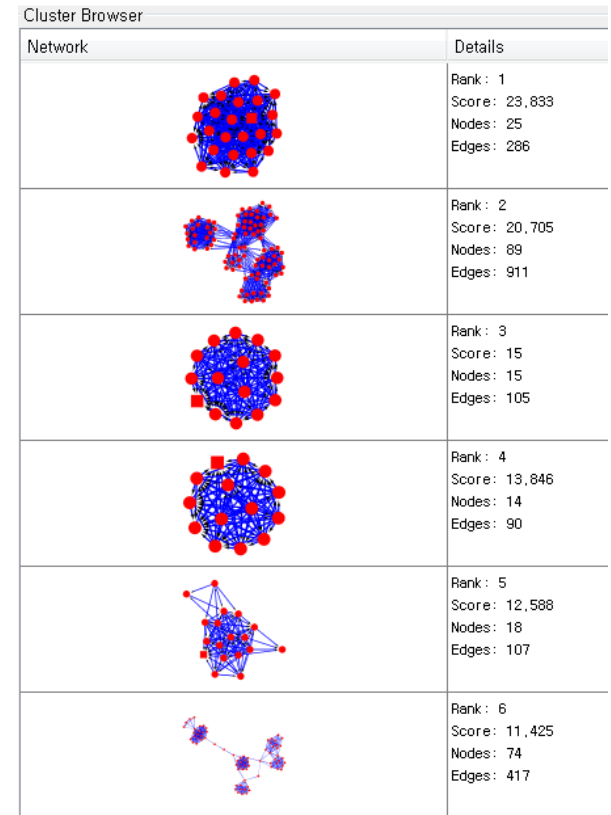


Supplementary Figures for Decipher
reliable biomarkers of brain aging through
integrating literature-based evidence with
interactome data

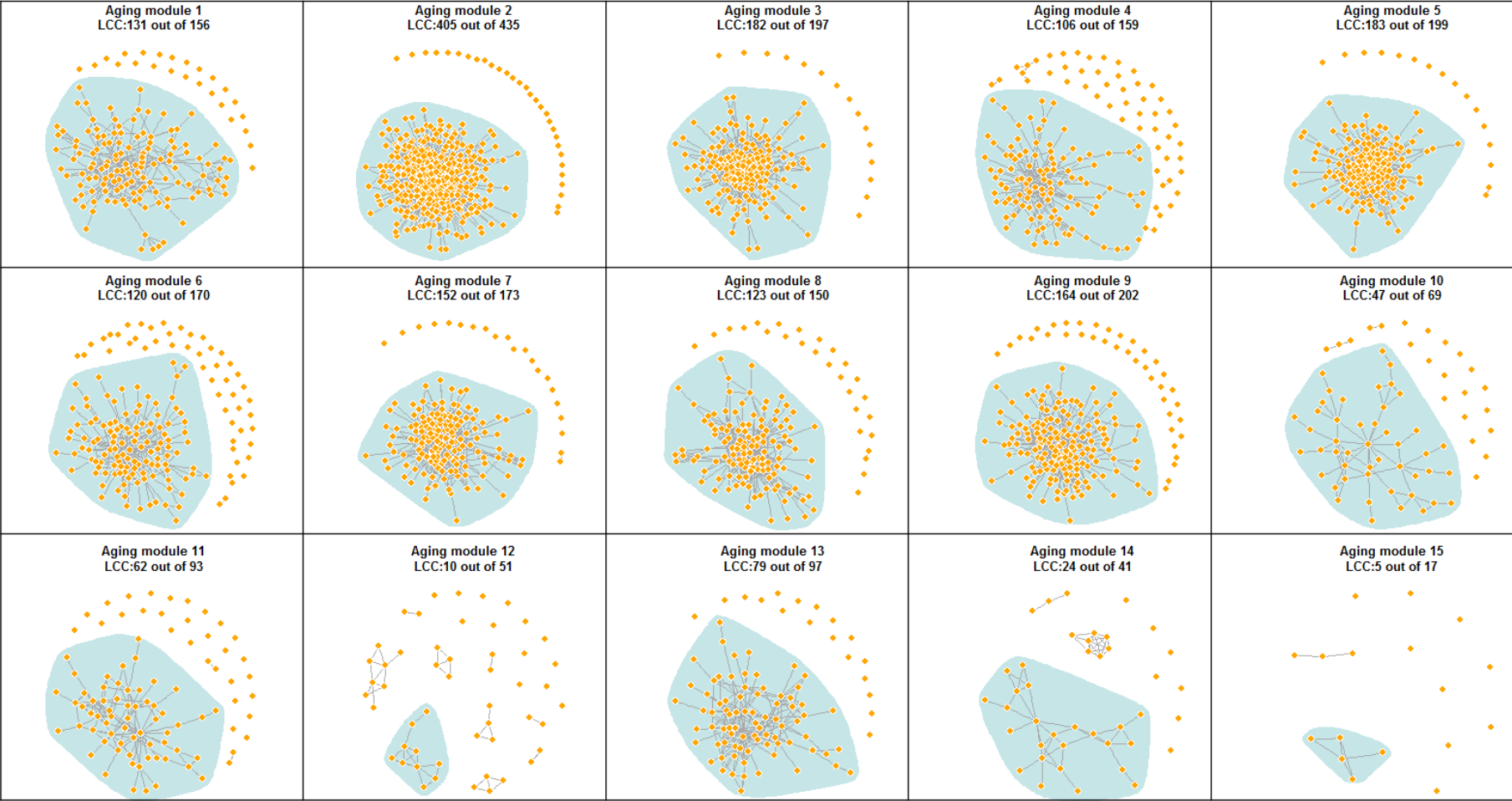
Modules rank



m code analysis

Supplementary Figure S2
(related to Figure 1)

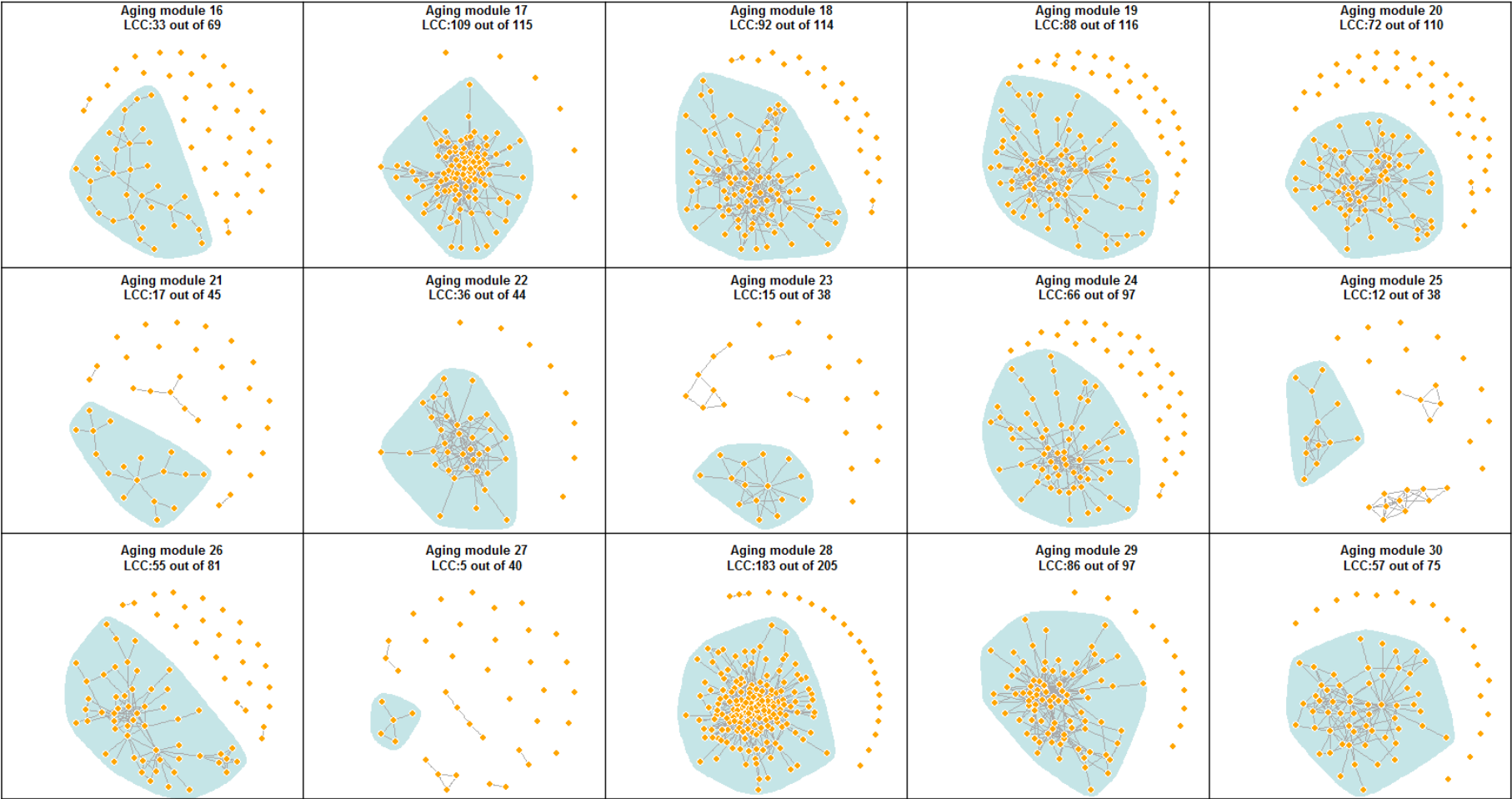
a



Supplementary Figure S2. (a) The largest connected component (LCC) of aging module 1-15. In the case of aging module 1, 131 genes (out of 156 genes) are connected among them, forming the LCC(colored regions).

