SUPPLEMENTAL MATERIALS

Participants

Thirteen participants with Tourette syndrome and 21 age-matched healthy control participants were scanned. Sample characteristics are detailed in Table 1 of the main text and Supplemental Table I. Tourette participants were recruited through the Tic Disorder Clinic of the Yale Child Study Center. Controls were recruited through phone calls to randomly selected names on a telemarketing list of families in the local community.

Six of the 13 Tourette participants had comorbid obsessive compulsive disorder (OCD), and one of these also had attention-deficit hyperactivity disorder (ADHD). All participants were free of other lifetime Axis I diagnoses. Six of Tourette participants were taking medications (4 SRIs, 1 clonazepam, 1 clonidine), and all of the Tourette patients reported experiencing premonitory urges prior to their tics. All participants had in common tics that were facial in location. However, they also had tics located in other body parts (Supplemental Table I). To ensure the validity and generalizability of our findings, we established rigorous criteria for the recruitment of the participants with Tourette syndrome. We required that Tourette participants had no large-amplitude tics of their head, neck, or shoulders that were of sufficient forcefulness to cause head movement during the scan. Tourette participants were required to have a right facial tic to ensure a comparable location in the body and brain for at least one tic across participants, thereby aiding identification of circuit-based activity for tics, and to ensure comparability of the behavior during spontaneous and mimicked tics. We also required them to be able to mimic their right-sided facial tic on demand so that we could identify voluntary motor circuit-based activity in the same body region involving the motor tic. This requirement, together with the general requirements to remain still during the scan, prompted us to study adults with Tourette syndrome, rather than participants of any age, so that we could be confident in their ability to comply with task instructions during the scan. In addition, we required that all Tourette participants reported experiencing premonitory urges prior to

their tics. Finally, they were not allowed to have comorbid lifetime Axis I disorders other than OCD or ADHD.

Neuropsychiatric diagnoses were established using the Schedule for Tourette and Other Behavioral Syndromes (1). The Yale Global Tic Severity Scale (2), Yale-Brown Obsessive Compulsive Scale (3), and ADHD rating scale (4) were used to rate current and worst-ever severity of tic, OCD, and ADHD symptoms, respectively. Socioeconomic status was quantified using the Hollingshead Four-Factor Index of Social Status (5).

Experimental Paradigms

Each experiment consisted of 2 types of scanning runs acquired in alternation, with 4 runs acquired for each, giving a total of 8 runs for each participant. For the participants with Tourette syndrome, odd numbered runs consisted of spontaneously generated tics on the right side of their face. Participants were instructed to let their tics happen of their own accord, and not to force or suppress them. They were also instructed not to move their heads, neck, or shoulders, and to refrain from making any other voluntary movements or behaviors during the scanning run. We termed these runs *"spontaneous tics"*. During even numbered runs, which we termed *"voluntary tics"*, participants with Tourette syndrome were instructed to produce a facial tic intentionally and at a self-selected pace that would eliminate as much as possible any urge to perform a tic in that muscle group. They were also instructed to allow any other tics to happen as they would normally, without actively suppressing them, and they were instructed to refrain from making any other voluntary movements or behaviors during scanning. The aim of this instruction was to generate voluntary motor activity for the same behavior as the spontaneously generated tics (similar in force, amplitude, and duration during the two conditions). Visual inspection ensured that voluntary tics were indistinguishable from spontaneous tics. Comparing neural activity during spontaneous vs. voluntary tics allowed discrimination of the neural mechanisms that generate tics, but not voluntary movements, and vice-versa, for the same motor behavior.

For the healthy control participants, odd numbered runs consisted of mimicking a facial tic of their right cheek. They were instructed before the scan by example on how to do this, and they practiced it

until the behavior was visually indistinguishable from a real tic. During the scan, they performed this mimicked facial tic with the instruction to perform a single tic at a time of their choosing (self-paced) but separated in time, so that they occurred roughly every 15-45 seconds. We termed these runs *"self-paced mimicked tics"*. The timing of these mimicked tics was recorded during each run (see below) and converted into a sound file that, during even runs, produced beeps at exactly the same times as the participant produced them during the self-paced mimicking tic runs. These sounds were presented to the participant during the next even numbered run via MRI-compatible headphones. Participants were instructed to mimic the facial tic whenever they heard a beep (cue-paced), but they were not told that the timing of the beeps had been determined by the timing of the mimicked tics in the previous odd numbered run. We termed these runs *"cue-paced mimicked tics"*. These runs were designed to identify the neural systems that subserve voluntary behaviors in healthy individuals but that are indistinguishable behaviorally from the real tics that affect persons with Tourette syndrome, and in a body region in which tics most commonly occur (facial tics were present in each of the participants with Tourette syndrome). We included this condition so that we could attempt to distinguish the circuits that produce a spontaneous, internally generated, voluntary behavior (i.e., 'self-paced' behavior) from an identical behavior that is dictated by a sensory cue (i.e., 'cue-paced' behavior). This contrast in healthy controls was thus intended to isolate the circuits involved in the internal generation of a simple normal voluntary motor behavior. We speculated that if the contrast of activity during spontaneous and voluntary tics in the Tourette group identified similar circuits as the contrast of activity during self-paced and cue-paced mimicked tics in healthy controls, we would have evidence that tic behaviors are generated by sensory cues.

Participants were trained prior to scanning so that the voluntary tics (in the Tourette group) and the mimicked tics (in controls) were visually indistinguishable from spontaneous tics. During the scanning session for each participant, a member of the study team was positioned at the bore of the scanner to observe the face both directly and through the mirror located on the bird cage RF head coil to ensure that during the even and odd numbered runs, the patterns of muscle contractions of the mimicked

tics produced by the controls were indistinguishable from the patterns of the real tics produced by the patients.

The odd and even numbered runs were intended to distinguish circuits that produce behaviors that are either self-generated or dictated by sensory cues. Neural activity in control participants during the generation of normal, spontaneous, and voluntary behaviors that look like tics compared with activity of the same behaviors that are generated by a sensory cue can identify circuits that generate behavior in response to sensory cues. Comparing activity in those circuits of healthy controls with activity during the generation of spontaneous tics in the Tourette group should aid our identification and understanding of the role of sensory systems in generating tics and voluntary tic-like behaviors.

