Peer-reviewed version available at Mar. Drugs 2019, 17, 671; doi:10.3390/md17120671

Article

Marine Bacteria from Rocas Atoll: A Promising Biotechnological Resource

Karen Y. Velasco-Alzate[1#], Anelize Bauermeister[1,2#], Marcelo M. P. Tangerina[3], Tito M. C. Lotufo[4], Marcelo J. P. Ferreira[3], Paula C. Jimenez[5], Gabriel Padilla[6], Norberto P. Lopes[2], Letícia V. Costa-Lotufo[1*]

[1] Departamento de Farmacologia, Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo/SP, Brazil; karenvelasco20@gmail.com (K.Y.V.A.); ane.meister@usp.br (A.B.); costalotufo@usp.br (L.V.C.L.)
[2] NPPNS, Departamento de Física e Química, Faculdade de Ciências Farmacêuticas de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto/SP, Brazil; npelopes@fcfrp.usp.br (N.P.L.)
[3] Departamento do Botânica, Instituto de Biociências, Universidade de São Paulo, São Paulo/SP, Brazil; marcelotangerina@usp.br (M.M.P.T.); marcelopena@ib.usp.br (M.J.P.F.)
[4] Departamento de Oceanografia Biológica, Instituto Oceanográfico, Universidade de São Paulo, São Paulo/SP, Brazil; tmlotufo@gmail.com (T.M.C.L.)
[5] Departamento de Ciências do Mar, Universidade Federal de São Paulo, Santos/SP, Brazil; paulacjimenez@gmail.com (P.C.J.)
[6] Departamento de Microbiologia, Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo/SP, Brazil; gpadilla@icb.usp.br (G.P.)
* Correspondence: costalotufo@usp.br ; Tel.: +55-11-30917316
#authors with equal contribution

Abstract: Rocas Atoll is a unique environment in the Equatorial Atlantic Ocean, hosting a large number of endemic species and studies on the chemical diversity emerging from this biota are rather scarce. Therefore, the present work aims to assess the metabolomic diversity and pharmacological potential of the microbiota from Rocas Atoll. A total of 76 bacteria were isolated and cultured in liquid culture media to obtain crude extracts. About one third (34%) of these extracts were considered cytotoxic against human colon adenocarcinoma HCT-116 cell line. 16S rRNA gene sequencing analysis revealed that the bacteria producing cytotoxic extracts are mainly from the Actinobacteria phylum, including *Streptomyces*, *Salinispora*, *Nocardiopsis* and *Brevibacterium* genera, and in a smaller proportion from Firmicutes phylum (*Bacillus*). The search in the GNPS spectral library unveiled a high chemodiversity being produced by these bacteria, including rifamycins, antimycins, desferrioxamines, ferrioxamines, surfactins, surugamides, staurosporine and saliniketals, along with several unidentified compounds. Using an original approach, molecular network successfully highlighted groups of compounds responsible for the cytotoxicity of crude extracts. DEREPLICATOR+, a recently developed in silico tool (GNPS), allowed the identification of derivatives of the macrolide novonestimycin, as the cytotoxic compounds into the extracts produced by *Streptomyces* BRB-298 and BRB-302. Overall, these results highlighted the pharmacological potential of bacteria from this singular Atoll.

Keywords: Secondary metabolites; microbial diversity; metabolomics; molecular network; marine bacteria.

1. Introduction

Marine environments are a rich resource of genetic diversity, further translated through the biosynthesis of unique and complex chemical structures with biotechnological and pharmacological interest [1, 2]. During the last decade, around 1,000 new natural products have been described each

Peer-reviewed version available at Mar. Drugs 2019, 17, 671; doi:10.3390/md17120671

year and, in 2016 alone for instance, 1277 new compounds were reported in 432 papers [1]. This chemical diversity arises from a great variety of organisms, such as microorganisms, algae, tunicates, sponges, mollusks, bryozoans, cnidarians, among many others, making marine environments particularly interesting to be deeply investigated in order to discover new compounds with promising therapeutic properties.

Brazil has the second largest continuous coastline in the world, with circa 8.500 km of extension and over 3.5 million km2 of economic exclusive zone. It also includes five sets of tropical oceanic islands – Fernando de Noronha Archipelago (FNA), Trindade Island and Martim Vaz Archipelago, Rocas Atoll (RA) and Saint Pieter and Saint Paul Archipelago (SPSPA) – and therefore, favorable to the development of a rich and diverse biota. SPSPA, RA and FNA are located in the equatorial zone of the Atlantic, and, although they form a set of islands with a great number of overlapping biotic components, each hosts peculiar faunas and cases of endemism [3, 4].

Atolls are oceanic islands with volcanic and biogenic formation. Its structure is usually ring-like, forming a lagoon inside with little water exchange, resulting in the development of biological reefs. Such conditions make these environments favorable to host a diverse biota, and consecutively, a rich source of marine natural products. This kind of natural formation is rather common in the Pacific and Indian Oceans, however, in the Atlantic Ocean, there are merely the RA, in Northeastern Brazil, and a few more in the Caribbean Sea. Rocas Atoll was the first acknowledged marine biological reserve in Brazil, granted in 1979, being a highly preserved marine environment [5]. Located 260 Km off the coast, RA is a unique environment, with a considerable biodiversity. For instance, Netto, et al. [6] described 95 taxa from the meiofauna and 79 macrofaunal taxa from the sediments of the tidal pools alone. A previous study from our research group described the occurrence of 12 ascidian species from RA, out of which 5 were recognized as new species, accenting the high degree endemism in this region [3].

Ascidians and sponges are filter-feeding invertebrates found in most marine environments, including Rocas Atoll. These organisms have yielded a wide variety of chemical structures, including antitumor compounds, with more than 6,000 natural products described so far [7-9]. However, the production of most of these metabolites was latter attributed to associated bacteria living in a syntrophic relationship with the invertebrate. These symbiotic bacteria produce different chemical structures, generally in response to interactions (antagonistic or beneficial) with their host [10, 11]. A high microbial diversity has been described from these holobionts, comprising predominantly proteobacteria, but also several other groups of bacteria and archaea [12-16].

Therefore, the goal of the present work is to assess culturable microorganisms from Rocas Atoll and analyze the cytotoxicity and the metabolomic profile of organic extracts produced by these bacteria, considering the enormous pharmacological potential described for marine bacteria. To accomplish this goal, the following proposal was designed: 1) to recover culturable bacteria from ascidians, sponges and sediments from Rocas Atoll; 2) to test the cytotoxicity of the crude extracts of cultured bacteria in colon carcinoma cells (HCT-116); and 3) to investigate the chemical content looking to contribute with chemical diversity characterization and to define possible relations with cytotoxicity properties.

2. Results and Discussion

2.1. Bacteria recovered from Rocas Atoll: cytotoxicity and identification

Considering the great pharmacological and biotechnological potential of the metabolites produced by marine bacteria, in this work we intended to prospect for the culturable bacteria from Rocas Atoll, focusing, largely, in actinomycete-like strains and their resulting metabolomic diversity. Nine ascidians and four sponges along with seven samples or marine sediments were collected, aiming at broadening the assortment of the assessed microbiota. Figure S1 shows the percentage of recovered bacteria from ascidians (39.5%), sponges (35.5%) and sediments (25%). The higher percentage of culturable bacteria was found in the invertebrates, especially from *Trididemnum maragoii*, an endemic ascidian from Rocas Atoll, and from the sponge *Chondrilla* cf. *nucula*, thus

Peer-reviewed version available at Mar. Drugs 2019, 17, 671; doi:10.3390/md17120671

highlighting these holobionts as rich microbial microenvironments. Table S1 describes the bacterial strains obtained, including their origin (ascidian, sponge or sediment), collection site and taxonomic identification, when attained.

A crude ethyl acetate extract obtained from each strain was tested for cytotoxicity against HCT-116 cells. Approximately 34% of the crude extracts were considered cytotoxic at 50 µg/mL (Figure 1a), while 10% were also active at 5 µg/mL (Figure 1b). These numbers are similar to those obtained from other studies performed with marine bacteria recovered from sediments collected at Atlantic islands [17, 18].

