A phylogenetically-conserved axis of thalamocortical connectivity in the human brain

Stuart Oldham[1] & Gareth Ball[1,2]
1. Developmental Imaging, Murdoch Children's Research Institute, The Royal Children's Hospital Melbourne
2. Department of Paediatrics, University of Melbourne, Melbourne

Corresponding author
Dr Stuart Oldham
Developmental Imaging,
Murdoch Children's Research Institute
Parkville 3052 VIC
Australia
stuart.oldham@mcri.edu.au

Abstract

The thalamus enables key sensory, motor, emotive, and cognitive processes via connections to the cortex. These projection patterns are traditionally considered to originate from discrete thalamic nuclei, however recent work showing gradients of molecular and connectivity features in the thalamus suggests the organisation of thalamocortical connections occurs along a continuous dimension. By performing a joint decomposition of densely sampled gene expression and non-invasive diffusion tractography in the adult human thalamus, we define a principal axis of genetic and connectomic variation along a medial-lateral thalamic axis. Projections along this axis correspond to an anterior-posterior cortical pattern and are aligned with electrophysiological properties of the cortex. The medial-lateral axis demonstrates phylogenetic conservation, reflects transitions in neuronal subtypes, and shows associations with neurodevelopment and common brain disorders. This study provides evidence for a supra-nuclear axis of thalamocortical organisation characterised by a graded transition in molecular properties and anatomical connectivity.

Introduction

The thalamus is central to the structure and function of the mammalian brain. Enabled by widespread connections to the cortex, basal ganglia and the peripheral nervous system, the thalamus is engaged in a range of functions from sensory processing and motor control, through to attention and memory[1–5]. Traditionally, this breadth of functionality has been assigned to the diverse nuclear structure of the thalamus, with around 50-60 distinct nuclei gathered into 4-6 functional groups and defined on the basis of cytoarchitecture and patterns of anatomical connections[6–8].

While efforts to understand the organisation and function of the thalamus have often focused on the properties of its nuclei, a consensus on a precise nomenclature and nuclear grouping remain elusive[9]. Further, while the specificity of projections from certain thalamic nuclei to distinct cortical targets is clearly evident, not all thalamic connections are circumscribed by nuclear boundaries[10,11]. Neurons within thalamic nuclei may project to multiple cortical regions and their patterning of cortical projections can overlap significantly[9,10,12]. In contrast to nuclear divisions, recent evidence from single-cell RNA sequencing (RNA-seq) studies has refocused attention on the significant cellular diversity present within the thalamus as a potential substrate for functional diversity[13,14]. While distinct subclasses of neurons in the thalamus have long been recognised[10,15,16], the degree of cellular heterogeneity recently revealed through genetic profiling of both cortical and subcortical structures is a striking and potentially defining feature of mammalian brain organisation[17,18].

At the cellular level, neuronal subtypes can form subunits supporting a diversity of functions, the spatial arrangement of which are dictated by organisational molecular gradients[14,19–21]. In the thalamus, as with other cortical and subcortical structures, early circuit formation is scaffolded by the differential areal patterning of morphogenetic gradients during development[22–24]. These early developmental sequences are reflected by concerted variation of structural and functional properties along spatial axes in the adult cortex and subcortex[25–32] Indeed, studies have found evidence for gene expression gradients and variations in cytoarchitecture both across and within thalamic nuclei[13,14,20,21,33].

In the mouse, Phillips et al. observed that thalamic nuclei can be arranged along an axis of gene expression running along a medial to lateral direction[14]. The position of neuronal subtypes along this axis was associated with continuous variations in gene expression accompanied by differences in cortical projections, axonal morphology, and laminar targets, as well as electrophysiological properties. This work highlights how delineation of spatial gradients can give key insight into underlying principles of thalamic organisation and function[13,14,20,22]. Preliminary evidence suggests the same medial-lateral axis of thalamic organisation extends to the human thalamus[14], however it has not been extensively characterised.

Non-invasive neuroimaging is a key tool for studying thalamocortical organisation in humans[9,34–37]. While MRI can be used to target specific subcortical nuclei with a high degree of[38,39] accuracy, attempts to resolve the nuclear structure of the thalamus based on patterns of cortical connectivity or correlated BOLD activation often fail to align to previously defined cytoarchitectural boundaries[40–42]. In contrast to delineating discrete brain regions, recent efforts in neuroimaging have focused on the definition of continuous axes of spatial variation based on measures of brain microstructure, anatomy, and/or function[25–27]. This approach has proven insightful, identifying continuous and overlapping patterns of anatomical variation that converge with patterns of gene expression and ontogenetic timing in the developing brain[43–45]; align with hierarchies of cortical function[29,46], and are disrupted in neurodevelopmental disorders and psychopathology[29,30,47,48]. As such, smooth transitions in anatomy, cytoarchitecture and function supported by spatially-varying gradients of gene expression may be considered a hallmark organisational motif of the mammalian brain[25,28,32,49].

In this study, we present evidence for a supra-nuclear axis of thalamocortical organisation characterised by a graded transition in molecular properties and anatomical connectivity in humans. By performing a joint decomposition of densely sampled gene expression and non-invasive diffusion tractography in the adult thalamus, we define a principal axis of genetic variation along the medial-lateral thalamic axis that corresponds to anterior-posterior patterns of thalamocortical connectivity and electrophysiological properties of the cortex. Using a large, single-cell RNA-seq survey of the brain, we demonstrate how continuous transitions in neuronal subtypes along the medial-lateral axis reflect the developmental origins of excitatory projection and inhibitory interneurons in the thalamus. We also report associations between axis-enriched thalamic genes, and genes associated with neurodevelopment and common brain disorders, and test the phylogenetic conservation of the principal axis through comparison to patterns of neuronal tracing in the adult mouse. Taken together, this study highlights a novel organisational axis in the thalamus, that exists over and above classical nuclear boundaries, is conserved across species and associated with distinct anatomical, functional, molecular, and developmental properties.

Results

A principal axis of thalamocortical connectivity is present in humans and conserved across species

We hypothesised that a primary organisational axis of thalamocortical connectivity exists in the human brain that spans specific nuclear boundaries and is demarcated by patterns of thalamic gene expression. We performed an unsupervised joint decomposition of post-mortem gene expression[50] (*n* = 3702 samples across 6 donor brains[51]; 2,233 genes enriched in brain tissue[52,53]) and non-invasive estimates of cortical connectivity to 250 cortical regions (averaged over *n* = 74 healthy adults aged 22 to 36) from 921 thalamic seed points (**Fig. 1**). This procedure resulted in a set of overlapping yet orthogonal components in the thalamus that sum together to reconstruct the full data matrix. Each component is represented by a set of PC scores, one per thalamic seed, defining the dominant axes of variation in both gene expression and cortical connectivity across the thalamus and a set of PC loadings that capture how strongly connections to particular cortical regions and expression of particular genes contribute to the component.

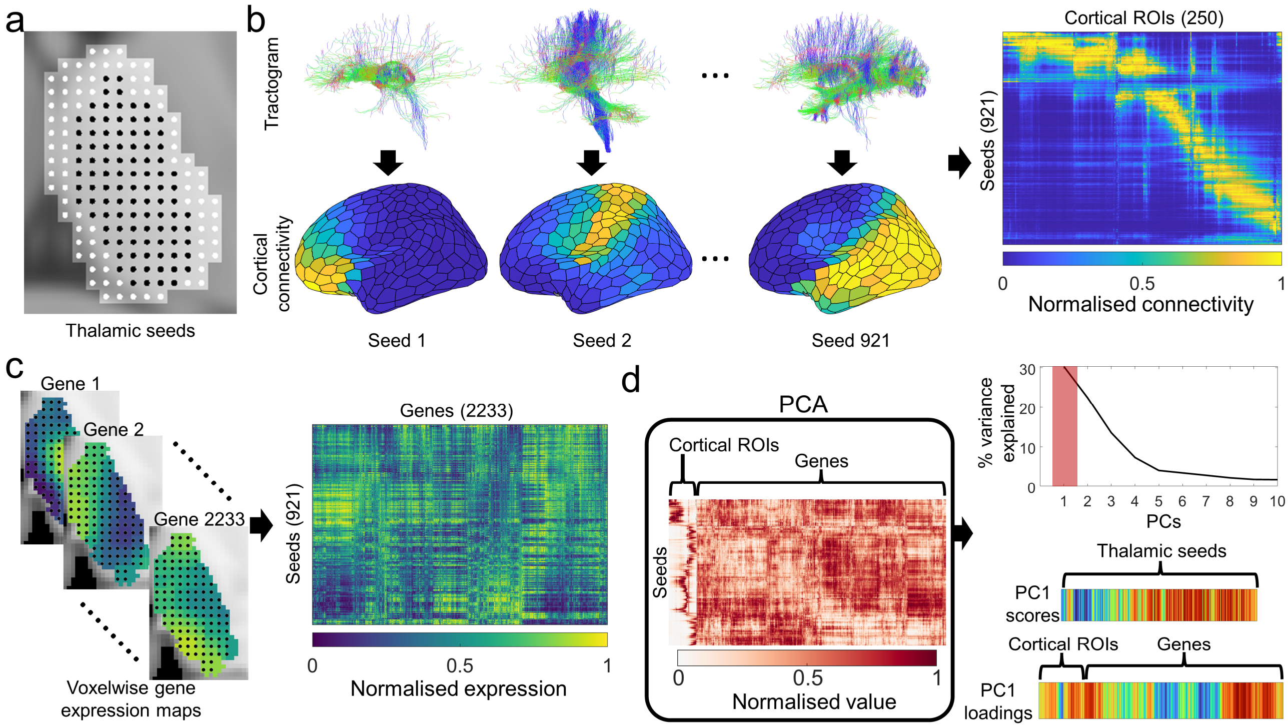

Figure 1. Schematic of methodology. a, Defining thalamic seeds. Throughout the thalamic volume (white shaded area), a set of seeds 1.75mm apart are defined. Only those which were consistently localised (see **Methods** for details) across participants were used as seed points (black: consistent seeds; white: inconsistent seeds). **b,** Thalamic seed connectivity. Probabilistic tractography was conducted from each seed to 250 left hemisphere cortical targets based on a random parcellation. Connectivity between seeds and cortical regions was averaged across participants to produce a 921 thalamic seed -by-250 cortical target matrix of thalamocortical connectivity. Connectivity to cortical regions was scaled to the unit interval using a sigmoid transformation. **c,** Assigning transcriptomic data to thalamic seeds. Voxelwise estimates of post-mortem gene expression for 2,233 genes with differential expression in brain tissue were extracted for the thalamus. For each gene, each seed point is assigned the expression value of the voxel it is located within to produce a 921-by-2233