Image Acquisition

Imaging was performed using a GE 1.5T Signa Scanner (Milwaukee, WI). Head positioning in the magnet was standardized using the canthomeateal line. A T_1-weighted sagittal localizing scan was used to position the axial images. In all participants, a high-resolution T_1-weighted structural image with 10 axial slices (acquisition matrix 256x256; field of view 20x20 cm) was positioned oriented parallel to the anterior commissure-posterior commissure (AC-PC) line. Functional images consisted of 10 axial slices acquired using a gradient-recalled single-shot echo-planar pulse sequence with repetition time (TR)=1200ms, echo time (TE)=60msec, flip angle=60^0, matrix=128 x 64, FOV=40x20 cm, providing an in-plane resolution of 3.125 x 3.125mm, with 102 images acquired per run, and 4 runs acquired for each run type (for Tourette participants, we had two run types, *spontaneous tics* and *voluntary tics*; for control participants, we had two matched run types, *self-paced mimicked tics* and *cue-paced mimicked tics*). The slice thickness was a constant of 7 mm, while the skip among slices varied from 0.5 to 2 mm to maintain a strict correspondence with the Talairach coordinate system, such that each axial slice represented the same plane within stereotaxic space across participants. Our sequence emphasized temporal resolution at the expense of spatial resolution, because short TRs are necessary for Granger causality analyses. We acquired a total of four types of functional imaging datasets corresponding to the four run types for the

two groups of participants (healthy controls and Tourette patients), each with total 408 functional images from four repeated runs for each participant.

Image Analysis

Image analysis consisted of five procedures: a) preprocessing of anatomical and functional imaging data; b) spatial Independent Component Analysis (ICA) of the preprocessed data; c) identification of reproducible ICA components; d) statistical comparison of the processed ICA components; and e) Granger causality analysis of the ICA components.

Preprocessing We performed standard preprocessing of the echoplanar images using our GUI-based batch-mode software package with integrated functions drawn from both SPM and FSL, including slice timing correction to correct phase shifts between slices caused by interleaved scans (phase shifts were corrected using sinc interpolation between time points at each voxel); motion correction (via a rigid-body transform with 3 translation and 3 rotation parameters to minimize mean square errors between each image and the template image); spatial normalization onto the Montreal Neurological Institute (MNI) template ICBM152; and spatial filtering to remove spatial noise (via a Gaussian kernel function with a full width at half maximum of 8mm).

For spatial normalization, the echo planar imaging (EPI) data for each participant were first coregistered with the high resolution anatomical image of the same participant, and then the anatomical image was normalized to the MNI template by combining an affine and a discrete cosine transform for global linear and local nonlinear normalization. Finally, the EPI data were normalized to the same MNI template, with a voxel size of 3x3x3 mm^3 and an image size of 53x63x46 voxels, using the parameters generated when normalizing the anatomical image to the MNI template.

Independent Component Analysis ICA was then performed on each of the four types of functional imaging datasets that had undergone the above preprocessing procedures to explore the neural activity associated with generating or mimicking tics. The post-processing of ICA components included spatially clustering components across datasets to identify components that were reproducible in their spatial configuration both across individuals and across run types.

We concatenated together, in ascending order, the 4 repeated runs of data within each run type for each participant to form a matrix X with a size of 408x153594 voxels, where 408 = 4(runs)x102(images per run) and 153594=53x63x46 (voxels per image). We excluded those voxels outside of the brain within this matrix, thereby reducing memory demands for subsequent analyses, by generating a masked image Y of X using a procedure for mask estimation that was modified from an SPM2 function (spm_spm) (http://www.fil.ion.ucl.ac.uk/spm/). We then applied Principal Component Analysis (PCA) to reduce the temporal dimension of Y from 408 time points to *N* time points, where *N* represented the number of independent components to be generated (We describe in the next paragraph how to determine an optimal *N*). Spatial ICA was then performed on the reduced *N* time-point dataset for each run type (spontaneous tics, voluntary tics, self-paced mimicked tics, or cue-paced mimicked tics) in each participant using the ICA algorithm "FastICA" (6). Briefly, we described how we calculated PCA and ICA as follows:

Let Y denote a matrix with *N* rows and *Q* columns, where *N* represents the number of time points (which was 408 for our raw data) and *Q* represents the number of total valid voxels after applying a spatial mask to the raw data. If we treat each of the N rows of Y as a spatial signal, then each of those spatial signals represents a spatial pattern of BOLD intensity at time t (where t varies from 1 to 408). Because of the slow fluctuations in BOLD signal intensity, these spatial patterns are highly similar to one another (i.e., they are statistically redundant over time), and thus we can compress Y along the time dimension N using PCA. Nevertheless, the spatial patterns of BOLD signal intensity will still vary over time with changes in underlying neural activity, albeit with a much poorer temporal resolution than the actual underlying neural activity. Moreover, the BOLD signal measured at any one point in the brain may derive not only from the neurons closest to that point, but from other neurons nearby. The BOLD signal thus represents a superposition or mixture of neuronal sources. It also contains an admixture of varying sources of noise. We can formulize the contributions to the spatially varying BOLD signal Y with the following linear model:

$$Y = AS \qquad \text{(a)}$$

where S denotes a matrix with *M* rows and *Q* columns representing *u* independent neural sources and *v* independent noise sources (*u*+*v* = *M*), and *A* denotes a matrix with *N* rows and *M* columns representing the coefficients that govern how these neural and noise sources are linearly mixed to generate Y. Our goal is to find S from Y. However, we know only Y, and both A and S are unknown. Finding S in this circumstance has been termed a blind source separation (BSS) problem, and ICA has been used to solve it (7).

As noted above, prior to ICA we performed PCA on the mean-removed functional imaging data. PCA was used not only for the purposes of data reduction, but also to reduce the complexity of the problem of determining the matrix A and thereby the independent components S (and therefore the *u* independent neural sources) for an iterative ICA algorithm, as we describe now briefly. Assume that we have removed the mean in the spatially varying BOLD signal Y of Equation (a). Applying PCA to Y is equivalent to applying a linear transform to Y. Let U denote such a linear transform, and let us apply U to both sides of Equation (a), such that UY = UAS = BS (i.e., B=UA). Let Z = UY. After PCA, the autocorrelation matrix of Z, $E\{ZZ^T\}$, equals *I*, where *I* denotes a unitary matrix and $E\{.\}$ denotes the mathematical expectation of a random variable. In signal processing theory, a random variable that has a unitary autocorrelation matrix is termed a "whitened" random variable. Because PCA ensures that Z has a unitary autocorrelation matrix, PCA is often viewed as a data whitening procedure. Now, we have

$$E\{\mathbf{ZZ}^T\} = E\{\mathbf{BSS}^T\mathbf{B}^T\} = \mathbf{B}E\{\mathbf{SS}^T\}\mathbf{B}^T = \mathbf{BB}^T = \mathbf{I} \qquad \text{(b)}$$

where $E\{\mathbf{SS}^T\} = \mathbf{I}$, since we have assumed that components in S are independent of each other and thus S should also have an unitary autocorrelation matrix. Thus, B is an orthogonal matrix, because $\mathbf{BB}^T = \mathbf{I}$. We have therefore reduced the difficulty of finding an arbitrary matrix A in Equation (a) to finding an orthogonal matrix B such that

$$Z = BS \qquad \text{(c)}$$

Thus, $S = B^TZ = B^TUY$, where the superscript ^{T} denotes a matrix transpose, and matrix U is a PCA transform (or a whitening matrix). To find B and thereby to allow us to determine S, a number of ICA

algorithms have been proposed, one of which is a fast fixed-point algorithm designated *FastICA* (6). We used the FastICA algorithm in our data analysis.

