Interneuron loss and microglia activation by transcriptome analyses in the basal ganglia of Tourette syndrome

Yifan Wang[1], Liana Fasching[2], Feinan Wu[2], Anita Huttner[3], Sabina Berretta[4], Rosalinda Roberts[5], James F. Leckman[2], Alexej Abyzov[1], Flora M. Vaccarino[2,6,7]

[1]Department of Quantitative Health Sciences, Center for Individualized Medicine, Mayo Clinic, Rochester, MN 55905, USA
[2]Child Study Center, Yale University, New Haven, CT 06520, USA
[3]Department of Pathology, Yale University, New Haven, CT 06520, USA
[4]McLean Hospital, Harvard Medical School, Belmont, MA, 02478, USA
[5]Department of Psychiatry and Behavioral Neurobiology, University of Alabama at Birmingham, Birmingham, AL 35294, USA
[6]Department of Neuroscience, Yale University, New Haven, CT 06520, USA
[7]Yale Kavli Institute for Neuroscience, New Haven, CT 06520, USA

*Correspondence: abyzov.alexej@mayo.edu (A.A.) and flora.vaccarino@yale.edu (F.M.V.)

Abstract
Tourette syndrome (TS) is a disorder of high-order integration of sensory, motor, and cognitive functions afflicting as many as 1 in 150 children and characterized by motor hyperactivity and tics. Despite high familial recurrence rates, a few risk genes and no biomarkers have emerged as causative or predisposing factors. The syndrome is believed to originate in basal ganglia, where patterns of motor programs are encoded. Postmortem immunocytochemical analyses of brains with severe TS revealed decreases in cholinergic, fast-spiking parvalbumin, and somatostatin interneurons within the striatum (caudate and putamen nuclei). Here, we performed single cell transcriptomic and chromatin accessibility analyses of the caudate nucleus from 6 adult TS and 6 control post-mortem brains. The data reproduced the known cellular composition of the adult human striatum, including a majority of medium spiny neurons (MSN) and small populations of GABAergic and cholinergic interneurons. Comparative analysis revealed that interneurons were decreased by roughly 50% in TS brains, while no difference was observed for other cell types. Differential gene expression analysis suggested that mitochondrial function, and specifically oxidative metabolism, in MSN and synaptic function in interneurons are both impaired in TS subjects, while microglia display strong activation of immune response pathways. Our data explicitly link gene expression changes to changes in cis-regulatory activity in the corresponding cell types, suggesting de-regulation as a factor for the etiology of TS. These findings expand on previous research and suggest that impaired modulation of striatal function by interneurons may be the origin of TS symptoms.

Introduction

Tourette syndrome (TS) is a common disorder afflicting as many as 1 in 150 children[1]. The main symptoms are motor and vocal tics, such as eye-blinking or throat clearing, that are performed out-of-context and in a repetitive fashion[2]. TS has one of the highest familial recurrence rates among complex neuropsychiatric disorders[3]; yet few risk genes meeting genome-wide statistical thresholds and no biomarkers have emerged as causative or predisposing factors, despite the efforts of international consortia and the combined examination of nearly 5,000 TS and 10,000 NC blood samples[3].

TS is a disorder of high-order integration of sensory, motor, and cognitive functions. Patterns of motor programs are encoded in the basal ganglia (BG), a set of interconnected nuclei situated deep within the cerebral hemispheres (**Fig. 1A**). The BG are activated during tics in TS[4, 5]. Neuroimaging data suggest that the caudate and putamen (referred to as striatum) are smaller in TS[6, 7]. Notably caudate volumes in childhood are predictive of TS severity in adulthood [8] and is reduced in volume in the most severely affected twin in monozygotic twin discordant for TS [9]. Our postmortem immunocytochemical analyses in adult individuals with severe TS revealed decreases in cholinergic interneurons (CH-IN) and in two GABAergic interneurons subtypes: PV-IN (PV-IN) and somatostatin/nitric oxide synthase/neuropeptide Y^+ (SST/NOS/NPY-IN) in the striatum, suggesting decreased generation, loss or hypofunction of these cell types[10-12].

The BG tonically inhibit the ventral and intralaminar thalamic nuclei through their common output, the globus pallidus pars interna/substantia nigra reticulata (GPi/SNr). The striatum receives converging inputs from the cortex, thalamus and midbrain dopamine neurons (**Fig. 1B**). The striatal medium spiny neurons (MSNs) form two major pathways, the indirect and direct pathway, which have opposite behavioral effects[13]. The direct pathway inhibits the GPi/SNr, resulting in disinhibition of thalamic neurons and facilitation of motor behavior; conversely, the indirect pathway excites the GPi/SNr, inhibiting motor behavior. Striatal interneurons exert important modulatory roles on these two pathways. The fast-spiking, parvalbumin$^+$ interneurons (PV-IN) respond with short latency to cortical stimulation[14-16] and exert feed-forward inhibition preferentially on MSNs of the direct pathway[17, 18]. The cholinergic interneurons (CH-IN) receive input from the intralaminar nuclei of the thalamus[19] and modulate the activity of INT and MSN PV-IN[20] (**Fig. 1B**). Hence, the reported decreases in interneuron number in TS[10, 11] should hinder the inhibitory influence of thalamus and cortex upon the activity of the striatal direct pathway. This notion is supported by work in animal models, as selective ablation of CH-IN, or combined ablation of PV-IN and CH-IN in mouse, causes repetitive stereotypic behavior and increases functional connectivity between frontal cortical regions and the motor region of the striatum[21-23]. Since previous work on striatal interneurons in TS was carried out primarily by immunocytochemical analyses, here we sought to advance our understanding of the pathophysiology of TS by performing single cell transcriptomic and chromatin accessibility analyses to both confirm the previous findings on the RNA level and perform a comprehensive characterization of the cellular composition and gene expression of the basal ganglia in TS individuals.

