OPEN

Metabolomic study of soft corals from the Colombian Caribbean: PSYCHE and ^{1}H-NMR comparative analysis

Liliana Santacruz[1], Diana X. Hurtado[1], Roisin Doohan[2], Olivier P. Thomas[2], Mónica Puyana[3] & Edisson Tello[1]*

Marine organisms have evolved to survive against predators in complex marine ecosystems via the production of chemical compounds. Soft corals (Cnidaria, Anthozoa, Octocorallia) are an important source of chemically diverse metabolites with a broad spectrum of biological activities. Herein, we perform a comparative study between high-resolution proton nuclear magnetic resonance (^{1}H-NMR) and pure shift yielded by chirp excitation (PSYCHE) experiments to analyze the metabolic profile of 24 soft corals from the Colombian Caribbean to correlate chemical fingerprints with their cytotoxic activity against three cancer cell lines (human cervical carcinoma (SiHa), human prostatic carcinoma (PC3) and human lung adenocarcinoma (A549)). All data obtained were explored using multivariate analysis using principal components analysis (PCA) and orthogonal partial least squares (OPLS) analysis. The results did not show a significant correlation between clusters using ^{1}H-NMR data in the PCA and OPLS-DA models and therefore did not provide conclusive evidence; on the other hand, a metabolomic analysis of PSYCHE data obtained under the same parameters revealed that when a decoupled experiment is performed, it was possible to establish a statistically valid correlation between the chemical composition of soft corals and their cytotoxic activity against the PC3 cancer cell line, where the asperdiol and plexaurolone markers were putatively identified and related to the cytotoxic activity presented by extracts of *Plexaurella* sp. and *Plexaura kukenthali*, respectively. These results increase the speed, effectiveness and reliability of analyses for the study of this type of complex matrices.

Metabolomics studies allow a complete analysis of a set of metabolites that are the substrates and products of metabolism driving essential cellular functions in a given biological system[1]. This research has applications in different fields, such as pharmacology, environmental sciences, chemotaxonomy, nutrition and medicine[2]. Recently, metabolomic approaches have allowed the understanding of complex biological systems and the biochemical composition of organisms that live in diverse environments, such as marine areas[3]. Goulitquer *et al.* demonstrated that metabolites are important links between genotype and phenotype and are important for studying several biological processes and for analyzing interactions between organisms within communities via mass spectrometry (MS)-based metabolomics[4]. In addition, Mohamed A. Farag *et al.* compared metabolomics results obtained with liquid chromatography coupled to mass spectrometry (LC-MS) with those obtained by nuclear magnetic resonance (NMR) to investigate the metabolism of 16 Sarcophyton species in the context of their genetic diversity and growth habitats[3].

The importance of studying marine invertebrates lies in their extraordinary ability to produce a broad variety of chemical compounds with unique chemical structures that in most cases have been correlated with significant biological activities, which has led to the successful development of commercial drugs. One of the best-known compounds is ziconotide (Prialt®), a synthetic derivative of a peptide originally isolated from the venom of the marine snail *Conus magus*; this compound is a neuronal calcium antagonist useful for treating severe chronic

[1]Bioprospecting Research Group and Biosciences Doctoral Program, Faculty of Engineering, Campus Puente del Común, Universidad de la Sabana, 250001, Chía, Colombia. [2]Marine Biodiscovery, School of Chemistry and Ryan Institute, National University of Ireland Galway (NUI Galway), University Road, H91 TK33, Galway, Ireland. [3]Departamento de Ciencias Biológicas y Ambientales, Universidad Jorge Tadeo Lozano, Carrera 4 # 22-61, 110311, Bogotá, Colombia. *email: edisson.tello@unisabana.edu.co

pain[5]. Trabectedin ET-743 (Yondelis™), a complex peptide originally isolated from *Ecteinascidia turbinata*, is effective in the treatment of many types of cancer, including melanoma, sarcoma, and lung, breast, ovarian, endometrial and prostate cancer[6]. Brentuximab vedotin®, originally isolated from the marine opisthobranch *Dolabella auricularia*, is used against breast and Hodgkin's lymphomas[7]. Eribulin mesylate (Halaven®), originally isolated from the marine sponge *Halichondria okadai*, is a potent microtubule inhibitor used in the treatment of breast cancer[8], and vidarabine (Vira-A®), a modified nucleoside originally isolated from the Caribbean sponge *Cryptotethya crypta*, is an inhibitor of viral DNA polymerases and other enzymes and is used against varicella zoster and herpes viruses[9]. This information allows us to infer that marine organisms are a prolific source of bioactive and novel molecules that can be used as potential agents against different human illnesses.

Bioprospecting marine research in the Colombian Caribbean has established organisms with various biological activities, mainly anti-inflammatory, cytotoxic and antiviral activities, e.g., gorgonian *Antillogorgia (Pseudopterogorgia) elisabethae* collected at the Islands of Providencia and San Andrés showed anti-inflammatory properties[10]; additionally, the octocorals *Eunicea laciniata* and *Eunicea asperula* have been evaluated for their cytotoxic and antiviral activities, where dolabellane diterpenes isolated from the soft corals *E. laciniata* and *E. asperula* showed anti HSV-1 activity[11]. In addition, the prostaglandins isolated from the soft coral *Plexaura homomalla* presented anti-inflammatory activity[12].

There is no single analytical technology or protocol to analyze the overall metabolome of an organism and obtain a complete metabolic profile[13]. However, metabolomics approaches use hyphenated and high-throughput techniques to perform chromatographic separation of metabolites using either liquid chromatography (LC) or gas chromatography (GC) coupled with mass spectrometry or nuclear magnetic resonance (NMR) to analyze complex mixtures of metabolites from different organisms[14]. Furthermore, each technique has advantages and drawbacks, and although the analytical technique most often used for metabolite profiling is liquid chromatography-mass spectrometry, due to its high sensitivity and the wide range of molecules that can be analyzed, the use of NMR experiments has expanded rapidly over the past ten years[15]. Regardless of the low sensitivity of NMR, this technique presents advantages, such as being nondestructive and requiring no sample preparation and relatively short acquisition times. NMR is a highly reproducible method for metabolomics studies. At present, it is possible to record NMR spectra from crude extracts[16] to perform preliminary studies of the metabolic composition of marine organisms. NMR analyses give a global overview of all metabolites present in complex biological samples such as soft coral extracts that produce compounds such as prostaglandins, sterols and a wide range of terpenes, and these compounds represent the main chemical defense of these organisms against predators[17,18]. On the other hand, metabolomic studies using ^{1}H-NMR on mixtures may experience signal overlap, especially in samples that contain significant amounts of fatty acids and terpenoids, hampering comparative metabolomic studies of extracts from soft corals[19,20]. The Table S1 in Supplementary Information shows the advantages and disadvantages of the most commonly used methods for metabolomics analysis.

The overlapping signals in ^{1}H-NMR experiments are a complex issue, and to greatly improve spectral resolution and the gathering of metabolomics data, new NMR techniques such as Pure Shift Yielded by Chirp Excitation (PSYCHE) are proposed. This method proposed by Morris and coworkers[21] contains two low flip angle (β) swept-frequency pulses in the presence of a weak pulsed field gradient. The advantage of this technique over ^{1}HNMR is that it resolves overlapping ^{1}H-^{1}H scalar coupling multiplets, which improves chemical shift analysis of complex natural products[22]. PSYCHE suppresses the effects of homonuclear coupling and allows observation of decoupled ^{1}H-NMR spectra with chemical shifts only, helping in the identification of potential biomarkers in metabolomic studies[16].

Here, we used ^{1}H and PSYCHE NMR experiments to perform a metabolomic comparison of 24 soft coral extracts (complex biological samples) to examine whether there is a correlation between the chemical composition of the extracts and their cytotoxicity against SiHa, PC3 and A549 tumor cancer cell lines[23]. When large data sets are analyzed, a multivariate analysis (MVA) is a valuable approach for the identification of potential key markers in complex mixtures.