We briefly describe now the principles upon which the FastICA algorithm is based. Let W and $\hat{\mathbf{S}}$ denote the estimations of B^T and S, respectively. We then have

$$\hat{\mathbf{S}} = \mathbf{WZ} \tag{d}$$

In a practical implementation of this iterative ICA algorithm, we estimate W row-by-row, and thus we estimate the independent components S one-by-one. Because B is an orthogonal matrix, the matrix W (which is as an estimation of B^T) must be also orthogonal. Thus, a Gram-Schmidt orthogonalization procedure is usually embedded into the iterative ICA algorithm. Let w_k denote a vector taken from the *k*-th row of W, and let $\hat{s}_k$ denote the corresponding the *k*-th estimated independent component. Then,

$$\hat{s}_k = w_k \mathbf{Z} \tag{e}$$

Substituting Z in Equation (e) with Z=BS from Equation (c), we have

$$\hat{s}_k = w_k \mathbf{Z} = w_k \mathbf{BS} = v_k \mathbf{S} \tag{f}$$

where $v_k = w_k \mathbf{B}$. We now describe how we estimate w_k based on the relations described in Equation (f). From $\hat{s}_k = v_k \mathbf{S}$, we know that $\hat{s}_k$ is a linear combination of multiple independent components given by S. According to the *central limit theorem* from probability theory, which states that the distribution of a sum of independent random variables tends toward a Gaussian distribution, we know that $\hat{s}_k$ tends to be Gaussian distributed. Based on this theorem, the sum of multiple independent random variables is more Gaussian than is any one of them individually. Thus, the distribution of $\hat{s}_k$ is maximally non-Gaussian only if v_k has one nonzero value and hence if $\hat{s}_k$ equals one of S — one of its independent components. Therefore, by maximizing the non-Gaussianity of $\hat{s}_k$, or equivalently by maximizing $w_k \mathbf{Z}$ since $\hat{s}_k = w_k \mathbf{Z}$, we can find w_k based on Z. We can use the fourth-order cumulant, *kurtosis*, to measure quantitatively the non-Gaussianity of the random variable $\hat{s}_k$:

$$kurt(\hat{s}_k) = kurt(w_k \mathbf{Z}) = E\{(w_k \mathbf{Z})^4\} - 3[E\{(w_k \mathbf{Z})^2\}]^2 \quad \text{(g)}$$

To maximize $kurt(\hat{s}_k)$, we first calculate the first-order derivative of the absolute value of $kurt(\hat{s}_k)$:

$$\frac{\partial |kurt(\hat{s}_k)|}{\partial w_k} = 4sign(kurt(w_k \mathbf{Z}))[E\{\mathbf{Z}(w_k \mathbf{Z})^3\} - 3\|w_k\|_2^2 w_k] \quad \text{(h)}$$

Based on this equation, we can estimate w_k via an iterative procedure:

$$w_k^{j+1} = E\{\mathbf{Z}(w_k^j \mathbf{Z})^3\} - 3\|w_k^j\|_2^2 w_k^j \quad \text{(i)}$$

where *j* denotes the *j*-th iteration. If w_k^j is divided by its 2-norm in each iterative step, we can rewrite (i) as

$$w_k^{j+1} = E\{\mathbf{Z}(w_k^j \mathbf{Z})^3\} - 3w_k^j \quad \text{(j)}$$

By randomly initializing w_k^j at *j*=0, we can iteratively estimate w_k^{j+1} using Equation (j) until w_k converges. These procedures summarize the basic principles of the FastICA algorithm (6) that we used to generate the independent components on each of our 4 run types of fMRI data.

Identification of Corresponding Independent Components across Datasets and across Participants Our aim was to compare neural activity in independent components that were reproducible across run types (i.e. spontaneous tics, voluntary tics, self-paced mimicked tics, or cue-paced mimicked tics) and across groups of participants (Tourette and control groups). We have previously reported validation of an algorithm, which we termed "Partner Matching" (PM) (8), that identified reproducible components across individuals, assuming that the number of components to be generated in each person had already been determined. One cannot know *a priori*, however, the single number (*N*) of components to generate with ICA that is "optimal" for the identification of reproducible components across run types and across individuals. We therefore used the principles of *Information Criteria* to determine the number of sets of independent components, *N*, to generate from each run type. We combined *Minimum Description Length* (MDL) (9, 10) and *Akaike's Information Criterion* (AIC) (11) to estimate the lower and upper bounds, respectively, of the numbers of independent components to generate for all 4 run types. Thus, these lower

and upper bounds defined an interval for the number *N* of components to generate. For these datasets, that interval was [20:130]. Unlike conventional ICA, however, in which only one set of independent components is generated (we term this "single-set ICA", or SS-ICA), we generated multiple sets of components for each run type in each participant, which we term "multiple-set ICA" (MS-ICA). For each participant, with an increment of 10 over the interval [20:130], we used the FastICA algorithm to generate 12 sets of independent components, with each of the 12 sets containing either 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, or 130 components, respectively. We then applied our PM algorithm to identify the independent components that corresponded in their spatial locations (i.e., that were "reproducible") across each of these 12 sets of independent components. We then combined the matched components across the 12 sets so that we could identify those independent components that were most reproducible across run types and participants, independent of the choice of the number of independent components that were generated. We term this modified algorithm "Hierarchical Partner Matching" (H-PM).