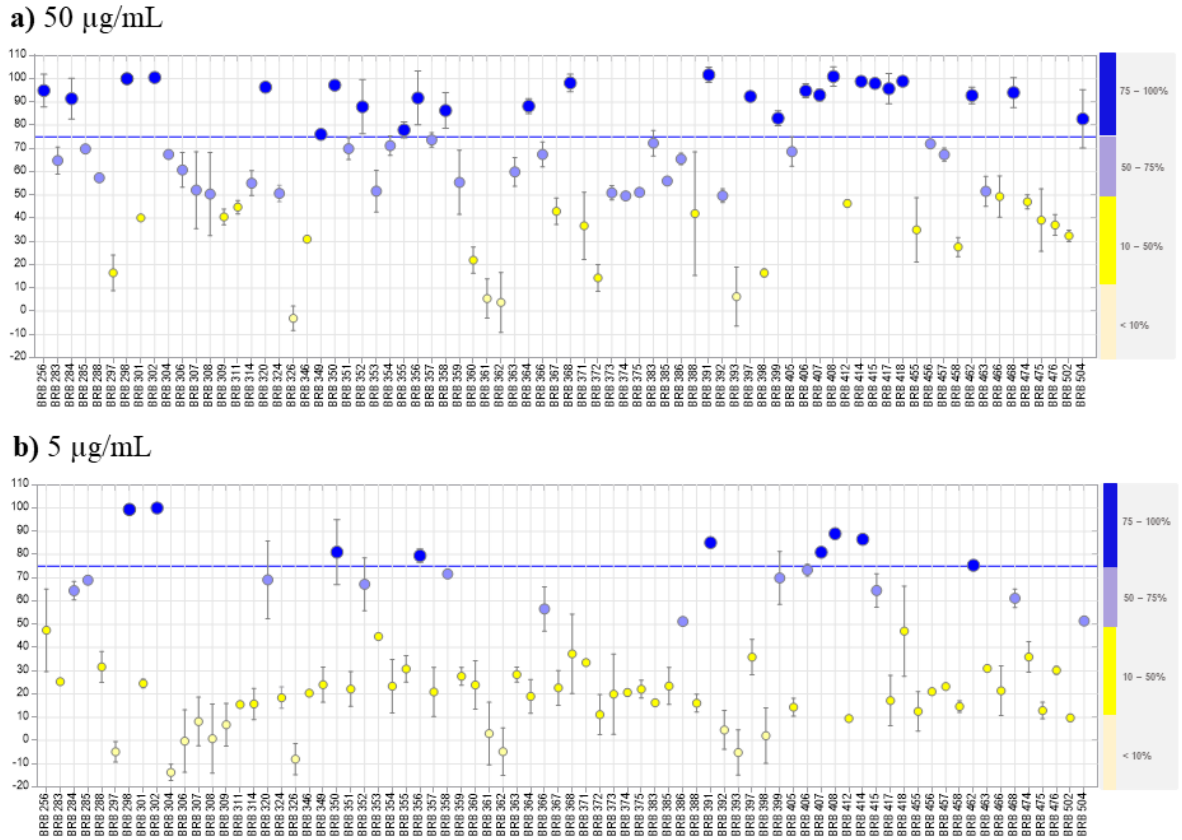

Figure 1. Cellular viability. Cytotoxicity of crude extracts produced by the recovered bacteria against HCT-116 carcinoma cell lines. Among 76 extracts tested, (a) 26 (34%) inhibited over 75% of cell growth at 50 µg/mL, and (b) 9 (12 %) at 5 µg/mL.

In a previous work that assessed culturable actinomycete isolated from sediments collected at the Saint Peter and Saint Paul's archipelago, 268 strains were isolated from 21 sediments samples, while 94 strains were selected for cytotoxicity analysis of their crude extract against HCT-116 cell lines [17]. Among them, 26 (27.6%) produced bioactive extracts, and the chemical analysis by LC-MS suggested the production of several renowned cytotoxic compounds by these strains, such as staurosporines and piericidins [17].

Herein, only strains that produced cytotoxic extracts were selected for taxonomic identification through 16S rRNA gene sequencing. The bacteria were found to belonged to two phyla: Actinobacteria and Firmicutes (Endobacteria) (Figure 2). Among the identified bacteria, Actinobacteria, comprised nearly 87%, being the dominant phylum among the producers of cytotoxic extracts. Within Actinobacteria, Micromonosporales and Streptomycetales were the most abundant taxa (40.0% and 35.0%, respectively, across all classified bacteria), followed by Streptosporangiales (20.0%) and Actinomycetales (5.0%). The remaining classified bacteria (13%) were included in the genus *Bacillus* (Firmicutes). It is important to bear in mind that the bacterial isolation procedures used

Peer-reviewed version available at Mar. Drugs 2019, 17, 671; doi:10.3390/md17120671

presently were, in fact, directed to favor Actinobacteria, guided by the phenotypic characteristics of the colonies, therefore, the strategy applied herein has demonstrated to be effective.

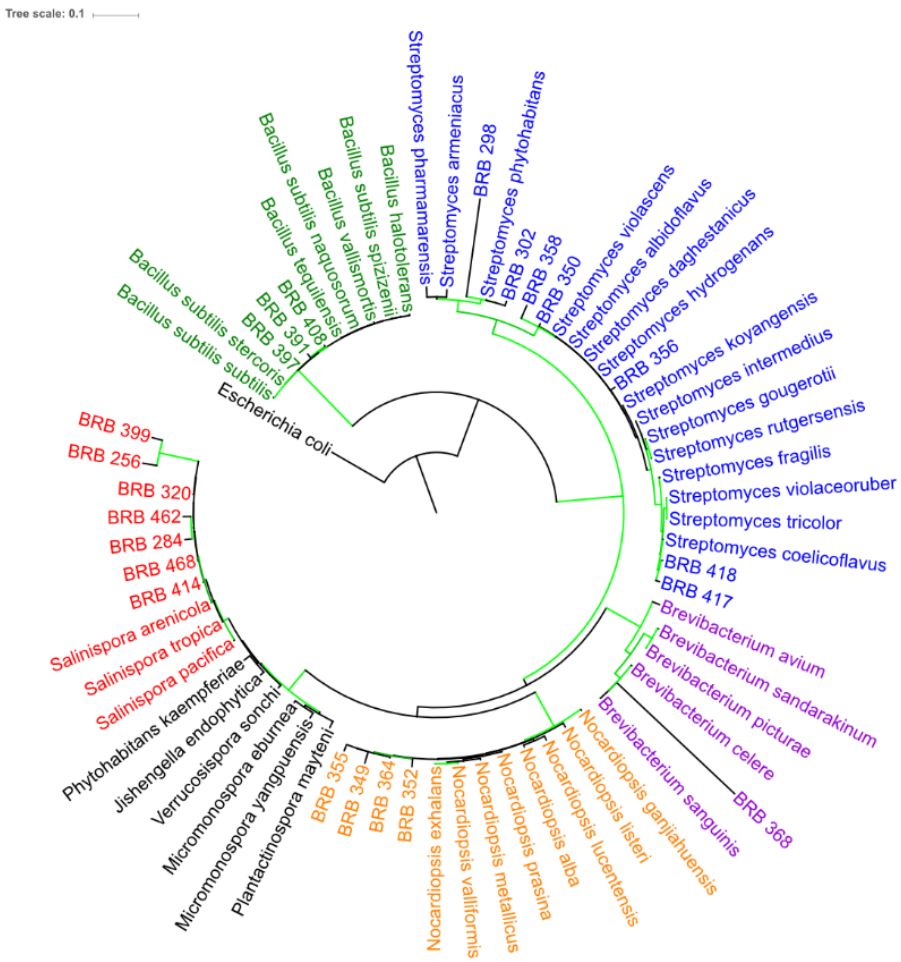

Figure 2. Phylogenetic relationships among recovered strains and most similar type species of bacteria obtained through Maximum Likelihood (RaxML) with 1000 bootstrap replications. Green lines indicate bootstrap support values >70% for the branch. Terminals are colored according to the genus of bacteria and follow the same pattern throughout this article. Accession numbers showed on table S1.

There are many studies describing the diversity of cultivable Actinobacteria from marine sediments [19-21], as well as those recovered from marine invertebrates, including ascidians and sponges [12, 22]. Indeed, all these previous studies pointed out the genera found in Rocas Atoll among the most common ones, especially in studies guided by cytotoxicity. The genus Streptomyces and the obligate marine genus *Salinispora* were greatly recurrent in the Rocas Atoll samples. The genus *Streptomyces* was recovered from all sample categories, and was also the most diverse in terms of numbers of distinct lineages (Figure 2). On the other hand, out of the eight *Salinispora* strains, six were recovered from ascidians, while the other two were associated with sponges. Interestingly, these strains were more related to each other than with the three known species of the genus. The presence of *Salinispora* was recently described in Brazilian waters along with the associated chemical diversity expected for this genus, including the production of highly cytotoxic compounds such as saliniketals and rifamycins, within strains isolated from sediments from Saint Peter and Saint Paul's Archipelago [23]. It is worth mentioning that *Streptomyces* bacteria are indigenous to most environments, and it

Peer-reviewed version available at Mar. Drugs 2019, 17, 671; doi:10.3390/md17120671

has been recognized as one of the most skilled genera in terms of secondary metabolism, producing numerous clinically used drugs [24, 25].

2.2. Chemical diversity produced by bacteria from Rocas Atoll

2.2.1. Molecular networking evaluation

Such great microbial diversity found in RA should be connected to an equivalent diversity of chemical structures. The dereplication process of natural crude extracts has been performed mostly by interpretation of UV, MS, MS/MS and NMR data, and in some cases, it is necessary to isolate the chemical components in order to obtain more specific data, which is laborious and time-consuming [26, 27]. Mass spectrometry (MS), in turn, is a powerful analytical technique capable of generating valuable structure information throughout gas phase reactions, thus accelerating the identification of compounds in a mixture [28, 29]. The coupling of liquid chromatography and MS techniques allowed the study of complex mixtures, such as extracts from natural sources, generating a large number of MS/MS. Also, the emergence of capable computational infrastructures allowed the use of reference libraries for compounds annotation [30]. Thus, for a large number of samples, such as the amount of extracts assessed here, the manual interpretation process becomes impractical and the use of computational support is a necessity. In this context, Molecular Networks from the GNPS platform has numerous advantages and have been widely applied in investigations of natural products [31-33]. This tool is able to compare each MS/MS spectrum from all data set and group them by spectral similarity, as similar chemical structures commonly present similar fragmentation pattern. In other words, it generates a network with clusters of parent ions (nodes) connected by edges/line. In addition, GNPS has also a library of MS/MS spectra that assists in the identification of known compounds present in the samples.