Results

Identification and annotation of major cell types

Human brain basal ganglia frozen brain samples from the head of the caudate nucleus were obtained from 12 individuals (6 patients with Tourette syndrome, 6 controls, paired by age and sex; **Supplementary Table 1**) and nuclear fractions were processed for single cell transcriptome analyses. For 3 pairs, we used stand-alone single-nuclei RNA-seq (snRNA-seq) and single nuclei ATAC-seq (snATAC-seq), while the remaining 3 pairs were processed using the multiome snRNA-seq and snATAC-seq platform from 10x, which allows deriving transcriptomic and epigenomic data from the same cell. We uniformly processed all 12 samples and concatenated the data from both techniques for downstream analysis (**Methods**).

After data QC, we retained 43,713 high-quality single nuclei whose expression profiles were clustered by similarity into 26 clusters and visualized via UMAP (**Figure 1C**). Clusters were grouped into 6 major cell types using known markers and public datasets[24, 25] (**Figure 1D, C**; **Methods**). We annotated 18,498 (42%) medium spiny neurons (MSN), 1,492 (3%) interneurons (INT), 5,394 (12%) astrocytes (AST), 15,563 (36%) oligodendrocytes (OL), 1,496 (3%) microglia (MG), 719 (2%) endothelial cells (EC), and 551 (1%) unidentified (UN) (**Supplementary Table 2,3, Supplementary Figure 1**). Overall, this cell distribution roughly corresponded