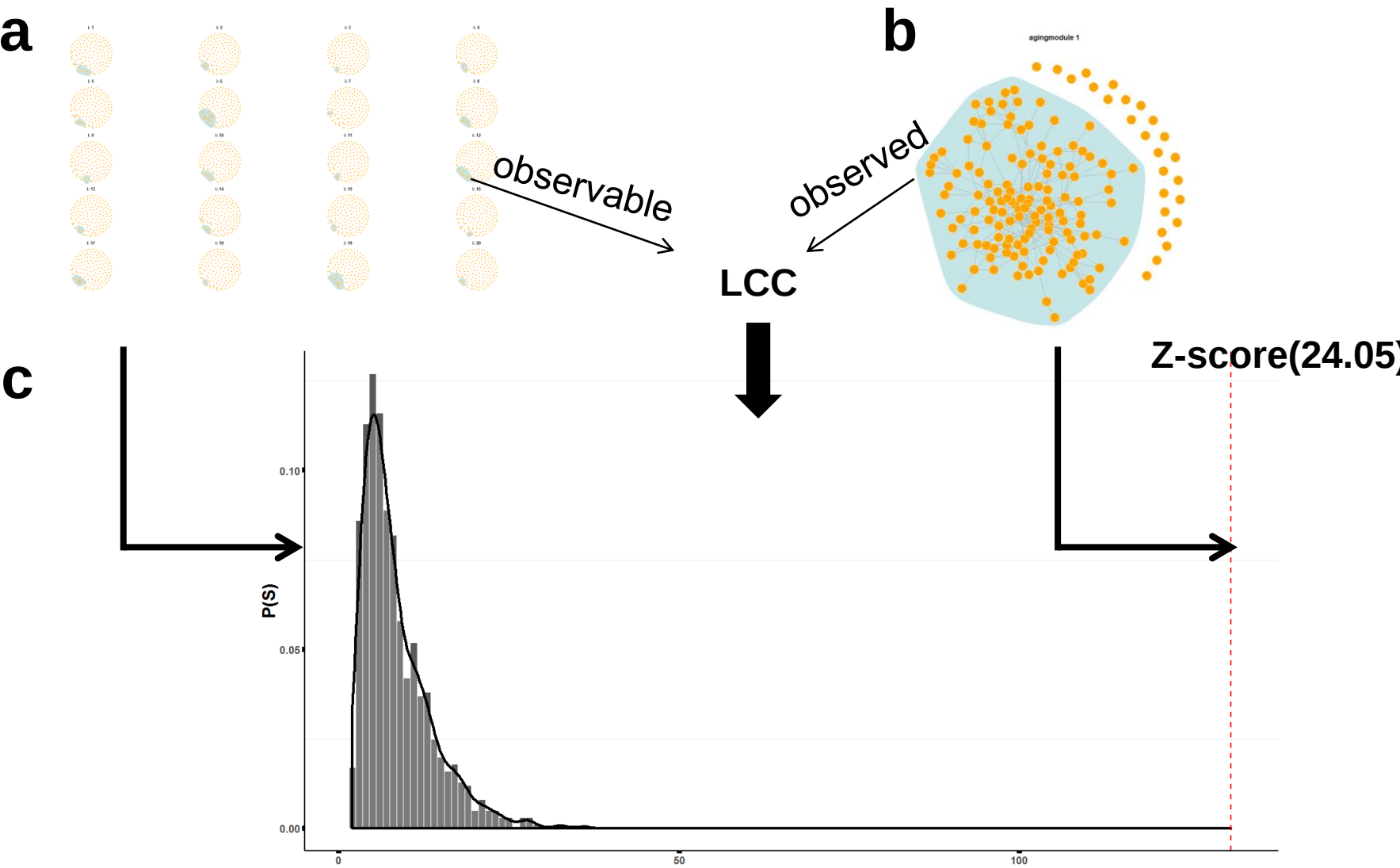
Supplementary Figure S2
(related to Figure 1)

b



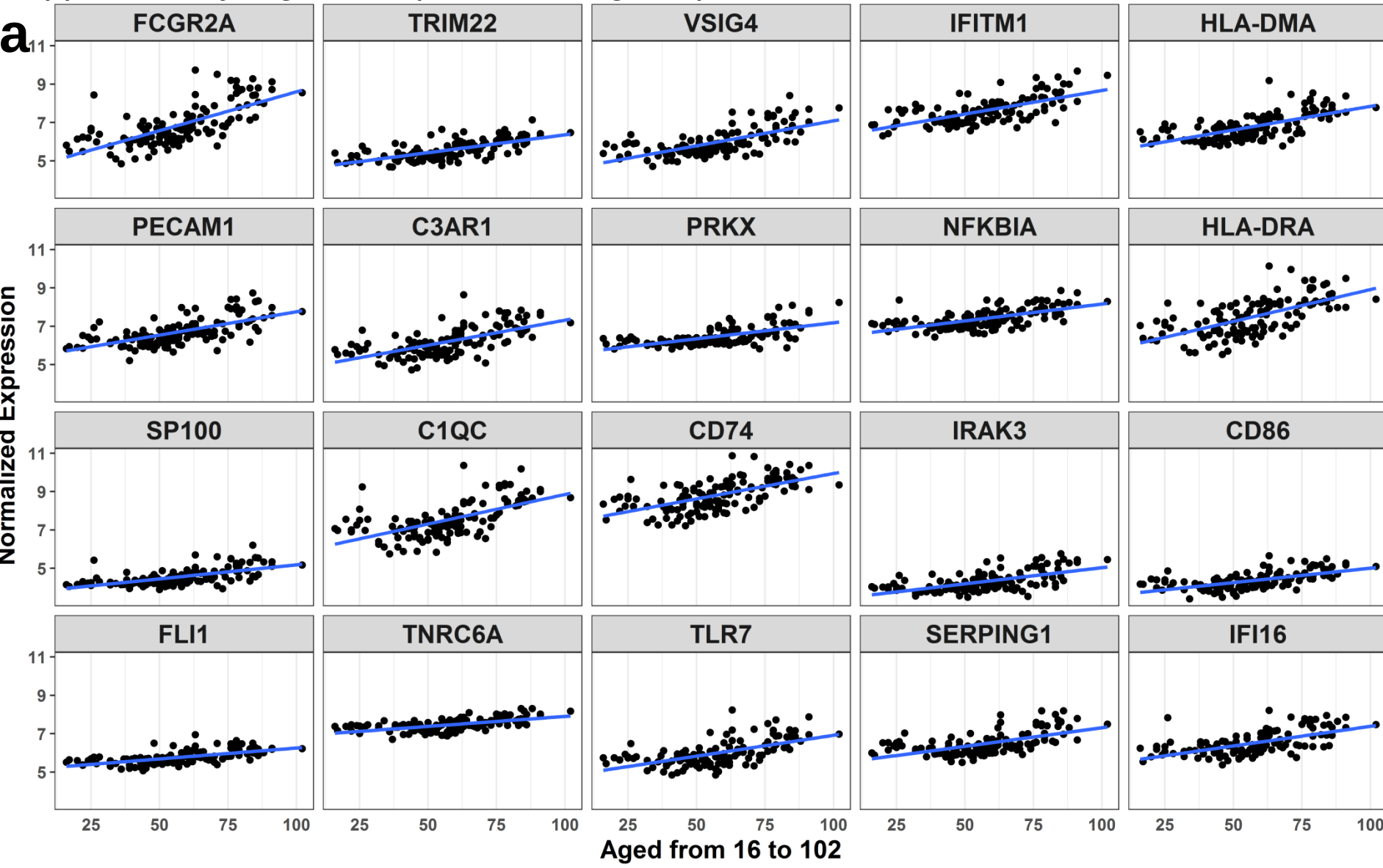
Supplementary Figure S2. (b) The largest connected component (LCC) of aging module 16-30. The colored regions are the LCC.

Supplementary Figure S3 (related to Figure1)



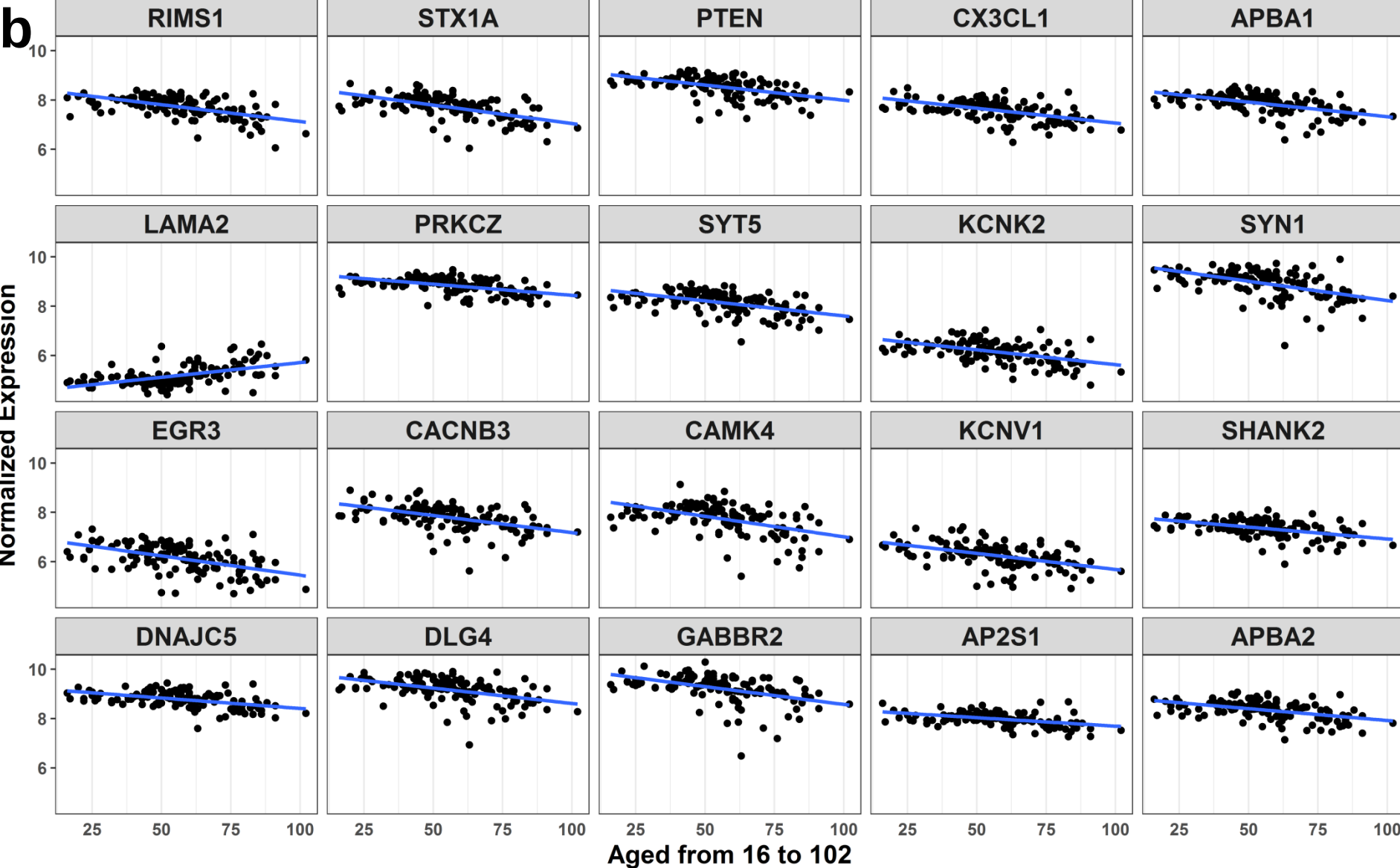
Supplementary Figure S3. Calculating significances of LCC of aging modules. (a) randomly generated module (observable). (b) Aging modules 1 (observed). (c) Density for the LCC numbers of the randomly generated modules. Red dotted line corresponds to the size of the LCC of aging module 1.

