Supplementary Material

Figure S1: Overview of the upstream (A) and downstream (B) analysis pipeline. A detailed description of the single analysis steps is provided in the Materials and Methods chapter.

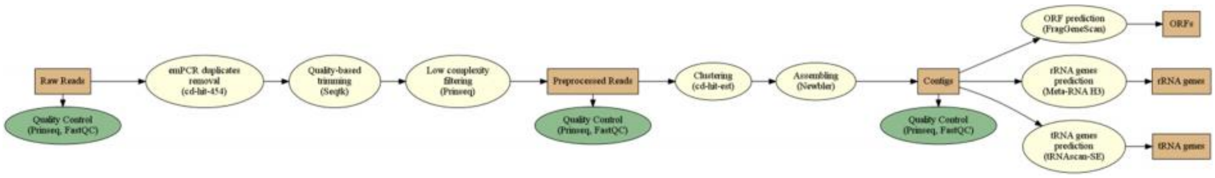
Figure S2: Collectors curves and Shannon diversity indices for the tested metagenomes from Sisters Peak (SP), Juan de Fuca (JdF), Mariana Trough (MT) and Brazos Trinity Basin (BTB) based on taxonomic assignments at species, genus, family, order, class and phylum level.

Figure S3: Analysis of the variability of SP metagenome sequences recruited to the reference genomes of *Aciduliprofundum* sp. MAR08-339 and *Aciduliprofundum boonei* T469.

Figure S4: Recruitment plots of SP metagenome fragments to the reference genomes of *Aciduliprofundum* sp. MAR08-339 (A), *Aciduliprofundum boonei* T469 (B), *Hippea maritima* DSM10411 (C), *Caldisericum exile* AZM16c01 (D) and *Niastella koreensis* GM20-10 (E). From the outside to the inside: the outer two tracks (track 1 and 2) are protein encoding genes, color-coded according to COG categories, on the forward and reverse strand of the reference genome. Track 3 denotes tRNAs (blue) and rRNAs (red) genes in the reference genome. Track 4 indicates the GC content of the reference genome, where red and blue coloring mark GC content, respectively, above or below the overall GC content of the reference genome. Green bars in track 5 indicate IslandViewer genomic island predictions. The scatter plot in track 6 represents the fragment recruitment of the SP metagenome assembly to the reference genome. Each dot represents the maximum identity value among the hits for a window of 30 bp. Red and blue dots represent blastn hits and tblastx hits, respectively. The y-axis goes from 75% (inside) to 100% identity (outside); a green line is drawn at the 95% level.

Figure S1

A


Raw Reads
Quality Control (Prinseq, FastQC)
emPCR duplicates removal (cd-hit-454)
Quality-based trimming (Seqtk)
Low complexity filtering (Prinseq)
Preprocessed Reads
Quality Control (Prinseq, FastQC)
Clustering (cd-hit-est)
Assembling (Newbler)
Contigs
Quality Control (Prinseq, FastQC)
ORF prediction (FragGeneScan)
ORFs
rRNA genes prediction (Meta-RNA H3)
rRNA genes
tRNA genes prediction (tRNAscan-SE)
tRNA genes

B

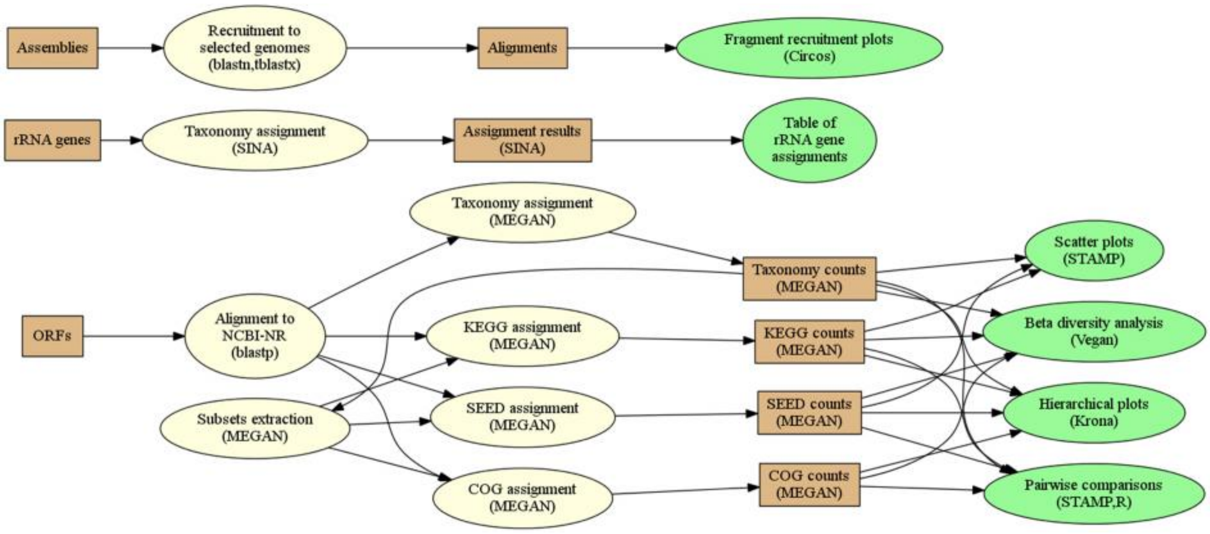
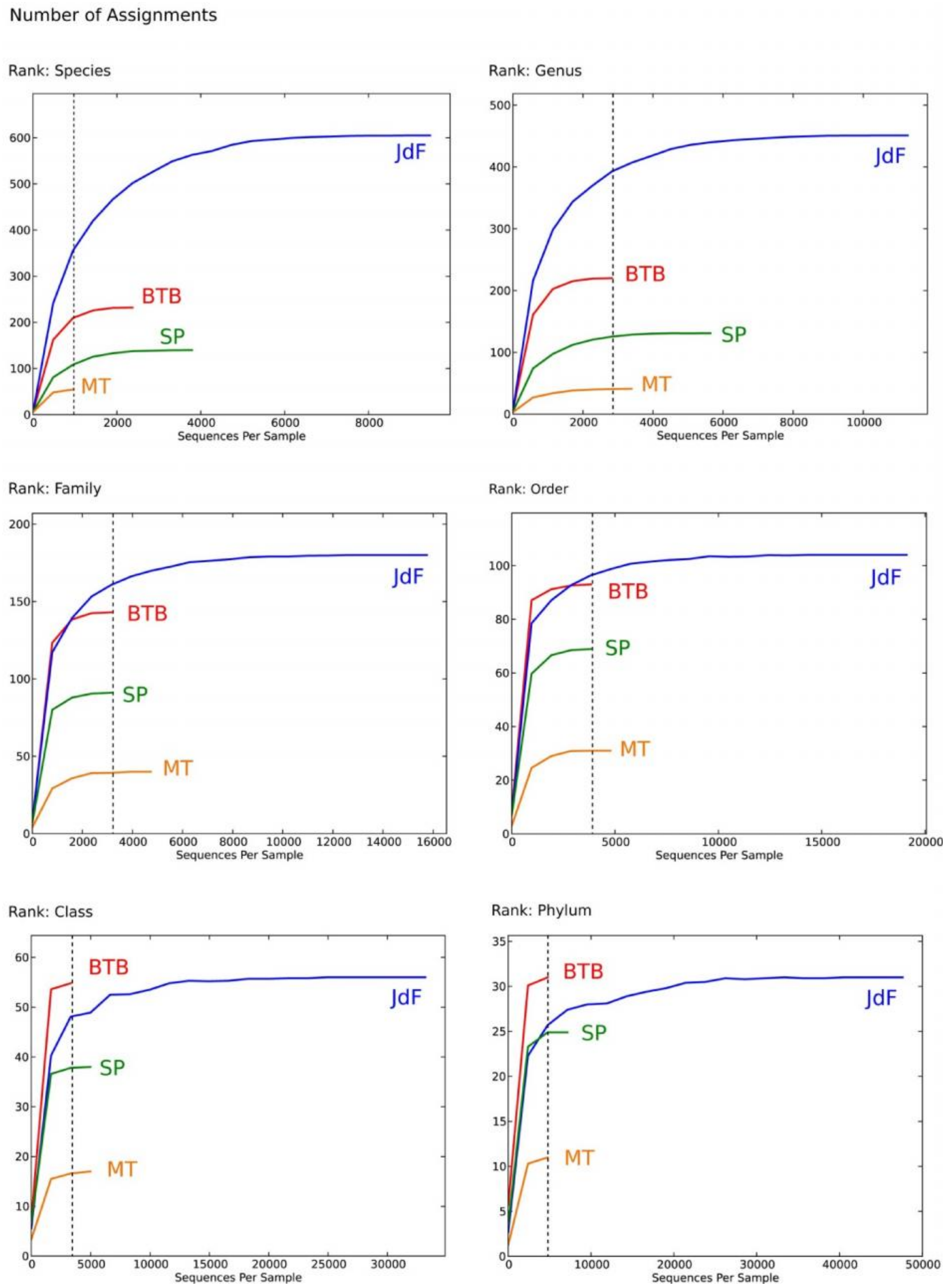

