Marine Genomics 36 (2017) 57–75

Contents lists available at ScienceDirect

Marine Genomics

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

Method paper

How to boost marine fungal research: A first step towards a multidisciplinary approach by combining molecular fungal ecology and natural products chemistry

Marlis Reich[a,*], Antje Labes[b]

[a] University of Bremen, BreMarE, NW2 B3320, Leobener Str. 5, D-28359 Bremen, Germany
[b] Flensburg University of Applied Sciences, Kanzleistr. 91-93, D-24943 Flensburg, Germany

ARTICLE INFO

Keywords:
"Omic" techniques
Taxonomy
Metabolomics
Molecular marker
Phylogeny
Fungi
Barcoding

ABSTRACT

Marine fungi have attracted attention in recent years due to increased appreciation of their functional role in ecosystems and as important sources of new natural products. The concomitant development of various "omic" technologies has boosted fungal research in the fields of biodiversity, physiological ecology and natural product biosynthesis. Each of these research areas has its own research agenda, scientific language and quality standards, which have so far hindered an interdisciplinary exchange. Inter- and transdisciplinary interactions are, however, vital for: (i) a detailed understanding of the ecological role of marine fungi, (ii) unlocking their hidden potential for natural product discovery, and (iii) designing access routes for biotechnological production. In this review and opinion paper, we describe the two different "worlds" of marine fungal natural product chemists and marine fungal molecular ecologists. The individual scientific approaches and tools employed are summarised and explained, and enriched with a first common glossary. We propose a strategy to find a multidisciplinary approach towards a comprehensive view on marine fungi and their chemical potential.

1. Introduction

Marine fungi have a broad diversity in terms of species richness, phylogenetic distribution, and their natural products (NPs). Compared to other marine organisms, the diversity of marine fungi and their NPs had been largely neglected and only recent research has started to provide in-depth data on this understudied organism group (Bhadury et al., 2006; Richards et al., 2015; Imhoff, 2016; Rämä et al., 2016; Taylor and Cunliffe, 2016).

Cultivation-dependent and molecular studies demonstrate that marine macroorganisms, such as sponges and algae, are a rich source for fungi (Debbab et al., 2010). In consequence of this recognition, an increasing number of new NPs from marine fungi have recently been characterised and reported. There is no doubt that these organisms produce a large number of interesting secondary metabolites, which often show pharmaceutically relevant bioactivities and may be candidates for the development of new drugs. The last few years have seen an explosion of research interest in marine fungi and their NPs, as reflected in the formation of the MaFNap consortium (http://mfnpc.vre3.upei.ca/, a network of research groups with interests in the diversity of marine fungi and investigation of their NPs in the context of drug discovery). When reviewing these recent developments (Bhadury et al., 2006; Saleem et al., 2007; Imhoff, 2016; Silber et al., 2016), it became obvious that current research has some limitations: The diversity of marine fungi is not adequately represented in investigations on their secondary metabolites; most of the published work on secondary metabolites of marine fungi has focused on just a few genera, mainly *Penicillium*, *Aspergillus*, *Fusarium* and *Cladosporium* (Imhoff, 2016). Although next-generation DNA sequencing (NGS) methods are now standard approaches for the detection of multiple genes of interest (Culligan et al., 2014) and provision of genomewide information is largely employed for several fungi (Grigoriev et al., 2014), genome and genetic data on marine fungi and especially secondary metabolite biosynthesis genes are poorly represented in the current databases.

Another drawback for marine fungal research is that data collected for the single taxa are far from being comprehensive and often not publicly available. However, the data would be of utmost importance for ecological, taxonomic and biotechnological research. Furthermore, many studies use either culture-dependent *or* culture-independent approaches and available datasets are not interlinked. The availability of data and tools is additionally imbalanced between the two "worlds" of NP research and marine fungal molecular ecology: several tools and

* Corresponding author.
E-mail addresses: reich@uni-bremen.de (M. Reich), antje.labes@hs-flensburg.de (A. Labes).

http://dx.doi.org/10.1016/j.margen.2017.09.007
Received 2 March 2017; Received in revised form 22 September 2017; Accepted 23 September 2017
Available online 12 October 2017

Box I
How to name a fungal taxon according to the *Code*?

For many situations in the daily work of a fungal researcher, knowledge of nomenclature is essential for producing comparable research results. Examples are the literature research, BLASTn output with several best hits or phylogenetic analysis, where a clade can be comprised by several different named fungal taxa. In all these cases, different taxon names could be the consequence of dual naming rather than standing for different taxa or bad marker gene resolution. Further, the taxon names of the lab's own culture collections should follow the rules of the *International Code of Nomenclature for algae, fungi, and plants* (McNeill et al., 2012) to ensure good communication among scientists. Basic knowledge on the rules of the *Code* is thus critical.

In general, a name must have been validly published and the denomination has to follow the rules of the *Code* to become the correct name. In the case that several names for the same taxon fulfil the demand of being validly published and legitimate, the correct name is the one, which was published earliest ("priority"). However, several exceptions exist. To understand the *Code*, the most important vocabularies should be known:

Basionym:	The original (legitimate and previously published) name on which a new name at new rank is based. This can be a helpful starting point for solving nomenclatural problems.
Conserved name:	*(nomen conservandum)* A name that has become the correct and legitimate name although it might have been illegitimate or lack priority. Only names on species, genus and family level can become conserved.
Correct name:	The name that has to be used for a taxon in accordance with the rules.
Legitimate name:	A validly published name that is in accordance to the rules (opposite: illegitimate).
Replacement name:	*(nomen novum)* A name created to replace another scientific name, but only when this other name cannot be used for technical, nomenclatural reasons.
Rejected name:	*(nomen rejiciendum)* A name ruled out for use. This can be the consequence of a formal action (over a proposal), where the rejection is in favor of a conserved name or it does simply not follow the *Code*.
Suppressed name:	*(nomen utique rejiciendum)* A name that has been ruled out for use. Such an outright rejection can be done for names of any rank.
Synonyms:	Validly published names other than the correct name.
Validly published:	The name has correctly been published and the minimal required information on the taxon is provided with the publication.

The *Code* offers different options how to name a fungal taxon. Here, we provide examples of the two most frequently occurring cases:

1. Example: Priority rule

In general, if several validly published names exist for the same taxon and all the names are legitimate, the rule of priority is applied, thus the oldest published name will supersede the other names.

Example: The family *Ophiocordycipitaceae* comprises several genera; among them are *Tolypocladium, Chaunopycnis* and *Elaphocordyceps*. *Elaphocordyceps* taxa are mainly described as teleomorphic forms and the few described asexual morphs showed similarity to *Tolypocladium* taxa (Sung et al., 2007). For most of the *Tolypocladium* taxa, no sexual morphs have been identified. However, a high similarity of the conidiogenesis between *Tolypocladium* and the asexually typified genus *Chaunopycnis* was recongised (Gams, 1980). The results of a multiple-locus phylogenetic tree indicated that these three genera form one monophyletic clade (Quandt et al., 2014). As all names were validly published and followed the *Code*, they are synonyms and the rule of priority has to be applied.

The entries of *Mycobank* (Crous et al., 2004) can help to find the correct name (discussion of pros and cons of nomenclatural repositories see section 2.1). Under "summary", the name and the publication, which introduced the name, is listed. "Remarks" indicate possible changes to the name. The row "synonyms" lists all possible synonyms to the taxon name indicated under "summary".

Chaunopycnis W. Gams, Persoonia 11 (1): 75 (1979); Remarks: Listed as synonym of *Tolypocladium* in Quandt et al. (2014).
Elaphocordyceps G.H. Sung & Spatafora, Studies in Mycology 57: 36 (2007); Remarks: Listed as synonym of *Tolypocladium* in Quandt et al. (2014).
Tolypocladium W. Gams, Persoone 6 (2): 185 (1975); Synonyms: *Chaunopycnis* and *Elaphocordyceps.*

Conclusion: *Tolypocladium* is the oldest name among the three validly published and legitimated names. Thus, it has priority over the two other names and taxa have to be named accordingly. The other names have been proposed to become a *nomen utique rejiciendum* (Quandt et al., 2014), a name not to be used any longer. Such an outright rejection is possible for a name at any rank.

2. Example: Conservation over rejection rule

The purpose of conserving a name is "to avoid disadvantageous nomenclatural changes entailed by the strict application of the rules, and especially of the principle of priority [...]" (Art. 14.1) (McNeill et al., 2012). Conservation is possible only for names at the rank of family, genus or species.

Penicillium chrysogenum belongs to the most important penicillin producing filamentous taxa. Phylogenetic analysis showed that its taxonomy can be subdivided by four clades including isolates of other related taxa (Scott et al., 2004) and some of them were identified to be conspecific.

The entries in *Mycobank* indicate the following: *P. chrysogenum* is the conserved name (*nomen conservandum*) and thus, naming has not

followed the priority rule. The name *P. chrysogenum* has been introduced to a later timepoint (1910) than the other species names. During the action of conservation, the other synonyms were proposed to be rejected (Kozakiewicz et al., 1992). The current names of *P. brunneorubrum* and *P. griseoroseum* are thus *P. chrysogenum*.

Penicillium brunneorubrum Dierckx, Annales de la Société Scientifique de Bruxelles 25 (1): 88 (1901); Current name: *P. chrysogenum.*
Penicillium chrysogenum Thom, U.S.D.A. Bureau of Animal Industry Bulletin 118: 58 (1910); Synonyms: *P. brunneorubrum, P. griseoroseum*, among others; Name status: *nomen conservandum*.
Penicillium griseoroseum Dierckx, Annales de la Société Scientifique de Bruxelles 25 (1): 86 (1901); Current name: *P. chrysogenum.*

Helpful links:

Trying to find the appropriate rule/name for a special case is very difficult. As one rule is normally regulated by several paragraphs in the *Code*, help is needed. First orientations can be found at the following websites:

- Basics on botanical nomenclature: https://botany.si.edu/references/botlinks/dhntyp.htm (accessed Dec. 2016)
- Access to the appendices of the *Code* including:
  Appendix IIA, conserved and rejected names of families of fungi
  Appendix III, conserved and rejected names of genera of fungi
  Appendix IV, conserved and rejected names of species and infraspecific taxa
  Appendix V, suppressed names http://botany.si.edu/references/codes/props/index.cfm (accessed Dec. 2016)
- To check, which names are under revision for conservation or rejection: http://botany.si.edu/references/codes/props/index.cfm (accessed Dec. 2016)
- Glossary of the *Code*: http://www.iapt-taxon.org/nomen/main.php?page=glo (accessed Dec. 2016)

datasets for fungal classification have been developed that allow relatively straightforward classification in a user-friendly manner. In contrast, data of NP research that have the potential to stimulate ecological studies are rarely processed and accessible in a way that they can be directly used. All these points significantly limit potential outcomes of marine fungal research. To improve the situation, the two "worlds" of NP researchers and ecologists, currently two very different scientific communities with their own languages and priorities, have to be brought together.

We compiled this review to address these imbalances with the aim of providing the necessary guidelines for using the tools, explaining the approaches and to open the roads for future information exchange and collaboration. Based on the existing data and tools, we have developed a first roadmap with the aim to share language, resources and knowledge.

2. The "world" of molecular fungal ecology

2.1. Obstacles for fungal identification

Marine fungal research faced some limitations in its earliest phase due to morphology-based approaches in taxonomy, which are limited by the scarcity and plasticity of discriminatory morphological characteristics. Morphological plasticity can be caused by environmental stimuli (Rayner and Coates, 1987) including substrate type, temperature or light cycle (Slepecky and Starmer, 2009). Some *Candida* species exhibit more than seven different morphotypes in culture comprising yeast forms, chlamydospores and (pseudo-)mycelium (Andrews, 1992). Morphological plasticity can also be the consequence of a di- or pleiomorphic life cycle including diverse anamorphic (asexual) and teleomorphic (sexual) stages. A prominent example is the marine yeast

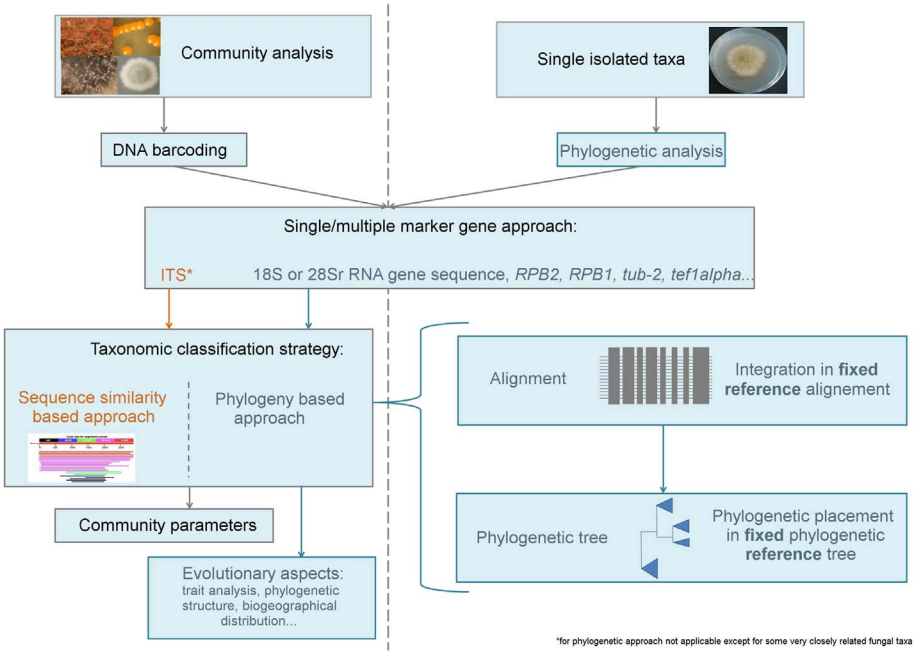

Fig. 1. Workflow schemes for the different classification strategies for community surveys and isolated taxa (e.g. from culture collections). In most cases, a phylogeny-based approach should be favoured except when monitoring known communities/isolates. Phylogenetic classification of undescribed taxa is obligatory. Note: not all marker genes are suitable for phylogenetic analysis. Orange, sequence similarity based approach; blue, phylogenetic analysis. Graphic of sequence similarity based approach taken from the output of the BLASTn program (Altschul et al., 1990) of the NCBI webpage. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Box II

The variable regions of the fungal 18S rRNA gene sequence.

The 18S rRNA gene sequence is a very important marker gene in fungal research. In total, eight variable regions span the sequence and some of them are important targets for barcoding. Although the sequence is widely used and some intensive studies have been done on its primary and secondary structure (Kappe et al., 1996; Wuyts et al., 2002), the positions and variability of the different regions have never been described and linked to an accession number of a publicly available sequence.

Here, we defined the variable positions in our manually curated fungal nearly full length 18S rRNA gene sequence alignment containing 13,501 non-redundant fungal sequences and spanning over a wide range of phyla and subphyla. The sequences were downloaded from the release 111 of the SILVA SSU Ref database (Quast et al., 2013) and aligned using the ARB software (Ludwig et al., 2004). Positions of the variable regions were manually fixed using alignment and primary and secondary structure information (Kappe et al., 1996, Wuyts et al., 2002). Indicated positions refer to the aligned *Saccharomyces cerevisiae* 18S rRNA gene sequence (acc. No. J01353). The variability of the diverse regions was defined by using the positional variability filter calculated on the curated alignment and visualised in the ARB editor window.

Table

Description of the variable regions of the fungal 18S rRNA gene sequences using the *S. cerevisiae* sequence acc. No. J01353 as the reference sequence. The eukaryotic V6 is a conserved region in fungi. A highly variable region is defined by a sequence of at least ten continuous bases with non-conserved positions.

Variable region	Position (bp)	Length (bp)	Comments on variability
V1	69–80	11	Highly variable region
V2	126–292	166	Three subareas with high variability: 126–136, 177–203, 227–242
V3	478–510	32	Conserved area, except region 490–497
V4	643–850	207	Highly variable region 643–742
V5	1048–1070	22	Highly variable region 1052–1063
V7	1350–1400	50	Variable region 1350–1375, rest is conserved
V8	1480–1531	51	Conserved area, except 1489–1504
V9	1674–1730	56	Highly variable region

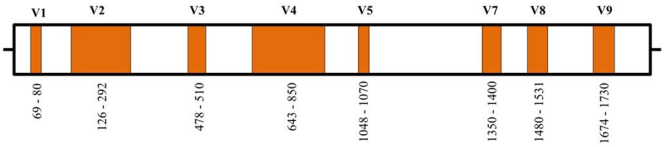

Fig. A. Schematic representation of the primary structure of the 18S rRNA gene sequence of *S. cerevisiae* (acc. No. J01353). Numbers indicate start and stop positions (nt) of variable regions (in orange).

Rhodotorula, reported from marine sediments, surface water and detritus (Lai et al., 2007; Krause et al., 2013; Panzer et al., 2015b; Richards et al., 2015) and used for industrial purposes (Zhao et al., 2010; Wang et al., 2016a). Initially described as one genus, DNA-based studies phylogenetically classified *Rhodotorula* as the anamorphic stage of diverse fungal species belonging to two different basidiomycete subphyla, three classes and at least six different orders (Kurtzman et al., 2011).

Molecular biology, thus, revolutionised marine fungal identification, but also underpinned the necessity to revise the *International Code of Nomenclature for algae, fungi, and plants* (simply called the "*Code*") ending the modus that one taxon can have several names. Hawksworth (2012b) assumed that up to 10,000–20,000 fungal species are affected by a dual nomenclature, including several marine fungi, like *Alternaria* (synonym *Lewia*) or *Cladosporium* (*Davidiella*) within the *Dothideomycetes* (Wijayawardene et al., 2014) and *Chaetomium* (*Botryotrichum/Trichocladium/Humicola*) and *Corollospora* (*Varicosporina/Halosigmoidea*) within the *Sordariomycetes* (Reblova et al., 2016). Consequently, when revising redundant fungal names, also higher taxonomic levels will be affected: *Dothideomycetes* have been described to comprise 23 orders and 110 families. From those, only 75 families are supported by molecular data, while 35 families lack molecular support (Wijayawardene et al., 2014).