Supplementary Figure S5 (related to Figure3)



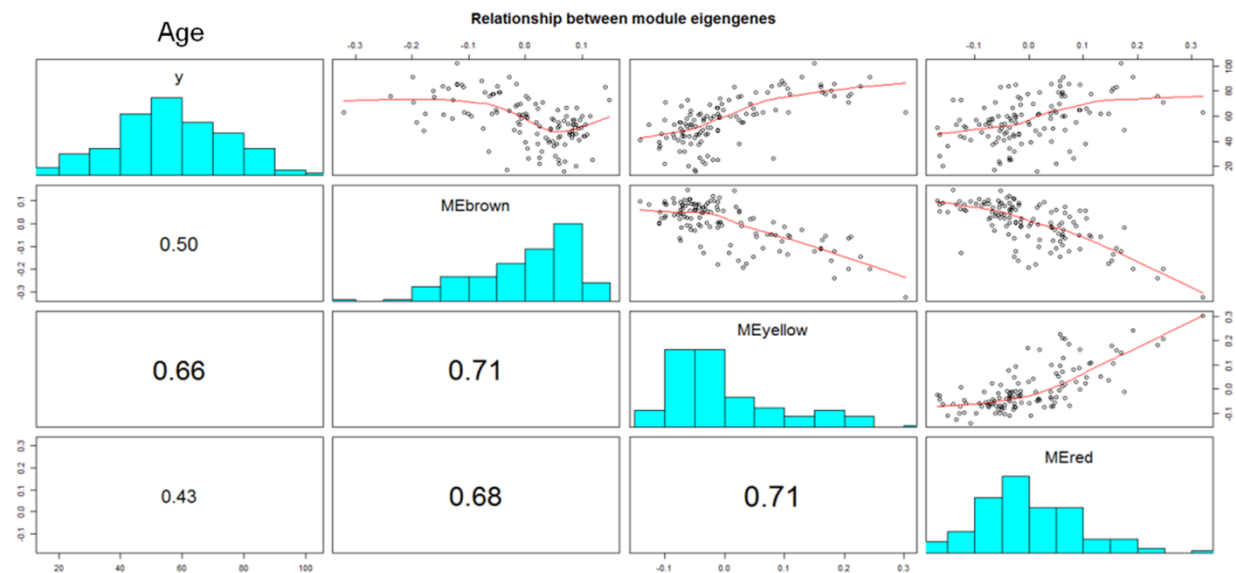
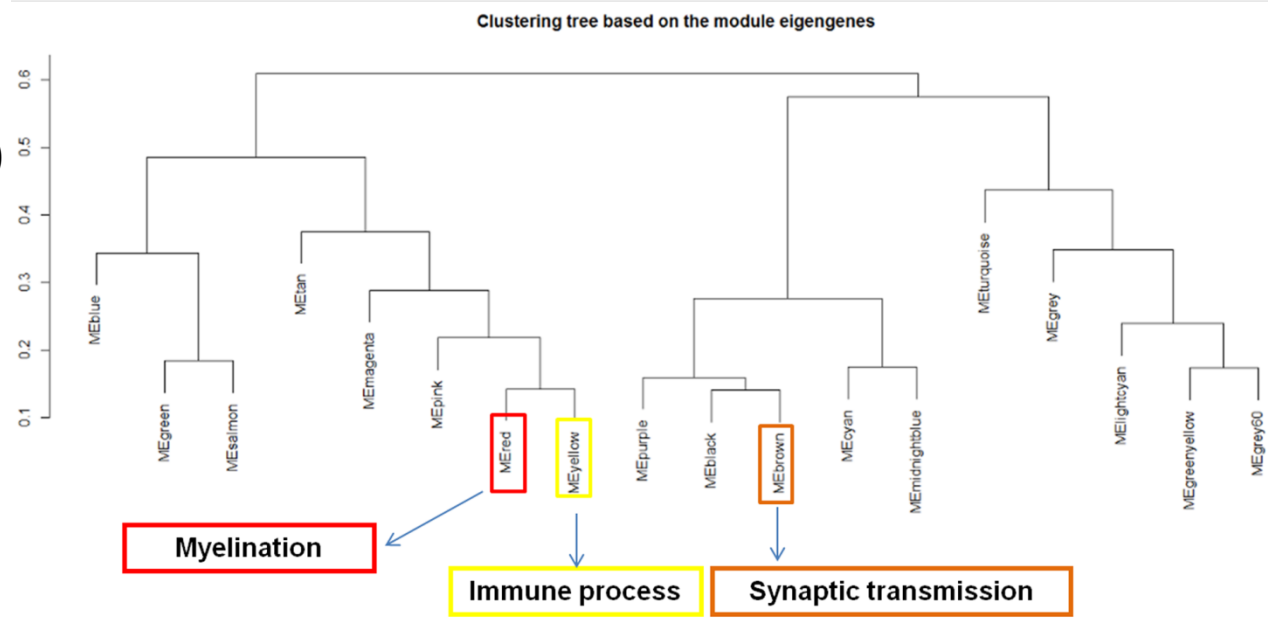
Supplementary Figure S5. Expression of individual genes of yellow and brown modules. (a) Expression of yellow module genes enriched in immune process (20 representative example). The genes of yellow module showed a tendency of an increase in expression with age.

Supplementary Figure S5 (related to Figure3)



(b) Expression of brown module genes enriched in synaptic transmission (20 representative example). The genes of brown module showed a tendency of a decrease in expression with age.

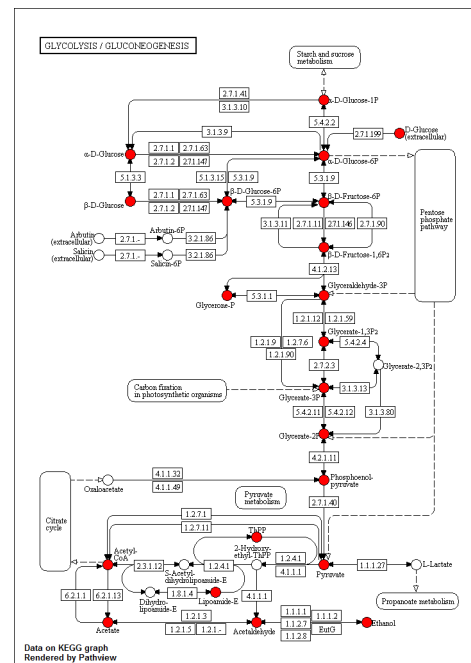
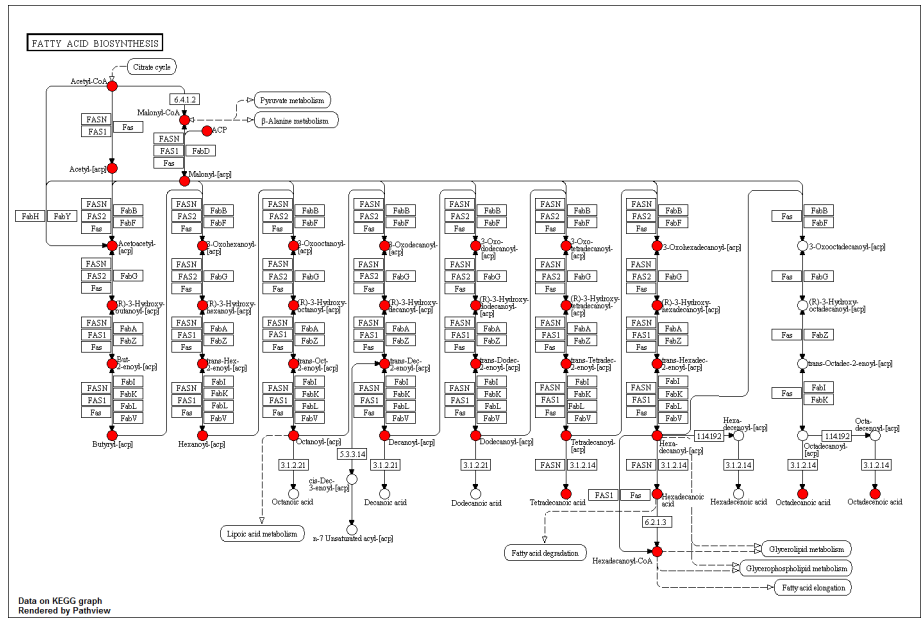
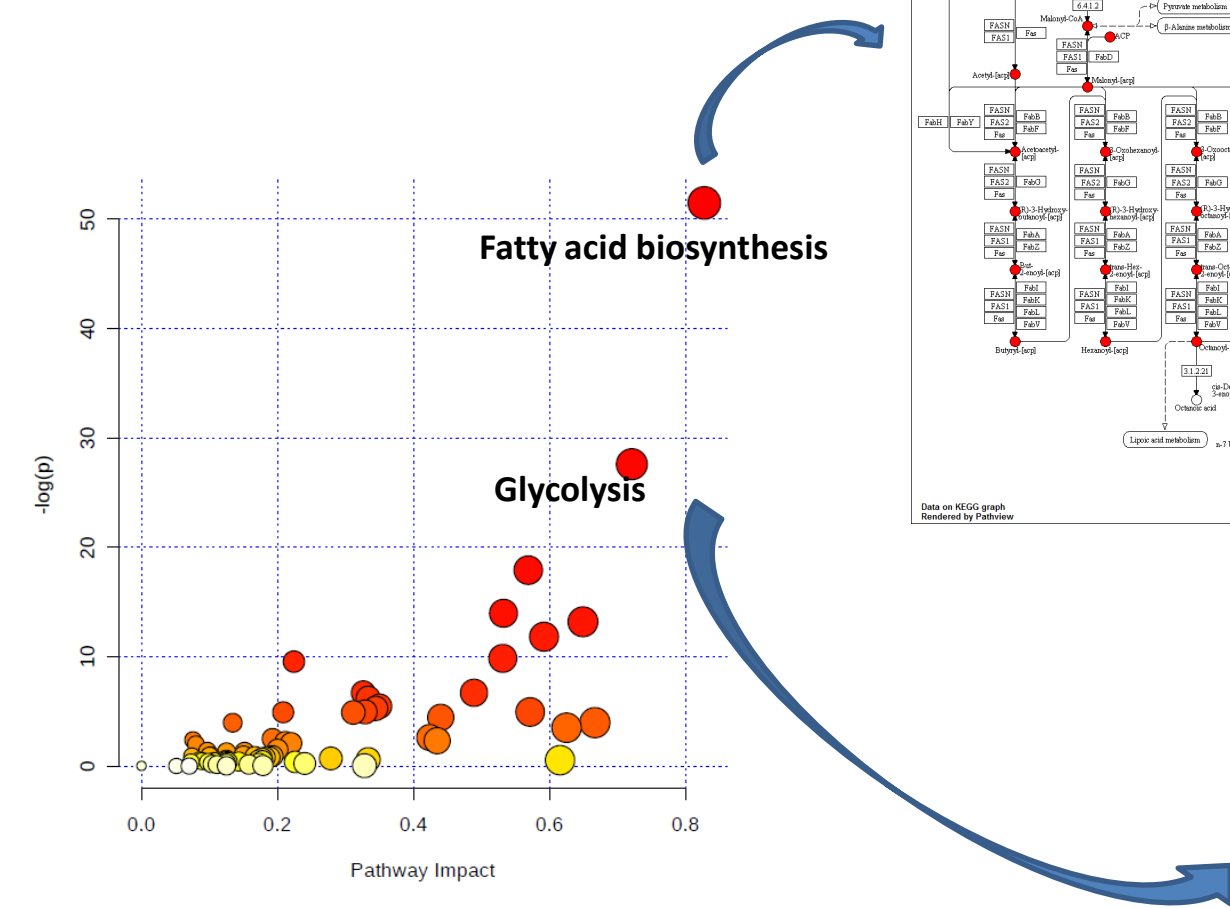
Supplementary
Figure S6
(related to Figure3)



