Micromonosporaceae Biosynthetic Gene Cluster Diversity Highlights the Need for Broad Spectrum Investigations

Imraan Alas[1], Doug R. Braun[1], Spencer S. Ericksen[2], Rauf Salamzade[3,4], Lindsay Kalan[3,4], Scott R. Rajski[1] and Tim S. Bugni[1,2,5]*

[1]Pharmaceutical Sciences Division, University of Wisconsin–Madison, Madison, Wisconsin, United States.

[2]Small Molecule Screening Facility, UW Carbone Cancer Center, Madison, Wisconsin, United States.

[3]Department of Medical Microbiology & Immunology, University of Wisconsin–Madison, Madison, Wisconsin, United States.

[4]Department of Biochemistry & Biomedical Sciences, McMaster University, Health Sciences Centre, Hamilton, Ontario, Canada.

[5]Lachman Institute for Pharmaceutical Development, University of Wisconsin–Madison, Madison, Wisconsin, United States.

* To whom correspondence should be addressed. Tel.: 1-608-263-2519. E-mail: tim.bugni@wisc.edu.

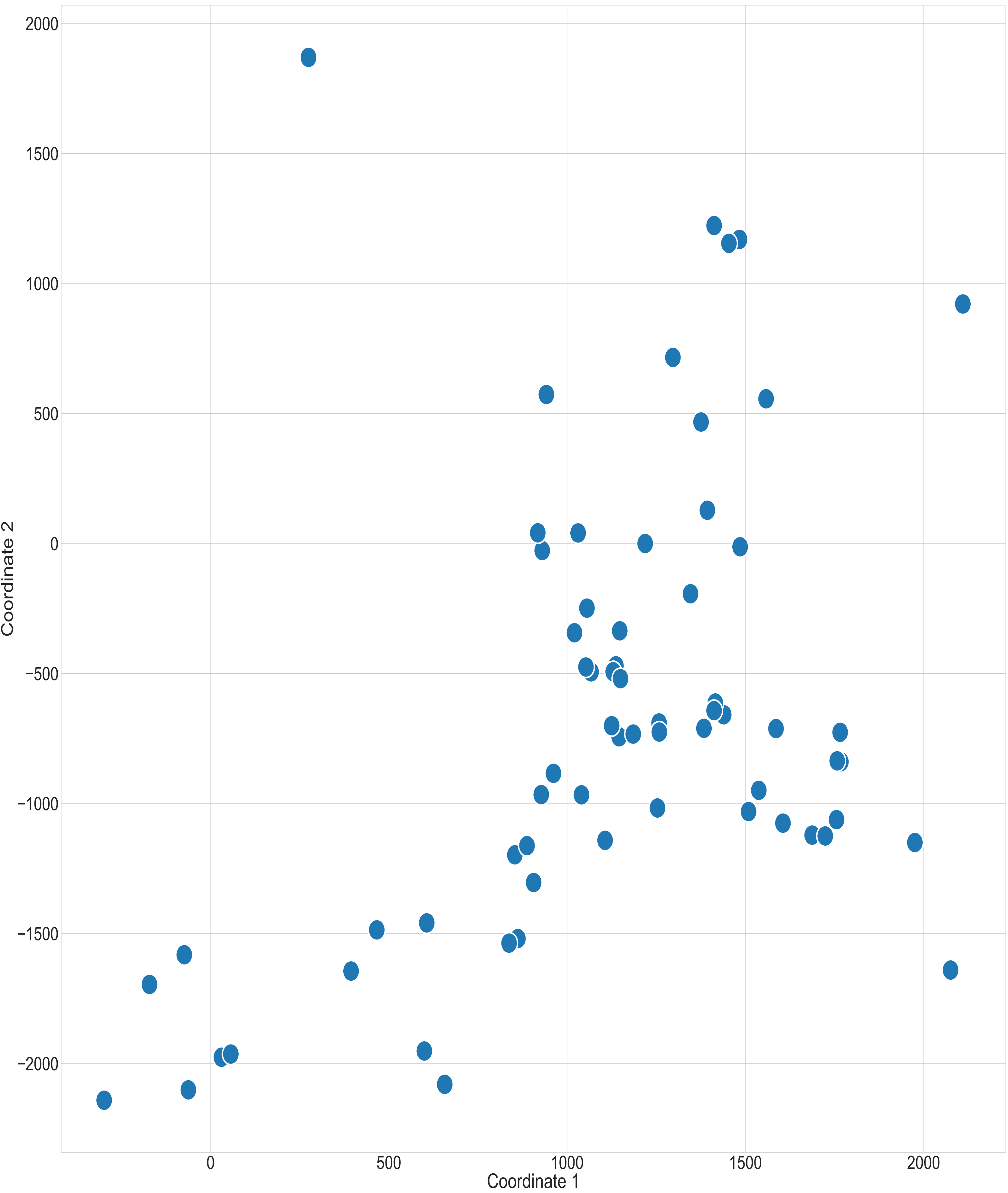
Supporting Information Table of Contents

Contents **Page**

Table S1. Statistical information regarding the distribution of *Micromonospora* and *Micromonospora_E* BGCs by the distance of the BGC to the closest GCF in BiG-SLiCE (T=900, GCFs=29,955). Mann-Whitney U-test results are reported as well.

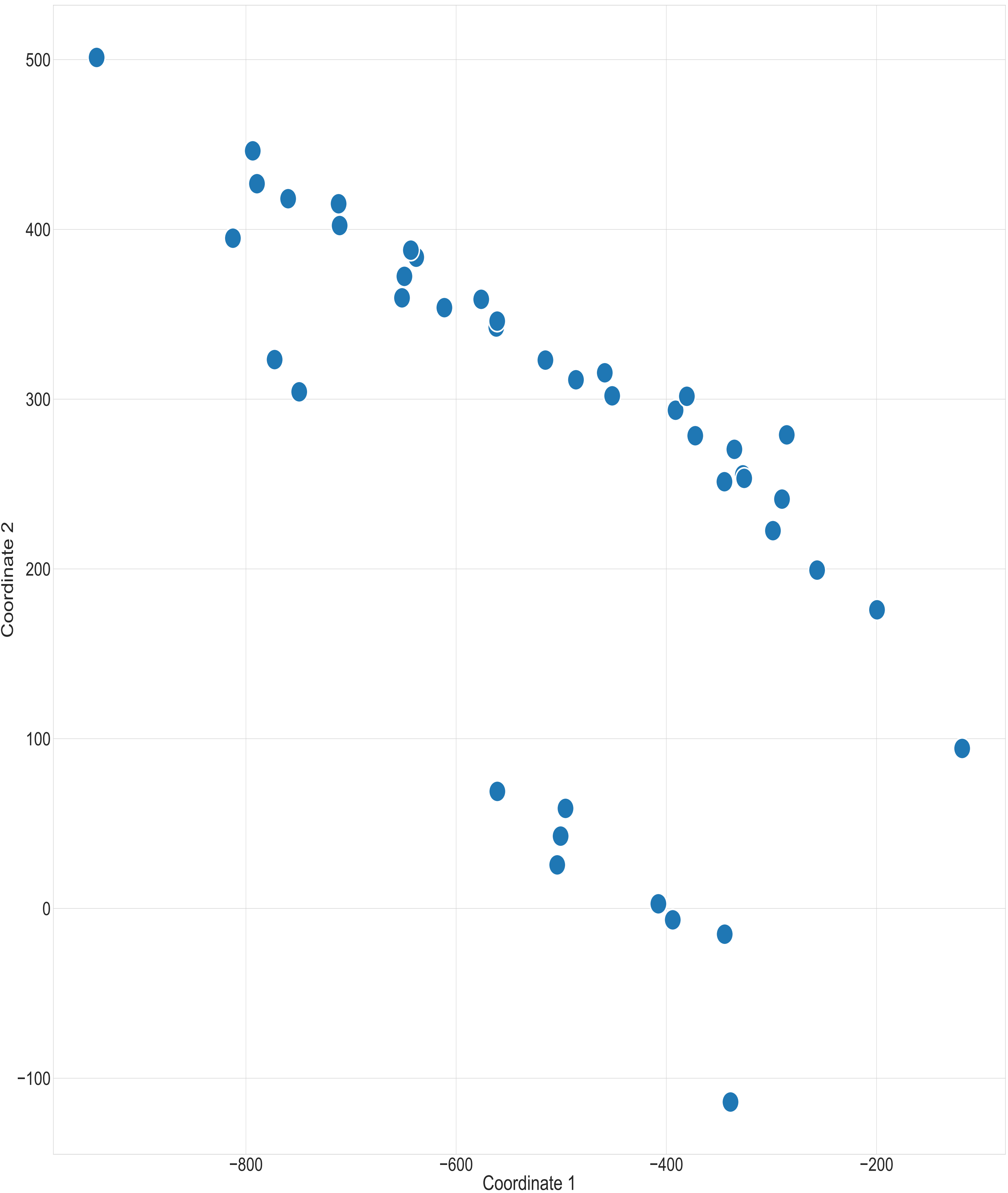
Statistics:	***Micromonospora:***	***Micromonospora_E:***
Mean:	1006.01	1174.01
Median:	908	1004
Standard Deviation:	486.34	555.8
Standard Error:	20.37	45.84
P-value:	6.842e-04	

Figure – S1 :



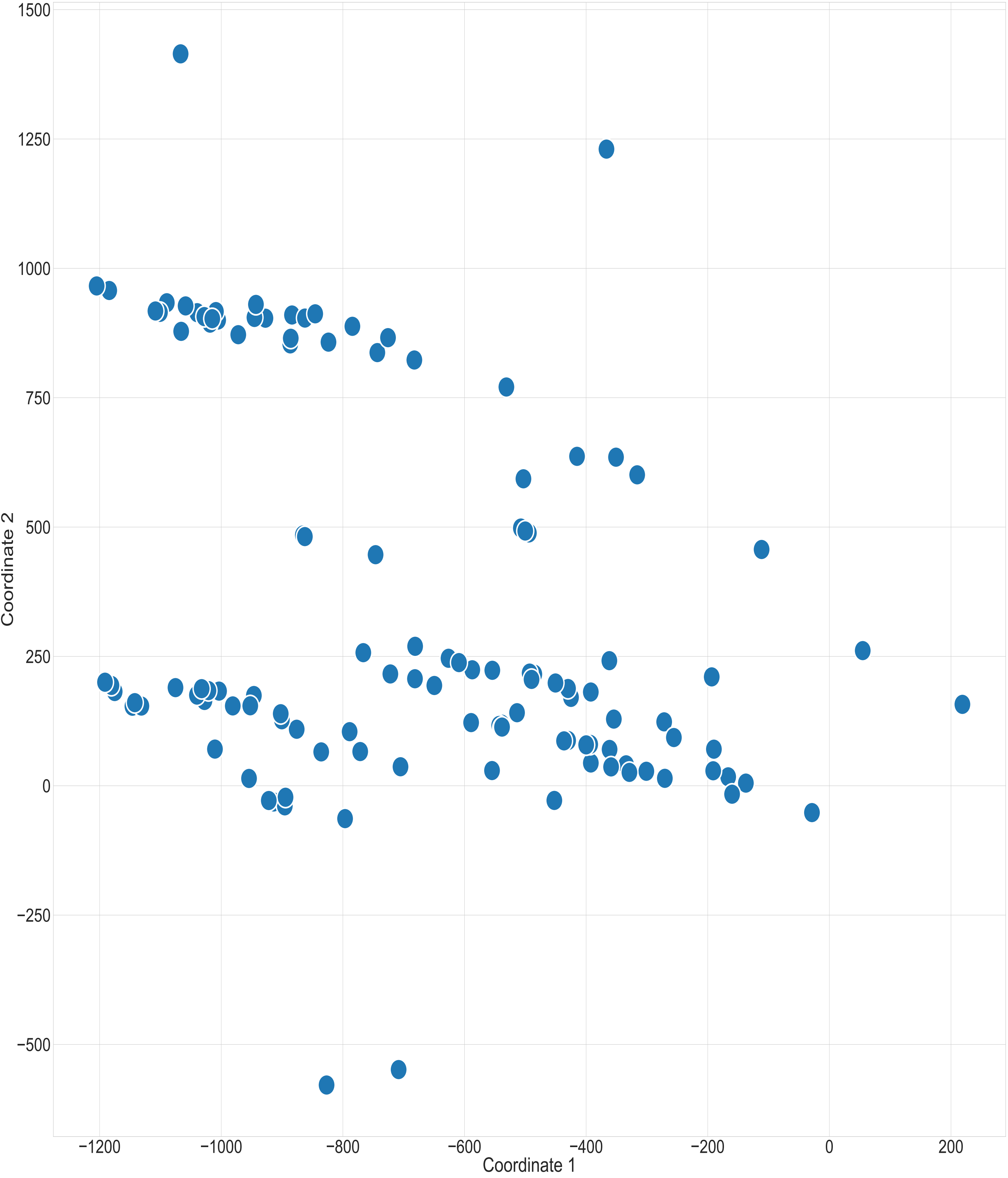
Scatter plot of PKS-NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using Chebyshev pairwise distance. Each dot represents an individual BGC. Distance between BGCs is associated with GCFs in BiG-SLiCE they were most similar to.

Figure – S2 :



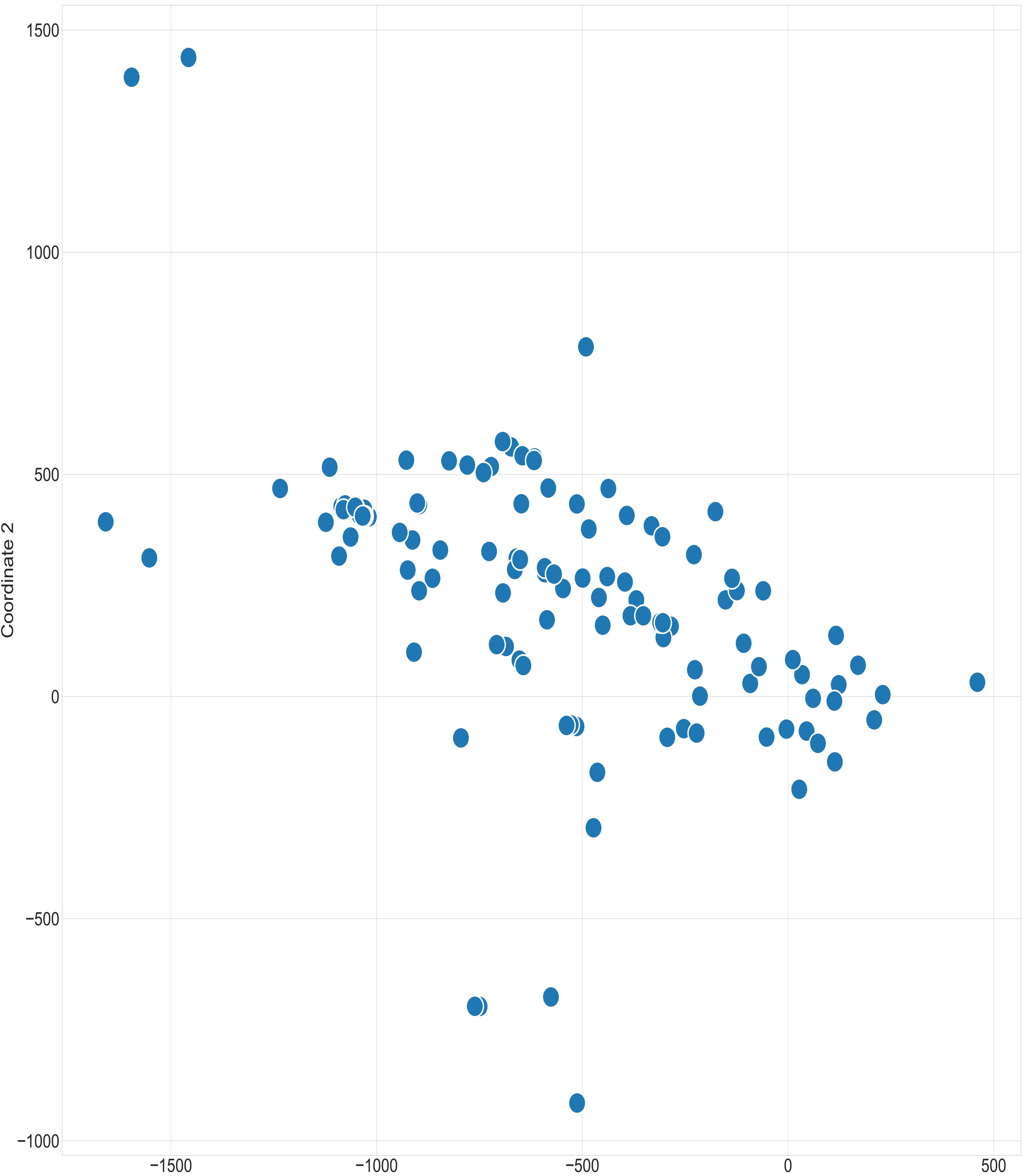
Scatter plot of Siderophore *Micromonosporaceae* BGCs analyzed via multidimensional scaling using Chebyshev pairwise distance. Each dot represents an individual BGC. Distance between BGCs is associated with GCFs in BiG-SLiCE they were most similar to.

Figure – S3 :



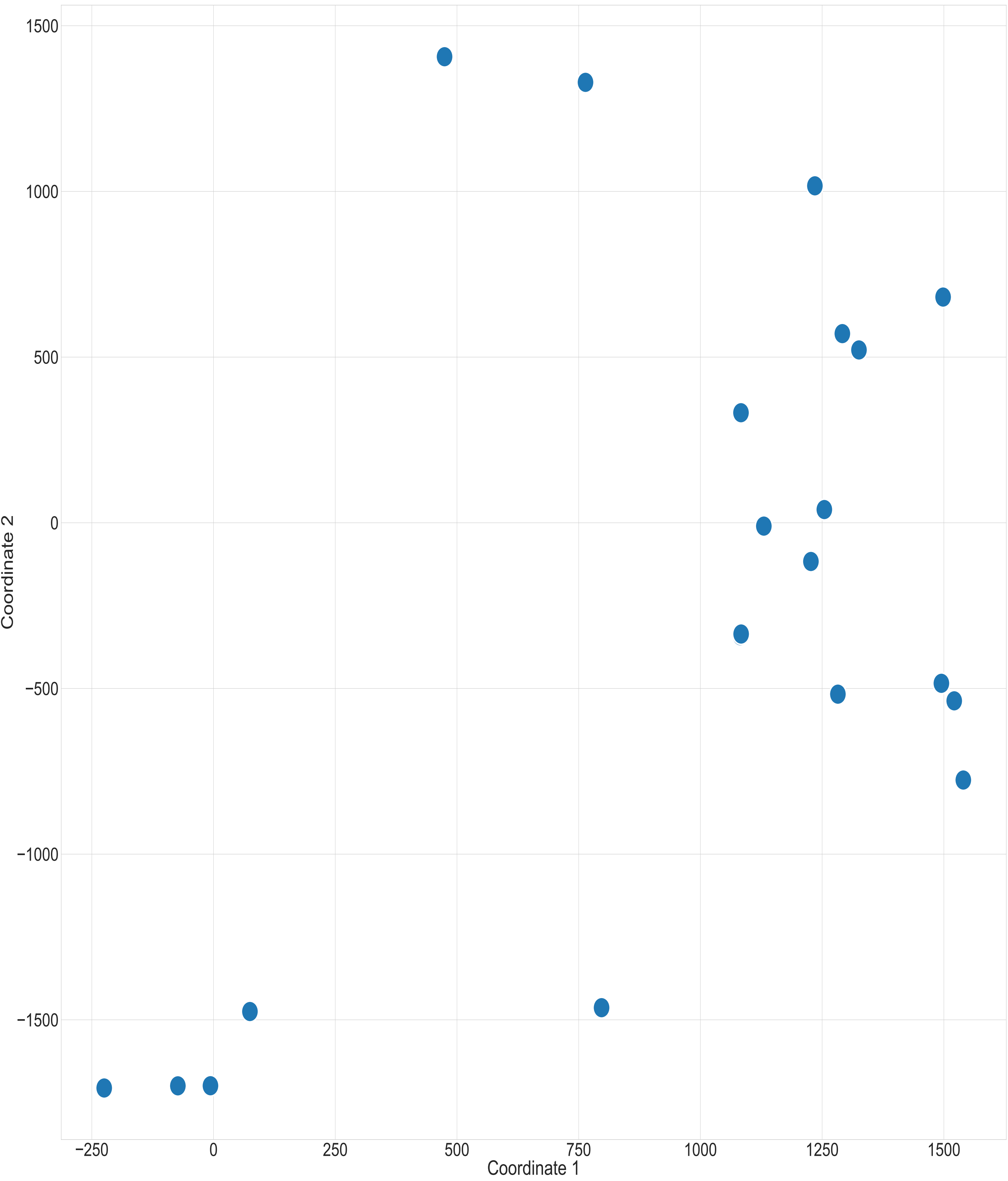
Scatter plot of Terpene *Micromonosporaceae* BGCs analyzed via multidimensional scaling using Chebyshev pairwise distance. Each dot represents an individual BGC. Distance between BGCs is associated with GCFs in BiG-SLiCE they were most similar to.

Figure – S4 :



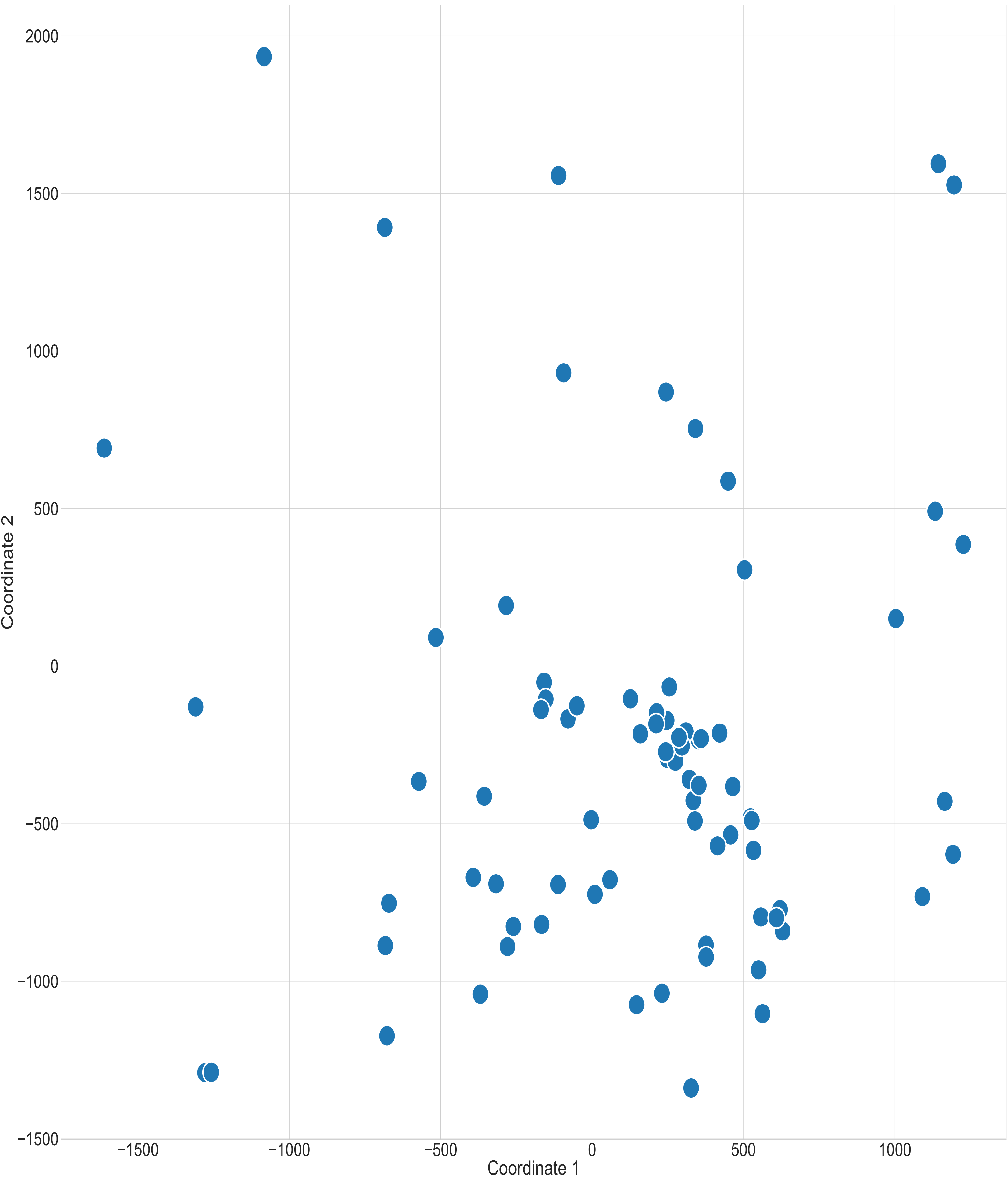
Scatter plot of RIPP *Micromonosporaceae* BGCs analyzed via multidimensional scaling using Chebyshev pairwise distance. Each dot represents an individual BGC. Distance between BGCs is associated with GCFs in BiG-SLiCE they were most similar to.

Figure – S5 :



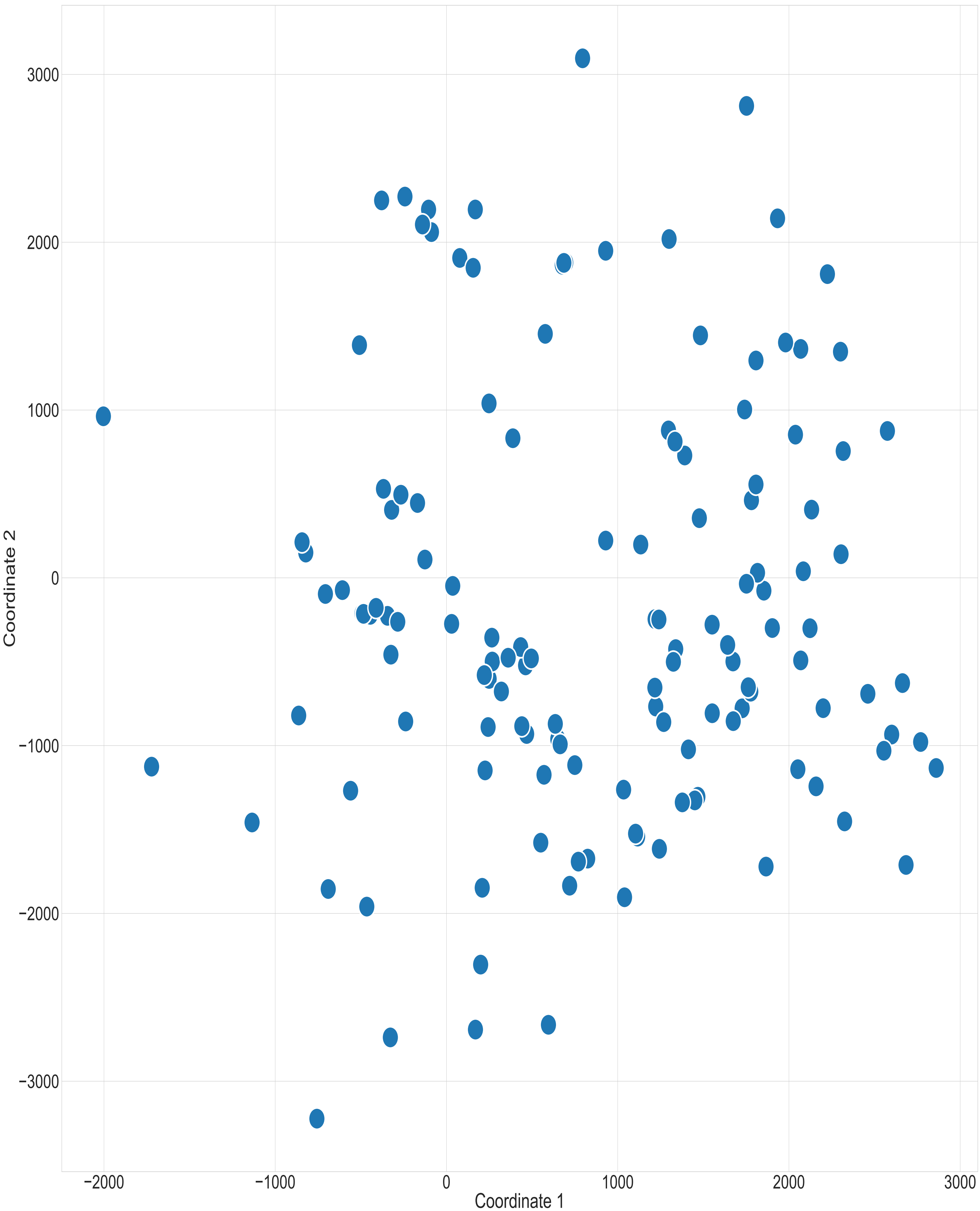
Scatter plot of PKS-1 *Micromonosporaceae* BGCs analyzed via multidimensional scaling using Chebyshev pairwise distance. Each dot represents an individual BGC. Distance between BGCs is associated with GCFs in BiG-SLiCE they were most similar to.

Figure – S6 :



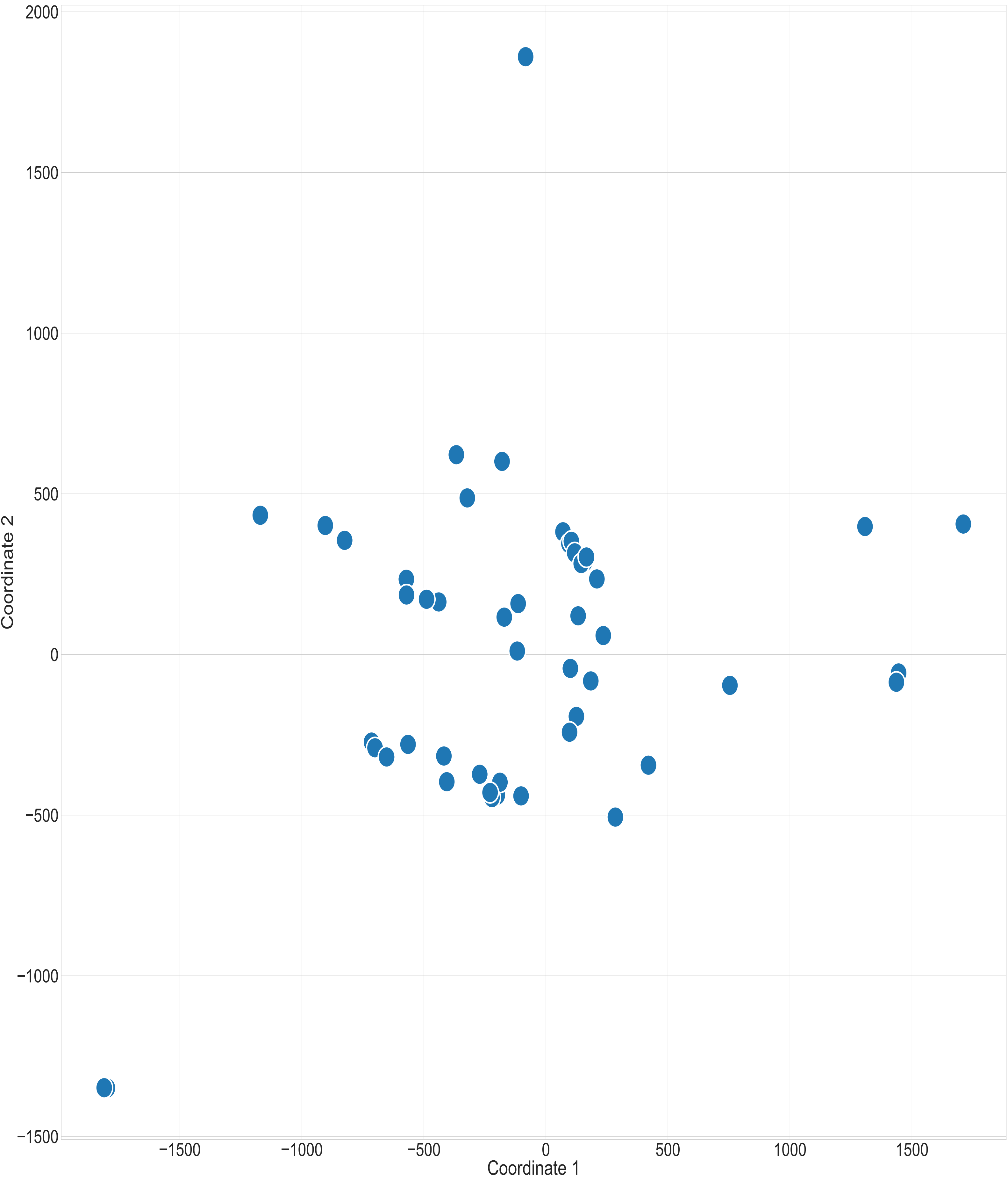
Scatter plot of NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using Chebyshev pairwise distance. Each dot represents an individual BGC. Distance between BGCs is associated with GCFs in BiG-SLiCE they were most similar to.

Figure – S7 :



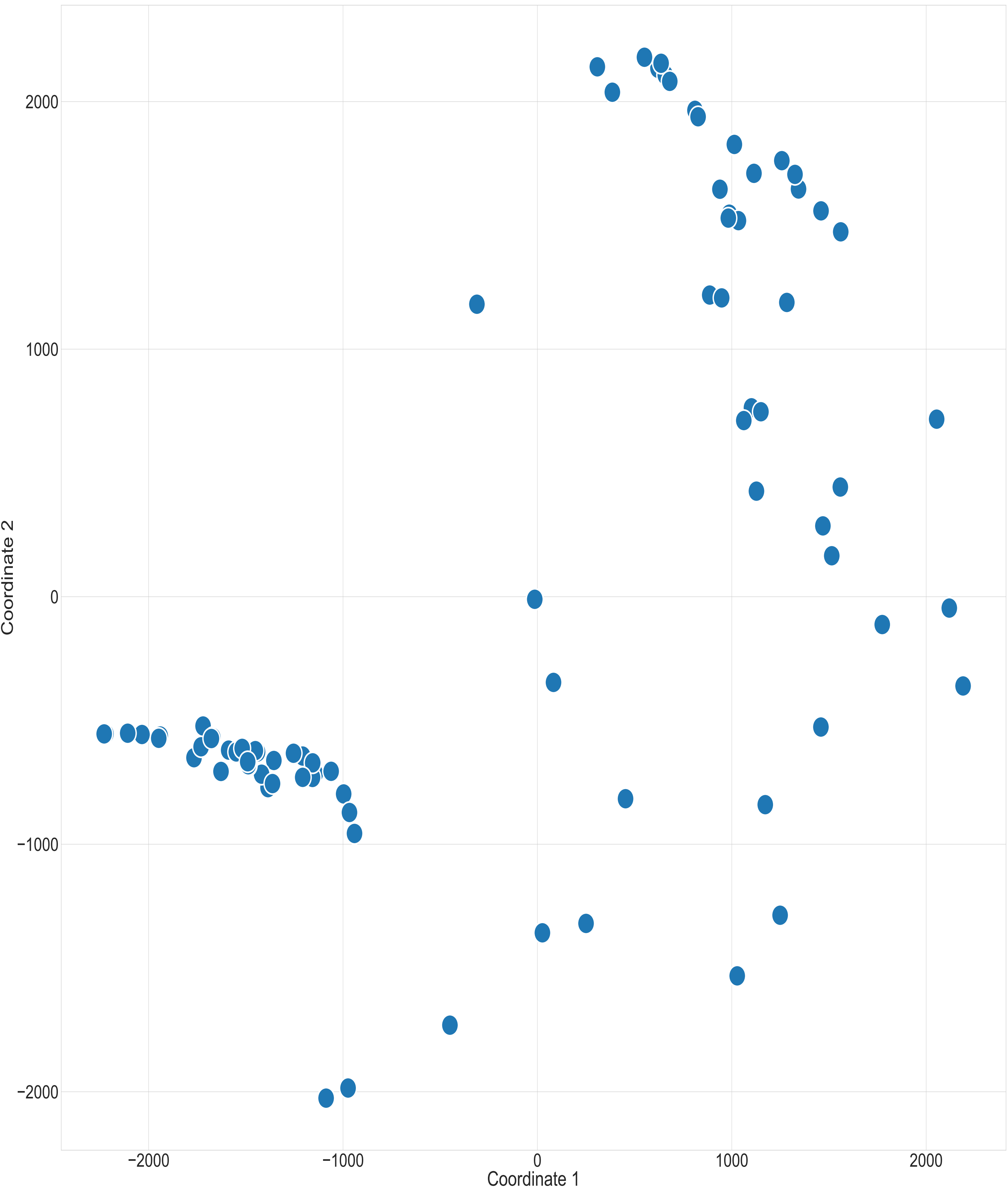
Scatter plot of Hybrid-(Not-PKS-NRPS) *Micromonosporaceae* BGCs analyzed via multidimensional scaling using Chebyshev pairwise distance. Each dot represents an individual BGC. Distance between BGCs is associated with GCFs in BiG-SLiCE they were most similar to.

Figure – S8 :



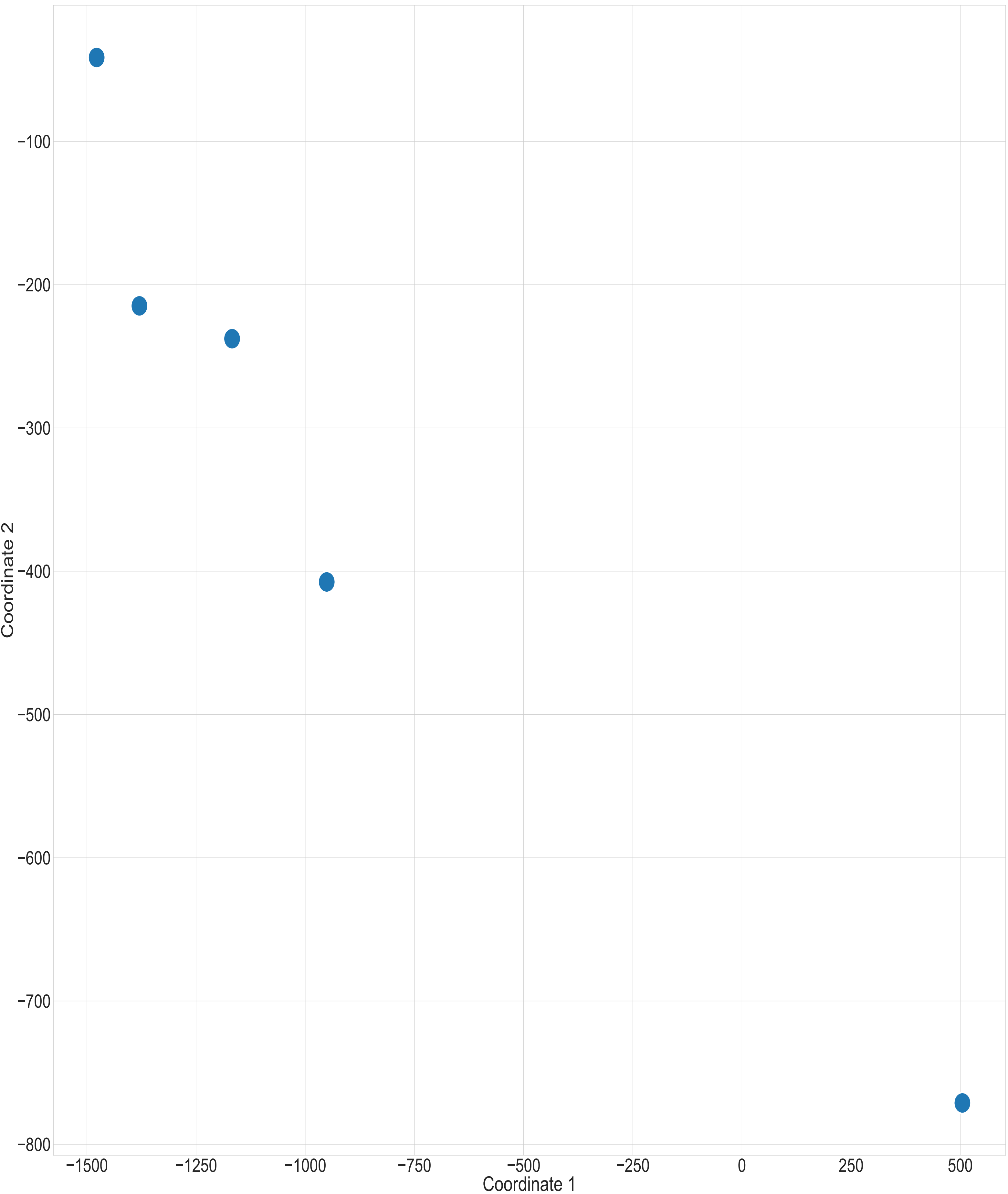
Scatter plot of Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using Chebyshev pairwise distance. Each dot represents an individual BGC. Distance between BGCs is associated with GCFs in BiG-SLiCE they were most similar to.

Figure – S9 :



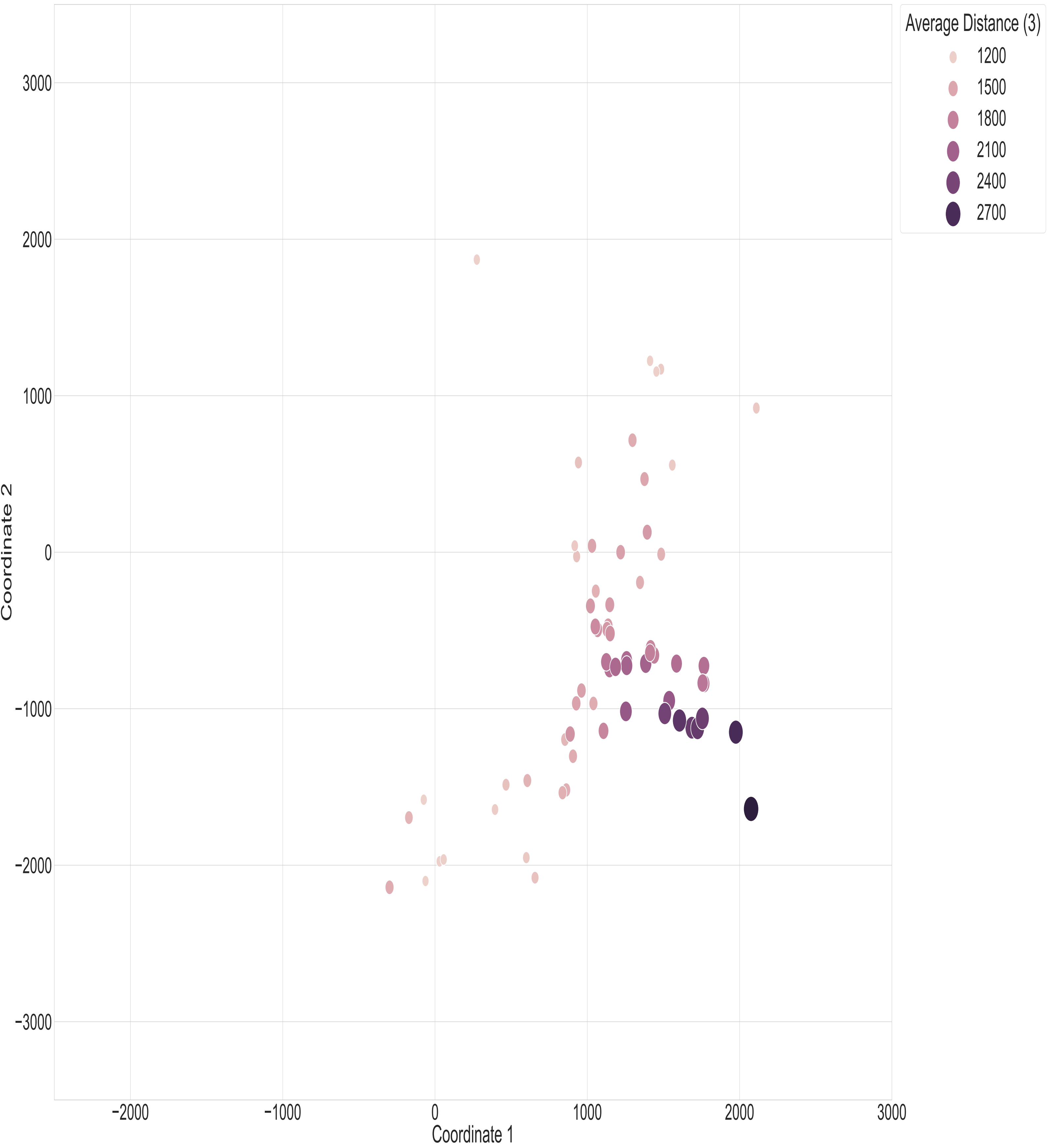
Scatter plot of PKS-Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using Chebyshev pairwise distance. Each dot represents an individual BGC. Distance between BGCs is associated with GCFs in BiG-SLiCE they were most similar to.

Figure – S10 :



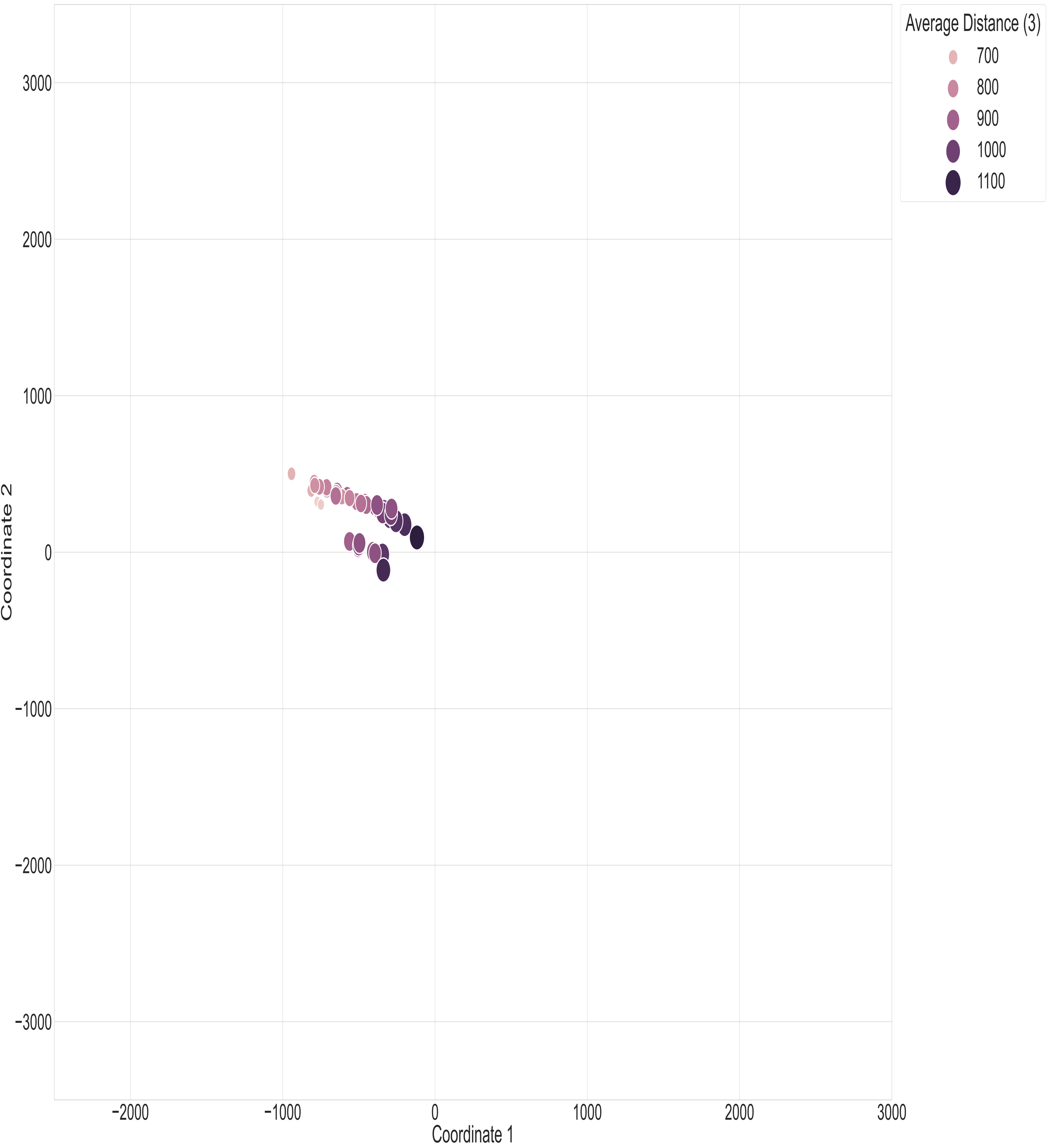
Scatter plot of Saccharides *Micromonosporaceae* BGCs analyzed via multidimensional scaling using Chebyshev pairwise distance. Each dot represents an individual BGC. Distance between BGCs is associated with GCFs in BiG-SLiCE they were most similar to.

Figure – S11 :



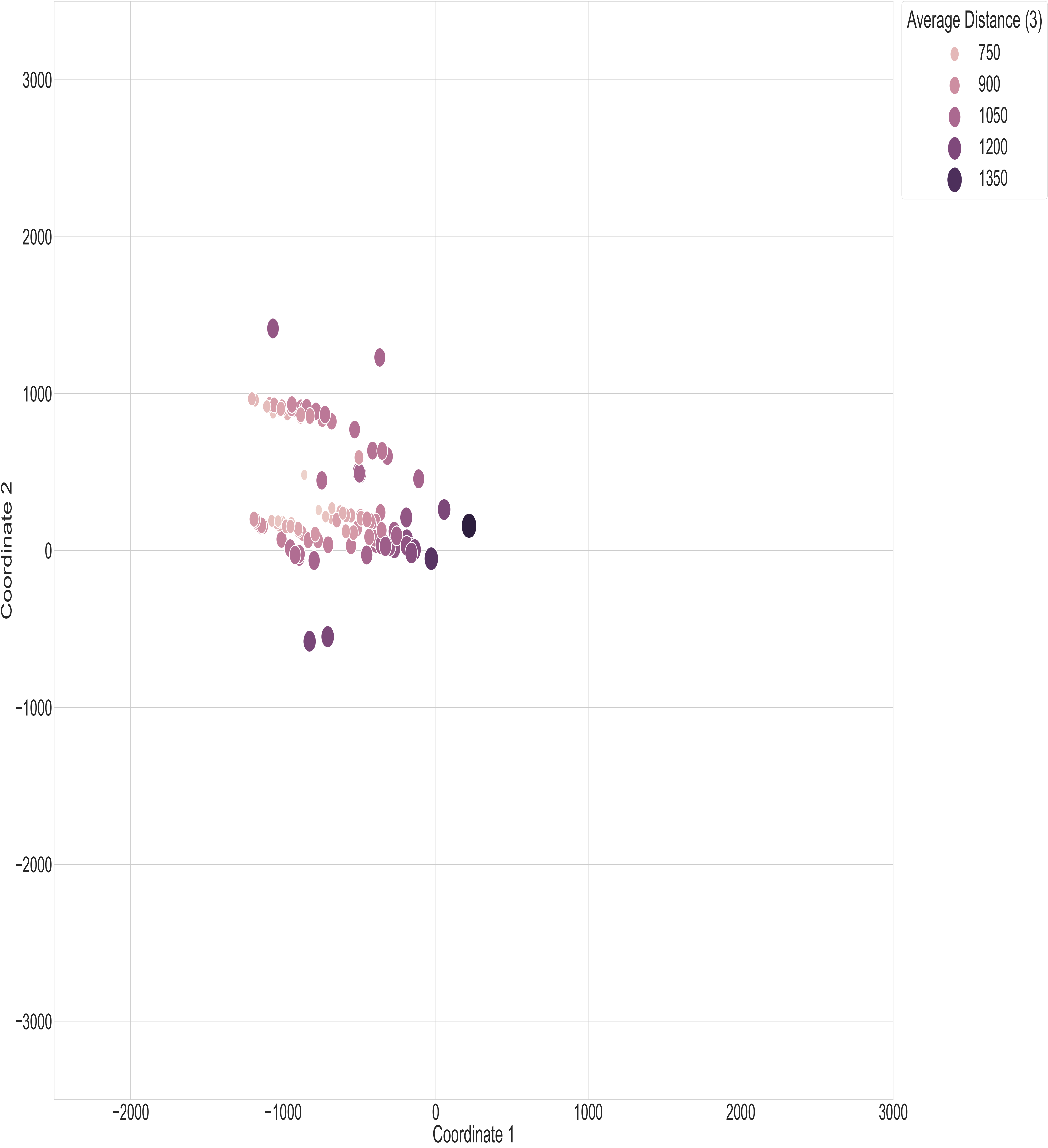
Scatter plot of PKS-NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using Chebyshev pairwise distance. Each dot represents an individual BGC. The hue and size of the dots were scaled based on the average distance of the individual BGC, across the entire dataset, to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S12 :



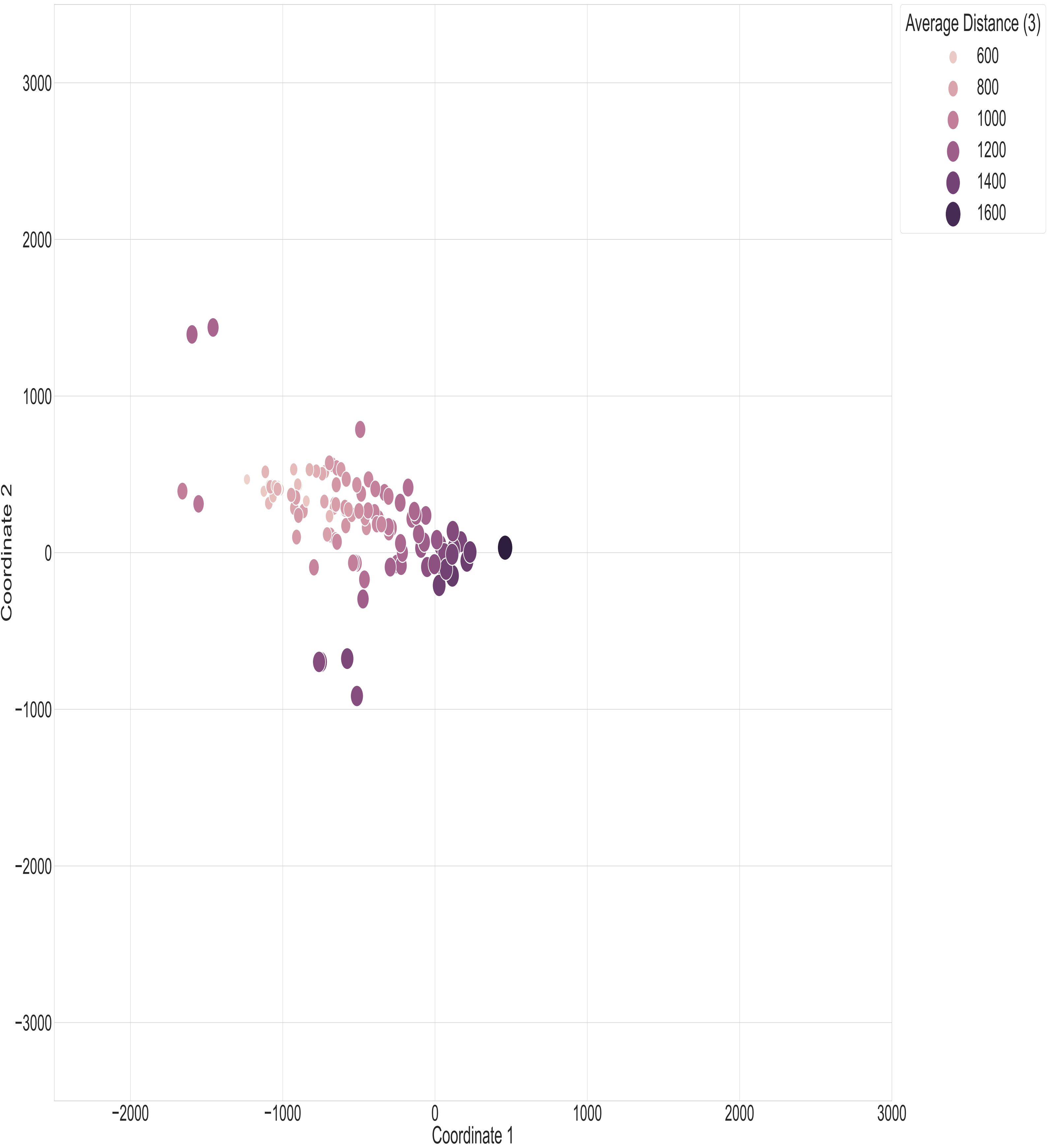
Scatter plot of Siderophore *Micromonosporaceae* BGCs analyzed via multidimensional scaling using Chebyshev pairwise distance. Each dot represents an individual BGC. The hue and size of the dots were scaled based on the average distance of the individual BGC, across the entire dataset, to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S13 :



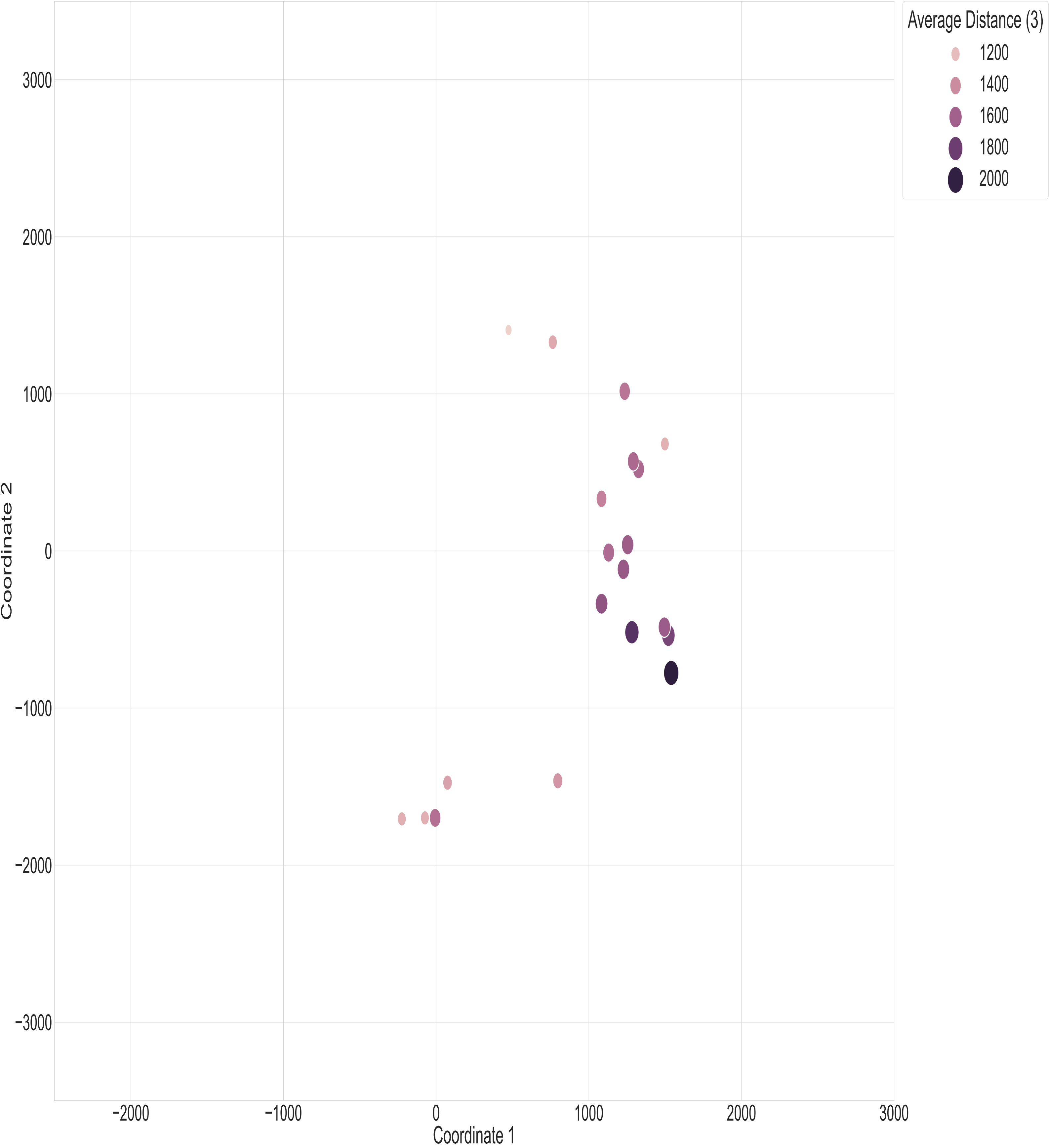
Scatter plot of Terpene *Micromonosporaceae* BGCs analyzed via multidimensional scaling using Chebyshev pairwise distance. Each dot represents an individual BGC. The hue and size of the dots were scaled based on the average distance of the individual BGC, across the entire dataset, to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S14 :



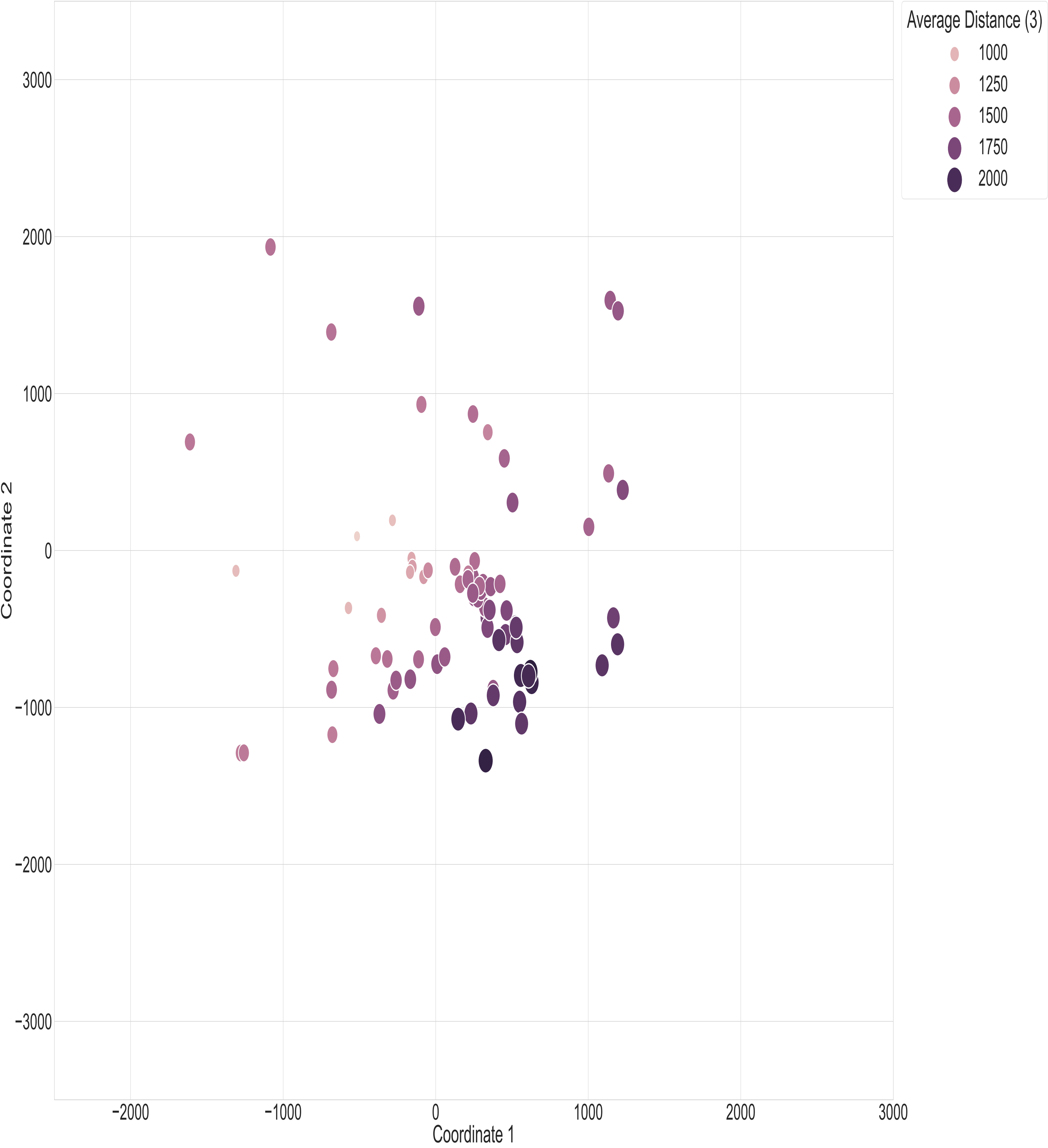
Scatter plot of RIPP *Micromonosporaceae* BGCs analyzed via multidimensional scaling using Chebyshev pairwise distance. Each dot represents an individual BGC. The hue and size of the dots were scaled based on the average distance of the individual BGC, across the entire dataset, to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S15 :



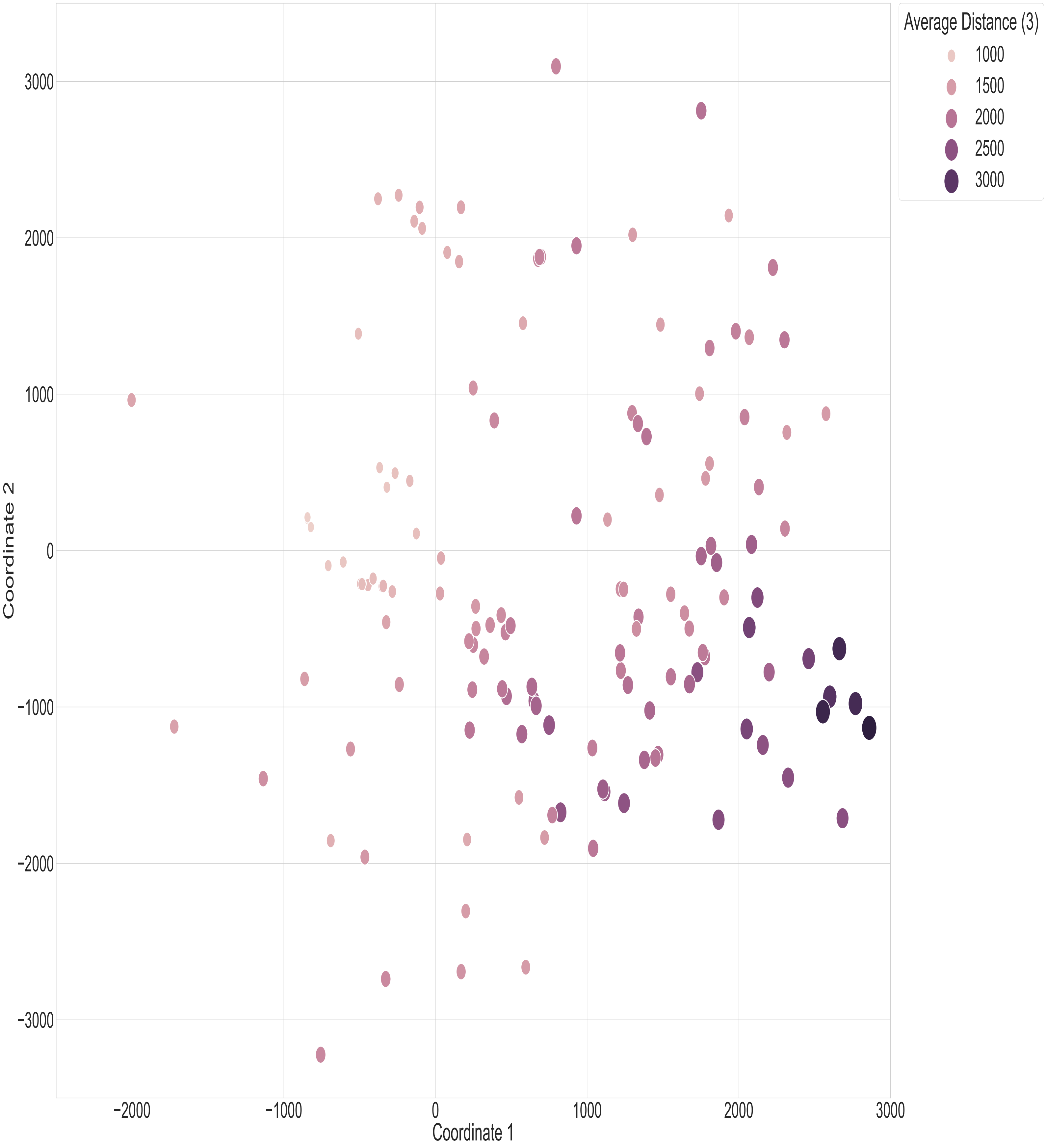
Scatter plot of PKS-1 *Micromonosporaceae* BGCs analyzed via multidimensional scaling using Chebyshev pairwise distance. Each dot represents an individual BGC. The hue and size of the dots were scaled based on the average distance of the individual BGC, across the entire dataset, to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S16 :



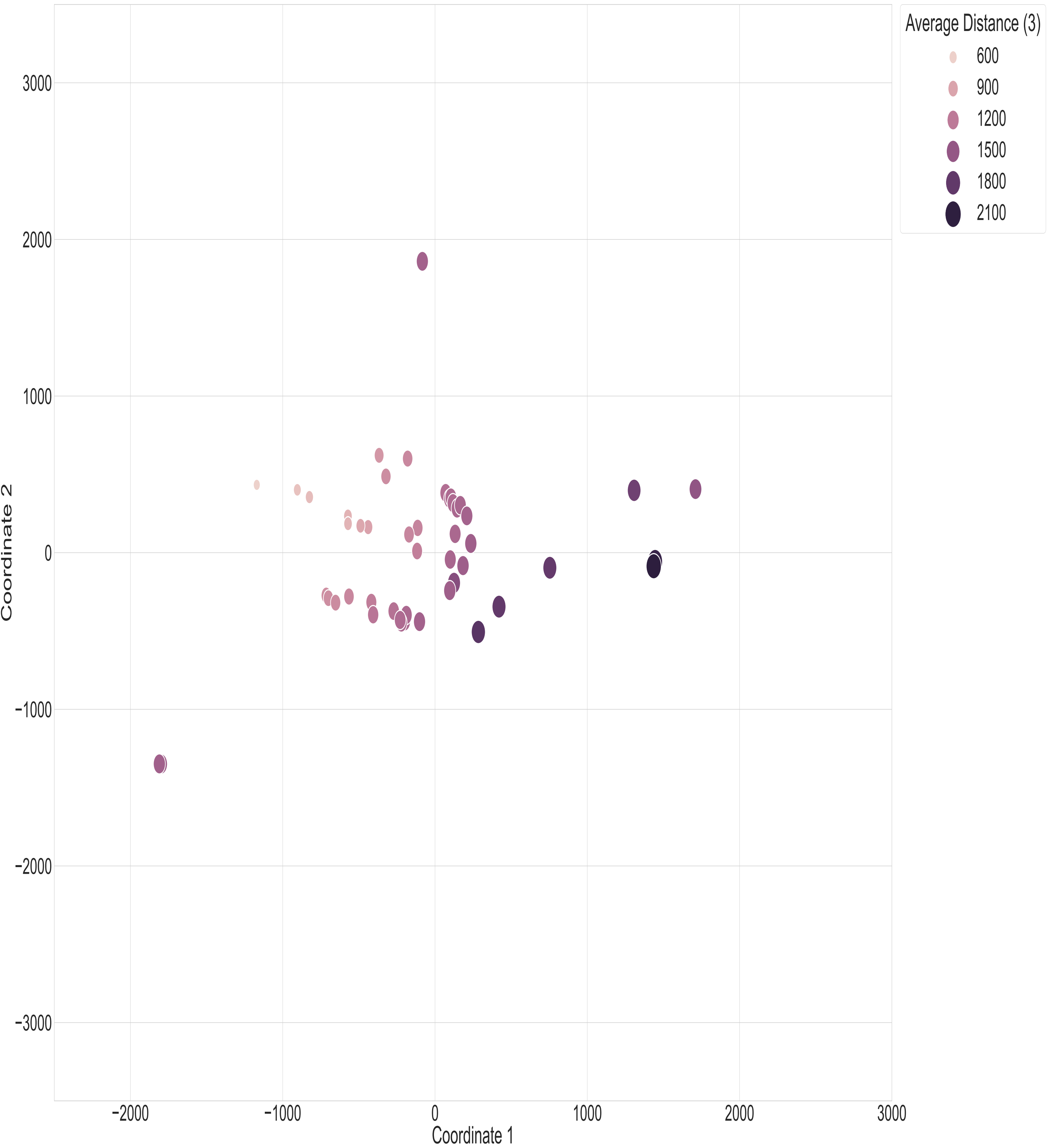
Scatter plot of NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using Chebyshev pairwise distance. Each dot represents an individual BGC. The hue and size of the dots were scaled based on the average distance of the individual BGC, across the entire dataset, to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S17 :



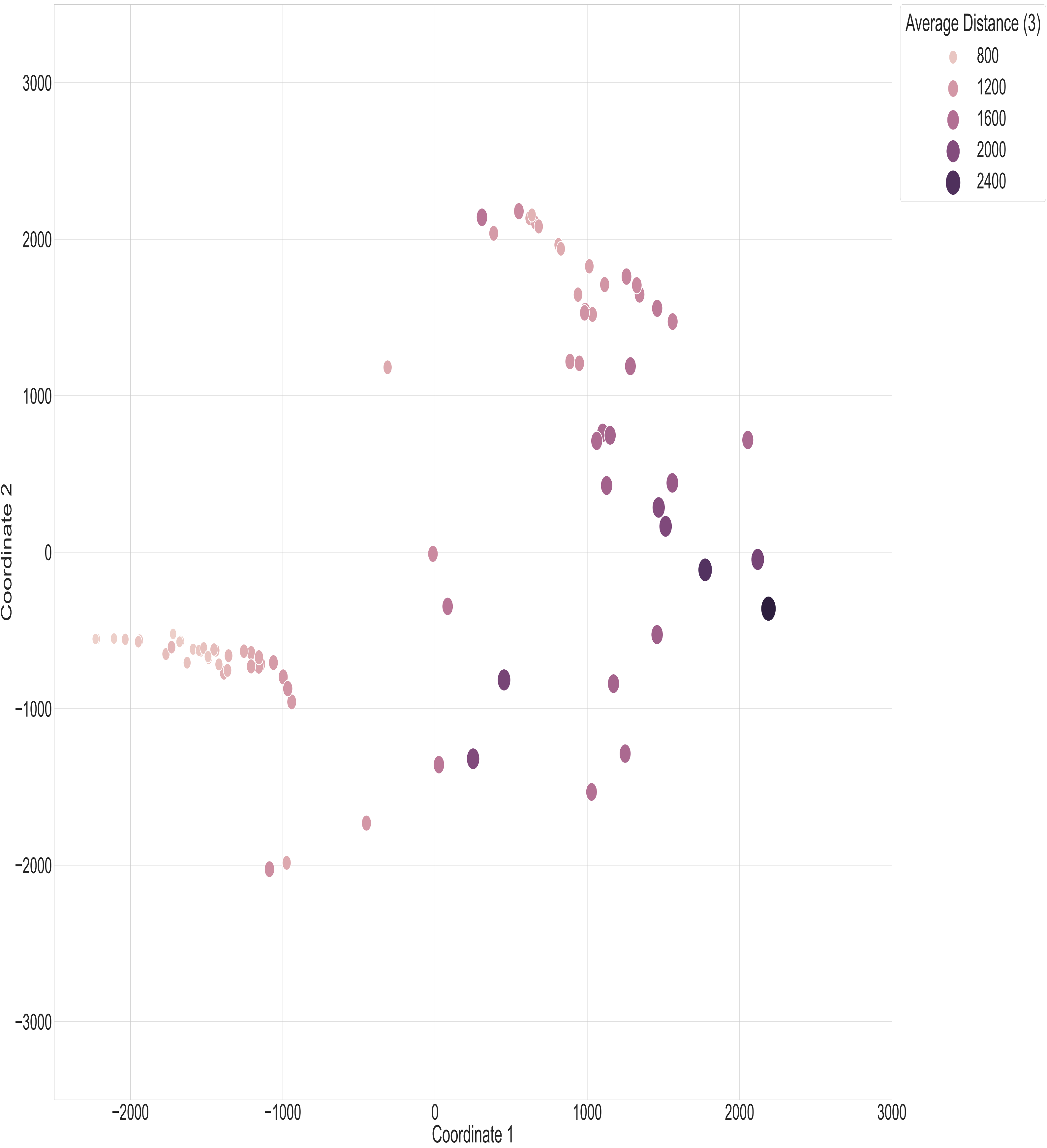
Scatter plot of Hybrid-(Not-PKS-NRPS) *Micromonosporaceae* BGCs analyzed via multidimensional scaling using Chebyshev pairwise distance. Each dot represents an individual BGC. The hue and size of the dots were scaled based on the average distance of the individual BGC, across the entire dataset, to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S18 :



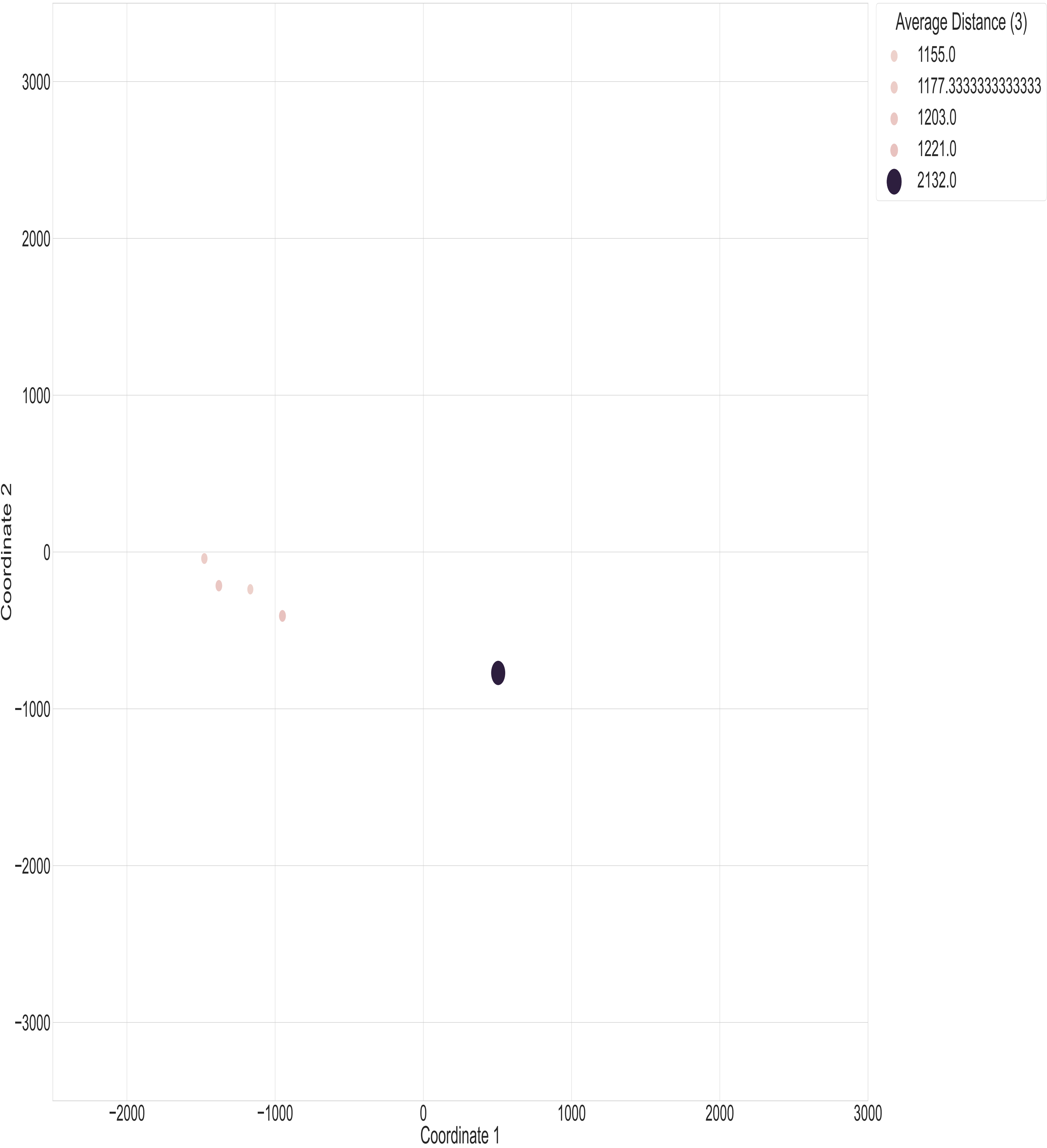
Scatter plot of Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using Chebyshev pairwise distance. Each dot represents an individual BGC. The hue and size of the dots were scaled based on the average distance of the individual BGC, across the entire dataset, to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S19 :



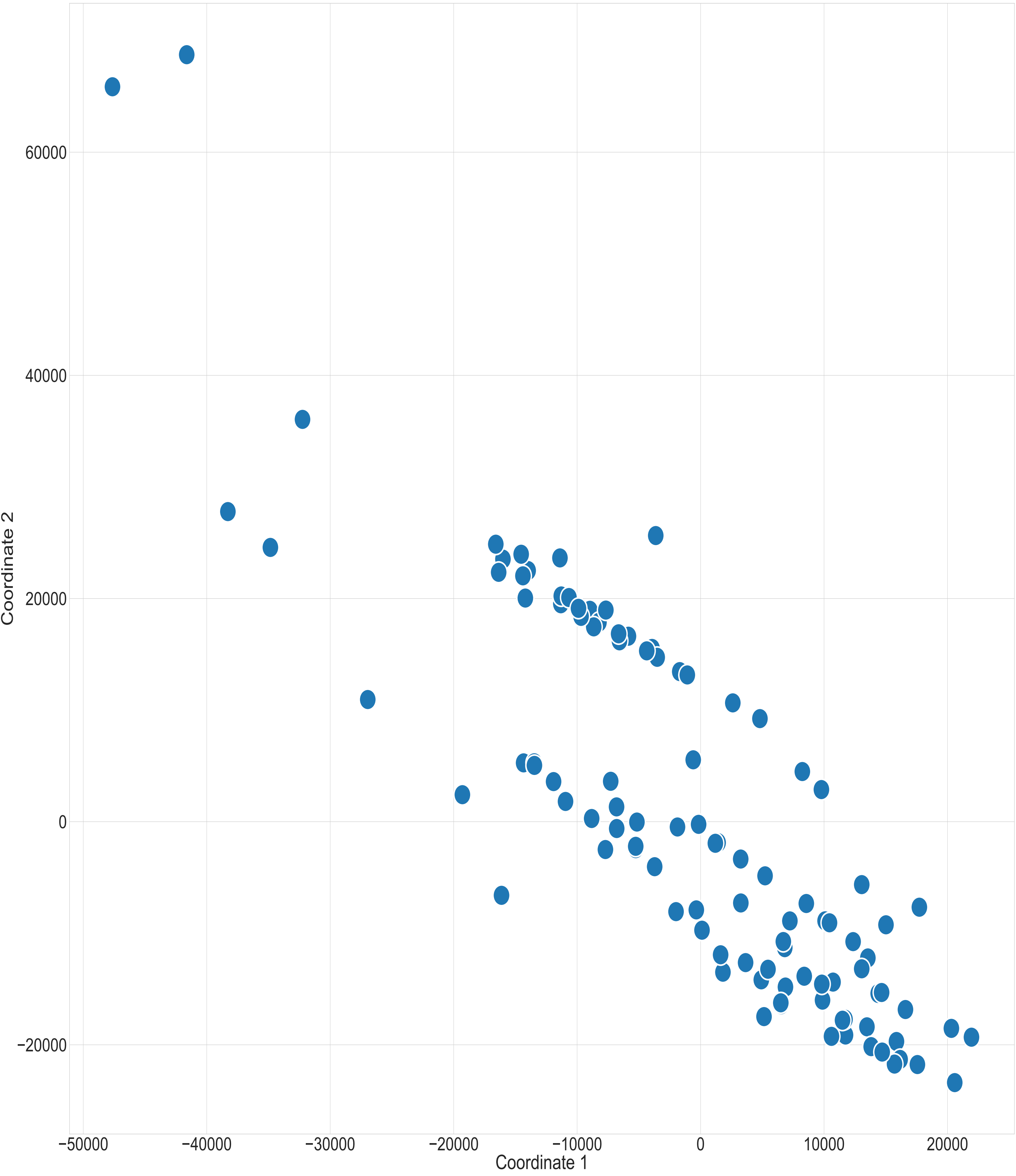
Scatter plot of PKS-Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using Chebyshev pairwise distance. Each dot represents an individual BGC. The hue and size of the dots were scaled based on the average distance of the individual BGC, across the entire dataset, to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S20 :



