Subject- and behavior-specific signatures extracted from fMRI data using whole-brain effective connectivity

Pallarés V*[1], Insabato A*[1,2], Sanjuán A[1], Kühn S[3,4], Mantini D[5,6,7], Deco G**[1,8], Gilson M** [1]

* Equal contribution
** Equal contribution

1. Center for Brain and Cognition, Computational Neuroscience Group, Department of Information and Communication Technologies, Universitat Pompeu Fabra, Carrer de Ramon Trias Fargas, 25-27, Barcelona, 08005, Spain
2. The Italian Academy, Center for Theoretical Neuroscience, Columbia University, 1161 Amsterdam Ave., New York NY 10027, USA
3. Max Planck Institute for Human Development, Center for Lifespan Psychology, Lentzeallee 94, 14195 Berlin, Germany
4. University Clinic Hamburg-Eppendorf, Clinic and Policlinic for Psychiatry and Psychotherapy, Martinistraße 52, 20246 Hamburg, Germany
5. Research Center for Motor Control and Neuroplasticity, KU Leuven, 101 Tervuursevest, 3001 Leuven, Belgium
6. Department of Health Sciences and Technology, ETH Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland
7. Department of Experimental Psychology, Oxford University, 15 Parks Road, Oxford OX1 3PH, United Kingdom
8. Institució Catalana de la Recerca i Estudis Avanats (ICREA), Universitat Pompeu Fabra, Passeig Lluís Companys 23, Barcelona, 08010, Spain

Abstract: 213 words
Main text: ~3900 words
References: 51

ABSTRACT:
The idea of fMRI-based personalized medicine is emerging to characterize brain disorders at the patient level and to develop tailored therapeutic protocols. A main limitation in this direction, as well as with fundamental studies of brain function, is the variability observed in fMRI data. In practice, such measurements exhibit a mixture of session-to-session variability with subject- and condition-related information. Disentangling these contributions is crucial for real-life applications where only a few sessions per subject are available. The present study aims to define a new reliable standard for the extraction of signatures from fMRI data. These signatures correspond to subnetworks of directed interactions between brain regions (typically 100 covering the whole brain) supporting the subject and condition identification for single fMRI sessions. The key is to step from functional connectivity to effective connectivity, which is estimated using a whole-brain dynamic model. Our method shows that as few as 4 sessions per subject are sufficient to perfectly identify more than 40 other sessions of 6 subjects. We also demonstrate the good generalization capability for 30 subjects. Using another dataset with resting state and movie viewing, we show that the two extracted signatures correspond to distinct subnetworks, suggesting some sort of orthogonality. Our results set solid foundations to follow longitudinally a subject's condition from fMRI data.

INTRODUCTION

Blood-oxygen-level dependent (BOLD) signals in functional magnetic resonance imaging (fMRI) have been used for more than two decades to observe human brain activity and relate it to functions [1,7]. Even at rest, the brain exhibits patterns of correlated activity between distant areas [4,40]. The functional connectivity (FC) measures the statistical dependencies between the BOLD activities of brain regions, which has then been studied for subjects performing tasks and compared with the resting state. Recently, interest has grown about the temporal BOLD structure for both individual regions [27] and between areas (via the cross-covariance lags at the scale of seconds) [36]; the 'dynamic FC' was defined to quantify the BOLD correlations at the scale of minutes [23,39]. Following fundamental discoveries about brain functions, fMRI has been increasingly used to complement clinical diagnostic for neuropathologies [34]. Resting-state fMRI has also been found to be informative about neuropsychiatric disorders [25]: alterations in FC correlate with and can predict the clinical scores of several diseases [30,41].

Recent studies have focused on the reliability of these FC measures recorded from the same subject over successive sessions [44,37,6,38]. Consistent differences between subjects (with individual stability) allow subject identification using recorded FC as a "fingerprint" [15]. Moreover, this subject specificity may even be enhanced in task-evoked activity [16]. Another model-based approach used linear-regression coefficients of BOLD signals instead of FC to identify the subjects [35]. A recent prospective study about the evolution of psychiatric disorders emphasized individual specificities in the FC stabilization during childhood (irrespective of the disease) [32], whereas traditional group-averaging aims to remove the individual differences to obtain task-specific [49] or pathology-specific [11] signatures. The mixture of session-to-session, subject-specific and condition-related variability in FC is a crucial issue for real-life applications where only a few sessions per subject can be recorded, such as clinical diagnostic. Because previous studies [15,35,16] were limited to datasets with at most 3 resting-state sessions per subject, we aim to rigorously assess the generalization capability of prediction methods to future (unseen) data.

Distributed signatures in FC across the whole brain have been observed in memory tasks [42] or when the subject experiences psychological pain [5]. Moreover, the etiology of many mental disorders is unknown: they are suspected to arise from network dysfunction, as reported for large-scale FC alterations in patients with schizophrenia [31]. These examples strongly point in favor of whole-brain approaches to study high-level cognition [9] and brain diseases [10]; in contrast, focusing on a few cortical areas only to test hypotheses [22,2] may not capture sufficient information and network effects. Such whole-brain approaches typically involve a large number of parameters to estimate, which may impair the robustness. One aim of the present study is to provide a practical answer to this trade-off dilemma.

The idea underlying the study of FC – in the broad sense – lies in that it reflects how brain areas dynamically bind to exchange and process information [18,28,3]. To move beyond a phenomenological description of FC, our method relies on a model inversion [20] to interpret FC by decomposing it into changes in network connectivity and local fluctuating activity. Note that we borrow the terminology of effective connectivity (EC) to describe the interactions between brain regions from dynamic causal model [19], although our model implies simplifications in comparison. As with FC, a crucial issue for EC is whether the estimated model parameters are reliable across several sessions for the same subject [17], which determines whether they can predict the subjects' identities in practice [35].

The present study aims to set a new standard to extract multivariate signatures from fMRI data, discriminative against subjects or behavioral conditions. The paper is organized in two parts. First, we couple whole-brain EC estimation with adequate machine learning tools to identify subjects from resting state fMRI (i.e., classify single sessions to the corresponding participant) capitalizing on previous studies [15,35]. The focus is on the comparison between EC and FC in their generalization capabilities. To do so, we rely on datasets with large numbers of sessions and healthy subjects to control the effect of session-to-session variability. Second, we predict both the subject's identity and condition (rest versus movie viewing) to verify that EC can disentangle the two types of signatures. Doing so, we examine the topological distribution of the EC links supporting the twofold classification. Although we test the method on behavioral conditions, we benchmark it keeping in mind future applications to clinical diagnostic.

RESULTS

Functional and effective connectivity as measures of the brain network dynamics

In this study we used fMRI data from three datasets described in Table 1. Classical functional connectivity (corrFC) was calculated using the pairwise Pearson correlation coefficient (PCC) between the time courses of the regions of interest (ROIs), obtaining an N×N symmetric matrix for each recorded session (N=116 for Datasets A and B, N=66 for Dataset C); see Eq. (2) in Methods. In parallel, we used the whole-brain dynamic model [20] in Figure 1B: each ROI is a node in a noise-diffusion network whose topology (skeleton) is determined by the structural connectivity (SC) obtained from diffusion tensor imaging (DTI) or similar techniques. In the model, the global pattern of FC arises from the local variability Σ_i that propagates via the network connections EC_{ij} (from j to i). To fit each fMRI session, all relevant EC_{ij} and Σ_i parameters are iteratively tuned such that the model spatio-temporal FC – as measured by FC0 (0-lag covariances) and FC1 (1-lag shifted covariances) – best reproduces the empirical counterpart. A detailed description of the model and the maximum-likelihood estimation procedure is provided in Methods. In essence, the model inversion decomposes the empirical matrices (FC0,FC1) into two estimates EC and Σ, which can be seen as multivariate biomarkers for the brain dynamics in each fMRI session.