Supplementary Figure S6. Clustering tree on module eigengenes(upper) and correlations between eigengenes and aging information (lower).

Supplementary Figure S8

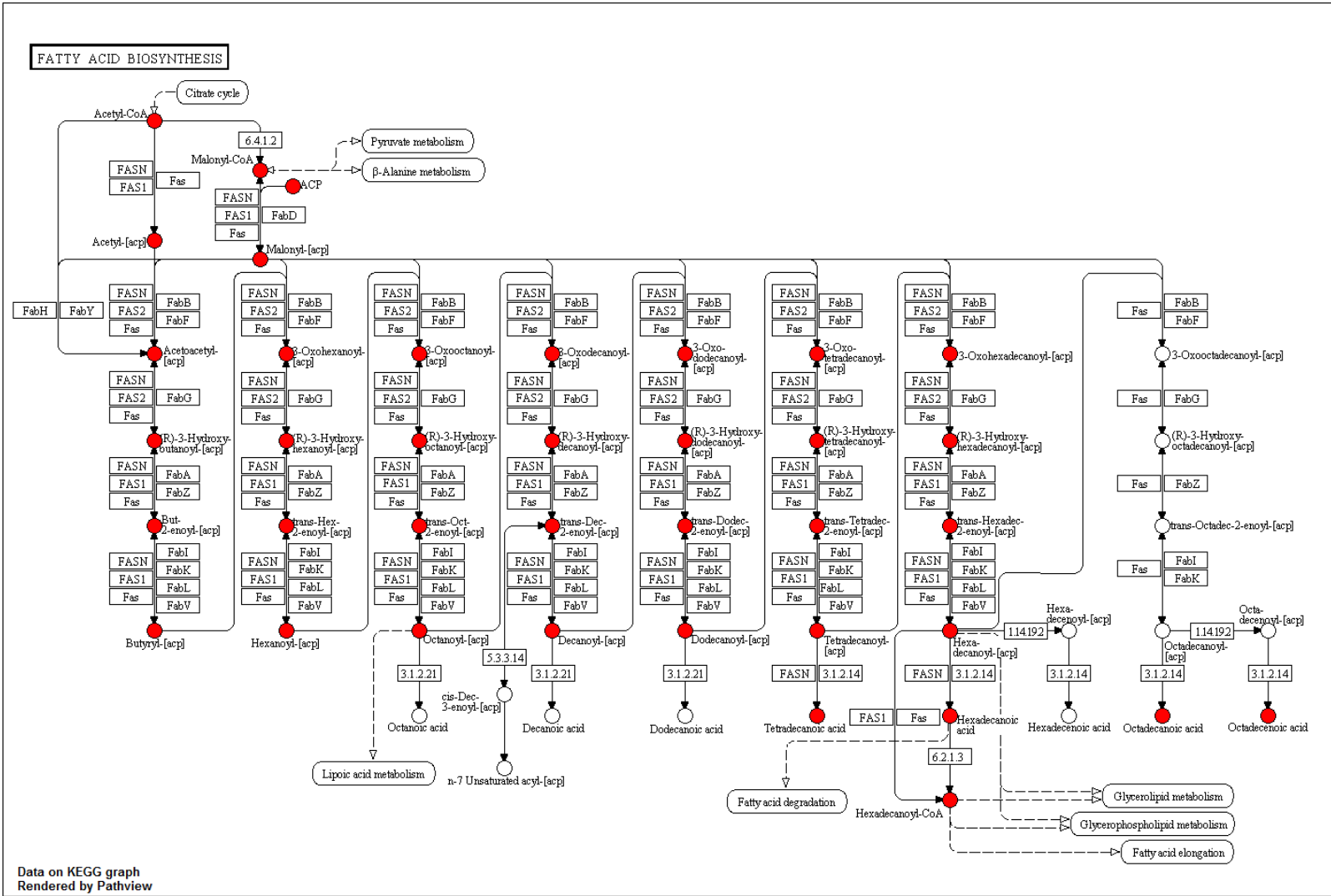
a



Supplementary Figure S8. Enrichment analysis of metabolites linked to genes of brown modules (associated with synaptic transmission). (a) Fatty acid biosynthesis and glycolysis were enriched. X axis corresponds to relative-betweenness centrality of the matched metabolites and Y axis to minus logarithm of raw p value.

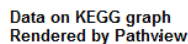
Supplementary Figure S8

b



Supplementary Figure S8 (b). Fatty acid biosynthesis. The matched metabolites or compounds were highlighted as red circles.

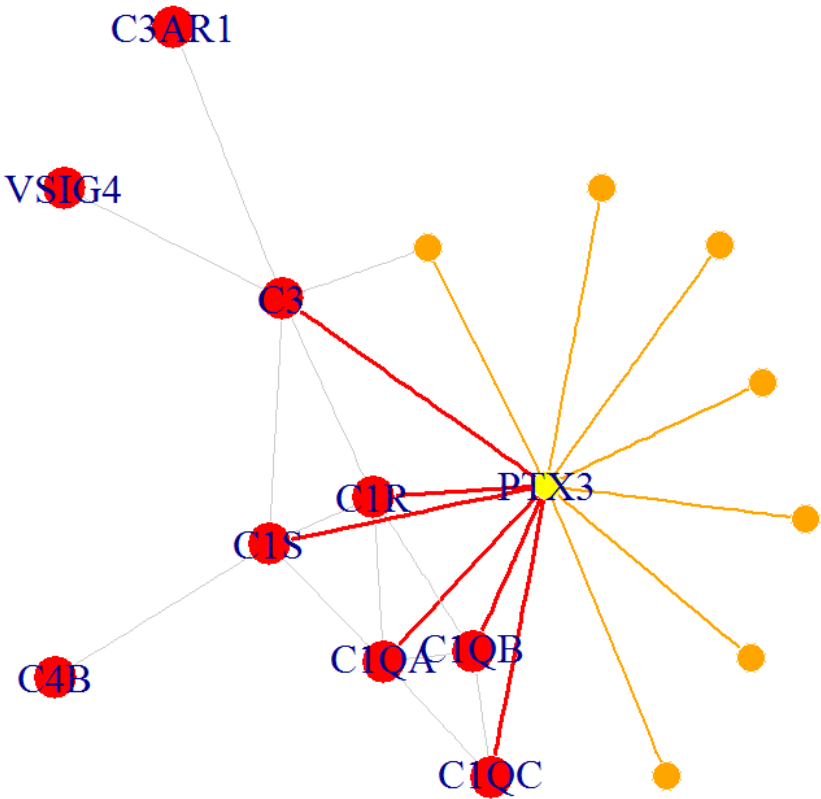
C



Supplementary Figure S8.
(c) Glycolysis. The matched metabolites or compounds were highlighted as red circles.

