

## Supporting Information

### Investigation of the ecological roles of putative keystone taxa during tailing revegetation

Xiaoxu Sun<sup>1,2</sup>, Rui Xu<sup>1,2</sup>, Yiran Dong<sup>3</sup>, Fangbai Li<sup>1,2</sup>, Wan Tao<sup>1,2</sup>, Tianle Kong<sup>1,2</sup>, Miaomiao Zhang<sup>1,2</sup>, Lang Qiu<sup>1,2</sup>, Xiaoyu Wang<sup>1,2</sup>, Weimin Sun<sup>1,2,\*</sup>

1. National-Regional Joint Engineering Research Center for Soil Pollution Control and Remediation in South China, Guangdong Key Laboratory of Integrated Agro-environmental Pollution Control and Management, Guangdong Institute of Eco-environmental Science & Technology, Guangdong Academy of Sciences, Guangzhou 510650, China
2. Guangdong-Hong Kong-Macao Joint Laboratory for Environmental Pollution and Control, Guangzhou 510650, China
3. School of Environmental Studies, China University of Geosciences (Wuhan), Wuhan 430074, China

\*Correspondence to: Dr. Weimin Sun

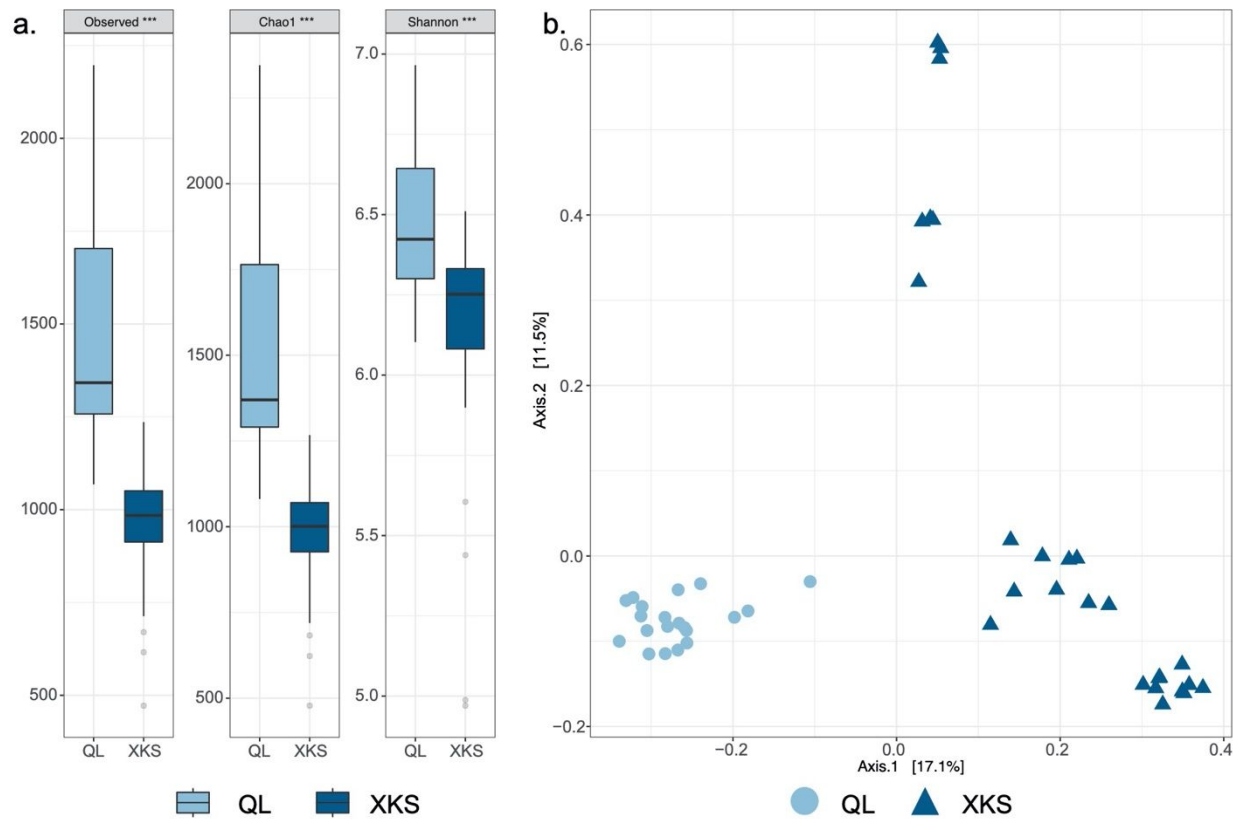
808 Tianyuan Road, Guangzhou, Guangdong, China

