# Supplementary Information

# Dereplication of Peptidic Natural Products Through Database Search of Mass Spectra

**Hosein Mohimani[1], Alexey Gurevich[2], Alla Mikheenko[2], Neha Garg[3], Louis-Felix Nothias[3], Akihiro Ninomiya[4], Kentaro Takada[4], Pieter C. Dorrestein[3,5], and Pavel A. Pevzner[1]**

[1]Department of Computer Science and Engineering, University of California, San Diego, La Jolla, California, USA
[2]Center for Algorithmic Biotechnology, Institute of Translational Biomedicine, St. Petersburg State University, St. Petersburg, Russia
[3]Collaborative Mass Spectrometry Innovation Center, Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California, San Diego, La Jolla, California, USA
[4]Laboratory of Aquatic Natural Products Chemistry, Graduate School of Agricultural and Life Sciences, The University of Tokyo, Tokyo, Japan.
[5]Department of Pharmacology, University of California, San Diego, La Jolla, California, USA

## Supplementary Results

## Supplementary Tables

**Supplementary Table 1.** Number of peptides with the specified number of generalized peptide bonds in AntiMarin, StreptomeDB, MIBiG and DNP.

| dataset | #compounds | 0-bond | 1-bond | 2-bonds | 3-bonds | 4-bonds | 5-bonds | 6-bonds | 7-bonds | 8-bonds | 9-bonds or more |
|---|---|---|---|---|---|---|---|---|---|---|---|
| AntiMarin | 60908 | 36337 | 12107 | 4767 | 2204 | 1207 | 790 | 793 | 560 | 359 | 1784 |
| StreptomeDB | 3992 | 2287 | 796 | 399 | 116 | 101 | 69 | 65 | 56 | 28 | 75 |
| MIBiG | 963 | 443 | 190 | 86 | 44 | 30 | 29 | 28 | 19 | 16 | 78 |
| DNP | 254735 | 172100 | 49259 | 18154 | 6719 | 3099 | 2100 | 1191 | 773 | 413 | 913 |

**Supplementary Table 2.** Information about spectral datasets. The accuracy of peaks was set to 0.01 Da for precursor ion and 0.02 Da for product ions for high-precision datasets and 0.05 Da for precursor ion and 0.5 Da for product ions for low-precision datasets.

| dataset | number of spectra | GNPS IDs |
|---|---|---|
| $Spectra_{GNPS}$ | ≈ 72,000,000 | 201 datasets |
| $Spectra_4$ (S. roseosporus) | 29768 | MSV000078577 |
| $Spectra_4$ (Bacillus and Pseudomonas) | 54472 | MSV000078552 |
| $Spectra_4$ (Cubist strains) | 32434 | MSV000078607 |
| $Spectra_4$ (Chinese marine strains) | 412079 | MSV000078557 |
| $Spectra_{Fungi}$ | 2475907 | MSV000079098 |
| $Spectra_{Acti}$ | 5598651 | MSV000078787, MSV000078936, MSV000078937, MSV000078836, MSV000078839 |
| $Spectra_{Pseu}$ | 3529556 | MSV000078803, MSV000078817, MSV000078635, MSV000078606 |
| $Spectra_{Cyan}$ | 11921457 | MSV000078568 |
| $Spectra_{Acti36}$ | 974988 | MSV000078836, MSV000078839 |

**Supplementary Table 3.** The list of 129 PNPs (in the increasing order of their p-values) identified by DEREPLICATOR. search of the GNPS molecular network against AntiMarin database with p-value threshold $10^{-11}$. The maximum precursor mass error was set to 0.05Da. The columns specify the ID of the GNPS dataset, the PNP from AniMarin, structure (cyclic, branch cyclic, or linear), category (peptide, lipopeptide, NRPS/PKS) , p-value, SPCscore, the number of peaks in the spectrum, and the number of generalized peptide bonds.

| # | GNPS ID | PNP | structure | category | p-value | SPC score | # peaks | # bonds |
|---|---|---|---|---|---|---|---|---|
| 1 | 78710 | Cyclosporin-A | cyc | peptide | 2.3E-39 | 61 | 142 | 11 |
| 2 | 79050 | hormothamnin-A | cyc | peptide | 1.8E-32 | 66 | 137 | 11 |
| 3 | 79050 | laxaphycin-B | cyc | lipo | 2.4E-30 | 63 | 126 | 12 |
| 4 | 78567 | Cyclosporin-C | cyc | peptide | 2.5E-29 | 39 | 91 | 11 |
| 5 | 78567 | Cyclosporin-I | cyc | prptide | 2.4E-27 | 68 | 90 | 11 |
| 6 | 78567 | cyclosporin | cyc | peptide | 4.6E-27 | 43 | 103 | 11 |
| 7 | 78552 | Ayfivin | bcyc | peptide | 2.0E-26 | 26 | 100 | 11 |
| 8 | 79050 | Laxaphycin-E | cyc | lipo | 4.3E-25 | 62 | 91 | 11 |
| 9 | 78635 | Tolaasin-I | bcyc | lipo | 6.5E-23 | 28 | 138 | 18 |
| 10 | 78803 | OrfamidE-C | bcyc | lipo | 2.0E-22 | 33 | 97 | 10 |
| 11 | 79050 | Laxaphycin-D | cyc | lipo | 1.4E-21 | 38 | 76 | 12 |
| 12 | 78552 | Tolaasin-B | bcyc | lipo | 2.5E-21 | 21 | 155 | 18 |
| 13 | 78830 | ID=16908 | linear | peptide | 2.5E-20 | 9 | 71 | 6 |
| 14 | 79050 | laxaphycin-B3 | cyc | lipo | 2.8E-20 | 39 | 129 | 12 |
| 15 | 79029 | Puwainaphycin-D | cyc | NRPS-PKS | 1.7E-19 | 36 | 157 | 10 |
| 16 | 78568 | dolastatin-12 | cyc | peptide | 6.1E-19 | 25 | 139 | 8 |
| 17 | 78577 | A-21978-C0 | bcyc | lipo | 6.3E-19 | 26 | 125 | 13 |
| 18 | 78635 | Bacillomycin-D2 | cyc | lipo | 9.6E-19 | 26 | 85 | 8 |
| 19 | 78853 | ID=8780 | cyc | lipo | 9.8E-19 | 21 | 70 | 7 |
| 20 | 78708 | Cycloaspeptide-A | cyc | peptide | 1.1E-18 | 13 | 139 | 5 |
| 21 | 78635 | massetolide-A | bcyc | lipo | 1.5E-18 | 23 | 73 | 9 |
| 22 | 78787 | WS-7338-B | cyc | peptide | 1.6E-18 | 15 | 45 | 5 |
| 23 | 78602 | Bacirtsin-2 | cyc | peptide | 1.8E-18 | 23 | 70 | 7 |
| 24 | 78567 | FR-901459 | cyc | peptide | 2.5E-18 | 54 | 87 | 11 |
| 25 | 78557 | C14-Surfactin | cyc | lipo | 5.6E-18 | 24 | 152 | 7 |
| 26 | 78937 | pumilacidin-B | cyc | lipo | 1.3E-17 | 29 | 62 | 7 |
| 27 | 78892 | majusculamide-C | cyc | peptide | 1.7E-17 | 28 | 162 | 8 |
| 28 | 78552 | Tolaasin-C | linear | lipo | 1.9E-17 | 17 | 154 | 19 |
| 29 | 78787 | Surfactin-A1 | cyc | lipo | 2.1E-17 | 26 | 77 | 7 |
| 30 | 78787 | cotteslosin-A | cyc | peptide | 2.2E-17 | 15 | 48 | 5 |
| 31 | 78787 | mojavenisn-A | cyc | lipo | 5.4E-17 | 23 | 93 | 8 |
| 32 | 78936 | WS-7338-A | cyc | peptide | 6.9E-17 | 12 | 44 | 5 |
| 33 | 78936 | Surfactin-C15ai | cyc | lipo | 7.9E-17 | 23 | 75 | 7 |
| 34 | 78635 | Pseudophomin-A | bcyc | lipo | 9.6E-17 | 24 | 82 | 9 |
| 35 | 78606 | Orfamide-A | bcyc | lipo | 1.4E-16 | 36 | 96 | 10 |
| 36 | 78552 | Mycosubtilin-III | cyc | lipo | 1.4E-16 | 15 | 75 | 8 |
| 37 | 78577 | Stenothricin-IV | bcyc | lipo | 1.7E-16 | 22 | 89 | 9 |
| 38 | 78556 | Surfactin-C15 | cyc | lipo | 2.1E-16 | 22 | 119 | 7 |
| 39 | 78937 | Surfactin-C13ai | cyc | lipo | 2.2E-16 | 23 | 57 | 7 |
| 40 | 78797 | valinomycin | cyc | peptide | 2.9E-16 | 30 | 128 | 6 |
| 41 | 79029 | Calophycin | cyc | peptide | 3.4E-16 | 45 | 138 | 10 |
| 42 | 78556 | Surfactin-C14i | cyc | lipo | 3.7E-16 | 19 | 92 | 7 |
| 43 | 78602 | SNA-60-367-19 | bcyc | lipo | 3.9E-16 | 24 | 115 | 10 |
| 44 | 78604 | Isariin | cyc | lipo | 6.1E-16 | 8 | 55 | 5 |
| 45 | 78936 | Ilamycin-B1 | cyc | peptide | 7.5E-16 | 22 | 72 | 7 |
| 46 | 79050 | laxaphycin-B2 | cyc | lipo | 9.1E-16 | 48 | 95 | 12 |
| 47 | 79054 | Bacillopeptin-B | cyc | lipo | 1.0E-15 | 27 | 81 | 8 |
| 48 | 78557 | Glumamycin | bcyc | lipo | 1.2E-15 | 26 | 90 | 12 |
| 49 | 78830 | versicoloritide-A | cyc | peptide | 1.4E-15 | 16 | 45 | 5 |
| 50 | 79050 | almiramide-A | linear | lipo | 1.6E-15 | 7 | 38 | 6 |
| 51 | 78937 | Surfactin-C14i | cyc | lipo | 1.8E-15 | 22 | 77 | 7 |
| 52 | 78787 | Kurstakin-2 | bcyc | lipo | 2.9E-15 | 9 | 68 | 7 |
| 53 | 78635 | Viscosinamide | bcyc | lipo | 3.7E-15 | 22 | 95 | 9 |
| 54 | 78830 | stylissamide-B | cyc | peptide | 5.0E-15 | 24 | 44 | 7 |
| 55 | 78635 | massetolide-E | bcyc | lipo | 5.1E-15 | 22 | 82 | 9 |
| 56 | 78617 | Amychelin | linear | peptide | 1.0E-14 | 6 | 55 | 6 |
| 57 | 78602 | SNA-60-367-8 | bcyc | lipo | 1.2E-14 | 25 | 118 | 10 |
| 58 | 78936 | Surfactin-C13ai | cyc | lipo | 1.6E-14 | 24 | 76 | 7 |
| 59 | 78891 | Anabaenopeptin | bcyc | peptide | 1.6E-14 | 14 | 57 | 7 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 60 | 78936 | Ilamycin-B2 | cyc | peptide | 1.8E-14 | 17 | 36 | 7 |
| 61 | 78602 | Surfactin-D | cyc | lipo | 2.3E-14 | 23 | 75 | 7 |
| 62 | 78937 | pumilacidin-E | cyc | lipo | 2.7E-14 | 25 | 68 | 7 |
| 63 | 78577 | A-21978-C2 | bcyc | lipo | 2.8E-14 | 24 | 140 | 13 |
| 64 | 78577 | Stenothricin-I | bcyc | lipo | 3.0E-14 | 20 | 90 | 9 |
| 65 | 78803 | massetolide-F | bcyc | lipo | 3.0E-14 | 24 | 95 | 9 |
| 66 | 79029 | Puwainaphycin-C | cyc | NRPS-PKS | 3.5E-14 | 30 | 174 | 10 |
| 67 | 78584 | Anthranicine | cyc | peptide | 4.2E-14 | 12 | 58 | 6 |
| 68 | 78577 | A-21978-C3 | bcyc | lipo | 4.3E-14 | 20 | 120 | 13 |
| 69 | 78803 | Massetolide-E | bcyc | lipo | 4.4E-14 | 24 | 94 | 9 |
| 70 | 78577 | Stenothricin-III | bcyc | lipo | 5.2E-14 | 25 | 95 | 9 |
| 71 | 78803 | massetolide-C | bcyc | lipo | 5.2E-14 | 16 | 79 | 9 |
| 72 | 78937 | Surfactin-C14i | cyc | lipo | 5.4E-14 | 21 | 66 | 7 |
| 73 | 78577 | A-21978-C1 | bcyc | lipo | 5.7E-14 | 28 | 140 | 13 |
| 74 | 78936 | Surfactin-C13ai | cyc | lipo | 5.8E-14 | 24 | 74 | 7 |
| 75 | 78568 | dragonamide-E | linear | lipo | 7.2E-14 | 5 | 51 | 5 |
| 76 | 78892 | carriebowmide | cyc | peptide | 8.2E-14 | 18 | 101 | 6 |
| 77 | 78891 | homophymine-E1 | bcyc | peptide | 9.7E-14 | 19 | 45 | 11 |
| 78 | 78568 | almiramide-C | linear | lipo | 1.2E-13 | 7 | 52 | 6 |
| 79 | 78602 | Surfactin-C14 | cyc | lipo | 1.3E-13 | 23 | 65 | 7 |
| 80 | 78830 | Aureobasidin-R | cyc | peptide | 1.3E-13 | 20 | 75 | 8 |
| 81 | 78805 | Amphibactin-E | linear | lipo | 1.5E-13 | 8 | 46 | 5 |
| 82 | 78635 | bacillomycin-D2 | cyc | lipo | 1.5E-13 | 26 | 85 | 8 |
| 83 | 78568 | dolastatin-14 | cyc | peptide | 1.6E-13 | 16 | 69 | 7 |
| 84 | 78787 | Kurstakin-4 | bcyc | lipo | 1.8E-13 | 8 | 64 | 7 |
| 85 | 78557 | nocardamine | cyc | peptide | 2.2E-13 | 9 | 89 | 4 |
| 86 | 78789 | Aureobasidin-S2b | cyc | peptide | 2.3E-13 | 12 | 25 | 8 |
| 87 | 78797 | Surfactin-C15 | cyc | lipo | 2.7E-13 | 17 | 116 | 7 |
| 88 | 78937 | fijimycin-A | bcyc | peptide | 2.8E-13 | 19 | 75 | 7 |
| 89 | 78552 | Bacitracin-B3 | bcyc | peptide | 3.5E-13 | 23 | 115 | 11 |
| 90 | 78805 | Amphibactin-F | linear | lipo | 3.9E-13 | 6 | 96 | 5 |
| 91 | 78568 | dragonamide | linear | lipo | 3.9E-13 | 6 | 57 | 5 |
| 92 | 78803 | massetolide-G | bcyc | lipo | 3.9E-13 | 19 | 95 | 9 |
| 93 | 78936 | pumilacidin-C | cyc | lipo | 4.4E-13 | 24 | 72 | 7 |
| 94 | 78688 | Lichenysin-G5a | cyc | lipo | 4.6E-13 | 20 | 66 | 7 |
| 95 | 78936 | Etamycin-VI-2 | bcyc | peptide | 5.7E-13 | 20 | 70 | 7 |
| 96 | 78584 | Fungisporin | cyc | peptide | 6.2E-13 | 34 | 49 | 8 |
| 97 | 78805 | Amphibactin-B | linear | lipo | 6.3E-13 | 9 | 55 | 5 |
| 98 | 78607 | a-Substance-Ib | linear | peptide | 7.5E-13 | 8 | 45 | 6 |
| 99 | 78635 | Plipastatin-A1 | bcyc | lipo | 8.3E-13 | 25 | 97 | 10 |
| 100 | 78607 | Kurstakin-1 | bcyc | lipo | 8.7E-13 | 7 | 60 | 7 |
| 101 | 78635 | massetolide-H | bcyc | lipo | 8.7E-13 | 20 | 89 | 9 |
| 102 | 78853 | [Ile2,4,7]Surfactin | cyc | lipo | 1.0E-12 | 22 | 142 | 7 |
| 103 | 78568 | almiramidE-B | linear | lipo | 1.0E-12 | 8 | 67 | 6 |
| 104 | 79054 | Surfactin-C1 | cyc | lipo | 1.2E-12 | 20 | 59 | 7 |
| 105 | 79029 | Microcystin-LR | cyc | peptide | 1.3E-12 | 20 | 119 | 7 |
| 106 | 78830 | Proferrioxamine-T7 | cyc | peptide | 1.3E-12 | 9 | 19 | 4 |
| 107 | 78853 | pumilacidin-A | cyc | lipo | 1.4E-12 | 22 | 75 | 7 |
| 108 | 78557 | Lavendomycin | linear | peptide | 1.4E-12 | 11 | 50 | 6 |
| 109 | 78936 | Lichenysin-A | cyc | lipo | 1.4E-12 | 23 | 69 | 7 |
| 110 | 78568 | dolastatin-10 | linear | peptide | 1.5E-12 | 5 | 94 | 5 |
| 111 | 78892 | normajusculamide | cyc | peptide | 1.8E-12 | 20 | 102 | 8 |
| 112 | 78839 | Desferrioxamine-X5 | cyc | peptide | 1.8E-12 | 10 | 160 | 4 |
| 113 | 78568 | majusculamide-C | cyc | peptide | 2.1E-12 | 20 | 96 | 8 |
| 114 | 78787 | Plipastatin-B2 | bcyc | lipo | 2.1E-12 | 27 | 123 | 10 |
| 115 | 78787 | Pumilacidin-C | cyc | lipo | 2.8E-12 | 24 | 69 | 7 |
| 116 | 78602 | SNA-60-367-23 | bcyc | lipo | 3.3E-12 | 25 | 114 | 10 |
| 117 | 78577 | Stenothricin-II | bcyc | lipo | 3.4E-12 | 24 | 90 | 9 |
| 118 | 78602 | SNA-60-367-17 | bcyc | lipo | 3.8E-12 | 25 | 115 | 10 |
| 119 | 79048 | Pepstatin-AC | linear | peptide | 3.9E-12 | 7 | 42 | 6 |
| 120 | 78797 | Surfactin-C13ai | cyc | lipo | 4.3E-12 | 15 | 116 | 7 |
| 121 | 79048 | Pepstatin-B | linear | peptide | 4.7E-12 | 7 | 33 | 6 |
| 122 | 78903 | Scytonemide-A | cyc | peptide | 5.0E-12 | 9 | 64 | 6 |
| 123 | 78552 | Plipastatin-A2 | bcyc | lipo | 5.8E-12 | 24 | 120 | 10 |
| 124 | 78936 | Amonabactin-P-750 | linear | lipo | 6.1E-12 | 5 | 21 | 6 |
| 125 | 78620 | Desferriferrichrome | cyc | peptide | 6.9E-12 | 21 | 45 | 6 |
| 126 | 78936 | Lichenysin-G1a | cyc | lipo | 7.7E-12 | 19 | 41 | 7 |
| 127 | 78803 | Orfamide-B | bcyc | lipo | 7.8E-12 | 26 | 107 | 10 |
| 128 | 78891 | Nostopeptolide-A1 | bcyc | lipo | 8.3E-12 | 17 | 80 | 9 |
| 129 | 79029 | Nostoginin-BN741 | linear | lipo | 8.4E-12 | 6 | 97 | 5 |