Hierarchical Partner Matching Algorithm for the Identification of Reproducible Multiple-Set ICA Components In our original PM algorithm (8), we defined a *family* as a basic dataset on which we performed ICA. In this study, we acquired 4 repeated runs of functional data for each of the 4 run types (2 run types per group: spontaneous tics and voluntary tics in Tourette participants, and self-paced mimicked tics and cue-paced mimicked tics in controls). We concatenated the 4 repeated runs for each run type, and then we applied ICA to each of these concatenated data types. We call the concatenated data for each run type a basic dataset or *family*. Note that it is not meaningful to assess whether an independent component is reproducible among the other components that are generated from the same family, as these components by definition and construction are independent of one another. Instead, we identify reliable components using a strategy to find those components that are significantly reproducible among different families of datasets acquired either from different participants or from different run types within the same participant. In the present study, we had data from 13 participants in the Tourette group and 21 participants in the control group. Each of the participants provided 2 run types corresponding to 1 of the 4 tasks, yielding a total of 13*2+21*2 = 68 different families (Supplemental Figure 2a). For each of these

68 families, we generated 12 sets of independent components, with each of the 12 sets containing either 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, or 130 components, as described above. We have demonstrated previously that assessing whether a matching is bidirectional (i.e., assessing a *"Partner Matching"* -- see Supplemental Figure 1) is a useful and valid way to test whether an independent component is reproducible between families (8).

Briefly, our Partner Matching algorithm works as follows. Given a component *i* from family A, we calculate indices of spatial similarity between independent component *i* from family A and each independent component in family B. We use Tanimoto Distance to quantify this spatial similarity between components (8). We then select the component from family B, say component *j*, that is determined to have the maximum similarity index to component *i* among all the components of family B. This identifies component *j* in family B as being the component that maximally matches component *i* from family A in its spatial configuration. However, this simply provides a unidirectional matching, because we have not yet investigated all the similarity indices between component *j* from family B and each of the components in family A. Thus, component *i* from family A may not necessarily be the maximal match for component *j* from family B. To test whether component *i* and component *j* match maximally in both directions, we also need to calculate all the similarity indices between component *j* from family B and each of the components in family A. We then select the component from family A, say component *k*, that is determined to have the maximum similarity index to component *j* among all the components of family A. If $k=i$, then the matching is bidirectional, and we consider this component in family A and this component in family B to be partner-matched (Supplemental Figure 1). We repeat this procedure to find all pairs of components that are bidirectionally matched between families A and B. When components are partner-matched between families A and B, the matching results are independent of the order of matching from A to B or from B to A. This matching strategy identifies similar components between families much more reliably than can a simple unidirectional matching. Bidirectional partner matching ensures that each component in family A is only matched with one component in family B (and vice-versa), because components are only matched if they are maximally

similar in both directions. Different components of family A (e.g., components a1, a2, ...) may match different components of family B (e.g., components b1, b2, ...). For example, components *i* and *p* of family A may match components *j* and *q* from family B.

We apply our algorithm to identify matching components not only across two families, but across multiple families. We call a collection of components that match across families a *cluster* (see Supplemental Figure 2a). For example, a cluster may contain component *i* of family A, component *j* of family B, component *m* of family C, etc., all of which are matched to one another according to the partner matching procedure described above. Therefore, each of these clusters contains the information of component correspondences between families. This correspondence information is not provided by ICA itself.

Supplemental Figures 1 and 2

Because we did not know *a priori* the correct number of components to generate for each participant and each run type, and because that number may in fact vary across participants and run types, we extended the partner matching algorithm described above to work with multiple sets of components, each of which had a different number of components (20, 30, 40, . . ., 130). We first applied partner matching to the independent components within each set to identify clusters of components that were reproducible across families, and then we applied it across sets to identify the most reproducible clusters. We designate this procedure *Hierarchical Partner Matching* (H-PM), to emphasize the hierarchically organized, two-step procedure to identify the independent components that match across participants and across run types (Supplemental Figure 2). In the first step of our H-PM algorithm, we applied PM separately within each set of independent components that were generated for each of the 68 families so as to identify multiple collections of components that were significantly reproducible across families of components. Each collection of such components, which we term a *cluster*, consisted of those components that were identified as being significantly similar to one another (i.e., they were significantly "reliable" or "reproducible") in their spatial configuration (Supplemental Figure 2a). We evaluated the significance of each cluster of independent components using a standard measure of cluster consistency,

Cronbach's Alpha (CA), which measures simultaneously the similarity and reproducibility (or the "consistency") of the independent components within the cluster (8). CA uses two parameters in evaluating cluster consistency, which in our case were (1) the number of independent components identified as similar to one another within a cluster, and (2) the average of the similarity indices between every pair of these components. As either of these two parameters increases – i.e., when a component partner-matches across more families, or when the similarity indices for that matching increases - the CA also increases. Therefore, the CA simultaneously measures the degree of similarity and the reproducibility across families of the independent components within the cluster. Following the standard convention for thresholding of this metric, we considered those clusters of components for which $CA>0.6$ as being sufficiently similar in their spatial configuration and reproducible across families (12).

For each component, ICA produces a set of *coefficients* – one per voxel – that represent the extent to which the temporal profile of activation of that voxel resembles the temporal profile captured by the independent component. The values of these coefficients, however, need to be converted to *z* scores so that their significance can be assessed (7, 8). We therefore converted the coefficients of each identified independent component into a z-score map by subtracting the global mean of the coefficients (across all the voxels of the brain) from each voxel's coefficients and then dividing the global standard deviation of the coefficients (7, 8). The z-score maps of independent components are equivalent to the contrast maps detected using conventional GLM-based analyses in that both maps indicate which brain regions are activated under the tasks (13). We then performed a 1-sample t-test on the z-score maps of independent components within each reliable cluster for each set to obtain the map of significant voxels for that cluster, which we then thresholded at an uncorrected $p<0.001$. This map, which we designate a "cluster map," represented a spatial pattern that tended to be present across families for this cluster (Supplemental Figure 2a).

Supplemental Figure 2

In the second step of our H-PM algorithm (Supplemental Figures 2b), we applied PM to these averaged component maps (cluster maps) across the 12 sets of clusters (each of which identified

components that were reproducible across families), to identify corresponding clusters across the different sets. This procedure effectively identified "clusters of clusters," or clusters that were partner-matched across sets. Whereas partner matching in the first step (or level) matched components across families, thereby identifying clusters of corresponding components, partner matching in this second step matched clusters across sets, thereby identifying clusters of corresponding clusters. For each cluster of clusters, we selected as output the cluster that had the highest CA (the index that we used to measure similarity and reproducibility of components within a cluster). As mentioned above, the appropriate number of independent components may vary across families (i.e., across participants and run types). Thus, if the cluster with the highest CA did not include all families, we determined whether the missing families were present in clusters that were partner-matched with this cluster (i.e., that were in the same cluster of clusters). If so, we added the corresponding component information to the optimal cluster. Following these procedures, we detected 15 reliable clusters (collections of independent components that were similar to each other in the configurations of their spatial maps) across each of the four run types (spontaneous tics and voluntary tics in the Tourette group, and self-paced mimicked tics and cue-paced mimicked tics in the control group).