Since the 1st of January 2013, the *Code* prescribes that one fungus has just one name regardless of the number of stages in its life cycle. Now, the correct name is typically the earliest name validly published – i.e. following the rules laid down in the *Code* (McNeill et al., 2012). Such a name is "legitimate". Several exceptions exist, as "illegitimate" names can be proposed to be included in the "Lists of accepted names"; for example commonly used names, which were introduced at a later time point (Hawksworth, 2012b). Further changes to the *Code* are expected in 2017 and will, for example, be related to conditions for the citation of sanctioned (adopted) names, the treatment of conspecific names with the same epithet (second part of the two-part name for a species) or the ending preference for sexually typified names. Explanations and the detailed proposals were summarised by Hawksworth (2015). (For more detailed explanation on the *Code*, we refer to Box I.)

In this situation, it is very difficult for a non-taxonomist to implement the *Code* in daily life, for instance by adjusting the designation of the taxa in the lab-owned culture collections. The correct denomination of fungal species is a prerequisite for all kinds of comparative research and to facilitate fruitful exchange between researchers independent of their research focus. Accordingly, the proper and consequent use of taxonomy is a key for secondary metabolite chemistry. Taxonomy is used as an indicator during chemical dereplication, the basis for chemodiversity studies and for intellectual property protection. Several tools exist guiding the way to a correct naming of a fungal taxon: The two most common ones are *Mycobank* (Crous et al., 2004) and *Index Fungorum* (www.indexfungorum.org, accessed May 2017) with the first one currently providing most of the names published in "non-mainstream" publications and "grey" literature. When querying a name for example in *Mycobank*, several items of taxonomic information are displayed such as synonyms, the basionym (original name with a useful description), the status of the name (validly published and conforming to the *Code*), but also information on morphology, physiology and ecology (Robert et al., 2016). (See further explanation in Box I). Users

have to keep in mind that the quality of these repositories relies on curators, who need time to verify and include changes in the database. Therefore, double-checking and comparing the entries with other repositories or recently published primary literature is highly recommended. Unfortunately, some journals in the NP field still do not have rules for taxonomic description and thereby allow publications without or with incomplete taxonomic classification. Best practice examples, such as the Journal of Natural Products (ACS Publications, guide for authors, http://pubs.acs.org/paragonplus/submission/jnprdf/jnprdf_authguide.pdf), which ask for detailed information, collection permits etc., will help to increase the pressure to comply with taxonomic requirements.

2.2. The toolkit for fungal taxon identification

Current estimates of global species numbers for Fungi suggest up to 5.1 million (Hawksworth, 2012a) but only about 100,000 are described (Kirk et al., 2008). Thus, whenever isolating or describing fungal taxa from environmental samples, the probability of dealing with one of those undescribed taxa is very high. This point is very critical for the classification process: it is highly recommendable to try to get a robust and reliable classification even if it is only to a higher taxonomic level rather than forcing classification down to (an erroneous) genus or species level. Many of the undescribed taxa actually belong to undescribed taxon groups on a higher taxonomic level: prominent examples are sequences clustering in the newly discovered groups of *Cryptomycota* (Jones et al., 2011) or *Archaeorhizomycetes* (Rosling et al., 2011).

Marker gene sequencing is nowadays the standard for fungal taxon identification. This technique is based on finding a genetic region which is easily extractable and amplifiable, therefore allowing sequencing with low DNA/tissue amount. The intertaxa sequence variation of a marker gene should exceed intrataxa variation and - under optimal conditions - the sequence is constant and unique to one taxon (group). If the aim is a fungal community survey, the variable sites of the marker gene should be flanked by highly conserved regions offering binding sites for primers shared by all fungal groups. In some cases, a polyphasic taxonomic approach can be advantageous as it exploits both phenotypic and sequence-based identification techniques simultaneously (Das et al., 2014). Currently, only a few databases offer such a service such as the webpage of the Westerdijk Fungal Biodiversity Institute (Utrecht, Netherlands, www.westerdijkinstitute.nl, accessed May 2017) including fungal genera regularly detected from marine realms like *Penicillium*, *Aspergillus* or *Fusarium*.

2.2.1. Fungal marker genes

An array of diverse fungal marker genes exist: they all differ in length, resolution (Aguileta et al., 2008; Schoch et al., 2012), phylogenetic power and number of publicly available sequences (Panzer et al., 2015b). Thus, the decision on a marker gene depends on the fungal target group and the research question (Fig. 1) and no true standard exists. The marker genes in current use can be separated into two groups: (i) genetic regions of the nuclear ribosomal DNA (nuc rDNA) cluster and (ii) protein-coding regions. In general, the PCR amplification of sequences in the nuc rDNA cluster seems to be more reliable than the protein-coding markers (Schoch et al., 2012). Additionally, the nuc rDNA cluster exists as tandem repeats in fungal genomes making the sequences easy to extract and amplify.

2.2.1.1. Nuclear ribosomal DNA cluster

2.2.1.1.1. The Internal Transcribed Spacer (ITS). The Internal Transcribed Spacer (ITS) is spliced during the post-transcriptional process of the nuc rRNA cistron. Two ITS regions exist in eukaryotes: the ITS1 located between the 18S and 5.8S rRNA gene sequence and the ITS2 located between the 5.8S and the 28S rRNA gene sequence. The segment of the two ITS regions together is roughly between 400 to 600 bp in length but can be up to 1000 bp as reported for some *Dothideomycetes* taxa (Seifert, 2009). The ITS has been proposed as the primary fungal DNA barcode for two main reasons: the interspecies variability exceeds that of intraspecies variability in a broad range of fungi and the PCR success rate is high (Schoch et al., 2012). However, the power of resolution differs between fungal groups and many environmental sequences can only be identified to the level of phylum or kingdom (Nilsson et al., 2016). Thus, the ITS successfully discriminates between many *Basidiomycota* and ascomycete *Pezizomycotina* species (Bruns and Gardes, 1993; Eberhardt et al., 1999; Tedersoo et al., 2006; Reich et al., 2009) allowing the detection of new species (Vralstad et al., 2000) and potentially cryptic species (Koljalg et al., 2002). In contrast, species resolution power is low within *Saccharomycotina* (Schoch et al., 2012) and ascomycete genera that possess a short ITS sequence (ca. 400 bp) including the marine fungi *Cladosporium*, *Penicillium* or *Fusarium* (O'Donnell and Cigelnik, 1997; Skouboe et al., 1999; Schubert et al., 2007). Also for basal fungal lineages, the ITS variation is often insufficient for unequivocal species identification. This has critical consequences as basal fungal lineages are a prominent part of marine fungal communities (Panzer et al., 2015b; Gutierrez et al., 2016; Taylor and Cunliffe, 2016) and can dominate communities of ocean surface water (Richards et al., 2015).

As well as the regularly observed intraspecific ITS variation (Smith et al., 2007; Simon and Weiss, 2008) a further obstacle for ecological studies is the very low phylogenetic power of the ITS: although the ITS can resolve the phylogeny for some closely related fungal species (Pryor and Gilbertson, 2000; Koljalg et al., 2002; den Bakker et al., 2004), ITS-based alignments of distantly related fungal taxa are not reliable due to high sequence and length variability. In contrast, the secondary structure of the ITS2 has a certain phylogenetic power (Schultz et al., 2006) but has so far been not applied for marine fungal taxon classification.

2.2.1.1.2. The 28S rRNA gene sequence. The fungal 28S rRNA gene sequence is ~2900 bp in length and contains two hypervariable regions, D1 (~130 bp) and D2 (~210 bp). Both are flanked by relatively conserved sequence regions highlighting the 28S rRNA gene sequence as a very suitable candidate for phylogenetic analysis. This marker gene has also DNA barcoding capacity when targeting the D1 or D2 region. The average accuracies across the two hypervariable regions have been reported to be 78% (D1) and 80% (D2) at genus level using a barcode length of 400 bp (Liu et al., 2012). The probability of correct classification is highest for basal fungal lineages and *Basidiomycota*, while the ascomycete *Pezizomycotina* and especially *Saccharomycotina* are less well resolved (Schoch et al., 2012).

The 28S rRNA gene sequence became the standard marker for basidiomycete yeasts (Kurtzman et al., 2011). For example, 28S–based phylogenetic analysis resolved the marine yeast genera *Cryptococcus*, *Rhodotorula* and *Bullera*, *Filobasidium*, *Leucosporidium* as poly- and monophyletic groups, respectively (Fell et al., 2000). However, for marine fungal research, the 28S rRNA gene sequence is a less used marker with the lowest number of publicly available reference sequences for marine fungi (Panzer et al., 2015b), though its successful application led to the classification of fungal isolates from the brown algae *Fucus serratus* (Zuccaro et al., 2003), the tropical seagrass *Enhalus acoroides* (Sakayaroj et al., 2010), or from wooden substrate (Rämä et al., 2014).

2.2.1.1.3. The 18S rRNA gene sequence. The 18S rRNA gene sequence is ~1900 bp and displays different variability levels with highly conserved to relatively variable regions. While the bacterial homologue (16S) possesses nine hypervariable regions, the eukaryotic 18S consists of eight variable regions as the region homologous to the bacterial V6 does not vary between eukaryotes (Kappe et al., 1996) (Box II).

A big advantage of the 18S rRNA gene sequence is its possible use both for phylogenetic analysis and DNA barcoding. The barcode should at least cover one of the variable regions (V1, V4, V5 and V9, Box II).

The longest one is V4 with ~200 bp in length (Box II) and this is the main target region for marine eukaryotic community assessments (Stoeck et al., 2010; Orsi et al., 2013; Tanabe et al., 2016). However, its application as a fungal barcode is contradictory: while it resolves *Pezizomyoctina* and *Basidiomycota* to species level only poorly, it performs very well for *Saccharomycotina* and basal fungal lineages compared to other commonly used fungal barcodes (Schoch et al., 2012). Thus, it has mainly been applied as a barcode for freshwater fungal community surveys due to its good resolution for basal fungal lineages (Lefevre et al., 2007; Monchy et al., 2011) and is gaining attention for general marine community surveys (Arfi et al., 2012; Redou et al., 2014; Reich et al., 2017). Currently, it is the prevalent marker gene in public databases for marine fungi (Panzer et al., 2015b) and the only marker gene for which phylogenetic reference trees exist (Panzer et al., 2015a; Yarza et al., 2017).

For phylogenetic analysis, the 18S rRNA gene sequence is a very robust marker. It was thus used to resolve the phylogeny around *Rozella* and the LKM11 clade (Lara et al., 2010), to identify yeasts as the dominant morphological group in the deep ocean (Bass et al., 2007), and to explore marine fungal diversity in general (Richards et al., 2012; Panzer et al., 2015b). Phylogenetic analyses become more reliable when (nearly) full-length sequences are taken. The sequence alignment can be improved by consulting the secondary structure organised as loops and stems. Such a secondary structure model exists so far only for *Saccharomyces cerevisiae* and *Coprinus cinereus* and these are superimposed with the overall eukaryotic variabilities (Wuyts et al., 2000) (http://bioinformatics.psb.ugent.be/webtools/rRNA/secmodel/index.html, accessed May 2017).

2.2.1.2. Protein-coding sequences. A variety of protein-coding sequences have been used in marine fungal research but not for DNA barcoding purpose with the exception of the *RPB1* and *RPB2* gene (Stockinger et al., 2014; Vetrovsky et al., 2016). All these genes exist in most of the fungal groups as single-copy genes, which sometimes causes problems in their extraction and amplification but avoids problems with paralogues. An advantage of protein-coding genes in phylogenetic analysis is the possibility of using the respective amino acid sequence as backbone for correction of sequencing errors. The necessary sequence variability among taxa is mainly found in intron sequences and at the third codon position allowing certain variability on the DNA level.

2.2.1.2.1. RPB1: the gene encoding for the largest subunit of RNA polymerase II. RNA polymerase II is the enzyme that transcribes protein-coding genes into pre-mRNA transcripts and consists of several subunits. The *RPB1* gene codes for the largest subunit and possesses an exon-intron structure with several variable and nine more conserved regions (Matheny et al., 2002). *RPB1* has the most significant barcoding gap compared to other fungal marker genes, which means that in terms of genetic distance nearly no overlap of intra- and interspecific sequence variation can be observed. It outperforms even the ITS in reliable species identification. For *Glomeromycotina*, it has been shown that the barcode gap of *RPB1* separates almost all morphospecies correctly (Stockinger et al., 2014). *RPB1* has low performance on basal fungal lineages (Schoch et al., 2012).

Another drawback, beside a reliable PCR amplification, is the current lack of primers targeting a wide range of fungal groups. Primers exist for *Agaricales* (Matheny et al., 2002), *Lecanoromycetidae* and *Peltigerales* (Hofstetter et al., 2007), and *Verrucariaceae* (Gueidan et al., 2007). Consequently, the selective studies on a small number of fungal groups are responsible for the bias of publicly available sequences towards mainly ascomycete species. However, *RPB1* is nowadays a standard in multiple-locus phylogenetic studies (Lutzoni et al., 2004; James et al., 2006) independent of its low performance compared to other single-copy genes (Aguileta et al., 2008).

2.2.1.3. RPB2: the gene encoding for the second largest subunit of the RNA polymerase II. The *RPB2* gene encodes the second largest subunit of the RNA polymerase II. As for *RPB1*, it is widely used for improving the phylogenetic inference on lower and higher taxonomic levels (Matheny, 2005; James et al., 2006; Matheny et al., 2006). Even shorter sequence reads with a barcode length (400–800 bp) allow very accurate identification. Therefore, *RPB2* has been suggested as an alternative fungal DNA barcode (Vetrovsky et al., 2016). The ability to work with *RPB2* in a phylogenetic context is a significant advantage over ITS. However, the lack of primers targeting fungi is also an issue for this gene. The existing primers were designed to be group-specific (Liu et al., 1999; Matheny, 2005; Matheny et al., 2007), although some of them show co-amplification towards other fungal groups and slight modifications might adapt them to be more suitable candidates for community analysis (Vetrovsky et al., 2016).

2.2.1.3.1. Other protein-coding markers. Further protein-coding marker genes have been in use for fungal research but to a much lower extent and always as phylogenetic markers rather than as fungal DNA barcodes. One reason is the low number of publicly available sequences and appropriate primers. However, there is an ongoing trend towards the inclusion and use of protein-coding, single-copy genes in fungal research due to the drawbacks of markers from the *nuc* rDNA cluster. Several of the markers discussed in this subsection (see below) are nowadays a standard in multiple-locus phylogenetic analysis (Feau et al., 2011).

Tubulin is the major constituent of microtubules and consists as a dimer of α- and β- chains. The gene *tub-2* encodes for the β-tubulin chain and is the third most utilized gene in fungal multi-locus phylogenies (Feau et al., 2011). *Tub-2* has been of special interest in studies resolving the systematics of *Penicillium* and *Phaeoacremonium* (Dupont et al., 2002; Samson et al., 2004; Rezaei-Matehkolaei et al., 2014), therefore adopted sequence databases have been developed (www.westerdijkinstitute.nl, accessed May 2017). One drawback of the *tub-2* gene sequence is the existence of paralogues, which are regularly misinterpreted as *tub-2* gene sequences (Hubka and Kolarik, 2012).

Another commonly used marker gene is *tef1α* coding for the translation elongation factor 1 (EF1)-α with a chain length of about ~400 amino acids. It is usually present as single-copy but has some paralogues in certain fungal genomes (Keeling and Inagaki, 2004). The partial sequence of *tef1α* is often used for successful identification of *Fusarium* species and can be accessed over the Westerdijk Institute-based *Fusarium* identification site (www.westerdijk.nl/Fusarium). The *act* gene encodes γ-actin, a cytoskeletal filament and has been found to be single-copy in the majority of fungi tested with exceptions (Tarkka et al., 2000). Both gene sequences have helped to resolve the evolutionary relationship of fungi on higher taxonomic level (Baldauf, 1999; Hirt et al., 1999; Baldauf et al., 2000; O'Donnell et al., 2001; Voigt and Wostemeyer, 2001; Helgason et al., 2003; Kristensen et al., 2005). Even if commonly used, the topology scores of phylogenetic trees inferred from data of *tub-2*, *tef1α* or *act* are poor compared to other single-copy genes, even when a concatenated alignment was used (Aguileta et al., 2008).

Further protein-coding genes used as fungal markers include those coding for heat shock proteins, chitinases, chitin synthases, dehydrogenases, histones or ATP6 but these have so far played only minor roles in fungal multigene phylogenies (Feau et al., 2011).

The most interesting marker candidates for the future are the genes *MCM7* and *MS277*. By testing 246 single-copy orthologous genes, these two were the only ones with a high phylogenetic performance within the group of *Ascomycota* and *Basidiomycota* demonstrating topological congruence with the overall fungal consensus tree (Aguileta et al., 2008). *MCM7* codes for a protein that is part of the MCM complex, which is important for priming replication origins. *MS277* codes for the ribosome biogenesis protein TSR1 required for processing of 20S pre-rRNA in the cytoplasm. Both have been successfully tested for revealing the evolutionary history of *Kickxellomycotina* (Tretter et al., 2013) and *Pezizomycotina* (Schmitt et al., 2009).

Table 1
Datasets and databases useful for molecular identification and chemotaxonomy of marine fungi.

