

Explore Functional Connectivity in Human Brain

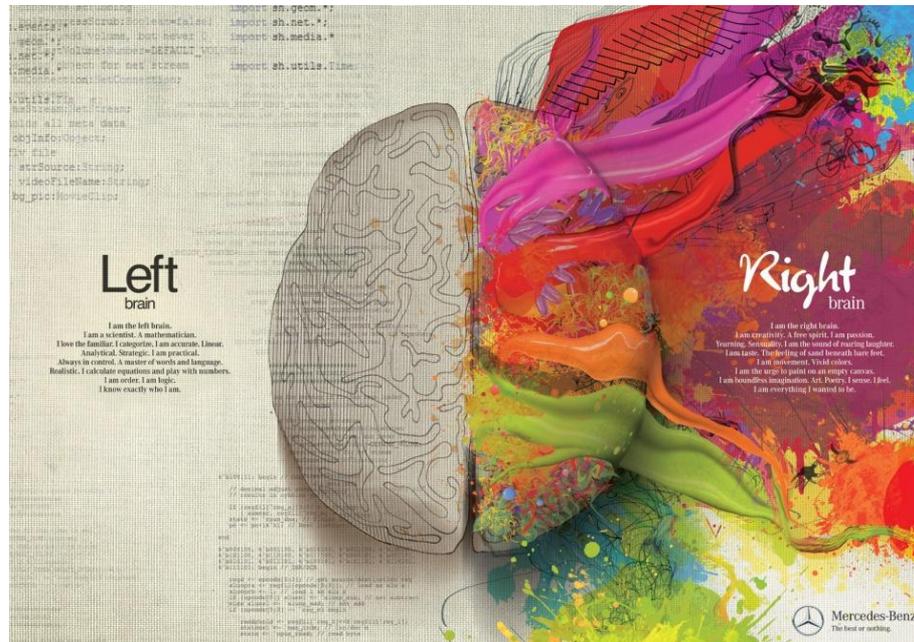
A Brief Survey

SongQi (Richard)

November 11, 2012

Why Functional Connectivity?

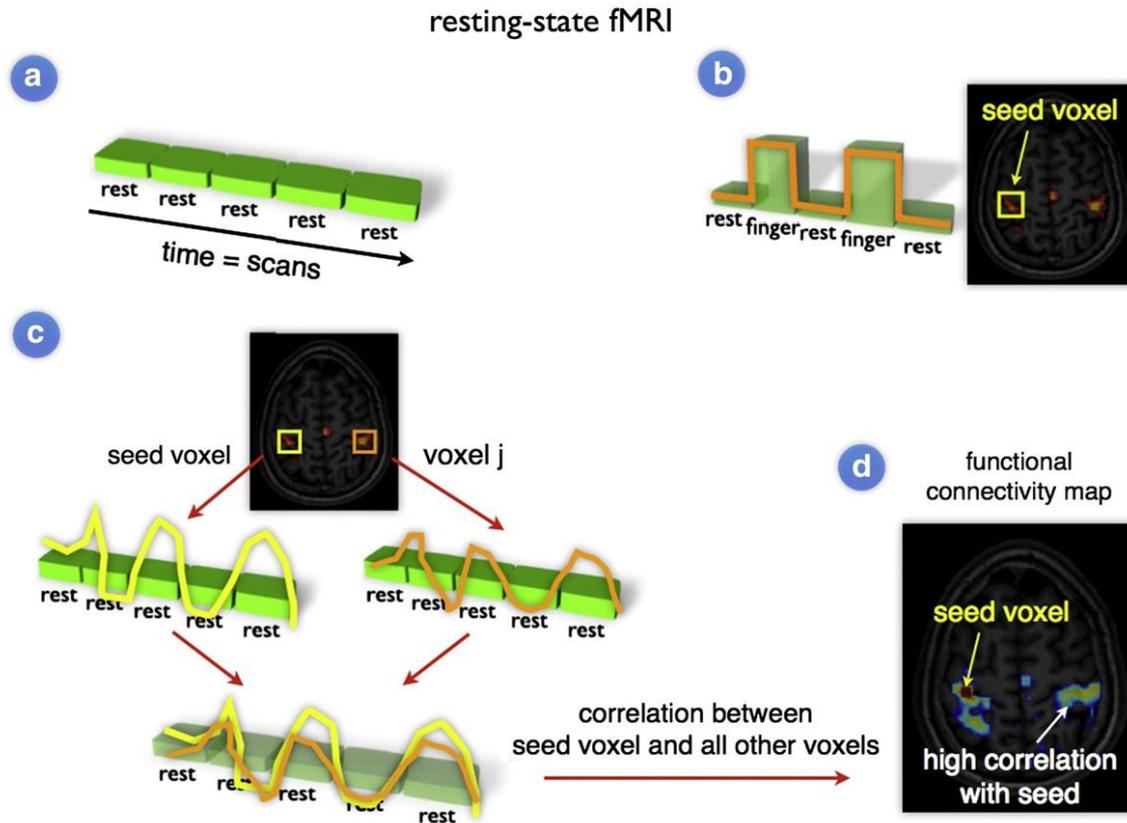
- Our Brain is a network consisting of *spatially distributed*, but *functionally linked* regions that continuously share information with each other.



What do we have for a Functional Connectivity?

- Resting-state fMRI time-series of each *ROI*
- Probability theories on *correlation*
- Graph theory and *complex network* analysis

How do we construct the network?



(Adopted from Martijn et al. 2010)

Time series: Preprocessing

Slice Timing Correction



Head Motion Correction



Distortion Correction



Coregistration & Normalization

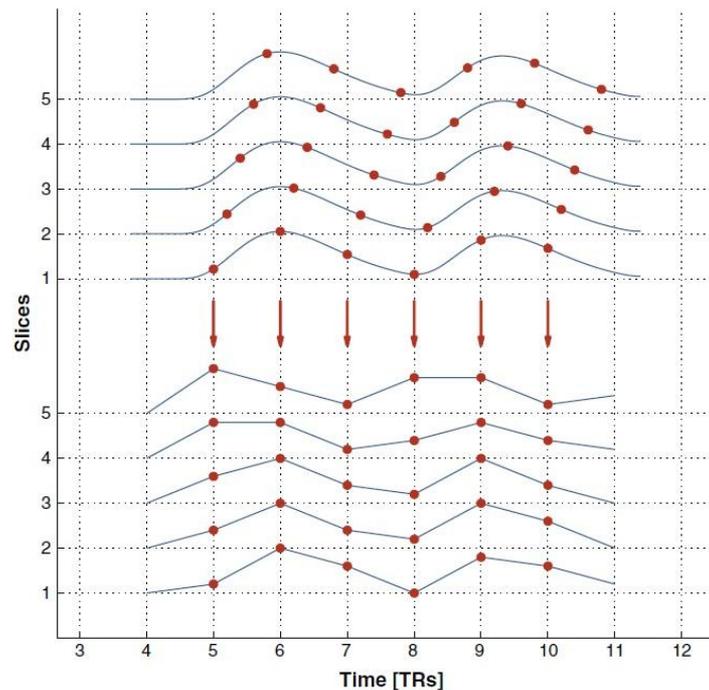


Spatial/Temporal Filtering

Preprocessing: Slice Timing

- 24 (or more) slices acquired within a single TR (1.5s-3.0s).....

Need some **Shifts!**



Slices acquired later reaches peak more quickly. (Sladky et al. 2011, *Neuroimage*)

Preprocessing: Slice Timing

- Shift the slices to the same time point: A **phase shift** in the **Fourier Domain**.

