

```
load(file = 'data/abudat.RData')
lapply(abudat, head)

## $DOMAIN
## # A tibble: 6 x 5
## # Groups:   site, wwtp, cont [6]
##   domain site wwtp cont abund
##   <chr> <chr> <chr> <chr> <dbl>
## 1 Archaea CA1 LACO int 148
## 2 Archaea CA10 ORCO int 50
## 3 Archaea CA11 SDCI int 70
## 4 Archaea CA12 ORCO clo 104
## 5 Archaea CA13 LACI far 96
## 6 Archaea CA14 LACI int 139
```

```
##
## $PHYLUM
## # A tibble: 6 x 6
## # Groups:   site, wwtp, cont [6]
##   phylum site wwtp cont abund domain
##   <chr> <chr> <chr> <chr> <dbl> <chr>
## 1 AC1 CA1 LACO int 2 Bacteria
## 2 AC1 CA10 ORCO int 0 Bacteria
## 3 AC1 CA11 SDCI int 0 Bacteria
## 4 AC1 CA12 ORCO clo 0 Bacteria
## 5 AC1 CA13 LACI far 0 Bacteria
## 6 AC1 CA14 LACI int 0 Bacteria
##
## $CLASS
## # A tibble: 6 x 7
## # Groups:   site, wwtp, cont [6]
##   class site wwtp cont abund domain phylum
##   <chr> <chr> <chr> <chr> <dbl> <chr> <chr>
## 1 028H05-P-BN-P5 CA1 LACO int 0 Bacteria Planctomycetes
## 2 028H05-P-BN-P5 CA10 ORCO int 0 Bacteria Planctomycetes
## 3 028H05-P-BN-P5 CA11 SDCI int 0 Bacteria Planctomycetes
## 4 028H05-P-BN-P5 CA12 ORCO clo 0 Bacteria Planctomycetes
## 5 028H05-P-BN-P5 CA13 LACI far 0 Bacteria Planctomycetes
## 6 028H05-P-BN-P5 CA14 LACI int 0 Bacteria Planctomycetes
##
## $ORDER
## # A tibble: 6 x 8
## # Groups:   site, wwtp, cont [6]
##   order site wwtp cont abund domain phylum
##   <chr> <chr> <chr> <chr> <dbl> <chr> <chr>
## 1 028H05-P-BN-P5_or CA1 LACO int 0 Bacteria Planctomycetes
## 2 028H05-P-BN-P5_or CA10 ORCO int 0 Bacteria Planctomycetes
## 3 028H05-P-BN-P5_or CA11 SDCI int 0 Bacteria Planctomycetes
## 4 028H05-P-BN-P5_or CA12 ORCO clo 0 Bacteria Planctomycetes
## 5 028H05-P-BN-P5_or CA13 LACI far 0 Bacteria Planctomycetes
## 6 028H05-P-BN-P5_or CA14 LACI int 0 Bacteria Planctomycetes
## # ... with 1 more variables: class <chr>
##
## $FAMILY
## # A tibble: 6 x 9
## # Groups:   site, wwtp, cont [6]
##   family site wwtp cont abund domain phylum class
##   <chr> <chr> <chr> <chr> <dbl> <chr> <chr> <chr>
## 1 01D2Z36 CA1 LACO int 0 Bacteria Verrucomicrobia Spartobacteria
```

```
## 2 01D2Z36 CA10 ORCO int 0 Bacteria Verrucomicrobia Spartobacteria
## 3 01D2Z36 CA11 SDCI int 0 Bacteria Verrucomicrobia Spartobacteria
## 4 01D2Z36 CA12 ORCO clo 0 Bacteria Verrucomicrobia Spartobacteria
## 5 01D2Z36 CA13 LACI far 0 Bacteria Verrucomicrobia Spartobacteria
## 6 01D2Z36 CA14 LACI int 1 Bacteria Verrucomicrobia Spartobacteria
## # ... with 1 more variables: order <chr>
##
## $GENUS
## # A tibble: 6 x 10
## # Groups:   site, wwtp, cont [6]
##      genus site wwtp cont abund domain phylum
##      <chr> <chr> <chr> <chr> <dbl> <chr> <chr>
## 1 01D2Z36_ge CA1 LACO int 0 Bacteria Verrucomicrobia
## 2 01D2Z36_ge CA10 ORCO int 0 Bacteria Verrucomicrobia
## 3 01D2Z36_ge CA11 SDCI int 0 Bacteria Verrucomicrobia
## 4 01D2Z36_ge CA12 ORCO clo 0 Bacteria Verrucomicrobia
## 5 01D2Z36_ge CA13 LACI far 0 Bacteria Verrucomicrobia
## 6 01D2Z36_ge CA14 LACI int 1 Bacteria Verrucomicrobia
## # ... with 3 more variables: class <chr>, order <chr>, family <chr>
##
## $SPECIES
## # A tibble: 6 x 11
## # Groups:   site, wwtp, cont [6]
##      species site wwtp cont abund domain phylum class order
##      <chr> <chr> <chr> <chr> <dbl> <chr> <chr> <chr> <chr>
## 1 0tu00001 CA1 LACO int 4993 Bacteria Firmicutes Bacilli Bacillales
## 2 0tu00001 CA10 ORCO int 7001 Bacteria Firmicutes Bacilli Bacillales
## 3 0tu00001 CA11 SDCI int 2573 Bacteria Firmicutes Bacilli Bacillales
## 4 0tu00001 CA12 ORCO clo 703 Bacteria Firmicutes Bacilli Bacillales
## 5 0tu00001 CA13 LACI far 842 Bacteria Firmicutes Bacilli Bacillales
## 6 0tu00001 CA14 LACI int 1578 Bacteria Firmicutes Bacilli Bacillales
## # ... with 2 more variables: family <chr>, genus <chr>
```

