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

Nashwa F. Tawfik1,2, Ahmed F. Tawfike1, Randa Abdou1,3, Grainne Abbott2, Usama R. Abdelmohsen4, RuAngelie Edrada-Ebel2 and Eman G. Haggag1*

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

*Corresponding author: Eman G. Haggag
1Department 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 Scenecio stapeliiformis (E.Philips) 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 trees1. 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 ulcers2,3. Moreover, some studies mentioned the cytotoxic activity of different species of Senecio4. 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\textsuperscript{5}. Endophytic fungi are a highly diverse group of fungi capable of living symbiotically inside plant tissue without causing apparent symptoms of diseases\textsuperscript{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\textsuperscript{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 \textit{Trypanosoma} which are transmitted to humans by tsetse fly (\textit{Glossina} genus) bites\textsuperscript{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\textsuperscript{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\textsuperscript{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\textsuperscript{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\textsuperscript{12}.

**MATERIALS AND METHODS**

**General instruments**

\textsuperscript{1}H-, \textsuperscript{13}C- and 2D-NMR spectra were recorded at 25°C in DMSO-d\textsubscript{6} 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 Exactive system instrument, Germany. Crude extracts were initially fractionated using medium pressure liquid chromatography (MPLC) from BUCHI, MPLC instrument was the Sepacore Purification System with Versaflash column stand. The Reveleris\textsuperscript{®} 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\textsubscript{254}), reverse phase TLC plates (TLC silica gel 60 RP-18 F\textsubscript{254S}) and fractionated using preparative TLC plates (TLC silica gel 60 F\textsubscript{254S} on 20x20 cm aluminium sheets) from Merck KGaA, Germany. Spots were visualized under UV lamp (\( \lambda 254 \) nm and \( \lambda 380 \) 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\textsuperscript{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\textsuperscript{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 Labotechnik, Germany.

**Plant material**

Fresh plant (\textit{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 (A2780) 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 μg/mL penicillin/streptomycin solution (all Invitrogen, Paisley, UK). All cells were maintained in a humidified incubator at 37 ºC in the presence of 5% CO2. 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 30°C. After 3-4 weeks, hyphal tips of the fungi were removed and transferred to fresh MA medium. Plates were prepared in duplicates to eliminate bacterial growth. Culture media and mycelia 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 H2O and extracted with EtOAc (3 x 200 mL) using a separating funnel.

**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. 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 (A2780), 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 H2O and extracted with EtOAc (3 x 200 mL) using a separating funnel.

**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.

Antitrypanosomal activity
Antitrypanosomal activity was tested following the protocol of Huber and Koella. Briefly, 10⁴ trypanosomes per ml of the Trypanosoma brucei brucei strain TC 221 were cultivated in Complete Baitz 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 simultaneously 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% CO₂ 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 IC₅₀ 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 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), Amphoterin 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). ¹H-NMR (DMSO, 400 MHz) ¹³C-NMR (DMSO, 100 MHz) data presented in table 1; ESIHRMS(pos): m/z 121.0296[M+H]⁺ (calcd. for C₇H₆O₂) Thus compound 3 was assigned in accordance to the reported data as p-hydroxybenzaldehyde.

Compound 2
Colorless needles (9 mg); ¹H-NMR (DMSO, 400 MHz) ¹³C-NMR (DMSO, 100 MHz) data presented in table 1; ESIHRMS(pos): m/z 113.03 [M+H]⁺ (calcd. for C₇H₆N₂O₂). Thus compound 3 was assigned in accordance to the reported data as Uracil.

Compound 3
Colourless needles (9 mg); ¹H-NMR (DMSO, 400 MHz) ¹³C-NMR (DMSO, 100 MHz) data presented in table 1; ESIHRMS(pos): m/z 261.1598 [M+H]⁺ (calcd. for C₁₃H₁₅N₂O₂). Thus compound 3 was assigned in accordance to the reported data as 3-benzyl-6-isobutyl piperazine-2,5-dione.

Compound 4
Colourless needles (7mg); ¹H-NMR (DMSO, 400 MHz) ¹³C-NMR (DMSO, 100 MHz) data presented in table 1; ESIHRMS(pos): m/z 185.1286 [M+H]⁺ (calcd. for C₇H₁₀N₂O₂). Thus compound 4 was assigned in accordance to the reported data as Cyclo(L-Alanin-L-leucine).