**Supplementary Table 4.** The list of 67 variants of PNPs (in the increasing order of their p-values) identified by DEREPLICATOR in the variable dereplication mode of the GNPS molecular network against the AntiMarin database with p-value threshold $10^{-11}$. 18 of them (27%) have characteristic offsets 14, 28, and 42 Da. Some of these variants could be adduct of known PNPs.

| PNP | spectrum mass | charge | mod mass | p-value |
|---|---|---|---|---|
| hormothamnin-A | 1182.7 | 1 | -13.9 | 4.0E-33 |
| laxaphycin-B2 | 705.4 | 2 | 30.0 | 1.7E-29 |
| hormothamnin-A | 1210.7 | 1 | 14.0 | 2.1E-29 |
| Tolaasin-C | 1005.1 | 2 | 3.9 | 8.0E-27 |
| hormothamnin-A | 1181.2 | 1 | -15.9 | 1.1E-23 |
| massetolide-A | 1253.8 | 1 | 113.0 | 2.4E-22 |
| none(ID=8780) | 1051.33 | 1 | 15.0 | 2.8E-22 |
| Tolaasin-B | 994.6 | 2 | 14.1 | 8.4E-22 |
| majusculamide-C | 971.6 | 1 | -13.9 | 8.8E-21 |
| desmethoxymajusculamide-C | 969.6 | 1 | 14.0 | 1.1E-20 |
| Tolaasin-B | 1001.7 | 2 | 28.0 | 1.8E-19 |
| dolastatin-12 | 999.5 | 1 | 29.9 | 1.9E-19 |
| Massetolide-E | 1239.8 | 1 | 127.1 | 6.6E-19 |
| valinomycin | 1125.2 | 1 | 14.0 | 8.4E-19 |
| dolastatin-12 | 955.5 | 1 | -14.0 | 3.3E-18 |
| hormothamnin-A | 1190.6 | 1 | -6.0 | 3.4E-18 |
| Desferriferrichrome | 689.3 | 1 | 1.0 | 6.8E-18 |
| Leu7-Surfactin-C14i-dimethyl-ester | 1054.6 | 1 | 3.9 | 1.4E-17 |
| Lichenysin-G5a | 1008.4 | 1 | -14.0 | 3.7E-17 |
| laxaphycin-B2 | 682.5 | 2 | -15.9 | 1.6E-16 |
| valinomycin | 1128.6 | 1 | 17.0 | 4.3E-16 |
| cotteslosin-A | 608.3 | 1 | -28.0 | 6.3E-16 |
| Desferrioxamine-X5 | 617.3 | 1 | 17.9 | 4.6E-15 |
| Stenothricin-component-I;-Stenothricin | 574.1 | 2 | 28.0 | 6.1E-15 |
| laxaphycin-B | 703.6 | 2 | 6.0 | 6.8E-15 |
| pumilacidin-B | 1022.6 | 1 | -14.0 | 6.9E-15 |
| Massetolide-E | 1126.7 | 1 | 14.0 | 7.2E-15 |
| Massetolide-E | 1225.7 | 1 | 113.1 | 9.1E-15 |
| Tolaasin-B | 1013.5 | 2 | 51.9 | 2.2E-14 |
| laxaphycin-B2 | 697.8 | 2 | 14.0 | 2.3E-14 |
| Bacirtsin-2;-Bacircine-3;-Surfactin-B;-S | 1008.6 | 1 | -12.0 | 3.0E-14 |
| almiramide-B | 743.6 | 1 | 18.0 | 3.5E-14 |
| Laxaphycin-E | 1168.7 | 1 | -56.0 | 7.8E-14 |
| pumilacidin-B | 1078.7 | 1 | 42.0 | 1.1E-13 |
| Pseudophomin-A | 1253.7 | 1 | 113.0 | 1.2E-13 |
| Bacillomycin-D2;-Raubitschek-substance | 1045.5 | 1 | 13.9 | 1.5E-13 |
| hormothamnin-A | 1201.8 | 1 | 5.0 | 1.6E-13 |
| almiramide-C | 743.6 | 1 | 16.0 | 2.0E-13 |
| Kurstakin-1 | 920.6 | 1 | 42.0 | 2.0E-13 |
| desmethoxymajusculamide-C | 985.6 | 1 | 29.9 | 2.4E-13 |
| Lichenysin-A | 1022.2 | 1 | 15.0 | 3.0E-13 |
| Pepstatin-AC | 702.3 | 1 | 58.0 | 4.0E-13 |
| dragonamide | 652.4 | 1 | -2.0 | 4.0E-13 |
| Pumilacidin-C;-(Surfactin-5) | 1064.7 | 1 | -13.9 | 4.1E-13 |
| Stenothricin-component-II | 566.8 | 2 | 27.7 | 4.2E-13 |
| Pseudophomin-A | 1094.7 | 1 | -45.9 | 6.9E-13 |
| valinomycin | 1121.5 | 1 | 10.0 | 6.9E-13 |
| Lichenysin-A | 994.4 | 1 | -14.0 | 7.4E-13 |
| dolastatin-10 | 800.1 | 1 | 16.0 | 1.2E-12 |
| carriebowmide | 847.5 | 1 | -33.9 | 1.2E-12 |
| majusculamide-C | 1003.6 | 1 | 18.0 | 1.4E-12 |
| dehydroxynocardamine | 601.8 | 1 | 16.9 | 1.4E-12 |
| dolastatin-10 | 814.8 | 1 | 27.9 | 1.5E-12 |
| Ile2,Val7-Surfactin-C15ai-monomethyl-e | 1064.7 | 1 | 28.0 | 1.8E-12 |
| laxaphycin-B | 686.1 | 2 | -22.0 | 2.0E-12 |
| majusculamide-C | 980.2 | 1 | -5.9 | 2.2E-12 |
| massetolide-F | 1154.2 | 1 | 26.0 | 2.4E-12 |
| massetolide-C | 1154.2 | 1 | -14.0 | 3.2E-12 |
| Desferrioxamine-X5 | 631.3 | 1 | 31.9 | 3.8E-12 |
| massetolide-A | 1126.4 | 1 | -14.0 | 4.0E-12 |
| valinomycin | 1100.6 | 1 | -11.0 | 4.6E-12 |
| Desferrioxamine-X5 | 517.3 | 1 | -82.0 | 5.0E-12 |
| Bacirtsin-2;-Bacircine-3;-Surfactin-B;-S | 1064.7 | 1 | 42.0 | 5.0E-12 |

| | | | | |
|---|---|---|---|---|
| pumilacidin-A | 1052.7 | 1 | 2.0 | 6.4E-12 |
| Lichenysin-G5a | 1036.3 | 1 | 14.0 | 7.8E-12 |
| desmethoxymajusculamide-C | 950.4 | 1 | -3.9 | 8.8E-12 |
| pumilacidin-C | 1096.7 | 1 | 18.0 | 9.5E-12 |

**Supplementary Table 5.** (A) Number of unique PNPs identified by DEREPLICATOR in $Spectra_{High}$ dataset at different p-value cut-offs, categorized by the number of generalized peptide bonds in the PNP. In the default mode, Dereplicator considers isotopic shifts +1 Da and +2 Da of the precursor mass and spectra with +1, +2 and +3 charge states. At a high FDR of 15%, DEREPLICATOR identified 23 peptides with three bonds, 26 peptides with four bonds, 76 peptides with five bonds and 325 peptides with six or more bonds. Entries in the table are colored green if they fall within 3% FDR, blue if they fall within 15% FDR, and red if they fall outside the 15% FDR. The entries with small number of PNPs for which FDR can not be reliably computed are shown in gray. (B) DEREPLICATOR results after removing isotopic shifts of the precursor mass (+1 Da and +2 Da) from consideration. Since considering isotopic shifts is counter-productive in the case of short peptides, the number of identifications at FDR of 15% increases to 30, 28 and 84 peptides with three bonds, four bonds and five bonds, respectively. (C) DEREPLICATOR results after removing spectra with +2 and +3 charge states. Since considering multicharged spectra is counter-productive in the case of short peptides, the the number of identifications at FDR of 15% increases to 47, 36 and 110 peptides with three bonds, four bonds and five bonds, respectively.

(A)

| p-value cut-off | #1-bond target | #1-bond decoy | #2-bond target | #2-bond decoy | #3-bond target | #3-bond decoy | #4-bond target | #4-bond decoy | #5-bond target | #5-bond decoy | #6-more target | #6-more decoy |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1.00E-04 | 701 | 2588 | 752 | 2121 | 604 | 1293 | 345 | 746 | 380 | 601 | 1448 | 1943 |
| 1.00E-05 | 621 | 2358 | 411 | 1396 | 326 | 786 | 194 | 421 | 279 | 422 | 998 | 1095 |
| 1.00E-06 | 184 | 632 | 97 | 295 | 174 | 281 | 117 | 203 | 204 | 264 | 710 | 635 |
| 1.00E-07 | 28 | 78 | 52 | 113 | 107 | 131 | 76 | 120 | 163 | 141 | 529 | 323 |
| 1.00E-08 | 10 | 14 | 28 | 52 | 64 | 44 | 51 | 40 | 132 | 69 | 432 | 161 |
| 1.00E-09 | 7 | 7 | 11 | 7 | 41 | 10 | 39 | 13 | 108 | 30 | 365 | 87 |
| 1.00E-10 | 0 | 0 | 6 | 0 | 30 | 5 | 28 | 6 | 93 | 15 | 325 | 42 |
| 1.00E-11 | 0 | 0 | 5 | 0 | 23 | 2 | 26 | 2 | 76 | 9 | 303 | 18 |
| 1.00E-12 | 0 | 0 | 4 | 0 | 16 | 1 | 22 | 1 | 67 | 5 | 283 | 7 |
| 1.00E-13 | 0 | 0 | 3 | 0 | 11 | 0 | 19 | 0 | 63 | 2 | 256 | 1 |
| 1.00E-14 | 0 | 0 | 1 | 0 | 8 | 0 | 16 | 0 | 57 | 2 | 238 | 1 |
| 1.00E-15 | 0 | 0 | 0 | 0 | 5 | 0 | 13 | 0 | 51 | 0 | 213 | 0 |

(B)

| p-value cut-off | #1-bond target | #1-bond decoy | #2-bond target | #2-bond decoy | #3-bond target | #3-bond decoy | #4-bond target | #4-bond decoy | #5-bond target | #5-bond decoy | #6-more target | #6-more decoy |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1.00E-04 | 521 | 1981 | 558 | 1654 | 450 | 1039 | 244 | 569 | 296 | 424 | 1448 | 1943 |
| 1.00E-05 | 462 | 1816 | 297 | 1060 | 223 | 576 | 134 | 272 | 216 | 236 | 998 | 1095 |
| 1.00E-06 | 139 | 442 | 86 | 214 | 127 | 203 | 85 | 126 | 163 | 132 | 710 | 635 |
| 1.00E-07 | 22 | 50 | 48 | 87 | 79 | 88 | 56 | 68 | 130 | 63 | 529 | 323 |
| 1.00E-08 | 8 | 10 | 27 | 40 | 52 | 21 | 44 | 16 | 112 | 27 | 432 | 161 |
| 1.00E-09 | 5 | 5 | 11 | 5 | 39 | 8 | 35 | 7 | 97 | 15 | 365 | 87 |
| 1.00E-10 | 0 | 0 | 6 | 0 | 30 | 4 | 28 | 4 | 84 | 10 | 325 | 42 |
| 1.00E-11 | 0 | 0 | 5 | 0 | 23 | 2 | 26 | 1 | 73 | 6 | 303 | 18 |
| 1.00E-12 | 0 | 0 | 4 | 0 | 16 | 1 | 22 | 1 | 65 | 4 | 283 | 7 |
| 1.00E-13 | 0 | 0 | 3 | 0 | 11 | 0 | 19 | 0 | 63 | 0 | 256 | 1 |
| 1.00E-14 | 0 | 0 | 1 | 0 | 8 | 0 | 16 | 0 | 57 | 0 | 238 | 1 |
| 1.00E-15 | 0 | 0 | 0 | 0 | 5 | 0 | 13 | 0 | 51 | 0 | 213 | 0 |

(C)

| p-value cut-off | #1-bond target | #1-bond decoy | #2-bond target | #2-bond decoy | #3-bond target | #3-bond decoy | #4-bond target | #4-bond decoy | #5-bond target | #5-bond decoy | #6-more target | #6-more decoy |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1.00E-04 | 350 | 1383 | 422 | 1159 | 338 | 716 | 172 | 371 | 208 | 234 | 1448 | 1943 |
| 1.00E-05 | 306 | 1237 | 141 | 385 | 130 | 183 | 90 | 108 | 154 | 97 | 998 | 1095 |
| 1.00E-06 | 22 | 50 | 56 | 67 | 97 | 86 | 61 | 53 | 121 | 37 | 710 | 635 |
| 1.00E-07 | 3 | 7 | 35 | 36 | 60 | 21 | 43 | 19 | 110 | 10 | 529 | 323 |
| 1.00E-08 | 0 | 0 | 16 | 1 | 47 | 5 | 36 | 2 | 99 | 3 | 432 | 161 |
| 1.00E-09 | 0 | 0 | 10 | 0 | 38 | 2 | 32 | 0 | 90 | 2 | 365 | 87 |
| 1.00E-10 | 0 | 0 | 6 | 0 | 29 | 0 | 26 | 0 | 80 | 2 | 325 | 42 |
| 1.00E-11 | 0 | 0 | 5 | 0 | 22 | 0 | 24 | 0 | 73 | 2 | 303 | 18 |
| 1.00E-12 | 0 | 0 | 4 | 0 | 16 | 0 | 21 | 0 | 65 | 1 | 283 | 7 |
| 1.00E-13 | 0 | 0 | 3 | 0 | 11 | 0 | 16 | 0 | 63 | 0 | 256 | 1 |
| 1.00E-14 | 0 | 0 | 1 | 0 | 8 | 0 | 14 | 0 | 56 | 0 | 238 | 1 |
| 1.00E-15 | 0 | 0 | 0 | 0 | 5 | 0 | 11 | 0 | 50 | 0 | 213 | 0 |