## 1 Community structure

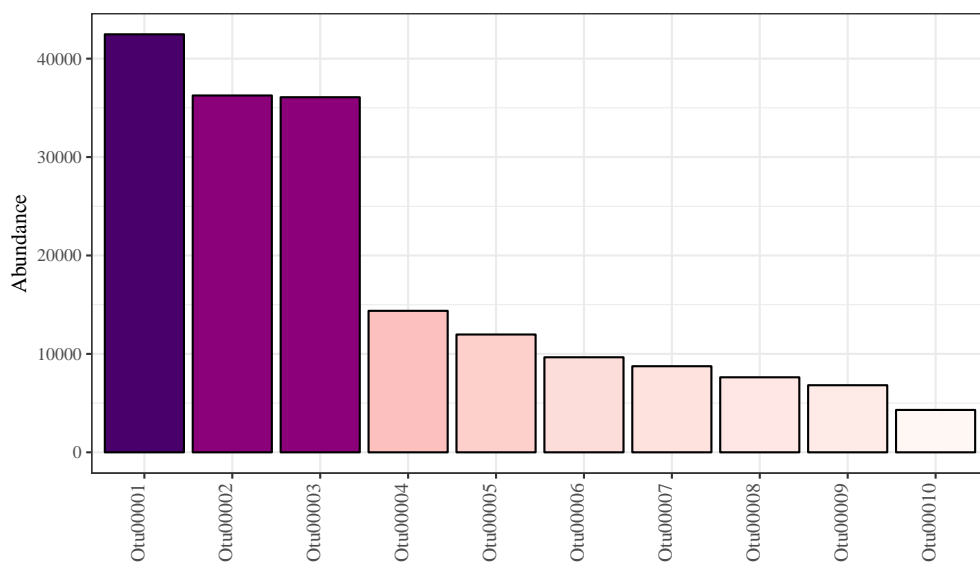


Figure 1: Top fifty most abundant species across all study sites.

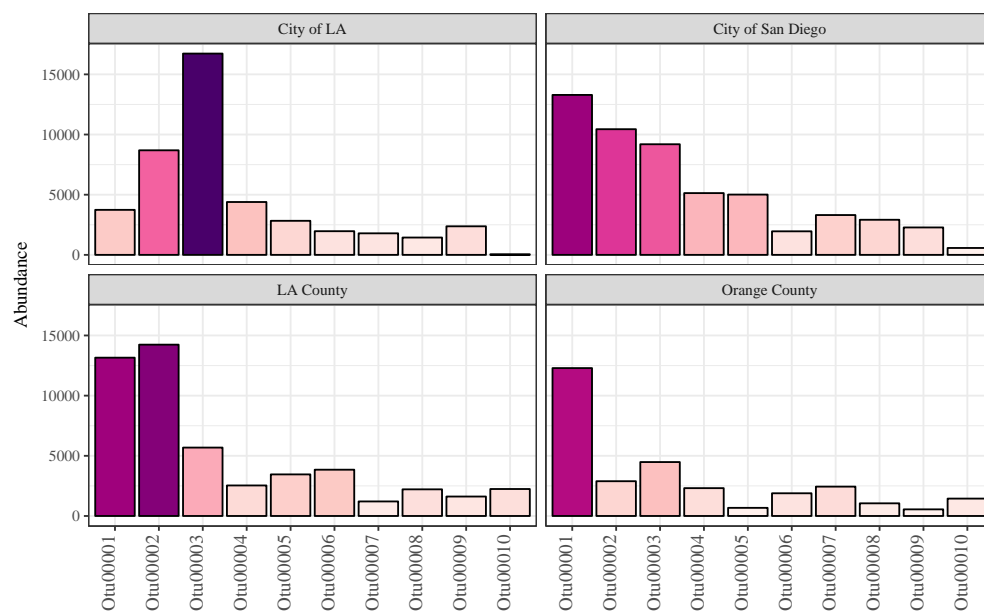


Figure 2: Species abundances by municipal locations using the top fifty most abundant across all study sites in fig. 1.

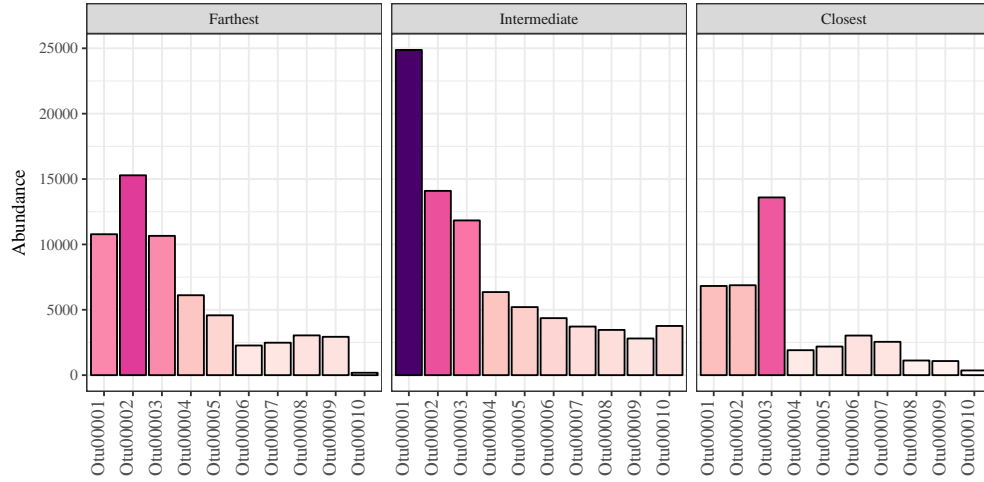


Figure 3: Species abundances by distance from pipe for the top fifty most abundant across all study sites in fig. 1. Treatments are based on relative distances from an outflow pipe (farthest, intermediate, closest) for each municipality.

