Supplementary to Optimized Combination of Multiple Graphs with Application to the Integration of Brain Imaging and (epi)Genomics Data
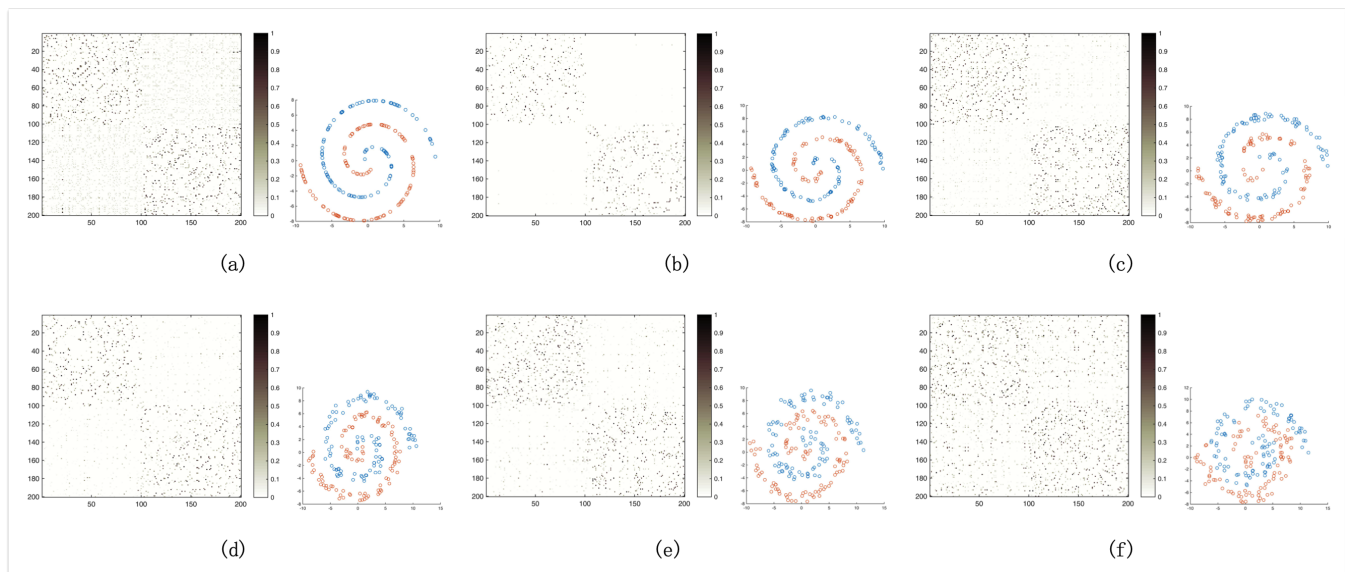

(a) (b) (c)

(d) (e) (f)

Figure S1: The simulated two-spiral pattern toy dataset. From (a) to (f): the noise levels L are successively 0,0.5,1,1.5,2,and 2.5; left figures are the similarity matrix calculated using Equation (11), and right figures are visualizations of the two-dimensional data, where two different colors represent the binary phenotypes.
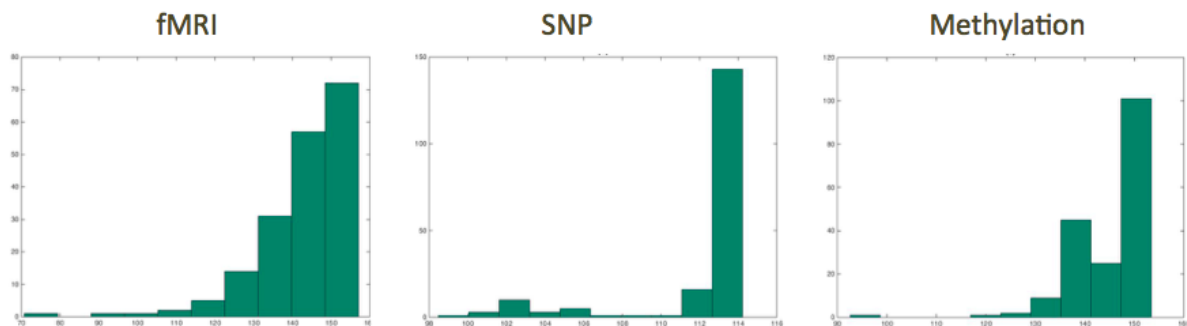

