

Supporting Information: *Micromonosporaceae* Biosynthetic Gene Cluster Diversity Highlights the Need for Broad Spectrum Investigations

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

Imraan Alas¹, Doug R. Braun¹, Spencer S. Ericksen², Rauf Salamzade^{3,4}, Lindsay Kalan^{3,4}, Scott R. Rajski¹ and Tim S. Bugni^{1,2,5*}

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

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

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

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

⁵Lachman 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

<u>Contents</u>	<u>Page</u>
1) Table S1. Statistical comparison of <i>Micromonospora</i> vs <i>Micromonospora_E</i>	10
2) Figure S1. MDS scatter plot of PKS-NRPS BGCs using Chebyshev pairwise distance	11
3) Figure S2. MDS scatter plot of Siderophore BGCs using Chebyshev pairwise distance	12
4) Figure S3. MDS scatter plot of Terpene BGCs using Chebyshev pairwise distance	13
5) Figure S4. MDS scatter plot of RIPP BGCs using Chebyshev pairwise distance	14
6) Figure S5. MDS scatter plot of PKS-1 BGCs using Chebyshev pairwise distance	15
7) Figure S6. MDS scatter plot of NRPS BGCs using Chebyshev pairwise distance	16
8) Figure S7. MDS scatter plot of Hybrid-(Not-PKS-NRPS) BGCs using Chebyshev pairwise distance	17
9) Figure S8. MDS scatter plot of Other BGCs using Chebyshev pairwise distance	18
10) Figure S9. MDS scatter plot of PKS-Other BGCs using Chebyshev pairwise distance	19
11) Figure S10. MDS scatter plot of Saccharides BGCs using Chebyshev pairwise distance	20
12) Figure S11. Annotated MDS scatter plot of PKS-NRPS BGCs using Chebyshev pairwise distance	21
13) Figure S12. Annotated MDS scatter plot of Siderophore BGCs using Chebyshev pairwise distance	22
14) Figure S13. Annotated MDS scatter plot of Terpene BGCs using Chebyshev pairwise distance	23
15) Figure S14. Annotated MDS scatter plot of RIPP BGCs using Chebyshev pairwise distance	24
16) Figure S15. Annotated MDS scatter plot of PKS-1 BGCs using Chebyshev pairwise distance	25
17) Figure S16. Annotated MDS scatter plot of NRPS BGCs using Chebyshev pairwise distance	26
18) Figure S17. Annotated MDS scatter plot of Hybrid-(Not-PKS-NRPS) BGCs using Chebyshev pairwise distance	27
19) Figure S18. Annotated MDS scatter plot of Other BGCs using Chebyshev pairwise distance	28
20) Figure S19. Annotated MDS scatter plot of PKS-Other BGCs using Chebyshev pairwise distance	29
21) Figure S20. Annotated MDS scatter plot of Saccharides BGCs using Chebyshev pairwise distance	30
22) Figure S21. MDS scatter plot of Terpene BGCs using Euclidean pairwise distance	31

23) Figure S22. MDS scatter plot of Terpene BGCs using Cosine pairwise distance	32
24) Figure S23. MDS scatter plot of Terpene BGCs using Cityblock pairwise distance.....	33
25) Figure S24. MDS scatter plot of Terpene BGCs using Chebyshev pairwise distance	34
26) Figure S25. MDS scatter plot of Terpene BGCs using L2 pairwise distance	35
27) Figure S26. MDS scatter plot of Terpene BGCs using Braycurtis pairwise distance	36
28) Figure S27. MDS scatter plot of Terpene BGCs using Canberra pairwise distance	37
29) Figure S28. MDS scatter plot of Terpene BGCs using Correlation pairwise distance.....	38
30) Figure S29. MDS scatter plot of PKS-1 BGCs using Euclidean pairwise distance	39
31) Figure S30. MDS scatter plot of PKS-1 BGCs using Cosine pairwise distance	40
32) Figure S31. MDS scatter plot of PKS-1 BGCs using Cityblock pairwise distance.....	41
33) Figure S32. MDS scatter plot of PKS-1 BGCs using Chebyshev pairwise distance	42
34) Figure S33. MDS scatter plot of PKS-1 BGCs using L2 pairwise distance	43
35) Figure S34. MDS scatter plot of PKS-1 BGCs using Braycurtis pairwise distance	44
36) Figure S35. MDS scatter plot of PKS-1 BGCs using Canberra pairwise distance	45
37) Figure S36. MDS scatter plot of PKS-1 BGCs using Correlation pairwise distance.....	46
38) Figure S37. MDS scatter plot of RIPP BGCs using Euclidean pairwise distance	47
39) Figure S38. MDS scatter plot of RIPP BGCs using Cosine pairwise distance	48
40) Figure S39. MDS scatter plot of RIPP BGCs using Cityblock pairwise distance.....	49
41) Figure S40. MDS scatter plot of RIPP BGCs using Chebyshev pairwise distance	50
42) Figure S41. MDS scatter plot of RIPP BGCs using L2 pairwise distance	51
43) Figure S42. MDS scatter plot of RIPP BGCs using Braycurtis pairwise distance	52
44) Figure S43. MDS scatter plot of RIPP BGCs using Canberra pairwise distance	53
45) Figure S44. MDS scatter plot of RIPP BGCs using Correlation pairwise distance.....	54
46) Figure S45. MDS scatter plot of PKS-Other BGCs	

using Euclidean pairwise distance	55
47) Figure S46. MDS scatter plot of PKS-Other BGCs using Cosine pairwise distance	56
48) Figure S47. MDS scatter plot of PKS-Other BGCs using Cityblock pairwise distance.....	57
49) Figure S48. MDS scatter plot of PKS-Other BGCs using Chebyshev pairwise distance	58
50) Figure S49. MDS scatter plot of PKS-Other BGCs using L2 pairwise distance	59
51) Figure S50. MDS scatter plot of PKS-Other BGCs using Braycurtis pairwise distance	60
52) Figure S51. MDS scatter plot of PKS-Other BGCs using Canberra pairwise distance	61
53) Figure S52. MDS scatter plot of PKS-Other BGCs using Correlation pairwise distance.....	62
54) Figure S53. MDS scatter plot of Siderophore BGCs using Euclidean pairwise distance	63
55) Figure S54. MDS scatter plot of Siderophore BGCs using Cosine pairwise distance	64
56) Figure S55. MDS scatter plot of Siderophore BGCs using Cityblock pairwise distance.....	65
57) Figure S56. MDS scatter plot of Siderophore BGCs using Chebyshev pairwise distance	66
58) Figure S57. MDS scatter plot of Siderophore BGCs using L2 pairwise distance	67
59) Figure S58. MDS scatter plot of Siderophore BGCs using Braycurtis pairwise distance	68
60) Figure S59. MDS scatter plot of Siderophore BGCs using Canberra pairwise distance	69
61) Figure S60. MDS scatter plot of Siderophore BGCs using Correlation pairwise distance.....	70
62) Figure S61. MDS scatter plot of PKS-NRPS BGCs using Euclidean pairwise distance	71
63) Figure S62. MDS scatter plot of PKS-NRPS BGCs using Cosine pairwise distance	72
64) Figure S63. MDS scatter plot of PKS-NRPS BGCs using Cityblock pairwise distance.....	73
65) Figure S64. MDS scatter plot of PKS-NRPS BGCs using Chebyshev pairwise distance	74
66) Figure S65. MDS scatter plot of PKS-NRPS BGCs using L2 pairwise distance	75
67) Figure S66. MDS scatter plot of PKS-NRPS BGCs using Braycurtis pairwise distance	76
68) Figure S67. MDS scatter plot of PKS-NRPS BGCs using Canberra pairwise distance	77
69) Figure S68. MDS scatter plot of PKS-NRPS BGCs using Correlation pairwise distance.....	78

70) Figure S69. MDS scatter plot of NRPS BGCs using Euclidean pairwise distance	79
71) Figure S70. MDS scatter plot of NRPS BGCs using Cosine pairwise distance	80
72) Figure S71. MDS scatter plot of NRPS BGCs using Cityblock pairwise distance.....	81
73) Figure S72. MDS scatter plot of NRPS BGCs using Chebyshev pairwise distance	82
74) Figure S73. MDS scatter plot of NRPS BGCs using L2 pairwise distance	83
75) Figure S74. MDS scatter plot of NRPS BGCs using Braycurtis pairwise distance	84
76) Figure S75. MDS scatter plot of NRPS BGCs using Canberra pairwise distance	85
77) Figure S76. MDS scatter plot of NRPS BGCs using Correlation pairwise distance.....	86
78) Figure S77. MDS scatter plot of Other BGCs using Euclidean pairwise distance	87
79) Figure S78. MDS scatter plot of Other BGCs using Cosine pairwise distance	88
80) Figure S79. MDS scatter plot of Other BGCs using Cityblock pairwise distance.....	89
81) Figure S80. MDS scatter plot of Other BGCs using Chebyshev pairwise distance	90
82) Figure S81. MDS scatter plot of Other BGCs using L2 pairwise distance	91
83) Figure S82. MDS scatter plot of Other BGCs using Braycurtis pairwise distance	92
84) Figure S83. MDS scatter plot of Other BGCs using Canberra pairwise distance	93
85) Figure S84. MDS scatter plot of Other BGCs using Correlation pairwise distance.....	94
86) Figure S85. MDS scatter plot of Saccharides BGCs using Euclidean pairwise distance	95
87) Figure S86. MDS scatter plot of Saccharides BGCs using Cosine pairwise distance	96
88) Figure S87. MDS scatter plot of Saccharides BGCs using Cityblock pairwise distance.....	97
89) Figure S88. MDS scatter plot of Saccharides BGCs using Chebyshev pairwise distance	98
90) Figure S89. MDS scatter plot of Saccharides BGCs using L2 pairwise distance	99
91) Figure S90. MDS scatter plot of Saccharides BGCs using Braycurtis pairwise distance	100
92) Figure S91. MDS scatter plot of Saccharides BGCs using Canberra pairwise distance	101
93) Figure S92. MDS scatter plot of Saccharides BGCs	

using Correlation pairwise distance.....	102
94) Figure S93. MDS scatter plot of Hybrid-(Not-PKS-NRPS) BGCs using Euclidean pairwise distance	103
95) Figure S94. MDS scatter plot of Hybrid-(Not-PKS-NRPS) BGCs using Cosine pairwise distance	104
96) Figure S95. MDS scatter plot of Hybrid-(Not-PKS-NRPS) BGCs using Cityblock pairwise distance.....	105
97) Figure S96. MDS scatter plot of Hybrid-(Not-PKS-NRPS) BGCs using Chebyshev pairwise distance	106
98) Figure S97. MDS scatter plot of Hybrid-(Not-PKS-NRPS) BGCs using L2 pairwise distance	107
99) Figure S98. MDS scatter plot of Hybrid-(Not-PKS-NRPS) BGCs using Braycurtis pairwise distance	108
100) Figure S99. MDS scatter plot of Hybrid-(Not-PKS-NRPS) BGCs using Canberra pairwise distance	109
101) Figure S100. MDS scatter plot of Hybrid-(Not-PKS-NRPS) BGCs using Correlation pairwise distance.....	110
102) Figure S101. Annotated MDS scatter plot of Terpene BGCs using Euclidean pairwise distance	111
103) Figure S102. Annotated MDS scatter plot of Terpene BGCs using Cosine pairwise distance	112
104) Figure S103. Annotated MDS scatter plot of Terpene BGCs using Cityblock pairwise distance.....	113
105) Figure S104. Annotated MDS scatter plot of Terpene BGCs using Chebyshev pairwise distance	114
106) Figure S105. Annotated MDS scatter plot of Terpene BGCs using L2 pairwise distance	115
107) Figure S106. Annotated MDS scatter plot of Terpene BGCs using Braycurtis pairwise distance	116
108) Figure S107. Annotated MDS scatter plot of Terpene BGCs using Canberra pairwise distance	117
109) Figure S108. Annotated MDS scatter plot of Terpene BGCs using Correlation pairwise distance.....	118
110) Figure S109. Annotated MDS scatter plot of PKS-1 BGCs using Euclidean pairwise distance	119
111) Figure S110. Annotated MDS scatter plot of PKS-1 BGCs using Cosine pairwise distance	120
112) Figure S111. Annotated MDS scatter plot of PKS-1 BGCs using Cityblock pairwise distance.....	121
113) Figure S112. Annotated MDS scatter plot of PKS-1 BGCs using Chebyshev pairwise distance	122
114) Figure S113. Annotated MDS scatter plot of PKS-1 BGCs using L2 pairwise distance	123
115) Figure S114. Annotated MDS scatter plot of PKS-1 BGCs using Braycurtis pairwise distance	124
116) Figure S115. Annotated MDS scatter plot of PKS-1 BGCs using Canberra pairwise distance	125

117) Figure S116. Annotated MDS scatter plot of PKS-1 BGCs using Correlation pairwise distance.....	126
118) Figure S117. Annotated MDS scatter plot of RIPP BGCs using Euclidean pairwise distance	127
119) Figure S118. Annotated MDS scatter plot of RIPP BGCs using Cosine pairwise distance	128
120) Figure S119. Annotated MDS scatter plot of RIPP BGCs using Cityblock pairwise distance.....	129
121) Figure S120. Annotated MDS scatter plot of RIPP BGCs using Chebyshev pairwise distance	130
122) Figure S121. Annotated MDS scatter plot of RIPP BGCs using L2 pairwise distance	131
123) Figure S122. Annotated MDS scatter plot of RIPP BGCs using Braycurtis pairwise distance	132
124) Figure S123. Annotated MDS scatter plot of RIPP BGCs using Canberra pairwise distance	133
125) Figure S124. Annotated MDS scatter plot of RIPP BGCs using Correlation pairwise distance.....	134
126) Figure S125. Annotated MDS scatter plot of PKS-Other BGCs using Euclidean pairwise distance	135
127) Figure S126. Annotated MDS scatter plot of PKS-Other BGCs using Cosine pairwise distance	136
128) Figure S127. Annotated MDS scatter plot of PKS-Other BGCs using Cityblock pairwise distance.....	137
129) Figure S128. Annotated MDS scatter plot of PKS-Other BGCs using Chebyshev pairwise distance	138
130) Figure S129. Annotated MDS scatter plot of PKS-Other BGCs using L2 pairwise distance	139
131) Figure S130. Annotated MDS scatter plot of PKS-Other BGCs using Braycurtis pairwise distance	140
132) Figure S131. Annotated MDS scatter plot of PKS-Other BGCs using Canberra pairwise distance	141
133) Figure S132. Annotated MDS scatter plot of PKS-Other BGCs using Correlation pairwise distance.....	142
134) Figure S133. Annotated MDS scatter plot of Siderophore BGCs using Euclidean pairwise distance	143
135) Figure S134. Annotated MDS scatter plot of Siderophore BGCs using Cosine pairwise distance	144
136) Figure S135. Annotated MDS scatter plot of Siderophore BGCs using Cityblock pairwise distance.....	145
137) Figure S136. Annotated MDS scatter plot of Siderophore BGCs using Chebyshev pairwise distance	146
138) Figure S137. Annotated MDS scatter plot of Siderophore BGCs using L2 pairwise distance	147
139) Figure S138. Annotated MDS scatter plot of Siderophore BGCs using Braycurtis pairwise distance	148
140) Figure S139. Annotated MDS scatter plot of Siderophore BGCs	

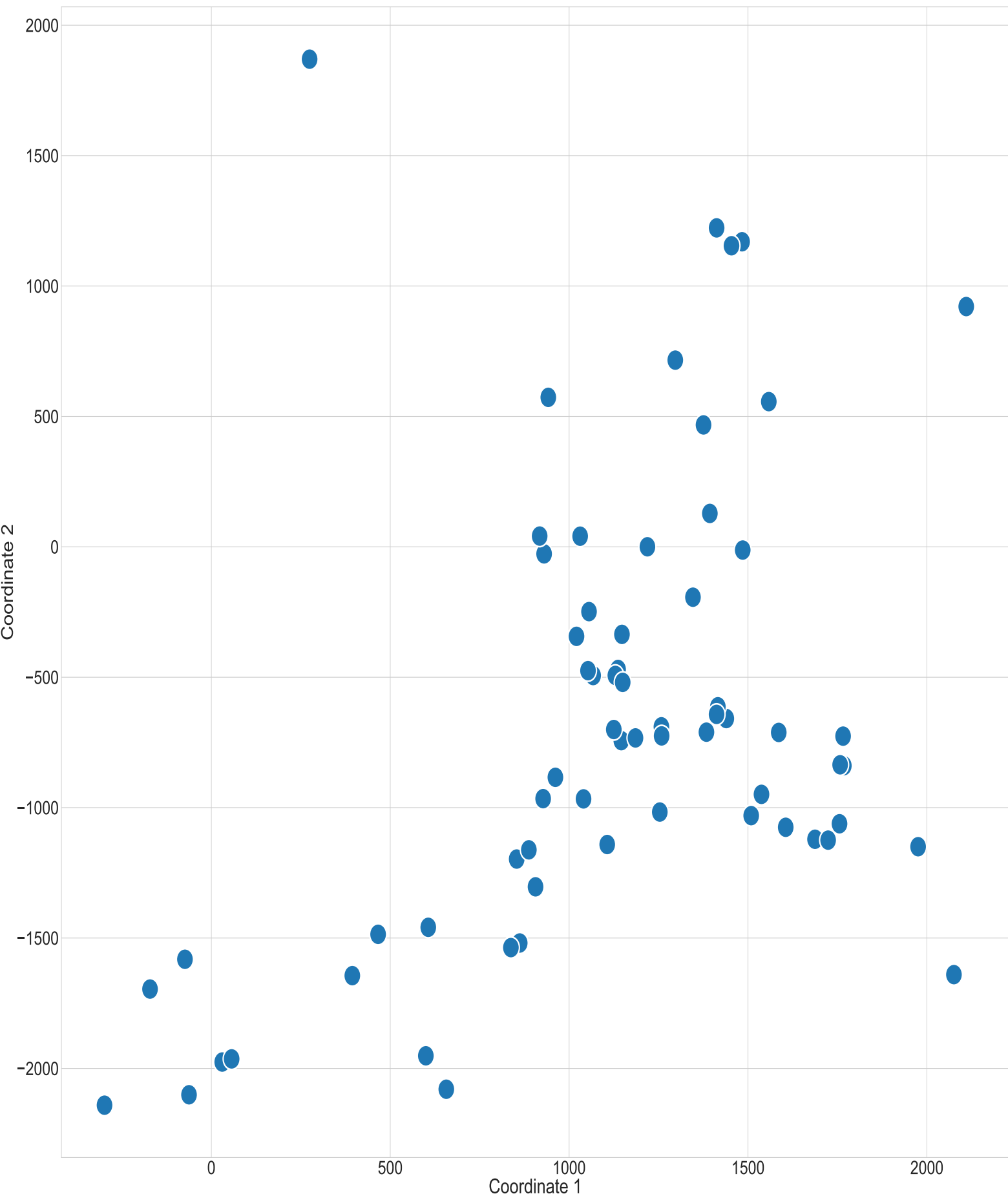
using Canberra pairwise distance	149
141) Figure S140. Annotated MDS scatter plot of Siderophore BGCs using Correlation pairwise distance.....	150
142) Figure S141. Annotated MDS scatter plot of PKS-NRPS BGCs using Euclidean pairwise distance	151
143) Figure S142. Annotated MDS scatter plot of PKS-NRPS BGCs using Cosine pairwise distance	152
144) Figure S143. Annotated MDS scatter plot of PKS-NRPS BGCs using Cityblock pairwise distance.....	153
145) Figure S144. Annotated MDS scatter plot of PKS-NRPS BGCs using Chebyshev pairwise distance	154
146) Figure S145. Annotated MDS scatter plot of PKS-NRPS BGCs using L2 pairwise distance	155
147) Figure S146. Annotated MDS scatter plot of PKS-NRPS BGCs using Braycurtis pairwise distance	156
148) Figure S147. Annotated MDS scatter plot of PKS-NRPS BGCs using Canberra pairwise distance	157
149) Figure S148. Annotated MDS scatter plot of PKS-NRPS BGCs using Correlation pairwise distance.....	158
150) Figure S149. Annotated MDS scatter plot of NRPS BGCs using Euclidean pairwise distance	159
151) Figure S150. Annotated MDS scatter plot of NRPS BGCs using Cosine pairwise distance	160
152) Figure S151. Annotated MDS scatter plot of NRPS BGCs using Cityblock pairwise distance.....	161
153) Figure S152. Annotated MDS scatter plot of NRPS BGCs using Chebyshev pairwise distance	162
154) Figure S153. Annotated MDS scatter plot of NRPS BGCs using L2 pairwise distance	163
155) Figure S154. Annotated MDS scatter plot of NRPS BGCs using Braycurtis pairwise distance	164
156) Figure S155. Annotated MDS scatter plot of NRPS BGCs using Canberra pairwise distance	165
157) Figure S156. Annotated MDS scatter plot of NRPS BGCs using Correlation pairwise distance.....	166
158) Figure S157. Annotated MDS scatter plot of Other BGCs using Euclidean pairwise distance	167
159) Figure S158. Annotated MDS scatter plot of Other BGCs using Cosine pairwise distance	168
160) Figure S159. Annotated MDS scatter plot of Other BGCs using Cityblock pairwise distance.....	169
161) Figure S160. Annotated MDS scatter plot of Other BGCs using Chebyshev pairwise distance	170
162) Figure S161. Annotated MDS scatter plot of Other BGCs using L2 pairwise distance	171
163) Figure S162. Annotated MDS scatter plot of Other BGCs using Braycurtis pairwise distance	172

164) Figure S163. Annotated MDS scatter plot of Other BGCs using Canberra pairwise distance	173
165) Figure S164. Annotated MDS scatter plot of Other BGCs using Correlation pairwise distance.....	174
166) Figure S165. Annotated MDS scatter plot of Saccharides BGCs using Euclidean pairwise distance	175
167) Figure S166. Annotated MDS scatter plot of Saccharides BGCs using Cosine pairwise distance	176
168) Figure S167. Annotated MDS scatter plot of Saccharides BGCs using Cityblock pairwise distance.....	177
169) Figure S168. Annotated MDS scatter plot of Saccharides BGCs using Chebyshev pairwise distance	178
170) Figure S169. Annotated MDS scatter plot of Saccharides BGCs using L2 pairwise distance	179
171) Figure S170. Annotated MDS scatter plot of Saccharides BGCs using Braycurtis pairwise distance	180
172) Figure S171. Annotated MDS scatter plot of Saccharides BGCs using Canberra pairwise distance	181
173) Figure S172. Annotated MDS scatter plot of Saccharides BGCs using Correlation pairwise distance.....	182
174) Figure S173. Annotated MDS scatter plot of Hybrid-(Not-PKS-NRPS) BGCs using Euclidean pairwise distance	183
175) Figure S174. Annotated MDS scatter plot of Hybrid-(Not-PKS-NRPS) BGCs using Cosine pairwise distance	184
176) Figure S175. Annotated MDS scatter plot of Hybrid-(Not-PKS-NRPS) BGCs using Cityblock pairwise distance.....	185
177) Figure S176. Annotated MDS scatter plot of Hybrid-(Not-PKS-NRPS) BGCs using Chebyshev pairwise distance	186
178) Figure S177. Annotated MDS scatter plot of Hybrid-(Not-PKS-NRPS) BGCs using L2 pairwise distance	187
179) Figure S178. Annotated MDS scatter plot of Hybrid-(Not-PKS-NRPS) BGCs using Braycurtis pairwise distance	188
180) Figure S179. Annotated MDS scatter plot of Hybrid-(Not-PKS-NRPS) BGCs using Canberra pairwise distance	189
181) Figure S180. Annotated MDS scatter plot of Hybrid-(Not-PKS-NRPS) BGCs using Correlation pairwise distance.....	190
182) Figure S181. Sequence similarity network (SSN) produced by BiG-SCAPE when analyzing <i>Micromonospora</i> strains as visualized and annotated with CytoSCAPE ...	191

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.

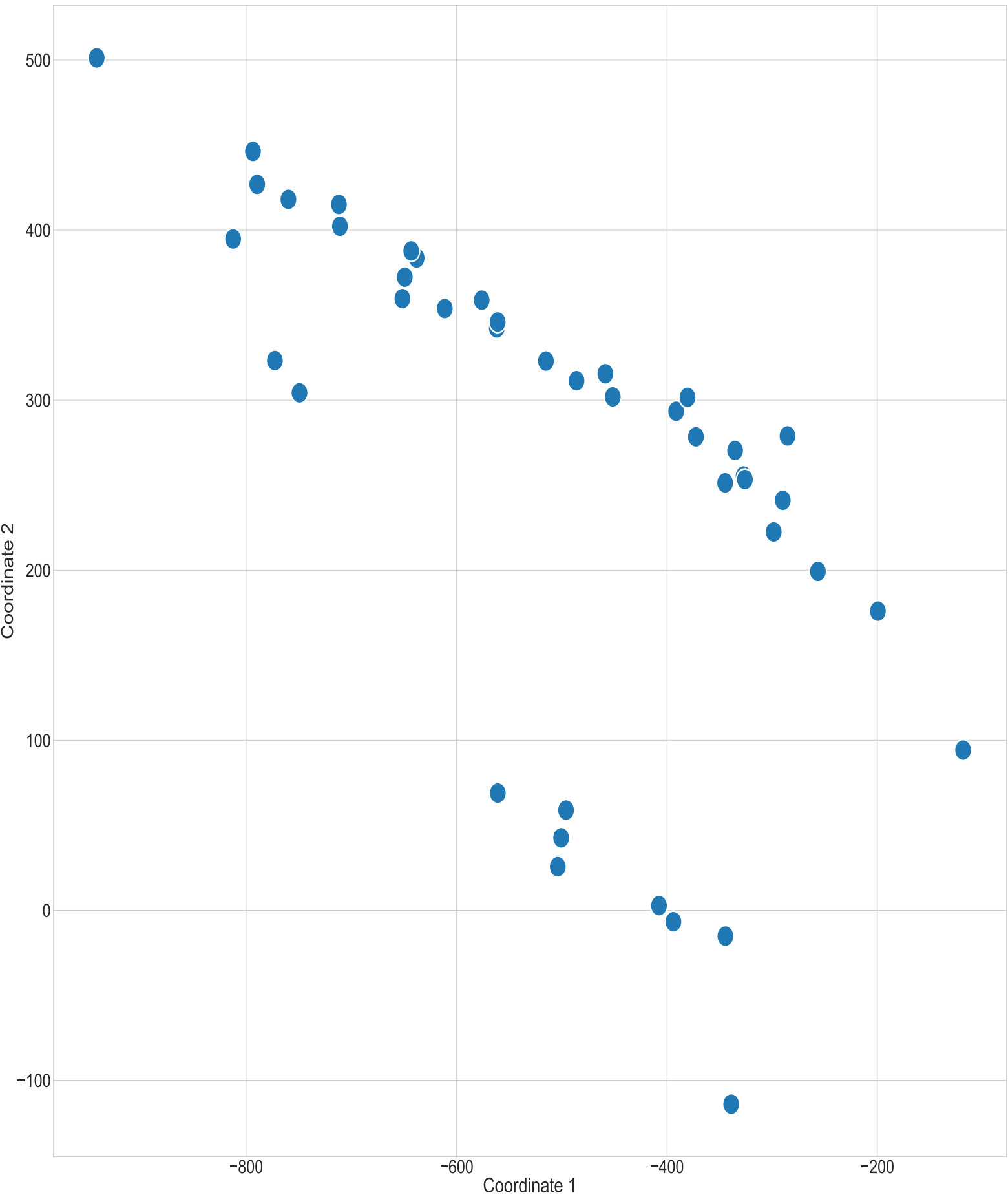
Statistics:	<i>Micromonospora:</i>	<i>Micromonospora_E:</i>
Mean:	1006.01	1174.01
Median:	908	1004
Standard Deviation:	486.34	555.8
Standard Error:	20.37	45.84
P-value:	6.842e-04	

Figure – S1 :



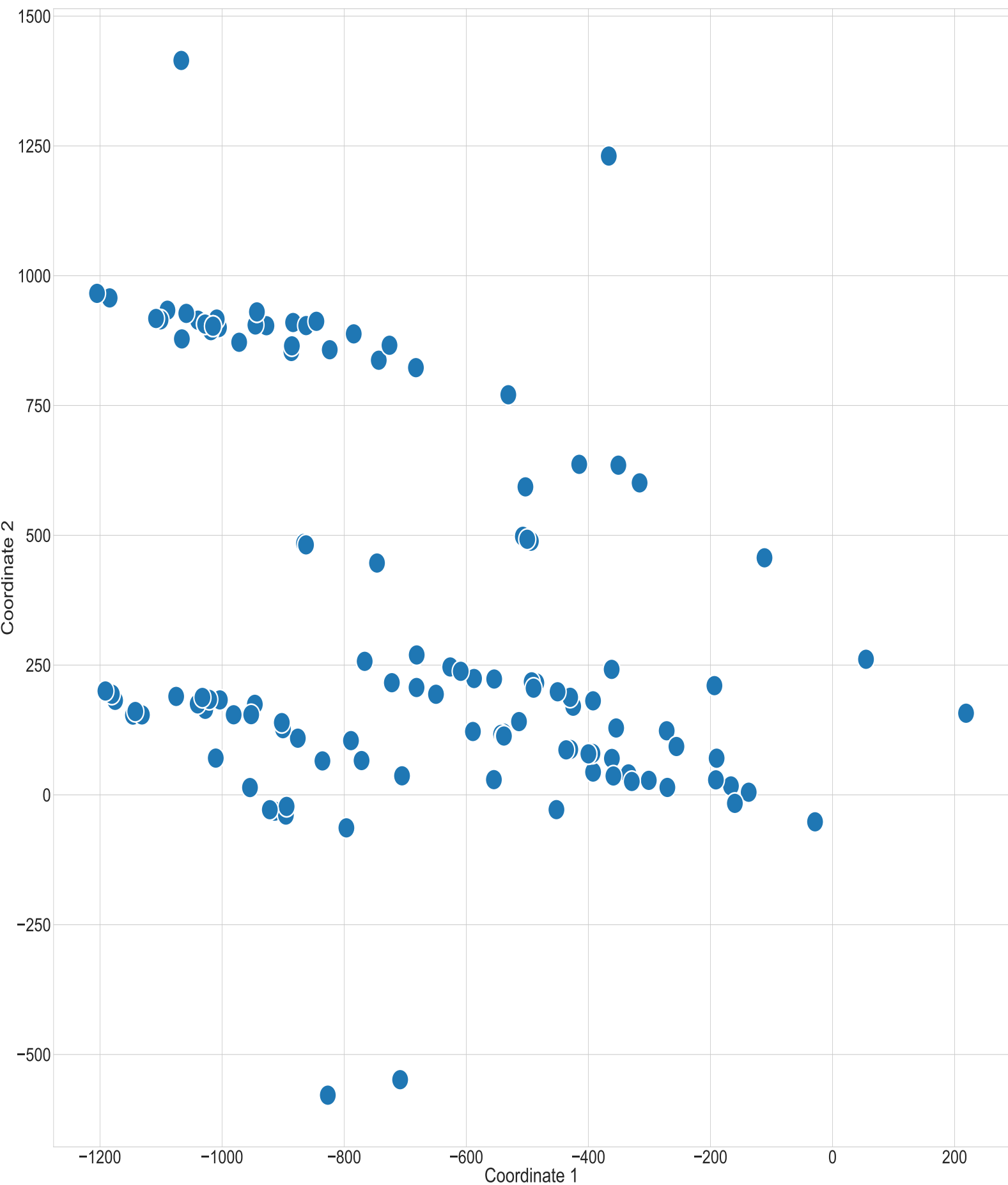
Scatter plot of PKS-NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using Chebyshev pairwise distance. Each dot represents an individual BGC. Distance between BGCs is associated with GCFs in BiG-SLICE they were most similar to.

Figure – S2 :



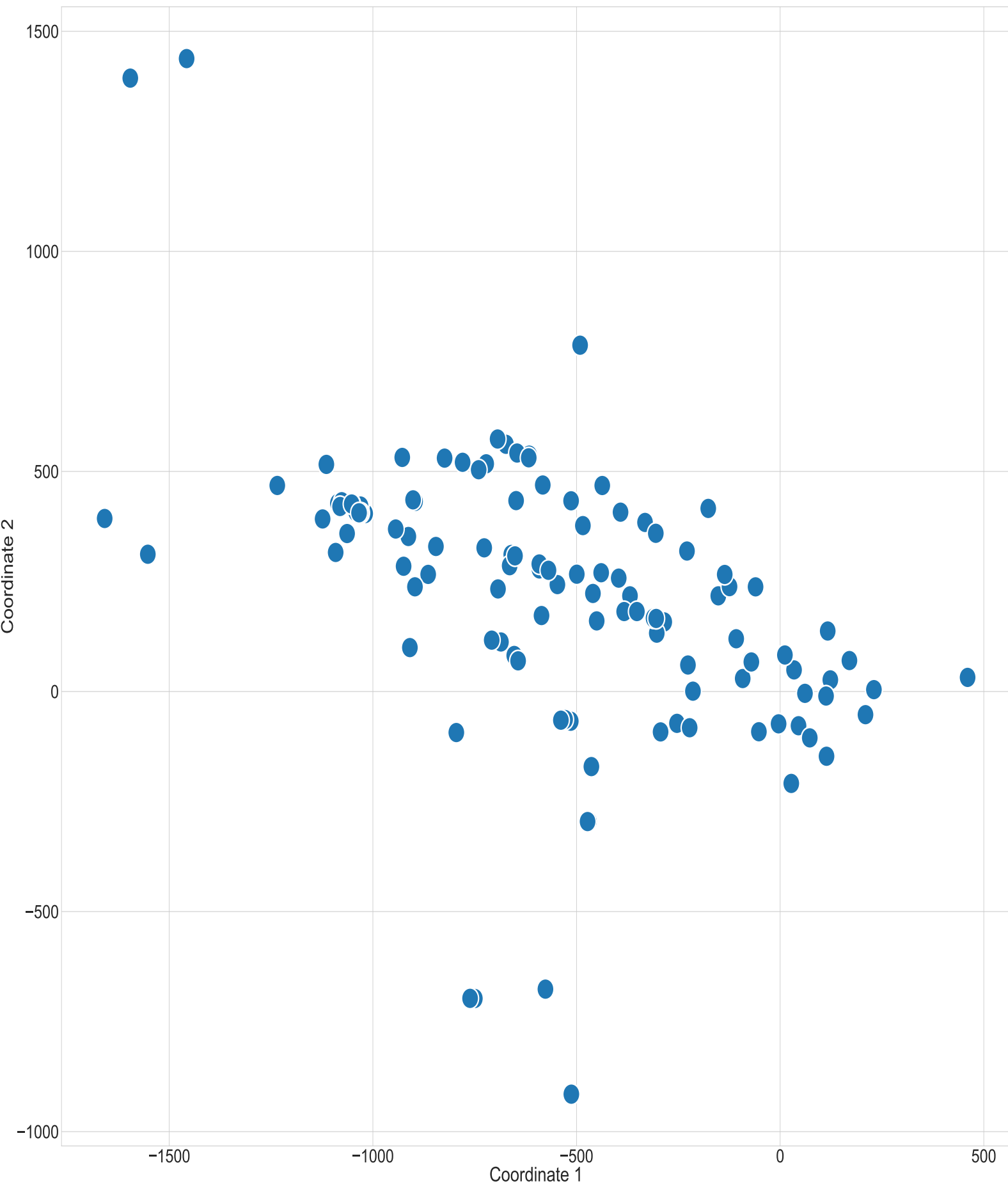
Scatter plot of Siderophore *Micromonosporaceae* BGCs analyzed via multidimensional scaling using Chebyshev pairwise distance. Each dot represents an individual BGC. Distance between BGCs is associated with GCFs in BiG-SLiCE they were most similar to.

Figure – S3 :



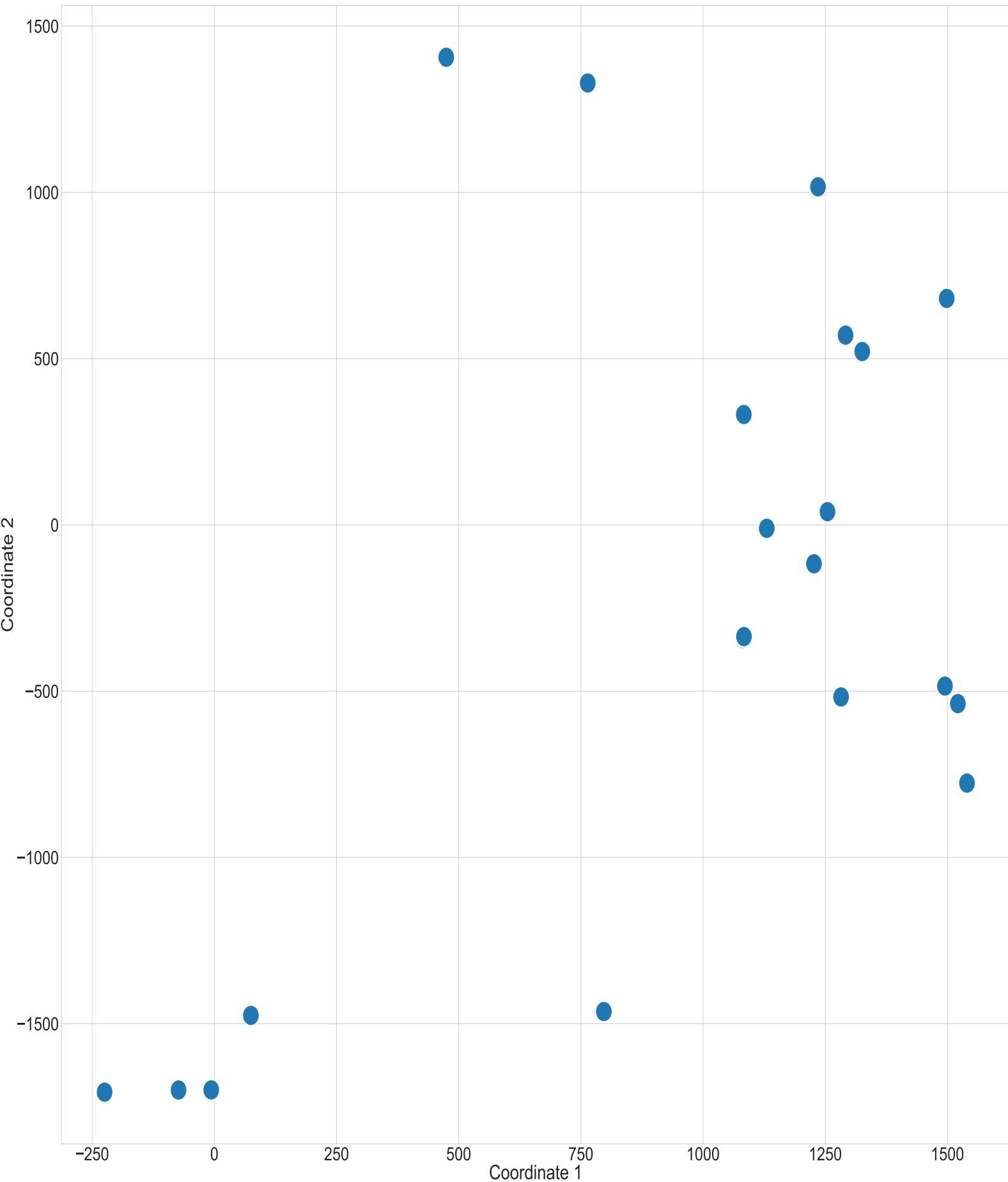
Scatter plot of Terpene *Micromonosporaceae* BGCs analyzed via multidimensional scaling using Chebyshev pairwise distance. Each dot represents an individual BGC. Distance between BGCs is associated with GCFs in BiG-SLiCE they were most similar to.

Figure – S4 :



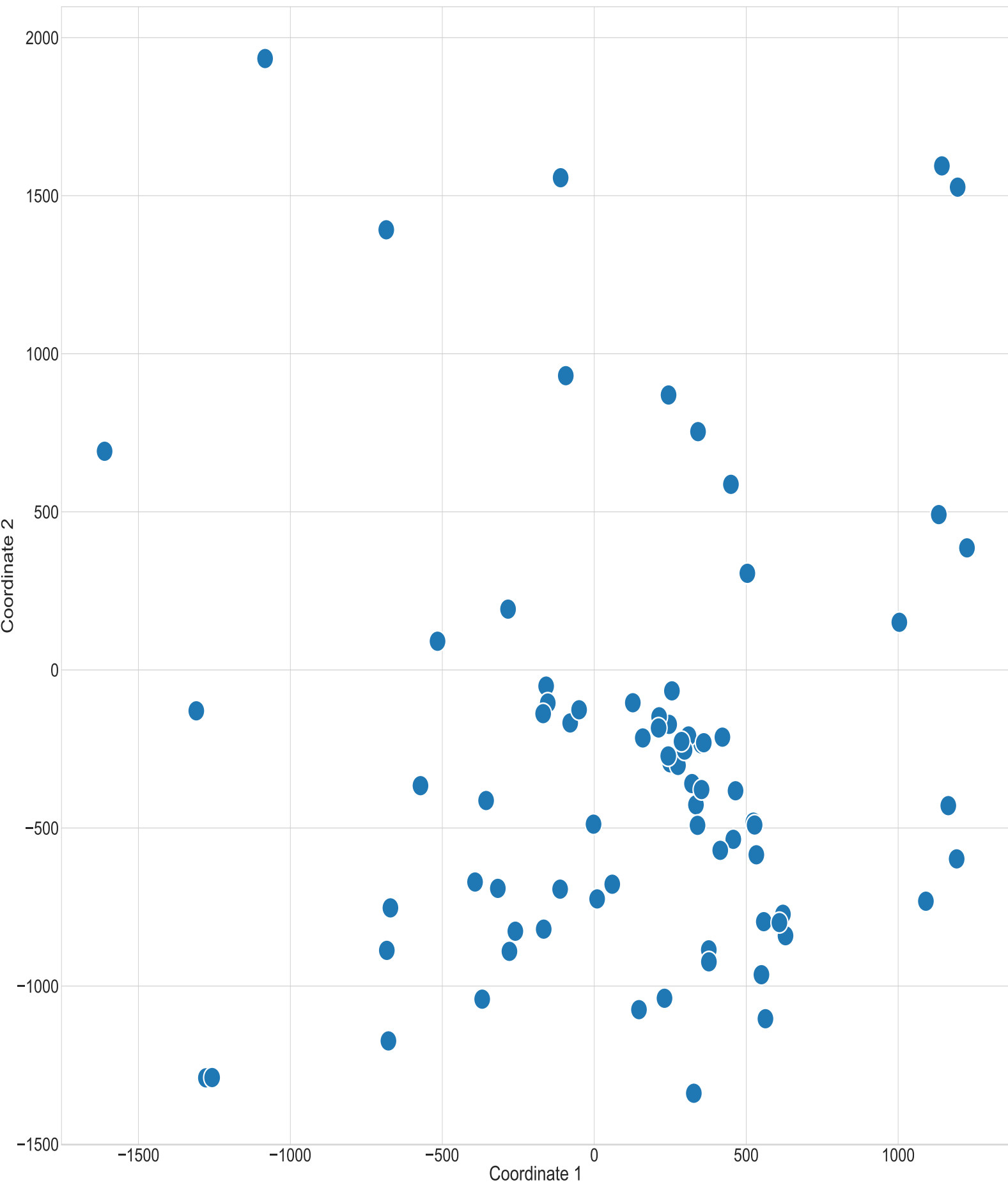
Scatter plot of RIPP *Micromonosporaceae* BGCs analyzed via multidimensional scaling using Chebyshev pairwise distance. Each dot represents an individual BGC. Distance between BGCs is associated with GCFs in BiG-SLiCE they were most similar to.

Figure – S5 :



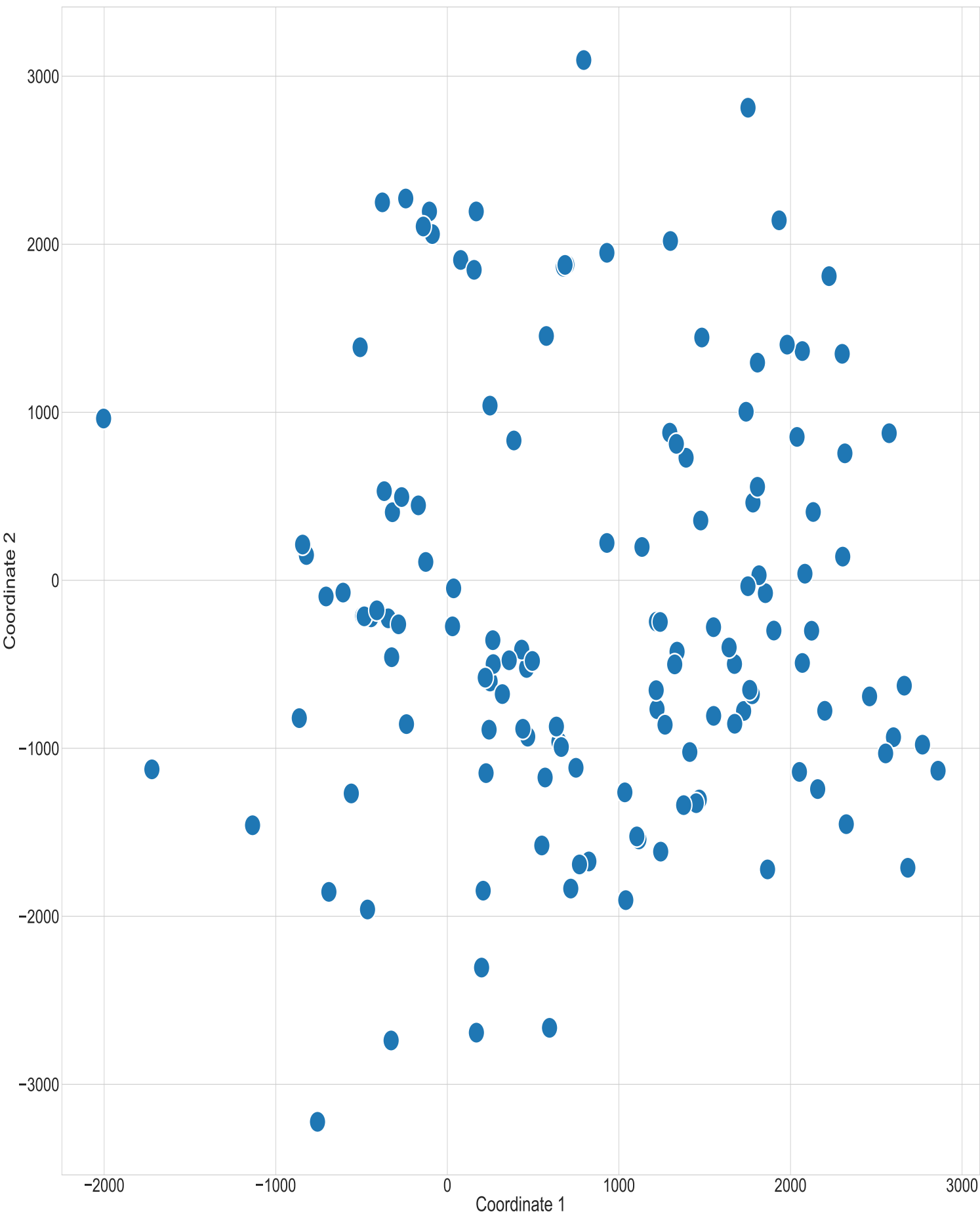
Scatter plot of PKS-1 *Micromonosporaceae* BGCs analyzed via multidimensional scaling using Chebyshev pairwise distance. Each dot represents an individual BGC. Distance between BGCs is associated with GCFs in BiG-SLiCE they were most similar to.

Figure – S6 :



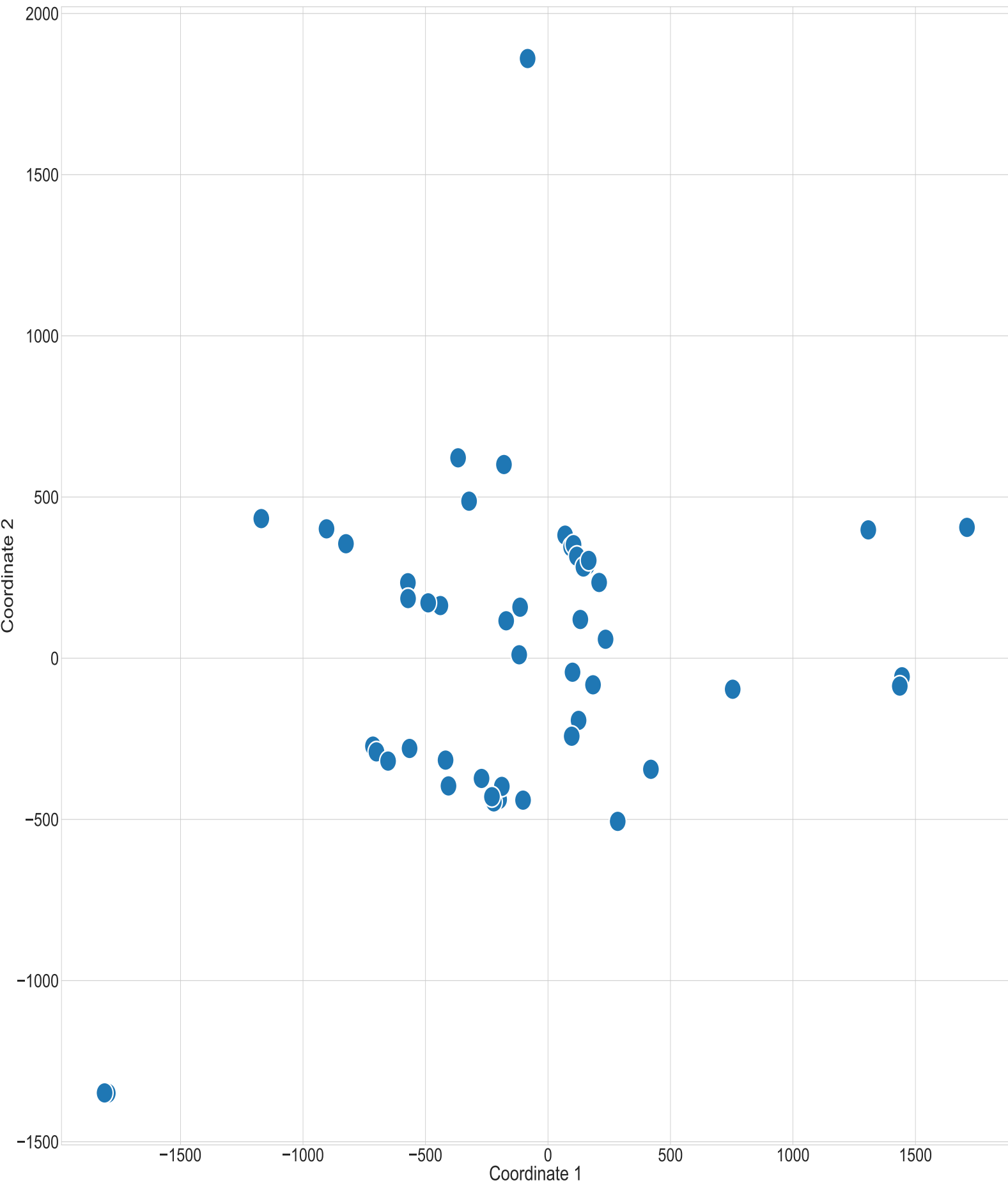
Scatter plot of NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using Chebyshev pairwise distance. Each dot represents an individual BGC. Distance between BGCs is associated with GCFs in BiG-SLiCE they were most similar to.

Figure – S7 :



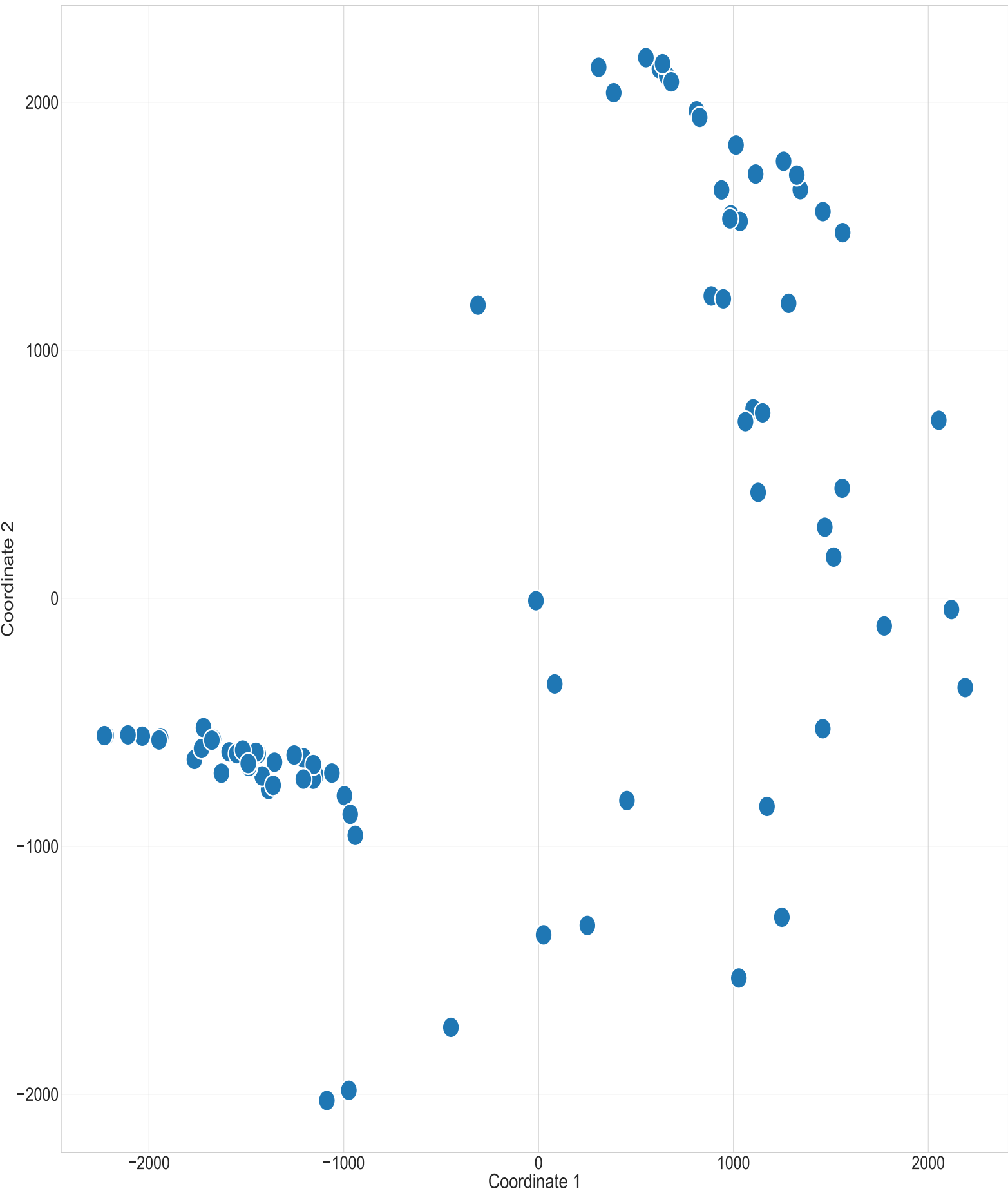
Scatter plot of Hybrid-(Not-PKS-NRPS) *Micromonosporaceae* BGCs analyzed via multidimensional scaling using Chebyshev pairwise distance. Each dot represents an individual BGC. Distance between BGCs is associated with GCFs in BiG-SLiCE they were most similar to.

Figure – S8 :



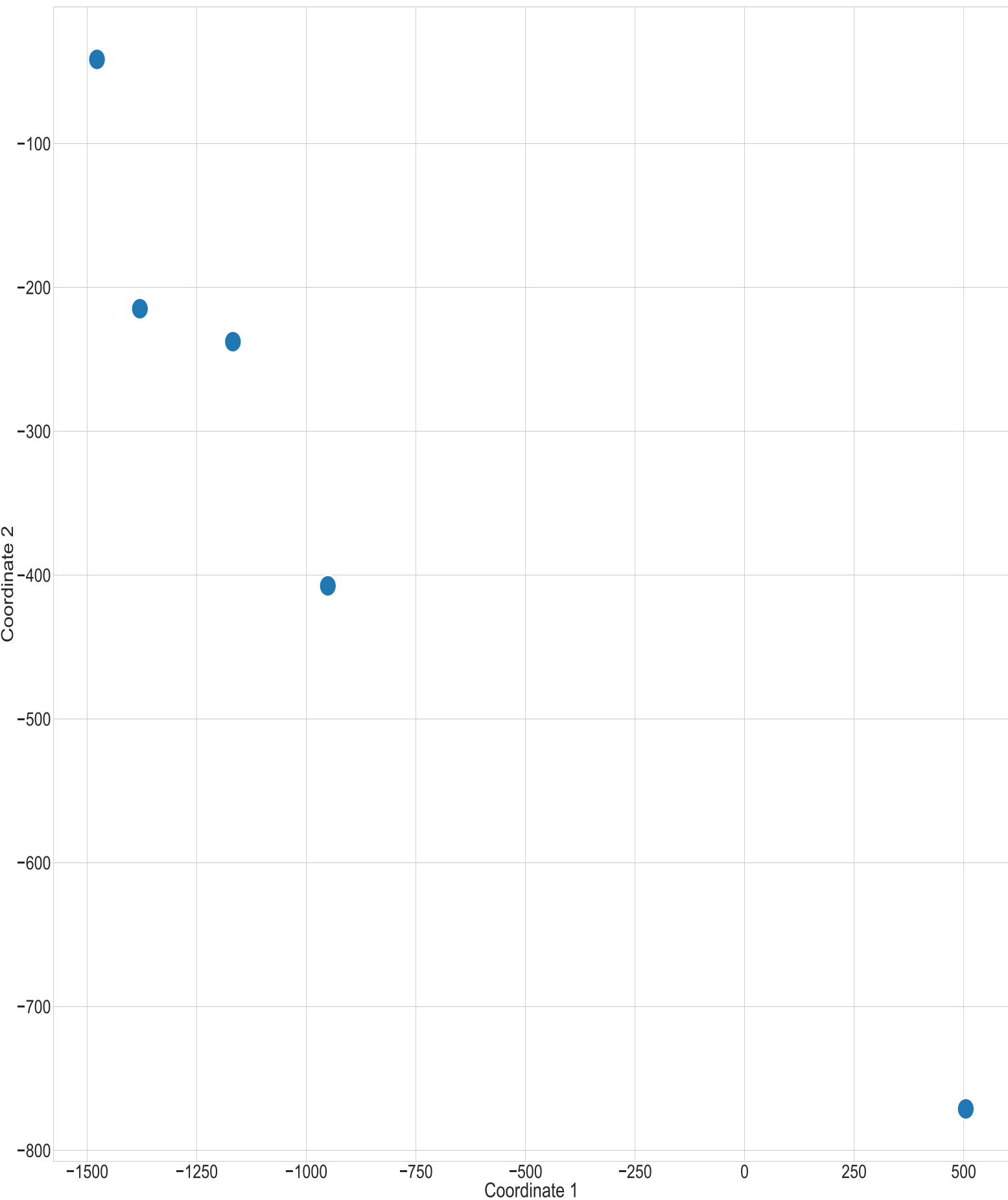
Scatter plot of Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using Chebyshev pairwise distance. Each dot represents an individual BGC. Distance between BGCs is associated with GCFs in BiG-SLiCE they were most similar to.

Figure – S9 :



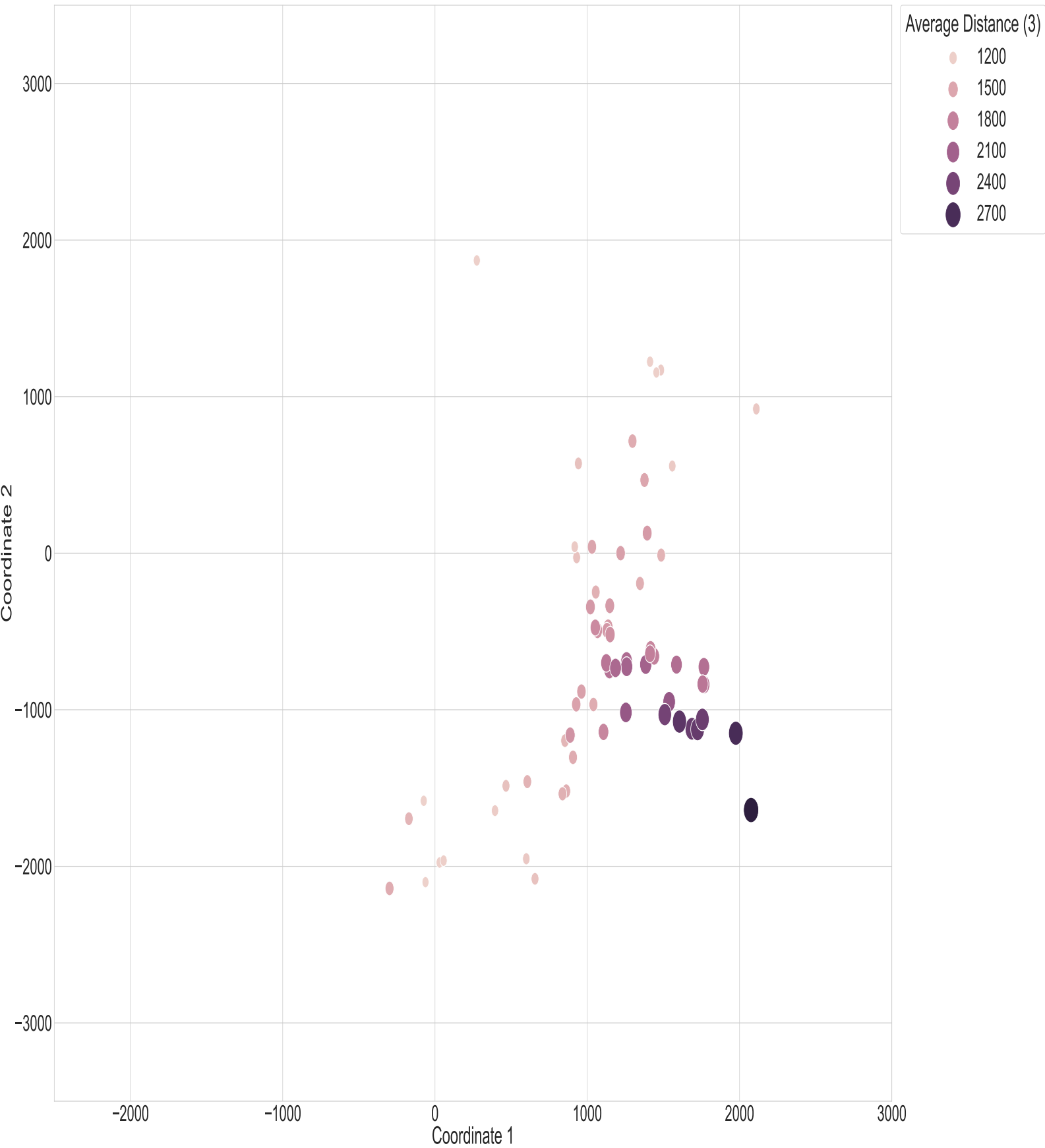
Scatter plot of PKS-Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using Chebyshev pairwise distance. Each dot represents an individual BGC. Distance between BGCs is associated with GCFs in BiG-SLiCE they were most similar to.

Figure – S10 :



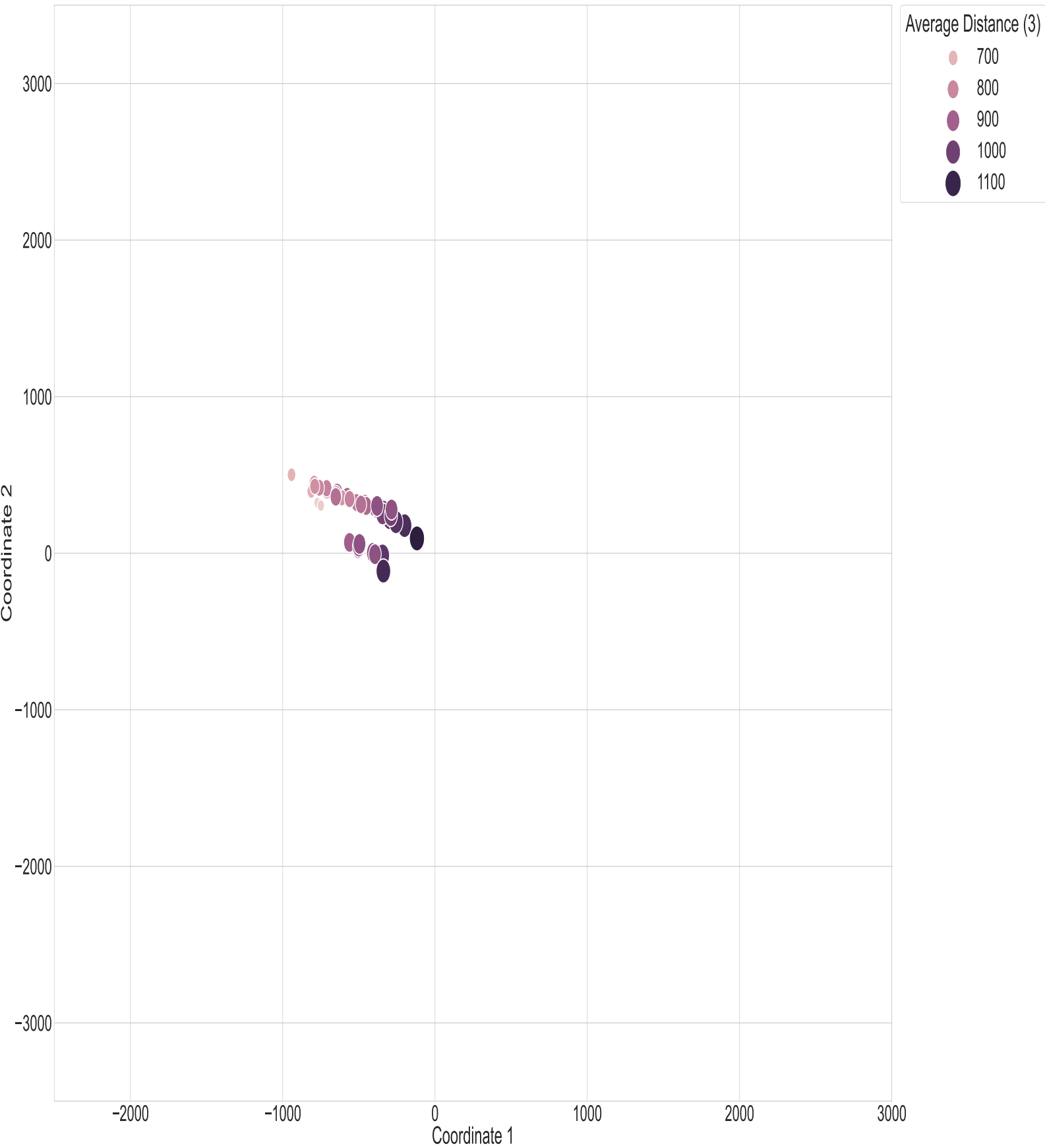
Scatter plot of Saccharides *Micromonosporaceae* BGCs analyzed via multidimensional scaling using Chebyshev pairwise distance. Each dot represents an individual BGC. Distance between BGCs is associated with GCFs in BiG-SLiCE they were most similar to.

Figure – S11:



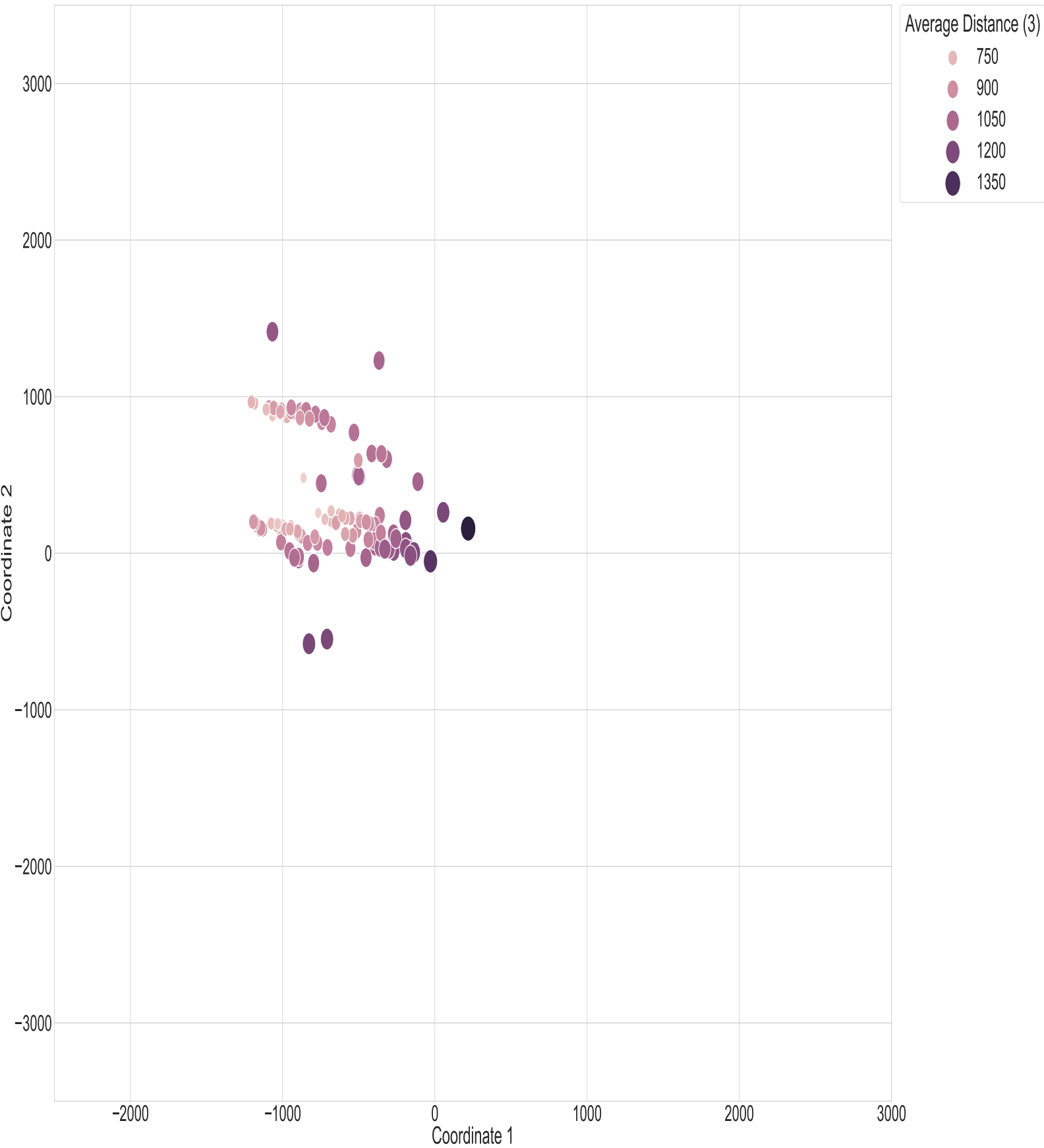
Scatter plot of PKS-NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using Chebyshev pairwise distance. Each dot represents an individual BGC. The hue and size of the dots were scaled based on the average distance of the individual BGC, across the entire dataset, to the closest 3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

Figure – S12 :



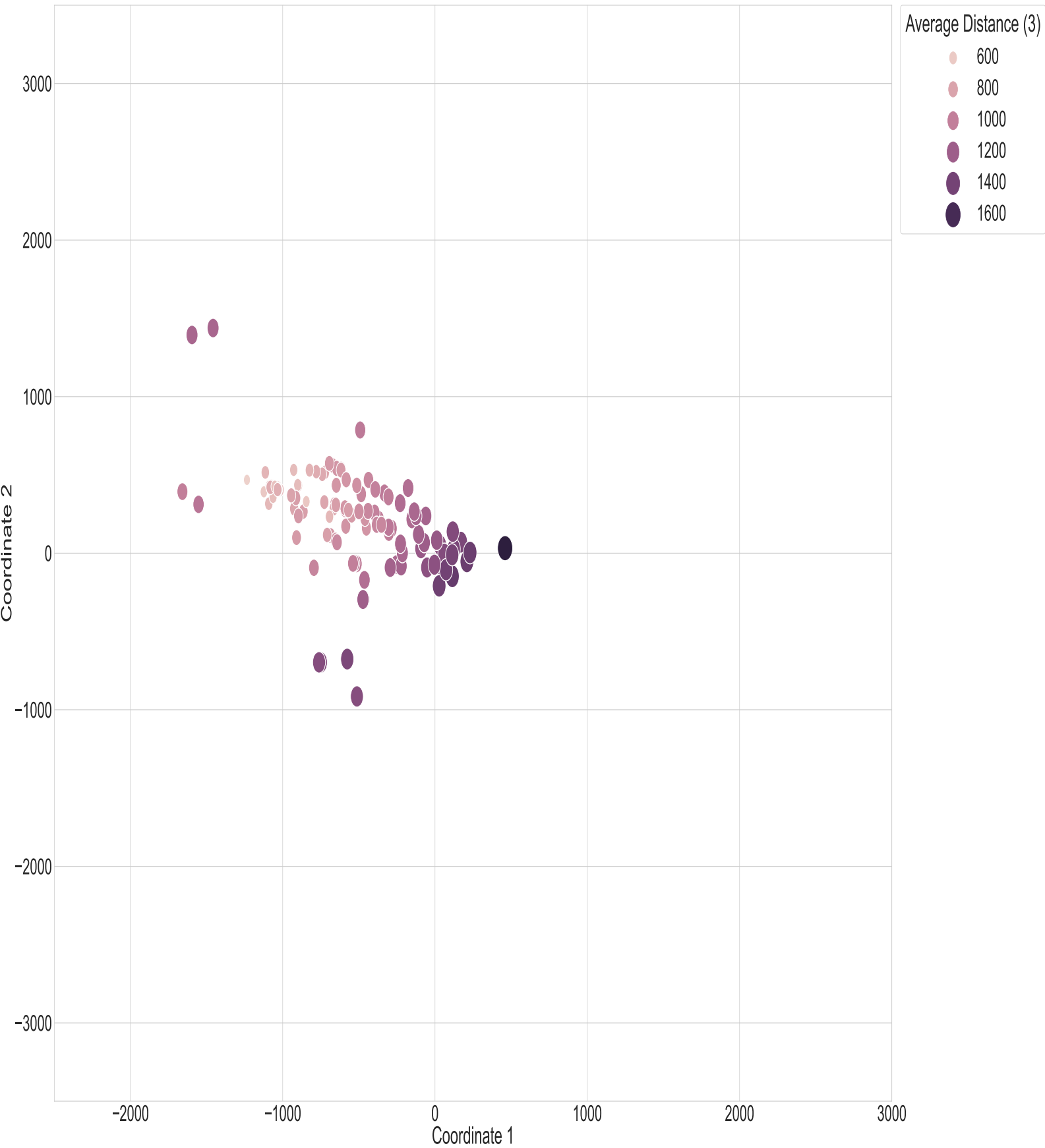
Scatter plot of Siderophore *Micromonosporaceae* BGCs analyzed via multidimensional scaling using Chebyshev pairwise distance. Each dot represents an individual BGC. The hue and size of the dots were scaled based on the average distance of the individual BGC, across the entire dataset, to the closest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

Figure – S13 :



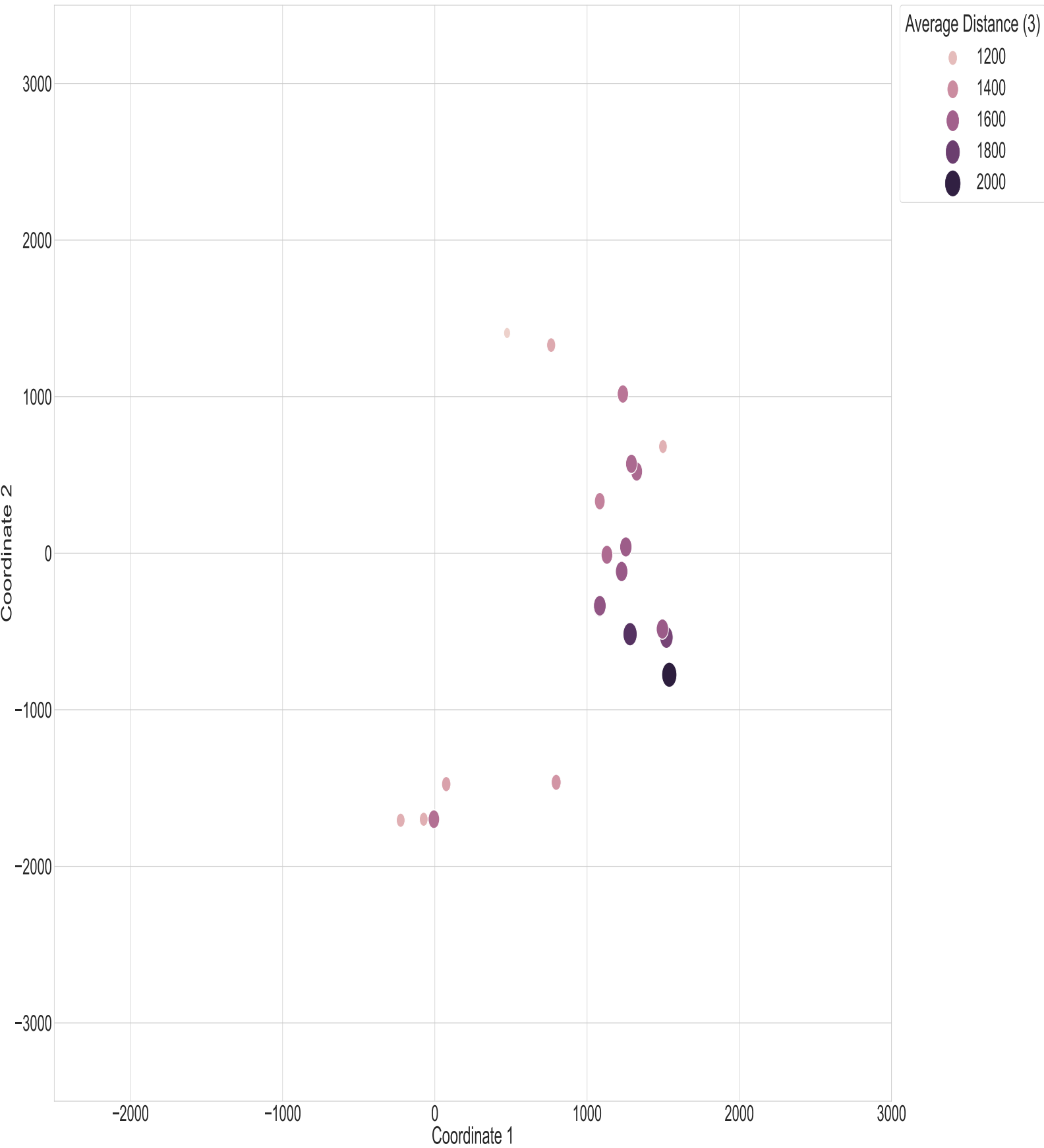
Scatter plot of Terpene *Micromonosporaceae* BGCs analyzed via multidimensional scaling using Chebyshev pairwise distance. Each dot represents an individual BGC. The hue and size of the dots were scaled based on the average distance of the individual BGC, across the entire dataset, to the closest 3 GCFs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

Figure – S14 :



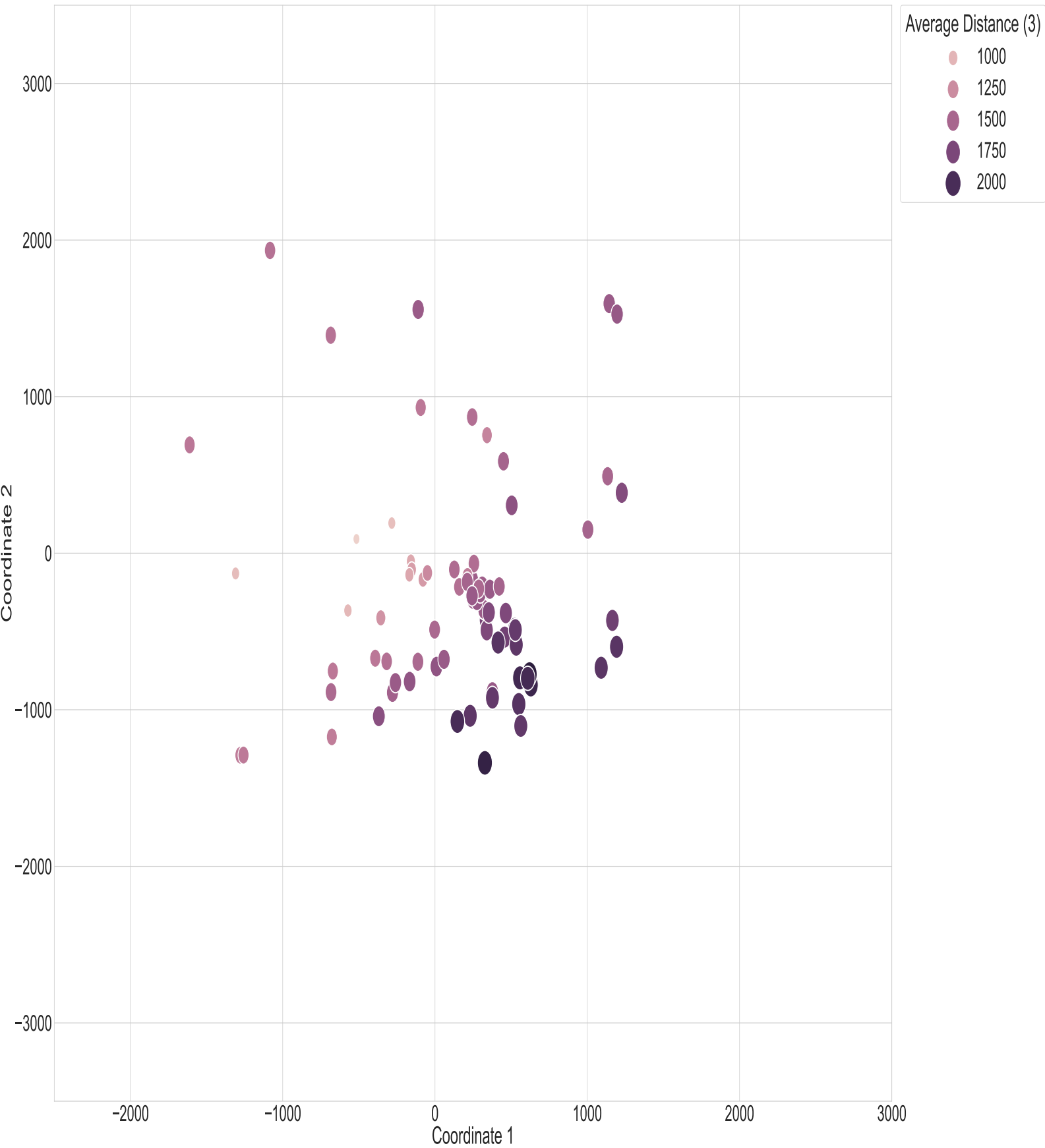
Scatter plot of RIPP *Micromonosporaceae* BGCs analyzed via multidimensional scaling using Chebyshev pairwise distance. Each dot represents an individual BGC. The hue and size of the dots were scaled based on the average distance of the individual BGC, across the entire dataset, to the closest 3 GCFs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

Figure – S15 :