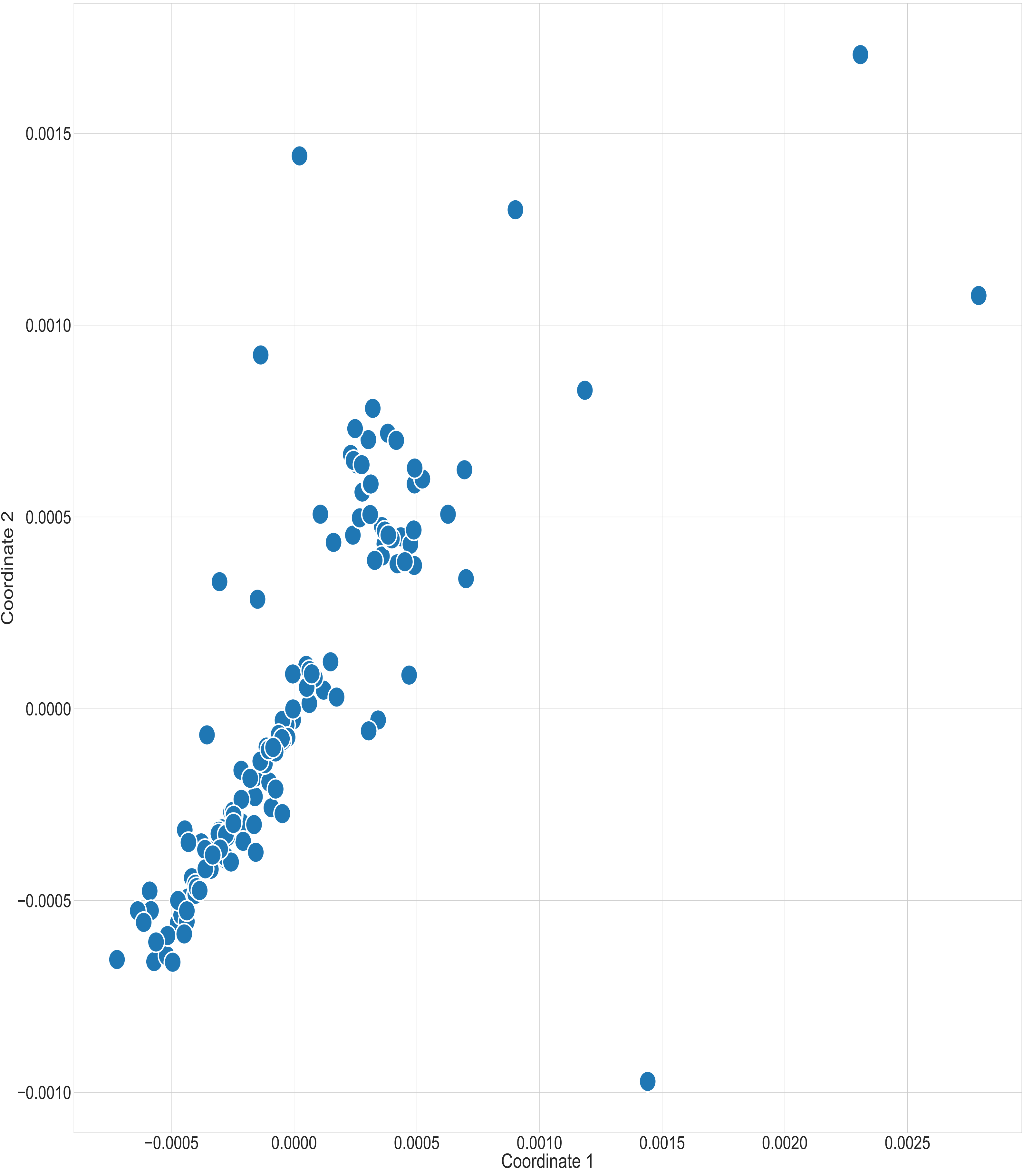
Scatter plot of Saccharides *Micromonosporaceae* BGCs analyzed via multidimensional scaling using Chebyshev pairwise distance. Each dot represents an individual BGC. The hue and size of the dots were scaled based on the average distance of the individual BGC, across the entire dataset, to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S21 :



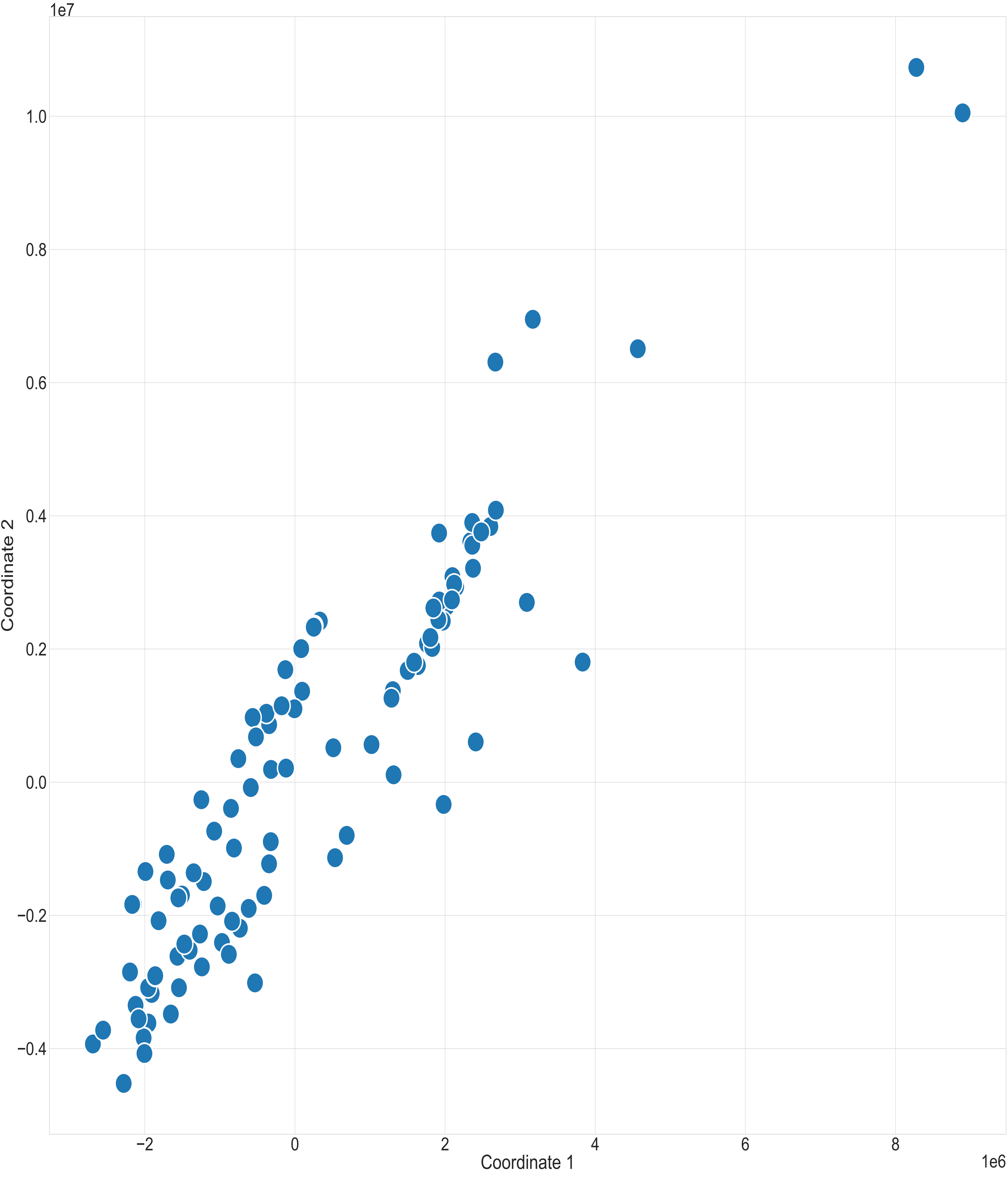
Scatter plot of Terpene *Micromonosporaceae* BGCs analyzed via multidimensional scaling using euclidean pairwise distance.
Only Terpene BGCs were considered for this analysis. Each dot represents an individual BGC.
Distance between Terpene BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S22 :



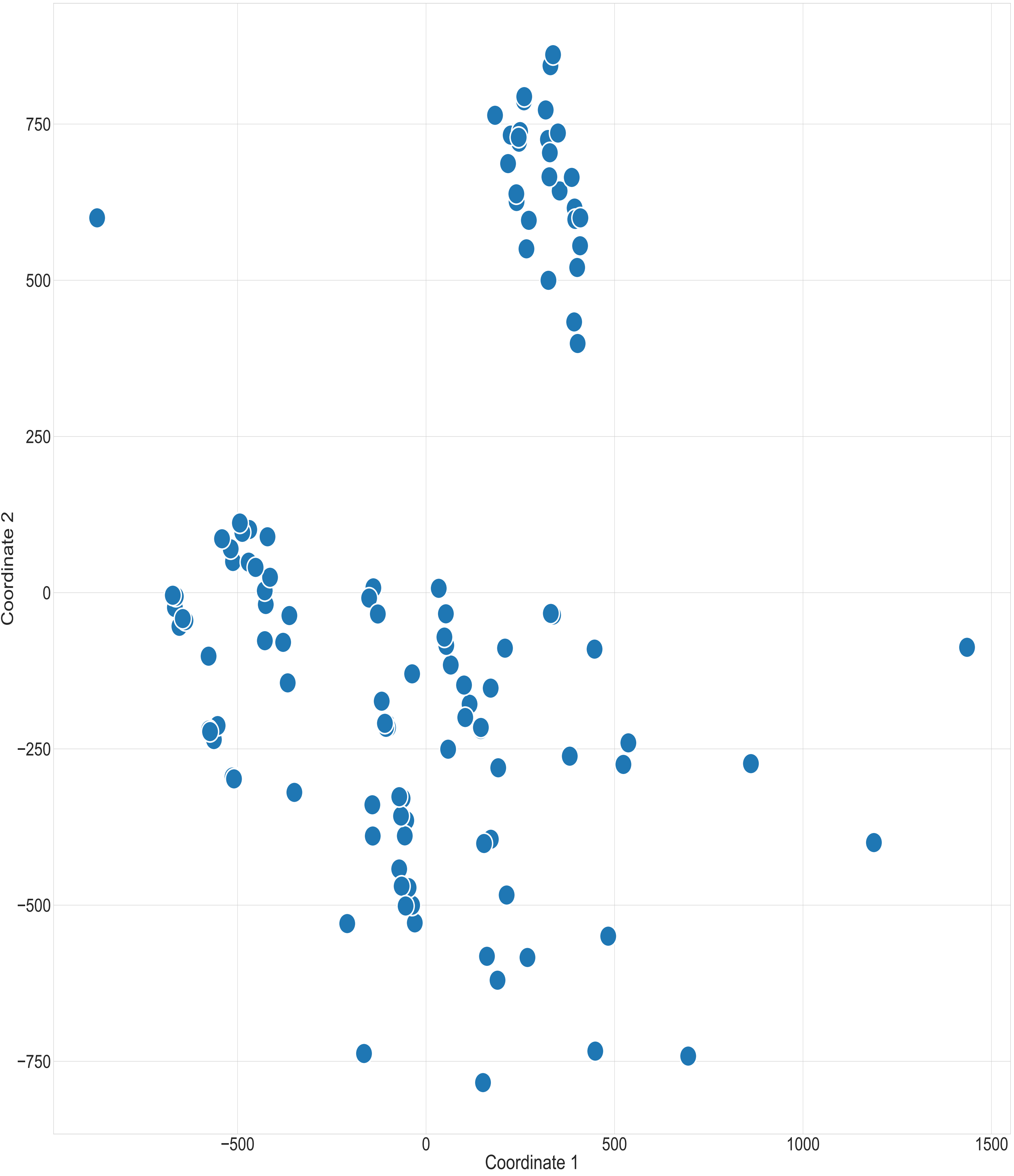
Scatter plot of Terpene *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cosine pairwise distance. Only Terpene BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Terpene BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S23 :



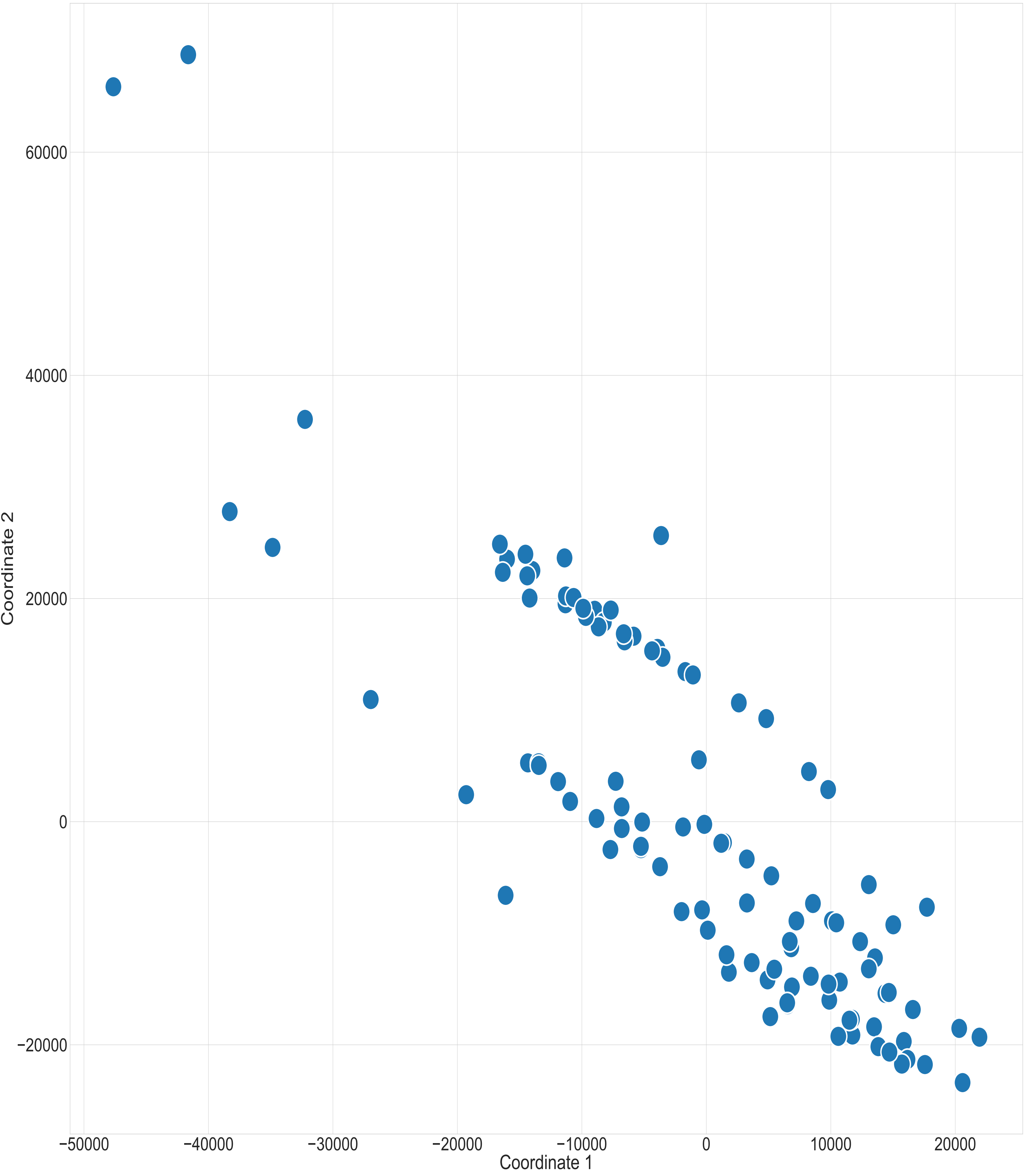
Scatter plot of Terpene *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cityblock pairwise distance.
Only Terpene BGCs were considered for this analysis. Each dot represents an individual BGC.
Distance between Terpene BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S24 :



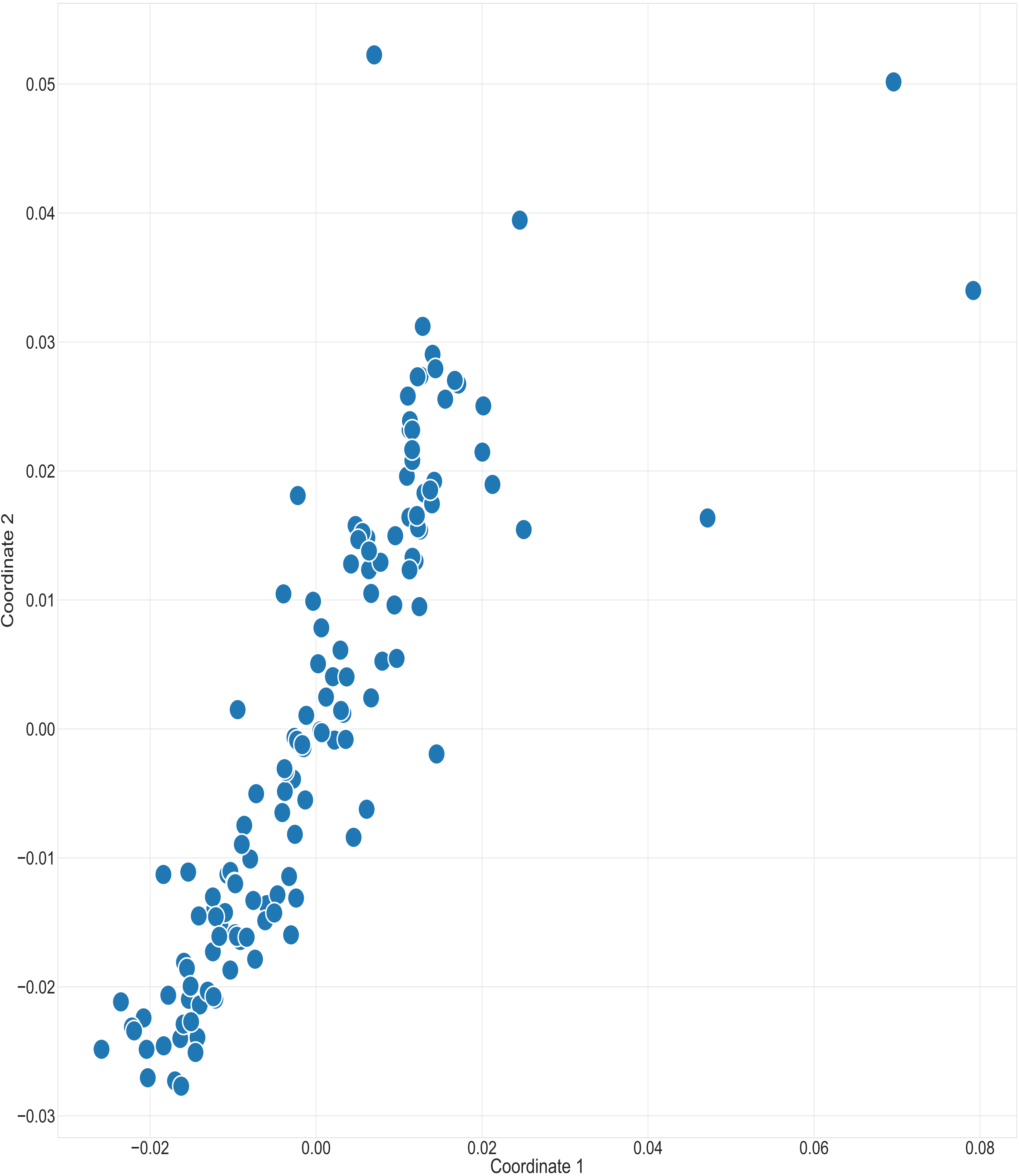
Scatter plot of Terpene *Micromonosporaceae* BGCs analyzed via multidimensional scaling using chebyshev pairwise distance. Only Terpene BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Terpene BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S25 :



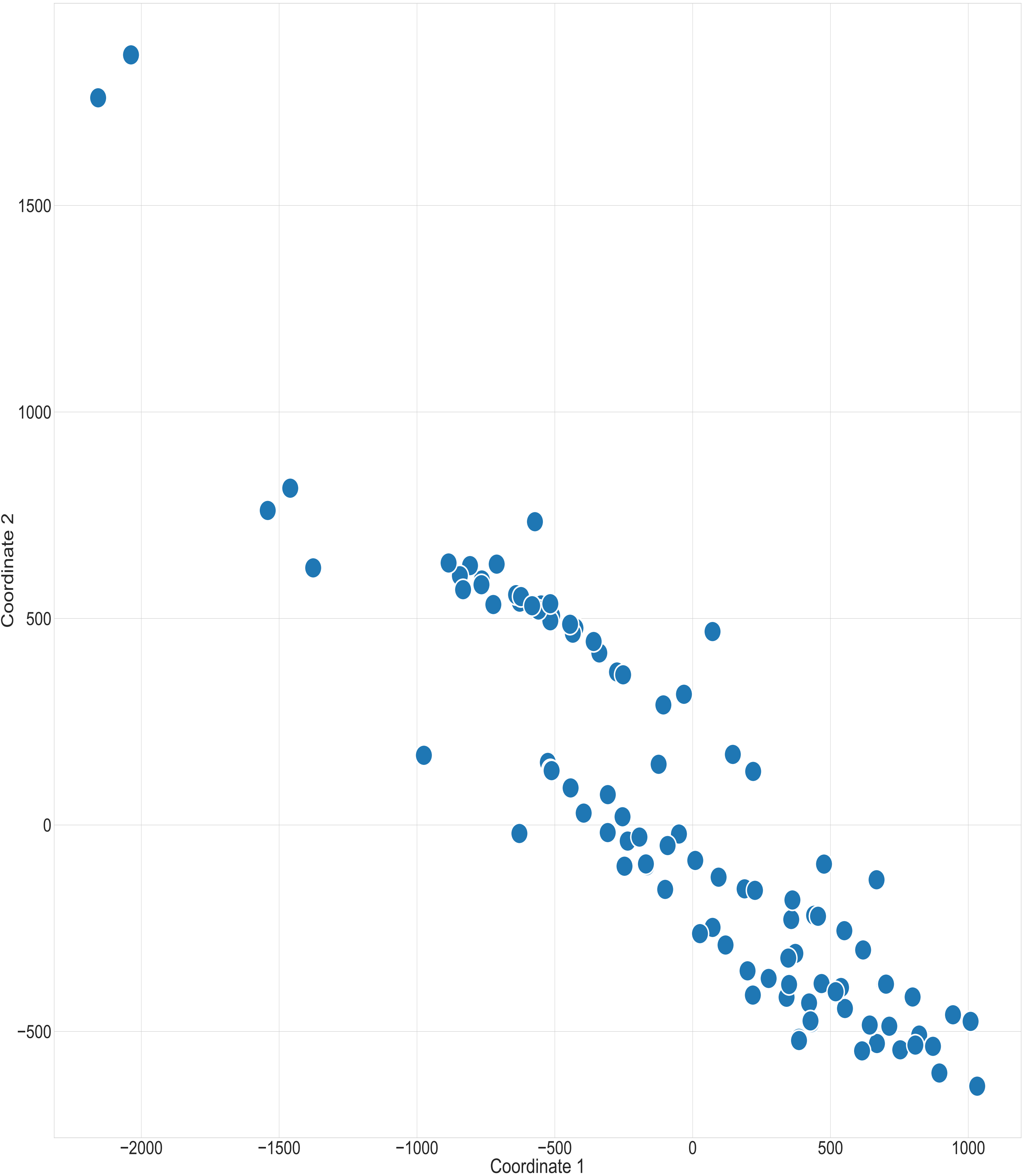
Scatter plot of Terpene *Micromonosporaceae* BGCs analyzed via multidimensional scaling using l2 pairwise distance. Only Terpene BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Terpene BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S26 :



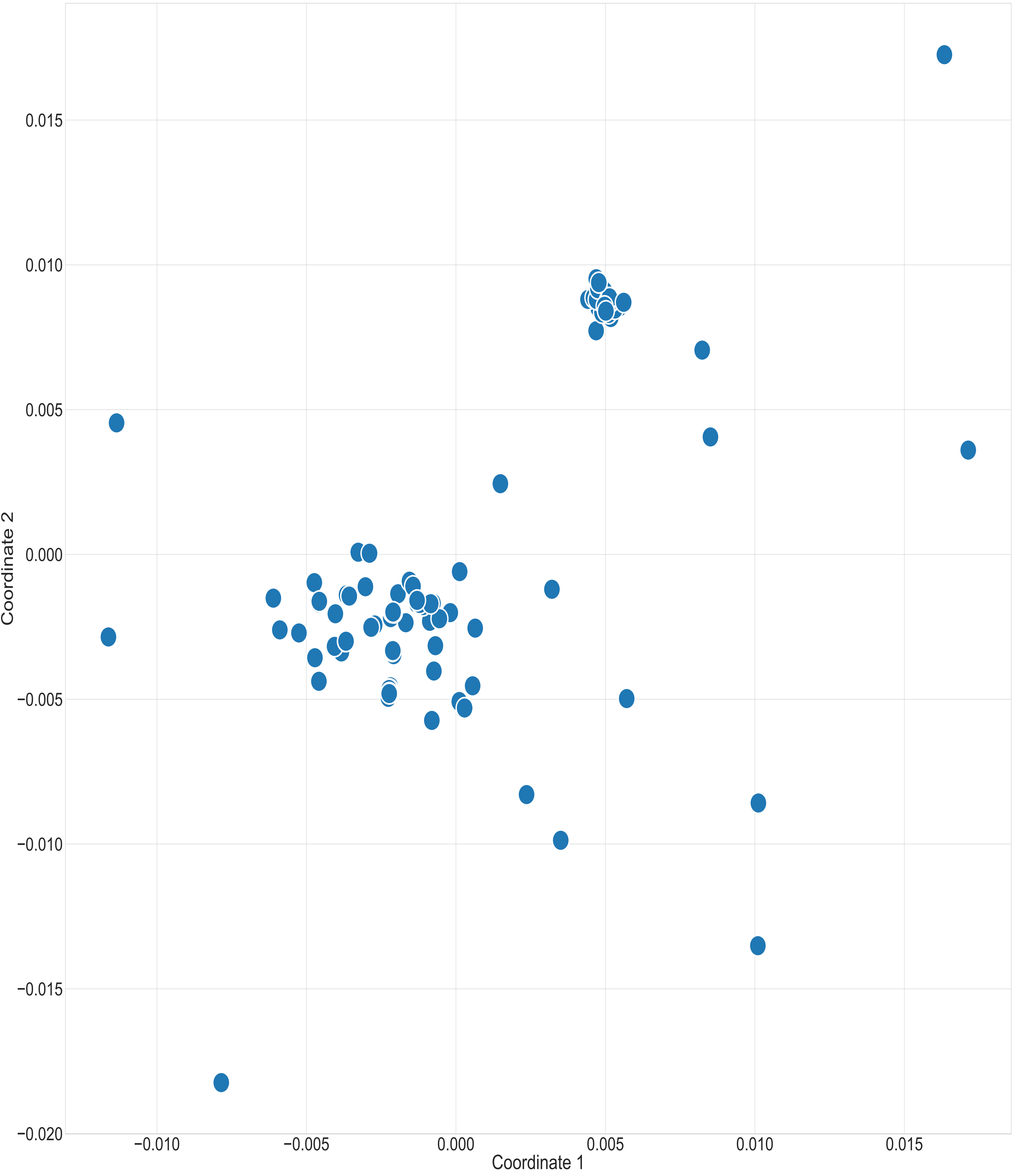
Scatter plot of Terpene *Micromonosporaceae* BGCs analyzed via multidimensional scaling using braycurtis pairwise distance. Only Terpene BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Terpene BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S27 :



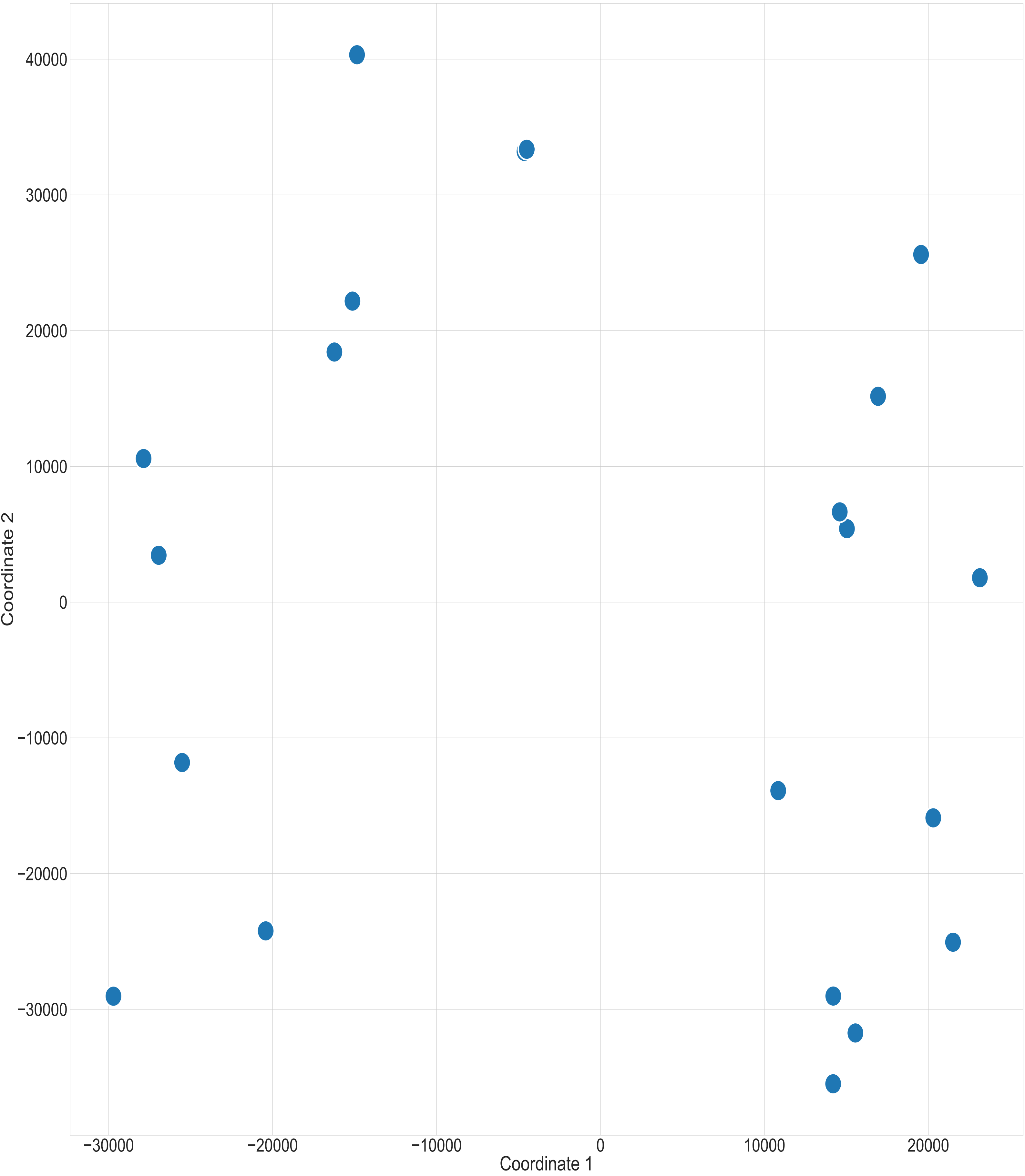
Scatter plot of Terpene *Micromonosporaceae* BGCs analyzed via multidimensional scaling using canberra pairwise distance.
Only Terpene BGCs were considered for this analysis. Each dot represents an individual BGC.
Distance between Terpene BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S28 :



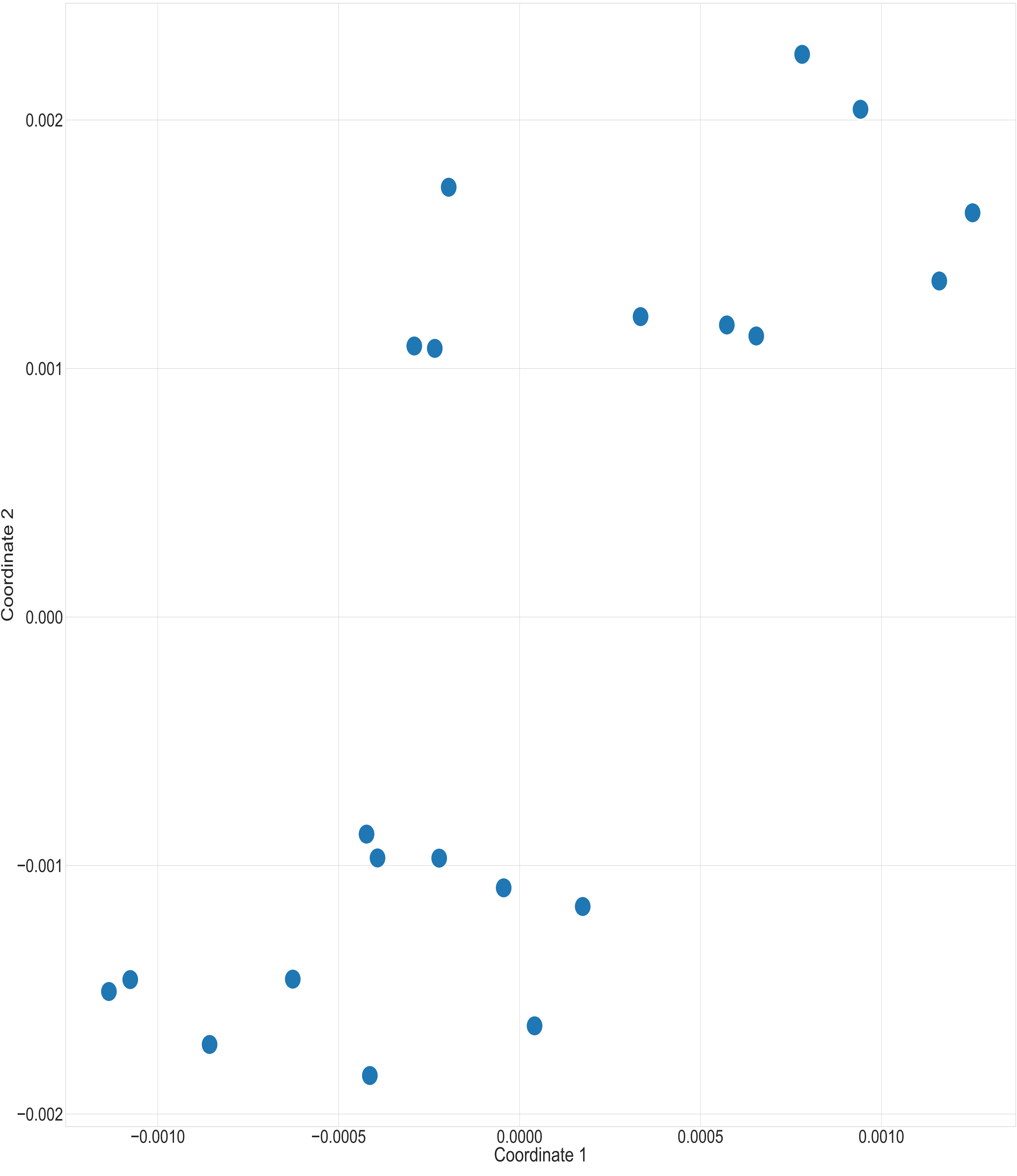
Scatter plot of Terpene *Micromonosporaceae* BGCs analyzed via multidimensional scaling using correlation pairwise distance.
Only Terpene BGCs were considered for this analysis. Each dot represents an individual BGC.
Distance between Terpene BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S29 :



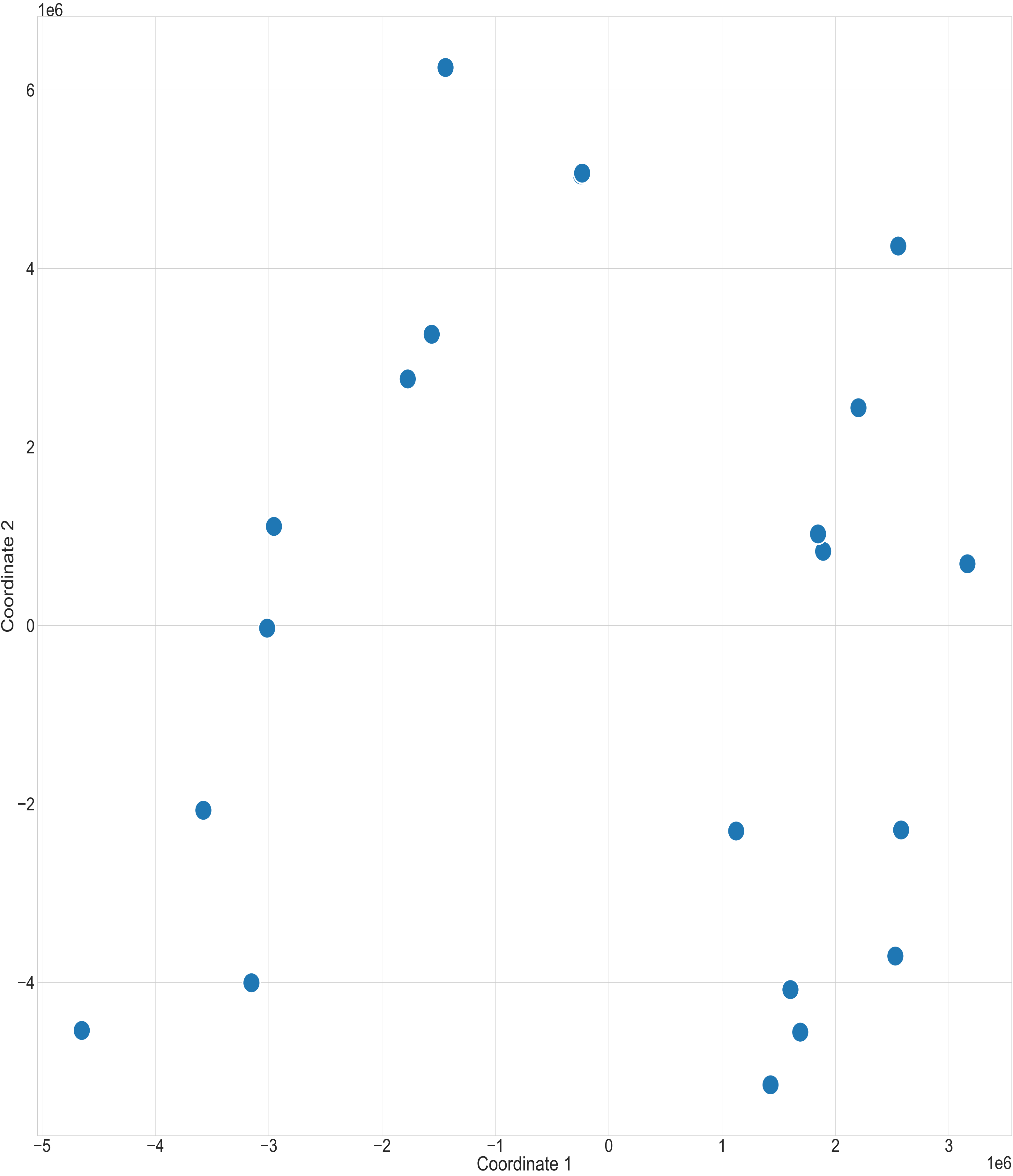
Scatter plot of PKS-1 *Micromonosporaceae* BGCs analyzed via multidimensional scaling using euclidean pairwise distance. Only PKS-1 BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between PKS-1 BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S30 :



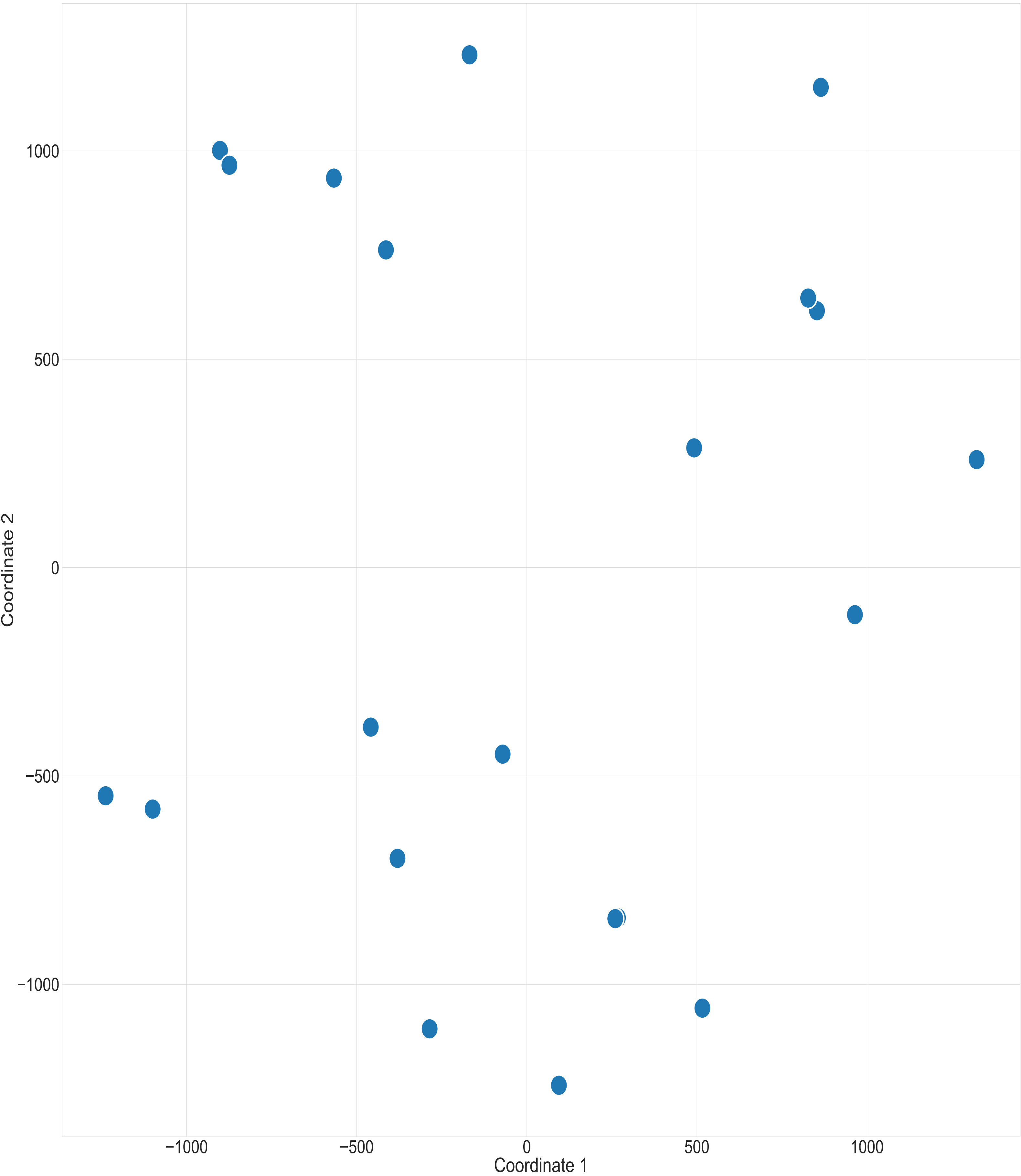
Scatter plot of PKS-1 *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cosine pairwise distance.
Only PKS-1 BGCs were considered for this analysis. Each dot represents an individual BGC.
Distance between PKS-1 BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S31 :



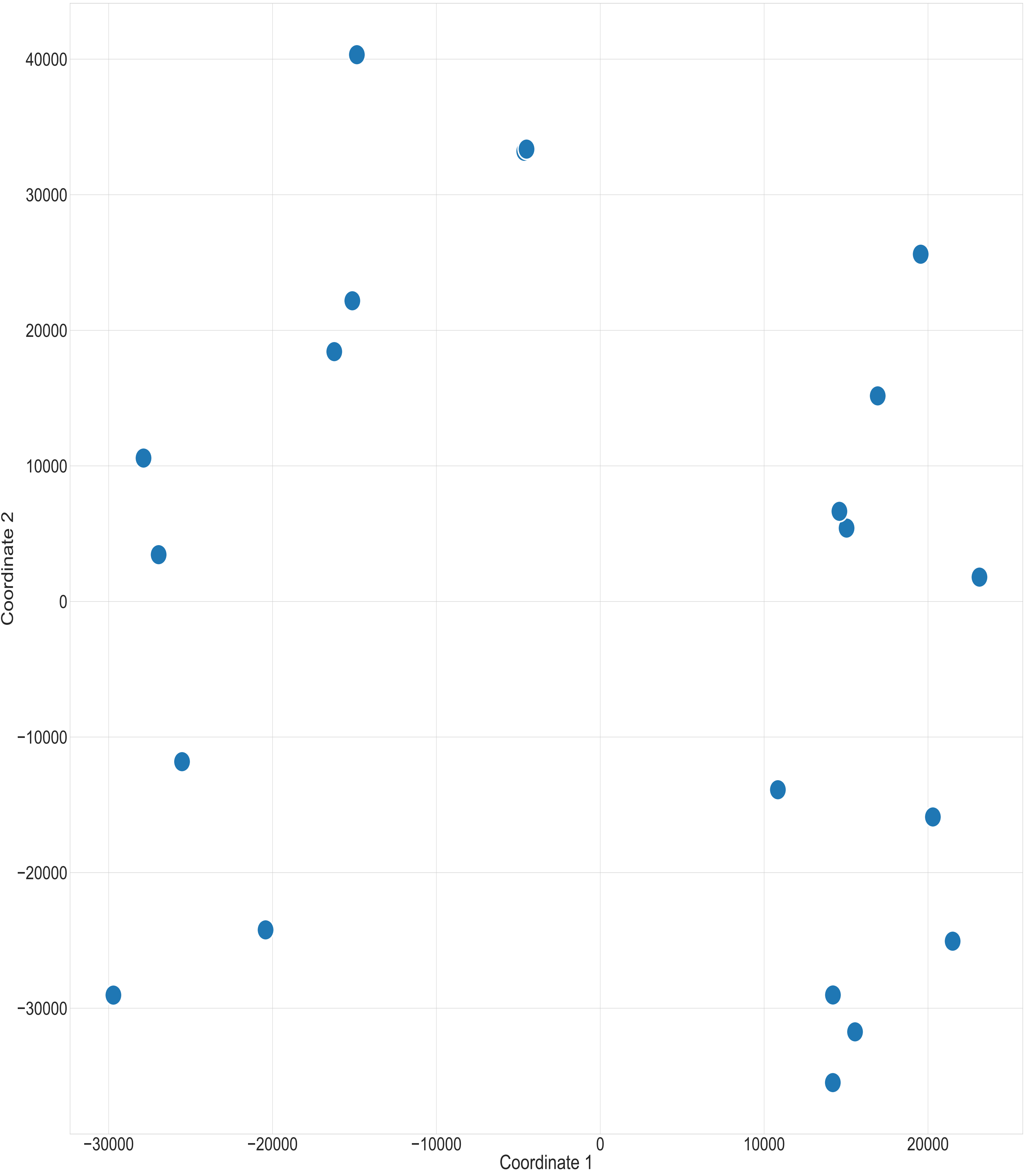
Scatter plot of PKS-1 *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cityblock pairwise distance. Only PKS-1 BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between PKS-1 BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S32 :



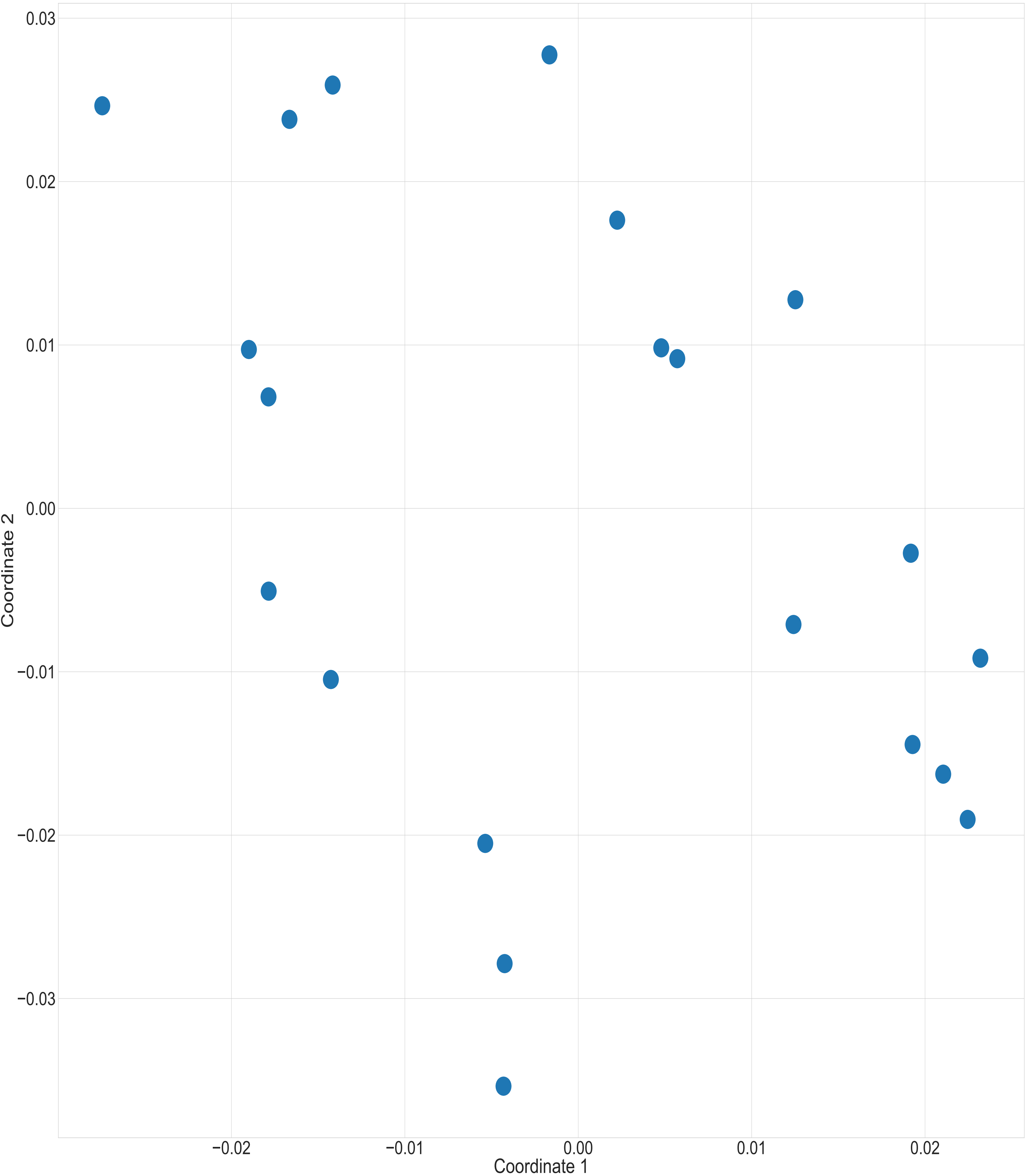
Scatter plot of PKS-1 *Micromonosporaceae* BGCs analyzed via multidimensional scaling using chebyshev pairwise distance. Only PKS-1 BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between PKS-1 BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S33 :



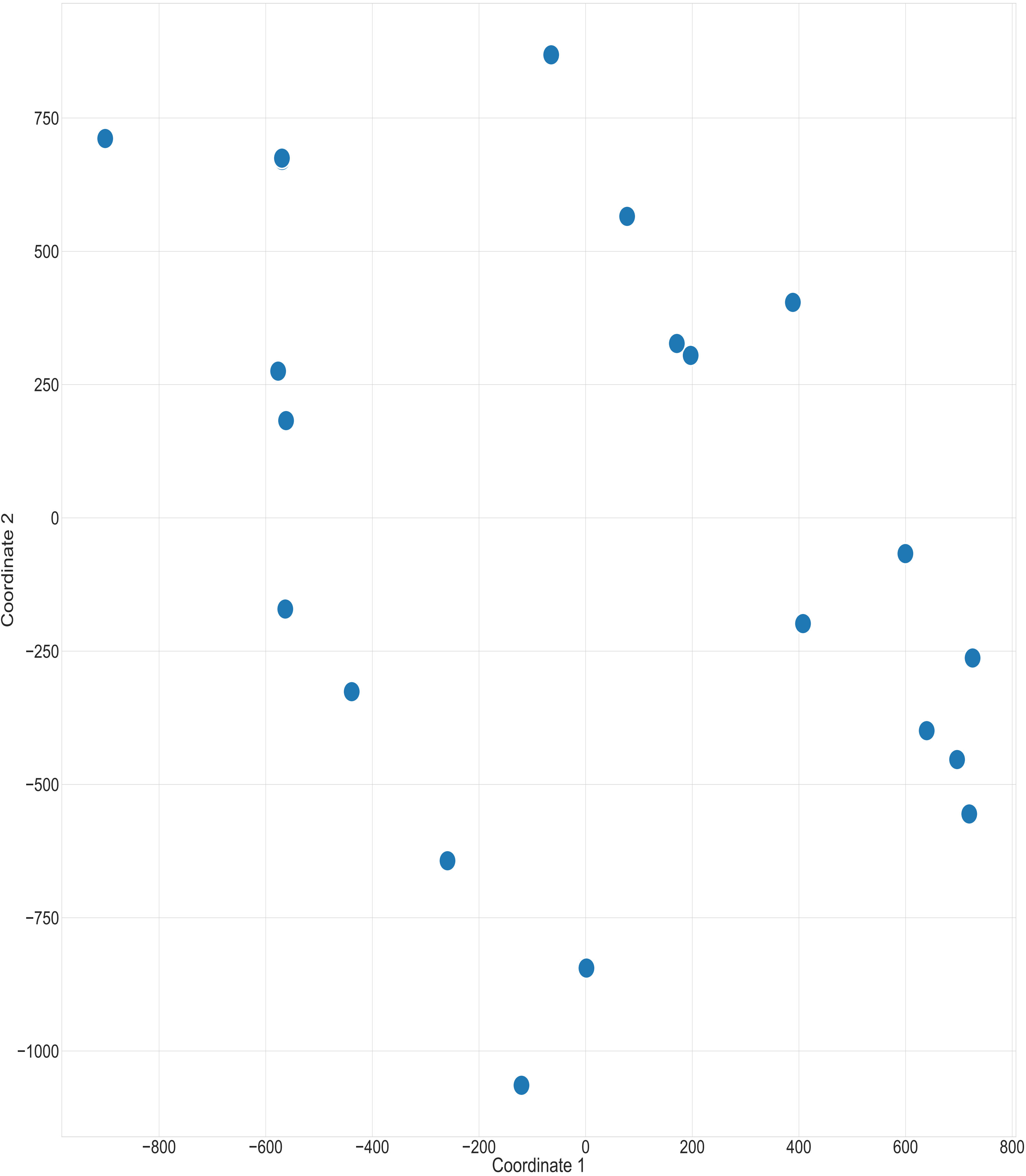
Scatter plot of PKS-1 *Micromonosporaceae* BGCs analyzed via multidimensional scaling using l2 pairwise distance.
Only PKS-1 BGCs were considered for this analysis. Each dot represents an individual BGC.
Distance between PKS-1 BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S34 :



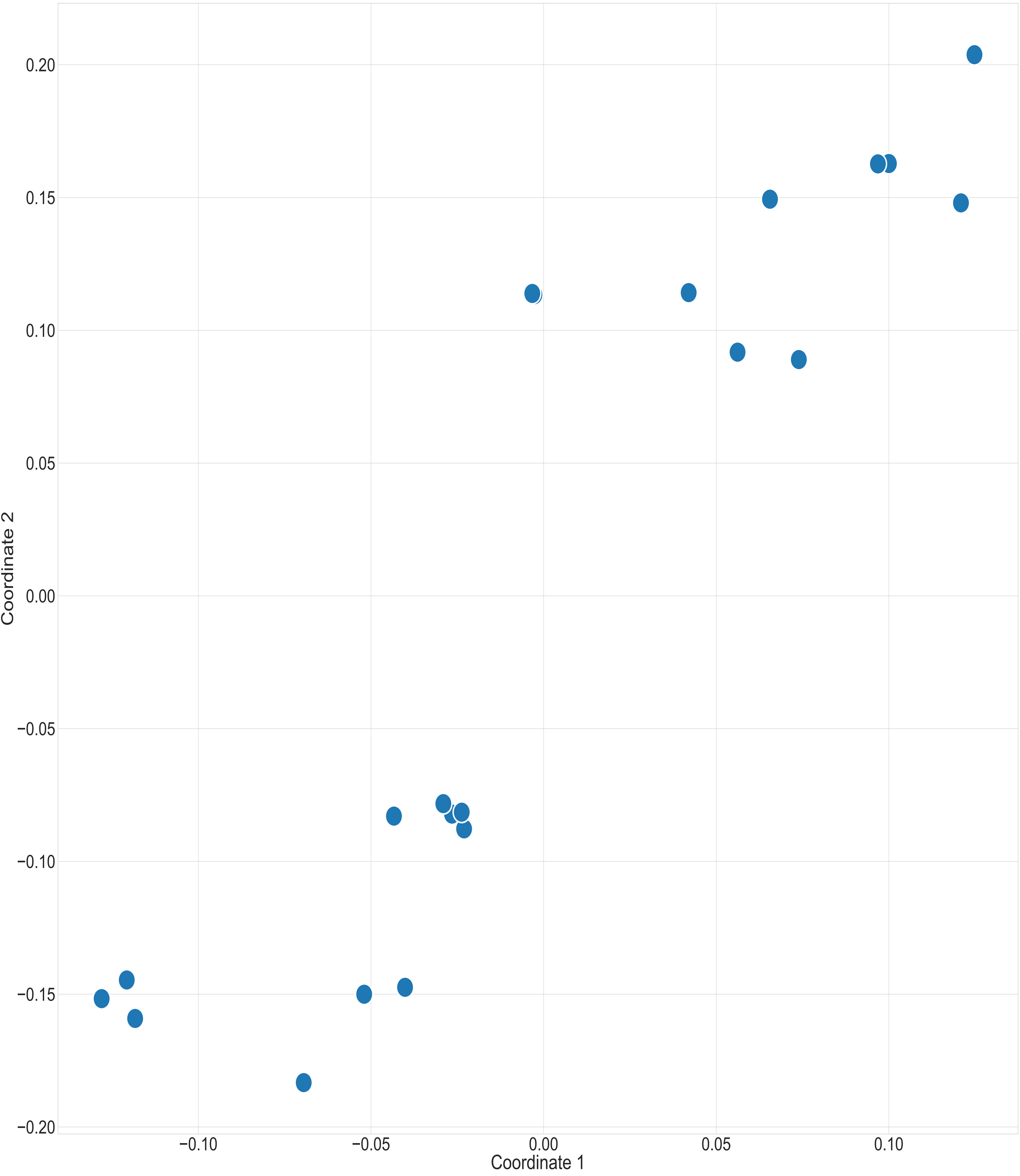
Scatter plot of PKS-1 *Micromonosporaceae* BGCs analyzed via multidimensional scaling using braycurtis pairwise distance. Only PKS-1 BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between PKS-1 BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S35 :



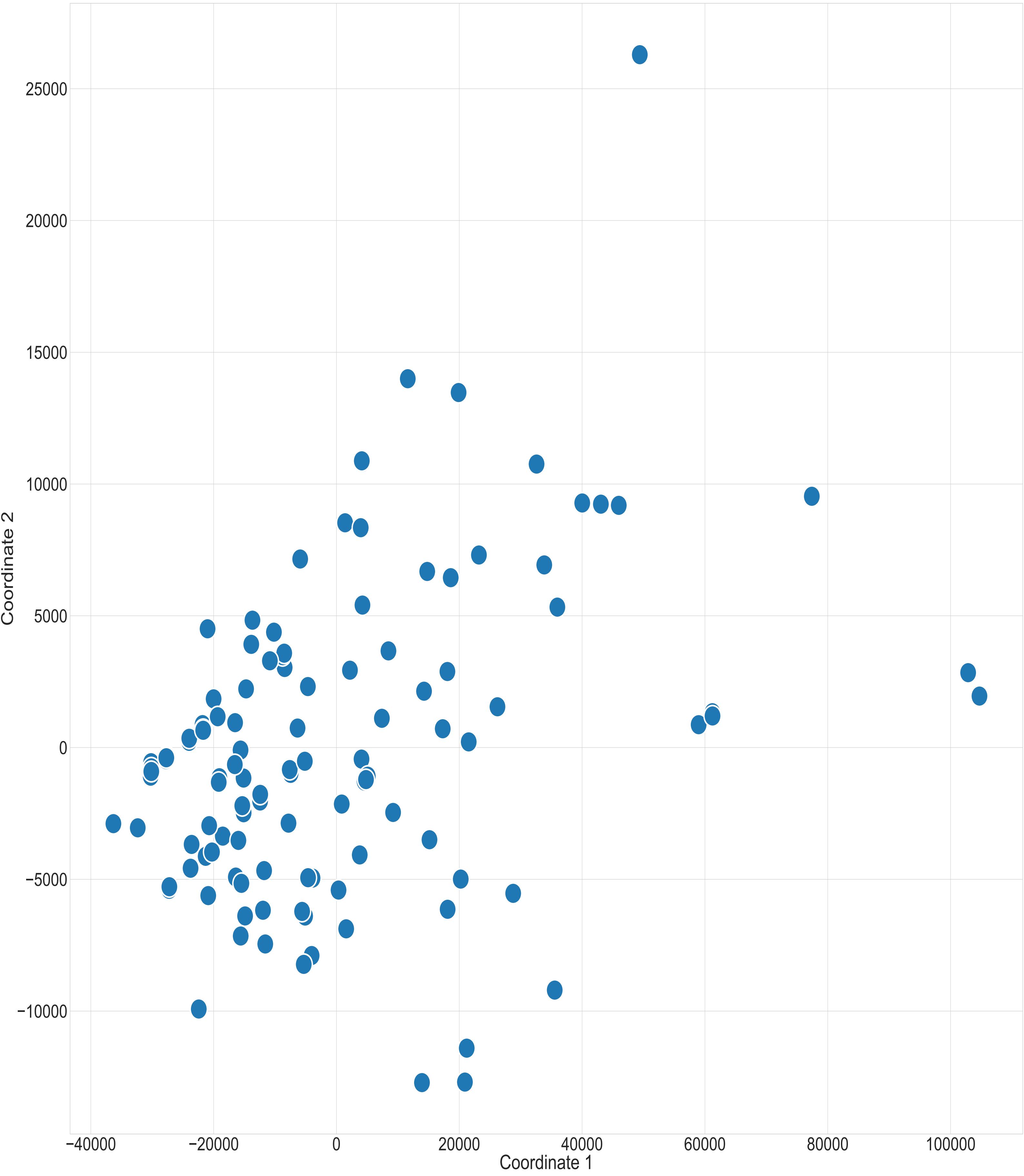
Scatter plot of PKS-1 *Micromonosporaceae* BGCs analyzed via multidimensional scaling using canberra pairwise distance. Only PKS-1 BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between PKS-1 BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S36 :



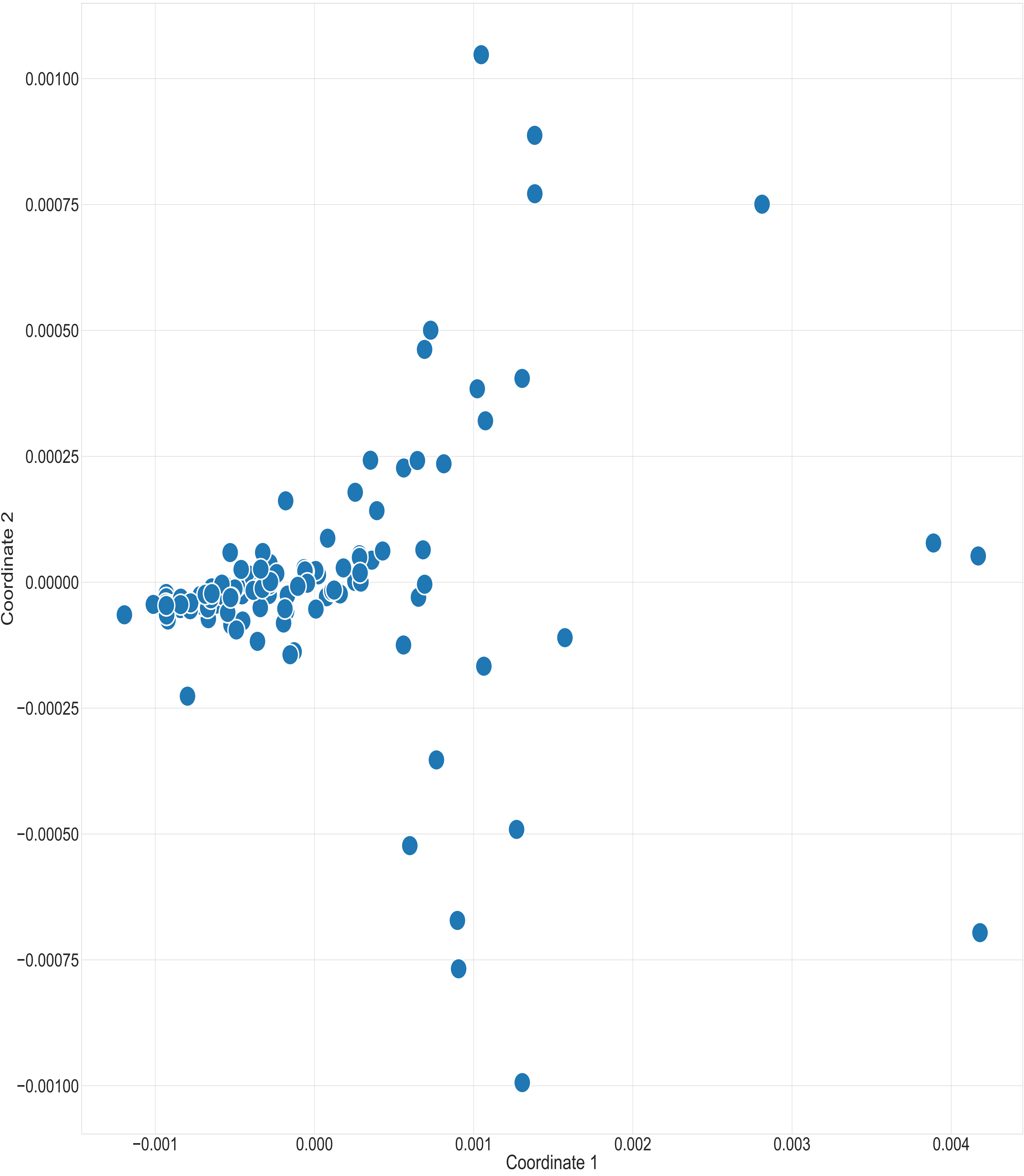
Scatter plot of PKS-1 *Micromonosporaceae* BGCs analyzed via multidimensional scaling using correlation pairwise distance. Only PKS-1 BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between PKS-1 BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S37 :



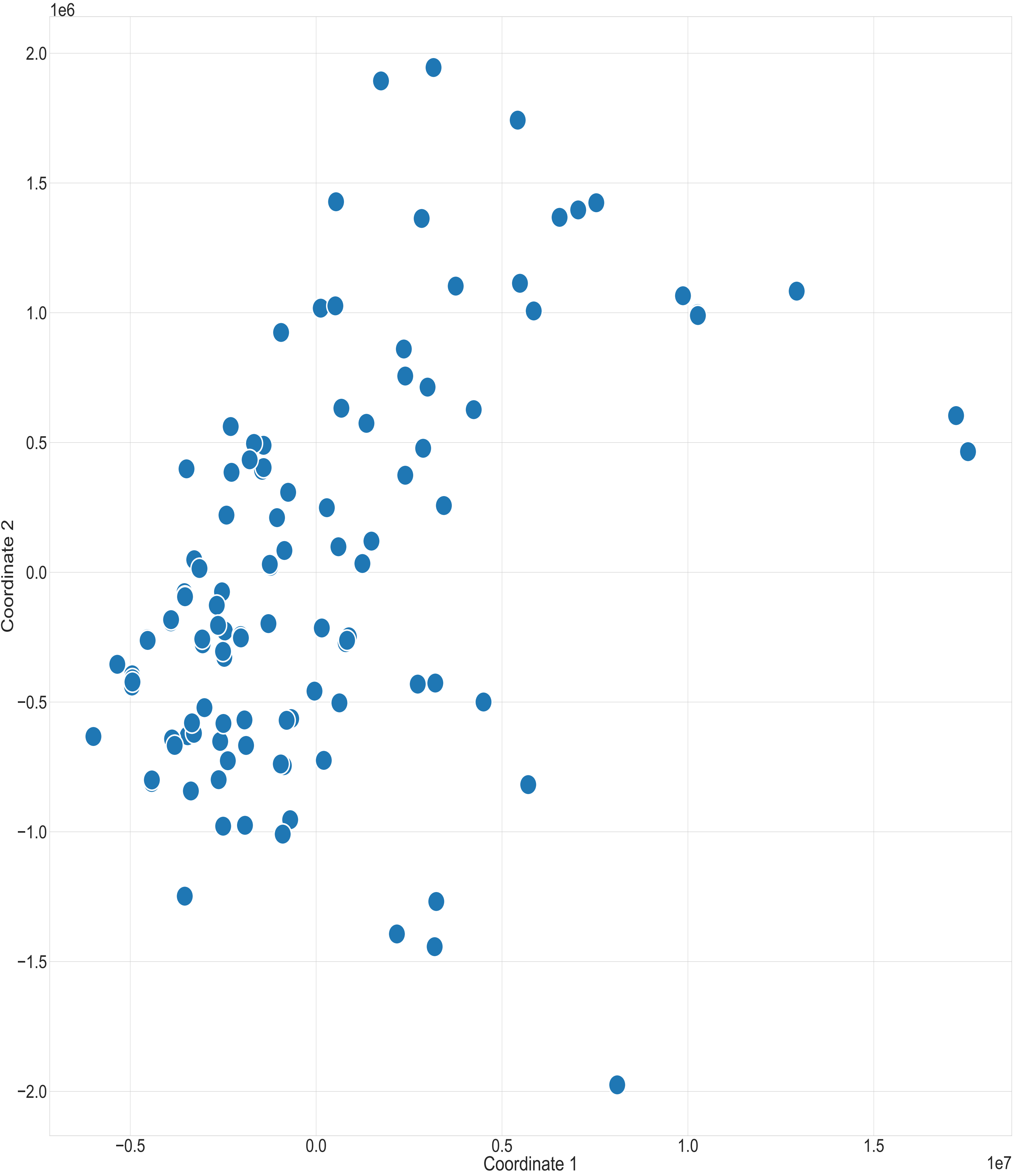
Scatter plot of RIPP *Micromonosporaceae* BGCs analyzed via multidimensional scaling using euclidean pairwise distance.
Only RIPP BGCs were considered for this analysis. Each dot represents an individual BGC.
Distance between RIPP BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S38 :



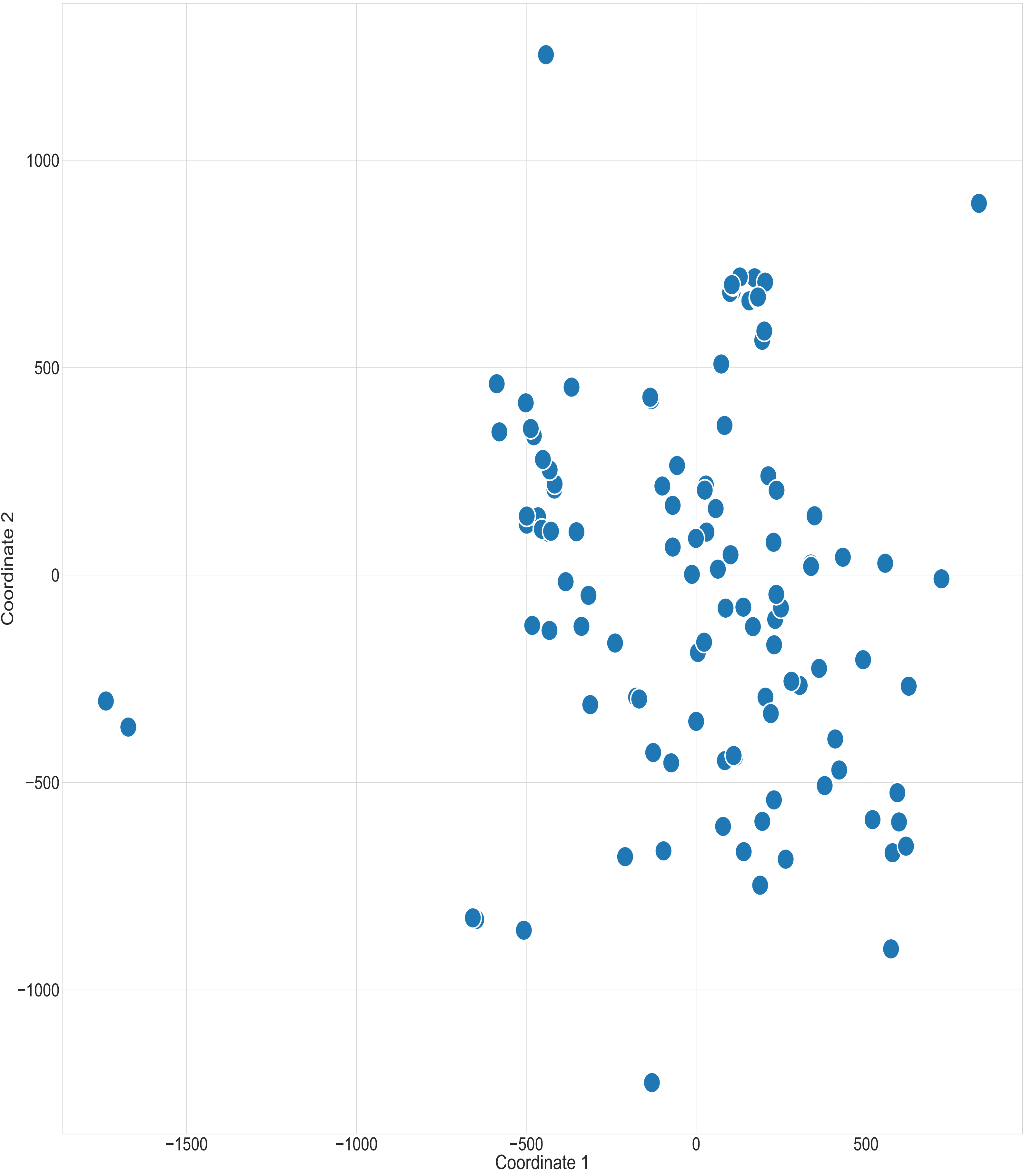
Scatter plot of RIPP *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cosine pairwise distance. Only RIPP BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between RIPP BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S39 :



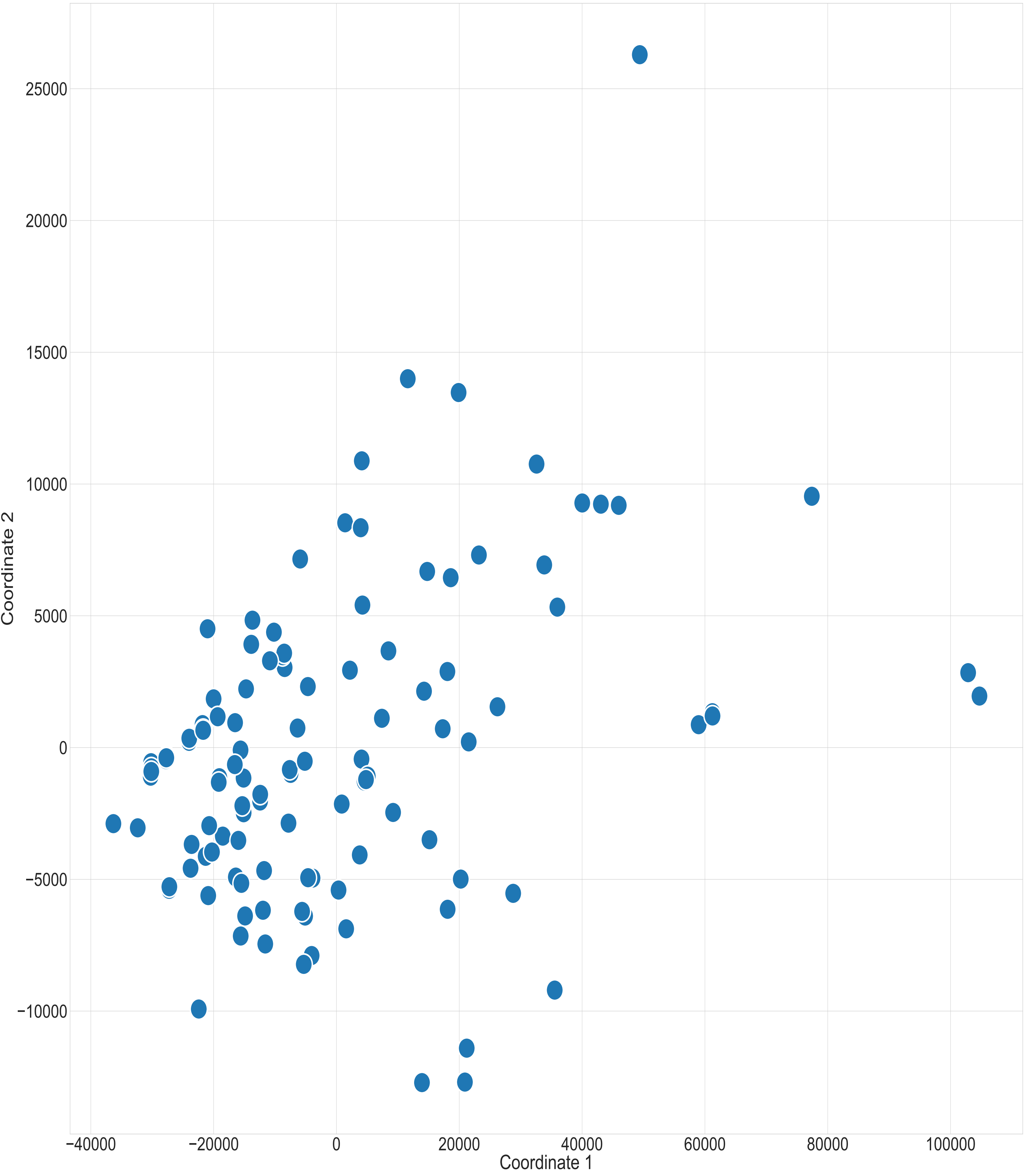
Scatter plot of RIPP *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cityblock pairwise distance.
Only RIPP BGCs were considered for this analysis. Each dot represents an individual BGC.
Distance between RIPP BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S40 :



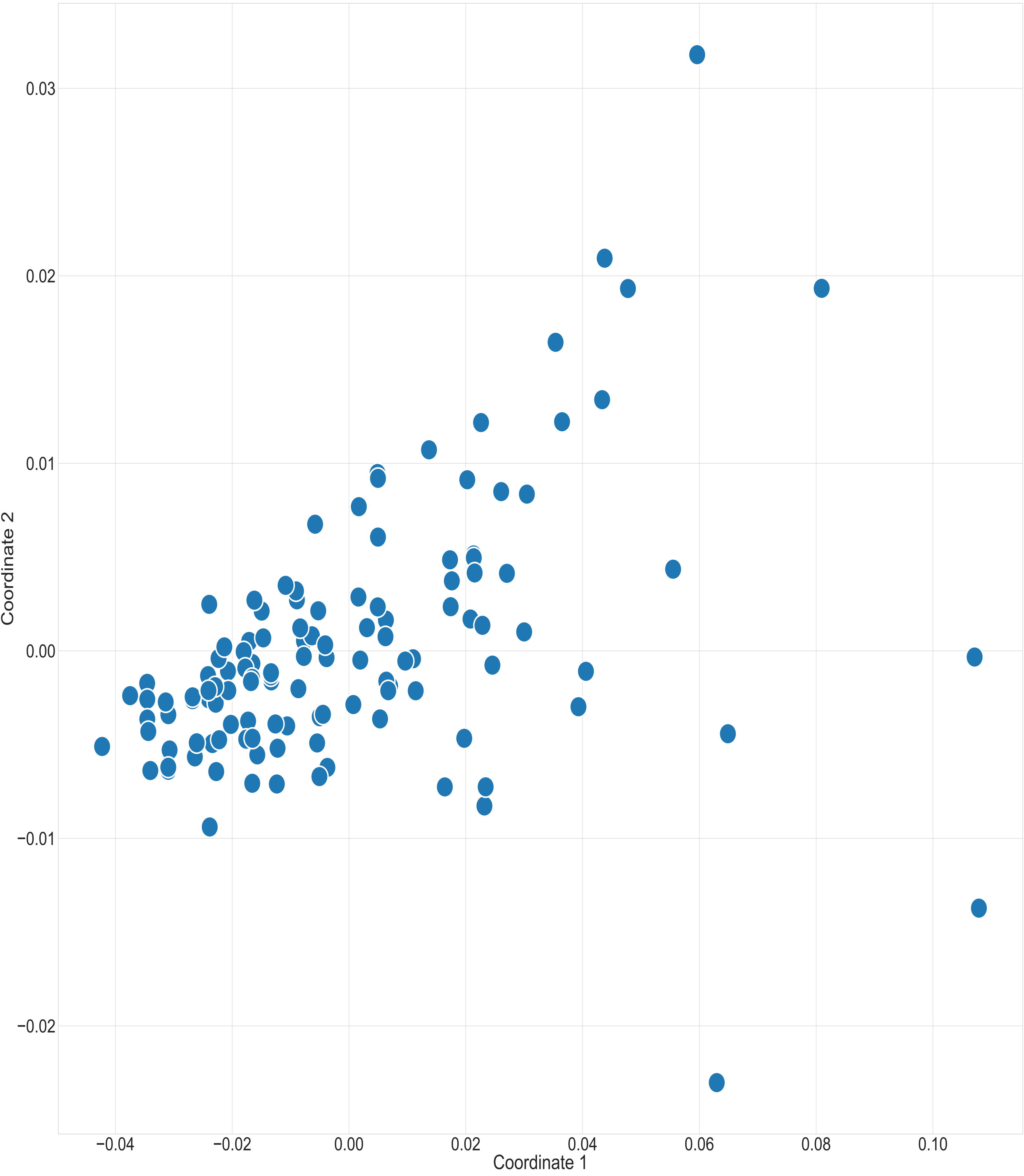
Scatter plot of RIPP *Micromonosporaceae* BGCs analyzed via multidimensional scaling using chebyshev pairwise distance.
Only RIPP BGCs were considered for this analysis. Each dot represents an individual BGC.
Distance between RIPP BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S41 :



