



读书报告

黄建蓉

2016.10.15

Illumina Miseq

Diversity



Environ Sci Pollut Res (2015) 22:10788–10799
DOI 10.1007/s11356-015-4186-3

RESEARCH ARTICLE

Illumina MiSeq sequencing investigation on the contrasting soil bacterial community structures in different iron mining areas

Chen Hong · Yanxiao Si · Yi Xing · Yang Li

Received: 10 October 2014 / Accepted: 29 January 2015 / Published online: 12 March 2015
© Springer-Verlag Berlin Heidelberg 2015

借助Illumina Miseq 测序技术概述不同铁矿区土壤细菌结构的区别



IF 2.760

Environmental Science and Pollution Research

ISSN: 0944-1344 (Print) 1614-7499 (Online)

Description

Environmental Science and Pollution Research (ESPR) serves the international community in all areas of Environmental Science and related subjects with emphasis on chemical compounds. It reports from a broad interdisciplinary outlook. Apart from the strictly scientific contributions as research articles (short and full papers) and reviews, ESPR publishes: news & views from research and technology, legislation and regulation, h ... [show all](#)



[Browse Volumes & Issues](#)

Impact Factor	Available
2.760	1994 - 2016
Volumes	Issues

CONTENTS

1、 Introduction

2、 Methods

3、 Results

4、 Inspiration



Introduction



采矿活动对周围土壤的质量和地下蓄水层产生不利影响，威胁人类健康和生态系统。



在铁矿开采地区，大量的尾矿直接堆放和废铁的沉积，借助风力和水力侵蚀周围的土壤，使其带来大量的不同种类有害重金属元素（即Cu、Zn、Cr、Cd、Pb）。



Introduction



由采矿活动造成的污染分布更广泛，更温和且神秘。同时微生物也广泛分布于重金属污染地区，可能会出现独特以及微妙的群落变化。



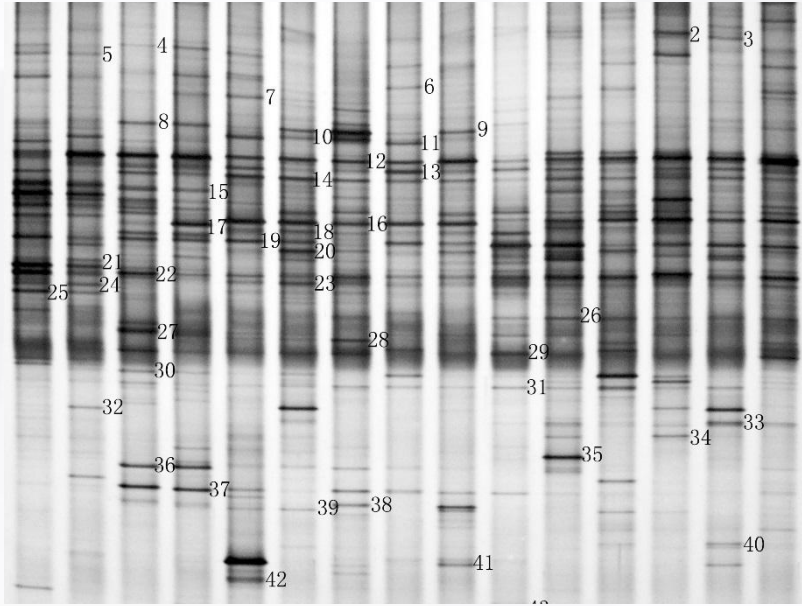
研究该类土壤微生物群落结构不仅能指导我们探索受影响土壤的潜在风险，并且找到一些土著具有抵抗功能的细菌为土壤修复提供参考。



Introduction

分子指纹图谱技术飞速发展，如变性梯度凝胶电泳（DGGE），末端限制性片段长度多态性，实时定量PCR，使得研究者能够探索重金属污染地区土壤的微生物多样性。

分子指纹图谱技术具有很高的分辨率，然而图谱的复杂性受到丰富度高的土著微生物群落影响，可能会掩盖由重金属污染引起的变化。具有重要生态意义的稀有种类，可能由于DGGE检测的限制而受到影响。





