REVIEW

NPR
www.rsc.org/npr

Phenotypic taxonomy and metabolite profiling in microbial drug discovery

Thomas O. Larsen, Jørn Smedsgaard, Kristian F. Nielsen, Michael E. Hansen and Jens C. Frisvad

Center for Microbial Biotechnology, BioCentrum, Technical University of Denmark, DK-2800, Kgs. Lyngby, Denmark

Received (in Cambridge, UK) 30th September 2005
First published as an Advance Article on the web 7th November 2005

Covering: up to October 2005

Microorganisms and in particular actinomycetes and microfungi are known to produce a vast number of bioactive secondary metabolites. For industrially important fungal genera such as *Penicillium* and *Aspergillus* the production of these compounds has been demonstrated to be very consistent at the species level. This means that direct metabolite profiling techniques such as direct injection mass spectrometry or NMR can easily be used for chemotyping/metabolomics of strains from both culture collections and natural samples using modern informatics tools. In this review we discuss chemotyping/metabolomics as part of intelligent screening and highlight how it can be used for identification and classification of filamentous fungi and for the discovery of novel compounds when used in combination with modern methods for dereplication. In our opinion such approaches will be important for future effective drug discovery strategies, especially for dereplication of culture collections in order to avoid redundancy in the selection of species. This will maximize the chemical diversity of the microbial natural product libraries that can be generated from fungal collections.

Thomas O. Larsen was born in Slagelse, Denmark in 1963. He studied chemical engineering at the Technical University of Denmark (DTU) with focus on organic chemistry and microbiology. After a job at Danisco A/S working on aroma synthesis he returned to DTU to study fungal volatile production, supervised by Professor Jens C. Frisvad and received his PhD in 1994. After this he joined the Marine Chemistry Group at the University of Copenhagen, headed by Reader Carsten Christophersen where he was trained as a classical natural product chemist. In 2001 he stayed for seven months with Dr Nigel Perry, Crop & Food Research, University of Otago, Dunedin, New Zealand. Currently he holds a position as associate professor at BioCentrum-DTU, at the Center for Microbial Biotechnology (CMB).

Jørn Smedsgaard was born just outside Copenhagen, Denmark, in 1959. He studied at the DTU focusing on analytical chemistry. After a master's project focusing on enzyme kinetics studied by NMR in 1986, JS took a detour around a commercial environmental analytical lab, before he set out to establish a mass spectrometry facility at the Agricultural University as a research assistance. In 1991 he returned to DTU to set-up a mass spectrometric facility at the Department of Biotechnology (now BioCentrum) and at the same time given the opportunity to do a PhD study supervised by Professor Jens Frisvad. JS received a PhD degree in 1996 for a chemotaxonomic study of Penicillium *species by direct infusion mass spectrometry. Currently JS holds a position as associate professor at BioCentrum-DTU and is acting manager for the analytic metabolite profiling facility.*

Thomas O. Larsen

Jørn Smedsgaard

DOI:10.1039/b404943h

Kristian F. Nielsen was born just outside Copenhagen, Denmark in 1972. He studied chemical and biotech engineering at the DTU with focus on analytical chemistry and microbiology. In collaboration with the Danish Building Research Institute he studied fungal growth and mycotoxin production in water damaged buildings, supervised by Ulf Thrane and Suzanne Gravesen and received his PhD in 2001. In 2004 he was awarded the research price of the Society of Mycotoxin Research for his PhD work. After holding a position as associate research professor at BioCentrum-DTU, he has just joined Neurosearch A/S.

Michael E. Hansen was born in Nykøbing Falster, Denmark in 1970. He received an MSc degree in engineering from the Department of Informatics and Mathematical Modelling at DTU, Kgs. Lyngby, in 1998 and a PhD degree in mathematics and statistics from the DTU in 2004. He then joined CMB as an assistant research professor. His current research interests include developing statistical methods for the analysis, data mining and data fusion of data extracted from microbiological organisms used within biotechnological applications, e.g. *metabolite profiling, drug discovery and biosystematics.*

Jens C. Frisvad was born in Kgs. Lyngby, Denmark in 1952. He studied chemical engineering at the DTU with focus on organic chemistry and mycology. He received his PhD in 1982 and defended his dr. techn. thesis on "Secondary metabolites and species concepts in Penicillium *and* Aspergillus*" in 1998. He became a full professor in Industrial Mycology at Biocentrum-DTU in 2001. In 1994 he stayed for six months with Professor Martha Christensen, Department of Botany, University of Wyoming, Laramie to study* Penicillium *and* Aspergillus *and has been visiting Dr Robert A. Samson, Centraalbureau voor Schimmelcultures, Utrecht, the Netherlands, extensively to collaborate on mycology and taxonomy.*

Kristian F. Nielsen

Michael E. Hansen

Jens C. Frisvad

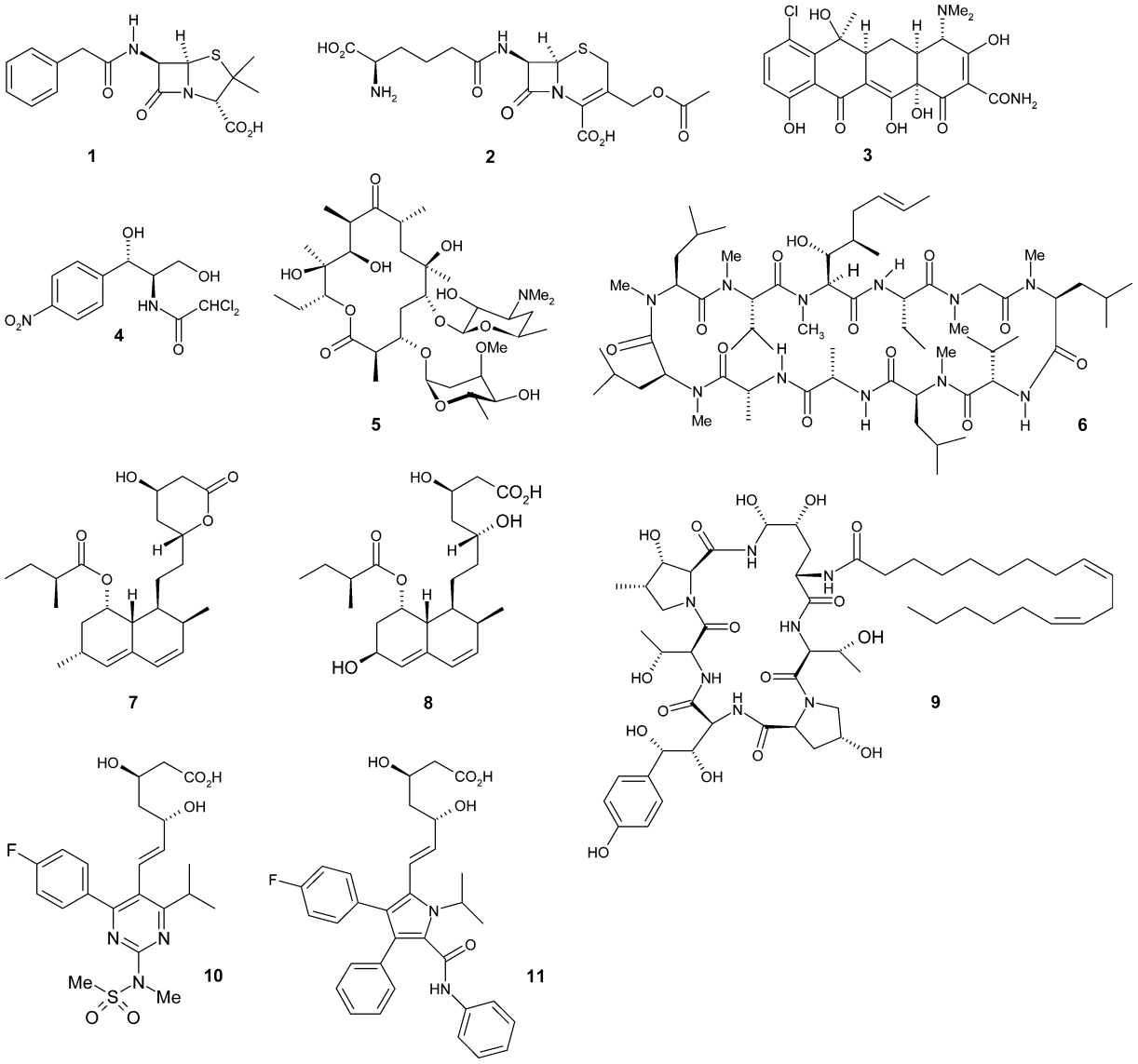
1. The potential of microbial natural products in drug discovery

Microbial natural products (NPs, or secondary metabolites) have played a pivotal role as sources for drug lead compounds during the last century. However, in order for natural product chemistry to continue to be competitive with purely synthetic based discovery methods, natural product research needs to continually improve the efficiency of the selection, screening, dereplication, isolation and structure elucidation processes.[1,2] The main intention of this review is to discuss how taxonomy and information on biodiversity can be used for selection of talented microbial strains to be included in a screening programme and how this together with the use of spectroscopic methods in combination with chemoinformatics can be used as part of an effective dereplication strategy. In this review we use the terms natural products or secondary metabolites (but not the terms specific metabolites or idiolites). We occasionally use the broader term extrolites, which comprise any metabolite that is outwards directed in an ecological sense (they could be extracellular or in the cell wall of the organism). Extrolites can be secondary metabolites, accumulated acids, extracellular enzymes *etc.*[3]

NPs are produced by all organisms but are mostly known from plants, insects, fungi, algae and prokaryotes. All of these organisms coexist in ecosystems and interact with each other in various ways in which chemistry plays a major role. Williams *et al.*[4] proposed that all secondary metabolites serve the producing organisms by improving their survival fitness—"by acting at specific receptors in competing organisms". On the contrary to primary metabolites that are common in all living cells and are involved in the formation of biomass and generation of energy, secondary metabolites are often only produced by one or few species.

Many NPs are biologically active and have been used by man for thousands of years as traditional medicines and as natural poisons. However, it was not until the discovery of penicillin G **1** from a *Penicillium* species about 80 years ago that fungi, actinomycetes and other microorganisms suddenly became a hunting ground for novel drug leads.[5] Hence many pharmaceutical companies were motivated to start sampling and screening large collections of microorganisms especially for antibiotics. About 20 years after the discovery of penicillin several other antibacterial agents such as cephalosporin C **2**, chlortetracycline **3**, chloramphenicol **4**, and erythromycin **5** had been discovered.[1,6] It has been estimated that drugs that trace their heritage to secondary metabolites have more than doubled the lifespan of human beings.[7] Apart from the use of antibiotics to combat bacterial infectious diseases this spans heterogeneous fields such as the use of fungal NPs as immunosuppressants during organ transplantations (cyclosporine **6**), as aids for neurological diseases (asperlicin), for fungal infections (semisynthetic lead compounds derived from echinocandin B **9**),[1] for cardiovascular and metabolic diseases (natural statin compounds such as lovastatin **7** and pravastatin **8** and synthetic analogue compounds such as the major selling synthetic statins crestor **10** and lipitor **11**).[1]

For the past 10–20 years there has been a tendency in drug discovery programmes to favour programmes using combinatorial chemistry for generation of chemical diversity.[1,8,9] The major reason for this has been the tremendous development of high throughput screens based on molecular targets in combination with automated instrument systems, robot technologies *etc.* However, apparently there is a growing opinion, at least among many natural product chemists, that combinatorial chemistry has failed to supplant NPs as the primary source of broad chemical diversity.[1,8,9] One argument is based on the fact that the number of new active substances has been declining during the last 20 years,[9] and the fact that a significant number of the top 35 worldwide selling drugs in the years 2000–2003 are natural product derived compounds (*e.g.* lipitor and other synthetic antilipidemic statins).[1] On the other hand combinatorial chemistry has achieved significant success in more specific discovery programs used to generate focused libraries centered



on core structures with desired activities, rather than finding these from an initial lead compound. More attention is now placed on the quality and diversity of combinatorial libraries, and provided a starting point, such as a natural product scaffold is available, it is clear that combinatorial chemistry is sufficiently advanced to accomplish parallel synthesis.[10] A major advantage of using natural products as drug leads is their often extremely complex structure, making a total synthesis, and the synthesis of their analogues, a daunting task, even when cost and yield are not important.[7] Structurally natural products are more likely to be rich in stereochemistry, concatenated polycyclic rings and reactive functional groups, than structures generated by combinatorial chemistry. For example it is doubtful that the β-lactam ring of the penicillins would have been discovered by synthetic chemists just making molecules at random.[11] In addition to the great chemical diversity produced in *Nature*, there are several other good reasons to choose a natural product based drug discovery strategy. First of all it makes a lot of sense to choose NPs since such compounds have "a biological history" selected by *Nature* during the evolution to serve a function in specific biological systems like binding to proteins.[8,12] In other words NPs that are biologically active in assays are generally small molecules with drug-like properties such as being capable of being absorbed and metabolised by the human body. At the same time *Nature* almost always produces chiral molecules and with the tendency of shifting towards the patenting and marketing of chiral drugs, NPs have the natural advantage of being enantiomeric.[6]

1.1 Antibiotics

Even though more than 30 000 diseases are clinically described today less than one-third of these can be treated symptomatically and even a fewer can be cured.[13] Hence there is an urgent need for new therapeutic agents, with infectious disease control as a striking example. The increasing occurrence of multi-resistant pathogenic strains has limited the effect of traditional antimicrobial treatment, and it has created a global concern that we may soon be facing a post-antibiotic era with reduced capabilities to combat microbes.

One very promising new approach for antibiotics is based on the fact that bacterial colonization and pathogenesis is facilitated by the ability of the bacteria to communicate and thereby coordinate the behaviour of the entire population. Population activity such as biofilm formation is coordinated by simple communication systems which in many Gram-negative bacteria is based on homoserine lactone (HSL) signals, which have been described in numerous pathogens.[14] HSL systems are referred to as quorum sensing (QS) systems, *i.e.* they express target genes in relation to the quorum size (or density) of the population. In most known cases QS systems control expression of virulence factors such as biofilm formation by *Pseudomonas aeruginosa* in the lungs of cystic fibrosis patients.[15] A screening strategy aiming at inhibition of QS is therefore not targeting bacterial growth but instead at blocking the coordination of bacterial population activity. This means that a quorum sensing inhibiting (QSI) drug is not generating a selective pressure on the bacteria, and it is therefore

unlikely that bacteria will develop resistance towards a given QSI compound. Several natural quorum sensing inhibitors have been described within the past *ca.* 10 years some of the latest being the two well known mycotoxins patulin and penicillic acid.[16] Because the *P. aeruginosa* genome has now been fully sequenced, DNA microarray technologies are being used in order to study the effects of potential new QSI hits at the transcriptional level.[16]

Altogether the increasing information on microbial pathogen genomes as well as the completion of the Human Genome Project will provide thousands of disease related targets to be used in future screening for novel drug leads.[13]

2. Microbial biodiversity

A major potential of NPs is the fact that many natural product resources are largely unexplored, and many environmental samples for isolation of interesting microorganisms have often been collected without a defined strategy.[17–19] Diverse habitats are tropical forests soils, the deep sea,[20,21] sites of extreme temperature, salinity or pH, since such habitats often generate novel microorganisms and therefore provide the potential for novel metabolic pathways and compounds.[13] However, at the same time temperate ecosystems should not be excluded especially if novel isolation strategies such as metagenome cloning (see below) is used. Among others the cyclosporins and penicillins were isolated from fungi collected in temperate regions.[13] Even cold regions can be rich in fungal diversity leading to a high hit-rate of novel psychrophilic or psychrotolerant species.[22,23] A number of these species have recently been investigated and found to produce several bioactive cyclic peptides.[24–26] These findings support the hypothesis that fungi from colder climates may be just as chemically prolific (and perhaps just as diverse) as those from tropical climates, the latter which are much more often cited as targets for biodiversity sought in screening programs. In general we find that only relatively few species appear to be dominant in a certain habitat, leading to the isolation of high numbers of strains of the same species. For fungi this is often referred to as the *associated funga*.[27]

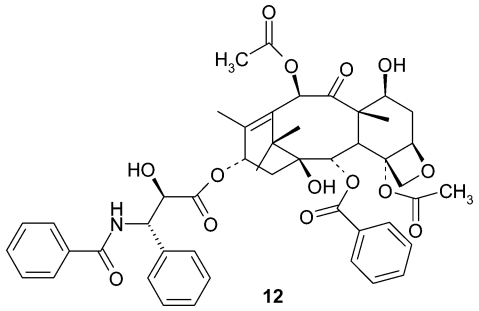
2.1 Marine microorganisms

With more than 70% of the planet's surface covered by water the oceans are probably the most promising habitat to explore for novel microbial biodiversity. It has been estimated that the biological diversity in marine ecosystems, such as the deep sea floor, is higher than in tropical rain forests.[17,28–30] Since the 1970's more than 15 000 NPs have been isolated from marine microbes, algae and invertebrates. It seems clear that many microorganisms such as actinomycetes and fungi are washed from the shore or blown by the air into the sea.[31] On the other hand specific populations of *e.g.* marine adapted actinomycetes such as *Salinospora* and *Marinophilus* have been discovered and described recently.[31] Many such organisms produce marine NPs that possess unique structural features as compared to terrestrial metabolites. Marine NPs often include a myriad of functional groups, which more than make up for some of their disadvantages. A major problem is that many promising bioactive marine compounds can only be isolated in extreme low yields, because many source organisms are difficult to culture by standard fermentation procedures. This is because some of these compounds are only produced as a result of symbiosis between *e.g.* an invertebrate and a microorganism.[32] The advances in molecular genetics are expected to have a great impact on marine natural product chemistry as cloning of polyketides (PKS) or nonribosomal peptide synthetases from "difficult sources" into more amendable bacterial hosts potentially can give an unlimited supply of target compounds.[32–34]

2.2 Terrestrial and insecticidal microorganisms

Hawksworth[35,36] has estimated that approximately 1.5 million fungal species are present on Earth. Out of this number it is suggested that around 100 000 valid species have been described implying that only about 7% of the world's fungi have been described today. Hawksworth and Rossman[37] speculated where to find all the undescribed species and suggested that many of them are likely to be; 1) fungi in tropical forests, and in particular endophytes that can be isolated in enormous numbers; 2) fungi in unexplored habitats such as insects; and 3) lost or hidden species, of which many isolates previously were considered a single species but when studied by modern molecular or biochemical methods, prove to comprise several biological species.