Supplementary Figure S9 (related to Figure 6b)

a



Ks: Ks links to seed genes(red edges)

S0:the number of seed genes (red nodes)

K:Degree of a node (PTX3) of interest; red edges +orange edges

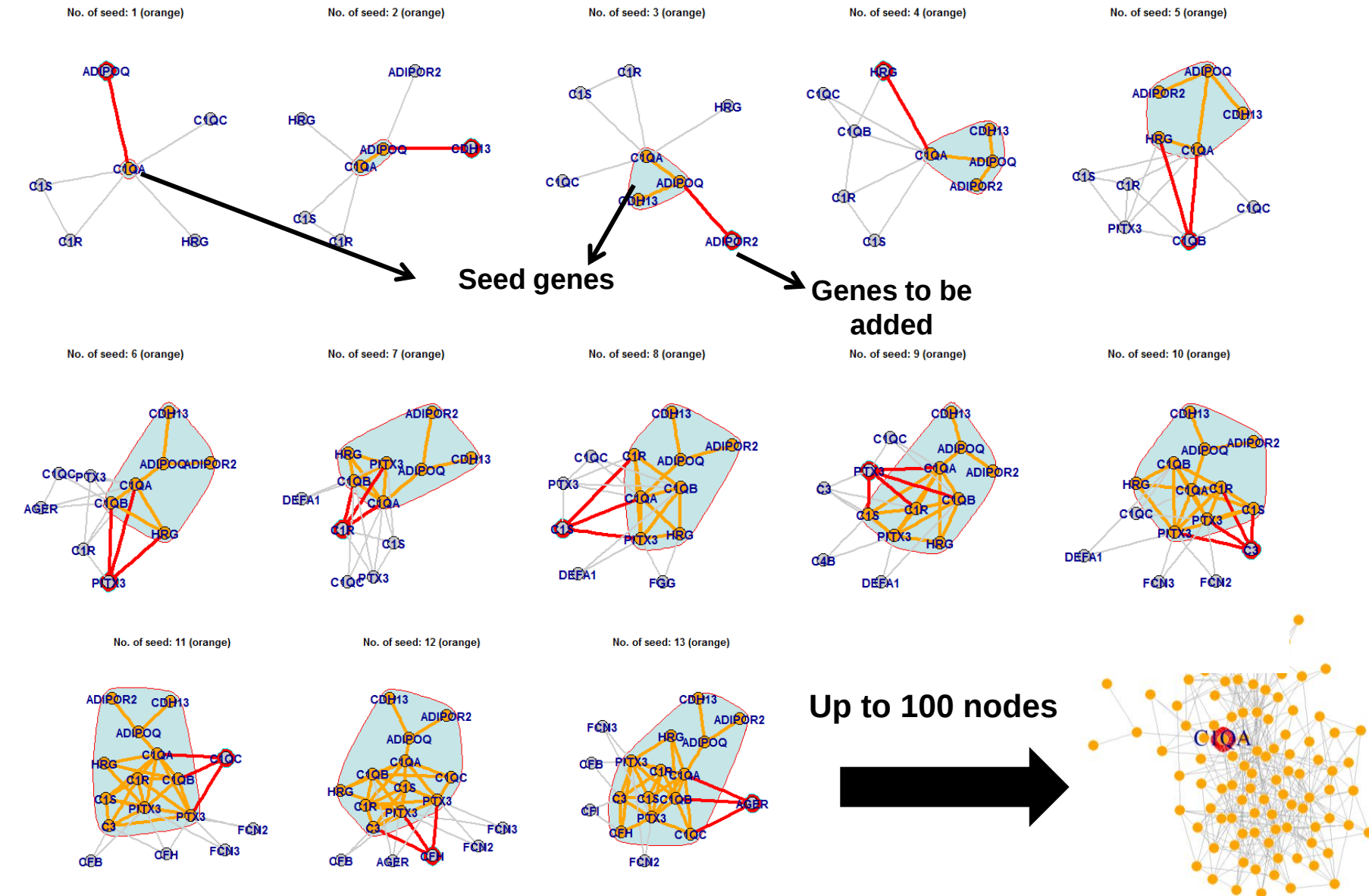
N:Total nodes

$$p(k, k_s, k_{s_0}) = \frac{\binom{s_0}{k_s} \binom{N - s_0}{k - k_s}}{\binom{N}{k}}$$

Supplementary Figure S9. Connectivity significance. (a) Calculating connectivity significance with hypergeometric distribution. Significance of links from PTX 3 (a yellow node) to seed genes (red nodes) can be calculated as the equation.

Supplementary Figure S9 (related to Figure 6b)

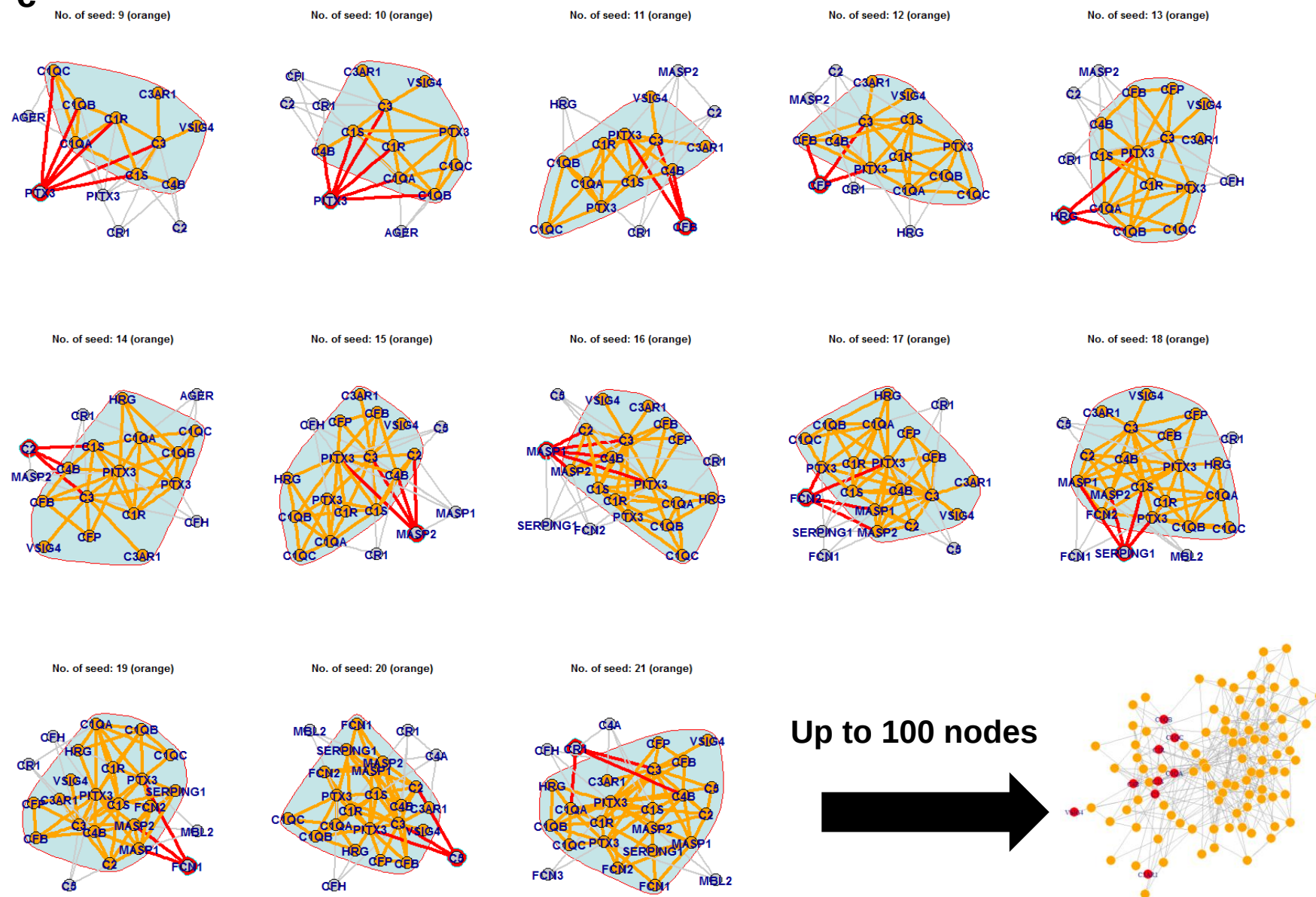
b



Supplementary Figure S9. (b) A module extended from an individual gene of the seed genes (A example of starting from C1QA of CCCh1).

Supplementary Figure S9 (related to Figure 6b)

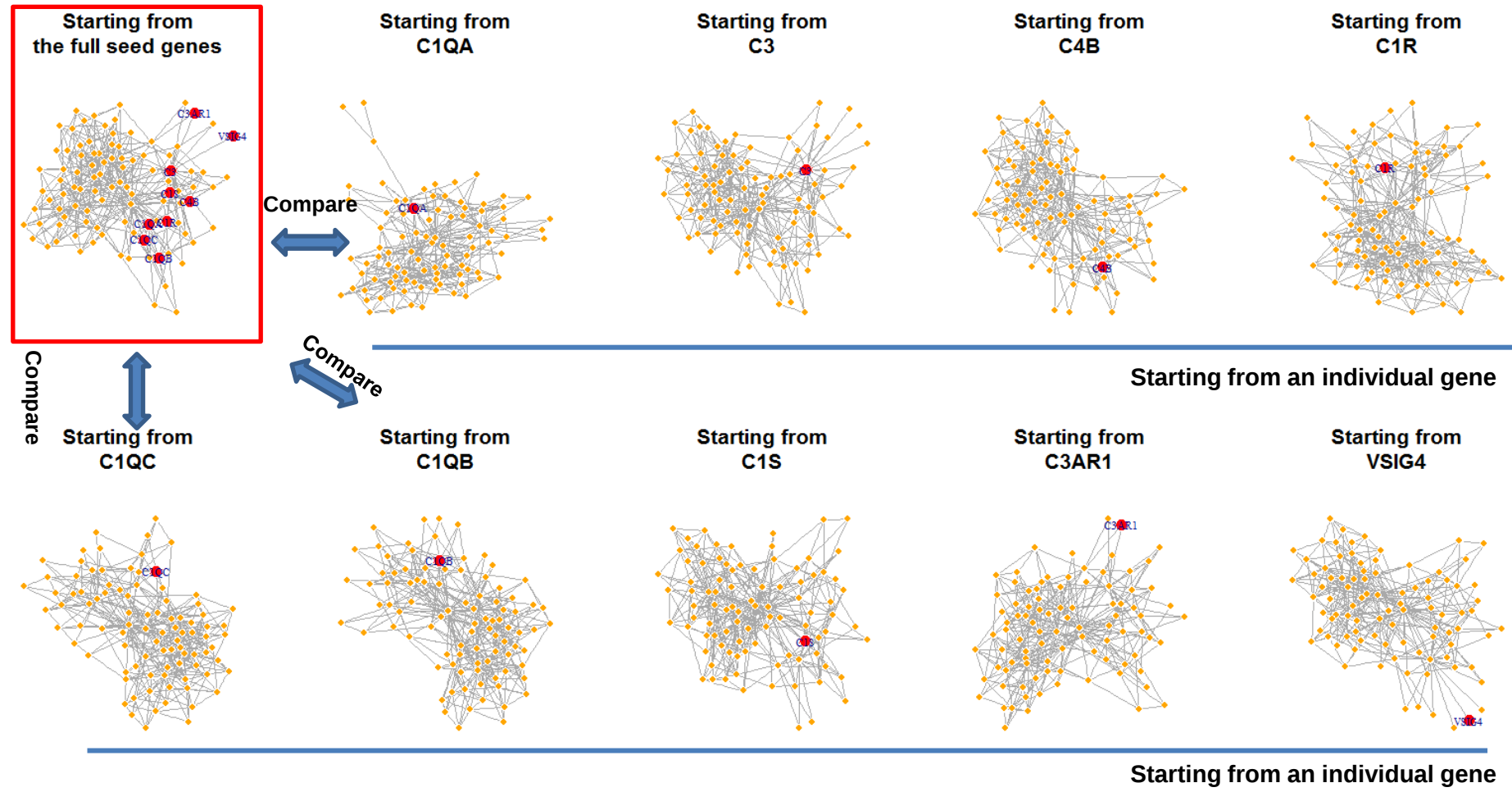
C



Supplementary Figure S9. (c) A module extended from full seed genes (A example of starting from CCCh1). Red edges correspond to the links from the added node to the seed nodes(colored region).

