Appendix

Table A-I: Percentage of cultivable bacterial phyla (bacterial isolates) retrieved from the RDP database of 16S rRNA gene inventory (https://rdp.cme.msu.edu/hierarchy/hb_intro.jsp) with a search keyword of "subsurface".

	Counts	Percentage
Total bacteria	**1637**	**100**
Proteobacteria	941	57.483
Firmicutes	329	20.098
Actinobacteria	305	18.632
Bacteroidetes	34	2.077
Aquificae	11	0.672
unclassified_Bacteria	6	0.367
Thermotogae	4	0.244
Chloroflexi	2	0.122
Deinococcus-Thermus	2	0.122
Fusobacteria	1	0.061
Tenericutes	1	0.061
Verrucomicrobia	1	0.061

Note: The RDP database search with word "Subsurface" and data set options selected as follow: 1) Strain: Both (type and non-type) 2) Source: Isolates 3) Size: Both (≥1200, <1200) 4) Quality: Good. Data retrieved on date 15-06-2020.

Table A-II: Percentage of unculturable bacterial phyla retrieved from the RDP database of 16S rRNA gene inventory (https://rdp.cme.msu.edu/hierarchy/hb_intro.jsp) with a search keyword of "subsurface".

	Counts	Percentage
Total bacteria	**34119**	**100**
Bacteroidetes	16867	49.436
Proteobacteria	9610	28.166
Firmicutes	2549	7.471
unclassified_Bacteria	1788	5.240
Acidobacteria	847	2.482
Actinobacteria	558	1.635
Chloroflexi	522	1.530
Planctomycetes	347	1.017
Aquificae	45	0.132
Thermotogae	9	0.026
Deinococcus-Thermus	48	0.141
Fusobacteria	9	0.026
Tenericutes	12	0.035
Verrucomicrobia	100	0.293
Ignavibacteriae	130	0.381
Nitrospirae	102	0.299
Spirochaetes	85	0.249
Gemmatimonadetes	75	0.220
Aminicenantes	66	0.193
Candidatus Saccharibacteria	54	0.158
Cyanobacteria/Chloroplast	50	0.147
Armatimonadetes	34	0.100
Elusimicrobia	24	0.070
Parcubacteria	24	0.070
Synergistetes	20	0.059
Latescibacteria	20	0.059
Caldiserica	18	0.053
SR1	16	0.047
Lentisphaerae	13	0.038
Deferribacteres	11	0.032
Microgenomates	11	0.032
Chlorobi	6	0.018
BRC1	6	0.018
Marinimicrobia	6	0.018
Fibrobacteres	5	0.015
Candidate division ZB3	5	0.015
Thermodesulfobacteria	4	0.012
Omnitrophica	4	0.012
Candidate division WPS-1	4	0.012
Acetothermia	3	0.009
Cloacimonetes	3	0.009
Chlamydiae	2	0.006
Hydrogenedentes	2	0.006
Nitrospinae	2	0.006
Chrysiogenetes	1	0.003
Atribacteria	1	0.003
Candidate division WPS-2	1	0.003

Note: The RDP database search with word "Subsurface" and data set options selected as follow: 1) Strain: Both (type and non-type) 2) Source: Unculturable 3) Size: Both (≥1200, <1200) 4) Quality: Good; data retrieved on date 15-06-2020.

Table A-III: Summary table of sequence similarity analysis of clones used in qPCR standard curve preparation

Insert (clone) sequence identity	**Best match sequence from NCBI database (with accession number)**	**Query cover**	**E-value**	**Percentage identity**
Archaeal 16S rRNA gene	Uncultured archaeon clone FAIR_SPR_A08E (KC437286.1)	100%	0.0	97.45%
NirS gene	*Pseudomonas stutzeri* strain SLG510A3-8, complete genome- From 3832039 to 3832431 nucleotide [cytochrome cd1 nitrite reductase (nirS) gene Sequence] (CP011854.1)	92%	0.0	99.22%
DsrB gene	Uncultured bacterium clone SJ20-22 dissimilatory sulfite reductase beta subunit (dsrB) gene (MF068050.1)	98%	6e-174	99.13%
CzcA gene	Acidovorax sp. KKS102, complete genome; From 1265099 to 1265301 nucleotide (encode heavy metal efflux pump, czca family protein)(CP003872.1)	87%	8e-62	88.67%