Phone: 86-020-87024633

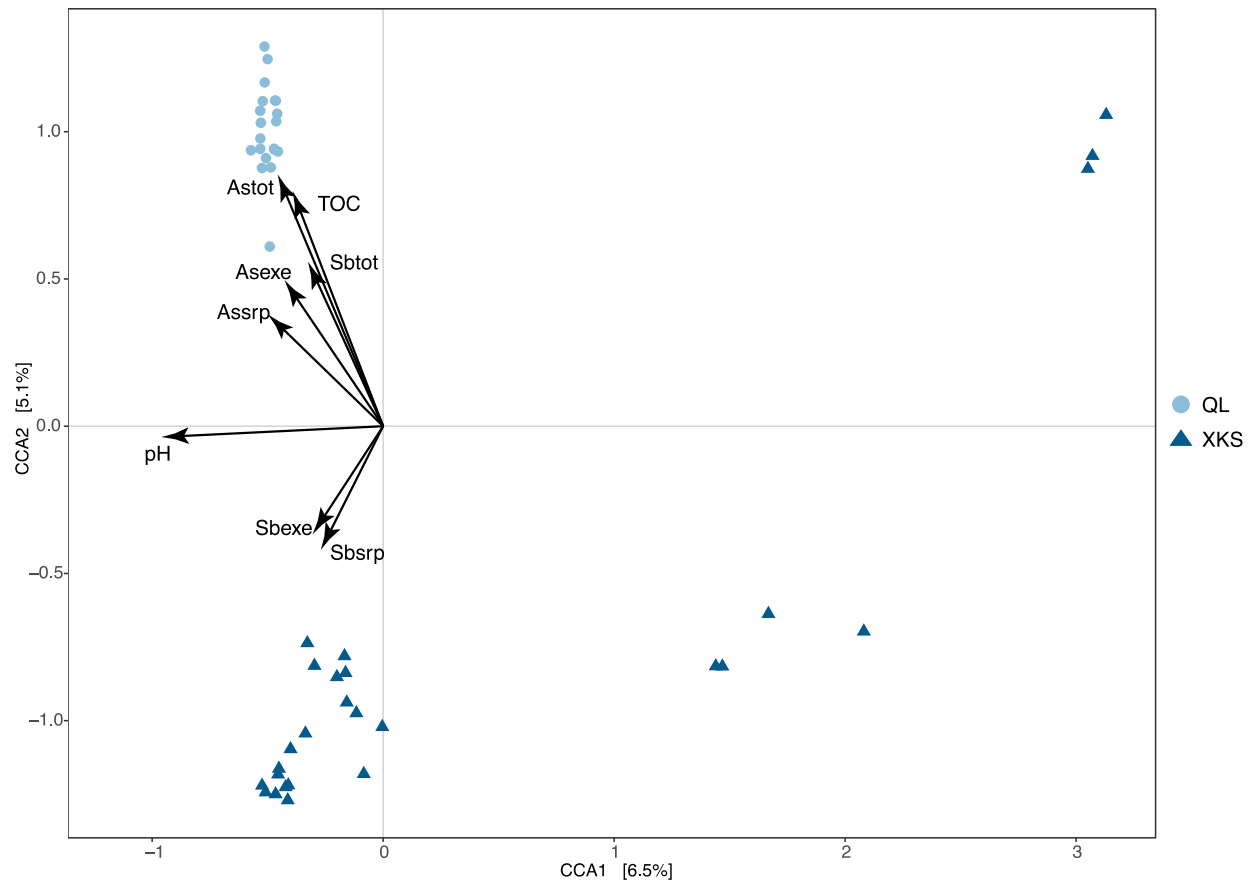
Fax: 86-020-87024123

Email: [wmsun@soil.gd.cn](mailto:wmsun@soil.gd.cn)