Due to this, MVA was developed using the principal component analysis (PCA) algorithm to compress a dataset onto a lower-dimensional feature subspace maintaining most of the relevant information allowing to stablish chemical differences in the NMR experiments[24]. This approach was followed by the orthogonal projection to latent structures discriminant analysis (OPLS-DA), which is a supervised model that filters out orthogonal metabolite variables that are not related to categorical variables to discriminate and separate predictive from nonpredictive (orthogonal) variation. A response matrix Y (containing toxicity data) was correlated to a descriptor matrix X (containing spectral data) that is orthogonal (noncorrelated) to Y to identify the markers that contributed to the discrimination of cytotoxic activity and to determine outliers[25]. The reliability of the model was verified by the cross validation method, and the parameters for the OPLS model, R2 and Q2, were calculated (varying from 0 to 1), where R2 corresponded to the fraction of the variance explained by the model, Q2 suggests the predictive capability of the model[26], and it cross validation reproducibility Q2/R2, was also considered indicative of relevant associations[27].

Furthermore, the variable importance in projection (VIP) was calculated by using partial least squares-discriminant analysis (PLS-DA) model[28], which describes a quantitative estimation of the discriminatory power of each individual feature and summarizes the contribution a variable makes to the model[29]. A combination of univariate and multivariate statistical approaches (VIP > 3, p < 0.05) was used as criteria to discriminate the key markers[30].

OPLS-DA allowed to discriminate the soft coral extracts by separating predictive and nonpredictive data[31]. Finally, MVA and its loading and score plots, which are closely linked such that features (chemical shifts) that are highly loaded in a specific direction in the loading plot contribute to an increased degree to the observations (soft corals) that are located in that direction in the score plot. This allowed us to establish that the key markers B2_5118, B4_4965 and B4_7686 were mainly responsible of the cytotoxic activity exhibit by the extracts of the soft corals *Plexaurella* sp. and *P. kukenthali* against the tumor cancer cell line PC3[23]. This article describes a
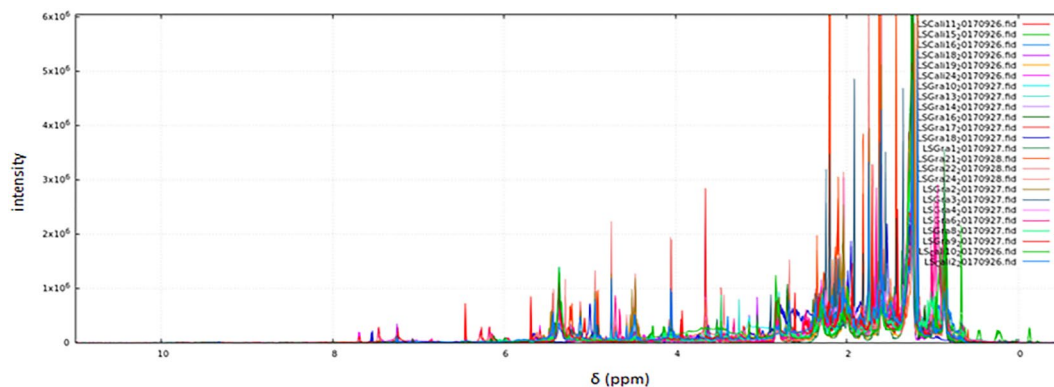
Figure 1. ¹H-NMR data of 24 soft coral extract samples using the open access software NMRProcFlow v1.2 25. Spectra of soft coral extracts are presented in different colors.

workflow that helped to estimate the confidence levels of compound annotations in reported metabolites that were used as established by the Metabolomics Society at 2017[32]. In summary, this work achieved a confidence level of 2 for features (VIPs) most responsible for cytotoxic activity considering that the feasible structure was compared with data or databases of literature by diagnostic evidence[33].

Results

¹H-NMR metabolomic fingerprints.
Cytotoxic effect correlation between PC3, A549 and SiHa cancer cell lines and chemical composition using ¹H-NMR spectra of the 24 crude extracts of soft corals (in the genera *Plexaura, Pseudopterogorgia*, *Eunicea*, *Plexaurella* and *Pseudoplexaura* Table S2 in Supplementary Information) were explored over 168 spectral bins (variables) describing metabolic profiles using the open access NMRProcFlow v1.2 software[34] developed by INRA Science & Impact in France. The complexity of the spectra can be visualized in some regions due to overlapping of signals (Fig. 1). The data processing script is described in the Supplementary Information (Table S4).

Exploratory data analysis was performed by PCA using the MetaboAnalyst version 3.0 web application[35] (Supplementary Fig. S1). For the tumor cell lines A549, PC3 and SiHa, extracts of *P. kukenthali, E. asperula* and *Plexaurella* sp. were most distant from those groups that exhibited the highest cytotoxic activity against the three tumor cell lines. Additionally, we used partial least squares-discriminate analyses (OPLS-DA) (Fig. 2) to visualize increased separation of the groups that presented cytotoxic activity from those that did not. Finally, to demonstrate that the observed PCA was valid, the results from the plots (loading and observations) were contrasted, with the results obtained in the PLS-DA model. Additionally, the statistical parameters for the PLS-DA models for the classification of the ¹H-NMR data were determined, as shown in Table 1.

In this study, an extract was considered as active, if it showed inhibition of tumor cell lines ≥40% at 20 μg/mL. This designation included extracts that showed moderate to strong cytotoxic activity[36] (Table S3).

With the purpose of discriminating between metabolite profiles (analysis of a large group of metabolites which are related to a class of compounds[37]) and cytotoxicity of all the extracts analyzed, a validation of the model was carried out using a PLS-DA based on the PLS algorithm. The discriminating variable was cytotoxic activity.

A cross validation test was developed for the classification model using the three tumor cell lines, the models were analyzed using R2 and Q2 metrics and the Q2/R2 ratio. Predictive relevance is considered when values are greater than 0.5[38]. The resulting models for each tumor cell line revealed an overfitting in the separation of the metabolic profiles. Discrimination of the samples according to their cytotoxic activity is shown in Table 1.

Pure shift experiments (PSYCHE) in metabolomic fingerprints.
To compare the results (cytotoxic activity vs. metabolic profile) between the ¹H and PSYCHE NMR experiments, the same protocol and the same number of samples were used for both experiments. The PSYCHE data was analyzed over 113 spectral bins (variables) describing metabolic profiles using the NMRProcFlow v1.2 software[34]. The PSYCHE spectra of the 24 soft coral extracts analyzed are shown in Fig. 3a. Figure 3b shows the same 24 spectra but separated into two groups by color: cytotoxic extracts appear in red, and noncytotoxic extracts appear in green. To establish bins responsible for group separation (VIP), it was necessary to perform a statistical analysis to show the relevant chemical shift that allowed the discrimination.

PCA[35] showed that the cytotoxic extracts of *P. kukenthali* and *E. asperula* are separated in PC2 for the tumor line A549, as shown in Supplementary Fig. S2a. On the other hand, the PC2 shown in the model for the PC3 tumor line showed that cytotoxic extracts of *P. kukenthali*, *E. asperula* and *Plexaurella* sp., were the most separated, which agrees with the results obtained for this same tumor cell line using the ¹H-NMR experiment (See Supplementary Fig. S2b). Finally, PC1 and PC2 from PCA generated for the SiHa tumor cell line did not show a clear separation between the extracts and their cytotoxic activity, as shown in Supplementary Fig. S2c.