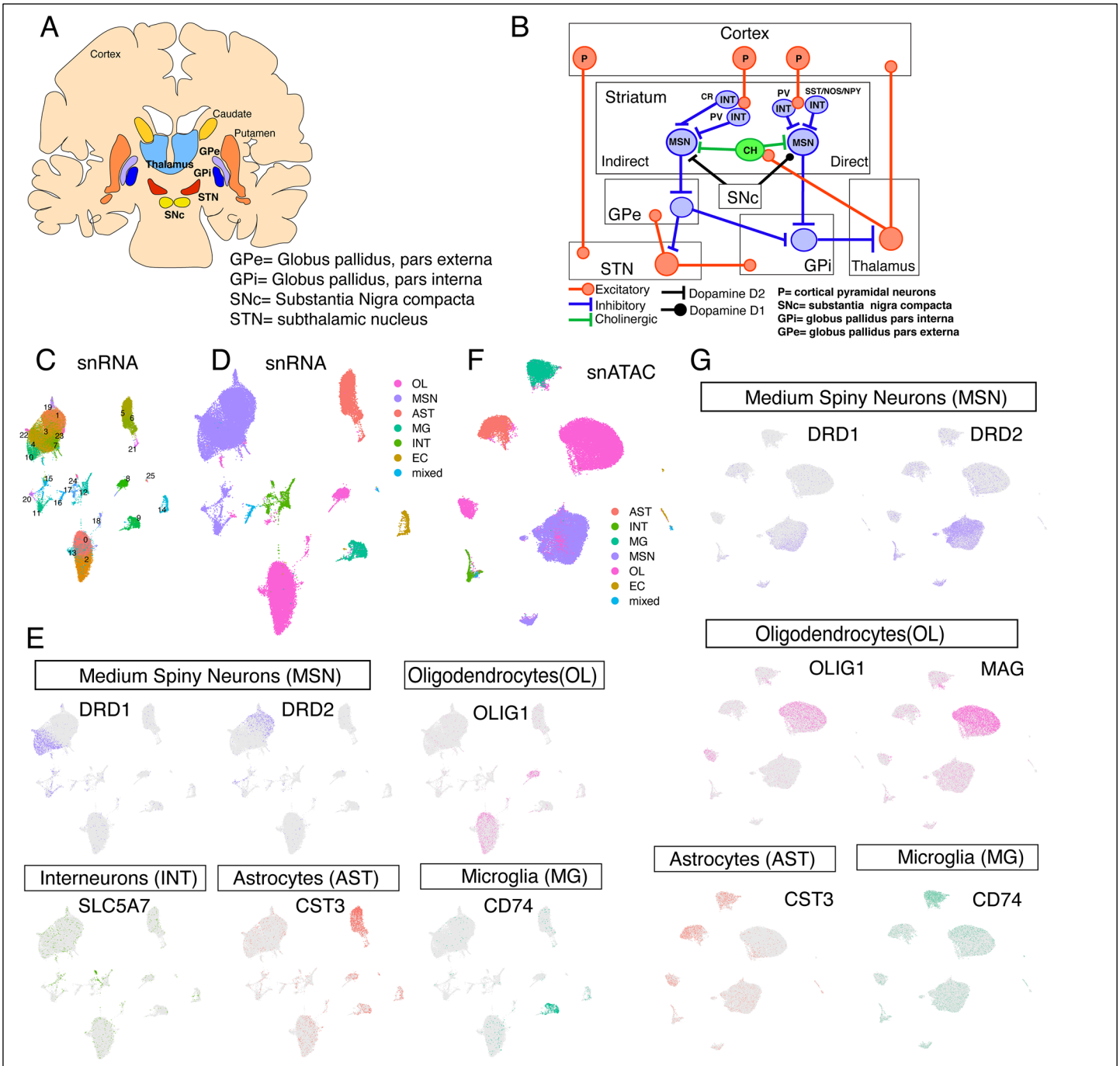

Figure 1. snRNA-seq and snATAC-seq clustering, cell types and genetic markers. A) Outline of Basal ganglia (BG) anatomy in a coronal human brain section. **B)** Wiring diagram of BG circuitry. **C)** UMAP clustering of nuclei from 12 brains based on snRNA-seq data. **D)** Annotation of snRNA-seq clusters with cell types. **E)** UMAP plots colored by marker expression for 5 cell types in snRNA-seq data. **F)** UMAP clustering of nuclei from 12 brains based on snATAC-seq data. Clusters are annotated based on label transfer from snRNA-seq clusters. **G)** UMAPs colored by marker expression for 5 cell types in snATAC-seq data. DRD1 and DRD2, dopamine receptor 1 and 2, respectively; OLIG1, Oligodendrocyte Transcription Factor 1; MAG, myelin associated glycoprotein 1; SLC5A7, high-affinity choline transporter expressed in cholinergic neurons; NPY, neuropeptide Y; CST3, Cystatin 1, a protease inhibitor expressed in astrocytes; CD74, a major histocompatibility complex class II (MHCII) interacting molecule expressed in microglia.

to known proportions of cell types in human striatum, e.g., neurons account for ~45% of all cells, while MSN account for ~93% of all neurons, and IN account for ~7% of all neurons. The MSN were sub-clustered into dopamine receptor 1 (DRD1)-MSN (45%) and dopamine receptor 2 (DRD2)-MSN (44%) (**Supplementary Figure 2**) which represent projection neurons of the direct and indirect pathways, respectively, distinguishable by specific gene expression profiles and each projecting to a separate set of targets, the globus pallidus pars interna and globus pallidus pars externa (**Fig. 1B**)[26].

We further profiled single-nuclei chromatin accessibility in the 11 brains (sample from brain TS2 was excluded due to low number of nuclei passing QC) and annotated the 6 major cell types using the label transfer method from the snRNA-seq cell types (**Methods**). Out of 39,728 single nuclei which passed QC, 33% were classified as MSN, 2% as IN, 8% as AST, 47% as OL, 9% as MG, 1% as EN, and 1% as unidentified (**Figure 1F, G, Supplementary Table 4,5, Supplementary Figure 3**). Label transfer accuracy was 97.6%, as estimated using the multiome snRNA-seq and snATAC-seq data (**Supplementary Figure 4**). Together, the data indicate that using snRNA-seq and scATAC-seq we were able to map the principal cell types of the human striatum, including the relatively sparse populations of striatal interneurons.

