Universidade de Aveiro
2023

Ana Rita
Balsa Gomes

Metodologias para produção, extração e caracterização de compostos antibacterianos

Methodologies for production, extraction, and characterization of antibacterial compounds

Universidade de Aveiro Departamento de Química
2023

Ana Rita Balsa Gomes

Metodologias para produção, extração e caracterização de compostos antibacterianos

Methodologies for production, extraction, and characterization of antibacterial compounds

Universidade de Aveiro Departamento de Química
2023

Ana Rita Balsa Gomes

Metodologias para produção, extração e caracterização de compostos antibacterianos

Methodologies for production, extraction, and characterization of antibacterial compounds

Tese apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Doutor em Química, realizada sob a orientação científica do Doutora Teresa Rocha Santos, Investigadora Principal com Agregação do Centro de Estudos do Ambiente e do Mar (CESAM) e do Departamento de Química da Universidade de Aveiro e do Doutor João Pinto da Costa, Investigador Auxiliar do Centro de Estudos do Ambiente e do Mar (CESAM) e do Departamento de Química da Universidade de Aveiro.

Apoio financeiro ao CESAM pela FCT/MCTES (UIDP/50017/2020+UIDB/50017/2020+ LA/P/0094/2020), através de fundos nacionais.

Apoio financeiro da Fundação para a Ciência e a Tecnologia (FCT), no âmbito do projeto BioPlasMar (PTDC/CTA-AMB/0934/2021).

Dedico este trabalho a todos aqueles que acreditaram em mim, mesmo quando eu deixei de acreditar.

o júri

presidente

Doutor Nuno Miguel Gonçalves Borges de Carvalho
Professor Catedrático, Universidade de Aveiro

Doutor Armando Costa Duarte
Professor Catedrático Aposentado, Universidade de Aveiro

Doutor José Carlos Márcia Andrade
Professor Associado, Cooperativa de Ensino Superior Politécnico e Universitário

Doutora Maria Elizabeth Tiritan
Professora Auxiliar, Universidade do Porto

Doutor João Miguel Martins Novaes Pinto da Costa (Coorientador)
Investigador Auxiliar, Universidade de Aveiro

Doutora Cristina Marisa Ribeiro de Almeida
Investigadora Auxiliar, Universidade do Porto

agradecimentos

A concretização do presente trabalho só foi possível graças à importante colaboração e contribuição de diversas personalidades às quais não poderia deixar de agradecer.

Aos meus orientadores, Doutora Teresa Santos e Doutor João Pinto da Costa, pelo apoio, orientação, conselhos, sugestões e paciência disponibilizados na realização deste trabalho. Ao Professor Doutor Armando Duarte, pela confiança, disponibilidade e colaboração prestada nas mais diversas fases do trabalho.

Aos meus colegas, Ana Paço e Micael Gonçalves pela disponibilidade e apoio no trabalho laboratorial.

Aos meus familiares e amigos por todo o suporte dado nos bons e menos bons momentos ao longo de todos estes anos.

A todas as outras pessoas que de uma forma direta ou indireta contribuíram para a realização e finalização deste trabalho.

A todos, o meu muito obrigada!

palavras-chave

Fungos, compostos bioactivos, técnicas de extracção, caracterização química, bioactividade.

resumo

Este trabalho de doutoramento é baseado em duas partes, uma usando mineração de dados e texto e outra em laboratório, fornecendo estratégias baseadas em metodologias analíticas avançadas para a descoberta de compostos bioativos (CB), com potenciais aplicações em ferramentas farmacêuticas e biotecnológicas de organismos fúngicos.

O objetivo principal deste projeto foi usar uma técnica de mineração de texto para identificar os tópicos de um texto científico relevante relacionado com compostos bioativos de fungos e a conexão evolutiva entre esses tópicos, bem como usar ferramentas de visualização para apresentar os tópicos e a associação entre eles, como uma maneira conveniente de ajudar os utilizadores a determinar tópicos relevantes.

Na segunda parte desta tese, inicialmente, as espécies de fungos foram cultivadas em laboratório usando um reator descontínuo. Em seguida, esses organismos fúngicos foram extraídos usando técnicas de extração verdes, como a Extração de Alta Pressão (HPE).

Em segundo lugar, os compostos extraídos foram caracterizados por análise elementar (CNHS), ressonância magnética nuclear (RMN) e espectroscopia de infravermelho por transformada de Fourier (FTIR).

Em terceiro lugar, a bioatividade foi avaliada através de ensaios antimicrobianos baseados na farmacopeia europeia.

Finalmente, as conclusões do projeto fornecerão dados relevantes às partes interessadas e envolvidas no plano de recuperação económica através da gestão de recursos sustentáveis.

keywords

Fungi, bioactive compounds, extraction techniques, chemical characterization, bioactivity.

abstract

This doctoral project is based on two parts, one using text and data mining and another in laboratory providing strategies based on advanced analytical methodologies for the discovery of bioactive compounds (BC), with potential applications on pharmaceutical and biotechnological tools from fungi organisms.

The primary goal of this project was using a text mining technique for identifying the topics of a relevant scientific text related to fungi bioactive compounds and evolutionary connection among these topics, as well as using visualization tools for presenting both the topics and the association among them, as a convenient way to help users to determine relevant topics.

In the second part of this thesis, firstly, fungi species were cultivated in lab using a batch reactor. Then, these fungi organisms were extracted by using environmentally friendly extraction techniques, such as High-Pressure Extraction (HPE).

Secondly, the extracted compounds were characterized using elemental analysis (CNHS), nuclear magnetic resonance (NMR) and Fourier transform infrared spectroscopy (FTIR).

Thirdly, the bioactivity was evaluated through antimicrobial assays based on European pharmacopeia.

Finally, the project findings will provide relevant data to stakeholders involved in the economic recovery plan through the sustainable resource management.

TABLE OF CONTENTS

Chapter 2

Chapter 3

Chapter 4

Chapter 2

Chapter 3

Chapter 4

5-MTHF	5-methyltetrahydrofolate
ABTS	2,2'-azinobis (3-ethylbenzothiazoline-6-sulfonic acid
AI	Artificial Intelligence
AML	Amoxicillin
ANOVA	One-way analysis of variance
API	Application program interface
ASE	Accelerated solvent extraction
BC	Bioactive compounds
BHA	Butylated hydroxyanisole
BHT	Butylated hydroxytoluene
BLAST	Basic local alignment search tool
CH	Methine
CH_2	Methylene
CH_3	Methyl
CNHS	Carbon, nitrogen, hydrogen, sulfur elemental analysis
CO_2	Carbon dioxide
CPE	Centrifugal partition extraction
CPU	Central process unit
CRM	Customer relationship management
DEPT	Distortionless enhancement by polarization transfer
DES	Deep eutectic solvents
DNA	Deoxyribonucleic acid
DOI	Digital object identified
DPPH	2,2-diphenyl-1-picrylhydrazyl
EACC	Ehrlich's ascites carcinoma cells
EAE	Enzyme-assisted extraction
EID	Enforcement integrated database
ETL	Extract, transform and load
EV71	Enterovirus 71
FDA	Food and drug administration
FRAP	Ferric reducing antioxidant power
FTIR	Fourier transform infrared
FTIR-ATR	Fourier transform infrared - Attenuated total reflectance
GC-MS	Gas chromatography-mass spectroscopy
GPx	Glutathione peroxidase
H1N1	Influenza A virus
HIV-1	Human immunodeficiency virus 1
HMGCR	3-hydroxy-3-methylglutaryl

HPE	High-pressure extraction
HTML	Hyper-text markup language
IAV	Influenza A virus
IC_{50}	50% inhibitory concentration
IDF	Inverse document frequency
IDE	Integrated development environment
IF	Impact factor
IL	Ionic liquids
IL-1β	Interleukin 1 beta
IR	Information retrieval
ITS	Internal transcribed spacer
LC-MS	Liquid chromatography-Mass Spectrometry
LDL	Low-density lipoprotein
LPS	Lipopolysaccharides
MAE	Microwave-assisted extraction
ML	Machine learning techniques
MO	Moringa oleifera
MRSA	Methicillin-resistant *Staphylococcus aureus*
MS	Mass spectrometry
MSSA	Methicillin-sensitive *Staphylococcus aureus*
NDES	Natural deep eutectic solvents
NIH	National institutes of health
NLP	Natural language processing
NMR	Nuclear magnetic resonance
NO	Nitric oxide
NOE	Nuclear overhauser effect
OPC	Ortho-phenolic content
ORM	Object relational mapping
PBP	Phycobiliprotein
PCR	Polymerase chain reaction
PDA	Potato dextrose agar
PDF	Portable document format
PEF	Pulsed electric field extraction
PG	Propyl gallate
PGE2	Prostaglandin E2
PHS	Public health service
PII	Personally identifiable information
PLE	Pressurized liquid extraction

RNA	Ribonucleic acid
ROS	Reactive oxygen species
RSA	Radical scavenging activity
RSC	Radical scavenging capacity
$ScCO_2$	Supercritical carbon dioxide
SD	Standard deviation
SJR	Scimago journal rank
SLE	Solid-liquid extraction
SFE	Supercritical fluid extraction
SOD	Superoxide dismutase
SWE	Subcritical water extraction
TFC	Total flavonoids content
TGC	Tigecycline
TNF	Tumor necrosis factor
TPA	*12-O-tetradecanoyl-phorbol-13-acetate*
TPC	Total phenolic content
TRC	Total reducing capacity
UAE	Ultrasound-assisted extraction
US	United States
USA	United States of America
UV-Vis	Ultraviolet-Visible
XRF	X-ray fluorescence

Chapter 1

OBJECTIVES

This section outlines a general description of the main objectives of the present doctoral project, providing an overview of the key topics developed throughout the research, as well as the structure of the dissertation. Thus, this doctoral project had the following scientific objectives:

I. Data mining for the assessment of the current landscape of fungal compounds research.

II. Extraction of compounds with antibacterial and antioxidant activities from fungi:

a) Growth and isolation of fungal samples, in laboratorial settings, with potential applications as pharmaceuticals and biotechnological tools;

b) Application of environmentally friendly methodology for the extraction of target bioactive compounds, such as High-Pressure Extraction (HPE);

c) Assessment of bioactivity through antibacterial and antioxidant assays;

d) Characterization of bioactive extracts via elemental analysis (CNHS), Fourier transform infrared (FTIR) and nuclear magnetic resonance (NMR) spectroscopies.

Within the framework of this doctoral thesis, compounds extracted from fungal biomass were obtained, with potential applications as pharmaceuticals and biotechnological tools. In order to accomplish the declared objectives, special attention was given to the species of fungi possible to be easily grown and kept in laboratory, in order to ensure a sustainable, environmentally friendly, industrial production of these bioactive compounds.

Therefore, this thesis is organized in five chapters, whereas chapter 1 provides the general and specific objectives of this thesis, as well as its organization in several chapters with reference to their associated contents. Chapter 2 gives an up-to-date introduction on the theoretical relevance of the various topics explored within this thesis summarizing the main scientific studies available on each topic. The chapter 3, describes the process of sorting through large data sets to identify patterns, predict trends and relationships that can help data analysis interpretation. Chapter 4 presents the high-pressure extraction method on biological properties of a selection of six understudied mushrooms. The selection of species, their elemental composition, as well as FTIR and NMR characterization is deeply presented and discussed. Finally, the chapter 5 outlines the concluding remarks concerning the various phases of this PhD work, highlighting the main achievements and discussing potential future directions.

Chapter 2

INTRODUCTION

2.1 General Overview

Nature has been a source of medicinal products for millennial and a powerful pharmacy harboring a vast source of biodiversity, as well as providing a wide abundance of still undisclosed compounds with enormous potential bioactivity (Carroll et al., 2021; Newman & Cragg, 2020). Since ancient times, natural products, mainly plants, have been used to treat many diseases. These products, well known for their unique chemical diversity and bioactivity, continue to offer templates for the development of novel scaffolds of drugs (Newman & Cragg, 2020; Sarker & Nahar, 2012).

Natural products have served as a major source of drugs for centuries, and about half of the pharmaceuticals in use today are derived from these products. The innovation of novel drugs is constantly encouraged; thus, academic and industry researchers are striving to discover new potentially bioactive molecules from new sources (Gomes et al., 2016a; Newman & Cragg, 2020; Pye et al., 2017).

Interest in natural products research continues to increase and can be attributed to several factors, including unmet therapeutic needs, often due to the growing numbers of drug-resistant diseases, the remarkable diversity of both chemical structures and biological activities of naturally occurring secondary metabolites, the utility of bioactive natural products as biochemical and molecular probes, the development of novel and sensitive techniques to detect biologically active natural products, improved techniques to isolate, purify, and structurally characterize these active constituents, and advances in solving the demand for supply of complex natural products (Calixto, 2019; Pye et al., 2017).

Currently, some natural products can also be prepared through chemical synthesis, providing challenging synthetic targets to research and high-value commercial with application on industry, cosmetics, food, and dietary supplements (Krause & Tobin, 2013; Wells et al., 2018). In some cases, this also paves the way to a more streamlined and industrial-level production of key products of interest, allowing obtaining products of higher purity at a comparatively low cost to their naturally-isolated counterparts.

For many years, investigation has mainly focused on plants and terrestrial organisms, due not only relative greater availability, but also the popular traditions of these organisms. However, research focusing on microorganisms has increased significantly over the last decades (Figure 2.1), leading to a discovery of unique biological molecules from these organisms (Gomes et al., 2016a; Patridge et al., 2016; Singh et al., 2019).

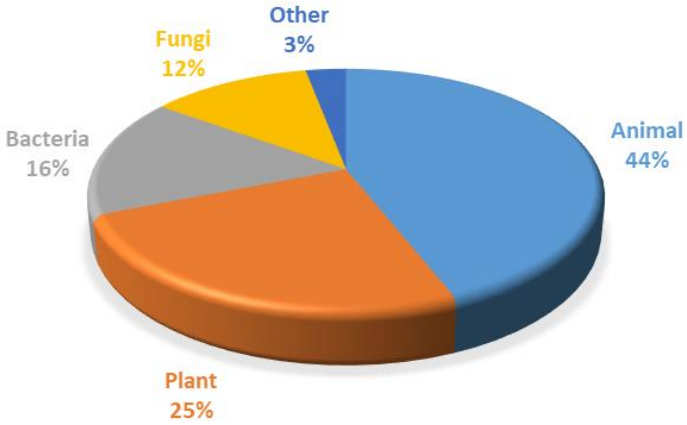

Figure 2.1. Main sources of new natural products. Animalia (44%), Plants (25%), Bacteria (16%), Fungi (12%), other (3%). Adapted from Patridge et al., 2016.

The natural compounds are known to present numerous advantageous, such as high chemical diversity, biochemical specificity, binding efficiency, and propensity to interact with biological targets (Martins et al., 2014). Due to their broad panel of bioactivities, such as anti-tumor, anti-microtubule, anti-proliferative, antibiotic, and anti-infective, natural products are exceptionally interesting high-value compounds for applications in the pharmaceutical industry (Calixto, 2019; Carroll et al., 2021; Pye et al., 2017). Following the same trend, other companies are investing in this field, such as cosmetics industry (Wells et al., 2018).

The natural products are currently an exclusive source of structurally diverse bioactive metabolites and have produced some of the most important bioactive compounds (BC) known today (Carrol et al., 2020; Gomes et al., 2016a; Gomes et al., 2017; Mayer et al., 2021). About one third of the best-selling drugs in the world are natural products or their derivatives. Of the 520 new drugs approved by Food and Drug Administration (FDA) between 1983 and 1994, 39% were natural products or derived from natural products and about 60 - 80% of antibiotics and anti-cancer drugs are derived from natural products (Calixto, 2019). On 2019, Newman and Cragg assessed the role of natural products in the drugs approved by the FDA between 1981 and 2019. They found that in this period the FDA approved 1881 drugs, 71 (3.8%) were unaltered natural products, 14 (0.8%) were botanical drugs (mixture), 356 (18.9%) were natural product derivatives and 463 (24.6%) were synthetic drugs, but with natural products pharmacophore (Figure 2.2) (Newman & Cragg, 2020).

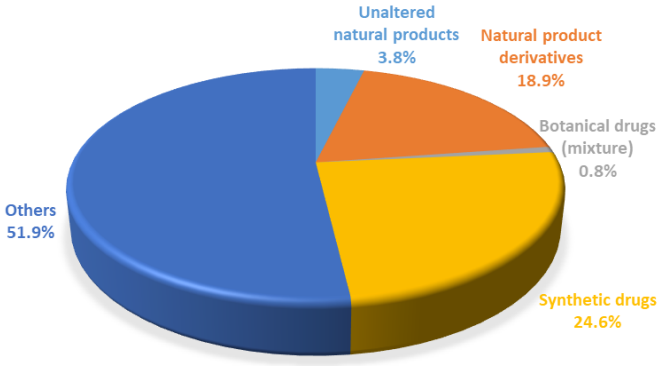

Figure 2.2. All new FDA-approved drugs between 1981−2019. Adapted from Newman & Cragg, 2020.

2.2 Bioactive Compounds

Bioactive compounds refer to both essential and nonessential compounds that are found in nature or even created during the processing of medicinal plants or foods, being capable of modulating different biological activities and ultimately benefiting health. Vitamins, more precisely vitamins A, B, C, D, and E, are one of the most studied essential compounds. As a matter of fact, vitamin C has been widely studied under distinct conditions, demonstrating a therapeutic effect in several diseases such as allergies, conversion of cholesterol to bile acids, synthesis of neurotransmitters, wound healing, bone formation, and maintenance of connective tissue (Biesalski et al., 2009; Caritá et al., 2020; Chambial et al., 2013).

On the other hand, fatty acids have also been widely studied as one of essential compounds, mainly due to their presence in several pathophysiological and physiological processes, with great benefits in addition to their nutritional advantages. Among these compounds, the most beneficial ones are the monounsaturated and the polyunsaturated, the latter of which includes the widely publicized omega 3, which has several health benefits, namely due to its bioactivities, namely anti-inflammatory, neuroprotective, protective in metabolic and cardiovascular diseases (Denis et al., 2015; Farley et al., 2021; Sokoła-Wysoczańska et al., 2018).

Regarding nonessential compounds, phytochemicals are major bioactive compounds, as well as extremely diverse, based on the classification of their chemical structures, more precisely: polyphenols, terpenoids, phytosterols, organosulfur compounds, and alkaloids. Overall, phytochemicals are found in vegetables, fruits, cereal grains, algae, sponges, cnidarians, and plants, being synthesized by the secondary metabolism of plant cells. The most common fruits that have these phytochemicals are blueberries, wild berries, blackberries, strawberries, and pomegranates; the most common vegetables are red peppers, spinach, asparagus, and broccoli; the most common cereals are black rice, red rice, and purple rice; and the most common medicinal plants that have

phytochemicals are *Dioscorea bulbifera*, *Camellia sinensis*, and *Ephedra sinica* (Liu, 2013; Somani et al., 2015; Zhang et al., 2015).

Lastly, the third class of bioactive compounds refers to those that result from food processing, consisting in a particularly important class, especially from the perspective of food science, given the fact that their use in food products has a dual benefit: on one hand, they protect the food, and on the other hand, they transform the product into a functional food, the latter consisting in a type of food that offers a health benefit that goes beyond the nutritional quality. For instance, biopeptides refer to peptide fractions of 3-20 amino acids, possessing a biological activity and resulting from the controlled enzymatic hydrolysis, especially under in vitro conditions; from bacterial fermentation, which occurs directly on the food; from the gastrointestinal digestion *in vivo*; and/or from food preparation. The main biological activities of these bioactive compounds are associated with antioxidant, hypolipidemic, anti-inflammatory, hypoglycemic, and antihypertensive activities, offering health benefits that are related to blood pressure control, the cardiometabolic disease control, the reduction of the risk of cardiovascular events, and to antiaging (Nasri, 2016; Walther & Sieber, 2011).

2.2.1 Health Benefits of Bioactive Compounds

Bioactive compounds may also be designated as nutraceuticals, given the fact that they combine both the nutrients and the pharmaceutical benefits when present in food products, which are directly isolated from medicinal plants and/or nutrients, hence offering a nutritional value that can also be used as a drug (Lordan et al., 2011; Nasri et al., 2014). In the following paragraphs, some health benefits of bioactive compounds are presented, aiming to better understand some of the therapeutical applications of these bioactive compounds, as well as their main bioactivities.

a) Antibacterial activity

Finding new antibacterial drugs able to combat antibiotic-resistant bacteria is an urgency in the scientific research area (Martinez & Baquero, 2014). Many of the infections that we are currently subjected to are difficult to treat with existing antibiotics, which are quickly losing their effectiveness because of antibiotic-resistant bacteria (Alos, 2015). In this context, some natural chemical substances from plants have emerged as unique potential therapeutic tools (Alvarez-Martinez et al., 2020a, 2020b; Kokoska et al., 2019).

Plants generate an incalculable source of secondary metabolites as a reaction to environmental cues and triggers, such as attacks by herbivores animals, interspecific relations, or even abiotic stress (Yang et al., 2018). Currently, the well-known and researched mechanisms of action among antimicrobial drugs are associated with a wide diversity of bacterial targets and processes, such as inhibition of metabolic pathways, bacterial membrane lysis, interference in cell-wall synthesis,

inhibition of protein synthesis, inhibition of deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) synthesis, among others. On the other hand, the most frequent and explored processes of bacterial resistance to antibacterial drugs are antibiotics modified by enzymes, antibiotic inactivation, and the expression effect of efflux pumps (Chandra et al., 2017; Khameneh et al., 2019). There are various reviews regarding plant-based antimicrobials drugs in the last years (Chandra et al., 2017).

Nevertheless, there are some examples showing promising activity, like the phenolic essential oil compounds constituted by hinokitiol, carvacrol, thymol, and menthol that can serve as a guideline for the use these phenolic essential oil compounds for oral health care products and food preservation, since dental caries and periodontitis represent the major oral infectious diseases. Bacterial plaques composed of native oral flora accumulate on dental surfaces and are the primary etiological agents of periodontal disease and dental caries (Botelho et al., 2007).

The extract of *Oxalis corniculata* is an example of a compound exhibiting intense antimicrobial activity against *Escherichia coli*, multi-drug resistant - *Salmonella typhi*, *Klebsiella pneumoniae*, and *Citrobacter koseri* (Manandhar et al., 2019). For instance, five local plants that have been used by indigenous populations in Iraq were investigated for their antimicrobial activity against *Staphylococcus epidermidis* and *Klebsiella pneumoniae*, showing that the solvents used, and the concentration of the extract strongly impacted their activity, being *Punica granatum* L. the most active extract (Al-Sa'ady, 2020). Another example is the sanguinarine substance extracted from root and aerial parts of *Chelidonium majus* L. which showed a strong effect against *Staphylococcus aureus* strains with a minimum inhibitory concentration of 1.9 mg/L (Cushnie et al., 2014).

Most of antibacterial drugs from natural origin derive from a set of terrestrial bacteria, nonetheless, the marine bacteria have molecular, biochemical, and physiological characteristics that are distinct from their terrestrial counterparts and, therefore, they can produce different bioactive molecules (Butler et al., 2013; Siddharth & Vittal, 2018). The major bacteria species discovered in seawater belong to the genera *Achromobacter sp*., *Flavobacterium sp.*, *Micrococcus sp.*, *Pseudomonas sp*., and *Vibrio sp*. (Baharum et al., 2010). However, the genus *Streptomyces* was the one that most significantly provided new molecules (Blunt et al., 2018).

Al-Dhabi et al. (2019) isolated from the Saudi Arabian marine-derived actinomycete *Streptomyces sp.* Al-Dabhi 90 extracts showed considerable antimicrobial activity against drug resistant pathogens, namely *Escherichia coli*, *Enterococcus faecium*, *Klebsiella pneumoniae*, *Proteus mirabilis, Pseudomonas aeruginosa* and *Staphylococcus aureus*. Furthermore, they noticed that the principal substances of the *Streptomyces* extracts were octadecanoic acid, n-hexadecanoic acid, indazol-4-one, 3-methylpyridazine and 3a-methyl-6-((4-methylphenyl) sul. Thus, *Streptomyces sp.* Al-Dhabi-90 can be considered a powerful source for the development of new antibiotics to act against multidrug resistant clinical pathogens (Al-Dhabi et al., 2019).

Additionally, innovative experiments on some tunicates species and their associated bacteria, like *Streptomyces* species, demonstrated increase in secondary metabolites also with antibacterial activity when cultured simultaneously with human pathogens, such as *Pseudomonas aeruginosa*,

methicillin-sensitive *Staphylococcus aureus* (MSSA), and methicillin-resistant *S. Aureus* (MRSA) (Stincone & Brandelli, 2020; Sung et al., 2017).

A marine actinobacterium, isolated from sea sediment samples was identified as *Streptomyces sp.* S2A. One of the extracts of this organism demonstrated high antimicrobial activity against pathogen bacteria and fungi in general, antioxidant and cytotoxic activities against various cell lines and inhibitory activity against *α*-amylase and *α*-glucosidase enzymes. Moreover, the main compound of the extract in study was identified as pyrrolo[1–*α*] pyrazine-1,4-dione, hexahydro-3-(2-methylpropyl), which correspond to a peptide derived from diketopiperazine (Siddharth & Vittal, 2018).

Galaviz-Silva et al. (2018) also isolated from marine habitats in Mexico six bacteria, identified as *Bacillus aerius*, *B. altitudinis*, *B. boroniphilus*, *B. oryzicola*, *B. safensis*, and *Virgibacillus senegalensis* with severe antimicrobial activity against the food borne poisoning strains *Staphylococcus aureus* and *Vibrio parahaemolyticus*. Apparently, everything indicates that these marine bacteria can serve as a potential alternative against other clinically determining bacteria for the development of new antimicrobials (Galaviz-Silva et al., 2018). Equally, Zhang et al. (2018a) isolated a new marine actinobacterium from a marine mud sample, identified as *Streptomyces* sp. ZZ745. This new microorganism permitted identified five bagremycin analogues, including two new ones. These two new bagremycins exhibited antibacterial activity against *E. coli* (Zhang et al., 2018a).

b) Antioxidant activity

The food industry, in general is working hard towards the development of antioxidants from natural sources. Antioxidants are substances that have low concentrations that significantly inhibit or delay oxidation. For instance, food deterioration occurs due to the oxidation of lipids, resulting in the spoilage of food commodities. Since lipid oxidation by reactive oxygen species (ROS) decreases the nutritional properties of lipid enriched food, several synthetics are used to reduce such oxidation and to maintain important nutrients (Su et al., 2019). However, many researchers are continuously investigating natural antioxidants as safer alternatives, with marine organisms playing a key role is this matter, considering that they provide several bioactive compounds (He et al., 2017). Despite the deterioration of food products, ROS in excess is also associated to several diseases, such as cancer, inflammatory and neurodegenerative diseases (Assi, 2017; Collin, 2019; Forrester et al., 2018).

According to the literature, the natural antioxidants from plant materials are mainly polyphenols (phenolic acids, flavonoids, anthocyanins, lignans and stilbenes), carotenoids (xanthophylls and carotenes) and vitamins (vitamin E and C) (Amarowicz & Pegg, 2019). Generally, among all these natural antioxidants, especially polyphenols and carotenoids, exhibit a wide range of biological effects, such as anti-inflammatory, antibacterial, antiviral, anti-aging, and anticancer (Abotaleb et al., 2019; Kopustinskiene et al., 2020; Malekia et al., 2019; Rosa et al., 2019). In addition, flavonoids are

also highly effective scavengers of most oxidizing molecules, including singlet oxygen and free radicals that are implied in several diseases (Karak, 2019; Montoro et al., 2005; Tungmunnithum et al., 2018).

Flavonoids are classified according to seven different categories, more precisely: 1) chalcones; 2) anthocyanins; 3) flavones; 4) isoflavones; 5) flavanones; 6) flavononols; and 7) flavanols. Among these, anthocyanins are the largest group of phenolic compounds in the human diet, with their strong antioxidant activity being especially important to maintain health, considering that these flavonoids reduce the incidence of heart diseases and cancer (Abotaleb et al., 2019; Ignat et al., 2011; Kopustinskiene et al., 2020; Vicente & Boscaiu, 2018).

Vitamin E or also named tocopherols, as well as tocotrienols, are widely distributed in nature. Vitamin E is the most frequent name that is given to a group of lipid-soluble compounds, with *a*-tocopherol being the most well-known. This specific compound can be found in lipoproteins and membranes that block the chain reaction of lipid peroxidation, namely by scavenging the intermediate peroxyl radicals that are generated. It is important to add that the highly steric *a*-tocopherol radical is less reactive when attacking fatty acid side chains, converting back to its parent phenol through ascorbic acid, hence breaking the chain reaction (Azzi, 2018; Fritsche et al., 2017). In turn, ascorbic acid, also designated as vitamin C, is vastly known for its antioxidant activity, being frequently used in cosmetics and to treat degenerative diseases. This specific compound has several physiological functions, being important to emphasize its high antioxidant ability to recycle vitamin E in membrane and lipoprotein lipid peroxidation (Silva et al., 2019).

Carotenoids have a significant antioxidant activity mainly due to their ability to delocalize unpaired electrons through their structure of conjugated double bonds. Even though they are not very efficient as quenchers of peroxyl radicals, carotenoids are able to quench singlet oxygen, hence aiding in the protection of lipids against the peroxidative damage. *β*-carotene is the most abundant of carotenoids, being highly reactive with oxidants and electrophiles and frequently used in several therapies. Moreover, several studies have demonstrated the *β*-Carotene inhibition of lipid auto-oxidation in both biological tissues and food, which corroborates the significant antioxidant activity of this specific compound (Alves et al., 2010; Young & Lowe, 2018).

Regarding to marine bioactive compounds with antioxidant activity, stand out the peptides, proteins, pigments, polyphenols, and carbohydrates. The main benefits of bioactive peptides include the prevention of lipid peroxidation and the scavenging of ROS. Over the past years, several studies have adequately characterized, isolated, and purified bioactive peptides from distinct marine sources that have an antioxidant potential, namely: Pacific hake, cod, hoki, mackerel, jumbo squid, Alaska pollack, blue mussel, conger eel, oyster, scad, yellow stripe trevally, tuna, yellow fin sole, capelin, and microalgae (Suleria et al., 2016). According to Qian et al. (2008), and more recently by Nikoo et al. (2014), the peptides that derive from marine fish present a higher antioxidant potential, especially when compared to *α*-tocopherol in different oxidative contexts and settings. Even though the exact mechanisms involved in these antioxidant activities are not known, some aromatic amino acids do play a significant role in such activity (Nikoo et al., 2014; Qian et al., 2008).