Figure S2: Histograms of vector D calculated from similarity matrices extracted from MCIC data constructed with bandwidth $\epsilon_{\text{connected}}$. From left to right is histogram from fMRI, SNP and DNA methylation data respectively. Denoting the minimum value in D as D_{lb} and maximum value as D_{ub}, the range of values for parameter δ(which equals $1/c$) to choose from is $0.5D_{lb} \sim 0.5(D_{lb} + D_{ub})$. The value is then determined from this range based on test accuracy in cross validation.

Table S1: ROIs (defined by automated anatomical labeling (AAL) template) selected for further investigation on Schizophrenia classification.

| Region ID	Anatomical Region
3	Left Superior Frontal Gyrus
61	Left Inferior Parietal Lobule
35, 36	Bilateral Posterior Cingulate Gyrus
78	Right Thalamus
91 $\sim$ 108	Cerebellum
26	Right Medial Orbitogrontal Cortex
29, 30	Bilateral Insula
57	Left Postcentral Gyrus
67, 68	Bilateral Precuneus
81, 82	Bilateral Superior Temporal Gyrus
85, 86	Bilateral Middle Temporal Gyrus
89, 90	Bilateral Inferior Temporal Gyrus
41	Left Amygdala
47, 48	Bilateral Lingual Gyrus
73, 74	Bilateral Putamen
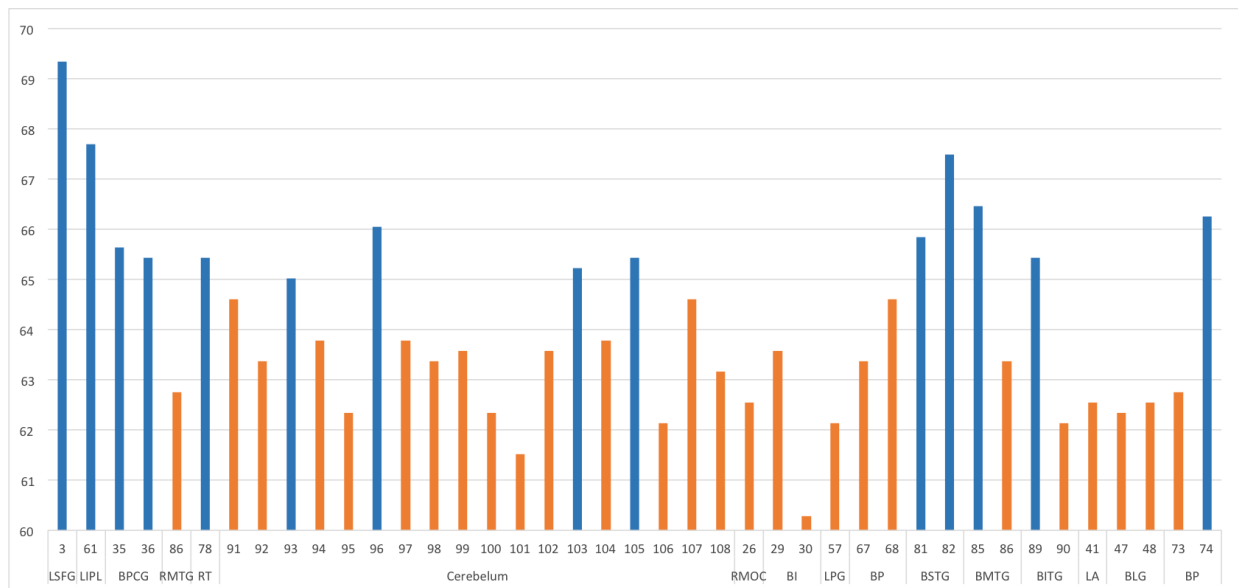
Figure S3: Classification accuracy of using single fMRI derived node data. X axis lists the abbreviation and Region ID of included ROIs (Table S1). Y axis is the classification accuracy with disconnected graphs constructed from fMRI data on single nodes. Blue ones are nodes that yield accuracy higher than 65%.