Fungi and particularly endophytes indeed are a very promising source of novel biological active compounds as reviewed by Schulz *et al.*,[38] who found a large hit-rate of novel compounds among the approx. 6500 endophytic fungi that they screened for biological activities. The pharmaceutical potential of endophytic fungi was truly verified with the finding of the taxol **12** producing endophytic fungus *Taxomyces andreanae*.[39,40]



12

The potential of finding new microorganisms associated with insects seems to be immense as illustrated by the discovery of over 200 new species of yeasts from a total of 650 isolates from the guts of beetles.[36] It has been estimated that up to 30 million species of insects exists.[41] Fungi have existed and coevolved with insects some millions of years, and much longer than with mammals,[42,43] so from an evolutionary point of view, it seems likely that a major part of the fungal biologically active metabolites are part of the ecological and in particular chemical defence system directed towards insects.[44] Fungi are generally more nutritious than plant tissue, due to higher levels of proteins, making them potentially desirable sources of foods for predatory insects.[43] Thus predation has without doubt been one of the selective forces shaping the chemical defence systems in fungi.[44] Evidence for this has been several studies on fungal sclerotium producing species of *Aspergillus* which have demonstrated that these compartments contain a variety of sclerotial compounds that cause feeding deterrence or have insecticidal effects.[45–47]

There are many examples of hidden species suddenly being "discovered" especially within taxonomically well studied genera such as *Penicillium*, *Aspergillus* and *Fusarium*. One recent example is the description of the two new species *P. tulipae* and *P. radicicola* within the series *Corymbifera* among the terverticillate *Penicillia*. Both these species produce different profiles of secondary metabolites than the other members of the series: *P. hirsutum*, *P. albocoremium*, *P. allii*, *P. hordei* and *P. venetum*.[48] Another important aspect concerning hidden species is the food safety issue of mycotoxin production. Thus the description of novel species closely related to already known ones might clarify inconsistent literature information about mycotoxin production. This has clearly been the case for *P. roqueforti* from which the two new species *P. carneum* and *P. paneum* have recently been described among others based on their ability to produce secondary metabolites.[49–51] Importantly, only the latter two species can produce the mycotoxin patulin, whereas *P. roqueforti*, applied as a starter culture in food production, will not. The issue of hidden species in combination with the often very difficult task for mycologists to identify a given fungal culture to the species level has meant that the literature is full of data on misidentified species. When chemistry is also reported this leads to erroneous postulates about metabolite production as described in detail in

the next section. Because of these problems many researchers have, in our opinion, come to the wrong conclusion of rejecting the usability of metabolites for chemotyping.[52]

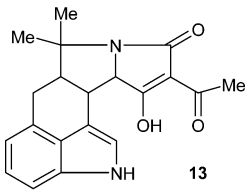
2.3 Non-culturable organisms

The fact that many microorganisms have not been discovered so far is a major challenge for future research since only little effort has been addressed to the isolation and cultivation of organisms difficult to culture. The optimal conditions for growth and secondary metabolite production vary a lot from species to species. Beside the general factors such as carbon, nitrogen, trace metals, temperature, aeration, time of cultivation (see section 3.4), some microorganisms may require stimulation by signal molecules from other organisms in order to grow even when provided with the proper nutrients. Thus the addition of factors such as pyruvate, cyclic AMP and homoserine lactones have all been demonstrated to increase the generation of greater numbers of microorganisms.[13] A second approach to increase the diversity is the use of oligotrophic isolation media, such as the use of seawater based media for marine organisms,[31] allowing only growth of a selected group of strains and at the same time inhibiting the majority of the natural population.[53]

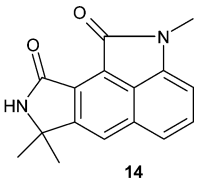
An alternative approach to access unculturable organisms, and in particularly prokaryotic species, is to access their DNA directly by cloning the metagenome.[54–56] Isolated DNA is ligated into bacterial artificial chromosome (BAC) vectors, which are low copy plasmids that can contain relatively large DNA inserts. The BAC vectors are then subsequently transformed into host microorganisms such as *E. coli.* The resulting clones can then be screened for biological activity or alternatively be probed for sequences of interest. This approach is expected to become a powerful resource in the future from which new chemical entities can be accessed for lead discovery.[57]

2.4 Molecular genetics and metabolic engineering

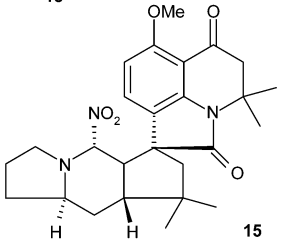
The genes coding for many natural products and in particular polyketide synthethase (PKS) genes are modular and produce multifunctional enzymes. This has lead to new possibilities to diversify unnatural microbial NPs since it is now possible to shuffle genes around within these clusters, or even to include genes from other pathways, thereby generating hybrid enzymes capable of synthesizing an unlimited set of new molecules that are difficult to make by traditional chemical methods.[58–60]



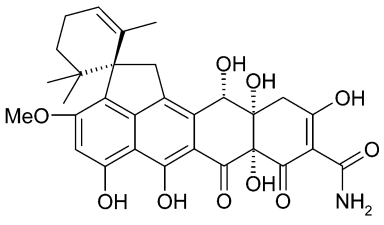
13



14



15



16

Polyketides such as the important compounds erythromycin and lovastatin have been manipulated with great success.[61] In order to fully explore the potential of genetic engineering for industrial strain development methods such as comparison of genomic microarrays, transcription profiles and metabolic profiles are now being used to guide yield improvement. Genetic engineering and modification of targeted pathways will without doubt be very important in future work for construction of novel pathways and NPs.[62–63] Alternatively to directed genetic modifications novel non-natural products can also be achieved more randomly from hybrid organisms generated by cell fusion techniques.[64]

3. Chemotaxonomy and chemo-consistency

"*The production of antibiotic substances by microorganisms is not a property characteristic of specific groups of organisms or even of given species within such groups, but of a few selected strains within a given species*".[65]

"*Thus, the search for novel secondary metabolites from fungi belonging to the group* Nodulisporium *appears to be a random walk in a random forest, at least once one has covered the more common metabolites produced by the genus*".[66]

"*Production of similar metabolic products does not provide an adequate basis for recognition of a new taxon*".[67]

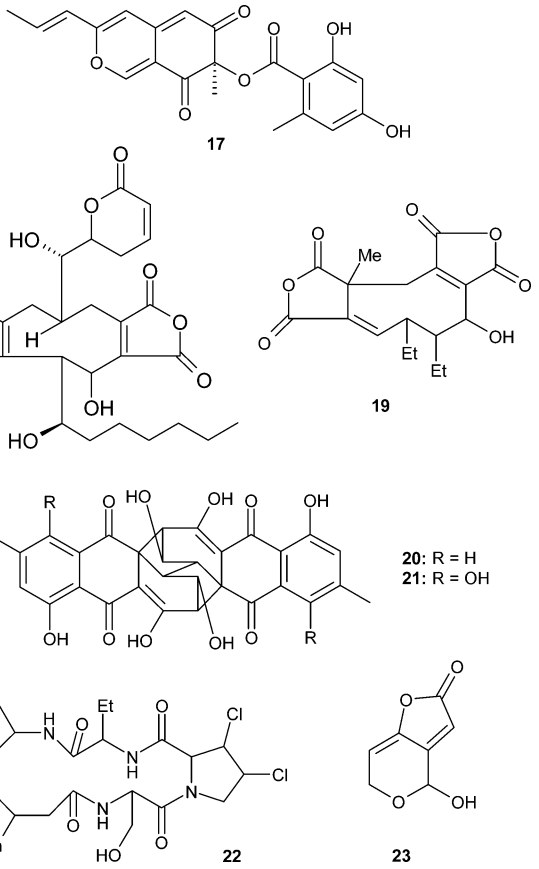
3.1 Fungal species specific production of profiles of NPs

When reading the three citations above, it seems to be a hopeless task to use chemotaxonomy in the classification of filamentous fungi. Opinions on the species specificity of NPs are diverse. Some biologists claim that metabolite production is strain specific,[65,68] some biologists claim that a few NPs may be species specific, but that most of them are strain specific,[66] and yet other biologists claim that most if not all NPs are species specific and even essential features of anyone species.[69] However, in plants NPs appear to be species specific, yet individual NPs often occur among widely different species that are not phylogenetically related according to DNA sequence data.[70] In the fungi chemotaxonomy based on NPs is an extremely effective and scientifically well-founded part of fungal taxonomy, even though it is only used extensively in the Lichens,[71] and few selected fungal genera such as the *Penicillium*, *Aspergillus*, and *Fusarium* and their perfect states.[72–75] The application of chemotaxonomy based on NPs in other ascomycetous fungal genera such as *Xylaria* and *Hypoxylon* also give very good results,[76] and in general NPs have always clarified and greatly improved fungal classifications, when included in revisions of species of different genera in both ascomycetous and basidiomycetous fungi.[72]

It is important to emphasize that species in *Penicillium*, *Aspergillus*, *Fusarium*, *Xylaria*, *Hypoxylon etc.* are identifiable, at least for experts, using traditional micromorphological and macromorphological characters. Species identified based on such features have later been shown to produce consistent profiles of NPs. The reason for species specificity has occasionally been questioned and is often based on compilation of both correctly identified and misidentified producers of particular compounds. It is well known that species in *Penicillium* and *Aspergillus* may be difficult to identify,[77–79] and misidentifications are unfortunately very common.[75,80] One example is cyclopiazonic acid **13**, cyclopiamide **14** and cyclopiamine **15**, which were originally reported from (and named after) *P. cyclopium*. However, the original producer was actually a *P. griseofulvum*.[80] Another well known example is the producer of viridicatumtoxin **16**, which was first identified as *P. viridicatum*,[81] hence the name, later as *P. expansum*,[82] and finally it was realized that the isolate was representing a new species *P. aethiopicum*.[75,83] In this case the misidentification was understandable as the new species was "hidden" at the time of isolation and structure elucidation of viridicatumtoxin, but the name of the compound is misleading now.

In traditional identifications secondary metabolites have only been used indirectly *via* the colour of diffusible pigments, odour of cultures, the KOH test, filter paper methods *etc.*[3,76,78] With the advent of separation methods such as thin-layer chromatography (TLC), high performance liquid chromatography (HPLC), and gas chromatography (GC), and advanced detectors it is now possible to identify the individual NPs (see section 6).

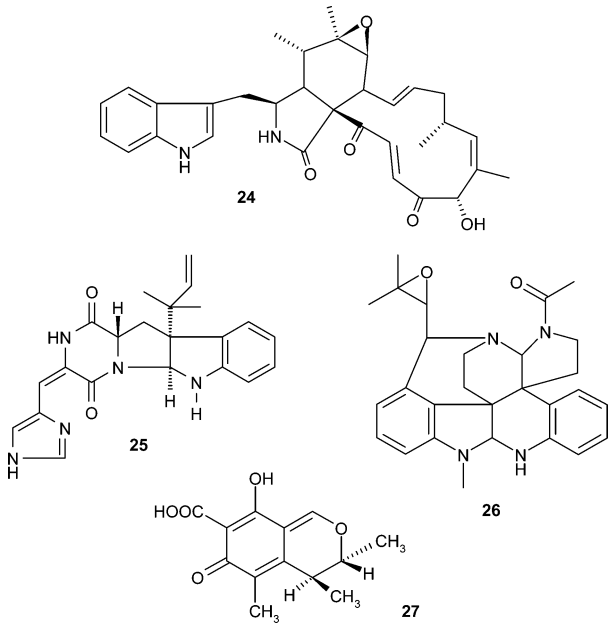
The genus *Penicillium* has been particularly well studied concerning NPs. This genus contains more than 225 accepted species of which 166 belong to the ascomycete genus *Eupenicillium* and 59 belong to the phylogenetically unrelated ascomycete genus *Talaromyces*.[84] Species of *Talaromyces* and their anamorphic states in *Penicillium* subgenus *Biverticillium* produce metabolite biosynthetic families such as mitorubrins **17**, rubratoxins **18**, glauconic acids **19**, rugulosins **20**, luteoskyrins **21**, and cyclochlorotines **22** in species specific combinations,[85] whereas *Eupenicillium* species and associated anamorphs in the subgenera *Aspergilloides*, *Furcatum* and *Penicillium* produce different extrolite families in different species specific combinations.[75,83,86,87]



In the *Penicillium* species examined so far all species produced a large number of already known or not yet structurally characterised NPs.[75] The profile of biosynthetic families of NPs is always species specific, while individual metabolite biosynthetic families have been found in both phylogenetically closely related and distantly related species. For example, the series *Urticicolae* in *Penicillium* subgenus *Penicilllium* section *Penicillium* contains three species, all characterized by very short phialides. These species produce cyclopiazonic acid **13** and patulin **23** in common, but else they produce different combinations of NPs.

3.2 Chemo-consistency

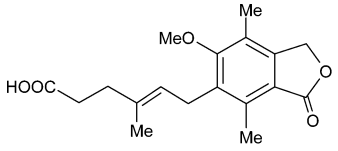
Usually the rather high number of biosynthetic families of NPs detectable in each species is sufficient to unequivocally classify strains into species in *Penicillium* and *Aspergillus*, despite an occasional lack of phenotypical expression of one or two metabolites. Most often the NPs characteristic of any one species are consistently expressed. For example of 85 isolates of *Penicillium expansum* examined, 83 produced patulin **23**, 85 produced chaetoglobosin A **24**, 85 produced roquefortine C **25**, 85 produced communesins **26** and 73 produced citrinin **27**.[88] Likewise a pronounced chemoconsistency was found in *P. crustosum*.[89] It is important to emphasize that chemotaxonomy needs to be based on profiles of NPs rather than the individual NPs, as those individual metabolites may in some cases be absent because of mutations in regulatory or other important genes in the gene clusters responsible for their accumulation.[90] Some of the early claims that antibiotic production may be strain specific,[65,91–93] were probably based on the use of few or suboptimal production media and quantitative rather than qualitative differences. At suboptimal conditions, different strains may produce widely different amounts of NPs, while at optimal conditions, extrolite production is much more consistent.[75,83,94]



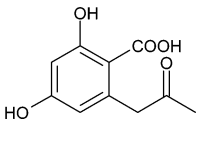
Of 241 strains of the *P. chrysogenum* series, 24 did not produce detectable penicillin G **1**.[95] Results of Brundidge *et al.*[96] also indicated that penicillin production may be less than consistent. However, reexamination of such strains on other media may show that all strains of *P. chrysogenum* produce penicillin. In the closely related species *P. nalgiovense*, all strains examined produced penicillin.[97] In *E. nidulans*, most strains produced penicillin,[98,99] but strains of mating type F did not produce penicillin and the strains apparently lacked the whole penicillin gene cluster.[100] However, these type F strains have never been examined since, and they may also represent other species in the genus *Emericella*, which are not necessarily all producers of penicillin. Again only examination for the presence of genes of the penicillin pathway in these non-producing isolates will tell us if the genes for a whole pathway may be lost completely in some strains of a species.

Mycophenolic acid **28** was cited as being produced by most isolates (12/15) of *P. brevicompactum* by Clutterbuck *et al.*[101] and furthermore the Raistrick phenols (2,4-dihydroxy-6-(2-oxopropyl, benzoic acid)) **29** were produced by 14/15 isolates of strains in the *P. brevicompactum* series (now series *Olsonii* in section *Coronata* of *Penicillium* subgenus *Penicillium*). The one strain producing neither mycophenolic acid nor the Raistrick phenols was *P. aurantiogriseum* var. *poznaniense*, now regarded to be a synonym of *P. aurantiogriseum* in series *Viridicata*.[3] The two strains claimed not to produce mycophenolic acid by Clutterbuck *et al.*,[101] the ex type cultures of *P. stoloniferum* (now *P. brevicompactum*) and *P. biourgeianum* (now *P. bialowiezense*) have later been shown to produce large amounts of mycophenolic acid.[75] Thus in fact the production of mycophenolic acid and the Raistrick phenols was entirely consistent. Later Frisvad and Filtenborg examined 124 strains of *P. brevicompactum* and

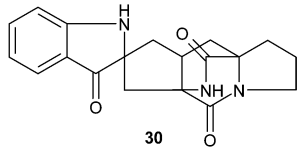
P. bialowiezense, at that time both included in *P. brevicompactum*, and found that all 124 strains produced both mycophenolic acid and Raistrick phenols.[83] It was later realized that while *P. brevicompactum* produced brevianamide A **30**, *P. bialowiezense* never did, but instead produced quinolactacin A **31**.[75] However, both species produce mycophenolic acid, the Raistrick phenols and asperphenamate **32** consistently.[75,86]



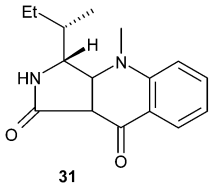
28



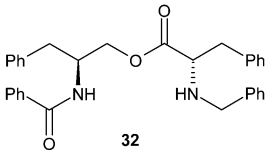
29



30



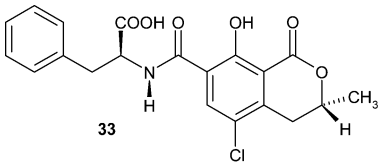
31



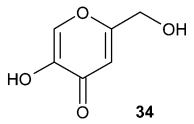
32

Several factors may influence the production of NPs by many strains of a species. 1) All the strains have to be correctly identified, using other identification features than NPs in order not to make a circular argument. 2) Some NPs are only produced under certain environmental conditions and if all trace metals, phosphate and other medium factors are present in certain ranges of concentrations, thus several good media need to be tried out. 3) The seeding medium may influence the production of NPs in the final production medium (the prehistory of the inoculum). 4) The strain needs to be in good condition and not deteriorated because of repeated transfer *etc.* 5) Accumulation of carbon dioxide may inhibit metabolite production, 6) The extraction solvent may also influence the success of detection of the particular extrolite. 7) An appropriate analytical chemical method is needed in order ensure that data are correct. However, genetic factors also play a major role. Often a single point mutation in a regulatory gene is sufficient to make an isolate a non-producer of a metabolite, for which it has the remaining genetic apparatus. In other cases genes may be silent for other reasons, so the metabolite is not expressed.

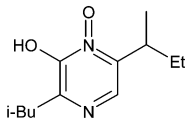
The best known examples of non-production of expected NPs is absence of aflatoxin accumulation in some strains of *A. flavus* and absence of ochratoxin A **33** accumulation in some strains of *P. verrucosum* and most strains of *A. niger* (only 6% positive). While some NPs, such as kojic acid **34** and aspergillic acid **35**, are consistently produced by all strains examined of *A. flavus*, aflatoxin B_1 **36** is only produced by a certain proportion of all strains. For example the culture ex type of *A. flavus* does not produce aflatoxins. On the other hand there have been indications that in so-called non-aflatoxin producing strains, a minority of conidia (single spore inoculations) may give rise to colonies producing aflatoxin anyway.[102] It has been shown that while aflatoxin has never been detected in *A. oryzae*, which is the domesticated form of *A. flavus*, and *A. sojae* which is the domesticated form of *A. parasiticus*, a major part of the genes needed for aflatoxin production are present in *A. oryzae* and *A. sojae* strains.[103–110] So apparently a major part of the genes required for aflatoxin B_1 biosynthesis is present in all strains of *A. flavus*, *A. parasiticus* and their domesticated forms, but they are not expressed because of silent or defective genes in the gene cluster responsible for aflatoxin production. The endeavour to use many different culture conditions have not resulted in detection of aflatoxin in any strain of *A. oryzae* or *A. sojae*.[111]



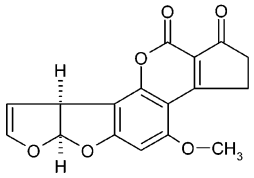
33



34



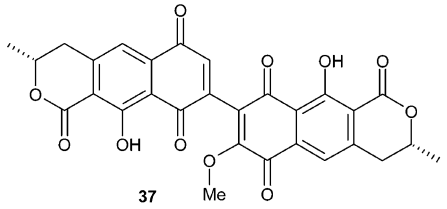
35



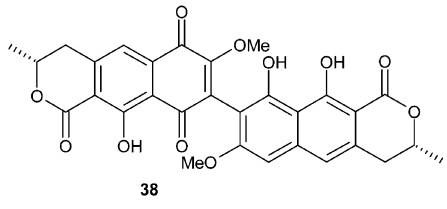
36

3.3 Phylogeny and classification: is production of natural products homoplastic?

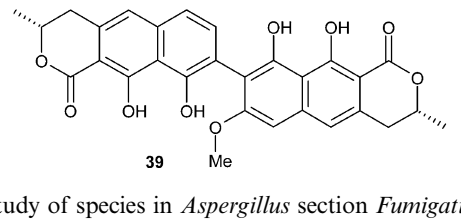
Results from Frisvad *et al.*[75] show that natural series of species often contain species that share several NPs, while some others are only produced by a certain number of the species in the series. In the series *Viridicata* of *Penicillium* subgenus *Penicillium* series *Viridicata*, xanthomegnin **37**, viomellein **38**, vioxanthin **39** and other minor compounds of this naphthoquinone biosynthetic family is produced by five of the nine species known,[112] and could thus either have been lost during evolution in the four remaining species or gained several times. A search for parts of the genes or gene cluster in the non-producing strains will show which hypothesis is correct, but this can only be done when all the genes involved in xanthomegnin biosynthesis have been sequenced. However, production of xanthomegnin by some quite unrelated species of *Penicillium* from subgenus *Furcatum* section *Divaricatum* series *Janthinella* such as *P. janthinellum* and *P. mariaecrucis*,[87] *Aspergillus* section *Circumdati*,[73] and the very distantly related genus *Trichophyton* shows that these naphthoquinones have evolved several times.[113–115]



37



38

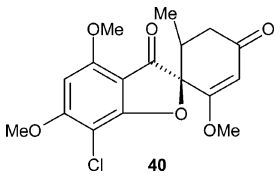


39

In a study of species in *Aspergillus* section *Fumigati* and the related teleomorphic genus *Neosartorya* it was shown that a phylogeny based on partial β-tubulin and hydrophobin gene sequences was not at all congruent with phylogenies suggest by morphology or NPs.[116] Thus functional characters such as profiles of NPs, morphology and physiology are effective in classifications of fungi, but are maybe less well suited for phylogenetic analysis because of many homoplasies or even

analogies in such functional characters. Some examples will be given below.