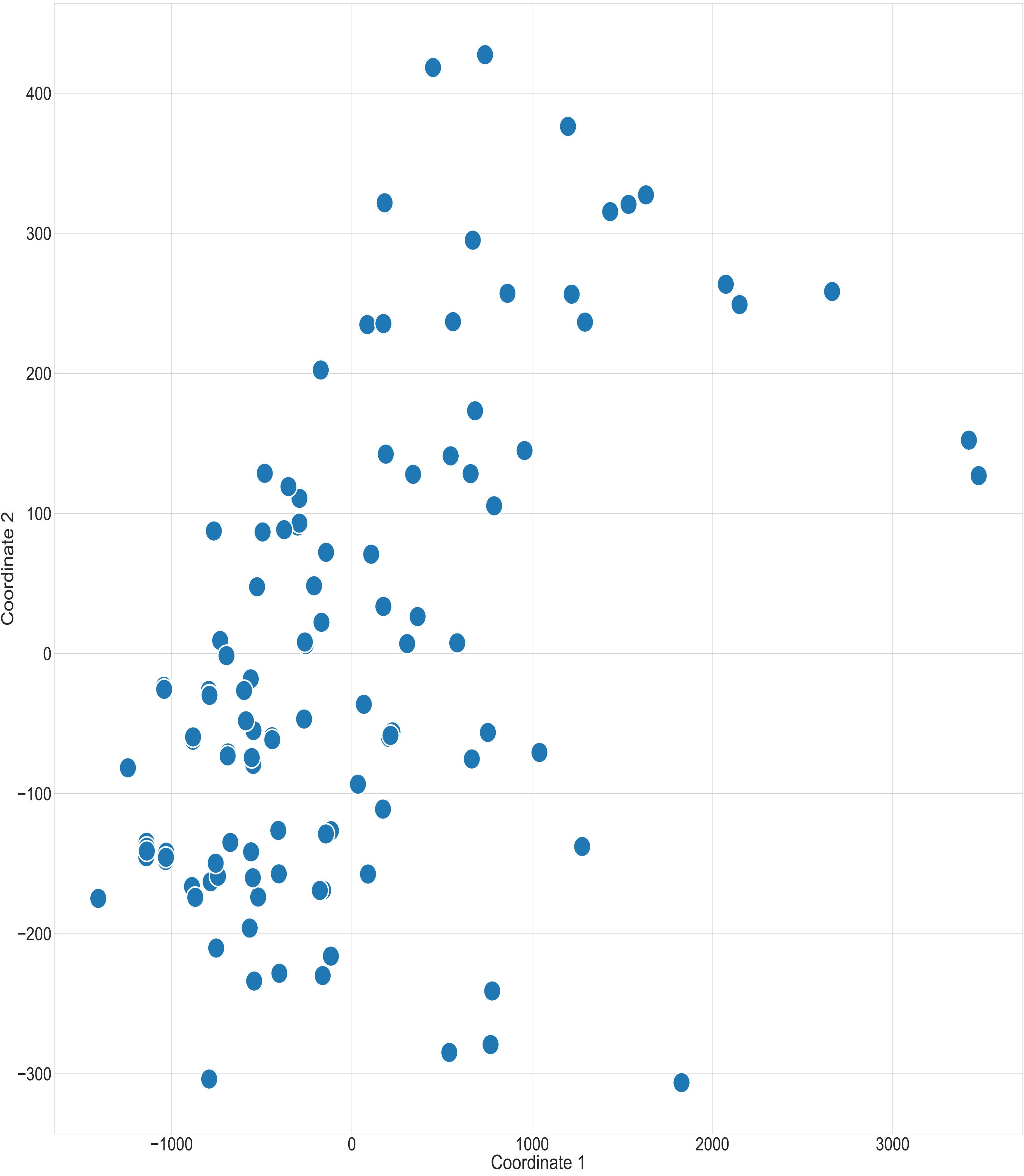
Scatter plot of RIPP *Micromonosporaceae* BGCs analyzed via multidimensional scaling using l2 pairwise distance. Only RIPP BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between RIPP BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S42 :



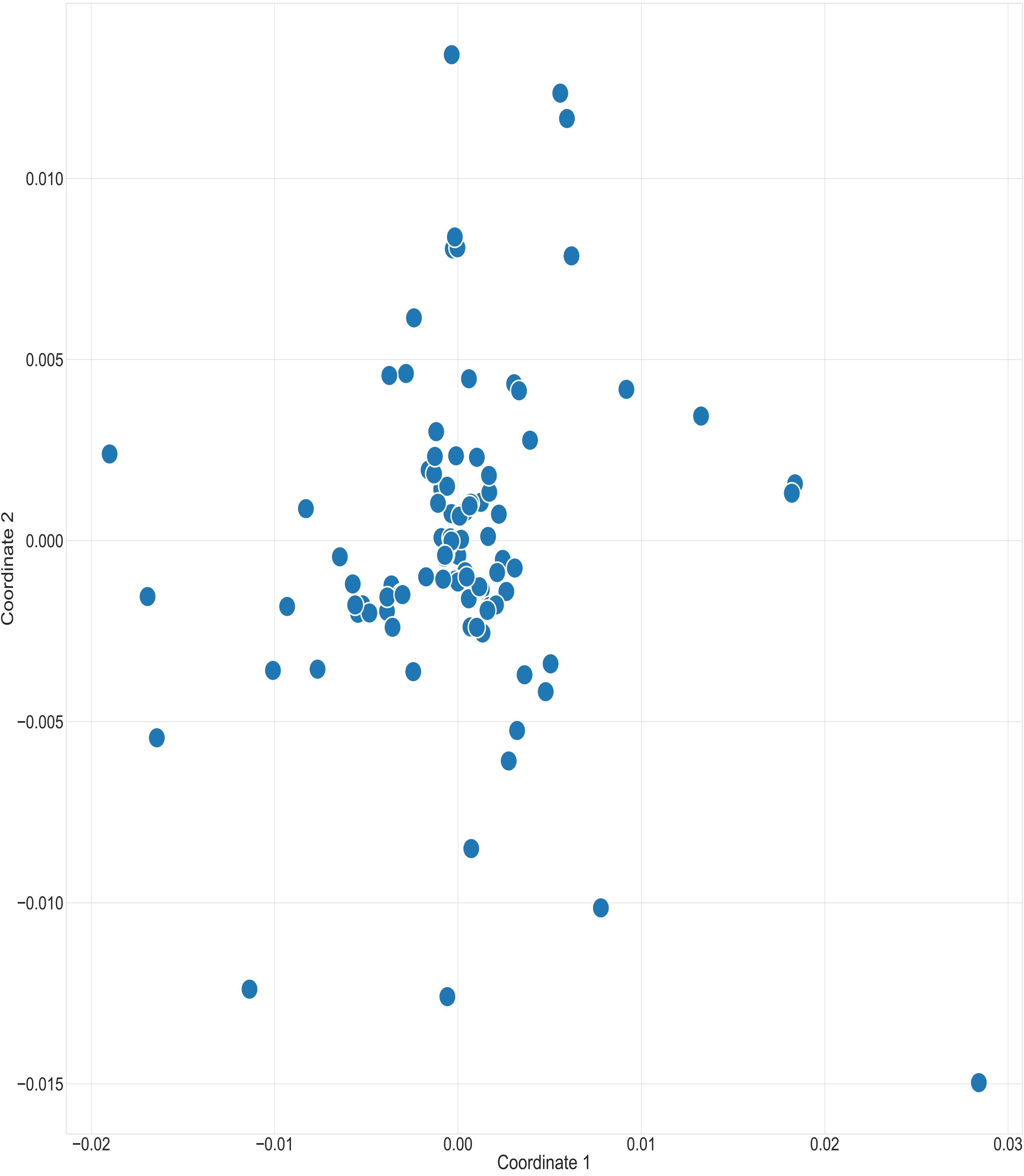
Scatter plot of RIPP *Micromonosporaceae* BGCs analyzed via multidimensional scaling using braycurtis pairwise distance. Only RIPP BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between RIPP BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S43 :



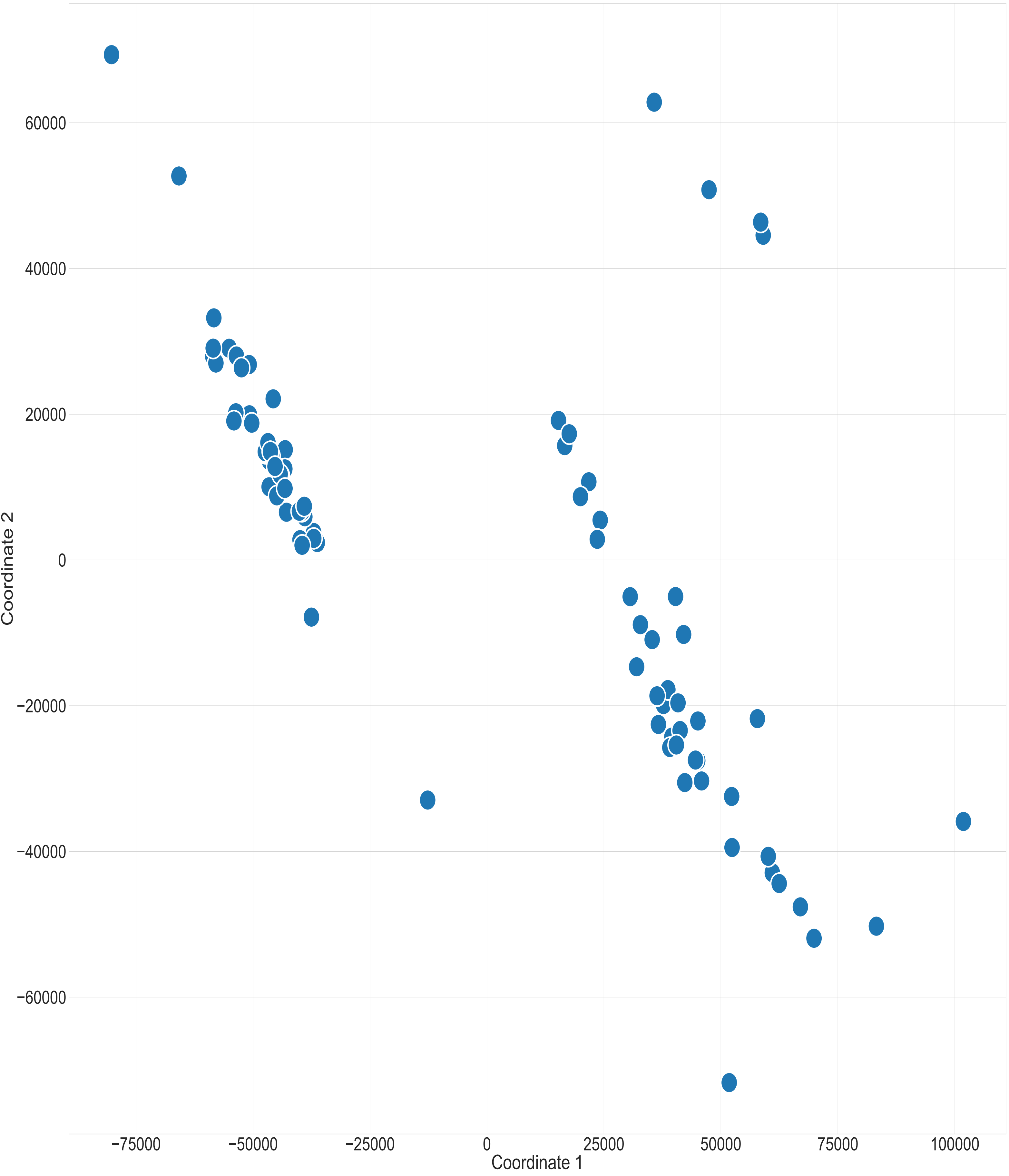
Scatter plot of RIPP *Micromonosporaceae* BGCs analyzed via multidimensional scaling using canberra pairwise distance. Only RIPP BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between RIPP BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S44 :



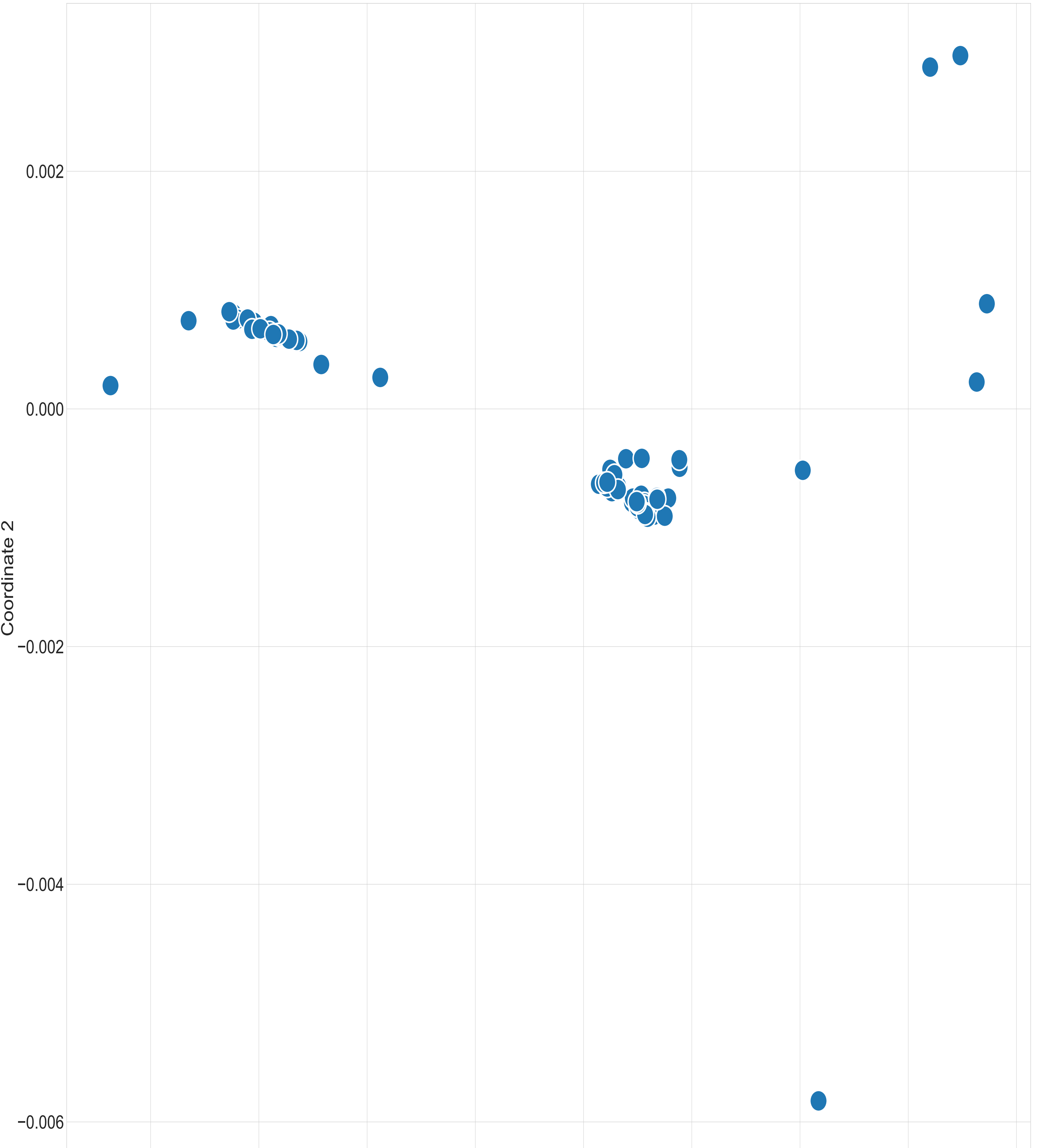
Scatter plot of RIPP *Micromonosporaceae* BGCs analyzed via multidimensional scaling using correlation pairwise distance.
Only RIPP BGCs were considered for this analysis. Each dot represents an individual BGC.
Distance between RIPP BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S45 :



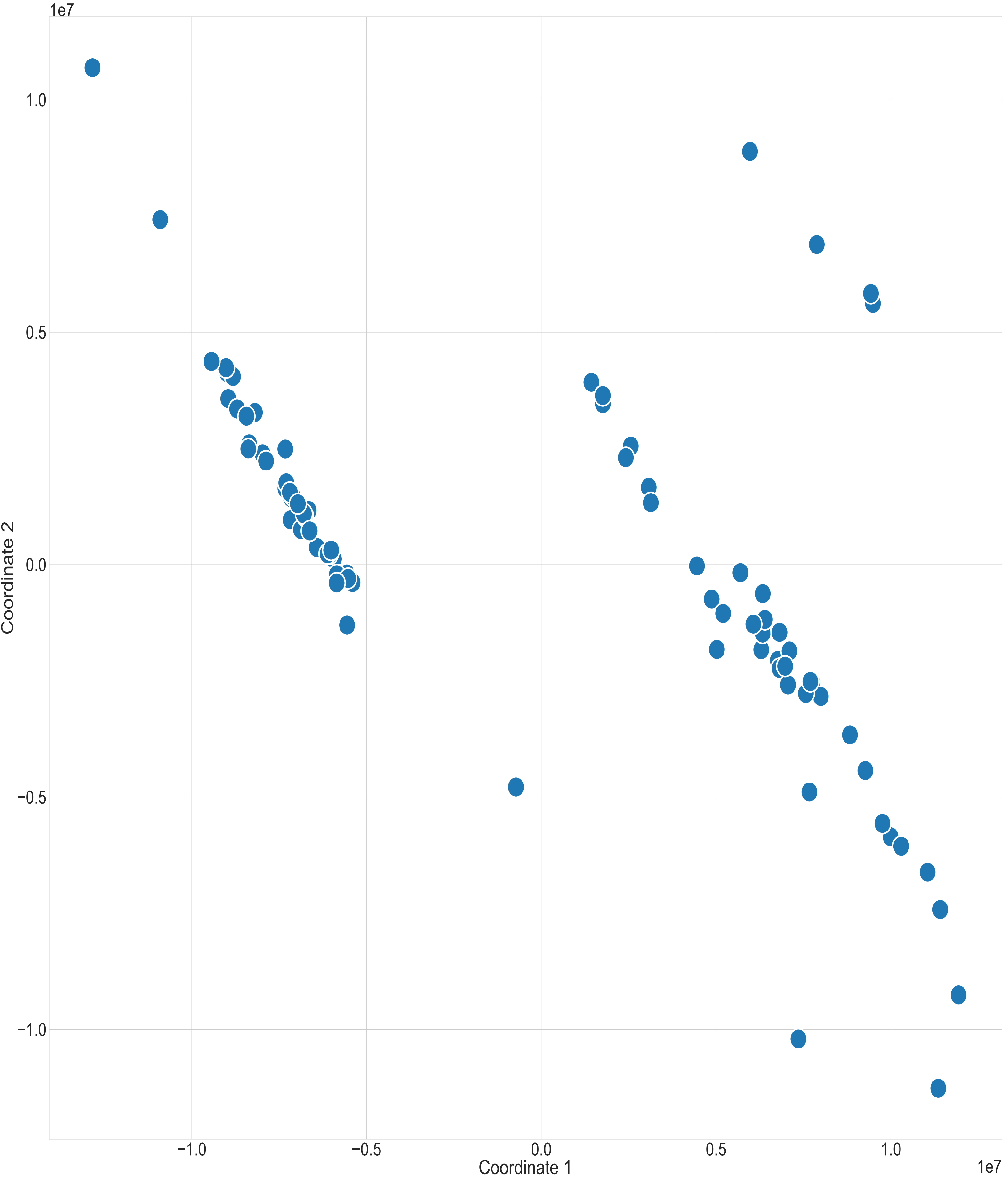
Scatter plot of PKS-Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using euclidean pairwise distance.
Only PKS-Other BGCs were considered for this analysis. Each dot represents an individual BGC.
Distance between PKS-Other BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S46 :



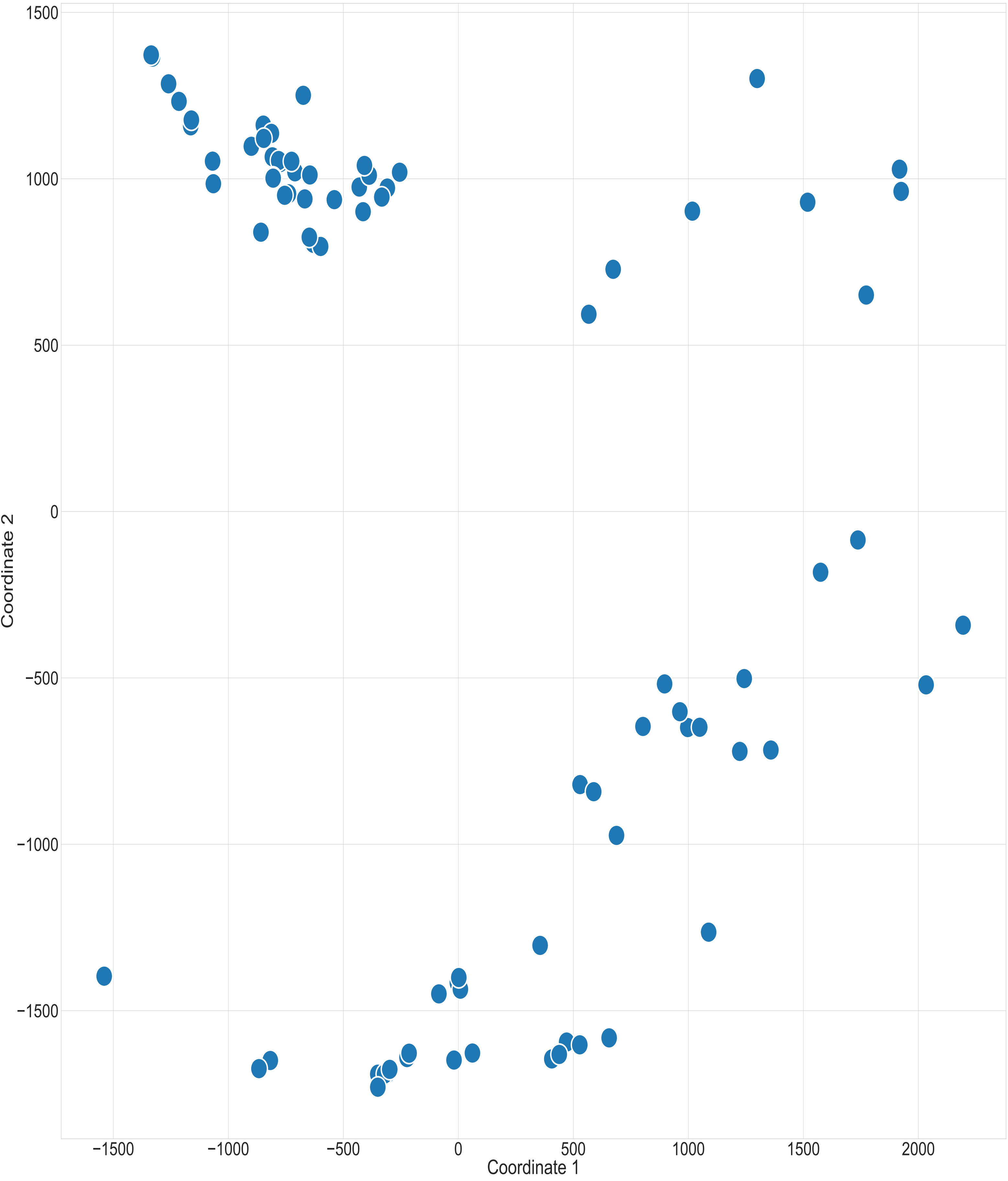
Scatter plot of PKS-Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cosine pairwise distance.
Only PKS-Other BGCs were considered for this analysis. Each dot represents an individual BGC.
Distance between PKS-Other BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S47 :



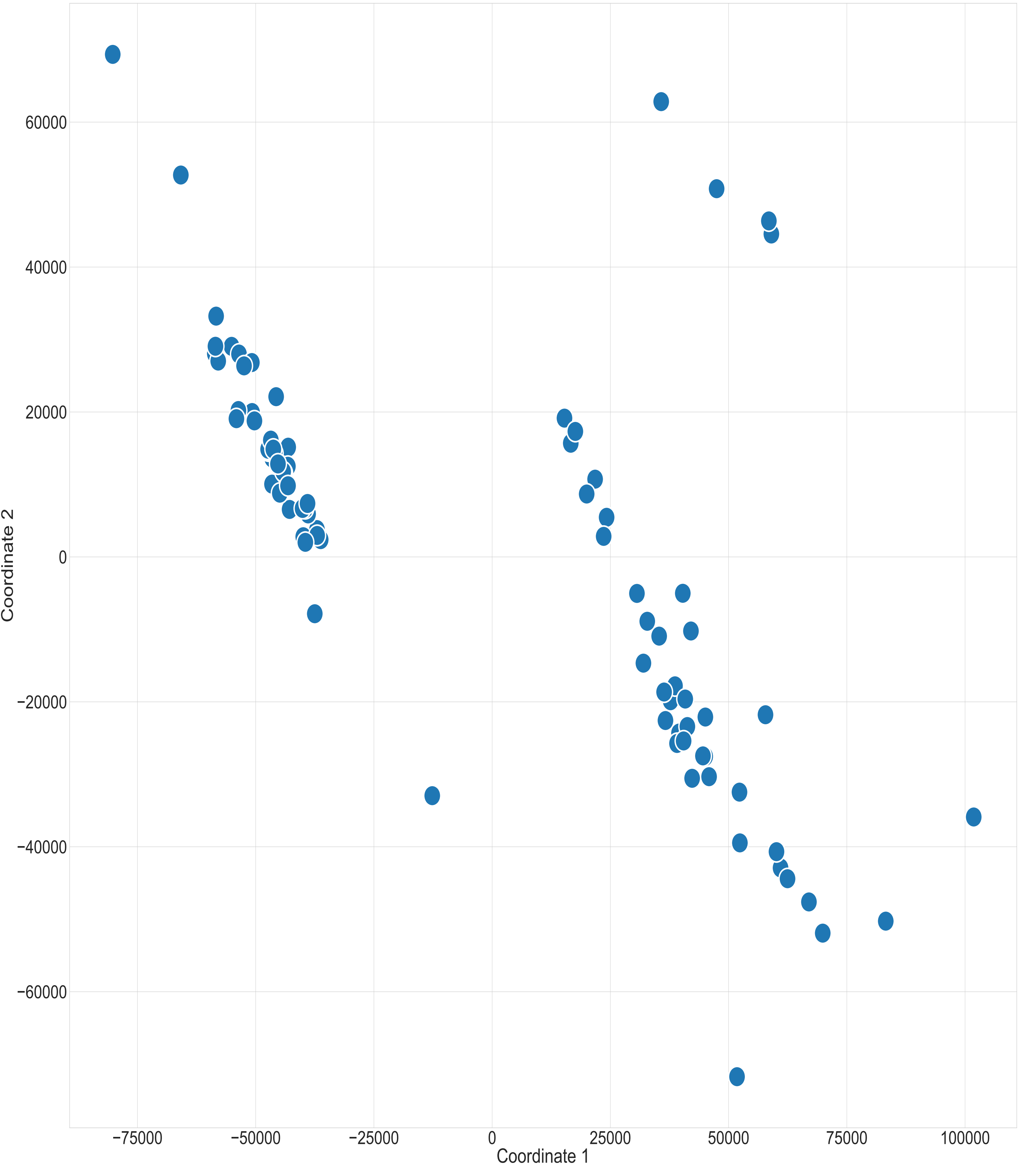
Scatter plot of PKS-Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cityblock pairwise distance.
Only PKS-Other BGCs were considered for this analysis. Each dot represents an individual BGC.
Distance between PKS-Other BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S48 :



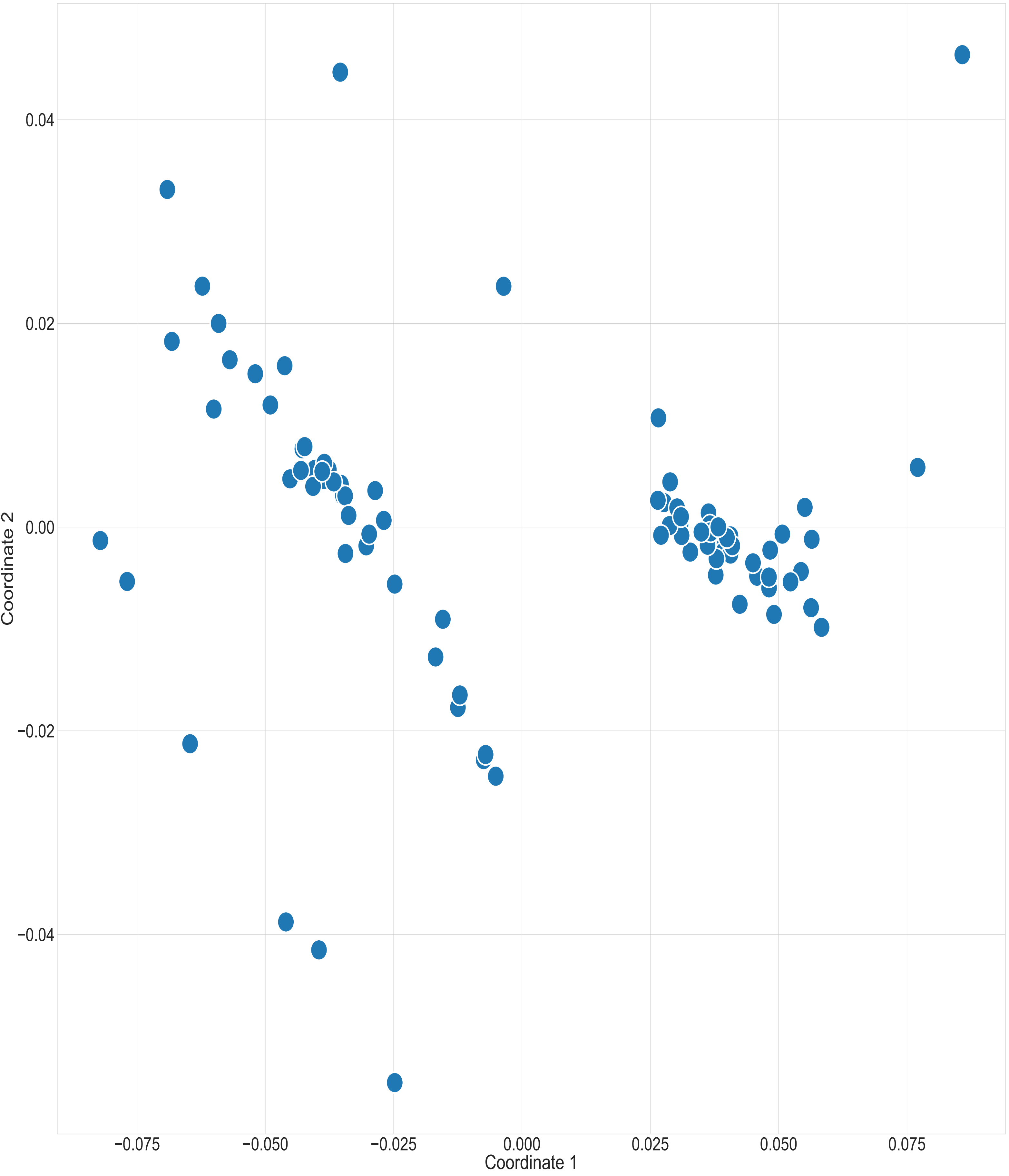
Scatter plot of PKS-Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using chebyshev pairwise distance. Only PKS-Other BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between PKS-Other BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S49 :



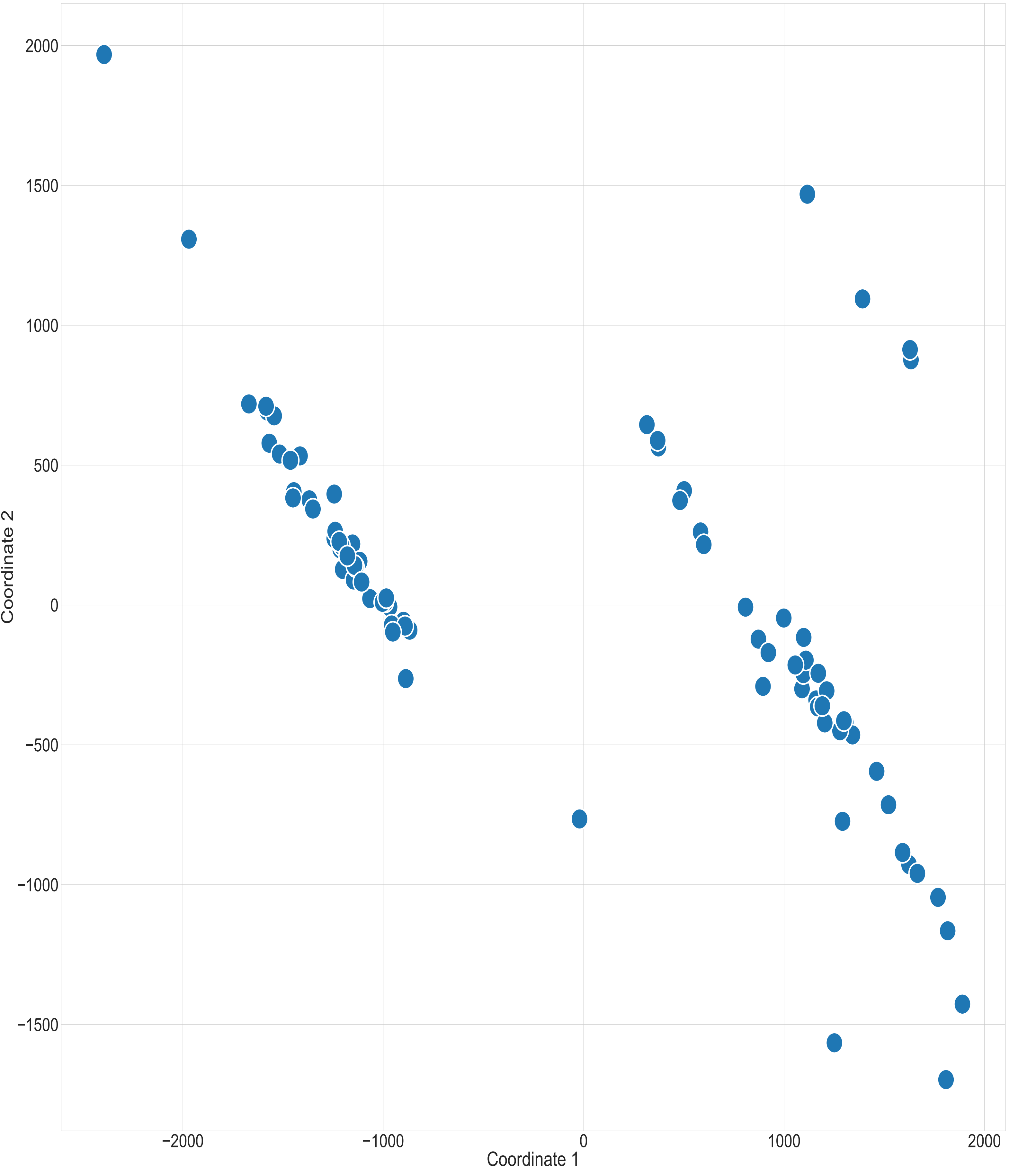
Scatter plot of PKS-Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using l2 pairwise distance.
Only PKS-Other BGCs were considered for this analysis. Each dot represents an individual BGC.
Distance between PKS-Other BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S50 :



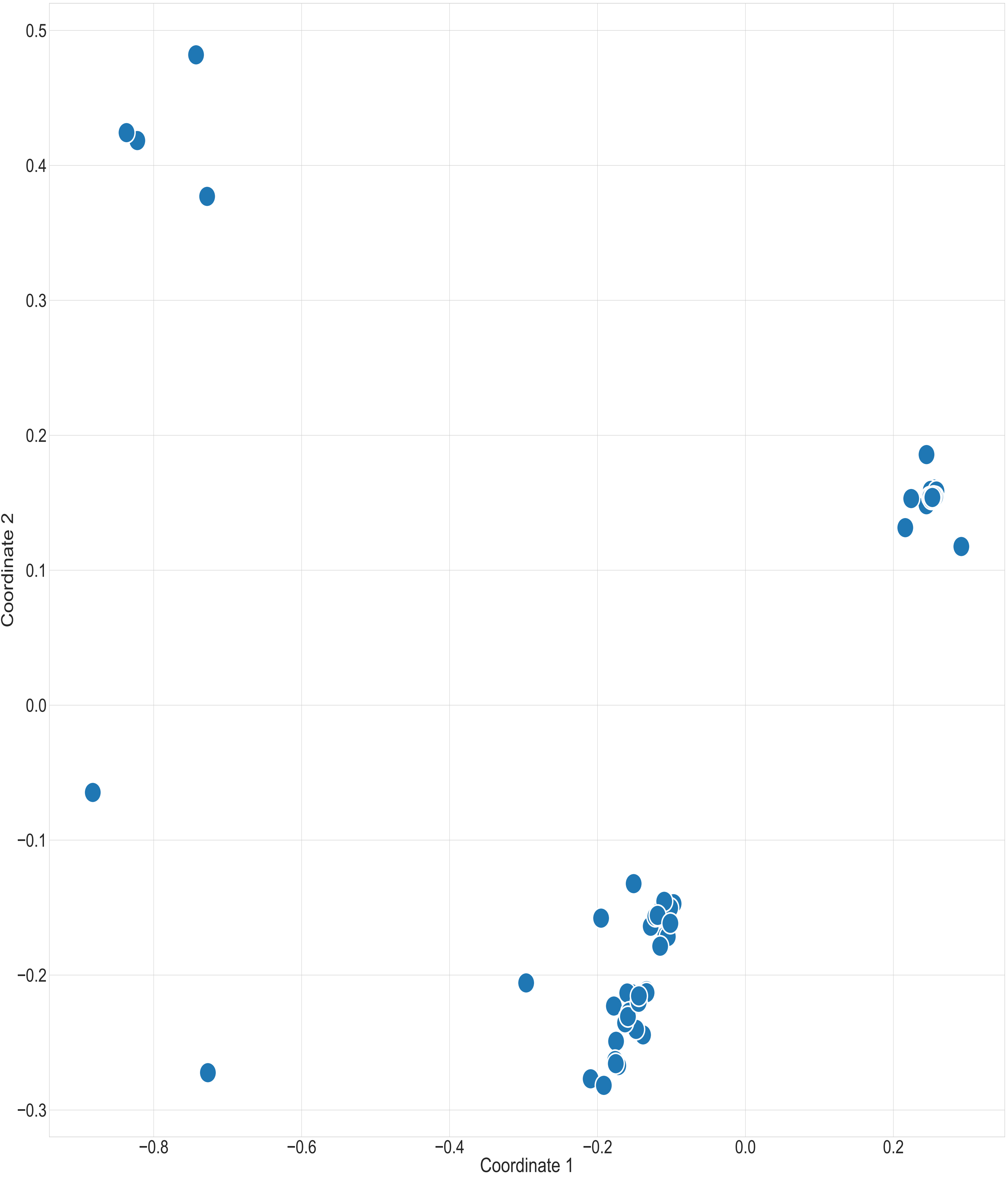
Scatter plot of PKS-Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using braycurtis pairwise distance.
Only PKS-Other BGCs were considered for this analysis. Each dot represents an individual BGC.
Distance between PKS-Other BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S51 :



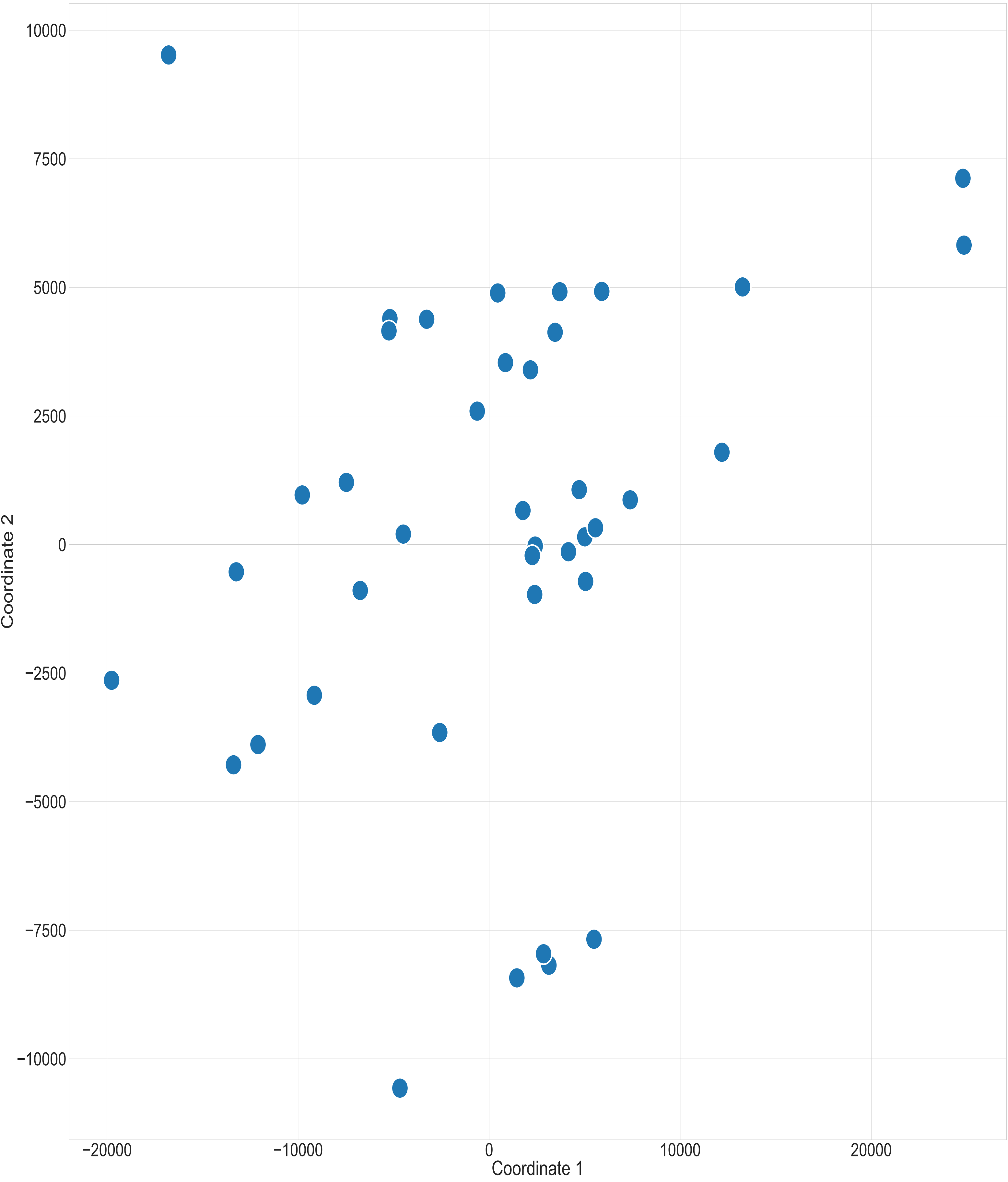
Scatter plot of PKS-Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using canberra pairwise distance.
Only PKS-Other BGCs were considered for this analysis. Each dot represents an individual BGC.
Distance between PKS-Other BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S52 :



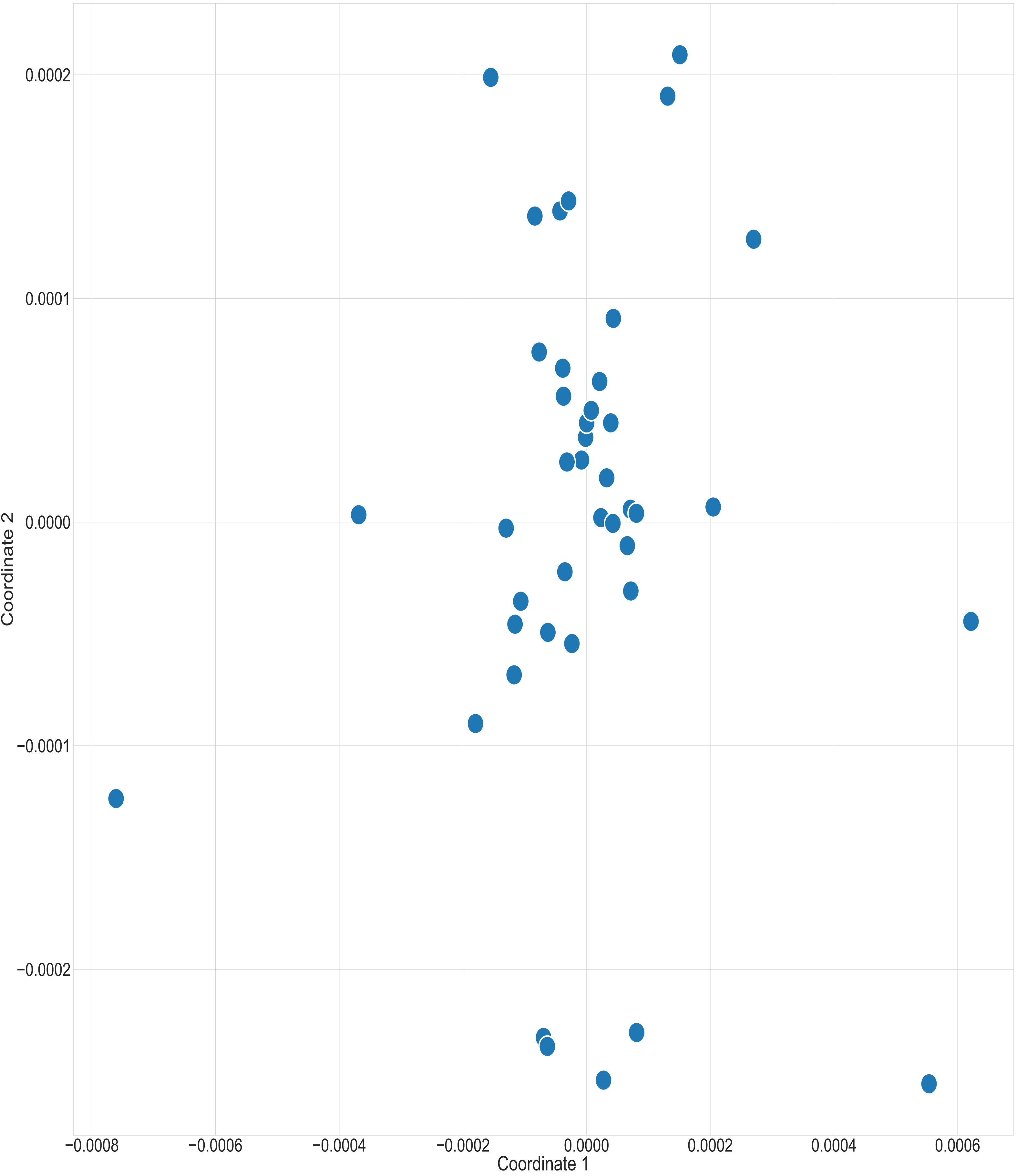
Scatter plot of PKS-Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using correlation pairwise distance. Only PKS-Other BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between PKS-Other BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S53 :



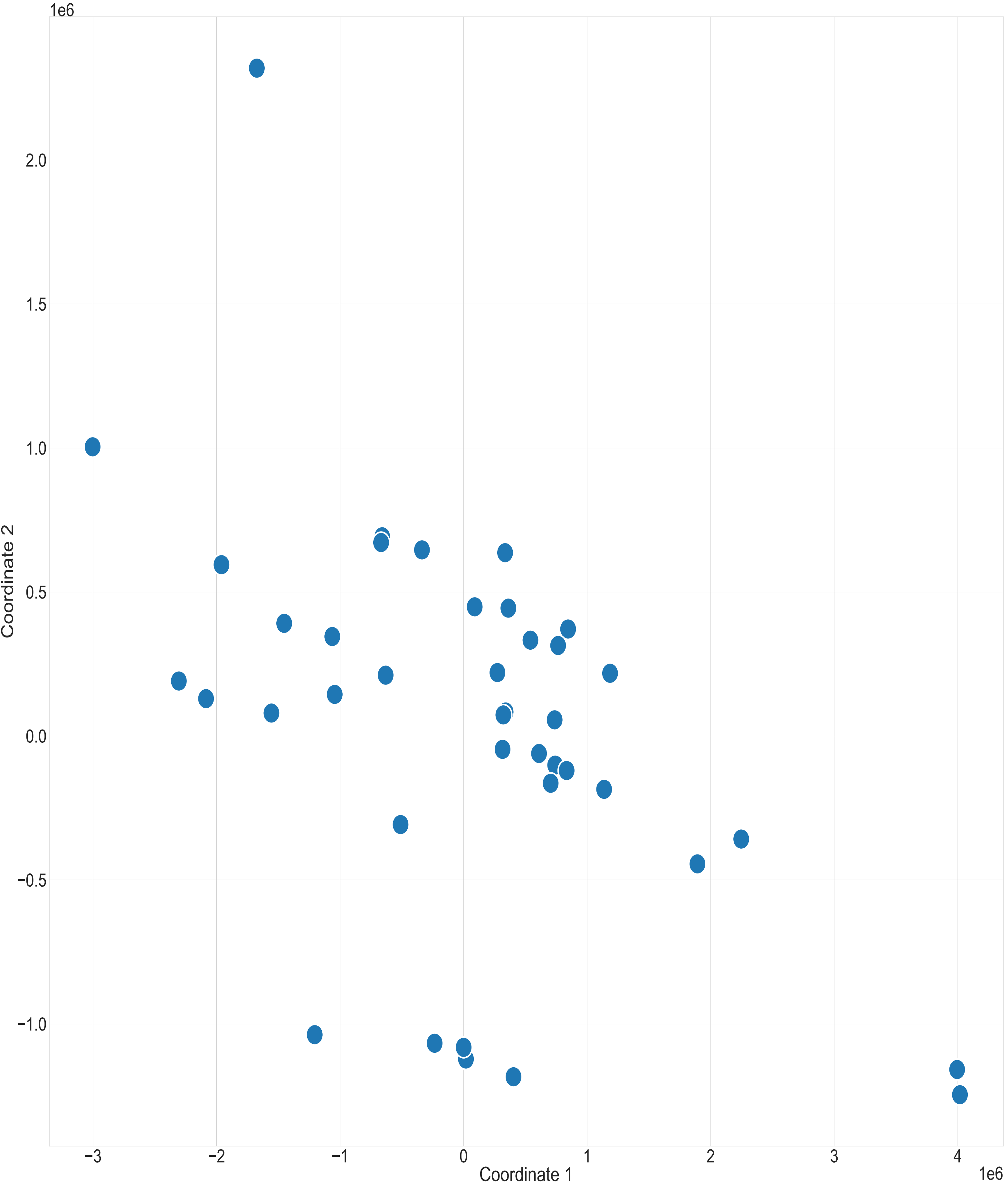
Scatter plot of Siderophore *Micromonosporaceae* BGCs analyzed via multidimensional scaling using euclidean pairwise distance.
Only Siderophore BGCs were considered for this analysis. Each dot represents an individual BGC.
Distance between Siderophore BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S54 :



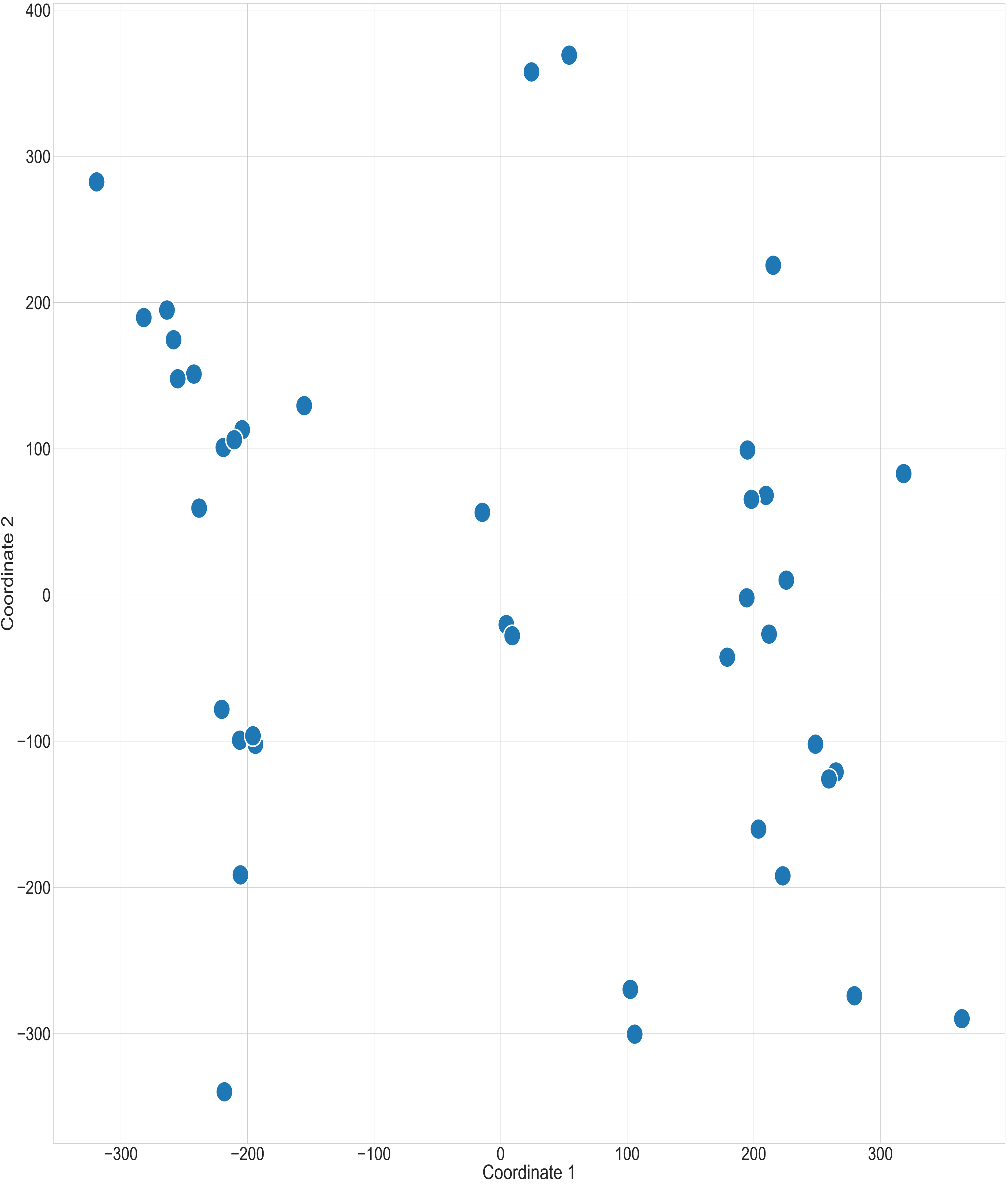
Scatter plot of Siderophore *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cosine pairwise distance. Only Siderophore BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Siderophore BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S55 :



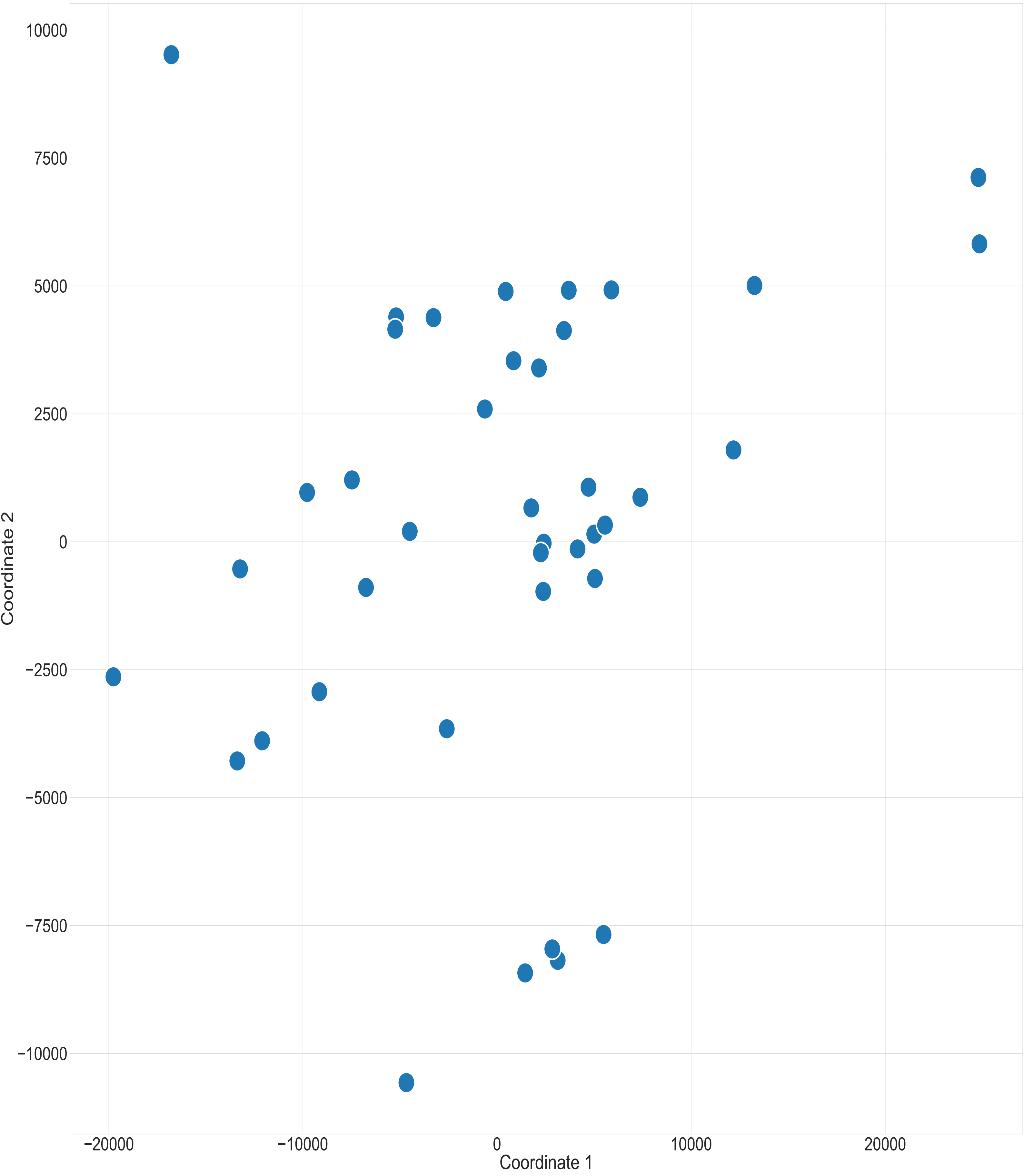
Scatter plot of Siderophore *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cityblock pairwise distance. Only Siderophore BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Siderophore BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S56 :



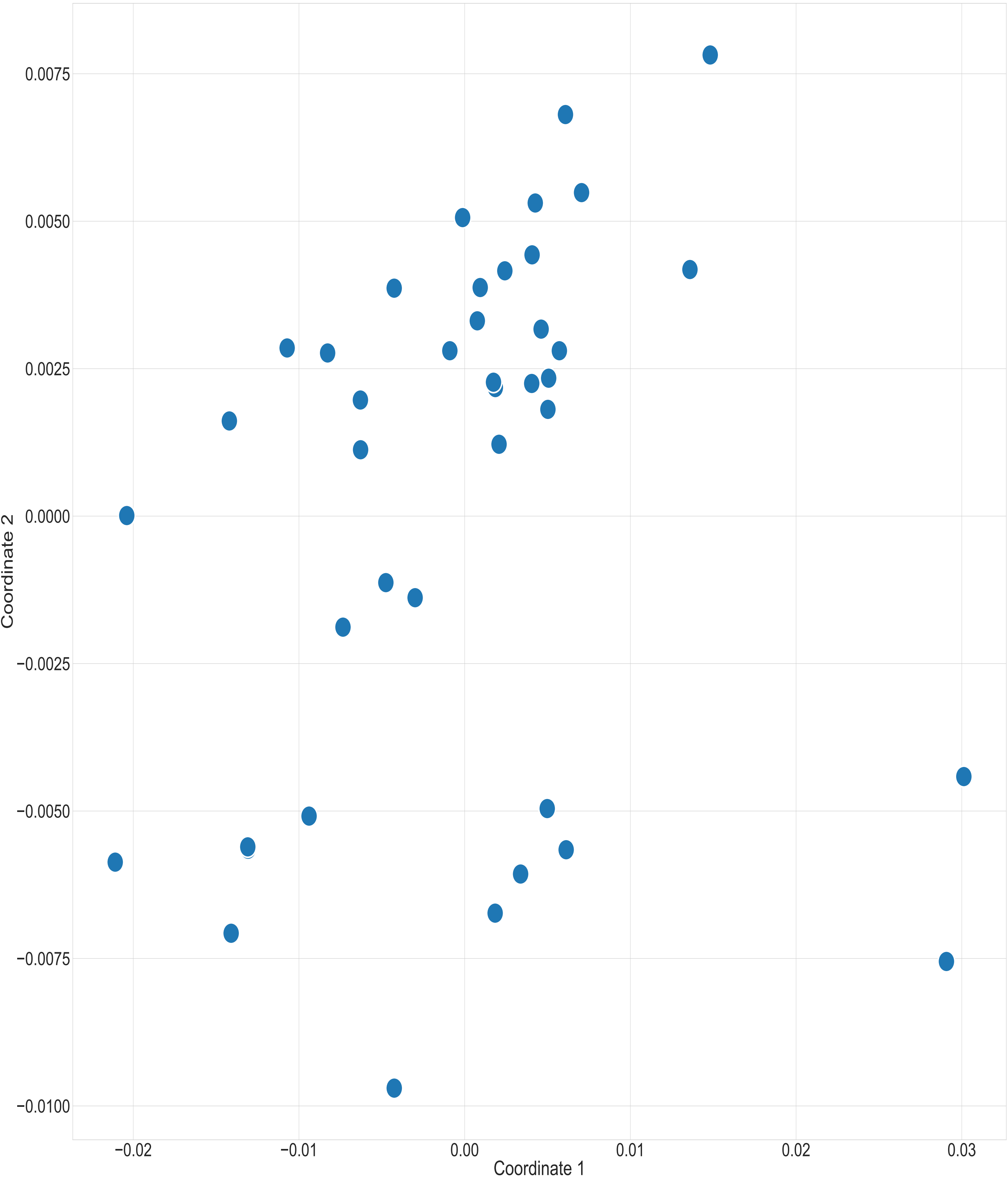
Scatter plot of Siderophore *Micromonosporaceae* BGCs analyzed via multidimensional scaling using chebyshev pairwise distance.
Only Siderophore BGCs were considered for this analysis. Each dot represents an individual BGC.
Distance between Siderophore BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S57 :



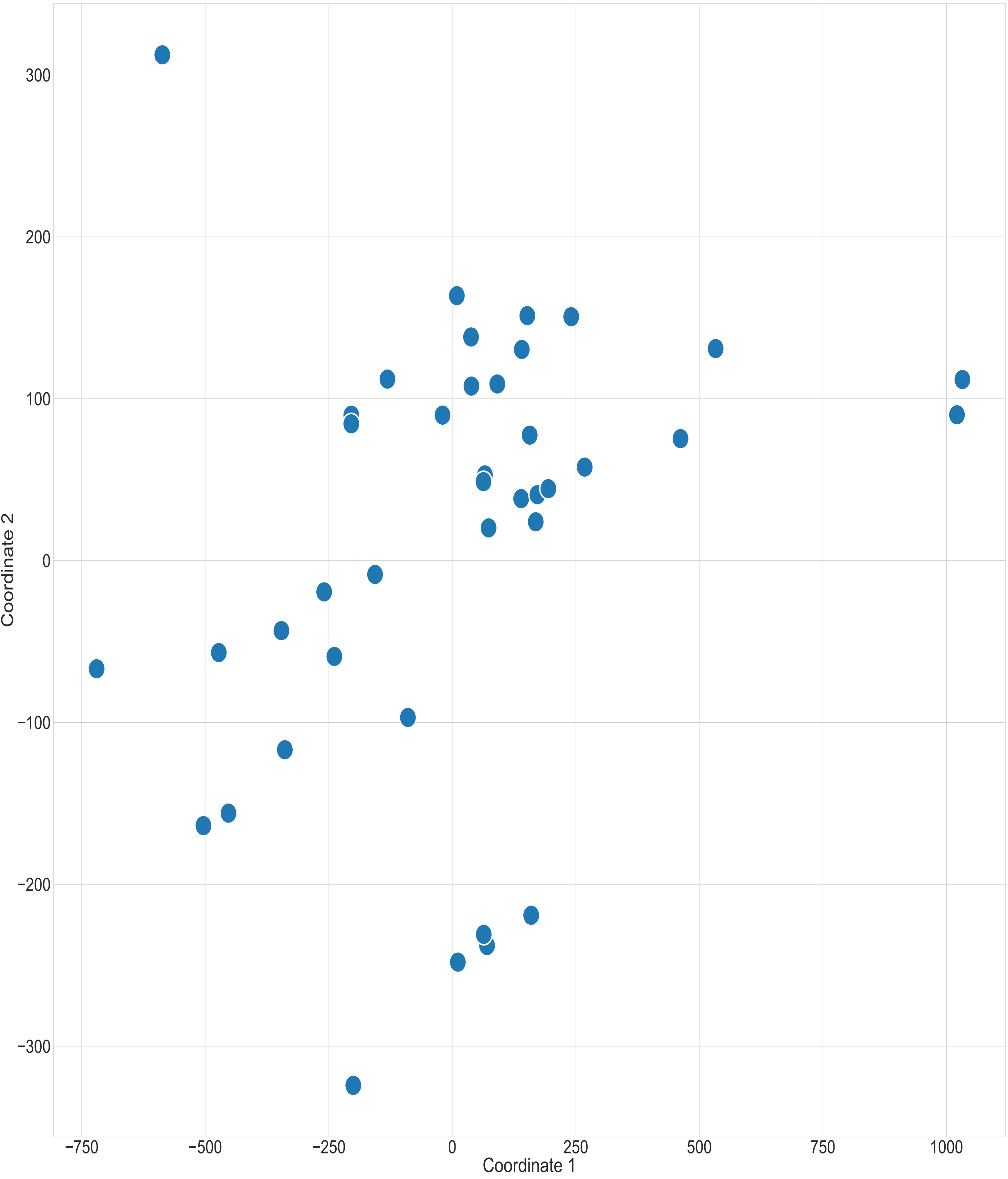
Scatter plot of Siderophore *Micromonosporaceae* BGCs analyzed via multidimensional scaling using l2 pairwise distance. Only Siderophore BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Siderophore BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S58 :



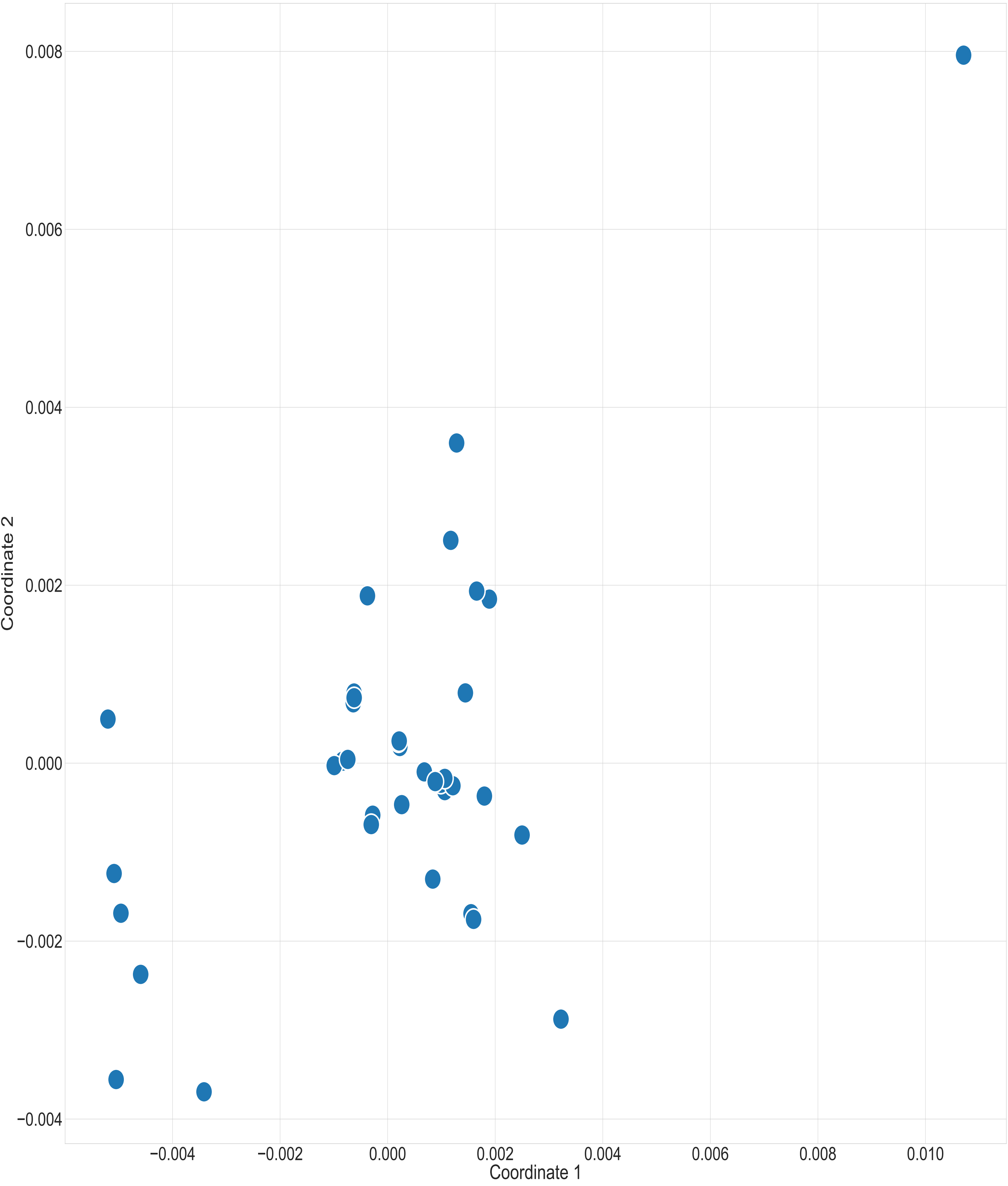
Scatter plot of Siderophore *Micromonosporaceae* BGCs analyzed via multidimensional scaling using braycurtis pairwise distance.
Only Siderophore BGCs were considered for this analysis. Each dot represents an individual BGC.
Distance between Siderophore BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S59 :



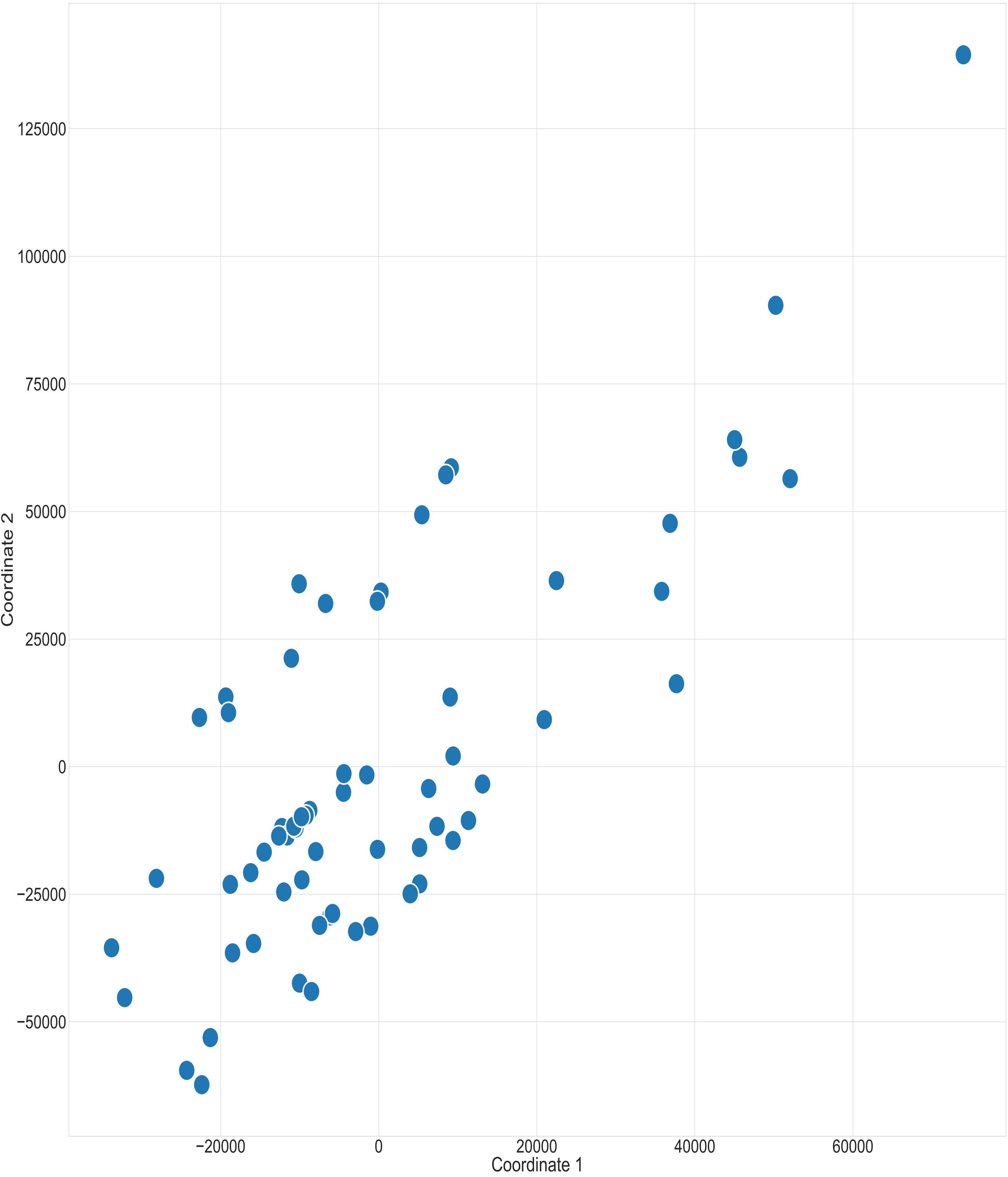
Scatter plot of Siderophore *Micromonosporaceae* BGCs analyzed via multidimensional scaling using canberra pairwise distance. Only Siderophore BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Siderophore BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S60 :



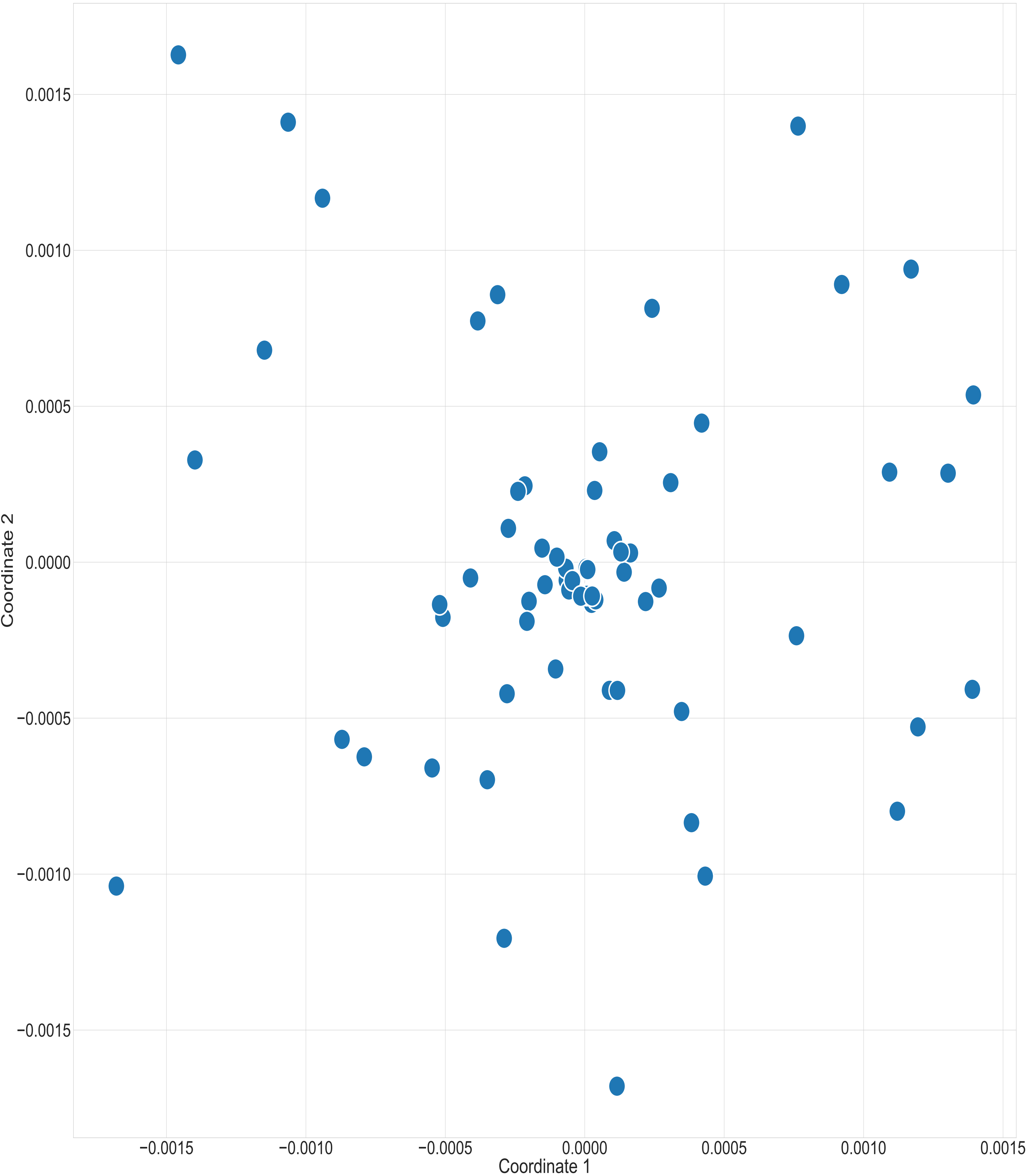
Scatter plot of Siderophore *Micromonosporaceae* BGCs analyzed via multidimensional scaling using correlation pairwise distance. Only Siderophore BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Siderophore BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S61 :



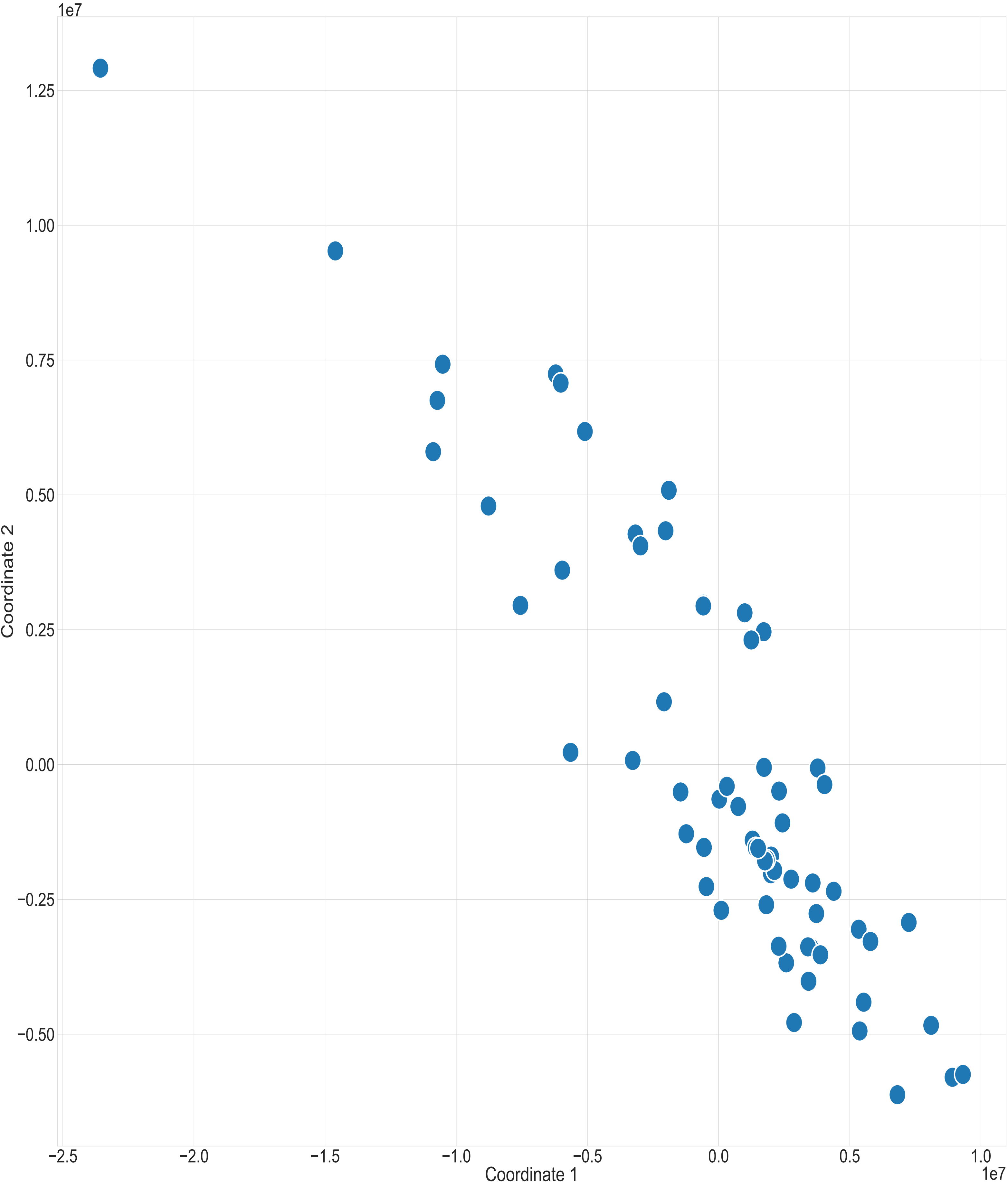
Scatter plot of PKS-NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using euclidean pairwise distance.
Only PKS-NRPS BGCs were considered for this analysis. Each dot represents an individual BGC.
Distance between PKS-NRPS BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S62 :



