ISSN: 2357-0547 (Print)
ISSN: 2357-0539 (Online)

Research Article / JAPR
Tawfik et al., **2017**, 1 (1), 66-74

Journal of Advanced Pharmacy Research

Metabolomics and Bioactivity Guided Isolation of Secondary Metabolites from the Endophytic Fungus *Chaetomium* sp.

Nashwa F. Tawfik[1,2], Ahmed F. Tawfike[1], Randa Abdou[1,3], Grainne Abbott[2], Usama R. Abdelmohsen[4], RuAngelie Edrada-Ebel[2] and Eman G. Haggag[1]*

[1]*Department of Pharmacognosy, Faculty of Pharmacy, Helwan University, Cairo, Egypt, 11795.*
[2]*Strathclyde Institute of Pharmacy and Biomedical Science, University of Strathclyde, 161 Cathedral street, Glasgow G4 0NR, Scotland, United Kingdom.*
[3]*Faculty of Pharmacy Umm Al Qura University, Mekkah, KSA.*
[4]*Faculty of Pharmacy, Minia University, Minia, Egypt.*

Corresponding author*: *Eman G. Haggag
[1]*Department of Pharmacognosy Faculty of Pharmacy, Helwan University, Cairo, 11795, Egypt, Tel.: +201000023022*
E-mail address: Eman.G.Haggag@pharm.helwan.edu

Submitted on: 18-11-2016; Revised on: 01-12-2016; Accepted on: 03-12-2016

ABSTRACT

Objectives: the aim of this study is to explore the secondary metabolites produced by the endophytic fungus *Chaetomium sp*. isolated from *Scencio stapeliiformis* (E.Phillips) as well as investigate the anticancer and antimicrobial activity of crude extracts, fractions and pure compounds. **Methods:** An endophytic fungus (*Chaetomium sp.*) was isolated from the arial part of *S. stapeliiformis* (from Giza, Egypt). DNA sequencing analysis, morphological and chemotaxonomy investigations were used for taxonomic identification. Metabolomics tools and dereplication studies were employed to choose the optimum growth medium and conditions that produce the most significant metabolites. The crude extract of the optimal fungal culture of *Chaetomium sp.* was then fractionated using flash chromatography and medium pressure liquid chromatography (MPLC). The structure of the isolated compounds was determined on the basis of 1D, 2D NMR and mass spectrometry (HR-ESIMS) analysis. **Results:** The Metabolomics and bioassay-guided isolation afforded five pure compounds; p-hydroxybenzaldehyde (1), Uracil (2), 3-benzyl-6-isobutyl piperazine-2,5-dione (3), Cyclo (L-Alanin-L-leucin) (4) and Cyclo-(L-proline-L-leucine) (5). Multivariate data analysis highlighted the most significant metabolites contributed to the measured bioactivity. All fungal extracts were tested for the anticancer activity but extract of 30 days liquid culture of *Chaetomium* showed the most anticancer activity. The pure compounds were tested for their anticancer and antimicrobial activities. Compounds 3 and 5 exhibited a significant anti-trypanosomal activity while compounds 1, 2 and 5 effectively inhibited the growth of *E-coli* and *Staphylococcus aureus.* **Conclusion:** A combination of metabolomic- and bioassay-guided protocol can efficiently predict the putative biologically active metabolites during the first stage of fractionation.

Keywords: Antimicrobial activity, Antitrypanosomal activity, *Chaetomium sp.*, Dereplication, Endophytes, Metabolomics, *Senecio stapeliiformis*.

INTRODUCTION

Senecio represents the largest genus of the family Asteraceae, and has more than 1500 species of herbs, shrubs, vines and trees[1]. *Senecio* species have been used in folk medicine in the treatment of wounds, chest pain, cough, fever and runny nose. It was reported to have a great gastrointestinal protective activity against ulcers[2,3]. Moreover, some studies mentioned the cytotoxic activity of different species of *Senecio*[4]. *Chaetomium*, an endophytic fungus isolated from *S. stapeliiformis*, belongs to Ascomycota of the family *Chaetomiaceae*. It is a large genus comprising over 100 species. Several strains of *Chaetomium* are found in the

soil, plants debris[5]. Endophytic fungi are a highly diverse group of fungi capable of living symbiotically inside plant tissue without causing apparent symptoms of diseases[6]. Endophytes might be involved in the biosynthesis of plant products; however, they might also be the producers themselves of many substances of potential use to the modern medicine, agriculture and pharmaceutical industry[7,8]. An area of major interest to us is to explore endophytic secondary metabolites as novel anticancer and antimicrobial agents.

Since sleeping sickness (Human African Trypanosomiasis "HAT") is an endemic disease in thirty African countries with the population at risk being about 60 million, this has driven us to search for a powerful antitrypanosome of natural source. HAT is a vector-borne parasitic disease caused by infection with protozoan parasites belonging to the genus *Trypanosoma* which are transmitted to humans by tsetse fly (*Glossina* genus) bites[9]. It has two stages: the first is the haemolymphatic stage which lasts for one to three weeks, followed by the chronic stage in which trypanosomes cross the blood–brain barrier to invade the central nervous system resulting to chronic meningo-encephalitis and eventually leads to encephalopathy[10].

Metabolomics is the technological tool designed to deliver general qualitative and quantitative profiles of metabolites in organisms exposed to various conditions. Plants and microorganisms produce many metabolites with different chemistry and bioactivity under stress conditions. Metabolomics displays extra information to figure out these complex relationships between the endophytes and their host plants which aids to discover novel bioactive natural components[11]. The metabolome is the complete set of small molecules found in a cell, tissue or organism at a certain point in time. Dereplication is the process of testing sample mixtures that are active in screening in order to recognize the novel compounds from the active substances that have already been studied. Dereplication was accomplished by employing differential expression analysis softwares like MZmine which involves dictionary of natural products database (DNP) to aid compound identification[11]. By using combinations of analytical, statistical and dereplication methods, the bioassay-guided isolation route is getting shorter and rapid dereplication of known activities is rapidly delivered[12].