Griseofulvin **40** was one of the first antifungal NPs found in filamentous fungi. This chlorine containing polyketide is produced by several ascomycetous species, some of them closely related, others very distantly related (Table 1). Griseofulvin production appears to be very consistent in all the species that have been examined systematically so far.[75] According to Table 1, the griseofulvin biosynthetic capability has been developed at least 13 independent times during evolution. All known species of *Khuskia* and its anamorphic state *Nigrospora* have been reported to produce griseofulvin, so this is apparently a monophyletic character of the genus. In contrast griseofulvin production has been developed 9 independent times in *Penicillium*, and thus this character is highly polyphyletic (Table 1). However, it is not known whether the four species in series *Lanosa* and the four species in series *Canescentia* that produce griseofulvin, are those that are most closely related according to a single or multi-gene phylogeny. In *Aspergillus*, griseofulvin production appears to be rare, as it has only been reported in *A. lanosus*, but in no other species in series *Flavi* of *Aspergillus* and its *Petromyces* teleomorph.[121] The production of griseofulvin by *Memnoniella echinata* is also autapomorphic, *i.e.* not shared with phylogenetically closely related species, no other species of *Memnoniella* or the closely related *Stachybotrys* produce this polyketide.[134,135] One strain of *Phomopsis* has also been reported to produce griseofulvin.[120]



40

Mycophenolic acid **28** is a strongly immunosuppressive extrolite and is used for organ transplantations and for treatment of autoimmune diseases (as the drug formulation mycophenolic acid mofetil).[136] It is also an antibacterial, antifungal and antiviral compound.[136] Mycophenolic acid is produced by five Penicillia, one *Aspergillus*, one *Byssochlamys* and one *Septoria* species (Table 2). Mycophenolic acid is thus produced by two out of three species in series *Olsonii*, two out of three species in series *Roqueforti*, but furthermore by 4 individual species in different genera (Table 2). According to β-tubulin sequencing and phylogenetic analysis *P. roqueforti* and *P. carneum*, both producing mycophenolic acid are more closely related to each other than to *P. paneum*, the only species in series *Roqueforti* not producing mycophenolic acid, and likewise *P. bialowiezense* and *P. brevicompactum*, both producing mycophenolic acid are more closely related to each other than to *P. olsonii*, the only species not producing mycophenolic acid in series *Olsonii*.[137] Despite this concurrence within the two series of *Penicillium*, mycophenolic acid biosynthesis seems to have been invented at least 6 times during evolution (Table 2). Most strains of *Penicillium roqueforti* produce mycophenolic acid,[137] although Engel *et al.* claimed the production of mycophenolic acid by this species was strain specific.[68] However, as mentioned above *P. roqueforti* has been shown to be consisting of three different species *P. carneum*, *P. paneum* and *P. roqueforti*.[3,49] *P. carneum* consistently produces mycophenolic acid,[49] whereas some strains of *P. roqueforti* may be non-producers. Again this could be a result of mutations in some of the genes coding for mycophenolic acid in *P. roqueforti* or finding an optimal medium for phenotypic expression of the metabolite.

In some fungal series, such as series *Urticicolae* in *Penicillium*, several NPs are common for all species (Table 3). For example all known members of this series produce patulin **23**, cyclopiazonic acid **13** and possibly penicillin G **1**. On the other hand griseofulvin **40** has only been found in *P. griseofulvum* and *P. dipodomyicola*, while fulvic acid **41** has been detected in *P. griseofulvum* and the new species in the series. In a cladistic sense griseofulvin and fulvic acid are incongruent. In the series *Viridicata* of *Penicillium* the production of verrucosidin **42**, xanthomegnins **37**, viridicatins **43**, puberulins **44**, terrestric acid **45** is neither consistent with the β-tubulin sequence based phylogeny nor are the many NPs produced congruently. Sequencing and comparison of the gene and nucleotide sequences of gene clusters for these NPs will eventually help finding out whether the genes are inherited, horizontally transferred or have evolved several times during evolution. In addition *P. dipodomyicola* produces four new kinds of alkaloids and four new polyketide derived extrolite types. The new species in the series produce

Table 1 Taxonomic placement of griseofulvin **40** producers

Species	Subgenus	Section	Series	Order	References
Khuskia oryzae	—	—	—	*Trichospaeriales*[a]	117,118
Khuskia sacchari	—	—	—	*Trichospaeriales*[a]	118
Nigrospora musae	—	—	—	*Trichospaeriales*[a]	117
Nigrospora sphaerica	—	—	—	*Trichospaeriales*[a]	117
Memnoniella echinata	—	—	—	*Sordariales*[b]	119
Phomopsis sp.	—	—	—	*Diaporthales*[c]	120
Aspergillus lanosus	*Circumdati*	*Flavi*	—	*Eurotiales*[d]	121
P. nodulum	*Aspergilloides*	*Aspergilloides*	*Implicata*	*Eurotiales*[d]	This report
P. aethiopicum	*Penicillium*	*Chrysogena*	*Aethiopica*	*Eurotiales*[d]	83
P. persicinum	*Penicillium*	*Chrysogena*	*Persicina*	*Eurotiales*[d]	122
P. coprophilum	*Penicillium*	*Penicillium*	*Claviformia*	*Eurotiales*[d]	83
P. dipodomyicola	*Penicillium*	*Penicillium*	*Urticicolae*	*Eurotiales*[d]	83
P. griseofulvum	*Penicillium*	*Penicillium*	*Urticicolae*	*Eurotiales*[d]	123,124,125
P. sclerotigenum	*Penicillium*	*Penicillium*	*Expansa*	*Eurotiales*[d]	126
P. jamesonlandense	*Furcatum*	*Ramosum*	*Lanosa*	*Eurotiales*[d]	23
P. lanosum	*Furcatum*	*Ramosum*	*Lanosa*	*Eurotiales*[d]	87
P. raistrickii	*Furcatum*	*Ramosum*	*Lanosa*	*Eurotiales*[d]	127,128
P. soppii	*Furcatum*	*Ramosum*	*Lanosa*	*Eurotiales*[d]	129
P. janczewskii	*Furcatum*	*Eladia*	*Canescentia*	*Eurotiales*[d]	125,130
P. murcianum	*Furcatum*	*Eladia*	*Canescentia*	*Eurotiales*[d]	This report
P. nigricans	*Furcatum*	*Eladia*	*Canescentia*	*Eurotiales*[d]	124,128,131
P. nodusitatum	*Furcatum*	*Eladia*	*Canescentia*	*Eurotiales*[d]	This report
P. yarmokense	*Furcatum*	*Eladia*	*Canescentia*	*Eurotiales*[d]	This report

[a] Sordariaceae, Sordariomycetidae, Ascomycetes. [b] Chaetosphariaceae, Sordariomycetidae, Ascomycetes. [c] Valsaceae, Sordariomycetidae, Ascomycetes. [d] Trichocomaceae, Eurotiomycetidae, Ascomycetes. Unsubstantiated reports, misidentified or reidentified culture: *Aspergillus versicolor*,[132] *Penicillium albidum*,[127] *P. brunneostoloniferum* (=*P. brevicompactum*),[131] *P. concentricum*;[128] *P. kapuscinskii*,[133] *P. melinii*,[128] *P. raciborskii*,[133] *P. verrucosum* var. *corymbiferum*,[128] reidentified to *P. aethiopicum*,[75] *P. viridicatum*,[81] reidentified to *P. aethiopicum*,[80] *P. viridicyclopium* (=*P. cyclopium*).[131]

Table 2 Taxonomic placement of mycophenolic acid **28** producers

	Subgenus	Section	Series	Order	References
Aspergillus unilateralis	*Fumigati*	*Fumigati*	—	*Eurotiales*[a]	This report
Byssochlamys nivea	—	—	—	*Eurotiales*[a]	139
Penicillium bialowiezens[c]	*Penicillium*	*Coronatum*	*Olsonii*	*Eurotiales*[a]	75
P. brevicompactum[d]	*Penicillium*	*Coronatum*	*Olsonii*	*Eurotiales*[a]	140–143
P. fagi[e]	*Furcatum*	*Furcatum*	—	*Eurotiales*[a]	87
P. carneum	*Penicillium*	*Roqueforti*	*Roqueforti*	*Eurotiales*[a]	83
P. roqueforti	*Penicillium*	*Roqueforti*	*Roqueforti*	*Eurotiales*[a]	68,138,144
Septoria nodorum	—	—	—	*Mycosphaerellales*[b]	145

[a] Trichocomaceae, Eurotiomycetidae, Ascomycetes. [b] Mycospaerellaceae, Dothideomycetidae, Ascomycetes. [c] Including the synonym *P. biourgeianum*. [d] Including the synonyms *P. hagemii*, *P. griseo-brunneum*, *P. patris-mei*, *P. scabrum*, *P. stoloniferum*. [e] Including the synonym *P. caerulescens*. Unsubstantiated reports or misidentified culture: *P. aurantiogriseum*,[146] *P. canescens*,[146] *P. carneolutescens*,[147] was a *P. brevicompactum*,[80] *P. expansum*,[146] *P. meleagrinum*,[148] *P. olivicolor*,[146] *P. paxilli*,[146] *P. rugulosum*,[146] *P. viridicatum*.[146,149]

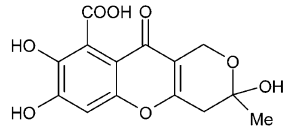
Table 3 Production of NPs by species in *Penicillium* section *Penicillium* series *Urticicolae*

Species	Patulin	Fulvic acid	Griseofulvin	Cyclopiazonic acid	Roquefortine C	Cyclopiamine	Penicillin	Cyclopiamide	Asteltoxin
P. griseofulvum	+	+	+	+	+	+	+	+	−
P. dipodomyicola	+	−	+	+	−	−	?	−	−
New species	+	+	−	+	−	−	?	−	+

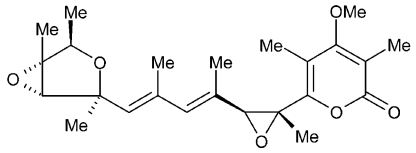
15 different groups of chromophore groups, previously seen in no other or few other *Penicillium* species. The three species in *Urticicolae* also produce many volatile NPs.[75,150]

Chaetoglobosins are cytotoxic mycotoxins that have also been considered as drugs.[151–153] Chaetoglobosin A **24** is active against the gastric ulcer involved bacterium *Helicobacter pylori* and chaetoglobosin K has been suggested for use in combination with other drugs in the treatment of RAS-induced cancers.[154–156] The chaetoglobosins are both phytotoxic and antifungal.[157–160] These very active compounds are produced by an array of very different species, that are not phylogenetically closely related (Table 4). The chaetoglobosins thus seem to have evolved at least 7 independent times.

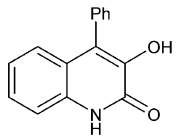
An overview of *Penicillium* subgenus *Penicillium* shows that all series contain species that have specific profiles of NPs.[75] There is a tendency that many NPs are common to several members of a series, so that each series is polythetic,[161] and in addition have autapomorphic NPs for each species. But even these autapomorphic NPs in one series may occur as an autapomorphy in another series. As an example brevianamide A **30** is produced only by two species in subgenus *Penicillium*: *P. brevicompactum* in series *Olsonii* and *P. viridicatum*



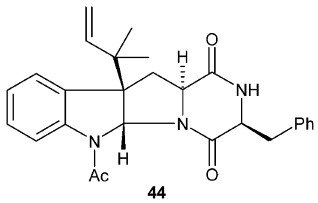
41



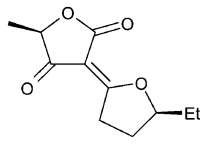
42



43



44



45

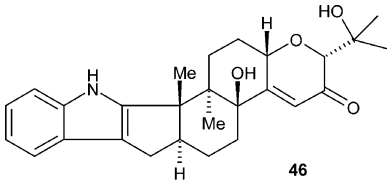
Table 4 Taxonomic placement of chaetoglobosin **24** producers

	Subgenus	Section	Series	Order	References
Calonectria morganii	—	—	—	*Hypocreales*[a]	158
Cylindrocladium floridanum	—	—	—	*Hypocreales*[a]	157
Chaetomium cochlioides	—	—	—	*Sordariales*[b]	167
C. globosum	—	—	—	*Sordariales*[b]	168,169
C. mollipileum	—	—	—	*Sordariales*[b]	167,170
C. rectum	—	—	—	*Sordariales*[b]	167,170
C. subaffine	—	—	—	*Sordariales*[b]	171,172
Diplodia macrospora	—	—	—	*Dothidiales*[c]	173
Discosia sp.	—	—	—	*? (Ascomycetes)*	174
Phomopsis lepstromiformis	—	—	—	*Diaporthales*[d]	175
Phomopsis sp.	—	—	—	*Diaporthales*[d]	120
Penicillium discolor	*Penicillium*	*Viridicata*	*Solita*	*Eurotiales*[e]	176
P. expansum	*Penicillium*	*Penicillium*	*Expansa*	*Eurotiales*[e]	83,88,177
P. marinum	*Penicillium*	*Penicillium*	*Expansa*	*Eurotiales*[e]	75,178,179

[a] Nectriaceae, Sordariomycetidae, Ascomycetes. [b] Chaetomiaceae, Sordariomycetidae, Ascomycetes. [c] Botryospaeriaceae, Sordariomycetidae, Ascomycetes. [d] Valcaceae, Sordariomycetidae, Ascomycetes. [e] Trichocomaceae, Eurotiomycetidae, Ascomycetes.

in the phylogenetically unrelated series *Viridicata*.[75,137] Brevianamide A has not been found in any other microorganism yet.

The accumulation of this anti-insecticidal extrolite[162] in the conidiophores of *P. brevicompactum*[163,164] has given rise to the hypothesis that brevianamide A may deter fungivorous arthropods from consuming the penicillus.[44] Since brevianamide A is also accumulated in the conidiophores of *P. viridicatum*, the same is probably the case for the latter fungus also. One hypothesis could be that brevianamide A has evolved twice, even in these rather closely related species, based on a strong interaction with the environment, *i.e.* selection at the species level. An alternative second hypothesis could be that brevianamide A genes are present in all 58 species, but have been lost in 56 species, based on loss of genes, gene silencing or mutations in the biosynthetic genes for brevianamide A. A third hypothesis could be that the whole gene cluster for brevianamide A has been horizontally transferred from one species to another. Based on the distributions of NPs in both phylogenetically closely and unrelated species, it is most probable that these metabolites have evolved several times during evolution. The second hypothesis would require an unrealistically large genome in all fungi, as most NPs are so widespread in several orders of fungi, that all fungal species should have all extrolite genes inherited from their ancestors if the hypothesis was correct. The tryptophan derived extrolite paxillin **46**, for example, is produced *via* a gene cluster of 17 genes (*ca.* 50 000 nucleotides in all),[165] and this gene cluster is not much larger or smaller than other gene clusters for NPs,[166] and with the very large number of NPs known in *Penicillium* subgenus *Penicillium* there is simply not space for all these gene clusters in any fungal genome. Horizontal gene transfer of such large gene clusters is also quite improbable. For present the hypothesis that NPs have been invented several times during evolution and that ecology has a major impact on the profiles of NPs seen in filamentous fungi seems most probable. If this is shown to be true, intelligent screening for these compounds should be based on ecology and systematics rather than on phylogeny based on household genes. Thus it is recommended to emphasize biodiversity at the species level and explore many different habitats, including a search for new species. Chemical stimulation by co-occurring species in the actual habitat is also a natural consequence of this ecologically based approach. Phylogenetic relatedness may, or may not, indicate that the same NPs are produced and will not give any ecological hints on stimulation of NP production.



46

3.4 Microbial physiology

It is well known that medium composition and culture conditions have great impact on growth and the production of secondary metabolites.[42,66,94,180,181] The physiology of secondary metabolism has often been neglected and still few of the regulatory features of secondary metabolism have been elucidated.[182] Thus depending on the diversity of the microorganisms to be studied it may be necessary to use several media and growth conditions, both during initial morphologically based investigations and later when "talented" strains are to be investigated (see Fig. 1) for their full metabolic potential using a one strain many compounds approach as suggested by Bode *et al.*[183]

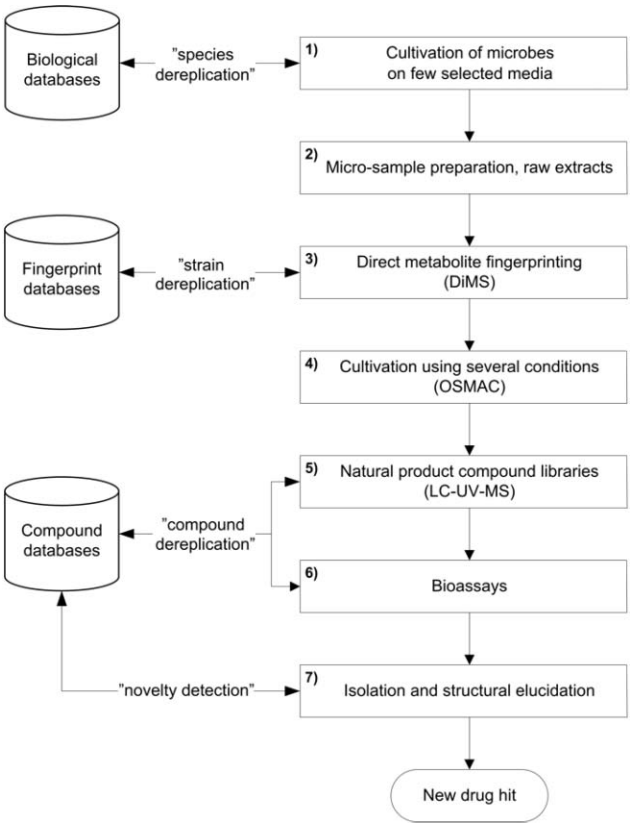


Fig. 1 *The chemotaxonomy based screening approach*. Step 1) Cultures from either natural samples or in-house collections are cultivated on a few media used for macro- and micromorphological identification purposes. Representative strains of the different species are selected by either experts or by automated image analysis methods.[196] Step 2) Extracts are made by micro-extraction of only a few agar plugs from one fungal culture and all metabolites in the extract are analysed simultaneously (step 3) using *e.g.* direct injection mass spectrometry (DiMS). Mass profiles of the extracts are clustered using chemometric methods in order to select representative chemotypes.[52] Step 4) The relatively few strains per species that have now been selected are grown on a larger number of media in order to generate conditions that will allow the expression of as broad a range of secondary metabolites as possible for a given strain, according to the one strain many compounds (OSMAC) philosophy.[183] Step 5) Extracts from the best media conditions can now be separated into microtiter plates generating natural product compound libraries. At the same time UV and MS data are obtained allowing compound dereplication of hits from bioassays (step 6) using comparison of UV and MS spectra with databases. Step 7) Active compounds are isolated and structure elucidated to generate novel drug candidates (step 8), altogether at a high-rate, allowing search for further similar but novel compounds by comparison of spectral data to information in databases using chemoinformatic tools (see section 6.5).

If possible it is very beneficial to know the genus being investigated as the general optimal media for good metabolite production changes.[66,184] For example YES and CYA are in general excellent media for metabolite production for *Penicillium* and *Aspergillus*, whereas they work very poorly for *Rhizopus*.[185] This is also the case for *Penicillium* sub-genus *Biverticillium* (teleomorph *Talaromyces*), where MEA and OAT give many more metabolites than other media.[181]

For initial identification/pre-selection studies solid substrates in Petri dishes are standard. This ensures easy assessment of contaminations of what were supposed to be pure cultures, something which can be difficult to control on media like rice and maize,[186,187] which on the other hand often give good sporulation and metabolite production. Point inoculated cultures furthermore have the advantage that mycelia at different ages are present and thus both intermediates and end-products can be extracted from the edge and the center respectively using *e.g.* the agar-plug-technique.[184]

Different and relatively easy to control conditions to investigate in a discovery programme include growing cultures at both solid and liquid conditions, incubation at two or more temperatures, incubation at two or more shaker speeds, incubation for at least two different time periods, media with at least two different pH levels, choosing carbon and nitrogen sources at different concentrations, high- or low phosphate content, adding trace minerals *etc.*[13] Authors like Hesseltine[188–190] Gloer[45] and Nielsen *et al.*[181] strongly argue in favour of using solid substrate fermentations in studies of fungal metabolites since fungi, unlike other microorganisms, typically grow in nature on solid substrates such as wood, roots, leaves of plants, and drier parts of animals such as fecal material that is low in moisture.

As mentioned earlier others argue that the production of some metabolites demand very specific "stimuli" *e.g.* certain precursors or "triggers" often present in the natural environment of the given fungus. This comes from the philosophy that most species are capable of inhabiting several environments as they would otherwise be too dependent on just one environment, and that a species will need a certain chemical profile in each of these environments. This calls for complementing general screening media with *e.g.* macerated plant tissue media for plant pathogens and endophytes as recently illustrated by the stimulation of some novel fungal phenolic metabolites using plant tissue media.[191,192]

Finally, some believe and argue that all metabolites can be expressed in liquid culture by varying carbohydrate composition, nitrogen source, oxygen tension, pH, redox potential, water activity, as the right conditions will produce intracellular conditions that will trigger production of a certain metabolite. Thus liquid conditions have been shown to be very successful in *Fusarium*.[193] Often metabolites associated to spore or sclerotia formation are produced under solid conditions whereas the production of others are enhanced under liquid conditions. This was demonstrated for *P. solitum* where alkaloids such as viridicatol and cyclopenol analogues where produced in relatively high amounts on semi-solid media either still or absorbed in lightweight expanded Clay aggregates (LECA), whereas the often targeted compactin polyketides were the most dominant compounds produced under submerged conditions.[181]

4. The chemotaxonomy based drug discovery process

As argued above the chemical diversity and the resources of NPs are immense and nowhere near fully exploited. Combined with the fact that fungi (and probably also other types of microorganisms) produce very species specific profiles of NP's that can be used as efficient tools to select one (or a few) representative strain(s) for biological testing, and with the revolution in molecular genomics several new strategies for a NPs based drug discovery programme are being opened:

-*Knowing the biodiversity and ecology*;
-*Using a metabolite profiling approach*;
-*Targeting certain ecological niches*;
-*Using a genome based approach*.