On the other hand, phycobiliproteins (PBPs) refer to a specific class of marine proteins with antioxidant activity that is found quite abundantly in cyanobacteria, being considered as very potent therapeutics due to their antioxidant capacity, which is associated to the different side chains of constituent amino acids. Furthermore, PBPs are also used as natural dyes in both cosmetic and food industries, being applied to change the color of many food products, such as milk shakes, desserts, fermented milk, and ice creams (Sonani et al., 2014). In sum, several marine-derived bioactive peptides and proteins have demonstrated a potent antioxidant activity, which proves that they truly have an important role in nutraceuticals and pharmaceuticals industries.

Marine algae are one of the main and richest sources of antioxidants in marine biota. Indeed, both microalgae and macro-algae (seaweed) were shown to decrease the ROS due to their abundance in bioactive compounds such as polyphenols, pigments, and vitamins. Some investigations proved that chlorophyll a, as well as related compounds that derive from brown algae, present antioxidant activities in methyl linolenate systems. However, chlorophyll b derivatives presented a stronger antioxidant activity, which clearly suggests that the presence of an aldehyde group in chlorophyll b provides a better antioxidant activity than the methyl group in chlorophyll a (Cornish & Garbary, 2010).

Fucoxanthin is one of the main antioxidant molecules that has a potential to scavenge for free radicals, which appears to be related to the presence of unusual double allenic bonds at the C-70 position. Fucoxanthin, as well as the metabolites fucoxanthinol and halocynthiaxanthin, have their antioxidative activity derived from *Undaria pinnatifida*, despite presenting different levels of efficiency. Indeed, fucoxanthin has the highest antioxidant activity, being immediately followed by the fucoxanthinol and by the halocynthiaxanthin due to the presence of an allenic bond in both the fucoxanthin and the fucoxanthinol (Liu et al., 2020; Sachindra et al., 2007).

Phlorotannins, which derive from the marine brown algae, have demonstrated a strong antioxidant activity against the free radical mediated oxidation, a potent antioxidant and protective effect against $H_2O_2^-$induced cell damage, an antioxidant activity in phospholipid peroxidation, and an antioxidant activity against 2,2-diphenyl-1-picrylhydrazyl (DPPH) and superoxide. Essentially, all phlorotannins present the potential to be natural antioxidants for both the food and the pharmaceutical industries (Shibata et al., 2008). Thus, it is possible to conclude that polyphenolic compounds and pigments that derive from marine algae present strong antioxidant activities, which is why they have a great potential as therapeutics in nutraceutical and pharmaceutical industries, as well as preservatives in the food industry.

Polycarbohydrates or polysaccharides are the most abundant carbohydrates found in food that are derived from marine organisms have an antioxidant activity due to their scavenging effect on ROS. As a matter of fact, fucoidans derived from the seaweed *Fucus vesiculosus* were shown to prevent the generation of hydroxyl and superoxide radicals and to decrease the lipid peroxidation, while specific fucoidan fractions derived from *Lonicera japonica* were demonstrated to have an excellent superoxide radical scavenging capacity (RSC), with low molecular weight fucoidan fractions possessing a greater inhibitory effect on low-density lipoprotein (LDL) oxidation induced by Cu^{2+}. In

sum, these findings support the beneficial effect of algal fucoidans as antioxidants, also highlighting the great potential of such molecules for alleviating oxidative stress, directly associated to diseases (Micheline et al., 2007; Zhao et al., 2011).

Even though sulfated polysaccharides are different from fucoidins, they are also very potent antioxidants. Their antioxidant capacity depends on the degree of sulfation, the molecular weight, the glycosidic branching and on the major saccharide unit. Essentially, some studies have shown that sulfated polysaccharides with different glycosidic branching, degrees of sulfation, and molecular weight are more useful, being considered as effective, non-toxic substances with a strong antioxidant activity/capacity (Qi et al., 2005; Sun et al., 2009).

c) Other bioactivities

Inflammation is a process that involves several biological pathways, which can be directed by either external or internal stimuli and consists in a particularly important part of the host's responses to many stimuli, including injury, immune reactions, or even microbial invasion. Its biological pathways can be reduced, inhibited, or modulated by specific compounds, often designated as anti-inflammatories. However, the main role of anti-inflammatories is to modulate macrophages, referring to an important player of the inflammation process, given the fact that they are the predominant source of the inflammatory mediators, such as IL-6, prostaglandin E2 (PGE2), interleukin-1β (IL-1β), pro-inflammatory cytokines, nitric oxide (NO), and a few other ROS (Block et al., 2007). Inflammation predisposes to the development of cancer and promotes all stages of tumorigenesis. Cancer cells, as well as surrounding stromal and inflammatory cells, engage in well-orchestrated reciprocal interactions (Fridlender et al., 2015).

According to some studies, it has been demonstrated that grains play defensive roles in inflammation, that hydrolysates or peptides derived from cereals have anti-inflammatory effects, and that peptides derived from several legumes can modulate inflammatory processes (Lee et al., 2015), which is the case of protein hydrolysates produced from rice (*Oryza sativa* L.) on the inflammatory response of RAW 264.7 macrophages, suggesting that they attenuate the inflammatory markers. In their study, these authors also emphasize the effects of a fraction that is designated as RPHs-C-7-3, since it suppressed the release of NO and Tumor Necrosis Factor (TNF)-α by the cells, as well as the phagocytic ability of the activated macrophages (Wen et al., 2016). Regarding other types of cereal, Montoya-Rodríguez and González De Mejía (2015) proved that peptides derived from amaranth proteins suppressed some inflammatory markers on THP-1 cells (Montoya-Rodríguez & González De Mejía, 2015). Concerning to legumes, which are another type of grains that have anti-inflammatory properties and bioactivities, López-Barrios et al. (2016) established that protein hydrolysates derived from common dry bean (*Phaseolus vulgaris* L.) significantly inhibit the NO production by RAW 264.7 cells induced with lipopolysaccharides (LPS) (López-Barrios et al., 2016).

It has been recognized that several bioactive compounds after being isolated and purified from their natural sources, can be used to boost immunity.

Moringa oleifera (MO), for instance, consists in an indigenous tree from the north of India and Nepal with all its components (bark, flowers, seeds, and leaves) being considered as medicinal. In more detail, the MO's leaves is the most frequently used component to treat fever, malaria, high blood pressure, arthritis, parasitic diseases, diabetes, HIV/AIDS, and skin lesions, the main active metabolites are quercetin, myricetin, isorhamnetin, kaempferol, caffeic acid, rutin, ellagic acid, chlorogenic acid, gallic acid, ferulic acid, N *α*-L-rhamnopyranosyl vincosamide, benzyl, 4-(*α*-L-rhamnopyranosyloxy) phenylacetonitrile (Niazirin), and sinalbin (Kou et al., 2018; Leone et al., 2015; Rani et al., 2018).

Another interesting example is the sulfated polysaccharides derived from seaweeds that have been proven effects on innate immunity, especially by modulating the ability of the immune cells to produce nitric oxide, consequently reducing the existing inflammation. Similarly, fucoidans derived from marine algae were shown to inhibit inflammatory responses in several *in vitro* studies and can be used for cancer immunotherapy, considering that they influence the activation and maturation of human monocyte-derived dendritic cells (Manikandan et al., 2020; Park et al., 2011).

Extracts of *Oscillatoria* and *Nostocmuscorum* sp. also have anti-tumor activity, mainly due to their inhibitory effect on the human hepatocellular cancer cell line (HepG2) and on Ehrlich's Ascites Carcinoma Cells (EACC). Curacin-A is a peptide that was isolated from *Lyngbya majuscule* and is also known for its anti-proliferative properties in several tumor cell lines, such as breast, colon and renal. Some of the most important discoveries refer to cryptophycin 1 and 8, cyanovirin, and borophycin, the latter being a boron-containing metabolite which is frequently purified from cyanobacterial strains, hence having cytotoxic effects against human colorectal cells (Alves et al., 2018; Kiruba et al., 2018, Suleria et al., 2016; Wali et al., 2019).

Finally, it is pertinent to mention that some algae are known and distinguished for their inter-conversion of fatty acids, from their simple arachidonic acid form to complex eicosanoids, which can play a major role in maintaining homeostasis and curing cancer, asthma, ulcers, heart disease, arteriosclerosis, and psoriasis. Moreover, different mechanisms of marine algae have been reported, especially due to their capacity of promoting anti-cancer activity, such as immune stimulation, apoptotic cell death, and anti-oxidation (Carrol et al., 2020; Mayer et al., 2021; Sithrangaboopathy & Kathiresan, 2010).

2.3 Bioactive Compounds Isolated from Fungi

Fungi are one of the most important organisms in the world, with about 611 000 fungal species estimated worldwide. Fungi have a long, well documented history as medicines, with their use in Chinese traditional and Ayurvedic medicines, as well as in Western medicine (Harvey et al., 2015). Fungi have produced several medicinally important compounds, including penicillin, mevinolin

(lovastatin), fingolimod, and caspofungin (VanderMolen et al., 2013). For example, atorvastatin, a fungal metabolite-based anti-cholesterol drug, was the top-grossing pharmaceutical in the world (Beekman & Barrow, 2014; Calixto, 2019).

Since the discovery of penicillin in 1928 by Fleming, several studies mainly focused on soil-derived fungi, demonstrated that these microorganisms are a rich and unique source of bioactive compounds (Letek, 2020; Tan & Tatsumura, 2015). Throughout the years, extensive screening programs were developed all around the world, and great efforts have been dedicated with the aim of the isolation of new bioactive compounds from fungi (Berman & Krysan, 2020; Pye et al., 2017). Reported in several studies, extracts from fungus have a broad spectrum of biological activity (Hyde et al., 2019).

Later, in 1948, Guiseppe Brotzu, while screening microbes for antibiotic activity, examined fungi cultured from a sewer outflow in the Mediterranean off Sardinia. Brotzu had observed that the seawater was being purified around this sewer and believed a microorganism to be responsible. The fungus *Cephalosporium acremonium* demonstrated potent antibiotic activity and Brotzu found that extracts from this fungus displayed activity towards both gram-positive and gram-negative bacteria. Subsequently, a sample of *C. acremonium* was examined by researchers at Oxford, who isolated several antibacterial compounds (Brotzu, 1948). Many of the isolated compounds possessed insignificant activity, but notably Abraham and Newton isolated penicillin N (originally named cephalosporin N), which was active against both gram-positive and gram-negative bacteria (Newton & Abraham, 1953). Additionally, they isolated cephalosporin C, which is closely related to penicillin N, but showed a greater ability to withstand b-lactamases and possessed lower toxicity (Newton & Abraham, 1955). Cephalosporin C was intended to be marketed as an antibiotic for penicillin-resistant bacteria, but the introduction of methicillin, a more potent antibiotic with similar b-lactamase stability, halted this. Subsequently, the investigation of cephalosporin analogues was performed in the hope of discovering greater activity (Beekman & Barrow, 2014). Nevertheless, although the first cephalosporin C antibiotic was obtained from a fungus in 1948, the discovery rate of biologically active metabolites from microorganisms remained very slow until the 1980s. Nevertheless, since 1985, more than 15 000 new metabolites have been identified from fungi organisms, many of which have displayed interesting bioactivities, such as antibacterial (Figure 2.3) (Ameen et al., 2021).

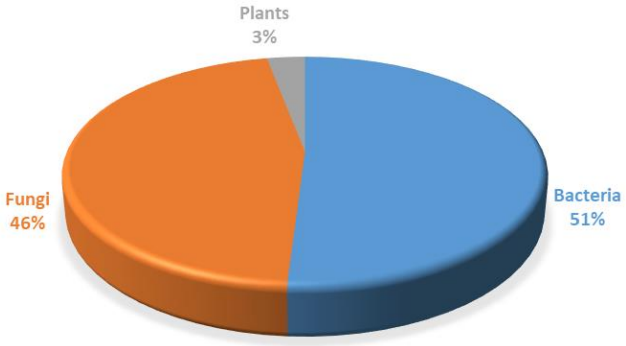

Figure 2.3. Principal sources of FDA-approved antibacterial natural products. Adapted from Patridge et al., 2016.

Fungi, more specifically endophytic fungi, have been widely used throughout the years to create new drugs with different biological activities and to treat several conditions/diseases. Endophytes refer to the chemical compound that is produced inside plants, with many of them being able to synthesize bioactive compounds that can be used as budding sources of pharmaceutical leads. As a matter of fact, these fungi have been considered as extremely useful for innovative drug discovery, with several investigations confirming the production of novel antifungal, antibacterial, anti-inflammatory, antiviral, anti-tumor, and other substances that belong to the steroid, alkaloid, terpenoid, and flavonoid extracts (Selvakumar & Panneerselvam, 2018).

Paclitaxel, more frequently called taxol, is a highly functionalized diterpenoid that is found in each yew (*Taxus*) species, even though it was originally isolated from *Taxus brevifolia*. Overall, this bioactive compound is the world's first billion-dollar anticancer drug, being used to treat breast and ovarian cancers, as well as other human tissue-proliferating diseases (Sekhon et al., 2020; Stage et al., 2018). Its high cost makes it unavailable to most people, given the fact that yew trees might support certain endophytic microorganisms that can also synthesize Taxol. Nonetheless, it was found that taxol is also produced by several other *Pestalotiopsis microspora* isolates, especially by those obtained from the bald cypress in South Carolina. This was the first indication that taxol could be produced by several other fungi, rather than just by the original yew trees (Gallego-Jara et al., 2020; Naik, 2019; Strobel et al., 1997).

Broader studies include probing for paclitaxel-generating fungi from *Taxus* and other associated plant species. Indeed, relatively 19 genera of fungi were demonstrated to generate paclitaxel in addition to its counter path, more precisely the following: *Alternaria, Aspergillus, Botryodiplodia, Botrytis, Cladosporium, Ectostroma, Fusarium, Metarhizium, Monochaetia, Mucor, Ozonium, Papulaspora, Periconia, Pestalotia, Pestalotiopsis, Phyllosticta, Pithomyces, Taxomyces*, and *Tubercularia* (Gallego-Jara et al., 2020; Naik, 2019; Strobel et al., 1997).

Another example is camptothecin which consists in a pentacyclic quinoline alkaloid that was firstly isolated from the wood of *Camptotheca acuminate*. This bioactive compound acts as an antineoplastic agent, with its primary action mechanism referring to the inhibition of the intranuclear enzyme topoisomerase-1, which is required in the DNA transcription and replication during molecular events. Both the hycamtin (topotecan) and camtostar (irinotecan), which consists of two of the most famous semi-synthetic drugs, have already been used to treat small lung, ovarian, and refractory ovarian cancers. *Fusarium solani*, isolated from the *Camptotheca acuminate* fungus, produces CPT, 9-methoxycamptothecin, and 10-hydroxycamptothecin (Kusari et al., 2009).

According to Selvakumar and Panneerselvam (2018), more recent studies demonstrated that two fungi *Fusarium solani* strains, MTCC9667 and MTCC9668, can generate CPT (cytidine triphosphate), 9-methoxycamptothecin, and 10-hydroxycamptothecin, while *Entrophospora infrequens* can also produce camptothecin (Selvakumar and Panneerselvam, 2018).

Both, the vinblastine and the vincristine, which refer to the terpenoid indole alkaloid derivative as of the combination of vindoline and catharanthine monomers, are anticancer agents (Zhang et al., 2018c). Regarding vincristine, its primary action mechanism is through the interference with microtubule formation and mitotic spindle dynamics, disruption of the intracellular transport, and decreased tumor blood flow, with the latter probably resulting from the anti-angiogenesis. The fungus *Alternaria* sp. was isolated commencing the phloem of *Catharanthus roseus*, presenting the ability to produce vinblastine, while the endophytic *Fusarium oxysporum* was isolated from the phloem of *C. roseus*, presenting the ability to produce vincristine. Nevertheless, an unidentified fungus from the leaves of *C. roseus* might be able to produce both vincristine and vinblastine (Alam, et al. 2017a; Qu et al., 2018; Ramezani et al., 2018; Taher et al., 2019).

Enniatins have insecticidal, phytotoxic, and antibiotic activities that are directly linked to their ionophoric properties. Commencing G, which refers to a new compound with a similar structure to the cyclohexapeptide, was also isolated from the culture broth of mangrove fungus *Fusarium* sp., displaying an anti-tumor activity. Essentially, enniatins are cyclohexadepsipeptides, consisting of three *D*-2-hydroxyisovaleric acid residues alternatively linked with L-amino acids or *N*-methyl-*L*-amino acids to result in an 18-membered cyclic skeleton (Gautier et al., 2020; Lin & Zhou, 2003; Olleik et al., 2019).

Trichodermin derives from the *Trichoderma harzianum*, which consists of a fungus that survives in *Ilex cornuta*. Overall, trichodermin was announced towards guard facing the solanaceous plant germs *Alternaria solani* and *Rhizoctonia solani ex vivo* status (Chen et al., 2007; Cheng et al., 2010; Poveda et al., 2019). Barúa et al., (2019) developed the hemisynthesis of trichodermin and trichodermol derivatives and demonstrated their antimicrobial and cytotoxic activities (Barúa et al., 2019). Table 2.1 summarizes the main bioactive compounds, as well as their fungi sources and biological activities.

Table 2.1. Bioactive compounds are derived from fungi and respective activities. Adapted from Selvakumar & Panneerselvam, 2018.

Fungi	Bioactive compounds	Bioactivities
Acremonium zeae	Pyrrocidines A and B	Antifungal activity against *A. flavus* and *F. verticillioides*
Aspergillus fumigatus CY018	Asperfumoid, fumigaclavine C, fumitremorgin C, physcion and helvolic acid	Restrict *Candida albicans*
Cephalosporium sp. IFB-E001	Graphislactone A	Free radical scavenging
Cephalotheca faveolata	Sclerotiorin	Antibacterial
Chaetomium globosum	Chaetoglobosins A and C	Inhibit *Mucor miehei*
Chaetomium globosum IFB-E019	Chaetoglobosin	Cytotoxic
	U, C, F, E	
Cladosporium sp.	Brefeldin A	Antifungal activity
Cryptosporiopsis cf. *quercina*	Cryptocin and Cryptocandian	Antifungal activity against
		Human pathogens:
		Candida albicans
		Trichophyton sp.
		Plant pathogens:
		Sclerotinia sclerotiorum
		Botrytis cinerea
		Pyricularia oryzae
Discosia sp.	*Amylase*	Enzymatic
Entrophospora infrequens	Camptothecin	Anticancer
Fusarium oxysporum	Vinca alkaloids	Leukaemia
Fusarium solani	Berberine	Anticancer
Fusarium solani LCPANCF01	Taxol	Anticancer
Fusidium sp.	Fusidikactones	Antifungal activity
Hypericum perforatum	Hypericin and emodin	Antibacterial activity:
		Staphylococcus aureus
		Pseudomonas aeruginosa
		Klebsiella pneumoniae
		Salmonella enterica

Fungi	Bioactive compounds	Bioactivities
		Antifungal activity:
		A. niger
		C. albicans
Penicillium janthinellum	Citrinin	Antimicrobial
Penicillium sp.	Alkaloid A, B	Bacteriostatic effect:
		E.coli
		Staphylococcus aureus
		Pseudomonas aeruginosa
		Bacillus
Periconia sp.	Periconicin A	Antibacterial
	Piperine	Inhibit *M. tuberculosis* and *M. smegmatis*
Pestalotiopsis adusta	Pestalachlorides	Antifungal activity against
		Plant pathogens
		Fusarium culmorum
		Gibberella zeae
		Verticillium albo-atrum
Pestalotiopsis jester	Jesterone	Antifungal activity against plant pathogens
Pestalotiopsis microspora	Ambuic acid and pestalotiopsins A and B	Antifungal activity
	Pestaloside	Phytotoxic properties
	Taxol	Anticancer
Phomopsis cassia	Ethyl 2,4 dihydroxy-5,6-dimethyl benzoate and Phomopsis lactone	Antimycotic action facing *Cladosporium cladosporioides*
		C. sphaerospermum
Phomopsis sp.	Cytosporone B and C	Inhibit *Candida albicans, F. oxysporum*
Rhizoctonia sp.	Rhizoctonic acid	Anti-helicobacter pylori activity
Talaromyces sp.	7-epiaustdiol	Inhibit multidrug-resistant opportunistic pathogen *P. aeruginosa*
Xylaria sp.	Griseofulvin	To treat human and animal mycotic diseases

Many fungal metabolites in the pharmaceutical market reveal the potential of fungi as valuable sources of lead drugs. Song et al. (2018) observed that the fungus *Penicillium* sp. ZZ380, collected from a wild crab, generated new seven pyrrospyrone alkaloids, one of which exhibited high anti-glioma activity, and the other three showed antimicrobial activity against MRSA and *E. coli* with MIC values between 2.0-5.0 µg/mL.

Agrawal et al. (2018) isolated five new marine fungal species, identified as *Aspergillus sydowii*, *Leptosphaerulina* sp., *Penicillium citrinum*, *Penicillium chrysogenum,* and *Simplicillium lamellicola* with great antimicrobial activity against the acne-inducing bacteria *Cutibacterium acnes* and *Staphylococcus epidermidis* with a MIC ranging from 0.8 to 1.0 mg/mL (Agrawal et al., 2018).

Pang et al. (2018) found three new and six already known compounds from the rice cultivation extract of the sponge-derived fungus *Trichoderma* sp. SCSIO41004. However, only one already known isolated compound, named 5-acetyl-2-methoxy-1,4,6-trihydroxy-anthraquinone demonstrated significant antiviral activity against the human Enterovirus 71 (EV71).

Li et al. (2018) discovered six new peniciphenalenins and the previously identified compounds (+)-sclerodin, (+)-sclerodione, (+)-scleroderolide, and physcion from cultured extracts of the marine-derived fungus *Penicillium* sp. ZZ901. The known (+)-scleroderolide metabolite demonstrated inhibition of glioma cell growth and showed antibacterial activity against the pathogenic bacteria *S. aureus* and *E. coli*, with a MIC of 7.0 µg/mL and 9.0 µg/mL, respectively.

Luo et al. (2018a) isolated a unique anti-methicillin-resistant *S. aureus* (MRSA) compound from the deep marine fungus *Fusarium* sp. 152, identified as equisetin and showing a MIC value of 1 µg/mL against MRSA.

Luo et al. (2018b) identified twenty-eight novel aromatic polyketides from the mangrove associated fungus *Diaporthe* sp. SCSIO 41011. Four of these isolated compounds demonstrated high antiviral activity against three variations of the influenza A virus (IAV). These new compounds were identified as 2 pestalotiopsone, 3,8-dihydroxy-6-methyl-9- oxo-9H-xanthene-1-carboxylate and 5-chloroisorotiorin.

Wang et al. (2018) found four novel chlorinated compounds (chaetovirides A, B, and C, and chaephilone C) and four known ones derived from azaphilone (cochliodone A, chaetoviridin A, chaetomugilin D, and chaetoviridine E) isolated from the deep-sea derived fungus *Chaetomium* sp. NA-S01-R1. Chaetoviride A and B presented antibacterial activity against Vibrio rotiferianus and Vibrio vulnificus. Additionally, chaetoviride B and C and, chaephilone C showed anti-MRSA activity. Finally, chaetoviride A presented cytotoxic activity against HepG2 tumor cells and chateoviride B and chaephilone C against HeLa tumor cells.

Zhao et al. (2018) isolated eight novel isoprenylated cyclohexanols and fourteen known analogous compounds from solid cultures of the sponge-associated fungus *Truncatella angustata*. From these new isoprenylated cyclohexanols identified, one of them showed inhibition against both human immunodeficiency virus 1 (HIV-1) and swine-origin influenza A virus (H1N1) and another one towards the HIV-1 virus.

Marine fungi have gained a growing interest as sources of bioactive compounds, especially for biotechnological applications, and due to their capacity of producing secondary metabolites with potential biological and pharmacological activities (Imhoff, 2016; Mora et al., 2011). Compared to the average number of new bioactive compounds reported for the past three years, new bioactive compounds from marine fungi increased by 85%. Furthermore, it is important to add that marine-derived fungi are significant sources for new bioactive secondary metabolites, such as glycosides, alkaloids, lipids, peptides, polyketides, proteins, and terpenoids, many of which present anti-tumor, antifungal, antiviral, and antibacterial properties (Ameen et al., 2021; Carroll et al., 2021; Wang et al., 2015).

In the last years, new compounds were identified, being isolated from marine-derived fungi from the following sources: marine animals (30.1%); mangrove (25.5%); sediment (22.9%); algae (14.4%); seawater (4.6%); and other (2.6%) (Figure 2.4) (Carroll et al., 2021; Jin et al., 2016; Mayer et al., 2021).

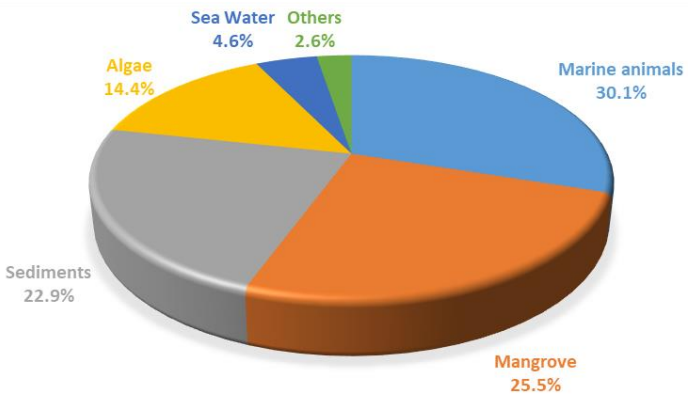

Figure 2.4. New bioactive compounds from marine-derived fungi according to their respective sources. Adapted from Jin et al., 2016.

Furthermore, the analysis of the new chemical structures, which resulted from the isolation of marine-derived fungi, demonstrated that both the alkaloids (27.0%) and the polyketides (25.7%) refer to the main chemical classes, being immediately followed by the peptides (13.8%), the terpenes (9.9%), the lactones (3.9%), and the steroids (3.3%) (Figure 2.5) (Carroll et al., 2021; Jin et al., 2016; Mayer et al., 2021).

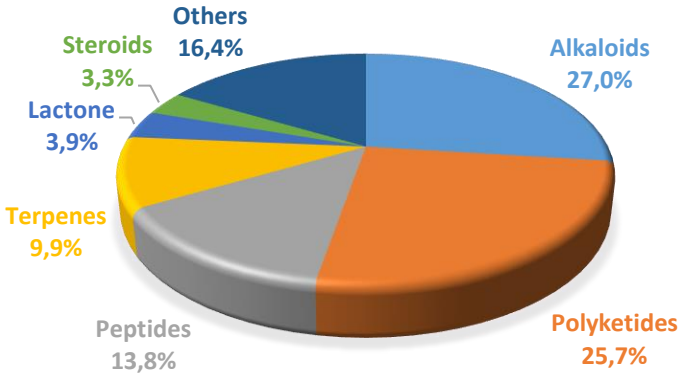

Figure 2.5. New bioactive compounds from marine-derived fungi according to their structural types. Adapted from Jin et al., 2016.

Lastly, among the biological activities that most of the new compounds demonstrated, it is noteworthy to mention the following ones: cytotoxicity (37.5%), antimicrobial activity (33.5% - with antibacterial representing 18.4%, antifungal 7.9%, and antiviral 7.2%), antioxidant activity (5.3%), lipid-lowering activity (5.3%), and other activities (18.4%) (Figure 2.6) (Carroll et al., 2021; Jin et al., 2016; Mayer et al., 2021).

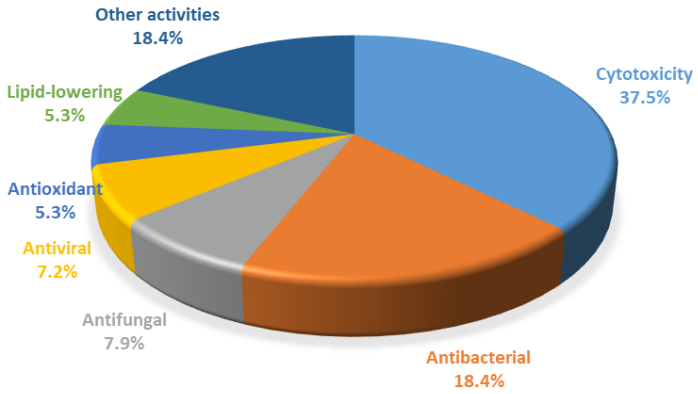

Figure 2.6. Bioactive categories of the new compounds isolated from marine-derived fungi between 2014 and 2015. Adapted from Jin et al., 2016.

The following table (Table 2.2) presents important information regarding the new compounds that were isolated from the marine-derived fungi, namely the sources of the fungal strains, their structural types, and their bioactive categories.

Table 2.2. New compounds isolated from marine-derived fungi according to their sources, structural type, and bioactive category. Adapted from Jin et al., 2016.