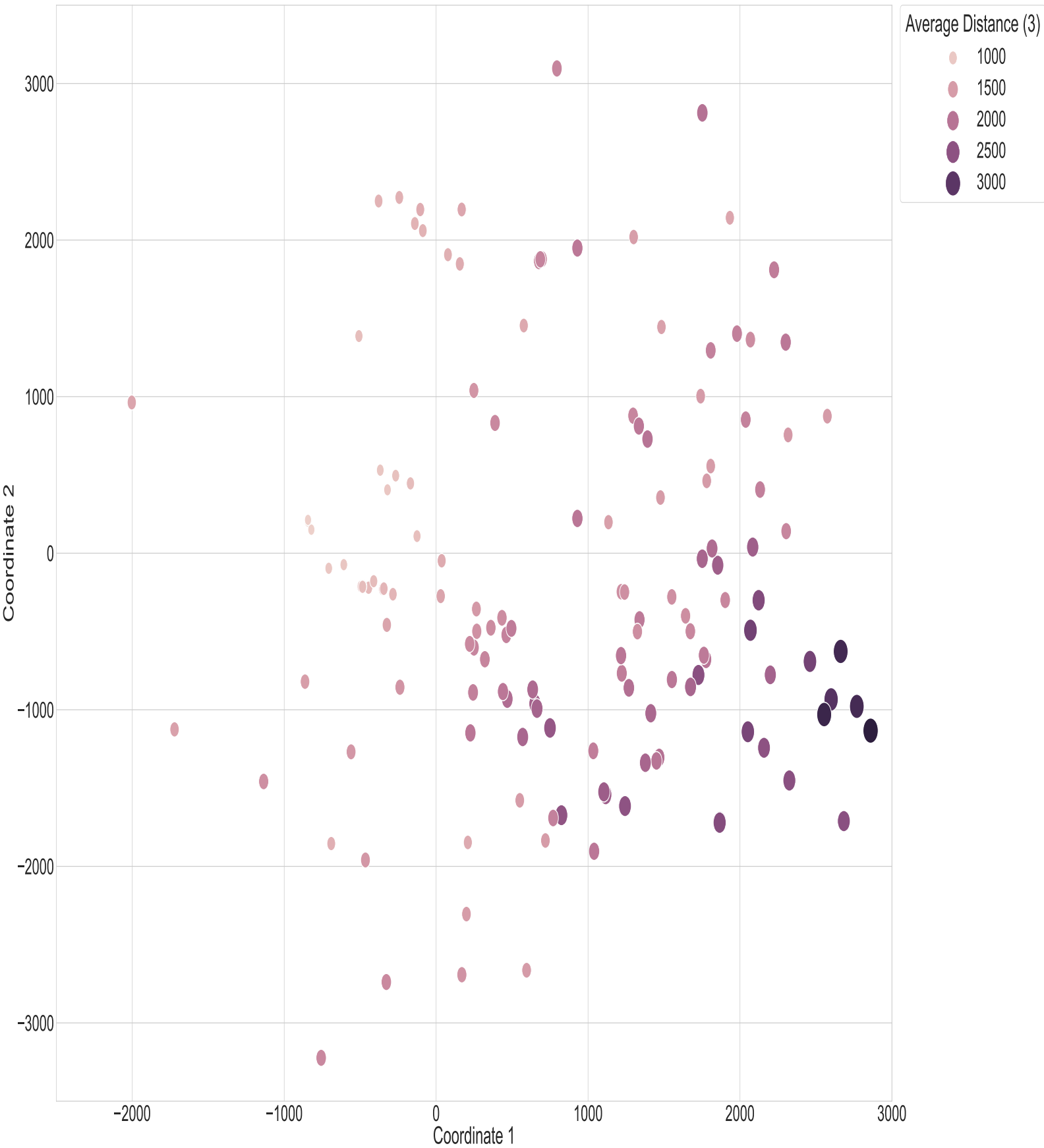
Scatter plot of PKS-1 *Micromonosporaceae* BGCs analyzed via multidimensional scaling using Chebyshev pairwise distance. Each dot represents an individual BGC. The hue and size of the dots were scaled based on the average distance of the individual BGC, across the entire dataset, to the closest 3 GCFs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

Figure – S16 :



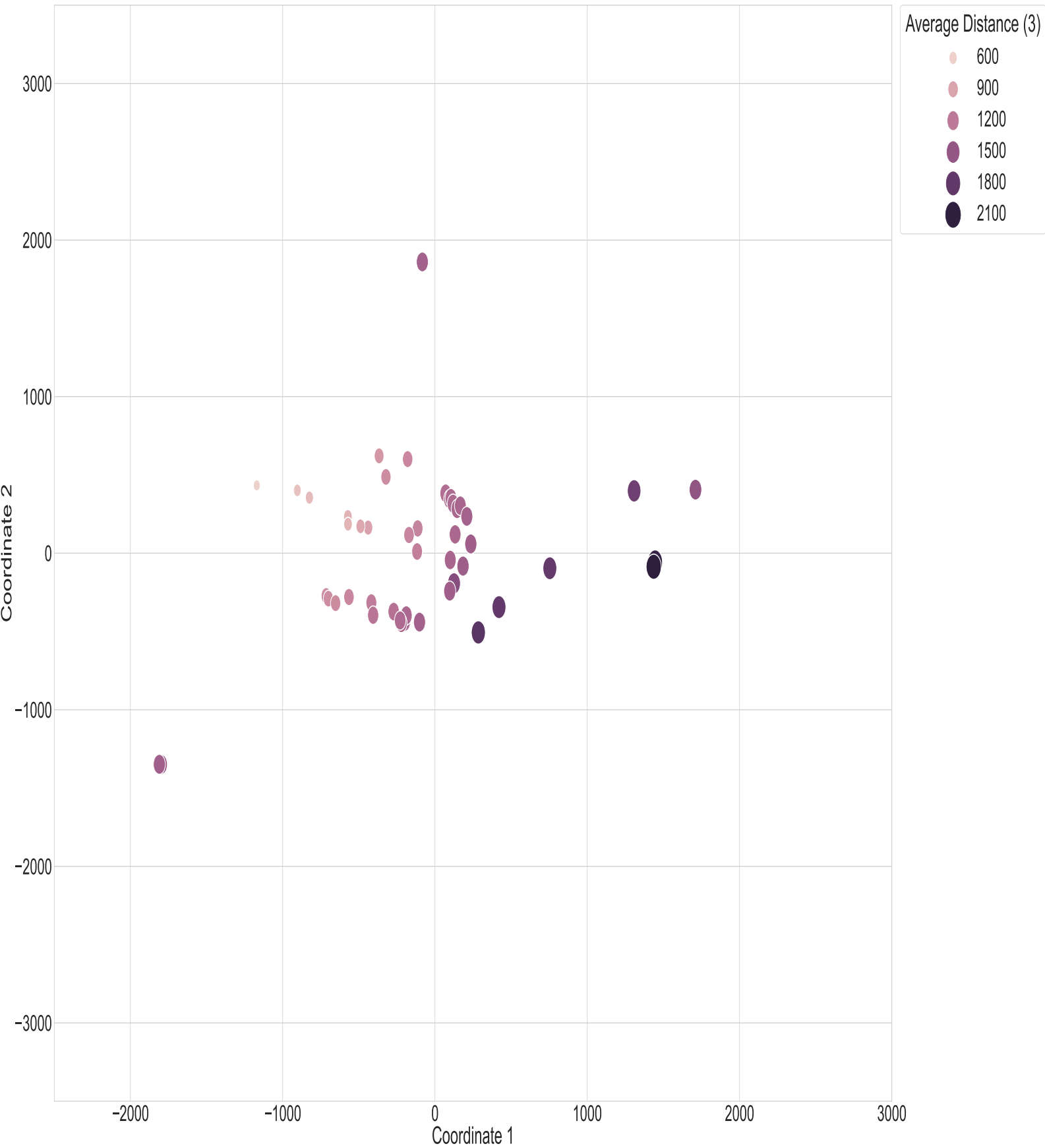
Scatter plot of NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using Chebyshev pairwise distance. Each dot represents an individual BGC. The hue and size of the dots were scaled based on the average distance of the individual BGC, across the entire dataset, to the closest 3 GCFs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



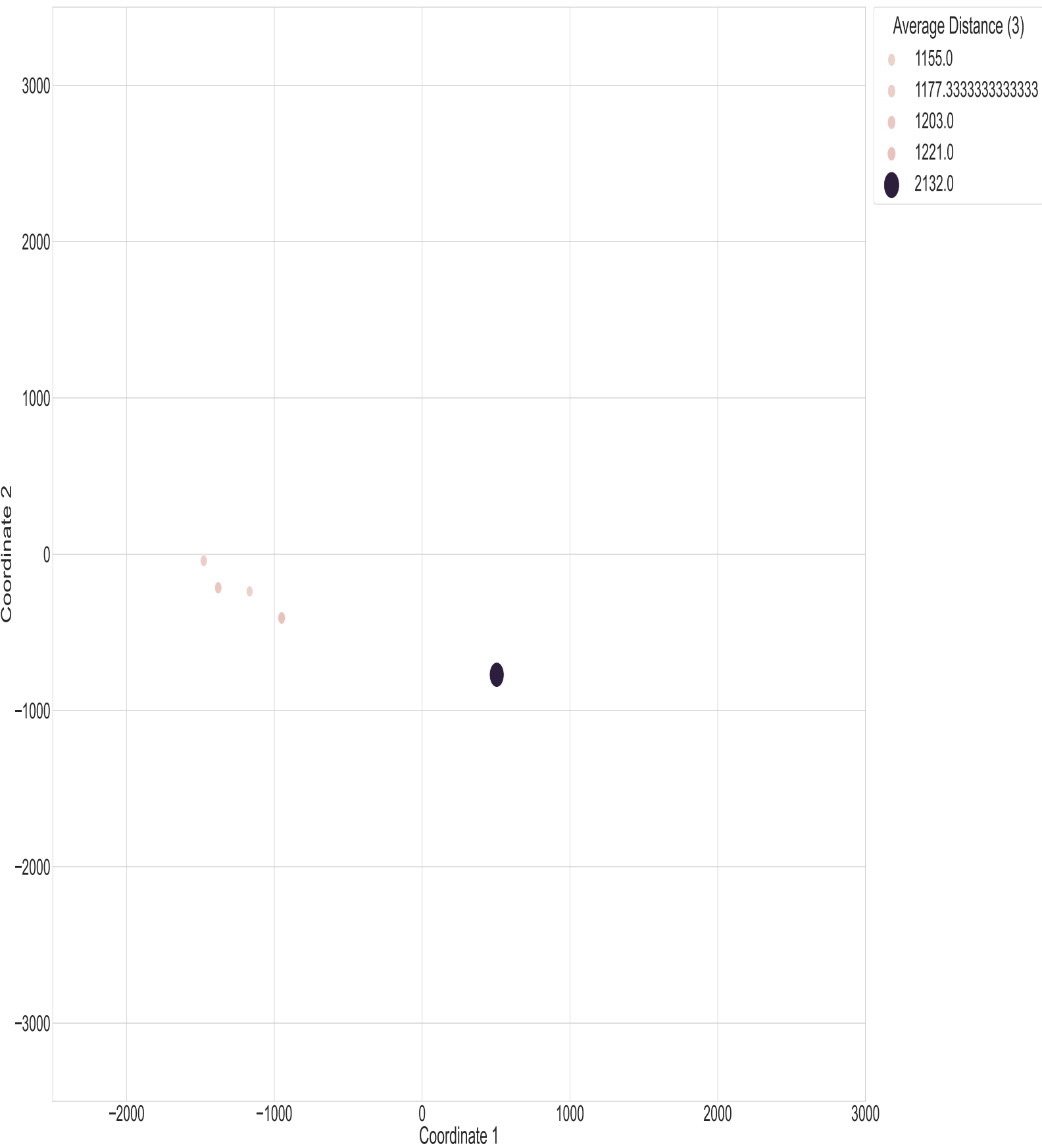
Scatter plot of Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using Chebyshev pairwise distance. Each dot represents an individual BGC. The hue and size of the dots were scaled based on the average distance of the individual BGC, across the entire dataset, to the closest 3 GCFs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

Figure – S19:



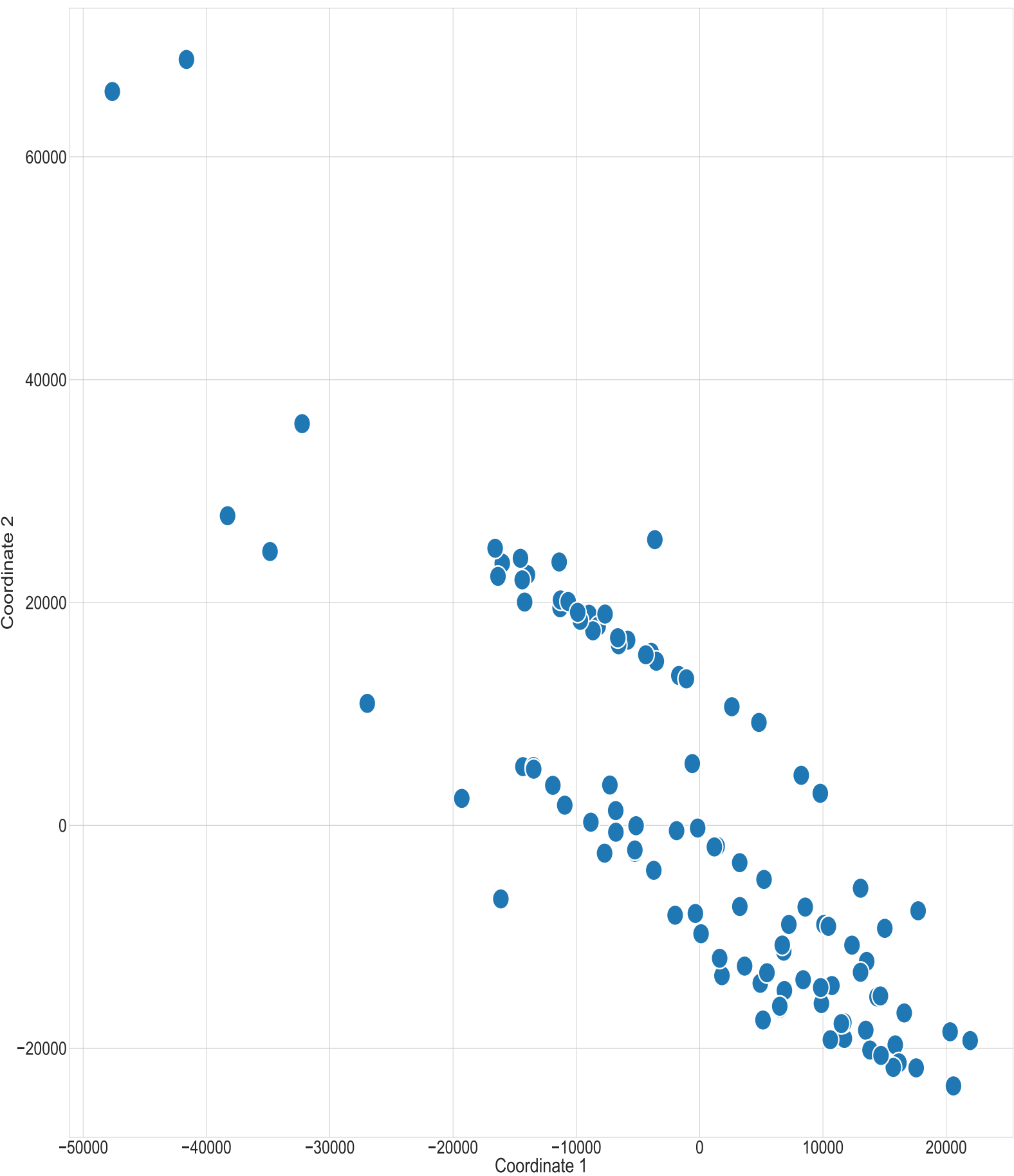
Scatter plot of PKS-Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using Chebyshev pairwise distance. Each dot represents an individual BGC. The hue and size of the dots were scaled based on the average distance of the individual BGC, across the entire dataset, to the closest 3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

Figure – S20 :



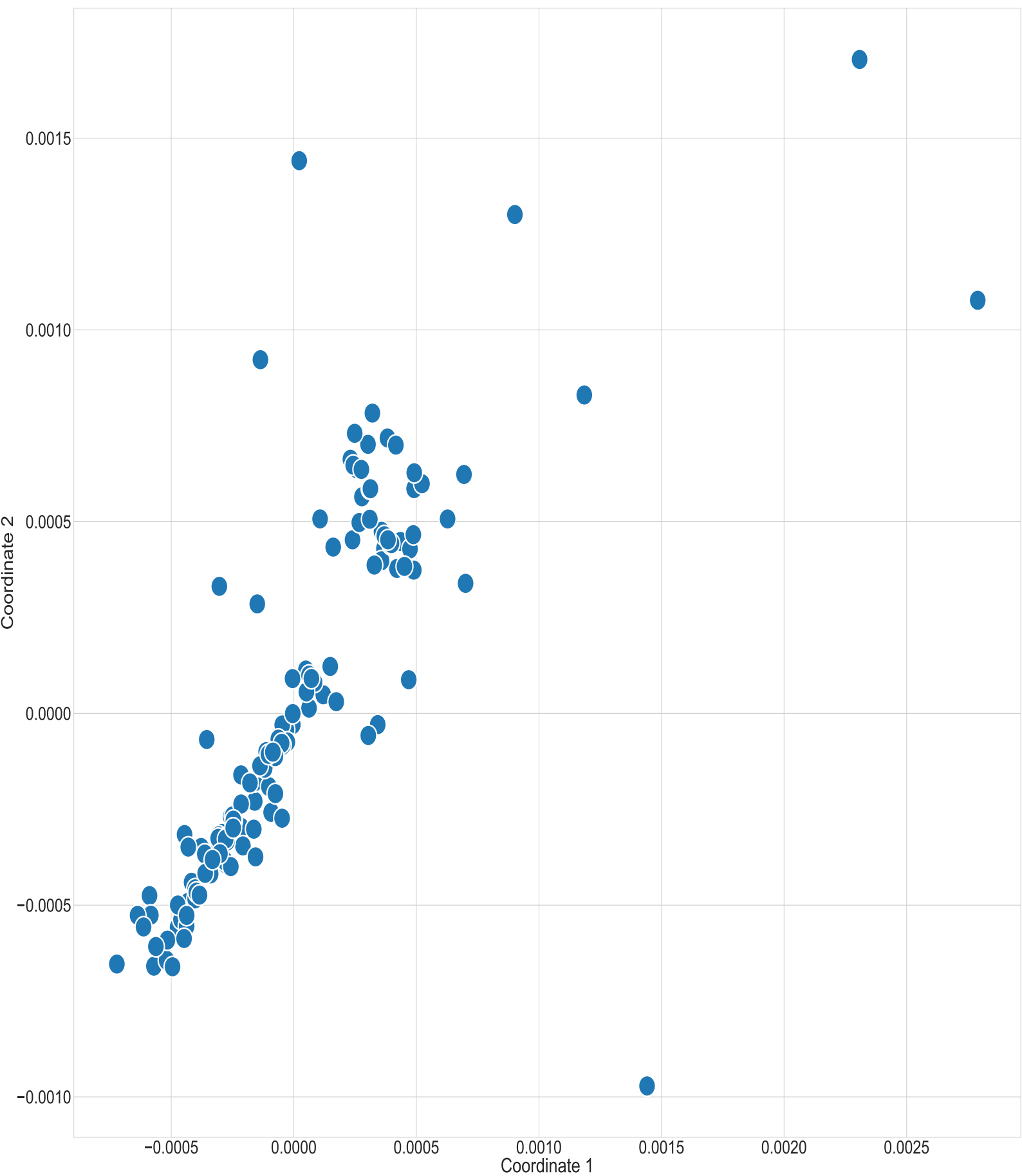
Scatter plot of Saccharides *Micromonosporaceae* BGCs analyzed via multidimensional scaling using Chebyshev pairwise distance. Each dot represents an individual BGC. The hue and size of the dots were scaled based on the average distance of the individual BGC, across the entire dataset, to the closest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

Figure – S21 :



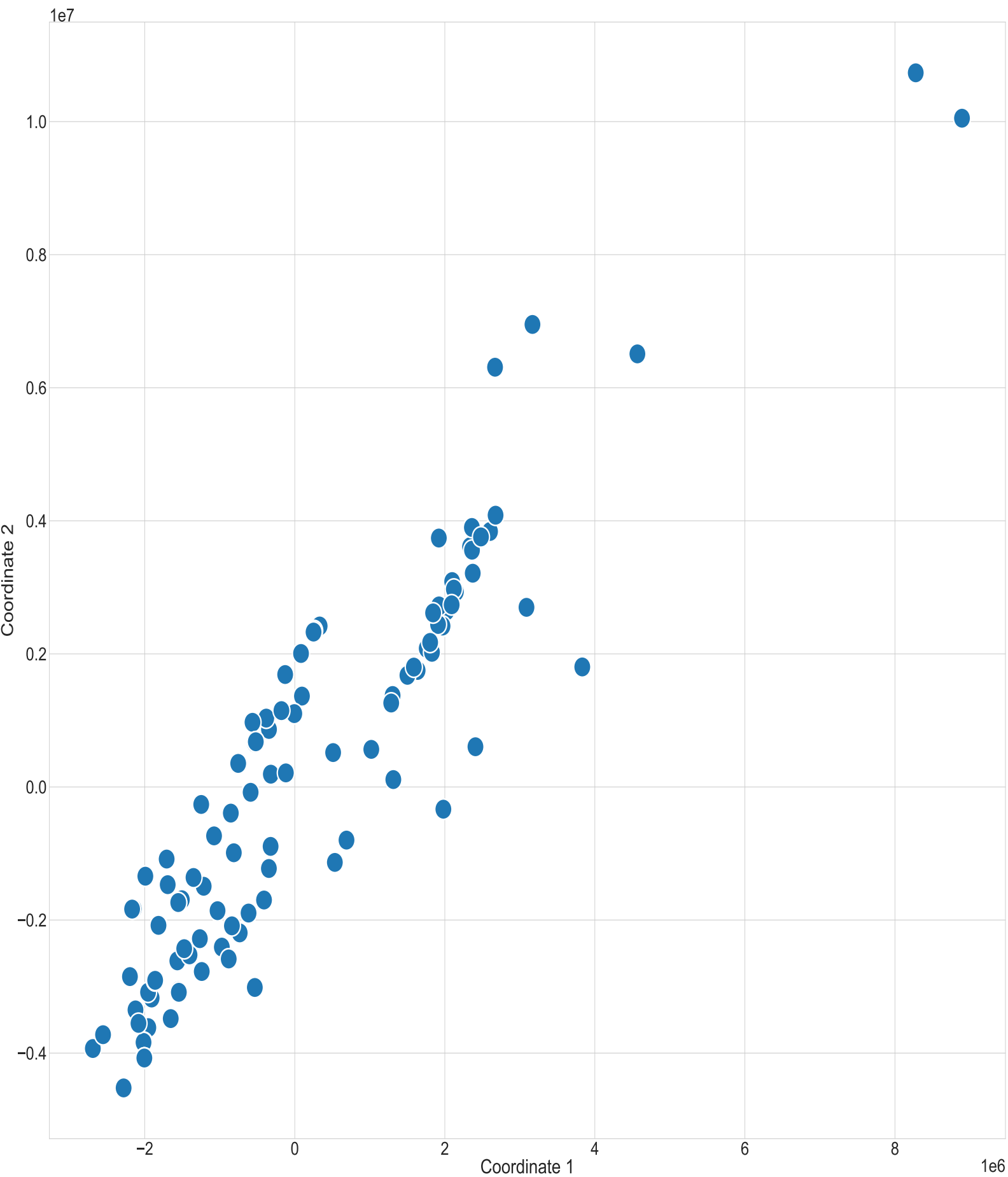
Scatter plot of Terpene *Micromonosporaceae* BGCs analyzed via multidimensional scaling using euclidean pairwise distance. Only Terpene BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Terpene BGCs is associated with the GCFs in BiG_SLiCE they were most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

Figure – S22 :



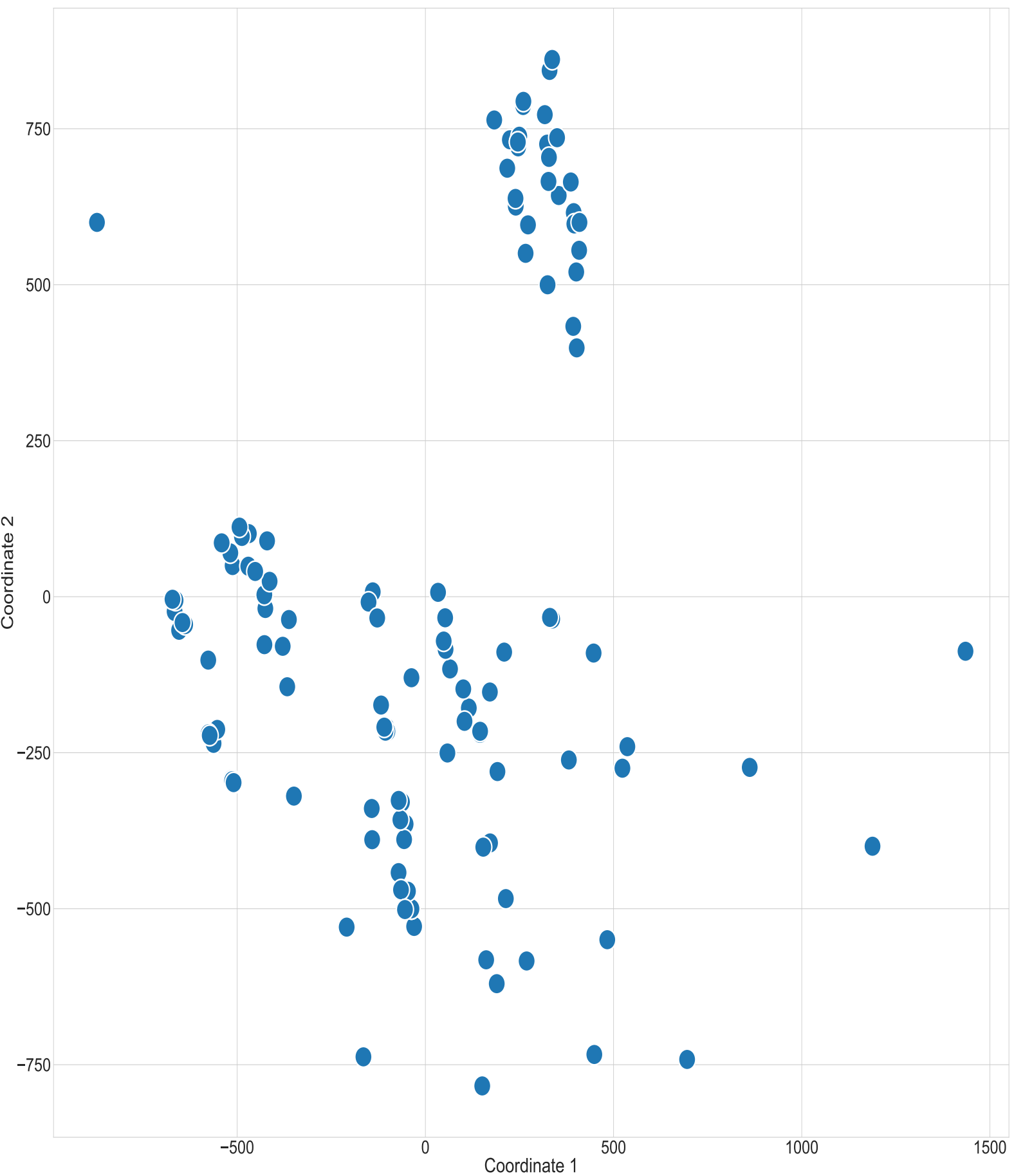
Scatter plot of Terpene *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cosine pairwise distance. Only Terpene BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Terpene BGCs is associated with the GCFs in BiG_SLICE they were most similar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

Figure – S23 :



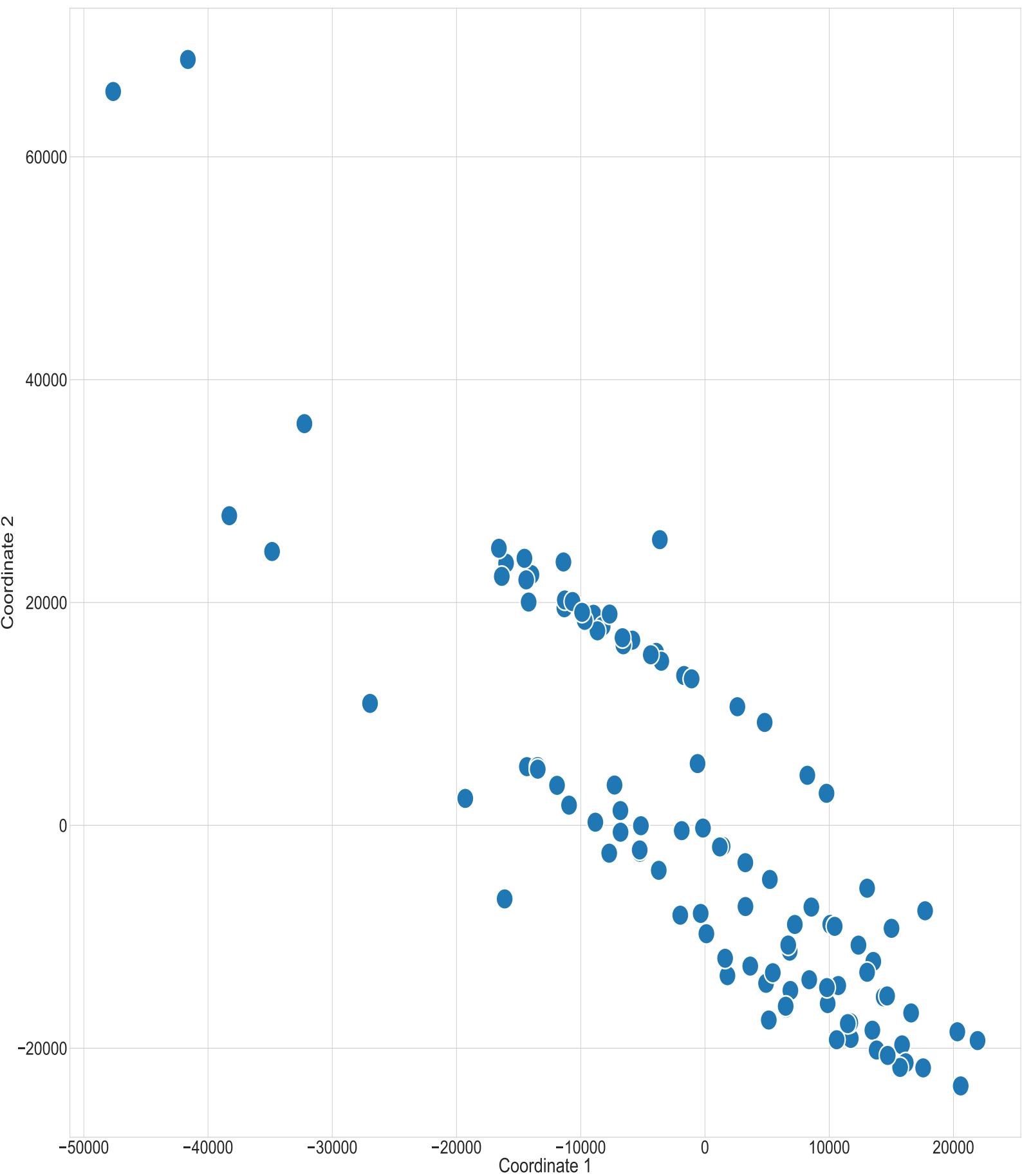
Scatter plot of Terpene *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cityblock pairwise distance. Only Terpene BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Terpene BGCs is associated with the GCFs in BiG_SLiCE they were most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

Figure – S24 :



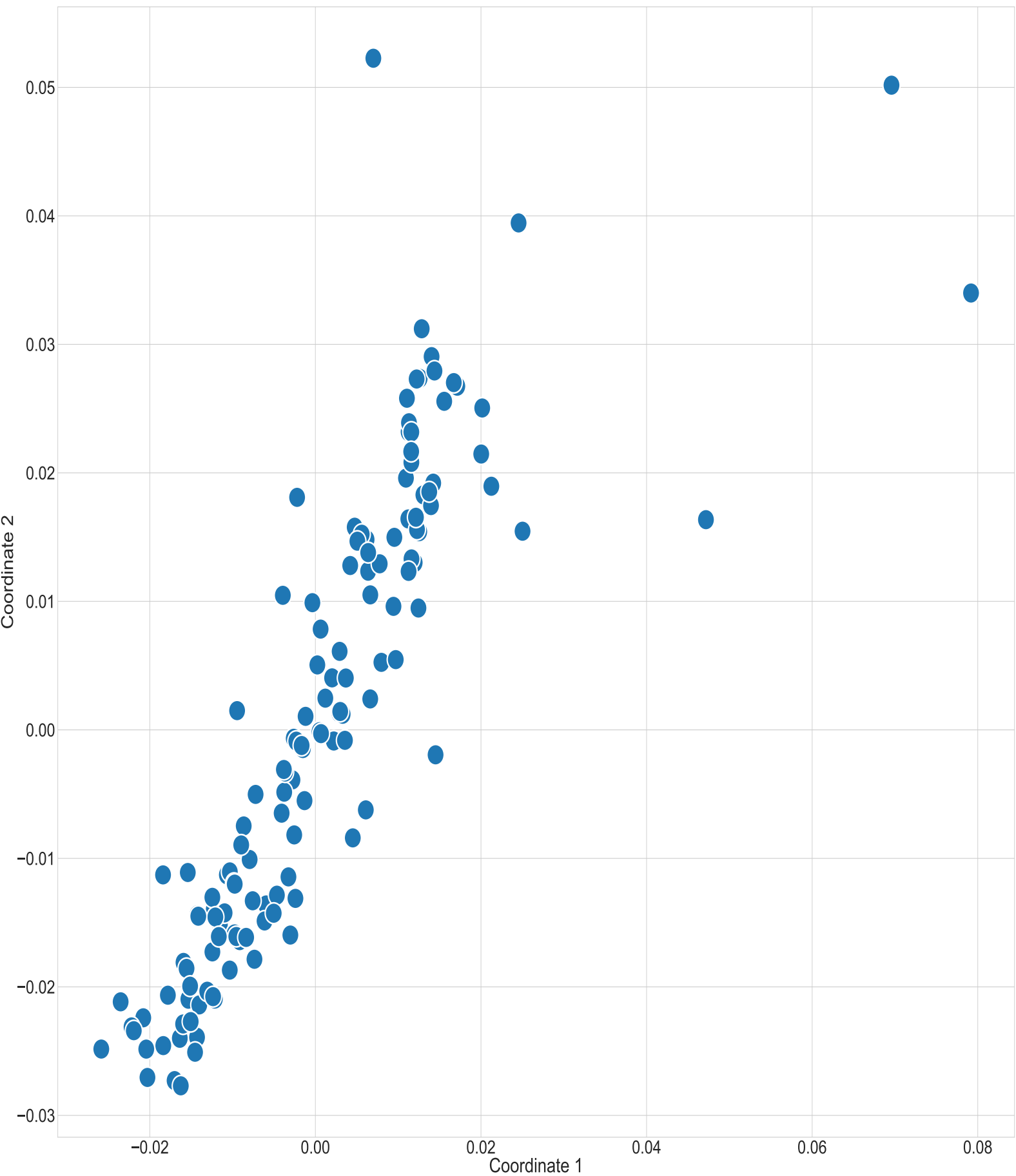
Scatter plot of Terpene *Micromonosporaceae* BGCs analyzed via multidimensional scaling using chebyshev pairwise distance. Only Terpene BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Terpene BGCs is associated with the GCFs in BiG_SLiCE they were most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

Figure – S25 :



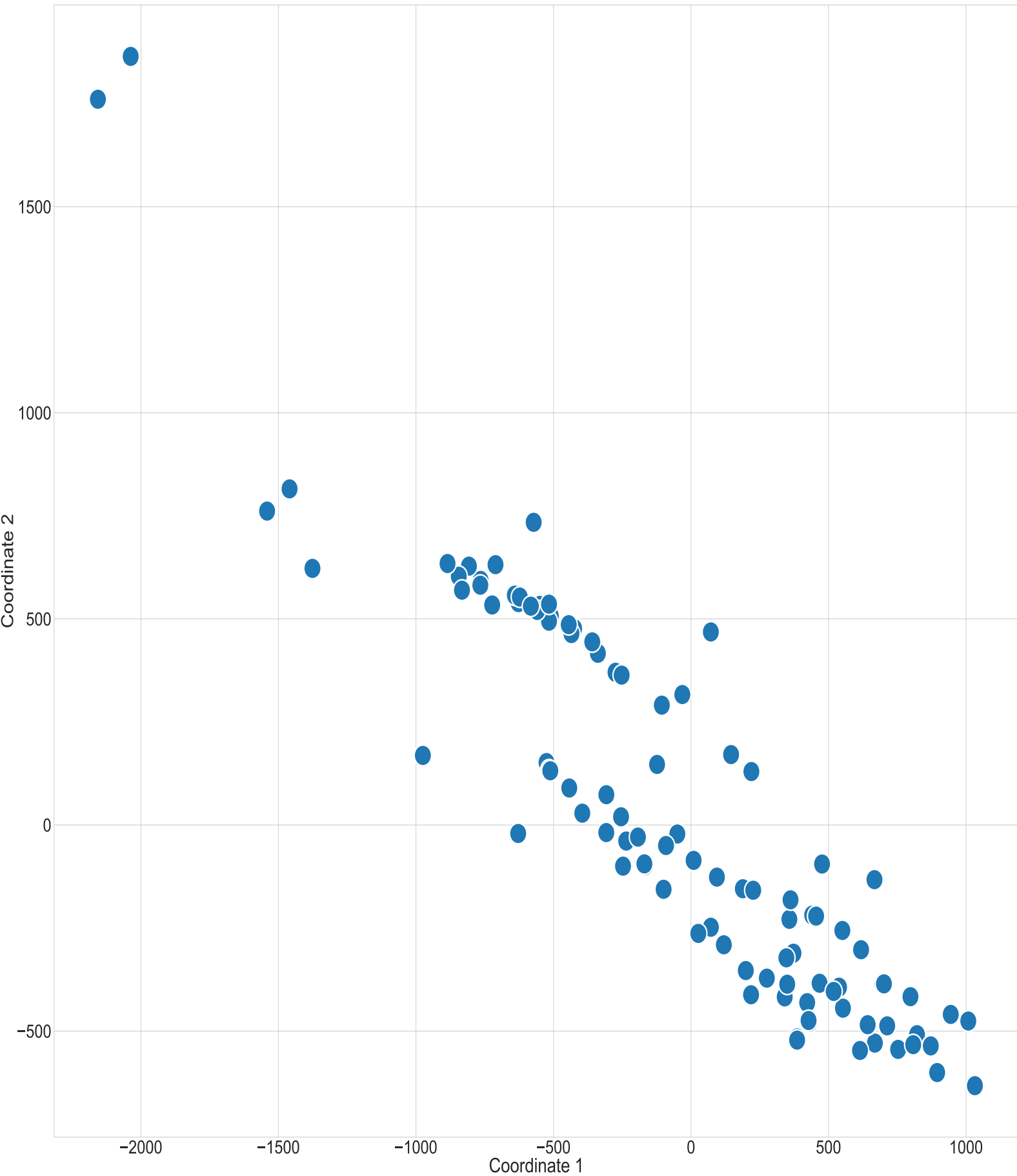
Scatter plot of Terpene *Micromonosporaceae* BGCs analyzed via multidimensional scaling using I2 pairwise distance. Only Terpene BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Terpene BGCs is associated with the GCFs in BiG_SLICE they were most similar to (T=900).

Figure – S26 :



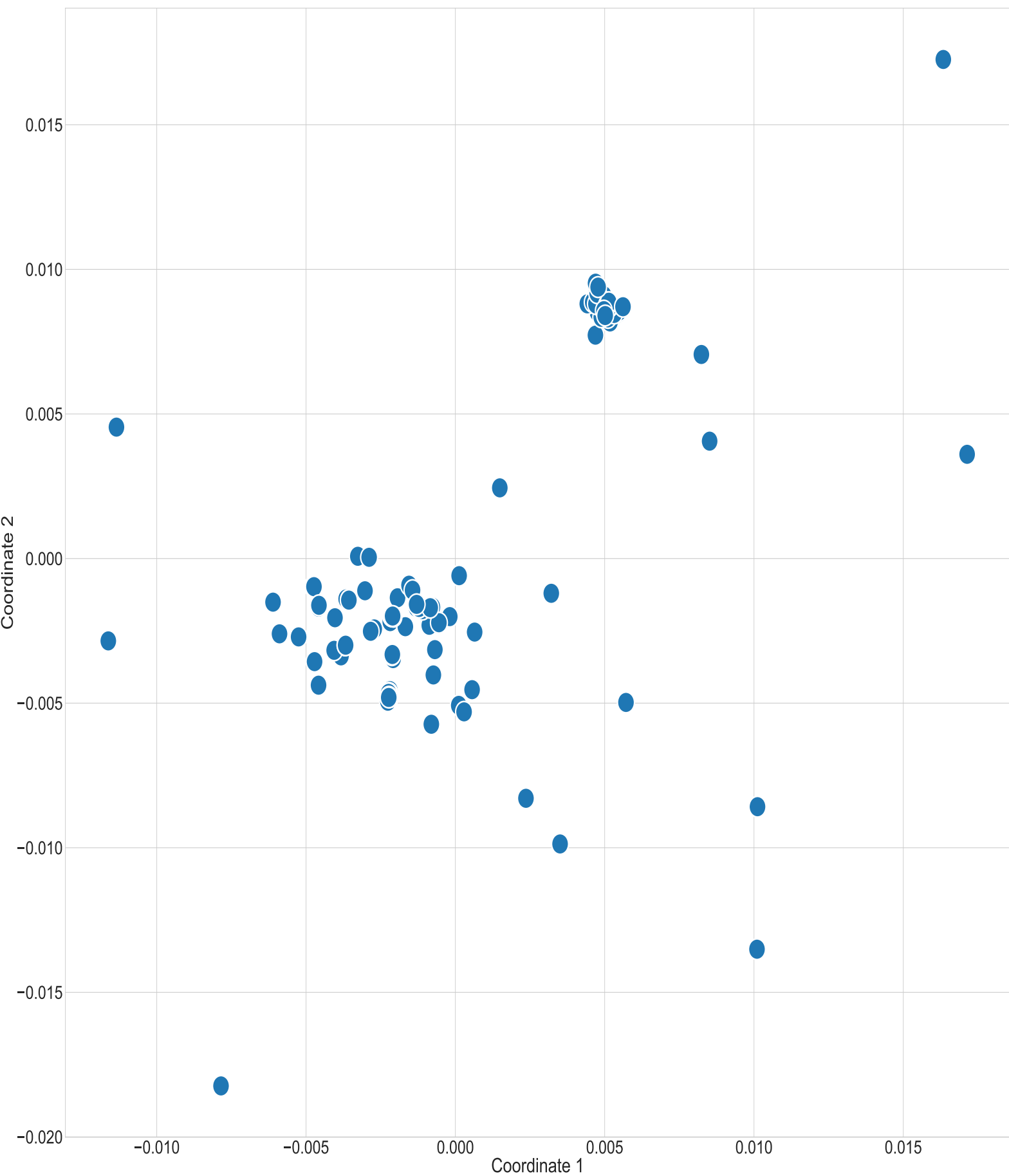
Scatter plot of Terpene *Micromonosporaceae* BGCs analyzed via multidimensional scaling using braycurtis pairwise distance. Only Terpene BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Terpene BGCs is associated with the GCFs in BiG_SLiCE they were most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

Figure – S27 :



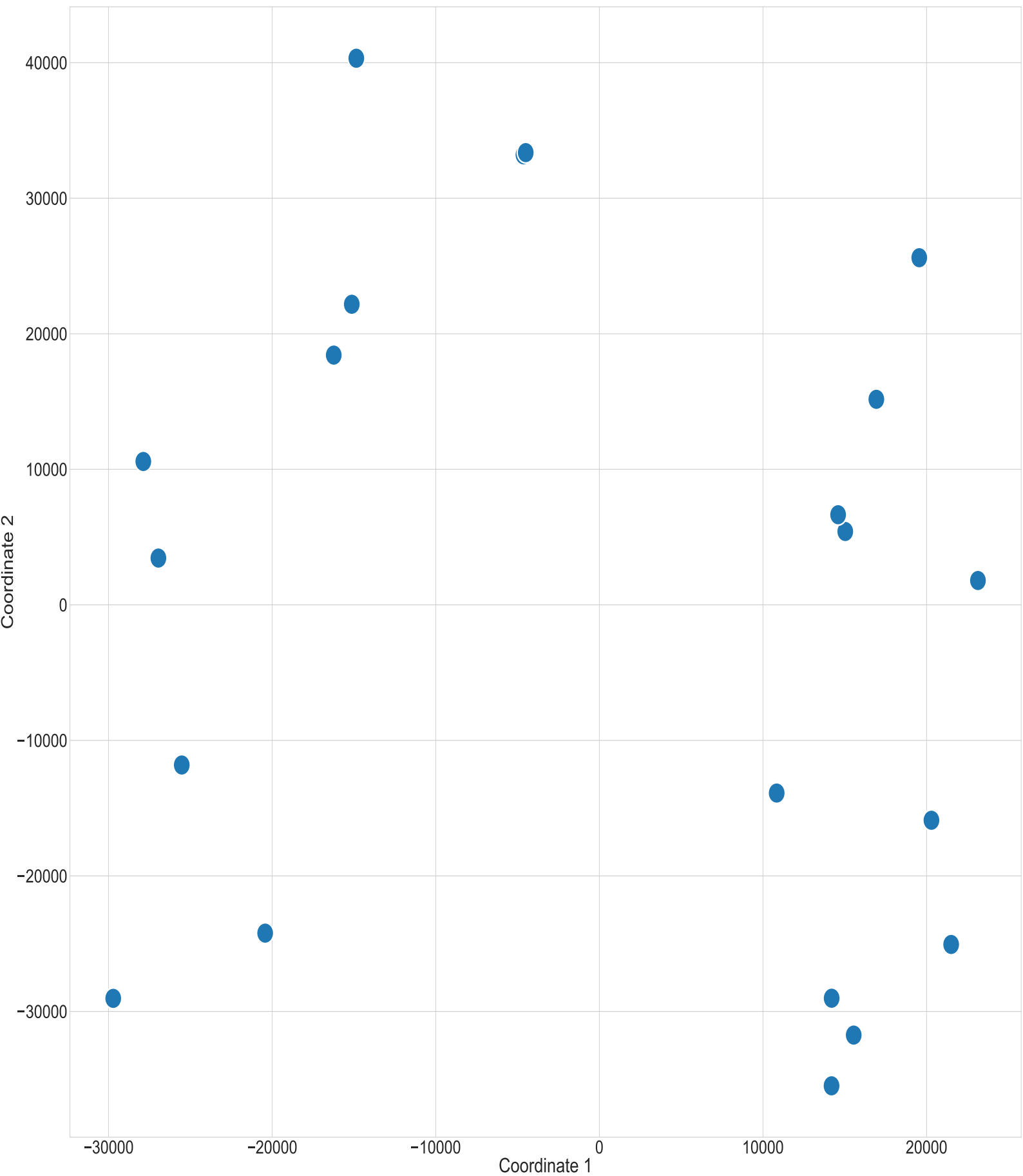
Scatter plot of Terpene *Micromonosporaceae* BGCs analyzed via multidimensional scaling using canberra pairwise distance. Only Terpene BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Terpene BGCs is associated with the GCFs in BiG_SLiCE they were most similar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

Figure – S28 :



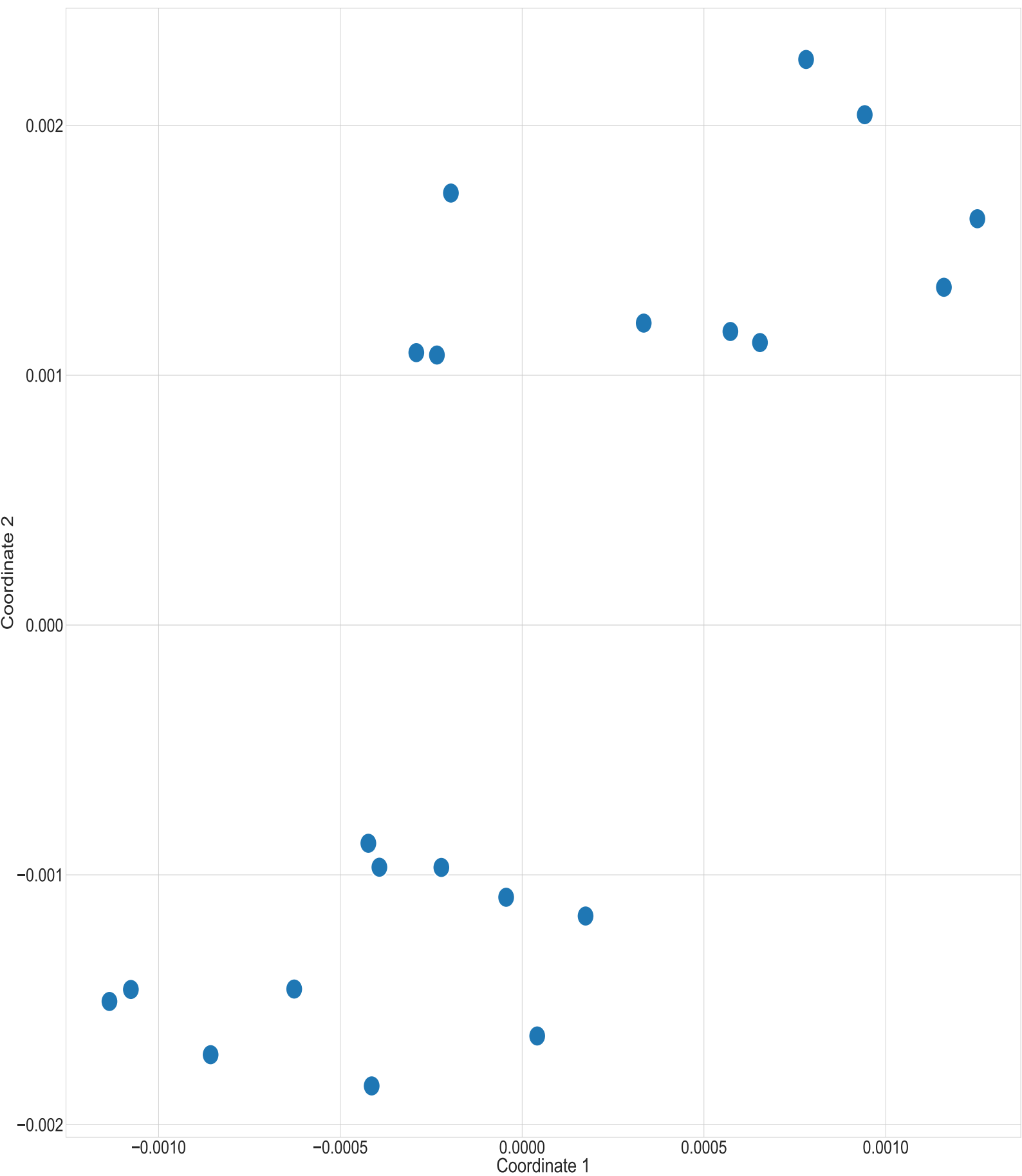
Scatter plot of Terpene *Micromonosporaceae* BGCs analyzed via multidimensional scaling using correlation pairwise distance. Only Terpene BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Terpene BGCs is associated with the GCFs in BiG_SLiCE they were most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

Figure – S29 :



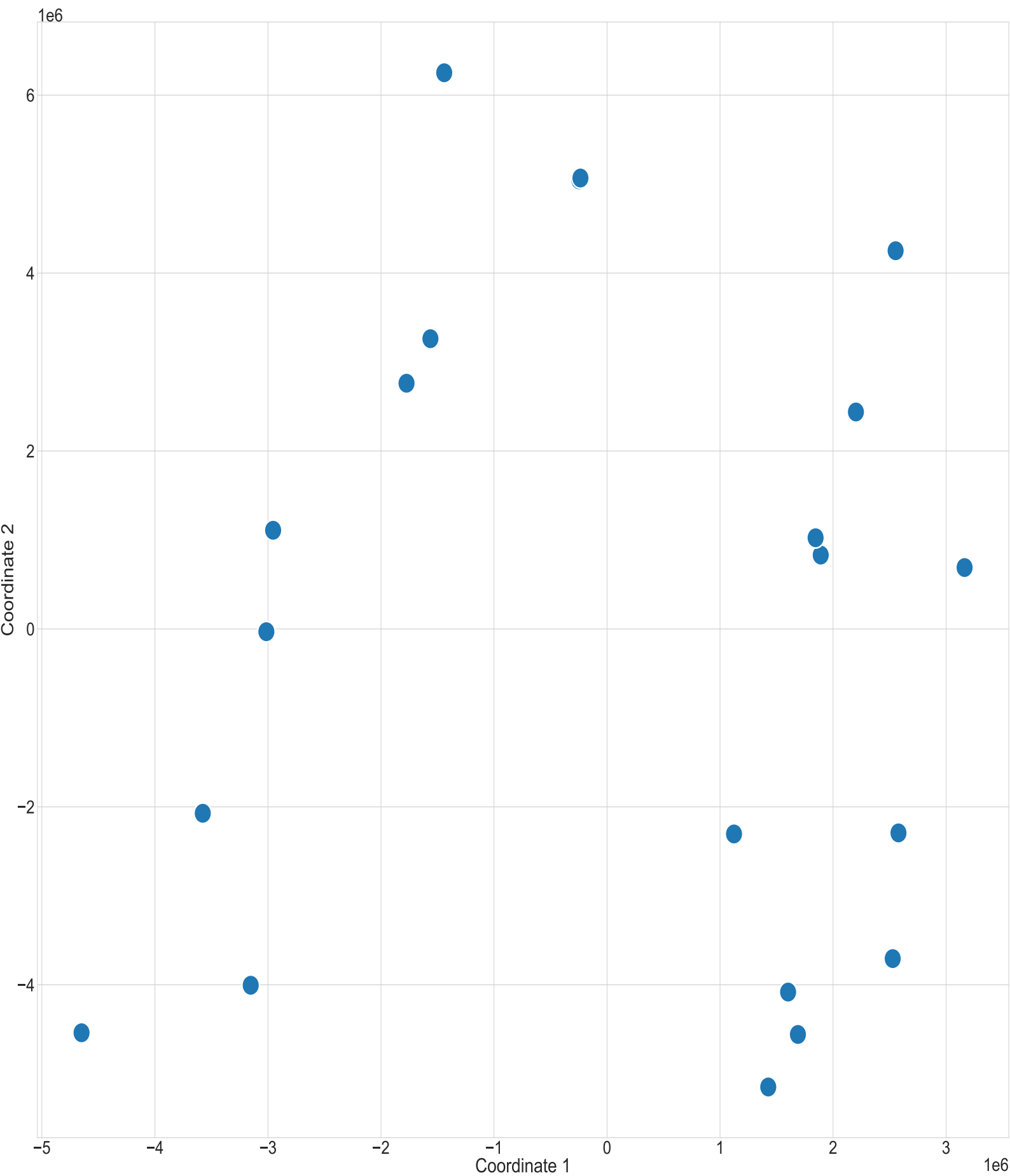
Scatter plot of PKS-1 *Micromonosporaceae* BGCs analyzed via multidimensional scaling using euclidean pairwise distance. Only PKS-1 BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between PKS-1 BGCs is associated with the GCFs in BiG_SLiCE they were most similar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

Figure – S30 :



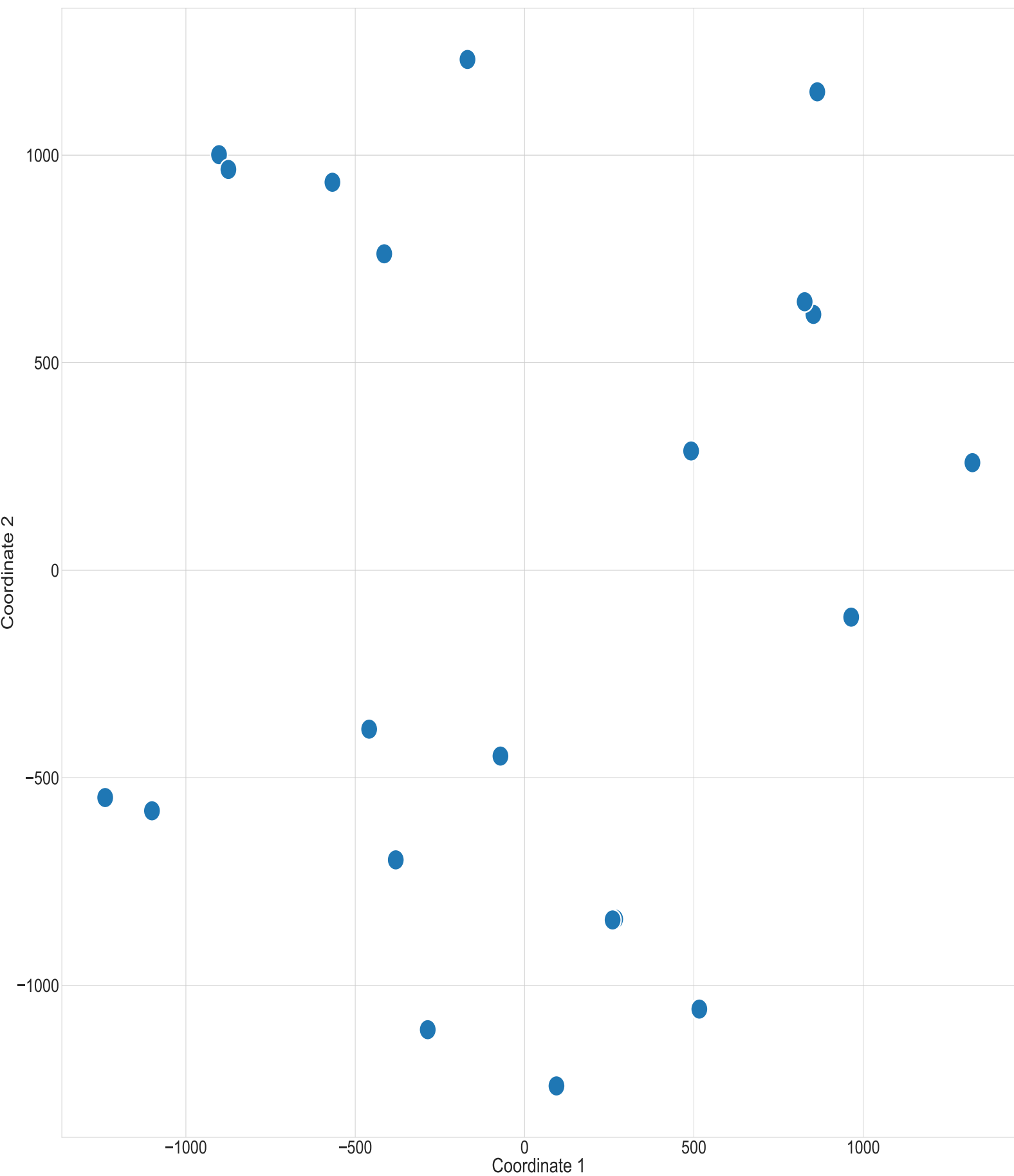
Scatter plot of PKS-1 *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cosine pairwise distance. Only PKS-1 BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between PKS-1 BGCs is associated with the GCFs in BiG_SLiCE they were most similar to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

Figure – S31 :



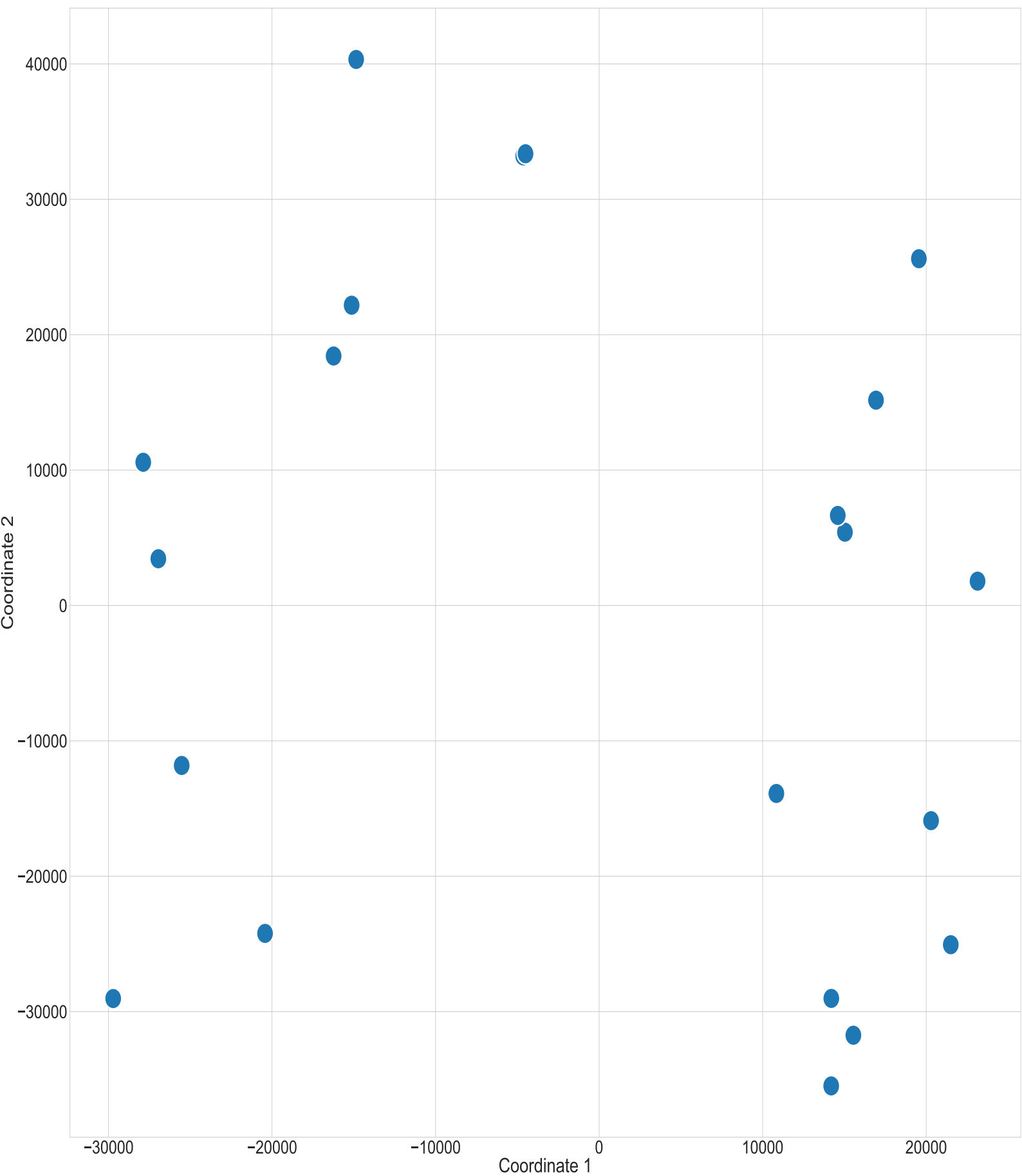
Scatter plot of PKS-1 *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cityblock pairwise distance. Only PKS-1 BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between PKS-1 BGCs is associated with the GCFs in BiG_SLiCE they were most similar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

Figure – S32 :



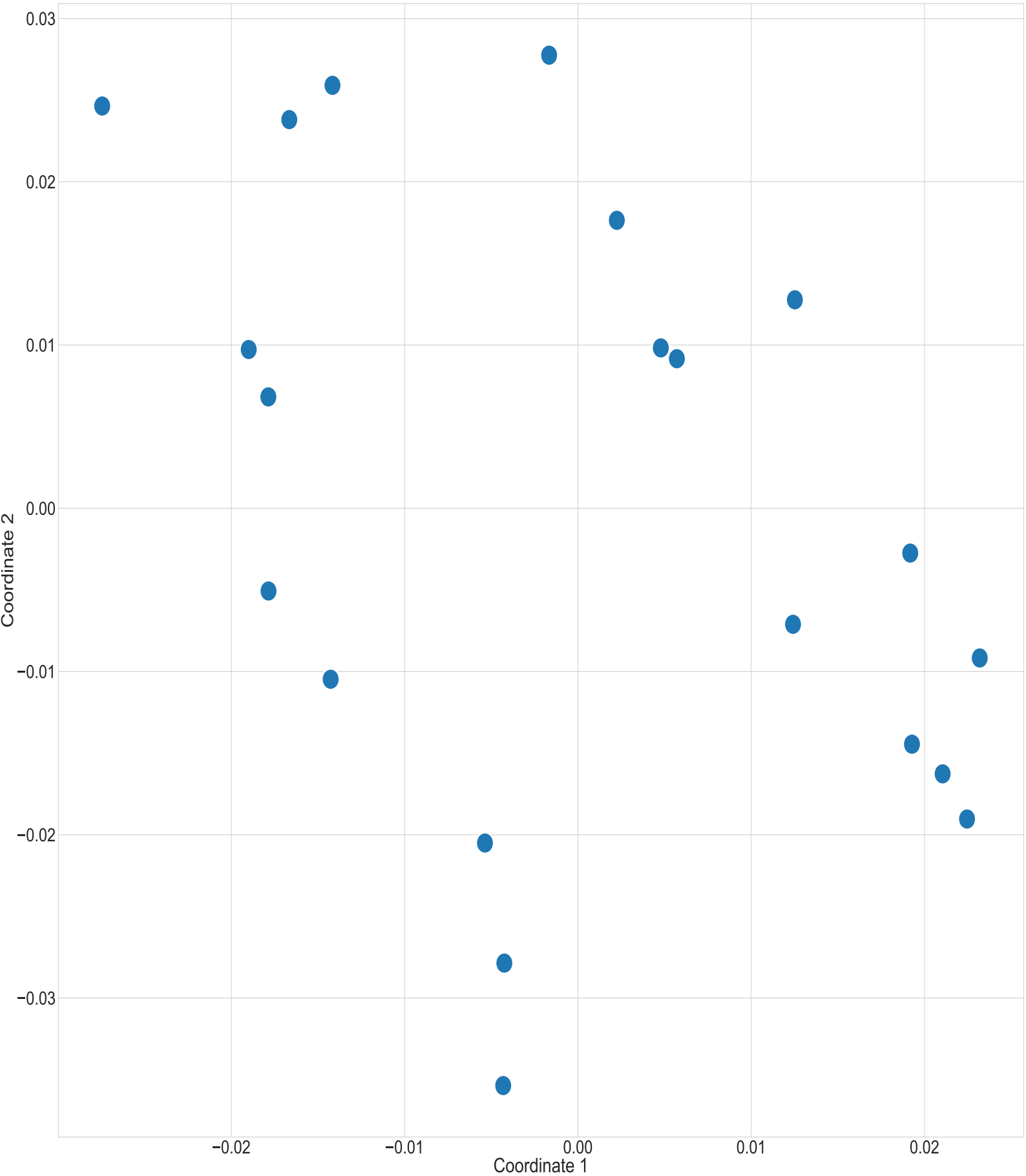
Scatter plot of PKS-1 *Micromonosporaceae* BGCs analyzed via multidimensional scaling using chebyshev pairwise distance. Only PKS-1 BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between PKS-1 BGCs is associated with the GCFs in BiG_SLiCE they were most similar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

Figure – S33 :



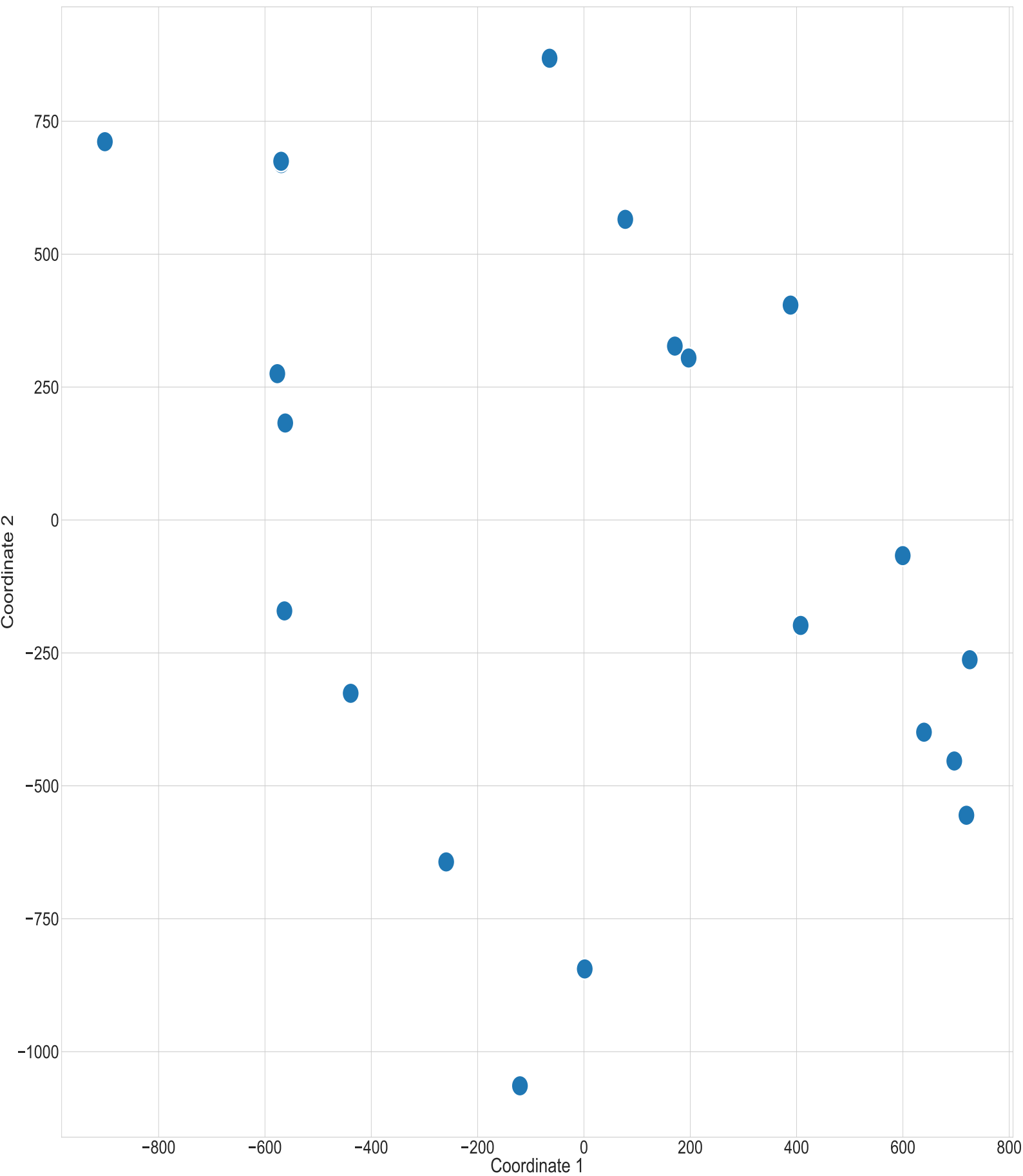
Scatter plot of PKS-1 *Micromonosporaceae* BGCs analyzed via multidimensional scaling using I2 pairwise distance. Only PKS-1 BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between PKS-1 BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S34 :



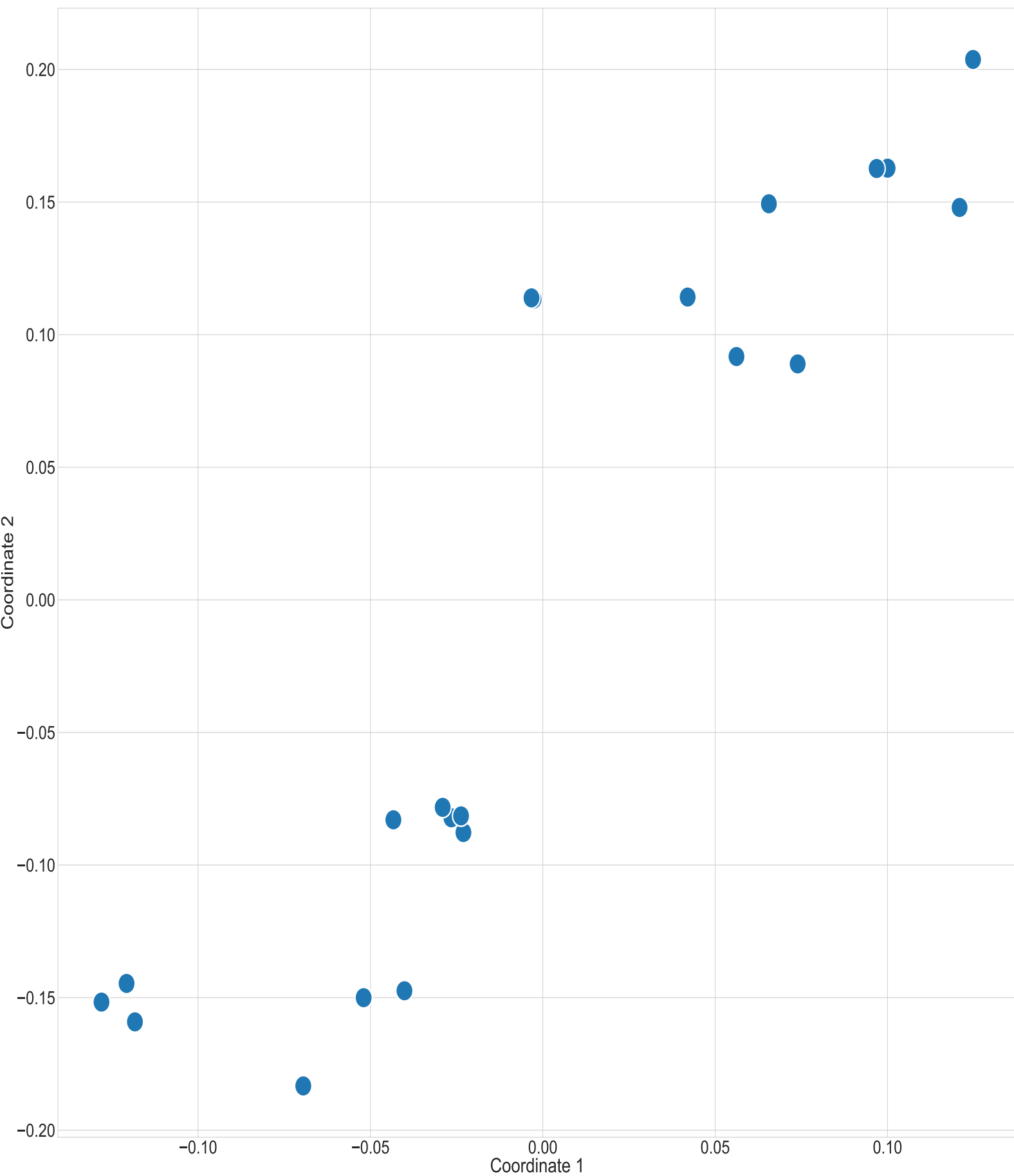
Scatter plot of PKS-1 *Micromonosporaceae* BGCs analyzed via multidimensional scaling using braycurtis pairwise distance. Only PKS-1 BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between PKS-1 BGCs is associated with the GCFs in BiG_SLiCE they were most similar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

Figure – S35 :



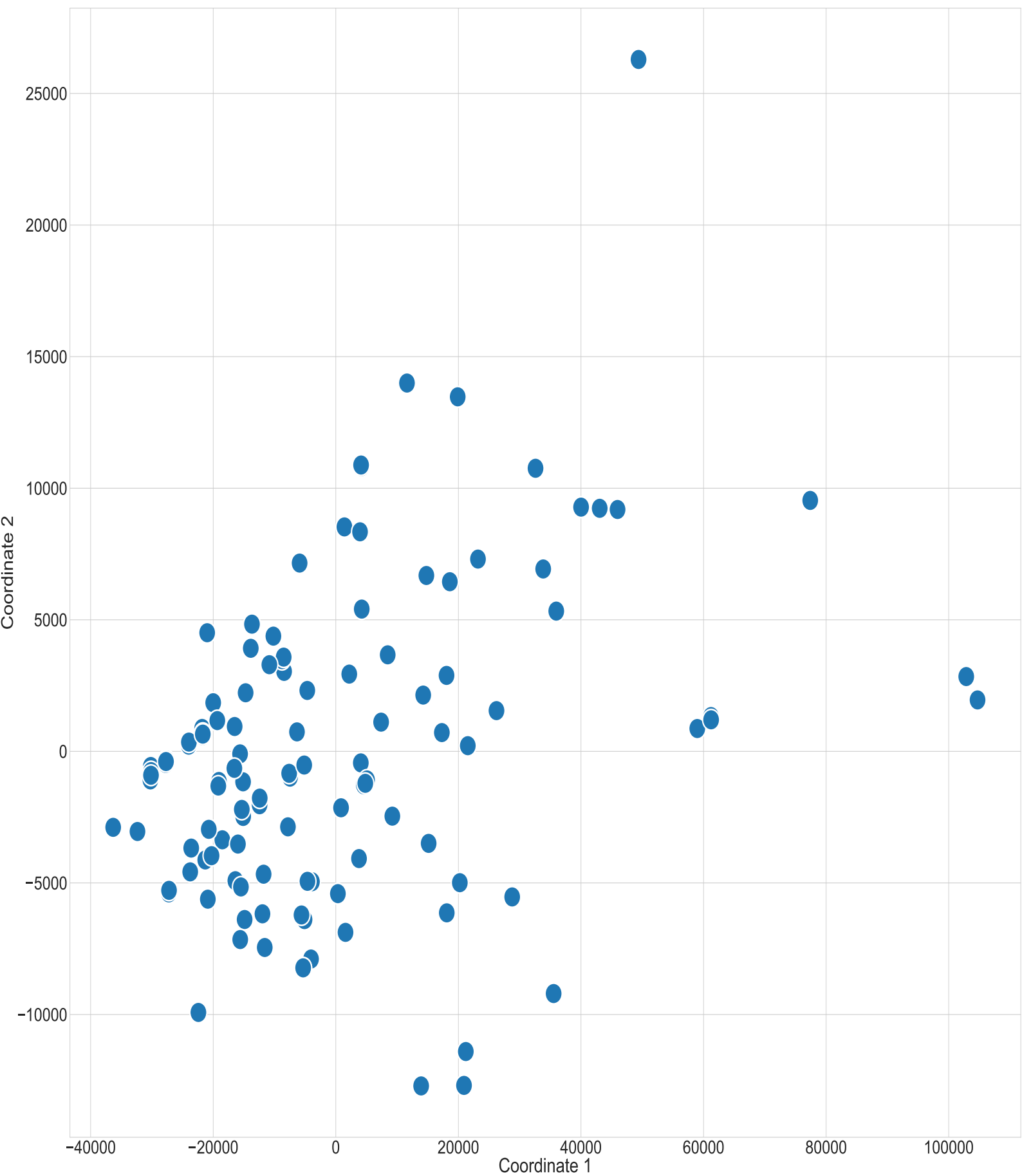
Scatter plot of PKS-1 *Micromonosporaceae* BGCs analyzed via multidimensional scaling using canberra pairwise distance. Only PKS-1 BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between PKS-1 BGCs is associated with the GCFs in BiG_SLiCE they were most similar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

Figure – S36 :



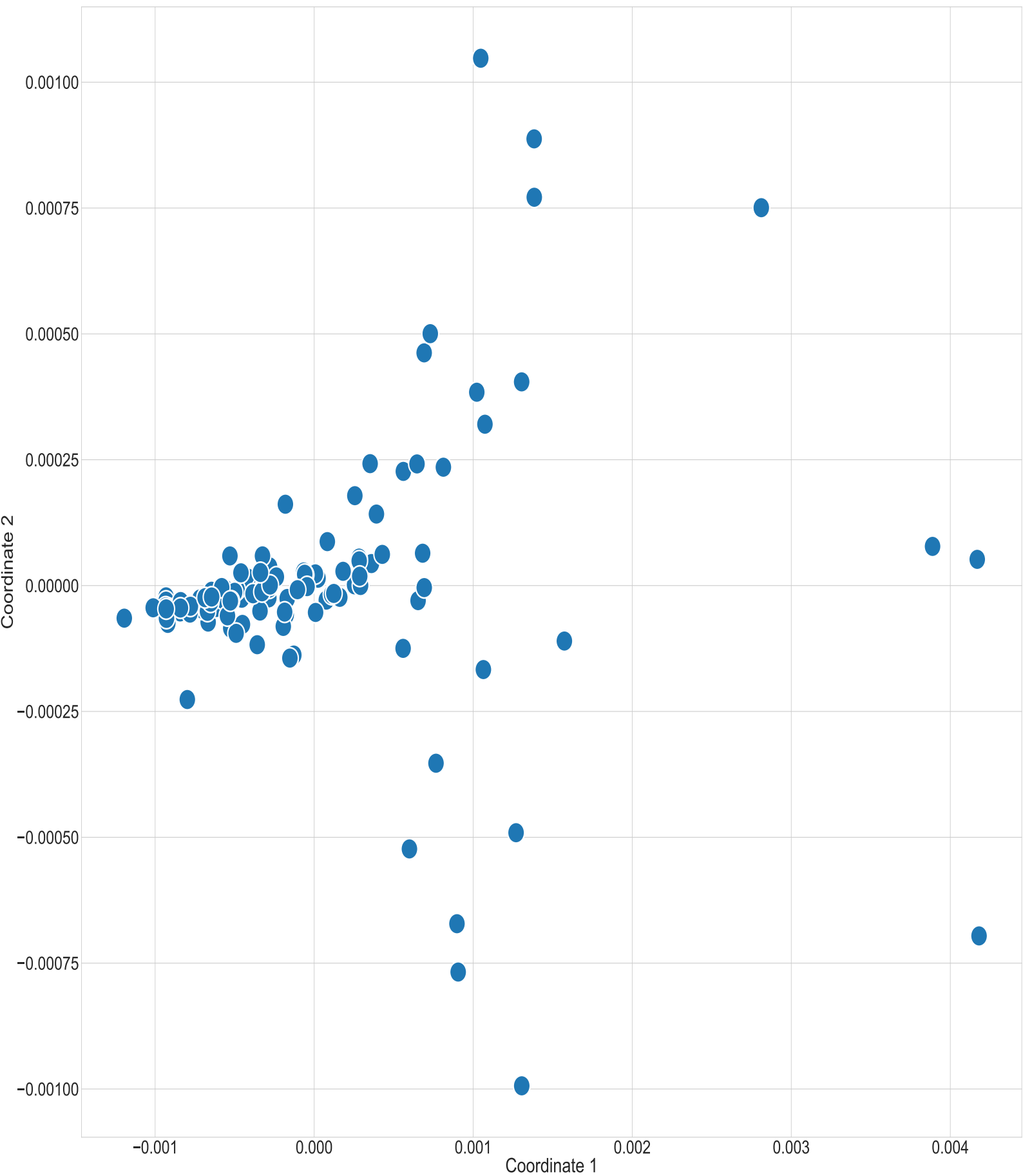
Scatter plot of PKS-1 *Micromonosporaceae* BGCs analyzed via multidimensional scaling using correlation pairwise distance. Only PKS-1 BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between PKS-1 BGCs is associated with the GCFs in BiG_SLiCE they were most similar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

Figure – S37 :



