

## **Diverse alkane hydroxylase genes in microorganisms and environments**

Yong Nie<sup>1,2</sup>, Chang-Qiao Chi<sup>1</sup>, Hui Fang<sup>1</sup>, Jie-Liang Liang<sup>1</sup>, She-Lian Lu<sup>1</sup>, Guo-Li Lai<sup>1,2</sup>, Yue-Qin Tang<sup>1,†</sup>, Xiao-Lei Wu<sup>1,\*</sup>

<sup>1</sup> College of Engineering, Peking University, Beijing 100871, P. R. China

<sup>2</sup> Institute of Engineering (Baotou), College of Engineering, Peking University, Baotou 014030, China

\* Corresponding author: Xiao-Lei Wu

Department of Energy and Resources Engineering,

College of Engineering, Peking University

Beijing 100871, P. R. China

Tel/Fax: +86-10-62759047

Email: [xiaolei\\_wu@pku.edu.cn](mailto:xiaolei_wu@pku.edu.cn)

**Running title:** Diverse alkane hydroxylase genes

---

<sup>†</sup> Current address: College of Architecture and Environment, Sichuan University, Chengdu, 610065, China

**Table S1** Distribution of multiple-domain alkane hydroxylases

			Accession	Source
2-domain hydroxylase	AlkB-like	alkane	ZP_07715778.1	<i>Aeromicrobium marinum</i> DSM 15272
			ZP_05041651.1	<i>Alcanivorax</i> sp. DG881
			YP_006552789.1	<i>Amycolatopsis mediterranei</i> S699
			YP_003768535.1	<i>Amycolatopsis mediterranei</i> U32
			ZP_10049740.1	<i>Amycolatopsis</i> sp. ATCC 39116
			ZP_08022271.1	<i>Dietzia cinnamea</i> P4
			ZP_05061580.1	gamma proteobacterium HTCC5015
			ZP_09213880.1	<i>Gordonia amarae</i> NBRC 15530
			YP_003275257.1	<i>Gordonia bronchialis</i> DSM 43247
			ZP_09787302.1	<i>Gordonia otitidis</i> NBRC 100426
			ZP_09795945.1	<i>Gordonia sputi</i> NBRC 100414
			ZP_00996652.1	<i>Janibacter</i> sp. HTCC2649
			ZP_08195602.1	Nocardiodaceae bacterium Broad-1
			ZP_08198697.1	Nocardiodaceae bacterium Broad-1
			YP_921354.1	<i>Nocardioides</i> sp. JS614
			ZP_08122407.1	<i>Pseudonocardia</i> sp. P1
			ZP_09743302.1	<i>Saccharomonospora marina</i> XMU15
			ZP_07282558.1	<i>Streptomyces</i> sp. AA4
ZP_07282765.1	<i>Streptomyces</i> sp. AA4			
3-domain hydroxylase	AlkB-like	alkane	YP_001963774.1	<i>Leptospira biflexa</i> serovar Patoc strain 'Patoc 1 (Ames)'
			YP_001840157.1	<i>Leptospira biflexa</i> serovar Patoc strain 'Patoc 1 (Paris)'
			ZP_01913733.1	<i>Limnobacter</i> sp. MED105
			YP_973211.1	<i>Polaromonas naphthalenivorans</i> CJ2
3-domain hydroxylase	CYP153	alkane	ZP_09213539.1	<i>Gordonia araii</i> NBRC 100433
			YP_005284867.1	<i>Gordonia polyisoprenivorans</i> VH2
			ZP_09268583.1	<i>Gordonia polyisoprenivorans</i> NBRC 16320

**Table S2** Genomes containing multiple copies of *alkB* genes

Phylum	Genus	Strain	No.of <i>alkB</i>
Proteobacteria (Alphaproteobacteria)	<i>Citricella</i>	<i>Citricella</i> sp. 357	2
	<i>Jannaschia</i>	<i>Jannaschia</i> sp. CCS1	2
	<i>Maritimibacter</i>	<i>Maritimibacter alkaliphilus</i> HTCC2654	2
	<i>Novosphingobium</i>	<i>Novosphingobium</i> sp. PP1Y(-2)	2
	<i>Oceanibulbus</i>	<i>Oceanibulbus indolifex</i> HEL-45	2
	<i>Oceanicola</i>	<i>Oceanicola batsensis</i> HTCC2597	2
		<i>Oceanicola granulosus</i> HTCC2516	2
	<i>Rhodobacter</i>	<i>Rhodobacter</i> sp. SW 2	2
		Rhodobacterales bacterium HTCC2150	2
	<i>Roseobacter</i>	<i>Roseobacter</i> sp. CCS2	2
		<i>Roseobacter</i> sp. GAI101	2
		<i>Roseobacter</i> sp. SK209-2-6	3
	<i>Roseovarius</i>	<i>Roseovarius nubinhibens</i> ISM	2
		<i>Roseovarius</i> sp. 217	2
		<i>Roseovarius</i> sp. TM1035	2
	<i>Ruegeria</i>	<i>Ruegeria pomeroyi</i> DSS-3	2
<i>Sagittula</i>	<i>Sagittula stellata</i> E-37	3	
<i>Sulfitobacter</i>	<i>Sulfitobacter</i> sp. EE-36	2	
	<i>Sulfitobacter</i> sp. NAS-14.1	2	
Proteobacteria (Betaproteobacteria)	<i>Burkholderia</i>	<i>Burkholderia</i> sp. 383	2
		<i>Burkholderia cepacia</i> GG4	2
		<i>Burkholderia</i> sp. Ch1-1	2
	<i>Limnobacter</i>	<i>Limnobacter</i> sp. MED105	3
Proteobacteria (Gammaproteobacteria)	<i>Acinetobacter</i>	<i>Acinetobacter calcoaceticus</i> RUH2202	2
		<i>Acinetobacter haemolyticus</i> ATCC 19194	2
		<i>Acinetobacter junii</i> SH205	2
		<i>Acinetobacter oleivorans</i> DR1	2
		<i>Acinetobacter</i> sp. ATCC 27244	2
		<i>Acinetobacter</i> sp. NBRC 100985	2
		<i>Acinetobacter</i> sp. P8-3-8	2
	<i>Alcanivorax</i>	<i>Alcanivorax borkumensis</i> SK2	2
		<i>Alcanivorax</i> sp. DG881	2
	<i>Legionella</i>	<i>Legionella pneumophila</i> subsp. pneumophila	2
	<i>Marinobacter</i>	<i>Marinobacter adhaerens</i> HP15	2
		<i>Marinobacter aquaeolei</i> VT8	3
		<i>Marinobacter hydrocarbonoclasticus</i> ATCC 49840	2
		<i>Marinobacter manganoxydans</i> Mni7-9	2

	<i>Pseudomonas</i>	<i>Pseudomonas aeruginosa</i> 2192	2
		<i>Pseudomonas aeruginosa</i> 39016	2
		<i>Pseudomonas aeruginosa</i> C3719	2
		<i>Pseudomonas aeruginosa</i> DK2	2
		<i>Pseudomonas aeruginosa</i> LESB58	2
		<i>Pseudomonas aeruginosa</i> M18	
		<i>Pseudomonas aeruginosa</i> NCGM2.S1	2
		<i>Pseudomonas aeruginosa</i> PA7	2
		<i>Pseudomonas aeruginosa</i> PAb1	2
		<i>Pseudomonas aeruginosa</i> PACS2	2
		<i>Pseudomonas aeruginosa</i> PAO1	2
		<i>Pseudomonas aeruginosa</i> UCBPP-PA14	2
		<i>Pseudomonas mendocina</i> ymp	3
		<i>Pseudomonas</i> sp. 2_1_26	2
Unclassified Gammaproteobacteria	<i>gamma proteobacterium</i> BDW918	3	
Actinobacteria	<i>Aeromicrobium</i>	<i>Aeromicrobium marinum</i> DSM 15272	2
	<i>Amycolicicoccus</i>	<i>Amycolicicoccus subflavus</i> DQS3-9A1	3
	<i>Gordonia</i>	<i>Gordonia amarae</i> NBRC 15530	2
	<i>Mycobacterium</i>	<i>Mycobacterium chubuense</i> NBB4	2
		<i>Mycobacterium gilvum</i> PYR-GCK	2
		<i>Mycobacterium gilvum</i> Spyr1	2
		<i>Mycobacterium kansasii</i> ATCC 12478	2
		<i>Mycobacterium marinum</i> M	2
		<i>Mycobacterium</i> sp. JDM601	2
		<i>Mycobacterium thermoresistibile</i> ATCC 19527	2
		<i>Mycobacterium tusciae</i> JS617	2
		<i>Mycobacterium vanbaalenii</i> PYR-1	2
	<i>Nocardia</i>	<i>Nocardia cyriacigeorgica</i> GUH-2	4
		<i>Nocardia farcinica</i> IFM 10152	3
	Unclassified Nocardioideae	Nocardioideae bacterium Broad-1	2
	<i>Rhodococcus</i>	<i>Rhodococcus erythropolis</i> PR4	4
		<i>Rhodococcus erythropolis</i> SK121	6
<i>Rhodococcus opacus</i> B4		2	
<i>Streptomyces</i>	<i>Streptomyces</i> sp. AA4	2	

**Table S3** Genomes containing multiple copies of CYP153 genes

Phylum	Genus	Strain	No. of CYP153
Proteobacteria (Alphaproteobacteria)	<i>Bradyrhizobium</i>	<i>Bradyrhizobium</i> sp. BTAi1	2
		<i>Bradyrhizobium</i> sp. ORS 278	2
		<i>Bradyrhizobium</i> sp. ORS 285	2
		<i>Bradyrhizobium</i> sp. ORS 375	2
		<i>Bradyrhizobium</i> sp. WSM471	3
		<i>Bradyrhizobium</i> sp. WSM1253	3
		<i>Bradyrhizobium</i> sp. S23321	2
		<i>Bradyrhizobium japonicum</i> USDA 110	3
		<i>Bradyrhizobium japonicum</i> USDA 6	3
		<i>Bradyrhizobium</i> sp. STM 3809	2
	<i>Bradyrhizobium</i> sp. STM 3843	2	
	<i>Caulobacter</i>	<i>Caulobacter</i> sp. K31	2
	<i>Erythrobacter</i>	<i>Erythrobacter litoralis</i> HTCC2594	3
		<i>Erythrobacter</i> sp. SD-21	2
<i>Novosphingobium</i>	<i>Novosphingobium aromaticivorans</i> DSM 12444	2	
	<i>Novosphingobium pentaromativorans</i> US6-1	2	
<i>Parvibaculum</i>	<i>Parvibaculum lavamentivorans</i> DS-1	6	
Proteobacteria (Gammaproteobacteria)	<i>Alcanivorax</i>	<i>Alcanivorax borkumensis</i> SK2	2
		<i>Alcanivorax</i> sp. DG881	2
	<i>Marinobacter</i>	<i>Marinobacter hydrocarbonoclasticus</i> ATCC 49840	2
	<i>Congregibacter</i>	<i>Congregibacter litoralis</i> KT71	2
	Unclassified Gammaproteobacteria	gamma proteobacterium BDW918	2
		gamma proteobacterium HIMB55	3
		gamma proteobacterium NOR5-3	2
		marine gamma proteobacterium HTCC2080	3
		marine gamma proteobacterium HTCC2143	4
	marine gamma proteobacterium HTCC2148	2	
Actinobacteria	<i>Aeromicrobium</i>	<i>Aeromicrobium marinum</i> DSM 15272	2
	<i>Amycolicococcus</i>	<i>Amycolicococcus subflavus</i> DQS3-9A1	2
	<i>Dietzia</i>	<i>Dietzia cinnamea</i> P4	2
	<i>Gordonia</i>	<i>Gordonia polyisoprenivorans</i> NBRC 16320	2
	<i>Rhodococcus</i>	<i>Rhodococcus erythropolis</i> PR4	2

**Table S4** Genomes containing both *alkB* and CYP153 genes

Phylum	Genus	Strain	No. of <i>alkB</i>	No. of CYP153
Actinobacteria	<i>Aeromicrobium</i>	<i>Aeromicrobium marinum</i> DSM 15272	2	2
	<i>Amycolicococcus</i>	<i>Amycolicococcus subflavus</i> DQS3-9A1	3	2
	<i>Dietzia</i>	<i>Dietzia cinnamea</i> P4	1	2
	<i>Gordonia</i>	<i>Gordonia araii</i> NBRC 100433	1	1
		<i>Gordonia neofelifaecis</i> NRRL B-59395	1	1
		<i>Gordonia</i> sp. KTR9	1	1
		<i>Gordonia terrae</i> NBRC 100016	1	1
	<i>Mycobacterium</i>	<i>Mycobacterium abscessus</i> ATCC 19977	1	1
		<i>Mycobacterium chubuense</i> NBB4	2	1
		<i>Mycobacterium gilvum</i> PYR-GCK	2	1
		<i>Mycobacterium intracellulare</i> MOTT-02	1	1
		<i>Mycobacterium marinum</i> M	2	1
		<i>Mycobacterium massiliense</i> str. GO 06	1	1
		<i>Mycobacterium parascrofulaceum</i> ATCC BAA-614	1	2
		<i>Mycobacterium rhodesiae</i> NBB3	1	1
		<i>Mycobacterium</i> sp. JLS	1	1
		<i>Mycobacterium</i> sp. KMS	1	1
		<i>Mycobacterium</i> sp. MCS	1	1
	<i>Nocardia</i>	<i>Nocardia cyriacigeorgica</i> GUH-2	2	2
		Nocardioideae bacterium Broad-1	2	1
<i>Patulibacter</i>	<i>Patulibacter</i> sp. II1	1	1	
<i>Rhodococcus</i>	<i>Rhodococcus erythropolis</i> PR4	4	2	
	<i>Rhodococcus erythropolis</i> SK121	6	1	
<i>Saccharomonospora</i>	<i>Saccharomonospora marina</i> XMU15	1	1	
Proteobacteria (Alphaproteobacteria)	<i>Caulobacter</i>	<i>Caulobacter</i> sp. K31	1	2
	<i>Oceanicola</i>	<i>Oceanicola batsensis</i> HTCC2597	2	1
Proteobacteria (Betaproteobacteria)	<i>Burkholderia</i>	<i>Burkholderia xenovorans</i> LB400	1	1
	<i>Limnobacter</i>	<i>Limnobacter</i> sp. MED105	3	1
	<i>Acinetobacter</i>	<i>Acinetobacter</i> sp. NBRC 100985	2	1
Proteobacteria (Gammaproteobacteria)	<i>Alcanivorax</i>	<i>Alcanivorax borkumensis</i> SK2	2	2
		<i>Alcanivorax</i> sp. DG881	2	2
		gamma proteobacterium BDW918	3	2
		gamma proteobacterium HdN1	1	1
		gamma proteobacterium NOR5-3	1	2

	<i>Marinobacter</i>	<i>Marinobacter adhaerens</i> HP15	2	1
		<i>Marinobacter aquaeolei</i> VT8	3	1
		<i>Marinobacter hydrocarbonoclasticus</i> ATCC 49840	2	2
		<i>Marinobacter manganoxydans</i> MnI7-9	2	1

**Table S5** Bacteria containing *alkB* or CYP153 genes that were not usually considered to be alkane degraders

Phylum/Class	Genus
Alphaproteobacteria	<i>Acidiphilium, Dinoroseobacter, Parvularcula, Pseudovibrio, Rhodospirillum, Ruegeria, Jannaschia, Sagittula, Loktanella, Thalassiobium, Ahrensia, Phaeobacter, Hoeflea, Octadecabacter, Pelagibaca, Hirschia, Citromicrobium</i>
Betaproteobacteria	<i>Limnobacter, Methylibium</i>
Gammaproteobacteria	<i>Glaciacola, Hahella, Congregibacter, Dickeya</i>
Deltaproteobacteria	<i>Plesiocystis</i>
Actinobacteria	<i>Actinoplanes, Catenulispora, Conexibacter, Segniliparus, Thermomonospora, Tsukamurella</i>
Bacteroidetes	<i>Belliella, Haliscomenobacter, Maribacter, Marivirga, Owenweeksia, Runella</i>
Spirochaetes	<i>Leptospira, Turneriella</i>
Planctomycetes	<i>Planctomyces</i>