seed-by-gene matrix. As above, each gene's expression levels were normalised to the unit interval according to a scaled sigmoid. **d,** Joint decomposition. The seed-by-cortical connectivity and seed-by-gene matrices were concatenated and decomposed into a set of orthogonal factors by Principal Component Analysis (PCA). From the resulting principal components (PCs), the first PC (PC1) explained 30.2% of the variance in the concatenated data matrix. For each PC, the scores, one per thalamic seed, describe the representation of each component in the thalamus and the loadings, describe the contribution to the PC of connectivity strength and gene expression level for each of the cortical regions and genes, respectively.

The principal component (PC1) represents the primary source of variation in gene expression and thalamocortical connectivity across thalamic seeds, accounting for 30.2% of variance in total (**Fig. 1d**). The spatial projection of PC1 scores varied primarily along the medial-lateral axis of the brain (**Fig. 2a**) with the PC1 score of each thalamic seed encoding position along the x-axis of MNI standard space ($r = -0.86$) more closely than the other Cartesian axes (anterior-posterior: $r = -0.44$; inferior-superior: $r = 0.38$; **Fig. S1**). The second and third principal components (PC2 and PC3) explained less variance overall (22.4% and 13.5%) and were aligned along superior-inferior and anterior-posterior axes, respectively (**Fig. S1**; **Fig. S2d-i**). We conducted a series of sensitivity analyses and found that this characterisation of the principal axis was largely unaffected when applying alternative, nonlinear decomposition techniques[26,54], using a different subset of genes[14], or performing the decomposition using only gene expression or connectivity data (**Fig. S3-S4**).

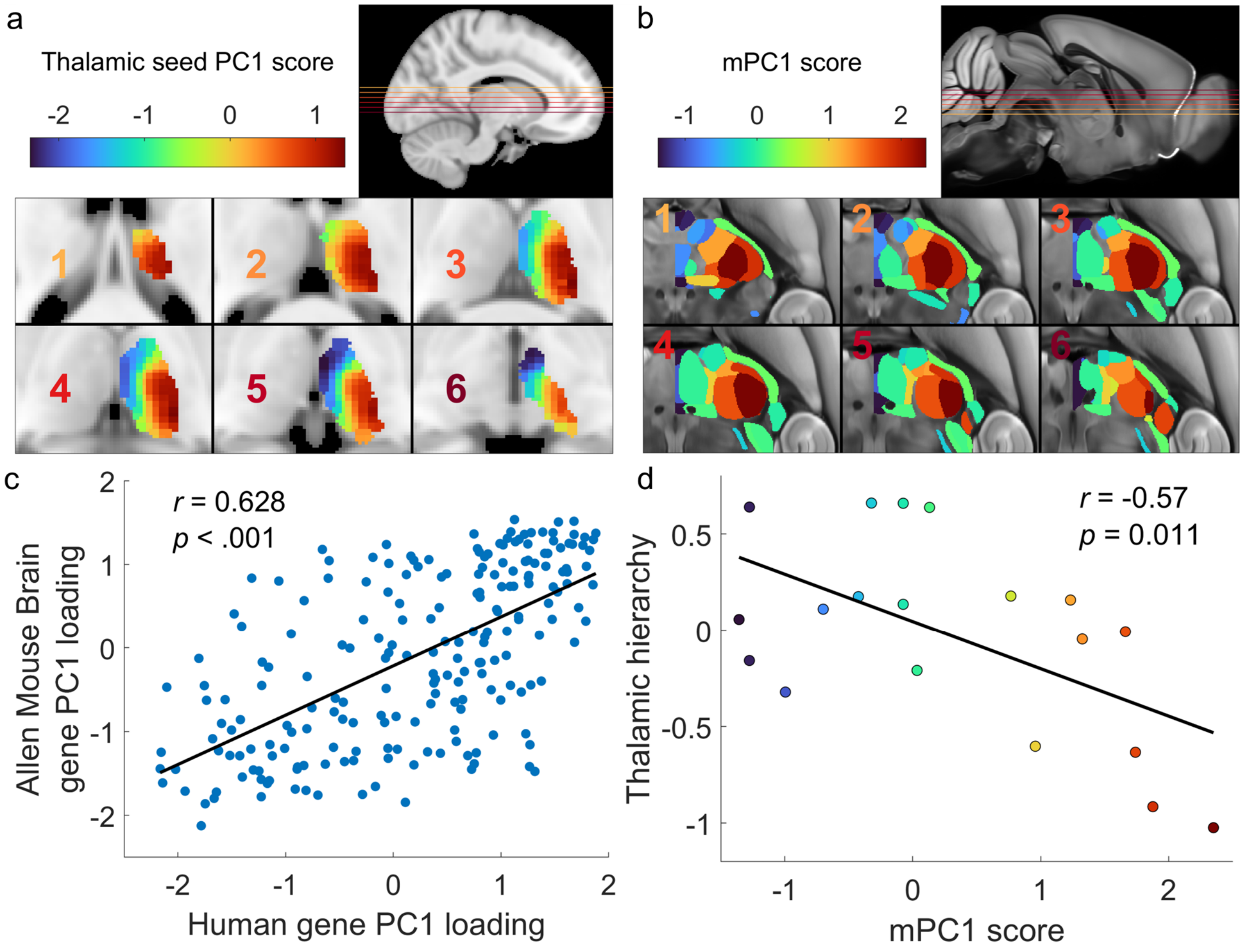

Figure 2. The principal thalamic axis runs medial-laterally in the human and mouse. a, Projection of PC1 scores of the human data onto thalamic voxels. PC1 scores for each seed are projected onto the closest voxels in the thalamic mask, overlaid on six axial sections (inset). PC1 scores vary along a medial-lateral direction. **b**, Projection of PC1 scores of the Allen Mouse Brain Atlas (AMBA) data (mPC1) onto respective thalamic nuclei. Note that while axonal-tracing data was obtained in the right hemisphere, we project the PC1 scores onto the left to enable straightforward comparison with the human data. **c**, Scatter plot showing the relationship between PC1 loadings for homologous genes in the human and AMBA datasets. **d**, Relationship between mouse thalamic nuclei PC1 (mPC1) score and a measure of hierarchical organisation[55].

Graded variation in patterns of connectivity and gene expression along the medial-lateral thalamic axis is consistent with previous findings in the mouse[14]. To replicate this finding more directly, we repeated our analysis using the same framework applied to anatomical tract tracing data and transcriptomic data from the Allen Mouse Brain Atlas (AMBA; **Fig. 2b**)[56–59]. This revealed a primary component (mPC1) that explained 30.6% of variance in total and was also largely aligned along a medial-lateral orientation. As in the human, mPC1 scores correlated with the x-coordinate in standard CCFv3 space ($r = 0.58$; $p < .001$), though with some divergence in nuclei of the ventral thalamus (reticular nucleus, ventral division of the lateral geniculate)[7] and in the medial geniculate body (**Fig. S5**). PC1 loadings and mPC1 loadings of homologous genes common to both datasets (n = 212) were also highly correlated ($r = 0.63$, $p < .001$; **Fig. 2c**). We confirmed these observations in an independent mouse dataset[14], finding highly correlated PC1 loadings of homologous mouse and human genes in the thalamus (**Fig. S6**).

Through detailed mapping of feedforward and feedback thalamocortical connections in the mouse, a recent study defined a detailed model of organisational hierarchy across thalamic nuclei[55]. Using these data, we find that medial-lateral position, defined by mPC1 scores, was significantly correlated with position in the hierarchy of feedforward–feedback interareal laminar projections[55] (**Fig. 2d**), providing evidence that medial-lateral orientation reflects key organisational properties of the thalamus.

Cortical patterning of thalamic connectivity follows key patterns of functional organisation

The thalamus has topographical projections to the cerebral cortex[7]; therefore, we reasoned that connections seeded along the medial-lateral gradient of the thalamus would vary along a corresponding spatial gradient in the cortex. To test this, we plotted the human PC1 loadings for each cortical region onto the cortical surface, revealing an anterior-posterior gradient (**Fig. 4a**). Anterior cortical regions were negatively loaded, displaying preferential connectivity to medial thalamic regions, posterior regions were positively loaded with preferential connectivity to lateral regions (**Fig. 4b**). Using the mPC1 loadings from the AMBA mouse data, we observed a similar projection pattern which varied from sensorimotor regions, to visual, and then to frontal/lateral cortex (**Fig. 4c**), following a hierarchical gradient[28,55].