Assemblies
Recruitment to selected genomes (blastn,tblastx)
Alignments
Fragment recruitment plots (Circos)
rRNA genes
Taxonomy assignment (SINA)
Assignment results (SINA)
Table of rRNA gene assignments
Taxonomy assignment (MEGAN)
Taxonomy counts (MEGAN)
ORFs
Alignment to NCBI-NR (blastp)
KEGG assignment (MEGAN)
KEGG counts (MEGAN)
Subsets extraction (MEGAN)
SEED assignment (MEGAN)
SEED counts (MEGAN)
COG assignment (MEGAN)
COG counts (MEGAN)
Scatter plots (STAMP)
Beta diversity analysis (Vegan)
Hierarchical plots (Krona)
Pairwise comparisons (STAMP,R)

Figure S2


Number of Assignments
Rank: Species
0
100
200
300
400
500
600
0
2000
4000
6000
8000
Sequences Per Sample
JdF
BTB
SP
MT
Rank: Genus
0
100
200
300
400
500
0
2000
4000
6000
8000
10000
Sequences Per Sample
JdF
BTB
SP
MT
Rank: Family
0
50
100
150
200
0
2000
4000
6000
8000
10000
12000
14000
16000
Sequences Per Sample
JdF
BTB
SP
MT
Rank: Order
0
20
40
60
80
100
0
5000
10000
15000
20000
Sequences Per Sample
JdF
BTB
SP
MT
Rank: Class
0
10
20
30
40
50
60
0
5000
10000
15000
20000
25000
30000
Sequences Per Sample
BTB
JdF
SP
MT
Rank: Phylum
0
5
10
15
20
25
30
35
0
10000
20000
30000
40000
50000
Sequences Per Sample
BTB
JdF
SP
MT

Shannon index

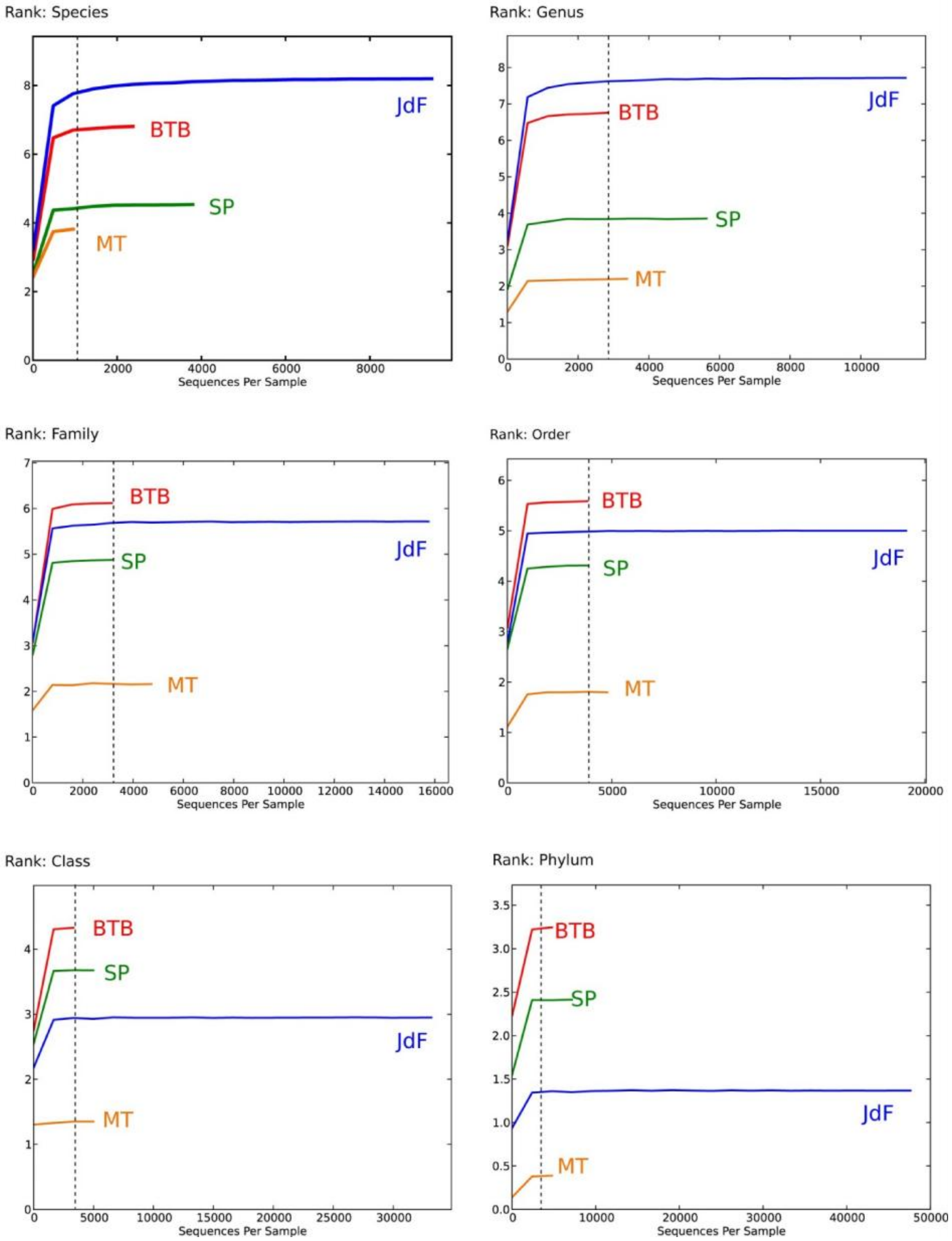
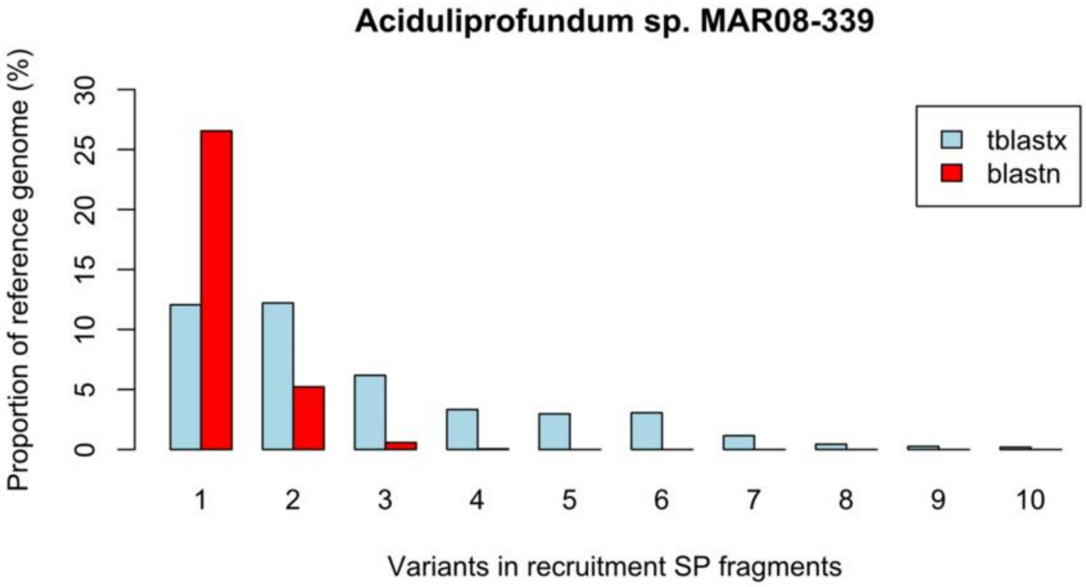