**Table S6** The reference sequences of AlkB and CYP153 with function identified used in this study

AlkB			CYP153		
Accession	Source	Reference	Accession	Source	Reference
AB049411	<i>Acinetobacter</i> sp. Strain M-1	1	AJ311718	<i>Acinetobacter</i> sp. EB104	2
AJ002316	<i>Acinetobacter</i> sp. Strain ADP1	3	AAS80270	<i>Alcanivorax borkumensis</i> SK2	4
JF747237	<i>Alcanivorax hongdengensis</i> A-11-3	5	AJ844908	<i>Alcanivorax borkumensis</i> AP1	4
GQ980249	<i>Alcanivorax dieselolei</i> B-5	6	AJ844909	<i>Alcanivorax borkumensis</i> AP1	4
GQ980251	<i>Alcanivorax dieselolei</i> B-5	6	NP_768493	<i>Bradyrhizobium japonicum</i> USDA 110	4
FJ905614	<i>Alcanivorax</i> sp. strain 2B5	7	NP_773882	<i>Bradyrhizobium japonicum</i> USDA 110	4
AJ295164	<i>Alcanivorax borkumensis</i> Ap1	8	NP_773883	<i>Bradyrhizobium japonicum</i> USDA 110	4
FJ744758	<i>Dietzia</i> sp. E1	9	NP_418882	<i>Caulobacter crescentus</i> CB15	4
HQ850582	<i>Dietzia</i> sp. DQ12-45-1b	10	JQ410793	<i>Dietzia</i> sp. DQ12-45-1b	11
HQ026811	<i>Gordonia</i> sp. SoCg	12	AJ783967	<i>Mycobacterium</i> sp. HXN-1500	4
AB250939	<i>Mycobacterium</i> sp. TY-6	13	ZP_00094181	<i>Novosphingobium aromaticivorans</i> Saro 178	4
O05895	<i>Mycobacterium tuberculosis</i> H37Rv	8	ZP_00096754	<i>Novosphingobium aromaticivorans</i> Saro 178	4
AF350429	<i>Nocardioides</i> sp. strain CF8	14	ZP_11192	<i>Rhodopseudomonas palustris</i> CGA009	4
AE004685	<i>Pseudomonas aeruginosa</i> PAO1	8	AJ850057	<i>Sphingomonas</i> sp. strain HXN-200	4
AE004581	<i>Pseudomonas aeruginosa</i> PAO1	8	AJ850057	<i>Sphingomonas</i> sp. strain HXN-200	4
AJ245436	<i>Pseudomonas putida</i> GPo1	8	AJ850058	<i>Sphingomonas</i> sp. strain HXN-200	4
AJ233397	<i>Pseudomonas putida</i> P1	15	AJ850059	<i>Sphingomonas</i> sp. strain HXN-200	4
AJ009587	<i>Pseudomonas rugosa</i> NRRL B-2295	8	AJ850060	<i>Sphingomonas</i> sp. strain HXN-200	4
AJ009579	<i>Pseudomonas fluorescens</i> CHA0	8			
FJ705347	<i>Rhodococcus</i> sp. SP2B	16			
AJ009586	<i>Rhodococcus erythropolis</i> NRRL B-16531	17			

---

	<i>Rhodococcus</i>	
AJ297269	<i>erythropolis</i> NRRL B-16531	17
	<i>Rhodococcus</i>	
AJ301876	<i>erythropolis</i> NRRL B-16531	17
	<i>Rhodococcus</i>	
AJ301877	<i>erythropolis</i> NRRL B-16531	17
AF388179	<i>Rhodococcus</i> sp. Q15	17
AF388180	<i>Rhodococcus</i> sp. Q15	17
AF388181	<i>Rhodococcus</i> sp. Q15	17
AF388182	<i>Rhodococcus</i> sp. Q15	17

---

**Table S7** The GenBank accession numbers of alkane hydroxylase genes found within microbial genomes in this study

<b>AlkB</b>	<p> NP_250216.1, NP_251264.1, NP_217769.1, NP_337878.1, NP_856925.1, NP_962298.1,  NP_969039.1, YP_046098.1, YP_096989.1, YP_108945.1, YP_102417.1, YP_114083.1,  YP_119532.1, YP_120829.1, YP_120833.1, YP_128249.1, YP_125370.1, YP_166583.1,  YP_168645.1, YP_251715.1, YP_260041.1, YP_334185.1, YP_351510.1, YP_371980.1,  YP_368326.1, YP_435857.1, YP_442346.1, YP_480358.1, YP_508336.1, YP_509155.1,  YP_554098.1, YP_614640.1, YP_620386.1, YP_638501.1, YP_691842.1, YP_694427.1,  YP_702497.1, YP_712218.1, YP_772734.1, YP_791728.1, YP_790621.1, YP_834626.1,  YP_883357.1, YP_886209.1, YP_906529.1, YP_915886.1, YP_921354.1, YP_937352.1,  YP_952571.1, YP_953908.1, YP_957728.1, YP_957898.1, YP_960105.1, YP_973211.1,  YP_993683.1, YP_979365.1, YP_001023605.1, YP_001028856.1, YP_001069662.1,  YP_001059702.1, YP_001081233.1, YP_001066986.1, YP_001042010.1, YP_001084670.1,  YP_001118743.1, YP_001134633.1, YP_001135977.1, YP_001185946.1, YP_001188029.1,  YP_001188237.1, YP_001234392.1, YP_001252544.1, YP_001280943.1, YP_001284638.1,  YP_001289208.1, YP_001348003.1, YP_001349162.1, YP_001506534.1, YP_001531377.1,  YP_001672212.1, YP_001704327.1, YP_001707231.1, YP_001713880.1, YP_001764240.1,  YP_001807560.1, YP_001840157.1, YP_001849604.1, YP_001851790.1, YP_001846325.1,  YP_001889129.1, YP_001892637.1, YP_001963774.1, YP_002232139.1, YP_002296646.1,  YP_002319225.1, YP_002325761.1, YP_002440314.1, YP_002441386.1, YP_002527219.1,  YP_002646323.1, YP_002764193.1, YP_002765609.1, YP_002765850.1, YP_002768905.1,  YP_002776786.1, YP_002779449.1, YP_002873097.1, YP_002897546.1, YP_002912304.1,  YP_002984139.1, YP_003033295.1, YP_003112137.1, YP_003275257.1, YP_003298195.1,  YP_003397515.1, YP_003455842.1, YP_003578527.1, YP_003620379.1, YP_003647687.1,  YP_003658078.1, YP_003732372.1, YP_003732938.1, YP_003768535.1, YP_003837040.1,  YP_003809668.1, YP_003855036.1, YP_003862088.1, YP_004008018.1, YP_004016090.1,  YP_004052990.1, YP_004077166.1, YP_004078475.1, YP_004084103.1, YP_004283531.1,  YP_004379711.1, YP_004361423.1, YP_004448035.1, YP_004493362.1, YP_004493370.1,  YP_004495037.1, YP_004499590.1, YP_004504542.1, YP_004524245.1, YP_004525252.1,  YP_004654946.1, YP_004724904.1, YP_004746682.1, YP_004758559.1, YP_004871798.1,  YP_004930483.1, YP_005082016.1, YP_004988211.1, YP_004995277.1, YP_004998267.1,  YP_005101791.1, YP_005172733.1, YP_005187321.1, YP_005264130.1, YP_005264231.1,  YP_005265730.1, YP_005265733.1, YP_005339605.1, YP_005344758.1, YP_005350034.1,  YP_005361799.1, YP_005462597.1, YP_005514696.1, YP_005525783.1, YP_005798769.1,  YP_005886266.1, YP_005886584.1, YP_005910795.1, YP_005914423.1, YP_005918340.1,  YP_005975004.1, YP_005976106.1, YP_005980655.1, YP_005981818.1, YP_006024004.1,  YP_006267869.1, YP_006274025.1, YP_006289939.1, YP_005430622.1, YP_005430734.1,  YP_006307744.1, YP_006323977.1, YP_006331968.1, YP_006408858.1, YP_005309505.1,  YP_006439912.1, YP_006451832.1, YP_006453223.1, YP_006474778.1, YP_006475911.1,  YP_006482457.1, YP_006483597.1, ZP_00439582.1, ZP_00948812.1, ZP_00954267.1,  ZP_00955262.1, ZP_00957064.1, ZP_00959368.1, ZP_00959618.1, ZP_00964204.1,  ZP_00996652.1, ZP_00998265.1, ZP_01000906.1, ZP_01002389.1, ZP_01011524.1,  ZP_01015838.1, ZP_01035050.1, ZP_01037964.1, ZP_01049702.1, ZP_01052921.1,  ZP_01055480.1, ZP_01155421.1, ZP_01157350.1, ZP_01201250.1, ZP_01225325.1, </p>
-------------	--

ZP\_01307000.1, ZP\_01364874.1, ZP\_01365940.1, ZP\_01442436.1, ZP\_01689499.1,  
ZP\_01738706.1, ZP\_01740604.1, ZP\_01741446.1, ZP\_01744806.1, ZP\_01746661.1,  
ZP\_01748311.1, ZP\_01748906.1, ZP\_01751972.1, ZP\_01753800.1, ZP\_01754375.1,  
ZP\_01755711.1, ZP\_01764629.1, ZP\_01878984.1, ZP\_01880441.1, ZP\_01895378.1,  
ZP\_01900149.1, ZP\_01901011.1, ZP\_01912185.1, ZP\_01913733.1, ZP\_01914093.1,  
ZP\_01916228.1, ZP\_02152773.1, ZP\_02153440.1, ZP\_02161687.1, ZP\_02166109.1,  
ZP\_02270228.1, ZP\_02356342.1, ZP\_02363472.1, ZP\_02374120.1, ZP\_02380481.1,  
ZP\_02387962.1, ZP\_02403680.1, ZP\_02412212.1, ZP\_02448321.1, ZP\_02456517.1,  
ZP\_02464124.1, ZP\_02472061.1, ZP\_02482535.1, ZP\_02498859.1, ZP\_02506862.1,  
ZP\_02889567.1, ZP\_02904755.1, ZP\_03451903.1, ZP\_03574223.1, ZP\_03579376.1,  
ZP\_03586616.1, ZP\_03700804.1, ZP\_03702922.1, ZP\_03795488.1, ZP\_03823182.1,  
ZP\_03824539.1, ZP\_04383196.1, ZP\_04383566.1, ZP\_04385381.1, ZP\_04386125.1,  
ZP\_04387164.1, ZP\_04388098.1, ZP\_04608296.1, ZP\_04661203.1, ZP\_04815249.1,  
ZP\_04884216.1, ZP\_04890888.1, ZP\_04897834.1, ZP\_04902275.1, ZP\_04905683.1,  
ZP\_04911982.1, ZP\_04927188.1, ZP\_04927943.1, ZP\_04928883.1, ZP\_04933196.1,  
ZP\_04934328.1, ZP\_04939380.1, ZP\_04950987.1, ZP\_04964412.1, ZP\_04974059.1,  
ZP\_04981921.1, ZP\_05041651.1, ZP\_05042146.1, ZP\_05051648.1, ZP\_05061580.1,  
ZP\_05064795.1, ZP\_05079274.1, ZP\_05083049.1, ZP\_05090796.1, ZP\_05099213.1,  
ZP\_05101918.1, ZP\_05122263.1, ZP\_05128075.1, ZP\_05142802.1, ZP\_05218167.1,  
ZP\_05342453.1, ZP\_05361594.1, ZP\_05741459.1, ZP\_05786269.1, ZP\_05826167.1,  
ZP\_05827357.1, ZP\_05842058.1, ZP\_05845010.1, ZP\_05847237.1, ZP\_06056754.1,  
ZP\_06057458.1, ZP\_06061712.1, ZP\_06065934.1, ZP\_06066074.1, ZP\_06069784.1,  
ZP\_06072466.1, ZP\_06188150.1, ZP\_06191942.1, ZP\_06417258.1, ZP\_06434577.1,  
ZP\_06438684.1, ZP\_06451706.1, ZP\_06456198.1, ZP\_06506448.1, ZP\_06511324.1,  
ZP\_06514753.1, ZP\_06518770.1, ZP\_06522818.1, ZP\_06690500.1, ZP\_06725872.1,  
ZP\_06727425.1, ZP\_06799499.1, ZP\_06850622.1, ZP\_06878434.1, ZP\_06879583.1,  
ZP\_06962020.1, ZP\_07014137.1, ZP\_07226035.1, ZP\_07238173.1, ZP\_07239643.1,  
ZP\_07282558.1, ZP\_07282765.1, ZP\_07375042.1, ZP\_07415918.1, ZP\_07420639.1,  
ZP\_07424446.1, ZP\_07428814.1, ZP\_07433315.1, ZP\_07437513.1, ZP\_07482009.1,  
ZP\_07486361.1, ZP\_07490579.1, ZP\_07495126.1, ZP\_07668175.1, ZP\_07668473.1,  
ZP\_07673680.1, ZP\_07715778.1, ZP\_07717360.1, ZP\_07775830.1, ZP\_07792758.1,  
ZP\_07793670.1, ZP\_07817114.1, ZP\_08022271.1, ZP\_08122407.1, ZP\_08155372.1,  
ZP\_08195602.1, ZP\_08198697.1, ZP\_08205256.1, ZP\_08276444.1, ZP\_08330953.1,  
ZP\_08414255.1, ZP\_08433663.1, ZP\_08436667.1, ZP\_08442348.1, ZP\_08461977.1,  
ZP\_08553310.1, ZP\_08634611.1, ZP\_08666479.1, ZP\_04750514.1, ZP\_04751264.1,  
ZP\_02147281.1, ZP\_08718102.1, YP\_006507153.1, YP\_006510216.1, YP\_006521713.1,  
YP\_006552789.1, YP\_006574713.1, YP\_006564389.1, YP\_006616353.1, YP\_006616644.1,  
YP\_006669705.1, ZP\_08861138.1, ZP\_09031888.1, ZP\_09052683.1, ZP\_09053916.1,  
ZP\_09081537.1, ZP\_09085074.1, ZP\_09129119.1, ZP\_09141990.1, ZP\_09142519.1,  
ZP\_09160710.1, ZP\_09161266.1, ZP\_09170021.1, ZP\_09211718.1, ZP\_09213880.1,  
ZP\_09215772.1, ZP\_09221131.1, ZP\_09221448.1, ZP\_09259201.1, ZP\_09273984.1,  
ZP\_09307381.1, ZP\_09310147.1, ZP\_09411055.1, ZP\_09469041.1, ZP\_09503270.1,  
ZP\_09514613.1, ZP\_09517910.1, ZP\_09684781.1, ZP\_09684970.1, ZP\_09743302.1,  
ZP\_09787302.1, ZP\_09795945.1, ZP\_09797891.1, ZP\_09903272.1, ZP\_09919770.1,