Fungi sources	Bioactive compounds	Structural type	Bioactivities
Sponge	Arthpyrones	Alkaloids	Cytotoxicity
	Chartarlactams	Phenylspirodrimanes	Lipid-lowering
	5-methoxydihydrosterigmatocystin	Secondary metabolites	Antimicrobial
	Diaporthalasin	Pentacyclic cytochalasin	Antibacterial
	Chevalone E	Chevalone derivative	Antibiotic
	Xylarianaphthol-1	Dinaphthofuran derivative	Cytotoxicity
	Lindgomycin	Polyketide	Antibiotic
Coral	Territrem derivatives	Lactones	Inhibitory, antifouling, antiviral
	Butyrolactone derivative		Antiviral
	Oxalicumones	Dihydrothiophene-condensed chromones	Cytotoxicity
	Nigrospins B and C	Citrinins	Antifungal
	Nucleoside derivatives	Nuclesoside	Antibacterial
	Eurothiocin A and B	Benzofuran derivatives	Inhibitory
	Chondrosterin	Metabolite	Cytotoxicity
	Steroid derivative	Steroid	Inhibitory
	Talaromycin A	Diphenyl ether derivative	Antifouling
Starfish	Glioxotin derivative	Gliotoxin	Inhibitory, cytotoxicity
	Pseudaboydins A	Isobenzofuranone derivative	Cytotoxicity
Bryozoan	Isaridin G, Desmethylisaridin G, and Desmethylisaridin C1	Cyclohexadepsipeptides	Antibacterial
	Iso-isariin D	Cyclodepsipeptide	Inhibitory
	Felinones A and B	*O*-containing heterocyclic compounds	Inhibitory
Sea urchin	2*E*,4*Z*-tanzawaic acid D	Tanzawaic acid derivative	Anti-inflammatory
Fish	Rubrolides R and S	Rubrolides	Antioxidant, antiviral

Fungi sources	Bioactive compounds	Structural type	Bioactivities
Prawn	Indole-diterpenoids derivatives	Indole-diterpenoids	Antibacterial
Mangrove	Peaurantiogriseols A, B, C, D, E, and F	Polyketides	Inhibitory
	Aspergifuranone	Isocoumarin derivatives	Inhibitory
	Isocoumarin derivative		
	Brocaeloid B	Alkaloid	Inhibitory
	Phomazines B	Thiodiketopiperazine	Cytotoxicity
	Brocazines	Diketopiperazine derivatives	Cytotoxicity
	(3*R*,4*S*)-3,4-dihydro-8-hydroxy-4-methoxy-3-methylisocoumarin	Isochroman	Acceleration of growth
	Penicibrocazines A, B, C, D, and E	Diketopiperazine derivatives	Antimicrobial
	Biogenetically related compounds	Biogenetically related	Inhibitory
	Isso-monodictyphenone	Benzophenone	Inhibitory
	Penikellides A and B	Diphenyl ether derivatives	Antibacterial, inhibitory
	Vaccinal A	Naphthalene derivative	Inhibitory
	Phthalide	New compound	Antibacterial
	Dihydroisocoumarin	New compound	Cytotoxicity
	Pestalamine A	Aromatic amine	Cytotoxicity
	Flavipesins A	Aromatic butyrolactone	Antibacterial, antibiofilm
	Vaccinol I	Prenylated phenol	Inhibitory
	Penicibilaenes A and B	Sesquiterpenes	Selective
	Resveratrodehydes A, B, and C	Resveratrol derivatives	Inhibitory, cytotoxicity, antioxidant
	Cyclo(*D*-Pro-*L*-Tyr-*L*-Pro-*L*-Tyr), cyclo(Gly-*L*-Phe-*L*-Pro-*L*-Tyr), and cyclo(*L*-leucyl-*trans*-4-hydroxy-*L*-prolyl-*D*-leucyl-*trans*-4-hydroxy-*L*-proline)	Cyclic tetrapeptides	Antifungal
Sediment	Sorbicatechols A and B	Sorbicillinoids	Antiviral

Fungi sources	Bioactive compounds	Structural type	Bioactivities
	Penicillipyrones B	Meroterpenoid	Induction of quinone reductase
	N-methyl-pretrichodermamide B	Epidithiodiketopiperazine	Cytotoxicity
	New polyketide	Polyketide	Cytotoxicity
	13-*O*-prenyl-26-hydroxyverruculogen	Indolediketopiperazine peroxide	Inhibitory
	Eleganketal A	Aromatic polyketide	Antiviral
	Speradines G and H	Alkaloids	Cytotoxicity
	Psychrophilins	Peptide	Lipid-lowering
	Penipalines B and C	Alkaloids	Cytotoxicity
	New versicamide H	Versicamide	Inhibitory
	Penicimutanolone, penicimutanin A, penicimutanin B, and penicimutatin	Modified diethyl sulphate mutagenesis	Cytotoxicity
	New C25 steroids	Steroids	Cytotoxicity
	New lipopeptides	Lipopeptides	Cytotoxicity
	(+)-6-*O*-Demethylpestalotiopsin A and (+)-6-*O*-demethylpestalotiopsin C	Caryophyllene derivatives	Inhibitory
	Cladosporin A and Cladosporin B	Diketopiperazines	Cytotoxicity
	14-hydroxy-cyclopeptine	Cyclic dipeptide	Inhibitory
	Trichobotrysins	Tetramic acid derivatives	Cytotoxicity, antiviral
Algae	Isocyathisterol	Ergosteroid derivative	Antibacterial
	New polyketides	Polyketides	Transcriptional, cytotoxicity, scavenging
	New austalide meroterpenoids	Austalide meroterpenoids	Transcriptional, inhibitory
	Arisugacin K	Meroterpene	Inhibitory
	6b,9a-dihydroxy-14-*p*-nitrobenzoylcinnamolide	Nitrobenzoyl sesquiterpenoid	Cytotoxicity, antiviral
	Varioxepine A	3*H*-oxepine-containing alkaloid	Antibacterial, inhibitory

Fungi sources	Bioactive compounds	Structural type	Bioactivities
	Butyrolactone IX and aspulvinone O	Butenolides	Scavenging
	Butanoate	Benzamide derivative	Scavenging
	Varioloids A and B	Alkaloids	Antifungal
	Eudesma-4(15),7-diene-5,11-diol	Eudesmane sesquiterpenoid	Antifungal
Sea water	Trichodin A	Pyridone	Antibiotic
	New metabolites	Metabolites	Cytotoxicity
	Novel spirocyclic drimane	Spirocyclic drimane	Antibacterial
	Penicilliumine	New structure	Acetylcholinesterase
	2-(2-hydroxypropanamido) benzoic acid	Benzoic acid	Anti-inflammatory, analgesic
Others	(±)-asperlone A and B	Racemic dinaphthalenone derivatives	Inhibitory
	Microsphaerol	Polychlorinated triphenyl diether	Antibacterial, antialgal, antifungal
	Seimatorone	Naphthalene derivative	Antibacterial, antialgal, antifungal
	Questinol	Anthraquinone derivative	Inhibitory, anti-inflammatory
	Aspochalasin V	Aspochalasin	Cytotoxicity
	Penicillosides A and B	Cerebrosides	Antifungal, antibacterial

2.4 Fungal Importance

Notably, fungi are a valuable natural source of bioactive compounds, with important applications in industry, technology, and medicine. Considering the extraordinary relevance of the fungi-associated bioactive compounds, due to their antioxidant, anti-inflammatory, anti-diabetic, anti-cancer, antibacterial, antifungal, and immunomodulatory effects, it is of utmost importance to understand its common characteristics and its several organisms.

Fungi refer to eukaryotic organisms with a relatively simple morphology, characterized by being spore-bearing, filamentous, heterotrophic, and for having absorptive nutrition and decomposing function. These organisms lack chlorophyll in opposition to plants, and may reproduce sexually and asexually (McLaughlin et al., 2009; Willey et al., 2011).

Despite these general characteristics, several attempts have been made, to attain a genetic, molecular or cellular character that specifically defines this kingdom, i.e., a synapomorphy (unifying character to the Fungi) that allows distinguishing Fungi from another organism like protists, although without great success (Choi & Kim, 2017; Richards et al., 2017).

Fungal species have extraordinary importance for humans, either due to beneficial or deleterious effects. For instance, together with other microbes, in particular chemoorganotrophic organisms, fungi may act as decomposers in several ecosystems, triggering a crucial feature in the balance of the ecosystems and also displaying essential roles in industry and technology (Liang & Gadd, 2017). Several yeasts are key elements in fermentation processes, enabling the production of several food products like cheese, wine, beer, and bread. Another relevant beneficial aspect is the application of the yeast, such as *Saccharomyces cerevisae* as a study model for the eukaryotic cell. Despite being well known for its functions in baking and brewing, the budding yeast *Saccharomyces cerevisiae* is found in a wide range of habitats and has been adapted for use in the laboratory. *S. cerevisiae*'s reputation as a strong genetic model system stem in part from its remarkably efficient homologous recombination, which allows researchers to easily modify yeast genes (Hanson, 2018). Fungi are also important in the production of various organic acids, such as citric acid and gallic acid, immunosuppressors (cyclosporine), and antibiotics, like penicillin, or griseofulvin.

Nonetheless, fungi are also a major cause of disease, not only for plants, which have a tremendous impact on the agriculture, crops, and garden plants and promote a disbalance in the ecosystems but also in animals and humans, as is the case of *Aspergillus* sp. (Aflatoxicosis), *Candida albicans* (candidiasis) and *Cryptococcus* sp. (Hall & Noverr, 2017; Köhler et al., 2017). Fungi may be pathogenic due to four major features: ability to grow at or above 37 ºC (thermal resistance), capacity to reach internal tissues through (or circumventing) host barriers, lysis, and absorption of animal tissues, and resistance to the host's immune defenses (Köhler et al., 2017).

2.4.1 Fungal Distribution and Ecology

Fungi species are predominantly present in terrestrial environments, although freshwater or marine systems are also generous sources. They can be found everywhere, from the stratosphere to depths of the Dead Sea or oceanic sediments, from Antarctic glaciers to torrid deserts, as well as the gut of flies. Fungi are globally distributed in all ecosystems, predominantly in dark and moist habitats, interact with a broad array of organisms, and may be pathogenic to plants or animals, in particular humans, originating mycoses. Nowadays, approximately 135 000 species have been described, although the estimated diversity surpasses 1 million species (Naranjo-Ortiz & Gabaldón, 2019; Taylor et al., 2014; Willey et al., 2011).

One of the most exciting aspects concerning fungal ecology is the relationships they form with other organisms. Fungi form beneficial symbiotic relationships, as in the case of the associations with vascular plant roots, called mycorrhizae, or in the case of lichens, associations of fungi and photosynthetic protists or cyanobacteria (Naranjo-Ortiz & Gabaldón, 2019). This happens because of the fungal osmotrophic function and loss of phagotrophy, where nutrients are processed in the extracellular environment, allowing metabolites processed outside the cell to be accessible to any other organisms that occupy the nearby ecosystem (Martínez-García et al., 2017; Richards et al., 2017).

Evolutionary, fungi developed two common biological features to maximize osmotrophic public good functions, namely the presence of hyphal networks and a wide array of secondary metabolisms. Fungal growth as multicellular hyphal networks enables to potentiate of the three-dimensional colonization of the surrounding area, minimizing the loss of metabolites to competitor organisms and obtaining further nutritional sources. In addition, their diversified metabolism provides several antimicrobial toxins to exclude neighboring microbes that do not possess the cognate gene cluster encoding certain detoxification phenotypes from an environment where the public good is present and the toxin is produced (Keller et al., 2005; Richards & Talbot, 2013; Richards et al., 2017).

2.4.2 Fungal Physiology

Fungal physiology encompasses the structure and morphology, metabolism, nutrition, growth, reproduction, and death of fungal cells (Walker & White, 2017). In terms of structure, the body or vegetative structure of a fungus is called a thallus, with a multiplicity of formats and sizes. It may range from a unicellular and microscopic form, as in the case of yeasts, to multicellular mycelia molds composed by filaments of hyphae, macroscopic puffballs, and the paradigmatic mushrooms (Walker & White, 2017; Willey et al., 2011).

Yeasts, particularly important in technologic and industrial processes, are unicellular and mononuclear fungus, able to reproduce asexually (budding and transverse division) or sexually (forming spores). With variable sizes, but mainly with spherical — oval appearance, these organisms

possess most of the eukaryotic organelles, being *Saccharomyces cerevisiae* and *Candida albicans* relevant examples (Copetti, 2019; Legrend et al., 2019; Rai et al., 2019). Spores are quintessential in the fungal physiology. They permit fungal survival in extreme conditions, as the case of nutrient scarcity, lack of water, extreme temperatures, and pH. They are easily disseminated, contributing to the wide and easy distribution of these organisms (Dijksterhuis, 2019).

Molds are characterized for having long and branched filaments of cells named hyphae, whose aggregates form mycelium, macroscopic masses growing in soil and other microenvironments. Hyphae can be septate, i.e., have septa (cross walls with one or several pores) or can be aseptate (coenocytic). The presence of hyphal networks is useful, allowing increasing the surface area to nutrient absorption (Copetti, 2019; Willey et al., 2011). Some fungi forms are dimorphic, with the two forms, and may switch from the yeast form (for instance in animals) to the mold or mycelial form in soil. *Candida albicans* and *Blastomyces dermatitidis* are known examples of such dimorphic organisms. Dimorphic switching requires the fungus to sense and respond to the host environment and is essential for its pathogenicity (Boyce & Andrianopoulos, 2015; Wall, 2020).

Another important aspect to be mentioned about fungal physiology concerns metabolism and nutrition. Fungi may be found in every ecosystem if organic material is available. They use organic compounds to originate carbon, electrons and energy and use carbohydrates (glucose or maltose) and nitrogenous compounds to produce their own amino acids and proteins (Walker & White, 2017; Willey et al., 2011).

These organisms are usually saprophytes, as they obtain their nutrients from dead organic material. Fungal species can digest external substrates in their surrounding area, by hydrolytic exoenzymes, allowing its subsequent absorption, a process called osmotrophy, already approached in previous subsection. Although predominantly aerobic, there are some species known for being facultative anaerobes or obligatory anaerobes, as those inhabiting the rumen of cattle like the class of Neocallimastigomycetes (Spatafora et al., 2017; Watkinson et al., 2015).

In fungi, reproduction can occur either asexually or sexually. After a period of intensive growth, fungi undergo a reproductive phase that culminates in the formation of spores, which can be produced sexually or asexually. Regarding asexual reproduction, a single individual cell originates a genetic duplicate progeny without contribution from another organism through several manners. A progenitor cell can divide into two daughter cells after mitosis, by transverse fission, occurring a central constriction and consequent new cell wall formation occurs. Another simple method for asexual reproduction is by fragmentation of the thallus or by budding, quite common in yeasts and in some filamentous fungi, where a bud develops on the surface of the fungi, with shared cytoplasm between the bud and the parent cell. Then, the nucleus of the parent cell divides with one of the daughter nuclei emerging into the bud. Several buds can be produced from the same origin, producing a chain of fungi cells. Finally, other way of asexual reproduction is by spore production. This can be present in hyphal fragmentation that results in each fragmenting hypha to behave as a spore (arthrospores or arthroconidia). Also, if the spores inside a sac (sporangium), in the tip of a hypha, they are called sporangiospores. In the case when the spores are not enclosed inside the

sac, although still produced in the hyphal tip, they are named conidiospores. Alternatively, if the cells are surrounded by a wall prior to the separation phase, the spores are called chlamydospores. Buds that are pinched off a hypha behave as spores - blastopores (Ahmadjian et al., 2020; Watkinson et al., 2015; Willey et al., 2011).

Concerning sexual reproduction, its importance can be attributed to the genetic variability. Indeed, sexual reproduction can accelerate adaptation to novel or changing ecosystems or to stressful conditions, generating novel strains. This type of reproduction involves the fusion of compatible nuclei, generally between mycelia of opposite mating types, and consists in three phases. First, plasmogamy involves the fusion of the two protoplasts, bringing together the two compatible haploid nuclei. This sexual fusion occurs in several ways, as between gametes, gametangia (gamete-producing bodies) or hyphae. After, in the karyogamy phase, those two nuclei fuse, hence forming a diploid nucleus, originating the zygote. Finally, a meiosis phase occurs, restoring the haploid phase, whose nuclei are incorporated in spores (meiospores as zygospores, ascospores or basidiospores) (Ahmadjian et al., 2020; Ashu & Xu, 2015; Willey et al., 2011).

Regarding to fungal cell walls, they are dynamic structures essential for cell viability, morphogenesis, and pathogenesis. The wall is much more than an outer layer of the fungus, it is also a dynamic organelle whose composition considerably influences the ecology of the fungus and whose composition is highly regulated in response to environmental conditions and forced stresses (Gow et al., 2017). Fungi undergo cell death mechanisms with similarity to apoptosis, pyroptosis and necrotopsis. Meiotic drive elements also cause death of sensitive spores in processes, like toxin-antidote systems (Rico-Ramírez et al., 2022).

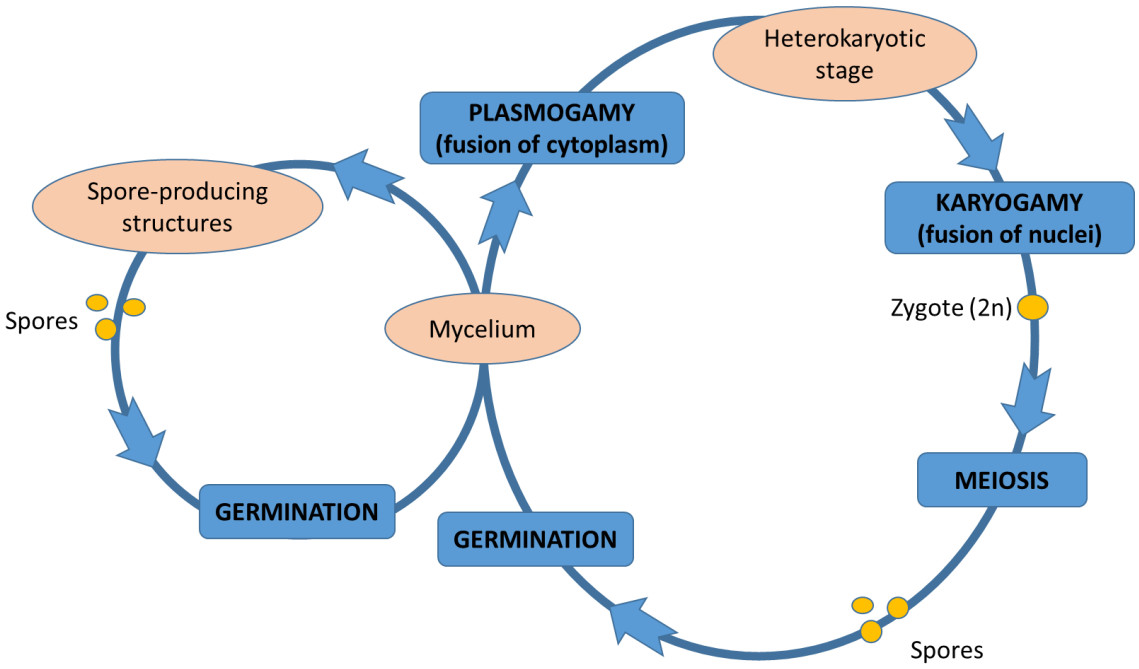

Figure 2.7. Fungi reproduction processes.

2.5 Technologies/Techniques for the Extraction of Fungi Bioactive Compounds

The extraction of bioactive compounds from fungi follows a multi-step process (Figure 2.8). One of the first steps to be considered is the culture conditions since fungi must grow in a nutritional media at the laboratory. Therefore, special attention should be considered for the temperature, incubation time, aeration, media composition, and pH, since less favorable conditions can affect not only the fungal biomass production but also the yield of the bioactive compound(s) of interest, as these are often produced in response to specific environmental stressors or triggers. The next step includes extraction of active samples, followed by the structural characterization of the bioactive compounds and the biological activity screening. The bioactivity assay is an indispensable test in the research of new molecules, usually conducted to measure the activity of a potential new biopharmaceutical (Gomes et al., 2016a).

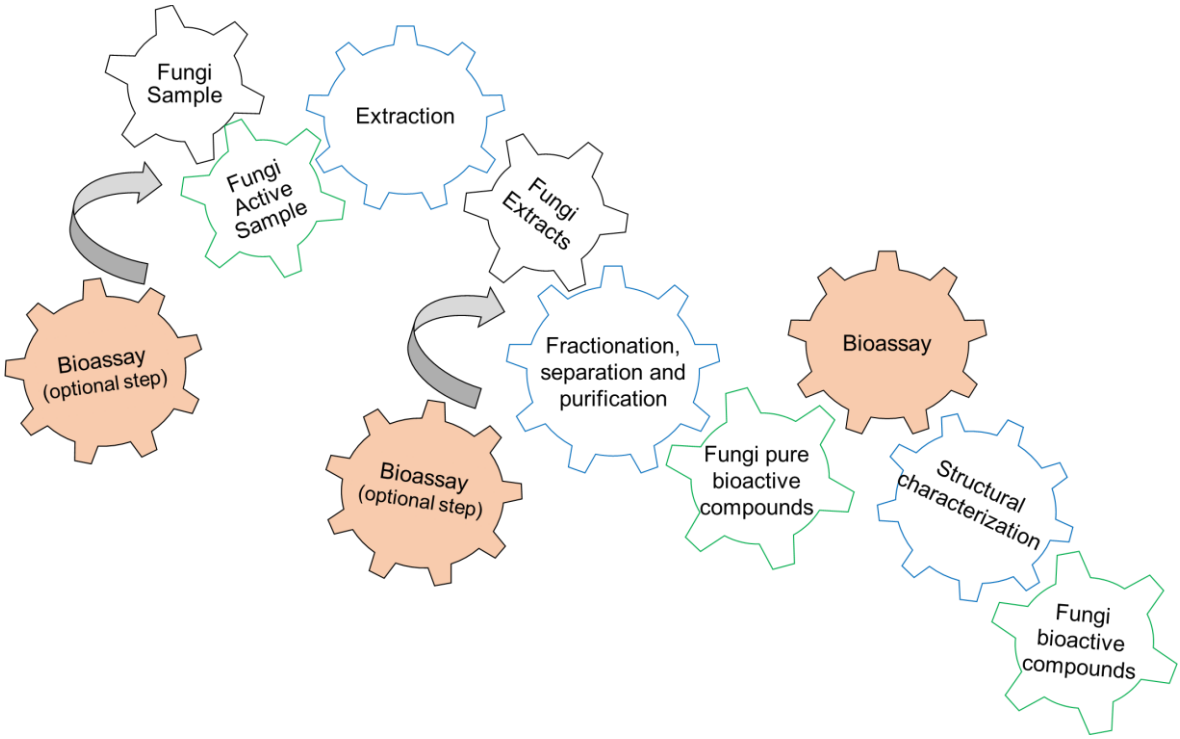

Figure 2.8. Typical process for obtaining bioactive compounds and optional bioassay steps. Adapted from Rocha Santos & Duarte, 2014.

Extraction refers to the first step that must occur to separate the desired natural products from the raw materials/sources. Some of the extraction techniques include solvent extraction, the distillation method, sublimation and pressing, according to the extraction principle, with the solvent extraction being the most widely used method. Overall, the process of extraction of natural products comprehends 4 distinct phases, more precisely: 1) the penetration of the solvent into the solid matrix;

2) the dissolution of the solute in the solvents; 3) the diffusion of the solute out of the solid matrix; and 4) the collection of the extracted solutes. Nevertheless, it is important to mention that the extraction's efficiency depends on several factors, namely on the particle size of the raw materials, the properties of the extraction, the extraction temperature, the solvent-to-solid ration, and the extraction duration (Zhang et al., 2018b).

The traditional extraction methods frequently use organic solvents and require a large volume of solvents, as well as a long extraction time. Some modern/greener extraction methods, which are emerging based on the use of more sustainable technologies and methods, are also applied in the extraction of natural products, offering several advantages such as shorter extraction time, lower organic solvent consumption, and higher selectivity (Grosso et al., 2015; Zhang et al., 2018b). Table 2.3 presents a summary of the existing extraction methods for natural products, both traditional and emerging techniques, and is based on the study conducted by Zhang et al. (2018b).

In the following paragraphs, the traditional extraction method, and the emerging technologies for extraction, which comprehend different techniques of extraction, are presented, aiming to provide a deeper understanding of their main characteristics and benefits.

2.5.1 Traditional Extraction

The traditional extraction, also designated as solid-liquid extraction (SLE) or solvent extraction, is the most frequently used extraction technique and it can be done in different ways, more precisely: 1) boiling the solvent and the sample with or without stirring for a certain duration; 2) refluxing by using a Soxhlet; and 3) maceration with continuous stirring, with a long duration. Some of the main solvents, or mixture of solvents, that are commonly used in this specific technique are ethanol, methanol, acetone, trichloromethane, ethyl acetate, a mixture of water with organic solvents (acetone, ethanol, methanol, and acetonitrile) at different mixing ratios (Catarino et al., 2019; Justino et al., 2014).

Nevertheless, traditional extraction methods often involve the use of a large volume of solvents, high temperatures, and longer extraction times, which might result in the hydrolysis or in the oxidation of the compounds. Furthermore, the use of these techniques at an industrial level would be extremely difficult, especially due to their demands in terms of energy, practicality, economic factor, and environmental considerations (Ojha et al., 2020). Therefore, emerging extraction technologies are being often introduced, aiming to overcome such challenges.

2.5.2 Emerging Technologies for Extraction

Regarding to the emerging technologies for extraction, these are essentially based on the used mechanism/source of energy when conducting the extraction of natural products and, among them, there is one that stands out, named High-Pressure Extraction (HPE).

The HPE (Figure 2.9) consists of a new technique that uses high pressure to extract active ingredients mostly from plant materials, with its application being extremely useful to obtain compounds that will ultimately help develop healthier products. It has been established that HPE showed an effective remedy at a low extraction rate and improved the quality and the concentration of the obtained compounds. This specific technique can increase the mass transfer rate by changing the concentration diffusivity and gradient, resulting in the damage of the cell membrane, and increasing its permeability, despite also enhancing the permeating of the extraction solvent into the cells. Thus, these properties result in a shorter processing time, a reduction in costs, an improvement in the safety of the process, and the achievement of higher compound yields (Huang et al., 2013).

Some other important characteristics of the HPE technique that must be mentioned are directly associated with some of its benefits, more precisely: 1) its operation is very quick, easy, and safe; 2) it maintains the natural structures of the bioactive compounds; 3) it consumes very little energy while simultaneously being very efficient; and 4) it has a reduced potential of polluting the environment and has no adverse effects on human health, therefore promoting health and safety both to the environment and to the overall population. In sum, it is possible to conclude that the HPE consists of a technology that allows the extraction efficiency to be significantly increased, with high yields being obtained in a few minutes. Since it can be operated at room temperature, the bioactivity of compounds with low thermal stability can be adequately protected, and it also protects the environment, because there is no volatilization of the solvent. HPE is an environmentally friendly, rapid, and highly efficient extraction method is suitable and beneficial to both the pharmaceutical and food industries (Huang et al., 2013).

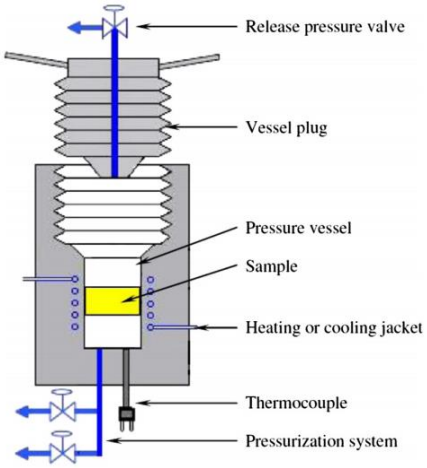

Figure 2.9. Schematic representation of the High-Pressure Extraction (HPE) system. Extracted from Huang et al., 2013 and reprinted with permission of Elsevier.

2.5.3 Other Emerging Technologies

The need for an efficient and green method to recover marine bioactive compounds has encouraged researchers to develop innovative techniques, with some of them still under development. Essentially, these new techniques are based on the energy sources that they use to extract bioactive compounds, such as enzyme-assisted extraction (EAE), ultrasound-assisted extraction (UAE), microwave-assisted extraction (MAE), pressurized liquid extraction (PLE), Supercritical fluid extraction (SFE), pulsed electric field extraction (PEF), and centrifugal partition extraction (CPE). Nonetheless, some new techniques are based on the use of surfactants to facilitate the extraction (SME), as well as on the type and medium of extraction, such as ionic liquids (IL), natural deep eutectic solvents (NDES), and deep eutectic solvents (DES) (Getachew et al., 2020).

Still, there are also other technologies used for this purpose though less frequently used, such as ohmic heating and pulsed electric field-assisted extraction (PEF). Additionally, over the past few years, new types of solvents have been emerging, including deep eutectic solvents and ionic liquids, which use different melting points (Garcia-Vaquero et al., 2020; Getachew, Jacobsen & Holdt, 2020).

The EAE, when compared to conventional extraction methods, the EAE stands out due to its main advantages, directly related to the following aspects: 1) overall efficacy; 2) high selectivity; 3) quick extraction; 4) low-energy consumption; 5) eco-friendly procedures; 6) minimal usage of harsh chemicals; 7) low/no wasteful protection/deprotection steps; 8) maximal yield; 9) process recyclability; and 10) facile recovery (Alam et al., 2017b; Xu et al., 2017). Moreover, enzyme-based pre-treatments, including cellulolytic, ligninolytic, and proteolytic enzymes, have been used as catalysts also help to induce the mass transfer phenomena, ultimately facilitating the release of bioactive compounds, as well as of other secondary metabolites in an efficient manner (Getachew et al., 2020; Sosa-Hernández et al., 2018).

The UAE refers to an emerging technology that can accelerate both the heat and the mass transfer, being successfully used in the extraction of bioactive compounds. Overall, the ultrasound waves change the chemical and physical properties after interacting with the exposed material. The cavitation effect of these waves facilitates the release of all extractable compounds, simultaneously enhancing the mass transfer by disrupting the cell walls (Duarte et al., 2014; Getachew et al., 2020).

The MAE, also known as microwave heating, is widely used to extract valuable materials from both animal and plant resources, being generated by dipole rotation of a polar solvent and ionic condition of dissolved ions (Zhang et al., 2018b). MAE is a green technology due to the reduction of the use of organic solvents. In this regard, it is pertinent to mention that MAE has two different methods: solvent-free extraction and solvent extraction, and that it is used in several industrial practices for the extraction of high-value bioactive phenolic compounds, functional foods, phytonutrients, and active pharmaceutical grade constituents from biomaterials (Chaturvedi, 2018; Sosa-Hernández et al., 2018).

The PLE, also referred to as subcritical water extraction (SWE) or accelerated solvent extraction (ASE), is the technique that is used when the solvent is only water and at a temperature range of

100 ºC up to 374 ºC. Indeed, the PLE takes a short time to complete the extraction, uses a lower number of solvents, requires a minimum consumption of solvents, and it comprehends a low degradation of the compounds. In addition, in the PLE it is possible to change the water's polarity by tuning the extraction temperature (Getachew et al., 2020; Sosa-Hernández et al., 2018).