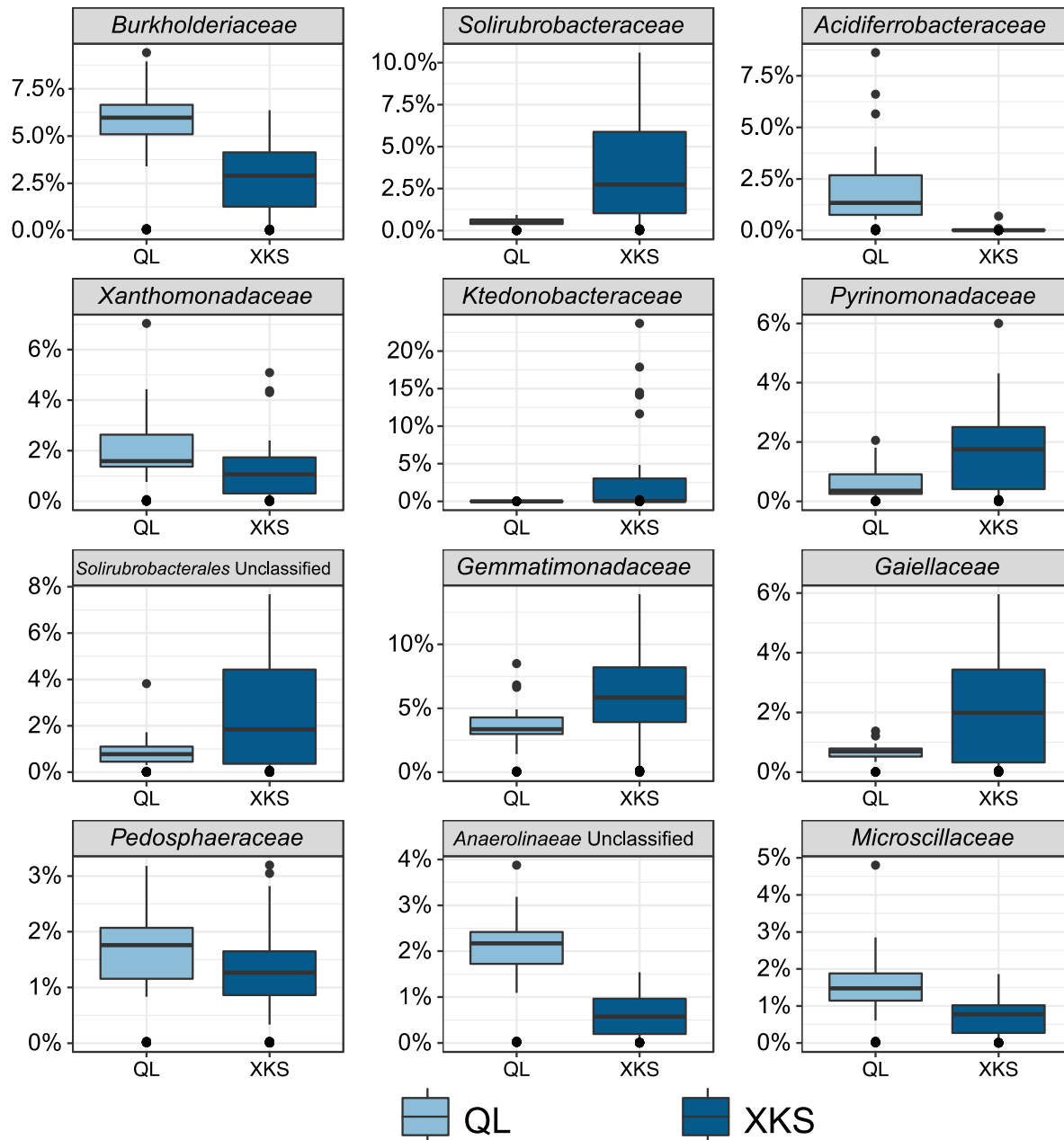
Supporting Information Available: Figure S1-S5 and Table S1-S8.