It is appropriate to apply PCA as a first step for exploratory studies where differences between experimental groups may be unknown or unpredictable; however, the spectral noise and high within-group variation do not show a separation between groups in many cases[39]; therefore, to overcome this problem, a supervised model OPLS-DA was constructed for classification of the samples and to improve discrimination between metabolic
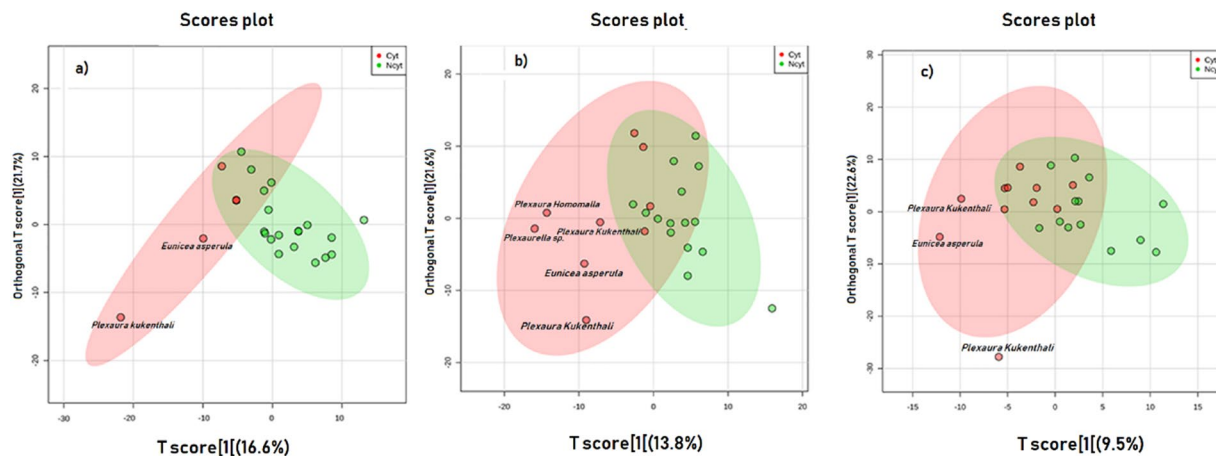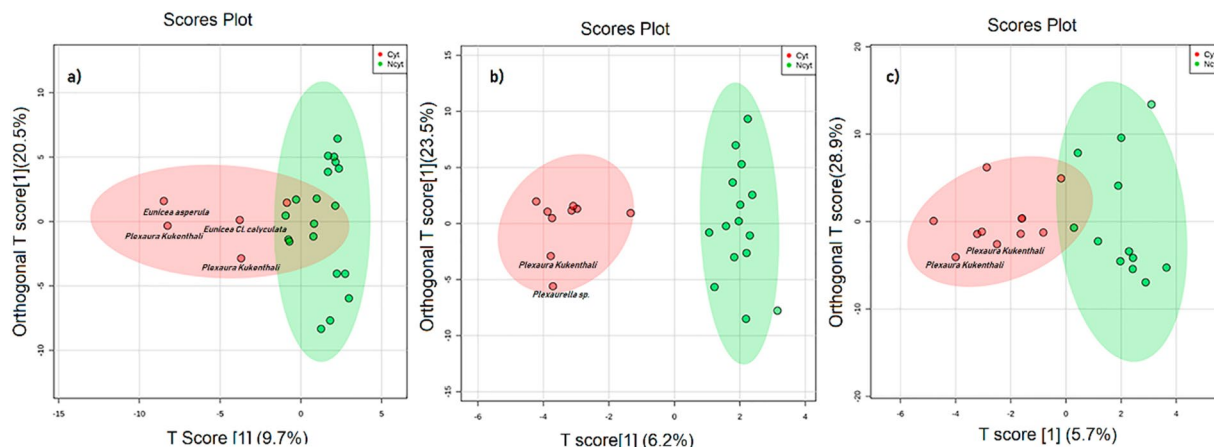
Figure 2. OPLS-DA Score plots of ^{1}H-NMR metabolomic data of 24 soft coral extracts based on their cytotoxic activity against three tumor cell lines. (**a**) human lung adenocarcinoma, A549 (**b**) human prostatic carcinoma, PC3 and (**c**) human cervical carcinoma, SiHa. Red dots represent active extracts, and green dots represent extracts that were not active. The ellipse represents the 95% confidence interval of the PLS-DA model[71].
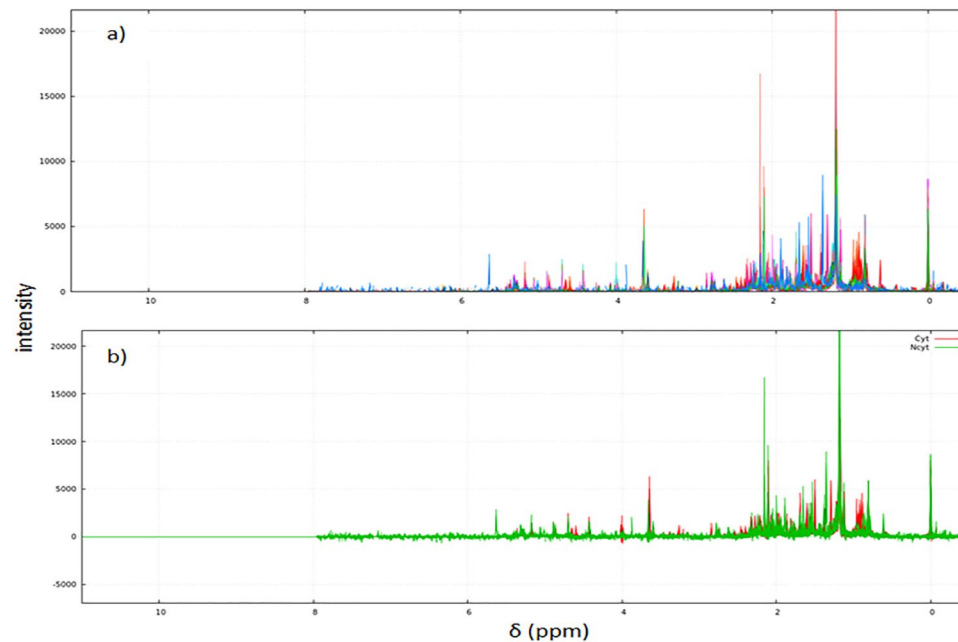
Figure 3. (**a**) PSYCHE spectra of the 24 soft coral extracts analyzed. The spectrum of each soft coral extract is represented in a different color. (**b**) Spectra separated into two groups, by color. Red (cytotoxic extracts), green (not cytotoxic extracts).

PLS-DA parameters			
Cell lines	Q2	R2	Q2/R2
A549	0.14	0.49	0.28
PC3	−0.02	0.49	−0.04
SiHa	0.03	0.59	0.05

Table 1. Statistical parameters of PLS-DA models for classification of ^{1}H-NMR experiments according to cytotoxic and noncytotoxic groups from 24 extracts of soft corals tested against three tumor cell lines; A549: human lung adenocarcinoma, PC3: human prostatic carcinoma and SiHa: human cervical carcinoma. The analysis was done using a 10-fold cross-validation method. Q2: predictive capability, R2: correlation coefficient.

Figure 4. OPLS-DA Scores Plot[35] of PSYCHE experiment of metabolomics data from 24 soft coral extracts based on their cytotoxicity against three tumor cell lines. (**a**) human lung adenocarcinoma A549, (**b**) human prostatic carcinoma PC3 and (**c**) human cervical carcinoma SiHa. Red dots represent active extracts, and green dots represent extracts that were not active. Red dots represent cytotoxic extracts, green dots represent extracts that were not cytotoxic. An extract was considered active if exhibited inhibition of the tumor cell line $\geq$ 40% at 20 µg/mL)[72]. The ellipse represents the 95% confidence interval of the model[71].

PLS-DA parameters			
Cell lines	**Q2**	**R2**	**Q2/R2**
A549	0.14	0.76	0.19
PC3	0.39	0.66	0.59
SiHa	0.07	0.72	0.09

Table 2. Statistical parameters of PLS-DA models for classification of PSYCHE experiment according to cytotoxic and non-cytotoxic groups from 24 extracts of soft corals tested against three tumor cell lines; A549: human lung adenocarcinoma, PC3: human prostatic carcinoma and SiHa: human cervical carcinoma. The analysis was done using a10-fold cross-validation method Q2: predictive capability, R2: correlation coefficient.

profiles and their cytotoxic activity (Fig. 4) because the OPLS-DA algorithm is normally applied when there are only two classes, improving the class discrimination and robustness of important feature identification[40].