Introduction

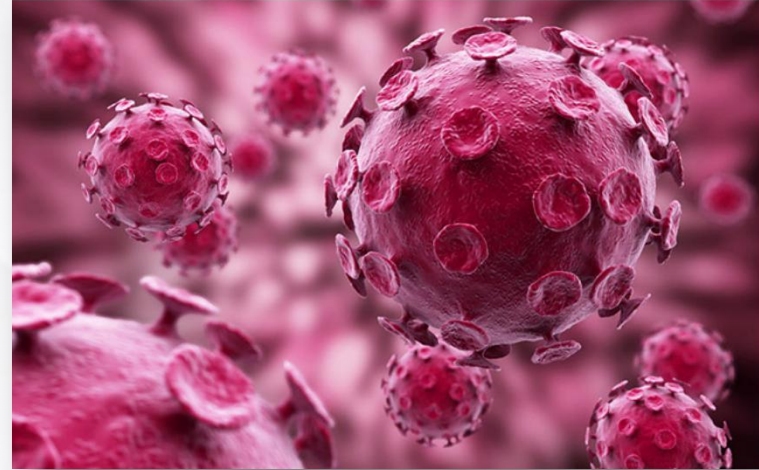


二代测序如454焦磷酸测序和 Illumina的应用，可提供一个直接方法来检测微生物的类群，特别是那些低丰富度类群。

此外，Illumina 测序成本更低，并且比454 焦磷酸测序能获取十倍或者更多的序列信息。



Introduction



本研究选择北京密云县鞍子沟铁矿区作为采样点。为了获得全面的细菌群落结构变化，作者对比了三个地区土壤细菌群落结构的差异。

目的是在重金属污染下，来源于不同土壤类型中的微生物类群详细的变化差异。



样品采集

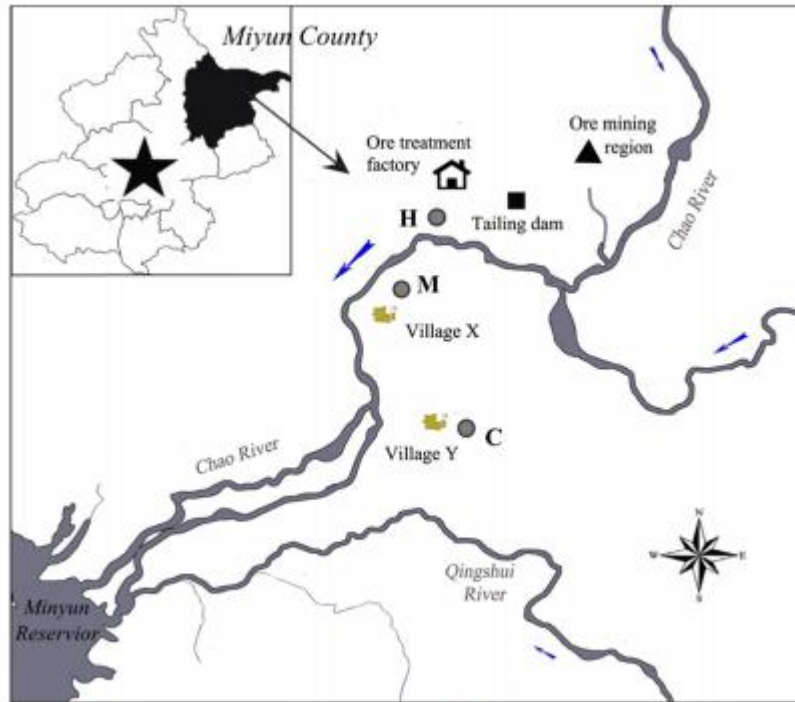


Fig. 1 The study area and sampling sites location

Three H samples (矿石处理厂, 矿开采区和尾矿堆)

Three M samples (农田, 人工林和菜田) 1 Km

Three C samples (农田, 人工林和菜田) 5 Km

*Sample at each site was collected in triplicates.



Methods

指标检测

Water content

pH value

Organic matter

Available phosphorus

NH_4^+

NO_3^-

Heavy metal contents (Cd, Cr, Cu, Zn, Fe)

Nemerow Index (内梅罗水污染指数)



Methods

DNA提取, PCR扩增, Illumina 测序

0.5g soil sample

Power Soil DNA kit



The V3–V4 regions of the bacterial 16S rRNA gene

338F 5'-barcode-**ACTCCTACGGGAGGCAGCA**-3'

806R 5'-**GGACTACHVGGGTWTCTAAT**-3'

Illumina MiSeq platform (Majorbio, Shanghai)





测序数据的处理与分析

Raw FASTQ files were de-multiplexed and quality-filtered using **QIIME** (version 1.17)

Operational taxonomic units (OTUs) with 97 % similarity cutoff were clustered using **UPARSE** (version 7.1)

chimeric sequences were identified and removed using **UCHIME**.



测序数据的处理与分析

The rarefaction analysis based on **Mothur** v.1.21.1 diversity indices (Chao, ACE, and Shannon diversity indices)

Principal component analysis (PCA) using the community ecology package, **R-forge**

Venn diagrams were implemented by **Venn Diagram**

Mantel test, Redundancy analysis (RDA), and Heatmap figures were performed in **Vegan** packages in R



Table 1 pH value and heavy metal contents of soil samples

Sampling sites	pH	Water contents (%)	Organic matter (g/kg)	Available phosphorus (mg/kg)	NH ₄ ⁺ (mg/kg)	NO ₃ ⁻ (mg/kg)	Heavy metal contents					Nemerow index
							Cd (mg/kg)	Cr (mg/kg)	Cu (mg/kg)	Zn (mg/kg)	Fe (g/kg)	
HF	7.23 (±0.12)	7.69 (±0.35)	0.77 (±0.001)	4.12 (±0.11)	2.09 (±0.05)	0.14 (±0.01)	0.69 (±0.01)	280.60 (±10.09)	47.32 (±1.09)	101.23 (±3.48)	89.27 (±3.25)	3.01
HM	7.69 (±0.20)	9.10 (±0.19)	0.91 (±0.001)	5.54 (±0.14)	7.11 (±0.29)	0.57 (±0.01)	0.75 (±0.03)	324.44 (±9.20)	62.11 (±2.31)	79.12 (±3.29)	104.01 (±4.08)	3.17
HT	7.81 (±0.35)	8.34 (±0.31)	0.55 (±0.002)	3.79 (±0.32)	1.25 (±0.04)	6.03 (±0.04)	0.68 (±0.02)	294.25 (±7.28)	61.29 (±3.37)	102.30 (±3.24)	125.70 (±2.75)	3.02
Mean value	7.58	8.38	0.72	4.48	3.48	2.24	0.71	299.76	59.91	94.22	106.33	3.06
MC	7.26 (±0.11)	19.72 (±0.98)	1.01 (±0.004)	60.27 (±2.92)	1.09 (±0.03)	0.22 (±0.01)	0.49 (±0.03)	156.29 (±5.09)	35.76 (±0.97)	95.52 (±4.12)	29.36 (±1.02)	2.05
MF	7.19 (±0.42)	18.39 (±1.02)	1.14 (±0.004)	27.54 (±1.22)	0.78 (±0.01)	19.13 (±1.88)	0.50 (±0.02)	134.76 (±4.29)	50.67 (±4.05)	117.45 (±4.38)	31.83 (±0.96)	2.12
MV	7.56 (±0.18)	21.95 (±0.68)	1.39 (±0.002)	58.54 (±2.16)	0.21 (±0.01)	24.02 (±1.27)	0.49 (±0.01)	199.30 (±8.68)	47.44 (±3.55)	106.83 (±3.27)	30.02 (±1.88)	2.14
Mean value	7.34	20.02	1.18	48.78	0.69	14.45	0.49	163.45	44.62	106.60	30.40	2.10
CC	7.27 (±0.20)	19.31 (±0.87)	0.87 (±0.001)	5.24 (±0.41)	8.57 (±0.02)	5.47 (±0.02)	0.19 (±0.01)	110.49 (±3.94)	49.01 (±2.04)	87.10 (±4.03)	24.14 (±0.52)	0.98
CF	7.71 (±0.36)	17.91 (±1.05)	0.72 (±0.002)	26.12 (±1.52)	0.98 (±0.01)	12.42 (±0.54)	0.17 (±0.02)	98.75 (±2.05)	32.49 (±1.56)	76.38 (±2.40)	27.67 (±0.81)	0.88
CV	5.07 (±0.45)	22.42 (±1.51)	1.07 (±0.005)	13.88 (±0.79)	1.56 (±0.01)	51.33 (±1.88)	0.09 (±0.02)	132.37 (±5.91)	36.70 (±1.95)	92.39 (±4.57)	28.22 (±0.74)	0.76
Mean value	6.68	19.88	0.89	15.08	3.70	23.07	0.15	113.84	39.40	85.29	26.68	0.87
Background heavy metal contents in Beijing							0.12	29.8	18.7	57.5		
Background in China							0.20	90	35	100		

