Electronic Supplementary Information (ESI)

How different are marine microbial natural products compared to their terrestrial counterparts?

Tanja M. Voser [ab], Max D. Campbell [ac] and Anthony R. Carroll [*ab]

[a.] *School of Environment and Science, Griffith University, Gold Coast, Australia.*
[b.] *Griffith Institute for Drug Discovery, Griffith University, Brisbane, Australia.*
[c.] *Australian Rivers Institute - Coasts and Estuaries, Griffith University, Nathan, Australia.*

Comparison Between the Three Different Types of Analysis

Overlap Between Marine and Terrestrial Microbial NPs

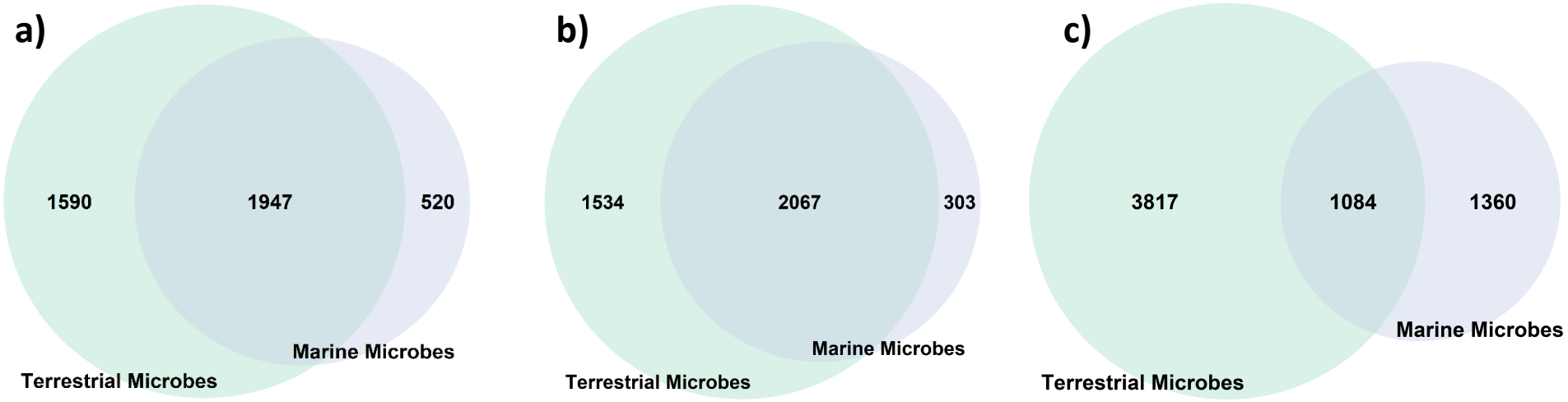


Fig. S1 a) Overlap between marine (blue) and terrestrial (green) microbial molecular clusters. The three different graphs represent the different analysis performed. a) shows the results from the extended fingerprint cluster analysis (which resulted in a total of 5000 clusters), b) shows the results of the fingerprint cluster analysis using the PubChem fingerprint (with the total of 5000 clusters) and c) shows the results of the scaffold analysis, where the NPs were divided into 9,195 scaffolds.

Overlap between the Three Different Groups (Clusters)

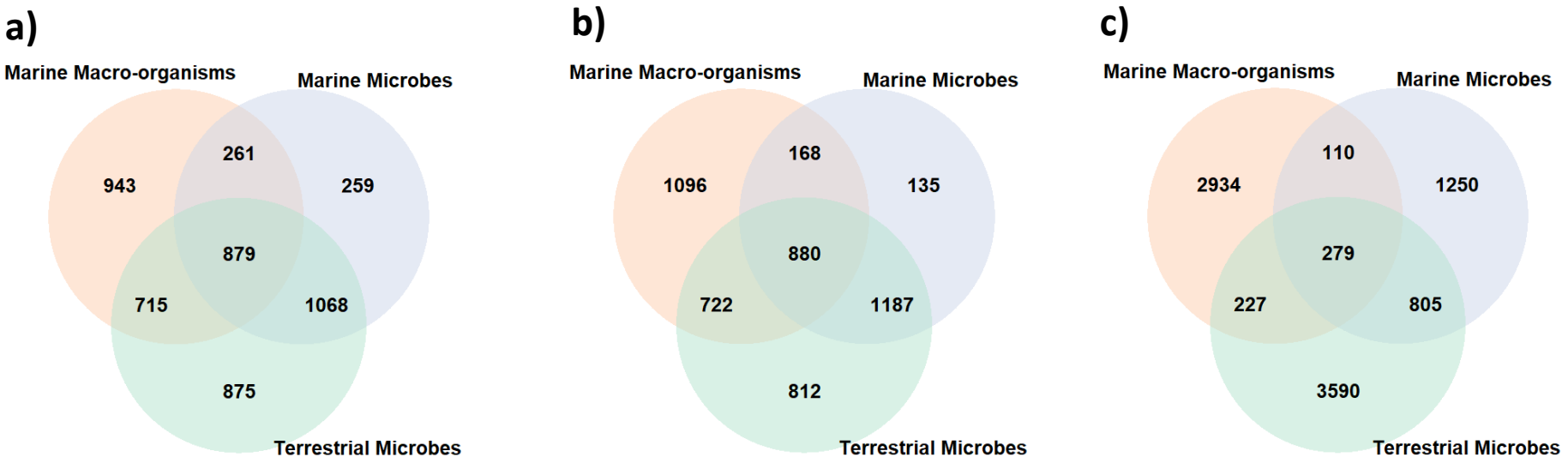


Fig. S2 a) Nesting of NPs from marine macro-organisms (orange), marine microorganisms (blue) and terrestrial microorganisms (green). Total number of clusters in each section is given in black. The three different graphs represent the different analysis performed. a) shows the results from the extended fingerprint cluster analysis (resulting in a total of 5000 clusters), b) shows the results of the fingerprint cluster analysis using the PubChem fingerprint (with the total of 5000 clusters) and c) shows the results of the scaffold analysis, where the NPs were divided into 9,195 scaffolds.

Cluster Analysis on Extended Fingerprint (5000 Clusters)

Overlap Between Marine and Terrestrial Microbial NPs

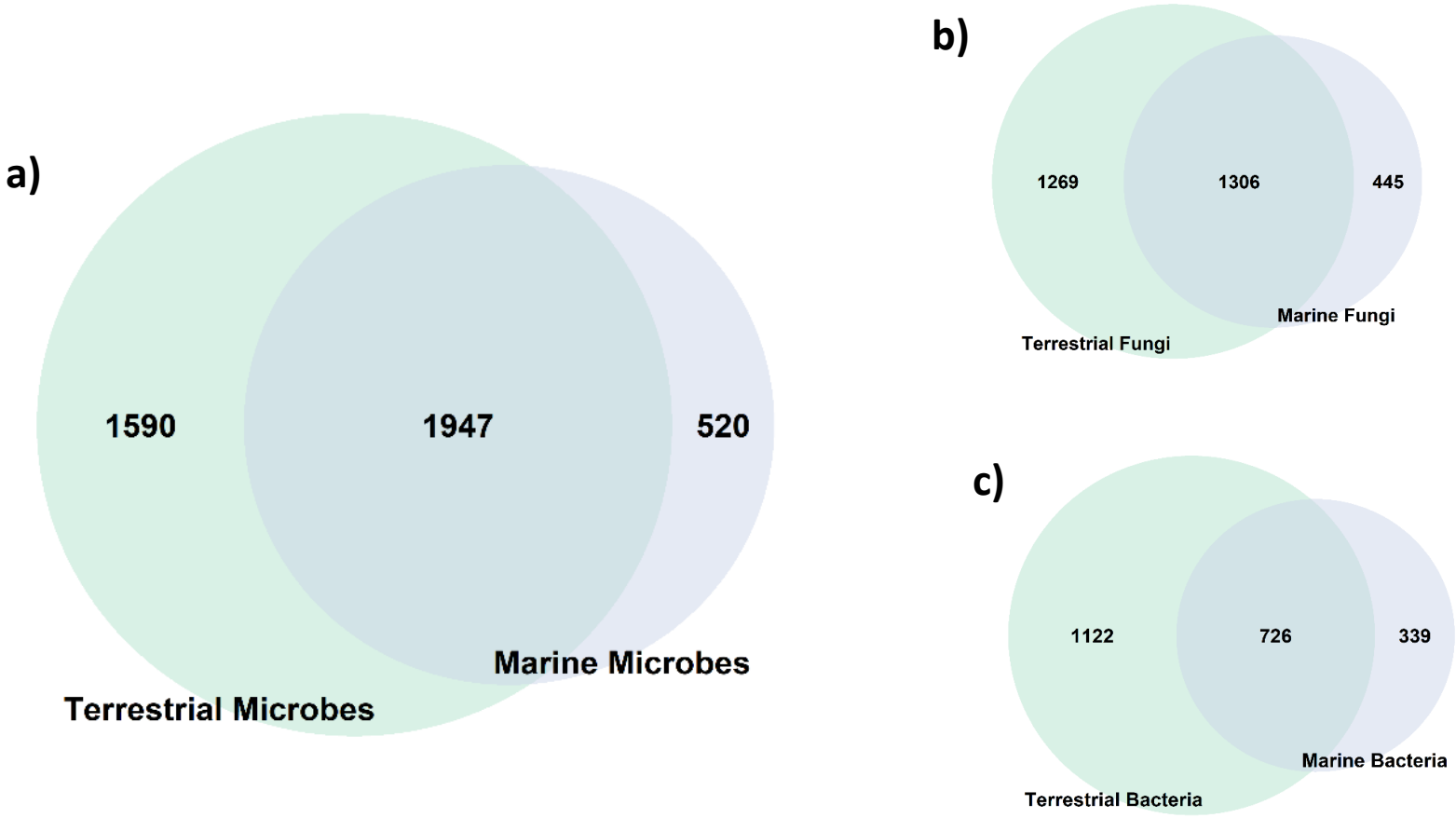


Fig. S3 a) Overlap between marine (blue) and terrestrial (green) microbial molecular clusters. It shows that 78.9% of the marine natural products clusters are nested amongst terrestrial clusters. Numbers indicate the total number of clusters in each section. b) Overlap between the marine (blue) and terrestrial (green) fungal NPs, with numbers representing clusters. c) Overlap of the marine (blue) and terrestrial (green) bacterial NPs.

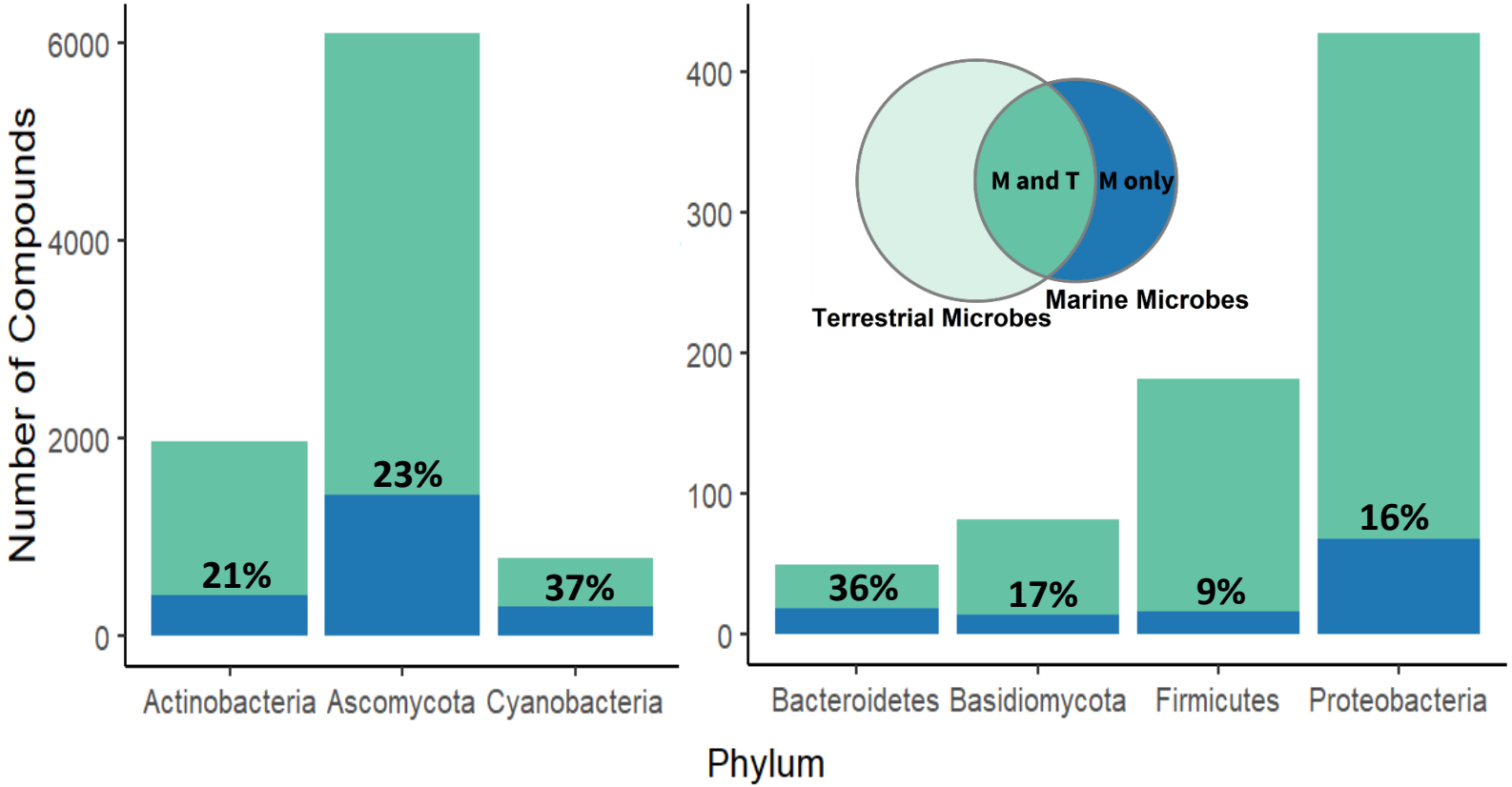


Fig. S4 Total Number and Percentage of uniquely marine microbial NPs (blue) produced by organisms in different phyla *vs* marine microbial NPs that overlap with terrestrial microbial NPs (green).

Overlap between the Three Different Groups (Clusters)

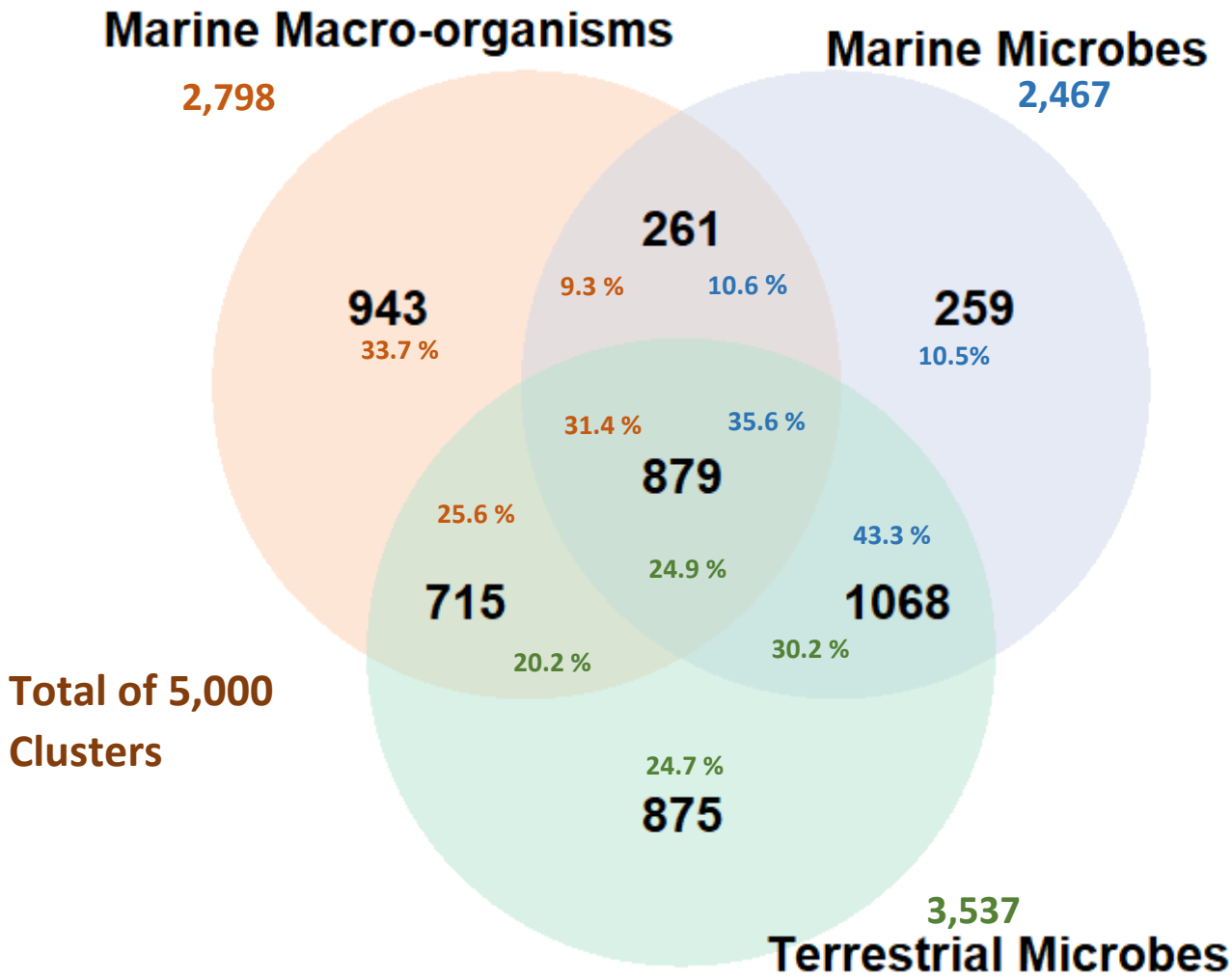


Fig. S5 Nesting of NPs from marine macro-organisms (orange), marine microorganisms (blue) and terrestrial microorganisms (green). Total number of clusters in each section is given in black. Smaller numbers represent percentages of total clusters for each biota group. Blue numbers for example relate to the percentage of marine microbial NPs clusters in each section as a proportion of the total 2467 marine microbial NP clusters

Overlap between the Three Different Groups (Compounds)