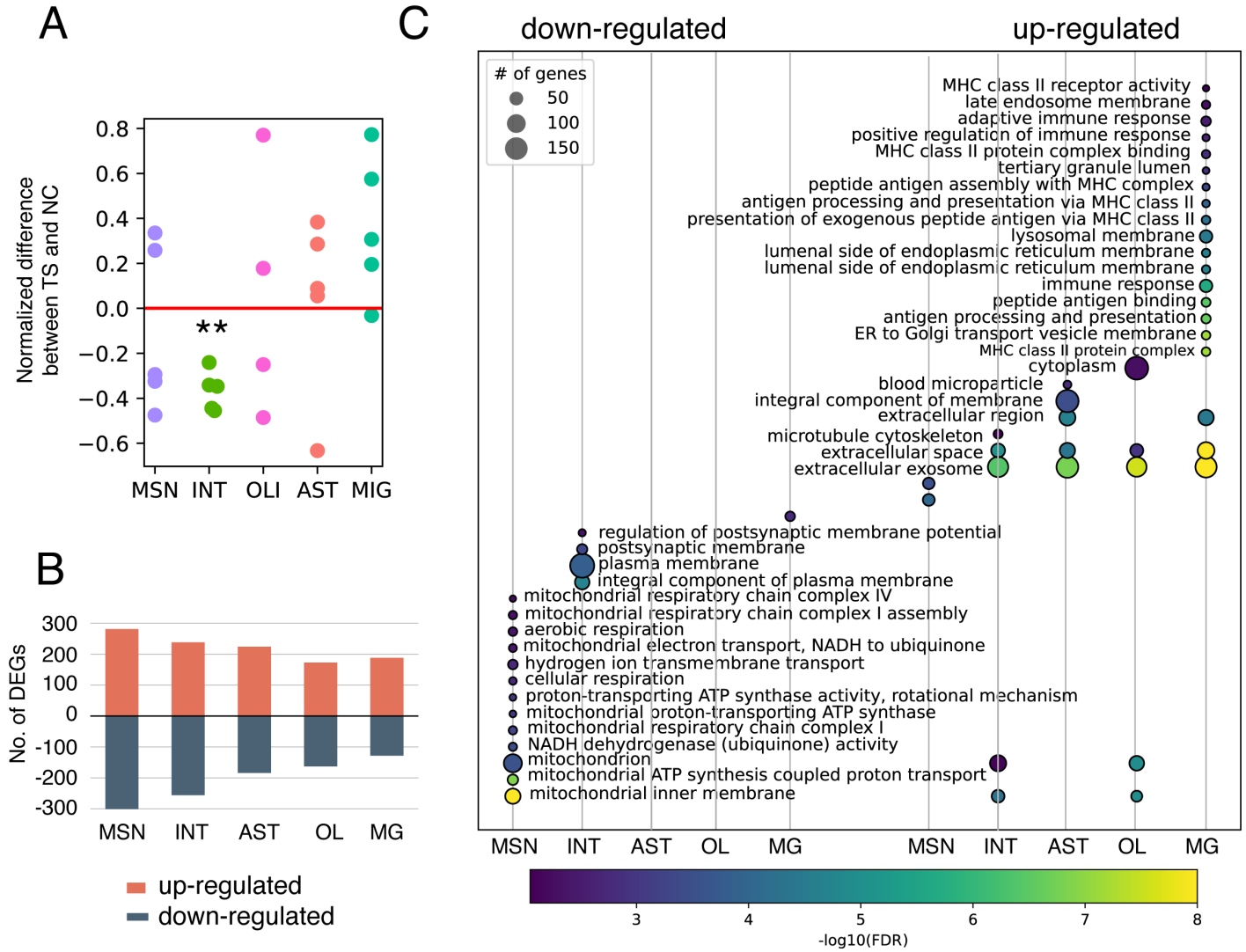

Figure 2. Tourette syndrome brains transcriptomic profiles compared to normal controls. A) Difference in number of cells in each cell types comparing TS to NC using snRNA-seq data. Y axis: normalized percentage difference between TS and NC; x-axis: cell types. **B)** Number of differentially expressed genes in each cell types comparing TS to NC brains. ***p-value < 0.001 by paired t-test. **C)** GO terms enrichment for differentially expressed genes (up-regulated and down-regulated) in TS vs NC brains. y-axis: enriched GO terms; x-axis: cell types. The size of a circle indicates the number of genes in the enriched GO term, while color corresponds to -log10(FDR) of the enriched GO term. GO terms were ordered based on the p-value in each cell type.

Differential cell type composition and differentially expressed genes (DEGs) in patients with Tourette Syndrome

Next, we considered the 6 TS/NC pairs matched for age, sex, and PMI (**Supplemental Table 1**). One pair was excluded because of the low total count of nuclei passing QC in the paired control (**Supplemental Table 2**). We then used a paired t-test to compare proportions in each cell type between TS and NC samples. We observed significantly fewer INT in the TS brains (2.4% vs 5.5% in TS vs NC, respectively, p-value<0.001), with no significant difference in proportion for other cell types (**Figure 2A, Supplementary Figure 1, Supplementary Table 3**). This observation is concordant with observations made previously by immunocytochemistry in a different cohort of postmortem brains[10-12].

To identify genes potentially involved in TS neurodevelopment, we assessed differentially expressed genes (DEGs) by comparing combined data from all single nuclei in TS and control brains in each annotated cell type (**Methods**). We observed between 300 and 600 DEGs in each cell type, with no statistically significant difference in the number of up and down-regulated DEGs. Down-regulated DEGs in MSN were enriched in gene ontology (GO) annotation for mitochondrial function related pathways, such as mitochondrial inner membrane and ATP synthesis. Interestingly, down-regulated DEGs in IN were enriched in plasma membrane/postsynaptic membrane annotations comprising GABA and AMPA/kainate glutamate post-synaptic receptors as well as several protocadherins, including PCDH17, PCDH19, PCDH7). In contrast, up-regulated DEGs in INT (as well as in AST, OL and MG) were enriched in GO annotations for extracellular exosomes, while those upregulated in MG were enriched in immune response-related annotations (**Fig. 2C**). Downregulation in synaptic transmission-related and protocadherin-related gene network modules and the upregulation in immune response-related modules were also observed in a prior study of bulk RNA-seq expression performed in the basal ganglia of an independent cohort of TS subjects[12]. Here we confirm these major signatures in single cells and assess their specific cellular origins – INT for synaptic transmission and protocadherin, MG for immune activation. Overall, the data suggest compromised mitochondrial function in TS neurons and activation of TS microglia.

Differentially accessible regions in Tourette syndrome patients and co-accessibility networks

Across the 5 cell types we identified 339,857 peaks from snATAC-seq data, of which 89,234 were cell type specific. Among all peaks, 33% were residing within the exonic regions, 45% within intronic regions, 2% within the 2kb regions upstream and downstream of transcriptional start sites, and 20% within the intergenic regions (**Figure 3B**). We next performed GO annotation considering the genes in proximity to peaks. The GO biological process terms enriched for each cell type were consistent with the known function of the cell type; for example, synaptic transmission-related terms for MSN and IN, axon ensheathment and myelination for OL, and immune response signaling for MG (**Supplementary Figure 5**). Using the snATAC-seq data, we observed the same trend as in RNA-seq data of fewer INT in TS brains, however the trend didn't reach statistical significance (**Fig. 3A**).

To investigate the differentially accessible regions, we compared the snATAC-seq peaks in Tourette syndrome brains to controls in each major cell type using the FindMarkers function in Seurat [27]. AST and MG exhibited the highest number of differentially accessible regions in TS vs NC (**Fig. 3C; Methods**). We further analyzed the differentially accessible peaks (DAP) between TS and NC and identified different GO biological process terms enriched for different cell types (**Fig. 3D**). Analysis of GO terms for DAP-proximal genes of revealed consistency with analogous analyses for DEGs. There was an enrichment of immune system related terms with more accessible DAPs in microglia, cilium function related terms with more accessible DAPs in astrocytes, and the inactivation of MAPK cascade with less accessible DAPs in MSN (**Fig. 3D**).