Values represent average±standard deviation

HF heavily polluted sites from ore treatment factory, *HM* heavily polluted sites from ore mining region, *HT* heavily polluted sites from tailing dam, *MC* moderately polluted sites from croplands, *MF* moderately polluted sites from planted forest, *MV* moderately polluted sites from vegetable fields from village X, *CC* clear sites from croplands, *CF* clear sites from planted forest, *CV* clear sites from vegetable fields from village Y

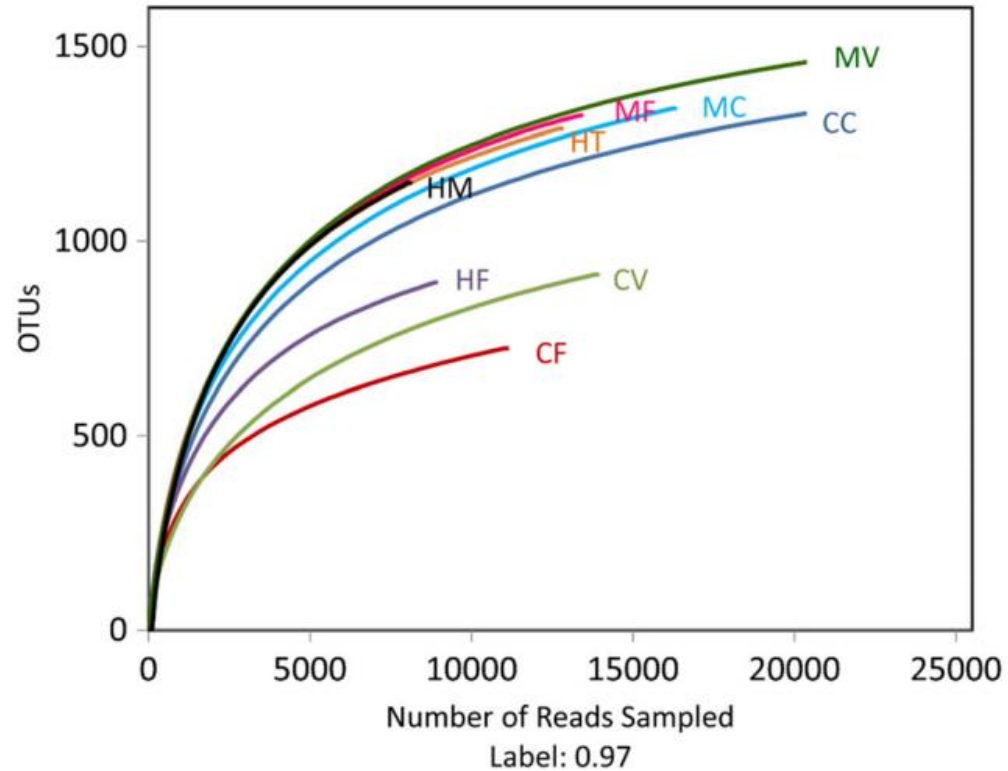


Fig. 2 Rarefaction curves of the OTU number at 97 % similarity boxplot for every samples

所有的稀疏曲线接近饱和状态，表明测序数据是合理的。

发现数量多的reads却只贡献了OTU总数的小部分。



Table 2 MiSeq sequencing results and diversity estimates for each sampling site

Sample	Sequencing results		Diversity estimates ^a		
	Total sequences	Total OTUs ^a	ACE	Chao	Shannon
HF	8877	894	1081±57	1091±78	5.97±0.03
HM	7957	1150	1407±65	1426±88	6.23±0.03
HT	12,721	1289	1492±48	1522±82	6.24±0.02
Mean value	9851	1111	1327	1346	6.15
MC	16,209	1340	1577±60	1603±89	6.16±0.02
MF	13,384	1323	1522±53	1534±73	6.08±0.03
MV	22,327	1483	1633±44	1670±71	6.08±0.02
Mean value	17,306	1382	1577	1602	6.11
CC	20,324	1327	1519±53	1535±75	5.98±0.02
CF	11,014	724	931±68	1029±138	5.43±0.03
CV	13,805	914	1157±69	1164±93	5.3±0.03
Mean value	15,048	978	1202	1242	5.57