There are several other steps to deal with when running a natural product discovery programme. According to Cordell[194] these key steps can be summarized: (i) collection, selection and cultivation of organisms; (ii) extraction and biological evaluation; (iii) dereplication; (iv) isolation and structure elucidation of metabolites; (v) biological evaluation; and (vi) information management.

With the large numbers of already known microbial (approx. 50 000) and plant (approx. 600 000) metabolites,[195] one of the major challenges in modern natural product discovery is to detect already known and trivial compounds rapidly, a process known as dereplication. However, in our opinion dereplication (or avoiding redundancy) can be implemented in several steps of the discovery process. This relies on information management and data mining of the enormous amounts of biological, chromatographic and spectroscopic data generated and which has become a bottleneck in modern drug discovery. To optimize the drug discovery process dereplication should be implemented at an early stage, where the cultures of a given microbial collection (culture collection or natural samples) are selected for extraction and biological evaluation. Traditionally, microbial strains have been selected based on morphology, rather than on more powerful approaches such as automated image analysis of pure cultures,[196] phenotypic characters including production of secondary metabolite profiles or based on genotyping.[13,197] One reason for the large redundancy in isolation of already known compounds in screening programs is due to the redundancy in selection and screening of multiple strains of the same species already studied by others.

It is therefore very relevant to develop an array of simple analytical methods and combine these methods with informatics, to select representative and promising strains to screen rapidly in a bioassay. As will be discussed in detail below, these methods are based on efficient use of MS, UV and NMR data in combination with modern informatics tools to characterize the nature of mixtures of compounds, in crude extracts. When some representative chemotypic strains have been selected, then their metabolic potential should be investigated using multiple growth conditions as discussed in section 3.4 and argued by Bode *et al.*[183]

Having found both some promising strains and their optimal growth conditions for metabolite production, natural product compound libraries (NPLs) can be prepared using an integrated setup for analysis and automated micro-fractionation into *e.g.* microtiter plates.[16,198] Such NPLs can then be screened in various bioassays and when hits have been generated these can be correlated to spectroscopic data and again to databases,[199] in order to dereplicate, thus avoiding the finding of trivial compounds as early as possible. Subsequently, potential novel compounds can be isolated on a larger scale and their structure elucidated. In the case of new compounds more analogues can be generated either by combinatorial or traditional chemistry. Alternatively an overall screening for new but similar compounds may be accomplished by an automated UV-guided search discussed in more detail in section (6).

The overall combined approach using integrated analytical and informatics techniques were recently presented as an *intelligent screening strategy* by Smedsgaard and Nielsen.[52] In this review we also include image analysis of fungal cultures, other direct profiling techniques and the concept of natural product libraries (NPLs) as part of a slightly modified *intelligent screening system* as illustrated in Fig. 1.

5. Morphology based strain selection—image analysis

In many cases manual inspection by expert microbiologists or mycologists of a strain collection could obviously lead to the dischargement of at least some of these strains based on high morphological similarities. However, dereplication based on macromorphological phenotypic characters can also be automated. Using a special camera system, colours as they appear from the surface of the fungal cultures, can be mapped into discrete arrays of pixel values representing a digital image. Images (micro and/or macro) have to be acquired under totally standardised conditions with a colour and geometry calibrated camera set up, so that absolute colour measurements and comparison can be made. The camera system consists of an integrating sphere (a so called Ulbricht sphere) combined with a photometrically calibrated camera system based on the 3-CCD colour camera. The sphere has a diameter of 36 cm. The inside of the sphere is covered with a faint titanium dioxide paint to create optimum light conditions. Light is then brought into the system through light diodes inside the sphere, giving the sample diffuse

and homogeneous illumination. In a standard camera system the spectral resolution of an image is normally 8 bits/pixel for each colour-channel. This camera system is capable of generating images with a higher spectral resolution for each colour-channel. The pixel resolution is 32-bit RGB *i.e.* a bit depth of 2^{32} = 4 294 967 296 per channel and the full pixel resolution is used in the retrieval process.

Based on the above image acquisition procedure, it has been shown, that image analysis of fungal cultures can be used to identify isolates within certain terverticillate *Penicillium* species.[200] In addition, results from DNA fingerprinting were recently compared with the results obtained from the image analysis.[196,201] The objective of this study was to investigate if image analysis could support or maybe serve as a substitute for subjective phenotyping methods and to substantiate the DNA fingerprinting of *P. commune* isolates; one of the most difficult species to identify within this genus. Fig. 2 illustrates the diversity of the cultures in appearance.

The figure shows four different clones of the *P. commune* after digitization. Although the principles described in Hansen *et al.*,[196] and Dörge *et al.*[200] were slightly different, the overall scenario of the methods were the same. First of all the obtained images contain a large amount of information. To reduce the computational complexity, the regions of interest (ROI) have to be detected before further analysis can be done;[202] that is detection of the Petri dish, followed by the colonies and inoculation points. After having detected the ROI's, features can be extracted from the images. The features used are based on calibrated (RGB) colour measurements extracted from each of the pixels inside the colonies. Even though colours (colour intensities) constitute the most important factor for the human visual system of identification,[203] spatial distribution is also important for the perception and understanding of a scene. Different isolates may have the same global content of pixels having certain colours, but the spatial relation of the pixels determines how we identify them. Therefore textual information has to be extracted as well.

Based on the colour and textural features, statistical models can be created for each of the images representing an isolate.[201] Using these models distances can be calculated in such a way, that the distances between visually similar cultures are low, whereas the distances between visually different cultures are high. This comparison enables the possibility of making queries in a database containing visual (phenotypic) information, as illustrated in Fig. 3. Based on the visual information obtained from an "unknown" isolate it is possible to calculate distances to known isolates in the database. Through different classification methods, such as *e.g.* the nearest neighbour classification rule, it is possible to assign species information to the unknown sample. The studies showed that it was possible to obtain a "leave-one-out" cross-validation identification rate of approximately 93–98% when compared with the identification results based on DNA fingerprinting. The method described by Hansen and Carstensen[201] has additionally been validated on small-spored *Alternaria* species proving to have a good performance when compared to the traditional identification methods.[204]

6. Strategies and methodologies for metabolite profiling and target analysis

Rapid profiling techniques have been desired for many years, which ultimately may determine all metabolites produced by a microorganism. In the current age of *'omics* this quest is now a part of what is known as the metabolomics,[205,206] which aims to detect *all* small metabolites in a cell or organism. In general terms these techniques are segregated into: fingerprinting, footprinting, profiling or target analysis. Fingerprinting aims to get a "chemical picture" of the sample where the signals cannot necessarily be used to detect/identify specific metabolites and depends strongly on the technique used. Profiling techniques require that signals in the profile (*e.g.* peaks in a chromatogram) can be assigned to a specific metabolite whether it is of known structure or not. Finally, target analysis aims to determine and quantify specific metabolites. Fingerprinting, profiling or target analysis can be performed by *e.g.* TLC screening,[184] by mass-profiling using direct infusion ESI-MS,[207] by NMR,[208,209] or more by doing elaborate profiling and target analysis, using hyphenated analytical methodologies *e.g.* GC-MS(-MS),

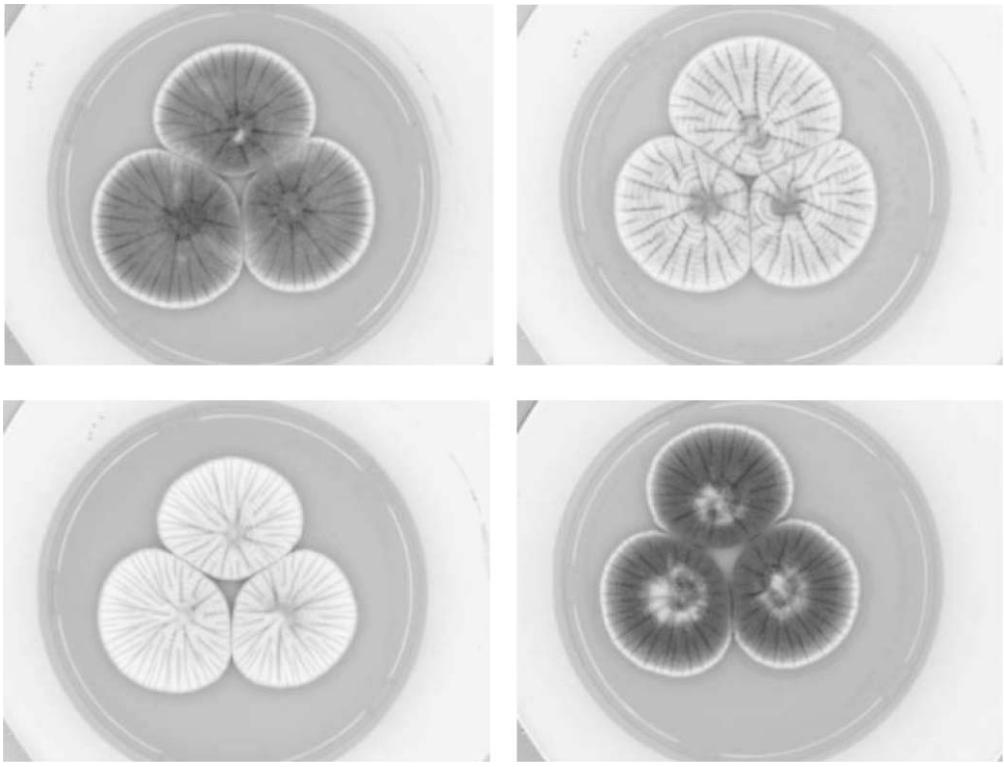


Fig. 2 Example of four Petri dishes after digitization. All isolates are different clones of *Penicillium commune.*

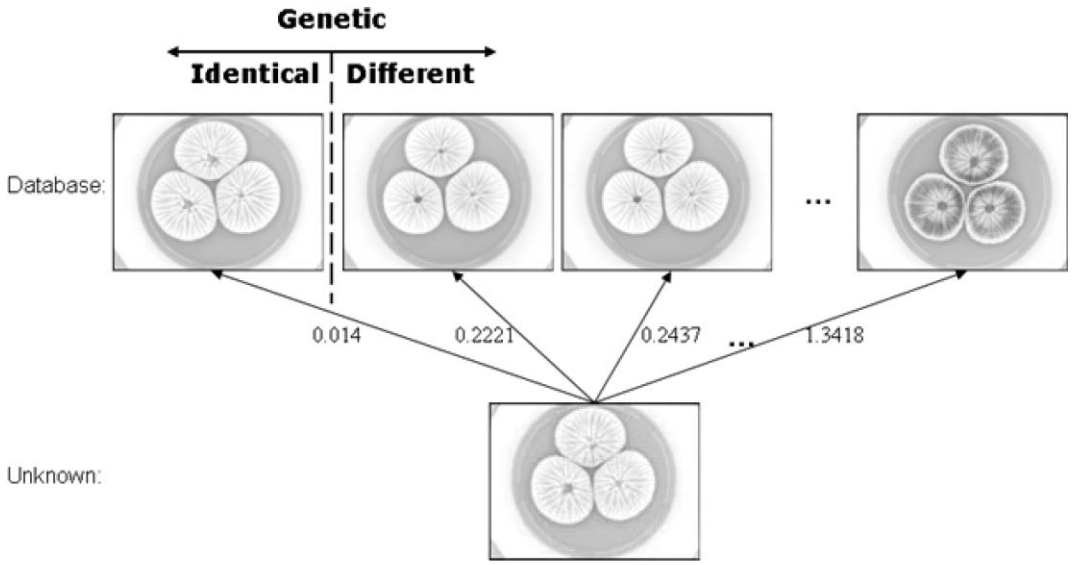


Fig. 3 Statistical models are generated based on visual information from the cultures. Using these models distances can be calculated in such a way, that visually similar cultures have small distances, whereas the visually different cultures have large distances. This comparison enables the possibility of making queries in a database containing visual (phenotypic) information. Through different classification methods it is possible to assign class (species, mutant, *etc.*) information to the unknown sample. Reprinted from Pattern Recognition, 37, Michael Edberg Hansen and Jens Michael Carstensen, Density-based retrieval from high-similarity image databases, pp. 2155–2164, Copyright (2004), with permission from Elsevier.

LC-UV(spectrometric), LC-MS(-MS), LC-NMR and other combinations. All these approaches are relevant in NP discovery.

Most analytical approaches begin with preparation of culture extracts, which can be anything from simple to daunting. Screening fungal cultures can be done in a high throughput screening manner by adapting a rapid plug extraction procedure.[210] However, extraction is (like cultivation) not trivial and consideration should be given to discrimination between metabolites due to extraction procedure and to ensure that the sample matrix does not interfere with the subsequent analyses.

The following sections introduce a few of the methodologies used for metabolite profiling in NP searches in microbes. The reader is referred to text books and original literature for detail on how the analytical work and procedures are carried out.

6.1 Chemical profiling—TLC

The easiest profiling technique to study fungal natural product in extracts are the *agar-plug-TLC* technique developed more than 20 years ago by Filtenborg and Frisvad,[211] and Filtenborg *et al.*,[212] which allow rapid and simple profiling (fingerprinting) of metabolites almost directly from cultures. By this simple technique metabolites are extracted "on the fly" by placing a drop of solvent on the small mycelium plug cut from the culture, where after the plug is placed on a TLC plate with the wetted side down for a few seconds and then removed. The plate is eluted and visualized under UV light by fluorescence or after selective spraying. While the method may seem primitive, it has proved efficient for classification, identification and metabolite detection even under primitive conditions.[213] Also, the TLC technique may be very useful for a first cultural dereplication method in combination with morphological inspection of cultures on fieldtrips collecting new biodiversity. One standard TLC plate can accommodate up to 20 extracts/cultures eluted in two solvent systems, giving a very visual representation of the chemistry. The information is similar in structure but not as detailed as obtained from HPLC analysis as discussed later.

6.2 Direct infusion electrospray mass spectrometry

With the arrival of electrospray ionization mass spectrometry (ESI-MS) and the associated techniques about 15 years ago the scientific community obtained a marvellous tool for studies of NPs and other bio-molecules. ESI-MS has the advantage of being a soft and sensitive ionization technique which can be optimized to produce mainly protonated or sodiated ions (assuming positive ESI) from a very broad range of NPs.[207] Taking advantage of limited fragmentation in ESI-MS Smedsgaard and Frisvad[207] developed a rapid fingerprinting technique where *mass profiles* are determined by infusing crude fungal extract directly into the electrospray source. Later both Julian *et al.*[214] and Higgs *et al.*[215] used similar approaches as rapid methods to differentiate and estimate the presence of secondary metabolites in microbial extracts. The advantages of fingerprinting by direct infusion mass spectrometry (DiMS) are: fingerprints can be made within minutes; metabolite and chemical structure information can be predicted/extracted; data can easily be stored in databases;[216] and data processing is relative easy to automate.[217–218] However, a warning note: infusing complex samples with many components directly into ESI-MS may lead to serious discrimination due to what is known as matrix effects. These matrix effects can discriminate the spectra so that metabolites (or co-extracted media components *e.g.* PEG and TWEEN) may "steal" all charges thereby suppressing other metabolites. Keeping the concentration within a suitable range, using nano-ESI techniques and selected solvent compositions can reduce these effects. As shown in the following examples mass profiles can, despite these problems, be used through chemometric methods to classify the samples (fungi) thus grouping the strains based on their chemical similarities (NP-profile). As fungal species normally produce stable and often quite unique profiles of secondary metabolites, as discussed in the previous section, these mass profiles contain species-specific information.

This was demonstrated almost 10 years ago by Smedsgaard and Frisvad in a study of a large group of fungal species (43 species on two media, approx. 293 strains).[219] They found that more than 80% of these species could be classified into chemical classes from the mass profiles which corresponded to species as determined by classical phenotypic identification. Storing these spectra in the normal mass spectra library included in

the instrument software many species could be identified semi-automatically.[216] These findings have been confirmed in a recent study now including 57 species and about 500 strains on two cultivation media.[218]

Examples of accurate mass profile fingerprints from direct infusion ESI-MS analyses of extracts from three different, but related species of *Penicillium* (associated to dung) are shown in Fig. 4 (see Smedsgaard *et al.*[218] for experimental details). As can be seen, they share similarities such as an intense ion at m/z 468.25, as well as differences *e.g.* only two of the species have ions at m/z 448.19.

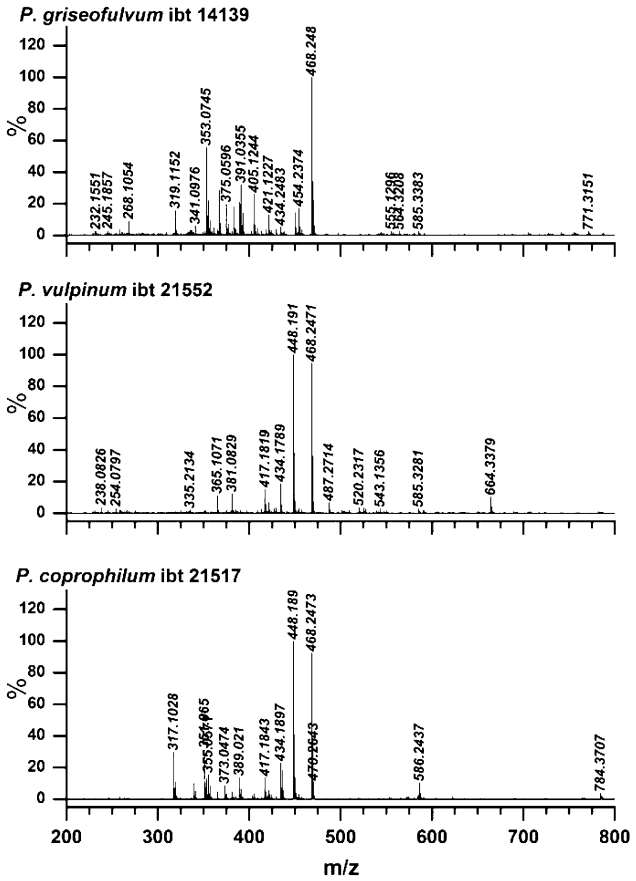


Fig. 4 Mass profile (fingerprint) from direct infusion of crude extract from three different *Penicillium* species as described by Smedsgaard *et al.*[218]

These mass profiles contain species specific information as discussed in the previous sections and can be used for classification/identification. To illustrate this, 72 DiMS mass profiles (including the three showed in Fig. 4) from strains of the 8 major *Penicillium* species associated to dung (from the series *Claviformia* and *Urticicolae*) and two outliers were classified by cluster analysis as shown by the dendrogram in Fig. 5a. The classification was done by a binning technique: the ions in each spectrum were binned into m/z 0.5 wide bins; empty bins are removed from the analysis. The (*bin*, *ion-count*) pairs was sorted according to the ion count in each (*bin*, *ion-count*) and the $N = 50$ bins with the largest ion count are used to represent each sample. The bins are given a score according to the *ion-count* in that bin. The result is a *bin-score* vector calculated from each spectrum. Within each of the class (species) some ions are present in all the isolates, others may be present in only a few, and therefore only bins with a score lying within a certain interval are used. From these selected bins (after they are centered and scaled) the distances are calculated between all samples and clustered using WPGMA (weighted average distance) linkage giving the dendrogram shown in Fig. 5a, a more detailed discussion can be found in Hansen and Smedsgaard,[217] and Smedsgaard *et al.*[218]

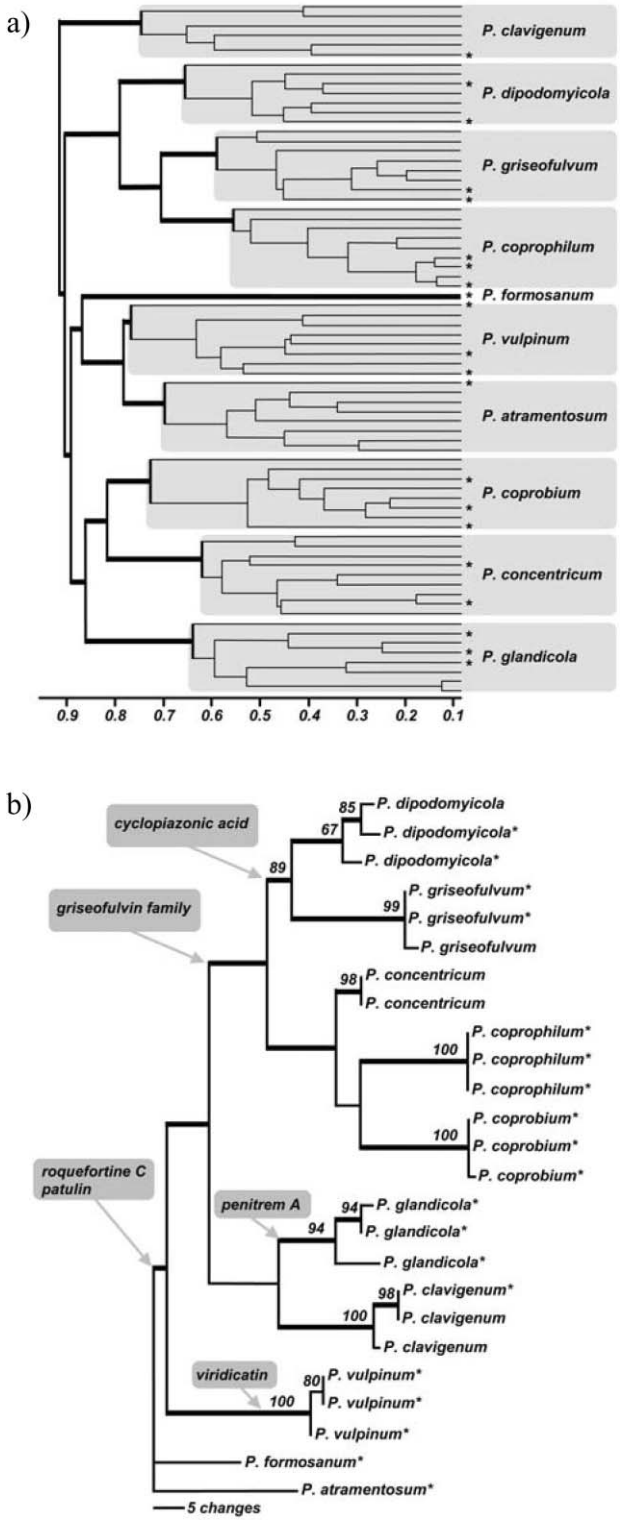


Fig. 5 Upper (a). Classification of the 7 species in *Penicillium* series *Claviformia* and two outliers (*P. formosanum* and *P. atramentosum*) based on cluster analysis of the raw mass profiles from direct infusion ESI mass spectrometry of crude extract.[218] The DiMS classification point to the same species groups as classification based on partial beta-tubulin sequences (lower, b) see Samson *et al.*[137] Strains marked with * are included in both analyses.

