Journal of Chromatography B, 910 (2012) 103–113

Contents lists available at SciVerse ScienceDirect

Journal of Chromatography B

journal homepage: www.elsevier.com/locate/chromb

Multivariate data analysis to evaluate the fingerprint peaks responsible for the cytotoxic activity of *Mallotus* species☆

C. Tistaert[a], G. Chataigné[b], B. Dejaegher[a], C. Rivière[b], N. Nguyen Hoai[c], M. Chau Van[c], J. Quetin-Leclercq[b], Y. Vander Heyden[a,*]

[a] *Department of Analytical Chemistry and Pharmaceutical Technology, Center for Pharmaceutical Research, Vrije Universiteit Brussel – VUB, Laarbeeklaan 103, B-1090 Brussels, Belgium*
[b] *Analytical Chemistry, Drug Analysis and Pharmacognosy Unit, Université Catholique de Louvain, Avenue E. Mounier 72, B-1200 Brussels, Belgium*
[c] *Institute of Marine Biochemistry, Vietnam Academy of Science and Technology, 18 Hoang Quoc Viet Road, Hanoi, Viet Nam*

ARTICLE INFO

Article history:
Received 8 November 2011
Accepted 1 October 2012
Available online 26 October 2012

Keywords:
Fingerprints
Cytotoxicity
Multivariate calibration
Indication of peaks

ABSTRACT

The *Mallotus* genus comprises numerous species used as traditional medicines in oriental countries and provides scientists a broad basis in the search for pharmacologically active constituents. In this paper, the cytotoxicity of 39 *Mallotus* extracts, different in species, part of the plant used, origin, and harvest season, is evaluated combining cytotoxicity assays with fingerprint technology and data handling tools. At first, the antiproliferative activity of the plant extracts is analyzed both on a non-cancerous cell line (WI-38 – human lung fibroblast) and on a cancerous cell line (HeLa human cervix carcinoma). The results are linked to a data set of high-performance liquid chromatographic fingerprint profiles of the samples using multivariate calibration techniques. The regression coefficients of the multivariate model are then evaluated to indicate those peaks potentially responsible for the cytotoxic activity of the *Mallotus* extracts. In a final step, the cytotoxic extracts are analyzed by HPLC–MS and the indicated peaks identified.

1. Introduction

For hundreds, even thousands of years, traditional medicines (TM) have been practiced in the larger part of the world for the prevention and treatment of human diseases. To date, up to 80% of the population in Africa, Latin America and Asia still relies on TM to meet their primary health care needs because of historical and cultural influences, and their accessibility and affordability compared to the expensive and scarce allopathic drugs [1,2]. In the developed countries, TM and especially herbal derived products have been of great importance for the identification and development of numerous pharmaceutical compounds [3–5]. However, the interest in TM does not remain limited to a source of new constituents, but is also growing rapidly because of concerns about the adverse affects of chemical drugs and the increasing amount of information available on TM.

The fast growing industry in TM and their gaining importance in health care systems worldwide require extensive quality control criteria. Numerous factors, including climate differences, cultivation conditions, time of harvest and sample process procedures, as well as coincidental or deliberate adulterations, may cause large variations in active ingredients and require strict regulations to assure the quality of the drug prior to administration to the patient [6–8]. Even though the first reports on the use of traditional medicines by ancient health care practitioners are dated almost 5000 years ago, it was not until 1978 before the first indication of the need to regulate traditional medicines was made by the World Health Organization (WHO) [9]. Since then, many monographs on the quality control of herbal substances have been published in official pharmacopeias worldwide allowing quality control, usually based on macroscopic and microscopic techniques and the analyses of single marker compounds [10,11].

Although the above methods have demonstrated to be useful, they cannot cope with the quality control standards set nowadays. The use of marker compounds might not always be suitable due to the lack of unique chemical compounds and ignores the synergic effects between the constituents of complex biological samples [12,13]. To counter the uprising problems, the WHO has introduced fingerprints analysis as a methodology for the assessment of natural products [1]. A fingerprint is defined as a characteristic profile reflecting the complex chemical composition of the sample and is obtained by chromatographic, spectroscopic or electrophoretic analysis. The resulting fingerprint profiles can then be used as unique identification utilities to ensure the quality and authenticity of the samples. Furthermore, the combination with data analytical

☆ This paper belongs to the Special Issue Chemometrics in Chromatography, Edited by Pedro Araujo and Bjørn Grung.

* Corresponding author. Tel.: +32 2 477 47 34; fax: +32 2 477 47 35.
E-mail address: yvanvdh@vub.ac.be (Y. Vander Heyden).

1570-0232/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jchromb.2012.10.001

tools does not only allow the establishment of classification models for herbal samples taking into account minor differences that may not be observed bare-eyed, but additionally may enable the prediction of important pharmacological activities and the indication of peaks potentially responsible for a measured activity [14–19].

One of the omnipresent herbal genera in TM is the *Mallotus* genus (family *Euphorbiaceae*). The genus comprises over 140 species of which many are used as traditional medicines in Asia [20]. Over the years, the genus has provided scientists a broad basis in their research for new pharmacological compounds. Numerous publications can be found in the literature about the identification and purification of compounds possessing activities as diverse as antioxidant, hepato-protective, cytotoxic, antimicrobial, anti-inflammatory and retroviral [21–27]. However, most studies focus on the analysis of a very limited number of compounds within one particular species, an approach which is hardly representative for the complex nature of herbal products.

The focus of the present study is the indication and identification of compounds responsible for the cytotoxic activity of 39 *Mallotus* samples based on their entire fingerprint profiles recorded by high-performance liquid chromatography coupled to UV detection. The cytotoxic activity on a cancerous cell line (HeLa – human cervix carcinoma) and a non-cancerous cell line (WI-38 – human lung fibroblast) is measured by the tetrazolium salt MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide) colorimetric method as described in Refs. [28,29]. Combining the information on the antiproliferative capacity of the herbal extracts with the chromatographic fingerprints allows constructing a multivariate calibration model using Orthogonal Projections to Latent Structures (O-PLS) [30]. The regression coefficients of the model can then be evaluated indicating the peaks important for its construction, i.e. the peaks potentially responsible for the modeled activity [15]. In a final step, LC–MS analyses are performed to analyze and identify the compounds underlying the indicated peaks.

2. Theory

Before applying chemometric techniques, the chromatographic data needs to be organized in an $n \times p$ data matrix **X**, where the n objects (herbal samples) constitute the rows and the p variables (time points) the columns.

2.1. Data pretreatment

Appropriate pretreatment of the data has a major influence on the outcome of the chemometric analysis and is considered to be an absolute prerequisite. In chromatographic fingerprinting, the alignment or warping of chromatographic data has a particularly important place. Along the time axis, peak shifts may occur due to instrument instability, column aging and small variations in mobile phase composition. Many techniques have been developed and amongst them correlation optimized warping (COW) [31–33], dynamic time warping (DTW) [31], parametric time warping (PTW) [33] and fuzzy warping [34] are the most frequently applied.

Other frequently used pretreatment techniques include normalization and column centering, both of which were applied in the presented study. Normalization removes undesired effects caused by unequal amounts of injected samples. It divides each row, corresponding to a fingerprint, by its norm [35,36]. Additionally, as the useful information resides in the between-sample variation of the variables, column centering is a generally applied preprocessing method. By removing the column mean from each corresponding element, a centered variable has a mean of zero.

2.2. Exploratory analysis: Principal Component Analysis

With Principal Component Analysis (PCA) [36,38] latent variables, called principal components, are defined. They are linear combinations of the original variables and describe the largest remaining variation in **X**. The different PCs are orthogonal and can be defined until a maximal number of PCs equal to $n-1$ (with $n<p$) is reached. The projections of the n objects from the original data space on a PC are the scores on this PC, while the contribution of each original variable to the PC is reflected by its loading. Both scores and loadings can be used for exploratory analysis of the original data. A score plot, representing the scores on two PCs, gives information regarding the (dis)similarity of the objects, while a loading plot provides information on the contribution of the original variables to the considered PCs.