Dataset name	Acquisition / Previous analysis	Number of subjects in our analysis	Sessions per subject	Session duration
Dataset A1	Day2day project [14]	6	40-50	5 minutes
Dataset A2	Day2day project [14]	50	1	5 minutes
Dataset B	CoRR [51]	30	10	10 minutes
Dataset C	[33,21]	19	3 for resting state and 2 for movie viewing	10 minutes

Table 1: Details of the datasets used in the present study. Dataset A (A1+A2) was used to test the robustness of subject identification to session-to-session variability. Dataset B was used to test the generalization capability of the identification procedure for a larger number of subjects. Dataset C was used to extract both individualized and behavioral signatures. After applying a standard preprocessing pipeline to the BOLD signals (see Methods for details), we parcellated the brain into 116 regions of interest (ROIs) by using an anatomical space [47] for Datasets A1, A2 and B (see Figure 1A). Dataset C was parcellated into 66 ROIs covering the cortex [26].

FIGURE 1

Similarity of connectivity measures across sessions and subjects

Using Datasets A and B, we compared the capability of the 4 connectivity measures (corrFC, FC0, FC1 and EC), as well as Σ, in terms of within- and between-subject similarity (WSS and BSS, respectively), as a first step toward subject identification. For each pair of sessions, the similarity S_X was calculated using the PCC between two z-scored vectorized connectivity measures X in Figure 1C (non-zero elements for EC, low-triangle elements for corrFC). In the matrix of S_{EC} values for Dataset A1 (Figure 2A), 6 diagonal blocks with larger values corresponding to the WSS can be noticed; the remaining matrix elements correspond to BSS. Figure 2B compares the distributions of S_{EC} and S_{corrFC}: WSS and BSS distributions are better separated for EC than for corrFC. In other words, sessions from the same subject are more similar to one another, and more different from those of other subjects, viewed from the EC than the corrFC viewpoint. This suggests a better capability of EC to discriminate between subjects. Note that the BSS from Datasets A1 (6 subjects) and A2 (50 subjects) remarkably overlap for both corrFC and EC, showing that BSS for 6 subjects generalizes well to larger numbers.

These qualitative observations are confirmed by Table 2 that summarizes the Kolmogorov-Smirnov (KS) distance between the similarity distributions (blue versus red and blue versus green in Figure 2B). EC gave

larger KS distance than corrFC and all other FC-related measures. Note that we also calculated KS distance using only the links in corrFC and FC0 corresponding to the 4056 existing connections in EC (determined by SC), in order to compensate for the (relative) sparsity of EC links as compared to corrFC and FC0; this did not change the results. Last, the diagonal elements of Σ showed the smallest distances. In the following, we focus on EC and corrFC. Supplementary Figures S1 and S2 show the similarity distributions for FC0, FC1, corrFC/SC and Σ using Datasets A1 and B.

Measure	Description	Dataset A1	Dataset A1 & A2
EC	EC weights estimated with the model	0.6440	0.6581
corrFC	FC computed as Pearson correlation	0.4517	0.5477
FC0	FC computed as 0-lag covariance matrix	0.3685	0.4888
FC1	FC computed as 1-lag covariance matrix	0.4139	0.5118
corrFC/SC	FC computed as Pearson correlation and masked with SC	0.4769	0.5729
FC0/SC	FC computed as 0-lag covariance matrix and masked with SC	0.3770	0.4644
Σ	Local variability for each ROI estimated with the model	0.3778	0.3287

Table 2: Kolmogorov-Smirnov (KS) distance between WSS and BSS distributions. Comparison of the connectivity measures (corrFC, FC0, FC1) and model estimates (EC, Σ). The third column corresponds to the distance between the blue and red distributions in Figure 2B, and the fourth column to the blue and green distributions.

FIGURE 2

Structure of individual session-to-session variability for EC and corrFC

The high dimensionality of our connectivity measures may reduce their predictive power, known as Hughes phenomenon [29]. This is especially important in our case where the number p of dimensions for the multivariate measures (Dataset A1: p=4056 for EC, p=6670 for corrFC) exceeds the number n of samples (n=6 subjects × 50 sessions=300 samples). To further characterize individual variability over sessions, we performed a reduction of dimensionality using principal component analysis (PCA) on the sessions of Dataset A1. By the naked eye, the colored clouds representing all sessions for each subject exhibit smaller overlap for EC than corrFC in Figure 2C, where data are projected onto the first 6 principal components (PCs).

We quantified the clustering degree of these clouds using a silhouette coefficient for each session, ranging from -1 (poor clustering) to 1 (perfect clustering); see Eq. (15) in Methods. As shown in Figure 2D, EC produced larger (almost all positive) silhouette values than corrFC, confirming the visual impression of Figure 2C. Silhouette coefficients were calculated on the data projected onto the first 6 PCs, i.e. the number of PCs that maximized the silhouette coefficient (see Figure S3). As can be seen in Figure 2E, the silhouette coefficients for the data in the original link space (left violin plots) are smaller than those for data in the PCs space (right); for Dataset B (in blue), the first 30 PCs are used. Consequently, PCA may facilitate the identification of subjects by reducing the dimensionality of the data.

Subject identification using EC is more robust than using corrFC

Now we turn to the main goal of our study: the classification of single sessions – attributing them to subjects – based on EC or corrFC. Robust subject identification for ~100 subjects was pioneered by a recent publication [15], relying on a k-nearest-neighbor (kNN) classifier with k=1 and PCC as metric. In order to classify a *target* session, the PCC between the target and 1 known reference session for each subject *(*called *database*) is calculated; the predicted identity for the target is that of the subject corresponding to the closest (most similar) session (see Figure S4 and Methods for details). In contrast with previous studies using 1NN [15,16,32], our method relies on multinomial logistic regression (MLR) classifier, a classical tool in machine learning. MLR uses a linear model to predict the probability that an input sample belongs to a class (subject here). A technical comparison of both approaches is further detailed in Methods.

In classification algorithms the problem of overfitting describes the situation where the algorithm performs very well with the data it is trained with, but fails to generalize to new samples. Due to the high dimensionality of the connectivity measures [29], it is essential to control for overfitting with an appropriate training and test procedure. Our train-test procedure and the use of large test-retest datasets – unlike previous studies [15,16,48] – aims to provide a trustworthy characterization of the quality of the classifiers. Figure 3A describes the train-test procedure for the identification of subjects: 1) fMRI sessions (EC in the figure) are randomly split in training and test datasets; 2) after preprocessing (orange arrows) involving within-session z-score (see Eq. (16) in Methods) followed – or not – by PCA, the classifier is optimized as illustrated for the MLR with boundaries that best predict the training dataset; 3) test set is used to verify the generalization capability of the classifier (blue arrows), by measuring to which extent the classifier boundaries, estimated with the train set, correctly classify single sessions from the test set.

We first used Dataset A1: we increased the number of training sessions per subject from 1 to 40 to evaluate how many training sessions are necessary for satisfactory accuracy. As shown in Figure 3B, EC (in red) outperformed corrFC (in blue) by more than one standard deviation (shaded area around the curve), for both MLR and 1NN. Moreover, almost perfect classification was reached with MLR for 5 training sessions only, whereas 10-15 were necessary for 1NN. This is important when only a few training sessions per subject are available, as expected with clinical applications. Figure 3C displays the classification accuracy on Dataset B, used to verify the robustness with respect to the number of subjects to be classified. We trained the classifiers with 1 session per subject and evaluated the performance varying the number of subjects from 2 to 30 (test set comprised the remaining 9 sessions per subject). Again, EC is more robust than corrFC: while performance with corrFC rapidly deteriorates as the number of subjects is increased, classification using EC is barely affected by the number of subjects. This is our core technical result: EC and MLR outperform by large corrFC and 1NN, respectively. Other connectivity measures such as FC0 showed similar performance to corrFC (not shown).