ACE abundance-based coverage estimator, *Chao* Chao's species richness estimator, *Shannon* Shannon-Weiner Index

^a Species level, 97 % similarity threshold used to define operational taxonomic units (OTUs)

平均的香农指数值

H组6.15, M组6.11, C组5.57。

这些结果表明, 受污染的H区域和M区域多样性高于未受污染的C区。



Taxonomy composition

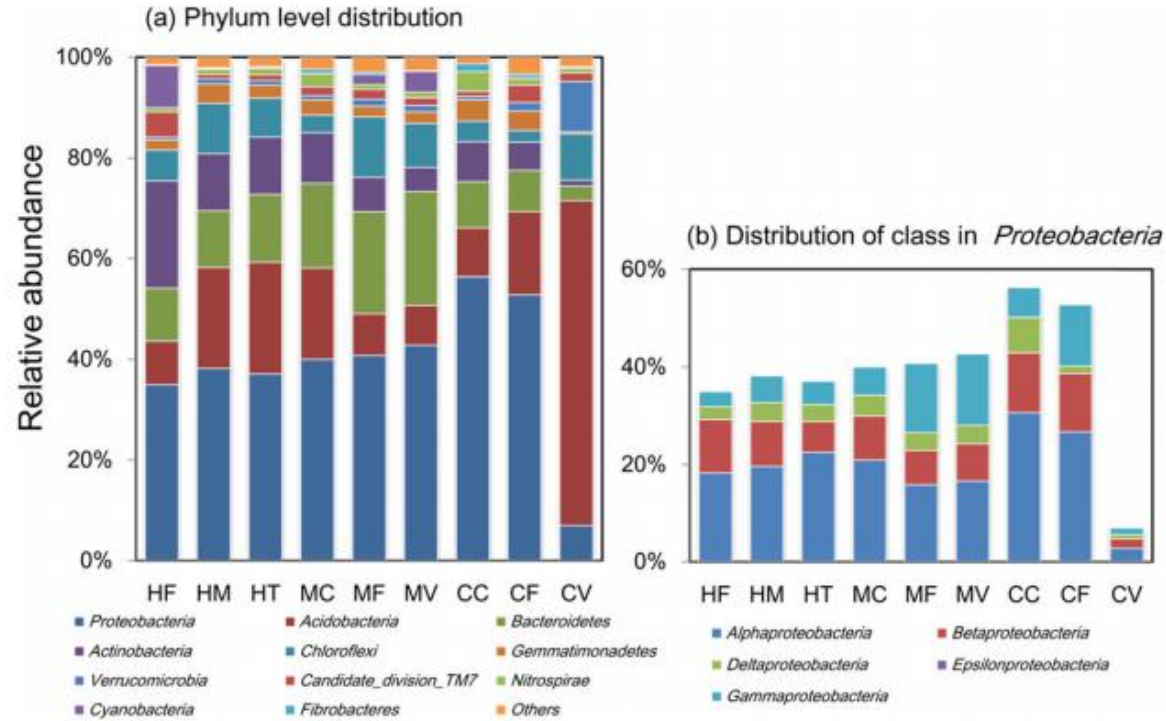


Fig. 3 Taxonomic distribution of soil samples.

a Phylum distribution of all samples.

b Class distribution of Proteobacteria phylum.