MATERIALS AND METHODS

General instruments

^{1}H-, ^{13}C- and 2D-NMR spectra were recorded at 25^{o}C in DMSO-d6 on JNM-LA400 NMR spectrometer, JEOL, Japan and the magnet NMR AS400 model EUR0034 from Oxford Instruments, England at Strathclyde Institute of Pharmacy and Biomedical Science and an AVANCE-III 600 instrument with a 14.1 T Bruker UltraShield magnet at Chemistry Department, Faculty of science, Strathclyde University. ESI-HRMS was measured using FTHRMS-Finnigan LTQ Orbitrap or Exactive mass spectrometer (Thermo Scientific). HPLC analysis was carried out using Dionex UltiMate 3000-ThermoScientific Exactive system instrument, Germany. Crude extracts were initially fractionated using medium pressure liquid chromatography (MPLC) from BÜCHI, MPLC instrument was the Sepacore Purification System with Versaflash column stand. The Reveleris® Flash Forward system of Grace Davison Discovery Sciences (Illinois, United States) was also used for further isolation, which is characterized of having two detectors, an evaporative light scattering detector (ELSD) and a UV detector (wavelength range: 200-500 nm). The fractions were investigated on normal phase thin layer chromatography plates (TLC silica gel 60 F_{254}), reverse phase TLC plates (TLC silica gel 60 RP-18 F_{254S}) and fractionated using preparative TLC plates (TLC silica gel 60 F_{254} on 20x20 cm aluminium sheets) from Merck KGaA, Germany. Spots were visualized under UV lamp (λ 254 nm and λ 380 nm) and after spraying with anisaldehyde and heating chromatograms till colour development.

LC-MS spectra were viewed using Thermo Xcalibur 2.1 (Thermo Scientific, Germany). To convert the raw data into separate positive and negative ionization files, Ms converter software was used. The files were then imported to the data mining software MZmine 2.10 forpeak picking, deconvolution, deisotoping, alignment and formula prediction[11]. Macro file with built in databases was written in Excel, used to combine positive and negative MS files and for further clean-up of media components[13]. The databases used for the identification of compounds were the Dictionary of Natural Products (DNP) 2015, MestReNova (MNova) 2.10 by Mestrelab Research, S.L, (Santiago de Compostela, Spain) was used to process all NMR data and SIMCA 14(Umetrics AB, Umeå, Sweden) was used for multivariate data analysis.

For microbiological work, the laminar flow hood (BioMAT2) was purchased from Medical Air Technology, UK. The stand incubator (Incu-160S) used for agar plates was from SciQuip Ltd., Shropshire. The homogenizer (IKA T18 Basic Ultra-Turrax) and handheld homogenizer (Ultra-Turrax T8) were obtained from IKA Labortechnik, Germany.

Plant material

Fresh plant (*Senecio stapeliiformis* E.Phillips) was collected from the Orman Botanical Garden in Giza, Egypt and identified by; Dr. Therese L. Yousef, senior taxonomist and Engineer Mervat

A. Hasan, herbarium curator at Orman Botanical Garden. Fresh plant materials including all arial parts of the plant were collected a day before isolation of fungal strains, kept in zipped plastic bags under 4° C for the isolation work.

Culture media for isolated endophytes

Wickerham liquid medium (yeast extract 3.0 gm, malt extract 3.0 gm, peptone 5.0 gm, glucose 10.0 gm and distilled water to 1L with pH adjusted at 7.4) and solid medium composed of 100 gm Rice and 100 ml distilled water, were used as culture media for the isolated endophytes.

Cell lines and culture media for cytotoxic assay

Lung cancer cells (A549), Prostatic cancer cells (PC3), breast cancer (ZR75), ovarian cancer cells (A278O) and normal epithelial cells derived from human prostate (PNT2 cells) were purchased from ECACC (Sigma-Aldrich, Dorset, UK). A549 and PC3 cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM), while PNT2, ZR75 and A2780 cells were cultured in RPMI 1640 media; both were supplemented with 10% (v/v) fetal bovine serum, 2 mM glutamine and 50 µ g/mL penicillin/streptomycin solution (all Invitrogen, Paisley, UK). All cells were maintained in a humidified incubator at 37 ºC in the presence of 5% CO_2. Cells were routinely passaged at 90%–95% confluence.

Isolation of the endophytes

The Arial part of the plant was rinsed with sterilized distilled water twice. In order to eliminate surface contaminating microbes, sterilization was carried out by immersing leaves and stems in 70% isopropanol (2 min x 2) followed by rinsing again twice with sterilized distilled water. Using a sterile scalpel a small segment of leave and stem tissue (1 cm in length) was cleaned from outer tissue, the inner tissues were carefully dissected under sterile conditions and placed on malt agar plate containing antibiotic to suppress bacterial growth (medium composition: 15 gm agar (Oxoid), 15 gm malt extract (Oxoid) and chloramphenicol (Acros organics, purity> 98%) in distilled water), pH was adjusted to (7.4-7.8) and incubated at 30ºC. After 3-4 weeks, hyphal tips of the fungi were removed and transferred to fresh MA medium. Plates were prepared in duplicates to eliminate the possibility of contamination. Pure strains were isolated by repeated inoculation. The purified fungus was later transferred to the liquid medium for scaling up[14].

Identification of fungal strain

The isolated fungal strains were identified according to molecular biological procedure by DNA extraction, amplification and sequencing of the ITS region[15]. BLAST search of the FASTA sequence was performed with the option "nr", including GenBank, Ref Seq Nucleotides, EMBL, DDBJ and PDB sequences on the BLAST homepage, (NCBI, Bethesda, USA) using accession number-**KC427016.1.**

Seven endophytic fungi were isolated from different parts of *S. stapeliiformis* identified as *Trichospherical sp., Chaetomium sp., Chaetomium megalocarpum, Asperagillus sp., Rhizopus sp., Ceratobasidium sp* and *Microascus sp.*

Small-scale extraction for screening, metabolomics profiling and dereplication

A plate of each fungal species was transferred into 250 ml flask, then macerated with ethyl acetate (200 ml) overnight followed by homogenization and filtration. The filtrate was then dried under vacuum. One mg of each extract was subjected to HRMS analysis and 8-10 mg for NMR analysis for metabolomics profiling and dereplication studies. A sample of 1 mg/mL concentration of each fungal extract was prepared in duplicate and sent to Strathclyde Institute for Drug Research SIDR for bioassay screening against ovarian cancer (A278O), lung cancer (A549), prostatic cancer (PC3) and breast Cancer (ZR75) cell lines.