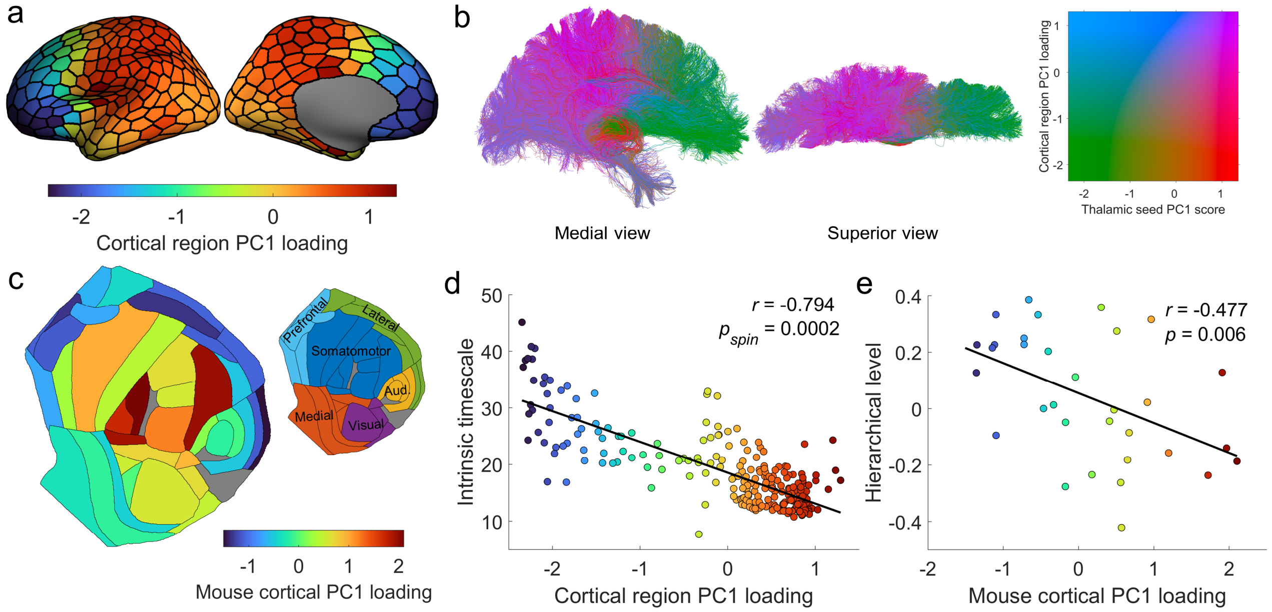

Figure 3. Cortical patterning of thalamic connections corresponds to key cortical gradients. a, The PC1 loadings for cortical regions are shown projected onto the cortical surface, revealing an anterior-posterior gradient of thalamocortical connectivity. **b**, Representative tractogram where each streamline is coloured according to its PC1 score and loading of the corresponding seed point and cortical region it traverses between. **c**, Projection of PC1 loadings of the Allen Mouse Brain Atlas (AMBA) data onto respective cortical regions, displayed as a flat map. The smaller flat map plot indicates major cortical divisions (prefrontal, lateral, somatomotor, visual, medial, and auditory)[55]. Grey regions indicate cortical

areas which no gene expression and/or connectivity data was available **d**, Correlation between cortical region PC1 loadings (for the human data) and MEG intrinsic timescales ($p_{spin} < .05$). Points are coloured according to their PC1 loading. **e**, Correlation between mouse cortical PC1 loadings and hierarchical level. Points are coloured according to their PC1 loading.

Recently, large-scale spatial gradients have been used to frame variation of a range of microstructural, connectomic, and functional properties across the cortex and subcortex[25–27]. Given the prominence of features which follow an anterior-posterior spatial arrangement in the cortex[29,46,60], we expected that PC1 loadings would vary in parallel with other cortical properties. To test this, we compared PC1 loadings to 72 cortical feature maps from the *neuromaps* toolbox[60]. Statistical significance was established using spin-tests (see **Methods**) to ensure observed correlations were not induced by low-order spatial autocorrelations[61]. Significant associations between cortical PC1 loadings were observed with several markers of functional organisation, including the primary functional gradient[26], sensorimotor-association axis[30], electrophysiological properties, and several neurotransmitter gradients (**Fig. S7**). The strongest association was observed with neuronal intrinsic timescales defined using MEG ($r = -0.79$, $p_{spin} = .0002$; **Fig 4d**), with projections from the medial thalamus preferentially connected to regions with slower cortical intrinsic timescales compared to the lateral thalamic regions. Performed a similar analysis in the mouse, we observed that cortical mPC1 loadings were correlated with multiple properties that together characterise cortical hierarchical organisation (**Fig. 4e**; **Fig. S8**).

Cellular and molecular composition varies as a function of the medial-lateral thalamic axis

Neuronal cell types are distributed non-uniformly in the thalamus[62,63] with distinct cellular subtypes differentiated by graded variations in gene expression both across and within discrete thalamic nuclei[13,14,20]. Therefore, we hypothesised that cellular composition, evidenced by differential gene expression, would vary systematically along the medial-lateral axis. Using genes with the largest positive and negative PC1 loadings (**Supplementary Table 1**; $n = 100$ each; after accounting for spatial autocorrelation across thalamic seeds using spin tests; see **Methods**[64]), we queried a comprehensive Drop-seq analysis of 89,027 cells from the adult mouse thalamus[33] to test if genes with medial-lateral patterns of expression were enriched for different thalamic cell type markers.

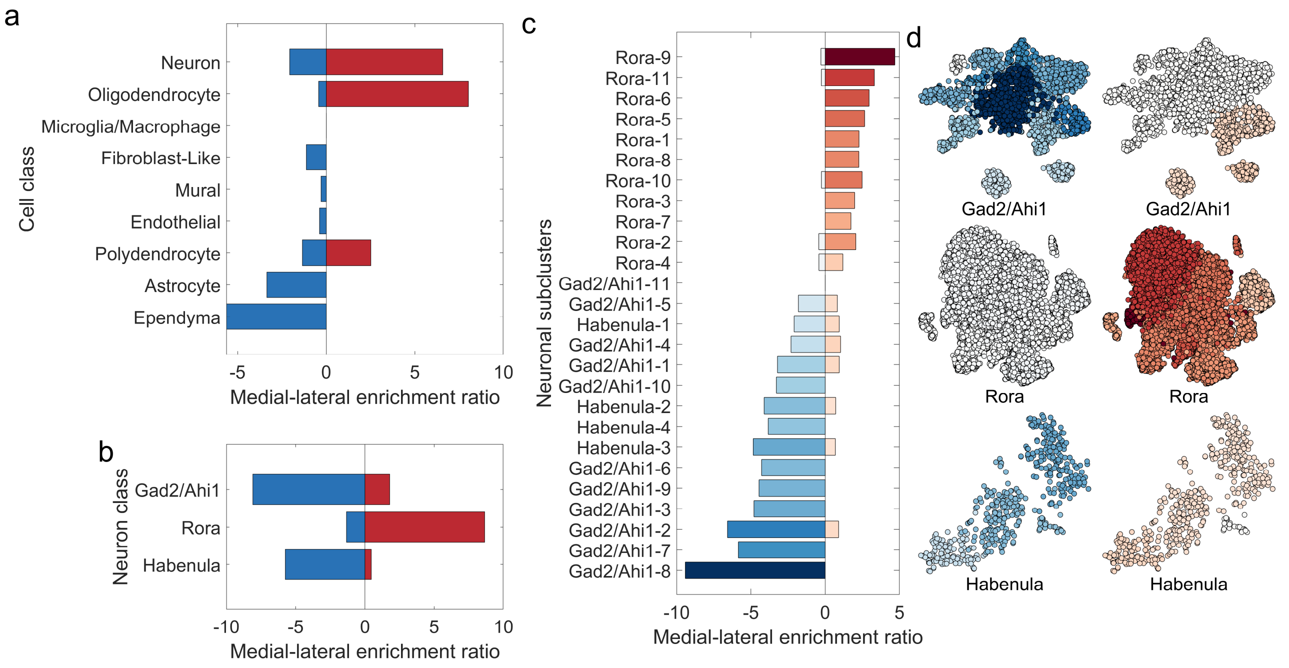

Figure 4. PC1 reflects graded distinctions in cytoarchitecture. a, Enrichment ratio for medial- and lateral-genes expressed by cell class. The medial-lateral enrichment ratio is positive when that cell class is enriched for lateral-genes and negative when enriched for medial-genes. **b**, Enrichment ratio for medial- and lateral-genes expressed by each neuron subtype[33]. **c**, Enrichment ratio for medial- and lateral-genes expressed by each neuronal subcluster. Subclusters were defined by clustering cells within each neuron subtype by their pattern of gene expression. In the bar plot, neuronal subclusters are ordered by their summed medial-lateral enrichment ratio. **d**, Enrichment ratios projected onto t-SNE plots for the different neuron type subclusters. Cells belonging to each subcluster are coloured according to their enrichment value as indicated in **c** for medial- (left) and lateral-genes (right).