Assessment of Effective Connectivity Between Independent Components Using Granger Causality

We used the Granger Causality Index (GCI) to assess connectivity for a set of connections identified a priori based on known neuroanatomical connections (Fig. 4). Let X(t) denote the time course of an independent component within region X and Y(t) denote the time course of another independent component within region Y. $X(t)$ was estimated using an autoregressive (AR) model as follows:

$$X(t) = a_1 X(t-1) + a_2 X(t-2) + \cdots + a_d X(t-d) + e_X = \sum_{m=1}^{d} a_m X(t-m) + e_X \qquad (1)$$

where $a_1, a_2, \cdots, a_d$ denote d AR coefficients; d denotes the order of the AR model, which was estimated based on the first local minima of the values of Bayesian Information Criterion (BIC) (14, 15) (which was defined as $BIC = k\ln(n) - 2\ln(L)$, where k denotes the model order; n denotes the length of X(t), and L

denotes the maximized value of likelihood function for the estimated model) ; e_X denotes the error term. We then added $Y(t)$ into the nonlinear AR model given in (1), yielding:

$$X(t) = \sum_{m=1}^{d} a_m X(t-m) + \sum_{n=1}^{d} b_n Y(t-n) + e_{XY} \quad (2)$$

where $b_1, b_2 \cdots, b_d$ denote *d* AR coefficients, e_{XY} denotes the new error term. To assess whether the addition of $Y(t)$ improves the prediction compared with the use of *X* alone, we calculate the causality index, $GCI_{Y->X} = 1 - Var(e_{XY}) / Var(e_X)$, where $Var(e_X)$ and $Var(e_{XY})$ denote the variance of the estimation errors e_X and e_{XY}, respectively. If $GCI_{Y->X}$ is greater than 0, then the addition of the previous values of $Y(t)$ into the right side of Equation (1) significantly improves the prediction of the current values of *X(t)* and deem that $Y(t)$ Granger caused $X(t)$, that is, region Y has a direct causal influence to region X. An indirect influence of region Y to region X via the third region Z was assessed by the conditional causality index, $GCI_{Y->X|Z} = 1 - Var(e_{XY|Z}) / Var(e_{XY})$ where $Var(e_{XY})$ and $Var(e_{XY|Z})$ denote the variance of the estimation errors e_{XY} and $e_{XY|Z}$, respectively. And e_{XY} was the error term when auto-regressively modeling *X(t)* using *d* previous values of *X(t)* and *Y(t)*; $e_{XY|Z}$ was the error term when auto-regressively modeling *X(t)* using *d* previous values of *X(t)*, *Y*(t) and *Z(t)*. If $GCI_{Y->X|Z} > 0$, an indirect influence is exerted from region Y to region X through region Z.

Because GCIs were not normally distributed, we used non-parametric methods to conduct statistical comparisons of the GCIs. We used the Wilcoxon signed rank test to test whether the medians of the GCIs were significantly different from zero within each run type for each diagnostic group; we used Wilcoxon rank sum test to test whether the medians of the GCIs were significantly different between two diagnostic groups on the same run type or between two different run types within the same diagnostic group. The data were represented by the median and the inter-quartile range (IQR). A probability of p smaller than 0.05 was considered significant to reject the null hypothesis that the medians are zeros in Wilcoxon signed rank test or that the medians are equal in Wilcoxon rank sum test. We also computed the

Spearman's rank correlation ρ between the GCIs and the symptom severity scores of tics within the Tourette's group, where a correlation was considered indicative of a significant association between the GCIs and the severity scores when the corresponding p-values were smaller than 0.05.

Supplemental References

1. Pauls DL,Hurst CR: Schedule for Tourette and Other Behavioral Syndromes, New Haven, Conn., Yale University Child Study Center, 1996

2. Leckman JF, Riddle MA, Hardin MT, et al: The Yale Global Tic Severity Scale: initial testing of a clinician-rated scale of tic severity. J Am Acad Child Adolesc Psychiatry 1989; 28:566-573

3. Goodman WK, Price LH, Rasmussen SA, et al: The Yale-Brown Obsessive Compulsive Scale. I. Development, use, and reliability. Arch Gen Psychiatry 1989; 46:1006-1011

4. Barkley RA, DuPaul GJ,McMurray MB: Attention deficit disorder with and without hyperactivity: clinical response to three dose levels of methylphenidate. Pediatrics 1991; 87:519-531

5. Hollingshead AB: Four factor index of social status, Yale University, 1975

6. Hyvarinen A: Fast and robust fixed-point algorithms for independent component analysis. IEEE Trans Neural Netw 1999; 10:626-634

7. McKeown MJ, Makeig S, Brown GG, et al: Analysis of fMRI data by blind separation into independent spatial components. Hum Brain Mapp 1998; 6:160-188

8. Wang Z,Peterson BS: Partner-matching for the automated identification of reproducible ICA components from fMRI datasets: algorithm and validation. Hum Brain Mapp. 2008; 29:875-893

9. Calhoun VD, Adali T, Pearlson GD, et al: A method for making group inferences from functional MRI data using independent component analysis. Hum Brain Mapp. 2001; 14:140-151

10. Rissanen J: A universal prior for integers and estimation by minimum description length. Ann Statics 1983; 11:416-431

11. Akaike H: A new look at the statistical model identification. IEEE Trans Automatic Control 1974; 19:716- 723

12. Hair JE, Anderson, R. E., Tatham, R. L. and Black, W. C.: Multivariate Data Analysis, 5. New Jersey, Prentice-Hall, 1998

13. Quigley MA, Haughton VM, Carew J, et al: Comparison of independent component analysis and conventional hypothesis-driven analysis for clinical functional MR image processing. AJNR Am J Neuroradiol 2002; 23:49-58

14. Roebroeck A, Formisano E,Goebel R: Mapping directed influence over the brain using Granger causality and fMRI. Neuroimage 2005; 25:230-242

15. Schwarz G: Estimating the dimension of a model. Annals of Statistics 1978; 6:461-464

SUPPLEMENTAL TABLES

Supplemental Table I: Prevalence of the Various Types of Tics in the 13 Participants with Tourette Syndrome