As performed for the ^{1}H-NMR experiments, the validation of the PSYCHE models (Table 2) was carried out using PLS-DA, where the models were evaluated using R2, Q2 and Q2/R2 metrics. A model is considered predictive if the Q2/R2 ratio is greater than 0.5[41]. When the results of the models obtained for the ^{1}H-NMR experiments were compared with those from the PSYCHE NMR experiment (Tables 1 and 2), an improvement in the total amount of variance (R2) and in the accuracy (Q2) was observed for the latter experiment. This may be because the soft coral extracts contain many overlapping peaks from multiple compounds and that such overlaps obscure peak assignments and compromise the quantitative analysis, which explains the overfitting observed in the OPLS-DA model using ^{1}H-NMR data compared to that in the decoupled PSYCHE model; all this is reflected in the total reproducibility explained by the ratio Q2/R2[42]. However, a reasonable separation between the groups (cytotoxic and noncytotoxic) was only evident in the model generated for the PC3 tumor cell line, with a Q2/R2 value of 0.59 (Table 2). These results confirm that cluster separations in the OPLS-DA score plot shown in Fig. 4b for the PC3 tumor cell line were statistically significant.

The cytotoxic activity of each extract against the tested tumor cell lines showed that the extracts of *P. kukenthali* (code G18P) and *Plexaurella* sp. (code G22P) exhibited inhibition percentages of 64.0% and 63.5% against the PC3 tumor cell line, respectively. In Fig. 4b, the OPLS-DA of PSYCHE experiments show that those extracts were part of the group that presented cytotoxic activity against the tumor line PC3.

Information on the metabolites responsible for the separation between cytotoxic and noncytotoxic groups (VIP) for the tumor cell line PC3 (statistically significant model) was extracted from the PLS-DA. This process allowed us to identify key discriminatory metabolites through a VIP analysis, which revealed several distinguishing patterns. High values indicate the increased discriminatory power of particular metabolites. Variables with VIP > 1.0 are considered potential biomarker candidates for group discrimination[43]. However, for this study, only variables with VIP > 3.0 were considered because these variables play important roles in the discrimination of cytotoxic activity, as three markers were the major variables responsible for group separation and the chemical shifts corresponding to these VIP markers in the NMR experiments were well resolved, which allowed an improved interpretation of the results.

The three markers that showed the highest scores in the VIP analysis (Fig. 5a) were B2_5118 (score = 5.1), marker B4_4965 (score = 4.8) and marker B4_7686 (score = 3.5). Additionally, in the OPLS S-plot (Fig. 5b), each of the three markers selected as VIPs can be clearly distinguished, with each coordinate representing a single
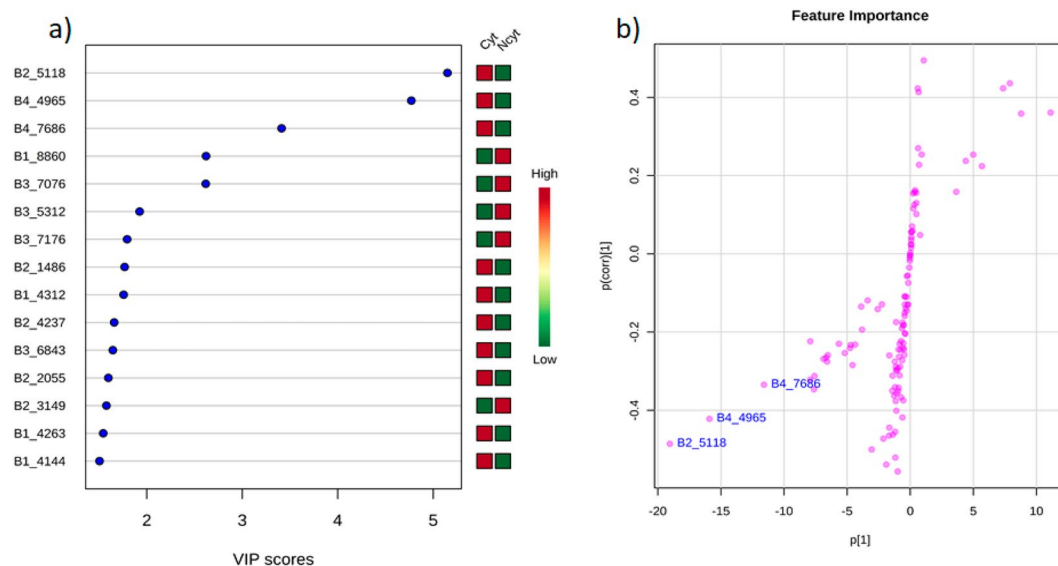
Figure 5. (**a**) VIP (variable importance in the projection) scores/p < 0.05[35] obtained from the PSYCHE PLS-DA model (**b**). Feature importance from the OPLS S-plot[71].

NMR signal (contributing variables to the classification). This result shows that features (VIP) correlated with extracts of *P. kukenthali* (Gra 18) and *Plexaurella* sp. (Gra 22) have cytotoxic activity against the PC3 tumor cell line, in accordance with the analysis obtained from the OPLS-DA (Fig. 4b).

The previously described markers B2_5118, B4_4965 and B4_7686 showed a high correlation with 25 additional features according to "Pattern Hunter" a tool from MetaboAnalyst software[35] (See Supplementary Fig. S3). This analysis allowed a putative identification of the possible compounds that corresponded to those features. In addition, PeakForest 2.0.1® software functions were used with a correlation coefficient of ±0.5 to help in the identification of metabolites.

Confidence annotation of compound identification.

Confidence annotation consisted of three steps. First, a literature review of the compounds that have been isolated from the species was performed using the SciFinder® database. Cembrane-type diterpenoids have been isolated from *Plexaurella* sp. and *P. kukenthali*[44]. From a biomedical perspective, some of those cembranes have shown cytotoxic activity against several tumor cell lines and have also been reported as anti-inflammatory, neuroprotective and antimicrobial compounds[45].

In a second step, a verification of chemical formulas and exact masses was developed using both LC-MS and LC-MS/MS with the purpose of yielding a putative identity of the compounds. The presence of the compound asperdiol in the extract of *Plexaurella* sp., with a molecular formula $C_{20}H_{32}O_3$ corresponding to the ion *m/z* 321.2426 [M + H], was confirmed (Supplementary Fig. S5a,b). In addition, analysis of MS/MS data fragmentation using MetFrag software[46] matched the data for the compound asperdiol (score 7.47). On the other hand, from a high-resolution electrospray ionization mass spectrometry (HRESIMS) analysis of the extract from *P. kukenthali*, the formula $C_{20}H_{34}O_3$ was calculated, which confirmed the possible presence of the compound plexaurolone with an ion *m/z* of 323.25860 [M + H] and an *m/z* of 345.24038 [M + Na] (Supplementary Fig. S6a,b)[47].

In the third step, chemical shifts were analyzed using the AntiMarin® database to check signal patterns that corresponded to the kind of compounds (diterpenoid-type metabolites) expected to be present in the mixture (diterpenoid-type compounds)[48]. As a consequence, in the 1H-NMR spectra of the extracts from *Plexaura kukenthali* and *Plexaurella* sp. (Fig. S4 in Supplementary Information), singlets were found at a range of chemical shifts from approximately 0.5 to 2.0 ppm, indicating the presence of several methyl groups, which is characteristic of terpenoid-type compounds and in agreement with the literature[49]. Additionally, chemical signals of VIP markers B4_4965 and B4_7686 were found in the 1H-NMR spectrum of the extract from *Plexaurella* sp. When comparing these data with those reported in the literature, we detected similarities with the NMR data published for the anti-neoplastic compound asperdiol[50]. Additionally, an analysis of the marker B2_5118 in the 1H-NMR spectrum of the extract from *P. kukenthali* and further comparisons with NMR data reported in the literature allowed us to make a putative identification of this compound as plexaurolone. Table S5 shows the summarizing the key results of this study.