Supplementary Figure S9 (related to Figure 6b)

d



Supplementary Figure S9. (d) Overlap between the module starting from the full seed (c) genes and the module starting from an individual gene (b) from seed genes.

Supplementary Figure S10 (related to Figure 6b)

a CCCh1

Removed gene	Iteration needed for recovery
C1QA	2
C3	2
C4B	8
C1R	2
C1QC	12
C1QB	2
C1S	2
C3AR1	29
VSIG4	greater than 100

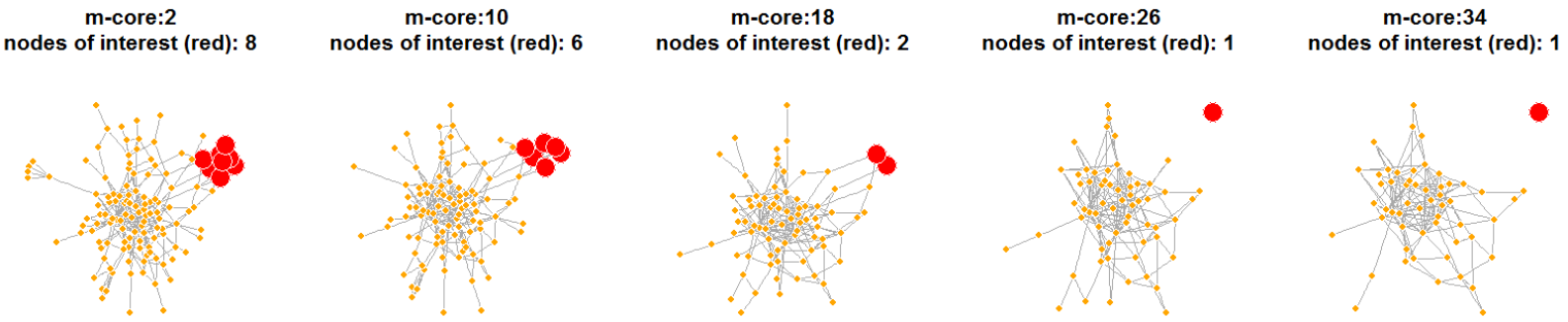
b CCCh2

Removed gene	Iteration needed for recovery
NFKBIA	greater than 100
BCL6	39
RELA	15
IRF8	32
IL18	greater than 100
TLR1	greater than 100
CASP1	greater than 100
TGFB1	greater than 100
A2M	greater than 100
MYD88	29
CCR1	greater than 100
TAP1	greater than 100
IL1B	11
CCL2	48
TLR5	greater than 100
CD14	greater than 100
TLR9	greater than 100

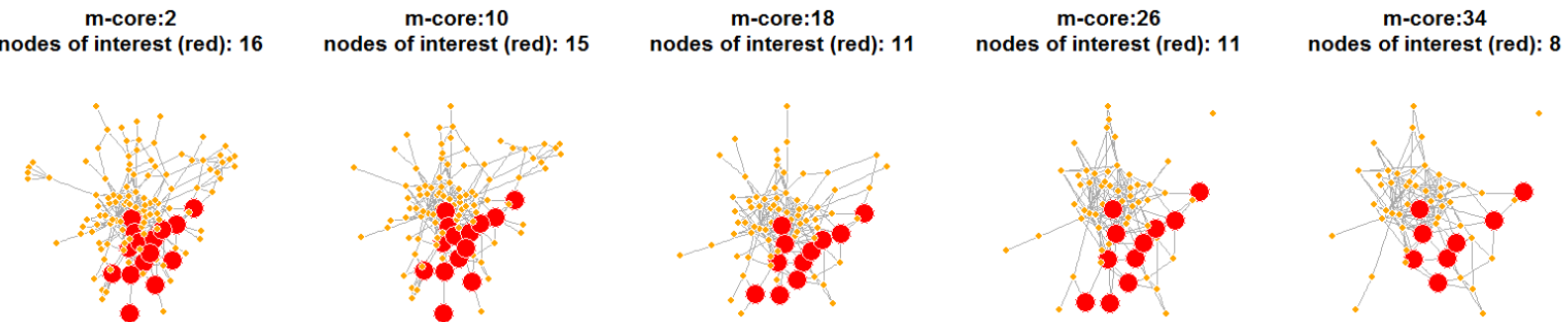
Supplement Figure S10. Measuring recovery after removing one gene.
Measure How many iterations it will take for the removed gene
to be added as a gene showing the highest significant connection to seed genes.
The iteration number needed to recover a removed gene was measured in CCCh1 (a) and in CCCh2 (b).

Supplementary Figure S11 (related to Figure 6c)