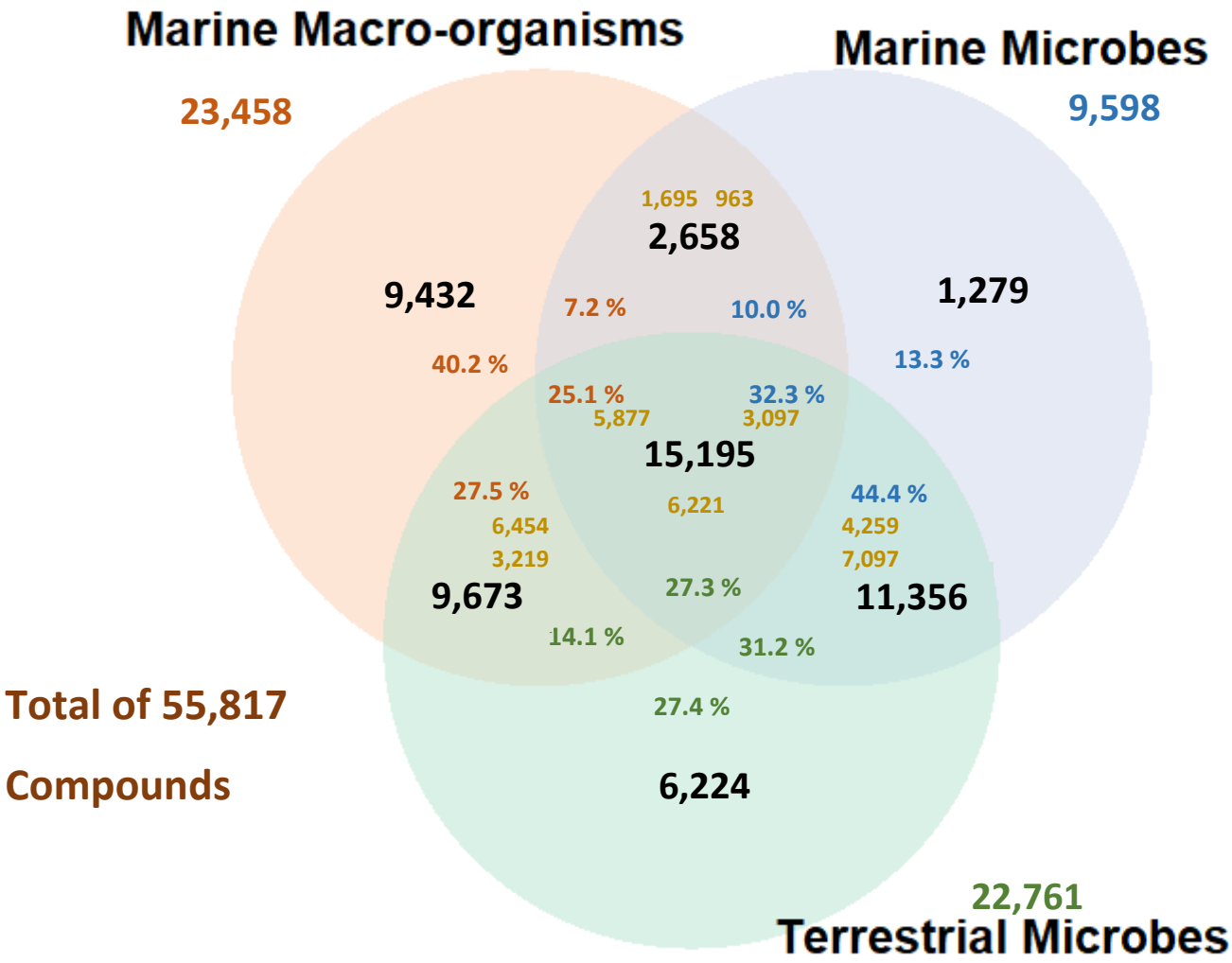


Fig. S6 Nesting of NPs from marine macro-organisms (orange), marine microorganisms (blue) and terrestrial microorganisms (green). Total number of NPs in each section is given in black. Smaller numbers represent percentages of total NPs for each biota group. Blue numbers for example relate to the percentage of marine microbial NPs in each section as a proportion of the total 9598 marine microbial NPs. The small orange numbers represent total NPs from each biota group in each intersection. For example, 3219 in the terrestrial microorganism/marine macro-organism intersection represents terrestrial microbial NPs.

Overlap between the Three Different Groups (Compounds & Accounting for Halogenation)

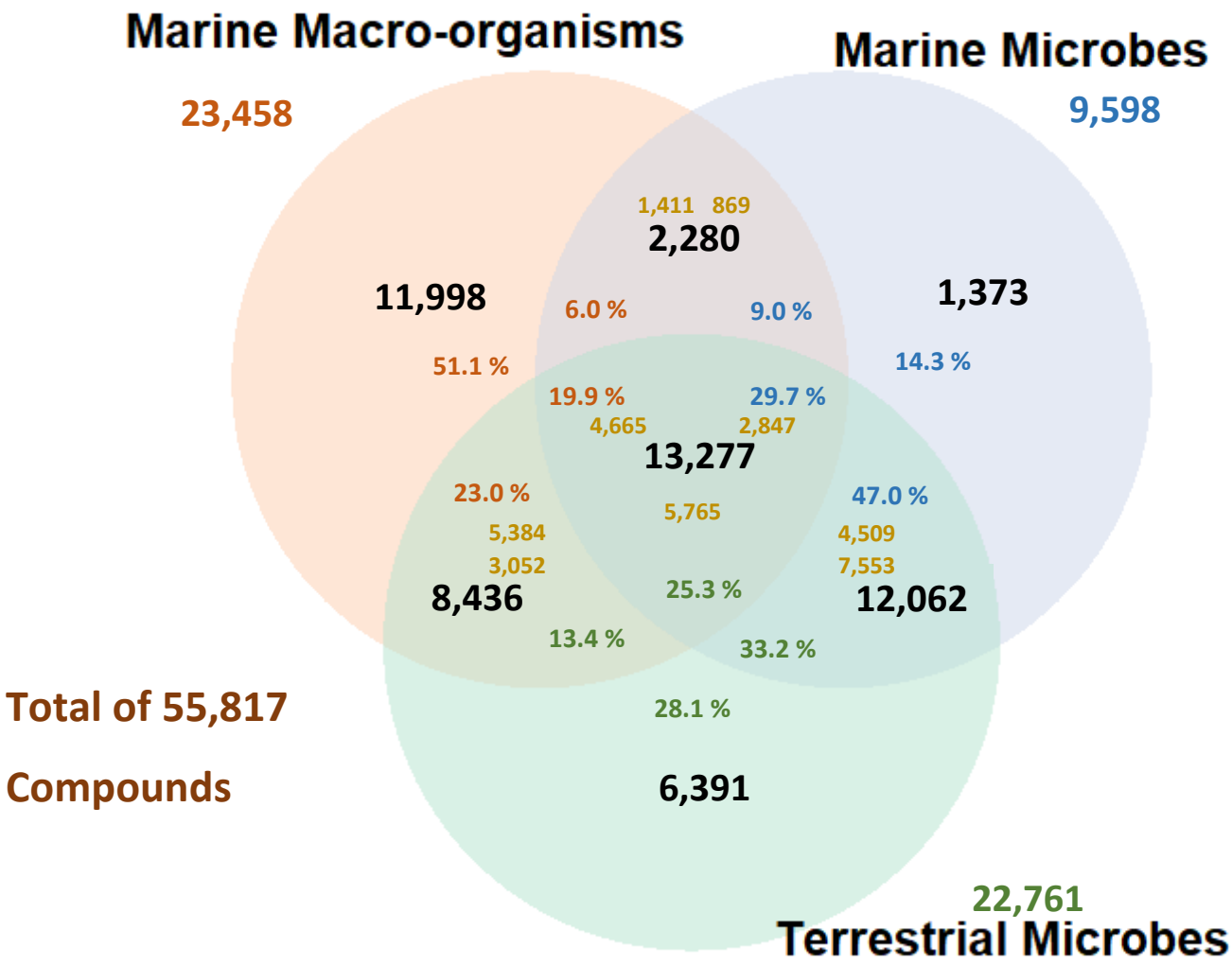


Fig. S7 Nesting of NPs from marine macro-organisms (orange), marine microorganisms (blue) and terrestrial microorganisms (green) with halogenated NPs are removed from common clusters. Total number of NPs in each section is given in black. Smaller numbers represent percentages of total NPs for each biota group. Blue numbers for example relate to the percentage of marine microbial NPs in each section as a proportion of the total 9598 marine microbial NPs. The small orange numbers represent total NPs from each biota group in each intersection. For example, 3052 in the terrestrial microorganism/marine macro-organism intersection represents terrestrial microbial NPs.

Overlap of NPs from Marine Macro-organisms

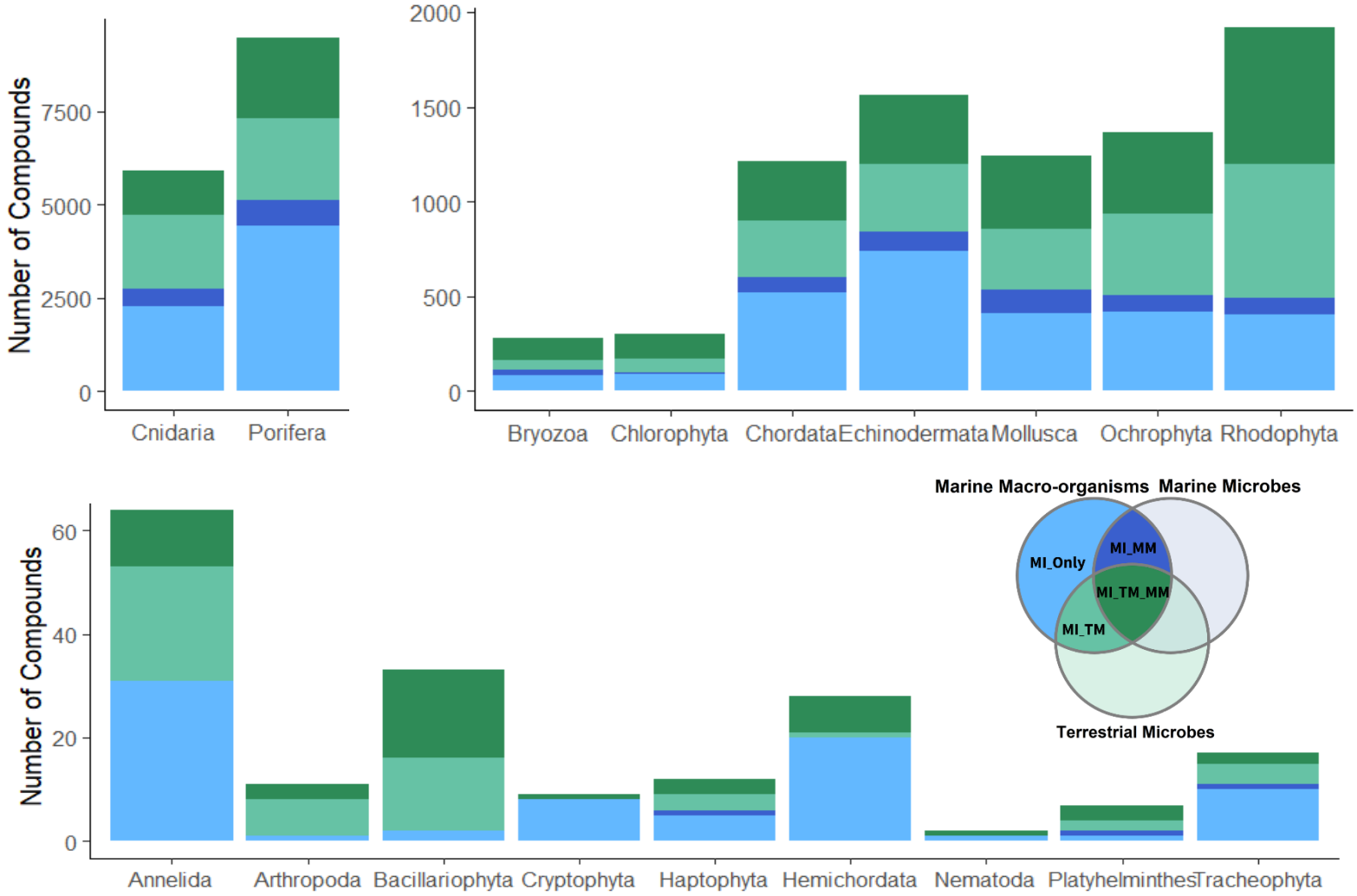


Fig. S8 Number of NPs isolated from marine macro-organisms, divided into their respective phyla. Colours indicate clusterings with NPs from other groups. NPs that cluster only amongst other marine macro-organsims NPs are represented in light blue, NPs that are nested amongst terrestrial microbial NPs are light green, NPs that are nested among marine microbial NPs are dark blue, and NPs that are nested amongst marine and terrestrial microbial NPs are shown in dark green.

Overlap of NPs from Marine Macro-organisms (Accounting for Halogenation)

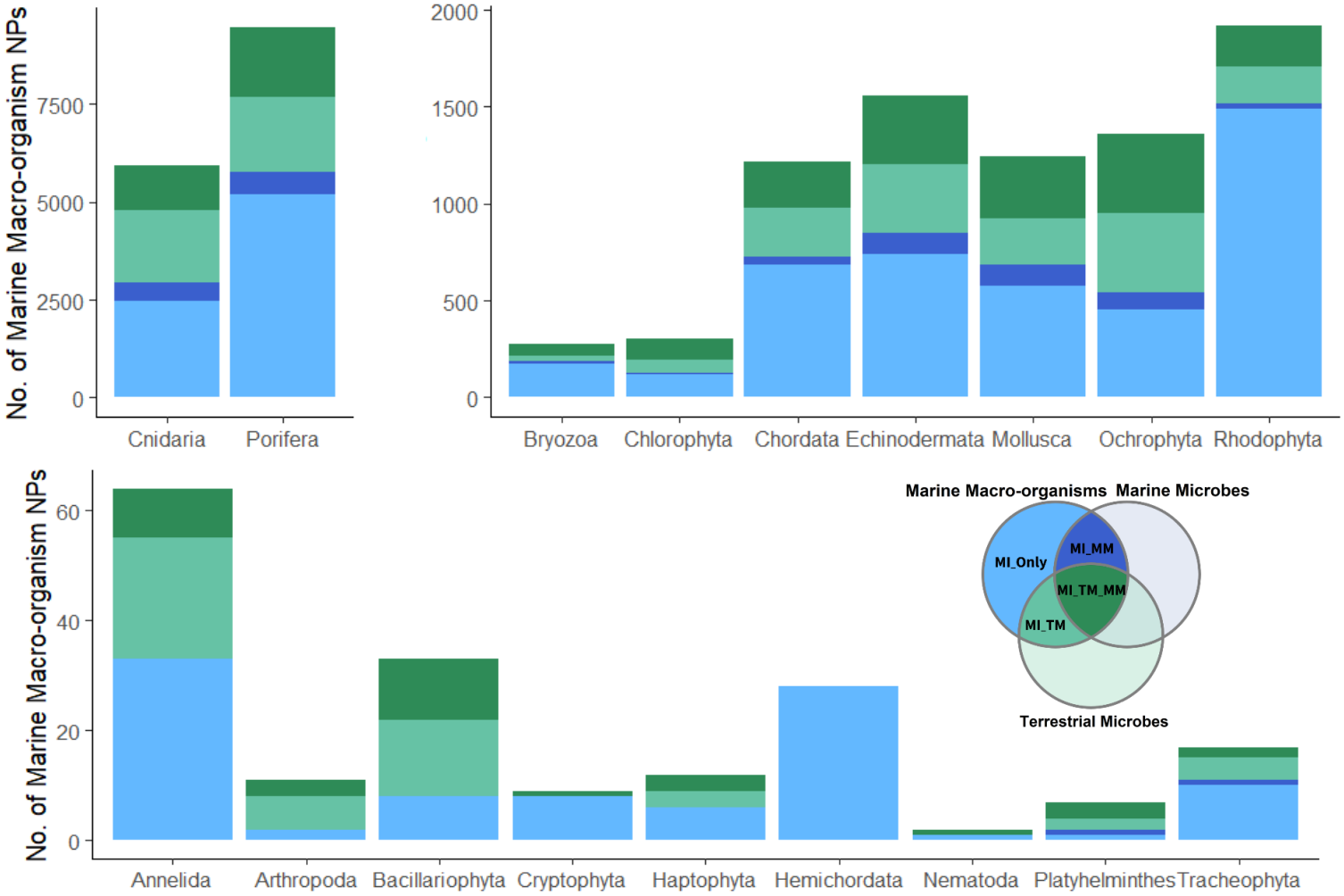


Fig. S9 Number of NPs isolated from marine macro-organisms, divided into their respective phyla with halogenated NPs removed from common clusters. Different colours represent clusterings with NPs from other groups. NPs that cluster only amongst other marine macro-organsims NPs are represented in light blue, NPs that are nested amongst terrestrial microbial NPs are light green, NPs that are nested among marine microbial NPs are dark blue, and NPs that are nested amongst marine and terrestrial microbial NPs are shown in dark green.

Overlap of the Marine Microbial NPs

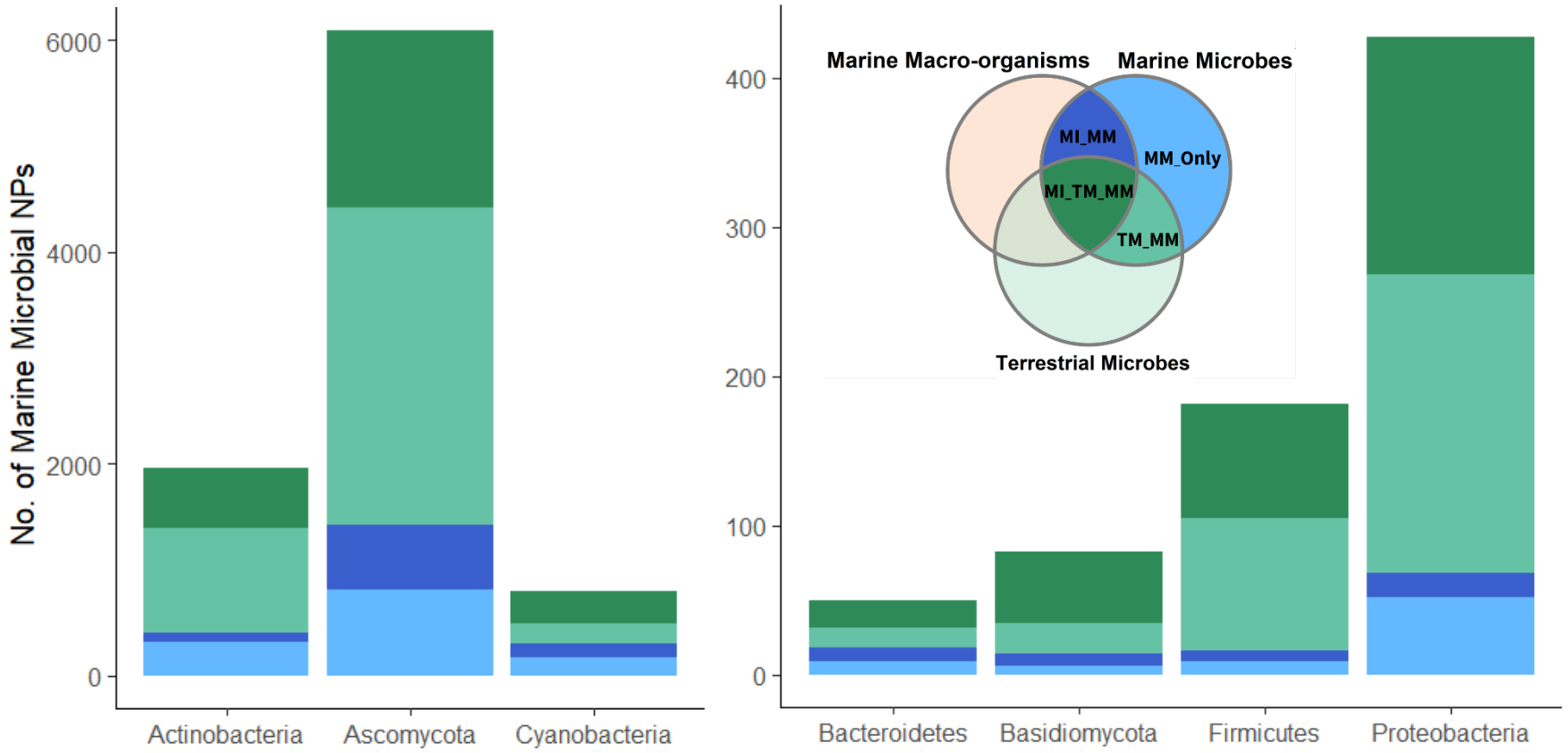


Fig. S10 Number of NPs isolated from marine microorganisms, divided into their respective phyla. Different colours represent clusterings with NPs from other groups. NPs that cluster only amongst other marine microbial NPs are represented in light blue, NPs that are nested amongst terrestrial microbial NPs are light green, NPs that are nested among NPs from marine macro-organisms are dark blue, and NPs that are nested amongst terrestrial microbial NPs and NPs from marine macro-organisms are shown in dark green.

