Supplementary Materials

Functional gene array-based analysis of microbial community structure in groundwaters with a gradient of contaminant levels

Patricia J. Waldron[1,2], Liyou Wu[1,2], Joy D. Van Nostrand[1,2], Chris W. Schadt[2,3], David B. Watson[2,3], Philip M. Jardine[2,3], Anthony V. Palumbo[2,3], Terry C. Hazen[2,4], Jizhong Zhou[1,2]*

[1]Institute for Environmental Genomics, Department of Botany and Microbiology, 770 Van Vleet Oval, University of Oklahoma, Norman, OK 73019

[2]Virtual Institute for Microbial Stress and Survival, http://vimss.lbl.gov

[3]Oak Ridge National Laboratory, Oak Ridge, TN

[4]Lawrence Berkeley National Laboratory, Berkeley, CA

Content:

TABLE S1. Number of unique and overlapping genes detected in wells

TABLE S2. Mantel test r-value results for environmental variables and gene diversity detected by GeoChip in FRC groundwater

TABLE S3. Summary of CCA Results from CANOCO

FIGURE S1. Relative abundance of all functional gene categories detected by the GeoChip.

FIGURE S2. Hierarchical cluster analysis of gene diversity based on hybridization signal intensities for selected functional gene groups. (A) Nitrite reductases (*nirK* and *nirS)*, (B) Genes involved in carbon fixation, (C) Sulfite-reductases (*dsrAB*), (D) Genes involved in organic contaminant degradation, (E) Genes involved in metal transport and resistance.

22 **Supplementary Figures**

TABLE S1. Number of unique and overlapping genes detected in wells

No. (%) of unique or overlapping genes[a]	FW300	FW003	FW010	FW021	FW024	TPB16
FW300	61 (20)[b]	189 (36)	80 (21)	174 (35)	111 (23)	91 (26)
FW003		25 (11)[b]	61 (17)	144 (35)	84 (20)	68 (24)
FW010			6 (5)[b]	64 (20)	118 (37)	108 (70)
FW021				10 (5)[b]	90 (24)	69 (26)
FW024					30 (16)[b]	120 (59)
TPB16						0 (0)[b]
Total no. of genes detected	302	219	130	192	190	133

[a] All values are for overlapping genes, unless indicated otherwise

[b] Values for unique genes

23

24

TABLE S2. Mantel test r-value results for environmental variables and gene diversity detected by GeoChip in FRC groundwater.

Geochemical Variable	P-Value[a]	R-value
pH	0.11111	-0.1538
aluminum	0.74775	-0.0603
Chloride	0.08008	-0.1451
nickel	0.1982	0.113
nitrate	**0.02703**	-0.2552
sulfate	0.36937	-0.1028
uranium	**0.04905**	-0.233
Tc	0.07908	-0.1731
all geochemical variables	0.07508	-0.1855

[a] Bolded values are significant (p-value<0.05)

25

26

TABLE S3. Summary of CCA Results from CANOCO.

Axes	1	2
Eigenvalues	0.490	0.148
Species-environment correlations	0.980	0.994
Cumulative percentage variance:		
of species data	47.4	61.6
of species-env relation	64.6	84.0
Sum of all eigenvalues	1.035	
Sum of all canonical eigenvalues	0.759	
Summary of Monte Carlo test		
Test of significance of first canonical axis	0.490	
F-ratio	1.800	
P-value	0.077	
Test of significance of all canonical axes	0.759	
F-ratio	1.833	
P-value	0.041	

27

FIGURE. S1. Relative abundance of all functional gene categories detected by the GeoChip. The total number of genes detected for each well was used to calculate the relative abundance of each category. The percentage of genes detected is standardized by dividing by the number of probes printed on the slide.