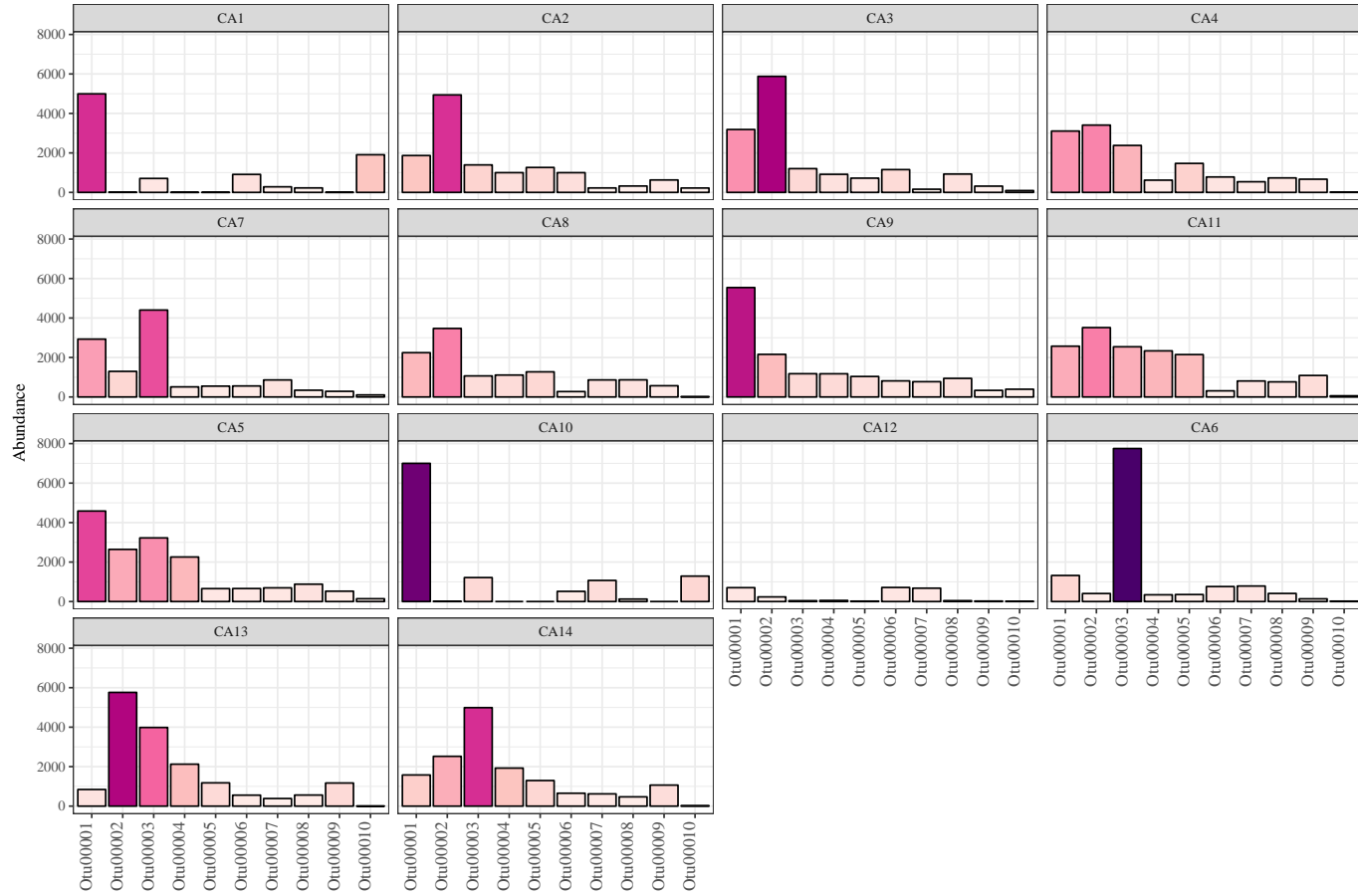


Figure 4: Species abundances by sites using the top fifty most abundant across all study sites in fig. 1.