A *Fourier Transform* transforms signals in the **time domain** to the **frequency domain**, and is given by

$$F(f) = \int_{-\infty}^{\infty} f(t) \cdot e^{-2\pi f i t} dt$$

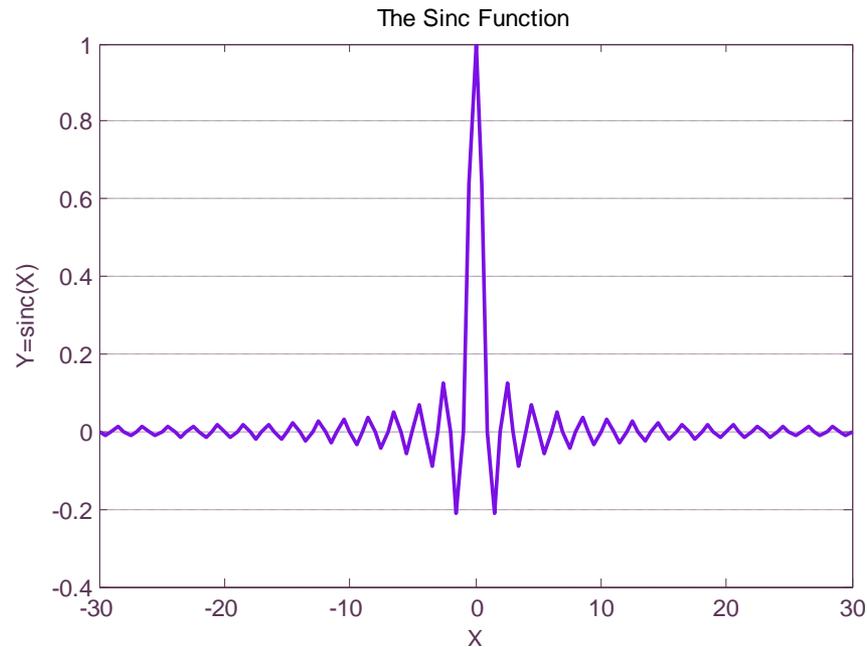
The *Reverse Fourier Transform* then transform the signal back to the time domain after a **phase shift**.

$$f(t) = \int_{-\infty}^{\infty} F(f) \cdot e^{2\pi f i t} df$$

Preprocessing: Slice Timing

- Method used in SPM8: *Linear interpolation* and *Sinc interpolation*

- A Sinc function is defined as $\frac{\sin x}{x}$



Preprocessing: Slice Timing

- The linear interpolation can be formally defined as

$$y_n^{(r)} = \frac{(t(r) - t(n-1))y_n + (t(n) - t(r))y_{n-1}}{t(n) - t(n-1)}$$

Where $y(r)$ is the corrected data and $t(r)$ is the **reference slice**.

- On the other hand, Sinc interpolation (which is believed to be more efficient), is formally defined as

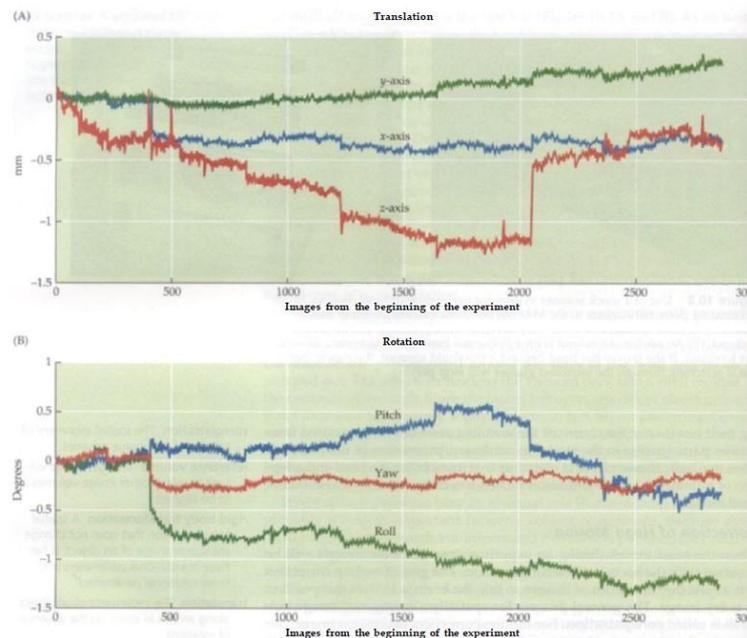
$$y_n^{(r)} = \sum_{i=-\infty}^{\infty} x_i \operatorname{sinc}\left(\frac{\pi}{\text{TR}}(r - i\text{TR})\right)$$

Which results from a phase shift in the frequency domain.

Preprocessing: Head Motion & Distortion Correction

- Subjects move their heads during scanning.....

Need a **Realignment!**



(Adopted from Scott et al. *Functional Magnetic Resonance Imaging*)

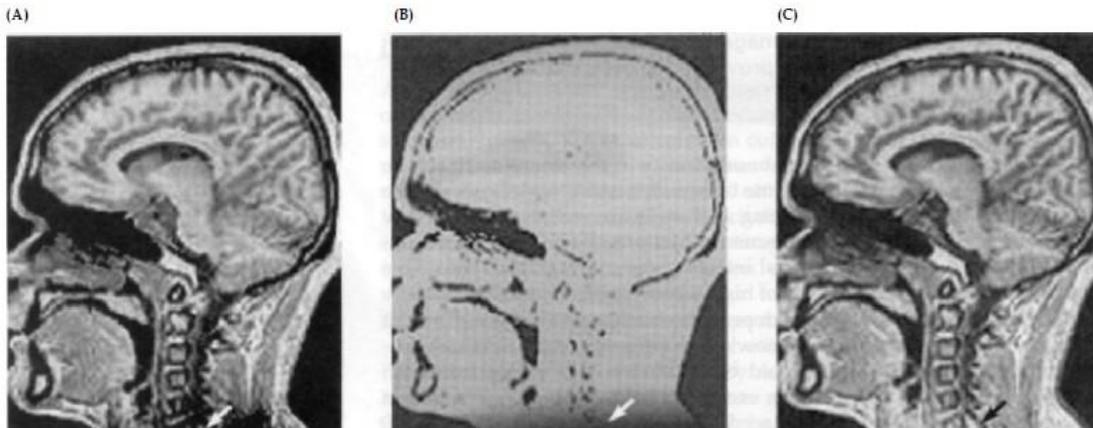
Preprocessing: Head Motion & Distortion Correction

- *Spatial Interpolation*

Similar to the method used in time slicing, while it's **2-dimensional**.

- *Magnetic Field Mapping*

- *Bias Field Estimation*



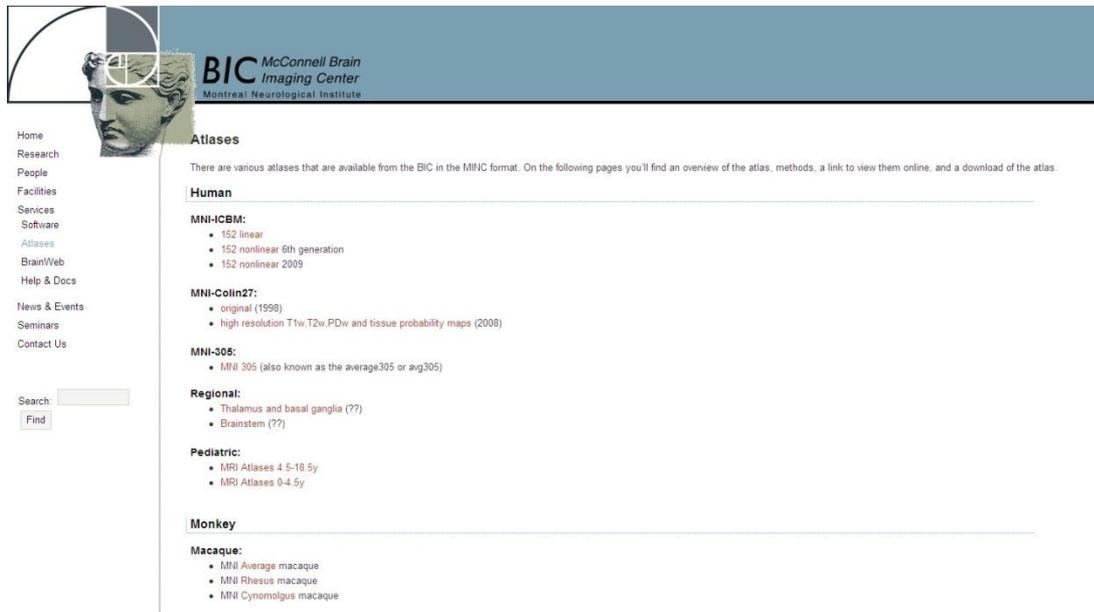
Preprocessing: Coregistration & Normalization