Datasets/databases	Kind of (sequence) information (listed only those of importance for fungal research)	Related publication or webpage	Comments
Fungal-specific datasets			
AFTOL (Assembling the Fungal Tree of Life)	Alignments for selected fungal groups: RPB1 & 2; 5.8S, 18S & 28S rRNA gene sequence; concatenated sequences	www.aftol.org/index.php Structural and biochemical database: https://aftol.umn.edu	Manually curated alignments on high-quality sequences, primer information, links to structural and biochemical databases
Aquatic fungal phylogenetic tree	18S rRNA gene sequences	(Panzer et al., 2015a)	Based on quality-checked sequences and KSMP Kiel culture collection integrated sequences
Faces of Fungi	Details for genera and species: molecular data, morphology, physiology, ecology, biosecurity, industrial relevance, biochemistry, phylogeny	(Jayasiri et al., 2015)	Open to all mycologists, data accepted from all mycologists
Fungal 18S rRNA gene sequence reference tree	18S rRNA gene sequences	(Yarza et al., 2017)	Based on 9239 manually aligned high-quality sequences; package includes alignment, phylogenetic tree, up-to-date taxonomy and ARB files as ready-to-use tools
Fusarium MLST	*tef1α*, RPB1, RPB2	www.westerdijk.nl/Fusarium	*Fusarium* specific, Pairwise DNA alignments, polyphasic identification
International subcommission on *Trichoderma* and *Hypocrea* Taxonomy	ITS sequences, partial *tef1α* and RPB2	www.isth.info/index.php	*Hypocrea/Trichoderma*-specific, polyphasic identification, sequence library
PHYMYCO database	18S rRNA gene sequence, *tef1α*	(Mahe et al., 2012)	Quality-checked sequences, manually cross-checked alignments
Q-Bank	Molecular decision schemes and sequences to identify selected fungal species: ITS, *tef1α*, *act*, *CAL*, *COI*, *tub-2*	www.q-bank.eu/Fungi/	Quality-controlled sequences, polyphasic identification, protocols, recommendations
UNITE	ITS sequences	(Koljalg et al., 2013)	Manual curated datasets, species hypothesis, reference and representative sequence
Other interesting datasets			
Internal Transcribed Spacer 2 ribosomal RNA database	ITS2 sequences and secondary structure	(Ankenbrand et al., 2015)	Detection and secondary structure prediction of ITS2 sequence for phylogenetic analysis
Open Tree of Life	Phylogenetic trees	(Hinchliff et al., 2015)	Dynamic and digitally-available tree of life by synthesising published phylogenetic trees
Protist ribosomal reference database (PR2)	Eukaryotic 18S rRNA gene sequence, some sequences from organelle SSU	(Guillou et al., 2013)	Manually curated
Ribosomal database project (RDP)	Fungal 28S rRNA gene sequence alignment, fungal ITS	(Cole et al., 2014)	Quality-controlled sequences, alignments and analysis tools
SILVA database	Eukaryotic 18S and 28S rRNA gene sequence	(Quast et al., 2013)	Quality-checked sequences, manually curated alignments, NGS pipeline
TreeBase	Phylogenetic trees and data used to generate them	https://treebase.org/treebase-web/home.html	Repository of phylogenetic information
(Marine) fungal culture collections[a]			
American type culture collection (ATCC)	Filamentous fungi and yeasts, representing over 7600 fungal species	www.lgcstandards-atcc.org/en/Products/Cells_and_Microorganisms/Fungi_and_Yeast.aspx	Offers storage and management of biomaterials
Biological Resource Center, NITE (NBRC)	Filamentous fungi and yeasts	http://www.nite.go.jp/en/nbrc/	General deposit possible
Canadian Collection of Fungal Cultures (CCFC)	Over 350,000 fungal and fungal plant disease specimens	http://res.agr.ca/brd/ccc	Small subset of national culture collection comprises marine isolates
Kiel collection of marine microbes	Marine fungi	www.geomar.de	Collection of marine fungi at Helmholtz Centre for Ocean Research, GEOMAR, not public
Marbank	Marine organisms from Norwegian waters	www.imr.no/marbank/marine_products/en	Offers storage of samples, marine products and customised sampling
National Center for Genetic Engineering and Biotechnology (BIOTEC)	2486 fungal species	www.thai2bio.net	Small subset of marine isolates
United Kingdom National Culture Collection	Central access point for a variety of collections: fungi, yeast, pathogenic fungi	www.ukncc.co.uk	Small subset of national culture collection comprises marine isolates
Westerdijk Fungal Biodiversity Institute	More than 50,000 fungal strains	http://www.westerdijkinstitute.nl/	Small subset of national culture collection comprises marine isolates
Natural product databases			
Dictionary of Natural Products (DNP)	Descriptive and numerical data on chemical, physical and biological properties of compounds; systematic and common names of compounds; literature references; structure diagrams and their associated connection tables	http://dnp.chemnetbase.com/dictionary-search.do?method=view&id=12018701&si=	The Chapman & Hall/CRC Chemical Database is a chemical structure-based database holding information on chemical substances.
Dictionary of Marine Natural Products	As DNP	http://dmnp.chemnetbase.com/intro/index.jsp	Contains over 30,000 compounds. DMNP is a subset of the DNP.
Marine Lit	Article records, collection site data, full taxonomy of producing organism, compound records and properties, cross referenced with binding data in BindingDB	http://pubs.rsc.org/marinlit	Searchable features and powerful dereplication tools

(*continued on next page*)

Table 1 (*continued*)

Datasets/databases	Kind of (sequence) information (listed only those of importance for fungal research)	Related publication or webpage	Comments
Global Natural Products Social Molecular Networking	Mass spectrometry data	https://gnps.ucsd.edu/	Open-access knowledge base for community-wide organization and sharing of raw, processed or identified tandem mass (MS/MS) spectrometry data
Natural products database	2D structures, physicochemical properties, predicted toxicity class and potential vendors	http://zinc.docking.org/	Commercially available natural products and natural product derivatives
Super Natural II	Compound information and the corresponding 2D structures, physicochemical properties, predicted toxicity class and potential vendors	http://bioinf-applied.charite.de/supernatural_new/	Database of natural products; 325,508 natural compounds, most entries plant based

[a] http://www.wfcc.info/home provides culture collections around the world. The WFCC pioneered the development of an international database on culture resources worldwide. The result is the WFCC World Data Center for Microorganisms (WDCM). This data resource has records of nearly 476 culture collections from 62 countries. The records contain data on the organization, management, services and scientific interests of the collections. Alternatively, the Global Registry of Biodiversity Repositories (GRBio), http://grbio.org/, can be used as a starting point for searches.

2.2.1.4. Combined marker approach. As shown for all marker genes, conclusions based on a single marker gene might be highly misleading. Accordingly, larger phylogenetic projects nowadays are based on a multiple-locus approach. For example, the "Assembling the Fungal Tree of Life" (AFTOL) project resolved the early evolution of fungi with the help of six different loci (James et al., 2006). In a multiple-locus approach, both concatenated and single marker analysis should be run in parallel, as nodes not recovered in a multiple-marker reference tree might appear well-supported in single-gene trees or vice versa.

Multiple-marker gene approach is also the most reliable one for DNA barcoding. According to Q-Bank (http://www.q-bank.eu/Fungi/, accessed Dec. 2016) the combination of *tef1α* or *act* alongside the ITS is the best technique for the identification of isolates belonging to the genus *Mycosphaerella* associated to algae (Fries, 1979; Garbary and Macdonald, 1995) or *Phoma* inhabiting marine sediment (Yang et al., 2005), respectively. The same holds true for community surveys. Marine fungal communities are often dominated by yeasts, filamentous ascomycetes or basal fungal lineages, for which species identification is limited with ITS. Tandem amplification of a second phylogenetic marker consequently allows the classification of unknown species and groups to a taxonomic level beyond kingdom. Another advantage is the possibility to implement ecological evolution-driven questions into the analysis (Fig. 1). A combination of ITS and the 18S rRNA gene sequence has been proposed (Arfi et al., 2012; Reich et al., 2017) but in the future, other combinations might be favoured like the 28S rRNA gene sequence and *RPB2* or *MCM7* with *MS277*. This will be highly dependent on the development of sequence databases and sequencing techniques. However, due to the wide diversity within the kingdom of Fungi, the standard markers might enable only coarse identification of some groups. Therefore, follow up analyses with other target regions might become necessary, when species-level identification has to be achieved.

2.2.2. Molecular-based fungal identification tools

Molecular fungal identification tools are mainly based on two different approaches to finally annotate the query sequence by using sequence similarity or phylogenetic trees. The main difference is that the first classifies taxa with reference to existing classifications bound to a sequence similarity threshold, while the latter one classifies over a phylogenetic hypothesis. Thus, the results and the conclusions, which can be driven from one or the other, can be quite different. A more recently developed tool, multispecies coalescent (MSC) approach, can currently only fulfil species boundary fine-tuning, but has great potential to delimit fungal species most correctly based on whole-genome coalescence.

2.2.2.1. Sequence similarity-based approach. Fungal taxon identification in a sequence similarity-based approach is based on finding a sequence in a reference database showing a high sequence similarity to the query. Basic Local Alignment Search Tool (BLAST) (Altschul et al., 1990) is one of the most widely used similarity search programs. The program works in two steps. First, it identifies high-quality short matches between reference and query sequences over a pattern search (default three nucleotides). From the best matches, it attempts to start alignments and extends them under a given score. Diverse alignments are scored and compared to each other until the best alignment(s) is found.

Once the "best hit" is identified, the query is taxonomically classified by transferring the classification from the reference sequence to the query provided that their similarity is high enough. Thus, a correct classification depends highly on the reference dataset. A good sequence reference database is characterised by high-quality sequences with a reliable annotation (Nilsson et al., 2006) and where changes in fungal taxonomy are quickly incorporated. An example for the latter case is the nucleotide sequence Genbank database of the National Center for Biotechnology Information (NCBI) (https://www.ncbi.nlm.nih.gov/genbank/) that is linked to the manually curated NCBI taxonomy database (Federhen, 2012). Here, fungal classification is quickly adapted based on the results of the newest published phylogenetic analyses.

However, some cons exist for this approach: Uneven taxonomic sampling of most reference datasets lowers the probability of classifying all fungal groups. This problem regularly occurs when working on ecological systems that are a rich source of new fungal species and groups such as the marine realm (Manohar and Raghukumar, 2013). For these sample types, the risk is high that fungi under-represented in the reference dataset are not identified. Here, a phylogeny-based approach is recommended. Another problem arises, when a genetic region possesses only low resolving power for the fungal group of interest: this leads to a long list of "best hits" all with acceptable sequence similarity values. In the case of several "best hits", the lowest common ancestor rule (LCA) is normally applied. The LCA of the "best hits" from conserved regions often unifies at a higher taxonomic level such as phylum or even kingdom. Researchers have also to deal with the undescribed fungal species regularly detected in environmental samples. The lack of a species name hinders discussion and communication regarding this taxon. A good step to overcome this obstacle is the use of the digital object identifiers (DOIs) developed by the UNITE initiative (Koljalg et al., 2016) allowing the identification of the same unnamed taxon occurring in diverse habitats and systems.

The similarity threshold classifying a BLAST hit as a "best hit" reflects the percentage that separates intra- from interspecific sequence variability. However, no single threshold value can capture these boundaries for all fungi. For ITS for example, a standard threshold of 97% is normally applied. 3% intraspecific variability, however, is much too high for genera like *Aspergillus* or *Penicillium* and for others too low (Nilsson et al., 2008). To overcome this problem, the UNITE initiative introduced a species hypothesis (SH) concept for fungal ITS sequences. A taxon is hereby identified over sequence clustering at diverse

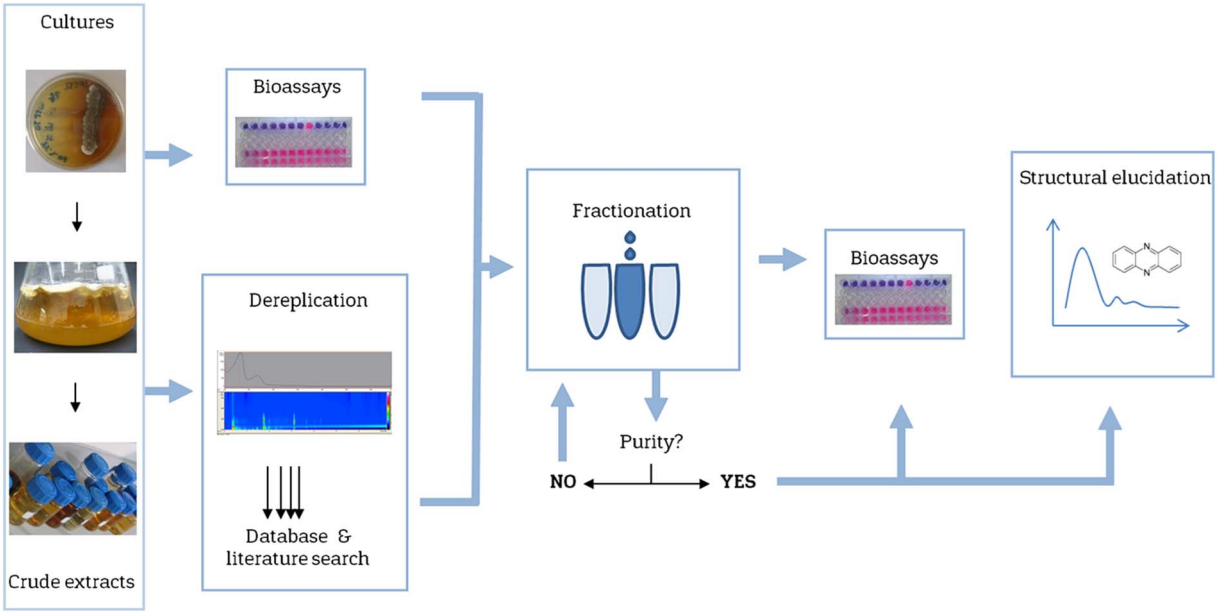

Fig. 2. Bioassay guided isolation and structural elucidation of natural products, the most common approach in natural product discovery.

similarity thresholds ranging from 0 to 3% (Koljalg et al., 2013). Currently, the SH is just based on a very tight threshold (up to 3%) though some fungal taxa possess much higher intraspecific variability. Another problem is that many marine fungal genera are under-represented in the UNITE database (Panzer, 2013). However, the ongoing work of the UNITE initiative will lead to an improvement of the database. Thus, mycologists should include the already existing SHs in their analyses allowing an adjustment of the sequence similarity thresholds to the fungal taxa of interest.

A sequence similarity based approach is often applied together with DNA barcoding but also longer sequences can be used. The sequence similarity based approach leads to the best results when applied for monitoring known fungal taxa, or for community analysis consisting of fungal taxa well represented in the reference database. Identification to species level is only achieved with a good resolving marker gene and the appropriate sequence similarity threshold. For analyses that do not reflect these optimal conditions, a manual readjustment is necessary and the addition of a phylogenetic marker.

2.2.2.2. Phylogeny-based approach. Fungal classification is structured on phylogenetic (evolutionary) hypotheses, which are ideas about the pattern of genealogical relationships linking the different fungal clades. The phylogenetic hypotheses are represented by tree-like diagrams called phylogenetic trees. As the phylogenetic tree reflects the process of evolution including the origin and diversification of fungal life, the big advantage of a phylogeny-based approach over a sequence similarity based one is the possibility of studying convergent evolution, diversification or transition processes. Phylogenies can also be used to track the geographic movements of species.

The first step is the alignment of the sequences identifying specific characters used to infer the phylogenetic tree. Adjacent positions in the alignment are presumed to be homologous meaning that they are derived from the same nucleotide position in a common ancestor. Therefore manual inspection and curation of a processed alignment is obligatory, as any errors will result in misleading inferences.

The utility of a molecular marker for a phylogenetic analysis depends on its evolutionary rate and degree of conservation. Thus, a given marker gene can be effectively used to resolve a soft polytomy if it evolves at an optimal rate at a relevant time scale (Townsend, 2007). However, genes are known to differ in their information content over historical time (Graybeal, 1994) and among fungal groups. Thus, the selection of the appropriate marker genes depends on the taxonomic level to be resolved and the taxonomic group.

For a reliable phylogenetic analysis, high-quality alignments and trees are needed. As inferring phylogenies is a science in itself and experience is needed, it might be better for those who are not experts to use an existing alignment/phylogenetic tree rather than a quickly generated one. Unfortunately, one obstacle here is the low number of published fungal phylogenies deposited in publicly accessible databases. It is estimated that about 70% of fungal alignments/trees published are no longer available (Drew et al., 2013). Some fungal-specific high-quality alignments can be downloaded over the AFTOL project website (http://wasabi.lutzonilab.net/pub/alignments/download_alignments, accessed Sept. 2017) including single marker and concatenated alignments. They are constructed for diverse fungal groups, such as *Ascomycota* (Miadlikowska et al., 2006), *Basidiomycota* (Matheny et al., 2006; Matheny et al., 2007), and alignments with 214 fungal taxa (Lutzoni et al., 2004; James et al., 2006). The database PHYMYCO-DB holds fungal-specific 18S rRNA gene sequence and *tef1α* alignments (Mahe et al., 2012). Another option is the use of eukaryotic alignments that are adjusted to the purpose of the particular study by deleting non-target groups. Several curated eukaryotic reference alignment databases exist (for more details see Table 1). Existing phylogenetic trees can also be used for inferring the phylogeny of the taxa of interest: Several programs exist that place a raw query sequence within a reference tree by phylogenetic placement (Ludwig et al., 2004; Matsen et al., 2010; Berger et al., 2011): Hereby, the placement is based on a phylogenetic criterion without changing the overall tree topology. Thus, the classification of the taxon of interest can be inferred from the placement within the reference tree like the fungal 18S rRNA gene sequence tree (https://www.arb-silva.de/download/archive/publications/fungi_18S/, accessed Sept. 2017) (Yarza et al., 2017). TreeBASE is a repository of user-submitted phylogenetic trees and associated data and deposited trees can be downloaded as a NEXUS-file (https://treebase.org/treebase-web/home.html9; accessed Dec. 2016). Another source of phylogenetic trees is the “Open Tree of Life” project, which assembles published trees into a complete tree of life synthesised with taxonomic data. Each node of the tree can be explored and discussed by setting comments. The upload and download of trees is possible (Hinchliff et al., 2015).

The phylogeny-based approach is the right choice, whenever new species and groups need to be correctly classified or evolutionary and

ecological forces shaping communities are of interest (Horner-Devine and Bohannan, 2006; Emerson et al., 2011). For fungi that are part of a culture collection, a phylogeny-based approach should be favoured to ensure an accurate classification.

2.2.2.3. Coalescent-based species delimitation. Phylogenetic trees, however, sometimes produce conflicting results as gene trees and species trees can overlap to a very significant degree without being identical (Yarza et al., 2017). The observed differences result from biological processes such as gene duplications and losses, horizontal gene transfer across species boundaries, hybridization events or are simply the product of misidentification of orthologs (Xu and Yang, 2016). MSC methods have the potential to overcome these shortcomings, as population parameters and processes inferred from DNA sequence data are used to define species boundaries. Thus, input for MSC tools are sequence alignments (up to several loci), group (population/species) membership and in some cases a phylogenetic tree. Supplying bad-quality alignments and phylogenetic trees will lead to misleading estimations of species numbers. In recent years, various MSC methods have been developed which differ in the model of species delimitation. For a detailed description of the underlying theory and methods, we would like to refer the reader to excellent reviews on this topic such as Carbone and Kohn (2004), Fujita et al. (2012) or Xu and Yang (2016).