Closer Look at the Different Overlaps in each Phyla and Genera

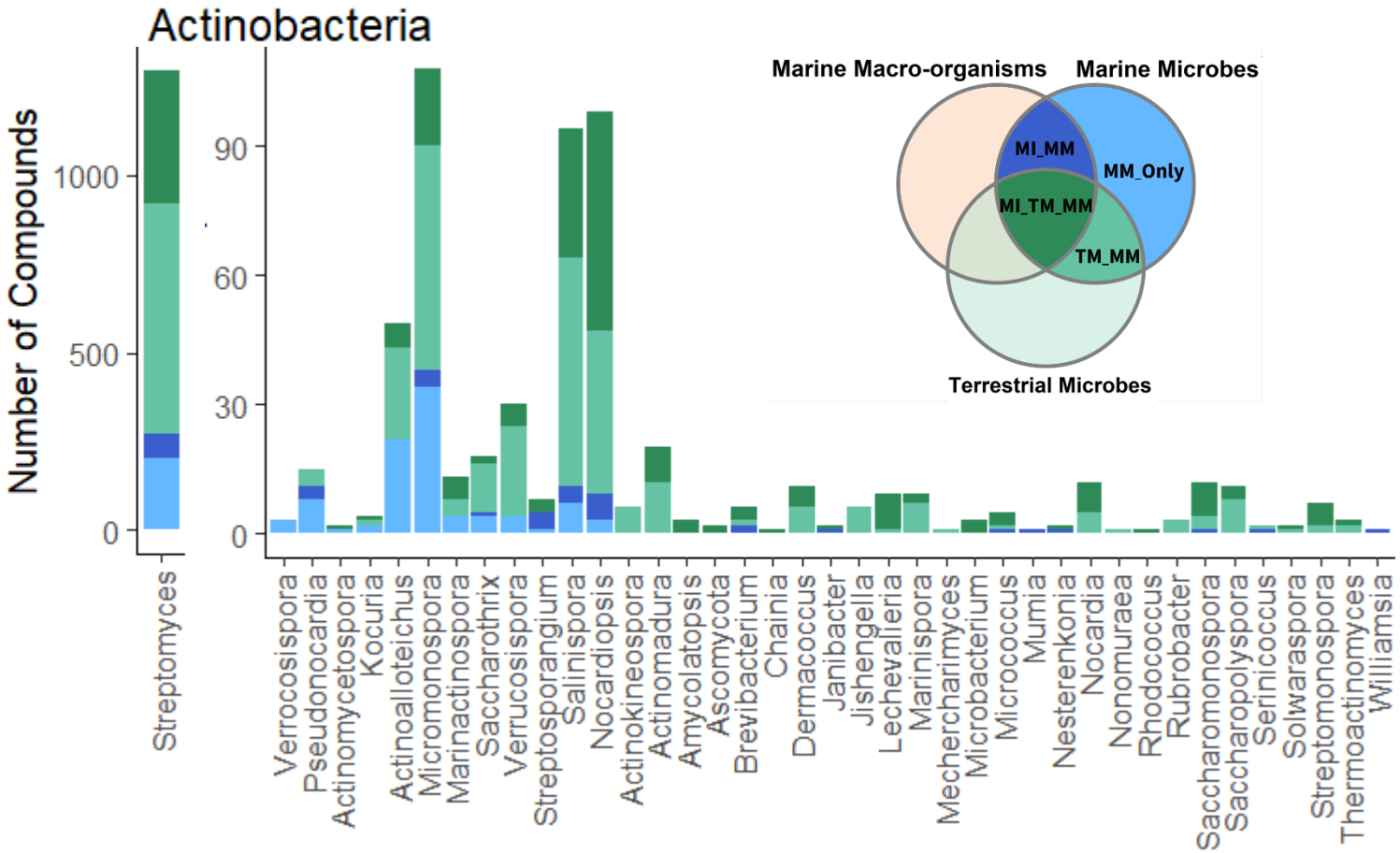


Fig. S11 Number of NPs isolated from **marine Actinobacteria**. Different colours represent clusterings with NPs from other groups. NPs that cluster only amongst other marine microbial NPs are represented in light blue, NPs that are nested amongst terrestrial microbial NPs are light green, NPs that are nested among NPs from marine macro-organisms are dark blue, and NPs that are nested amongst terrestrial microbial NPs and NPs from marine macro-organisms are shown in dark green.

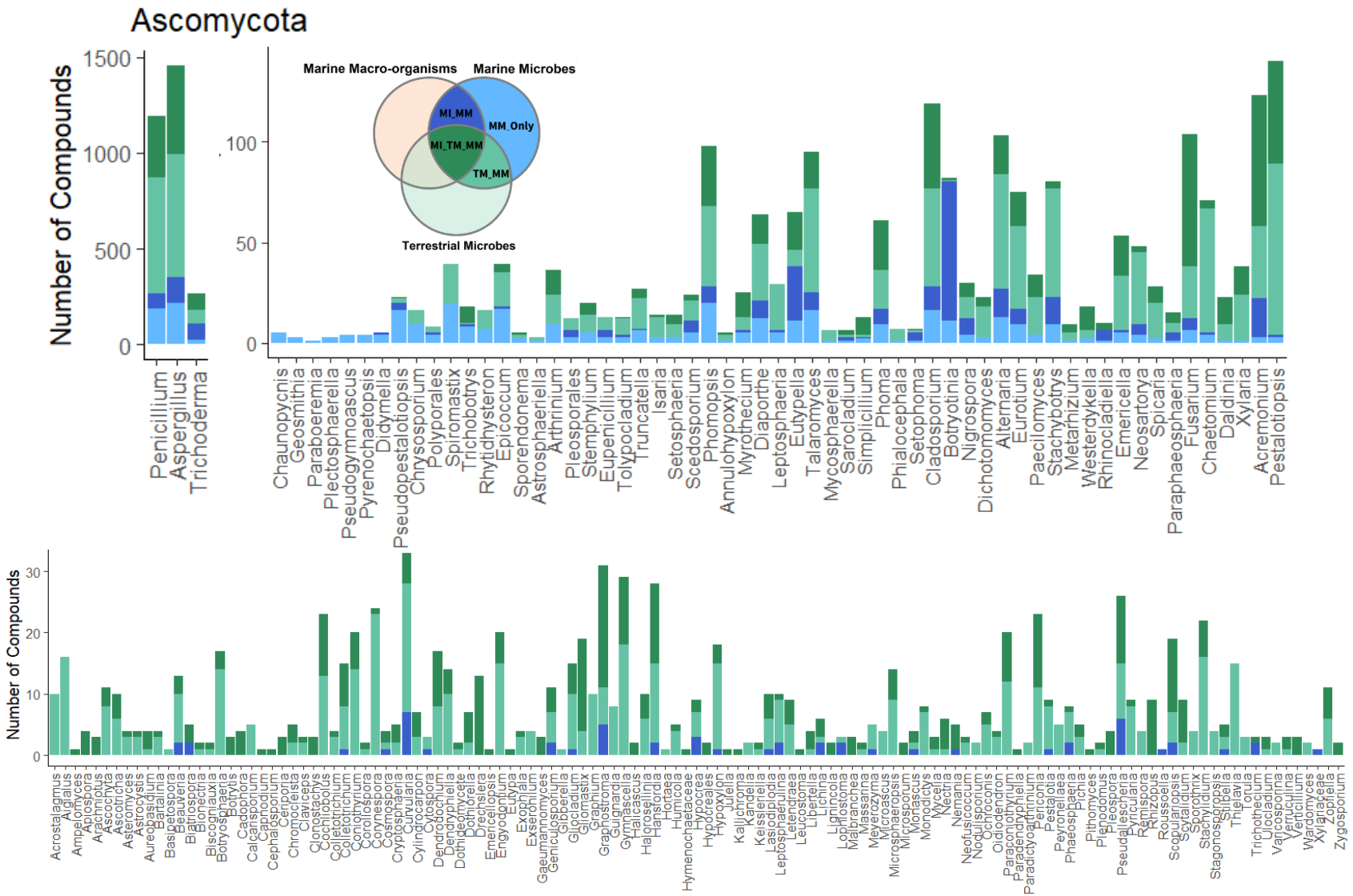


Fig. S12 Number of NPs isolated from **marine Ascomycota**. Different colours represent clusterings with NPs from other groups.. NPs that cluster only amongst other marine microbial NPs are represented in light blue, NPs that are nested amongst terrestrial microbial NPs are light green, NPs that are nested among NPs from marine macro-organisms are dark blue, and NPs that are nested amongst terrestrial microbial NPs and NPs from marine macro-organisms are shown in dark green.

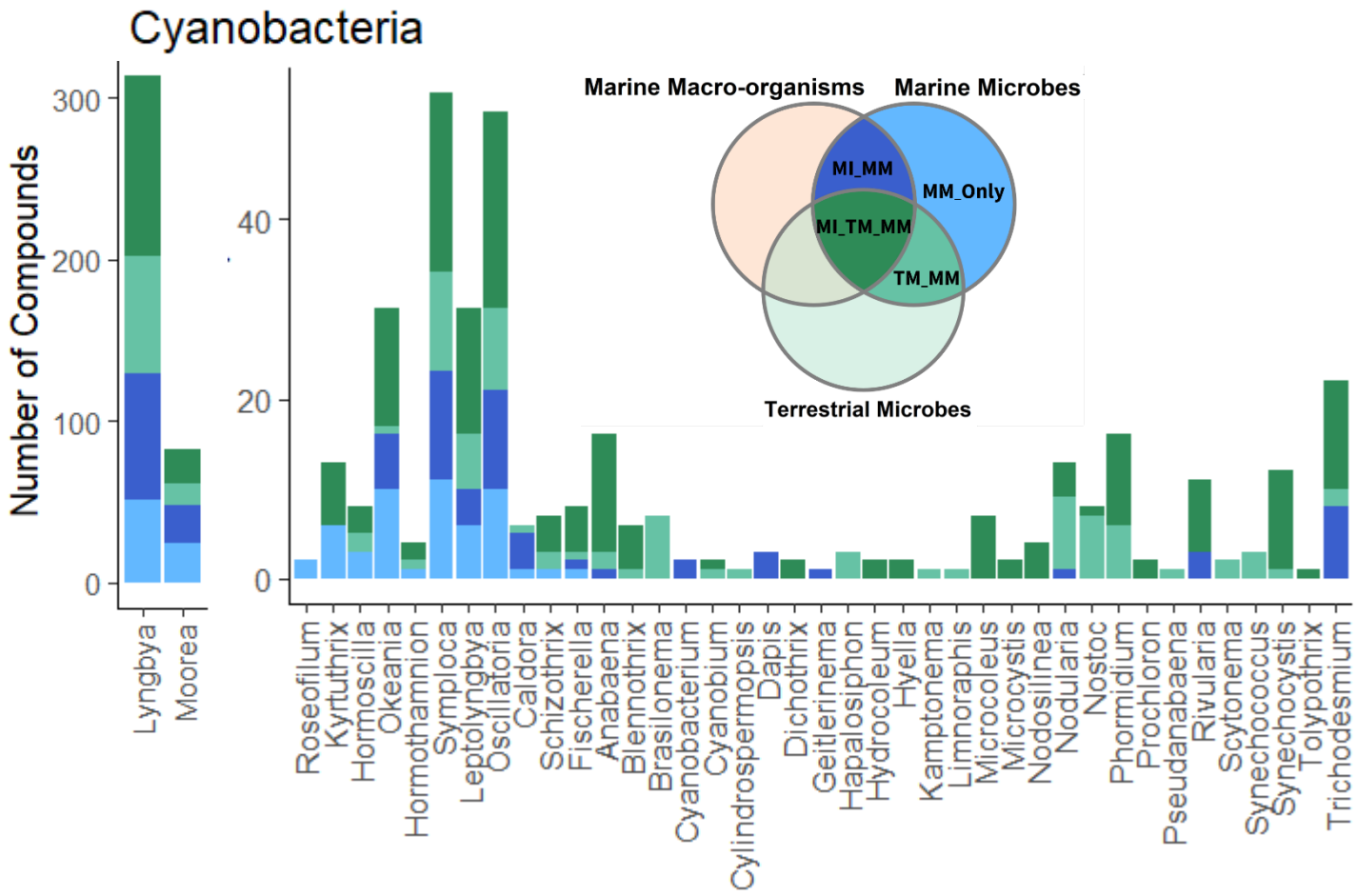


Fig. S13 Number of NPs isolated from **marine Cyanobacteria**. Different colours represent clusterings with NPs from other groups.. NPs that cluster only amongst other marine microbial NPs are represented in light blue, NPs that are nested amongst terrestrial microbial NPs are light green, NPs that are nested among NPs from marine macro-organisms are dark blue, and NPs that are nested amongst terrestrial microbial NPs and NPs from marine macro-organisms are shown in dark green.

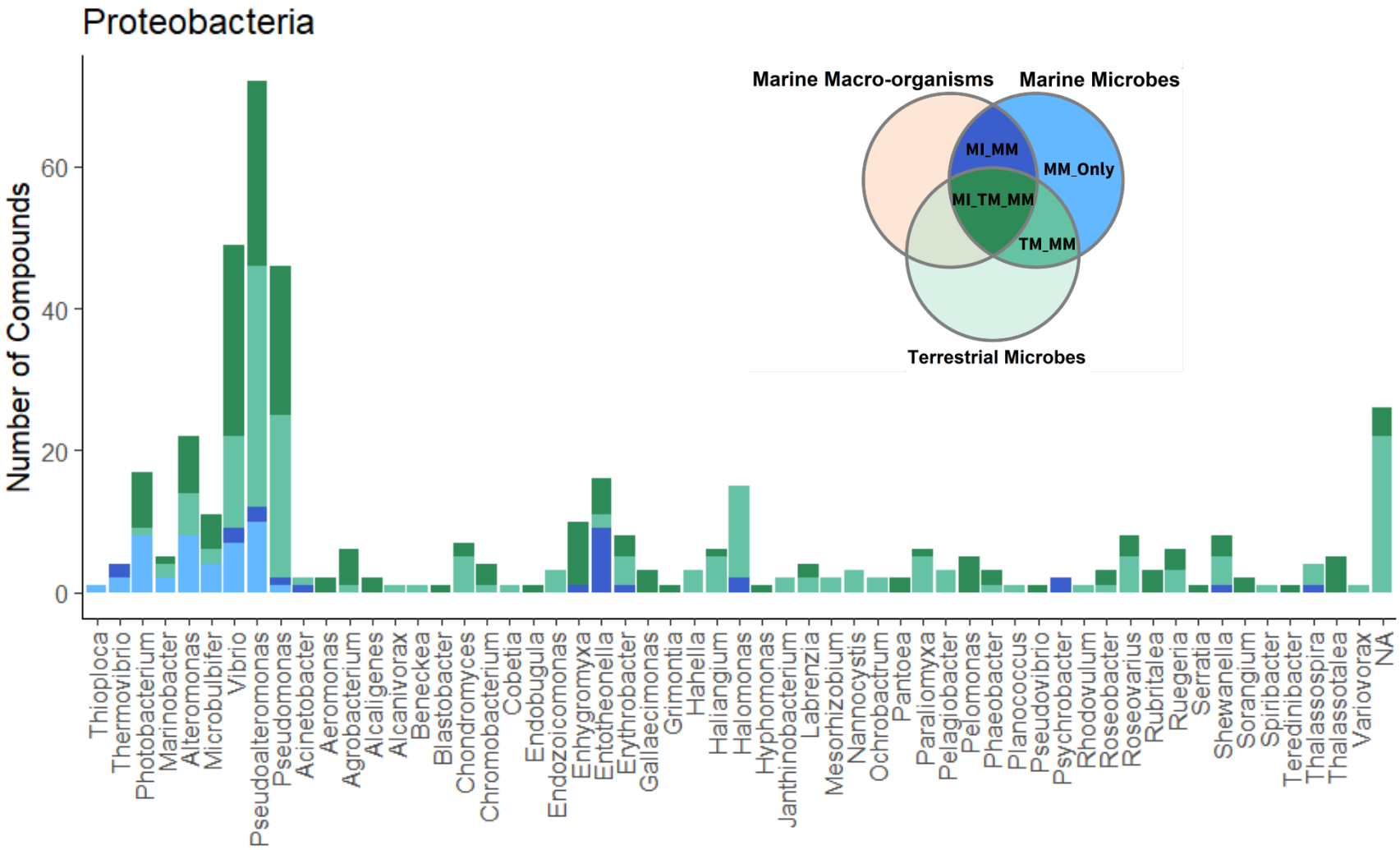


Fig. S14 Number of NPs isolated from **marine Proteobacteria**. Different colours represent clusterings with NPs from other groups.. NPs that cluster only amongst other marine microbial NPs are represented in light blue, NPs that are nested amongst terrestrial microbial NPs are light green, NPs that are nested among NPs from marine macro-organisms are dark blue, and NPs that are nested amongst terrestrial microbial NPs and NPs from marine macro-organisms are shown in dark green.

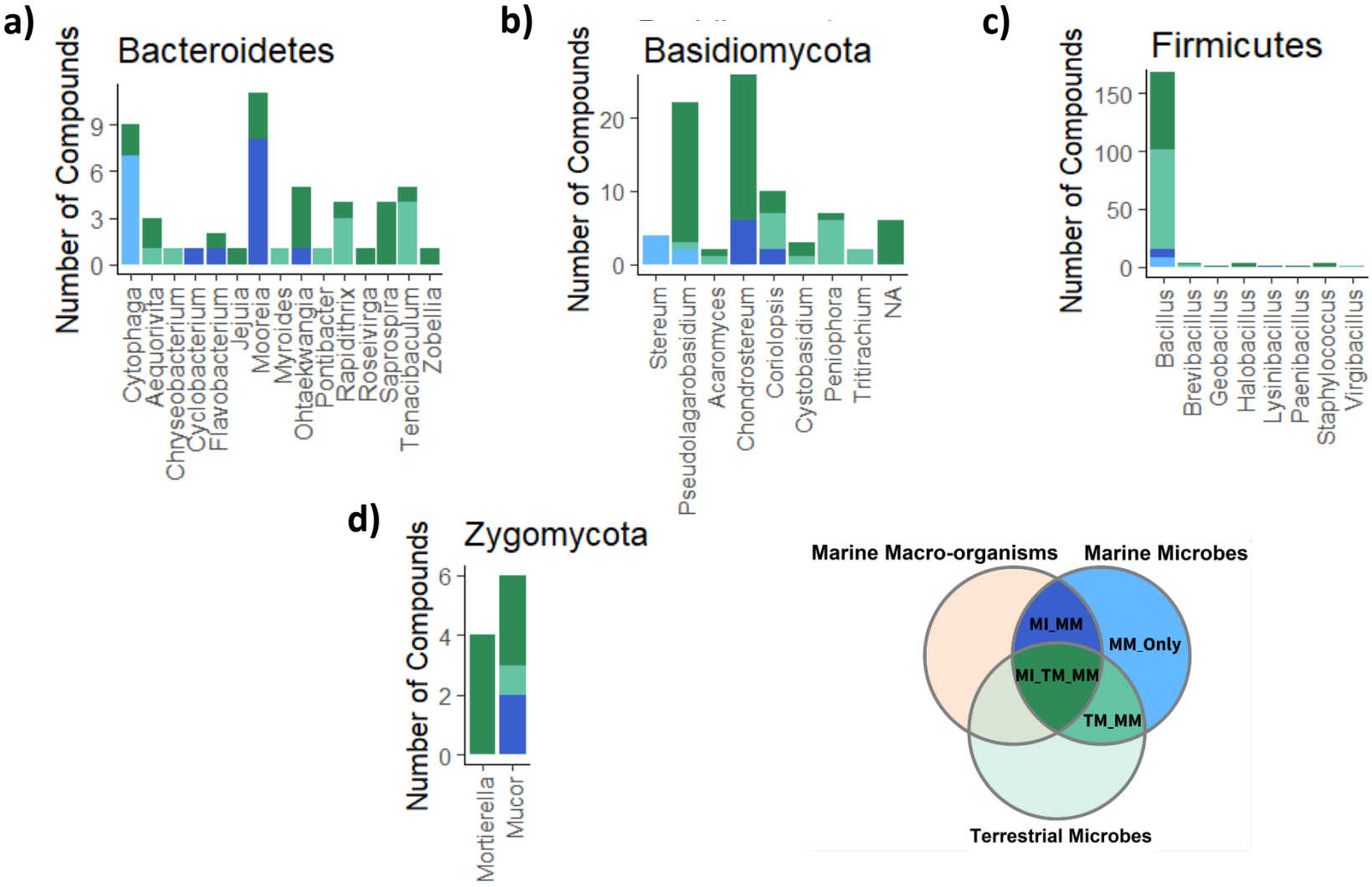


Fig. S15 Number of NPs isolated from **marine Bacteriodetes a), Basidiomycota b), Frimicutes c) and Zygomycota d)**. Different colours represent clusterings with NPs from other groups. NPs that cluster only amongst other marine microbial NPs are represented in light blue, NPs that are nested amongst terrestrial microbial NPs are light green, NPs that are nested among NPs from marine macro-organisms are dark blue, and NPs that are nested amongst terrestrial microbial NPs and NPs from marine macro-organisms are shown in dark green.