- A Common Normalization Scheme: A **Stereotaxic Space**

Talairach Space

MNI Template

(Developed by the Montreal Neurological Institute)



The screenshot shows the website for the BIC McConnell Brain Imaging Center at the Montreal Neurological Institute. The header features the BIC logo and a classical bust of a man's head. A navigation menu on the left includes links for Home, Research, People, Facilities, Services, Software, Atlases, BrainWeb, Help & Docs, News & Events, Seminars, and Contact Us. Below the menu is a search bar with a 'Find' button. The main content area is titled 'Atlases' and contains the following text: 'There are various atlases that are available from the BIC in the MINC format. On the following pages you'll find an overview of the atlas, methods, a link to view them online, and a download of the atlas.' Below this text are several sections: 'Human' (with sub-sections MNI-ICBM, MNI-Colin27, and MNI-305), 'Regional', 'Pediatric', and 'Monkey' (with sub-section Macaque). Each section lists specific atlas types and their characteristics.

BIC McConnell Brain Imaging Center
Montreal Neurological Institute

Atlases

There are various atlases that are available from the BIC in the MINC format. On the following pages you'll find an overview of the atlas, methods, a link to view them online, and a download of the atlas.

Human

MNI-ICBM:

- 152 linear
- 152 nonlinear 6th generation
- 152 nonlinear 2009

MNI-Colin27:

- original (1998)
- high resolution T1w, T2w, PDw and tissue probability maps (2008)

MNI-305:

- MNI 305 (also known as the average305 or avg305)

Regional:

- Thalamus and basal ganglia (??)
- Brainstem (??)

Pediatric:

- MRI Atlases 4.5-18.5y
- MRI Atlases 0-4.5y

Monkey

Macaque:

- MNI Average macaque
- MNI Rhesus macaque
- MNI Cynomolgus macaque

Correlation: PPMCC

- PPMCC (*Pearson's Product-Moment Correlation Coefficient*) is a measurement of **Linear dependence** between two variables.
- It is defined by (*for population*)

$$\rho_{X,Y} = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{E[(X - \mu_x)(Y - \mu_y)]}{\sigma_X \sigma_Y}$$

Or equivalently (*for a sample*)

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}}$$

Correlation: PPMCC

- The absolute value of PPMCC is less than or equal to 1.
- **Invariant** to linear transformations.
- A value of 1/-1 indicates **maximum** positive/negative correlation.

Correlation: Partial Correlation

- *Partial Correlation* measures the degree of association between two variables, with the effect of a set of **controlling variables** removed .
- Formally, given two variables of interest X and Y, together with a controlling variable Z, the partial correlation can be defined as the **residues** resulting from the **linear regression** of X with Y and X with Z.

That is

$$\rho_{XY \cdot Z} = \frac{\text{cov}(R_x, R_y)}{\sigma_{R_x} \sigma_{R_y}}$$

Correlation: Partial Correlation

- Solving the proposed linear regression problems amounts to finding the n-dimensional vectors

$$\mathbf{w}_X^* = \arg \min_{\mathbf{w}} \left\{ \sum_{i=1}^N (x_i - \langle \mathbf{w}, \mathbf{z}_i \rangle)^2 \right\}$$
$$\mathbf{w}_Y^* = \arg \min_{\mathbf{w}} \left\{ \sum_{i=1}^N (y_i - \langle \mathbf{w}, \mathbf{z}_i \rangle)^2 \right\}$$

- The Residues are then

$$r_{X,i} = x_i - \langle \mathbf{w}_X^*, \mathbf{z}_i \rangle$$
$$r_{Y,i} = y_i - \langle \mathbf{w}_Y^*, \mathbf{z}_i \rangle$$

Correlation: Partial Correlation

- Finally, the sample Partial correlation is defined as

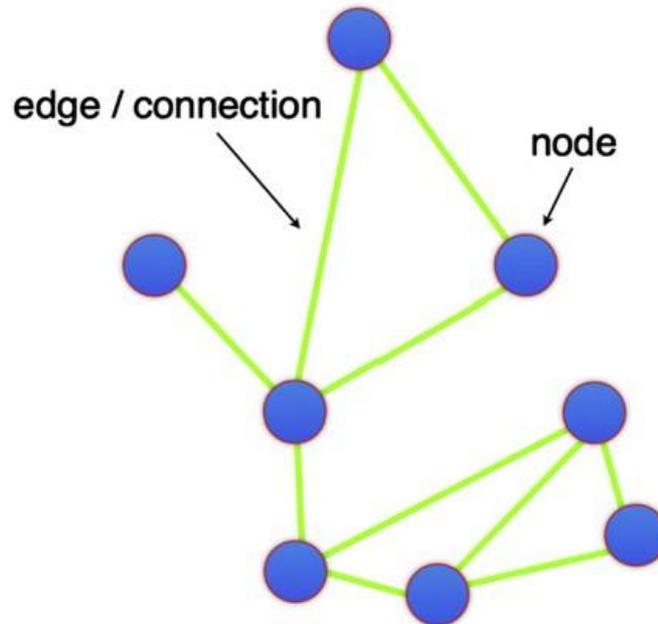
$$\hat{\rho}_{XY \cdot \mathbf{Z}} = \frac{N \sum_{i=1}^N r_{X,i} r_{Y,i} - \sum_{i=1}^N r_{X,i} \sum_{i=1}^N r_{Y,i}}{\sqrt{N \sum_{i=1}^N r_{X,i}^2 - \left(\sum_{i=1}^N r_{X,i} \right)^2} \sqrt{N \sum_{i=1}^N r_{Y,i}^2 - \left(\sum_{i=1}^N r_{Y,i} \right)^2}}.$$

Graph Theory: Basics

- Brain functional network is an **undirected** graph consisting of various **nodes** (ROIs) and **edges** (Functional connectivity)

(Effective connectivity map can be a directed one, though.)

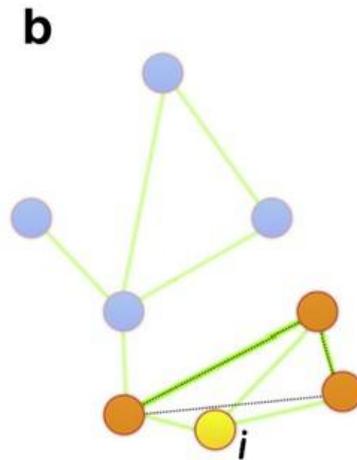
a



Graph Theory: Clustering-Coefficient

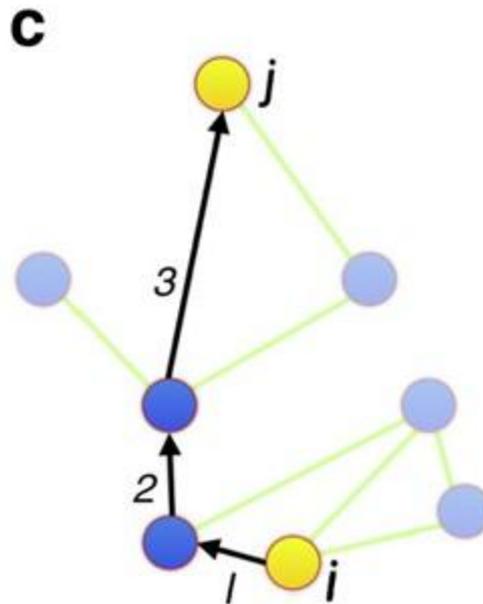
- Clustering-Coefficient of node i characterized the **local connectedness** and is defined by

$$\frac{\text{Number of connections between } i \text{ and its neighbors}}{\text{Maximum number of possible connections between } i \text{ and its neighbors}}$$



