

# Aggressive dereplication using UHPLC–DAD–QTOF: screening extracts for up to 3000 fungal secondary metabolites

Andreas Klitgaard · Anita Iversen · Mikael R. Andersen · Thomas O. Larsen · Jens Christian Frisvad · Kristian Fog Nielsen

Received: 11 September 2013 / Revised: 3 December 2013 / Accepted: 14 December 2013 / Published online: 18 January 2014  
© The Author(s) 2014. This article is published with open access at Springerlink.com

**Abstract** In natural-product drug discovery, finding new compounds is the main task, and thus fast dereplication of known compounds is essential. This is usually performed by manual liquid chromatography–ultraviolet (LC–UV) or visible light–mass spectroscopy (Vis–MS) interpretation of detected peaks, often assisted by automated identification of previously identified compounds. We used a 15 min high-performance liquid chromatography–diode array detection (UHPLC–DAD)–high-resolution MS method (electrospray ionization (ESI)<sup>+</sup> or ESI<sup>-</sup>), followed by 10–60 s of automated data analysis for up to 3000 relevant elemental compositions. By overlaying automatically generated extracted-ion chromatograms from detected compounds on the base peak chromatogram, all major potentially novel peaks could be visualized. Peaks corresponding to compounds available as reference standards, previously identified compounds, and major contaminants from solvents, media, filters etc. were labeled to differentiate these from compounds only identified by elemental composition. This enabled fast manual evaluation of both known peaks and potential novel-compound peaks, by manual verification of: the adduct pattern, UV–Vis, retention time compared with log D, co-identified biosynthetic related compounds, and elution order. System performance, including adduct patterns, in-source fragmentation, and ion-cooler bias,

was investigated on reference standards, and the overall method was used on extracts of *Aspergillus carbonarius* and *Penicillium melanoconidium*, revealing new nitrogen-containing biomarkers for both species.

**Keywords** Metabolomics · Mycotoxin · NRPS · LC–MS · UPLC · Polyketide · Nonribosomal peptide

## Introduction

Fungi are an immense source of diverse natural products that can be used as drugs, food and feed additives, and industrial chemicals [1, 2]. Unfortunately fungi also have a negative side, producing mycotoxins which include some of the most immunotoxic, estrogenic, cytotoxic, and carcinogenic compounds known [3, 4].

Fast and accurate dereplication of previously described compounds is an essential and resource-saving aspect of working with natural products [1, 5–9]. The alternative, isolation and subsequent NMR-based structure elucidation, is time consuming and costly [7], and is thus primarily used in important cases, e.g. for compounds with known bioactivity.

Currently, dereplication is mainly performed by liquid chromatography–mass spectrometry (LC–MS) analysis of extracts, followed by a search of all ions of interest performed by entering the monoisotopic mass into appropriate databases. For microbial compounds, the most comprehensive database is AntiBase (Wiley–VCH, Weinheim, Germany) the 2012 version of which contains 41,000 recorded compounds. In dereplication, obtaining an elemental composition is the most efficient first step because it reduces the number of hits from a database search 3–10-fold compared with searching for a nominal mass [9–11]. For compounds below 400–600 Da, high-resolution MS (HRMS) instruments can often provide the elemental composition unambiguously if they have < 0.5–

**Electronic supplementary material** The online version of this article (doi:10.1007/s00216-013-7582-x) contains supplementary material, which is available to authorized users.

A. Klitgaard · A. Iversen · M. R. Andersen · T. O. Larsen · J. C. Frisvad · K. F. Nielsen (✉)  
Department of Systems Biology, Søtofts Plads, Technical University of Denmark, 2800 Kgs., Lyngby, Denmark  
e-mail: kfn@bio.dtu.dk

A. Iversen  
Current address: Danish Emergency Management Agency,  
Universitetsparken 2, 2100 Copenhagen, Denmark

1.5 ppm mass accuracy. In addition, time of flight (TOF)-based mass spectrometers can now provide an accurate isotope pattern, enabling an even higher degree of certainty for identification of elemental compositions [9, 12, 13].

An important extra detector is the UV–Vis diode array detection (DAD) detector, which provides information on the conjugated double-bond systems found in most secondary metabolites. This can be used to confirm or reject candidates from a database search [14, 15]. Finally, log D-based calculations can be used to predict the chromatographic elution order of compounds of interest [9].

Dereplication of peaks in extracts from genera, including *Aspergillus*, *Penicillium*, and *Fusarium*, which are known to produce many different compounds often results in many hits (1724, 1726, and 611 compounds, respectively, listed in AntiBase). Because of this, identifying compounds on the basis of UV–Vis, chromatographic retention, elution order, and comparison to biosynthetically related compounds is a slow (0.5–3 h per extract) and tedious task.

A solution could be to use MS–MS libraries [16] to identify compounds automatically. This is the preferred strategy in forensic science and toxicology, for which subjects commercial compound libraries are available [17]. However, no natural-product MS–MS libraries are currently available, because including an MS–MS spectrum for future dereplication is unfortunately not a prerequisite for publishing new structures. Because of this, only a few percent of described compounds from fungi are commercially available, and therefore only small in-house databases are available [9, 18, 19].

Another complication is that the compound adduct pattern and possible fragmentations need to be correctly interpreted, because unnoticed loss of water or addition of sodium or ammonium ions will invalidate a subsequent database search. Unambiguous determination of the accurate mass of fungal metabolites on the basis of adduct formation, dimers, and mutably charged ions can be challenging [9], but software including ACDs IntelliXtract [19] and some instrument vendor software packages have algorithms for this.

To reduce the analysis time for known fungal compounds in complex extracts, we decided to test the TargetAnalysis software from Bruker Daltonics (similar software available from Waters, Thermo, Agilent, and Advanced Chemical Developments). The program was originally developed for pesticide [20] and forensic analysis [21]. TargetAnalysis can screen an extract for 3000 compounds, on the basis of mass accuracy, isotope fit, and retention time (RT), within 10–60 s, depending on how small peaks are integrated. The screening software was interfaced with our internal compound database, containing approximately 7100 compounds [9], via an in-house-built Excel application that generated automatic search lists for TargetAnalysis, and made it possible to search for the most likely adduct and/or fragment ions and to only include taxonomically relevant compounds if wanted.

Using this approach, we are able to rapidly screen extracts from several different fungi, and to annotate chromatographic peaks corresponding to known compounds. The approach makes it possible to easily identify chromatographic peaks that do not correspond to known compounds, thereby enabling one to quickly ascertain which compounds might be novel.

## Materials and methods

### Chemicals

Solvents were LC–MS grade, and all other chemicals were analytical grade. All were from Sigma-Aldrich (Steinheim, Germany) unless otherwise stated. Water was purified using a Milli-Q system (Millipore, Bedford, MA). ESI–TOF tune mix was purchased from Agilent Technologies (Torrance, CA, USA).

Reference standards of mycotoxins and microbial metabolites (approximately 1500, 95 % of fungal origin) had been collected over the last 30 years [9, 22, 23], either from commercial sources, as gifts from other research groups, or from our own projects. Approximately one-third of the standards were purchased from Sigma-Aldrich, Axxora (Bingham, UK), Cayman (Ann Arbor, MI), TebuBio (Le-Perray-en-Yvelines, France), Biopure (Tulln, Austria), Calbiochem, (San Diego, CA), and ICN (Irvine, CA). Standards were maintained dry at  $-20\text{ }^{\circ}\text{C}$ , and were compared with original UV–VIS data, accurate mass, and relative RT from previous studies [22].

Culture extracts in the examples originated from three-point cultures on solid media, incubated for seven days in darkness at  $25\text{ }^{\circ}\text{C}$ , and extracted using a (3:2:1) (ethyl acetate:dichloromethane:methanol) mixture [24]. *Penicillium melanoconidium* IBT 30549 (IBT culture collection, author's address) was grown on CYA, and *A. carbonarius* IBT 31236 (ITEM5010) was grown on YES [24].