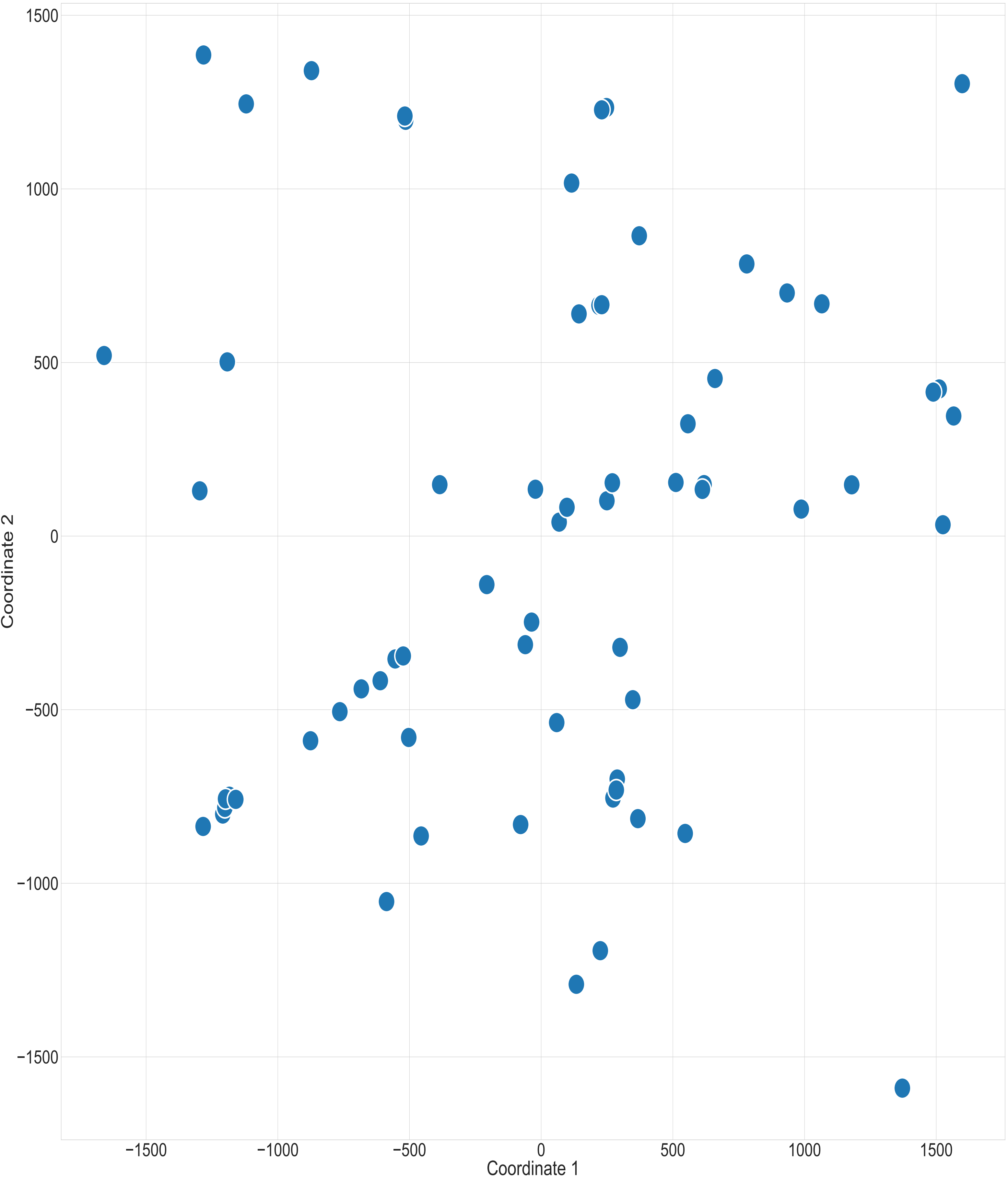
Scatter plot of PKS-NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cosine pairwise distance. Only PKS-NRPS BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between PKS-NRPS BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S63 :



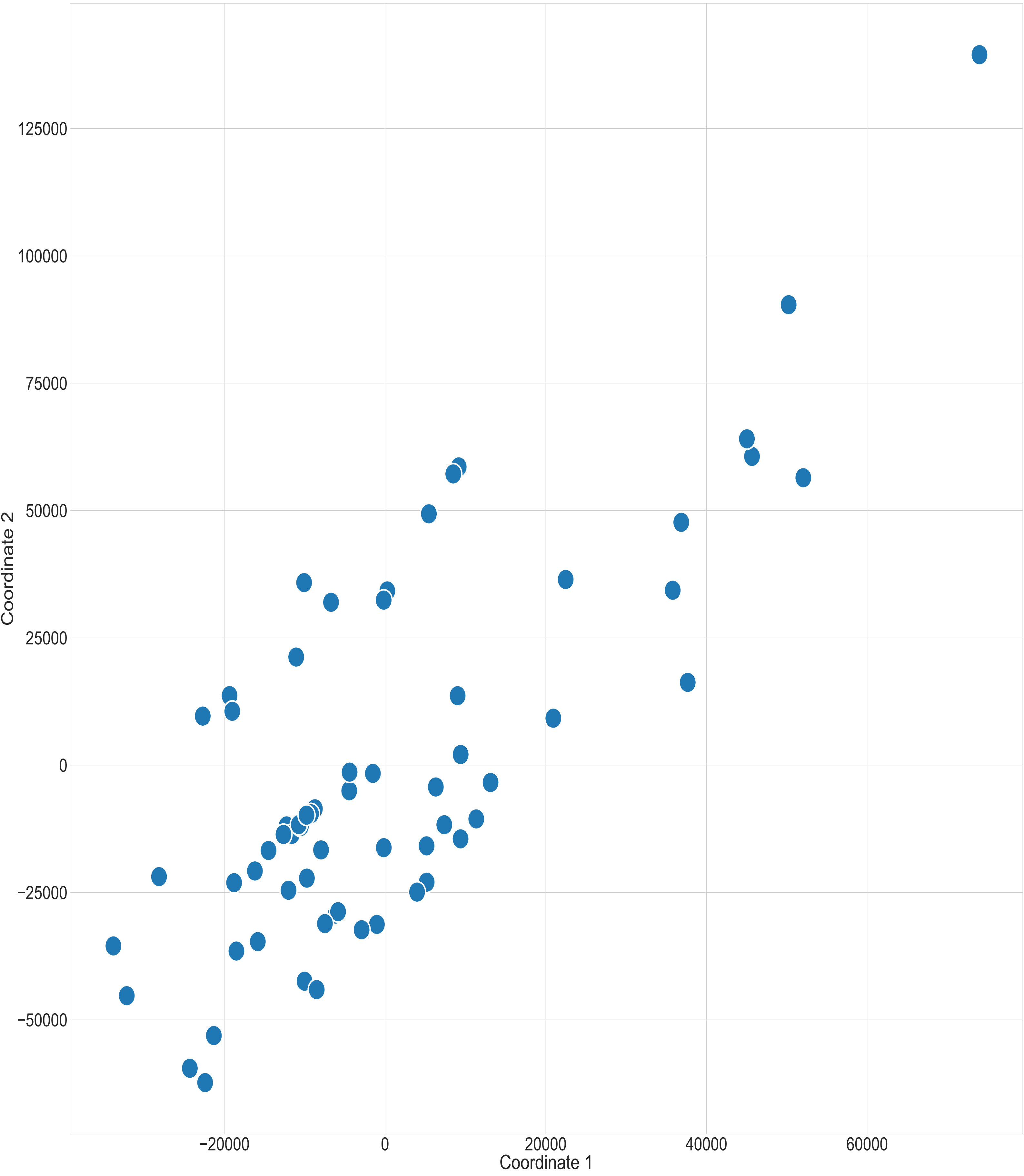
Scatter plot of PKS-NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cityblock pairwise distance. Only PKS-NRPS BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between PKS-NRPS BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S64 :



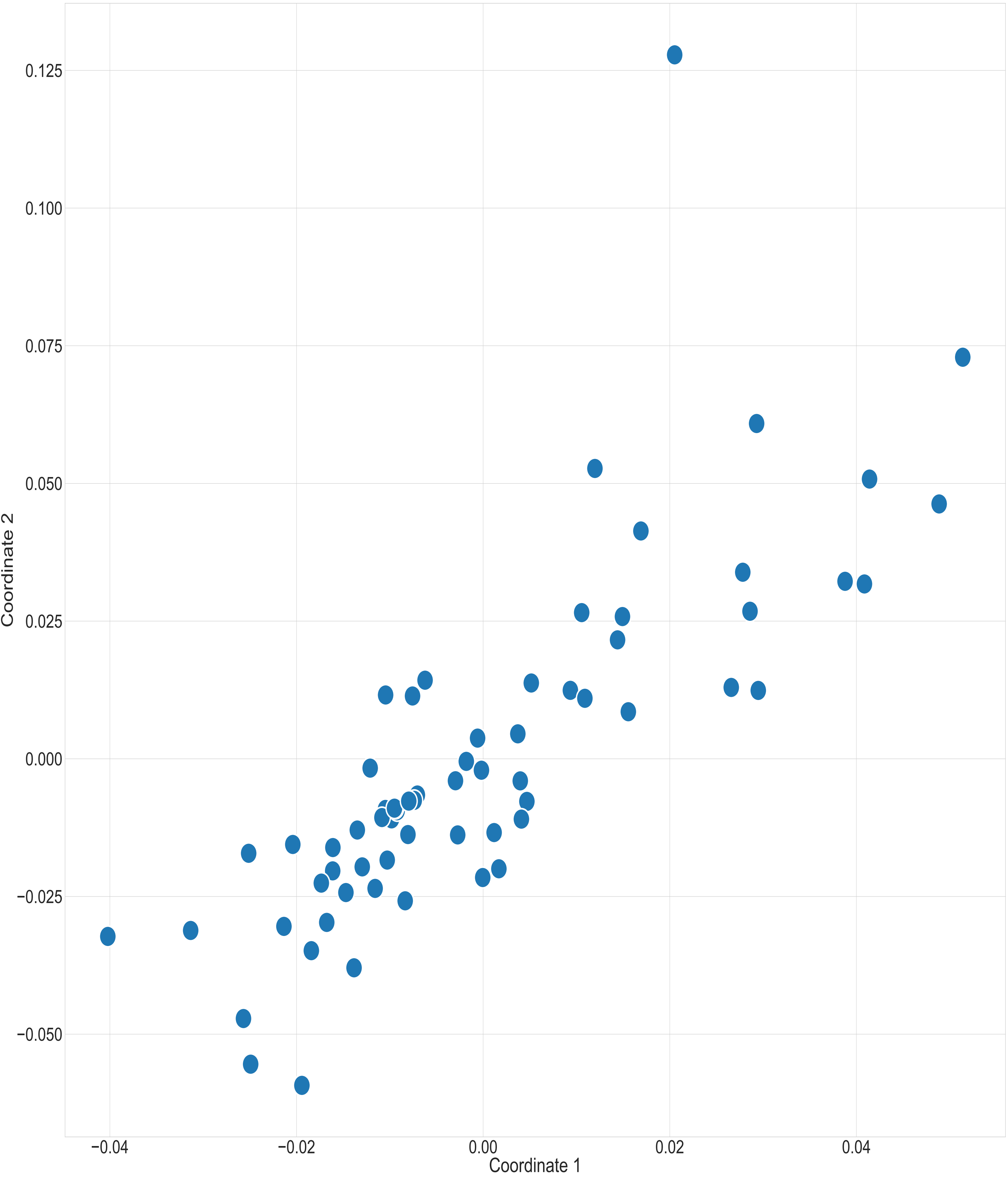
Scatter plot of PKS-NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using chebyshev pairwise distance. Only PKS-NRPS BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between PKS-NRPS BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S65 :



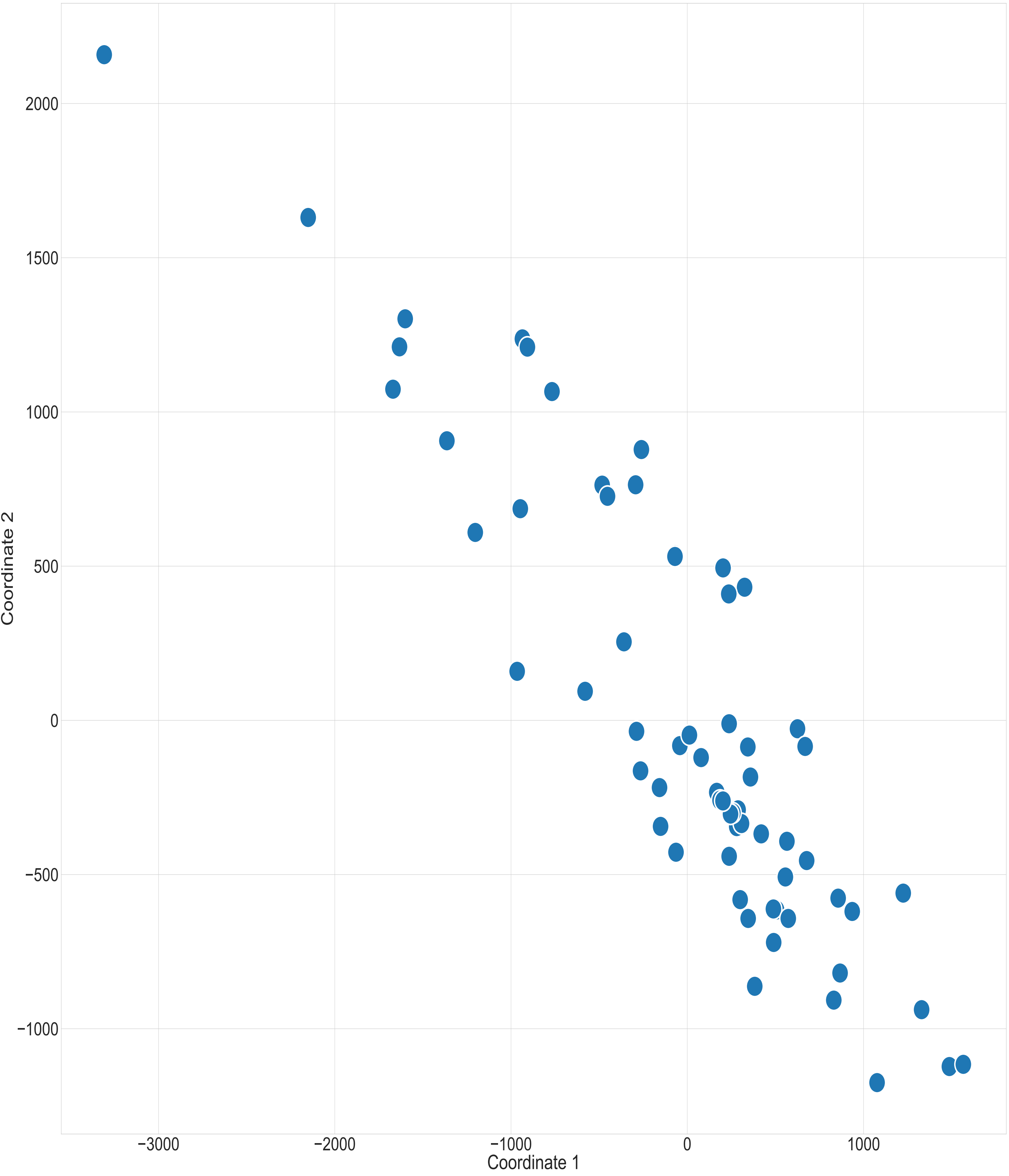
Scatter plot of PKS-NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using l2 pairwise distance.
Only PKS-NRPS BGCs were considered for this analysis. Each dot represents an individual BGC.
Distance between PKS-NRPS BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S66 :



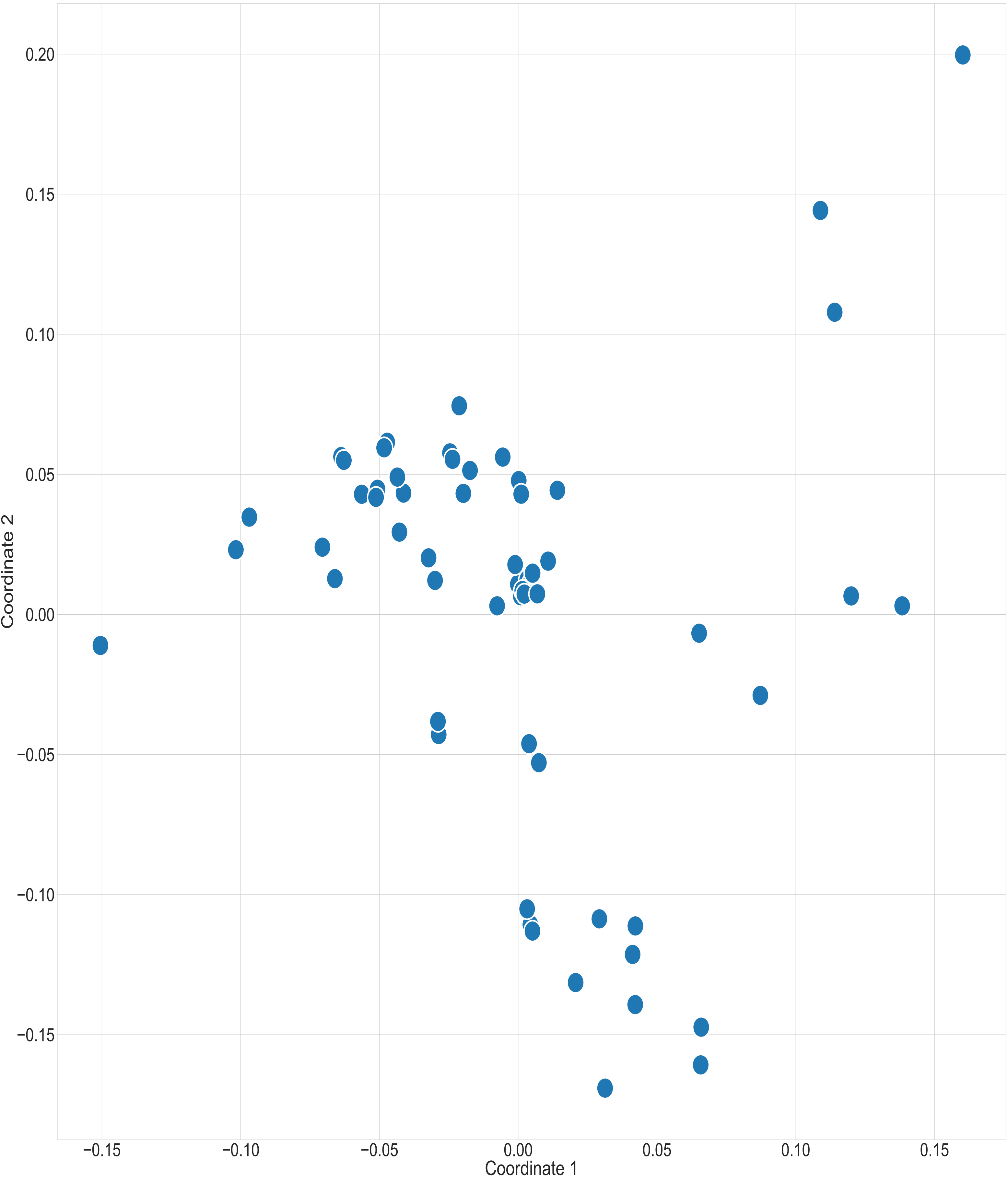
Scatter plot of PKS-NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using braycurtis pairwise distance. Only PKS-NRPS BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between PKS-NRPS BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S67 :



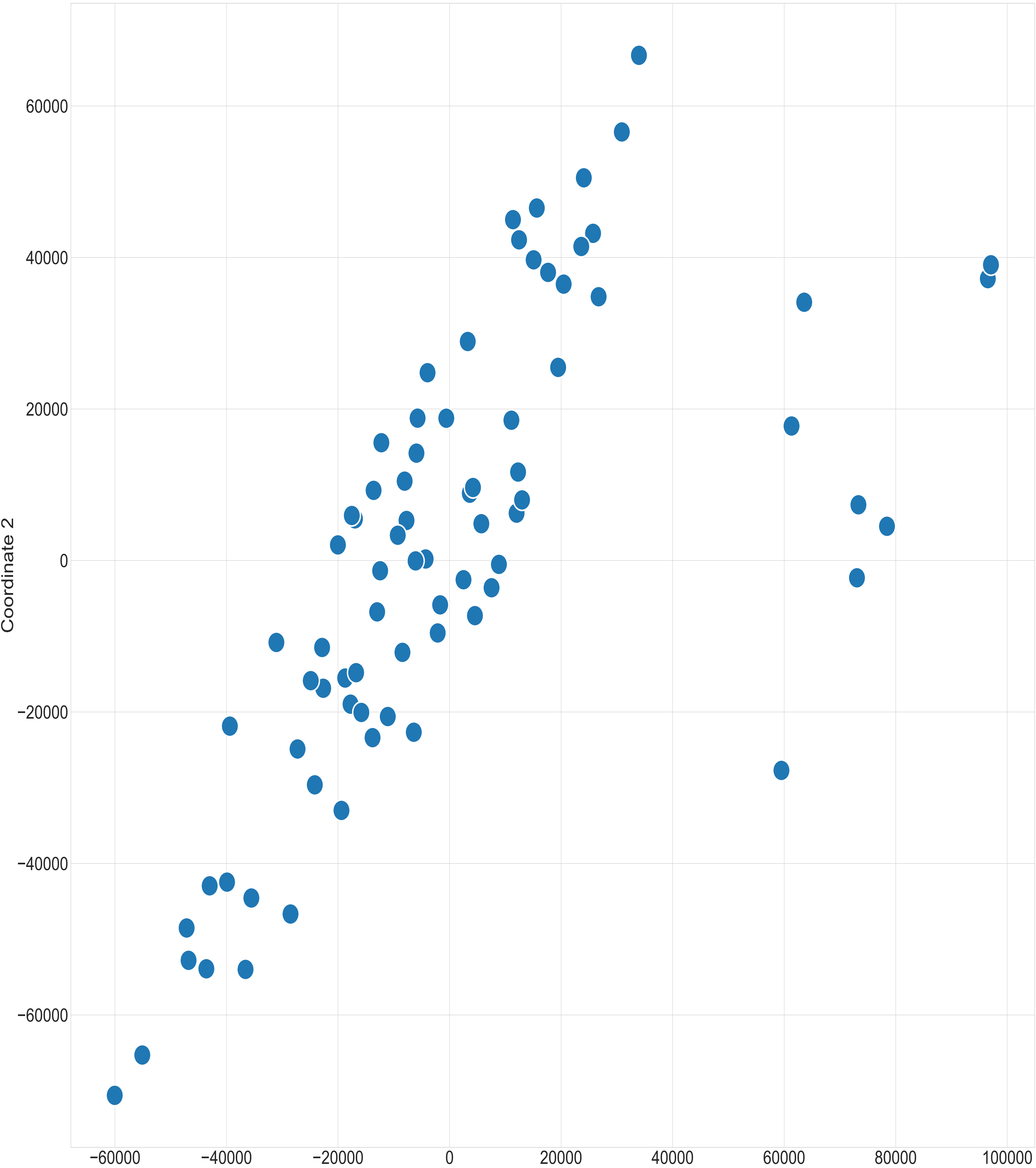
Scatter plot of PKS-NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using canberra pairwise distance.
Only PKS-NRPS BGCs were considered for this analysis. Each dot represents an individual BGC.
Distance between PKS-NRPS BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S68 :



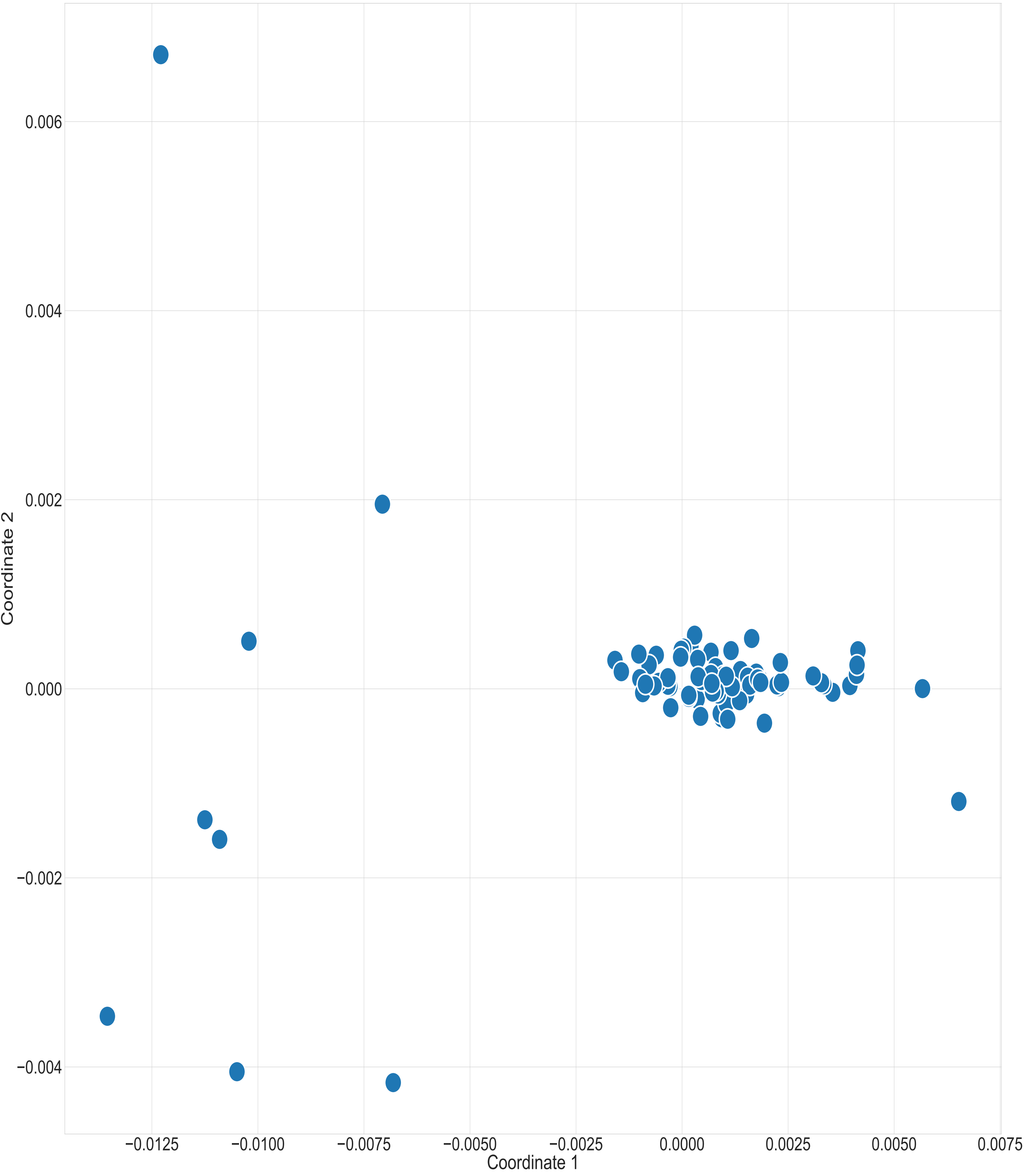
Scatter plot of PKS-NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using correlation pairwise distance. Only PKS-NRPS BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between PKS-NRPS BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S69 :



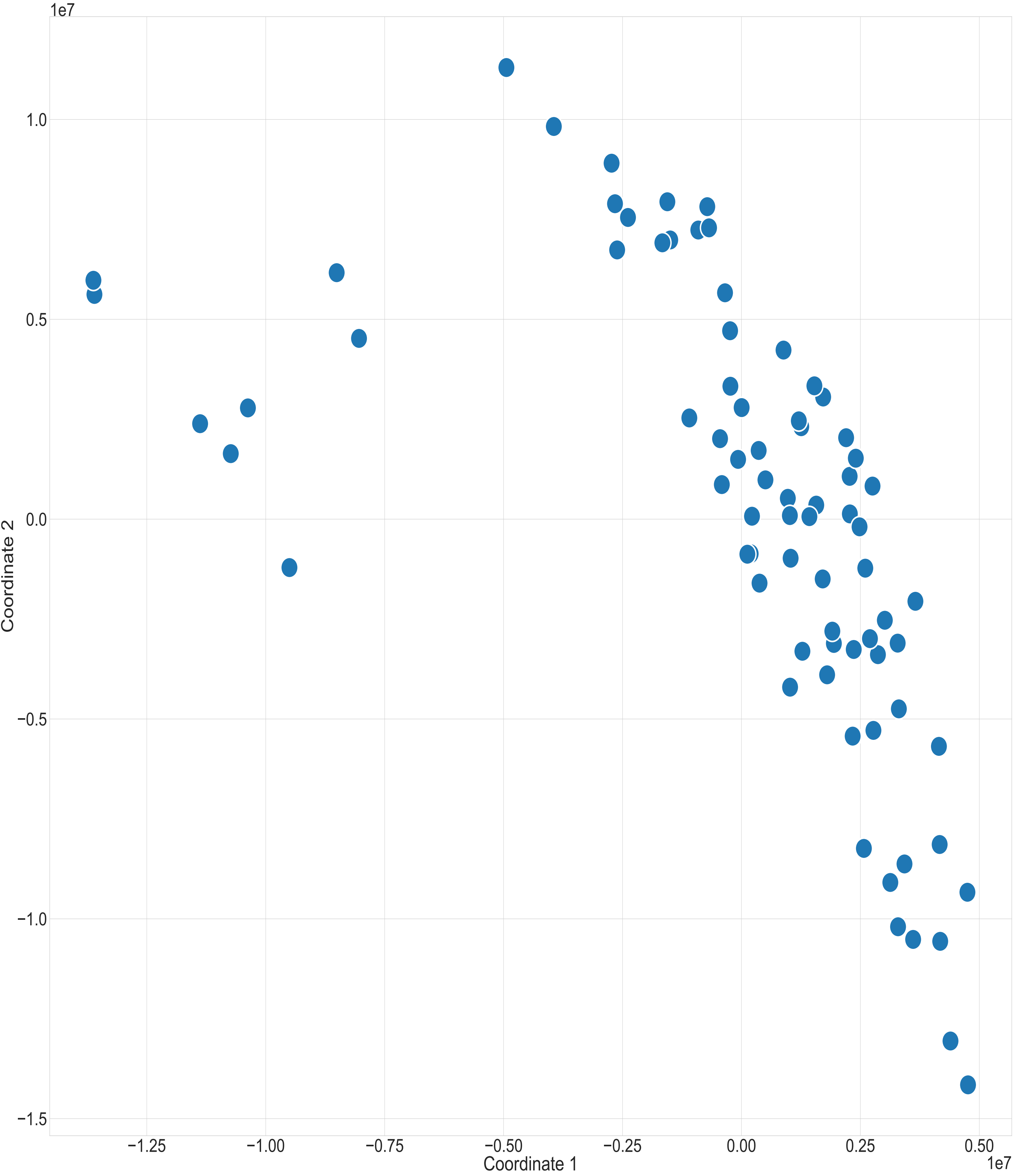
Scatter plot of NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using euclidean pairwise distance.
Only NRPS BGCs were considered for this analysis. Each dot represents an individual BGC.
Distance between NRPS BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S70 :



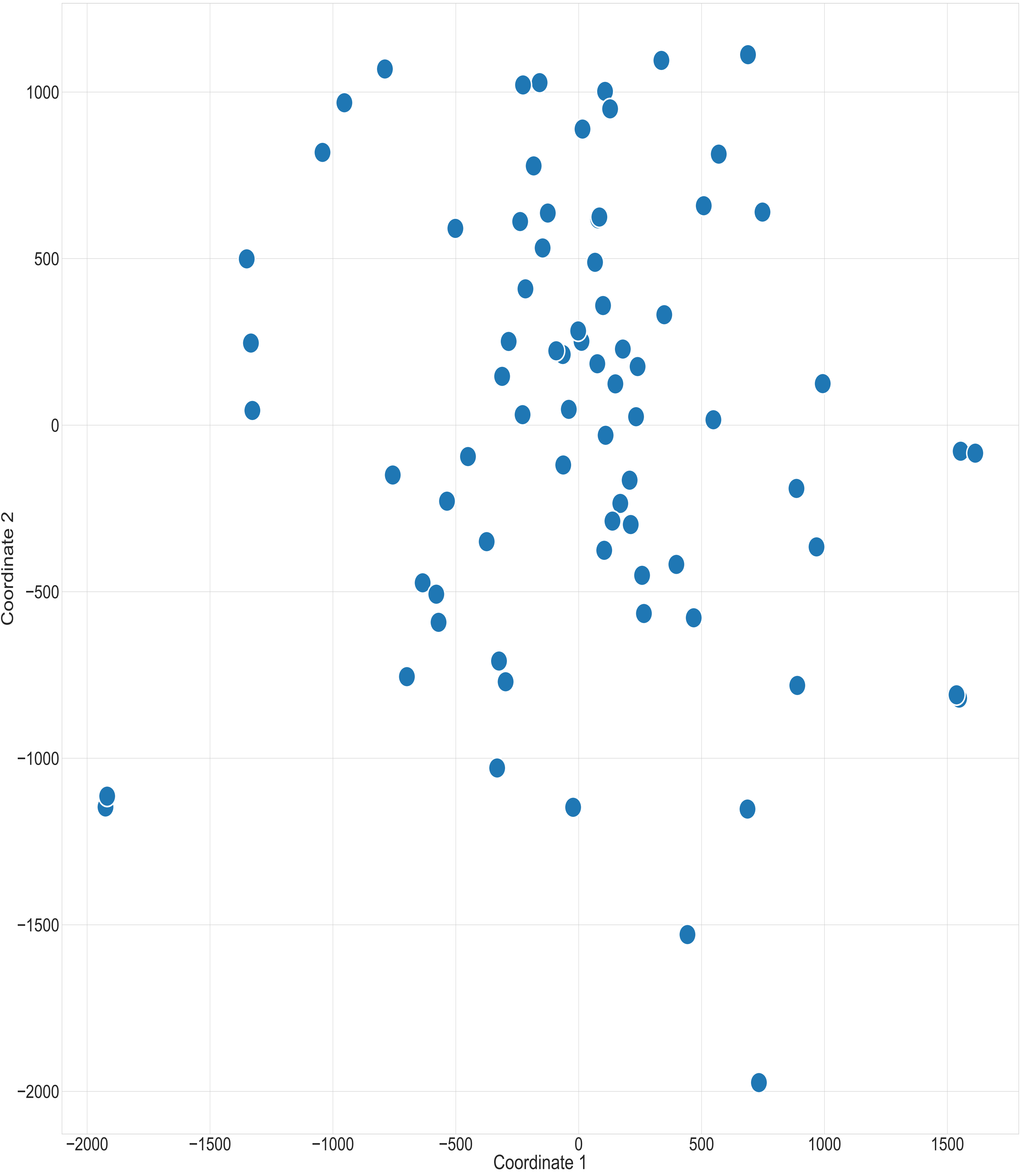
Scatter plot of NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cosine pairwise distance. Only NRPS BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between NRPS BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S71 :



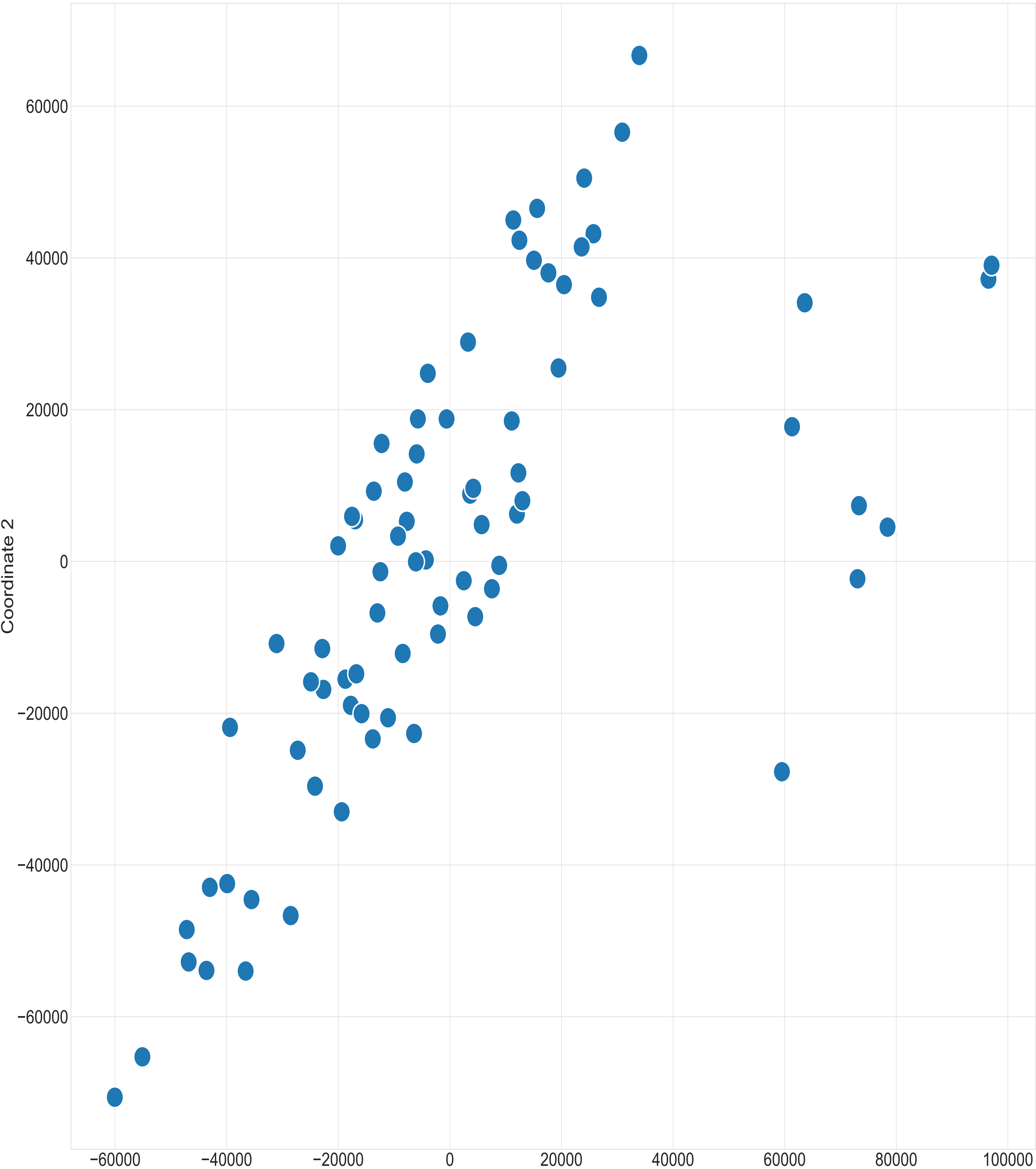
Scatter plot of NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cityblock pairwise distance. Only NRPS BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between NRPS BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S72 :



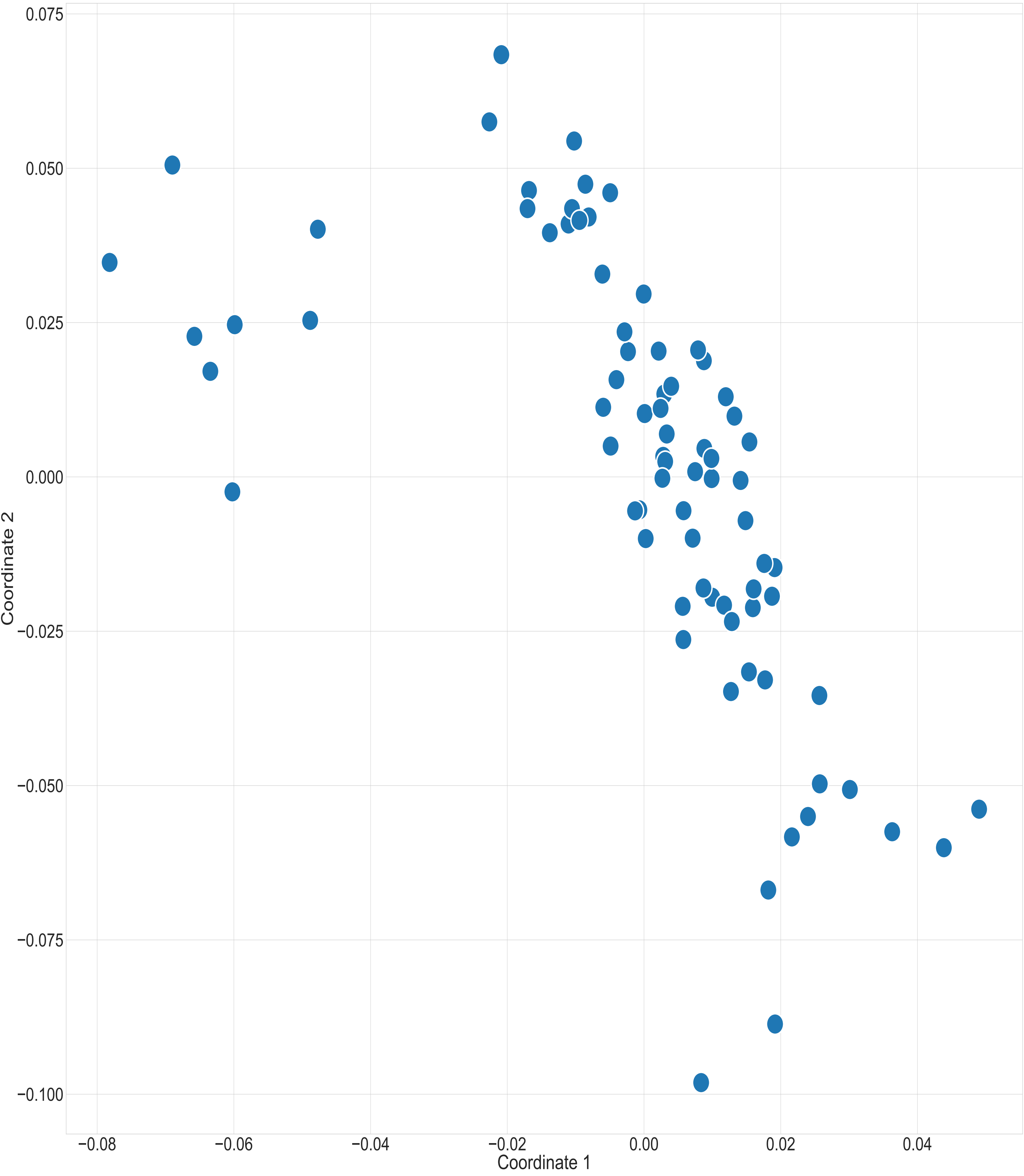
Scatter plot of NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using chebyshev pairwise distance.
Only NRPS BGCs were considered for this analysis. Each dot represents an individual BGC.
Distance between NRPS BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S73 :



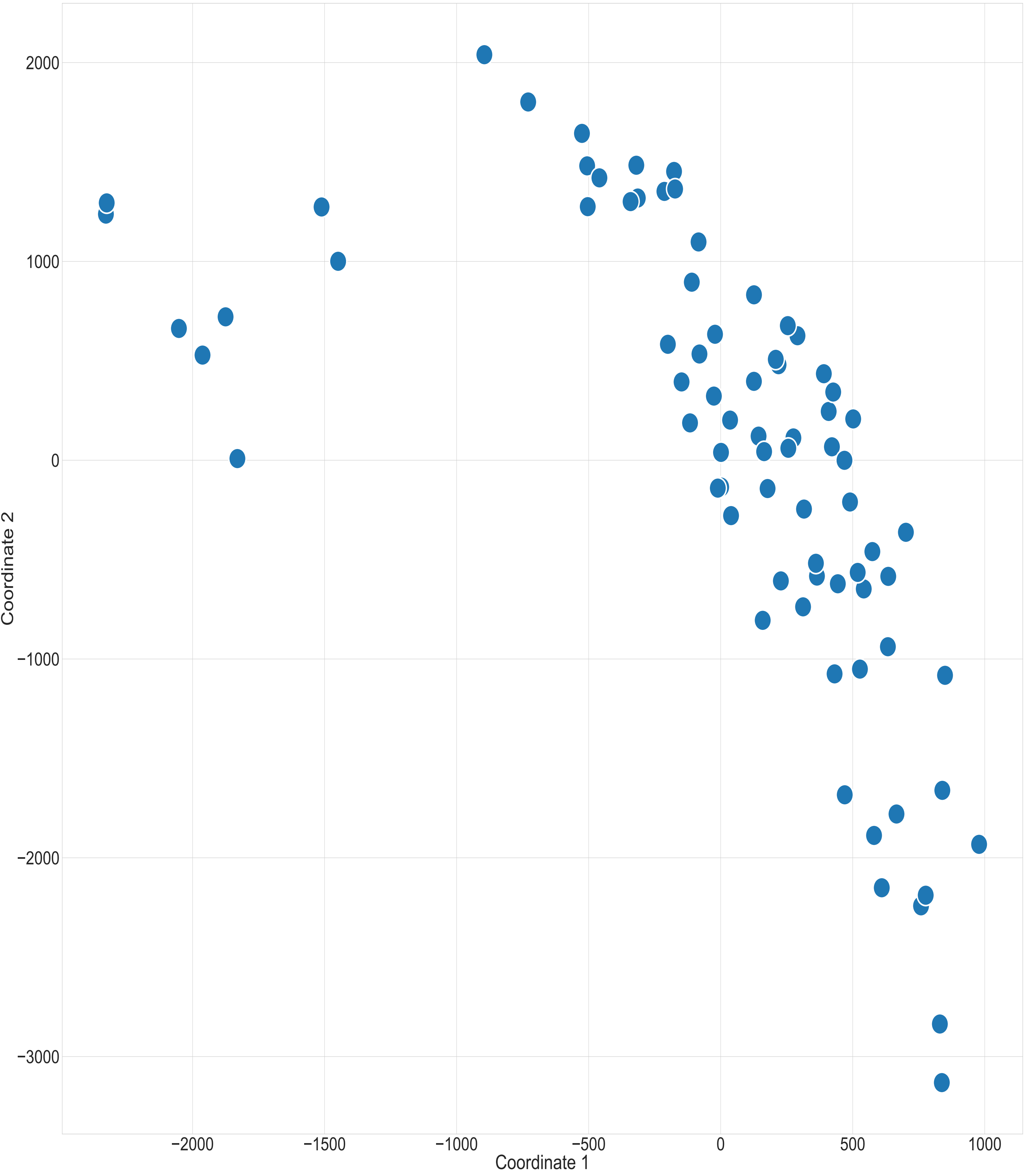
Scatter plot of NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using l2 pairwise distance. Only NRPS BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between NRPS BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S74 :



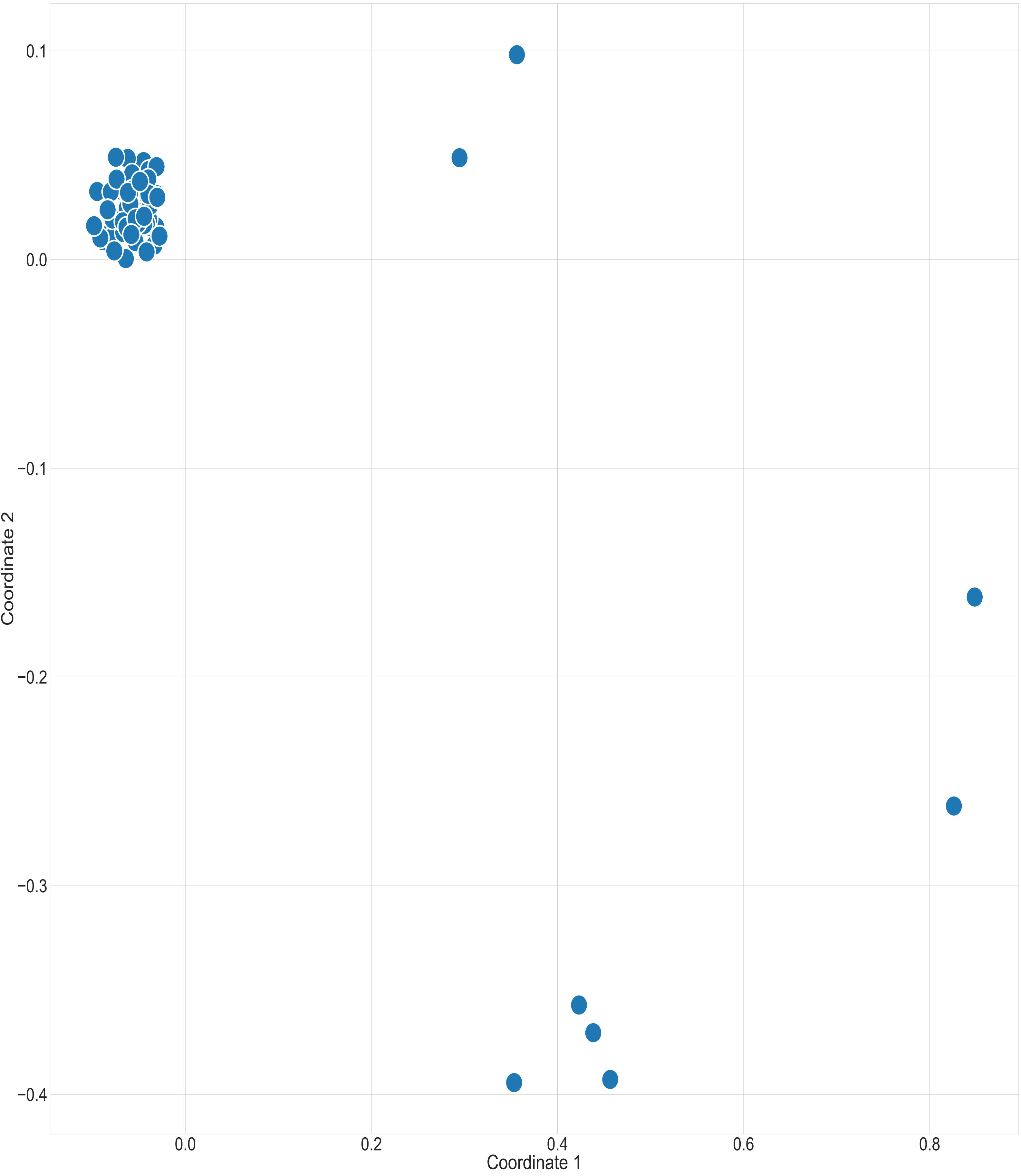
Scatter plot of NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using braycurtis pairwise distance.
Only NRPS BGCs were considered for this analysis. Each dot represents an individual BGC.
Distance between NRPS BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S75 :



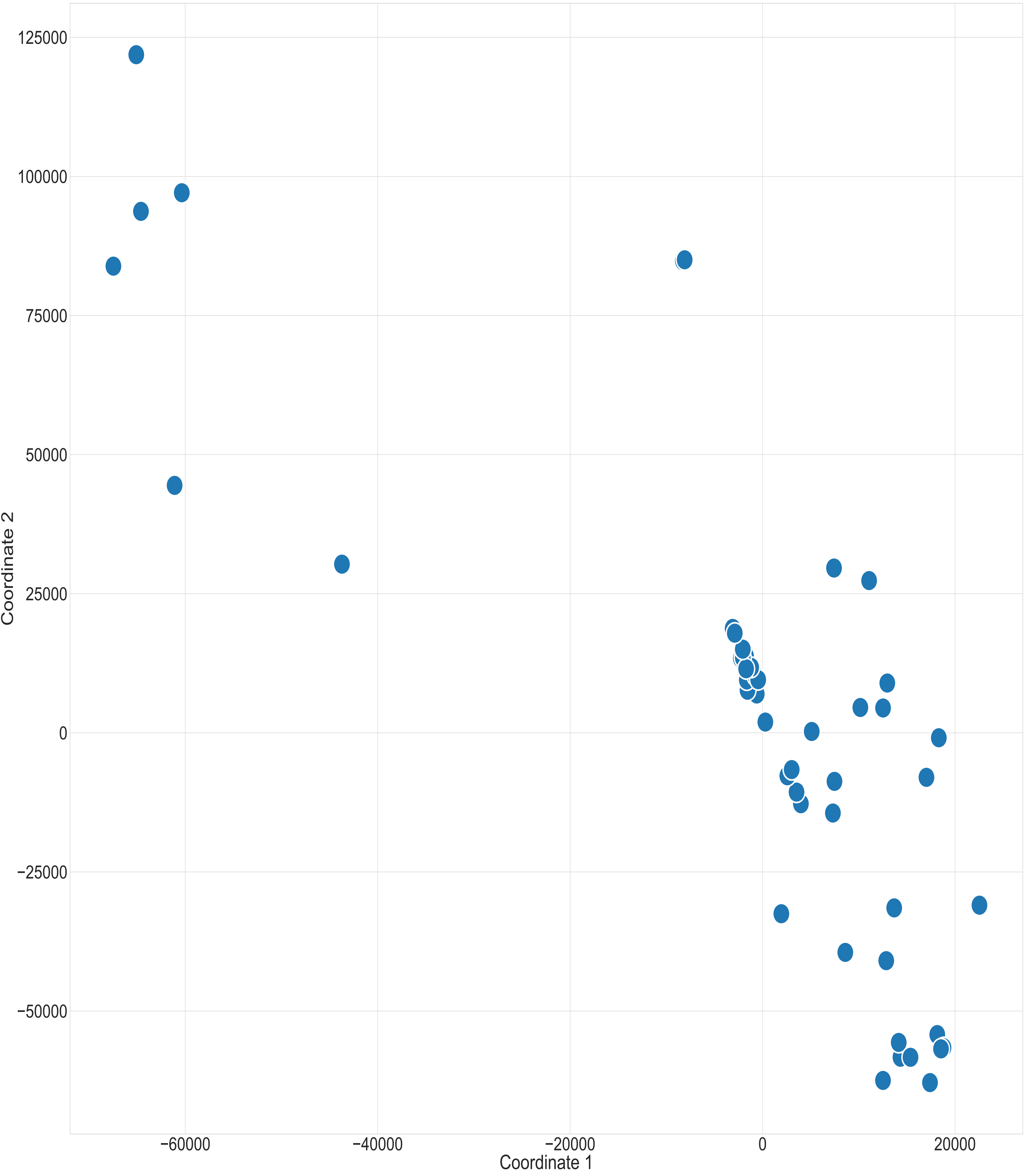
Scatter plot of NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using canberra pairwise distance.
Only NRPS BGCs were considered for this analysis. Each dot represents an individual BGC.
Distance between NRPS BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S76 :



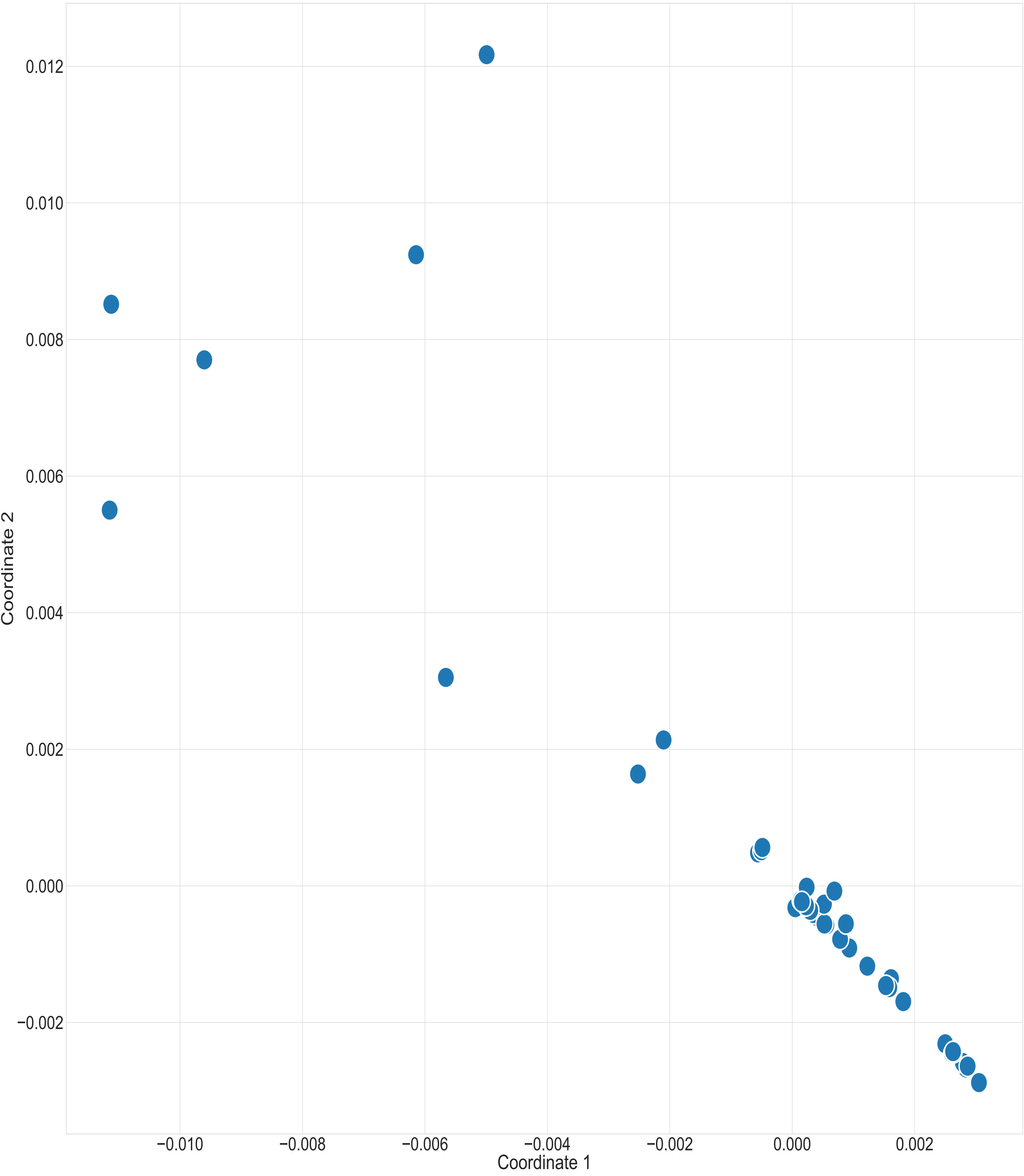
Scatter plot of NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using correlation pairwise distance. Only NRPS BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between NRPS BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S77 :



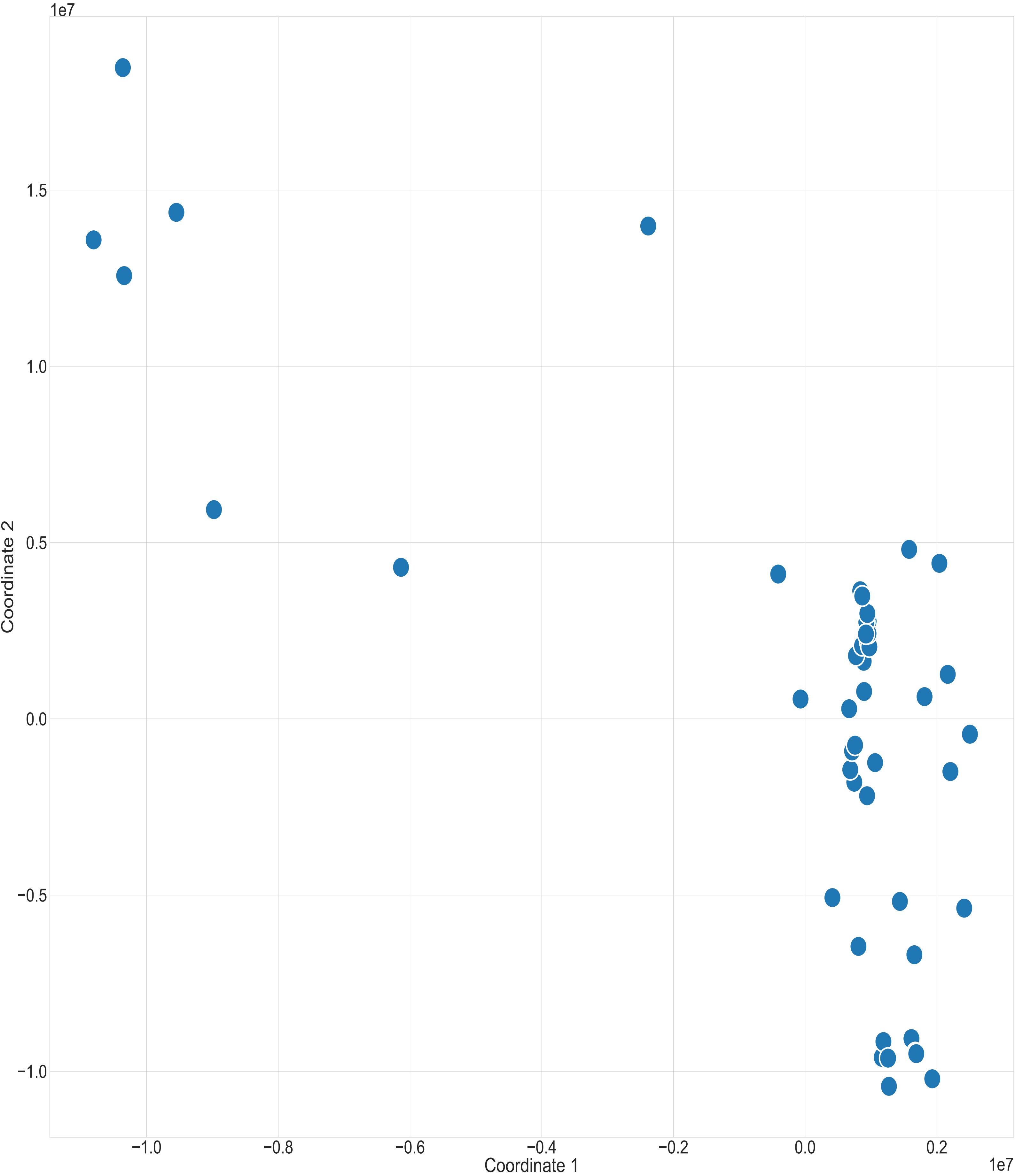
Scatter plot of Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using euclidean pairwise distance.
Only Other BGCs were considered for this analysis. Each dot represents an individual BGC.
Distance between Other BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S78 :



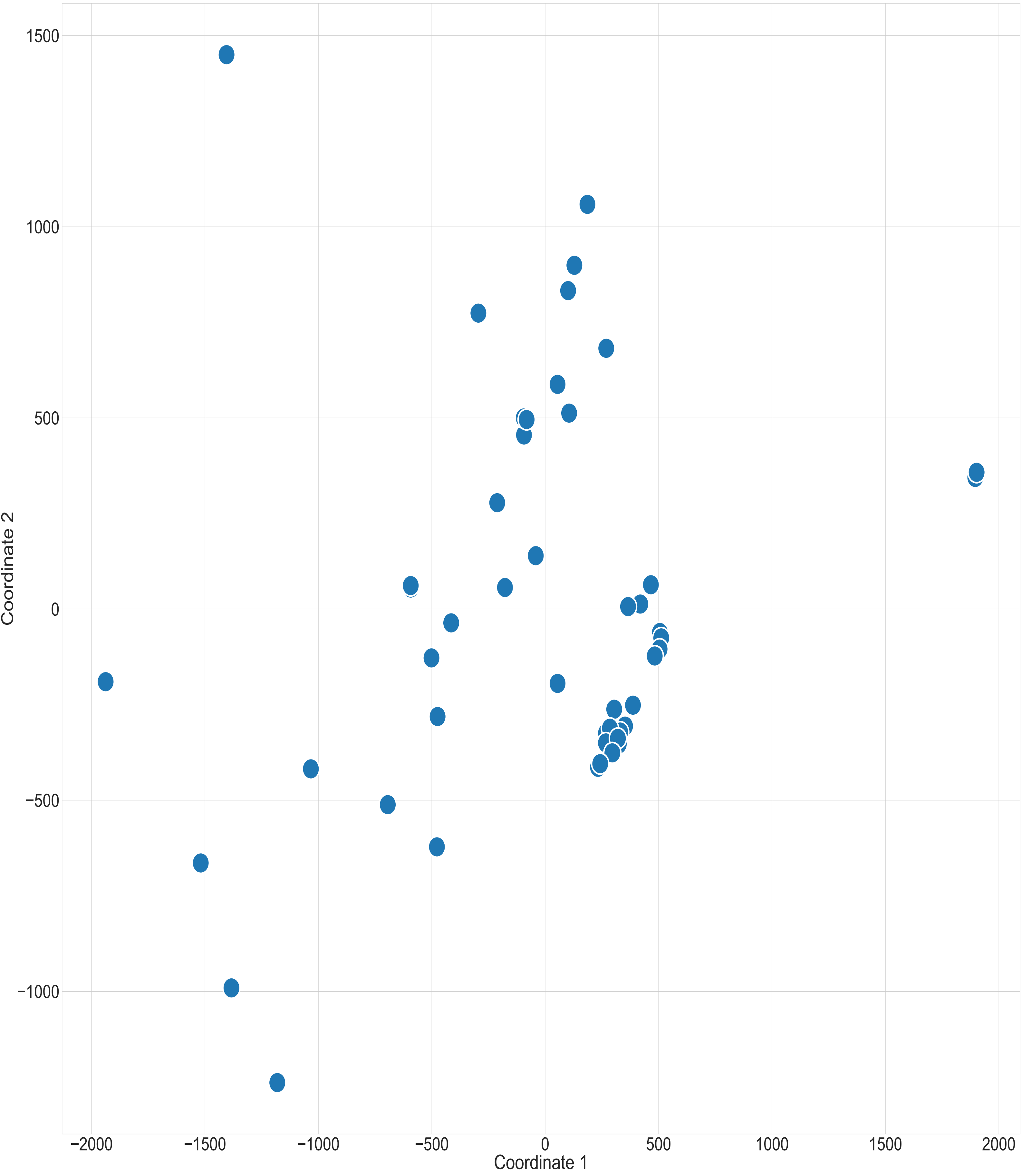
Scatter plot of Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cosine pairwise distance. Only Other BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Other BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S79 :



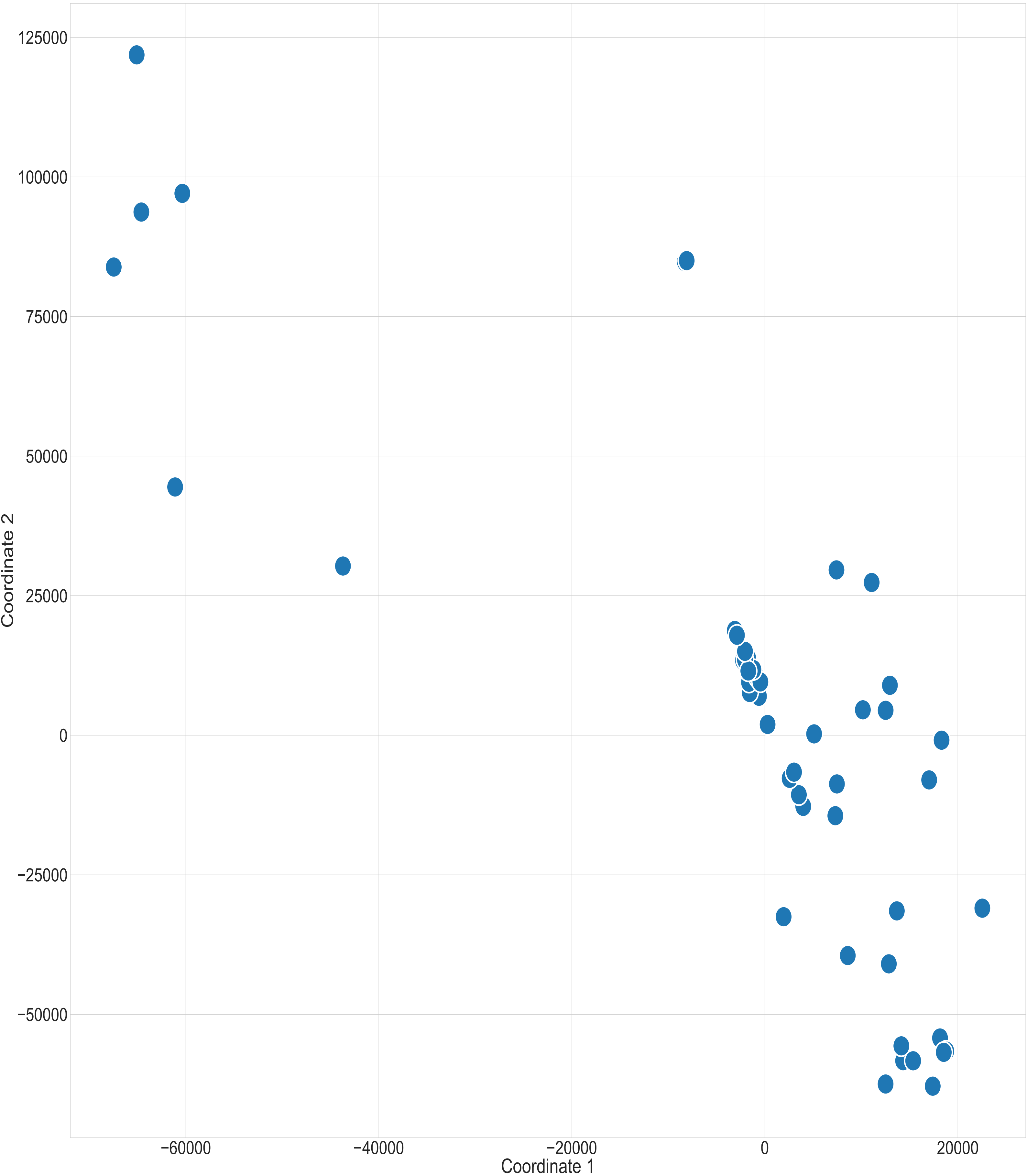
Scatter plot of Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cityblock pairwise distance. Only Other BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Other BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S80 :



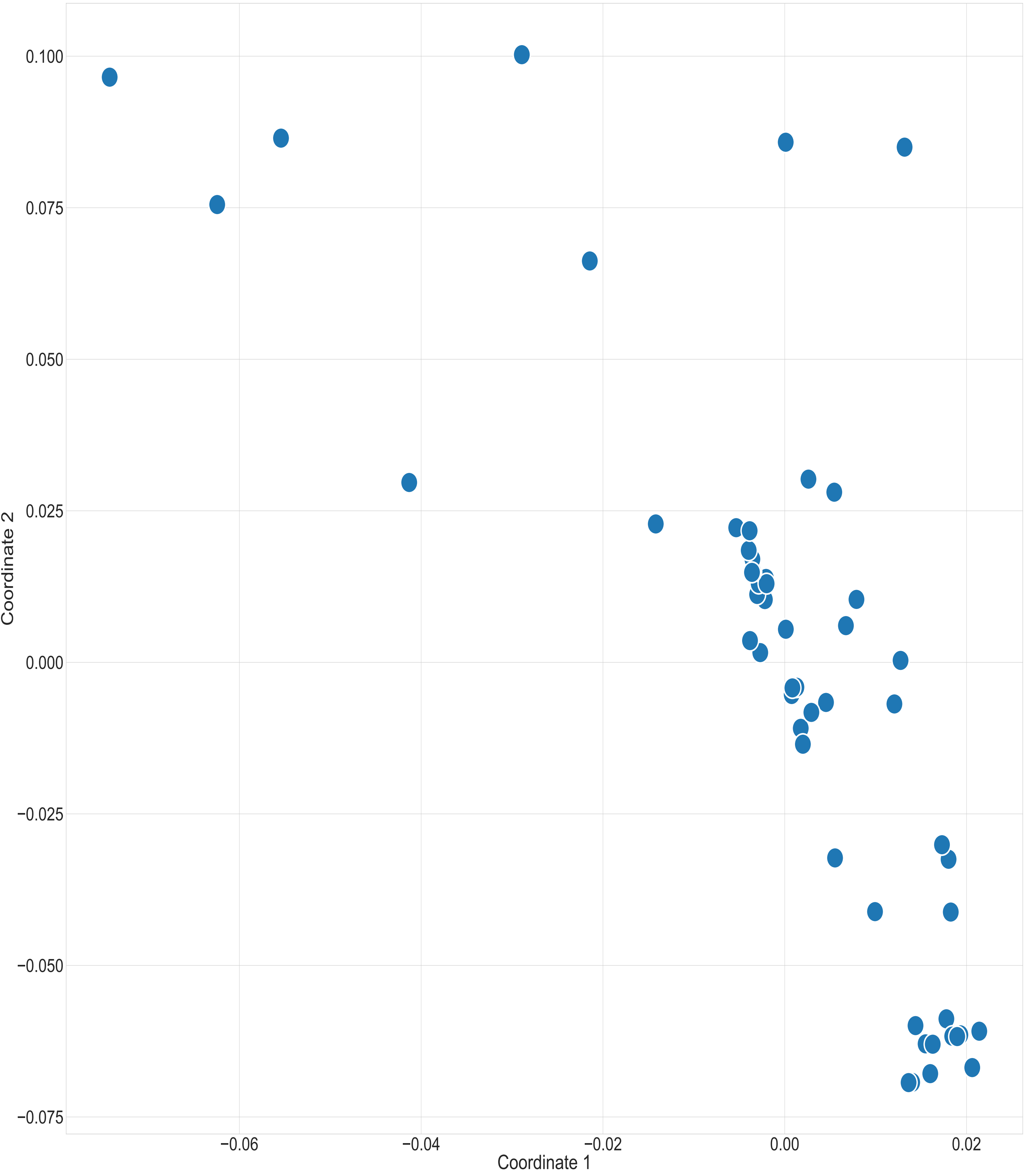
Scatter plot of Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using chebyshev pairwise distance.
Only Other BGCs were considered for this analysis. Each dot represents an individual BGC.
Distance between Other BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S81 :



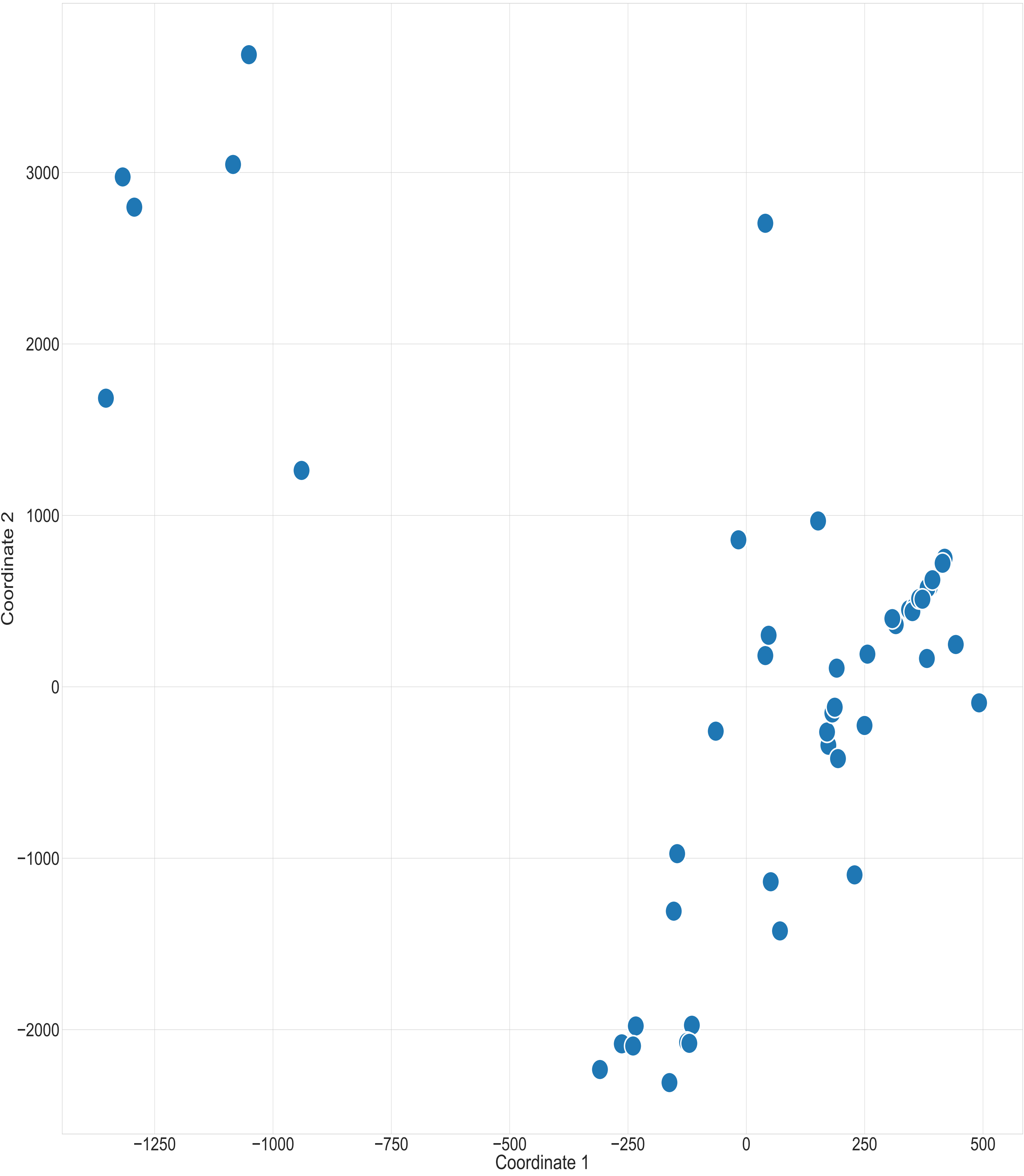
Scatter plot of Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using l2 pairwise distance.
Only Other BGCs were considered for this analysis. Each dot represents an individual BGC.
Distance between Other BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S82 :



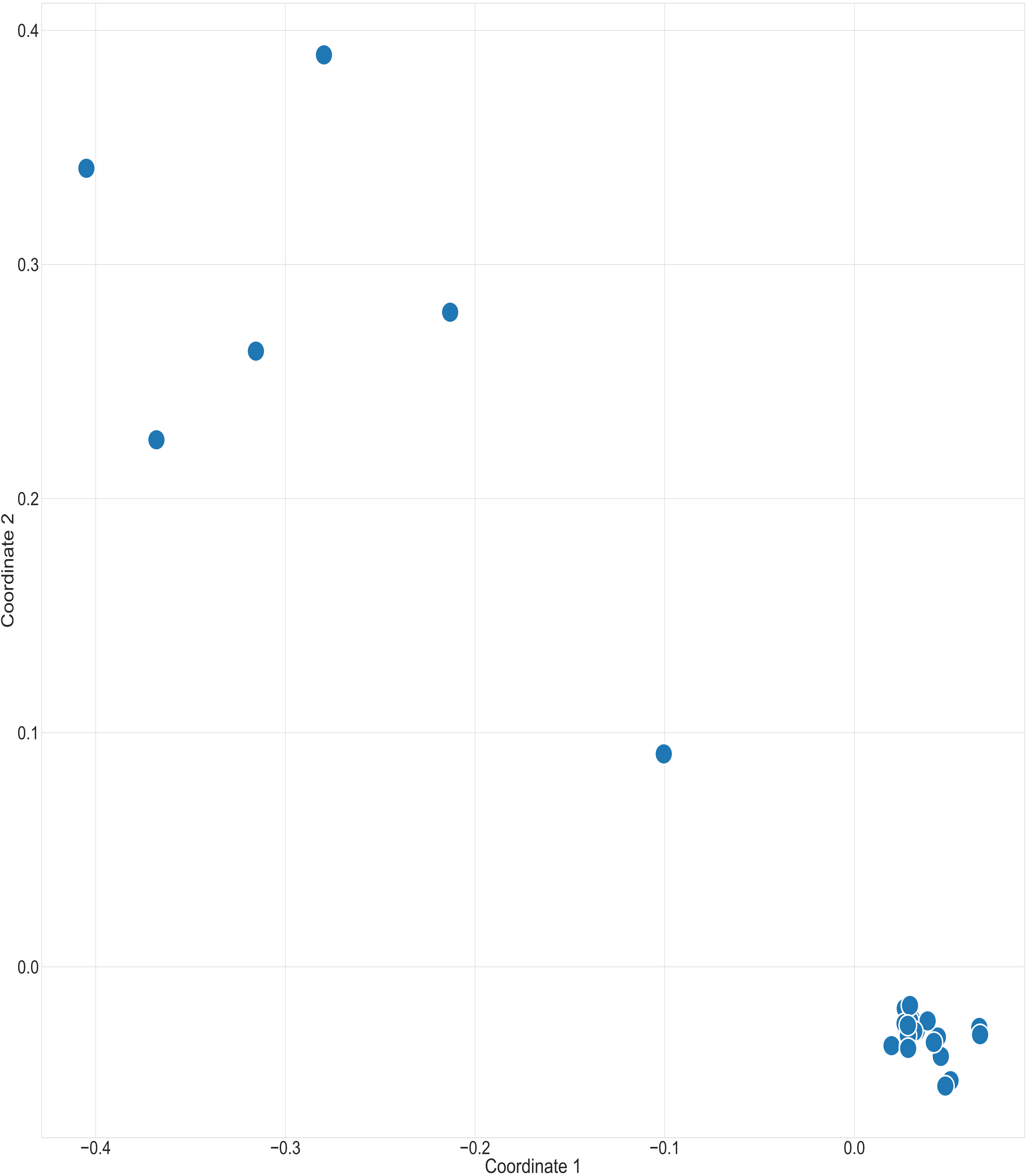
Scatter plot of Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using braycurtis pairwise distance. Only Other BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Other BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S83 :



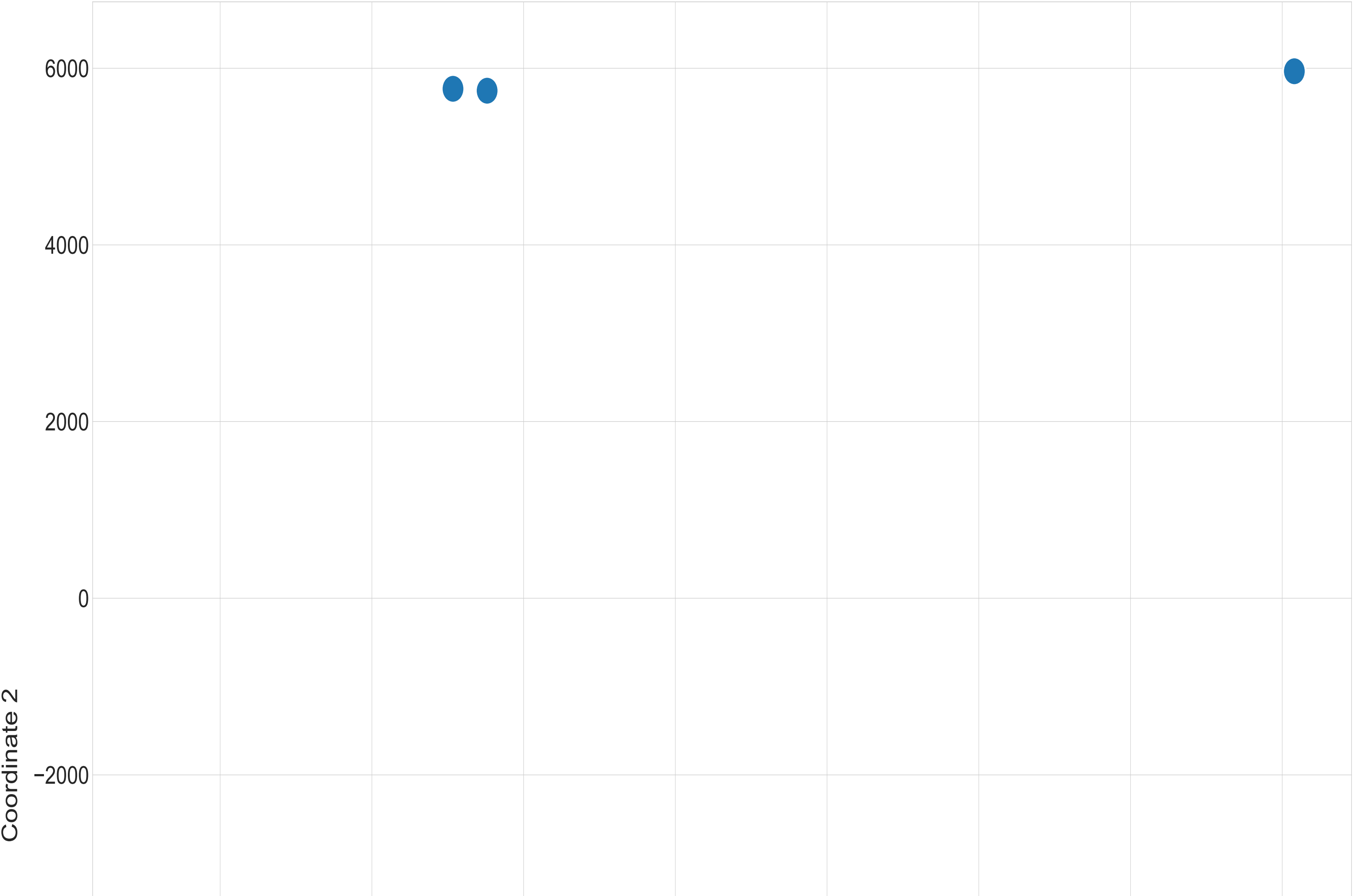
Scatter plot of Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using canberra pairwise distance.
Only Other BGCs were considered for this analysis. Each dot represents an individual BGC.
Distance between Other BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S84 :



