LC-MS-based Urinary Metabolite Signatures in Idiopathic Parkinson's Disease

Hemi Luan[1#], Liang-Feng Liu[2#], Nan Meng[3], Zhi Tang[1], Ka-Kit Chua[2], Lei-Lei Chen[2], Ju-Xian Song[2], Vincent CT Mok[4], Li-Xia Xie[5], Min Li[2]*, Zongwei Cai[1]*

[1] Department of Chemistry, Hong Kong Baptist University, Science Tower T1304, Hong Kong SAR, China

[2] School of Chinese Medicine, Hong Kong Baptist University, 7 Baptist University Road, Hong Kong SAR, China

[3] Beijing Genomics Institute-Shenzhen (BGI-Shenzhen), Build 11, Beishan Industrial Zone, Shenzhen 518083, China

[4] Department of Medicine and Therapeutics, Faculty of Medicine, The Chinese University of Hong Kong, 9/F., Lui Che Woo Clinical Sciences Building, Hong Kong SAR, China

[5] Department of Pharmacy, Nanshan Affiliated Hospital of Guangdong Medical College, 89 Taoyuan Road, Shenzhen 518052, China

* To whom correspondence may be addressed

[#] These two authors contributed equally to this work.

Name and complete address for correspondence:

Dr. Min LI

Tel: +852 3411 2919; Fax: +852 3411 2461; E-mail: limin@hkbu.edu.hk

Address: No. 7 Hong Kong Baptist University Road, Kowloon Tong, Hong Kong SAR, China

Prof. Zongwei CAI

Tel: +852-34117070; Fax: +852-34117348; E-mail: zwcai@hkbu.edu.hk

Address: 224, Waterloo Road, Kowloon Tong, Hong Kong SAR, China

Materials and Methods

Data analysis

Permutation Multivariate Analysis of Variance (PERMANOVA), a permutation-based version of the multivariate analysis of variance, was employed to test the statistical significant differences between metabolic profiles and individuals' phenotypes. PERMANOVA analyses were performed in R using the "vegan" package [1]. Orthogonal projections to latent structures discriminant analysis (OPLS-DA), was performed by using SIMCA 13 software (Umetrics, Umea, Sweden). Coupling the receiver operating characteristic curve with its area under the curve (AUC), a widely used method to estimate the diagnosis potential of a classifier in clinical applications, were performed in R using the "pROC" package[2]. The Random Forest has previously been used to diagnose Alzheimer and Parkinson's Diseases, which suggested that the Random Forest gives the ideal prediction accuracy. Random Forest has been employed in many research fields, including metabolites selection and cancer classification. The Random Forest analyses were performed in R using the "randomForest" package [3].

References

1. Dixon, P., VEGAN, a package of R functions for community ecology. *Journal of Vegetation Science* **2003,** 14, (6), 927-930.
2. Robin, X.; Turck, N.; Hainard, A.; Tiberti, N.; Lisacek, F.; Sanchez, J.-C.; Müller, M., pROC: an open-source package for R and S+ to analyze and compare ROC curves. *Bmc Bioinformatics* **2011,** 12, (1), 77.
3. Liaw, A.; Wiener, M., Classification and Regression by randomForest. *R news* **2002,** 2, (3), 18-22.

Table S1 Results of PERMANOVA for PD patients and Controls

	***P* value**	
	PD × Controls	PD
Group	1.00×10^{-4}	-
Gender	3.75×10^{-1}	4.21×10^{-1}
Age	4.75×10^{-2}	9.32×10^{-2}
H & Y Grade	-	1.00×10^{-4}
UPDRS motor score	-	3.91×10^{-2}

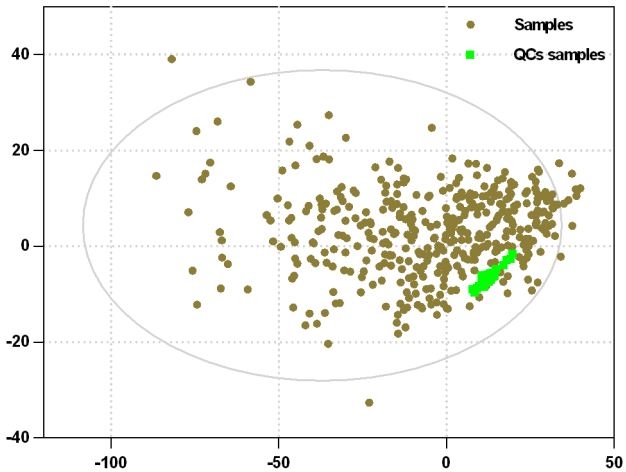

Figure S1 two dimensional principal component analysis (2D PCA) scores plot for consecutively analyzed QC samples. (Green indicates QC samples; Brown, samples)

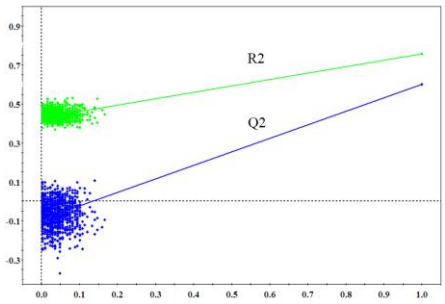

Figure S2 the permutation test (n=999) for the OPLS-DA model.

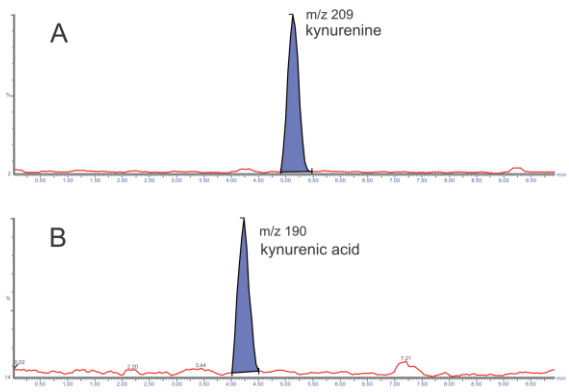

Figure S3 the MRM chromatograms of a representative sample of fly, (A) Kynurenine; (B) Kynurenic acid.