### UHPLC–DAD–QTOFMS

A UHPCL–DAD–QTOF method was set up for screening, with typical injection volumes of 0.1–2  $\mu\text{l}$  extract. Separation was performed on a Dionex Ultimate 3000 UHPLC system (Thermo Scientific, Dionex, Sunnyvale, California, USA) equipped with a  $100\times 2.1\text{ mm}$ ,  $2.6\text{ }\mu\text{m}$ , Kinetex C<sub>18</sub> column, held at a temperature of  $40\text{ }^{\circ}\text{C}$ , and using a linear gradient system composed of A:  $20\text{ mmol L}^{-1}$  formic acid in water, and B:  $20\text{ mmol L}^{-1}$  formic acid in acetonitrile. The flow was  $0.4\text{ ml min}^{-1}$ , 90 % A graduating to 100 % B in 10 min, 100 % B 10–13 min, and 90 % A 13.1–15 min.

Time-of-flight detection was performed using a maXis 3G QTOF orthogonal mass spectrometer (Bruker Daltonics, Bremen, Germany) operated at a resolving power of  $\sim 50000$  full

width at half maximum (FWHM). The instrument was equipped with an orthogonal electrospray ionization source, and mass spectra were recorded in the range  $m/z$  100–1000 as centroid spectra, with five scans per second. For calibration, 1  $\mu\text{l}$  10  $\text{mmol L}^{-1}$  sodium formate was injected at the beginning of each chromatographic run, using the divert valve (0.3–0.4 min). Data files were calibrated post-run on the average spectrum from this time segment, using the Bruker HPC (high-precision calibration) algorithm.

For  $\text{ESI}^+$  the capillary voltage was maintained at 4200 V, the gas flow to the nebulizer was set to 2.4 bar, the drying temperature was 220  $^{\circ}\text{C}$ , and the drying gas flow was 12.0  $\text{L min}^{-1}$ . Transfer optics (ion-funnel energies, quadrupole energy) were tuned on HT-2 toxin to minimize fragmentation. For  $\text{ESI}^-$  the settings were the same, except that the capillary voltage was maintained at  $-2500$  V. Unless otherwise stated, ion-cooler settings were: transfer time 50  $\mu\text{s}$ , radio frequency (RF) 55 V peak-to-peak (Vpp), and pre-pulse storage time 5  $\mu\text{s}$ . After changing the polarity, the mass spectrometer needed to equilibrate the power supply temperature for 1 h to provide stable mass accuracy.

#### Construction of the compound database

The database was constructed in ACD/ChemFolder (Advanced Chemistry Development, Toronto, Canada) from:

1. reference standards (~1500) [9];
2. tentatively identified compounds (~500) [25–27];
3. compound peaks appearing in blank samples; and
4. all compounds in AntiBase2012 listed as coming from: *Aspergillus*, *Fusarium*, *Trichoderma*, *Penicillium*, *Chaetomium*, *Stachybotrys*, *Alternaria*, and *Cladosporium*.

A detailed description of the database construction can be found in the Electronic Supplementary Material, Section “Introduction”.

For each compound, the known or suspected major adducts were registered as:  $[\text{M}+\text{H}]^+$ ,  $[\text{M}+\text{Na}]^+$ ,  $[\text{M}+\text{NH}_4]^+$ ,  $[\text{M}+\text{K}]^+$ ,  $[\text{M}+\text{H}+\text{CH}_3\text{CN}]^+$ ,  $[\text{M}+\text{Na}+\text{CH}_3\text{CN}]^+$ ,  $[\text{M}+\text{H}-\text{H}_2\text{O}]^+$ ,  $[\text{M}+\text{H}-2\text{H}_2\text{O}]^+$ ,  $[\text{M}+\text{H}-\text{H}_2]^+$  (sterols),  $[\text{M}+\text{H}-\text{HCOOH}]^+$ ,  $[\text{M}+\text{H}-\text{CH}_3\text{COOH}]^+$ ,  $[\text{M}+2\text{H}]^{2+}$ ,  $[\text{M}+\text{Na}+\text{H}]^{2+}$  or  $[\text{M}+2\text{Na}]^{2+}$  or “No ionization” in  $\text{ESI}^+$ , and in  $\text{ESI}^-$ :  $[\text{M}-\text{H}]^-$ ,  $[\text{M}-\text{H}+\text{HCOOH}]^-$ , and  $[\text{M}+\text{Cl}]^-$ .

#### Creating search lists for targetanalysis

A Microsoft Excel application was created for sorting the ChemFolder database into a taxonomically relevant search-list for TargetAnalysis (elemental composition and charge state of desired adduct, and name of compound).

For labeling peaks in Bruker DataAnalysis 4.0 (DA), compounds that were available as reference standards were labeled “S-x” in front of the name. A description of the database creation procedure can be found in the Electronic Supplementary Material, Section “Introduction”.

#### Automated screening of fungal samples

TargetAnalysis 1.2 (Bruker Daltonics, Bremen, Germany), was used to process data-files, with the following typical settings:

- A) retention time (if known) as  $\pm 1.2$  min as broad, 0.8 min as medium, and 0.3 min as narrow range;
- B) SigmaFit; 1000 (broad) (isotope fit not used), 40 (medium), and 20 (narrow); and
- C) mass accuracy of the peak assessed at 4 ppm (broad), 2.5 ppm (medium), and 1.5 ppm (narrow).

Area cut-off was set to 3000 counts as default, but was often adjusted for very concentrated or dilute samples.

The software DataAnalysis (DA) from Bruker Daltonics was used for manual comparison of all extracted-ion chromatograms (EIC) generated by TargetAnalysis to the base peak chromatograms (BPC), to identify non-detected major peaks.

## Results and discussion

#### The database

The database used for screening comprised 7100 compounds, of which 1500 were available reference standards and 500 were tentatively identified compounds. The database was handled in ACD/ChemFolder, using a custom interface shown in Fig. S1, Electronic Supplementary Material. The database also contained legacy data from older HPLC–DAD [22], HPLC–DAD–TOFMS [9, 23], and pKa data [9] if available. Records from AntiBase needed proofreading, because we found that approximately 2–3 % of the structures had incorrect elemental compositions. We also estimate that approximately 5 % of structures published annually are not indexed.

Because TargetAnalysis could not extract both targeted and untargeted data and combine them, the fastest workflow was to overlay all the identified compounds from TargetAnalysis on the BPC chromatograms. All major non-identified peaks could then easily be observed visually (as shown in Fig. 1), dereplicated, and added to the database as a tentatively identified [9, 25] or unknown compound. Subsequently it was clear that the signals from compounds originating from filters, media blanks etc. were most efficiently handled by including them in the database, so that they would be annotated and

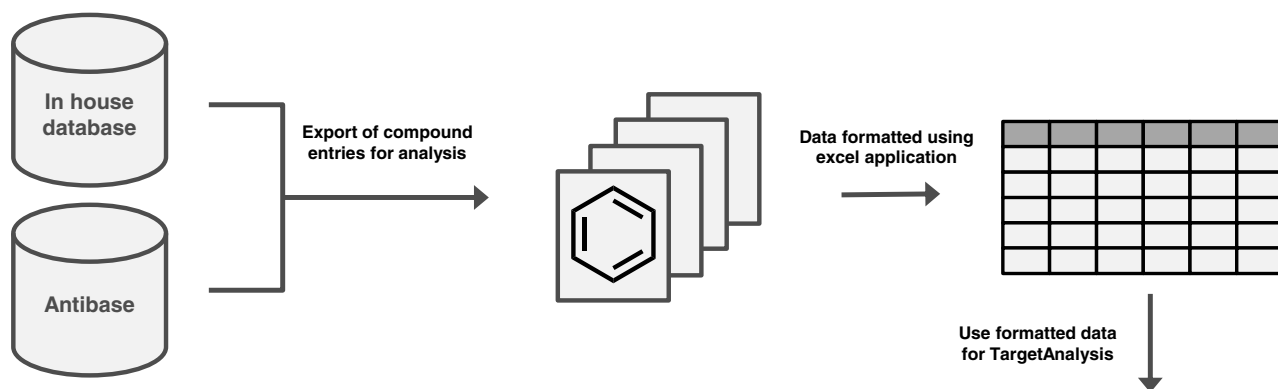
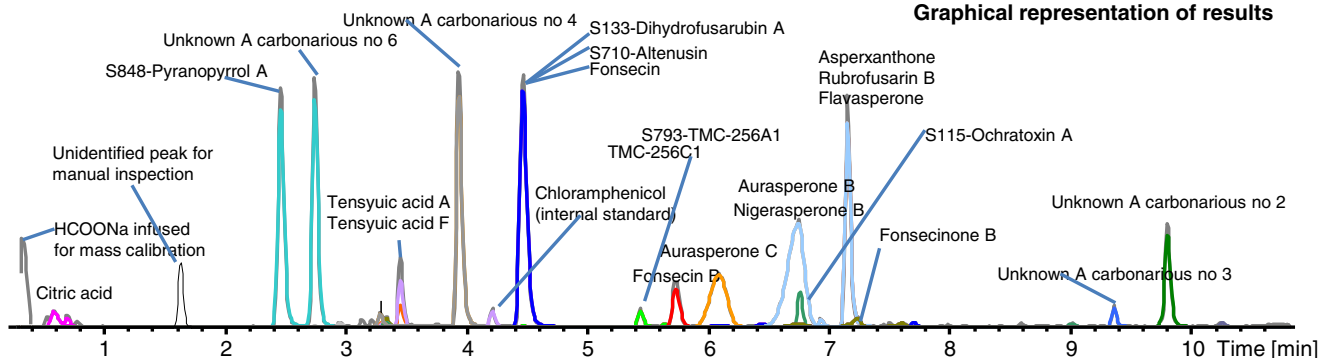