2.3. Orthogonal Projections to Latent Structures

Orthogonal Projections to Latent Structures (O-PLS) [30,35] is a modification of the partial least squares (PLS) algorithm [35,37,38] studying the relationship between an $n \times p$ data matrix **X** and an $n \times 1$ response vector **y**. It removes the information that is not correlated to the response by subtracting orthogonal components from the original data. Consequently, the data is split into two data sets containing the **y**-relevant information and the **y**-orthogonal information. By removing the orthogonal information from the original data, the model complexity can be reduced to a single factor, improving the interpretability of the regression coefficients without compromising the predictive power of the model.

3. Experimental

3.1. Preparation of the herbal extracts

39 *Mallotus* samples, belonging to at least 17 different species (5 samples were unidentified), were collected in different Vietnamese regions (Table 1) and authenticated by Professor Nguyen Nghia Thin (Hanoi National University, Viet Nam). The samples were different in the time of harvest, their origin, or the part of the plant that was used.

To prepare the extracts, 2.5 g crude plant material was weighed and extracted three times for 1 h with 25 mL methanol in an ultrasonic bath (Branson Ultrasonic Corporation, CT, USA) at a temperature between 40° and 50 °C. The extract was filtered through a 240 nm pore size filter paper (Whatman, Hanoi, Viet Nam) and solvent evaporated at reduced pressure (60 Pa) and elevated temperature (50° C). The obtained crude extract was divided over three sample tubes, i.e. one for the cytotoxicity assay, one for HPLC and MS analysis, and one as a voucher specimen. The voucher specimens were deposited at the Institute of Natural Products Chemistry, Hanoi, Viet Nam.

3.2. Cytotoxicity assay

The antiproliferative capacity of the *Mallotus* extracts was evaluated on HeLa and WI-38 cell lines using the tetrazolium MTT (Sigma–Aldrich, Bornem, Belgium) colorimetric method as described in Refs. [28,29]. The method is based on the cleavage of the reagent by succinyl dehydrogenase in viable cells, which is measured by recording the absorbances of the medium at 570 nm (with background subtraction at 620 nm) against a background control and a positive control. By calculating the difference between the absorbences at both wavelengths, the value is corrected for non-specific background interferences. The obtained value is inversed to the antiproliferative activity of the extract.

Table 1
The *Mallotus* samples with their voucher number, species, origin, collection time, used part of the plant and viability of the cancerous (HeLa) and non cancerous (WI-38) cell lines incubated with the sample extracts (50 μg/ml). All experiments were performed in triplicate and are shown as the %viability ± SEM (standard error of the mean). The cytotoxic samples are marked in bold, while the samples of which MS spectra are available are indicated with *.

Sample number	Voucher	Species	Origin	Collection time	Part of plant	%Viability WI-38	%Viability HeLa
1	01	*Mallotus luchenensis*	Son La	July 2006	Leaves	118.0 ± 2.0	62.3 ± 1.4
2	02	*Mallotus microcarpus*	Son La	July 2006	Leaves	61.0 ± 4.0	45.3 ± 2.6
3*	03	*Mallotus barbatus*	Son La	July 2006	Leaves	64.5 ± 1.0	84.0 ± 1.7
4*	**MA07**	***Mallotus sp1***	**Van Hoa**	**April 2006**	**Leaves**	**14.0 ± 0.7**	**47.7 ± 1.5**
5*	NT01	*Mallotus barbatus*	Hagiang	November 2006	Leaves	114.0 ± 1.7	61.1 ± 0.4
6	NT02	*Mallotus paniculatus*	Hagiang	November 2006	Leaves	109.3 ± 4.9	61.0 ± 1.4
7	NT03	*Mallotus metcalfianus*	Hagiang	November 2006	Leaves	59.3 ± 3.2	82.3 ± 1.2
8*	**MA01**	***Mallotus apelta* (*Ma1*)**	**Tam Dao**	**July 2006**	**Leaves**	**16.0 ± 0.8**	**43.0 ± 0.6**
9*	**MA02**	***Mallotus apelta* (*Ma2*)**	**Tam Dao**	**December 2006**	**Leaves**	**13.0 ± 0.6**	**50.2 ± 2.0**
10*	MA03	*Mallotus paniculatus*	Tam Dao	April 2006	Leaves	99.3 ± 2.8	89.0 ± 4.4
11	SP4	*Mallotus sp2*	Langson	March 2006	Leaves	87.0 ± 1.2	69.6 ± 2.6
12	SP5	*Mallotus philippinensis*	Langson	March 2006	Leaves	97.3 ± 2.0	71.1 ± 2.8
13*	**MA11**	***Mallotus macrostachyus***	**Langson**	**March 2006**	**Leaves**	**13.3 ± 2.0**	**41.3 ± 0.3**
14*	**MA12**	***Mallotus microcarpus***	**Quangbinh**	**March 2006**	**Leaves**	**24.3 ± 2.6**	**52.7 ± 1.0**
15	MA13	*Mallotus pallidus*	Quangbinh	March 2006	Leaves	110.0 ± 1.2	77.7 ± 0.9
16*	MA14	*Mallotus oblongifolius*	Quangtri	March 2006	Leaves	96.0 ± 2.3	72.0 ± 1.0
17*	MA15	*Mallotus floribundus*	Langson	November 2006	Leaves	74.3 ± 1.5	73.8 ± 2.8
18	**MA16**	***Mallotus cuneatus***	**Langson**	**November 2006**	**Leaves**	**13.5 ± 1.0**	**46.7 ± 1.5**
19*	MA17	*Mallotus cuneatus*	Quangbinh	December 2006	Leaves	61.0 ± 6.0	48.3 ± 1.2
20	**MA18**	***Mallotus sp3***	**Quang tri**	**December 2006**	**Leaves**	**12.8 ± 1.0**	**43.7 ± 0.7**
21	**MA19**	***Mallotus yunnanensis***	**Lang Son**	**November 2006**	**Leaves**	**15.3 ± 0.9**	**61.3 ± 1.8**
22	MA20	*Mallotus poilanei*	Ke Bang	March 2006	Leaves	61.3 ± 1.5	56.8 ± 1.5
23	MA22	*Mallotus hookerianus*	Dakrong	March 2006	Leaves	98.7 ± 2.9	79.3 ± 2.3
24	MA23	*Mallotus nanus*	Daclak	March 2006	Leaves	101.0 ± 2.3	67.0 ± 2.6
25	MA24	*Mallotus sp4*	Daclak	March 2006	Leaves	91.0 ± 1.2	86.7 ± 0.9
26	M25	*Mallotus oreophilus*	LaoCai	June 2006	Leaves	94.0 ± 2.9	82.7 ± 2.3
27*	MA28	*Mallotus philippinensis*	Cucphuong	December 2006	Leaves	103.3 ± 4.3	86.0 ± 2.9
28*	MA29	*Mallotus barbatus*	Cucphuong	December 2006	Leaves	86.5 ± 1.5	86.3 ± 5.2
29*	MP31L	*Mallotus paniculatus*	VQG Pumat	September 2006	Leaves	79.3 ± 2.0	76.1 ± 2.2
30*	MP32R	*Mallotus paniculatus*	VQG Pumat	September 2006	Roots	76.8 ± 2.8	73.9 ± 1.6
31*	MP33L	*Mallotus paniculatus*	Bach Ma-TTH	October 2006	Leaves	84.7 ± 4.3	82.7 ± 0.9
32*	MP34R	*Mallotus paniculatus*	Bach Ma-TTH	October 2006	Roots	96.0 ± 2.1	91.3 ± 0.9
33*	MP35R	*Mallotus paniculatus*	Cucphuong	December 2006	Roots	77.8 ± 2.9	77.3 ± 2.1
34*	MP36L	*Mallotus paniculatus*	Cucphuong	December 2006	Leaves	103.7 ± 0.7	85.3 ± 2.2
35*	MN37R	*Mallotus nanus*	VQG-Bachma	May 2006	Roots	93.0 ± 2.3	66.9 ± 0.3
36*	MN37L	*Mallotus nanus*	VQG-Bachma	May 2006	Leaves	81.0 ± 0.6	74.8 ± 2.0
37*	MN39C	*Mallotus nanus*	VQG-Bachma	May 2006	Bark	87.0 ± 2.6	68.6 ± 0.9
38	M40L	*Mallotus sp5*	VQG Bavi	August 2006	Leaves	95.3 ± 4.8	79.3 ± 2.0
39	M41C	*Mallotus sp6*	VQG Bavi	August 2006	Bark	90.0 ± 3.5	79.3 ± 1.5