OTU 可分为35个门 251个科 381个属，其中6个门在9个样点中是常见的类群，在每个文库中占85%。

变形菌门是最丰富的类群，大约占有所有样品中OTU的29.45%，reads 的39.95%，而酸杆菌门是第二大丰富的类群。

在变形门中，Alphaproteobacteria 占主导地位。



Mantel 检验结果表明土壤细菌与**重金属含量**呈显著相关，而土壤的属性并没有明显影响细菌。

表明重金属离子在形成当地细菌群落结构中发挥更重要的角色。

为了进一步查明影响土壤微生物群落结构的主要变量，构建了RDA图。



Results

Correlation between community structure and environmental factors

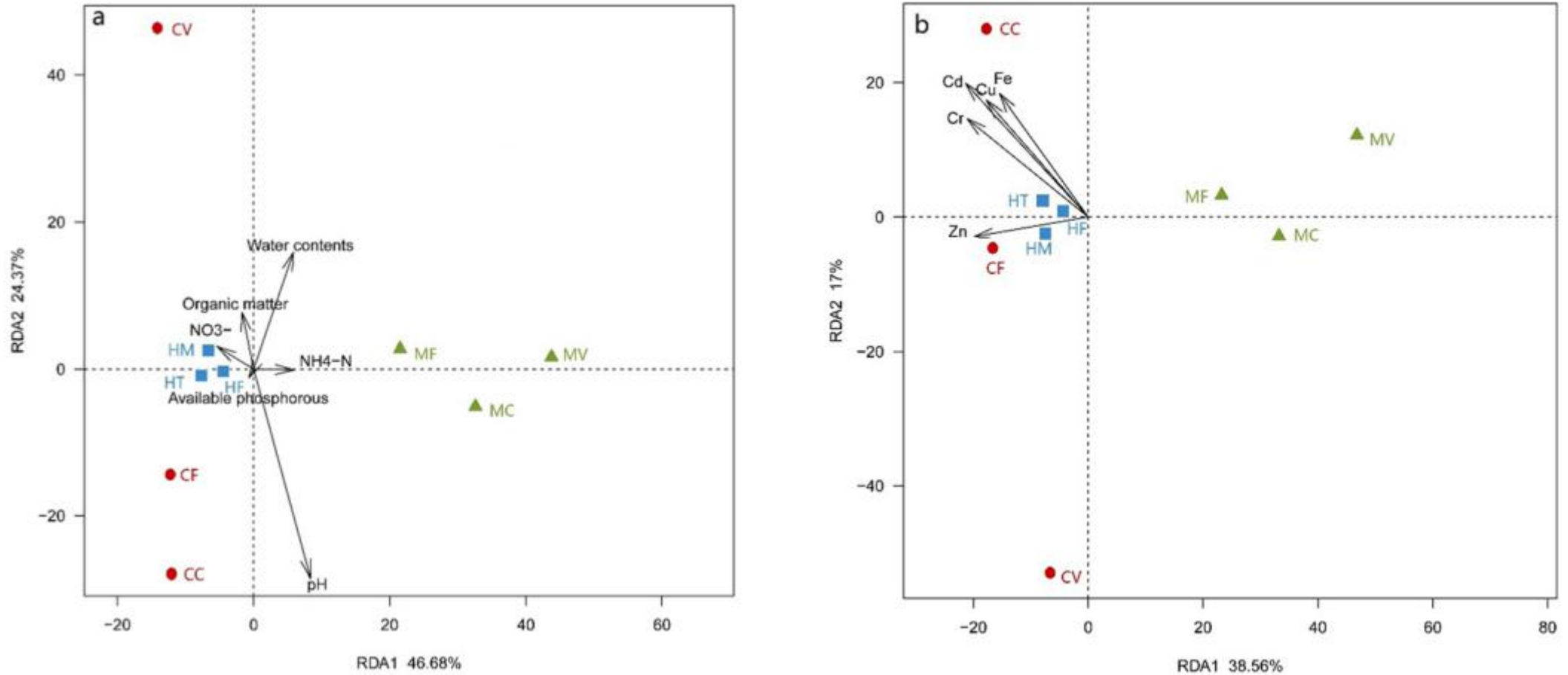


Fig. 4 RDA ordination plot showed the relationship between the bacteria community structure and soil properties (a) or heavy metal (b)



Results

Table 3 Significance of environmental variables in explaining the bacteria community structure obtained from the RDA results

	r^2	<i>P</i> value
Soil		
Water contents	0.309	0.315
pH	0.952	0.001**
Organic matter	0.066	0.826
Available phosphorus	0.002	0.994
NH ₄ ⁺	0.038	0.899
NO ₃ ⁻	0.042	0.832
Heavy metal		
Cd	0.862	0.001**
Cr	0.674	0.011*
Cu	0.622	0.033*
Zn	0.375	0.232
Fe	0.576	0.046*

r^2 indicates the decision coefficient of environmental variables on the community structure

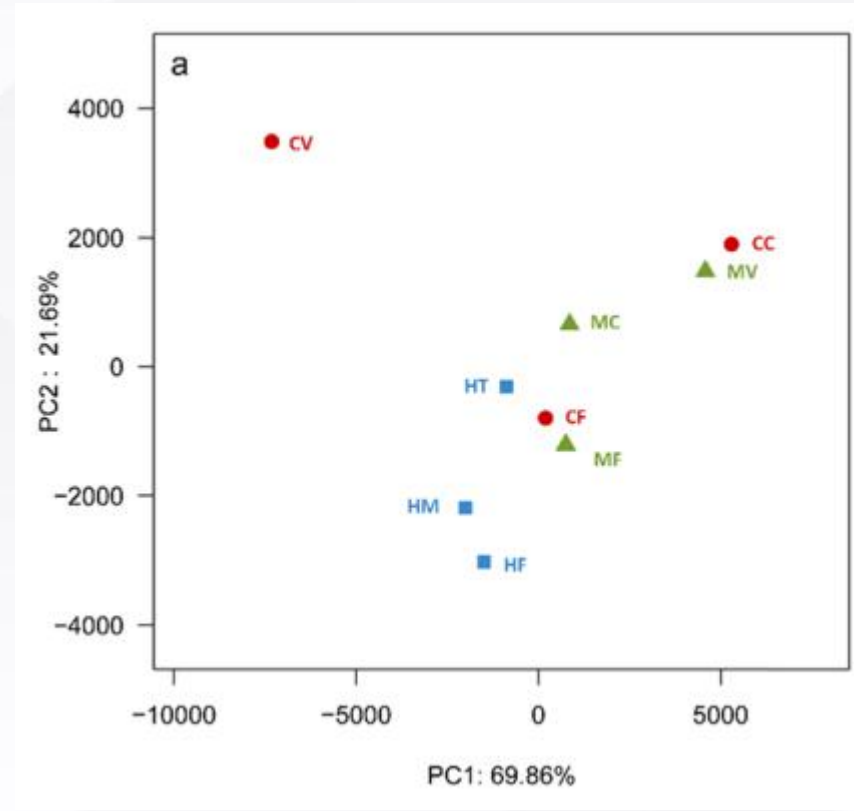
通过RDA的结果可以得到影响细菌群落的参数。

结果表明, Cd, Cr, Cu, Fe 和土壤pH 值是主导群落组成变化的最主要因素。



Comparison of bacterial community structure among groups

PCA



Multivariate analysis based on information at phylum, genus, and OTU levels.

