



Kamal, Nurkhalida and Viegelmann, Christina V. and Clements, Carol J. and Edrada-Ebel, RuAngelie (2016) Metabolomics-guided isolation of anti-trypanosomal metabolites from the endophytic fungus *Lasiodiplodia theobromae*. *Planta Medica*. ISSN 0032-0943 , <http://dx.doi.org/10.1055/s-0042-118601>

This version is available at <https://strathprints.strath.ac.uk/59826/>

Strathprints is designed to allow users to access the research output of the University of Strathclyde. Unless otherwise explicitly stated on the manuscript, Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Please check the manuscript for details of any other licences that may have been applied. You may not engage in further distribution of the material for any profitmaking activities or any commercial gain. You may freely distribute both the url (<https://strathprints.strath.ac.uk/>) and the content of this paper for research or private study, educational, or not-for-profit purposes without prior permission or charge.

Any correspondence concerning this service should be sent to the Strathprints administrator: strathprints@strath.ac.uk

Metabolomics-guided isolation of anti-trypanosomal metabolites from the endophytic fungus *Lasiodiplodia theobromae*

Nurkhalida Kamal^{1,2}, Christina V. Viegelmann¹, Carol J. Clements¹, RuAngelie Edrada-Ebel¹

Affiliation

¹Strathclyde Institute of Pharmacy and Biomedical Sciences, University of Strathclyde, The John Arbuthnott Building, 161 Cathedral Street, Glasgow G4 0RE, United Kingdom

²Faculty of Pharmacy, Cyberjaya University College of Medical Sciences (CUCMS), Cyberjaya, 63000, Selangor, Malaysia

Correspondence Authors:

RuAngelie Edrada-Ebel,

Email: ruangelie.edrada-ebel@strath.ac.uk

Telephone: +44-(0)-141-548-5968 Fax: +44-(0)-141-552-2562

Nurkhalida Kamal

Email: minnur_27@yahoo.com Telephone: +44-(0)-141-548-3728

Abstract

Fungal endophytes offer diverse and unique secondary metabolites, making these organisms potential sources of promising drug leads. The application of High Resolution-Liquid Chromatography Mass Spectrometry (HR-LCMS) and Nuclear Magnetic Resonance (NMR) -based metabolomics to fungal endophytes is practical in terms of dereplication studies and the mining of bioactive compounds. In this paper we report the application of metabolomics in parallel with anti-trypanosomal assays to determine the ideal conditions for the medium-scale fermentation of the endophyte *Lasiodiplodia theobromae*. The ¹H NMR comparison between the active versus inactive fractions identified several unique chemical fingerprints belonging to the active fractions. Furthermore, by integrating HR-LCMS data with chemometric data analysis, such as partial least squares-discriminant analysis (PLS-DA), orthogonal PLS-DA (OPLS-DA), and S-plots and the bioactivity results of the fractions of *L. theobromae*, the most predictive anti-trypanosomal agents were easily discerned. With available databases such as Antibase and MarinLit coupled to MZmine by *in-house* algorithms optimized in our laboratory, the predictive metabolites were readily identified prior isolation. Fractionation was performed on one of the active fractions and three known compounds were isolated, namely cladospirone B, desmethyl-lasiodiplodin, and R-(-) mellein. Cladospirone B and desmethyl-lasiodiplodin were among the predictive compounds generated by the S-plot, and these compounds exhibited good activity against *Trypanosoma brucei brucei* with MICs of 17.8 μM and 22.5 μM, respectively. Therefore, HR-LCMS and NMR-based metabolomics proved to be a powerful decision-making tool in mining active metabolites of *L. theobromae* against *T. b. brucei*.

Key words

Fungal endophytes, *Lasiodiplodia theobromae*, Botryosphaeriaceae, anti-trypanosomal activity, HR-LCMS, NMR, metabolomics, chemometrics

Introduction

Despite the increasing disinclination of the pharmaceutical industry to pursue natural products in their pipelines, statistical findings show that natural products still play a major role in drug discovery with more than 50% of FDA-approved drugs derived from natural products/mimics [1]. However, natural products research has been found to be too laborious, time-consuming and uneconomical, which may have led to the declining trend. Nevertheless, with the emergence of new and more advanced technologies such as genomics, transcriptomics, proteomics, metabolomics and bioinformatics, natural products research has become more competent in finding promising novel drugs for the pipeline [1-4].

The utilization of metabolomics in natural products research is increasingly powerful in several perspectives. Metabolomics is defined as a global study of all or a subset of chemical entities including either or both primary and secondary metabolites, that are present in living organisms (cells or tissues) under certain growth conditions [5-7]. Metabolomics, or metabolome mining, in natural products research has been used for dereplication studies of both known and new compounds in crude plant, marine or microbial extracts [8-18], in differentiating biologically active natural products (NPs) from non-active fractions [19-23], optimising the production of bioactive secondary metabolites, as well as in developing cultivation processes for large-scale fermentation and understanding their biosynthetic pathways [7,24].

Endophytic fungi are microorganisms that mutually live inside plant tissues without causing any immediate negative effects towards the host plant for at least a part of

fungal life cycle [25]. The total population of endophytic fungi species has been estimated to be up to 1.3 million [26]. Another study places this estimate at approximately 1.5 million [27]. However, as of the year 2000, only 75 000 fungal species have been identified; the remainder are still untapped and unexplored [28]. Bioactive NPs derived from endophytic fungi display wide ranges of activities, such as (–)-oxysporidinone, (2,6-dihydroxyphenyl)pentan-1-one and (Z)-1-(2-(2-butyryl-3-hydroxyphenoxy)-6-hydroxyphenyl)-3-hydroxybut-2-en-1-one exhibited antimicrobial [29,30], pullularin A and hinnuliquinone displayed antiviral [31,32], and spiropreussione A and 9-deacetyfumigaclavine C for anticancer activity [33,34]. Compounds like cochlioquinone A, isocochlioquinone A and cercosporin exhibited activity against neglected tropical diseases [35,36].

Human African trypanosomiasis or sleeping sickness is a fatal vector-borne parasitic disease caused by *Trypanosoma b. brucei* and transmitted by the tsetse fly (*Glossina* spp.). This neglected tropical disease occurs only in rural areas of sub-Saharan Africa (Simarro, 2011). To date only a few drugs have been approved for the treatment of Human African trypanosomiasis. These include suramin, pentamidine, melarsoprol, eflornithine and the combination of nifurtomox/eflornithine. Most of the drugs are old, having been discovered in the 1940s and 1950s, and have **adverse effects** such as nausea, vomiting, fatigue, seizures, fever, diarrhea, hypoglycemia, abdominal cramping, peripheral neuropathy, hypertension, heart damage and neutropenia on the patients. For this reason, mining and developing new Human African trypanosomiasis drugs from natural products is crucial and essential **because various natural sources including plants, microorganisms, animals, and marine organisms offer high number of NPs with**

diverse chemical structures and novel pharmacological mechanism of action (Jacobs, 2011).

The aim of this study is to adopt an untargeted HR-LCMS (High Resolution-Liquid Chromatography Mass Spectrometry) and NMR-based metabolomics approach to determine the optimal conditions of *Lasiodiplodia theobromae* for medium-scale fermentation and also to capture and trace active anti-trypanosomal metabolites by using statistical multivariate data analysis of generated HR-LCMS data, such as Principal Component Analysis (PCA), Orthogonal Partial Least Squares-Discriminant and Analysis (OPLS-DA). To verify the dereplication results obtained from HR-LCMS data, 1D and 2D ¹H NMR data were utilized. Finally, the isolation of active metabolites was performed based on the outcome of HR-LCMS and NMR-based metabolomics profile data.