The chemical classification based mass profile fingerprinting shown in Fig. 5a point to the same groups as cladification based on partial β-tubulin sequencing.[137] These classes have also been confirmed by classical phenotypic classification by experts.[69] This confirms that NPs are closely linked to species, hence selecting one or a few single isolates within each group rather

than all isolates for further studies will give maximal chemical diversity in a screening programme.

As mentioned above, direct infusion ESI-MS can be optimized to produce mostly protonated and sodiated molecular ions from many different chemical compounds. Using the more recent high resolution and high accuracy mass spectrometers each ion seen in these spectra is likely to represent one or just a few metabolites and the chemical structure can with sufficient accuracy be limited to only a few formulae. However, to obtain the best possible accuracy an internal mass reference is needed. Since most dung associated *Penicillia* will produce roquefortine C **25** with a $[M + H]^+$ at *m/z* 390.193, easily seen in positive mass profiles, the accurate protonated mass of this compound can be used to correct the mass scale of the bottom spectrum in Fig. 4 to get the accurate spectrum from *P. coprophilum* shown in Fig. 6.

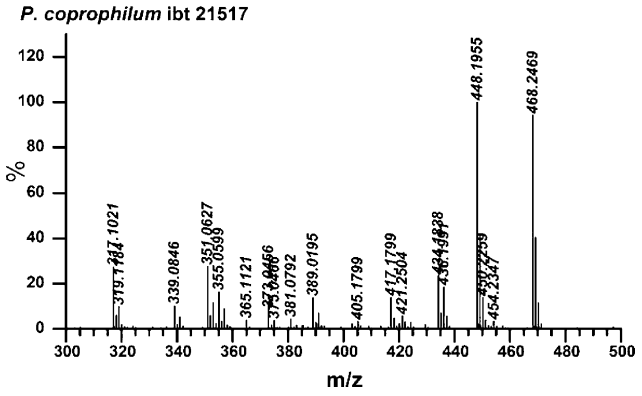
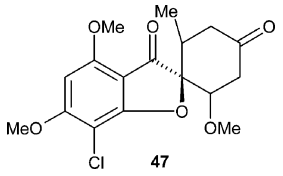


Fig. 6 Mass profile (fingerprint) of *P. coprophilum* (IBT 21517) cultivated on YES after the mass scale have been corrected using the accurate mass of roquefortine C as internal reference. 37 ions above 3% base peak intensity are listed in Table 5.

Above 3% of the base peak intensity we find 37 ions, of these 19 correspond to the formula of metabolites known to be produced by *P. coprophilum*, most within a few ppm (Table 5). Five significant unknown ions are included, however these correspond most likely to three metabolites as the difference between *m/z* 417.1798 and 418.1843 ions correspond closely to the difference between ^{12}C and ^{13}C in both mass and intensity.

Looking closer at the griseofulvin group of ions around *m/z* 350–358, as shown in Fig. 7 bottom spectrum, a series of ion patterns corresponding to one chlorine atom can be seen. If we calculate the ion pattern for dehydro-griseofulvin shown at the top of Fig. 7 and griseofulvin in the middle of Fig. 7 we can see a good match. However, the intensity of *m/z* 355.0599 is higher than expected indicating the presence of another metabolite in this peak. One could speculate that dihydro-griseofulvin **47** might also be present accounting for the 357 peak.



47

To validate the findings indicated by the mass profiles in Fig. 7, the same sample was analyzed by LC-UV-accurate MS and narrow ion traces were extracted for all these ions shown in Fig. 7. As the mass spectrometer was overloaded profiles for both ^{12}C and ^{13}C were extracted. Multi-plate channel detectors (MCP) with time to digital conversion (TDC), a commonly used detector in a TOF mass spectrometer suffers from dead time which in case of saturation will give too low mass values, as seen in the following section. Therefore, the mass traces shown in Fig. 8 are all made ±0.01 Da e^{-1} except in case of serious overload where the trace window is extending to include lower masses. Looking at the traces corresponding to the mass of the ^{13}C isotopes of dehydro-griseofulvin and griseofulvin, chromatographic peaks about 0.2 minutes apart can easily be seen in traces from both chlorine isotopes. The area for the two chlorine traces corresponds nicely to the expected natural occurrence and there is about three times more dehydro-griseofulvin than griseofulvin. These finding can be confirmed by looking at the UV spectra.

Table 5 37 ions are found above 3% base peak height in Fig. 5. Of these can 19 ions be assigned to metabolites known to be produced by *P. coprophilum* within a few ppm mass accuracy. Roquefortine C was used as natural occurring internal mass reference. Five significant unassigned ions are included, X1–X5, see text

Metabolite	Formula	Ion	Calculated	Measured
Dehydro-dechloro-griseofulvin	$C_{17}H_{16}O_6$	H^+	317.1025	317.1021
^{13}C-dehydro-dechloro-griseofulvin	$^{13}CC_{16}H_{16}O_6$	H^+	318.1058	318.1072
Dechloro-griseofulvin	$C_{17}H_{18}O_6$	H^+	319.1181	319.1184
Dehydro-dechloro-griseofulvin	$C_{17}H_{16}O_6$	Na^+	339.0845	339.0846
Dechloro-griseofulvin	$C_{17}H_{18}O_6$	Na^+	341.1001	341.1006
^{35}Cl-dehydro-griseofulvin	$C_{17}H_{15}O_6{}^{35}Cl$	H^+	351.0635	351.0627
^{35}Cl–^{13}C-dehydro-griseofulvin	$^{13}CC_{16}H_{15}O_6{}^{35}Cl$	H^+	352.0669	352.0667
^{37}Cl-dehydro-griseofulvin or	$C_{17}H_{15}O_6{}^{37}Cl$	H^+	353.0606	353.0630
^{35}Cl-griseofulvin	$C_{17}H_{17}O_6{}^{35}Cl$	H^+	353.0792	
^{37}Cl-^{13}C-dehydro-griseofulvin	$^{13}CC_{16}H_{15}O_6{}^{37}Cl$	H^+	354.0639	354.0663
^{35}Cl-^{13}C-griseofulvin	$^{13}CC_{16}H_{17}O_6{}^{35}Cl$	H^+	354.0825	
^{35}Cl-^{13}C-griseofulvin	$C_{17}H_{17}O_6{}^{35}Cl$	H^+	355.0762	355.0599
^{37}Cl-^{13}C-griseofulvin	$^{13}CC_{16}H_{17}O_6{}^{37}Cl$	H^+	356.0796	356.0634
^{35}Cl-dehydro-griseofulvin	$C_{17}H_{15}O_6{}^{35}Cl$	Na^+	373.0455	373.0456
^{35}Cl-griseofulvin	$C_{17}H_{17}O_6{}^{35}Cl$	Na^+	375.0611	375.0466
X1				389.0195
Roquefortine C	$C_{22}H_{23}N_5O_2$	H^+	390.1930	390.1930
X2				417.1798
X3				418.1843
Meleagrin	$C_{23}H_{23}N_5O_4$	H^+	434.1828	434.1838
Meleagrin	$^{13}CC_{22}H_{23}N_5O_4$		435.1862	435.1876
X4				436.1991
X5				437.2033
Oxaline	$C_{24}H_{25}N_5O_4$	H^+	448.1985	448.1955
Oxaline	$^{13}CC_{23}H_{25}N_5O_4$	H^+	449.2018	449.2027
Cyclopiamine		H^+		468.2469
^{13}C-cyclopiamine		H^+		469.2549

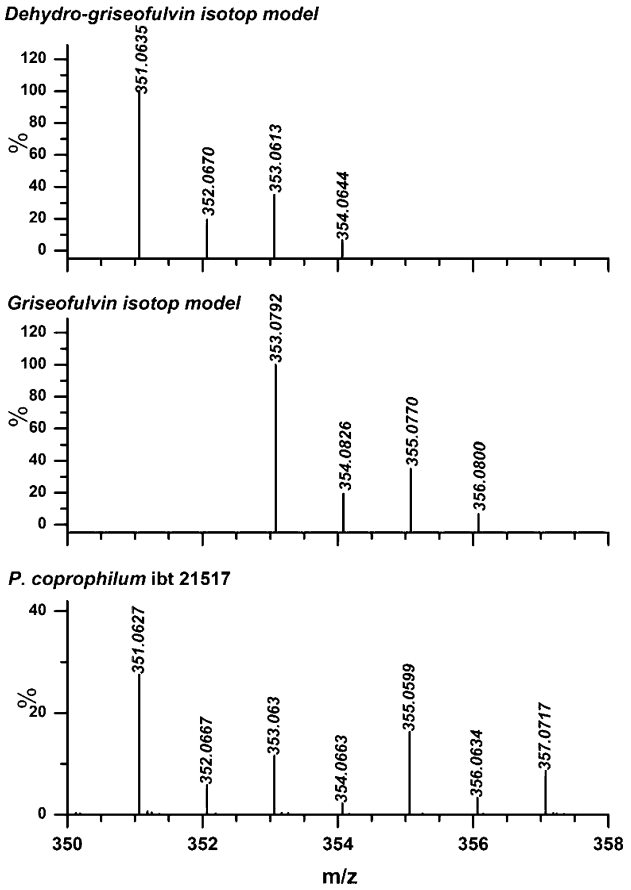


Fig. 7 Prediction of metabolite production from the mass profile of *P. coprophilum* (IBT 21517). Shown at the bottom is a small section of the full mass profile from Fig. 6. The ions found in this part of the spectrum correspond closely to spectra calculated for protonated dehydro-griseofulvin and protonated griseofulvin.

The formula of the unknown ions shown in Table 5 can be predicted under the assumption that only carbon, hydrogen, oxygen and nitrogen are present and given a mass accuracy. The mass accuracy can in this case be predicted from the known ions, in this case about 5 ppm. If we furthermore assume that 417 is the ^{12}C isotope and 418 is the ^{13}C isotope (require that exactly one ^{13}C is present) we can calculate all possible formulae. The results are shown in Table 6. Only two formulae fit within the limits for both ions: $C_{23}H_{23}N_5O_3$, $C_{25}H_{25}N_2O_4$ assuming that these ions are protonated we need to subtract a proton to get the real formula. The first of these formulae will then have an even number of hydrogens which is not possible with an odd number of nitrogens, hence the second formula is the most likely. Several databases may help getting a candidate for this formula, see next section. A similar analysis can be done for the unknown ion pair X4 and X5 ion pair and *m/z* 436 and 437 respectively.

Table 6 Calculating the elemental composition of the unknown X2 and X3 ions from Table 5. Mass accuracy 5.0 ppm, double bond equivalent min = −0.5, max = 50.0 considering both odd and even electron ions. Search limits $^{12}C < 500$; $H < 1000$; $N < 7$; $O < 13$. First search X2 of *m/z* 417.1798 assuming all ^{12}C carbons: 481 formulae were evaluated with 3 results within limits. Second search X3 of *m/z* 418.1843 requiring *one* ^{13}C: 4536 formulae were evaluated with 5 results within limits

Mass	Calc. mass	mDa	ppm	DBE	Formula
417.1798	**417.1801**	**−0.2**	**−0.6**	**15.0**	**$C_{23}H_{23}N_5O_3$**
	417.1788	1.1	2.6	10.0	$C_{22}H_{27}NO_7$
	417.1814	**−1.6**	**−3.8**	**14.5**	**$C_{25}H_{25}N_2O_4$**
418.1843	**418.1848**	**−0.5**	**−1.1**	**14.5**	**$^{13}C^{12}C_{24}H_{25}N_2O_4$**
	418.1834	**0.9**	**2.1**	**15.0**	**$^{13}C^{12}C_{22}H_{23}N_5O_3$**
	418.1853	−1.0	−2.3	2.0	$^{13}C^{12}C_{10}H_{27}N_7O_{10}$
	418.1861	−1.8	−4.3	19.5	$^{13}C^{12}C_{25}H_{21}N_6$

6.3 Target analysis and dereplication by MS

Effective dereplication methodologies should aim to ensure that isolation, structure elucidation and pharmacological investigations are focused on novel compounds. In this scenario, mass spectrometry and especially high resolution mass spectrometry is the core technology for dereplication used in combination with databases such as *Antibase*, and *MarinLit*.[199,214,220–224] It can be argued whether dereplication should be done before or after a bio-assay step (see overview Fig. 1).

The key issue in dereplication is to assign the ions to chromatographic peaks in the LC-MS spectrum correctly, a task that can vary from being very easy to impossible. In positive mode alkaloids usually ionize well and amides reasonably well, whereas sugars and other polyols ionise very poorly. Carboxylic acids have a reputation of ionizing poorly in positive mode, however this is not always the case, although it is certainly true for phenolics and aromatic carboxylic acids.[225] Even small structural changes between molecules can induce significant changes in response factors, a good example of this scenario is the type A and B trichothecenes. The keto group in the Type B (presumably tautomerisation into a triol), are poorly detected using ESI^+, however very good ionization is seen when using ESI^-, whereas the opposite is the case for the type A ones. Therefore it may be advantageous to use on-line positive/negative ionisation during the chromatographic run. Ion-traps and quadrupole instruments can easily do this whereas TOF instruments have problems doing this and still run in accurate mass mode. Moreover buffering of the solvent system is even more restricted using polarity switching so for trace analysis or for very dirty samples it is still advantageous to run the samples in separate runs with different solvent systems.

It should be noted that the ions formed in both positive and negative mode are even-electron ions compared to classic positive electron impact ionization (EI^+) mass spectrometry (MS) where odd-electron ions are formed. This means that the fragmentation mechanisms are fundamentally different.

ESI^+ is the most versatile ionization mode and when working with low resolution data a deconvolution analysis of the data file is the fastest. However, this is not possible using the full accuracy of modern TOF instruments as no commercial software can handle this at the moment. In this case the approach is to: i) obtain a background subtracted average spectrum of the peak of interest; ii) deconvolute all the ions from this spectrum to make sure that they are not coming from co-eluting compounds. The spectrum is first investigated for the presence of A + 2 and A + 4 ions, showing chlorine, bromine, and many (>3) sulfur atoms just as known from classic EI^+ mass spectrometry. The deconvoluted mass spectrum is then assessed for ions seen in Table 7, this means in practice jumps of 22 (H^+ to Na^+), 5 (NH_4^+ to Na^+), 63 (H^+ to $Na^+ + CH_3CN$), 18 (loss of water), 36 (loss of 2 waters) and other jumps calculated from Table 7. Spectra should also be examined for dimers $[2M + H]^+$ and $[2M + Na]^+$, which for some compounds can be even more predominant than the monomeric ions. Acids can exchange H^+ to Na^+ by a simple ion exchange mechanism and subsequent adduct formation with Na^+ gives $[M–H + 2Na]^+$, which very strongly indicates that the target compound is a carboxylic acid.

Sodium adduct ions are very stable compared with both protonated and ammoniated ions and we have found that identification of the sodium adduct ion(s) is very important for a confident molecular mass determination. To make it easier to find the sodiated adduct ion(s) it is advantageous to use ion-source CID fragmentation as pseudo MS/MS, as it will enhance their abundance along with fragment ions whereas the relative abundance of the protonated and ammoniated ions will decrease. This is done by using two "simultaneous" scan

Table 7 Common ions observed during LC-MS analyses of small molecules[a]

	Positive	Negative
Common adducts	$[M + H]^+$, $[M + NH_4]^+$, $[M + Na]^+$ $[M + CH_3CN + H]^+$, $[M + CH_3CN + Na]^+$, $[M + K]^+$	$[M - H]^-$, $[M + HCOO]^-$, $[M + CH_3COO]^-$
Less common adducts	$[M–H + 2Na]^+$	$[M - 2H + Na]^-$, $[M + Na - 2H + CH_3COOH]^-$
Common fragments	$[M + H - H_20]^+$, $[M + H - 2(H_2O)]^+$, $[M + H - 3(H_2O)]^+$, $[M + H - CO_2]^+$	$[M - H - CO_2]^-$
Less common fragments	$[M + H - CH_4O]^+$, $[M + H–CH_3COOH]^+$, $[Fragment + NH_3]^+$ $[Fragment + CH_3CN]^+$	
Polymeric ions	$[2M + H]^+$, $[2M + NH_4]^+$, $[2M + Na]^+$	$[2M - H]^-$

[a] (<1500 Da).

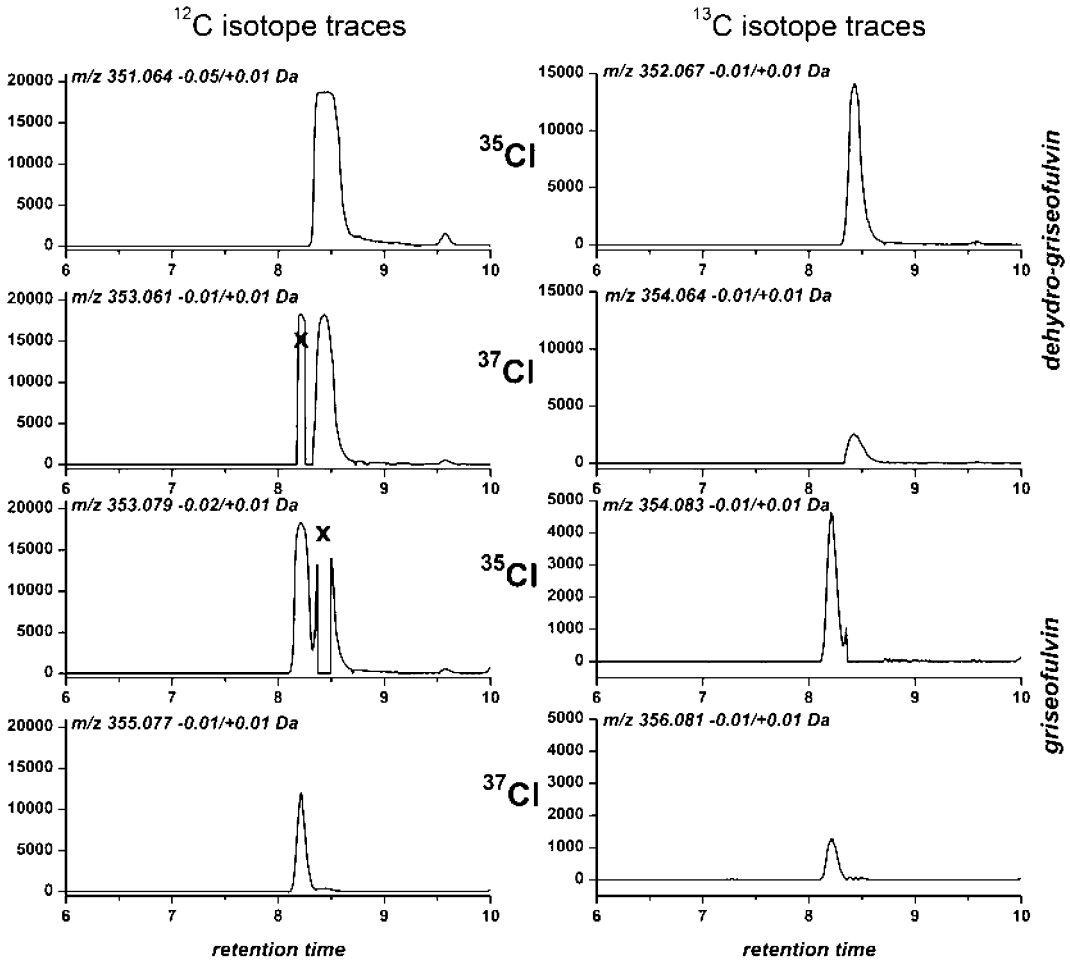


Fig. 8 Confirmation of some of the findings shown in Fig. 7 using narrow ion traces (±0.01 Da except in two chromatograms) from accurate LC-MS analysis of the same sample. The detector was overloaded for some of the ions (the ^{12}C isotopes of dehydro-griseofulvin and griseofulvin) giving too low mass values due to the detector dead time error. The two peaks marked with an X are due to the high intensity error.

functions where one is set with a very low potential difference between the two skimmers in the ion-source and one with a high potential difference. The deconvoluted mass spectra from the two scan functions are then compared.