Unlike with the analysis of the data structure (Figure 2E), PCA only marginally increased the classification performance here (Figure S5). The classification accuracy is already very high with ~95% for MLR and EC with 1 training session, which partly explains this limitation. As further discussed in Supplementary Material (Section S4), the distribution of subject-specific information across the PCs is broad. This supports the use of proper machine learning tools to extract it in order to discriminate the subjects.

FIGURE 3

Signature network of links supporting the classification

Another important advantage of the MLR over kNN is its efficiency in characterizing the links that contribute to the classification. We used recursive feature elimination (RFE, see Methods for details) to rank the links according to their weight in the classification and chose the lowest number of links that achieved the maximum classification performance. In comparison, the same procedure with kNN would require many more computations, recalculating the closest neighbor for all combinations of links (here the number of links is $p>>1000$; see Methods for further discussion). The resulting support network for dataset A1 had 18 links, to be compared with 44 links for dataset B. In both cases, subject identification using only those links achieved perfect accuracy (90% of all available sessions were used for training and 10% for testing, see Figure S9). The two support networks are shown in Figure 3D in the same matrix: remarkably, the networks are very sparse and non-uniformly distributed across the whole brain. This is the signature of the most subject-discriminative ROIs: frontal and cingulate cortices, as well as the temporal and occipital regions, seem to play a major role here. It is worth noting that the adjacency matrix is not symmetric, which implies different roles for nodes as receivers (especially frontal ROIs) or senders (cingulate).

The sparsity of the signature in Figure 3D hides the fact that the rankings for Datasets A1 and B are close (PCC=0.59, p-value$<<10^{-50}$), indicating that similar neural networks characterize individuals in two disjoint sets of subjects; see also Figure S10 that illustrates the correspondence at the level of anatomical groups (in color in Figure 3D). To further measure the overlap between these networks, we selected for each dataset the subset of links with the highest ranking and computed the number of common links. Figure 3E shows that the proportion of common links exceeds by far its expectation under the hypothesis of random rankings (shaded gray area). This indicates a good agreement between the support networks from the two datasets even at the single-link level.

Twofold classification of subject identity and behavioral condition

Finally, we used Dataset C to extract a signature for the subject identity and another for the behavioral condition. This is schematically depicted in Figure 4A, with three fictive dimensions: the information about subject identity corresponds to the x-axis and information about the condition to the z-axis; the session-to-session variability, that should be ignored, spreads along the y-axis. In this idealized scenario, it is possible to classify a session with respect to both subjects and conditions using different dimensions of the data. In the high dimensional case, this occurs when different sets of links support the two classifications.
Using MLR and EC, we achieved very high performance (accuracy > 90%) for subject identification and perfect classification for the condition.

We then sought the smallest subsets of links that achieve the maximum performance of each classification using RFE (Figure 4B), as done before (see Methods for details). Both support networks were again very sparse and distributed across the brain, as can be seen in their adjacency matrix (Figure 4C). More links are necessary to identify the subjects (57) than the behavioral conditions (13), indicating a higher complexity for the former.

As before, we used RFE to rank the links according to their contribution to the classification. We computed the number of common links for the subject and condition identifications, which fell within the expected values with the null hypothesis (Figure 4D). This indicated that distinct subsets of links are relevant for the subjects' identity and behavioral condition.

FIGURE 4

Similar to Datasets A1 and B, subject identification of Dataset C largely concerns the frontal and cingulate systems. Condition identification is also supported by occipital and temporal cortices, which are expected to have the strongest activity modulations during movie viewing. The top panels in Figures 5A and B represent the two support networks such that the directed nature of links can be appreciated. Apart from two small components, the subject network is almost fully connected with several central nodes (hubs, indicated by their large size) located in frontal and cingulate regions. On the contrary, the network for condition is segregated with small isolated components. The bottom plots in Figure 5 show the lateralization of the support links, stressing the asymmetries between the two hemispheres: most of the important links are ipsilateral (i.e. within the same hemisphere) and many belong to the left hemisphere for the subject network, whereas they are mainly contralateral for the condition network.

FIGURE 5

DISCUSSION:

In this study, we have proposed a framework to predict the identity of subjects as well as their condition from fMRI time series, by robustly extracting discriminative signatures about subject/condition differences. We obtain very sparse signatures, supposedly because of the datasets used (30 subjects maximum; 2 conditions). Their size is expected to increase with the complexity of the "environment" to represent (many subjects, many tasks); resources are becoming available to test this quantitatively [50 ,24]. Importantly, we have proven that such EC-based signatures are robust to the session-to-session variability, and can be obtained relying on a limited number of sessions (4-5 recordings of 5 minutes each). Proper machine-learning tools such as MLR are necessary to efficiently extract those signatures. We now discuss specific points.

The fundamental advancement of our study is the development of a reliable and well-benchmarked method, extending the previously published proofs of concept [35,15]. EC discriminates subjects better than corrFC (Figure 2), as well as with both classifiers MLR and kNN (Figure 3). In particular, the generalization capability for EC is much more robust than FC when the classification becomes harder (few sessions per subject or many subjects to identify, Figure 3B-C). This confirms that the BOLD temporal structure – captured by the EC after the bandpass filtering of the BOLD signals – reflects the identity of the subject [35], as previously shown for a task involving (or not) attention [27] or for wake versus sleep [36]. The use of z-scores in the classification shows that the EC ranking (i.e., which brain connections have large weights among all) conveys individual information. The quasi-perfect identification shows that these EC patterns are very stable, as was found previously with FC across tasks [15,16]. Here the focus was on EC because it performed better than Σ estimates for the resting-state fMRI, but it has been recently shown that Σ is strongly affected when engaging a task condition [21], so Σ might further improve the classification for conditions, in particular involving sensory stimuli.

Our whole-brain dynamic model is a continuous-time network with linear feedback that incorporates topological constraints from SC. EC corresponds to a maximum-likelihood estimate and can be very efficiently calculated for the whole brain with ~100 ROIs and each session with ~300 time points per ROI [20,21]. Our results show that, although the dynamic model and estimation procedure are a simplification compared to the dynamic causal model with hemodynamics and Bayesian machinery [45], it nonetheless provides powerful signatures that can be used for discrimination between subjects and conditions. Our study has focused on two coarse parcellations covering the whole brain [47] or cortex [26]. Although the two parcellations where applied to different datasets we did not observe significant difference in the performance of the classifiers. Much work has been done recently to correct the bias due to the use of specific parcellations [8]; for our purpose, more refined parcellations may entail better discriminability in higher-dimensional spaces, but raise issues for the EC estimation robustness. On another ground, preprocessing using PCA was not found to significantly enhance the performance here. Nonetheless, PCA may be useful for datasets with larger number of subjects and conditions [39]. Eventually, the generalization capability is *the* criterion for the classification performance and further work is needed to define a suitable scale for applications with more subjects and conditions.

We have found that very few links (<4%) were sufficient to classify perfectly 30 subjects from Dataset B (Figure 3C) and both subjects and conditions in Dataset C (Figure 4D). For a larger cohort [15] and more tasks [24], we expect this number to grow and the infra/supra-linear dependency with the subject number should be addressed carefully. Those support networks for the twofold classification (subject and condition) show several noticeable differences (Figure 5). The subject network is large, almost fully connected, distributed over the two hemispheres (with more links within the left one) and concentrated in the cingulate and frontal areas. This suggests subject-specific dynamics within areas involved in high-level functions and overlapping with the default mode network [40]. This interpretation of EC in terms of brain communication comes from the directed nature of EC, which considers the propagation of BOLD activity. It follows that the discriminative EC patterns may reflect heterogeneities in the interactions between the different neural subsystems (e.g., frontal to cingulate in Figure 5) and the propagation of information between them [12,13]. We also found a much higher percentage of contralateral links for condition than subject. This is in line with strong inter-hemispheric interactions observed for the same dataset with community analysis [21]. As expected with the movie viewing condition studied here, links in the visual and temporal areas are discriminative.