Next, we called possible physical links between the identified peaks by analyzing their co-accessibility across sub-cluster of nuclei of local cell neighborhoods based on dimension reduction. We focused on pairs of peaks where at least one peak corresponded to the TSS of a gene (**Methods**). Out of all 275,573 identified links, 77,263 were between pairs of peaks where only one corresponded to TSS, while the other one was (by exclusion) a putative enhancer. Consequently, such pairs likely represented enhancer-promoter cis-regulatory interactions. This set of links allowed us to investigate the correspondence between differences in gene expression and enhancer activity in TS vs control. Remarkably, we observed a significant correlation of the links between DEGs and DAPs (**Fig. 3E**; p-value < 0.001 by Fisher's exact test), revealing a connection between expression and regulatory changes in TS.

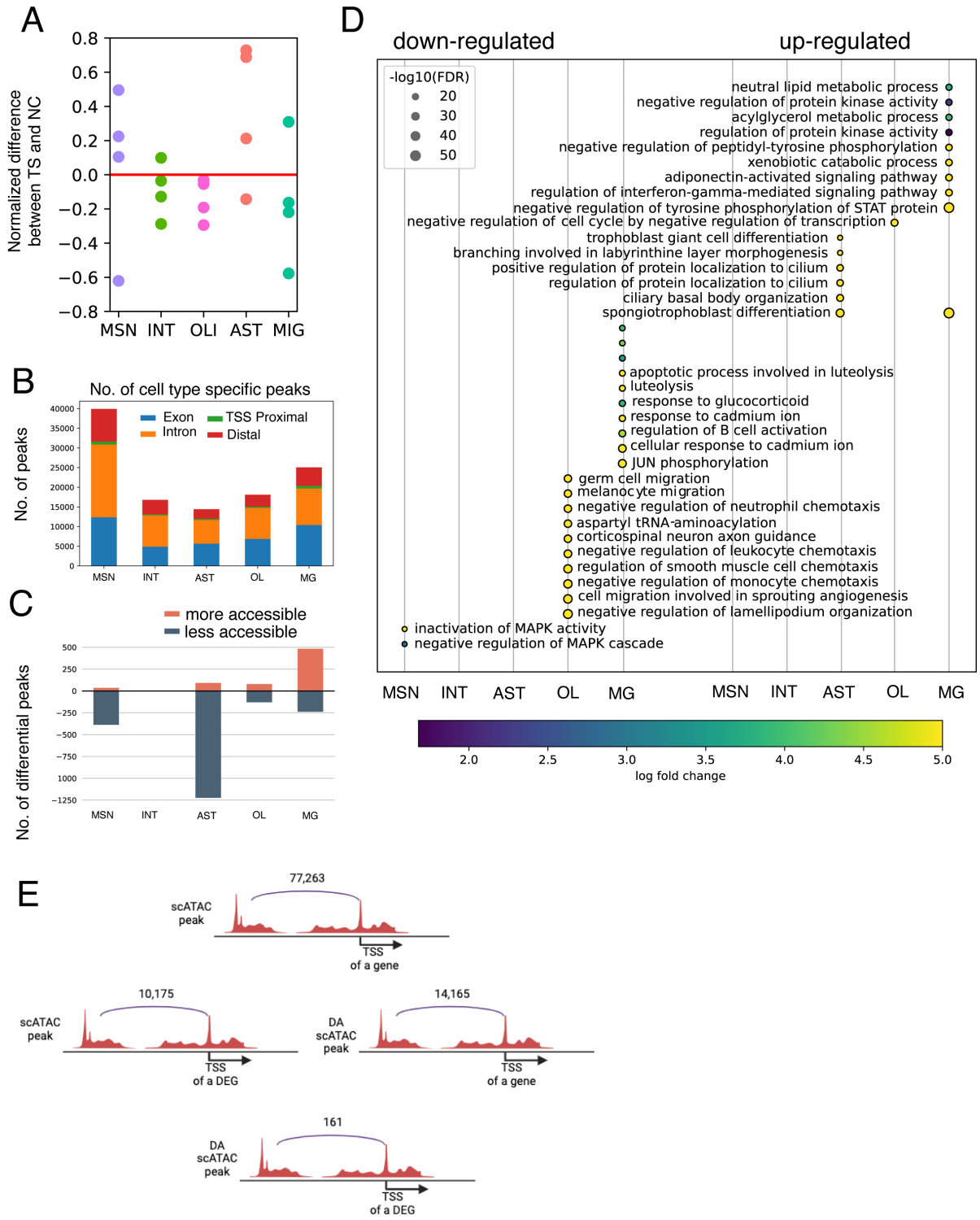

Figure 3. Tourette syndrome brains chromatin accessibility profiles compared to normal controls. A) Difference in number of cells in 5 cell types comparing TS to NC in snATAC-seq data. Y axis: normalized percentage difference between TS and NC; x-axis: cell types. **B)** Number of snATAC-seq peaks across genome annotations. Y-axis: number of peaks; x-axis: cell types. **C)** Number of differentially accessible peaks in Tourette syndrome brains. Y-axis: number of peaks; x-axis: cell types. **D)** GO biological process terms enrichment with the differentially accessible peaks in TS vs NC brains for each cell type. Size of the dots corresponds to -log10(FDR) of the enriched GO term, while color corresponds to log2 fold change of peak intensity. **E)** Number of co-accessible peaks discovered in snATAC-seq data connecting transcription start site (TSS) of genes (DEGs) with potential enhancers including differentially accessible (DA) peaks.