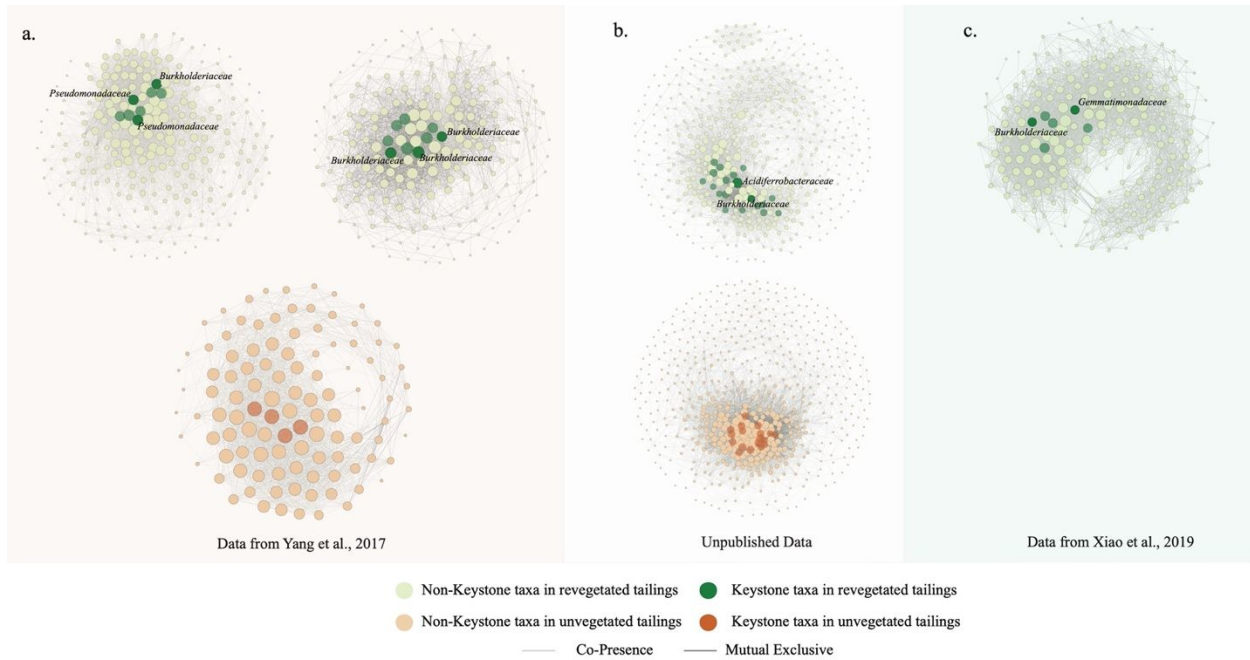
**Figure S1.** The alpha diversity (a) measured in three indices of the microbial community between the abandoned (QL) and active (XKS) tailings. Samples were rarefied to the number of counts in the lowest libraries. The PCoA plots of beta similarities (b) measured as Bray-Curtis distances for microbial community compositions between the abandoned (QL) and active (XKS) tailings.