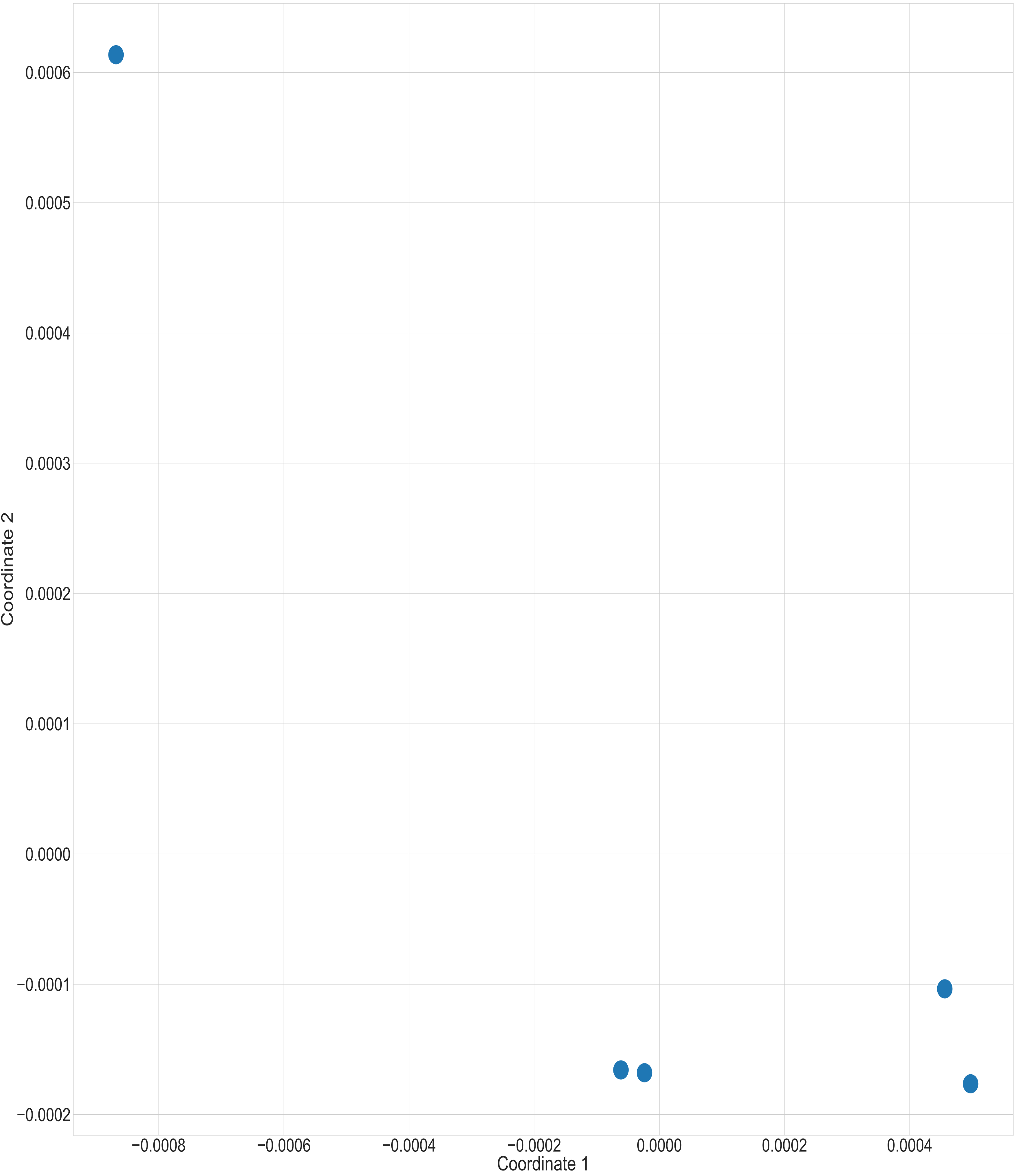
Scatter plot of Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using correlation pairwise distance.
Only Other BGCs were considered for this analysis. Each dot represents an individual BGC.
Distance between Other BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S85 :



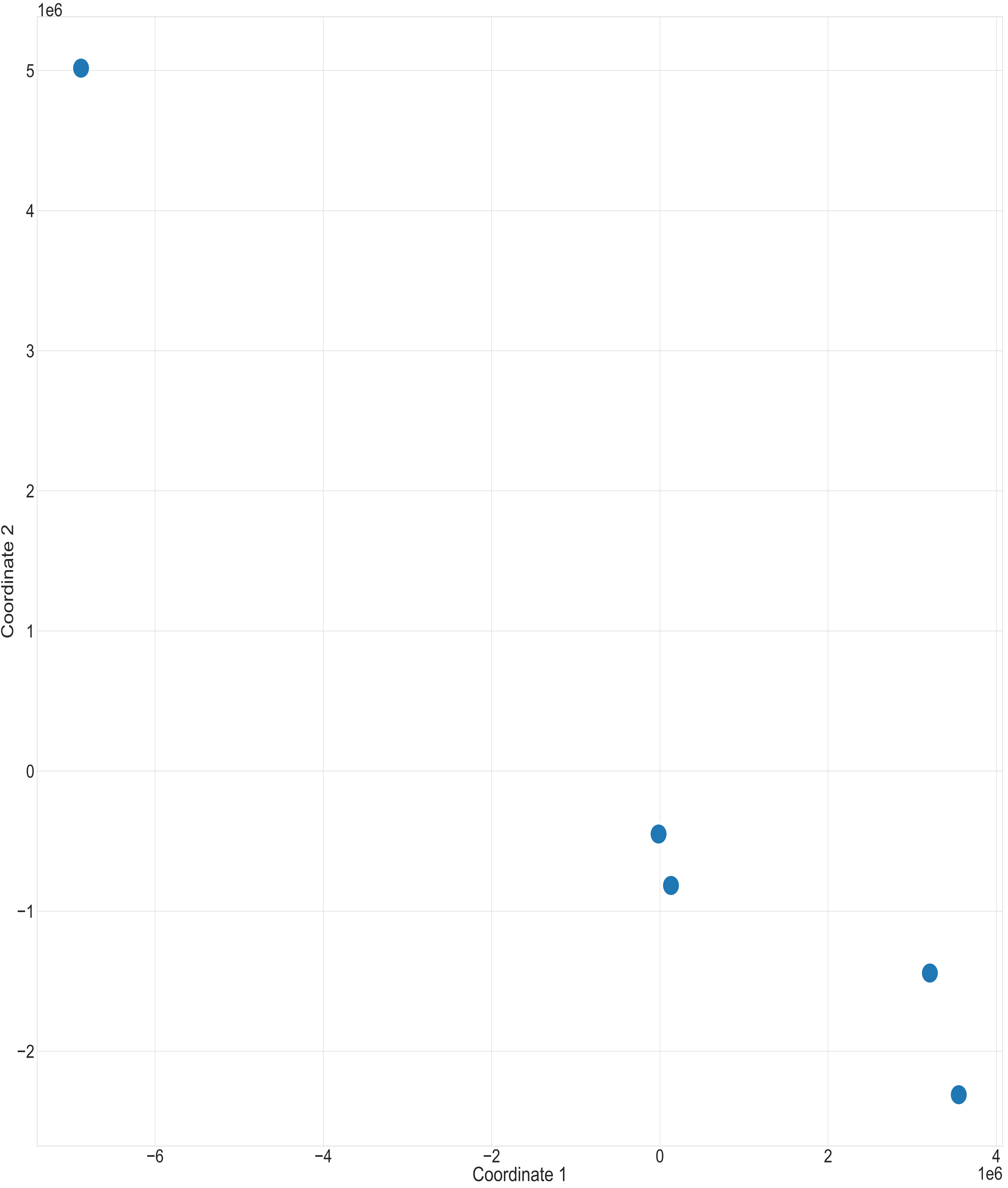
Scatter plot of Saccharides *Micromonosporaceae* BGCs analyzed via multidimensional scaling using euclidean pairwise distance. Only Saccharides BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Saccharides BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S86 :



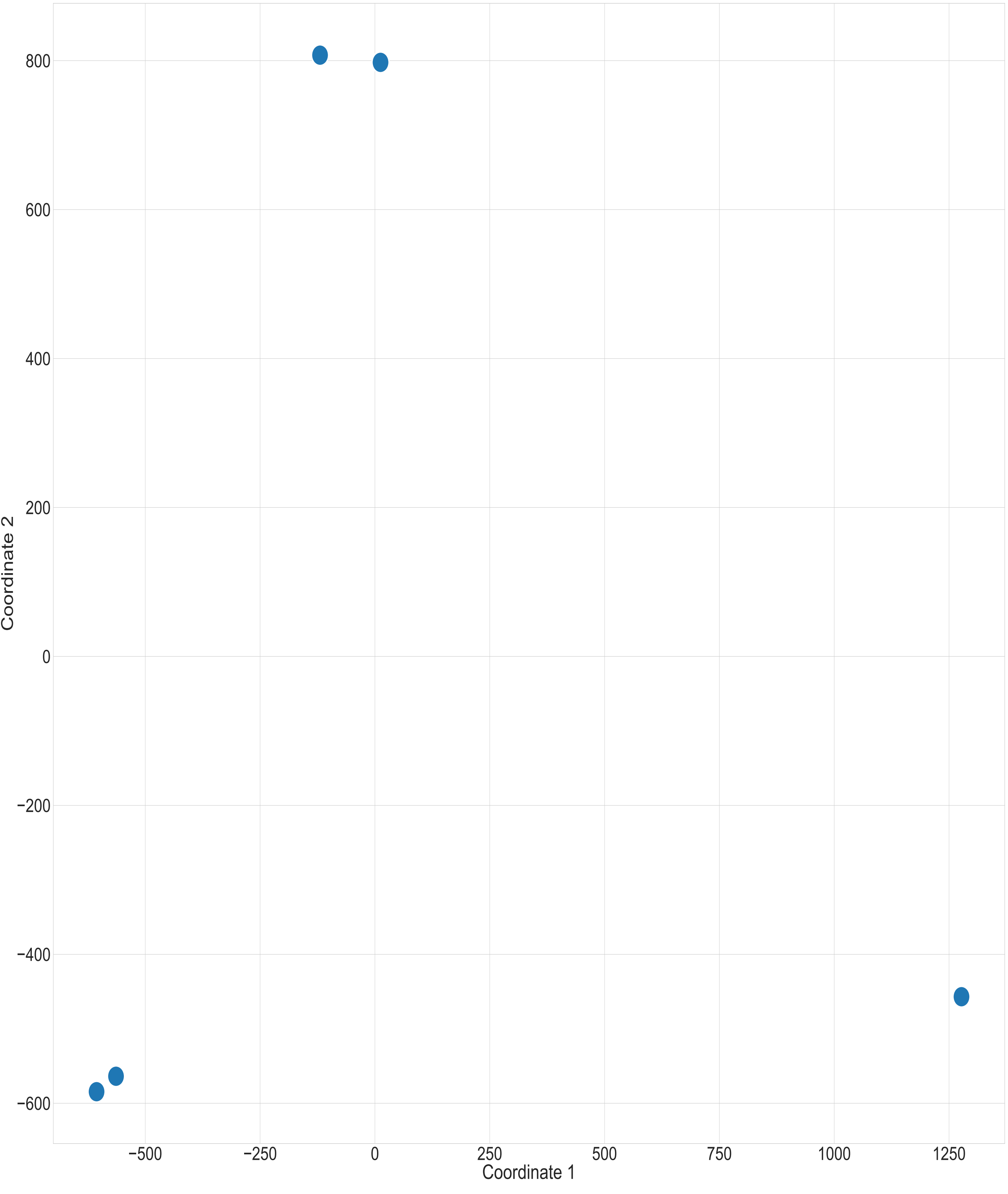
Scatter plot of Saccharides *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cosine pairwise distance.
Only Saccharides BGCs were considered for this analysis. Each dot represents an individual BGC.
Distance between Saccharides BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S87 :



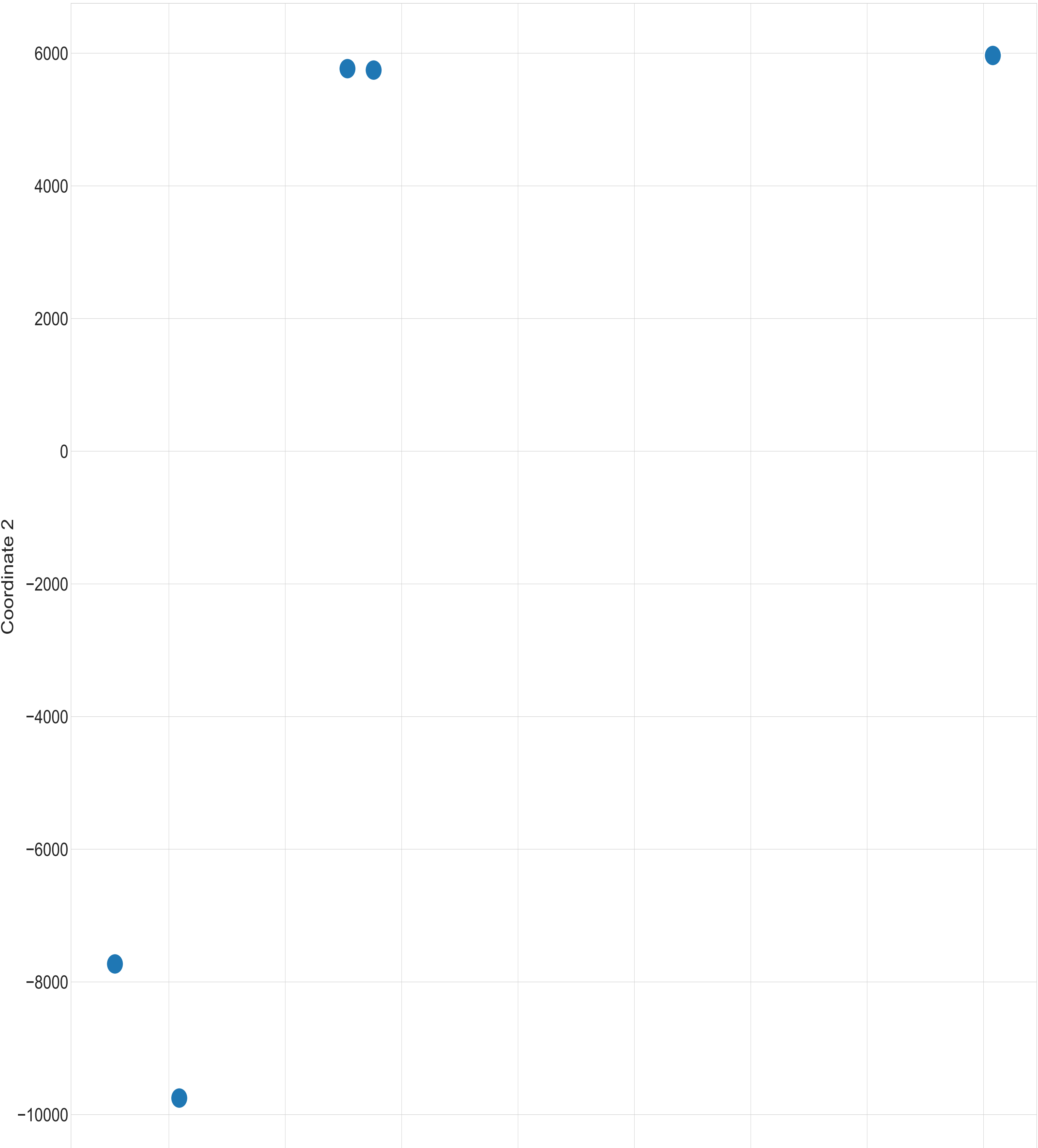
Scatter plot of Saccharides *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cityblock pairwise distance.
Only Saccharides BGCs were considered for this analysis. Each dot represents an individual BGC.
Distance between Saccharides BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S88 :



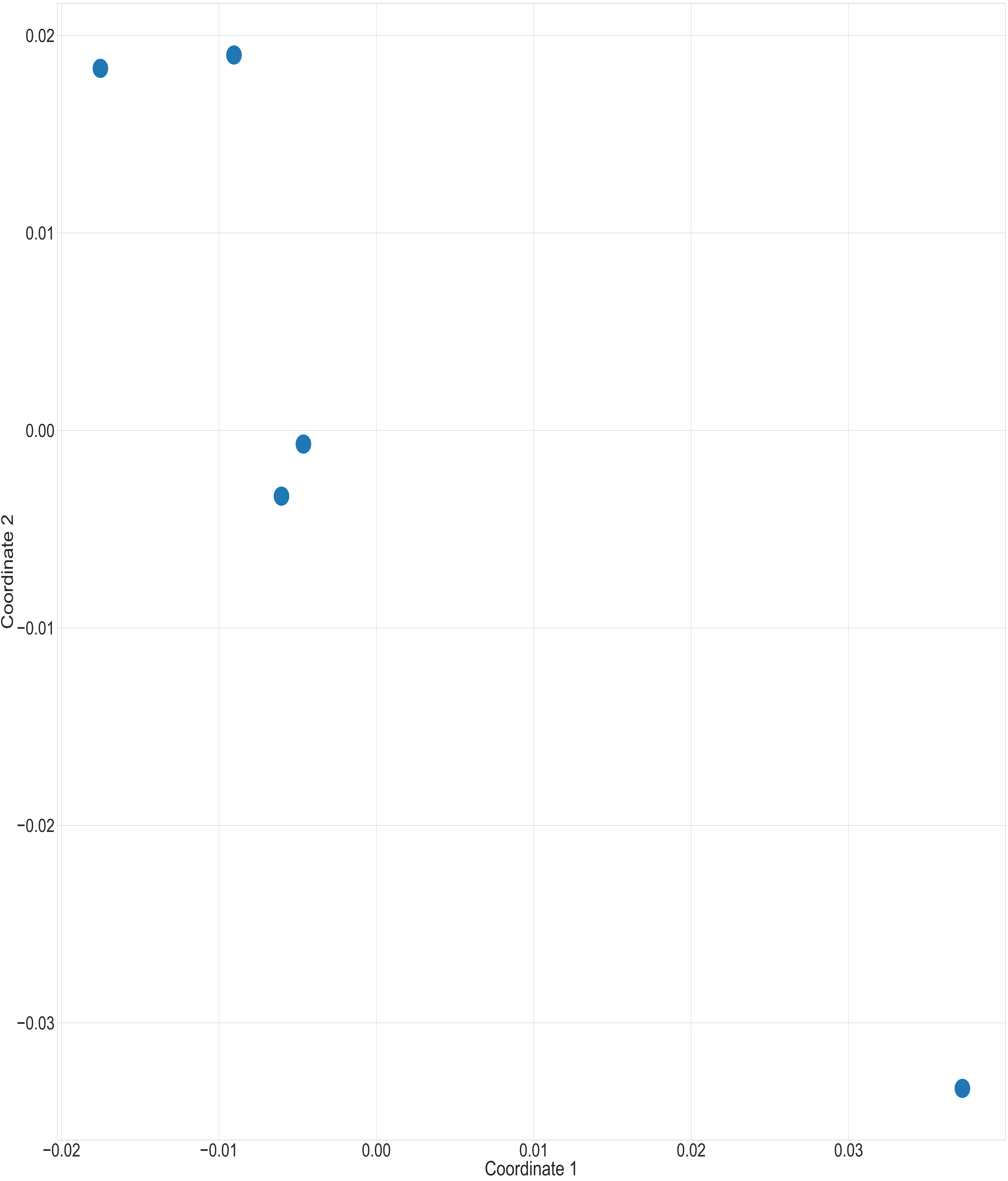
Scatter plot of Saccharides *Micromonosporaceae* BGCs analyzed via multidimensional scaling using chebyshev pairwise distance.
Only Saccharides BGCs were considered for this analysis. Each dot represents an individual BGC.
Distance between Saccharides BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S89 :



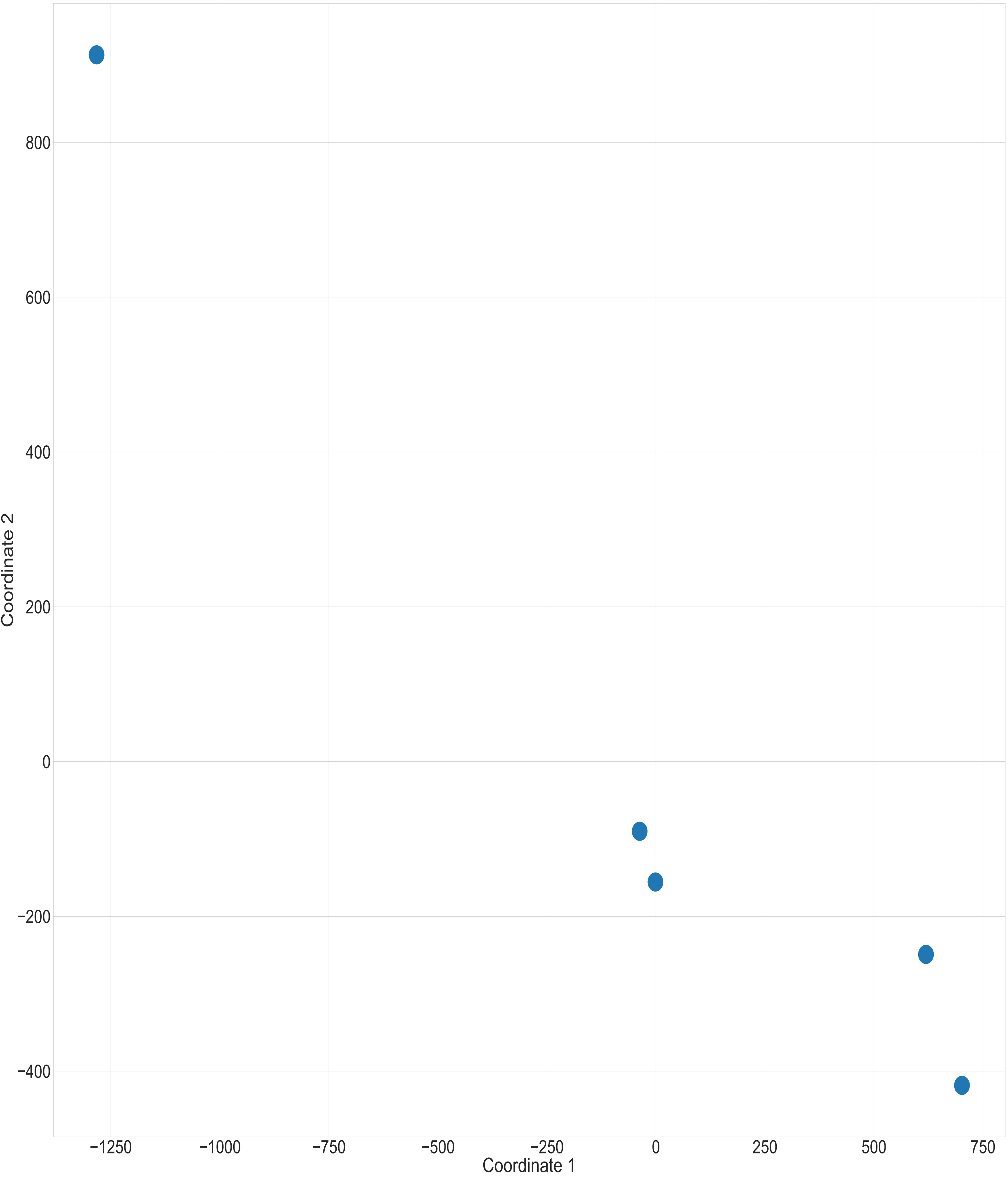
Scatter plot of Saccharides *Micromonosporaceae* BGCs analyzed via multidimensional scaling using l2 pairwise distance. Only Saccharides BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Saccharides BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S90 :



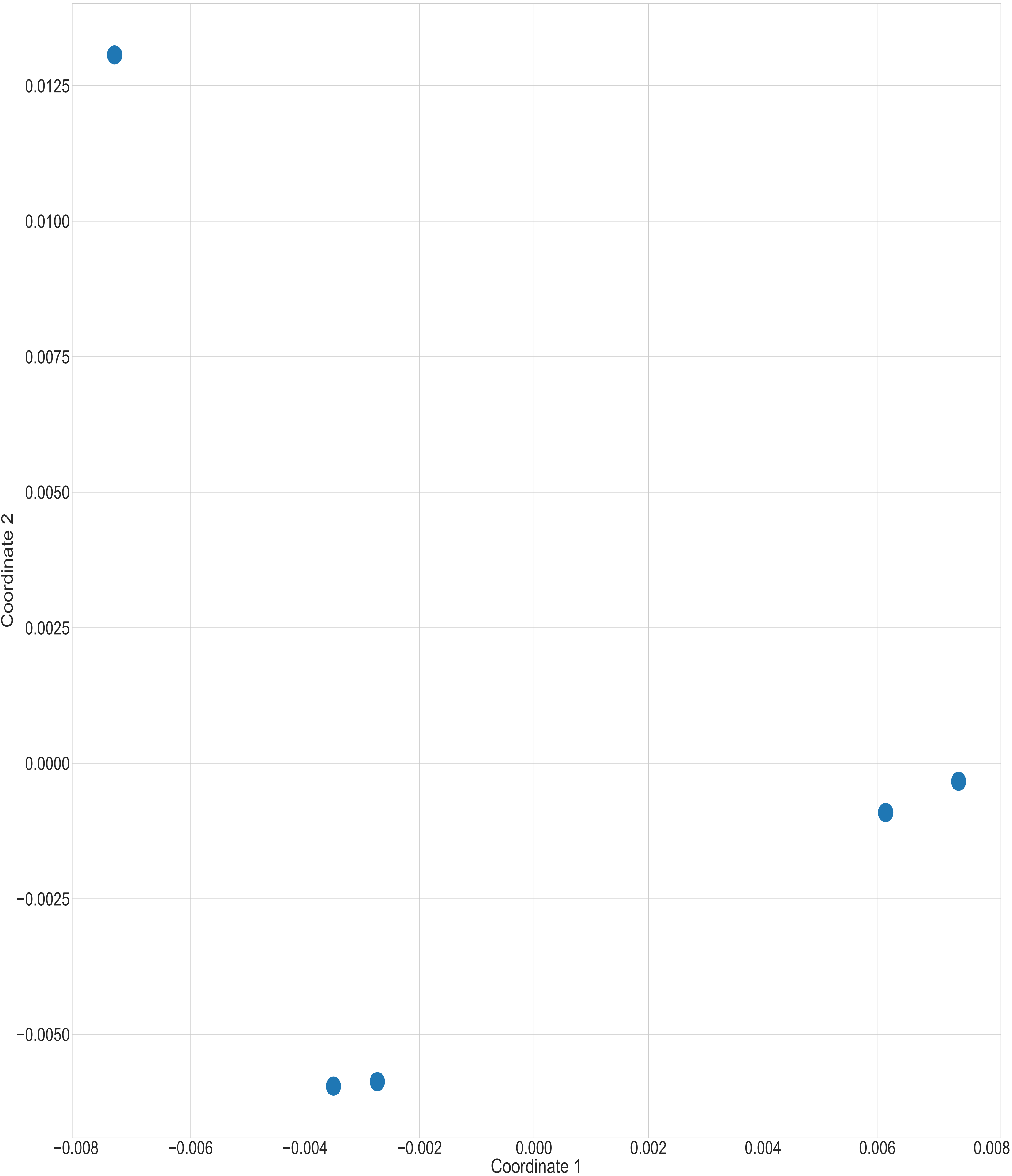
Scatter plot of Saccharides *Micromonosporaceae* BGCs analyzed via multidimensional scaling using braycurtis pairwise distance. Only Saccharides BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Saccharides BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S91 :



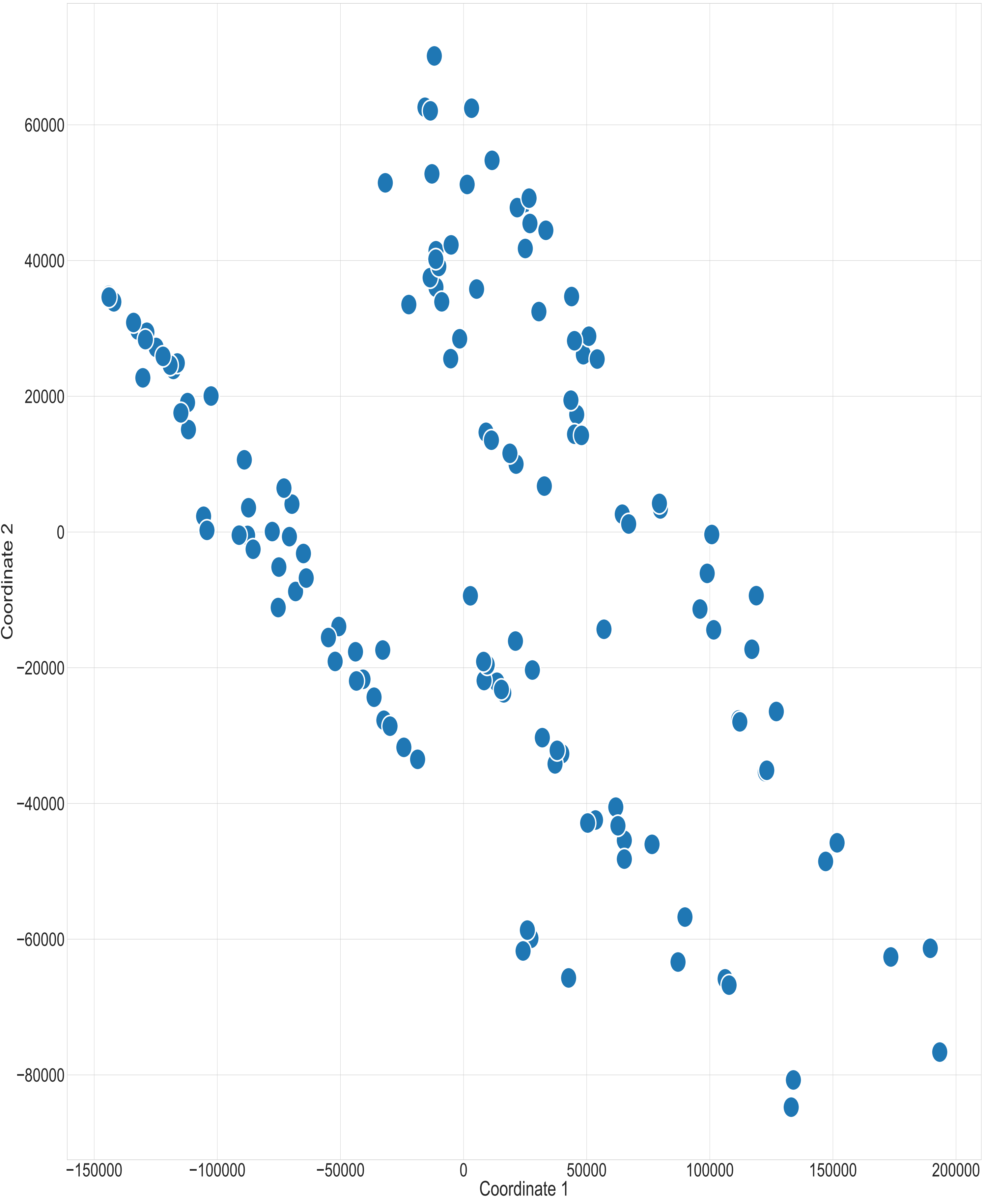
Scatter plot of Saccharides *Micromonosporaceae* BGCs analyzed via multidimensional scaling using canberra pairwise distance.
Only Saccharides BGCs were considered for this analysis. Each dot represents an individual BGC.
Distance between Saccharides BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S92 :



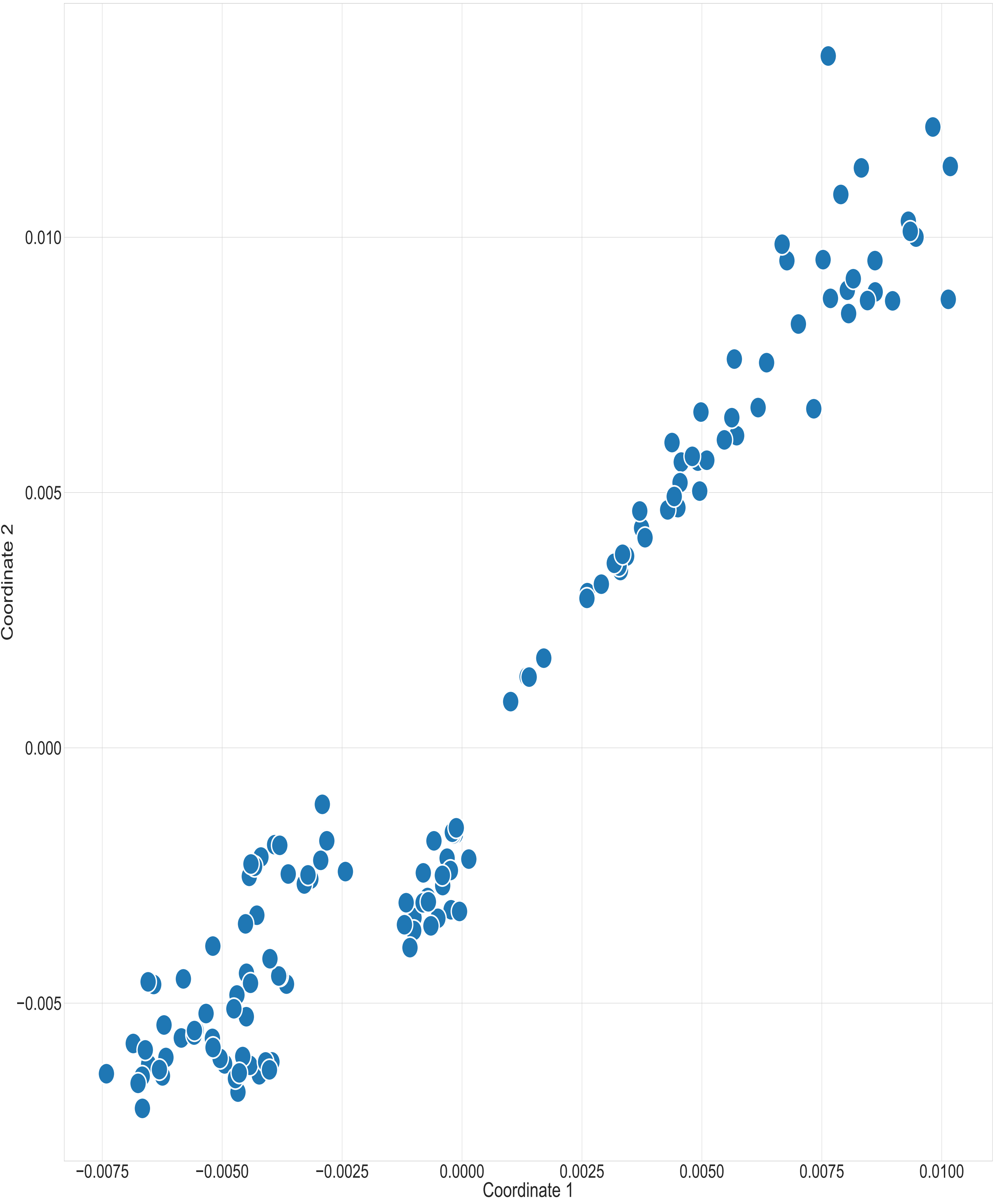
Scatter plot of Saccharides *Micromonosporaceae* BGCs analyzed via multidimensional scaling using correlation pairwise distance. Only Saccharides BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Saccharides BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S93 :



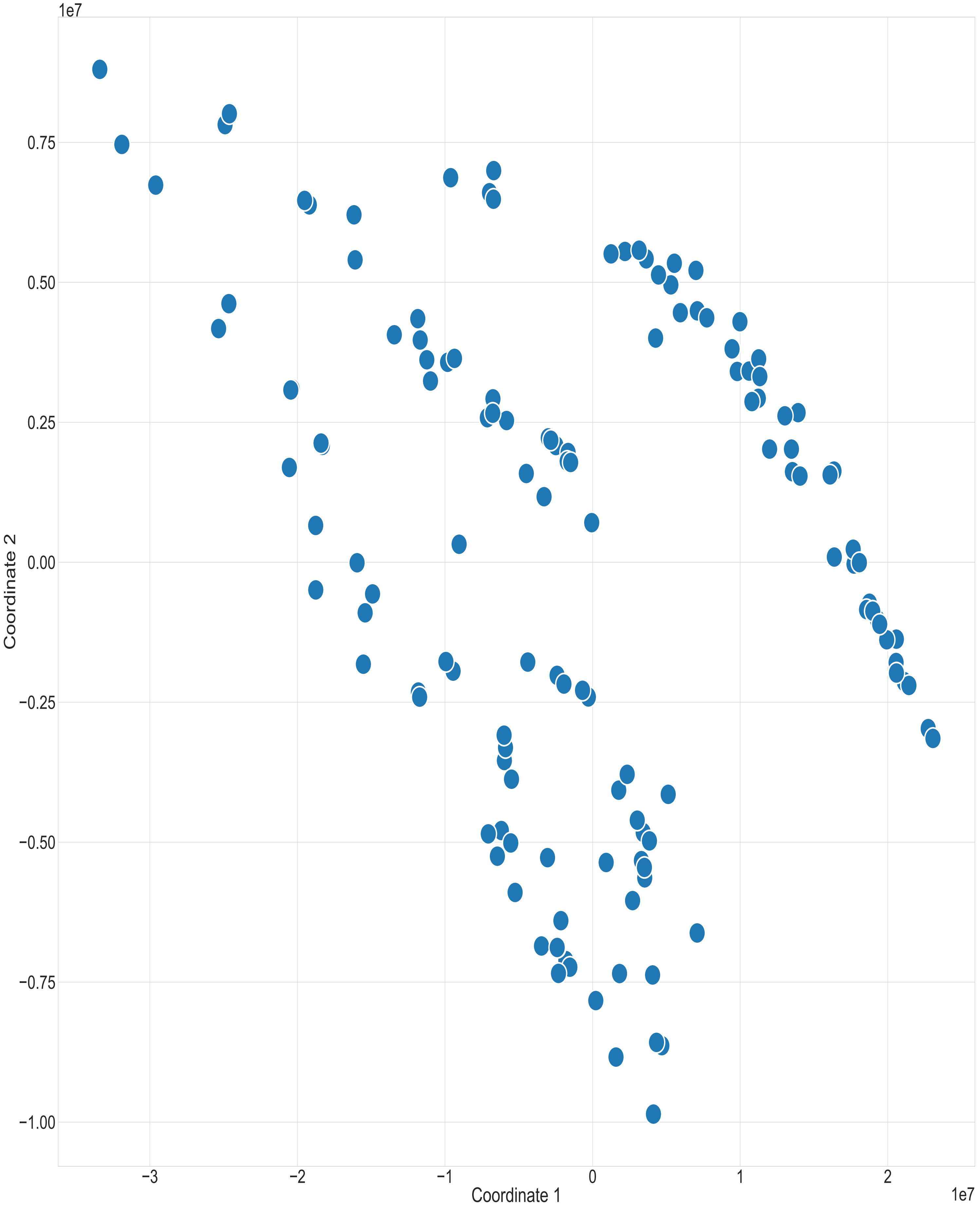
Scatter plot of Hybrid-(Not-PKS-NRPS) *Micromonosporaceae* BGCs analyzed via multidimensional scaling using euclidean pairwise distance. Only Hybrid-(Not-PKS-NRPS) BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Hybrid-(Not-PKS-NRPS) BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S94 :



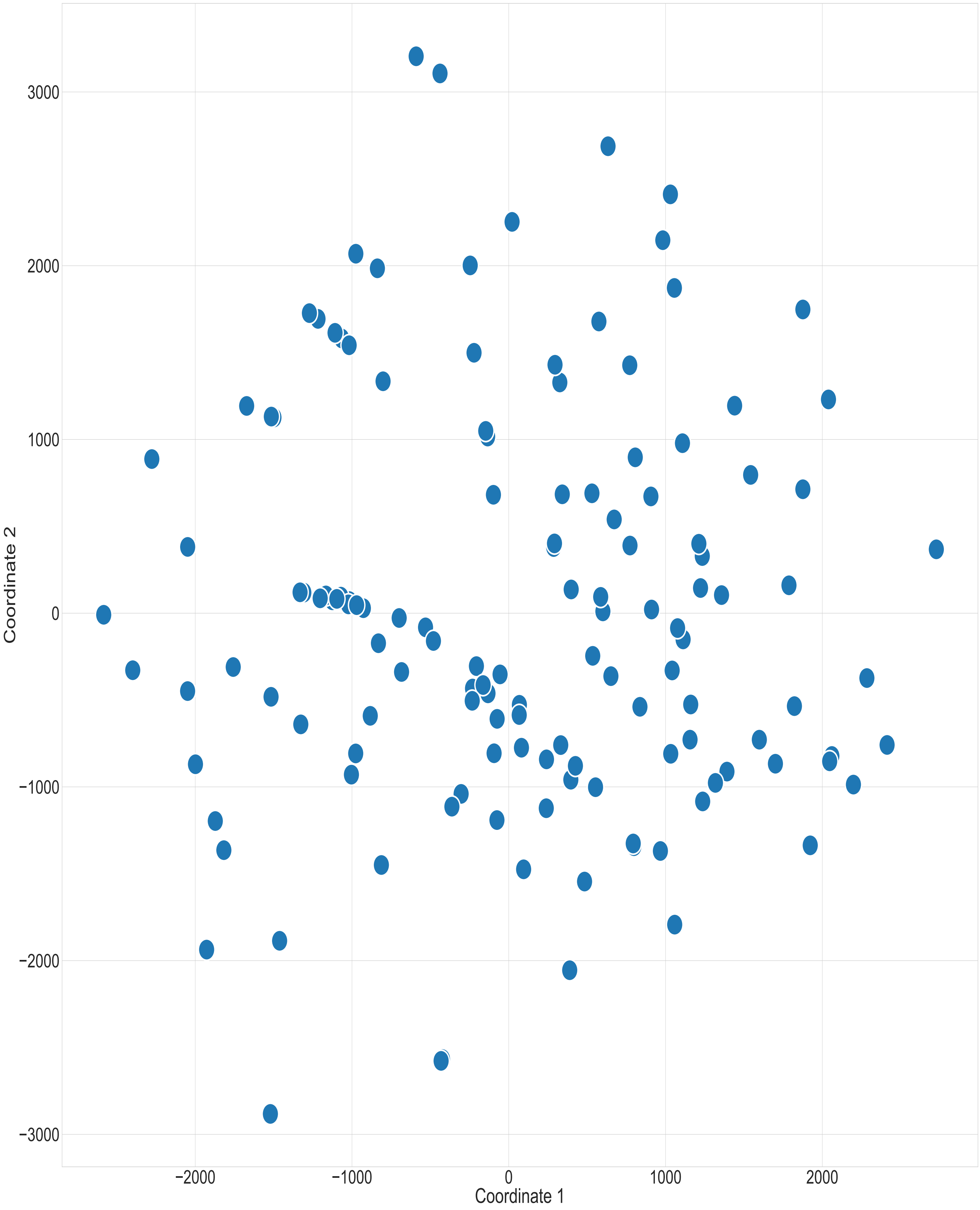
Scatter plot of Hybrid-(Not-PKS-NRPS) *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cosine pairwise distance.
Only Hybrid-(Not-PKS-NRPS) BGCs were considered for this analysis. Each dot represents an individual BGC.
Distance between Hybrid-(Not-PKS-NRPS) BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S95 :



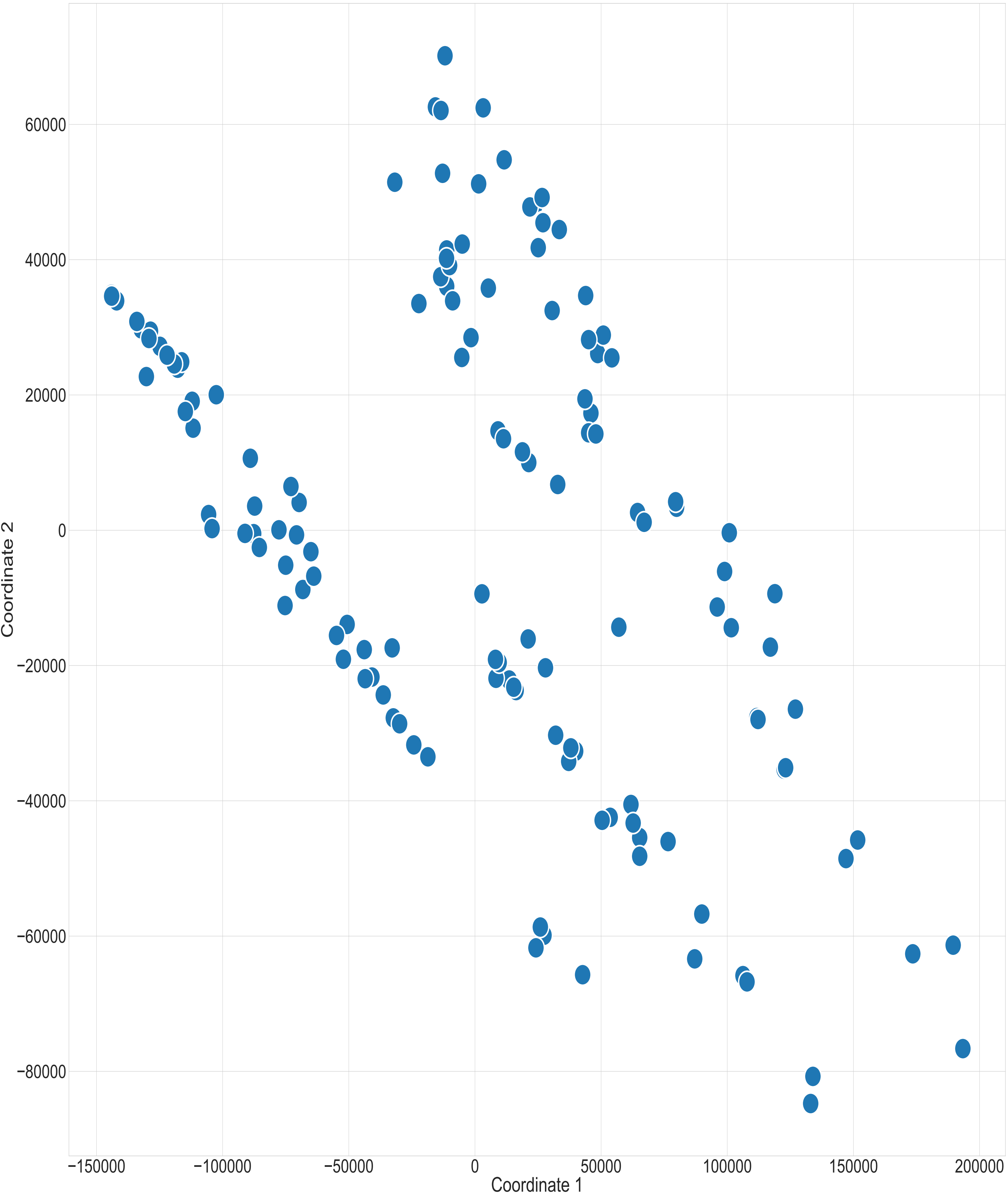
Scatter plot of Hybrid-(Not-PKS-NRPS) *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cityblock pairwise distance. Only Hybrid-(Not-PKS-NRPS) BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Hybrid-(Not-PKS-NRPS) BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S96 :



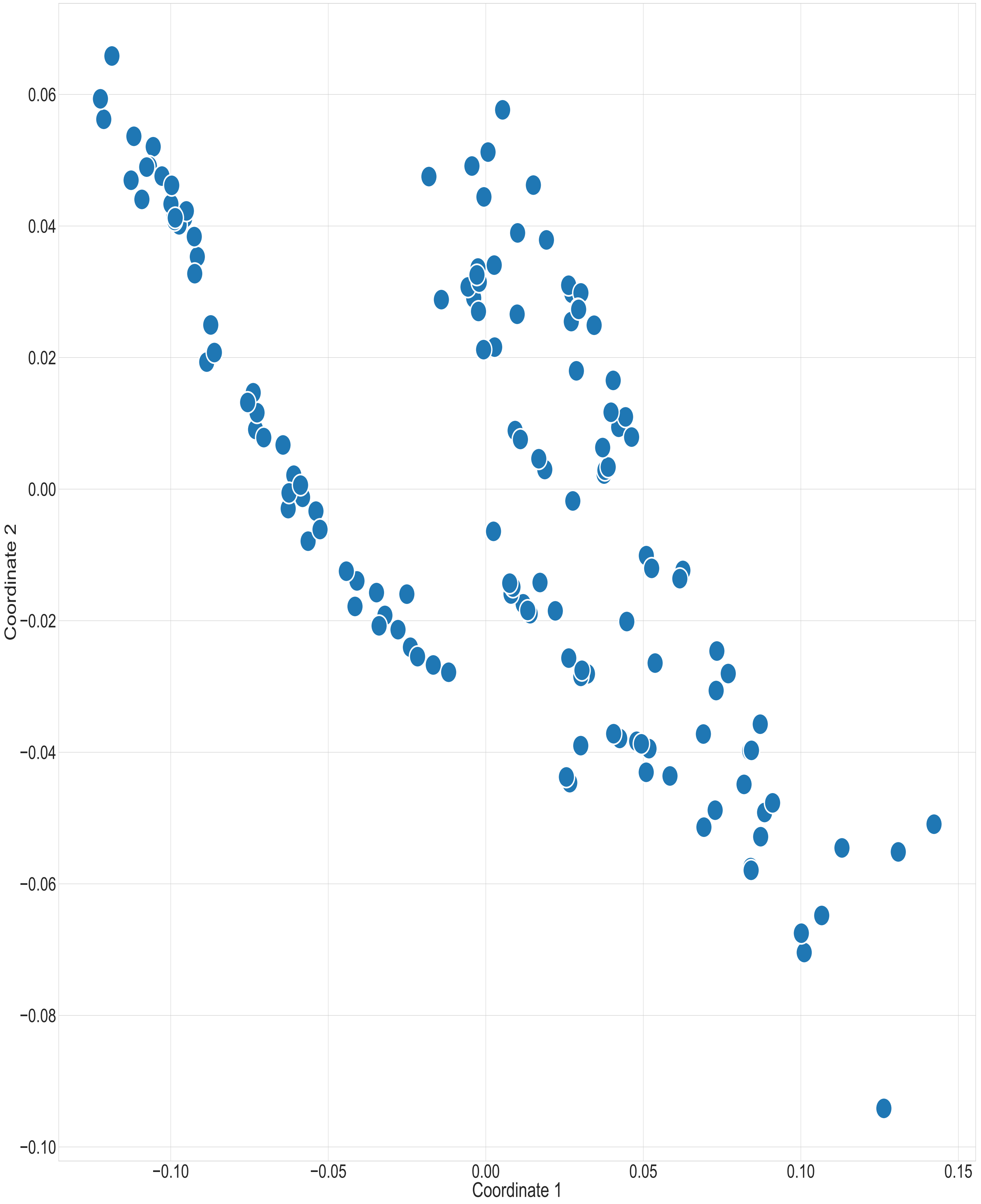
Scatter plot of Hybrid-(Not-PKS-NRPS) *Micromonosporaceae* BGCs analyzed via multidimensional scaling using chebyshev pairwise distance.
Only Hybrid-(Not-PKS-NRPS) BGCs were considered for this analysis. Each dot represents an individual BGC.
Distance between Hybrid-(Not-PKS-NRPS) BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S97 :



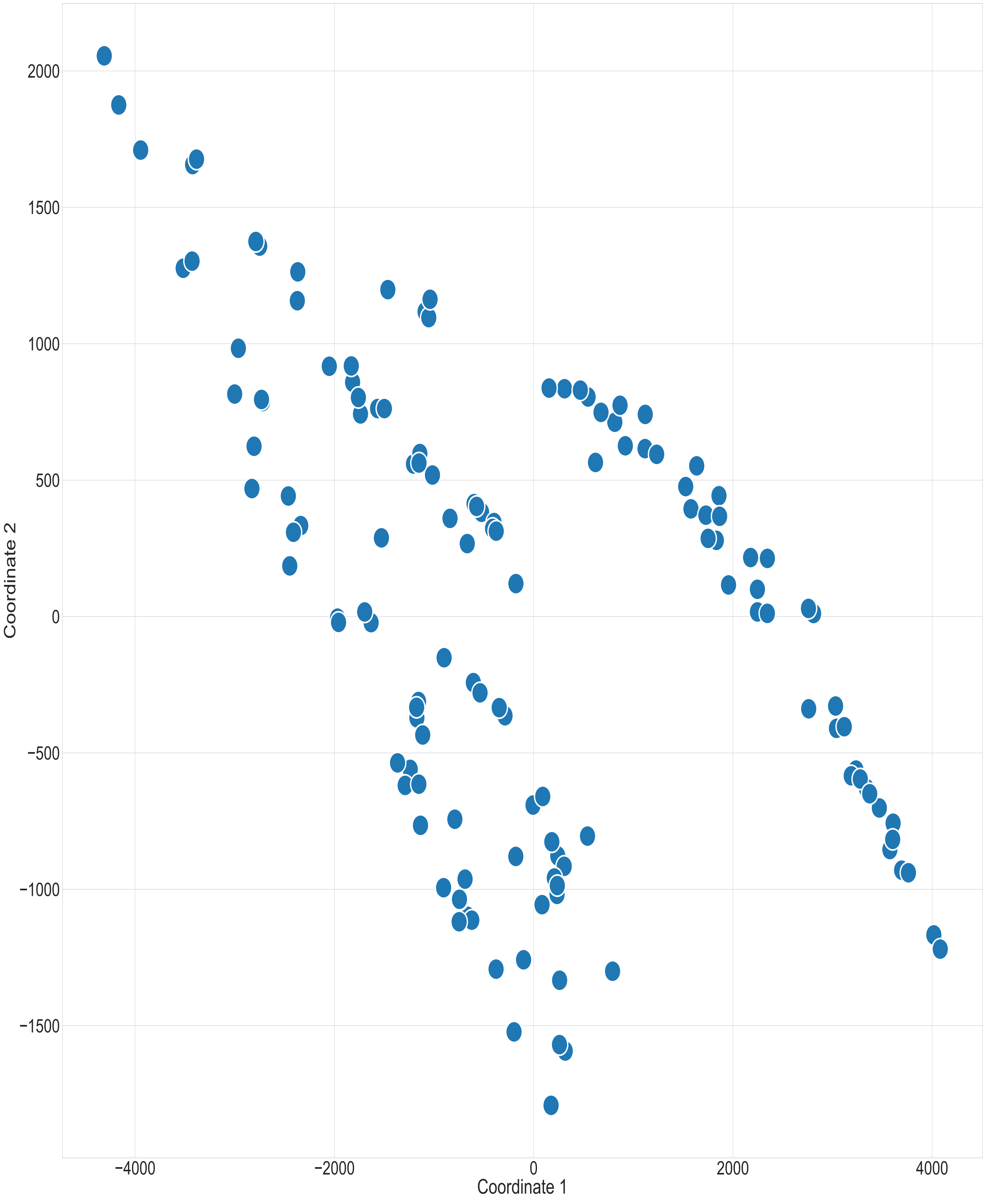
Scatter plot of Hybrid-(Not-PKS-NRPS) *Micromonosporaceae* BGCs analyzed via multidimensional scaling using l2 pairwise distance. Only Hybrid-(Not-PKS-NRPS) BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Hybrid-(Not-PKS-NRPS) BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S98 :



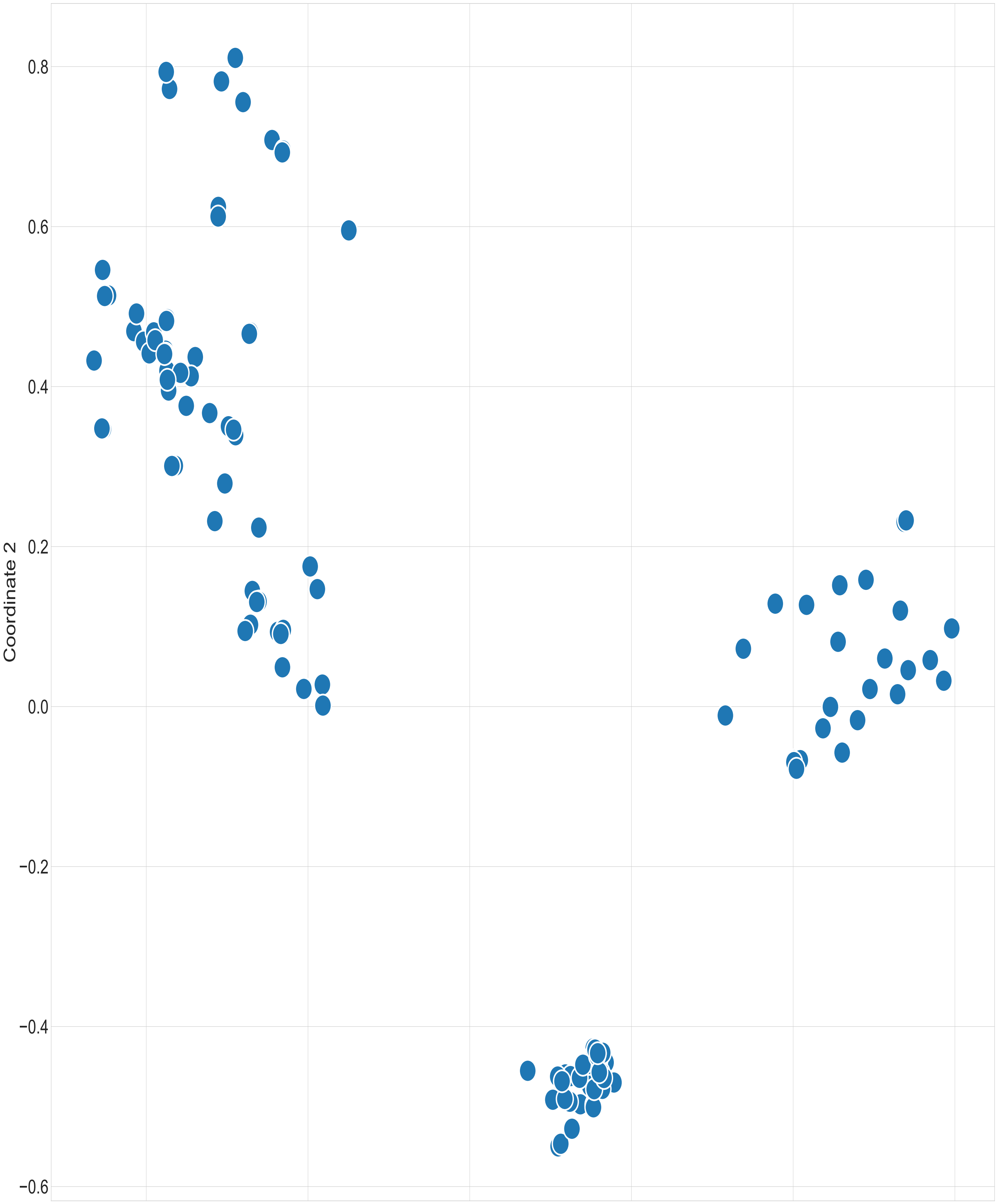
Scatter plot of Hybrid-(Not-PKS-NRPS) *Micromonosporaceae* BGCs analyzed via multidimensional scaling using braycurtis pairwise distance. Only Hybrid-(Not-PKS-NRPS) BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Hybrid-(Not-PKS-NRPS) BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S99 :



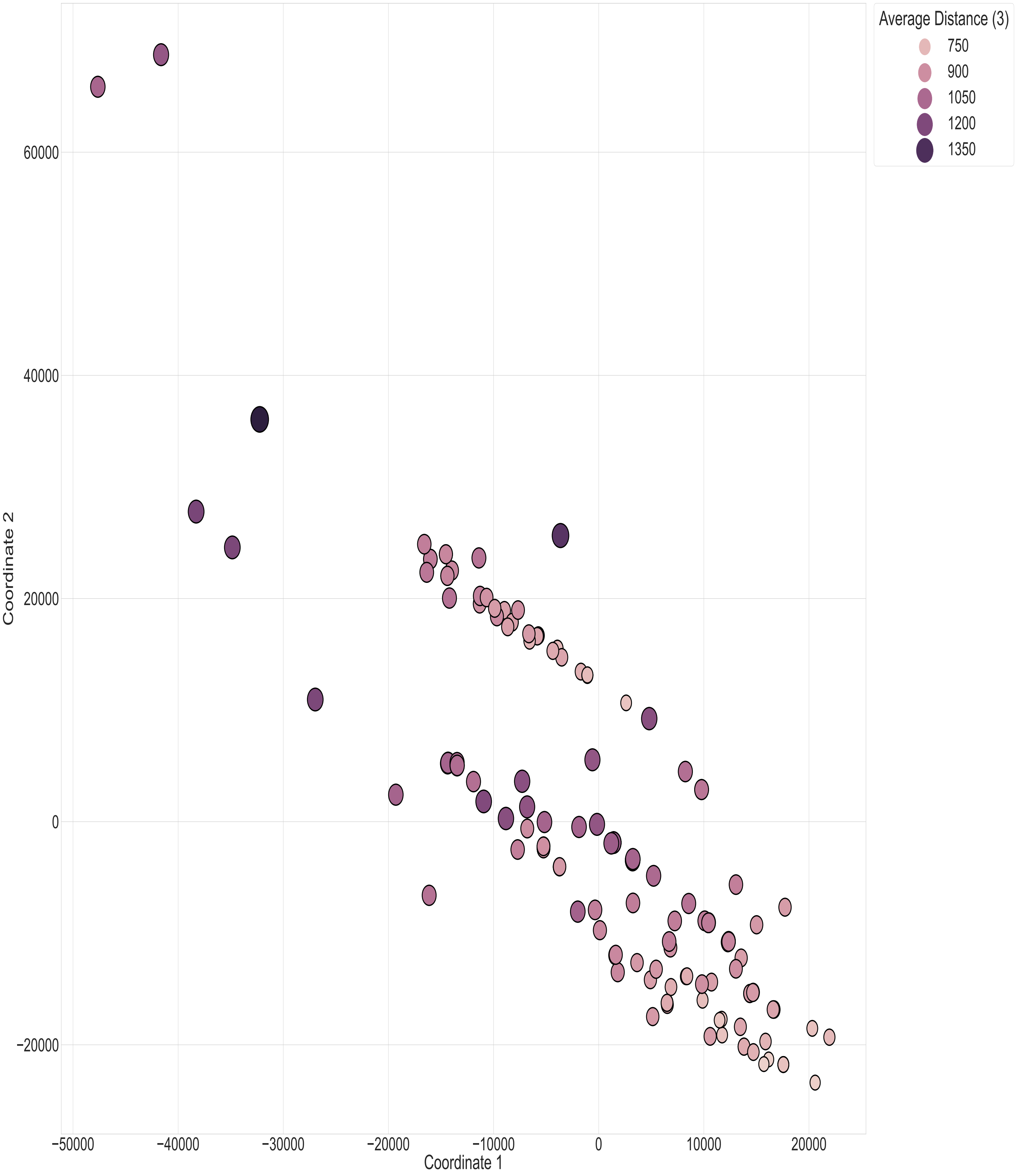
Scatter plot of Hybrid-(Not-PKS-NRPS) *Micromonosporaceae* BGCs analyzed via multidimensional scaling using canberra pairwise distance. Only Hybrid-(Not-PKS-NRPS) BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Hybrid-(Not-PKS-NRPS) BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S100 :



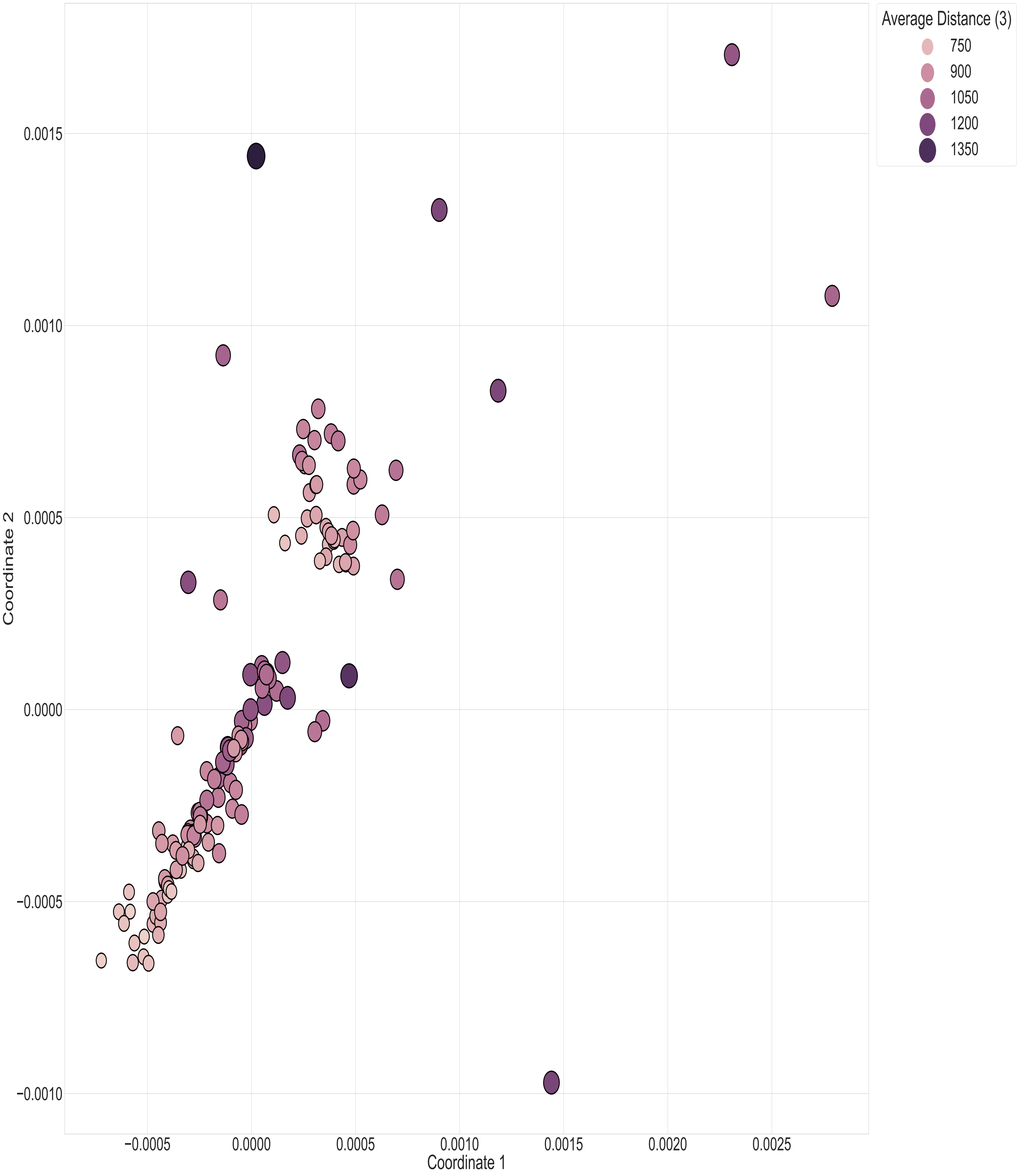
Scatter plot of Hybrid-(Not-PKS-NRPS) *Micromonosporaceae* BGCs analyzed via multidimensional scaling using correlation pairwise distance.
Only Hybrid-(Not-PKS-NRPS) BGCs were considered for this analysis. Each dot represents an individual BGC.
Distance between Hybrid-(Not-PKS-NRPS) BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S101 :



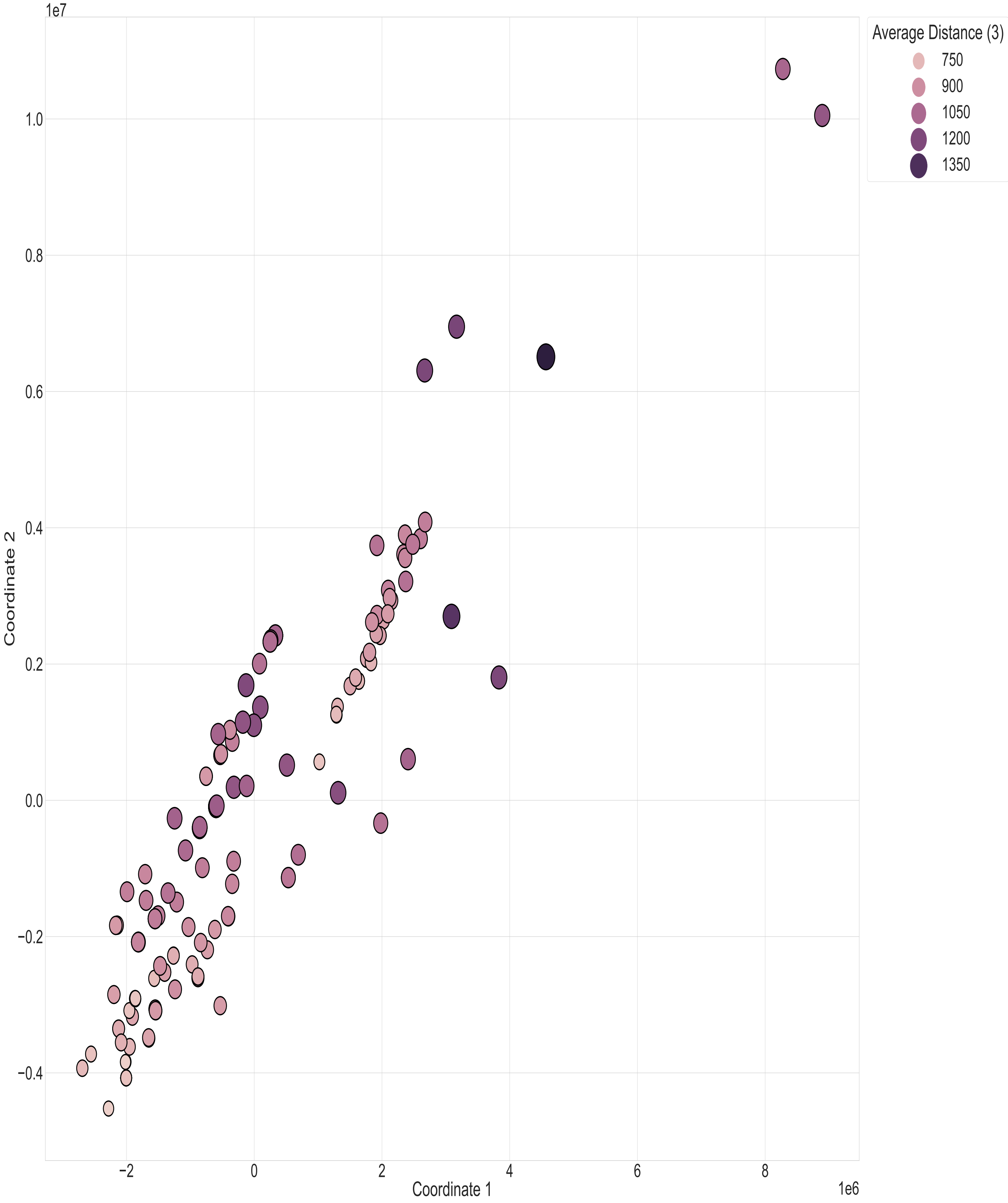
Scatter plot of Terpene *Micromonosporaceae* BGCs analyzed via multidimensional scaling using euclidean pairwise distance.
Only Terpene BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual Terpene BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S102 :



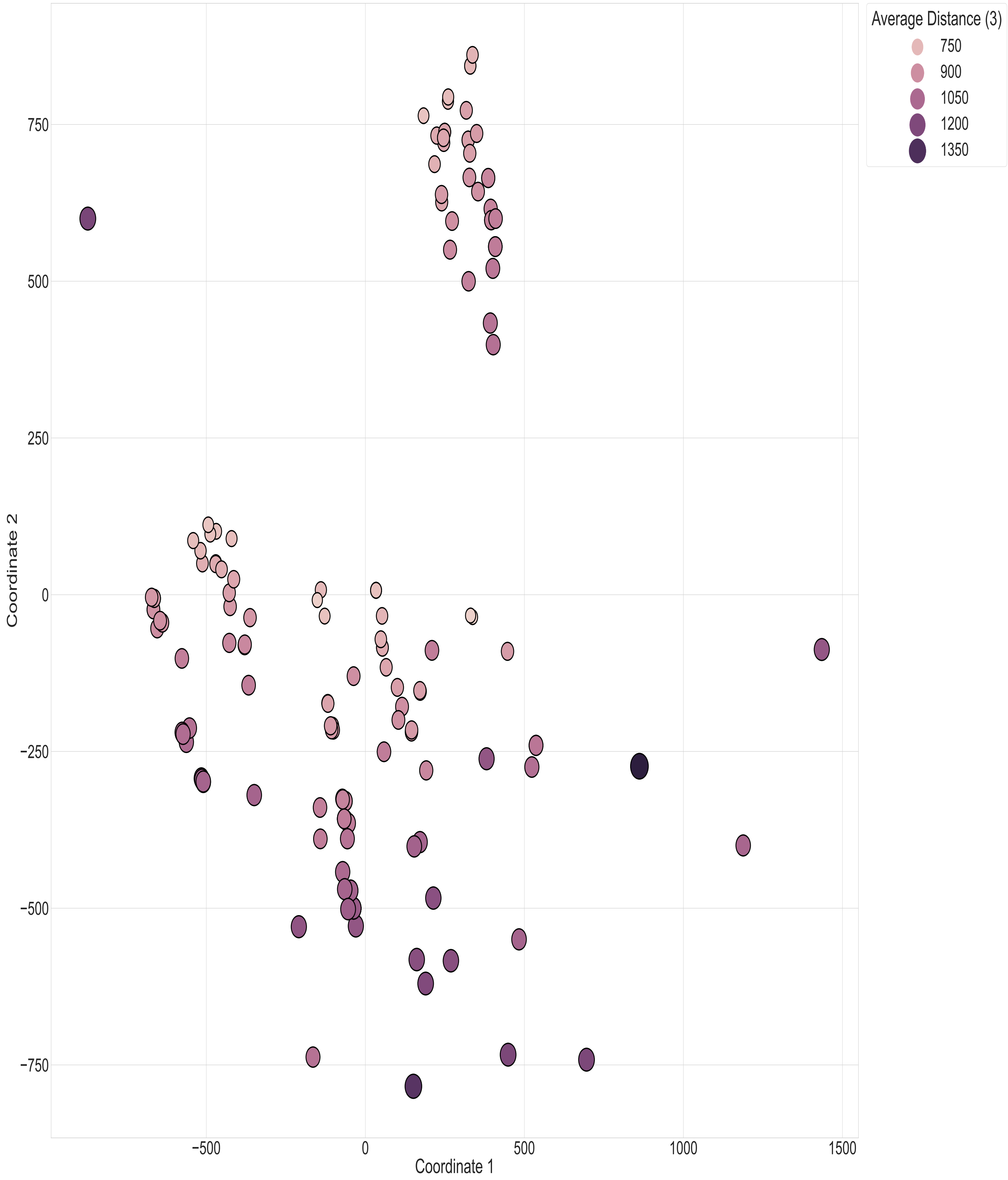
Scatter plot of Terpene *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cosine pairwise distance.
Only Terpene BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual Terpene BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S103 :



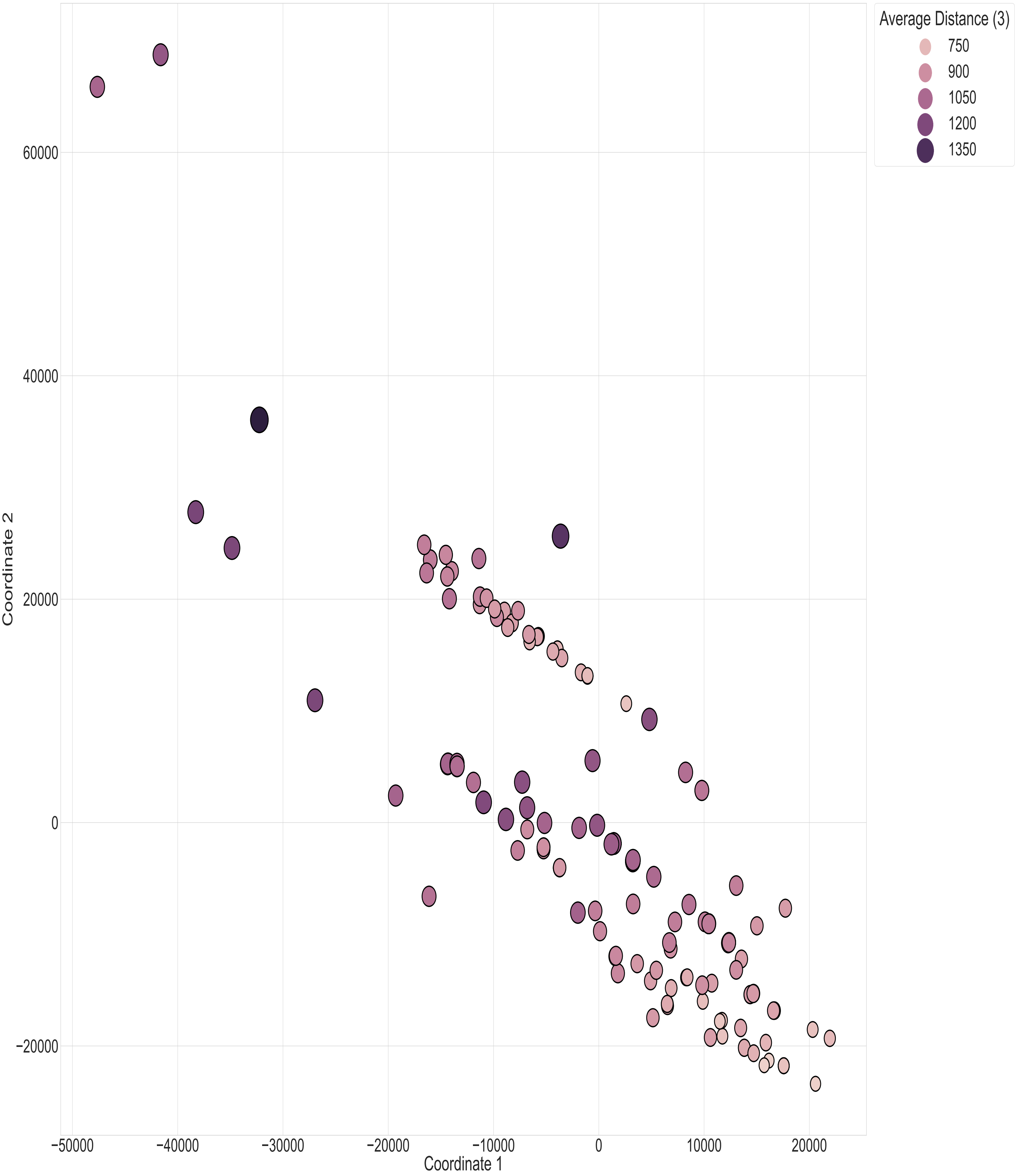
Scatter plot of Terpene *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cityblock pairwise distance.
Only Terpene BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual Terpene BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S104 :



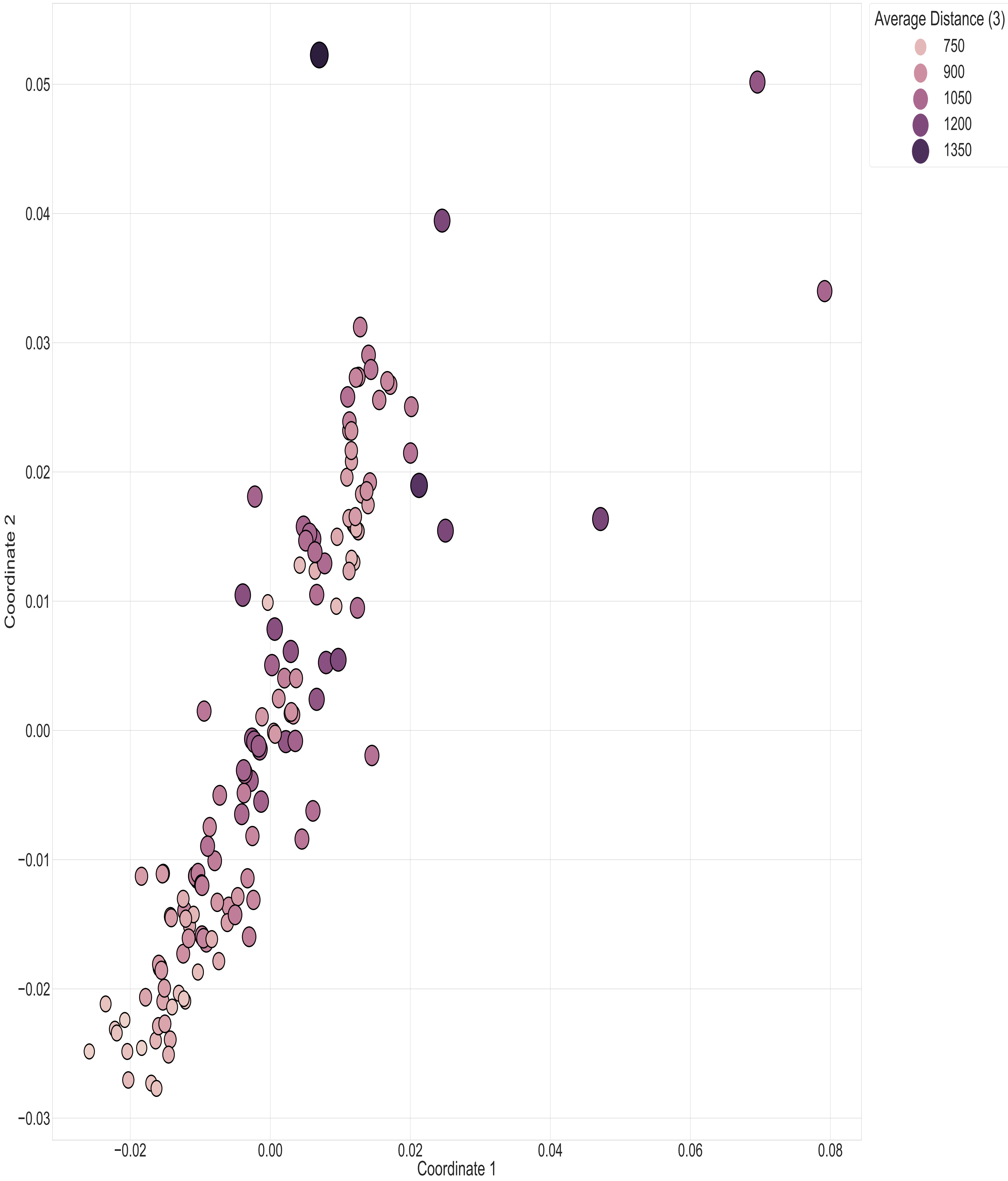
Scatter plot of Terpene *Micromonosporaceae* BGCs analyzed via multidimensional scaling using chebyshev pairwise distance.
Only Terpene BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual Terpene BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S105 :