Both lateral- (n = 100) and medial-genes (n = 100) were significantly enriched for neuron class markers (lateral-genes: enrichment = 6.57, $p_{FDR} = 3.43 \times 10^{-28}$; medial-genes: enrichment = 2.06, $p_{FDR} = .02$); while lateral-genes were additionally enriched for oligodendrocyte markers (enrichment = 8.04, $p_{FDR} = 3.41 \times 10^{-11}$), and medial-genes for glial (astrocyte: enrichment = 3.35, $p_{FDR} = .007$; ependymal: enrichment = 5.63, $p_{FDR} = 7.62 \times 10^{-11}$) markers (**Fig. 3a; Supplementary Table 2**). To further differentiate between medial and lateral gene sets, we focused on neuronal classes and tested enrichment of three previously identified neuronal subtypes[33]: *Rora* (excitatory neurons expressing *Slc17a6*), *Gad2/Ahi1* (neurons largely expressing inhibitory markers *Gad1* and *Gad2*), and *Habenula* (cholinergic and glutamatergic neurons in the habenula). Lateral- and medial-genes showed differential patterns of enrichment, with genes expressed by *Rora* subtypes and enriched in lateral genes (enrichment = 8.64, $p_{FDR} = 3.11 \times 10^{-26}$) encoding glutamate receptors (*Grid1*/*Grm1*), voltage-gated channels (*Scn1b*/*Kcna2*/*Scn8a*) and calcium transporters (*Slc24a2*) with medial-genes where enriched for *Habenula* (enrichment = 5.74, $p_{FDR} = 8.33 \times 10^{-6}$) and *Gad2/Ahi1* (enrichment = 8.09, $p_{FDR} = 1.28 \times 10^{-7}$) makers, including those for forebrain interneurons (*Dlx1*[65]) and GABAergic neurons (*Cnr1*) (**Fig. 3b; Supplementary Table 3**).

Each neuronal subtype comprised several closely related clusters revealed through two-dimensional embedding of gene expression in each subtype[33]. Though distinct in identity, the close proximity of neuronal subtypes, and clusters within subtypes, suggested overall similar patterns of gene expression across cells with graded transitions across borders between both clusters and subtypes[13,33]. Based on this, we reasoned that a graded pattern of gene enrichment would be evident across adjacent clusters within each neuronal subtype, rather than discrete clusters with or without enrichment. We repeated the overrepresentation analysis focusing on neuronal cell subclusters (**Supplementary Table 2**). We found the degree of medial and lateral enrichment in transitioned smoothly across subclusters (**Fig. 3c**), thereby encoding position of individual cell types across the thalamus. This smooth transition across subclusters is evident when projecting enrichment ratios onto embedded cells (**Fig. 3d**). and supports previous work demonstrating that cell types near the border of adjacent thalamic nuclei display similar patterns gene expression to each other and do not belong to separate populations[14].

Genes expressed along the medial-lateral axis are associated with thalamic development and disease

Forebrain development is founded on early canonical molecular gradients during gestation, which are reflected in the differential timing of key developmental processes across structures[24,32]. Spatial gradients in adult gene expression data vary along developmental axes[29] and recent evidence has shown transcriptional profiles of neurons retain a persistent marker of their developmental origins[66]. Based on this evidence, we examined if the medial-lateral axis captured differences in developmental timing across the thalamus. Using a database of post-mortem RNA-seq data acquired from multiple brain regions across the human lifespan[67], we identified a set of genes differentially expressed over nine developmental windows (**Methods**). We observed divergence across the medial-lateral axis, with medial genes enriched during both the prenatal and postnatal developmental periods, while lateral genes are largely enriched postnatally (**Fig. S9**). Differential expression of several prenatally-enriched genes was identified across the medial-lateral axis (**Supplementary Table 4**, with medially-enrichment of genes involved in interneuron migration from the forebrain ganglionic eminence (*DLX1*[65] and *GLRA2*[68]) and lateral enrichment of gene involved in thalamocortical outgrowth (*DSCAML1*[69], *SLIT1*[70], and *FZD3*[71,72]), and functional maturation of synapses (*ADCY1*[73]).

Multiple neurodevelopmental and neurodegenerative disorders are associated with thalamic dysfunction[74,75], we examined if genes expressed along the medial-lateral axis were also differentially associated with disease. We found that medial-genes were significantly enriched ($p_{FDR} < .05$) for genes associated with alcohol-related disorders (alcoholism, alcoholic intoxication, alcohol withdrawal seizures), eating-related disorders (eating disorders, hyperphagia, and anorexia), and psychosis-related disorders (psychotic disorders, schizoaffective disorder, and schizophrenia; **Supplementary Table 5**). Genes encoding for gamma-aminobutyric acid (*GABRB1*/*GABRG1*/*GABRG3*), cannabinoid (*CNR1*), serotonin (*HTR2R*/*HTR2C*), and neuropeptide (*NPY2R*/*NPY5R*) receptors, in addition to oxytocin (*OXT*), were commonly associated with the disorders showing enrichment for medial-genes. For lateral-genes, these showed enrichment for epilepsy-related (generalised and myoclonic), bipolar disorder, schizophrenia and other generalised movement/neurological conditions (**Supplementary Table 5**), with genes for sodium/potassium transport and channels (*ATP1A3*/*KCNA1*/*KCNA2*/*KCNC1*/*SCN1b*/*SCN8A*), glutamate receptors (*GRM1*/*GRM4*), and parvalbumin (*PVALB*) implicated across disorders.

Discussion

The thalamus and its reciprocal connections are crucial to shaping whole brain neural activity, yet the underlying organisational principles of these connections are not well understood[7,13,14]. In this study, we have defined a principal organisational axis of the adult human thalamus, characterised by a graded transition in aggregated gene expression, that is conserved across species and aligned to functional cortical organisation via a medial-lateral to anterior-posterior mapping.

Previous work identified an axis of gene expression running in a medial-lateral orientation at the level of individual thalamic nuclei in the mouse[14]. We extend these observations, demonstrating continuous variation in structural and molecular properties along this spatial axis. By decomposing bulk tissue microarray data into constituent cell types using markers from a comprehensive single-cell database[33], we identify potential differences in cell class distribution across the principal thalamic axis. Lateral areas, adjacent to the white matter, were enriched for oligodendrocyte markers, while medial areas, adjacent to ventricles and developmental sources of radiating glial cells, were enriched for other glial markers. Similar divergence was observed across neuronal classes suggesting that composition of neuronal populations also varies along the thalamic axis, with lateral areas preferentially enriched for excitatory markers, and medial areas for inhibitory ones.

Traditionally, the identity and functional role of thalamic nuclei have been in part determined through examination of afferent connections and cortical projection targets[1,7]. We examined patterns of cortical connectivity with the thalamus using non-invasive diffusion tractography, validating our findings using anatomical tracing data in the mouse. Cortical projections from the medial-lateral thalamic axis mapped to an anterior-posterior cortical gradient, aligning with cytoarchitectural and function markers of cortical hierarchy[25,29,32,46]. We find this cortical projection gradient mirrored variation in neuronal intrinsic timescales, fundamental patterns of oscillatory activity and neural dynamics in the cortex[76]. This cortical patterning is likely conserved across species. The gradient of thalamic projections in the mouse cortex was also related to measures of cortical hierarchies[28,55] and the thalamic medial-lateral to cortical anterior-posterior mapping has been observed in tract-tracing experiments in rodents and primates[77,78]. In the mouse, electrophysiological properties of neurons including action potential threshold/width, ion channel and receptor profiles vary as a function of position along the primary axis[14]. We additionally find that medial-lateral position encodes flow of hierarchical information across the mouse connectome[55], further highlighting potential functional importance of the thalamic axis. Overall, our results indicate a close alignment of thalamocortical connectivity with cortical organisational principles and suggest a conserved thalamus-to cortical mapping across species.

Prior studies have observed that measures of functional connectivity between the thalamus and cortex in humans form large-scale gradients[37] that partially align with spatial distribution of core/matrix neuronal cell types and are correlated with timescales of neural activity in the cortex[79]. Electrophysiological signals in the cortex are shaped by a balance of excitatory and inhibitory inputs, which the thalamus is involved in the regulation of[80]. Systematic variation in thalamic electrophysiological properties also influence functional activity in the cortex[11,14]. Taken alongside previous findings, our work supports the notion that intrinsic cortical dynamics, at least in part, are constrained by a principal mapping of thalamic projections to the cortex, potentially underwritten by distributional variations of specific neuronal populations[13] along the medial-lateral axis.

Expression gradients across and within nuclei may offer a more parsimonious account of function attributed to thalamic nuclei[13]. However, while prominent, several aspects of thalamic organisation remain unaccounted for by the medial-lateral axis. Different projection systems of the thalamus have distinct patterns of gene expression[14] and cytoarchitectonic boundaries between nuclei are well-documented[7,13]. As an example, in the mouse data, we found the reticular nucleus and geniculate bodies diverged from the primary axis. The reticular nucleus and ventral division of the lateral geniculate form part of the ventral thalamus, or prethalamus, a structure distinct from the rest of the (dorsal) thalamus in terms of development, connectivity and function[7,22,81]. In our human data, however, due to the position at the extreme lateral border of the thalamus, these structures were only sparsely sampled and were unable to be fully characterised.

We note that, as with cortical arealisation, the presence of supra-areal organisational gradients does not preclude functional localisation to discrete regions[25]. Localisation of functions to discrete nuclei is well supported by an extensive body of evidence. Specific thalamic nuclei are directly involved in specific cognitive[8], sensory[82], and motor[83] activity and display preferential connectivity to focal cortical targets[7]. However, these focal projections vary in a spatially systematic way which we and others have observed[14,35,37] and the notion of discrete thalamic units cannot account for all functional heterogeneity observed in the thalamus. Reconciling how continuous and discrete patterns of thalamic organisation interact to support whole brain dynamics, and what specific functional roles each may have, is a point of keen interest going forward.