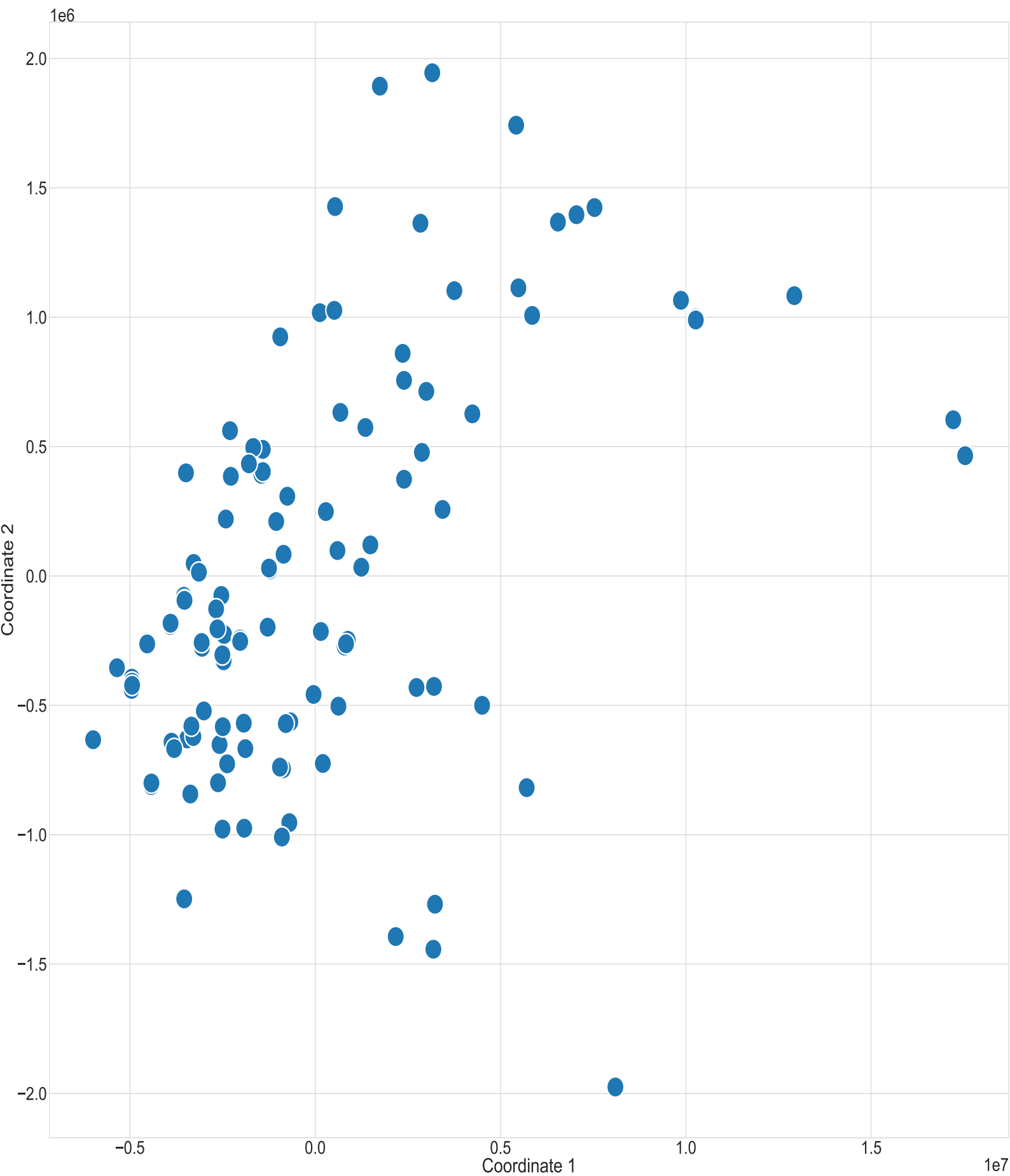
Scatter plot of RIPP *Micromonosporaceae* BGCs analyzed via multidimensional scaling using euclidean pairwise distance. Only RIPP BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between RIPP BGCs is associated with the GCFs in BiG_SLiCE they were most similar to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

Figure – S38 :



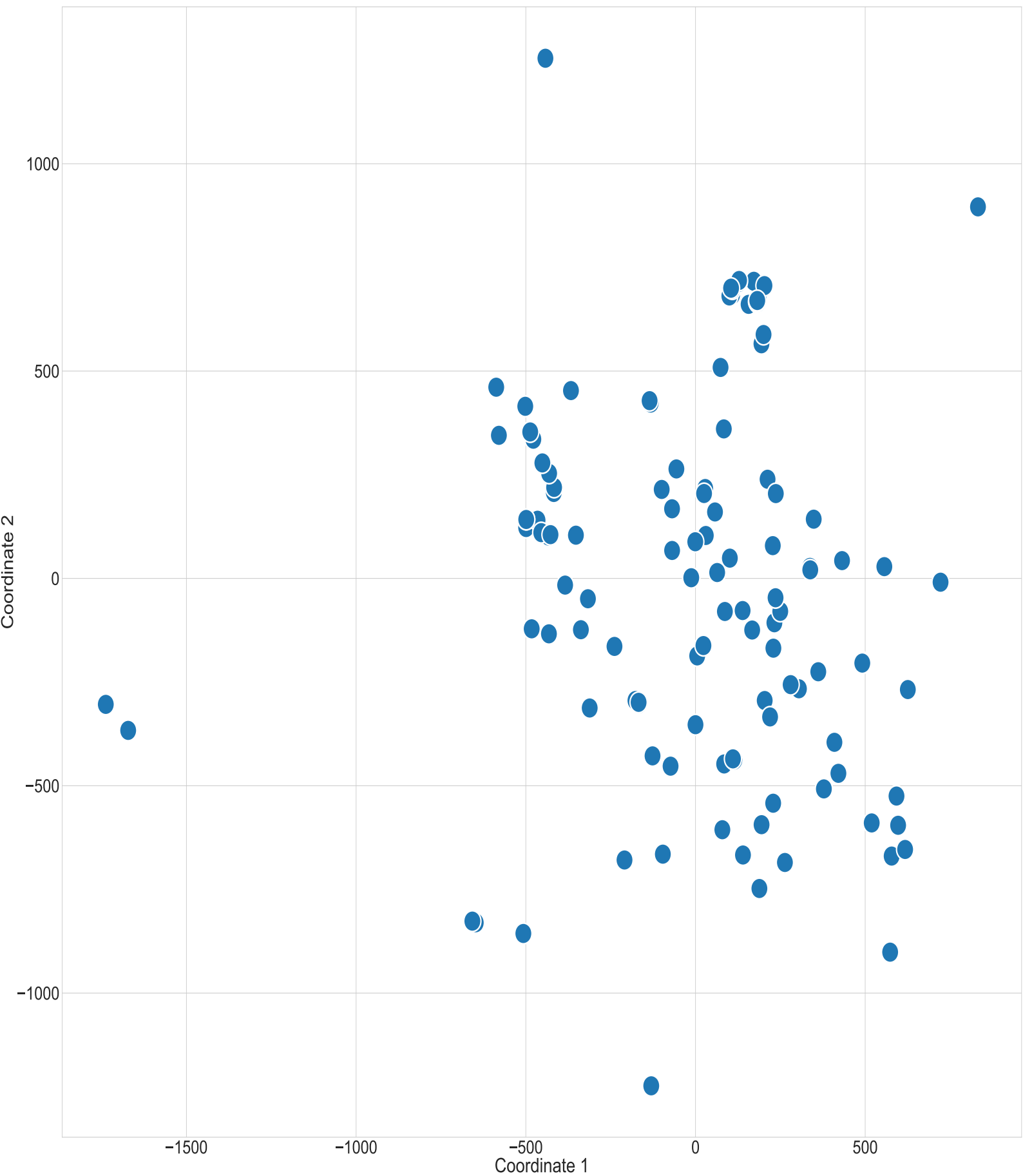
Scatter plot of RIPP *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cosine pairwise distance. Only RIPP BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between RIPP BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S39 :



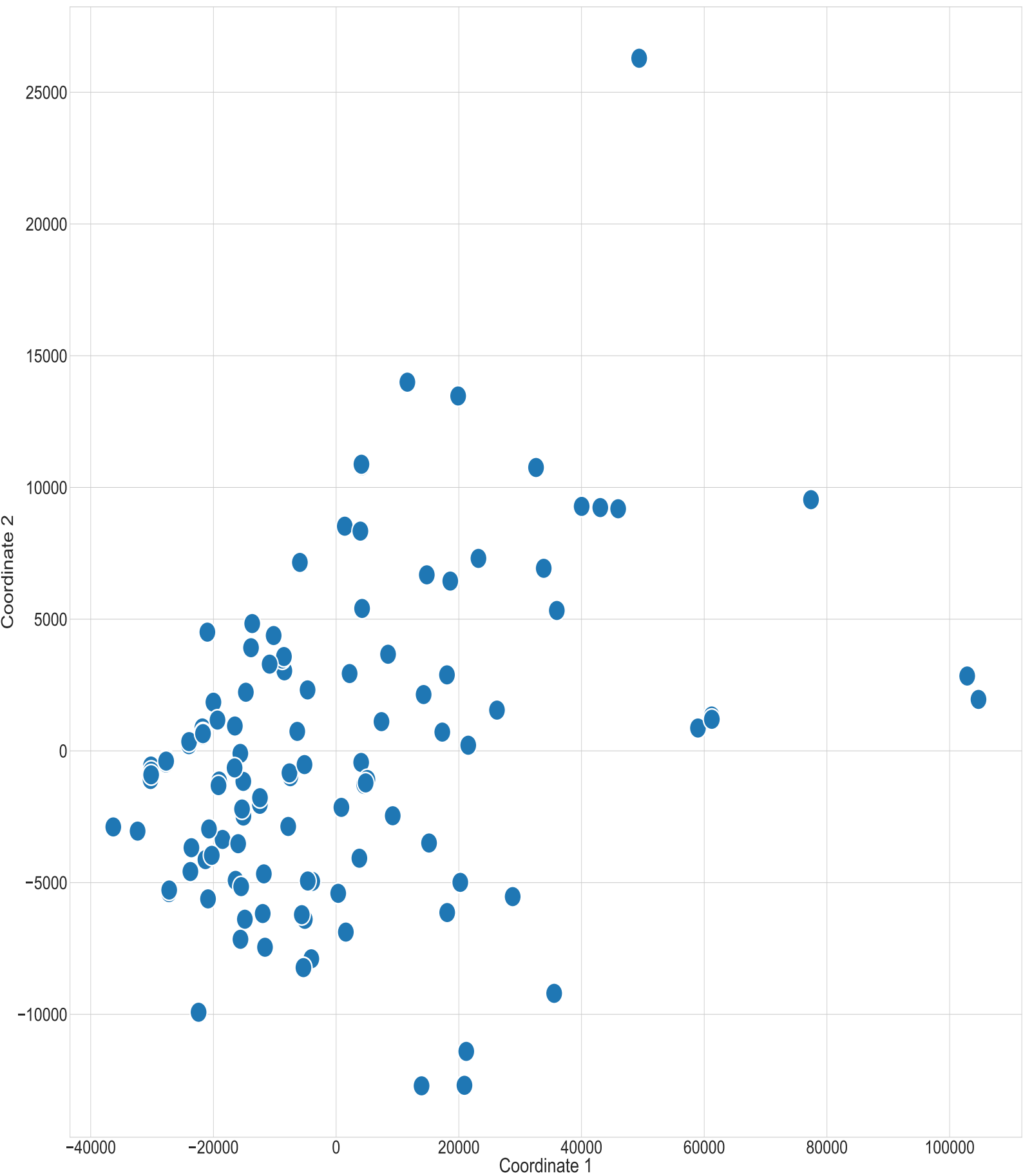
Scatter plot of RIPP *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cityblock pairwise distance. Only RIPP BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between RIPP BGCs is associated with the GCFs in BiG_SLICE they were most similar to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

Figure – S40 :



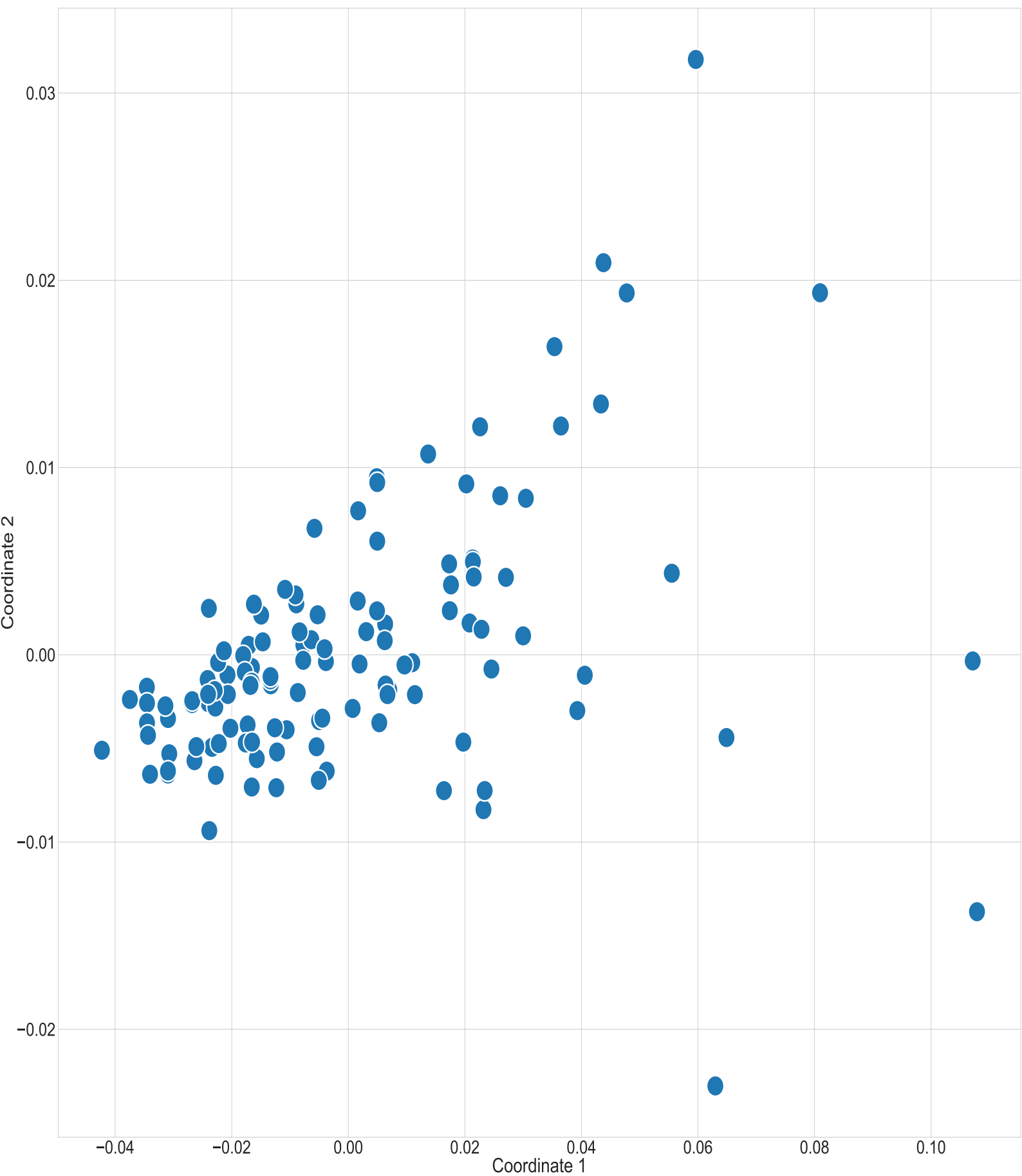
Scatter plot of RIPP *Micromonosporaceae* BGCs analyzed via multidimensional scaling using chebyshev pairwise distance. Only RIPP BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between RIPP BGCs is associated with the GCFs in BiG_SLiCE they were most similar to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

Figure – S41 :



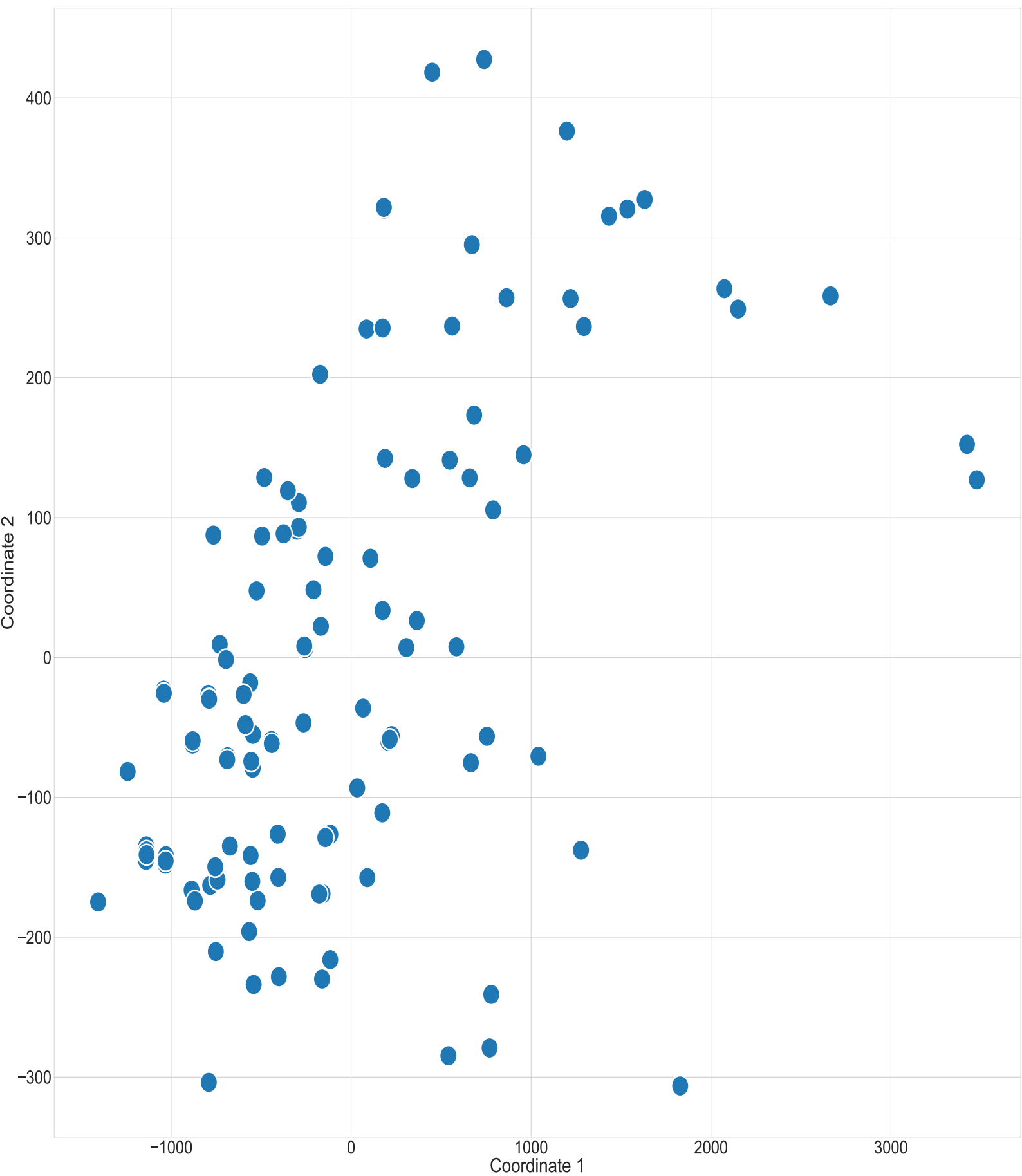
Scatter plot of RIPP *Micromonosporaceae* BGCs analyzed via multidimensional scaling using l2 pairwise distance. Only RIPP BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between RIPP BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S42 :



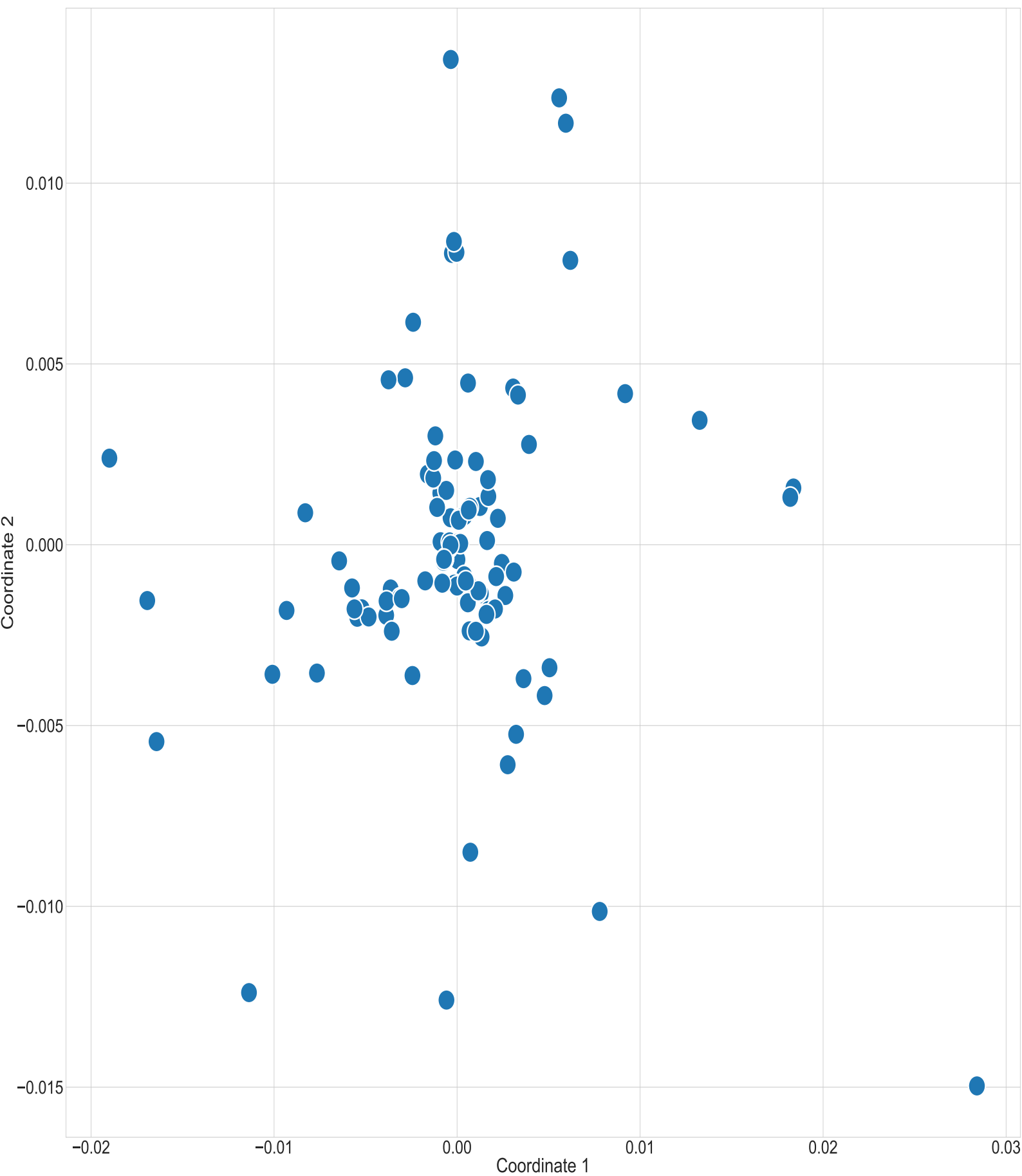
Scatter plot of RIPP *Micromonosporaceae* BGCs analyzed via multidimensional scaling using braycurtis pairwise distance. Only RIPP BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between RIPP BGCs is associated with the GCFs in BiG_SLiCE they were most similar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

Figure – S43 :



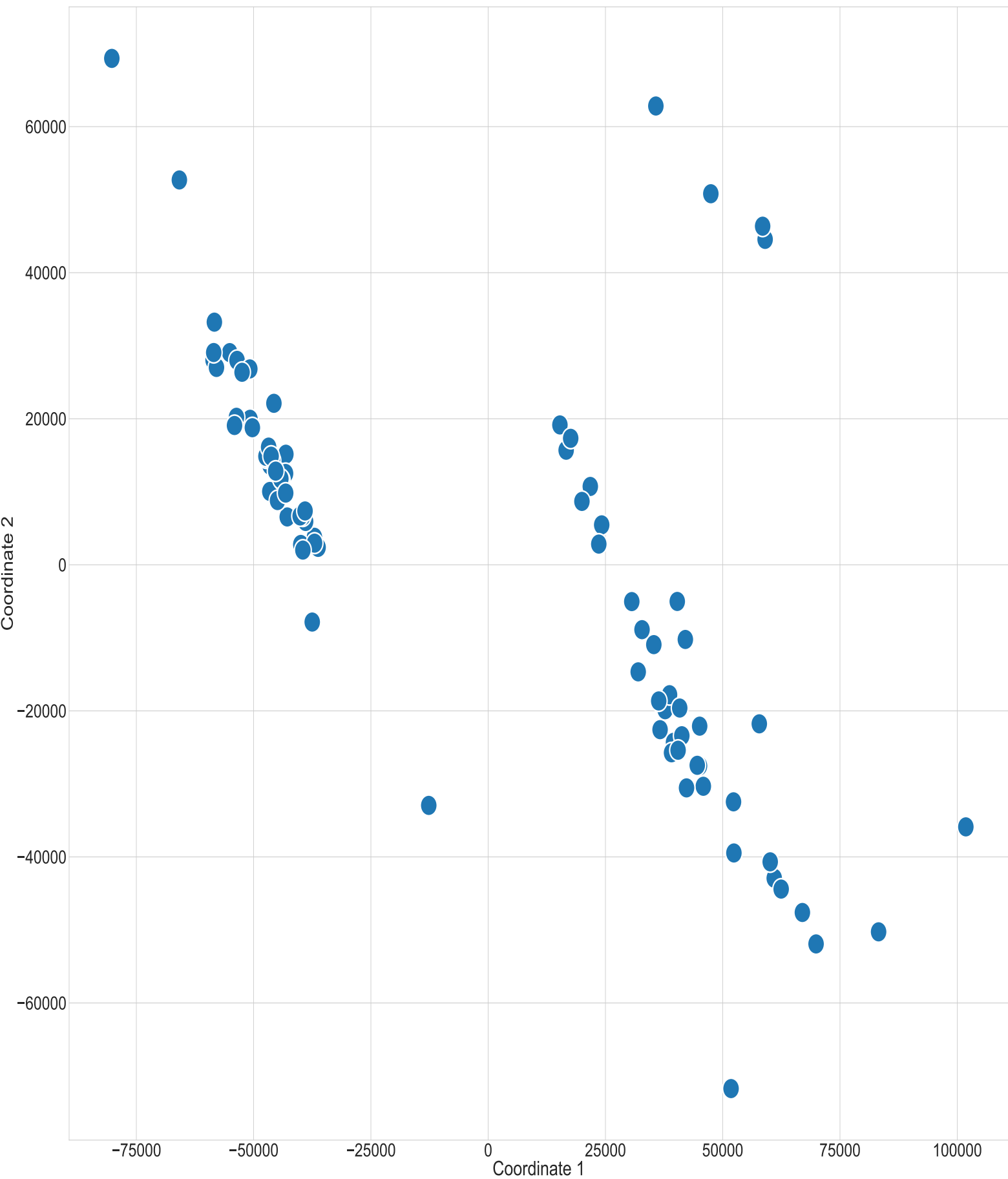
Scatter plot of RIPP *Micromonosporaceae* BGCs analyzed via multidimensional scaling using canberra pairwise distance. Only RIPP BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between RIPP BGCs is associated with the GCFs in BiG_SLiCE they were most similar to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

Figure – S44 :



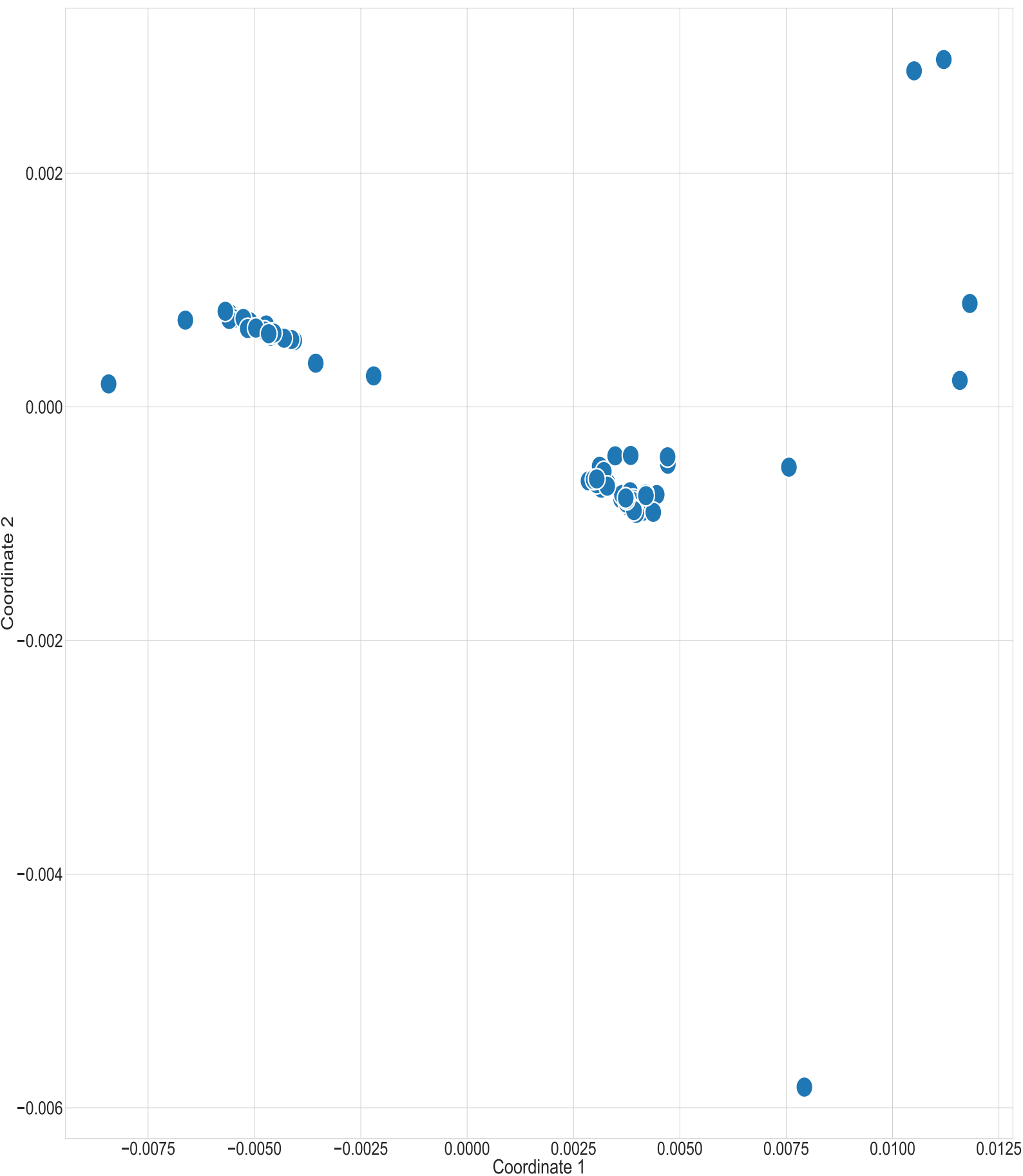
Scatter plot of RIPP *Micromonosporaceae* BGCs analyzed via multidimensional scaling using correlation pairwise distance. Only RIPP BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between RIPP BGCs is associated with the GCFs in BiG_SLiCE they were most similar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

Figure – S45 :



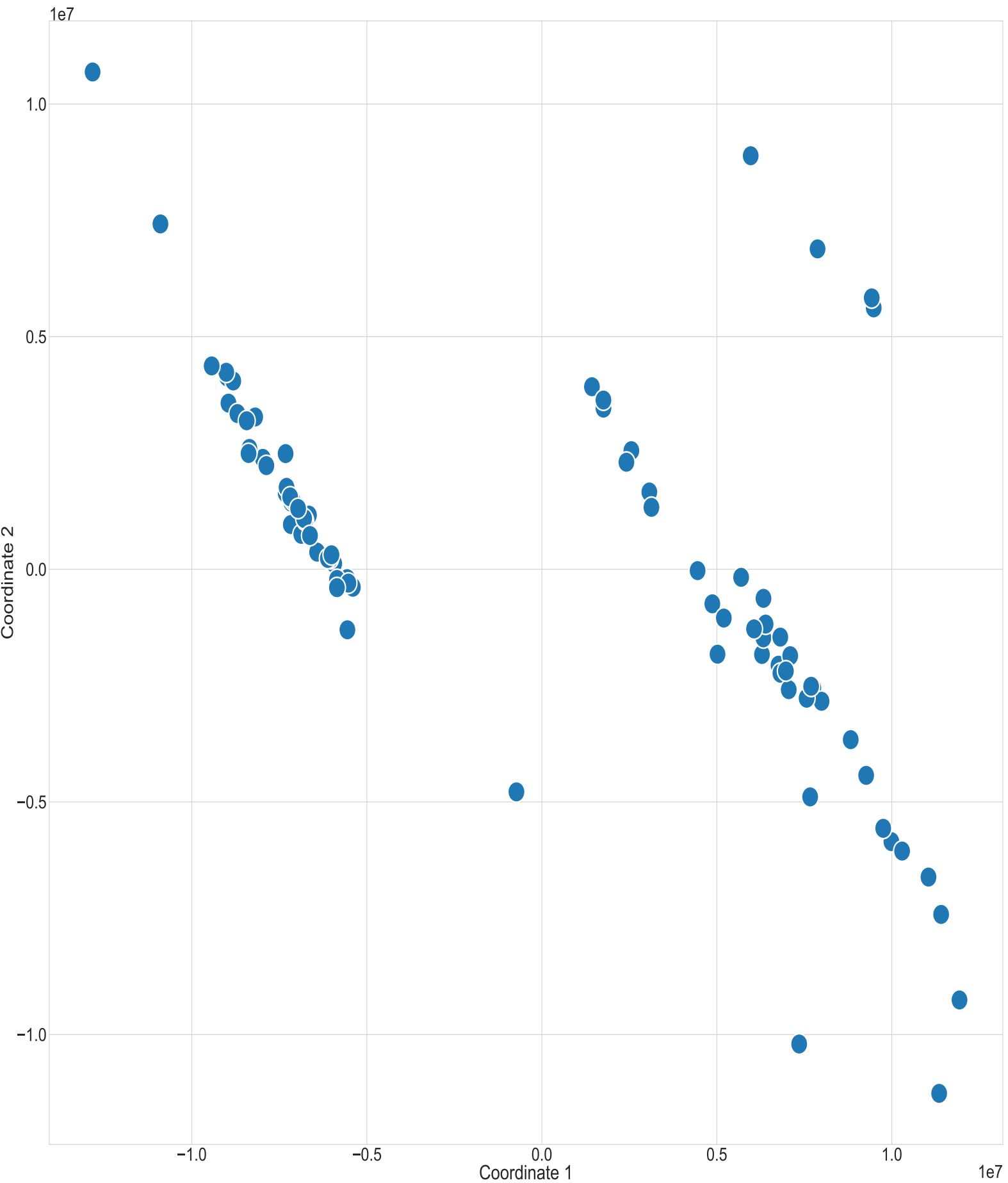
Scatter plot of PKS-Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using euclidean pairwise distance. Only PKS-Other BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between PKS-Other BGCs is associated with the GCFs in BiG_SLiCE they were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

Figure – S46 :



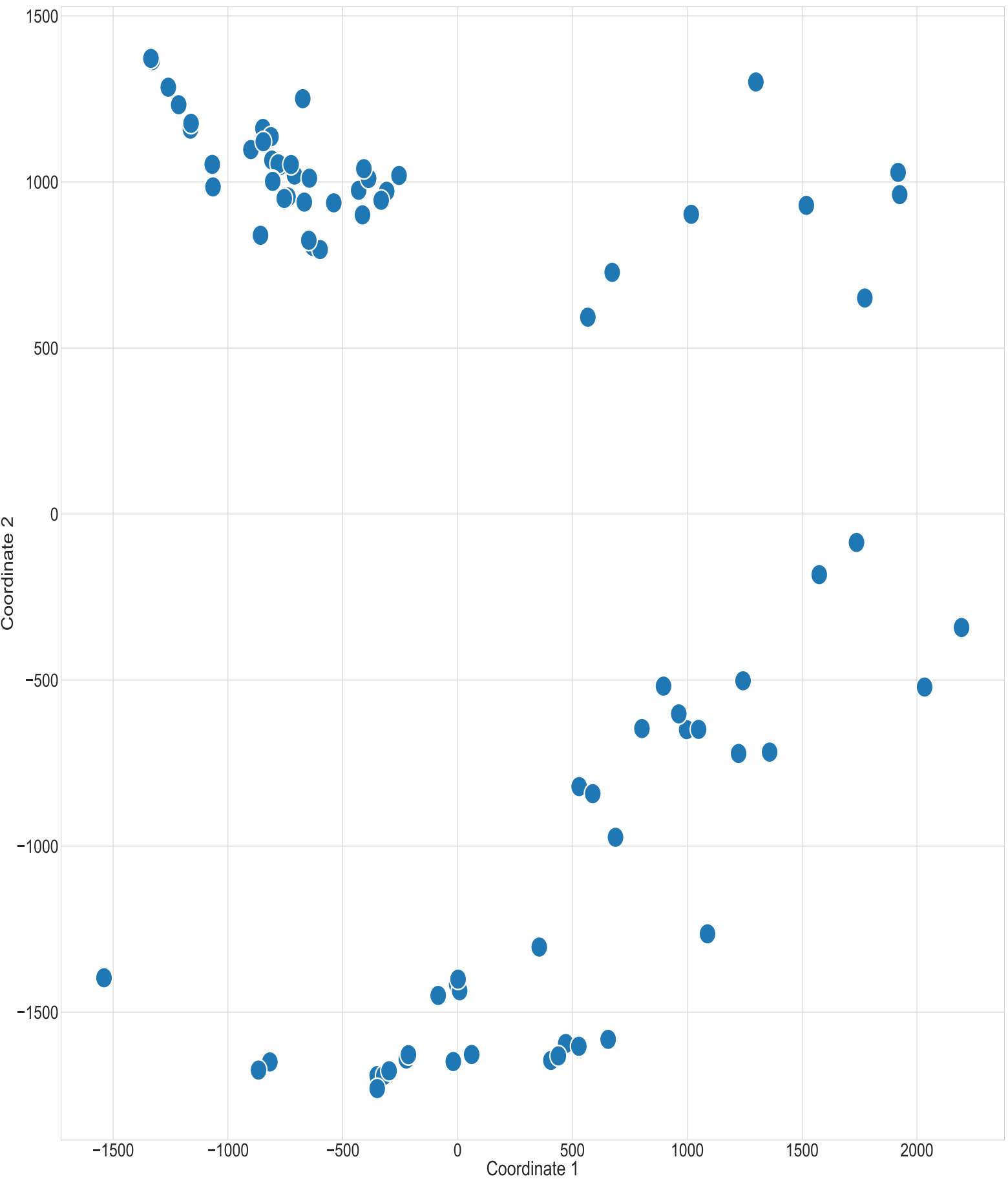
Scatter plot of PKS-Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cosine pairwise distance. Only PKS-Other BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between PKS-Other BGCs is associated with the GCFs in BiG_SLICE they were most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

Figure – S47 :



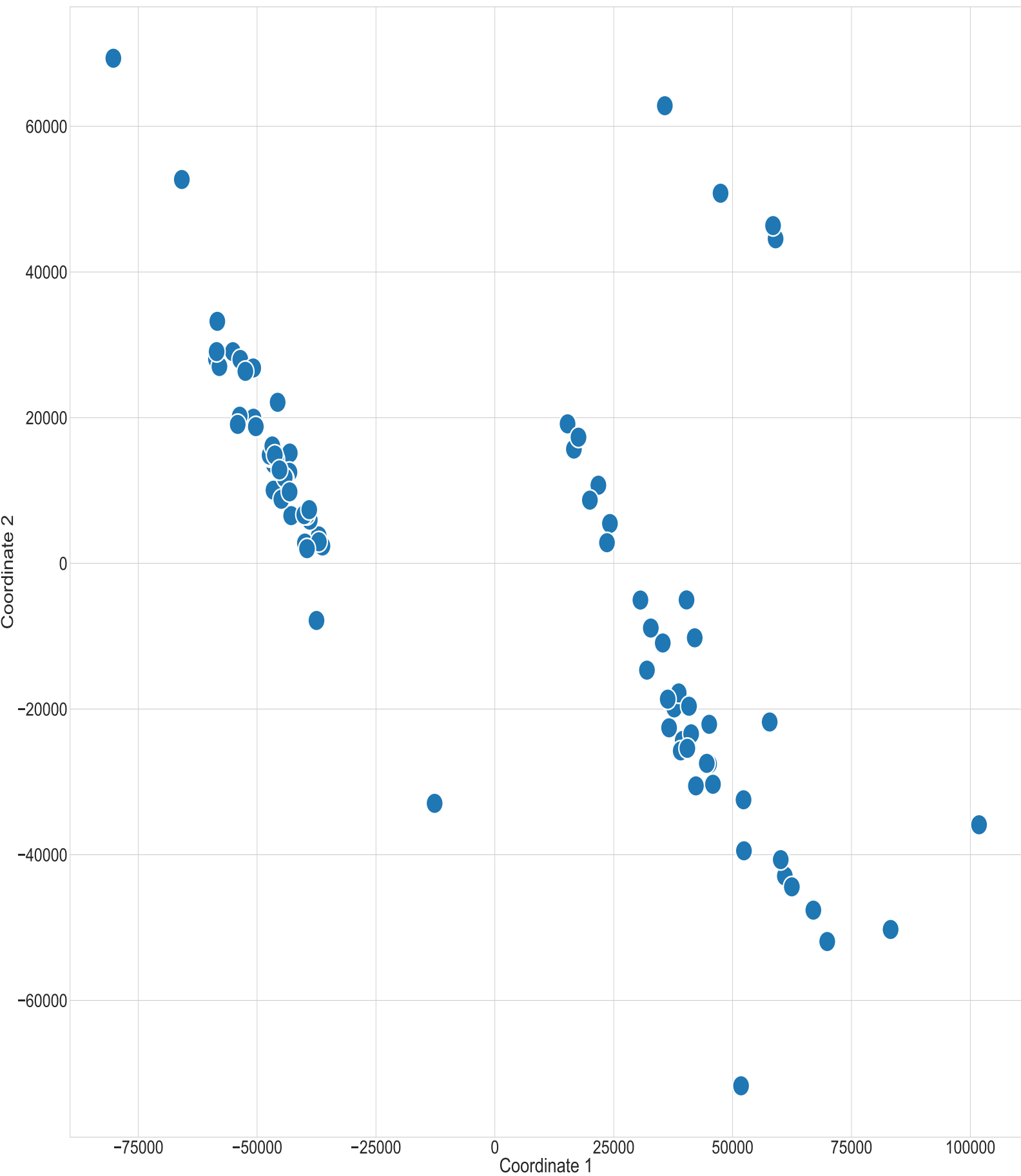
Scatter plot of PKS-Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cityblock pairwise distance. Only PKS-Other BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between PKS-Other BGCs is associated with the GCFs in BiG_SLICE they were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

Figure – S48 :



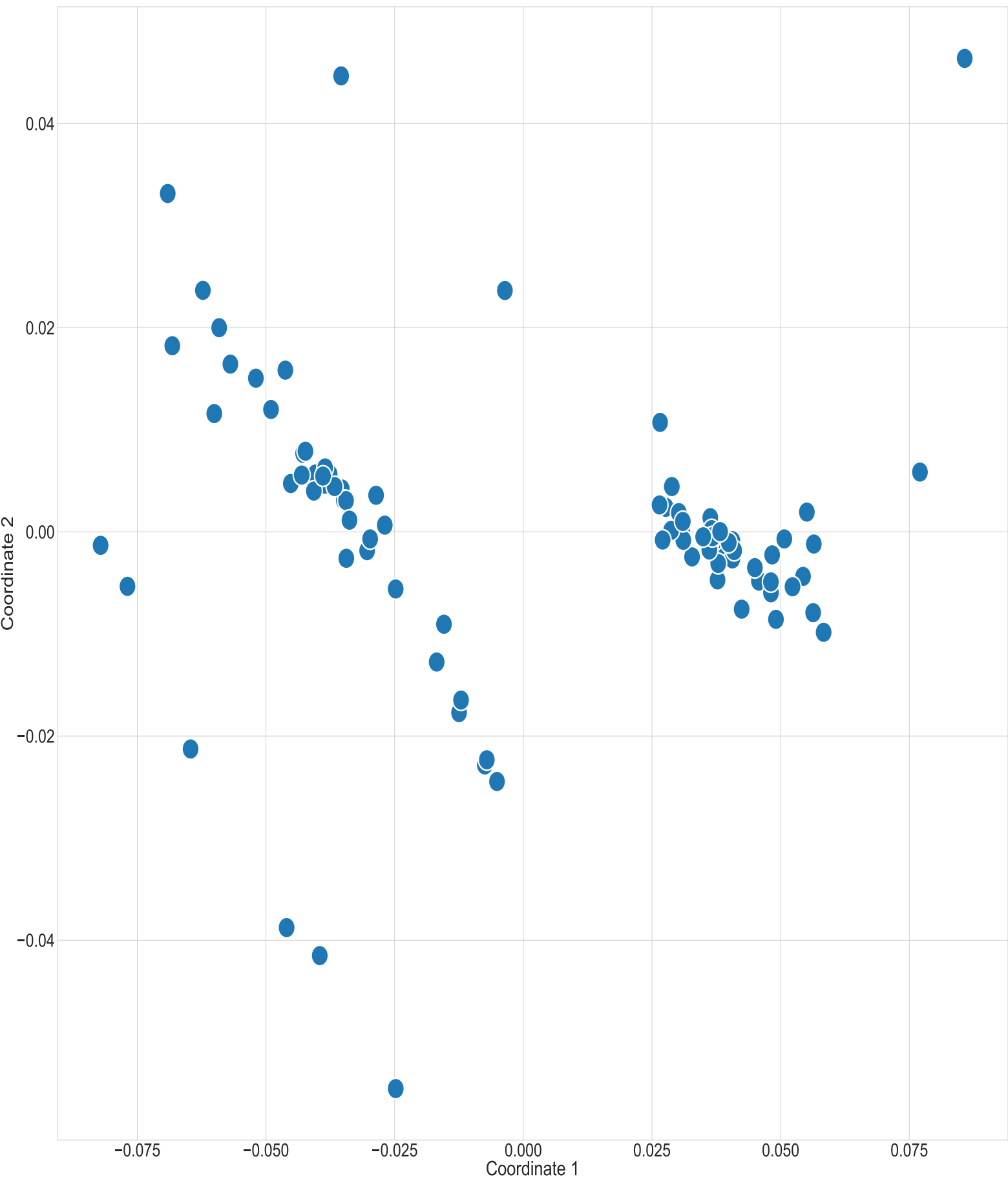
Scatter plot of PKS-Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using chebyshev pairwise distance. Only PKS-Other BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between PKS-Other BGCs is associated with the GCFs in BiG_SLiCE they were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

Figure – S49 :



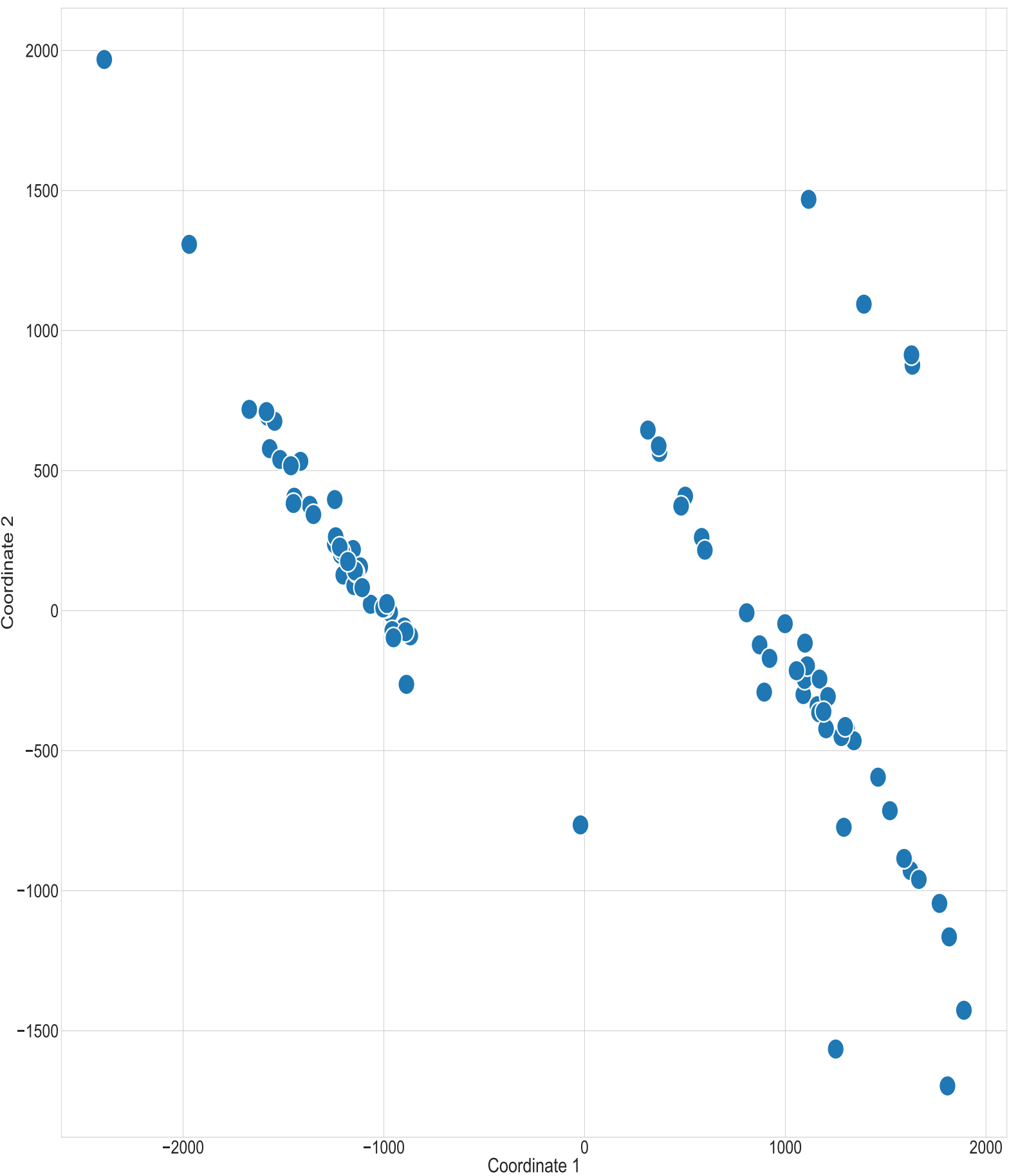
Scatter plot of PKS-Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using l2 pairwise distance. Only PKS-Other BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between PKS-Other BGCs is associated with the GCFs in BiG_SLiCE they were most similar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

Figure – S50 :



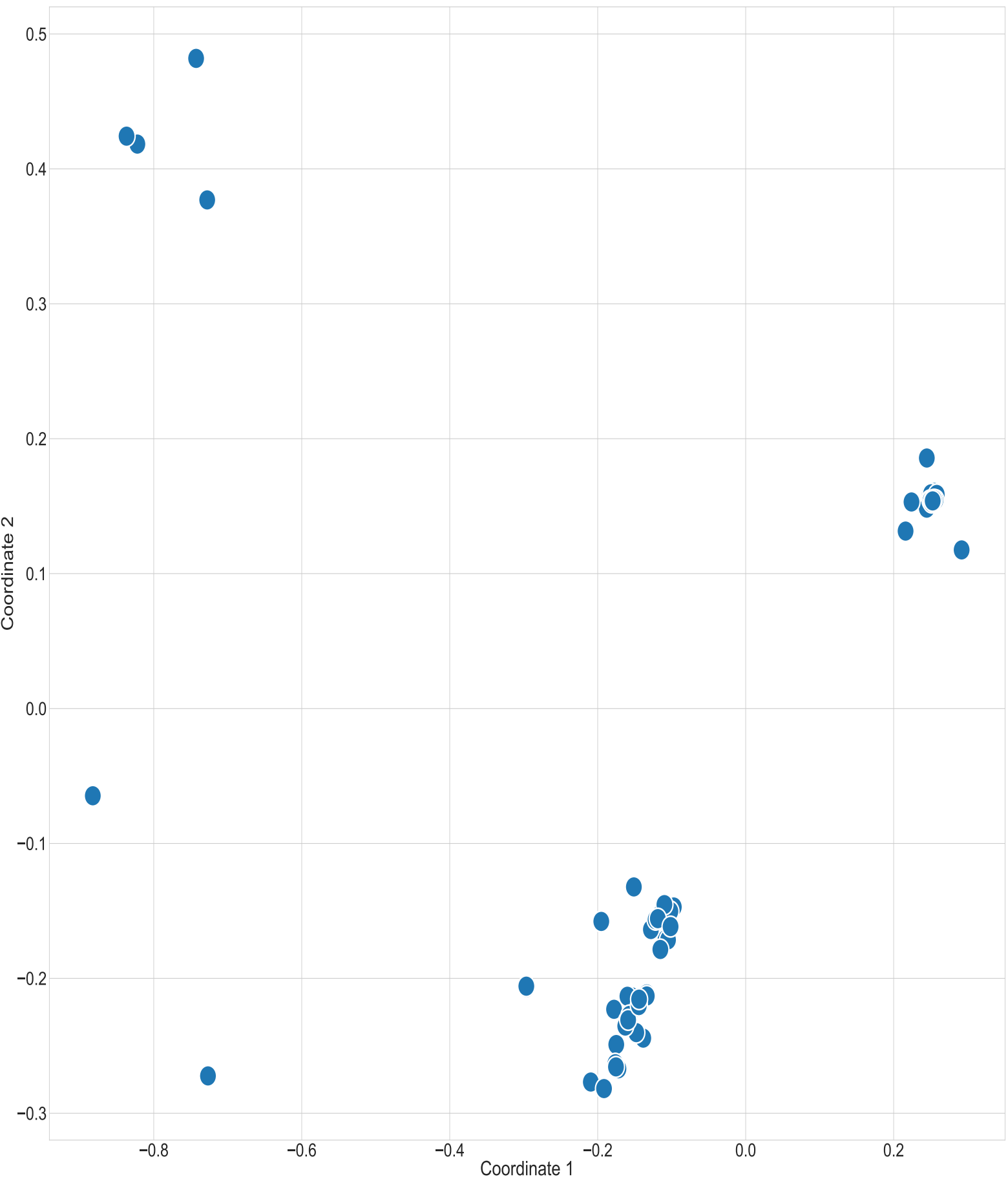
Scatter plot of PKS-Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using braycurtis pairwise distance. Only PKS-Other BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between PKS-Other BGCs is associated with the GCFs in BiG_SLICE they were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

Figure – S51 :



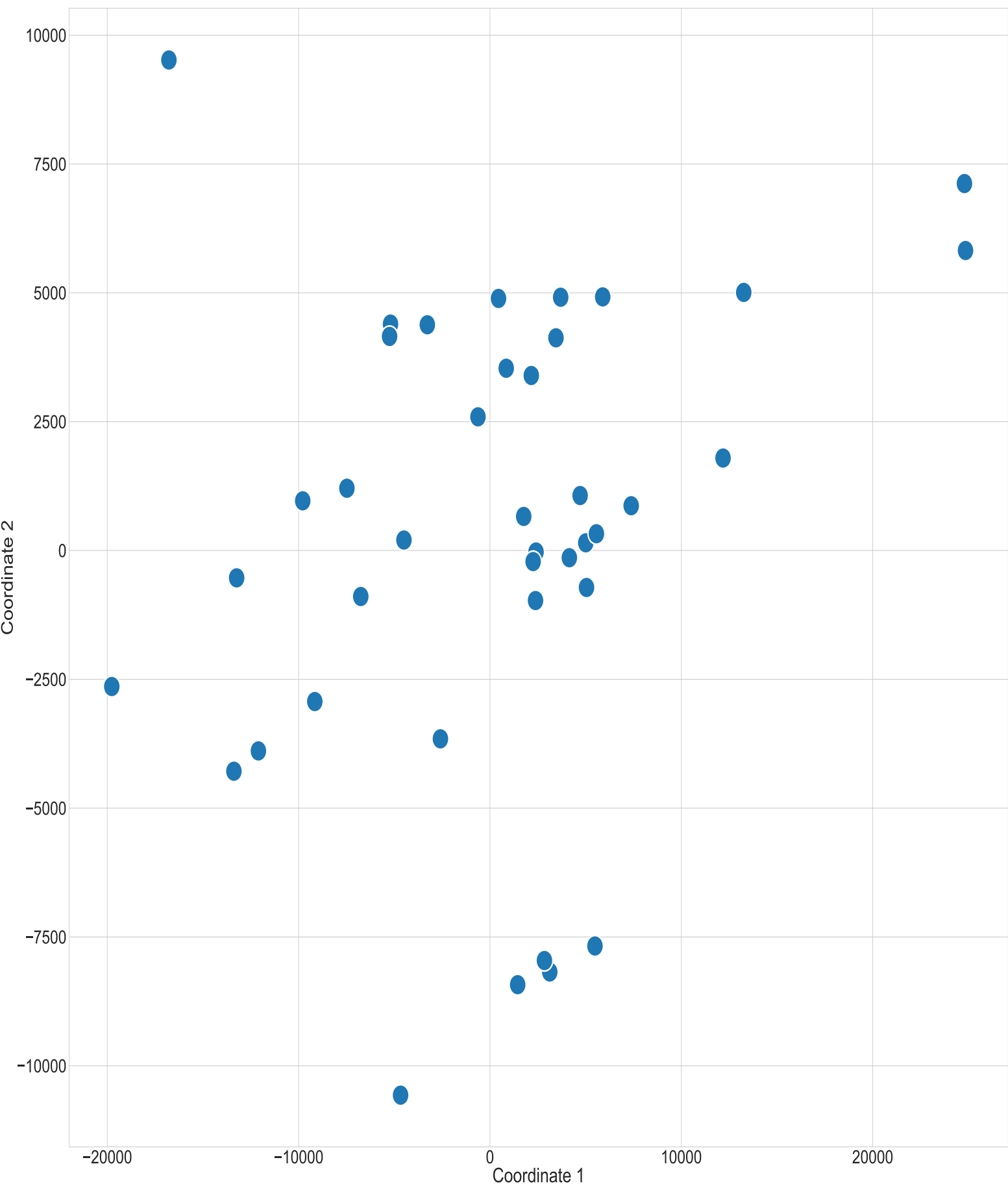
Scatter plot of PKS-Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using Canberra pairwise distance. Only PKS-Other BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between PKS-Other BGCs is associated with the GCFs in BiG_SLiCE they were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

Figure – S52 :



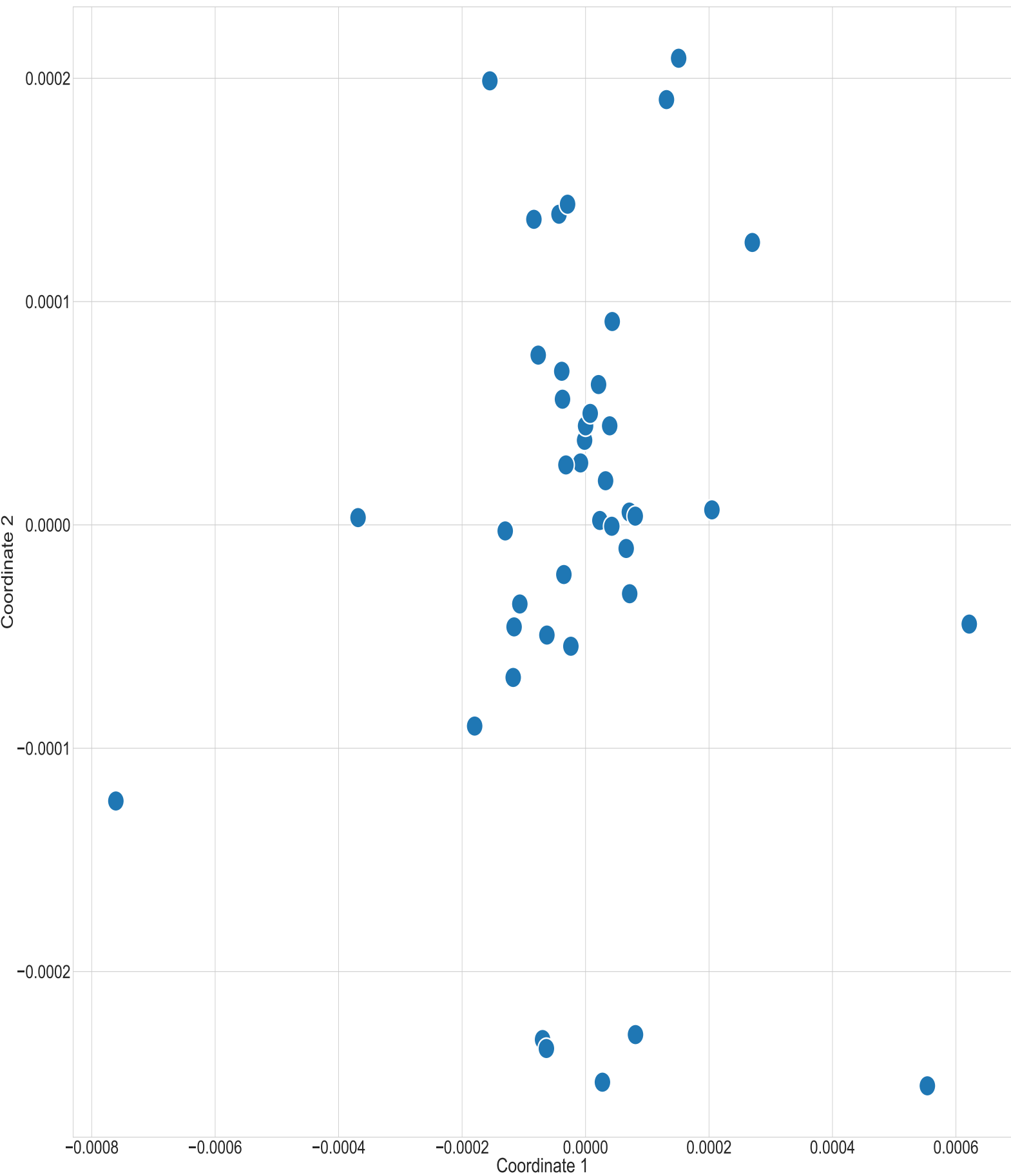
Scatter plot of PKS-Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using correlation pairwise distance. Only PKS-Other BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between PKS-Other BGCs is associated with the GCFs in BiG_SLiCE they were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

Figure – S53 :



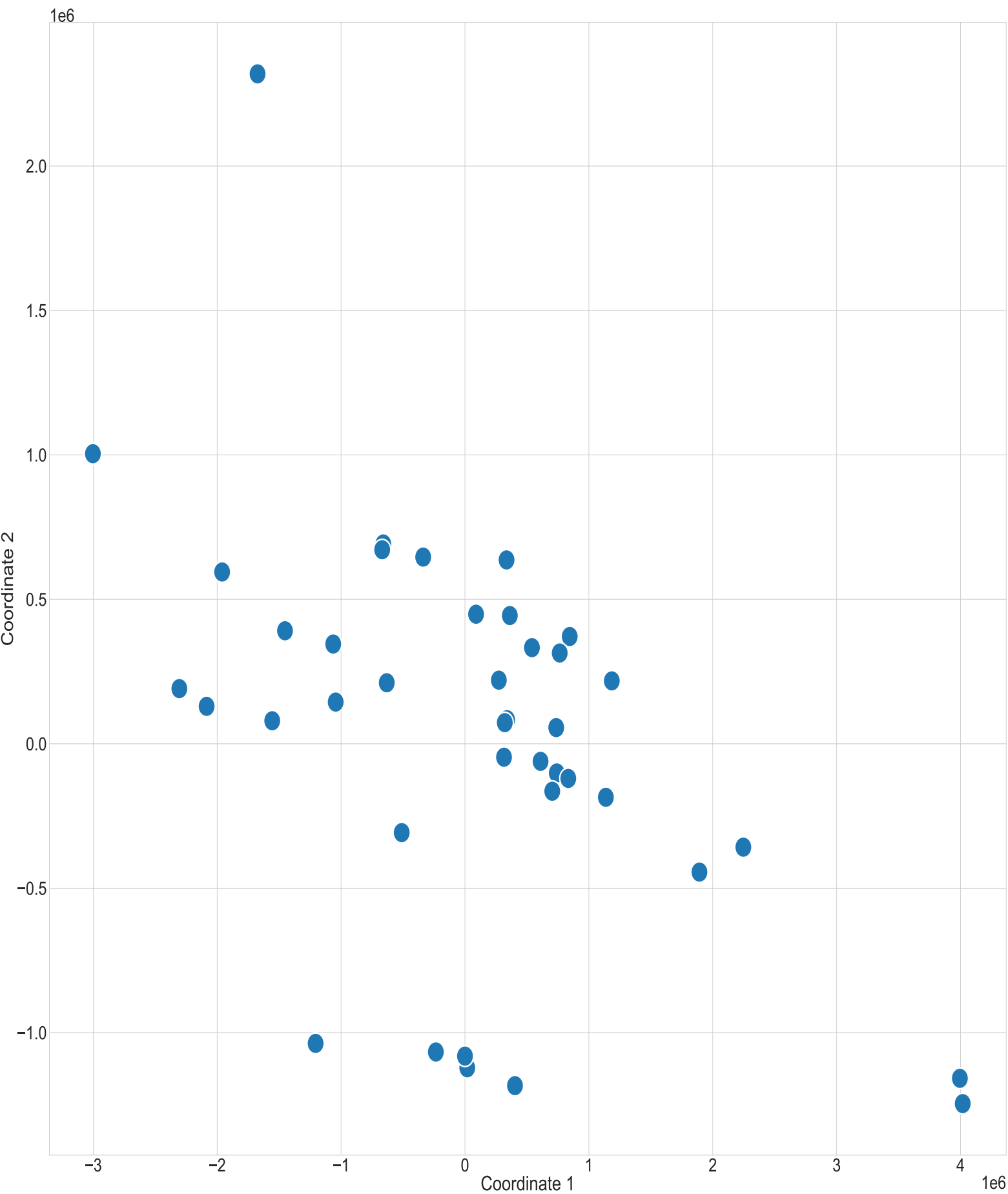
Scatter plot of Siderophore *Micromonosporaceae* BGCs analyzed via multidimensional scaling using euclidean pairwise distance. Only Siderophore BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Siderophore BGCs is associated with the GCFs in BiG_SLICE they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

Figure – S54 :



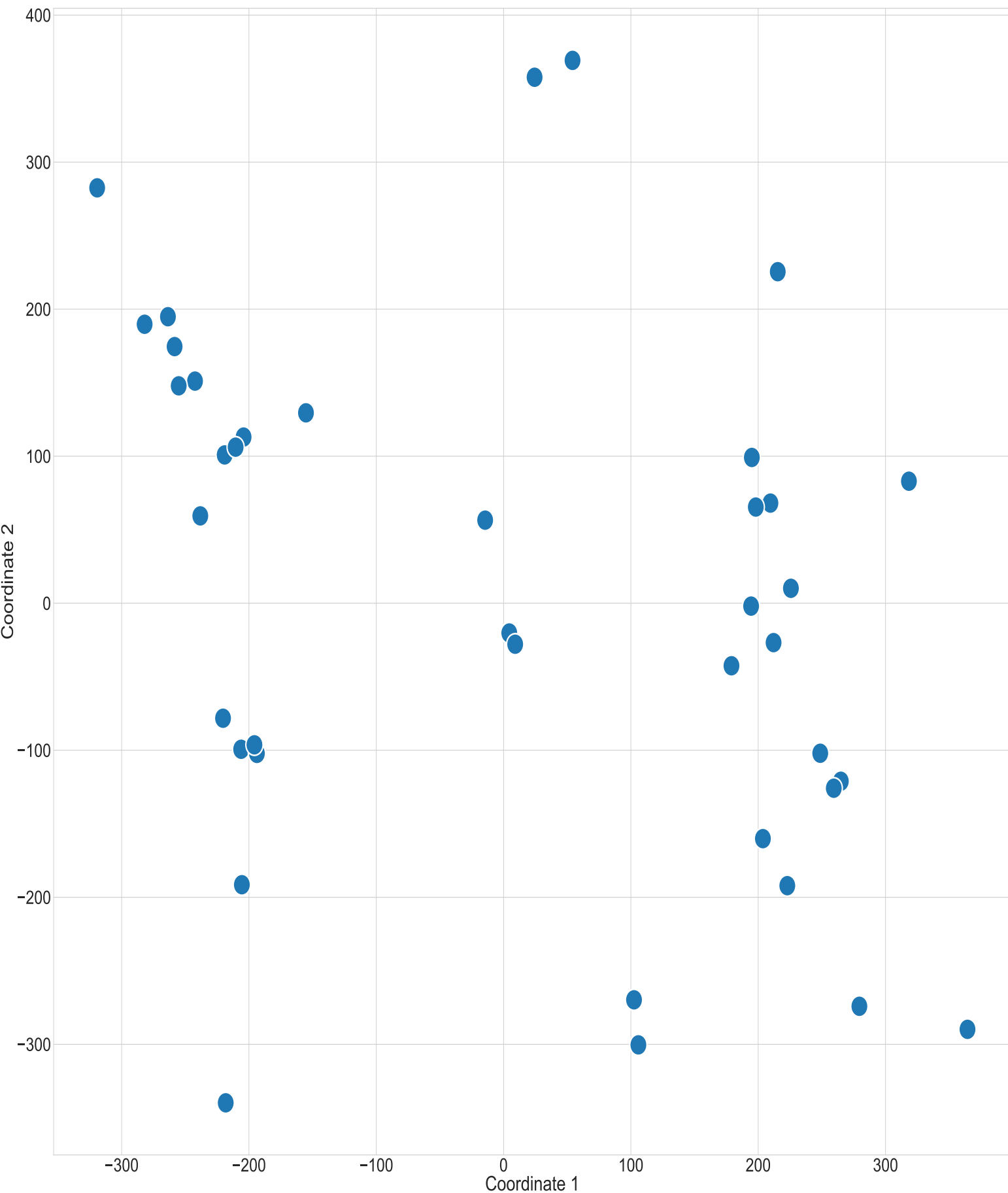
Scatter plot of Siderophore *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cosine pairwise distance. Only Siderophore BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Siderophore BGCs is associated with the GCFs in BiG_SLiCE they were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

Figure – S55 :



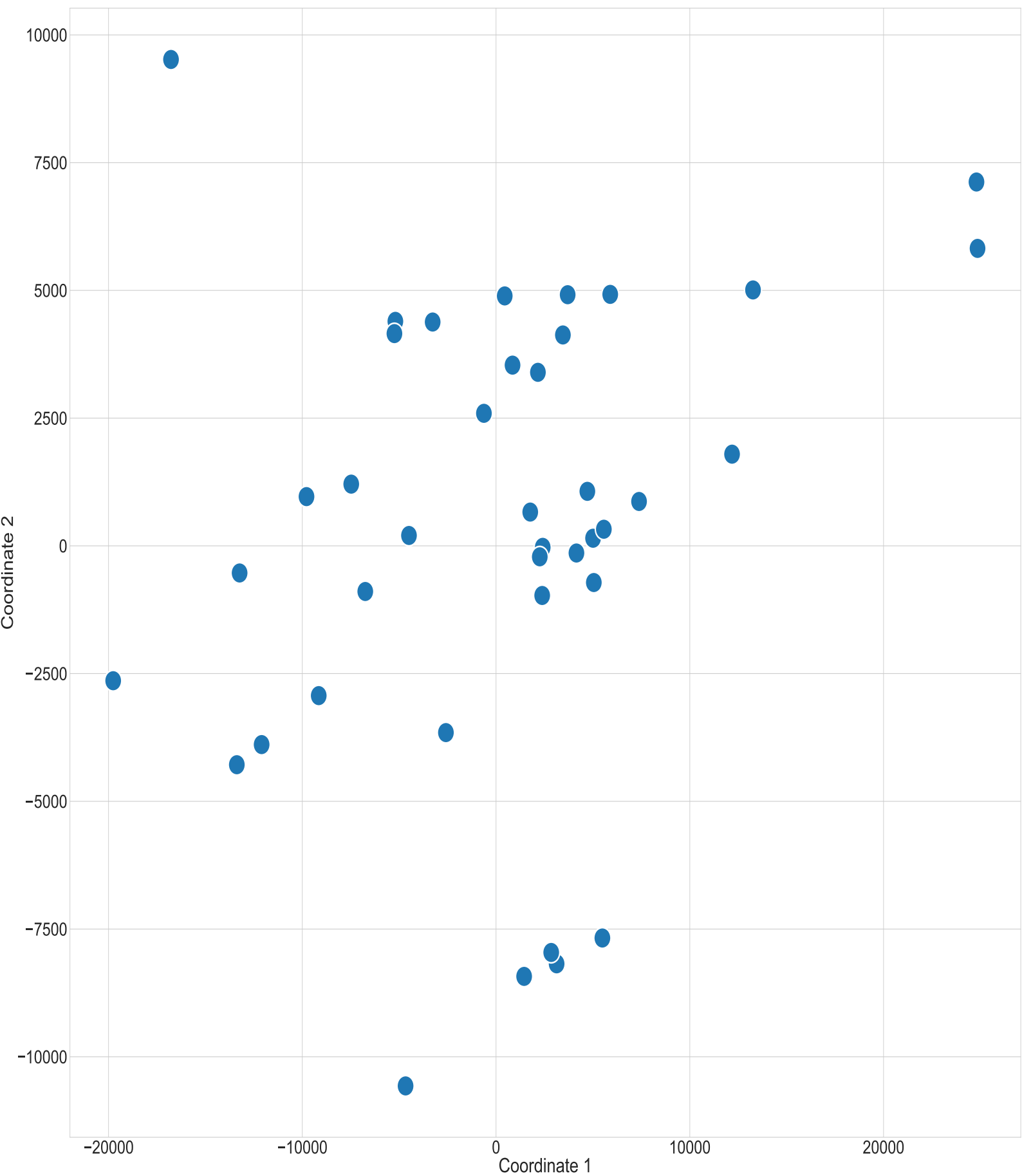
Scatter plot of Siderophore *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cityblock pairwise distance. Only Siderophore BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Siderophore BGCs is associated with the GCFs in BiG_SLiCE they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

Figure – S56 :



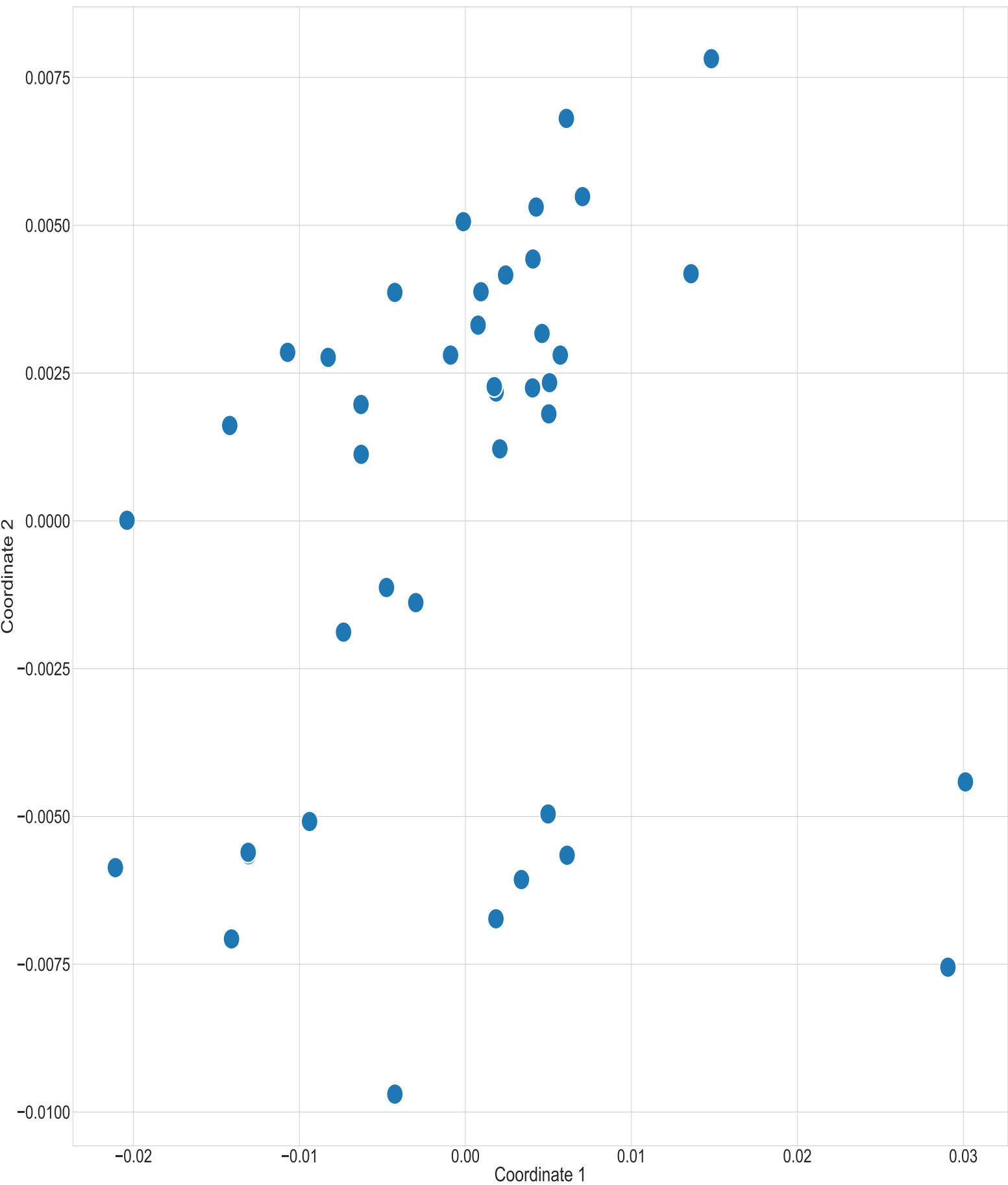
Scatter plot of Siderophore *Micromonosporaceae* BGCs analyzed via multidimensional scaling using chebyshev pairwise distance. Only Siderophore BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Siderophore BGCs is associated with the GCFs in BiG_SLiCE they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

Figure – S57 :



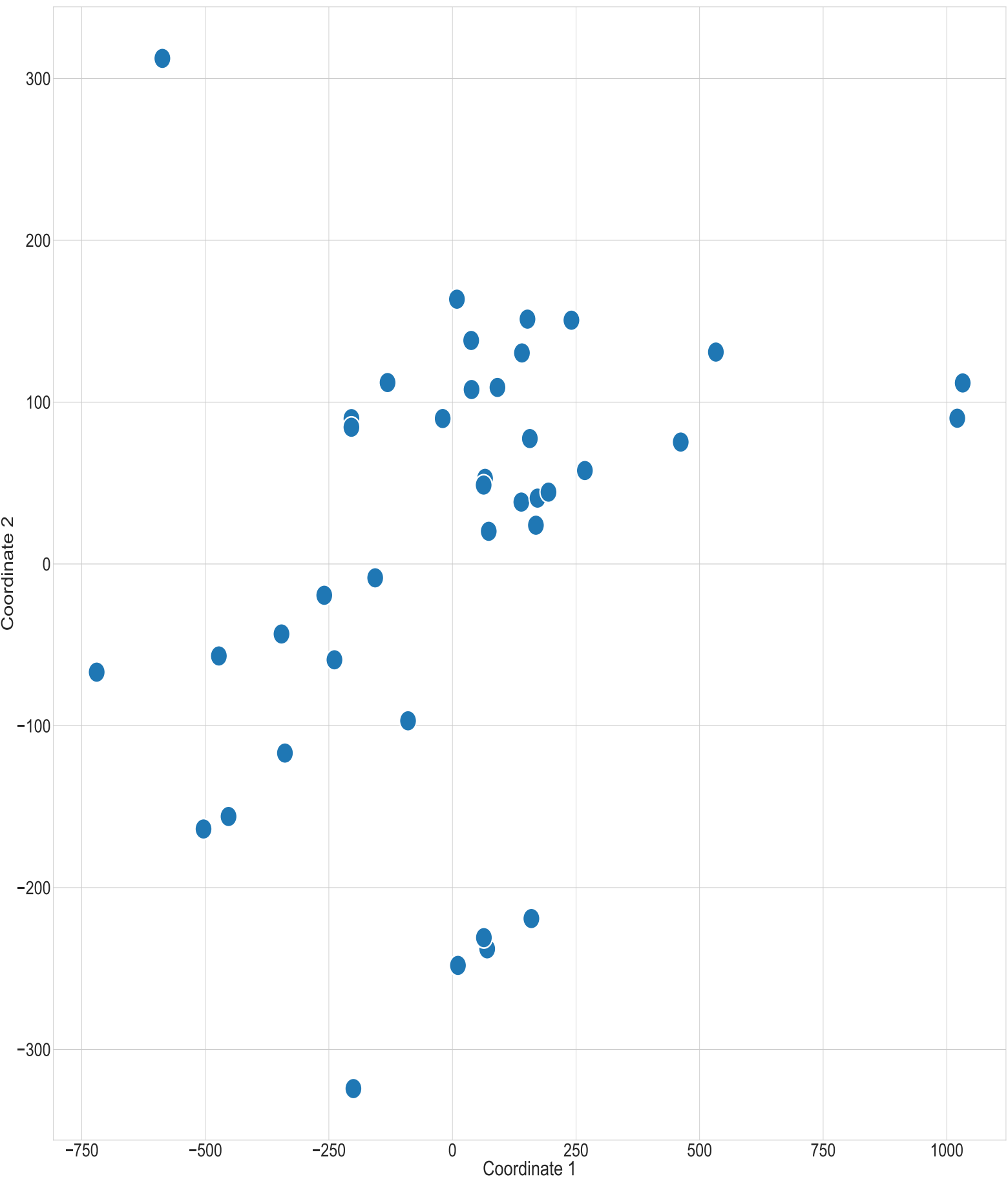
Scatter plot of Siderophore *Micromonosporaceae* BGCs analyzed via multidimensional scaling using I2 pairwise distance. Only Siderophore BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Siderophore BGCs is associated with the GCFs in BiG_SLICE they were most similar to (T=900).