HeLa and WI-38 cells were grown in Dulbecco's modified Eagle's medium (DMEM, Gibco® BRL, Invitrogen, Antwerp, Belgium) containing L-glutamine, D-glucose and sodium pyruvate, supplemented with 10% fetal calf serum (FBS, Gibco® BRL, Invitrogen, Antwerp, Belgium) and antibiotics (100 IU penicillin/mL and 100 μg streptomycin/mL). The cells were incubated in a humidified atmosphere containing 5% CO_2. Stock solutions of the extracts were prepared at 5 mg/mL in dimethyl sulfoxide (DMSO, Sigma, Bornem, Belgium) and stored at −4 °C. Briefly, 5000 HeLa or WI-38 cells per well were seeded in 100 μL of DMEM with 10% FBS in 96-well microculture plates for 24 h. After 24 h, the medium was removed and 200 μL of fresh medium containing 50 μg/mL extract were added to each well, while control cells received fresh medium containing analogous DMSO concentrations. Each concentration was tested in at least 6 wells. After 72 h incubation, the medium was replaced by 100 μL DMEM containing 10 μL of MTT solution (3 mg/mL in PBS). After 45 min, the medium was removed and 100 μL DMSO was added to each well. The plates were shaken and the absorbencies were recorded at 570 nm and 620 nm on a microplate reader (SpetraMax 190, Sopachem, Ochten, The Netherlands) against a background control (100 μL of pure DMSO) as blank. Camptothecin (25–0.00025 μg/mL) (Sigma–Aldrich, Bornem, Belgium) was used as positive control. The %viability was expressed as the difference in absorbance at both wavelengths of the treated cells relative to that of the control cells, which was considered as 100% (Eq. (1)):

$$\%\,\text{viability} = \frac{A_T}{A_{NT}} \times 100 \quad (1)$$

where A is the difference in absorbance at both wavelengths for the control cells (NT) and the treated cells (T). All experiments were made at least in triplicate and the results were averaged.

3.3. Sample preparation

Samples for HPLC and MS analysis were prepared diluting 50.0 mg crude extract in 2.0 mL methanol. The mixture was shaken during 15 min at 250 rpm on a shaking bath (Edmund Bühler, Hechingen, Germany), filtered through a 2 μm pore size filter (Schleicher & Schuell, Dassel, Germany) followed by filtration trough a 25 mm syringe polypropylene membrane with 0.2 μm pore size (VWR International, Leuven, Belgium).

The recently isolated cytotoxic benzopyrans 6-[1′-oxo-3′(R)-hydroxy-butyl]-5,7-dimethoxy-2,2-dimethyl-2H-1-benzopyran (Malloapelta A) and 6-[1′-oxo-3′(R)-methoxy-butyl]-5,7-dimethoxy-2,2-dimethyl-2H-1-benzopyran (Malloapelta B) [24] were provided by the Institute of Natural Products Chemistry (Vietnamese Academy of Science and Technology, Hanoi, Viet

Nam). For the standards 1.0 mg was weighed and dissolved in 10.0 mL methanol. Then, the same procedure as for the crude extracts was followed.

3.4. HPLC

3.4.1. Equipment, chemicals and reagents

The experiments were performed on an Agilent HPLC system (Waldbronn, Germany) equipped with an auto sampler, vacuum degasser, quaternary pump, column oven and a variable wavelength UV detector. The stationary phase consisted of two coupled Chromolith™ Performance RP-18e (100 mm × 4.6 mm I.D.) with a Chromolith guard column RP-18e (5 mm × 4.6 mm I.D.) purchased from Merck (Darmstadt, Germany). HPLC grade acetonitrile (Fisher Scientific, Leicestershire, UK), trifluoroacetic acid (TFA) (Sigma–Aldrich, Steinheim, Germany), and MilliQ water obtained from a MilliQ water purification system (Millipore, Bedford, MA) were used to prepare the mobile phases. All solvents were degassed during 15 min on an ultrasonic bath (Branson Ultrasonic Corporation, Danbury, CT) prior to analysis. The data acquisition and processing was done with Chemstation for LC (Agilent).

3.4.2. Chromatographic conditions

The chromatographic conditions developed in Ref. [34] were applied. The mobile phase consisted of (A) 0.05% TFA in MilliQ water, and (B) 0.05% TFA in ACN. Gradient elution was applied. The gradient program for the two coupled Chromolith™ Performance RP-18e with guard column was 5–20% B in 0–25 min, 20–95% B in 25–50 min and 95% B in the 50–60 min interval. Furthermore, the column temperature was 25 °C, the flow rate 1.0 mL/min, the injection volume 10 μL, and the detection wavelength 254 nm.

3.5. LC–MS

All experiments were executed on an Alliance HPLC (Waters, Milford, MA, USA) equipped with an auto sampler and column oven. MS-detection was conducted using an ion trap LCQ-advantage system (Thermo Fisher Scientific, Waltham, MA, USA) equipped with an APCI interface. All MS analyses were performed with a mass precision of 0.5 atomic mass units (amu). The MS acquisitions were performed in both positive and negative atmospheric pressure ionization modes.

The following APCI inlet conditions were used. Nitrogen was used both as a nebulizing gas at 450 °C with an arbitrary flow of 70, and as a drying gas at 450 °C with an arbitrary flow of 30. The capillary temperature was set at 200 °C. In the positive mode, the capillary voltage was set to 26 V, the source voltage to 6 kV and the source current to 5 μA. In the negative mode, the capillary voltage was set to −4 V, the source voltage to 4.5 kV and the source current to 80 μA. In both modes 25 V of collision energy was applied.

3.6. Data analysis

Computations were performed on a PC with an Intel Core 2 Duo E6750 processor containing 2 gigabyte RAM and running Microsoft Windows XP Pro and Matlab™ 7.1 (The Mathworks, Natick, MA). All data analyses were performed using m-files written for Matlab 7.1.

4. Results and discussion

4.1. Cytotoxicity assay

The samples are arbitrarily divided into three different categories based on the %viability of the cell lines. Extracts resulting in a %viability < 30 are considered highly cytotoxic, samples with %viability between 30 and 60 are intermediately cytotoxic, and samples with %viability > 60 do not demonstrate cytotoxic activity. The results (Table 1) show eight samples to possess high cytotoxic activity against the non-cancerous WI-38 cell line, while all other samples are not active. None of the samples presents high antiproliferative activity against the cancerous HeLa cell line, but ten samples are determined having intermediate activity. When comparing the antiproliferative capacities toward both cell lines, the WI-38 cells are more sensitive than HeLa. Of the 39 samples, eight *Mallotus* extracts seem to be promising for the search of cytotoxic compounds. The eight samples belong to following species: *Mallotus apelta* (MA01 and MA02), *Mallotus macrostachyus* (MA11), *Mallotus microcarpus* (MA12), *Mallotus cuneatus* (MA16), *Mallotus yunnanensis* (MA19), and two *Mallotus* samples of unknown species (MA07 and MA18). Additionally, the IC_{50} of the positive control camptothecin was determined to be 0.7 μM for WI-38 and 0.5 μM for HeLa. For the standard compound Malloapelta B, isolated and purified from the leaves of *M. apelta*, the IC_{50} was determined to be 13 μM.

4.2. Evaluation of the cytotoxic activity based on HPLC fingerprint profiles

HPLC fingerprints of the 39 *Mallotus* samples have been optimized and developed as described in Ref. [38]. Because of the diversity of the samples, belonging to at least 17 different *Mallotus* species, harvested in different geographic regions, in different seasons (Table 1), the fingerprints are very different (Fig. 1).

In the following sections, the recorded fingerprints are linked to the cytotoxic activity by O-PLS. A calibration model for the cytotoxic activity as a function of the fingerprints is constructed. By studying the regression coefficients of the resulting model, the peaks important for the cytotoxic activity of the samples can be indicated and, in a next step, identified by, e.g. LC–MS. Prior to the construction of the model, the chromatographic data is preprocessed and an exploratory analysis is performed.