Top Producers of Marine Only Compounds

Table 1 List of the top producers of marine microbial NPs that do not nest with terrestrial NPs, according to the fingerprint cluster analysis. On the left are the top 20 genera according to the absolute number of marine only NPs and on the right are the top 20 genera ranked by proportion of NPs that cluster only with marine NPs versus total compounds produced in the marine environment by this genus. For proportion calculations, genera that produced less than 5 compounds in total were removed from the list.

Genus	# of Marine Only
Aspergillus	217
Streptomyces	200
Penicillium	188
Moorea	58
Micromonospora	34
Trichoderma	27
Actinoalloteichus	22
Phomopsis	20
Spiromastix	19
Epicoccum	17
Lyngbya	17
Cladosporium	16
Pseudopestalotiopsis	16
Talaromyces	16
Alternaria	13
Diaporthe	12
Botryotinia	11
Eutypella	11
Symploca	11
Okeania	10

Genus	# of Marine Only	Total # of Molecules	Proportion of Marine Only
Cytophaga	7	9	0.8
Pseudopestalotiopsis	16	23	0.7
Chrysosporium	9	16	0.6
Pseudonocardia	8	15	0.5
Polyporales	4	8	0.5
Spiromastix	19	39	0.5
Photobacterium	8	17	0.5
Kyrtuthrix	6	13	0.5
Actinoalloteichus	22	49	0.4
Trichobotrys	8	18	0.4
Rhytidhysteron	7	16	0.4
Epicoccum	17	39	0.4
Hormoscilla	3	8	0.4
Alteromonas	8	22	0.4
Microbulbifer	4	11	0.4
Okeania	10	30	0.3
Micromonospora	34	108	0.3
Marinactinospora	4	13	0.3
Arthrinium	9	36	0.3
Pleosporales	3	12	0.3

World Maps

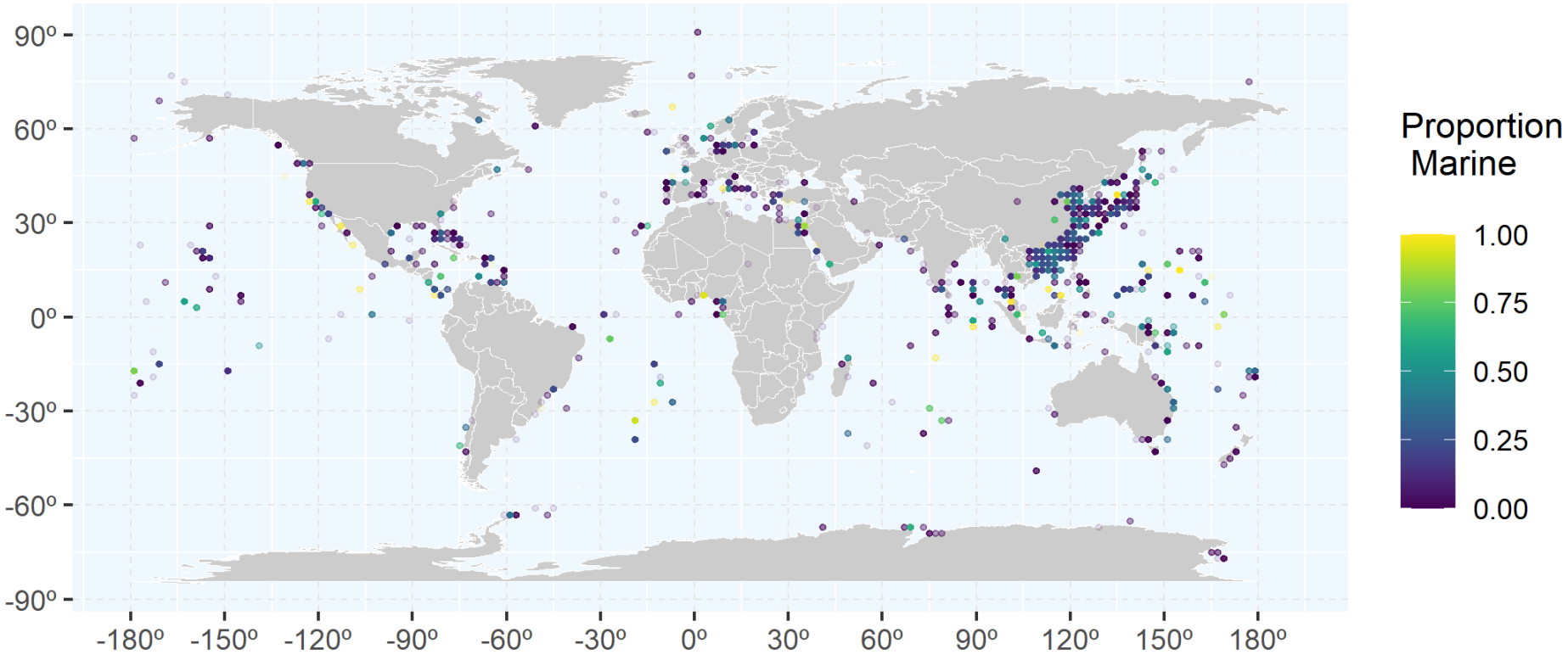


Fig. S16 World map, showing the distribution of the 8,577 marine microbial NPs that have reported coordinates for their microorganism collection. The colours show the proportion of marine microbial NPs found at a specific location (2° x 2° grid cell) that are unique to the marine environment and don't cluster with terrestrial NPs, according to the fingerprint cluster analysis. The number of NPs found at each location varied substantially, and to represent low numbers of NPs at a location, more transparent dots (log10 scale) have been used.

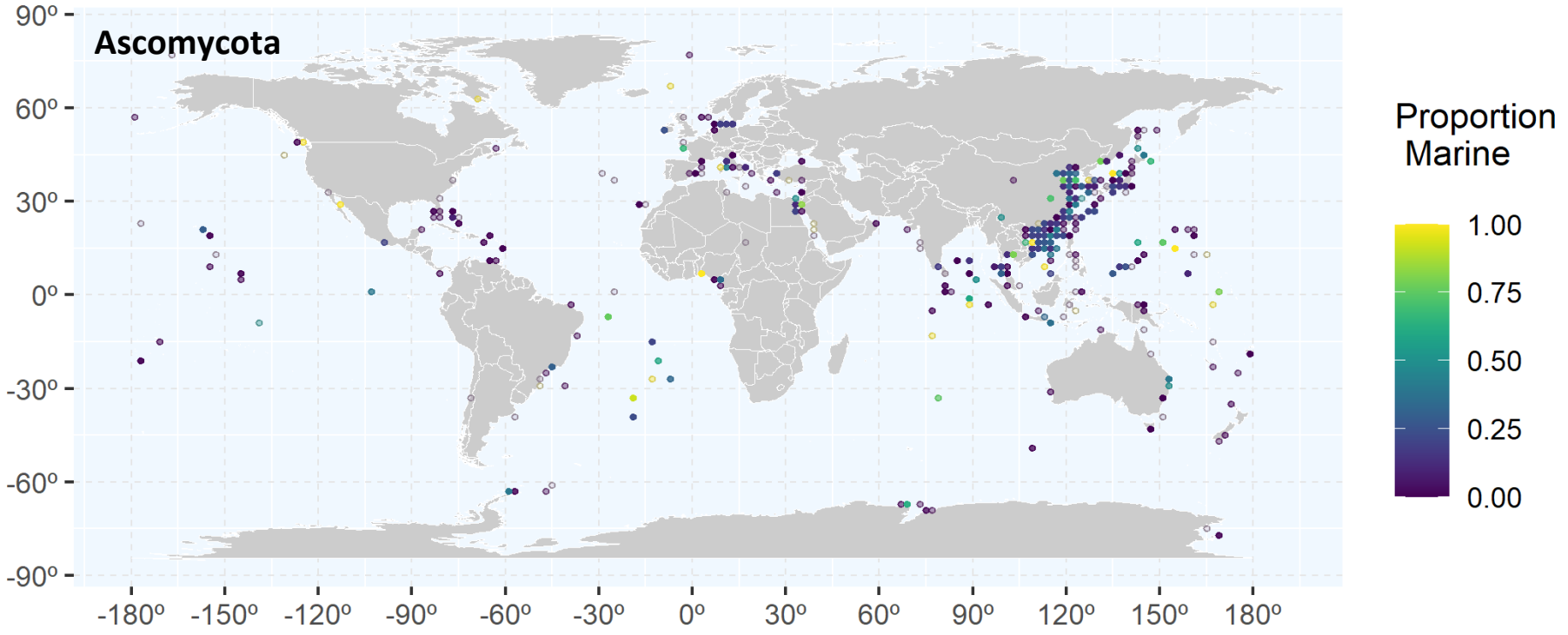


Fig. S17 World map, showing the distribution of the 5,626 marine microbial NPs isolated from Ascomycota that have collection coordinates reported for the microorganism. The colours show the proportion of marine microbial NPs found at a specific location (2° x 2° grid cell) that are unique to the marine environment and don't cluster with terrestrial NPs, according to the fingerprint cluster analysis. The number of NPs found at each location varied substantially, and to represent low numbers of NPs at a location, more transparent dots (log10 scale) have been used.

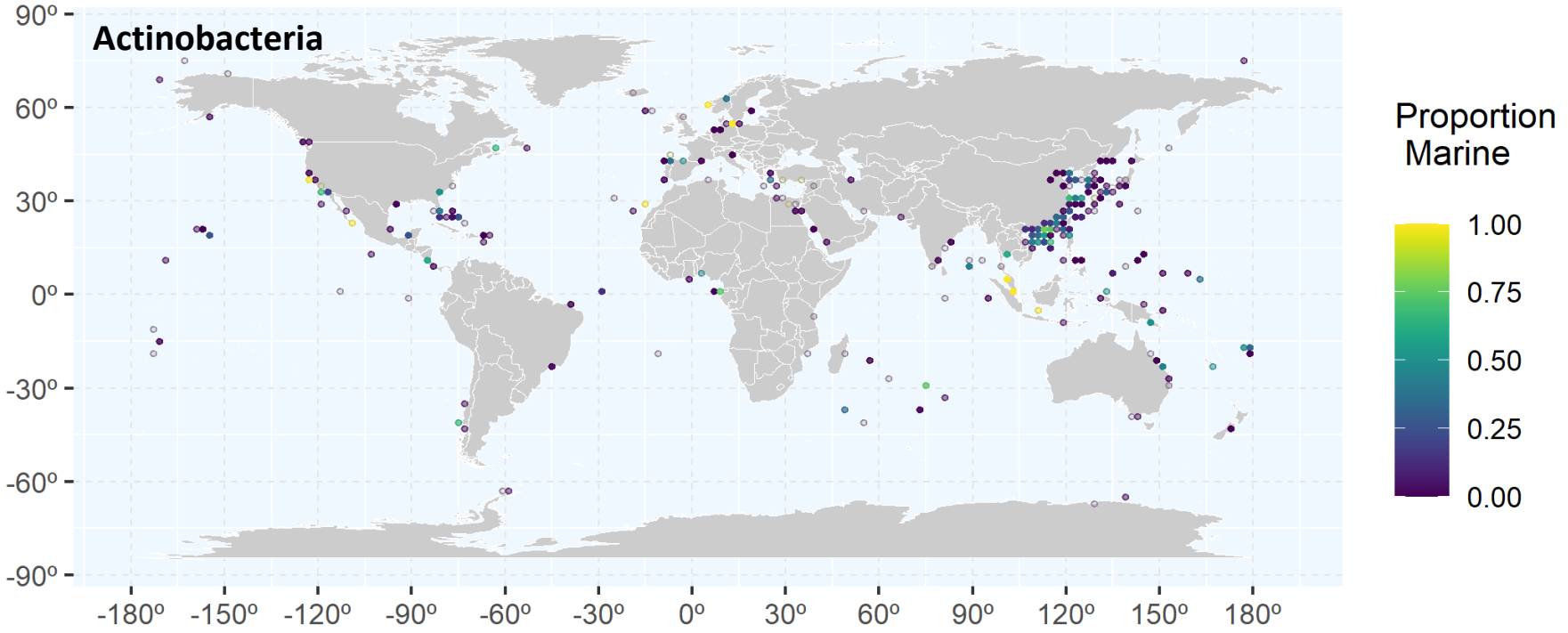


Fig. S18 World map, showing the distribution of the 1,659 marine microbial NPs isolated from Actinobacteria that have coordinates reported for the microorganism collection. The colours show the proportion of marine microbial NPs found at a specific location (2° x 2° grid cell) that are unique to the marine environment and don't cluster with terrestrial NPs, according to the fingerprint cluster analysis. The number of NPs found at each location varied substantially, and to represent low numbers of NPs at a location, more transparent dots (log10 scale) have been used.

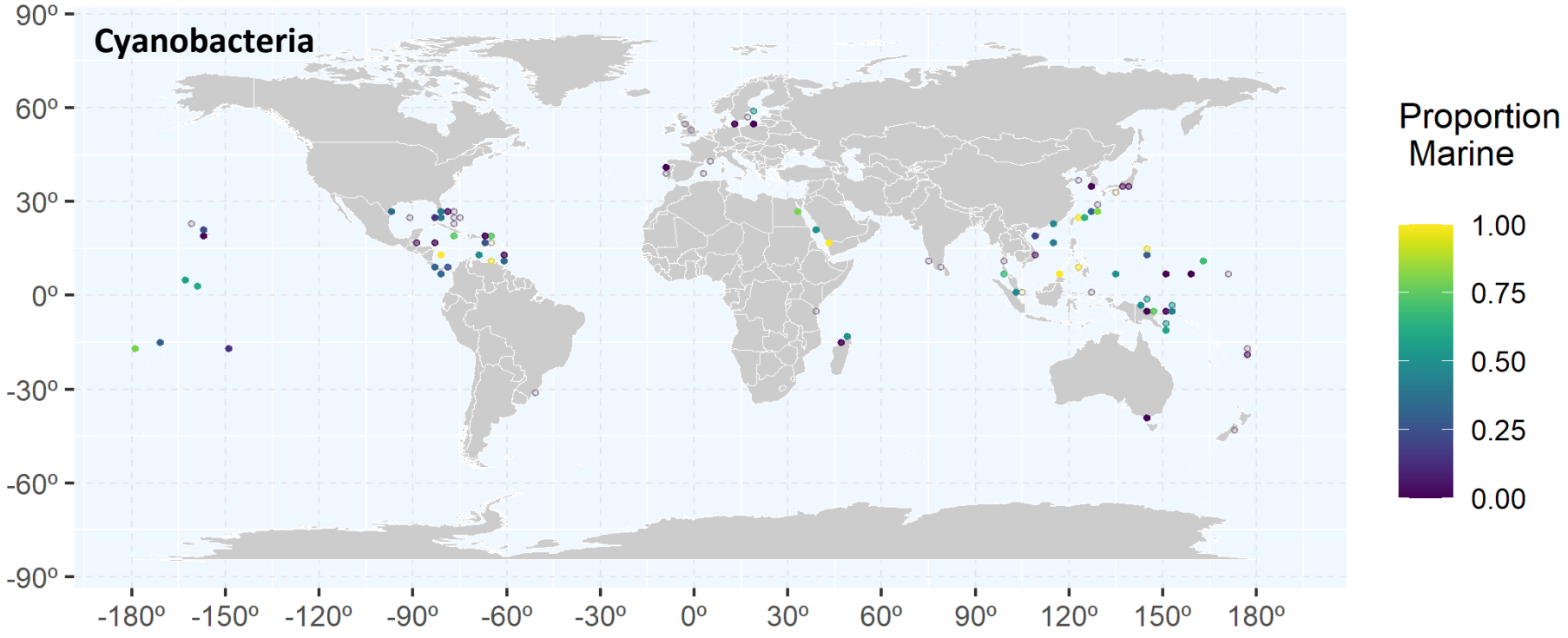


Fig. S19 World map, showing the distribution of the 734 marine microbial NPs isolated from Cyanobacteria that have reported coordinates for microorganism collection. The colours show the proportion of marine microbial NPs found at a specific location (2° x 2° grid cell) that are unique to the marine environment and don't cluster with terrestrial NPs, according to the fingerprint cluster analysis. The number of NPs found at each location varied substantially, and to represent low numbers of NPs at a location, more transparent dots (log10 scale) have been used.