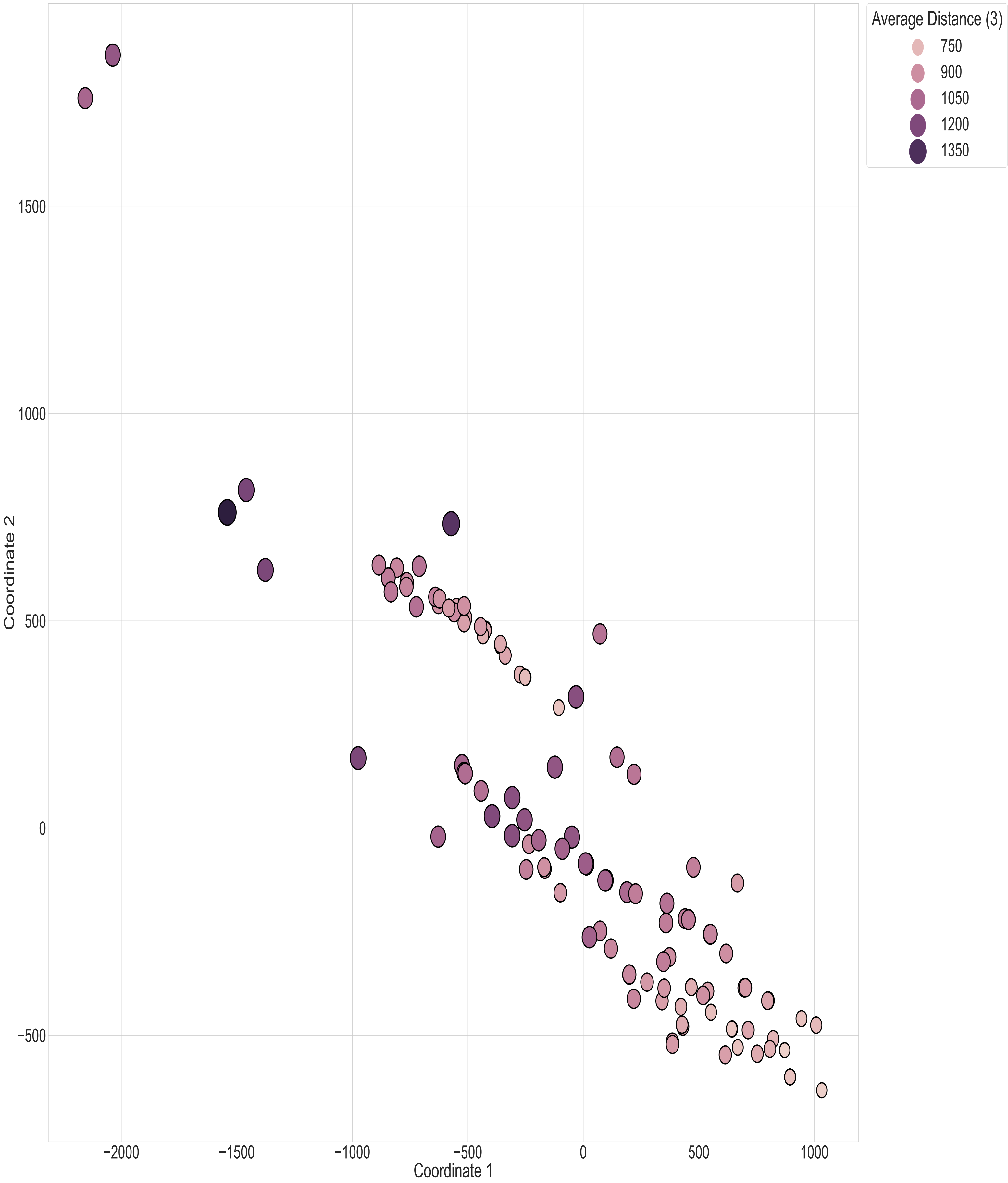
Scatter plot of Terpene *Micromonosporaceae* BGCs analyzed via multidimensional scaling using l2 pairwise distance.
Only Terpene BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual Terpene BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S106 :



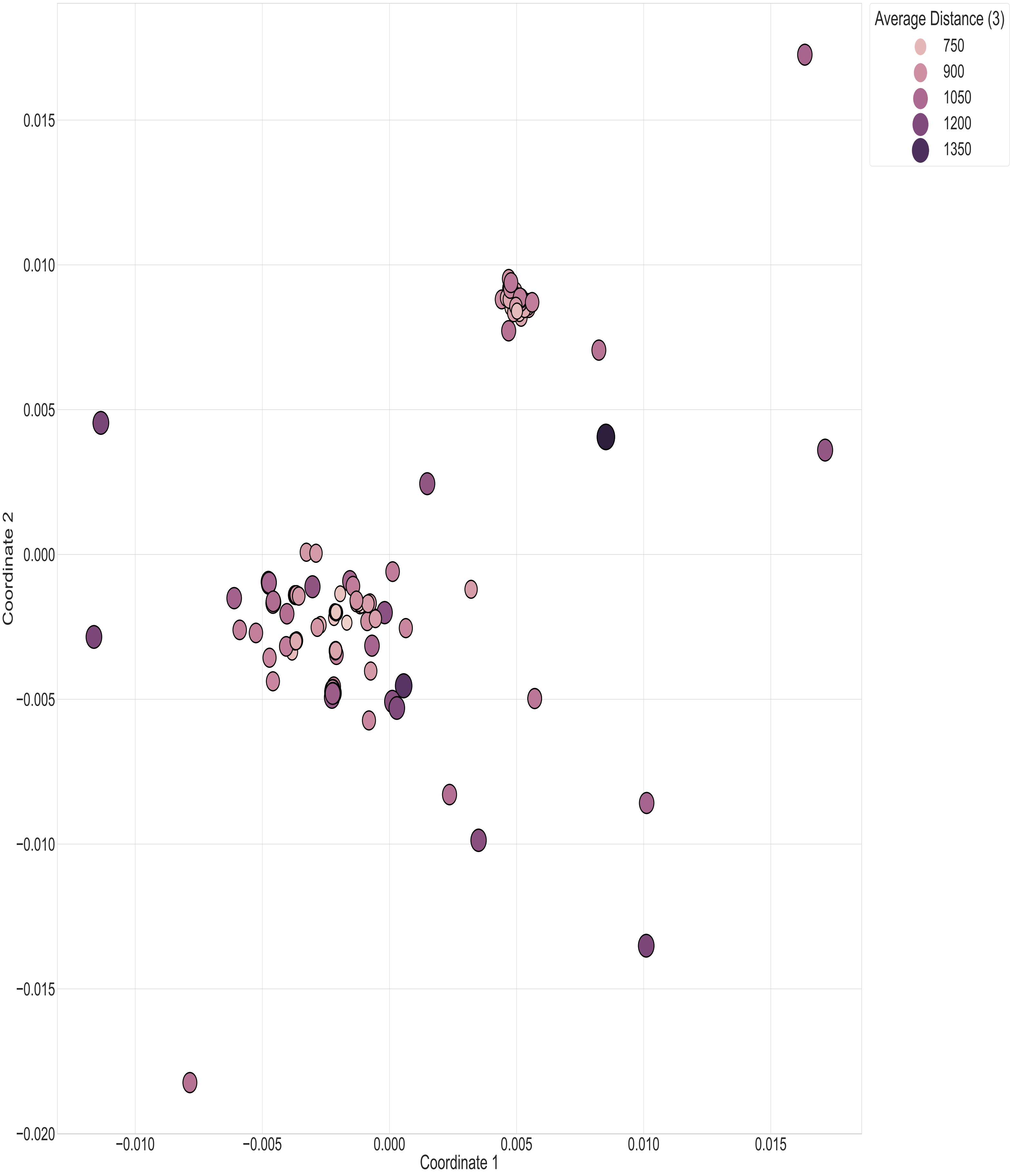
Scatter plot of Terpene *Micromonosporaceae* BGCs analyzed via multidimensional scaling using braycurtis pairwise distance.
Only Terpene BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual Terpene BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S107 :



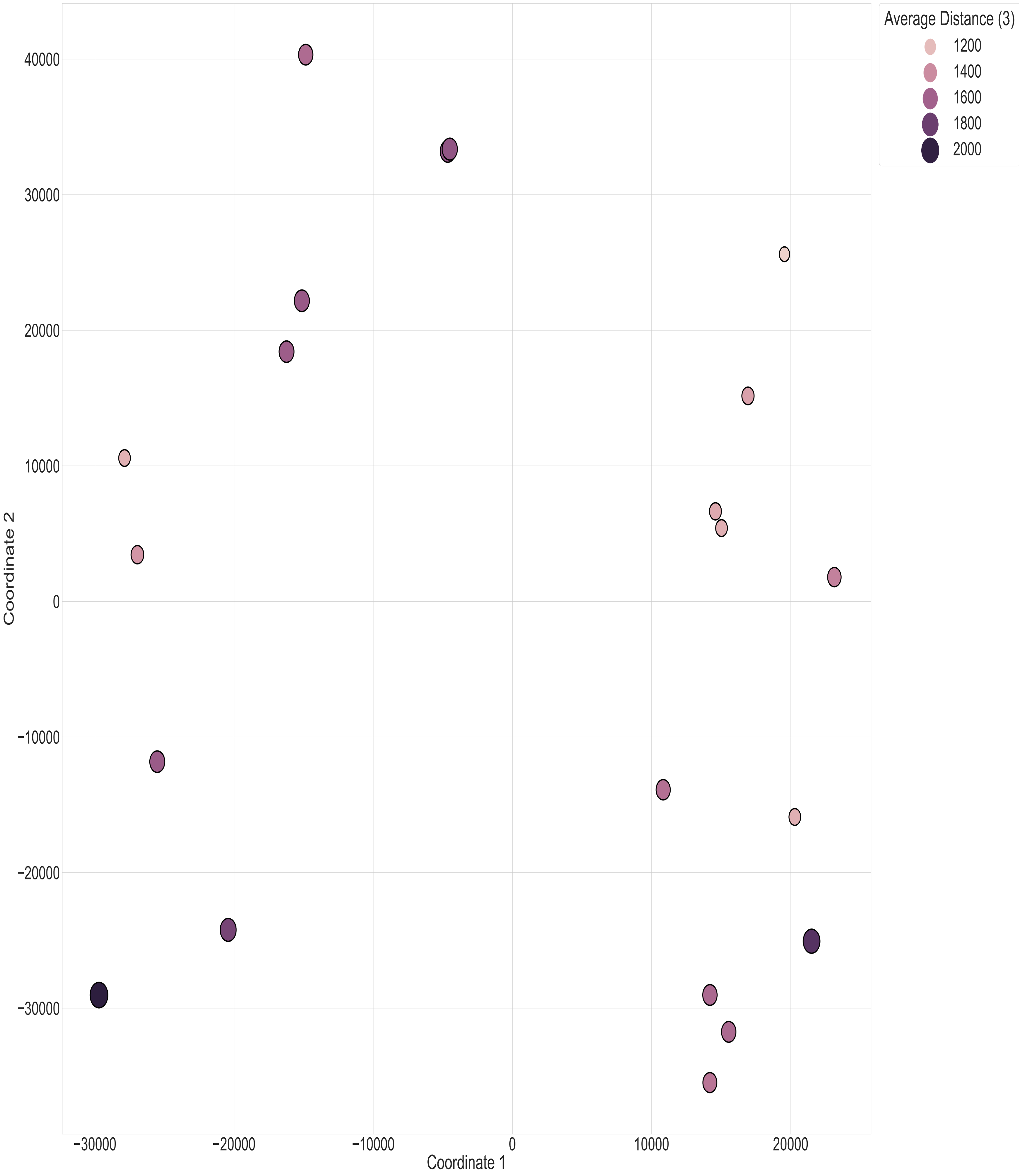
Scatter plot of Terpene *Micromonosporaceae* BGCs analyzed via multidimensional scaling using canberra pairwise distance.
Only Terpene BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual Terpene BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S108 :



Scatter plot of Terpene *Micromonosporaceae* BGCs analyzed via multidimensional scaling using correlation pairwise distance.
Only Terpene BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual Terpene BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S109 :

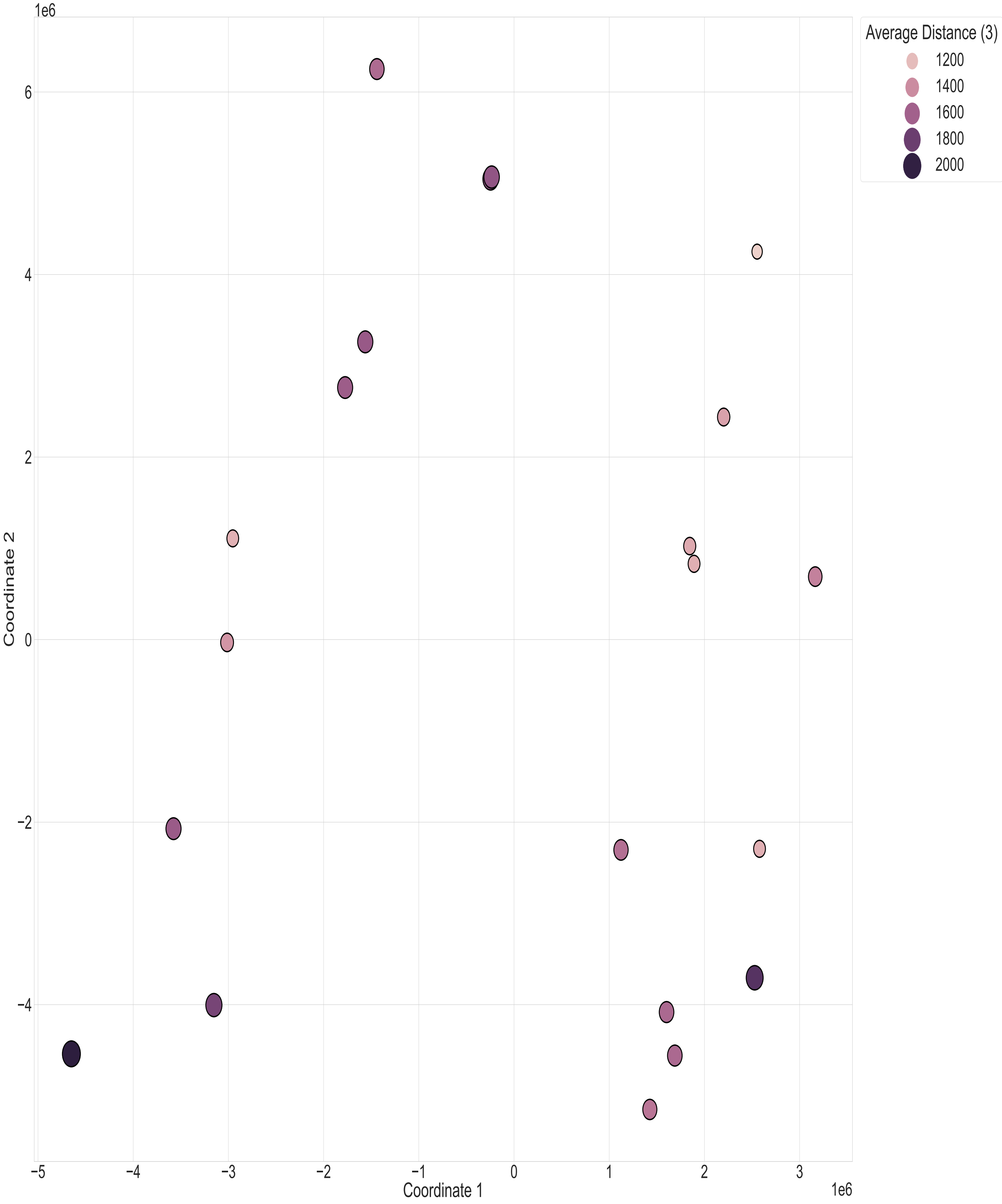


Scatter plot of PKS-1 *Micromonosporaceae* BGCs analyzed via multidimensional scaling using euclidean pairwise distance.
Only PKS-1 BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual PKS-1 BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S110 :

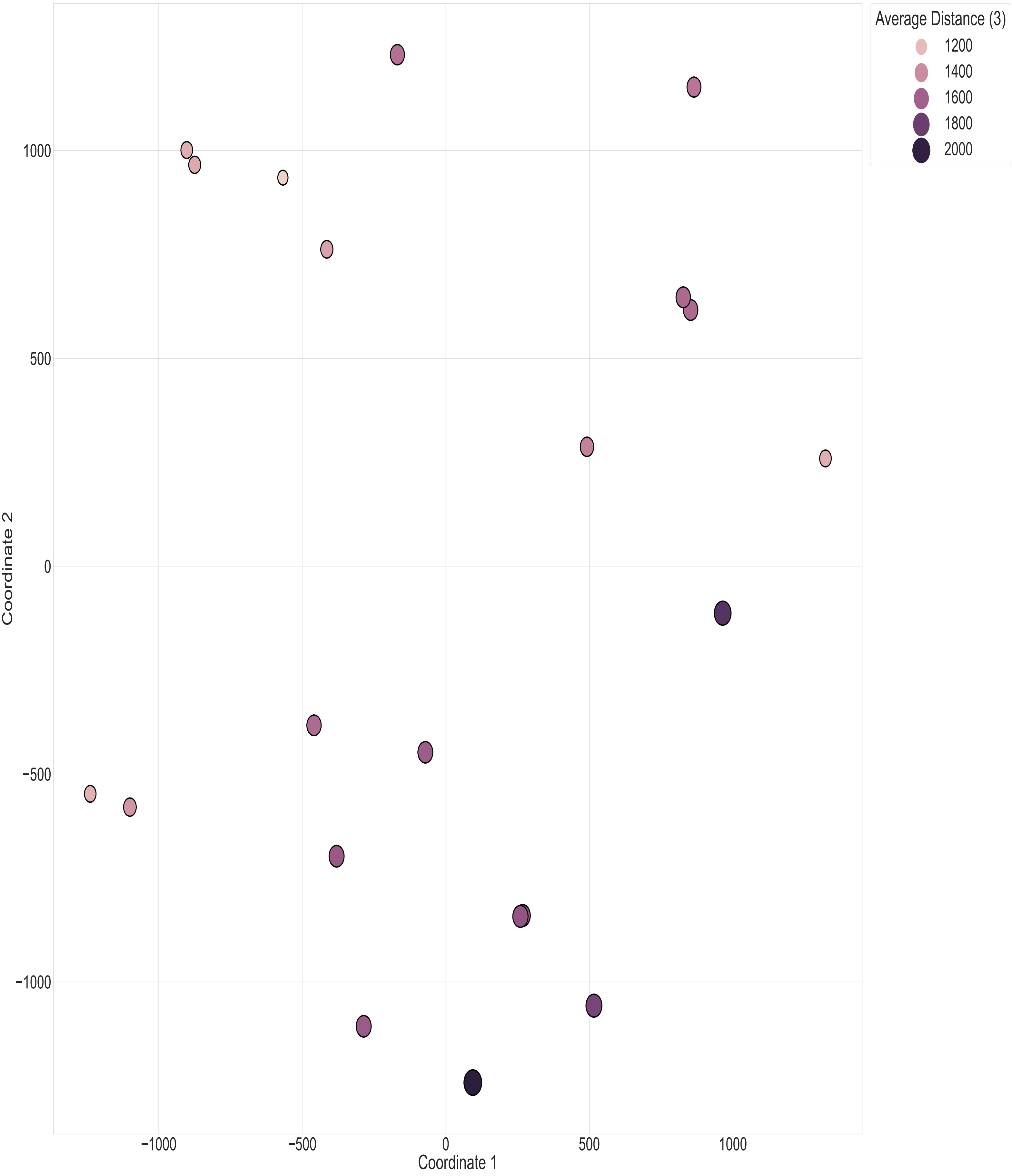
Scatter plot of PKS-1 *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cosine pairwise distance.
Only PKS-1 BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual PKS-1 BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S111 :



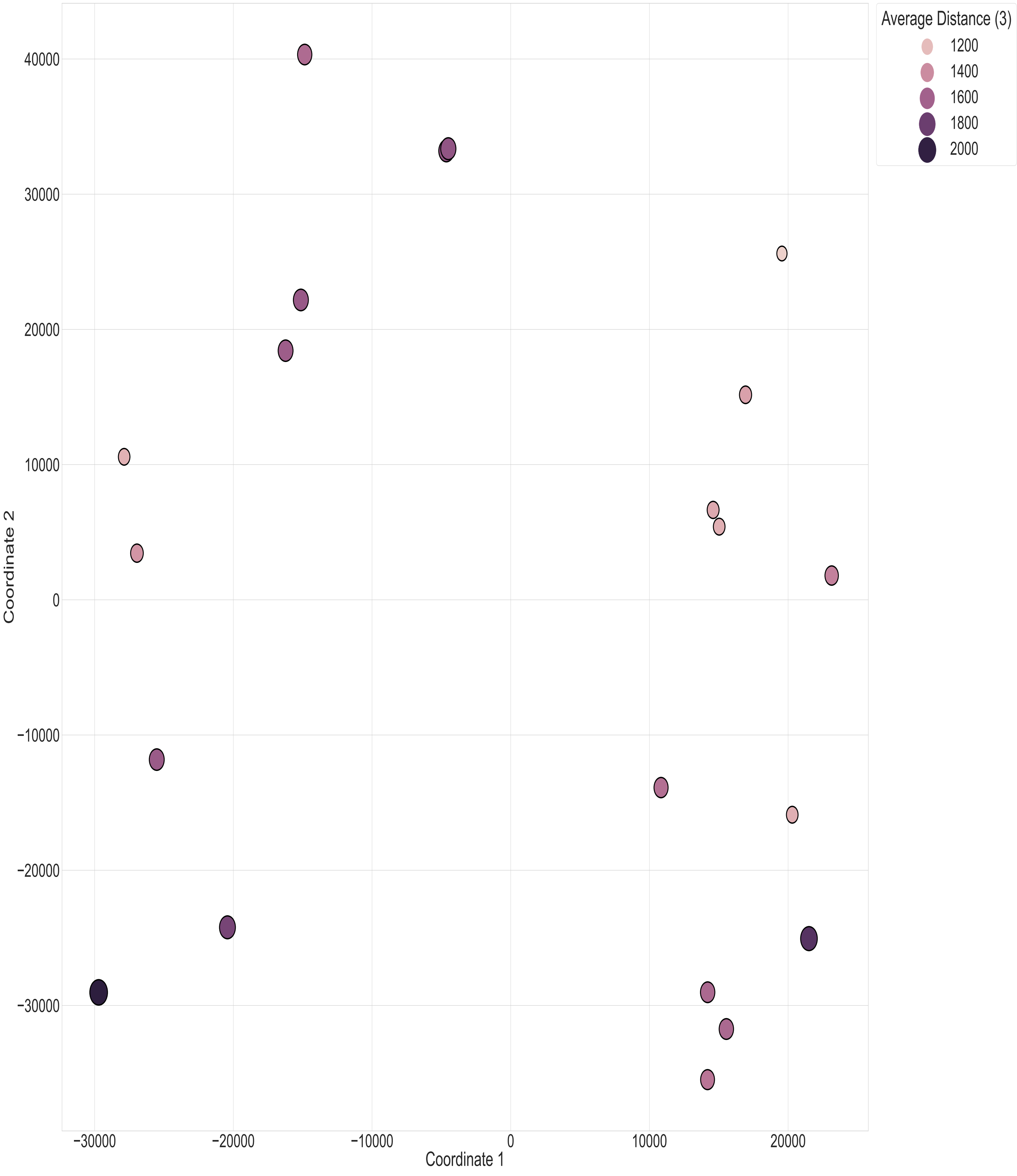
Scatter plot of PKS-1 *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cityblock pairwise distance. Only PKS-1 BGCs were considered for this analysis. Each dot represents an individual BGC. The hue and size of the dots were scaled based on the average distance of the individual PKS-1 BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S112 :



Scatter plot of PKS-1 *Micromonosporaceae* BGCs analyzed via multidimensional scaling using chebyshev pairwise distance.
Only PKS-1 BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual PKS-1 BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S113 :

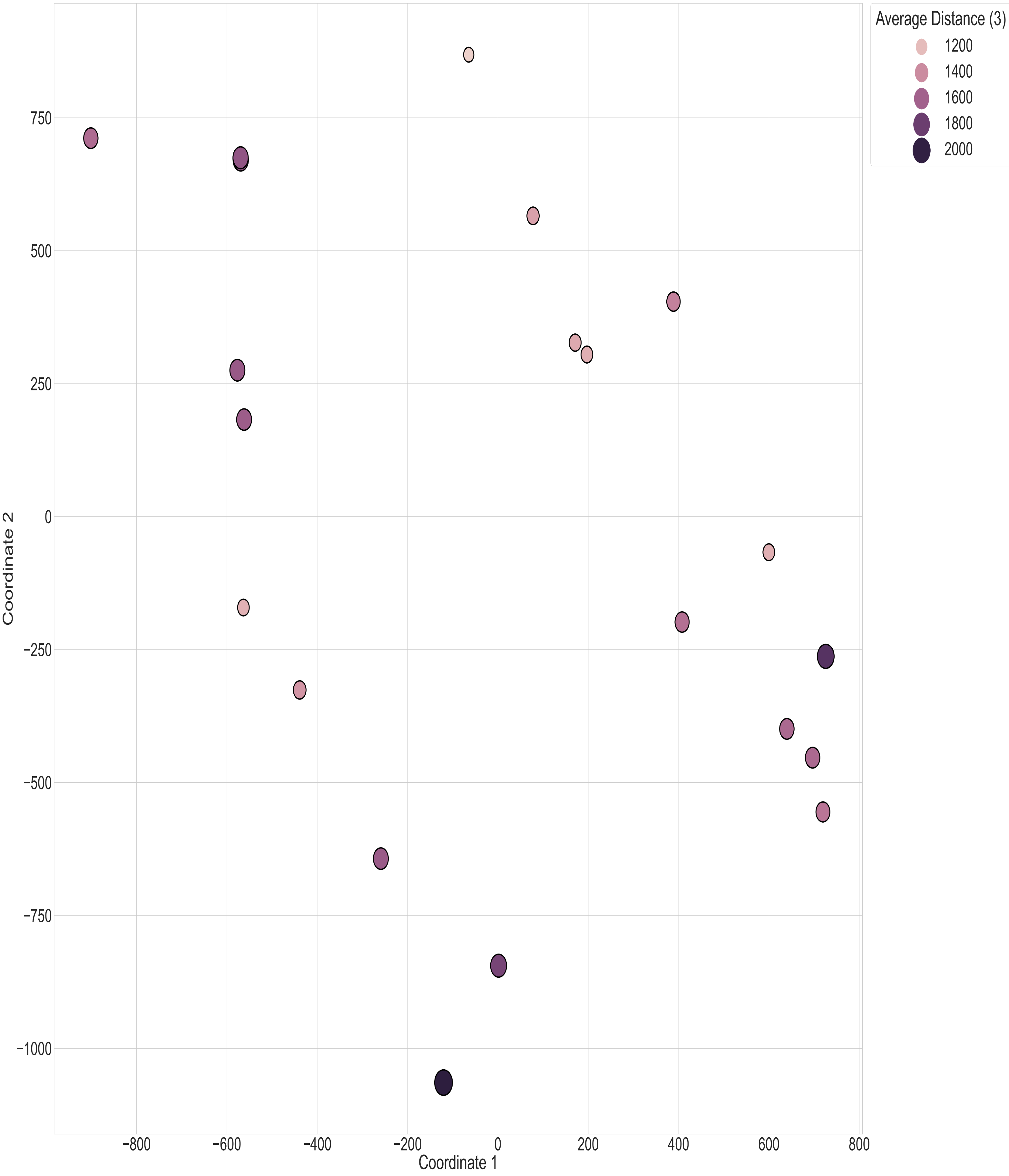


Scatter plot of PKS-1 *Micromonosporaceae* BGCs analyzed via multidimensional scaling using l2 pairwise distance.
Only PKS-1 BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual PKS-1 BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S114 :

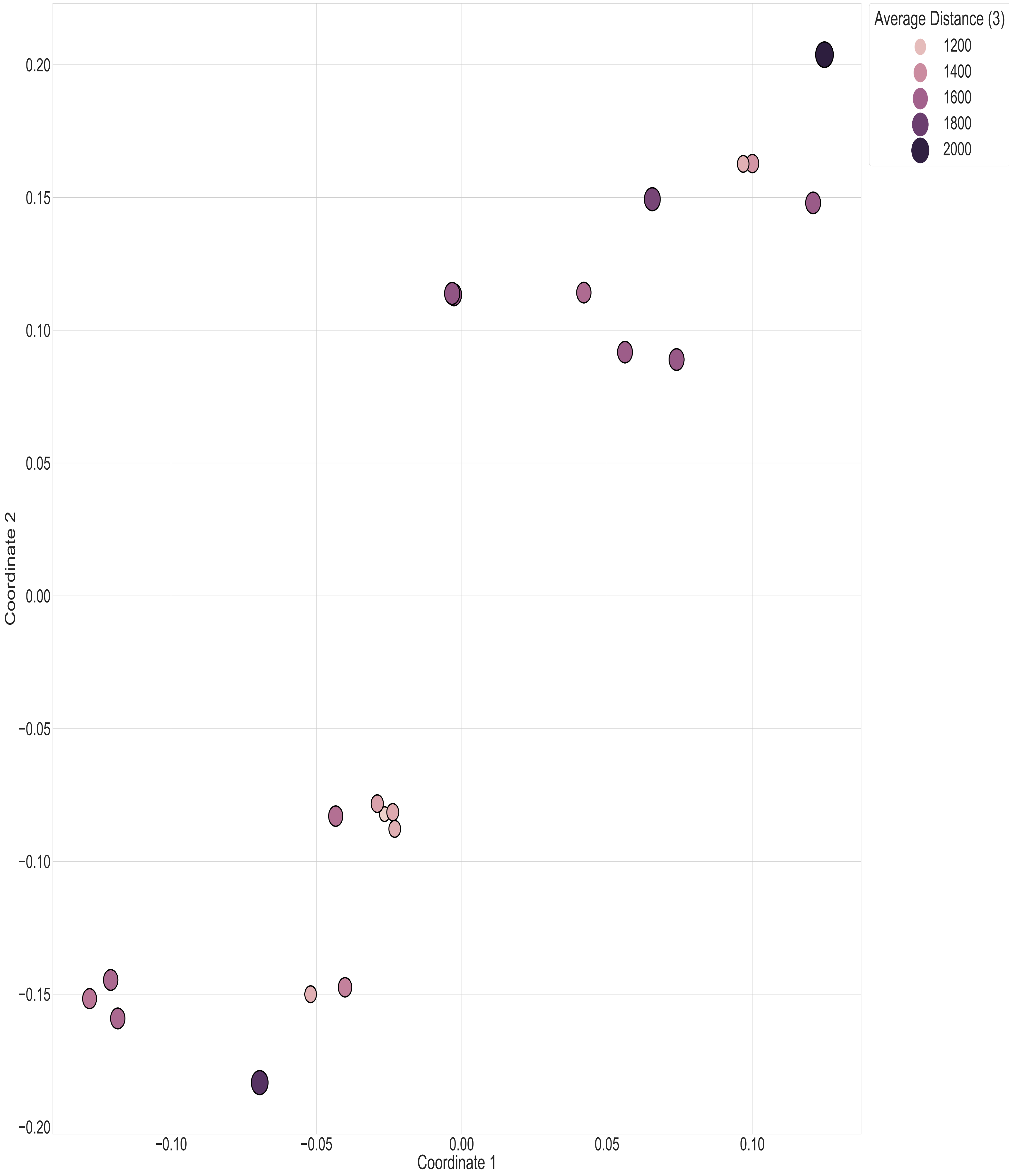
Scatter plot of PKS-1 *Micromonosporaceae* BGCs analyzed via multidimensional scaling using braycurtis pairwise distance.
Only PKS-1 BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual PKS-1 BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S115 :



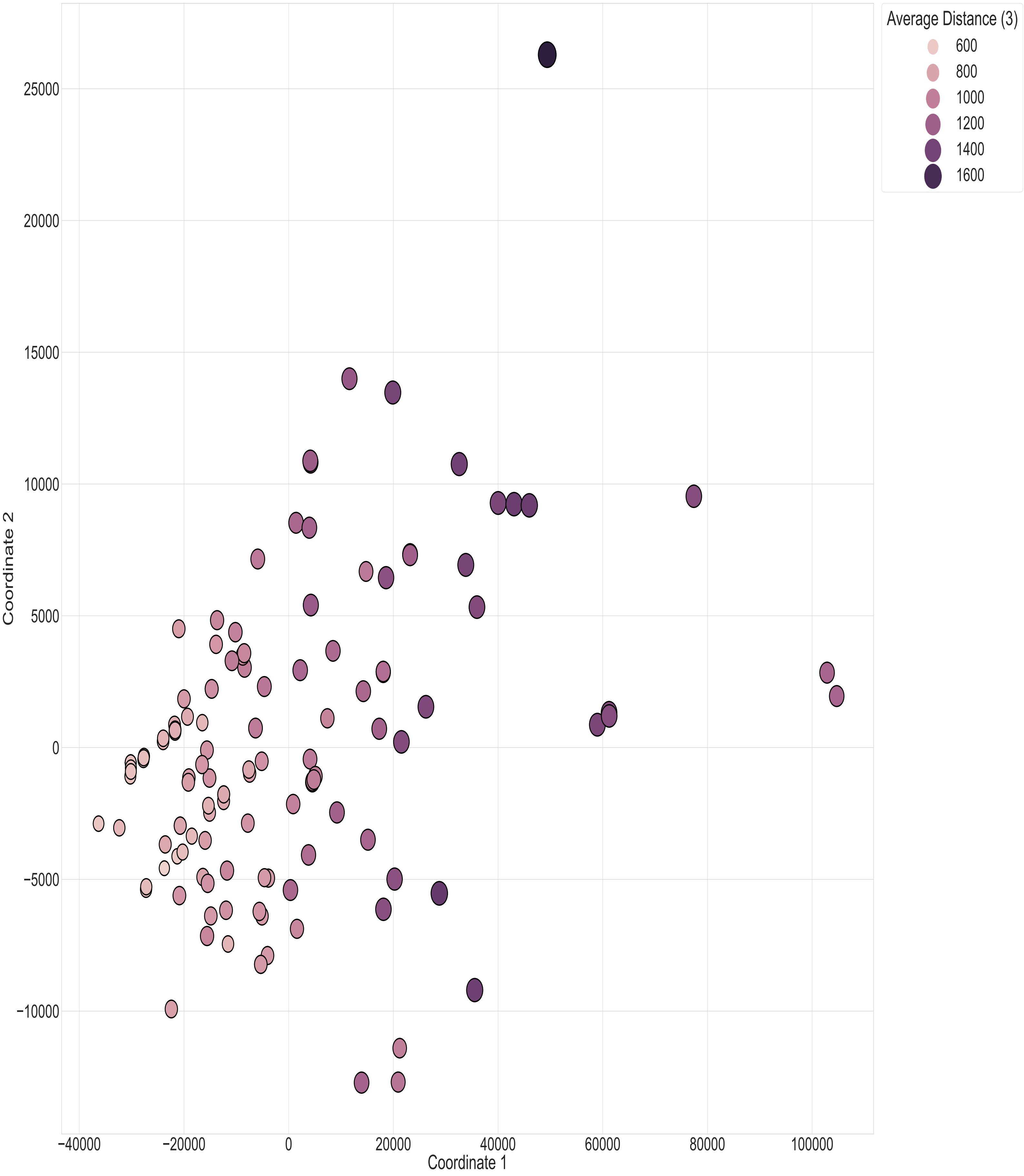
Scatter plot of PKS-1 *Micromonosporaceae* BGCs analyzed via multidimensional scaling using canberra pairwise distance.
Only PKS-1 BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual PKS-1 BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S116 :



Scatter plot of PKS-1 *Micromonosporaceae* BGCs analyzed via multidimensional scaling using correlation pairwise distance.
Only PKS-1 BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual PKS-1 BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S117 :

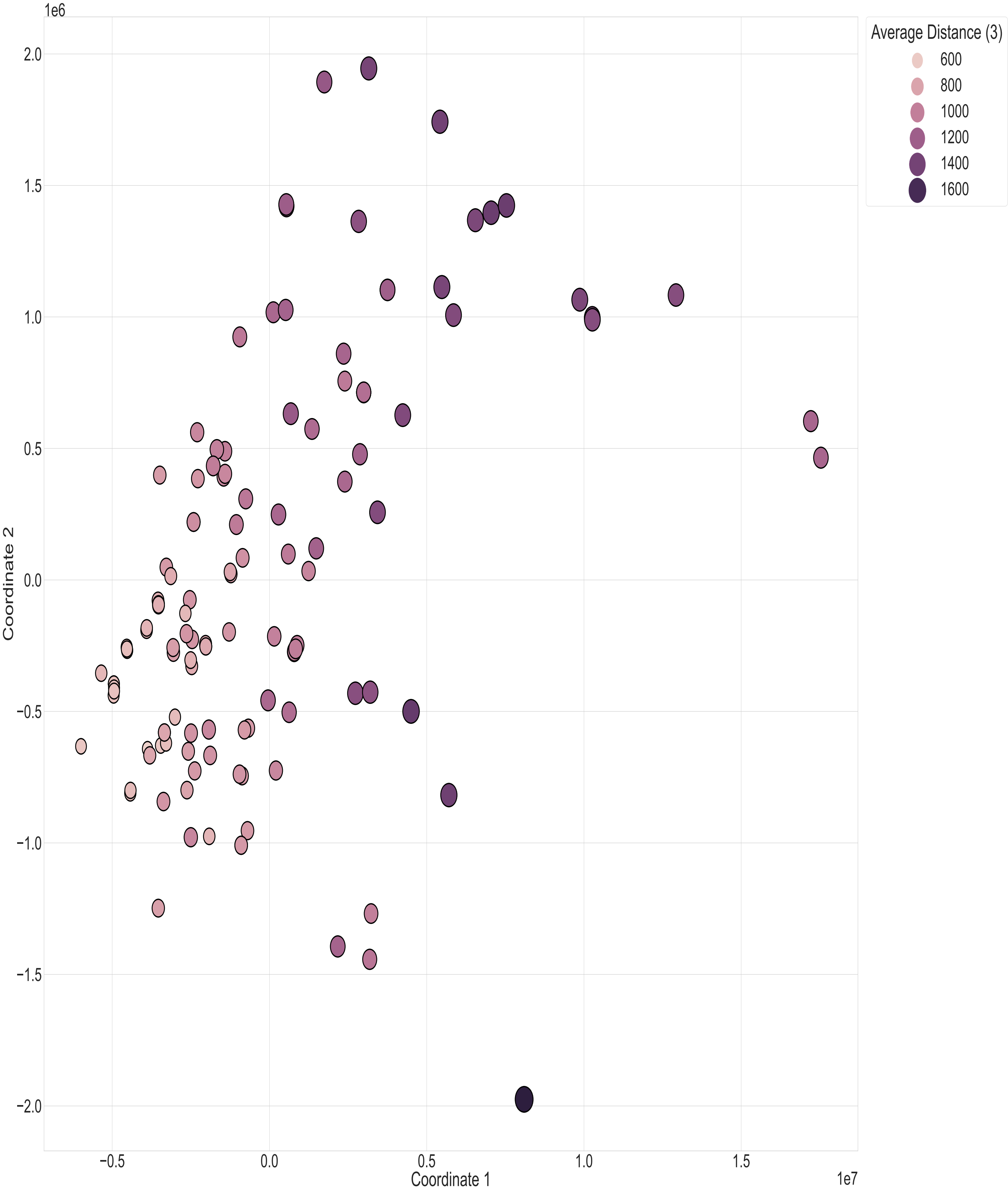


Scatter plot of RIPP *Micromonosporaceae* BGCs analyzed via multidimensional scaling using euclidean pairwise distance.
Only RIPP BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual RIPP BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S118 :

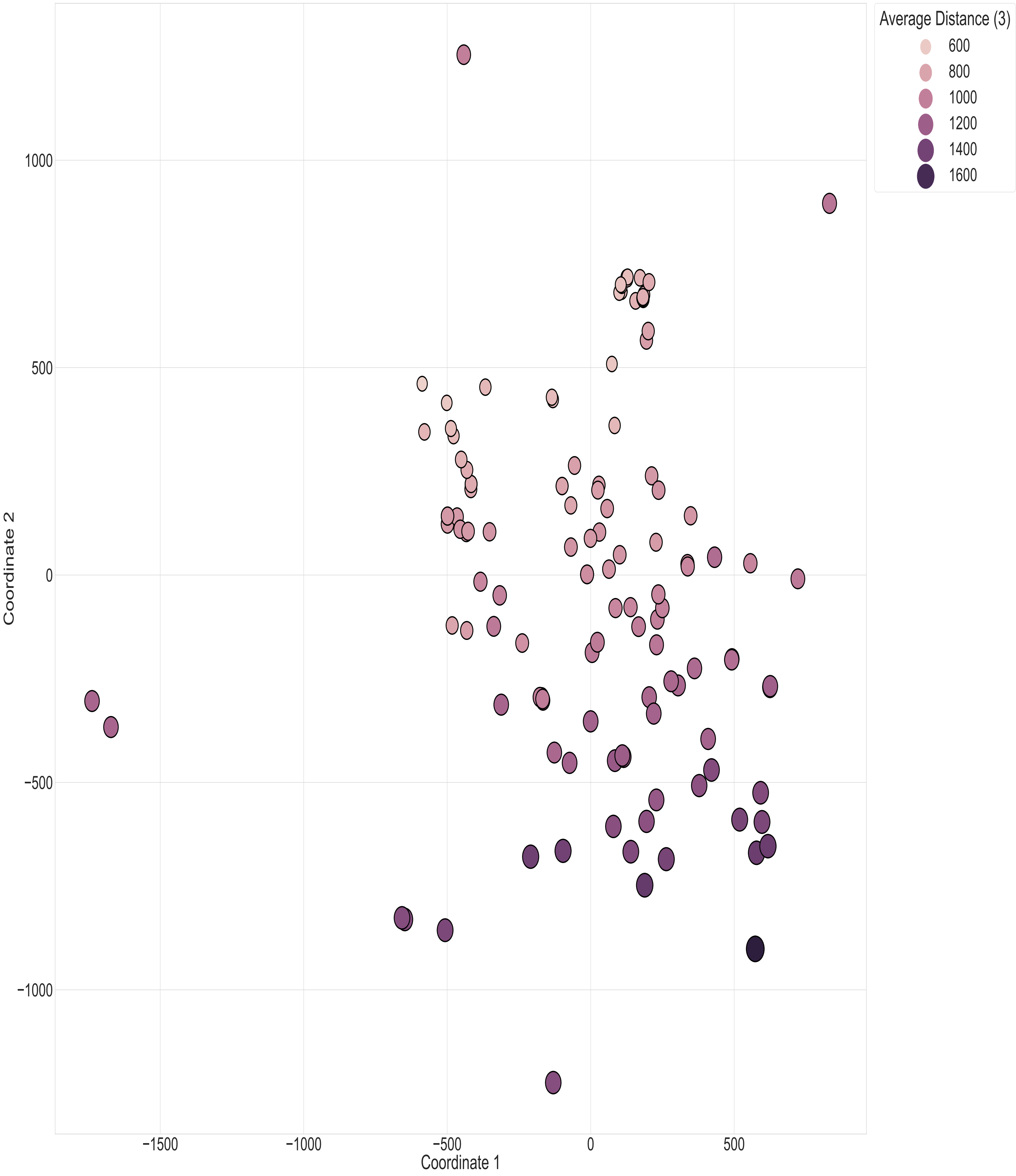
Scatter plot of RIPP *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cosine pairwise distance.
Only RIPP BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual RIPP BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S119 :



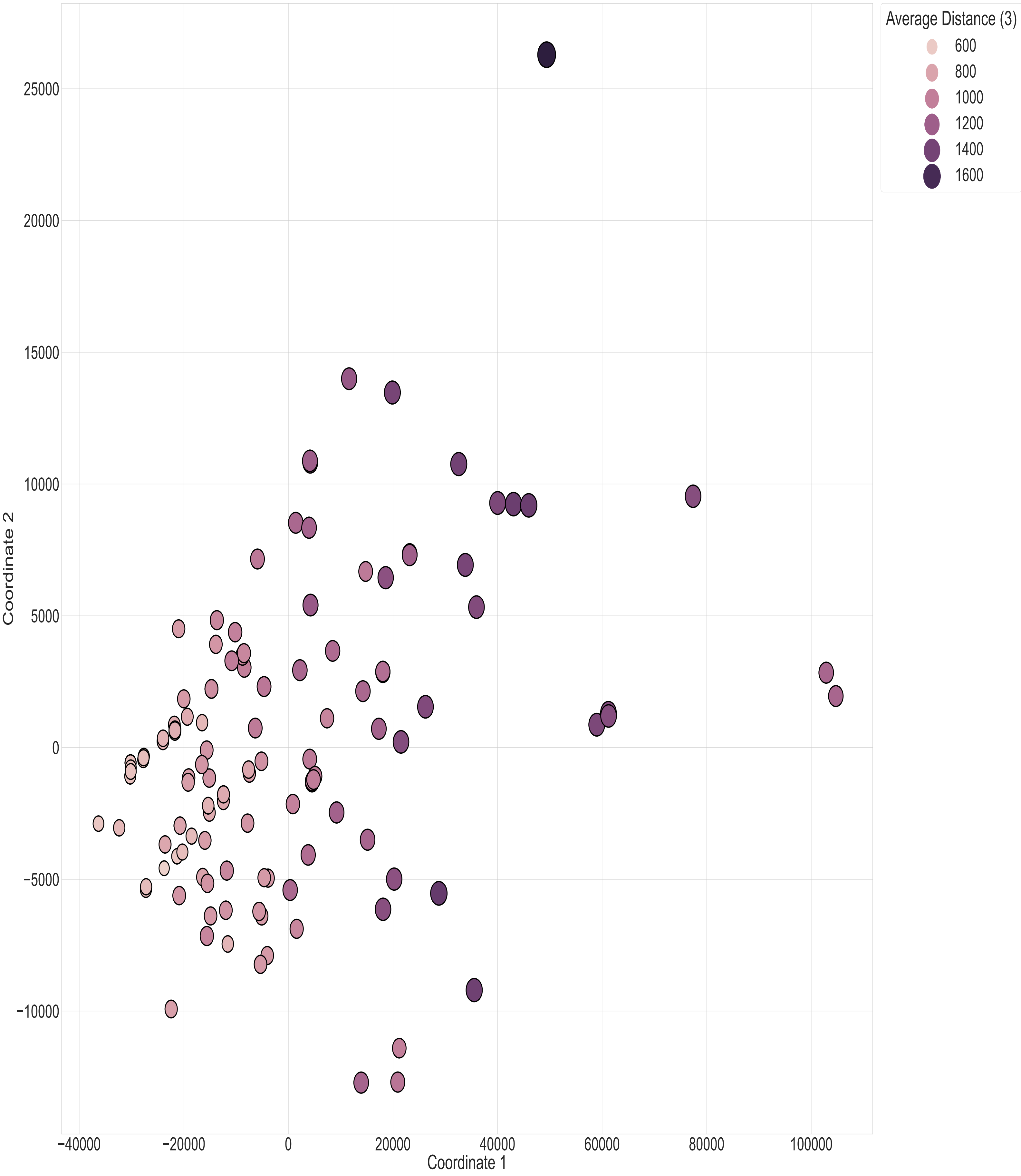
Scatter plot of RIPP *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cityblock pairwise distance. Only RIPP BGCs were considered for this analysis. Each dot represents an individual BGC. The hue and size of the dots were scaled based on the average distance of the individual RIPP BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S120 :



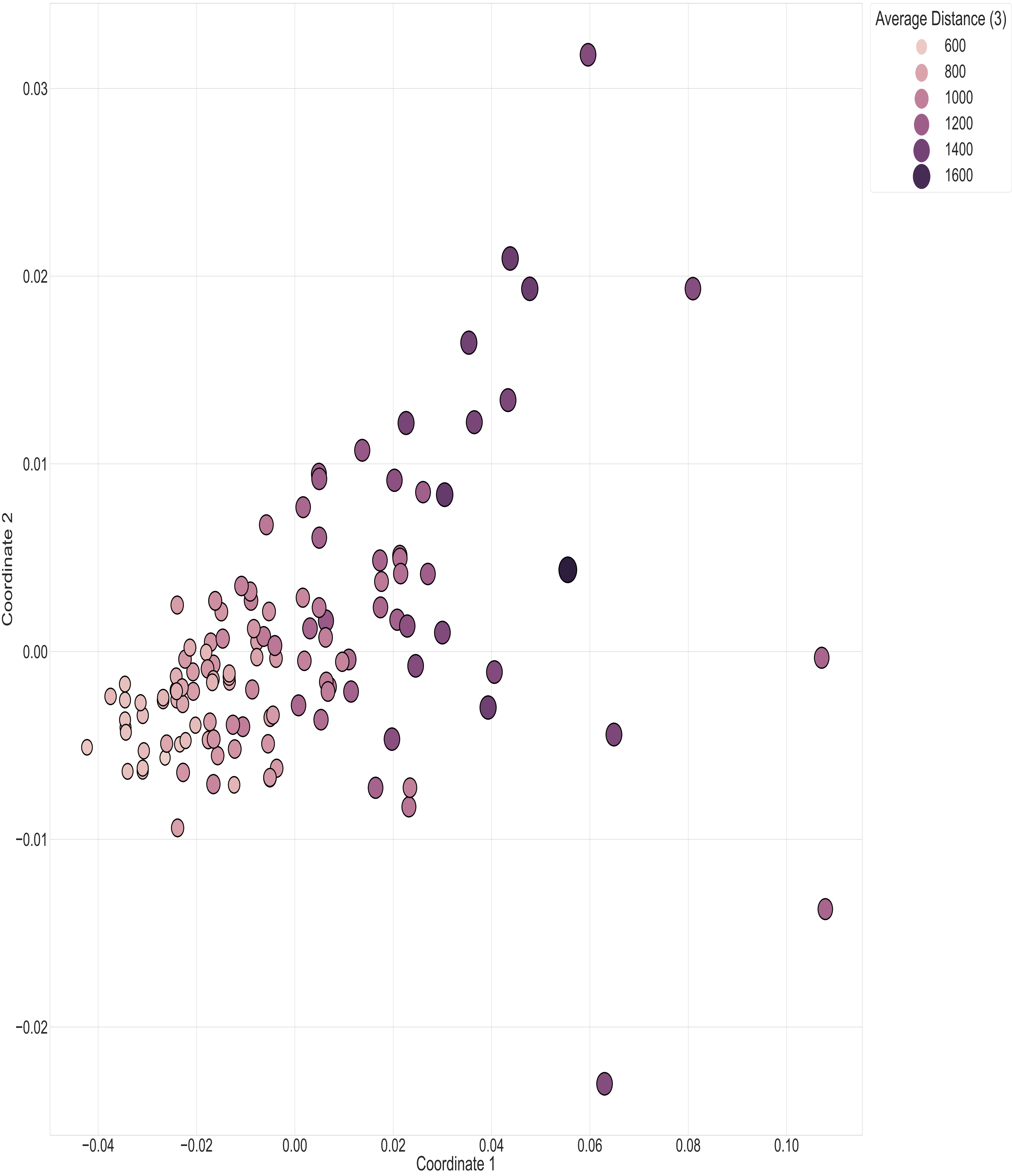
Scatter plot of RIPP *Micromonosporaceae* BGCs analyzed via multidimensional scaling using chebyshev pairwise distance.
Only RIPP BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual RIPP BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S121 :



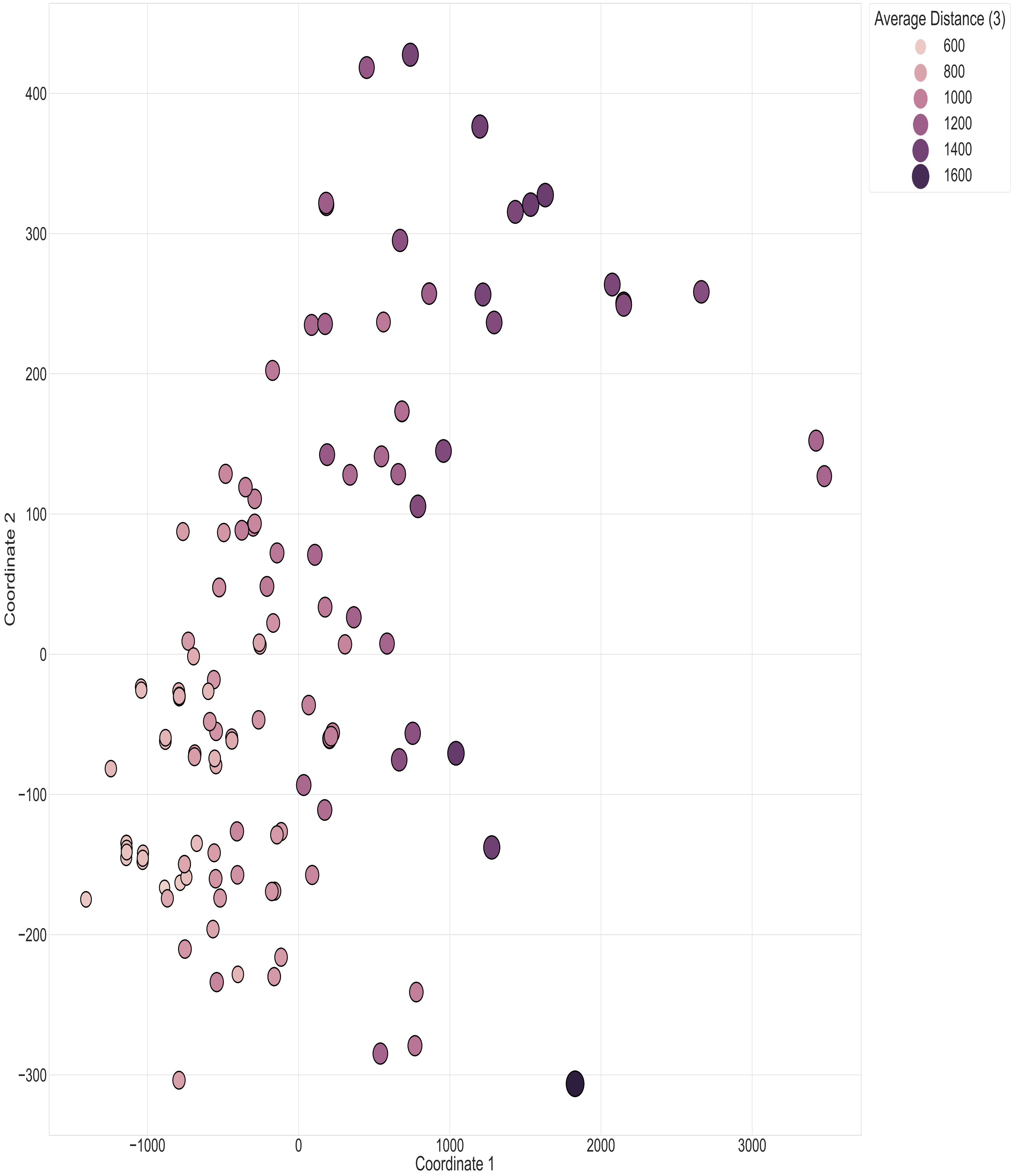
Scatter plot of RIPP *Micromonosporaceae* BGCs analyzed via multidimensional scaling using l2 pairwise distance. Only RIPP BGCs were considered for this analysis. Each dot represents an individual BGC. The hue and size of the dots were scaled based on the average distance of the individual RIPP BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S122 :



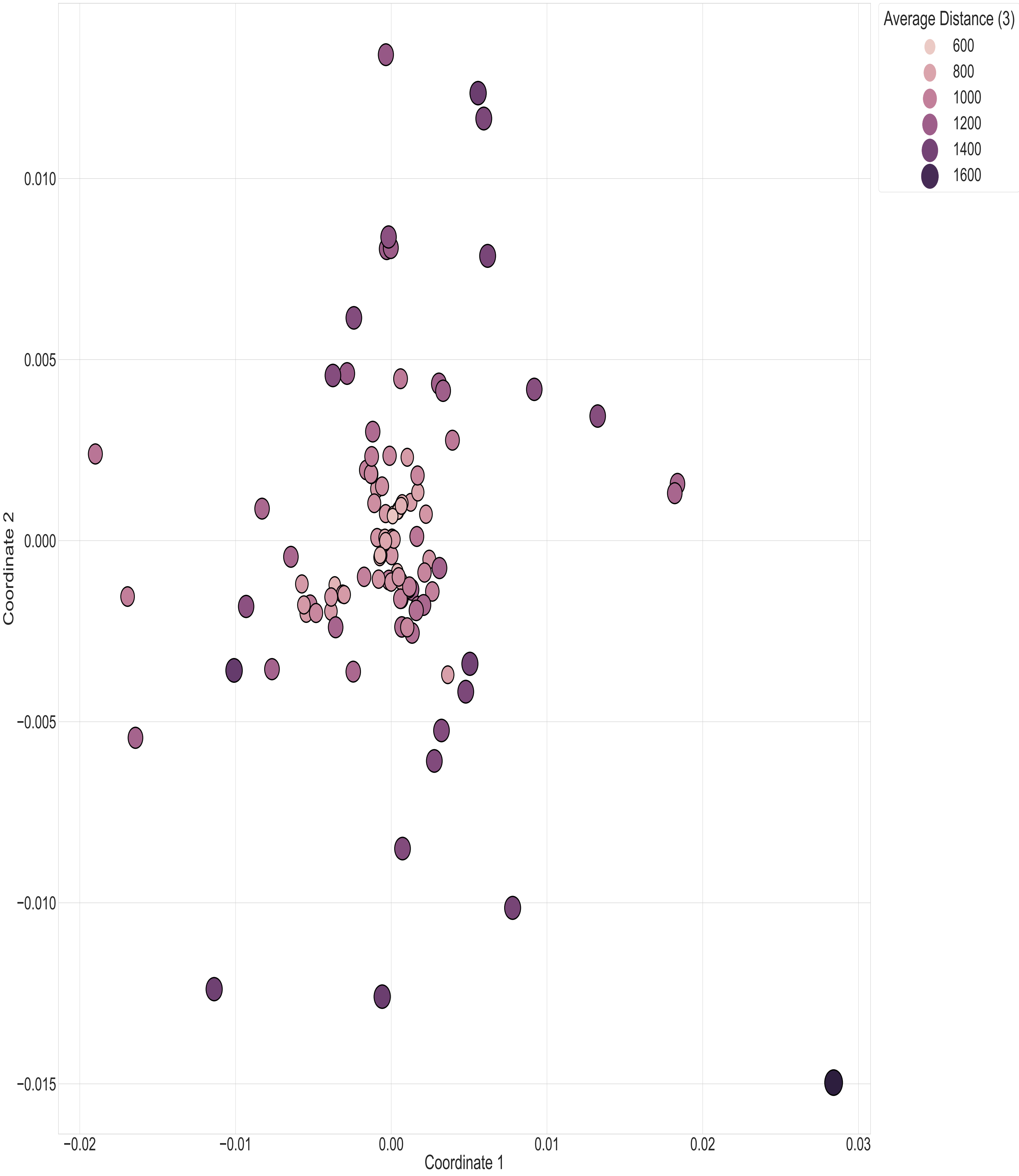
Scatter plot of RIPP *Micromonosporaceae* BGCs analyzed via multidimensional scaling using braycurtis pairwise distance.
Only RIPP BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual RIPP BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S123 :



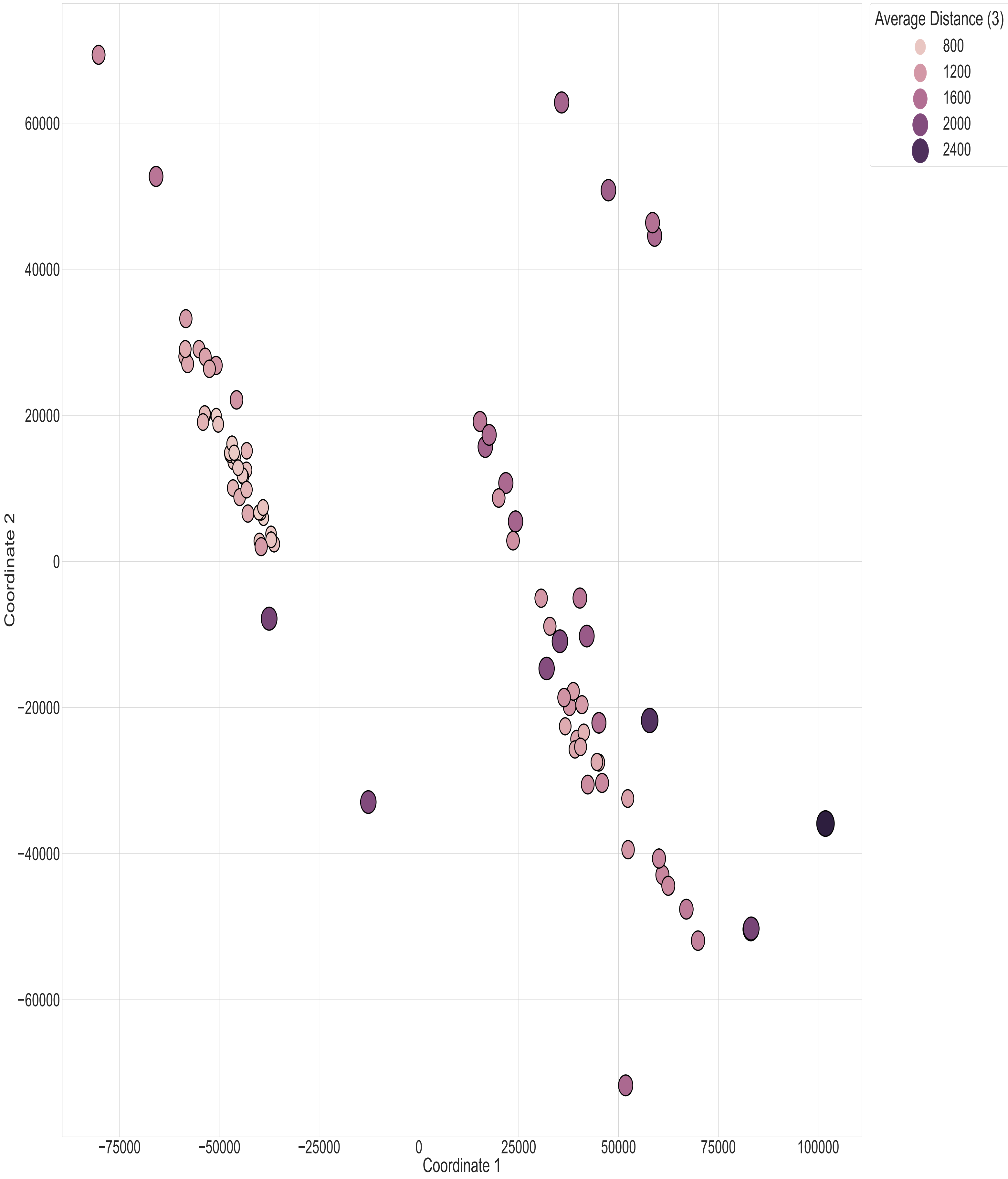
Scatter plot of RIPP *Micromonosporaceae* BGCs analyzed via multidimensional scaling using canberra pairwise distance.
Only RIPP BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual RIPP BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S124 :



Scatter plot of RIPP *Micromonosporaceae* BGCs analyzed via multidimensional scaling using correlation pairwise distance.
Only RIPP BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual RIPP BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S125 :



Scatter plot of PKS-Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using euclidean pairwise distance.
Only PKS-Other BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual PKS-Other BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

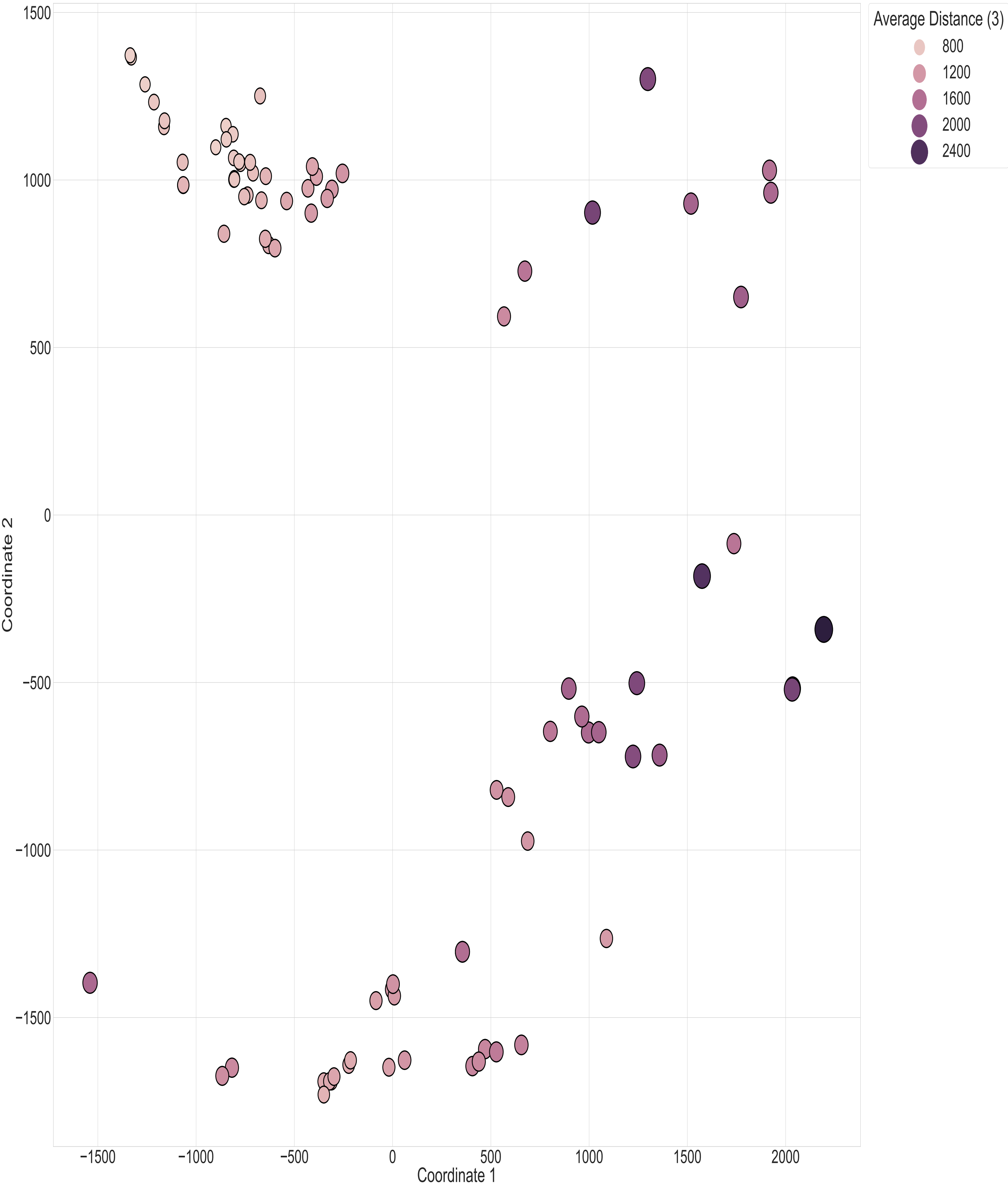
Figure – S126 :

Scatter plot of PKS-Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cosine pairwise distance.
Only PKS-Other BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual PKS-Other BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S127 :

Scatter plot of PKS-Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cityblock pairwise distance.
Only PKS-Other BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual PKS-Other BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S128 :

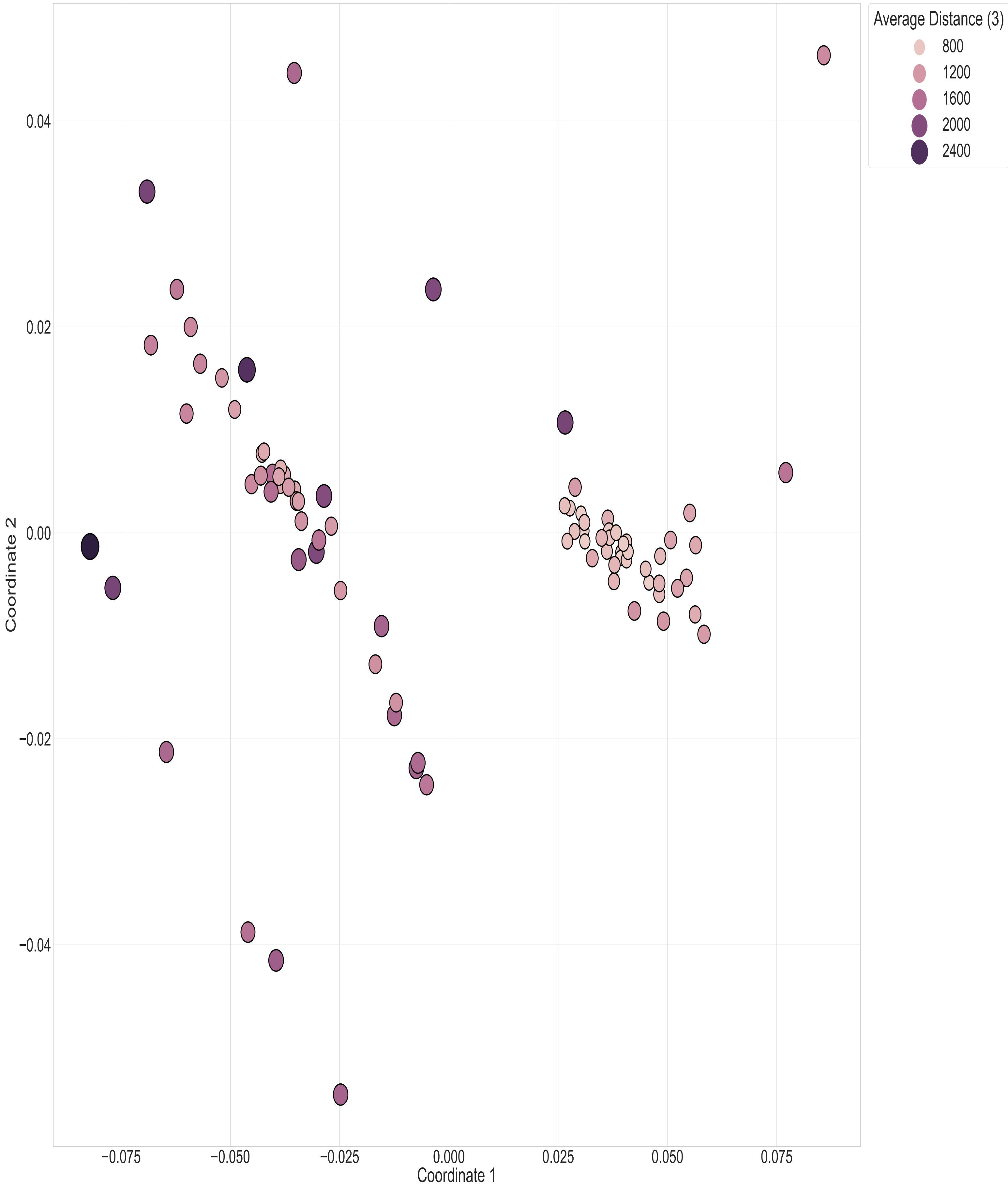


Scatter plot of PKS-Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using chebyshev pairwise distance.
Only PKS-Other BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual PKS-Other BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S129 :

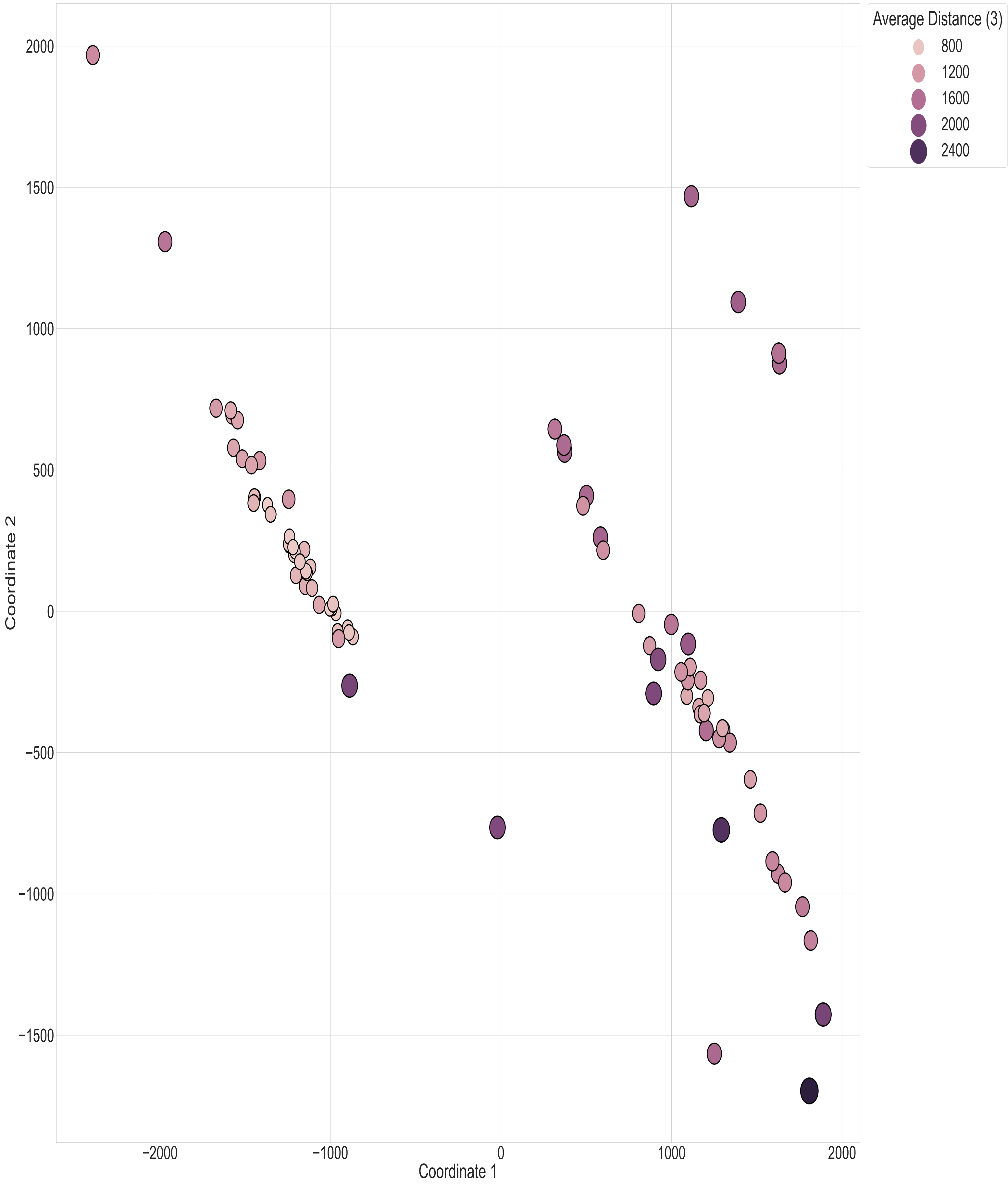
Scatter plot of PKS-Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using l2 pairwise distance.
Only PKS-Other BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual PKS-Other BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S130 :



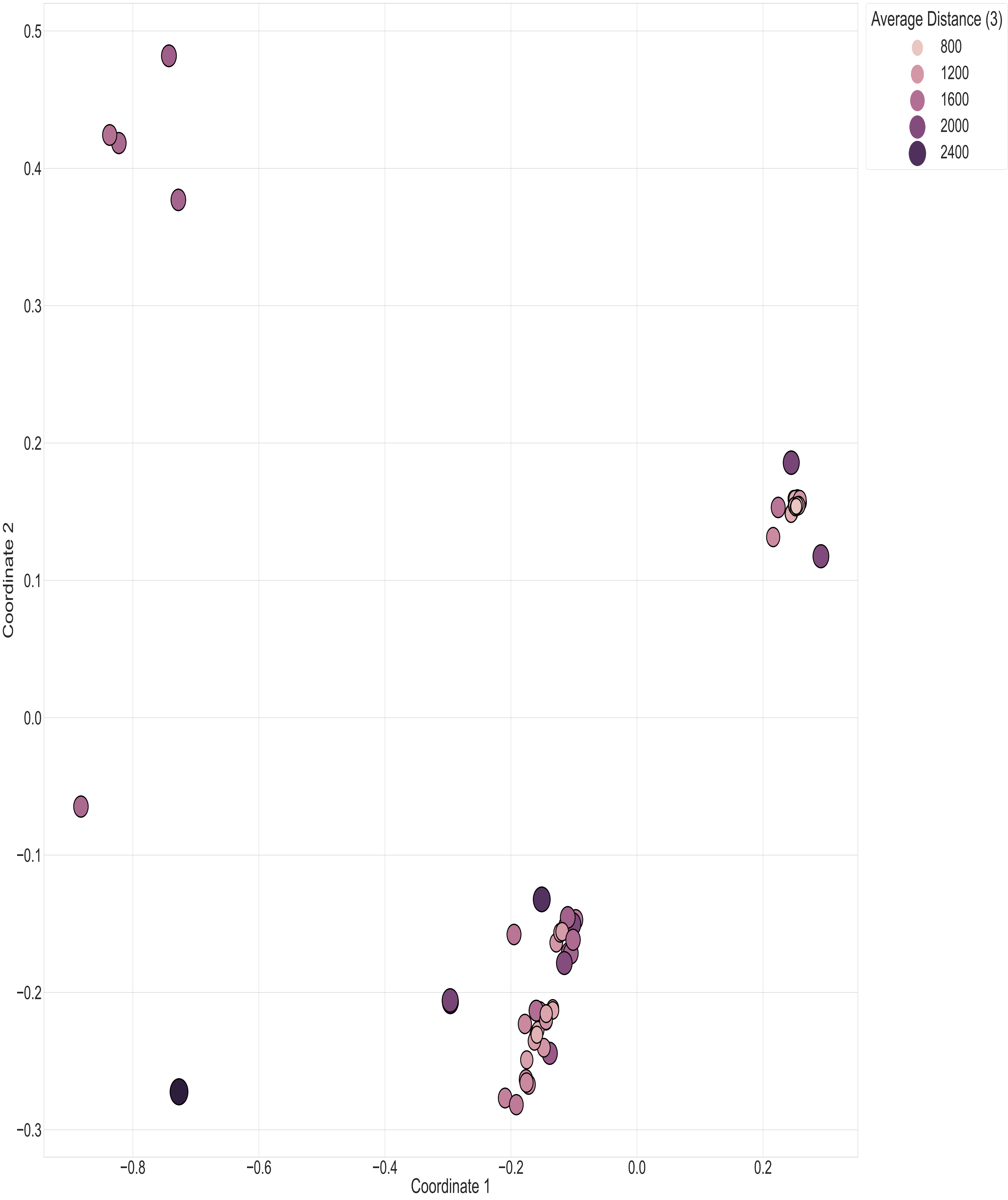
Scatter plot of PKS-Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using braycurtis pairwise distance.
Only PKS-Other BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual PKS-Other BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S131 :



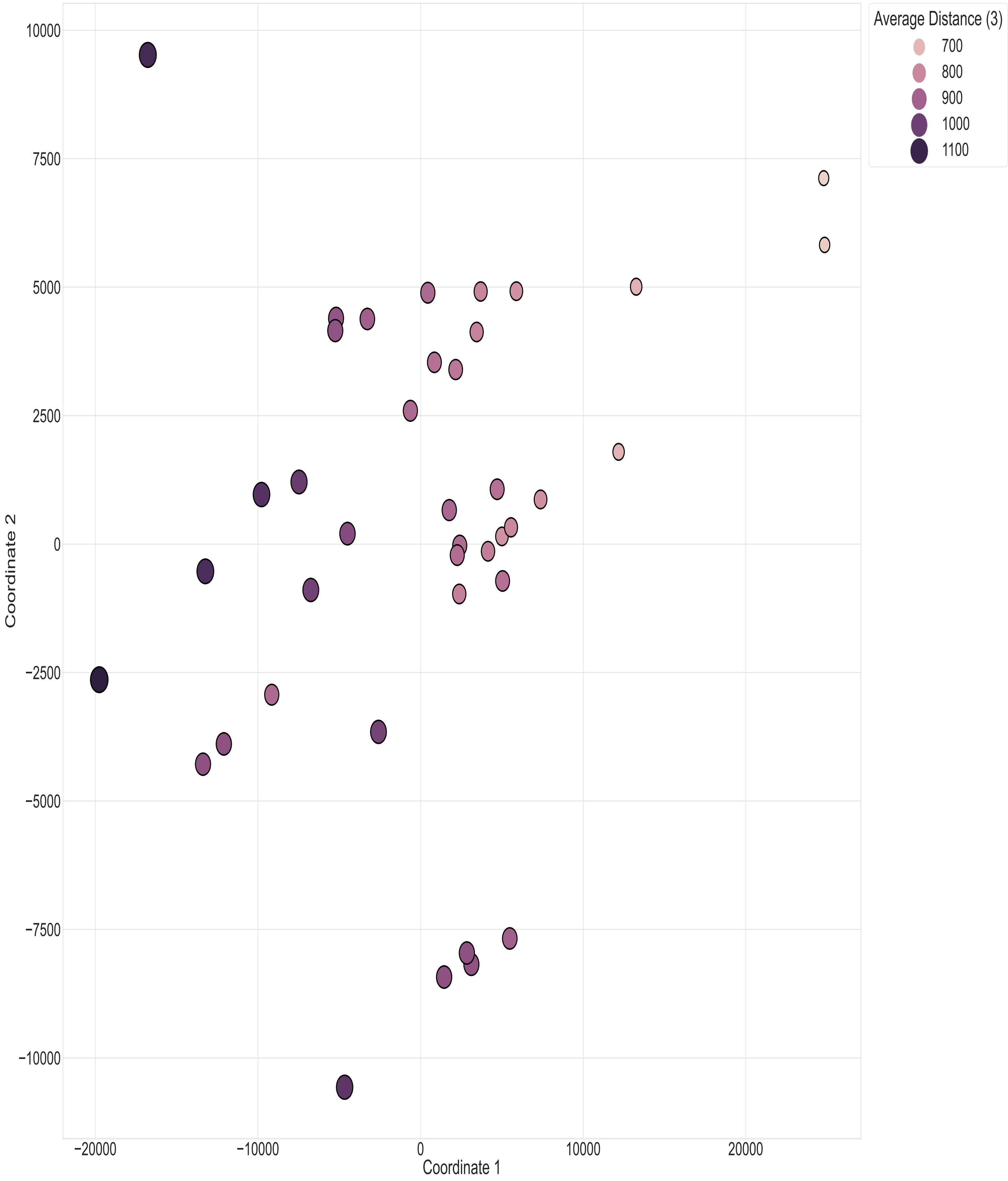
Scatter plot of PKS-Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using canberra pairwise distance.
Only PKS-Other BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual PKS-Other BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S132 :



Scatter plot of PKS-Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using correlation pairwise distance. Only PKS-Other BGCs were considered for this analysis. Each dot represents an individual BGC. The hue and size of the dots were scaled based on the average distance of the individual PKS-Other BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S133 :