**Supplementary Table 6.** The list of 325 PNPs (in the increasing order of their p-values) identified by DEREPLICATOR search of $Spectra_{High}$ against AntiMarin database with p-value threshold $10^{-10}$. 180, 77, 40, and 64 PNPs were identified in $Spectra_{Fungi}$, $Spectra_{Acti}$, $Spectra_{Pseu}$, and $Spectra_{Cyan}$ datasets, respectively. The columns specify the ID of the GNPS dataset, the PNP from AniMarin, the p-value computed in the high-accuracy (HH) mode, the score computed in the high-accuracy (HH) mode, the p-value computed in the low-accuracy (HL) mode, number of peaks in the spectrum, PNP mass, and the number of generalized peptide bonds in the PNP. Double-cyclic structures are listed as "dcyc", Y-shape branched structures are listed as "y-branch", and branch-cyclic peptides with y-shape branch are listed as ybcyc. The maximum precursor mass error was set to 0.01Da. The HH mode refers to high accuracy of both MS/MS peaks and the precursor mass, while the HL mode refers to low accuracy of MS/MS peaks but high accuracy of the precursor mass,

| # | GNPS ID | PNP | structure | HH p-value | HH score | HL p-value | # peaks | PNP mass | # bonds |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 79098 | Cyclosporin-A;-Ramihyphin-A;-S-7481-F1 | cyc | 4.1e-70 | 41 | 2.7e-31 | 111 | 1201.8 | 11 |
| 2 | 78568 | hormothamnin-A | cyc | 2.1e-63 | 39 | 1.4e-26 | 124 | 1195.7 | 11 |
| 3 | 78606 | Plipastatin-A2;-SNA-60-367-6 | bcyc | 5.8e-59 | 29 | 1.4e-23 | 83 | 1476.8 | 11 |
| 4 | 78568 | dolastatin-12 | cyc | 4.8e-55 | 32 | 4.5e-15 | 103 | 968.6 | 9 |
| 5 | 79098 | Sch-217048 | cyc | 3.0e-53 | 30 | 3.0e-20 | 83 | 1136.7 | 10 |
| 6 | 78606 | Plipastatin-A1;-SNA-60-367-3 | bcyc | 3.0e-53 | 28 | 1.3e-19 | 103 | 1462.8 | 11 |
| 7 | 79098 | Cyclosporin-L | cyc | 1.0e-50 | 31 | 3.3e-19 | 98 | 1187.8 | 11 |
| 8 | 79098 | Cyclosporin-N-B;-Cyclosporin-B;-S-7481-F | cyc | 4.1e-50 | 28 | 3.5e-19 | 67 | 1187.8 | 11 |
| 9 | 79098 | Cyclosporin-E;-5-L-Valinecyclosporine-A | cyc | 5.9e-50 | 30 | 1.1e-22 | 90 | 1187.8 | 11 |
| 10 | 79098 | Cyclosporin-F | cyc | 1.8e-49 | 27 | 6.0e-22 | 60 | 1185.8 | 11 |
| 11 | 79098 | Cyclosporin-C;-7-L-Threoninecyclosporin | cyc | 3.8e-48 | 29 | 3.9e-24 | 90 | 1217.8 | 11 |
| 12 | 78635 | Orfamide-A | bcyc | 1.3e-46 | 26 | 5.6e-19 | 108 | 1294.8 | 11 |
| 13 | 78936 | Esperin | bcyc | 7.2e-45 | 22 | 3.7e-21 | 79 | 1035.7 | 8 |
| 14 | 78787 | SNA-60-367-23 | bcyc | 1.9e-44 | 24 | 2.1e-15 | 101 | 1488.8 | 11 |
| 15 | 78568 | desmethoxymajusculamide-C | cyc | 9.5e-44 | 26 | 1.9e-12 | 91 | 954.6 | 9 |
| 16 | 78937 | Pumilacidin-F | cyc | 1.3e-43 | 23 | 5.2e-19 | 72 | 1049.7 | 8 |
| 17 | 78787 | SNA-60-367-17 | bcyc | 2.0e-42 | 23 | 1.7e-14 | 98 | 1474.8 | 11 |
| 18 | 79098 | Cyclosporin-T | cyc | 3.6e-42 | 24 | 9.8e-17 | 62 | 1187.8 | 11 |
| 19 | 79098 | Cyclosporin-U | cyc | 3.7e-42 | 23 | 1.9e-16 | 52 | 1187.8 | 11 |
| 20 | 78937 | pumilacidin-D | cyc | 4.7e-42 | 22 | 9.0e-20 | 68 | 1063.7 | 8 |
| 21 | 78635 | Surfactin-D | cyc | 9.1e-42 | 22 | 3.2e-18 | 69 | 1049.7 | 8 |
| 22 | 79098 | Leucinostatin-B;-SF-1907II;-Paecilotoxin | linear | 1.2e-41 | 19 | 3.8e-18 | 93 | 1203.8 | 10 |
| 23 | 79098 | Leucinostatin-A;-CC-1014;-ICI-13959;-M-1 | linear | 1.5e-41 | 19 | 4.6e-18 | 95 | 1217.8 | 10 |
| 24 | 78937 | [Val7]-Surfactin-C15ai | cyc | 3.5e-41 | 22 | 3.6e-20 | 71 | 1021.7 | 8 |
| 25 | 79098 | Leucinostatin-D;-Paecilotoxin-D | linear | 6.0e-41 | 16 | 1.7e-23 | 37 | 1117.8 | 10 |
| 26 | 79098 | Cyclosporin-Y | cyc | 6.9e-41 | 23 | 2.4e-15 | 58 | 1201.8 | 11 |
| 27 | 79098 | SCH-378161 | cyc | 1.6e-40 | 26 | 3.4e-17 | 111 | 1122.6 | 10 |
| 28 | 79098 | [Leu(4)]-Cyclosporin-A | cyc | 2.4e-40 | 23 | 4.0e-17 | 60 | 1187.8 | 11 |
| 29 | 78568 | Ibu-epidemethoxylyngbystatin-3 | cyc | 1.0e-39 | 24 | 9.1e-11 | 92 | 982.6 | 9 |
| 30 | 78936 | pumilacidin-B | cyc | 1.4e-39 | 22 | 3.4e-16 | 83 | 1035.7 | 8 |
| 31 | 79098 | Heptaibin | linear | 2.4e-39 | 19 | 7.4e-19 | 105 | 1558.9 | 15 |
| 32 | 78937 | Bacirtsin-2;-Bacircine-3;-Surfactin-B;-S | cyc | 4.9e-39 | 21 | 1.3e-16 | 70 | 1021.7 | 8 |
| 33 | 79098 | Cyclosporin-X | cyc | 8.6e-39 | 24 | 1.0e-17 | 82 | 1201.8 | 11 |
| 34 | 78787 | SNA-60-367-19 | bcyc | 1.4e-38 | 19 | 6.5e-13 | 57 | 1460.8 | 11 |
| 35 | 79098 | Compound-1 | cyc | 6.3e-38 | 22 | 7.1e-17 | 69 | 1125.7 | 10 |
| 36 | 79098 | Emerimicin-IV;-Stilbellin-II;-Samarospor | linear | 8.0e-38 | 18 | 2.2e-18 | 95 | 1572.9 | 15 |
| 37 | 78606 | none | bcyc | 1.1e-37 | 21 | 1.6e-16 | 78 | 1035.7 | 9 |
| 38 | 78937 | Surfactin-C1;-Analysin;-Staphylococca;-S | cyc | 1.7e-37 | 21 | 3.4e-15 | 82 | 1035.7 | 8 |
| 39 | 78635 | bacillomycin-D2 | cyc | 3.4e-37 | 21 | 6.4e-15 | 85 | 1044.5 | 8 |
| 40 | 78936 | Ilamycin-B1;-Rufomycin-I;-Rufomycin-B1 | cyc | 4.7e-37 | 19 | 5.8e-18 | 68 | 1011.6 | 7 |
| 41 | 78787 | Lichenysin-G5a | cyc | 1.2e-36 | 19 | 1.4e-17 | 56 | 1020.7 | 8 |
| 42 | 78936 | [Val7]-Surfactin-C14i | cyc | 2.2e-36 | 19 | 4.9e-16 | 57 | 1007.6 | 8 |
| 43 | 79098 | Cyclosporin-I;-4-L-Leucine-7-L-valinecyc | cyc | 4.6e-36 | 21 | 5.6e-15 | 60 | 1201.8 | 11 |
| 44 | 78787 | Telomycin;-A-128-HYP;-A128-OP;-C-159;-Ne | bcyc | 5.3e-36 | 22 | 6.5e-14 | 106 | 1271.5 | 11 |
| 45 | 78568 | majusculamide-C | cyc | 5.8e-36 | 23 | 4.5e-12 | 107 | 984.6 | 9 |
| 46 | 79098 | Trichogin-A-IV;-GA-IV | linear | 7.4e-36 | 17 | 1.7e-17 | 83 | 1065.7 | 11 |
| 47 | 79098 | Leucinostatin-R | linear | 9.7e-36 | 15 | 2.0e-19 | 52 | 1201.8 | 10 |
| 48 | 78937 | halobacillin | cyc | 1.6e-35 | 20 | 2.6e-14 | 80 | 1034.7 | 8 |
| 49 | 78787 | Ayfivin;-Bacitracin-A | bcyc | 4.1e-35 | 18 | 3.6e-12 | 75 | 1421.8 | 11 |
| 50 | 78936 | [Val7]-Surfactin-C13ai | cyc | 2.8e-34 | 19 | 3.2e-15 | 70 | 993.6 | 8 |
| 51 | 79098 | Leucinostatin-K;-Paecilotoxin-K | linear | 4.1e-33 | 15 | 2.2e-18 | 76 | 1233.8 | 10 |
| 52 | 79098 | Acremostatin-B | linear | 5.8e-33 | 15 | 9.5e-20 | 80 | 1275.9 | 10 |
| 53 | 79098 | TA-IIIc | linear | 6.9e-33 | 16 | 5.3e-17 | 92 | 1923.1 | 19 |
| 54 | 78936 | WS-7338-B;-BE-18257-B | cyc | 8.0e-33 | 15 | 1.4e-16 | 43 | 612.3 | 5 |
| 55 | 79098 | Efrapeptin-F | linear | 8.5e-33 | 15 | 1.5e-16 | 67 | 1634.1 | 16 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 56 | 78635 | Bacillomycin-D2;-Raubitschek-substance | cyc | 9.3e-33 | 19 | 8.2e-15 | 85 | 1030.5 | 8 |
| 57 | 78937 | fijimycin-A | bcyc | 2.3e-32 | 18 | 1.1e-16 | 75 | 878.5 | 8 |
| 58 | 78635 | Bacillomycin-D4;-Raubitschek-substance | cyc | 2.6e-32 | 18 | 2.6e-14 | 72 | 1058.6 | 8 |
| 59 | 79098 | Acremostatin-C | linear | 5.6e-32 | 14 | 8.8e-17 | 64 | 1291.9 | 10 |
| 60 | 78803 | massetolide-F | bcyc | 6.2e-32 | 18 | 7.5e-13 | 91 | 1125.7 | 10 |
| 61 | 79098 | Isocyclosporin-D | cyc | 7.0e-32 | 20 | 5.9e-16 | 74 | 1215.9 | 11 |
| 62 | 78787 | Lichenysin-G4 | cyc | 1.2e-31 | 18 | 1.3e-12 | 75 | 1020.7 | 8 |
| 63 | 78787 | tumescenamide-A | bcyc | 1.9e-31 | 15 | 2.9e-15 | 53 | 699.4 | 6 |
| 64 | 79098 | FR-901459 | cyc | 2.0e-31 | 19 | 1.2e-15 | 63 | 1217.8 | 11 |
| 65 | 78635 | ”APD-I-component-a;-Surfactin-A1;-””standa” | cyc | 2.5e-31 | 18 | 2.7e-12 | 77 | 1007.6 | 8 |
| 66 | 79098 | Hypomurocin-B-3b | linear | 4.5e-31 | 16 | 3.7e-15 | 112 | 1733.0 | 18 |
| 67 | 78803 | massetolide-E | bcyc | 9.8e-31 | 17 | 8.6e-14 | 80 | 1111.7 | 10 |
| 68 | 79098 | Leucinostatin-A2 | linear | 1.3e-30 | 13 | 3.8e-18 | 49 | 1199.8 | 10 |
| 69 | 79098 | efrapeptin-H | linear | 1.5e-30 | 15 | 4.2e-17 | 93 | 1662.1 | 16 |
| 70 | 78787 | SNA-60-367-14 | bcyc | 1.6e-30 | 15 | 1.6e-10 | 46 | 1446.8 | 11 |
| 71 | 79098 | Leucinostatin-B2 | linear | 2.8e-30 | 13 | 8.4e-18 | 51 | 1185.8 | 10 |
| 72 | 78936 | Lichenysin-A | cyc | 3.4e-30 | 16 | 4.2e-13 | 52 | 1006.7 | 8 |
| 73 | 78936 | Ilamycin-B2;-Rufomycin-B2 | cyc | 3.5e-30 | 14 | 4.8e-15 | 36 | 1027.6 | 7 |
| 74 | 78787 | Lipopeptide-NO | cyc | 4.2e-30 | 16 | 5.1e-13 | 52 | 993.6 | 8 |
| 75 | 78787 | cotteslosin-A | cyc | 1.2e-29 | 14 | 1.5e-14 | 48 | 635.3 | 5 |
| 76 | 78787 | Lichenysin-G3 | cyc | 1.5e-29 | 15 | 1.8e-13 | 42 | 1006.7 | 8 |
| 77 | 79098 | g-Hydroxy-Meleu4-cyclosporin | cyc | 1.7e-29 | 18 | 1.2e-12 | 62 | 1217.8 | 11 |
| 78 | 78803 | massetolide-H | bcyc | 2.3e-29 | 17 | 2.0e-11 | 98 | 1153.7 | 10 |
| 79 | 79098 | Cycloaspeptide-A | cyc | 2.5e-29 | 13 | 3.2e-15 | 35 | 641.3 | 5 |
| 80 | 79098 | Trikoningin-KB-II | linear | 4.4e-29 | 14 | 2.3e-15 | 75 | 1051.7 | 11 |
| 81 | 79098 | Trichoaureocin-4 | linear | 5.8e-29 | 15 | 4.0e-14 | 111 | 1965.1 | 20 |
| 82 | 78803 | Orfamide-C | bcyc | 6.1e-29 | 16 | 4.6e-13 | 68 | 1266.8 | 11 |
| 83 | 78937 | Grividomycin-I | bcyc | 1.2e-28 | 14 | 5.6e-15 | 37 | 864.4 | 8 |
| 84 | 78937 | pumilacidin-A | cyc | 1.4e-28 | 16 | 1.3e-11 | 66 | 1049.7 | 8 |
| 85 | 79098 | Trichoaureocin-6 | linear | 1.4e-28 | 14 | 3.3e-12 | 82 | 1937.1 | 20 |
| 86 | 78635 | massetolide-A | bcyc | 1.6e-28 | 16 | 1.4e-11 | 82 | 1139.7 | 10 |
| 87 | 79098 | aspergillicin-B | bcyc | 1.8e-28 | 15 | 6.6e-15 | 59 | 726.4 | 7 |
| 88 | 78787 | Plipastatin-B2;-SNA-60-367-12 | bcyc | 2.1e-28 | 16 | 1.7e-07 | 85 | 1504.8 | 11 |
| 89 | 79098 | Trichopolyn-I | linear | 2.7e-28 | 13 | 3.4e-17 | 71 | 1205.8 | 10 |
| 90 | 78936 | pumilacidin-C | cyc | 4.0e-28 | 14 | 4.4e-13 | 39 | 1077.7 | 8 |