Table containing screening results

| Found | Compound Name                        | Reg.No. | Mol.Formula            | PMI                | d RT [...] | Err [...] | Err [...] | mSi... | Area    | Intens. | RT,e... | RT,m... | m/z,calc. | m/z,me... |
|-------|--------------------------------------|---------|------------------------|--------------------|------------|-----------|-----------|--------|---------|---------|---------|---------|-----------|-----------|
| +++   | Chloramphenicol IS                   | 0       | C 11 H 12 Cl 2 N 2 ... | [M+H] <sup>+</sup> | 0.069      | 0.9       | 0.3       | 3.3    | 519133  | 125123  | 4.120   | 4.051   | 323.0196  | 323.0193  |
| +++   | Pestalamide B                        | 0       | C 18 H 18 N 2 O 5      | [M+H] <sup>+</sup> | 0.136      | 0.0       | -0.0      | 12.8   | 138063  | 29033   | 4.455   | 4.319   | 343.1288  | 343.1288  |
| +     | Carbonarone B                        | 0       | C 13 H 11 N 1 O 3      | [M+H] <sup>+</sup> | -4.422     | 0.3       | -0.1      | 1.0    | 419988  | 100203  | 0.000   | 4.422   | 230.0812  | 230.0812  |
| +     | Carbonarone A                        | 0       | C 13 H 11 N 1 O 3      | [M+H] <sup>+</sup> | -4.422     | 0.3       | -0.1      | 1.0    | 419988  | 100203  | 0.000   | 4.422   | 230.0812  | 230.0812  |
| +++   | Tensidol A                           | 0       | C 13 H 11 N 1 O 3      | [M+H] <sup>+</sup> | 0.138      | 0.3       | -0.1      | 1.0    | 419988  | 100203  | 4.560   | 4.422   | 230.0812  | 230.0812  |
| +     | 7-Hydroxy-4-methoxy-5-methylcoumarin | 0       | C 11 H 10 O 4          | [M+H] <sup>+</sup> | -4.469     | 0.7       | 0.1       | 11.7   | 24707   | 6182    | 0.000   | 4.469   | 207.0652  | 207.0650  |
| +++   | Unknown in A. niger 22 MW 186        | 0       | C 12 H 10 O 2          | [M+H] <sup>+</sup> | 0.146      | 0.0       | 0.0       | 3.5    | 278120  | 66933   | 4.616   | 4.470   | 187.0754  | 187.0754  |
| +++   | Orlandin                             | 0       | C 22 H 18 O 8          | [M+H] <sup>+</sup> | 0.101      | 0.1       | -0.0      | 5.7    | 286078  | 67889   | 4.850   | 4.749   | 411.1074  | 411.1075  |
| +     | Bicumagrigin                         | 0       | C 22 H 18 O 8          | [M+H] <sup>+</sup> | -4.749     | 0.1       | -0.0      | 5.7    | 286078  | 67889   | 0.000   | 4.749   | 411.1074  | 411.1075  |
| +++   | S847-Funalenone                      | 0       | C 15 H 12 O 6          | [M+H] <sup>+</sup> | 0.046      | 0.3       | -0.1      | 3.2    | 3229445 | 206891  | 4.816   | 4.770   | 289.0707  | 289.0707  |

Graphical representation of results



**Fig. 1** Example of workflow for screening of fungal extracts, in this case an extract from *Aspergillus niger*. The database maintained at our center contains 7100 records, comprising reference standards and their associated MS and UV data. For a specific analysis it is possible to export

relevant entries from the database and, via an in-house-built Excel application, convert these to a format that can be imported into TargetAnalysis. Analysis via TargetAnalysis then yields both a graphical interpretation of the results and a table of the data

labeled by TargetAnalysis. This led to labeling peaks with the reference standard number (Fig. 1), indicating whether a compound was available as a reference standard for subsequent reanalysis.

The results from the analysis of an extract from *A. niger* are depicted in Fig. 1, illustrating the major disadvantage of the method. It can be seen that several compounds have been annotated to the same chromatographic peak, because numerous compounds in the search list had the same elemental composition and unknown RT. This is the major reason for not including, e.g., all 41,000 compounds from AntiBase2012 in the search list, because it contains up to 130 compounds with the same elemental composition [9]. For each experiment it is therefore important to use a search list from which highly unlikely compounds, for example metabolites from other organisms, are restricted. If no compounds are found, reanalysis

can be conducted using a list of all elemental compositions in the database of choice.

#### Handling adducts and in-source fragmentation

Early analytical work (results not shown), using atmospheric-pressure chemical ionization (APCI)<sup>+</sup>, APCI<sup>-</sup>, ESI<sup>+</sup> and ESI<sup>-</sup> ionization for analysis of extracts from *A. niger* and *A. nidulans*, did not reveal superior ionization by APCI over ESI for any compound. Thus APCI was not further pursued, although there must be some apolar and/or semi-volatile compounds that are better ionized by APCI.

Adduct formation on the maXis 3G ion-source was surprisingly different from that observed on our 10-years-old Waters Micromass LCT (z-spray source) [9], even though exactly the same eluents were used. In ESI<sup>+</sup> mode we

observed many compounds using the maXis, e.g. chloramphenicol and several anthraquinones, which were not previously detected by the LCT system using ESI<sup>+</sup>. It remains to be investigated whether this was caused by the grounded needle (and thus a potential of –42000 V over the source), the ion-funnel, or other changes in the source. Ammonium adducts were also far less abundant on the maXis, and formation seemed to be efficiently suppressed by the drying gas, leading to spectra with abundant [M+H]<sup>+</sup> and [M+Na]<sup>+</sup>, because most compounds with high affinity for ammonium also have a high affinity for sodium [9].

An interesting phenomenon observed with ESI<sup>+</sup> was that in the end of the gradient, when the acetonitrile content was close to 100 %, ionization seemed to favor formation of [2M+Na]<sup>+</sup> ions. For such analytes as the variecoxanthones and emericellin (Fig. S2, Electronic Supplementary Material) the [2M+Na]<sup>+</sup> ion (*m/z* 839.3766) had a 5–10-fold-higher intensity than [M+H]<sup>+</sup>. This was presumably caused by the high acetonitrile content, which would have facilitated fast evaporation, and acidic compounds may thus hold the residual Na<sup>+</sup> by ion exchange before evaporation from the droplet.

Macrocyclic trichothecenes in extracts from *Baccharis megapotamica* [28] revealed that the adduct pattern was concentration-dependent, with the highest intensity [M+Na]<sup>+</sup> occurring at low concentrations of the analyte (Fig. S3, Electronic Supplementary Material). This is probably the result of limited Na<sup>+</sup>, and thus [M+H]<sup>+</sup> is most abundant when Na<sup>+</sup> is depleted. On full-scan instruments this phenomenon can be regarded as *adduct displacement*, whereas it will be observed as ion suppression on MS–MS instruments if only one of [M+H]<sup>+</sup> or [M+Na]<sup>+</sup> is measured. For MS–MS characterization of compounds that favor sodium adducts, we have in several applications used ammonium formate as buffer to depress sodium adduct formation. In one example we also changed the sodium formate calibration solution to a polyethylene glycol mixture, and switched the glass water-solvent bottle to plastic.