In the SFE, it is very important to control the extraction parameters (diffusibility, viscosity, density, dielectric constant, and surface tension), especially to maximize the extraction yield and minimize the operation cost. The most frequently used supercritical fluid in cosmetic, pharmaceutical, and food industries is carbon dioxide ($ScCO_2$), mainly due to its low critical temperature and pressure. (Getachew et al., 2020; Lourenço et al. 2019). For instance, by increasing the extraction pressure the density of the solvent is also increased, as well as the solvating power of the $ScCO_2$, which results in an easier penetration into the sample matrix, thus facilitating the extraction rate. Nonetheless, extremely high pressure is not recommended, since it might compact the extraction bed, restricting the flow of CO_2 and reducing the diffusivity of the solvent, which ultimately reduces the extraction yield itself (Zhang et al., 2018b).

In PEF applications, high voltages are applied in the pulses of short duration to cause an electro-permeabilization and ultimately destroy the cell membranes, the latter aiming to accelerate the extraction rate. This specific technique is widely used, even though there is limited information about its application in marine resources. Nonetheless, and considering that this method is quite effective in the extraction of multiple compounds from plants, it should also be replicated to extract marine bioactive compounds (Yan et al., 2017).

The CPE, on the other hand, refers to a multi-stage liquid-liquid extraction technique that is conducted under a centrifugal field, where compounds are extracted based on the partition coefficients between the two liquid phases. Moreover, this technique is widely used for the purification of marine bioactive compounds, even though more investigations are needed to assess more alternative methods, aiming to optimize the operating conditions and extract more compounds from distinct marine sources (Getachew et al., 2020).

The use of SME is also a very promising technique for the extraction of marine bioactive compounds. Essentially, surfactants can form monomolecular layers on the liquid's surface, hence decreasing the interfacial tension between the two liquids and allowing the miscibility of both. These characteristics might enable SME to isolate compounds with a wide range of polarities, as well as complex chemical structures (Sharma et al, 2015).

In turn, ILs are types of simple molten salts, containing a relatively large organic cation and an inorganic anion. When compared to common organic solutions, ILs present several advantages, such as a low melting point, a broad liquid temperature, extended, specific, solvent properties, and negligible vapor pressure. Recently, ILs-based extraction techniques were applied to extract several bioactive compounds from plants, with studies approaching its application in marine resources being very limited (Getachew et al., 2020).

Finally, DESs refers to a system that is constituted by a mixture of two or more Lewis acids and bases, or by Bronsted-Lowry acids and bases that has the lowest freezing point. DESs result from

the complexation of a halide salt, which acts as a hydrogen-bond acceptor and a hydrogen-bond donor, having a similar physical structure to the ILs. Nevertheless, DESs are different in terms of the source of the starting ingredients and of the chemical formation process, which justifies why their applications are very distinct. In sum, DESs are used to extract bioactive compounds from terrestrial plants, with no studies published regarding the application of this technique to extract bioactive compounds from marine resources (Getachew et al., 2020).

After the extraction methodology, the characterization techniques following the principles of green chemistry should also be considered adequate choices for providing multi-analyte detection, reduction of solvents and reagents, and a decrease in time of analysis. Methodologies, such as elemental analysis (CNHS), Fourier transform infrared (FTIR), and nuclear magnetic resonance (NMR), are expected to play an important role in multi-analyte characterization and direct measurements with minimal sample preparation. After the isolation of a “pure” compound from extracts of organisms, the elucidation of its chemical structure is required, and sometimes being challenging. The structural characterization is an essential step to obtaining a potential new bioactive compound, since it allows the clarification of the skeleton of the new molecule, classifying the respective compound in the proper chemical class (Gomes et al., 2016b; Zhang & Elser, 2017). Bioactivity is the desired final goal in the entire process of extraction, and purification of fungi organisms. Thus, the designed bioassay is critical for the detection of compounds with potential therapeutic applications. The bioassays can be performed *in vitro* and/or *in vivo*, and they must be compatible with the natural product being tested. Therefore, screening systems must include a wide range of biological assays to uncover potential substance-related activities (Gomes et al., 2016b).

Table 2.3. Summary of the several extraction methods for fungi bioactive compounds. Adapted from Zhang et al., 2018b.

Method	Solvent	Temperature	Pressure	Time	Volume of organic solvent consumed	Polarity of natural products extracted
Maceration	Water, aqueous and non-aqueous solvents	Room temperature	Atmospheric	Long	Large	Dependent on extracting solvente
Percolation	Water, aqueous and non-aqueous solvents	Room temperature, often under heat	Atmospheric	Long	Large	Dependent on extracting solvent
Decoction	Water	Under heat	Atmospheric	Moderate	None	Polar compounds
Reflux extraction	Aqueous and non-aqueous solvents	Under heat	Atmospheric	Moderate	Moderate	Dependent on extracting solvent
Soxhlet extraction	Organic solvents	Under heat	Atmospheric	Long	Moderate	Dependent on extracting solvent
Pressurized liquid extraction	Water, aqueous and non-aqueous solvents	Under heat	High	Short	Small	Dependent on extracting solvent
Supercritical fluid extraction	Supercritical carbon dioxide fluid ($ScCO_2$), often with modified	Near room temperature	High	Short	None or small	Nonpolar to moderate polar compounds
Ultrasound assisted extraction	Water, aqueous and non-aqueous solvents	Room temperature, or under heat	Atmospheric	Short	Moderate	Dependent on extracting solvent
Microwave assisted extraction	Water, aqueous and non-aqueous solvents	Room temperature	Atmospheric	Short	None or moderate	Dependent on extracting solvent
Pulsed electric field extraction	Water, aqueous and non-aqueous solvents	Room temperature, or under heat	Atmospheric	Short	Moderate	Dependent on extracting solvent
Enzyme assisted extraction	Water, aqueous and non-aqueous solvents	Room temperature, or heated after enzyme treatment	Atmospheric	Moderate	Moderate	Dependent on extracting solvent
Hydro distillation and steam distillation	Water	Under heat	Atmospheric	Long	None	Essential oil (often non-polar)

2.6 References

Abotaleb, M., Samuel, S. M., Varghese, E., Varghese, S., Kubatka, P., Liskova, A., & Büsselberg, D. (2019). Flavonoids in Cancer and Apoptosis. *Cancers*, 11, 28.

Agrawal, S., Adholeya, A., Barrow, C. J., & Deshmukh, S. K. (2018). *In-vitro* evaluation of marine-derived fungi against Cutibacterium acnes. *Anaerobe*, 49, 5–13.

Ahmadjian, V., Alexopoulos, C. J., & Moore, D. (2020). Fungus. *Britannica*. Available at https://www.britannica.com/science/fungus. Consulted on January 16, 2021.

Alam, M. M., Naeem, M. Khan, M. M. A., & Uddin, M. (2017a). Vincristine and Vinblastine anticancer *Catharanthus* alkaloids: Pharmacological applications and strategies for yield improvement. *Catharanthus roseus*, 277-307.

Alam, M., Sarker, M., Ghafoor, K., Happy, R. A., & Ferdosh, S. (2017b). Bioactive compounds and extraction techniques. *Recovering Bioactive Compounds from Agricultural Wastes*, 33-53.

Al-Dhabi, N. A., Ghilan, A.-K.M., Esmail, G. A., Arasu, M. V., Duraipandiyan, V., Ponmurugan, K. (2019). Bioactivity assessment of the Saudi Arabian Marine *Streptomyces* sp. Al-Dhabi-90, metabolic profiling and its *in vitro* inhibitory property against multidrug resistant and extended-spectrum beta-lactamase clinical bacterial pathogens. *Journal of Infection and Public Health*, 12, 549-556.

Alos, J. I. (2015). Antibiotic resistance: a global crisis. *Enfermedades Infecciosas y Microbiologia Clinica*, 33, 692-699.

Al-Sa'ady, A. T. (2020). Antibacterial screening for five local medicinal plants against nosocomial pathogens: Klebsiella pneumoniae and Staphylococcus epidermidis. *EurAsian Journal of BioSciences*, 14, 553–555.

Alvarez-Martinez, F. J., Barrajon-Catalan, E., Encinar, J. A., Rodriguez-Diaz, J. C., & Micol, V. (2020a). Antimicrobial Capacity of Plant Polyphenols against gram-positive Bacteria: A Comprehensive Review. *Current Medicinal Chemistry*, 27, 2576–2606.

Alvarez-Martinez, F. J., Barrajon-Catalan, E., Micol, V. (2020b). Tackling Antibiotic Resistance with Compounds of Natural Origin: A Comprehensive Review. *Biomedicines*, 8, 405.

Alves, C., Silva, J., Pinteus, S., Gaspar, H., Alpoim, M. C., Botana, L. M., & Pedrosa, R. (2018). From marine origin to therapeutics: The anti-tumor potential of marine algae-derived compounds. *Frontiers in Pharmacology*, 1-51.

Alves, C. Q., David, J. M., David, J. P., Bahia, M. V., & Aguiar, R. M. (2010). Methods for determination of *in vitro* antioxidant activity for extracts and organic compounds. *Química Nova*, 33, 2202-2210.

Amarowicz, R., & Pegg, R. B. (2019). Chapter One - Natural antioxidants of plant origin. *Advances in Food and Nutrition Research*, 90, 1-81.

Ameen, F., AlNadhari, S., & Al-Homaidan, A. A. (2021). Marine microorganisms as an untapped source of bioactive compounds. *Saudi Journal of Biological Sciences*, 28, 224–23.

Ashu, E. E., & Xu, J. (2015). The roles of sexual and asexual reproduction in the origin and dissemination of strains causing fungal infectious disease outbreaks. *Infection, Genetics and Evolution*, 36, 199-209.

Assi, M. (2017). The differential role of reactive oxygen species in early and late stages of cancer. *American Journal of Physiology*, 313, R646–R653.

Azzi, A. (2018). Many tocopherols, one vitamin E. *Molecular Aspects of Medicine*, 61, 92-103.

Baharum, S. N., Beng, E. K., & Mokhtar, M. A. A. (2010). Marine microorganisms: potential application and challenges. *Journal of Biological Sciences*, 10, 555-564.

Barúa, J. E., Cruz, M., Pedro, N., Cautain, B., Hermosa, R., Cardoza, R. E., Gutiérrez, S., Monte, E., Vicente, F., & Collado, I. G. (2019). Synthesis of Trichodermin derivatives and their antimicrobial and cytotoxic activities. *Molecules*, 24, 3811.

Beekman, A. M., & Barrow, R. A. (2014). Fungal Metabolites as Pharmaceuticals. *Australian Journal of Chemistry*, 67, 827-843.

Berman, J., & Krysan, D. (2020). Drug resistance and tolerance in fungi. *Nature Reviews Microbiology*, 18, 319-331.

Bhargava, R., Chasen, M., Elten, M., & MacDonald, N. (2020). The effect of ginger (Zingiber officinale Roscoe) in patients with advanced cancer. *Support Care Cancer*, 28; 3279–3286.

Biesalski, H. K., Dragsted, L. O., Elmadfa, I., Grossklaus, R., Müller, M., Schrenk, D., & Weber, P. (2009). Bioactive compounds: Definition and assessment of activity. *Nutrition*, 25, 1202-1205.

Block, M., Zecca, L., & Hong, J. (2007). Microglia-mediated neurotoxicity: Uncovering the molecular mechanisms. *Nature Reviews Neuroscience,* 8, 57-69.

Blunt, J. W., Carroll, A. R., Copp, B. R., Davis, R. A., Keyzers, R. A., & Prinsep, M. R. (2018). Marine natural products. *Natural Product Reports*, 35, 8-53.

Botelho, M. A., Nogueira, N. A., Bastos, G. M., Fonseca, S. G., Lemos, T. L., Matos, F. J., Montenegro, D., Heukelbach, J., Rao, V. S., & Brito, G. A. C. (2007). Antimicrobialactivity of the essential oil from Lippia sidoides, carvacrol and thymol against oral pathogens. *Brazilian Journal of Medical and Biological Research*, 40, 349–356.

Boyce, K. J., & Andrianopoulos, A. (2015). Fungal dimorphism: The switch from hyphae to yeast is a specialized morphogenetic adaptation allowing colonization of a host. *FEMS Microbiology Reviews*, 39, 797-811.

Brotzu, G. (1948). Ricerche su di un nuovo antibiotico. *Labori dell'stituto d'lgiene di Cagliari*. 4-18.

Butler, M. S., Blaskovich, M. A., & Cooper, M. A. (2013). Antibiotics in the clinical pipeline in 2013. *The Journal of Antibiotics*, 66, 571-591.

Caritá, A. C., Fonseca-Santos, B., Shultz, J. D. Michniak-Kohn, B., Chorilli, M., & Leonardi, G. R. (2020). Vitamin C: One compound, several uses. Advances for delivery, efficiency and stability. *Nanomedicine: Nanotechnology, Biology and Medicine*, 24, 102117.

Calixto, J. (2019). The role of natural products in modern drug discovery. *Annals of the Brazilian Academy of Sciences*, 91, e20190105.

Carroll, A. R., Copp, B. R., Davis, R. A., Keyzers, R. A., & Prinsep, M. R. (2021). Marine natural products. *Natural Product Reports*, 38, 362-413.

Catarino, M., Silva, A., Mateus, N., & Cardoso, S. (2019). Optimization of phlorotannins extraction from fucus vesiculosus and evaluation of their potential to prevent metabolic disorders. *Marine Drugs,* 17, 162.

Chambial, S., Dwivedi, S., Shukla, K. K., John, P. J., & Sharma, P. (2013). Vitamin C in disease prevention and cure: An overview. *Indian Journal of Clinical Biochemistry*, 28, 314-328.

Chandra, H., Bishnoi, P., Yadav, A., Patni, B., Mishra, A. P., & Nautiyal, A. R. (2017). Antimicrobial Resistance and the Alternative Resources with Special Emphasis on Plant-Based Antimicrobials-A Review. *Plants*, 6, 16.

Chaturvedi, A. K. (2018). Extraction of neutraceuticals from plants by microwave assisted extraction. *Systematic Reviews in Pharmacy,* 9, 31-35.

Chen, L., Chen, J., Zheng, X., Zhang, J., & Yu, X. (2007). Identification and antifungal activity of the metabolite of endophytic fungi isolated from *Ilex cornuta*. *Nongyaoxue Xuebao*, 9, 143-150.

Cheng, J. L., Zhou, Y., Zhao, J. H. Zhang, C., & Lin, F. C. (2010). Synthesis and antifungal activity of trichodermin derivatives. *Chinese Chemical Letters*, 21, 1037-1040.

Choi, J., & Kim, S. (2017). A genome Tree of Life for the Fungi kingdom. *PNAS*, 114, 9391-9396.

Collin, F. (2019). Chemical basis of reactive oxygen species reactivity and involvement in neurodegenerative diseases. *International Journal of Molecular Sciences*, 20, 1-17.

Copetti, M. V. (2019). Yeasts and molds in fermented food production: an ancient bioprocess. *Current Opinion in Food Science*, 25, 57-61.

Cornish, M., & Garbary, D. (2010). Antioxidants from macroalgae: Potential applications in human health and nutrition. *Algae,* 25, 155-171.

Cragg, G. M., & Newman D. J. (2013). Natural products: A continuing source of novel drug leads. *Biochimica et Biophysica Acta*, 1830, 3670-3695.Denis, I., Potier, B., Heberden, C., & Vancassel, S. (2015). Omega-3 polyunsaturated fatty acids and brain aging. *Current Opinion in Clinical Nutrition and Metabolic Care*, 18, 139-146.

Cushnie, T. P. T., Cushnie, B., & Lamb, A. J. (2014). Alkaloids: an overview of their antibacterial, antibiotic-enhancing and antivirulence activities. International Journal of Antimicrobial Agents, 44, 377–386.

Duarte, K., Justino, C. I. L., Gomes, A. M., Rocha-Santos, T., & Duarte, A. C (2014). Green analytical methodologies for preparation of extracts and analysis of bioactive compounds. In T. Rocha-Santos, & A. Duarte (Eds), *Analysis of Marine Samples in Search of Bioactive Compounds* (pp. 35-57). Elsevier.

Farley, G., Riggs, D. W., Bhatnagar, A., & Hellmann, J. (2021). Omega-3 polyunsaturated fatty acids modify the inverse association between systemic inflammation and cardiovascular fitness. *Clinical Nutrition*, article in press.

Forrester, S. J., Kikuchi, D. S., Hernandes, M. S., Xu, Q., & Griendling, K. K. (2018). Reactive oxygen species in metabolic and inflammatory signaling. *Circulation Research*, 122, 877–902.

Fridlender, M., Kapulnik, Y., & Koltai, H. (2015). Plant derived substances with anti-cancer activity: From folklore to practice. *Frontiers in Plant Science*, 6, 1-9.

Fritsche, S., Wang, X., & Jung, C. (2017). Recent advances in our understanding of tocopherol biosynthesis in plants: An overview of key genes, functions, and breeding of vitamin E improved crops. *Antioxidants*, 6, 99.

Galaviz-Silva, L., Iracheta-Villarreal, J. M., Molina-Garza, Z. J. (2018). *Bacillus* and *Virgibacillus* strains isolated from three Mexican coasts antagonize *Staphylococcus aureus* and *Vibrio parahaemolyticus*. *FEMS microbiology letters 365*, fny202.

Gallego-Jara, J., Lozano-Terol, G., Sola-Martínez, R. A., Cánovas-Díaz, M., & Puente, T. D. (2020). A compressive review about Taxol®: History and future challenges. *Molecules*, 25, 5986.

Garcia-Vaquero, M., Ummat, V., Tiwari, B., & Rajauria, G. (2020). Exploring ultrasound, microwave and ultrasound-microwave assisted extraction technologies to increase the extraction of bioactive compounds and antioxidants from brown macroalgae. *Marine Drugs*, 18, 172.

Gautier, C., Pinson-Gadais, L., & Richard-Forget, F. (2020). *Fusarium Mycotoxins* Enniatins: An updated review of their occurrence, the producing *Fusarium* Species, and the abiotic determinants of their accumulation in crop harvests. *Journal of agricultural and Food Chemistry*, 68, 4788-4798.

Getachew, A. T., Jacobsen, C., & Holdt, S. L. (2020). Emerging technologies for the extraction of marine phenolics: Opportunities and challenges. *Marine Drugs*, 18, 389.

Gomes, A. R., Duarte, A. C., & Rocha-santos, T. (2016a). Medical applications of marine natural products. In O. P. Jenkins (Ed.), *Advances in animal science and zoology* (pp. 71-97). New York: Nova Science Publishers, Inc.

Gomes, A. R., Duarte, A. C., & Rocha-santos, T. A. P. (2016b). Analytical techniques for discovery of bioactive compounds from marine fungi. In J. M. Merillon, & K. Ramawat (Eds), *Fungal Metabolites* (1-20). Springer, Cham.

Gomes, A. R., Freitas, A. C., Duarte, A. C., & Rocha-Santos, T. A. P. (2017). Clinical trials for deriving bioactive compounds from marine invertebrates. *Natural Products in Clinical Trials,* 1, 3-36.

Gow, N. A. R., Latge, J.-P., & Munro, C. A. (2017). The Fungal Cell Wall: Structure, Biosynthesis, and Function. *The Fungal Kingdom*, 267-292.

Grosso, A., Valentão, P., Ferreres, F., & Andrade, P. B. (2015). Alternative and efficient extraction methods for marine-derived compounds. *Marine Drugs*, 13, 3182-3230.

Hall, R. A., & Noverr, M. C. (2017). Fungal interactions with the human host: exploring the spectrum of symbiosis. *Current Opinion in Microbiology*, 40, 58-64.

Hanson, P. K. (2018). *Saccharomyces cerevisiae*: A Unicellular Model Genetic Organism of Enduring Importance. *Current Protocols Essential Laboratory Techniques*, 16, e21.

Harvey, A. L., Edrada-Ebel, R., & Quinn, R. J. (2015). The re-emergence of natural products for drug discovery in the genomics era. Nature Reviews Drug Discovery, 14, 111-119

He, L., He, T., Farrar, S., Ji, L., Liu, T., & Ma, X. (2017). Antioxidants maintain cellular redox homeostasis by elimination of reactive oxygen species. *Cellular Physiology and Biochemistry*, 44, 532-553.

Huang, H.-W., Hsu, C.-P., Yang, B. B., & Wang, C.-Y. (2013). Advances in the extraction of natural ingredients by high pressure extraction technology. *Trends in Food Science & Technology,* 33, 54-62.

Hyde, K. D., Xu, J., Rapior, S., Jeewon, R., Lumyong, S., Niego, A. G. T., Abeywickrama, P. D., … & Stadler M. (2019). The amazing potential of fungi: 50 ways we can exploit fungi industrially. *Fungal Diversity*, 97, 1-136.

Ignat, I., Volf, I., & Popa, V. I. (2011). A critical review of methods for characterization of polyphenolic compound in fruits and vegetables. *Food Chemistry*, 126, 1821-1835.

Imhoff, J. F. (2016). Natural products from marine fungi: Still an underrepresented resource. *Marine Drugs*, 14, 1-19.

Jin, L., Quan, C., Hou, X., & Fan, S. (2016). Potential pharmacological resources: Natural bioactive compounds from marine-derived fungi. *Marine Drugs,* 14, 1-25.

Justino, C. I. L., Duarte, K., Freitas, A. C., Duarte, A. C., & Rocha-Santos, T. (2014). Classical Methodologies for Preparation of Extracts and Fractions. In T. Rocha-Santos, & A. Duarte (Eds), *Analysis of Marine Samples in Search of Bioactive Compounds* (pp. 35-57). Elsevier.

Karak, P. (2019). Biological activities of flavonoids: An overview. *International Journal of Pharmaceutical Sciences and Research*, 10, 1567-1574.

Khameneh, B., Iranshahy, M., Soheili, V., & Fazly Bazzaz, B. S. (2019). Review on plant antimicrobials: a mechanistic viewpoint. *Antimicrobial Resistance and Infection Control*, 8, 118.

Kiruba, N. J. M., Pradeep, M. A., & Thatheyus, A. J. (2018). Discovering promising anti-cancer drug candidates from marine algae. *Science International*, 6, 44-50.

Köhler, J. R., Hube, B., Puccia, R., Casadevall, A., & Perfect, J. R. (2017). Fungi that infect humans. *The Fungal Kingdom*, 811-843.

Kokoska, L., Kloucek, P., Leuner, O., & Novy, P. (2019). Plant-derived products as antibacterial and antifungal agents in human health care. *Current Medicinal Chemistry*, 26, 5501–5541.

Kopustinskiene, D. M., Jakstas, V., Savickas, A., & Bernatoniene, J. (2020). Flavonoids as Anticancer Agents. *Nutrients*, 12, 457.

Kou, X., Li, B., Olayanju, J., Drake, J., & Chen, N. (2018). Nutraceutical or pharmacological potential of *Moringa oleífera* Lam. *Nutrients*, 10, 1-12.

Krause, J., & Tobin, G. (2013). Discovery, development, and regulation of natural products. In M. Kulka (Ed.), *Using old solutions to new problems: Natural drug discovery in the 21st century* (pp. 3-36). s.l.; InTech.

Kusari, S., Zühlke, S., & Spiteller, M. (2009). An endophytic fungus from Camptotheca acuminata that produces Camptothecin and analogues. *Journal of Natural Products*, 72, 2-7.

Lee, Y. M., Han, S. I., Song, B. C., & Yeum, K. J. (2015). Bioactives in commonly consumed cereal grains: Implications for oxidative stress and inflammation. *Journal of Medicinal Food*, 18, 1-8.

Leone, A., Spada, A., Battezzati, A., Schiraldi, A., Aristil, J., & Bertoli, S. (2015). Genetic, ethnopharmacology, phytochemistry and pharmacology of *Moringa oleifera* leaves: An overview. *International Journal of Molecular Sciences*, 16, 12791-12835.

Letek, M. (2020). Alexander Fleming, The discoverer of the antibiotic effects of Penicillin. *Frontiers for Young Minds*, 7.

Li, Q., Zhu, R., Yi, W., Chai, W., Zhang, Z., & Lian, X.-Y. (2018). Peniciphenalenins A-F from the culture of a marine-associated fungus Penicillium sp. ZZ901. *Phytochemistry*, 152, 53-60.

Liang, X., & Gadd, G. M. (2017). Metal and metalloid biorecovery using fungi. *Microbial Biotechnology*, 10, 1199-1205.

Lin, Y. C., & Zhou, S. N. (2003). Marine microorganisms and its metabolites. *Chemical Industry Press*, 426-427.

Liu, M., Li, W., Chen, Y., Wan, X., & Wang, J. (2020). Fucoxanthin: A promising compound for human inflammation-related diseases. *Life Sciences*, 255, 117850.

López-Barrios, L., Antunes-Ricardo, M., & Gutiérrez-Uribe, J. A. (2016). Changes in antioxidant and antiinflammatory activity of black bean (*Phaseolus vulgaris* L.) protein isolates due to germination and enzymatic digestion. *Food Chemistry*, 203, 417-424.

Lordan, S., Ross, R. P., & Stanton, C. (2011). Marine bioactives as functional food ingredients: Potential to reduce the incidence of chronic diseases. *Marine Drugs*, 9, 1056-1100.

Lourenço, S. C., Moldão-Martins, M., & Alves, V. D. (2019). Antioxidants of Natural Plant Origins: From Sources to Food Industry Applications. *Molecules*, 24, 4132.

Luo, M., Ming, Y., Wang, L., Li, Y., Li, B., Chen, J., & Shi, S. (2018a). Local delivery of deep marine fungus-derived equisetin from polyvinylpyrrolidone (PVP) nanofibers for anti-MRSA activity. *Chemical Engineering Journal*, 350, 157-163.

Luo, X., Yang, J., Chen, F., Lin, X., Chen, C., Zhou, X., Liu, S., & Liu, Y. (2018b). Structurally diverse polyketides from the mangrove-derived fungus *diaporthe* sp. SCSIO 41011 with their anti-influenza A virus activities. *Frontiers in Chemistry*, 6, 282.

Malekia, S. J., Crespo, J., F., & Cabanillas, B. (2019). Anti-inflammatory effects of flavonoids. Food Chemistry, 299, 1-11.

Manandhar, S., Luitel, S., & Dahal, R. K. (2019). *In vitro* antimicrobial activity of some medicinal plants against human pathogenic bacteria. *Journal of Tropical Medicine*, 2019, 1-5.

Martins, A., Vieira, H., Gaspar, H., & Santos, S. (2014). Marketed marine natural products in the pharmaceutical and cosmeceutical industries: Tips for success. *Marine Drugs*, 12, 1066-1101.

Martinez, J. L., & Baquero, F. (2014). Emergence and spread of antibiotic resistance: setting a parameter space. *Upsala Journal of Medicinal Sciences*, 119, 68–77.

Martínez-García, L. B., De Deyn, G. B., Pugnaire, F. I., Kothamasi, D., & van der Heijden, M. G. A. (2017). Symbiotic soil fungi enhance ecosystem resilience to climate change. *Global Change Biology*, 23, 5228-5236.

Mayer, A. M. S., Guerrero, A. J., Rodríguez, A. D., Taglialatela-Scafati, O., Nakamura, F., & Fusetani, N. (2021). Marine Pharmacology in 2016-2017: Marine Compounds with Antibacterial, Antidiabetic, Antifungal, Anti-Inflammatory, Antiprotozoal, Antituberculosis and Antiviral Activities; Affecting the Immune and Nervous Systems, and Other Miscellaneous Mechanisms of Action. *Marine Drugs*, 19, 1-75.

McLaughlin, D. J., Hibbett, D. S., Lutzoni, F., Spatafora, J. W., & Vilgalys, R. (2009). The search for the fungal tree of life. *Trends in Microbiology,* 17, 488-497.

Montoro, P., Braca, A., Pizza, C., & De Tommasi, N. (2005). Structure-antioxidant activity relationships of flavonoids isolated from different plant species. *Food Chemistry*, 92, 349-355.

Montoya-Rodríguez, A., & González de Mejía, E. (2015). Pure peptides from amaranth (*Amaranthus hypochondriacus*) proteins inhibit LOX-1 receptor and cellular markers associated with atherosclerosis development *in vitro*. *Food Research International*, 77, 204-214.

Mora, C., Tittensor, D. P., Adl, S., Simpson, A. G. B., & Worm, B. (2011). How many species are thre on Earth and in the ocean? *PLoS Biology*, 9, 1-8.

Naik, B. S. (2019). Developments in taxol production through endophytic fungal biotechnology: a review. *Oriental Pharmacy and Experimental Medicine*, 19, 1-13.

Naranjo-Ortiz, M. A., & Gabaldón, T. (2019). Fungal evolution: diversity, taxonomy and phylogeny of the Fungi. *Biological Reviews*, 94, 2101-2137.

Nasri, H., Baradaran, A., Shirzad, H., & Rafieian-Kopaei, M. (2014). New concepts in nutraceuticals as alternative for pharmaceuticals. *International Journal of Preventive Medicine*, 5, 1487-1499.

Nasri, M. (2016). Protein hydrolysates and biopeptides: Production, biological activities, and applications in foods and health benefits. A review. *Advances in Food and Nutrition Research*, 81, 109-159.

Newman, D. J., & Cragg, G. M. (2020). Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019. *Journal of Natural Products*, 83, 770-803.

Newton, G. G. F., & Abraham, E. P. (1953). Isolation of Penicillaminic Acid and D-*α*-Aminoadipic Acid from Cephalosporin *N*. Nature, 172, 395.

Newton, G. G. F., & Abraham, E. P. (1955). Cephalosporin *C*, a New Antibiotic containing Sulphur and D-*α*-Aminoadipic Acid. *Nature*, 175, 548.

Nikoo, M., Benjakul, S., Ehsani, A., Li, J., Wu, F., Yang, N., Xu, B., Jin, Z., & Xu, X. (2014). Antioxidant and cryoprotective effects of a tetrapeptide isolated from Amur sturgeon skin gelatin. *Journal of Functional Foods*, 7, 609–620.

Ojha, K. S., Aznar, R., O'Donnell, C., & Tiwari, B. K. (2020). Ultrasound technology for the extraction of biologically active molecules from plant, animal and marine sources. *TrAC Trends in Analytical Chemistry*, 122, 115663.

Olleik, H., Nicoletti, C., Lafond, M., Courvoisier-Dezord, E., Xue, P., Hijazi, A., Baydoun, E., Perrier, J., & Maresca, M. (2019). Comparative structure–activity analysis of the antimicrobial activity, cytotoxicity, and mechanism of action of the fungal cyclohexadepsipeptides Enniatins and Beauvericin. *Toxins*, 11, 514.