Figure – S58 :



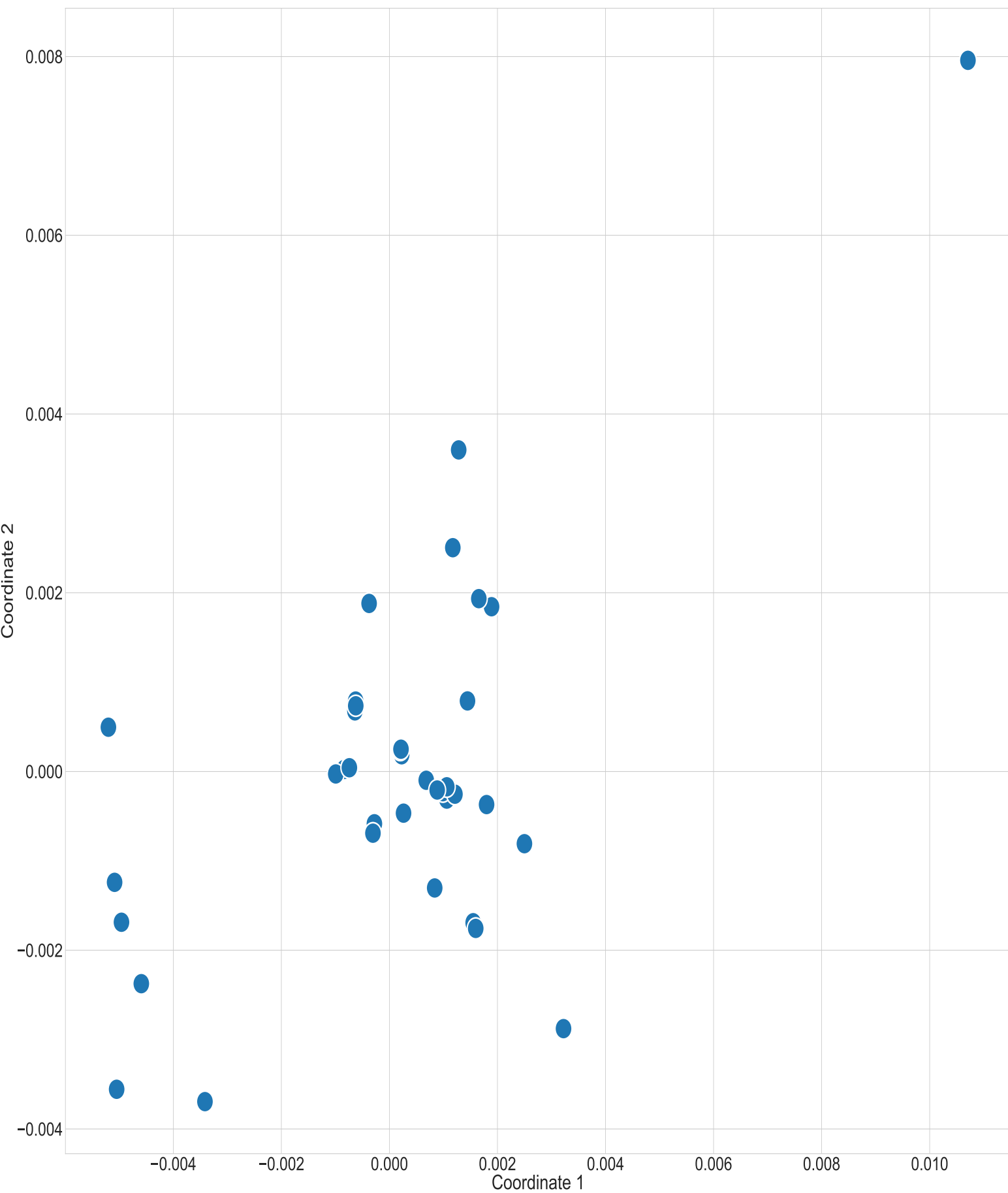
Scatter plot of Siderophore *Micromonosporaceae* BGCs analyzed via multidimensional scaling using braycurtis pairwise distance. Only Siderophore BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Siderophore BGCs is associated with the GCFs in BiG_SLiCE they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

Figure – S59 :



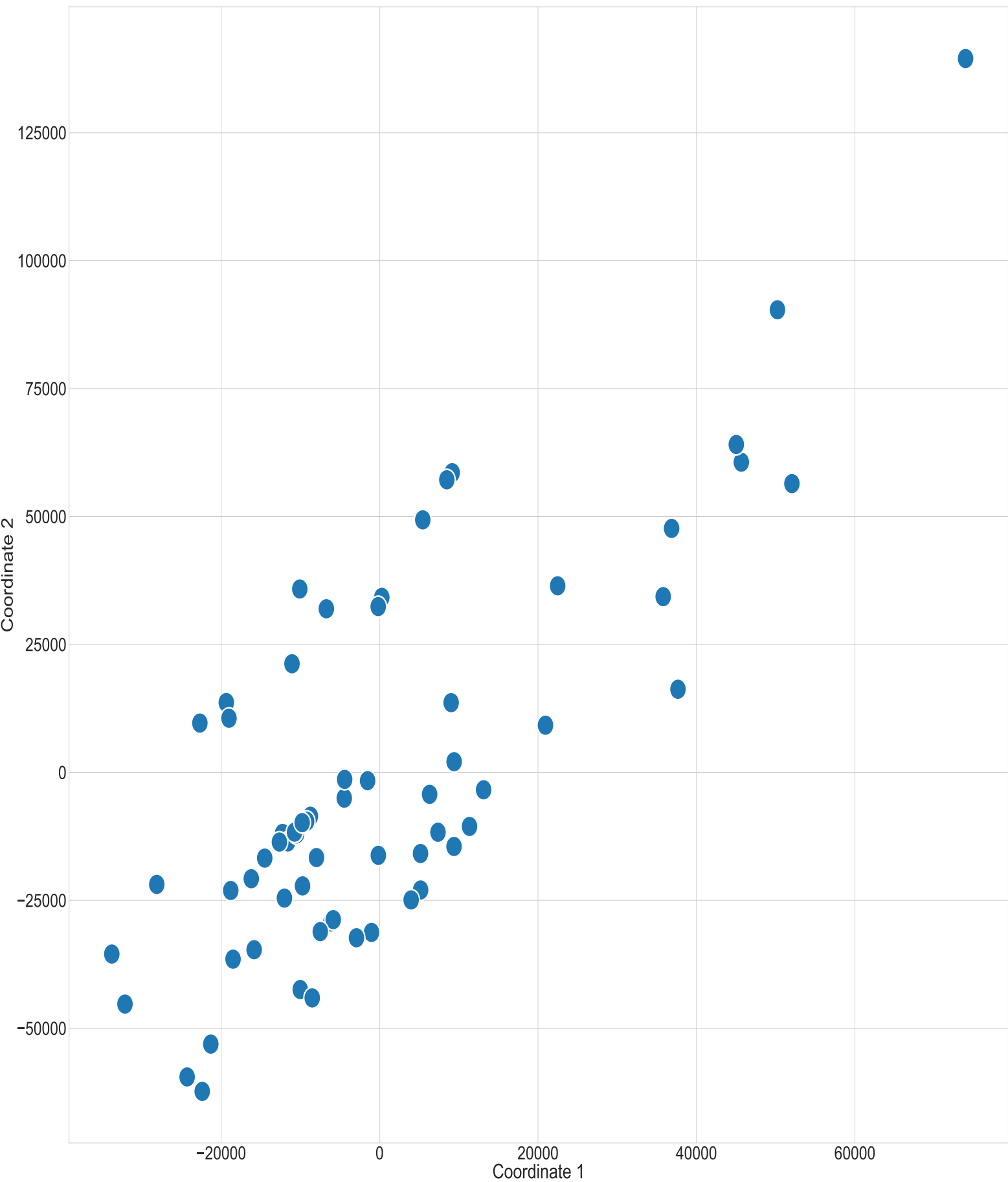
Scatter plot of Siderophore *Micromonosporaceae* BGCs analyzed via multidimensional scaling using canberra pairwise distance. Only Siderophore BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Siderophore BGCs is associated with the GCFs in BiG_SLiCE they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

Figure – S60 :



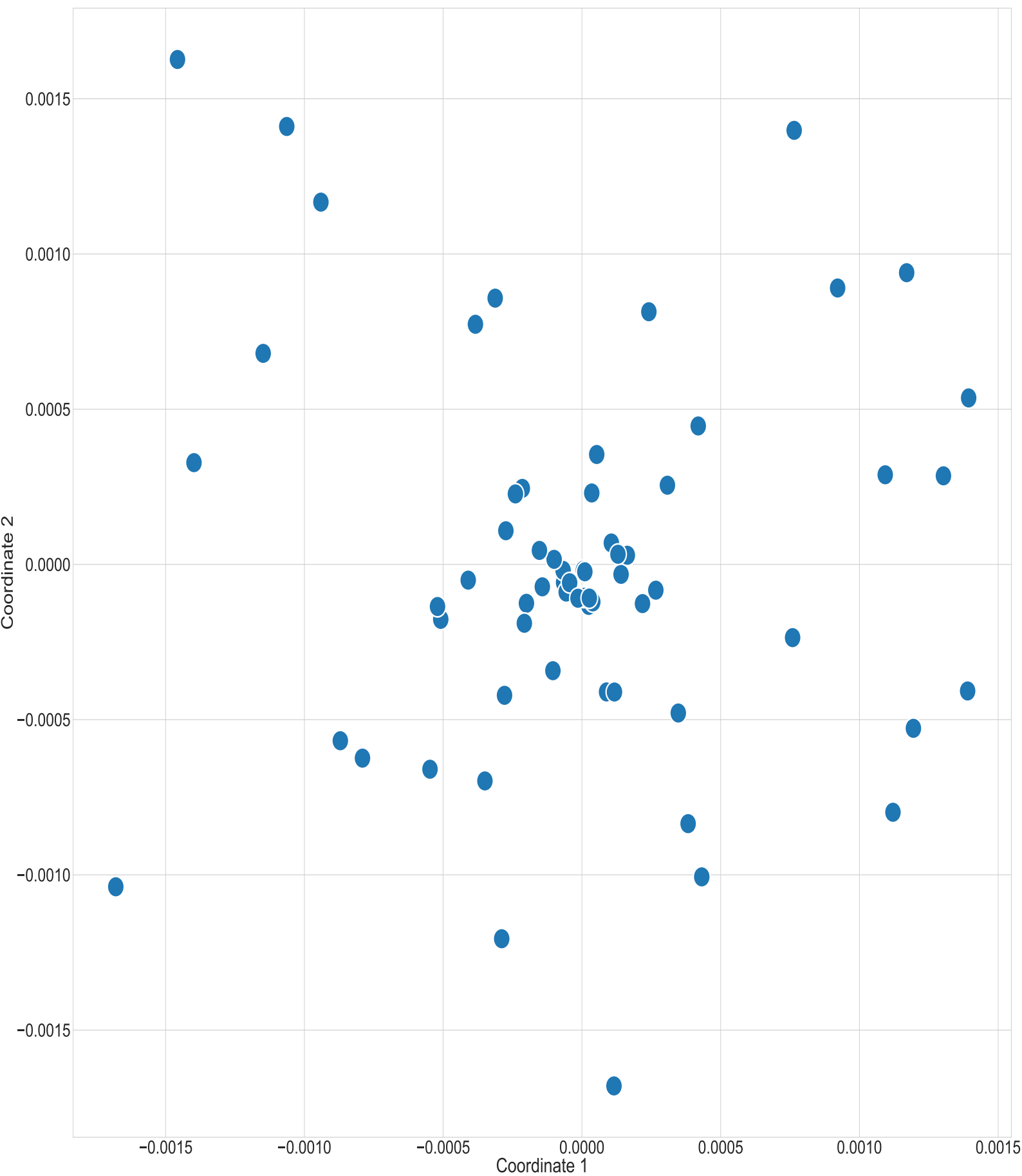
Scatter plot of Siderophore *Micromonosporaceae* BGCs analyzed via multidimensional scaling using correlation pairwise distance. Only Siderophore BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Siderophore BGCs is associated with the GCFs in BiG_SLiCE they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

Figure – S61 :



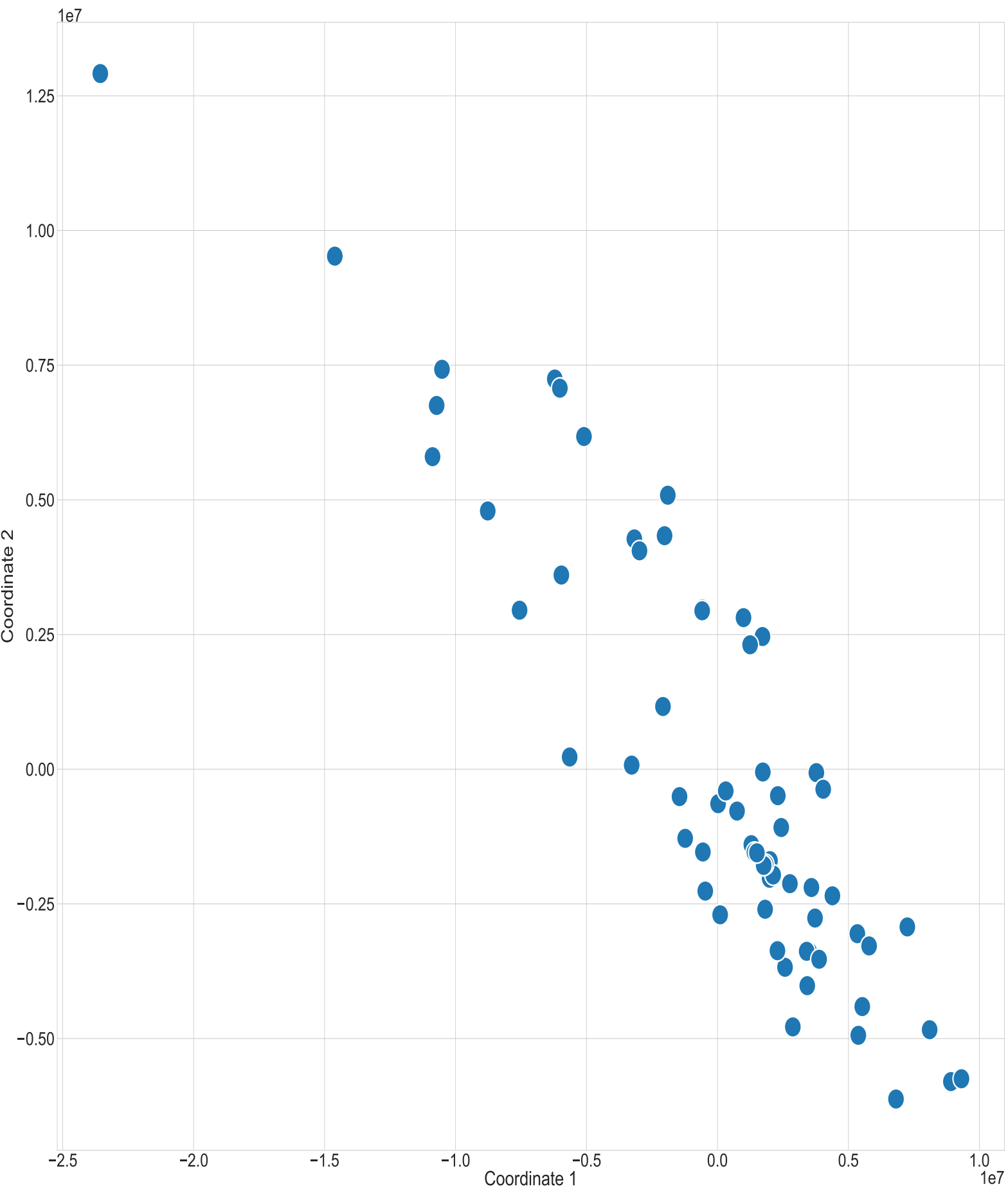
Scatter plot of PKS-NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using euclidean pairwise distance. Only PKS-NRPS BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between PKS-NRPS BGCs is associated with the GCFs in BiG_SLiCE they were most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

Figure – S62 :



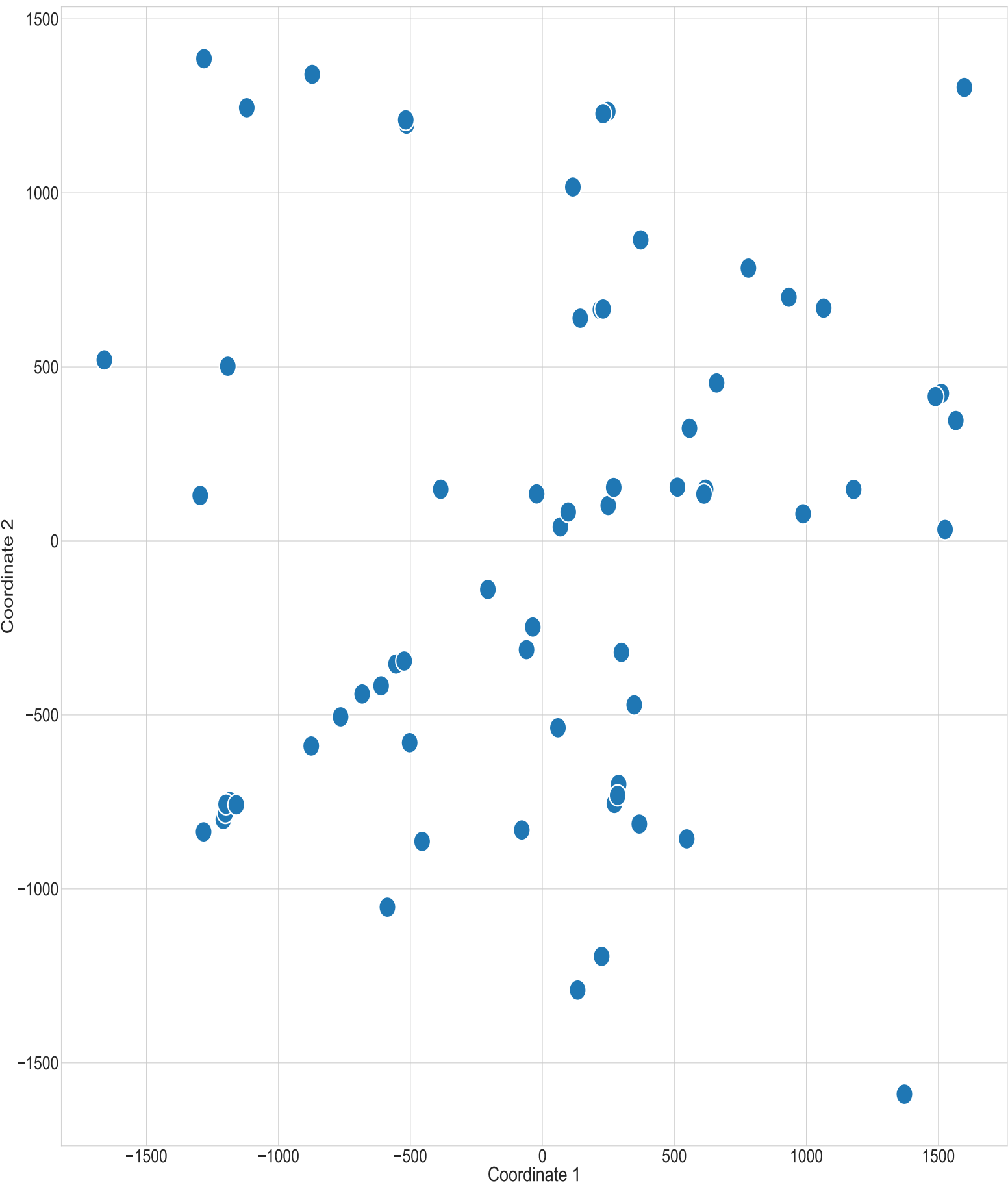
Scatter plot of PKS-NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cosine pairwise distance. Only PKS-NRPS BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between PKS-NRPS BGCs is associated with the GCFs in BiG_SLiCE they were most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

Figure – S63 :



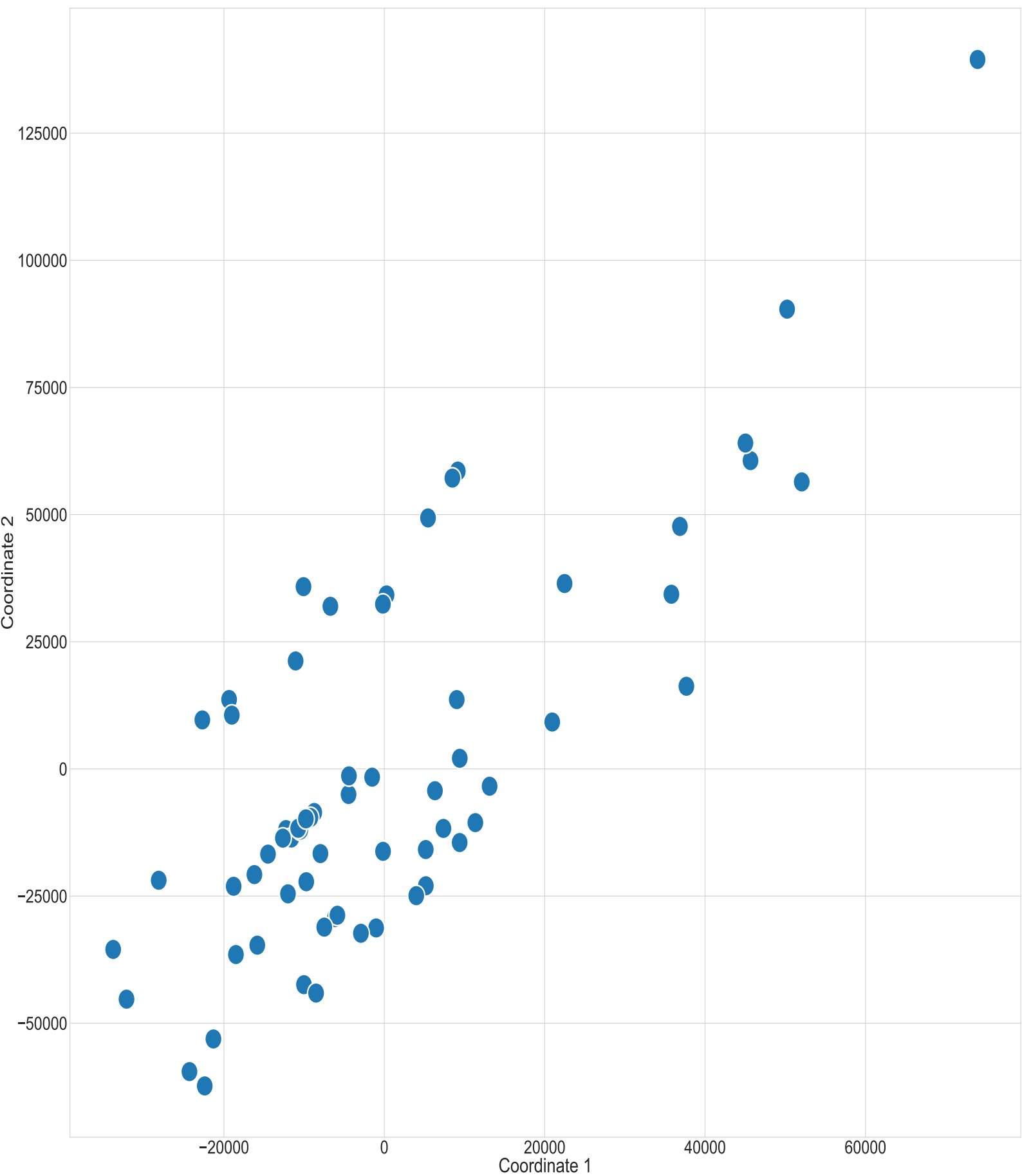
Scatter plot of PKS-NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cityblock pairwise distance. Only PKS-NRPS BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between PKS-NRPS BGCs is associated with the GCFs in BIG_SLICE they were most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

Figure – S64 :



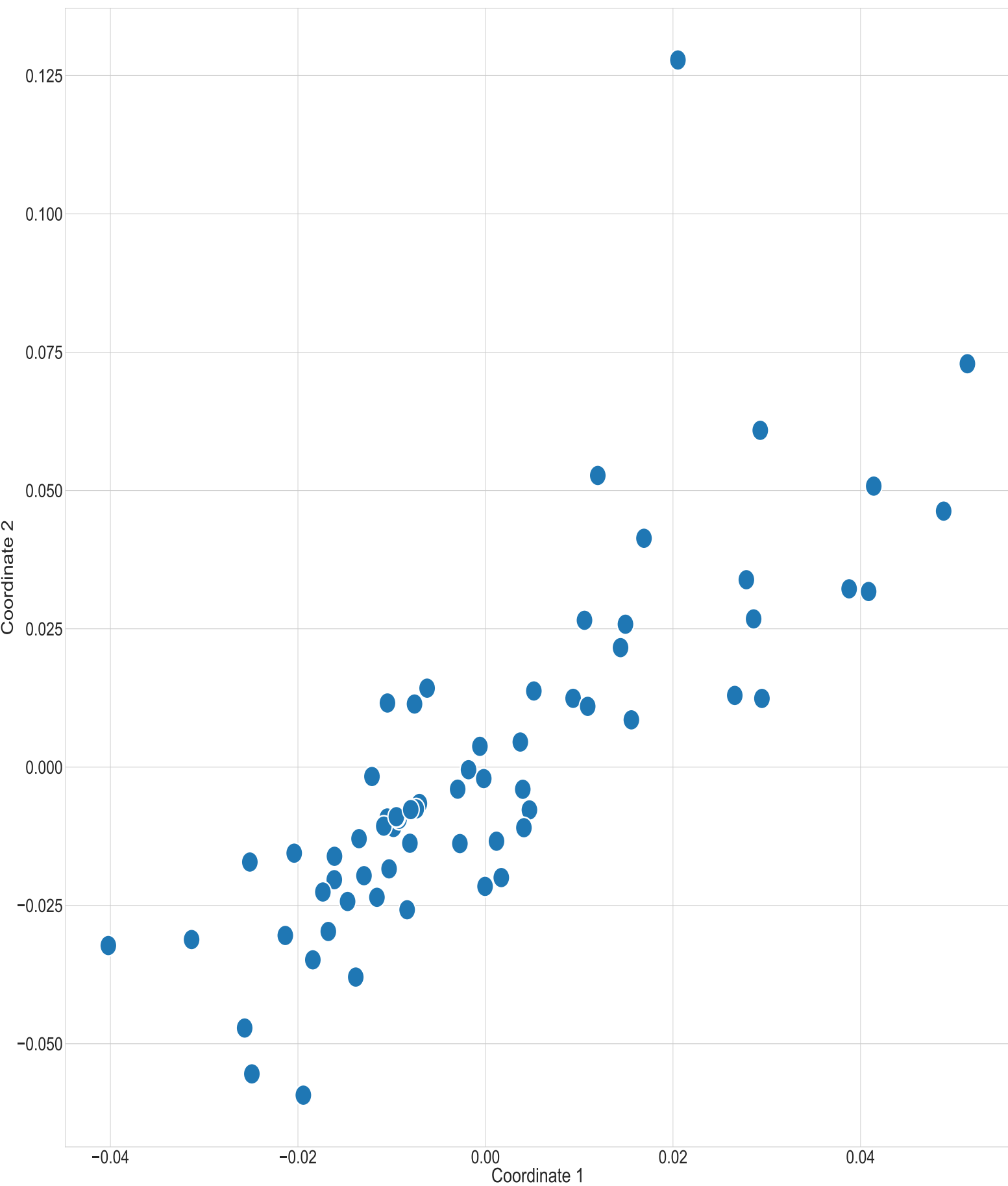
Scatter plot of PKS-NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using chebyshev pairwise distance. Only PKS-NRPS BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between PKS-NRPS BGCs is associated with the GCFs in BiG_SLiCE they were most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

Figure – S65 :



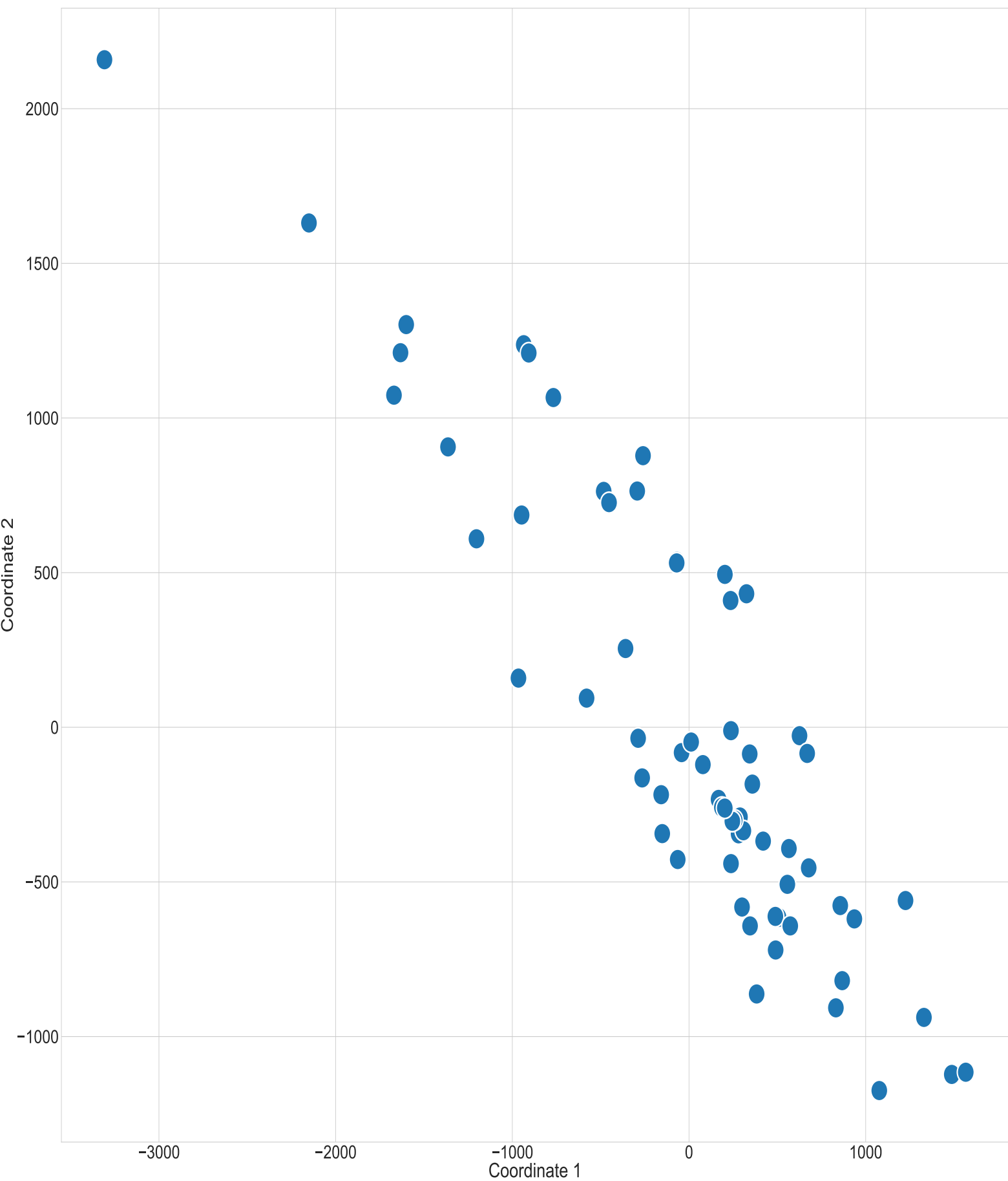
Scatter plot of PKS-NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using I2 pairwise distance. Only PKS-NRPS BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between PKS-NRPS BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S66 :



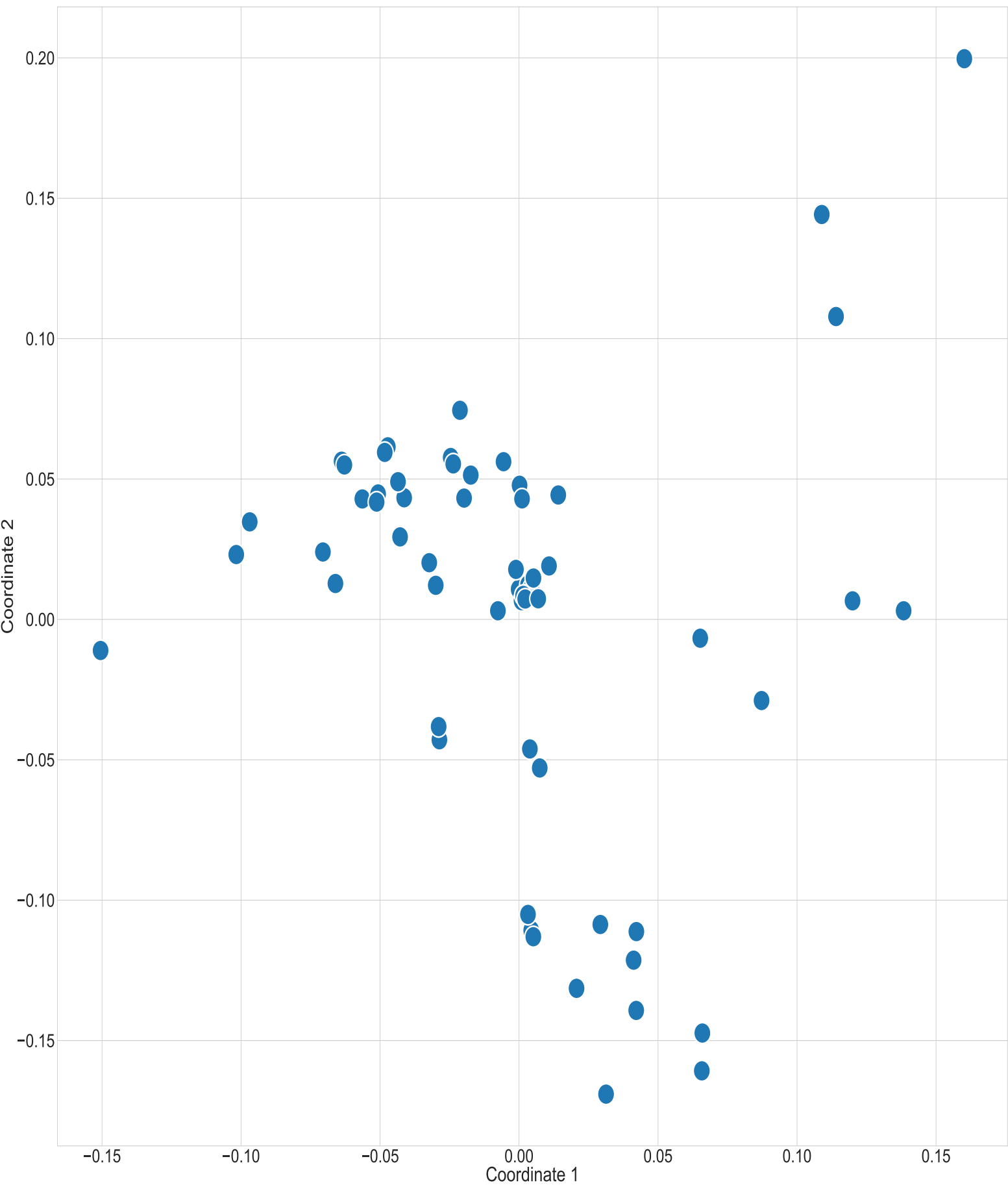
Scatter plot of PKS-NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using braycurtis pairwise distance. Only PKS-NRPS BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between PKS-NRPS BGCs is associated with the GCFs in BIG_SLICE they were most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

Figure – S67 :



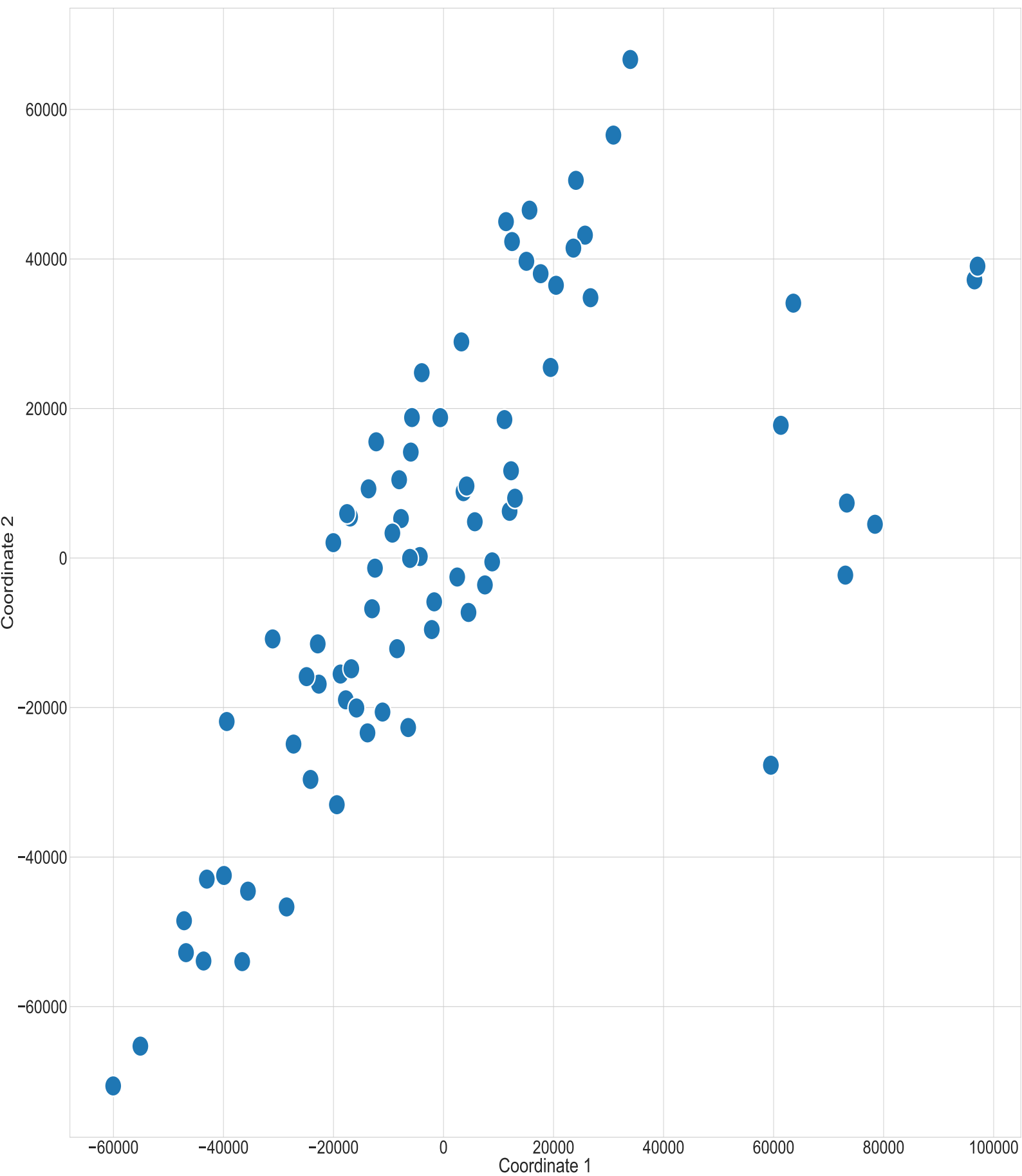
Scatter plot of PKS-NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using canberra pairwise distance. Only PKS-NRPS BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between PKS-NRPS BGCs is associated with the GCFs in BiG_SLiCE they were most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

Figure – S68 :



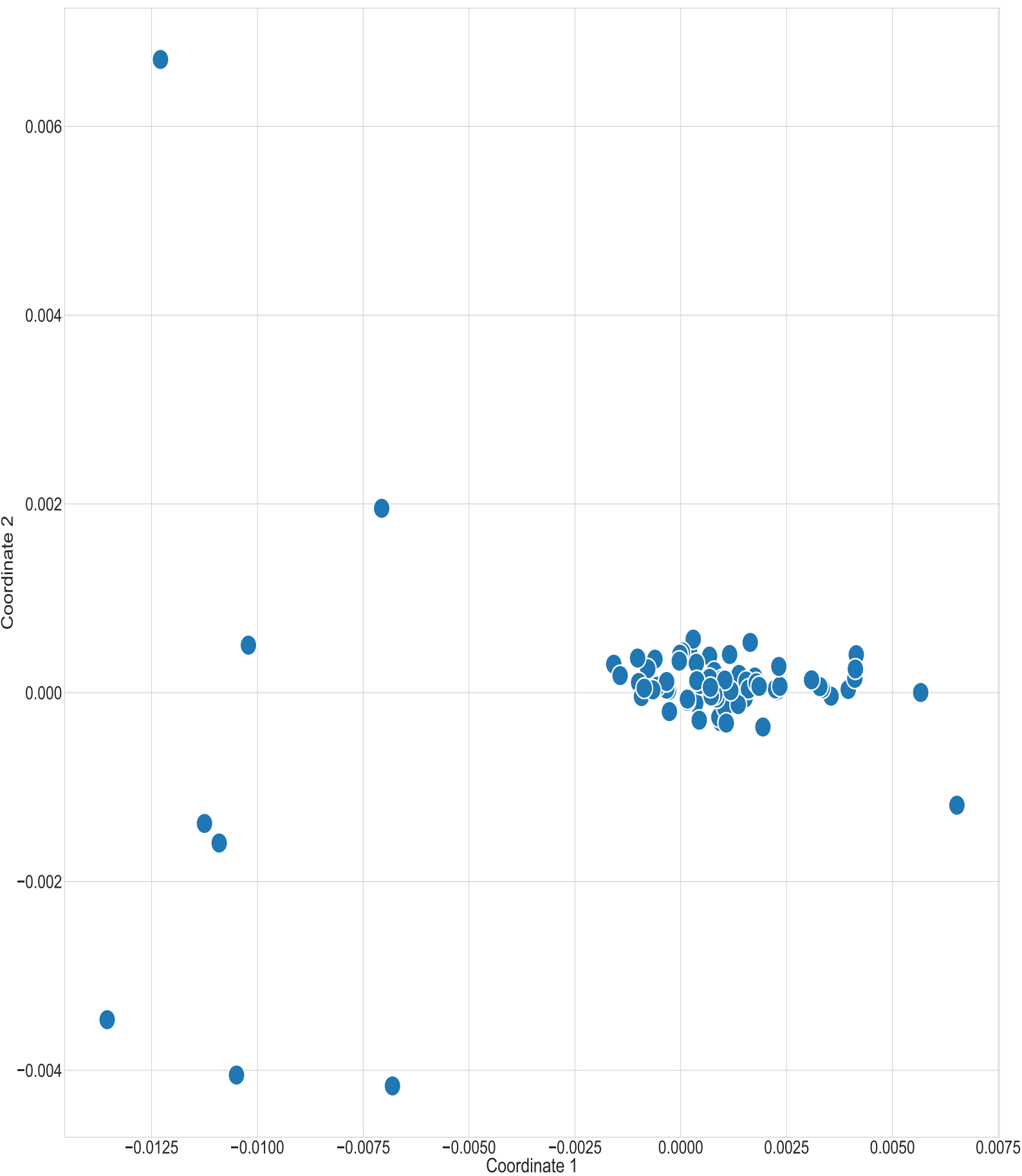
Scatter plot of PKS-NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using correlation pairwise distance. Only PKS-NRPS BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between PKS-NRPS BGCs is associated with the GCFs in BiG_SLiCE they were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

Figure – S69 :



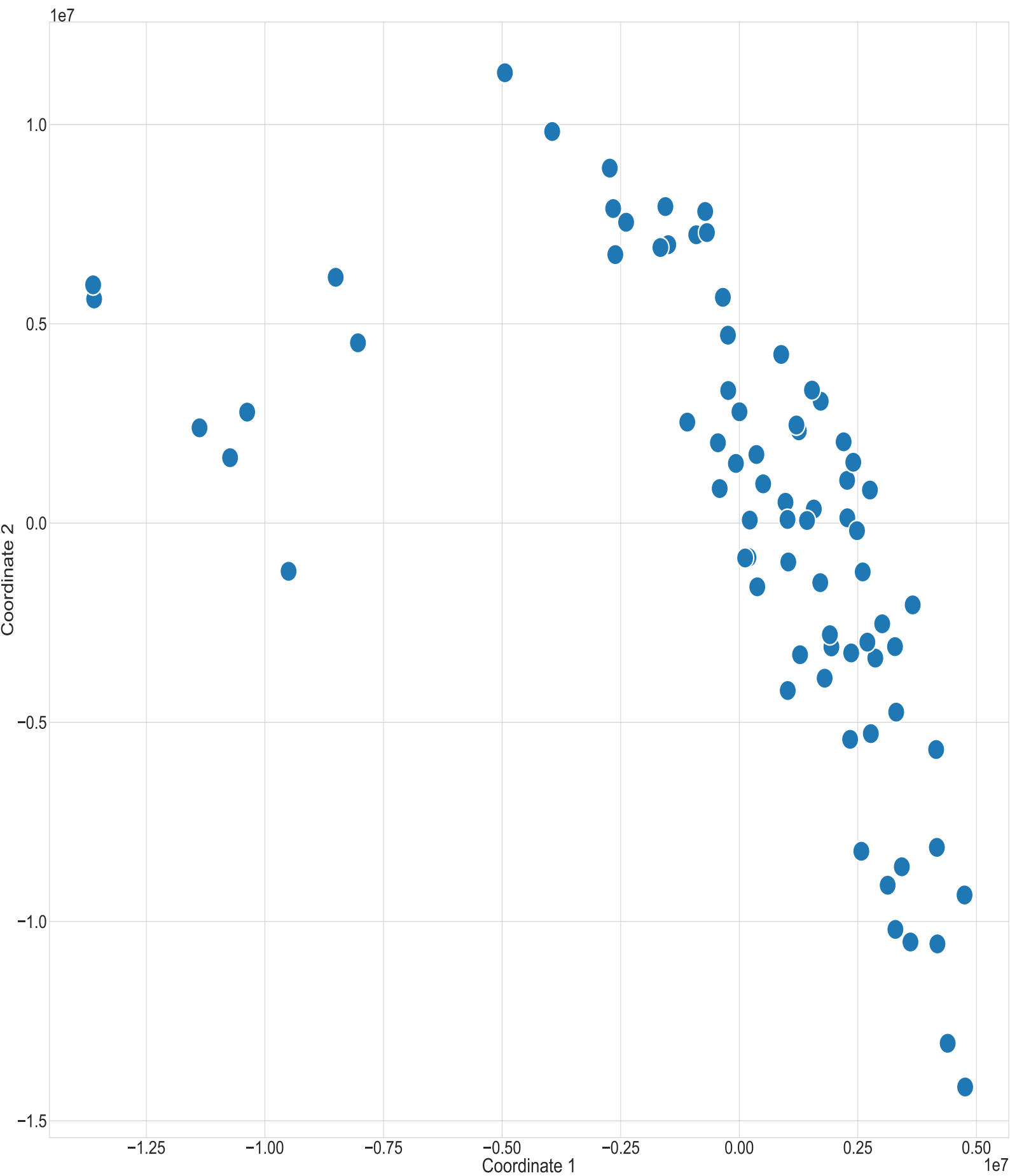
Scatter plot of NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using euclidean pairwise distance. Only NRPS BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between NRPS BGCs is associated with the GCFs in BiG_SLiCE they were most similar to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

Figure – S70 :



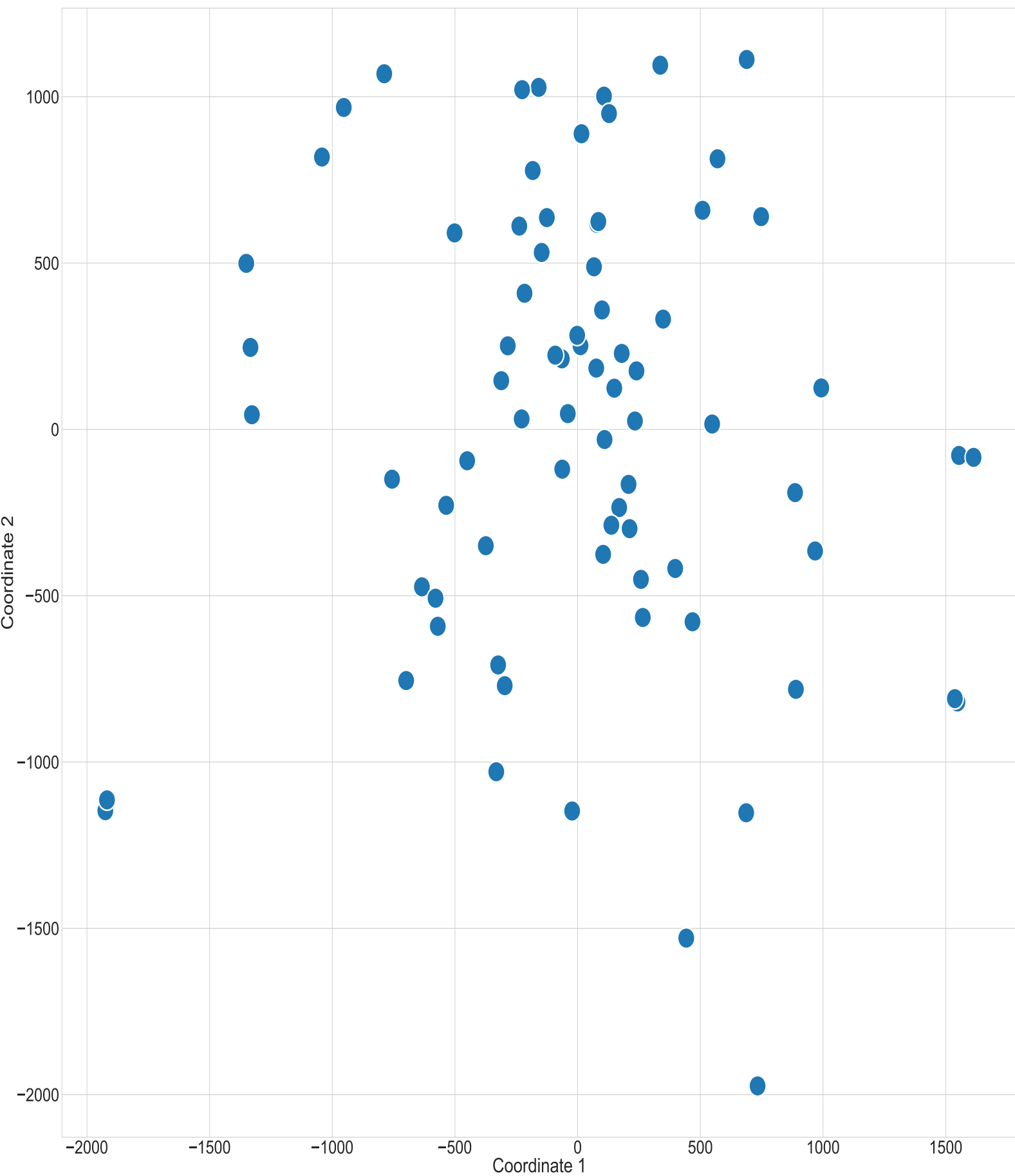
Scatter plot of NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cosine pairwise distance. Only NRPS BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between NRPS BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S71 :



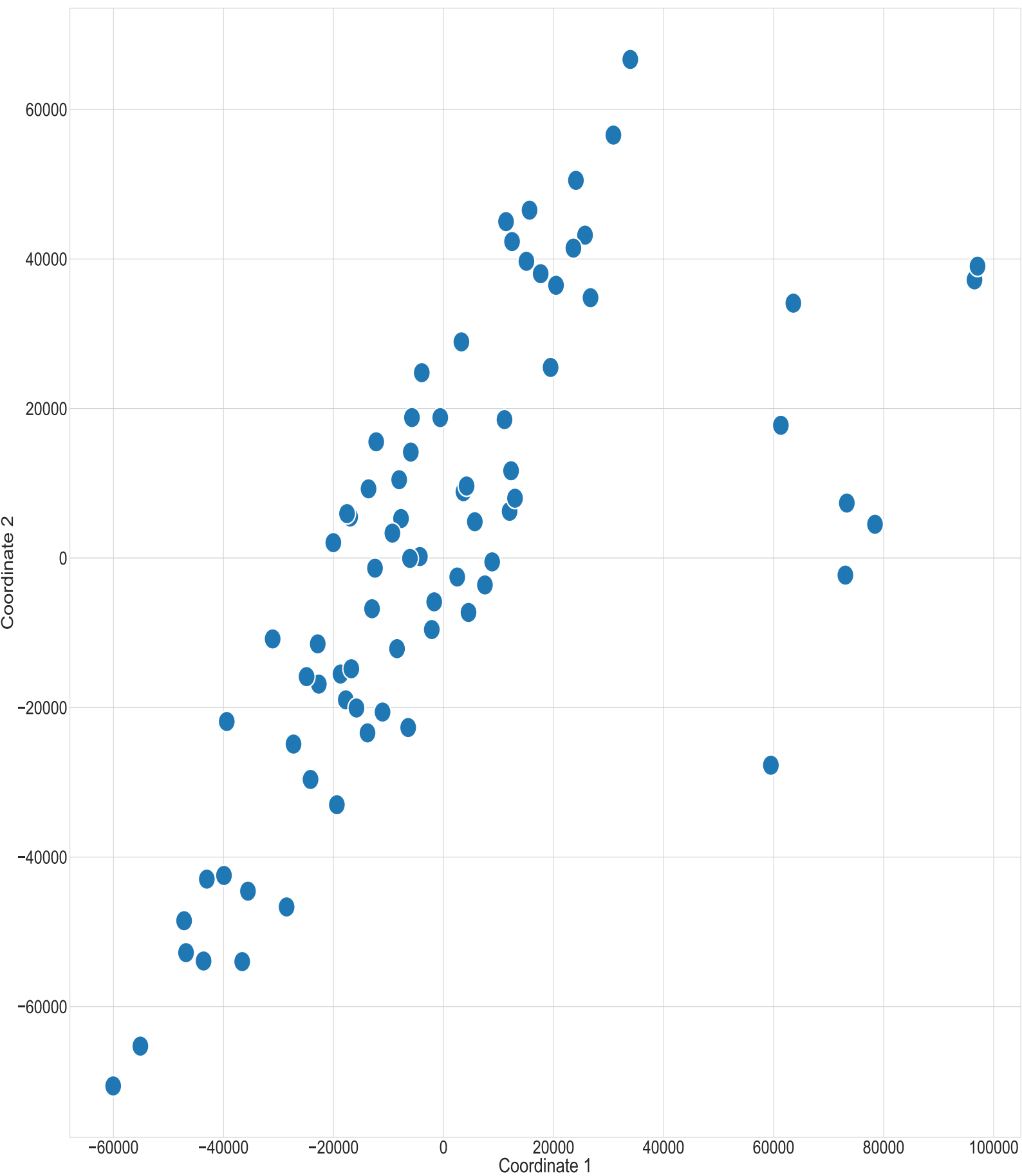
Scatter plot of NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cityblock pairwise distance. Only NRPS BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between NRPS BGCs is associated with the GCFs in BiG_SLiCE they were most similar to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

Figure – S72 :



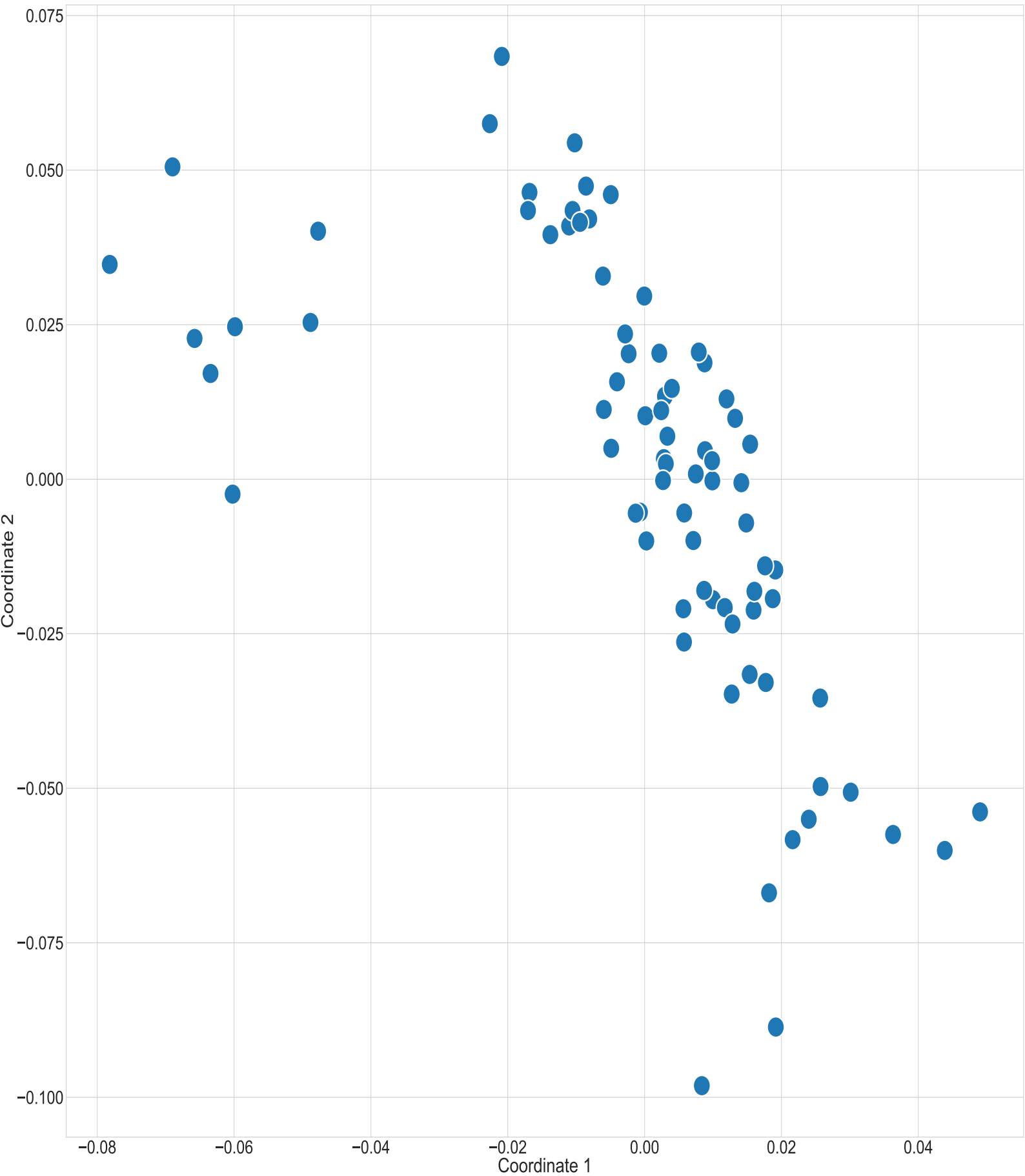
Scatter plot of NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using chebyshev pairwise distance. Only NRPS BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between NRPS BGCs is associated with the GCFs in BiG_SLiCE they were most similar to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

Figure – S73 :



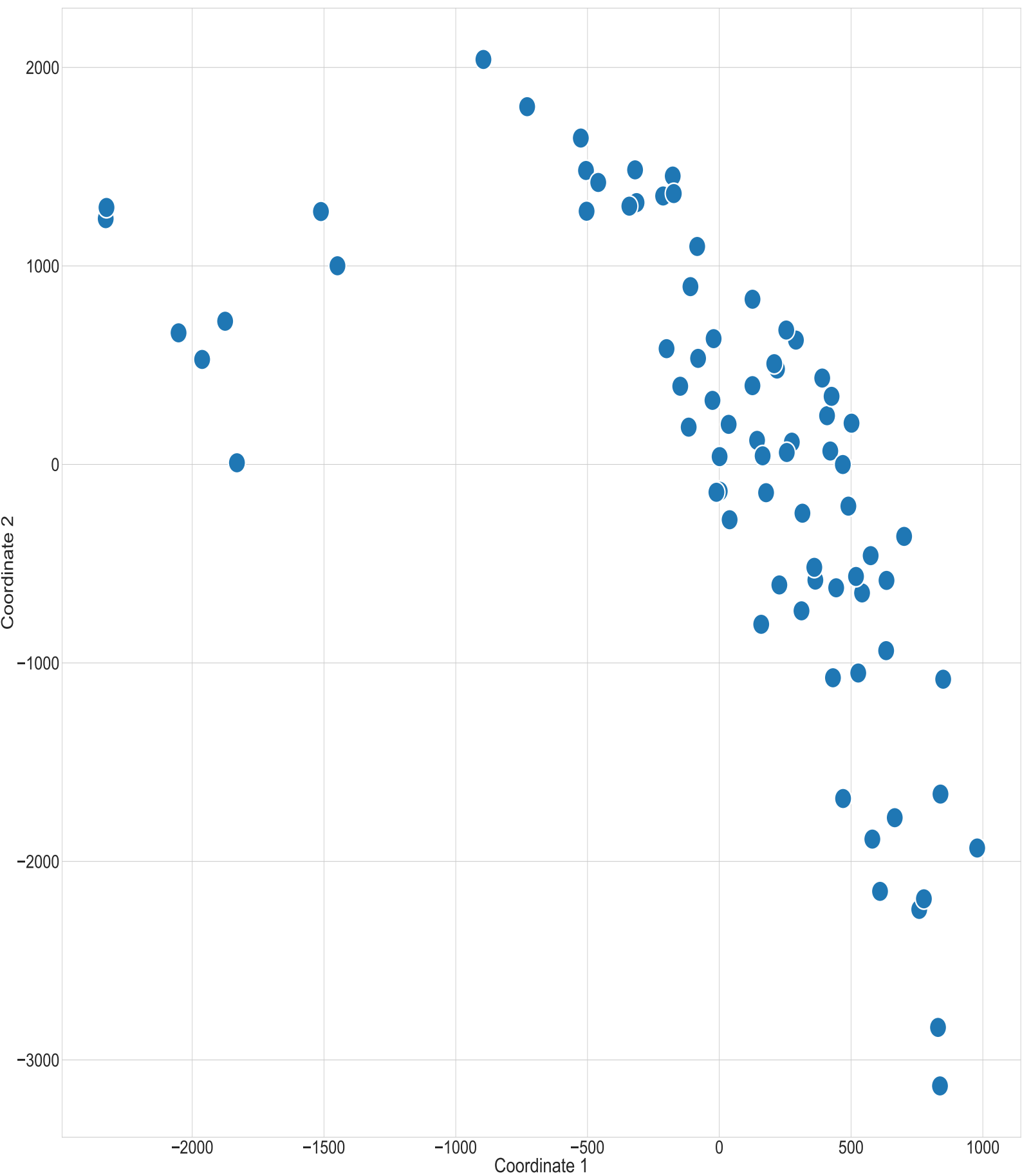
Scatter plot of NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using I2 pairwise distance. Only NRPS BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between NRPS BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S74 :



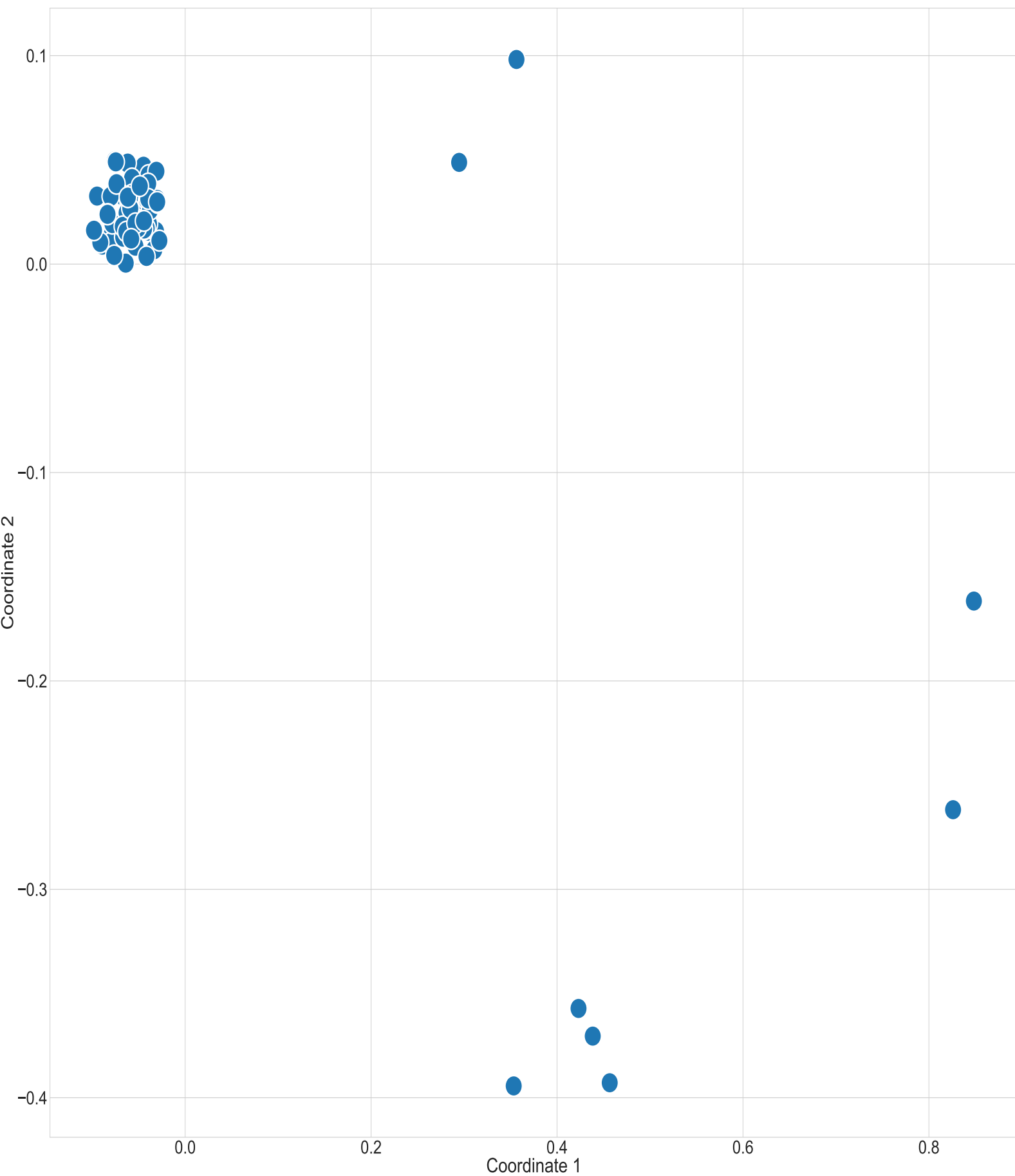
Scatter plot of NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using braycurtis pairwise distance. Only NRPS BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between NRPS BGCs is associated with the GCFs in BiG_SLiCE they were most similar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

Figure – S75 :



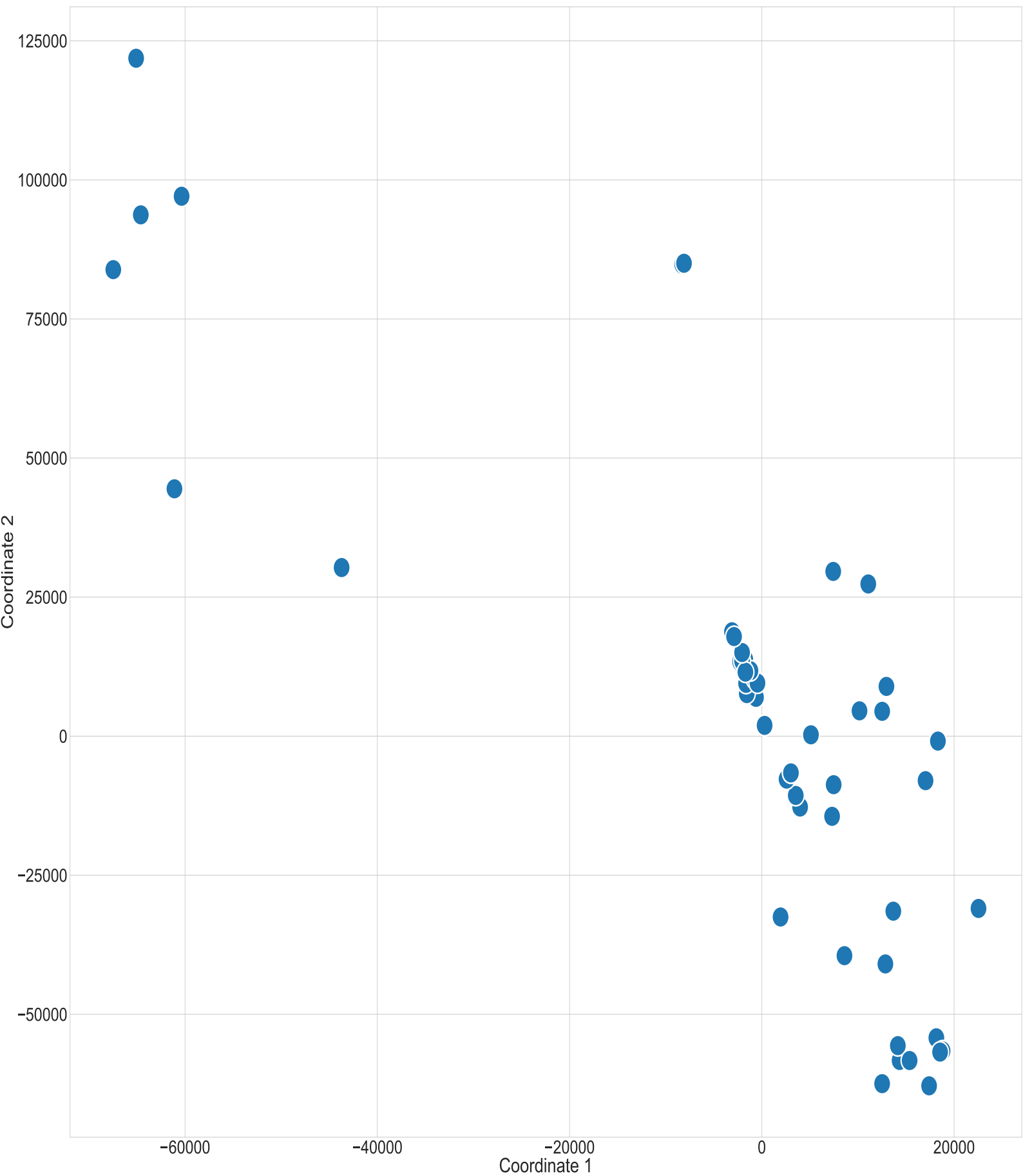
Scatter plot of NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using canberra pairwise distance. Only NRPS BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between NRPS BGCs is associated with the GCFs in BiG_SLiCE they were most similar to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

Figure – S76 :



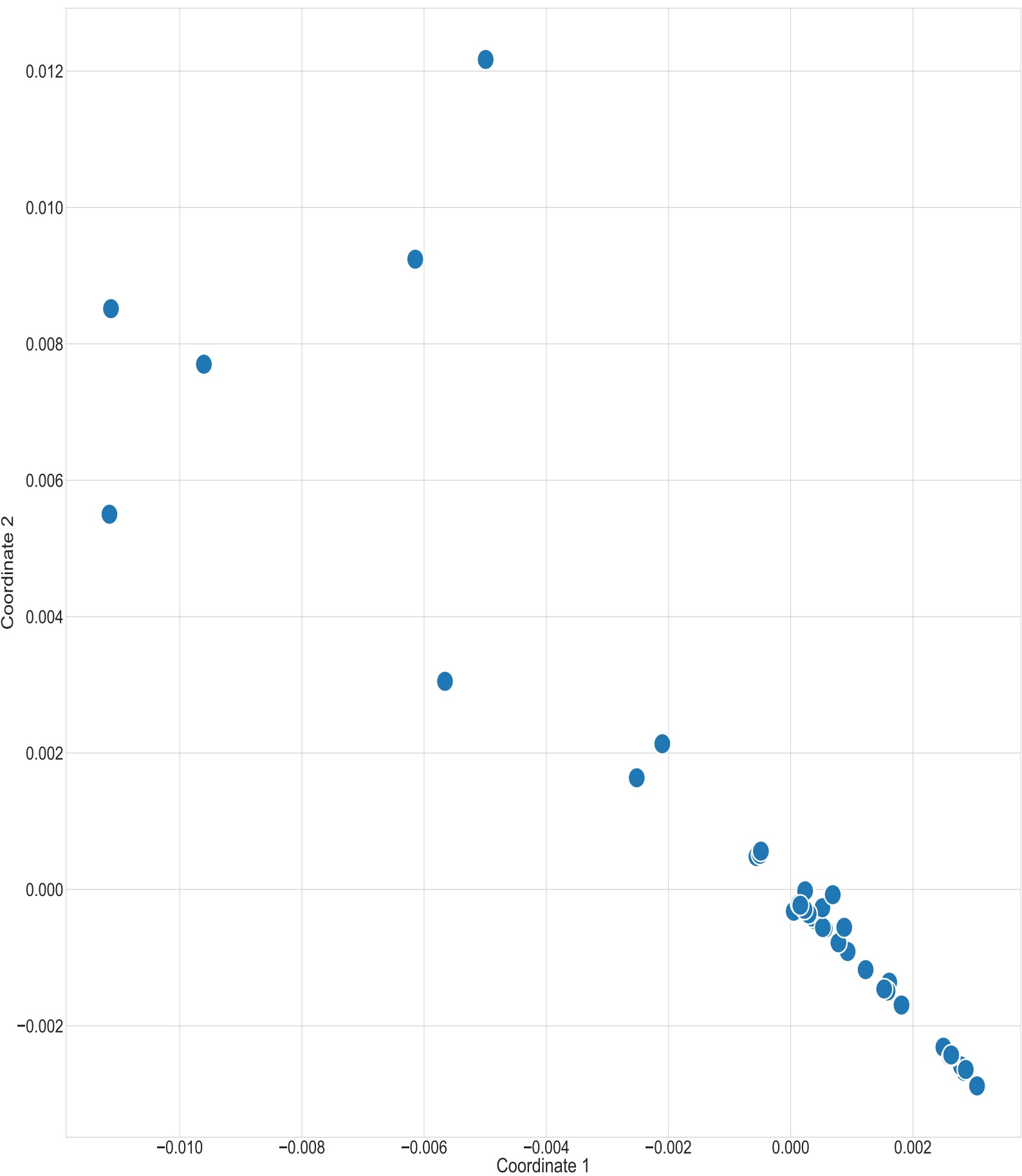
Scatter plot of NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using correlation pairwise distance. Only NRPS BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between NRPS BGCs is associated with the GCFs in BiG_SLiCE they were most similar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

Figure – S77 :



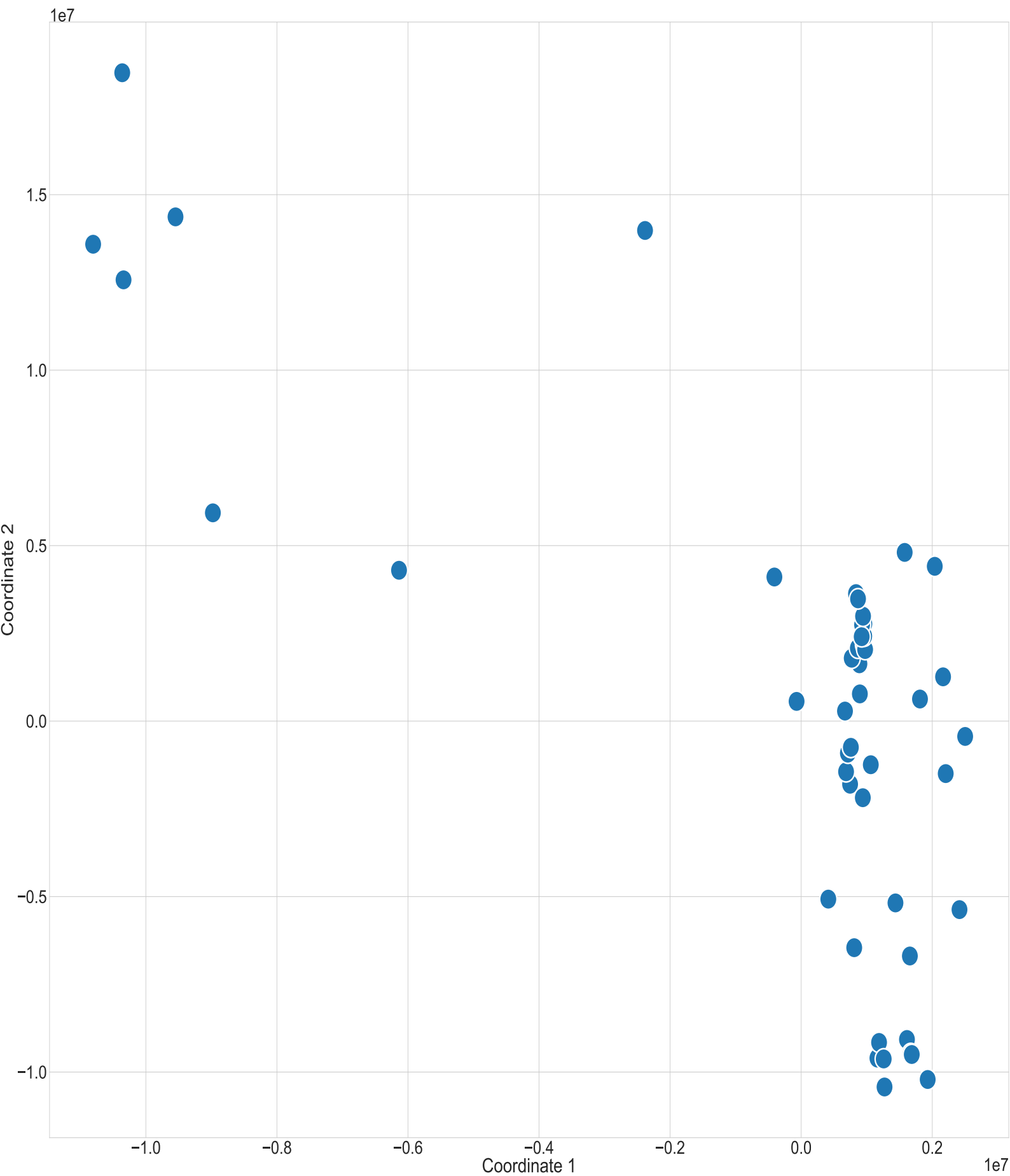
Scatter plot of Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using euclidean pairwise distance. Only Other BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Other BGCs is associated with the GCFs in BiG_SLiCE they were most similar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

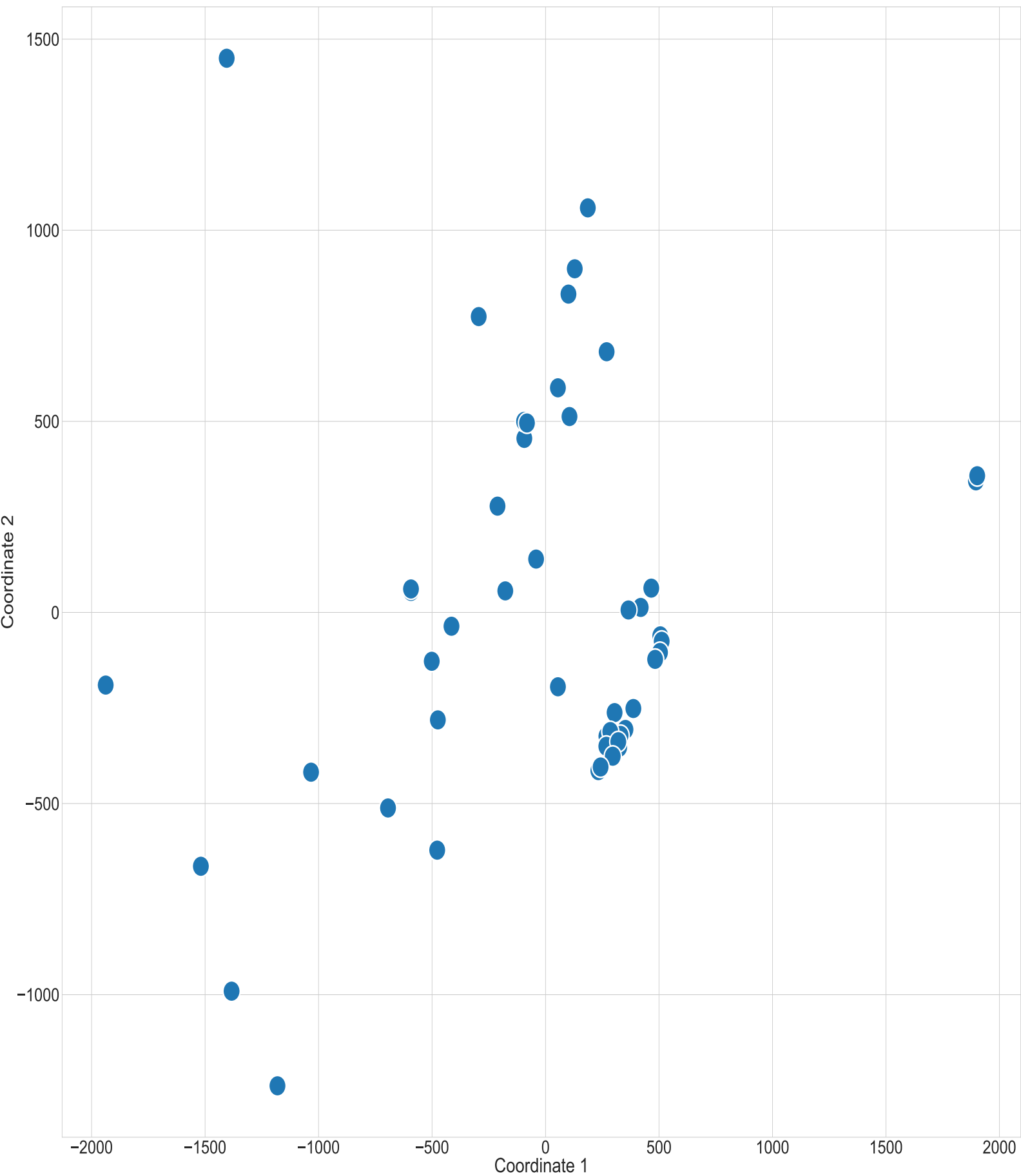
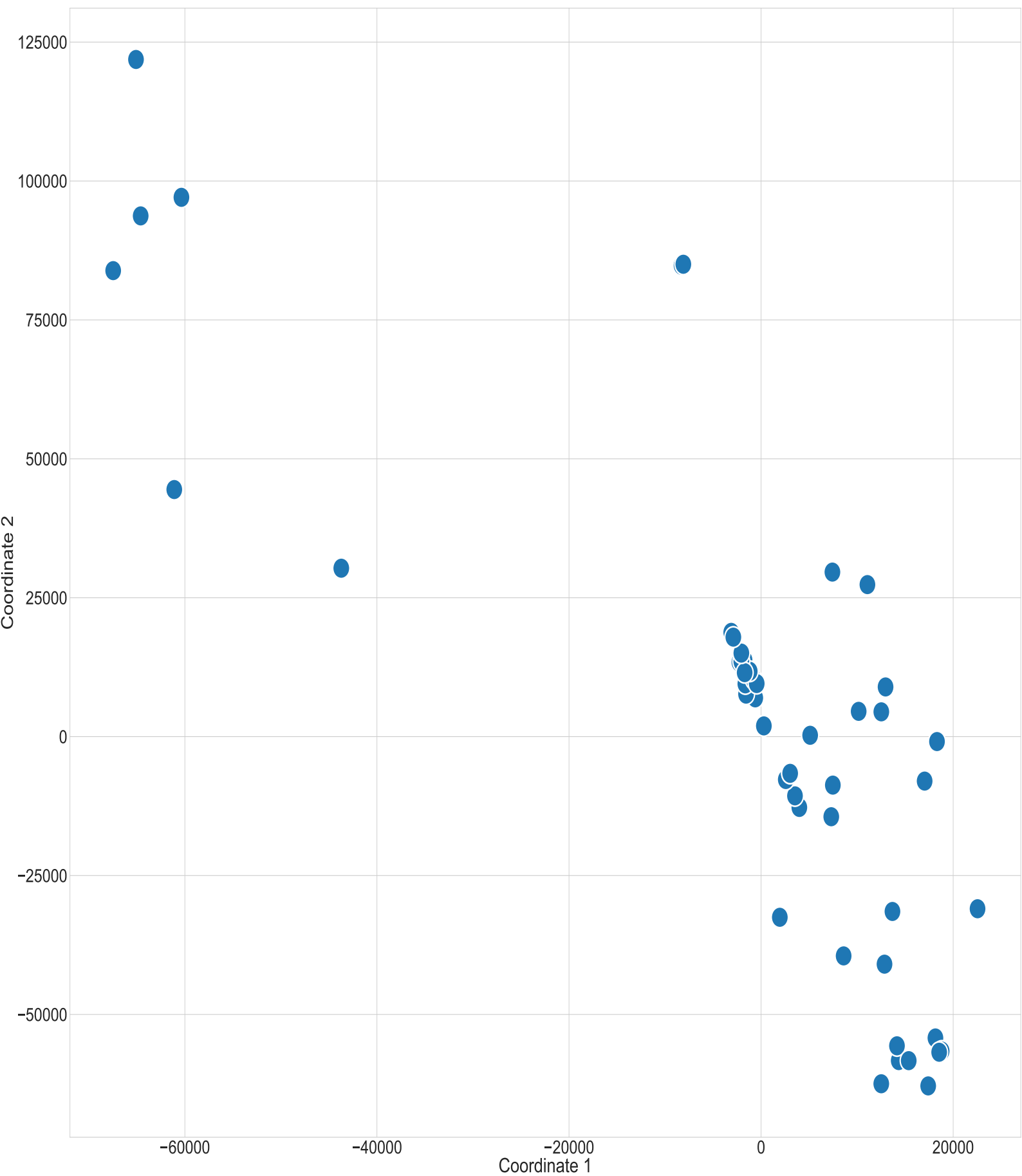
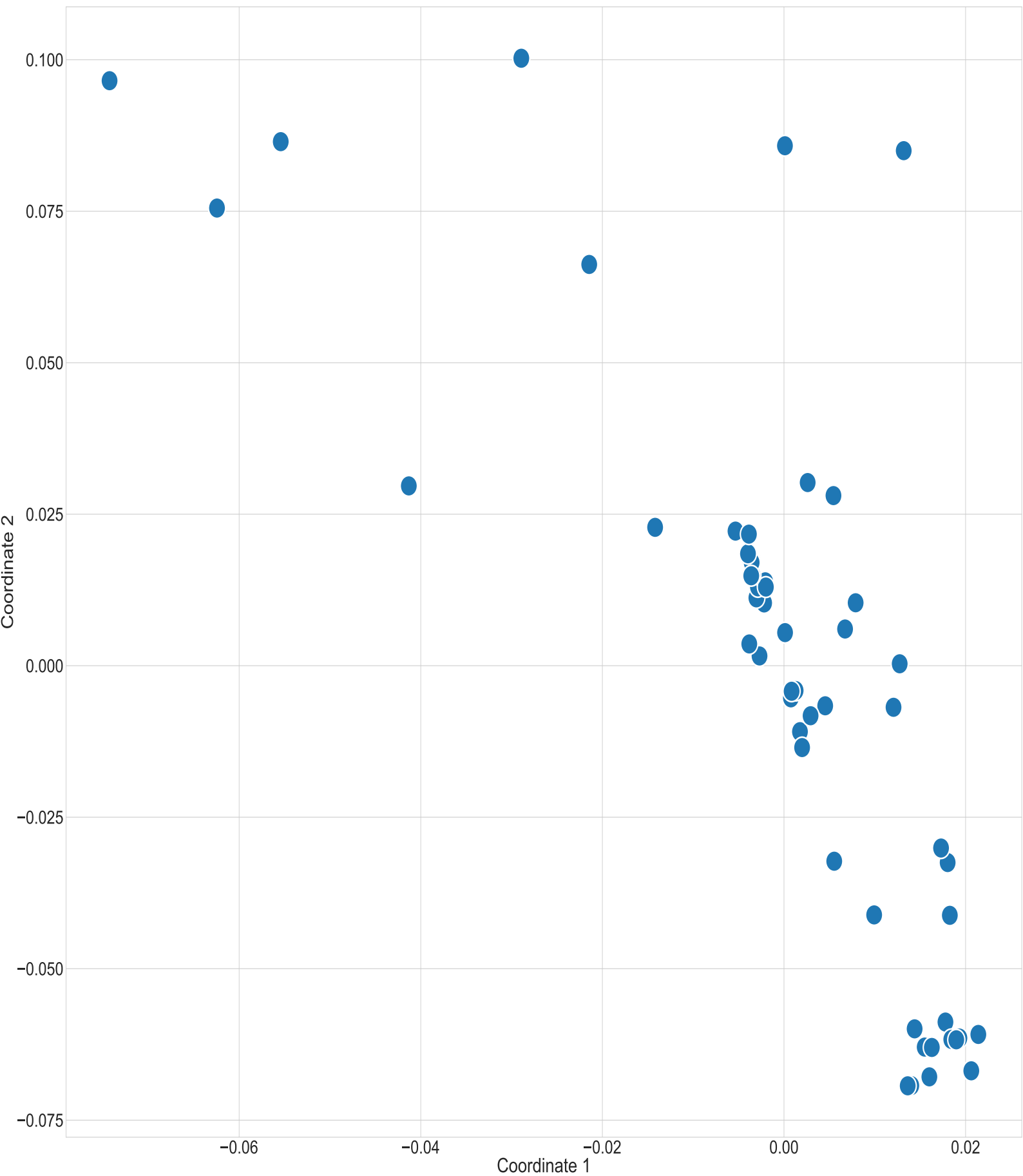