4.2.1. Data preprocessing

Prior to chemometric treatment of the data, the fingerprints are aligned by correlation optimized warping to correct for retention time shifts. Additionally, normalization followed by column centering is applied to evaluate the between-sample variation of the

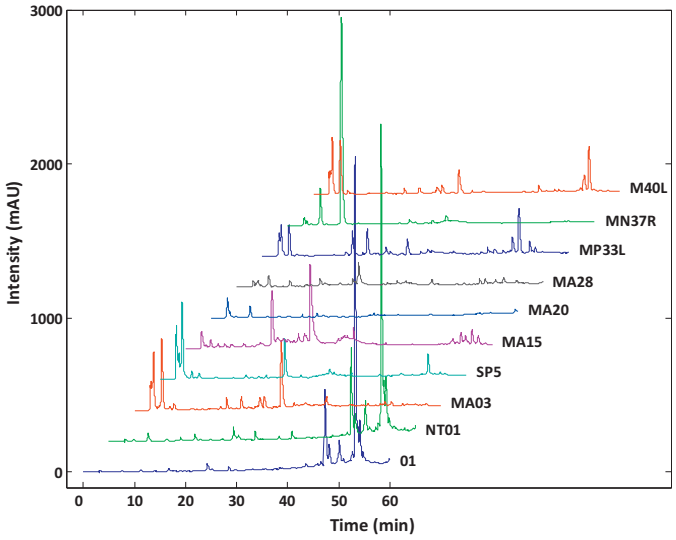


Fig. 1. 60 min fingerprints of the *Mallotus* extracts on the coupled Chromolith™ Performance RP-18e columns demonstrating the diversity of the samples.

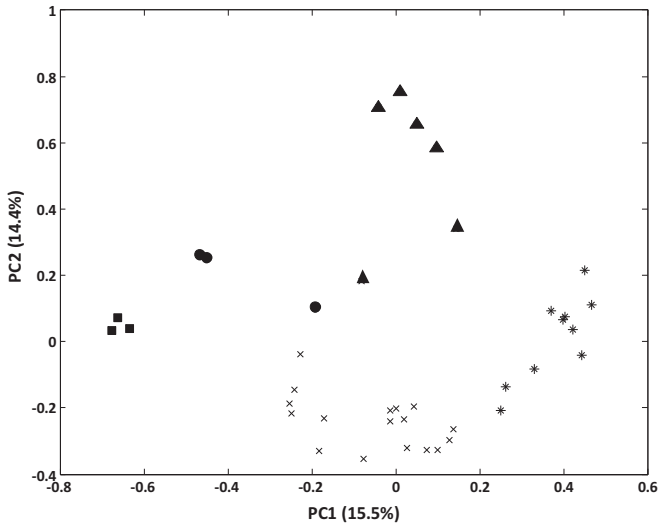


Fig. 2. PC1–PC2 score plot of the unaligned fingerprints of the 39 *Mallotus* samples. The available a priori information on the species, part of the plant used, the fingerprint profiles, combined with the proximity on the score plot resulted in the distinction of five subgroups used to develop a warping solution for the data set.

variables and to remove undesired effects due to unequal concentrations in the injected sample, respectively.

Because of the great divergence in the fingerprints of the different *Mallotus* species, peak alignment turned out to be far from evident. Since no diode array detector (DAD) or mass spectrometry (MS) data were available for all samples, a strategy to align the fingerprints was developed based on the available MS spectra (Table 1 – marked with *) and the clustering tendency of the samples as visualized by Principal Component Analysis. In a first step, all samples belonging to the same species are organized in groups and aligned. Next, all available a priori information on the species, the part of the plant used and the fingerprint profiles were combined with the proximity on the PC1–PC2 score plot of the unaligned data. This resulted in the distinction of five subgroups (Fig. 2): group (■) consisting of *M. apelta* and an unknown sample, group (●) consisting of *Mallotus nanus* samples, group (▲) consisting of *Mallotus paniculatus* and two unknown samples, and groups (×) and (∗) based on the similarity of the fingerprint profiles. To align the samples, the correlation coefficients between the fingerprints within each group are determined and the sample with the highest average correlation coefficient is selected as the target profile for COW. Finally, the individually aligned groups are recombined and a final warping solution was constructed. During the entire strategy, the alignment is carefully monitored using the available MS spectra to avoid the mismatching of compounds. To evaluate whether or not changes to the interpretation of the data occur by the time consuming alignment procedure, the exploratory analysis and the construction of the model are performed for both the unaligned and aligned data sets.

4.2.2. Exploratory analysis: Principal Component Analysis

To verify whether groups of samples, occasionally with similar cytotoxic activity, could be distinguished, Principal Component Analysis (PCA) is applied on the normalized and centered data for both the unaligned and aligned fingerprints (Fig. 3). When examining the PC1–PC2 score plot of the unaligned fingerprints (Fig. 3a), the samples with cytotoxic activity (marked in bold) are not densely clustered. However, the combined information on the proximity of the samples on the score plot, the fingerprint profiles, and the a priori knowledge on the species with the cytotoxic activity, allows the assumption that *Mallotus sp1* (sample 4) can be identified as *M. apelta* since similar fingerprints are obtained for its nearest neighbors (samples 8 and 9), which are *M. apelta* samples with similar cytotoxic activity. Furthermore, the unidentified samples *Mallotus sp5* and *Mallotus sp6* (samples 38 and 39) most probably belong to *M. paniculatus* based on their similarity with the *M. paniculatus* cluster (samples 6–10–29–30–31–32–33–34). Similar conclusions are obtained when evaluating the PC1–PC2 score plot of the aligned fingerprints (Fig. 3b), no additional or diverging information is observed.

4.2.3. Indication of the potentially cytotoxic compounds

To indicate the peaks with cytotoxic activity, the O-PLS model was constructed, relating the fingerprints to the cytotoxic activity. Because of the limited number of samples and as the model is not intended for the prediction of new samples, the data was not split into calibration and validation sets. The multivariate models are optimized following a leave-one-out cross validation procedure and their abilities to predict the cytotoxic activity of the available samples, and thus accurately indicate the peaks of interest, is evaluated prior to analyzing the regression coefficients and performing LC–MS experiments to identify the indicated peaks. The

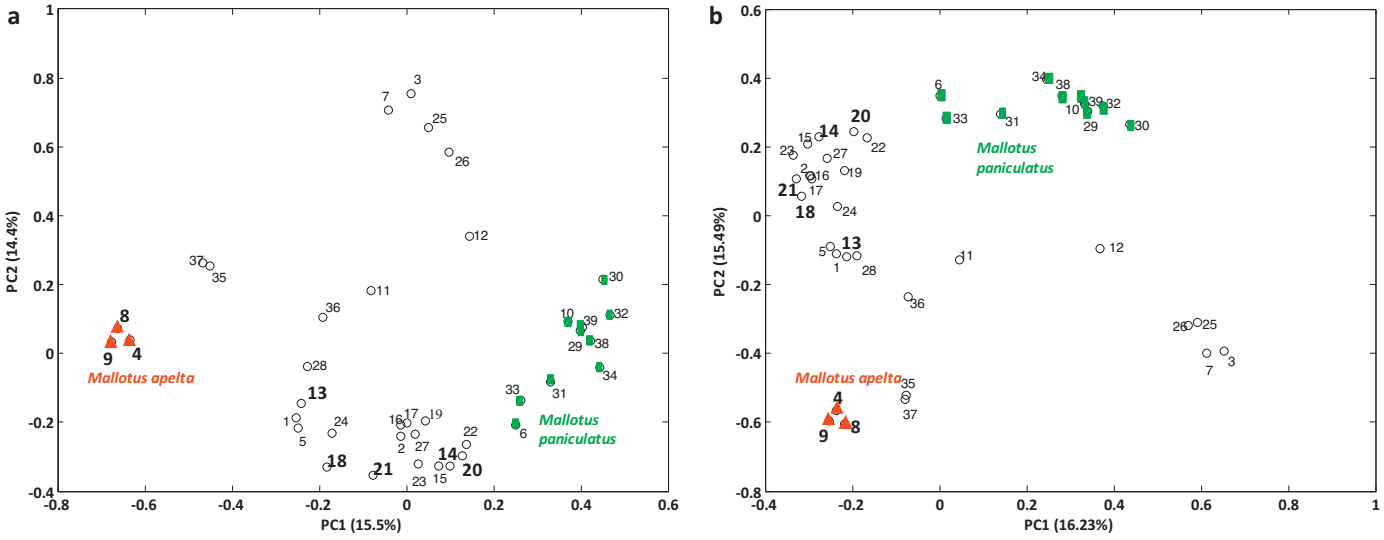


Fig. 3. (a) PC1–PC2 score plot of the unaligned fingerprints did not cluster the cytotoxic samples (bold), but allowed for the identification of the unknown sample 4 as *Mallotus apelta* and the unknown samples 38 and 39 as *Mallotus paniculatus*. (b) Similar conclusions are drawn from the PC1–PC2 score plot of the aligned fingerprints.