Results and discussion

In search of the best condition for scaling up the culture of the endophyte *L. theobromae* obtained from the leaves of *V. pinnata*, HR-LCMS and NMR-based metabolomics along with the bioassay data were utilized. The fungus was grown in solid rice cultures and liquid Wickerham cultures for seven, fifteen and thirty days, after which the metabolites were extracted. The three different incubation times were chosen based on the fungal life-cycle [39]. In this case, the first seven days represent the germination phase, while the fifteen-day and thirty-day cultures cover the hyphal growth stage and sporing phase, respectively, of *L. theobromae*. The production of secondary metabolites was monitored using HR-LCMS and NMR at each of the growth phases parallel to the bio-

assay results. Based on the bioassay results (Table 1), the 30-days rice culture extract exhibited the strongest activity against *T. b. brucei* with an MIC of less than 25 µg/ml. The HR-LCMS raw data was processed using MZMine 2.10 [40]. The results of the assay were reproducible between the scale-up batches. Metabolite production and distribution between cultures were analyzed through ion peak scatter plots (Fig. 1). Based on the MS data, the occurrences of the metabolites on the 7th and 15th days were similar while a decrease in metabolite production was observed on the 30th day. In addition, the ion chromatogram both in positive and negative mode revealed a different set of metabolites on the 30th day to those of the 7th and 15th-day extracts (Fig. 1). Moreover, the ¹H NMR data revealed findings similar to the MS data (Fig. 2). Therefore, the 30-day rice culture condition was chosen for scale-up and further isolation work.

Table 1

Fig. 1

Fig. 2

The medium-scale 30-day rice culture extract of *L. theobromae* was fractionated yielding 19 fractions (LT-1 until LT-19). These were submitted for the anti-trypanosomal activity and ¹H NMR metabolomics. The non-polar fractions LT-2 to LT-8 exhibited strong bioactivity, except for LT-4 which showed only moderate activity against *T. b. brucei* (Fig. 3). LT-1 was excluded from the bioassay screening because it contained only fatty acids as indicated by its ¹H NMR data (Fig. 4). The ¹H NMR data of the 19 fractions (Fig. 4) were analyzed, and unique chemical fingerprints of the active frac-

tions were detected. Among these active fractions, two distinctive sub-groups, LT-2 to LT-4 and LT-6 to LT-8, clustered together as they shared similar spectral data. Fraction LT-5 was a mixture of both groups. The ^1H NMR spectra of fractions LT-2 to LT-5 displayed a pair of *meta*-coupled aromatic protons at δ_{H} 6.23 and 6.28 ($J= 2.7$ Hz) and a multiplet peak at δ_{H} 5.15 which may be an olefinic or oxygenated methine. The upfield shift of the *meta*-coupled aromatic protons at the 6 ppm region suggested the presence of an electron-withdrawing group such as a hydroxyl or halogen substituent. In fractions LT-5 to LT-8 two *meta*-coupled doublets at δ_{H} 6.18 and 6.25 ($J=2.6$ Hz) were observed (see red arrow in Fig. 4). The proton signals between 4.2- 4.8 ppm revealed the presence of oxygenated methines while proton signals between 6.6 to 7.7 ppm indicated the presence of aromatic compounds in fractions LT 6 to LT8.

Fig. 3

Fig. 4

Supervised methods of multivariate data analysis were used to analyze the similarity of the data sets between 19 samples. Principal component analysis (PCA) was used in earlier step to observe an overview of variance between the fractions and metabolites generated from MS data and also to identify any outliers. The distribution difference of the type of metabolites between active vs inactive fractions of *L. theobromae* against *T. b. brucei* was analyzed by subjecting the data to OPLS-DA. The results of the analysis led to the prediction of compounds that contribute towards the anti-trypanosomal activity of the fractions. For the OPLS-DA models (Fig. 5a), the MS-based metabolomics data set was assigned as the X independent variable while the fractions' anti-trypanosomal response was the Y dependent variable. The quality of OPLS-DA model was measured by

two parameters, R² (goodness of fit), and Q² (predictive ability) and the model's R²X was 0.261, R²Y was 0.955 and Q² was 0.733. These results were interpreted that 26.1% of the X variables could be used to describe 95.5% of the variation between active fractions and inactive fractions while this model has 73.3% of the average of prediction ability. The value of R²Y and Q² was greater than 50%, indicating a **well-fitted** model exhibiting good prediction (Robotti, 2016). The quality and robustness of OPLS-DA model was validated by a permutation test (n=100) with the Q²-intercept value was -0.368 (below 0.05) shows that the original model is statistically effective (Fig. 5b) {Haoula, 2015 #1012}. For the OPLS-DA **scores** plot (Fig. 5a), the active fractions were **clustered** together versus the inactive ones. Under the active group, fractions LT-2 to LT-5 (Group 1) and LT-6 to LT-8 (Group 2) clustered together indicating a shared set of metabolites while fractions LT-9 to LT-19 were observed as outliers of the group. The generated S-plot (Fig. 5d) determined the “end point” or unique compounds for each of the respective groups, indicating the metabolites that are potentially responsible for the bioactivity against *T. b. brucei* which discriminated the active from the inactive fractions of *L. theobromae*. Eight metabolites were identified from Antibase and MarinLit as shown in Table 2. **The end point** compounds were targeted for bioassay-guided isolation work with three active compounds in Group 1 and the remaining five active compounds in Group 2.

Fig. 5

Table 2

Fig. 6

In Group 1, the active metabolites listed from the dereplication step of the HR-LCMS data were identified as palmarumycin CP17 (**5**), cladospirone B (**6**) and desmethyl-lasiodiplodin (**8**). Further analysis of the ^1H - ^1H COSY NMR spectrum of fraction LT-3 revealed substructures belonging to the predicted active metabolites. For example, the correlation of the *meta*-coupled aromatic protons at δ_{H} 6.23 and 6.28, as well as that of the methyl doublet at δ_{H} 1.35 with one oxygenated methine proton at δ_{H} 5.15 which further correlated with the aliphatic chain were characteristics of the desmethyl-lasiodiplodin structure (**8**) (see Fig. 7). Substructures of palmarumycin CP17 and cladospirone B were elucidated in the same manner.

Fig. 7

A similar approach was used for Group 2. Fraction LT-7 was selected based on the OPLS-DA data which indicated the bioactive metabolite. Among these metabolites, the structure of the compound 6-oxo-de-O-methyl-lasiodiplodin (**2**) was confirmed by its COSY spectral data, which exhibited correlations similar to desmethyl-lasiodiplodin (**8**), as shown in Figure 8a. In the aromatic region, the COSY spectrum revealed correlations for preussomerin-C (Figure 8b).

Fig. 8

Isolation and purification of the compounds in Group 1 was performed by high-throughput MPLC. Three known compounds were isolated and elucidated based on their NMR and MS data as cladospirone B (**6**) [43], desmethyl-lasiodiplodin (**8**) [44], and R-(-)-mellein (**9**) [45]. The isolation of cladospirone B (**6**) and desmethyl-lasiodiplodin (**8**) confirmed the putative identification of the metabolites in Group 1 predicted earlier by the S-plot of the OPLS-DA model. R-(-)-mellein (**9**) is structurally close analogue of 6,8-dihydroxy-3-methylisocoumarin (**1**). All isolated compounds were tested against *T. b. brucei*, and cladospirone B (**6**) and desmethyl-lasiodiplodin (**8**) had MICs of 17.8 and 22.5 μM , respectively. All three metabolites were checked again in the S-plot and R-(-)-mellein was located in the middle of the plot suggesting that the anti-trypanosomal activity would be less (Fig. 9); this is indeed the case as confirmed by the bioassay results. Due to the limited amount of sample, however, palmarumycin CP17 was not isolated. Desmethyl-lasiodiplodin (**8**) has been known to exhibit anti-cancer activity against MCF-7 via apoptosis with an IC_{50} seven-fold more potent than its toxicity on normal cells [46]. On the other hand, cladospirone B (**6**) has been reported to be inactive in antibacterial and antifungal assays [45]. Further isolation and purification of five other active metabolites from Group 2 is still on-going.