Figure S3


Aciduliprofundum sp. MAR08-339
Proportion of reference genome (%)
0
5
10
15
20
25
30
tblastx
blastn
1
2
3
4
5
6
7
8
9
10
Variants in recruitment SP fragments

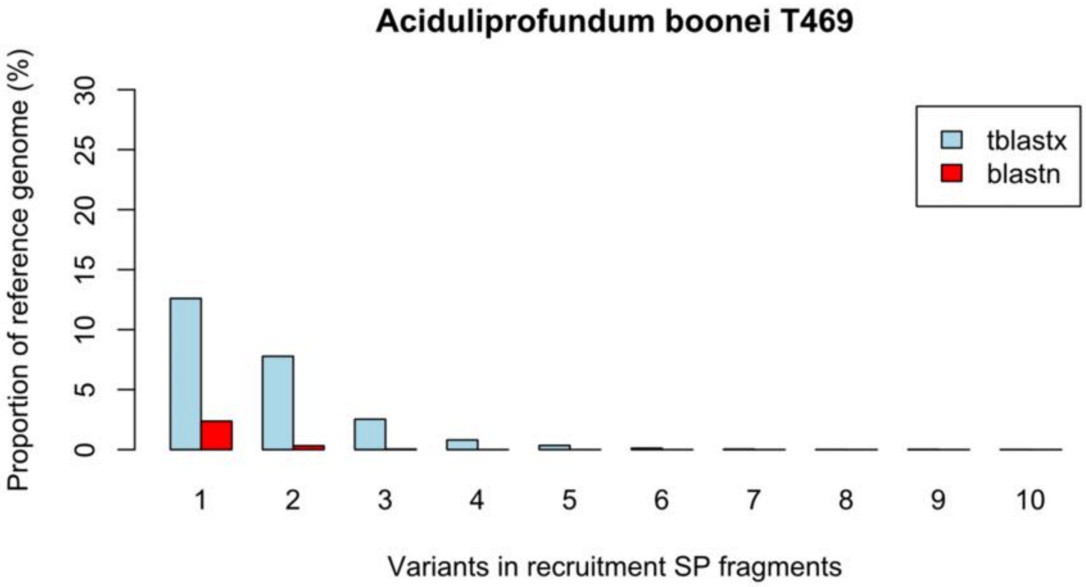
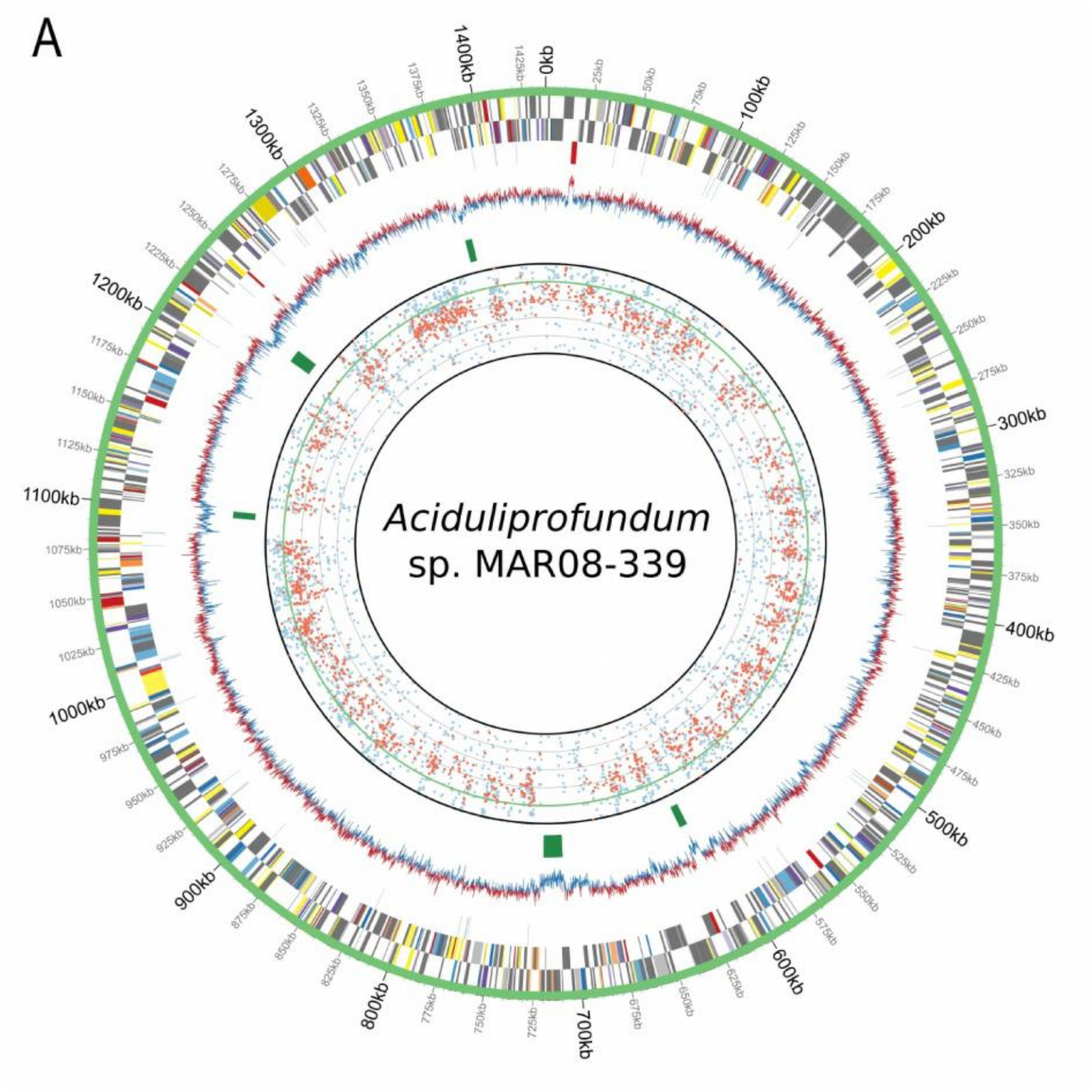
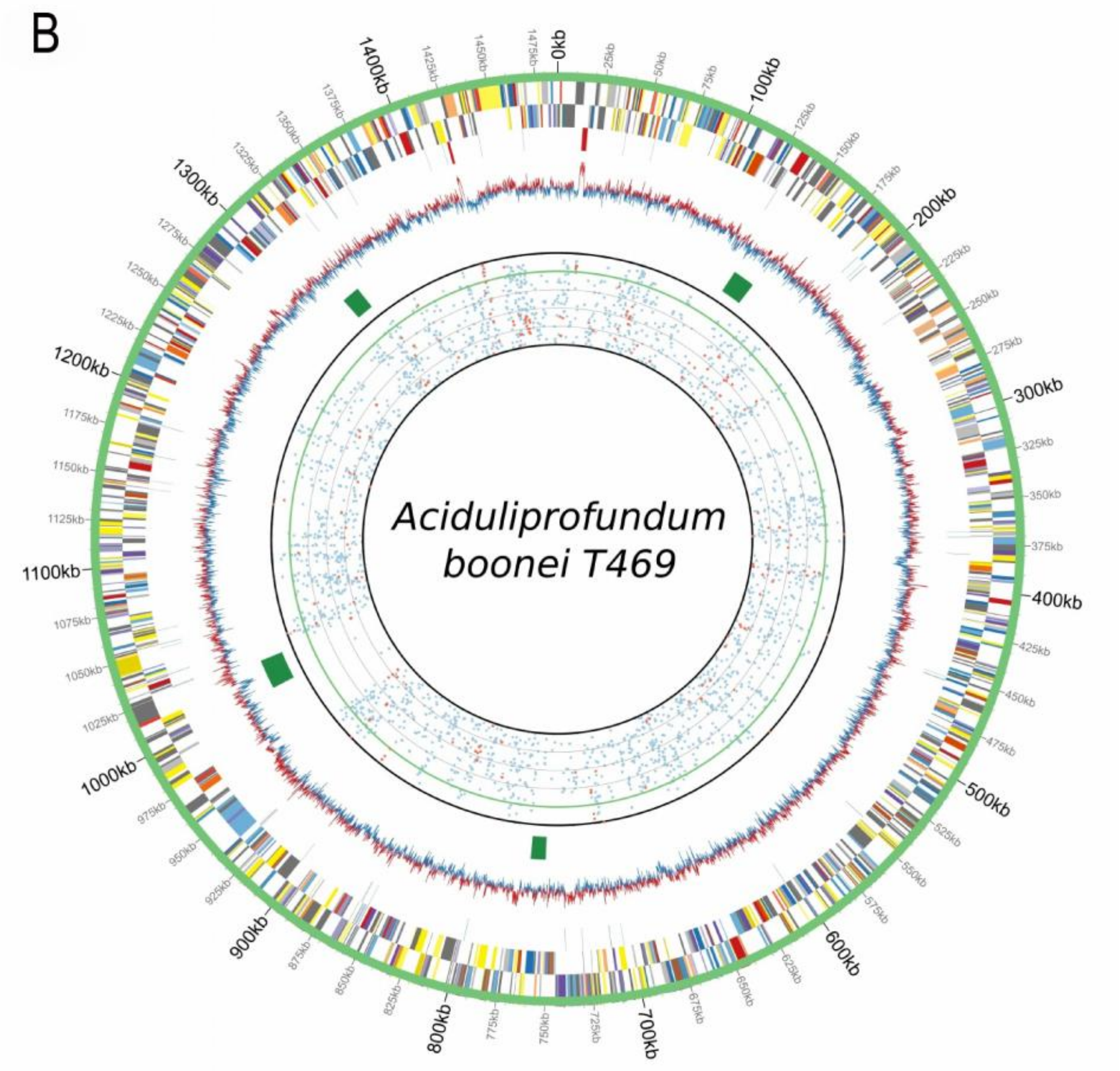
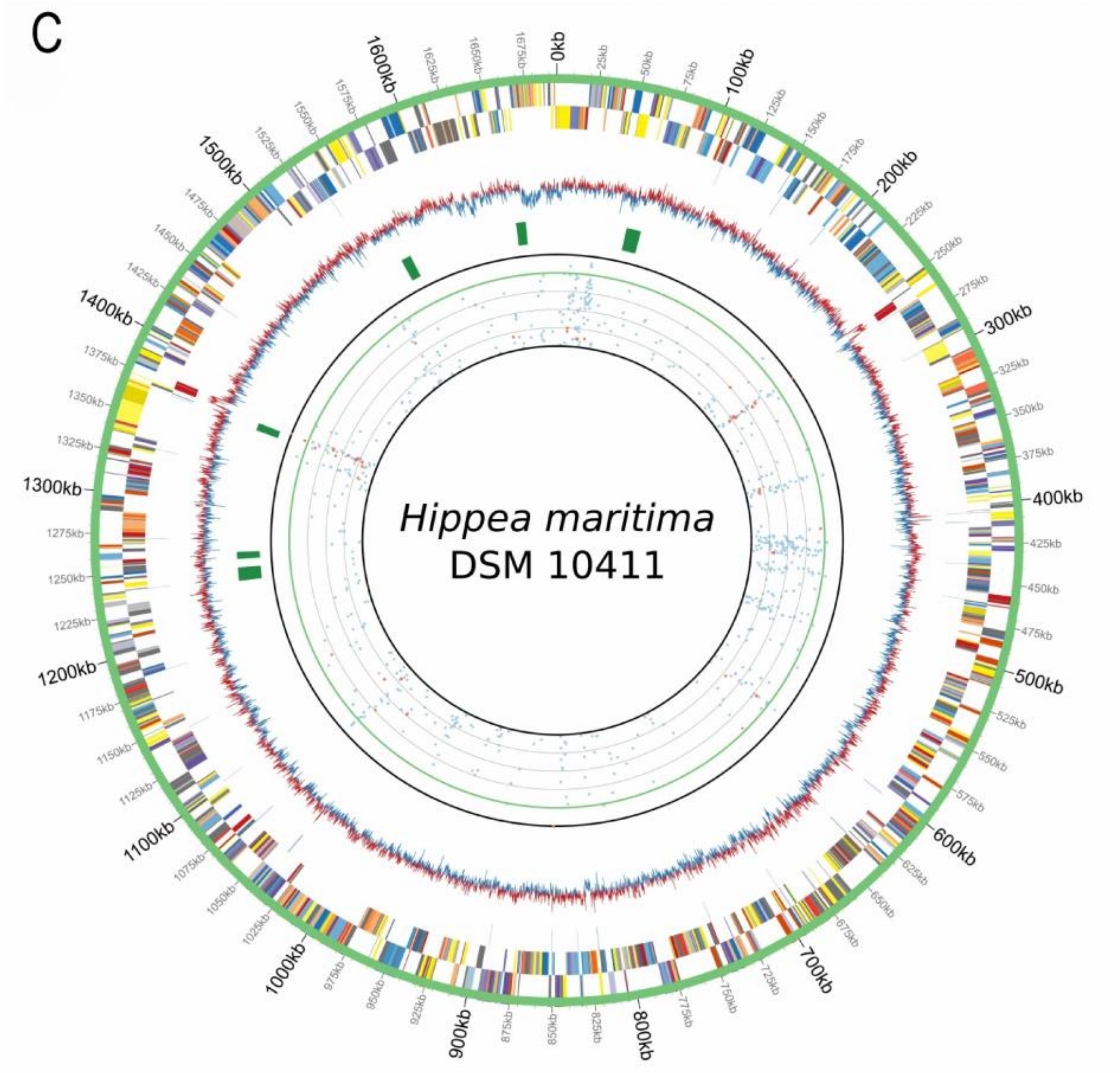