Cluster Analysis on PubChem Fingerprint (5000 Clusters)

Overlap Between Marine and Terrestrial Microbial NPs

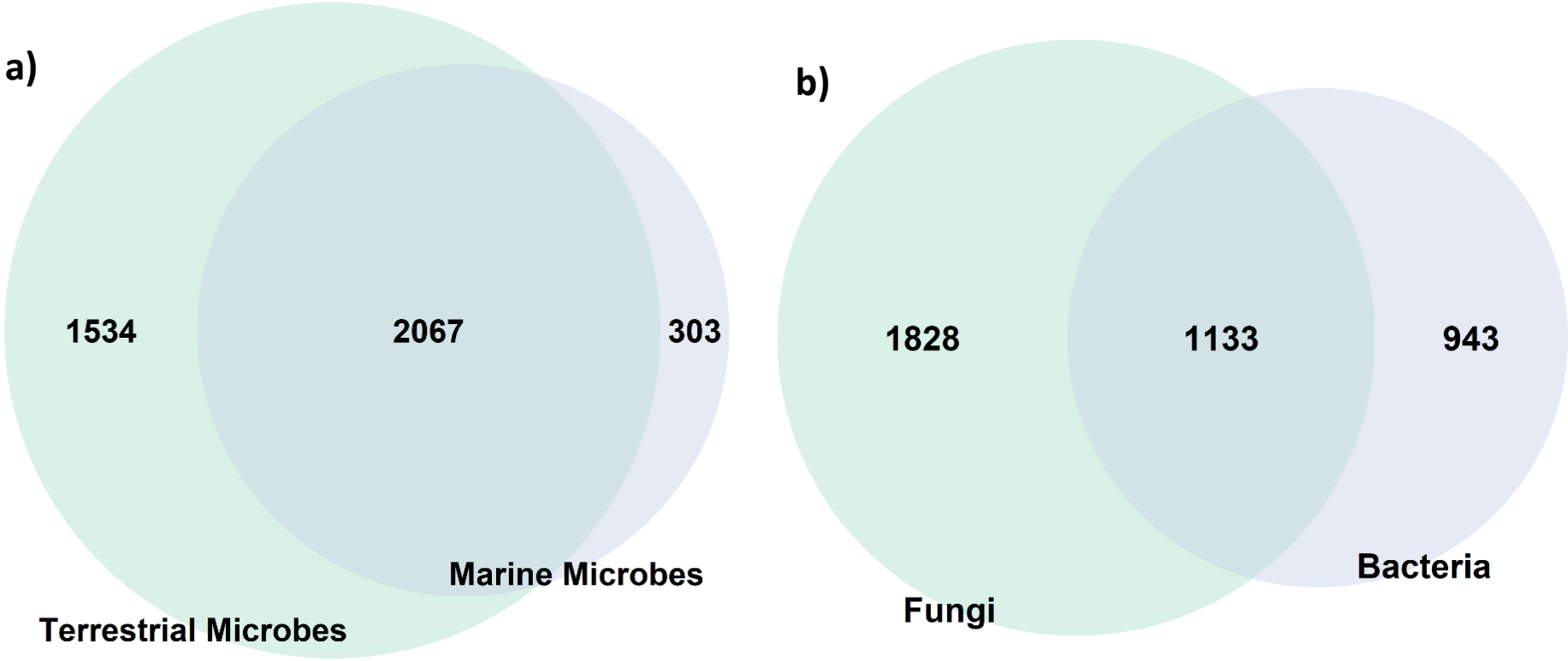


Fig. S20 a) Overlap between marine (blue) and terrestrial (green) microbial molecular clusters. It shows that 78.9% of the marine natural products clusters are nested amongst terrestrial clusters. Numbers indicate the total number of clusters in each section. b) Overlap between the marine and terrestrial fungal NPs (green) and bacterial NPs (blue).

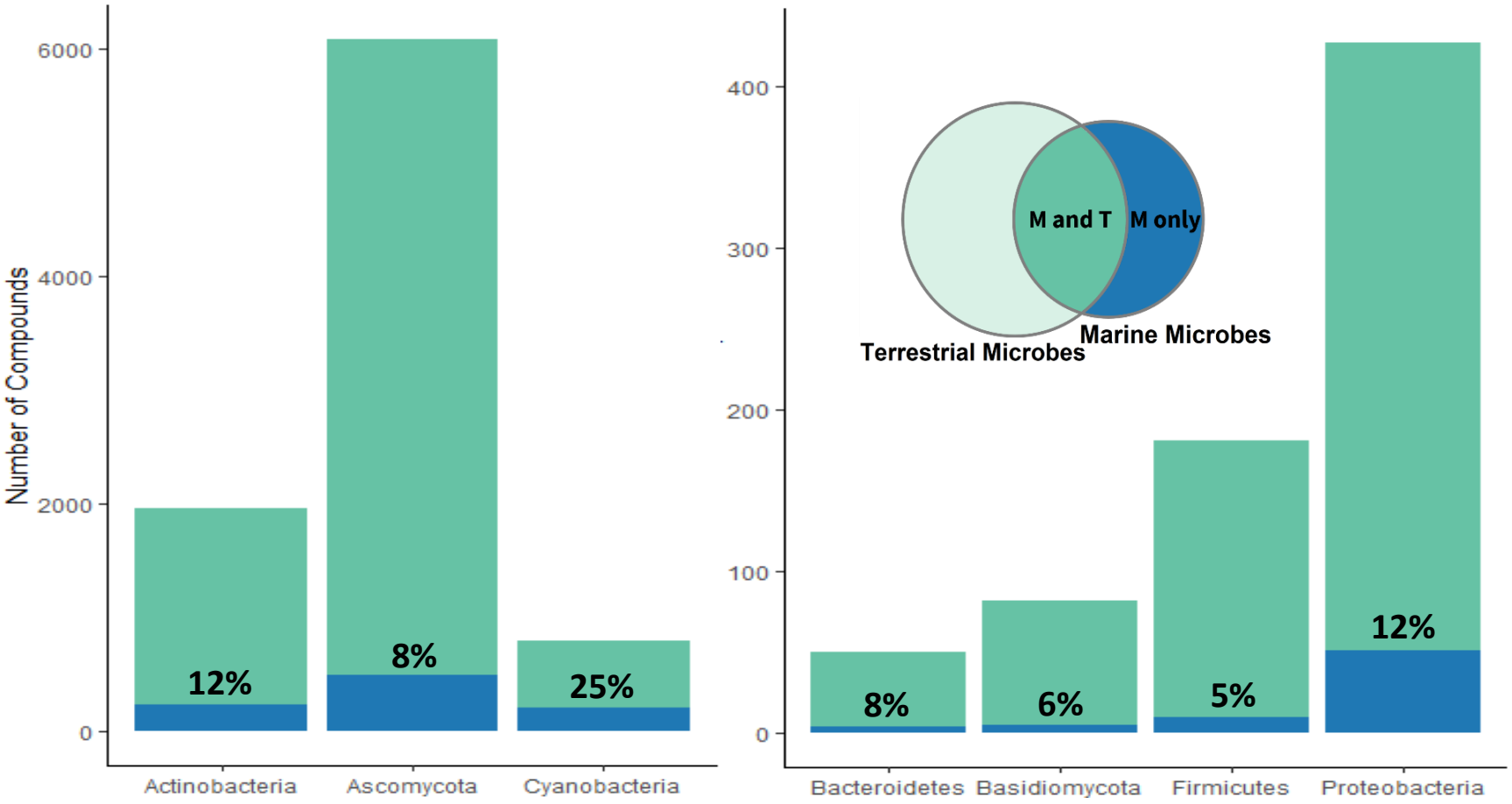


Fig. S21 Total number and percentage of uniquely marine microbial NPs (blue) produced by organisms in different phyla *vs* marine microbial NPs that overlap with terrestrial microbial NPs (green).

Overlap between the Three Different Groups (Clusters)

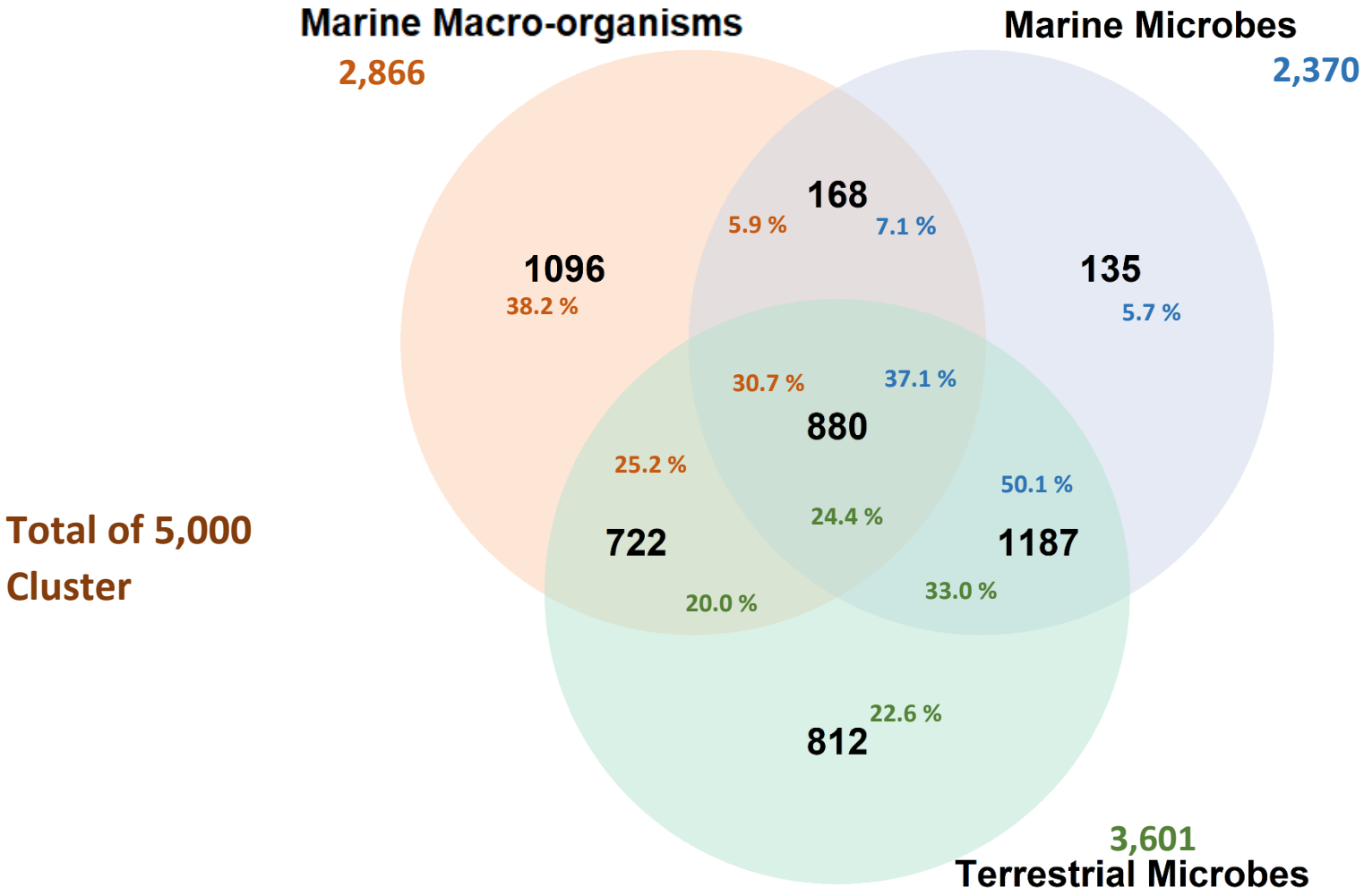


Fig. S22 Nesting of NP clusters from marine macro-organisms (orange), marine microorganisms (blue) and terrestrial microorganisms (green). Total number of clusters in each section is given in black. Smaller numbers represent percentages of total clusters for each biota group. Blue numbers for example relate to the percentage of marine microbial NPs clusters in each section as a proportion of the total 2370 marine microbial NP clusters

Overlap between the Three Different Groups (Compounds)

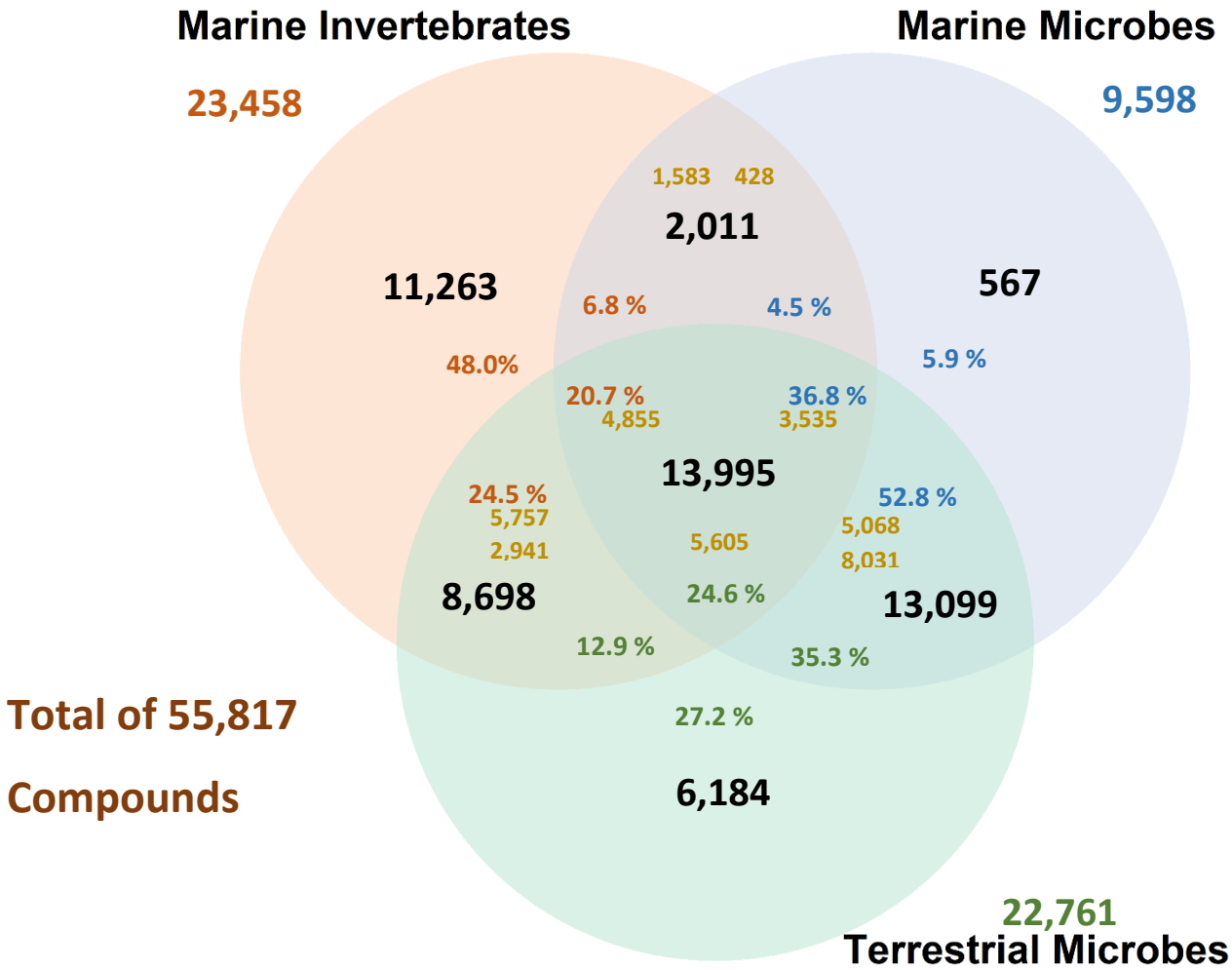


Fig. S23 Nesting of NPs from marine macro-organisms (orange), marine microorganisms (blue) and terrestrial microorganisms (green). Total number of NPs in each section is given in black. Smaller numbers represent percentages of total NPs for each biota group. Blue numbers for example relate to the percentage of marine microbial NPs in each section as a proportion of the total 9598 marine microbial NPs. The small orange numbers represent total NPs from each biota group in each intersection. For example, 2941 in the terrestrial microorganism/marine macro-organism intersection represents terrestrial microbial NPs.

Murcko Scaffold Analysis (9195 Scaffolds)

Overlap Between Marine and Terrestrial Microbial NPs

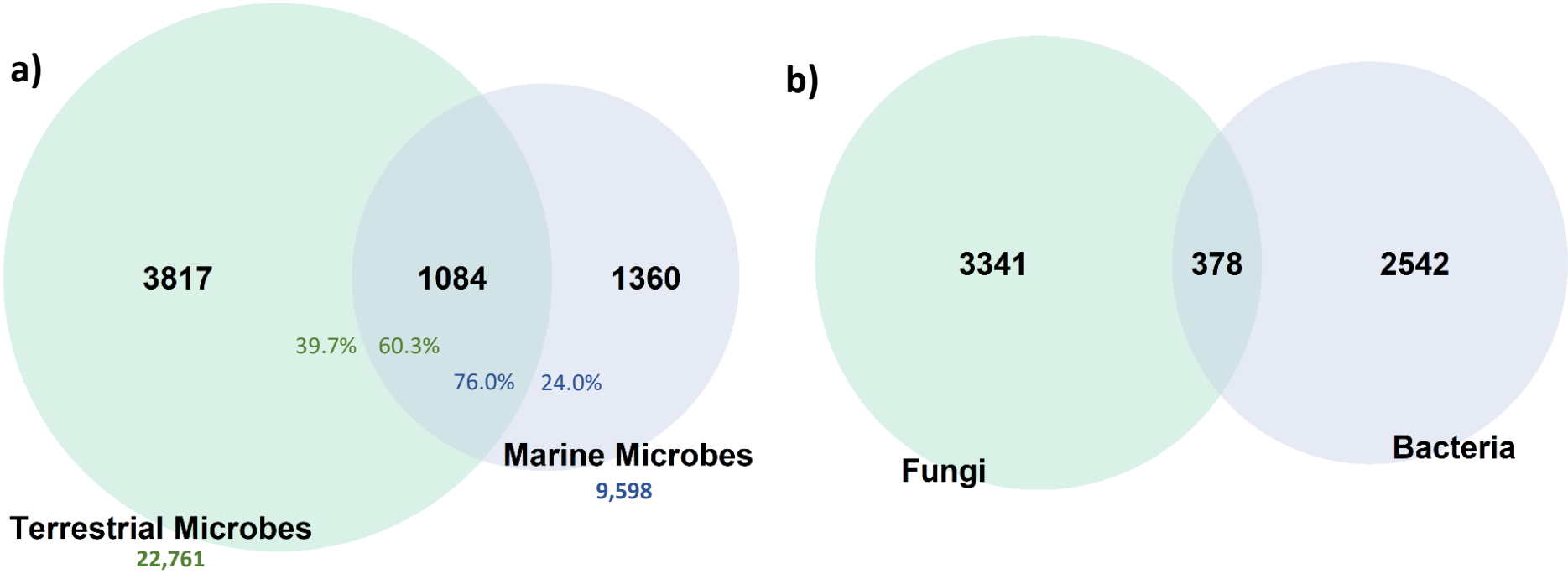


Fig. S24 a) Overlap between marine (blue) and terrestrial (green) microbial molecular scaffolds. It shows that 78.9% of the marine NP scaffolds are nested amongst terrestrial scaffolds. Numbers indicate the total number of scaffolds in each section. b) Overlap between the marine and terrestrial fungal NPs (green) and bacterial NPs (blue). Jaccard index shows that there is only a 6% overlap between bacterial and fungal NPs.