The crude extracts produced by the recovered bacteria were analyzed by HPLC-MS/MS in order to create a molecular network to compare samples and to investigate the chemical diversity by dereplication of known compounds. In this step, all the crude extracts were included in the analyses, both cytotoxic and non-cytotoxic. This approach was employed intending for a prompt identification of molecular families present exclusively in cytotoxic extracts, focusing on metabolites that may be the responsible for the biological activity. Besides, the higher the number of spectra, the greater the chance of annotation of known compounds. Metabolites that also occur in non-cytotoxic extracts might have decreased levels of cytotoxicity, or, even, be present at low concentrations in the crude extracts. The complete molecular network obtained for the bacterial extracts from RA is shown in Supplementary Figure S2 (after solvent blank removal) in which each node/sphere represents a parent ion of an MS/MS spectrum. The color code refers to bacteria identification according to the legend. All identified clusters of molecular family are amplified in Supplementary Figures S3-S11. To better visualize the data, Figure 3 highlights some clusters of molecular families that have at least one node annotated by the GNPS library. Due to some false-positive results that molecular network can generate, every spectrum from data set that matched some spectrum from GNPS library was manually checked.

Molecular network allowed the identification of many different chemical classes of compounds, such as diketopiperazines, lipopeptides (including surfactins and esperin), staurosporins, surugamides, sphinganines, erythromycins, TAN antibiotics, rifamycins, and the metal complexing agent (siderophore) desferrioxamine. These results pointed the bacteria from Rocas Atoll as a rich source of compounds with biotechnological and pharmacological interest.

Diketopiperazines (DKPs) (Supplementary Figure S3), which were detected in extracts from all examined strains, are low-molecular-weight compounds known to display a large range of biological activities, comprising immunosuppression, antibacterial, antifungal, nematicidal, insecticidal and cytotoxicity against many different types of cancer cell lines [34, 35], including a multidrug resistant colon carcinoma, in which a DKP derivative was described as reverting this phenotype [36]. Furthermore, DKPs have also been reported to be involved in chemical signaling/communication

Peer-reviewed version available at Mar. Drugs 2019, 17, 671; doi:10.3390/md17120671

between plants and plant-pathogenic bacteria [37], in which the production of these bacterial metabolites generate a range of functional responses in plants.

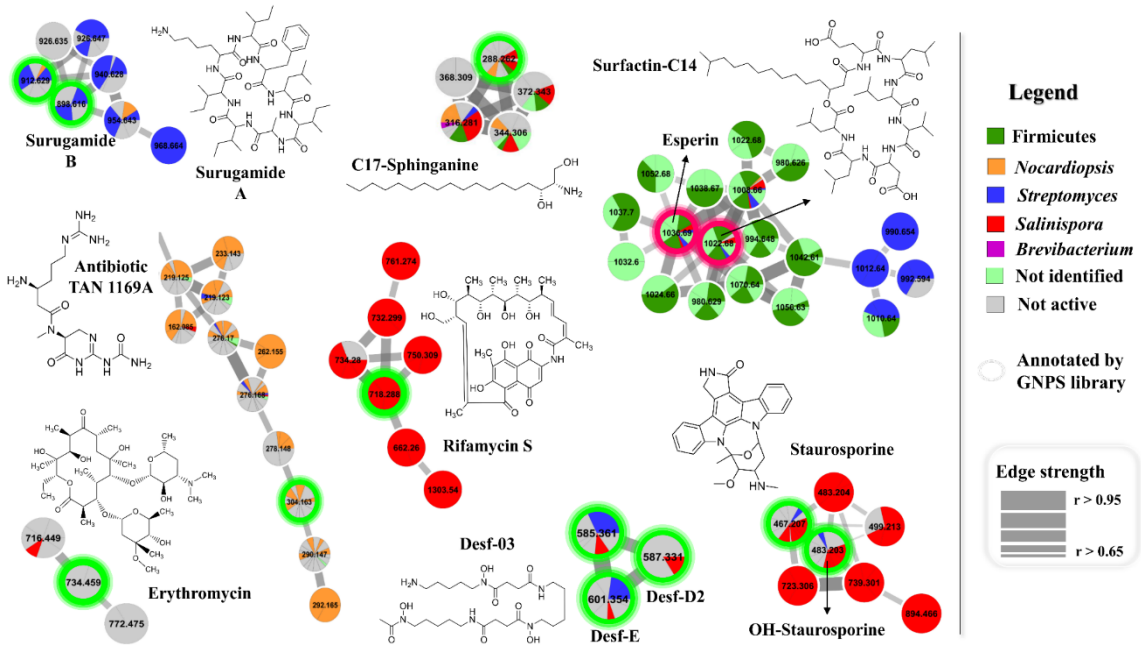

Figure 3. Molecular Network of extracts produced by bacteria recovered from Rocas Atoll, considering the positive ionization mode (ESI+) data after removal of solvent blank. Nodes represent parent masses, and their color are according to the legend, where the non-cytotoxic extracts (< 75%) are in gray. Extracts that showed inhibition of cell growth above 75% are divided by taxonomic group in Actinobacteria (*Streptomyces*, *Salinispora*, *Nocardiopsis* and *Brevibacterium*) and Firmicutes (*Bacillus*). Extracts considered cytotoxic produced by bacteria non-identified are depicted in light blue. Only clusters containing at least two nodes are shown.

Although the antimicrobial activity was not evaluated here, many of the chemical classes identified in this work are described in the literature with this biological property. For instance, rifamycin antibiotics, like rifampin and rifapentine, are antimycobacterial substances produced by actinobacteria, and have been specifically used in the effective treatment of tuberculosis [38]. Here, rifamycin S, hydroxyrifamycin S and dihydroxyrifamycin S, along with three unknown rifamycin analogs (*m/z* 662, 732 and 761) (Figure 3 and detailed in Supplementary Figure S4) were produced only by *Salinispora* strains,. On the other hand, lipopeptides, such as surfactins, are potent surfactant agents produced by members of *Bacillus* genus that present a wide range of applications as antibiotic agent due to the ability to penetrate cell membranes, and it has been applied against Gram-positive and Gran-negative bacteria, and fungi [39]. In this study, lipopeptides (Figure 3 and detailed in Supplementary Figure S5) were detected mainly in extracts produced by members of Bacilli, but also, in smaller proportion, in extracts from *Streptomyces*, *Salinispora* and *Nocardiopsis* strains. Interestingly, in the molecular cluster of surfactins, there are some ions (*m/z* 980, 1024, 1032, 1042, 1052 and 1056) referring to molecular masses not described for surfactin class, which shows that bacteria from RA produces unknown surfactin analogs. . Erythromycin is a macrolide widely employed as antibiotic agent to treat a number of bacterial infections [40]. It is usually produced by actinobacteria like *Streptomyces* and *Saccharopolyspora*. In this work, this chemical class (erythromycin A and 15-(2-Propynyl)erythromycin A) (Figure 3 and detailed in Supplementary Figure S6) was detected mainly in non-cytotoxic samples (BRB-407 and BRB-504), but also, one unknown analog (*m/z* 716) was found in one extract produced by *Salinispora* (BRB-415). The potent class of antibiotic TAN was also detected in our samples (Figure 3 and detailed in Supplementary Figure S7), produced mainly by *Nocardiopsis*

Peer-reviewed version available at Mar. Drugs 2019, 17, 671; doi:10.3390/md17120671

and some non-cytotoxic extracts, but derivatives were also produced by members of Firmicutes, *Streptomyces*, *Salinispora* and non-identified strains. Members of this dipeptide antibiotic family have presented activity against Gram-positive and Gram-negative bacteria, especially against drug-resistant strains [41].

Still regarding peptides, the detection of surugamides within the RA bacterial extracts should be highlighted (Figure 3 and detailed in Supplementary Figure S8). These cyclic octapeptides produced by marine *Streptomyces* were only recently described and are still poorly studied. In the generated molecular network, a cluster containing 7 members of surugamides were detected mainly in non-cytotoxic extracts and extracts produced by *Streptomyces* strains, but also by one *Nocardiopsis* (BRB-352). These compounds are described as inhibitors of cathepsin B, a cysteine peptidase overexpressed in many pathological events, such as inflammation and cancer. Cathepsin B inhibitors are promising compounds to be applied in cancer treatments [42]. To date, there are only 5 surugamides described in the literature. Our results indicate the presence of unknown surugamide analogs in the RA samples, thus suggesting that these bacteria are an interesting source of new chemicals to be further investigated.

Concerning the compounds previously described with cytotoxic activity, staurosporines (Figure 3 and detailed in Supplementary Figure S9) and saliniketals (as single node in Supplementary Figure S2) were also identified in this work. The putative staurosporine and derivatives have been extensively studied for years regarding their anticancer properties [43, 44]. Staurosporine and derivatives were previously detected by our research group and isolated from the ascidian *Eudistoma vannamei* collected on the west coast of Ceará state, Brazil [45, 46]. [45, 46]. This chemical class was detected mainly in extracts from *Salinispora* strains, but also in Streptomyces and non-cytotoxic strains. Moreover, saliniketals A and B, two unusual bicyclic polyketides, produced by the marine actinobacteria *Salinispora arenicola*, has shown inhibition of ornithine decarboxylase, a target for cancer chemoprevention therapies [47]. Here, saliniketal A was detected in most extracts produced by *Salinispora* strains. This chemical class was not present as a cluster in the molecular network, however these compounds were detected and annotated by the GNPS library.

