Electronic Supplementary Material

Diversity and abundance of n-alkane degrading bacteria in the near surface soils of a Chinese onshore oil and gas field

Kewei Xu[1, 2*], Yuping Tang[2], Chun Ren[2], Kebin Zhao[2], Yongge Sun[1*]

1. Department of Earth Science, Zhejiang University, Hangzhou 310027, China.

2. Wuxi Research Institute of Petroleum Geology, Research Institute of Petroleum Exploration & Production, SINOPEC, Wuxi, Jiangsu 214151, China.

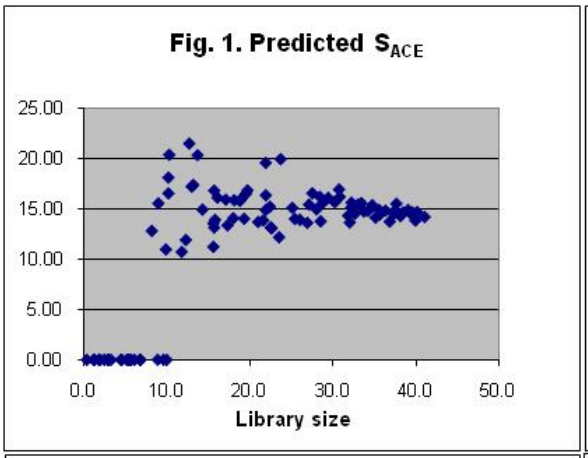
Table S1. Phylogenetic affiliation of deduced *alkB* sequences (182 amino acids) from different soil samples compared with public database

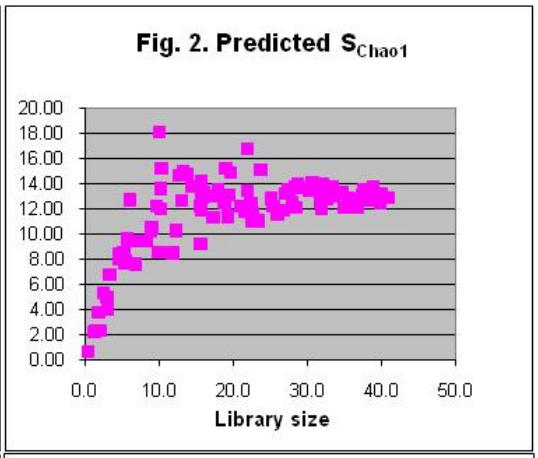
Phylogenetic cluster	Representative OTU	Number of clone[a]			t-RF length (bp)	Closest uncultured relative			Closest cultured relative		
		RS	OS	GS		Accession number	Similarity (%)	Habitat	Accession number	Similarity (%)	Taxon
Actinobacteria	SL-alkB-OTU-15	3		1	133	HM441092	96	HC contaminated soil	ZP08718102	81	*Mycobacterium colombiense*
	SL-alkB-OTU-14	4	1	3	133	HM441130	98	HC contaminated soil	JN616332	88	*Rhodococcus* sp. MS272a1
	SL-alkB-OTU-5	2			120	GQ261089	71	Fuel-contaminated soil	AJ301870	69	*Rhodococcus erythropolis*
Proteobacteria	SL-alkB-OTU-6	8	5	4	33	GU184277	70	Sandy soil	EU853327	79	*Salinisphaera shabanensis*
	SL-alkB-OTU-9		2		70	GU184277	82	Sandy soil	EU853342	80	*Alcanivorax dieselolei*
	SL-alkB-OTU-11	2	1	2	70	GQ261057	84	Fuel-contaminated soil			
	SL-alkB-OTU-3	1	7	9	70	JN106044	99	Oil field soil			
	SL-alkB-OTU-4		1		70	JN106044	99	Oil field soil			
	SL-alkB-OTU-18	1	2		74	JN986868	99	Fuel incubation	NC017506	97	*Marinobacter adhaerens*
	SL-alkB-OTU-10		1	2	74	JN986868	93	Fuel incubation	NC017506	92	*Marinobacter adhaerens*
	SL-alkB-OTU-16	1	4		74	JN986868	92	Fuel incubation	NC017506	92	*Marinobacter adhaerens*
	SL-alkB-OTU-12			2	74	JN986868	92	Fuel incubation	NC017506	89	*Marinobacter adhaerens*
	SL-alkB-OTU-8		5	4	74	JN986868	91	Fuel incubation	NC017506	91	*Marinobacter adhaerens*
	SL-alkB-OTU-7	3	13	10	74	JN986868	91	Fuel incubation	NC017506	90	*Marinobacter adhaerens*
	SL-alkB-OTU-17		3	2	74	GQ184417	98	Hydrocarbon seep			
Cluster A	SL-alkB-OTU-13	2	1	1	340	DQ288068	74	Grassland soil			
	SL-alkB-OTU-1	3			142	DQ288068	77	Grassland soil			
	SL-alkB-OTU-2	11	2	7	133	DQ288068	79	Grassland soil			