| 91 | 78936 | pumilacidin-E | cyc | 5.4e-28 | 16 | 4.5e-11 | 72 | 1063.7 | 8 |
| 92 | 79098 | Hypomurocin-B-1 | linear | 8.4e-28 | 14 | 2.4e-15 | 91 | 1719.0 | 18 |
| 93 | 78787 | Lichenysin-G9a | cyc | 1.4e-27 | 14 | 1.4e-12 | 41 | 1048.7 | 8 |
| 94 | 79098 | aspergillicin-A | bcyc | 2.4e-27 | 15 | 2.1e-11 | 70 | 740.4 | 7 |
| 95 | 78937 | Lichenysin-G1a | cyc | 2.9e-27 | 15 | 2.6e-11 | 56 | 992.7 | 8 |
| 96 | 79098 | Acremostatin-A | linear | 3.1e-27 | 13 | 1.9e-17 | 88 | 1261.9 | 10 |
| 97 | 78635 | Plipastatin-B1;-SNA-60-367-7 | bcyc | 3.4e-27 | 15 | 3.5e-10 | 74 | 1490.8 | 11 |
| 98 | 79098 | Trikoningin-KA-V;-KA-V | linear | 4.9e-27 | 13 | 2.2e-14 | 73 | 1889.1 | 19 |
| 99 | 78803 | massetolide-G | bcyc | 5.1e-27 | 15 | 2.5e-13 | 75 | 1139.7 | 10 |
| 100 | 79098 | Trichorzin-HA-3 | linear | 6.3e-27 | 14 | 2.2e-14 | 104 | 1717.0 | 18 |
| 101 | 78937 | Peptidolipin-NA | cyc | 2.1e-26 | 14 | 1.1e-12 | 45 | 963.7 | 8 |
| 102 | 79098 | [MeVal]5-cyclosporin | cyc | 3.2e-26 | 15 | 2.3e-10 | 43 | 1215.9 | 11 |
| 103 | 78937 | [Val7]-Surfactin-C14i-dimethyl-ester | cyc | 4.2e-26 | 15 | 2.6e-11 | 64 | 1035.7 | 10 |
| 104 | 79098 | cotteslosin-B | cyc | 7.3e-26 | 13 | 6.6e-12 | 60 | 649.3 | 5 |
| 105 | 78787 | Aurantin-II | dcyc | 9.1e-26 | 15 | 5.9e-10 | 106 | 1254.6 | 12 |
| 106 | 79098 | Longibrachin-BIII | linear | 2.1e-25 | 13 | 3.9e-09 | 94 | 1951.1 | 20 |
| 107 | 79098 | Leucinostatin-V | linear | 3.7e-25 | 11 | 2.0e-16 | 49 | 1187.8 | 10 |
| 108 | 78568 | dolastatin-14 | cyc | 3.8e-25 | 14 | 2.1e-11 | 60 | 1088.7 | 8 |
| 109 | 78787 | Bacitracin-B3 | bcyc | 6.0e-25 | 11 | 5.1e-13 | 27 | 1407.7 | 11 |
| 110 | 79098 | Leucinostatin-F | linear | 7.9e-25 | 10 | 9.6e-16 | 29 | 1103.8 | 10 |
| 111 | 78787 | [Leu7]-Surfactin-C14i-monomethyl-ester | bcyc | 9.5e-25 | 14 | 6.2e-10 | 59 | 1035.7 | 9 |
| 112 | 79098 | Trikoningin-KB-I | linear | 9.6e-25 | 12 | 8.0e-12 | 68 | 1037.7 | 11 |
| 113 | 79098 | scopularide-A | cyc | 1.1e-24 | 13 | 8.5e-11 | 52 | 671.4 | 6 |
| 114 | 78937 | [Leu7]-Surfactin-C13ai-dimethyl-ester | cyclic | 1.5e-24 | 14 | 9.5e-10 | 59 | 1035.7 | 10 |
| 115 | 79098 | Trichovirin-II-1a | linear | 1.9e-24 | 13 | 1.4e-10 | 107 | 1703.0 | 18 |
| 116 | 78568 | pitiprolamide | cyc | 2.2e-24 | 16 | 1.5e-09 | 100 | 904.5 | 8 |
| 117 | 78937 | Glumamycin;-Amphomycin;-A-6786-52;-A-899 | bcyc | 4.7e-24 | 16 | 2.0e-08 | 90 | 1289.7 | 12 |
| 118 | 78568 | OHB;-cPHB;-Oligo-(hydroxybutyric-acid), | linear | 5.6e-24 | 11 | 4.4e-12 | 52 | 792.3 | 8 |
| 119 | 78817 | massetolide-C | bcyc | 6.0e-24 | 14 | 3.3e-09 | 88 | 1167.7 | 10 |
| 120 | 78787 | Bacitracin-B1 | bcyc | 8.9e-24 | 11 | 6.8e-12 | 33 | 1407.7 | 11 |
| 121 | 79098 | Trichocellin-A-VI | linear | 1.2e-23 | 11 | 9.2e-12 | 55 | 1964.2 | 20 |
| 122 | 78568 | carriebowmide | cyc | 2.1e-23 | 13 | 1.4e-08 | 60 | 880.5 | 7 |
| 123 | 78937 | mojavenisn-A | cyc | 2.8e-23 | 14 | 1.5e-11 | 79 | 1083.6 | 8 |
| 124 | 78568 | lyngbyastatin-3 | cyc | 3.5e-23 | 15 | 1.1e-07 | 80 | 1012.6 | 9 |
| 125 | 79098 | aspergillicin-E | bcyc | 4.9e-23 | 12 | 1.1e-12 | 49 | 754.4 | 7 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 126 | 79098 | N-methylsansalvamide | cyc | 5.1e-23 | 12 | 2.2e-11 | 61 | 600.4 | 5 |
| 127 | 78568 | PHB;-sPHB;-Poly-(hydroxybutyric-acid),-m | linear | 6.6e-23 | 11 | 4.7e-11 | 63 | 878.4 | 9 |
| 128 | 78635 | Orfamide-B | bcyc | 7.4e-23 | 12 | 4.8e-10 | 45 | 1280.8 | 11 |
| 129 | 79098 | Destruxin-Ed(1) | cyc | 9.0e-23 | 12 | 3.7e-11 | 46 | 625.4 | 6 |
| 130 | 79098 | [8'-Hydroxy-MeBmf]1-cyclosporin | cyc | 9.8e-23 | 14 | 3.4e-13 | 52 | 1217.8 | 11 |
| 131 | 79098 | Cyclosporin-V | cyc | 1.0e-22 | 14 | 2.6e-14 | 52 | 1215.9 | 11 |
| 132 | 79098 | aspergillicin-C | bcyc | 1.7e-22 | 12 | 7.2e-11 | 51 | 710.4 | 7 |
| 133 | 78936 | WS-7338-A;-BE-18257-A | cyc | 1.8e-22 | 11 | 1.2e-10 | 44 | 598.3 | 5 |
| 134 | 79098 | Cordyheptapeptide-A | cyc | 2.3e-22 | 11 | 6.7e-12 | 33 | 879.5 | 7 |
| 135 | 79098 | Isariin-G1 | cyc | 2.9e-22 | 12 | 7.7e-12 | 49 | 609.4 | 6 |
| 136 | 79098 | JBIR-113 | bcyc | 3.0e-22 | 11 | 2.0e-10 | 42 | 595.3 | 6 |
| 137 | 79098 | Trichorzin-MA-1;-Trichoharzin-MA-1 | linear | 3.5e-22 | 12 | 2.0e-09 | 110 | 1733.0 | 18 |
| 138 | 79098 | Trichosporin-B-IIIc | linear | 4.9e-22 | 11 | 3.8e-14 | 74 | 1936.1 | 20 |
| 139 | 79098 | Trichokonin-VI;-Gliodeliquescin;-Gliodel | linear | 6.7e-22 | 11 | 1.2e-12 | 76 | 1936.1 | 20 |
| 140 | 79098 | [§-Me-Pro]-destruxin-E-chlorohydrin | cyc | 7.1e-22 | 12 | 4.0e-09 | 55 | 643.3 | 6 |
| 141 | 78606 | Viscosinamide | bcyc | 1.5e-21 | 12 | 6.5e-10 | 63 | 1124.7 | 10 |
| 142 | 78787 | Actinomycin-X2;-Actinomycin-V;-Actinomyc | dcyc | 1.7e-21 | 12 | 7.6e-10 | 77 | 1268.6 | 12 |
| 143 | 79098 | [Phe3,N-MeVal5]-Destruxin-B | cyc | 1.8e-21 | 12 | 9.3e-09 | 60 | 655.4 | 6 |
| 144 | 79098 | Trichokindin-IVa | linear | 2.1e-21 | 11 | 7.7e-11 | 84 | 1747.1 | 18 |
| 145 | 78568 | palmyramide-A | cyc | 4.7e-21 | 12 | 2.2e-08 | 66 | 671.4 | 6 |
| 146 | 79098 | Trichovirin-II-5 | linear | 5.1e-21 | 11 | 3.8e-08 | 90 | 1731.1 | 18 |
| 147 | 78937 | Actinomycin-K1c | dcyc | 9.4e-21 | 12 | 4.1e-08 | 88 | 1268.6 | 12 |
| 148 | 78937 | Grividomycin-III;-5-(4-Methylisoleucine) | bcyc | 1.2e-20 | 10 | 9.4e-13 | 27 | 850.4 | 8 |
| 149 | 79098 | Trichorozin-IV | linear | 1.4e-20 | 10 | 3.8e-11 | 68 | 1188.8 | 11 |
| 150 | 78937 | tumescenamide-B | bcyc | 1.4e-20 | 9 | 8.2e-11 | 26 | 741.5 | 6 |
| 151 | 79098 | Trichorozin-II | linear | 1.5e-20 | 10 | 8.8e-10 | 68 | 1174.8 | 11 |
| 152 | 79098 | Enniatin-J2 | cyc | 1.6e-20 | 11 | 5.8e-10 | 46 | 625.4 | 6 |
| 153 | 78568 | pitipeptolide-E | cyc | 2.1e-20 | 13 | 8.4e-08 | 89 | 793.5 | 7 |
| 154 | 79098 | Trichosporin-B-III-a | linear | 3.8e-20 | 10 | 1.1e-10 | 68 | 1950.1 | 20 |
| 155 | 79098 | Destruxin-E1 | cyc | 4.8e-20 | 11 | 1.3e-10 | 49 | 607.4 | 6 |
| 156 | 79098 | Cyclosporin-P;-6-((2S,3R,4R,6E)-2-Amino- | cyc | 5.0e-20 | 14 | 1.2e-10 | 77 | 1203.8 | 11 |
| 157 | 79098 | Sporidesmolide-I | cyc | 5.8e-20 | 11 | 2.2e-09 | 52 | 638.4 | 6 |
| 158 | 79098 | Destruxin-B1;-NSC-236580;-SB-242536 | cyc | 6.2e-20 | 11 | 1.7e-10 | 50 | 607.4 | 6 |
| 159 | 79098 | Trichorozin-I | linear | 6.5e-20 | 10 | 3.5e-09 | 77 | 1160.8 | 11 |
| 160 | 79098 | CNC457.1424 | linear | 8.2e-20 | 10 | 6.5e-13 | 76 | 1423.9 | 14 |
| 161 | 79098 | Destruxin-A1 | cyc | 8.3e-20 | 11 | 3.1e-09 | 50 | 591.4 | 6 |
| 162 | 78568 | veraguamide-I | cyc | 8.9e-20 | 11 | 4.0e-08 | 58 | 692.5 | 6 |
| 163 | 79098 | Isariin;-Isariin-A | cyc | 1.2e-19 | 11 | 4.4e-09 | 55 | 637.4 | 6 |
| 164 | 78936 | WS-9320-A | bcyc | 1.3e-19 | 11 | 5.0e-09 | 61 | 1036.5 | 8 |
| 165 | 79098 | [Ala(2),Val(11)]Cyclosporin | cyc | 1.4e-19 | 12 | 5.3e-08 | 44 | 1173.8 | 11 |
| 166 | 78937 | valinomycin | cyc | 1.5e-19 | 15 | 5.4e-06 | 84 | 1110.6 | 12 |
| 167 | 78937 | Etamycin-VI-2 | bcyc | 1.7e-19 | 9 | 9.0e-10 | 21 | 864.4 | 8 |
| 168 | 79098 | Aselacin-C | bcyc | 2.5e-19 | 10 | 1.2e-08 | 64 | 906.5 | 7 |
| 169 | 79098 | Sporidesmolide-II | cyc | 2.5e-19 | 11 | 6.4e-11 | 60 | 652.4 | 6 |
| 170 | 79098 | Enniatin-B4 | cyc | 3.1e-19 | 11 | 1.2e-07 | 61 | 653.4 | 6 |
| 171 | 78568 | dolastatin-10 | linear | 6.7e-19 | 8 | 2.3e-05 | 53 | 784.5 | 5 |
| 172 | 79098 | Integramide-A | linear | 6.7e-19 | 10 | 3.0e-08 | 93 | 1631.0 | 16 |
| 173 | 79098 | laxaphycin-B2 | cyc | 7.8e-19 | 13 | 6.0e-08 | 67 | 1378.8 | 12 |
| 174 | 79098 | Trichosporin-Bs-f;-Trichosporin-B-I-a | linear | 8.0e-19 | 9 | 1.3e-06 | 54 | 1966.1 | 20 |
| 175 | 79098 | Isariin-D | cyc | 9.7e-19 | 10 | 2.9e-09 | 37 | 539.3 | 6 |
| 176 | 79098 | Suzukacillin-B | linear | 1.1e-18 | 10 | 3.7e-09 | 94 | 1965.1 | 20 |
| 177 | 78568 | symplostatin-1 | linear | 1.2e-18 | 9 | 7.3e-05 | 110 | 798.5 | 5 |
| 178 | 79098 | Trichotoxin-A; | linear | 1.5e-18 | 10 | 5.1e-09 | 95 | 1731.1 | 18 |
| 179 | 79098 | Trichorzin-HA-5 | linear | 1.5e-18 | 10 | 3.7e-10 | 95 | 1731.1 | 18 |
| 180 | 79098 | Isariin-E | cyc | 1.8e-18 | 10 | 5.7e-09 | 40 | 553.3 | 6 |
| 181 | 78937 | Actinomycin-Au6a;-Actinoleucin-A;-Actino | dcyc | 2.0e-18 | 11 | 6.5e-07 | 94 | 1282.7 | 12 |
| 182 | 79098 | Beauvericin-D | cyc | 2.3e-18 | 10 | 9.6e-08 | 55 | 769.4 | 6 |
| 183 | 79098 | psuedodestruxin-C | cyc | 2.6e-18 | 10 | 8.5e-09 | 49 | 669.4 | 6 |
| 184 | 78937 | Kurstakin-4 | bcyc | 2.8e-18 | 9 | 1.3e-08 | 57 | 905.5 | 8 |
| 185 | 78937 | A-1437-M | bcyc | 4.6e-18 | 12 | 8.8e-06 | 71 | 1275.6 | 12 |
| 186 | 78787 | SNA-60-367-4 | bcyc | 4.9e-18 | 11 | 6.5e-05 | 84 | 1476.8 | 11 |
| 187 | 79098 | Enniatin-B2 | cyc | 5.1e-18 | 10 | 1.7e-08 | 49 | 625.4 | 6 |
| 188 | 79098 | Isariin-C2 | cyc | 5.3e-18 | 10 | 1.7e-08 | 45 | 567.4 | 6 |
| 189 | 79098 | Trichokindin-IIa | linear | 1.1e-17 | 9 | 4.6e-08 | 70 | 1747.1 | 18 |
| 190 | 79098 | Enniatin-A1 | cyc | 1.2e-17 | 10 | 3.9e-08 | 56 | 667.4 | 6 |
| 191 | 78568 | veraguamide-G | cyc | 1.2e-17 | 10 | 4.3e-07 | 58 | 690.5 | 6 |
| 192 | 79098 | sansalvamide | cyc | 1.7e-17 | 9 | 7.1e-08 | 46 | 586.4 | 5 |
| 193 | 79098 | Cyclosporin-R | cyc | 2.2e-17 | 10 | 6.7e-08 | 31 | 1173.8 | 11 |
| 194 | 78568 | veraguamide-E | cyc | 2.4e-17 | 10 | 8.0e-07 | 64 | 716.5 | 6 |
| 195 | 78568 | cocosamide-B | cyc | 2.5e-17 | 9 | 1.0e-07 | 41 | 741.4 | 6 |