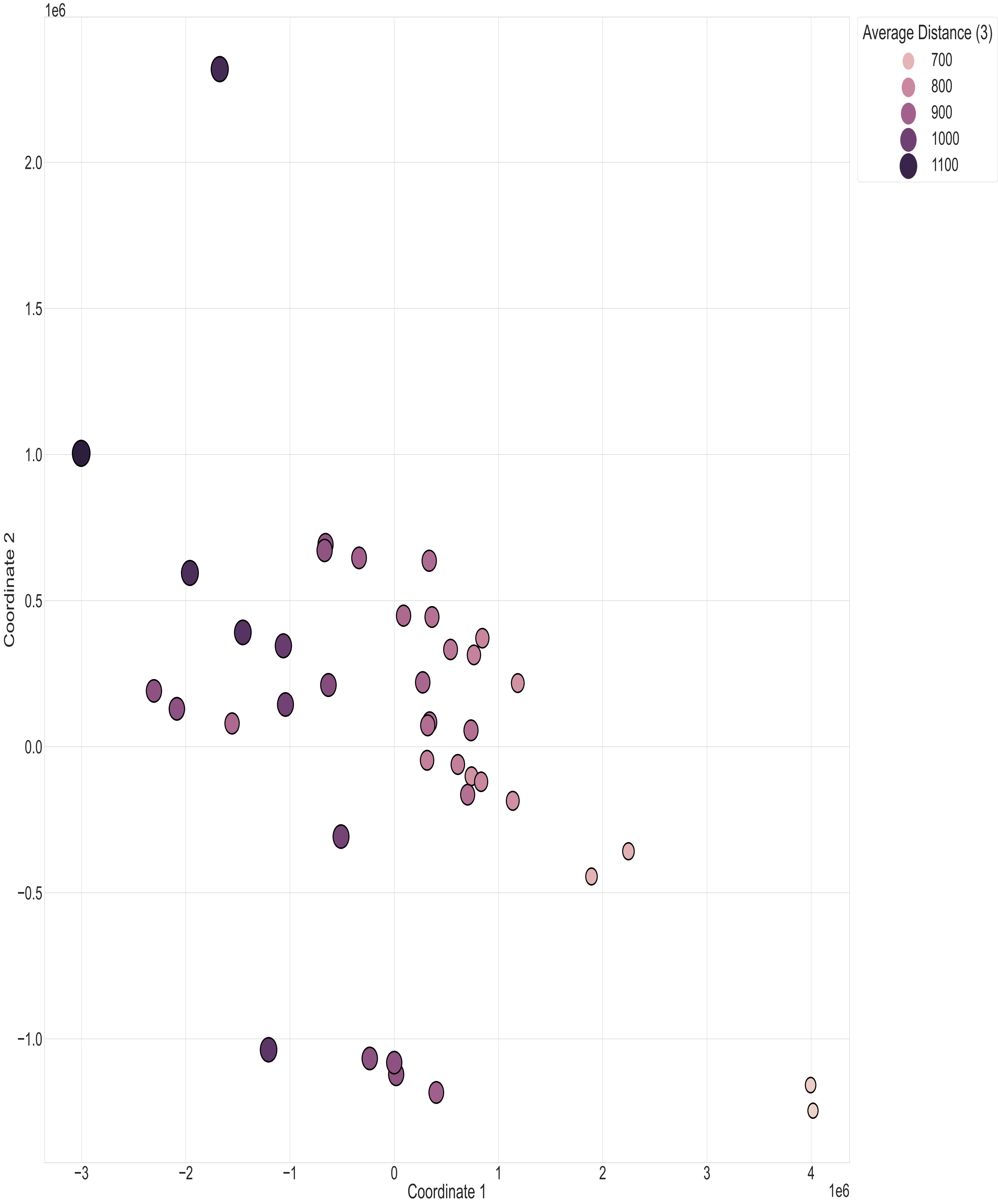


Scatter plot of Siderophore *Micromonosporaceae* BGCs analyzed via multidimensional scaling using euclidean pairwise distance.
Only Siderophore BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual Siderophore BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S134 :

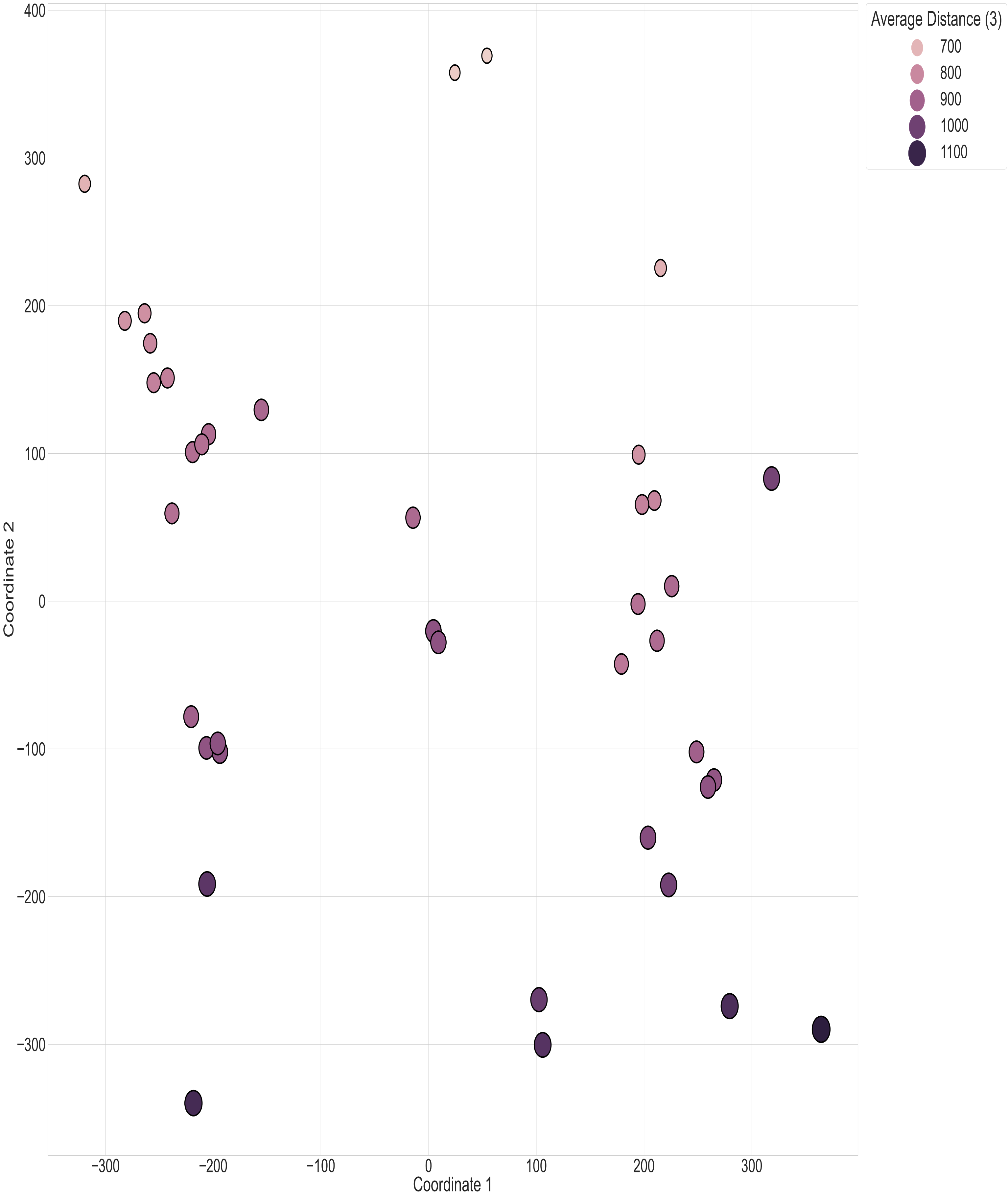
Scatter plot of Siderophore *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cosine pairwise distance.
Only Siderophore BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual Siderophore BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S135 :



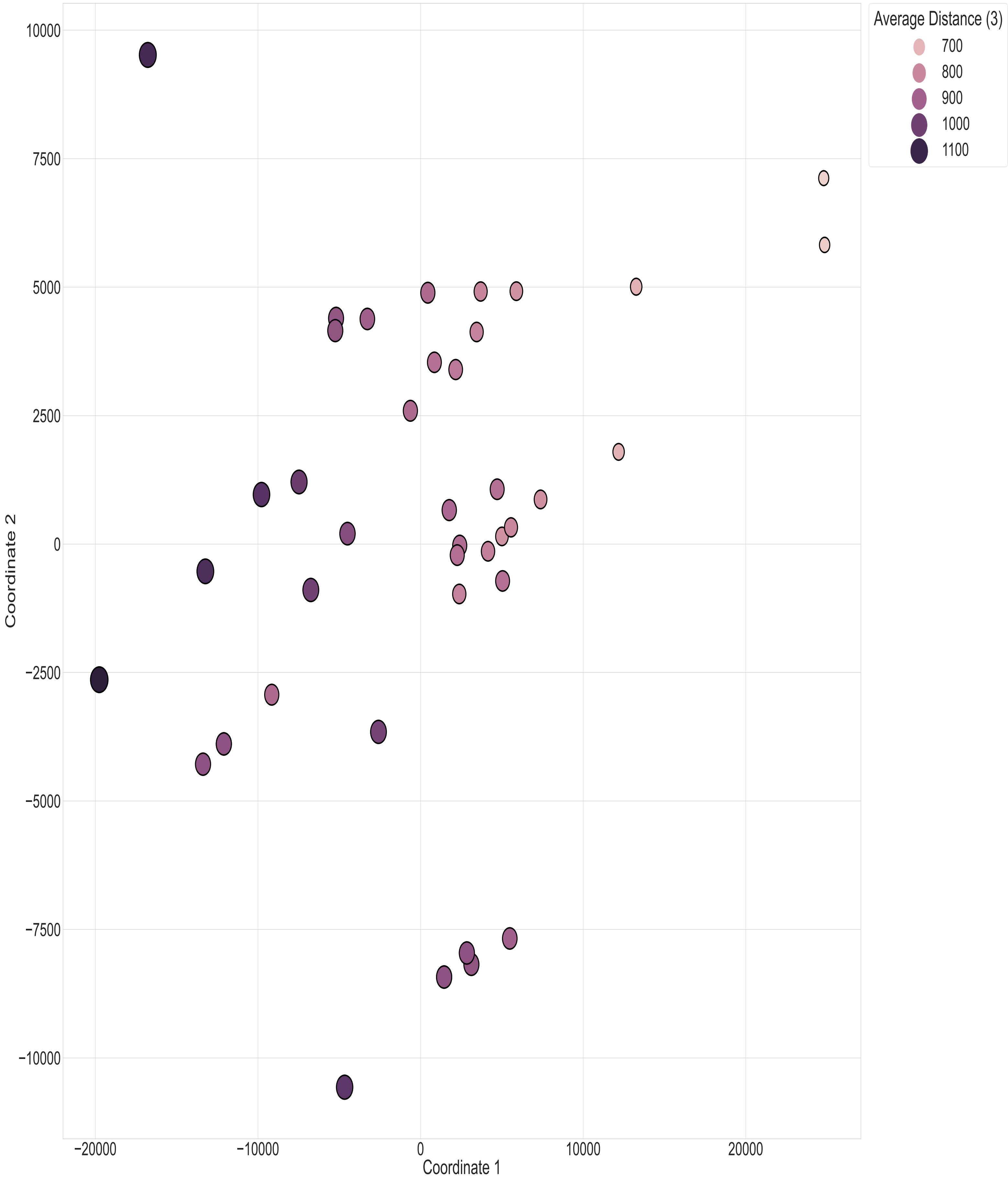
Scatter plot of Siderophore *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cityblock pairwise distance.
Only Siderophore BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual Siderophore BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S136 :



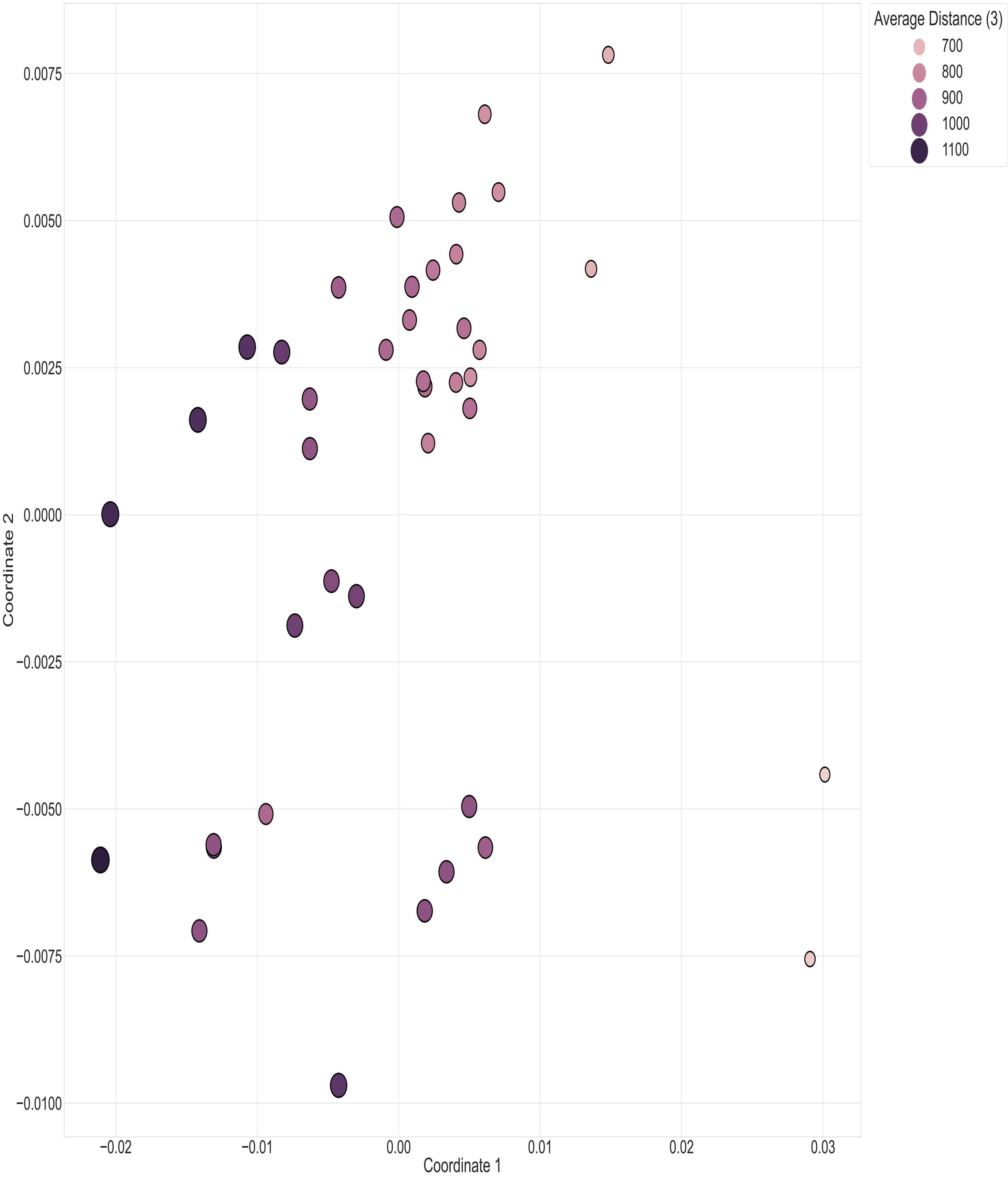
Scatter plot of Siderophore *Micromonosporaceae* BGCs analyzed via multidimensional scaling using chebyshev pairwise distance.
Only Siderophore BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual Siderophore BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S137 :



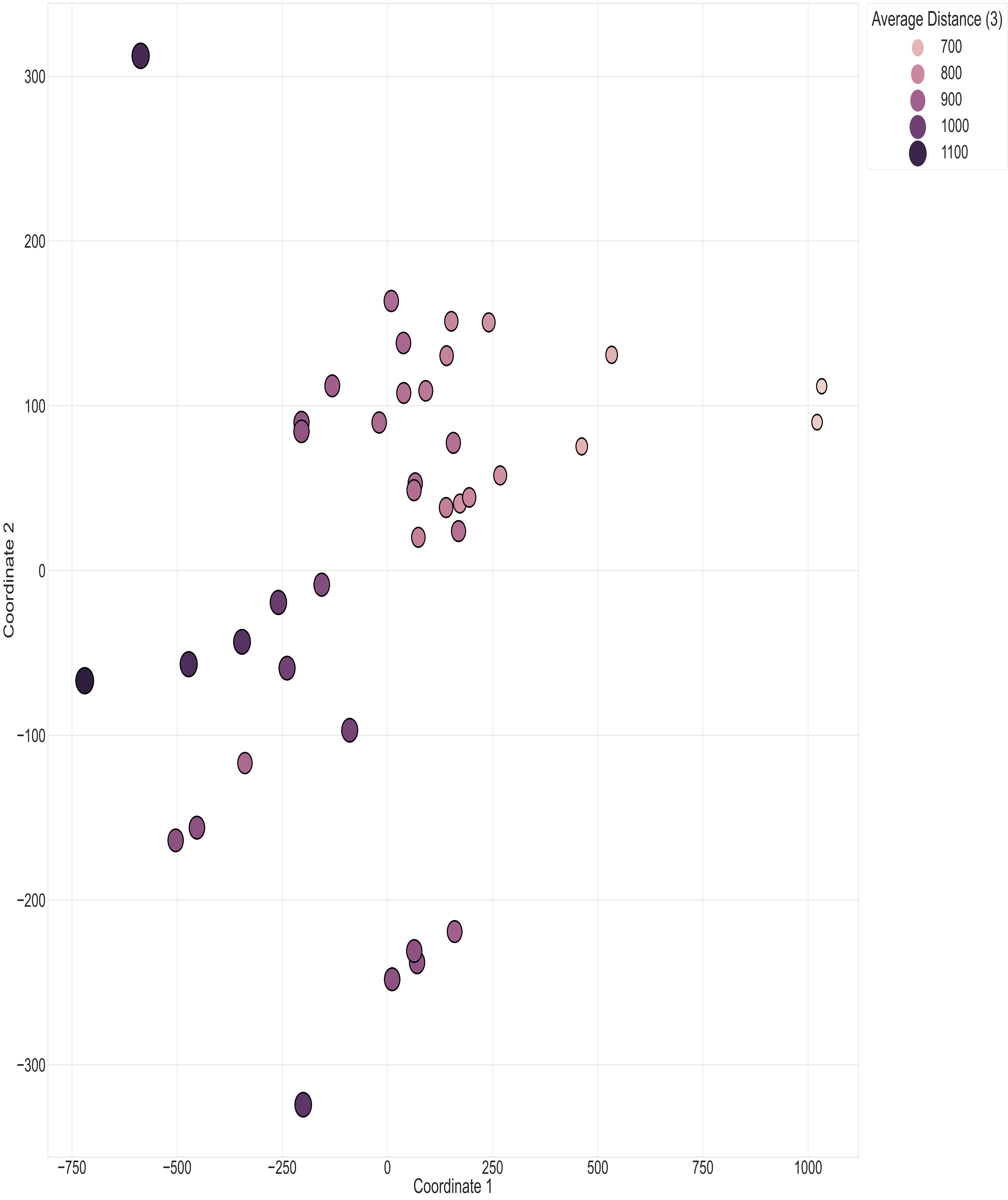
Scatter plot of Siderophore *Micromonosporaceae* BGCs analyzed via multidimensional scaling using l2 pairwise distance.
Only Siderophore BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual Siderophore BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S138 :



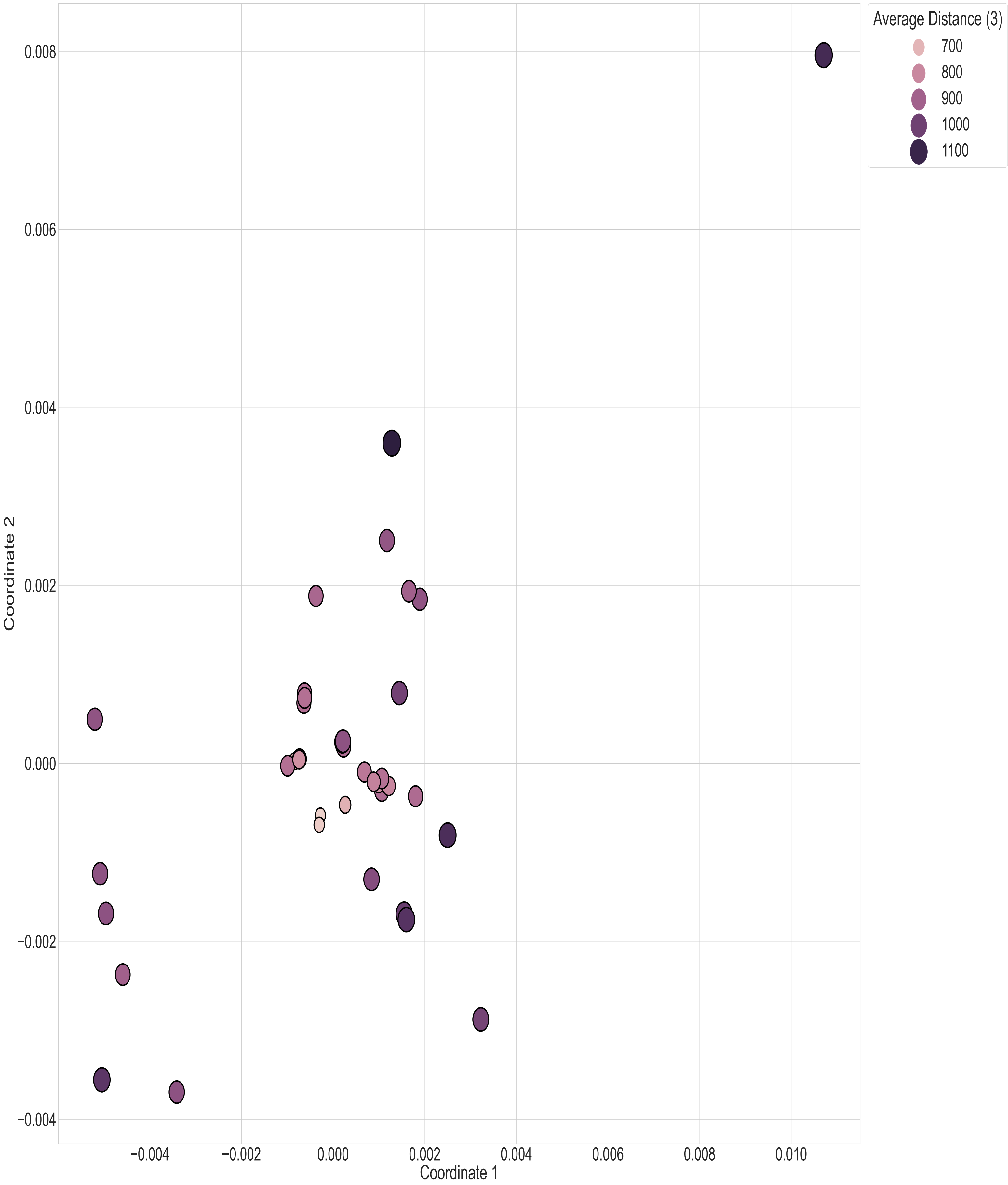
Scatter plot of Siderophore *Micromonosporaceae* BGCs analyzed via multidimensional scaling using braycurtis pairwise distance.
Only Siderophore BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual Siderophore BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S139 :



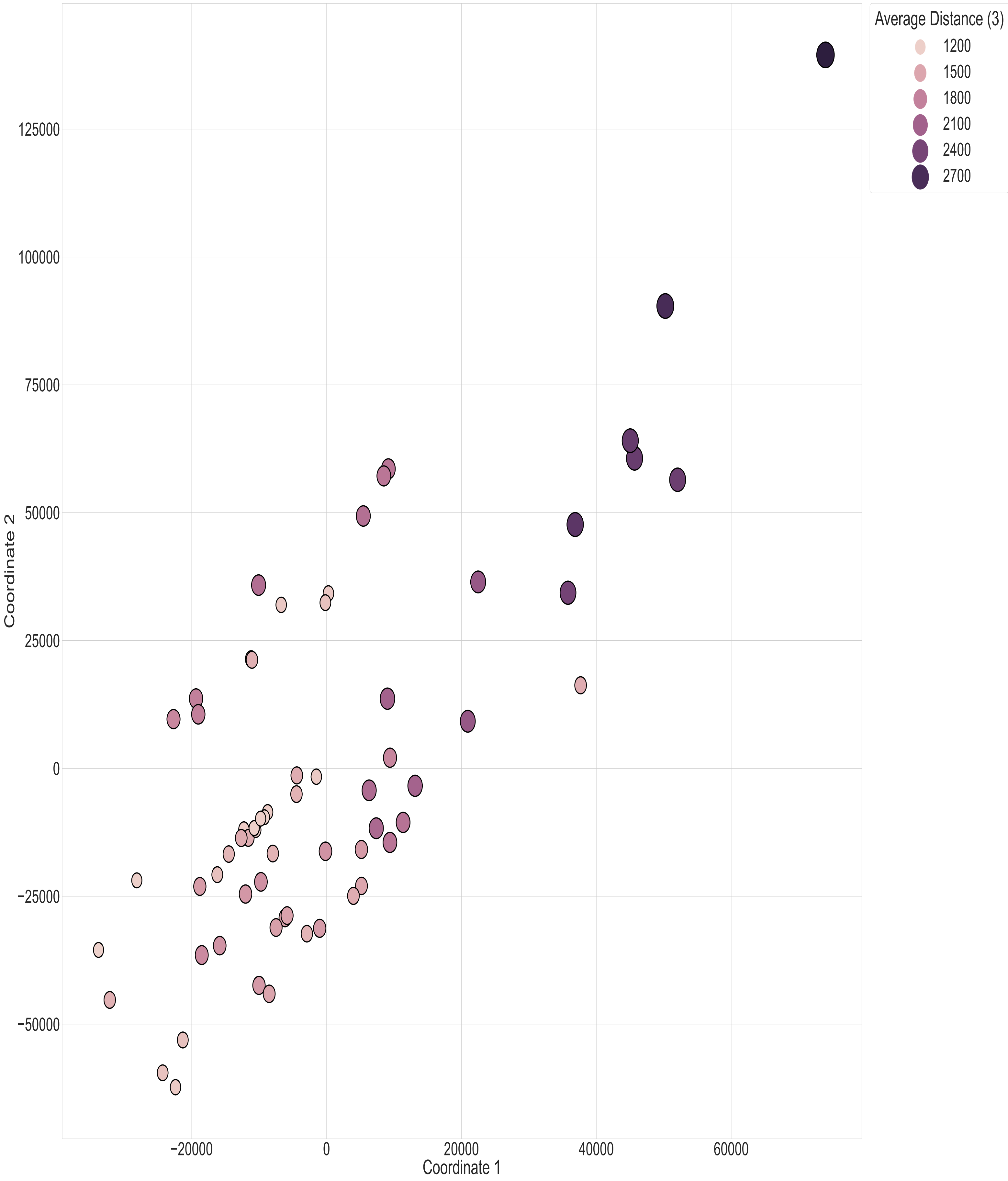
Scatter plot of Siderophore *Micromonosporaceae* BGCs analyzed via multidimensional scaling using canberra pairwise distance.
Only Siderophore BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual Siderophore BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S140 :



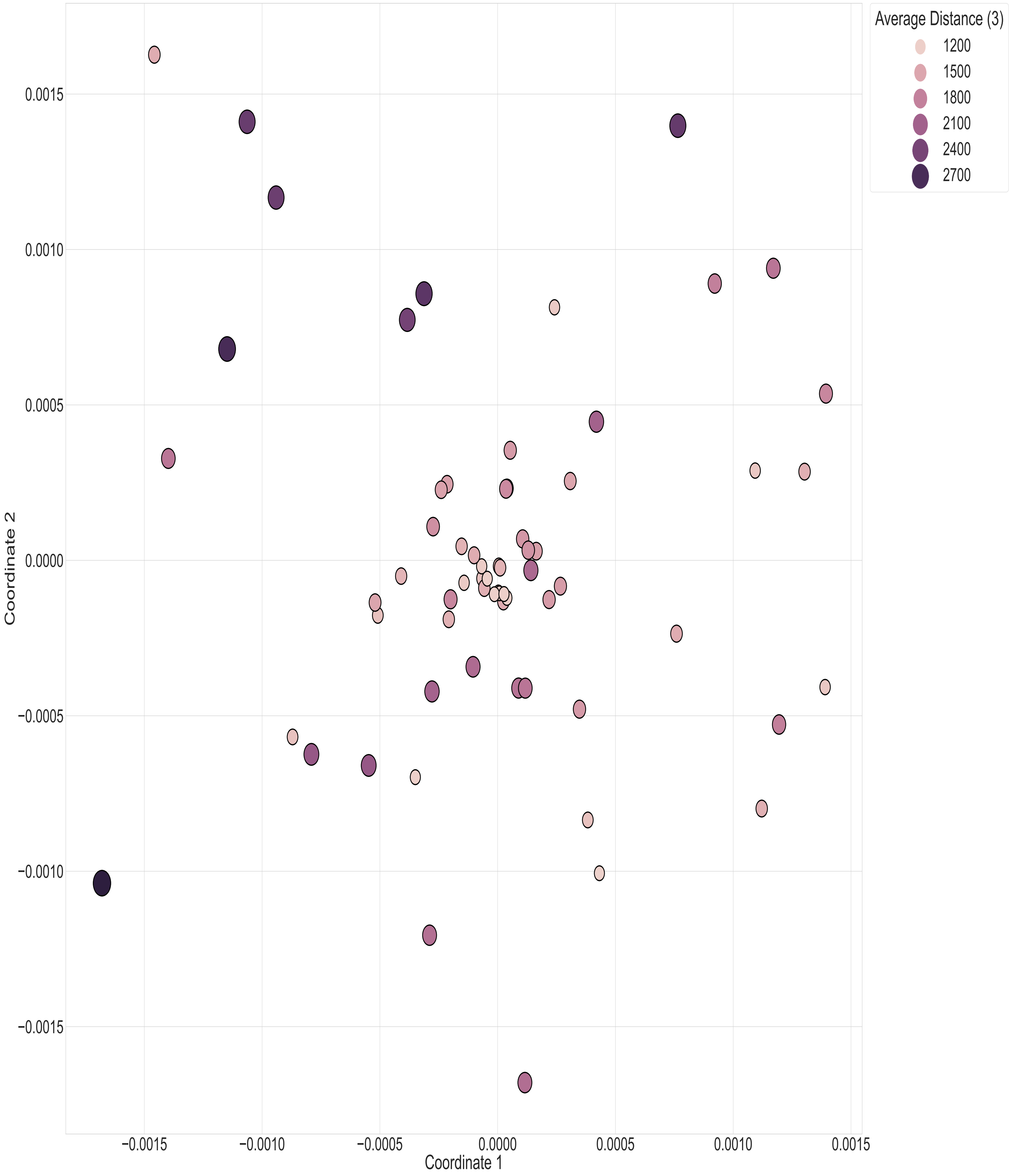
Scatter plot of Siderophore *Micromonosporaceae* BGCs analyzed via multidimensional scaling using correlation pairwise distance.
Only Siderophore BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual Siderophore BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S141 :



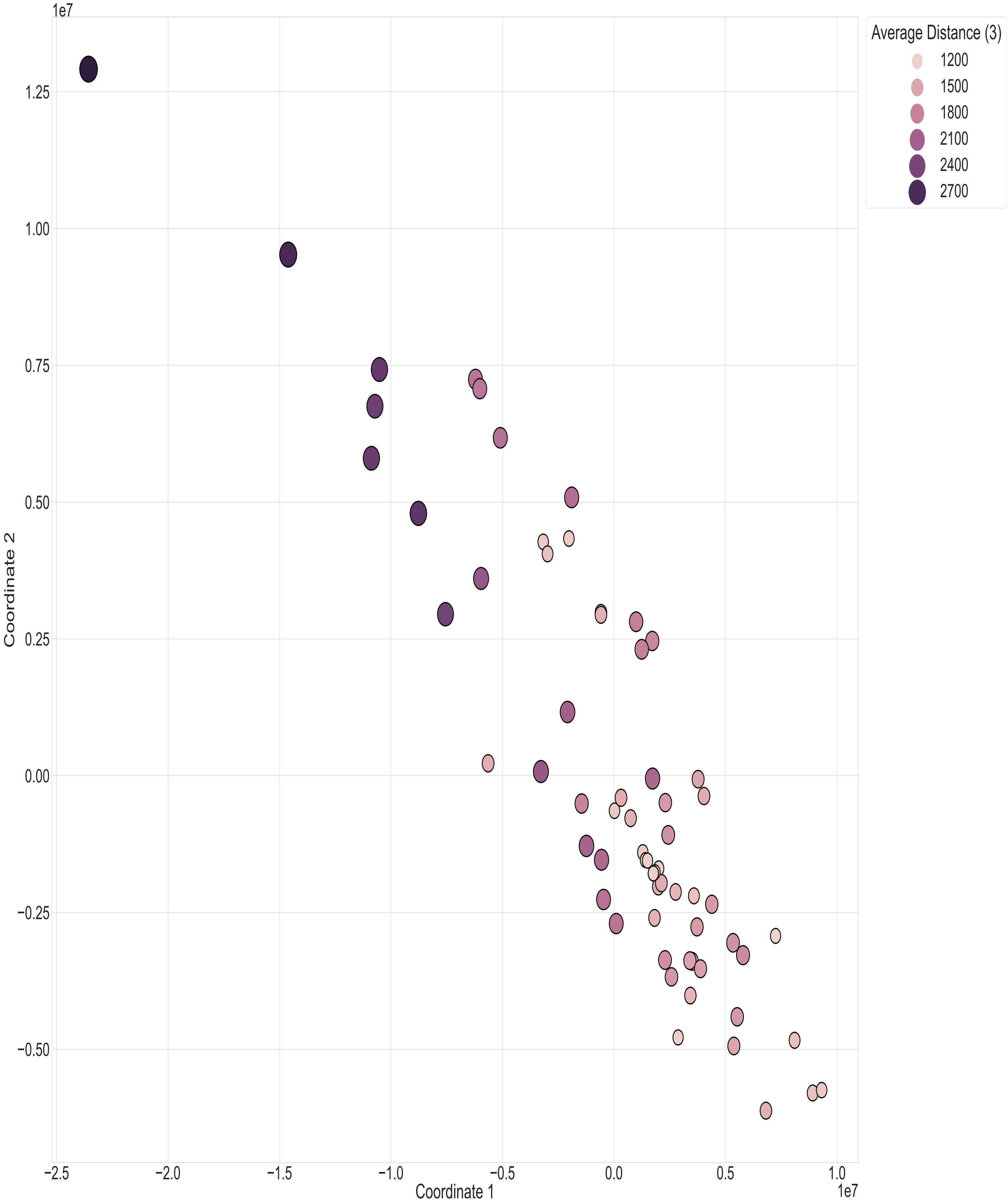
Scatter plot of PKS-NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using euclidean pairwise distance.
Only PKS-NRPS BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual PKS-NRPS BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S142 :



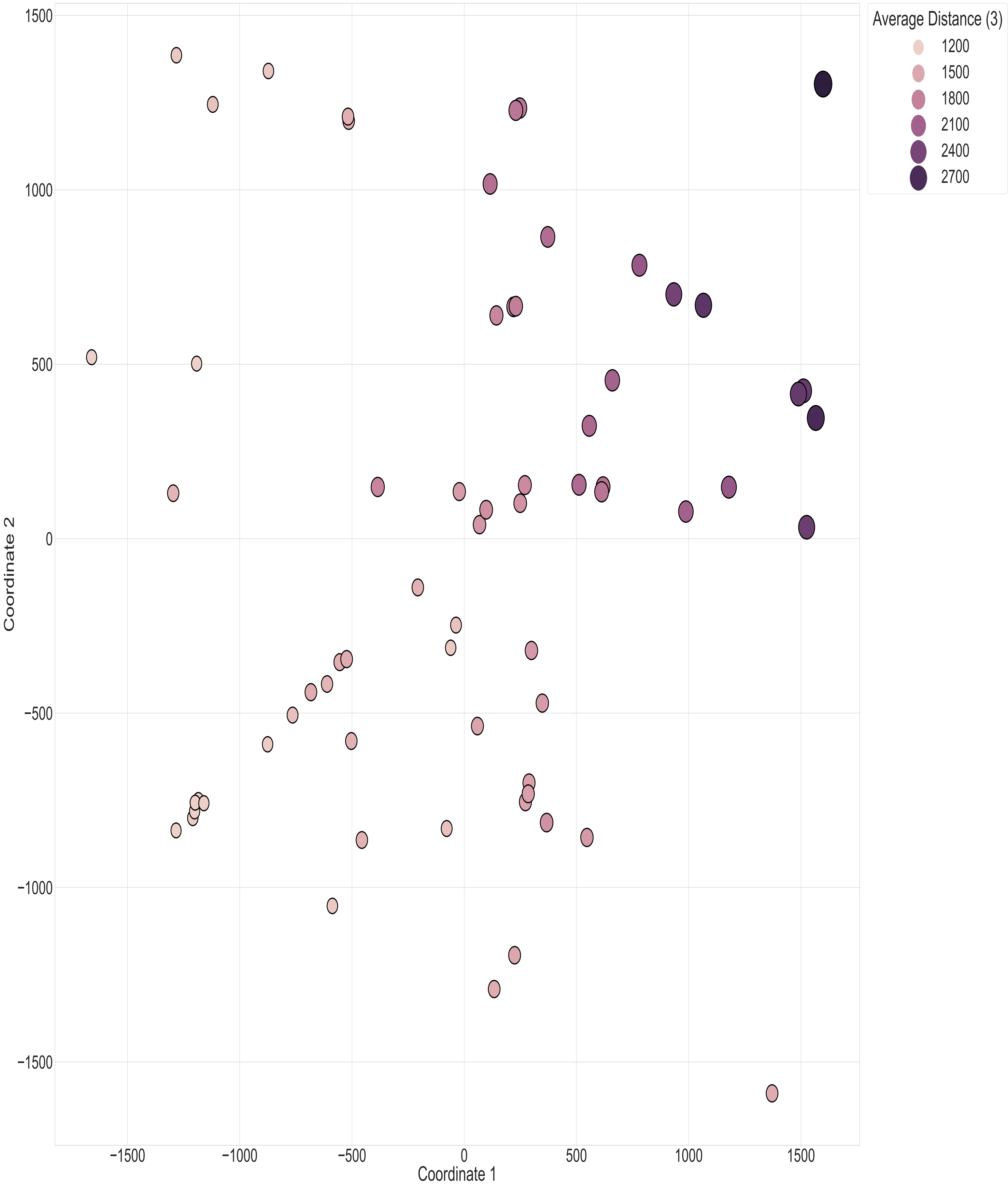
Scatter plot of PKS-NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cosine pairwise distance.
Only PKS-NRPS BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual PKS-NRPS BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S143 :



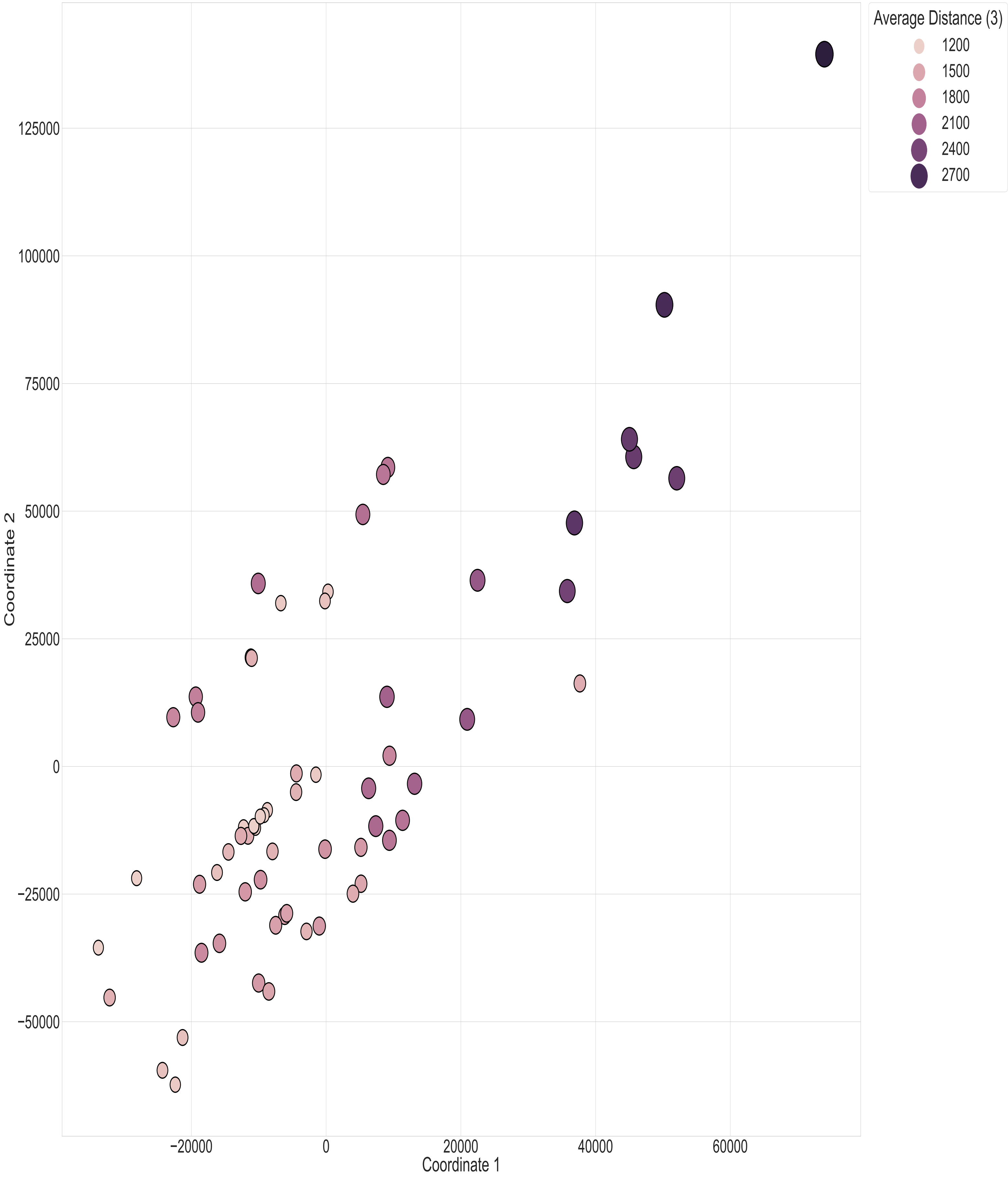
Scatter plot of PKS-NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cityblock pairwise distance.
Only PKS-NRPS BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual PKS-NRPS BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S144 :



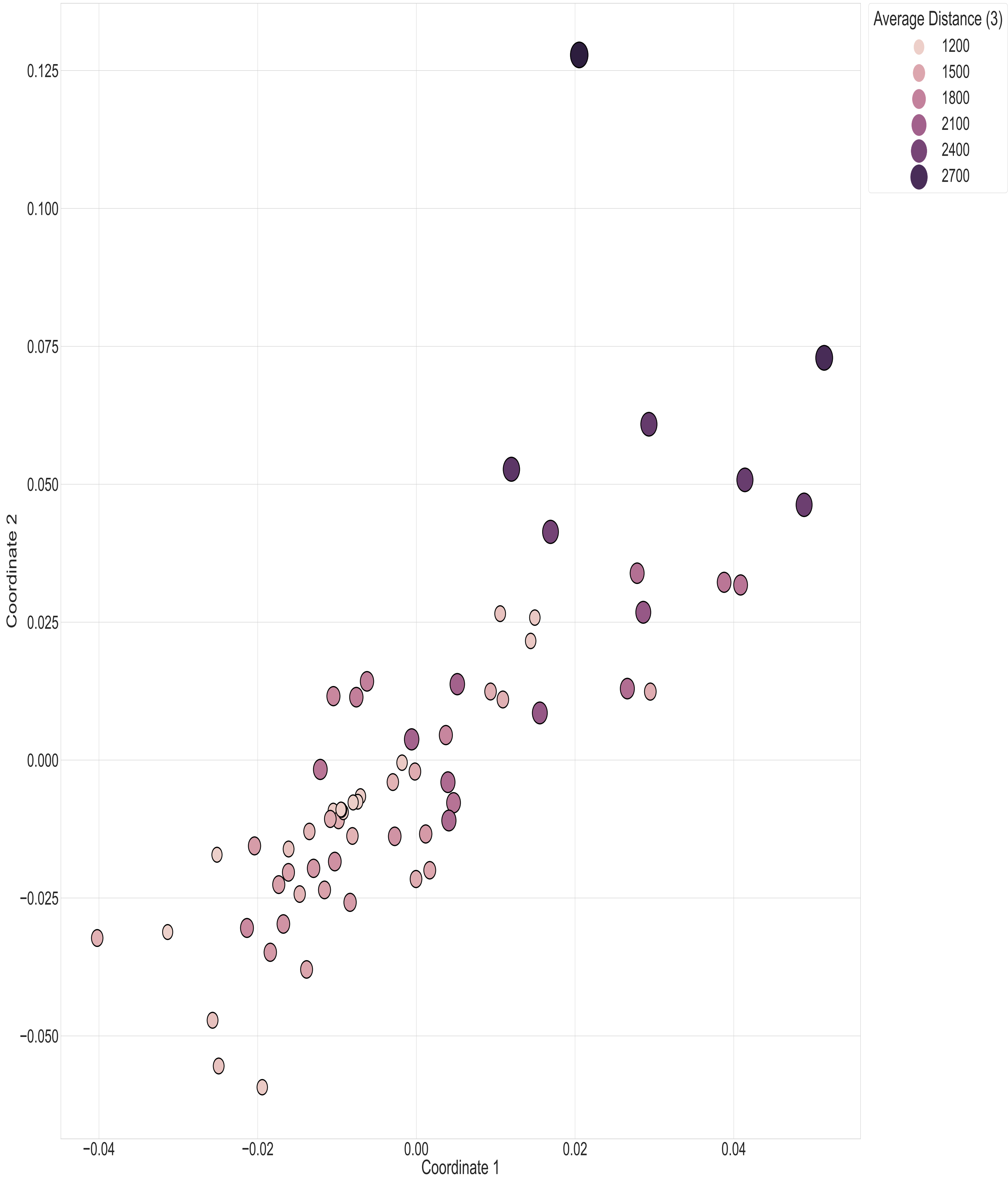
Scatter plot of PKS-NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using chebyshev pairwise distance.
Only PKS-NRPS BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual PKS-NRPS BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S145 :



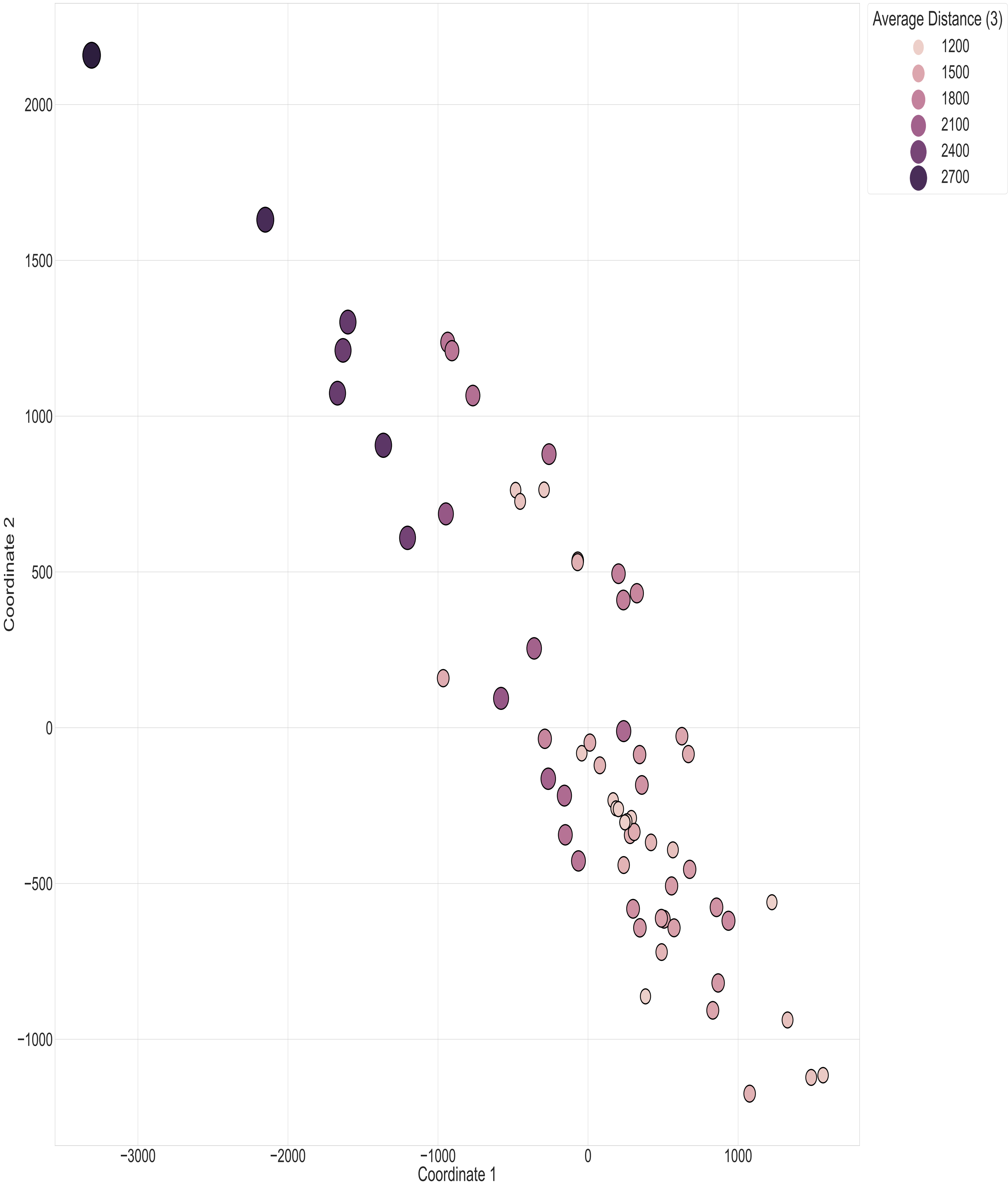
Scatter plot of PKS-NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using l2 pairwise distance.
Only PKS-NRPS BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual PKS-NRPS BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S146 :



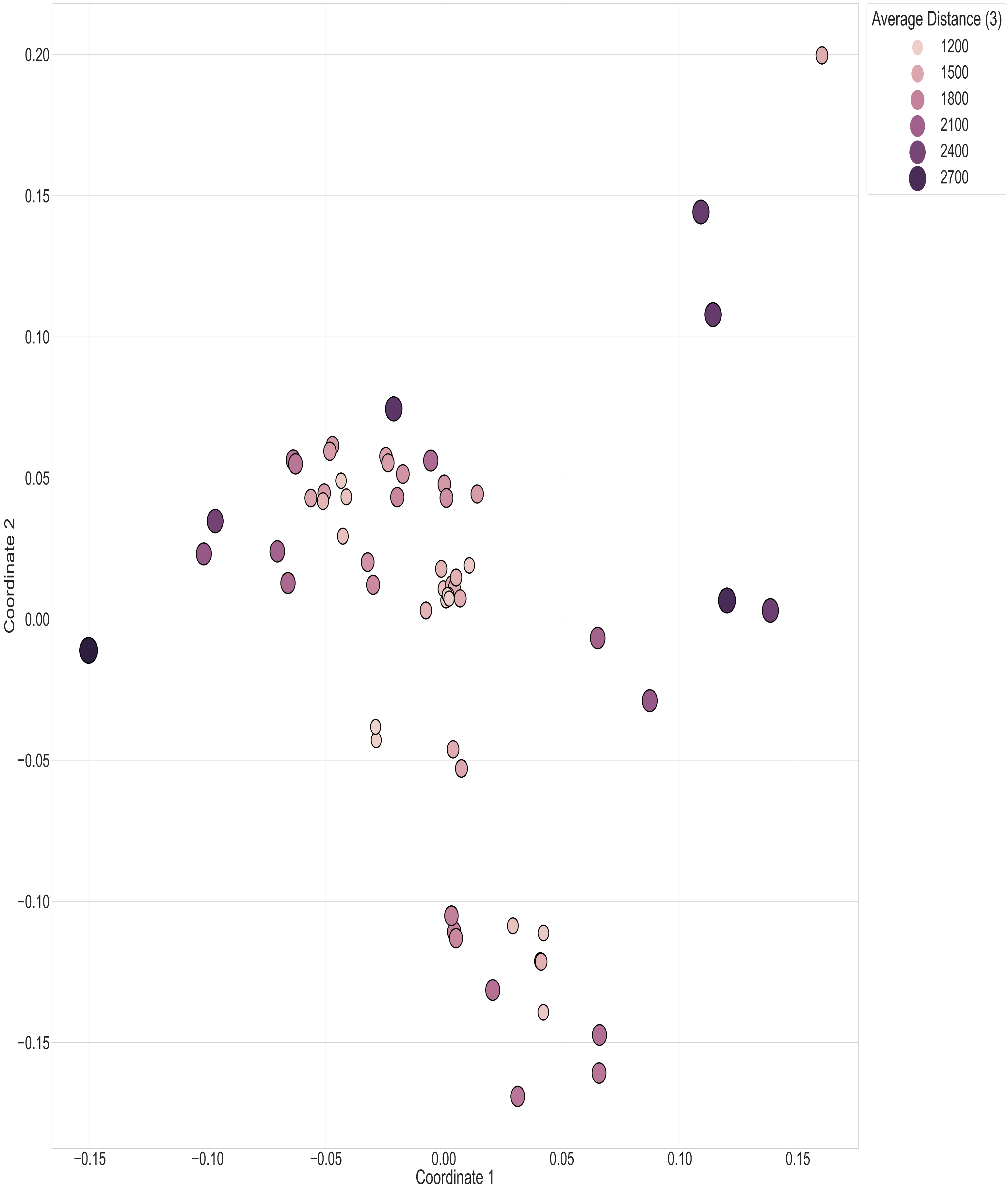
Scatter plot of PKS-NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using braycurtis pairwise distance.
Only PKS-NRPS BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual PKS-NRPS BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S147 :



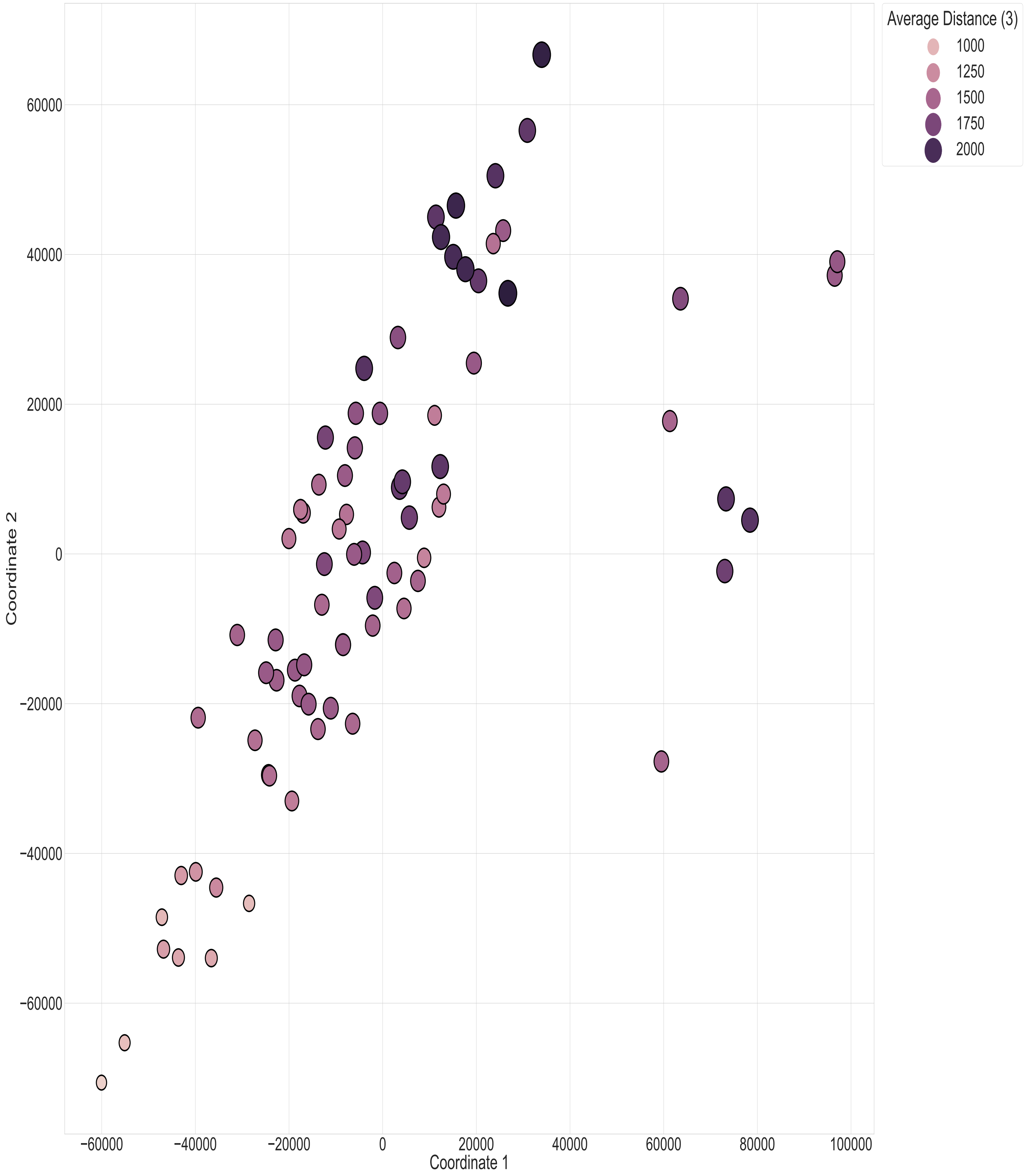
Scatter plot of PKS-NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using canberra pairwise distance.
Only PKS-NRPS BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual PKS-NRPS BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S148 :



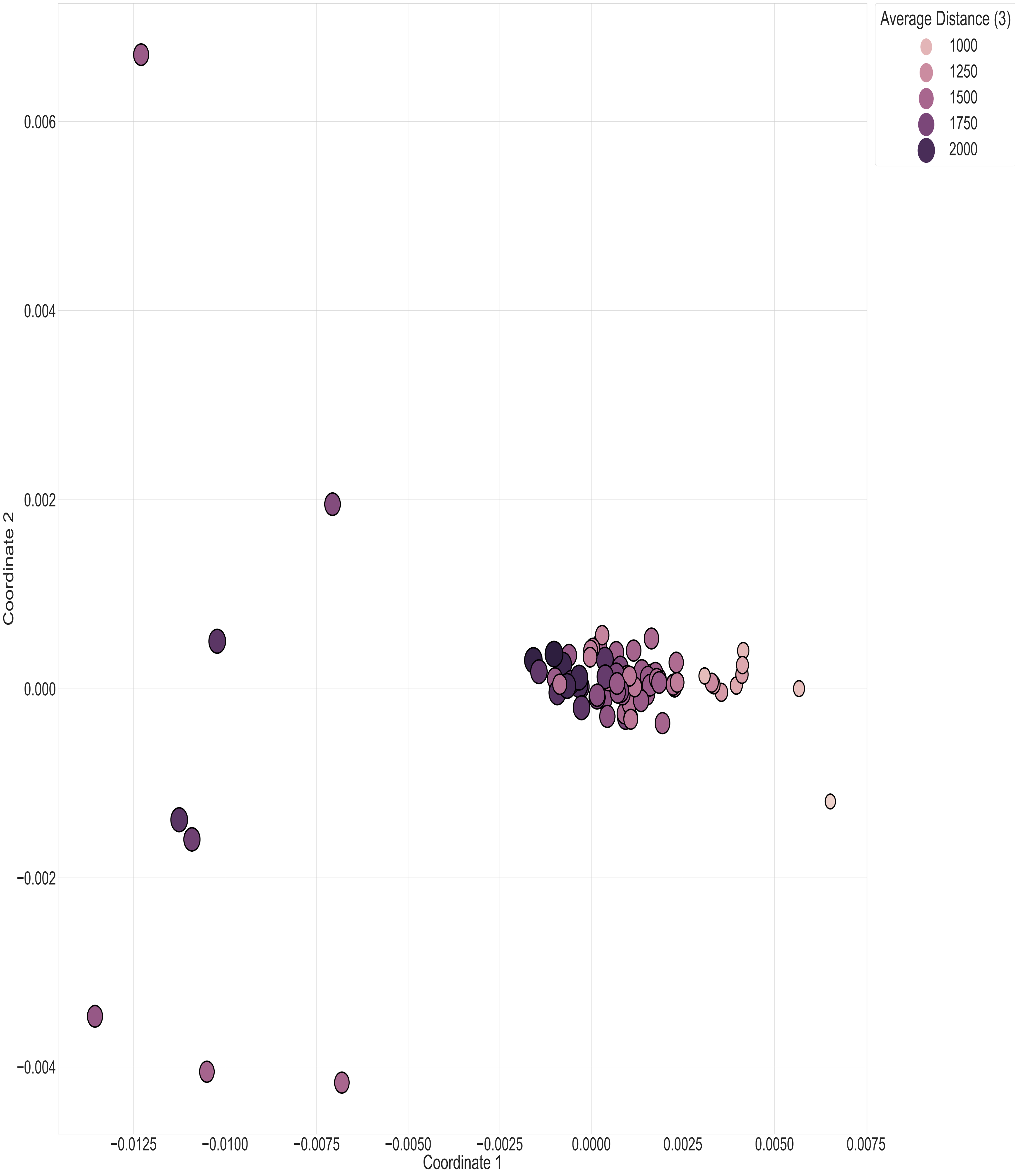
Scatter plot of PKS-NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using correlation pairwise distance.
Only PKS-NRPS BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual PKS-NRPS BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S149 :



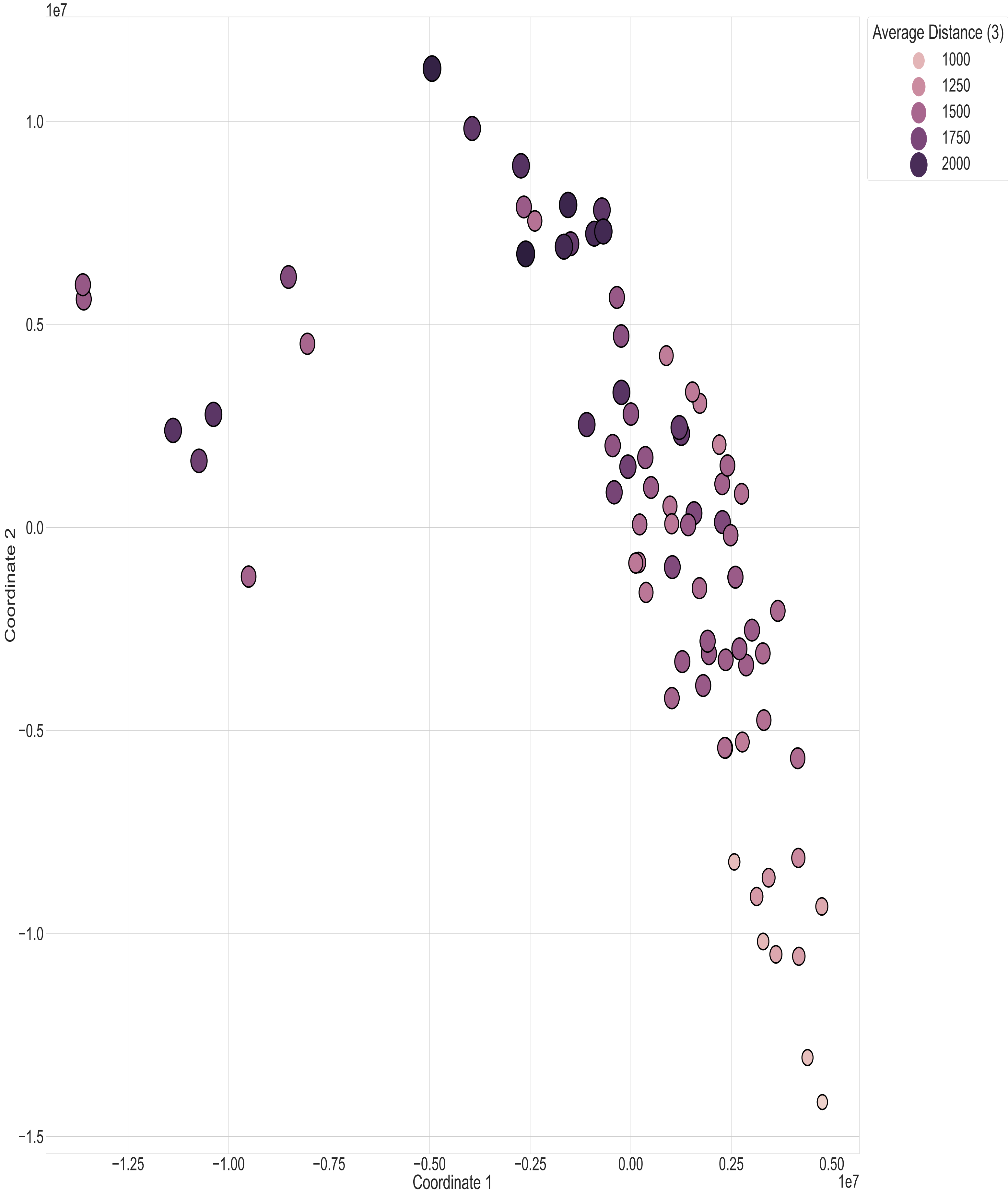
Scatter plot of NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using euclidean pairwise distance.
Only NRPS BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual NRPS BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S150 :



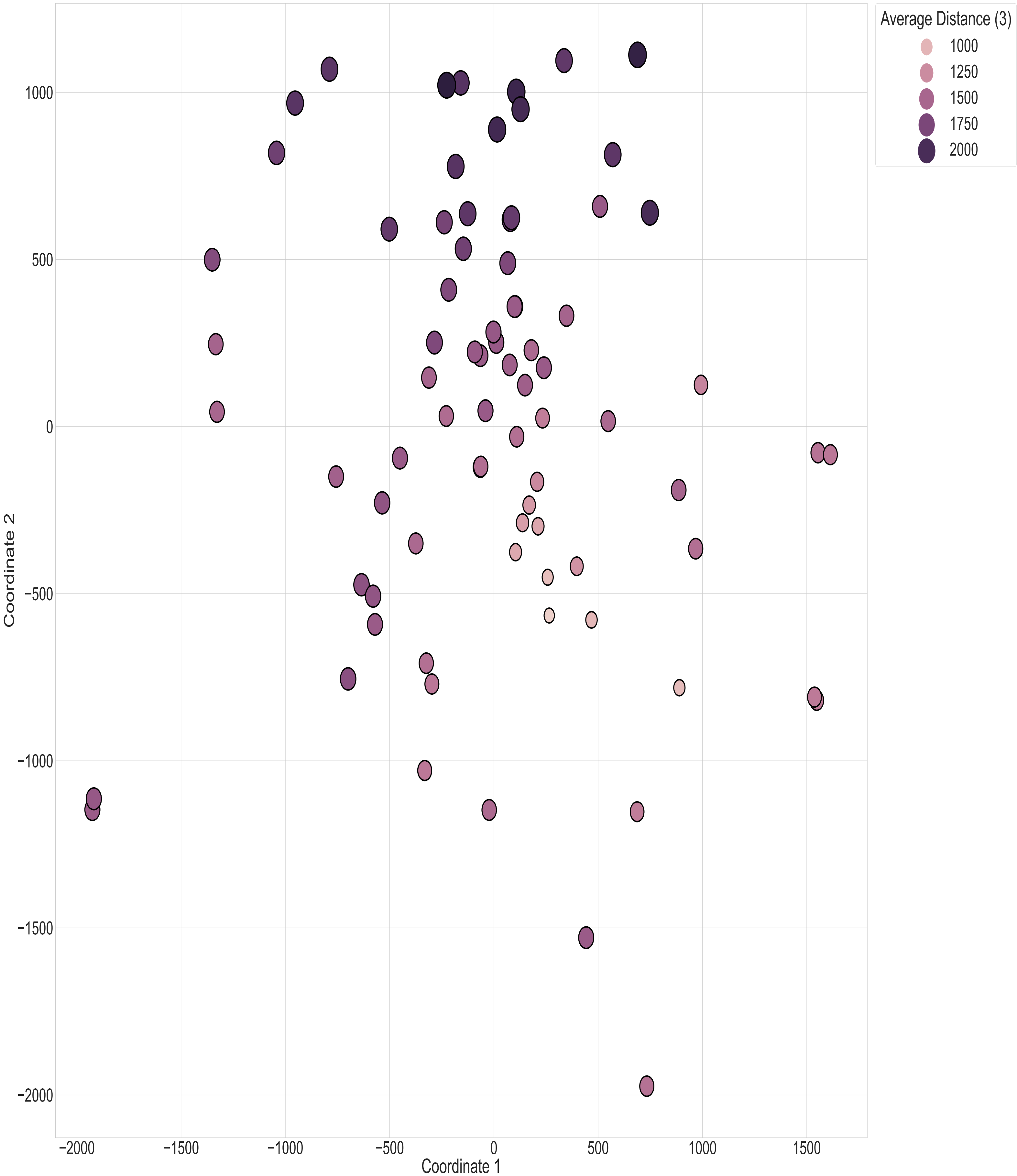
Scatter plot of NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cosine pairwise distance.
Only NRPS BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual NRPS BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S151 :



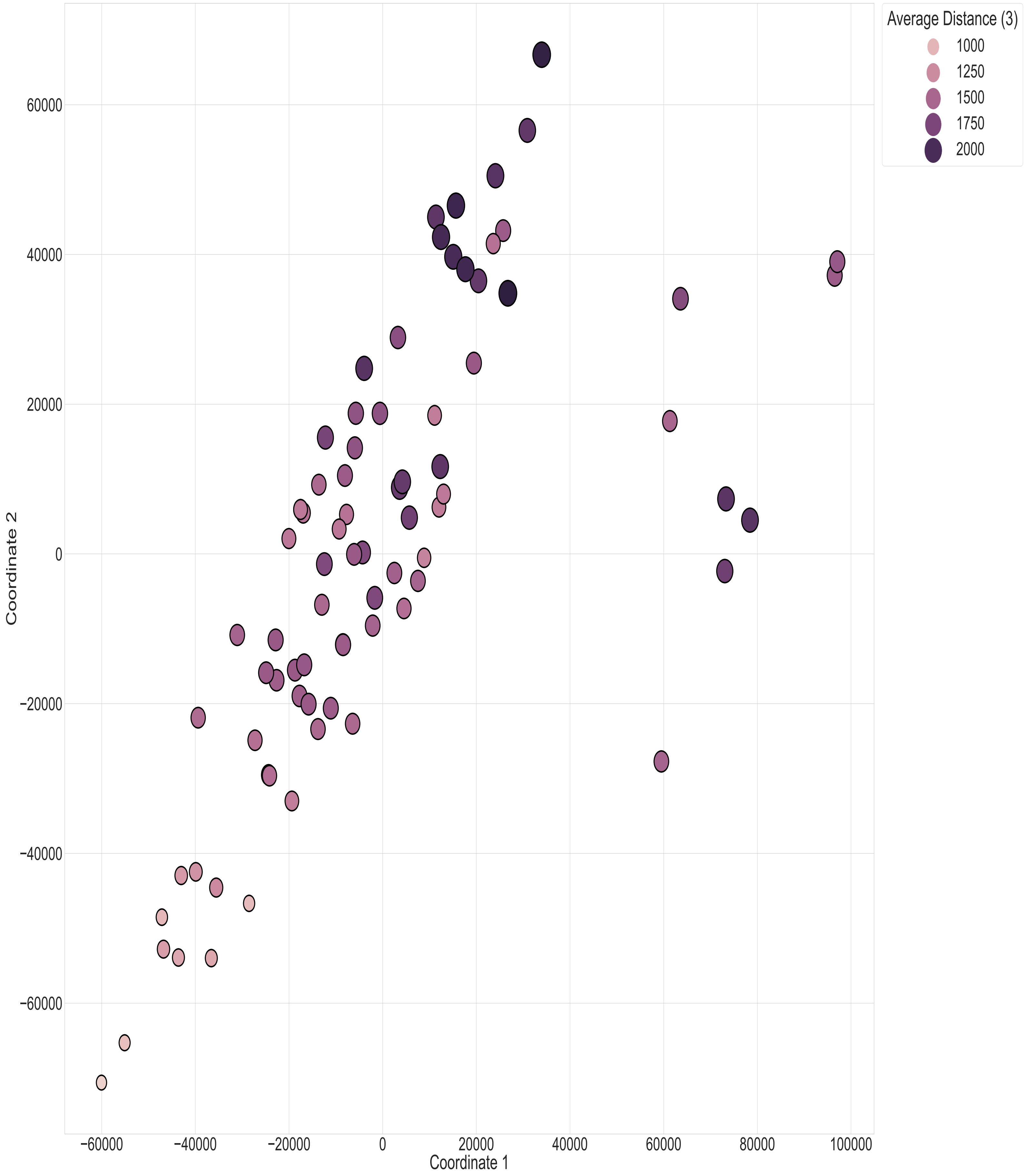
Scatter plot of NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cityblock pairwise distance.
Only NRPS BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual NRPS BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S152 :



Scatter plot of NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using chebyshev pairwise distance.
Only NRPS BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual NRPS BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S153 :

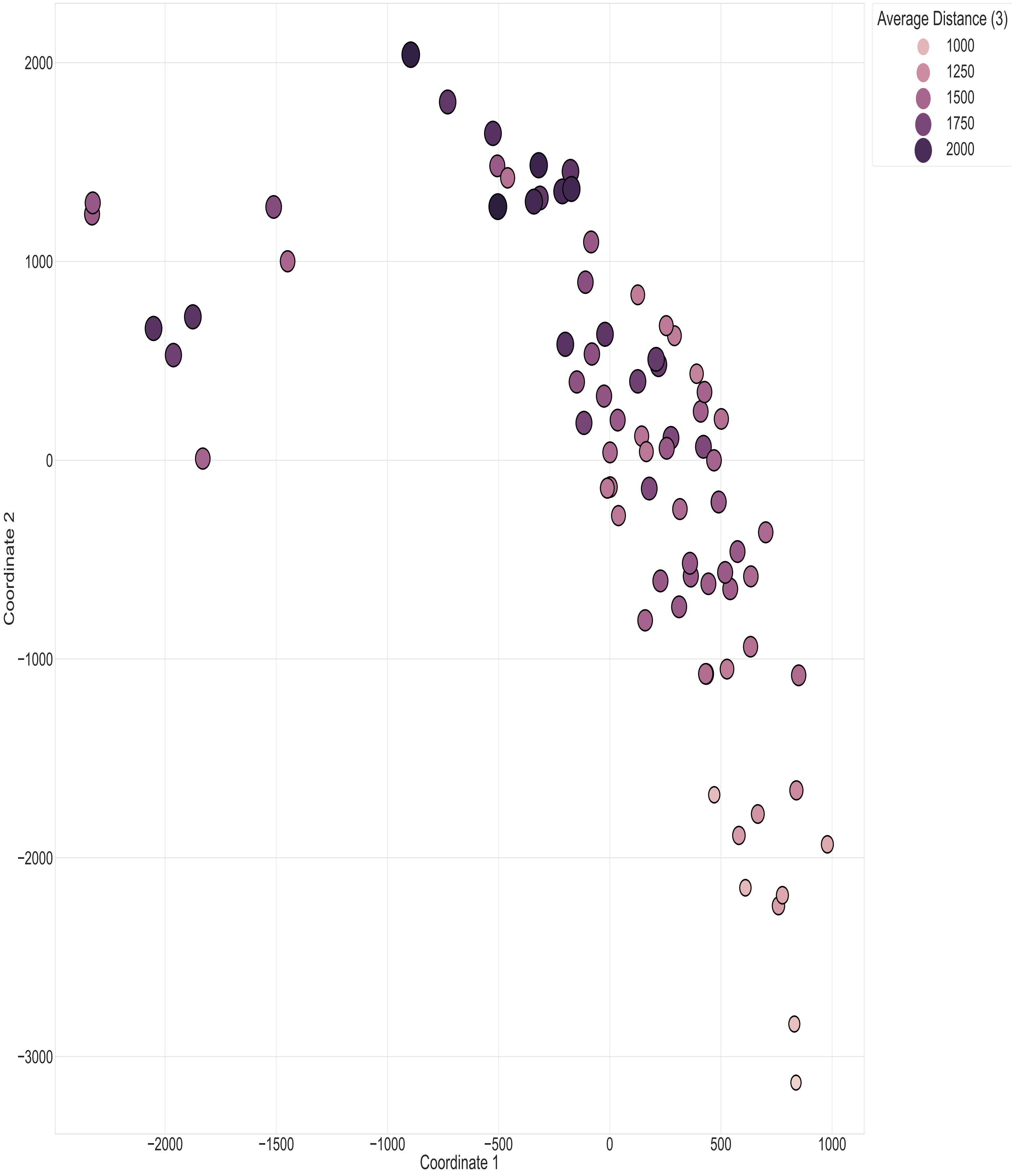


Scatter plot of NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using l2 pairwise distance.
Only NRPS BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual NRPS BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S154 :

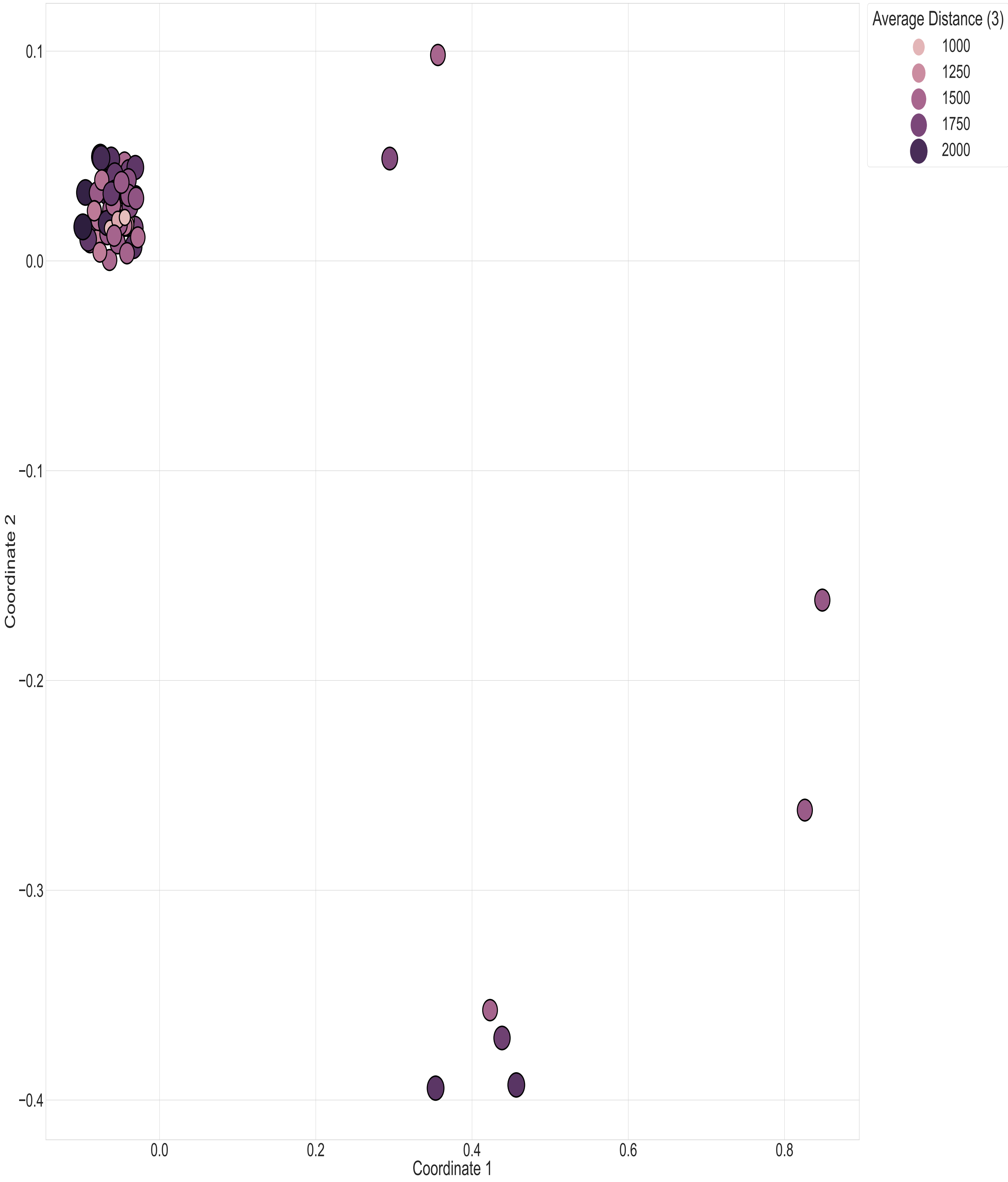
Scatter plot of NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using braycurtis pairwise distance.
Only NRPS BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual NRPS BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S155 :



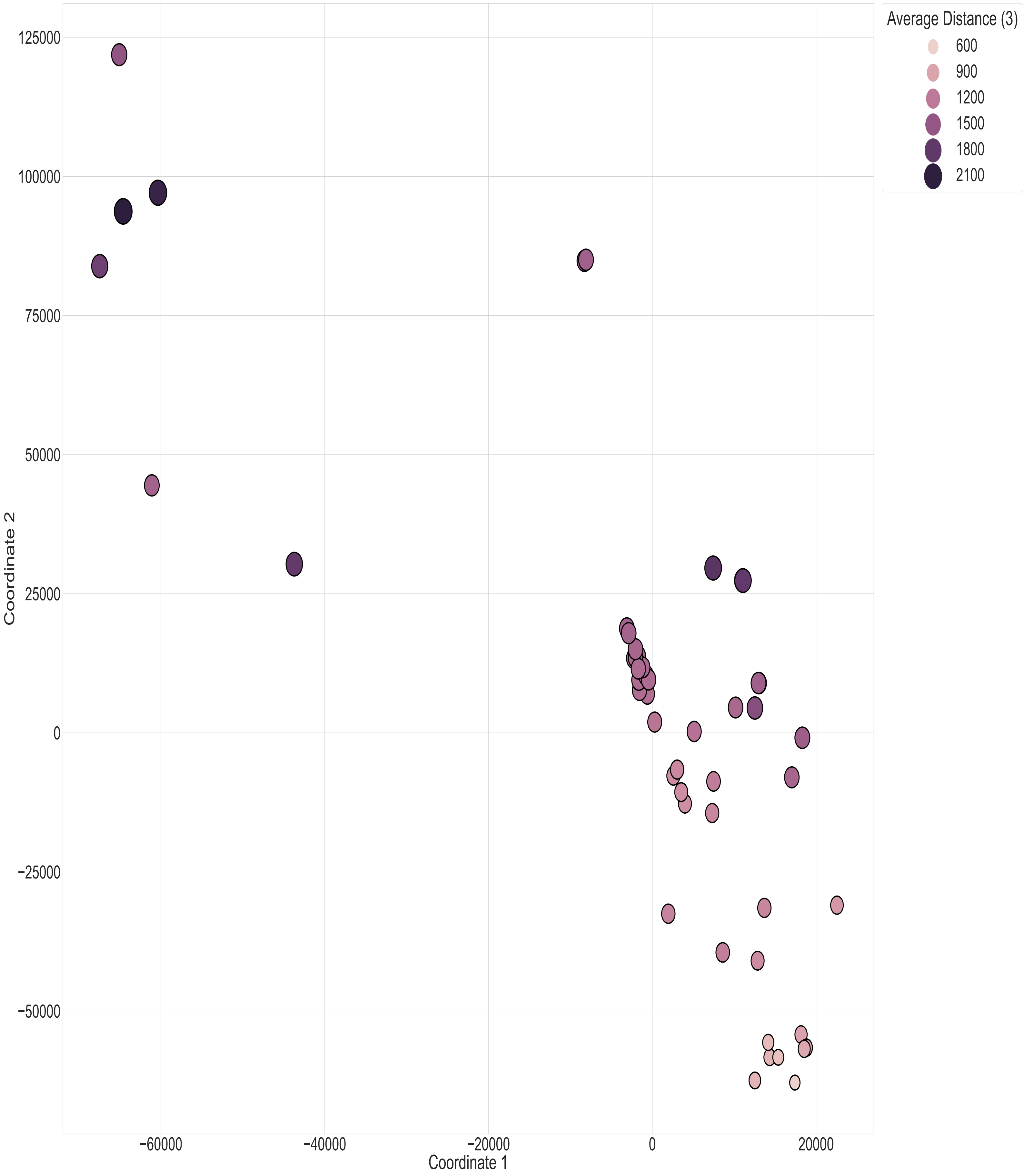
Scatter plot of NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using canberra pairwise distance.
Only NRPS BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual NRPS BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S156 :