Simple Motor Tics	**# Tourette Participants**
1 eye blinking	8
2 eye movements	5
3 nose movements	6
4 mouth movements	5
5 facial grimace	7
6 head jerks/movements	5
7 shoulder shrugs	5
8 arm movements	4
9 hand movements	3
10 abdominal tensing	5
11 leg foot or toe movements	4
12 other	3
Complex Motor Tics	
1 Eye movements	2
2 Mouth movements	1
3 Facial movements or expression	2
4 Head gestures or movements	3
5 Shoulder movements	2
6 Arm movements	1
7 Hand movements	2
8 Writing tics	0
9 Dystonic postures	0
10 Bending or gyrating	0
11 Rotating	0
12 Leg or foot or toe movements	1
13 Blocking	0
14 Tic related compulsive behavior (touching, tapping, grooming, evening-up)	5
15 Copropraxia	0
16 Self-abusive behavior	1
17 Paroxysms of tics (displays)	1
18 Disinhibited behavior	0

Supplemental Table II: Granger Causality Indices of the Interregional Connections Defined in Figure 4

	NC Self-paced Mimicked Tics	NC Cue-paced Mimicked Tics	TS Spontaneous Tics	TS Voluntary Tics
PFC→PMA	0.013 (0.002-0.056), p=5.96E-5	0.008 (0.002-0.023), p=8.86E-5,	0.009 (0.007-0.075), p=2.44E-4,	0.028 (0.015-0.054), p=2.44E-4,
PFC→SMA	0.020 (0.007-0.027), p=5.96E-5,	0.012 (0.004-0.019), p=5.96E-5,	0.035 (0.015-0.057), p=2.44E-4,	0.037 (0.019-0.072), p=2.44E-4,
PFC→Caudate	0.014 (0.007-0.045), p=5.96E-5,	0.008 (0.002-0.024), p=5.96E-5,	0.019 (0.004-0.055), p=2.44E-4,	0.012 (0.007-0.062), p=2.44E-4,
PMA→SMA	0.008 (0.004-0.034), p=5.96E-5,	0.006 (0.002-0.036), p=8.86E-5,	0.031 (0.024-0.088), p=2.44E-4,	0.046 (0.024-0.162), p=2.44E-4,
PMA→M1	0.012 (0.003-0.031), p=5.96E-5,	0.024 (0.002-0.043), p=1.32E-4,	0.021 (0.013-0.106), p=2.44E-4,	0.069 (0.041-0.176), p=2.44E-4,
SMA→PMA	0.011 (0.004-0.018), p=5.96E-5,	0.018 (0.006-0.043), p=8.86E-5,	0.023 (0.009-0.067), p=2.44E-4,	0.016 (0.013-0.043), p=2.44E-4,
SMA→M1	0.012 (0.005-0.024), p=5.96E-5,	0.010 (0.005-0.028), p=8.86E-5,	0.027 (0.007-0.043), p=2.44E-4,	0.024 (0.007-0.091), p=2.44E-4,
M1→Putamen	0.013 (0.004-0.042), p=5.96E-5,	0.014 (0.008-0.055), p=1.32E-4,	0.031 (0.007-0.093), p=2.44E-4,	0.039 (0.015-0.125), p=2.44E-4,
S1→M1	0.012 (0.004-0.067), p=5.96E-5,	0.019 (0.006-0.045), p=8.86E-5,	0.071 (0.021-0.146), p=2.44E-4,	0.065 (0.017-0.201), p=2.44E-4,
S1→Putamen	0.012 (0.007-0.027), p=5.96E-5,	0.015 (0.007-0.042), p=1.32E-4,	0.022 (0.009-0.144), p=2.44E-4,	0.065 (0.015-0.152), p=2.44E-4,
Caudate→Pallidum	0.031 (0.014-0.057), p=5.96E-5,	0.024 (0.006-0.069), p=5.96E-5,	0.085 (0.010-0.220), p=2.44E-4,	0.033 (0.011-0.064), p=2.44E-4,
Putamen→Pallidum	0.022 (0.009-0.065), p=5.96E-5,	0.020 (0.004-0.028), p=1.32E-4,	0.052 (0.014-0.127), p=2.44E-4,	0.022 (0.004-0.114), p=2.44E-4,
Thalamus[a] Pallidum → PFC	0.010 (0.004-0.023), p=5.96E-5,	0.019 (0.005-0.039), p=5.96E-5,	0.025 (0.005-0.045), p=2.44E-4,	0.015 (0.006-0.032), p=2.44E-4,
Thalamus Pallidum → PMA	0.016 (0.008-0.033), p=5.96E-5,	0.017 (0.006-0.026), p=8.86E-5,	0.046 (0.015-0.216), p=2.44E-4,	0.023 (0.006-0.075), p=2.44E-4,
Thalamus Pallidum → SMA	0.022 (0.014-0.045), p=5.96E-5,	0.019 (0.006-0.049), p=5.96E-5,	0.066 (0.034-0.100), p=2.44E-4,	0.021 (0.012-0.118), p=2.44E-4,
Thalamus Pallidum → M1	0.017 (0.009-0.047), p=5.96E-5,	0.021 (0.015-0.045), p=8.86E-5,	0.054 (0.011-0.206), p=2.44E-4,	0.059 (0.016-0.188), p=2.44E-4,
Thalamus Pallidum → S1	0.016 (0.010-0.022), p=5.96E-5,	0.018 (0.007-0.038), p=5.96E-5,	0.031 (0.002-0.064), p=2.44E-4,	0.028 (0.016-0.052), p=2.44E-4,
Pallidum→Thalamus	0.024 (0.010-0.066), p=5.96E-5,	0.019 (0.010-0.042), p=5.96E-5,	0.028 (0.014-0.056), p=2.44E-4,	0.049 (0.016-0.070), p=2.44E-4,
SN→Caudate	0.015 (0.006-0.043), p=1.32E-4,	0.016 (0.005-0.026), p=1.32E-4,	0.069 (0.041-0.112), p=4.88E-4,	0.064 (0.025-0.073), p=2.44E-4,
SN→Putamen	0.011 (0.005-0.023), p=1.32E-4,	0.018 (0.004-0.032), p=2.93E-4,	0.028 (0.020-0.056), p=4.88E-4,	0.013 (0.009-0.068), p=2.44E-4,
SN→Pallidum	0.014 (0.003-0.038), p=1.32E-4,	0.026 (0.008-0.044), p=1.32E-4,	0.014 (0.010-0.058), p=4.88E-4,	0.017 (0.013-0.063), p=2.44E-4,