Graph Theory: Characteristic Path Length

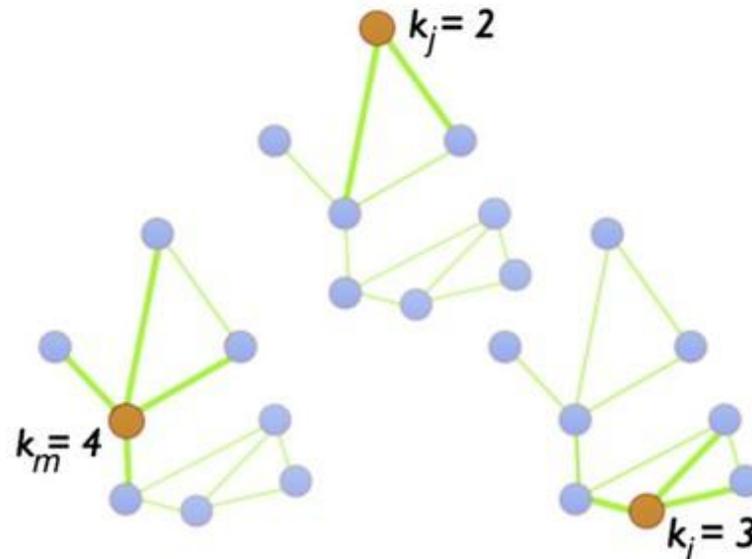
- *Characteristic Path Length* provides information about how close node i is connected to all other nodes. It is defined as $d(i,j)$, the distance between node i and node j .



Graph Theory: Centrality

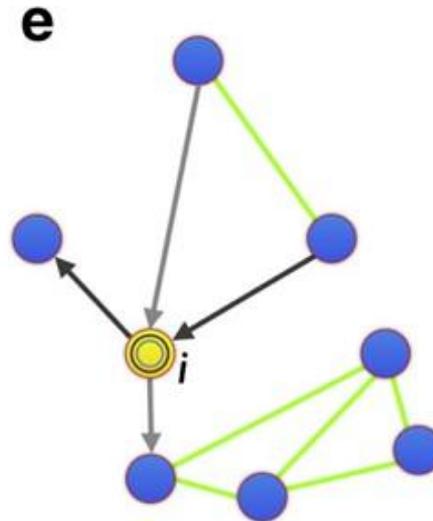
- Centralities of a graph are called *hubs*. A hub has a **small** value of characteristic length and a **large** value of clustering-coefficient.

d



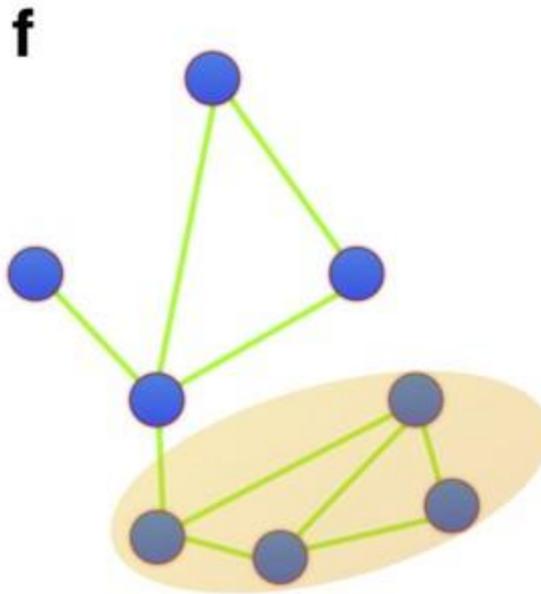
Graph Theory: Degree

- *Degree* of a node characterizes its connectedness and is defined by the **total number of connections** around that particular node.



Graph Theory: Modularity

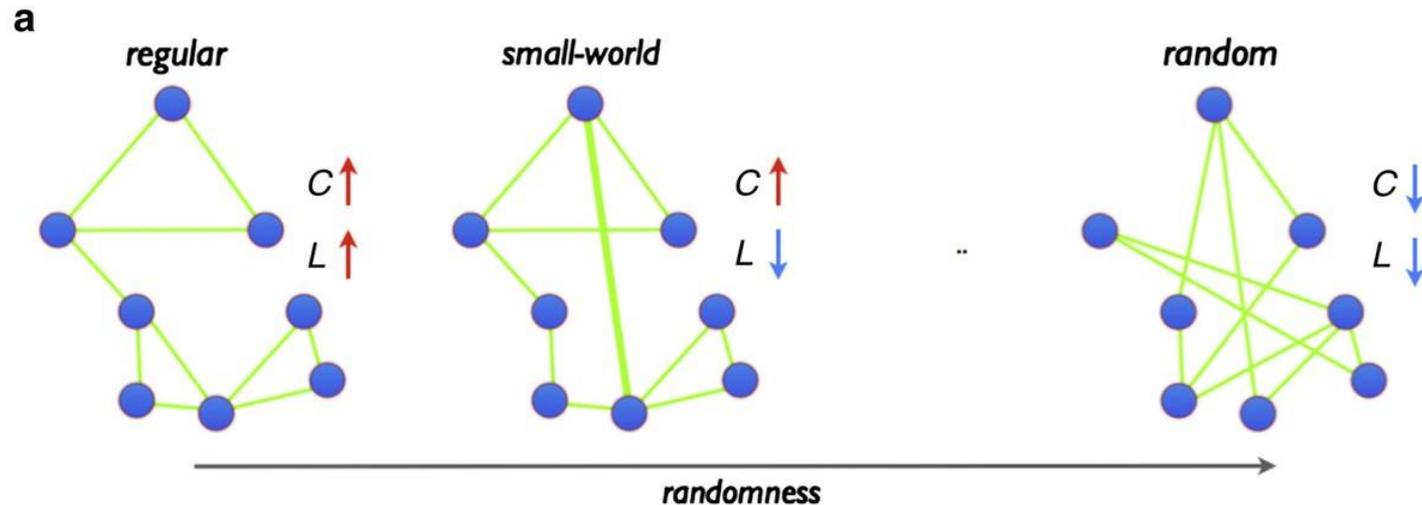
- *Modularity*, or **network communities** refer to groups of nodes within which the connecting links are **dense** but between which they are **sparse**.



Complex Network: Small-World Networks

- Our Brain is a *small-world network*,

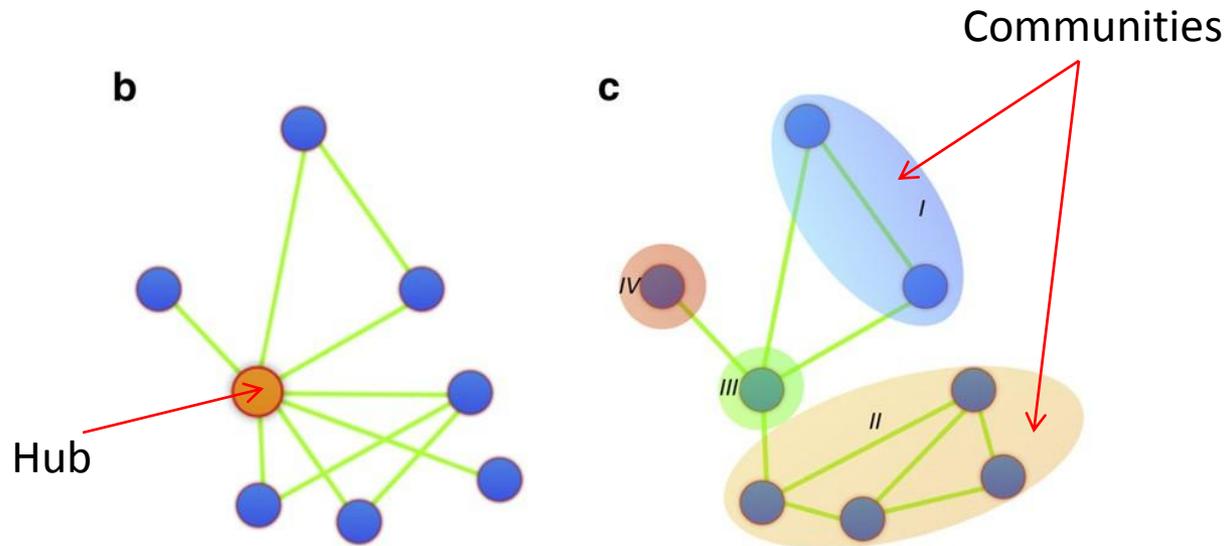
i.e. A network with **high** level of **local connectedness**
and **Low** level of **average characteristic length**