The idea of personalized medicine using neuroimaging to characterize brain disorders at the patient level is emerging [34,50]. The development of tailored therapeutic protocols [43] – to optimize recovery and minimize adverse effects – requires quantitative tools that deliver a precise diagnostic of the patient's evolution. Our proposed scheme is to follow a patient's trace over time in the (high-dimensional) EC space: extending the diagram in Figure 4A, the classification should be extended to a three-fold space (subject, task and pathology), the latter dimension corresponding to healthy versus pathological states. One (or several) pathology-specific signature(s) will be extracted from resting-state [25] or task-evoked fMRI; specific tasks may indeed reveal powerful signatures for certain pathologies, e.g., memory exercises for Alzheimer [30]. We expect these signatures to be much more complex [5] than that for movie viewing (Figure 5).The generalization capability of prediction methods to future (unseen) data [29] is crucial in this clinical context. To this end, our method disentangles the diverse signatures, while properly conditioning out the day-to-day fMRI variability (as uninformative intrinsic noise). This provides a practical solution to the recent criticism that “a major reason for disappointing progress of psychiatric diagnostics and nosology is the lack of tests which enable mechanistic inference on disease processes within individual patients” [46]. Note that the goal is more to prevent individual signatures from mixing with those for pathologies than to discriminate between subjects.

REFERENCES

1. Amunts K, Hawrylycz MJ, Van Essen DC, Van Horn JD, Harel N, Poline JB, De Martino F, Bjaalie JG, Dehaene-Lambertz G, Dehaene S, Valdes-Sosa P, Thirion B, Zilles K, Hill SL, Abrams MB, Tass PA, Vanduffel W, Evans AC, Eickhoff SB (2014) Interoperable atlases of the human brain. Neuroimage, 1: 525-532. doi: 10.1016/j.neuroimage.2014.06.010

2. Bastos-Leite AJ, Ridgway GR, Silveira C, Norton A, Reis S, Friston KJ (2015) Dysconnectivity within the default mode in first-episode schizophrenia: a stochastic dynamic causal modeling study with functional magnetic resonance imaging. Schizophr Bull, 41: 144-153; doi: 10.1093/schbul/sbu080

3. Betti V, Della Penna S, de Pasquale F, Mantini D, Marzetti L, Romani GL, Corbetta M (2013) Natural scenes viewing alters the dynamics of functional connectivity in the human brain. Neuron, 79: 782-797

4. Biswal B, Yetkin FZ, Haughton VM, Hyde JS (1995) Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. Magn Reson Med, 34:537-541; doi: 10.1002/mrm.1910340409

5. Chang LJ, Gianaros PJ, Manuck SB, Krishnan A, Wager TD (2015) A Sensitive and Specific Neural Signature for Picture-Induced Negative Affect. PLoS Biol, 13: e1002180; doi:10.1371/journal.pbio.1002180

6. Chen B, Xu T, Zhou C, Wang L, Yang N, Wang Z, Dong HM, Yang Z, Zang YF, Zuo XN, Weng XC (2015) Individual Variability and Test-Retest Reliability Revealed by Ten Repeated Resting-State Brain Scans over One Month. PLoS One, 10: e0144963. doi: 10.1371/journal.pone.0144963

7. Cordes D, Haughton VM, Arfanakis K, Wendt GJ, Turski PA, Moritz CH, Quigley MA, Meyerand ME (2000) Mapping functionally related regions of brain with functional connectivity MR imaging. AJNR Am J Neuroradiol 21: 1636-1644

8. Da Mota B, Fritsch V, Varoquaux G, Banaschewski T, Barker GJ, Bokde ALW, Bromberg U, Conrod P, Gallinat J, Garavan H, Martinot J-L, Nees F, Paus T, Pausova Z, Rietschel M, Smolka MN, Ströhle A, Frouin V, Poline J-B, Thirion B (2014) Randomized parcellation based inference. NeuroImage 89: 203-215; http://dx.doi.org/10.1016/j.neuroimage.2013.11.012

9. Deco G, Jirsa V, McIntosh A (2011) Emerging concepts for the dynamical organization of resting-state activity in the brain. Nat Rev Neurosci, 12: 43-56; doi:10.1038/nrn2961

10. Deco G, Kringelbach ML (2014) Great expectations: using whole-brain computational connectomics for understanding neuropsychiatric disorders. Neuron, 84: 892-905; doi: 10.1016/j.neuron.2014.08.034

11. Drysdale AT, Grosenick L, Downar J, Dunlop K, Mansouri F, Meng Y, Fetcho RN, Zebley B, Oathes DJ, Etkin A, Schatzberg AF, Sudheimer K, Keller J, Mayberg HS, Gunning FM, Alexopoulos GS, Fox MD, Pascual-Leone A, Voss HU, Casey BJ, Dubin MJ, Liston C (2017) Resting-state connectivity biomarkers define neurophysiological subtypes of depression. Nat Med, 23: 28-38. doi: 10.1038/nm.4246

12. Ekstrom A (2010) How and when the fMRI BOLD signal relates to underlying neural activity: the danger in dissociation. Brain Res Rev, 62: 233-244. doi: 10.1016/j.brainresrev.2009.12.004

13. Engel AK, Gerloff C, Hilgetag CC, Nolte G (2013) Intrinsic coupling modes: multiscale interactions in ongoing brain activity. Neuron, 80: 867-886. doi: 10.1016/j.neuron.2013.09.038.

14. Filevich E, Lisofsky N, Becker M, Butler O, Lochstet M, Martensson J, Wenger E, Lindenberger U, Kühn S (2017) Day2day: investigating daily variability of magnetic resonance imaging measures over half a year. BMC Neurosci, 18: 65. doi: 10.1186/s12868-017-0383-y

15. Finn ES, Shen X, Scheinost D, Rosenberg MD, Huang J, Chun MM, Papademetris X, Constable RT (2015) Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. Nat Neurosc, 18: 1664-1671; doi: 10.1038/nn.4135

16. Finn ES, Scheinost D, Finn DM, Shen X, Papademetris X, Constable RT (2017) Can brain state be manipulated to emphasize individual differences in functional connectivity? Neuroimage, S1053-8119. doi: 10.1016/j.neuroimage.2017.03.064

17. Frässle S, Stephan KE, Friston KJ, Steup M, Krach S, Paulus FM, Jansen A (2015) Test-retest reliability of dynamic causal modeling for fMRI. Neuroimage, 15: 56-66. doi: 10.1016/j.neuroimage.2015.05.040

18. Fries P (2005) A mechanism for cognitive dynamics: Neuronal communication through neuronal coherence. Trends Cogn Sci, 9: 474-480

19. Friston KJ (2011) Functional and effective connectivity: A review. Brain Connect, 1: 8; doi: 10.1089/brain.2011.0008

20. Gilson M, Moreno-Bote R, Ponce-Alvarez A, Ritter P, Deco G (2016) Estimation of Directed Effective Connectivity from fMRI Functional Connectivity Hints at Asymmetries of Cortical Connectome. PLoS Comput Biol, 12: e1004762; doi: 10.1371/journal.pcbi.1004762

21. Gilson M, Deco G, Friston K, Hagmann P, Mantini D, Betti V, et al. (2017): Effective connectivity inferred from fMRI transition dynamics during movie viewing points to a balanced reconfiguration of cortical interactions. BioRxiv.110015