Ergosterol and related sterols were, surprisingly, detected as [M+H–H<sub>2</sub>]<sup>+</sup> ions, whereas, e.g., cholesterol was detected as [M+H–H<sub>2</sub>O]<sup>+</sup>.

ESI<sup>–</sup> ionized acidic compounds (carboxylic acids, enols and phenols) well, because of easy disassociation of H<sup>+</sup>, and also proved superior to ESI<sup>+</sup> unless the target compounds also contained amine or amide functionalities. Compounds without acidic protons, that were observed as [M+HCOO]<sup>–</sup> on both Waters LCT z-spray source instrumentation [9] and an Agilent 6550 QTOF, were often not detected at all using the maXis system.

Ion-source fragmentation was unavoidable for very fragile molecules, but was mainly observed as water loss for compounds that formed sodium adducts: jumping from [M+Na]<sup>+</sup> to [M+H–H<sub>2</sub>O]<sup>+</sup>, with *m/z* 39.9925, and occasionally also to [M+H–2H<sub>2</sub>O]<sup>+</sup>, with *m/z* 58.0031. Thus the sodium adducts

could be an advantage when screening fragile compounds. Cases where [M+H]<sup>+</sup> was not observed were much more predominant on the maXis than on the Waters LCT (z-spray source). In-source fragmentation could be minimized by lowering the potential of the quadrupole and between the funnels, but could not be abolished because this would lead to >10 % loss of sensitivity. We therefore included [M+H–H<sub>2</sub>O]<sup>+</sup> and [M+H–2H<sub>2</sub>O]<sup>+</sup> in the database of compounds losing H<sub>2</sub>O during ESI<sup>+</sup> (often an alcohol group with α-carbon was available for elimination via double-bond formation) [9].

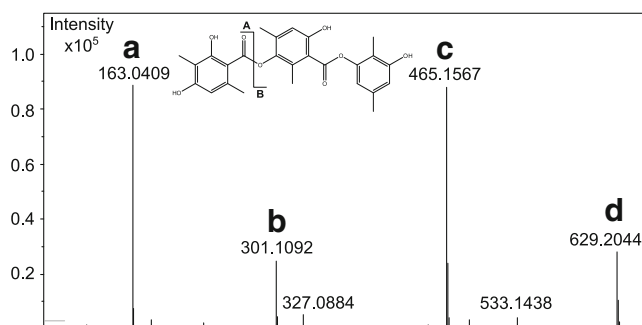
The screening process was also performed, using similar samples, on an Agilent 1290 UHPLC–6550 QTOF system, using Agilent Masshunter's *Find By Formula* option. This function could handle different adducts and simple losses, for example water loss, theoretically ensuring that no compounds were overlooked. This, however, also resulted in many more false positives, because all peaks are believed to correspond to, e.g., an [M+H–H<sub>2</sub>O]<sup>+</sup> ion, even if the peaks also fit the [M+H]<sup>+</sup> of another compound. ACD's MS Workbook Suite intelliXtract function (v. 12) was also tested. The software could assign the whole adduct, multimer and fragment pattern for a peak, but required the presence of a [M+H]<sup>+</sup> or [M–H]<sup>–</sup> ion. This software was approximately 50–100 times more time-consuming than Bruker's TargetAnalysis for a list of 3000 compounds, but does work for smaller databases [19].

Molecules with masses above 1000 Da, which include many NRPs (e.g. lipopeptides and peptaibols), all produced doubly and often also triply charged ions, thus appearing in the scan window of *m/z* 100–1000. The only two exceptions were special cyclic peptides, for example cereulide and valinomycin, which are very strong K<sup>+</sup>-ionophores and therefore only produced [M+Na]<sup>+</sup> and [M+K]<sup>+</sup> ions [29].

The adduct formation behavior of some compounds can however be hard to predict. This was observed for an extract of *Phoma levellei* [30] (incorrectly identified as *Cladosporium uredinicola*), for which the ESI<sup>–</sup> spectrum of 3-Hydroxy-2,5-dimethylphenyl 3-[(2,4-Dihydroxy-3,6-dimethylbenzoyl)oxy]-6-hydroxy-2,4-dimethylbenzoate (Fig. 2) indicated the presence of several co-eluting compounds. Deconvolution of the ions revealed that ions labeled A–D came from the same compound. Ion C corresponded to [M–H]<sup>–</sup>, A and B were fragments, and D was a composite ion of [M–H]<sup>–</sup> and one fragment-ion A.

#### Ion-cooler bias

The maXis 3G is equipped with a hexapole ion-cooler, which collects the ions, reduces their kinetic energy, and ejects them into the orthogonal accelerator in the TOF mass analyzer. Our results reveal that the ion cooler settings have a significant effect on the intensities of the ions in the measured mass range (Fig. S4, Electronic Supplementary Material).



**Fig. 2** ESI<sup>-</sup> spectrum of 3-Hydroxy-2,5-dimethylphenyl 3-[(2,4-Dihydroxy-3,6-dimethylbenzoyl)oxy]-6-hydroxy-2,4-dimethylbenzoate, showing  $M-H$ <sup>-</sup> (C) and fragment ions a and b. d is a composite of ions a and c

Three variables were important:

1. the *ion-cooler radio frequency* (RF), which sets the voltage for the ion-cooler;
2. the transfer time, which is the time window wherein ions are transmitted into the TOF; and
3. the *pre-pulse storage time*, which will apply a low mass limit and is a delay between the transfer time and the TOF pulser. Higher values favored the transfer of higher  $m/z$  ions, but also discriminated low  $m/z$  ions.

Figure S4 (Electronic Supplementary Material) shows selected results from analysis using seven different transfer times. The results revealed that the ion-cooler “window” for low mass compounds is narrow, and the settings used to obtain an optimum signal for lower  $m/z$  ions resulted in low intensities of higher  $m/z$  ions, and vice versa. For analytes with  $m/z$  lower than 100 (data not shown), the optimum settings excessively discriminated the signal intensity of higher  $m/z$  values. At an ion cooler RF value of 30 V<sub>pp</sub>, the signal of  $m/z$  91 was highly suppressed at all transfer times.

Our in-house database contained 7100 compounds with a  $[M+H]^+$  in the range  $m/z$  100–1000. Of these, 14 % will have a  $[M+H]^+ < 226$   $m/z$  and will reach only 30 % of their maximum intensity using standard screening settings. For ions smaller than  $m/z$  130 the signal suppression will be extensive, but luckily less than 1 % of the compounds in our in-house database and AntiBase have masses this low [9]. If a target compound was in the mass range below  $m/z$  130, the optimum ion-cooler settings resulted in an intensity of less than 10 % for compounds with an  $m/z > 226$ , and of only 5 % of the signal from compounds with an  $m/z > 600$ . It is important to be aware of this signal discrimination in some mass ranges under different ion-cooler settings.

Effect of detector overload on isotope pattern and mass accuracy

Because fungal extracts contain many different compounds with varying concentrations and ionization efficiencies,

screening of extracts routinely resulted in analysis of compounds with intensities higher than  $2-3 \times 10^6$  counts, which overloaded the detector of the maXis QTOF (this problem was much more severe on older TOF instruments [9]). This caused an  $m/z$  shift to higher values, which in the worst case resulted in an increase of up to 3–4 ppm. This also led to a distorted isotopic pattern, where the A+1, A+2 isotopomers were too intense relative to the A isotopomer. To avoid false negative results in TargetAnalysis, it was thus crucial to set a wide range (5 ppm) on the isotope fit and mass accuracy. However, these high-intensity peaks could be easily spotted by the peak height in the results table, after which data for the chromatographic peak could be examined from scans where the detector was not overloaded. The isotope fit was highly dependent on a weekly detector tuning, and the medium and narrow-range settings had to be increased twofold when the detector had not been tuned within the week.

Aggressive dereplication reveals new metabolites from highly toxic spoilage fungus *Aspergillus carbonarius*

*A. carbonarius* is a physiologically very well investigated species because of its contamination of grapes, and the subsequent contamination of wine and raisins, with ochratoxin A [31]. However, other compounds from the fungus have attracted little attention. As well as this toxin, it is capable of producing carbonarones and pestalamide A (former tensidol B) [32], pyranonigrins, carbonarins, organic acids, and aurasperones [26].