Medium scale fermentation, extraction and isolation

Fresh fungal cultures were transferred into Erlenmeyer flasks (1L each) containing 500 ml of Wickerham medium for liquid cultures prepared as stated per materials and methods. The cultures were then incubated at room temperature in static form for 30 days. Medium scale cultivation was carried out using 20 One-L Erlenmeyer flasks for liquid cultures, then 250 mL EtOAc was added to Erlenmeyer flasks containing 500 ml culture medium and left overnight to stop cell growth. Culture media and mycelia were then homogenized in the Ultraturrax for 10 min for cell destruction, followed by vacuum filtration using a Buchner funnel. The mycelium residue was discarded while EtOAc culture filtrates were collected, pooled, dried under vacuum, suspended in 200 mL H_2O and extracted with EtOAc (3 x 200 mL) using a separating funnel [14].

Cytotoxic activity

Cells were seeded in clear 96 flat-bottomed plates and allowed to adhere overnight. Thereafter, metabolite extracts and fractions were added at a final concentration of 30 µg/mL, while for the pure compounds at a concentration of 10 mM/mL and allowed to incubate for 48 hours. Viability was determined using Alamar Blue® (Thermo Fisher, Paisley, UK), according to the manufacturer's

instructions and incubated for a further 6 h. The resulting fluorescence was measured using a Wallac Victor 2 1420 multi-label counter (Perkin Elmer, Beaconsfield, UK), in fluorescence mode: excitation 560, emission 590. Vehicle treated control cells (media with 0.3% DMSO) were considered 100% viable against which metabolite extract treated cells (at a concentration of 30 µg/mL, at least n = 2) and the pure compounds treated cells (at concentration 10 mM/mL) were compared. All results were confirmed microscopically[16].

Antitrypanosomal activity

Antitrypanosomal activity was tested following the protocol of Huber and Koella[17]. Briefly, 10^4 trypanosomes per ml of the *Trypanosoma brucei brucei* strain TC 221 were cultivated in Complete Baltz Medium. Trypanosomes were tested in 96-well plate chambers against different concentrations of test substances at 0.25–50 µM in 1% DMSO to a final volume of 200 µL. For controls, 1% DMSO as well as parasites without any test compound were used simultanously in each plate to show that DMSO did not perturb the results. The plates were then incubated at 37 °C in an atmosphere of 5% CO_2 for 24 h. After addition of 20 µL of Alamar Blue, the activity was measured after 48 and 72 h by light absorption using an MR 700 Microplate Reader (Dynatech, Chantilly, United States) at a wavelength of 550 nm with a reference wavelength of 650 nm. The IC_{50} values of the test compound were quantified by linear interpolation of three independent measurements.

Anti-microbial activity

The *in vitro* antimicrobial activity assessment was carried out using a modified Kirby-Bauer disk diffusion assay[18, 19] against various pathogenic bacterial strains (*Staphylococcus aureus* strain 12600, *Escherichia coli* strain 11775, and Fungi (*Candida albicans* strain 7102). Standard discs of Ampicillin (Antibacterial agent), Amphotericin B (Antifungal agent) served as positive controls, while a filter discs impregnated with 10 µL of solvent (DMSO) was used as a negative control.

RESULTS

Compound 1

Brown sugary substance (8mg), ^{1}H-NMR (DMSO, 400 MHz) ^{13}C-NMR (DMSO, 100 MHz) data presented in table 1; ESIHRMS(pos): m/z 121.0296$[M+H]^+$ (calcd. for $C_7H_6O_2$) Thus compound 3 was assigned in accordance to the reported data [20] as **p-hydroxybenzaldehyde.**

Compound 2

Colorless needles (9 mg); ^{1}H-NMR (DMSO, 400 MHz) ^{13}C-NMR (DMSO, 100 MHz) data presented in table 1; ESIHRMS(pos): m/z 113.03 $[M+H]^+$ (calcd. for $C_4H_4N_2O_2$). Thus compound 3 was assigned in accordance to the reported data [21] as **Uracil**.

Compound 3

Colourless needles (9 mg); ^{1}H-NMR (DMSO, 400 MHz) ^{13}C-NMR (DMSO, 100 MHz) data presented in table 1; ESIHRMS(pos): m/z 261.1598 $[M+H]^+$ (calcd. for $C_{15}H_{20}N_2O_2$). Thus compound 3 was assigned in accordance to the reported data [22] as **3-benzyl-6-isobutyl piperazine-2,5-dione**.

Compound 4

Colourless needles (7mg); ^{1}H-NMR (DMSO, 400 MHz) ^{13}C-NMR (DMSO, 100 MHz) data presented in table 1; ESIHRMS(pos): m/z 185.1286 $[M+H]^+$ (calcd. for $C_9H_{16}N_2O_2$). Thus compound 4 was assigned in accordance to the reported data [21] as **Cyclo(L-Alanin-L-leucine)**

Compound 5

White crystals (17.8 mg), ^{1}H-NMR (DMSO, 400 MHz) ^{13}C-NMR (DMSO, 100 MHz) data presented in table 1; ESIHRMS(pos): m/z 211.1448 $[M+H]^+$ (calcd. for $C_{11}H_{18}N_2O_2$). Thus compound 5 was assigned in accordance to the reported data [23] as **Cyclo-(L-proline-L-leucine**).