22. Goebel R, Roebroeck A, Kim D, Formisano E (2003) Investigating directed cortical interactions in time-resolved fMRI data using vector autoregressive modeling and Granger causality mapping. Magn Reson Imaging, 21: 1251-1261; doi: 10.1016/j.mri.2003.08.026

23. Gonzalez-Castillo J, Bandettini PA (2017) Task-based dynamic functional connectivity: Recent findings and open questions. Neuroimage. 2017 Aug 3. pii: S1053-8119(17)30653-5. doi: 10.1016/j.neuroimage.2017.08.006

24. Gordon EM, Laumann TO, Gilmore AW, Newbold DJ, Greene DJ, Berg JJ, Ortega M, Hoyt-Drazen C, Gratton C, Sun H, Hampton JM, Coalson RS, Nguyen AL, McDermott KB, Shimony JS, Snyder AZ, Schlaggar BL, Petersen SE, Nelson SM, Dosenbach NUF (2017) Precision Functional Mapping of Individual Human Brains. Neuron, xxx; doi: 10.1016/j.neuron.2017.07.011

25. Greicius MD (2008) Resting-state functional connectivity in neuropsychiatric disorders. Curr Opin Neurology, 21: 424

26. Hagmann P, Cammoun L, Gigandet X, Meuli R, Honey CJ, Wedeen VJ, Sporns O (2008) Mapping the structural core of human cerebral cortex. PLoS Biol, 6: e159. doi: 10.1371/journal.pbio.0060159

27. He BJ (2011) Scale-free properties of the functional magnetic resonance imaging signal during rest and task. J

Neurosci 31 (2011) 13786–13795Honey CJ, Thivierge JP, Sporns O (2010) Can structure predict function in the human brain? Neuroimage 52: 766-76. doi: 10.1016/j.neuroimage.2010.01.071

28. Hipp JF, Engel AK, Siegel M (2011) Oscillatory synchronization in large-scale cortical networks predicts perception. Neuron, 69: 387-396

29. Hughes GF (1968) On the mean accuracy of statistical pattern recognizers. IEEE Trans Inf Theory, 14: 55–63. doi: 10.1109/TIT.1968.1054102

30. Kurth S, Moyse E, Bahri MA, Salmon E, Bastin C (2015) Recognition of personally familiar faces and functional connectivity in Alzheimer's disease. Cortex, 67: 59-73. doi: 10.1016/j.cortex.2015.03.013

31. Hoptman MJ, Zuo X-N, D'Angelo D, Mauro CJ, Butler PD, Milham MP, Javitt DC (2012) Decreased interhemispheric coordination in schizophrenia: A resting state fMRI study. Schizophr Res, 141: 1-7. doi: 10.1016/j.schres.2012.07.027

32. Kaufmann T, Alnæs D, Doan NT, Brandt CL, Andreassen OA, Westlye LT (2017) Delayed stabilization and individualization in connectome development are related to psychiatric disorders. Nat Neurosci, 20: 513-515. doi: 10.1038/nn.4511

33. Mantini D, Hasson U, Betti V, Perruci MG, Romani GL, Corbetta M, Orban GA, Vanduffel W (2012) Interspecies activity correlations reveal functional correspondence between monkey and human brain areas. Nat Meth, 9: 277-282. doi:10.1038/nmeth.1868

34. Matthews PM and Hampshire A (2016) Clinical Concepts Emerging from fMRI Functional Connectomics. Neuron, 91, 511-528; doi: 10.1016/j.neuron.2016.07.031

35. Miranda-Dominguez O, Mills BD, Carpenter SD, Grant KA, Kroenke CD, Nigg JT, Fair DA (2014) Connectotyping: model based fingerprinting of the functional connectome. PLoS One, 9: e111048. doi: 10.1371/journal.pone.0111048

36. Mitra A, Snyder AZ, Tagliazucchi E, Laufs H, Raichle ME (2015) Propagated infra-slow intrinsic brain activity reorganizes across wake and slow wave sleep. Elife, 4: e10781; doi: 10.7554/eLife.10781

37. Mueller S, Wang D, Fox MD, Pan R, Lu J, Li K, Sun W, Buckner RL, Liu H (2015) Reliability correction for functional connectivity: Theory and implementation. Hum Brain Mapp, 36: 4664-4680. doi: 10.1002/hbm.22947

38. Pannunzi M, Hindriks R, Bettinardi RG, Wenger E, Lisofsky N, Martensson J, Butler O, Filevich E, Becker M, Lochstet M, Kühn S, Deco G (2017) Resting-state fMRI correlations: From link-wise unreliability to whole brain stability. Neuroimage. 157: 250-262; doi: 10.1016/j.neuroimage.2017.06.006

39. Pretia MG, Bolton TAW, Van De Ville D (2016) The dynamic functional connectome: State-of-the-art and perspectives (in press) Neuroimage; https://doi.org/10.1016/j.neuroimage.2016.12.061

40. Raichle ME, MacLeod AM, Snyder AZ, Powers WJ, Gusnard DA, Shulman GL (2001) A default mode of

brain function. Proc Nat Acad Sci, 98: 676-682; doi:10.1073/pnas.98.2.676.

41. Rahim M, Thirion B, Bzdok D, Buvat I, Varoquaux G (2017) Joint prediction of multiple scores captures better individual traits from brain images. NeuroImage, 158: 145-154; doi: 10.1016/j.neuroimage.2017.06.072.

42. Rissman J, Wagner AD (2012) Distributed representations in memory: insights from functional brain imaging. Annu Rev Psychol, 63: 101-128. doi: 10.1146/annurev-psych-120710-100344

43. Shen H (2014) Neuroscience: Tuning the brain. Nature, 507: 290-292

44. Shehzad Z, Kelly AM, Reiss PT, Gee DG, Gotimer K, Uddin LQ, Lee SH, Margulies DS, Roy AK, Biswal BB, Petkova E, Castellanos FX, Milham MP (2009) Cereb Cortex, 19: 2209-2229. doi: 10.1093/cercor/bhn256

45. Stephan K, Harrison L, Penny W, Friston K (2004) Biophysical models of fMRI responses, Curr Opin Neurol 14: 629-635

46. Stephan KE, Mathys C (2014) Computational approaches to psychiatry. Curr Opin Neurobiol, 25: 85-92; doi: 10.1016/j.conb.2013.12.007

47. Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M (2002) Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage, 15: 273-289

48. Vanderwal T, Eilbott J, Finn ES, Craddock RC, Turnbull A, Castellanos FX (2017) Individual differences in functional connectivity during naturalistic viewing conditions. Neuroimage (in press)

49. Xie H, Calhoun VD, Gonzalez-Castillo J, Damaraju E, Miller R, Bandettini PA, Mitra S (2017) Whole-brain connectivity dynamics reflect both task- specific and individual-specific modulation: A multitask study, Neuroimage; doi:10.1016/j.neuroimage.2017.05.050

50. Yahata N, Kasai K, Kawato M (2017) Computational neuroscience approach to biomarkers and treatments for mental disorders. Psychiatry Clin Neurosci, 71: 215-237. doi: 10.1111/pcn.12502