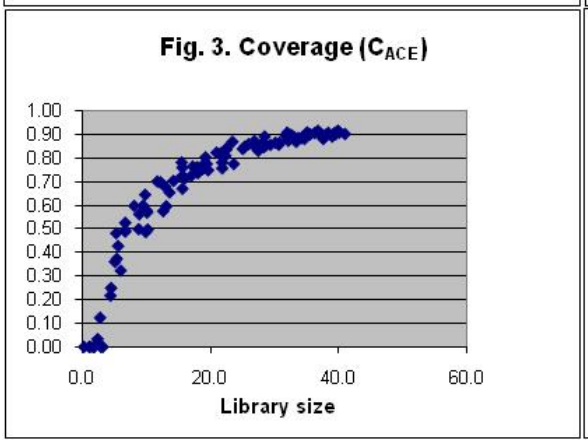
[a] OS, oilfield surface soil; GS, gasfield surface soil, RS, reference soil.

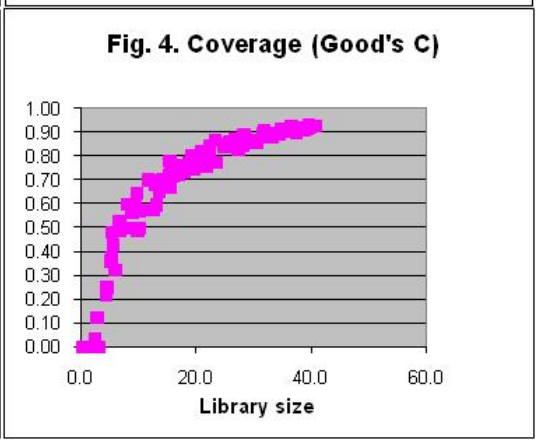
① The clone library of the reference soil:

DATA SUMMARY	
Number of clones in library	41
Number of phylotypes observed	12
Predicted value of S_{ACE}	14.45
Predicted value of S_{Chao1}	13.05
Observed phylotypes / predicted S_{ACE}	0.84
Observed phylotypes / predicted S_{Chao1}	0.95