## 1.1 Multivariate analyses with archaea

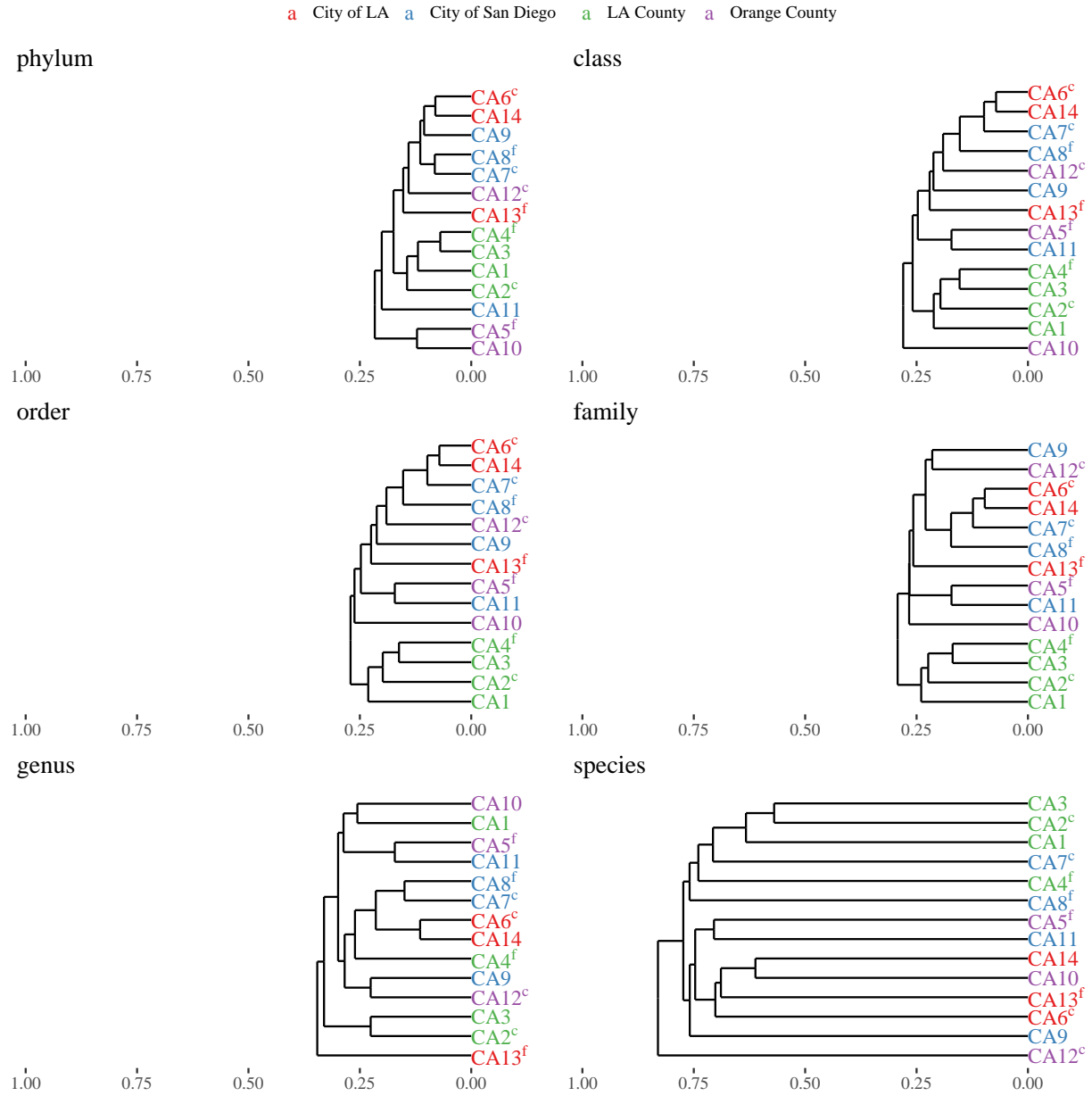


Figure 5: Site clusters of archaea by taxonomic resolution. Colors indicate municipality and superscripts indicate distance categories from an outflow pipe at each site (‘f’ is farthest, ‘c’ is closest, none is intermediate). Clustering was based on a Bray-Curtis dissimilarity matrix of abundance data and sorting using the unweighted pair group method.