Complex Network: Scale-Free Networks

- Our Brain is a *Scale-Free network*,

i.e. A network with **Hubs** and **Communities**



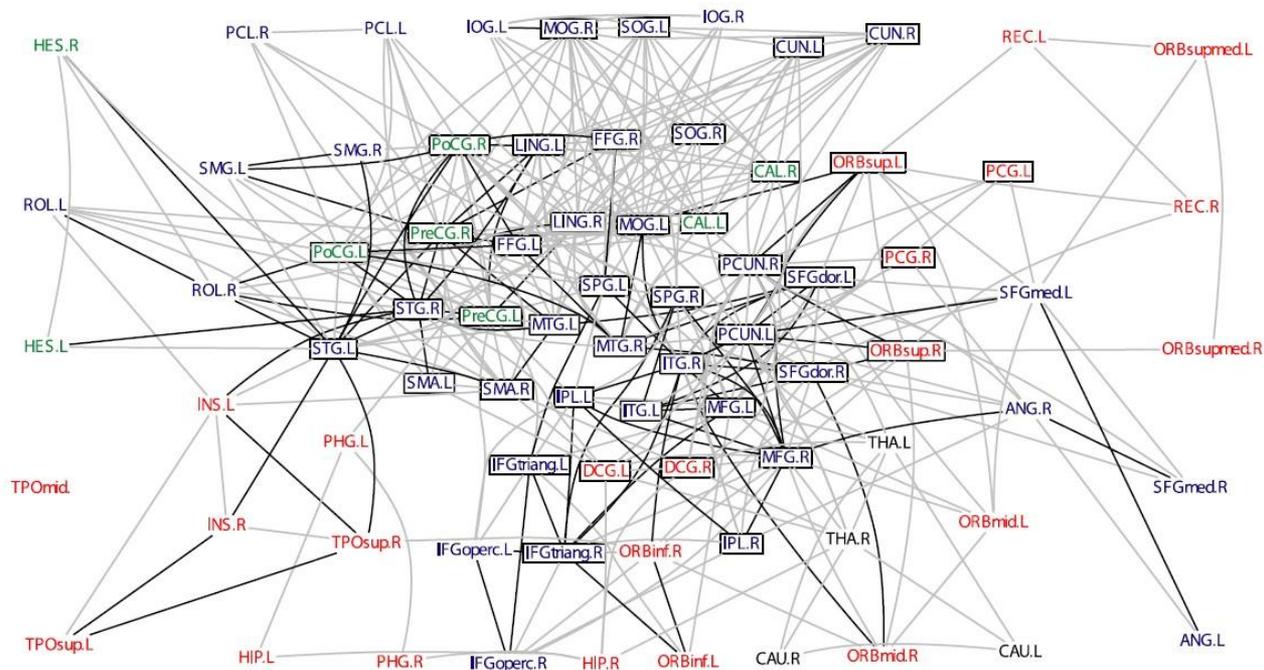
Robust against **random** attacks but **vulnerable** to specialized attacks on hub nodes

Sample Work: Construction of a whole brain network

Sophie et al. 2006, J.Neurosci.

A Resilient, Low-Frequency, Small-World Human Brain Functional Network with Highly Connected Association Cortical Hubs

(Brain mapping Unit and Wolfson Brain Imaging Centre, University of Cambridge)



Sample Work: Construction of a whole brain network

- Preprocessing of time series data
- Correlation Analysis: **MODWT** (*Maximal Overlap Discrete Wavelet Transform*)

$$W_{j,t}^{(X)} = \sum_{l=0}^{L_j-1} \tilde{h}_{j,l} X_{t-l \bmod N}$$

The scale-dependent Covariance of two time series is the given by

$$\gamma_{XY}(\lambda_j) = \frac{1}{N_j} \sum_{l=L_j-1}^{N-1} W_{j,l}^{(X)} W_{j,l}^{(Y)}$$

Sample Work: Construction of a whole brain network

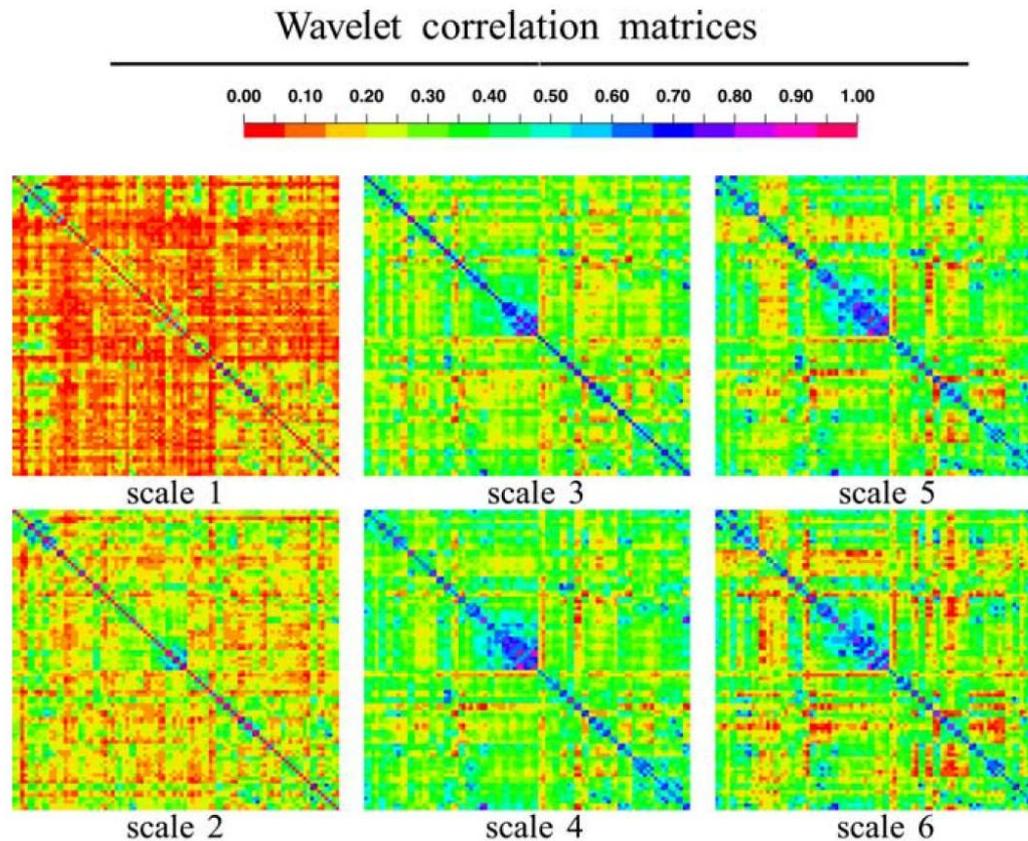
- Finally, the cross-correlation between two time series is defined by

$$\rho_{XY}(\lambda_j) = \frac{\gamma_{XY}(\lambda_j)}{v_X(\lambda_j)v_Y(\lambda_j)}$$

- A filter is then applied to the transformed signals to capture the **very low frequency** (< 0.1Hz) signal components, by which **resting-state correlations** are predominately subtended by.