51. Zuo X-N, Anderson JS, Bellec P, Birn RM, Biswal BB, Blautzik J, Buckner RL, Calhoun VD, Castellanos FX, Chen A, Chen B, Chen J, Chen X, Colcombe SJ, Courtney W, Craddock RC, Di Martino A, Dong HM, Fu X, Gong Q, Gorgolewski KJ, Han Y, He Y, He Y, Ho E, Holmes A, Hou XH, Huckins J, Jiang T, Jiang Y, Kelley W, Kelly C, King M, LaConte SM, Lainhart JE, Lei X, Li HJ, Li K, Li K, Lin Q, Liu D, Liu J, Liu X, Liu Y, Lu G, Lu J, Luna B, Luo J, Lurie D, Mao Y, Margulies DS, Mayer AR, Meindl T, Meyerand ME, Nan W, Nielsen JA, O'Connor D, Paulsen D, Prabhakaran V, Qi Z, Qiu J, Shao C, Shehzad Z, Tang W, Villringer A, Wang H, Wang K, Wei D, Wei GX, Weng XC, Wu X, Xu T, Yang N, Yang Z, Zang YF, Zhang L, Zhang Q, Zhang Z, Zhang Z, Zhao K, Zhen Z, Zhou Y, Zhu XT, Milham MP (2014) An open science resource for establishing reliability and reproducibility in functional connectomics. *Scientific Data*, *1*, 140049. http://doi.org/10.1038/sdata.2014.49

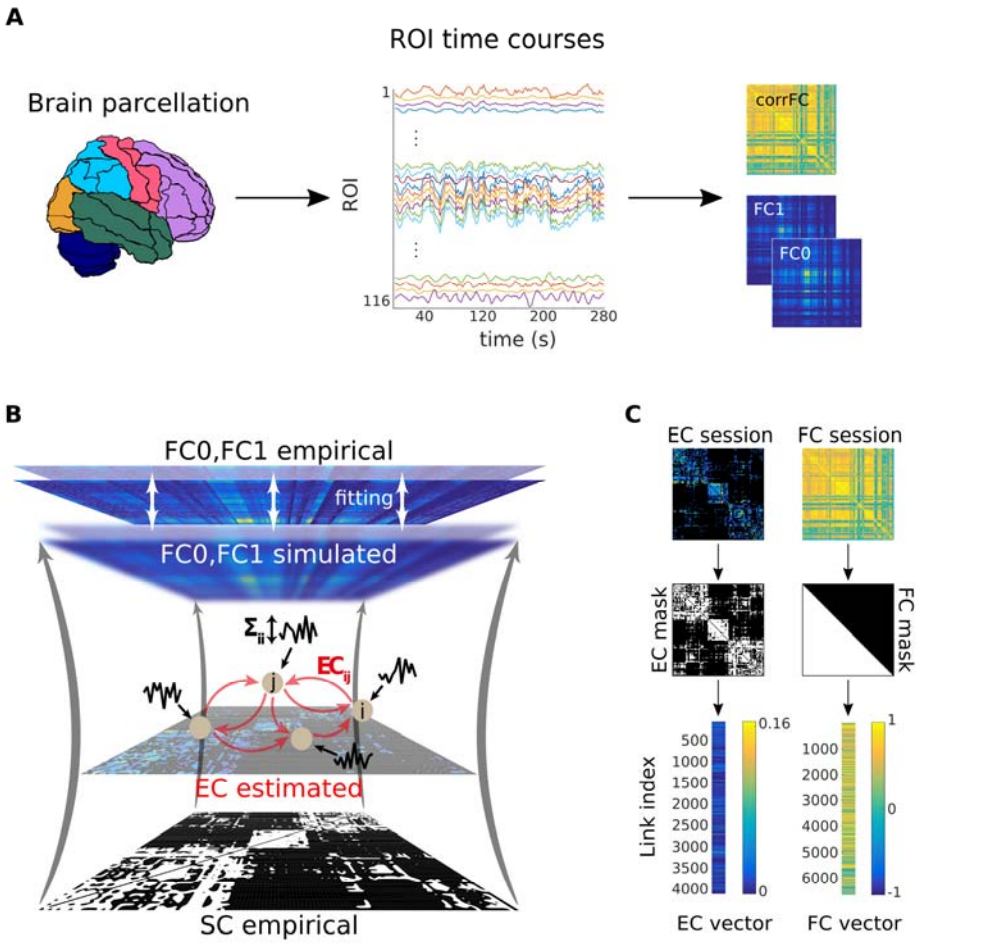

Figure 1: Workflow for the calculation of the connectivity measures from fMRI measurements. A) After a standard pre-processing pipeline, a parcellation covering the whole-brain is applied to extract BOLD time series: 116 ROIs for the AAL used here and 66 ROIs for the Hagmann parcellation; each color represents an anatomical subsystem of several ROIs. Here we consider two versions of functional connectivity: the classical corrFC corresponding to the Pearson correlation coefficient (PCC) between pairs of time series; the spatiotemporal FC embodied by the two covariance matrices FC0 and FC1 without and with time shift, respectively; see Eqs. (1) and (2) in Methods for details. **B)** Whole-brain network model to interpret fMRI data. The local fluctuating activity (where Σ_{ii} is the variance of the input to region i) propagates via the recurrent EC to generate the correlation patterns at the network level. Structural connectivity (SC, bottom) obtained using DTI determines the skeleton of EC. The fitting procedure iteratively tunes EC and Σ such that the model best reproduces the empirical FC0 and FC1. **C)** Each corrFC matrix is symmetric and has all diagonal elements equal to 1, so only 6670 independent links are retained for identification/classification (lower triangle). Likewise, the EC matrix has 4056 non-zero elements that are used in the classification (density of 30%).