Although MSC is now more and more used to define species within various organism groups (Fujita et al., 2012), its application for fungal species delimitation remains challenging. Fungi often do not adhere to the biological species concept, for example by having offspring through interspecific mating or asexual reproduction (Stukenbrock, 2013). This hinders the sampling of sequence data within the frame of populations. Further drawbacks are the high demand on computational power and the availability of sequences from different loci for the same species (Zhang et al., 2011). Therefore, MSC can currently only be considered as an additional tool for fine-tuning species boundaries of well-defined taxon groups rather than a standard approach for fungal species identification. However, with the increasing number of sequenced fungal genomes, comparative genomics will unlock genomic-based mechanisms on content and structure that drives fungal lineage divergence and species boundaries. Once species divergence in less well defined fungal taxa are understood, whole-genome coalescent approaches will become a powerful tool for estimating genomic relationships and speciation times.

2.2.3. Fungal identification in the era of "omics"

The use of NGS techniques opened a new era for marine fungal biodiversity research, tapping an unexpected fungal diversity and richness in the marine realm (e.g. Amend et al., 2012, Redou et al., 2014, Richards et al., 2015, Hassett and Gradinger, 2016, Rämä et al., 2016). It further scrutinized the dynamic of fungal communities, deciphering influencing factors and interactions (Orsi et al., 2013; Taylor and Cunliffe, 2016; Reich et al., 2017). However, metabarcoding currently relies on the amplification of (relatively short) marker gene regions and thus is biased by limited taxon resolution and PCR errors. A multiple-marker gene approach improves resolution power but remains PCR-dependent. With the characteristics of single-molecule template based sequencing and generation of read length up to (several) kb, third-generation sequencing technologies are good candidates to overcome the limitations of metabarcoding (Schadt et al., 2010). Unfortunately, the current sequencing error rate is still far too high (Koren et al., 2012) to support a reliable fungal identification.

Thus, the roadmap for future fungal taxon identification indicates some directions: Single-molecule sequencing (SMS) technologies with no need of PCR amplification will eradicate an entire suite of PCR-induced errors. SMS in a biodiversity context is similar to a metagenomic approach where the pool of scattered genomes in a sample are fragmented and sequenced. As marine fungi often represent a smaller part of the microeukaryotic community, intensive sequencing is necessary to receive sufficient amounts of reads for a reliable classification. In contrast, future single-taxon identification will likely be based on the sequence of the whole genome, which has the power to resolve the phylogenetic placement of fungal lineages within a phylogenomic context (Wang et al., 2009) – and will deliver at the same time the full genetic background including secondary metabolite gene clusters. The outcome of both approaches relies highly on the quality of downstream bioinformatic analyses. So far, no standards have been developed and no well-surveyed tools are available for easy fungal identification. The need of developing such methods and standards demands once again a global, integrative, and interdisciplinary programme for marine fungal research.

2.3. Ecological "omics"

"omics" applied in ecological studies can go far beyond simple fungal identification by addressing questions of nutrient cycling, stress responses or ecological niche construction (Dupont et al., 2007). Therefore, ecological "omics" provide information that is unobtainable by any other means, for example the biological capacities of marine fungi that underlie the ecology of marine ecosystems. WGS of key organisms or meta-analysis of natural communities shed light on the biodiversity that supports ecosystem function. Genomics further fuel the understanding of taxa relationships, speciation events, life-history patterns and physiological ecology. One example is the WGS-based identification of shared genomic elements between the fungus *Rozella allomycis* and closely-related endoparasitic microsporidia species coding for a nucleotide transporter that is used for stealing energy in the form of ATP from the host. At the same time, the *Rozella* genome lacks many genes for primary metabolism but is enriched for signal-transduction genes, elucidating a host-dependent life-history (James et al., 2013).

Transcriptomics tackle the active part of a community or the expressed genes and thus can study factors responsible for ecological niche limitations: For example, fungi thriving in the subseafloor are exposed to extreme energy limitations. One adaptation strategy seems to be an elevated autophagy that reduces and recycles unnecessary proteins for new protein synthesis (Orsi et al., 2015). Finally, proteomics reflects the fungal response to the environment and provides complementary information to what could be explained by genome and transcriptome variation. Being unaffected by post-translational and -transcriptional modifications or phosphorylation events, proteomics is a genuine measure for the phenotype. It is the key to understand adaptation and colonisation processes for fungi that possess very similar genomic elements (de Vries et al., 2017) but show variations in life-history patterns. Up to now, proteomics has not been applied for marine fungal ecology research. One reason might be the lack of adequate databases that link the identified protein classes/groups to a functional pathway. The link between genetic loci, expressed genes and proteins with functionality is currently the largest obstacle for all "omics". However, NP research has the potential to identify many of the missing links and expand the knowledge on physiological pathways in marine fungi. Transferring this knowledge into databases accessible and easily readable for ecologists is an important step towards interdisciplinary research on marine fungi.

3. The "world" of natural products discovery and functional studies

3.1. Natural products discovery from marine fungi

Marine fungi are a rich and promising source of novel anticancer, antibacterial, antiplasmodial, anti-inflammatory and antiviral agents, and bioactive compounds isolated from them have been comprehensively reviewed (Bugni and Ireland, 2004; Bhadury et al., 2006; Saleem

et al., 2007; Blunt et al., 2016). These metabolites are produced as so called “secondary metabolites”, organic compounds that are not directly involved in the normal growth, development, or reproduction of an organism, but provide long-term enhancement of the organism's survival prospects. Secondary metabolites are often bioactive and are usually of low molecular weight. Although chemically diverse, all secondary metabolites in fungi are produced by a few common biosynthetic strategies based on variations and combinations of a small number of fundamental pathways, often in conjunction with morphological development (Keller et al., 2005).

Studies of NPs from marine fungi continue to rise: considering mangrove-associated fungi, the number of studies published is increasing every year, with the majority coming from endophytic species. Most of these metabolites are derivatives of those discovered previously from terrestrial fungi but derivatives might exhibit significant differences in their biological activities, potency and specificity. In recent years, the fungal sources of new metabolites have been broadened from saprophytic terrestrial strains and living plants with their endophytes to marine habitats. While by far most of the metabolites described are produced by asexual stages of ascomycetes e.g. *Penicillium* and *Aspergillus* species, the interest in basidiomycetes and the number of metabolites derived from them has grown considerably (Imhoff, 2016). The fungal species of the former group of *Zygomycota,* so far easily isolated and cultured, seem to be less prolific producers (Schueffler and Anke, 2014).

Considering their enormous untapped biodiversity and their almost ubiquitous distribution, the diversity of marine fungi is not adequately represented in investigations on their secondary metabolites. NP discovery from marine fungi usually follows the classical discovery approach; combining chemical and biological screening for a bioassay-guided isolation of NPs from fungi isolated from marine habitats by classical culture based isolation techniques (Fig. 2).

Metabolomic approaches have only recently been adopted in the field of NP drug discovery. Thereby, NP research is changing and rapidly adopting cutting-edge tools with the potential to radically transform the characterisation of small molecule-containing extracts. Early integration of deep metabolome annotation information can efficiently guide the isolation of valuable NPs. In parallel, the new approaches generate massive metadata sets which can be used for various study perspectives, such as ecological questions (Allard et al., 2017). Still, most of these studies focus on the discovery of the novel chemistry rather than the biology of the producer. The motivation to search for NPs in most cases differs from the motivation of gene miners or even fungal ecologists, being interested in the genetic structure, evolutionary questions or functional system analyses. However, analyses of secondary metabolites based on the genetic background of the producing fungus face the difficulty of combining phenotypes and genotypes and thereby understanding the biological background of the synthesis of the metabolite of interest. Here, the disciplines could come together to allow pioneering research in all fields. The increasing number of fungal genome sequences indicates that their biosynthetic potential is far from being exploited (Grigoriev et al., 2014).

Ecology and lifestyle have a big impact on the secondary metabolite profile of fungi. The behaviour of a fungus isolated on a petri dish differs significantly from that in its natural habitat. Only a small part of the marine fungal metabolic potential is observed under classical experiments consisting of cultivating one strain in a standard medium over a definite period of time in artificial lab conditions. This greatly limits the potential of drug discovery from fungi (Brakhage, 2013). To overcome this issue, many research teams have then investigated ways to unravel cryptic biosynthetic pathways to access a wider chemodiversity (Pettit, 2011; Gram, 2015; Reen et al., 2015). Thus far, different methods were applied to awake silent pathways using elicitors and epigenetic modifiers, as well as co-cultivation of organisms (Scharf and Brakhage, 2013). However, most of these approaches are based on random selection of conditions rather than on prediction, so called OSMAC approaches (*o*ne *s*train *ma*ny *c*ompounds, coined by Bode et al., 2002.

3.2. Limitations in prediction

There are several classes of secondary metabolites including polyketides, non-ribosomal peptides, terpenes, indole alkaloids and “hybrid molecules”. Genes for their biosynthetic pathways are grouped in clusters including transcription factors, enabling relatively easy identification and prediction of the clusters in genome sequences. A number of fungal genomes are currently being sequenced and those already completed demonstrate a tremendous overall biosynthetic capacity of fungi. On average, about 30 to 40 biosynthetic gene clusters coding for secondary metabolites are found in a single genome (Sanchez et al., 2012). The majority of these gene clusters have not been linked to their corresponding NPs but improved understanding will facilitate phylogenetic considerations of biosynthetic pathways. These gene clusters are accessible for bioinformatic prediction of the genes themselves, but as many of these newly discovered secondary metabolism gene clusters are silent under standard laboratory conditions, a connection between the gene cluster and the product cannot be made in many situations. Furthermore, the biosynthetic pathways can be assembled in many combinations and sometimes in iterative reactions, resulting in a variety of possible products, making prediction of the final secondary metabolite structure challenging (Keller et al., 2005). For instance, polyketides, the most abundant fungal secondary metabolites, are produced by enormous multi-enzyme complexes. Fungal polyketides are synthesised by type I polyketide synthases (PKSs), short-chain carboxylic acids are condensed to form carbon chains of varying lengths. In the fungal PKSs, the ketoacyl CoA synthase (KS), acyltransferase (AT) and acyl carrier (ACP) domains are essential for polyketide synthesis. Fungal PKSs have one module, which can carry out repeated biosynthetic reactions, and are therefore called ‘iterative PKSs’. The diversity of fungal polyketide structures results from the number of iteration reactions, the number of reduction reactions, which extender unit is used and, in the case of aromatic polyketides, cyclisation of the nascent polyketide chain. Further variety is achieved by the introduction of many different post-polyketide-synthesis steps (Keller et al., 2005). This complexity complicates the prediction of the resulting chemical structure from genetic data. The prediction of the biological activity or function of the secondary metabolite/NP is a challenging task. In contrast to proteins, where similarities often implicate similar activities, belonging to a chemical class does not necessarily allow prediction of biological activity. There are different reasons for this: Firstly, small structural differences in the molecules may lead to different bioactivities (as illustrated e.g. in (Silber et al., 2016)) – only structure-function-analyses can help to overcome this prediction gap. Another reason is the limited availability of ecologically relevant bioactivity data for most of the molecules. Many compounds are described based on their chemical novelty or with respect to a specifically targeted activity, such as medically relevant activities (Roemer et al., 2011; Wohlleben et al., 2016).

This is another limitation explaining the observation of the number of biotechnological processes described for large-scale production from marine fungi compared to number of newly-discovered NPs. The current developments in metabolic engineering and marine microbiology have not yet matured into broadly-applicable processes, but offer numerous options for improvement of production processes and establishment of new process chains (Silber et al., 2016).

3.3. Inclusion of “omics” for improved prediction and production

In order to establish the overall potential of marine fungi for NP biosynthesis, it is important to understand the phylogenetic diversity of marine fungi, the biosynthetic potential of the species and strains and the phylogeny of NPs biosynthesis. Therefore, much emphasis has to be

given to the determination of the phylogenetic position of the fungi under investigation in order to enable correlation of the phylogenetic relationship to secondary metabolite production or vice versa (Kramer et al., 2016). The approach must be carefully selected with respect to the targeted phylogenetic group, as shown in the discussion on genetic markers above.

Comprehensive knowledge from the genome to the metabolome level contributes not only to a general understanding of the potential of fungi in drug discovery, but also helps to crystallise optimisation strategies for biotechnological processes. The analyses on proteome level (proteomics) may especially deliver valuable insights into the producing cell, underlying regulatory processes and angles for metabolic engineering. Transcriptomics, proteomics, and secretomics can be applied to elucidate the metabolic state of a cell on all levels of gene regulation and to indicate regulation sites on DNA, RNA, and protein level. Based on such knowledge, conditions required to induce expression of the full biosynthetic potential of an organism can be established and further be controlled. A current example shows how powerful these tools can be: a comparative proteome study on a marine *Microascus brevicaulis* strain revealed how the biotechnological fermentation process should be controlled in order to increase the production of the anticancer compounds scopularides A and B (Kramer et al., 2015). Furthermore, fluxomics, which determine the metabolic flux of primary molecules during primary and secondary metabolism in a quantitative manner, is a powerful tool to display the conversion of nutrient source into products or by-products (Knuf and Nielsen, 2012).

This knowledge can be used to design fermentation conditions or to engineer the underlying pathways by means of genetic modification. Metabolomics has the potential to deliver quantitative analyses of the metabolite profile (Silber et al., 2016).

3.4. Chemotaxonomy as the key for defining best practice

Secondary metabolite profiling has also been assessed as one of the markers for differentiation among filamentous fungi (chemotaxonomy). Moreover, there have been efforts to classify fungi based on secondary metabolite profiles (Frisvad et al., 2008). The major objective of any taxonomic study includes systematic grouping of taxa of interest through generation of robust natural classification based on constant characteristics, which reveal their factual evolutionary record and development of trustworthy identification for uncomplicated taxon determination. As outlined, molecular taxonomic methods are the most widely accepted tools for identification of the appropriate taxonomic level of microbes. The delivered taxonomic information often has some limitations with respect to resolution: species or even strain level cannot always be elucidated. As a complementary method, chemotaxonomy may utilize secondary metabolites to distinguish fungi at the species or even strain level (Frisvad et al., 2008), which can be applied to strains available in culture. With the progress of modern chromatographic techniques, metabolite profiling is easy, fast and reliable, which has resulted in increasing secondary metabolite data. Conversely, this method has a drawback as the production of these secondary metabolites can be affected by different factors such as environmental conditions, temperature and pH (Guarro et al., 1999). Adaptation to identical environmental conditions as well as phylogenetic relatedness can lead to consistent secondary metabolite profiles.

The presence of genes for the biosynthesis of a given type of secondary metabolite may or may not be limited to phylogenetic groups, as known e.g. for the species-specific secondary metabolite production in the genus *Penicillium* (Smedsgaard and Frisvad, 1996). Several studies on the chemotaxonomy of a number of fungal genera have been reported (Kang et al., 2011; Rai et al., 2014); however, these studies did not identify the metabolites that determined different clustering patterns. Metabolite profiling is often used to identify unknown compounds from complex data because it makes the simultaneous observation of all kinds of metabolites possible, rather than affiliating specific compounds to specific taxonomic questions. Chemotaxonomy could also be quite useful for researchers in the field of NPs: incorporating chemotaxonomy into fungal NPs research is very useful for comparing strains and dereplication of known metabolites (Bugni and Ireland, 2004; Overy et al., 2014). Accordingly, modern chemotaxonomy might develop into a field, where a strong interaction between phylogenists and chemists would be advantageous for both disciplines. Kramer et al. provide an example of how data can be integrated: chemical profiling of the anticancer active compound scopulariside and its derivatives in *Microascus/Scopulariopsis* strains indicated that all strains obviously share the gene cluster for scopulariside A production (2016). No other genera are known from the literature producing the scopularides. This suggests an almost mutually exclusive distribution of the scopulariside production in the genus of *Scopulariopsis/Microascus*. Hence, this molecule could serve as phylogenetic marker for the *Scopulariopsis/Microascus* group. The putative biosynthetic gene cluster of scopulariside A was identified based on the sequenced genome of strain LF580 (Lukassen et al., 2015). Furthermore, the molecule could be related to the asexual lifestyle as high production titers were found only in a clade comprising strains that show only asexual behaviour under lab conditions. This could link further studies to functional ecology.

3.5. "omic" techniques can improve biotechnology of marine fungal natural products

Biotechnology of marine fungal NPs can take advantage of genome mining; varying from the simple search for compounds with predicted physicochemical properties up to methods that exploit a probable interaction of microorganisms. Until now, the majority of successful approaches have utilized molecular biology in the forms of gene knockouts, promoter exchange, overexpression of transcription factors or other pleiotropic regulators. Moreover, strategies based on epigenetics opened a new avenue for the elucidation of the regulation of secondary metabolite formation and will certainly continue to play a significant role in the characterisation of cryptic NPs. The conditions under which a given gene cluster is naturally expressed are largely unknown. One technique is to simulate the natural habitat by co-cultivation of microorganisms from the same ecosystem. This has already led to the activation of silent gene clusters and the identification of novel compounds in *Aspergillus nidulans* (Bergmann et al., 2010). These simulation strategies will help discover new NPs in the future, and may also provide fundamental new insights into microbial communication. Metabolic engineering has the potential to be used for large scale production of these compounds using rational biochemical designs. There are reports of implementation of metabolic engineering for production of marine fungal secondary metabolites (Yanai et al., 2004; Silber et al., 2016).

4. Integrating fungal molecular ecology and natural product research

4.1. Wanted: strategies to merge the two "worlds"

Bringing together fungal molecular ecology and NP research in the field of marine fungi will benefit both fields and will facilitate new research directions. The integration, however, lies in the future. The fields need to find a multidisciplinary approach allowing a comprehensive view on marine fungi and their chemical potential. The establishment of the MaFNAp Consortium as a network of research groups with interests in the development of this evolving field of research in August 2014 was a first important step in this direction. The consortium has led to the development of collaborative projects, accompanied by biannual conferences. However, additional steps must follow, such as agreements on a common language. Such consortia could be the starting points for development of common rules that enable access to strains, compounds and data in formats needed for a comprehensive

Box III
Glossary

Chemical profiling (metabolite profiling) Connected to dereplication and defined as the systematic detection and/or identification and/or quantification of a range of natural compounds. Depending on the final purpose and on the analytical technique available, chemical profiling can be performed as a fingerprint analysis or as a targeted profile. A fingerprint analysis represents a global analysis where all detected metabolites are not necessarily identified. In a targeted chemical profiling approach, a predefined number of compounds or a particular chemical class of compounds is investigated and the molecular structures are identified.