Finally, in accordance with recent discussions in the metabolomics community[51], the confidence grade in the identification of these metabolites (asperdiol and plexaurolone shown in the Supplementary Information) is level 2 because this level reveals probable structure using fragmentation data from literature and/or libraries and databases.

Discussion

There are two analytical techniques commonly used in metabolomics, each with advantages and disadvantages: mass spectrometry (MS) and nuclear magnetic resonance (NMR) spectroscopy. There is an increasing growth in NMR-based metabolomics over the last few years due to the advantages that this technique offers: high reproducibility and quantification ability over a wide dynamic range of compounds and support for the identification of unknown structures; moreover, NMR is nondestructive and can be utilized *in vivo*[48].

NMR is a very suitable method for analysis that allows the simultaneous detection of diverse groups of secondary metabolites in complex matrices, as evidenced by a study conducted by Mohamed *et al*. where some cembrane diterpenoids were isolated from the *Sarcophyton ehrenbergi* soft coral and identified using NMR techniques[52]. In another study, Zhang *et al*., also using NMR, found cytotoxic diterpenoids from the soft coral *Sinularia microclavata*[53].

One-dimensional (1D) ¹H-NMR is the most widely used NMR approach in metabolomics. However, this approach has some limitations being key signal identification and accurate peak integration the most challenging, due to hundreds of overlapping signals[54]. These challenging conditions can be solved if the density of signals in a spectrum is decreased by almost an order of magnitude, whereby a pure shift experiment (PSYCHE) appear as an alternative that offers greatly improved signal resolution and spectral purity. Thus, it has the potential to find wide application in the NMR spectroscopy of small molecules and NMR-based metabolomics[21,55], as we demonstrate in this research by using the PSYCHE experiment in soft coral metabolomics for the first time.

Here, the application of OPLS-DA to 1D ¹H-NMR data (Table 1) revealed an excessive data adjustment in the model that could be attributed to signal overlap in the ¹H-NMR spectra of soft-coral extracts, which was principally due to olefinic protons of diterpenes and the high-field region between 3.2 and 0.8 ppm having a high density of signals due to terpene methyl, methylene, and methine resonances, revealing that soft coral extracts contain large amounts of terpenoids[56]. The noise present in each spectrum of this type of biological sample hinders the alignment and creation of the bins used to reduce the data dimensionality in the NMR spectrum from several thousand points[34]. Therefore, the results obtained from this experiment were not conclusive, which agrees with the reports by Farag *et al*. 2017.

Validations using OPLS-DA of PSYCHE NMR experiment data (Table 2), on the other hand, showed an improvement in the total amount of variance, accuracy and reproducibility of the models, particularly for the PC3 tumor cell line, which was predictive according to the validation results. This result showed that metabolomic studies using PSYCHE NMR experiments allowed us to obtain a more reliable correlation between the chemical composition and the cytotoxic activity of soft coral extracts against the PC3 tumor cell line than metabolomic studies using solely ¹H-NMR data. The results obtained throughout the metabolomic process always showed a correlation in the data that was identified by the PCA, OPLS-DA, the markers displayed in the VIP table, and an S-plot obtained from PLS-DA of the extracts from *P. kukenthali* and *Plexaurella* sp. These organisms were highlighted in all the analyses as those responsible for the separation of the groups that presented cytotoxic activity and those indicating that the variables B2_5118, B4_4965 and B4_7686 (chemical shifts) that were present in these species were the main markers related to the activity shown; also, we can derive from this analysis that these extracts, in addition to exhibiting the greatest cytotoxicity, were the closest in the orthogonal t score, indicating they may have a common chemical composition. This was evidenced from MS and MS/MS spectra (Supplementary Figs. S5 and S6), which showed that the chemical formulas $C_{20}H_{32}O_3$ (with an *m/z* 321.2426 [M + H]) and $C_{20}H_{34}O_3$ (with an *m/z* of 323.25860 [M + H]) corresponded to the compounds asperdiol and plexaurolone, respectively. They also showed that both compounds were present in the two soft coral species. However, the ¹H-NMR signals corresponding to asperdiol were not clearly observed in the extract of *P. kukenthali*, and in the same way, the chemical shifts of plexaurolone could not be observed in *Plexaurella* sp. extracts due to plexaurolone abundance.

Here, the fact that the extracts from *P. kukenthali* and *Plexaurella* sp. were mainly responsible for the separation of the group that presented cytotoxic activity was statistically validated using the cross-validation method (Table 2), which agreed with the cytotoxicity presented by these species as reported by Honda *et al*.[57] and Rueda *et al*.[58], who showed that some terpenoids obtained from *Plexaurella grisea* and *P. kukenthali* exhibited cytotoxic activity against myeloid leukemia and P-388 cancer cell lines.

From analyses of the VIP and the OPLS score plot (Fig. 5), it was established that the most important characteristics in the separation of extracts that showed cytotoxic activity were B2_5118, B4_4965 and B4_7686. Based on their ¹H-NMR chemical signals, it was possible to make a putative identification of asperdiol in the extract of *Plexaurella* sp. due to the signals evident in the extract; these signals agree with the chemical shifts for this compound reported by Weinheimer and Matson (1977)[59]. Chemical shifts at 4.76 and 4.95 ppm correspond to an exomethylene group characteristic of the cembrane skeleton. Signals located at 4.50 and 4.05 ppm are characteristic of carbinolic methine and methylene groups. Additionally, three signals, 1.20, 1.62 and 1.77 ppm, were assigned to three methyl groups, and the signal located at 2.70 ppm corresponds to an epoxy proton characteristic of asperdiol. The remaining chemical shifts were consistent with those reported in the literature[50]. In addition, chemical shifts assigned to the marker B2_5118 confirmed that the compound plexaurolone was present in the extract of *P. kukenthali*, in agreement with NMR data reported for this compound[47]. The signal located at 2.53 ppm was assigned to a typical methine that supports the isopropenyl group in the cembrane skeleton. Additionally, the signals at 0.99, 1.03 and 1.04 ppm were in agreement with three methyl groups, and the signal at 4.72 ppm together with the signal 1.68 ppm were assigned to an isopropenyl group typical in this kind of nucleus. All the remaining signals were consistent with data reported in the literature[47].

Finally, the main prospective application of metabolomic analyses using PSYCHE NMR experiments in soft corals was identifying metabolites with potential cytotoxic activity. This metabolomic approach may be useful in various scenarios but mostly in untargeted studies in which the identification of compounds is a challenge that involves a long time isolating and identifying biologically active metabolites. In addition, the advantage of using the decoupled PSYCHE experiments over the ¹H-NMR experiments due to the overlapping of signals present in the latter is corroborated.

Conclusion

In this study, it was possible to correlate the chemical compositions of extracts from soft coral with their cytotoxic activity against the tumor cell line PC3 using a metabolomics workflow and PSYCHE NMR experiments. By using this approach, it was possible to resolve the overlapping of ¹H-¹H scalar coupling multiplets, yielding an adequate matrix for reliable statistical and chemical shift analyses of complex natural products. Additionally, a preliminary identification of features responsible for the separation of the groups of extracts was possible due

to the chemical shifts observed in the VIP analysis, which were deemed the most important projection variables. The PSYCHE NMR experiment, combined with metabolomics studies, allowed the development of a procedure/methodology to establish which extracts from complex biological samples were most active and allowed the identification of compounds responsible for the activity. Extracts from *P. kukenthali* and *Plexaurella* sp. were the most cytotoxic and were responsible for separation between the groups. Asperdiol and plexaurolone were the compounds responsible for the cytotoxicity exhibited by the most-active extracts.

Materials and Methods

Materials.   Methanol and dichloromethane used for extraction were bought from Merck (Darmstadt, Germany). Cell culture reagents, D-MEM (Dulbecco's Modified Eagle Medium (1×), RPMI 1640 Roswell Park Memorial Institute, Darmstadt, Germany were made by Gibco/Invitrogen, Paisley, UK. Other reagents were Fetal bovine serum (FBS), Eurobio brand (Les Ulis, France), trypticase soy broth (TSB) and trypticase soy agar (TSA) Scharlau Co. brand (Barcelona, Spain). All cancer cell lines were acquired from ATCC, PC3 human prostatic carcinoma (ATCC® CRL1435™), SiHa human cervical carcinoma (ATCC® HTB-35™) and A549 human lung adenocarcinoma (ATCC® CCL-185™).