Figure – S81 :



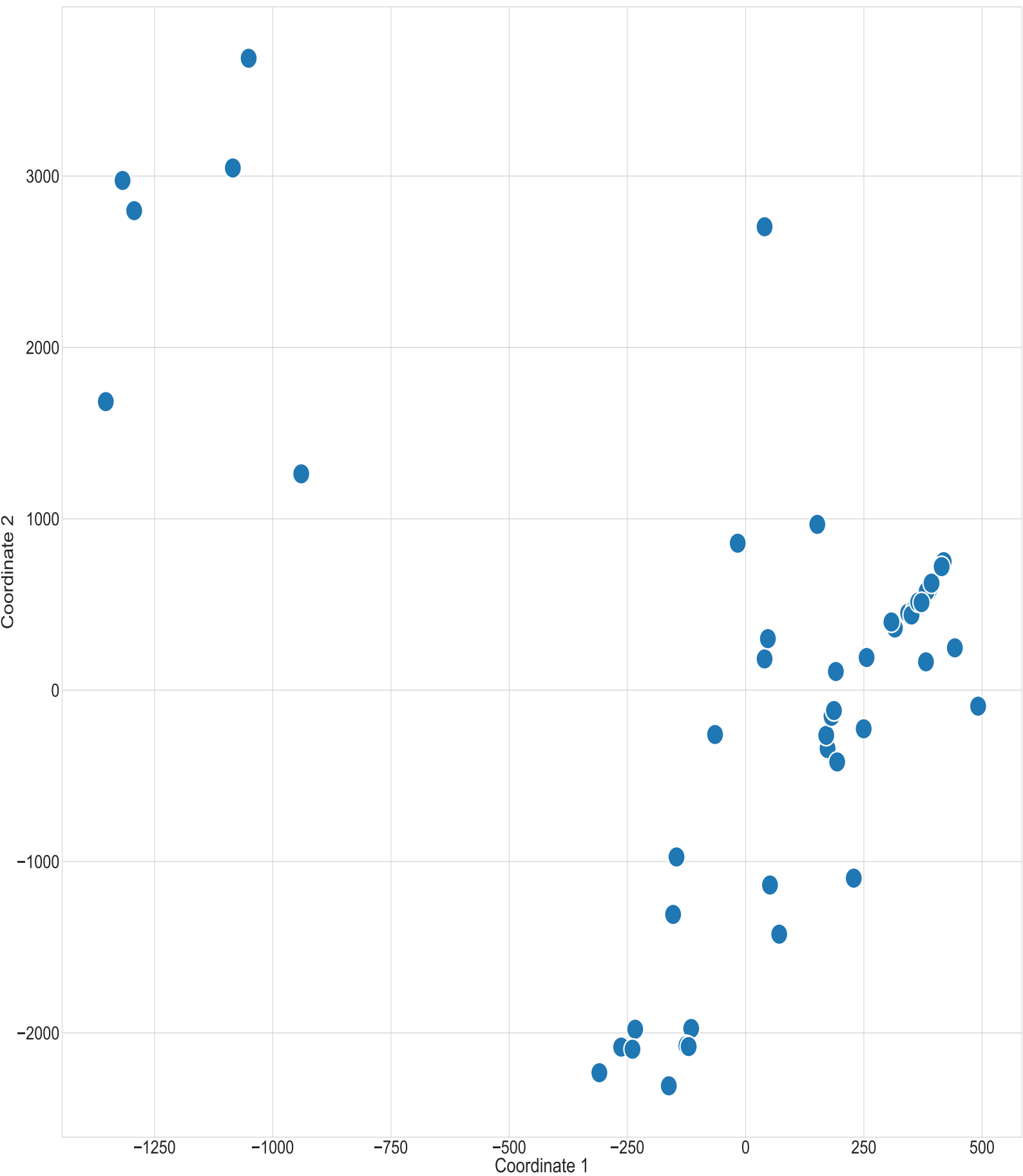
Scatter plot of Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using I2 pairwise distance. Only Other BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Other BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S82 :



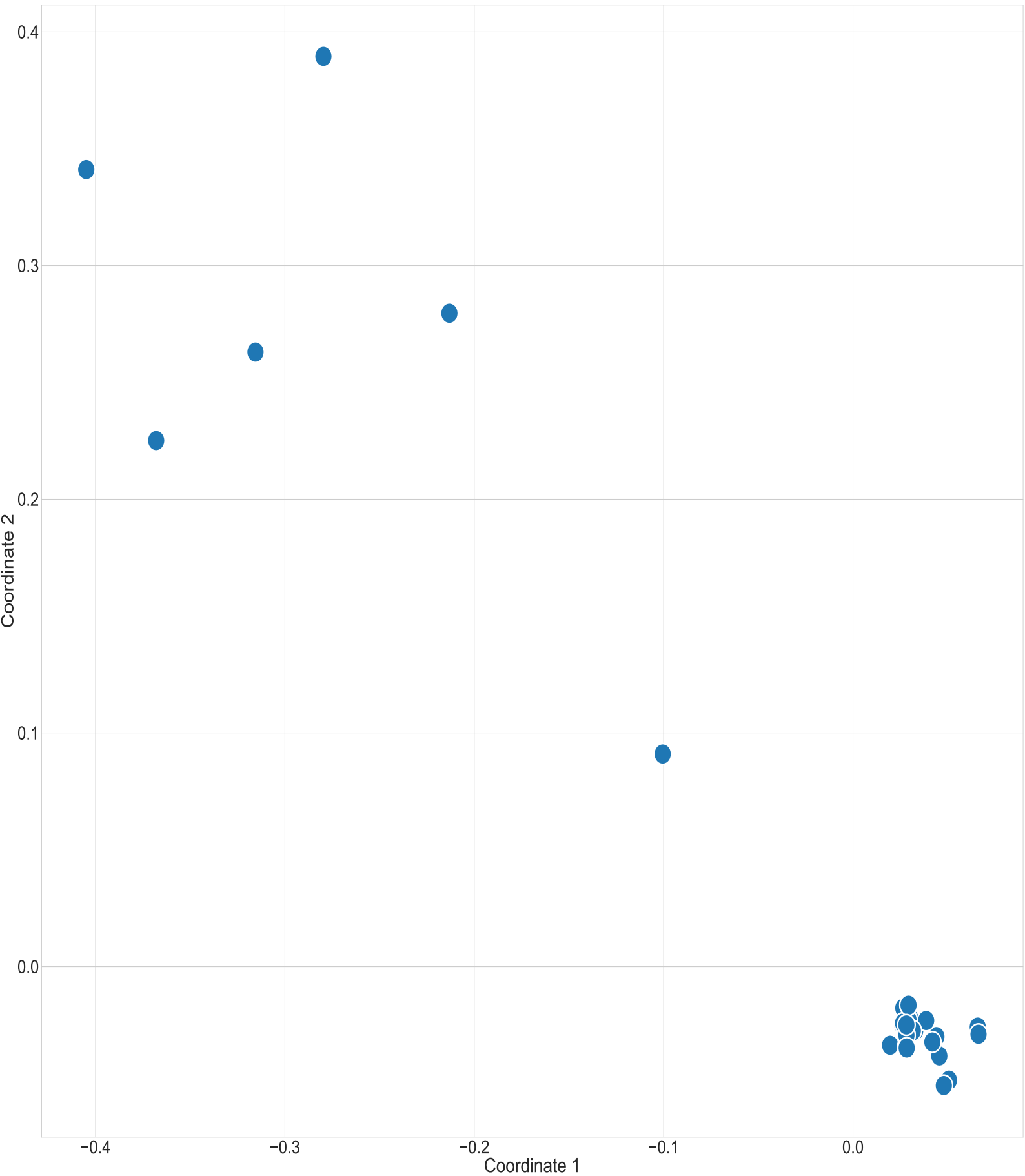
Scatter plot of Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using braycurtis pairwise distance. Only Other BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Other BGCs is associated with the GCFs in BiG_SLiCE they were most similar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

Figure – S83 :



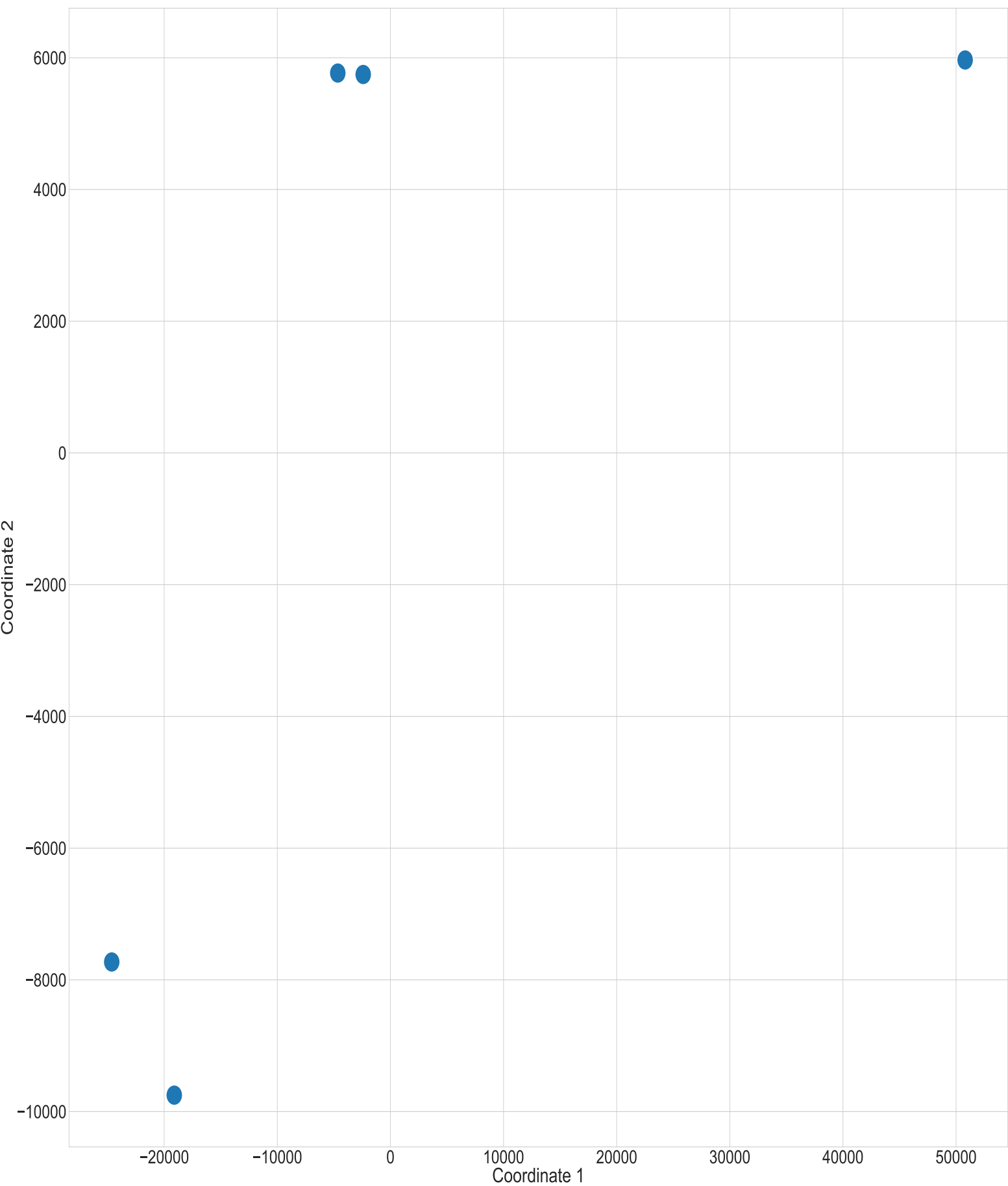
Scatter plot of Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using canberra pairwise distance. Only Other BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Other BGCs is associated with the GCFs in BiG_SLiCE they were most similar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

Figure – S84 :



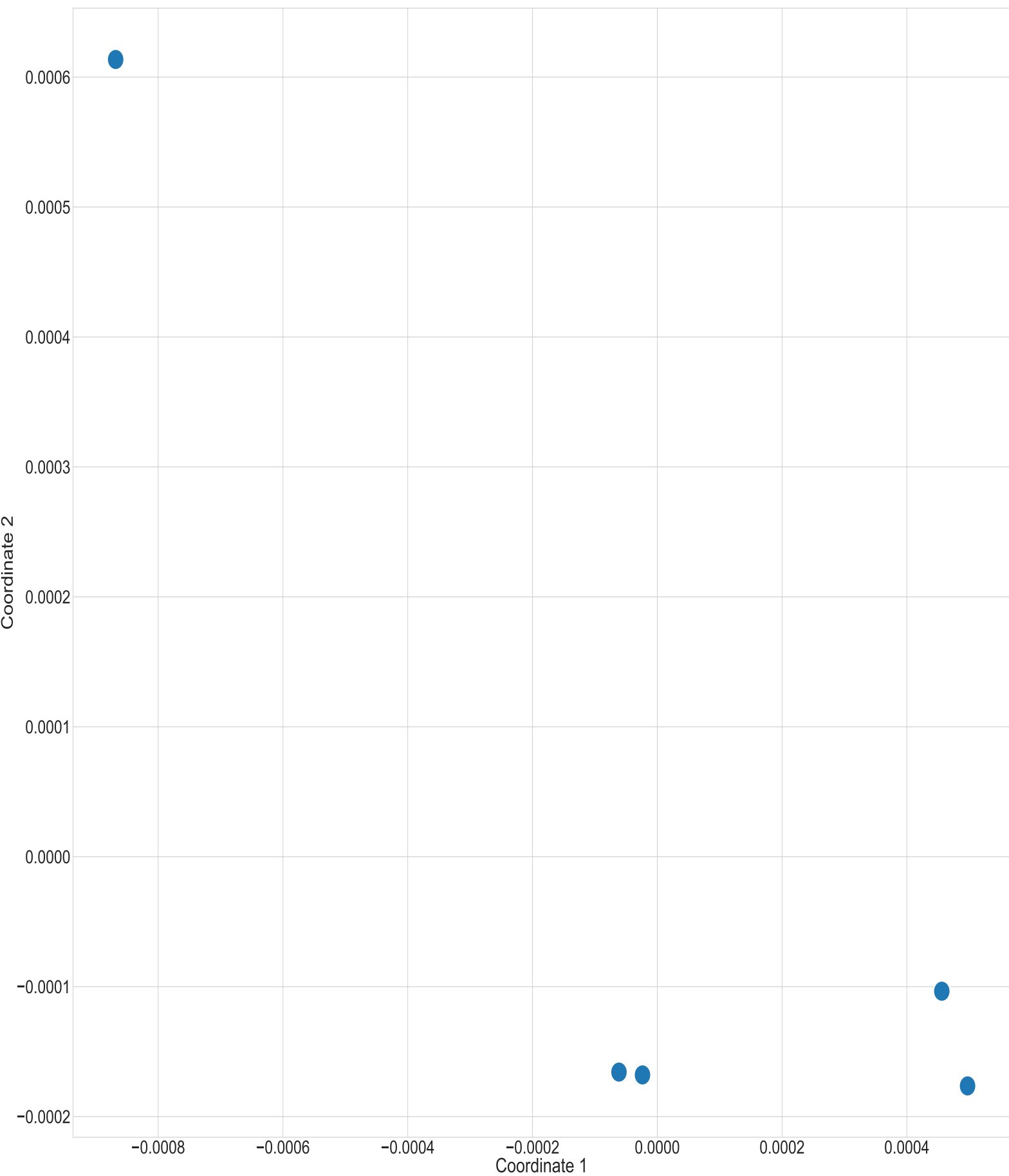
Scatter plot of Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using correlation pairwise distance. Only Other BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Other BGCs is associated with the GCFs in BiG_SLICE they were most similar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

Figure – S85 :



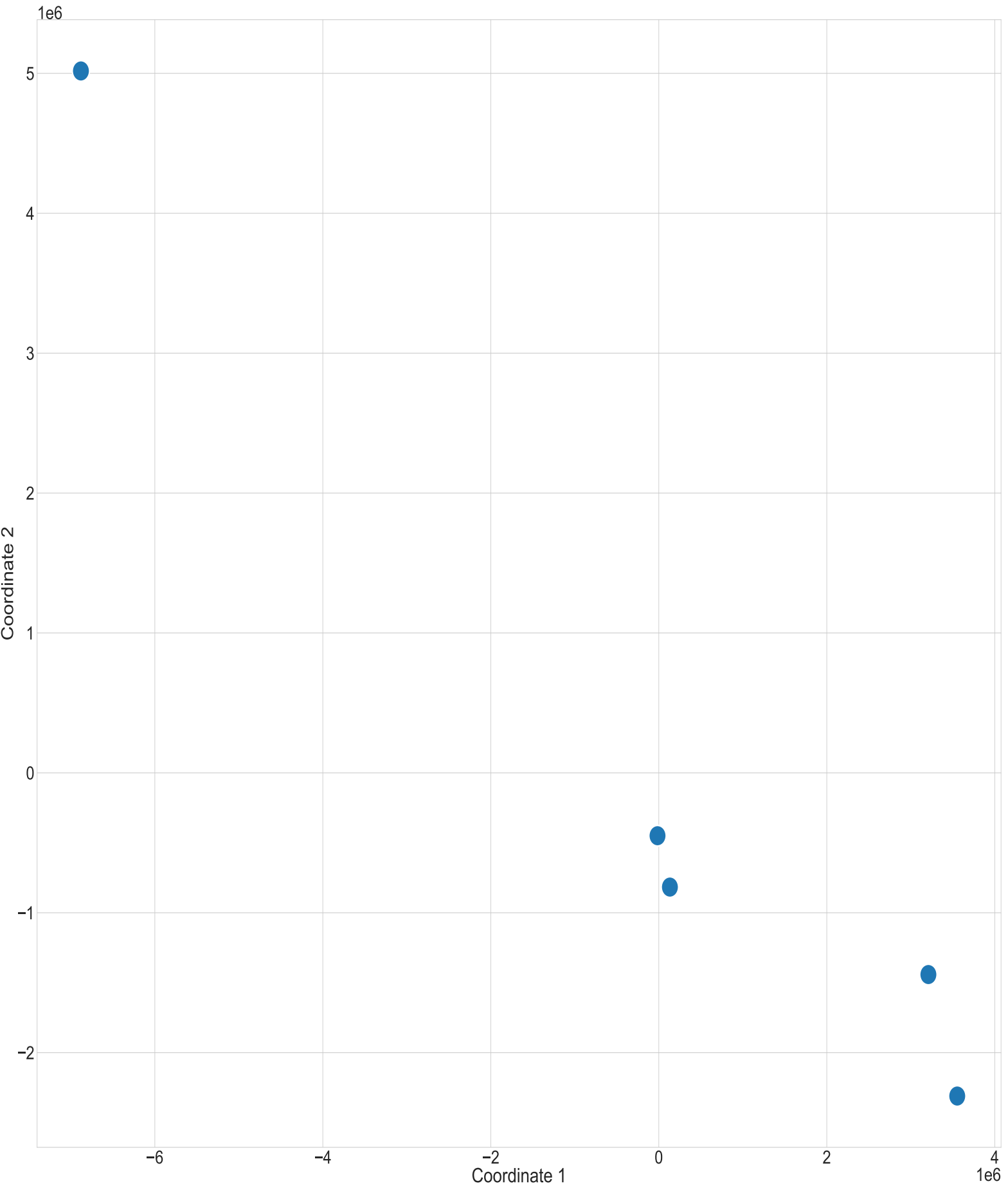
Scatter plot of Saccharides *Micromonosporaceae* BGCs analyzed via multidimensional scaling using euclidean pairwise distance. Only Saccharides BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Saccharides BGCs is associated with the GCFs in BiG_SLICE they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

Figure – S86 :



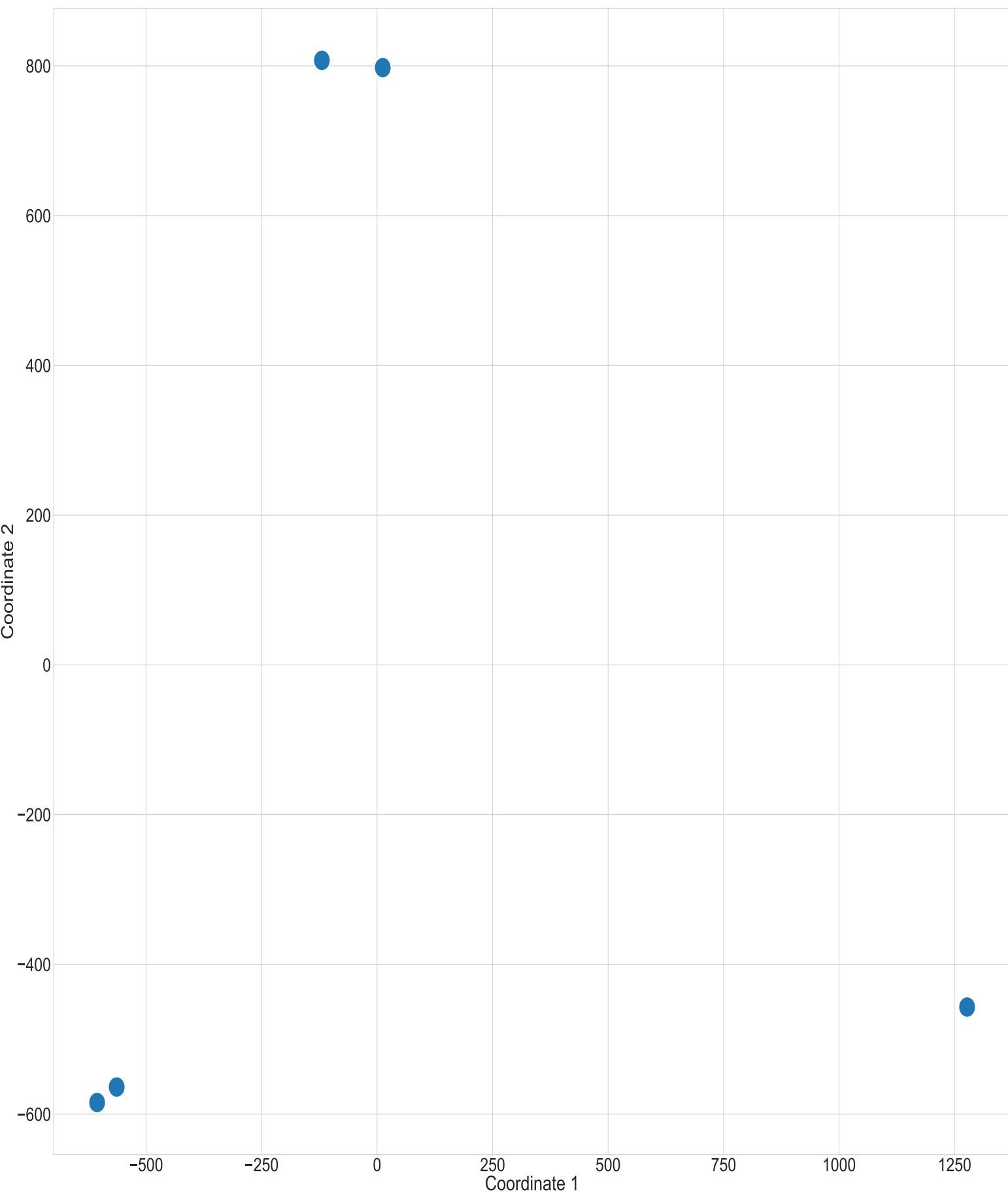
Scatter plot of Saccharides *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cosine pairwise distance. Only Saccharides BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Saccharides BGCs is associated with the GCFs in BiG_SLiCE they were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

Figure – S87 :



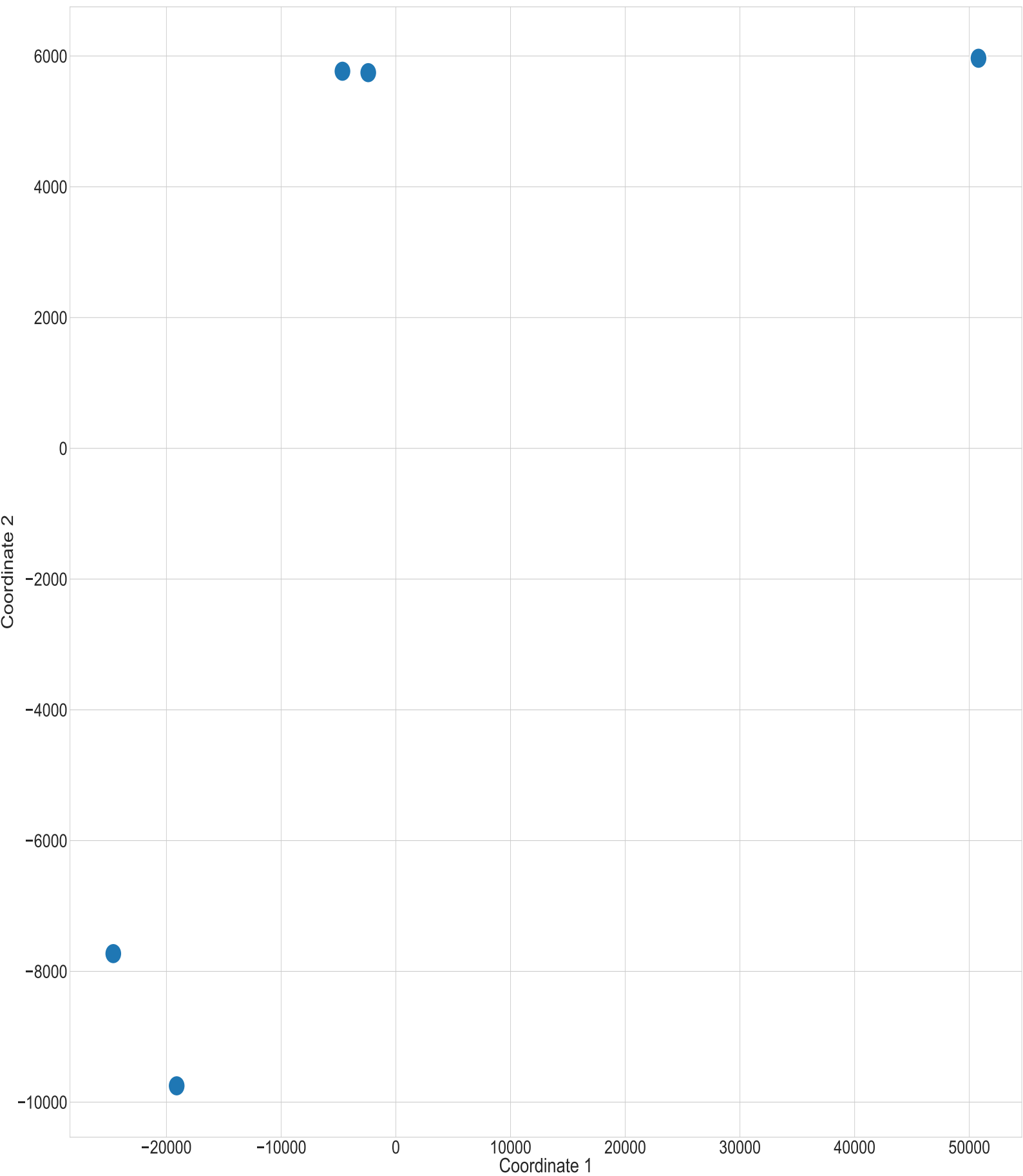
Scatter plot of Saccharides *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cityblock pairwise distance. Only Saccharides BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Saccharides BGCs is associated with the GCFs in BiG_SLiCE they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

Figure – S88 :



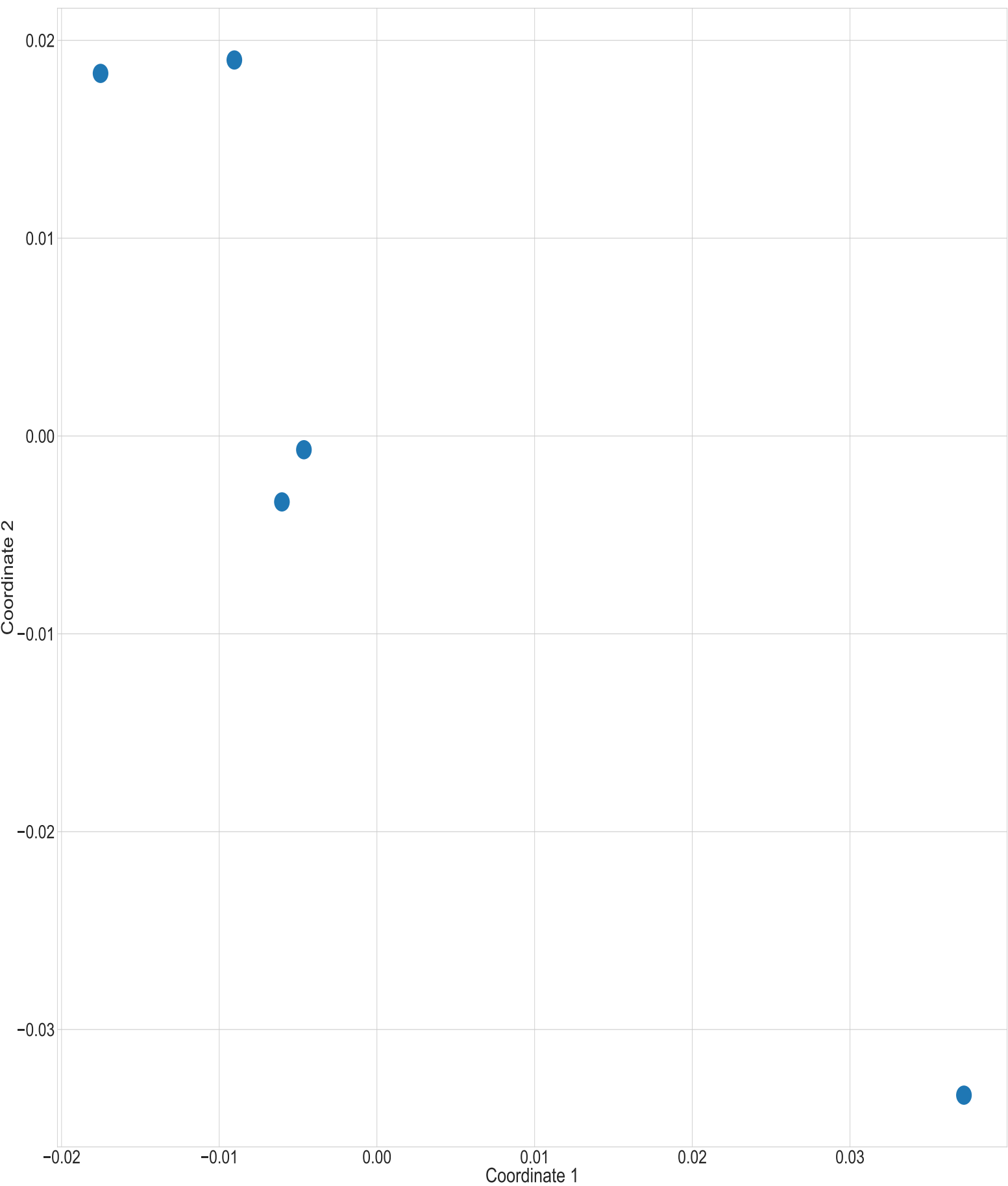
Scatter plot of Saccharides *Micromonosporaceae* BGCs analyzed via multidimensional scaling using chebyshev pairwise distance. Only Saccharides BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Saccharides BGCs is associated with the GCFs in BiG_SLICE they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

Figure – S89 :



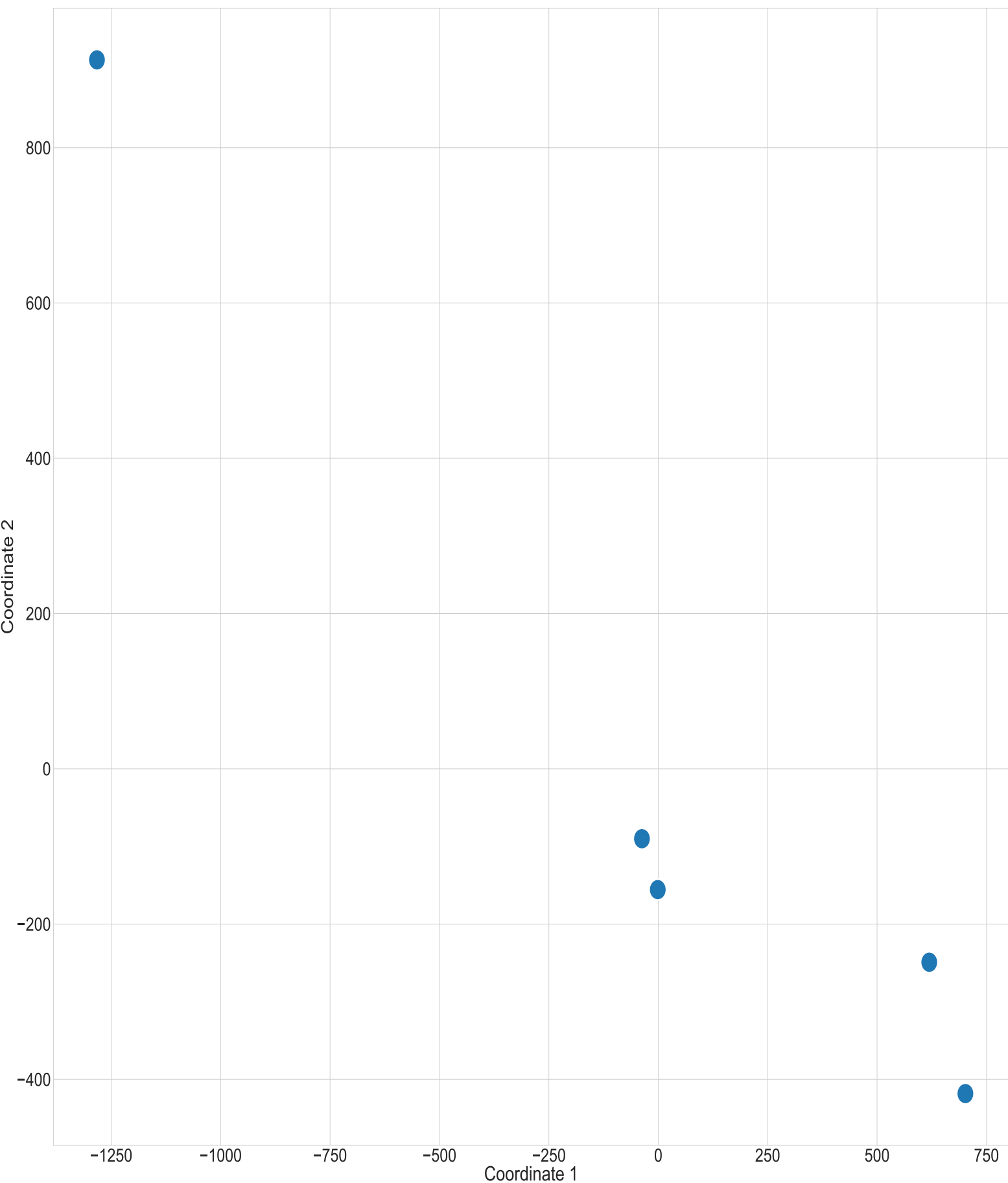
Scatter plot of Saccharides *Micromonosporaceae* BGCs analyzed via multidimensional scaling using I2 pairwise distance. Only Saccharides BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Saccharides BGCs is associated with the GCFs in BiG_SLiCE they were most similar to (T=900).

Figure – S90 :



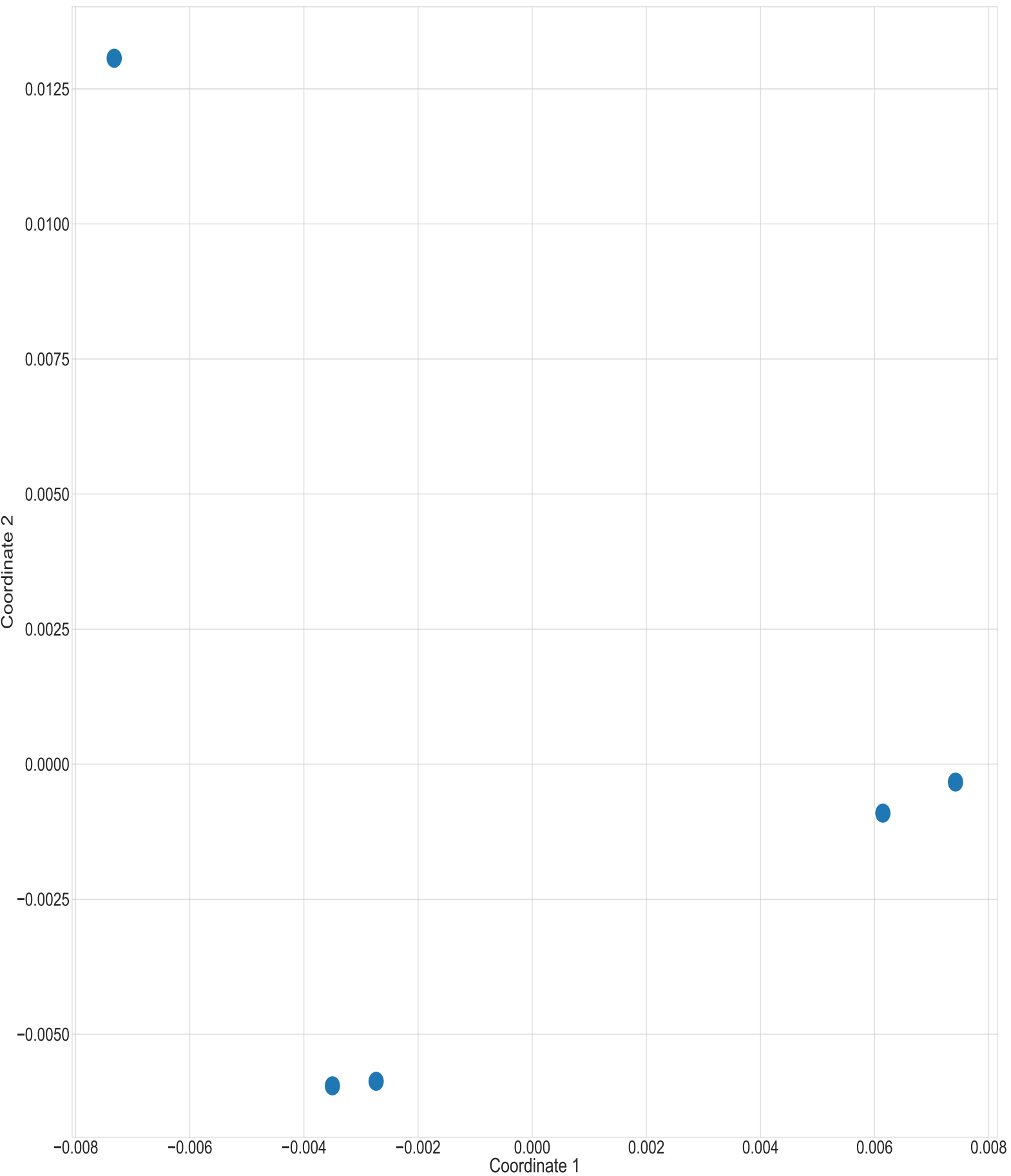
Scatter plot of Saccharides *Micromonosporaceae* BGCs analyzed via multidimensional scaling using braycurtis pairwise distance. Only Saccharides BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Saccharides BGCs is associated with the GCFs in BiG_SLiCE they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

Figure – S91 :



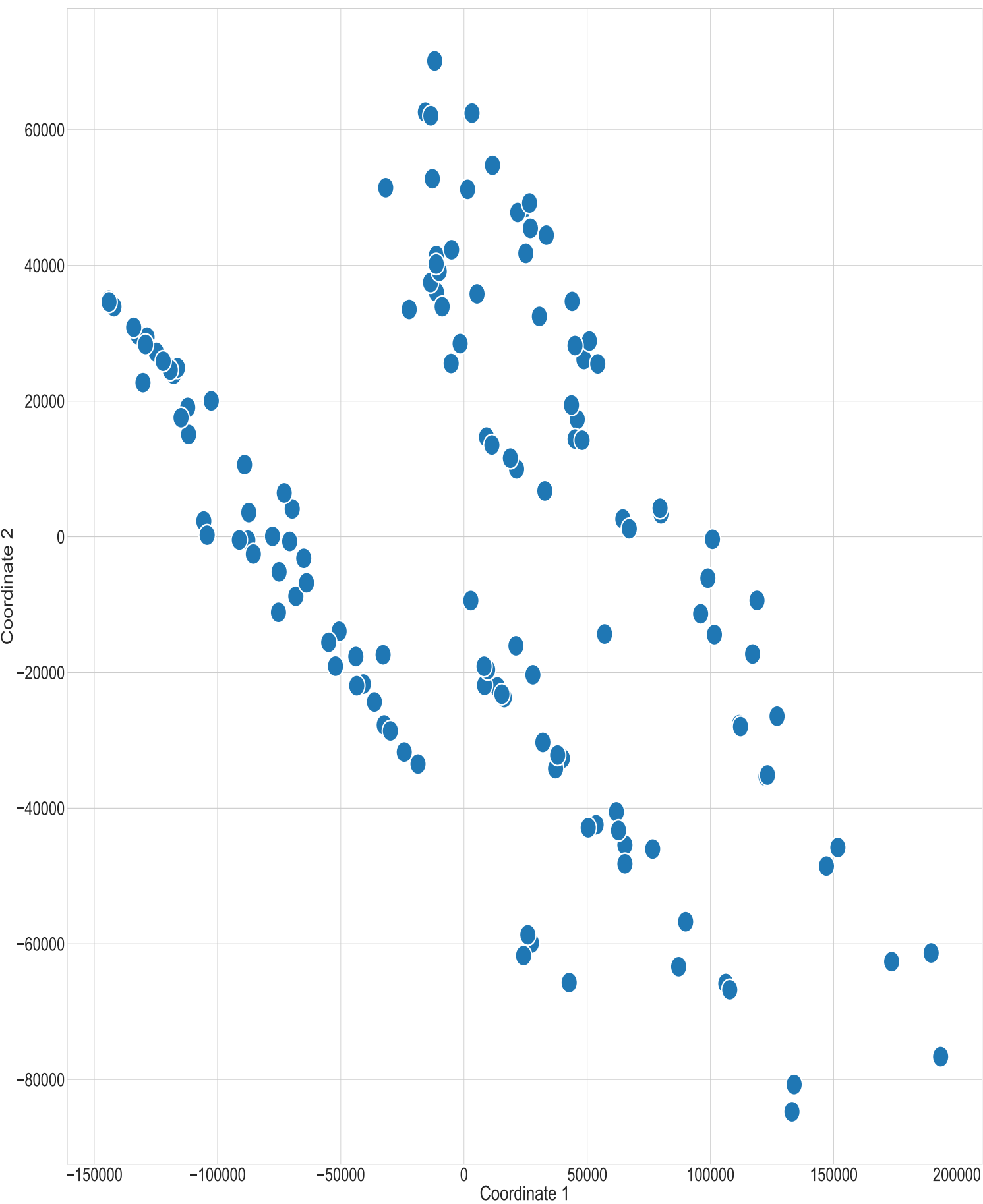
Scatter plot of Saccharides *Micromonosporaceae* BGCs analyzed via multidimensional scaling using canberra pairwise distance. Only Saccharides BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Saccharides BGCs is associated with the GCFs in BiG_SLiCE they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

Figure – S92 :



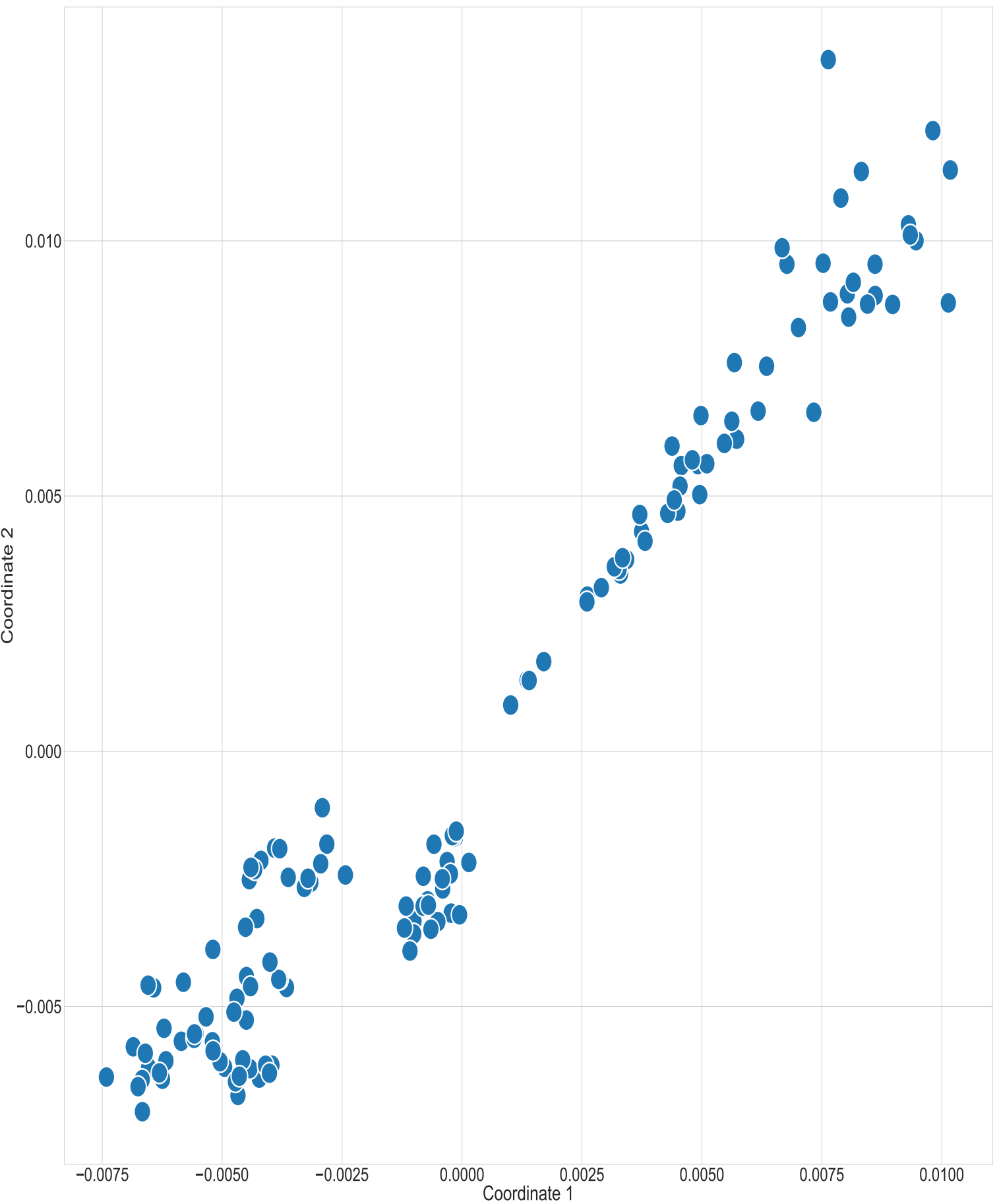
Scatter plot of Saccharides *Micromonosporaceae* BGCs analyzed via multidimensional scaling using correlation pairwise distance. Only Saccharides BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Saccharides BGCs is associated with the GCFs in BiG_SLiCE they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

Figure – S93 :



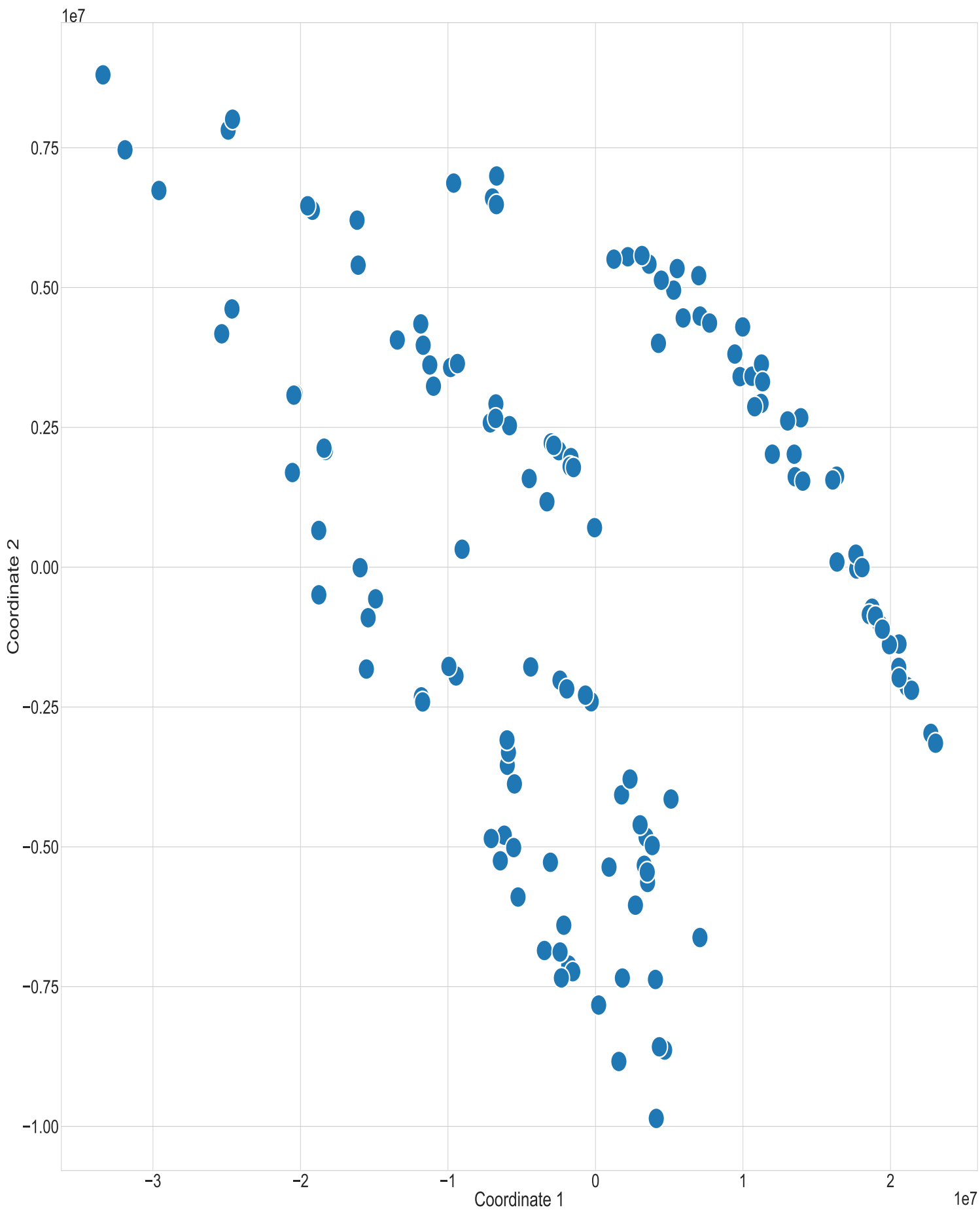
Scatter plot of Hybrid-(Not-PKS-NRPS) *Micromonosporaceae* BGCs analyzed via multidimensional scaling using euclidean pairwise distance. Only Hybrid-(Not-PKS-NRPS) BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Hybrid-(Not-PKS-NRPS) BGCs is associated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

Figure – S94 :



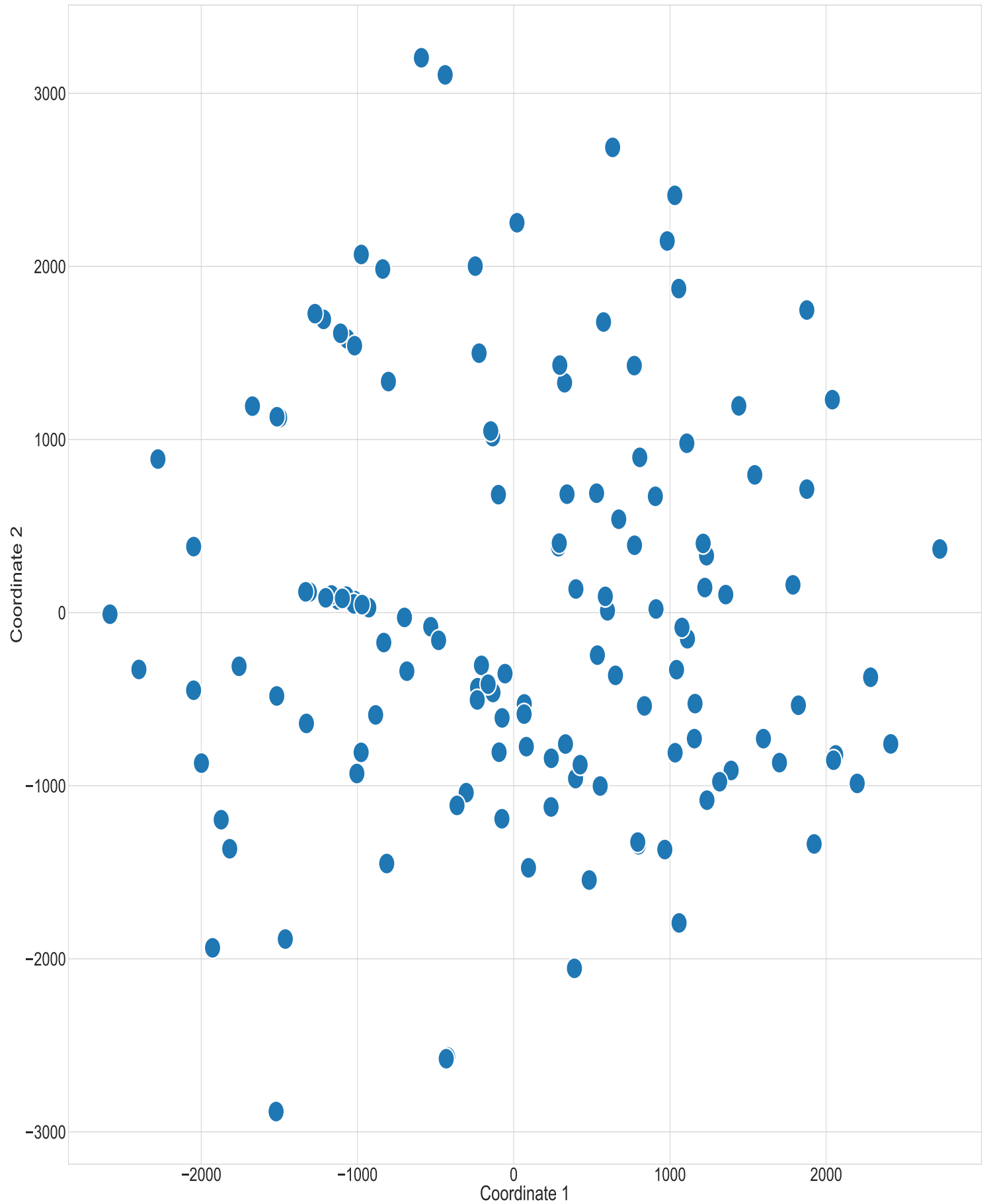
Scatter plot of Hybrid-(Not-PKS-NRPS) *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cosine pairwise distance. Only Hybrid-(Not-PKS-NRPS) BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Hybrid-(Not-PKS-NRPS) BGCs is associated with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

Figure – S95 :



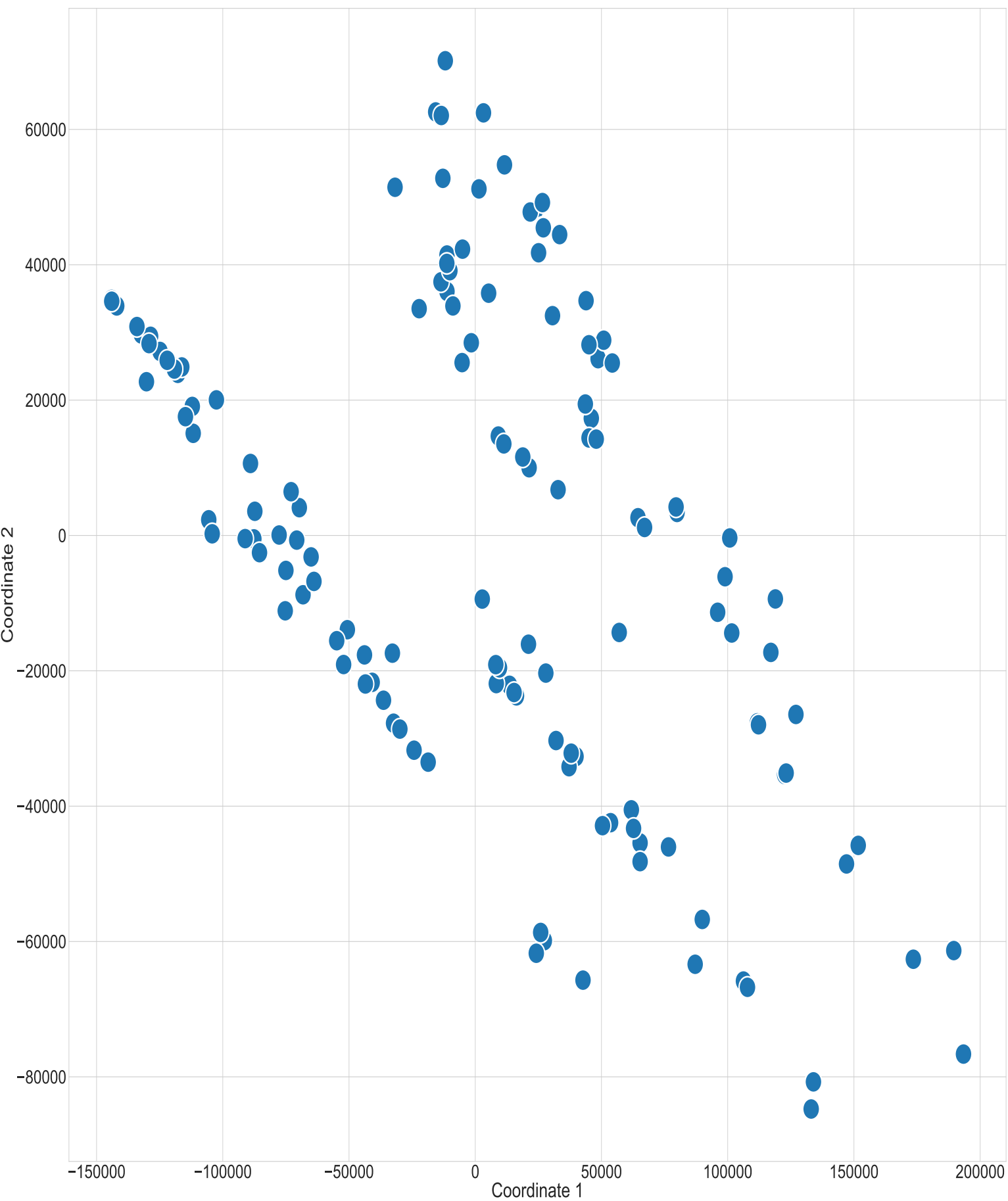
Scatter plot of Hybrid-(Not-PKS-NRPS) *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cityblock pairwise distance. Only Hybrid-(Not-PKS-NRPS) BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Hybrid-(Not-PKS-NRPS) BGCs is associated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

Figure – S96 :



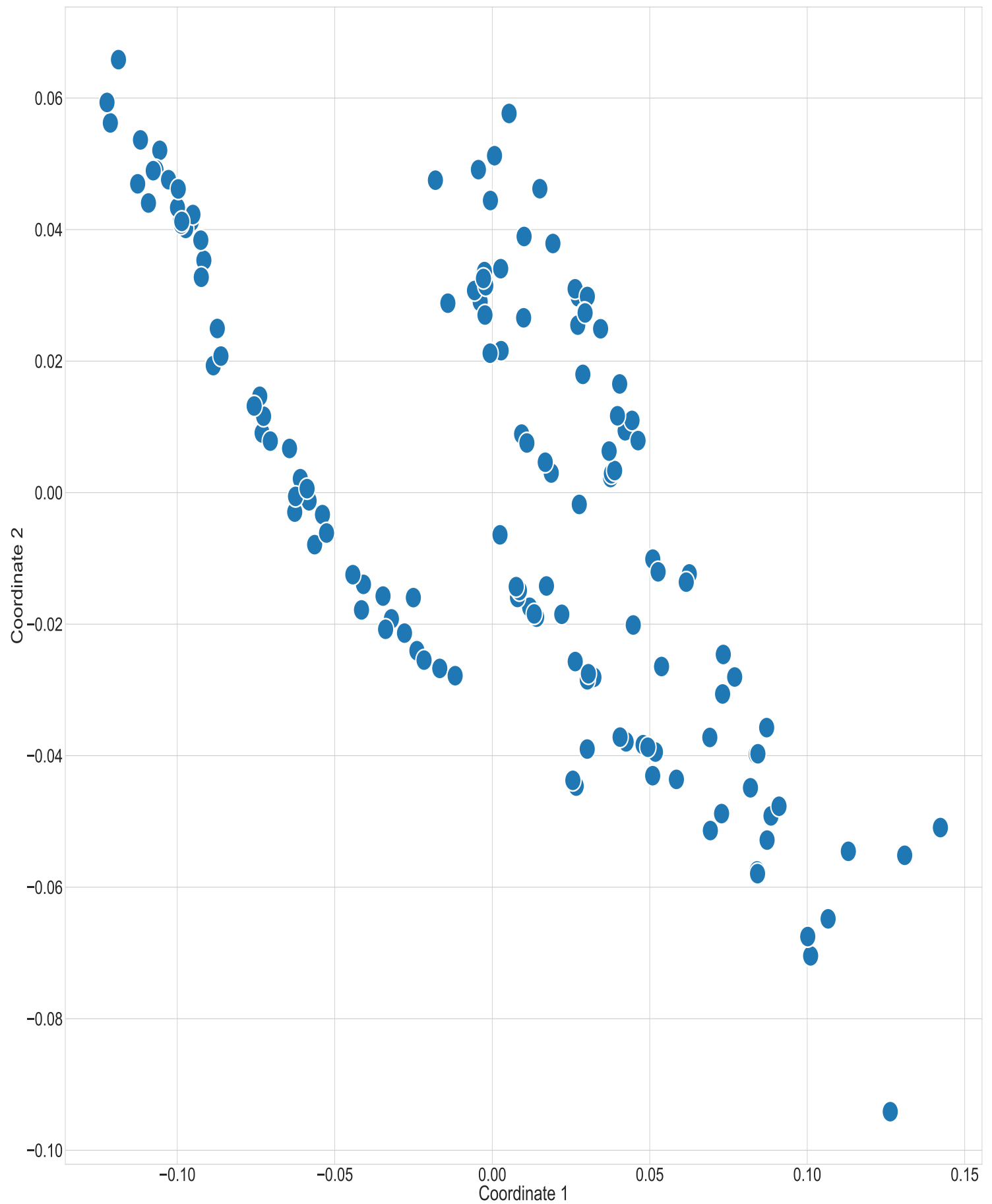
Scatter plot of Hybrid-(Not-PKS-NRPS) *Micromonosporaceae* BGCs analyzed via multidimensional scaling using chebyshev pairwise distance. Only Hybrid-(Not-PKS-NRPS) BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Hybrid-(Not-PKS-NRPS) BGCs is associated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

Figure – S97 :



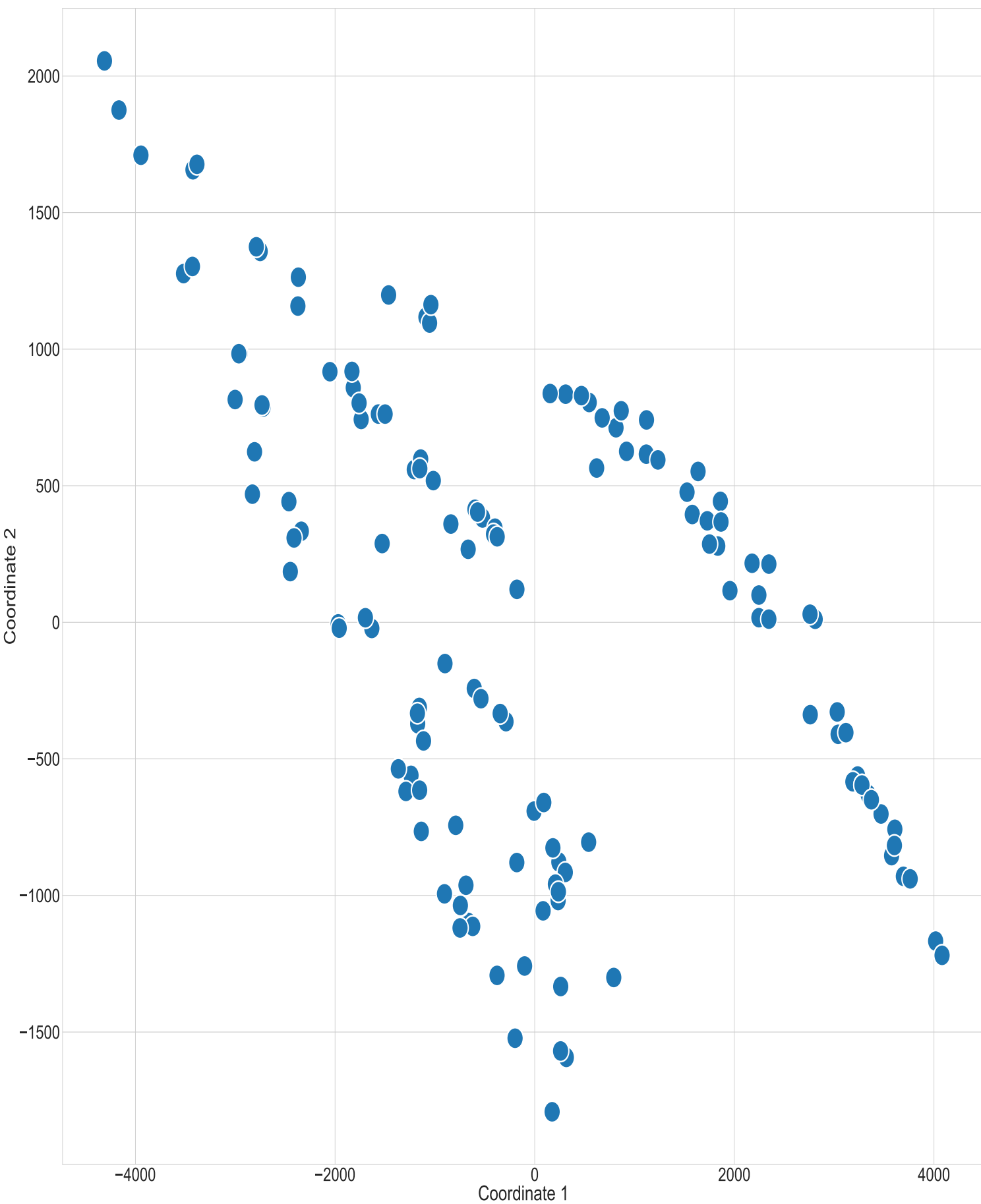
Scatter plot of Hybrid-(Not-PKS-NRPS) *Micromonosporaceae* BGCs analyzed via multidimensional scaling using l2 pairwise distance. Only Hybrid-(Not-PKS-NRPS) BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Hybrid-(Not-PKS-NRPS) BGCs is associated with the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

Figure – S98 :



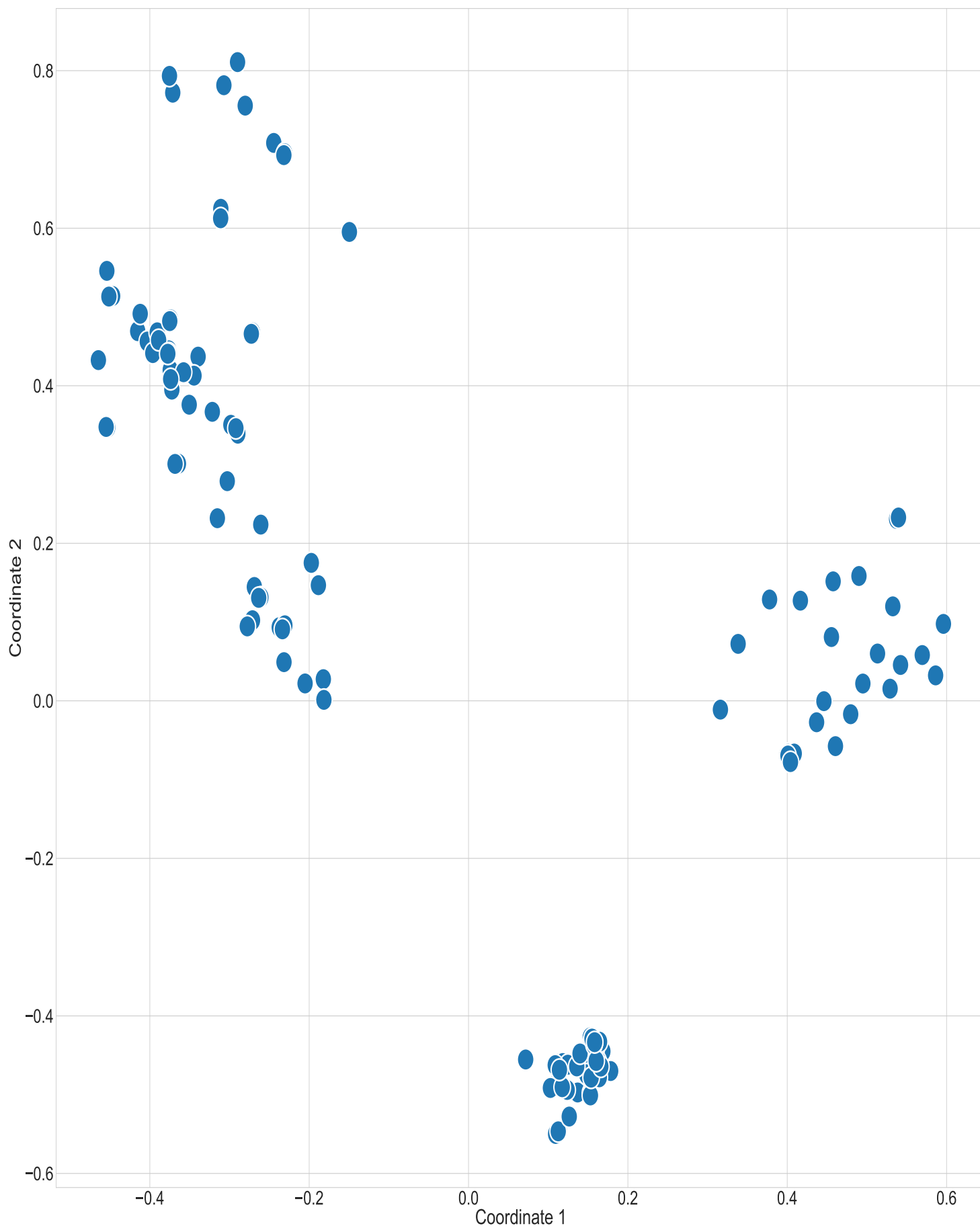
Scatter plot of Hybrid-(Not-PKS-NRPS) *Micromonosporaceae* BGCs analyzed via multidimensional scaling using braycurtis pairwise distance. Only Hybrid-(Not-PKS-NRPS) BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Hybrid-(Not-PKS-NRPS) BGCs is associated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

Figure – S99 :



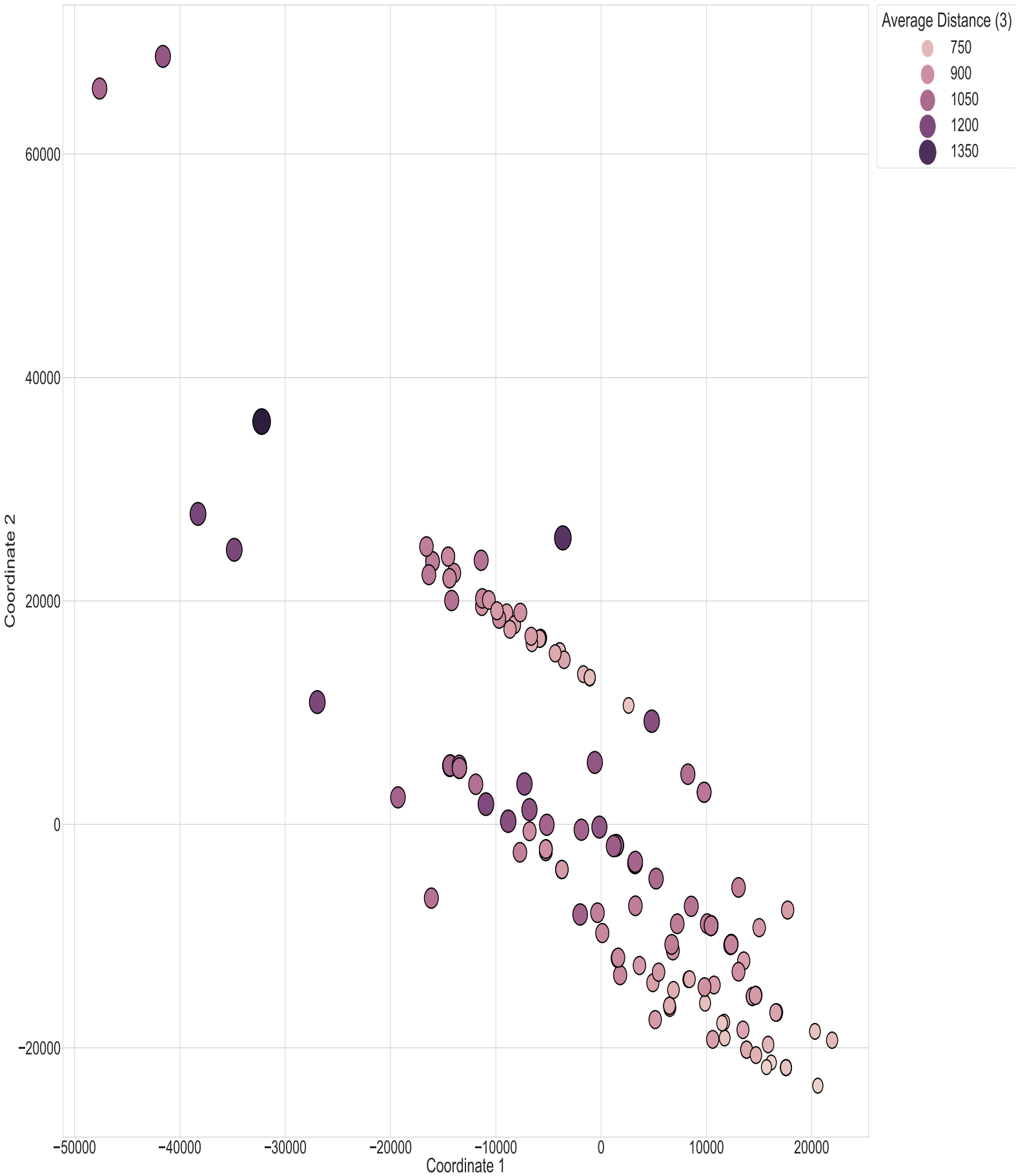
Scatter plot of Hybrid-(Not-PKS-NRPS) *Micromonosporaceae* BGCs analyzed via multidimensional scaling using canberra pairwise distance. Only Hybrid-(Not-PKS-NRPS) BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Hybrid-(Not-PKS-NRPS) BGCs is associated with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

Figure – S100 :



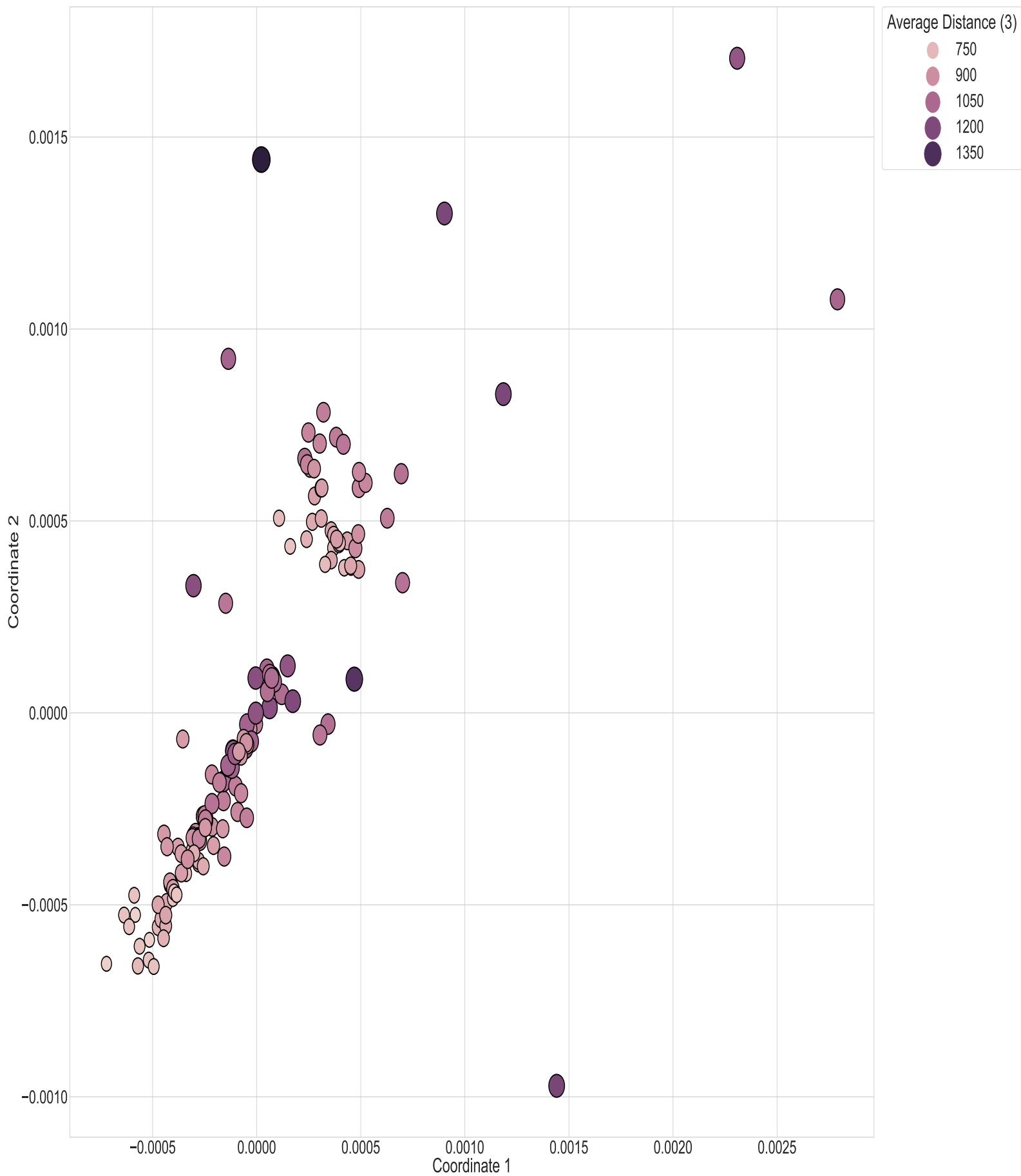
Scatter plot of Hybrid-(Not-PKS-NRPS) *Micromonosporaceae* BGCs analyzed via multidimensional scaling using correlation pairwise distance. Only Hybrid-(Not-PKS-NRPS) BGCs were considered for this analysis. Each dot represents an individual BGC. Distance between Hybrid-(Not-PKS-NRPS) BGCs is associated 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 :

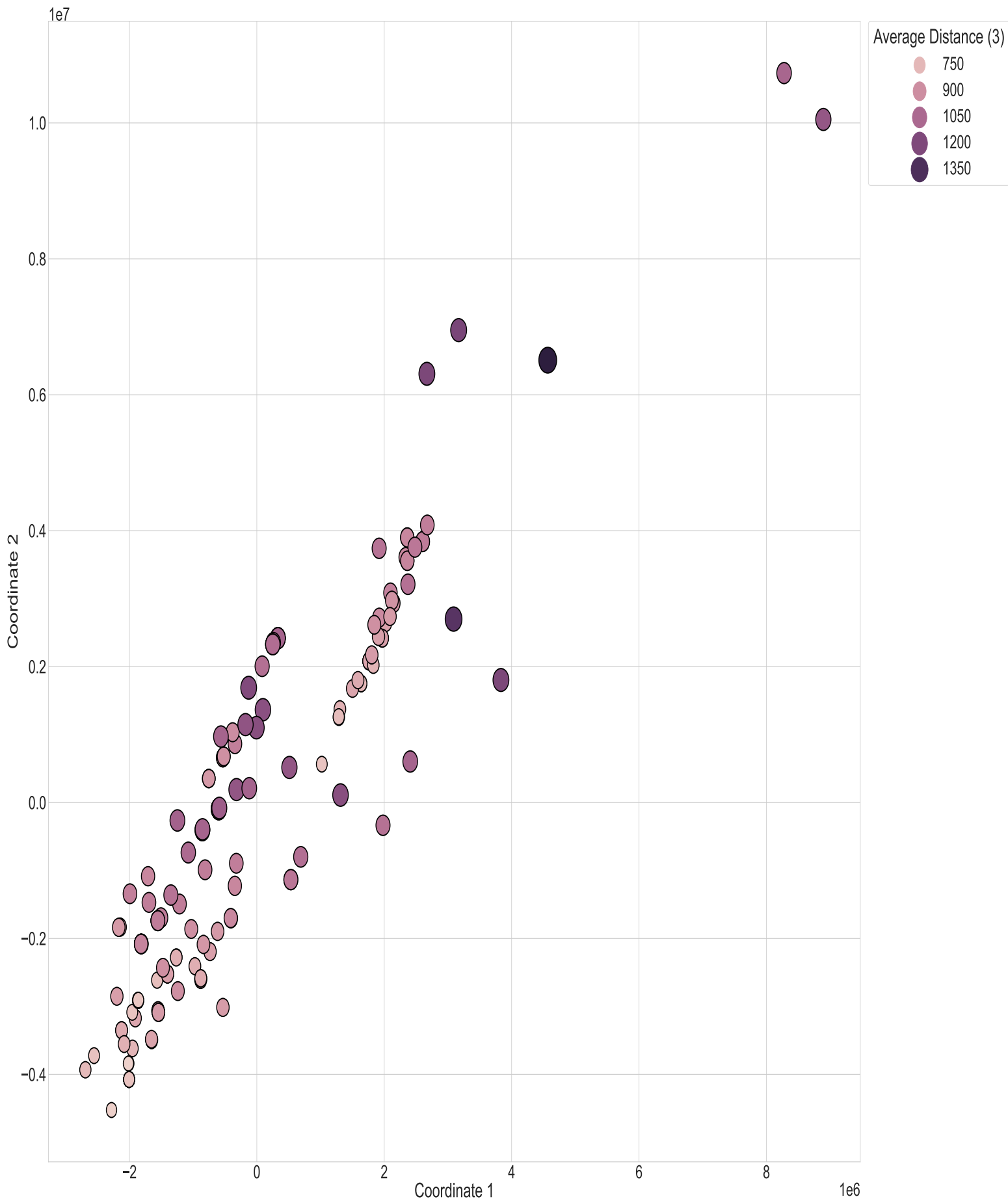


Scatter plot of Terpene *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cosine pairwise distance.