Extracts from *A. carbonarius* cultivated on YES agar were screened for 3000 compounds:

1. compounds from *Aspergillus* (with an emphasis on *Aspergillus* section Nigri compounds) and *Penicillium*;
2. all standards available in our collection; and
3. all unidentified peaks registered in our database.

With a high area cut-off of 10,000 counts, 66 peaks were integrated (Table 1); however, 16 of these compounds were from peaks assigned to several compounds (up to five) and thus only 45 true peaks were annotated. The major peaks in the sample are displayed in Fig. 3.

Citric acid was detected as the sodium adduct and as two peaks because of poor retention on the column, which occurred because the LC–MS method is not well suited to such polar compounds. Kojic acid was incorrectly identified as another compound with the same elemental composition, because neither the RT nor the characteristic UV spectrum matched a reference standard.

Three interesting nitrogen-containing biomarkers for this species, with elemental compositions  $C_{11}H_{11}NO_5$  and

**Table 1** Results from the aggressive dereplication of an extract of *Aspergillus carbonarius* grown on YES agar

| Peak | Class | Comment                            | Compound name              | Molecular formula   | Err (ppm) | mSigma | Area (arbitrary units) | RT measured (min) | RT expected (min) |
|------|-------|------------------------------------|----------------------------|---|-----------|--------|------------------------|-------------------|-------------------|
| A    | +++   | OK double peak caused by injection | Citric acid                | C <sub>6</sub> H <sub>7</sub> NaO <sub>7</sub>                                | 0.1       | 8      | 351577                 | 0.609             | 0.61              |
| B    | +++   | OK double peak caused by injection | Citric acid                | C <sub>6</sub> H <sub>7</sub> NaO <sub>7</sub>                                | 0.1       | 3      | 256614                 | 0.719             | 0.72              |
| C    | +++   |                                    | BL-UK Cla no 60 pos. blank | C <sub>10</sub> H <sub>13</sub> N <sub>5</sub> O <sub>4</sub>                 | 0.9       | 7      | 22958                  | 0.722             | 0.72              |
| D    | +     | Wrong, UV and RT do not fit        | S96-Kojic acid             | C <sub>6</sub> H <sub>6</sub> O <sub>4</sub>                                  | 0.9       | 9      | 14965                  | 0.791             | 1.2               |
| E    | +++   |                                    | BL-UK Cla no 72 pos. blank | C <sub>10</sub> H <sub>16</sub> N <sub>2</sub> O <sub>2</sub>                 | 0.2       | 11     | 15379                  | 1.807             | 1.75              |
| F    | +++   |                                    | BL-UK Cla no 95 pos. blank | C <sub>7</sub> H <sub>14</sub> N <sub>2</sub> O <sub>3</sub>                  | 1.2       | 6      | 15141                  | 2.243             | 2.1               |
| G    | +++   | OK                                 | S848-Pyranonigrin A        | C <sub>10</sub> H <sub>9</sub> N <sub>1</sub> O <sub>5</sub>                  | 0.9       | 19     | 5428853                | 2.475             | 2.36              |
| H    | +++   |                                    | UK in A. ni 2              | C <sub>10</sub> H <sub>9</sub> N <sub>1</sub> O <sub>4</sub>                  | 0.4       | 17     | 24641                  | 2.756             | 2.906             |
| I    | +++   | Interesting new biomarker          | UK A car no 6              | C <sub>11</sub> H <sub>11</sub> N <sub>1</sub> O <sub>5</sub>                 | 0.6       | 17     | 5203919                | 2.756             | 2.751             |
| J    | +++   |                                    | UK in A. ni 19             | C <sub>18</sub> H <sub>37</sub> NaO <sub>10</sub>                             | 0.2       | 10     | 13945                  | 2.892             | 2.844             |
| K    | +++   |                                    | BL-UK Cla no 11 pos. blank | C <sub>11</sub> H <sub>18</sub> N <sub>2</sub> O <sub>2</sub>                 | 1.3       | 10     | 29484                  | 2.912             | 3.09              |
| L    | +++   |                                    | UK in A. ni 2              | C <sub>10</sub> H <sub>9</sub> N <sub>1</sub> O <sub>4</sub>                  | 1.2       | 1      | 90082                  | 2.962             | 2.906             |
| M    | +++   |                                    | BL-UK Cla no 12 pos. blank | C <sub>11</sub> H <sub>18</sub> N <sub>2</sub> O <sub>2</sub>                 | 0.2       | 5      | 44764                  | 3.14              | 3.09              |
| N    | +++   | Interesting new biomarker          | UK A car no 4              | C <sub>18</sub> H <sub>21</sub> N <sub>1</sub> O <sub>2</sub>                 | 0.1       | 16     | 350827                 | 3.295             | 3.288             |
| O    | +++   |                                    | UK in A. ni 16             | C <sub>22</sub> H <sub>45</sub> NaO <sub>12</sub>                             | 0.6       | 18     | 13611                  | 3.299             | 3.25              |
| P    | +     | No confused by the A isomer        | Tensyucic acid A           | C <sub>11</sub> H <sub>16</sub> O <sub>6</sub>                                | 0.2       | 7      | 96858                  | 3.344             | 0                 |
| P    | +     | Presumably OK                      | Tensyucic acid F           | C <sub>11</sub> H <sub>16</sub> O <sub>6</sub>                                | 0.2       | 7      | 96858                  | 3.344             | 0                 |
| Q    | ++    |                                    | UK A car no 4              | C <sub>18</sub> H <sub>21</sub> N <sub>1</sub> O <sub>2</sub>                 | 0.1       | 15     | 48785                  | 3.592             | 3.288             |
| Q    | ++    |                                    | UK A car no 1              | C <sub>18</sub> H <sub>21</sub> N <sub>1</sub> O <sub>2</sub>                 | 0.1       | 15     | 48785                  | 3.592             | 3.923             |
| R    | +++   |                                    | UK in A. ni 5              | C <sub>21</sub> H <sub>44</sub> O <sub>11</sub>                               | 0.3       | 14     | 10039                  | 3.63              | 3.581             |
| S    | +     | OK but may be the C isomer         | Pyranonigrin B             | C <sub>11</sub> H <sub>11</sub> N <sub>1</sub> O <sub>6</sub>                 | 0.5       | 9      | 55596                  | 3.76              | 0                 |
| S    | +     | OK but may be the B isomer         | Pyranonigrin C             | C <sub>11</sub> H <sub>11</sub> N <sub>1</sub> O <sub>6</sub>                 | 0.5       | 9      | 55596                  | 3.76              | 0                 |
| T    | +++   |                                    | UK in A. ni 7              | C <sub>23</sub> H <sub>47</sub> NaO <sub>12</sub>                             | 0.4       | 37     | 17040                  | 3.767             | 3.72              |
| U    | ++    |                                    | UK A car no 4              | C <sub>18</sub> H <sub>21</sub> N <sub>1</sub> O <sub>2</sub>                 | 0.7       | 15     | 5265217                | 3.944             | 3.288             |
| U    | +++   |                                    | UK A car no 1              | C <sub>18</sub> H <sub>21</sub> N <sub>1</sub> O <sub>2</sub>                 | 0.7       | 15     | 5265217                | 3.944             | 3.923             |
| V    | +     |                                    | Pyranonigrin D             | C <sub>11</sub> H <sub>9</sub> N <sub>1</sub> O <sub>5</sub>                  | 0.2       | 9      | 17070                  | 3.946             | 0                 |
| W    | +++   | Internal standard                  | Chloramphenicol IS         | C <sub>11</sub> H <sub>12</sub> Cl <sub>2</sub> N <sub>2</sub> O <sub>5</sub> | 0.2       | 31     | 326301                 | 4.219             | 4.12              |
| X    | +++   | No confused by Fonsecin            | S133-Dihydrofusarubin A    | C <sub>15</sub> H <sub>14</sub> O <sub>6</sub>                                | 1.1       | 25     | 6829770                | 4.47              | 4.75              |
| X    | ++    | Wrong, UV and RT do not fit        | S710-Altenucin             | C <sub>15</sub> H <sub>14</sub> O <sub>6</sub>                                | 1.1       | 25     | 6829770                | 4.47              | 4.908             |
| X    | +++   | OK                                 | Fonsecin                   | C <sub>15</sub> H <sub>14</sub> O <sub>6</sub>                                | 1.1       | 25     | 6829770                | 4.47              | 4.45              |
| Y    | +     | OK but one must be a new isomer    | Tensyucic acid B           | C <sub>12</sub> H <sub>18</sub> O <sub>6</sub>                                | 1.1       | 24     | 21361                  | 4.554             | 0                 |
| Z    | +     | OK but one must be a new isomer    | Tensyucic acid B           | C <sub>12</sub> H <sub>18</sub> O <sub>6</sub>                                | 1         | 22     | 10189                  | 4.681             | 0                 |
| AA   | +++   | OK                                 | S133-Dihydrofusarubin A    | C <sub>15</sub> H <sub>14</sub> O <sub>6</sub>                                | 1         | 46     | 10340                  | 5.031             | 4.75              |
| AA   | +++   | Wrong, UV and RT do not fit        | S710-Altenucin             | C <sub>15</sub> H <sub>14</sub> O <sub>6</sub>                                | 1         | 46     | 10340                  | 5.031             | 4.908             |
| AB   | ++    | No confused by Dihydrofusarubin A  | Fonsecin                   | C <sub>15</sub> H <sub>14</sub> O <sub>6</sub>                                | 1         | 46     | 10340                  | 5.031             | 4.45              |
| AC   | ++    |                                    | Aurasperone C              | C <sub>31</sub> H <sub>28</sub> O <sub>12</sub>                               | 0.5       | 37     | 15414                  | 5.249             | 5.94              |
| AD   | +++   | No confused by TMC-256A1           | TMC-256C1                  | C <sub>15</sub> H <sub>12</sub> O <sub>5</sub>                                | 0.6       | 18     | 349791                 | 5.437             | 5.67              |
| AD   | +++   | OK                                 | S793-TMC-256A1             | C <sub>15</sub> H <sub>12</sub> O <sub>5</sub>                                | 0.6       | 18     | 349791                 | 5.437             | 5.37              |
| AE   | ++    |                                    | Aurasperone C              | C <sub>31</sub> H <sub>28</sub> O <sub>12</sub>                               | 0.4       | 41     | 19423                  | 5.494             | 5.94              |
| AF   | +++   | OK                                 | TMC-256C1                  | C <sub>15</sub> H <sub>12</sub> O <sub>5</sub>                                | 0.3       | 7      | 65429                  | 5.641             | 5.67              |
| AF   | +++   | No confused by TMC-256C1           | S793-TMC-256A1             | C <sub>15</sub> H <sub>12</sub> O <sub>5</sub>                                | 0.3       | 7      | 65429                  | 5.641             | 5.37              |