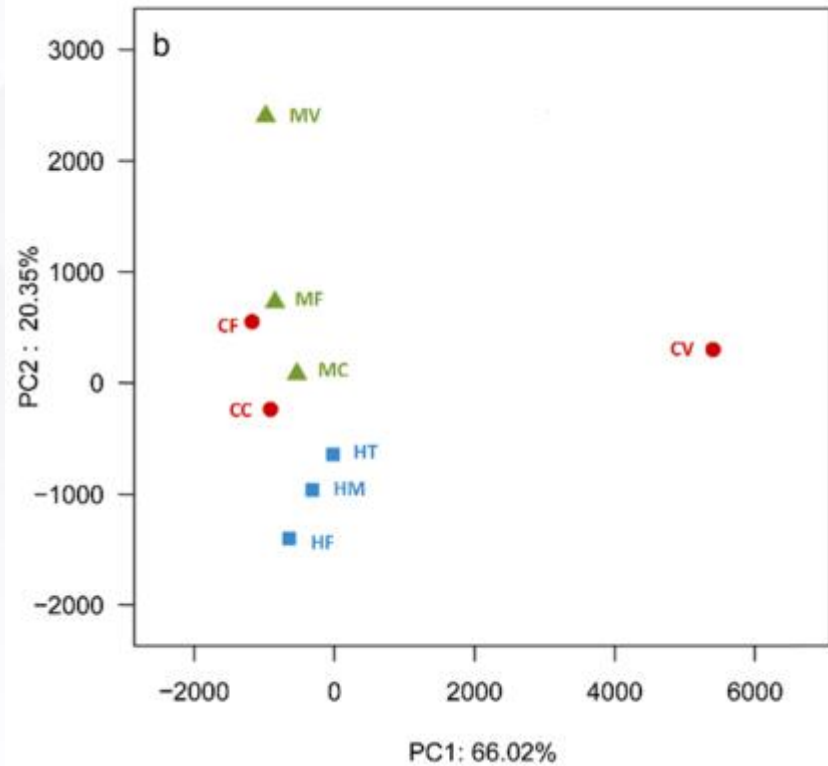
a PCA based on phylum distribution

除了CV组与其余样点在门水平不同，表明其余样点在门水平多样性组成相似。



Comparison of bacterial community structure among groups

PCA



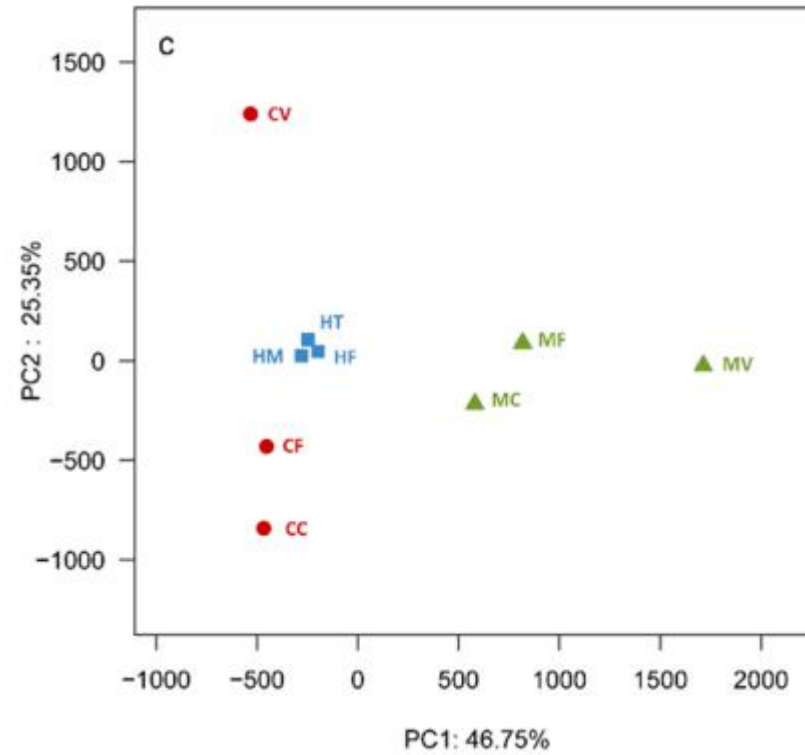
Multivariate analysis based on information at phylum, genus, and OTU levels.

b, PCA based on genus distribution

H组的聚集在一块，而其他组却没有明确的界限。



PCA



Multivariate analysis based on information at phylum, genus, and OTU levels.

c PCA based on total OTU level information

而在OTU水平上，PCA分析表明，中度污染的M组和H组聚集在一起，和C组距离较远。H组的样点相比较其他组更集中，从而指示在污染严重的地区，细菌群落结构在很大程度上很相似。



这一发现和上述alpha多样性的结果一致。

然而，土壤性质对群落结构的影响不能忽视。每个水平的PCA分析都表明，未受污染区的CV组和其他两组存在较大的差距就是和CV中酸杆菌门的存在密切相关，由于在样点具有较强酸性的性质。许多研究已经正式发现，这种现象与较强酸性土壤的性质有密切的关系。