Only Terpene BGCs were considered for this analysis. Each dot represents an individual BGC.

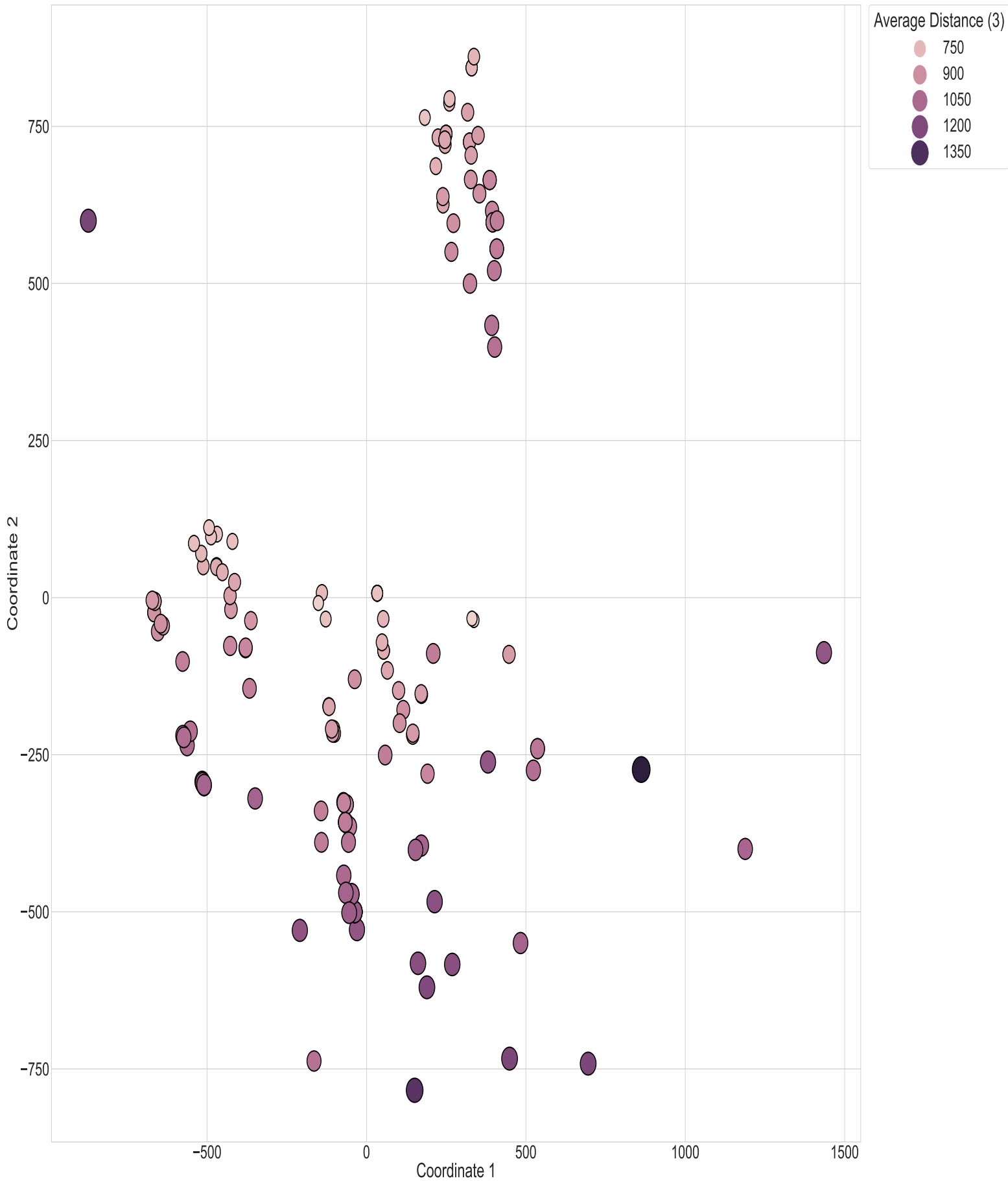
The hue and size of the dots were scaled based on the average distance of the individual Terpene BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

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 :

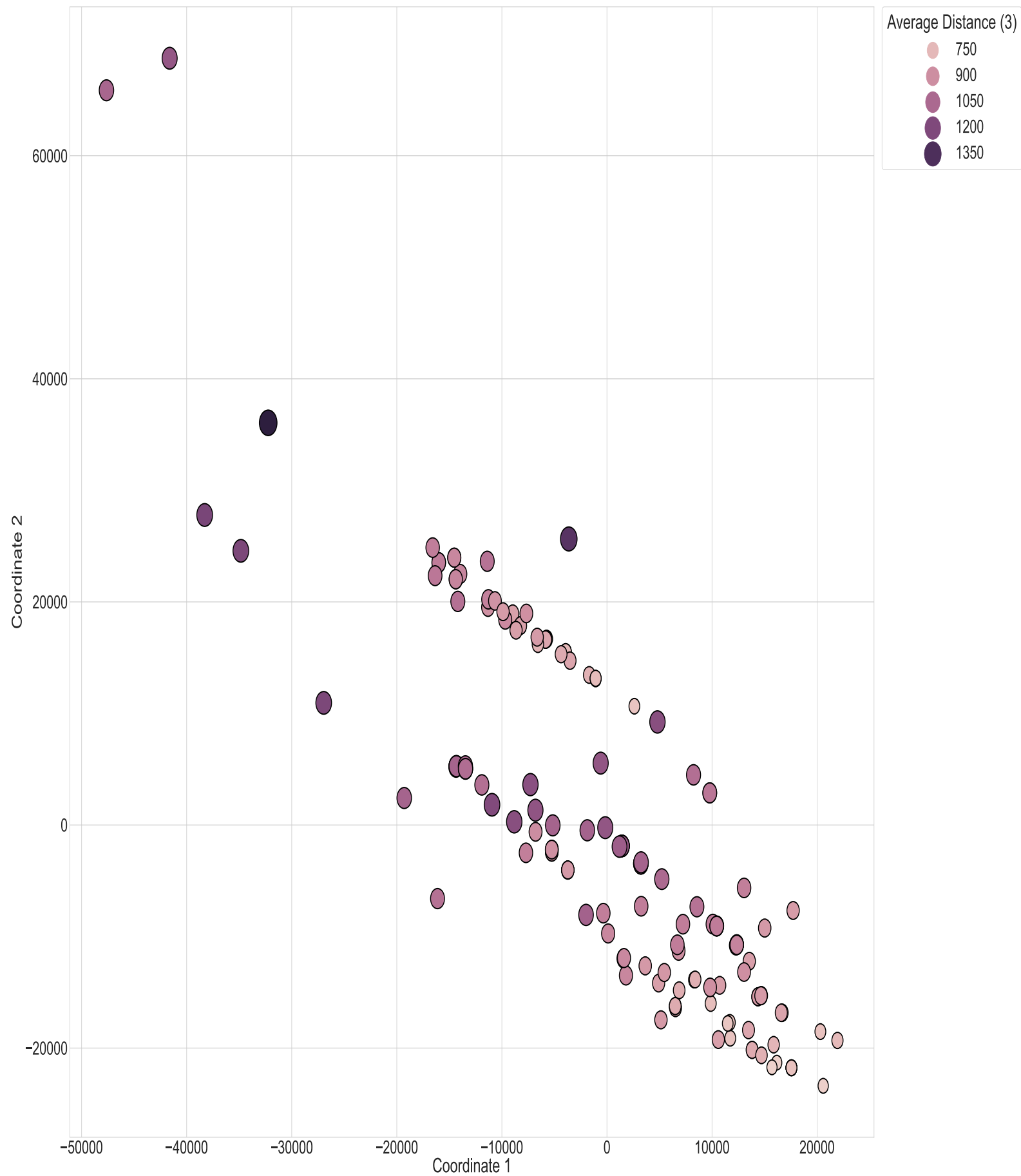


Scatter plot of Terpene *Micromonosporaceae* BGCs analyzed via multidimensional scaling using chebyshev pairwise distance.

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 :

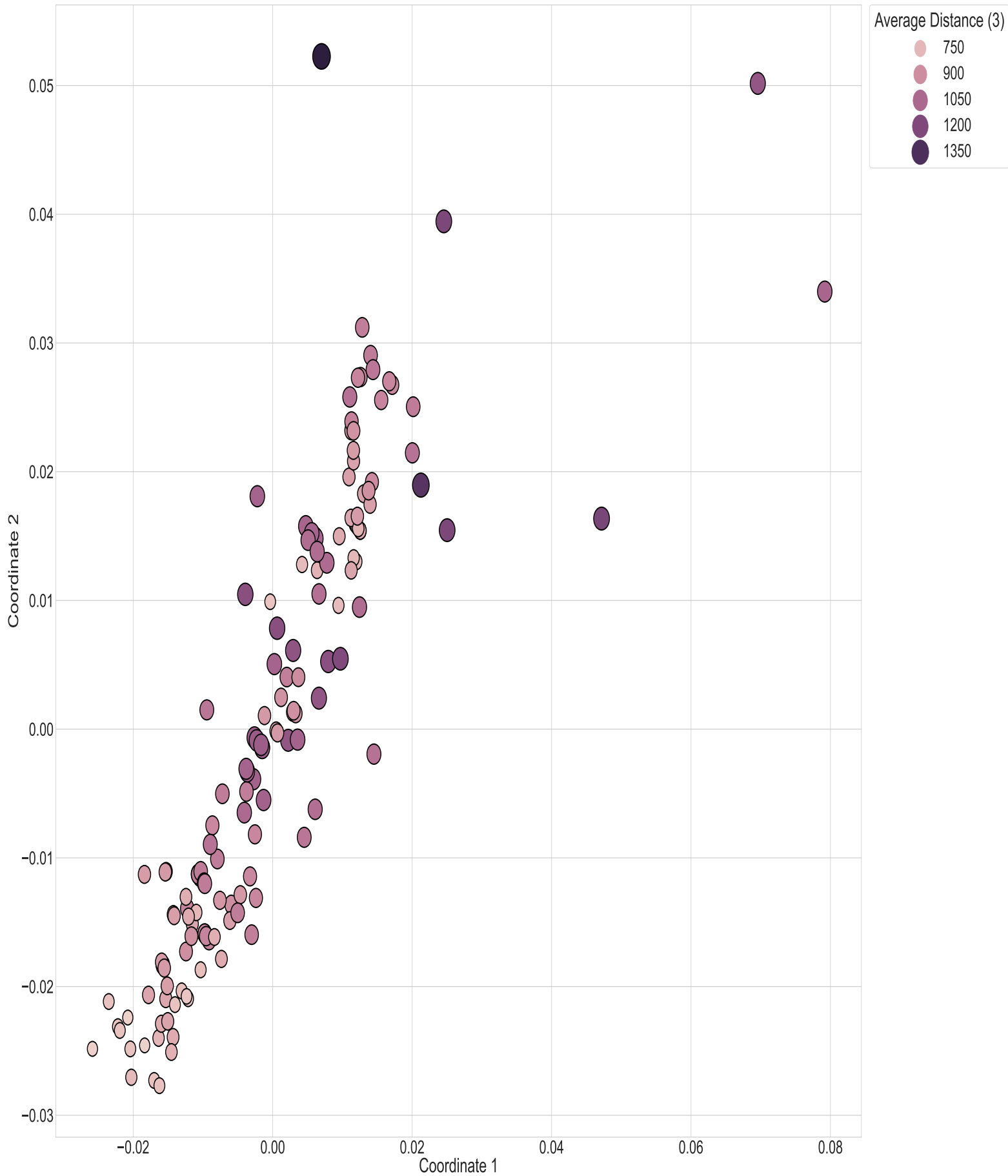


Scatter plot of Terpene *Micromonosporaceae* BGCs analyzed via multidimensional scaling using l2 pairwise distance.

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 :

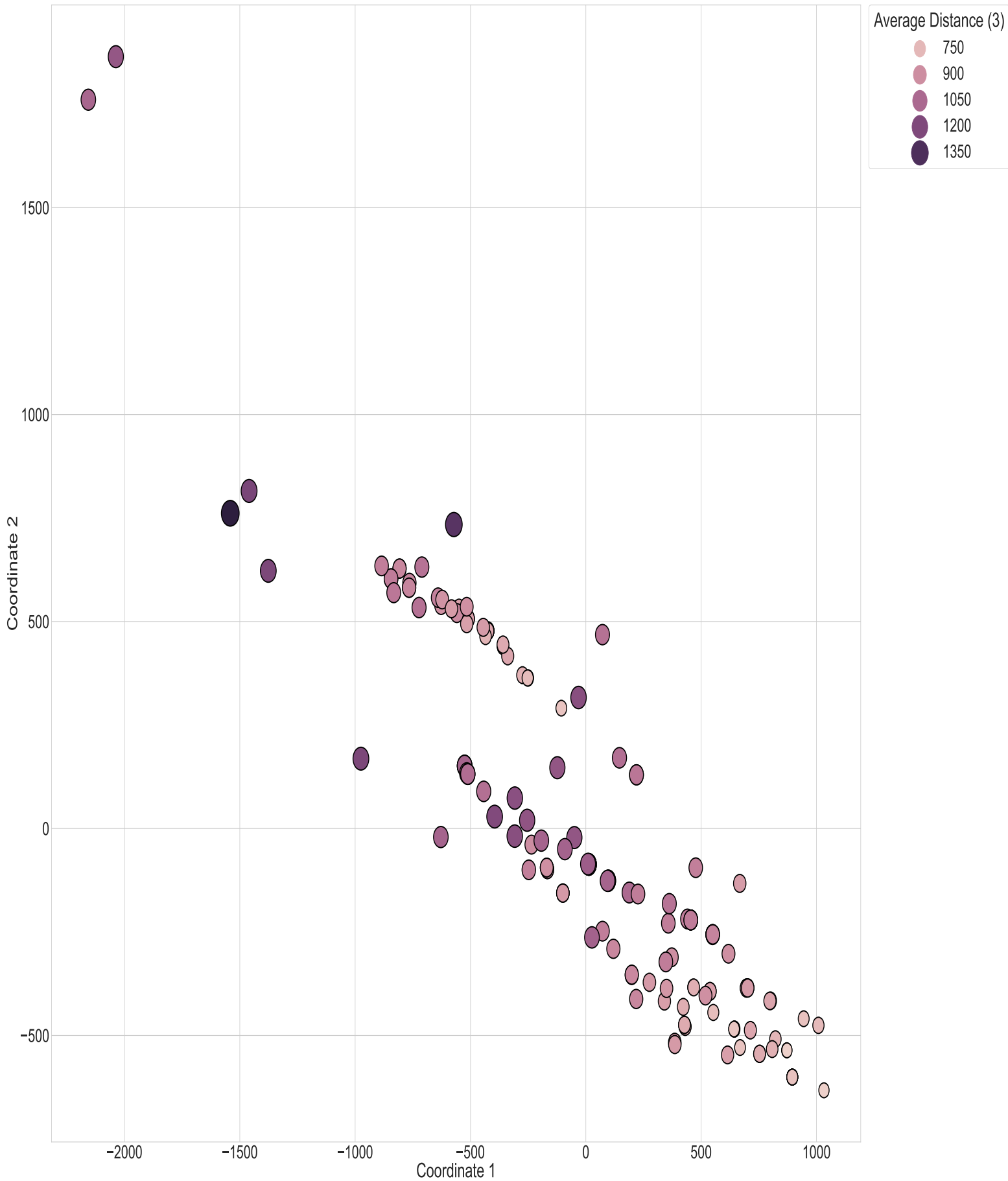


Scatter plot of Terpene *Micromonosporaceae* BGCs analyzed via multidimensional scaling using braycurtis pairwise distance.

Only Terpene BGCs were considered for this analysis. Each dot represents an individual BGC.

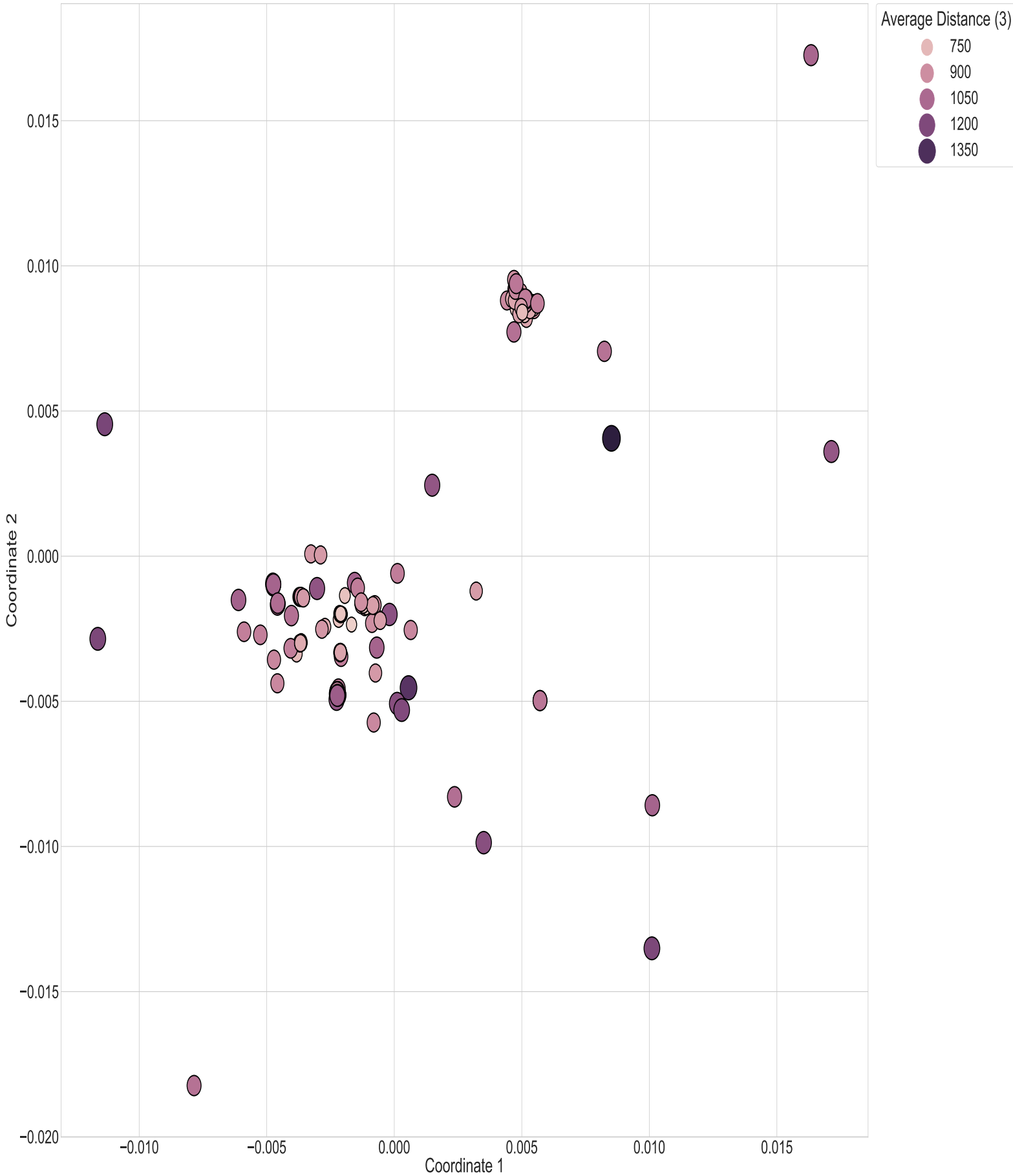
The hue and size of the dots were scaled based on the average distance of the individual Terpene BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

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 :

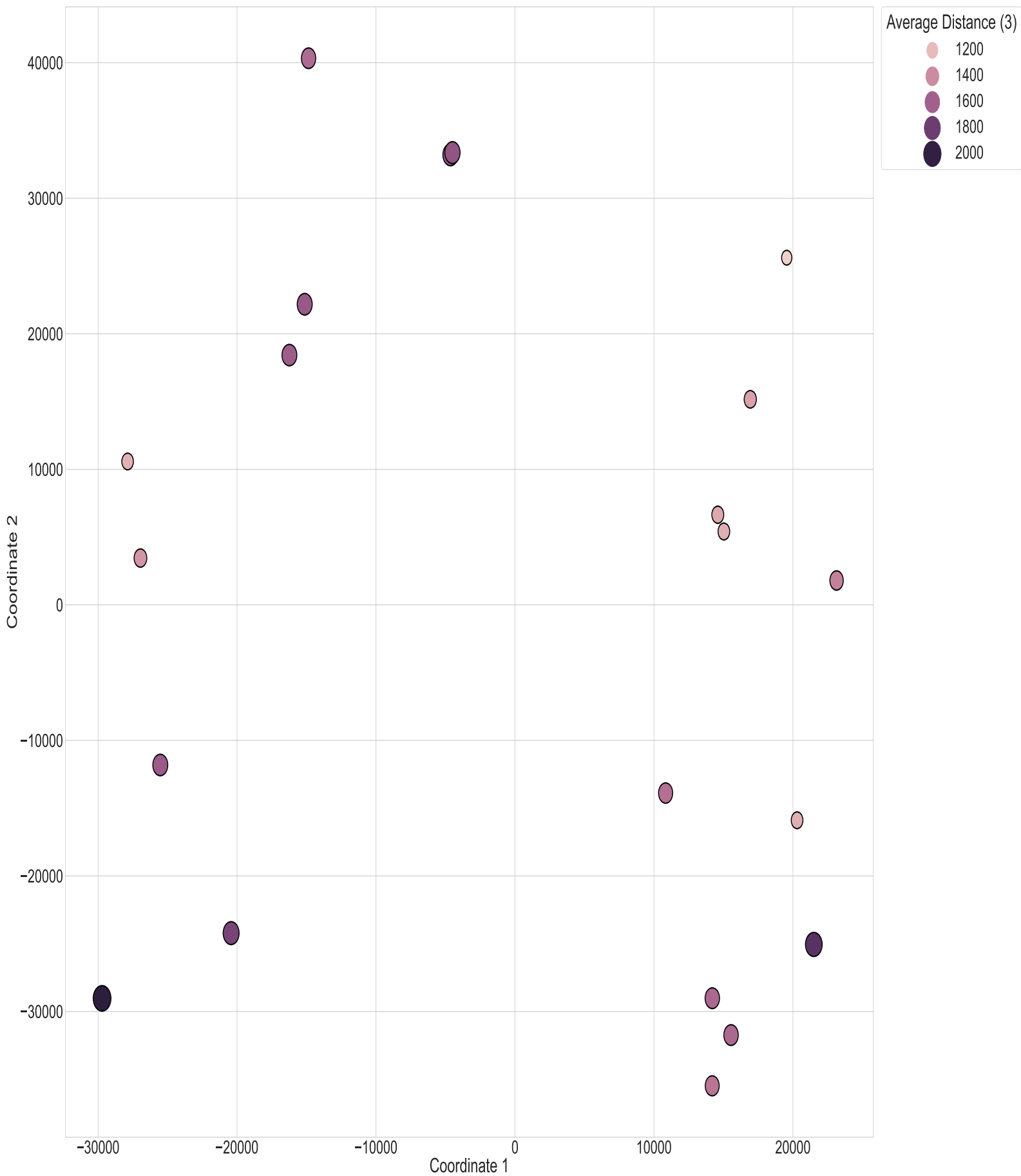


Scatter plot of Terpene *Micromonosporaceae* BGCs analyzed via multidimensional scaling using correlation pairwise distance.

Only Terpene BGCs were considered for this analysis. Each dot represents an individual BGC.

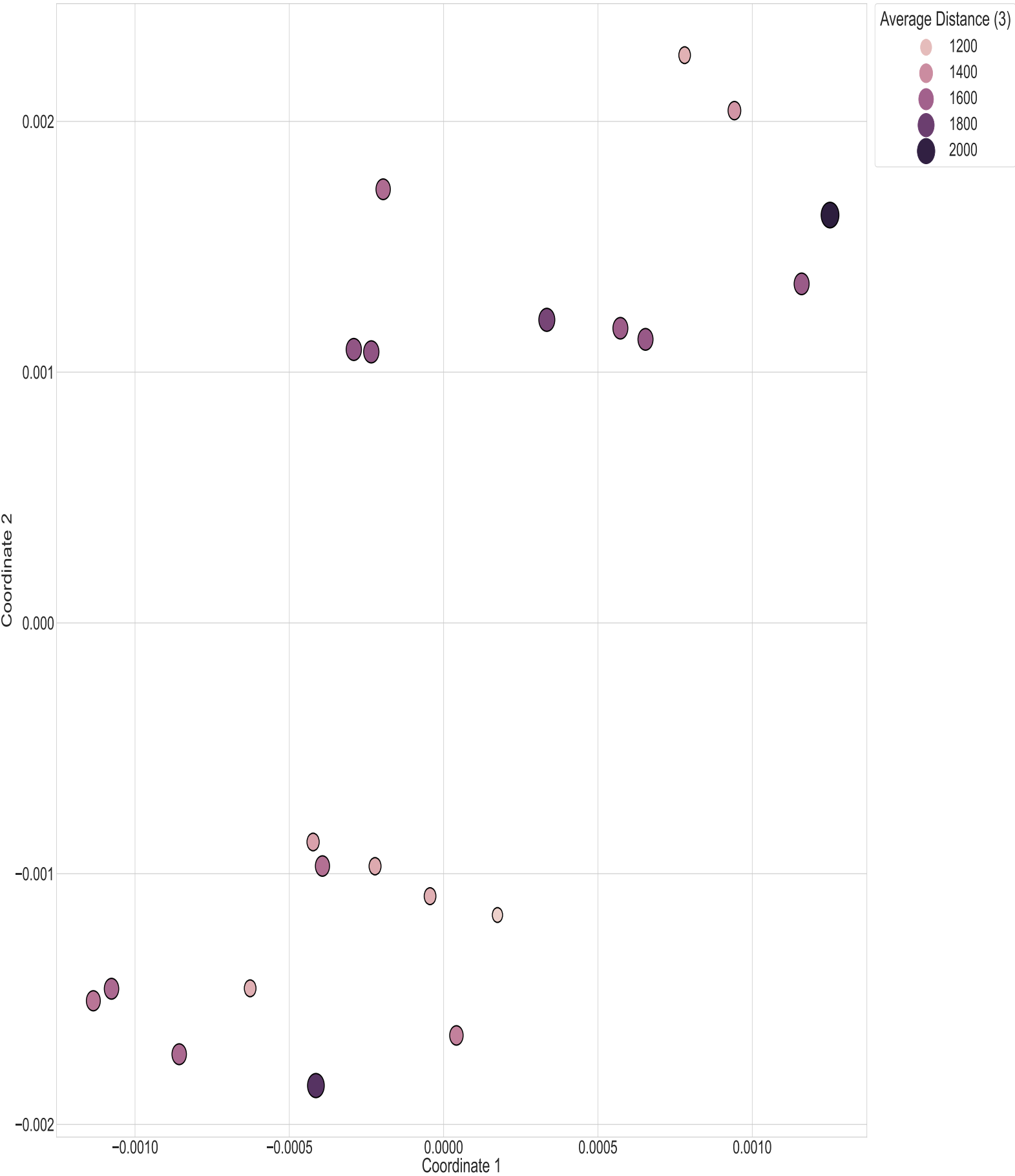
The hue and size of the dots were scaled based on the average distance of the individual Terpene BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

Figure – S109 :



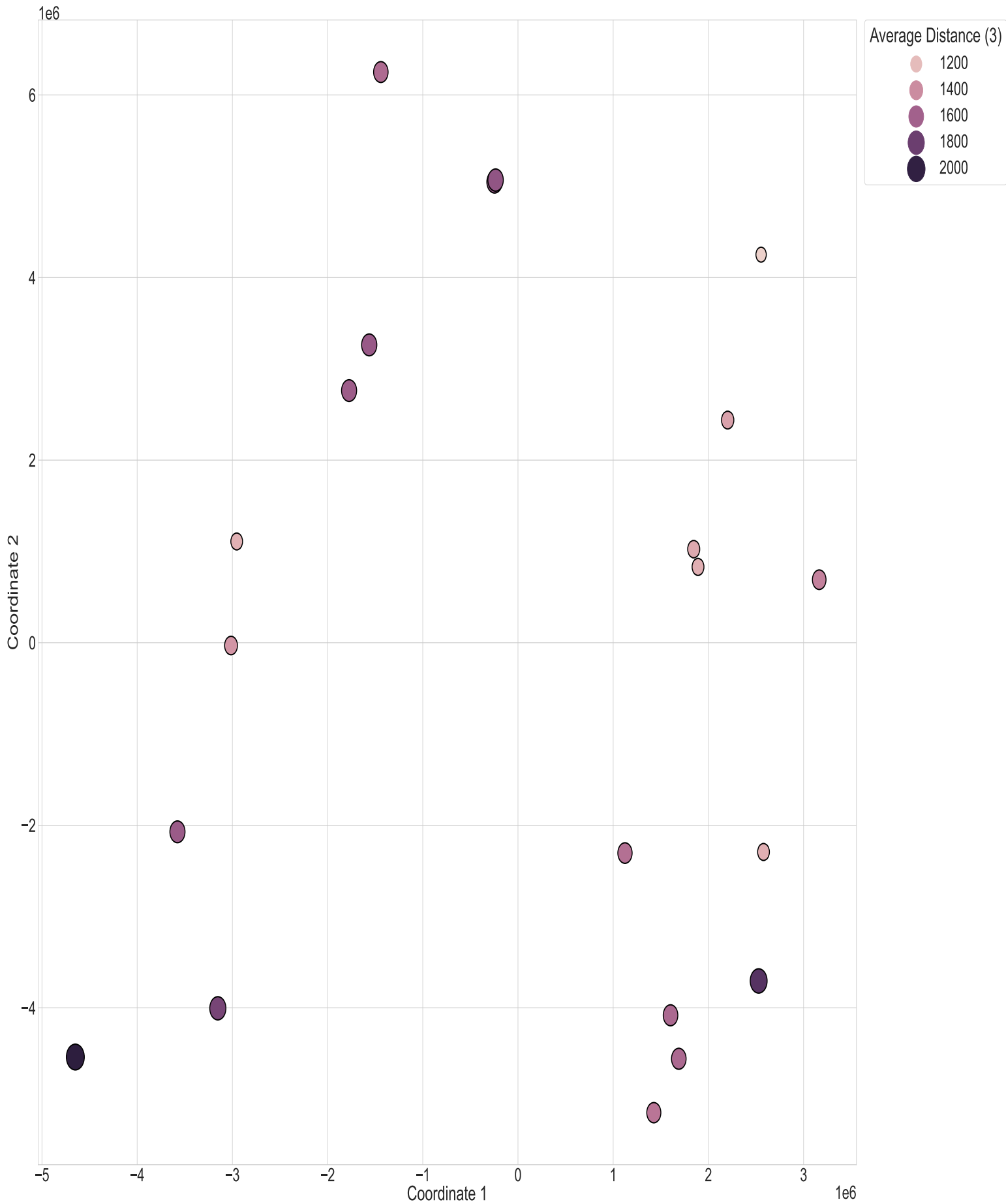
Scatter plot of PKS-1 *Micromonosporaceae* BGCs analyzed via multidimensional scaling using euclidean pairwise distance. Only PKS-1 BGCs were considered for this analysis. Each dot represents an individual BGC. The hue and size of the dots were scaled based on the average distance of the individual PKS-1 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.

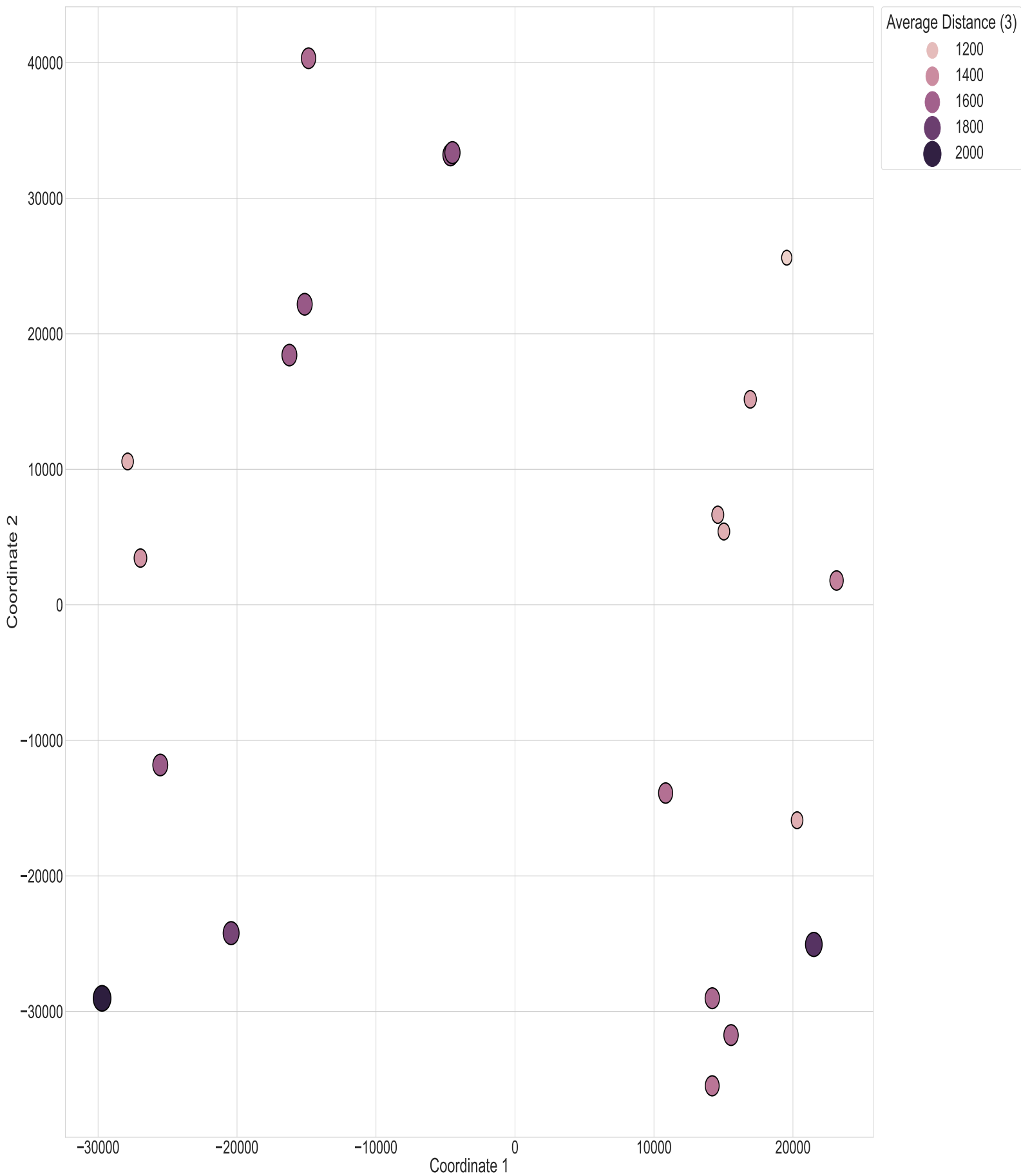
The hue and size of the dots were scaled based on the average distance of the individual PKS-1 BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

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.

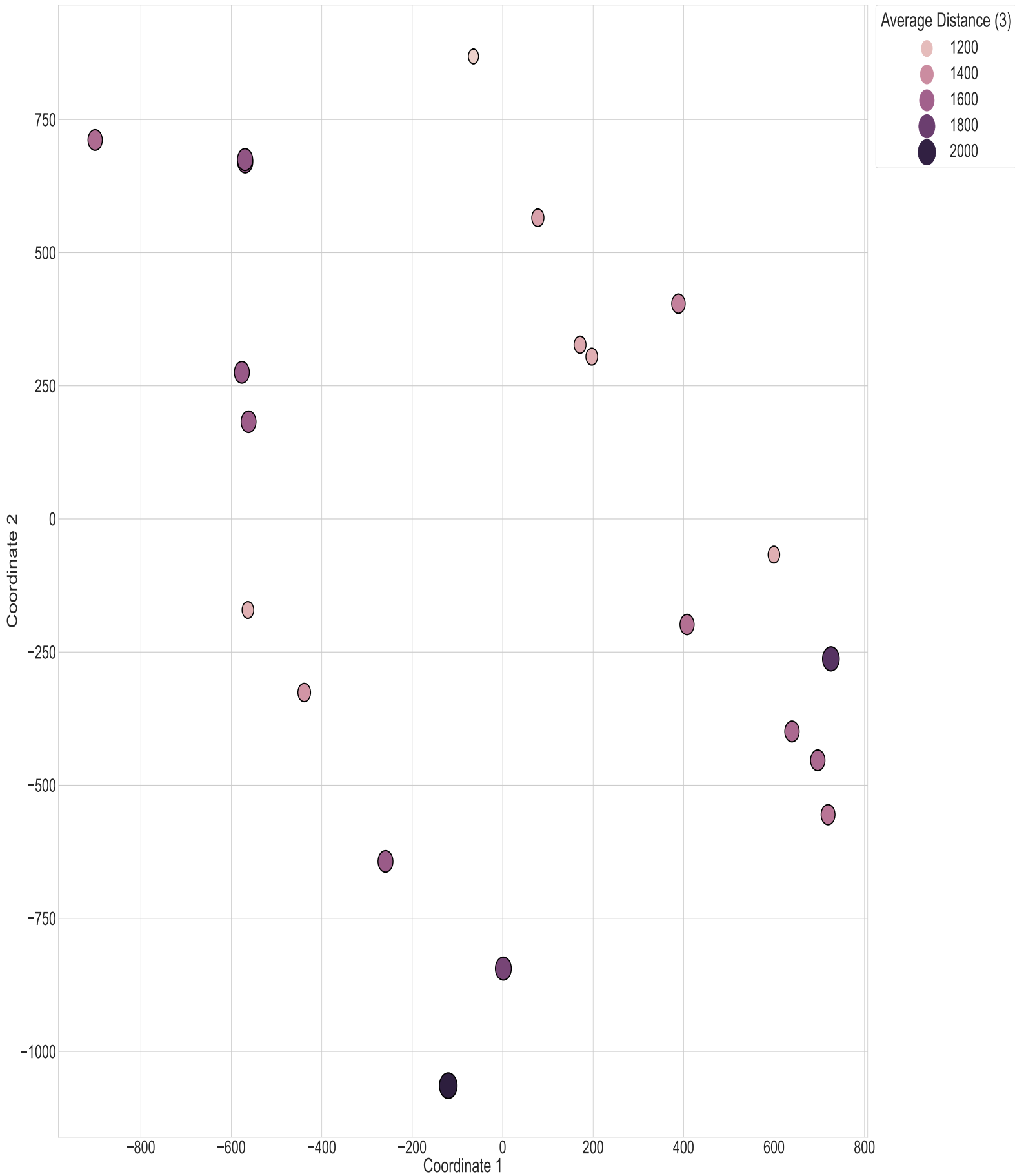
The hue and size of the dots were scaled based on the average distance of the individual PKS-1 BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

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 :

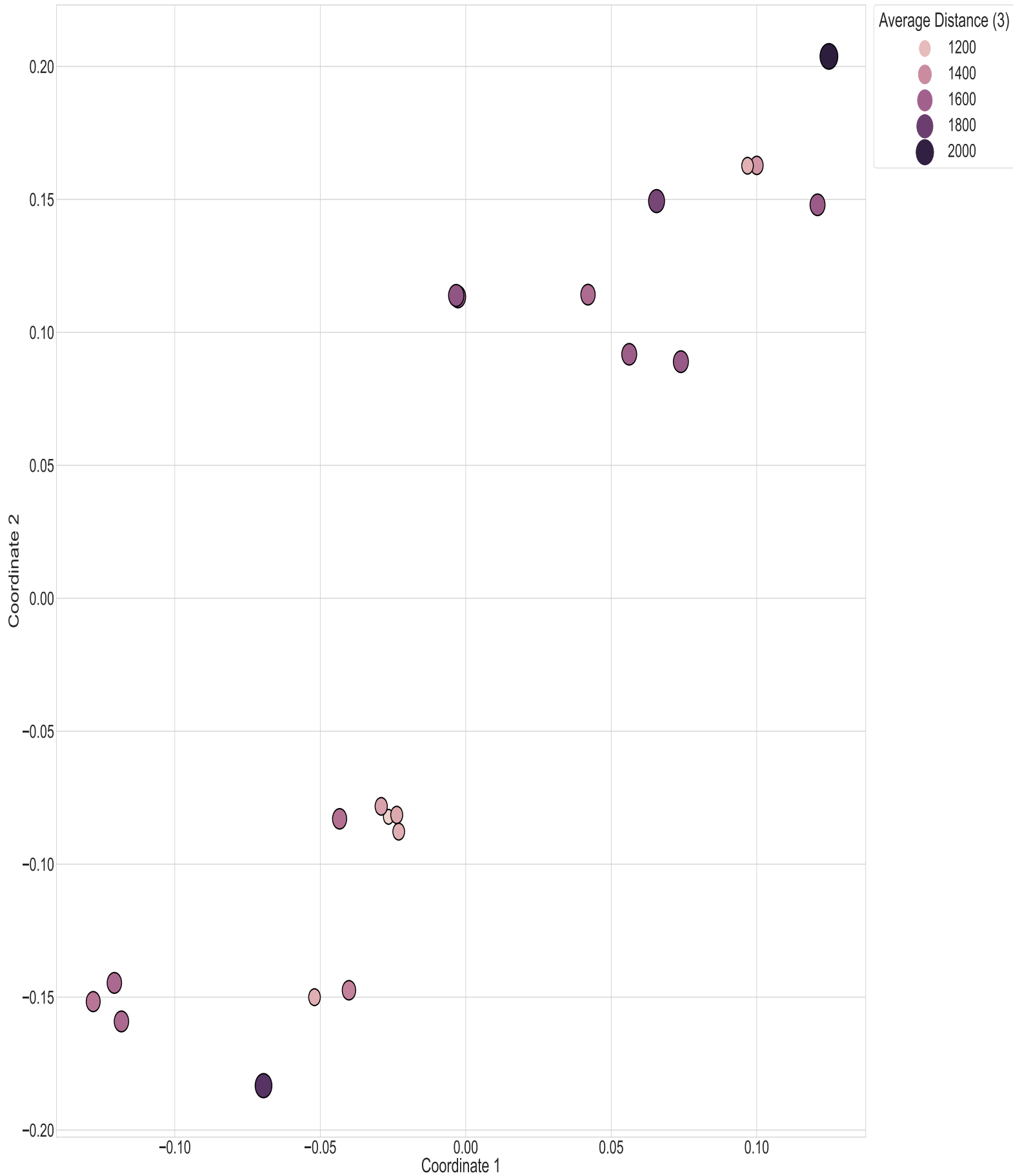


Scatter plot of PKS-1 *Micromonosporaceae* BGCs analyzed via multidimensional scaling using canberra pairwise distance.

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 :

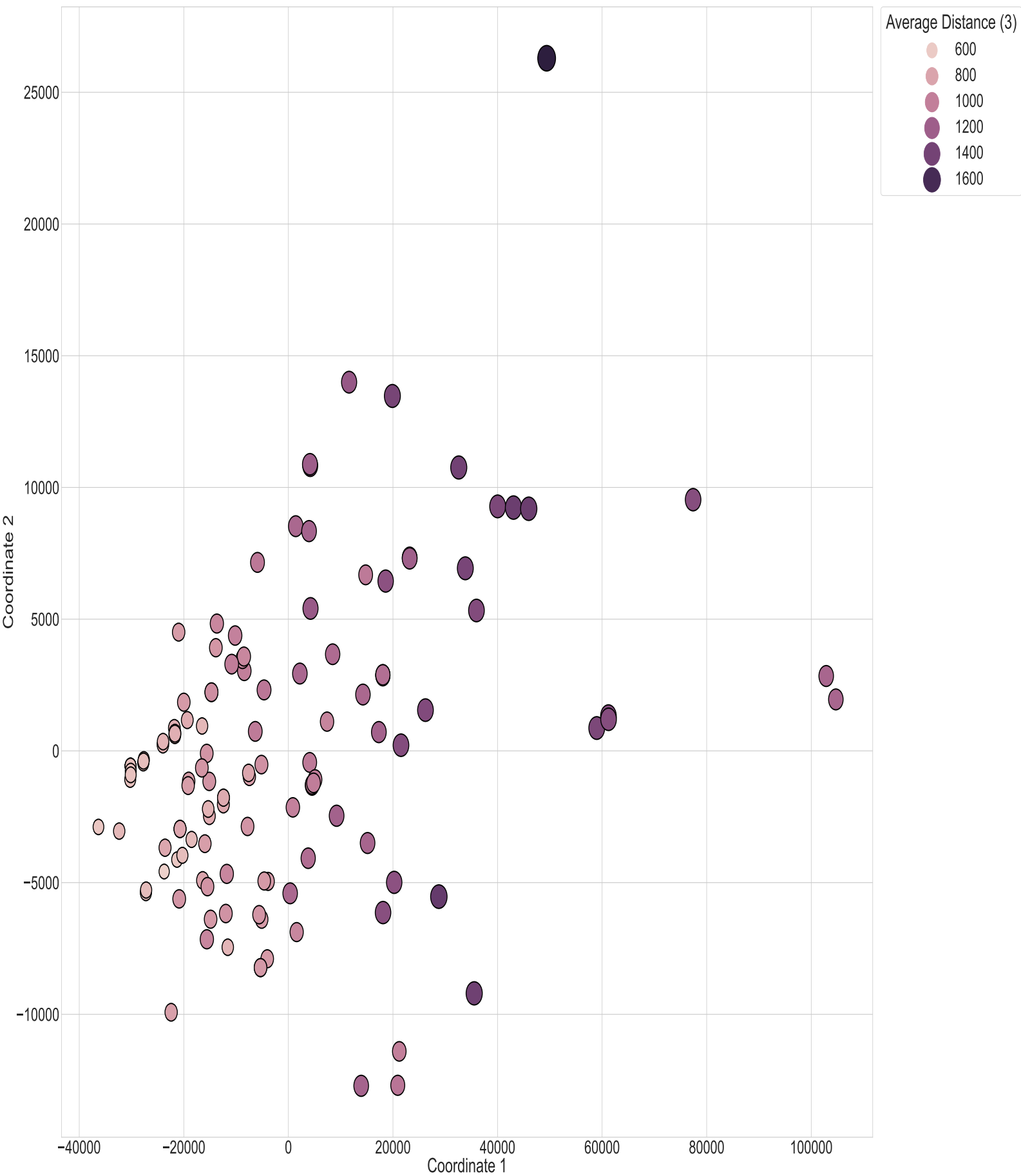


Scatter plot of PKS-1 *Micromonosporaceae* BGCs analyzed via multidimensional scaling using correlation pairwise distance.

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 :



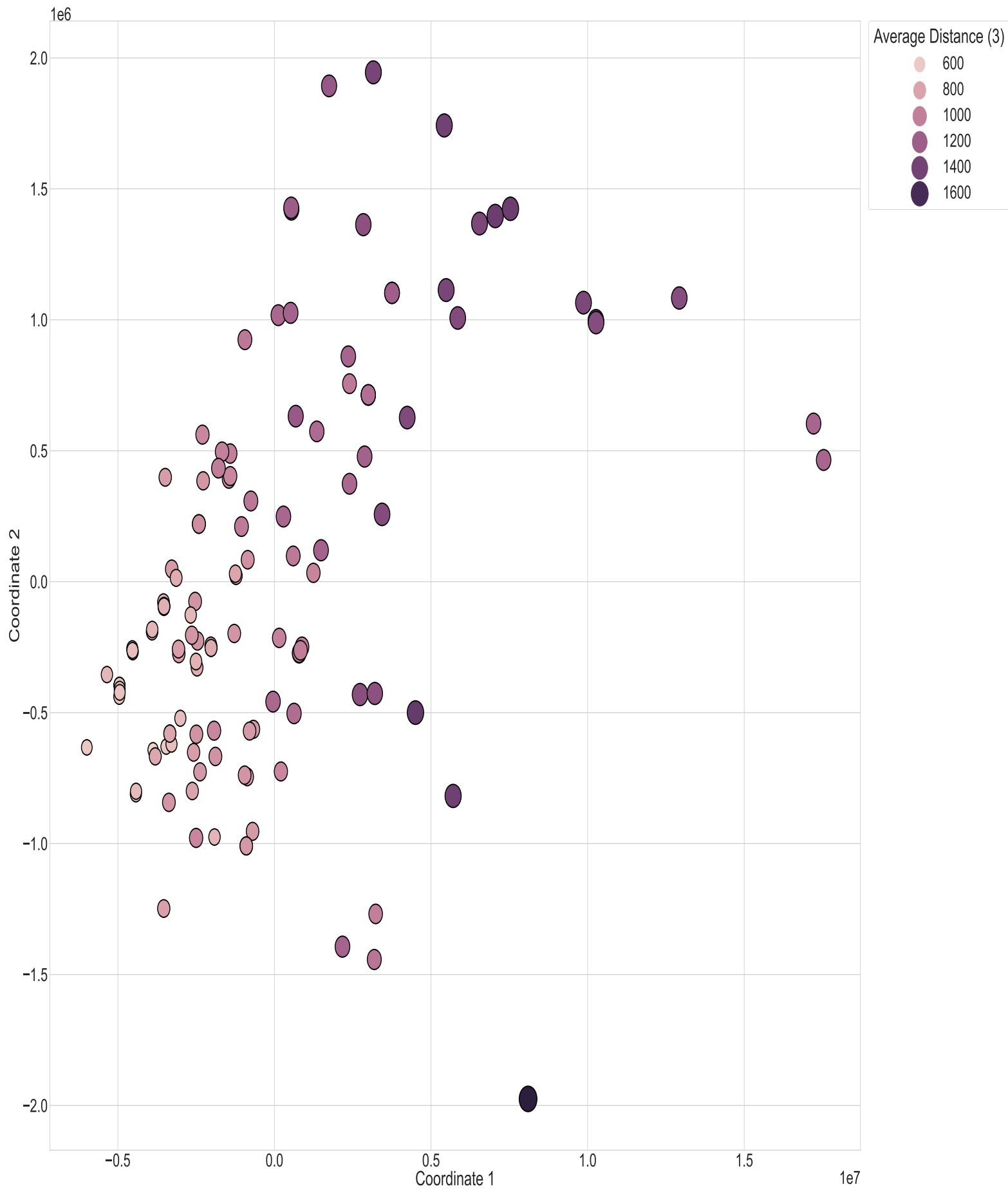
Scatter plot of RIPP *Micromonosporaceae* BGCs analyzed via multidimensional scaling using euclidean pairwise distance. Only RIPP BGCs were considered for this analysis. Each dot represents an individual BGC. The hue and size of the dots were scaled based on the average distance of the individual RIPP 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 :

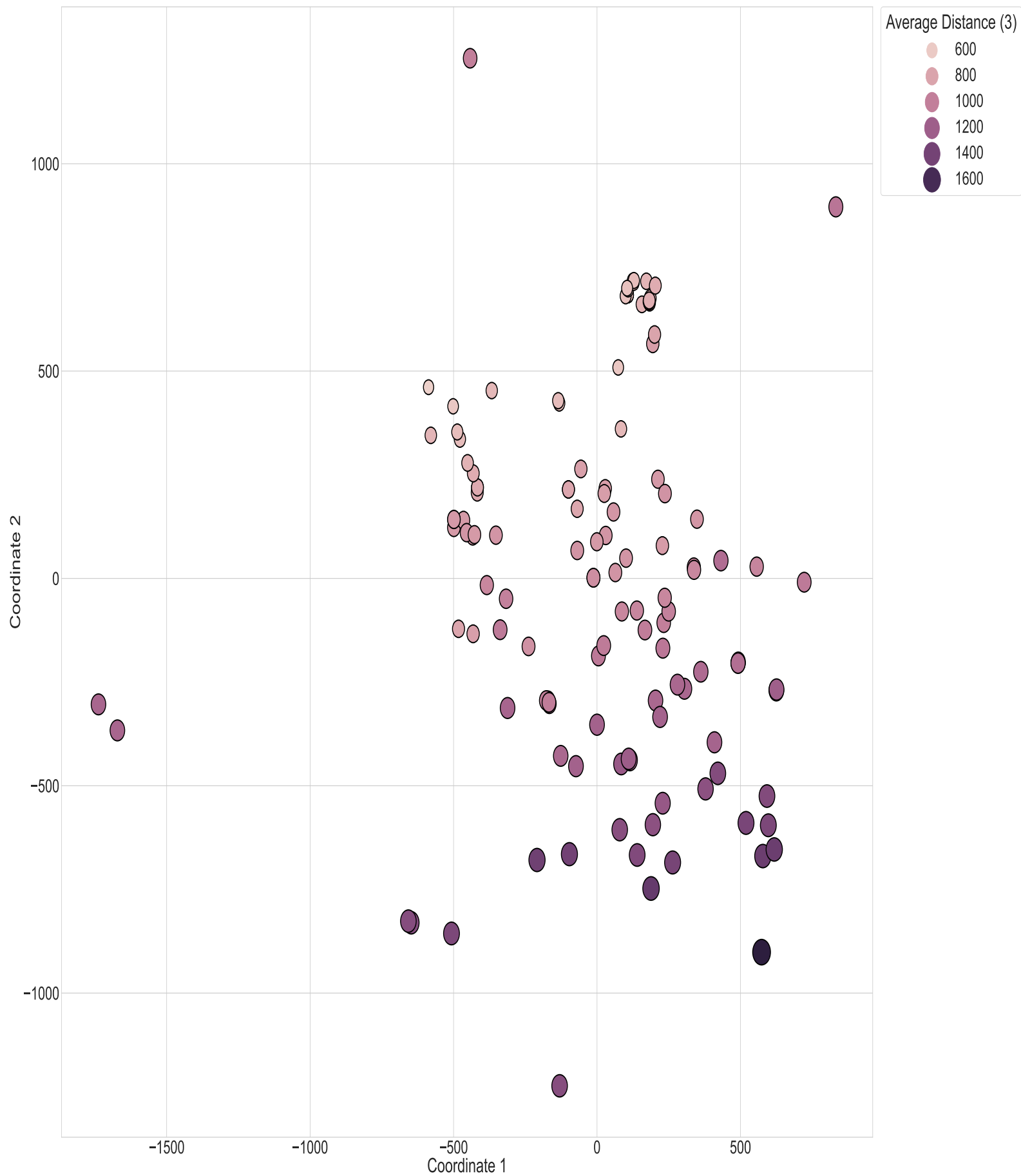


Scatter plot of RIPP *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cityblock pairwise distance.

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 :

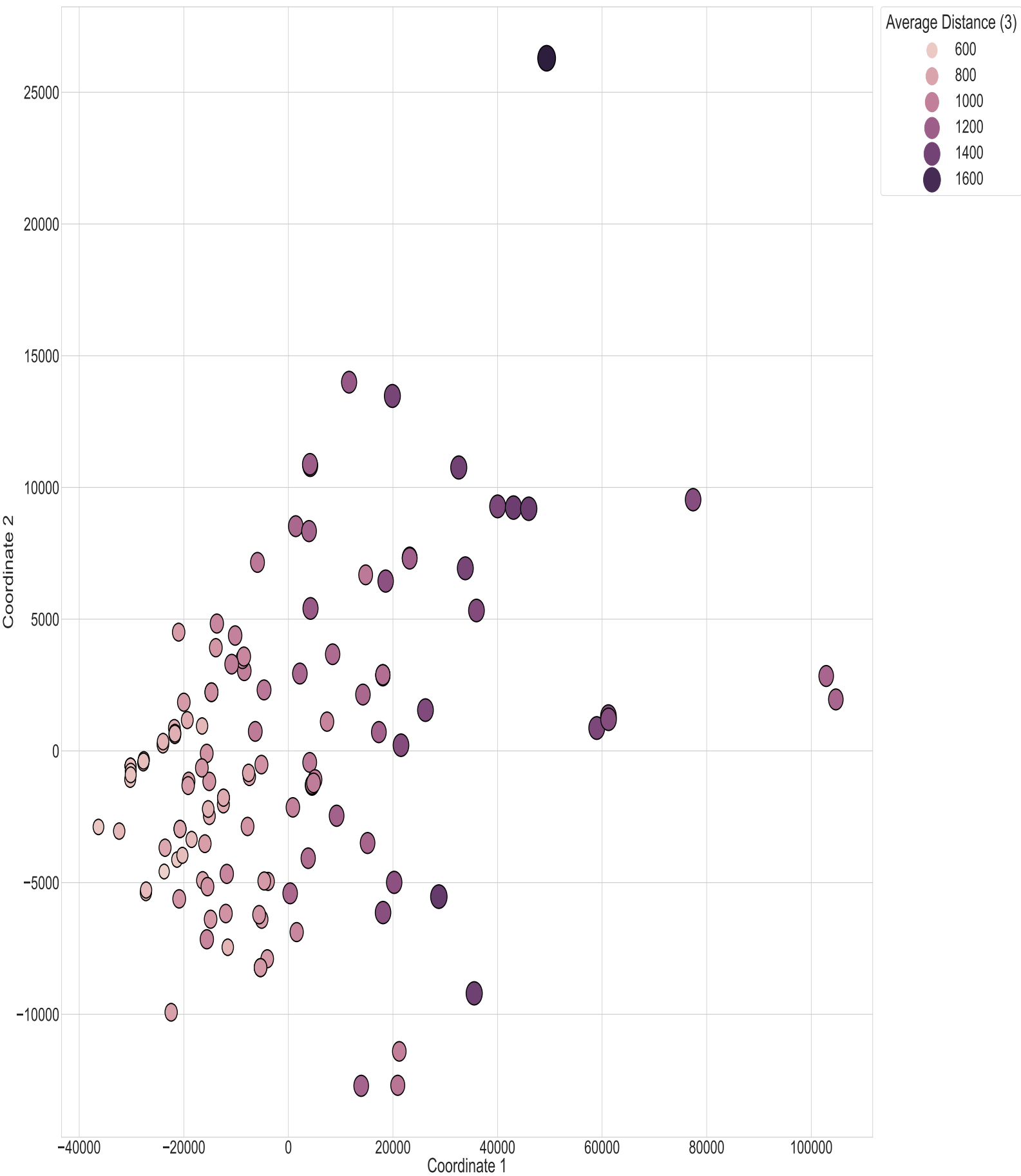


Scatter plot of RIPP *Micromonosporaceae* BGCs analyzed via multidimensional scaling using chebyshev pairwise distance.

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 :

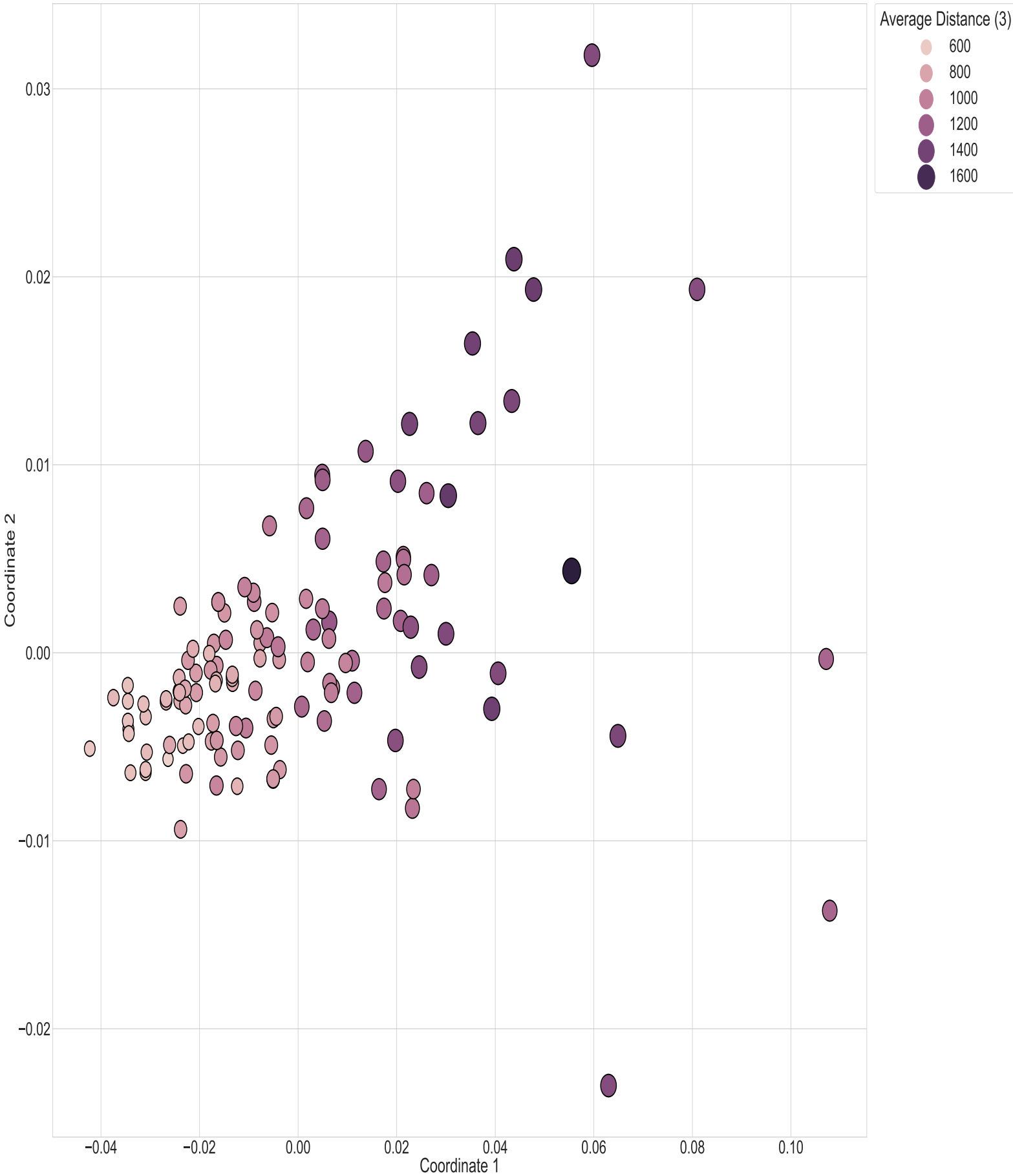


Scatter plot of RIPP *Micromonosporaceae* BGCs analyzed via multidimensional scaling using 12 pairwise distance.

Only RIPP BGCs were considered for this analysis. Each dot represents an individual BGC.

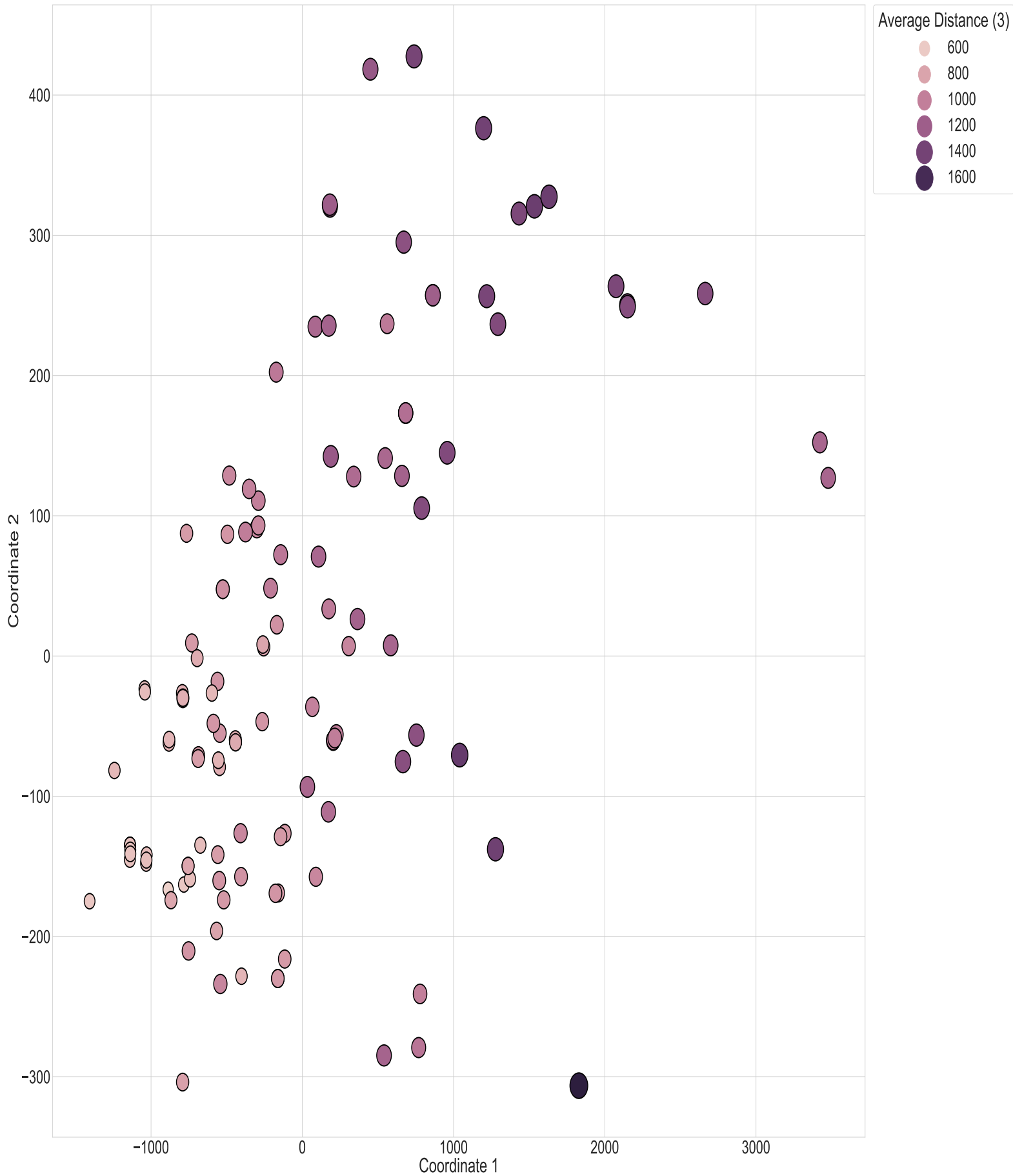
The hue and size of the dots were scaled based on the average distance of the individual RIPP BGC to the closest 3 GCFs in BiG-SLICE (T=900).

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 :



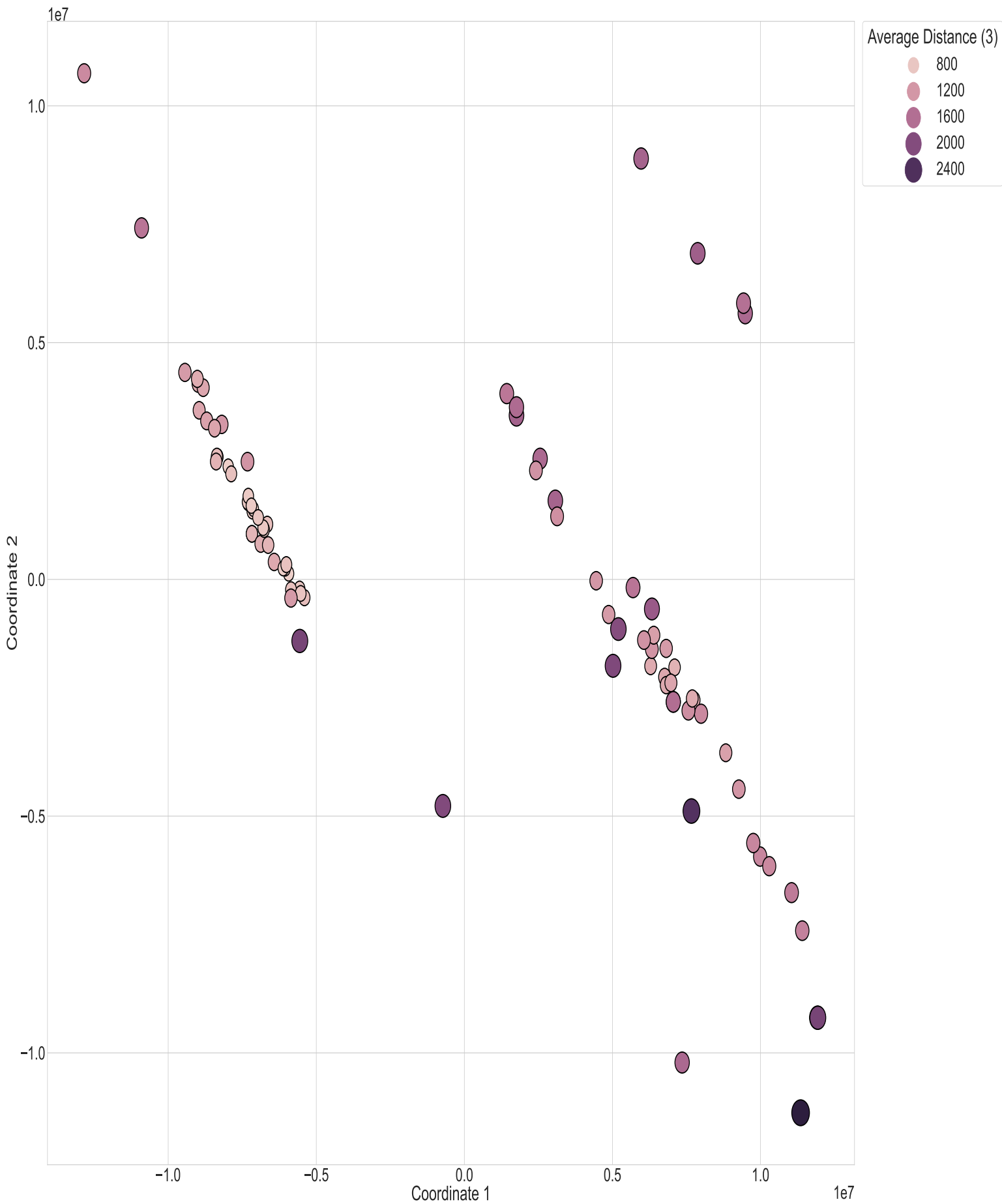
Scatter plot of PKS-Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using euclidean pairwise distance. Only PKS-Other BGCs were considered for this analysis. Each dot represents an individual BGC. The hue and size of the dots were scaled based on the average distance of the individual 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 :

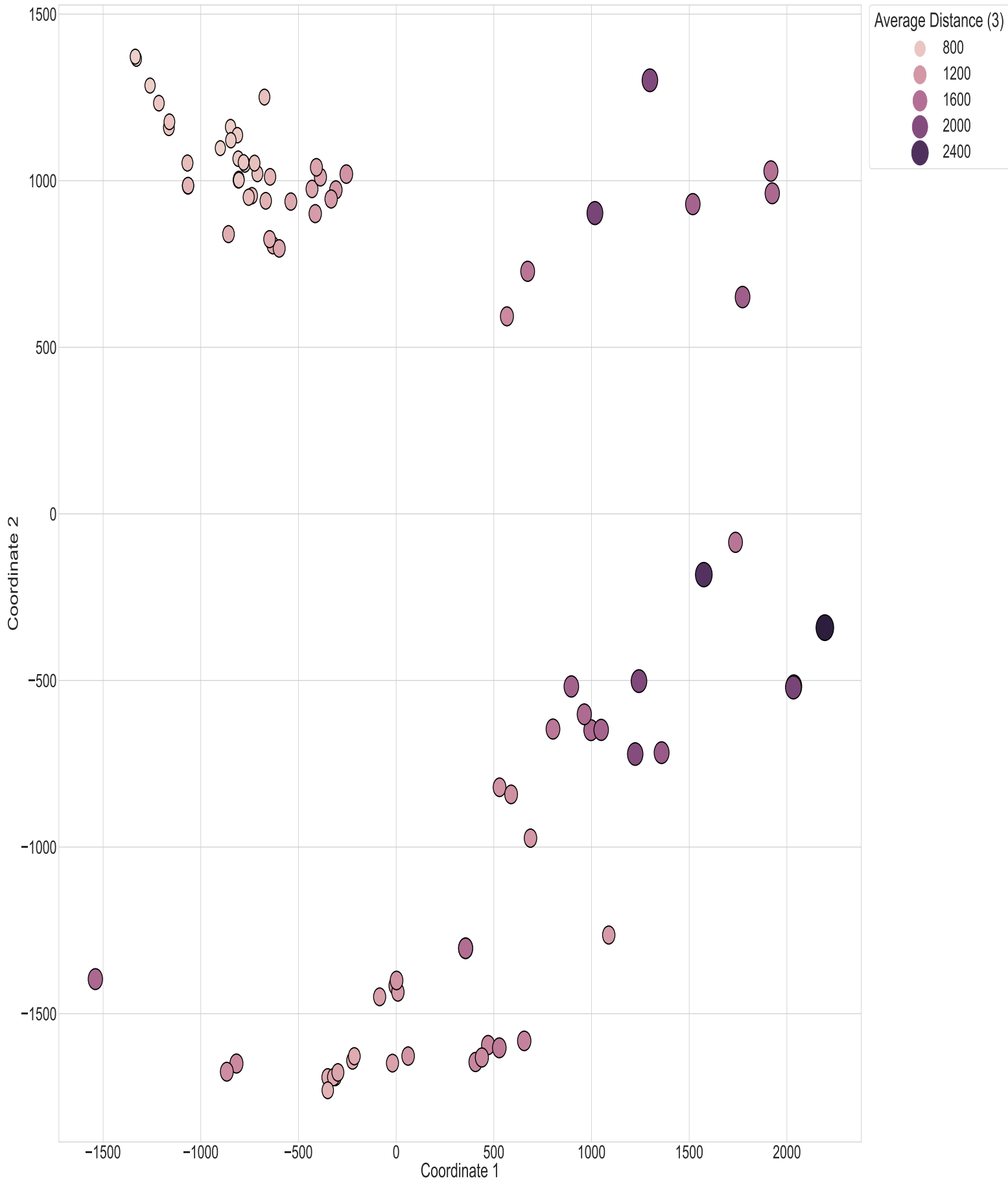


Scatter plot of PKS-Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cityblock pairwise distance.

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 :

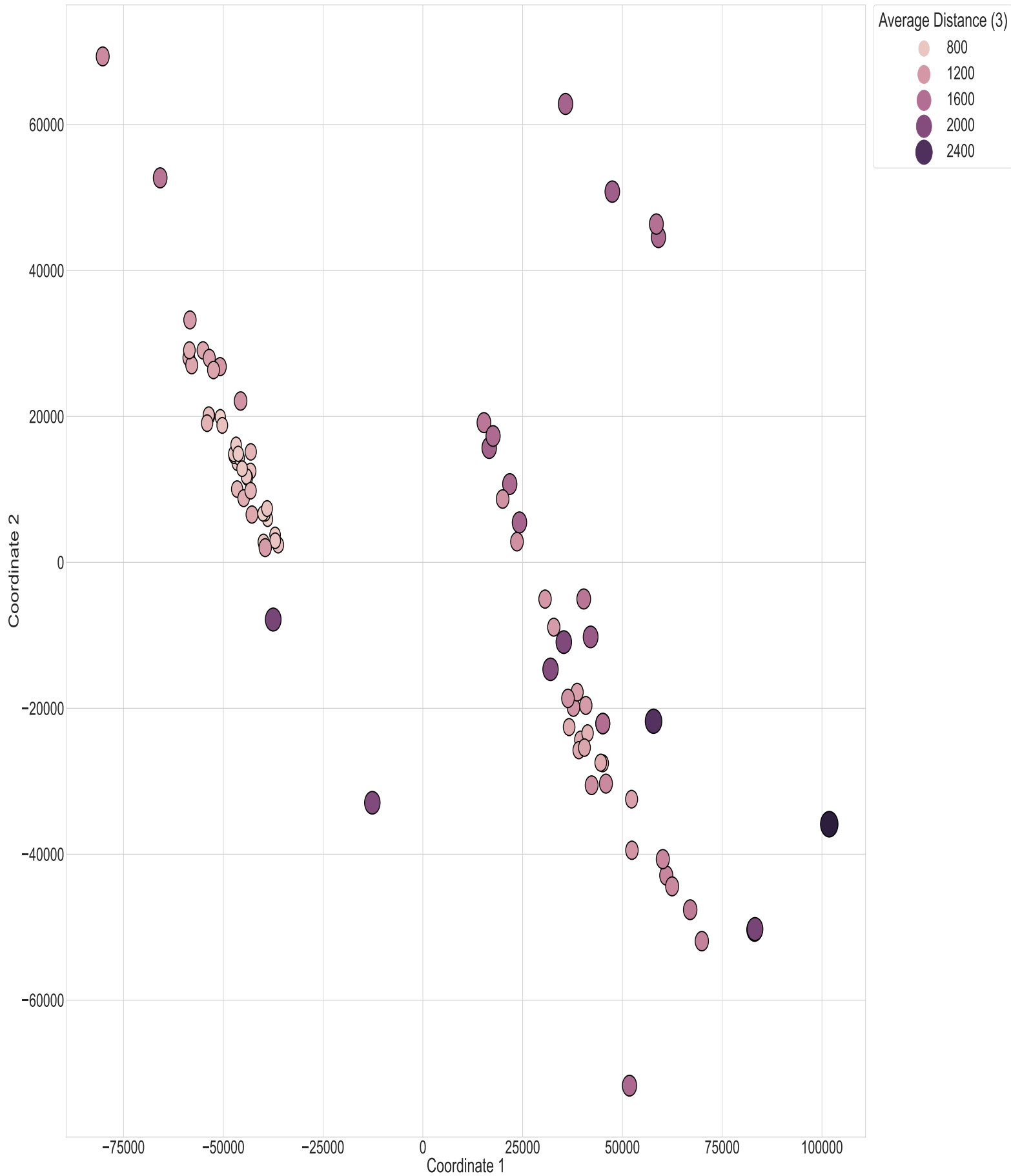


Scatter plot of PKS-Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using chebyshev pairwise distance.

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 :

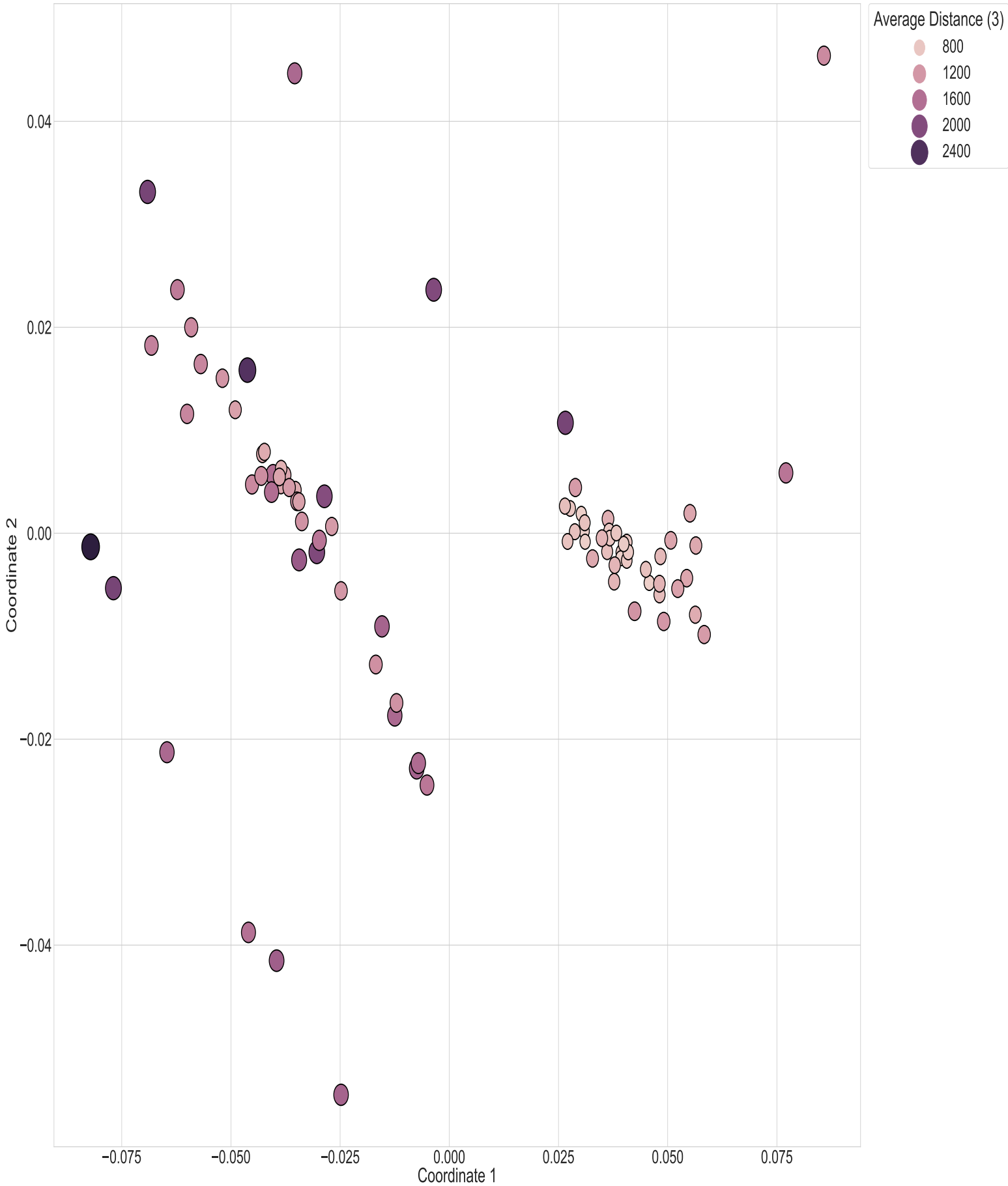


Scatter plot of PKS-Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using 12 pairwise distance.