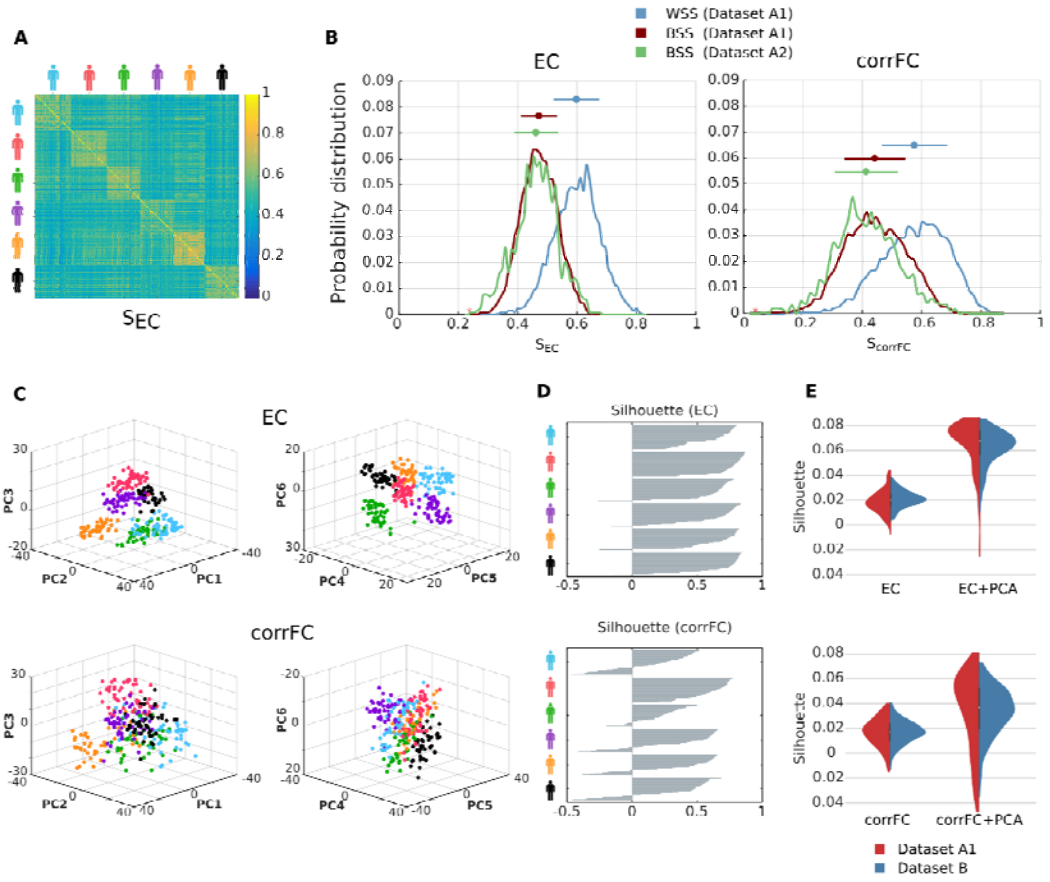

Figure 2: Within- and between-subject similarity (WSS and BSS, respectively) for EC and corrFC. A) Matrix of similarity values for EC between all pairs of sessions from Dataset A1. Practically, ECs from two sessions are transformed into two vectors (as illustrated in Figure 1C), from which the PCC is calculated to obtain S_{EC} (see Eq. 13 in Methods). The sessions are grouped by subjects, as indicated by the colored symbols. **B)** The left panel shows that distributions of WSS (blue) and BSS (red) values for Datasets A1 – corresponding to diagonal and off-diagonal blocks in panel A, respectively – and of BSS (green) for Dataset A2. The right panel shows the corresponding distributions for corrFC. The above error bars indicate the means and standard deviations, indicating smaller overlap between WSS and BSS for EC. **C)** Visualization of the sessions of Dataset A1 in the space of the first 6 PCs (split into the left and right panels) obtained from PCA for EC (top row) and corrFC (bottom row). Each point corresponds to a session and each color to one of the 6 subjects, as in panel A. **D)** Silhouette coefficients of each session in panel C (see main text and Eq. 15 in Methods for further details). **E)** Distribution of the silhouette coefficients for EC (top panel) and corrFC (bottom panel): comparison between the original link space (left) and the PCA space (right, corresponding to panel D). Both Datasets A1 (6 subjects with 6 PCs, in red) and B (30 subjects with 30 PCs, in blue) are represented by the violin plots; see also Figure S3 about the choice for the number of PCs. Note the larger silhouette coefficients for EC than corrFC.

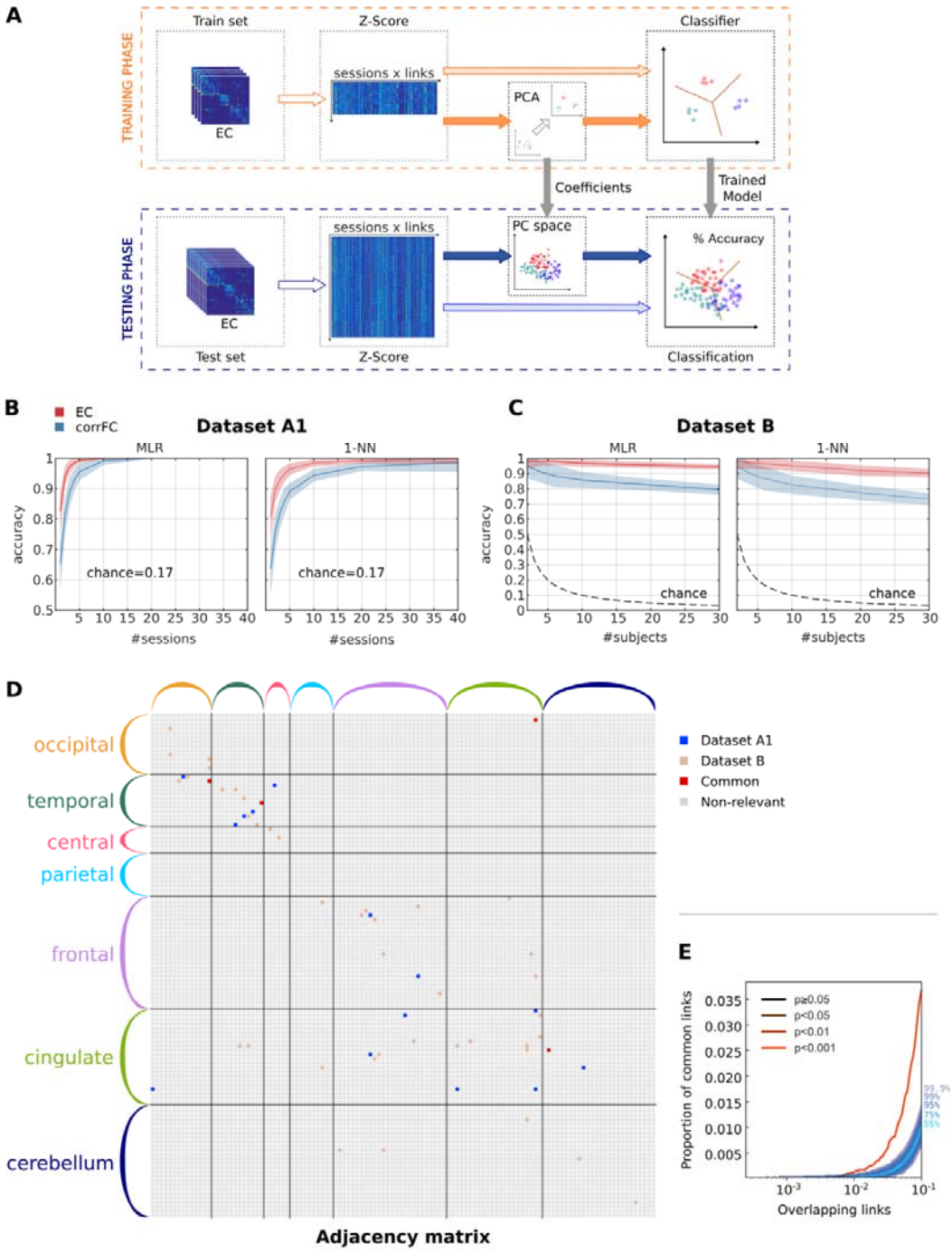

Figure 3: Subject identification using EC and FC. A) Classification pipeline used to assess the generalization performance. The full set of connectivity measures (here EC) over all fMRI sessions was split into two groups: a train set and a test set. We applied a standardization using z-score over the elements of each session matrix (see Eq. 16 in Methods). We trained the classifier – with or without previously applying PCA – and evaluated the classification accuracy on the test set. **B)** Performance of multinomial logistic regression (MLR, left panel) and 1-nearest-neighbor (1NN, right panel) classifiers when increasing the number of sessions per

subject used as training set with Dataset A1. The mean (solid curve) and standard deviation (colored area) were calculated for 100 repetitions with cross-validation. **C)** Same as B when varying the number of subjects using Dataset B, using a single training session per subject (leaving 9 sessions per subject as test test). **D)** Extracted links that contribute to the classification with both datasets, obtained using recursive feature elimination (RFE). The ROIs are grouped in anatomical pools, as detailed in Supplementary Table S1. **E)** Overlap between the two signatures for Datasets A1 and B as a function of selected links. The curve represents the amount of common links in the data. Shaded areas represent different quantiles of the surrogate distribution of common links under the null-hypothesis of random rankings. The color of the curve indicates the probability of the corresponding amount of common links under the null-hypothesis (here p-value < 0.001 when considering more than 1% of the total links, namely 40 links).