Methods.   *Soft coral collection and identification.*   Small portions (30 cm) of soft corals (N = 24) (Table S2, Supplementary Information) were collected by SCUBA diving at Punta Venado (11°16.26′ 87″N, 74°12.24′58″W), Santa Marta, Colombian Caribbean. Samples were collected at a depth range between 10–20 m. Collected samples were stored in dry ice and transported to the laboratory. These were kept frozen at −80 °C until extraction.

Samples were identified by colony morphology and sclerite shape, dimensions and distribution. Sclerites were obtained from a distal fragment of each soft coral portion that was treated with 5% sodium hypochlorite. Once the organic material was dissolved, sclerites were observed under the microscope. A voucher of each sample is stored at the collection of the Instituto de Ciencias Naturales (ICN) of Universidad Nacional de Colombia (Bogotá, Colombia) (Table S2 in Supplementary Information).

General experimental procedures.   All NMR data were acquired using nonspinning samples on a 600 MHz Agilent DD2 NMR at 25 °C equipped with a 5 mm C13 enhanced HCN cold probe. Signals were referenced in ppm in reference to the residual solvent signals (CDCl$_3$, at δH 7.26). The ^{1}H NMR spectra were acquired with 32 transients, a 1 s relaxation delay (d1), a 90-degree pulse of 6.70 μs and a 9.6 kHz (16 ppm) spectral window (sw). The transmitter offset was set to 6 ppm.

The PSYCHE (pure shift yielded by chirp excitation) spectra were acquired with 6 transients, a 1 s relaxation delay (d1), steady state scans of 2 (ss) and a tauPS of 6.3 ms, 1/sw1 = 0.12 μsec. The spectral window was set to 6 kHz (10 ppm) with the transmitter offset at 3 ppm. A wurst 180 pulse with a 6.0 deg flip angle, a 531(1.0 G/cm) gradient and a pulse width of 30.0 ms was used. A Grad-90-Grad option for the PSYCHE steady state with a G-strength of 6372 and G-time of 2 ms was used along with the Echo gradients (Encode Grad (5.0 G/cm)) using a 1.0 ms time, a 1.5 ms recovery time and a strength of 2659.

Sample extraction.   Prior to extraction, each soft coral fragment was ground. Subsequently, 10 g of dried powder of each soft coral sample was extracted at room temperature with a mixture of solvents (DCM/MeOH, 1:1) three times (300 mL) using an ultrasonic bath for 20 minutes. Remaining debris was removed by centrifugation twice at 12,000×g for 5 minutes, the solvent was then evaporated, and the dried samples were passed through a C18 cartridge with MeOH to remove salts. Extracts were concentrated under vacuum using rotary evaporation.

Cytotoxicity assay.   Tumor cell lines A549, SiHa and PC3 were cultured in DMEM supplemented with 10% heat-inactivated fetal bovine serum (FBS) and penicillin-streptomycin (1%) at 37 °C in a 5% CO$_2$ humidified atmosphere until 100% cell confluence was achieved[46].

The *in vitro* cytotoxicity of soft coral extracts was evaluated using the MTT method following Mosmann (1983) with modifications by Denizot and Lang (1986). This is a colorimetric assay based on the capacity of mitochondrial succinate dehydrogenase enzymes, found in living cells, to reduce the yellow, water-soluble substrate 3-(4,5-dimethyl thiazol-2-yl)-2,5-diphenyl tetrazolium bromide (MTT) into an insoluble, colored, formazan product, which is measured spectrophotometrically. Cells were grown and left to attach in 96-well plates (4.0E + 04 cells/well) for 48 hours.

A preliminary screening was performed using the 24 soft coral extracts evaluated at 20 μg/mL and DMSO at 0.1%, which was not cytotoxic to any of the cell lines. Then, the supernatant was removed, and 100 μL of 12 mM MTT solution in sterile PBS was added to each well and incubated at 37 °C for 4 hours. The solution was removed, and dimethyl sulfoxide (DMSO) was added to each well, followed by incubation at 37 °C for 15 minutes. Cell density was read in an iMarkTM Microplate Reader at a wavelength of 595 nm. Controls consisted of cells cultured without extract, and cells exposed to doxorubicin (25 ppm) were a positive control[60,61]. All tests were performed in triplicate. The cell viability percentage was calculated with Eq. (1), and the cell inhibition percentage was calculated with Eq. (2)[62]:

$$\%viability = \frac{\left(Abs_{sample}\right)}{Abs_{control}} * 100$$

(1)

where Abs$_{sample}$ is the absorbance of cells treated with the test extract, and Abs$_{control}$ is the absorbance of untreated cells.

$$\%cell\ inhibition = 100 - Cell\ Survival \qquad (2)$$

Metabolomic analyses. ¹H and PSYCHE NMR experiments of the 24 soft coral extracts were run in the NMRProcFlow v1.2 program[34]. This software performs all the spectral processing steps, including baseline correction, chemical shift calibration and alignment, and allows metabolic fingerprinting and targeted metabolomics[63].

All spectral data were compressed with a basic Zip format and processed in the NMRProcFlow v1.2 software, which allows the input of raw data in FID format. Data processing comprises calibration of the PPM scale, baseline correction, alignment, binning and scaling.

1. Calibration of the ppm scale was performed to adjust chemical shifts according to a known reference compound. The reference compound used for the calibration of the ¹H-NMR spectra was tetramethylsilane (TMS), with a chemical shift of $\delta = 0.00$ ppm, and for the PSYCHE experiment, the reference compound was the solvent $CDCl_3$, with a chemical shift of $\delta = 7.26$ ppm.
2. Baseline correction was performed using the global baseline correction algorithm[64], where the correction level was chosen as the 'soft' level. For increased efficiency of the method, spectral level noise ranges between 10.2 and 10.5 ppm for the ¹H-NMR experiment and between 7.7 and 8.0 ppm for the PSYCHE experiment were considered.
3. Alignment steps were very tedious to solve. Misalignments are a result of changes in the chemical shifts of NMR peaks largely due to differences in pH, ionic strength or other physicochemical interactions[65]. To perform this step, we use the algorithm based on least squares.
4. Binning. An NMR spectrum may contain several thousands of variables. Binning is used to reduce data dimensionality. When binning, spectra are divided into bins (so-called buckets), and the total area (sum of each resonance intensity) within each bin is calculated to represent the original spectrum. The approach we chose was the adaptive 'Intelligent Binning' method[66]. This allowed us to split the spectra so that each area, common to all spectra, contained the same single resonance, i.e., belonged to the same metabolite.
5. Normalization. Before bucket data export, to make all spectra comparable with each other, variations in the total concentrations of samples must be considered. We used the constant sum normalization, which consists of normalizing the total intensity of each individual spectrum to the same value[67]. After the data matrices of each of the selected regions were exported, they were used for statistical analysis using MetaboAnalyst version 3.0[35].

Multivariate data analysis. All variables of the data matrix were scaled using a self-scaling algorithm prior to multivariate data analysis. Principal component analysis (PCA), an unsupervised pattern recognition tool that explains the maximum amount of variation inherent in a multidimensional dataset[68], was carried out to detect patterns in the variables matrix and to detect outliers.

Orthogonal partial least squares discriminate analysis (OPLS-DA) is a supervised pattern recognition technique that aims at finding the maximum separation between a priori established groups[69]. Therefore, OPLS-DA was applied to discriminate spectral data obtained from the ¹H and PSYCHE NMR experiments and to discriminate between groups that exhibited or did not exhibit cytotoxicity.

The resulting models were evaluated using both R2 and Q2 metrics. R2 values reported the total amount of variance explained by the model in both the data (R2X) and independent variables (R2Y). Q2 reported the model accuracy and was calculated by cross-validation[70].

