## **Electronic Supplementary Information:**

## Marine indole alkaloid diversity and bioactivity. What do we know and what are we missing?

Darren C. Holland <sup>ab</sup> and Anthony R. Carroll <sup>\*ab</sup>

<sup>o</sup>School of Environment and Science, Griffith University, Southport, Queensland, 4222, Australia <sup>b</sup>Griffith Institute for Drug Discovery, Griffith University, Nathan, Queensland, 4111, Australia \*corresponding author email - <u>a.carrol@griffith.edu.au</u>

## **Table of Contents**

Figure S1	MIAs reported yearly from 1972 to the end of 2021.
Figure S2	Worldwide distribution of collecting sites for MIAs coloured by their producing phyla (up
	to the end of 2021).
Figure S3	MIAs reported 2002 to the end of 2021 from marine fungi (green), marine bacteria (red),
	and marine invertebrates (yellow).
Figure S4	Indoline (C-3 saturated indole) containing MIAs reported per major producing phyla (up
	to the end of 2021).
Figure S5	Chemical diversity of MIAs (n = 2048) visualised in a 50 x 50 cell self-organising map
	(SOM) using the SkelSpheres chemical descriptor scaled to producing phyla.
Figure S6	A: Proportion of not-tested, inactive, weak, moderate, and potent MIAs. B: Underlying
	proportion of unknown potential MIA bioactivity (combined not tested, inactive and
	weak), moderate, and potent MIAs.
Figure S7	Number of different disease targets examined per MIA scaled for bioactivity and
	trendline for proportion of actives (moderate to potent).
Figure S8	Heatmap (log scale) of MIAs tested against assay or disease target per major MIA
	producing phyla (increasing frequency of testing from blue to red).
Figure S9	Overall potency of bioactivity reported for fungi and Porifera MIAs.
Figure S10	Average number of disease targets examined per level of bioactivity for fungi vs Porifera
	MIAs.
Figure S11	Scequinadoline D, E and J (11-13) related fumiquinazoline alkaloids from diabetes mode
	of action SOM analysis (Fig. 6) including fumiquinazoline F and G (14 and 15),
	cladoquinazoline (16), epi-cladoquinazoline (17), norquinadoline A (18), 3-
	hydroxyglyantrypine (19), oxoglyantrypine (20), versiquinazoline I (21), sceuinadoline G
	(22), aspertoryadin F and G (23 and 24), and quinadoline D (25).
Figure S12	Pyrrolocarbazole alkaloids dictyodendrins A-J (29-38) (related to the dictyodendrins 25-
5	<b>28</b> ) highlighted in diabetes mode of action SOM ( <b>Fig. 6</b> ).
Figure S13	SOM of MIA chemical diversity scaled for CNS mode of action activities displaying cluster
	of MAO-A and B inhibitory N-2' and N-3' aplysinopsins; aplysinopsin ( <b>39</b> ), 6-
	bromoapiysinopsin (40), methylaplysinopsin (41), and 6-bromo- <i>N</i> -methylaplysinopsin
	(42), and unexamined analogues; 8E-3'-deimino-3'-oxoaplysinopsin (43), <i>N</i> -propionyl 6-
	bromoaplysinopsin (44), <i>N</i> -propionyl aplysinopsin (45), and <i>N</i> -3'-ethylaplysinopsin (46).

**Figure S14** 50 x 50 SOM (SphereFP descriptor) of MIA and approved indole drugs chemical diversity with red circles identifying clusters of shared MIA/synthetic chemical space: A. approved drug midostaurin (**47**) and MIA fradcarbazole A (**48**), B. approved drug lurbinectedin (**49**) and MIA ecteinascidin-736 (**50**), and C. approved drug daptomycin (**51**) and MIA taromycin B (**52**).



Fig S1. MIAs reported yearly from 1972 to the end of 2021.



Fig S2. Worldwide distribution of collecting sites for MIAs coloured by their producing phyla



**Fig S3.** MIAs reported 2002 to the end of 2021 from marine fungi (green), marine bacteria (red), and invertebrates (yellow).



**Fig S4.** Indoline (C-3 saturated indole) containing MIAs reported per major producing phyla (up to the end of 2021).



**Fig S5.** Chemical diversity of MIAs (n = 2048) visualised in a 50 x 50 cell self-organising map (SOM) using the SkelSpheres chemical descriptor scaled to producing phyla.



**Fig S6.** A: Proportion of not-tested, inactive, weak, moderate, and potent MIAs. B: Underlying proportion of unknown potential MIA bioactivity (combined not tested, inactive and weak), moderate, and potent MIAs.



**Fig S7.** Number of different disease targets examined per MIA scaled for bioactivity and trendline for proportion of actives (moderate to potent).



**Fig S8.** Heatmap (log scale) of MIAs tested against assay or disease target per major MIA producing phyla (increasing frequency of testing from blue to red).



Fig S9. Potency of bioactivity reported for Ascomycota and Porifera MIAs.



**Fig S10.** Average number of disease targets examined per level of bioactivity for fungi *vs* Porifera MIAs.



**Fig S11.** Scequinadoline D, E and J (**11-13**) related fumiquinazoline alkaloids from SOM analysis including fumiquinazoline F and G (**14** and **15**), cladoquinazoline (**16**), epi-cladoquinazoline (**17**), norquinadoline A (**18**), 3-hydroxyglyantrypine (**19**), oxoglyantrypine (**20**), versiquinazoline I (**21**), sceuinadoline G (**22**), aspertoryadin F and G (**23** and **24**), and quinadoline D (**25**).



**Fig S12.** Pyrrolocarbazole alkaloids dictyodendrins A-J (**29-38**) (related to the dictyodendrins **25-28**) highlighted in diabetes mode of action SOM (**Fig. 6**).



**Fig S13.** SOM of MIA chemical diversity scaled for CNS mode of action activities displaying cluster of MAO-A and B inhibitory N-2' and N-3' aplysinopsins; aplysinopsin (**39**), 6-bromoaplysinopsin (**40**), methylaplysinopsin (**41**), and 6-bromo-*N*-methylaplysinopsin (**42**), and unexamined analogues; 8E-3'-deimino-3'-oxoaplysinopsin (**43**), *N*-propionyl 6-bromoaplysinopsin (**44**), *N*-propionyl aplysinopsin (**45**), and *N*-3'-ethylaplysinopsin (**46**).



**Fig S14.** 50 x 50 SOM (SphereFP descriptor) of MIA and approved indole drugs chemical diversity with red circles identifying clusters of shared MIA/synthetic chemical space: A. approved drug midostaurin (**47**) and MIA fradcarbazole A (**48**), B. approved drug lurbinectedin (**49**) and MIA ecteinascidin-736 (**50**), and C. approved drug daptomycin (**51**) and MIA taromycin B (**52**).

Α