Chemotaxonomy The use of chemical profiling as classification tool enabling identification and classification of organisms based on profiles of secondary metabolites.

Coalescent theory A probabilistic theory to identify the evolutionary history of alleles as a basis for population genetic processes.

Dereplication The rapid identification of known compounds present in a mixture. Historically, the first definition of the term "dereplication" was given by Beutler et al. (1990) as "a process of quickly identifying known chemotypes". Their goal was to evaluate the activity of a range of terrestrial and marine plant extracts by using a simple phorbol dibutyrate (PDBu) receptor binding assay and to rapidly identify compounds responsible for this activity without investing time in traditional bioassay- guided fractionation or full structure elucidation procedures. Today used in bioactivity screening campaigns for the identification of known compounds based on a combination of biological activity and physicochemical properties and/or the grouping of samples containing similar compounds likely to be responsible for the biological activity detected.

DNA Barcoding A common approach in biodiversity assessments, where a short marker gene (in the range of 400–800 bp) is used for fungal taxon identification. The proposed genetic region for fungal identification is the Internal Transcribed Spacers (ITS) (Schoch et al., 2012) though wide differences in its resolution power for the diverse fungal groups exist.

Fluxomics The determination of the metabolic flux of primary molecules during primary and secondary metabolism, in a quantitative manner.

Marker gene A marker gene is a precise genetic region used to identify and/or classify the queried fungal organisms. Marker genes differ in the observed sequence variability among fungal groups and thus are suitable for the resolution on different taxonomic levels. Fungal classification has to be based on phylogenetic markers to infer a robust phylogeny.

Metabarcoding High-throughput- and DNA-based identification of multiple species from a complex and/or environmental sample.

Metabolomics Defined as the qualitative and quantitative analysis of the whole set of low molecular weight metabolites present in a biological system (Fukushima and Kusano, 2013), also falls within the scope of chemical profiling. However, more specifically, metabolomics is an interdisciplinary field that combines high resolution analytical systems, multivariate statistics and data mining tools, chemical and biological knowledge, and sometimes metabolic network modeling in an attempt to understand metabolic pathways, gene-function relationships, or states of an organism in response to environmental changes, drug perturbations, phylogeny, genotypic or phenotypic variabilities. Metabolomics workflows are strongly context-dependent and have been used over the last years in so many different fields that a range of experimental design methodologies have been developed with specific lines regarding sample preparation, data analysis or metabolite identification strategies (Hubert et al., 2015).

Metagenomics A functional and sequence-based analysis of the collective microbial genomes contained in an environmental sample. The goal is the assembly of whole organism genomes and the finding of most common functional genes in a taxonomic context.

Natural product (NP) The extent to which the term NP has been characterised is both limited and debatable. Therefore, a common definition that is accepted by all involved disciplines will remain a moving target. In the simplest of terms, a NP is a small molecule that is produced by a biological source. As a central theme of exploration bordering chemistry and biology, NP research focuses on the chemical properties, biosynthesis and biological functions of secondary metabolites. In this context, the task of defining "natural" is more straight forward and encompasses isolation from a native organism, synthesis in a laboratory, biosynthesis in vitro, or isolation from a metabolically engineered organism whereby the chemical structure has been determined and the resultant compound is chemically equivalent to the original NP. In summary, a NP is a pharmacologically or biologically active chemical compound or substance, which is found in nature and produced by a living organism and can even be considered as such if it can be prepared by a totally synthetic approach (Natural Product. Feb 7, 2013. Available from: http://www.thefreedictionary.com/Natural+product.). In most cases, NPs are considered as small molecules with molecular weights below appr. 1000 Da, however, high molecular weight substances such as polysaccharides and proteins do also show therapeutic potential.

"omic" techniques Comprehensive approaches for the description of the different biological levels of a cell (DNA - genome, RNA - transcriptome, proteins - proteome, metabolites - metabolome). They all share novel and rapid methods for the broad analysis of the given level (the aim is to describe the totality of e.g. all genes, all proteins, etc.) and the generation of big datasets. A combination of these approaches will lead to a deep understanding of the metabolic potential and the regulation on cellular level.

Phylogenetics The study of evolutionary relatedness of organism groups, or molecular data based on observed differences (morphological/physical/genetic). The evolutionary relatedness can be inferred by diverse phylogenetic methods and tools. The outcome is a phylogenetic tree. However, inferring a phylogeny is always an estimation procedure that makes a best estimate of an evolutionary history based on incomplete information.

Phylogeny-based approach A phylogeny-based approach uses phylogenetic metrics to analyse differences between communities. Thus, it can account for highly divergent and phylogenetically unique taxonomic units and recognizes closely related groups of species that would be treated equivalently in a sequence similarity-based approach (Martin, 2002; Lozupone and Knight, 2005). This is of importance as environmental parameters often constrain/support phylogenetic structure (Enquist et al., 2002) and non-random processes occur (Emerson et al., 2011). For

example, phylogenetic clustering and overdispersion are often caused by environmental filtering or competitive exclusion (Horner-Devine and Bohannan, 2006; Webb et al., 2008).

Phylogenetic placement A method used for integrating sequences into an existing phylogenetic reference tree. For this method, the topology of the tree is not changed and integration is based on phylogenetic methods.

Phylogenetic tree A branching diagram in form of a tree demonstrating the evolutionary relationship among e.g. of species. Branches of the tree interconnect nodes. The nodes represent species while the branch determines the ancestor relationship among nodes. Dependent on the method chosen to calculate the phylogenetic tree, the branch length can indicate the amount of evolution such as mutation events.

Secondary metabolite The term "natural product" is generally regarded as being synonymous with "secondary metabolite" (Kinghorn et al., 2009). Secondary metabolites are organic compounds in the correct chiral (i.e. stereochemical) configuration to exert biological activity, but have no "primary" function directly involved in the normal growth, development or reproduction of an organism (Zähner, 1979). Secondary metabolites are usually relatively small molecules with a molecular weight below 3000 Da and exhibit considerable structural diversity (Krause and Tobin, 2013).

Sequence similarity-based approach A sequence similarity-based approach is used to identify a sequence in a reference database that shows a high sequence similarity to the query relatively quickly. This search is conducted over pairwise alignments on very short stretches of the query sequence, which is iteratively elongated. The pairwise alignments are scored by sequence similarity without incorporation of an evolutionary model (for example mutation rates, base occurrence etc.).

Species delimitation The process of determining the boundaries of a species concept from empirical data.

Structural elucidation Determination of the chemical structure of a secondary metabolite by means of nuclear magnetic resonance (NMR) spectroscopic data, exact mass measurements, polarimetry etc., in many contexts this includes the determination of the stereostructure of the molecule, which often determines biological activity.

Taxonomy Taxonomy is the science of classifying organisms into groups that are evolutionarily-related and form a taxon. The classification is based on shared characteristics. Nowadays, molecular data are used for fungal taxonomy where a character is defined as nucleotides of the same alignment column of a multiple sequence alignment. Thus, taxonomy is defined on the basis of phylogenetic analysis where a taxon has to be represented as a monophyletic clade (Yarza et al., 2014). Due to the evolving number of publicly available sequences fungal taxonomy is under regular reconsideration. This can have even consequences on a higher taxonomic rank such as a phylum or subphylum (Spatafora et al., 2016).

analysis of the data. Training and education of the next generation of marine fungal researchers could help to sustain the collaboration and enable true transdisciplinary approaches.

4.2. Common vocabulary and repositories

As outlined, the two disciplines benefit from closer interaction but the languages, quality criteria and directions for scientific investigation differ, often without appreciation by the researchers affected. Limited communication will always lead to limited outcomes; both scientific communities will learn that they need to understand each other. Only rare examples, as the Hubert et al. (2015) review on dereplication show the way for best practices, how such vocabularies can lead to a consequent, standardised use of terms being developed in a variety of different research fields. Therefore, a common vocabulary further developed during common conferences and the use of the same sources and databases would enable progress.

4.2.1. Common vocabulary

Integration of the two "worlds" in a transdisciplinary matter does not mean learning all the techniques and knowledge of the other party to expert level, but to understand when the complementary expertise is needed. A prerequisite is the use of the correct vocabulary when communicating with each other. This compliance will be of benefit, for subsequent publication, for compliance to regulatory frames including the Nagoya protocol (www.cbd.int/abs/, accessed May 2017) and – most important - to frontier research. In Box III we propose a first glossary of terms to support the development of a common vocabulary.

4.2.2. Databases

For marine fungal research, most often, general fungal resources are used independent of whether the outcome should be an ecological, taxonomic or functional annotation of the queries: an exhaustive list can be found in Hibbett et al. (2016) with databases and tools for sequence-based classification and identification, genomic databases and tools and data standards. Data treatment and databases of metabolic studies, including dereplication were recently reviewed by Allard et al. (2017). However, most of the specific databases, as *Marine Lit* (http://pubs.rsc.org/marinlit/, accessed May 2017) and *Dictionary of Natural Products* (http://dnp.chemnetbase.com/dictionary-search.do?method=view&id=12018701&si=, accessed May 2017), are not open access. First steps were undertaken to combine mass spectra data for comprehensive networking studies. Such examples will pave the way for future databases (Wang et al., 2016b). The website '*omics tools*' can be used as a starting point to access the different databases (https://omictools.com/, accessed May 2017).

In any case, all databases are biased towards terrestrial fungi. Until now, no marine fungi-specific database exists that combines the diverse information that could be theoretically compiled, based on the data and knowledge owned by the diverse researcher's laboratories. A few attempts exist such as the *Freshwater Ascomycete Database* (http://fungi.life.illinois.edu/; accessed Dec. 2016) devoted solely to taxa in the phylum of *Ascomycota* and providing a general taxon background, lists reports over freshwater *Ascomycota*, and offers illustrated profiles among other services (Shearer and Raja, 2010). Another database is the *Indian Marine Fungal Database* (www.fungifromindia.com/fungiFromIndia/databases/IMFD/nextPage.php?id=home.php, accessed Dec. 2016), owning 233 entries of marine fungi found in India and is linked to *Mycobank*. The *World Register of Marine Species* (WoRMS) (www.marinespecies.org/, accessed Dec. 2016) aims to provide an authoritative and comprehensive list of names of marine organisms. A more advanced database called *Faces of Fungi* (www.facesoffungi.org/, accessed Dec. 2016) is set up to assemble all kind of information of fungi and fungal-like organisms: The included fungal profiles consist of information including industrial relevance, quarantine status, chemistry, links to sequence and culture collections, morphological description, phylogenetic information, ecology and human impacts (Jayasiri et al., 2015). Currently, the database is sparsely populated and the information provided is scarce, with a high bias towards terrestrial fungi. Additionally, connected datasets, such as habitat data, are not interlinked and must be searched elsewhere, for example in *Pangaea* (www.pangaea.de/,

accessed Dec. 2016), which is an open access library to archive, publish and distribute global georeferenced data.

4.2.3. Culture collections

Culture collections play a vital role in the conservation and sustainable use of microbial resources including fungi (Daniel and Prasad, 2010). They also provide the authentic biological material needed for referencing not only the taxonomic but also the NP research in the form of reference strains, reagents for quality control, etc. Unfortunately, only a limited number of marine fungi are currently deposited in culture collections. As a consequence, only a few voucher sequences are available although they are a prerequisite for an accurate sequence-based identification of marine fungi. Especially under the light of the continuous discovery of new microbial taxa and strains, the need to preserve and make them accessible to other researchers for research, teaching and for biotechnological exploitation, is indisputable. Individual laboratories are unable to do this due to lack of financial support and manpower. This role is thus played by culture collections (Smith, 2003) and adopted software to store the different kinds of data related to a cultured taxon (Pena and Malm, 2012). Some small specific collections of marine fungi exist (Table 1) but again, finding comprehensive resources and even more important, the compliance of researchers to deposit their strains in publicly accessible culture collections is rather rare in the field.

5. Outlook

Considering the existing knowledge and recent approaches, the following factors seem to be essential for unifying the two “worlds” of fungal molecular ecologists and NP researchers:

- Improving and expanding existing datasets on marine fungi in terms of bio- and chemo-diversity with the help of “omic” tools.
- Adherence to the existing rules for taxon naming, which is the very basis for communication and comparison of research results.
- Use of adapted identification tools: the ITS being the announced fungal DNA barcode does most often not deliver the necessary resolution for marine fungi: extension of identification protocols towards a double-marker gene approach including a phylogenetic marker.
- A phylogenetic context is mandatory for the classification of undescribed fungal taxa.
- Ecological studies should take advantage of interlinking metabolomic screening and taxonomic analysis: results can be passed on to network analysis elucidating the ecological role of secondary metabolites.
- Comprehensive databases are needed: either existing databases must be interlinked or researchers from both “worlds” have to provide substantial support to set up a new database where all results on marine fungi are made available in a comprehensive manner.
- Emphasising the deposition of marine fungal isolates in culture collections to extend the fungal biodiversity accessible by biological material.
- True transdisciplinarity: the transdisciplinary and integrative approach of developmental projects should encompass research projects. Thus, important gaps in the knowledge of marine fungi relevant for the production of fungal NPs should be actively addressed and made relevant. Successfully addressing these challenges will require the combined effort of multidisciplinary teams.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

The authors are grateful to Katrin Panzer and Stefanos Banos for help in the manual curation of the 18S rRNA gene sequence alignment and in defining the variable regions, respectively. We would like to thank two anonymous reviewers helping us to improve the quality of the manuscript and fueling the internal discussion on the manuscript.

References

Aguileta, G., Marthey, S., Chiapello, H., Lebrun, M.H., Rodolphe, F., Fournier, E., Gendrault-Jacquemard, A., Giraud, T., 2008. Assessing the performance of single-copy genes for recovering robust phylogenies. Syst. Biol. 57, 613–627.
Allard, P.M., Genta-Jouve, G., Wolfender, J.L., 2017. Deep metabolome annotation in natural products research: towards a virtuous cycle in metabolite identification. Curr. Opin. Chem. Biol. 36, 40–49.
Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410.
Amend, A.S., Barshis, D.J., Oliver, T.A., 2012. Coral-associated marine fungi form novel lineages and heterogeneous assemblages. ISME J. 6, 1291–1301.
Andrews, J.H., 1992. Fungal life-history strategies. In: Carroll, G.C., Wicklow, D.T. (Eds.), The Fungal Community: Its Organization and Role in the Ecosystem. Marcel Dekker, Inc., New York.
Ankenbrand, M.J., Keller, A., Wolf, M., Schultz, J., Forster, F., 2015. ITS2 database V: twice as much. Mol. Biol. Evol. 32, 3030–3032.
Arfi, Y., Buee, M., Marchand, C., Levasseur, A., Record, E., 2012. Multiple markers pyrosequencing reveals highly diverse and host-specific fungal communities on the mangrove trees *Avicennia marina* and *Rhizophora stylosa*. FEMS Microbiol. Ecol. 79, 433–444.
Baldauf, S.L., 1999. A search for the origins of animals and fungi: comparing and combining molecular data. Am. Nat. 154, S178–S188.
Baldauf, S.L., Roger, A.J., Wenk-Siefert, I., Doolittle, W.F., 2000. A kingdom-level phylogeny of eukaryotes based on combined protein data. Science 290, 972–977.
Bass, D., Howe, A., Brown, N., Barton, H., Demidova, M., Michelle, H., Li, L., Sanders, H., Watkinson, S.C., Willcock, S., Richards, T.A., 2007. Yeast forms dominate fungal diversity in the deep oceans. Proc. R. Soc. B 274, 3069–3077.
Berger, S.A., Krompass, D., Stamatakis, A., 2011. Performance, accuracy, and web server for evolutionary placement of short sequence reads under maximum likelihood. Syst. Biol. 60, 291–302.
Bergmann, S., Funk, A.N., Scherlach, K., Schroeckh, V., Shelest, E., Horn, U., Hertweck, C., Brakhage, A.A., 2010. Activation of a silent fungal polyketide biosynthesis pathway through regulatory cross talk with a cryptic nonribosomal peptide synthetase gene cluster. AEM 76, 8143–8149.
Beutler, J.A., Alvarado, A.B., Schaufelberger, D.E., Andrews, P., Mccloud, T.G., 1990. Dereplication of phorbol bioactives - *Lyngbya majuscula* and *Croton cuneatus*. J. Nat. Prod. 53, 867–874.
Bhadury, P., Mohammad, B.T., Wright, P.C., 2006. The current status of natural products from marine fungi and their potential as anti-infective agents. J. Ind. Microbiol. Biotechnol. 33, 325–337.
Blunt, J.W., Copp, B.R., Keyzers, R.A., Munro, M.H.G., Prinsep, M.R., 2016. Marine natural products. Nat. Prod. Rep. 33, 382–431.
Bode, H.B., Bethe, B., Hofs, R., Zeeck, A., 2002. Big effects from small changes: possible ways to explore nature's chemical diversity. Chembiochem 3, 619–627.
Brakhage, A.A., 2013. Regulation of fungal secondary metabolism. Nat. Rev. Microbiol. 11, 21–32.
Bruns, T.D., Gardes, M., 1993. Molecular tools for the identification of ectomycorrhizal fungi - taxon-specific oligonucleotide probes for suilloid fungi. Mol. Ecol. 2, 233–242.
Bugni, T.S., Ireland, C.M., 2004. Marine-derived fungi: a chemically and biologically diverse group of microorganisms. Nat. Prod. Rep. 21, 143–163.
Carbone, I., Kohn, L., 2004. Inferring process from pattern in fungal population genetics. In: Khachatourians, G.G., Arora, D.K. (Eds.), Applied Mycology and Biotechnology, Book 1. Elsevier Science, Amsterdam, Netherlands.
Cole, J.R., Wang, Q., Fish, J.A., Chai, B.L., McGarrell, D.M., Sun, Y.N., Brown, C.T., Porras-Alfaro, A., Kuske, C.R., Tiedje, J.M., 2014. Ribosomal database project: data and tools for high throughput rRNA analysis. Nucleic Acids Res. 42, D633–D642.
Crous, P.W., Gams, W., Stalpers, J.A., Robert, V., Stegehuis, G., 2004. MycoBank: an online initiative to launch mycology into the 21st century. Stud. Mycol. 19–22.
Culligan, E.P., Sleator, R.D., Marchesi, J.R., Hill, C., 2014. Metagenomics and novel gene discovery: Promise and potential for novel therapeutics. Virulence 5, 399–412.
Daniel, H.M., Prasad, G.S., 2010. The role of culture collections as an interface between providers and users: the example of yeasts. Res. Microbiol. 161, 488–496.
Das, S., Dash, H.R., Mangwani, N., Chakraborty, J., Kumari, S., 2014. Understanding molecular identification and polyphasic taxonomic approaches for genetic relatedness and phylogenetic relationships of microorganisms. J. Microbiol. Methods 103, 80–100.
de Vries, R.P., Riley, R., Wiebenga, A., Aguilar-Osorio, G., Amillis, S., Uchima, C.A., Anderluh, G., Asadollahi, M., Askin, M., Barry, K., Battaglia, E., Bayram, O., Benocci,

