Supporting Information S3. The process of how to select selecting the key classifiers from the elementary classifiers

Suppose the distance between any two elementary classifiers $\mathbb{C}^0(i)$ and $\mathbb{C}^0(j)$ is measured by the following equation (Liu, et al., 2017):

$$\text{Distance}\left(\mathbb{C}^0(i), \mathbb{C}^0(j)\right) = 1 - \frac{1}{2m}\sum_{k=1}^{m}\left(d_{ik}\Delta d_{jk}\right) \qquad \text{(S1)}$$

where m represents the number of training samples, d_{ik} represents the misclassification probability of classifier $\mathbb{C}^0(i)$ on the k-th sample, and $d_{ik}\Delta d_{jk}$ can be calculated by:

$$d_{ik}\Delta d_{jk} = \begin{cases} d_{ik} + d_{jk}, & \text{if } \mathbb{C}^0(i) \text{ and } \mathbb{C}^0(j) \text{ have the same results on the } k\text{th sample} \\ 0, & \text{otherwise} \end{cases} \qquad \text{(S2)}$$

Please note that Eq.S2 is different from Eq.**3** in (Liu, et al., 2017). By using this equation, the distance between any two elementary classifiers can be more accurately measured. The range of Distance($\mathbb{C}^0(i)$, $\mathbb{C}^0(j)$) is from 0 to 1, where 1 indicates the predictive results of two classifiers are completely complementary, and 0 means that their results are identical. Based on the distance between any two classifiers, all of the classifiers in each layer were clustered by the affinity propagation clustering algorithm (Frey and Dueck, 2007). The preference values for the first layer clustering and second layer clustering were 0.72 and 0.70, respectively. Finally, for each cluster, the classifier with the highest accuracy value was selected as the key individual classifier for further usage.

References

Frey, B.J. and Dueck, D. (2007) Clustering by passing messages between data points, *science*, **315**, 972-976.

Lin, C.*, et al.* (2014) LibD3C: Ensemble Classifiers with a Clustering and Dynamic Selection Strategy, *Neurocomputing*, **123**, 424-435.

Liu, B., Long, R. and Chou, K.-C. (2016) iDHS-EL: identifying DNase I hypersensitive sites by fusing three different modes of pseudo nucleotide composition into an ensemble learning framework, *Bioinformatics*, **32**, 2411-2418.

Liu, B.*, et al.* (2017) iRSpot-EL: identify recombination spots with an ensemble learning approach, *Bioinformatics* **33**, 35-41.

Liu, B., Yang, F. and Chou, K.-C. (2017) 2L-piRNA: A Two-Layer Ensemble Classifier for Identifying Piwi-Interacting RNAs and Their Function, *Molecular Therapy-Nucleic Acids*, **7**, 267-277.