2.2.2. Chemometric analysis

A total of 3,985 ions were detected, being 1,642 common to both active and non-active extracts, and 1,475 exclusives to active ones (Supplementary Figure S12 A). In general, there are more ions in the cytotoxic sample than non-cytotoxic, and, while the cytotoxic strains produced an average of 471 40 ions/strain, non-cytotoxic ones produced 224 24 putative metabolites, representing a rate of approximately 2:1 (Supplementary Figure S12 B). The Shannon-Wiener index (H') (Supplementary Figure S12 C) shows that the cytotoxic extracts (H' = 5.7 0.1) display a more diverse metabolomic content than non-cytotoxic extracts (H' = 4.9 ± 0.1).

In addition to what was observed through analysis of the molecular network, a relationship between the metabolomic profile and taxonomy became evident. The production of molecular families, such as lipopeptides, can be clearly associated with *Bacillus*. These findings are somewhat expected, since differences in the genomes lead to the expression of different enzymes and biosynthetic pathways, which in turn generate distinct chemical structures. In the heatmap (Figure 4), there are clearly associations of some chemical groups with a particular genus, such as the antibiotic TAN and other unknown molecular family that were detected in all *Nocardiopsis* strains; as well as rifamycin, staurosporine and an unknown molecular family that were detected in most of *Salinispora* strains.

Peer-reviewed version available at Mar. Drugs 2019, 17, 671; doi:10.3390/md17120671

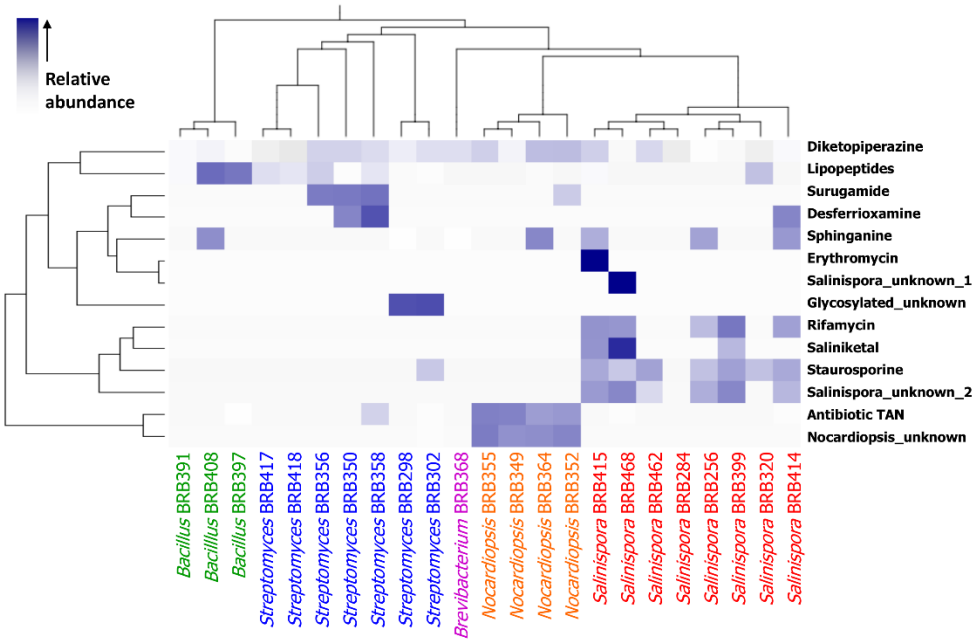

Figure 4. Heatmap with metabolites classes distribution on the bacteria phylogenetic tree. It was considered only the bacteria that produced cytotoxic samples. The chemical classes used here were obtained from molecular network.

2.3. Molecular networking vs Cytotoxicity

The molecular network analysis was further applied expecting to highlight clusters of ions more likely to be accountable for the biological activity of the extracts, as those present in both cytotoxic and non-cytotoxic samples will hardly be bioactive metabolites. Such strategy allowed for directing efforts towards some few clusters, decreasing the number of possibilities and pinpointing the most interesting groups of compounds to be investigated. Among the identified molecular families are rifamycin, lipopeptides and staurosporine. Besides those, one large cluster stood out and is circled in light blue in Supplementary Figure S2. It (Supplementary Figure S13) includes compounds produced exclusively by Streptomyces sp., more specifically, the strains BRB-298 and BRB-302. This cluster comprises parent masses in the range of 800 to 1400 m/z, approximately. The MS and MS/MS data showed these ions correspond to glycosylated structures. Moreover, an in-source fragmentation was observed, generating many parent masses for the same compound, which is characteristic of compounds with many glycosyl moieties. The MS/MS spectra of the ions from this cluster found no matches in the GNPS library or even other public libraries that were searched, such as "Dictionary of Natural Products" and "SciFinder", indicating a high probability of being new compounds. Therefore, BRB-302 was selected to be cultured in large scale and fractionated in order to investigate the cytotoxicity of the fractions and to check if the compounds highlighted here in the molecular network are indeed the active compounds in the crude extract.

2.4. Identification of glycosylated compounds produced by BRB-302

The approach presented in this work allowed selection of one crude extract among 76, and further, point out the active compounds in the crude extract without the need of isolation. Baring this in mind, the crude extract was fractionated and the collected fractions were evaluated through cytotoxic assay. It was observed a high activity in fractions 28-32 (Figure S13), however these fractions have low absorption in 280 nm. Thus, to continue the analysis, these fractions were reanalyzed using GNPS platform. The MS spectra for these compounds were not found in the spectral library of GNPS. Therefore, in order to identify at least the chemical class of those unknown compounds produced by the *Streptomyces* sp. BRB-302 without isolating them, we decided to continue exploring our data employing other metabolomics tools available at GNPS. Molecular network is a tool at GNPS based

Peer-reviewed version available at Mar. Drugs 2019, 17, 671; doi:10.3390/md17120671

on a spectral library of known compounds; however, GNPS also has other tools based on in silico fragmentation that helps on identification of known compounds that do not present available MS/MS spectra, or even unknown compounds. In general, predicting tools have been successfully applied on the identification of several challenging structure from different chemical classes [48, 49]. DEREPLICATOR+ is an algorithm able to construct theoretical MS spectra of compounds from natural products, including peptides, polyketides, flavonoids, terpenes, alkaloids, among other classes of natural products, through cleaving the bridges and measuring the masses of such components [50]. Although it has been only recently introduced, in silico computational tools for MS data has been developed and applied for more than a decade, and the developers of DEREPLICATOR+ have described over 70% of right identification.

Therefore, we analyzed the MS/MS data of the extract produced by the *Streptomyces* sp. BRB-302 by DEREPLICATOR+. The tool was able to indicate a structure for only one component of the entire cluster, the novonestmycin A (molecular mass 1,228.6605 Da) [51], which was designated for the ion m/z 1,229.6635 ([M+H]+, error 3.4 ppm). UV(MeOH) λ_{max} described for novonestmycins are 222, 260 and 290 nm, while the observed for the compound produced by *Streptomyces* sp. BRB-302 are 222, 261 and 290 nm (Figure 5). Moreover, novonestmycins are produced by *Streptomyces phytohabitans*, and both *Streptomyces* sp. BRB-298 and BRB-302 strains are on the same branch of *S. phytohabitans* in the phylogenetic tree (Figure 2). Furthermore, novonestmycins were reported as high cytotoxicity against tumor cell lines, as well as the compounds produced by *Streptomyces* sp. BRB-302. Checking the fragmentation spectrum of m/z 1,229.6635 it is possible to observe the neutral loss of 280 a.m.u. which refers to the loss of the phenolic glycoside moiety, followed by successive water losses, which is pattern for polyketide macrolide chains [52, 53]. Therefore, as our data presents several parent masses with different masses than described in literature for novonestmycins, we can assume that *Streptomyces* sp. BRA-302 and BRB-298 produce several novel derivatives of the novonestmycins family, which will be the aims of future studies concerning isolation, structure characterization and pharmacological assays. Here, we focused on the chemical richness produced by the recovered bacteria from Rocas Atoll. Furthermore, metabolomics approaches allowed to point out the possible chemical class responsible for the biological activity observed for *Streptomyces* sp. BRB-302 and BRB-298 strains, which is a motivating strategy to be employed in screening of natural crude extracts.