Discussion

In this study we perform a single cell-level transcriptome and epigenome analysis of the postmortem caudate nucleus and show that inhibitory interneurons are decreased, while projection neurons show signs of metabolic stress and mitochondrial dysfunction. At the same time, microglial cells are activated, displaying increased expression of immune system-related genes. These results fit with a scenario where decreased tonic inhibition by interneurons upon MSN causes hyperactivity of the latter, along with signs of metabolic stress.

The picture is consistent with the established role of the basal ganglia in the storage of motor program and modulation of motor behavior, where the globus pallidus (GPi) exerts a tonic inhibitory tone upon the thalamus and cortex, suppressing motor activity, and MSNs of the direct and indirect pathways integrate many inputs to modulate the GPi. The interneurons are a crucial component of such modulation as they receive from the cortex and thalamus and exert prominent feed-forward and feed-back inhibitory modulation upon the MSNs[26]. Hence the decrease in interneuron number observed in TS propagating through the basal ganglia circuitry is likely to be a crucial factor in the generation of tics.

Decreased number of interneurons (both cholinergic and GABAergic) has been previously reported in a small number of TS individuals (n=5 pairs) by immmunocytochemical studies[11]. A subsequent study of bulk RNA-seq performed in the striatum of 9 TS/NC pairs[12] suggested decrease inhibitory synaptic function (including a decrease in a number of synaptic channels and receptors) and increased immune activation but was unable to pinpoint the cellular origin of these alterations.

The current study on the single cell level, together with the previous evidence, provides a coherent picture of basal ganglia dysfunction by linking interneuron losses with potential hyper-excitability and stress of MSN. Downregulation of membrane/synapses related genes in interneurons additionally suggest decreased arborization and connectivity, further impairing the function of these cells. The correlation between changes in gene expression and chromatin state suggests that some of these changes in transcriptome are driven by altered regulatory element activity. Further explorations of noncoding element activity in TS may provide insights into upstream causes of dysregulated gene expression in patient's brains.

An important caveat of all postmortem studies is their descriptive nature, and particularly their inability to distinguish causes from consequences, which is noteworthy in chronic brain dysfunction. One pertinent question is whether interneurons are dying or fail to be generated in TS. Our prior studies in iPSC-derived brain organoids from living TS and NC individuals suggested that BG interneurons are not generated in normal numbers due an aberrant response to the ventralizing morphogen Sonic hedgehog (SHH) in embryonic development [28], although we cannot exclude further losses later in life.

Implicating loss and dysfunction of BG interneurons in TS has implications for treatment. Current treatment strategies with antipsychotic drugs are not addressing the pathophysiology of the disorder. Our data suggest that compounds that increase interneuron neurotransmission may represent promising candidates.

Acknowledgments
We acknowledge Livia Tomasini for technical advice on sample processing; the Tourette Association of America (TAA) for spearheading, organizing and supporting the postmortem brain collection for patients with Tourette syndrome. We thank the personnel at the Harvard Brain Tissue Resource Center/NIH NeuroBioBank for processing brain tissue; and Nancy Thompson for help with the clinical phenotyping. This work was supported in part by a fellowship grant from the Tourette Association of America (TAA) to L.F. and the National Institutes of Mental Health (R01 MH118453 to FMV). The Harvard Brain Tissue Resource Center is supported in part by the National Institutes of Health.

Methods

Subjects. TS brain samples of cases with refractory symptoms in adulthood, collected by the Tourette Association of America (TAA), were obtained from the Harvard Brain Tissue Resource Center. Samples of normal control subjects were obtained from the Harvard Brain Tissue Resource Center and from the Department of Pathology, Yale University (**Table S1**). Cases with known or suspected neurological diseases and prolonged hypoxia were excluded. There were no statistical differences in age, postmortem interval, or RNA quality between the TS and control groups (**Supplementary Table 1**).

Sample preparation and nuclei isolation. Bulk brain tissue (caudate nucleus and putamen) was cut into 2mm cubes and was lysed by mechanical dissociation in 6ml lysis buffer (5.47g sucrose, 250ul $CaCl_2$ (1M), 150ul $Mg(Ac)_2$(1M), 10ul EDTA, 500ul Tris-HCL (1M), 17ul DTT (3M), 50ul Triton X in 50ml RNAse and DNAse free water). Nuclei were passed through a 40um diameter strainer and the lysate was transferred into a 15ml ultracentrifuge tube (Beckman). 9ml of ice cold sucrose solution (30.78g Sucrose, 150ul $Mg(Ac)_2$ (1M), 500ul Tris-HCL (1M), 17ul DTT (3M) in 50ml RNAse and DNAse free water) was slowly added to the tube bottom. Nuclei were isolated by gradient ultracentrifugation under following conditions: 25,000rpm for 1h at 4° C.

Standalone single nuclei RNA-seq and single nuclei ATAC-seq library preparation and sequencing. For single nuclei RNA-seq: After ultracentrifugation, supernatant was aspirated and the nuclei resuspended in 1x PBS and RNAse inhibitor (Roche, RNAse inhibitor concentration according to suppliers' recommendation). For single nuclei ATAC-seq: After ultracentrifugation, supernatant was aspirated and the nuclei (3,080-7,700 nuclei/ul as recommended by 10X Genomics) were resuspended in 1X nuclei buffer (10X genomics). Samples were submitted to the Yale Center for Genome Analysis core for library preparation (10,000 nuclei/sample) using Chromium Single Cell ATAC reagent kit (10X Genomics).