References

1. Johnson, C. H., Ivanisevic, J. & Siuzdak, G. Metabolomics: Beyond biomarkers and towards mechanisms. *Nat. Rev. Mol. Cell Biol.* **17**, 451–459 (2016).
2. Pontes, J. G. M., Brasil, A. J. M., Cruz, G. C. F., de Souza, R. N. & Tasic, L. NMR-based metabolomics strategies: plants, animals and humans. *Anal. Methods* **9**, 1078–1096 (2017).
3. Farag, M. A. *et al.* Soft Corals Biodiversity in the Egyptian Red Sea: a Comparative MS and NMR Metabolomics Approach of Wild and Aquarium Grown Species. *J. Proteome Res.* **15**, 1274–1287 (2016).
4. Goulitquer, S., Potin, P. & Tonon, T. Mass spectrometry-based metabolomics to elucidate functions in marine organisms and ecosystems. *Mar. Drugs* **10**, 849–80 (2012).
5. Miljanich, G. P. Ziconotide: neuronal calcium channel blocker for treating severe chronic pain. *Curr. Med. Chem.* **11**, 3029–3040 (2004).
6. Ebada, S. S., Edrada, R. A., Lin, W. & Proksch, P. Methods for isolation, purification and structural elucidation of bioactive secondary metabolites from marine invertebrates. *Nat. Protoc.* **3**, 1820–1831 (2008).
7. Younes, A. *et al.* Brentuximab vedotin (SGN-35) for relapsed CD30-positive lymphomas. *N. Engl. J. Med.* **363**, 1812–1821 (2010).
8. Cortes, J., Montero, A. & Glück, S. Eribulin mesylate, a novel microtubule inhibitor in the treatment of breast cancer. *Cancer Treat. Rev.* **38**, 143–151 (2012).
9. Sagar, S., Kaur, M. & Minneman, K. P. Antiviral lead compounds from marine sponges. *Mar. Drugs* **8**, 2619–2638 (2010).
10. Correa, H., Valenzuela, A. L., Ospina, L. F. & Duque, C. Anti-inflammatory effects of the gorgonian Pseudopterogorgia elisabethae collected at the Islands of Providencia and San Andrés (SW Caribbean). *J. Inflamm. (Lond).* **6**, 5 (2009).
11. Amaya García, F. *et al.* Dolabellane diterpenes from the Caribbean soft corals Eunicea laciniata and Eunicea asperula and determination of their anti HSV-1 activity. *Rev. Colomb. Química* **46**, 5 (2017).
12. Reina, E., Ramos, F. A., Castellanos, L., Aragón, M. & Ospina, L. F. Anti-inflammatory R-prostaglandins from Caribbean Colombian soft coral Plexaura homomalla. *J. Pharm. Pharmacol.* **65**, 1643–1652 (2013).

13. Van Der Greef, J., Stroobant, P. & Van Der Heijden, R. The role of analytical sciences in medical systems biology. *Curr. Opin. Chem. Biol.* **8**, 559–565 (2004).
14. Patel, K., Patel, J., Patel, M., Rajput, G. & Patel, H. Introduction to hyphenated techniques and their applications in pharmacy. *Pharm. Methods* **2**, 1–13 (2010).
15. Pettersson, L. & Elvingson, K. Studies of Vanadate-Organic Ligand Systems Using Potentiometry and NMR Spectroscopy. *ACS Symp. Ser.* 30–50, https://doi.org/10.1021/bk-1998-0711.ch002 (1998).
16. Panda, A., Parida, A. K. & Rangani, J. Advancement of Metabolomics Techniques and Their Applications in Plant Science: Current Scenario and Future Prospective. in Plant Metabolites and Regulation Under Environmental Stress 1–36, https://doi.org/10.1016/B978-0-12-812689-9.00001-7 (Elsevier Inc., 2018).
17. Lages, B. G., Fleury, B. G., Ferreira, C. E. L. & Pereira, R. C. Chemical defense of an exotic coral as invasion strategy. *J. Exp. Mar. Bio. Ecol.* **328**, 127–135 (2006).
18. Han, A. R. *et al.* Cytotoxic constituents of the octocoral *Dendronephthya gigantea*. *Arch. Pharm. Res.* **28**, 290–293 (2005).
19. Mahrous, E. A. & Farag, M. A. Two dimensional NMR spectroscopic approaches for exploring plant metabolome: A review. *J. Adv. Res.* **6**, 3–15 (2015).
20. Farag, M. A., Fekry, M. I., Al-hammady, M. A., Khalil, M. N. & Wessjohann, L. A. Cytotoxic Effects of *Sarcophyton* sp. Soft Corals — Is There a Correlation to Their NMR Fingerprints? *Mar. Drugs* **15**, 1–13 (2017).
21. Foroozandeh, M. *et al.* Ultrahigh-resolution NMR spectroscopy. *Angew. Chemie - Int. Ed.* **53**, 6990–6992 (2014).
22. Straßburger, D. *et al.* Hadamard homonuclear broadband decoupled TOCSY NMR: Improved effica- cy in detecting long-range chemical shift correlations. *ChemPhysChem* **17**, 4037–4042 (2016).
23. Santacruz, L., Thomas, O. P., Duque, C., Puyana, M. & Tello, E. Comparative analyses of metabolomic fingerprints and cytotoxic activities of soft corals from the Colombian Caribbean. *Mar. Drugs* **17**, 1–14 (2019).
24. De Ketelaere, B., Hubert, M. & Schmitt, E. Overview of PCA-based statistical process-monitoring methods for time-dependent, high-dimensional data. *J. Qual. Technol.* **47**, 318–335 (2015).
25. Tistaert, C. *et al.* Multivariate data analysis to evaluate the fingerprint peaks responsible for the cytotoxic activity of Mallotus species. *J. Chromatogr. B* **910**, 103–113 (2012).
26. Triba, M. N. *et al.* Molecular BioSystems PLS/OPLS models in metabolomics: the impact of permutation of dataset rows on the K-fold. *Mol. Biosyst.* **11**, 13–19 (2014).
27. Chiu, C. Y. *et al.* Metabolomics reveals dynamic metabolic changes associated with age in early childhood. *PLoS One* **11**, 14 (2016).
28. Szymanska, E., Saccenti, E., Smilde, A. K. & Westerhuis, J. A. Double-check: validation of diagnostic statistics for PLS-DA models in metabolomics studies. *Metabolomics* **8**, 14 (2012).
29. Cho, H. W. *et al.* Discovery of metabolite features for the modelling and analysis of high-resolution NMR spectra. *Int. J. Data Min. Bioinform.* **2**, 176–192 (2008).
30. Abdullah, M. *et al.* Non-targeted metabolomics analysis of Golden Retriever Muscular Dystrophy-affected muscles reveals alterations in arginine and proline metabolism, and elevations in glutamic and oleic acid *in vivo. Metabolites* **7**, 1–19 (2017).
31. Rantalainen, M., Cloarec, O., Nicholson, J. K., Holmes, E. & Trygg, J. OPLS discriminant analysis: Combining the strengths of PLS-DA and SIMCA classification. *J. Chemiometrics* **20**, 341–351 (2006).
32. Blaženović, I., Kind, T., Ji, J. & Fiehn, O. Software tools and approaches for compound identification of LC-MS/MS data in metabolomics. *Metabolites* **8**, 31 (2018).
33. Rochat, B. Proposed Confidence Scale and ID Score in the Identification of Known-Unknown Compounds Using High Resolution MS Data. *J. Am. Soc. Mass Spectrom.* **28**, 709–723 (2017).
34. Jacob, D., Deborde, C., Lefebvre, M., Maucourt, M. & Moing, A. NMRProcFlow: a graphical and interactive tool dedicated to 1D spectra processing for NMR-based metabolomics. *Metabolomics* **13**, 36 (2017).
35. Xia, J., Sinelnikov, I., Han, B. & Wishart, D. MetaboAnalyst 3.0–making metabolomics more meaningful. *Nucleic Acids Res.* **43**, 251–257 (2015).
36. Olsen, E. K. *et al.* Cytotoxic activity of marine sponge extracts from the sub-Antarctic Islands and the Southern Ocean. *S. Afr. J. Sci.* **112**, 1–5 (2016).
37. Rasmussen, S., Lane, G. A. & Villas-bo, S. G. Metabolomics or metabolite profiles? Sci. *Direct* **23**, 385 (2005).
38. Moltu, S. J. *et al.* Urinary metabolite profiles in premature infants show early postnatal metabolic adaptation and maturation. *Nutrients* **6**, 1913–1930 (2014).
39. Powers, R. The current state of drug discovery and a potential role for NMR metabolomics. *J. Med. Chem.* **57**, 5860–70 (2014).
40. Worley, B. & Powers, R. Multivariate Analysis in Metabolomics. *Curr. Metabolomics* **1**, 92–107 (2012).
41. Fay, M. P. & Shih, J. H. Permutation Tests Using Estimated Distribution Functions. *J. Am. Stat. Assoc.* **93**, 387–396 (2014).
42. Gromski, P. S. *et al.* A tutorial review: Metabolomics and partial least squares-discriminant analysis - a marriage of convenience or a shotgun wedding. *Anal. Chim. Acta* **879**, 10–23 (2015).
43. Yang, Z. *et al.* Exploring potential biomarkers of early stage esophageal squamous cell carcinoma in pre-and post-operative serum metabolomic fingerprint spectrum using 1 H-NMR method. *Am. J. Transl. Res.* **11**, 819–831 (2019).
44. Sponholtz, W. R., Bianco, M. A. & Gribble, G. W. Isolation and structure determination of the cembranoid eunicin from a new genus of octocoral, Pseudoplexaura. *Nat. Prod. Res.* **22**, 441–448 (2008).
45. Sawant, S. S. *et al.* Bioactive rearranged and halogenated semisynthetic derivatives of the marine natural product sarcophine. *J. Nat. Prod.* **67**, 2017–2023 (2004).
46. Yang, S.-F., Weng, C.-J., Sethi, G. & Hu, D.-N. Natural Bioactives and Phytochemicals Serve in Cancer Treatment and Prevention. *Evidence-Based Complementary and Alternative Medicine* **2013**, 112 (2013).
47. Tello, E. *et al.* Absolute stereochemistry of antifouling cembranoid epimers at C-8 from the Caribbean octocoral *Pseudoplexaura flagellosa*. Revised structures of plexaurolones. *Tetrahedron* **67**, 9112–9121 (2011).
48. Markley, J. L. *et al.* The future of NMR-based metabolomics. *Curr. Opin. Biotechnol.* **43**, 34–40 (2017).
49. Rodríguez, A. D. The natural products chemistry of West Indian gorgonian octocorals. *Tetrahedron* **51**, 4571–4618 (1995).
50. Weinheimer, A. & Matson, J. Marine Anticancer Agents: Asperdiol, Cembranoid Prom Gorgonians, Eunicea Asperula And E. Tourneforti. *Tetrahedron Lett.* **18**, 1295–1298 (1977).
51. Creek, D. *et al.* Metabolite identification: are you sure? And how do your peers gauge your confidence? *Metabolomics* **10**, 350–353 (2014).
52. Hegazy, M. E. F. *et al.* Cembrene diterpenoids with ether linkages from sarcophyton ehrenbergi: An anti-proliferation and molecular-docking assessment. *Mar. Drugs* **15**, 2–15 (2017).
53. Zhang, C. *et al.* Cytotoxic diterpenoids from the soft coral Sarcophyton crassocaule. *J. Nat. Prod.* **69**, 1476–1480 (2006).
54. Kim, H. K., Choi, Y. H. & Verpoorte, R. NMR-based metabolomic analysis of plants. *Nat. Protoc.* **5**, 536–549 (2010).
55. Huang, Y. *et al.* A Pure Shift-Based NMR Method for Transverse Relaxation Measurements on Complex Samples. *IEEE Trans. Instrum. Meas.* **PP**, 1–11 (2019).
56. He, Q. *et al.* NMR-based metabolomic analysis of spatial variation in soft corals. *Mar. Drugs* **12**, 1876–1890 (2014).
57. Honda, A., Mori, Y., Iguchi, K. & Yamada, Y. Antiproliferative and cytotoxic effects of newly discovered halogenated coral prostanoids from the Japanese stolonifer Clavularia viridis on human myeloid leukemia cells in culture. *Mol. Pharmacol.* **32**, 530–535 (1987).