	ZP_09973373.1, ZP_10003334.1, ZP_09979645.1, ZP_10003769.1, ZP_10019707.1, ZP_10021373.1, ZP_10031959.1, ZP_10037453.1, ZP_10049740.1, ZP_10063874.1, ZP_10063965.1, ZP_10065940.1, ZP_10088994.1, ZP_10111660.1, ZP_10987683.1, ZP_09843791.1, ZP_09843795.1, YP_001945311.1, ZP_02181409.1
<b>CYP153</b>	YP_006520827.1, YP_006671729.1, ZP_09194131.1, ZP_09195429.1, ZP_09691619.1, ZP_09692272.1, ZP_09692836.1, ZP_09158760.1, ZP_09470676.1, ZP_09221884.1, ZP_09213539.1, ZP_09268583.1, ZP_09270150.1, ZP_09798600.1, ZP_09472550.1, ZP_09472551.1, ZP_09424714.1, ZP_09424715.1, ZP_09425828.1, ZP_09425829.1, ZP_09433453.1, ZP_09433454.1, ZP_09646153.1, ZP_09650890.1, ZP_09650891.1, ZP_10082979.1, ZP_10082980.1, ZP_10085171.1, ZP_10062948.1, ZP_10065575.1, YP_691921.1, YP_694008.1, YP_004494060.1, YP_004495520.1, YP_001242465.1, YP_001242466.1, NP_768493.1, NP_773882.1, NP_773883.1, YP_005607016.1, YP_005607017.1, YP_005612806.1, YP_001203574.1, YP_001203575.1, YP_005453168.1, YP_005453169.1, YP_557448.1, NP_418882.1, YP_002515436.1, YP_001672075.1, YP_001681656.1, YP_003591161.1, YP_003883668.1, YP_003332953.1, YP_458852.1, YP_459378.1, YP_459878.1, YP_003810988.1, YP_005284867.1, YP_003059227.1, YP_760740.1, YP_005888020.1, YP_957888.1, YP_005430725.1, YP_005431650.1, YP_001702784.1, YP_006442774.1, YP_001135848.1, YP_005342987.1, YP_001071498.1, YP_939264.1, YP_001851443.1, YP_640381.1, YP_005002875.1, YP_005262975.1, YP_495502.1, YP_497095.1, YP_004535688.1, YP_001411305.1, YP_001413041.1, YP_001413057.1, YP_001413223.1, YP_001413399.1, YP_002129656.1, YP_552058.1, YP_345695.1, YP_345718.1, YP_783213.1, YP_534107.1, YP_570816.1, NP_946959.1, YP_004110202.1, YP_487538.1, YP_001990805.1, YP_617903.1, ZP_01041003.1, ZP_01102434.1, ZP_01102687.1, ZP_01615522.1, ZP_01618279.1, ZP_01618281.1, ZP_01618489.1, ZP_01625037.1, ZP_01626700.1, ZP_01627262.1, ZP_01858113.1, ZP_01863452.1, ZP_01864930.1, ZP_01913735.1, ZP_07715822.1, ZP_07716433.1, ZP_04384919.1, ZP_06860085.1, ZP_08271781.1, ZP_08022914.1, ZP_08025219.1, ZP_08206912.1, ZP_08700267.1, ZP_08207422.1, ZP_08764507.1, ZP_00997728.1, ZP_08631162.1, ZP_05042596.1, ZP_05043097.1, ZP_05095005.1, ZP_05095535.1, ZP_05126641.1, ZP_05129284.1, ZP_06847816.1, ZP_08199554.1

**Table S8** The Gene ID of alkane hydroxylase genes found within metagenomes from IMG in this study

<p><b>AlkB from freshwater metagenomes</b></p>	<p>2152912120, 2153066070, 2153127702, 2153155321, 2153184800, 2153232323, 2153236786, 2199956256, 2199957283, 2199959501, 2199980691, 2199992636, 2200001422, 2200049462, 2200051662, 2200055129, 2200056815, 2200094276, 2200096633, 2200102244, 2199042235, 2199096441, 2199253809, 2199377124, 2199441716, 2007429959, 2007448411, 2007479190, 2007496894, 2007550655, 2110018361, 2110069610, 2146024190, 2146044713, 2146062026, 2146074999, 2146112658, 2146112659, 2082684756, 2082693304, 2082694834, 2082705798, 2082729702, 2082733452, 2082753217, 2110214773, 2110285135, 2110313491, 2110375343, 2110397959, 2110416595, 2110503688, 2110517070, 2147883958, 2147903161, 2147917038, 2147944308, 2147977363, 2148066441, 2148089738, 2148170759, 2148226069, 2148317721, 2010483029, 2199890927, 2199897727, 2199899501, 2200107347, 2200114277, 2200114984, 2200116083, 2200116243, 2200117507, 2200118551, 2200124805, 2200124883, 2200137080, 2200139921, 2200152227, 2200158405, 2200161836, 2200173753, 2200193092, 2200195534, 2200220844, 2200229725, 2200230174, 2200240816, 2200244335, 2099135427, 2199551028, 2199612732, 2199677650, 2199779248, 2199824497, 2199854856, 2007098840, 2007103172, 2155871294, 2155871295, 2155977052, 2091446552, 2091508638, 2091532092, 2091629931, 2091666092, 2091669390, 2091695484, 2091718016, 2091739609, 2091751830, 2091811338, 2091843978, 2091847535, 2091860619, 2091911350, 2091913147, 2091919334, 2091925842, 2091962507, 2091987132, 2091987133, 2091991340, 2092063056, 2092096976, 2092164989, 2092219133, 2092255995, 2092258914, 2092328625, 2092368024, 2092374476, 2092386543, 2092390813, 2092401330, 2092452497, 2092468746, 2092515239, 2092540804, 2092550047, 2092557248, 2092561734, 2092574952, 2092584914</p>
<p><b>CYP153 from freshwater metagenomes</b></p>	<p>2153128639, 2200049077, 2200065807, 2199258204, 2007553160, 2007571680, 2146087201, 2082768974, 2082772809, 2091483348, 2091603010, 2091654133, 2091754804, 2091779054, 2091786426, 2091792952, 2091792953, 2091826037, 2091826038, 2091839789, 2091912774, 2091927638, 2092058632, 2092146923, 2092156155, 2092163238, 2092190235, 2092190718, 2092197684, 2092208081, 2092252838, 2092264416, 2092303637, 2092337780, 2092353830, 2092377372, 2092383966, 2092389982, 2092406710, 2092506549, 2092507591, 2092542583, 2092546486</p>
<p><b>AlkB from marine metagenomes</b></p>	<p>2001432130, 2001441346, 2001448879, 2001450388, 2001452645, 2001460866, 2001475876, 2001479249, 2001482387, 2001497719, 2001500492, 2001512768, 2001514417, 2001515382, 2004208250, 2014619960, 2014619962, 2014621037, 2014630392, 2014652732, 2014656780, 2014666703, 2014670873, 2014692467, 2014700866, 2014704638, 2014713745, 2017296663, 2017296670, 2017302191, 2017313637, 2079130576, 2079267324, 2079809231, 2079814040, 2079828514, 2079842491, 2079866558, 2079913685, 2079914104, 2079953958, 2079997357, 2080007911, 2080025718, 2080046538, 2080090589, 2080104259, 2080128464, 2080132150, 2080161122, 2080284914, 2080313102, 2080357523, 2080447082, 2086380661, 2094781363, 2096821728, 2096830541, 2096835172, 2096841236,</p>

2099043894, 2099099723, 2103059495, 2103071925, 2103079376, 2103123565, 2103151160, 2103160283, 2103402355, 2103405849, 2103418446, 2103418447, 2103433170, 2103445802, 2103481956, 2103604471, 2103609008, 2103641907, 2103670696, 2103721458, 2103764513, 2103774990, 2103829796, 2105890463, 2106530136, 2108229128, 2108591045, 2108603436, 2109154077, 2141541937, 2141828326, 2142034601, 2162007421, 2162010203, 2162017346, 2162020460, 2162026532, 2162043452, 2162049665, 2162056127, 2162072581, 2162085321, 2162092136, 2162110014, 2162116680, 2162151750, 2162154648, 2162162822, 2162168927, 2162173274, 2162184588, 2162202271, 2162218454, 2162220229, 2162220798, 2162241212, 2162248020, 2162271306, 2162291073, 2162292121, 2162294397, 2162307433, 2162316385, 2162343393, 2162343666, 2162345406, 2162346134, 2162360769, 2162360955, 2162366241, 2162367470, 2162375335, 2162379066, 2162379217, 2162381860, 2162382003, 2162384634, 2162391712, 2162393050, 2162394966, 2162397309, 2162401655, 2162401809, 2162403985, 2162404137, 2162404600, 2162404914, 2162405739, 2162423566, 2162438424, 2162439172, 2162439326, 2162440127, 2162440628, 2162441187, 2162442605, 2162442765, 2162442839, 2162444929, 2162444996, 2162450656, 2162456128, 2162466659, 2162471267, 2162473117, 2162476553, 2162476699, 2162476847, 2162552361, 2162553847, 2162560521, 2162563635, 2162564651, 2162566591, 2162571756, 2162578607, 2162584829, 2162584848, 2162586236, 2162597218, 2162600322, 2162602007, 2162605871, 2162618350, 2162624203, 2162625559, 2162628283, 2162631845, 2162634335, 2162636475, 2162643663, 2162645448, 2162645832, 2162646132, 2162651453, 2162651780, 2162655884, 2162656859, 2162658404, 2162669989, 2162673199, 2162674397, 2162677746, 2162679644, 2162682481, 2162683564, 2162683771, 2162684208, 2162684370, 2162684868, 2162685168, 2162689180, 2162691023, 2162695177, 2162697836, 2162700905, 2162706804, 2162711686, 2162714505, 2162715869, 2162717012, 2162725180, 2162736612, 2162739243, 2162751832, 2162752570, 2162754117, 2162755933, 2162756785, 2162756958, 2162758772, 2162763909, 2162765452, 2162767732, 2162771399, 2162783172, 2162785691, 2162798702, 2162801544, 2162807816, 2162812839, 2162822923, 2162835067, 2162841956, 2162859312, 2162862290, 2163753504, 2163754414, 2163774932, 2163801300, 2163803586, 2163803885, 2163804793, 2163805225, 2163819979, 2163848441, 2163888715, 2163891999, 2163902975, 2163903434, 2163909706, 2163924620, 2163930125, 2163935099, 2163939590, 2163944512, 2163946449, 2163947256, 2163951838, 2163953415, 2163961453, 2163985239, 2163987717, 2163997581, 2164014667, 2164023963, 2164036723, 2164041745, 2164043857, 2164057731, 2164063613, 2164068369, 2164083553, 2164084234, 2164101215, 2164102827, 2164111420, 2164121421, 2164147095, 2164165073, 2164186548, 2164195381, 2164197759, 2164199894, 2164218533, 2164219524, 2164225713, 2164241788, 2164246302, 2164250827, 2164251627, 2164259619, 2164260319, 2164266229, 2164267156, 2164273302, 2164277457, 2164286343, 2164286354, 2164297583, 2164322728, 2164322995, 2164325950, 2164335953, 2164335997, 2164336939, 2164337971, 2164342454, 2164344949, 2164376719, 2164399044, 2164407824, 2164417231, 2164423077,

	<p>2164436262, 2164437328, 2164437752, 2164440543, 2164440910, 2164449794, 2164453845, 2164457771, 2164459249, 2164462369, 2164463259, 2164465346, 2164471828, 2164475320, 2164492827, 2164493116, 2164493227, 2164494766, 2164500916, 2164513253, 2164518194, 2164531685, 2164548161, 2164549546, 2164557407, 2164573051, 2164577726, 2164596007, 2164597434, 2164600492, 2164603470, 2164612124, 2164620125, 2164651628, 2164666327, 2164669092, 2164669994, 2164670749, 2164673839, 2164675678, 2164683498, 2164689764, 2164691762, 2164695642, 2164697365, 2164697379, 2164699476, 2164701986, 2164702653, 2164703148, 2164707131, 2164708440, 2164720706, 2164722533, 2164736344, 2164736457, 2164737067, 2164740993, 2164748420, 2164753040, 2164761203, 2164764315, 2164765294, 2164766840, 2164770590, 2164770712, 2164792685, 2164794626, 2164794905, 2164800669, 2164814688, 2164820601, 2164843967, 2164846844, 2195652973, 2195654196, 2195658154, 2195659483, 2195661767, 2195661998, 2195669072, 2195672581, 2195705514, 2195736793, 2195769611, 2195802107, 2195817592, 2195848340, 2195851701, 2195855993, 2195858492, 2195860684, 2195866935, 2195868573, 2195869422, 2195882488, 2195883111, 2195888120, 2195891976, 2195897390, 2195897500, 2195898628, 2195900438, 2195904291, 2195908253, 2195911522, 2195922116, 2195924068, 2195930140, 2195935300, 2195945397, 2195949953, 2195959008, 2195960315, 2195970135, 2195972110, 2195972906, 2196101050, 2196111001, 2196112354, 2196117866, 2196122362, 2196122628, 2196126858, 2196126992, 2196127114, 2196135001, 2196135215, 2196150304, 2196151769, 2196168863, 2196173500, 2196178172, 2196189534, 2196205399, 2196207047, 2196212385, 2196218048, 2196218543, 2196229160, 2196234423, 2196243765, 2196247867, 2196264840, 2196273104, 2196281163, 2196286696, 2196298168, 2196308175, 2196315808, 2196322642, 2196344748, 2196347304, 2196355951, 2196356889, 2196363467, 2196368381, 2196369652, 2196372029, 2196400168, 2196401234, 2196403886, 2196412137, 2196421016, 2196421140, 2196626924, 2196627115, 2196630541, 2196651952, 2196658704, 2196659370, 2196667400, 2196668006, 2196711621, 2196727526, 2196732530, 2196733935, 2196748809, 2196749838, 2196756723, 2196757714, 2196772292, 2196777337, 2196801954, 2196802113, 2196807957, 2196811651, 2196996140, 2197020524, 2197021928, 2197023436, 2197030929, 2197078638, 2197089723, 2197110789, 2197132384, 2197152915, 2197158938, 2197160725, 2197164385</p>
<p><b>CYP153 from marine metagenomes</b></p>	<p>2001401131, 2001403949, 2001421464, 2001421471, 2001421655, 2001427901, 2001433768, 2001448252, 2001451474, 2001451477, 2001454418, 2001456585, 2001458611, 2001458612, 2001463317, 2001471176, 2001475656, 2001481059, 2001483897, 2001486798, 2001492130, 2001515938, 2014621232, 2014621233, 2014631380, 2014633014, 2014642947, 2014656349, 2014668094, 2014668095, 2014668100, 2014681047, 2079226637, 2079852292, 2079877946, 2079877947, 2079929002, 2079941068, 2079946074, 2079947107, 2079954290, 2079954291, 2079982206, 2079990807, 2079990808, 2079999522, 2080029905, 2080120778, 2080148224, 2080152475, 2080215897, 2080218024, 2080236646, 2080238468, 2080241775, 2080282744, 2080284604, 2080298663, 2080472402, 2086525020,</p>