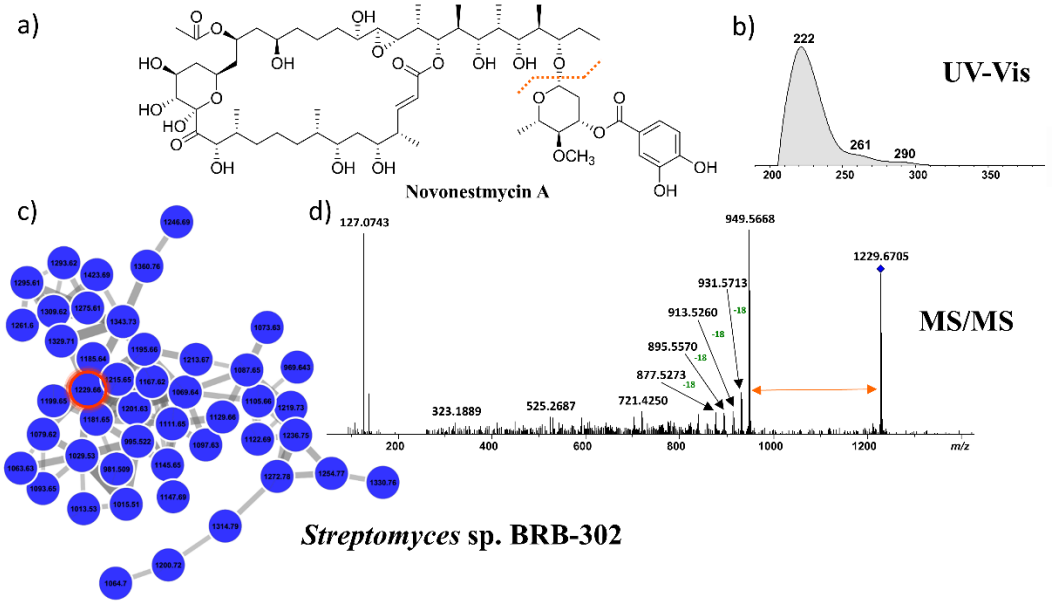

Figure 5. (a) Molecular structure of novonestmycin A; (b) UV spectrum referring to the novonestmycin derivative produced by the *Streptomyces* sp. BRB-302; (c) Cluster of an active molecular family produced exclusively by *Streptomyces* sp. BRB-298 and BRB-302 strains in which the parent ion circled in red is the annotated by DEREPLICATOR+; (d) MS/MS spectrum of the ion *m/z* 1229 annotated as novonestmycin derivative.

Peer-reviewed version available at Mar. Drugs 2019, 17, 671; doi:10.3390/md17120671

3. Materials and Methods

3.1 Sample Collection

Different specimens of ascidians and sponges and also samples of sediments were manually collected at Rocas Atoll (03°51'00"S and 33°49'00"W). The samples were immediately frozen in liquid nitrogen for transportation and kept in the laboratory at – 20oC until they were processed. The project was developed under permission from the Instituto Chico Mendes de Conservação da Biodiversidade (SISBIO # 44435-1), as well as from the Conselho Gestor do Patrimônio Genético (SISGEN # 010170/2015-4).

3.2 Bacteria Isolation and Cultivation

Under sterile conditions, invertebrate specimens were, initially, sorted for epibionts and debris then dipped in ethanol 70% for decontamination of the surface. Next, all samples assessed (ascidians, sponges and sediments) were diluted in sterile seawater (1:1 w/v), heated to 55 °C during 10 min and inoculated onto 90 mm Petri dishes filled with one of three types of medium: a nutrient-rich medium (A1); trace metals agar (TMA), a nutrient-poor medium with trace metals; and seawater agar (SWA), a minimum medium. All media were prepared with reconstituted seawater and added with Cycloheximide (0.1 mg/mL) to reduce fungal contamination. The use of these three different culture media was important to recover bacterial strains with different metabolic characteristics [54].

Sediment-inoculated dishes were left between 12 – 90 days at 26-28 °C to allow for bacterial growth. Distinguishable colonies were selected based on their phenotypic characteristics, including color, brightness, shape and texture, then transferred to fresh A1 (prepared in the laboratory with soluble starch, yeast extract and peptone) agar dishes using a sterile toothpick, for isolation and purification. Pure strains were further grown in liquid media, supplemented with 25% glycerol, codified and distributed into cryovials for preservation at -80 °C.

3.3 Crude extract production

Bacterial strains were grown in Erlenmeyer flasks (250 mL) containing 50 mL of A1 liquid medium. The cultures were maintained at 28 °C under 170 rpm between 3 and 25 days, depending on proliferation rate of each strain under the offered conditions. The cultures were extracted with ethyl acetate (50 mL) during 2 h and the organic fraction was collected. Solvent was then removed under reduced pressure (rotary evaporator) to yield the respective crude extracts.

3.4 Cultivation and Fractionation of BRB-302 Strain

Cryovials of strain BRB-302 were inoculated into two Erlenmeyers (2 L) containing 250 mL of A1 medium and cultivated for 7 days at 180 rpm and 28 °C. Cultivated broths were extracted twice with 200 mL of ethyl acetate, shaken at 180 rpm during 1 h. Crude extracts were combined (55 mg), diluted to a concentration of 50 mg/mL and fractionated through analytical HPLC-PDA.

Chromatographic separation was carried out through a chromatograph Agilent 1200 containing a fraction collector and using a Zorbax® Eclipse Plus C18 4.6x150 mm, 3.5 µm column at 45 °C. The method consisted of solvents A (0.1% acetic acid in H_2O) and B (MeOH), starting at 5% up to 100% of B in 30 min followed by a hold of 100% of B for 5 min and conditioning with 5% of B for 7 min. Fractions were collected at 0.75 min intervals over 30 min from the beginning of the chromatographic run, and one at the end of the fractionation (100% MeOH), to a total of 41 fractions collected.

3.5 Cytotoxic Assay

The crude extracts and fractions were dissolved in DMSO and evaluated regarding their cytotoxicity at two different concentrations, 5 and 50 µg/mL, against a colon carcinoma cell line (HCT-116 ATCC CCL-247), using the MTT assay [55]. Briefly, cells cultured in RPMI media supplemented with 10% SBF were plated into 96 well plates (10,000 cells per well), left to adhere during 24 h and exposed to test samples, in duplicates, during 72 h. Doxorubicin and DMSO were used as positive

Peer-reviewed version available at Mar. Drugs 2019, 17, 671; doi:10.3390/md17120671

and negative controls, respectively. After this period, wasted media was substituted with fresh media containing MTT at 0.5 mg/mL. After 3 h, media was removed, reduced MTT (formazan) was dissolved in DMSO and absorbance of each well was determined at 570 nm. Single well data were transformed to percentage of growth inhibition after normalization with positive (100% of inhibition) and negative (0% of inhibition) controls. Extracts were considered cytotoxic when inhibited over 75% cell growth at 50 μg/mL.

3.6 DNA Extraction, 16S rRNA Amplification and Sequencing

Strains producing cytotoxic crude extracts were selected for taxonomic identification. Firstly, a phenotypic analysis, including the morphology of bacterial colony and a Gram test, was performed. Then, the genomic DNA from the bacterial culture was isolated, following the Wizard® Genomic DNA Purification Kit, with some modifications – such as the use of Proteinase K, a higher concentration of RNase and a greater incubation period during the extraction step [54, 56]. Partial amplification of the 16S rRNA gene was performed with Promega's GoTaq Green Master Mix, and universal primers for eubacteria 27F (5'-AGAGTTTGATCCTGGCTCAG-3') and 1492R (5'-TACGGCTACCTTGTTACGACTT-3'). PCR products were purified with a purification kit (QIAGEN Inc., Valencia, CA) and sequenced using the ABI 3730 DNA Analyzer, a 48-capillary DNA analysis system with Life Technologies technology - Applied Biosystems. Sequencing reactions were performed using the BigDye® Terminator v3.1 Cycle Sequencing Kit (code 4337456) and runs carried out in capillaries (36 cm) using the POP7 polymer. Sequences were analyzed using ChromasPro software (version 2.0, TECHNELSIUM DNA Sequencing Software, Australia - http://technelysium.com.au/wp/chromaspro/), compared to the sequences present in the Ezbiocloud database for type species and also at NCBI (http://www.ncbi.nlm.nih.gov/) using Basic Local Alignment Search Tool (BLAST).

3.7 Phylogenetic analysis

The forward and reverse sequences from the isolated bacteria strains were aligned and the consensus sequences were obtained using Geneious 7 software (Biomatters Ltd). All the sequences reported in this study were deposited in GenBank under accession numbers MK720152 through MK720174 (Table S1) (http://www.ncbi.nlm.nih.gov/genbank). Sequences were compared to the Ez-Taxon database (http://www.ezbiocloud.net), and most similar sequences from type species were included in the analysis. The full set of sequences was then aligned with MAFFT version 7 using the L-INS-i strategy [57]. For the phylogenetic reconstruction, we used the GTR+G nucleotide substitution model. The phylogenetic relationships were inferred through a maximum likelihood approach using RaxML [58] with 1000 bootstrap iterations. The resulting tree was exported and edited using the iTOL v4 online tool [59, 60].

3.8 Metabolomic fingerprint by HPLC-MS/MS

For metabolomic evaluation, all crude extracts and fractions were diluted in methanol at 1.0 mg/mL and analyzed by HPLC-MS/MS. The analyses were developed on a HPLC system (Shimadzu) coupled to a micrOTOF QII QqTOF mass spectrometer (Bruker Daltonics) fitted with an electrospray ionization source, operating in positive ionization mode. The chromatographic separation occurred on a Supelco Ascentis Express C18 (5 μm, 150 x 3.0 mm) column, using water (phase A) and acetonitrile (phase B) as the mobile phase, both containing 0.1% formic acid. It was employed a gradient from 5% to 100% phase B over 25 min followed by 100% phase B for 6 min, with a flow rate of 0.7 mL/min. The column temperature was set to 40 °C. Ion source: ESI, endplate: 4,500 V, capillary voltage: 3,500 V, dry temp.: 220 °C, dry gas: 9.0 L/min, *m/z* range: 50-1,500, gas pressure: 40 psi. The injection volume was 15 μL. An untargeted method was used in the mass spectrometer, in which the analyzer selects the higher intensity ions to fragment. A ramp from 20 to 75 eV as collision energy was employed.