Park, H. Y., Han, M. H., Park, C., Jin, C. Y., Kim, G. Y., Choi, I. W., Kim, N. D., Nam, T. J., Kwon, T. K., & Choi, Y. H. (2011). Antiinflammatory effects of fucoidan through inhibition of NF-B, MAPK and Akt activation in lipopolysaccharide-induced BV2 microglia cells. *Food and Chemical Toxicology,* 49, 1745-1752.

Patridge, E., Gareiss, P., Kinch, M. S., & Hoyer, D. (2016). An analysis of FDA-approved drugs: natural products and their derivatives. *Drug Discovery Today*, 21, 204-207.

Pang, X., Lin, X., Tian, Y., Liang, R., Wang, J., Yang, B., Zhou, X., Kaliyaperumal, K., Luo, X., Tu, Z., others, (2018). Three new polyketides from the marine spongederived fungus *Trichoderma* sp. SCSIO41004. *Natural Product Research*, 32, 105-111.

Poveda, J., Hermosa, R., Monte E., & Nicolás, C. (2019). *Trichoderma harzianum* favours the access of arbuscular mycorrhizal fungi to non-host Brassicaceae roots and increases plant productivity. *Scientific Reports*, 9, 11650.

Pye, C. R., Bertinb, M. J., Lokeya, R. S., Gerwickb, W. H., & Linington, R. G. (2017). Retrospective analysis of natural products provides insights for future discovery trends. *PNAS*, 144, 5601-5606.

Qi, H., Zhang, Q., Zhao, T., Chen, R., Zhang, H., Niu, X., & Li, Z. (2005). Antioxidant activity of different sulfate content derivatives of polysaccharide extracted from Ulva pertusa (Chlorophyta) invitro. *International Journal of Biological Macromolecules,* 37, 195-199.

Qian, Z. J., Jung, W. K., Byun, H. G., & Kim, S. K. (2008). Protective effect of an antioxidative peptide purified from gastrointestinal digests of oyster, Crassostrea gigas against free radical induced DNA damage. *Bioresource Technology*, 99, 3365– 3819.

Qu, Y., Safonova, O., & Luca, V. (2018). Completion of the canonical pathway for assembly of anticancer drugs vincristine/vinblastine in *Catharanthus roseus*. *The Plant Journal*, 97, 257–266.

Ramezani, A., Haddad, R., Sedaghati, B., & Jafari, D. (2018). Effects of fungal extracts on vinblastine and vincristine production and their biosynthesis pathway genes in *Catharanthus roseus*. *South African Journal of Botany*, 199, 163-171.

Rani, N. Z. A., Husain, K., & Kumolosasi, E. (2018). Moringa genus: A review of phytochemistry and pharmacology. *Frontiers in Pharmacology*, 9, 1-26.

Richards, T. A., Leonard, G., & Wideman, J. G. (2017). What defines the "kingdom" fungi? *The Fungal Kingdom*, 57-77.

Richards, T. A., & Talbot, N. J. (2013). Horizontal gene transfer in osmotrophs: Playing with public goods. *Nature Reviews Microbiology*, 11, 720-727.

Rico-Ramírez, A. M., Gonçalves, A. P., & Glass, N. L. (2022). Fungal cell death: The beginning of the end. *Fungal Genetics and Biology*, 159, 10361.

Rocha-Santos, T., & Duarte, A. C. (2014). Chapter 1 - Introduction to the Analysis of Bioactive Compounds in Marine Samples. *Comprehensive Analytical Chemistry*, 65, 1-13.

Rosa, L.A., Moreno-Escamilla, J. O., Rodrigo-García, J., & Alvarez-Parrilla, E. (2019). Chapter 12 - Phenolic Compounds. *Postharvest Physiology and Biochemistry of Fruits and Vegetables*, 253-271.

Sachindra, N. M., Sato, E., Maeda, H., Hosokawa, M., Niwano, Y., Kohno, M., & Miyashita, K. (2007). Radical scavenging and singlet oxygen quenching activity of marine carotenoid fucoxanthin and its metabolites. *Journal of Agricultural and Food Chemistry,* 55, 8516-8522.

Sarker, S. D., & Nahar, L. (2012). An introduction to natural products isolation. *Methods in Molecular Biology*, 864, 1-25.

Sekhon, A., Wang, J. Y. F., Tan, J. C. H., Holland, S. P., & Yeung, S. N. (2020). Limbal stem cell deficiency secondary to systemic paclitaxel (Taxol) for breast cancer: a case report. *BMC Ophthalmology*, 20, 2-4.

Selvakumar, V., & Panneerselvam, A. (2018). Bioactive compounds from endophytic fungi. In P. Gehlot, & J. Singh (eds.), *Fungi and their role in sustainable development: Current perspectives* (pp. 699-717). Singapore: Springer.

Sharma, S., Kori, S., & Parmar, A. (2015). Surfactant mediated extraction of total phenolic contents (TPC) and antioxidants from fruits juices. *Food Chemistry,* 185, 284-288.

Shibata, T., Ishimaru, K., Kawaguchi, S., Yoshikawa, H., & Hama, Y. (2008). Antioxidant activities of phlorotannins isolated from Japanese Laminariaceae. *Journal of Applied Phycology*, 20, 705-711.

Siddharth, S., & Vittal, R. R. (2018). Evaluation of antimicrobial, enzyme inhibitory, antioxidant and cytotoxic activities of partially purified volatile metabolites of marine *Streptomyces* sp. S2A. *Microorganisms*, 6, 72.

Silva, S., Ferreira, M., Oliveira, A. S., Magalhães, C., Sousa, M. E., Pinto, M., Lobo, J. M. S., & Almeida, I. F. (2019). Evolution of the use of antioxidants in anti-ageing cosmetics. *International Journal of Cosmetic Science*, 41, 378-386.

Singh, B. (2007). Psyllium as therapeutic and drug delivery agent. *International Journal of Pharmaceutics*, 334, 1-14.

Singh, B., Mal, G., Gautam, S. K., Mukesh, M. (2019). Nutraceuticals from Bioengineered Microorganisms. *Advances in Animal Biotechnology*, 59-69.

Sithrangaboopathy, N., & Kathiresan, K. (2010). Anticancer drugs from marine flora: an overview. *Journal of Oncology,* 1-18.

Sokoła-Wysoczańska, E., Wysoczański, T., Wagner, J., Czyż, K., Bodkowski, R., Lochyński, S., & Patkowska-Sokoła, B. (2018). Polyunsaturated fatty acids and their potential therapeutic role in cardiovascular system disorders - A review. *Nutrients*, 10, 1561.

Somani, S. J., Modi, K. P., Majumdar, A. S., & Sadarani, B. N. (2015). Phytochemicals and their potential usefulness in inflammatory bowel disease. *Phytotherapy Research*, 29, 339-350.

Sonani, R. R., Singh, N. K., Kumar, J., Thakar, D., & Madamwar, D. (2014). Concurrent purification and antioxidant activity of phycobiliproteins from Lyngbya sp. A09DM: An antioxidant and anti-aging potential of phycoerythrin in Caenorhabditiselegans. *Process Biochemistry*, 49, 1757-1766.

Song, T., Chen, M., Chai, W., Zhang, Z., & Lian, X.-Y. (2018). New bioactive pyrrospirones C- I from a marine-derived fungus *Penicillium sp*. ZZ380. *Tetrahedron*, 74, 884-891.

Sosa-Hernández, J. E., Escobedo-Avellaneda, Z., Iqbal, H. M. N., & Welti-Chanes, J. (2018). State-of-the-art extraction methodologies for bioactive compounds from algal biome to meet bio-economy challenges and opportunities. *Molecules*, 23, 2953.

Spatafora, J. W., Aime, M. C., Grigoriev, I. V., Martin, F., Stajich, J. E., & Blackwell, M. (2017). The fungal tree of life: From molecular systematics to genome-scale phylogenies. *The Fungal Kingdom,* 1-34.

Stage, T. B., Bergmann, T. K., & Kroetz, D. L. (2018). Clinical pharmacokinetics of paclitaxel monotherapy: An updated literature review. *Clinical Pharmacokinetics*, 57, 7-19.

Stincone, P., & Brandelli, A. (2020). Marine bacteria as source of antimicrobial compounds. *Critical Reviews in Biotechnology*, 40, 306-319.

Strobel, G., Hess, W. M., Li, J. Y., Ford, E., Sears, J., Sidhu, R. S., & Summerell, B. (1997). *Pestalotiopsis guepinii*, a taxol producing endophyte of the Wollemi Pine, Wollemia nobilis. *Australian Journal of Botany*, 45, 1073-1082.

Su, L.-J., Zhang, J.-H., Gomez, H., Murugan, R., Hong, X., Xu, D., Jiang, F, & Peng, Z.-Y. (2019). Reactive oxygen species-induced lipid peroxidation in apoptosis, autophagy, and ferroptosis. *Oxidative Medicine and Cellular Longevity*, 2019, 1-13

Suleria, H. A. R., Gobe, G., Masci, P., & Osborne, S. A. (2016). Marine bioactive compounds and health promoting perspectives; innovation pathways for drug discovery. *Trends in Food Science & Technology,* 50, 44-55.

Sun, L., Wang, C., Shi, Q., & Ma, C. (2009). Preparation of different molecular weight polysaccharides from Porphyridium cruentum and their antioxidant activities. *International Journal of Biological Macromolecules,* 45, 42-47.

Sung, A. A., Gromek, S. M., & Balunas, M. J. (2017). Upregulation and identification of antibiotic activity of a marine derived *Streptomyces* sp. via co-cultures with human pathogens. *Marine Drugs*, 15, 250.

Taher, Z. M., Agouillal, F., Lim J. R., Marof, A. Q. Dailin, D. J., Nurjayadi, M., Razif, E. N. M., Gomaa, S., Enshasy, H. A. (2019). Anticancer molecules from Catharanthus roseus. *Indonesian Journal of Pharmacy*, 30, 147–156.

Tan, S. Y., Tatsumura, Y. (2015). Alexander Fleming (1881–1955): Discoverer of penicillin. *Singapore Medical Journal*, 56, 366-367.

Taylor, D. L., Hollingsworth, T. N., McFarland, J. W., Lennon, N. J., Nusbaum, C., & Ruess, R. W. (2014). A first comprehensive census of fungi in soil reveals both hyperdiversity and fine-scale niche partitioning. *Ecological Monographs,* 84, 3-20.

Tungmunnithum, D., Thongboonyou, A., Pholboon, A., & Yangsabai, A. (2018). Flavonoids and other phenolic compounds from medicinal plants for pharmaceutical and medical aspects: An overview. *Medicines*, 5, 93.

VanderMolen, K. M., Raja, H. A., El-Elimat, T. & Oberlies, N. H. (2013). Evaluation of culture media for the production of secondary metabolites in a natural products screening program. *AMB Express*, 3, 1-7.

Vicent, O., & Boscaiu, M. (2018). Flavonoids: Antioxidant compounds for plant defence... and for a healthy human diet. *Notulae Botanicae Horti Agrobotani Cluj-Napoca*, 46, 14-21.

Wali, A., F., Majid, S., Rasool, S., Shehada, S. B., Abdulkareem, S. K., Firdous, A., Beigh, S., Shakeel, S., Mushtaq, S., Akbar, I., Madhkali, H., & Rehman, M. U. (2019). Natural products against cancer: Review on phytochemicals from marine sources in preventing cancer. *Saudi Pharmaceutical Journal*, 27, 767-777.

Walker, G. M., & White, N. A. (2017). Introduction to fungal physiology. In K. Kavanagh (Ed.), *Fungi: Biology and Applications* (pp. 1-35), Wiley Online Library, Inc.

Wall, S. (2020). Blastomycosis: review of a dimorphic fungus in central Pennsylvania. *Clinical Microbiology Newsletter*, 42, 157-161.

Walther, B., & Sieber, R. (2011). Bioactive proteins and peptides in foods. *International Journal for Vitamin and Nutrition Research*, 81, 181-191.

Wang, W., Liao, Y., Chen, R., Hou, Y., Ke, W., Zhang, B., Gao, M., Shao, Z., Chen, J., & Li, F. (2018). Chlorinated azaphilone pigments with antimicrobial and cytotoxic activities isolated from the deep sea derived fungus *Chaetomium* sp. NA-S01-R1. *Marine Drugs*, 16, 61.

Wang, Y.-T., Xue, Y.-R., & Liu, C.-H. (2015). A brief review of bioactive metabolites derived from deep-sea fungi. *Marine drugs*, 13, 4594-4616.

Watkinson, S. C., Boddy, L., & Money, N. (2015). *The fungi*. s.l.: Academic Press.

Wells, R., Truong, F., Adal, A. M., Sarker, L. S., Mahmoud, S. S. (2018). *Lavandula* essential oils: A current review of applications in medicinal, food, and cosmetic industries of lavender. *Natural Product Communications*, 13, 1403-1417.

Wen, L., Chen, Y., Zhang, L., Yu, H., Xu, Z., You, H., & Cheng, Y. (2016). Rice protein hydrolysates (RPHs) inhibit the LPS-stimulated inflammatory response and phagocytosis in RAW 264.7 macrophages by regulating the NF-kB signaling pathway. *RSC Advances*, 6, 71295-71304.

Willey, J. M., Sherwood, L., & Woolverton, C. J. (2011). *Prescott's microbiology* (Vol. 7). New York: McGraw-Hill.

Xu, D.-P., Li, Y., Meng, X., Zhou, T., Zhou, Y., Zheng, J., Zhang, J.-J., & Li, H.-B. (2017). Natural Antioxidants in Foods and Medicinal Plants: Extraction, Assessment and Resources. International Journal of Molecular Sciences, 18, 96.

Yan, L.-G., He, L., & Xi, J. (2017). High intensity pulsed electric field as an innovative technique for extraction of bioactive compounds: A review. *Critical Reviews in Food Science and Nutrition,* 57, 2877-2888.

Yang, L., Wen, K. S., Ruan, X., Zhao, Y. X., Wei, F., & Wang, Q. (2018). Response of plant secondary metabolites to environmental factors. *Molecules*, 23, 762.

Young, A. J., & Lowe, G. L. (2018). Carotenoids - Antioxidant Properties. *Antioxidants*, 7, 28.

Zhang, D., Shu, C., Lian, X., & Zhang, Z. (2018a). New antibacterial bagremycins F and G from the marine-derived *Streptomyces* sp. ZZ745. *Marine drugs*, 16, 330.

Zhang, J., & Elser, J. J. (2017). Carbon:Nitrogen:Phosphorus stoichiometry in fungi: A meta-analysis. *Frontiers in Microbiology*, 8, 1281.

Zhang, Q.-W., Lin, L.-G., & Ye, W.-C. (2018b). Techniques for extraction and isolation of natural products: A comprehensive review. *Chinese Medicine,* 13, 1-26.

Zhang, Y., Gan, R., Li, S., Zhou, Y., Li, A., Xu, D., & Li, H.-B. (2015). Antioxidant phytochemicals for the pevention and treatment of chronic diseases. *Molecules*, 20, 21138-21156.

Zhang, Y.-W., Kong, X.-Y. Wang, J.-H., & Du, G.-H. (2018c). Vinblastine and Vincristine. Natural Small Molecule Drugs from Plants, 551-557.

Zhao, C.-L., Cui, B.-K., Song, J., & Dai, Y.-C. (2015). *Fragiliporiaceae*, a new family of Polyporales (Basidiomycota). *Fungal Diversity*, 70, 115-126.

Zhao, D.-L., Wang, D., Tian, X.-Y., Cao, F., Li, Y.-Q., & Zhang, C.-S. (2018). Anti-phytopathogenic and cytotoxic activities of crude extracts and secondary metabolites of marine-derived fungi. *Marine Drugs*, 16, 36.

Zhao, X., Wang, J. F., & Xue, C. H. (2011). The inhibitory effects of fucoidans from laminaria japonica on oxidation of human low-density lipoproteins. *Advanced Materials Research,* 2067-2071.

Chapter 3

DATA MINING FOR THE ASSESSMENT OF THE CURRENT LANDSCAPE OF FUNGAL COMPOUNDS RESEARCH

3.1 Introduction

Data can be defined as information, including statistics or measurements, used as a basis for reasoning, discussion, or calculation and from which knowledge may be generated. Much of the information existing in different sources is currently preserved in electronic documents which contains semi-structured or even unstructured data. Discovering patterns and trends from large volumes of data is a significant challenge, being the main objective of data mining finding out correctly the unknown trends and patterns from databases (Krassmann et al., 2017). The first step in any data mining research project is to collect relevant data for analysis. Data can be extracted from different types of formats, for example, simple documents, webpages, user comments, reviews, journals, books. When parts of documents are collected and then analysed, these are called document features (Padhy et al., 2012). This suggests that a part of the document could be representative of the entire document, since the use of all the information available in each document tends to generate ambiguities and it becomes unfeasible due the large quantities of data contained in text files (Feldman & Sanger, 2006). Thus, data pre-processing is essential before applying any other methodology. Many approaches, such as clustering, classification, and decision trees are implicated in data mining (Allahyari et al., 2017). Many data or textual information is usually stored electronically, either on personal computers or on a web server. Due to the increasing growth in hardware storage devices, any computer or laptop can store a large volume of data. Generating new information can be simple but discovering relevant information from an enormous quantity of data is already a challenge (Jo, 2019; Salloum et al., 2018).

A study by Salloum et al. (2017) stated that the data mining methodology has become one of the trendy fields that has been integrated in several research areas, like computational linguistics, information recovery and text mining (Salloum et al., 2017). Data mining is focused on discovering patterns from large databases, while text mining focuses on textual information (Gupta & Lehal, 2009; Kobayashi et al., 2018; Kroeze et al., 2003). Information retrieval methodologies, such as text indexing techniques, have been developed to deal with unstructured data. In standard research, it is assumed that a user is mainly searching for familiar terms, which have already been used or written by someone else. The potential problem is that search results may not be relevant to the user requirements. One solution is to use text mining to discover relevant information, which is not clearly indicated or written so far (Aggarwal & Zhai, 2013; Justicia de la Torre et al., 2018). The text mining procedure starts with collecting documents through different resources. A specific document is retrieved through the text mining tool and by verifying its format and character sets. Text mining aims to detect relevant information that is not commonly recognized by automatically extracting it from different and several text-based sources (Jiang, 2013; Nenkova & McKeown, 2012). Text clustering is an example based on the cluster hypothesis that intends, that relevant documents should have more similarities to each other than non-relevant ones (Huang, 2008). In a cluster, similar terms or patterns are grouped extracted from various documents. The clustering methodology is a valuable and reliable method that is generally employed to analyse large volumes of data. It has been proven

that text clustering is one of the most effective tools used for analysing text terms. Furthermore, it simplifies the topic analysis method in which named entities with similar occurrences are grouped together, followed by submission them to the clustering process in such a way that frequent item is placed in sets (Talib et al., 2016; Wagstaf et al., 2001). Different techniques of clustering are hierarchical, distribution, density, centroid, and k-mean (Narayana & Kumar, 2015).

In a document, the “abstract” is an example of unstructured text field, while examples of structured data in a document are, for example, Digital Object Identified (DOI), author’s name, publication date, journal name, title, category and can be handled through data mining tools (Gaikwad et al., 2014). Overall, structured data (Figure 3.1) frequently refers to sources of information that rely on some processes of storing or delivering data in a structured manner. Data is well-defined, both in terms of structure and form, and is completely categorized and ready for human or machine language analysis since it is easy and quite quick to search and manipulative this type of data. As a matter of fact, some examples of current systems that do provide this specific type of representation are relational databases, developed application programming interfaces, and any Customer Relationship Management (CRM) software available, which is a tool designed to help project organization by providing a complete picture of all customer interactions (McLachlan & Krishnan, 2007; Mukhopadhyay, 2019).

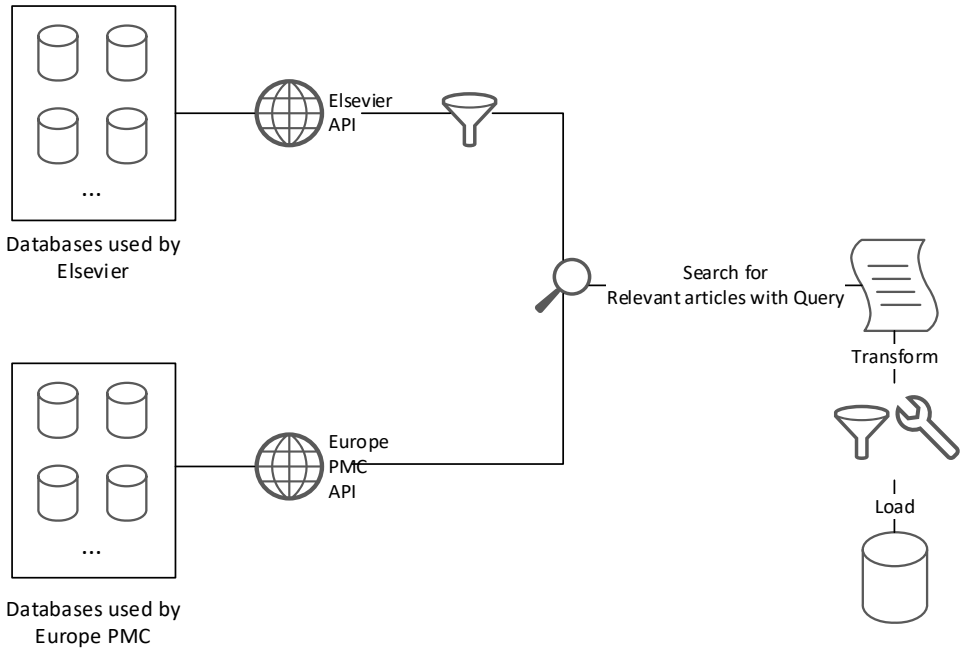

Figure 3.1. System design of the structured data collection process.
(Cylinder – database/warehouse; Sphere – Service; Funnels – Filters; Magnifier – Search; Document – Data; Screwdriver – Data transformation).

On the other hand, unstructured data (Figure 3.2) refers to a different type of data, namely to one that any individual interacts the most these days, whether it is text from local news, social media, videos, and audio, or from any other type of representation that one cannot see any structure around it. Moreover, unstructured data like articles, webpages, full-text documents, comments, emails, and books are a qualitative type of data, considering that it cannot be easily processed or analysed by conventional methods, especially due to the lack of a good and pre-defined model, but can be handled through text mining and, usually, this information can be maintained in a natural form known as text (Miner et al., 2012; Mukhopadhyay, 2019; Sulova & Nacheva, 2017). This text, then pass through a text analysis stage. Text analysis includes semantic evaluation aimed to obtain high-quality information through text. Different text analysis methods are available and can be used based on the organization's objective. The results can be stored in an administration information system that provides a large quantity of relevant information for the user of that system (Gupta et al., 2005).

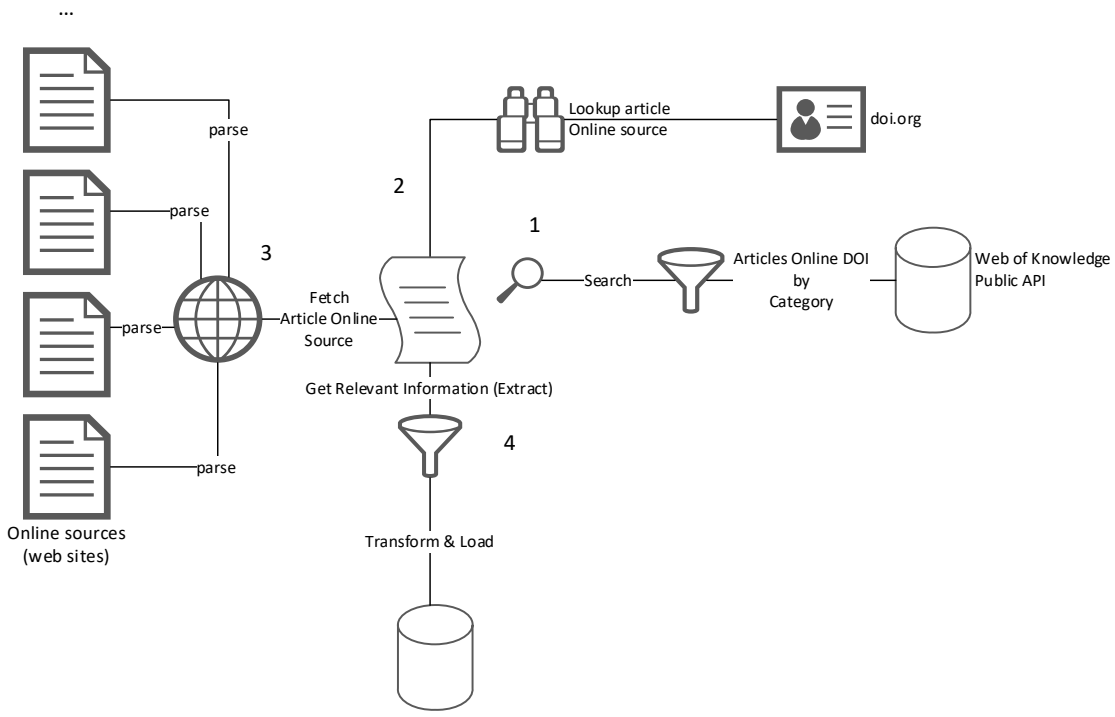

Figure 3.2. System design regarding the unstructured data collection process.
(Cylinder – database/warehouse; Sphere – Service; Funnels – Filters; Magnifier – Search; Binoculars – Search and DOI validation).

To extract significant information, knowledge, or patterns from several and different sources that are in unstructured form, text mining technique can be employed. Graphics, tables, or structured data illustration are some examples of how to structure the extracted information. There are several research areas, techniques, and models involved in different research domains. The hottest topics of the research domains are the principal focus of many research papers. Search results of a specific

domain may influence other search domains, since some these search domains may have similar topics. These research topics always discuss an area of research so promising that it is worth studying (Gupta et al., 2005; Hassani et al., 2016). Therefore, the main goal of the present chapter is to assess the current landscape of fungi compounds research, namely by using a data mining tool. In other words, the objective is to find out if there are several investigations/studies within this specific field, or if there is a lack of information and data regarding fungi compounds. Therefore, the present chapter intends to either support existing studies or to add pertinent and relevant data to the field of expertise, aiming to contribute to the literature pertaining fungi compounds, especially bioactive compounds from terrestrial and marine sources. The concept of text mining is defined, and the data mining processing framework is detailed.

3.1.1 Data Mining

Data mining, also referred to as knowledge discovery in databases, is the process of analysing large information repositories and of discovering implicit, still potentially useful, information (Han et al., 2011). According to Sumathi and Sivanandam (2006), data mining can uncover hidden relationships and reveal unknown trends and patterns, namely by digging into large amounts of data and information.

The most frequent and common data mining process (Figure 3.3) refers to an interactive sequence of steps, which usually starts with the integration of raw data from several data sources and from different formats. Afterwards, these raw data are cleansed, aiming to remove both the noise and the inconsistent and duplicated data. Furthermore, these cleansed data are adequately transformed into appropriated formats, which are understood by other data mining tools, with multiple aggregation and filtration techniques being applied to the data to extract summarized data (Dang & Ahmad, 2015; Fan et al., 2006). The resulting transformed data is ultimately analysed to identify the genuinely interesting patterns, with the knowledge being visualized and understood by the user (Han et al., 2011).

At last, it is also important to mention that data mining techniques are applied in a distinct range of domains, where large amounts of data are indeed available for the identification of either hidden or unknown information. Therefore, and in this specific context, the data mining techniques that are used in the “WWW” are designated as web mining, the data mining techniques that are used in the text are called text mining, and the data mining techniques that are used in libraries are designated by bibliomining (Girija & Srivatsa, 2006).

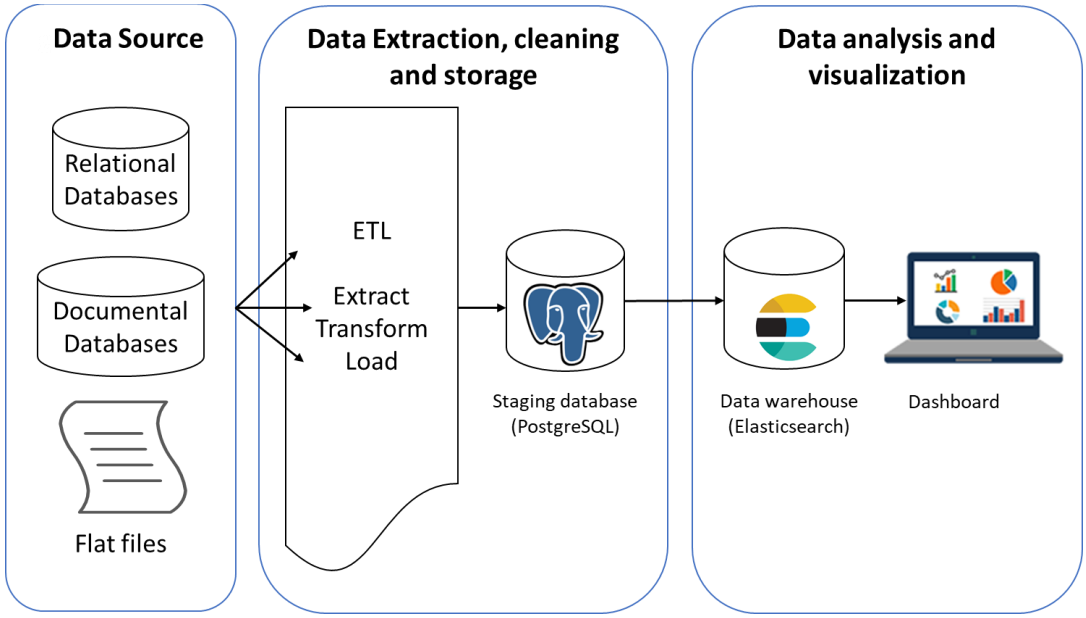

Figure 3.3. Schematic data mining process. Adapted from Siguenza-Guzman et al., 2015.

