FeCO3, constructing the network biomarkers using the inter-feature correlation coefficients

Shenggeng Lin1,2,*, Yuqi Lin1,*, Kexin Wu1,*, Yueying Wang3, Zixuan Feng1, Meiyu Duan1, Shuai Liu1, Jingxuan Huang1, Zhao Wang1, Yusi Fan4, Lan Huang1, Fengfeng Zhou1,#.

1 College of Computer Science and Technology, and Key Laboratory of Symbolic Computation and Knowledge Engineering of Ministry of Education, Jilin University, Changchun, Jilin 130012, China.

2 State Key Laboratory of Microbial Metabolism, and School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, China.

3 Department of Epidemiology and Biostatistics, School of Public Health, Jilin University, Changchun, Jilin Province, China.

4 College of Software, and Key Laboratory of Symbolic Computation and Knowledge Engineering of Ministry of Education, Jilin University, Changchun, Jilin 130012, China.

* Contributed equally to this study.

# Correspondence may be addressed to Fengfeng Zhou: FengfengZhou@gmail.com or ffzhou@jlu.edu.cn. Lab web site: http://www.healthinformaticslab.org/. Phone: +86-431-8516-6024. Fax: +86-431-8516-6024.

Supplementary Figure S1
Comparison of the classification metric mAcc and the number of features between the original features and the FeCO3 features using the feature selection algorithm McTwo2. The series “original” gives the data of the original features, and the series “FeCO3” illustrates the status of the FeCO3 features. The series “Improvement” is the mAcc value of the FeCO3 features minus the mAcc value of the original features. The classification accuracy is calculated using the 5-fold cross validation strategy. The data was calculated using the feature selection algorithms

Comparison of the classification metric mAcc and the number of features between the original features and the FeCO3 features using the feature selection algorithm MI. The series “original” gives the data of the original features, and the series “FeCO3” illustrates the status of the FeCO3 features. The series “Improvement” is the mAcc value of the FeCO3 features minus the mAcc value of the original features. The classification accuracy is calculated using the 5-fold cross validation strategy. The data was calculated using the feature selection algorithms
Comparison of the classification metric mAcc and the number of features between the original features and the FeCO3 features using the feature selection algorithm Wrank.

The series “original” gives the data of the original features, and the series “FeCO3” illustrates the status of the FeCO3 features. The series “Improvement” is the mAcc value of the FeCO3 features minus the mAcc value of the original features. The classification accuracy is calculated using the 5-fold cross validation strategy. The data was calculated using the feature selection algorithms.

Comparison of the classification metric mAcc and the number of features between the original features and the FeCO3 features using the feature selection algorithm Trank. The series “original” gives the data of the original features, and the series “FeCO3” illustrates the status of the FeCO3 features. The series “Improvement” is the mAcc value of the FeCO3 features minus the mAcc value of the original features. The classification accuracy is calculated using the 5-fold cross validation strategy. The data was calculated using the feature selection algorithms.
Comparison of the classification metric mAcc and the number of features between the original features and the FeCO3 features using the feature selection algorithm DT. The series “original” gives the data of the original features, and the series “FeCO3” illustrates the status of the FeCO3 features. The series “Improvement” is the mAcc value of the FeCO3 features minus the mAcc value of the original features. The classification accuracy is calculated using the 5-fold cross validation strategy. The data was calculated using the feature selection algorithms.

Comparison of the classification metric mAcc and the number of features between the original features and the FeCO3 features using the feature selection algorithm L2-LR. The series “original” gives the data of the original features, and the series “FeCO3” illustrates the status of the FeCO3 features. The series “Improvement” is the mAcc value of the FeCO3 features minus the mAcc value of the original features. The classification accuracy is calculated using the 5-fold cross validation strategy. The data was calculated using the feature selection algorithms.
Comparison of the classification metric mAcc and the number of features between the original features and the FeCO3 features using the feature selection algorithm L2-SVM. The series “original” gives the data of the original features, and the series “FeCO3” illustrates the status of the FeCO3 features. The series “Improvement” is the mAcc value of the FeCO3 features minus the mAcc value of the original features. The classification accuracy is calculated using the 5-fold cross validation strategy. The data was calculated using the feature selection algorithms.

Comparison of the classification metric mAcc and the number of features between the original features and the FeCO3 features using the feature selection algorithm XGB. The series “original” gives the data of the original features, and the series “FeCO3” illustrates the status of the FeCO3 features. The series “Improvement” is the mAcc value of the FeCO3 features minus the mAcc value of the original features. The classification accuracy is calculated using the 5-fold cross validation strategy. The data was calculated using the feature selection algorithms.
References