Aciduliprofundum boonei T469
Proportion of reference genome (%)
0
5
10
15
20
25
30
tblastx
blastn
1
2
3
4
5
6
7
8
9
10
Variants in recruitment SP fragments

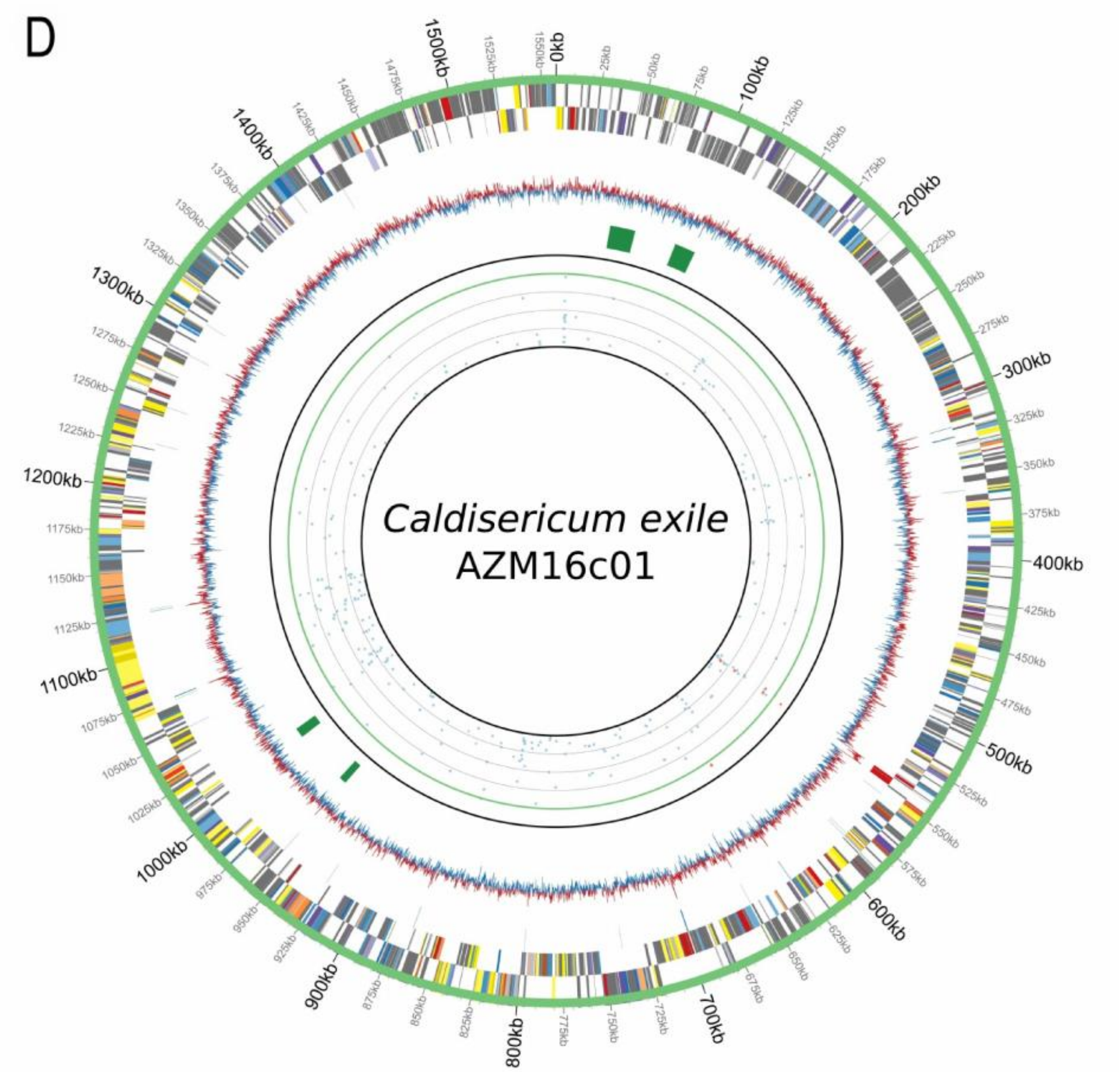
Figure S4


A
Aciduliprofundum sp. MAR08-339
0kb
25kb
50kb
75kb
100kb
125kb
150kb
175kb
200kb
225kb
250kb
275kb
300kb
325kb
350kb
375kb
400kb
425kb
450kb
475kb
500kb
525kb
550kb
575kb
600kb
625kb
650kb
675kb
700kb
725kb
750kb
775kb
800kb
825kb
850kb
875kb
900kb
925kb
950kb
975kb
1000kb
1025kb
1050kb
1075kb
1100kb
1125kb
1150kb
1175kb
1200kb
1225kb
1250kb
1275kb
1300kb
1325kb
1350kb
1375kb
1400kb
1425kb


B
Aciduliprofundum boonei T469
0kb
25kb
50kb
75kb
100kb
125kb
150kb
175kb
200kb
225kb
250kb
275kb
300kb
325kb
350kb
375kb
400kb
425kb
450kb
475kb
500kb
525kb
550kb
575kb
600kb
625kb
650kb
675kb
700kb
725kb
750kb
775kb
800kb
825kb
850kb
875kb
900kb
925kb
950kb
975kb
1000kb
1025kb
1050kb
1075kb
1100kb
1125kb
1150kb
1175kb
1200kb
1225kb
1250kb
1275kb
1300kb
1325kb
1350kb
1375kb
1400kb
1425kb
1450kb
1475kb


C
Hippea maritima
DSM 10411
0kb
25kb
50kb
75kb
100kb
125kb
150kb
175kb
200kb
225kb
250kb
275kb
300kb
325kb
350kb
375kb
400kb
425kb
450kb
475kb
500kb
525kb
550kb
575kb
600kb
625kb
650kb
675kb
700kb
725kb
750kb
775kb
800kb
825kb
850kb
875kb
900kb
925kb
950kb
975kb
1000kb
1025kb
1050kb
1075kb
1100kb
1125kb
1150kb
1175kb
1200kb
1225kb
1250kb
1275kb
1300kb
1325kb
1350kb
1375kb
1400kb
1425kb
1450kb
1475kb
1500kb
1525kb
1550kb
1575kb
1600kb
1625kb
1650kb
1675kb

43


D
Caldisericum exile
AZM16c01
0kb
25kb
50kb
75kb
100kb
125kb
150kb
175kb
200kb
225kb
250kb
275kb
300kb
325kb
350kb
375kb
400kb
425kb
450kb
475kb
500kb
525kb
550kb
575kb
600kb
625kb
650kb
675kb
700kb
725kb
750kb
775kb
800kb
825kb
850kb
875kb
900kb
925kb
950kb
975kb
1000kb
1025kb
1050kb
1075kb
1100kb
1125kb
1150kb
1175kb
1200kb
1225kb
1250kb
1275kb
1300kb
1325kb
1350kb
1375kb
1400kb
1425kb
1450kb
1475kb
1500kb
1525kb
1550kb