A properly Extract, Transform and Load (ETL) designed system extracts data from one or more sources, enforces data consistency and quality standards, conforms data so that separate sources can be used together, and lastly delivers data into a format for use. The first step of an ETL process involves extracting the data from the source system. In many cases, this represents the most important characteristic of ETL, once the data extraction correctly sets the stage for the success of subsequent processes. Most data-warehousing's combine data from different source systems (Navathe & Elmasri, 2000; Sullivan, 2001). Each single system may also use a distinct data organization and/or format. The component of extraction retrieves data from several sources with different formats, converting it into a single format to be suitable for the next action of the process, the transformation phase (Feldman & Sanger, 2007). Moreover, the data extraction process is performed in two different phases: full extraction, when the data is extracted for the first time; and incremental extraction, when modified or new data are retrieved from the sources (Kimball & Caserta, 2004).

In terms of the data transformation, on the other hand, a series of rules or functions are applied to the extracted data to prepare it for loading into the end target. An important function of transformation is the data cleansing, which aims to reformat, and integrate only "proper" data, to better suit the format of the model of a target data warehouse. A data warehouse is generally developed by a multidimensional data structure, named "data cube", in which each dimension corresponds to an attribute or a set of attributes in the structure, and each cell of this structure stores the value of some aggregate measure like count or sum. A data cube provides a multidimensional view of data and allows the precomputation and quick access of summarized data (Farooqui & Mehra, 2018). The main goals of the transformation phase aim to prevent the transformation of "dirty data", namely by cleaning the data through the identification or solving/removing the existing

problems in the data for further integration (Barateiro & Galhardas, 2005), as well as make the data conform to the target format, especially by applying a set of transformation rules provided by the data warehouse designers (Kimball & Caserta, 2004).

The challenge is when different systems interact in the relevant systems' interfacing and communicating. Character sets that may be available in one system may not be so in others. In other cases, one or more of the transformation types may be required to meet the business and technical needs of the server or data warehouse, applying any form of data validation. Failed validation may result in a full rejection of the data, partial rejection, or no rejection at all, and thus none, some, or all of the data is handed over to the next step depending on the rule design and exception handling. Many of transformations may result in exceptions, for example, when a code translation parses an unknown code in the extracted data (Reddy & Venkatadri, 2011).

Lastly, the component of loading writes the extracted and transformed data from the staging area to the target data warehouse. This last phase varies according to the organizational requirements. Indeed, some data warehouses can overwrite pre-existing data with new data on a daily, weekly, or monthly basis, while others can keep the history of data by adding new data at regular intervals. The load component is frequently implemented by using loading jobs that either fully or incrementally transform data to the data warehouse (Feldman & Sanger, 2007; Mathen, 2010).

3.1.2 Text Mining Processing Framework

The implementation of the data mining tool comprises three different steps: Step 1 – pre-processing stage; Step 2 – text mining operations stage; and Step 3 – postprocessing stage. In turn, all these steps are based on different processes, which must be conducted to successfully achieve the data mining tool’s goal. Hence, during the pre-processing stage (Step 1) it is important to collect the text (text data), to treat the obtained text data, and to transform such data. This transformation of the data occurs in Step 2, which refers to the several text mining operations, namely to word clouds, keywords extraction, affiliations extraction, and word frequency count. Lastly, Step 3, referring to the postprocessing stage, is based on the analysis of the obtained data, as well as on its subsequent interpretation and discussion (Figure 3.4).

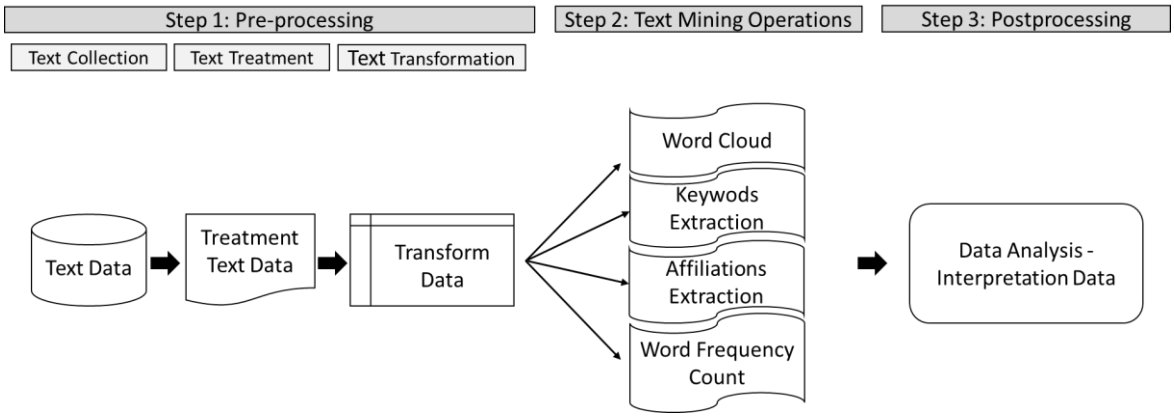

Figure 3.4. Steps and operations of the process of implementation of the data mining tool. Adapted from Kobayashi et al., 2018.

3.1.2.1 Pre-processing

The pre-processing stage of the data mining method is directly associated with three specific processes: text collection or extraction, text treatment, and text transformation (Figure 3.4). This is an extremely important stage of the data mining methodology, especially after collecting the necessary documents. At this stage, the collected documents will be converted into a format suitable for the subsequent core mining process. It includes all the relevant procedures to gradually improve the structure of the documents until applying a representation of the features that are the input of the core mining algorithms (Feldman & Sanger, 2006; Tang et al., 2014). Natural Language Processing (NLP) technique is used to allow these data manipulations (Webster & Kit, 1992). NLP is described as the automatic processing of human language, which requests a computer through an algorithm to convert free-form text into structured understandable data. This algorithm works according to Artificial Intelligence (AI) and Machine Learning techniques (ML) (Ananiadou et al., 2006).

This process generally combines linguistic concepts, as part of speech, such as nouns, adjectives, verbs, among others, as well as grammatical structure, considering grammatical ambiguities. To perform such a sophisticated task, NLP makes use of various representations of knowledge, including lexicon of words, their meaning and grammatical rules. The described method can be combined with other resources, like ontology of entities and actions, or a glossary of synonyms or abbreviations (Ananiadou et al., 2006; Sun et al., 2017).

Pre-processing starts with the tokenization task, where each phrase or sentence is divided into words or tokens. In this step, the tokenization process converts the phrases at the word level, the idea behind keeping more than one word is to recover knowledge from the grouping of two or more words that can lead to a different perception instead of the separated words, although the use of more than one word per token increases the data dimension (Sun et al., 2017). The next step is

concerned with the treatment process of all the characters where each document title is created in a lower case, followed by the stop words. Lastly, the final task of pre-processing is the representation of data, where the text is transformed to be 'understandable' for the next algorithms, which are initialized by building a text representation model (Boulis & Ostendorf, 2005). The key methods that are used in this pre-processing stage are the following:

1) **Extraction** is used to tokenize the text's content into individual words (Kononenko & Kukar, 2007).
2) **Stop words elimination**, where stop words are a division of natural language. They should be removed from the text since they make it look quite heavier and less important to further analysis. The most frequent words in texts are usually prepositions, articles, and pronouns, which do not give any meaning to the text itself. Therefore, these specific words are considered stop words, being removed from the text because they are not measured as keywords in text mining applications (Sarica & Luo, 2021).
3) **Stemming** is used to identify the root/stem of a word, aiming to remove several suffixes, reduce the number of words, accurately match stems, and save memory space and time (Jivani, 2011; Ramasubramanian & Ramya, 2013; Sharma, 2012).

3.1.2.2 Text Mining Operations

Regarding text mining operations, which refer to the second step of the data mining tool, these comprehend four distinct processes, all of them analysing different aspects of the collected documents/texts. In more detail, the four processes are the following: word clouds, keyword extraction, affiliations extraction, and word frequency count.

a) Word clouds

Word clouds are exceedingly popular for both text and website analysis. Word clouds, also known as text clouds or tag clouds are graphical representations of word frequency that give greater prominence to words that appear more frequently in a source textual data. A word cloud is a collection, or cluster, of words represented in different sizes. The bigger and bolder the word appears, the more often it is mentioned in a certain text and the more important it is (DePaolo & Wilkinson, 2014). Word cloud is among the most frequently used technique for presenting text data in a graphical style, making it useful for analysing different forms of text data, such as essays and written opinions or short answers for a particular survey or questionnaire (Sinclair & Cardew-Hall, 2008; Viegas et al., 2007). In common text analyses, the words of interest are placed in a rectangular form, with the font size and the colour of words representing frequency and usefulness, respectively. However, it is possible to choose other options, such as font style and layout, aiming to enhance the

visual appeal of the word clouds. Furthermore, in typical word clouds, the tags from websites or words from documents are packed into a rectangular shape in which the font size indicates the tag's popularity or word frequency, while the font colour indicates useful information that must be retained (Cui et al., 2010). The larger the text size, the more frequent the word is in the given text/website. Finally, the main goal of word clouds is to summarize the most important concepts in a visual presentation, hence helping in the synthesis of the big ideas that are inserted in the document's content (Hamm, 2011). For example, the word cloud helps to find significance between given text and essential and required information. Nevertheless, the technique has some disadvantages as well. One of the major obstacles is not considering the linguistic knowledge about the words and their respective connection with the subject in question or under study while providing a purely statistical resume of the isolated words. As a result, in most systems, the word clouds are frequently applied statistically to summarize text, offering a visual correlation of data. It is understood that this technique can be one of the most influential visualization paradigms for most analysis conditions (Salloum et al, 2018).

b) Keyword extraction

Keyword extraction is based on the automatic identification of a set of concepts that best describe the subject of a specific document. The concepts that represent the most relevant information within the document can be defined by different terminologies, such as key segments, key phrases, keywords, or key terms. However, all these terminologies have the same function, which refers to the characterization of the topics under discussion in a document. The extraction of a small set of units, which can be composed of one or more concepts, from a document is an important process in text mining, NLP, and Information Retrieval (IR) (Berry & Kogan, 2010; Abilhoa & Castro, 2014).

Overall, keywords are used to enable queries within IR systems, considering that they are quite easy to define, remember, revise, and share. Moreover, keywords are independent of any corpus, being able to be applied across multiple IR systems and corpora and can improve the functionality of IR systems in general (Berry & Kogan, 2010). Therefore, when relevant keywords are extracted, they can be used to build an automatic index for a document collection, or even for document representation in both classification and categorization processes (Allahyari et al. 2017, Baba & Kumar, 2016). Despite often working on single documents, keyword extraction is also used in more complex tasks, such as keyword extraction for entire collections or an automatic web summarization (Fan et al., 2006). Moreover, with the appearance of big data, the construction of an effective model for text representation is vital and demanding simultaneously (Beliga et al., 2015).

c) Affiliation's extraction

Affiliations refer to a list of parsed associations of the authors to the document, in the order that is given in the document. A single affiliation often contains 1) raw text of the affiliation; 2) organization name; 3) address, and 4) country. It is also very frequent for both the authors and affiliations to be extracted, with the relationship between them being hence determined.

d) Word frequency count

Lastly, the word frequency count, also designated as the term frequency method, where the terms are weighted to indicate their importance for document representation, assuming that term relevance is proportional to the number of documents that contain the respective term. The term frequency method then assesses each term's importance by the number of times it occurs in a specific document. Still, this method is limited to the term occurrence within a single document, despising the term occurrence among a collection of documents. In this case, the inverse document frequency is used, aiming to measure the importance of a term, which is inversely proportional to the number of documents that contain such term (Sebastiani, 2002).

3.1.2.3 Postprocessing

In addition to pre-processing, as well as to the text mining operations that were previously mentioned, the postprocessing step is extremely important, comprehending both the data analysis and the interpretation of data. It is important to establish that the text mining operations produce long listings with plenty of results, with some information not being useful for the process. Therefore, the postprocessing step is vital to select the most adequate information, as well as to provide important results and an adequate interpretation of data. In more detail, it is quite essential to:

1) Identify the most relevant information from the data sources, depending on the aims of the analysis.
2) Find the best way to present the selected results to the user, considering that everything must be directly and objectively understandable, even to those who do not know the technical details of the text mining operations that were conducted (Gilbert et al., 2008).

According to the same authors, in this specific step of the data mining process, it is possible, as well as useful, to use several tools to analyse and present results, contributing to their interpretation. In their study, Gilbert et al. (2008) suggest the usage of several techniques, such as decision trees, statistical modelling, neural nets, and graphics, all corresponding to visualization techniques that assist both the user and the reader to understand the obtained results and their contributions objectively and easily.

3.2 Materials and Methods

3.2.1 Extract, Transform, and Load (ETL)

The implementation of the data mining tool comprised the use of several technologies, which assisted during the entire process, helping to conduct all three steps (pre-processing, text mining operations, and postprocessing) already addressed in the previous section. Hence, the technologies that were used to implement the data mining tool were the following ones:

1) C# /.Net Core (v 2.2).
2) Entity Framework Core (v 2.2).
3) PostgreSQL.
4) Dapper.
5) Elasticsearch.
6) Visual Studio / Azure data Studio.
7) React.
8) R Core Team.
9) RStudio.

The ETL method was the process implemented to extract data from different sources of interest (both structured and unstructured information), then transform the data into a common model (PostgreSQL), and finally load the data into a data warehouse (Elasticsearch). This method is a multi-stage process of the text mining processing framework conducted using the various programming software tools mentioned from 1) to 9), based on the time range between 1786 and 2020. The date 1786 was chosen, not because it was the date of the first publication containing the topics of interest, but rather, the first possible paper to be extracted to a database, that is, older papers are often found digitally in Portable Document Format (PDF), which is impossible to apply text mining. The analysed data was extracted in February of 2021.

To provide a deeper knowledge about the specificities of the existing research so far, as well as a more sense of its evolution throughout the last few years, the timeframe between 2010 and 2020 was also evaluated in more detail. Thus, the objective was not only to assess the same specificities that were previously assessed in the first-time range, but also to compare them, establishing a sense of evolution in terms of the conducted research within this field of expertise.

The ETL process was used in the present study to extract data using related to “fungi”, “fungus” or “fungal” keywords from papers title and abstract. At a first analysis, it was in our best interest to extract the information with the goal of obtaining summaries for the words contained in the title and abstract documents or to compute summaries for the abstract documents based on the words contained in them, using programming tools to research scientific papers in major online libraries,

such as Google scholar, Web of Science, Science Direct, Europe PMC, among others. However, the data extraction process itself seemed like a challenge. Nonetheless, after several challenges (detailed bellow in section 3.2.1.4, items 1 and 2), it was possible to find a suitable platform for the purpose of this work. Thus, using Europe PMC Collection to extract the data using C# language and the frameworks Entity Framework Core and Dapper, it was possible to analyse the words, clusters of words in documents, or to examine documents and determine the existing similarities between all of them. The Europe PMC Collection website provides subscription-based access to several databases that provide comprehensive citation data for many different academic subjects, so it was used as the primary database for the present work. Moreover, to obtain comparative information in order to validate our results, the “Advanced Search” tool available on the “Europe PMC” online website was used (https://europepmc.org, consultation carried out in February 2021) (Europe PMC, 2020, 2021).

To gather information from the scientific articles title and respective abstracts, a crawler (Entity Framework Core and Dapper) was used to fetch results from well-known sources, like Europe PMC, which would then iterate throughout the source’s item collection, and to fetch its relevant information/metadata avoiding documentation duplication, namely by querying DOI from the relevant platform DOI system (Manku et al., 2007), extracting data from C# language format, to then be stored in a database (PostgreSQL) for a transformation and later processing and sending using Dapper to a final data warehouse (Elasticsearch) (Figure 3.5) (Elasticsearch, 2021a).

Visual Studio, Azure data Studio, and RStudio are an integrated development environment (IDE), which is a software application that provides comprehensive facilities to computer programmers for software development.

Other tools that were used for drawing figures or schemes were Microsoft Excel®, Microsoft Visio® Professional 2021, and Nivo.

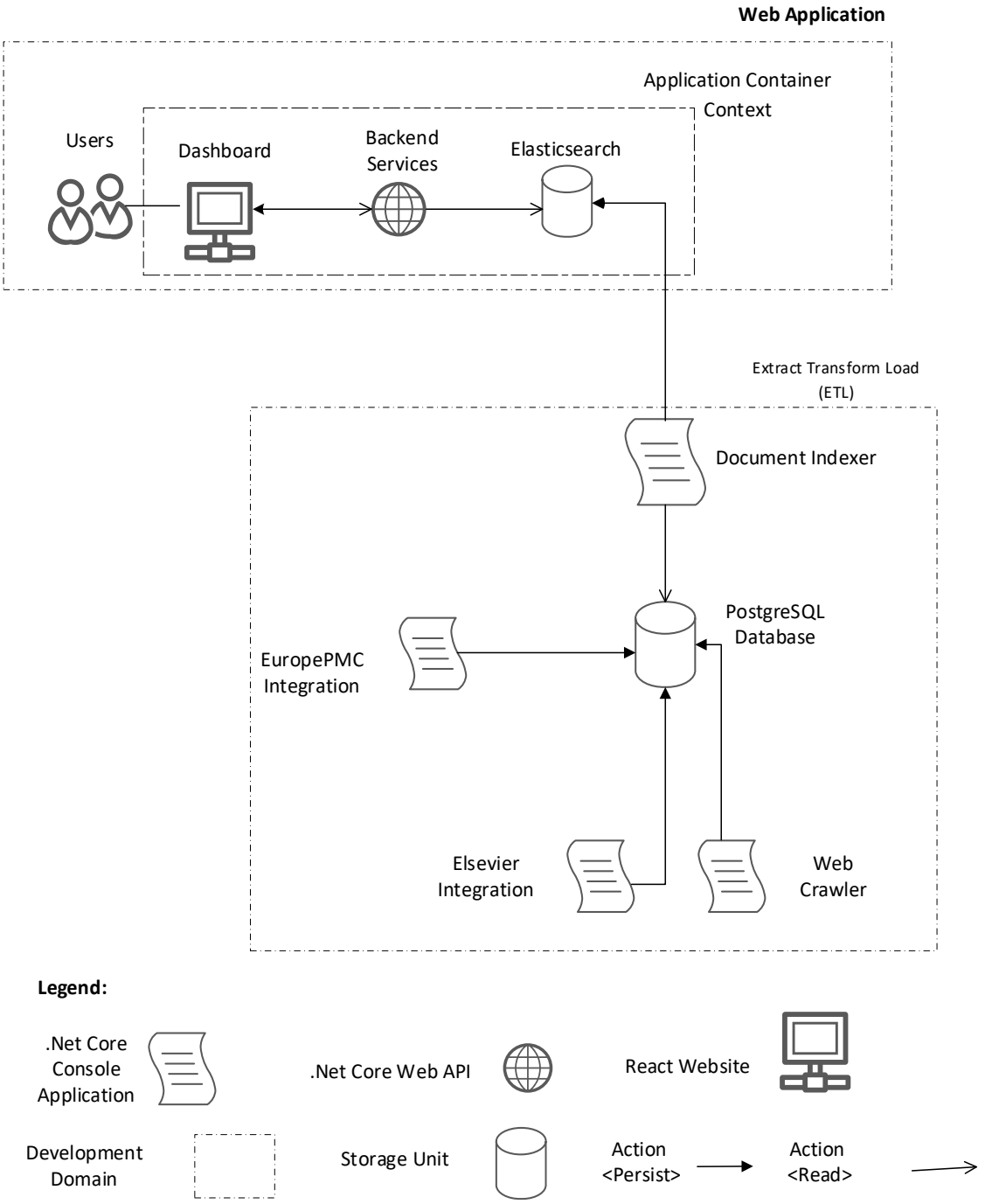

Figure 3.5. Schematic representation of the entire extraction, transformation, loading and web application process.

3.2.1.1 Structured and Unstructured Data Extraction and Transformation

Taking into consideration that the main goal of the present study is to find/search documents throughout specific queries related to the document's content and area of expertise (e.g., fungus,

marine/terrestrial fungi, and antibacterial/antioxidant activity), the main processing tasks were focused on these fields, establishing an evolutionary connection between these topics of interest over the years. To this end, the fields (features and affiliations) that were transformed/created were the following:

1) Feature extraction:
   a. Detection of the language used in the title and abstract of documents.
   b. Extraction of keywords used in the title and abstract of documents.
   c. Extraction of recognized entities name in the documents.
   d. Extraction of linked entities in the documents.

2) Affiliation:
   a. Country.
   b. City.
   c. Authors' name.
   d. Institution.
   e. Department.
   f. Journals' name.
   g. Documents' type.
   h. Documents' format.

Since these topics of interest have been indexed as of type **Text**, each term can be matched against any of the specified terms used in the search. Therefore, it was specified an analyser (Elasticsearch), which performs the tokenization of the Abstract's content, lowercase it, and remove any stop words (Elasticseach, 2021f). In terms of the document's titles, and since they are generally small and composed by simpler words, the goal was to merely tokenize, lowercase and remove stop words. Regarding the keywords and the affiliations, they were getting all that, in addition to an Edge N-gram (from 3 to 15 characters length), which is a constructing method of composition of parts of the one same word, aiming to make the process of searching and autocompletion easier, considering that it will break words into smaller subsets (already incorporated in the Elasticsearch analytics engine) (Tromp & Pechenizkiy, 2011). Figures 3.6 and 3.7 presents the code containing the queries to feature extraction and queries with multi-match field search (aggregations).

The 2010-2020 population estimated was based on Worldometer elaboration of the latest United Nations data.

```
GetDetectedLanguage(string document, string country = null);

GetKeywords(string document, string language = null, CancellationToken ct = default);

GetNamedEntitiesRecognized(string document, string language = null, CancellationToken
ct = default);

GetLinkedEntities(string document, string language = null, CancellationToken ct =
default);

GetNamedEntitiesRecognizedInBatch(IEnumerable<DocumentInput> documents,
CancellationToken ct = default);

GetLinkedEntitiesInBatch(IEnumerable<DocumentInput> documents, CancellationToken ct =
default);

GetKeywordsDocumentsInBatch(IEnumerable<DocumentInput> documents, CancellationToken ct
= default);
}
```

Figure 3.6. Code sample for extract the desired features (language, keywords, recognized entities name and linked entities).

```
public async Task<ISearchResponse<Document>> SearchDocuments(string search)
{
    Func<QueryContainerDescriptor<Document>, QueryContainer> qd = null;

    if (string.IsNullOrWhiteSpace(search))
    {
        qd = q => q.MatchAll();
    }
    else
    {
        qd = q => q
            .MultiMatch(mm => mm
                .Fields(f => f.Field(p => p.Title).Field(p => p.Abstract))
                .Query(search)
                .MinimumShouldMatch(1)
                .AutoGenerateSynonymsPhraseQuery(false)
                .Type(TextQueryType.CrossFields)
            );
    }

    return await _elasticClient.SearchAsync<Document>(sd => sd
        .Query(qd)
        .Aggregations(ag => ag
          .DateHistogram("hist_pubs_per_year", date => date
               .Field(p => p.Date)
               .CalendarInterval(DateInterval.Year)
               .MinimumDocumentCount(50)
               .Order(HistogramOrder.KeyAscending)
          )
          .Terms("topn_pub_types", f => f
              .Field(ff => ff.PublishedAs.Suffix("keyword"))
              .ExecutionHint(TermsAggregationExecutionHint.Map)
              .Order(d => d.CountDescending())
              .Size(10)
          )
          .Terms("topn_pub_formats", f => f
              .Field(ff => ff.PublishFormat.Suffix("keyword"))
              .ExecutionHint(TermsAggregationExecutionHint.Map)
              .Order(d => d.CountDescending())
              .Size(10)
```

```
            )
            .Terms("distrib_pubs_agg_language", f => f
                .Field(ff => ff.Language.Suffix("keyword"))
                .ExecutionHint(TermsAggregationExecutionHint.Map)
                .Order(d => d.CountDescending())
                .Size(100)
            )
            .Terms("topn_languages", f => f
                .Field(ff => ff.Language.Suffix("keyword"))
                .ExecutionHint(TermsAggregationExecutionHint.Map)
                .Order(d => d.CountDescending())
                .Size(5)
            )
            .Terms("topn_authors", f => f
                .Field(ff => ff.Authors.Suffix("keyword"))
                .ExecutionHint(TermsAggregationExecutionHint.Map)
                .Order(d => d.CountDescending())
                .Size(5)
            )
        )
        .Sort(d => d.Descending(document => document.Date))
        .Scroll("5m")
        .TrackTotalHits()
    );
}
```

Figure 3.7. Code/full-text queries for aggregations multi-match.

3.2.1.2 Aggregations and Data Visualization

The facets navigation makes it ease narrowing down results for the user with respect to the most relevant properties of the articles. Therefore, aspects as languages, keywords, journals, authors, country, and city will most definitely provide a more concise and straightforward navigation for the user to find the results that are more relevant for his search. This is accomplished in code (Figure 3.8, Appendix I) by calculating the relevant aggregations for all the desired fields and without really running the query, considering that the important result is the total number of documents where these fields are analysed.

Regarding to the different types of aggregation that were done, overall, histograms allow a global overview of the data on a time basis (Nie & Sun, 2017). In this specific case, the idea was to analyse whether during the last two centuries the area of interest had any trends or how deep it has focused on developing research on that field of expertise. The code that handles date histograms (publications during the “two past centuries”) is represented in Figure 3.9 (Appendix I).

Aggregations occur after a query is processed, acting upon the obtained results. Assessing Figure 3.9 (Appendix I), it shows a DateHistogram aggregation designated “hist_pubs_per_year”, using the document publication date as the field for aggregation. Even though the publication date was indexed in its full format (YYYY-MM-DD, HH:MM:SS), for this specific aggregation the goal was to analyse it yearly (**DateInterval.Year**). However, and considering that the idea is to demonstrate the progression historically, the results were also organized ascending by year

(**HistogramOrder.KeyAscending**) (Elasticsearch, 2021c, 2021d). Figure 3.10 (Appendix I) presents the code sample for general and search aggregations that were used in the present study.

Another example is demonstrated as “TopN” in the analysis of the top 5 countries within the top 5 journals and the corresponding top 5 authors, emphasizing in Figure 3.11 (Appendix I) the code that was applied to obtain the necessary data. The data is temporally bounded in approximately 230 years, between 1786 and 2020.

In terms of data visualization, the following graphs were developed to provide a better cognitive understanding and comprehension of the analysis conducted throughout the present work and around the different levels of data transformation and gathering:

1) Search operations:
    a. Facets.
    b. List views.
    c. Pagination/scrolling.

2) Temporal data:
    a. Global distribution.
    b. Historical/Histograms.
    c. Frequency (TopN).
    d. Comparisons.

3) High term frequencies:
    a. Word cloud.
    b. Top N publications.
    c. Geo spatial distribution.

3.2.1.3 Dashboard Analytics – Information Website Visualization

Information visualization places large textual foundations in a visual hierarchy or plan and offers navigation abilities as well as general search. This technique offers improved and faster comprehensive knowledge, which helps us to mine documents of huge accumulation of information. Users can easily distinguish and interpret results by colours and associations, as well as detect related gaps. The variety of documents can be demonstrated as a structured layout using indexing or vector space model. Thus, once the data has been preprocessed (step 1), the following stages of the process are related to the preparation of the data for different forms of visualization (steps 2 and 3), which are available in the website interface. To develop and complete this stage of the process, the chosen tools for the present study were Elasticsearch and React. Elasticsearch consists in a distributed open-source data search and analysis engine that facilitates several forms of processing

and making available big loads of data for visualization platforms in an optimized manner, to provide full-text query capabilities, namely by enabling them to specify *keywords,* which may be present in either the document's title or in its abstract. The technique that was used to accomplish this in Elasticsearch is called multi-match field search. On the other hand, React is an open source library focused on creating user interfaces on web pages. React can be used as a base in the development of single-page, mobile, or server-rendered applications with frameworks. It is maintained by Facebook, Instagram, and a community of individual developers, and it is used on the websites of Netflix, Imgur, Feedly, Airbnb, SeatGeek, HelloSign, and Walmart among others. The only concerned using React was with the state management, where creating React applications usually requires the use of additional libraries.

3.2.1.4 ETL Challenges and Database Performance

It is tempting to think that creating a data warehouse is simply extracting data from multiple sources and loading into database of a data warehouse. This is far from the truth and requires a complex ETL process. The ETL process requires active inputs from various stakeholders including several steps and it is technically challenging.

The database performance has posed some challenges in the present study, being defined as the optimization of resource use to increase throughput, and minimize contention, enabling identifies bottlenecks to largest possible workload to be processed. In more detail, the challenges faced followed the following aspects: 1) crawling in different website providers; 2) the Elsevier integration of Application Program Interface (APIs); 3) the data volume and 4) the articles' metadata normalization. These challenges will be thoroughly and individually assessed in the present subsection.

1) **Problems related to crawling different website providers when fetching results from Web of Science website.**

Although there are many websites being supported by the ETL method and respective software tools, the ratio of fully gathered and well-formatted documents and the number of documents available in these websites in general is approximately 0.0025%. Using initially the Web of Science Core Collection, as the preferred data extraction platform, it was only possible extracted approximately 5 600 publications; however, it is announced in the Web of Science website that there are 79 million publications available. Even though, this comparison is not totally fair, since in this case, the scope of this collection was not limited to publications related just to the main scope of the present work (fungi, fungus and fungal), but in general to evaluate the platform and its receptivity in data extraction.

However, the main problem faced in this step was the fact that most of the information scattered on the website was mostly unstructured data. No obvious structured or processes are in place to make information retrieval easier or easily methodical. In fact, it was only in most recent years that several attempts have been made to make web more semantically accessible and readable (namely, semantic content/website language, structuring hyper-text markup language (html) with meaningful tags and attributes, use of semantic engines, among other facilities) (Markov & Larose, 2007; Sikos, 2015).

Since Web of Science works as a content aggregator, most of the information is just briefly shown in its website or merely Citations from the Web of Science Collection. The main content is hosted in scientific renowned platforms such as Springer, Elsevier, Wiley, which are easily redirected through the DOI (https://www.doi.org/) to main the website for a more detailed reading, but often pays. Hence, there are many websites working in partnership with the Web of Science platform, making it easier to host and structure content from any of their publications, but no open access.

2) Elsevier integration APIs did not enable an easy integration for data collection.

