Semi-supervised learning for potential human microRNA-disease associations inference

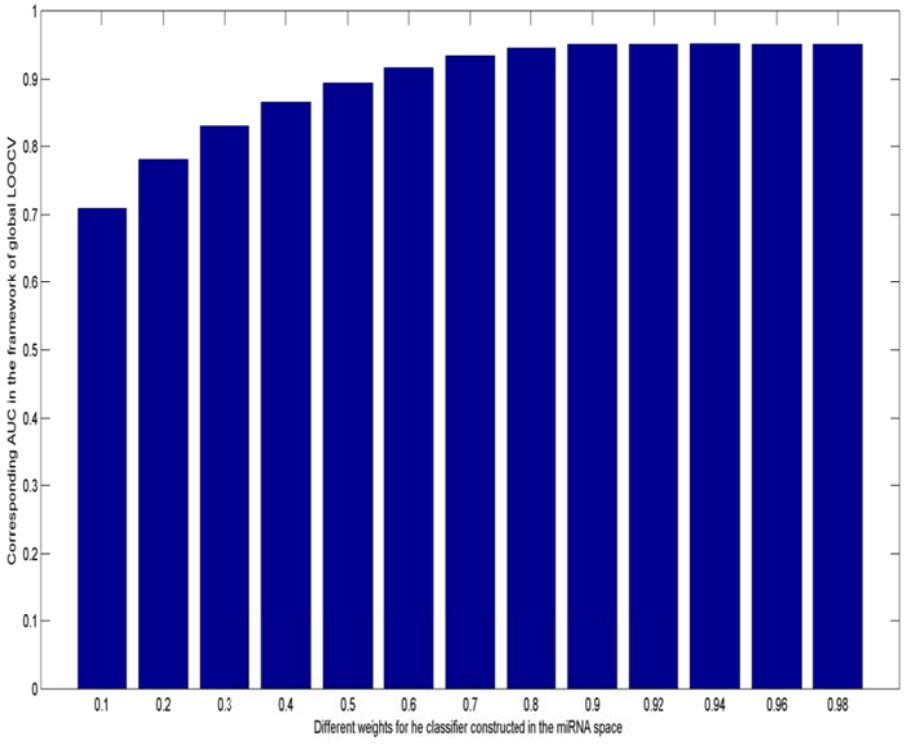
Xing Chen[1, 2, *] and Gui-Ying Yan [1, 2, *]

[1]National Center for Mathematics and Interdisciplinary Sciences, Chinese Academy of Sciences, Beijing, 100190, China

[2]Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, 100190, China

*Corresponding authors

Supplementary Information



Supplemental Figure 1. Different weights have been assigned to the classifier constructed in the miRNA space and corresponding AUCs have been shown. It could

be observed that a higher weight can improve the final performance of RLSMDA.

Supplemental Table 1. The top 50 potential colonic cancer related miRNAs and confirmation evidences for the associations were listed.

Supplemental Table 2. The top 100 potential disease-miRNA associations were shown and verified based on various databases and literatures.

Supplemental Table 3. The top 50 potential HCC related miRNAs when the information about known HCC related miRNAs are removed and evidences for the associations with HCC were listed

Supplemental Table 4. The top 50 potential colon cancer related miRNAs when the information about known colon cancer related miRNAs are removed and evidences for the associations with colon cancer were listed.

Supplemental Table 5. RLSMDA is further applied to predict potential human disease-miRNAs associations after confirming the reliable performance of RLSMDA in the term of cross validation and case studies. All the known disease-miRNA associations in the gold-standard dataset were used as positive samples. We publicly released potential human disease-miRNA association list to facilitate the biological experimental validation.

Supplemental Table 6. Known human miRNA-disease association dataset.

Supplemental Table 7. miRNA functional similarity scores used in this paper.

Supplemental Table 8. Disease semantic similarity used in this paper.