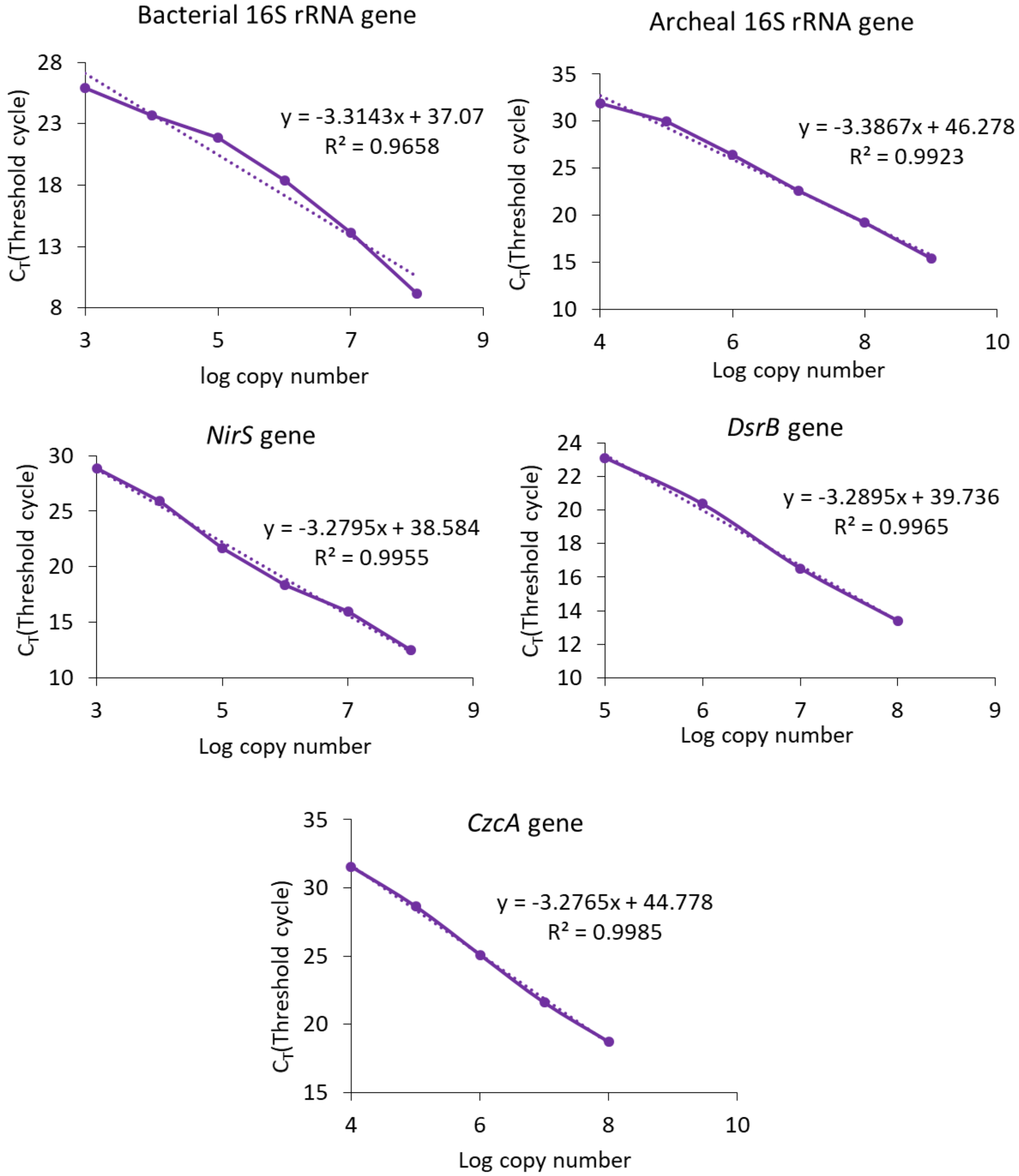
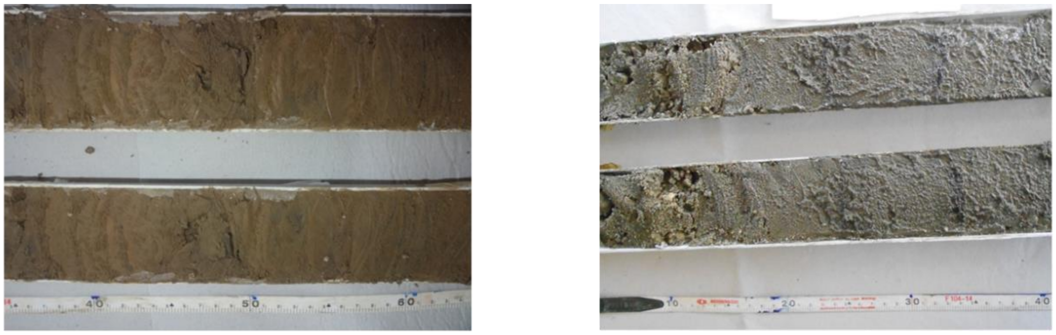