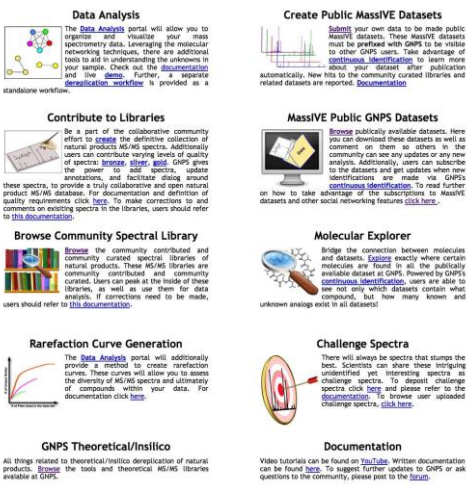
| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 196 | 78568 | veraguamide-C | cyc | 3.8e-17 | 9 | 1.0e-08 | 40 | 688.4 | 6 |
| 197 | 78568 | almiramide-C | linear | 5.1e-17 | 7 | 4.5e-11 | 32 | 726.5 | 5 |
| 198 | 79098 | Trichokindin-VI;-(Trichokindin-IX) | linear | 5.6e-17 | 7 | 7.8e-13 | 23 | 1775.1 | 18 |
| 199 | 78568 | apramide-A | linear | 9.5e-17 | 8 | 5.9e-09 | 69 | 976.6 | 8 |
| 200 | 79098 | Integramide-B | linear | 9.8e-17 | 8 | 7.3e-07 | 52 | 1645.0 | 16 |
| 201 | 79098 | Beauvericin-A | cyc | 1.6e-16 | 9 | 5.8e-07 | 53 | 797.4 | 6 |
| 202 | 79098 | Enniatin-I | cyc | 1.6e-16 | 8 | 6.5e-08 | 26 | 667.4 | 6 |
| 203 | 78568 | veraguamide-A | cyc | 1.9e-16 | 9 | 5.7e-08 | 52 | 766.4 | 6 |
| 204 | 78787 | [Ile2,4,7]Surfactin | cyc | 2.3e-16 | 9 | 7.9e-07 | 39 | 1021.7 | 8 |
| 205 | 79098 | Beauvericin-B | cyc | 3.4e-16 | 9 | 1.1e-06 | 58 | 811.4 | 6 |
| 206 | 78937 | SNA-60-367-21 | bcyc | 3.6e-16 | 8 | 2.3e-06 | 30 | 1474.8 | 11 |
| 207 | 78936 | Bamylocin-A | cyc | 3.8e-16 | 9 | 5.8e-10 | 41 | 1021.6 | 8 |
| 208 | 79098 | emericellamide-B | cyc | 4.7e-16 | 9 | 1.1e-08 | 49 | 651.5 | 6 |
| 209 | 79098 | Hypomurocin-B-2 | linear | 6.3e-16 | 9 | 1.8e-08 | 107 | 1733.0 | 18 |
| 210 | 79098 | Leucinostatin-N | linear | 6.6e-16 | 6 | 2.3e-09 | 15 | 1119.8 | 10 |
| 211 | 79098 | Trichokindin-IIb;-(Trichokindin-Vb) | linear | 9.0e-16 | 9 | 9.2e-13 | 113 | 1761.1 | 18 |
| 212 | 78937 | Stendomycin;-Stendomycin-A;-Stendomycin- | bcyc | 9.1e-16 | 7 | 3.4e-08 | 20 | 1628.0 | 15 |
| 213 | 79098 | Verrucamide-D | cyc | 9.1e-16 | 11 | 2.2e-03 | 46 | 1386.8 | 14 |
| 214 | 79098 | Hypomurocin-A-2 | linear | 1.0e-15 | 8 | 1.4e-09 | 71 | 1174.8 | 11 |
| 215 | 78937 | WS-7338-D | cyc | 1.0e-15 | 7 | 3.8e-08 | 21 | 584.3 | 5 |
| 216 | 79098 | LP-237-F7 | linear | 1.1e-15 | 8 | 2.7e-08 | 80 | 1310.8 | 11 |
| 217 | 79098 | Enniatin-J1 | cyc | 1.2e-15 | 9 | 3.3e-08 | 51 | 611.4 | 6 |
| 218 | 79098 | JBIR-114 | bcyc | 1.3e-15 | 8 | 2.9e-08 | 38 | 581.3 | 6 |
| 219 | 78568 | almiramide-B | linear | 1.3e-15 | 7 | 4.9e-08 | 49 | 724.5 | 5 |
| 220 | 78568 | dolastatin-D | cyc | 1.4e-15 | 9 | 3.9e-06 | 71 | 573.3 | 5 |
| 221 | 79098 | Verrucamide-C | cyc | 1.4e-15 | 11 | 3.2e-05 | 48 | 1400.8 | 14 |
| 222 | 79098 | arenamide-A | cyc | 1.4e-15 | 8 | 6.7e-06 | 33 | 671.4 | 6 |
| 223 | 79098 | Sporidesmolide-V | cyc | 1.4e-15 | 9 | 4.2e-07 | 56 | 666.5 | 6 |
| 224 | 79098 | Trichoaureocin-3 | linear | 1.4e-15 | 8 | 5.0e-07 | 68 | 1951.1 | 20 |
| 225 | 79098 | Trichorzin-MA-2;-Trichoharzin-MA-2 | linear | 1.4e-15 | 9 | 1.9e-12 | 118 | 1747.1 | 18 |
| 226 | 78936 | Ilamycin;-Ilamycin-A;-Rufomycin-A;-Rufom | cyc | 1.7e-15 | 8 | 2.1e-09 | 37 | 1041.5 | 7 |
| 227 | 79098 | Destruxin-D1 | cyc | 1.7e-15 | 9 | 5.2e-07 | 55 | 637.4 | 6 |
| 228 | 78936 | Radamicin;-Radamycin | bcyc | 2.3e-15 | 8 | 6.8e-04 | 49 | 1105.3 | 13 |
| 229 | 79098 | Leucinostatin-S | linear | 2.3e-15 | 6 | 7.9e-09 | 19 | 1203.9 | 10 |
| 230 | 78568 | almiramide-A | linear | 2.6e-15 | 7 | 9.4e-08 | 55 | 742.5 | 5 |
| 231 | 79098 | Trichorovin-IIIb | linear | 2.6e-15 | 7 | 9.5e-08 | 40 | 1160.8 | 11 |
| 232 | 79098 | RHM-1 | linear | 3.0e-15 | 6 | 1.0e-08 | 21 | 1035.7 | 8 |
| 233 | 79098 | Pleofungin-A;-F-15078-A | bcyc | 3.2e-15 | 10 | 1.4e-06 | 65 | 1094.7 | 10 |
| 234 | 78568 | nordolastatin-G | cyc | 4.1e-15 | 9 | 1.3e-07 | 54 | 1058.7 | 8 |
| 235 | 79098 | Fungisporin | cyc | 6.0e-15 | 7 | 4.6e-06 | 17 | 984.5 | 8 |
| 236 | 79098 | efrapeptin-J | linear | 6.5e-15 | 6 | 7.2e-14 | 19 | 1620.1 | 16 |
| 237 | 79098 | Leucinostatin-H;-Paecilotoxin-H | linear | 6.6e-15 | 6 | 4.2e-10 | 21 | 1133.8 | 10 |
| 238 | 79098 | Stilbellin-I;-Samarosporin-I;-Emerimicin | linear | 8.0e-15 | 6 | 5.1e-10 | 20 | 1558.9 | 15 |
| 239 | 79098 | RHM-2 | linear | 1.1e-14 | 6 | 3.5e-08 | 25 | 1021.7 | 8 |
| 240 | 78568 | veraguamide-K | linear | 1.1e-14 | 7 | 3.6e-07 | 61 | 812.4 | 6 |
| 241 | 79098 | Beauvericin-G2 | cyc | 1.7e-14 | 8 | 4.9e-06 | 49 | 755.4 | 6 |
| 242 | 79098 | Trichosporin-Bs-e | linear | 1.9e-14 | 7 | 6.2e-07 | 48 | 1936.1 | 20 |
| 243 | 79098 | Peptaivirin-A | linear | 2.4e-14 | 8 | 6.6e-06 | 98 | 1909.1 | 19 |
| 244 | 79098 | Enniatin-M1 | cyc | 2.4e-14 | 7 | 7.8e-07 | 25 | 683.4 | 6 |
| 245 | 78635 | Laxaphycin-D | cyc | 2.8e-14 | 9 | 7.0e-11 | 38 | 1366.8 | 12 |
| 246 | 79098 | Trichovirin-II-2c | linear | 2.8e-14 | 8 | 5.6e-09 | 95 | 1717.0 | 18 |
| 247 | 78568 | pitipeptolide-B | cyc | 3.6e-14 | 9 | 1.4e-06 | 69 | 809.5 | 7 |
| 248 | 78568 | dragonamide | linear | 4.4e-14 | 6 | 4.1e-09 | 39 | 653.5 | 4 |
| 249 | 79098 | Isariin-II;-Factor-II;-Isariin-B | cyc | 4.5e-14 | 8 | 1.0e-07 | 44 | 595.4 | 6 |
| 250 | 79098 | emericellamide-A | cyc | 5.3e-14 | 7 | 9.2e-08 | 25 | 609.4 | 6 |
| 251 | 79098 | arenamide-C | cyc | 5.9e-14 | 7 | 2.3e-02 | 27 | 655.4 | 6 |
| 252 | 79098 | Pimaydolide | cyc | 6.0e-14 | 7 | 1.1e-07 | 27 | 653.4 | 6 |
| 253 | 78568 | hantupeptin-C | cyc | 6.3e-14 | 8 | 1.7e-06 | 56 | 740.5 | 6 |
| 254 | 78635 | versicoloritide-A | cyc | 6.3e-14 | 7 | 2.0e-06 | 34 | 559.3 | 5 |
| 255 | 78937 | Actinomycin-K2c | dcyc | 6.5e-14 | 9 | 1.9e-05 | 109 | 1282.7 | 12 |
| 256 | 79098 | Boletusin | linear | 8.7e-14 | 8 | 1.6e-10 | 114 | 1895.1 | 19 |
| 257 | 78568 | veraguamide-J | cyc | 1.1e-13 | 7 | 3.1e-06 | 30 | 678.5 | 6 |
| 258 | 78568 | veraguamide-D | cyc | 1.1e-13 | 8 | 3.1e-07 | 57 | 702.5 | 6 |
| 259 | 78936 | Montanastatin;-cyclo-(D-Val-L-Lac-L-Val- | cyc | 1.1e-13 | 7 | 8.8e-09 | 19 | 740.4 | 8 |
| 260 | 79098 | Pseudodestruxin-B | cyc | 1.5e-13 | 7 | 2.9e-07 | 31 | 669.4 | 6 |
| 261 | 78787 | Oxachelin | linear | 1.5e-13 | 6 | 1.7e-06 | 46 | 635.3 | 5 |
| 262 | 79098 | Trichovirin-I-7A | linear | 1.7e-13 | 6 | 5.2e-07 | 31 | 1427.9 | 14 |
| 263 | 79098 | W-493-A | cyc | 1.8e-13 | 8 | 4.0e-05 | 54 | 873.5 | 7 |
| 264 | 78568 | veraguamide-H | cyc | 1.8e-13 | 7 | 5.2e-06 | 32 | 674.4 | 6 |
| 265 | 78568 | Tyrocidine-A;-Graminic-acid;-(Tyrothrici | cyc | 1.9e-13 | 8 | 4.9e-06 | 38 | 1269.7 | 10 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 266 | 78568 | carmabin-B | linear | 1.9e-13 | 6 | 5.9e-07 | 54 | 721.4 | 4 |
| 267 | 78568 | coibamide-A | bcyc | 1.9e-13 | 9 | 5.3e-05 | 109 | 1286.8 | 11 |
| 268 | 79098 | Trichorovin-Ia | linear | 2.7e-13 | 6 | 8.1e-07 | 34 | 1146.7 | 11 |
| 269 | 78839 | A-21978-C3 | bcyc | 2.8e-13 | 11 | 8.7e-04 | 160 | 1661.8 | 14 |
| 270 | 79098 | Enniatin-L | cyc | 3.0e-13 | 7 | 6.3e-07 | 34 | 669.4 | 6 |
| 271 | 79098 | RHM3 | linear | 3.1e-13 | 5 | 8.7e-08 | 15 | 1007.7 | 8 |
| 272 | 78568 | wewakazole | cyc | 3.5e-13 | 9 | 2.4e-06 | 73 | 1140.5 | 12 |
| 273 | 78568 | carmabin-A | linear | 3.9e-13 | 6 | 1.2e-06 | 59 | 703.4 | 4 |
| 274 | 78936 | cyclomarin-A | cyc | 4.6e-13 | 6 | 1.4e-06 | 19 | 1042.6 | 7 |
| 275 | 78568 | apramide-B | linear | 4.7e-13 | 6 | 1.5e-07 | 48 | 962.6 | 8 |
| 276 | 78568 | hantupeptin-A | cyc | 5.4e-13 | 8 | 1.0e-04 | 72 | 736.4 | 6 |
| 277 | 78568 | majusculamide-D | linear | 5.5e-13 | 7 | 1.4e-05 | 105 | 811.5 | 7 |
| 278 | 79098 | Verrucamide-B | cyc | 5.5e-13 | 11 | 1.9e-06 | 80 | 1400.8 | 14 |
| 279 | 78568 | somamide-B | ybcyc | 6.1e-13 | 8 | 1.2e-04 | 112 | 918.4 | 8 |
| 280 | 79098 | Efrapeptin-C | linear | 8.2e-13 | 6 | 1.5e-10 | 39 | 1606.0 | 16 |
| 281 | 78568 | dragonamide-E | linear | 8.4e-13 | 6 | 2.4e-06 | 62 | 651.4 | 4 |
| 282 | 78568 | kulomo'opunalide-2 | cyc | 8.8e-13 | 8 | 3.3e-06 | 73 | 702.5 | 6 |
| 283 | 78568 | apramide-C | linear | 9.9e-13 | 6 | 1.7e-08 | 55 | 978.6 | 8 |
| 284 | 78606 | Pseudodesmin-B | bcyc | 1.0e-12 | 6 | 2.9e-06 | 20 | 1110.7 | 10 |
| 285 | 78568 | symplostatin-3 | linear | 1.2e-12 | 6 | 7.3e-04 | 75 | 746.5 | 4 |
| 286 | 78568 | cocosamide-A | cyc | 1.2e-12 | 7 | 2.8e-06 | 45 | 743.4 | 6 |
| 287 | 78568 | lyngbyastatin-1 | cyc | 1.3e-12 | 8 | 8.4e-02 | 47 | 998.6 | 9 |
| 288 | 79098 | Destruxin-D | cyc | 1.4e-12 | 7 | 3.3e-06 | 39 | 623.4 | 6 |
| 289 | 79098 | Cyclosporin-Z | cyc | 1.7e-12 | 7 | 2.6e-07 | 22 | 1173.8 | 11 |
| 290 | 78568 | YM-47142 | bcyc | 1.7e-12 | 6 | 9.8e-03 | 31 | 900.5 | 8 |
| 291 | 79098 | Trichorovin-IXa | linear | 1.8e-12 | 6 | 5.1e-06 | 47 | 1174.8 | 11 |
| 292 | 79098 | Destruxin-E-diol;-SB-285968 | cyc | 2.8e-12 | 7 | 6.3e-05 | 42 | 611.4 | 6 |
| 293 | 79098 | Harzianin-HC-13 | linear | 3.0e-12 | 6 | 7.9e-06 | 49 | 1443.9 | 14 |
| 294 | 79098 | Trichovirin-I-3C | linear | 3.4e-12 | 5 | 9.1e-07 | 19 | 1429.9 | 14 |
| 295 | 78635 | Tolaasin-I | bcyc | 3.6e-12 | 5 | 3.4e-05 | 16 | 1986.2 | 19 |
| 296 | 78787 | Mer-N5075-A | linear | 5.2e-12 | 5 | 1.4e-06 | 29 | 597.3 | 4 |
| 297 | 79098 | Beauvericin-J | cyc | 5.3e-12 | 6 | 1.4e-05 | 29 | 799.4 | 6 |
| 298 | 78568 | dragomabin | linear | 5.8e-12 | 6 | 1.5e-05 | 86 | 661.4 | 4 |
| 299 | 78635 | lyngbyazothrin-B | cyc | 6.1e-12 | 7 | 9.3e-04 | 28 | 1282.7 | 11 |
| 300 | 79098 | MK-1688 | cyc | 7.5e-12 | 7 | 1.5e-04 | 53 | 681.5 | 6 |
| 301 | 78568 | AO-341B;-LL-A-0341-B | bcyc | 7.6e-12 | 7 | 3.0e-02 | 37 | 1140.5 | 10 |
| 302 | 78568 | hantupeptin-B | cyc | 8.3e-12 | 7 | 2.1e-05 | 58 | 738.5 | 6 |
| 303 | 79098 | Enniatin-B3 | cyc | 9.8e-12 | 6 | 2.4e-05 | 25 | 611.4 | 6 |
| 304 | 79098 | Sporidesmolide-III;-N-Desmethyl-sporides | cyc | 1.2e-11 | 7 | 3.0e-05 | 52 | 624.4 | 6 |
| 305 | 79098 | Sch-486058 | bcyc | 1.7e-11 | 7 | 4.4e-02 | 31 | 1530.8 | 14 |
| 306 | 79098 | Bassianolide | cyc | 1.9e-11 | 7 | 3.4e-04 | 44 | 908.6 | 8 |
| 307 | 79098 | Sch-218157 | cyc | 2.0e-11 | 7 | 4.8e-05 | 35 | 1135.7 | 10 |
| 308 | 79098 | WIN-66306 | cyc | 2.4e-11 | 6 | 4.4e-06 | 27 | 800.4 | 7 |
| 309 | 79098 | apratoxin-G | cyc | 2.5e-11 | 6 | 5.8e-05 | 55 | 813.5 | 5 |
| 310 | 78936 | Amonabactin-P-750 | y-branch | 2.7e-11 | 4 | 6.2e-07 | 11 | 750.3 | 5 |
| 311 | 78803 | Clavariopsin-A | cyc | 2.8e-11 | 9 | 1.8e-05 | 96 | 1153.7 | 10 |
| 312 | 78568 | pahayokolide-B | cyc | 3.1e-11 | 7 | 5.5e-02 | 35 | 1302.7 | 11 |
| 313 | 78568 | malevamide-B | cyc | 3.5e-11 | 8 | 2.4e-01 | 47 | 1428.9 | 13 |
| 314 | 78937 | Neotelomycin;-Neotelomycin-P;-A-128P | bcyc | 3.9e-11 | 6 | 5.5e-07 | 31 | 1271.5 | 11 |
| 315 | 78568 | dragonamide-B | linear | 3.9e-11 | 5 | 9.7e-06 | 43 | 605.5 | 4 |
| 316 | 79098 | SPF-5506-A4 | linear | 3.9e-11 | 5 | 9.8e-06 | 30 | 1443.9 | 14 |
| 317 | 79098 | Harzianin-HC-15 | linear | 5.9e-11 | 5 | 6.3e-07 | 32 | 1427.9 | 14 |
| 318 | 79098 | Trichorovin-Xa | linear | 6.6e-11 | 4 | 1.7e-08 | 10 | 1174.8 | 11 |
| 319 | 79098 | Protodestruxin | cyc | 6.8e-11 | 6 | 1.4e-05 | 32 | 579.4 | 6 |
| 320 | 79098 | Trichorovin-IVa | linear | 6.8e-11 | 5 | 1.7e-05 | 34 | 1160.8 | 11 |
| 321 | 79098 | Desmethyldestruxin-A | cyc | 8.0e-11 | 6 | 1.6e-05 | 32 | 563.3 | 6 |
| 322 | 78568 | kulomo'opunalide-1 | cyc | 8.0e-11 | 6 | 1.4e-03 | 40 | 716.5 | 6 |
| 323 | 79098 | Beauvericin | cyc | 8.8e-11 | 7 | 1.2e-03 | 85 | 783.4 | 6 |
| 324 | 78568 | theonellapeptolide-VI | bcyc | 9.3e-11 | 7 | 2.3e-01 | 41 | 1389.9 | 14 |
| 325 | 79098 | Trichorovin-IIb | linear | 9.8e-11 | 5 | 1.2e-06 | 36 | 1146.7 | 11 |