2095063124, 2096824297, 2096841889, 2096842399, 2096845216, 2096848258, 2096861319, 2099029949, 2099063112, 2103092180, 2103140936, 2103389175, 2103398524, 2103398525, 2103398526, 2103445100, 2103458071, 2103494659, 2103584332, 2103587993, 2103615551, 2103619096, 2103703988, 2103704069, 2103719945, 2103743845, 2103748063, 2103755966, 2109038312, 2109161684, 2109748251, 2109774313, 2109851023, 2162000997, 2162007439, 2162029874, 2162049559, 2162066271, 2162117089, 2162172995, 2162175310, 2162175321, 2162202220, 2162231570, 2162271641, 2162272021, 2162314461, 2162319846, 2162332409, 2162351148, 2162351642, 2162362593, 2162374832, 2162379787, 2162379896, 2162382296, 2162391951, 2162399252, 2162404593, 2162409559, 2162422768, 2162424078, 2162424945, 2162429134, 2162445570, 2162463962, 2162470917, 2162476762, 2162477591, 2162482213, 2162551207, 2162564204, 2162602725, 2162607630, 2162616781, 2162641430, 2162654366, 2162674235, 2162689863, 2162760603, 2162763612, 2162789648, 2162802765, 2162812825, 2162814909, 2162850539, 2162869532, 2162874081, 2163743708, 2163783151, 2163801455, 2163824416, 2163905331, 2163912112, 2163913038, 2163935048, 2163958946, 2164025190, 2164037097, 2164047365, 2164109478, 2164122310, 2164171570, 2164191573, 2164191574, 2164211572, 2164213445, 2164214453, 2164217303, 2164249183, 2164250325, 2164254560, 2164268400, 2164275956, 2164281789, 2164286521, 2164291139, 2164295231, 2164299353, 2164305720, 2164335936, 2164354841, 2164389340, 2164392275, 2164392276, 2164396154, 2164396155, 2164401284, 2164403772, 2164431005, 2164447391, 2164462774, 2164464918, 2164465125, 2164475223, 2164477551, 2164478973, 2164490195, 2164490255, 2164494799, 2164494800, 2164514469, 2164514830, 2164514831, 2164531630, 2164536580, 2164537000, 2164550914, 2164553920, 2164555475, 2164563586, 2164565831, 2164589304, 2164594675, 2164598315, 2164601674, 2164604414, 2164616436, 2164616437, 2164628664, 2164629965, 2164629966, 2164634482, 2164635709, 2164641420, 2164646703, 2164650565, 2164655268, 2164659123, 2164662021, 2164662022, 2164664918, 2164665113, 2164669485, 2164683083, 2164690003, 2164694593, 2164708283, 2164708565, 2164708566, 2164749878, 2164757387, 2164778567, 2164780869, 2164782208, 2164786311, 2164788288, 2164788518, 2164788519, 2164792423, 2164802981, 2164805135, 2164812894, 2164813762, 2164827606, 2164827607, 2164835463, 2164838071, 2164848821, 2164851087, 2164851088, 2195660937, 2195854590, 2195880739, 2195886590, 2195887866, 2195888763, 2195893764, 2195916259, 2195919609, 2195941875, 2195943106, 2196132280, 2196154418, 2196164550, 2196191042, 2196199868, 2196200346, 2196216812, 2196224421, 2196233802, 2196237519, 2196257211, 2196292014, 2196300504, 2196302636, 2196314552, 2196333329, 2196339269, 2196345196, 2196346320, 2196350255, 2196356068, 2196364401, 2196369256, 2196372278, 2196377991, 2196391278, 2196398487, 2196403635, 2196405224, 2196407247, 2196628530, 2196651257, 2196657622, 2196665544, 2196683294, 2196687272, 2196690883, 2196695849, 2196695850, 2196701889, 2196704042, 2196704387, 2196705794, 2196724724, 2196727822, 2196733524, 2196742219, 2196748093, 2196752167, 2196755993, 2196759346, 2196778803,

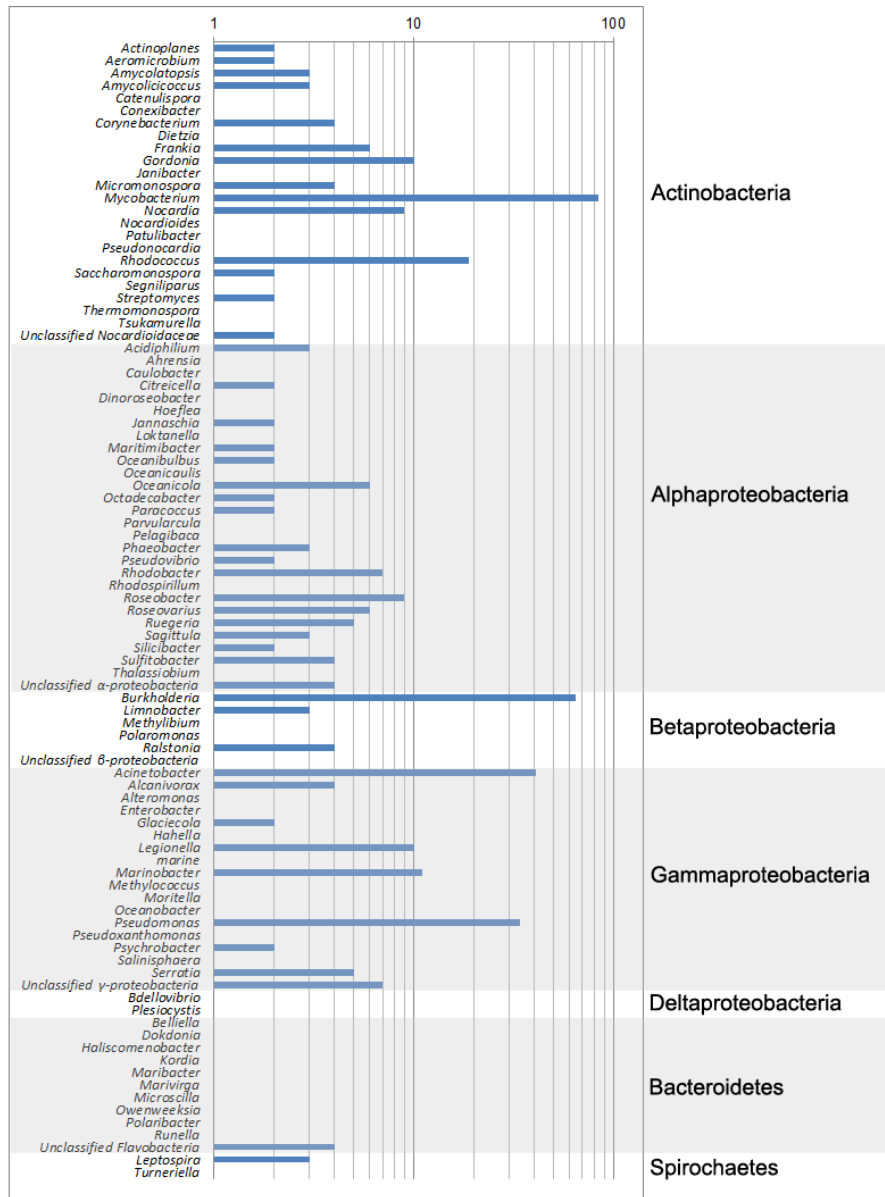
	2196797073, 2196809510, 2196810200, 2197000555, 2197029676, 2197062848, 2197076482, 2197147661
<b>AlkB from terrestrial metagenomes</b>	2044097886, 2044192625, 2138966160, 2137518635, 2137751467, 2142083522, 2142142659, 2142265303, 2033392830, 2033397823, 2033488086, 2033534823, 2033547419, 2033571370, 2033751883, 2033793281, 2033810236, 2033824712, 2033835587, 2033859262, 2034003778, 2034061695, 2034144450, 2034172207, 2034194073, 2034201828, 2034240979, 2021898468, 2021974252, 2165900571, 2165926876, 2165928888, 2166046783, 2122685362, 2122736624, 2122742278, 2122832070, 2122967859, 2123173135, 2123173992, 2123293021, 2123318239, 2123492521, 2001217786, 2001249707, 2001274721, 2001285122, 2001327087, 2001371637, 2001396247, 2136051439, 2136051440, 2136051988, 2121847312, 2121862381, 2121865704, 2121958690, 2121963234, 2121980556, 2122071687, 2122092627, 2122114140, 2122124808, 2122184031, 2122353024, 2122435506, 2122499380, 2122511080, 2122594968, 2139041515, 2139042916, 2139099814, 2139128499, 2139140302, 2139185345, 2137968763, 2138231929, 2138523247, 2138781332, 2124989147, 2125073275, 2125297867, 2125354067, 2125447820, 2125503195, 2125541197, 2125604435, 2125764432, 2125849661, 2125856888, 2126163169, 2126163658, 2041786030, 2041990873, 2042046256, 2036252947, 2036317017, 2036252947, 2036317017, 2032534351, 2032723864, 2032736648, 2032784488, 2032784505, 2032906534, 2032959357, 2032982747, 2032989697, 2033016936, 2033027367, 2037091775, 2037105532, 2037133248, 2037168796, 2037258412, 2037314484, 2033081260, 2033104302, 2033305047, 2034350059, 2034408742, 2034451174, 2034583464, 2131987043, 2132016602, 2132085688, 2132113132, 2132157127, 2132228797, 2132279567, 2132367809, 2132436766, 2132494896, 2132495018, 2132530687, 2132565300, 2132593989, 2132622146, 2132792464, 2132829836, 2132840226, 2132882733, 2132895758, 2130980536, 2131031943, 2131066086, 2131115878, 2131169924, 2131196835, 2131213019, 2131232795, 2131290872, 2131326636, 2131342609, 2131368734, 2131525401, 2131567704, 2131606037, 2131623043, 2131627322, 2131780748, 2131822297, 2131856143, 2131869255, 2131871764, 2131912567, 2131938037, 2129979333, 2130242426, 2130252357, 2130275153, 2130312463, 2130531530, 2130559631, 2130568242, 2130619783, 2130655440, 2130662709, 2130723047, 2130732873, 2130792184, 2130831203, 2130890918, 2129088363, 2129095285, 2129128191, 2129208549, 2129215712, 2129244909, 2129473576, 2129492863, 2129591482, 2129775754, 2129865409, 2129877170, 2121847312, 2121862381, 2121865704, 2121958690, 2121963234, 2121980556, 2122071687, 2122092627, 2122114140, 2122124808, 2122184031, 2122353024, 2122435506, 2122499380, 2122511080, 2122594968, 2209196193, 2209222452, 2209224582, 2209354615, 2209380262, 2209501071, 2209553309, 2209614992, 2209760518, 2209773098, 2209807142, 2209990709, 2210049141, 2210150160, 2210156651, 2120772088, 2120805680, 2120968337, 2120988190, 2121058068, 2121309987, 2121366478, 2121448666, 2121467755, 2121659521, 2123591061, 2123677147, 2123745057, 2123801225, 2123851307, 2124301143, 2124381370, 2124400709, 2124472096, 2124493978, 2124552969, 2131987043, 2132016602, 2132085688, 2132113132, 2132157127,

	<p>2132228797, 2132279567, 2132367809, 2132436766, 2132494896, 2132495018, 2132530687, 2132565300, 2132593989, 2132622146, 2132792464, 2132829836, 2132840226, 2132882733, 2132895758, 2122685362, 2122736624, 2122742278, 2122832070, 2122967859, 2123173135, 2123173992, 2123293021, 2123318239, 2123492521, 2209196193, 2209222452, 2209224582, 2209354615, 2209380262, 2209501071, 2209553309, 2209614992, 2209760518, 2209773098, 2209807142, 2209990709, 2210049141, 2210150160, 2210156651, 2021681900, 2021682418, 2021682752, 2021686808, 2021687273, 2021689320, 2021689709, 2021898468, 2021974252</p>
<p><b>CYP153 from terrestrial metagenomes</b></p>	<p>2044140713, 2138924488, 2138932531, 2138950332, 2138994944, 2139000238, 2142191092, 2033440037, 2033456434, 2033475663, 2033600287, 2033644092, 2033686702, 2033710083, 2033807286, 2033881815, 2033894132, 2033915504, 2033915681, 2033925351, 2033949508, 2034069109, 2034071204, 2034074633, 2034090461, 2034104922, 2034105578, 2034111166, 2034129966, 2034131664, 2034151128, 2034170717, 2034178394, 2034179358, 2034199168, 2034220872, 2021772396, 2021910894, 2021925129, 2021932823, 2165935941, 2166075969, 2166096375, 2122737178, 2122922483, 2122959077, 2122980582, 2123040404, 2123069833, 2123251491, 2123434932, 2001397730, 2121803935, 2121837583, 2121841340, 2121848343, 2121855017, 2121903220, 2121916472, 2121922593, 2121935341, 2121938710, 2121953727, 2121955827, 2121961750, 2121973663, 2121973664, 2121981037, 2121991891, 2122056893, 2122070502, 2122086773, 2122090294, 2122131080, 2122140034, 2122145709, 2122162687, 2122164466, 2122182952, 2122185286, 2122252419, 2122253389, 2122253625, 2122260641, 2122262114, 2122333309, 2122333381, 2122354834, 2122390457, 2122397806, 2122414792, 2122415035, 2122418580, 2122419263, 2122421230, 2122450910, 2122456658, 2122473567, 2122477736, 2122518098, 2122520578, 2122553623, 2122560833, 2122575071, 2122575593, 2122615848, 2139067055, 2139072915, 2139098005, 2139124800, 2139193300, 2139203577, 2139203578, 2138156095, 2125014928, 2125062833, 2125071726, 2125120844, 2125143276, 2125146548, 2125189074, 2125194629, 2125218532, 2125219649, 2125300441, 2125371751, 2125404094, 2125460003, 2125536253, 2125536794, 2125542355, 2125546122, 2125549773, 2125552151, 2125571248, 2125572453, 2125631475, 2125637262, 2125648280, 2125687936, 2125710060, 2125711155, 2125743925, 2125785832, 2125786529, 2125808660, 2125815138, 2125823618, 2125886057, 2125886899, 2125902968, 2125921569, 2125941082, 2125943646, 2125950363, 2125997225, 2126000400, 2126007841, 2126026262, 2126026263, 2126056959, 2126084671, 2126125136, 2126147123, 2041785978, 2041793274, 2041800271, 2041803681, 2041804748, 2041827004, 2041827470, 2041846259, 2041874866, 2041908436, 2041909922, 2041917382, 2041943827, 2041957807, 2041964037, 2041965913, 2041970484, 2041981528, 2041988147, 2041988809, 2041998948, 2042027747, 2042079593, 2042092639, 2042103451, 2042117654, 2042133292, 2036174980, 2036174981, 2036195851, 2036203368, 2036208278, 2036213780, 2036220857, 2036229666, 2036251464, 2036257081, 2036263021, 2036263614, 2036274698, 2036300827, 2036300828, 2036314221, 2036337522, 2036362678, 2036402784,</p>