Fig. 9

Table 3

In our screening program, we observed that the crude extract of *L. theobromae* from agar plates showed anti-trypanosomal activity against *T. b. brucei*. In order to find the

optimal conditions to grow *L. theobromae*, HR-LCMS-based metabolomics was applied, which resulted in the selection of solid rice culture for 30 days as the best conditions for medium-scale fermentation. Fractionation was performed on the crude extract, and based on the 1D ¹H NMR data comparison of 19 fractions several unique chemical fingerprints in the active fractions were highlighted. Furthermore, by utilizing the HR-LCMS data for chemometric analysis such as OPLS-DA, an S-plot and consequently, a set of the metabolites that were predicted to be active was generated. All predictive metabolites were easily identified with the aid of databases such as AntiBase and MarinLit, which were coupled to MZmine by *in-house* algorithms. The application of ¹H NMR and COSY allowed the detection of the predictive metabolites within the active fractions as well as the confirmation of the dereplication results obtained from the HR-LCMS data. Further isolation of secondary metabolites was performed on fraction LT-2 (active fraction in Group 1). Three known compounds were isolated and identified as cladospirone B (**6**), desmethyl-lasiodiplodin (**8**) and R-(-)-mellein (**9**). The isolation of cladospirone B (**6**) and desmethyl-lasiodiplodin (**8**) confirmed the putative identification of the metabolites in Group 1 predicted earlier by the S-plot of the OPLS-DA model. To our best knowledge, this is the first report of isolation of cladospirone B (**6**) from *L. theobromae*. It is also the first report to indicate the good anti-trypanosomal activity of cladospirone B (**6**) and desmethyl-lasiodiplodin (**8**) against *T. b. brucei* with MICs of 17.8 and 22.5 μM, in comparison with suramin (MIC value of 0.1 ± 0 μM) respectively. Our strategy of emphasizing HR-LCMS and NMR-based metabolomics to search for active anti-trypanosomal compounds has therefore proven to be effective. In conclusion, this study determined that the combination of HR-LCMS and NMR-based metabolomics is a powerful and advantageous decision-making tool in mining active metabolites of *L. theobromae*

against *T. b. brucei* and is also promising for implementation in other drug discovery programs elsewhere.

Materials and methods

Fungal sampling

The fungus *L. theobromae* was isolated from fresh healthy leaves of *Vitex pinnata* collected in April 2011 near Kuala Terengganu, Malaysia. The plant was identified by Dr. Nashriyah Mat from the Faculty of Bioresources and Food Industry, Universiti Sultan Zainal Abidin and voucher specimen was deposited (collection number VP 01). Samples were kept in zip lock bags and stored at 4°C until the isolation of endophytic fungi was performed four days later. The surfaces of the leaves and stems were sterilised with 70% iso-propanol for two minutes and subsequently rinsed in sterile water. Small tissue samples from inside the leaves and stems were cut aseptically and pressed onto agar plates (composition of isolation medium: 15 g/L malt extract, 15 g/L agar, and 0.2 g/L chloramphenicol in distilled water, pH 7.4-7.8, adjusted with 10% NaOH or 36.5% HCl). Chloramphenicol, ≥98% purity (Sigma Aldrich) was added to inhibit bacterial growth. The plates were left for a few days until fungal growth was observed. Reinoculation onto new malt agar plates was repeated several times until pure strains were attained.

Identification of fungal strains

The fungal strain was identified using DNA amplification and sequencing of the internal transcribed spacer (ITS) region as described previously [47]. The sequence data has been submitted to GenBank with accession number KC960898. The fungal strain

was identified as *Lasiodiplodia theobromae*. A voucher strain was submitted and kept at the Natural Product Metabolomics Laboratory, SIPBS.

Small and medium-scale fermentation

The fungal strain was cultivated on malt agar plates for seven days at 30°C. The colonies and agar were cut into small pieces and were placed in liquid or solid media. The liquid medium used was Wickerham medium, consisting of 3 g yeast extract, 3 g malt extract, 5 g peptone, 10 g glucose, distilled water added up to 1000 mL in 2L Erlenmeyer flasks, pH 7.2-7.4, adjusted with 10% NaOH or 36.5% HCl. The solid medium used was rice medium, with 100 g of long grain rice and 100 mL of distilled water autoclaved together in 1L Erlenmeyer flasks. The strain was grown for three different incubation times: seven, fifteen and thirty days, under static conditions at room temperature. For medium-scale fermentation, *L. theobromae* was cultivated in 1L flasks using the optimal conditions as determined by the small-scale cultures. In this case, the growth of the fungus on rice medium for 30 days under static conditions was determined to be the most favorable condition for the production of anti-trypanosomal metabolites.

Extraction and Isolation of Pure Compounds

The medium-scale rice cultures were extracted with ethyl acetate and homogenized as finely as possible using a T18 Basic Ultra-Turrax (IKA) at maximum speed. These were then kept overnight. The extract was subsequently dried under *vacuo* using a rotary evaporator (Buchi Labortechnik AG). The crude extract (5.0 g) was then partitioned with 10% *n*-hexane and 90% methanol to remove fatty acids. The extract containing methanol-soluble compounds (2.0 g) was collected for further isolation work. The

fractionation of the methanol extract was accomplished using Medium Pressure Liquid Chromatography (MPLC). Linear gradient elution was employed with hexane (A) and ethyl acetate (B) as the mobile phase at a flow rate of 20 mL/min. A pre-packed silica column (20–45 μm , 23 x 110 mm, Silica VersaPak cartridge) was used. It was connected to a Buchi Pump Manager C-615 coupled to binary pumps (Buchi Modules C-601). 100% A was run for 5 min, followed by 100% A to 100% B for 20 min, and finishing with 100% B for the last 5 min. The total run time was 30 min. Fractions were collected in collection tubes automatically every 2 mL using a fraction collector Frac 920 (GE Healthcare Bio-Sciences AB). Fractions with similar TLC profiles were pooled together, yielding a total of nineteen fractions. Fraction LT-2 (435 mg) was further subjected to MPLC on a pre-packed silica column (20–45 μm , 23 x 53 mm, Silica VersaPak cartridge) utilizing an isocratic gradient system (80% *n*-hexane : 20% ethyl acetate) for 30 min at a flow rate of 20 mL/min. In total, 300 collection tubes of fractions (2 mL in each tube) were collected automatically and using their TLC profiles. Similar fractions were pooled, resulting in 17 fractions and giving three known compounds **6** (3 mg), **8** (73 mg) and **9** (11 mg).

***In vitro* anti-trypanosomal assay**

The samples were prepared to a final concentration of 10 mg/mL (stock solution) by being dissolved in the appropriate amount of DMSO. To screen for *in-vitro* activity, a concentration of 200 $\mu\text{g/mL}$ was used. This was achieved by diluting the stock solution 1 in 10 with HMI-9 (drug solution). 4 μL of drug solution was transferred to 96 μL of HMI-9 in the 96-well plate. 100 μL of trypanosome suspension, consisting of *T. b. brucei* S427 blood stream form at 3×10^4 trypanosomes/mL, were then added to the 96-well plate to make the final concentration of the compounds ranged from 100 $\mu\text{g/mL}$ to

0.17 $\mu\text{g/mL}$. DMSO was used as the negative control (concentration of 1 to 0.002%) and suramin (Calbiochem-Novabiochem Co.) was selected as the positive control (concentration of 1 to 0.008 μM). The plate was incubated for 48 hours at 37°C, 5% CO_2 with a humidified atmosphere, after which 20 μL of Alamar blue was added. The plate was again incubated for another 24 hours under the same conditions. The fluorescence was measured using a Wallac Victor microplate reader (PerkinElmer) with excitation at 530 nm and emission at 590 nm. The results were calculated as percentages of control values. All samples that exhibited > 90% inhibition were selected for the minimum inhibitory concentration (MIC) assay to determine the MIC value.

NMR Instrumentation

One- and two-dimensional ^1H and ^{13}C NMR spectra were recorded at 400 MHz on a JEOL-LA400 FT-NMR spectrometer system with a 40TH5AT/FG probe (JEOL LTD.). Compounds **6** and **8** were reconstituted in deuterated chloroform (CDCl_3) while compound **9** was reconstituted in deuterated DMSO ($\text{DMSO-}d_6$).

HR-LCMS Procedure

HR-LCMS was measured using an Accela 600 HPLC pump with Accela autosampler and UV/Vis detector (Thermo Scientific) and Orbitrap Exactive mass spectrometer (Thermo Fisher Scientific Inc). Analysis of samples was done using similar protocols described previously [23,47,48].

HR-LCMS data processing

Initially the raw HR-LCMS data were sliced into two data sets according to ionization mode using the MassConvert tool from ProteoWizard (ProteoWizard Homepage) (<http://proteowizard.sourceforge.net/>). The sliced data were imported to MZMine 2.10 (<http://sourceforge.net/projects/mzmine/>), a software developed for the differential analysis of mass spectrometry data (Pluskal) The data processing step was performed in the same manner as explained by (Abdelmohsen, 2014) albeit with slightly modified parameters. In this analysis, the data set was crop-filtered from 0.1 to 35 min and the retention time normalizer was not applied because only one batch of data was used.