Compound 5
White crystals (17.8 mg), ¹H-NMR (DMSO, 400 MHz) ¹³C-NMR (DMSO, 100 MHz) data presented in table 1; ESIHRMS(pos): m/z 211.1448 [M+H]⁺ (calcd. for C₁₁H₁₅N₂O₂). Thus compound 5 was assigned in accordance to the reported data 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.
the resulting fractions were imported into

PCA score plot (Figure 4a) which was indicative of

The heatmap of all extracts showed that *Chaetomium*

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,

A2780 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 abundance of metabolites

in the 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

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

<table>
<thead>
<tr>
<th>Atom No.</th>
<th>Compound (1)</th>
<th>Compound (2)</th>
<th>Compound (3)</th>
<th>Compound (4)</th>
<th>Compound (5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>δ_H (m, f in Hz) ppm</td>
<td>δ_C</td>
<td>δ_H (m, f in Hz) ppm</td>
<td>δ_C</td>
<td>δ_H (m, f in Hz) ppm</td>
<td>δ_C</td>
</tr>
<tr>
<td>1</td>
<td>-</td>
<td>164.4</td>
<td>10.83</td>
<td>-</td>
<td>8.09</td>
</tr>
<tr>
<td>2</td>
<td>6.93(d, J=8.68 Hz)</td>
<td>132.7</td>
<td>-</td>
<td>151.1</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>7.77(d, J=8.68 Hz)</td>
<td>116.4</td>
<td>11.03</td>
<td>-</td>
<td>4.17</td>
</tr>
<tr>
<td>4</td>
<td>-</td>
<td>129.3</td>
<td>-</td>
<td>164.9</td>
<td>8.13</td>
</tr>
<tr>
<td>5</td>
<td>7.77(d, J=8.68 Hz)</td>
<td>116.4</td>
<td>5.45 (d, J=7.57Hz)</td>
<td>101.0</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>6.93(d, J=8.68 Hz)</td>
<td>132.7</td>
<td>7.39 (dd, J=7.01Hz)</td>
<td>143.1</td>
<td>3.47(m)</td>
</tr>
<tr>
<td>7</td>
<td>9.79(s)</td>
<td>191.8</td>
<td>10.83</td>
<td>-</td>
<td>2.83, 3.13 (m)</td>
</tr>
<tr>
<td>8</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>9</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>10</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>11</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>12</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>13</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>14</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>15</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>16</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>17</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

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](image)

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

<table>
<thead>
<tr>
<th>m/z</th>
<th>Retention time</th>
<th>M.wt</th>
<th>Name</th>
<th>Molecular formula</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>166.086</td>
<td>2.35</td>
<td>165.079</td>
<td>2-Acetyl-6-ethyl-3-hydroxypyridine</td>
<td>C₈H₁₂NO₂</td>
<td><em>Abelmoschus moschatus</em> (ambrette)</td>
</tr>
<tr>
<td>180.102</td>
<td>5.58</td>
<td>179.095</td>
<td>5-Butyl-2-pyridinecarboxylic acid</td>
<td>C₁₀H₁₂NO₂</td>
<td><em>Fusarium lycopersici, Fusarium oxysporum,</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><em>Fusarium vasinfectum</em> and Gibberella fujikuroi*</td>
</tr>
<tr>
<td>181.105</td>
<td>5.55</td>
<td>180.098</td>
<td>Unknown</td>
<td>C₆H₁₄N₂O₂</td>
<td>Aspergillus sojae</td>
</tr>
<tr>
<td>185.128</td>
<td>4.47</td>
<td>184.121</td>
<td>Cyclo(alanlylleucyl); (3S,6S)-form</td>
<td>C₉H₁₂NO₂</td>
<td><em>Aspergillusphoenicus</em> and <em>Nocardiopsis</em></td>
</tr>
<tr>
<td>187.081</td>
<td>7.92</td>
<td>188.088</td>
<td>Unknown</td>
<td></td>
<td></td>
</tr>
<tr>
<td>211.144</td>
<td>5.34</td>
<td>210.137</td>
<td>6-(1-Methylbutyl)-3-(2-methylpropyl)-2(H)-pyrazon; 1-Hydroxy</td>
<td>C₁₂H₁₄N₂O₂</td>
<td><em>Zosteramarina</em> (<em>Zosteraceae</em>) and <em>Rhizoclonium hieroglyphicum</em></td>
</tr>
<tr>
<td>284.295</td>
<td>31.22</td>
<td>283.287</td>
<td>Octadecanoic acid; Amide</td>
<td>C₁₈H₃₂NO</td>
<td><em>Fusarium oxysporum</em> GU250648</td>
</tr>
<tr>
<td>297.218</td>
<td>20.38</td>
<td>298.225</td>
<td>Unknown</td>
<td></td>
<td></td>
</tr>
<tr>
<td>329.21</td>
<td>11.63</td>
<td>330.217</td>
<td>Unknown</td>
<td></td>
<td></td>
</tr>
<tr>
<td>492.332</td>
<td>18.71</td>
<td>491.325</td>
<td>Oxysporidinone; 4’β-Alcohol</td>
<td>C₁₃H₁₈NO₆</td>
<td></td>
</tr>
</tbody>
</table>