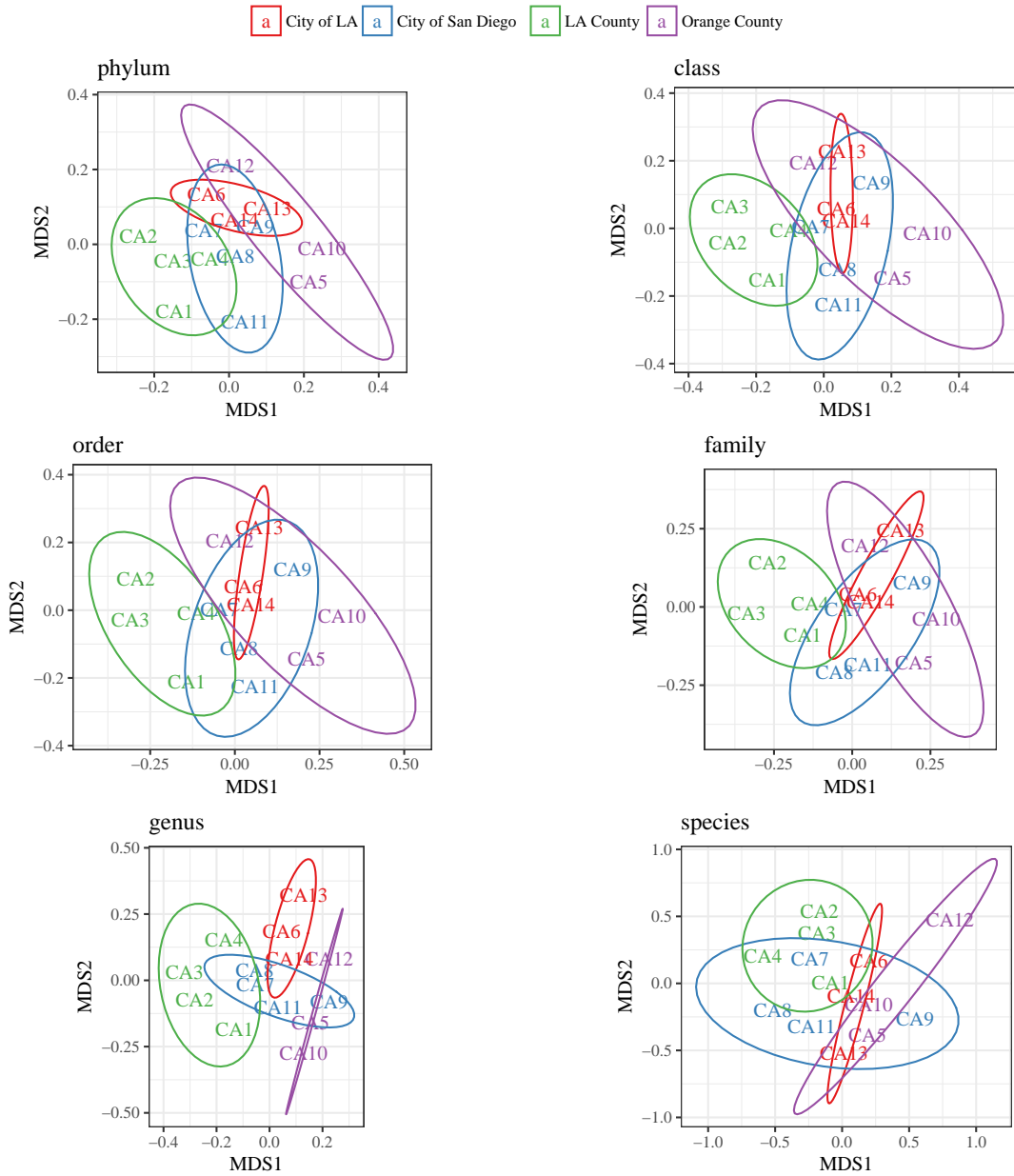


Figure 6: Site ordinations of archaea by taxonomic resolution. Colors indicate municipality with ellipses showing 95% bivariate confidence intervals. Ordinations were created using Nonmetric Multidimensional scaling with several random starts to find a stable solution of the configuration. A Bray-Curtis dissimilarity matrix was used as a measure of differences between sites. The ordination is the same as in fig. 7 but with different group categories in the plot.

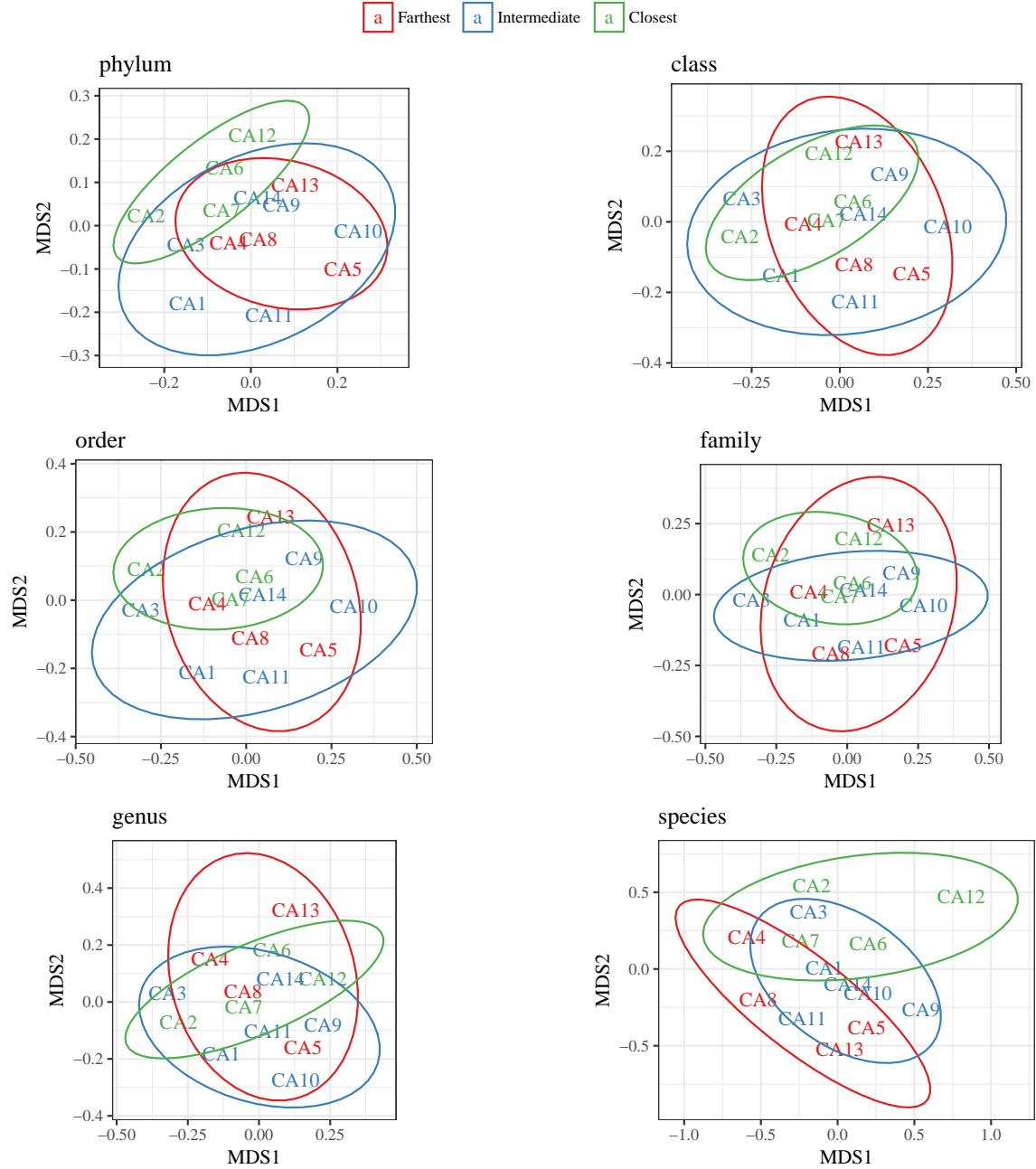


Figure 7: Site ordinations of archaea by taxonomic resolution. Colors indicate approximate distance from pipe with ellipses showing 95% bivariate confidence intervals. Ordinations were created using Nonmetric Multidimensional scaling with several random starts to find a stable solution of the configuration. A Bray-Curtis dissimilarity matrix was used as a measure of differences between sites. The ordination is the same as in fig. 6 but with different group categories in the plot.

Table 1: Alpha diversity of archaea by municipality, distance from outflow, and sites for each taxonomic level. Original data were taxonomic abundance aggregated by each grouping. Alpha was based on methods in Fisher et al. (1943) that measure diversity as a function of richness and abundance at a site.