Multiome single nuclei RNA-seq and ATAC-seq library preparation and sequencing. After ultracentrifugation, the supernatant was aspirated and the nuclei were resuspended in ice cold 1X nuclei buffer (10X nuclei buffer stock, provided by 10X genomics) and nuclei were handled according to supplier's recommendations. Samples were passed through a 40um diameter strainer. Nuclei were counted (Countess II, Thermo Fisher) and samples were submitted to Yale Center for Genome Analysis core for library preparation (10,000 nuclei/sample) using the Chromium Next GEM Single Cell Multiome ATAC + Gene Expression kit (10X Genomics) .

snRNA-seq data processing, clustering and annotation. The gene count matrices for standalone and multiome snRNA-seq data were calculated using cellranger-6.0.1 with gex-GRCh38 as reference genome and genome annotation. The ribosomal genes were then removed from the count matrices. The single nuclei libraries with an extremely low number of genes (<200) and an extremely high number of genes (>mean+2*standard deviation) were removed from the analysis. All 12 samples were then merged using Seurat 5.0.1 [27, 29, 30] merge function. Potential doublets in the snRNA-seq data were removed using solo [31]. The single nuclei libraries with too many mitochondrial gene expressions (>5%) were removed from the analysis. Expression of all genes for all the single nuclei were then normalized and integrated using the top 2000 expressed genes with Seurat normalization and integration functions. Dimension reduction was performed using UMAP with 50 dimensions as input features. Each cluster was annotated using literature [25, 32] and public datasets (https://www.brainspan.org) and clusters were combined into major cell types.

snATAC-seq data processing, clustering and label transfer. The fragment count matrices for standalone and multiome snATAC-seq data were calculated using cellranger ARC-2.0.0 with arc-GRCh38 as reference genome and genome annotation. The single nuclei libraries with an extremely low number of fragments (<3000) or an extremely high number of fragments (>20000) were removed. The low quality single nuclei with <15% fragments in the predicted peak regions and single nuclei with nucleosome signal < 2 or > 4 were removed based on the suggestion provided by Signac [33]. The count matrices of 12 samples were merged and integrated using features with more than 10 total counts. The integrated count matrix was then normalized and integrated for dimension reduction using UMAP with 30 dimensions as the input features. The common anchors (genes) were then identified between the snRNA-seq and snATAC-seq count matrices for label transferring using Signac[33].

snRNA-seq differential gene expression. Genes present in at least 10% of cells in either TS patients or controls were considered in the differential gene expression analysis. Differential gene expression analysis was performed using the glmgampoi [34] R package v1.13.2. For each cell type, pairs with one sample having an extremely low number of single nuclei were not considered in the analysis (pairs included in analyses in different cell types: MSN: all pairs; IN: pair 1,2,3,19; AST: all pairs; OLI: all pairs; MG: pair 2,3,4,19). Genes were considered as differentially expressed in a specific cell type when exhibiting an adjusted p-value < 0.01 and an absolute log2 fold change > 0.5. GO term analysis was performed using DAVID Functional Annotation[35].

snATAC-seq differentially accessible peaks and co-accessibility analysis. Peaks were called using CallPeaks function in Signac v1.12.0 [33] with MACS2 [36]. Differentially active peaks were then identified using the FindMarkers function by comparing each cell type to the rest of the cells as well as by comparing the single nuclei in TS patients to the single nuclei in controls. Peaks were considered as differentially accessible when their adjusted p-value was <= 0.05 and the log2 fold change was > 0.5. GO analysis was performed using GREAT [37, 38]. Co-accessible sites were identified using cicero v1.3.9 [39] with a score cutoff of 0.4 for considering significant connections between two peaks. The connections were then overlapped with transcriptional start sites (TSS) of genes to identify potential gene-enhancer links.

References

1. Leckman, J.F. Tourette's syndrome. *Lancet* **360**, 1577-1586 (2002).
2. Leckman, J.F. & Peterson, B.S. The pathogenesis of Tourette's syndrome: epigenetic factors active in early CNS development. *Biol Psychiatry* **34**, 425-427 (1993).
3. Davis, L.K.*, et al.* Partitioning the heritability of Tourette syndrome and obsessive compulsive disorder reveals differences in genetic architecture. *PLoS Genet* **9**, e1003864 (2013).
4. McNaught, K.S. & Mink, J.W. Advances in understanding and treatment of Tourette syndrome. *Nature reviews. Neurology* **7**, 667-676 (2011).
5. Leckman, J.F., Vaccarino, F.M., Kalanithi, P.S. & Rothenberger, A. Annotation: Tourette syndrome: a relentless drumbeat--driven by misguided brain oscillations. *J Child Psychol Psychiatry* **47**, 537-550 (2006).
6. Peterson, B.S.*, et al.* Basal Ganglia volumes in patients with Gilles de la Tourette syndrome. *Arch Gen Psychiatry* **60**, 415-424 (2003).
7. Peterson, B.S.*, et al.* Reduced basal ganglia volumes in Tourettes's syndrome using three-dimensional reconstruction techniques from magnetic resonance images. *Neurology* **43**, 941-949 (1993).
8. Bloch, M.H., Leckman, J.F., Zhu, H. & Peterson, B.S. Caudate volumes in childhood predict symptom severity in adults with Tourette syndrome. *Neurology* **65**, 1253-1258 (2005).
9. Hyde, T.M.*, et al.* Cerebral morphometric abnormalities in Tourette's syndrome: a quantitative MRI study of monozygotic twins. *Neurology* **45**, 1176-1182 (1995).
10. Kalanithi, P.S.*, et al.* Altered parvalbumin-positive neuron distribution in basal ganglia of individuals with Tourette syndrome. *Proc Natl Acad Sci U S A* **102**, 13307-13312 (2005).
11. Kataoka, Y.*, et al.* Decreased number of parvalbumin and cholinergic interneurons in the striatum of individuals with Tourette syndrome. *J Comp Neurol* **518**, 277-291 (2010).
12. Lennington, J.B.*, et al.* Transcriptome Analysis of the Human Striatum in Tourette Syndrome. *Biol Psychiatry* **79**, 372-382 (2016).