**Table 1** (continued)

| Peak | Class | Comment                          | Compound name        | Molecular formula   | Err (ppm) | mSigma | Area (arbitrary units) | RT measured (min) | RT expected (min) |
|------|-------|----------------------------------|----------------------|---|-----------|--------|------------------------|-------------------|-------------------|
| AG   | +++   |                                  | Fonsecin B           | C <sub>16</sub> H <sub>16</sub> O <sub>6</sub>                                | 0.8       | 30     | 1055089                | 5.729             | 5.66              |
| AH   | +     | Wrong water-loss ion of C isomer | Niasperone C         | C <sub>31</sub> H <sub>26</sub> O <sub>11</sub>                               | 1         | 9      | 76397                  | 6.08              | 0                 |
| AH   | +++   | Wrong water-loss ion of C isomer | Aurasperone F        | C <sub>31</sub> H <sub>26</sub> O <sub>11</sub>                               | 1         | 9      | 76397                  | 6.08              | 6.303             |
| AH   | +++   |                                  | Aurasperone C        | C <sub>31</sub> H <sub>28</sub> O <sub>12</sub>                               | 1.1       | 23     | 3247597                | 6.081             | 5.94              |
| AI   | ++    |                                  | UK in A. ni 23       | C <sub>15</sub> H <sub>33</sub> N <sub>17</sub> O <sub>6</sub>                | 0.2       | 62     | 39935                  | 6.344             | 6.23              |
| AJ   | ++    |                                  | UK in A. ni 20       | C <sub>28</sub> H <sub>36</sub> N <sub>4</sub> O <sub>5</sub>                 | 0.9       | 25     | 49747                  | 6.397             | 6.043             |
| AK   | +     | OK but may be a different isomer | Niasperone C         | C <sub>31</sub> H <sub>26</sub> O <sub>11</sub>                               | 0.8       | 11     | 115620                 | 6.434             | 0                 |
| AK   | +++   | OK but may be a different isomer | Aurasperone F        | C <sub>31</sub> H <sub>26</sub> O <sub>11</sub>                               | 0.8       | 11     | 115620                 | 6.434             | 6.303             |
| AL   | +++   | Wrong water-loss ion of B isomer | Aurasperone E        | C <sub>32</sub> H <sub>28</sub> O <sub>11</sub>                               | 0.9       | 23     | 186091                 | 6.728             | 6.62              |
| AL   | ++    | Wrong water loss ion of B isomer | Aurasperone E-isomer | C <sub>32</sub> H <sub>28</sub> O <sub>11</sub>                               | 0.9       | 23     | 186091                 | 6.728             | 7.104             |
| AL   | ++    | Wrong water loss ion of B isomer | Fonsecinone B        | C <sub>32</sub> H <sub>28</sub> O <sub>11</sub>                               | 0.9       | 23     | 186091                 | 6.728             | 7.472             |
| AL   | +     | OK but may be a different isomer | Niasperone B         | C <sub>32</sub> H <sub>30</sub> O <sub>12</sub>                               | 1.3       | 22     | 6659679                | 6.728             | 0                 |
| AL   | +++   | OK but may be a different isomer | Aurasperone B        | C <sub>32</sub> H <sub>30</sub> O <sub>12</sub>                               | 1.3       | 22     | 6659679                | 6.728             | 6.605             |
| AM   | +++   | OK                               | S115-Ochratoxin A    | C <sub>20</sub> H <sub>18</sub> Cl <sub>1</sub> N <sub>1</sub> O <sub>6</sub> | 0.7       | 50     | 693721                 | 6.75              | 6.62              |
| AN   | +     | OK but may be a different isomer | Niasperone C         | C <sub>31</sub> H <sub>26</sub> O <sub>11</sub>                               | 1.5       | 9      | 62334                  | 6.779             | 0                 |
| AN   | ++    | OK but may be a different isomer | Aurasperone F        | C <sub>31</sub> H <sub>26</sub> O <sub>11</sub>                               | 1.5       | 9      | 62334                  | 6.779             | 6.303             |
| AO   | ++    | No rubrofusarin                  | Flavasperone         | C <sub>16</sub> H <sub>14</sub> O <sub>5</sub>                                | 0.7       | 20     | 146028                 | 6.923             | 7.2               |
| AO   | +++   | OK                               | Rubrofusarin B       | C <sub>16</sub> H <sub>14</sub> O <sub>5</sub>                                | 0.7       | 20     | 146028                 | 6.923             | 7.029             |
| AP   | +++   | OK                               | Flavasperone         | C <sub>16</sub> H <sub>14</sub> O <sub>5</sub>                                | 0.6       | 14     | 4285585                | 7.145             | 7.2               |
| AP   | ++    | No flavasperone                  | Rubrofusarin B       | C <sub>16</sub> H <sub>14</sub> O <sub>5</sub>                                | 0.6       | 14     | 4285585                | 7.145             | 7.029             |
| AQ   | ++    | OK but may be a different isomer | Aurasperone E        | C <sub>32</sub> H <sub>28</sub> O <sub>11</sub>                               | 0.2       | 35     | 300587                 | 7.221             | 6.62              |
| AQ   | +++   | OK but may be a different isomer | Aurasperone E-isomer | C <sub>32</sub> H <sub>28</sub> O <sub>11</sub>                               | 0.2       | 35     | 300587                 | 7.221             | 7.104             |
| AQ   | +++   | OK but may be a different isomer | Fonsecinone B        | C <sub>32</sub> H <sub>28</sub> O <sub>11</sub>                               | 0.2       | 35     | 300587                 | 7.221             | 7.472             |
| AR   | +++   | OK but may be a different isomer | Fonsecinone B        | C <sub>32</sub> H <sub>28</sub> O <sub>11</sub>                               | 0.7       | 15     | 156648                 | 7.588             | 7.472             |
| AS   | +++   |                                  | S598-Linoleic acid   | C <sub>18</sub> H <sub>32</sub> O <sub>2</sub>                                | 0.6       | 11     | 104992                 | 10.23             | 10.17             |