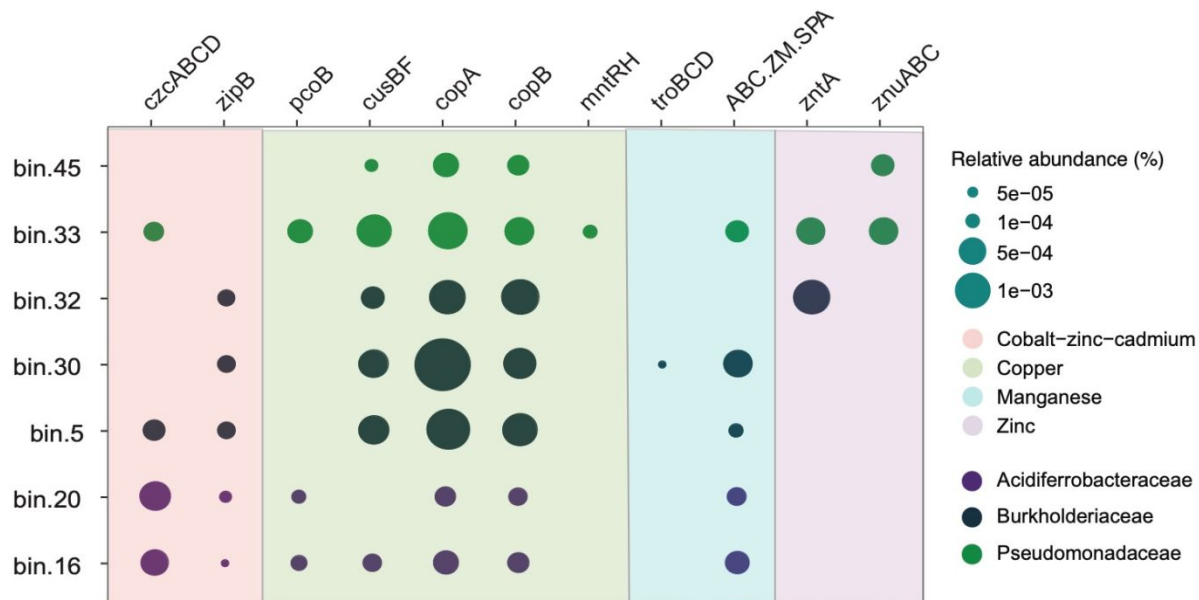
**Figure S2.** The CCA plot demonstrating the relative importance of geochemical parameters on the microbial community compositions. The length of the arrows is proportional to its relative importance.



**Figure S3.** Boxplot showing the relative abundance of the major families that were differentially distributed between the abandoned (QL) and active (XKS) tailings. The \* sign represents the statistical significance between the sample types (\* indicates  $p < 0.05$ , \*\* indicates  $p < 0.01$ , \*\*\* indicates  $p < 0.001$ ).



**Figure S4.** Co-occurrence network analysis showing the biological interactions in a) a revegetated tailing and a unvegetated tailing of a Cu mine; b) three tailings revegetated with *Pteris vittata* and their unvegetated tailings counterparts; and c) a tailing revegetated with *Biden pinnata*. Edges are shown only strong (Spearman correlation  $> |0.6|$ ) and significant ( $p < 0.05$ ) connections. Size of the nodes is proportional to the degree of connectedness. The thickness of the edges is proportional to the strongness of the correlation.



**Figure S5.** Percentage of reads assigned to pathways related to metal resistance in each bin. The size of the dot is proportion to the relative abundance. *czcABCD*, Cobalt-zinc-cadmium resistance; *zipB*, Zinc and cadmium transporter; *pcoB*, Copper resistance protein B; *cusBF*, CU(I)/Ag(I) efflux system periplasmic protein; *copA*,  $\text{Cu}^+$ -exporting ATPase; *copB*,  $\text{Cu}^{2+}$ -exporting ATPase; *mntRH*, Manganese transport regulator; *troBCD*, Manganese/zinc/iron transport system permease protein (); *ABC.ZM.SPA*, Zinc/manganese transport system permease protein; *zntA*,  $\text{Cd}^{2+}/\text{Zn}^{2+}$ -exporting ATPase (); *znuABC*, Zinc transport system permease protein.