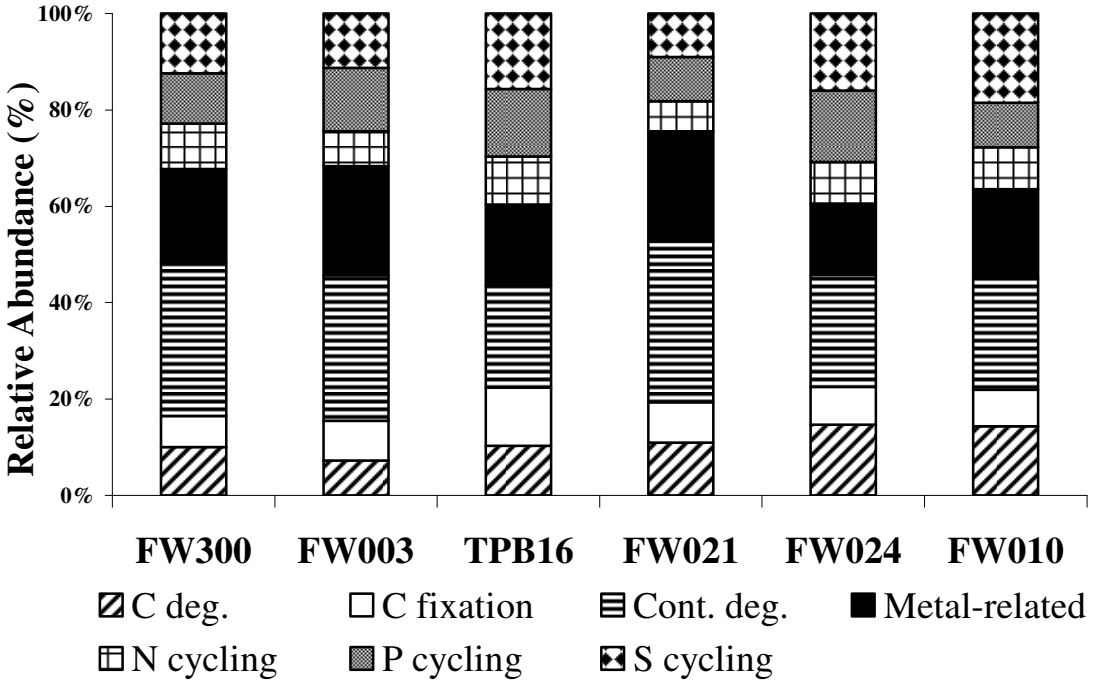
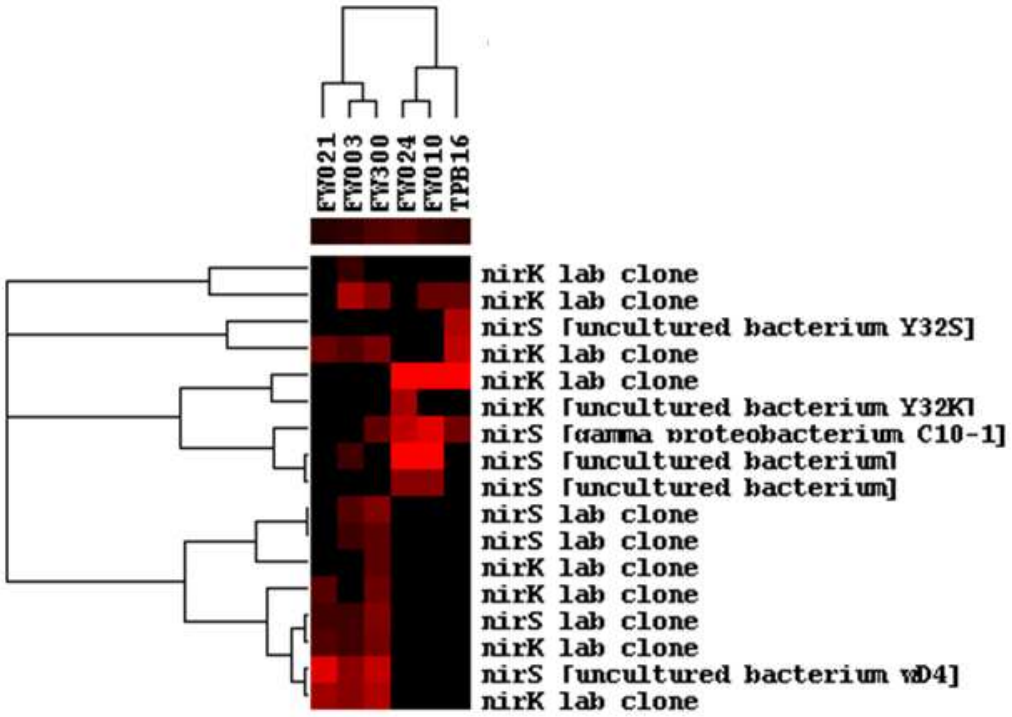
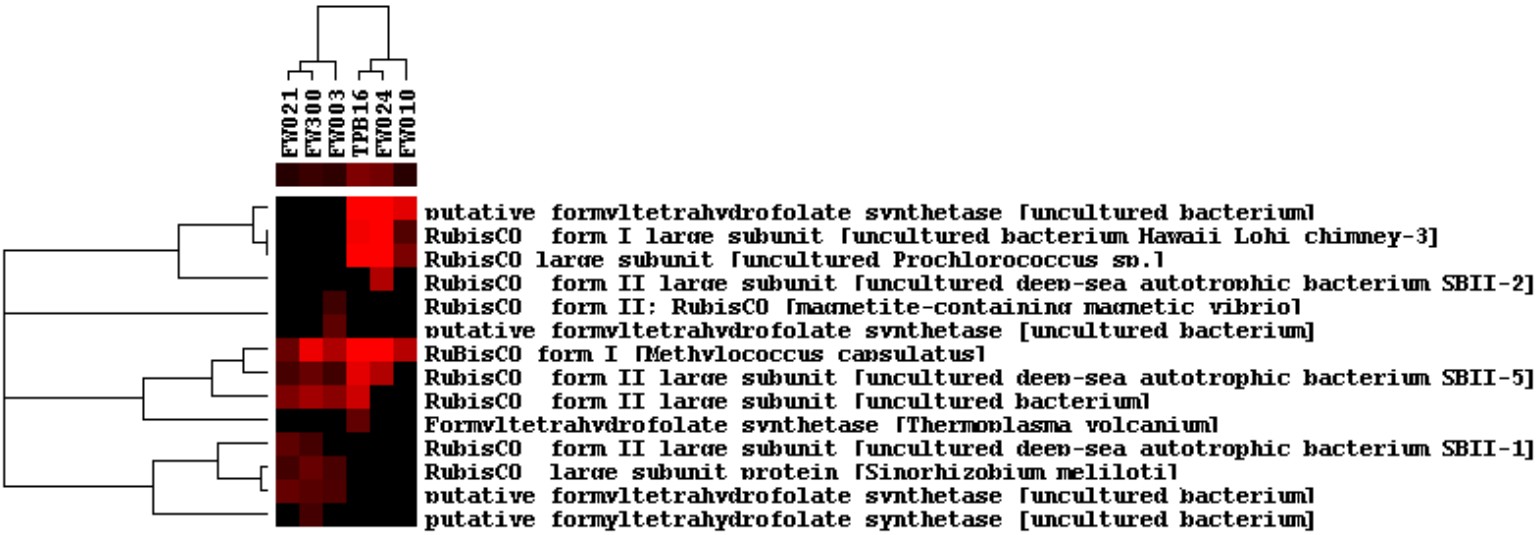