3.9 Molecular Networking

Peer-reviewed version available at Mar. Drugs 2019, 17, 671; doi:10.3390/md17120671

All HPLC MS/MS data were converted to mzXML file format by using MSConvert, and thereafter, uploaded to the Global Natural Products Social (GNPS) molecular networking server (gnps.ucsd.edu) [31]. On the GNPS platform, MS/MS spectra were combined with MSCluster algorithm considering cosine similarity values (higher than 0.95) to created consensus MS2 spectra. For spectral networks, parent mass and fragment ion tolerance of 0.08 Da and 0.1 Da were considered, respectively. For creating edges, a cosine score over than 0.65, more than four matched peaks, and two nodes at least in the top 10 cosine scores (K parameter) was fitted. The spectra were also searched against GNPS's spectral libraries, considering score above 0.65 and at least four matched peaks. Data are publicly available at http://gnps.ucsd.edu under accession number MSV000083601. The generated molecular network was imported and visualized as nodes and edges into Cytoscape versions 2.8.2 and 3.4.0 [61]. Nodes represent parent masses and edge thickness corresponds to cosine score between two nodes.

Data used to build the molecular networking was recovered as .csv table and analyzed to obtain the number and abundance of precursor ions among cytotoxic and non-cytotoxic extracts. The Shannon-Wiener diversity index was calculated based on the molecular abundance of the precursor ions using R software with the 'vegan' package. To verify the significance of observed value, data from cytotoxic and non-cytotoxic extracts were compared with Student's t test with a level of significance of 5% using GraphPad Prism.

4. Conclusions

Overall, Rocas Atoll emerge as an invaluable source of natural products with interesting pharmacological and biotechnological properties. The approach employed herein was efficient in allowing for the detection and identification of many different chemical classes produced by the bacteria. Particularly, such breakthrough was made possible due to the versatility of mass spectrometry, which generated a large amount of data on the crude extracts, including important structural information on organic compounds therein, which accelerated their identification in complex samples. Additionally, the GNPS platform, a prodigious novelty in Natural Products Research, fast-tracked dereplication of known and unknown compounds and, furthermore, highlighted chemical groups possibly responsible for the biological activity observed for the crude extracts. Therefore, we conclude that Brazilian islands host a treasured diversity in natural products, and the inhabitant species should be protected against anthropic impacts to prevent ecosystem degradation and the loss of such valuable resources.

Supplementary Materials: The following are available online, Figure S1: Origin of the 80 bacterial strains recovered from the Atlantic Rocas Atoll. (A) Bacteria recovered from ascidians (40%), sponges (35%) and sediments (25%); (B) Identification of ascidians, sponge and sites where sediments were collected. Figure S2: Molecular Network of crude extracts produced by bacteria recovered from Rocas Atol, using positive ionization mode (ESI+) data. Node colors represent the strains extracts accordingly to the legend. Nodes highlighted in black boxes represent parent ions that are identified by GNPS library, and blue boxes represent cluster of ions that are present only in active crude extracts. Figure S3: Cluster of the diketopiperazine family, observed as protonated adducts ([M+H]+), produced by all investigated strain. The node color represents the genus of the bacteria, and edge size is according to the cosine score. Figure S4: Cluster of the rifamycin family, observed as protonated adducts ([M+H]+), produced mainly by *Salinispora* (BRB-256, BRB-399, BRB-414, BRB-415 and BRB-468), but also by BRB-283, a not active strain. The node color represents the genus of the bacteria, and edge size is according to the cosine score. Figure S5: Cluster of lipopeptide family, including surfactins and esperin, observed as protonated adducts ([M+H]+), produced by *Bacillus* (Firmicutes) (BRB-391, BRB-397 and BRB-408), *Nocardiopsis* (BRB-364), *Streptomyces* (BRB-298, BRB-302, BRB-350, BRB-356, BRB-358, BRB-417 and BRB-418), *Salinispora* (BRB-320 and BRB-415) not identified (BRB-406, BRB-407 and BRB-504) and not active (BRB-346 and BRB-359) strains. The node color represents the genus of the bacteria, and edge size is according to the cosine score. Figure S6: Cluster of the erythromycin family, observed as protonated adducts ([M+H]+), produced mainly by not identified (BRB-346, BRB-374, BRB-375, BRB-405, BRB-455 and BRB-463) strains, but also by one *Salinispora* (BRA-415) strain. The node color represents the genus of the bacteria, and edge size is according to the cosine score. Figure S7: Cluster of the antibiotic TAN family, observed as protonated adducts ([M+H]+), produced mainly by *Nocardiopsis* (BRB-349, BRB-352, BRB-355 and BRB-364), but also Firmicutes (BRB-397), *Streptomyces*

Peer-reviewed version available at Mar. Drugs 2019, 17, 671; doi:10.3390/md17120671

(BRB-302 and BRB-358), *Salinispora* (BRB-414 and BRB-468), not identified (BRB-407) and not active (BRB-351, BRB-354, BRB-360, BRB-361, BRB-371, BRB-372, BRB-373, BRB-383, BRB-385, BRB-388, BRB-393 and BRB-455) strains. The node color represents the genus of the bacteria, and edge size is according to the cosine score. Figure S8: Cluster of the surugamide family, observed as protonated adducts ([M+H]+), produced mainly by *Streptomyces* strains (BRB-350, BRB356 and BRB-358), *Nocardiopsis* (BRB-352) and not active (BRB-353, BRB-366, BRB-375 and BRB-386). The node color represents the genus of the bacteria, and edge size is according to the cosine score. Figure S9: Cluster of the staurosporine family, observed as protonated adducts ([M+H]+), produced mainly by *Salinispora* (BRB-256, BRB-320, BRB-399, BRB-414, BRB-415, BRB-462 and BRB-468), *Streptomyces* (BRB-302) not active (BRB-285, BRB-304, BRB-306, BRB-307, BRB-308, BRB-309, BRB-324 and BRB-405) strains. The node color represents the genus of the bacteria, and edge size is according to the cosine score. Figure S10: Cluster of the desferrioxamine family, observed as protonated adducts ([M+H]+), produced mainly by not active (BRB-367, BRB-375, BRB-386, BRB-412, BRB-463 and BRB-476) strains, but also by *Streptomyces* (BRB-350 and BRB-358) and *Salinispora* (BRB-414). The node color represents the genus of the bacteria, and edge size is according to the cosine score. Figure S11: Cluster of the sphinganine family, observed as protonated adducts ([M+H]+), produced by many different strains, including *Bacillus* (Firmicutes) (BRB-408), *Brevibacterium* (BRB-368), *Nocardiopsis* (BRB-364), *Salinispora* (BRB-256, BRB-414 and BRB-415), *Streptomyces* (BRB-298), not identified (BRB-406 and BRB-504) and not active (BRB-297, BRB-304, BRB-306, BRB-307, BRB-308, BRB-309, BRB-346, BRB-374, BRB-398, BRB-405 and BRB-463). The node color represents the genus of the bacteria, and edge size is according to the cosine score. Figure S12: (A) Venn diagram and (B) distributions of samples according to the number of ions detected in cytotoxic and non-cytotoxic extracts; (C) Shannon-Wiener index (D). For these analyses, the .csv table obtained from the molecular network was considered. Table S1: Bacteria strains isolated from ascidians, sponges and sediments from the Rocas Atoll and their anticancer activity results against human colon carcinoma cell line HCT-116.

Author Contributions: All authors listed have made substantial, direct and intellectual contribution to the work, and approved it for publication.

Funding: This research was funded by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) grant number 2015/17177-6.

Acknowledgments: The authors are grateful to CAPES, CNPq and FAPESP (2017/17648-4 (AB); 2017/16606-6 (MMPT)) for the fellowships. The authors are further thankful to Secretaria da Comissão Interministerial dos Recursos do Mar (SECIRM) for providing all the logistic support disposed to the scientific expedition, as part of the Program PROARQUIPELAGO. TMCL, MJPF, NPL and LVCL are CNPq research fellows.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Blunt, J. W.; Carroll, A. R.; Copp, B. R.; Davis, R. A.; Keyzers, R. A.; Prinsep, M. R., Marine natural products. *Nat. Prod. Rep.* **2018,** 35, (1), 8-53.
2. Soldatou, S.; Baker, B. J., Cold-water marine natural products, 2006 to 2016. *Nat. Prod. Rep.* **2017,** 34, (6), 585-626.
3. Paiva, S. V.; Oliveira Filho, R. R. d.; Lotufo, T. M., Ascidians from Rocas Atoll, northeast Brazil. *Front. Mar. Sci.* **2015,** 2, (39), 1-20.
4. Azevedo, F.; Padua, A.; Moraes, F.; Rossi, A.; Muricy, G.; Klautau, M., Taxonomy and phylogeny of calcareous sponges (Porifera: Calcarea: Calcinea) from Brazilian mid-shelf and oceanic islands. *Zootaxa* **2017,** 4311, (3), 301-344.
5. Coimbra, J. C.; Carreño, A., Richness and palaeo-zoogeographical significance of the benthic Ostracoda (Crustacea) from the oceanic Island of Trindade and Rocas Atoll, Brazil. *Rev. Bras. Paleontolog* **2012,** 15, 189-202.