Statistical Analysis

MS spectral data were converted to an ASCII text file and imported to MS-Excel. The data was sorted to exclude background peaks that belonged to the MeOH blank. The sorted data were then exported to the SIMCA-P software 14.0 version (Umetrics). Pareto scaling was employed on the MS data set. Finally, Principal Component Analysis (PCA), orthogonal PLS-DA (OPLS-DA) and *S*-plot were performed.

Supporting information

The ¹H, ¹³C NMR and HMBC and ESI-MS data of isolated compounds are available as Supporting Information.

Acknowledgments

We thank Ministry of Education Malaysia for a scholarship for Nurkhalida Kamal.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Newman DJ, Cragg GM. Natural Products As Sources of New Drugs over the 30 Years from 1981 to 2010. *Journal of Natural Products* 2012; 75: 311-335
2. Challis GL. Mining microbial genomes for new natural products and biosynthetic pathways. *Microbiology* 2008; 154: 1555-1569
3. Gongora-Castillo E, Buell CR. Bioinformatics challenges in de novo transcriptome assembly using short read sequences in the absence of a reference genome sequence. *Natural Product Reports* 2013; 30: 490-500
4. Yang JY, Karr JR, Watrous JD, Dorrestein PC. Integrating ‘-omics’ and natural product discovery platforms to investigate metabolic exchange in microbiomes. *Current Opinion in Chemical Biology* 2011; 15: 79-87
5. Verpoorte R, Choi Y, Kim H. NMR-based metabolomics at work in phytochemistry. *Phytochemistry reviews* 2007; 6: 3-14
6. Moco S, Vervoort J, Bino RJ, De Vos RCH, Bino R. Metabolomics technologies and metabolite identification. *TrAC Trends in Analytical Chemistry* 2007; 26: 855-866
7. Sumner LW, Lei Z, Nikolau BJ, Saito K. Modern plant metabolomics: advanced natural product gene discoveries, improved technologies, and future prospects. *Natural Product Reports* 2015; 32: 212-229
8. Lang G, Mayhudin NA, Mitova MI, Sun L, van der Sar S, Blunt JW, Cole ALJ, Ellis G, Laatsch H, Munro MHG. Evolving Trends in the Dereplication of Natural Product Extracts: New Methodology for Rapid, Small-Scale Investigation of Natural Product Extracts. *Journal of Natural Products* 2008; 71: 1595-1599
9. Vasilev N, Ebel R, Edrada RA, Fuss E, Alfermann AW, Ionkova I, Petrova A, Replinger M, Schmidt TJ. Metabolic Profiling of Lignan Variability in *Linum* species of Section *Syllinum* native to Bulgaria. *Planta Med* 2008; 74: 273-280
10. Frisvad JC, Rank C, Nielsen KF, Larsen TO. Metabolomics of *Aspergillus fumigatus*. *Medical Mycology* 2008; 47: S53-S71
11. Wolfender J-L, Marti G, Ferreira Queiroz E. Advances in Techniques for Profiling Crude Extracts and for the Rapid Identification of Natural Products: Dereplication, Quality Control and Metabolomics. *Current Organic Chemistry* 2010; 14: 1808-1832

12. Funari CS, Eugster PJ, Martel S, Carrupt P-A, Wolfender J-L, Silva DHS. High resolution ultra high pressure liquid chromatography–time-of-flight mass spectrometry dereplication strategy for the metabolite profiling of Brazilian *Lippia* species. *Journal of Chromatography A* 2012; 1259: 167-178
13. Geiger M, Desanglois G, Hogeveen K, Fessard V, Leprêtre T, Mondeguer F, Guitton Y, Hervé F, Séchet V, Grovel O, Pouchus Y-F, Hess P. Cytotoxicity, Fractionation and Dereplication of Extracts of the Dinoflagellate *Vulcanodinium rugosum*, a Producer of Pinnatoxin G. *Marine Drugs* 2013; 11: 3350-3371
14. Zhang T, Omar R, Siheri W, Al Mutairi S, Clements C, Fearnley J, Edrada-Ebel R, Watson D. Chromatographic analysis with different detectors in the chemical characterisation and dereplication of African propolis. *Talanta* 2014; 120: 181-190
15. Hou Y, Braun DR, Michel CR, Klassen JL, Adnani N, Wyche TP, Bugni TS. Microbial Strain Prioritization Using Metabolomics Tools for the Discovery of Natural Products. *Analytical chemistry* 2012; 84: 4277-4283
16. Ellis GA, Hou Y, Braun DR, Wyche TP, Adnani N, Vazquez-Rivera E, Bugni TS. LC/MS Untargeted Metabolomics for Prioritizing Marine Invertebrate-Associated Bacteria for Discovery of Natural Products. *Planta Med* 2013; 79: PK7
17. Samat N, Tan PJ, Shaari K, Abas F, Lee HB. Prioritization of Natural Extracts by LC–MS-PCA for the Identification of New Photosensitizers for Photodynamic Therapy. *Analytical chemistry* 2014, DOI:
18. Macintyre L, Zhang T, Viegelmann C, Martinez IJ, Cheng C, Dowdells C, Abdelmohsen UR, Gernert C, Hentschel U, Edrada-Ebel R. Metabolomic Tools for Secondary Metabolite Discovery from Marine Microbial Symbionts. *Marine Drugs* 2014; 12: 3416-3448
19. Kim HK, Wilson EG, Choi YH, Verpoorte R. Metabolomics: A Tool for Anticancer Lead-Finding from Natural Products. *Planta Med* 2010; 76: 1094-1102
20. Ali K, Iqbal M, Yuliana N, Lee Y-J, Park S, Han S, Lee J-W, Lee H-S, Verpoorte R, Choi Y. Identification of bioactive metabolites against adenosine A1 receptor using NMR-based metabolomics. *Metabolomics* 2013; 9: 778-785
21. Bohni N, Cordero-Maldonado ML, Maes J, Siverio-Mota D, Marcourt L, Munck S, Kamuhabwa AR, Moshi MJ, Esguerra CV, de Witte PA. Integration

- of Microfractionation, qNMR and Zebrafish Screening for the In Vivo Bioassay-Guided Isolation and Quantitative Bioactivity Analysis of Natural Products. PLoS ONE 2013; 8: e64006
22. Chagas-Paula DA, Zhang T, Oliveira TB, Edrada-Ebel R, Da Costa FB. Discovery of plant anti-inflammatory biomarkers by machine learning algorithms and metabolomic studies. *Planta Med* 2013; 79: SL27
 23. Abdelmohsen U, Cheng C, Viegelmann C, Zhang T, Grkovic T, Ahmed S, Quinn R, Hentschel U, Edrada-Ebel R. Dereplication Strategies for Targeted Isolation of New Antitrypanosomal Actinosporins A and B from a Marine Sponge Associated-*Actinokineospora* sp. EG49. *Marine Drugs* 2014; 12: 1220-1244
 24. Harvey AL, Edrada-Ebel R, Quinn RJ. The re-emergence of natural products for drug discovery in the genomics era. 2015; 14: 111-129
 25. Bacon CW, White J. *Microbial endophytes*: CRC Press; 2000
 26. Dreyfuss M, Chapela IH. Potential of fungi in the discovery of novel, low-molecular weight pharmaceuticals. In, *The discovery of natural products with therapeutic potential*. In V. P. Gullo ed. London, United Kingdom: Butterworth-Heinemann; 1994: 49-80
 27. Hawksworth DL. The fungal dimension of biodiversity: magnitude, significance, and conservation. *Mycological Research* 1991; 95: 641-655
 28. Hawksworth DL. The magnitude of fungal diversity: the 1.5 million species estimate revisited. *Mycological Research* 2001; 105: 1422-1432
 29. Tsuchinari M, Shimanuki K, Hiramatsu F, Murayama T, Koseki T, Shiono Y. Fusapyridons A and B, novel pyridone alkaloids from an endophytic fungus, *Fusarium* sp. YG-45. *Zeitschrift Fur Naturforschung B* 2007; 62: 1203
 30. Zilla MK, Qadri M, Pathania AS, Strobel GA, Nalli Y, Kumar S, Guru SK, Bhushan S, Singh SK, Vishwakarma RA, Riyaz-Ul-Hassan S, Ali A. Bioactive metabolites from an endophytic *Cryptosporiopsis* sp. inhabiting *Clidemia hirta*. *Phytochemistry* 2013; 95: 291-297
 31. Isaka M, Berkaew P, Intereya K, Komwijit S, Sathitkunanon T. Antiplasmodial and antiviral cyclohexadepsipeptides from the endophytic fungus *Pullularia* sp. BCC 8613. *Tetrahedron* 2007; 63: 6855-6860
 32. Singh SB, Ondeyka JG, Tsipouras N, Ruby C, Sardana V, Schulman M, Sanchez M, Pelaez F, Stahlhut MW, Munshi S, Olsen DB, Lingham RB.