FIGURE. S2. Hierarchical cluster analysis of gene diversity based on hybridization signal intensities for selected functional gene groups. The color red indicates the level of signal intensity; black indicates no hybridization above the background level.

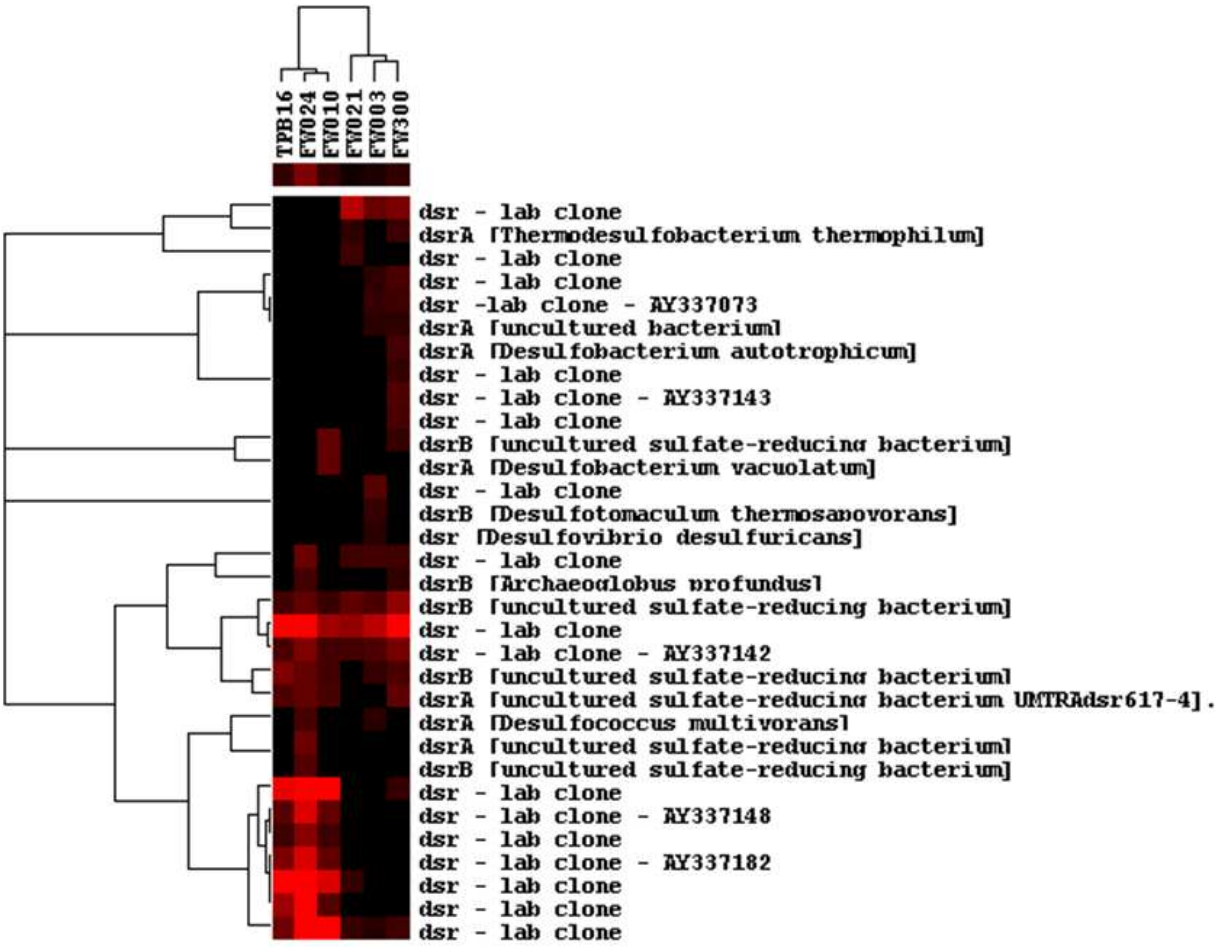
(A) Nitrite reductases (*nirK* and *nirS)*