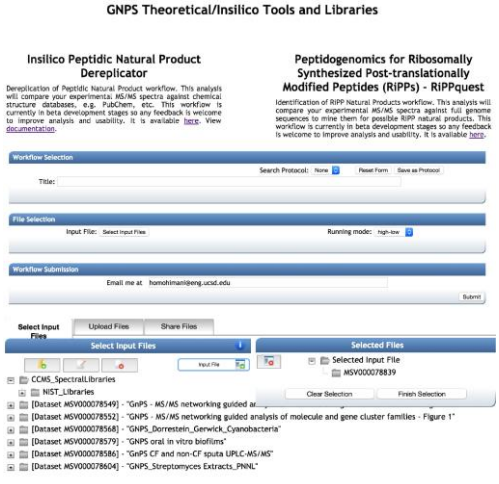
**Supplementary Table 7.** Annotation of genes in the surugamide biosynthetic gene cluster.

| name | length | predicted function | uniprot ID |
|---|---|---|---|
| orf-06544 | 680 | cell division protein FtsH | |
| orf-06547 | 201 | GTP cyclohydrolase I | SSHG-02505 |
| orf-06548 | 169 | integral membrane protein | SSHG-02504 |
| orf-06551 | 156 | 3-dehydroquinate dehydratase | SSHG-02503 |
| orf-06554 | 255 | dihydropteroate synthase | SSHG-02502 |
| orf-06557 | 388 | ABC transporter ATP-binding protein | SSHG-02501 |
| orf-06558 | 229 | ABC transporter permease | SSHG-02500 |
| orf-06559 | 253 | ABC transporter permease | SSHG-02499 |
| orf-06560 | 329 | amino acid ABC transporter substrate-binding protein | SSHG-02498 |
| orf-06562 | 395 | secreted protein | missing |
| orf-06564 | 61 | oxidoreductase | missing |
| orf-06565 | 524 | MFS transporter | SSHG-02496 |
| orf-06567 | 102 | hypothetical protein | missing |
| orf-06570 | 203 | TetR family transcriptional regulator | missing |
| orf-06571 | 452 | drug resistance transporter EmrB/QacA subfamily protein | missing |
| orf-06577 | 4114 | Nonribosomal peptide synthetase C-A-C-A-E-C-A-E | SSHG-02491-4 |
| orf-06580 | 7674 | Nonribosomal peptide synthetase C-A-C-A-E-C-A-C-A-E-C-C-A-C-A-E | SSHG-02490 |
| orf-06582 | 4285 | Nonribosomal peptide synthetase A-C-A-E-C-A-C-A | SSHG-02489 |
| orf-06587 | 5733 | Nonribosomal peptide synthetase A-C-A-E-C-A-C-A-E-C-A | SSHG-02487-8 |
| orf-06588 | 451 | alpha/beta hydrolase | SSHG-02486 |
| orf-06590 | 368 | membrane protein | SSHG-02485 |
| orf-06594 | 138 | GntR family transcriptional regulator | SSHG-02484 |
| orf-06595 | 123 | hypothetical protein | SSHG-02483 |
| orf-06597 | 313 | ABC transporter ATP-binding protein | SSHG-02482 |
| orf-06598 | 265 | ABC transporter permease | SSHG-02481 |
| orf-06601 | 91 | antibiotic synthesis protein MbtH | SSHG-02480 |
| orf-06604 | 269 | putative hydrolase or acyltransferase of alpha/beta superfamily | SSHG-02479 |

## Supplementary Figures



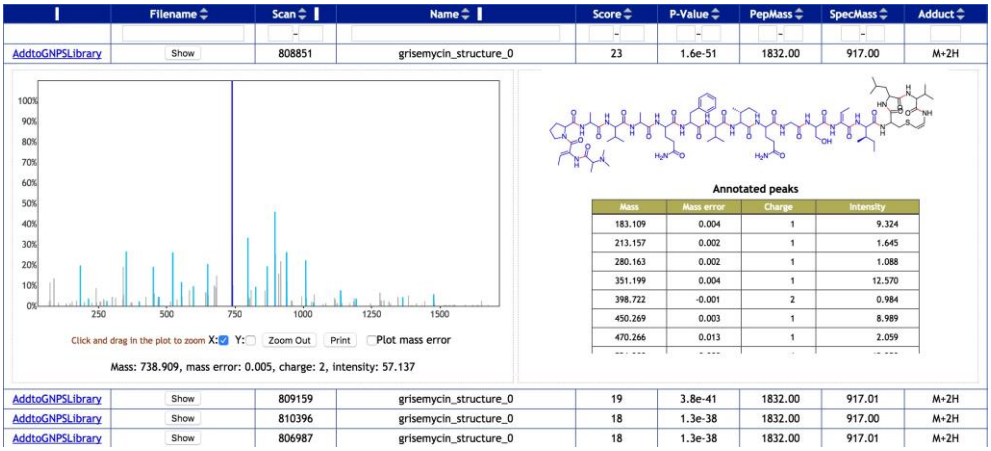

(a)




(b)

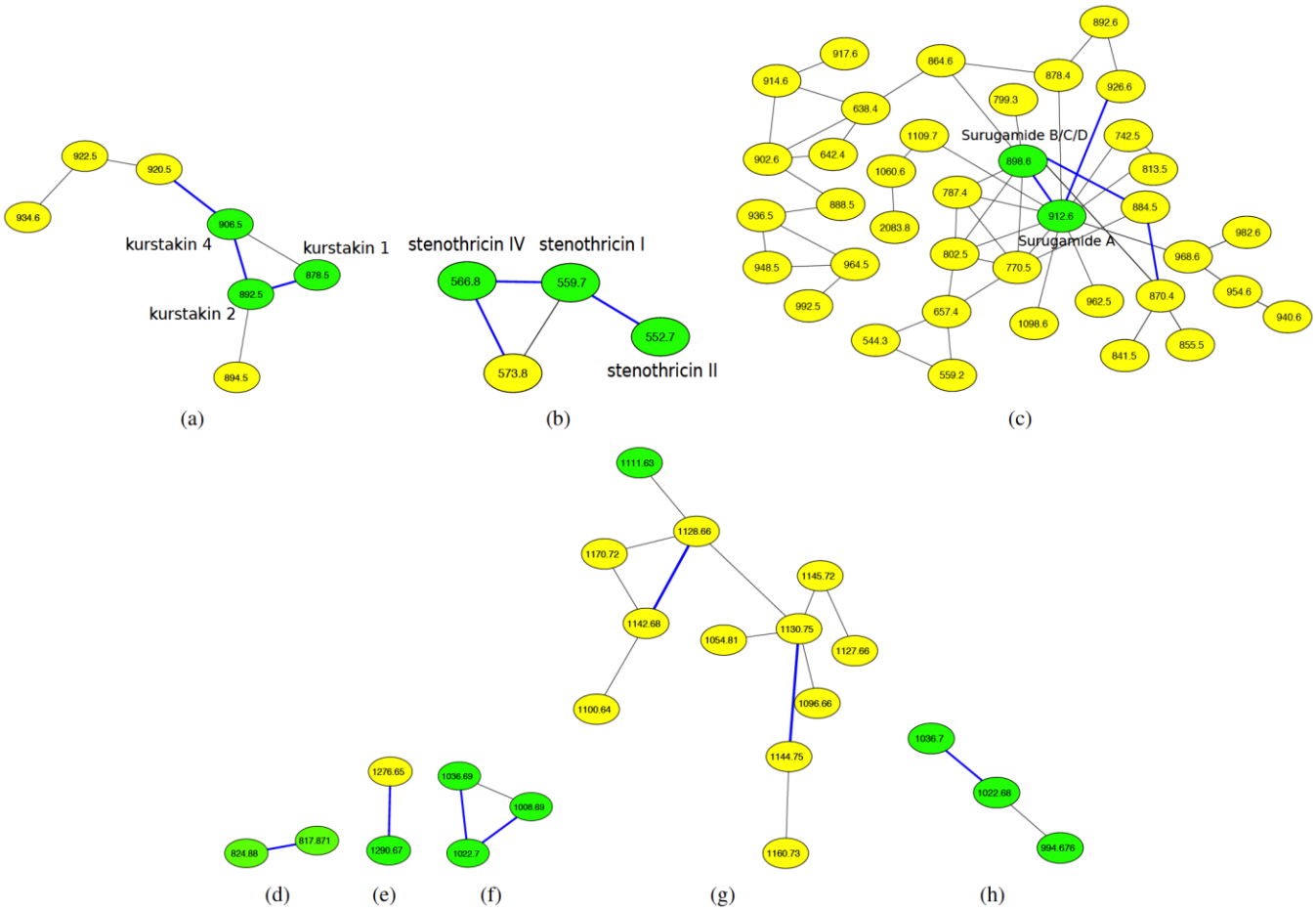
| | Filename | Scan | Name | Score | P-Value | PepMass | SpecMass | Adduct | PSM FDR % |
|---|---|---|---|---|---|---|---|---|---|
| AddtoGNPSLibrary | Show | 808851 | grisemycin_structure_0 | 23 | 1.6e-51 | 1832.00 | 917.00 | M+2H | 0.00 |
| AddtoGNPSLibrary | Show | 809159 | grisemycin_structure_0 | 19 | 3.8e-41 | 1832.00 | 917.01 | M+2H | 0.00 |
| AddtoGNPSLibrary | Show | 810396 | grisemycin_structure_0 | 18 | 1.3e-38 | 1832.00 | 917.00 | M+2H | 0.00 |
| AddtoGNPSLibrary | Show | 806987 | grisemycin_structure_0 | 18 | 1.3e-38 | 1832.00 | 917.01 | M+2H | 0.00 |
| AddtoGNPSLibrary | Show | 807227 | grisemycin_structure_0 | 16 | 1.2e-33 | 1832.00 | 917.01 | M+2H | 0.00 |
| AddtoGNPSLibrary | Show | 805890 | grisemycin_structure_0 | 14 | 9.1e-29 | 1832.00 | 917.01 | M+2H | 0.00 |
| AddtoGNPSLibrary | Show | 921817 | E'Surugamide_A' | 17 | 9.5e-28 | 911.62 | 912.62 | M+H | 0.00 |
| AddtoGNPSLibrary | Show | 923054 | E'Surugamide_A' | 17 | 9.5e-28 | 911.62 | 456.81 | M+2H | 0.00 |
| AddtoGNPSLibrary | Show | 864560 | E'Surugamide_A' | 17 | 9.5e-28 | 911.62 | 456.82 | M+2H | 0.00 |
| AddtoGNPSLibrary | Show | 863203 | E'Surugamide_A' | 17 | 9.5e-28 | 911.62 | 456.82 | M+2H | 0.00 |
| AddtoGNPSLibrary | Show | 811250 | grisemycin_structure_0 | 13 | 2.2e-26 | 1832.00 | 917.00 | M+2H | 0.00 |
| AddtoGNPSLibrary | Show | 812351 | grisemycin_structure_0 | 13 | 2.2e-26 | 1832.00 | 917.00 | M+2H | 0.00 |
| AddtoGNPSLibrary | Show | 804687 | grisemycin_structure_0 | 13 | 2.2e-26 | 1832.00 | 917.01 | M+2H | 0.00 |
| AddtoGNPSLibrary | Show | 863914 | E'Surugamide_A' | 16 | 9.2e-26 | 911.62 | 456.82 | M+2H | 0.00 |
| AddtoGNPSLibrary | Show | 816611 | grisemycin_structure_0 | 12 | 4.9e-24 | 1832.00 | 917.00 | M+2H | 0.00 |

(c)

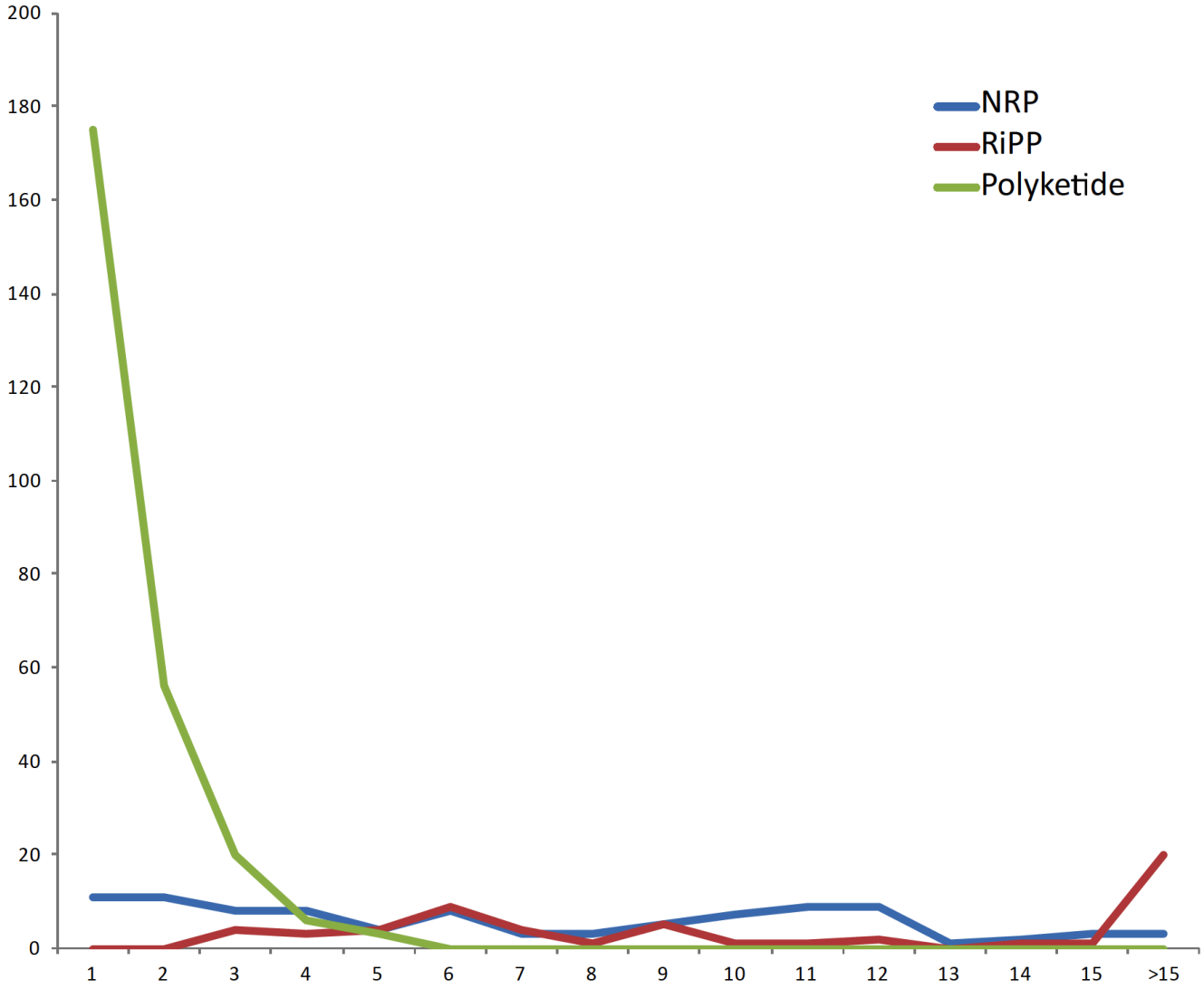



(d)

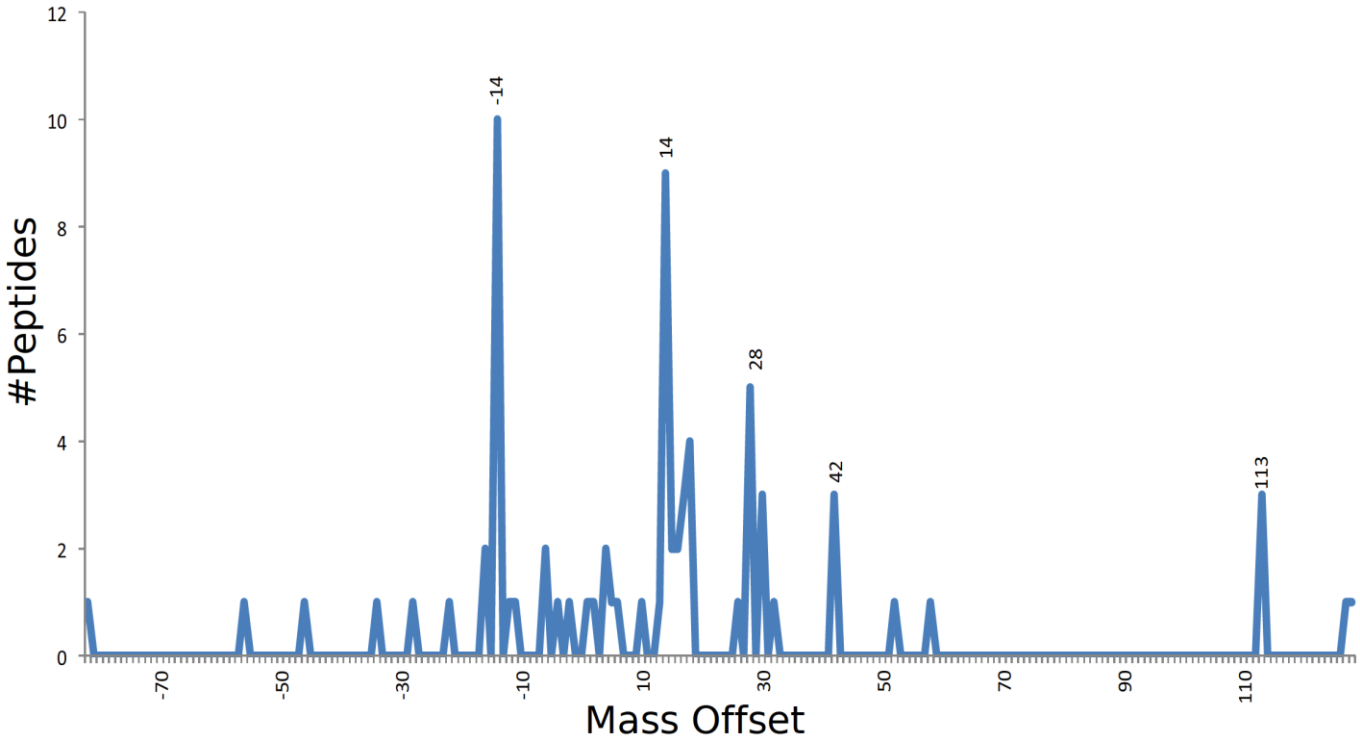
**Supplementary Figure 1.** Running DEREPLICATOR on the GNPS web server. (a) Image of the GNPS web page at www.gnps.edu. Dereplicator users click on ”Browse” in the section ”GNPS Theoretical/Insilico” (b) At ”Insilico Peptidic Natural Product Dereplicator” page, users click on ”here”. To select input files, users click on ”Select Input Files”. Users can either import one of the existing GNPS public datasets from the ”Share Files” section, or use their own data from ”Upload Files” section. Users select a title/email for the job and specify the DEREPLICATOR accuracy mode (low-low, high-low, or high-high). Users receive a notification emal after the job is completed. (c) When the job is completed, users click on ”View Significant Matches”. (d) After selecting a Peptide-Spectrum-Match, the users are presented with a vizualization of the peptide and the spectrum forming a PSM as well as the annotation of the spectrum. The annotated peaks in the spectrum are shown in blue. After selecting an annotated peak, the corresponding subpeptide will be shown in blue.