- Hinnuliquinone, a C₂-symmetric dimeric non-peptide fungal metabolite inhibitor of HIV-1 protease. *Biochemical and Biophysical Research Communications* 2004; 324: 108-113
33. Chen X, Shi Q, Lin G, Guo S, Yang J. Spirobisnaphthalene Analogues from the Endophytic Fungus *Preussia* sp. *Journal of Natural Products* 2009; 72: 1712-1715
 34. Ge HM, Yu ZG, Zhang J, Wu JH, Tan RX. Bioactive alkaloids from endophytic *Aspergillus fumigatus*. *Journal of Natural Products* 2009; 72: 753-755
 35. Campos FF, Rosa LH, Cota BB, Caligiorne RB, Rabello ALT, Alves TMA, Rosa CA, Zani CL. Leishmanicidal metabolites from *Cochliobolus* sp., an endophytic fungus isolated from *Piptadenia adiantoides* (fabaceae). *PLoS neglected tropical diseases* 2008; 2: e348
 36. Moreno E, Varughese T, Spadafora C, Arnold AE, Coley PD, Kursar TA, Gerwick WH, Cubilla-Rios L. Chemical constituents of the new endophytic fungus *Mycosphaerella* sp. nov. and their anti-parasitic activity. *Natural product communications* 2011; 6: 835
 37. Simarro PP, Diarra A, Postigo JAR, Franco JR, Jannin JG. The human African trypanosomiasis control and surveillance programme of the World Health Organization 2000–2009: the way forward. *PLoS neglected tropical diseases* 2011; 5: e1007
 38. Jacobs RT, Nare B, Phillips MA. State of the art in African trypanosome drug discovery. *Current topics in medicinal chemistry* 2011; 11: 1255
 39. Webster J, Weber R. *Introduction to Fungi*: Cambridge University Press; 2007
 40. Pluskal T, Castillo S, Villar-Briones A, Orešič M. MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics* 2010; 11: 395
 41. Robotti E, Marengo E. Chemometric multivariate tools for candidate biomarker identification: LDA, PLS-DA, SIMCA, Ranking-PCA. *2-D PAGE Map Analysis: Methods and Protocols* 2016, DOI: 237-267
 42. Haoula Z, Ravipati S, Stekel DJ, Ortori CA, Hodgman C, Daykin C, Raine-Fenning N, Barrett DA, Atiomo W. Lipidomic analysis of plasma samples from women with polycystic ovary syndrome. *Metabolomics* 2015; 11: 657-666

43. Schulz B, Sucker J, Aust H, Krohn K, Ludewig K, Jones P, Döring D. Biologically active secondary metabolites of endophytic *Pezizula* species. *Mycological Research* 1995; 99: 1007-1015
44. Aldridge D, Galt S, Giles D, Turner W. Metabolites of *Lasiodiplodia theobromae*. *Journal of the Chemical Society C: Organic* 1971, DOI: 1623-1627
45. Bode HB, Walker M, Zeeck A. Cladospirones B to I from *Sphaeropsidales* sp. F-24' 707 by Variation of Culture Conditions. *European Journal of Organic Chemistry* 2000; 2000: 3185-3193
46. Hazalin NAM, Lim SM, Cole AL, Majeed ABA, Ramasamy K. Apoptosis induced by desmethyl-lasiodiplodin is associated with upregulation of apoptotic genes and downregulation of monocyte chemotactic protein-3. *Anti-cancer drugs* 2013, DOI:
47. Tawfike A, Viegelmann C, Edrada-Ebel R. Metabolomics and Dereplication Strategies in Natural Products. In: Roessner U, Dias DA eds, *Metabolomics Tools for Natural Product Discovery*: Humana Press; 2013: 227-244
48. Bafor EE, Lim CV, Rowan EG, Edrada-Ebel R. The Leaves of *Ficus exasperata* Vahl (Moraceae) Generates Uterine Active Chemical Constituents. *Journal of Ethnopharmacology* 2012; 145: 803-812

Figure Legends

Fig. 1 Scatter plot of the ion peaks of the *L. theobromae* extracts from different days A. in positive ionization and B. in negative ionization.

Fig. 2 The ^1H NMR data of *L. theobromae* extracts obtained from solid rice culture at three different incubation periods (solvent a is DMSO-*d*6; b is Chloroform-*d*). The 30-day rice culture extract could only be fully dissolved in chloroform indicating the compounds occurring in this extract are semi non-polar.

Fig. 3 Anti-trypanosomal activity of *L. theobromae* fractions against *T. b. brucei*. LT: *L. theobromae* extract as positive control; LT2-LT19: *L. theobromae* fractions.

Fig. 4 Above: The ^1H NMR spectra of the 19 fractions; Below: The expansion of the ^1H NMR data of active anti-trypanosome fractions highlighting several unique chemical fingerprints found only in these fractions.

Fig. 5 Chemometric analysis of *L. theobromae* fractions and anti-trypanosomal activity data correlation. A) The score scatter plot of OPLS-DA shows the samples were grouped based on their bioactivity. B) The permutation test result of the OPLS-DA model. C) The loading scatter plot of OPLS-DA shows the *m/z* values of active metabolites D) The S-plot generated from the OPLS-DA model shows the end point compounds that are the most predictive metabolites responsible for the bioactivity (highlighted in red).

Fig. 6 Chemical structures of unique metabolites (**1-8**) predicted from the S-plot and further secondary metabolites (**6, 8-9**) isolated from an active fraction of *L. theobromae* (Group 1).

Fig. 7 The ^1H - ^1H COSY correlation found in fraction LT-3 (Group 1) confirmed the substructure of desmethyl-lasiodiplodin.

Fig. 8 The ^1H - ^1H COSY spectrum of fraction LT-7 (Group 2) shows partial correlations of a) 6-Oxo-de-O-methyl-lasiodiplodin and b) preussomerin-C to confirm the presence of these metabolites determined from the dereplication of HR-LCMS data.

Fig. 9 Three isolated compounds from *L. theobromae* labelled in the S-plot. R-(-)-mellein was in the middle of the plot, suggesting less anti-trypanosomal activity for this compound.

Table 1 Antitrypanosomal activity of *L. theobromae* extracts derived from *V. pinnata* in different types of media and incubation periods. MIC was only determined for the bioactive extracts

Sample	<i>T. b. brucei</i>	<i>T. b. brucei</i>
	20 µg/ml % of viability	MIC Average ± Std Dev (n=4)
LT-LC-7	112.3	Not tested
LT-LC-15	106.3	Not tested
LT-LC-30	126.0	Not tested
LT-RC-7	103.9	Not tested
LT-RC-15	94.3	Not tested
LT-RC-30	1.4	25 ± 1.3 µg/ml
Suramin	Not tested	0.1 ± 0 µM

*LT- *L. theobromae*

Table 2 List of unique metabolites of *L. theobromae* active fractions obtained from S-plot “end-point” data shown on Figure 3d. (P=positive mode; N=negative mode)

Ionization Mode	MS <i>m/z</i>	Rt (min)	Chemical Formula	Name	Group
N	191.035	11.32	C ₁₀ H ₈ O ₄	6,8-Dihydroxy-3-methylisocoumarin(1)	2
N	291.124	13.17	C ₁₆ H ₂₀ O ₅	6-Oxo-de- <i>O</i> -methyllasiodiplodin (2)	2
P	293.139	13.71	C ₁₆ H ₂₀ O ₅	6-Oxo-de- <i>O</i> -methyllasiodiplodin (2)	2
N	395.077	13.60	C ₂₁ H ₁₆ O ₈	Preussomerin-C (3)	2
N	363.051	14.61	C ₂₀ H ₁₂ O ₇	Preussomerin-H (4)	2
N	333.077	17.61	C ₂₀ H ₁₄ O ₅	Palmarumycin CP17 (5)	1
N	349.072	13.71	C ₂₀ H ₁₄ O ₆	Cladospirone B (6)	1
P	321.17	19.14	C ₁₈ H ₂₄ O ₅	Phomopsin B (7)	2
N	319.155	19.16	C ₁₈ H ₂₄ O ₅	Phomopsin B (7)	2
N	277.144	21.79	C ₁₆ H ₂₂ O ₄	Desmethyl-lasiodiplodin (8)	1
P	279.159	21.80	C ₁₆ H ₂₂ O ₄	Desmethyl-lasiodiplodin (8)	1
N	555.296	21.79		Complex of 277.144	1

Table 3 Antitrypanosomal activity of isolated compounds obtained from *L. theobromae* fermented in solid rice culture for thirty days.