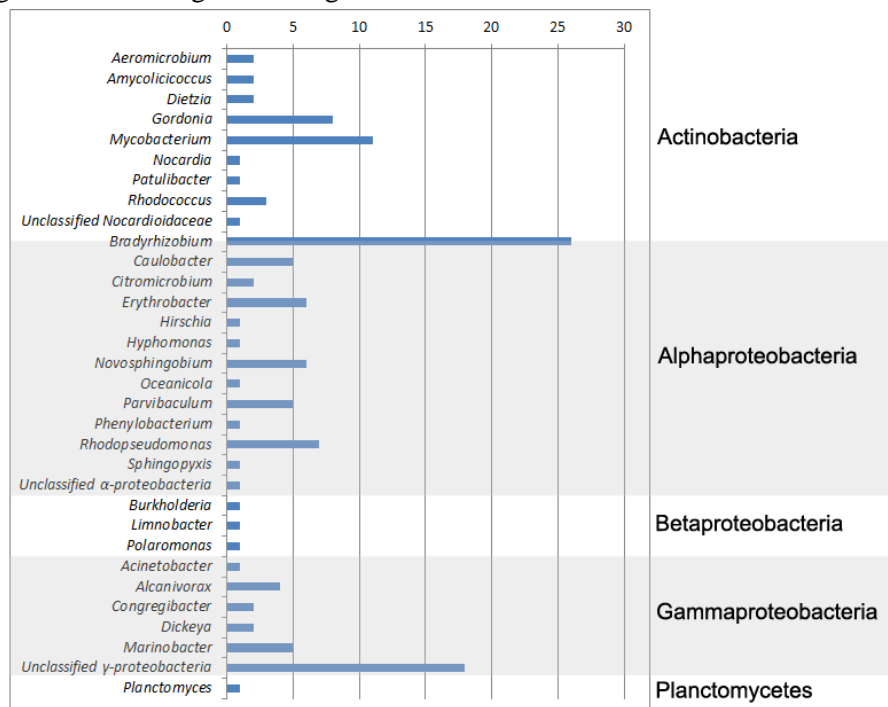
2036174980, 2036174981, 2036195851, 2036203368, 2036208278, 2036213780, 2036220857, 2036229666, 2036251464, 2036257081, 2036263021, 2036263614, 2036274698, 2036300827, 2036300828, 2036314221, 2036337522, 2036362678, 2036402784, 2032534309, 2032581254, 2032597030, 2032653500, 2032696916, 2032728825, 2032765602, 2032780629, 2032854971, 2032936998, 2033063266, 2037107177, 2037138142, 2037154581, 2037179837, 2037394055, 2037448908, 2037449996, 2033178847, 2033183226, 2033223943, 2033229708, 2033343250, 2034285153, 2034323546, 2034385820, 2034438878, 2034448696, 2034458409, 2034537374, 2034546594, 2034555525, 2034581818, 2034589544, 2034589871, 2034617687, 2132021991, 2132046120, 2132072591, 2132075662, 2132088228, 2132110790, 2132127901, 2132144776, 2132164041, 2132174430, 2132207627, 2132217601, 2132217602, 2132230237, 2132243718, 2132261837, 2132306050, 2132348575, 2132354306, 2132363618, 2132381486, 2132423919, 2132456074, 2132487370, 2132526462, 2132558888, 2132565014, 2132609728, 2132612352, 2132629190, 2132652034, 2132654600, 2132689753, 2132764522, 2132790145, 2132791008, 2132802399, 2132828174, 2132843893, 2132872442, 2132887103, 2132910838, 2132943215, 2132946593, 2132999915, 2133008881, 2133010884, 2130925585, 2130942053, 2130947408, 2130980854, 2131018552, 2131023378, 2131061023, 2131086319, 2131086347, 2131094704, 2131167875, 2131198422, 2131214097, 2131252904, 2131375910, 2131380446, 2131388934, 2131390487, 2131442710, 2131447763, 2131503097, 2131525883, 2131539508, 2131575981, 2131591937, 2131594987, 2131594988, 2131607834, 2131646461, 2131691510, 2131726129, 2131730297, 2131737470, 2131789722, 2131794695, 2131843898, 2131846208, 2131919786, 2131920508, 2131961937, 2129898841, 2129918672, 2129933735, 2129968752, 2129983928, 2130011770, 2130021888, 2130025477, 2130026692, 2130062266, 2130069369, 2130114788, 2130123745, 2130170246, 2130172289, 2130182749, 2130184638, 2130308228, 2130325054, 2130328278, 2130342511, 2130344235, 2130365045, 2130380518, 2130402580, 2130415024, 2130418511, 2130437435, 2130439213, 2130510033, 2130518687, 2130552976, 2130555827, 2130576053, 2130619689, 2130640172, 2130647744, 2130730351, 2130755110, 2130778645, 2130793978, 2130845435, 2130892291, 2130906571, 2130920159, 2128986319, 2129017442, 2129058729, 2129075337, 2129082236, 2129093614, 2129114006, 2129137124, 2129233993, 2129235474, 2129285193, 2129361414, 2129396862, 2129427828, 2129427920, 2129441286, 2129473231, 2129491082, 2129520172, 2129552329, 2129556589, 2129625406, 2129665294, 2129665862, 2129667132, 2129683692, 2129687505, 2129733270, 2129740206, 2129753864, 2129827369, 2129830276, 2121803935, 2121837583, 2121841340, 2121848343, 2121855017, 2121903220, 2121916472, 2121922593, 2121935341, 2121938710, 2121953727, 2121955827, 2121961750, 2121973663, 2121973664, 2121981037, 2121991891, 2122056893, 2122070502, 2122086773, 2122090294, 2122131080, 2122140034, 2122145709, 2122162687, 2122164466, 2122182952, 2122185286, 2122252419, 2122253389, 2122253625, 2122260641, 2122262114, 2122333309, 2122333381, 2122354834, 2122390457, 2122397806, 2122414792, 2122415035, 2122418580, 2122419263, 2122421230, 2122450910, 2122456658,

	2122473567, 2122477736, 2122518098, 2122520578, 2122553623, 2122560833, 2122575071, 2122575593, 2122615848, 2209169006, 2209230120, 2209304642, 2209348999, 2209349849, 2209491694, 2209807656, 2209824303, 2209824304, 2209848057, 2210058976, 2120790094, 2121253781, 2121253782, 2121263576, 2121286601, 2121381434, 2121635116, 2121667516, 2121684825, 2123631175, 2123751342, 2123770331, 2123839989, 2123929833, 2123931076, 2123950093, 2123972088, 2124221034, 2124241149, 2124446926, 2132021991, 2132046120, 2132072591, 2132075662, 2132088228, 2132110790, 2132127901, 2132144776, 2132164041, 2132174430, 2132207627, 2132217601, 2132217602, 2132230237, 2132243718, 2132261837, 2132306050, 2132348575, 2132354306, 2132363618, 2132381486, 2132423919, 2132456074, 2132487370, 2132526462, 2132558888, 2132565014, 2132609728, 2132612352, 2132629190, 2132652034, 2132654600, 2132689753, 2132764522, 2132790145, 2132791008, 2132802399, 2132828174, 2132843893, 2132872442, 2132887103, 2132910838, 2132943215, 2132946593, 2132999915, 2133008881, 2133010884, 2122737178, 2122922483, 2122959077, 2122980582, 2123040404, 2123069833, 2123251491, 2123434932, 2209169006, 2209230120, 2209304642, 2209348999, 2209349849, 2209491694, 2209807656, 2209824303, 2209824304, 2209848057, 2210058976, 2044673418, 2021772396, 2021910894, 2021925129, 2021932823
--	--

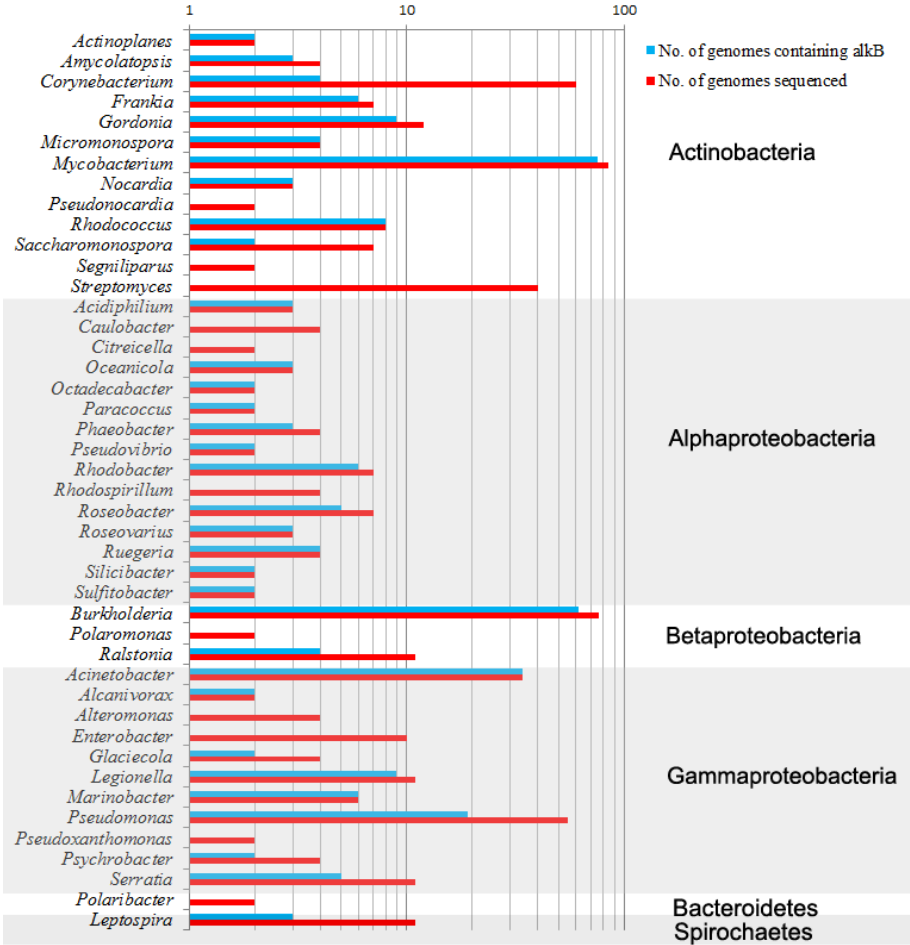
**Figure S1** General distribution of *alkB* genes in microbial genomes. The X-axis indicates the number of *alkB* genes found in sequence genomes of *alkB* harbouring genera. The Y-axis indicates the genera harbouring *alkB* genes.



**Figure S2** General distribution of CYP153 genes in microbial genomes. The X-axis indicates the number of CYP153 genes found in sequence genomes of CYP153 harbouring genera. The Y-axis indicates the genera harbouring CYP153 genes.

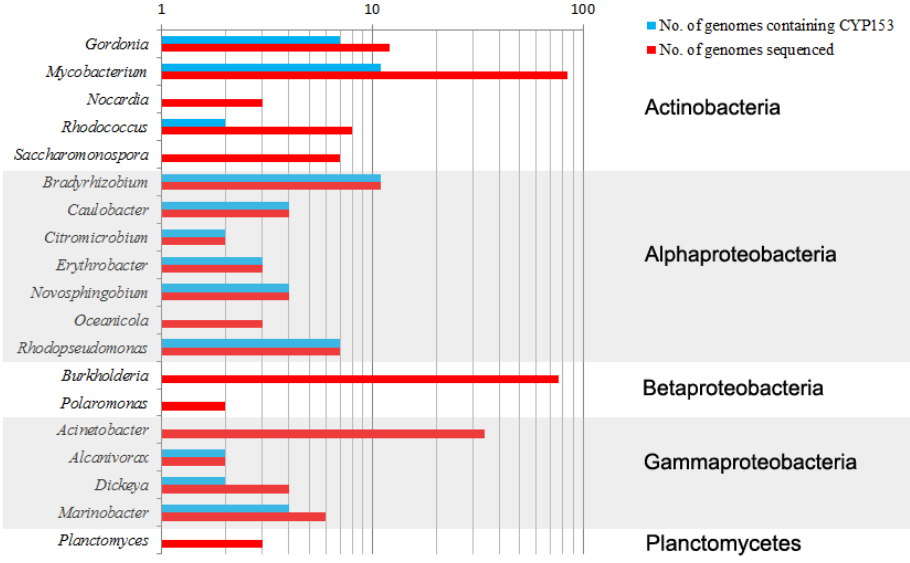


**Figure S3** Distribution of genomes containing *alkB* genes in the microbial genomes deposited in GenBank. The X-axis indicates the number of genomes in which the *alkB* genes were found.

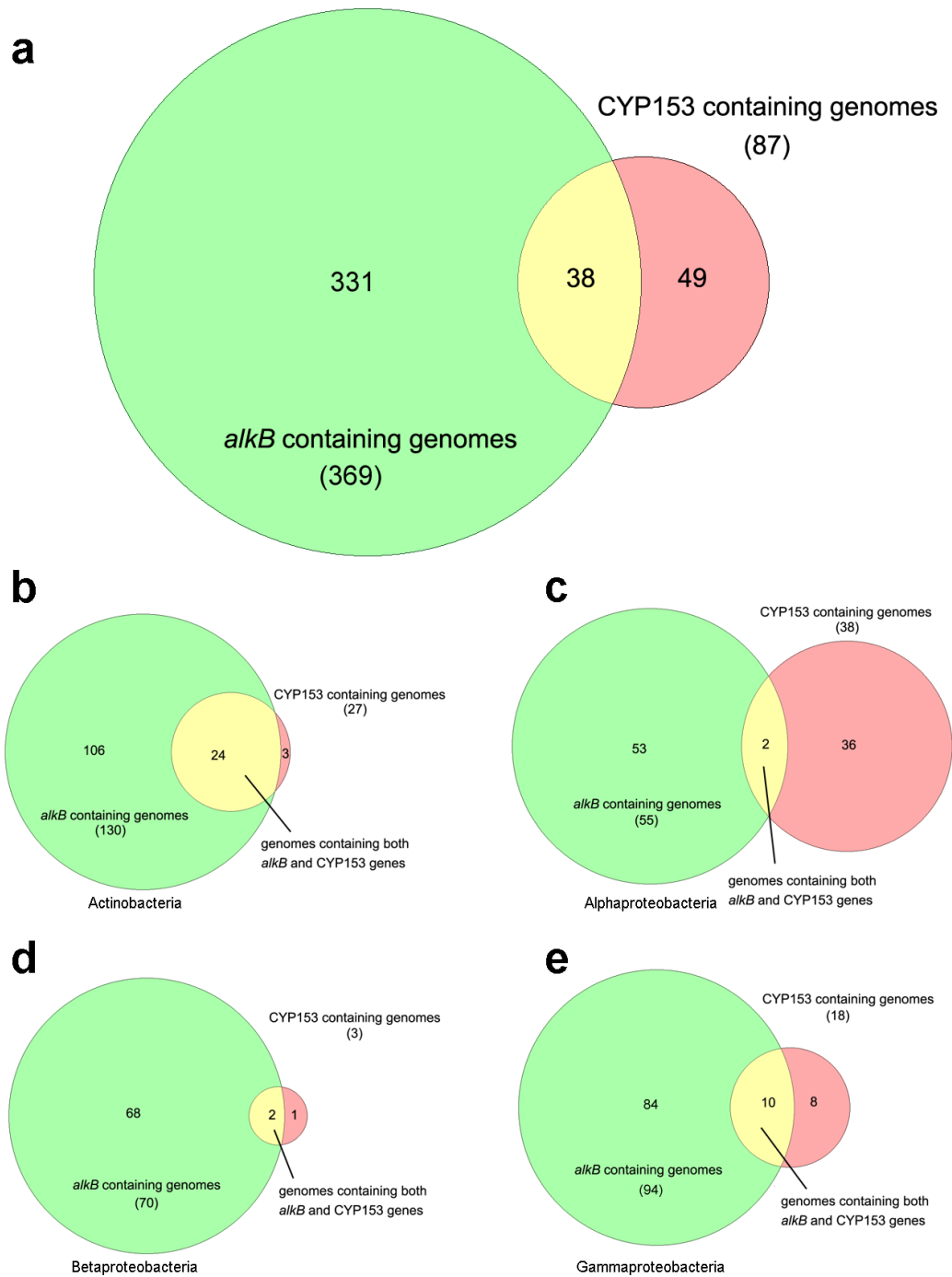




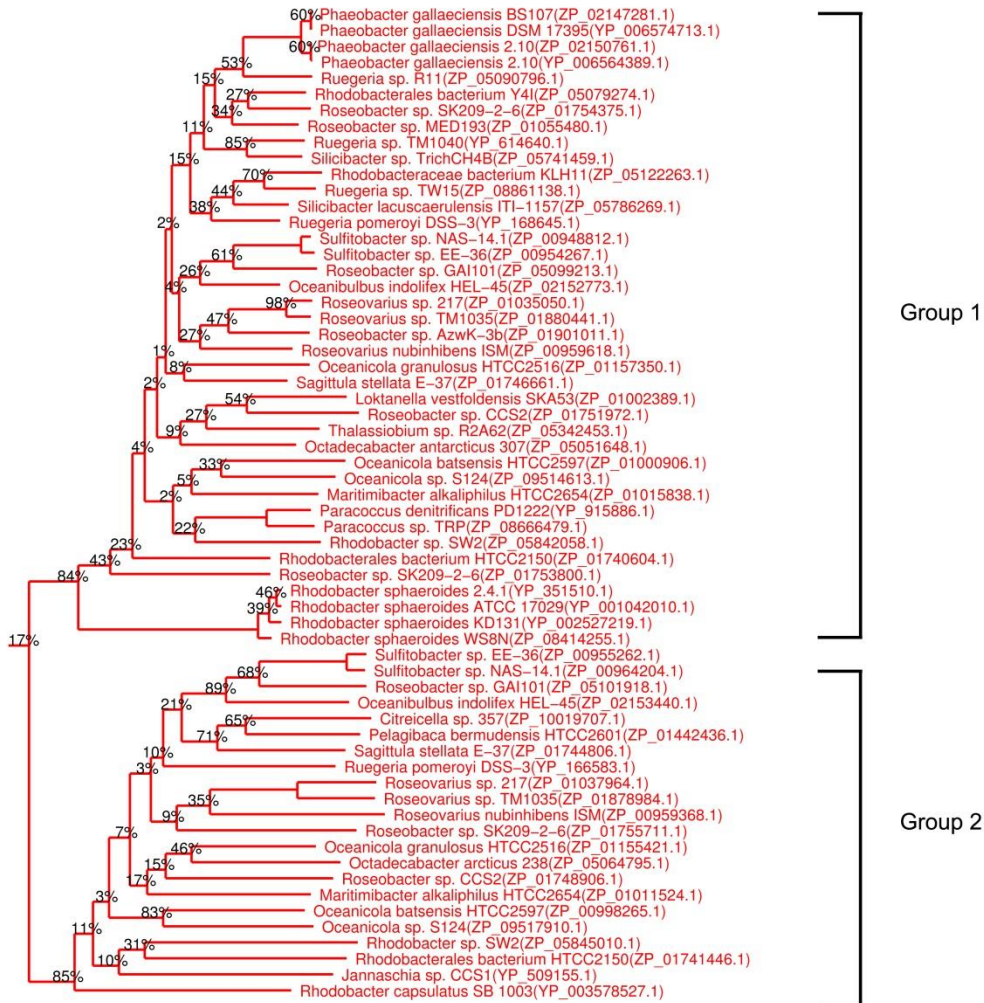
**Figure S4** Distribution of genomes containing CYP153 genes in the microbial genomes deposited in GenBank. The X-axis indicates the number of genomes in which the CYP153 genes were found.