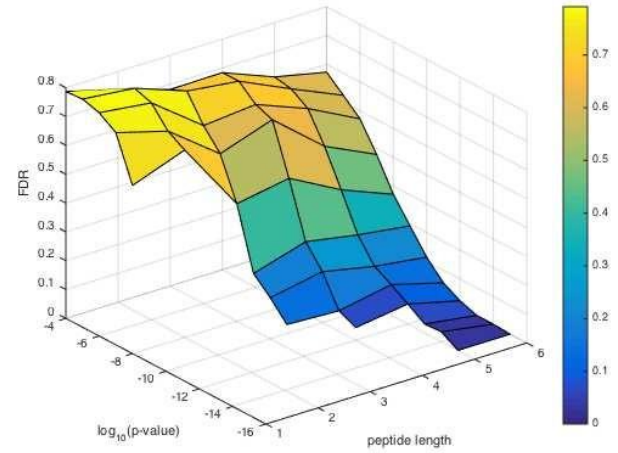

**Supplementary Figure 2.** Spectral networks of various PNPs. Spectral networks of kurstakins (a), stenothricins (b), surugamides from *S. albus* J1074 from dataset MSV000078604 (c), daptomycins from MSV000078577 (d), glumamycins from MSV000078557 (e), lichenycins from MSV000078557 (f), valinomycins from MSV000078557 (g), and surfactins from MSV000078552 (h). The spectral network of stenothricins is formed by doubly charged ions. Green (yellow) nodes represent known (unknown) PNPs. Blue edges in the network represent the most common mass shift 14 Da (7 Da in the case of doubly charged ions). Five variants of surugamide that differ only in methylations are represented by nodes in a blue path. Further relaxing the accuracy threshold reveals 11 methylated variants of surugamide with masses varying from 842.5 Da to 982.6 Da.



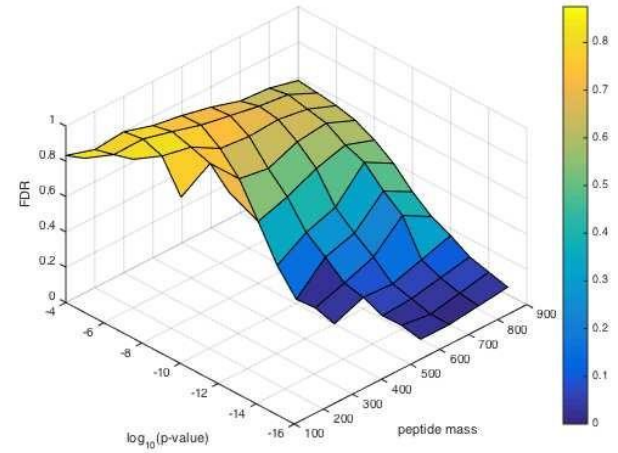

**Supplementary Figure 3.** Number of polyketides, NRPs, and RiPPs with a fixed number of amide bonds in the MiBIG dataset. The lion's share of compounds with less than 5 bonds are polyketides. Most compounds with less than five peptide bonds are classified as non-peptides.



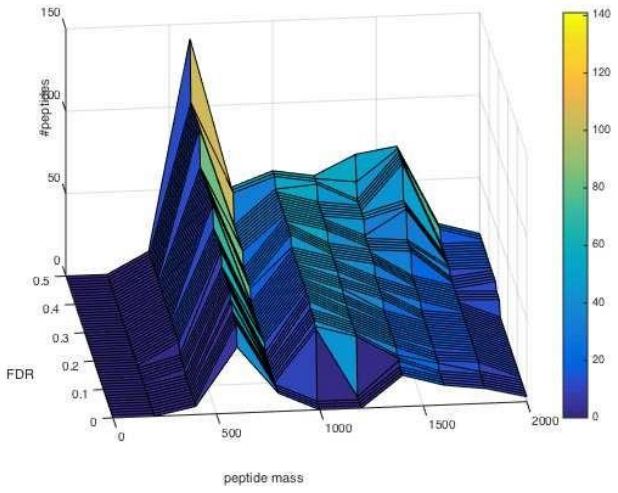

**Supplementary Figure 4.** Mass offsets identified by variable dereplication. Number of PSMs with specified mass offsets as identified by variable dereplications from PNPs listed in Supplementary Table 4. The histogram reveals the characteristic peaks at -14Da, 14Da, 28Da, 42Da (methylation/dimethylation/trimetylation), 17Da (ammonium adduct), 18Da (hydration), and 113Da (possible insertion or deletion of Leucine/Isoleucine).




(a)

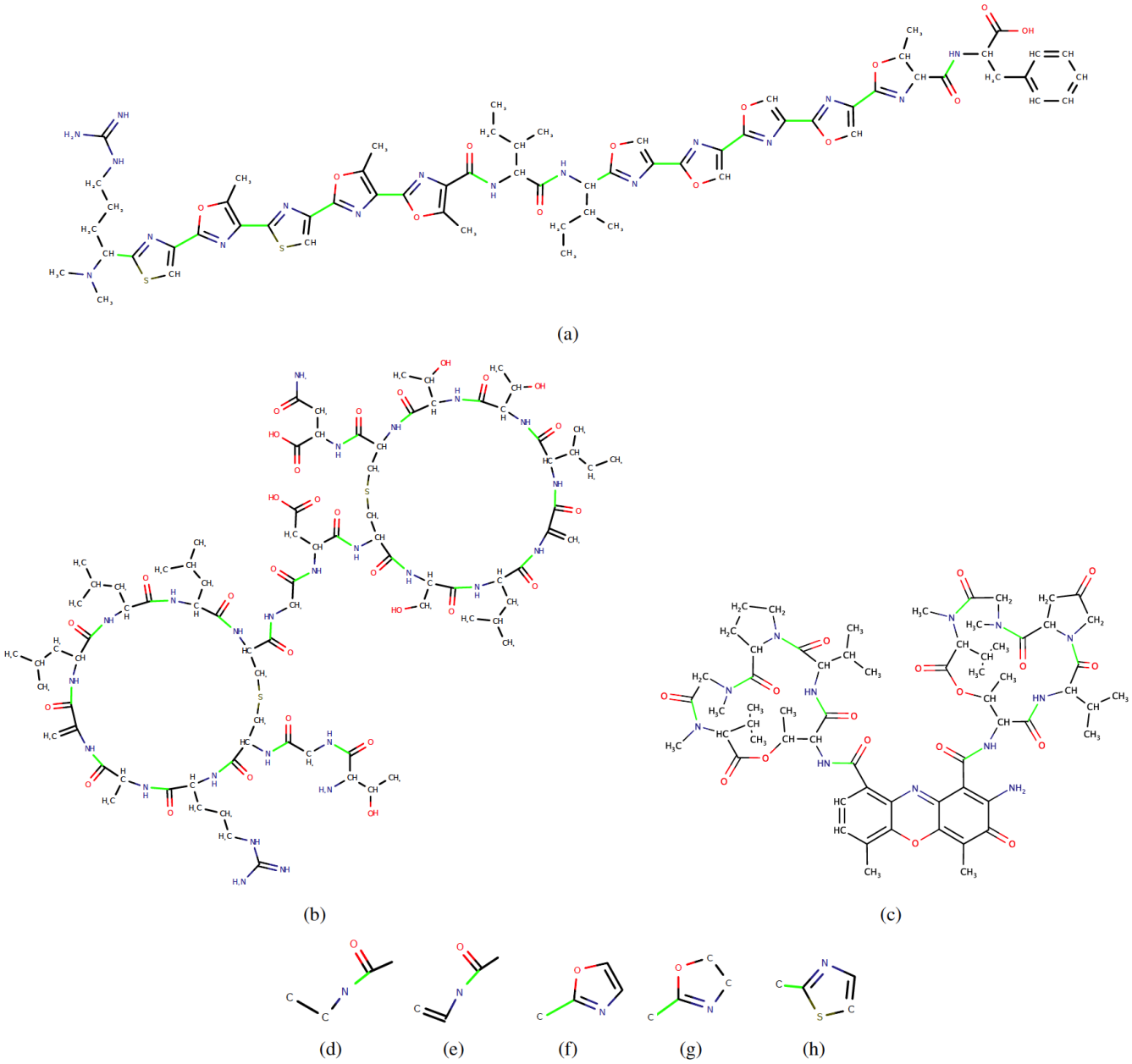



(b)

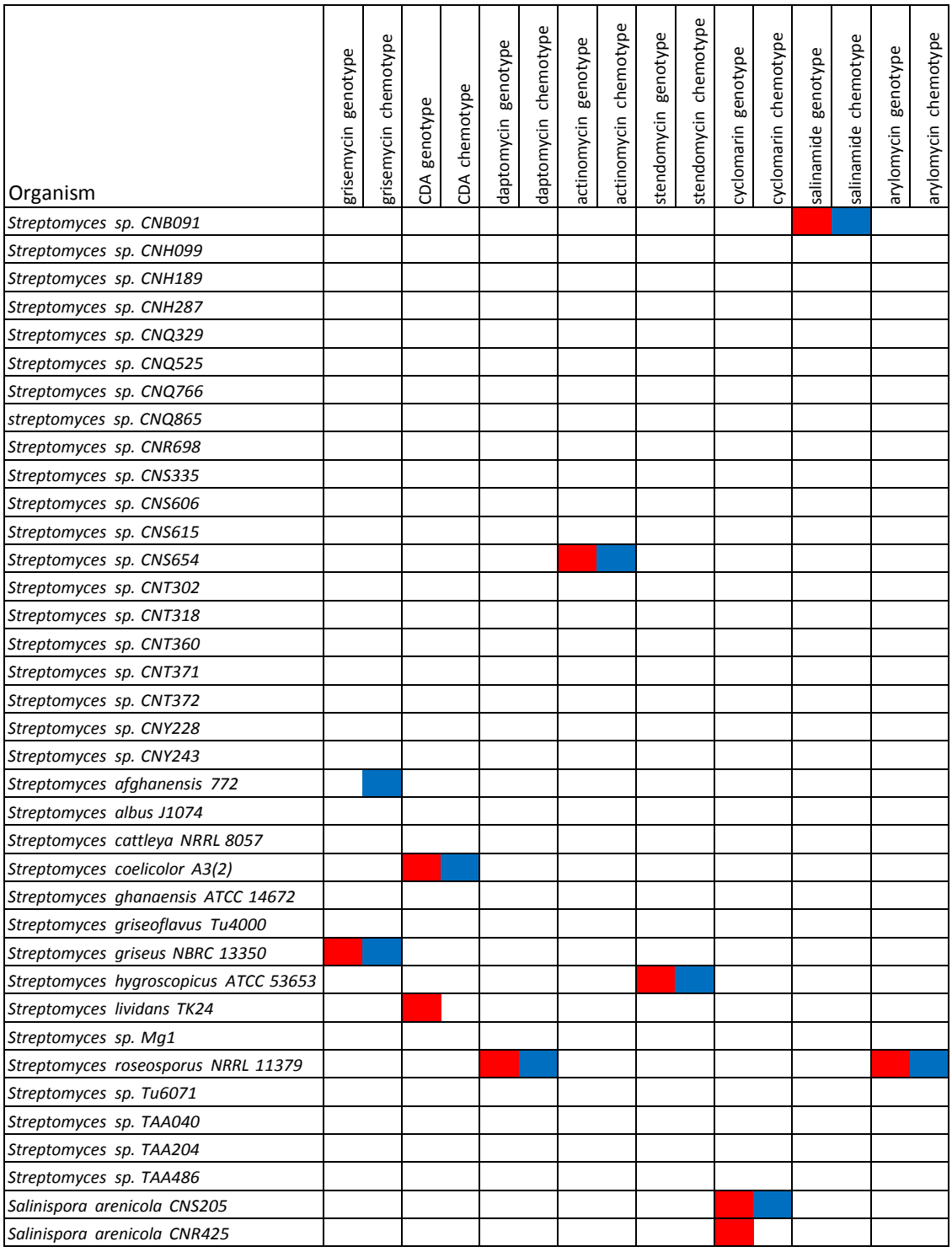



(c)

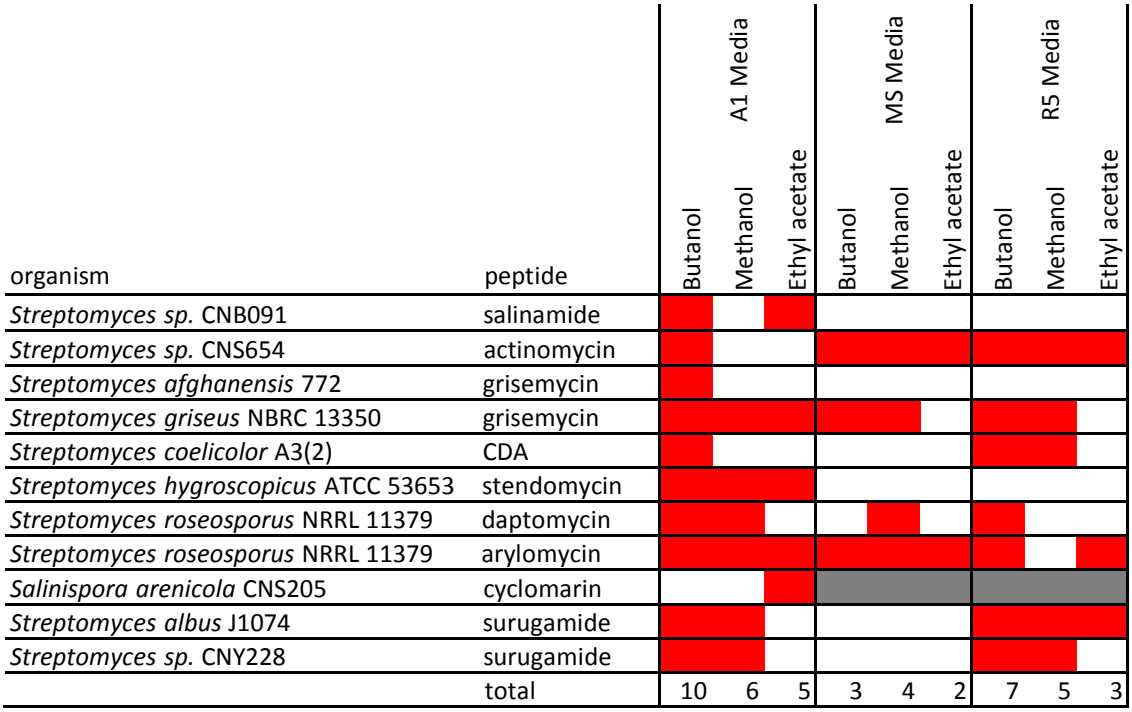
**Supplementary Figure 5.** Analyzing false discovery rates for various peptide lengths. (a) FDR as a function of the peptide length and the p-value threshold. (b) FDR as a function of the peptide mass and the p-value threshold. (c) Number of unique peptide identifications as a function of FDR and peptide mass.



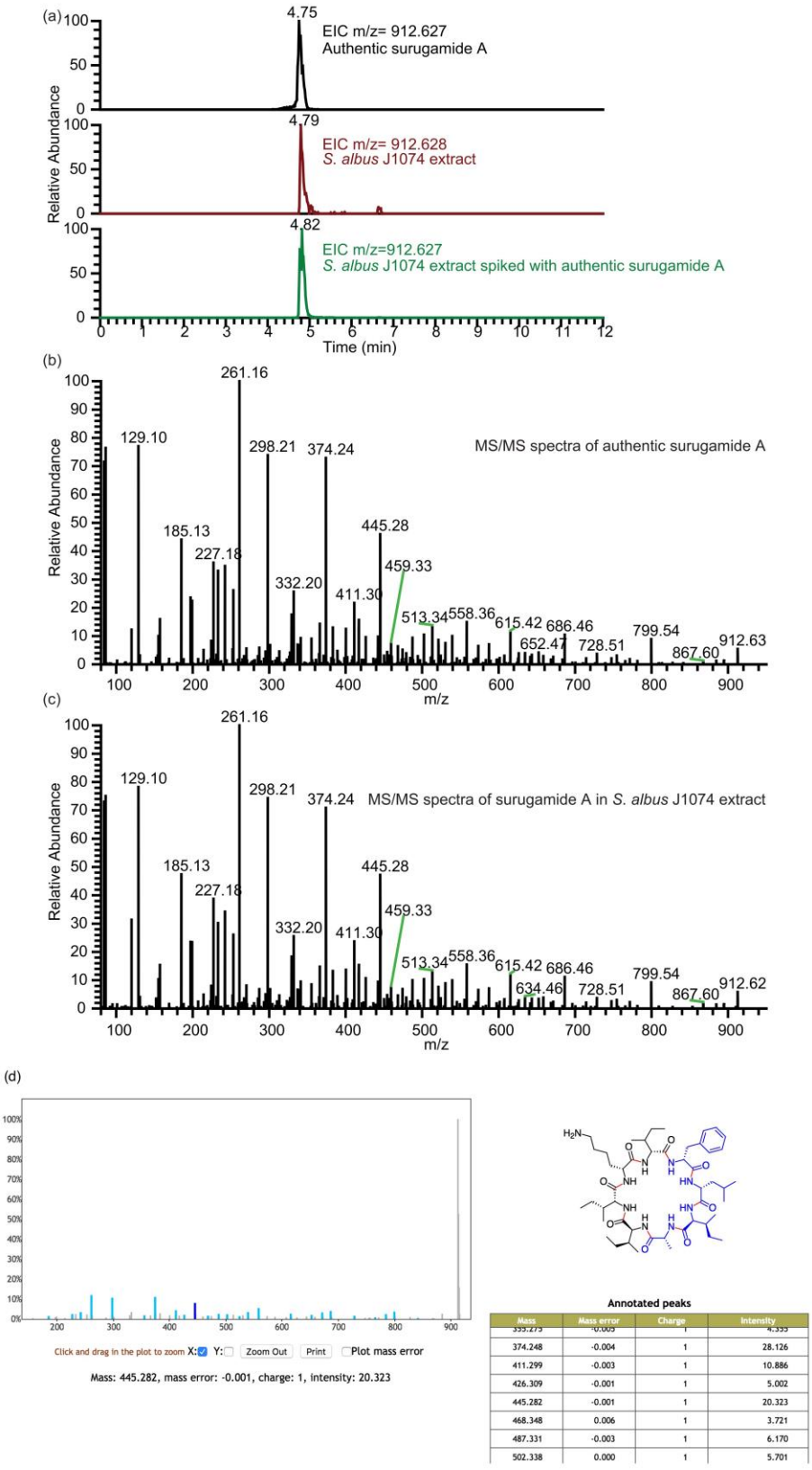

**Supplementary Figure 6.** Various chemical structures and peptide bonds. Chemical structures of (a) thiazole/oxazole containing peptide plantazolicin A from B. amyloliquefaciens, (b) lanthipeptide SapB from S. coelicolor A3, (c) actinomycin X2 from S. sp. CNS-654. Panels (d)-(h) show five types of substructures that are searched in the compound structure for identification of the generalized peptide bonds. These substructures include amide bonds (d), Dha/Dhb bonds like ones shown in SapB (e), oxazole/methyl-oxazole bonds (f), oxazoline/methyl-oxazoline bonds (g), thiazole bonds like ones shown in plantazolicin (h). Generalized peptide bonds are shown in green.