DISCUSSION

ESI-MS data produced by Excel-macro database file after combining positive and negative modes and removing the media effect, was sent to R software to apply the heatmap script.

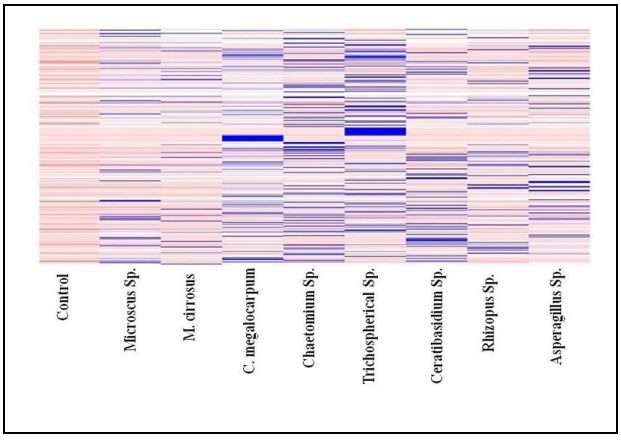

Figure 1. Heatmap of ESI-MS data of all endophytic extracts isolated from *S. stapeliiformis* in which the blue lines represented the produced metabolites.

Table 1. ^{1}HNMR and ^{13}CNMR data of isolated compounds (1-5)

Atom No.	Compound (1)		Compound (2)		Compound (3)		Compound (4)		Compound (5)	
	δ_H (m, *J* in Hz) ppm	δ_C	δ_H (m, *J* in Hz) ppm	δ_C	δ_H (m, *J* in Hz) ppm	δ_C	δ_H (m, *J* in Hz) ppm	δ_C	δ_H (m, *J* in Hz) ppm	δ_C
1	-	164.4	10.83	-	8.09	-	8.10	-	-	167.4
2	6.93(d, *J*=8.68 Hz)	132.7	-	151.1	-	166.7	-	168.8	-	-
3	7.77(d, *J*=8.68 Hz)	116.4	11.03	-	4.17	56.0	3.77(m)	53.1	3.35(2H,m)	45.0
4	-	129.3	-	164.9	8.13	-	8.12	-	1.92, 2.13(m)	28.0
5	7.77(d, *J*=8.68 Hz)	116.4	5.45 (d, *J*=7.57Hz)	101.0	-	168.1	-	169.3	1.84(2H, m)	22.7
6	6.93(d, *J*=8.68 Hz)	132.7	7.39 (dd, *J*=7.01Hz)	143.1	3.47(m)	52.7	3.86(q,,*J*=7.01 Hz)	50.4	4.21(t)	59.1
7	9.79(s)	191.8	10.83	-	2.83, 3.13 (m)	38.9	1.46, 1.61(m)	42.9	-	170.8
8	-	-	-	-	-	136.6	1.82(m)	24.0	8.00	-
9	-	-	-	-	7.27	128.6	0.87(3H,d,*J*=7 Hz)	22.4	4.01	52.9
10	-	-	-	-	7.13	130.0	0.87(3H,d,*J*=7 Hz)	23.5	1.37, 1.78 (m)	38.0
11	-	-	-	-	7.23	127.2	1.27(3H, d, *J*=7.01Hz)	20.1	1.89(m)	24.4
12	-	-	-	-	7.13	130.0	-	-	0.88(3H)	22.4
13	-	-	-	-	7.27	128.6	-	-	0.86(3H)	23.6
14	-	-	-	-	0.74, 0.09	44.1	-	-	-	-
15	-	-	-	-	1.42	23.3	-	-	-	-
16	-	-	-	-	0.61(3H)	23.5	-	-	-	-
17	-	-	-	-	0.61(3H)	21.8	-	-	-	-

The heatmap of all extracts showed that *Chatomium sp*. and *Trichospherical sp.* fungal extracts were the richest in metabolites of different mass range as shown in figure 1**.** The cytotoxicity assay showed that *Chaetomium* sp. was active against A549, A278O and PC3 cell lines and non-toxic for the normal cells PNT2A (Figure 2)**,** implying that *Chaetomium Sp.* could have a unique chemical and biological fingerprints.

Chaetomium was then cultivated on small scale solid and liquid cultures to test the optimum growth condition producing the highest amount of interesting metabolites. HRESI-MS data of crude extracts of both rice (RC) and liquid (LC) culture media of *Chaetomium* have been subjected to a metabolomics workflow which begun with data mining by MZmine. The heatmap for the processed ESI-MS data of both RC and LC extracts of *Chaetomium* showed more abundancy of metabolites inthe 30 days LC extract (Figure 3). Moreover, Multivariate data analysis (MVDA) of different culture extracts of *Chaetomium*, performed by SIMCA-P V.14 software, discriminated 30-days LC extracts from other fungal extracts as shown in the PCA score plot (Figure 4a) which was indicative of the unique nature of the metabolites produced in LC-30 extract.PCA loading plot (Figure 4b) illustrated the metabolites which could be contributed to the variation of 30 days LC extracts. These metabolites were dereplicated by searching DNP 2015 as shown in table 2. Most of these metabolites were reported previously in the literature however metabolites at m/z (retention time in minutes); 181.105 $[M+H]^+$ (5.55), 187.081 $[M-H]^-$ (7.92), 297.218 $[M-H]^-$ (20.38) and 329.210 $[M-H]^-$ (11.63) were not identified in the database. This was motivating to work further on 30 days LC extract. Since it was the most active against the selected cancer cell line and showed no toxicity toward normal cells, 30 days LC was chosen for scale up and further isolation work.

The thirty-day liquid culture extract of *Chaetomium* was subjected to fractionation using MPLC. The resulted fractions were imported into SIMCA for MVDA. The PCA score plot (Figure 5a) showed an outlying of fractions 35-37, 38-39, 40-41, 42-48 and 91-92. The PCA loading plot (Figure 5b) showed the metabolites corresponding to the outlier fractions which are further jack-knifed to remove the

insignificant features. The bioassay guided MPLC fractionation of the 30-days liquid culture extract sorted the active from the inactive fractions according to their activity toward cancer cell lines PC3, A549, ZR75 and A278O (Figure 8).