此外，与PCA在属和OTU水平不同，在门这一水平，三组没有明确的边界。



Venn diagram

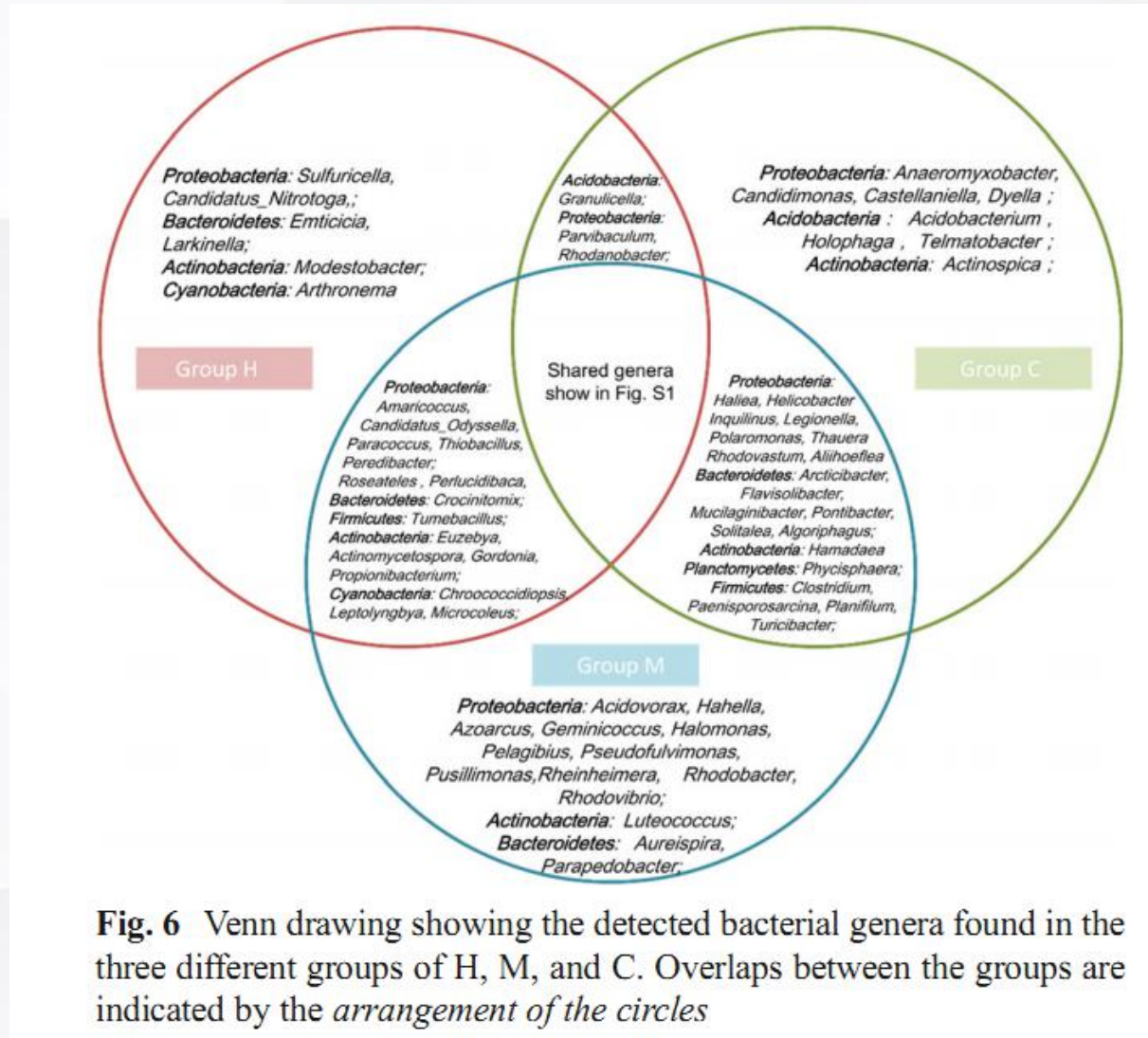


Fig. 6 Venn drawing showing the detected bacterial genera found in the three different groups of H, M, and C. Overlaps between the groups are indicated by the *arrangement of the circles*

共发现341个属，并且其中79.76%属于共享的属。



Results

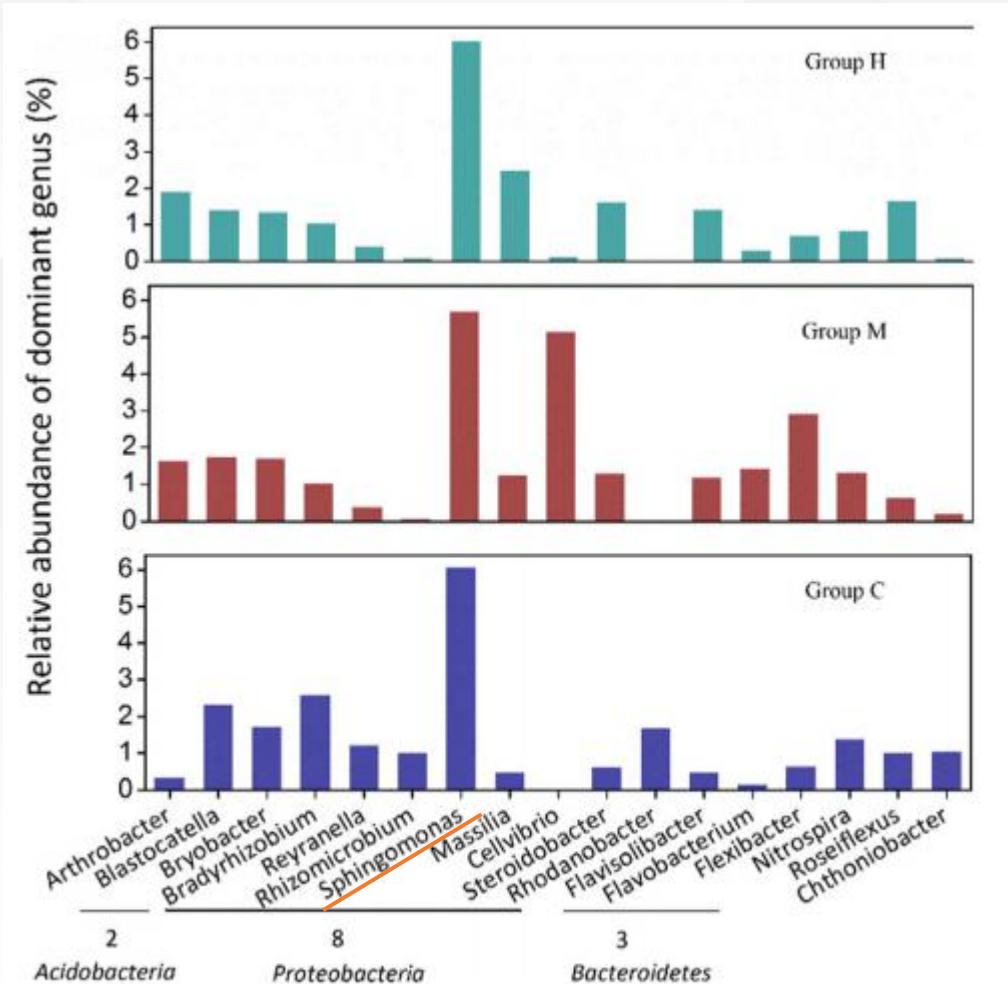


Fig. 8 Relative abundance of dominant genera in different groups. The genera with an average abundance of >1 % in at least one group were defined as dominant