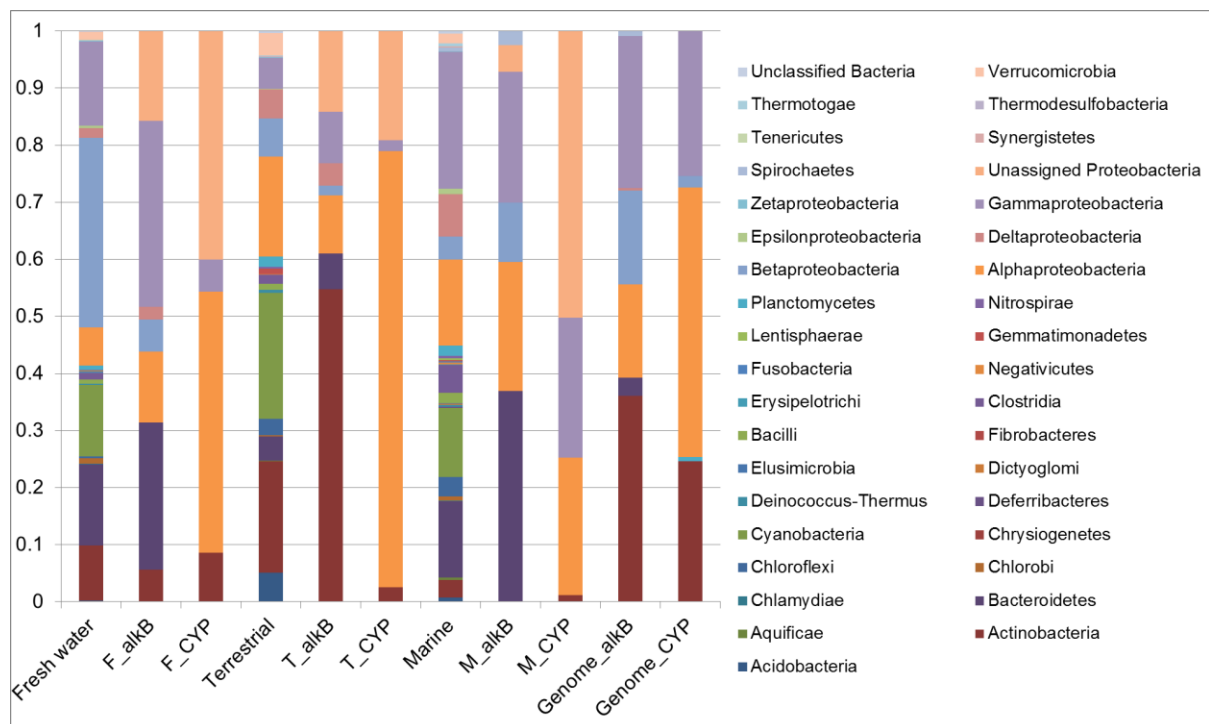
**Figure S5** The proportional Venn diagram of genomes containing *alkB* or CYP153 genes. a: The proportional Venn diagram of all the genomes containing *alkB* or CYP153 identified in this study. b: The proportional Venn diagram of Actinobacteria genomes identified in this study. c: The proportional Venn diagram of Alphaproteobacteria genomes identified in this study. d: The proportional Venn diagram of Betaproteobacteria genomes identified in this study. e: The proportional Venn diagram of Gammaproteobacteria genomes identified in this study. The number in brackets and circles indicates the number of genomes.



**Figure S6** Phylogenetic tree of *alkB* sequences in Cluster VII. Sequences in Cluster VII were all from Alphaproteobacteria and clustered into two groups.



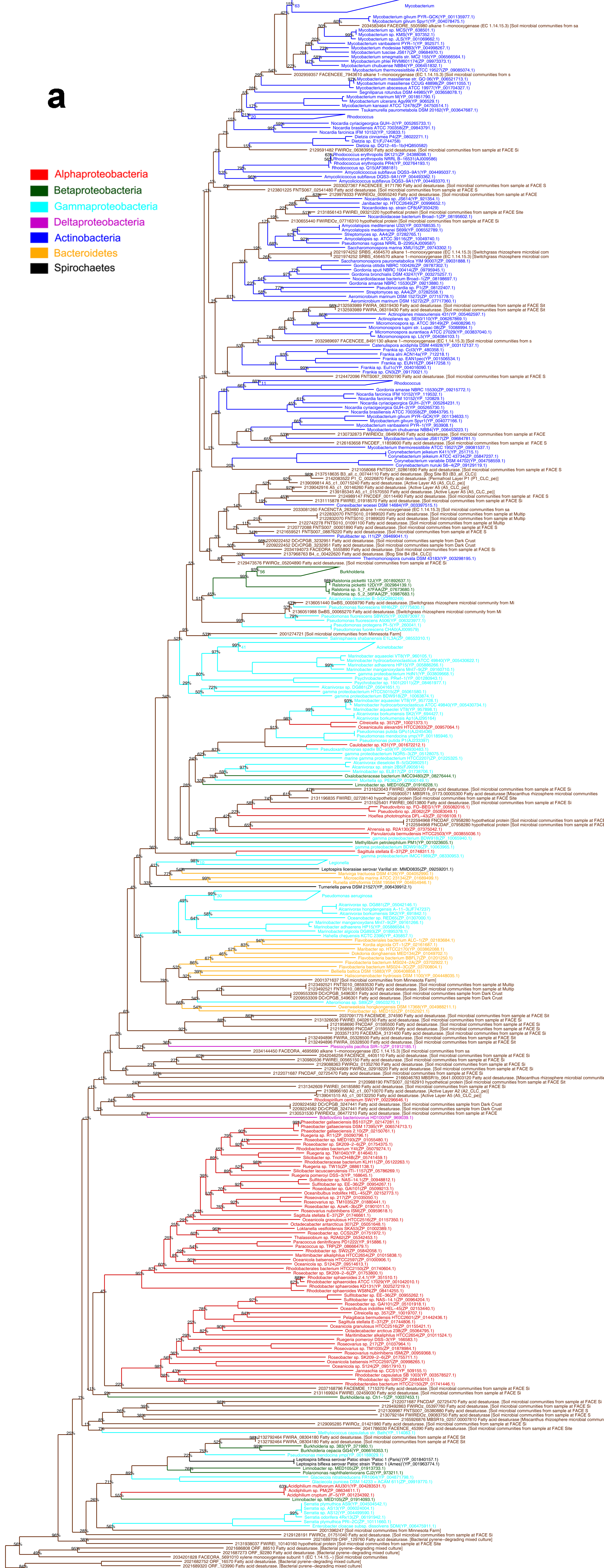
**Figure S7** Comparison of the taxonomic composition of metagenomes from three habitats and the taxonomic distribution of alkane hydroxylases in freshwater, marine and terrestrial habitats. Freshwater, the taxonomic composition of metagenomes from freshwater. F\_alkB, the taxonomic distribution of *alkB* genes in freshwater. F\_CYP, the taxonomic distribution of CYP153 genes in freshwater. Terrestrial, the taxonomic composition of metagenomes from terrestrial environments. T\_alkB, the taxonomic distribution of *alkB* genes in terrestrial environments. T\_CYP, the taxonomic distribution of CYP153 genes in terrestrial environments. Marine, the taxonomic composition of metagenomes from marine. M\_alkB, the taxonomic distribution of *alkB* genes in marine. M\_CYP, the taxonomic distribution of CYP153 genes in marine. Genome\_alkB, the taxonomic distribution of *alkB* genes among microbial genomes. Genome\_CYP, the taxonomic distribution of CYP153 genes among microbial genomes.



**Figure S8** Phylogenetic tree of *alkB* genes in metagenomes together with those from microbial genomes based on amino acid sequences analysis. a: AlkB from terrestrial metagenomes. b: AlkB from fresh water metagenomes. c: AlkB from marine metagenomes. Only the segments containing both the HYG motif and the third histidine box were used for alignment and to construct phylogenetic trees using ARB. The trees were bootstrapped with 1000 replicates, indicated at respective nodes. Genes from metagenomes are in brown.

a

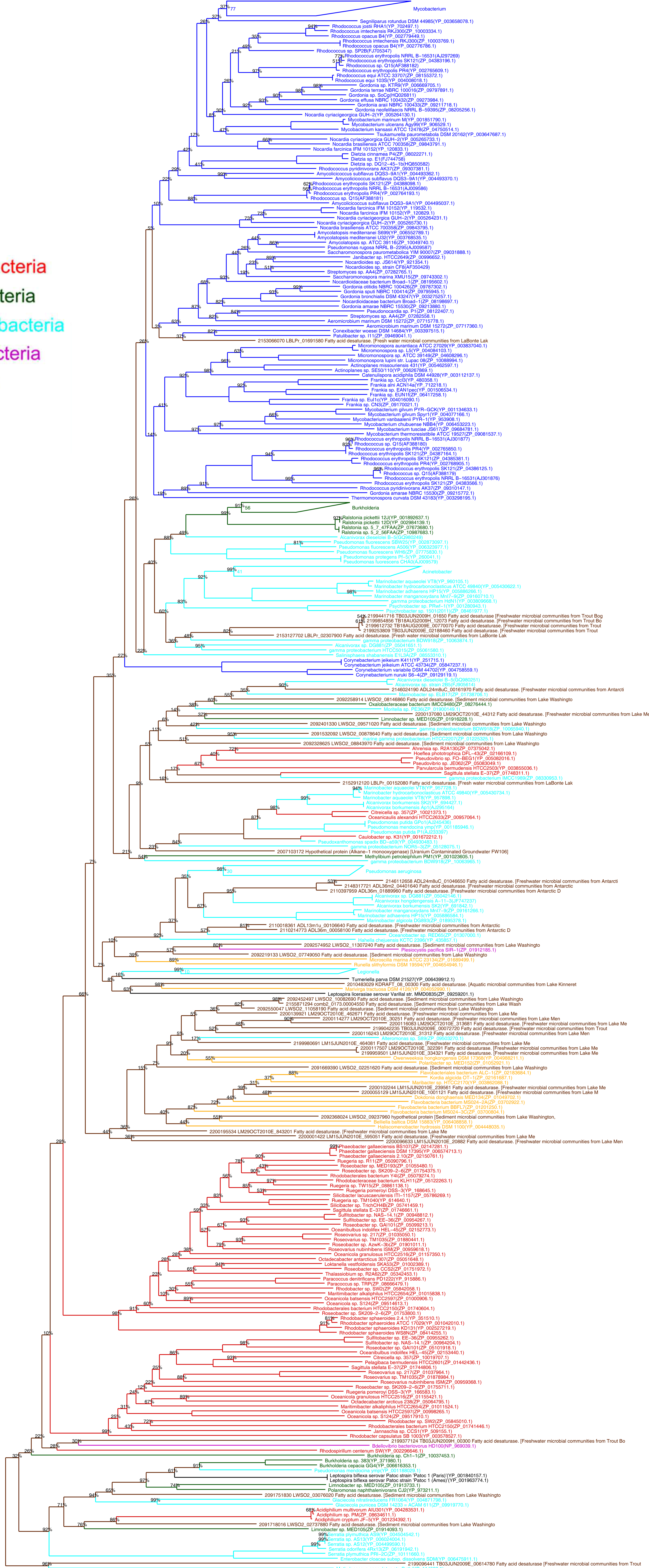
- Alphaproteobacteria
- Betaproteobacteria
- Gammaproteobacteria
- Deltaproteobacteria
- Actinobacteria
- Bacteroidetes
- Spirochaetes



0.10

b

- Alphaproteobacteria
- Betaproteobacteria
- Gammaproteobacteria
- Deltaproteobacteria
- Actinobacteria
- Bacteroidetes
- Spirochaetes



0.1

C

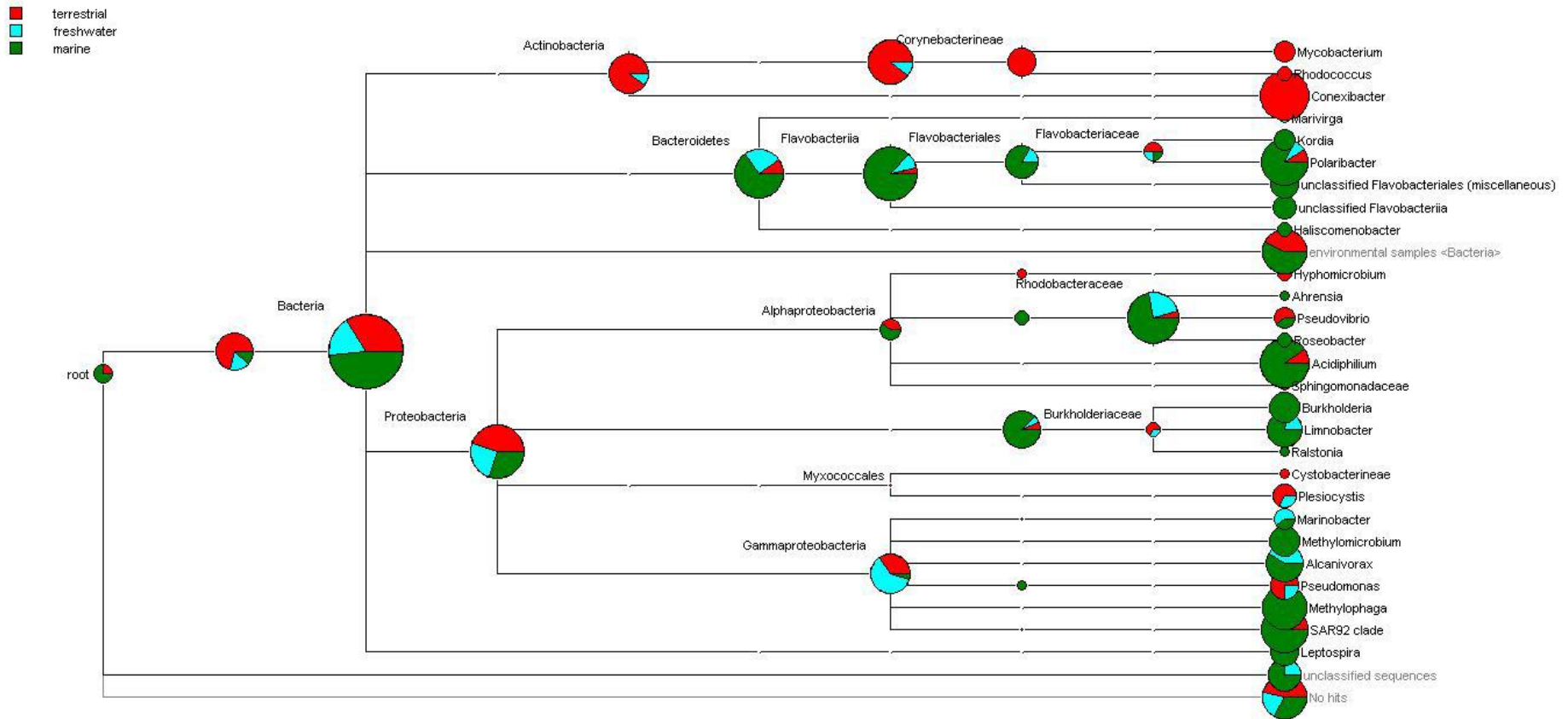
- Alphaproteobacteria
- Betaproteobacteria
- Gammaproteobacteria
- Deltaproteobacteria
- Actinobacteria
- Bacteroidetes
- Spirochaetes



