Supplemental Text

Text S1. Definitions of network metrics

Network strength. For a network (graph) G with N nodes and K edges, we calculated the strength of G as:

$$S_p(G) = \frac{1}{N}\sum_{i \in G} S(i)$$

where S(i) is the sum of the edge weights w_{ij} linking to node i. The strength of a network is the average of the strength across all of the nodes in the network.

Small-world properties. Small-world network parameters (clustering coefficient, C_p and shortest path length, L_p) were originally proposed by Watts and Strogatz (1998). In this study, we investigated the small-world properties of the weighted brain networks.

The clustering coefficient of a node i, C(i), which was defined as the likelihood of whether the neighborhoods were connected with each other or not, was computed as follows (Onnela et al., 2005):

$$C(i) = \frac{2}{k_i(k_i - 1)}\sum_{j,k} (\bar{w}_{ij}\bar{w}_{jk}\bar{w}_{ki})^{1/3}$$

where k_i is the degree of node i and $\bar{w}$ is the weight of edge, which is scaled by the largest weight of the network. The clustering coefficient is zero if the nodes are isolated or have just one connection, i.e., $k_i = 0$ or $k_i = 1$. The clustering coefficient, C_p, of a network is the average of the clustering coefficient over all nodes and indicates the extent of the local interconnectivity or cliquishness in a network (Watts and Strogatz, 1998).

The path length between any pair of nodes (e.g., node i and node j) is defined as the sum of the edge lengths along this path. For weighted networks, the length of each edge was assigned by computing the reciprocal of the edge weight, $1/w_{ij}$. The shortest path length, L_{ij}, is defined as the length of the path for node i and node j with the shortest length. The shortest path length of a network was computed as follows:

$$L_p(G) = \frac{1}{N(N-1)} \sum_{i \neq j \in G} L_{ij}$$

where N is the number of nodes in the network. The L_p of a network quantifies the ability for information to propagate in parallel.

To examine the small-world properties, the clustering coefficient, C_p, and the shortest path length, L_p, of the brain networks were compared with those of random networks. In this study, we generated 100 matched random networks, which had the same number of nodes, edges, and degree distribution as the real networks (Maslov and Sneppen, 2002). Of note, we retained the weight of each edge during the randomization procedure such that the weight distribution of the network was preserved. Furthermore, we computed the normalized L_p, $\lambda = L_p^{real} / L_p^{rand}$, and the normalized C_p, $\gamma = C_p^{real} / C_p^{rand}$, where L_p^{rand} and C_p^{rand} are the mean L_p and the mean C_p of 100 matched random networks, respectively. Importantly, two parameters correct the differences in the edge number and degree distribution of the networks across individuals. A real network would be considered small-world if $\gamma > 1$ and $\lambda \approx 1$ (Watts and Strogatz, 1998). Thus, a small-world network not only has a higher local interconnectivity, but it also has an approximately equivalent shortest path length compared with random networks. These two measurements can be summarized into a simple quantitative metric, small-worldness, $\sigma = \gamma / \lambda$, which is typically greater than 1 for small-world networks (Humphries and Gurney, 2008).

Network efficiency. The global efficiency of G measures the global efficiency of the parallel information transfer in the network (Latora and Marchiori, 2001), which can be computed as:

$$E_{glob}(G) = \frac{1}{N(N-1)} \sum_{i \neq j \in G} \frac{1}{L_{ij}}$$

where L_{ij} is the shortest path length between node i and node j in G.

The local efficiency of G reveals how much the network is fault tolerant and shows how efficient the communication is among the first neighbors of the node i when it is removed. The local efficiency of a graph is defined as:

$$E_{loc}(G) = \frac{1}{N} \sum_{i \in G} E_{glob}(G_i)$$

where G_i denotes the subgraph composed of the nearest neighbors of node i.

Text S2. Network-based statistic (NBS)

To localize the specific connected components in which the structural connectivity differed between each pair of groups, we used a NBS approach [34]. First, the significant non-zero connections (backbone) within each group were detected using a nonparametric one-tailed sign test ($p < 0.05$, corrected). Next, the non-zero connections within either the patient or control groups were established and combined into a connection mask. The NBS approach was then performed within the connection mask, where a primary threshold ($p = 0.05$) was first applied to a t-statistic (two-sample one-tailed t tests). This t-statistic was computed for each link to define a set of supra-threshold links among which any connected components and their size (number of links) could then be determined. To estimate the significance for each component, the null-distribution of the connected component size was empirically derived using a nonparametric permutation approach (10,000 permutations). For each permutation, all subjects were randomly reallocated into two groups, and the t statistic was computed independently for each link. Next, the threshold ($p = 0.05$) was used to generate suprathreshold links among which the maximal connected component size was recorded. Finally, for a connected component of size M detected in the right grouping of controls and patients, the corrected p value was determined by calculating the proportion of the 10,000 permutations for which the maximal connected component was larger than M. Notably, the effects of age and gender were removed using a regression analysis before the statistical analysis of the connections. For a detailed description, see Zalesky et al. [34].

Text S3. Group differences in global network metrics

Both the patients and control subjects showed a small-world topology (lambda ≈ 1, sigma > 1) of the brain structural network. Among the three groups, significant group effects in all global network metrics were observed (Figure S1). *Post hoc* comparisons showed reduced strength, decreased global and local efficiency, increased shortest path length and decreased clustering in both the MS and CIS patients compared with controls. MS patients also showed increased gamma and sigma relative to controls.

Compared with CIS patients, MS patients showed reduced strength, reduced global efficiency, increased shortest path length, and increased lambda, gamma and sigma (all $p < 0.05$).

Supplemental Figures

Figure S1. Group differences in the global network metrics. The bars and error bars represent the mean values and standard deviations of the network properties in each group after removing the effects of age and gender. Significantly reduced strength, global efficiency, local efficiency, clustering and increased shortest path length of the structural networks were observed in both CIS and MS patients relative to the controls (HC). *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.005$.

Figure S2. Reproducibility of the rich-club organization in L-AAL network. (A) Hub distribution of structural backbone network across all subjects. Network hubs are represented with nodes in red, with nodal size indicating the degree of the regions. (B) Mean normalized RC coefficient curve under a series of thresholds k for each group. (C) Group differences in the strength of the rich-club, feeder and local connections. The bars and error bars represent the mean values and standard deviations of the connection strength in each group after removing the effects of age and gender. *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.005$.