Another potentially interesting platform for data mining would be the ScienceDirect website. ScienceDirect is a website which provides access to a large bibliographic database of scientific and medical publications of the publisher Elsevier. It hosts over 18 million pieces of content from more than 4 000 academic journals and 30 000 e-books of this publisher. Nonetheless, during the analysis of database content for potential collection, Elsevier's integration APIs did not enable an easy integration for data collection, mainly due to 2 factors:

a. APIs usage is highly controlled through a given quota of requests per service. Search API is limited to 20k requests per week, whereas the number of results for the query at hand is in the order of approximately 500k results, which would lead to 25 weeks of data collection for the same amount of data.

b. API communication standards are not effectively used on response serialization, which prevents a complete data collection for analysis. Due to the difficulty of deserializing messages from Elsevier's API for any of the document's identifier (DOI, Enforcement Integrated Database (EID), Personally Identifiable Information (PII)), integration with such services was not complete, since it prevented obtaining the abstract for any of the results.

Considering the challenges encountered regarding to Elsevier integration API, after some research and trial error using C# programming language, it was ended up choosing an alternative tool and language, like the R programming language.

In general, being the literature review process a problem that is divided into several steps, which include the location, appraisal, and synthesis of the obtained information, it was at this stage, that the first challenge was encountered. Even using an alternative extraction tool and different permissions, it was extremely difficult to extract the abstracts from ScienceDirect data collection. Thus, it was only possible to recover and examine 2 750 refereed scientific articles from the ScienceDirect authentic database published between 2010-2020 (see Appendix II, Figure 3.12). Based on the most relevant 2 750 abstract research articles published by keyword group through the last 10 years, this study went through each article in the Elsevier website, retrieving DOI, author's name, publication date, journal name, title and abstract (see Appendix II, Figures 3.13 and 2.14).

Thus, once the abstracts were collected, both irrelevant and unnecessary information was removed, the data were adequately and thoroughly cleaned and analysed. Similarly, the words with low information content and that did not contribute much to the meaning in the text were also excluded from the dataset, considering that they were irrelevant and unnecessary as well (see Appendix II, Figure 3.15). The most frequent words were determined (Figure 3.16) and the main journals assigned (Figure 3.17).

3) Data volume → running data bulk operations always has a significant impact on both processes and speed of analysis.

It is good for transformations to extract and normalize data, considering that high data volumes have an incredibly significant weight on the machine's performance. This required some changes to the existing algorithms (e.g., affiliation transformer, database retrieval layer) since these represent 2 of the main bottlenecks in any system. First, due to the amount of data that is being loaded in memory, and then because of the operations done to transform data into the desired format.

a. Text indexing (implementation).

To fasten indexation, different approaches were required, mainly due to the big amount of data that were being retrieved from the staging area (PostgreSQL) to be indexed in the final storage unit (Elasticsearch index). The strategy that was used initially (Entity Framework Core) applied batching requests of data from database to transform and index small chunks of data, which leaded to higher Central Process Unit (CPU), but less in-memory used.

Due to the high data volume, and to the memory load it created, it was decided to replace the database Object Relational Mapping (ORM) by a more lightweight ORM (Dapper) that decreased the number of tasks that were being done in the background to load entities from the data warehouse. Both Dapper and Entity Framework Core are classified as ORMs Frameworks. In fact, Dapper is classified as a Micro ORM because it is smaller, has fewer features and is therefore lighter. In the case of Entity Framework Core, there are several other facilities, such as being more robust. In both,

the purpose is to facilitate the execution of operations on relational databases (Buse & Weimer, 2012; Maalej & Robillard, 2013).

Since extracting this amount of data, and transforming it in memory, has an associated heavy load on the system (even more if there are not dedicated servers for big data processing), data was retrieved in batches of 2.5k, transformed and indexed to minimize memory starvation. However, the amount of CPU bursts was obviously increased.

Indexation was done per field basis, even though the rules are almost the same for the fields being used. Text fields are mainly used when there is a need for full text search. Since there is a need to make almost every field searchable, it is required to make all fields get indexed to Text datatype to have the field analysed and split to individual terms (Figure 3.18) (Elasticsearch, 2021b).

Data was also indexed in small bulks of data (1000 documents/insert) but in a parallelized fashion, by making use of the machine specifications, namely, the CPU.

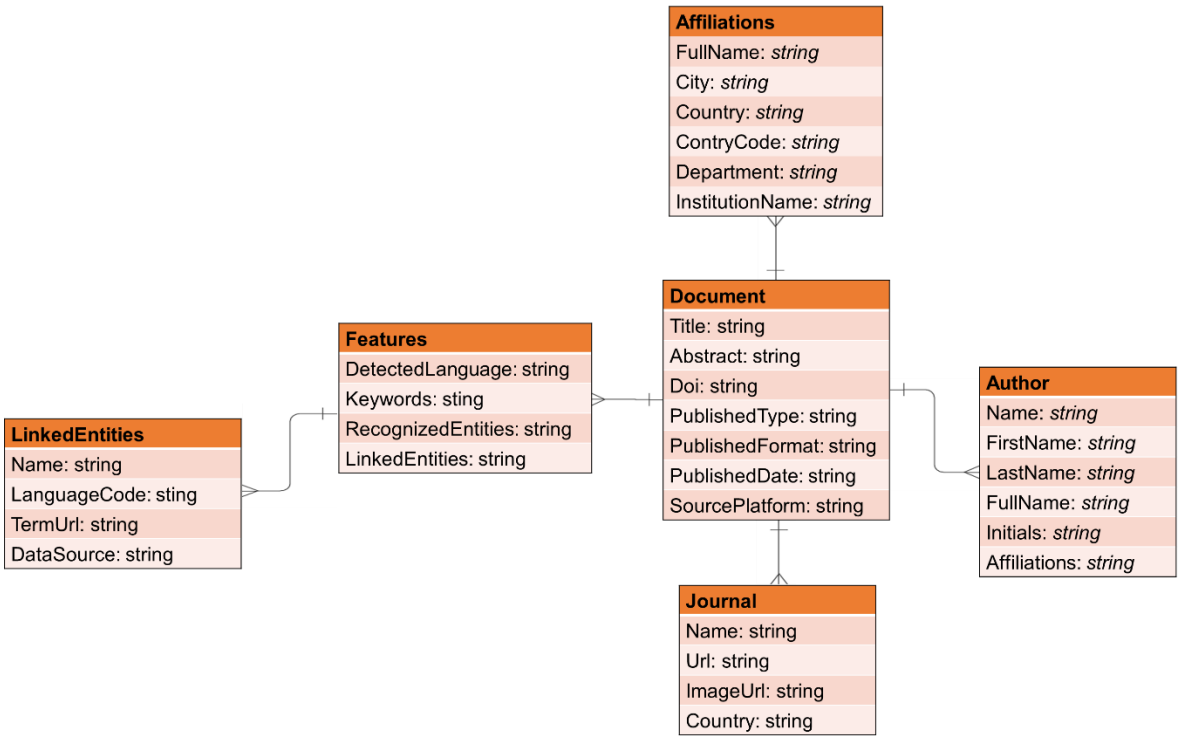

Figure 3.18. Schematic representation of the transformed data index model. Adapted from Feinerer et al., 2008.

b. Authors' naming formats and unique identification issue.

When saving the data from the articles' provider, one of the challenges was to only identify a person when attaching publications that were gathered by the provider throughout the years. This challenge is illustrated in the following table (Table 3.1).

Table 3.1. Problems in terms of the authors' naming formats and unique identification.

Name (Created field)	Affiliation (Journal Name)	First Name	Full Name	Initials	Last Name
Maiara I Costa	Laboratório de Micologia Médica, Departamento de Análises Clínicas e Biomedicina, Universidade Estadual de Maringá, Maringá, Brazil.	Maiara I	Costa MI	MI	Costa
Maiara Ignacio Costa		Maiara Ignacio	Costa MI	MI	Costa
Maiara Ignacio Costa	Department of Clinical Analysis and Biomedicine, Laboratory of Medical Mycology, State University of Maringá, Brazil, Avenida. Colombo, 5790, CEP: 87020-900, Maringá, Paraná, Brazil.	Maiara Ignacio	Costa MI	MI	Costa
Maiara Ignacio Costa		Maiara Ignacio	Costa MI	MI	Costa
Marliete Carvalho Costa		Marliete Carvalho	Costa MC	MC	Costa
Marliete Carvalho da Costa	Departamento de Microbiologia, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Pampulha, Belo Horizonte, Minas Gerais, Brazil.	Marliete Carvalho da	Costa MC	MC	Costa
Marliete C Costa		Marliete C	Costa MC	MC	Costa

Based on Table 3.1, it is possible to observe that the author's metadata differs depending on the journals providing it. Provided examples demonstrate some of the problems that were identified related to the presence of pronouns and name variations that include or are abbreviations, same affiliations being described in different languages or even the lack of any affiliation for the same author.

When saving this information, although these records are related to the same person, at least from the data it is possible to analyse and lead to a valid deduction, due to the many varying name

formats, they were not getting matched automatically. Then, to minimize the number of name variations, the composite field created as “Name” is now only from the author’s “First Name” and “Last Name” fields, with intermediary names being excluded.

E.g.,

1) Maiara I Costa → Maiara Costa
2) Maiara Ignacio Costa → Maiara Costa

This helped to reduce the name space vector and to aggregate affiliation variations per author, instead of creating a new author.

c. Normalizing the author’s affiliation into a structured representation through data transformation.

Extracting countries, cities, and institutions from a denormalized unique field is a complex task. However, it is even more complex when this field does not follow a standard representation. In fact, depending on the journal, time range, and region where the articles were published, different formats were used for the author’s affiliation. The lack of a unique standard used by authors makes it hard to have a consistent way of uniquely identifying the institution, city, country, or any other field through an algorithmic approach. In the context of this work, many “templates” were found when trying to process affiliation. Thus, considering the difficulties encountered in relation to institutions/departments, this query could not be indexed. The following are some of the examples found:

Department of Chemistry, Carleton University, 1125 Colonel by Drive, Ottawa, ON K1S 5B6, Canada.

Department of Chemistry, Carleton University, 1125 Colonel by Drive, Ottawa, ON, Canada K1S 5B6

Department of Chemistry, Carleton University, Canada.

d. City mismatch due to internationalization convention.

When trying to identify/extract cities from affiliation, one problem that rose was the internationalization mismatch. More explicitly, since this work’s scope is based on articles published and analysed worldwide, the affiliation’s normalization algorithm uses international dictionary representations of world countries and cities to allow for consistent search mechanisms (Jonnalagadda & Topham, 2010). Although this was an assumption, that international published

documents' metadata would be restricted to such formats, they may have the author's affiliation in one of the following formats:

1) Institution and/or countries' native language.
2) Institution internationally written (English), but city and institution in countries' native language.
3) Internationally written (English).

E.g.,

1) Laboratoire de Génétique Moléculaire et Cellulaire INRA-CNRS, Institut National Agronomique, Thiverval-Grignon, France
2) cE3c: Centre for Ecology, Evolution and Environmental Changes, Faculdade de Ciências, Universidade de Lisboa, Campo Grande, 1749-016 Lisbon, Portugal.
3) Hematology Unit, Capuchos Hospital, Lisbon, Portugal.

Considering the issues faced in points b., c. and d. trying to find the top co-cited authors and co-cited journals also proved to be an impossible task, due to different non-standard formats founded during this research. Certainly, a number of citations is a very relative indicator of the value, recognition, and importance of the published results. Despite the co-citation analysis is used as an effective method for identifying the intellectual structure of a research domain, must be taken into account it relies on simple co-citation counting and that does not take the citation content into consideration. Moreover, it should be aware that the document co-citation analysis is a method developed usually by bibliometric research, and is often automatically made available on the web, on tools already developed, such as CiteSpace.

3.3 Results and Discussion

Despite the text mining technique has been used over the years, it was only in 2013 by Song and Kim presented the first attempt to apply text mining methodologies to a large collection of full-text scientific articles aiming to analyse and discover the knowledge organization of the area. The results obtained by Song and Kim (2013), showed that most of the documents published in bioinformatics area were not cited by others, demonstrating poverty in information sharing and correlation. Nonetheless, additionally, a constant and linear increase was observed in the number of publications over the years of publication (Song & Kim, 2013).

Creating new information can be quite simple. The difficult and challenging part is finding out relevant information from a large amount of data. Therefore, and to extract patterns, relevant information or knowledge from several sources that are in an unstructured form, the text mining

technique is very efficient, being considered as the best tool to be employed (Miner et al., 2012). Still, any online proposed programs are quite complex, also requiring programming skills that must be learned and tested by trial-and-error. Hence, it was only after understanding the different languages and operations of each one that it was possible to know which one best suited the intended purpose of this dissertation.

For each action/result intended to obtain from the software, well-defined “instructions/orders – algorithms” had to be provided to obtain the desired result. Any programming/language error transmitted to the software does not deliver the desired results. Therefore, one of the first major challenges found referred to the choice of the most suitable software to the study’s purpose, considering that not all the tools available online are free or compatible with the windows system that is being used.

Taking into consideration that the main goal of the present study was to search, find and analyse the global scientific outputs of fungi research and show the progress, trends and hotspots, a systematic review was given approximately in the last two centuries based on data mining technology. Consequently, the second biggest challenge was choosing the proper platform for data extraction. Comprehensively, searches on scientific database from inception, around 1786 to 2020, and a total of 500 452 related literatures were identified. Based on the text mining analysis, information about keywords, publication format, document type, journals, authors, countries, cities, and language from all the literatures founded containing at least one of these keywords: "fungi” or “fungus” or “fungal” from papers title and abstract were systematically summarized.

3.3.1 ETL Programming Language

The global geographic distribution regarding to the number of publications in the first-time range (1786-2020) (Figure 3.19), showed considerable differences among the conducted research worldwide. Most countries have developed truly little research about fungi topic in general in the present study (1.0k – 13k (thousand)), with only one region presenting remarkable research during this period (about 130k – 140k), being it, the United States of America (USA). Be note, that the graphical representation is by states. So, some regions, such as Alaska, although it may not have developed as much research as shown schematically in the graph (black colour), but because it is considered a state belonging to the United States, it is represented with the same colour. When analyzing the last 10 years, there was no great variance in research around the world, therefore, and regarding the global geographic distribution of publications between 2010 and 2020, Figure 3.20 (see Appendix III) demonstrates that they occur worldwide, with no significant differences when compared to what was previously stated in the first-time range (most countries developed little research about the subject in study; 1.0k – 13k). Still, the overall effort of conducting more research about marine and terrestrial bioactive compounds was quite evident, with several countries presenting significant

numbers of conducted research. Nevertheless, it is quite outstanding the fact that every country worldwide has been developing some research about the subject throughout over the years.

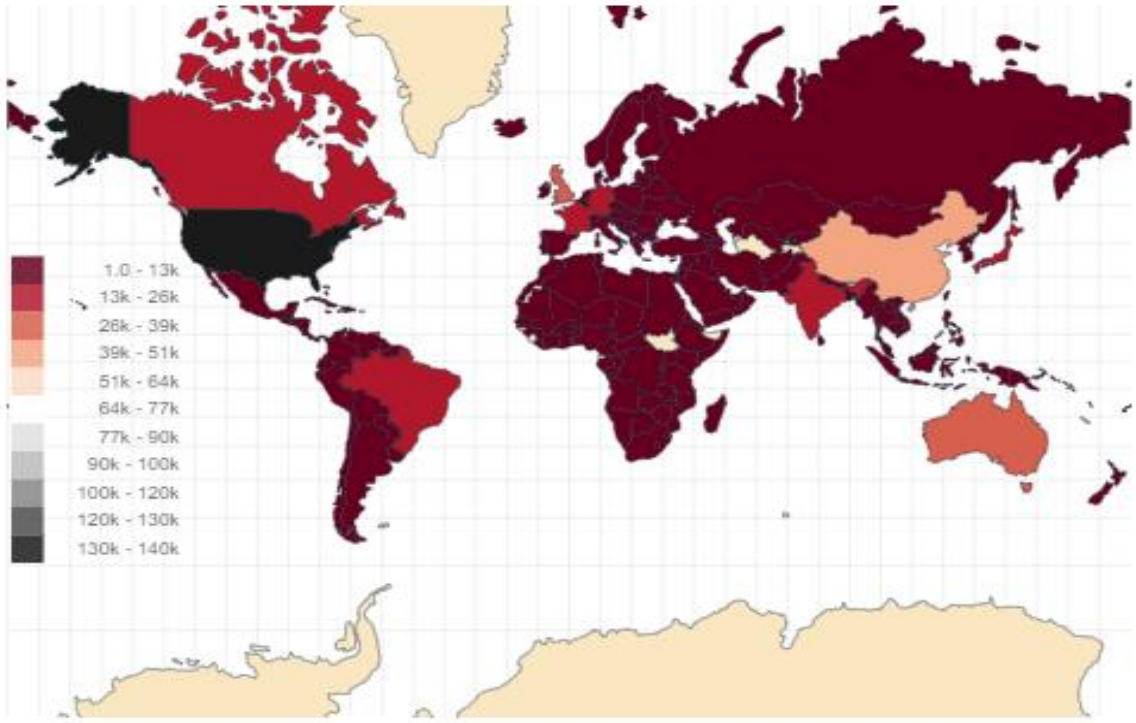

Figure 3.19. Global geographic distribution of scientific publications between 1786 and 2020. Number of scientific papers expressed in k (thousand) units with a colour gradient.

Considering the publication per year in a general viewpoint, it was possible to verify that there is a distinct ratio throughout the specified time range (Figure 3.21 and Table 3.2 - Appendix III). In more detail, at the beginning of the studied time range (1786) there were only two publications of research about the subject under analysis. This seems to be the tendency up until 1853, when there were for the first time 20 publications in total, this number being observed again only in 1893. Nonetheless, it was only in 1946 when the number of publications hits the three digits, with a total of 190 registered. As time went by, several other milestones were achieved: about 1 000 publications per year (in 1968 – 1 184 publications); 2 000 publications per year (in 1982 – 2 060 publications); 5 000 publications per year (in 1992 – 5 212); 10 000 publications per year (in 2004 – 10 307 documents); 25 000 publications per year (in 2015 – 25 978 articles); and mora than 33 300 publications per year (in 2019 – 33 330). Overall, and based on Figure 3.21 and Table 3.2, it was possible to verify that the number of publications per year started to increase after 1952, with a constant, and sometimes even substantial, increase up to the final year of 2020. Analysing in more detail, the last 10 years, in terms of the number of publications per year during the time range between 2010 and 2020, these have continuously increased until 2019, from around 16 884 in 2010

up to 33 330 in 2019. In 2020, the number of publications recorded was only 18 156 publications in that year (Figure 3.21). This can be justified because there is often a delay between the acceptance and publication date of scientific articles on the website and the updating of the platforms' database. In addition, some gaps were detected on specific dates (Table 3.2 - Appendix III), with no publication being recorded, this is believed to happen if the documents are not available in text format, but in PDF format or even unavailable.

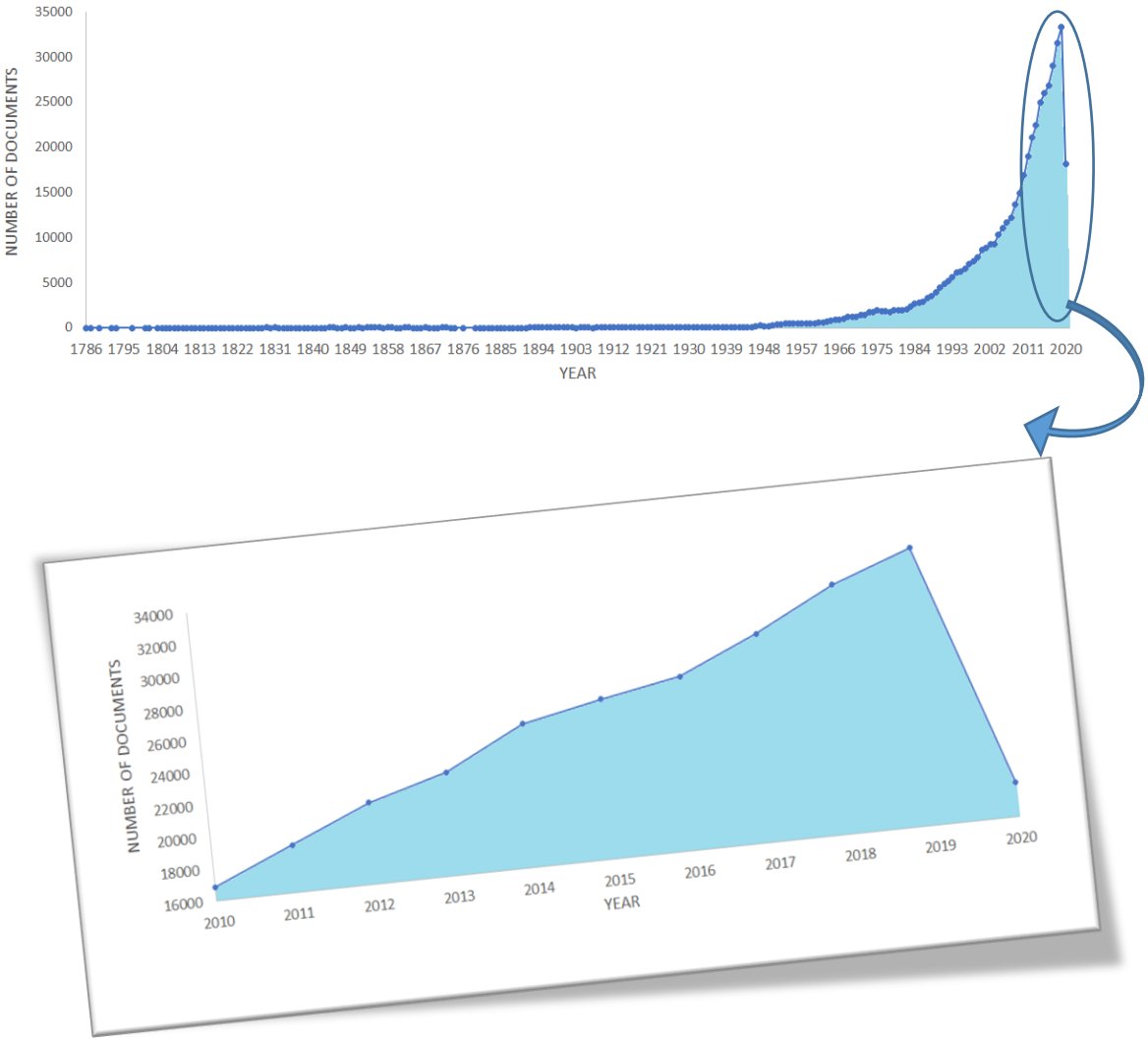

Figure 3.21. Publication per year between 1786 and 2020 (above) and in more detail the number of publications per year between 2010 and 2020 (below).

The absence of scientific articles with similar comparative studies makes it difficult to support this study with bibliography. Nonetheless, although the database used to extract the base

documentation of this study was through the Europe PMC platform (Europe PMC Collection), to validate our results, a comparative study was performed directly using the SEARCH tool provided by the website “Europe PMC”. The results obtained are expressed in Figure 3.22, showing a similar number of publications per year in last 10 years for specific keywords, namely “fungi”, “fungus” and “fungal” contained in the Title and Abstracts of the scientific papers, showing an average percentual difference between values of each platform of about 4.5%, corroborating our research. A discrepancy between the number of documents obtained and the comparative documents with a percentage difference of about 55% was observed for the year 2020, this was due to the fact, that updates/feeds were not being immediate on the website, respecting a certain periodicity of time for updates. By interest and to give more robustness to the work developed, a similar research attempt was made using the "Web of Science" website (see Appendix III, Table 3.3). The results obtained by this website compared to our study showed an average percentual difference between the data of about 9%. Additionally, when comparing the results between the two validated websites (Web of Science vs Europe PMC), the average percentage difference between values were about 6%, providing a low difference when compared to our results.

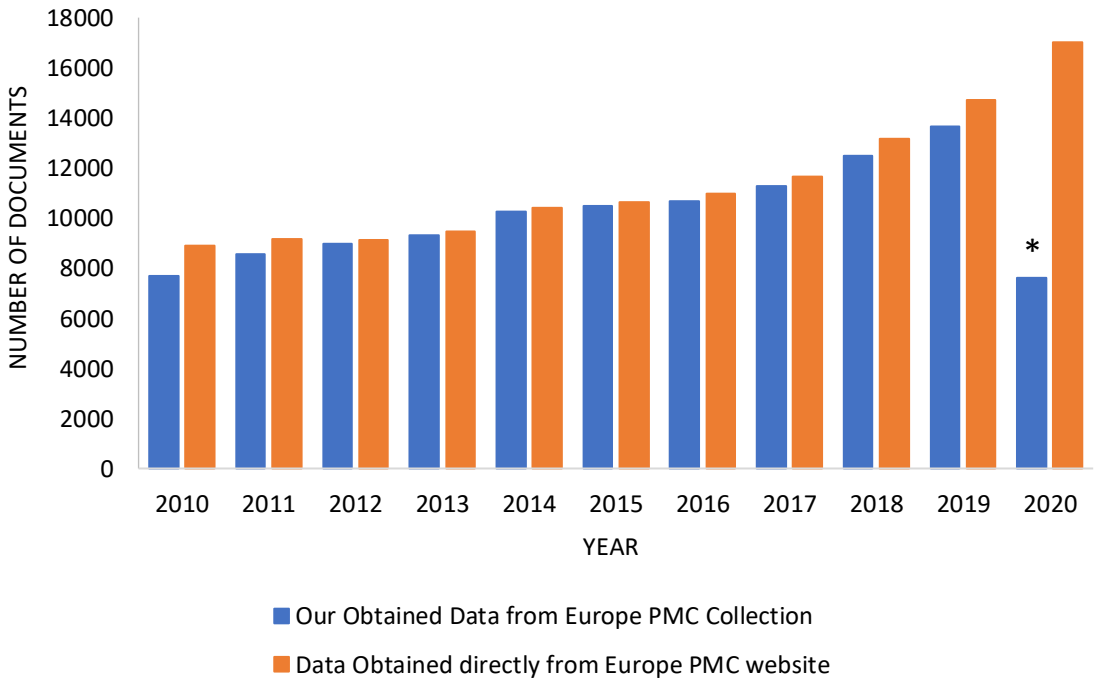

Figure 3.22. Comparative number of publications per year in last 10 years in the Title and Abstract between the results obtained using Data mining tools and the Europe PMC search website. (*Value with a percentual difference of 55% between the data obtained in this study and the data obtained directly from Europe PMC website).

Equally interesting additional research was to know how many relevant documents have been published over the last decades (between 1786 and 2020) mentioning at least one of the keywords of interest in the Title and Abstract (Figure 3.23), which will allow us to have an overview of the evolution of publications and, consequently, of the investigations that have been developed in the desired themes. As expected, regardless of the keyword, the Abstract was the location where the terms of interest were most often founded. Furthermore, “fungal” seems to be the most published term over the years, followed by “fungi”. Nonetheless, it should be noted that the keyword "fungal" was much broader than the term "fungi", since "fungi" was any member of the group of eukaryotic organisms that includes microorganisms, such as yeasts and moulds, which are characterized by a substance in their cell walls called chitin and the plural of "fungus", while the term "fungal" is often used to refer to diseases like "fungal infection" or "fungal diseases". In addition, it was possible to clearly observe that a weak relationship between these microorganisms and their potential antibacterial capacity is little investigated. Incredibly, contrary to expectations, studies related to the terms "marine" and "terrestrial" are still far from expected, since the environment being a sustainable source of bioactive compounds with a broad spectrum of activities (Carroll et al. 2021; Gladfelter et al., 2019; Jones et al., 2019; Mayer et al., 2021; Peyrat et al., 2019). An extra survey of the number of publications found in the Title or in the Abstract containing at least one of the relevant terms was also carried out and analysed (Table 3.4, see Appendix III). Please note that the number of publications obtained in “#Documents with term in Title OR Abstract” column is not the direct sum of the values obtained individually for “#Documents with term in Title" and “#Documents with term in Abstract". The sum of the values presented in the columns “Title" and "Abstract" cannot be coincident with the values of the column “#Documents with term in Title OR Abstract”, because there may be papers where the same keyword is present either in the Title or in the Abstract, thus considering that scientific article only once.

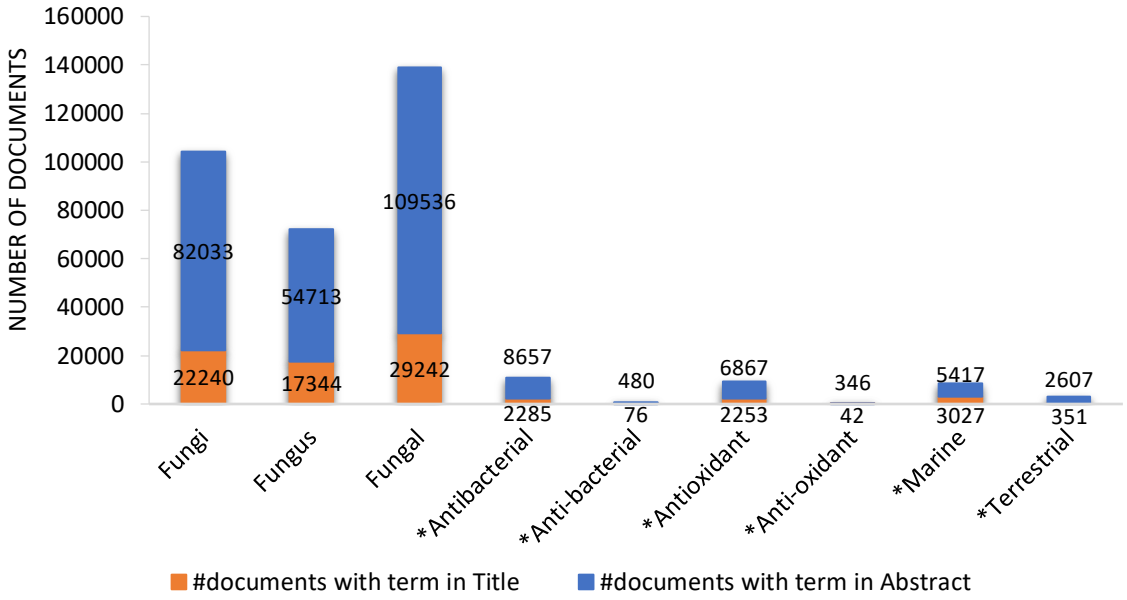

Figure 3.23. Number of publications by relevant terms searched in the Title or Abstract between 1786 and 2020. Keywords containing an "*" showed a combination with at least one of the following keywords"fungi" or "fungus" or "fungal".