As seen in Table 7 fragment ions with adducts of acetonitrile and NH_3 occasionally occur, however, we have rarely if ever seen fragments with sodium adducts, although they can be formed in an ion-trap.[226] In ion-source CID, small molecules are fragmented much more vigorously than larger molecules as the kinetic energy (potential difference between skimmers) is the same, meaning that the velocity is proportional to the square root of the potential difference between the skimmers, *e.g.* at *m/z* 150 Da ion will get double the velocity of a *m/z* 600 ion and thus be fragmented much more.

Except for scientists working with specific groups of compounds not ionizing very well in positive mode, positive ionization data will usually be available, and ESI^- will in many cases just be used to confirm what is suspected from positive mode unless "no ionization" was obtained by ESI^+. Nevertheless it is always safest to have a molecular mass assigned from analysis using both polarities indicating the same molecular weight (see for example data for chaetoviridin A **48** in Fig. 9). Usually ESI^- generated spectra are far simpler to analyse than ESI^+ spectra. When using ESI^- $[M - H]^-$ or an adduct with formate and/or acetate can be detected, depending on which of the acids is present in the solvent. These acid adducts can some times be so predominant that the deprotonated ions are not detected.

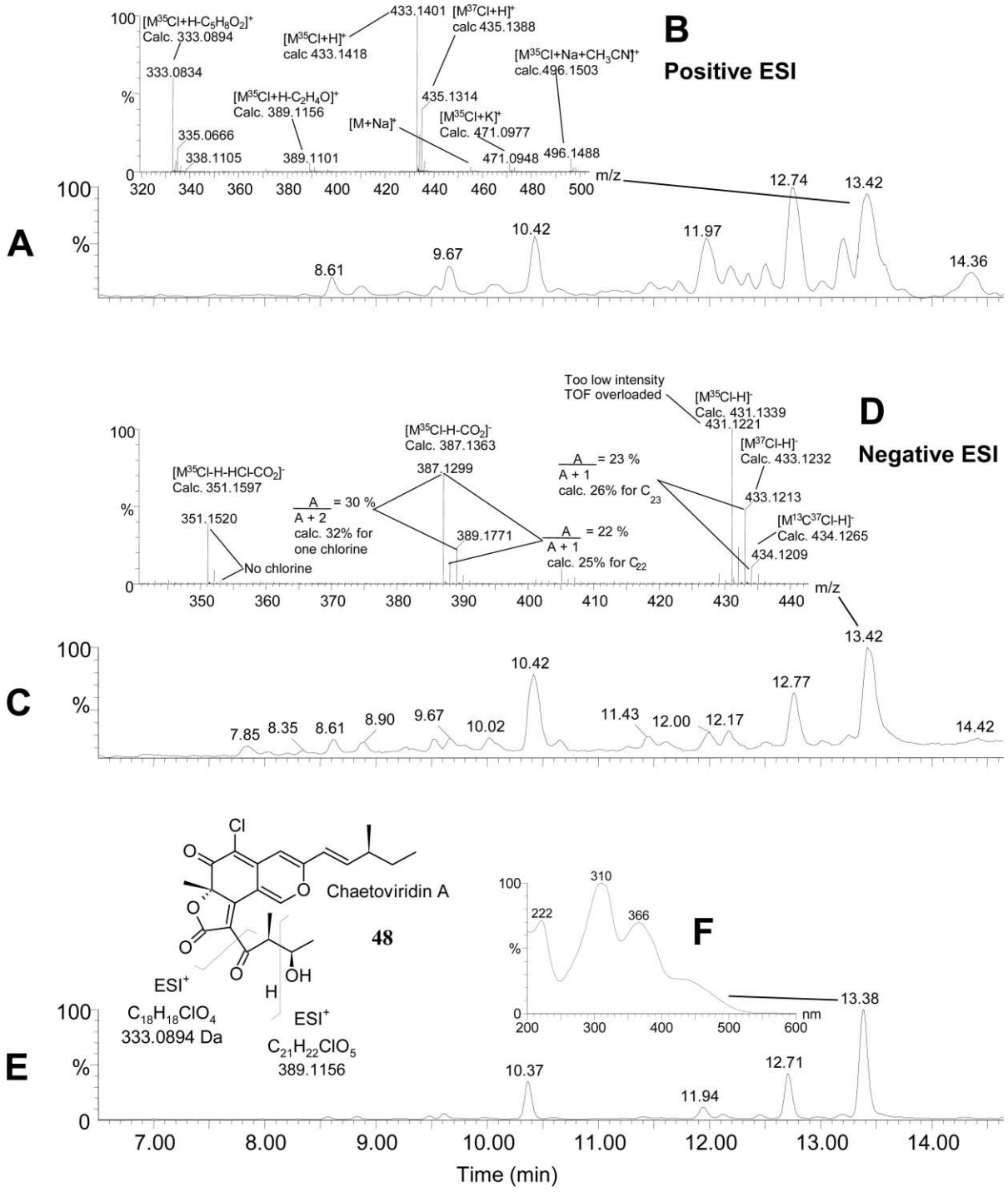


Fig. 9 LC-UV-MS analysis of a *Chaetomium globosum* extract where chaetoviridin A ($C_{23}H_{25}ClO_6$) was tentatively identified. **A** ESI^+ time-of-flight (TOF) MS total ion chromatogram; **B** ESI^+ TOF MS spectrum with annotated peaks; **C** ESI^+ time-of-flight (TOF) MS total ion chromatogram; **D** ESI^- TOF MS spectrum with annotated peaks; **E** UV-VIS chromatogram (200–700 nm) from same file as ESI^-; **F** UV-VIS spectrum of the target compound. Note that there is a time delay of *ca.* 0.04 min from the UV-VIS detector to the MS.

However, usually both ions can be detected. If in doubt it may be necessary to reanalyse the sample using another buffer, *e.g.* substitute formate for acetate or vice versa—or mix both to see two ions with m/z 14 (CH_2) in-between.

The jump from using nominal mass MS to accurate mass MS in dereplication when searching databases is very helpful. *E.g.* if Antibase 2003 (30 000 compounds) is searched for a nominal mass in the range 200 to 500 Da, an average of 78 candidates with even masses is found and 16 with an uneven mass are found. However, if the accurate mass is known, then for the elemental composition usually 3–10 candidates are found for an even mass compound, although there for a few compounds can be up to 40. For odd mass compounds usually 1 to 5 candidates are found.

However, this demands that the accurate mass is correct and on QTOF and TOF instrument with TDC detectors this is not always the case. There are some pitfalls, as also described above, as the number of ions in a single scan should not exceed a certain number (usually 50–100 ions), since the mass determination will else be too low (see also Fig. 8). In such a case the spectra used for calculating the elementary compositions must be taken in the front or tail of the chromatographic peak.[225]

With a mass accuracy of 3 ppm there are usually 2–3 candidates in the 400 to 500 Da region (calculated with max 100 C, 15 O, 8 N, 4 S, 100 H, and 50 DBE), assuming that Cl, Br and >4 S will be seen by the operator. If the mass accuracy is in the range of 3 to 6 ppm 4 to 8 candidates are usually found in the

400 to 500 Da region. To validate the elemental composition or narrow down the number of candidates, the elemental composition should be calculated for all the adducts, dimeric, and fragment ions which can be unambiguously assigned as demonstrated by Nielsen and Smedsgaard.[225] Subsequently or integrated in the latter process isotope ratios of A, A + 1, A + 2, A + 3, *etc.* can be investigated further *e.g.* eliminating sulfur and removing candidates where the A and A + 1 ratios does not support the number of calculated carbon atoms. In this process it is vital to know if there are interfering co-eluting compounds which may affect the measured isotope pattern. It is also vital to know how accurate the MS in use shows the true isotope pattern, which can easily be checked on some known compounds.

The final step in the dereplication process is to search either the candidate(s) from the elementary composition calculations or the molecular mass in appropriate databases, *i.e. Antibase*, *MARINLIT*, *NAPRALERT*, and *Dictionary of Natural Products*. The hit(s) from the database search are then matched against the UV-VIS data from the target peak (see next section), which usually can eliminate many candidates. After that the retention on the chromatographic system can be assessed and usually many very polar/apolar compounds can be eliminated based on such comparisons. The ionization efficiency in positive *versus* negative mode and relative UV-VIS also gives some valid information. *E.g.* amines will always ionize well so if a very small peak is seen in positive mode whereas large peaks are seen in negative and/or UV then the compound is not an alkaloid. Finally the fragments obtained by in-source fragmentation and/or data-depended MS/MS can be evaluated against the remaining structures.

6.4 Chemical image analysis—UV spectral analysis

With the advancement of HPLC and UPLC as well as much more stable and better columns for high resolution separation, combined with fast UV diode array detectors it has become easy to acquire the UV spectrum of practically every single component from an extract. Consequently the UV spectrum has turned into being one of the most readily accessible pieces of information related to structure of NPs why there is increased interest in exploiting its usefulness.[227]

However, the chromatographic data matrix can also be viewed as a landscape with the retention time being the first dimension, the spectrum the second, and finally the detected values the third. These topographic landscapes can be compared automatically using techniques from image analysis. However, it is crucial to correct for different artefacts due to differences in the instrumental acquisition procedure from analysis to analysis. Some of the most important artefacts include baseline drifts and shifts in retention time. Baseline correction is done through subtraction of an estimated baseline, wavelength by wavelength. The baseline drift can be regarded as a nonlinear function, but most often segments are detected from which the baseline is estimated by partially linear interpolation.

In order to compensate for the minor shifts in retention time, it is necessary to align the chromatograms. All methods are more or less derivatives of the same basis; to warp the time axis of a chromatogram in order to obtain the best match to a reference chromatogram. An early attempt of developing a retention time warping algorithm was made by Wang and Isenhour.[228] Since then, improvements have been made based on different optimization criteria.[229–231] But the probably most efficient way to do aligning is through the correlation optimized warping (COW) technique, using all the trace information available.[232,233] Although time consuming, COW has proven to be the method that gives the best results.[234] Whereas the aligning algorithms focus on solving the problem of shifts in the retention time by optimizing the trace profiles, spectral shape information can be added to the different aligning concepts.[235] In addition, the similarity between two HPLC matrices $\mathbf{X}_i$ and $\mathbf{X}_j$ is evaluated in two steps: 1) The similarity between the UV-spectra where $\mathbf{X}_i$ has peaks, and 2) the similarity between the UV-spectra where $\mathbf{X}_j$ has peaks. The peaks were found as the mean absorbance value from each aligned spectrum, followed by a smothering using a simple mean filtering over a window of 9 UV-spectra. This was found to be both fast and sufficient to remove the small spikes that could be present; however, other filtering techniques could be applied if needed.[236] The principle of the method is based on evaluating the spectral information present in the HPLC data matrices, and the algorithm consists of two parts:

1) Calculation of a local similarity, followed by

2) calculation of a global similarity between the HPLC data matrices.

For two chromatographic data matrices a vector of locally aligned full spectral similarities is calculated along the retention time axis. The vector depicts the evaluation of the alikeness between two fungal extracts based upon eluted compounds and corresponding UV-absorbance spectra. By comparing co-eluting components by their UV-spectra across samples, information about the (dis)similarity between actual compounds could be examined. The similarity was evaluated as a "distance" between the observed UV-spectra. Based on these spectra classification of the samples can be made.[204,235]

One of the major advantages of applying the method is that the chemical diversity can be calculated by selecting only a few input parameters involved in the process. 1) Which part of the chromatograms and which spectral range to include, 2) perform a simple baseline correction, 3) align the chromatograms by warping, 4) scale the chromatograms and finally 5) calculate the similarity. Therefore, the method removes the bias from comparisons and makes reproducibility possible between data files made at different periods. Most algorithms used for warping mainly rely on chromatographic traces, and problems may occur when several compounds are eluting within a short period of time. By using the full UV-spectral information in the aligning, the algorithm aligns peaks having the same UV-spectrum.

6.5 Dereplication and partial identification of NPs by UV-based techniques

Apart from exact structural formulae structural databases usually also contain physical chemical data including UV maxima and minima characteristics of the included compounds. Today only few organic chemists rely on UV as a primary tool for structural elucidation, one reason being that the absorption frequencies of the C–C, C–H and isolated C=C groups cannot be observed in the easily accessible region of the UV spectrum. However, many natural products such as polyketides and alkaloids derived from aromatic amino acids have characteristic UV spectra (Fig. 10) due to their polyunsaturated nature. In addition many such NPs often have one or more carbonyl groups as part of ketone, carboxylic acid, ester or amide functional groups.

One early report using a UV based library for dereplication of mycotoxins is the work by Frisvad and Thrane,[237] who used characteristic UV spectra of 187 mycotoxins and other secondary metabolites for chemotaxonomy and food safety studies.

Nielsen and Smedsgaard extended this fungal metabolite database to 474 metabolites now also including MS spectral data.[225] Other research groups have developed similar in-house HPLC-UV-VIS databases for dereplication purposes.[238] Thus Fiedler *et al.*[239] used their database containing approx. 750 UV spectra to investigate a set of 600 marine actinomycete strains for production of both a number of known antibiotics but also as an approach to discover possible new metabolites. Similarly Larsen *et al.* used a knowledge based UV-guided approach for detection and isolation of novel alkaloid compounds from fungal extracts.[240–242] In order to systematise and keep track of data and structural knowledge related to UV-spectra of

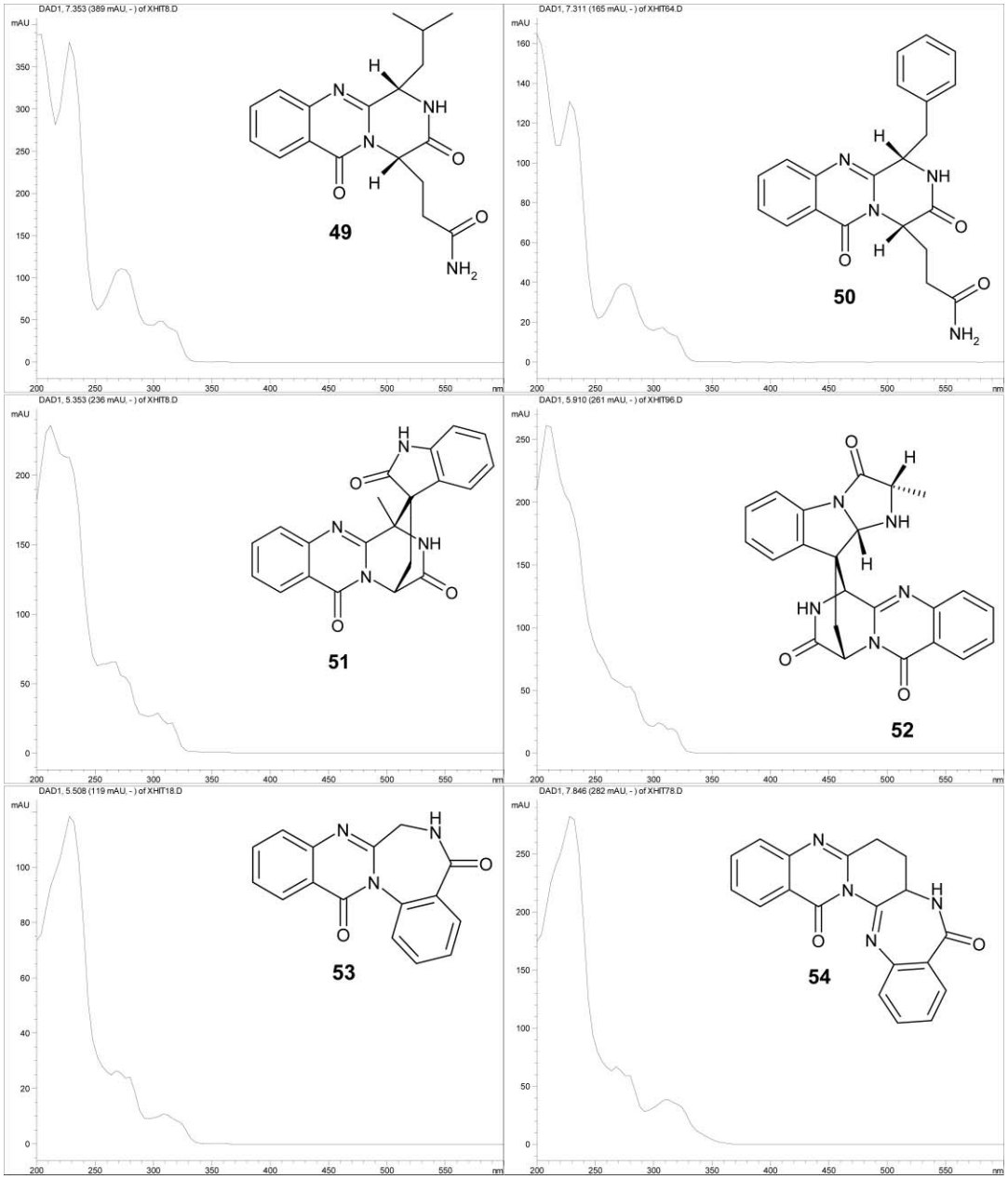


Fig. 10 Six similar but still slightly different and characteristic UV spectra of fungal alkaloids (anacine **49**, verrucine A **50**, alantrypinone **51**, lapatin A **52**, sclerotigenin **53** and auranthine **54**) that can be used for targeting of both known and new (but similar) natural products by X-hitting.[235,243]

both known and possible new fungal metabolites Hansen *et al.* recently developed a new algorithm called *X-hitting* to be used in combination with a learning-based database.[235] In *X-hitting* already known metabolites are tracked by the feature *cross-hitting*, whereas potential new metabolites are indicated by the *new-hitting* feature of *X-hitting*. The algorithm behind *X-hitting* compares the shape between two spectra and returns a similarity index describing how statistically similar the two spectra are. In order to capture information about the actual shape of each of the profiles at different scales (coarser or finer details), a linear combination of the correlation between higher order derivatives are used. A main criterium being retaining relations between "neighboring" values in spectra and in addition, that the algorithm has to be fast to compute. By using derivatives such relationships can be incorporated by measuring the differences and the positions of topologies and extreme points. Finally filtering techniques are applied to the derivatives in order to reduce the sensitivity to potential noise fragments in the spectra. Regarding the drug discovery process the great potential of cross-hitting is its ability to support the choice of the optimal organism for production of a given natural product. This is the case both in de-screening of strains producing mycotoxins, but maybe even more important also for the tracking of new species producing a desired compound (in this scenario a new drug lead). In this way cross-hitting of known compounds can be applied for new-hitting of organisms. Very often a certain compound is produced as a minor metabolite by one species (under certain conditions), while it appears as one of the major metabolites produced by another completely different species.

The scope of the new-hitting feature of "X-hitting" is in scenarios where a new lead such as a quinazoline (anacine **49**, verrucine **50**, alantrypinone **51**, lapatin A **52**) or a benzodiazepine (sclerotigenin **53**, auranthine **54**) (Fig. 10) have already been discovered by bio-guided screening. Due to the very similar structure of these two types of compounds (Fig. 10), that both incorporate anthranilic acid, they have very distinct and similar chromophore systems, and can easily be tracked down from large data sets of HPLC Diode array data by X-hitting.[235] Thus using the cross-hitting feature of X-hitting we can easily detect many of the known quinazolines and benzodiazepines (unpublished data) produced by *Penicillium* and *Aspergillus*.