Compound	Antitrypanosomal activity (<i>T. b. brucei</i>)
	MIC Average \pm Std Dev (n=4) (μM)
Cladospirone-B 6	17.8 \pm 0
Desmethyl-lasiodiplodin 8	22.5 \pm 1.50
R-(-)- mellein 9	>100 \pm 2.75
Suramin	0.1 \pm 0

Fig. 1

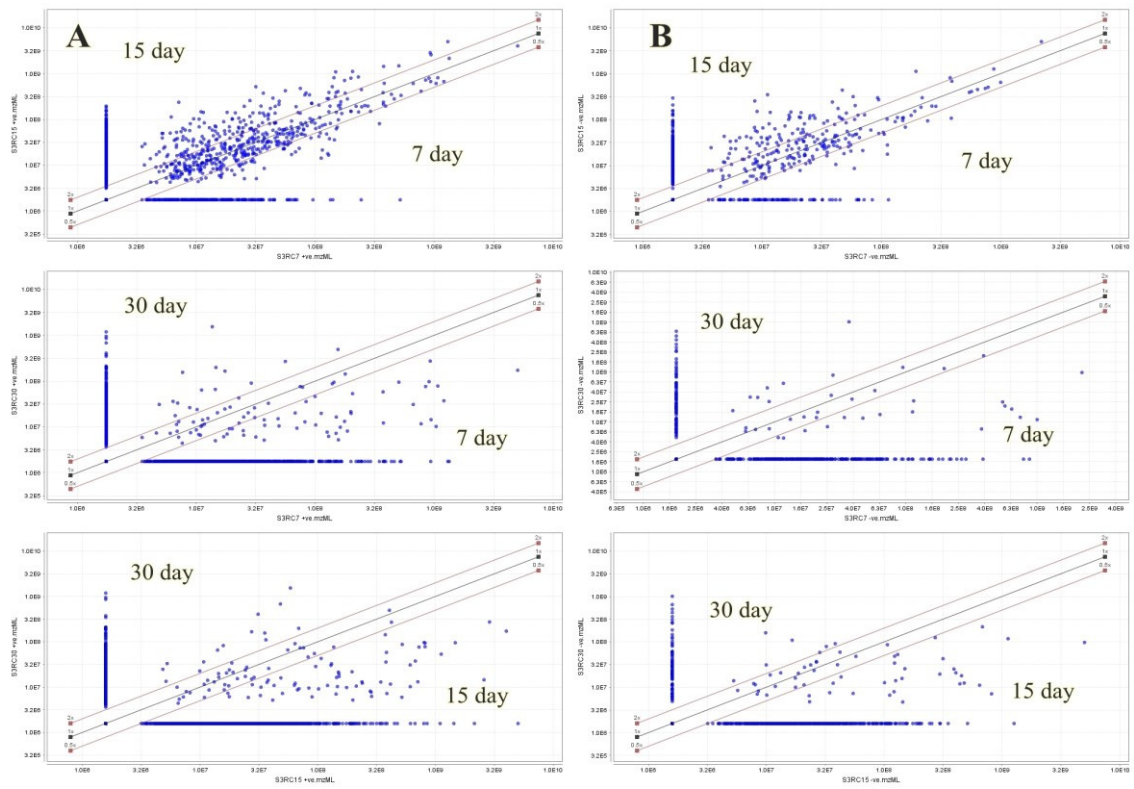


Fig. 2

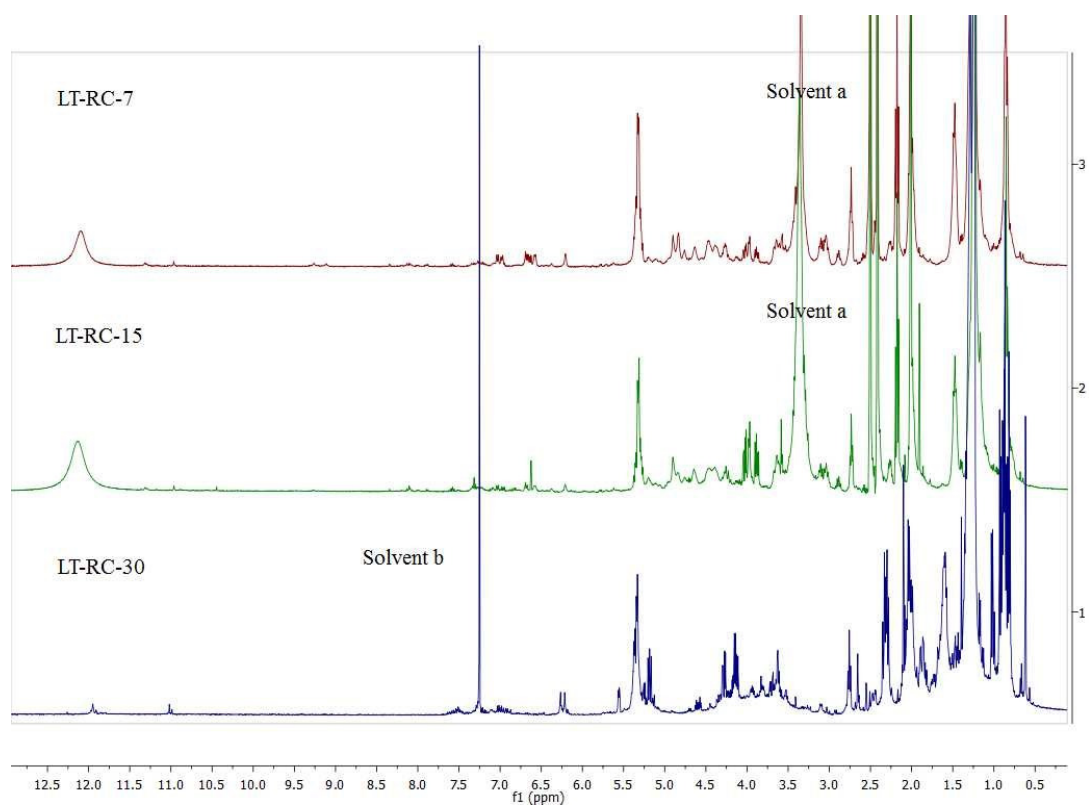


Fig. 3

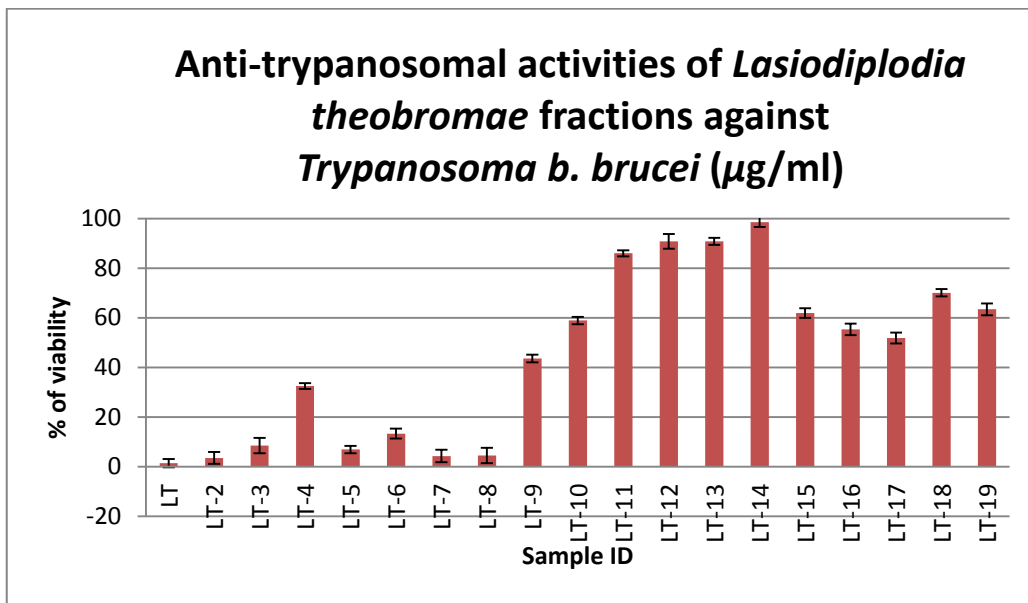


Fig. 4

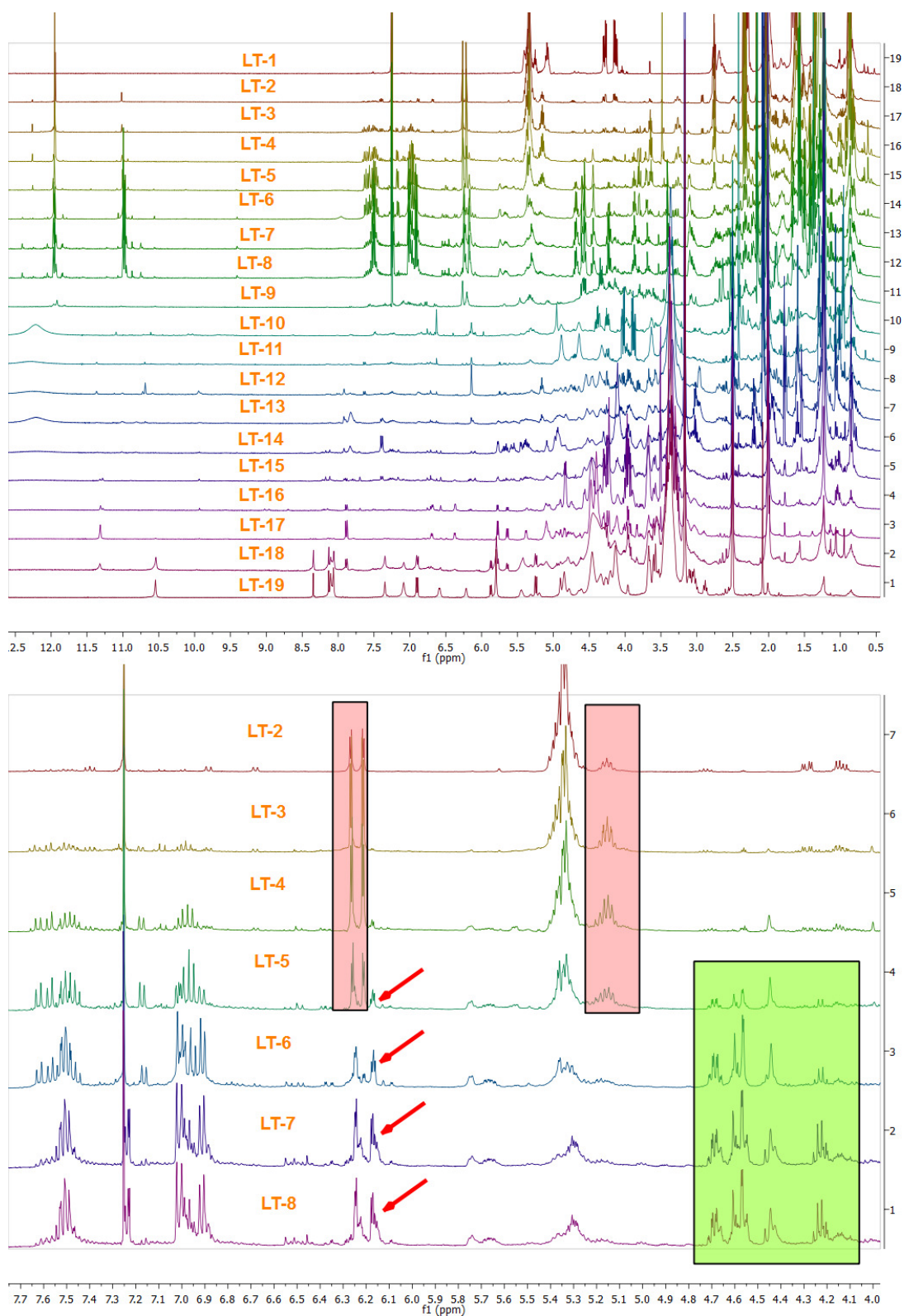