Only PKS-Other BGCs were considered for this analysis. Each dot represents an individual BGC.

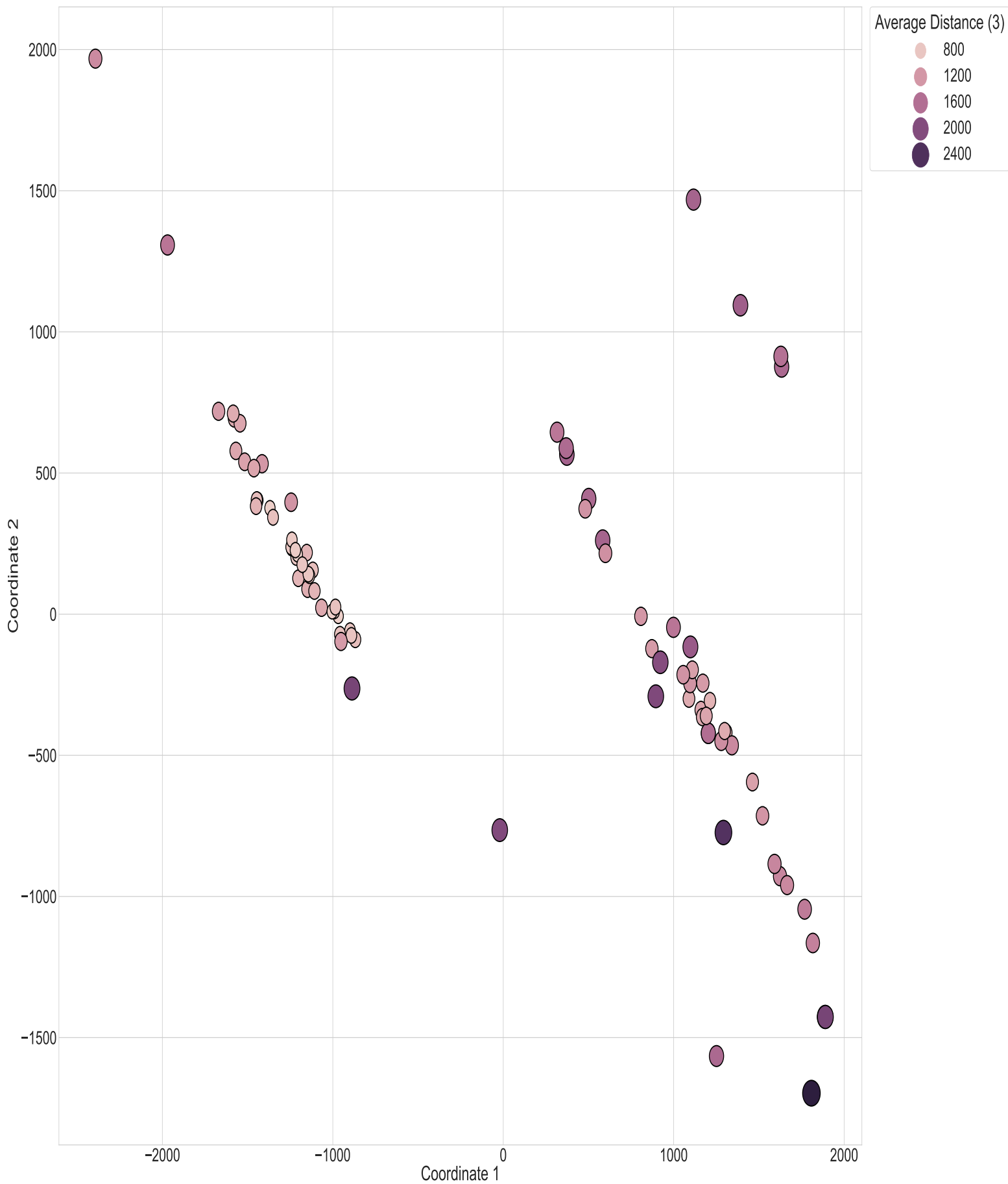
The hue and size of the dots were scaled based on the average distance of the individual PKS-Other BGC to the closest 3 GCFs in BiG-SLICE (T=900).

Figure – S130 :



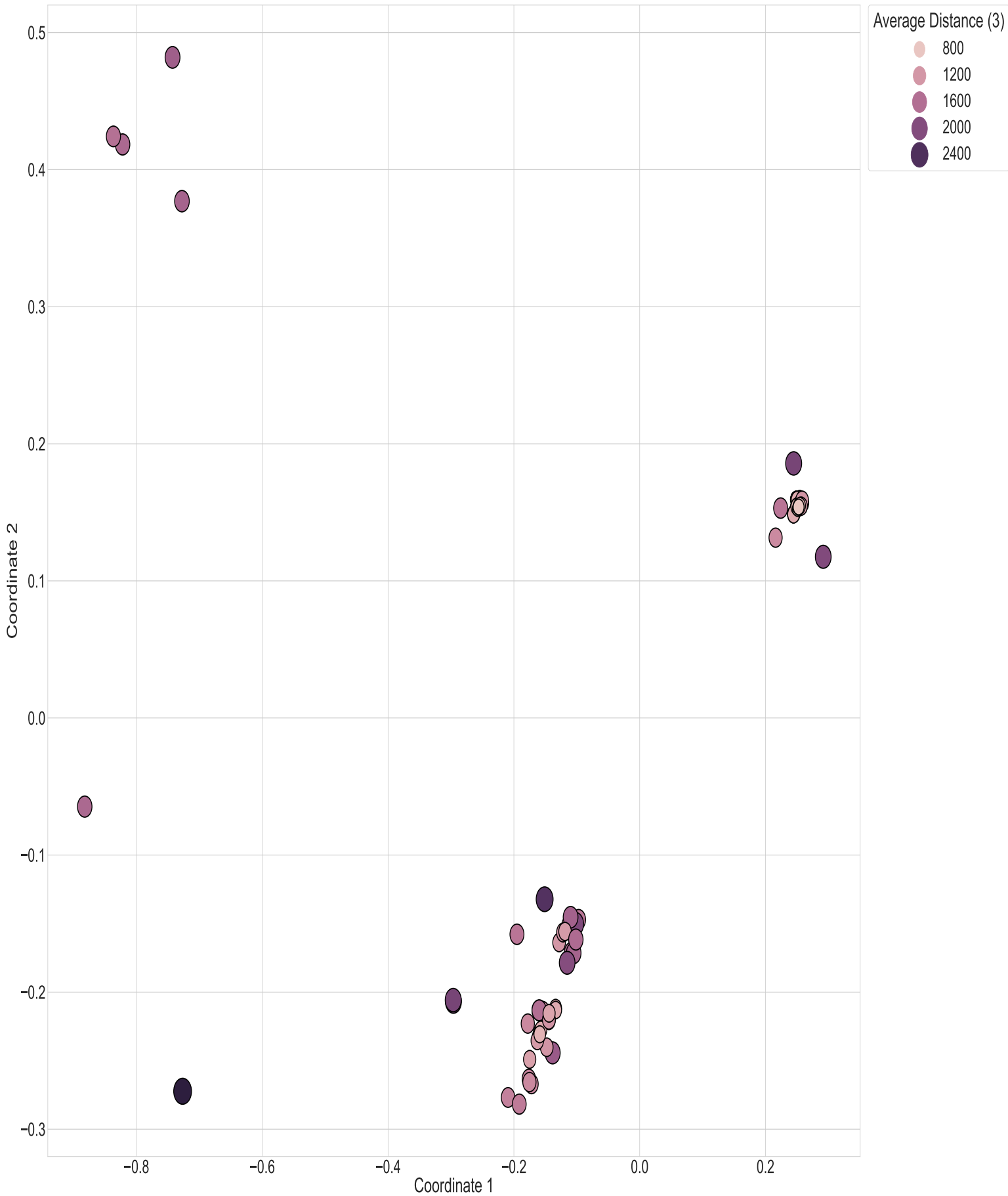
Scatter plot of PKS-Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using braycurtis pairwise distance. Only PKS-Other BGCs were considered for this analysis. Each dot represents an individual BGC. The hue and size of the dots were scaled based on the average distance of the individual 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 :

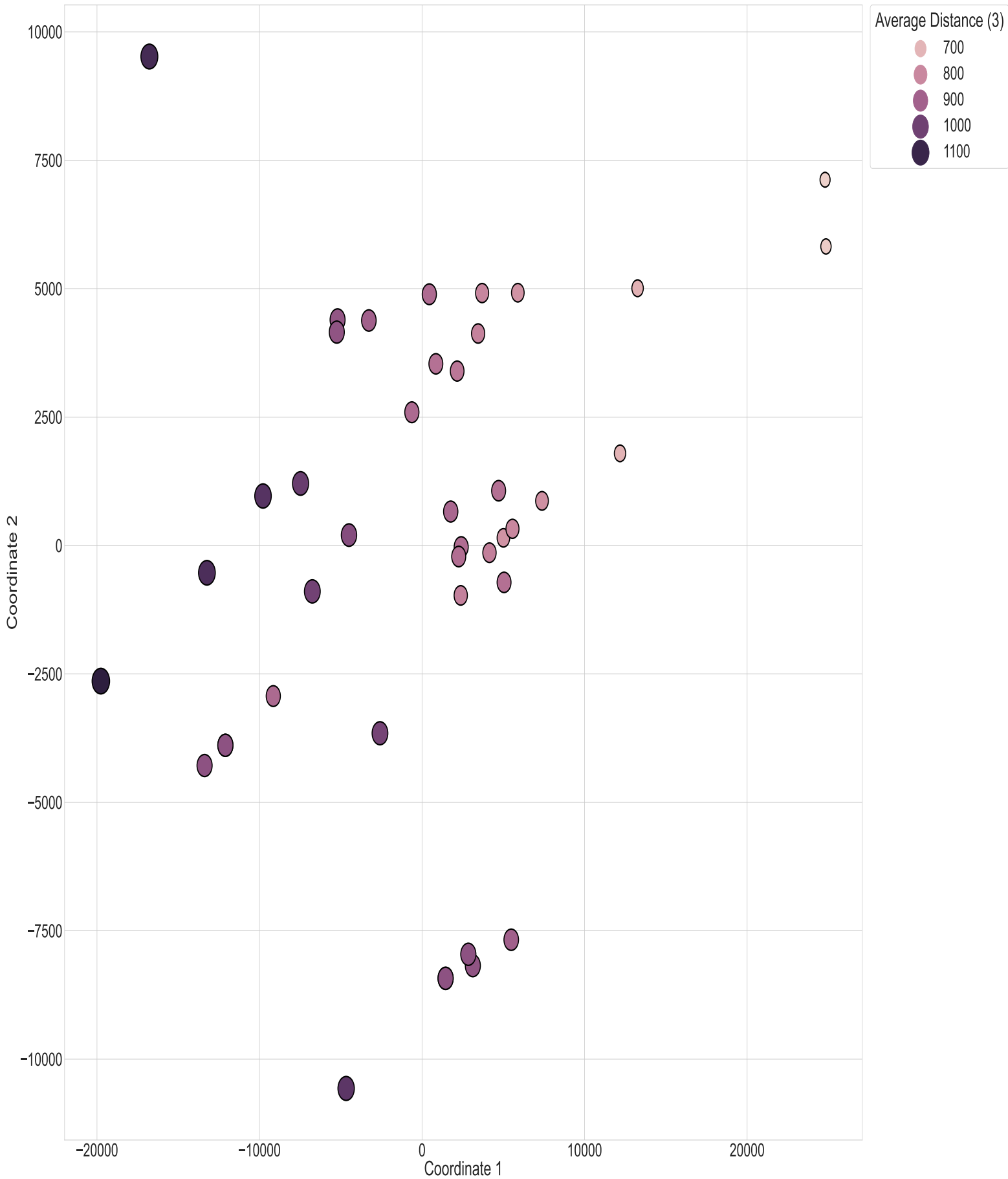


Scatter plot of PKS-Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using correlation pairwise distance.

Only PKS-Other BGCs were considered for this analysis. Each dot represents an individual BGC.

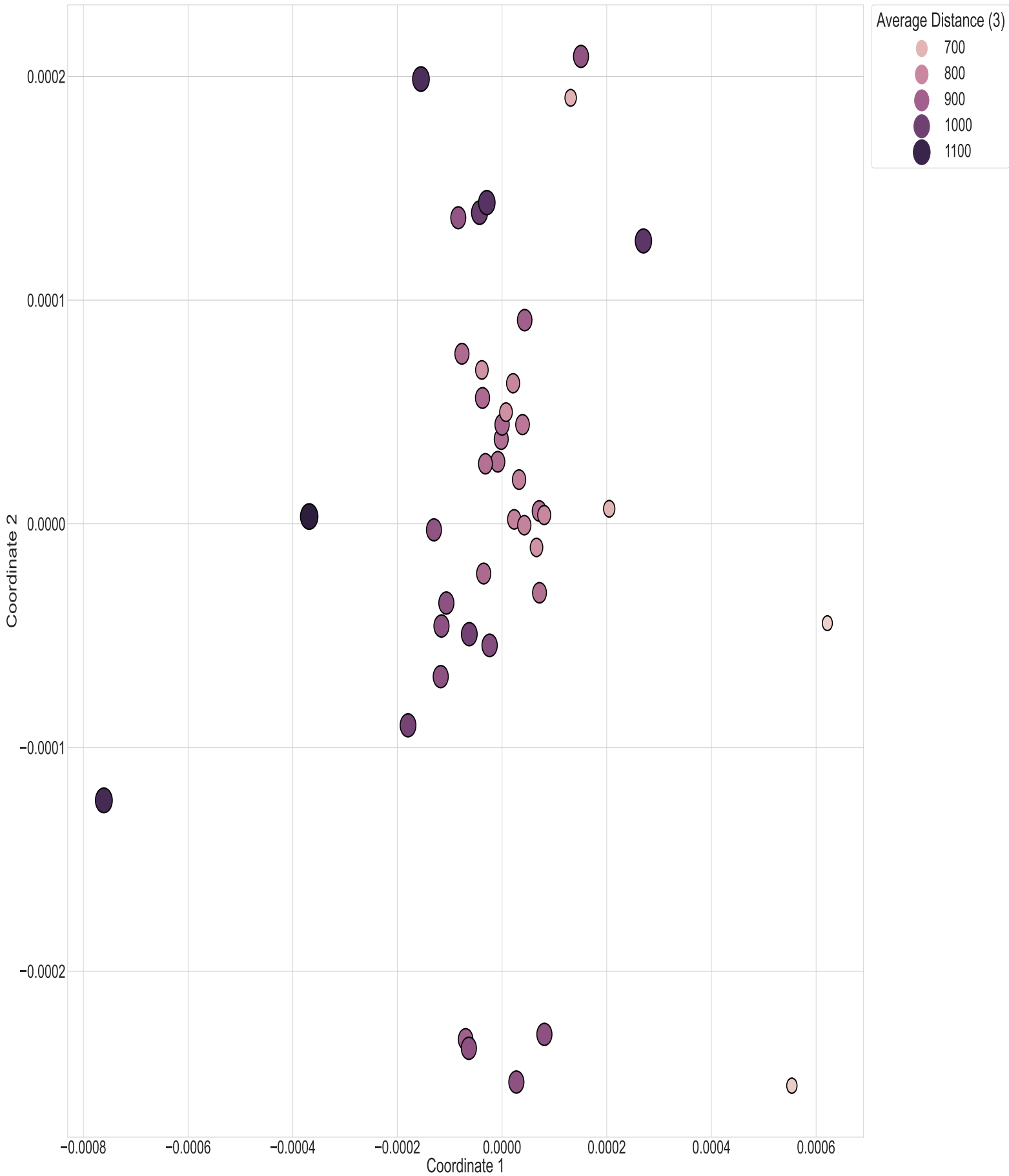
The hue and size of the dots were scaled based on the average distance of the individual PKS-Other BGC to the closest 3 GCFs in BiG-SLICE (T=900).

Figure – S133 :



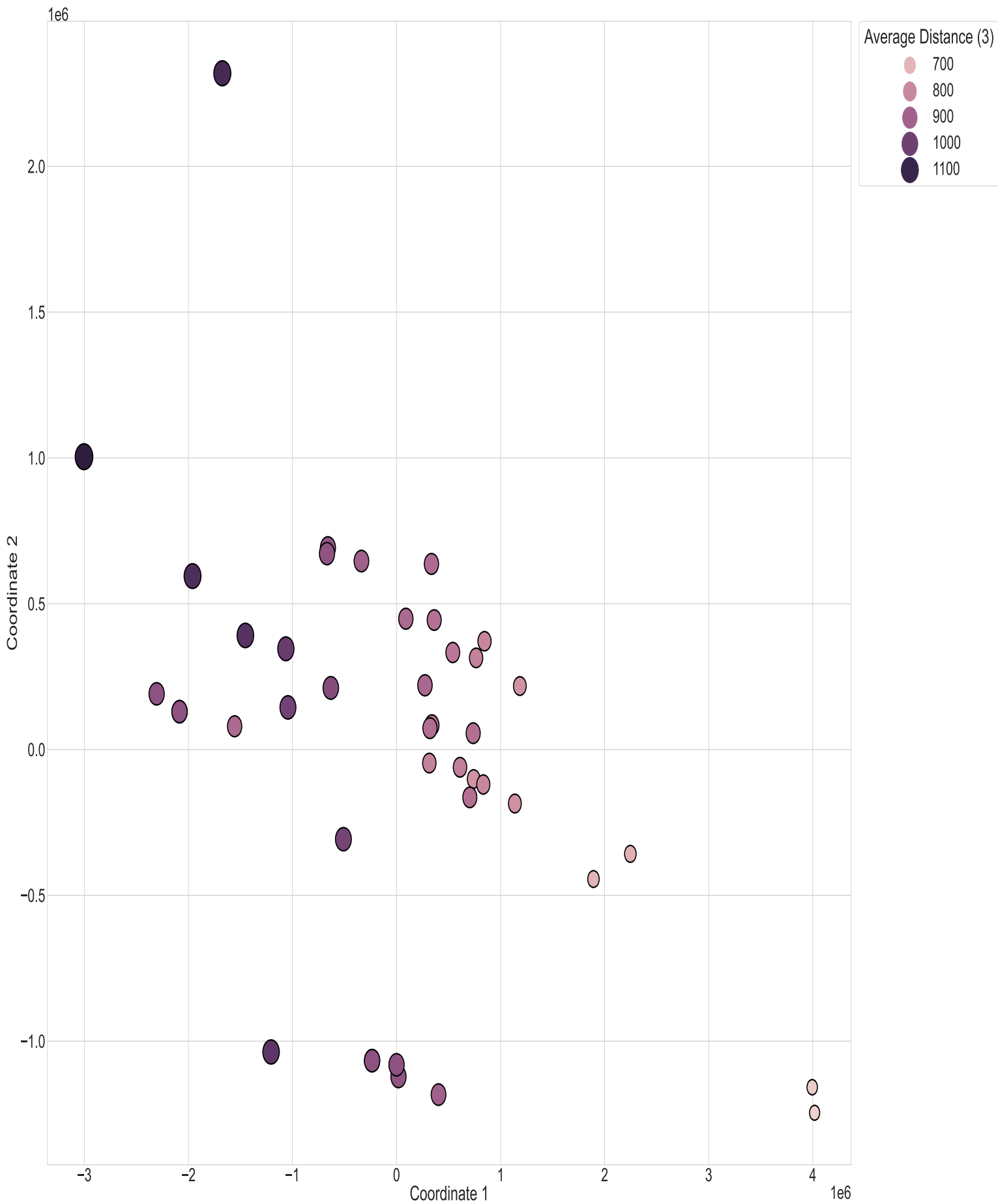
Scatter plot of Siderophore *Micromonosporaceae* BGCs analyzed via multidimensional scaling using euclidean pairwise distance. Only Siderophore BGCs were considered for this analysis. Each dot represents an individual BGC. The hue and size of the dots were scaled based on the average distance of the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

Figure – S134 :



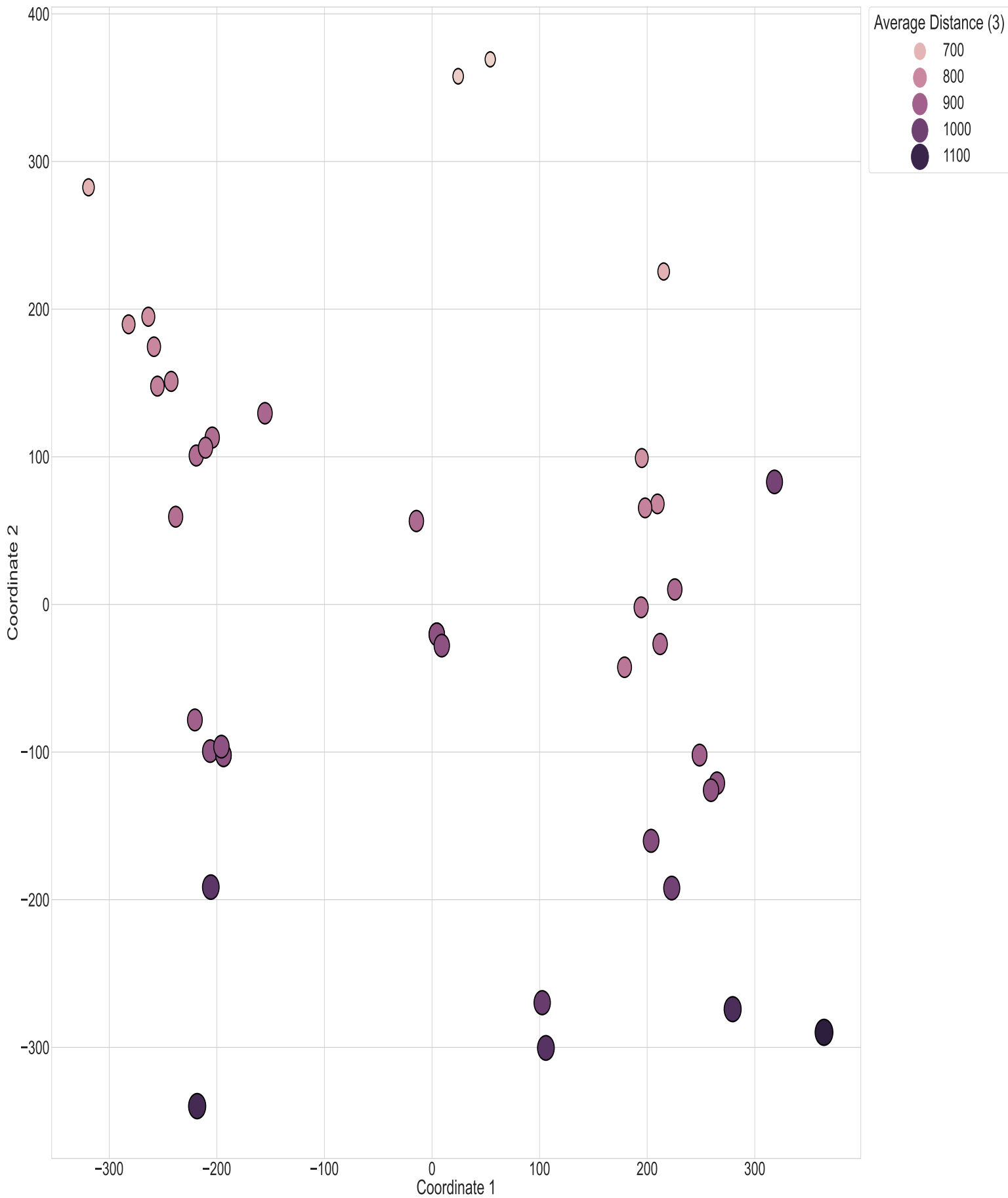
Scatter plot of Siderophore *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cosine pairwise distance. Only Siderophore BGCs were considered for this analysis. Each dot represents an individual BGC. The hue and size of the dots were scaled based on the average distance of the individual 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 :

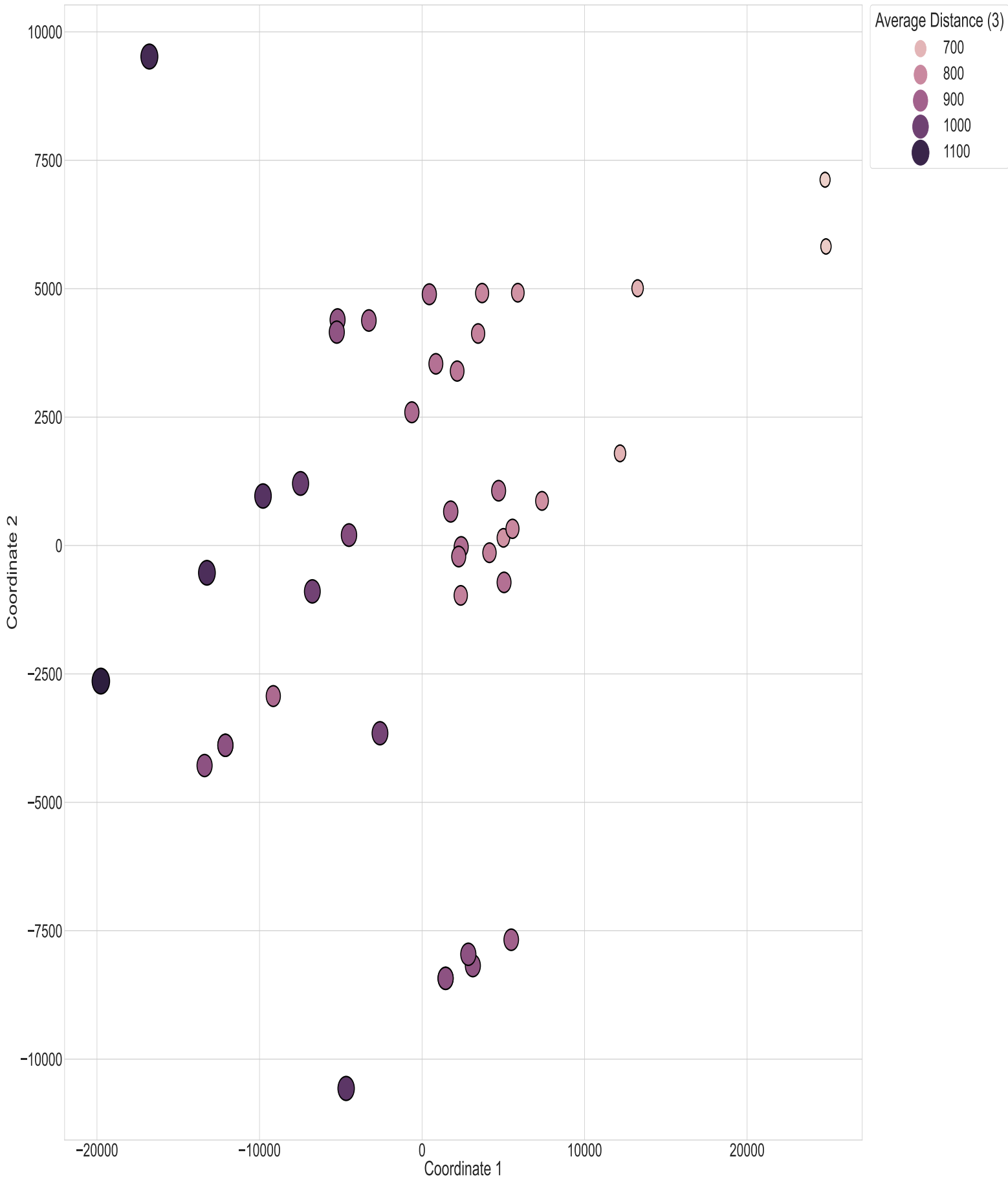


Scatter plot of Siderophore *Micromonosporaceae* BGCs analyzed via multidimensional scaling using chebyshev pairwise distance.

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 :

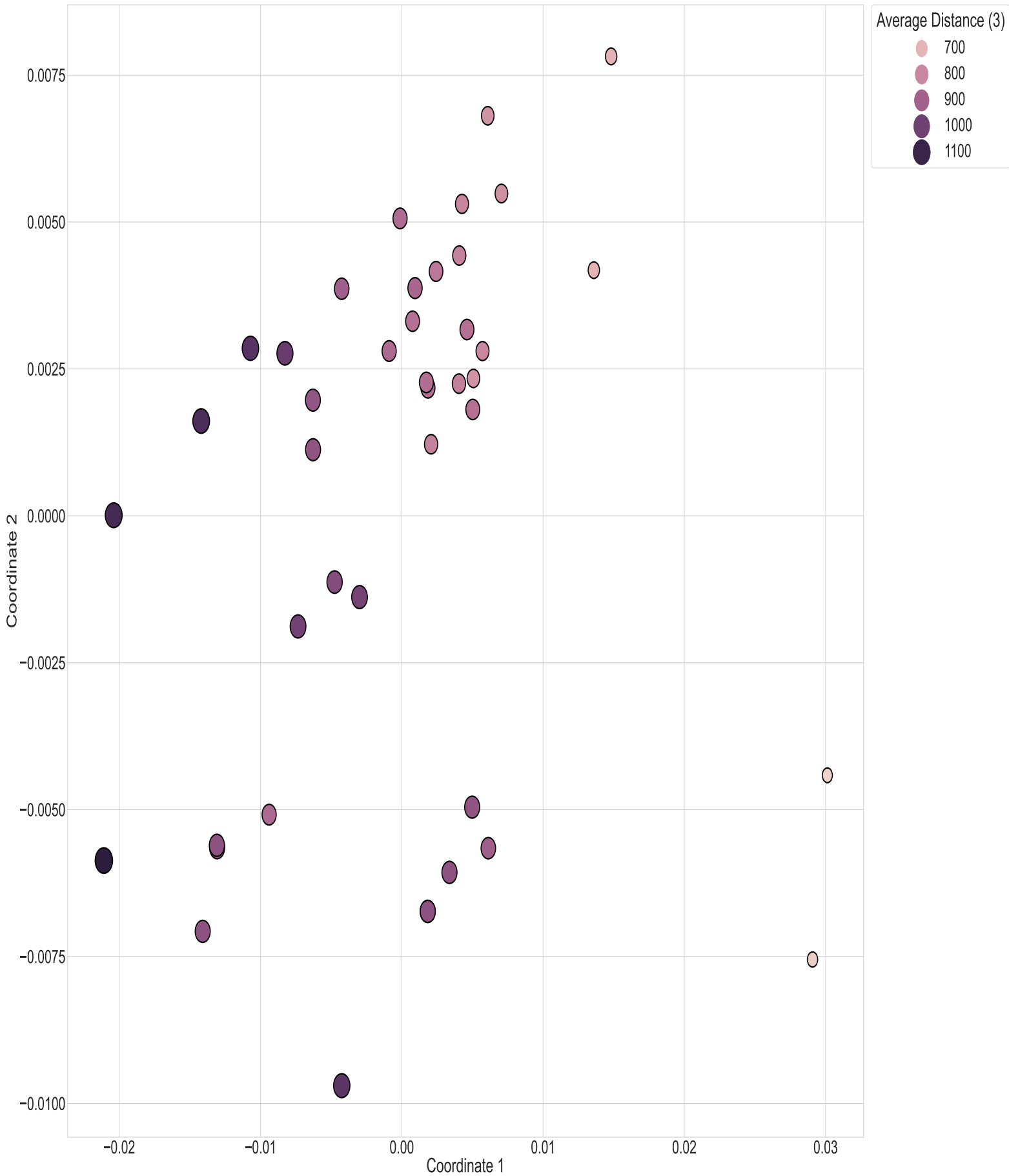


Scatter plot of Siderophore *Micromonosporaceae* BGCs analyzed via multidimensional scaling using l2 pairwise distance.

Only Siderophore BGCs were considered for this analysis. Each dot represents an individual BGC.

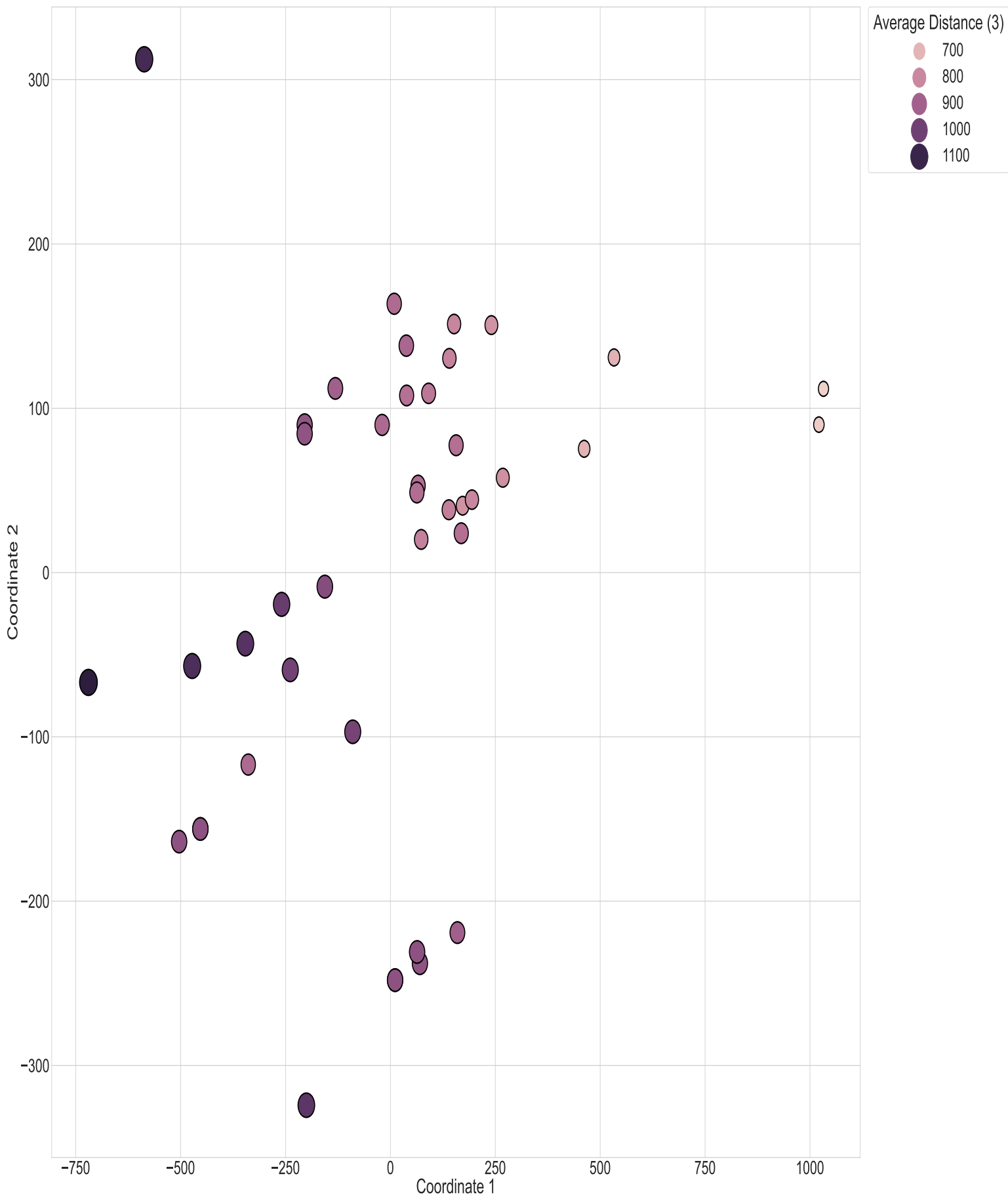
The hue and size of the dots were scaled based on the average distance of the individual Siderophore BGC to the closest 3 GCFs in BiG-SLICE (T=900).

Figure – S138 :



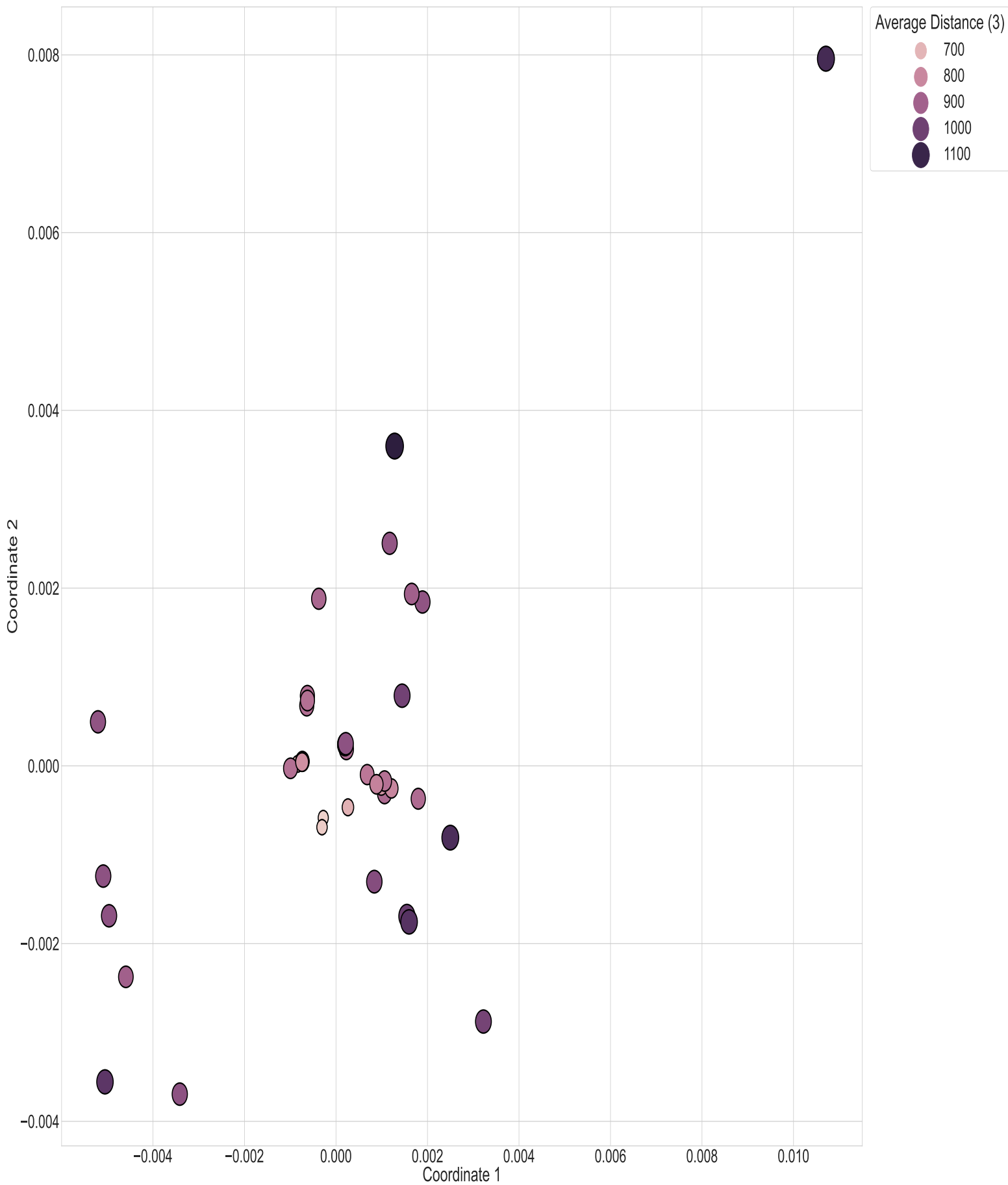
Scatter plot of Siderophore *Micromonosporaceae* BGCs analyzed via multidimensional scaling using braycurtis pairwise distance. Only Siderophore BGCs were considered for this analysis. Each dot represents an individual BGC. The hue and size of the dots were scaled based on the average distance of the 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 :

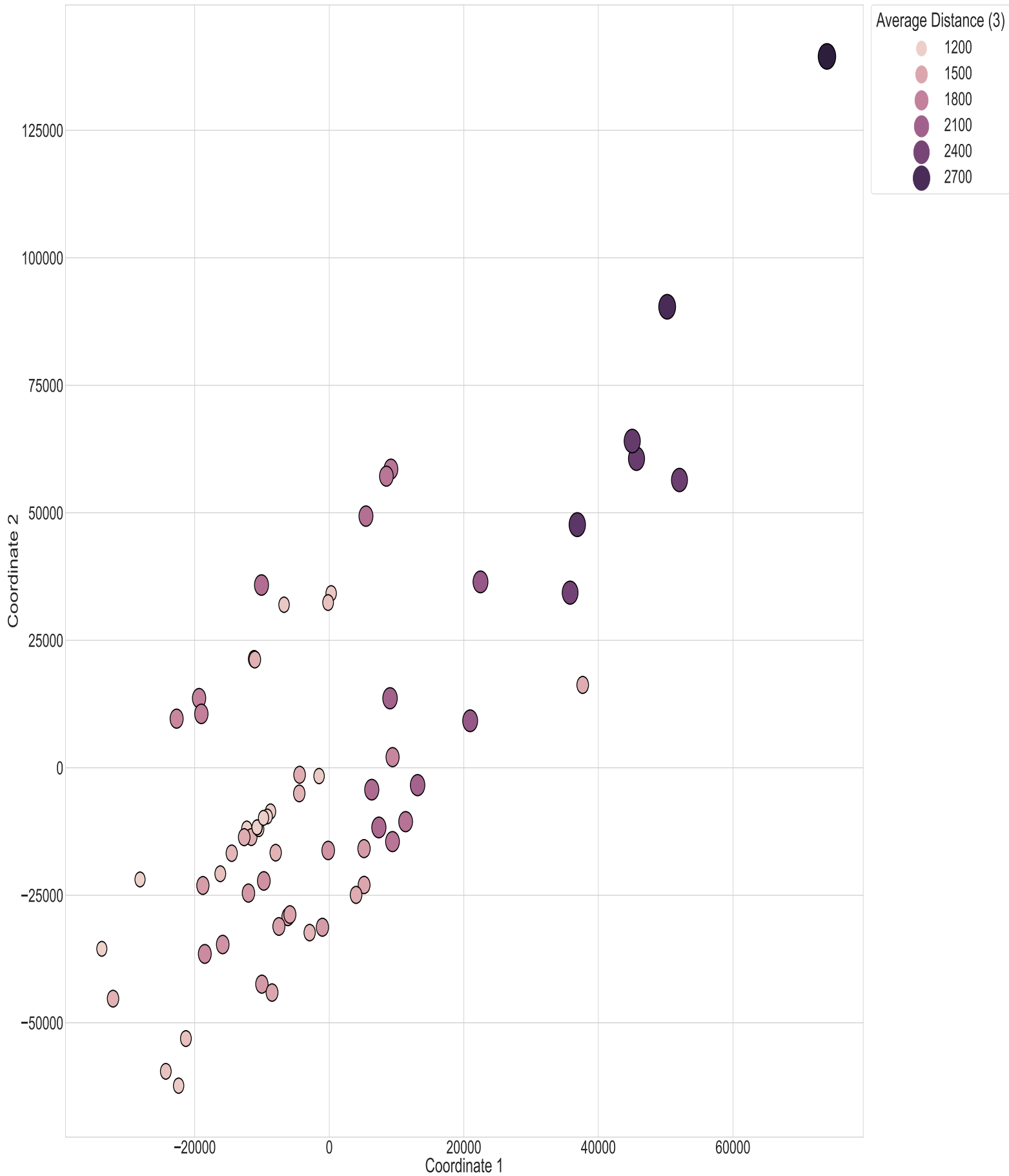


Scatter plot of Siderophore *Micromonosporaceae* BGCs analyzed via multidimensional scaling using correlation pairwise distance.

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 :

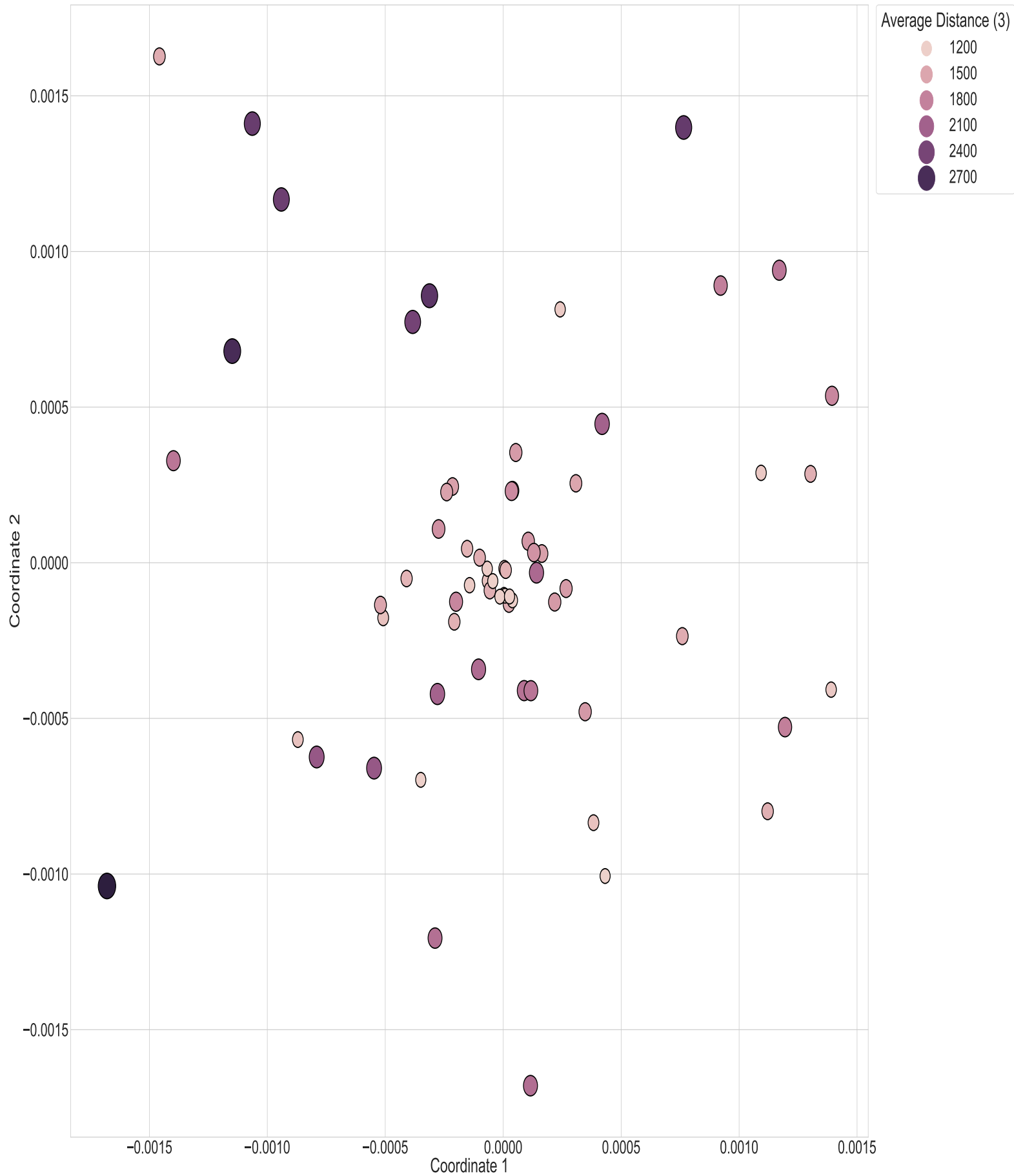


Scatter plot of PKS-NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using euclidean pairwise distance.

Only PKS-NRPS BGCs were considered for this analysis. Each dot represents an individual BGC.

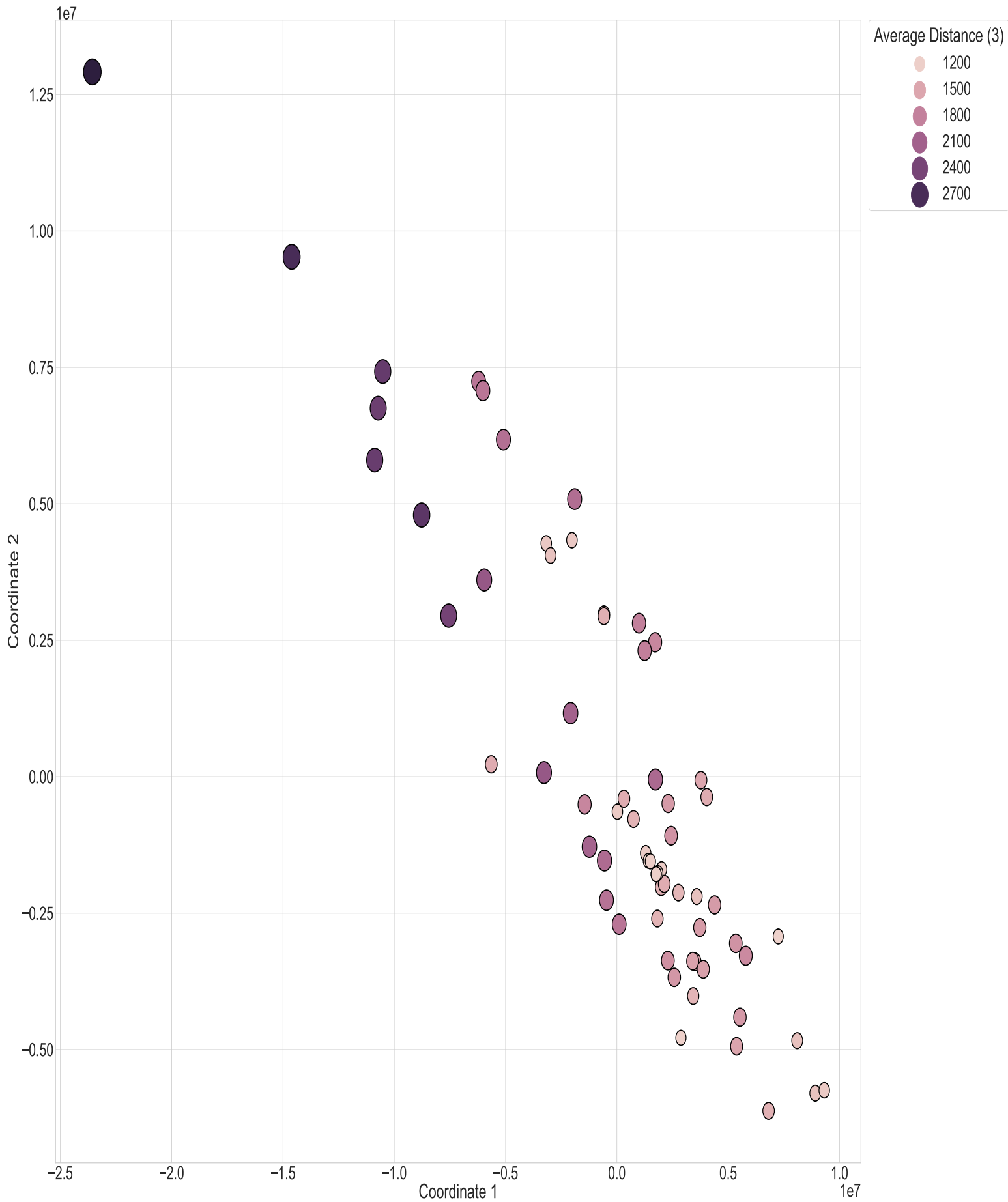
The hue and size of the dots were scaled based on the average distance of the individual PKS-NRPS BGC to the closest 3 GCFs in BiG-SLICE (T=900).

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 :

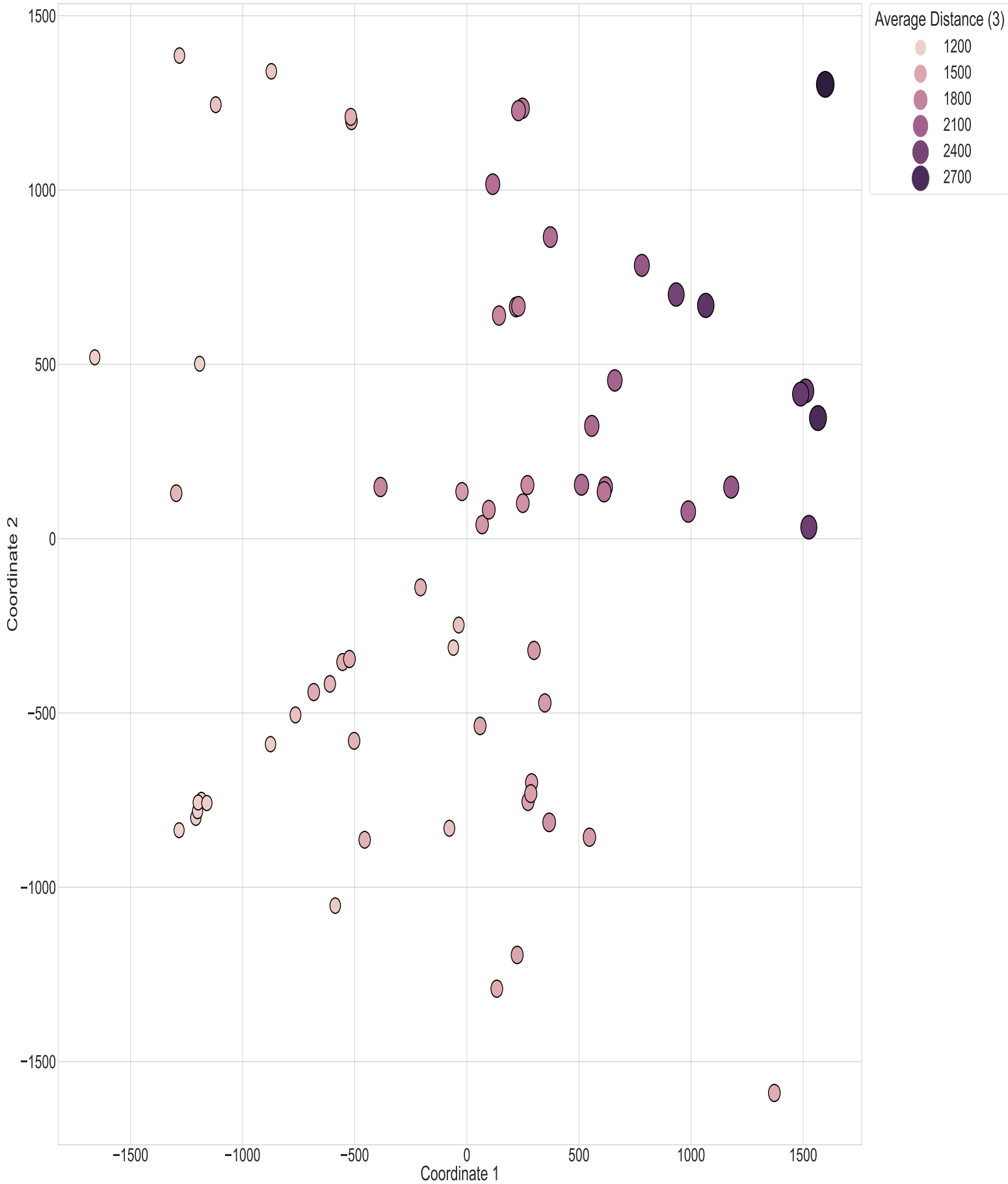


Scatter plot of PKS-NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cityblock pairwise distance.

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 :

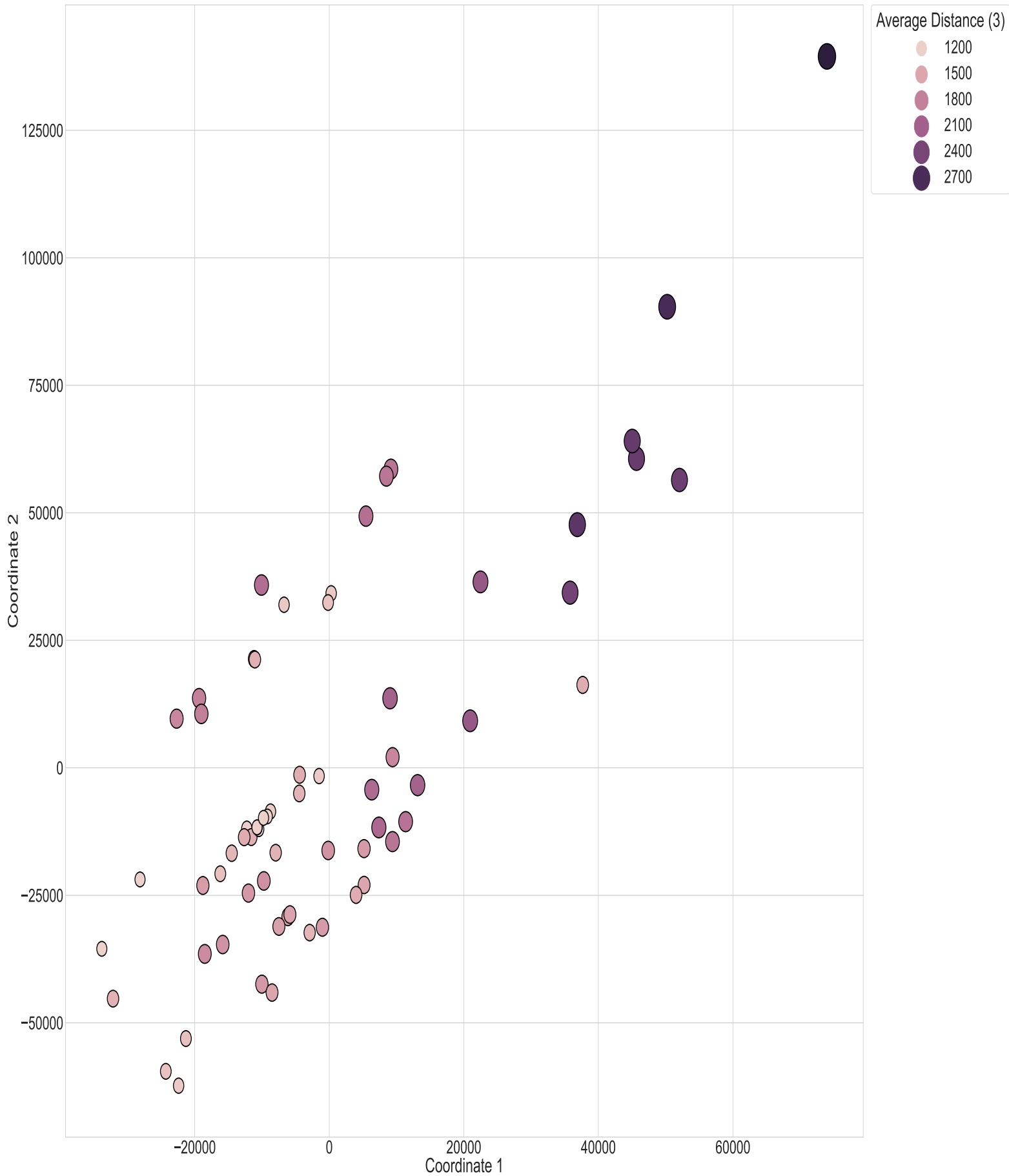


Scatter plot of PKS-NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using chebyshev pairwise distance.

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 :

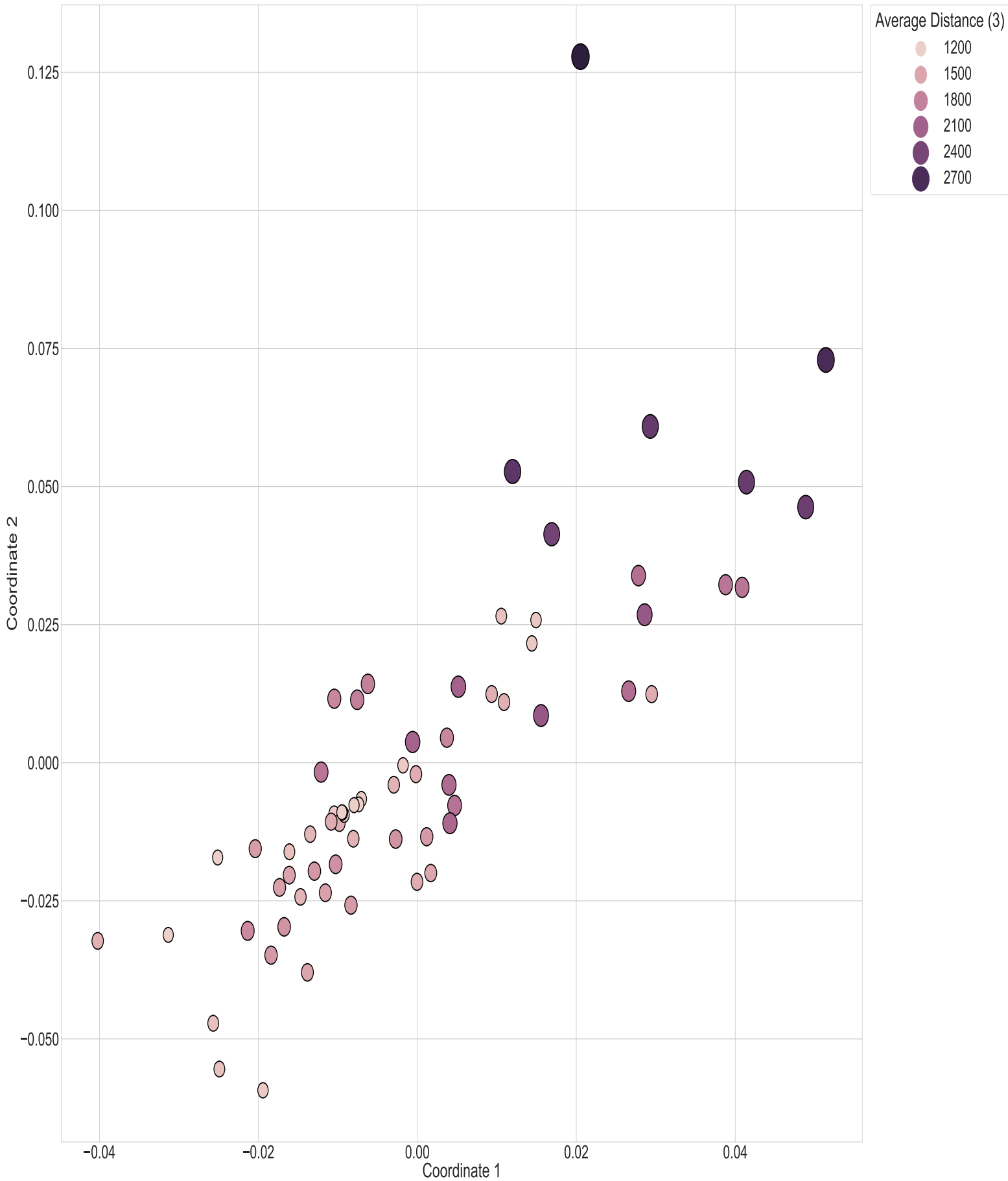


Scatter plot of PKS-NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using l2 pairwise distance.

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 :

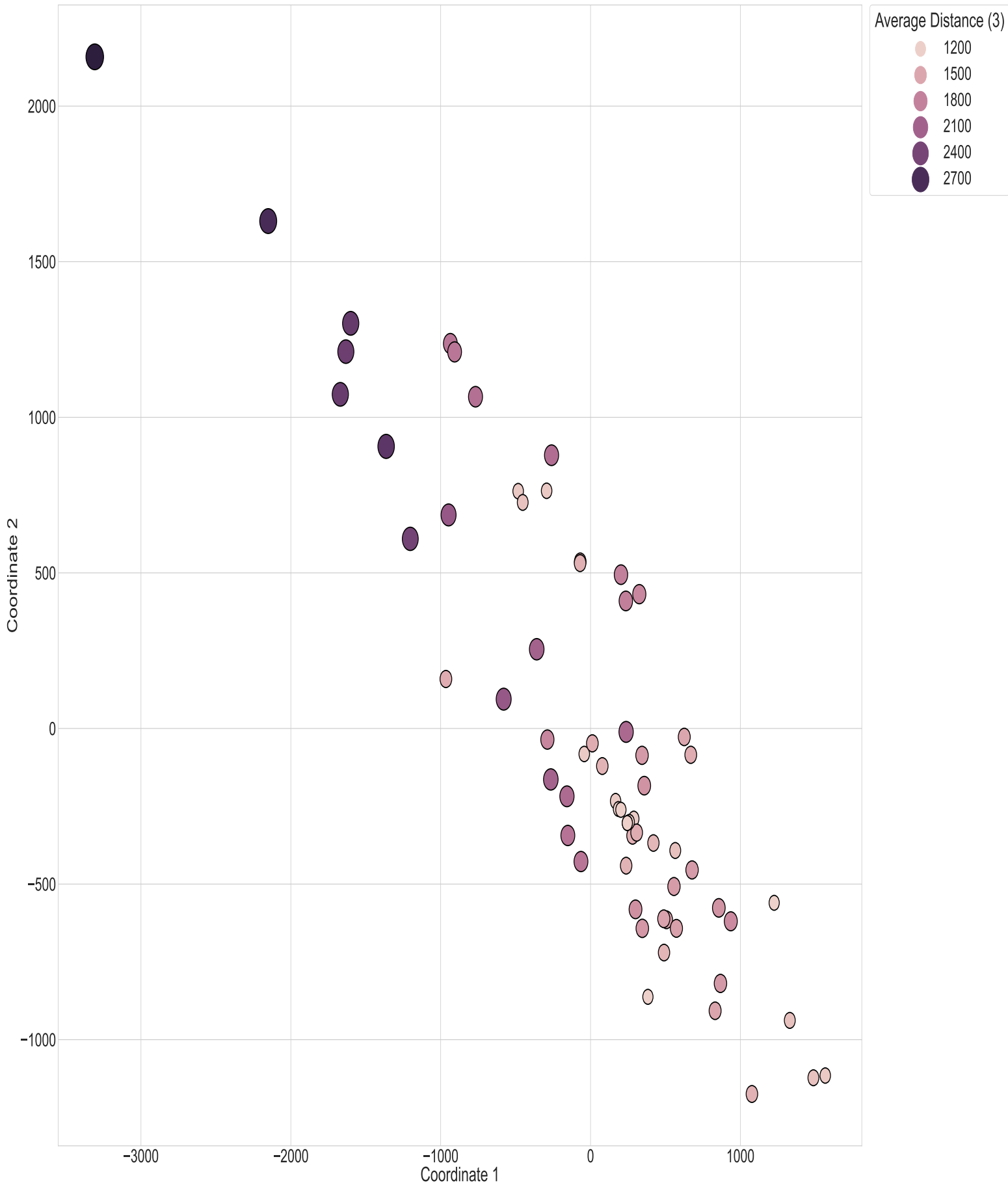


Scatter plot of PKS-NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using braycurtis pairwise distance.

Only PKS-NRPS BGCs were considered for this analysis. Each dot represents an individual BGC.

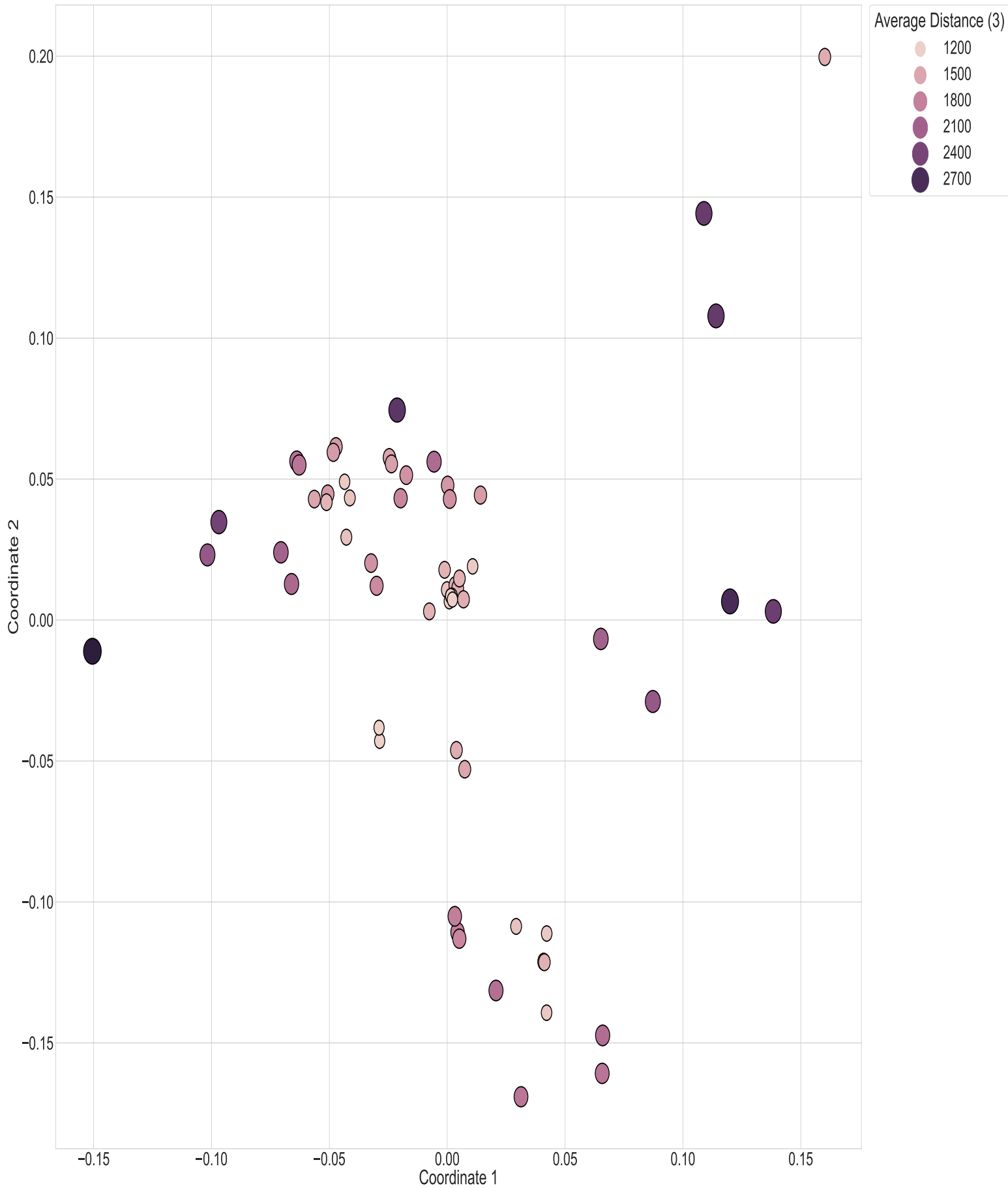
The hue and size of the dots were scaled based on the average distance of the individual PKS-NRPS BGC to the closest 3 GCFs in BiG-SLICE (T=900).

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 :

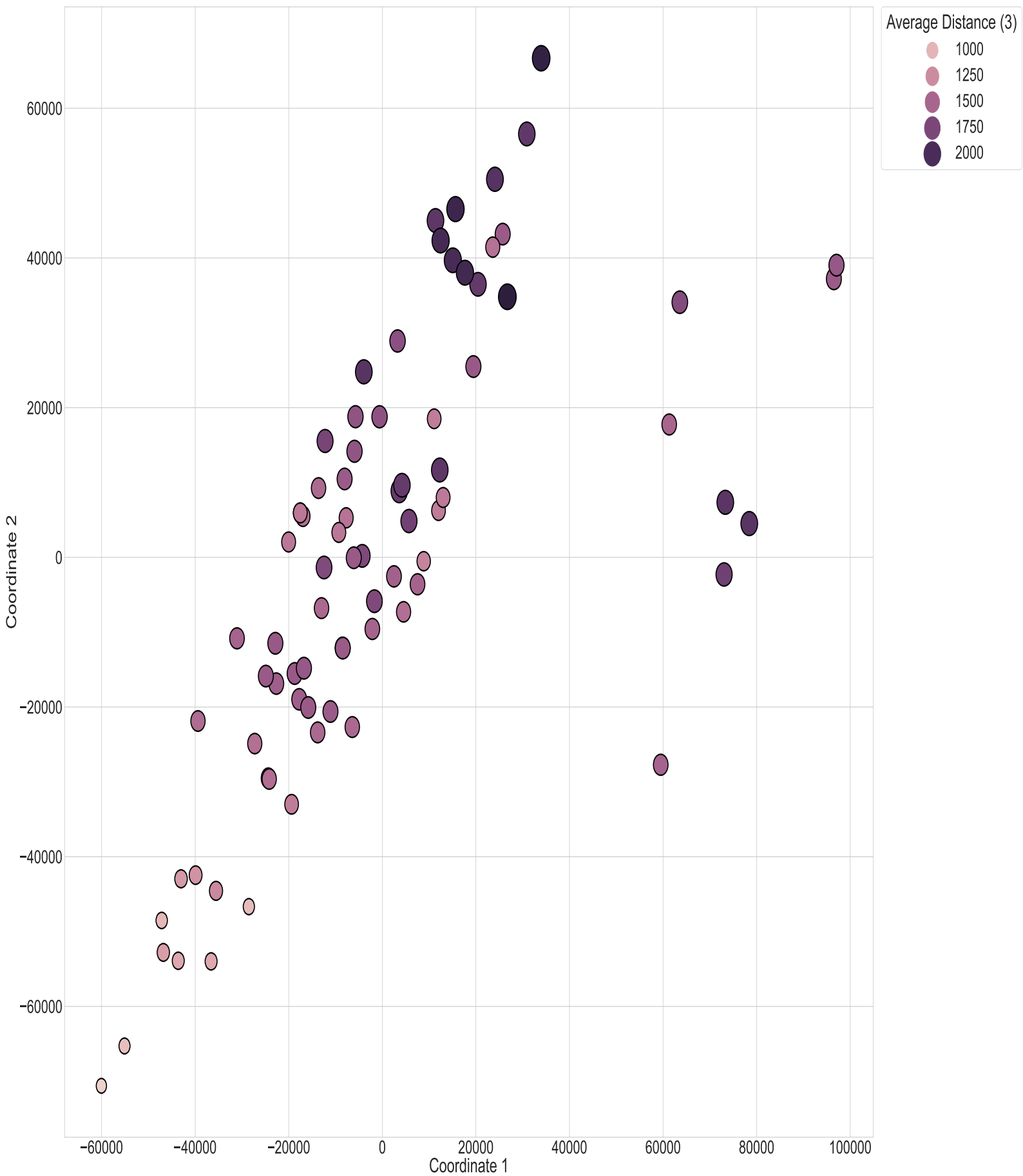


Scatter plot of PKS-NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using correlation pairwise distance.

Only PKS-NRPS BGCs were considered for this analysis. Each dot represents an individual BGC.

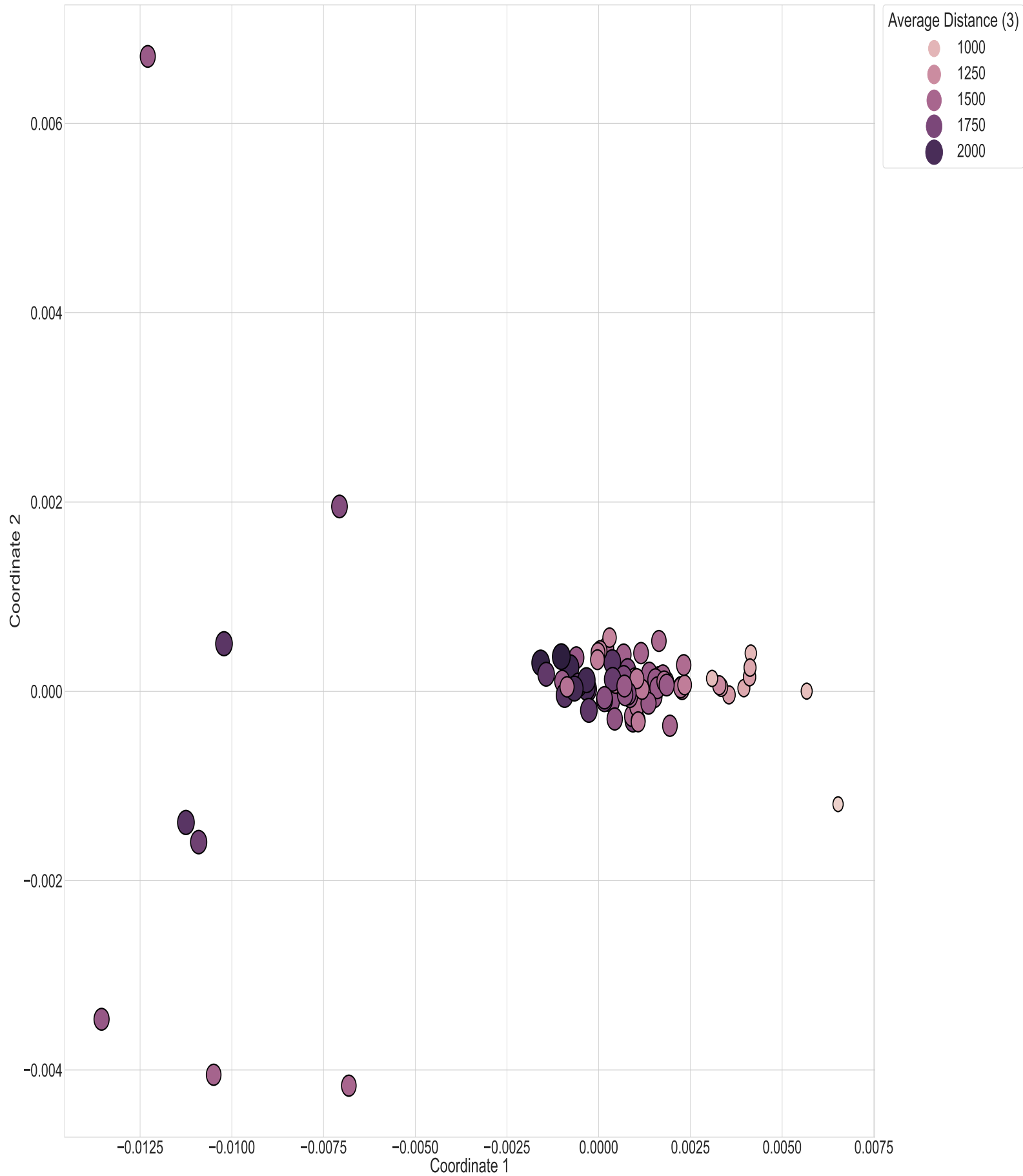
The hue and size of the dots were scaled based on the average distance of the individual PKS-NRPS BGC to the closest 3 GCFs in BiG-SLICE (T=900).