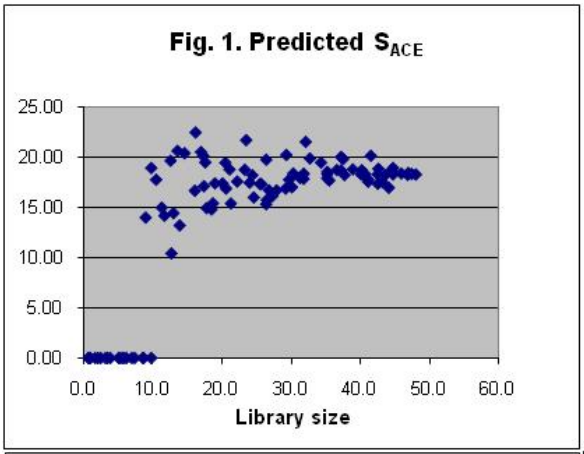


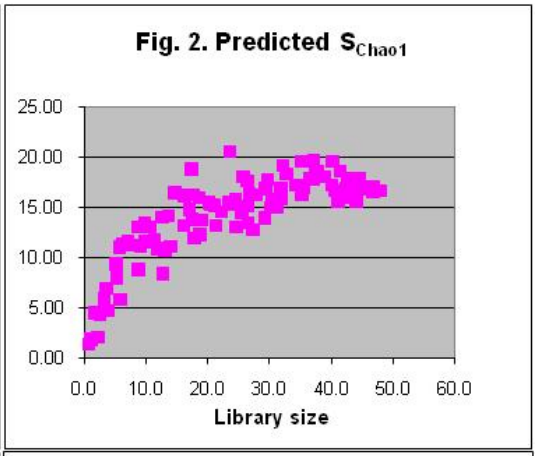


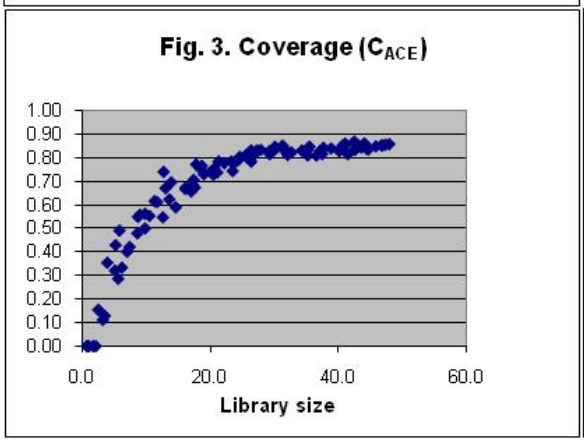


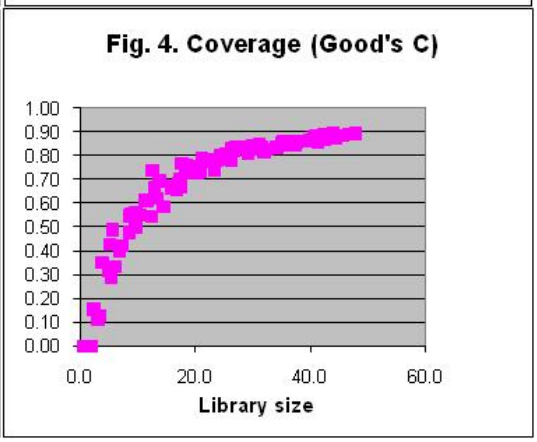
② The clone library of oilfield surface soil:

DATA SUMMARY	
Number of clones in library	48
Number of phylotypes observed	14
Predicted value of S_{ACE}	18.27
Predicted value of S_{Chao1}	16.77
Observed phylotypes / predicted S_{ACE}	0.83
Observed phylotypes / predicted S_{Chao1}	0.90









③ The clone library of gasfield surface soil:

DATA SUMMARY	
Number of clones in library	47
Number of phylotypes observed	12
Predicted value of S_{ACE}	13.39
Predicted value of S_{Chao1}	12.26
Observed phylotypes / predicted S_{ACE}	0.87
Observed phylotypes / predicted S_{Chao1}	0.95

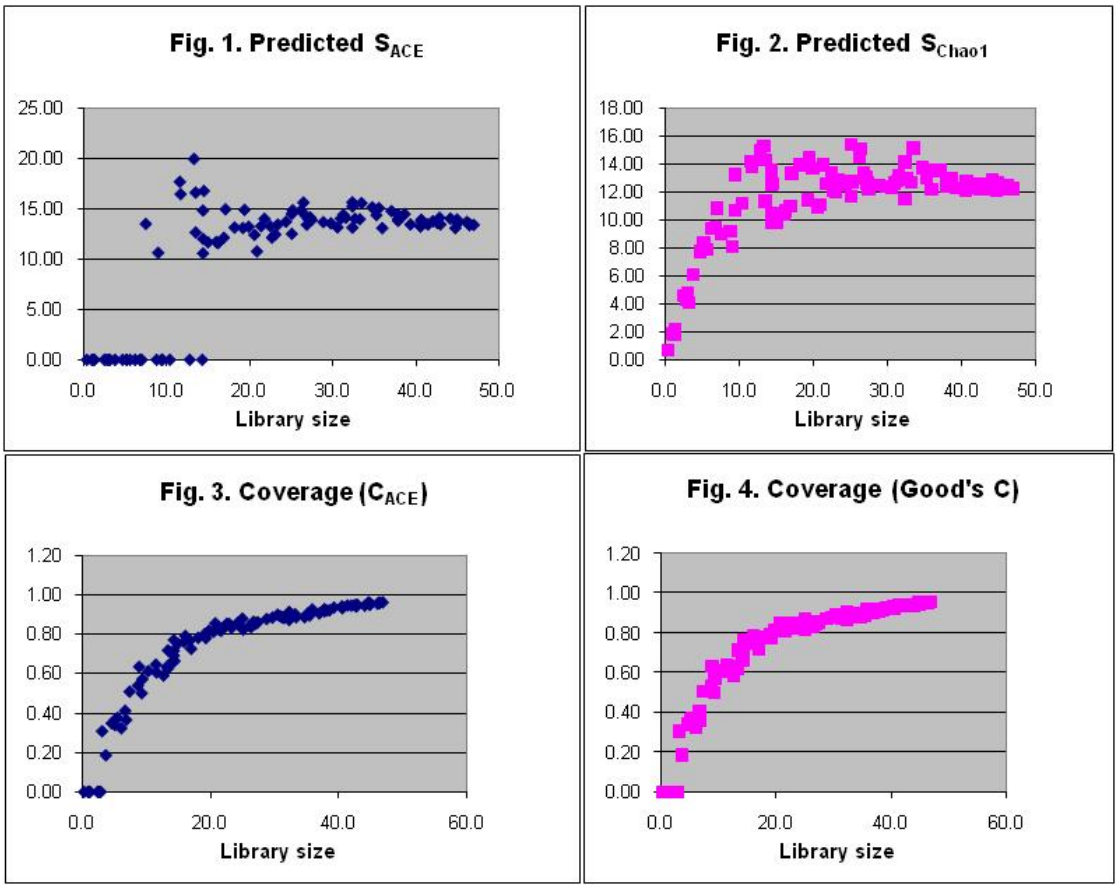


Figure S1. The final estimates of S_{ACE} and S_{Chao1}, and the fraction of estimated phylotype richness values actually recovered in the libraries of reference soil, oilfield surface soil and gasfield surface soil.