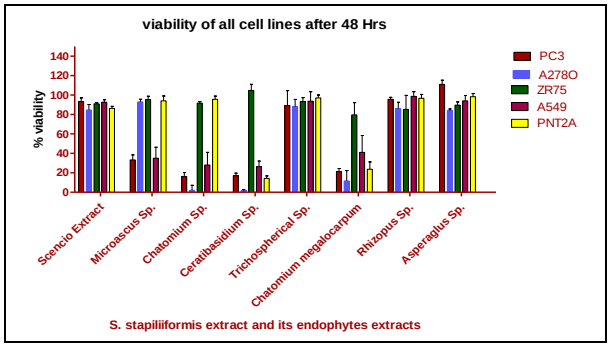

Figure 2. Cytotoxic activity of all endophytic extract isolated from *S. stapeliiformis*

Fractions have been classified into active and inactive in OPLS-DA analysis. OPLS-DA score plot (Figure 5c) displayed a clustering of fractions 38-39, 40-41 and 42-48 in the active side while fraction 4 was singled out because of its different chemical finger print. The respective OPLS-DA loading plot (Figure 5d) showed that fraction 38-39-, 40-41 and 42-48 were characterized by these metabolites m/z 178.08, 259.191 and 341.151 which were identified in DNP as Streptazone A, 3,11-Dihydroxytetradecanoic acid and Pancrimatine B respectively. Whereas metabolites m/z 214.025 and 410.125 recognized as (2-Amino-3-(3-chloro-4-hydroxyphenyl) propanoic acid and Cetocycline, respectively were characteristic for fraction 4. The S-plot of active versus inactive fractions showed the most significant metabolites highly correlated to the cytotoxicity of active fractions (Figure 6). These metabolites were dereplicated as shown in

Table 2. Dereplication of the metabolites contributed to the variation of 30 days LC extract of *Chaetomium* sp.

m/z	Retention time	M.wt	Name	Molecular formula	Source
166.086	2.35	165.079	2-Acetyl-6-ethyl-3-hydroxypyridine	$C_9H_{11}NO_2$	*Abelmoschusmoschatus* (ambrette)
180.102	5.58	179.095	5-Butyl-2-pyridinecarboxylic acid	$C_{10}H_{13}NO_2$	*Fusarium lycopersici, Fusarium oxysporum, Fusarium vasinfectum and Gibberellafujikuroi*
181.105	5.55	180.098	Unkown	$C_5H_{14}N_3O_4$	
185.128	4.47	184.121	Cyclo(alanylleucyl); (3S,6S)-form	$C_9H_{16}N_2O_2$	*Aspergillusphoenicis* and *Nocardiopsis sp.*
187.081	7.92	188.088	Unknown		
211.144	5.34	210.137	6-(1-Methylethyl)-3-(2-methylpropyl)-2(1H)-pyrazinone; 1-Hydroxy	$C_{11}H_{18}N_2O_2$	*Aspergillussojae*
284.295	31.22	283.287	Octadecanoic acid; Amide	$C_{18}H_{37}NO$	*Zosteramarina* (Zosteraceae) and *Rhizocloniumhieroglyphicum*
297.218	20.38	298.225	Unknown		
329.21	11.63	330.217	Unknown		
492.332	18.71	491.325	Oxysporidinone; 4'β-Alcohol	$C_{28}H_{45}NO_6$	*Fusariumoxysporum* GU250648

Table 3. Dereplication of the metabolites highly correlated to the activity of fractions from 30 days LC extract of *Chaetomium* sp.

m/z	Retention time	Molecular weight	Name	Molecular formula	Source
307.13	9.86	308.137	1,14-Diisothiocyanato-1,13-tetradecadiene	$C_{16}H_{24}N_2S_2$	
214.028	4.66	215.035	2-Amino-3-(3-chloro-4-hydroxyphenyl) propanoic acid; (S)-form	$C_9H_{10}ClNO_3$	
343.249	14.51	344.256	Tianshic acid	$C_{19}H_{36}O_5$	*Sambucus williamsii*
341.151	14.26	342.158	Pancrimatine B	$C_{19}H_{22}N_2O_4$	*Pancratium maritimum*
435.275	16.01	436.283	Pancherin A	$C_{25}H_{40}O_6$	*Pittosporum pancheri*
224.093	5.59	225.1	Pyridoxine; O1"-Me, O2"-Ac	$C_{11}H_{15}NO_4$	*Albizziatanganyicensis*
339.135	12.89	340.142	Mactanamide	$C_{19}H_{20}N_2O_4$	*Aspergillus sp.*
259.191	13.98	260.198	3,11-Dihydroxytetradecanoic acid; (3S,11S)-form	$C_{14}H_{28}O_4$	*Ipomoea purpurea*
178.086	4.20	177.079	Streptazone A	$C_{10}H_{11}NO_2$	*Streptomycetes*

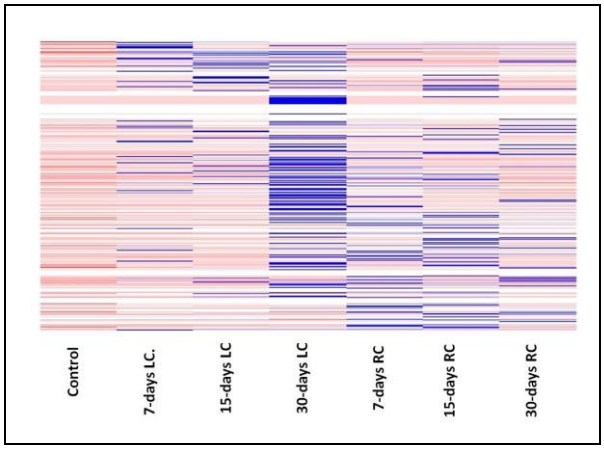

Figure 3. Heatmap for LC and RC culture extracts of *Chaetomium sp*. in which the blue lines represented the produced metabolites