Fig. 5

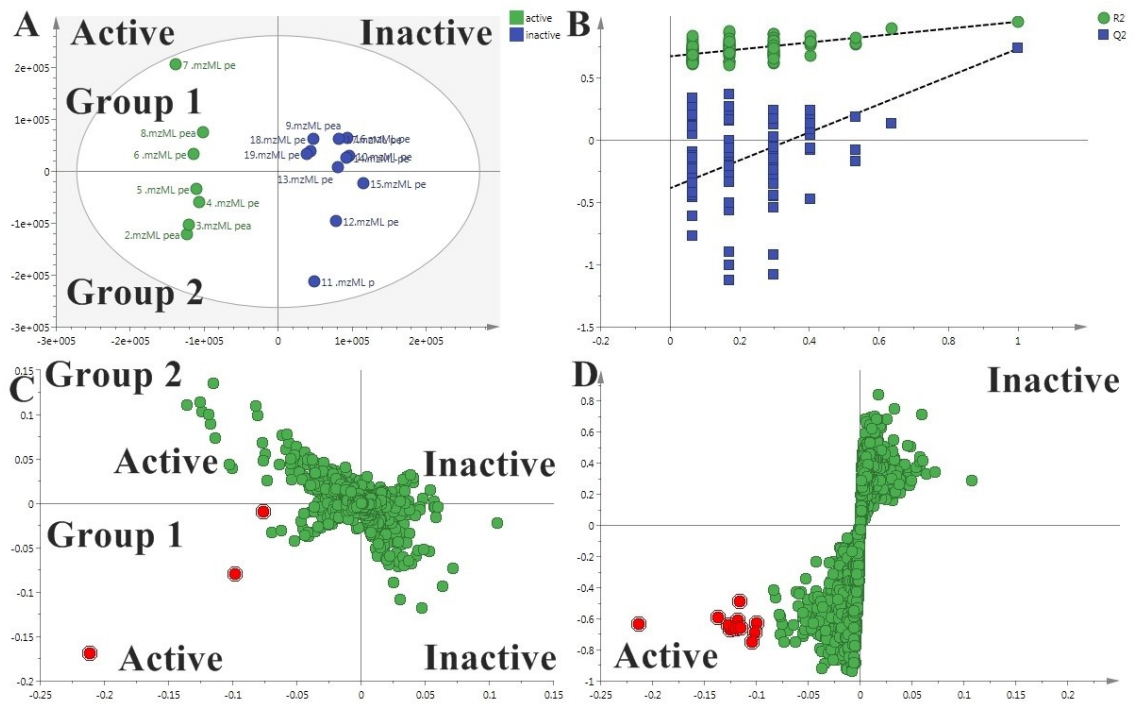


Fig. 6

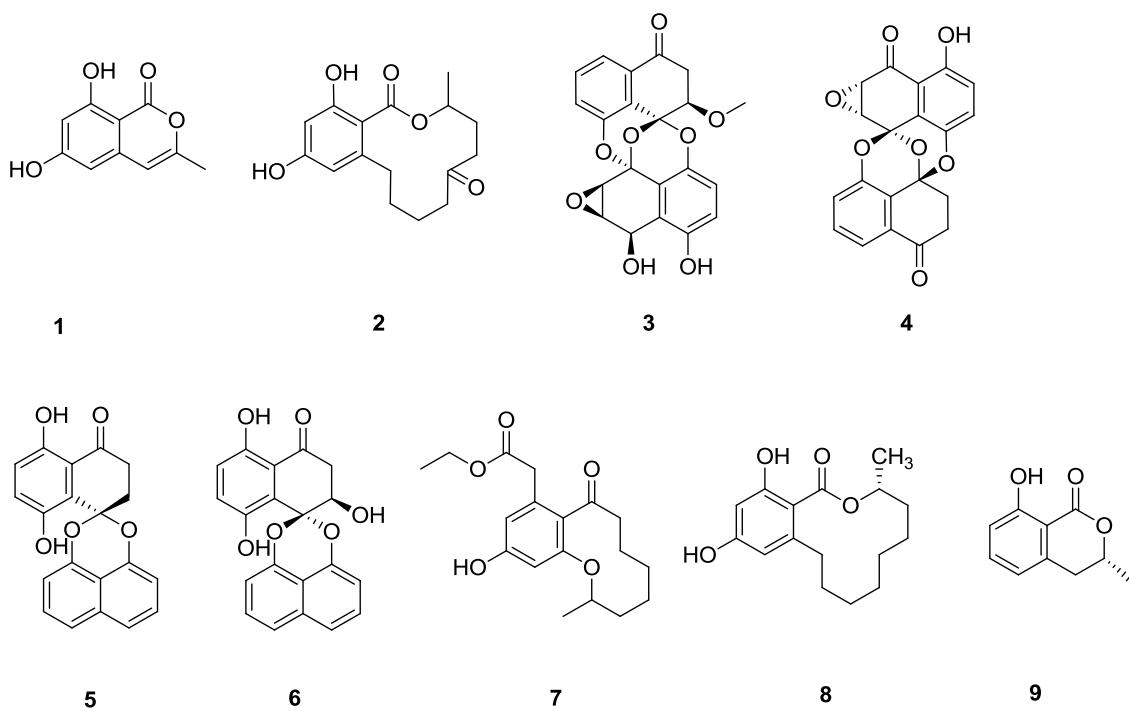


Fig. 7

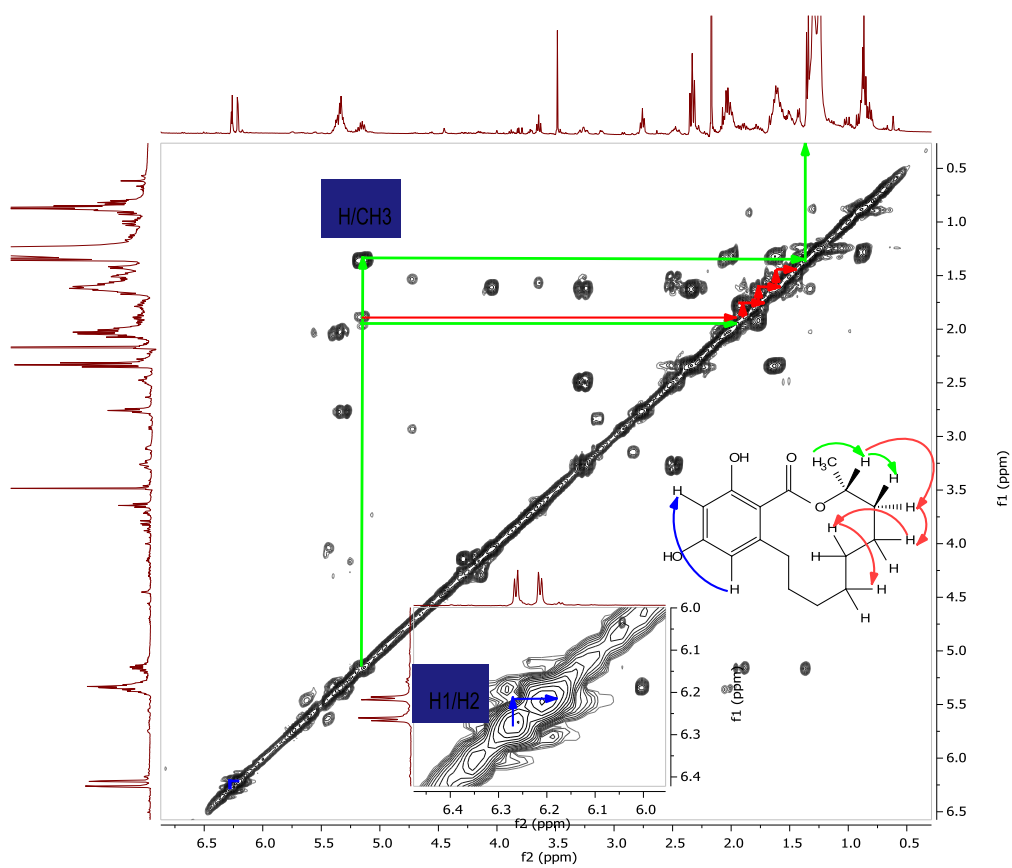


Fig. 8

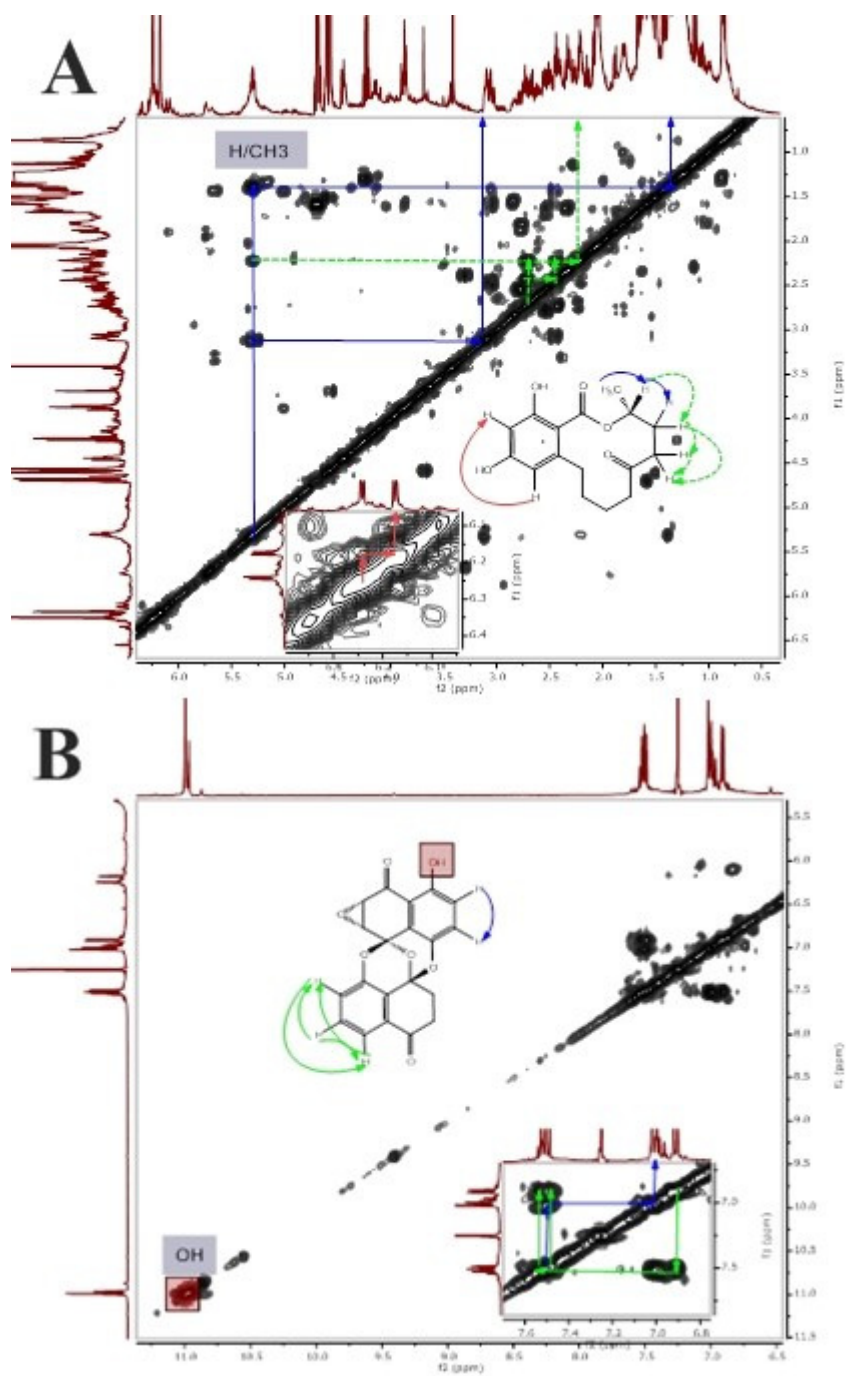


Fig. 9