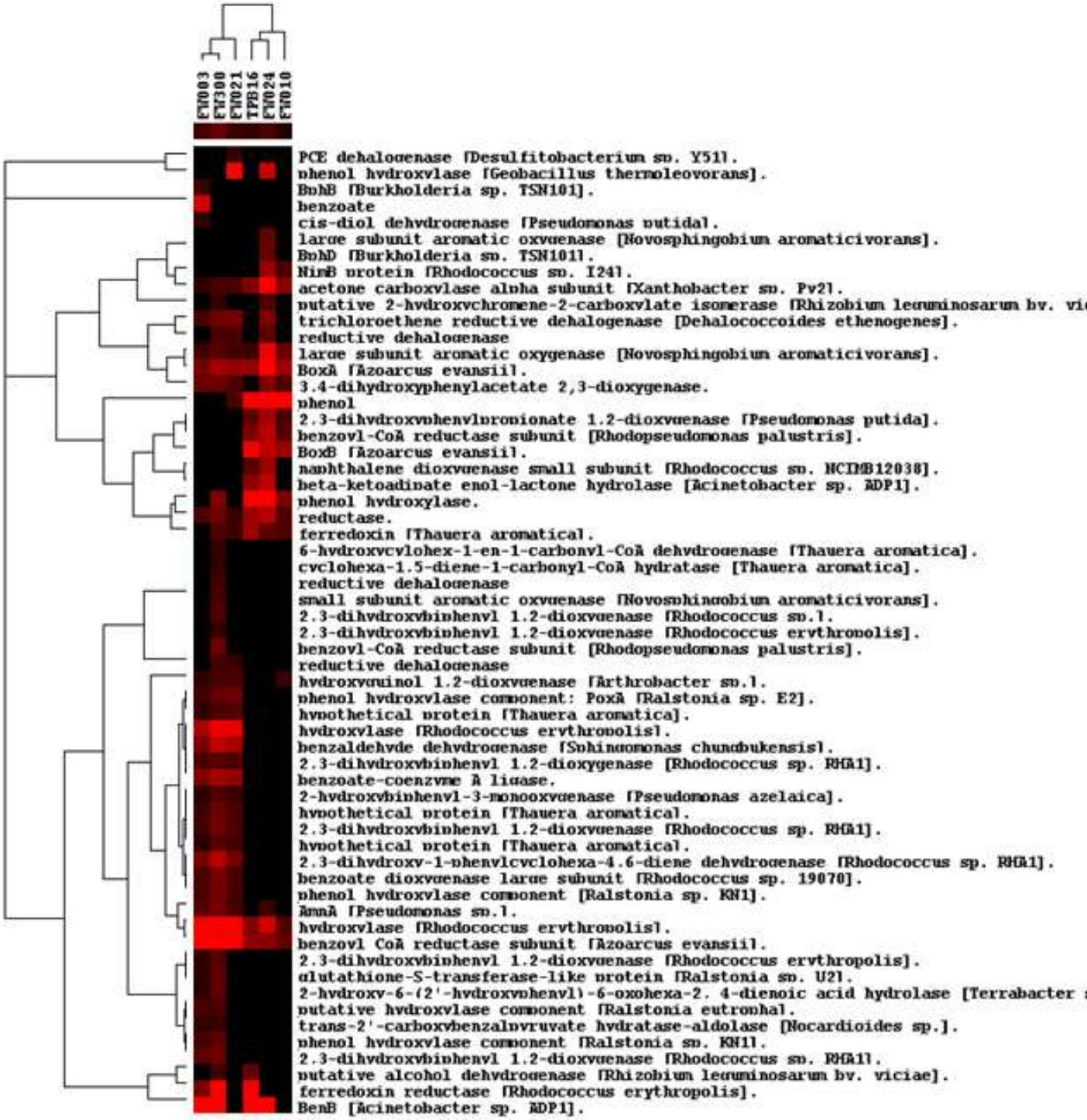
37 **(B)** Genes involved in carbon fixation