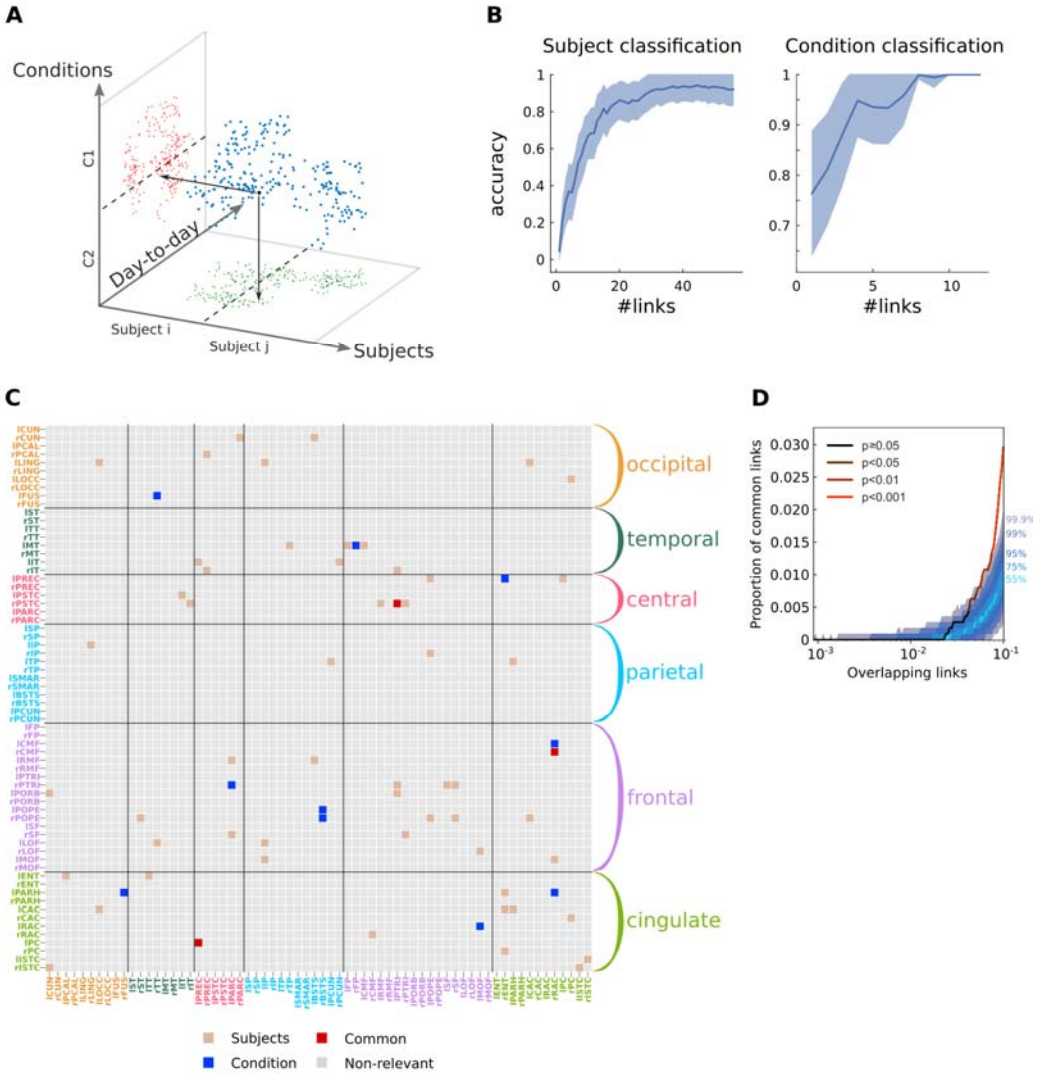

Figure 4: Two-fold discrimination between subjects and conditions using EC.
A) Idealized scheme of the twofold classification where each session (blue dots) is "projected" onto two subspaces, one for subjects (green) and one for conditions (red). In each subspace, classification can be performed efficiently. Depending on the subspaces orthogonality, the two signatures have more or less overlap. **B)** Performance of the classification for 19 subjects and 2 conditions using Dataset C as a function of number of links. Note the distinct scales for the y-axis, because the subject identification is a harder problem. **C)** Signatures of the most discriminative EC links (estimated with RFE, see text for details) for the twofold classification in B: 54 links for subject classification in brown, 10 for condition classification in blue, 3 common links in red. The ROIs are grouped in anatomical pools, as detailed in Supplementary Table S2. **D)** Proportion of common links between the subject and condition signatures as a function of selected links (in the order of the RFE ranking). Color coding is the same as in Figure 3E: the two signatures are significantly different, i.e., with a number of common links corresponding to the null hypothesis with p-value ≥ 0.05 (cf. legend) up to 4% of the total links.

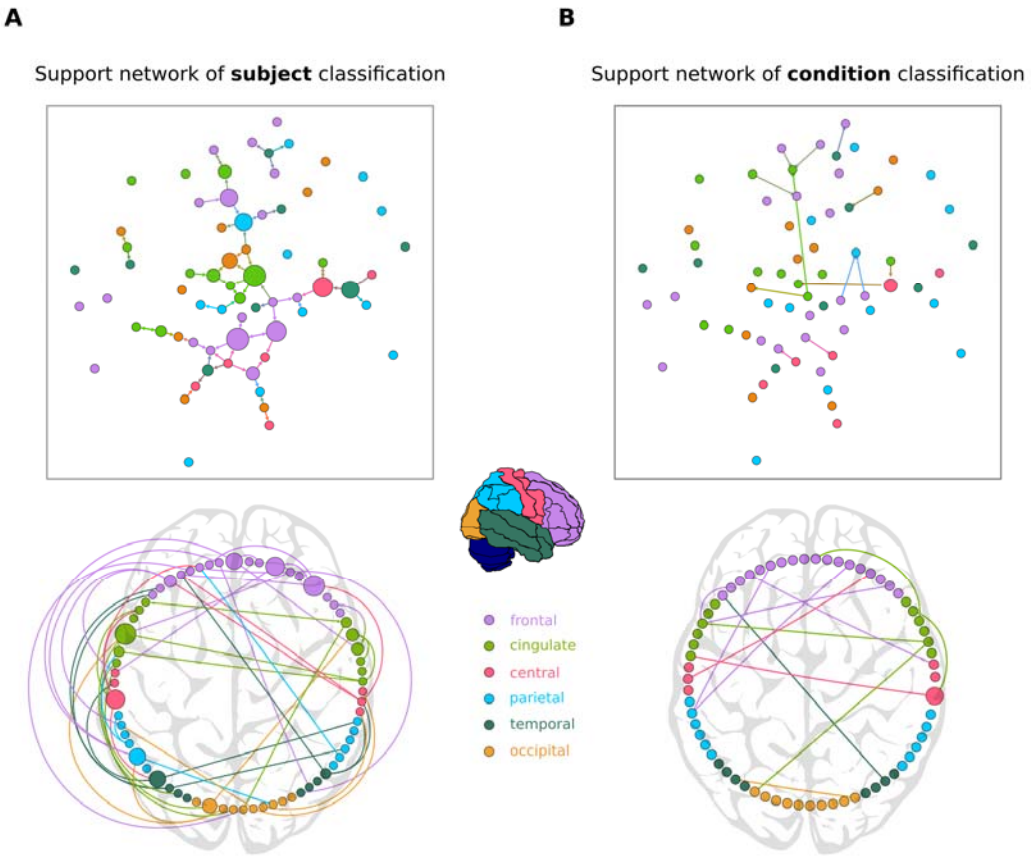

Figure 5: Support networks of subject and condition classification. A) The top graph plot represents the 57 most discriminative EC links supporting the classification of subjects (same as in Figure 3C). The size of each node represents its betweenness centrality in the extracted network. The most central regions are located mainly in the frontal and cingulate cortices. The bottom circular plot shows the asymmetry and lateralization of the network, with more links located in the left hemisphere. Links that are inside the circle correspond to contralateral connections, while links outside the circle are ipsilateral connections. **B)** Similar graph and circular plots as in A for the 13 links supporting the classification between the two conditions (resting versus movie viewing). Fewer links are required to reach high accuracy in the condition discrimination: they form a network with many disjoint components and are mainly contralateral, in comparison to the subject classification support network.