Marine Microbial NPs and their overlap with Terrestrial Microbial NPs

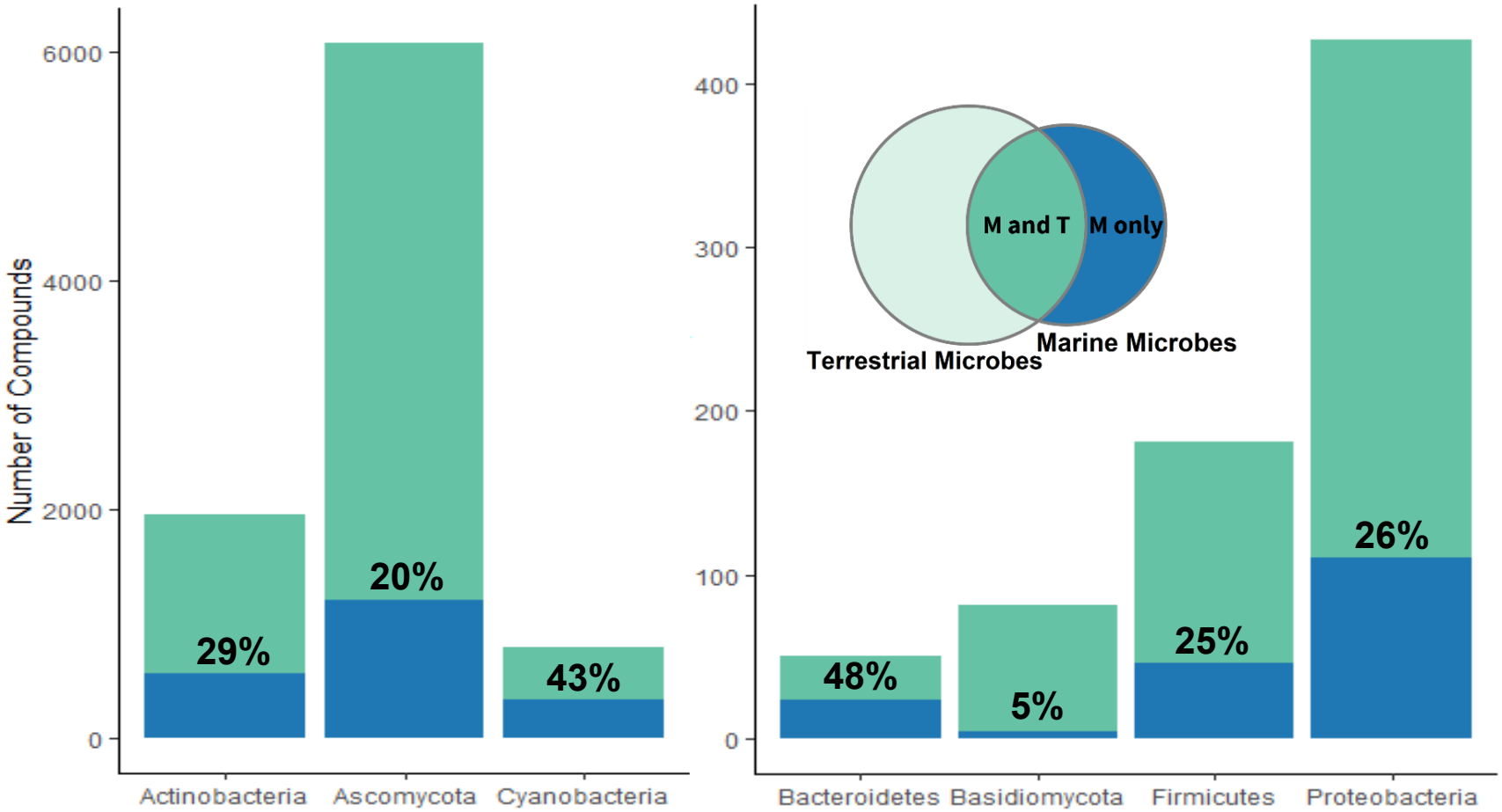


Fig. S25 Total number and percentage of uniquely marine microbial NPs (blue) produced by microbes from different phyla *vs* marine microbial NPs that overlap with terrestrial microbial NPs (green).

Overlap between the Three Different Groups (Scaffolds)

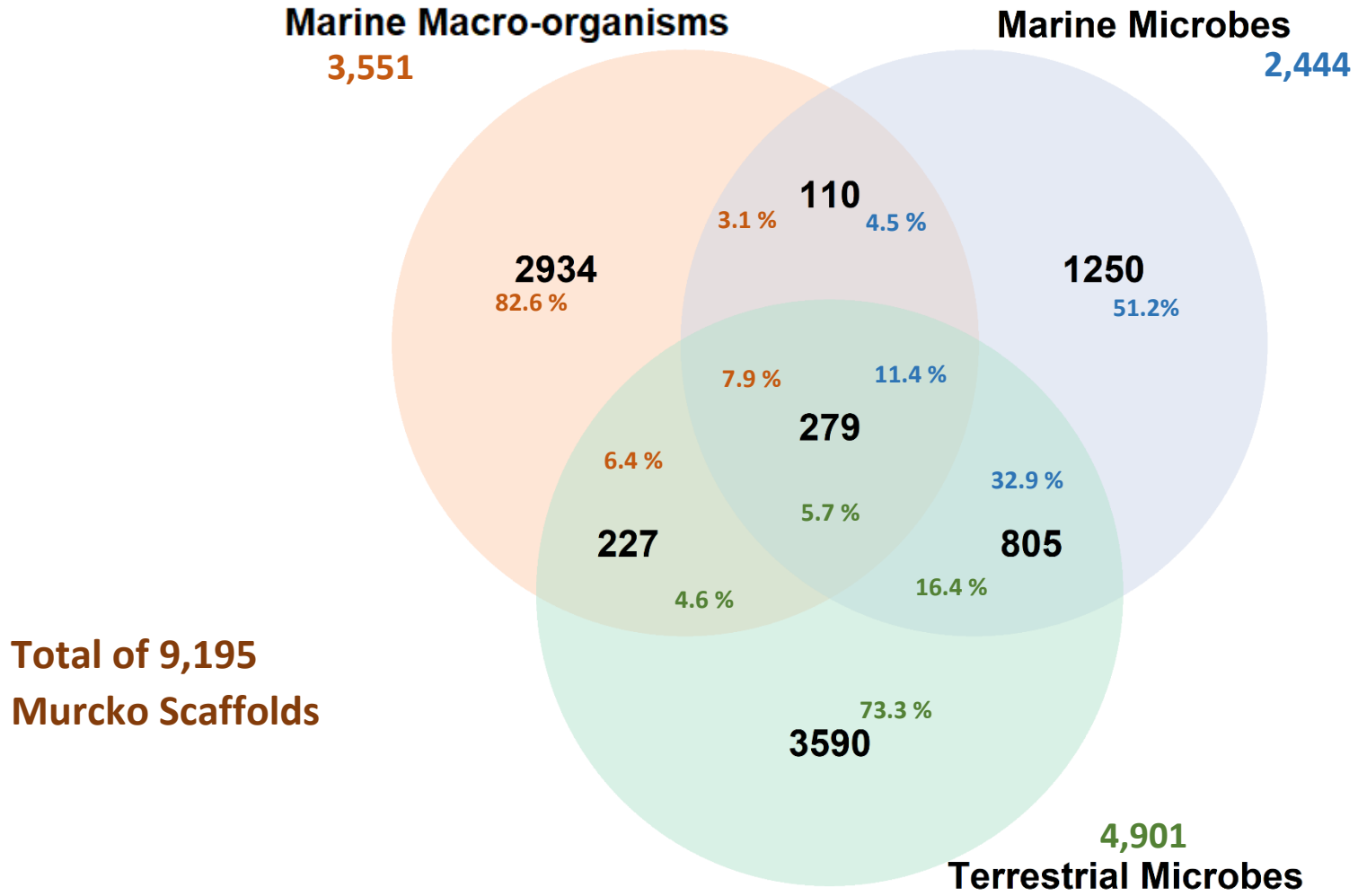


Fig. S26 Nesting of NP scaffolds from marine macro-organisms (orange), marine microorganisms (blue) and terrestrial microorganisms (green). Total number of scaffolds in each section is given in black. Smaller numbers represent percentages of total scaffolds for each biota group. Blue numbers for example relate to the percentage of marine microbial NPs scaffolds in each section as a proportion of the total 2444 marine microbial NP scaffolds

Overlap between the Three Different Groups (Compounds)

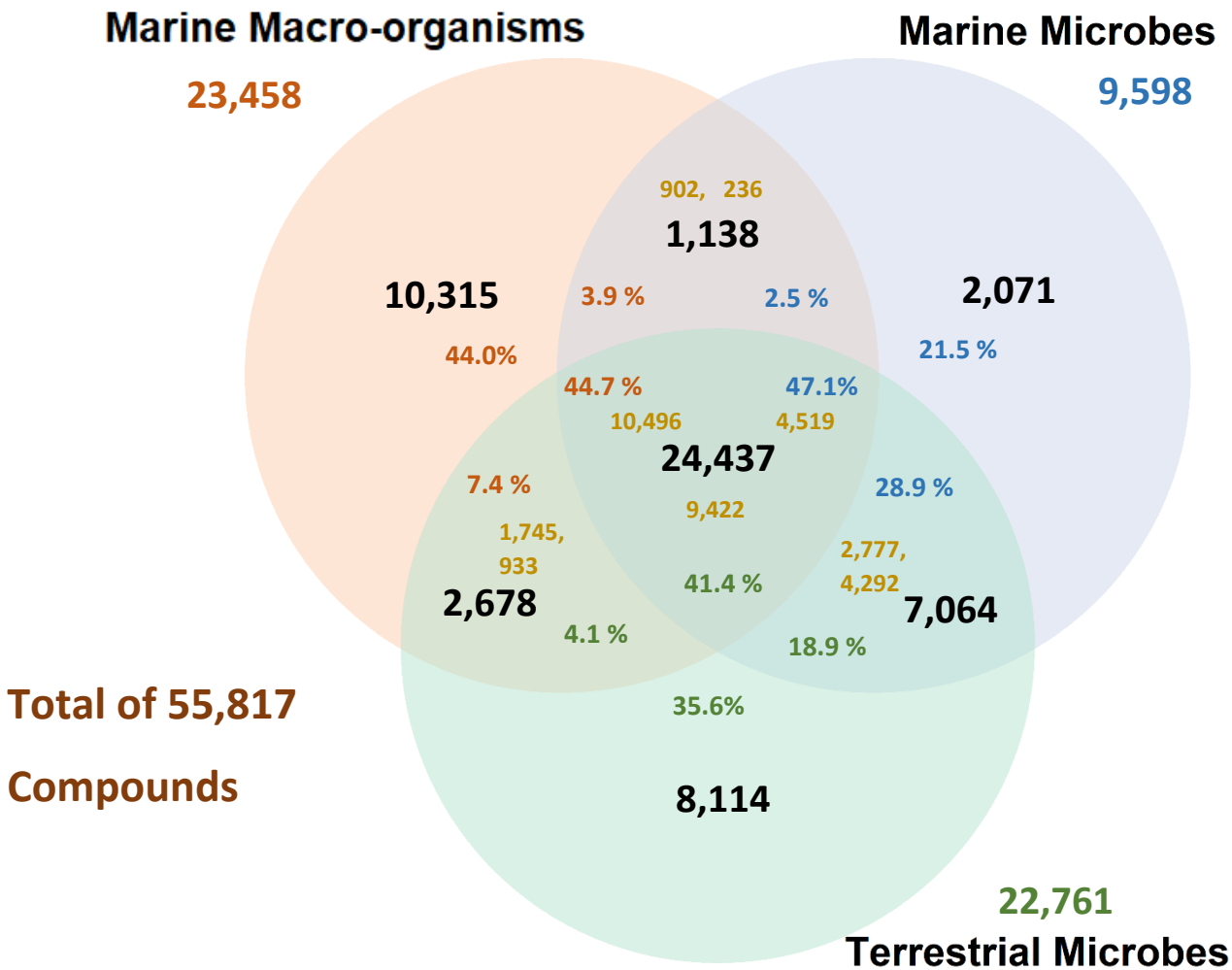


Fig. S27 Nesting of NP scaffolds from marine macro-organisms (orange), marine microorganisms (blue) and terrestrial microorganisms (green). Total number of NPs in each section is given in black. Smaller numbers represent percentages of total NPs for each biota group. Blue numbers for example relate to the percentage of marine microbial NPs in each section as a proportion of the total 9598 marine microbial NPs. The small orange numbers represent total NPs from each biota group in each intersection. For example, 933 in the terrestrial microorganism/marine macro-organism intersection represents terrestrial microbial NPs.

Overlap of the Marine Microbial NPs

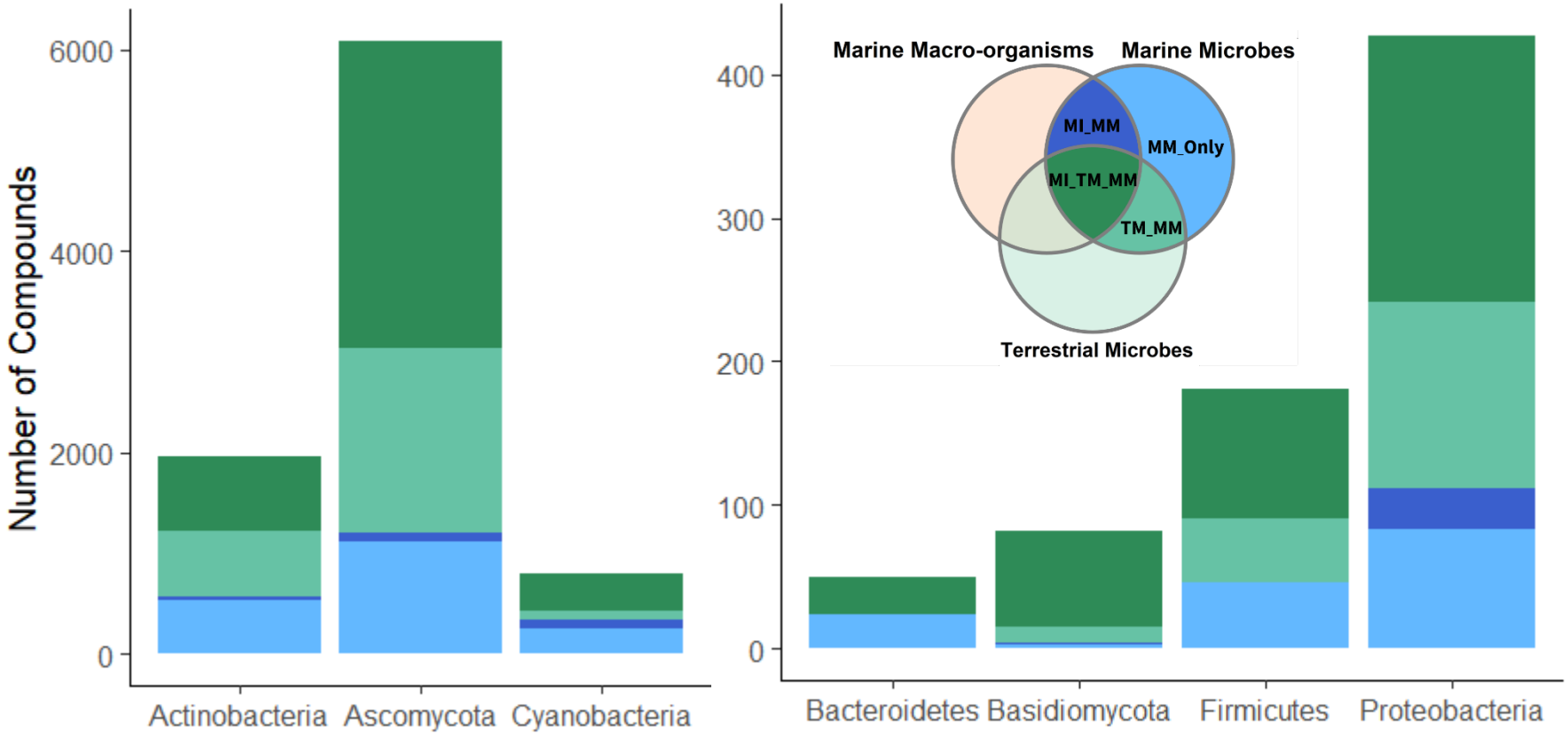


Fig. S28 Numbers of NPs isolated from marine microorganisms, divided into their respective phyla. Colours indicate if these NPs cluster with NPs from other groups. NPs that cluster only amongst other marine microbial NPs are represented in light blue, NPs that are nested amongst terrestrial microbial NPs are light green, NPs that are nested among NPs from marine macro-organisms are dark blue, and NPs that are nested amongst terrestrial microbial NPs and NPs from marine macro-organisms are shown in dark green.

Closer Look at the Different Overlaps in each Phyla and Genera

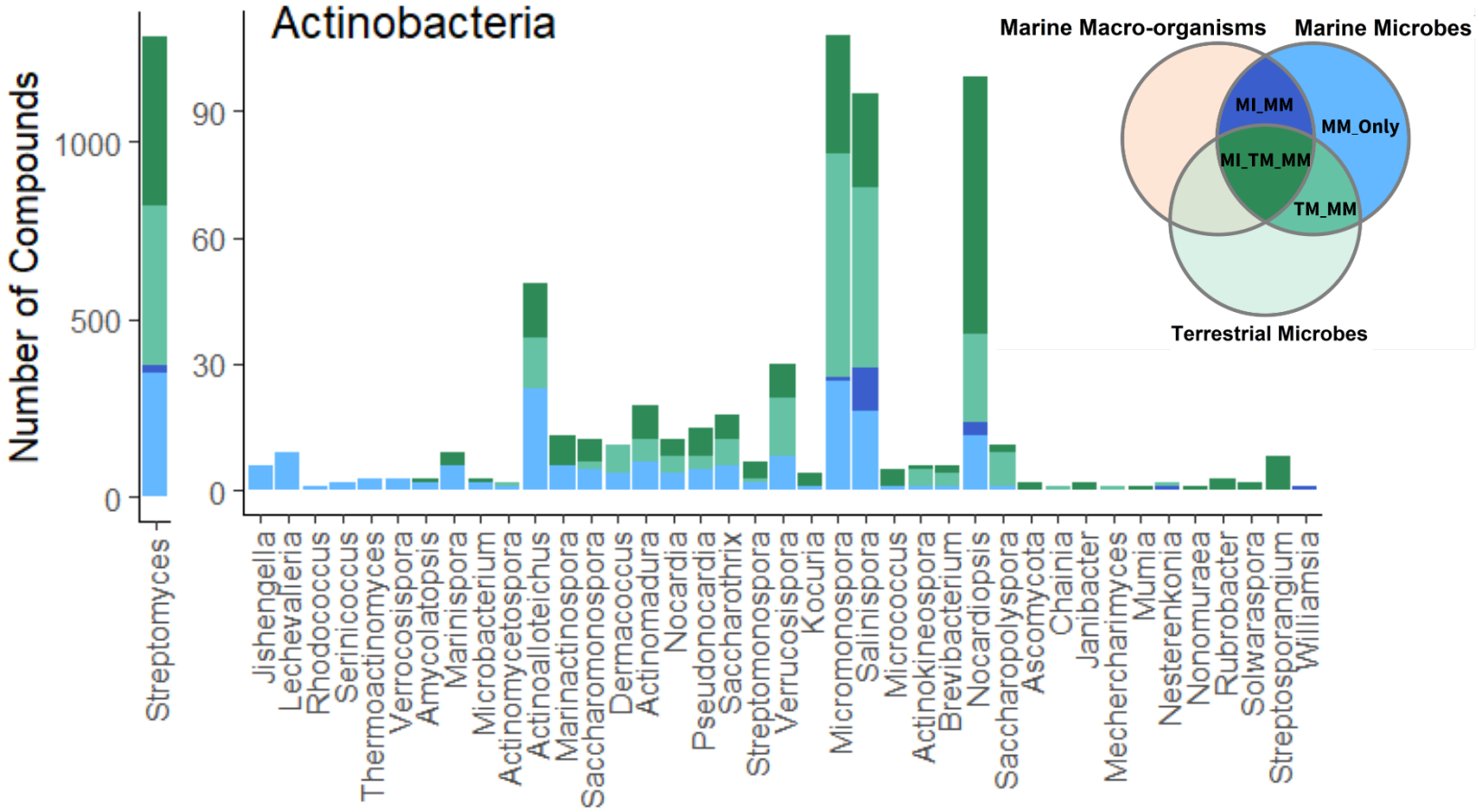


Fig. S29 All the NPs isolated from **marine Actinobacteria separated by genus**. Different colours represent clusterings with NPs from other groups, based on Murko scaffold analysis. NPs that cluster only amongst other marine microbial NPs are represented in light blue, NPs that are nested amongst terrestrial microbial NPs are light green, NPs that are nested among NPs from marine macro-organisms are dark blue, and NPs that are nested amongst terrestrial microbial NPs and NPs from marine macro-organisms are shown in dark green.