table 3. Searching literature for the bioactivity reported for these metabolites revealed that metabolites at m/z (retention time in minutes); 343.249 $[M-H]^-$ (14.52), 435.275 $[M-H]^-$ (16.01), 341.151 $[M+H]^+$ (14.26) and 178.086 $[M+H]^+$ (4.20) equivalent for $C_{19}H_{36}O_5$, $C_{25}H_{40}O_6$, $C_{19}H_{20}N_2O_4$ and $C_{10}H_{11}NO_2$ respectively, had cytotoxic activity against different types of cancer cell lines[24,25,26,27,28]. This confirmed the power of metabolomics in predicting the bioactive metabolites at first stage of fractionation. However, the rest of significant metabolites in table 3 were not reported to have anticancer activity hence further purification of the active fractions was fundamental to confirm the structure of the previously reported bioactive compounds and test the cytotoxicity for the unreported metabolites.

.

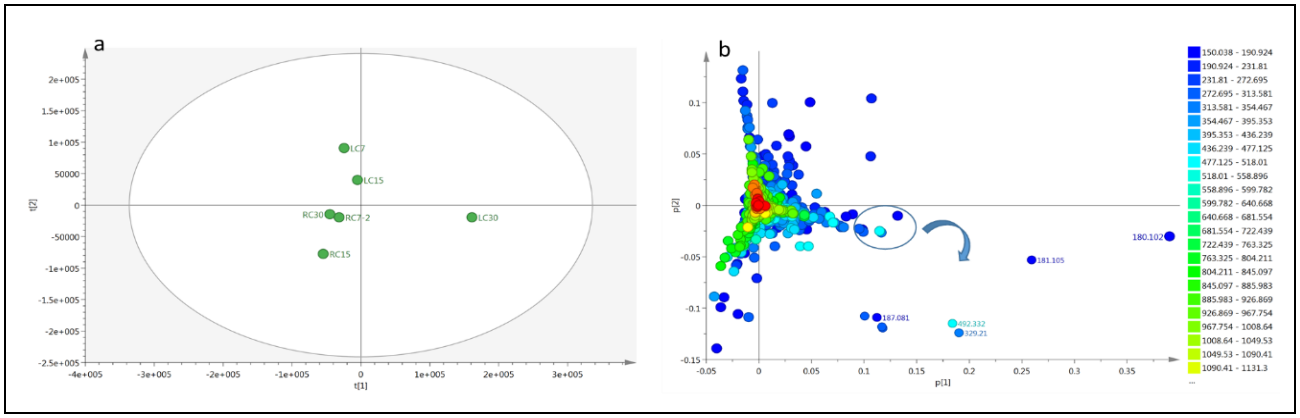

Figure 4. a: PCA score plot of different extracts from solid and liquid fungal culture of *Chaetomium* sp., b: PCA loading plot showing metabolites contributed in 30days LC of *Chaetomium*.

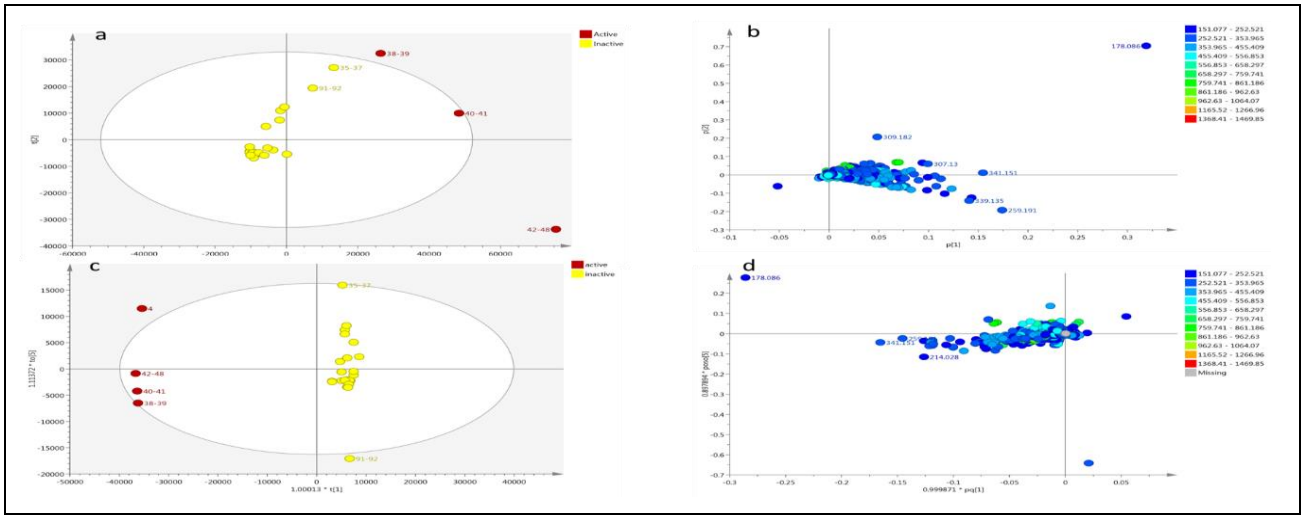

Figure 5. a: PCA score plot of fractions from 30 days LC extract of *Chaetomium* showing the outliers, b: PCA loading plot showing the metabolites contributes to the variation of outliers fraction, c: OPLS-DA score plot of active versus inactive fractions from 30 days LC extract of *Chaetomium*, d: OPLS-DA loading plot highlighting the features corresponding to the active fractions

Var ID	Probability	(m/z)	(Rt)	M.wt	p(corr)[1]
N_7701	7.90E-07	307.13	9.85816	308.137	-0.78497
N_6245	1.11E-06	214.028	4.66293	215.035	-0.77647
N_4209	3.15E-06	343.249	14.5186	344.256	-0.76038
N_6103	1.61E-05	341.151	14.2693	342.158	-0.72352
N_15506	6.23E-05	435.275	16.0142	436.283	-0.68892
N_4303	0.00012	343.249	14.3241	344.256	-0.66651
N_3778	0.001286	224.093	5.59619	225.1	-0.5813
N_979	0.006219	339.135	12.8954	340.142	-0.50832
N_854	0.010246	259.191	13.9815	260.198	-0.48281
P_420	0.046486	178.086	4.2043	177.079	-0.38389