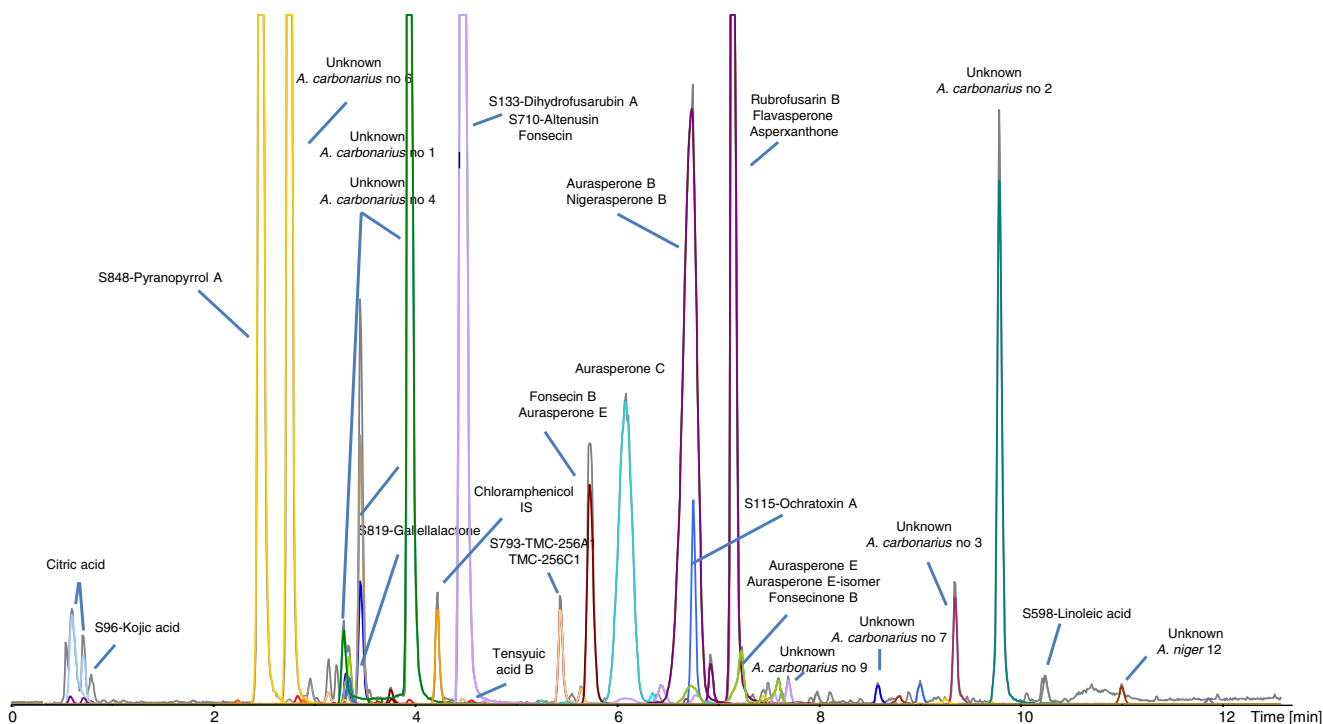
mSigma, fit of isotope pattern (see text for more details); RT, retention time

C<sub>18</sub>H<sub>21</sub>NO<sub>2</sub> (two isomers), were detected (unknown 1, 4, and 6), and these were not detected for other black *Aspergilli* (results not shown). Ochratoxin A, which was produced in very high amounts, is an interesting case because its precursors, ochratoxin  $\alpha$  and B, were not detected even in trace amounts, indicating that the biosynthetic enzymes are very efficient.

Several closely eluting same-elemental-composition groups were observed and needed manual verification. For example, the rationale for identifying peak AA, as seen in Table 1, was:

1. Altenusin C<sub>15</sub>H<sub>14</sub>O<sub>6</sub> was from *Alternaria* and thus taxonomically unlikely. RT was within the limits where a reference standard should be co-analyzed in the sequence for verification. Inspection of the UV–Vis data led to easy elimination, and so did the presence of a perfectly co-eluting [M+Na]<sup>+</sup> ion with M=C<sub>15</sub>H<sub>16</sub>O<sub>7</sub>.
2. Fonsecin could be eliminated by the same arguments.
3. Finally, dihydrofusarubin A was identified as the correct compound, on the basis of its perfectly matching UV–Vis spectrum and its [M+H–H<sub>2</sub>O]<sup>+</sup> and [M+Na]<sup>+</sup> ions. However, dihydrofusarubin A was only detected because





**Fig. 3** Analyzed fungal extract from *A. carbonarius* cultivated on YES media. The chromatogram is overlaid with EIC from detected compounds, facilitating easy dereplication. The chromatogram has been scaled to better illustrate the smaller peaks

it was registered in the database in the form  $[M+H-H_2O]^+$ .

The AL peak (Table 1) must be niasperone B or aurasperone B, but could not be differentiated without a reference standard. In that case, water-loss ions led to the peak being wrongly assigned to aurasperone E and one of its isomers, and to fonsecinone B.

The pair flavasperone and rubrofusarin B should both be produced when the dimeric naphtho- $\gamma$ -pyrones are produced, and a log D calculation revealed that rubrofusarin B should elute first.

Differentiating the tensyucic acids was more ambiguous, because the reported elution pattern from reversed phase is F, A, B, C, D, and E [33], with F and B having the same elemental composition, and A and B almost co-eluting. Manual inspection of the screening results was therefore necessary to attempt to distinguish between the isomers. This revealed that the first-eluting tensyucic acid was most probably the F isomer (1.3 min to the B isomer). However, the B isomer could not be unambiguously assigned as one of the two peaks Y or Z, because only one compound with  $C_{12}H_{18}O_6$  is described.

In conclusion, the method very quickly identified suspected compounds from *A. carbonarius*. Besides this, a novel group of nitrogen-containing compounds, and tensyucic acids and numerous other compounds from related species,

were detected. This indicated that, from a toxicological perspective, more compounds needed to be considered. A problem is that many of the closely related niasperones, aurasperones, and fonsecinones have identical elemental compositions and UV–Vis spectra and are very difficult to differentiate. To enable differentiation, we are currently considering an MS–HRMS library approach, as done for a toxic substance library [17]. However, TargetAnalysis does not presently have the capability to handle MS–HRMS data or pseudo-MS–MS data including MS-E, MS-All and/or All-Ions [21]. A further example of aggressive dereplication applied to *Penicillium melanoconidium* can be found in Electronic Supplementary Material Section “Materials and methods” and Tables S1 and S2. Here, several families of compounds not previously seen in the species were detected (Fig. S5, Electronic Supplementary Material). This included the highly toxic verrucosidins, and a presumed novel dideoxyverrucosidin. Chrysogine, a compound often detected in cereal-infecting Fusaria, was also detected, indicating that this may be an important virulence factor. The example shows how the aggressive dereplication procedure was used to detect known compounds not previously detected from the fungus. The results illustrate that all major peaks in the chromatogram were overlaid with an EIC, proving the effectiveness of the procedure and also indicating that it is a chemically very well characterized species.

## Conclusion

Screening fungal secondary metabolites on the basis of elemental composition and lists restricted to the same genus and related fungi was proved to be an efficient way to quickly investigate fungal extracts. By overlaying detected peaks and BPC chromatograms, the approach gives a visual overview of a sample and indicates whether it is a previously uninvestigated species by establishing how many peaks are unlabeled. This approach can also be used on other vendor instrumentations using analogous software packages, for example: TargetLynx (Waters), TraceFinder (Thermo), MassHunter Find By Formula (Agilent), and ACD intelliXtract (Advanced Chemical Developments).

Labeling of co-identified biosynthetic related compounds could also be directly identified from the peak, making it possible to quickly assess the elution order of such compounds.

However, adduct formation and simple fragmentations are still important challenges to address when working with analytes that do not only form  $[M+H]^+$  or  $[M-H]^-$ . Using a database approach and learning from the spectrometric behavior of reference standards can minimize problems with false-negative results. More efficient adduct-analysis software will further improve this setup [9, 21].