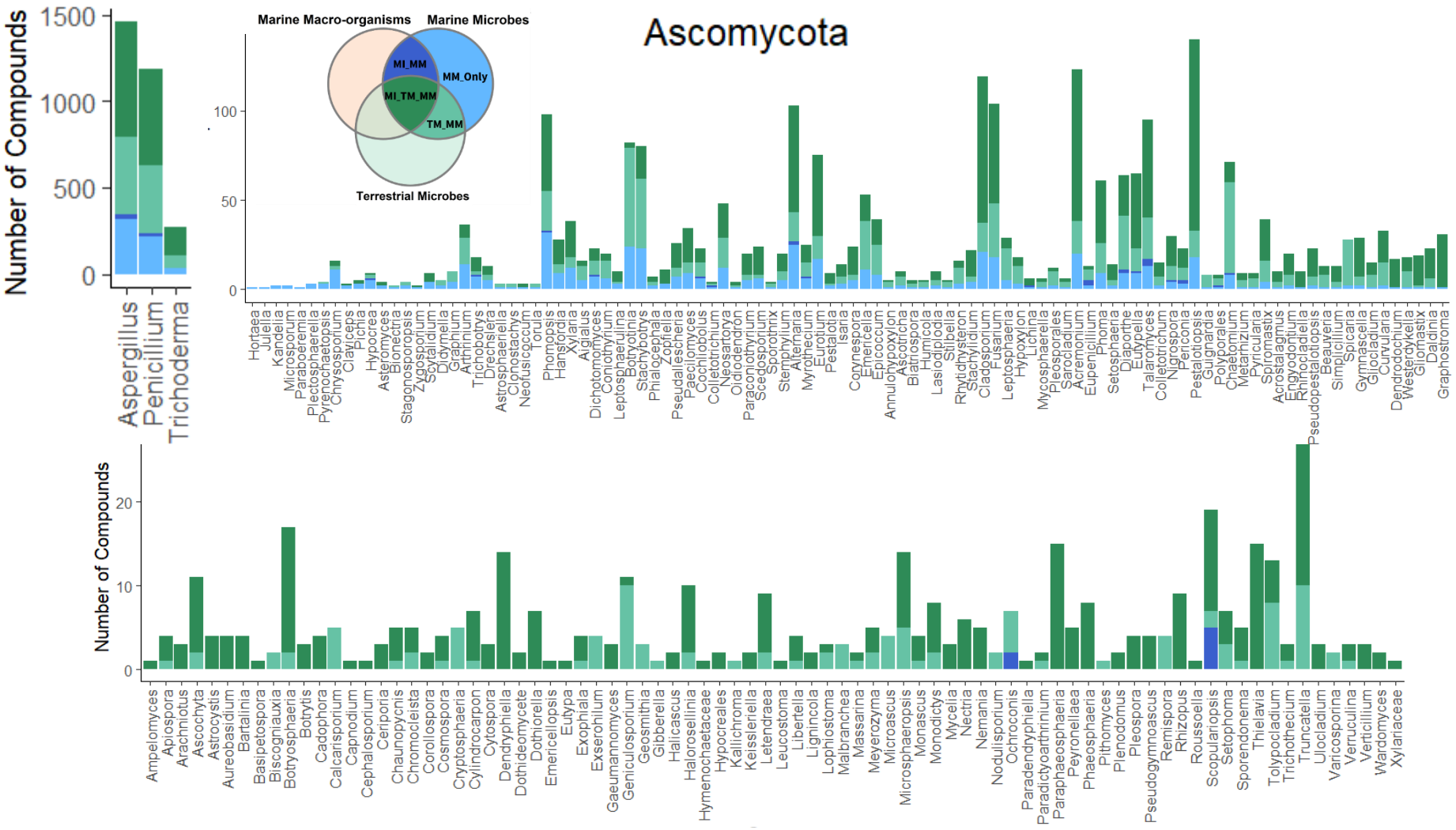


Fig. S30 All the NPs isolated from **marine Ascomycota separated by genus**. Different colours represent clusterings with NPs from other groups, based on Murko scaffold analysis. NPs that cluster only amongst other marine microbial NPs are represented in light blue, NPs that are nested amongst terrestrial microbial NPs are light green, NPs that are nested among NPs from marine macro-organisms are dark blue, and NPs that are nested amongst terrestrial microbial NPs and NPs from marine macro-organisms are shown in dark green.

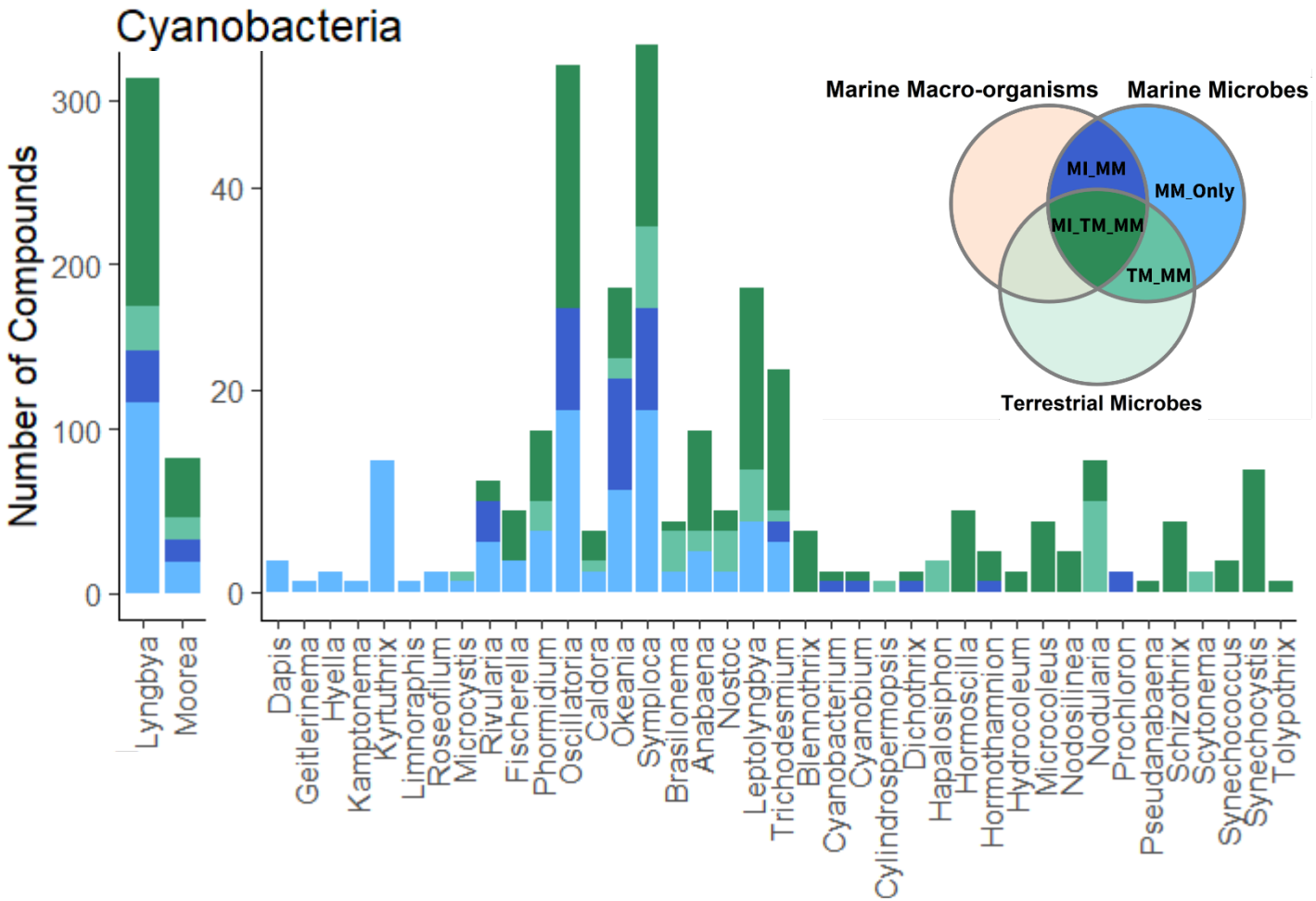


Fig. S31 All NPs isolated from **marine Cyanobacteria separated by genus**. Different colours represent clusterings with NPs from other groups, based on Murko scaffold analysis. NPs that cluster only amongst other marine microbial NPs are represented in light blue, NPs that are nested amongst terrestrial microbial NPs are light green, NPs that are nested among NPs from marine macro-organisms are dark blue, and NPs that are nested amongst terrestrial microbial NPs and NPs from marine macro-organisms are shown in dark green.

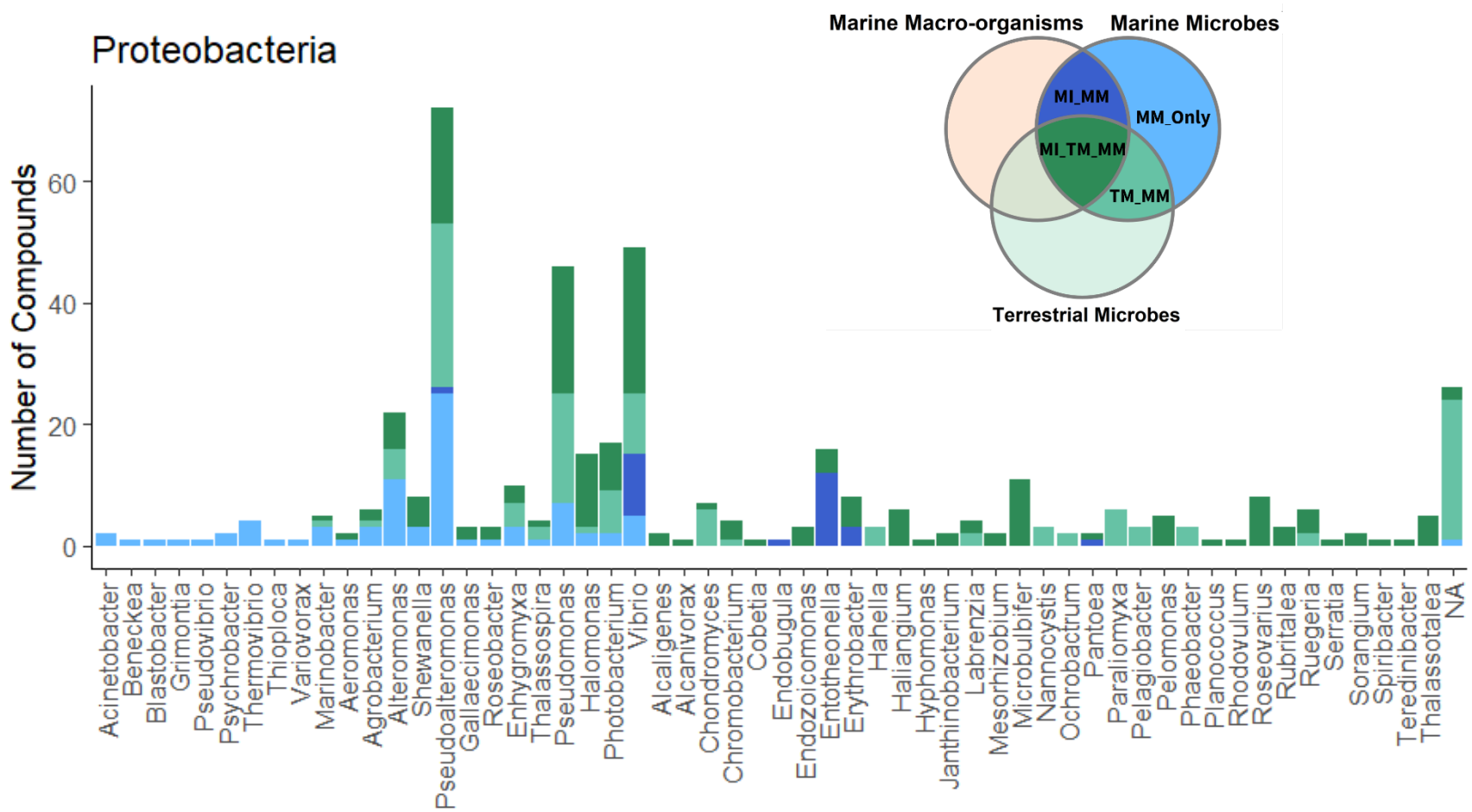


Fig. S32 All NPs isolated from **marine Proteobacteria separated by genus**. Different colours represent clusterings with NPs from other groups, based on Murko scaffold analysis. NPs that cluster only amongst other marine microbial NPs are represented in light blue, NPs that are nested amongst terrestrial microbial NPs are light green, NPs that are nested among NPs from marine macro-organisms are dark blue, and NPs that are nested amongst terrestrial microbial NPs and NPs from marine macro-organisms are shown in dark green.

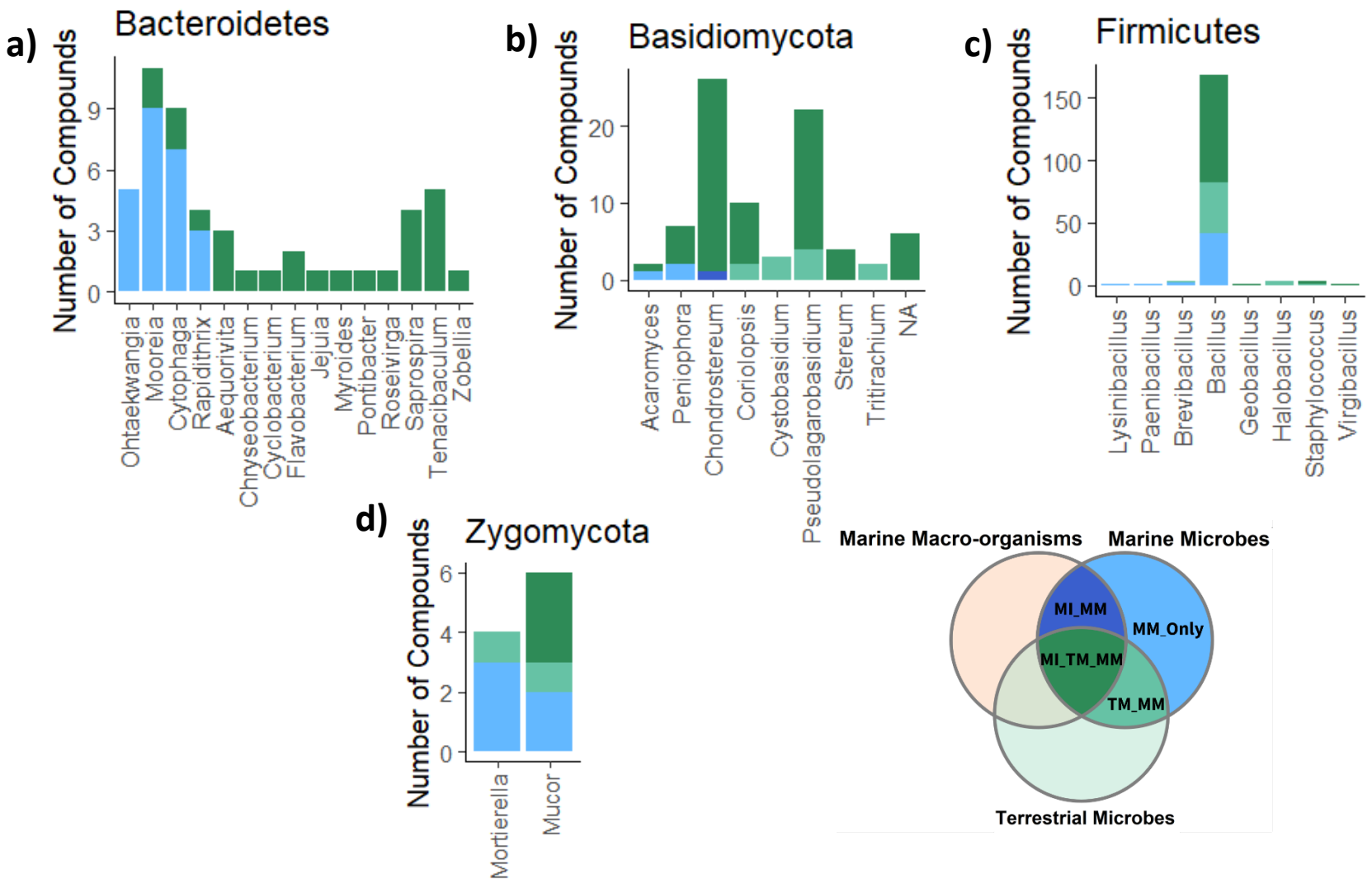


Fig. S33 All the NPs isolated from **marine Bacteriodeted a), Basidiomycota b), Firmicutes c) and Zygomycota d)**. Different colours represent clusterings with NPs from other groups, based on Murko scaffold analysis. NPs that cluster only amongst other marine microbial NPs are represented in light blue, NPs that are nested amongst terrestrial microbial NPs are light green, NPs that are nested among NPs from marine macro-organisms are dark blue, and NPs that are nested amongst terrestrial microbial NPs and NPs from marine macro-organisms are shown in dark green.

Top Producers of Marine Only Compounds

Table 2 List of the top producers of marine microbial NPs that do not nest with terrestrial NPs, according to the scaffold analysis. On the left are the top 20 genera according to the absolute number of marine only NPs and on the right are the top 20 genera ranked by proportion of NPs that cluster only with marine microbial NPs versus total compounds produced in the marine environment by this genus. For proportion calculations, genera that produced less than 5 compounds in total were removed from the list.