Table 2
The multivariate models for the non-cancerous cell line (WI-38) prior (model 1) and after (model 2) leaving out the badly predicted cytotoxic samples (MA11, MA12, MA16, MA18 and MA19). Only the cytotoxic samples are shown.

	Model 1			Model 2		
	#Factors	RMSECV	RMSE	#Factors	RMSECV	RMSE
Unaligned	1 (1)	25.8	23.6	1 (1)	13.7	19.0
Samples (measured)		Prediction			Prediction	
MA07 (14.0)		−8.9			22.2	
MA01 (16.0)		−10.1			18.5	
MA02 (13.0)		−13.4			19.2	
MA11 (13.3)		51.3			–	
MA12 (24.3)		70.3			–	
MA16 (13.5)		40.4			–	
MA18 (12.8)		70.1			–	
MA19 (15.3)		45.4			–	
	Model 1			Model 2		
	# Factors	RMSECV	RMSE	#Factors	RMSECV	RMSE
Aligned	1 (1)	26.1	24.1	1 (1)	14.2	19.1
Samples (measured)		Prediction			Prediction	
MA07 (14.0)		−5.2			23.9	
MA01 (16.0)		−6.1			20.3	
MA02 (13.0)		−9.4			21.0	
MA11 (13.3)		51.9			–	
MA12 (24.3)		69.3			–	
MA16 (13.5)		46.2			–	
MA18 (12.8)		68.7			–	
MA19 (15.3)		47.4			–	

peaks corresponding to potentially cytotoxic compounds are indicated by negative peaks in the regression coefficients plot since a strong antiproliferative activity corresponds to a low viability of the cell lines.

For both cell lines and for both the unaligned and aligned fingerprints, model optimization resulted in the removal of one orthogonal PLS component. Alignment of the fingerprints resulted in small changes of the root mean squared error of cross validation (RMSECV) and the root mean squared error (RMSE). While the modeling of the cancerous HeLa cell line resulted in acceptable prediction errors of 9.4% (unaligned) and 9.5% (aligned), the models for the non-cancerous cell line resulted in high prediction errors of 23.6% (unaligned) and 24.1% (aligned) (Table 2). A closer look at the prediction of the samples of interest reveals close approximations between the predicted and the experimental values for the cancerous cell line. However, for the non-cancerous cell line substantial deviations between both values are observed. Of the eight cytotoxic samples, samples MA11, MA12, MA16, MA18 and MA19 are predicted to have far less cytotoxic activity than experimentally determined. Only samples MA01, MA02 and MA07 are considered as highly cytotoxic by the model, but the predicted values exceed the experimental ones. For all models, none of the samples without experimental cytotoxic activity was predicted as cytotoxic.

Because of the largely deviating prediction of the antiproliferative activity against the non-cancerous cell line, the decision is made to eliminate samples MA11, MA12, MA16, MA18 and MA19 for the indication of the cytotoxic peaks and repeat the calculations. For both the unaligned and aligned fingerprints, optimization of the new model resulted in the removal of one orthogonal PLS component, while the alignment resulted in only small changes in the RMSECV (13.7% against 14.2%) and RMSE (19.0% against 19.1%). These results are a considerable improvement to the RMSECV and RMSE of the models with the five samples included. Additionally, the prediction of samples MA01, MA02 and MA07 is much more accurate (Table 4). Hence, the regression coefficients of the latter obtained models are also considered for evaluation. One should note that for the cancerous cell line, the models obtained without samples MA11, MA12, MA16, MA18 and MA19 are similar to the model with all samples included. All the above discussed results are summarized in Tables 2 (non-cancerous WI-38) and 3 (cancerous HeLa).

4.2.3.1. Regression coefficients. Despite the inaccurate predictions for five out of eight highly active samples on the non-cancerous cell line, no apparent reason for the deviating results was found. Consequently, the samples were not removed for the data but their reliability was questioned. This implied that the indication of the potentially cytotoxic compounds was evaluated by eight different models from two cell lines, prior and after alignment, built with 34 or 39 samples. Besides the indication of the potentially cytotoxic compounds, the influence of aligning such complex and diverse fingerprints for the indication of compounds of interest is discussed. Afterwards, within each cell line, the models obtained with 34 and 39 samples are compared, carefully monitoring the differences between the estimated regression coefficients and evaluating potential causes.

The regression coefficients prior and after alignment, for both cell lines, are evaluated first. Based on the full data for the non-cancerous cell line, evaluation of the O-PLS regression coefficients of the unaligned fingerprints results in the indication of six potentially cytotoxic peaks (38.5, 39.3, 42.5, 49.0, 49.8 and 54.5 min) in the fingerprints. The indicated peaks are found to be identical to those indicated by the model built with the aligned fingerprints taking into account small fluctuations in the retention times because of the alignment procedure (Fig. 4a). For the cancerous cell line, the evaluation of the regression coefficients prior and after alignment resulted in the indication of some peaks already indicated for the non-cancerous cell line. Again six peaks (38.5, 42.5, 49.0, 49.8, 53.1 and 54.5 min) were observed with negative regression coefficients (Fig. 4b). When comparing the results of the four models, seven different peaks are indicated of which five are found from every model (38.5, 42.5, 49.0, 49.8 and 54.5 min). The peak at 39.3 min is only indicated from the non-cancerous model and that at 53.1 min only from the cancerous model. The latter results may indicate selective

Table 3
The multivariate models for the cancerous cell line (HeLa) prior (model 1) and after (model 2) leaving out the badly predicted cytotoxic samples (MA11, MA12, MA16, MA18 and MA19). Only the cytotoxic samples are shown.

	Model 1			Model 2		
	#Factors	RMSECV	RMSE	#Factors	RMSECV	RMSE
Unaligned	1 (1)	11.1	9.4	1 (1)	9.9	10.0
Samples (measured)		Prediction			Prediction	
MA07 (47.7)		42.8			46.4	
MA01 (43.0)		43.0			45.7	
MA02 (50.2)		41.1			44.8	
MA11 (41.3)		51.5			–	
MA12 (52.7)		62.0			–	
MA16 (46.7)		50.2			–	
MA18 (43.7)		62.7			–	
MA19 (61.3)		55.1			–	
	Model 1			Model 2		
	#Factors	RMSECV	RMSE	#Factors	RMSECV	RMSE
Aligned	1 (1)	10.7	9.5	1 (1)	9.7	10.0
Samples (measured)		Prediction			Prediction	
MA07 (47.7)		45.0			48.5	
MA01 (43.0)		43.0			48.0	
MA02 (50.2)		41.1			47.1	
MA11 (41.3)		51.5			–	
MA12 (52.7)		62.0			–	
MA16 (46.7)		50.2			–	
MA18 (43.7)		62.7			–	
MA19 (61.3)		55.1			–	

cytotoxicity of the indicated compounds against non-cancerous or cancerous cells, respectively.

Additionally, as the predictions of the cytotoxic activity of samples MA11, MA12, MA16, MA18 and MA19 for the non-cancerous cell line are significantly different from the experimental results, these samples were removed from modeling. The models were rebuilt for both cell lines to compare both cytotoxic experiments. Similar to the results of the full data, no major differences are observed between the regression coefficients obtained prior and after alignment. However, less peaks are indicated: for the non-cancerous cell line two peaks remain indicated (38.5 and 42.5 min), while for the cancerous cell line three peaks are (38.5, 42.5 and 53.1 min).