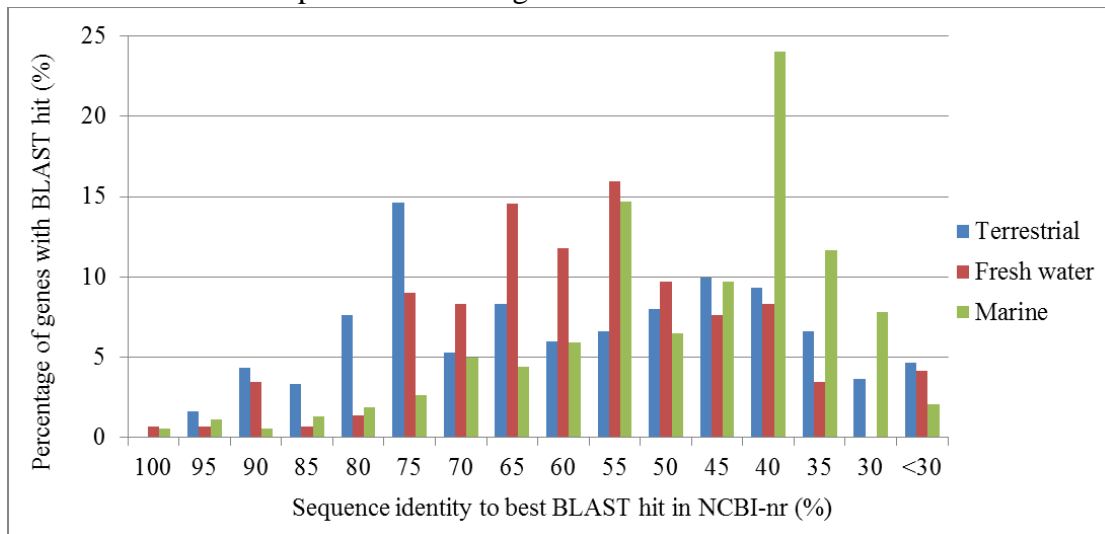
0.1



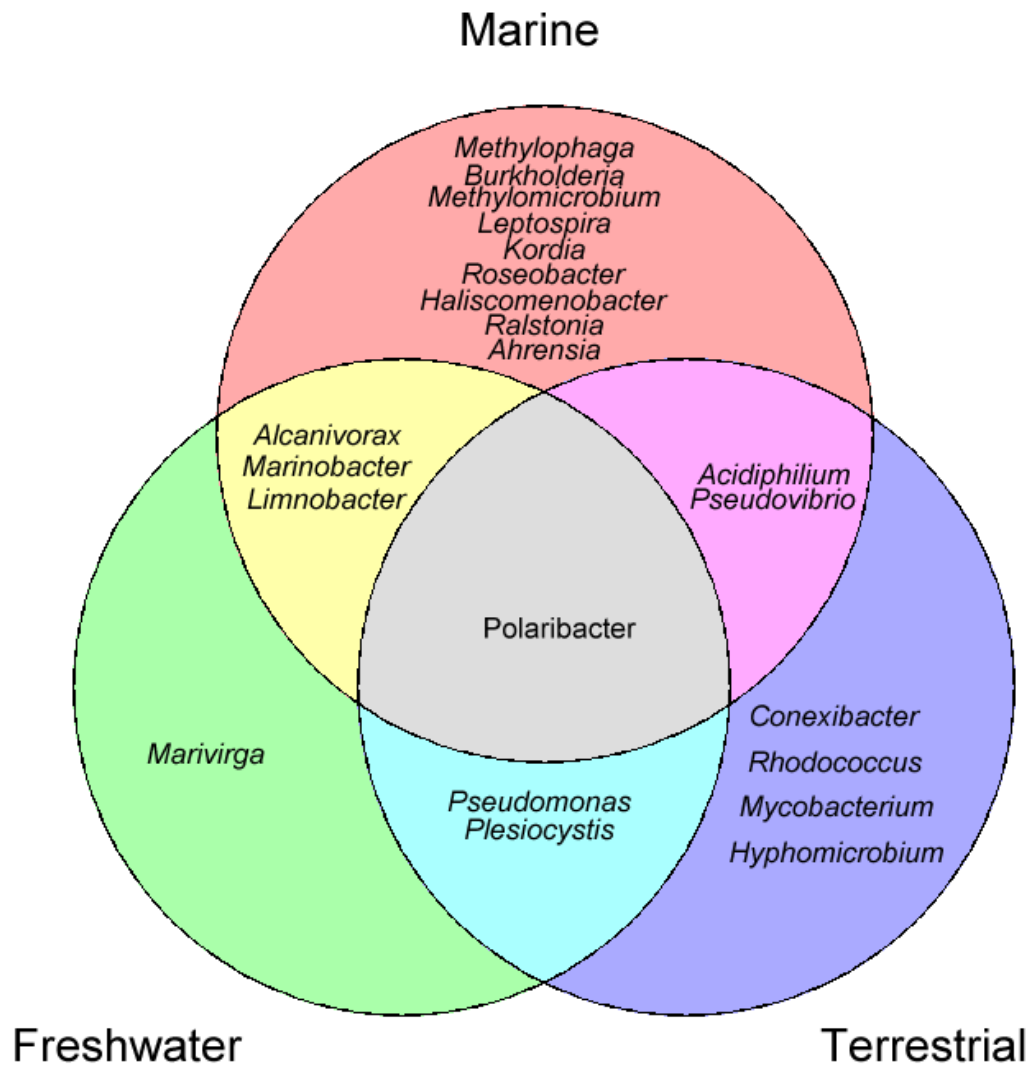
**Figure S9** Taxonomic distribution of candidate *alkB* genes from the three habitats metagenomic datasets. The blast results were imported into MEGAN soft for taxonomic analysis, using a last common ancestor (LCA) based algorithm.



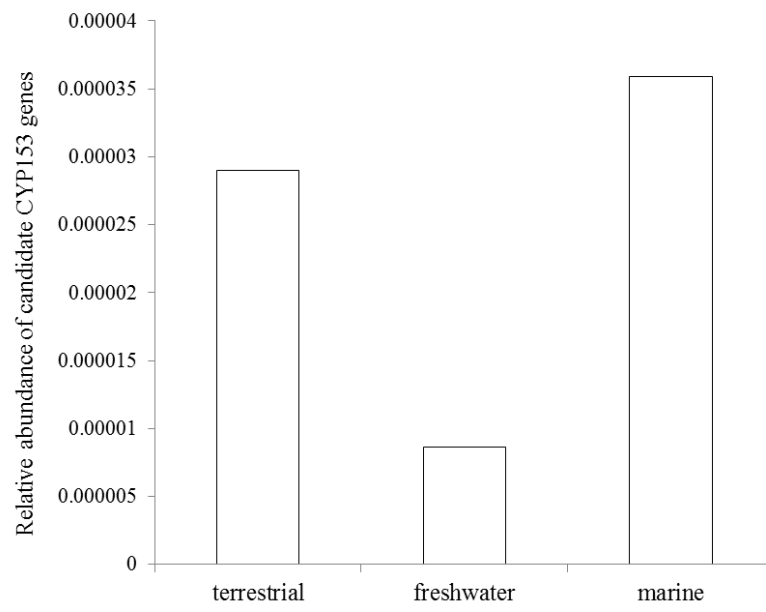
**Figure S10** Similarity distribution of candidate *alkB* genes in three habitats metagenomic datasets. All the sequences were investigated by BLASTP against non-redundant (NR) NCBI protein database (best BLAST hit, E-value  $\leq 1E-5$ ) and the result showed a wide spectrum of *alkB* genes in environments.



**Figure S11** Venn diagram of the distribution of candidate *alkB* sequences assigned to Genus in the three habitats metagenomic datasets.



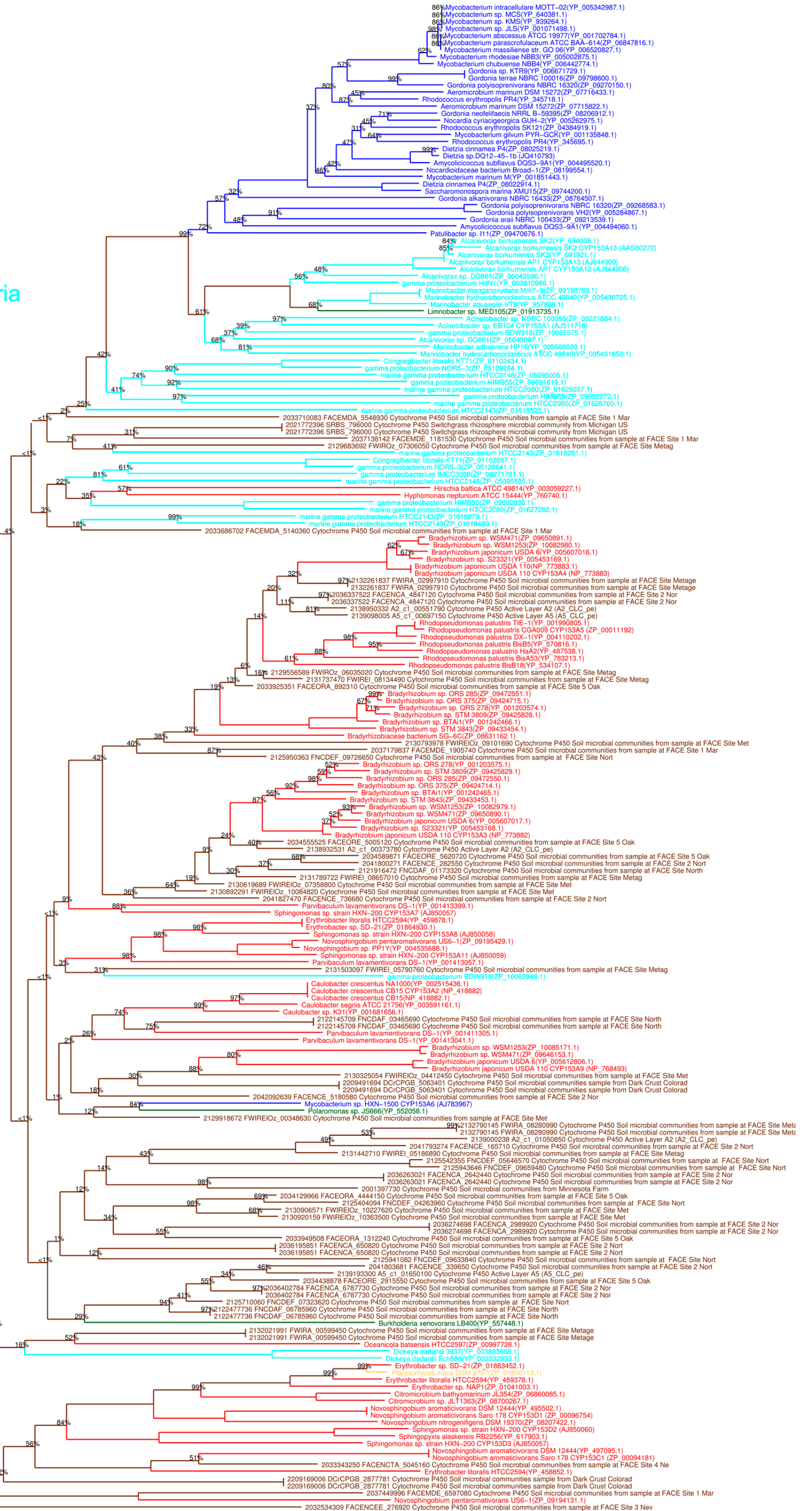
**Figure S12** The relative abundances of candidate CYP153 genes in metagenomes



**Figure S13** Phylogenetic tree of CYP153 genes in metagenomes together with those from microbial genomes based on amino acid sequences analysis. a: CYP153 from terrestrial metagenomes. b: CYP153 from fresh water metagenomes. c: CYP153 from marine metagenomes. Only the segments containing the conserved regions LL(I/V)VGGNDTTR(S/N) and R(W/Y)QTPLA(Y/H)MRR were used for alignment and to construct phylogenetic trees using ARB. The trees were bootstrapped with 1000 replicates, indicated at respective nodes. Genes from metagenomes are in brown

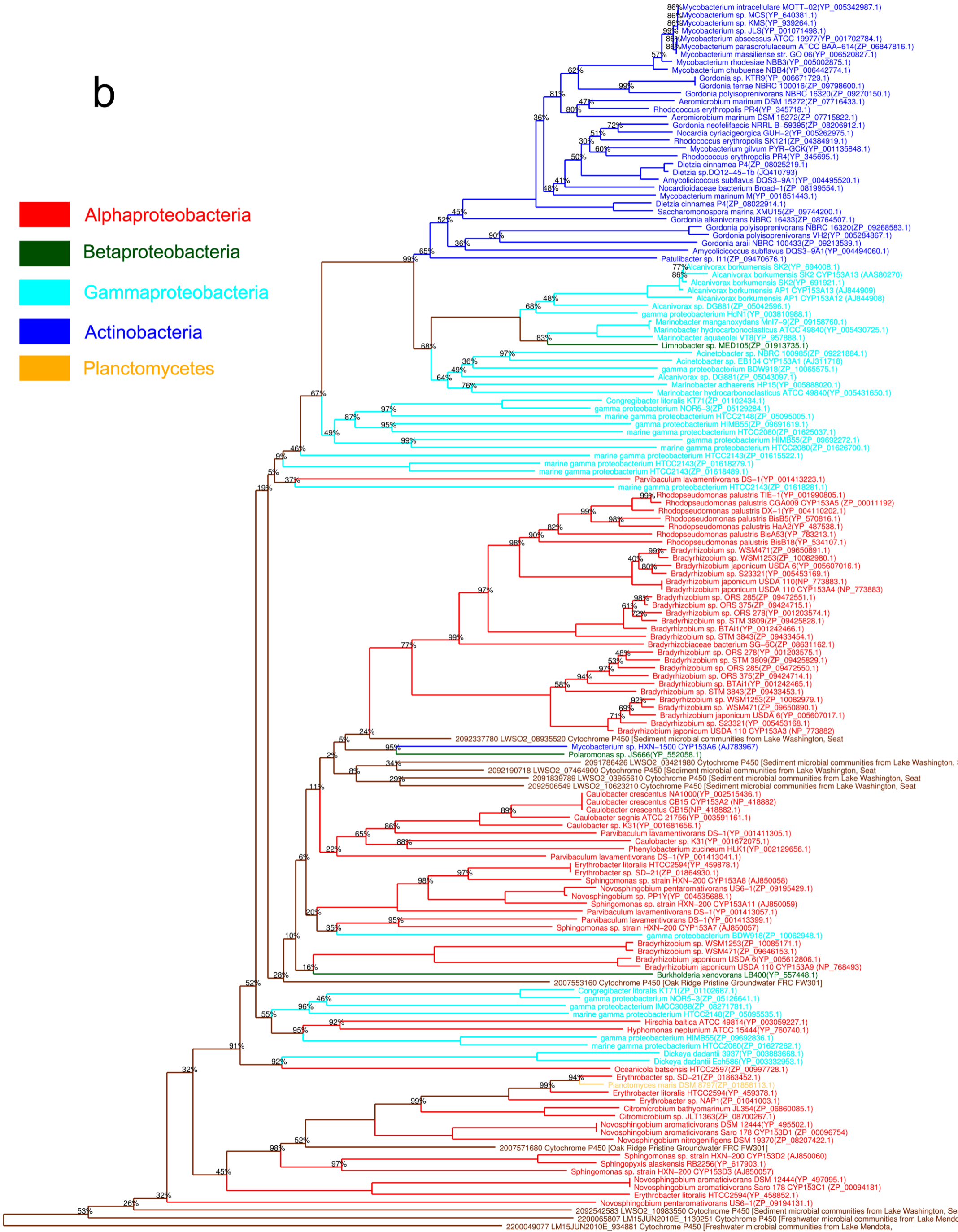
a

- Alphaproteobacteria
- Betaproteobacteria
- Gammaproteobacteria
- Actinobacteria
- Planctomycetes