Nodes placed closer to the periphery of network

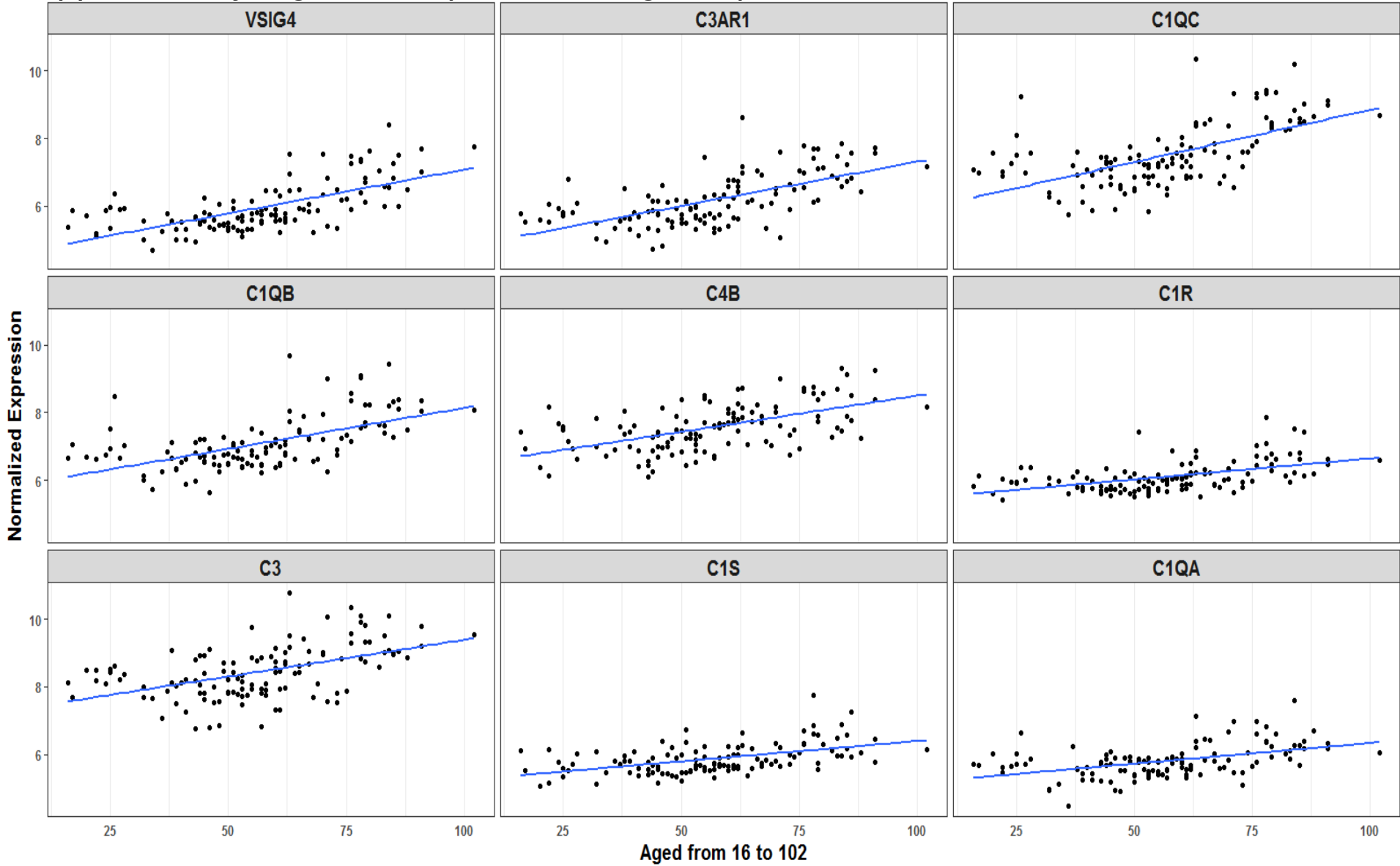


Nodes placed closer to the hub



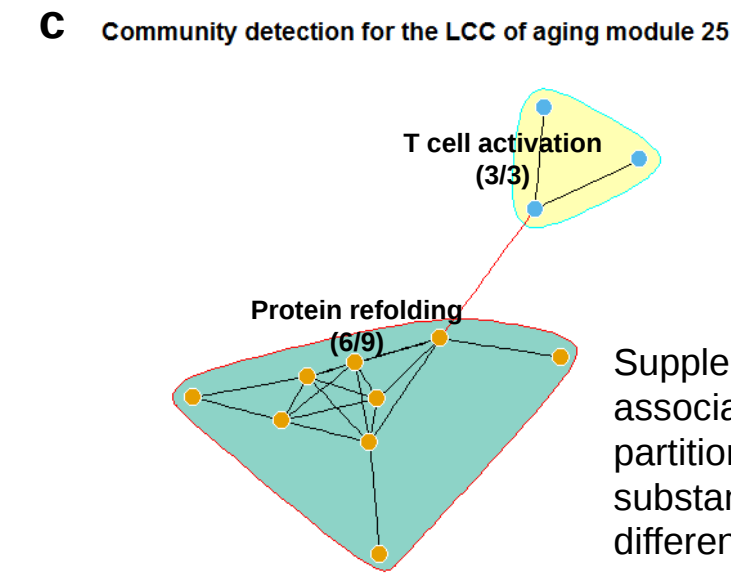
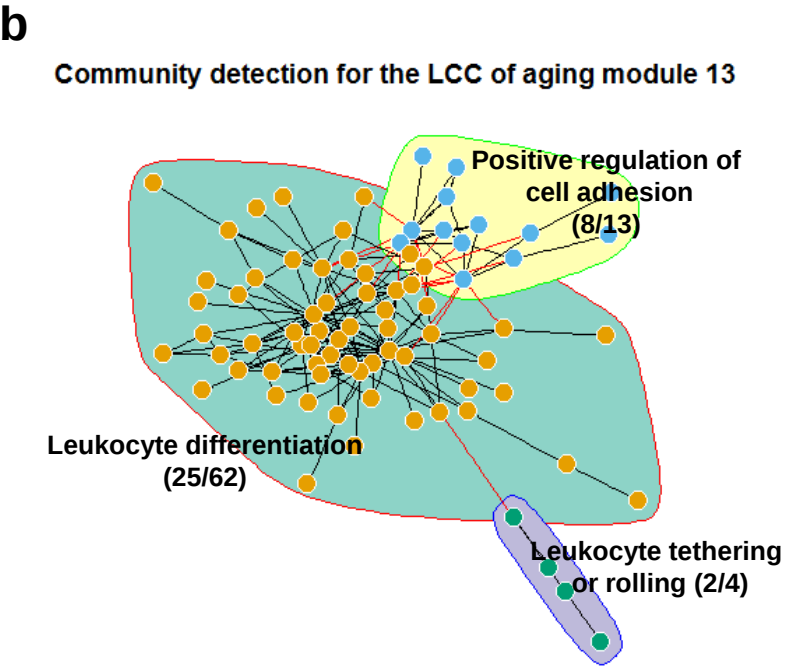
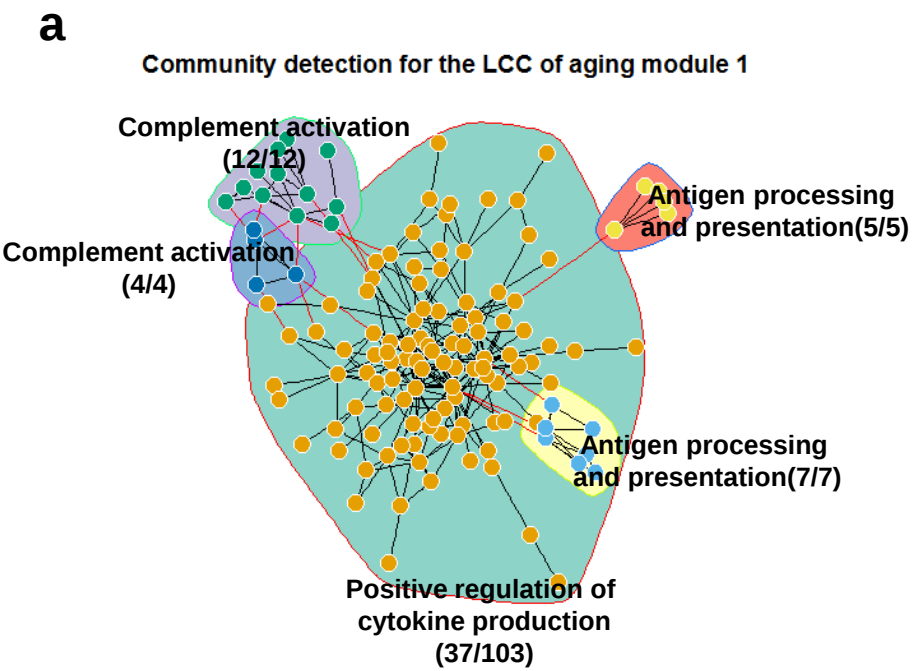
Supplement Figure S11. Process of Excess retention. The number of m-core is a cut-off degree. For example, m-core 2 means that the genes with less than 2 degree were removed. From left to right, genes with less than the m degree were removed in an iterative manner. Upper panel: nodes placed closer to the periphery are almost removed at the m-core of 18. Bottom panel: Over half of nodes placed closer to the hub remained even at the m-core of 18.

Supplementary Figure S12 (related to Figure 7)



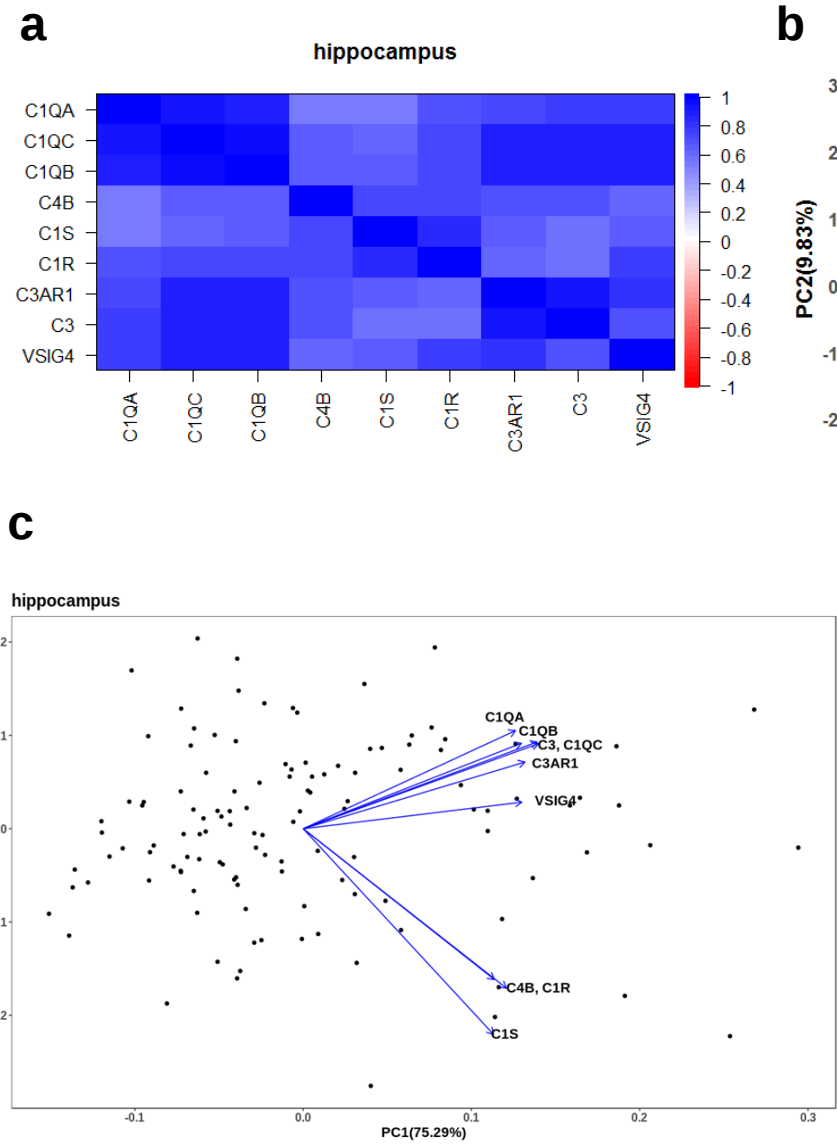
Supplement Figure S12. Expression of co-expressed genes involved in complement system (CCCh1). The 9 genes showed a time-dependent increase. X axis corresponds to age and Y axis to normalized expression of each gene.

Supplementary Figure S13



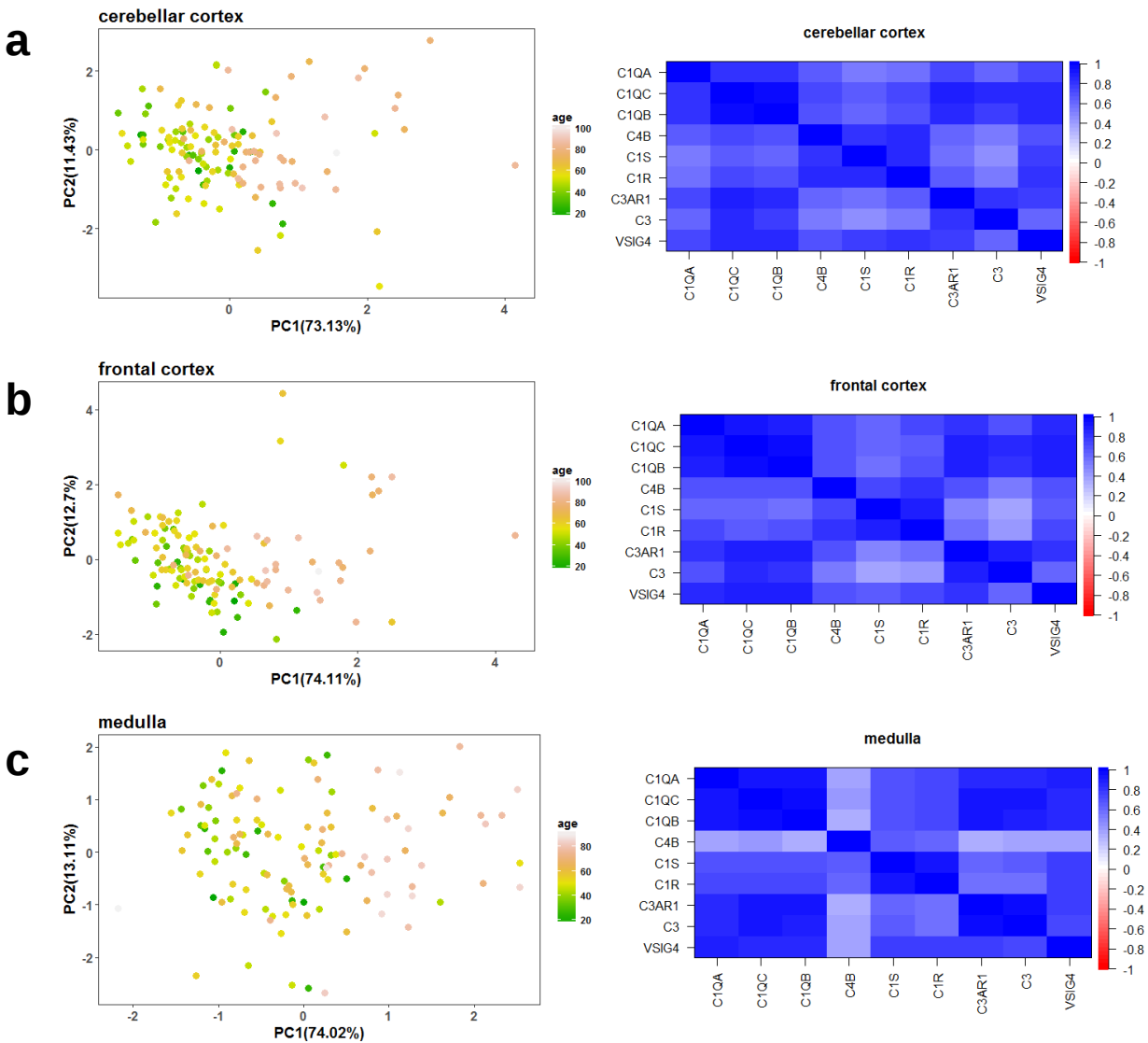
Supplement Figure S13. Three aging modules 1(a),13 (b) and 25 (c) are associated with immune system. However, with community detection to partition the vertices into communities, three modules are found to differ substantially from each other in terms of specific immune function. These differences may lead to different links to disease modules.