13. Albin, R.L., Young, A.B. & Penney, J.B. The functional anatomy of basal ganglia disorders. *TINS* **12**, 366-375 (1989).
14. Lapper, S.R., Smith, Y., Sadikot, A.F., Parent, A. & Bolam, J.P. Cortical input to parvalbumin-immunoreactive neurones in the putamen of the squirrel monkey. *Brain Res* **580**, 215-224 (1992).
15. Kita, H., Kosaka, T. & Heizmann, C.W. Parvalbumin-immunoreactive neurons in the rat neostriatum: a light and electron microscopy study. *Brain Res.* **536**, 1-15 (1990).
16. Kawaguchi, Y., Wilson, C.J., Augood, S.J. & Emson, P.C. Striatal interneurones: chemical, physiological and morphological characterization. *Trends Neurosci* **18**, 527-535 (1995).
17. Mallet, N., Le Moine, C., Charpier, S. & Gonon, F. Feedforward inhibition of projection neurons by fast-spiking GABA interneurons in the rat striatum in vivo. *J Neurosci* **25**, 3857-3869 (2005).
18. Gittis, A.H.*, et al.* Selective inhibition of striatal fast-spiking interneurons causes dyskinesias. *J Neurosci* **31**, 15727-15731 (2011).
19. Lapper, S.R. & Bolam, J.P. Input from the frontal cortex and the parafascicular nucleus to cholinergic interneurons in the dorsal striatum of the rat. *Neuroscience* **51**, 533-545 (1992).
20. Koos, T. & Tepper, J.M. Dual cholinergic control of fast-spiking interneurons in the neostriatum. *J Neurosci* **22**, 529-535 (2002).
21. Martos, Y.V., Braz, B.Y., Beccaria, J.P., Murer, M.G. & Belforte, J.E. Compulsive Social Behavior Emerges after Selective Ablation of Striatal Cholinergic Interneurons. *J Neurosci* **37**, 2849-2858 (2017).
22. Xu, M.*, et al.* Targeted ablation of cholinergic interneurons in the dorsolateral striatum produces behavioral manifestations of Tourette syndrome. *Proc Natl Acad Sci U S A* **112**, 893-898 (2015).
23. Rapanelli, M.*, et al.* Targeted Interneuron Depletion in the Dorsal Striatum Produces Autism-like Behavioral Abnormalities in Male but Not Female Mice. *Biol Psychiatry* **82**, 194-203 (2017).
24. Bandler, R.C.*, et al.* Single-cell delineation of lineage and genetic identity in the mouse brain. *Nature* **601**, 404-409 (2022).
25. Schmitz, M.T.*, et al.* The development and evolution of inhibitory neurons in primate cerebrum. *Nature* **603**, 871-877 (2022).
26. Albin, R.L., Young, A.B. & Penney, J.B. The functional anatomy of disorders of the basal ganglia. *Trends Neurosci* **18**, 63-64 (1995).
27. Hao, Y.*, et al.* Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat Biotechnol* (2023).
28. Brady, M.V.*, et al.* Mispatterning and interneuron deficit in Tourette Syndrome basal ganglia organoids. *Mol Psychiatry* **27**, 5007-5019 (2022).
29. Stuart, T.*, et al.* Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902 e1821 (2019).
30. Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol* **36**, 411-420 (2018).
31. Bernstein, N.J.*, et al.* Solo: Doublet Identification in Single-Cell RNA-Seq via Semi-Supervised Deep Learning. *Cell systems* **11**, 95-101 e105 (2020).
32. Bandler, R.C., Mayer, C. & Fishell, G. Cortical interneuron specification: the juncture of genes, time and geometry. *Curr Opin Neurobiol* **42**, 17-24 (2017).
33. Stuart, T., Srivastava, A., Madad, S., Lareau, C.A. & Satija, R. Single-cell chromatin state analysis with Signac. *Nat Methods* **18**, 1333-1341 (2021).
34. Ahlmann-Eltze, C. & Huber, W. glmGamPoi: fitting Gamma-Poisson generalized linear models on single cell count data. *Bioinformatics* **36**, 5701-5702 (2021).
35. Huang da, W.*, et al.* DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. *Nucleic Acids Res* **35**, W169-175 (2007).
36. Zhang, Y.*, et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137 (2008).
37. Tanigawa, Y., Dyer, E.S. & Bejerano, G. WhichTF is functionally important in your open chromatin data? *PLoS Comput Biol* **18**, e1010378 (2022).
38. McLean, C.Y.*, et al.* GREAT improves functional interpretation of cis-regulatory regions. *Nat Biotechnol* **28**, 495-501 (2010).

39. Pliner, H.A., *et al.* Cicero Predicts cis-Regulatory DNA Interactions from Single-Cell Chromatin Accessibility Data. *Mol Cell* **71**, 858-871 e858 (2018).