Evaluation of the different models (Table 4) leads to the indication of identical chromatographic peaks (38.5 and 42.5 min) for samples MA01, MA02 and MA07. Furthermore, for the non-cancerous cell line, the peaks indicated at 39.3, 49.0, 49.8 and 54.5 min correspond to compounds present in samples MA11, MA12, MA16, MA18 and MA19. This is consistent with the indicated peaks from the model on the reduced data set: removing the five samples results in the disappearance of the corresponding coefficient peaks. A similar situation occurs with the cancerous cell line data: the peaks at 49.0, 49.8 and 54.5 min correspond to compounds present in the five removed samples. Hence, they are not indicated by the regression coefficients when modeling the reduced data set. However, one should note the indication of

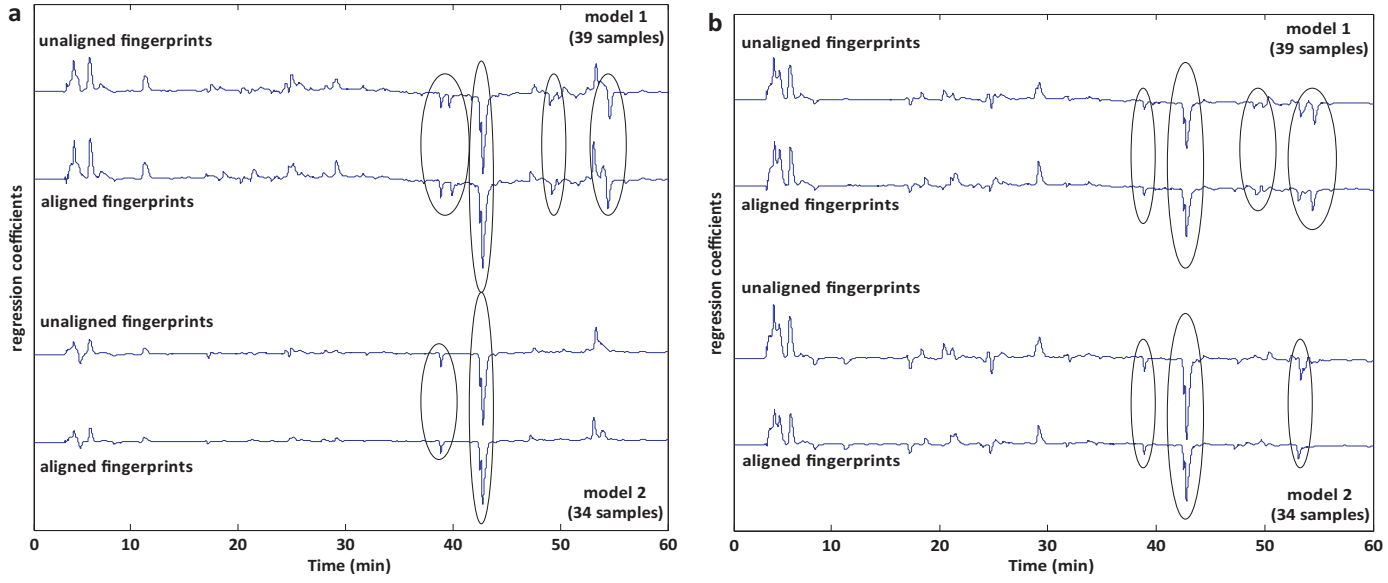


Fig. 4. The O-PLS regression coefficients prior and after alignment of the fingerprints for the non-cancerous (a) and cancerous (b) cell lines.

Table 4
Indicated peaks for the non-cancerous (WI-38) and cancerous (HeLa) cell lines prior and after alignment by COW. The peaks marked with an * are also observed in the models based on the reduced data set (leaving out samples MA11, MA12, MA16, MA18 and MA19).

WI-38		HeLa	
Unaligned	Aligned	Unaligned	Aligned
38.5*	38.5*	38.5*	38.5*
39.3	39.5	–	–
42.5*	42.7*	42.5*	42.7*
49.0	48.9	49.0	48.9
49.8	49.8	49.8	49.8
–	–	53.1*	53.0*
54.5	54.4	54.5	54.4

the peak at 53.1 min. Although the corresponding compound was identified in sample MA11, after removal of the sample the model still indicates a regression coefficient peak for the cancerous cell line. A closer look at the fingerprints reveals the presence of a corresponding chromatographic peak in sample MA22 (unidentified *Mallotus* sample), a sample with low cytotoxic activity (67.0%).

4.2.4. Identification of the indicated cytotoxic compounds

As no differences are observed between the peaks indicated prior and after alignment of the data, the discussed results use the regression coefficients from the unaligned fingerprints.

Once the peaks of interest are indicated based on the regression coefficients of the models, identification is performed by LC–MS. All MS analyses are performed in both the positive and negative modes. Because of the presence of TFA as additive, many of the analyzed compounds are bound to TFA in the MS spectra when analyzing in the negative mode, causing a difference of +113 amu. During analyses in the positive mode, this problem does not occur. In the negative mode, all reported values in this paper are corrected for the addition of TFA to avoid confusion. As for the non-cancerous cell line, the peaks indicated at 39.3, 49.0 and 54.5 min correspond to chromatographic peaks present in the samples with inaccurate prediction of the cytotoxic activity, the reliability of the indication of these samples is questioned (Section 4.2.3). Nonetheless, the samples are analyzed by LC–MS and the results reported for the eight cytotoxic samples against the regression coefficients obtained with the full data set as well as with the reduced data.

Fig. 5 visualizes the indication and identification of the peaks of interest for the samples presenting cytotoxic activity against the non-cancerous (Fig. 5a) and cancerous cell lines (Fig. 5b),

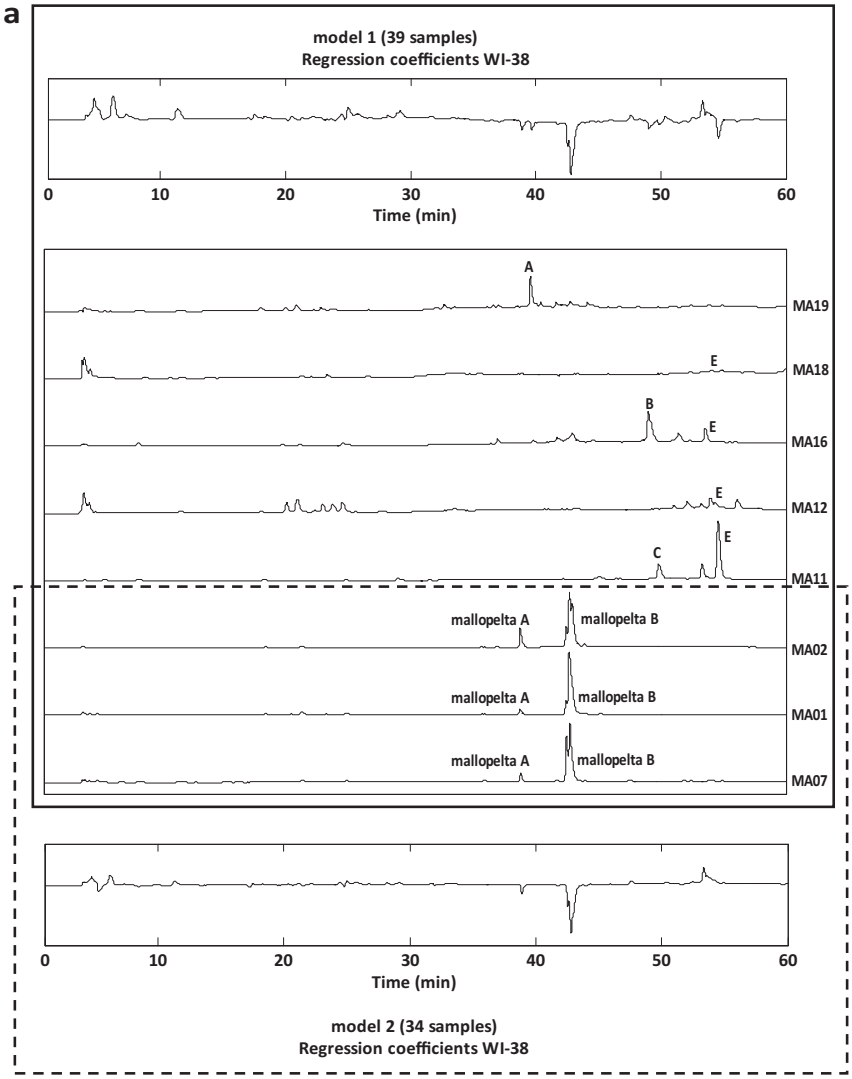


Fig. 5. Indication of the identified and unknown compounds. The fingerprints of the *Mallotus apelta* samples are plotted against the regression coefficients from the (a) non-cancerous and (b) cancerous cell lines.