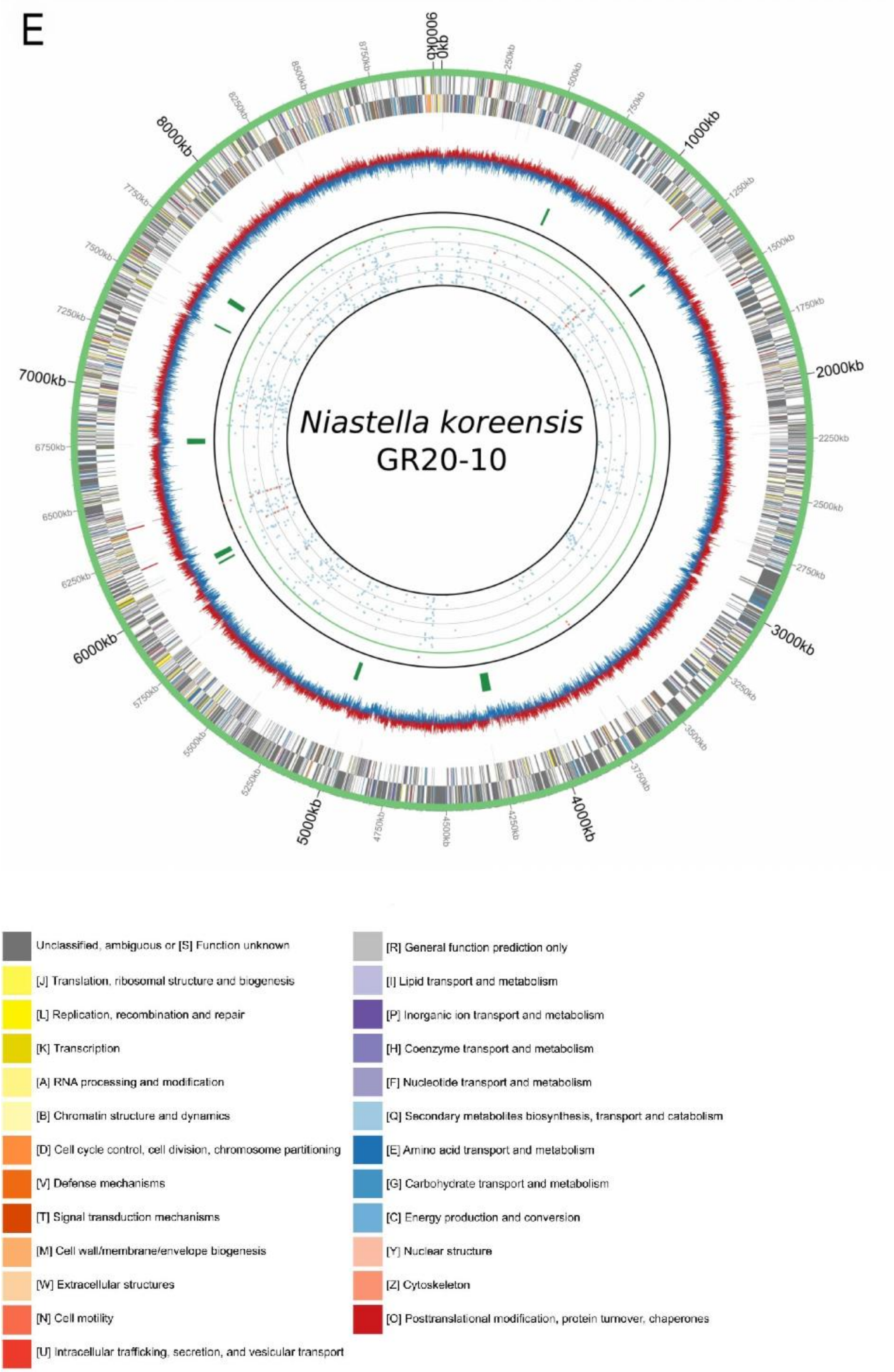

E
Niastella koreensis
GR20-10
0kb
9006kb
250kb
500kb
750kb
1000kb
1250kb
1500kb
1750kb
2000kb
2250kb
2500kb
2750kb
3000kb
3250kb
3500kb
3750kb
4000kb
4250kb
4500kb
4750kb
5000kb
5250kb
5500kb
5750kb
6000kb
6250kb
6500kb
6750kb
7000kb
7250kb
7500kb
7750kb
8000kb
8250kb
8500kb
8750kb
Unclassified, ambiguous or [S] Function unknown
[J] Translation, ribosomal structure and biogenesis
[L] Replication, recombination and repair
[K] Transcription
[A] RNA processing and modification
[B] Chromatin structure and dynamics
[D] Cell cycle control, cell division, chromosome partitioning
[V] Defense mechanisms
[T] Signal transduction mechanisms
[M] Cell wall/membrane/envelope biogenesis
[W] Extracellular structures
[N] Cell motility
[U] Intracellular trafficking, secretion, and vesicular transport
[R] General function prediction only
[I] Lipid transport and metabolism
[P] Inorganic ion transport and metabolism
[H] Coenzyme transport and metabolism
[F] Nucleotide transport and metabolism
[Q] Secondary metabolites biosynthesis, transport and catabolism
[E] Amino acid transport and metabolism
[G] Carbohydrate transport and metabolism
[C] Energy production and conversion
[Y] Nuclear structure
[Z] Cytoskeleton
[O] Posttranslational modification, protein turnover, chaperones

Table S1: Metadata of sequenced metagenomes used for comparison

Location	DNA source	NCBI SRA Accession	raw reads		after dereplication/ filtering/trimming		after 98%id clustering		assembly		gene prediction and alignment	
			total length	number of reads	total length	number of reads	total length	number of reads	total length	number of sequences	number of ORFs	with hits to NCBI-NR
<u>marine hydrothermal vent samples</u>												
Sisters Peak (southern Mid-Atlantic ridge)	MDA (REPLI-g Kit)	SRA047926	340 Mbp	875,069	113 Mbp	333,597	21 Mbp	60,929	8 Mbp	23,933	24,055	12,112 (50.4%)
Juan de Fuca (eastern Pacific ridge)	fosmid library (pCC1FOS)	SRR029255	133 Mbp	578,567	43 Mbp	197,775	36 Mbp	157,198	25 Mbp	101,011	100,478	70,157 (69.8%)
Mariana Trough (western Pacific back-arc basin)[1]	MDA (REPLI-g Kit)	SRR016610	56 Mbp	323,700	13 Mbp	67,820	6 Mbp	31,087	4 Mbp	17,170	16,230	7,132 (43.9%)
<u>marine non-vent samples</u>												
Brazos-Trinity Basin (Gulf of Mexico)	MDA (REPLI-g Kit)	SRR023396	106 Mbp	546,127	52 Mbp	254,571	29 Mbp	127,421	11 Mbp	45,818	44,516	19,948 (44.8%)

All the metagenomes were pyrosequenced using Titanium chemistry. The metagenomic data are derived from Sisters Peak (this study), Juan de Fuca (1), Mariana Trough (2) and Brazos Trinity Basin (3). The raw reads from all locations were clustered, processed and normalized as described in material and methods. [1]The metagenomic data is derived from the fluid 0.1 – 0.2 µm particle size fraction.

Table S2: Summary of the taxonomic and functional assignment results.

Metagenome	Genes with hits in NCBI-NR	Taxonomic assignments	SEED functional assignments	KEGG functional assignments	COG functional assignments
Sisters Peak	12,112	11,419 (94.3%)	4,842 (40.0%)	7,078 (58.4%)	8,844 (73.0%)
Juan de Fuca	70,157	65,899 (93.9%)	33,162 (47.3%)	45,187 (64.4%)	57,780 (82.3%)
Mariana Trough	7,132	6,597 (92.5%)	2,922 (41.0%)	3,729 (52.3%)	5,499 (77.1%)
Brazos-Trinity Basin	19,948	17,434 (87.4%)	8,885 (44.5%)	12,282 (61.6%)	15,028 (75.3%)

Taxonomic and functional assignments of genes in the analyzed metagenomes according to MEGAN. The relative values (percentage) refer to the total number of genes with at least one significant (E-value < 0.001) blastp hit in NCBI-NR.

Table S3: Taxonomic assignment of predicted rRNA genes.