A further improvement to be introduced is use of MS–MS [17, 19, 34] and/or pseudo-MS–MS (MS-All, MS-E, All Ions) [21] to obtain compound-specific fragment ions for confirmation of reference standards, reducing the need to run many thousands of reference standards on a daily basis. The addition of qualifier and/or fragment ions from libraries and literature data will help to minimize the number of wrongly annotated ions with the same elemental composition, which is the main disadvantage of this method.

**Acknowledgements** This work was supported by the Danish Research Agency for Technology and Production (grant 09-064967), and the EEC project MycoRed (KBBE-2007-222690-2). Dr Sven Meyer and Dr Verena Tellström from Bruker Daltonics are acknowledged for fruitful discussions and help on scripting and setting up TargetAnalysis.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

## References

- Zengler K, Paradkar A, Keller M (2009) in: Zhang L and Demain AL (Eds.) Natural Products: Drug Discovery and Therapeutic Medicine, Humana Press Inc., Totowa.
- Butler MS (2004) The Role of Natural Product Chemistry in Drug Discovery. *J Nat Prod* 67:2141–2153
- Miller JD (2008) Mycotoxins in small grains and maize: Old problems, new challenges. *Food Addit Contam* 25:219–230
- Shephard GS (2008) Impact of mycotoxins on human health in developing countries. *Food Addit Contam* 25:146–151
- Bitzer J, Kopcke B, Stadler M, Heilwig V, Ju YM, Seip S, Henkel T (2007) Accelerated dereplication of natural products, supported by reference libraries. *Chimia* 61:332–338
- Bobzin SC, Yang S, Kasten TP (2000) LC-NMR: A new tool to expedite the dereplication and identification of natural products. *J Ind Microbiol Biotechnol* 25:342–345
- Cordell GA, Shin YG (1999) Finding the needle in the haystack. The dereplication of natural products extracts. *Pure Appl Chem* 71:1089–1094
- Zhang L (2005) in: Zhang L and Demain AL (Eds.) Natural Products: Drug Discovery and Therapeutic Medicine, Humana Press Inc., Totowa.
- Nielsen KF, Månsson M, Rank C, Frisvad JC, Larsen TO (2011) Dereplication of microbial natural products by LC-DAD-TOFMS. *J Nat Prod* 74:2338–2348
- Bueschl C, Kluger B, Berthiller F, Lirk G, Winkler S, Krska R, Schuhmacher R (2012) MetExtract: A new software tool for the automated comprehensive extraction of metabolite-derived LC/MS signals in metabolomics, research. *Bioinformatics* 28:736–738
- Sleno L (2012) The use of mass defect in modern mass spectrometry. *J Mass Spectrom* 47:226–236
- Kind T, Fiehn O (2006) Metabolomic database annotations via query of elemental compositions: Mass accuracy is insufficient even at less than 1 ppm. *BMC Bioinforma* 7:234
- Erve JC, Gu M, Wang Y, DeMaio W, Talaat RE (2009) Spectral Accuracy of Molecular Ions in an LTQ/Orbitrap Mass Spectrometer and Implications for Elemental Composition Determination. *J Am Mass Spectr* 20:2058–2069
- Hansen ME, Smedsgaard J, Larsen TO (2005) X-Hitting: An Algorithm for Novelty Detection and Dereplication by UV Spectra of Complex Mixtures of Natural Products. *Anal Chem* 77:6805–6817
- Larsen TO, Petersen BO, Duus JO, Sørensen D, Frisvad JC, Hansen ME (2005) Discovery of New Natural Products by Application of X-hitting, a Novel Algorithm for Automated Comparison of Full UV Spectra, Combined with Structural Determination by NMR Spectroscopy. *J Nat Prod* 68:871–874
- Fredenhagen A, Derrien C, Gassmann E (2005) An MS/MS Library on an Ion-Trap Instrument for Efficient Dereplication of Natural Products. Different Fragmentation Patterns for  $[M+H]^+$  and  $[M+Na]^+$  Ions. *J Nat Prod* 68:385–391
- Broecker S, Herre S, Wust B, Zweigenbaum J, Pragst F (2011) Development and practical application of a library of CID accurate mass spectra of more than 2,500 toxic compounds for systematic toxicological analysis by LC-QTOF-MS with data-dependent acquisition. *Anal Bioanal Chem* 400:101–117
- Bijlsma L, Sancho JV, Hernandez F, Niessen WMA (2011) Fragmentation pathways of drugs of abuse and their metabolites based on QTOF MS/MS and MSE accurate-mass spectra. *J Mass Spectrom* 46:865–875
- El-Elimat T, Figueroa M, Ehrmann BM, Cech NB, Pearce CJ, Oberlies NH (2013) High-Resolution MS, MS/MS, and UV Database of Fungal Secondary Metabolites as a Dereplication Protocol for Bioactive Natural Products. *J Nat Prod* 76:1709–1716
- Meyer S, Ketterlinus R (2011) Confirming Multi-Target Screening Full Scan Workflows of Pesticides in Food. *Lc Gc Europe* S1:11
- Ojanpera S, Pelander A, Pelzing M, Krebs I, Vuori E, Ojanpera I (2006) Isotopic pattern and accurate mass determination in urine drug screening by liquid chromatography/time-of-flight mass spectrometry. *Rapid Commun mass sp* 20:1161–1167
- Frisvad JC, Thrane U (1987) Standardised High-Performance Liquid Chromatography of 182 mycotoxins and other fungal metabolites based on alkylphenone retention indices and UV-VIS spectra (Diode Array Detection). *J Chromatogr* 404:195–214

23. Nielsen KF, Smedsgaard J (2003) Fungal metabolite screening: database of 474 mycotoxins and fungal metabolites for de-replication by standardised liquid chromatography-UV-mass spectrometry methodology. *J Chromatogr A* 1002:111–136
24. Samson RA, Houbraken J, Thrane U, Frisvad JC, Andersen B (2010) *Food and Indoor Fungi*. CBS Laboratory Manual Series 2, CBS, Utrecht.
25. Månsson M, Phipps RK, Gram L, Munro MH, Larsen TO, Nielsen KF (2010) Explorative Solid-Phase Extraction (E-SPE) for Accelerated Microbial Natural Product Discovery, Dereplication, and Purification. *J Nat Prod* 73:1126–1132
26. Nielsen KF, Mogensen JM, Johansen M, Larsen TO, Frisvad JC (2009) Review of secondary metabolites and mycotoxins from the *Aspergillus niger* group. *Anal Bioanal Chem* 395:1225–1242
27. Frisvad JC, Rank C, Nielsen KF, Larsen TO (2009) Metabolomics of *Aspergillus fumigatus*. *Med Mycol* 47:S53–S71
28. Oliveira-Filho JC, Carmo PMS, Iversen A, Nielsen KF, Barros CLS (2012) Experimental poisoning by *Baccharis megapotamica* var. *weirii* in buffalo. *Pesquisa vet Brasil* 32:383–390
29. Thorsen L, Paulin A, Hansen BM, Rønsbo MH, Nielsen KF, Hounhouigan DJ, Jacobsen M (2011) Formation of cereulide and enterotoxins by *Bacillus cereus* in fermented African locust beans. *Food Microbiol* 28:1441–1447
30. de Medeiros LS, Murgu M, de Souza AQL, Rodrigues-Fo E (2011) Antimicrobial Depsides Produced by *Cladosporium uredinicola*, an Endophytic Fungus Isolated from *Psidium guajava* Fruits. *Helv Chim Acta* 94:1077–1084
31. Abarca ML, Accensi F, Bragulat MR, Castella G, Cabanes FJ (2003) *Aspergillus carbonarius* as the main source of ochratoxin A contamination in dried vine fruits from the Spanish market. *J Food Prot* 66: 504–506
32. Henrikson JC, Ellis TK, King JB, Cichewicz RH (2011) Reappraising the Structures and Distribution of Metabolites from Black *Aspergilli* Containing Uncommon 2-Benzyl-4H-pyran-4-one and 2-Benzylpyridin-4(1H)-one Systems. *J Nat Prod* 74: 1959–1964
33. Hasegawa Y, Fukuda T, Hagimori K, Tomoda H, Omura S (2007) Tensyuic acids, new antibiotics produced by *Aspergillus niger* FKI-2342. *Chem Pharm Bull* 55:1338–1341
34. Guthals A, Watrous JD, Dorrestein PC, Bandeira N (2012) The spectral networks paradigm in high throughput mass spectrometry. *Mol Biosyst* 8:2535–2544