Supplementary Figure S14



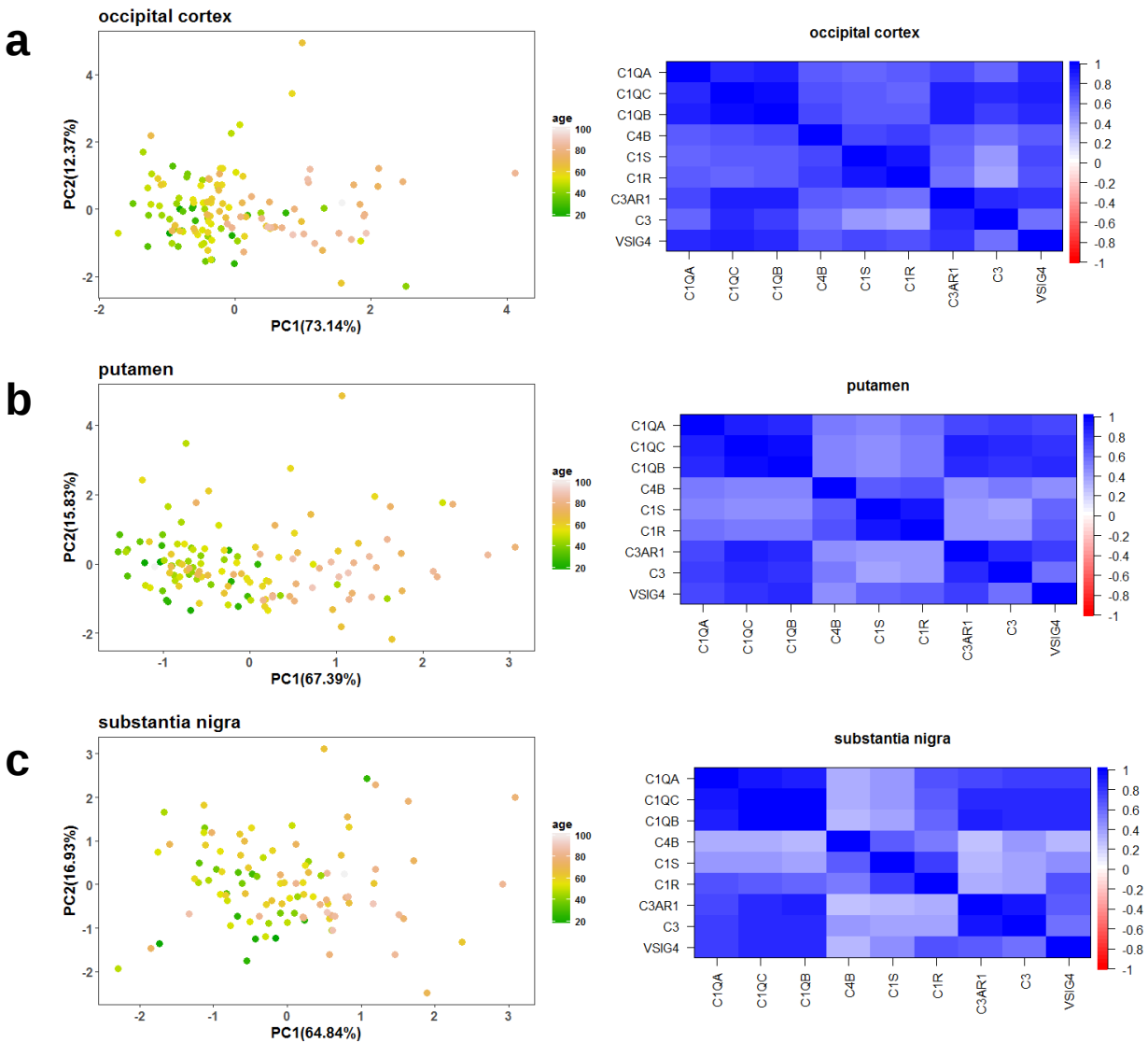
Supplement Figure S14. PCA (principal components analysis). Correlations between expressions of the 9 genes in the hippocampus were strong (a). Blue corresponds to positive correlation and red to negative correlation. All correlations are positive. Color intensity is proportional to the correlation coefficient. The PCA plot (b) showed that the older subjects tend to be located in the right side of the plot and the younger subjects in the left side of the plot. In the biplot (c), 9 genes (features or variable) are toward the positive direction of first principal components, indicating the 9 genes may tend to be up regulated with age. X (Y) value of each variable or feature (gene) in the biplot corresponds to the relative contribution of the original variable to the first (second) principal component.

Supplementary Figure S15



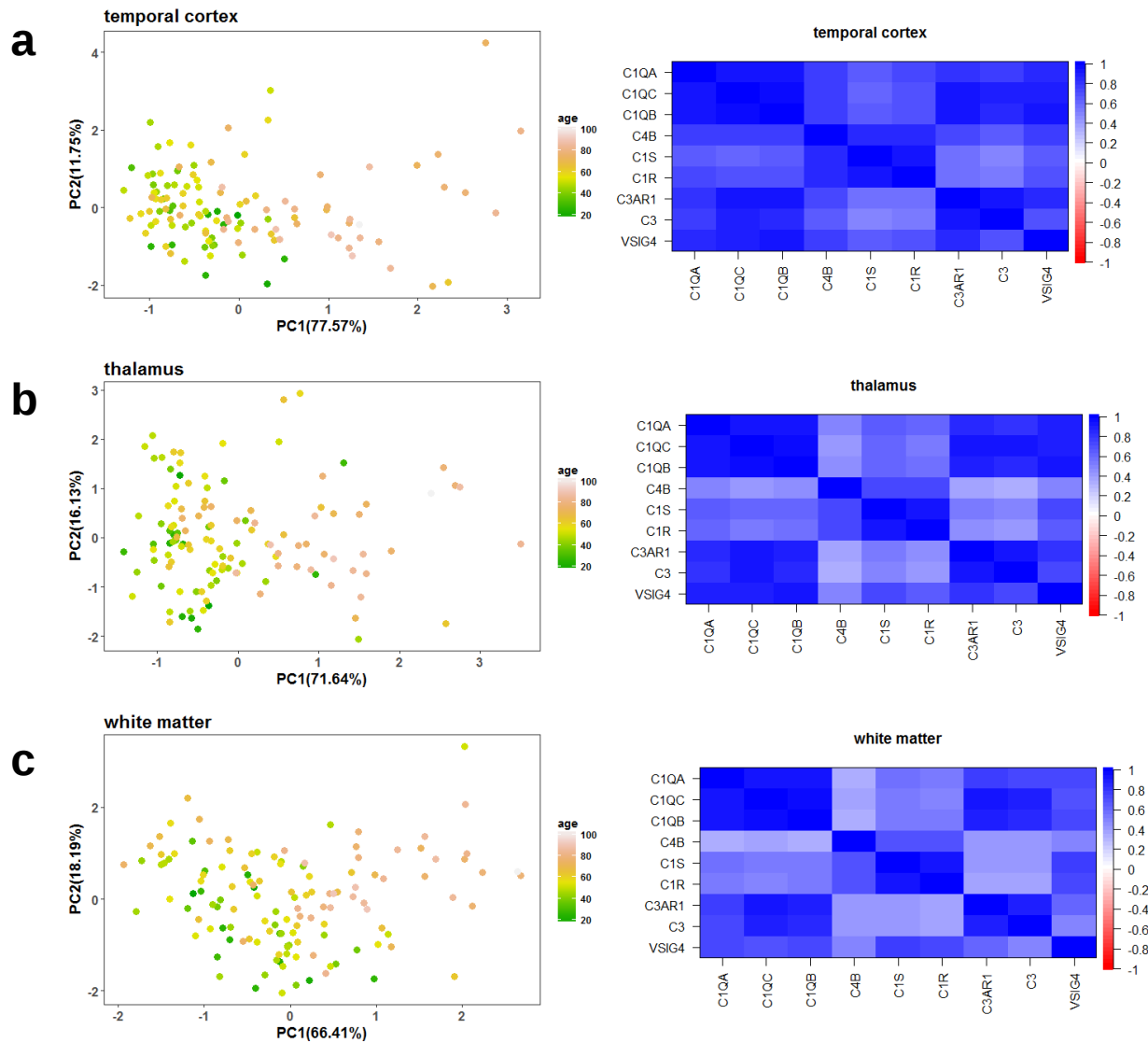
Supplement Figure S15. PCA analysis and correlation plot with the 9 genes' expressions as features in the cerebellar cortex (a),frontal cortex (b), medulla (c).

Supplementary Figure S16



Supplement Figure S16. PCA analysis and correlation plot with the 9 genes' expressions as features in the occipital cortex (a), putamen (b), substantia nigra (c).

Supplementary Figure S17



Supplement Figure S17. PCA analysis and correlation plot with the 9 genes' expressions as features in the temporal cortex (a),thalamus (b), white matter (c).