Figure A-I: Standard plots used for absolute quantification of different genes by qPCR.



Estuarine sequence showing thin alternating silt and clay layers.

Core pipe showing fine sand with silt horizons.

Gravel deposit at the base of estuarine sequence and the underlying fluvial sediment deposits

The lower fluvial sequence showing the semi-compacted nature of the sediments and palaeosol.

Figure A-II: Representative photographs of splited core liner showing the lithology of sediments present within CRD core.

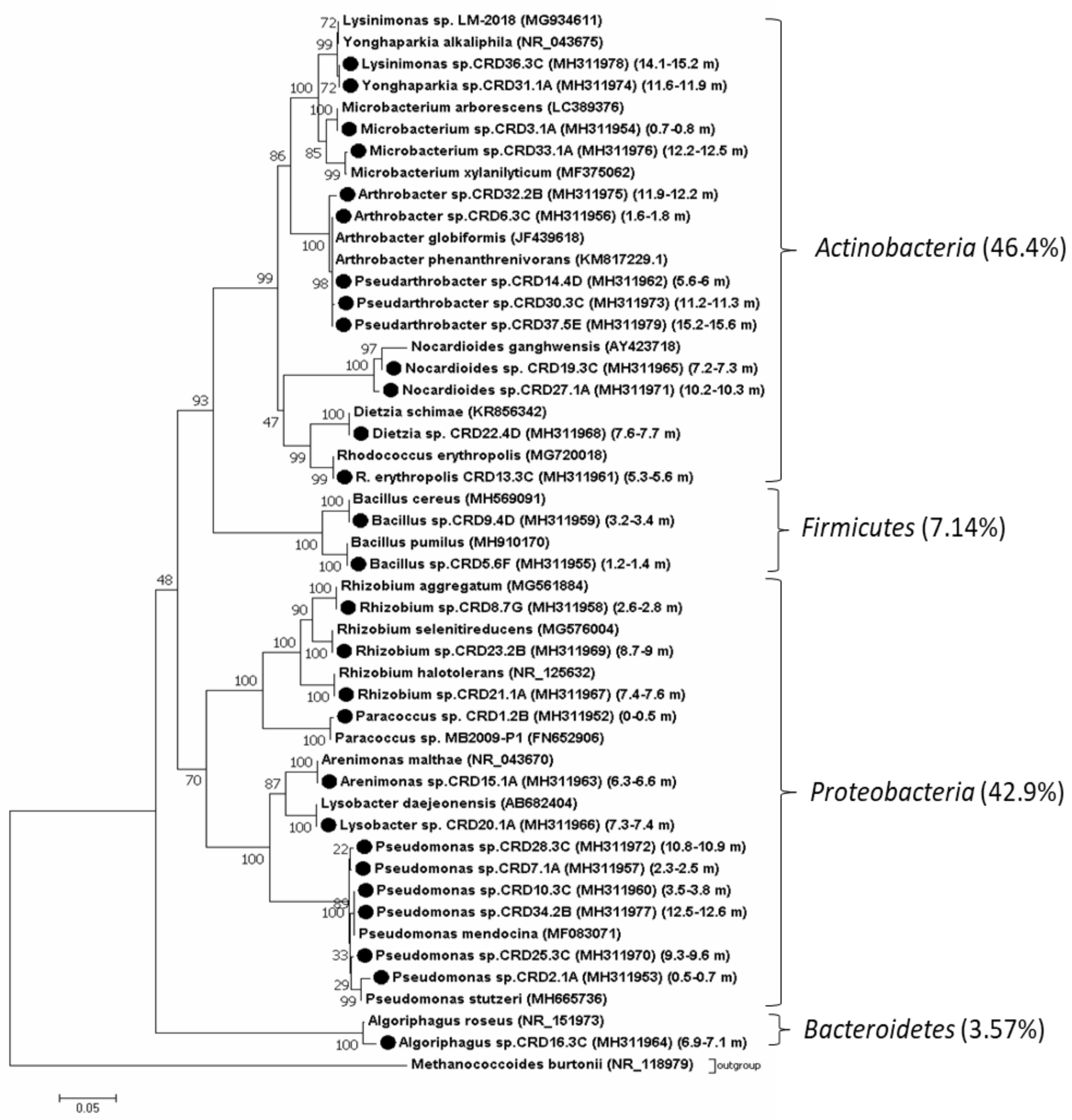


Figure A-III: The phylogenetic relatedness (based on 16S rRNA gene partial sequence) of heterotrophic isolates obtained from different subsurface horizons of the Chokari (CRD) core. The phylogenetic tree was contracted using the Neighbor-Joining method with the Jukes-Cantor algorithm. Isolates obtained in this study are shown in the black circle. Codes in left parentheses correspond to GenBank sequence accession numbers and codes in right parentheses correspond to depth intervals (cm) from which isolates were obtained. Numbers at nodes depict the bootstrap values analyzed with 1000 replicates. *Methanococcoides burtonii* (NR_ 118979) was used as an outgroup. Scale bar represents the number of base substitutions per site.

Silty sand and red palaeosol horizons observed within RYD core

Gravelly layer observed within RYD core

Sand and silt as well as red palaeosol horizons observed within RMD core

Gravelly layer and clay silt horizons observed within RMD core

Figure A-IV: Representative photographs of splited core liner showing the lithology of sediments present within RYD and RMD cores.