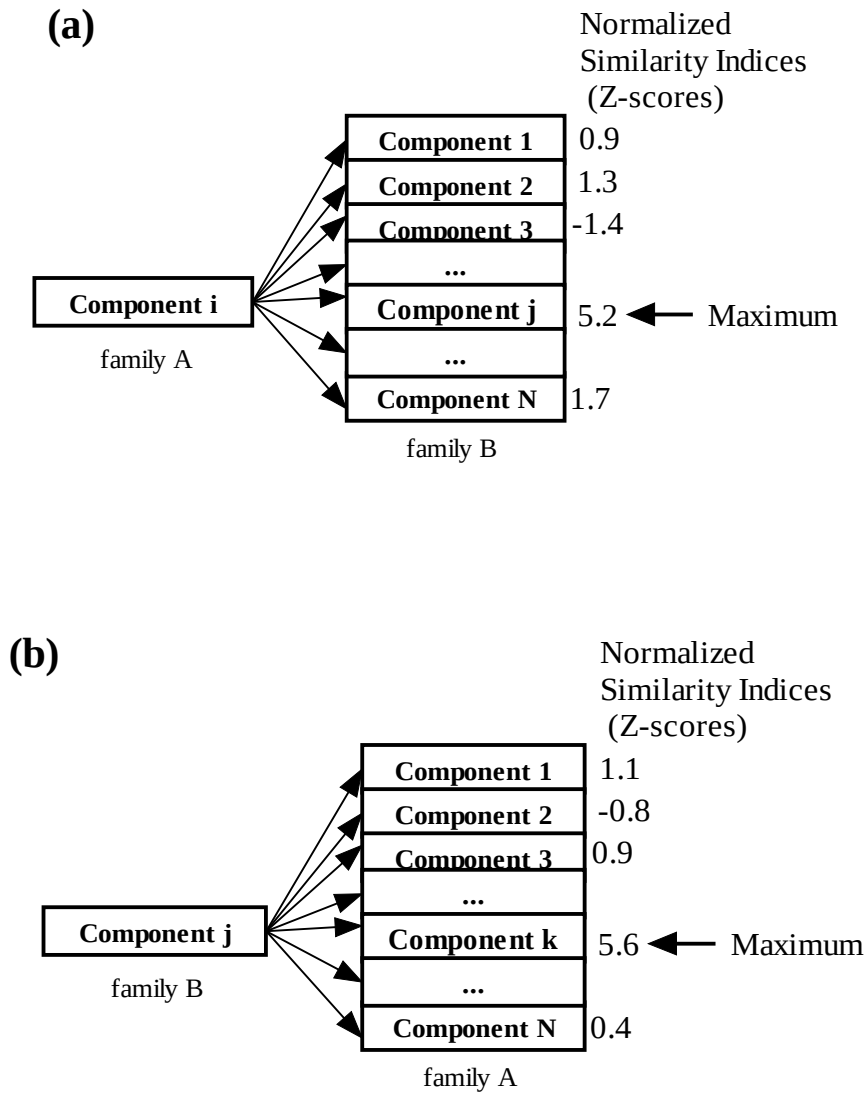
The data in each cell represent the median and interquartile range (the latter in brackets) of the Granger causality index. The p-values (uncorrected) indicate how significantly different the median was from zero (assessed using the Wilcoxon signed rank test). All p-values remain significant after Bonferroni correction. [a]Thalamus X→Y represents the connectivity between X and Y via the thalamus.

Supplemental Table III: Comparisons of Granger Causality Indices of the Interregional Connections Defined in Figure 4