Location grouping	Phylum	Class	Order	Family	Genus	Species
<b>Municipality</b>						
City of LA	1.83	2.77	3.02	3.80	4.62	70.34
City of San Diego	1.78	2.70	2.70	3.69	3.95	135.86
LA County	2.07	3.16	3.63	4.10	4.85	80.58
Orange County	1.65	2.47	2.47	3.06	3.37	88.71
<b>Pipe location</b>						
Farthest	2.06	3.02	3.28	4.34	5.19	109.98
Intermediate	2.11	3.00	3.47	3.95	4.70	88.42
Closest	1.63	2.44	2.65	3.31	4.00	127.81
<b>Site</b>						
CA1	2.11	2.75	2.75	2.75	2.75	23.76
CA2	2.00	2.79	3.07	3.35	3.64	46.23
CA3	2.19	3.58	3.96	3.96	3.96	27.14
CA4	1.67	2.33	2.33	2.68	3.44	44.31
CA7	1.49	2.35	2.35	2.66	2.98	86.56
CA8	1.22	1.75	1.75	2.33	2.33	53.49
CA9	1.50	2.59	2.59	3.00	3.44	29.57
CA11	3.19	3.67	3.67	4.17	4.17	47.43
CA5	2.10	3.01	3.01	3.51	3.51	45.68
CA10	1.78	2.69	2.69	2.69	2.69	23.95
CA12	1.69	2.73	2.73	3.51	3.92	59.12
CA6	1.42	2.23	2.23	2.23	3.14	34.14
CA13	2.43	4.06	4.51	6.00	7.10	38.20
CA14	1.28	2.15	2.15	2.47	3.15	30.15

Table 2: Beta diversity of archaea between municipalities, distance categories from outflow, and sites for each taxonomic level. Original data were taxonomic abundance aggregated by each grouping. Beta was estimated as total species richness across all categories in each grouping, divided by mean species richness within each category, minus one.

Location grouping	Phylum	Class	Order	Family	Genus	Species
Distance from outflow	0.18	0.27	0.35	0.45	0.48	1.39
Site	0.72	0.86	1.11	1.45	1.67	7.64
Municipality	0.30	0.36	0.49	0.60	0.69	2.05

## 1.2 Multivariate analyses with bacteria

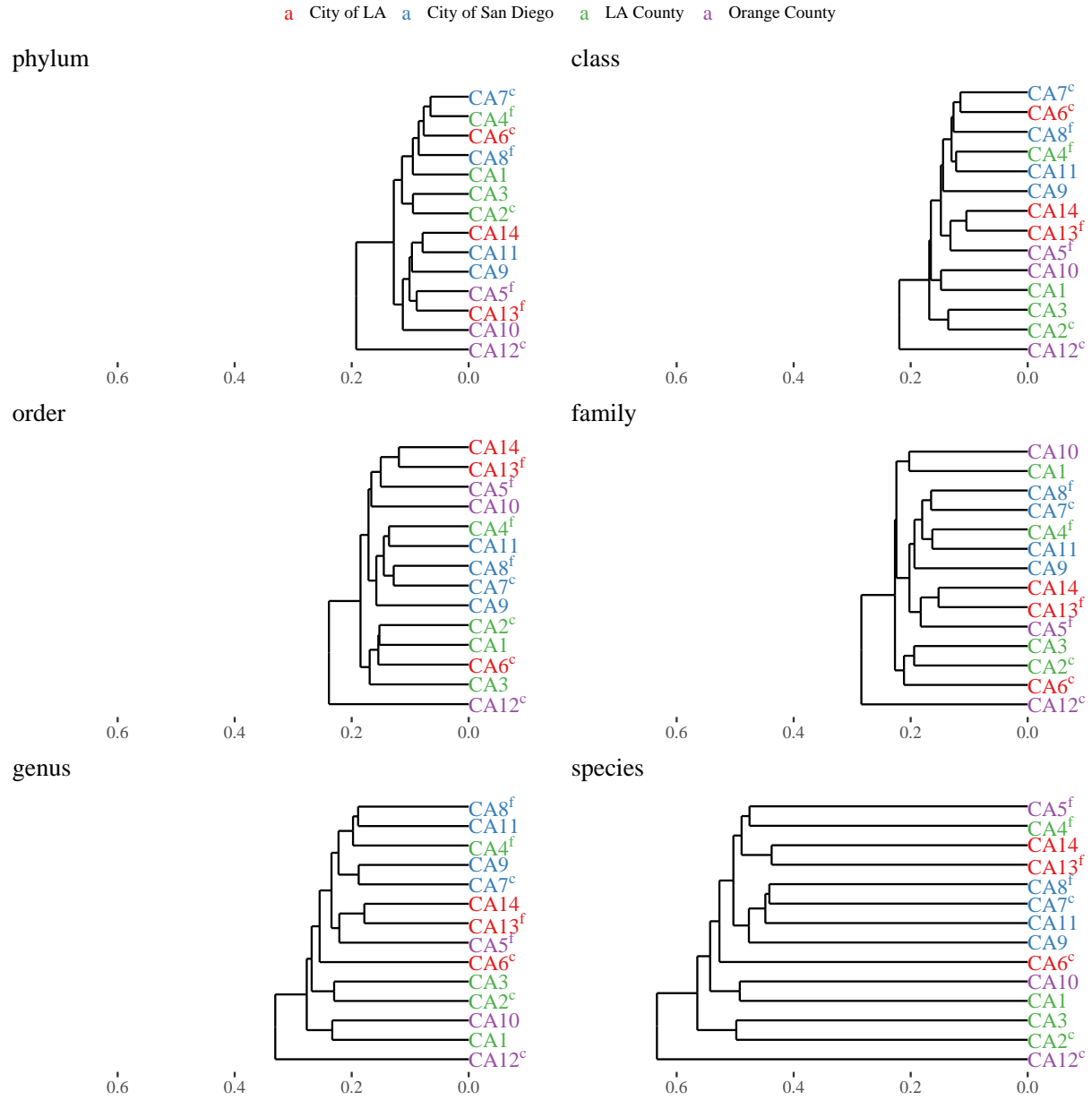


Figure 8: Site clusters of bacteria by taxonomic resolution. Colors indicate municipality and superscripts indicate distance categories from an outflow pipe at each site ('f' is farthest, 'c' is closest, none is intermediate). Clustering was based on a Bray-Curtis dissimilarity matrix of abundance data and sorting using the unweighted pair group method.