Recently, Larsen *et al.* demonstrated the potential of *new-hitting* by finding the two novel quinazolines lapatin A (**52**) and B.[243] Isolation of some new benzodiazepines by our group using new-hitting is in progress, where the use of sclerotigenin as target UV-spectrum gives several hits in an extract of *Aspergillus janus*. We predict that new-hitting will be a powerful alternative to finding/generating new leads by traditional organic synthesis or by combinatorial chemistry methodologies.

6.6 Metabolite fingerprinting, profiling and target analysis by NMR

Traditionally NMR has been used for structural elucidation of pure compounds or relatively simple mixtures of compounds generated from *e.g.* organic synthesis. The recent combination of high-performance liquid chromatography interfaced to NMR has eliminated the need to purify a compound before analysis. NMR spectra can either be recorded in continuous or stopped flow. Furthermore the output from an HPLC can be split into two pathways, one for NMR analysis and the other for MS, increasing the sensitivity by using a solid phase extraction (SPE) step between the HPLC column and the NMR spectrometer.[209] Thus several papers report the characterisation of numerous natural products directly from complex mixtures using LC-NMR.[244–247] When combined with computer assisted software programs for structural elucidation,[248,249] the task of identifying both known and new NPs by NMR have surely become a much easier task within the last 5–10 years.

More recently the analysis of complex mixtures and biofluids with the purpose of categorizing large numbers of individual spectra by NMR spectroscopy by computational methods has become an increasing application. This covers areas such as disease diagnosis, quality control of foods such as citrus juices, edible/olive oils, wine/beer.[208,250,251] Even though it is still a challenge, metabolite fingerprinting by NMR is considered a fast, convenient, and effective tool for discriminating between groups of related samples just based on a comparison on the distribution of intensity in the NMR rather than on the assignment of the individual signals.[209]

To our knowledge no attempts have been made to use *e.g.* ^{1}H NMR spectra as part of a microbial drug discovery process using *e.g.* ^{1}H NMR spectra with the purpose of clustering data of microbial extracts into chemotypes in a similar way as direct MS profiling has been used.[207] On the other hand NMR has been extensively used in plant metabolomics where some argue that the metabolome is a fundamentally important biochemical manifestation of the genome, each extract defining a metabolic phenotype that forms the basis for discriminating between plants of different genotypes.[209] These are altogether arguments that are in very good agreement with the major conclusions drawn in this review. In our opinion the potential for using NMR of crude extracts as part of a microbial drug discovery process is largely untapped even though issues such as stringent control of sample preparation, with precise control of the pH, and precise temperature control during acquisition need to be addressed.

7. Conclusions and future perspectives

With the genomic revolution the need for new pharmaceutical leads is enormous. Since there are still many unexplored resources in *Nature*, the potential for finding new organisms and thereby new metabolic pathways is also enormous. Therefore natural products still have a very important role to play in the drug discovery process, especially due to their often very complex and chiral nature. A large proportion of new leads/scaffolds in the future can therefore be expected to be NPs and natural product chemistry should therefore not be considered a competing discipline to combinatorial chemistry. Instead a combination of the two disciplines seems to be the right track to follow in the search for valuable new drugs.

In this review we have argued that microbial organisms, and at least eukaryotic filamentous fungi, are very consistent in their production of secondary metabolites at the species level when cultured under standardized conditions. In other words strains of different species represent different chemotypes, whereas strains of the same species represent the same chemotype. We have shown that these differences and similarities can easily be analyzed for using direct analytical profiling techniques such as DI-HRMS in combination with modern chemoinformatics tools. This might not be the case for other organisms such as *Actinomycetes* where horizontal gene transfer could be a pronounced event leading to potential transferring of biosynthetic pathways. Recent studies on the new marine genus *Salinospora*, however, showed a clear correlation between species and classes of compounds produced, indicating a large degree of chemoconsistency also within these types of organisms.[31] Additional types of organisms, including several genera and species need to be studied in order to prove our hypothesis that secondary metabolites in most cases will be very consistently produced at the species level also in other microorganisms than filamentous fungi.

We have argued that instead of including many identical strains with the same chemotypes in a drug discovery screening process, the effort should instead be addressed on the investigation of much fewer talented strains under multiple culture conditions for optimization of secondary metabolite production in order to fully explore the secondary metabolite potential of a given organism. Also a combined taxonomic and thereby ecological knowledge about a given species will point towards investigation of relevant stimulation studies such as the stimulation of fungal metabolite production by adding plant extracts or by co-cultivation of relevant natural co-existing organisms.

A chemical based drug discovery approach is not only important in the process of dereplicating of culture collections, but also in the dereplication of single compounds using hyphenated analytical techniques interphasing LC with MS, NMR and UV. The importance of high resolution MS in this context is of course substantial, with the possibility of using accurate masses for suggesting molecular formulae that can be searched in various databases. However, the fact that many biological active NPs contain conjugated systems and thereby more or less specific UV-chromophores could also be applied as an important aid not only in the dereplication process both also for the search and discovery of novel bioactives with similar structural features as already known bioactives. The different spectroscopic information that can be gained from MS, NMR and UV complement each other and are all expected to play and important role in future chemoinformatics based drug discovery approaches.

8. Acknowledgements

The authors wish to acknowledge the support from The Danish Research Councils for the *Program for Predictive Biotechnology* (2000–2004) (Grant No. 9901295) and the IVC *Center for Microbial Biotechnology* (2004–2008) (Grant No. 46–00–0005), during which many of the ideas, methodologies, and results presented here were generated.

9. References

1 M. S. Butler, *J. Nat. Prod.*, 2004, **67**, 2141.
2 M. S. Butler, *Nat. Prod. Rep.*, 2005, **22**, 162.
3 J. C. Frisvad and R. A. Samson, *Stud. Mycol.*, 2004, **49**, 1.
4 D. H. Williams, M. J. Stone, P. R. Hauck and S. K. Rahman, *J. Nat. Prod.*, 1989, **52**, 1189.
5 A. L. Demain, *Appl. Microbiol. Biotechnol.*, 1999, **52**, 455.
6 W. R. Strohl, *Drug Discovery Today*, 2000, **5**, 39.
7 G. L. Verdine, *Nature*, 1996, **384s**, 11.
8 L. Müller-Kuhrt, *Nat. Biotechnol.*, 2003, **21**, 602.

9 D. J. Newman, G. M. Cragg and K. M. Snader, *J. Nat. Prod.*, 2003, **66**, 1022.
10 A. Ganesan, *Curr. Opin. Biotechnol.*, 2004, **15**, 584.
11 G. A. Petsko, *Nature*, 1996, **384s**, 7.
12 L. J. Nisbet and M. Moore, *Curr. Opin. Biotechnol.*, 1997, **8**, 708.
13 V. Knight, J. J. Sanglier, D. DiTullio, S. Braccili, P. Bonner, J. Waters, D. Hughes and L. Zhang, *Appl. Microbiol. Biotechnol.*, 2003, **62**, 446.
14 J. W. Costerton, P. S. Stewart and E. P. Greenberg, *Science*, 1999, **284**, 1318.
15 D. G. Davies, M. R. Parsek, J. P. Pearson, B. H. Iglewski, J. W. Costerton and E. P. Greenberg, *Science*, 1998, **280**, 295.
16 T. B. Rasmussen, M. E. Skindersø, T. Bjarnsholt, R. K. Phipps, K. B. Christensen, J. B. Andersen, B. Koch, T. O. Larsen, M. Hentzer, N. Høiby and M. Givskov, *Microbiology*, 2005, **151**, 1325.
17 A. T. Bull, J. E. M. Stach, A. C. Ward and M. Goodfellow, *Antonie van Leeuwenhoek*, 2005, **87**, 65.
18 A. T. Bull, A. C. Ward and M. Goodfellow, *Microbiol. Mol. Biol. Rev.*, 2000, **64**, 573.
19 G. M. Müller, G. F. Bills and M. S. Foster, *Biodiversity of Fungi. Inventory and Monitoring Methods*, Elsevier Academic Press, Amsterdam, 2004, 777 pp.
20 M. Jaspars and L. Lawton, *Curr. Opin. Drug Discovery Dev.*, 1998, **1**, 77.
21 J. Blunt, B. R. Copp, M. H. G. Munro, P. T. Northcote and M. R. Prinsep, *Nat. Prod. Rep.*, 2005, **22**, 15.
22 J. C. Frisvad, in *Innovative Roles of Biological Resource Centers*, ed. M. M. Watanabe, K. Suzuki and T. Seki, Japan Society for Culture Collections & World Federation for Culture Collections, Tsukuba, 2004, p. 165.
23 J. C. Frisvad, T. O. Larsen, P. W. Dalsgaard, K. A. Seifert, G. Louis-Seize, E. K. Lyhne, B. B. Jarvis, J. C. Fettinger and D. P. Overy, *Int. J. Syst. Evol. Microbiol.*, 2005, in press.
24 P. W. Dalsgaard, T. O. Larsen and C. Christophersen, *J. Nat. Prod.*, 2004, **67**, 878.
25 P. W. Dalsgaard, J. W. Blunt, M. H. G. Munro, T. O. Larsen and C. Christophersen, *J. Nat. Prod.*, 2004, **67**, 1950.
26 P. W. Dalsgaard, T. O. Larsen and C. Christophersen, *J. Antibiot.*, 2005, **58**, 141.
27 O. Filtenborg, J. C. Frisvad and U. Thrane, *Int. J. Food Microbiol.*, 1996, **33**, 85.
28 B. Haefner, *Drug Discovery Today*, 2003, **8**, 536.
29 J. E. M. Stach and A. T. Bull, *Antonie van Leeuwenhoek*, 2005, **87**, 3.
30 U. Höller, A. D. Wright, G. F. Matthée, G. M. Konig, S. Draeger, H.-J. Aust and B. Schulz, *Mycol. Res.*, 2000, **104**, 1354.
31 P. R. Jensen, T. J. Mincer, P. G. Willaims and W. Fenical, *Antonie van Leeuwenhoek*, 2005, **87**, 43.
32 C. E. Salomon, N. A. Magarey and D. H. Sherman, *Nat. Prod. Rep.*, 2004, **21**, 105.
33 E. W. Schmidt, J. T. Nelson, D. A. Ratsko, S. Sudek, J. A. Eisen, M. G. Haygood and J. Ravel, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 7315.
34 P. F. Long, W. C. Dunlap, C. N. Battershill and M. Jaspars, *ChemBioChem*, 2005, **6**, 1760.
35 D. L. Hawksworth, *Mycol. Res.*, 1991, **95**, 641.
36 D. L. Hawksworth, *Stud. Mycol.*, 2004, **50**, 9.
37 D. L. Hawksworth and A. Y. Rossman, *Phytopathology*, 1997, **87**, 888.
38 B. Schulz, C. Boyle, S. Draeger, A.-K. Römmert and K. Krohn, *Mycol. Res.*, 2002, **106**, 996.
39 A. Stierle, G. Stobel and D. Stierle, *Science*, 1993, **260**, 214.
40 A. Stierle and G. Stobel, *J. Nat. Prod.*, 1995, **58**, 1315.
41 R. Verpoorte, *Drug Discovery Today*, 1998, **3**, 232–238.
42 P. F. Dowd, J. D. Miller and R. Greenhalgh, *Mycologia*, 1989, **81**, 646.
43 P. F. Dowd, in *Handbook of Applied Mycology, Vol. 5: Mycotoxins in Ecological Systems*, ed. D. Bhatnagar, E. B. Lillehøj and D. J. Arora, Marcel Dekker, New York, 1992, p. 333.
44 D. T. Wicklow, in: *Coevolution of Fungi with Plants and Animals*, ed. K. A. Pirozynski and D. L. Hawksworth, Academic Press, London, 1998, p. 173.
45 J. Gloer, *Can. J. Bot.*, 1995s1, **73**, 1265.
46 Y. Zhang, C. Li, D. C. Swenson, J. B. Gloer, D. T. Wicklow and P. F. Dowd, *Org. Lett.*, 2003, **5**, 773.
47 C. Li, J. B. Gloer, D. T. Wicklow and P. F. Dowd, *J. Nat. Prod.*, 2005, **68**, 319.
48 D. P. Overy and J. C. Frisvad, *Syst. Appl. Microbiol.*, 2003, **26**, 631.
49 M. Boysen, P. Skouboe, J. C. Frisvad and L. Rossen, *Microbiology*, 1996, **142**, 541.
50 K. Karlshøj and T. O. Larsen, *J. Agric. Food Chem.*, 2005, **53**, 708.
51 K. F. Nielsen, P. W. Dalsgaard, J. Smedsgaard and T. O. Larsen, *J. Agric. Food Chem.*, 2005, **53**, 2908.
52 J. Smedsgaard and J. Nielsen, *J. Exp. Bot.*, 2005, **56**, 273.
53 K. Zengler, G. Toledo, M. Rappe, J. Elkins, E. J. Mathur, J. M. Short and M. Keller, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 15681.
54 M. S. Osburne, T. H. Grossman, P. R. August and I. A. MacNeil, *Am. Soc. Microbiol. News*, 2000, **7**, 411.
55 J. Handelsman, M. R. Rondon, S. F. Brady, J. Clardy and R. M. Goodman, *Chem. Biol.*, 1998, **5**, 245.
56 D. Cowan, Q. Meyer, W. Stafford, S. Muyanga, R. Cameron and P. Wittwer, *Trends Biotechnol.*, **23**, 321.
57 N. Peric-Concha and P. F. Long, *Drug Discovery Today*, 2003, **8**, 1078.
58 J. Kennedy and C. R. Hutchinson, *Nat. Biotechnol.*, 1999, **17**, 538.
59 B. S. Moore, J. A. Kalaitzis and L. Xiang, *Antonie van Leeuwenhoek*, 2005, **87**, 49.
60 J. Staunton and K. J. Weissman, *Nat. Prod. Rep.*, 2001, **18**, 380.
61 B. A. Pfeifer and C. Khosla, *Microbiol. Mol. Biol. Rev.*, 2001, **65**, 106.
62 J. Nielsen, *Curr. Opin. Microbiol.*, 1998, **1**, 330.
63 J. Nielsen, *Appl. Microbiol. Biotechnol.*, 2001, **55**, 263.
64 S. Kosemura, *Tetrahedron*, 2003, **59**, 5055.
65 S. A. Waksman and E. Bugie, *Proc. Natl. Acad. Sci. U. S. A.*, 1943, **29**, 282.
66 R. L. Monaghan, J. D. Polishook, V. J. Pecore, G. F. Bills, M. Nallin-Olmstead and S. L. Streicher, *Can. J. Bot.*, 1995, **73s1**, 925, ISBN 87-88584-44-5.
67 A. Ciegler, D. I. Fennell, D. A. Sansing, R. W. Detroy and G. A. Bennett, *Appl. Microbiol.*, 1973, **26**, 271.
68 G. Engel, K. E. von Milczewskii, D. Prokopek and M. Teuber, *Appl. Environ. Microbiol.*, 1982, **43**, 1034.
69 J. C. Frisvad, *Dr. Tech. Dissertation*, Technical University of Denmark, 1998.
70 M. Wink, *Phytochemistry*, 2003, **64**, 3.
71 H. T. Lumbsch, in *Chemical Fungal Taxonomy*, ed. J. C. Frisvad, P. D. Bridge and D. K. Arora, Marcel Dekker, New York, 1998, p. 345.
72 J. C. Frisvad, U. Thrane and O. Filtenborg, in *Chemical Fungal Taxonomy*, ed. J. C. Frisvad, J. C., P. D. Bridge and D. K. Arora, Marcel Dekker, New York, 1998, p. 289.
73 J. C. Frisvad, J. M. Frank, J. A. M. Houbraken, A. F. A. Kuijpers and R. A. Samson, *Stud. Mycol.*, 2004, **50**, 23.
74 R. A. Samson, J. A. M. Houbraken, A. F. A. Kuijpers, J. M. Frank and J. C. Frisvad, *Stud. Mycol.*, 2004, **50**, 4561.
75 J. C. Frisvad, J. Smedsgard, T. O. Larsen and R. A. Samson, *Stud. Mycol.*, 2004, **49**, 201.
76 V. Hellwig, Y.-M. Ju, J. D. Rogers, Fournier and M. Stadler, *Mycol. Prog.*, 2005, **4**, 39.
77 C. Thom, *The Penicillia*, Williams and Wilkins, Baltimore, MD, 1930.
78 K. B. Raper and C. Thom, *A Manual of the Penicillia*, Williams and Wilkins, Baltimore, MD, 1949.
79 K. B. Raper and D. I. Fennell, *The genus Aspergillus*, Williams and Wilkins, Baltimore, MD, 1965.
80 J. C. Frisvad, *Arch. Environ. Contam. Toxicol.*, 1989, **18**, 331.
81 R. D. Hutchinson, P. S. Steyn and S. J. van Rensburg, *Toxicol. Appl. Pharmacol.*, 1973, **24**, 507.
82 A. E. de Jesus, W. E. Hull, P. S. Steyn, F. R. van Heerden and R. Vleggaar, *J. Chem. Soc., Chem. Commun.*, 1982, **16**, 902.
83 J. C. Frisvad and O. Filtenborg, *Mycologia*, 1989, **81**, 836.
84 J. I. Pitt, R. A. Samson and J. C. Frisvad, in *Integration of Modern Taxonomic Methods for Penicillium and Aspergillus Classification*, ed. R. A. Samson, and J. I. Pitt, Harwood Academic Publishers, Amsterdam, 2000, p. 9.
85 J. C. Frisvad, O. Filtenborg, R. A. Samson and A. C. Stolk, *Antonie van Leeuwenhoek*, 1990, **57**, 179.
86 J. C. Frisvad and O. Filtenborg, in *Modern Concepts in Penicillium and Aspergillus Classification*, ed. R. A. Samson, and J. I. Pitt, Plenum Press, New York, 1990, p. 373.
87 J. C. Frisvad and O. Filtenborg, in *Modern Concepts in Penicillium and Aspergillus Classification*, ed. R. A. Samson, and J. I. Pitt, Plenum Press, New York, 1990, p. 159.
88 B. Andersen, J. Smedsgaard and J. C. Frisvad, *J. Agric. Food Chem.*, 2004, **52**, 2421.
89 S. Sonjak, J. C. Frisvad and N. Gunde-Cimerman, *FEMS Microbiol. Ecol.*, 2005, **53**, 51.
90 J. F. Makins, G. Holt and K. D. MacDonald, *J. Gen. Microbiol.*, 1983, **116**, 3027.
91 S. A. Waksman and E. S. Hornung, *Mycologia*, 1943, **35**, 47–65.
92 S. A. Waksman and A. Schatz, *Proc. Natl. Acad. Sci. U. S. A.*, 1943, **29**, 74.