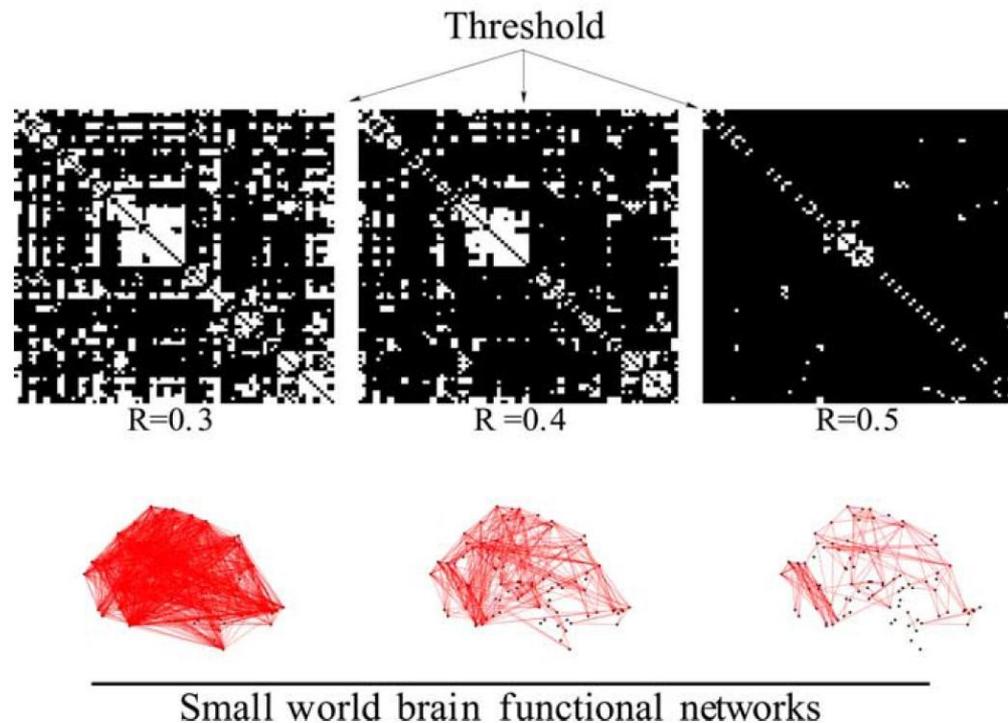
Sample Work: Construction of a whole brain network

- Correlation matrices of 90 X 90 brain regions at 6 different MODWT scales.



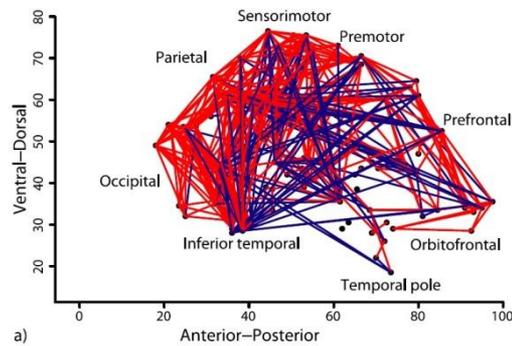
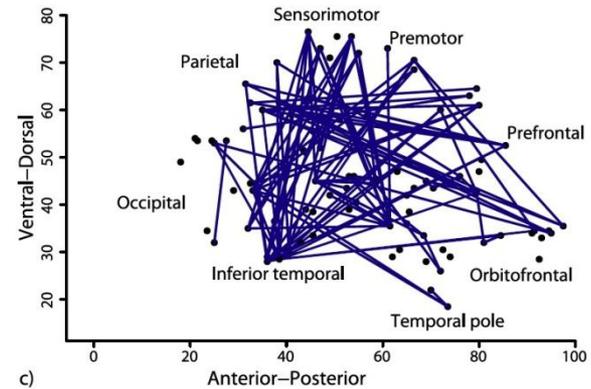
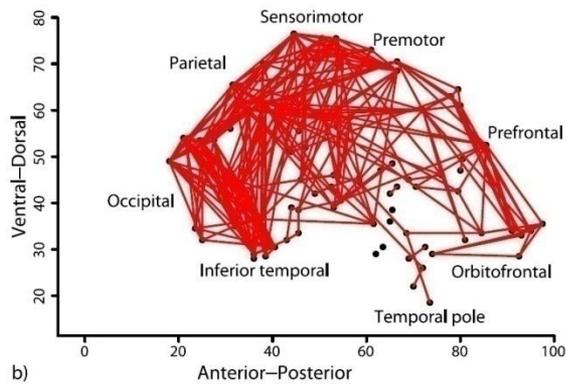
Sample Work: Construction of a whole brain network

- If a correlation value between two ROIs surpasses a certain **threshold**, then a functional connection is considered to be existent.



Sample Work: Construction of a whole brain network

- Anatomically near connections: Red
- Anatomically remote connections: Blue



Functional Connectivity and Brain Disorders

- Altered resting-state functional connectivity occur in various disorders.

Neurological/ **Psychiatric**

Depression uncouples brain hate circuit

H Tao^{1,6}, S Guo^{2,6}, T Ge^{3,6}, KM Kendrick^{4,6}, Z Xue¹, Z Liu¹ and J Feng^{3,5}

¹Institute of Mental Health, Second Xiangya Hospital, Central South University, Changsha, PR China; ²Mathematics and Computer Science College, Hunan Normal University, Changsha, PR China; ³Centre for Computational Systems Biology, School of Mathematical Sciences, Fudan University, Shanghai, PR China; ⁴Cognitive and Systems Neuroscience Group, The Babraham Institute, Cambridge, UK and ⁵Department of Computer Science, University of Warwick, Coventry, UK

Molecular Psychiatry (2011)

Functional Connectivity and Brain Disorders

- Construction of canonical template of healthy individuals.
Six discrete interconnected communities (Mined through the **LM algorithm**)
- Altered “Hate Circuit” in **FEMDD** (First-Episode Major Depressive Disorder) and **RMDD** (Resistant Major Depressive Disorder)
- Calculate **scores** of particular links, and observe changes in the connectivity
- Comparison, comparison, **comparison**.....

Preprocessing (Software-aided)

- Data preprocessing with SPM8 and DPARSF.
- What is expected: *.mat A 3D time series matrix

[Time * Region * Subject]

Where the *Time* dimension size corresponds to Length of the mean time series

Region

Number of ROIs

Subject

Number of subjects

Correlation Calculation

- Calculation of Pearson correlation (standard)

```
if any(data~=0 & data~=1)
    fprintf('*\tStart computing partial-correlation matrix ...\n');
    cm = zeros(reg, reg, sub); % correlation matrix
    pv = cm; % p-value
    tm = cm; % result of the test

    for n = 1:sub
        [cm(:, :, n), pv(:, :, n)] = corrcoef(data(:, :, n));
    end
```

Every time instant is an observation and different regions are the random variables.

- Correlations that are not strong enough to be identified as a functional connection can be eliminated using **pv**.