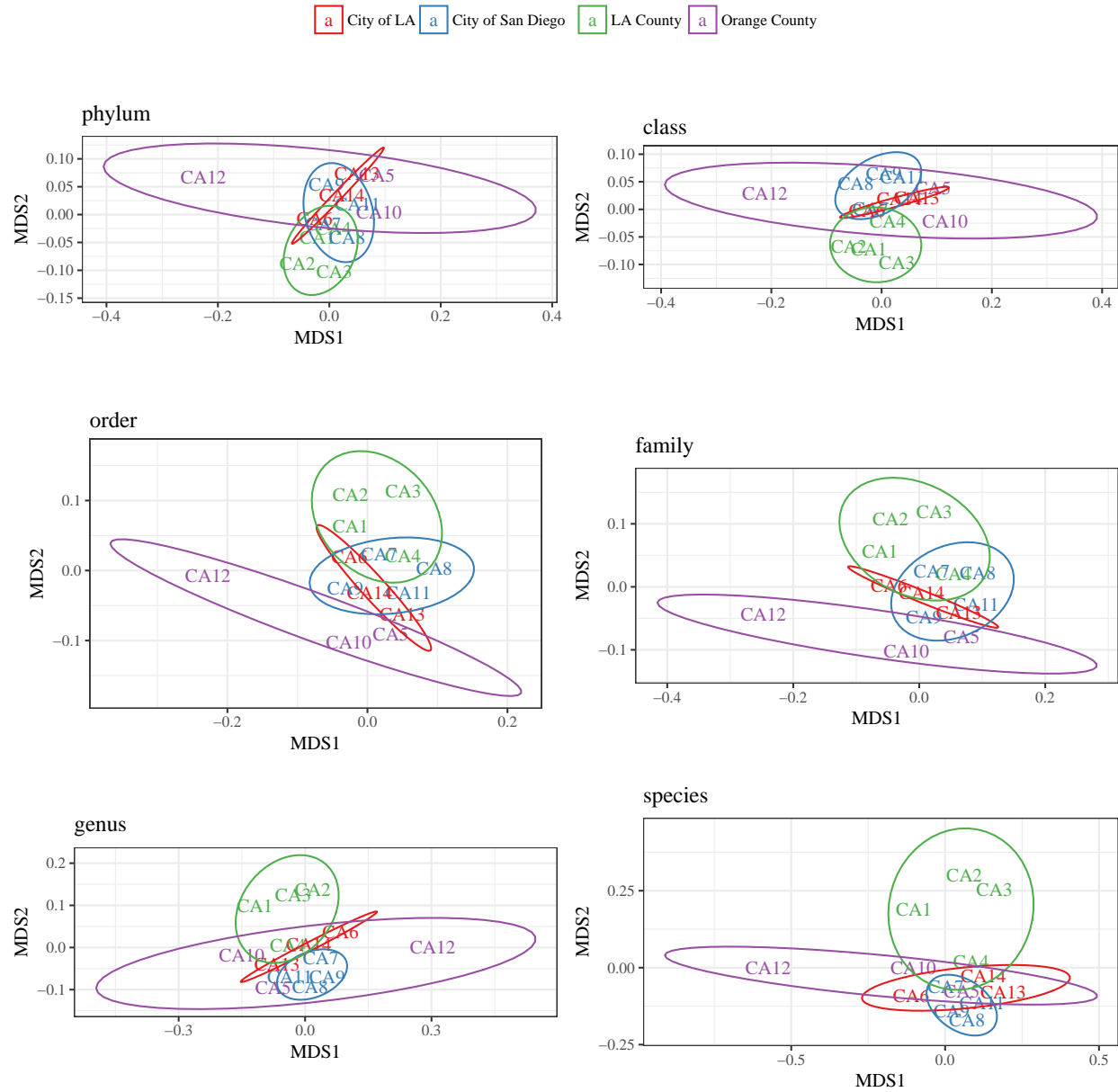


Figure 9: Site ordinations of bacteria by taxonomic resolution. Colors indicate municipality with ellipses showing 95% bivariate confidence intervals. Ordinations were created using Nonmetric Multidimensional scaling with several random starts to find a stable solution of the configuration. A Bray-Curtis dissimilarity matrix was used as a measure of differences between sites. The ordination is the same as in fig. 10 but with different group categories in the plot.