58. Rueda, A., Zubía, E., Ortega, M. J. & Salvá, J. Structure and cytotoxicity of new polyhydroxylated sterols from the Caribbean gorgonian Plexaurella grisea. *Steroids* **66**, 897–904 (2001).
59. Weinheimer, A. J. & Matson, J. A. Marine Anticancer Agents: Asperdiol, Cembranoid from Gorgonians, Eunicea Asperula And E. Tourneforti. *Tetrahedron* 2926 (1977).
60. Al-Ghamdi, S. S. Time and dose dependent study of doxorubicin induced DU-145 cytotoxicity. *Drug Metab. Lett.* **2**, 47–50 (2008).
61. Shaikh, K. S., Pawar, A., Aphale, S. R. & Moghe, A. S. Effect of vesicular encapsulation on *in-vitro* cytotoxicity of ciclopirox olamine. *Int. J. Drug Deliv.* **4**, 139–146 (2012).
62. Patel, S., Gheewala, N., Suthar, A. & Shah, A. *In-Vitro* cytotoxicity activity of *Solanum Nigrum* extract against Hela cell line and Vero cell line. Int. *J. Pharm. Pharm. Sci.* **1**, 38–47 (2009).
63. Sogin, E. M., Anderson, P., Williams, P., Chen, C. S. & Gates, R. D. Application of 1H-NMR metabolomic profiling for reef-building corals. *PLoS One* **9**, 4–8 (2014).
64. Bao, Q. *et al.* A new automatic baseline correction method based on iterative method. *J. Magn. Reson.* **218**, 35–43 (2012).
65. Vu, T. N. *et al.* An integrated workflow for robust alignment and simplified quantitative analysis of NMR spectrometry data. *BMC Bioinformatics* **12**, 14 (2011).
66. Anderson, P. E. *et al.* Dynamic adaptive binning: An improved quantification technique for NMR spectroscopic data. *Metabolomics* **7**, 179–190 (2011).
67. Kohl, S. M. *et al.* State-of-the art data normalization methods improve NMR-based metabolomic analysis. *Metabolomics* **8**, 146–160 (2012).
68. Jolliffe, I. T. & Cadima, J. Principal component analysis: a review and recent developments. *R. Soc.* **16** (2016).
69. Dona, A. C. *et al.* A guide to the identification of metabolites in NMR-based metabonomics/metabolomics experiments. *Comput. Struct. Biotechnol. J.* **14**, 135–153 (2016).
70. Lussu, M. *et al.* 1H NMR spectroscopy-based metabolomics analysis for the diagnosis of symptomatic E. coli-associated urinary tract infection (UTI). *BMC Microbiol.* **17**, 1–8 (2017).
71. Chong, J. *et al.* MetaboAnalyst 4.0: Towards more transparent and integrative metabolomics analysis. *Nucleic Acids Res.* **46**, W486–W494 (2018).
72. Hostettman, K. Methods in Plant Biochemistry. Assays for Bioactivity. *New Phytol.* **119**, 474 (1991).

Acknowledgements

Author contributions

L.S., O.P.T. and E.T. conceived and designed the experiments; L.S., M.P. and E.T. collected the biological material; M.P. identified the soft corals, L.S., R.D., D.X. and E.T. performed chemical experiments; L.S. and E.T. analyzed the data; L.S. and E.T. wrote the article; L.S., M.P., O.P.T. and E.T. edited and reviewed the article. All the authors have contributed to and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41598-020-62413-0.

Correspondence and requests for materials should be addressed to E.T.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.