Genus	# of Marine Only
Streptomyces	351
Aspergillus	314
Penicillium	218
Moorea	78
Lyngbya	57
Bacillus	42
Phomopsis	32
Trichoderma	31
Micromonospora	26
Alternaria	25
Pseudoalteromonas	25
Actinoalloteichus	24
Botryotinia	24
Stachybotrys	23
Cladosporium	21
Acremonium	20
Salinispora	19
Fusarium	18
Oscillatoria	18
Pestalotiopsis	18

Genus	# of Marine Only	Total # of Molecules	Proportion of Marine Only
Jishengella	6	6	1.0
Kyrtuthrix	13	13	1.0
Lechevalieria	9	9	1.0
Mooreia	9	11	0.8
Cytophaga	7	9	0.8
Chrysosporium	11	16	0.7
Marinispora	6	9	0.7
Hypocrea	5	9	0.6
Lyngbya	57	104	0.5
Agrobacterium	3	6	0.5
Alteromonas	11	22	0.5
Actinoalloteichus	24	49	0.5
Marinactinospora	6	13	0.5
Rivularia	5	11	0.5
Scytalidium	4	9	0.4
Saccharomonospora	5	12	0.4
Graphium	4	10	0.4
Arthrinium	14	36	0.4
Trichobotrys	7	18	0.4
Drechslera	5	13	0.4

Similarity Analysis of the Scaffolds

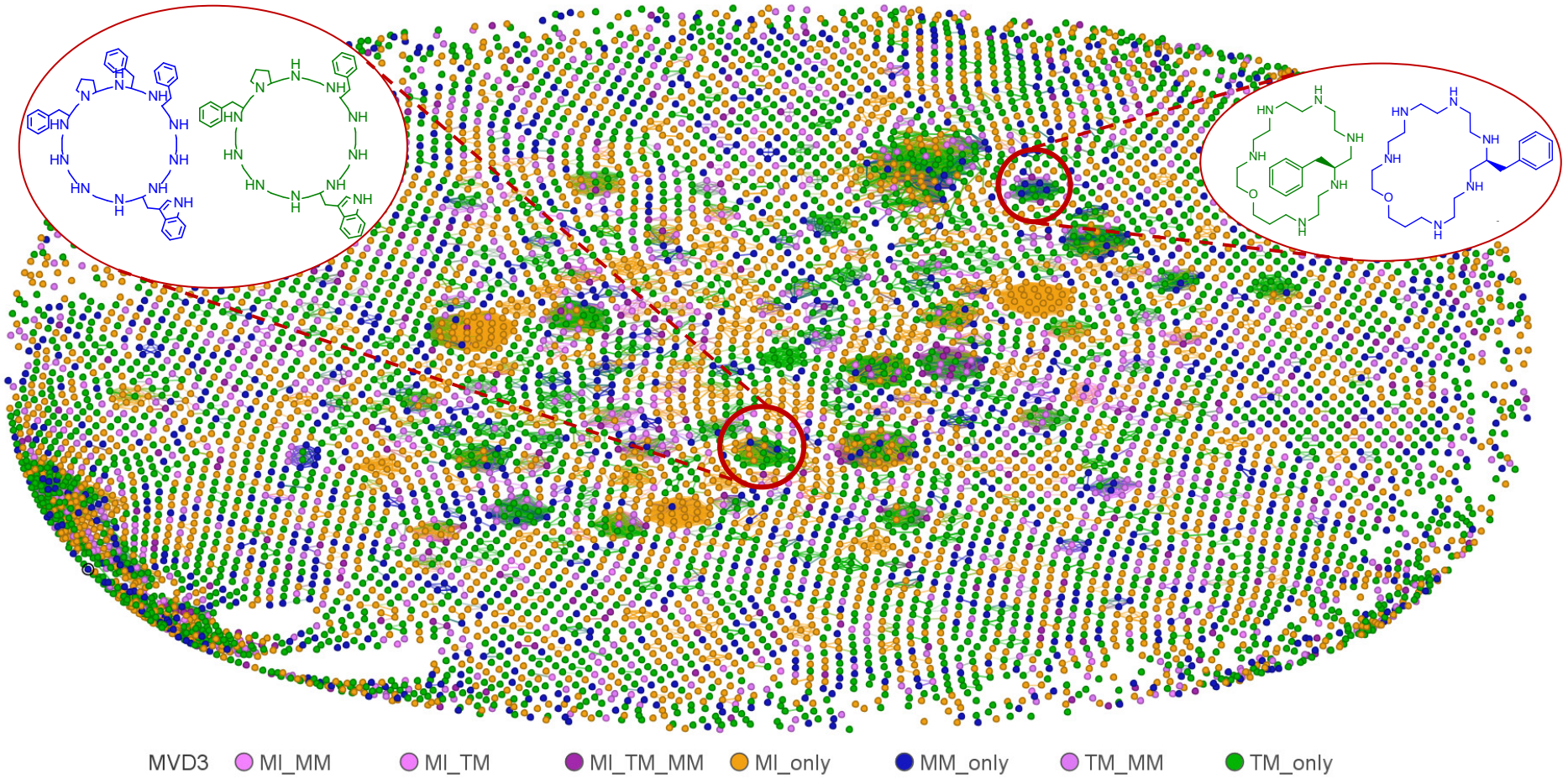


Fig. S34 Similarity analysis of the scaffolds. Scaffolds that cluster within a classification (same colour) are not a problem, they only inflate the scaffold count, which is eliminated once scaffolds are translated back into the actual number of NPs. However, cluster where there are mixed colours, can be problematic since they show a potential underestimate of the overlap between the different kingdoms. Clusters containing terrestrial only scaffolds (green) and marine only scaffolds (blue), are an example of where cyclic peptides have been sorted into different scaffolds but are structurally very alike (two examples are shown in the red circles).

World Map

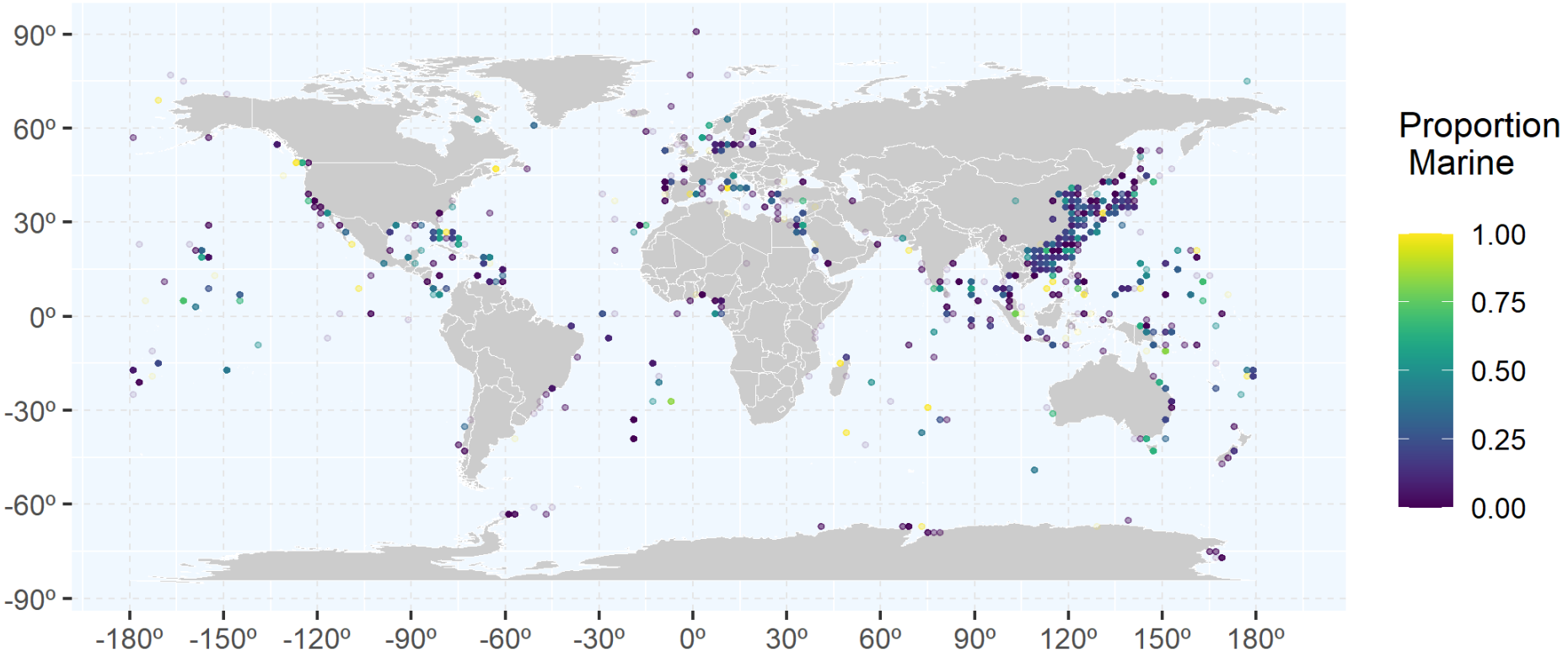
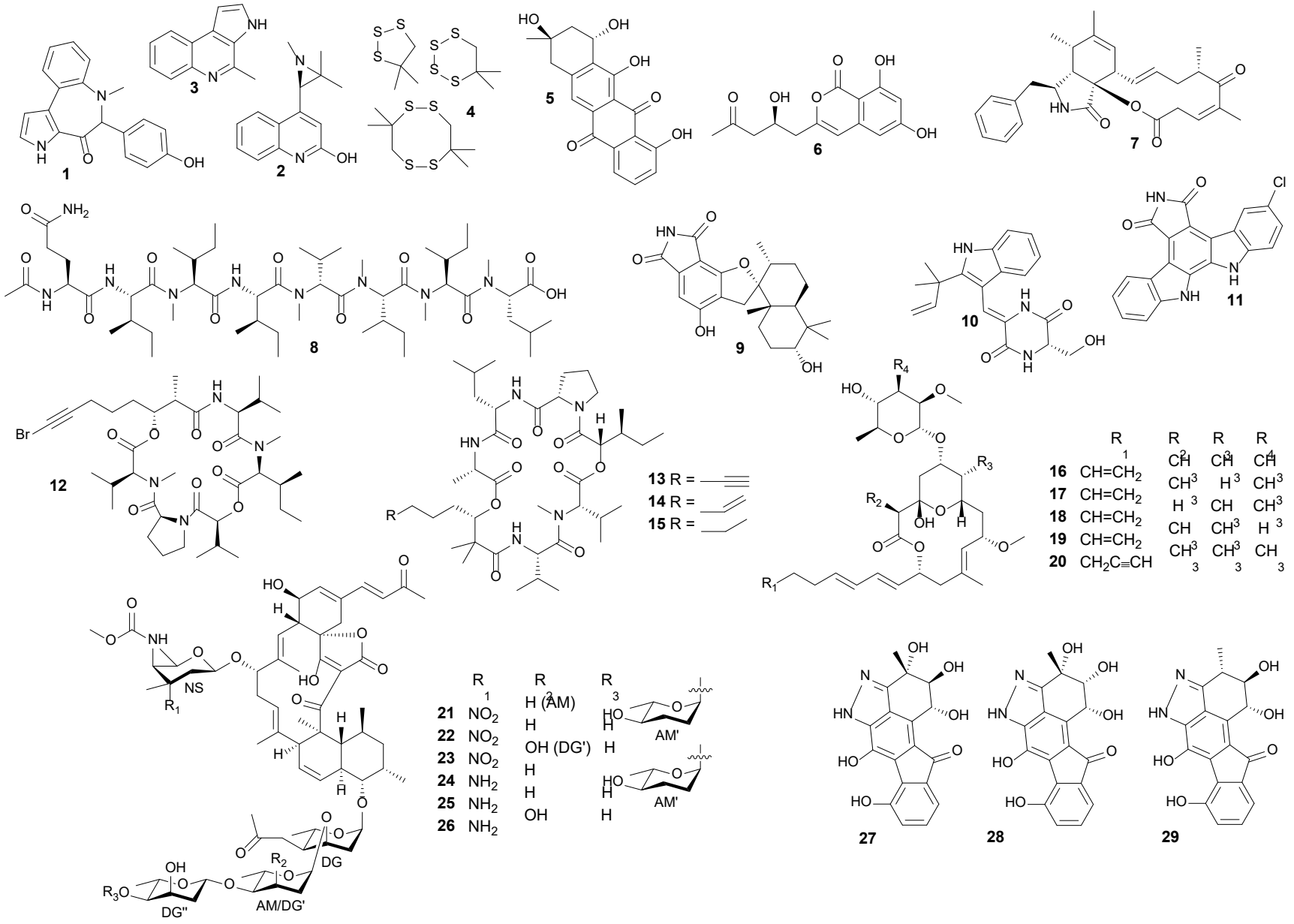
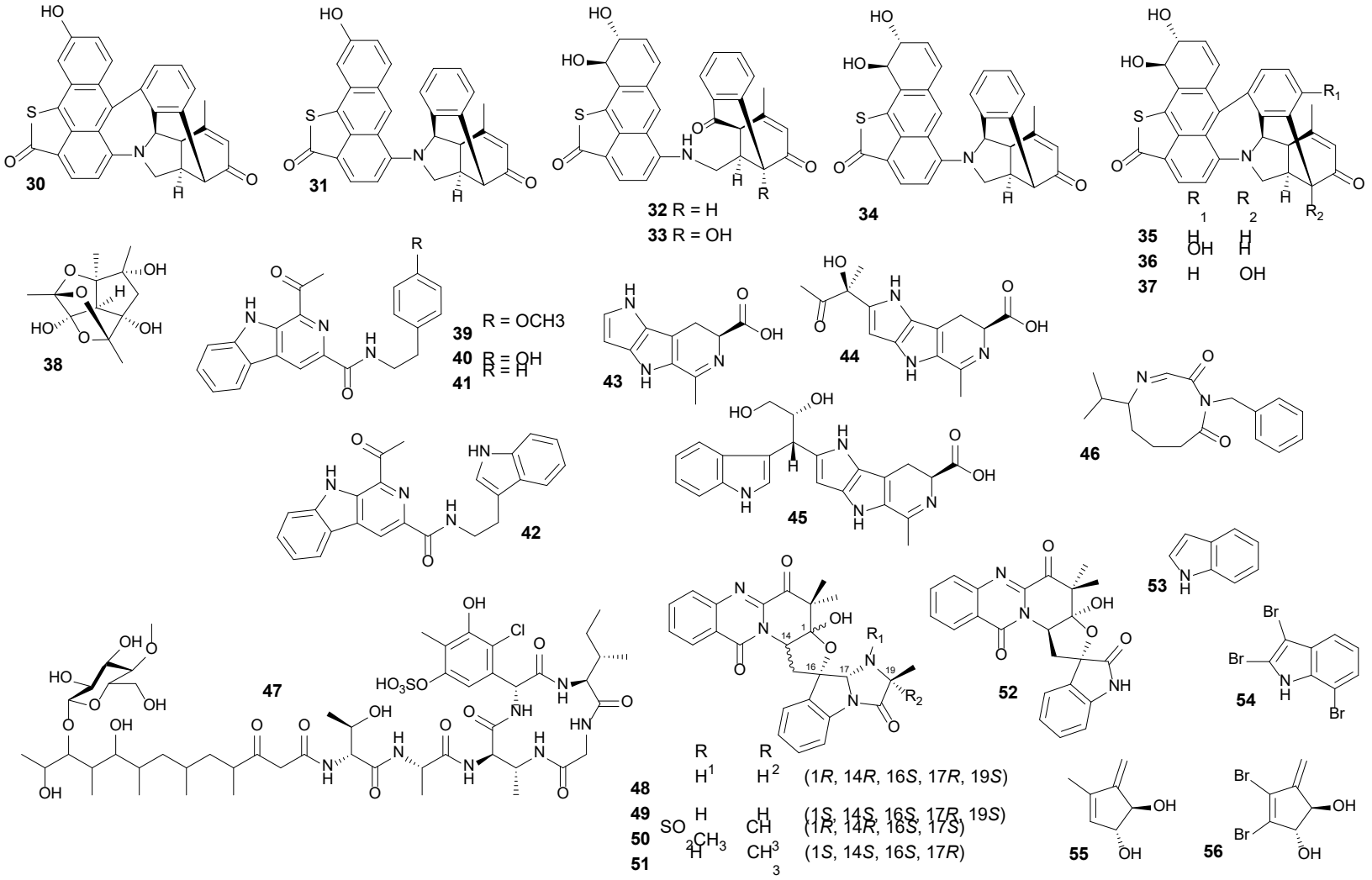
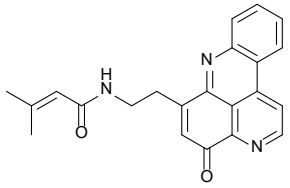


Fig. S35 World map, showing the distribution of the 8,577 marine microbial NPs that have reported coordinates for microorganism collection. The colours show the proportion of marine microbial NPs found at a specific location (2° x 2° grid cell) that are unique to the marine environment and don't cluster with terrestrial NPs, according to the scaffold analysis. The number of NPs found at each location varied substantially, and to represent low numbers of NPs at a location, more transparent dots (log10 scale) have been used.

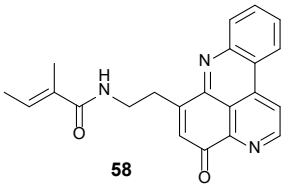
Chemical Structures mentioned in the Highlight



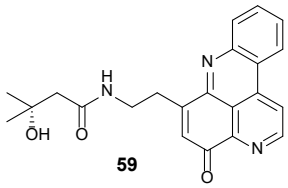

30
31
32 R = H
33 R = OH
34
35 H H
36 OH H
37 H OH
38
39 R = OCH3
40 R = OH
41 R = H
42
43
44
45
46
47
48 H H (1R, 14R, 16S, 17R, 19S)
49 H H (1S, 14S, 16S, 17R, 19S)
50 SO2CH3 CH3 (1R, 14R, 16S, 17S)
51 H CH3 (1S, 14S, 16S, 17R)
52
53
54
55
56



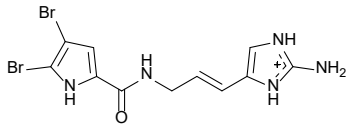
57



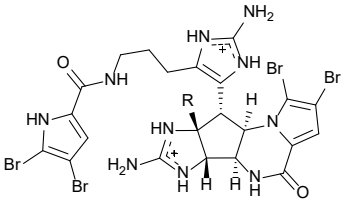
58



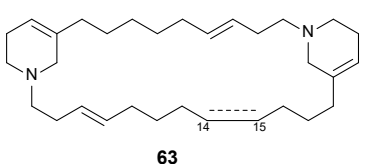
59



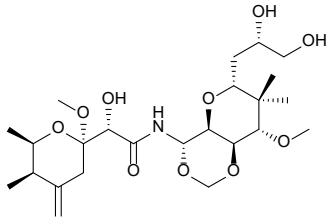
60



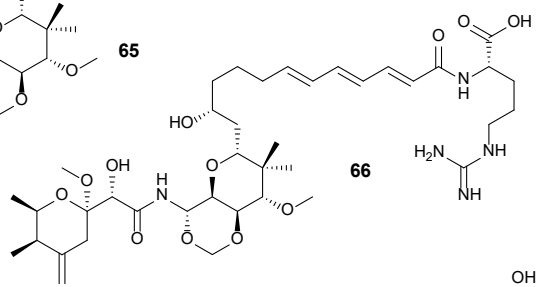
61 R ≡ OMe
62 R ≡ OH



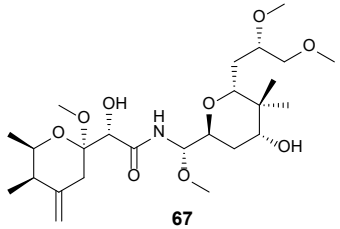
63
64 $\Delta^{14,15}$



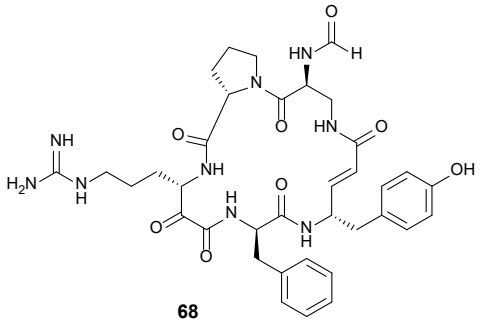
65



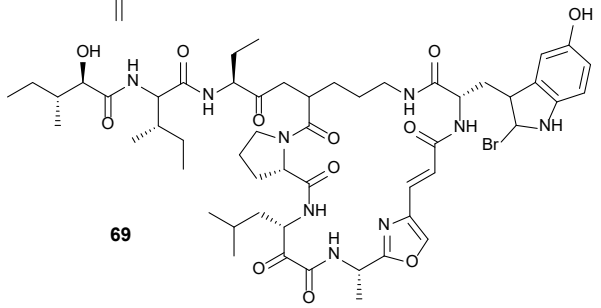
66



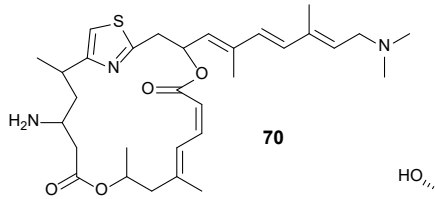
67



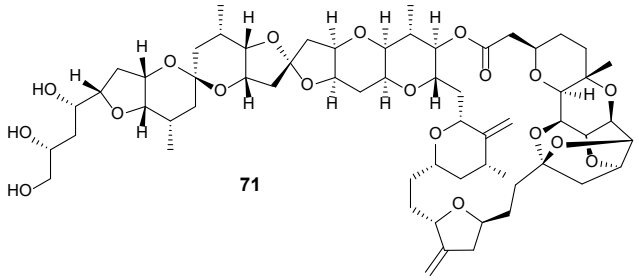
68



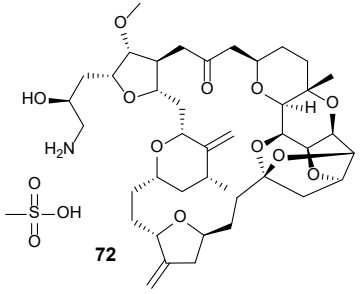
69



70



71



72