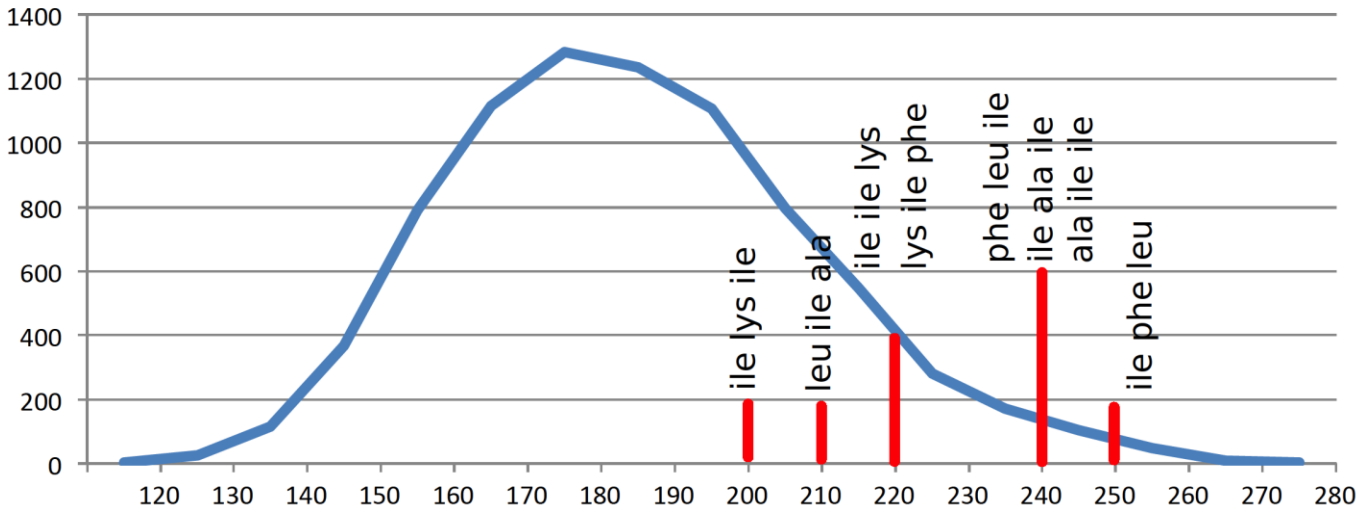

**Supplementary Figure 7.** Cross-validating genome mining and peptidogenomics results. Cross-validating chemotypes reported by DEREPLICATOR with genotypes predicted in genomes of 36 actinomycetes. Red (blue) squares stand for discovery of the genotype (chemotype). In majority of the cases, when DEREPLICATOR identifies a chemotype in a specific strain, genome mining predicts the corresponding genotype in the same strain. Only in the case of grisemycin, DEREPLICATOR predicted its production in both *S. griseus* and *S. afghanensis*, while genome mining predicted the genotype only in *S. griseus*. On the other hand, genome mining showed evidence for CDA in both *S. coelicolor* and *S. lividans*, while DEREPLICATOR identified CDA only in *S. coelicolor*.



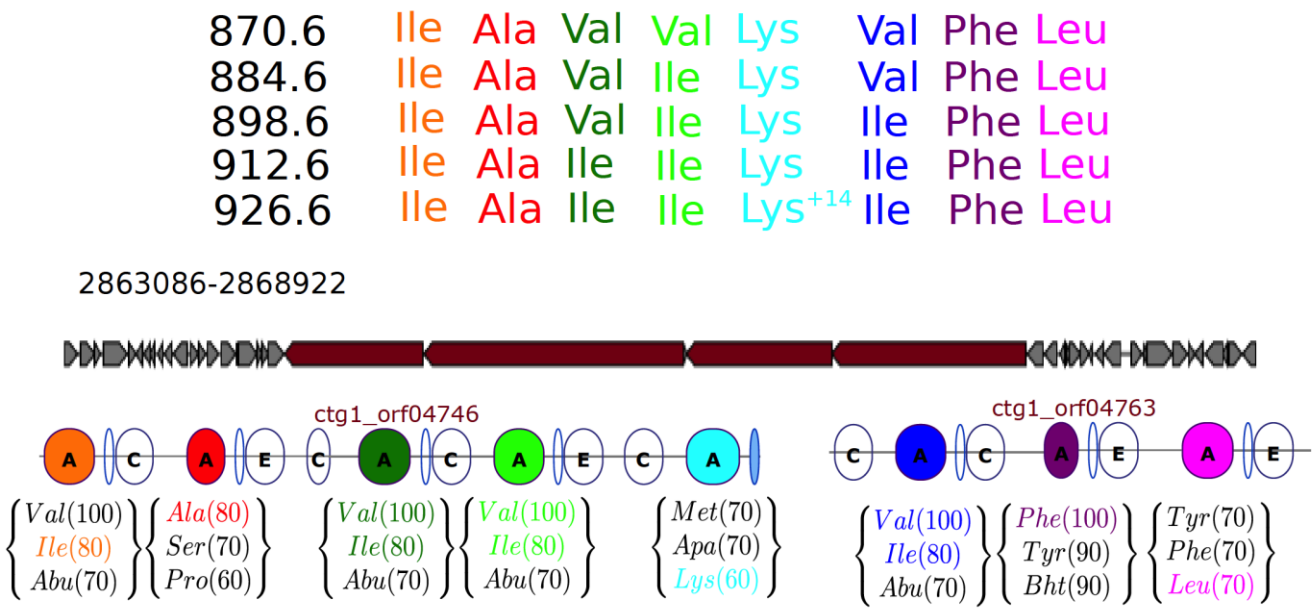

**Supplementary Figure 8.** Comparison of various combinations of media agars and extract conditions for PNP discovery in 36 actinomycetes. The dataset $Spectra_{Act\,i36}$ was collected under three different growth conditions (A1 agar, MS agar and R5 agar) and extracted in three different ways (ethyl acetate, butanol and methanol). A red square indicates that a PNP was identified in a given strain for a given sample preparation protocol (gray squares indicate that experimental data is not available). The p-value in this analysis was increased from $10^{-10}$ to $10^{-5}$. This reduction is statistically justified since it did not bring new peptides but only increased the number of conditions where the peptides were found. The number of recovered peptides varied from 2 to 10 for various combinations.



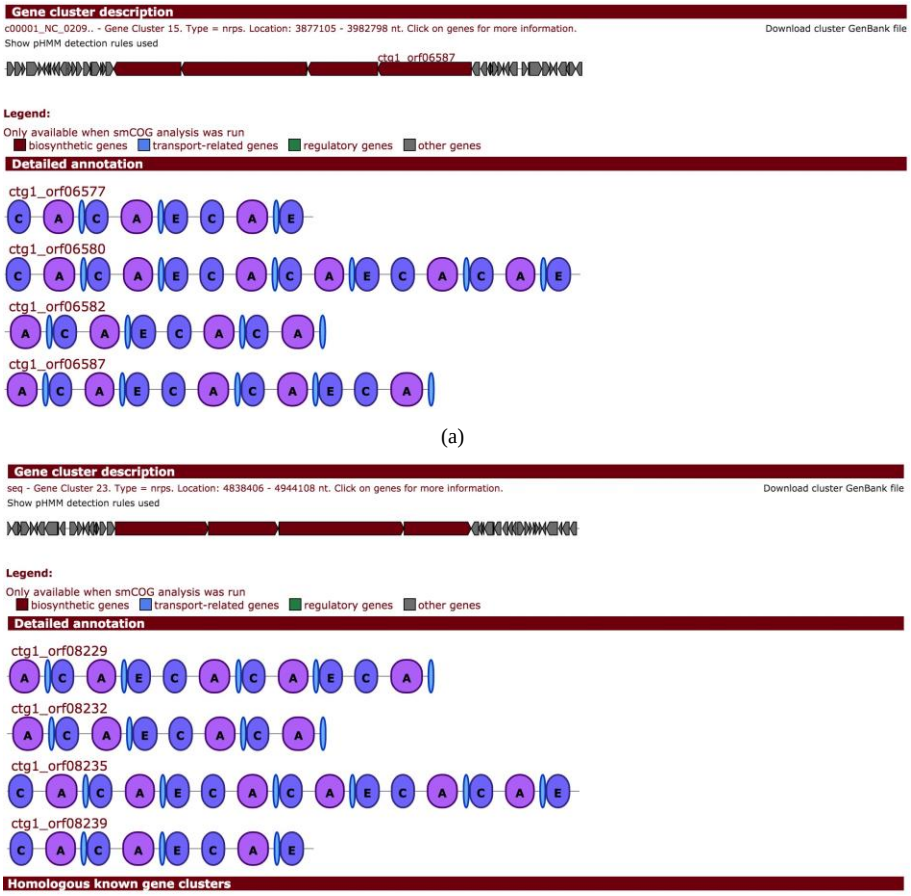

**Supplementary Figure 9.** Comparison of authentic surugamide A with the compound identified in *S. albus* J1074. (a) The extracted ion chromatogram for the authentic surugamide A (black), surugamide A in methanol extract of *S. albus* J1074 (red), and surugamide A in methanol extract of S. albus J1074 spiked with the authentic surugamide A standard (green). The peaks are labelled by the retention time. (b) MS/MS spectrum of the authentic surugamide. (c) MS/MS spectrum of the putative surugamide identified by Dereplicator in methanol extract of *S. albus* J1074. (d) Annotation of mass spectra of surugamide A




**Supplementary Figure 10.** Search for surugamide biosyntetic gene cluster in *S. albus* J1074. The score distribution for all 3-mers formed by the proteinogenic amino acids and all triples of consecutive adenylation domains in *S. albus* J1074. The scores of eight 3-mers from surugamide A (IAIIKIFL) are shown by red bars.




**Supplementary Figure 11.** Five surugamides (along with their masses) and the proposed biosyntetic gene cluster explaining them. The ORFs 4751 (4 adenylation domains) and 4755 (6 adenylation domains), separating ORFs 4748 and 4759 are not shown. Three most likely amino acids for each adenylation domain are shown along with their NRPS2predictor scores. The residues in these NRPs correlate with the residues predicted by NRPS2predictor (shown by the same colors). While most residues in the peptide IAIIK$^{+14}$IFL correlate with the proposed adenylation domains, the nature of K$^{+14}$ in this peptide remains unclear. The biosynthetic gene cluster was identified by antiSMASH. For simplicity, the residues Ile$^{-14}$ are shown as Val.




**Supplementary Figure 12.** Surugamide gene cluster revealed by antiSMASH in *S. albus* J1074 and *S. sp. CNY228*. The NRPS gene cluster in in *S. albus* J1074 (a) is inverted as compared to the cluster in *S. sp. CNY228* (b).

## Supplementary Note

**Revealing the biosynthetic gene cluster for surugamides.** Surugamide, champacyclin, and reginamide are recently discovered NRPs from marine *Streptomyces* that share the same amino acid residues. Regi- namide was computationally predicted based on mass spectrometry data, while champacyclin has an inversion of two amino acids in a region that had low NMR data signal. Thus, there is a possibility that a reversal of two amino acids in champacyclin may represent an artifact rather than an evolutionary change.

To reveal the biosyntetic gene cluster for surugamides, we focused attention on 5 consecutive adenylation domains (ORFs 4746 and 4749) and 3 consecutive adenylation domains (ORFs 4760 and 4763) in *S. albus* J1074. While these two clusters of adenylation domains are not consecutive in the genomic scaffold (there are 10 other adenylation domains separating them), Supplementary Figures 11 and 12 (combined with the fact that the *S. albus* J1074 assembly may be error-prone) suggest that they are likely to code for surugamides A and B as there are no other gene clusters that indicate the releveant amino acid specificity. Surugamide A has four amino acids in the D configuration and the location of the epimerization domains in the biosynthetic gene cluster match that of the structure, an additional evidence that biosynthetic gene cluster is correctly predicted. DEREPLICATOR also identified surugamide in *Streptomyces sp. CNY228* where there exists a very similar gene cluster with identical specificities.

In addition to surugamides A and B, the analysis of the spectral network in Supplementary Figure 2 revealed surugamide-like peptides $\mathrm{IAI}^{-14}\mathrm{IKI}^{-14}\mathrm{FL}$ and $\mathrm{IAI}^{-14}\mathrm{I}^{-14}\mathrm{KI}^{-14}\mathrm{FL}$ (labeled as IAVIKVFL and IAVVKVFL in Supplementary Figure 11) as well as $\mathrm{IAIIK}^{+14}\mathrm{IFL}$. As Supple- mentary Figure 11 illustrates, 4 out of 5 surugamide-like peptides (IAIIKIFL, IAVIKIFL, IAVIKVFL, and IAVVKVFL) revealed by the spectral network are likely to be generated through promiscuous adenylation domains, an additional argument supporting our computational analysis. Supplementary Table7 shows annotation of the genes in the surugamide biosynthetic gene cluster.

**Experimental methods for analyzing surugamides.** After 72 h of growth, 150 uL of 1 mM 13C6-labeled isoleucine was spread on the plate by swirling the plate gently. For incorporation of lysine, *S. albus* J1074 and *S. albus* ATCC 21838 were cultured on ISP2 media. After 24 h, a filter disk containing 0.1 mg of $^{13}C_6$ -labeled lysine in 10 uL was placed next to growing colony. Additional 0.1 mg of amino acid in 10 uL of sterile water was added to the same filter disk after 72 and 120 h of growth. The agar media was cut into small pieces and placed in 10 mL of butanol. The butanol containing the agar was sonicated for 15 min and incubated at room temperature for 1 h. The dried extracts were resuspended in same volume of 80% methanol in water and analyzed by a QExactive (Thermo Scientific) mass spectrometer with HESI-II probe source.

An external calibration was performed with Pierce LTQ Velos ESI positive ion calibration solution (Thermo Scientific) prior to spectral data acquisition. The following probe settings were used for flow aspiration and ionization: spray voltage of 3500 V, sheath gas (N2) pressure of 53 psi, auxiliary gas pressure (N2) of 14 psi, ion source temperature of 250 C, S-lens RF level of 50 Hz and auxiliary gas heater temperature at 440° C.

Data acquisition parameters were as follows: data was recorded starting at 0.5 min with data-dependent MS/MS acquisition mode. Full scan at MS1 level was performed with resolution of 35 K in profile mode. The 10 most intense ions with 1 m/z isolation window per MS1 scan were subjected to normalized Collision Induced Dissociation at 30 eV. MS2 scans were performed at 17.5 K resolution in profile mode. The injected samples were chromatographically separated using a Vanquish uHPLC (Thermo Scientific) using a KinetexTM 1.7 um C18 reversed phase (50 X 2.1 mm) chromatography column (Phenomenex) at a flow rate of 0.5 mL/min with mobile phase A as 0.1% formic acid (Fisher Scientific, Optima LC/MS) in water (J.T.Baker, LC-MS grade) and mobile phase B as 0.1% formic acid (Fisher Scientific, Optima LC/MS) in acetonitrile (J.T.Baker, LC-MS grade).

The separation was performed using the following gradient: 0-1 min at 5% solvent B followed by a linear gradient from 5% solvent B to 100% solvent B over 7 min. The column was then held at 100% solvent B for 3 min and equilibrated back to 5% solvent B for 2 min. The column compartment was maintained at 40° C throughout the runs.

The raw data files were converted to the .mzXML format using ProteoWizard (http://proteowizard.sourceforge.net/) and uploaded to the GNPS (MSV000079516 for labeled isoleucine

experiment and MSV000079517 for labeled lysine experiment). Molecular network was generated with a minimum cosine threshold of 0.75 and minimum matched peaks of 6. All other parameters are default.

Purified and NMR-characterized authentic surugamide A and the *S. albus* J1074 extract were analyzed using ultra high performance liquid chromatog- raphy coupled to mass spectrometry. The authentic surugamide A was also added to the crude extract of *S. albus* J1074 and was analyzed under identical conditions. Solutions of the authentic surugamide A. and the *S. albus* J1074 extract were dissolved in methanol at 2 µM and 2 mg/mL, respectively. The co-migration assay was performed by adding 1 µL of the authentic surugamide A solution into 100 µL of the crude extract solution of *S. albus* J1074. LC-MS/MS analysis was performed under the same analytical conditions.