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

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 fingerprint. 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 Pancratimine 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>
<thead>
<tr>
<th>m/z</th>
<th>Retention time</th>
<th>Molecular weight</th>
<th>Name</th>
<th>Molecular formula</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>307.13</td>
<td>9.86</td>
<td>308.137</td>
<td>1,14-Diisothiocyanato-1,13-tetradecadiene</td>
<td>C₁₆H₂₀N₂S₂</td>
<td><em>Sambucus williamsii</em></td>
</tr>
<tr>
<td>214.028</td>
<td>4.66</td>
<td>215.035</td>
<td>2-Amino-3-(3-chloro-4-hydroxyphenyl) propanoic acid; (S)-form</td>
<td>C₁₀H₁₄ClNO₂</td>
<td><em>Pancratium maritimum</em></td>
</tr>
<tr>
<td>343.249</td>
<td>14.51</td>
<td>344.256</td>
<td>Tianshic acid</td>
<td>C₁₈H₂₀NO₂</td>
<td><em>Pittosporum pancheri</em></td>
</tr>
<tr>
<td>341.151</td>
<td>14.26</td>
<td>342.158</td>
<td>Pancratimine B</td>
<td>C₁₈H₂₀NO₂</td>
<td><em>Albizziatananegycenis</em></td>
</tr>
<tr>
<td>435.275</td>
<td>16.01</td>
<td>436.283</td>
<td>Pancherin A</td>
<td>C₁₈H₂₀NO₂</td>
<td><em>Aspergillus sp.</em></td>
</tr>
<tr>
<td>224.093</td>
<td>5.59</td>
<td>225.1</td>
<td>Pyridoxine; O1’-Me, O2’-Ac</td>
<td>C₁₁H₁₈NO₄</td>
<td><em>Ipomoea purpurea</em></td>
</tr>
<tr>
<td>339.135</td>
<td>12.89</td>
<td>340.142</td>
<td>Mactanamide</td>
<td>C₁₈H₂₀NO₂</td>
<td><em>Streptomyces</em></td>
</tr>
<tr>
<td>259.191</td>
<td>13.98</td>
<td>260.198</td>
<td>3,11-Dihydroxytetradecanoic acid; (3S,11S)-form</td>
<td>C₁₄H₂₀NO₂</td>
<td><em>Streptomyces</em></td>
</tr>
<tr>
<td>178.086</td>
<td>4.20</td>
<td>177.079</td>
<td>Streptazone A</td>
<td>C₁₄H₁₈NO₂</td>
<td><em>Ipomoea purpurea</em></td>
</tr>
</tbody>
</table>

Table 3. Dereplication of the metabolites highly correlated to the activity of fractions from 30 days LC extract of *Chaetomium* sp.
Figure 3. Heatmap for LC and RC culture extracts of Chaetomium sp. in which the blue lines represented the produced 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 30 days 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

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 C₁₀H₅₀O₃, C₂₅H₄₀O₆, C₁₉H₂₀N₂O₄ and C₁₀H₁₅NO₂ respectively, had cytotoxic activity against different types of cancer cell lines²⁴,²⁵,²⁶,²⁷,²⁸. 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.
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.

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.

Conflict of Interest

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