Correlation Calculation

- Calculation of Partial correlation

```
for n = 1:sub
    fprintf('*\t--> %d <-- \n',n);

    tm(:,:,n) = (pv(:,:,n) < handles.val(2)); % Find significant links
    [idx, idy] = find(tril(tm(:,:,n), -1)); % Mark the lower triangle matrix
    for k = 1:length(idx)
        for r = 1:reg
            if r~=idx(k) && r~=idy(k)
                [pc, p] = partialcorr(data(:, idx(k), n), data(:, idy(k), n), data(:, r, n)); % partial correlation
                if p > handles.val(2)
                    tm(idx(k), idy(k), n) = 0; % Eliminate non-significant links
                    tm(idy(k), idx(k), n) = 0;
                    break;
                end
            end
        end
    end
end
```

Establish the Functional Connectivity

- A 90 X 90 connection matrix was obtained for **each subject** with a P-value < 0.01
- The population level network is obtained by summarizing individual networks in FEMDD, RMDD and Healthy subject groups, **respectively**.
- The population matrices are then **thresholded** into binary values (1/0). According to a reasonable **sparsity** value
- The final matrices: 1 stands for a connection
0 stands for no connection

Establish the Functional Connectivity

```
tmp = (1:sub) < (handles.val(1) + 0.0001);
pm = sum(tm(:, :, tmp), 3); % patient matrix
nm = sum(tm(:, :, ~tmp), 3); % normal matrix

% >> transform to 0-1 matrix
[pm01, per] = to01(pm, handles.val(3));
set(handles.editSpa1, 'String', num2str(per));
handles.val(3) = per;

[rm01, per] = to01(rn, handles.val(4));
set(handles.editSpa2, 'String', num2str(per));
handles.val(4) = per;
fprintf(' *tThe true sparsity of net is:\n');
fprintf(' *t\t(net1)%0.3f%%, (net2)%0.3f%%\n', 100*handles.val(3:4));
```

Binary operation based on user-specified sparsity value

Score Calculation

- Significance of change in a particular link is characterized by

$$S = \frac{L_h}{N_h} - \frac{L_p}{N_p}$$

Where $L(h)$ denotes number of the link presented in healthy individuals
 $L(p)$ denotes number of the link presented in Depressed individuals
 $N(h)$ denotes the **total number** of healthy subjects
 $N(p)$ denotes the **total number** of depressed subjects
(FEMDD and RMDD, respectively)

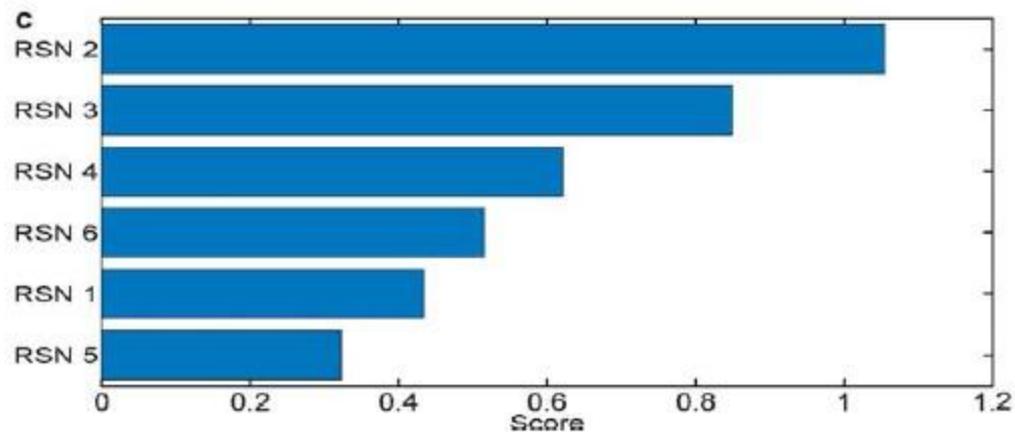
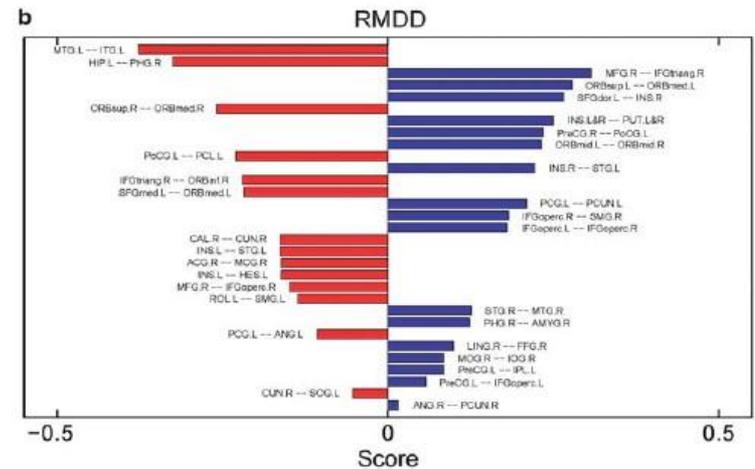
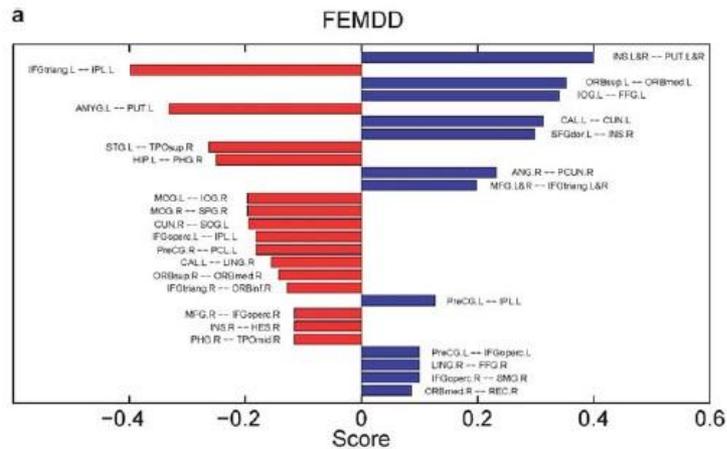
Score Calculation

```
% >> score of particular link
nump = fix(handles.val(1)); % number of patient
plink = xor(pm01,rm01); % particular link
score = (rm/(sub - nump) - pm/nump).*plink;
[II, JJ, VAL] = find(tril(score,-1));
score_link = sortrows([II JJ VAL abs(VAL)], -4);
score_link(:,end) = [];
fprintf('*\tAll %d particular-links: (net1)%d, (net2)%d \n\n',...
        size(II,1), sum(VAL>0), sum(VAL<0));

% >> compute path & score of them
fprintf('*\tCalculating score of path ...');
fprintf('\n*\tNet1 Path \n');
allpath_p = search_path(pm01.*plink);
score_path_p = path_score(allpath_p, score);
fprintf('\n*\tNet2 Path \n');
allpath_n = search_path(rm01.*plink);
score_path_n = path_score(allpath_n, score);
```