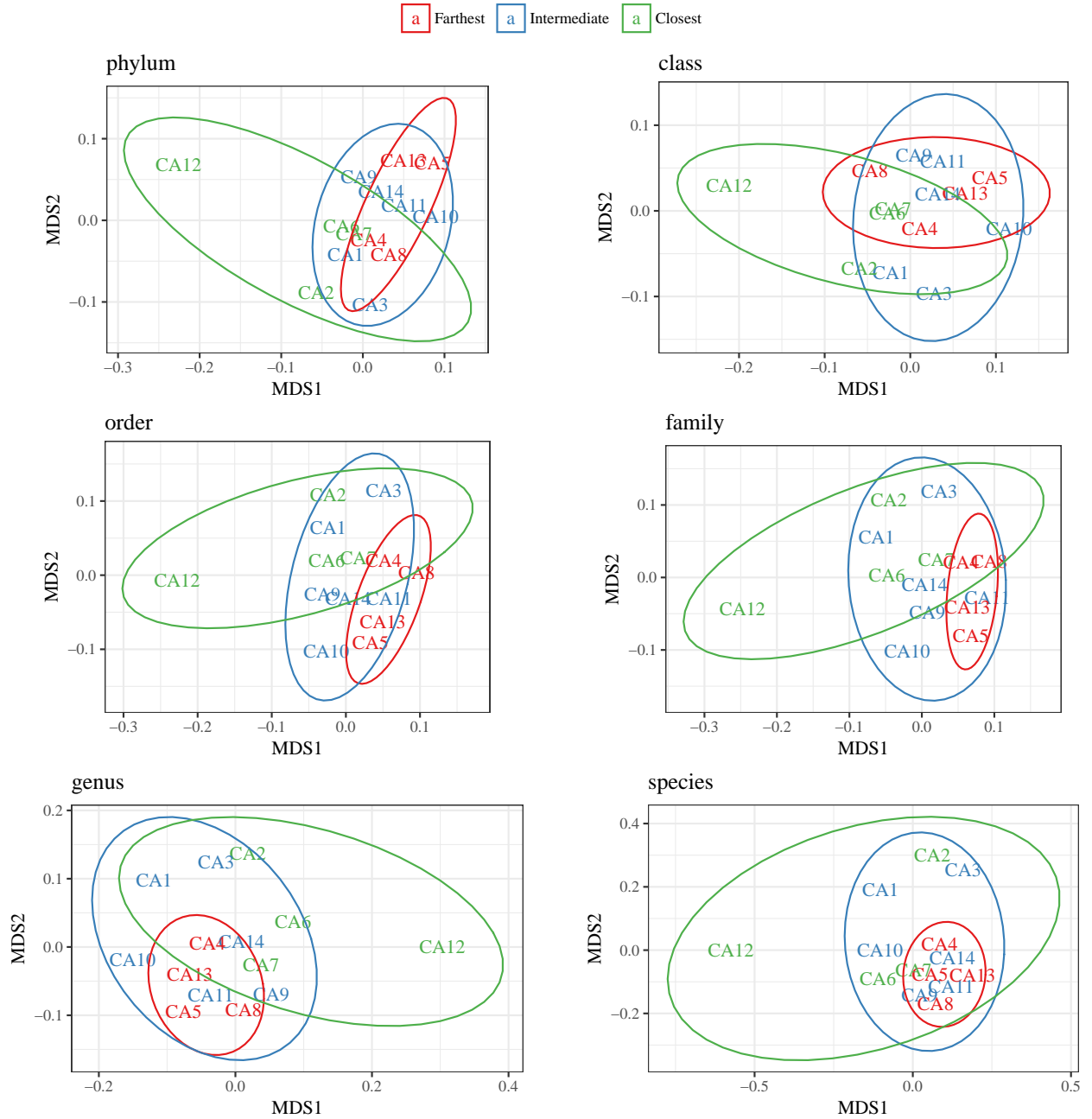


Figure 10: Site ordinations of bacteria by taxonomic resolution. Colors indicate approximate distance from pipe with ellipses showing 95% bivariate confidence intervals. Ordinations were created using Nonmetric Multidimensional scaling with several random starts to find a stable solution of the configuration. A Bray-Curtis dissimilarity matrix was used as a measure of differences between sites. The ordination is the same as in fig. 9 but with different group categories in the plot.

Table 3: Alpha diversity of bacteria by municipality, distance from outflow, and sites for each taxonomic level. Original data were taxonomic abundance aggregated by each grouping. Alpha was based on methods in Fisher et al. (1943) that measure diversity as a function of richness and abundance at a site.

Location grouping	Phylum	Class	Order	Family	Genus	Species
<b>Municipality</b>						
City of LA	5.28	16.57	30.81	54.53	93.88	1161.57
City of San Diego	5.45	16.99	31.80	57.76	98.51	1334.51
LA County	5.68	17.89	33.21	58.85	104.05	1626.50
Orange County	4.94	16.30	28.77	51.04	89.02	1039.74
<b>Pipe location</b>						
Farthest	5.11	16.34	30.26	52.45	83.41	1261.77
Intermediate	5.53	17.49	32.48	59.57	108.56	1680.25
Closest	5.45	17.90	34.49	62.40	117.12	1756.89
<b>Site</b>						
CA1	5.26	17.23	29.40	50.32	78.76	785.92
CA2	6.06	18.51	34.29	57.25	95.48	841.61
CA3	5.66	15.83	29.23	46.20	74.05	637.01
CA4	5.39	15.98	27.35	43.05	63.42	692.85
CA7	5.52	16.61	30.61	50.89	81.67	731.89
CA8	5.79	17.23	29.92	48.82	73.05	791.90
CA9	5.26	15.20	27.85	48.60	80.37	571.54
CA11	4.99	14.90	26.00	42.68	64.79	544.88
CA5	4.09	13.07	24.00	39.41	57.16	448.07
CA10	4.21	13.07	24.00	39.58	59.67	501.72
CA12	5.13	15.98	27.18	46.19	78.12	639.99
CA6	5.79	17.08	31.48	53.75	88.60	857.85
CA13	4.47	14.29	25.83	41.41	62.04	464.25
CA14	4.86	15.21	26.85	44.90	72.63	570.57

Table 4: Beta diversity of bacteria between municipalities, distance categories from outflow, and sites for each taxonomic level. Original data were taxonomic abundance aggregated by each grouping. Beta was estimated as total species richness across all categories in each grouping, divided by mean species richness within each category, minus one.

Location grouping	Phylum	Class	Order	Family	Genus	Species
Distance from outflow	0.07	0.13	0.18	0.23	0.37	1.02
Site	0.31	0.48	0.63	0.85	1.31	4.72
Municipality	0.11	0.18	0.26	0.33	0.51	1.50