Peer-reviewed version available at Mar. Drugs 2019, 17, 671; doi:10.3390/md17120671

6. Netto, S. A.; Attrill, M. J.; Warwick, R. M., The relationship between benthic fauna, carbonate sediments and reef morphology in reef-flat tidal pools of Rocas Atoll (north-east Brazil). *J Mar Biol Assoc Uk* **2003,** 83, (2), 425-432.
7. Schmidt, E. W.; Donia, M. S., Life in cellulose houses: symbiotic bacterial biosynthesis of ascidian drugs and drug leads. *Curr Opin Biotech* **2010,** 21, (6), 827-833.
8. Palanisamy, S. K.; Rajendran, N. M.; Marino, A., Natural products diversity of marine ascidians (Tunicates; Ascidiacea) and successful drugs in clinical development. *Natural products and bioprospecting* **2017,** 7, (1), 1-111.
9. Mehbub, M. F.; Lei, J.; Franco, C.; Zhang, W., Marine sponge derived natural products between 2001 and 2010: trends and opportunities for discovery of bioactives. *Marine drugs* **2014,** 12, (8), 4539-4577.
10. Molloy, E. M.; Hertweck, C., Antimicrobial discovery inspired by ecological interactions. *Current opinion in microbiology* **2017,** 39, 121-127.
11. Schmidt, E. W., The secret to a successful relationship: lasting chemistry between ascidians and their symbiotic bacteria. *Invertebrate biology : a quarterly journal of the American Microscopical Society and the Division of Invertebrate Zoology/ASZ* **2015,** 134, (1), 88-102.
12. Steinert, G.; Taylor, M. W.; Schupp, P. J., Diversity of actinobacteria associated with the marine ascidian *Eudistoma toealensis*. *Mar. Biotechnol.* **2015,** 17, (4), 377-385.
13. Evans, J. S.; Erwin, P. M.; Shenkar, N.; Lopez-Legentil, S., Introduced ascidians harbor highly diverse and host-specific symbiotic microbial assemblages. *Scientific reports* **2017,** 7, (1), 11033-11043.
14. Pita, L.; Rix, L.; Slaby, B. M.; Franke, A.; Hentschel, U., The sponge holobiont in a changing ocean: from microbes to ecosystems. *Microbiome* **2018,** 6, (46), 2-18.
15. Bauermeister, A.; Branco, P. C.; Furtado, L. C.; Jimenez, P. C.; Costa-Lotufo, L. V.; da Cruz Lotufo, T. M., Tunicates: A model organism to investigate the effects of associated-microbiota on the production of pharmaceuticals. *Drug Discovery Today: Disease Models* **2019**.
16. Chen, L.; Hu, J. S.; Xu, J. L.; Shao, C. L.; Wang, G. Y., Biological and chemical diversity of ascidian-associated microorganisms. *Marine drugs* **2018,** 16, (10).
17. Ferreira, E. G.; Torres, M. D. M.; da Silva, A. B.; Colares, L. L. F.; Pires, K.; Lotufo, T. M. C.; Silveira, E. R.; Pessoa, O. D. L.; Costa-Lotufo, L. V.; Jimenez, P. C., Prospecting anticancer compounds in actinomycetes recovered from the sediments of Saint Peter and Saint Paul's Archipelago, Brazil. *Chem Biodivers* **2016,** 13, (9), 1149-1157.
18. Prieto-Davo, A.; Dias, T.; Gomes, S. E.; Rodrigues, S.; Parera-Valadezl, Y.; Borralho, P. M.; Pereira, F.; Rodrigues, C. M. P.; Santos-Sanches, I.; Gaudencio, S. P., The Madeira Archipelago as a significant source of marine-derived actinomycete diversity with anticancer and antimicrobial potential. *Front Microbiol* **2016,** 7.
19. Maldonado, L. A.; Stach, J. E. M.; Pathom-aree, W.; Ward, A. C.; Bull, A. T.; Goodfellow, M., Diversity of cultivable actinobacteria in geographically widespread marine sediments. *Anton Leeuw Int J G* **2005,** 87, (1), 11-18.
20. Maldonado, L. A.; Fragoso-Yanez, D.; Perez-Garcia, A.; Rosellon-Druker, J.; Quintana, E., Actinobacterial diversity from marine sediments collected in Mexico. *Anton Leeuw Int J G* **2009,** 95, (2), 111-120.
21. Becerril-Espinosa, A.; Freel, K. C.; Jensen, P. R.; Soria-Mercado, I. E., Marine Actinobacteria from the Gulf of California: diversity, abundance and secondary metabolite biosynthetic potential. *Anton Leeuw Int J G* **2013,** 103, (4), 809-819.

Peer-reviewed version available at Mar. Drugs 2019, 17, 671; doi:10.3390/md17120671

22. Lackner, G.; Peters, E. E.; Helfrich, E. J. N.; Piel, J., Insights into the lifestyle of uncultured bacterial natural product factories associated with marine sponges. *P Natl Acad Sci USA* **2017,** 114, (3), E347-E356.
23. Bauermeister, A.; Velasco-Alzate, K.; Dias, T.; Macedo, H.; Ferreira, E. G.; Jimenez, P. C.; Lotufo, T. M. C.; Lopes, N. P.; Gaudencio, S. P.; Costa-Lotufo, L. V., Metabolomic fingerprinting of *Salinispora* from Atlantic oceanic islands. *Front Microbiol* **2018,** 9, 3021.
24. Kemung, H. M.; Tan, L. T. H.; Khan, T. M.; Chan, K. G.; Pusparajah, P.; Goh, B. H.; Lee, L. H., *Streptomyces* as a prominent resource of future anti-MRSA drugs. *Front Microbiol* **2018,** 9, 1-26.
25. Ser, H. L.; Tan, L. T. H.; Law, J. W. F.; Chan, K. G.; Duangjai, A.; Saokaew, S.; Pusparajah, P.; Ab Mutalib, N. S.; Khan, T. M.; Goh, B. H.; Lee, L. H., Focused review: cytotoxic and antioxidant potentials of mangrove-derived *Streptomyces*. *Front Microbiol* **2017,** 8, 1-11.
26. Gaudencio, S. P.; Pereira, F., Dereplication: racing to speed up the natural products discovery process. *Nat. Prod. Rep.* **2015,** 32, (6), 779-810.
27. Brunetti, A. E.; Carnevale Neto, F.; Vera, M. C.; Taboada, C.; Pavarini, D. P.; Bauermeister, A.; Lopes, N. P., An integrative omics perspective for the analysis of chemical signals in ecological interactions. *Chemical Society reviews* **2018,** 47, (5), 1574-1591.
28. Bauermeister, A.; Zucchi, T. D.; Moraes, L. A., Mass spectrometric approaches for the identification of anthracycline analogs produced by actinobacteria. *Journal of mass spectrometry : JMS* **2016,** 51, (6), 437-45.
29. Demarque, D. P.; Crotti, A. E.; Vessecchi, R.; Lopes, J. L.; Lopes, N. P., Fragmentation reactions using electrospray ionization mass spectrometry: an important tool for the structural elucidation and characterization of synthetic and natural products. *Nat. Prod. Rep.* **2016,** 33, (3), 432-55.
30. Aksenov, A. A.; da Silva, R.; Knight, R.; Lopes, N. P.; Dorrestein, P. C., Global chemical analysis of biology by mass spectrometry. *Nat Rev Chem* **2017,** 1, (7).
31. Wang, M.; Carver, J. J.; Phelan, V. V.; Sanchez, L. M.; Garg, N.; Peng, Y.; Nguyen, D. D.; Watrous, J.; Kapono, C. A.; Luzzatto-Knaan, T.; Porto, C.; Bouslimani, A.; Melnik, A. V.; Meehan, M. J.; Liu, W. T.; Crusemann, M.; Boudreau, P. D.; Esquenazi, E.; Sandoval-Calderon, M.; Kersten, R. D.; Pace, L. A.; Quinn, R. A.; Duncan, K. R.; Hsu, C. C.; Floros, D. J.; Gavilan, R. G.; Kleigrewe, K.; Northen, T.; Dutton, R. J.; Parrot, D.; Carlson, E. E.; Aigle, B.; Michelsen, C. F.; Jelsbak, L.; Sohlenkamp, C.; Pevzner, P.; Edlund, A.; McLean, J.; Piel, J.; Murphy, B. T.; Gerwick, L.; Liaw, C. C.; Yang, Y. L.; Humpf, H. U.; Maansson, M.; Keyzers, R. A.; Sims, A. C.; Johnson, A. R.; Sidebottom, A. M.; Sedio, B. E.; Klitgaard, A.; Larson, C. B.; P, C. A. B.; Torres-Mendoza, D.; Gonzalez, D. J.; Silva, D. B.; Marques, L. M.; Demarque, D. P.; Pociute, E.; O'Neill, E. C.; Briand, E.; Helfrich, E. J. N.; Granatosky, E. A.; Glukhov, E.; Ryffel, F.; Houson, H.; Mohimani, H.; Kharbush, J. J.; Zeng, Y.; Vorholt, J. A.; Kurita, K. L.; Charusanti, P.; McPhail, K. L.; Nielsen, K. F.; Vuong, L.; Elfeki, M.; Traxler, M. F.; Engene, N.; Koyama, N.; Vining, O. B.; Baric, R.; Silva, R. R.; Mascuch, S. J.; Tomasi, S.; Jenkins, S.; Macherla, V.; Hoffman, T.; Agarwal, V.; Williams, P. G.; Dai, J.; Neupane, R.; Gurr, J.; Rodriguez, A. M. C.; Lamsa, A.; Zhang, C.; Dorrestein, K.; Duggan, B. M.; Almaliti, J.; Allard, P. M.; Phapale, P.; Nothias, L. F.; Alexandrov, T.; Litaudon, M.; Wolfender, J. L.; Kyle, J. E.; Metz, T. O.; Peryea, T.; Nguyen, D. T.; VanLeer, D.; Shinn, P.; Jadhav, A.; Muller, R.; Waters, K. M.; Shi, W.; Liu, X.; Zhang, L.; Knight, R.; Jensen, P. R.; Palsson, B. O.; Pogliano, K.; Linington, R. G.; Gutierrez, M.; Lopes, N. P.; Gerwick, W. H.; Moore, B. S.; Dorrestein, P. C.; Bandeira, N., Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nature biotechnology* **2016,** 34, (8), 828-837.
32. Aron, A. T.; Gentry, E.; McPhail, K. L.; Nothias, L. F.; Nothias-Esposito, M. l.; Bouslimani, A.; Petras, D.; Gauglitz, J. M.; Sikora, N.; Vargas, F., Reproducible molecular networking of untargeted mass spectrometry data using GNPS. **2019**.