	Count	**Assignment**
16S rRNA genes	6	Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales; Deep Sea Hydrothermal Vent Gp 2 (DHVE2)
	1	Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales; Deep Sea Hydrothermal Vent Gp 2 (DHVE2); Candidatus Aciduliprofundum
	1	Bacteria
	1	Bacteria; Bacteroidetes
	1	Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae
	1	Bacteria; Proteobacteria; Epsilonproteobacteria; Nautiliales; Nautiliaceae;Caminibacter
	2	Bacteria; Proteobacteria; Epsilonproteobacteria; Nautiliales; Nautiliaceae
	Count	**Assignment**
23S rRNA genes	17	Unclassified
	3	Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales; Family Incertae Sedis; Candidatus Aciduliprofundum
	1	Bacteria
	2	Bacteria; Proteobacteria; Epsilonproteobacteria
	3	Bacteria; Proteobacteria; Epsilonproteobacteria; Nautiliales; Nautiliaceae
	2	Bacteria; Proteobacteria; Epsilonproteobacteria; Nautiliales; Nautiliaceae; Nitratifractor
	1	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales; Helicobacteraceae; Sulfurimonas

Table S4: Potential genomic islands of *A*. sp. MAR08-339, respective matches in the SP metagenome and other best hits

SP metagenome	***A*. sp. MAR08-339**				**others**		
number of hits	**genomic islands range**	**best matches**			**next best blast hits in NCBI NR**		
		locus tag	**predicted functions**	**e-values**	**hit taxon**	**predicted functions**	**e-value**
	potential genomic island 1 predicted range 611,663..617,498						
4	predicted	AciM339_0710	hypothetical protein	6,00E-35 to 2,00E-07	*M. alvus*	hypothetical protein MMALV_13000	3,00E-12
38	predicted	AciM339_0711	hypothetical protein	3,00E-89 to 1,00E-08	*M. alvus*	Exonuclease SBCC	3,00E-36
no hits	predicted	AciM339_0712	hypothetical protein	-	no hits		
16	predicted	AciM339_0713	putative ATPase	7,00E-51 to 6,00E-09	*M. alvus*	hypothetical protein MMALV_13020	0.0
20	predicted	AciM339_0714	NurA domain-containing protein	3,00E-47 to 6,00E-16	*M. alvus*	hypothetical protein MMALV_13030	5,00E-76
	extended range 611,629..620,879						
1	extended	AciM339_0715	site-specific recombinase XerD	6,00E-08	*A. boonei*	integrase family protein	1,00E-160
1	extended	AciM339_0716	putative nucleic-acid-binding protein containing a Zn-ribbon	6,00E-08	*A. boonei*	protein of unknown function DUF35	8,00E-84
16	extended	AciM339_0717	acetyl-CoA acetyltransferase	7,00E-38 to 8,00E-8	*A. boonei*	acetyl-CoA acetyltransferase	0.0
19	extended	AciM339_0718	hydroxymethylglutaryl-CoA synthase, putative	6,00E-25 to 2,00E-6	no hits		

potential genomic island 2

extended range 683,941..724,667

no hits	extended	AciM339_0781	phosphomannomutase	-	no hits		
4	extended	AciM339_0782	adenosine deaminase	7,00E-22 to 1,00E-7	*A. boonei*	adenine deaminase	5,00E-156
16	extended	AciM339_0783	UDP-N-acetylglucosamine diphosphorylase/glucosamine-1-phosphate N-acetyltransferase	2,00E-56 to 7,00E-10	*A. boonei*	Nucleotidyl transferase	0.0
no hits	extended	AciM339_0784	putative permease	-	*A. boonei*	hypothetical protein	1,00E-87
no hits	extended	AciM339_0785	amino acid transporter	-	*A. boonei*	amino acid transporter	0.0
no hits	extended	AciM339_0786	putative signal-transduction protein containing cAMP-binding and CBS domains	-	*A. boonei*	signal transduction protein with CBS domains	3,00E-35
1	extended	AciM339_0787	nucleoside-diphosphate-sugar epimerase	4,00E-09	*A. boonei*	NAD-dependent epimerase/dehydratase	1,00E-161
2	extended	AciM339_0788	ABC-type multidrug transport system, ATPase component	1,00E-13 and 5,00E-11	*T. zilligii*	daunorubicin ABC transporter ATP-binding protein	6,00E-81
8	extended	AciM339_0789	ABC-type multidrug transport system, permease component	2,00E-40 to 1,00E-6	*T. zilligii*	ABC transporter	3,00E-56
18	extended	AciM339_0790	methyltransferase, FkbM family	1,00E-39 to 2,00E-6	*A. boonei*	methyltransferase FkbM family	1,00E-123
no hits	extended	AciM339_0791	glycosyl transferase	-	*A. boonei*	glycosyl transferase family 2	3,00E-112
12	extended	AciM339_0792	glycosyltransferase	2,00E-53 to 7,00E-6	*A. boonei*	glycosyl transferase group 1	1,00E-105
13	extended	AciM339_0793	glycosyl/glycerophosphate transferase, teichoic acid biosynthesis	3,00E-70 to 3,00E-9	*A. boonei*	CDP-glycerol:Poly(glycero-phosphate) glycerophospho-transferase superfamily	1,00E-68
2	extended	AciM339_0794	putative sugar nucleotidyltransferase	9,00E-14 and 3,00E-24	*Maricaulis* sp. JL2009	nucleotidyltransferase	7,00E-45

20	extended	AciM339_0795	asparagine synthase, glutamine-hydrolyzing	2,00E-47 to 2,00E-7	*P. abyssi*	asparagine synthetase, glutamine-hydrolyzing	5,00E-146
6	extended	AciM339_0796	putative ATPase (AAA+ superfamily)	2,00E-61 to 2,00E-13	uncultured archaeon	conserved hypothetical protein	9,00E-146
19	extended	AciM339_0797	putative ATPase (AAA+ superfamily)	2,00E-63 to 1,00E-9	*M. acidiphilum*	conserved hypothetical protein	7,00E-123
10	extended	AciM339_0798	hypothetical protein	1,00E-39 to 4,00E-6	*A. boonei*	hypothetical protein	6,00E-12
9	extended	AciM339_0799	Transmembrane exosortase (Exosortase_EpsH)	8,00E-63 to 4,00E-8	*A. boonei*	Exosortase EpsH-related protein	6,00E-09
	predicted range 706,082..720,183						
10	predicted	AciM339_0800	membrane protein (export of O-antigen and teichoic acid)	1,00E-62 to 9,00E-8	*Methanobrevibacter* sp. AbM4	polysaccharide biosynthesis protein	6,00E-17
55	predicted	AciM339_0801	hypothetical protein	7,00E-70 to 1,00E-10	no hits		
8	predicted	AciM339_0802	hypothetical protein	2,00E-35 to 2,00E-10	*A. veneficus*	nucleotide pyrophosphatase	5,00E-09
29	predicted	AciM339_0803	glycosyltransferase	4,00E-58 to 4,00E-8	*N. lucentensis*	glucosyl transferase	8,00E-22
23	predicted	AciM339_0804	glycosyltransferase	2,00E-81 to 2,00E-6	*N. multiformis*	group 1 glycosyl transferase	4,00E-32
8	predicted	AciM339_0805	glycosyltransferase	1,00E-38 to 2,00E-12	*Thermococcus* sp. CL1	glycosyl transferase family protein 8	7,00E-76
6	predicted	AciM339_0806	UDP-N-acetylglucosamine 2-epimerase	1,00E-95 to 5,00E-10	*A. boonei*	UDP-N-acetylglucosamine 2-epimerase	1,00E-172
no hits	predicted	AciM339_0807	hypothetical protein	-	*M. fervidus*	hypothetical protein Mfer_0403	5,00E-28
no hits	predicted	AciM339_0808	thymidylate kinase	-	*M. yellowstonensis*	thymidylate kinase	6,00E-11
no hits	predicted	AciM339_0809	hypothetical protein	-	*M. liminatans*	hypothetical protein	3,00E-38