93 S. A. Waksman, E. S. Hornung and E. L. Spencer, *J. Bacteriol.*, 1943, **45**, 47.
94 O. Filtenborg, J. C. Frisvad and U. Thrane, in *Modern Concepts in Penicillium and Aspergillus Classification*, ed. R. A. Samson and J. I. Pitt, Plenum Press, New York, 1990, p. 433.
95 K. B. Raper, D. F. Alexander and R. D. Coghill, *J. Bacteriol.*, 1944, **48**, 639.
96 S. P. Brundidge, F. C. A. Gaeta, D. J. Hook, C. Jr. Sapino, R. P. Elander and R. B. Morin, *J. Antibiot.*, 1980, **33**, 1348.
97 S. J. Andersen and J. C. Frisvad, *Lett. Appl. Microbiol.*, 1994, **19**, 486.
98 E. L. Dulaney, *Mycologia*, 1947, **39**, 570.
99 D. S. Cole, G. Holt and K. D. MacDonald, *J. Gen. Microbiol.*, 1976, **96**, 423.
100 A. P. MacCabe, M. B. R. Riach, S. E. Unkles and J. R. Kinghorn, *EMBO J.*, 1990, **9**, 279.
101 P. W. Clutterbuck, A. E. Oxford, H. Raistrick and G. Smith, *Biochem. J.*, 1932, **24**, 1441.
102 J. Jacquet and A. Tantaoui-Eleraki, *Ann. Nutr. Aliment.*, 1977, **31**, 563.
103 P.-K. Chang, D. Bhatnagar, T. E. Cleveland and J. W. Bennett, *Appl. Environ. Microbiol.*, 1995, **61**, 40.
104 M. A. Klich, P.-K. Chang, E. J. Mullaney, D. Bhatnagar and T. E. Cleveland, *Appl. Microbiol. Biotechnol.*, 1995, **44**, 439.
105 K. Kusumoto, K. Mori, H. Nogata, H. Ohta and M. Manabe, *J. Ferment. Bioeng.*, 1996, **82**, 161.
106 K. Kusumoto, K. Yabe, Y. Nagata and H. Ohta, *FEMS Microbiol. Lett.*, 1998, **169**, 303.
107 K. Kusumoto, Y. Nogata and H. Ohta, *Curr. Genet.*, 2000, **37**, 104.
108 A. J. Watson, L. J. Fuller, D. J. Jeens and D. B. Archer, *Appl. Environ. Microbiol.*, 1999, **65**, 307.
109 K. Matsushima, P.-K. Chang, J. Yu, D. Bhatnagar and T. E. Cleveland, *Appl. Microbiol. Biotechnol.*, 2001, **55**, 585.
110 K. Matsushima, K. Yashiro, Y. Hanya, K. Abe, K. Yabe and T. Hamasaki, *Appl. Microbiol. Biotechnol.*, 2001, **55**, 771.
111 T. Takahashi, P.-K. Chang, K. Matsushima, J. Yu, K. Abe, D. Bhatnagar, T. E. Cleveland and Y. Koyama, *Appl. Environ. Microbiol.*, 2002, **68**, 3737.
112 F. Lund and J. C. Frisvad, *Mycol. Res.*, 1994, **98**, 481.
113 F. Blank, W. C. Day and G. Just, *J. Invest. Dermatol.*, 1963, **40**, 133.
114 J. G. Wirth, T. E. Beesley and S. R. Anand, *Phytochemistry*, 1965, **4**, 505.
115 A. S. Ng, G. Just and F. Blank, *Can. J. Chem.*, 1969, **47**, 1223.
116 D. A. Geiser, J. C. Frisvad and J. W. Taylor, *Mycologia*, 1998, **90**, 832.
117 K. Furuya, R. Emokita and M. Shirasaka, *Ann. Sankyo Res. Lab.*, 1967, **19**, 91.
118 D. Broadbent, H. G. Hemming and G. McGowern, *Trans. Br. Mycol. Soc.*, 1974, **62**, 625.
119 B. B. Jarvis, W. G. Sorenson, E.-L. Hintikka, M. Nikulin, Y. Zhou, J. Jiang, S. Wang, S. F. Hinkley, R. A. Etzel and D. G. Dearbom, *Appl. Environ. Microbiol.*, 1998, **64**, 3620.
120 M. Namikoshi, H. Kobayashi, T. Yoshimoto, S. Meguro and K. Akano, *Chem. Pharm. Bull.*, 2000, **48**, 1452.
121 J. C. Frisvad and R. A. Samson, *Stud. Mycol.*, 2000, **45**, 201.
122 L. Wang, H.-B. Zhou, J. C. Frisvad and R. A. Samson, *Antonie van Leeuwenhoek*, 2004, **86**, 173.
123 A. E. Oxford, H. Raistrick and P. Simonart, *Biochem. J.*, 1939, **33**, 240.
124 E. G. Jeffreys, P. W. Brian, H. G. Hemming and D. Lowe, *J. Gen. Microbiol.*, 1953, **9**, 314.
125 P. W. Brian, P. J. Curtis and H. G. Hemming, *Trans. Br. Mycol. Soc.*, 1949, **32**, 30.
126 S. M. Clarke and M. McKenzie, *Nature*, 1967, **213**, 504.
127 P. W. Brian, P. J. Curtis and H. G. Hemming, *Trans. Br. Mycol. Soc.*, 1955, **38**, 305.
128 L. Leistner and C. Eckardt, *Fleischwirtschaft*, 1979, **59**, 1892.
129 M. Christensen, J. C. Frisvad and D. I. Tuthill, *Mycol. Res.*, 1998, **103**, 527.
130 J. MacMillan, *Chem. Ind. (London)*, 1951, 719.
131 K. Udagawa and S. Abe, *J. Antibiot.*, 1961, **14**, 215.
132 D. Kingston, P. Chen and J. R. Vercellotti, *Phytochemistry*, 1976, **15**, 1037.
133 P. W. Brian, *Trans. Br. Mycol. Soc.*, 1960, **43**, 1.
134 B. Andersen, K. F. Nielsen and B. B. Jarvis, *Mycologia*, 2002, **94**, 392.
135 B. Andersen, K. F. Nielsen, U. Thrane, T. Szaro, J. W. Taylor and B. B. Jarvis, *Mycologia*, 2003, **95**, 1227.
136 R. Bentley, *Chem. Rev.*, 2000, **100**, 3801.
137 R. A. Samson, K. A. Seifert, A. F. A. Kuijpers, J. A. M. Houbraken and J. C. Frisvad, *Stud. Mycol.*, 2004, **49**, 175.
138 P. Lafont, J.-P. Debeaupuis, M. Gaillardin and J. Payen, *Appl. Environ. Microbiol.*, 1979, **37**, 365.
139 O. Puel, S. Tadrist, P. Galtier, I. P. Oswald and M. Delaforgem, *Appl. Environ. Microbiol.*, 2005, **71**, 550.
140 R. Gosio, *Rivista Igiene Sanita Publica Ann.*, 1896, **7**, 825.
141 C. L. Alsberg and O. F. Black, *U.S. Dept. Agric. Bur. Plant Industry Bull.*, 1913, **270**, 1.
142 A. E. Oxford and H. Raistrick, *Biochem. J.*, 1933, **27**, 1176.
143 A. E. Oxford and H. Raistrick, *Biochem. J.*, 1933, **27**, 1473.
144 J. C. Frisvad and O. Filtenborg, *Appl. Environ. Microbiol.*, 1983, **46**, 1301.
145 M. Devys, J. F. Bousquet, A. Kollman and M. Barbier, *Phytochemistry*, 1980, **19**, 2221.
146 N. G. Vinokurova, N. E. Ivanushkina, G. A. Kochina, M. U. Arinbasarov and S. M. Ozerskaya, *Appl. Biochem. Microbiol.*, 2005, **41**, 83.
147 Y. Fujimoto, M. Kamiya, H. Tsunoda, K. Ohtsubo and T. Tatsuno, *Chem. Pharm. Bull.*, 1980, **28**, 1062.
148 M. Umeda, T. Yamashita, M. Saito, S. Sekita, C. Takahashi, K. Yoshihira, S. Natori, H. Kurata and S. Udagawa, *Jpn. J. Exp. Med.*, 1974, **44**, 83.
149 H. S. Burton, *Br. J. Exp. Pathol.*, 1949, **30**, 151–158.
150 T. O. Larsen and J. C. Frisvad, *Mycol. Res.*, 1995, **99**, 1153.
151 M. Umeda, K. Ohtsubo, M. Saito, S. Sekita, K. Yoshihira, S. Natori, S. Udagawa, F. Sakabe and H. Kurata, *Experientia*, 1975, **15**, 435.
152 J. P. Springer, J. Clardy, J. M. Wells, R. J. Cole, J. W. Kirksey, R. D. Macfarlane and D. F. Torgerson, *Tetrahedron Lett.*, 1976, **17**, 1355.
153 D. Veselý, D. Veselá and R. Jelínek, *Mycopathologia*, 1995, **132**, 31.
154 J. Namgoong, S. Yeon, N. Paak, Y. Kim, C. Kim and K. Kim, *Sanop Misaengmul Hakhoechi*, 1998, **26**, 137.
155 A. Tikoo, H. Cutler, S. H. Lo, L. B. Chen and H. Maruta, *Cancer J.*, 1999, **5**, 293.
156 A. Tikoo, R. Shakri, L. Conolly, Y. Hirokawa, T. Shishido, B. Bowers, L. Ye, K. Kohama, R. J. Simpson and H. Maruta, *Cancer J.*, 2000, **6**, 162.
157 A. Ichihara, K. Katayama, H. Teshima, H. Oikawa and S. Sakamura, *Biosci., Biotechnol., Biochem.*, 1996, **60**, 360.
158 C. Von Wallbrunn, H. Luftmann, K. Bergander and F. Meinhardt, *J. Gen. Appl. Microbiol.*, 2001, **47**, 33.
159 Y. Amemiya, A. Kondo, K. Hirano, T. Hirukawa and T. Kato, *Chiba Daigaku Engeigakubu Gakujutsu Hokoku*, 1994, **48**, 13.
160 J. G. Kang, K. K. Kim and K. Y. Kang, *Agric. Chem. Biotechnol. (Engl. Ed.)*, 1999, **42**, 146.
161 P. H. A. Sneath and R. R. Sokal, in *Numerical taxonomy*, W. H. Freeman and Co., San Francisco, 1973.
162 R. R. M. Paterson, S. M. J. Simmonds and W. M. Blaney, *J. Invertebr. Path.*, 1987, **50**, 124.
163 B. A. Bird, A. T. Remaley and I. M. Campbell, *Appl. Environ. Microbiol.*, 1981, **42**, 521.
164 B. A. Bird and I. M. Campbell, *Appl. Environ. Microbiol.*, 1982, **43**, 345.
165 C. Young, L. McMillan, E. Telfer and B. Scott, *Mol. Microbiol.*, 2001, **39**, 754.
166 N. P. Keller and T. M. Hohn, *Fungal Genet. Biol.*, 1997, **21**, 17.
167 S. Udagawa, T. Muroi, H. Kurata, S. Sekita, K. Yoshihira, S. Natori and M. Umeda, *Can. J. Microbiol.*, 1979, **26**, 170.
168 S. Sekita, K. Yoshihira, S. Natori, S. Udagawa, F. Sakabe, H. Kurata and M. Umeda, *Chem. Pharm. Bull.*, 1982, **30**, 1609.
169 S. Sekita, K. Yoshihira, S. Natori and H. Kuwano, *Chem. Pharm. Bull.*, 1982, **30**, 1618.
170 S. Sekita, K. Yoshihira, S. Natori, S. Udagawa, T. Muroi, Y. Sugiyama, H. Kurata and M. Umeda, *Can. J. Microbiol.*, 1981, **27**, 766.
171 H. Oikawa, Y. Murakami and A. Ichihara, *Tetrahedron Lett.*, 1991, **32**, 4533.
172 K. Koyoma, K. Takahashi, S. Natori and S. Udagawa, *Proc. Jpn. Assoc. Mycotoxicol.*, 1991, **33**, 40.
173 C. Spöndlin and C. Tamm, *Helv. Chim. Acta*, 1988, **71**, 1881.
174 R. Donoso, A. Rivera-Sagredo, J. A. Hueso-Rodriguez and S. W. Elson, *Nat. Prod. Lett.*, 1997, **10**, 49.
175 O. Convert, A. Jellal, I. Correira, F. Dardoize, L. Menguy and J. C. Cherton, *Analysis*, 1994, **22**, 217.
176 J. C. Frisvad, R. A. Samson, B. Rassing, M. I. van der Horst, F. T. J. van Rijn and J. Stark, *Antonie van Leeuwenhoek*, 1997, **72**, 119.
177 T. O. Larsen, J. C. Frisvad, G. Ravn and T. Skaaning, *Food Addit. Contam.*, 1998, **15**, 671.
178 A. Numata, C. Takahashi, Y. Ito, K. Minoura, T. Yamada, C. Matsusa and K. Nomoto, *J. Chem. Soc.,Perkin Trans. 1*, 1995, 239.
179 C. Iwamoto, T. Yamada, Y. Ito, K. Monoura and A. Numata, *Tetrahedron*, 2001, **57**, 2997.
180 G. F. Bills, *Can. J. Bot.*, 1995, **73s1**, 33.

181 K. F. Nielsen, T. O. Larsen and J. C. Frisvad, *J. Antibiot.*, 2004, **57**, 29.
182 A. L. Demain, *Int. Microbiol.*, 1998, **1**, 259.
183 H. B. Bode, B. Bethe, R. Höfs and A. Zeeck, *ChemBioChem*, 2002, **3**, 619.
184 K. F. Nielsen, J. Smedsgaard, T. O. Larsen, F. Lund, U. Thrane and J. C. Frisvad, in *Fungal Biotechnology in Agricultural, Food, and Environmental Applications*, ed. D. K. Arora, New York, Marcel Dekker, 2003, p. 19.
185 J. Jennessen, K. F. Nielsen, J. Houbraken, J. Schnürer, E. K. Lyhne, J. C. Frisvad and R. A. Samson, *J. Agric. Food Chem.*, 2005, **53**, 1833.
186 D. Wild, G. Toth and H. U. Humpf, *J. Agric. Food Chem.*, 2002, **50**, 3999.
187 B. B. Jarvis, in *Mycotoxins and Phytoalexins*, ed. R. P. Sharma and D. K. Salunkhe, CRC Press, Boca Raton, 1990, 361.
188 C. W. Hesseltine, *Process Biochem. (Oxford, U. K.)*, 1977, **12**, 29.
189 C. W. Hesseltine, *Process Biochem. (Oxford, U. K.)*, 1997, **32**, 24.
190 C. W. Hesseltine, *J. Ind. Microbiol. Biotechnol.*, 1999, **22**, 482.
191 D. P. Overy and J. W. Blunt, *J. Nat. Prod.*, 2004, **67**, 1850.
192 D. P. Overy, C. H. Zidorn, B. O. Petersen, J. Ø. Duus, P. W. Dalsgaard, T. O. Larsen and R. K. Phipps, *Tetrahedron Lett.*, 2005, **46**, 3225.
193 D. R. Lauren, A. Ashley, B. A. Blackwell, R. Greenhalgh, J. D. Miller and G. A. Neish, *J. Agric. Food Chem.*, 1987, **35**, 884.
194 G. A. Cordell, *Phytochemistry*, 2000, **55**, 463.
195 J. Bérdy, *J. Antibiot.*, 2005, **58**, 1.
196 M. E. Hansen, F. Lund and J. M. Carstensen, *J. Microbiol. Methods*, 2003, **52**, 221.
197 R. D. Firn and C. G. Jones, *Mol. Microbiol.*, 2000, **37**, 989.
198 S. A. S. Mapari, K. F. Nielsen, T. O. Larsen, J. C. Frisvad, A. S. Meyer and U. Thrane, *Curr. Opin. Biotechnol.*, 2005, **16**, 1.
199 G. A. Cordell and Y. G. Shin, *Pure Appl. Chem.*, 1999, **71**, 1089.
200 T. Dörge, J. M. Carstensen and J. C. Frisvad, *J. Microbiol. Methods*, 2000, **41**, 121.
201 M. E. Hansen and J. M. Carstensen, *Pattern Recognition*, 2004, **37**, 2155.
202 M. Sonka, V. Hlavac and R. Boyle, in *Image Processing, Analysis and Machine Vision*, 2nd ed., Chapman & Hall, California, 1998.
203 M. S. Landy and N. Graham, *The Visual Neurosciences*, 2004, 1106.
204 B. Andersen, M. E. Hansen and J. Smedsgaard, *Phytopathology*, 2005, **95**, 1021.
205 O. Fiehn, *Comp. Funct. Genomics*, 2001, **2**, 155.
206 L. W. Sumner, P. Mendes and R. A. Dixon, *Phytochemistry*, 2003, **62**, 817.
207 J. Smedsgaard and J. C. Frisvad, *J. Microbiol. Methods*, 1996, **25**, 5.
208 G. P. Pierens, M. E. Palframan, C. J. Tranter and A. R. Carrol, *Magn. Reson. Chem.*, 2005, **43**, 359.
209 P. Krishnan, N. J. Kruger and R. G. Ratcliffe, *J. Exp. Bot.*, 2005, **56**, 255.
210 J. Smedsgaard, *J. Chromatogr., A*, 1997, **760**, 264.
211 O. Filtenborg and J. C. Frisvad, *Lebensm. Wiss. Technol.*, 1980, **13**, 128.
212 O. Filtenborg, J. C. Frisvad and A. Svendsen, *Appl. Environ. Microbiol.*, 1983, **45**, 581.
213 K. Singh, J. C. Frisvad, U. Thrane and S. B. Mathur, in *An Illustrated Manual on Identification of Some Seed-borne Aspergilli, Fusaria, Penicillia and their Mycotoxins*, Danish Government Institute of Seed Pathology for The Developing Countries, Denmark, 1991.
214 R. K. Julian, Jr., R. E. Higgs, J. D. Gygi and M. D. Hilton, *Anal. Chem.*, 1998, **70**, 3249.
215 R. E. Higgs, J. A. Zahn, J. D. Gygi and M. D. Hilton, *Appl. Environ. Microbiol.*, 2001, **67**, 371.
216 J. Smedsgaard, *Biochem. Syst. Ecol.*, 1997, **25**, 65.
217 M. E. Hansen and J. Smedsgaard, *J. Am. Soc. Mass Spectrom.*, 2004, **15**, 1173.
218 J. Smedsgaard, M. E. Hansen and J. C. Frisvad, *Stud. Mycol.*, 2004, **49**, 243.
219 J. Smedsgaard and J. C. Frisvad, *Biochem. Syst. Ecol.*, 1997, **25**, 51.
220 H. L. Constant and C. W. Beecher, *Nat. Prod. Lett.*, 1995, **6**, 193.
221 D. G. Corley and R. C. Durley, *J. Nat. Prod.*, 1994, **57**, 1484.
222 N. Shingematsu, *J. Mass Spectrom. Soc. Jpn.*, 1997, **45**, 295.
223 P. Waridel, K. Ndjoko, K. R. Hobby, H. J. Major and K. Hostettmann, *Analysis*, 2000, **28**, 895.
224 G. R. Eldridge, H. C. Vervoort, C. M. Lee, P. A. Cremin, C. T. Williams, S. M. Hart, M. G. Goering, M. O'Niel-Johnson and L. Zeng, *Anal. Chem.*, 2002, **74**, 3963.
225 K. F. Nielsen and J. Smedsgaard, *J. Chromatogr., A*, 2003, **1002**, 111.
226 A. Fredenhagen, C. Derrien and E. Gassmann, *J. Nat. Prod.*, 2005, **68**, 385.
227 F. VanMiddlesworth and R. J. P. Cannall, in *Methods in Biotechnology, Vol 4: Natural Products Isolation*, ed. R. J. P. Cannall, Humana Press Inc., Totowa, NJ, 1998, p. 1.
228 C. P. Wang and T. L. Isenhour, *Anal. Chem.*, 1987, **59**, 649.
229 A. J. Round, M. I. Aguilar and T. W. Hearn, *J. Chromatogr., A*, 1994, **661**, 61.
230 D. Bylund, R. Danielsson, G. Malmquist and K. E. Markides, *J. Chromatogr., A.*, 2002, **961**, 237.
231 G. Malmquist and R. Danielsson, *J. Chromatogr., A.*, 1994, **687**, 81.
232 N.-P. V. Nielsen, J. M. Carstensen and J. Smedsgaard, *J. Chromatogr., A.*, 1998, **805**, 17.
233 N.-P. V. Nielsen, J. Smedsgaard and J. C. Frisvad, *Anal. Chem.*, 1999, **71**, 727.
234 V. Pravdova, B. Walczak and D. L. Massart, *Anal. Chim. Acta*, 2002, **456**, 77.
235 M. E. Hansen, J. Smedsgaard and T. O. Larsen, *Anal. Chem.*, 2005, **77**, 6805.
236 P. H. C. Eilers, *Anal. Chem.*, 2003, **75**, 3631.
237 J. C. Frisvad and U. Thrane, *J. Chromatogr.*, 1987, **404**, 195.
238 H.-P. Fiedler, *Nat. Prod. Lett.*, 1993, **2**, 119.
239 H.-P. Fiedler, C. Bruntner, A. T. Bull, A. C. Ward, M. Goodfellow, O. Potterat, C. Puder and G. Mihm, *Antonie van Leeuwenhoek*, 2005, **87**, 37.
240 T. O. Larsen, K. Frydenvang, J. C. Frisvad and C. Christophersen, *J. Nat. Prod.*, 1998, **61**, 1154.
241 T. O. Larsen, H. Franzyk and S. R. Jensen, *J. Nat. Prod.*, 1999, **62**, 1578.
242 T. O. Larsen, K. Frydenvang and J. C. Frisvad, *Biochem. Syst. Ecol.*, 2000, **28**, 881.
243 T. O. Larsen, B. O. Petersen, D. Sørensen, J. Ø. Duus, J. C. Frisvad and M. E. Hansen, *J. Nat. Prod.*, 2005, **68**, 871.
244 S. C. Bobzin, S. Yang and T. P. Kasten, *J. Chromatogr., B*, **748**, 259.
245 O. Corcoran and M. Spraul, *Drug Discovery Today*, 2003, **8**, 624.
246 S. H. Hansen, A. G. Jensen, C. Cornett, I. Bjørnsdottir, S. Taylor, B. Wright and I. D. Wilson, *Anal. Chem.*, 1999, **71**, 5235.
247 L. Wolfender, S. Rodriguez and K. Hostettman, *J. Chromatogr., A*, 1998, **794**, 299.
248 C. Steinbeck, *Nat. Prod. Rep.*, 2004, **21**, 512.
249 M. E. Elyashberg, K. A. Blinov, A. J. Williams, E. R. Martirosian and S. G. Molodtsov, *J. Nat. Prod.*, 2002, **65**, 693–703.
250 W. El-Deredy, *NMR Biomed.*, 1997, **10**, 99.
251 J. C. Lindon, E. Holmes and J. K. Nicholson, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2001, **39**, 1.