuron*, 76(1), pp.33–50.

Van der Werf, Y.D., Witter, M.P. & Groenewegen, H.J., 2002. The intralaminar and midline nuclei of the thalamus. Anatomical and functional evidence for participation in processes of arousal and awareness. *Brain research. Brain research reviews*, 39(2-3), pp. 107–140.

Van Eden, C.G. & Uylings, H.B., 1985. Postnatal volumetric development of the prefrontal cortex in the rat. *The Journal of comparative neurology*, 241(3), pp.268–274.

van Groen, T. & Wyss, J.M., 1995. Projections from the anterodorsal and anteroventral nucleus of the thalamus to the limbic cortex in the rat. *The Journal of comparative neurology*, 358(4), pp.584–604.

Veening, J.G., Cornelissen, F.M. & Lieven, P.A., 1980. The topical organization of the afferents to the caudatoputamen of the rat. A horseradish peroxidase study. *NSC*, 5(7), pp.1253–1268.

Vertes, R.P., Hoover, W.B. & Rodriguez, J.J., 2012. Projections of the central medial nucleus of the thalamus in the rat: node in cortical, striatal and limbic forebrain circuitry. *Neuroscience*, 219, pp.120–136.

Volkow, N.D. et al., 2012. Addiction Circuitry in the Human Brain *. *Annual Review of Pharmacology and Toxicology*, 52(1), pp.321–336.

Voorn, P. et al., 2004. Putting a spin on the dorsal-ventral divide of the striatum. *Trends in neurosciences*, 27(8), pp.468–474.

Walker, A.E., 1938. *The Primate Thalamus*, University of Chicago Press.

Wall, N.R. et al., 2013. Differential innervation of direct- and indirect-pathway striatal projection neurons. *Neuron*, 79(2), pp.347–360.

Wang, Q. & Burkhalter, A., 2007. Area map of mouse visual cortex. *The Journal of comparative neurology*, 502(3), pp.339–357.

Wang, Q. et al., 2014. Systematic comparison of adeno-associated virus and biotinylated dextran amine reveals equivalent sensitivity between tracers and novel projection targets in the mouse brain. *The Journal of comparative neurology*.

Watson, C., Paxinos, G. & Puelles, L., 2012. *The Mouse Nervous System* 1st ed. C. Watson, G. Paxinos, & L. Puelles, eds., London, UK: Academic Press.

Webster, K.E., 1961. Cortico-striate interrelations in the albino rat. *Journal of anatomy*, 95, pp. 532–544.

Weiler, N. et al., 2008. Top-down laminar organization of the excitatory network in motor cortex. *Nature Neuroscience*, 11(3), pp.360–366.

Weinberger, D.R., 1993. A connectionist approach to the prefrontal cortex. *The Journal of neuropsychiatry and clinical neurosciences*, 5(3), pp.241–253.

Willuhn, I., Sun, W. & Steiner, H., 2003. Topography of cocaine−induced gene regulation in the rat striatum: relationship to cortical inputs and role of behavioural context. *The European journal of neuroscience*, 17(5), pp.1053–1066.

Wilson, C.J., 2000. *Brain dynamics and the striatal complex* R. Miller & J. Wickens, eds., CRC.

Wilson, C.J., 1987. Morphology and synaptic connections of crossed corticostriatal neurons in the rat. *The Journal of comparative neurology*, 263(4), pp.567–580.

Wise, R.A., 2009. Roles for nigrostriatal--not just mesocorticolimbic--dopamine in reward and addiction. *Trends in neurosciences*, 32(10), pp.517–524.

Yang, J.M. et al., 2012. Postnatal Development of 2 Microcircuits Involving Fast-Spiking Interneurons in the Mouse Prefrontal Cortex. *Cerebral Cortex*.

Yin, H.H., 2014. How Basal Ganglia Outputs Generate Behavior. *Advances in Neuroscience*, 2014(3), pp.1–28.

Yin, H.H. & Knowlton, B.J., 2006. The role of the basal ganglia in habit formation. *Nature reviews. Neuroscience*, 7(6), pp.464–476.

Zhang, Z.W. & Deschênes, M., 1998. Projections to layer VI of the posteromedial barrel field in the rat: a reappraisal of the role of corticothalamic pathways. *Cerebral cortex (New York, N.Y. : 1991)*, 8(5), pp.428–436.

Zingg, B. et al., 2014. Neural Networks of the Mouse Neocortex. *Cell*, 156(5), pp.1096–1111.