6	predicted	AciM339_0810	putative phosphoesterase, PHP family	4,00E-72 to 7,00E-6	*M. hungatei*	PHP-like protein	7,00E-28
10	predicted	AciM339_0811	putative ATP-grasp enzyme	2,00E-20 to 1,00E-6	*D. acetoxidans*	hypothetical protein Desac_1648	8,00E-83
no hits	extended	AciM339_0812	putative xylanase/chitin deacetylase	-	*K. racemifer*	polysaccharide deacetylase	9,00E-34
2	extended	AciM339_0813	UDP-N-acetylglucosamine-1-phosphate transferase	1,00E-22 and 4,00E-10	*A. boonei*	glycosyl transferase, family 4	8,00E-137
14	extended	AciM339_0814	nucleotide sugar dehydrogenase	2,00E-37 to 1,00E-6	*Thermococcus* sp. ES1	NDP-sugar dehydrogenase	0.0
3	extended	AciM339_0815	putative dehydrogenase	3,00E-26 to 1,00E-8	*M. luminyensis*	hypothetical protein	7,00E-82
3	extended	AciM339_0816	serine acetyltransferase	1,00E-14	no hits		
	potential genomic island 3						
	extended range 1,083,982..1,105,557						
25	extended	AciM339_1215	hypothetical protein	8,00E-54 to 1,00E-7	*A. boonei*	hypothetical protein	4,00E-10
29	extended	AciM339_1216	DNA/RNA helicase, superfamily II	3,00E-50 to 2,00E-6	*A. boonei*	hypothetical protein	0.0
69	extended	AciM339_1217	adenine specific DNA methylase Mod	6,00E-75 to 1,00E-6	*A. boonei*	DNA methylase domain protein	0.0
79	extended	AciM339_1218	DNA/RNA helicase, superfamily I	7,00E-61 to 4,00E-6	*A. boonei*	UvrD/REP helicase domain protein	0.0
7	extended	AciM339_1219	hypothetical protein	4,00E-54 to 5,00E-14	*M. hollandica*	molecular chaperone of HSP90 family	2,00E-07
2	extended	AciM339_1220	hypothetical protein	2,00E-26 and 7,00E-6	no hits		
	predicted range 1,096,394..1,101,344						
38	predicted	AciM339_1221	Lhr-like helicase	6,00E-57 to 4,00E-6	*C. saccharofermentans*	hypothetical protein	9,00E-174

32	predicted	AciM339_1222	ATPase	1,00E-49 to 1,00E-6	*C. saccharofermentans*	hypothetical protein	1,00E-83
44	predicted	AciM339_1223	unknown function	3,00E-52 to 2,00E-6	*C. saccharofermentans*	hypothetical protein	2,00E-159
35	predicted	AciM339_1224	hypothetical protein	2,00E-75 to 2,00E-7	*C. saccharofermentans*	hypothetical protein	9,00E-52
	extended range 1,083,982..1,105,557						
11	extended	AciM339_1225	hypothetical protein	7,00E-61 to 8,00E-8	no hits		
5	extended	AciM339_1226	hypothetical protein	7,00E-61 to 6,00E-16	no hits		
10	extended	AciM339_1227	hypothetical protein	1,00E-34 to 1,00E-7	no hits		
27	extended	AciM339_1228	hypothetical protein	2,00E-72 to 2,00E-7	no hits		
4	extended	AciM339_1229	hypothetical protein	9,00E-40 to 6,00E-11	no hits		
2	extended	AciM339_1230	hypothetical protein	1,00E-31 and 2,00E-8	no hits		
5	extended	AciM339_1231	hypothetical protein	4,00E-10 to 2,00E-6	no hits		
	potential genomic island 4						
	extended range 1,216,880..1,229,475						
13	extended	AciM339_1347	glutamate dehydrogenase/leucine dehydrogenase	1,00E-45 to 4,00E-7	no hits		
5	extended	AciM339_1349	glycosyl transferase	8,00E-51 to 7,00E-9	*A. boonei*	glycosyl transferase family 2	0.0
	predicted 1,219,892..1,229,434						
11	predicted	AciM339_1350	TIGR02688 family protein	5,00E-39 to 6,00E-8	*T. petrophila*	hypothetical protein	0.0

39	predicted	AciM339_1351	hypothetical protein	2,00E-28 to 1,00E-6	*T. petrophila*	hypothetical protein	0.0
56	predicted	AciM339_1352	hypothetical protein	1,00E-41 to 2,00E-6	*T. petrophila*	hypothetical protein	0.0
57	predicted	AciM339_1353	PAPS reductase/FAD synthetase family	2,00E-56 to 2,00E-6	*M. thermautotrophicus*	phosphoadenosine phosphosulfate reductase	0.0
	potential genomic island 5 predicted range 1,377,646..1,382,742						
14	predicted	AciM339_1513	DEAD/DEAH box helicase	1,00E-43 to 5,00E-6	*T. sibiricus*	hypothetical protein	0.0
3	predicted	AciM339_1514	CRISPR-associated protein	1,00E-62 to 7,00E-21	*T. sibiricus*	hypothetical protein	8,00E-80
4	predicted	AciM339_1515	CRISPR-associated protein	9,00E-29 to 6,00E-7	*T. sibiricus*	hypothetical protein	2,00E-141
12	predicted	AciM339_1516	hypothetical protein	3,00E-38 to 5,00E-6	*T. sibiricus*	hypothetical protein	1,00E-149
	extended range 1,377,325..1,384,042						
3	extended	AciM339_1517	CRISPR-associated endoribonuclease	2,00E-35 to 6,00E-30	*A. boonei*	CRISPR-associated protein	2,00E-35
14	extended	AciM339_1517	phosphoesterase	3,00E-61 to 2,00E-7	*T. thermarum*	CRISPR-associated protein	2,00E-31

Abbreviations used in the table denote: *Aciduliprofundum boonei, Archaeoglobus veneficus, Caldatribacterium saccharofermentans, Desulfobacca acetoxidans, Ktedonobacter racemifer, Metallosphaera yellowstonensis, Methanofollis liminatans, Methanomassiliicoccus luminyensis, Methanomethylophilus alvus, Methanomethylovorans hollandica, Methanospirillum hungatei, Methanothermobacter thermautotrophicus, Methanothermus fervidus, Micrarchaeum acidiphilum, Nitrosospira multiformis, Nocardiopsis lucentensis, Pyrococcus abyssi, Thermococcus sibiricus, Thermococcus zigli, Thermotoga thermarum* and *Thermotoga petrophila*.

References

1. **Xie W, Wang F, Guo L, Chen Z, Sievert SM, Meng J, Huang G, Li Y, Yan Q, Wu S, Wang X, Chen S, He G, Xiao X, Xu A.** 2010. Comparative metagenomics of microbial communities inhabiting deep-sea hydrothermal vent chimneys with contrasting chemistries. ISME J. **5:**414-426.

2. **Nakai R, Abe T, Takeyama H, Naganuma T.** 2011. Metagenomic analysis of 0.2-µm-passable microorganisms in deep-sea hydrothermal fluid. Mar. Biotechnol. **13:**900-908.

3. **Biddle JF, White JR, Teske AP, House CH.** 2011. Metagenomics of the subsurface Brazos-Trinity Basin (IODP site 1320): comparison with other sediment and pyrosequenced metagenomes. ISME J. **5:**1038-1047.