We found genes expressed along the medial-lateral gradient were associated with numerous disorders including eating-related, alcohol-related, psychosis, neurodevelopmental, and epilepsy related disorders, reflecting the diversity of thalamic function in human behaviour and neurophysiology. The specific genes enriched for these disorders included those which encode for GABA, cannabinoid, serotonin, glutamate, and neuropeptide receptors as well as sodium/potassium transport. Eating, alcohol-related, psychosis, neurodevelopmental and epilepsy disorders have been associated with both thalamic and neurotransmitter abnormalities[74,75,84–88]. As the thalamus has a key role in neuromodulation[89], disruptions to thalamic and/or neurotransmitter systems may be reflected by altered neurotransmitter signalling along the medial-lateral axis. Furthermore, several genes were also differentially enriched across early developmental windows, and temporal sequences of neurogenesis occur along a medial-lateral direction in the thalamus[90,91]. Thus, disruptions along the medial-lateral axis during development could cause substantial alterations in connectivity. Our study provides a potential framework for interrogating fundamental properties of thalamocortical organisation across species. Examining how principles of thalamocortical connectivity emerge and are potentially disturbed, remains an underexamined area of research which addressing would promote new insights into the course of healthy and abnormal development.

In summary, we find a principal axis of transcriptomic and structural connectivity in the human thalamus that is situated along a medial-lateral axis and conserved across species. Position along this axis encodes functional hierarchy, variations in cellular composition, and is mapped to key properties of cortical function, revealing a simple organisational principle of thalamocortical connectivity.

Methods

Human neuroimaging data

Minimally-processed structural and diffusion data from 100 unrelated participants were acquired from the Human Connectome Project[92,93]. Of these 76 (46 females, age mean ± SD: 28.39 ± 3.95) were retained after quality control (see below). Data were acquired on a customized Siemens 3T Connectome Skyra scanner at Washington University in St Louis, Missouri, USA. Structural T1-weighted data were acquired with 0.7 mm^3 voxels, TR = 2400 ms, TE = 2.14 ms, FOV of 224×224 mm. Diffusion data was acquired using a multi-shell protocol for the DWI (1.25 mm^3 voxel size, TR = 5520 ms, TE = 89.5 ms, FOV of 210×180 mm, 270 directions with b = 1000, 2000, 3000 s/mm^2 (90 per b value, and 18 b = 0 volumes). Image pre-processing is described in detail elsewhere[92–94]. Briefly, diffusion data were corrected for EPI susceptibility and signal outliers, eddy-current-induced distortions, slice dropouts, gradient-non-linearities and subject motion[94]. T1-weighted data were corrected for gradient and readout distortions prior to being processed with FreeSurfer. The left hemisphere was divided into 250 approximately equally sized regions for further analysis.

Human gene expression data

Gene expression levels were assigned to thalamic seeds using high-resolution maps of estimated gene expression in the thalamus[50]. Briefly, Gaussian Process Regression was used to estimate the spatial dependence of gene expression between neighbouring locations in the cortical and subcortical structures based on post mortem microarray data from the Allen Human Brain Atlas (AHBA)[51]. The AHBA contains 3,702 microarray samples of 58,692 probes across six brains. Microarray pre-processing is detailed elsewhere[50,51]. Using spatial models of gene expression estimated at discrete locations in volumetric space, voxelwise expression maps of 18,836 genes were generated[50]. In the present study, we focused on a previously identified list of 2,413 genes with differential expression in the human brain[52,53], of which 2,233 were present in the high-resolution dataset (**Fig. 1c**) which were downloaded from an online repository (http://www.meduniwien.ac.at/neuroimaging/mRNA.html). As few samples in the AHBA were obtained in the right hemisphere, we elicited to only use data from the left hemisphere.

Thalamic seed definition and quality control

To measure variation in connectivity and gene expression across the thalamus, we defined a set of seeds. Seeds are defined within an MNI152 thalamic mask (1,811 seeds total; 1.75mm apart) and are registered to each participant. This number of seeds was selected to maintain a balance between dense spatial coverage of the thalamus and minimising computational burden. Quality control was performed by defining a binary vector indicating if a seed was inside or outside) each participants own thalamic mask (defined using tissue segmentation with FreeSurfer[95]). Participants whose vector was not highly correlated with others (mean $r < 0.7$) were excluded (*n* = 24), as these participants likely have an inconsistent spatial distribution of seeds to other participants. Seeds present across 85% of the remaining participants were retained (*n* = 921; **Fig. 1a**).

Thalamic seed connectivity

To estimate connectivity between the thalamus and cortex, for the remaining 76 participants, 5,000 streamlines were generated from each of the 921 spatially-consistent thalamic seeds using the second-order integration over fibre orientation distributions tractography algorithm[96,97] (1.25mm step size, 45° maximum angle, 0.05 fibre orientation distribution cut-off) with Anatomically Constrained Tractography[98] applied (**Fig. 1b**). Streamlines were assigned to the nearest left hemisphere cortical targets within a 5mm radius of their endpoint. Cortical target regions were based on a random parcellation where each parcel had approximately equal surface area (**Fig. 1b**). Connectivity between seeds and cortical regions was averaged across participants to produce a 921-by-250 matrix of thalamocortical connectivity. Expression values were sampled from the voxelwise gene expression maps by assigning seeds the expression values of the voxel they resided in, producing a 921-by-2233 matrix of thalamic seed gene expression.

To ensure comparison between genes and cortical connection values, these data were normalised using a scaled sigmoid transformation to the interval [0,1]. This first involved applying a sigmoidal transformation to the raw data:

$$S(x) = \frac{1}{1 + \exp\left(-\frac{x - \langle x \rangle}{\sigma_x}\right)}, \tag{1}$$

where $S(x)$ is the normalised value of a gene/connection, x is the raw value, $\langle x \rangle$ is the mean and σ_x is the standard deviation of the values of that gene/connection across thalamic seeds. Following the sigmoidal transform, each gene or cortical connection was linearly scaled to the unit interval. This transformation was used to reduce the impact of outliers in the data[56,99] (**Fig. 1d**). The same concatenation and normalisation procedure was applied to the mouse axonal tracing and thalamic gene expression data.

Mouse data

We used gene expression[59] and anatomical connectivity[57] data for the mouse from the Allen Mouse Brain Atlas. Data processing has been detailed elsewhere[56,58]. Expression data were extracted for the 213-region mouse parcellation of Oh et al[57]. Of the 19,417 genes for which expression had been measured across all 213 regions, we extract those for which there was a) complete expression data for all 35 thalamic nuclei, and b) were part of the 500 most differentially expressed genes across mouse thalamic nuclei as identified by Phillips et al[14]. This resulted in a total of 447 genes being retained, producing a 35×447 thalamic-by-gene expression matrix.

Connectivity data was derived from the Allen Mouse Brain Connectivity Atlas, which consists of 469 anterograde viral microinjection experiments conducted on C57BL/6J male mice at age P56[57]. We extracted connections from 35 thalamic nuclei to 38 cortical targets, producing a 35x38 axonal thalamic-by-cortical region connectivity matrix for use in the decomposition. For visualisation, Common Coordinate Framework version 3 atlas for the Allen Mouse Brain Atlas was used[59,100].

Joint decomposition

We decomposed the concatenated 921-by-2483 ($n \times m$) data matrix, M, into a set of orthogonal components using Principal Component Analysis (PCA) via Singular Value Decomposition (SVD):

$$M = USV^T, \tag{2}$$

where, US is a 921 × k matrix represents the Principal Component (PC) scores, one per thalamic seed for each of k components; and V is a $2483 \times k$ matrix representing the PC loadings, or coefficients, that denote the contributions of each cortical region's (normalised) thalamic connectivity or each gene's (normalised) thalamic expression to each component (M was centred prior to SVD/PCA). The decomposition is normally truncated to $k < min(n, m)$ and the variance explained by each component, λ_k, is given by its singular values, s_k:

$$\lambda_k = \frac{s_k^2}{n-1} \tag{3}$$

This approach reduces the dimensionality of the data by finding components (axes which maximise the variance explained in the data) which are orthogonal to each other. We repeated this analysis in the mouse using the corresponding, concatenated 35×482 data matrix.

Sensitivity analyses

To assess the potential impact of the decomposition approach and the data used on our observations, we performed a series of sensitivity analyses:

- Performed the joint decomposition using a nonlinear alternative to PCA (diffusion embedding)[26,54] to test if our observations were limited by using a linear model.
- Performed the PCA using only human homologs of genes previously identified as differentially expressed along the medial-lateral axis in the mouse[14] to test if the results were consistent when a more restricted gene-set was used.
- Performed the PCA on the connectivity and gene expression data matrices separately to test if the decomposition was driven by connectivity or gene expression.
- Performed an initial PCA on the concatenated top ten components from the separate PCA of the connectivity and gene expression data to ensure each data type contributed the same number of features to the decomposition.
- Calculated the cosine affinity matrix for the connectivity and gene expression matrices separately, and performed diffusion embedding on the averaged affinity matrices so each data type contributed equally to the non-linear decomposition.

Maps of human cortical properties

We obtained maps of multiple cortical features from the *neuromaps* toolbox[60]. This dataset consists of 72 high-resolution maps of different cortical properties including measures of tissue microstructure, gene expression, metabolism, neurotransmitter receptor distribution, electrophysiology, and cortical expansion amongst others. All maps were transformed to the fsaverage 164k template. Vertex-wise properties were averaged within each parcel to get a single value for each region of the parcellation.