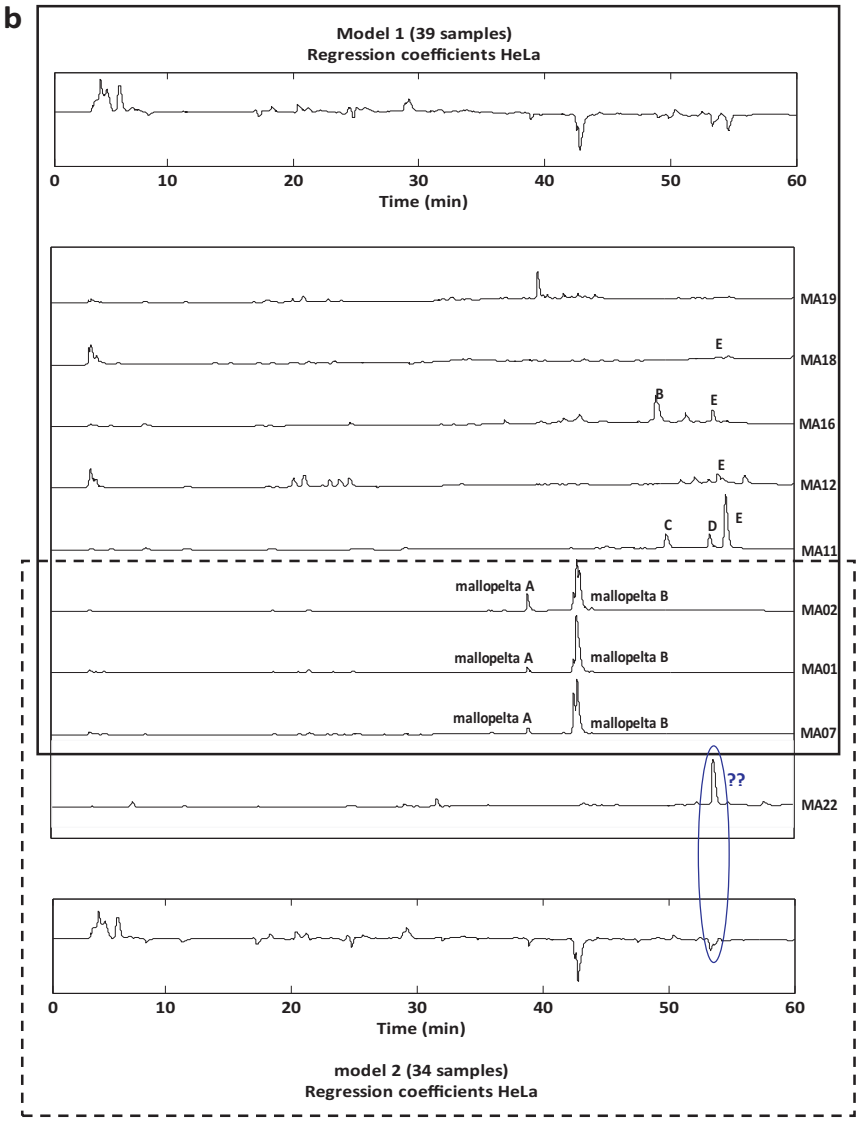


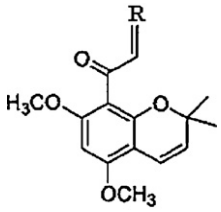
Fig. 5. (*Continued*).

respectively. The following sections briefly describe the results of the LC–MS analyses of the eight cytotoxic samples and discuss the indicated (un)identified compounds. More information can be found in Table 5.

4.2.4.1. Identified compounds. Analysis of the indicated chromatographic peaks of the cytotoxic samples resulted in the identification of two known compounds, i.e. Malloapelta A and Malloapelta B (Figs. 5 and 6). Both compounds are present in the three *M. apelta* samples (MA01, MA02 and MA07) and elute at 38.5 min and 42.5 min, respectively. The compounds are identified by LC–MS in the positive mode at *m*/*z* 301 $[M+H]^+$ and *m*/*z* 321 $[M+H]^+$ respectively, which is in accordance to the findings of Ref. [22] describing both compounds. The indicated antiproliferative activity toward both cell lines is confirmed based on earlier reports on the cytotoxic activity of the purified compounds [22,23]. Considering the proven cytotoxic activity of the compounds and the *M. apelta* species, it is not surprising to find both compounds indicated by the regression coefficients of all models, regardless the cell line considered.

Table 5
Results from the LC–MS analysis for the cytotoxic samples. Compounds detected in the negative mode are marked with $^-$, compounds detected in the positive mode with $^+$. The corresponding mass values are put between brackets.

Retention time (min)	Cell line	MA07	MA01	MA02	MA11	MA12	MA16	MA18	MA19
38.5	WI38 HeLa	Malloapelta A (307^+)	Malloapelta A (307^+)	Malloapelta A (307^+)					
39.3	HeLa								A (389^+)
42.5	WI38 HeLa	Malloapelta B (321^+)	Malloapelta B (321^+)	Malloapelta B (321^+)					
49.0	HeLa						B (711^+)		
49.8	HeLa				C (466^+)				
53.1	HeLa				D (275^+)				
54.5	HeLa	E (607^+)	E (607^+)	E (607^+)	E (607^+)	E (607^+)	E (607^+)	E (607^+)	E (607^+)



1. R = H
2. R = CH3

Fig. 6. Structure of the standard compounds Malloapelta A (1) and Malloapelta B (2).

4.2.4.2. Unidentified compounds. When considering the unidentified compounds, differences are found between the models. A grand total of five unidentified peaks are indicated at 39.3, 49.0, 49.8, 53.1 and 54.5 min. The negative regressions coefficients correspond to compound A from sample MA19 eluting at 39.3 min, compound B from sample MA16 eluting at 49.0 min, and compounds C, D and E from sample MA11 at 49.8, 53.1 and 54.5 min, respectively. For MA12, MA16, MA18 and MA19, compound E (54.4 min) is also identified.

Based on the above described results, structure elucidation and purification of the unidentified compounds might lead to interesting new cytotoxic molecules. Especially compounds A and D may be interesting as they demonstrate selective cytotoxic activities: compound A is only indicated for the non-cancerous cell line, while compound D is only active against the cancerous cell line. However, one should be careful when considering compound D. Although the compound was identified in sample MA11, removal of the sample from the cancerous model still indicated a regression coefficient peak at the same retention time corresponding to a chromatographic peak of the slightly cytotoxic sample MA22. If it turns out that the sample also contains compound D, it presents at best low cytotoxic activity and may not be interesting for further analysis. Unfortunately, no MS data on this sample is available neither can a literature study provide an outcome as the sample is unidentified.

5. Conclusions

To indicate and identify the compounds potentially responsible for the antiproliferative activity of some *Mallotus* species, HPLC fingerprints of 39 samples are developed and the antiproliferative capacity of the samples against non-cancerous (WI-38) and cancerous (HeLa) cell lines is evaluated. Prior to the construction of multivariate calibration models, the fingerprints are aligned based on a warping strategy combining correlation optimized warping with the knowledge retrieved from an exploratory analysis and from the LC–MS spectra.