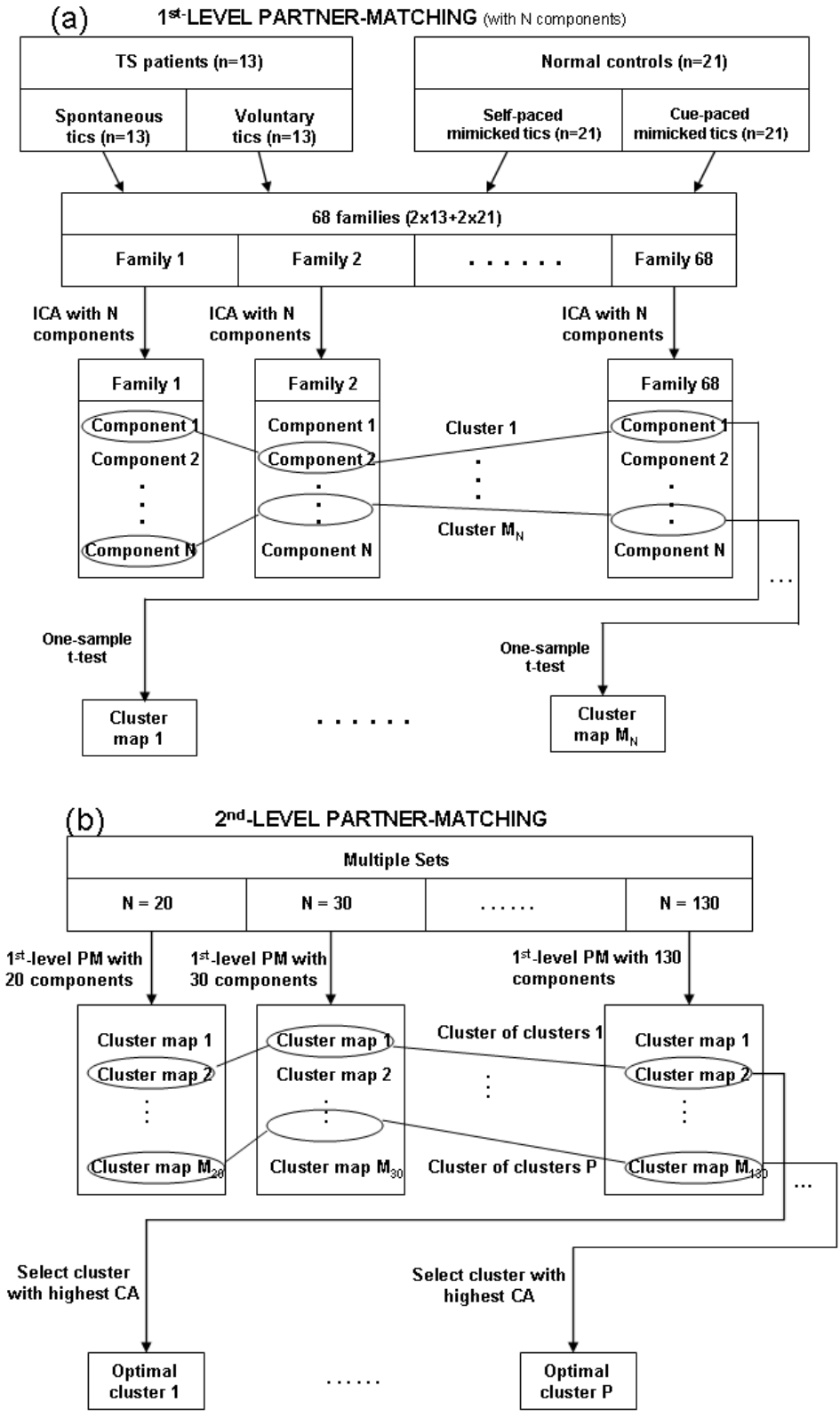
	NC Cue-paced Mimicked Tics vs. NC Self-paced Mimicked Tics	TS Voluntary Tics vs. TS Spontaneous Tics	TS Spontaneous Tics vs. NC Self-paced Mimicked Tics	TS Voluntary Tics vs. NC Cue-paced Mimicked Tics
PFC→PMA	z=-0.326, p=7.44E-1	z=1.333, p=1.82E-1	z=0.284, p=7.77E-1	**z=2.229, p=2.58E-2**
PFC→SMA	z=-1.711, p=8.72E-2	z=0.462, p=6.44E-1	z=0.957, p=3.39E-1	**z=2.764, p=5.71E-3**
PFC→Caudate	z=-1.585, p=1.13E-1	z=0.667, p=5.05E-1	z=-0.532, p=5.95E-1	z=1.630, p=1.03E-1
PMA→SMA	z=-0.143, p=8.86E-1	z=0.974, p=3.30E-1	**z=2.800, p=5.12E-3**	**z=2.192, p=2.84E-2**
PMA→M1	z=0.271, p=7.87E-1	z=1.333, p=1.82E-1	z=1.701, p=8.89E-2	**z=2.609, p=9.08E-3**
SMA→PMA	z=0.874, p=3.82E-1	z=-0.615, p=5.38E-1	**z=2.020, p=4.34E-2**	z=0.534, p=5.93E-1
SMA→M1	z=-0.065, p=9.48E-1	z=0.051, p=9.59E-1	z=1.524, p=1.28E-1	z=1.308, p=1.91E-1
M1→Putamen	z=0.975, p=3.30E-1	z=0.615, p=5.38E-1	z=1.205, p=2.28E-1	z=1.189, p=2.34E-1
S1→M1	z=0.691, p=4.89E-1	z=0.154, p=8.78E-1	**z=2.162, p=3.06E-2**	z=1.824, p=6.82E-2
S1→Putamen	z=0.785, p=4.32E-1	z=0.718, p=4.73E-1	z=1.488, p=1.37E-1	**z=2.302, p=2.13E-2**
Caudate→Pallidum	z=-0.755, p=4.50E-1	z=-0.667, p=5.05E-1	z=0.673, p=5.01E-1	z=0.602, p=5.47E-1
Putamen→Pallidum	z=-1.435, p=1.51E-2	z=-0.769, p=4.42E-1	z=0.886, p=3.76E-1	z=0.998, p=3.18E-1
Thalamus[a] Pallidum → PFC	z=1.233, p=2.18E-1	z=-0.308, p=7.58E-1	z=1.347, p=1.78E-1	z=0.00, p=1.00E+0
Thalamus Pallidum → PMA	z=-0.300, p=7.64E-1	z=-1.333, p=1.82E-1	z=1.914, p=5.57E-2	z=0.792, p=4.28E-1
Thalamus Pallidum → SMA	z=-0.604, p=5.46E-1	z=-0.718, p=4.73E-1	**z=2.197, p=2.80E-2**	z=0.815, p=4.15E-1
Thalamus Pallidum → M1	z=0.665, p=5.06E-1	z=0.103, p=9.18E-1	z=1.382, p=1.67E-1	z=1.087, p=2.77E-1
Thalamus Pallidum → S1	z=0.428, p=6.69E-1	z=0.359, p=7.20E-1	z=0.602, p=5.47E-1	z=1.205, p=2.28E-1
Pallidum→Thalamus	z=-0.604, p=5.46E-1	z=0.923, p=3.56E-1	z=0.106, p=9.15E-1	**z=2.091, p=3.65E-2**
SN→Caudate	z=-0.350, p=7.26E-1	z=0.789, p=4.30E-1	**z=2.535, p=1.13E-2**	**z=2.954, p=3.13E-3**
SN→Putamen	z=0.824, p=4.10E-1	z=0.734, p=4.63E-1	**z=2.494, p=1.26E-2**	z=0.628, p=5.30E-1
SN→Pallidum	z=1.022, p=3.07E-1	z=-0.571, p=5.68E-1	z=1.075, p=2.83E-1	z=0.729, p=4.66E-1

Note: **bold** indicates that the comparisons are significant ($p<0.05$, uncorrected, two-sided Wilcoxon rank sum test). [a]Thalamus X→Y represents the connectivity between X and Y via the thalamus.

SUPPLEMENTAL FIGURES



Supplemental Figure 1: Bidirectional matching ("Partner Matching") between families A and B. (a) For each component *i* in family A, we calculate the normalized similarity index with each component in family B, identifying component *j* in family B as having the maximum similarity index with component *i* in family A. We considere component *j* to be a candidate matching for component *i*. (b) Because component *j* is the component in B that has the maximum similarity to *i* does not necessarily mean that the reverse is also true – i.e., that component *i* is the component in family A that has maximum similarity with component *j* in family B. We therefore repeat the matching procedure, but in the reverse direction. Let component *k* be the component in family A that has the maximum similarity with component *j* in family B. If $k = i$, then the match is deemed to be bidirectional, and we consider that component *i* in family A and component *j* in family B match. If $k \neq i$, then we deem that no match is present.



Supplemental Figure 2: Hierarchical Partner Matching Procedures. (a) We applied 1st-level partner matching to the fMRI data collected from TS and control participants, who each performed two run types

(spontaneous and voluntary tics in the TS group; self-paced and cue-paced mimicked tics in the control group), yielding a total of 13*2+21*2 = 68 different families. We used ICA to generate N components for each family and then used partner matching to match these components across the families, yielding clusters of components that matched across families. For each of these clusters, we ran one-sample t-tests to generate cluster maps. (b) The main limitation of simple (1st-level) partner matching is that it assumes a fixed number N of independent components, but that number is not known *a priori*, and it may even vary across families (i.e., across participants and run types). We therefore used a hierarchical approach that includes a second level of partner matching. In this second level, we used various numbers of components (from 20 to 130 in increments of 10). For each of these sets of component numbers, we performed 1st-level partner matching to generate the cluster maps. We then applied partner matching again across these cluster maps (2^{nd} level), to determine which clusters mapped across sets containing differing numbers of components. This hierarchical procedure yielded clusters of clusters. For each cluster of clusters, we then selected for use in all further analyses the cluster that had the highest Cronbach alpha, or measure of similarity and reproducibility in the spatial configuration of the independent components within that cluster.