Maps of mouse cortical properties

We used maps of nine different cortical properties of the mouse brain which have previously been found to be reflective of mouse hierarchical organisation[28]. These included the ratio of T1-weighted to T2-weighted (T1w:T2w) images, mean cell density for parvalbumin-containing (PV) cells[101], cytoarchitectonic classification based on regional eulamination[102], cortical gene expression[59], intracortical axonal connectivity[57], and inferred hierarchy from feedforward–feedback laminar projection patterns between cortical and thalamic regions[55]. Further details of how these data were preprocessed are available elsewhere[28,55].

Additional gene expression datasets

Using the DropViz (http://dropviz.org/) database, we downloaded lists of differentially expressed (DE) genes for nine thalamic cell types: neurons, ependyma, astrocyte, polydendrocyte, endothelial, mural, fibroblast-like, microglia/macrophage, and oligodendrocyte[33]. DE genes were defined using the following parameters: minimum fold ratio of three; maximum p-value exponent of -50; minimum log expression in target of one; and a maximum mean expression in comparison of six). Where cell types constituted more than one cluster (neuron, *n* = 3; oligodendrocyte, *n* = 2), they were combined into a single 'target' cluster and compared to all other cell types. Neuron genes consisted of three clusters, *Rora*, *Gad2/Ahi1*, and *Habenula*, each of which had several subclusters (*n* = 11, 11, and 4 respectively). For enrichment of subclusters, each subcluster was compared to a reference set which consisted of all other subclusters in the different subtypes, and the background set was defined as all genes expressed by neurons.

Gene homologs were identified using Ensembl Biomart (https://www.ensembl.org/index.html; reference genomes: human GRCh38.p13; mouse GRCm39). Gene lists were filtered to only include: genes with identified mouse-human homologs; genes with protein expression in the human thalamus (list from https://www.proteinatlas.org/); and genes with expression in the top 75% based on aggregated unique molecular identifier across cell types in the mouse thalamus, ensuring that corresponding genes are expressed in both the mouse and human thalamus.

Overrepresentation analysis

To assess enrichment of genes across different gene sets, we used the hypergeometric statistic:

$$p = 1 - \sum_{i=0}^{x} \frac{\binom{K}{i}\binom{N-K}{n-i}}{\binom{N}{n}}, \tag{4}$$

where p is the probability of finding x or more genes from a gene list K in a set of N randomly selected genes drawn from a background set N. Enrichment was expressed as the ratio of the top n genes present in the gene list of interest, compared to the proportion in the full background.

To identify genes associated with disorders, we used the WEB-based GEne SeT AnaLysis Toolkit[103], using protein encoding genes from the human genome as a reference set and the DisGeNET[104], GLAD4U[105], and OMIM[106] databases as functional sets. A threshold of $p_{FDR} < .05$ was used to identify significant categories.

Spatial nulls

As cortical and subcortical features exhibit spatial autocorrelations, we implemented spatial-autocorrelation-preserving permutation tests to assess statistical significance (commonly known as "spin-tests") between pairs of brain maps and to correct for smooth spatial autocorrelation in thalamic maps of gene expression. We conducted two separate spin-tests, one for the cortex and the other for the thalamus.

The cortical spin-test was used to find which of the cortical maps from the *neuromaps* toolbox was significantly correlated with the PC1 loadings in the cortex. First the centroid of each cortical region on the FreeSurfer spherical projection was found. These coordinates are then rotated at three randomly generated angles. The Euclidean distance between each pair of rotated region centroids and original region centroids is the calculated. Regions are then iteratively assigned to rotated ones by finding which rotated region is closest on average to all original regions, and then mapping that rotated one to the most distant original region. This process is repeated until each rotated region is mapped to a unique original one. Based on this mapping, regional values can be mapped to a new region to preserve spatial contiguity. This procedure was repeated to produce 10,000 permutations. We calculated a spin-test derived *p*-value (p_{spin}) for a pair of brain maps by comparing the Pearson correlation between them to a distribution of correlations between one empirical map and 1,000 spatial permutations of the other (this was repeated such that each brain map was permuted and compared to the empirical pair, the mean of these runs *p*-values was then taken)[61]. Significance was determined at $p_{spin} < .05$.

To identify the top 100 genes that were positively/negatively correlated with PC1, while accounting for the smooth spatial variation of the gene expression maps, we performed a spin test to derive significance of the correlation between each gene's expression pattern across and PC1 score across thalamic seeds. For genes showing a significant effect, we extracted the 100 with the strongest positive correlation and 100 with the strongest negative correlation. Because thalamic data is represented as a volume rather than a surface, a separate spin-test was used. In this method the values of the thalamic seeds are randomly permuted, whereupon variogram modelling is used to smooth and rescale the data as to impose the original spatial autocorrelation[64]. We repeated this procedure 1,000 times for the PC1 scores and calculated the correlation between each of these permutations and expression of all 2,233 genes across thalamic seeds. This distribution of correlations for each gene-PC1 pair was then compared to the corresponding empirical correlation to establish significance ($p_{spin} < .05$). The top 100 positively and negatively correlated genes (as determined by correlation magnitude) which reached significance were then selected for further analysis.

Modelling gene expression trajectories

Using pre-processed PsychENCODE bulk tissue mRNA data[67], we first identified genes that were both expressed in the thalamus and differentially expressed across time. Differential expression was determined through pairwise comparison over nine developmental windows to determine genes that where enriched either prenatally or postnatally (a minimum of three significant pairwise differences were needed for a gene to be considered enriched for at least one of these timepoints).

To model the trajectory of medial- and lateral-genes, we used generalised additive models as previously described[43]. Briefly, genes expression was modelled as a nonlinear function of age with sex and RNA integrity number acting as fixed effects, along with a random intercept to account for sample-specific variation. The nonlinear function was specified to use a natural cubic spline with four knots evenly spaces across the age span for smoothness. AIC and BIC were used to evaluate model performance. Age-corrected relative gene-expressed was then calculated using the residuals of the best-fit nonlinear mixed model.

Data and code availability

All data for this project was obtained from open-source repositories detailed in **Table 1**. Code is available at https://github.com/StuartJO/ThalamicGradients. Data is available at https://doi.org/10.25374/MCRI.21556659.

Acknowledgements

The authors would like to thank Martina Arenella and Mac Shine for helpful discussions and feedback on the manuscript. Data were provided [in part] by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University.

Table 1. List of data sources

Dataset	Description	URL
Human Connectome Project	Diffusion and structural MRI data for 100 adults from the S1200 subjects data release[92–94]	https://www.humanconnectome.org/study/hcp-young-adult/data-releases
Allen Human Brain Atlas, voxelwise smoothed expression maps	Voxelwise expression maps for 2233 genes[50,51]	http://www.meduniwien.ac.at/neuroimaging/mRNA.html
Allen Mouse Brain Atlas	Gene expression data for 19419 genes across 213 regions in the mouse brain[56,58,59]	https://doi.org/10.5281/zenodo.4609603
Allen Mouse Connectivity Atlas	Axonal tracing data for 213 regions in the mouse brain[56–58]	https://doi.org/10.5281/zenodo.4609603
Allen Mouse Brain CCFv3 Atlas	NiFti volume and flat map for visualising mouse data[59,100]	https://scalablebrainatlas.incf.org/mouse/ABA_v3#downloads http://download.alleninstitute.org/publications/allen_mouse_brain_common_coordinate_framework/cortical_surface_views/ccf/annotation/
Mouse cortical properties	List of nine measures of cortical organisation in the mouse[28]	https://doi.org/10.6084/m9.figshare.7775684.v1
Mouse brain hierarchies	Measure of mouse brain hierarchy (based on cortico-cortical, thalamo-cortical, and cortico-thalamic connections)[55]	https://github.com/AllenInstitute/MouseBrainHierarchy/
DropViz	Drop-seq analysis of 89,027 cells in the adult mouse thalamus[33]	http://dropviz.org
The Human Protein Atlas	List of genes with protein expression in the human thalamus	https://www.proteinatlas.org/humanproteome/brain/thalamus
Ensembl Biomart human-mouse homologues	List of genes with identified mouse-human homologs	https://www.ensembl.org/index.html
PsychENCODE data	Developmental gene expression[67]	http://development.psychencode.org/

References

1. Sherman, S. M. & Guillery, R. W. The role of the thalamus in the flow of information to the cortex. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **357**, 1695–1708 (2002).
2. Sherman, S. M. Thalamus plays a central role in ongoing cortical functioning. *Nat. Neurosci.* **19**, 533–541 (2016).
3. Sommer, M. A. The role of the thalamus in motor control. *Curr. Opin. Neurobiol.* **13**, 663–670 (2003).
4. Mitchell, A. S. The mediodorsal thalamus as a higher order thalamic relay nucleus important for learning and decision-making. *Neurosci. Biobehav. Rev.* **54**, 76–88 (2015).
5. Jankowski, M. M. *et al.* The anterior thalamus provides a subcortical circuit supporting memory and spatial navigation. *Front. Syst. Neurosci.* **7**, (2013).
6. Morel, A. *Stereotactic atlas of the human thalamus and basal ganglia*. (CRC Press, 2007).
7. Jones, E. G. *The thalamus*. (Cambridge University Press, 2007).
8. Fama, R. & Sullivan, E. V. Thalamic structures and associated cognitive functions: Relations with age and aging. *Neurosci. Biobehav. Rev.* **54**, 29–37 (2015).
9. Mai, J. K. & Majtanik, M. Toward a Common Terminology for the Thalamus. *Front. Neuroanat.* **12**, 114 (2019).
10. Jones, E. G. Viewpoint: the core and matrix of thalamic organization. *Neuroscience* **85**, 331–345 (1998).
11. Jones, E. G. The thalamic matrix and thalamocortical synchrony. *Trends Neurosci.* **24**, 595–601 (2001).
12. McFarland, N. R. & Haber, S. N. Thalamic Relay Nuclei of the Basal Ganglia Form Both Reciprocal and Nonreciprocal Cortical Connections, Linking Multiple Frontal Cortical Areas. *J. Neurosci.* **22**, 8117–8132 (2002).