T., Braus-Stromeyer, S.A., Caldana, C., Canovas, D., Cerqueira, G.C., Chen, F.S., Chen, W.P., Choi, C., Clum, A., dos Santos, R.A.C., Damasio, A.R.D., Diallinas, G., Emri, T., Fekete, E., Flipphi, M., Freyberg, S., Gallo, A., Gournas, C., Habgood, R., Hainaut, M., Harispe, M.L., Henrissat, B., Hilden, K.S., Hope, R., Hossain, A., Karabika, E., Karaffa, L., Karanyi, Z., Krasevec, N., Kuo, A., Kusch, H., LaButti, K., Lagendijk, E.L., Lapidus, A., Levasseur, A., Lindquist, E., Lipzen, A., Logrieco, A.F., MacCabe, A., Makela, M.R., Malavazi, I., Melin, P., Meyer, V., Mielnichuk, N., Miskei, M., Molnar, A.P., Mule, G., Ngan, C.Y., Orejas, M., Orosz, E., Ouedraogo, J.P., Overkamp, K.M., Park, H.S., Perrone, G., Piumi, F., Punt, P.J., Ram, A.F.J., Ramon, A., Rauscher, S., Record, E., Riano-Pachon, D.M., Robert, V., Rohrig, J., Ruller, R., Salamov, A., Salih, N.S., Samson, R.A., Sandor, E., Sanguinetti, M., Schutze, T., Sepcic, K., Shelest, E., Sherlock, G., Sophianopoulou, V., Squina, F., Sun, H., Susca, A., Todd, R.B., Tsang, A., Unkles, S.E., van de Wiele, N., van Rossen-Uffink, D., Oliveira, J.V.D., Vesth, T.C., Visser, J., Yu, J.H., Zhou, M.M., Andersen, M.R., Archer, D.B., Baker, S.E., Benoit, I., Brakhage, A.A., Braus, G.H., Fischer, R., Frisvad, J.C., Goldman, G.H., Houbraken, J., Oakley, B., Pocsi, I., Scazzocchio, C., Seiboth, B., vanKuyk, P.A., Wortman, J., Dyer, P.S., Grigoriev, I.V., 2017. Comparative genomics reveals high biological diversity and specific adaptations in the industrially and medically important fungal genus *Aspergillus*. Genome Biol. 18.

Debbab, A., Aly, A.H., Lin, W.H., Proksch, P., 2010. Bioactive compounds from marine Bacteria and Fungi. Microb. Biotechnol. 3, 544–563.

den Bakker, H.C., Gravendeel, B., Kuyper, T.W., 2004. An ITS phylogeny of *Leccinum* and an analysis of the evolution of minisatellite-like sequences within ITS1. Mycologia 96, 102–118.

Drew, B.T., Gazis, R., Cabezas, P., Swithers, K.S., Deng, J.B., Rodriguez, R., Katz, L.A., Crandall, K.A., Hibbett, D.S., Soltis, D.E., 2013. Lost branches on the tree of life. PLoS Biol. 11.

Dupont, J., Magnin, S., Cesari, C., Gatica, M., 2002. ITS and beta-tubulin markers help delineate *Phaeoacremonium* species, and the occurrence of *P. parasiticum* in grapevine disease in Argentina. Mycol. Res. 106, 1143–1150.

Dupont, S., Wilson, K., Obst, M., Skold, H., Nakano, H., Thorndyke, M.C., 2007. Marine ecological genomics: when genomics meets marine ecology. Mar. Ecol. Prog. Ser. 332, 257–273.

Eberhardt, U., Walter, L., Kottke, I., 1999. Molecular and morphological discrimination between *Tylospora fibrillosa* and *Tylospora asterophora* mycorrhizae. Can. J. Bot. 77, 11–21.

Emerson, B.C., Cicconardi, F., Fanciulli, P.P., Shaw, P.J.A., 2011. Phylogeny, phylogeography, phylobetadiversity and the molecular analysis of biological communities. Phil. Trans. R. Soc. B 366, 2391–2402.

Enquist, B.J., Haskell, J.P., Tiffney, B.H., 2002. General patterns of taxonomic and biomass partitioning in extant and fossil plant communities. Nature 419, 610–613.

Feau, N., Decourcelle, T., Husson, C., Desprez-Loustau, M.L., Dutech, C., 2011. Finding single copy genes out of sequenced genomes for multilocus phylogenetics in non-model Fungi. PLoS One 6.

Federhen, S., 2012. The NCBI taxonomy database. Nucleic Acids Res. 40, D136–D143.

Fell, J.W., Boekhout, T., Fonseca, A., Scorzetti, G., Statzell-Tallman, A., 2000. Biodiversity and systematics of basidiomycetous yeasts as determined by large-subunit rDNA D1/D2 domain sequence analysis. Int. J. Syst. Evol. Microbiol. 50, 1351–1371.

Fries, N., 1979. Physiological characteristics of *Mycosphaerella ascophylli*, a fungal endophyte of the marine brown alga *Ascophyllum nodosum*. Physiol. Plant. 45, 117–121.

Frisvad, J.C., Andersen, B., Thrane, U., 2008. The use of secondary metabolite profiling in chemotaxonomy of filamentous fungi. Mycol. Res. 112, 231–240.

Fujita, M.K., Leache, A.D., Burbrink, F.T., McGuire, J.A., Moritz, C., 2012. Coalescent-based species delimitation in an integrative taxonomy. Trends Ecol. Evol. 27, 480–488.

Fukushima, A., Kusano, M., 2013. Recent progress in the development of metabolome databases for plant systems biology. Front. Plant Sci. 4.

Gams, W., 1980. *Chaunopycnis alba*, gen et sp nov, a soil fungus intermediate between *Moniliales* and *Sphaeropsidales*. Persoonia 11, 75–79.

Garbary, D.J., Macdonald, K.A., 1995. The *Ascophyllum polysiphonia Mycosphaerella* symbiosis .4. Mutualism in the *Ascophyllum mycosphaerella* interaction. Bot. Mar. 38, 221–225.

Gram, L., 2015. Silent clusters - speak up!. Microb. Biotechnol. 8, 13–14.

Graybeal, A., 1994. Evaluating the phylogenetic utility of genes - a search for genes informative about deep divergences among vertebrates. Syst. Biol. 43, 174–193.

Grigoriev, I.V., Nikitin, R., Haridas, S., Kuo, A., Ohm, R., Otillar, R., Riley, R., Salamov, A., Zhao, X.L., Korzeniewski, F., Smirnova, T., Nordberg, H., Dubchak, I., Shabalov, I., 2014. MycoCosm portal: gearing up for 1000 fungal genomes. Nucleic Acids Res. 42, D699–D704.

Guarro, J., Gene, J., Stchigel, A.M., 1999. Developments in fungal taxonomy. Clin. Microbiol. Rev. 12, 454.

Gueidan, C., Roux, C., Lutzoni, F., 2007. Using a multigene phylogenetic analysis to assess generic delineation and character evolution in *Verrucariaceae* (*Verrucariales, Ascomycota*). Mycol. Res. 111, 1145–1168.

Guillou, L., Bachar, D., Audic, S., Bass, D., Berney, C., Bittner, L., Boutte, C., Burgaud, G., de Vargas, C., Decelle, J., del Campo, J., Dolan, J.R., Dunthorn, M., Edvardsen, B., Holzmann, M., Kooistra, W.H.C.F., Lara, E., Le Bescot, N., Logares, R., Mahe, F., Massana, R., Montresor, M., Morard, R., Not, F., Pawlowski, J., Probert, I., Sauvadet, A.L., Siano, R., Stoeck, T., Vaulot, D., Zimmermann, P., Christen, R., 2013. The Protist Ribosomal Reference database (PR2): a catalog of unicellular eukaryote Small Sub-Unit rRNA sequences with curated taxonomy. Nucleic Acids Res. 41, D597–D604.

Gutierrez, M.H., Jara, A.M., Pantoja, S., 2016. Fungal parasites infect marine diatoms in the upwelling ecosystem of the Humboldt current system off central Chile. Environ. Microbiol. 18, 1646–1653.

Hassett, B.T., Gradinger, R., 2016. Chytrids dominate arctic marine fungal communities. Environ. Microbiol. 18, 2001–2009.

Hawksworth, D.L., 2012a. Global species numbers of fungi: are tropical studies and molecular approaches contributing to a more robust estimate? Biodivers. Conserv. 21, 2425–2433.

Hawksworth, D.L., 2012b. Managing and coping with names of pleomorphic fungi in a period of transition. Mycosphere 3, 143–155.

Hawksworth, D.L., 2015. Proposals to clarify and enhance the naming of fungi under the *International Code of Nomenclature for algae, fungi, and plants*. IMA Fungus 6, 199–205.

Helgason, T., Watson, I.J., Young, J.P.W., 2003. Phylogeny of the *Glomerales* and *Diversisporales* (Fungi: *Glomeromycota*) from actin and elongation factor 1-alpha sequences. FEMS Microbiol. Lett. 229, 127–132.

Hibbett, D.S., Abarenkov, K., Koljalg, U., Öpik, M., Chai, B., Cole, J.R., Wang, Q., Crous, P.W., Robert, V., Helgason, T., Herr, J.R., Kirk, P.M., Lueschow, S., O'Donnell, K., Nilsson, R.H., Oono, R., Schoch, C., Smyth, C., Walker, D.M., Porras-Alfaro, A., Taylor, J.W., Geiser, D.M., 2016. Sequence-based classification and identification of Fungi. Mycologia 108, 1049–1068.

Hinchliff, C.E., Smith, S.A., Allman, J.F., Burleigh, J.G., Chaudhary, R., Coghill, L.M., Crandall, K.A., Deng, J., Drew, B.T., Gazis, R., Gude, K., Hibbett, D.S., Katz, L.A., Laughinghouse, H.D., McTavish, E.J., Midford, P.E., Owen, C.L., Ree, R.H., Rees, J.A., Soltis, D.E., Williams, T., Cranston, K.A., 2015. Synthesis of phylogeny and taxonomy into a comprehensive tree of life. Proc. Natl. Acad. Sci. U. S. A. 112, 12764–12769.

Hirt, R.P., Logsdon, J.M., Healy, B., Dorey, M.W., Doolittle, W.F., Embley, T.M., 1999. Microsporidia are related to Fungi: Evidence from the largest subunit of RNA polymerase II and other proteins. Proc. Natl. Acad. Sci. U. S. A. 96, 580–585.

Hofstetter, V., Miadlikowska, J., Kauff, F., Lutzoni, F., 2007. Phylogenetic comparison of protein-coding versus ribosomal RNA-coding sequence data: A case study of the *Lecanoromycetes* (*Ascomycota*). Mol. Phylogenet. Evol. 44, 412–426.

Horner-Devine, M.C., Bohannan, B.J.M., 2006. Phylogenetic clustering and overdispersion in bacterial communities. Ecology 87, S100–S108.

Hubert, J., Nuzillard, J.M., Renault, J.H., 2015. Dereplication strategies in natural product research: How many tools and methodologies behind the same concept? Phytochem. Rev. 1–41.

Hubka, V., Kolarik, M., 2012. beta-tubulin paralogue *tubC* is frequently misidentified as the *benA* gene in *Aspergillus* section *Nigri* taxonomy: primer specificity testing ad taxonomic consequences. Persoonia 29, 1–10.

Imhoff, J.F., 2016. Natural products from marine fungi - still an underrepresented resource. Mar. Drugs 14.

James, T.Y., Kauff, F., Schoch, C.L., Matheny, P.B., Hofstetter, V., Cox, C.J., Celio, G., Gueidan, C., Fraker, E., Miadlikowska, J., Lumbsch, H.T., Rauhut, A., Reeb, V., Arnold, A.E., Amtoft, A., Stajich, J.E., Hosaka, K., Sung, G.H., Johnson, D., O'Rourke, B., Crockett, M., Binder, M., Curtis, J.M., Slot, J.C., Wang, Z., Wilson, A.W., Schüssler, A., Longcore, J.E., O'Donnell, K., Mozley-Standridge, S., Porter, D., Letcher, P.M., Powell, M.J., Taylor, J.W., White, M.M., Griffith, G.W., Davies, D.R., Humber, R.A., Morton, J.B., Sugiyama, J., Rossman, A.Y., Rogers, J.D., Pfister, D.H., Hewitt, D., Hansen, K., Hambleton, S., Shoemaker, R.A., Kohlmeyer, J., Volkmann-Kohlmeyer, B., Spotts, R.A., Serdani, M., Crous, P.W., Hughes, K.W., Matsuura, K., Langer, E., Langer, G., Untereiner, W.A., Lucking, R., Budel, B., Geiser, D.M., Aptroot, A., Diederich, P., Schmitt, I., Schultz, M., Yahr, R., Hibbett, D.S., Lutzoni, F., DJ, McLaughlin, Spatafora, J.W., Vilgalys, R., 2006. Reconstructing the early evolution of Fungi using a six-gene phylogeny. Nature 443, 818–822.

James, T.Y., Pelin, A., Bonen, L., Ahrendt, S., Sain, D., Corradi, N., Stajich, J.E., 2013. Shared signatures of parasitism and phylogenomics unite Cryptomycota and Microsporidia. Curr. Biol. 23, 1548–1553.

Jayasiri, S.C., Hyde, K.D., Ariyawansa, H.A., Bhat, J., Buyck, B., Cai, L., Dai, Y.C., Abd-Elsalam, K.A., Ertz, D., Hidayat, I., Jeewon, R., Jones, E.B.G., Bahkali, A.H., Karunarathna, S.C., Liu, J.K., Luangsa-Ard, J.J., Lumbsch, H.T., Maharachchikumbura, S.S.N., McKenzie, E.H.C., Moncalvo, J.M., Ghobad-Nejhad, M., Nilsson, H., Pang, K.L., Pereira, O.L., Phillips, A.J.L., Raspe, O., Rollins, A.W., Romero, A.I., Etayo, J., Selcuk, F., Stephenson, S.L., Suetrong, S., Taylor, J.E., Tsui, C.K.M., Vizzini, A., Abdel-Wahab, M.A., Wen, T.C., Boonmee, S., Dai, D.Q., Daranagama, D.A., Dissanayake, A.J., Ekanayaka, A.H., Fryar, S.C., Hongsanan, S., Jayawardena, R.S., Li, W.J., Perera, R.H., Phookamsak, R., De Silva, N.I., Thambugala, K.M., Tian, Q., Wijayawardene, N.N., Zhao, R.L., Zhao, Q., Kang, J.C., Promputtha, I., 2015. The faces of fungi database: fungal names linked with morphology, phylogeny and human impacts. Fungal Divers. 74, 3–18.

Jones, M.D.M., Forn, I., Gadelha, C., Egan, M.J., Bass, D., Massana, R., Richards, T.A., 2011. Discovery of novel intermediate forms redefines the fungal tree of life. Nature 474, 200–U234.

Kang, D., Kim, J., Choi, J.N., Liu, K.H., Lee, C.H., 2011. Chemotaxonomy of *Trichoderma* spp. using mass spectrometry-based metabolite profiling. J. Microbiol. Biotech. 21, 5–13.

Kappe, R., Fauser, C., Okeke, C.N., Maiwald, M., 1996. Universal fungus-specific primer systems and group-specific hybridization oligonucleotides for 18S rDNA. Mycoses 39, 25–30.

Keeling, P.J., Inagaki, Y., 2004. A class of eukaryotic GTPase with a punctate distribution suggesting multiple functional replacements of translation elongation factor 1 *alpha*. Proc. Natl. Acad. Sci. U S A 101, 15380–15385.

Keller, N.P., Turner, G., Bennett, J.W., 2005. Fungal secondary metabolism - from biochemistry to genomics. Nat. Rev. Microbiol. 3, 937–947.

Kinghorn, A.D., Chin, Y.W., Swanson, S.M., 2009. Discovery of natural product anticancer agents from biodiverse organisms. Curr. Opin. Drug Disc. 12, 189–196.

Kirk, P.M., Cannon, P.F., David, J.C., Stalpers, J.A., 2008. Ainsworth and Bisby's dictionary of the fungi. CABI Publishing, Wallingford, UK.

Knuf, C., Nielsen, J., 2012. Aspergilli: systems biology and industrial applications. Biotechnol. J. 7.