(C) Sulfite-reductases (*dsrAB*)

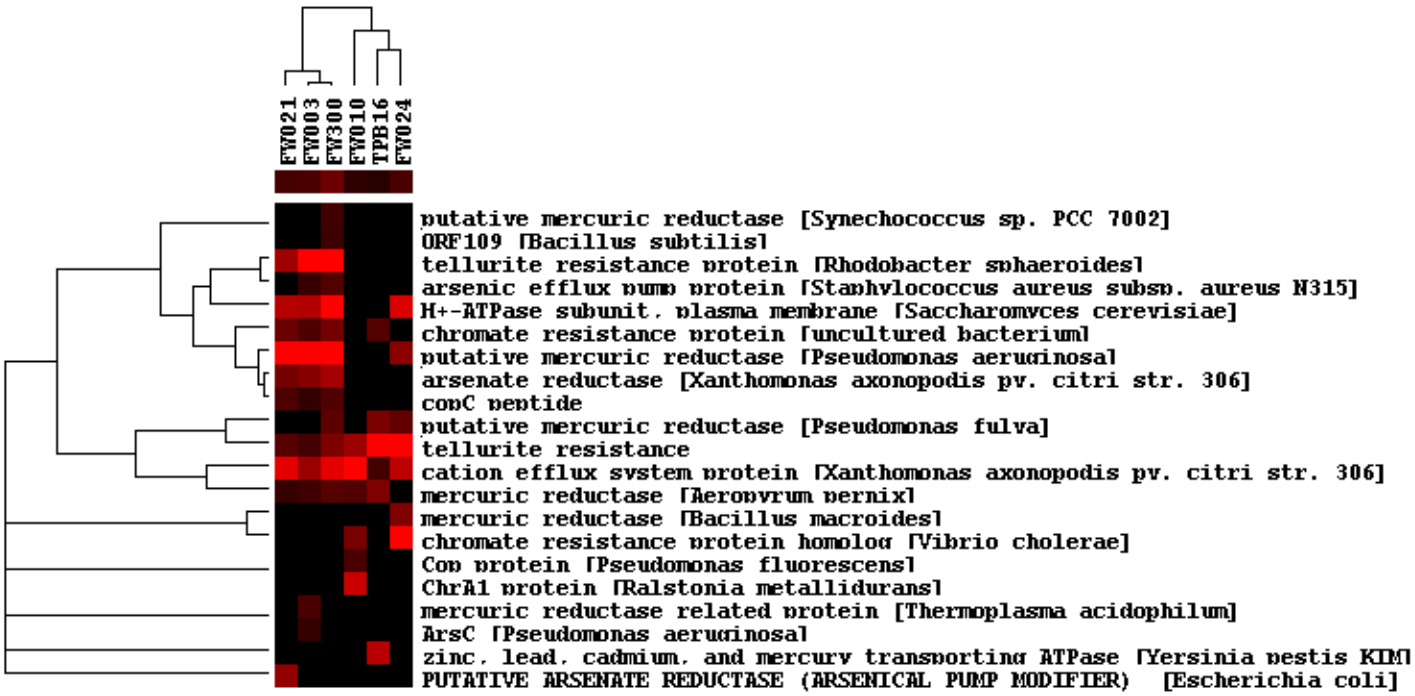


39 **(D)** Genes involved in organic contaminant degradation



40 **(E)** Genes involved in metal transport and resistance

41



42

43