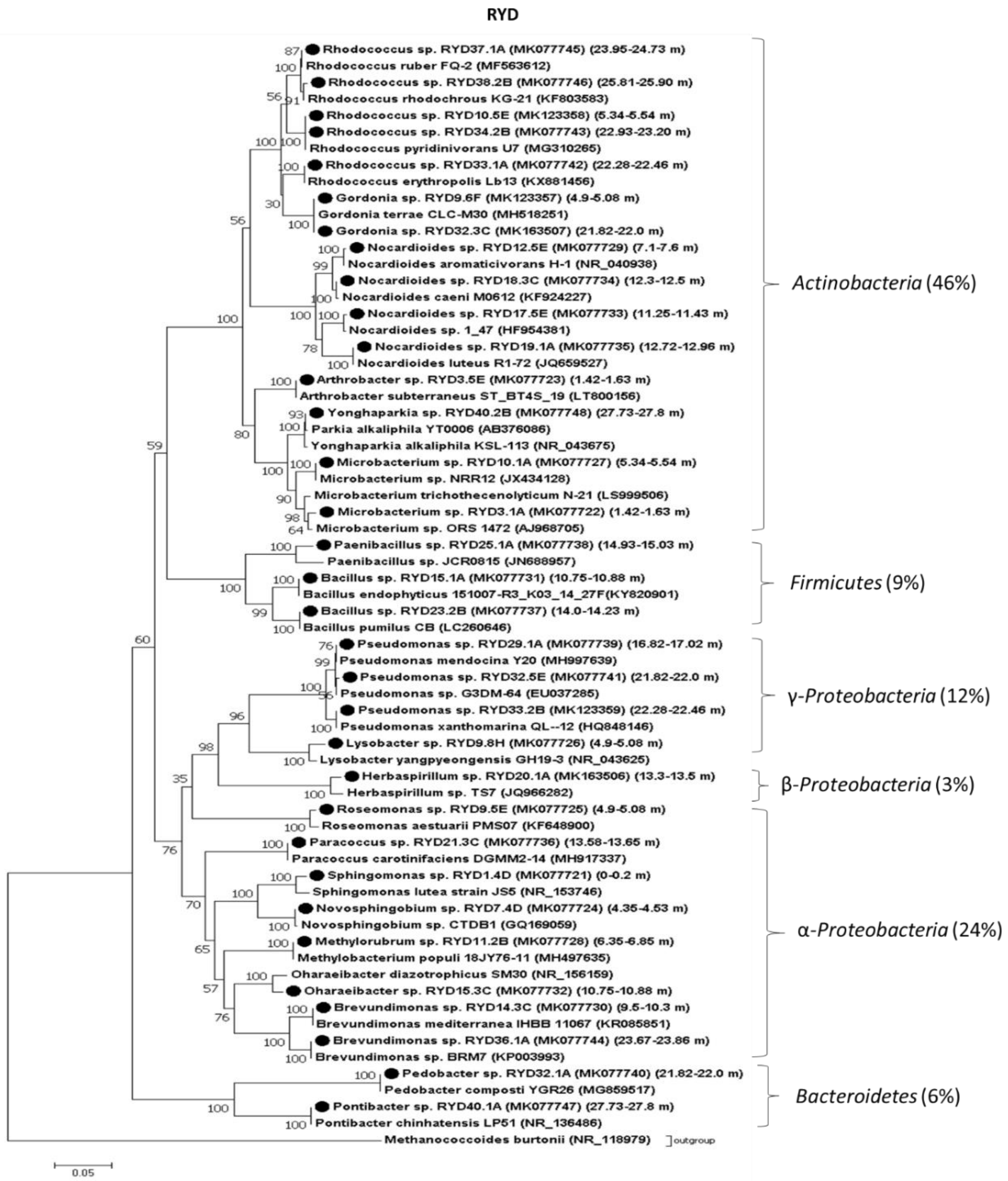


Figure A-V: The phylogenetic relatedness (based on 16S rRNA gene partial sequence) of heterotrophic isolates obtained from different subsurface horizons of the Rayka (RYD) core. The phylogenetic tree was contracted using the Neighbor-Joining method with the Jukes-Cantor algorithm. Isolates obtained in this study are shown in the black circle. Codes in left parentheses correspond to GenBank sequence accession numbers and codes in right parentheses correspond to the depth from which isolates were obtained. Numbers at nodes depict the bootstrap values analyzed with 1000 replicates. *Methanococcoides burtonii* (NR_ 118979) was used as an outgroup. Scale bar represents the number of base substitutions per site.