13. Roy, D. S., Zhang, Y., Halassa, M. M. & Feng, G. Thalamic subnetworks as units of function. *Nat. Neurosci.* **25**, 140–153 (2022).

14. Phillips, J. W. *et al.* A repeated molecular architecture across thalamic pathways. *Nat. Neurosci.* **22**, 1925–1935 (2019).

15. Majorossy, K. & Kiss, A. Types of interneurons and their participation in the neuronal network of the medial geniculate body. *Exp. Brain Res.* **26**, (1976).

16. Nakagawa, Y. & O'Leary, D. D. M. Combinatorial Expression Patterns of LIM-Homeodomain and Other Regulatory Genes Parcellate Developing Thalamus. *J. Neurosci.* **21**, 2711–2725 (2001).

17. BRAIN Initiative Cell Census Network (BICCN) *et al.* A multimodal cell census and atlas of the mammalian primary motor cortex. *Nature* **598**, 86–102 (2021).

18. Cembrowski, M. S. & Spruston, N. Heterogeneity within classical cell types is the rule: lessons from hippocampal pyramidal neurons. *Nat. Rev. Neurosci.* **20**, 193–204 (2019).

19. Clascá, F., Rubio-Garrido, P. & Jabaudon, D. Unveiling the diversity of thalamocortical neuron subtypes: Thalamocortical neuron diversity. *Eur. J. Neurosci.* **35**, 1524–1532 (2012).

20. Li, Y. *et al.* Distinct subnetworks of the thalamic reticular nucleus. *Nature* **583**, 819–824 (2020).

21. Gao, C. *et al.* Two genetically, anatomically and functionally distinct cell types segregate across anteroposterior axis of paraventricular thalamus. *Nat. Neurosci.* **23**, 217–228 (2020).

22. Nagalski, A. *et al.* Molecular anatomy of the thalamic complex and the underlying transcription factors. *Brain Struct. Funct.* **221**, 2493–2510 (2016).

23. Price, D. J., Clegg, J., Duocastella, X. O., Willshaw, D. & Pratt, T. The Importance of Combinatorial Gene Expression in Early Mammalian Thalamic Patterning and Thalamocortical Axonal Guidance. *Front. Neurosci.* **6**, (2012).

24. Teissier, A. & Pierani, A. Wiring of higher-order cortical areas: Spatiotemporal development of cortical hierarchy. *Semin. Cell Dev. Biol.* **118**, 35–49 (2021).

25. Huntenburg, J. M., Bazin, P. L. & Margulies, D. S. Large-Scale Gradients in Human Cortical Organization. *Trends Cogn. Sci.* 22, 21–31 (2018).
26. Margulies, D. S. *et al.* Situating the default-mode network along a principal gradient of macroscale cortical organization. *Proc. Natl. Acad. Sci. U. S. A.* 113, 12574–12579 (2016).
27. Paquola, C. *et al.* Microstructural and Functional Gradients are Increasingly Dissociated in Transmodal Cortices. *PLOS Biol.* 17, e3000284 (2019).
28. Fulcher, B. D., Murray, J. D., Zerbi, V. & Wang, X. J. Multimodal gradients across mouse cortex. *Proc. Natl. Acad. Sci.* 116, 4689–4695 (2019).
29. Vogel, J. W. *et al.* Conserved whole-brain spatiomolecular gradients shape adult brain functional organization. (2022) doi:10.1101/2022.09.18.508425.
30. Sydnor, V. J. *et al.* Neurodevelopment of the association cortices: Patterns, mechanisms, and implications for psychopathology. *Neuron* 1–27 (2021) doi:10.1016/j.neuron.2021.06.016.
31. Huntenburg, J. M., Yeow, L. Y., Mandino, F. & Grandjean, J. Gradients of functional connectivity in the mouse cortex reflect neocortical evolution. *NeuroImage* 225, 117528 (2021).
32. Cahalane, D. J., Charvet, C. J. & Finlay, B. L. Systematic, balancing gradients in neuron density and number across the primate isocortex. *Front. Neuroanat.* 6, 12 (2012).
33. Saunders, A. *et al.* Molecular Diversity and Specializations among the Cells of the Adult Mouse Brain. *Cell* 174, 1015-1030.e16 (2018).
34. Behrens, T. E. J. *et al.* Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging. *Nat. Neurosci.* 6, 750–757 (2003).
35. Lambert, C., Simon, H., Colman, J. & Barrick, T. R. Defining thalamic nuclei and topographic connectivity gradients in vivo. *NeuroImage* 158, 466–479 (2017).
36. Zhang, D., Snyder, A. Z., Shimony, J. S., Fox, M. D. & Raichle, M. E. Noninvasive Functional and Structural Connectivity Mapping of the Human Thalamocortical System. *Cereb. Cortex* 20, 1187–1194 (2010).

37. Yang, S. *et al.* The thalamic functional gradient and its relationship to structural basis and cognitive relevance. *NeuroImage* **218**, 116960 (2020).
38. Horn, A. *et al.* Lead-DBS v2: Towards a comprehensive pipeline for deep brain stimulation imaging. *NeuroImage* **184**, 293–316 (2019).
39. Isaacs, B. R. *et al.* Variability in subthalamic nucleus targeting for deep brain stimulation with 3 and 7 Tesla magnetic resonance imaging. *NeuroImage Clin.* **32**, 102829 (2021).
40. Iglehart, C., Monti, M., Cain, J., Tourdias, T. & Saranathan, M. A systematic comparison of structural-, structural connectivity-, and functional connectivity-based thalamus parcellation techniques. *Brain Struct. Funct.* **225**, 1631–1642 (2020).
41. Kumar, V. J., van Oort, E., Scheffler, K., Beckmann, C. F. & Grodd, W. Functional anatomy of the human thalamus at rest. *NeuroImage* **147**, 678–691 (2017).
42. Keifer, O. P., Gutman, D. A., Hecht, E. E., Keilholz, S. D. & Ressler, K. J. A comparative analysis of mouse and human medial geniculate nucleus connectivity: A DTI and anterograde tracing study. *NeuroImage* **105**, 53–66 (2015).
43. Ball, G. *et al.* Cortical morphology at birth reflects spatiotemporal patterns of gene expression in the fetal human brain. *PLOS Biol.* **18**, e3000976 (2020).
44. Valk, S. L. *et al.* Genetic and phylogenetic uncoupling of structure and function in human transmodal cortex. *Nat. Commun.* **13**, 2341 (2022).
45. Paquola, C. *et al.* Shifts in myeloarchitecture characterise adolescent development of cortical gradients. *eLife* **8**, 1–23 (2019).
46. Valk, S. L. *et al.* Shaping brain structure: Genetic and phylogenetic axes of macroscale organization of cortical thickness. *Sci. Adv.* **6**, eabb3417 (2020).
47. Park, B. *et al.* Multilevel neural gradients reflect transdiagnostic effects of major psychiatric conditions on cortical morphology. Preprint at https://doi.org/10.1101/2021.10.29.466434 (2021).

48. Seidlitz, J. *et al.* Transcriptomic and cellular decoding of regional brain vulnerability to neurogenetic disorders. *Nat. Commun.* **11**, 3358 (2020).
49. Goulas, A., Majka, P., Rosa, M. G. P. & Hilgetag, C. C. A blueprint of mammalian cortical connectomes. *PLOS Biol.* **17**, e2005346 (2019).
50. Gryglewski, G. *et al.* Spatial analysis and high resolution mapping of the human whole-brain transcriptome for integrative analysis in neuroimaging. *NeuroImage* **176**, 259–267 (2018).
51. Hawrylycz, M. J. *et al.* An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature* **489**, 391–399 (2012).
52. Burt, J. B. *et al.* Hierarchy of transcriptomic specialization across human cortex captured by structural neuroimaging topography. *Nat. Neurosci.* **21**, 1251–1259 (2018).
53. Fagerberg, L. *et al.* Analysis of the Human Tissue-specific Expression by Genome-wide Integration of Transcriptomics and Antibody-based Proteomics. *Mol. Cell. Proteomics* **13**, 397–406 (2014).
54. Coifman, R. R. *et al.* Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps. *Proc. Natl. Acad. Sci.* **102**, 7426–7431 (2005).
55. Harris, J. A. *et al.* Hierarchical organization of cortical and thalamic connectivity. *Nature* **575**, 195–202 (2019).
56. Fulcher, B. D. & Fornito, A. A transcriptional signature of hub connectivity in the mouse connectome. *Proc. Natl. Acad. Sci.* **113**, 1435–1440 (2016).
57. Oh, S. W. *et al.* A mesoscale connectome of the mouse brain. *Nature* **508**, 207–214 (2014).
58. Fulcher, B. D., Arnatkeviciute, A. & Fornito, A. Overcoming false-positive gene-category enrichment in the analysis of spatially resolved transcriptomic brain atlas data. *Nat. Commun.* **12**, 2669 (2021).
59. Lein, E. S. *et al.* Genome-wide atlas of gene expression in the adult mouse brain. *Nature* **445**, 168–176 (2007).