Then, multivariate models for the cytotoxic activity from the unaligned and aligned fingerprints are constructed by Orthogonal Projections to Latent Structures for both cell lines. The regression coefficients are studied to indicate the peaks potentially responsible for the activity. For both cell lines, no differences are found to indicate the peaks of interest prior and after alignment of the fingerprints. Additionally, the model for the activity on the non-cancerous cell line results in large differences between the predicted and experimental values for five out of eight highly active samples. The inaccurate prediction of these samples has led to questioning the reliability of the obtained regression coefficients. Therefore, these samples are excluded from the data set and the calculations repeated. The indication of the interesting peaks from both the full and reduced data sets as well as the unaligned and aligned data is evaluated. The unaligned and aligned data sets result in the indication of the same peaks, while the differences in the indicated peaks between the full and reduced data sets correspond to the chromatographic peaks observed in the removed fingerprints.

LC–MS analyses of the indicated peaks identify Malloapelta A and Malloapelta B as cytotoxic compounds against both cell lines, an observation that is confirmed from the literature. Furthermore, five unknown but potentially interesting compounds are observed underlying the indicated peaks. Especially compounds A and D prove to be interesting as they demonstrate selective cytotoxic activities: compound A is only indicated for the non-cancerous cell line, while compound D is only active against the cancerous cell line. Structure elucidation and purification of these compounds might lead to interesting new cytotoxic molecules.

Acknowledgements

The authors gratefully thank the Belgian Science Policy Office (BELSPO) Bilateral Project (BIL/03/V09) between Belgium and Viet Nam for financial support on this research. Bieke Dejaegher is a post-doctoral fellow of the FWO.

References

[1] Traditional Medicine Strategy 2002–2005, World Health Organization, Geneva, 2005, Available from: http://www.who.int/medicines/publications/traditionalpolicy/en/ (accessed 10.05.11).
[2] National Policy on Traditional Medicine and Regulation of Herbal Medicines, World Health Organization, Geneva, 2005, Available from: http://apps.who.int/medicinedocs/pdf/s7916e/s7916e.pdf (accessed 10.05.11).
[3] Y. Fujii, H. Fujii, M. Yamazaki, J. Chromatogr. 258 (1983) 147.
[4] R.J. Miller, C. Jolles, H. Rapoport, Phytochemistry 12 (1973) 597.
[5] G. Strobel, A. Stierle, W.M. Hess, Plant Sci. 101 (1994) 115.
[6] V. Cianchino, G. Acosta, C. Ortega, L.D. Martinez, M.R. Gomez, Food Chem. 108 (2008) 1075.
[7] S. Lopez-Feria, S. Cardenas, J.A. Garcıa-Mesa, M. Valcarcel, Talanta 75 (2008) 937.
[8] P.A. Whitton, A. Lau, A. Salisbury, J. Whitehouse, C.S. Evans, Phytochemistry 64 (2003) 673.
[9] Declaration of Alma-Ata, International Conference on Primary Health Care Alma-Ata, USSR, 6–12 September 1978, 1978, Available from: http://www.who.int/hpr/NPH/docs/declaration_almaata.pdf (accessed 10.05.11).
[10] United States Pharmacopoeia 33 – National Formulary 28, United States Pharmacopeia, Rockville, MD, USA, 2010.
[11] Pharmacopoeia of the People's Republic of China – English Edition, Chinese Pharmacopoeia Commission, People's Medical Publishing House, Beijing, China, 2010.
[12] M.K. Lee, Y.M. Ahn, K.R. Lee, J.H. Jung, O-S. Jung, J. Hong, Anal. Chim. Acta 633 (2009) 271.
[13] G. Chou, S.J. Xu, D. Liu, G.Y. Koh, J. Zhang, Z. Lui, J. Agric. Food Chem. 57 (2009) 1076.
[14] M. Dumarey, A.M. van Nederkassel, E. Deconinck, Y. Vander Heyden, J. Chromatogr. A 1192 (2008) 81.
[15] C. Tistaert, B. Dejaegher, N. Nguyen Hoai, G. Chataigné, C. Rivière, V. Nguyen Thi Hong, M. Chau Van, J. Quetin-Leclerq, Y. Vander Heyden, Anal. Chim. Acta 649 (2009) 24.
[16] F.T. Chau, H.Y. Chan, C.Y. Cheung, C.J. Xu, Y. Liang, Anal. Chem. 81 (2009) 7217.
[17] T. Rajalahti, R. Arneberg, F.S. Berven, K.M. Myhr, R.J. Ulvik, O.M. Kvalheim, Chemom. Intell. Lab. Syst. 95 (2009) 35.
[18] T. Rajalahti, R. Arneberg, A.C. Kroksveen, M. Berle, K.M. Myhr, O.M. Kvalheim, Anal. Chem. 81 (2009) 2581.
[19] S. Wiklund, E. Johansson, L. Sjostrom, E.J. Mellerowicz, U. Edlund, J.P. Shockcor, J. Gottfries, T. Moritz, J. Trygg, Anal. Chem. 80 (2008) 115.
[20] C. Rivière, V. Nguyen Thi Hong, Q. Tran Hong, G. Chataigné, N. Nguyen Hoai, B. Dejaegher, C. Tistaert, T. Nguyen Thi Kim, Y. Vander Heyden, M. Chau Van, J. Quetin-Leclercq, Phytochem. Rev. 9 (2010) 217.
[21] S.C. Thakur, S.S. Thakur, S.K. Chaube, Reprod. Toxicol. 20 (2005) 149.
[22] T. Tanaka, T. Ito, M. Linuma, Y. Takahashi, H. Naganawa, Phytochemistry 48 (1998) 1423.
[23] S.H. Fang, Y.K. Rao, Y.M. Tseng, J. Ethnopharmacol. 116 (2008) 333.
[24] W.H. Hui, M.M. Li, Phytochemistry 15 (1976) 985.
[25] P. Van Kiem, C.V. Minh, H.T. Huong, N.H. Nam, J.J. Lee, Y.H. Kim, Arch. Pharm. Res. 27 (2004) 1109.
[26] P. Van Kiem, N.H. Dang, H.V. Bao, H.T. Huong, M.V. Chau, L.M. Huong, J.J. Lee, Y.H. Kim, Arch. Pharm. Res. 28 (2005) 1131.
[27] C.V. Luu, M.V. Chau, J.J. Lee, S.H. Jung, Arch. Pharm. Res. 29 (2006) 840.
[28] T. Mosmann, J. Immunol. Methods 65 (1983) 55.
[29] S. Block, C. Baccelli, B. Tinant, L. Van Meervelt, R. Rozenberg, J.L. Habib Jiwan, M.C. De Pauw-Gillet, J. Quetin-Leclercq, Phytochemistry 65 (2004) 1165.
[30] J. Trygg, S. Wold, J. Chemom. 16 (2002) 119.

[31] L. Eriksson, E. Johansson, N. Kettaneh-Wold, J. Trygg, C. Wikström, S. Wold, Multi- and Megavariate Data Analysis, Part I Basic Principles and Applications, Umetrics AB, Umea, 2006.
[32] D.L. Massart, B.G.M. Vandeginste, L.M.C. Buydens, S. de Jong, P.J. Lewi, J. Smeyers-Verbeke, Handbook of Chemometrics and Qualimetrics. Part A, Elsevier, Amsterdam, 1997.
[33] G. Tomasi, F. van den Berg, C. Andersson, J. Chemom. 18 (2004) 231.
[34] N.-P. Vest Nielsen, J.M. Carstensen, J. Smedsgaard, J. Chromatogr. A 805 (1998) 17.
[35] A.M. van Nederkassel, M. Daszykowski, P.H.C. Eilers, Y. Vander Heyden, J. Chromatogr. A 1118 (2006) 199.
[36] B. Walczak, W. Wu, Chemom. Intell. Lab. Syst. 77 (2005) 173.
[37] H. Martens, M. Martens, Multivariate Analysis of Quality, An Introduction, Wiley, Chichester, 2001.
[38] N. Nguyen Hoai, B. Dejaegher, C. Tistaert, V. Nguyen Thi Hong, C. Rivière, C. van Minh, J. Quetin-Leclercq, Y. Vander Heyden, J. Pharm. Biomed. Anal. 50 (2009) 753.