Scatter plot of NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using correlation pairwise distance.
Only NRPS BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual NRPS BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S157 :

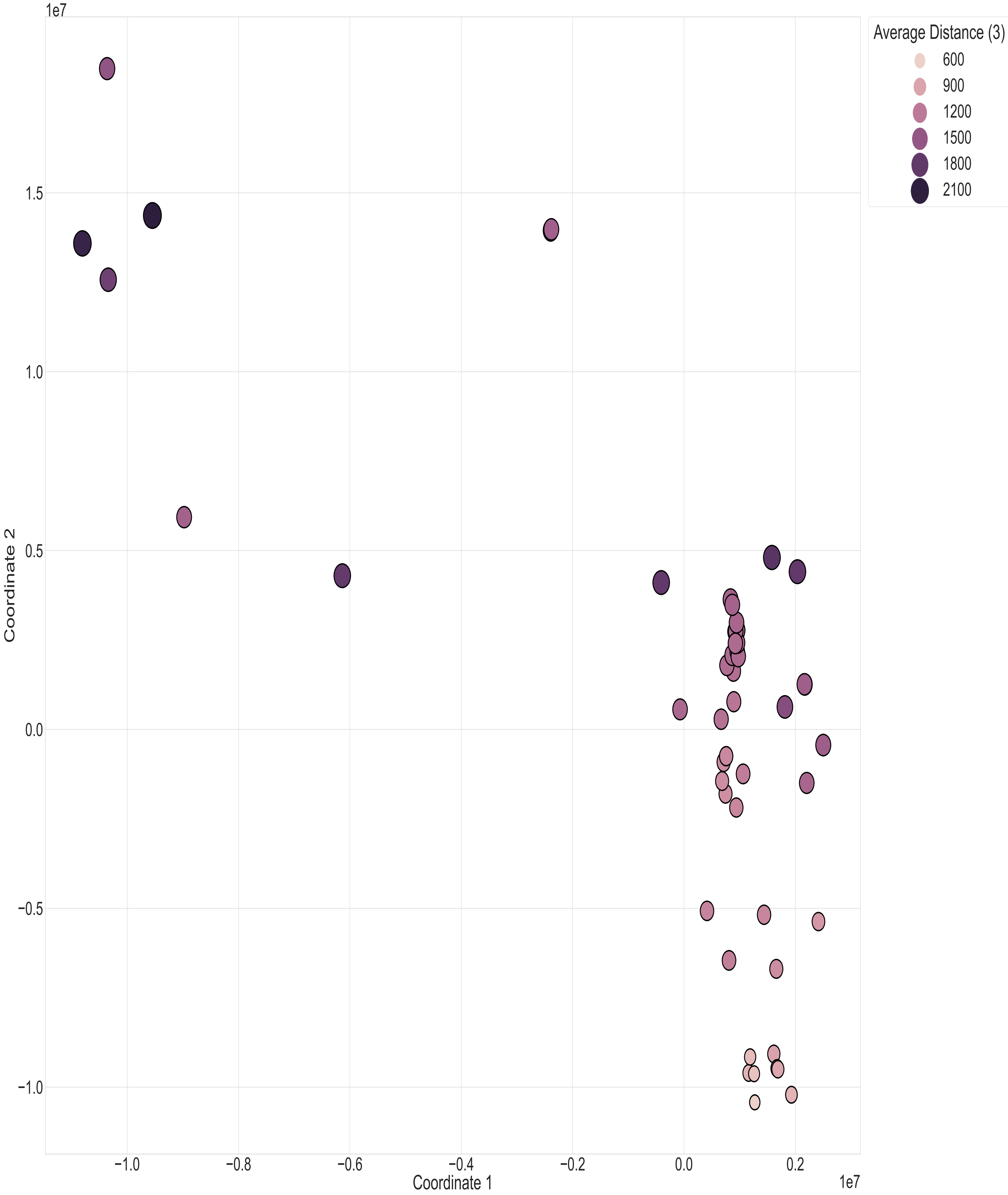


Scatter plot of Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using euclidean pairwise distance.
Only Other BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual Other BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S158 :

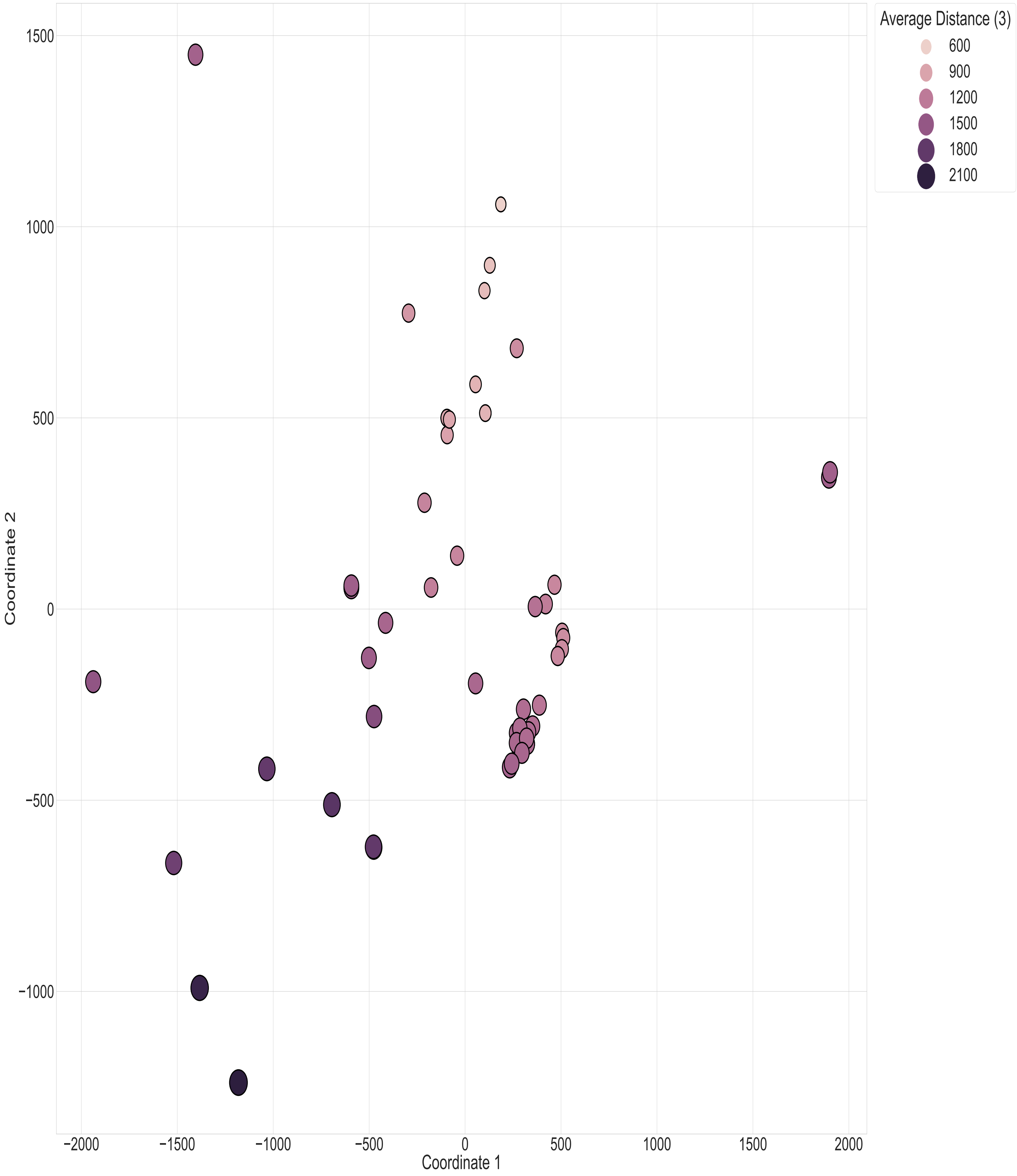
Scatter plot of Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cosine pairwise distance.
Only Other BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual Other BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S159 :



Scatter plot of Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cityblock pairwise distance.
Only Other BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual Other BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S160 :



Scatter plot of Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using chebyshev pairwise distance.
Only Other BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual Other BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

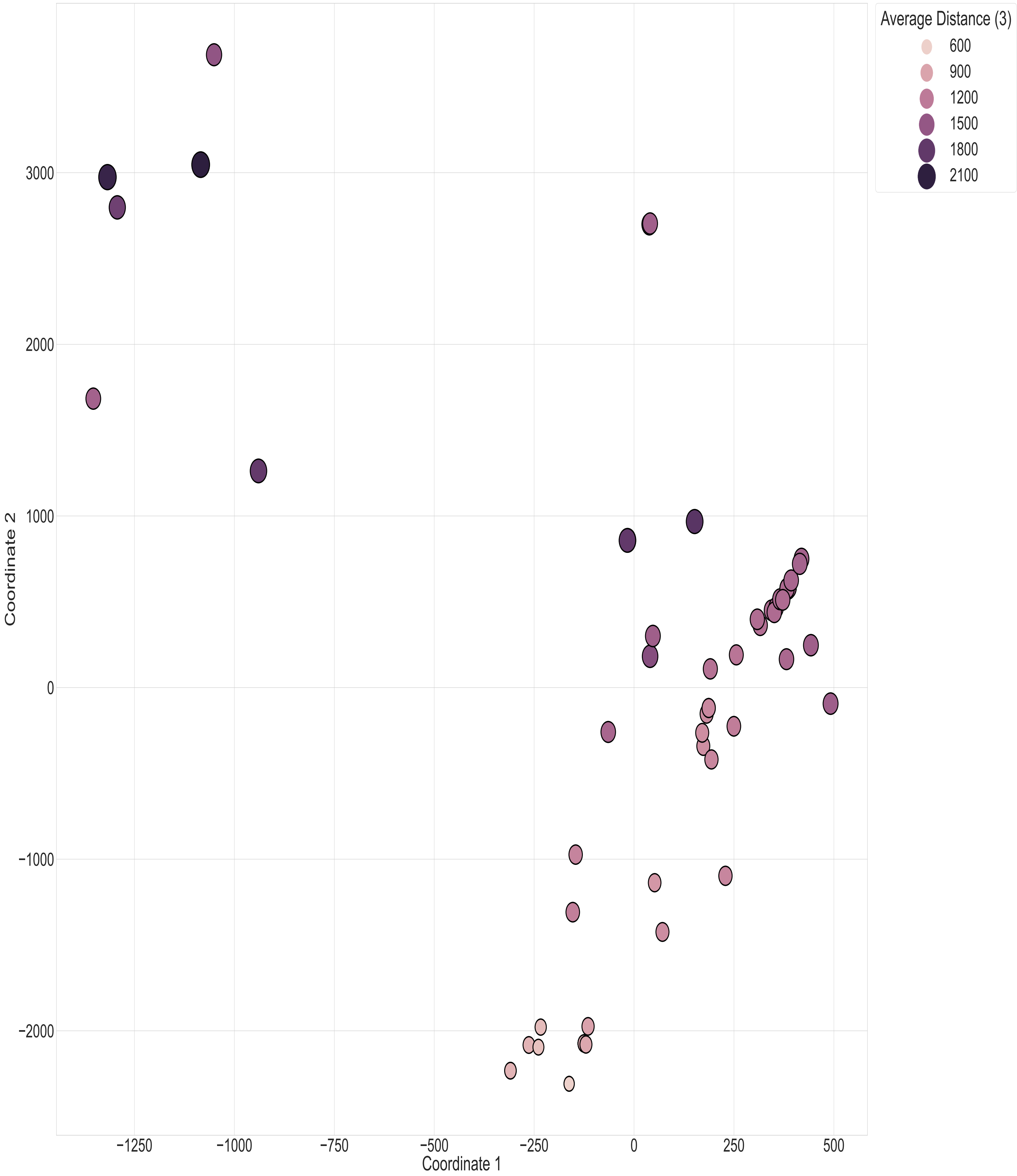
Figure – S161 :

Scatter plot of Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using l2 pairwise distance.
Only Other BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual Other BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S162 :

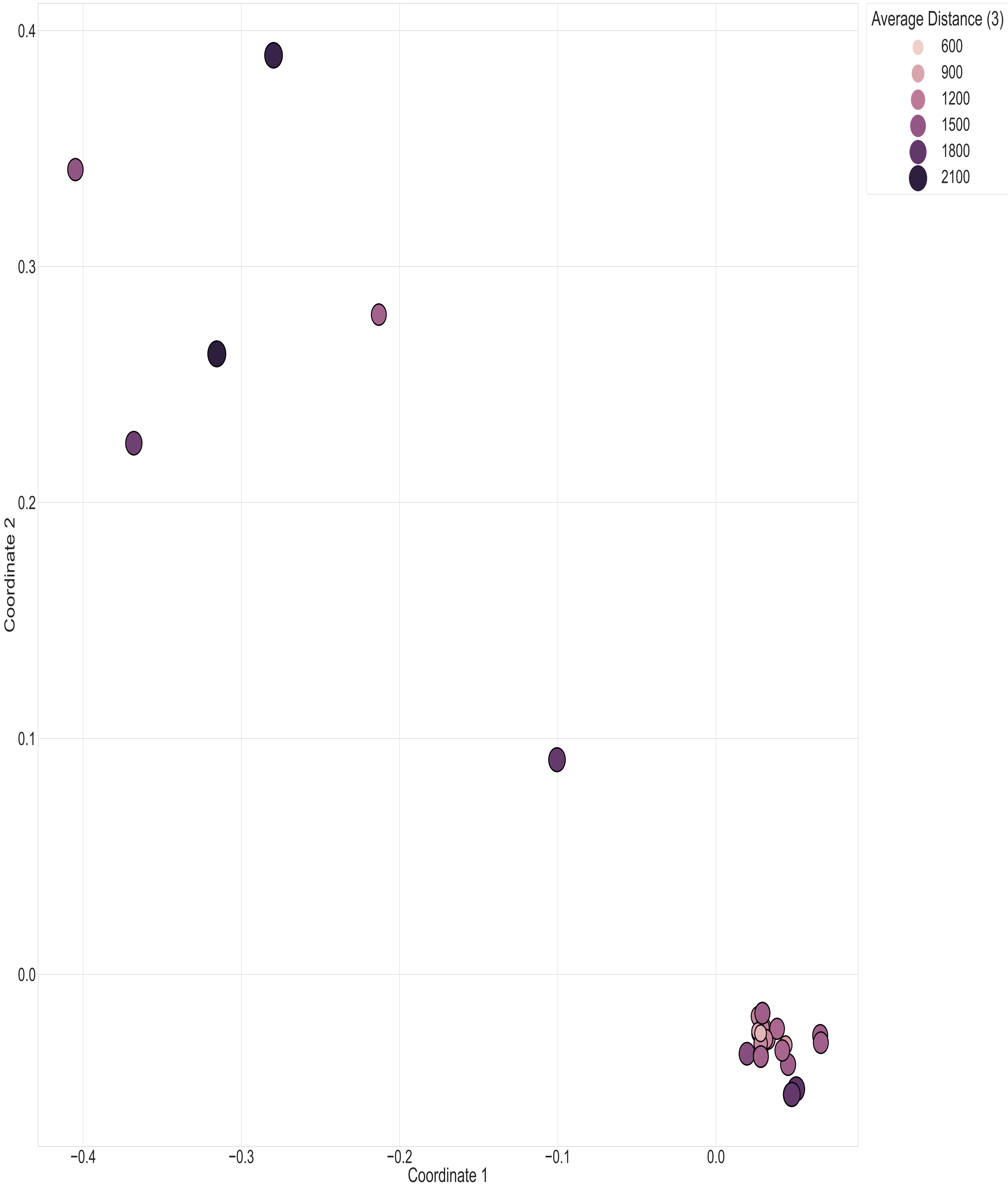
Scatter plot of Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using braycurtis pairwise distance.
Only Other BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual Other BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S163 :



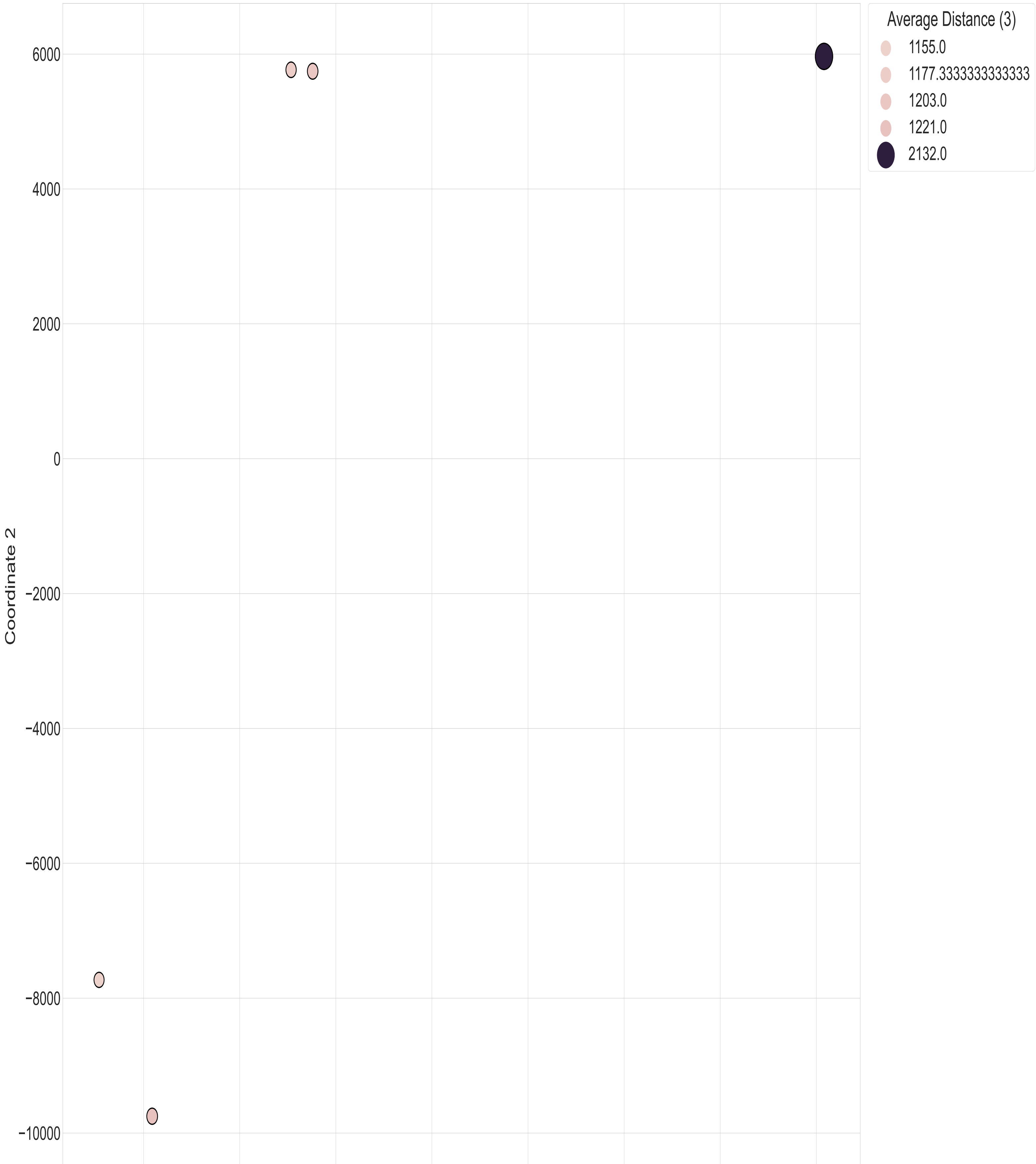
Scatter plot of Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using canberra pairwise distance.
Only Other BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual Other BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S164 :



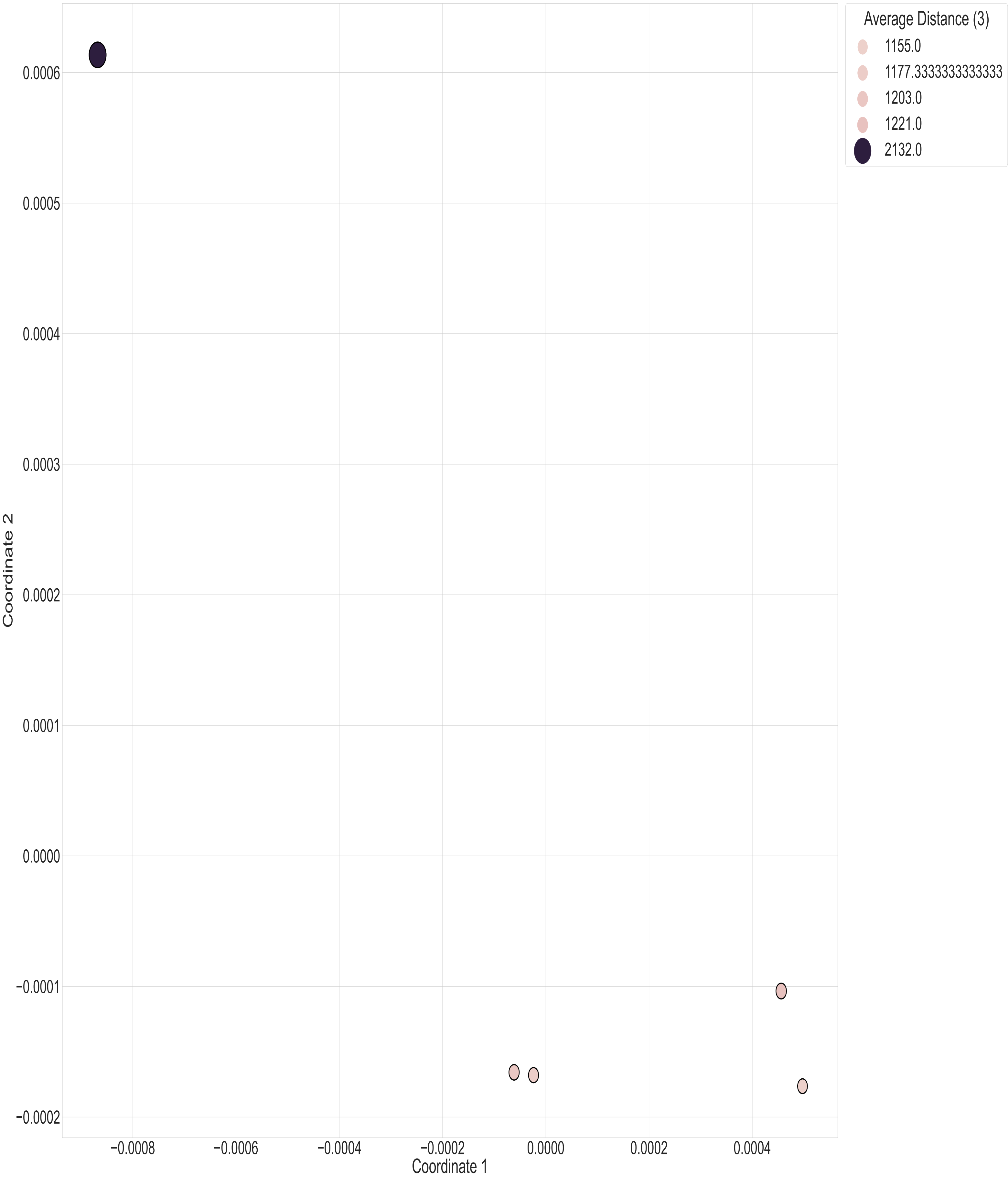
Scatter plot of Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using correlation pairwise distance. Only Other BGCs were considered for this analysis. Each dot represents an individual BGC. The hue and size of the dots were scaled based on the average distance of the individual Other BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S165 :



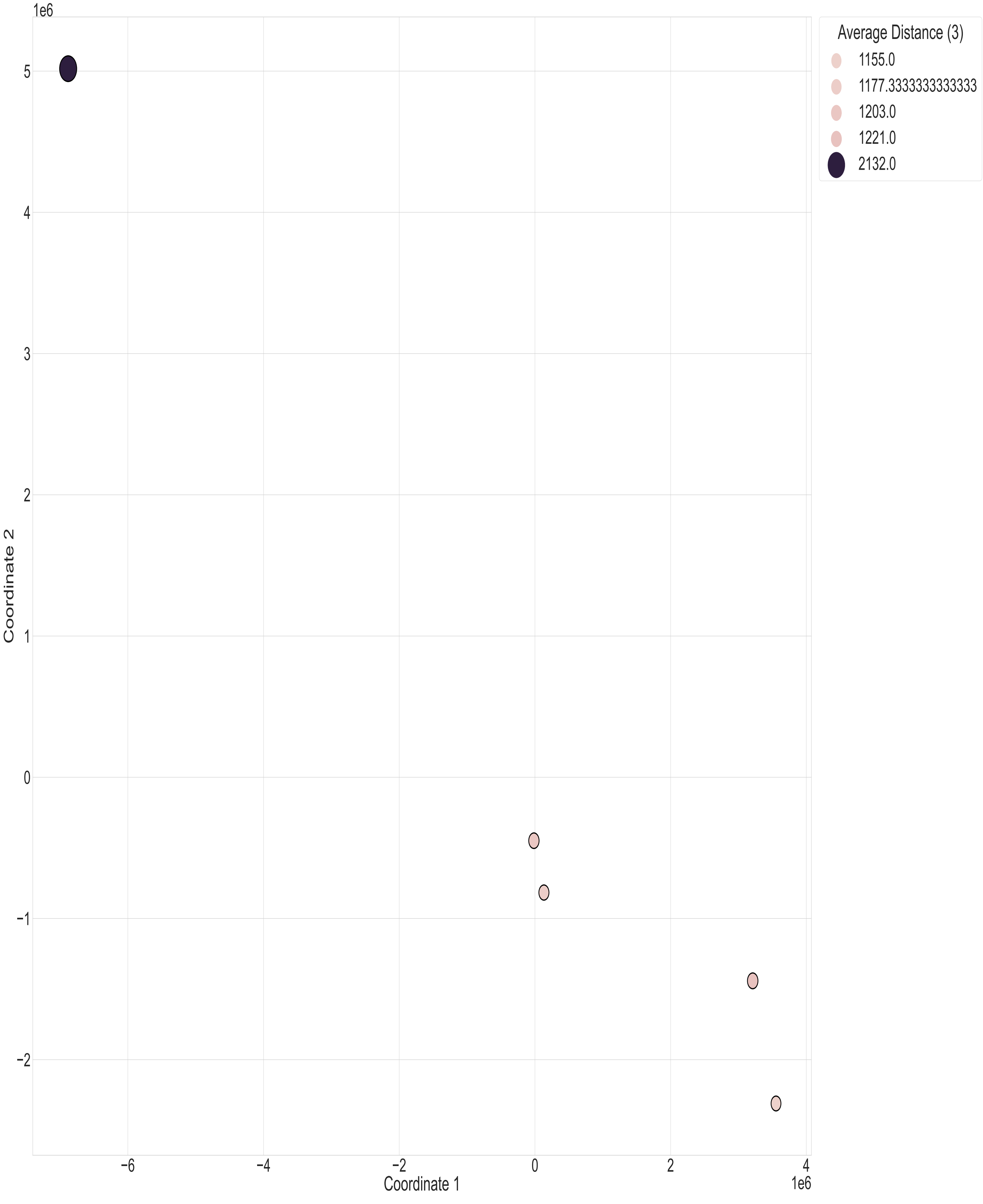
Scatter plot of Saccharides *Micromonosporaceae* BGCs analyzed via multidimensional scaling using euclidean pairwise distance. Only Saccharides BGCs were considered for this analysis. Each dot represents an individual BGC. The hue and size of the dots were scaled based on the average distance of the individual Saccharides BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S166 :



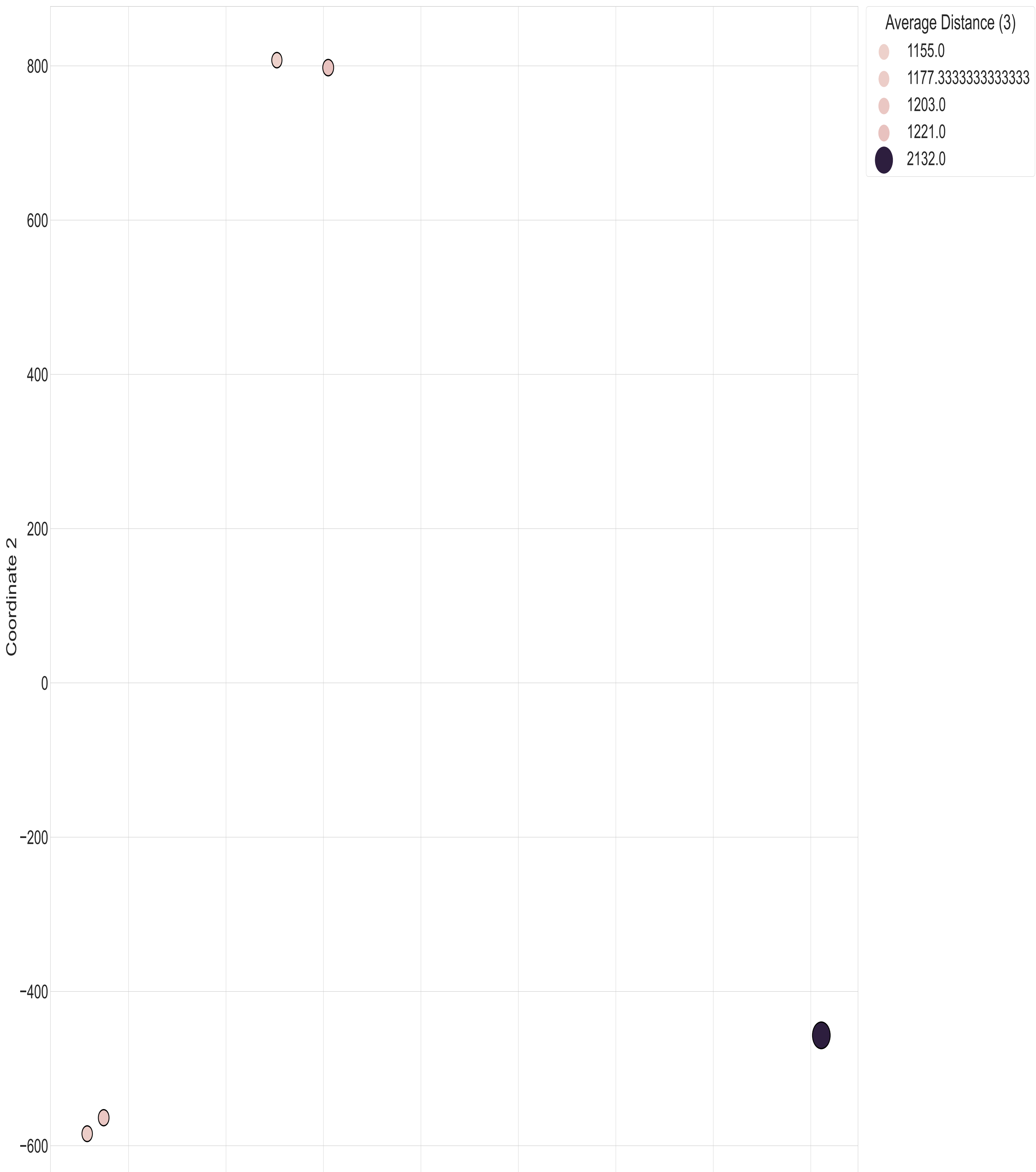
Scatter plot of Saccharides *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cosine pairwise distance.
Only Saccharides BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual Saccharides BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S167 :



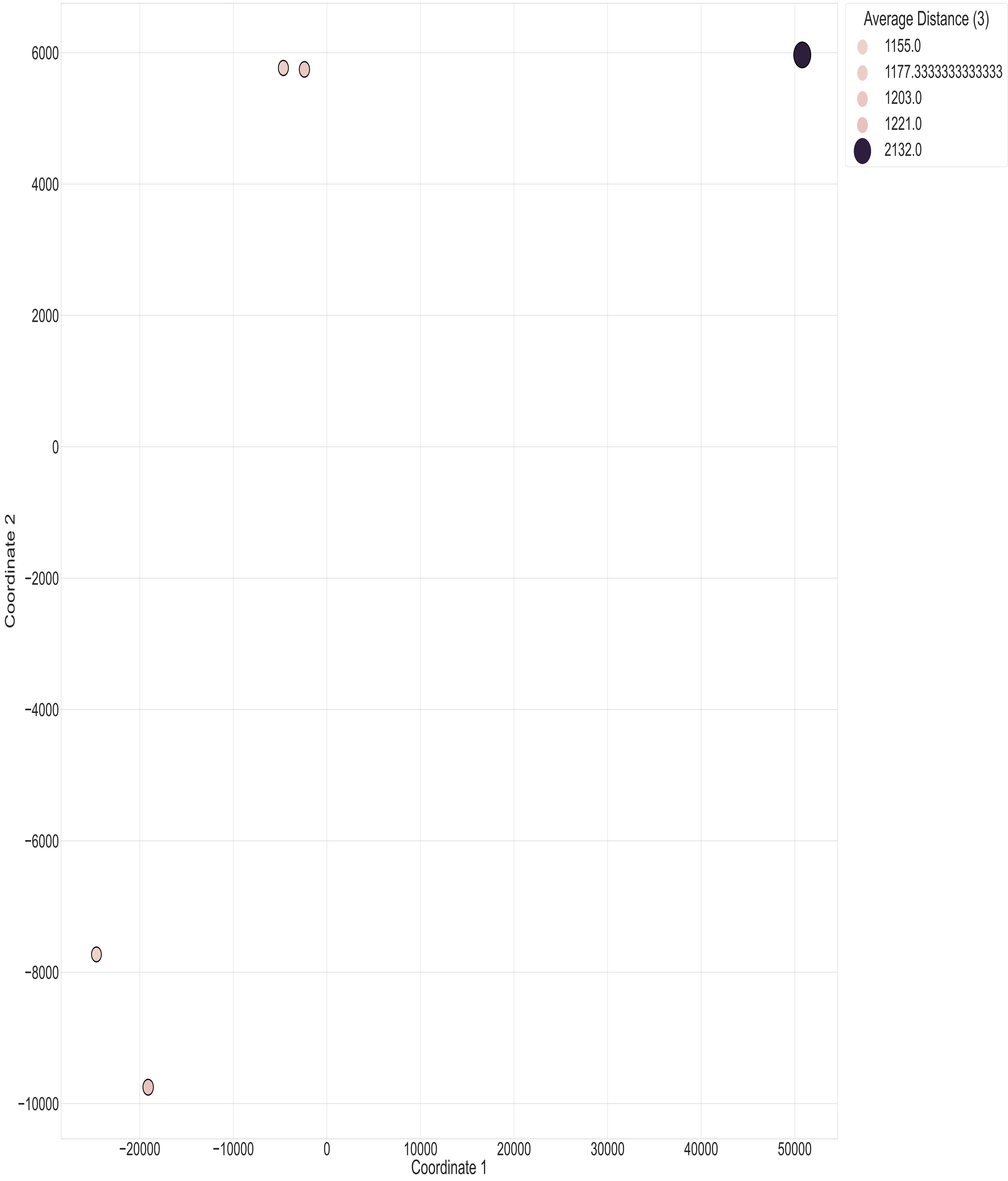
Scatter plot of Saccharides *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cityblock pairwise distance.
Only Saccharides BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual Saccharides BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S168 :



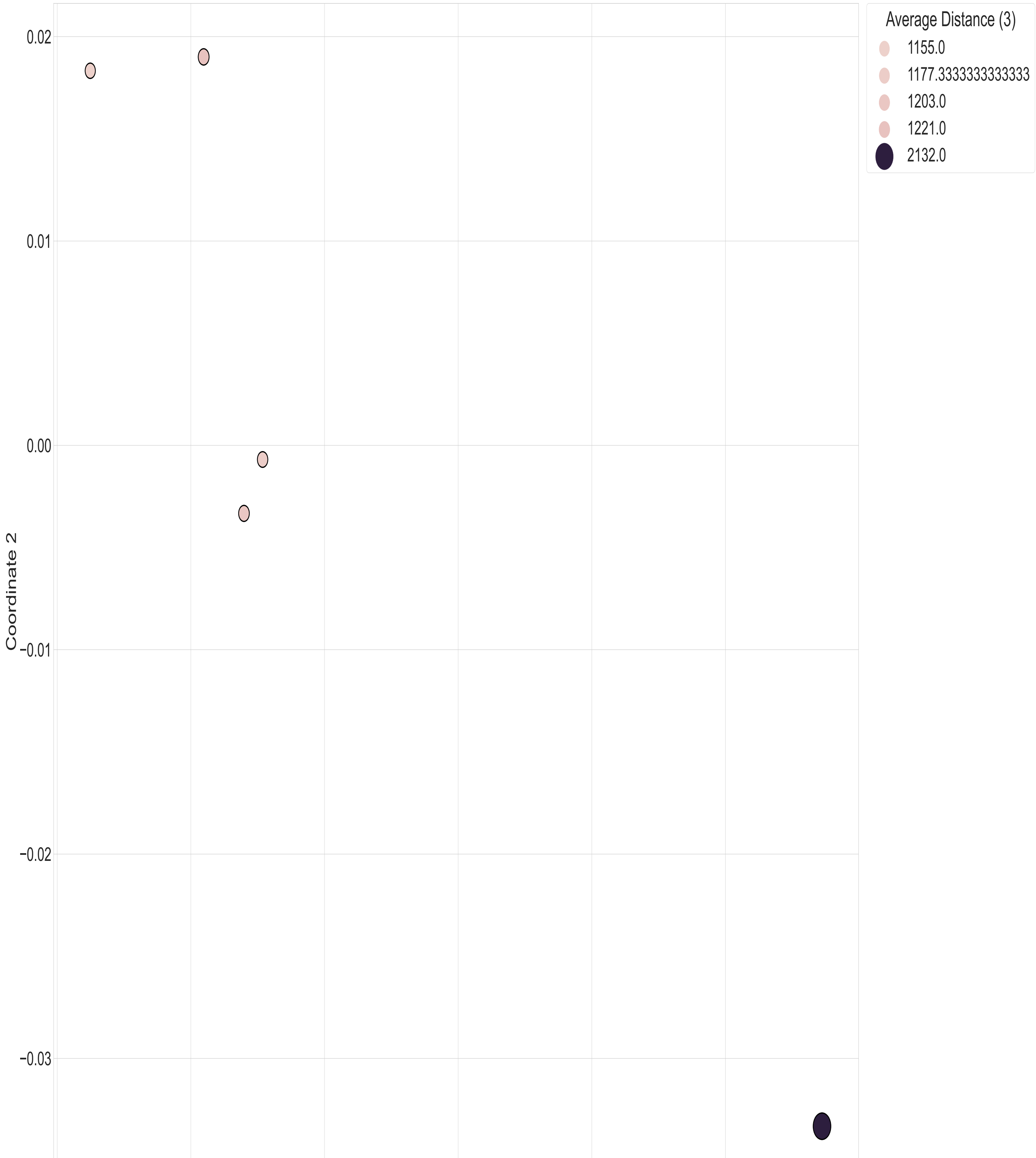
Scatter plot of Saccharides *Micromonosporaceae* BGCs analyzed via multidimensional scaling using chebyshev pairwise distance.
Only Saccharides BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual Saccharides BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S169 :



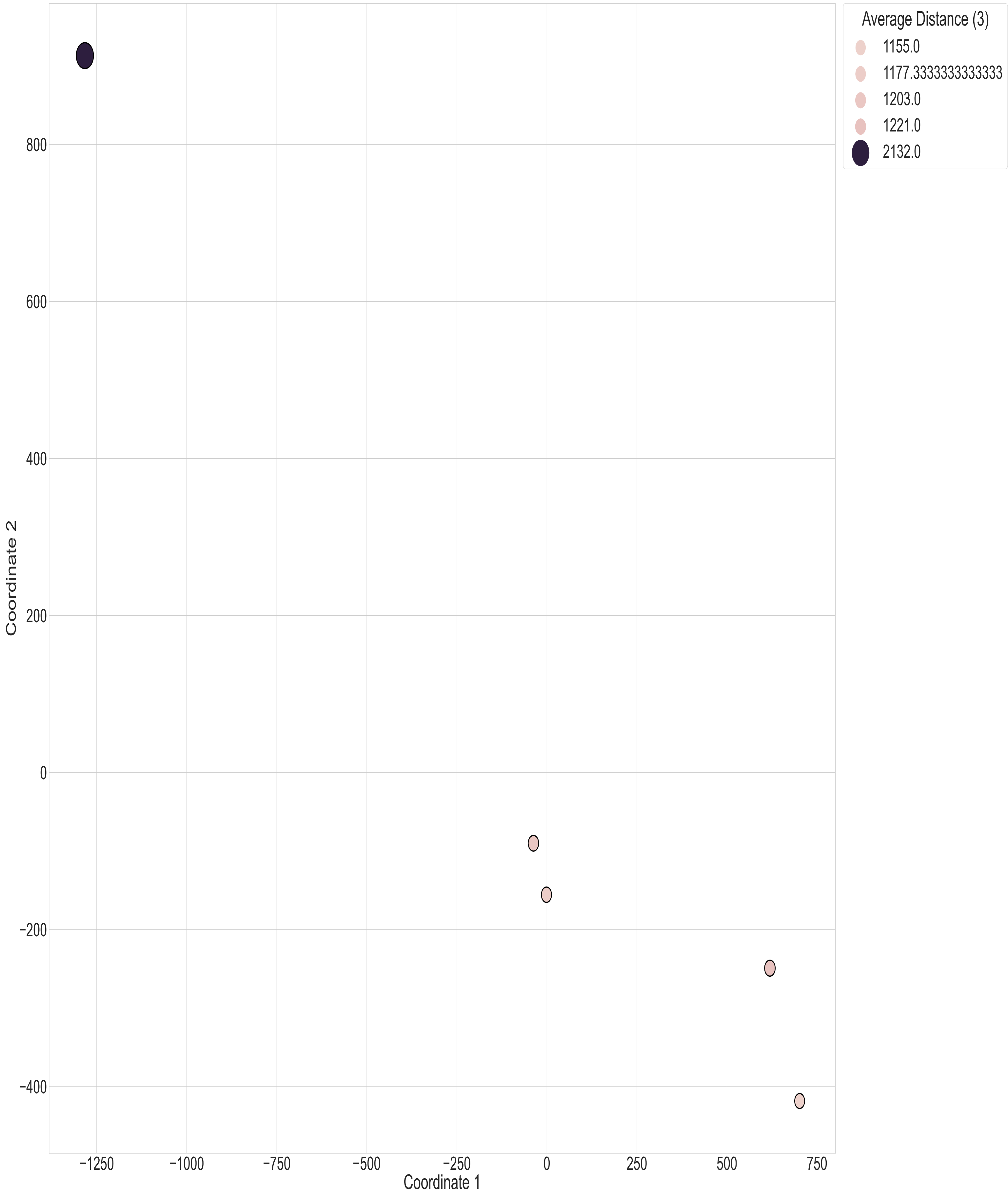
Scatter plot of Saccharides *Micromonosporaceae* BGCs analyzed via multidimensional scaling using l2 pairwise distance. Only Saccharides BGCs were considered for this analysis. Each dot represents an individual BGC. The hue and size of the dots were scaled based on the average distance of the individual Saccharides BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S170 :



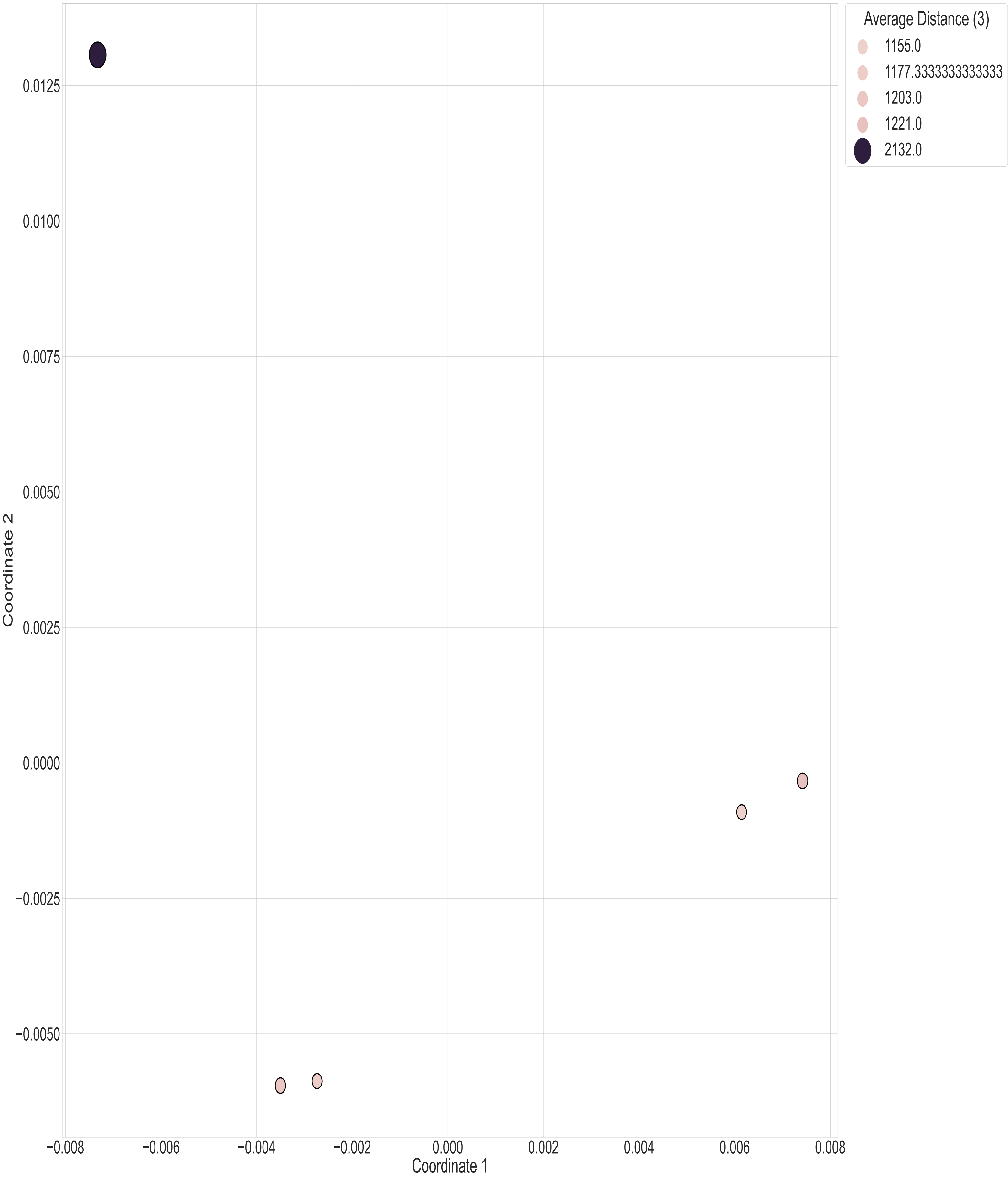
Scatter plot of Saccharides *Micromonosporaceae* BGCs analyzed via multidimensional scaling using braycurtis pairwise distance.
Only Saccharides BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual Saccharides BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S171 :



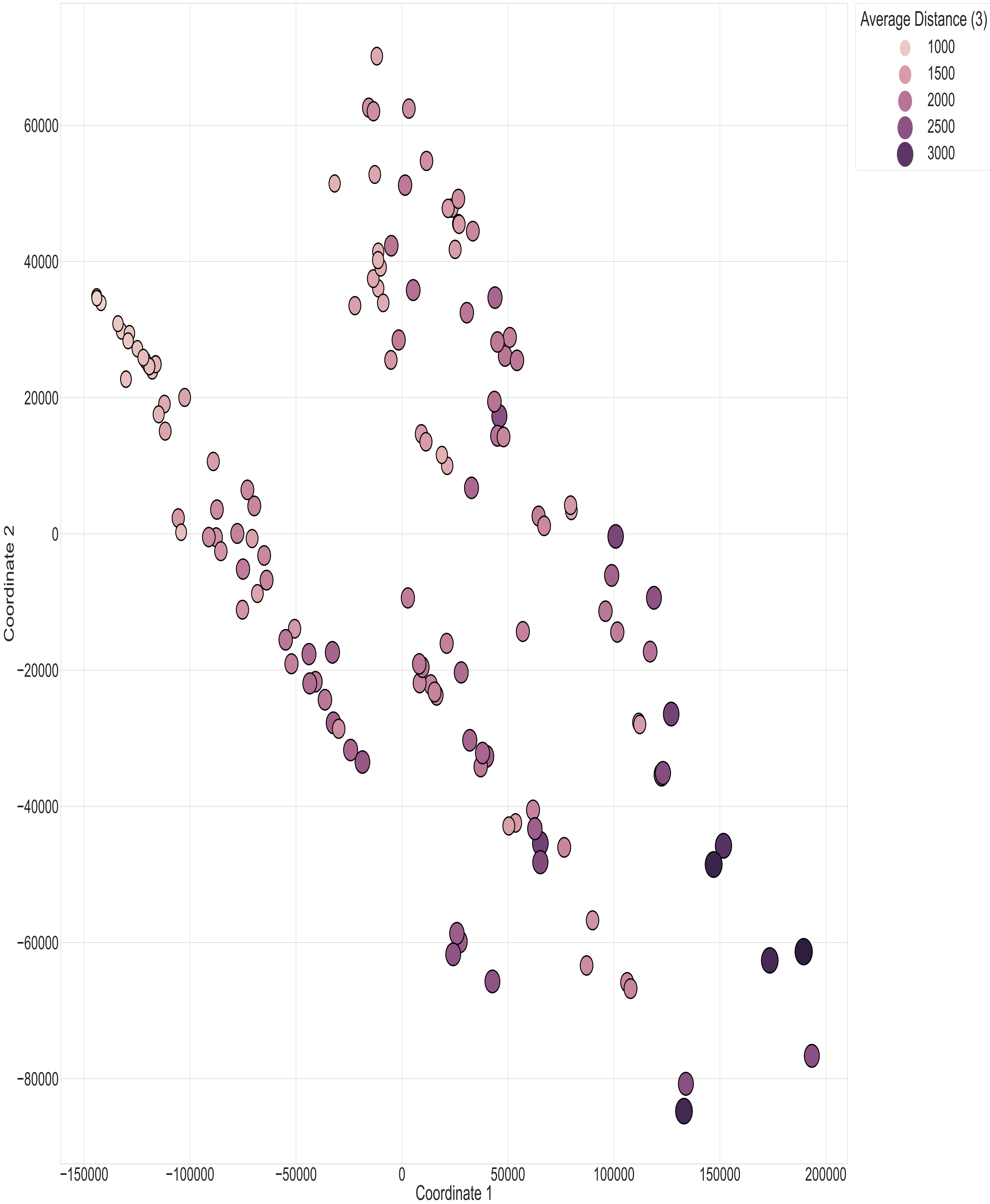
Scatter plot of Saccharides *Micromonosporaceae* BGCs analyzed via multidimensional scaling using canberra pairwise distance.
Only Saccharides BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual Saccharides BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S172 :



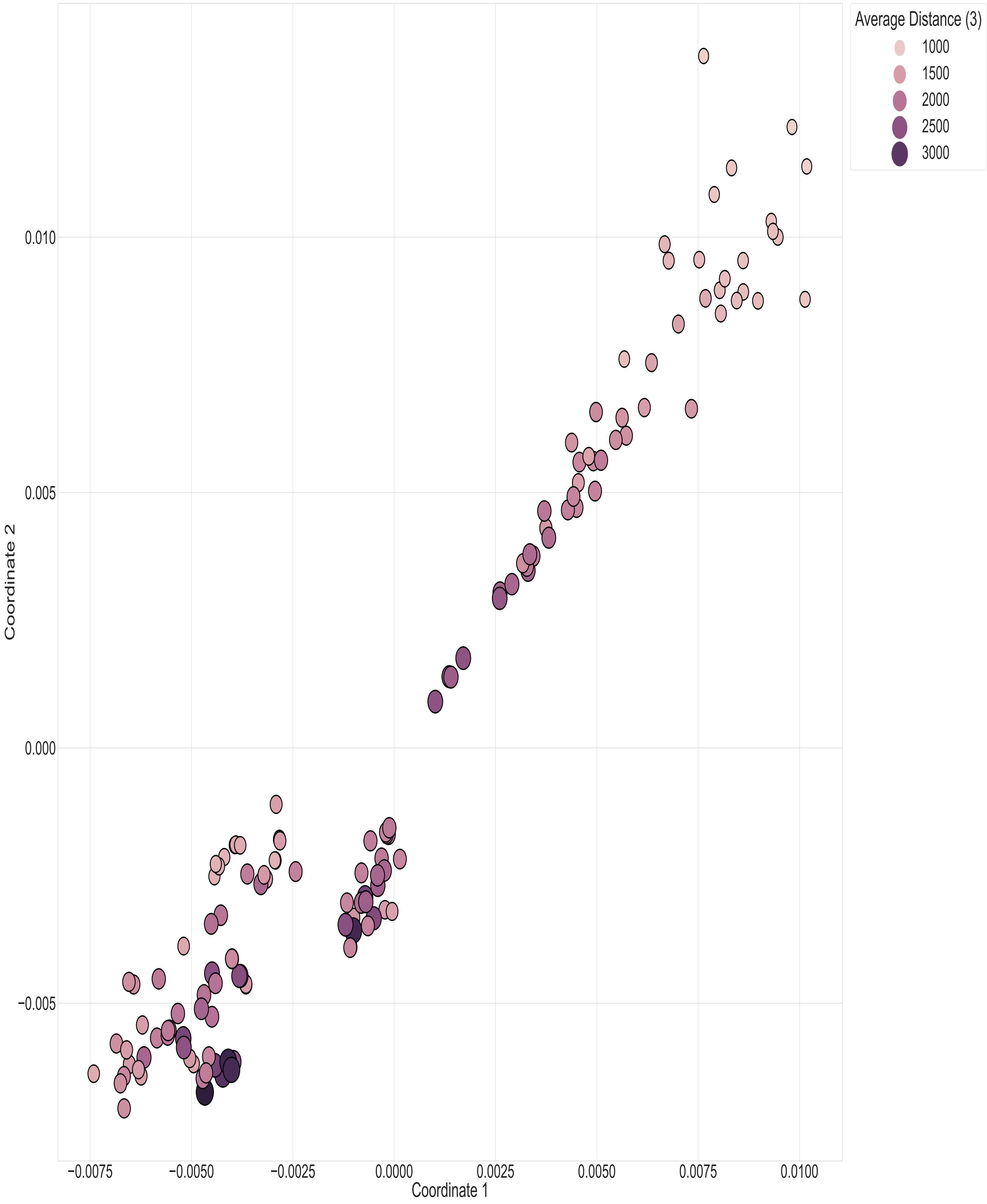
Scatter plot of Saccharides *Micromonosporaceae* BGCs analyzed via multidimensional scaling using correlation pairwise distance.
Only Saccharides BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual Saccharides BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S173 :



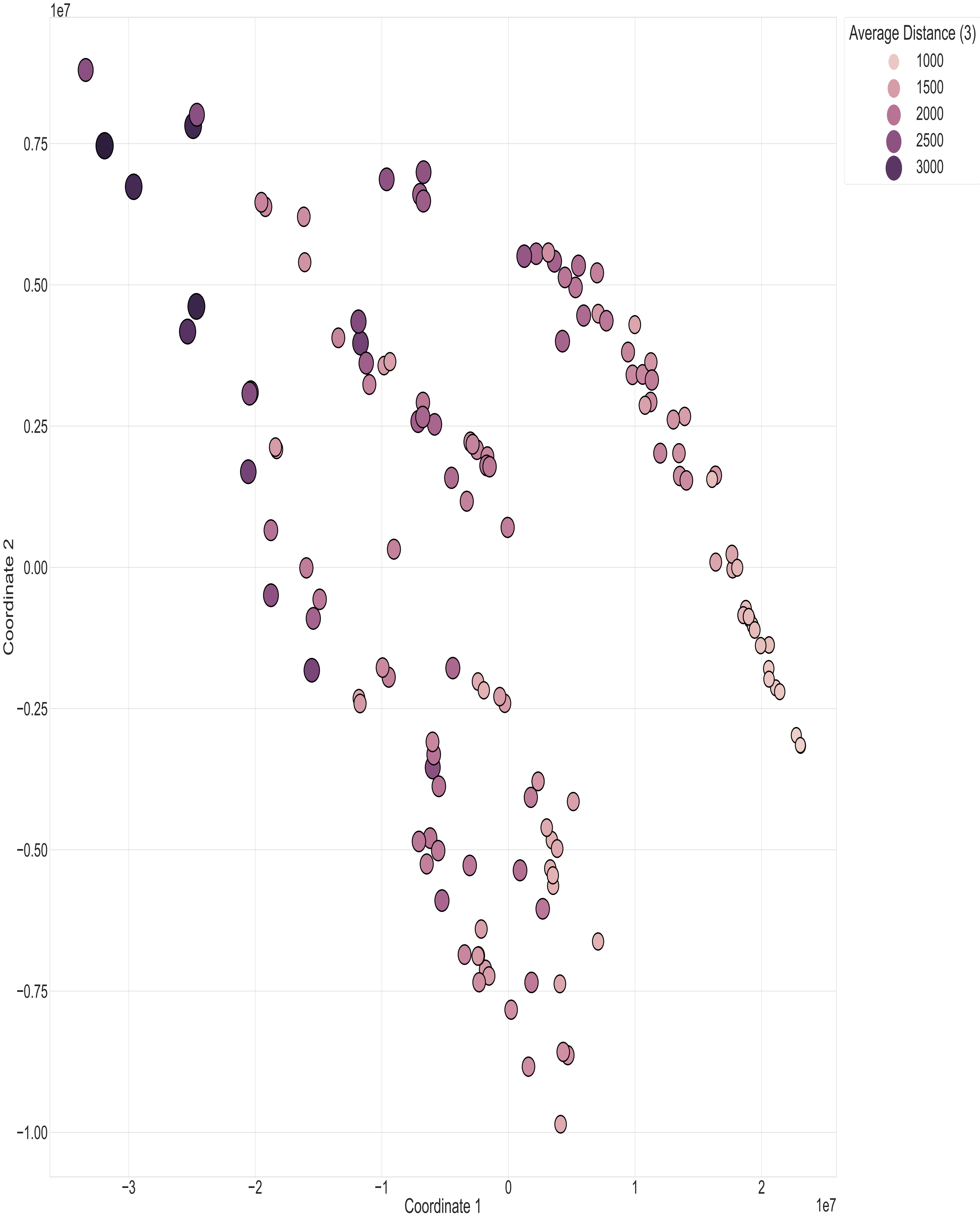
Scatter plot of Hybrid-(Not-PKS-NRPS) *Micromonosporaceae* BGCs analyzed via multidimensional scaling using euclidean pairwise distance.
Only Hybrid-(Not-PKS-NRPS) BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual Hybrid-(Not-PKS-NRPS) BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S174 :



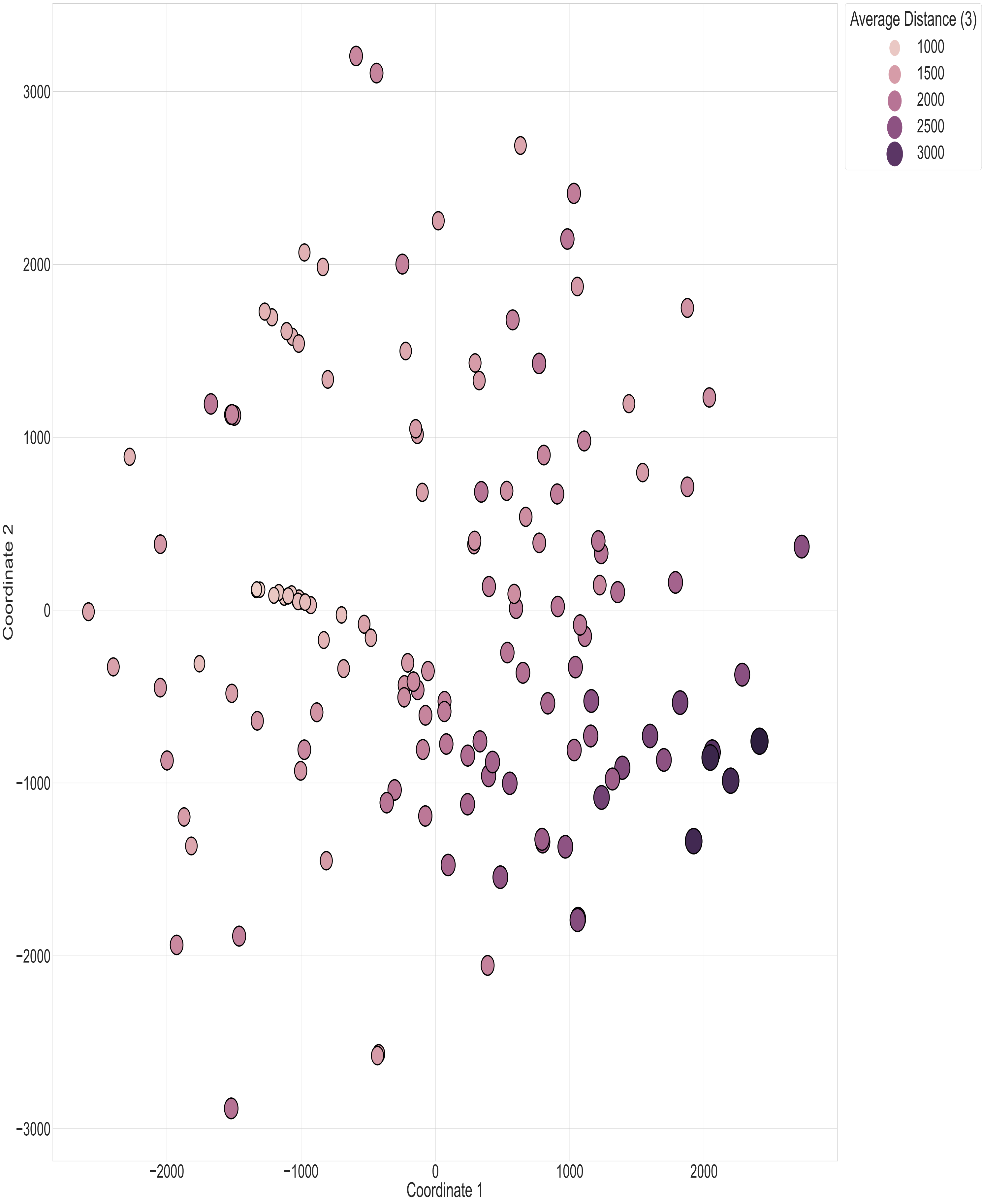
Scatter plot of Hybrid-(Not-PKS-NRPS) *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cosine pairwise distance.
Only Hybrid-(Not-PKS-NRPS) BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual Hybrid-(Not-PKS-NRPS) BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S175 :



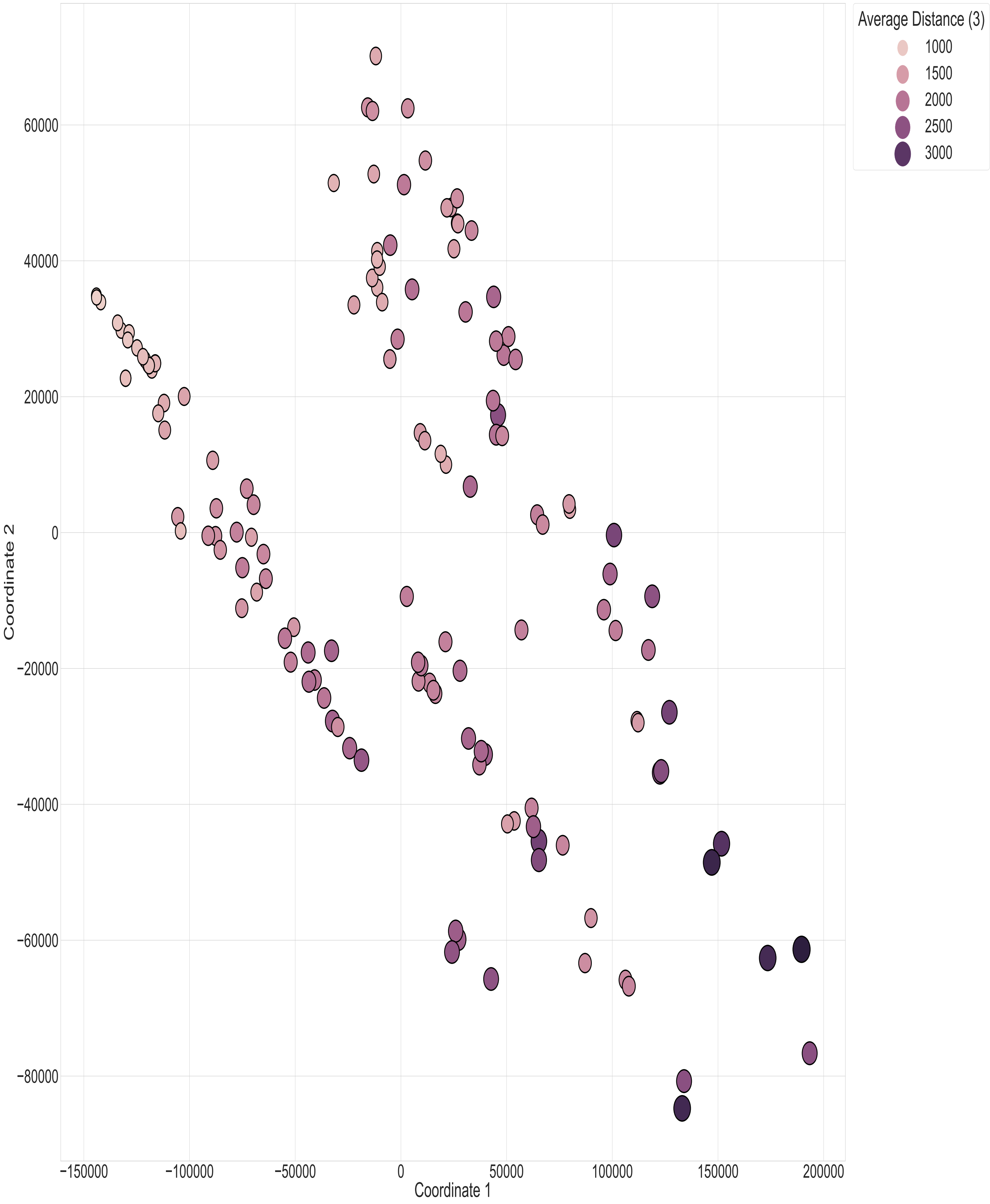
Scatter plot of Hybrid-(Not-PKS-NRPS) *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cityblock pairwise distance.
Only Hybrid-(Not-PKS-NRPS) BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual Hybrid-(Not-PKS-NRPS) BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S176 :



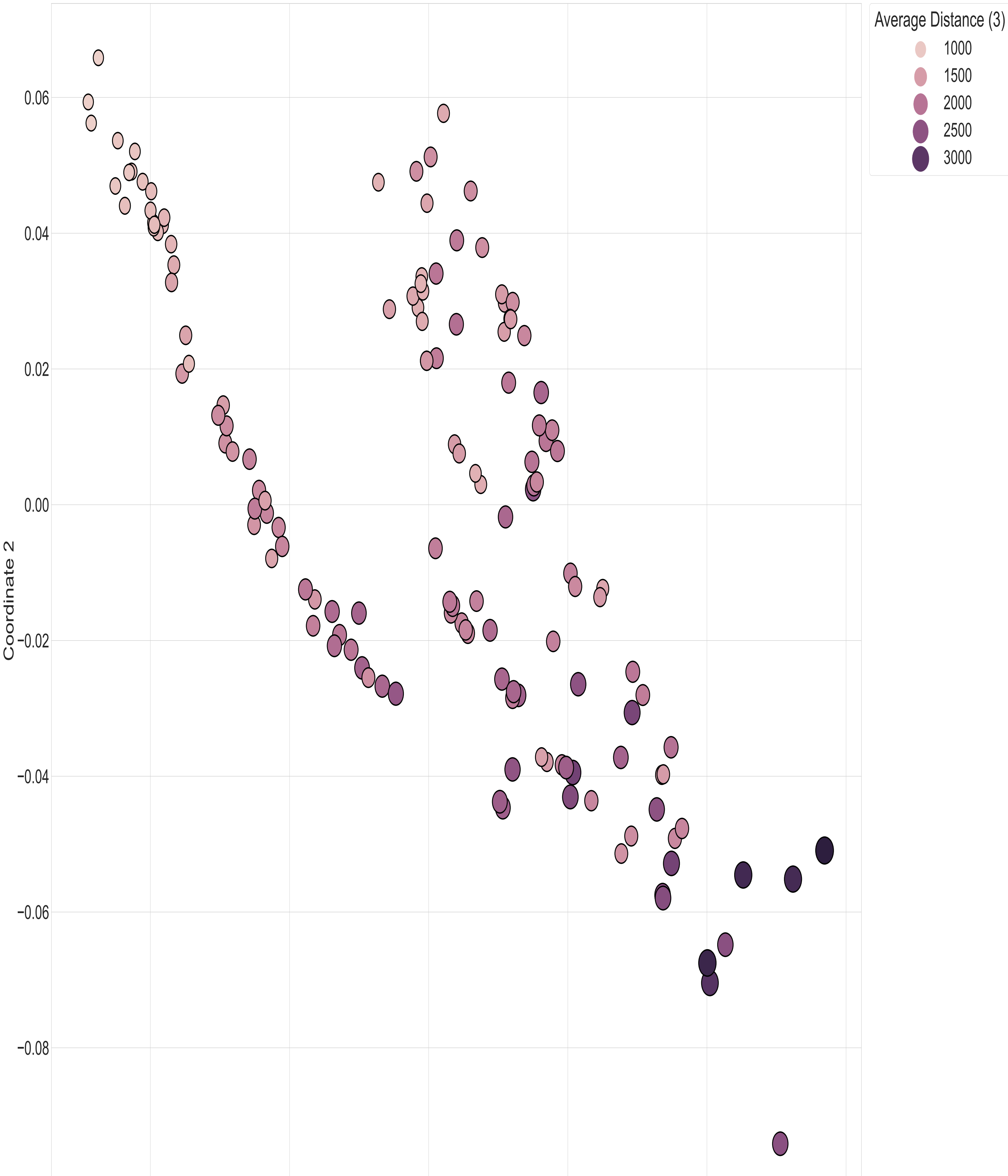
Scatter plot of Hybrid-(Not-PKS-NRPS) *Micromonosporaceae* BGCs analyzed via multidimensional scaling using chebyshev pairwise distance.
Only Hybrid-(Not-PKS-NRPS) BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual Hybrid-(Not-PKS-NRPS) BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S177 :



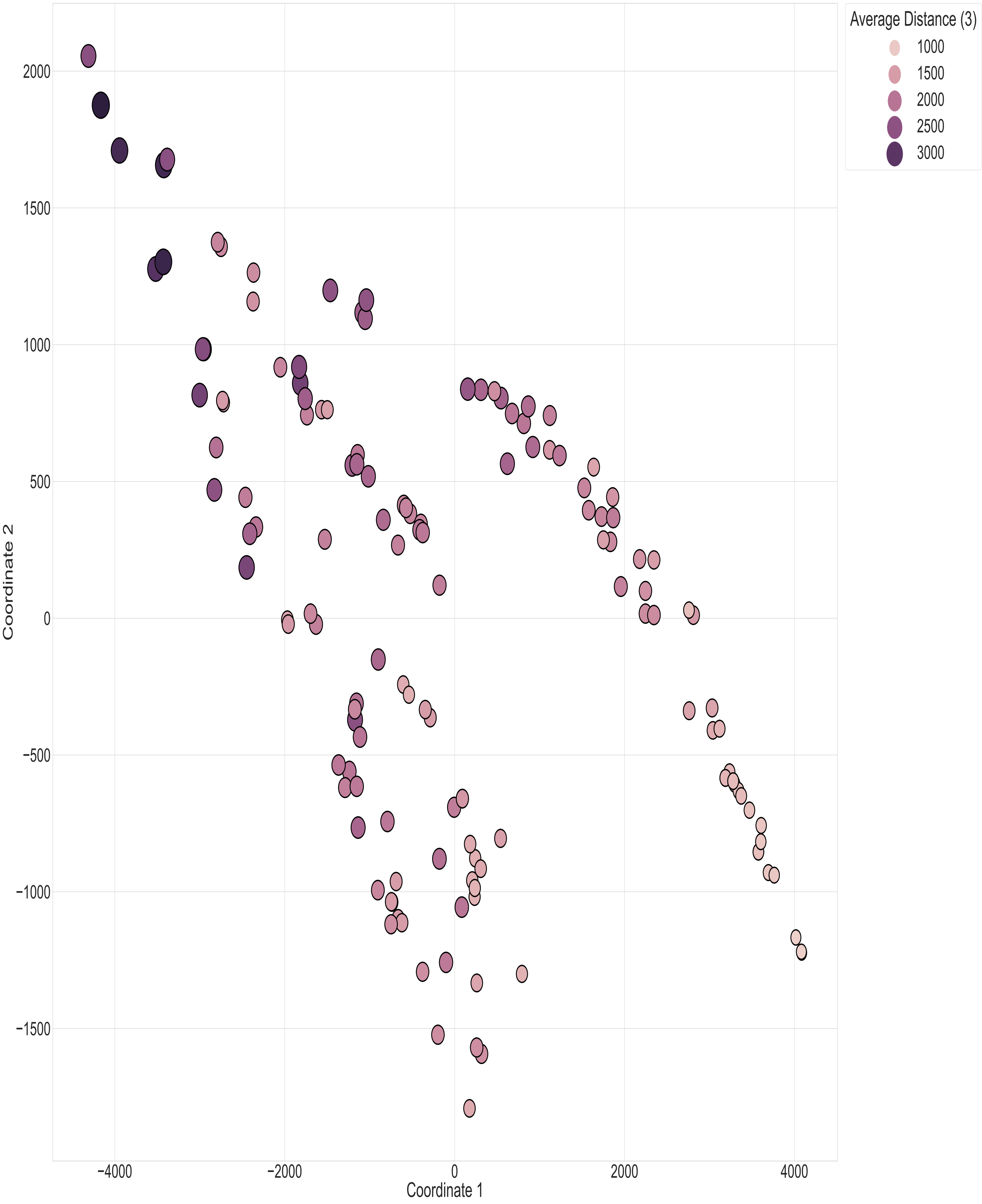
Scatter plot of Hybrid-(Not-PKS-NRPS) *Micromonosporaceae* BGCs analyzed via multidimensional scaling using l2 pairwise distance.
Only Hybrid-(Not-PKS-NRPS) BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual Hybrid-(Not-PKS-NRPS) BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S178 :



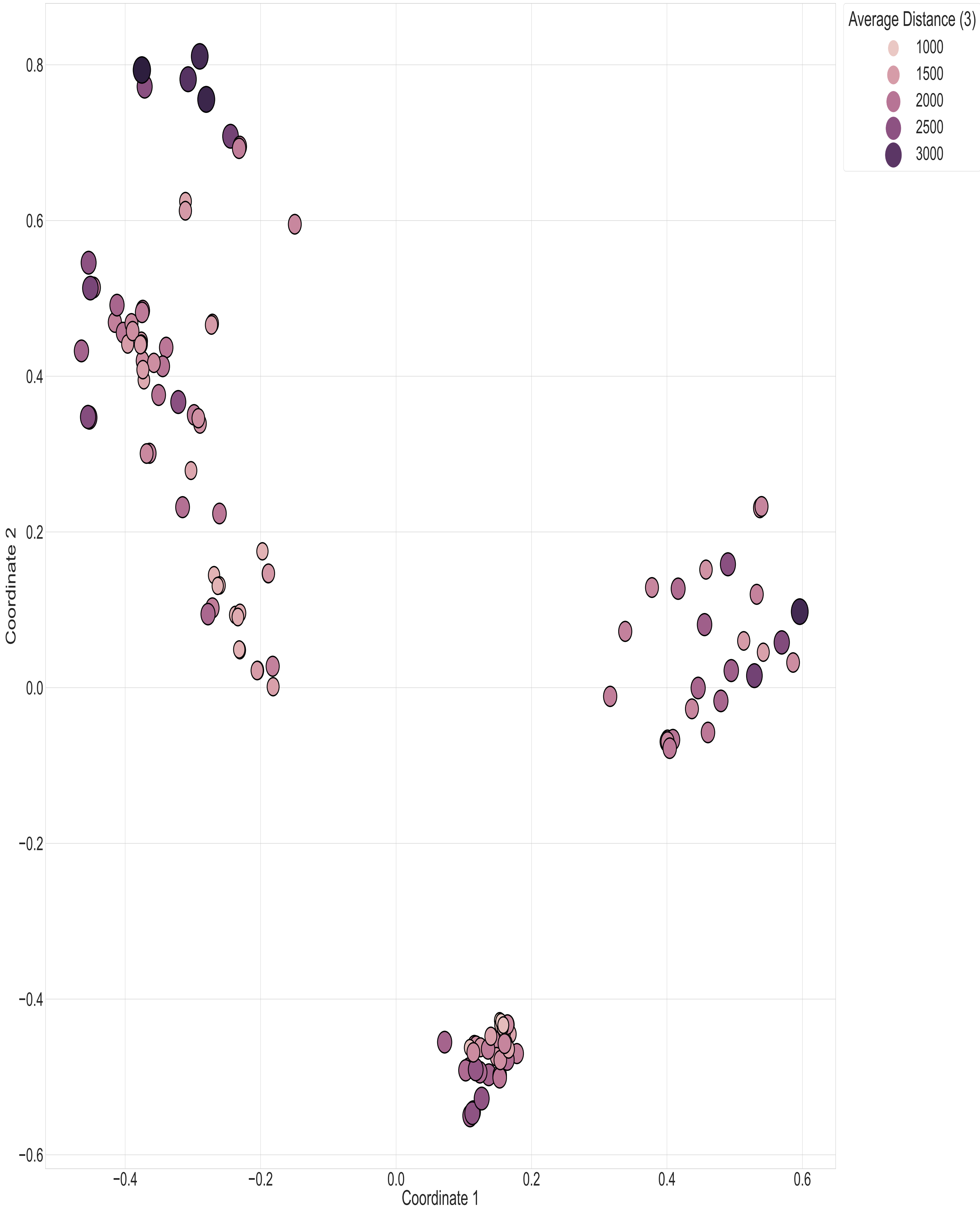
Scatter plot of Hybrid-(Not-PKS-NRPS) *Micromonosporaceae* BGCs analyzed via multidimensional scaling using braycurtis pairwise distance.
Only Hybrid-(Not-PKS-NRPS) BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual Hybrid-(Not-PKS-NRPS) BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S179 :



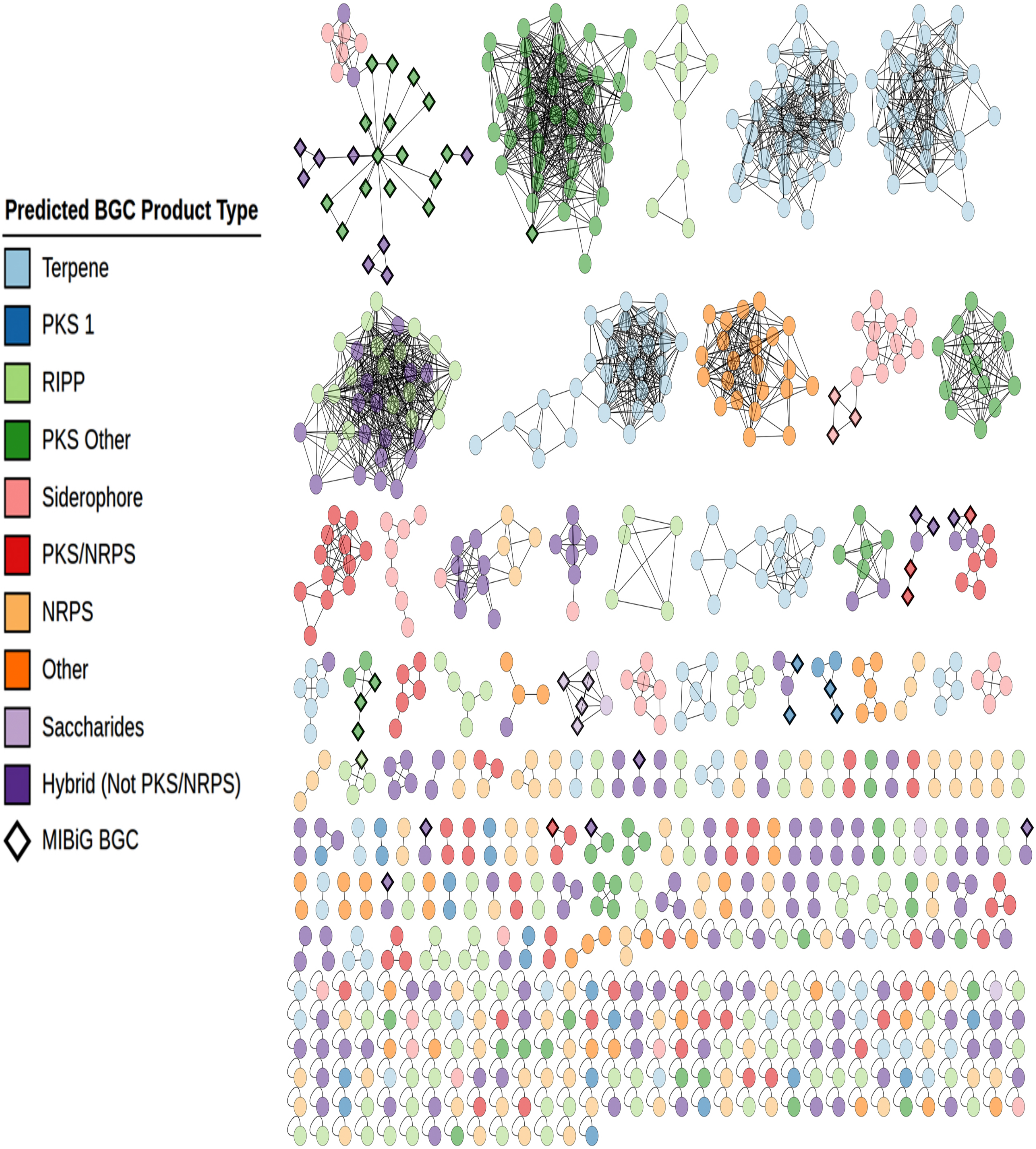
Scatter plot of Hybrid-(Not-PKS-NRPS) *Micromonosporaceae* BGCs analyzed via multidimensional scaling using canberra pairwise distance.
Only Hybrid-(Not-PKS-NRPS) BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual Hybrid-(Not-PKS-NRPS) BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S180 :



Scatter plot of Hybrid-(Not-PKS-NRPS) *Micromonosporaceae* BGCs analyzed via multidimensional scaling using correlation pairwise distance.
Only Hybrid-(Not-PKS-NRPS) BGCs were considered for this analysis. Each dot represents an individual BGC.
The hue and size of the dots were scaled based on the average distance of the individual Hybrid-(Not-PKS-NRPS) BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S181:



Sequence similarity network (SSN) produced by BiG-SCAPE when analyzing Micromonosporaceae strains, visualized and annotated with Cytoscape.Nodes represent individual BGCs. BGC types are colored according to the color legend. Micromonosporaceae BGCs are represented as circles, and MIBiG BGCs are represented as diamonds with a black border. The 196 singletons not visualized in Figure 2 can be seen here.