Figure – S149 :



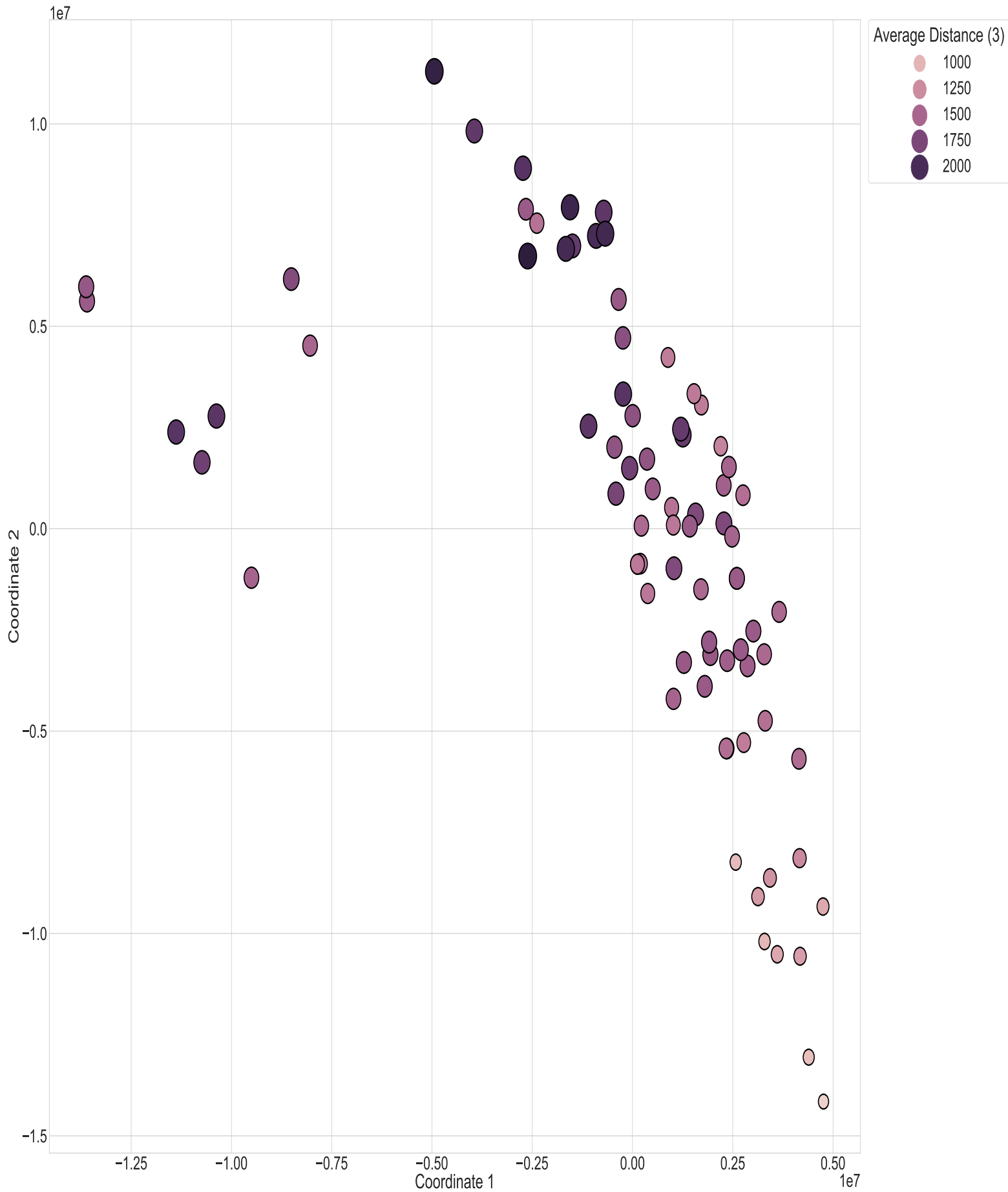
Scatter plot of NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using euclidean pairwise distance. Only NRPS BGCs were considered for this analysis. Each dot represents an individual BGC. The hue and size of the dots were scaled based on the average distance of the individual NRPS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

Figure – S150 :



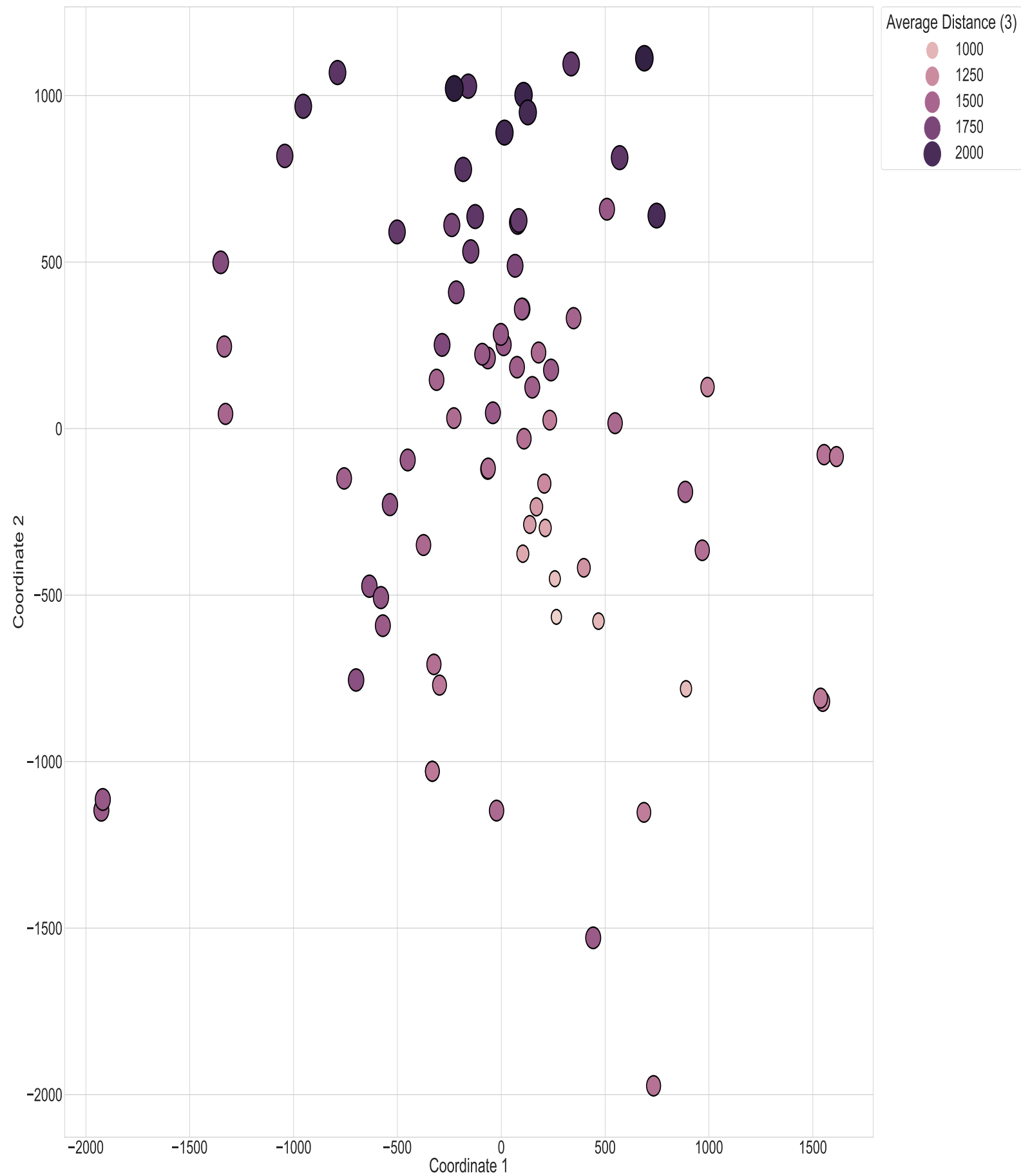
Scatter plot of NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cosine pairwise distance. Only NRPS BGCs were considered for this analysis. Each dot represents an individual BGC. The hue and size of the dots were scaled based on the average distance of the individual NRPS BGC 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 :

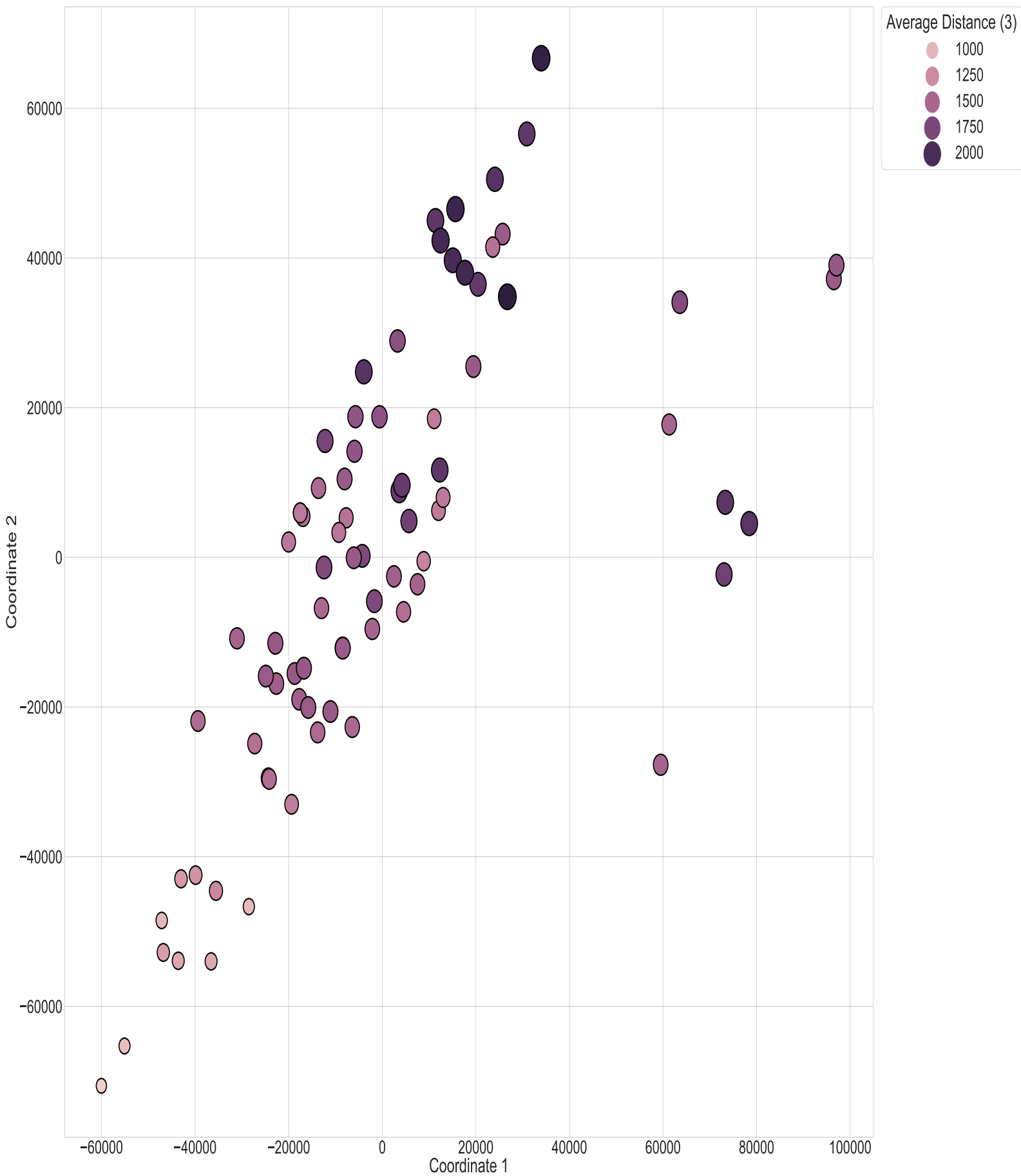


Scatter plot of NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using chebyshev pairwise distance.

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 :

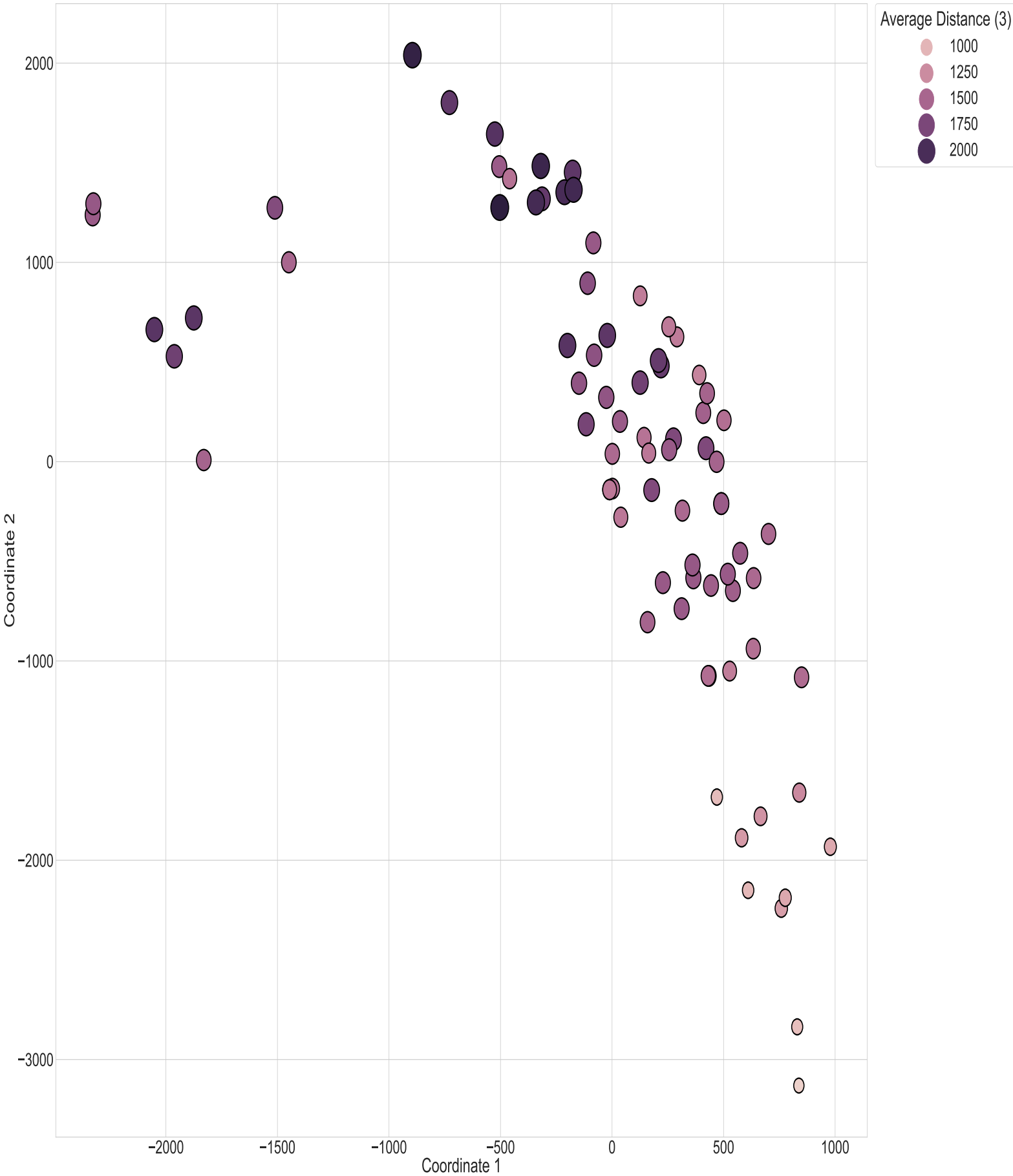


Scatter plot of NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using braycurtis pairwise distance.

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 :

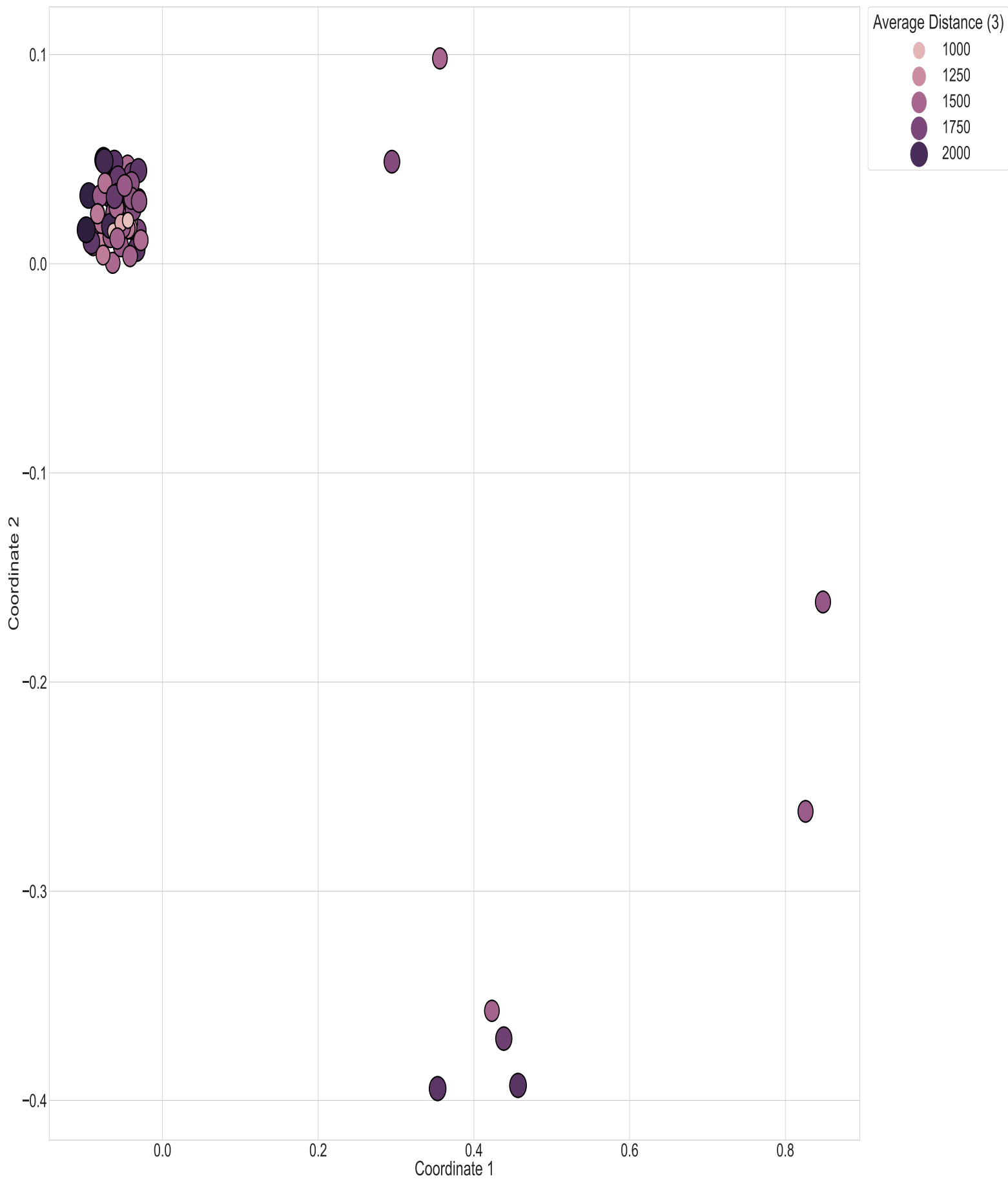


Scatter plot of NRPS *Micromonosporaceae* BGCs analyzed via multidimensional scaling using canberra pairwise distance.

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.

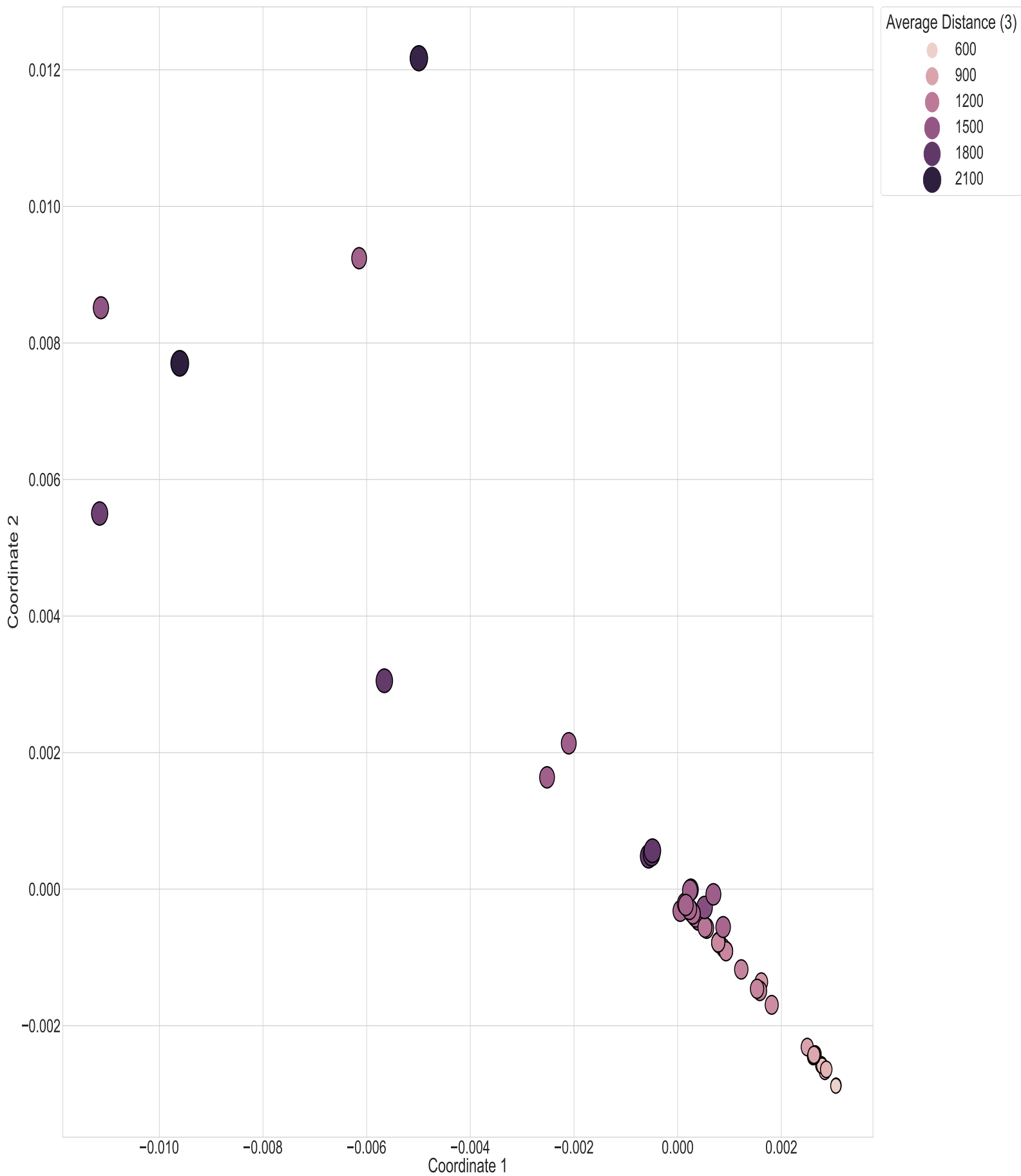
The hue and size of the dots were scaled based on the average distance of the individual NRPS BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

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 :

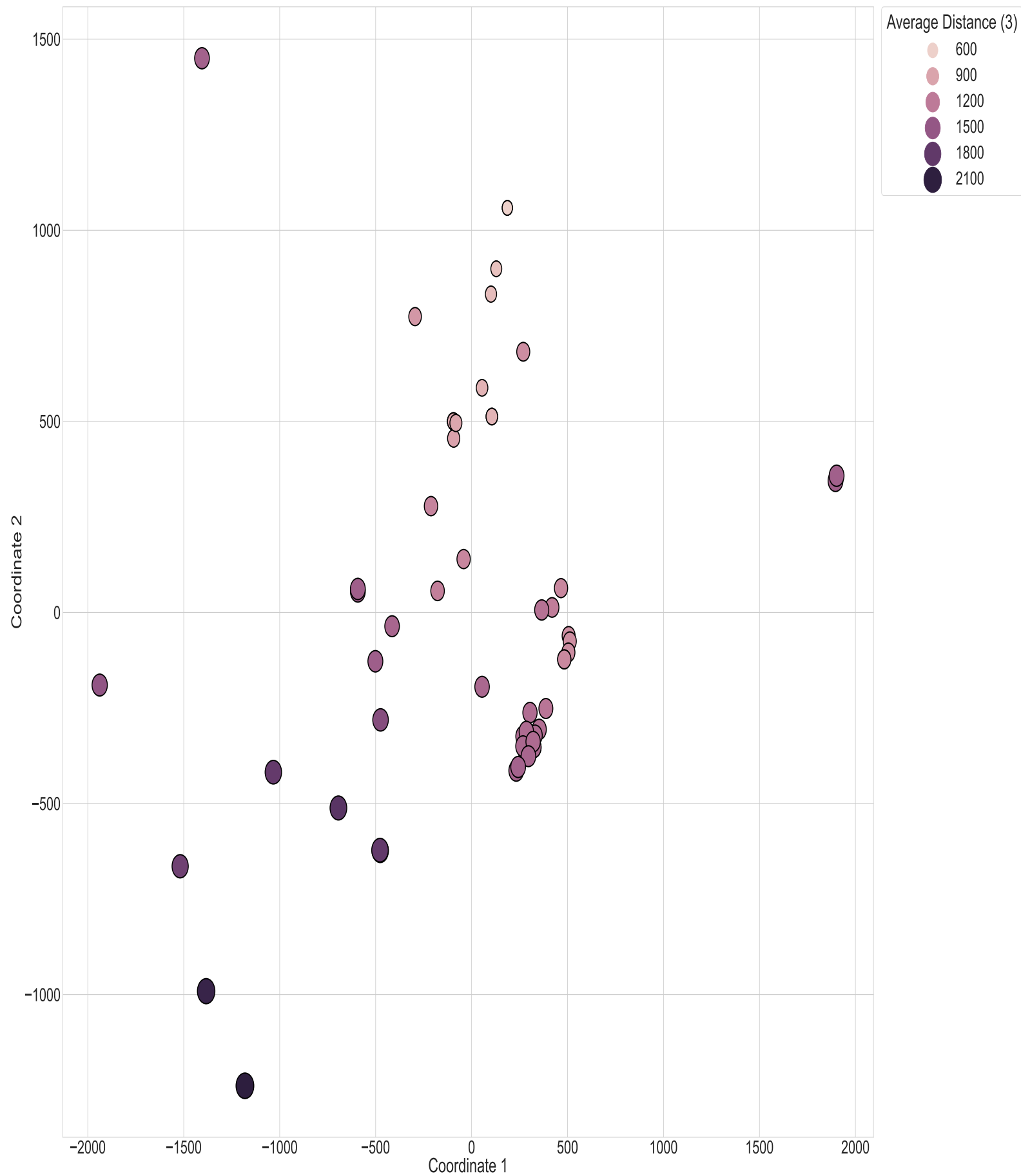


Scatter plot of Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cityblock pairwise distance.

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 I2 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 :

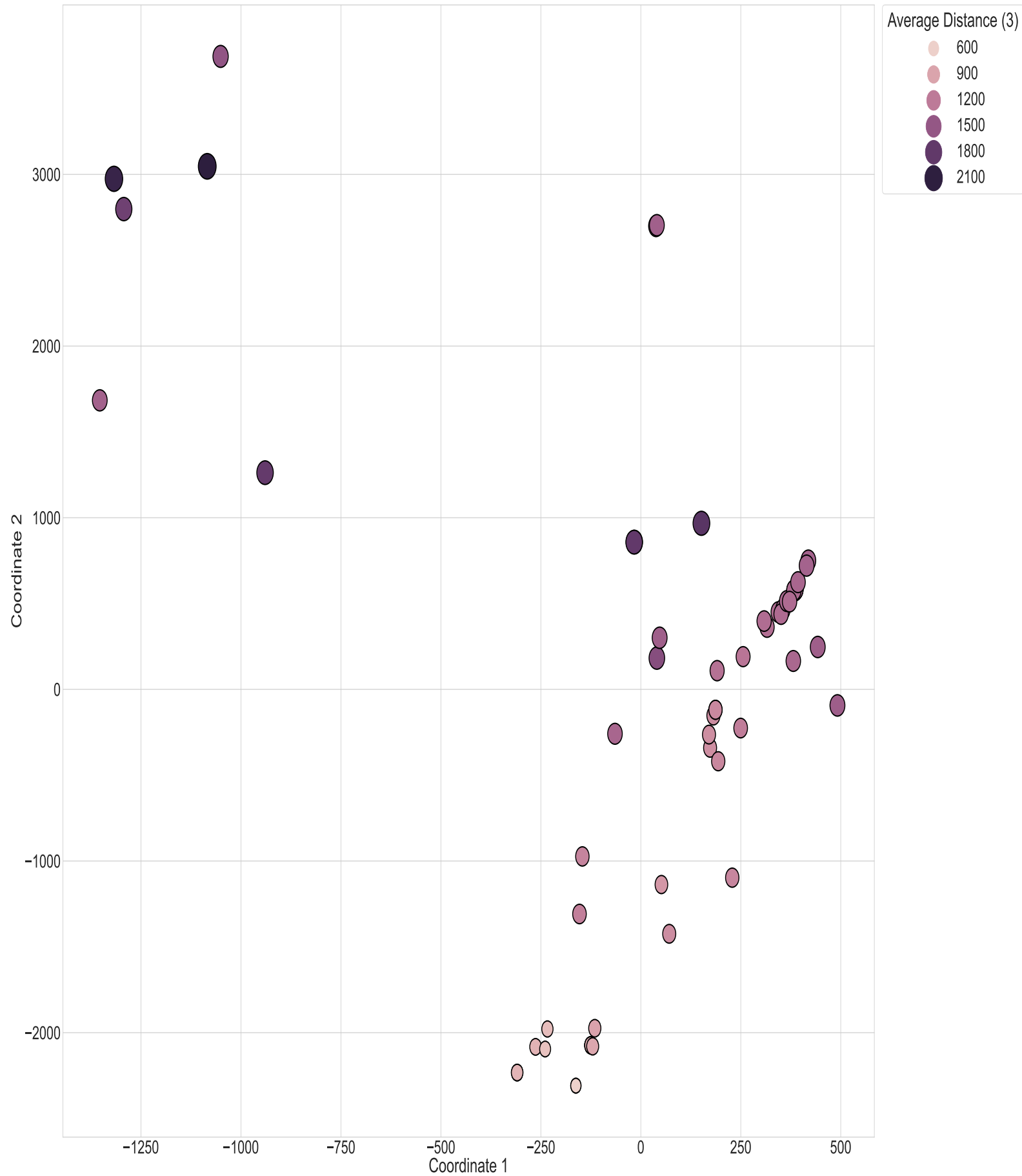


Scatter plot of Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using braycurtis pairwise distance.

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 :



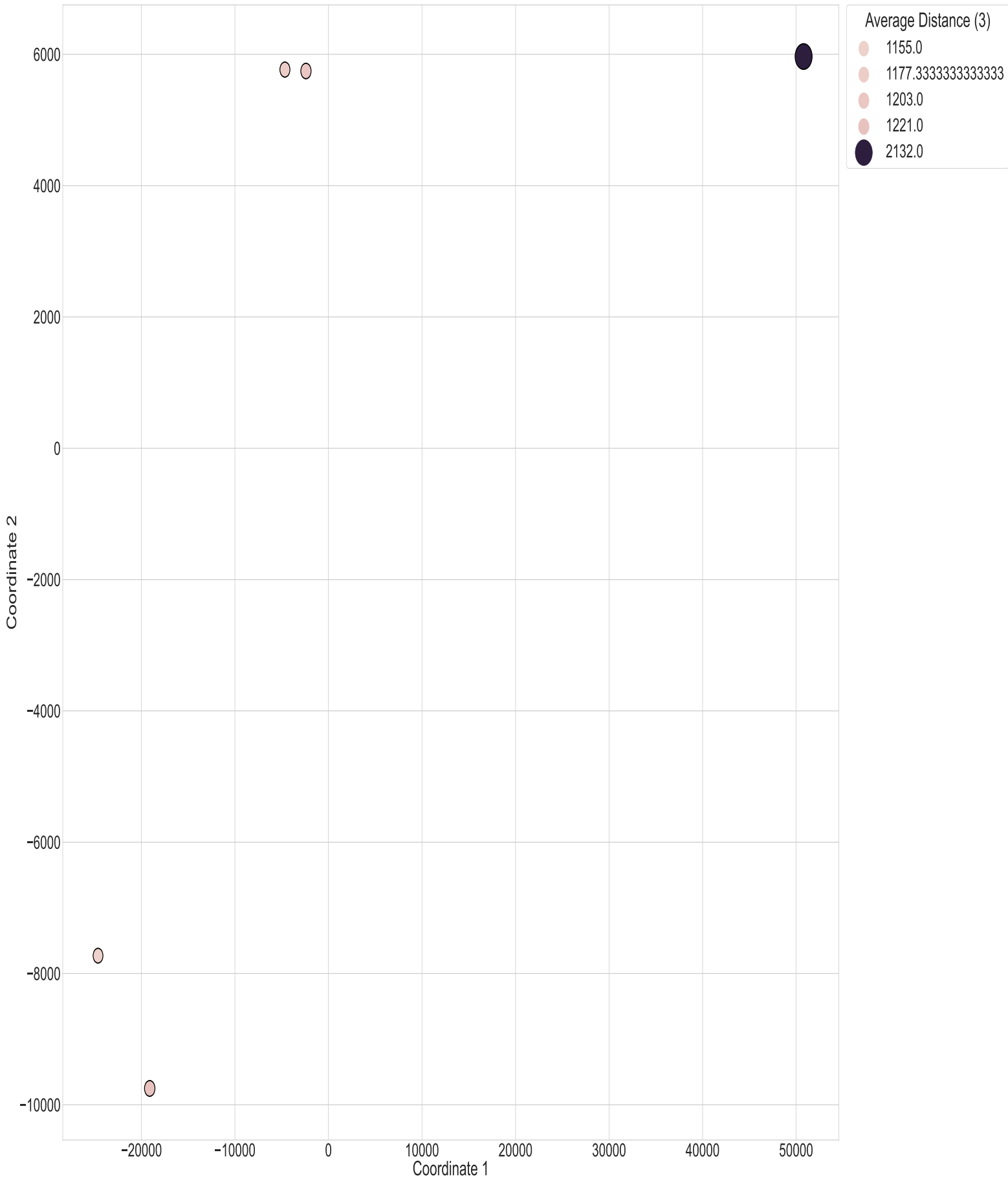
Scatter plot of Other *Micromonosporaceae* BGCs analyzed via multidimensional scaling using canberra pairwise distance. Only Other BGCs were considered for this analysis. Each dot represents an individual BGC. The hue and size of the dots were scaled based on the average distance of the individual Other 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 :

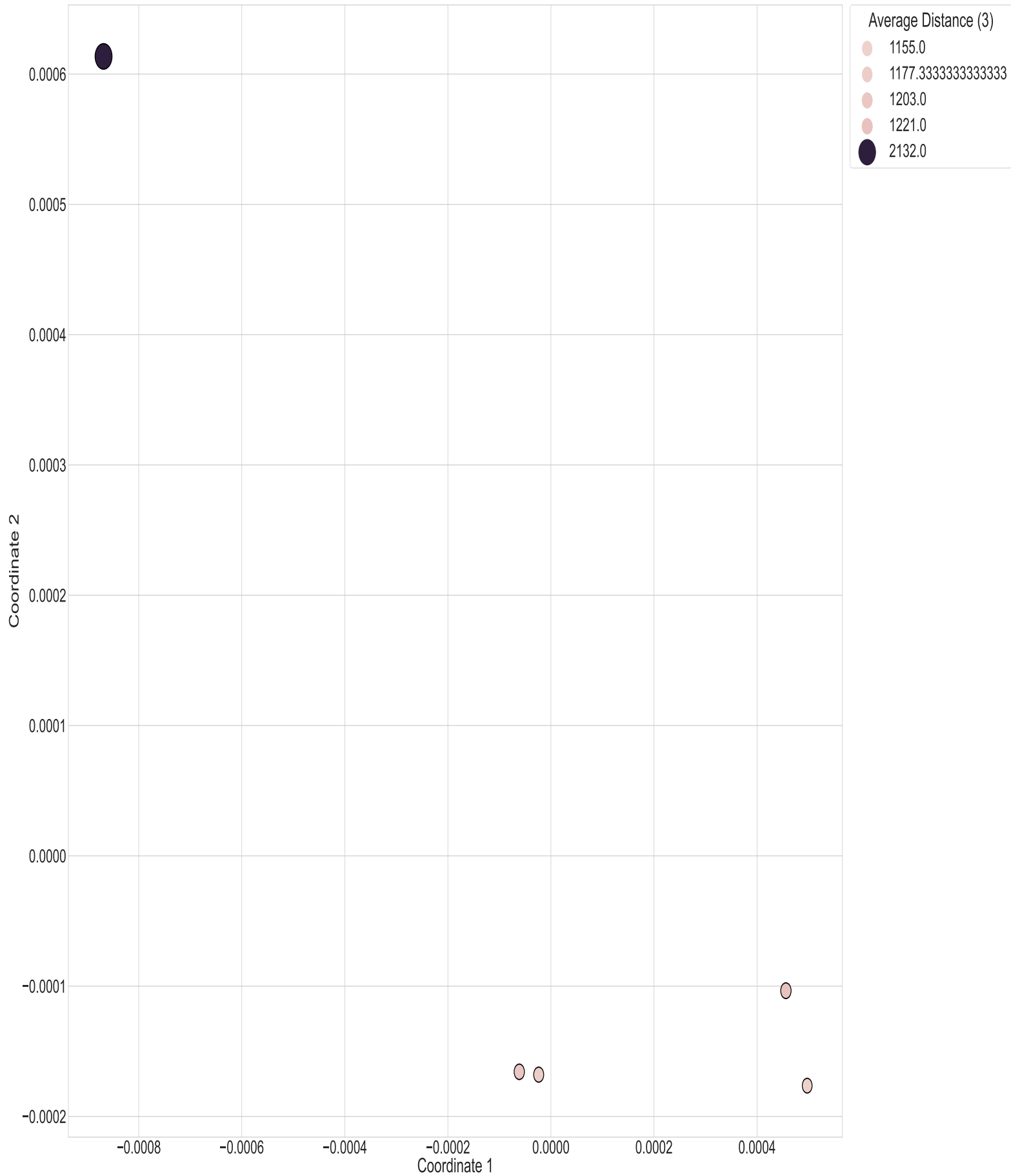


Scatter plot of Saccharides *Micromonosporaceae* BGCs analyzed via multidimensional scaling using euclidean pairwise distance.

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 :

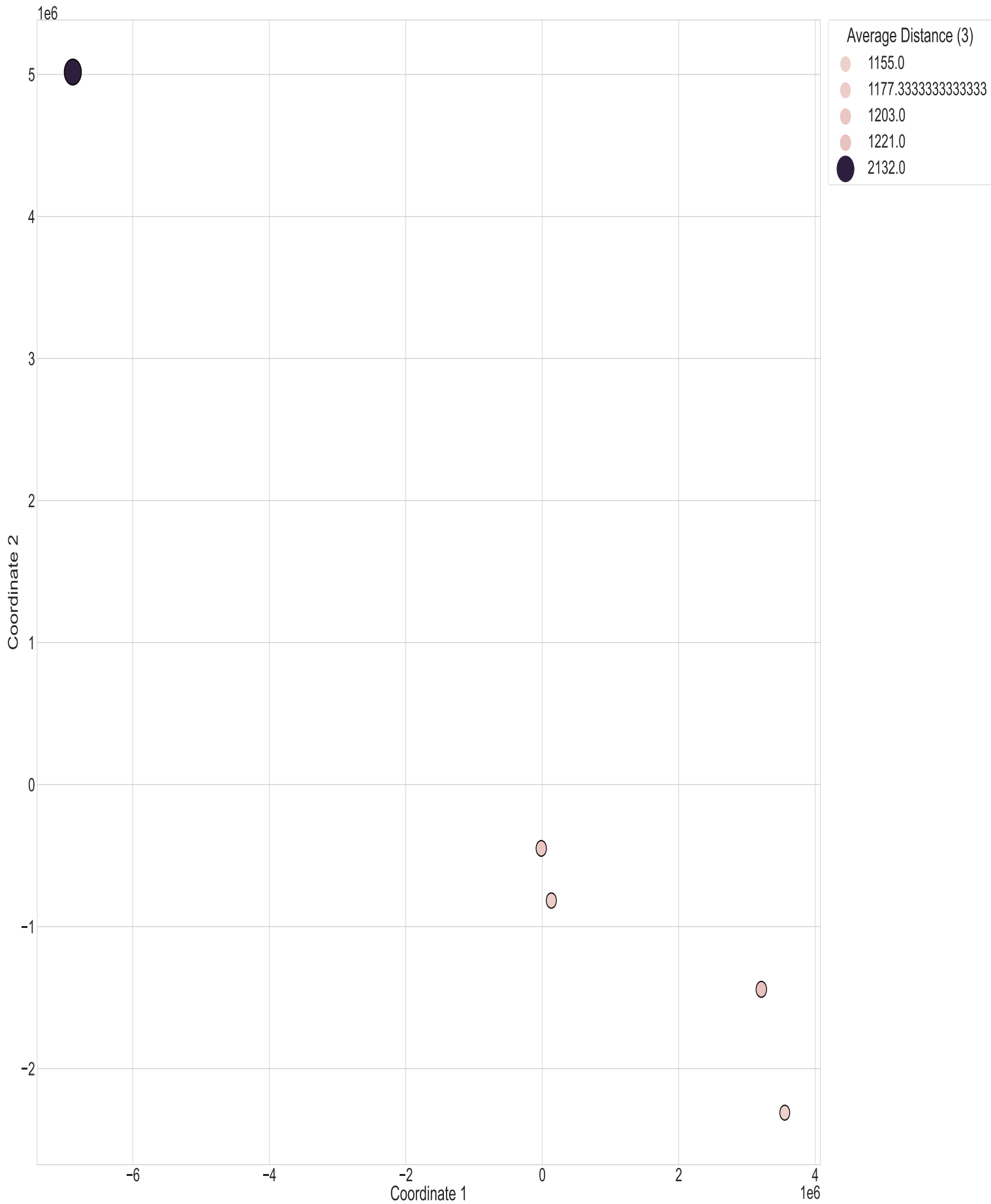


Scatter plot of Saccharides *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cosine pairwise distance.

Only Saccharides BGCs were considered for this analysis. Each dot represents an individual BGC.

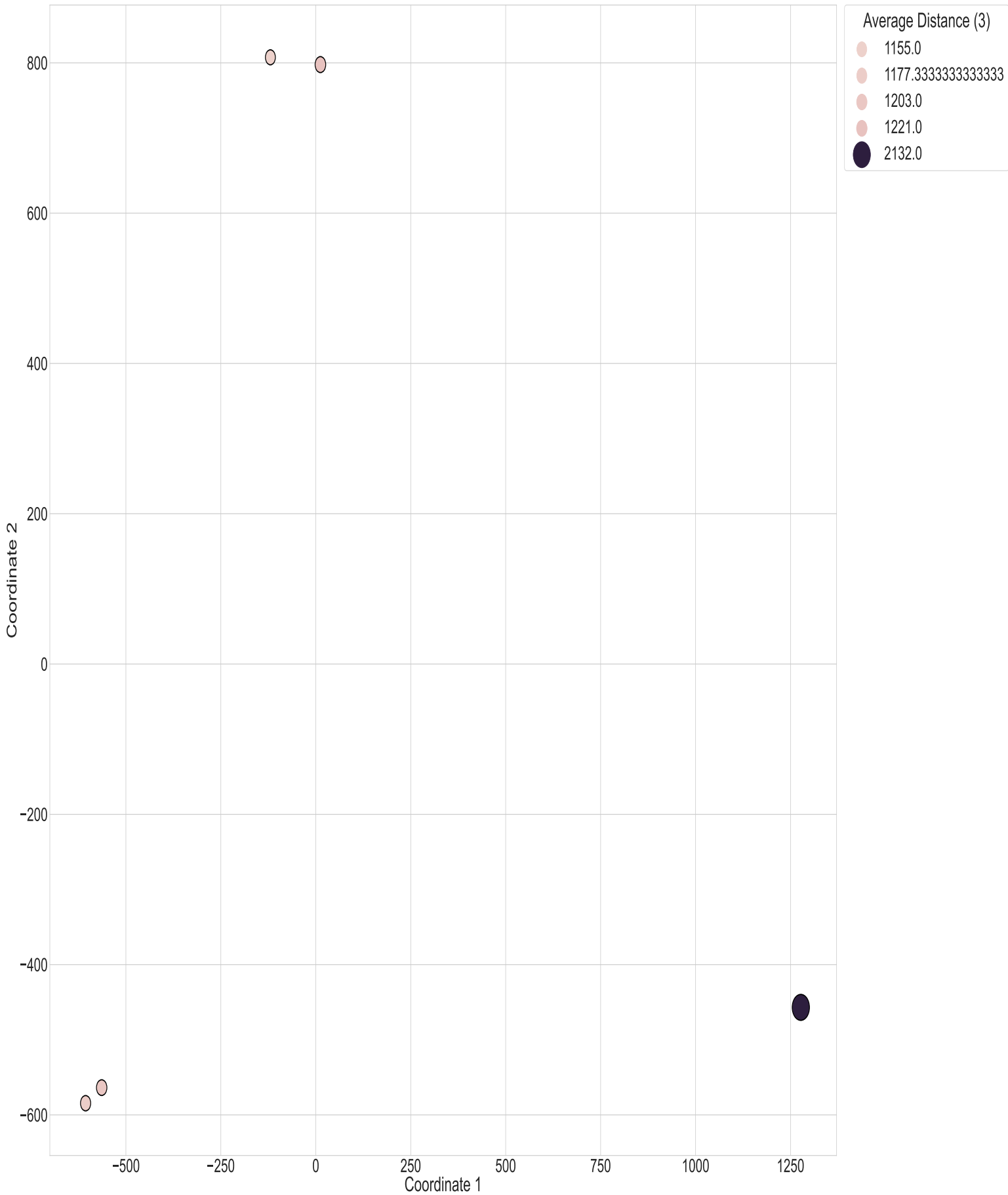
The hue and size of the dots were scaled based on the average distance of the individual Saccharides BGC to the closest 3 GCFs in BiG-SLICE (T=900).

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 :

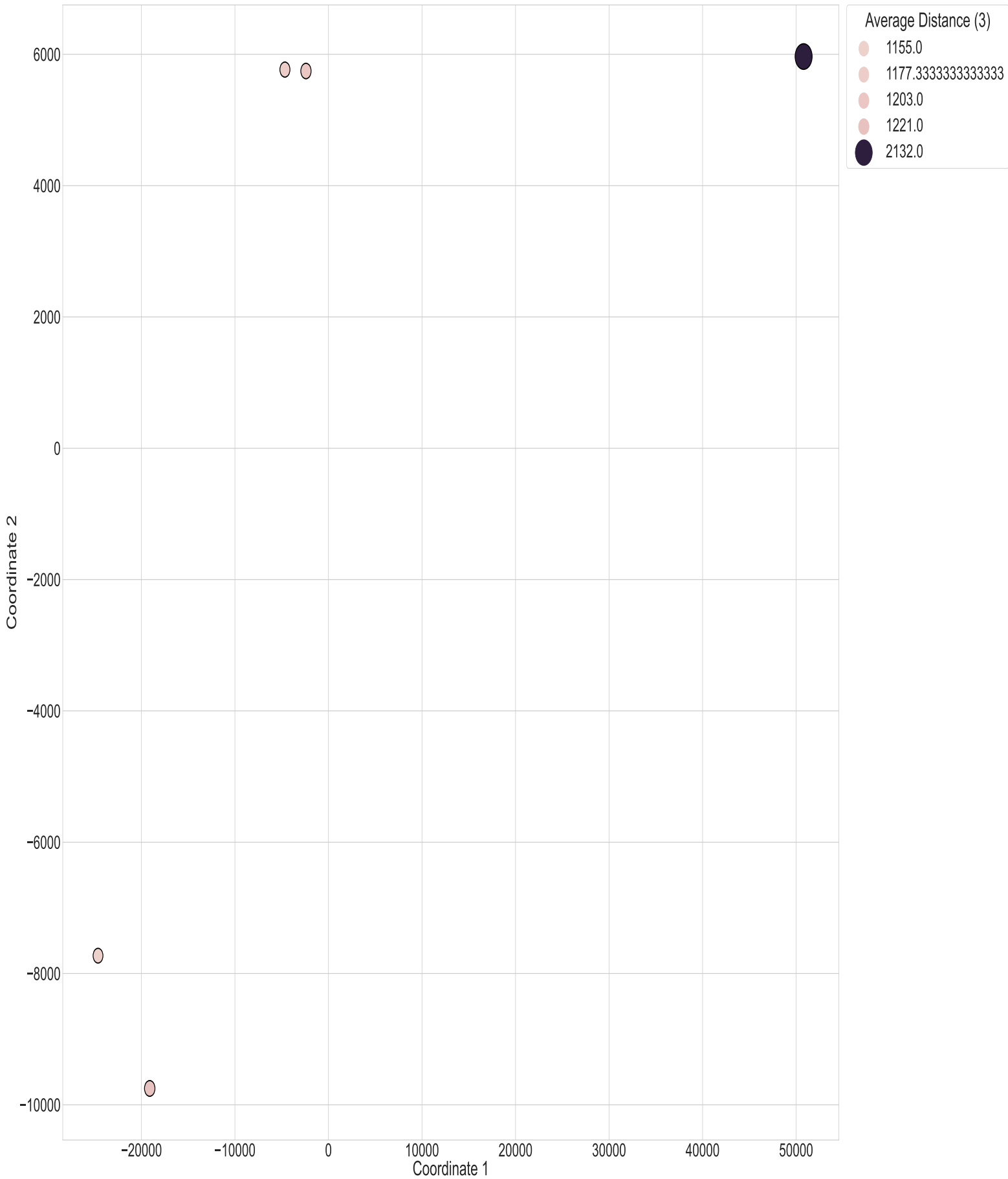


Scatter plot of Saccharides *Micromonosporaceae* BGCs analyzed via multidimensional scaling using chebyshev pairwise distance.

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 :

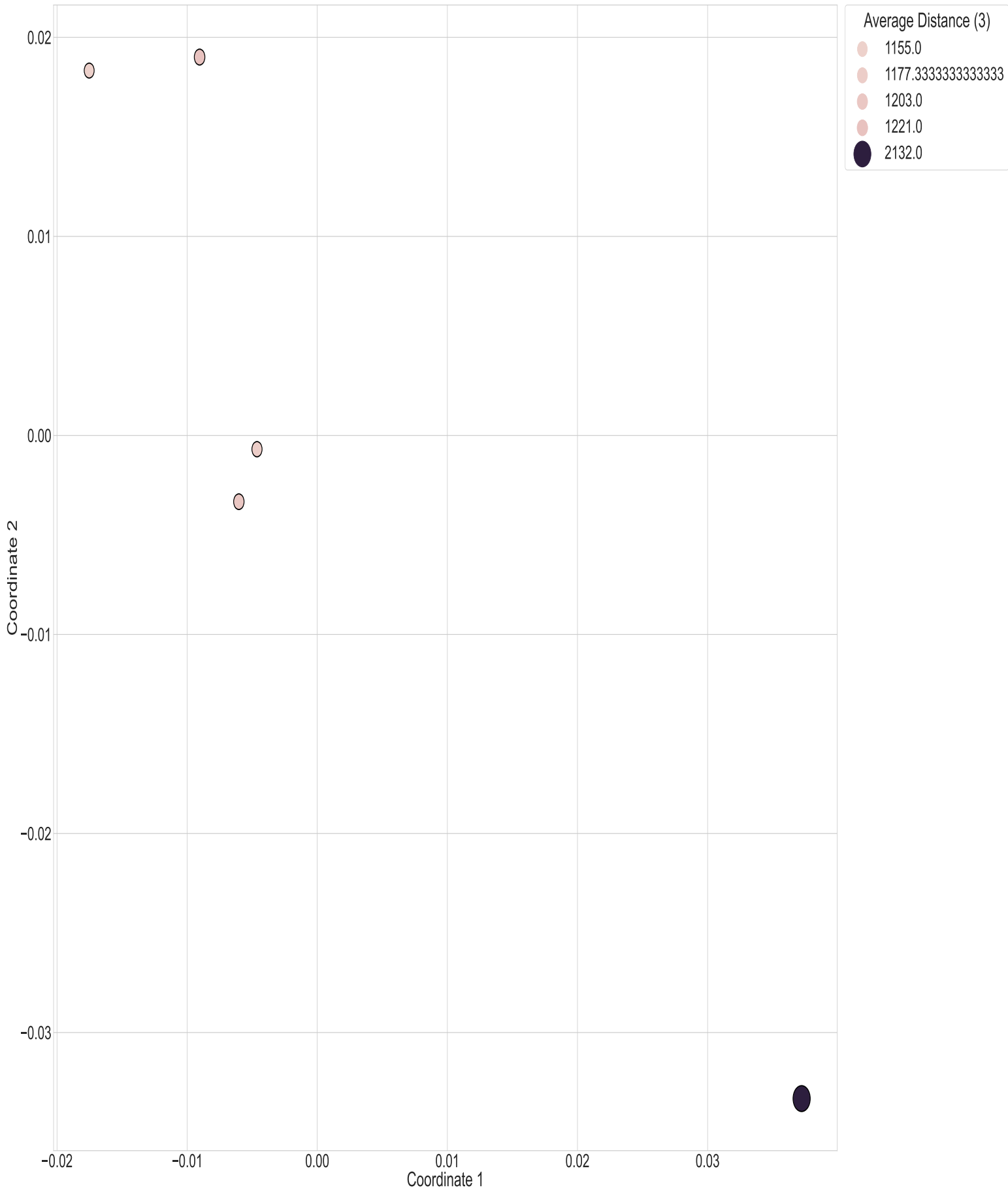


Scatter plot of Saccharides *Micromonosporaceae* BGCs analyzed via multidimensional scaling using l2 pairwise distance.

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 :

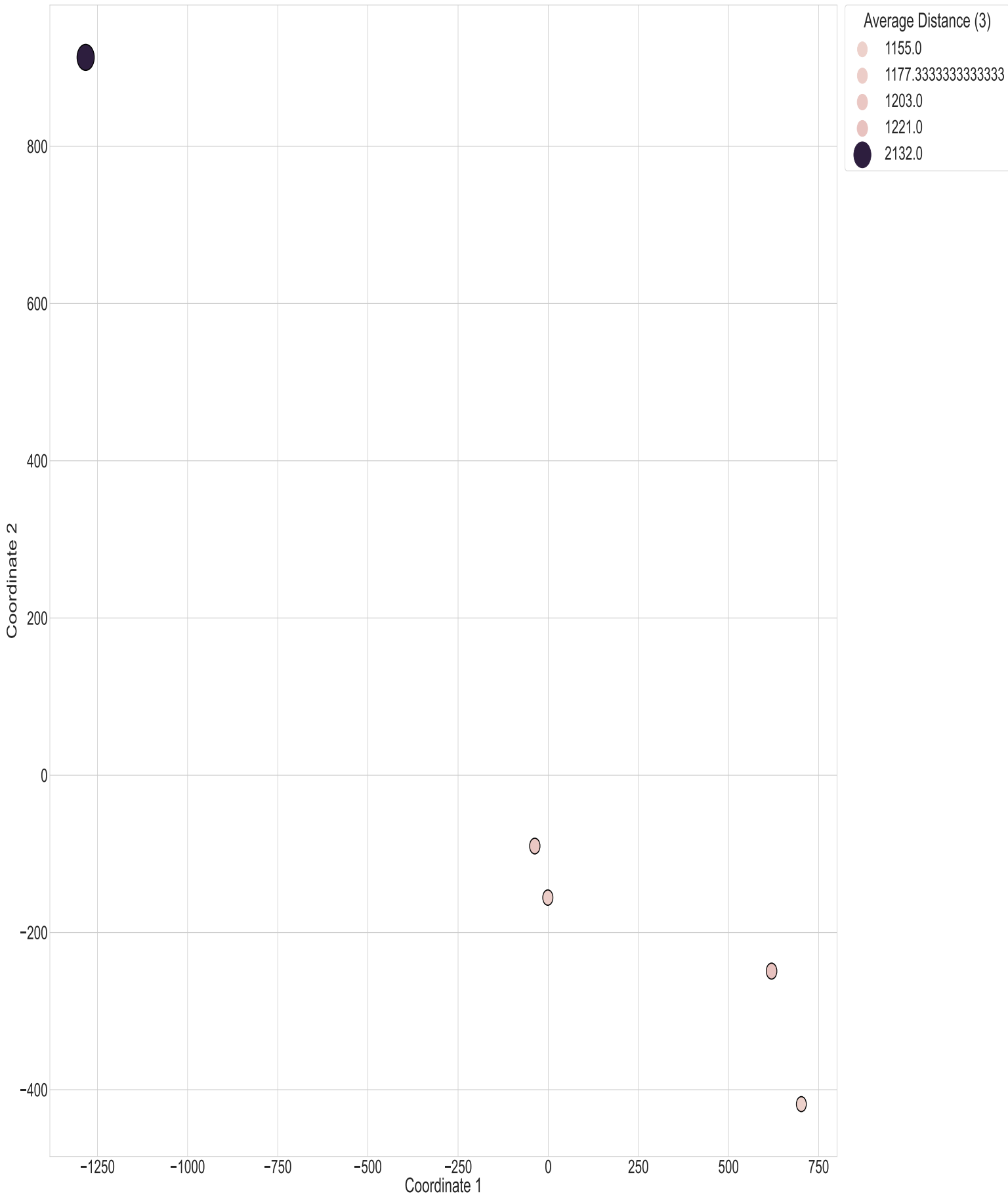


Scatter plot of Saccharides *Micromonosporaceae* BGCs analyzed via multidimensional scaling using braycurtis pairwise distance.

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:

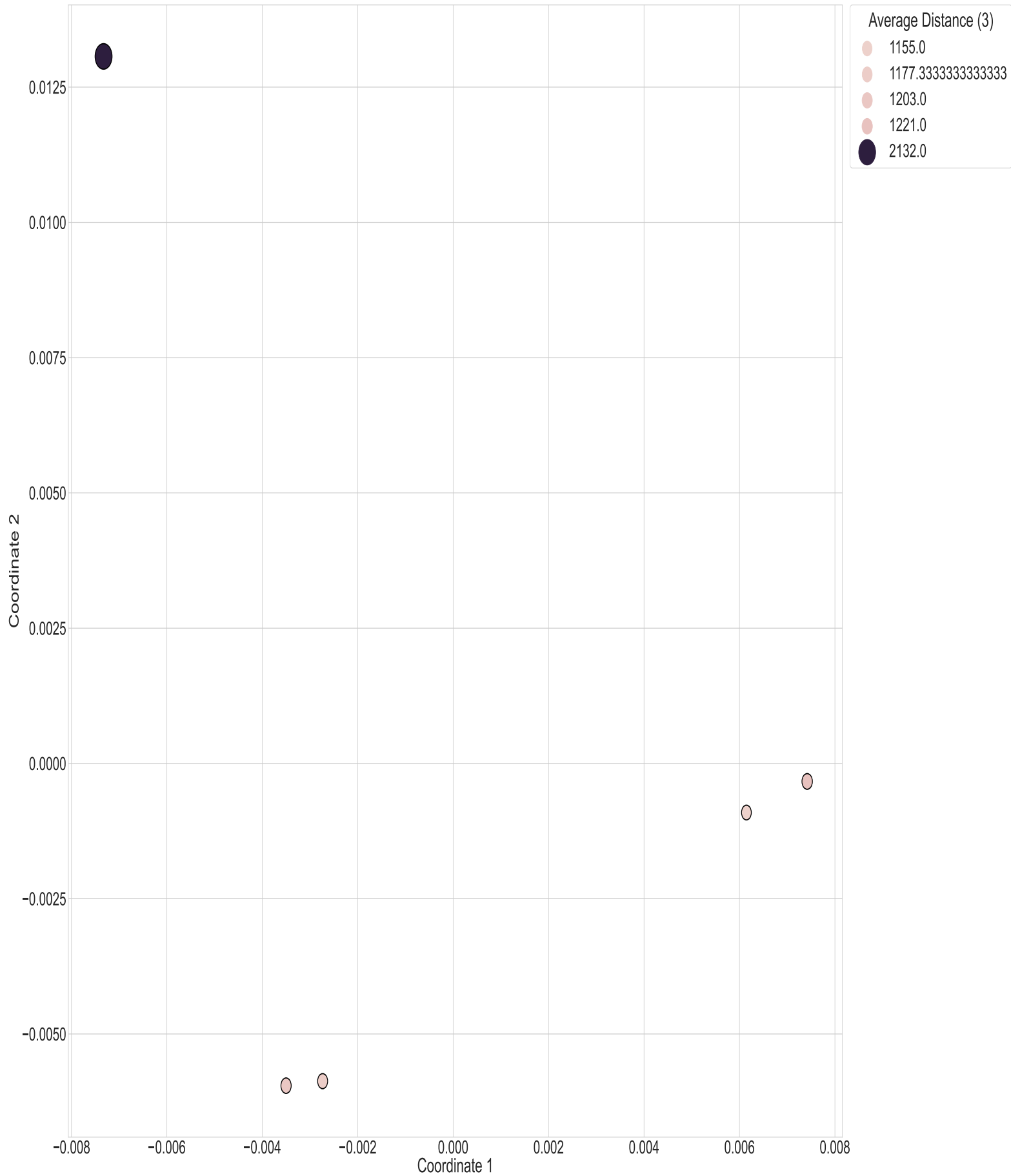


Scatter plot of Saccharides *Micromonosporaceae* BGCs analyzed via multidimensional scaling using canberra pairwise distance.

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 :

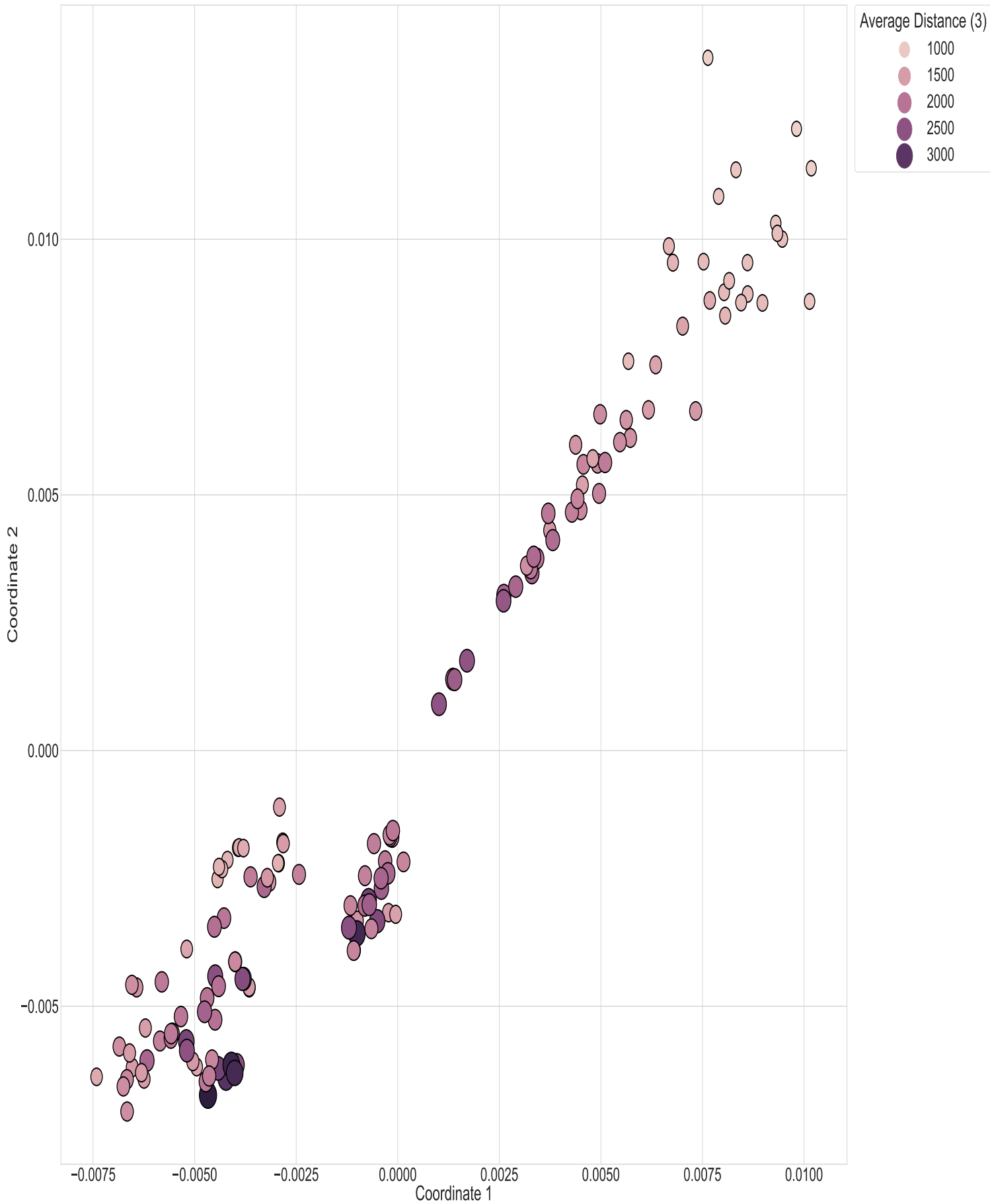


Scatter plot of Hybrid-(Not-PKS-NRPS) *Micromonosporaceae* BGCs analyzed via multidimensional scaling using euclidean pairwise distance.

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 :

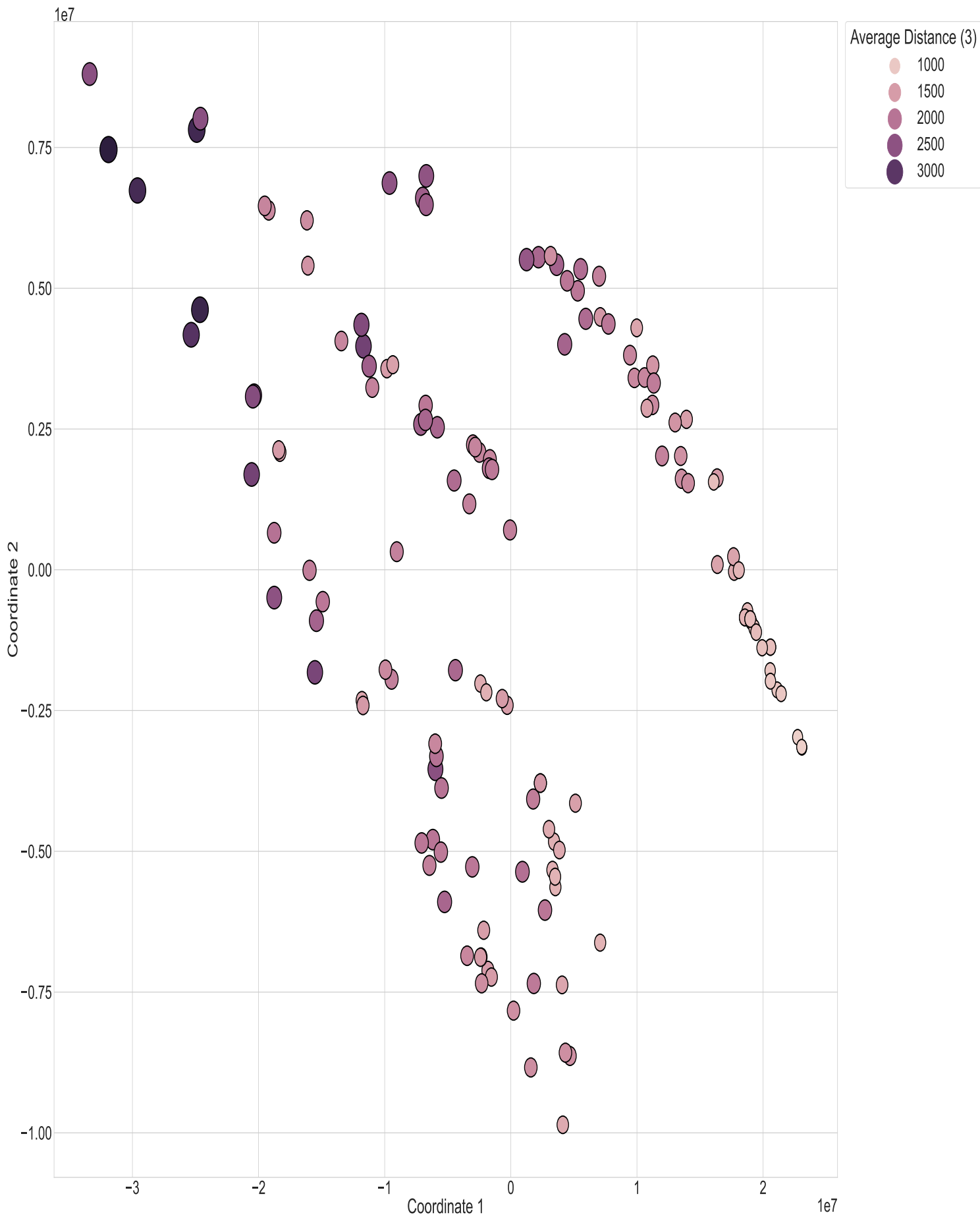


Scatter plot of Hybrid-(Not-PKS-NRPS) *Micromonosporaceae* BGCs analyzed via multidimensional scaling using cosine pairwise distance.

Only Hybrid-(Not-PKS-NRPS) BGCs were considered for this analysis. Each dot represents an individual BGC.

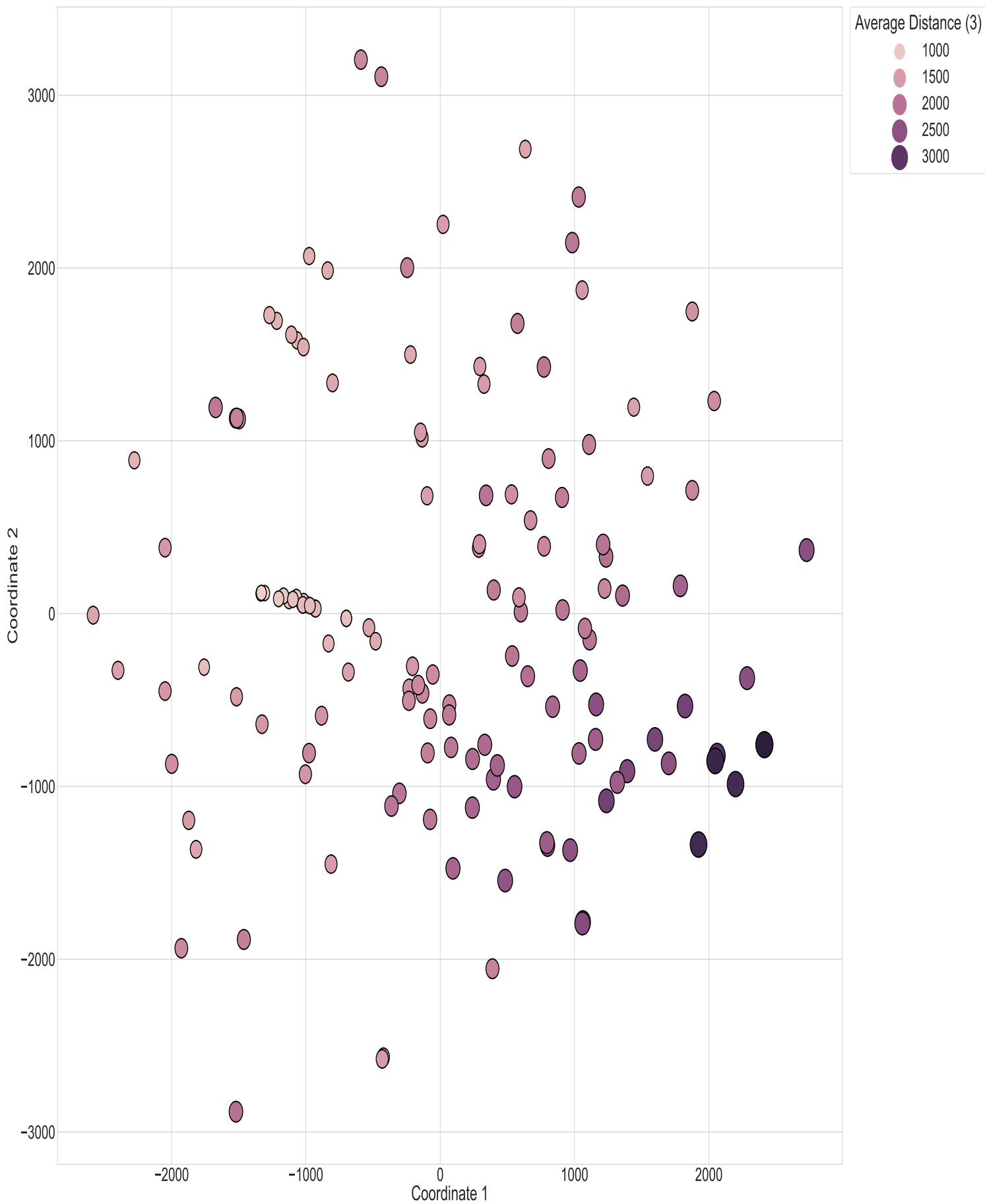
The hue and size of the dots were scaled based on the average distance of the individual Hybrid-(Not-PKS-NRPS) BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

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:

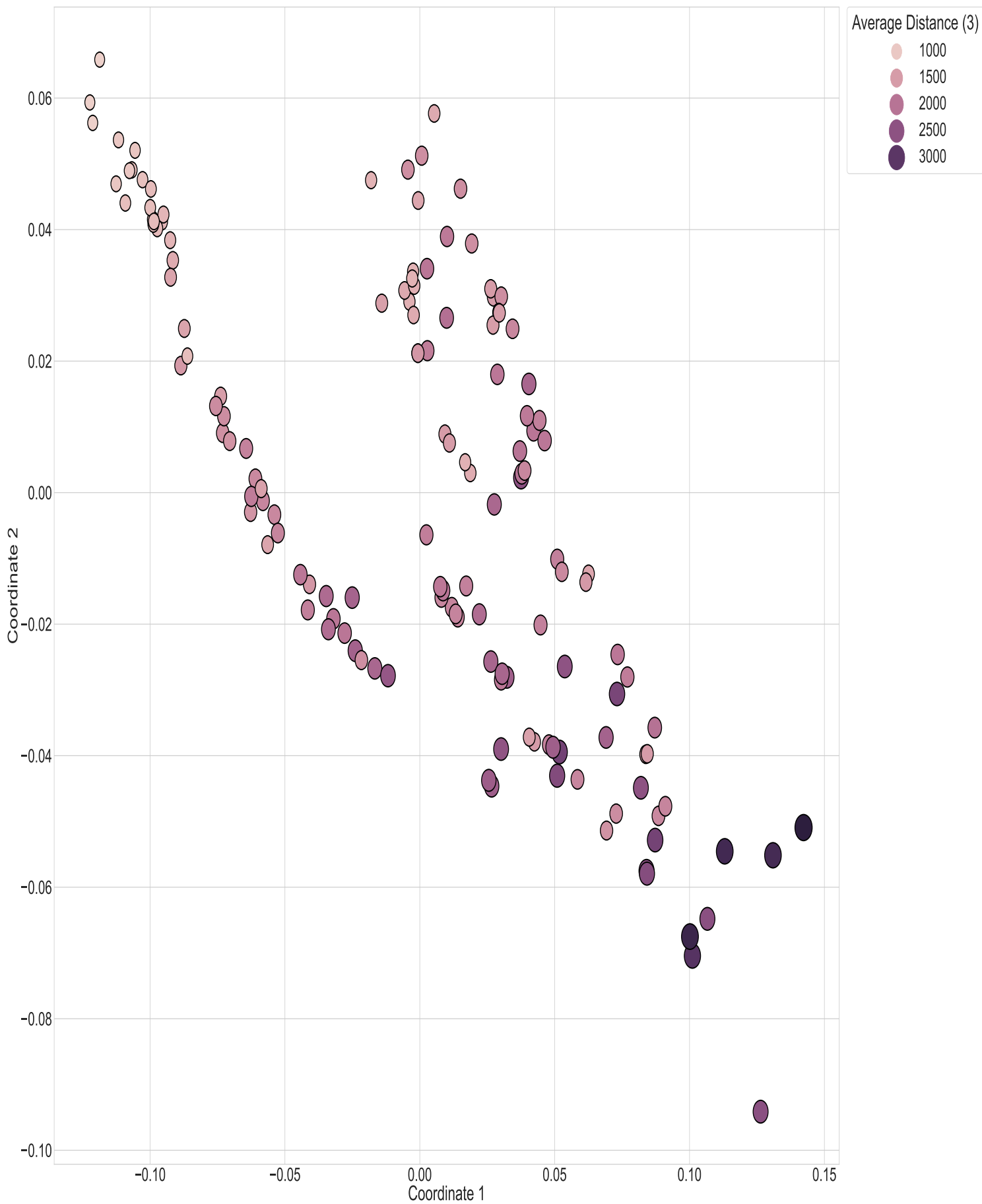


Scatter plot of Hybrid-(Not-PKS-NRPS) *Micromonosporaceae* BGCs analyzed via multidimensional scaling using l2 pairwise distance.

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:

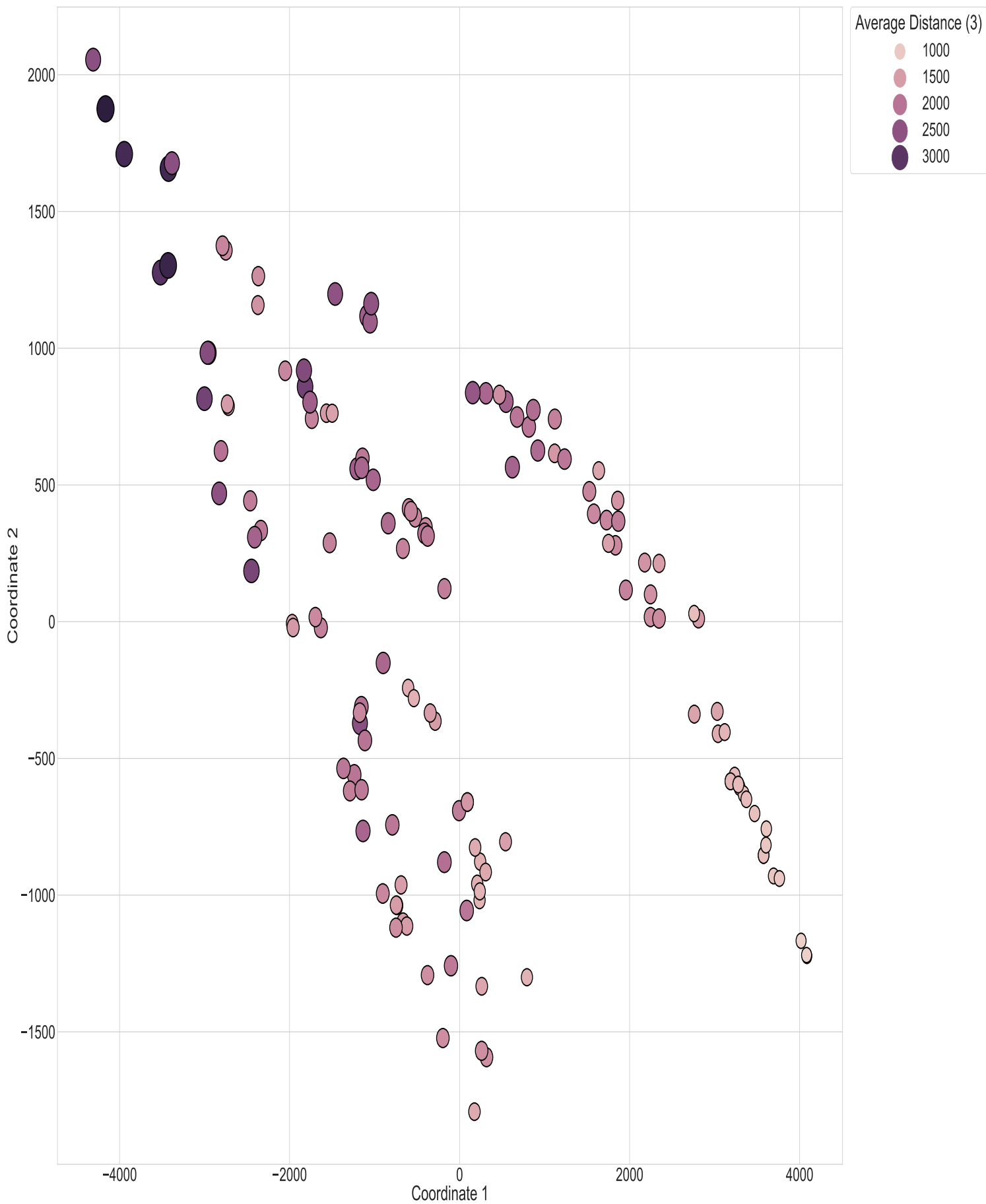


Scatter plot of Hybrid-(Not-PKS-NRPS) *Micromonosporaceae* BGCs analyzed via multidimensional scaling using braycurtis pairwise distance.

Only Hybrid-(Not-PKS-NRPS) BGCs were considered for this analysis. Each dot represents an individual BGC.

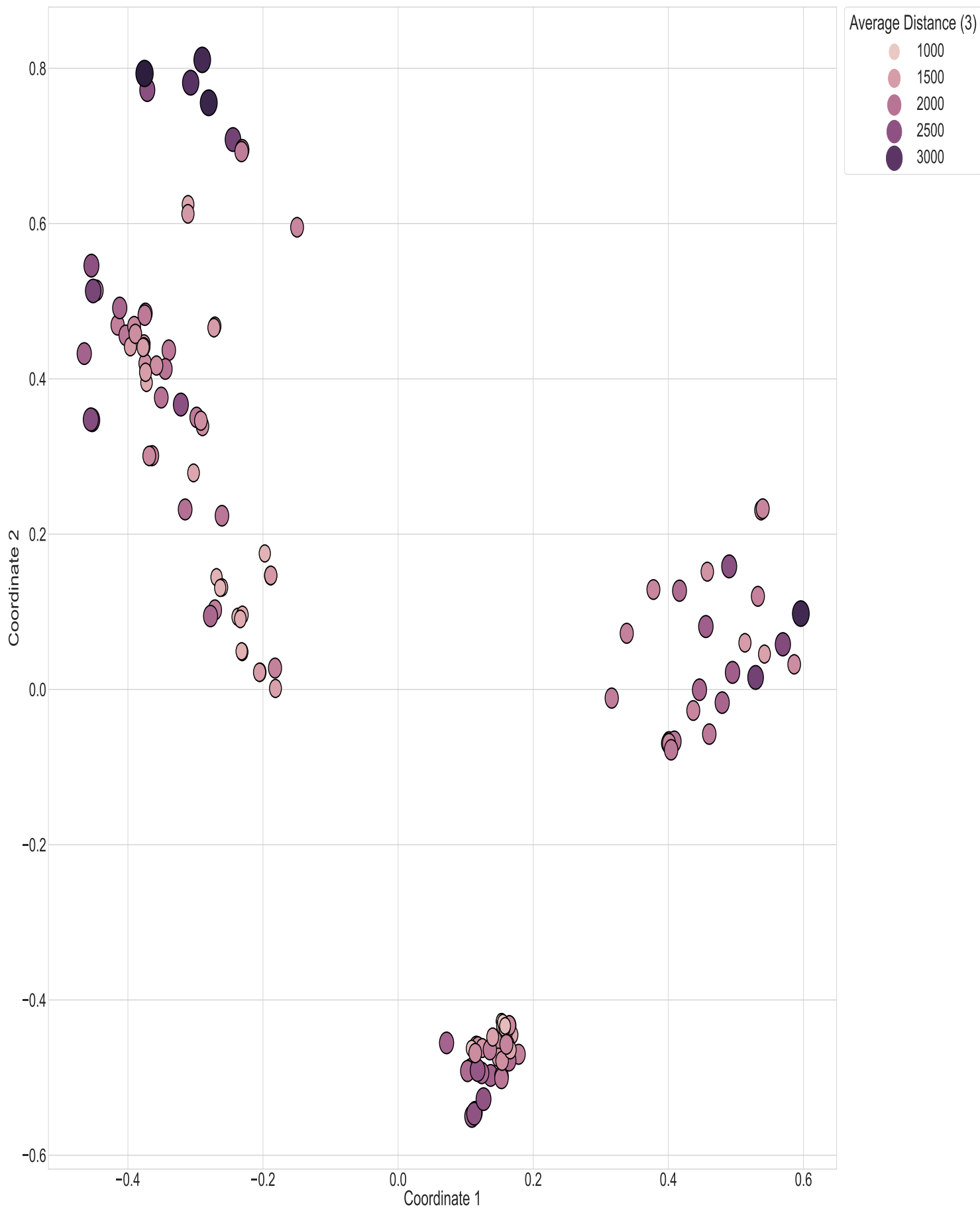
The hue and size of the dots were scaled based on the average distance of the individual Hybrid-(Not-PKS-NRPS) BGC to the closest 3 GCFs in BiG-SLiCE (T=900).

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 :

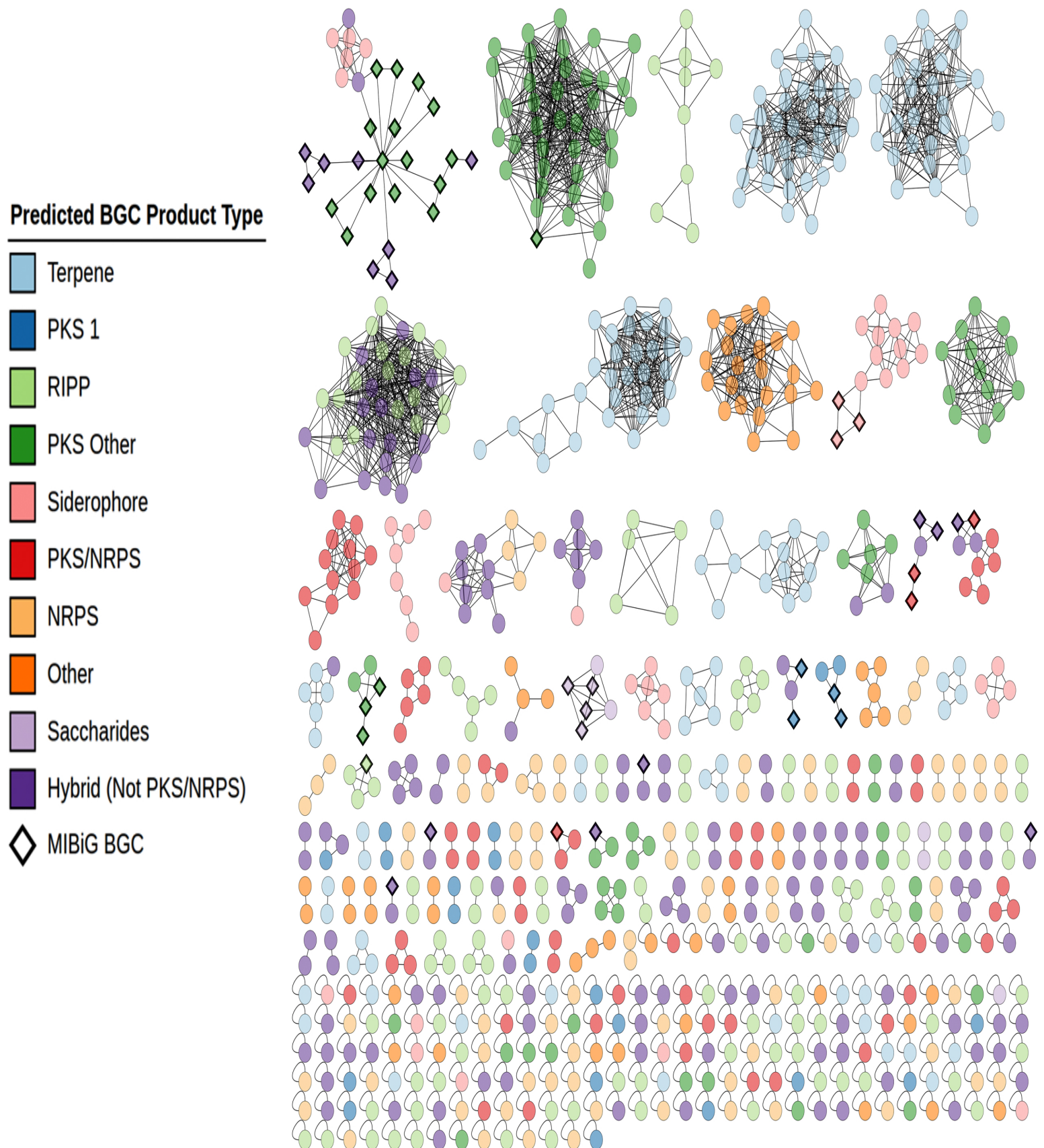


Scatter plot of Hybrid-(Not-PKS-NRPS) *Micromonosporaceae* BGCs analyzed via multidimensional scaling using correlation pairwise distance.

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.