Koljalg, U., Tammi, H., Timonen, S., Agerer, R., Sen, R., 2002. ITS rDNA nucleotide sequence-based phylogenetic analysis of *Tomentellopsis* species from boreal and

temperate forests, and the identification of pink-type ectomycorrhizas. Mycol. Prog. 1, 81–92.
Koljalg, U., Nilsson, R.H., Abarenkov, K., Tedersoo, L., Taylor, A.F.S., Bahram, M., Bates, S.T., Bruns, T.D., Bengtsson-Palme, J., Callaghan, T.M., Douglas, B., Drenkhan, T., Eberhardt, U., Duenas, M., Grebenc, T., Griffith, G.W., Hartmann, M., Kirk, P.M., Kohout, P., Larsson, E., Lindahl, B.D., Luecking, R., Martin, M.P., Matheny, P.B., Nguyen, N.H., Niskanen, T., Oja, J., Peay, K.G., Peintner, U., Peterson, M., Poldmaa, K., Saag, L., Saar, I., Schuessler, A., Scott, J.A., Senes, C., Smith, M.E., Suija, A., Taylor, D.L., Telleria, M.T., Weiss, M., Larsson, K.H., 2013. Towards a unified paradigm for sequence-based identification of fungi. Mol. Ecol. 22, 5271–5277.
Koljalg, U., Tedersoo, L., Nilsson, R.H., Abarenkov, K., 2016. Digital identifiers for fungal species. Science 352, 1182–1183.
Koren, S., Schatz, M.C., Walenz, B.P., Martin, J., Howard, J.T., Ganapathy, G., Wang, Z., Rasko, D.A., McCombie, W.R., Jarvis, E.D., Phillippy, A.M., 2012. Hybrid error correction and de novo assembly of single-molecule sequencing reads. Nat. Biotechnol. 30, 692.
Kozakiewicz, Z., Frisvad, J.C., Hawksworth, D.L., Pitt, J.I., Samson, R.A., Stolk, A.C., 1992. Proposals for *nomina specifica conservanda* and *rejicienda* in *Aspergillus* and *Penicillium* (*Fungi*). Taxon 41, 109–113.
Kramer, A., Beck, H.C., Kumar, A., Kristensen, L.P., Imhoff, J.F., Labes, A., 2015. Proteomic analysis of anti-cancerous scopularide production by a marine *Microascus brevicaulis* strain and its UV mutant. PLoS One 10.
Kramer, A., Labes, A., Imhoff, J.F., 2016. Phylogenetic relationship and secondary metabolite production of marine fungi producing the cyclodepsipeptides scopularide A and B. Mar. Biotechnol. 18, 466–474.
Krause, J., Tobin, G., 2013. Discovery, development, and regulation of natural products. In: Kulka, M. (Ed.), Using Old Solutions to New Problems - Natural Drug Discovery in the 21st Century. InTech, Rijeka, Croatia.
Krause, E., Wichels, A., Erler, R., Gerdts, G., 2013. Study on the effects of near-future ocean acidification on marine yeasts: a microcosm approach. Helgol. Mar. Res. 67, 607–621.
Kristensen, R., Torp, M., Kosiak, B., Holst-Jensen, A., 2005. Phylogeny and toxigenic potential is correlated in *Fusarium* species as revealed by partial translation elongation factor 1 alpha gene sequences. Mycol. Res. 109, 173–186.
Kurtzman, C.P., Fell, J.W., Boekhout, T., 2011. The Yeasts, a Taxonomic Study. vol. 3.
Lai, X.T., Cao, L.X., Tan, H.M., Fang, S., Huang, Y.L., Zhou, S.N., 2007. Fungal communities from methane hydrate-bearing deep-sea marine sediments in South China Sea. ISME J. 1, 756–762.
Lara, E., Moreira, D., Lopez-Garcia, P., 2010. The environmental clade LKM11 and Rozella form the deepest branching clade of fungi. Protist 161, 116–121.
Lefevre, E., Bardot, C., Noel, C., Carrias, J.F., Viscogliosi, E., Amblard, C., Sime-Ngando, T., 2007. Unveiling fungal zooflagellates as members of freshwater picoeukaryotes: evidence from a molecular diversity study in a deep meromictic lake. Environ. Microbiol. 9, 61–71.
Liu, Y.J.J., Whelen, S., Benjamin, D.H., 1999. Phylogenetic relationships among ascomycetes: evidence from an RNA polymerase II subunit. Mol. Biol. Evol. 16, 1799–1808.
Liu, K.L., Porras-Alfaro, A., Kuske, C.R., Eichorst, S.A., Xie, G., 2012. Accurate, rapid taxonomic classification of fungal large-subunit rRNA genes. AEM 78, 1523–1533.
Lozupone, C., Knight, R., 2005. UniFrac: a new phylogenetic method for comparing microbial communities. AEM 71, 8228–8235.
Ludwig, W., Strunk, O., Westram, R., Richter, L., Meier, H., Yadhukumar, Buchner A., Lai, T., Steppi, S., Jobb, G., Forster, W., Brettske, I., Gerber, S., Ginhart, A.W., Gross, O., Grumann, S., Hermann, S., Jost, R., Konig, A., Liss, T., Lussmann, R., May, M., Nonhoff, B., Reichel, B., Strehlow, R., Stamatakis, A., Stuckmann, N., Vilbig, A., Lenke, M., Ludwig, T., Bode, A., Schleifer, K.H., 2004. ARB: a software environment for sequence data. Nucleic Acids Res. 32, 1363–1371.
Lukassen, M.B., Saei, W., Sondergaard, T.E., Tamminen, A., Kumar, A., Kempken, F., Wiebe, M.G., Sorensen, J.L., 2015. Identification of the scopularide biosynthetic gene cluster in *Scopulariopsis brevicaulis*. Mar. Drugs 13, 4331–4343.
Lutzoni, F., Kauff, F., Cox, C.J., McLaughlin, D., Celio, G., Dentinger, B., Padamsee, M., Hibbett, D., James, T.Y., Baloch, E., Grube, M., Reeb, V., Hofstetter, V., Schoch, C., Arnold, A.E., Miadlikowska, J., Spatafora, J., Johnson, D., Hambleton, S., Crockett, M., Shoemaker, R., Hambleton, S., Crockett, M., Shoemaker, R., Sung, G.H., Lucking, R., Lumbsch, T., O'Donnell, K., Binder, M., Diederich, P., Ertz, D., Gueidan, C., Hansen, K., Harris, R.C., Hosaka, K., Lim, Y.W., Matheny, B., Nishida, H., Pfister, D., Rogers, J., Rossman, A., Schmitt, I., Sipman, H., Stone, J., Sugiyama, J., Yahr, R., Vilgalys, R., 2004. Assembling the fungal tree of life: progress, classification and evolution of subcellular traits. Am. J. Bot. 91, 1446–1480.
Mahe, S., Duhamel, M., Le Calvez, T., Guillot, L., Sarbu, L., Bretaudeau, A., Collin, O., Dufresne, A., Kiers, E.T., Vandenkoornhuyse, P., 2012. PHYMYCO-DB: a curated database for analyses of fungal diversity and evolution. PLoS One 7.
Manohar, C.S., Raghukumar, C., 2013. Fungal diversity from various marine habitats deduced through culture-independent studies. FEMS Microbiol. Lett. 341, 69–78.
Martin, A.P., 2002. Phylogenetic approaches for describing and comparing the diversity of microbial communities. AEM 68, 3673–3682.
Matheny, P.B., 2005. Improving phylogenetic inference of mushrooms with *RPB1* and *RPB2* nucleotide sequences (*Inocybe*; *Agaricales*). Mol. Phylogenet. Evol. 35, 1–20.
Matheny, P.B., Liu, Y.J.J., Ammirati, J.F., Hall, B.D., 2002. Using *RPB1* sequences to improve phylogenetic inference among mushrooms (*Inocybe, Agaricales*). Am. J. Bot. 89, 688–698.
Matheny, P.B., Gossmann, J.A., Zalar, P., Kumar, T.K.A., Hibbett, D.S., 2006. Resolving the phylogenetic position of the *Wallemiomycetes*: an enigmatic major lineage of *Basidiomycota*. Can. J. Bot. 84, 1794–1805.
Matheny, P.B., Wang, Z., Binder, M., Curtis, J.M., Lim, Y.W., Nilsson, R.H., Hughes, K.W., Hofstetter, V., Ammirati, J.F., Schoch, C.L., Langer, E., Langer, G., McLaughlin, D.J., Wilson, A.W., Froslev, T., Ge, Z.W., Kerrigan, R.W., Slot, J.C., Yang, Z.L., Baroni, T.J., Fischer, M., Hosaka, K., Matsuura, K., Seidl, M.T., Vauras, J., Hibbett, D.S., 2007. Contributions of *rpb2* and *tef1* to the phylogeny of mushrooms and allies (Basidiomycota, Fungi). Mol. Phylogenet. Evol. 43, 430–451.
Matsen, F.A., Kodner, R.B., Armbrust, E.V., 2010. pplacer: linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. BMC Bioinforma. 11.
McNeill, J., Barrie, F.R., Buck, W.R., Demoulin, V., Greuter, W., Hawksworth, D.L., Herendeen, P.S., Knapp, S., Marhold, K., Prado, J., Prud'homme van Reine, W.F., Smith, G.E., Wiersema, J.H., Turland, N.J., 2012. International Code of Nomenclature for Algae, Fungi, and Plants (Melbourne Code) Adopted by the Eighteenth International Botanical Congress Melbourne, Australia, July 2011. Koeltz Scientific Books.
Miadlikowska, J., Kauff, F., Hofstetter, V., Fraker, E., Grube, M., Hafellner, J., Reeb, V., Hodkinson, B.P., Kukwa, M., Lucking, R., Hestmark, G., Otalora, M.G., Rauhut, A., Budel, B., Scheidegger, C., Timdal, E., Stenroos, S., Brodo, I., Perlmutter, G.B., Ertz, D., Diederich, P., Lendemer, J.C., May, P., Schoch, C.L., Arnold, A.E., Gueidan, C., Tripp, E., Yahr, R., Robertson, C., Lutzoni, F., 2006. New insights into classification and evolution of the Lecanoromycetes (Pezizomycotina, Ascomycota) from phylogenetic analyses of three ribosomal RNA- and two protein-coding genes. Mycologia 98, 1088–1103.
Monchy, S., Sanciu, G., Jobard, M., Rasconi, S., Gerphagnon, M., Chabe, M., Cian, A., Meloni, D., Niquil, N., Christaki, U., Viscogliosi, E., Sime-Ngando, T., 2011. Exploring and quantifying fungal diversity in freshwater lake ecosystems using rDNA cloning/sequencing and SSU tag pyrosequencing. Environ. Microbiol. 13, 1433–1453.
Nilsson, R.H., Ryberg, M., Kristiansson, E., Abarenkov, K., Larsson, K.H., Käljalg, U., 2006. Taxonomic reliability of DNA sequences in public sequence databases: a fungal perspective. PLoS One 1, e59.
Nilsson, R.H., Kristiansson, E., Ryberg, M., Hallenberg, N., Larsson, K.H., 2008. Intraspecific ITS Variability in the kingdom Fungi as expressed in the international sequence databases and its implications for molecular species identification. Evol. Bioinforma. 4, 193–201.
Nilsson, R.H., Wurzbacher, C., Bahram, M., Coimbra, V.R.M., Larsson, E., Tedersoo, L., Eriksson, J., Ritter, C.D., Svantesson, S., Sanchez-Garcia, M., Ryberg, M., Kristiansson, E., Abarenkov, K., 2016. Top 50 most wanted fungi. Mycokeys 29–40.
O'Donnell, K., Cigelnik, E., 1997. Two divergent intragenomic rDNA ITS2 types within a monophyletic lineage of the fungus *Fusarium* are nonorthologous. Mol. Phylogenet. Evol. 7, 103–116.
O'Donnell, K., Lutzoni, F.M., Ward, T.J., Benny, G.L., 2001. Evolutionary relationships among mucoralean fungi (*Zygomycota*): evidence for family polyphyly on a large scale. Mycologia 93, 286–297.
Orsi, W., Biddle, J.F., Edgcomb, V., 2013. Deep sequencing of subseafloor eukaryotic rRNA reveals active fungi across marine subsurface provinces. PLoS One 8.
Orsi, W.D., Richards, T.A., Santoro, A.E., 2015. Cellular maintenance processes that potentially underpin the survival of subseafloor fungi over geological timescales. Estuar Coast Shelf 164, A1–A9.
Overy, D.P., Bayman, P., Kerr, R.G., Bills, G.F., 2014. An assessment of natural product discovery from marine (*sensu strictu*) and marine-derived fungi. Mycology 5, 145–167.
Panzer, K., 2013. A Complete View on the Aquatic-derived Fungal Phylogeny and Ecology. Master Master Thesis. University of Bremen, Bremen, Germany.
Panzer, K., Yilmaz, P., Weiß, M., Reich, L., Richter, M., Wiese, J., Schmaljohann, R., Labes, A., Imhoff, J.F., Glöckner, F.O., Reich, M., 2015a. Data from: Identification of habitat-specific biomes of aquatic fungal communities using a comprehensive nearly full-length 18S rRNA dataset enriched with contextual data. In: Dryad Digital Repository.
Panzer, K., Yilmaz, P., Weiß, M., Reich, L., Richter, M., Wiese, J., Schmaljohann, R., Labes, A., Imhoff, J.F., Glöckner, F.O., Reich, M., 2015b. Identification of habitat-specific biomes of aquatic fungal communities using a comprehensive nearly full-length 18S rRNA dataset enriched with contextual data. PLoS One 10, e0134377.
Pena, C., Malm, T., 2012. VoSeq: a voucher and DNA Sequence web application. PLoS One 7.
Pettit, R.K., 2011. Small-molecule elicitation of microbial secondary metabolites. Microb. Biotechnol. 4, 471–478.
Pryor, B.M., Gilbertson, R.L., 2000. Molecular phylogenetic relationships amongst *Alternaria* species and related fungi based upon analysis of nuclear ITS and mt SSU rDNA sequences. Mycol. Res. 104, 1312–1321.
Quandt, C.A., Kepler, R.M., Gams, W., Araujo, J.P.M., Ban, S., Evans, H.C., Hughes, D., Humber, R.A., Hywel-Jones, N., Li, Z.B., Luangsa-Ard, J.J., Rehner, S.A., Sanjuan, T., Sato, H., Shrestha, B., Sung, G.H., Yao, Y.J., Zare, R., Spatafora, J., 2014. Phylogenetic-based nomenclatural proposals for *Ophiocordycipitaceae* (*Hypocreales*) with new combinations in *Tolypocladium*. IMA Fungus 5, 121–134.
Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., Glockner, F.O., 2013. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. Nucleic Acids Res. 41, D590–D596.
Rai, M.K., Tiwari, V.V., Irinyi, L., Kovics, G.J., 2014. Advances in taxonomy of genus *Phoma*: Polyphyletic nature and role of phenotypic traits and molecular systematics. Indian J. Microbiol. 54, 123–128.
Rämä, T., Norden, J., Davey, M.L., Mathiassen, G.H., Spatafora, J.W., Kauserud, H., 2014. Fungi ahoy! Diversity on marine wooden substrata in the high North. Fungal Ecol. 8, 46–58.
Rämä, T., Davey, M., Norden, J., Halvorsen, R., Blaalid, R., Mathiassen, G., Alsos, I., Kauserud, H., 2016. Fungi sailing the Arctic Ocean: speciose communities in North Atlantic driftwood as revealed by high-throughput amplicon sequencing. Microb. Ecol. 72, 295–304.
Rayner, A.D.M., Coates, D., 1987. Regulation of mycelial organization and responses. In:

Rayner, A.D.M., Brasier, C.M., Moore, D. (Eds.), Evolutionary Biology of the Fungi. Cambridge Univ Press, Cambridge, UK.
Reblova, M., Miller, A., Rossman, A.Y., Seifert, K.A., Crous, P.W., Hawksworth, D.L., Abdel-Wahab, M.A., Cannon, P.F., Daranagama, D.A., De Beer, Z.W., Huang, S.K., Hyde, K.D., Jayawardena, R., Jaklitsch, W.M., Jones, E.B.G., YM, Ju, Judith, C., Maharachchikumbura, S.S., Pang, K.L., Petrini, L.E., Raja, H.A., Romero, A.I., Shearer, C., Senanayake, I.C., Voglmayr, H., Weir, B.S., Wijayawardene, N.N., 2016. Recommendations for competing sexual-asexually typified generic names in *Sordariomycetes* (except *Diaporthales, Hypocreales*, and *Magnaporthales*). IMA Fungus 7, 131–153.
Redou, V., Ciobanu, M.C., Pachiadaki, M.G., Edgcomb, V., Alain, K., Barbier, G., Burgaud, G., 2014. In-depth analyses of deep subsurface sediments using 454-pyrosequencing reveals a reservoir of buried fungal communities at record-breaking depths. FEMS Microbiol. Ecol. 90, 908–921.
Reen, F.J., Romano, S., Dobson, A.D.W., O'Gara, F., 2015. The sound of silence: activating silent biosynthetic gene clusters in marine microorganisms. Mar. Drugs 13, 4754–4783.
Reich, M., Kohler, A., Martin, F., Bu, E.M., 2009. Development and validation of an oligonucleotide microarray to characterise ectomycorrhizal fungal communities. BMC Microbiol. 9.
Reich, M., Wichels, A., Panzer, K., Krause, E., Gimenez, L., Gerdts, G., 2017. Impacts of a reduction in seawater pH mimicking ocean acidification on the structure and diversity of mycoplankton communities. Aquat. Microb. Ecol. 79, 221–233.
Rezaei-Matehkolaei, A., Mirhendi, H., Makimura, K., de Hoog, G.S., Satoh, K., Najafzadeh, M.J., Shidfar, M.R., 2014. Nucleotide sequence analysis of beta tubulin gene in a wide range of dermatophytes. Med. Mycol. 52, 674–688.
Richards, T.A., Jones, M.D.M., Leonard, G., Bass, D., 2012. Marine fungi: their ecology and molecular diversity. Annu. Rev. Mar. Sci. 4, 495–522.
Richards, T.A., Leonard, G., Mahé, F., del Campo, J., Romac, S., Jones, M.D.M., Maguire, F., Dunthorn, M., De Vargas, C., Massana, R., Chambouvet, A., 2015. Molecular diversity and distribution of marine fungi across 130 European environmental samples. Proc. R. Soc. Lond. [Biol] 282.
Robert, V., Vu, D., Amor, A.B.H., van de Wiele, N., Brouwer, C., Jabas, B., Szoke, S., Dridi, A., Triki, M., Daoud, S.B., Chouchen, O., Vaas, L., de Cock, A., Stalpers, J.A., Stalpers, D., Verkley, G.J.M., Groenewald, M., dos Santos, F.B., Stegehuis, G., Li, W., Wu, L., Zhang, R., Ma, J., Zhou, M., Gorjón, S.P., Eurwilaichitr, L., Ingsriswang, S., Hansen, K., Schoch, C., Robbertse, B., Irinyi, L., Meyer, W., Cardinali, G., Hawksworth, D.L., Taylor, J.W., Crous, P.W., 2016. MycoBank gearing up for new horizons. IMA Fungus 4, 371–379.
Roemer, T., DM, Xu, Singh, S.B., Parish, C.A., Harris, G., Wang, H., Davies, J.E., Bills, G.F., 2011. Confronting the challenges of natural product-based antifungal discovery. Chem. Biol. 18, 148–164.
Rosling, A., Cox, F., Cruz-Martinez, K., Ihrmark, K., Grelet, G.A., Lindahl, B.D., Menkis, A., James, T.Y., 2011. Archaeorhizomycetes: unearthing an ancient class of ubiquitous soil fungi. Science 333, 876–879.
Sakayaroj, J., Preedanon, S., Supaphon, O., Jones, E.B.G., Phongpaichit, S., 2010. Phylogenetic diversity of endophyte assemblages associated with the tropical seagrass *Enhalus acoroides* in Thailand. Fungal Divers. 42, 27–45.
Saleem, M., Ali, M.S., Hussain, S., Jabbar, A., Ashraf, M., Lee, Y.S., 2007. Marine natural products of fungal origin. Nat. Prod. Rep. 24, 1142–1152.
Samson, R.A., Seifert, K.A., Kuijpers, A.F.A., Houbraken, J.A.M.P., Frisvad, J.C., 2004. Phylogenetic analysis of *Penicillium* subgenus *Penicillium* using partial beta-tubulin sequences. Stud. Mycol. 175–200.
Sanchez, J.F., Somoza, A.D., Keller, N.P., Wang, C.C.C., 2012. Advances in *Aspergillus* secondary metabolite research in the post-genomic era. Nat. Prod. Rep. 29, 351–371.
Schadt, E.E., Turner, S., Kasarskis, A., 2010. A window into third-generation sequencing. Hum. Mol. Genet. 19, R227–R240.
Scharf, D.H., Brakhage, A.A., 2013. Engineering fungal secondary metabolism: a roadmap to novel compounds. J. Biotechnol. 163, 179–183.
Schmitt, I., Crespo, A., Divakar, P.K., Fankhauser, J.D., Herman-Sackett, E., Kalb, K., Nelsen, M.P., Nelson, N.A., Rivas-Plata, E., Shimp, A.D., Widhelm, T., Lumbsch, H.T., 2009. New primers for promising single-copy genes in fungal phylogenetics and systematics. Persoonia 23, 35–40.
Schoch, C.L., Seifert, K.A., Huhndorf, S., Robert, V., Spouge, J.L., Levesque, C.A., Chen, W., Bolchacova, E., Voigt, K., Crous, P.W., Miller, A.N., Wingfield, M.J., Aime, M.C., An, K.D., Bai, F.Y., Barreto, R.W., Begerow, D., Bergeron, M.J., Blackwell, M., Boekhout, T., Bogale, M., Boonyuen, N., Burgaz, A.R., Buyck, B., Cai, L., Cai, Q., Cardinali, G., Chaverri, P., Coppins, B.J., Crespo, A., Cubas, P., Cummings, C., Damm, U., de Beer, Z.W., de Hoog, G.S., Del-Prado, R.B.D., Dieguez-Uribeondo, J., Divakar, P.K., Douglas, B., Duenas, M., Duong, T.A., Eberhardt, U., Edwards, J.E., Elshahed, M.S., Fliegerova, K., Furtado, M., Garcia, M.A., Ge, Z.W., Griffith, G.W., Griffiths, K., Groenewald, J.Z., Groenewald, M., Grube, M., Gryzenhout, M., Guo, L.D., Hagen, F., Hambleton, S., Hamelin, R.C., Hansen, K., Harrold, P., Heller, G., Herrera, G., Hirayama, K., Hirooka, Y., Ho, H.M., Hoffmann, K., Hofstetter, V., Hognabba, F., Hollingsworth, P.M., Hong, S.B., Hosaka, K., Houbraken, J., Hughes, K., Huhtinen, S., Hyde, K.D., James, T., Johnson, E.M., Johnson, J.E., Johnston, P.R., Jones, E.B., Kelly, L.J., Kirk, P.M., Knapp, D.G., Koljalg, U., GM, K., Kurtzman, C.P., Landvik, S., Leavitt, S.D., Liggenstoffer, A.S., Liimatainen, K., Lombard, L., Luangsa-Ard, J.J., Lumbsch, H.T., Maganti, H., Maharachchikumbura, S.S., Martin, M.P., May, T.W., AR, McTaggart, Methven, A.S., Meyer, W., Moncalvo, J.M., Mongkolsamrit, S., Nagy, L.G., Nilsson, R.H., Niskanen, T., Nyilasi, I., Okada, G., Okane, I., Olariaga, I., Otte, J., Papp, T., Park, D., Petkovits, T., Pino-Bodas, R., Quaedvlieg, W., Raja, H.A., Redecker, D., T, R., Ruibal, C., Sarmiento-Ramirez, J.M., Schmitt, I., Schussler, A., Shearer, C., Sotome, K., Stefani, F.O., Stenroos, S., Stielow, B., Stockinger, H., Suetrong, S., Suh, S.O., Sung, G.H., Suzuki, M., Tanaka, K., Tedersoo, L., Telleria, M.T., Tretter, E., Untereiner, W.A., Urbina, H., Vagvolgyi, C., Vialle, A., Vu, T.D., Walther, G., Wang, Q.M., Wang, Y., Weir, B.S., Weiss, M., White, M.M., Xu, J., Yahr, R., Yang, Z.L., Yurkov, A., Zamora, J.C., Zhang, N., Zhuang, W.Y., Schindel, D., Consortium, F.B., 2012. Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. Proc. Natl. Acad. Sci. U. S. A. 109, 6241–6246.
Schubert, K., Groenewald, J.Z., Braun, U., Dijksterhuis, J., Starink, M., Hill, C.F., Zalar, P., De Hoog, G.S., Crous, P.W., 2007. Biodiversity in the *Cladosporium* herbarum complex (*Davidiellaceae, Capnodiales*), with standardisation of methods for *Cladosporium* taxonomy and diagnostics. Stud. Mycol. 58, 105–156.
Schueffler, A., Anke, T., 2014. Fungal natural products in research and development. Nat. Prod. Rep. 31, 1425–1448.
Schultz, J., Muller, T., Achtziger, M., Seibel, P.N., Dandekar, T., Wolf, M., 2006. The internal transcribed spacer 2 database - a web server for (not only) low level phylogenetic analyses. Nucleic Acids Res. 34, W704–W707.
Scott, J., Untereiner, W.A., Wong, B., Straus, N.A., Malloch, D., 2004. Genotypic variation in *Penicillium chrysogenum* from indoor environments. Mycologia 96, 1095–1105.
Seifert, K.A., 2009. Progress towards DNA barcoding of fungi. Mol. Ecol. Resour. 9, 83–89.
Shearer, C.A., Raja, H.A., 2010. Freshwater Ascomycetes Database. http://fungi.life.illinois.edu/.
Silber, J., Kramer, A., Labes, A., Tasdemir, D., 2016. From discovery to production: biotechnology of marine fungi for the production of new antibiotics. Mar. Drugs 14.
Simon, U.K., Weiss, M., 2008. Intragenomic variation of fungal ribosomal genes is higher than previously thought. Mol. Biol. Evol. 25, 2251–2254.
Skouboe, P., Frisvad, J.C., Taylor, J.W., Lauritsen, D., Boysen, M., Rossen, L., 1999. Phylogenetic analysis of nucleotide sequences from the ITS region of terverticillate *Penicillium* species. Mycol. Res. 103, 873–881.
Slepecky, R.A., Starmer, W.T., 2009. Phenotypic plasticity in fungi: a review with observations on *Aureobasidium pullulans*. Mycologia 101, 823–832.
Smedsgaard, J., Frisvad, J.C., 1996. Using direct electrospray mass spectrometry in taxonomy and secondary metabolite profiling of crude fungal extracts. J. Microbiol. Methods 25, 5–17.
Smith, D., 2003. Culture collections over the world. Int. Microbiol. 6, 95–100.
Smith, M.E., Douhan, G.W., Rizzo, D.M., 2007. Intra-specific and intra-sporocarp ITS variation of ectomycorrhizal fungi as assessed by rDNA sequencing of sporocarps and pooled ectomycorrhizal roots from a *Quercus* woodland. Mycorrhiza 18, 15–22.
Spatafora, J.W., Chang, Y., Benny, G.L., Lazarus, K., Smith, M.E., Berbee, M.L., Bonito, G., Corradi, N., Grigoriev, I., Gryganskyi, A., James, T.Y., O'Donnell, K., Roberson, R.W., Taylor, T.N., Uehling, J., Vilgalys, R., White, M.M., Stajich, J.E., 2016. A phylum-level phylogenetic classification of zygomycete fungi based on genome-scale data. Mycologia 108, 1028–1046.
Stockinger, H., Peyret-Guzzon, M., Koegel, S., Bouffaud, M.L., Redecker, D., 2014. The largest subunit of RNA polymerase II as a new marker gene to study assemblages of arbuscular mycorrhizal fungi in the field. PLoS One 9.
Stoeck, T., Bass, D., Nebel, M., Christen, R., Jones, M.D.M., Breiner, H.W., Richards, T.A., 2010. Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. Mol. Ecol. 19, 21–31.
Stukenbrock, E.H., 2013. Evolution, selection and isolation: a genomic view of speciation in fungal plant pathogens. New Phytol. 199, 895–907.
Sung, G.H., Hywel-Jones, N.L., Sung, J.M., Luangsa-Ard, J.J., Shrestha, B., Spatafora, J.W., 2007. Phylogenetic classification of *Cordyceps* and the clavicipitaceous fungi. Stud. Mycol. 5–59.
Tanabe, A.S., Nagai, S., Hida, K., Yasuike, M., Fujiwara, A., Nakamura, Y., Takano, Y., Katakura, S., 2016. Comparative study of the validity of three regions of the 18S-rRNA gene for massively parallel sequencing-based monitoring of the planktonic eukaryote community. Mol. Ecol. Resour. 16, 402–414.
Tarkka, M.T., Vasara, R., Gorfer, M., Raudaskoski, M., 2000. Molecular characterization of actin genes from homobasidiomycetes: two different actin genes from *Schizophyllum commune* and *Suillus bovinus*. Gene 251, 27–35.
Taylor, J.D., Cunliffe, M., 2016. Multi-year assessment of coastal planktonic fungi reveals environmental drivers of diversity and abundance. ISME J. 10, 2118–2128.
Tedersoo, L., Suva, T., Larsson, E., Käljalg, U., 2006. Diversity and community structure of ectomycorrhizal fungi in a wooded meadow. Mycol. Res. 110, 734–748.
Townsend, J.P., 2007. Profiling phylogenetic informativeness. Syst. Biol. 56, 222–231.
Tretter, E.D., Johnson, E.M., Wang, Y., Kandel, P., White, M.M., 2013. Examining new phylogenetic markers to uncover the evolutionary history of early-diverging fungi: comparing *MCM7*, *TSR1* and rRNA genes for single- and multi-gene analyses of the *Kickxellomycotina*. Persoonia 30, 106–125.
Vetrovsky, T., Kolarik, M., Zifcakova, L., Zelenka, T., Baldrian, P., 2016. The *rpb2* gene represents a viable alternative molecular marker for the analysis of environmental fungal communities. Mol. Ecol. Resour. 16, 388–401.
Voigt, K., Wostemeyer, J., 2001. Phylogeny and origin of 82 zygomycetes from all 54 genera of the *Mucorales* and *Mortierellales* based on combined analysis of actin and translation elongation factor EF-1 alpha genes. Gene 270, 113–120.
Vralstad, T., Fossheim, T., Schumacher, T., 2000. *Piceirhiza bicolorata* - the ectomycorrhizal expression of the *Hymenoscyphus ericae* aggregate? New Phytol. 145, 549–563.
Wang, H., Xu, Z., Gao, L., Hao, B.L., 2009. A fungal phylogeny based on 82 complete genomes using the composition vector method. BMC Evol. Biol. 9.
Wang, M.X., Carver, J.J., Phelan, V.V., Sanchez, L.M., Garg, N., Peng, Y., Nguyen, D.D., Watrous, J., Kapono, C.A., Luzzatto-Knaan, T., Porto, C., Bouslimani, A., Melnik, A.V., Meehan, M.J., Liu, W.T., Criisemann, M., Boudreau, P.D., Esquenazi, E., Sandoval-Calderon, M., Kersten, R.D., Pace, L.A., Quinn, R.A., Duncan, K.R., Hsu, C.C., Floros, D.J., Gavilan, R.G., Kleigrewe, K., Northen, T., Dutton, R.J., Parrot, D., Carlson, E.E., Aigle, B., Michelsen, C.F., Jelsbak, L., Sohlenkamp, C., Pevzner, P., Edlund, A., McLean, J., Piel, J., Murphy, B.T., Gerwick, L., Liaw, C.C., Yang, Y.L., Humpf, H.U., Maansson, M., Keyzers, R.A., Sims, A.C., Johnson, A.R., Sidebottom,

A.M., Sedio, B.E., Klitgaard, A., Larson, C.B., Boya, C.A., Torres-Mendoza, D., Gonzalez, D.J., Silva, D.B., Marques, L.M., Demarque, D.P., Pociute, E., O'Neill, E.C., Briand, E., Helfrich, E.J.N., Granatosky, E.A., Glukhov, E., Ryffel, F., Houson, H., Mohimani, H., Kharbush, J.J., Zeng, Y., Vorholt, J.A., Kurita, K.L., Charusanti, P., McPhail, K.L., Nielsen, K.F., Vuong, L., Elfeki, M., Traxler, M.F., Engene, N., Koyama, N., Vining, O.B., Baric, R., Silva, R.R., Mascuch, S.J., Tomasi, S., Jenkins, S., Macherla, V., Hoffman, T., Agarwal, V., Williams, P.G., Dai, J.Q., Neupane, R., Gurr, J., Rodriguez, A.M.C., Lamsa, A., Zhang, C., Dorrestein, K., Duggan, B.M., Almaliti, J., Allard, P.M., Phapale, P., Nothias, L.F., Alexandrovr, T., Litaudon, M., Wolfender, J.L., Kyle, J.E., Metz, T.O., Peryea, T., Nguyen, D.T., VanLeer, D., Shinn, P., Jadhav, A., Muller, R., Waters, K.M., Shi, W.Y., Liu, X.T., Zhang, L.X., Knight, R., Jensen, P.R., Palsson, B.O., Pogliano, K., Linington, R.G., Gutierrez, M., Lopes, N.P., Gerwick, W.H., Moore, B.S., Dorrestein, P.C., Bandeira, N., 2016a. Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. Nat. Biotechnol. 34, 828–837.

Wang, J.W., Luo, Z.H., Xu, W., Ding, J.F., Zheng, T.L., 2016b. Transformation of dimethyl phthalate esters (DMPEs) by a marine red yeast *Rhodotorula mucilaginosa* isolated from deep sea sediments of the Atlantic Ocean. Int. Biodeterior. Biodegrad. 109, 223–228.

Webb, C.O., Ackerly, D.D., Kembel, S.W., 2008. Phylocom: software for the analysis of phylogenetic community structure and trait evolution. Bioinformatics 24, 2098–2100.

Wijayawardene, N.N., Crous, P.W., Kirk, P.M., Hawksworth, D.L., Boonmee, S., Braun, U., Dai, D.Q., D'souza, M.J., Diederich, P., Dissanayake, A., Doilom, M., Hongsanan, S., Jones, E.B.G., Groenewald, J.Z., Jayawardena, R., Lawrey, J.D., Liu, J.K., Lucking, R., Madrid, H., Manamgoda, D.S., Muggia, L., Nelsen, M.P., Phookamsak, R., Suetrong, S., Tanaka, K., Thambugala, K.M., Wanasinghe, D.N., Wikee, S., Zhang, Y., Aptroot, A., Ariyawansa, H.A., Bahkali, A.H., Bhat, D.J., Gueidan, C., Chomnunti, P., De Hoog, G.S., Knudsen, K., Li, W.J., McKenzie, E.H.C., Miller, A.N., Phillips, A.J.L., Piatek, M., Raja, H.A., Shivas, R.S., Slippers, B., Taylor, J.E., Tian, Q., Wang, Y., Woudenberg, J.H.C., Cai, L., Jaklitsch, W.M., Hyde, K.D., 2014. Naming and outline of *Dothideomycetes*-2014 including proposals for the protection or suppression of generic names. Fungal Divers. 69, 1–55.

Wohlleben, W., Mast, Y., Stegmann, E., Ziemert, N., 2016. Antibiotic drug discovery. Microb. Biotechnol. 9, 541–548.

Wuyts, J., De Rijk, P., Van de Peer, Y., Pison, G., Rousseeuw, P., De Wachter, R., 2000. Comparative analysis of more than 3000 sequences reveals the existence of two pseudoknots in area V4 of eukaryotic small subunit ribosomal RNA. Nucleic Acids Res. 28, 4698–4708.

Wuyts, J., Van de Peer, Y., Winkelmans, T., De Wachter, R., 2002. The European database on small subunit ribosomal RNA. Nucleic Acids Res. 30, 183–185.

Xu, B., Yang, Z.H., 2016. Challenges in species tree estimation under the multispecies coalescent model. Genetics 204, 1353–1368.

Yanai, K., Sumida, N., Okakura, K., Moriya, T., Watanabe, M., Murakami, T., 2004. *Para*-position derivatives of fungal anthelmintic cyclodepsipeptides engineered with *Streptomyces venezuelae* antibiotic biosynthetic genes. Nat. Biotechnol. 22, 848–855.

Yang, X.B., Gao, X.D., Han, F., Tan, R.X., 2005. Sulfation of a polysaccharide produced by a marine filamentous fungus *Phoma herbarum* YS4108 alters its antioxidant properties *in vitro*. Bba-Gen Subj. 1725, 120–127.

Yarza, P., Yilmaz, P., Pruesse, E., Glockner, F.O., Ludwig, W., Schleifer, K.H., Whitman, W.B., Euzeby, J., Amann, R., Rossello-Mora, R., 2014. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. Nat. Rev. Microbiol. 12, 635–645.

Yarza, P., Yilmaz, P., Glöckner, F.O., Reich, M., 2017. A phylogenetic framework for the kingdom Fungi based on 18S rRNA gene sequences. Mar. Genomics (in this issue).

Zähner, H., 1979. What are secondary metabolites? Folia Microbiol. 24, 435–443.

Zhang, C., Zhang, D.X., Zhu, T.Q., Yang, Z.H., 2011. Evaluation of a bayesian coalescent method of species delimitation. Syst. Biol. 60, 747–761.

Zhao, C.H., Zhang, T., Li, M., Chi, Z.M., 2010. Single cell oil production from hydrolysates of inulin and extract of tubers of *Jerusalem artichoke* by *Rhodotorula mucilaginosa* TJY15a. Process Biochem. 45, 1121–1126.

Zuccaro, A., Schulz, B., Mitchell, J.I., 2003. Molecular detection of ascomycetes associated with *Fucus serratus*. Mycol. Res. 107, 1451–1466.