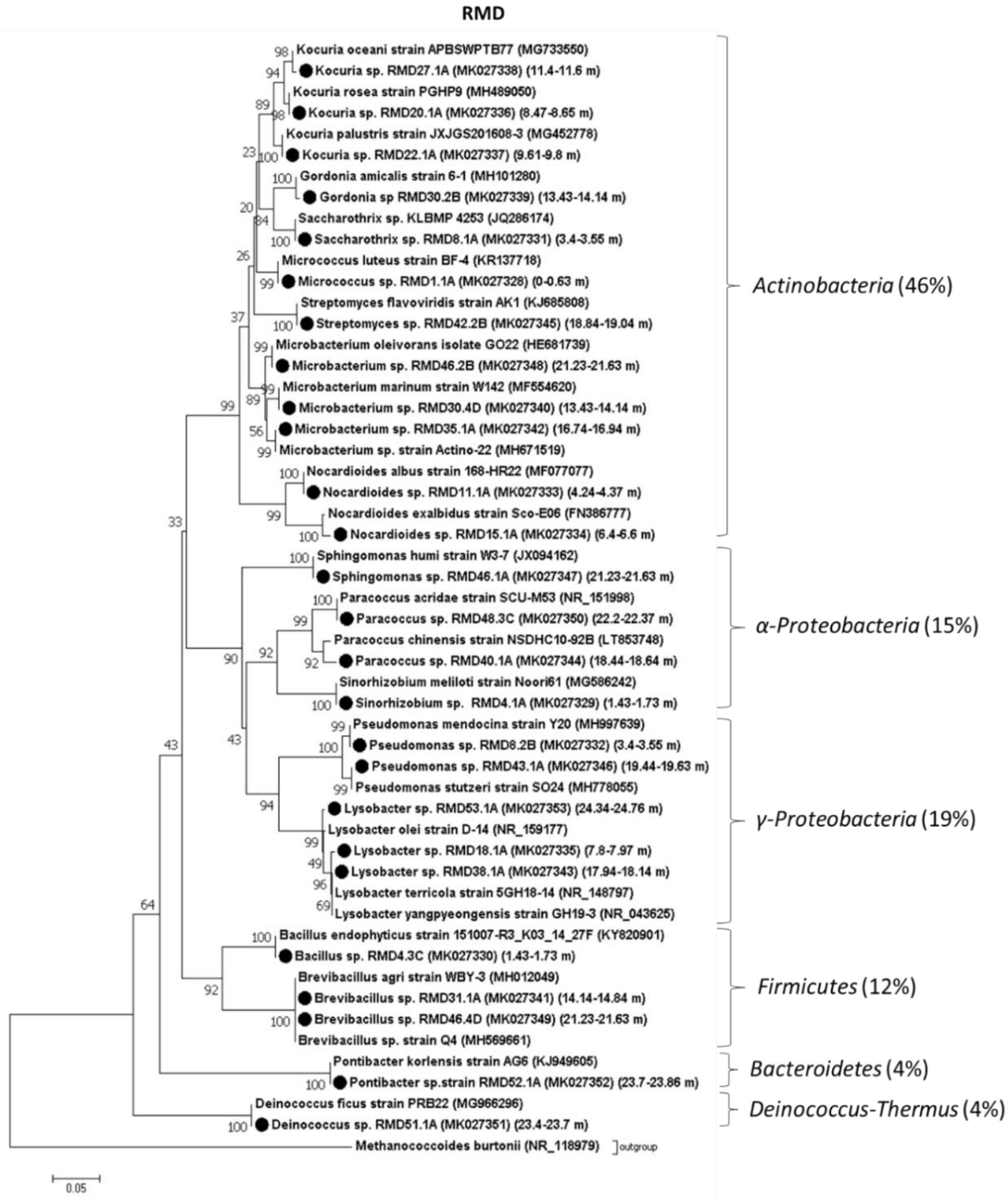


Figure A-VI: The phylogenetic relatedness (based on 16S rRNA gene partial sequence) of heterotrophic isolates obtained from different subsurface horizons of the Rampura (RMD) core. The phylogenetic tree was contracted using the Neighbor-Joining method with the Jukes-Cantor algorithm. Isolates obtained in this study are shown in the black circle. Codes in left parentheses correspond to GenBank sequence accession numbers and codes in right parentheses correspond to the depth from which isolates were obtained. Numbers at nodes depict the bootstrap values analyzed with 1000 replicates. *Methanococcoides burtonii* (NR_ 118979) was used as an outgroup. Scale bar represents the number of base substitutions per site.