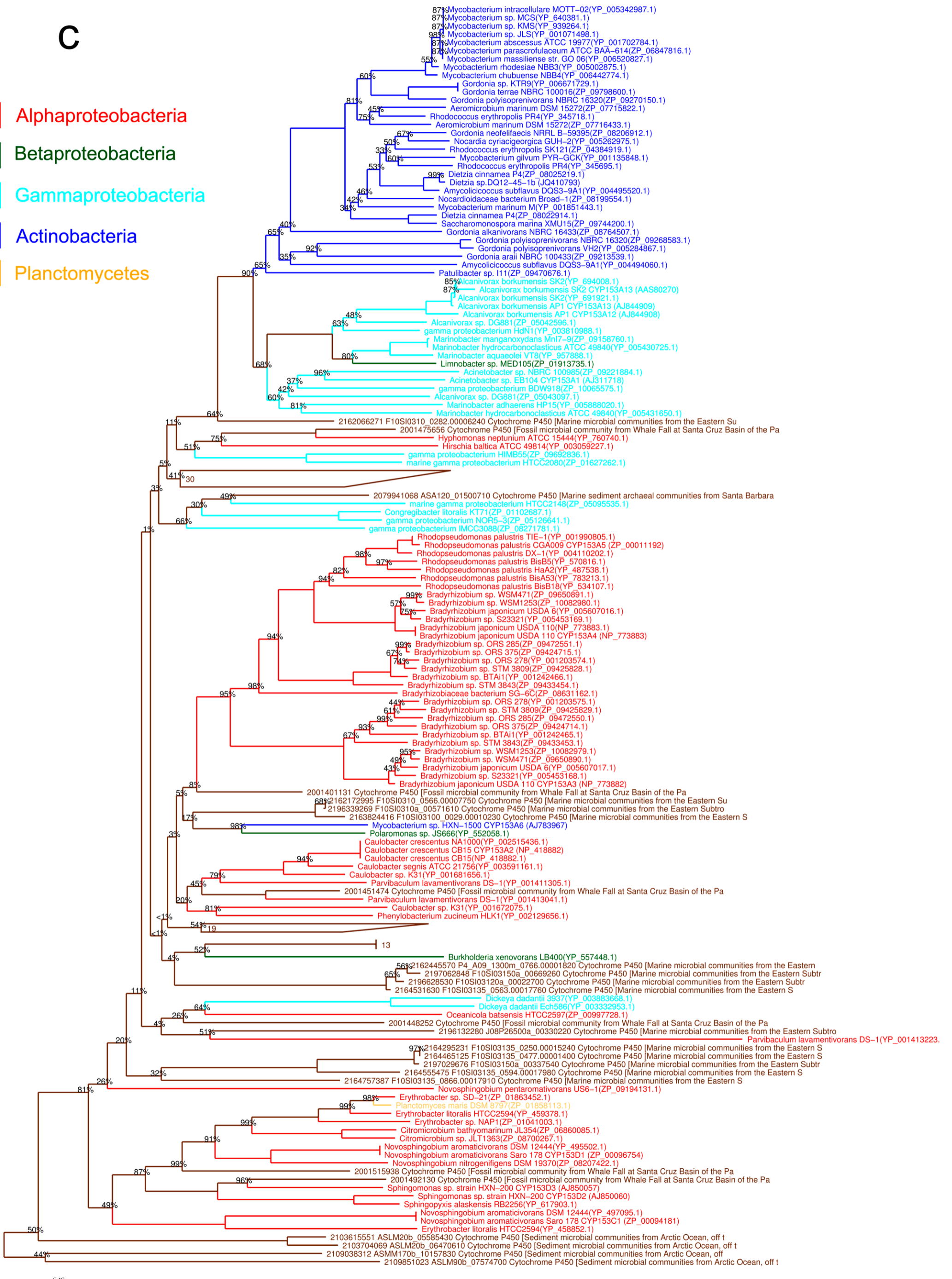
b

- Alphaproteobacteria
- Betaproteobacteria
- Gammaproteobacteria
- Actinobacteria
- Planctomycetes



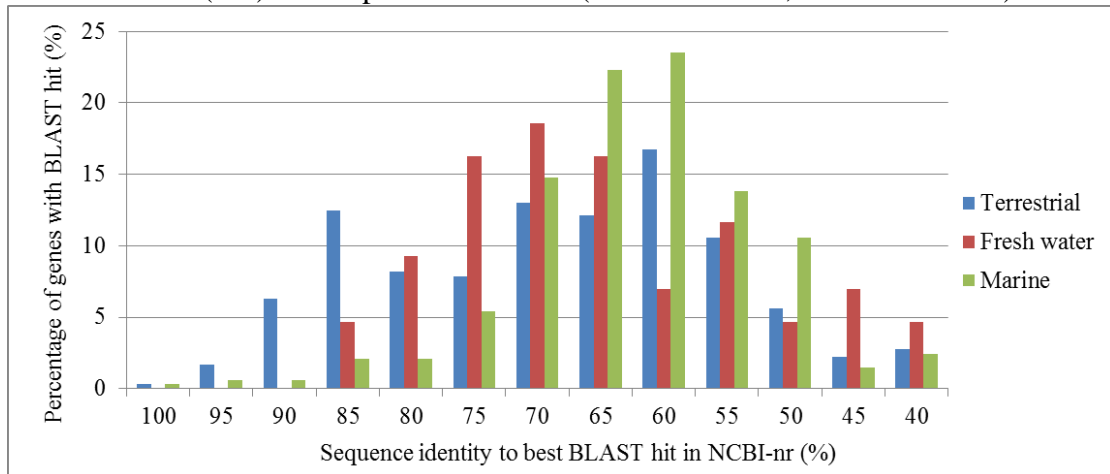
C

- Alphaproteobacteria
- Betaproteobacteria
- Gammaproteobacteria
- Actinobacteria
- Planctomycetes

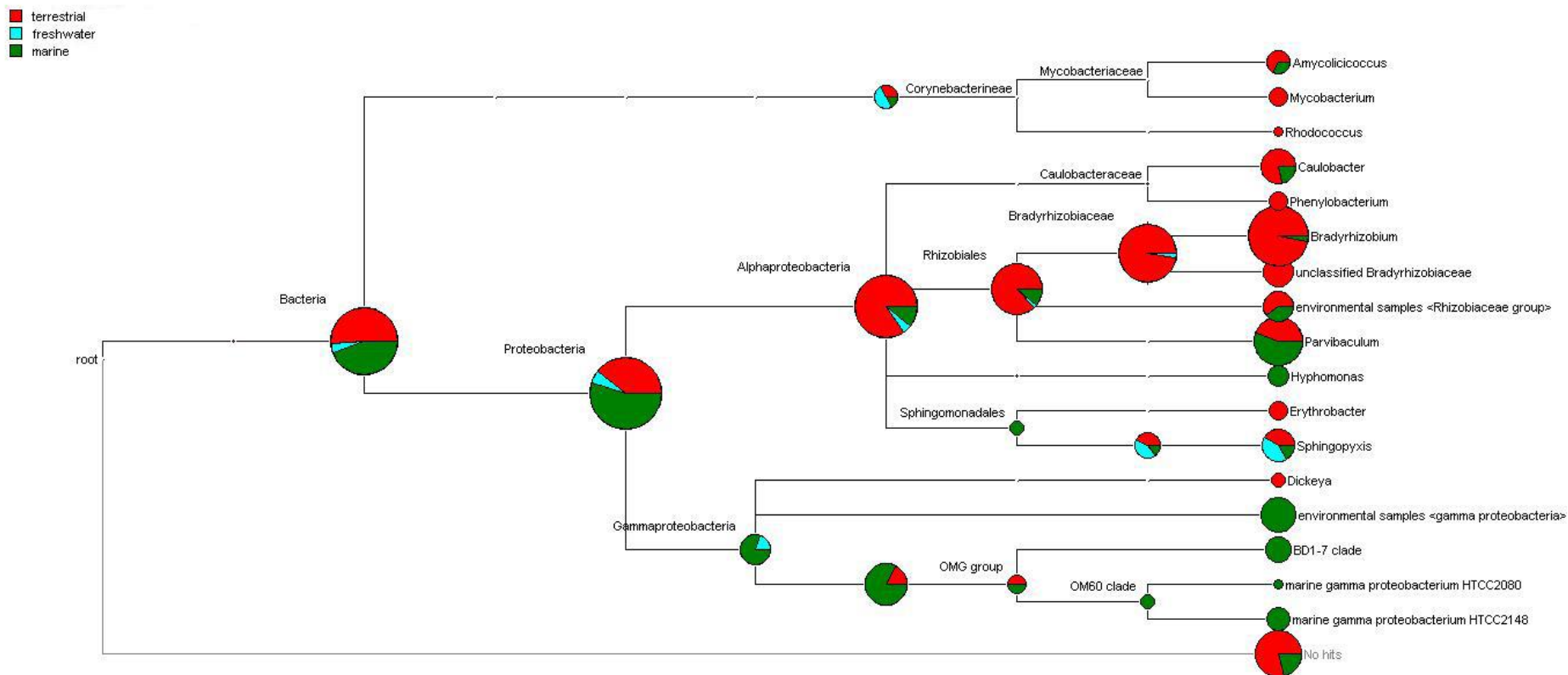




**Figure S14** Similarity distribution of candidate CYP153 genes in three habitats metagenomic datasets. All the sequences were investigated by BLASTP against non-redundant (NR) NCBI protein database (best BLAST hit, E-value  $\leq 1E-5$ ).



**Figure S15** Taxonomic distribution of candidate CYP153 genes from the three habitats metagenomic datasets. The blast results were imported into MEGAN soft for taxonomic analysis, using a last common ancestor (LCA) based algorithm



## References

- 1 Tani, A., Ishige, T., Sakai, Y. & Kato, N. Gene structures and regulation of the alkane hydroxylase complex in *Acinetobacter* sp. Strain M-1. *J. Bacteriol.* **183**, 1819-1823 (2001).
- 2 Maier, T., Forster, H. H., Asperger, O. & Hahn, U. Molecular characterization of the 56-kDa CYP153 from *Acinetobacter* sp. EB104. *Biochem. Biophys. Res. Commun.* **286**, 652-658 (2001).
- 3 Ratajczak, A., Geissdorfer, W. & Hillen, W. Alkane hydroxylase from *Acinetobacter* sp. strain ADP1 is encoded by alkM and belongs to a new family of bacterial integral-membrane hydrocarbon hydroxylases. *Appl. Environ. Microbiol.* **64**, 1175-1179 (1998).
- 4 van Beilen, J. B. *et al.* Cytochrome P450 alkane hydroxylases of the CYP153 family are common in alkane-degrading eubacteria lacking integral membrane alkane hydroxylases. *Appl. Environ. Microbiol.* **72**, 59-65 (2006).
- 5 Wang, W. & Shao, Z. Genes involved in alkane degradation in the *Alcanivorax hongdengensis* strain A-11-3. *Appl. Microbiol. Biotechnol.* **94**, 437-448 (2012).
- 6 Liu, C. *et al.* Multiple alkane hydroxylase systems in a marine alkane degrader, *Alcanivorax dieselolei* B-5. *Environ. Microbiol.* **13**, 1168-1178 (2011).
- 7 Liu, Y.-C. *et al.* Isolation of an alkane-degrading *Alcanivorax* sp. strain 2B5 and cloning of the alkB gene. *Bioresour. Technol.* **101**, 310-316, (2010).
- 8 Smits, T. H. M., Balada, S. B., Witholt, B. & van Beilen, J. B. Functional analysis of alkane hydroxylases from Gram-negative and Gram-positive bacteria. *J. Bacteriol.* **184**, 1733-1742 (2002).
- 9 Bihari, Z. *et al.* Functional analysis of long-chain n-alkane degradation by

- Dietzia* spp. *FEMS Microbiol. Lett.* **316**, 100-107 (2011).
- 10 Nie, Y., Liang, J., Fang, H., Tang, Y. Q. & Wu, X. L. Two novel alkane hydroxylase-rubredoxin fusion genes isolated from a *Dietzia* bacterium and the functions of fused rubredoxin domains in long-chain *n*-alkane degradation. *Appl. Environ. Microbiol.* **77**, 7279-7288 (2011).
- 11 Nie, Y., Liang, J.-L., Fang, H., Tang, Y.-Q. & Wu, X.-L. Characterization of a CYP153 alkane hydroxylase gene in a Gram-positive *Dietzia* sp. DQ12-45-1b and its “team role” with *alkW1* in alkane degradation. *Appl. Microbiol. Biotechnol.* **98**, 163-173 (2014).
- 12 Lo Piccolo, L., De Pasquale, C., Fodale, R., Puglia, A. M. & Quatrini, P. Involvement of an alkane hydroxylase system of *Gordonia* sp. strain SoCg in degradation of solid *n*-alkanes. *Appl. Environ. Microbiol.* **77**, 1204-1213 (2011).
- 13 Kotani, T., Kawashima, Y., Yurimoto, H., Kato, N. & Sakai, Y. Gene structure and regulation of alkane monooxygenases in propane-utilizing *Mycobacterium* sp. TY-6 and *Pseudonocardia* sp. TY-7. *J. Biosci. Bioeng.* **102**, 184-192 (2006).
- 14 Hamamura, N., Yeager, C. M. & Arp, D. J. Two distinct monooxygenases for alkane oxidation in *Nocardioides* sp. Strain CF8. *Appl. Environ. Microbiol.* **67**, 4992-4998 (2001).
- 15 Smits, T. H. M., Rothlisberger, M., Witholt, B. & van Beilen, J. B. Molecular screening for alkane hydroxylase genes in Gram-negative and Gram-positive strains. *Environ. Microbiol.* **1**, 307-317 (1999).
- 16 Amouric, A. *et al.* Identification of different alkane hydroxylase systems in *Rhodococcus ruberstrain* SP2B, an hexane-degrading actinomycete. *J. Appl.*

*Microbiol.* **108**, 1903-1916 (2009).

- 17 Whyte, L. G. *et al.* Gene cloning and characterization of multiple alkane hydroxylase systems in *Rhodococcus* strains Q15 and NRRL B-16531. *Appl. Environ. Microbiol.* **68**, 5933-5942 (2002).