除了未知属，优势类群相对丰富变化程度随着污染程度发生变化。

三组中最丰富的类群是*Sphingomonas*鞘氨醇单胞菌。



Results

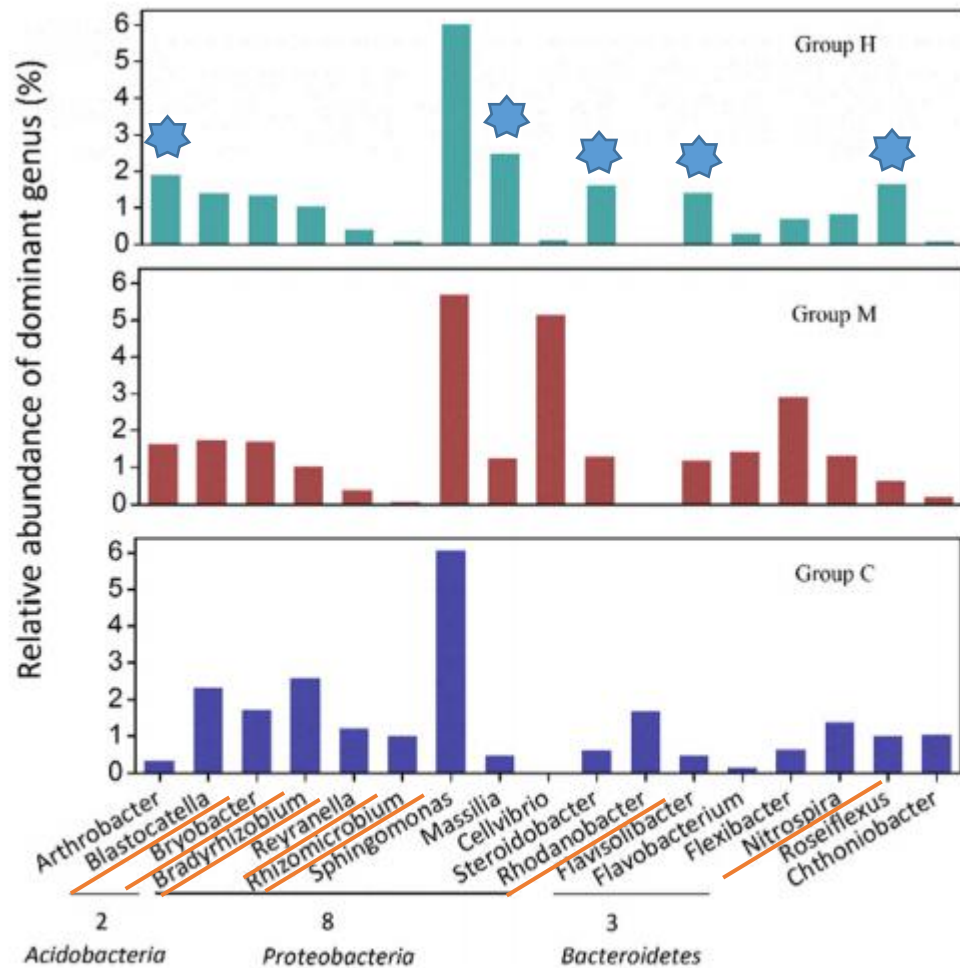


Fig. 8 Relative abundance of dominant genera in different groups. The genera with an average abundance of >1 % in at least one group were defined as dominant

Bradyrhizobium、*Rhodanobacter*、*Reyranelia* 和 *Rhizomicrobium* 显著下降（从组 C 至 H）。

同时，相对丰度上，*Blastocatella*，*Bryobacter* 和 *Nitrospira* 略有下降，但维持他们的优势地位。

下列属丰度H组仍明显高于其他组，*Steroidobacter*、*Massilia*、*Arthrobacter*、*Flavisolibacter* 和 *Roseiflexus*。



Bacteria with potential resistance to heavy metals

验证在重金属污染地区具有抗重金属潜力的菌

皮尔逊相关系数（ Pearson correlation ）。数值越接近1，越相关。

节杆菌是典型的Cr⁶⁺减少细菌，和Cr的含量密切相关，在H组中成为优势类群。(R²=0.726, P<0.05)

硫杆菌被称为与催化亚铁离子和氧化无机硫密切相关的属。硫杆菌相对丰度高，并且仅发现与污染区，其相对丰度比H组高了3倍。此外，硫杆菌的相对丰度和铁含量也相关。(R²=0.691, P<0.05)

大多数蓝藻也是可以产生胞外聚合物，主要是多糖，可以吸附散布在环境中的重金属元素。蓝藻门中的鞘藻属和Zn含量有很强的相关性。(R²=0.789, P < 0.05)。

红球菌属也被报道和抵抗几种重金属相关，表现为被动或积极吸收重金属。基于韦恩图，红球菌属只出现在污染严重的地区。皮尔逊相关分析表明，Zn含量和红球菌属有重大相关。(R²= 0.801, P < 0.01)。

抗重金属菌的数量在土著微生物群落中少于2%。

在抗重金属菌在污染的土壤中数量可以增加。



Conclusions

本研究借助高通量Illumina测序技术，展现了土壤重金属污染地区微生物类群在门和属水平的变化。测序结果和 α 多样性表明，重度污染地区H样点和中度污染地区M样点的微生物多样性显著高于未污染地区。

研究发现，一些属与重金属有着密切关系，可能是一种新的自然生物修复的类群。



个性化分析

你会如何处理？



THANKS

恳请各位老师同学批评指正！