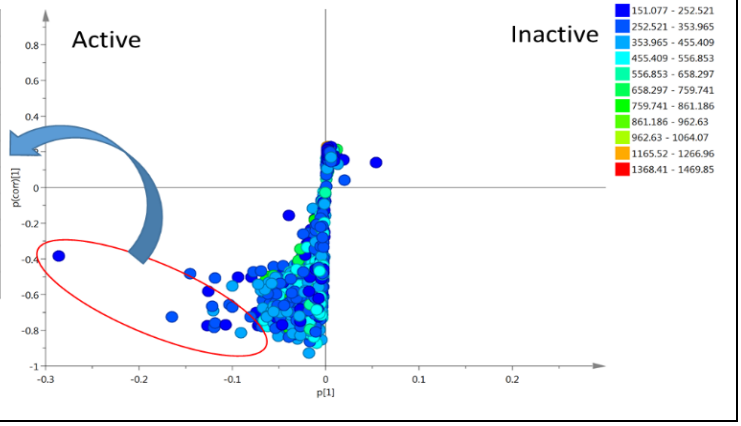

Figure 6. S-plot of active versus inactive fractions showing the metabolites highly correlated the anticancer activity of *Chaetomium*.

Metabolomics- and bioactivity guided studies were greatly focused on the anticancer activity due to the significant inhibition demonstrated by the crude extracts and fractions from 30 day LC fungal extract of *Chaetomium.* Since the putatively identified metabolites, which were highly correlated to this activity, are produced in a very small amount, it was not possible to purify them from the active fractions. The major compounds (1-5) isolated from the fractions of 30-days liquid culture were identified as p-hydroxybenzaldehyde (1), Uracil (2),3-benzyl-6-isobutyl piperazine-2,5-dione (3), Cyclo(L-Alanin-L-leucin) (4) and Cyclo-(L-proline-L-leucine) (5) (Figure 7), showed no activity against the tested cancer cell lines. Consequently, they were investigated for their antitrypanosomal and antimicrobial activity

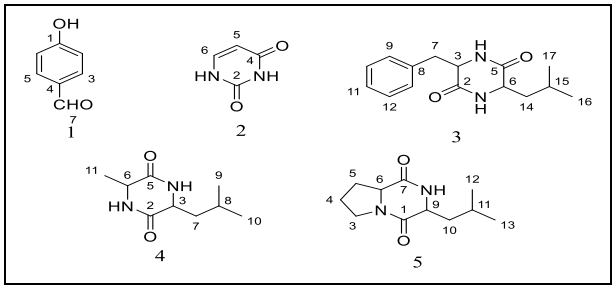

Figure 7. Structure of isolated compounds (1-5) from *Chaetomium* extract.

Testing the compounds for antitrypanosomal activity illustrated that compound 3 (3-benzyl-6-isobutyl piperazine-2,5-dione) and compound 5 (Cyclo-(L-proline-L-leucine) showed a significant activity against *T. brucei brucei* with IC_{50} value of 5.17 µM (48 hrs), 6.26 µM at (72 hrs) and 14.73 µM (48 hrs), 17.35 µM (72 hrs), respectively. The antitrypanosomal activity of these compounds is reported for the first time in this study. Furthermore, compound 1 (p-hydroxybenzaldehyde), compound 2 (Uracil) and compound 5 (Cyclo-(L-proline-L-leucine) exhibited antibacterial activity against *E-coli* and *S. aureus* with inhibition zones of (10, 11), (9, 9) and (10, 10) mm respectively.

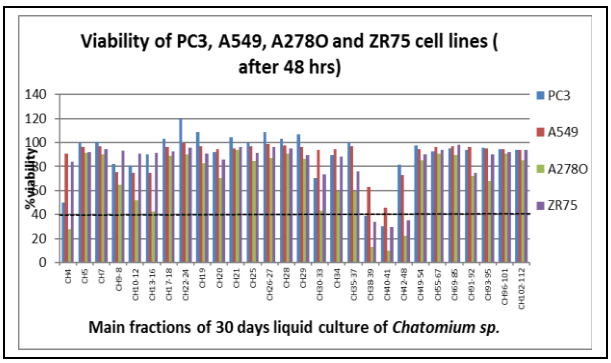

Figure 8. Cytotoxic activity of the MPLC fractions of 30 days liquid culture of *Chaetomium Sp.*

CONCLUSION

A combination of metabolomic- and bioassay-guided protocol could be used to recognize the putative biologically active metabolites during the first stage of fractionation. Metabolomics is a powerful tool employing a multivariate statistical approach which logically highlights the metabolites highly correlated to the biological activity. Metabolomics gave an access to predict the significant metabolites that might be produced in a very small amount, solving the mystery behind the fact that purified major compounds are not found active.

Conflict of Interest

The authors declare that they don't have any conflict of interest.

