Co-expression network analysis identified novel potential Signature Genes Associated with human left ventricle cardiomyopathies arises from different etiologies

Jiao Tian[1,2,*], Zhengyuan Wu[1,*], Yingying He[3,#], Shubai Liu[1,#]

[1.] State Key Laboratory of Phytochemistry and Plant Resources in West China, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, 650201 Yunnan, P. R. China;

[2.] School of Life Sciences, Yunnan University, Kunming, Yunnan 650091, China;

[3.] School of Chemical Science & Technology, Yunnan University, Kunming, Yunnan 650091, China;

* Contributed equally.

Correspondence to:

Dr. Shubai Liu, E-mail: liushubai@mail.kib.ac.cn. or

Dr. Yingying He, E-mail: yingying.he10@gmail.com

Supplementary Information

Additional file 1: Figure.S1. Sample dendrogram and Clinic Feature traits heatmap. Clustering dendrogram of samples based on their Euclidean distance. The clinical feature traits were histopathology, gender and age. The white color means a low value, red means a high value, and grey represents a missing entry.

Additional file 2: Figure.S2. Analysis of network topology for various soft-thresholding powers. In panel (A), the scale-free topology model fit index (signed R^2, y-axis) shows as a function of the soft-thresholding power (x-axis). In panel (B), the mean connectivity (ki, y-axis) displays as a function of the

soft-thresholding power (x-axis) under different weighting coefficients. The connectivity ki of node i equals the number of its direct connections to other nodes. P(k) indicates the frequency distribution of the connectivity. The higher the coefficient, the closer the network is to the distribution of the scale free network.

Additional file 3: **Figure.S3.** The scatterplots of Gene Significance (GS) for different Pathological trait vs. Module Membership (MM) in the all subtypes cases. There are seven subtype groups, including ID (A), IS (B), IDCM (C), FCM (D), PCM (E), ISCM (F), VCM (G). In each case group, there are several modules highly significant correlation between GS and MM. The higher Cor value, the module tighter correlated with clinical feature.

Additional file 4: **Figure.S4.** Clustering dendrogram of the eigengenes and adjacency with different cardiomyopathies trait dissimilarity based on topological overlap. There are seven subtype groups, including ID (A), IS (B), IDCM (C), FCM (D), PCM (E), ISCM (F), VCM (G). In each group, the panel (A) shows a hierarchical clustering dendrogram of the eigengenes in which the dissimilarity of eigengenes (EI, EJ is given by 1 – cor (EI, EJ). The heatmap in panel (B) shows the eigengene adjacency (AI J = (1 + cor (EI, EJ))/2).

Additional file 5: **Figure.S5.** Enrichment analysis & network blast of module significance genes in the different case groups. The enrichment analysis contains biological function, cellular components and molecular functions categories (A panel). The significance genes blast for literature network analysis for different subtypes cardiomyopathies. There are six subtype groups, including IS (A), IDCM (B), FCM (C), PCM (D), ISCM (E), VCM (F).

Additional file 6: **Figure.S6.** Module significance genes functions cluster in the different case groups. There are six subtype groups, including IS (A), IDCM (B), FCM (C), PCM (D), ISCM (E), VCM (F).

Additional file 7: **Table.S1.** The samples dataset information and case groups sorting.

All the used samples clustering to detect outlies.

Additional file 8: **Table.S2.** The pick soft threshold for Module.

Additional file 9: **Table.S3.** List of the significance genes in the most significant module of each case group. There are six subtype groups, including IS (A), IDCM (B), FCM (C), PCM (D), ISCM (E), VCM (F).

Additional file 10: **Table.S4.** Enrichment analysis of module Significance Genes for each case group. There are containing six subtype groups, including IS, IDCM, FCM, PCM, ISCM, VCM.