Peer-reviewed version available at Mar. Drugs 2019, 17, 671; doi:10.3390/md17120671

33. Bauermeister, A.; Pereira, F.; Grilo, I. R.; Godinho, C. C.; Paulino, M.; Almeida, V.; Gobbo-Neto, L.; Prieto-Davo, A.; Sobral, R. G.; Lopes, N. P.; Gaudencio, S. P., Intra-clade metabolomic profiling of MAR4 Streptomyces from the Macaronesia Atlantic region reveals a source of anti-biofilm metabolites. *Environ Microbiol* **2019,** 21, (3), 1099-1112.
34. Ma, Y. M.; Liang, X. A.; Kong, Y.; Jia, B., Structural diversity and biological activities of indole diketopiperazine alkaloids from fungi. *Journal of agricultural and food chemistry* **2016,** 64, (35), 6659-71.
35. da Silva, M. M.; Andrade, M. D.; Bauermeister, A.; Merfa, M. V.; Forim, M. R.; Fernandes, J. B.; Vieira, P. C.; da Silva, M. F. D. F.; Lopes, N. P.; Machado, M. A.; de Souza, A. A., A simple defined medium for the production of true Diketopiperazines in *Xylella fastidiosa* and their identification by ultra-fast liquid chromatography-electrospray ionization ion trap mass spectrometry. *Molecules* **2017,** 22, (6), 985-999.
36. Rabindran, S. K.; He, H. Y.; Singh, M.; Brown, E.; Collins, K. I.; Annable, T.; Greenberger, L. M., Reversal of a novel multidrug resistance mechanism in human colon carcinoma cells by fumitremorgin C. *Cancer Res.* **1998,** 58, (24), 5850-5858.
37. Bofinger, M. R.; de Sousa, L. S.; Fontes, J. E. N.; Marsaioli, A. J., Diketopiperazines as cross-communication quorum-sensing signals between *Cronobacter sakazakii* and *Bacillus cereus*. *Acs Omega* **2017,** 2, (3), 1003-1008.
38. Alfarisi, O.; Alghamdi, W. A.; Al-Shaer, M. H.; Dooley, K. E.; Peloquin, C. A., Rifampin vs. rifapentine: what is the preferred rifamycin for tuberculosis? *Expert Rev Clin Phar* **2017,** 10, (10), 1027-1036.
39. Meena, K. R.; Kanwar, S. S., Lipopeptides as the antifungal and antibacterial agents: applications in food safety and therapeutics. *Biomed Res Int* **2015,** 2015, 1-9.
40. Vazquez-Laslop, N.; Mankin, A. S., How macrolide antibiotics work. *Trends Biochem Sci* **2018,** 43, (9), 668-684.
41. Katayama, N.; Fukusumi, S.; Funabashi, Y.; Iwahi, T.; Ono, H., TAN-1057 A-D, new antibiotics with potent antibacterial activity against methicillin-resistant *Staphylococcus aureus*. Taxonomy, fermentation and biological activity. *The Journal of antibiotics* **1993,** 46, (4), 606-13.
42. Frlan, R.; Gobec, S., Inhibitors of cathepsin B. *Curr Med Chem* **2006,** 13, (19), 2309-2327.
43. Li, T.; Christensen, S. D.; Frankel, P. H.; Margolin, K. A.; Agarwala, S. S.; Luu, T.; Mack, P. C.; Lara, P. N., Jr.; Gandara, D. R., A phase II study of cell cycle inhibitor UCN-01 in patients with metastatic melanoma: a California Cancer Consortium trial. *Investigational new drugs* **2012,** 30, (2), 741-8.
44. Yadav, S. S.; Prasad, C. B.; Prasad, S. B.; Pandey, L. K.; Singh, S.; Pradhan, S.; Narayan, G., Anti-tumor activity of staurosporine in the tumor microenvironment of cervical cancer: An in vitro study. *Life sciences* **2015,** 133, 21-8.
45. Jimenez, P. C.; Wilke, D. V.; Ferreira, E. G.; Takeara, R.; de Moraes, M. O.; da Silveira, E. R.; da Cruz Lotufo, T. M.; Lopes, N. P.; Costa-Lotufo, L. V., Structure elucidation and anticancer activity of 7-oxostaurosporine derivatives from the Brazilian endemic tunicate *Eudistoma vannamei*. *Marine drugs* **2012,** 10, (5), 1092-102.
46. Andréo, M. A.; Jimenez, P. C.; Siebra, J. B.; Costa-Lotufo, L. V.; Vessecchi, R.; Niehues, M.; Lopes, J. L.; Lopes, N. P., Systematic UPLC-ESI-MS/MS study on the occurrence of staurosporine and derivatives in associated marine microorganisms from *Eudistoma vannamei*. *JBCS* **2012,** 23, (2), 335-343.
47. Williams, P. G.; Asolkar, R. N.; Kondratyuk, T.; Pezzuto, J. M.; Jensen, P. R.; Fenical, W., Saliniketals A and B, bicyclic polyketides from the marine actinomycete *Salinispora arenicola*. *Journal of natural products* **2007,** 70, (1), 83-8.
48. Hufsky, F.; Scheubert, K.; Bocker, S., New kids on the block: novel informatics methods for natural product discovery. *Nat Prod Rep* **2014,** 31, (6), 807-817.

Peer-reviewed version available at Mar. Drugs 2019, 17, 671; doi:10.3390/md17120671

664 49. Hufsky, F.; Bocker, S., Mining molecular structure databases: Identification of small molecules based on
665 fragmentation mass spectrometry data. *Mass Spectrom Rev* **2017,** 36, (5), 624-633.
666 50. Mohimani, H.; Gurevich, A.; Shlemov, A.; Mikheenko, A.; Korobeynikov, A.; Cao, L.; Shcherbin, E.; Nothias,
667 L. F.; Dorrestein, P. C.; Pevzner, P. A., Dereplication of microbial metabolites through database search of
668 mass spectra. *Nat Commun* **2018,** 9.
669 51. Wan, Z. Y.; Fang, W.; Shi, L. Q.; Wang, K. M.; Zhang, Y. N.; Zhang, Z. G.; Wu, Z. Y.; Yang, Z. W.; Gu, Y. C.,
670 Novonestmycins A and B, two new 32-membered bioactive macrolides from *Streptomyces phytohabitans*
671 HBERC-20821. *J Antibiot* **2015,** 68, (3), 185-190.
672 52. Crowe, M. C.; Brodbelt, J. S.; Goolsby, B. J.; Hergenrother, P., Characterization of erythromycin analogs by
673 collisional activated dissociation and infrared multiphoton dissociation in a quadrupole ion trap. *J Am Soc*
674 *Mass Spectr* **2002,** 13, (6), 630-649.
675 53. Wills, R. H.; Tosin, M.; O'Connor, P. B., Structural characterization of polyketides using high mass accuracy
676 tandem mass spectrometry. *Anal Chem* **2012,** 84, (20), 8863-8870.
677 54. Gontang, E. A.; Fenical, W.; Jensen, P. R., Phylogenetic diversity of gram-positive bacteria cultured from
678 marine sediments. *Applied and environmental microbiology* **2007,** 73, (10), 3272-82.
679 55. Mosmann, T., Rapid colorimetric assay for cellular growth and survival: application to proliferation and
680 cytotoxicity assays. *Journal of immunological methods* **1983,** 65, (1-2), 55-63.
681 56. Kumar, V.; Bisht, G. S.; Institu, S. B. S. P. G., An improved method for isolation of genomic DNA from
682 filamentous actinomycetes. *Journal of Sci. Engg. & Tech. Mgt. Vol* **2010,** 2, (2), 10-13.
683 57. Katoh, K.; Standley, D. M., MAFFT multiple sequence alignment software version 7: improvements in
684 performance and usability. *Molecular biology and evolution* **2013,** 30, (4), 772-80.
685 58. Stamatakis, A., RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies.
686 *Bioinformatics* **2014,** 30, (9), 1312-3.
687 59. Letunic, I.; Bork, P., Interactive tree of life (iTOL) v3: an online tool for the display and annotation of
688 phylogenetic and other trees. *Nucleic acids research* **2016,** 44, (W1), W242-5.
689 60. iTOL v4. In *Interative Tree of Life*.
690 61. Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N. S.; Wang, J. T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker,
691 T., Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome*
692 *research* **2003,** 13, (11), 2498-504.