REFERENCES

1. Nordestam, B. *The Biology and Chemistry of the Compositae*. Academic Press Inc., London, UK, **1977**.
2. Lakshmanan, H.; Raman J.; Nanjian R. Gastroprotective effect of *Senecio candicans* DC on experimental ulcer models. *J. Ethnopharmacol.* **2014**, *140*, 145-150.
3. York, T.; DeWet, H.; Van Vuuren S. F. Plants used for treating respiratory infections in rural maputaland, kwazulu-natal, South Africa. *J. Ethnopharmacol.* **2011**, *135*, 696-710.
4. Rosa, T.; Monica R. T.; Marco B.; Federica M.; Daiela D.; Micodemo G. P.; Giancarlo S.; Francesco M. In-vitro cytoxic effect of *Senecio stabianus lactia* (Astraceae) on human cancer cell lines. *Nat. Prod. Res*. **2009**, *23* (18), 1707-1718.
5. Rodriguez, K. S.; Guarro J., Mycologia. Three new species of *Chaetomium* from soli. *Mycologia* **2002**, *94*, 116-126.
6. Petrini, O. Fungal Endophytes of Tree Leaves. In: *Microbial Ecology of Leaves.* John, H.; Andrews, J. H.; Hirano, S. S.; Eds.; Springer-Verlag, New York, USA, **1991**.
7. Strobel, G. A. Endophytes as sources of bioactive products. *Microbes. Infect.* **2009**, *5* (6), 535-544.
8. Arnold, A. E. Understanding the diversity of foliar endophytic fungi: Progress, challenges, and frontiers. *Fungal Biol.* **2007**, *21*, 51-66.
9. Barrett, M. P. The fall and rise of sleeping sickness. *Lancet* **1999**, *353* (9159), 1113-1114.
10. Enanga, B.; Burchmore R. J.; Stewart M. L.; Barrett M. P. Sleeping sickness and the brain. *Cell Mol. Life Sci.* **2002**, *59* (5), 845-858.
11. Tawfike, A. F.; Viegelmann C.; and Ebel R. E. Metabolomics and dereplication strategies in natural products. *Methods Mol. Biol*. **2013**, *1055*, 227-244.
12. Ebadam, S. S.; Edrada R. E.; Lin W.; Proksch P. Methods for isolation, purification and structural elucidation of bioactive secondary metabolites from marine invertebrates. *Nat. Protoc.* **2008**, *3* (12), 1820-1831.
13. Macintyre, L.; Zhang T.; Viegelmann, C.; Martinez, I. J.; Cheng, C.; Dowdells, C.; Abdelmohsen, U. R; Gernert, C; Hentschel, U.; Ebel, R. E. Metabolomic tools for secondary metabolite discovery from marine microbial symbionts. *Mar. Drugs.* **2014**, *12* (6), 3416-3448.
14. Debbab, A.; Aly, A. H.; Ebel, E. R.; Mueller, W. E. G.; Mosaddak, M.; Hakikj, A.; Ebel, R.; Proksch, P. Bioactive secondary metabolites from the endophytic fungus *Chaetomium sp*. isolated from *Saliva officinalis* growing in Morocco. *Biotech. Agronomy Soc. Environ.* **2009**, *13* (2), 229-234.
15. Kjer, J.; Debbab A.; Aly, A. H.; Proksch, P. Methods for isolation of marine-derived endophytic fungi and their bioactive secondary products. *Nat. Protoc.* **2010**, *5* (3), 479-490.
16. Kevin, L. M.; Debra, B.; Guðmundur, Ó. H.; Eva, K.; Margrét, E. Á.; Louise, C. Y.; David, H. G.; Ebel, R. E.; Katherine R. D. Using molecular networking for microbial secondary metabolite bioprospecting. *Metabolites* **2016**, *6* (2), 1-18.
17. Huber, W.; Koella, J. C. A comparison of three methods of estimating EC_{50} in studies of drug resistance of malaria parasites. *Acta. Trop.* **1993**, *55* (4), 257-261.
18. Bauer, A. W.; Kirby, W. M.; Sherris, J. C.; Turck, M. Antibiotic susceptibility testing by a standardized single disk method. *Am. J. Clin. Pathol*. **1966**, *45* (4), 493-496.
19. Pfaller, A. M.; Burmeister, L.; Bartlett, M. S.; Rinaldi, M. G. Multicenter evaluation of four methods of yeast inoculum preparation. *J. Clin. Microbiol.* **1988**, *26* (8), 1437-1441.
20. Kim, H.; Ralph, J.; Lu, F.; Ralph, S. A.; Boudet, A. M.; MacKay, J. J.; Sederoff, R. R.; Ito, T.; Kawai, S.; Ohashi, H.; Higuchi, T. NMR analysis of lignins in cad-deficient plants. Part 1. Incorporation of hydroxycinnamaldehydes and hydroxy-benzaldehydes into lignins. *Org. Biomol. Chem*. **2003**, *1* (2), 268-281.
21. Zhang-Gui, D. J. Y. Z.; Pei-Wen Y.; Ming-Gang, L.; Rong, H.; Meng-Liang, W. ^{1}H and ^{13}C NMR assignments of eight nitrogen containing compounds from *Nocardia alba sp*. nov (YIM 30243t). *Magn. Reson. Chem.* **2009**, *47*, 366-370.
22. Estibaliz, S. F. S.; Jacqueline, J.; Ángel, M.; Aurelio, O. Diketopiperazines derivatives isolated from *Bacillus thuringiensis* and *Bacillus endophyticus*, establishment of their configuration by X-ray and their synthesis. *Tetrahedron Lett.* **2016**, *57*, 2604-2607.
23. Xujuan, Y.; Naili W.; Albert Sun-C.; Xinsheng, Y. A new eudesmane derivative and a new fatty acid ester from *Sambucus williamsii*. *Chem. Pharm. Bull.* **2006**, *54* (5) 676-678.
24. Véronique, É. O. T.; Hadjira, B.; Françoise, G.; Thierry, S.; Marc, L. Cytotoxic farnesyl glycosides from *Pittosporum pancheri*. *Phytochemistry* **2007**, *68* (5), 604-608.
25. Wang, F.; Guan Y. Cytotoxic nor-dammarane triterpenoids from Dysoxylum hainanense. *Fitoterapia* **2012**, *83* (1), 13-17.
26. Ibrahim, S. R.; Mohamed, G. A.; Shaala, L. A.; Youssef, D. T.; El Sayed, K. A. New alkaloids from *Pancratium maritimum*. *Planta Med*. ***2013***, *79* (15), 1480-1484.
27. Carsten, P. P. K.; Axel Z. Streptazones a, b1, b2, c, and d: New piperidine alkaloids from Streptomycetes. *J. Nat. Prod.* **2000**, *63* (9), 1258-1260.