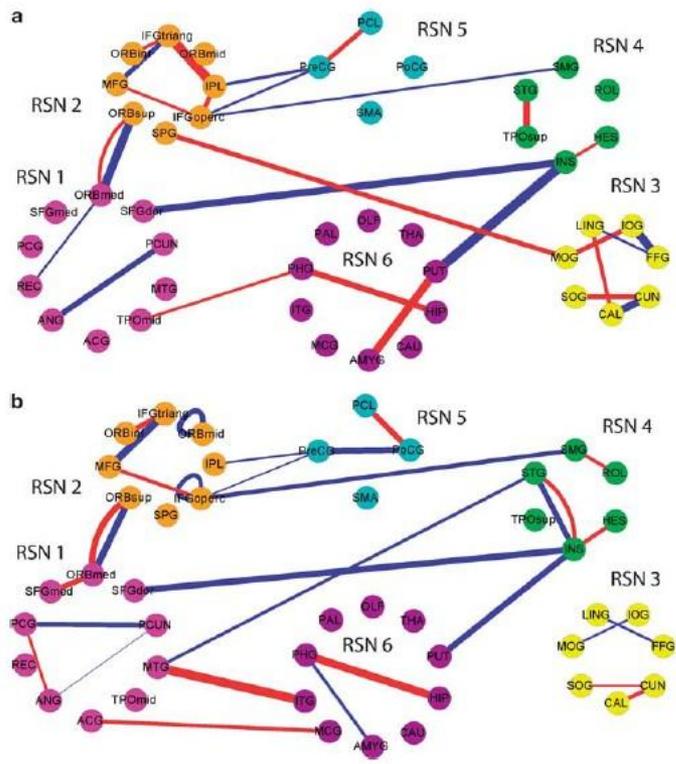


Score Calculation



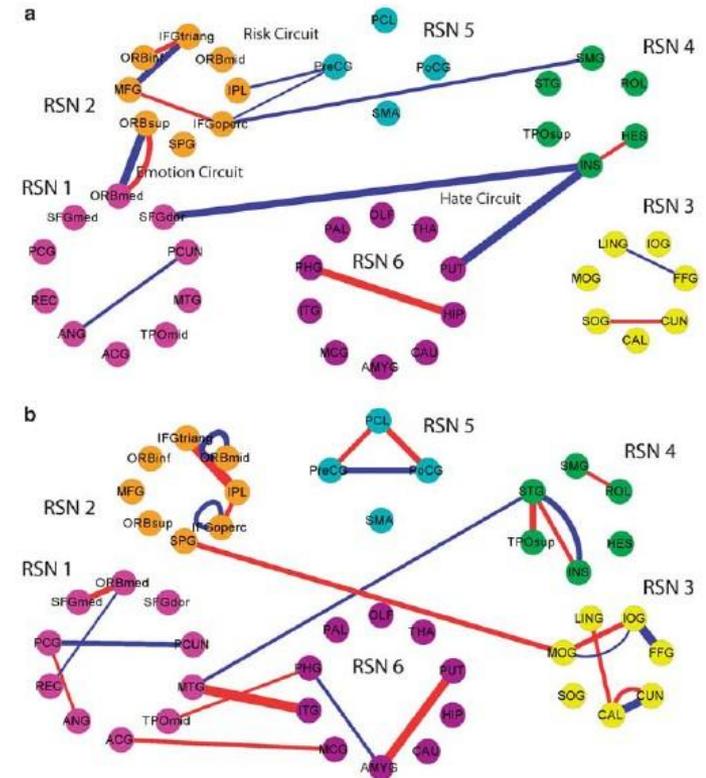
Change in the Network Communities

Patient (red) vs. Normal (blue)



a. FEMDD

b. RMDD



a. common

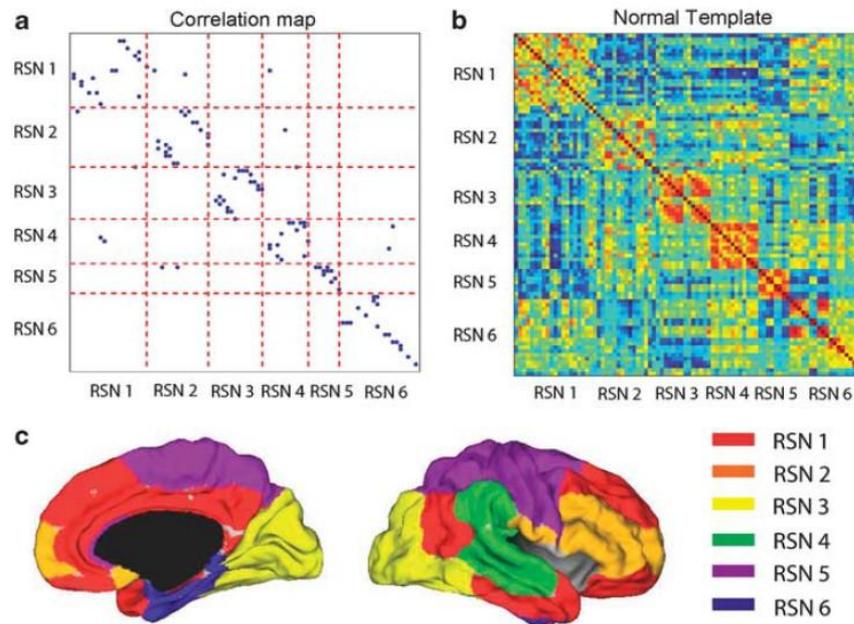
b. different

Brief Conclusions

- No significant gray/white matter changes: Disorders are not caused by reduced tissue volumes
- Changes in the “hate circuit” (superior frontal gyrus, insula and putamen) emotion and risk/action circuits
- Significant difference between FEMDD and RMDD
Depression should not be studied based on a single patient type

Canonical Template

- RSN 1: Default Mode Network
- RSN 2: Attention Network
- RSN 3: Visual Recognition Network
- RSN 4: Auditory Network
- RSN 5: Sensory-Motor Network
- RSN 6: Subcortical Network



The Community Mining Algorithm

How are the six communities mined?

- Consider a **Stochastic Process** defined on the network, in which one **agent** walks freely from one node to one of its neighbours.
- We define the probability of the agent's hitting node i after t steps to be

$$P(X_t = i, 1 \leq i \leq n)$$

- This process is actually a **Markov Chain**, in which the agent's current position depends solely on its previous position

$$P(X_t = i_t | X_0 = i_0, X_1 = i_1, \dots, X_{t-1} = i_{t-1}) = P(X_t = i_t | X_{t-1} = i_{t-1}).$$

The Community Mining Algorithm

- The **transition probability** from node i to node j satisfies

$$P(X_t = j | X_{t-1} = i) = p_{ij} \qquad p_{ij} = \frac{a_{ij}}{\sum_j a_{ij}}$$

Where $a(ij)$ is an element of the **adjacent matrix** of the network.

- The transition probability for the whole matrix can then be defined as

$$\mathbf{P} = \mathbf{D}^{-1} \mathbf{A}$$

Where $\mathbf{P} = (p_{ij})_{n \times n}$ $\mathbf{D} = \text{diag}\{d_1, d_2, \dots, d_n\}$ $d_i = \sum_j a_{ij}$

The Community Mining Algorithm

- Now let's assume that there are six clear-cut communities hidden in our network, but we do not know who's who.
- If we regard the agent's walking paths as connections, then the following statement is true:

The connections will first be establish **within** each community (**Local mixing states**). After all communities have been internally connected, the agent then starts connecting the six communities to a whole (**Global mixing state**)

(The Proof is skipped here 😊)

The Community Mining Algorithm

- We then define the time of each local mixing state as

$$T_s^{ext}, 1 \leq s \leq n$$

According to the *Large-Deviation Theory*, it can be estimated from the spectrum of the Markov Generator $Q = I - P$

$$T_s^{ext} \approx 1/\lambda_s, s = 1, \dots, n$$

Where the lambdas are the eigenvalues of the generator Q.

We further define the hitting time of the next stage to be

$$T_s^{hit} = T_{s+1}^{ext} \approx 1/\lambda_{s+1}$$

The Community Mining Algorithm

- By intuition, we can argue that if the agent spends little time walking internally within each communities, then the **cohesion** of the communities are good.

Thus, a small value of $C_K = T_K^{hit} \approx 1/\lambda_{K+1}$ indicates a good cohesion.

- Also, the **Separability** of the total network will be good, if the agent spends a long time connecting all the communities.

Hence, a large value of $S_K = T_K^{ext} - T_K^{hit} \approx 1/\lambda_K - 1/\lambda_{K+1}$ indicates a good separability

The Community Mining Algorithm

- Finally, we define the spectral signature of a network as

$$CQ_k = \frac{C_k}{C_k + S_k} = \frac{\lambda_k}{\lambda_{k+1}}, \quad 1 \leq k \leq n - 1$$

Clearly, a small value of it indicates a well-formed community network
(with both good within-community cohesion and inter-community separation)

The Community Mining Algorithm

The Traditional Algorithm

- 1) construct the transition matrix P of a given network;
- 2) calculate the spectrum of $I - P$ ($\lambda_1 \leq \dots \leq \lambda_n$);
- 3) calculate its spectral signature (CQ_1, \dots, CQ_{n-1}) and find the minimum CQ_K ($1 < K < n$);
- 4) characterize and analyze network' communities with the above spectral signature;
- 5) mine K communities from the matrix $P^{1/\lambda_{K+1}}$ if CQ_K is lower than a threshold.

The Community Mining Algorithm

The LM Algorithm

- 1) select an attractor with the maximum degree;
- 2) calculate an OTD regarding to this attractor;
- 3) bisection this OTD by its mean;
- 4) return if stopping criterion is satisfied; otherwise
- 5) bipartition the network and recursively manipulate two sub-networks.

Thanks!

*Ying Cai Experimental School
University of Electronic Science & Technology of China (UESTC)
No.4 Block 2 N.Jianshe Road, Chengdu, Sichuan, China
Postcode: 610054*

 +86 18696841639

 songqi.uest@gmail.com