60. Markello, R. D. *et al.* neuromaps: structural and functional interpretation of brain maps. *Nat. Methods* **19**, 1472–1479 (2022).

61. Váša, F. *et al.* Adolescent Tuning of Association Cortex in Human Structural Brain Networks. *Cereb. Cortex* **28**, 281–294 (2018).

62. Jager, P. *et al.* Dual midbrain and forebrain origins of thalamic inhibitory interneurons. *eLife* **10**, e59272 (2021).

63. Evangelio, M., García-Amado, M. & Clascá, F. Thalamocortical Projection Neuron and Interneuron Numbers in the Visual Thalamic Nuclei of the Adult C57BL/6 Mouse. *Front. Neuroanat.* **12**, 27 (2018).

64. Burt, J. B., Helmer, M., Shinn, M., Anticevic, A. & Murray, J. D. Generative modeling of brain maps with spatial autocorrelation. *NeuroImage* **220**, 117038 (2020).

65. Letinic, K. & Rakic, P. Telencephalic origin of human thalamic GABAergic neurons. *Nat. Neurosci.* **4**, 931–936 (2001).

66. Krienen, F. M. *et al.* A marmoset brain cell census reveals persistent influence of developmental origin on neurons. Preprint at https://doi.org/10.1101/2022.10.18.512442 (2022).

67. Li, M. *et al.* Integrative functional genomic analysis of human brain development and neuropsychiatric risks. *Science* **362**, eaat7615 (2018).

68. Avila, A. *et al.* Glycine Receptor α2 Subunit Activation Promotes Cortical Interneuron Migration. *Cell Rep.* **4**, 738–750 (2013).

69. Zhang, L., Huang, Y., Chen, J.-Y., Ding, Y.-Q. & Song, N.-N. DSCAM and DSCAML1 regulate the radial migration and callosal projection in developing cerebral cortex. *Brain Res.* **1594**, 61–70 (2015).

70. Bielle, F. *et al.* Emergent Growth Cone Responses to Combinations of Slit1 and Netrin 1 in Thalamocortical Axon Topography. *Curr. Biol.* **21**, 1748–1755 (2011).

71. Feng, J. *et al.* Celsr3 and Fzd3 Organize a Pioneer Neuron Scaffold to Steer Growing Thalamocortical Axons. *Cereb. Cortex* **26**, 3323–3334 (2016).
72. Qu, Y. *et al.* Genetic evidence that *Celsr3* and *Celsr2*, together with *Fzd3*, regulate forebrain wiring in a *Vangl*-independent manner. *Proc. Natl. Acad. Sci.* **111**, (2014).
73. Iwasato, T. *et al.* Cortical Adenylyl Cyclase 1 Is Required for Thalamocortical Synapse Maturation and Aspects of Layer IV Barrel Development. *J. Neurosci.* **28**, 5931–5943 (2008).
74. Elvsåshagen, T. *et al.* The genetic architecture of the human thalamus and its overlap with ten common brain disorders. *Nat. Commun.* **12**, 2909 (2021).
75. Peters, S. K., Dunlop, K. & Downar, J. Cortico-Striatal-Thalamic Loop Circuits of the Salience Network: A Central Pathway in Psychiatric Disease and Treatment. *Front. Syst. Neurosci.* **10**, (2016).
76. Murray, J. D. *et al.* A hierarchy of intrinsic timescales across primate cortex. *Nat. Neurosci.* **17**, 1661–1663 (2014).
77. Brysch, W., Brysch, I., Creutzfeldt, O. D., Schlingensiepen, R. & Schlingensiepen, K.-H. The topology of the thalamo-cortical projections in the marmoset monkey (Callithrix jacchus). *Exp. Brain Res.* **81**, 1–17 (1990).
78. Höhl-Abrahão, J. C. & Creutzfeldt, O. D. Topographical mapping of the thalamocortical projections in rodents and comparison with that in primates. *Exp. Brain Res.* **87**, 283–294 (1991).
79. Müller, E. J. *et al.* Core and matrix thalamic sub-populations relate to spatio-temporal cortical connectivity gradients. *NeuroImage* **222**, 117224 (2020).
80. Ferguson, B. R. & Gao, W.-J. Thalamic Control of Cognition and Social Behavior Via Regulation of Gamma-Aminobutyric Acidergic Signaling and Excitation/Inhibition Balance in the Medial Prefrontal Cortex. *Biol. Psychiatry* **83**, 657–669 (2018).

81. Fratzl, A. & Hofer, S. B. The caudal prethalamus: Inhibitory switchboard for behavioral control? *Neuron* 110, 2728–2742 (2022).

82. Guillery, R. W. & Sherman, S. M. Their Role in Corticocortical Communication: Generalizations from the Visual System. *Neuron* 33, 163–175.

83. Bosch-Bouju, C., Hyland, B. I. & Parr-Brownlie, L. C. Motor thalamus integration of cortical, cerebellar and basal ganglia information: implications for normal and parkinsonian conditions. *Front. Comput. Neurosci.* 7, (2013).

84. Borgan, F., Kokkinou, M. & Howes, O. The Cannabinoid CB1 Receptor in Schizophrenia. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* 6, 646–659 (2021).

85. Quiñones, G. M., Mayeli, A., Yushmanov, V. E., Hetherington, H. P. & Ferrarelli, F. Reduced GABA/glutamate in the thalamus of individuals at clinical high risk for psychosis. *Neuropsychopharmacology* 46, 1133–1139 (2021).

86. Akyuz, E. *et al.* Revisiting the role of neurotransmitters in epilepsy: An updated review. *Life Sci.* 265, 118826 (2021).

87. Pitel, A. L., Segobin, S. H., Ritz, L., Eustache, F. & Beaunieux, H. Thalamic abnormalities are a cardinal feature of alcohol-related brain dysfunction. *Neurosci. Biobehav. Rev.* 54, 38–45 (2015).

88. Banerjee, N. Neurotransmitters in alcoholism: A review of neurobiological and genetic studies. *Indian J. Hum. Genet.* 20, 20 (2014).

89. Varela, C. Thalamic neuromodulation and its implications for executive networks. *Front. Neural Circuits* 8, (2014).

90. Wong, S. Z. H. *et al.* In vivo clonal analysis reveals spatiotemporal regulation of thalamic nucleogenesis. *PLoS Biol.* 16, 1–24 (2018).

91. Nakagawa, Y. Development of the thalamus: From early patterning to regulation of cortical functions. *WIREs Dev. Biol.* 8, (2019).

92. Van Essen, D. C. *et al.* The Human Connectome Project: A data acquisition perspective. *NeuroImage* 62, 2222–2231 (2012).

93. Glasser, M. F. *et al.* The minimal preprocessing pipelines for the Human Connectome Project. *NeuroImage* 80, 105–124 (2013).

94. Sotiropoulos, S. N. *et al.* Advances in diffusion MRI acquisition and processing in the Human Connectome Project. *NeuroImage* 80, 125–143 (2013).

95. Fischl, B. FreeSurfer. *NeuroImage* 62, 774–781 (2012).

96. Tournier, J. D. *et al.* MRtrix3: A fast, flexible and open software framework for medical image processing and visualisation. *NeuroImage* 202, 116137 (2019).

97. Tournier, J. D., Calamante, F. & Connelly, A. Improved probabilistic streamlines tractography by 2nd order integration over fibre orientation distributions. *Proc. Int. Soc. Magn. Reson. Med.* 1670 (2010).

98. Smith, R. E., Tournier, J. D., Calamante, F. & Connelly, A. Anatomically-constrained tractography: Improved diffusion MRI streamlines tractography through effective use of anatomical information. *NeuroImage* 62, 1924–1938 (2012).

99. Parkes, L., Fulcher, B. D., Yücel, M. & Fornito, A. Transcriptional signatures of connectomic subregions of the human striatum. *Genes Brain Behav.* 16, 647–663 (2017).

100. Bakker, R., Tiesinga, P. & Kötter, R. The Scalable Brain Atlas: Instant Web-Based Access to Public Brain Atlases and Related Content. *Neuroinformatics* 13, 353–366 (2015).

101. Kim, Y. Brain-wide Maps Reveal Stereotyped Cell-Type-Based Cortical Architecture and Subcortical Sexual Dimorphism. 37.

102. Goulas, A., Uylings, H. B. M. & Hilgetag, C. C. Principles of ipsilateral and contralateral cortico-cortical connectivity in the mouse. *Brain Struct. Funct.* 222, 1281–1295 (2017).

103. Liao, Y., Wang, J., Jaehnig, E. J., Shi, Z. & Zhang, B. WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs. *Nucleic Acids Res.* 47, W199–W205 (2019).

104. Piñero, J. *et al.* The DisGeNET knowledge platform for disease genomics: 2019 update. *Nucleic Acids Res.* gkz1021 (2019) doi:10.1093/nar/gkz1021.

105. Jourquin, J., Duncan, D., Shi, Z. & Zhang, B. GLAD4U: deriving and prioritizing gene lists from PubMed literature. *BMC Genomics* 13, S20 (2012).

106. Hamosh, A. Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. *Nucleic Acids Res.* 33, D514–D517 (2004).