Within the shorter time range, namely between 2010 and 2020, another aspect was assessed due to its pertinence to the present study. More precisely, the ETL programming language was also used to assess a detailed evolution in terms of the number of publications for both marine and terrestrial bioactive compounds' research. A graphic representation was developed and according to Figure 3.24, it was possible to conclude that the number of publications for both topics have been gradually increasing through time in this last 10-year period, even though the number of publications for both topics stabilized slightly in 2020. Still, there are more publications regarding marine bioactive compounds throughout the years when compared to the publications about terrestrial bioactive compounds. During years, pharmaceutical research was mainly dedicated to terrestrial plants and microorganisms, not only because they were more easily accessible organisms, but also because popular traditions associated their use to beneficial effects (Ameen et al., 2021; Khattab & Farag, 2021). However, it appears that there has been a shift in the focus of researchers towards exploring the huge biodiversity of oceans and seas in the hope of identifying and taking advantage of new and unique biological compounds produced by marine organisms (Ameen et al., 2021; Khattab & Farag, 2021; Rai et al., 2018).

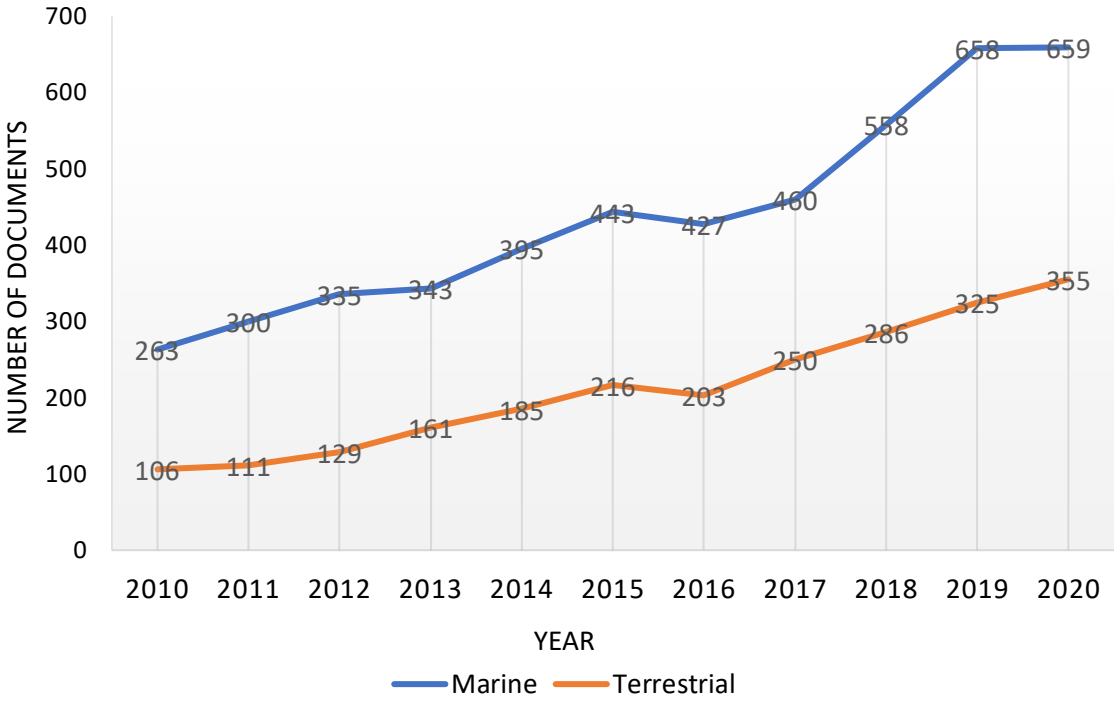

Figure 3.24. Marine vs. terrestrial fungi bioactive compounds' publications between 2010 and 2020.

The ETL process also allowed to investigate the main activities of both marine and terrestrial bioactive compounds (see Appendix III, Figure 3.25). Essentially, between 1932 and 2020 the most documents that were published demonstrated that the main activities studied related to fungi were the antibacterial activity, one of the activities demonstrated in bench studies in this dissertation in the next chapter (Chapter 4, section 4.3.3). The date of 1932 was the first record of publication relating fungi and their potential activities. Equally in agreement are the results obtained from this survey with our laboratory findings regarding to the antioxidant activity of these microorganisms (Chapter 4, section 4.3.4). Nonetheless, many more studies can be developed in this direction, considering the encouraging results obtained in the extraction of bioactive compounds with promising activities.

Largely, most of the published documents indicated that the main activities are the following ones, respectively:

1) Antibacterial activity.
2) Antimicrobial activity.
3) Antibacterial.
4) Antimicrobial.
5) Antifungal.
6) Cytotoxicity.
7) Antifungal activity.

8) Silver nanoparticles.
9) Antioxidant.
10) Antimicrobial peptides.
11) Antioxidant activity.
12) Essential oil.
13) Antiviral.
14) Anticancer.

After analysing in general the distribution of relevant terms (fungi, fungus, fungal, antibacterial, antioxidant, marine and terrestrial) studied, it was possible to observe in more detail some relevant parameters. Nowadays, and mainly due to the overwhelming amount of textual information presented in scientific literature, there is a need for effective automated processing that can help scientists to locate, gather, and use the knowledge that is encoded in the literature that is electronically available. Thus, the specificities that were assessed for each time range were based on the keywords (Top 10); most relevant terms related to top10 keywords founded; publication format and document type of the papers (Top 5 and Top 10, respectively); journals (Top 10); authors (Top 10); countries (Top 10); cities (Top 10) and languages (Top 5).

Regarding to the most popular keywords in the papers that were published between 1786 and 2020, these were essentially the following ones, as it is demonstrated in Figure 3.26 (see Appendix III):

1) “Fungi” (4908 publications).
2) “Candida albicans” (1894 publications).
3) “Inflammation” (1587 publications).
4) “Bacteria” (1584 publications).
5) “Infection” (1429 publications).
6) “Yeast” (1316 publications).
7) “Antimicrobial activity” (1282 publications).
8) “Microbiome” (1268 publications).
9) “Aspergillus” (1211 publications).
10) “Antifungal activity” (1158 publications).

A word cloud was built representing the most representative keywords between 1786 and 2020 (see Appendix III, Figure 3.27). Overall, this graphic representation reflects the most frequent terms and keywords among the published papers about fungi, which allows a quick comprehension of the most common terms analysed and searched for within the area of expertise. According to Jayashankar and Sridaran (2017), word clouds are defined as the visual representation of words for a particular written content, being structured according to their frequency, giving greater prominence to the words that appear more repeatedly in a text. Thus, the larger and visual a particular word, the

more common the word was in the abstracts. The word clouds tend to serve as a preliminary stage for in-depth analysis of specific text material, since they truly help to find the relevancy between the given text and the required information. The font size and the colour spectrum help to identify and visualize, quickly and easily, the words of interest, which in this case was clearly “fungi”. This observation seems contradictory, since "fungal" had been the predominant keyword in the title and abstract of the researched scientific articles. This phenomenon stems from the interpretation of what were highly relevant terms, in detriment of terms with high frequency. Thus, quickly, the Inverse Document Frequency (IDF) is a statistical analysis already incorporated in Elasticsearch that analyses the frequency with which a word appears throughout the document together with its inverse frequency to determine the relevance of the term for a more targeted, accurate and objective search. Jones (1972) conceived a statistical interpretation of the term IDF, which became a base concept for the weighting of terms. Documents with more frequent terms do not necessarily mean that they are more relevant to different searches. As mentioned above, “fungal” word was a more generic and encompassing word than “fungi”, that is more objective and directional and, the focus of this research work.

Figure 3.27. Word cloud representing the hundred most frequent words published between 1786 and 2020.

Equally interesting, the most popular keywords in publications between 2010 and 2020 were basically like those presented in the first-time range, even though there are some minor differences, such as different rankings for some keywords, with an exchange of positions between “Microbiome” and “Antimicrobial activity” and a move up to the “Antifungal activity” keyword, demonstrating greater

study interest in this activity in last 10 years. Despite maintaining the first position as the most popular keyword, “fungi” showed a decrease of almost 50% in the number of publications. In more detail, the top 10 ranking is the following one (see Appendix II, Figure 3.28):

1) “Fungi” (2844 publications).
2) “Candida albicans” (1865 publications).
3) “Inflammation” (1573 publications).
4) “Bacteria” (1382 publications).
5) “Infection” (1373 publications).
6) “Yeast” (1285 publications).
7) “Microbiome” (1268 publications).
8) “Antimicrobial activity” (1241 publications).
9) “Antifungal activity” (1124 publications).
10) “Antifungal” (1101 publications).

During this substantial period (1786-2020), it has been demonstrated that there are several terms that were relevant in relation to specific top keywords among the existing research about the subject under discussion (see Appendix II, Figure 3.29). Indeed, the retrieved data showed that in a total of 100 terms identified as relevant within the conducted research, for a perceptible representation only 35 of the terms were presented and all directly linked to the top keywords, namely: “bacteria”, “infection”, “yeast”, “antimicrobial activity”, “microbiome”, “aspergillus”, “antifungal activity”, “experimental lab study”, “plants”, “field crops”, “biofilm”, “candida glabrata”, “fluconazole”, “biofilms”, “innate immunity”, “cytokines”, “cancer”, “virus”, “viruses”, “archaea”, “sepsis”, “immunity”, “saccharomyces cerevisiae”, “mitochondria”, “transcription”, “antioxidant activity”, “cytotoxicity”, “essential oil”, “silver nanoparticles”, “microbiota”, “metagenomics”, “gut”, “candida”, “penicillium”, and “antibacterial activity”. A connection between keywords and related terms was presented, showing a world of interconnections and possible scientific research, evidencing the main research concern around microorganism and viral infections or diseases, like immune or cancerous related. Thus, it is urgent the constant demand for the unknown.

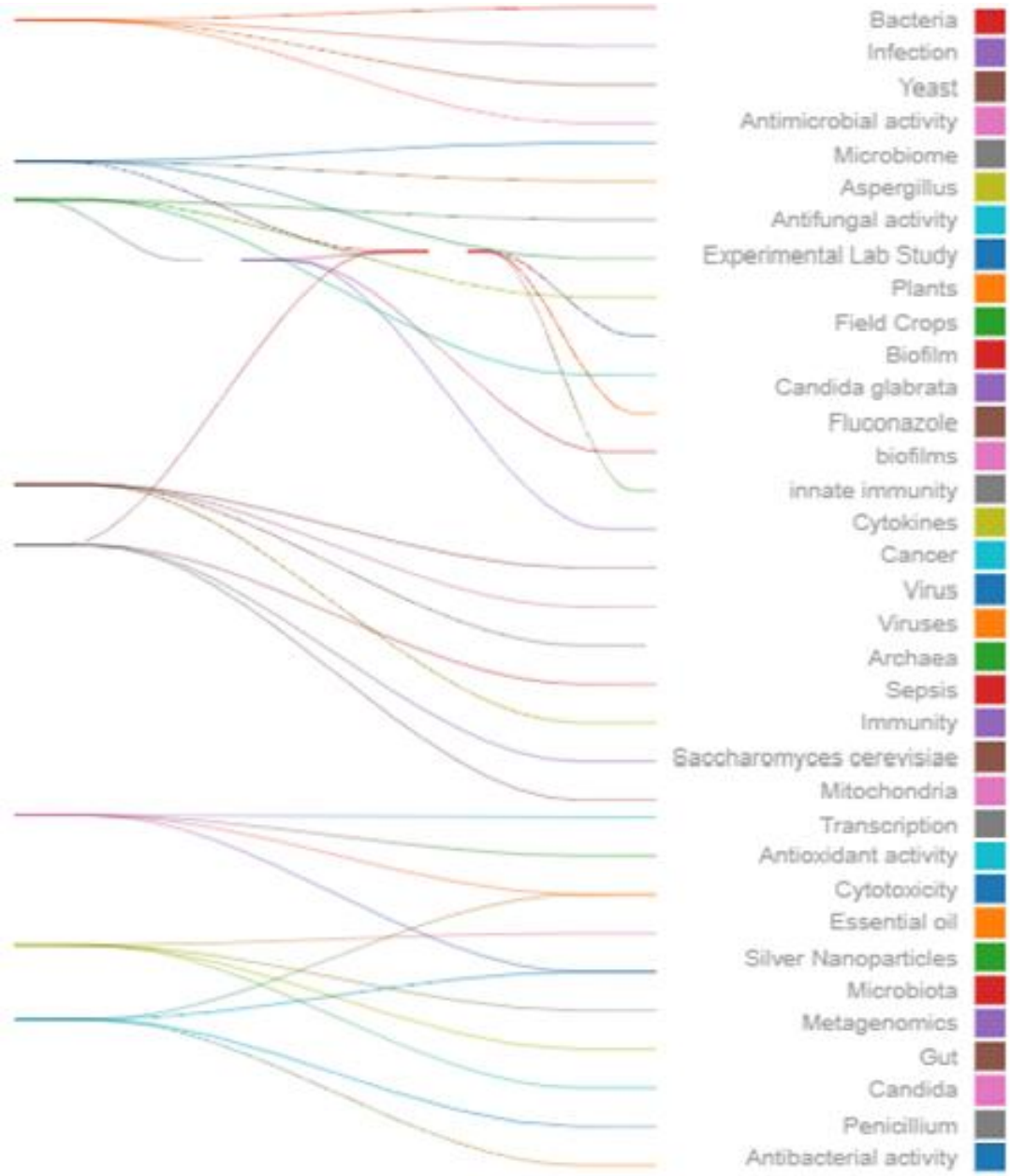

Figure 3.29. Most thirty-five relevant terms related to top10 keywords obtained in scientific articles between 1786 and 2020.

Another interesting representation was that evidenced in Figure 3.30 (see Appendix III). It summarizes the top 10 keywords based on the obtained data. Overall, this figure summarizes the most frequently used keywords in publications about the subject under analysis, consisting in a combination of all the previously conducted representations and analyses in both time ranges.

It has been previously demonstrated that text clustering is one of the most effective techniques to be used for text mining analysis (Salloum et al., 2018; Justicia de la Torre et al., 2018). Additionally,

this technique also facilitates the analysis of topics where concurrent occurrences were grouped, so that the most frequent items were put together by applying the graph-based method. Each occurrence that was grouped was represented by a cluster that relates to one of the relevant topics.

Regarding to the publication format between 1786 and 2020, it was possible to conclude that most publications were “print” (more than 200 000 publications), followed by “print-electronic” format (about 150 000 publications), which is understandable, given the time period under study, as in the 18th century there was no electronic technology whatsoever. With the appearance of the “print-electronic” format, it was possible to detect the evolution of the times and the appearance of technology and its use in the scientific area. Still, there were also publications with other formats, such as “electronic” (around 50 000 publications), “electronic-eCollection” (less than 45 000 publications), and about 30 000 publications with undetermined format (see Appendix III, Figure 3.31). Considering the publication format between 2010 and 2020, Figure 3.32 (see Appendix III) demonstrated that a significant number of publications were “print-electronic” (118 128 publications). Still, other formats are quite frequent, such as “print”, “electronic”, and “electronic-eCollection”. Comparing to the previous time range, it was possible to conclude that there was an evolution in terms of the publication format, with most publications being “print-electronic”, rather than “print”. This might be explained by the constant adoption of technology among several spheres of daily life, including in academic research and investigations. It is well known that electronic resources offer wider access, to more users and at any time, and they can offer greater functionality, i.e. full-text searching, combination searching, but for them to be a suitable choice, they are measured against several standards, such as accessible for all, ease of use, clean design, preservation of customer privacy to the greatest degree possible, sustainable cost of purchase and continued access, with future cancellation options, the supplier provides reliable service and quickly responsive to performance issues, among others. It is not always possible to meet all standards, so benefits must be balanced against drawbacks. On the other hand, “print” is often chosen for materials to be read cover-to-cover, for browsing or for use over a longer duration. Image quality in print may be a factor, also. Print is preferred when the electronic options fail to meet standards above described or where the cost of the electronic version, by contrast, is prohibitive. Regarding the publications in ebooks, more adherence to the format would be expected, however, it is necessary to keep in mind that electronic access to books is an evolving landscape and presents many challenges to libraries.

Another pertinent aspect that was assessed in terms of the gathered data refers to the document type of each publication. According to Figure 3.33 (see Appendix III), most of the publications were journal articles (481 265 publications), immediately followed by research articles (203 394 publications), and research support (Non-United States (U.S.) Government; 183 725 publications). Other types of documents that were also identified, but with less preponderance in the time range (1786-2020), were reviews (with 59 763 publications), research support (National Institutes of Health (N.I.H.), extramural; 34 969 publications), case reports (28 319 publications), research support (U.S. Government, Public Health Service (P.H.S.); 25 973 publications), research support (U.S. Government, Non-P.H.S.; 25 921 publications), review articles (25 912 publications), and

comparative studies (with 23 680 publications), respectively. On the other hand and considering the document type of the publications that were published more recently (between 2010 and 2020), most of them were journal articles, just like in the previous time range, accounting for 258 286 publications (see Appendix III, Figure 3.34). Other types that were identified were the following ones, and respectively: research articles (with 134 053 publications), research support (Non-U.S. Government; 99 363 publications), reviews (counting 39 295 publications), research support (N.I.H., extramural; 25 672 publications), review articles (23 223 publications), case reports (14 16 publications), research support (U.S. Government, Non-P.H.S.; 12 155 publications), case reports (9 456 publications), and comparative studies (5 216 publications). Interestingly, during our independent research, several types of documents, such as “research support” were found, which are distinguished by their funding support. Non-U.S. Government contribution can be from any non-US government agency, for example state and local governments, foreign governments, and private organizations. The N.I.H. office of extramural research supports the study of many aspects of minority health and health disparities from genetic, molecular, and biologic science to clinical, behavioural, and translational research, as well as research on health systems, workforce development, and environmental justice. On the other hand, U.S. Government, Non-P.H.S. acknowledges that funding support is from any US government agency other than the P.H.S., such as the National Science Foundation, NASA, Department of Energy, among others.

Comparing our results with those presented on the Europe PMC tool search available on the website, it was possible to observe that the largest number of scientific publications was presented only in three types such as “Research articles”, followed by “Reviews” and “Preprints” with around 400 808, 45 332 and 6 474 publications (between 1786 and 2020), and 250 723, 27 954 and 6 474 publications during 2010-2020, respectively. A slight discrepancy in the number of publications was observed between our survey and the comparison Europe PMC website (around 45% of percentual difference), this can be explained because in our research, the various types of documents are presented in raw and a document can be classified into several categories, that is, a unique scientific paper can be classified as "Journal article" and "Research article" for example

As per Berg et al. (2016), preprint is a complete scientific manuscript, often one also being submitted to a peer-reviewed journal, which is uploaded by the authors to a public server without formal review. After a brief examination to ensure that the work is scientific in nature, the posted scientific document can be viewed without charge on the Web. Thus, preprint servers facilitate the direct and open delivery of new knowledge and concepts to the worldwide scientific community before traditional validation through peer review (Berg et al., 2016; Vale, 2015). These manuscripts that have not undergone peer-reviewed were first adopted in physics sciences in the early of 1990s, to create an open online repository for scholarly documents. Nevertheless, only after 2013 that similar initiatives were adopted by the biological, followed by medical sciences, however, new publishing platforms continue to emerge. Schalkwyk et al. (2020) believe the potential for harm is outweighed by the benefits, but others have raised specific concerns regarding to medical preprints and mitigating the risk of harm to the public (Schalkwyk et al., 2020). These discussions have

intensified during the covid-19 pandemic, which has been accompanied by an explosion of preprint publications (Berg et al., 2016; Ginsparg, 2016; Hoy, 2020; Maslove, 2018). Another concept usually found in some scientific information repositories is “proceedings paper”. Proceeding’s papers are documents initially presented at a conference or workshop and later adapted for publication in a journal. In 2008, the document type “proceedings paper” was assigned in the Web of Science database to journal articles that were initially presented at a conference and later adapted for publication in a determined journal (González-Albo & Bordons, 2011). The study performed by González-Albo and Bordons (2011), including the period between 1990 and 2008, showed that the “proceedings paper” account for approximately 9% of all articles in this area, two thirds of which were published in monographic issues dedicated to conferences, which tend to focus on specific journals. Proceedings papers emerge as a heterogeneous set including proceedings paper in ordinary issues, similar to standard articles in structure and research impact; and proceedings paper in monographic editions, which appear to be less comprehensive and tend to receive fewer citations. The quicker publication of proceedings paper in monographic than in ordinary editions can hide differences in the review process undertaken by any type of paper. Furthermore, the use of two different labels (“article” and “proceedings paper”) can lead to inferring differences in their relevance and/or quality, as well as aggravated by the fact that most websites do not have a specific search tool or a query to apply this feature, it is quite possible that the labels “article” and “proceeding paper” are considered in many searches as a single category, not checking if the “proceeding paper” was initially presented at a conference or workshop (González-Albo & Bordons, 2011; WoS, 2009).

Therefore, and considering that most publications were journal articles, the top 10 journals were also identified in this analysis. The most prominent journal was the *PloS ONE* (13 732 publications), immediately followed by *Proceedings of the National Academy of Sciences of the United States of America* (7 184 publications), *The Journal of Biological Chemistry* (6 654 publications), *Applied and Environmental Microbiology* (6 483 publications), *Scientific Reports* (5 622 publications), *Plant Disease* (5 460 publications), *Journal of Bacteriology* (4 653 publications), *Frontiers in Microbiology* (4 255 publications), *Antimicrobial Agents and Chemotherapy* (4 178 publications), and *Journal of Clinical Microbiology* (4 028 publications) (see Appendix III, Figure 3.35). Considering the last decade, the popularity ranking of journals has changed a little, although the first place continues to be occupied by *PloS ONE* with 13 260 publications, followed by *Scientific Reports* (5 622 publications), *Frontiers in Microbiology* (4 255 publications), *Molecules* (3 191 publications), *Frontiers in Plant Science* (2 984 publications), *International Journal of Molecular Sciences* (2 774 publications), *Plant Disease* (2 482 publications), *Proceedings of the National Academy of Sciences of the United States of America* (2 290 publications), *The Journal of Biological Chemistry* (2 141 publications), and the *Applied Microbiology and Biotechnology* (1 542 publications). In comparison to the previous time range, it was possible to conclude that the top journal was the same (*PloS ONE*), and that in this latter time range (2010-2020) there were new journals that were not in the top 10 in the previous time range, such as the *International Journal of Molecular Sciences* and the *Molecules*, which may highlight new research trends (Figure 3.36, Appendix III). Regarding to this topic, since

there is no directly related field in the search tool of the Europe PMC website, it was not possible to compare the results obtained, but considering that *PLoS ONE* journal belongs to the first quartile (Q1) of SCImago Journal Rank (SJR) ranked journals and have an impact factor (IF) of 3.24, which may arouse interest and be considered for publication. Interestingly, according to MacCallum (2006), most science is not published in *Science*, *Nature*, *Cell*, or even *PLoS Biology* journals. Indeed, the increasing pressure from submissions, limited page budgets, and the existing reward system whereby the price of a paper is placed not on its content, but rather on where it is published, has led most journals to reject a considerable quantity of papers prior to peer review. But in 2006, in the opinion of the same author, the "journal," *PLoS ONE* (http://www.plosone.org/) was published, which began a radical departure from the suffocating restrictions of this existing rejection system (MacCallum, 2006).

Out of curiosity, 10 specific authors were highlighted due to their significant contribution in terms of conducted research in this first-time range, which comprises the years of 1786 and 2020. As a matter of fact, these 10 authors have published more than 250 publications during their careers, which is why they deserve a special emphasis in this area of expertise (see Appendix III, Figure 3.37). Hence, the top 10 is the following one:

1) Arturo Casadevall (383 publications).
2) Wei Wang (349 publications).
3) Joseph Heitman (341 publications).
4) G. de Hoog (339 publications).
5) Dimitrios Kontoyiannis (334 publications).
6) Wei Zhang (332 publications).
7) Thomas Walsh (323 publications).
8) Yan Li (317 publications).
9) Jing Li (302 publications).
10) Wei Li (283 publications).

A list of authors should only reflect those who have made substantial contributions to a research project and their draft manuscript, therefore, every author on the list should be credited equally, since it takes a team to successfully complete a project. A search by author, allowed us to have a perception of the quantity of scientific articles by department or laboratory, since it became difficult to reach a conclusion about the author's affiliation (see section 3.2.1.4, items 3 b-d)

Comparing the list of the top 10 authors in this second time range between 2010 and 2020, the ranking of those who published a significant number of publications within this time range was also established. Curiously, the greatest to the topic under study contribution has been recorded in the last 10 years, with special attention to USA authors, namely Wei Wang, Wei Zhang, Yan Li, Jing Li and Wei Li. The ranking goes as follows and according to Figure 3.38 (see Appendix III):

1) Wei Wang (322 publications).
2) Wei Zhang (308 publications).
3) Yan Li (290 publications).
4) Jing Li (264 publications).
5) Wei Li (259 publications).
6) Yang Liu (255 publications).
7) Dimitrios Kontoyiannis (236 publications).
8) Jacques Meis (236 publications).
9) G. de Hoog (222 publications).
10) Arturo Casadevall (205 publications).

Considering that there is no search field in the Europe PMC search tool to order the ranking of authors with the highest number of publications on the topic related to fungi, this was another parameter that it was not possible to compare the results obtained with the documentation available directly on the website. However, despite each website having its own database sources, a search on the Web of Science website helped to clarify the results obtained, highlighting names such as Y. Li with 594 publications; Y. Liu (458 publications) or even J. Li (413 publications).

In a broader scope, but also related to the vast number of publications that were published between 1786 and 2020, the top 10 countries were assessed by the ETL process (see Appendix III, Figure 3.39), which resulted in the following ranking:

1) United States of America (141 200 publications).
2) China (50 186 publications).
3) United Kingdom (UK, 27 979 publications).
4) Australia (27 200 publications).
5) Germany (23 023 publications).
6) Japan (22 230 publications).
7) Brazil (19 637 publications).
8) India (19 573 publications).
9) Canada (17 744 publications).
10) France (16 131 publications).

Regarding a specific scenario and continuing to lead the first two positions in the ranking, it was found the USA and China, respectively, which corroborates the results obtained for the authors with the largest publications. The United Kingdom gave place to Australia in the last decade, and the biggest surprise was for Germany and Japan, which dropped two positions in the ranking of countries with the highest number of publications related to fungi topic. The top 10 countries with the most significant numbers of publications are the following ones, respectively (see Appendix III, Figure 3.40):

1) United States of America (81 556 publications).
2) China (41 921 publications).
3) Australia (16 043 publications).
4) United Kingdom (15 649 publications).
5) India (14 798 publications).
6) Brazil (14 194 publications).
7) Germany (13 910 publications).
8) Japan (10 541 publications).
9) Mexico (10 184 publications).
10) Canada (9 786 publications).

For an out of curiosity and for a more realistic data interpretation, the publication coefficient per 100 thousand inhabitants between 2010 and 2020 was determined (Table 3.5). Hence, the following results showed a curious change in the ranking of countries with the highest number of publications, with Australia remaining in a leading position, but curiously followed by Canada, which occupied the last position in the top 10 of countries in last research results. The USA remains in the first three positions and the UK has not changed its position, proving to be the most balanced country in terms of the number of scientific populations versus general population per country. India occupies the last place, since it is necessary to keep in mind that it is a country with a high underprivileged population.

Table 3.5. Publication incidence coefficient per 100 thousand inhabitants between 2010 and 2020.

Country	Average population between 2010 and 2020	Incidence coefficient
Australia	24 362 250	65.85
Canada	36 502 435	26.81
United States of America	323 593 481	25.20
United Kingdom	66 414 747	23.56
Germany	82 555 999	16.85
Japan	127 476 061	8.27
Mexico	123 822 999	8.22
Brazil	206 751 507	6.87
China	1 415 926 411	2.96
India	1 329 527 432	1.11

Europe PMC support department was contacted, since there was no possibility of searching by country, as well as cities in the search tool available on the website. Unfortunately, they do not have this parameter captured separately on the website. The only alternative is look at the

"authorAffiliationDetailsList" to see manually the affiliations of the authors, which is unaffordable, given the many thousands of existing records.

Similarly, 10 top cities were highlighted in this time range (1786-2020), especially due to the significant number of publications that were published there (see Appendix III, Figure 3.41). As can be verified, the cities with the highest number of published documents are not in agreement with the countries that had more publications in recent decades. These data were influenced due to the challenge already described in section 3.2.1.4, 3-d). Anyway, the top 10 of cities with most publications between 1786 and 2020 are the following:

1) Beijing (15 371 publications).
2) Bambari (9 385 publications).
3) Shanghai (6 672 publications).
4) Boston (6 316 publications).
5) Guangzhou (6 227 publications).
6) Nanjing (6 148 publications).
7) London (5 584 publications).
8) Tokyo (5 470 publications).
9) Wuhan (5 291 publications).
10) Bethesda (5 180 publications).

In this second time range the top 5 cities that published the most publications between 2010 and 2020 were gathered. According to Figure 3.42 (see Appendix III), the top 5 cities are: Beijing (12 794 publications), Shanghai (5 609 publications), Guangzhou (5 305 publications), Nanjing (5 238 publications), and Bambari (4 959 publications). In comparison to the previous analysis, there are no new cities in this ranking, even though some of them are now represented in different ranks.

In terms of the language of the publications between 1786 and 2020, the top five, based on Figure 3.43 (Appendix III), was the following:

1) English (475 710 publications).
2) Mandarin (5 684 publications).
3) Russian (4 324 publications).
4) German (3 557 publications).
5) French (2 587 publications).

Regarding to the language of the publications between 2010 and 2020, the top 5 was: 1) English (265 317 publications); 2) Mandarin (1 534 publications); 3) Russian (559 publications); 4) Spanish (533 publications); and 5) French (388 publications) (see Appendix III, Figure 3.44). When compared to the previous time range, there is only one difference in this top 5, which is related to the substitution of the German by the Spanish language, both occupying the same ranking (4th place). However, it is

possible to conclude that most publications are written in English, as well as in both time ranges. This is because English has long been considered the lingua franca, especially within the scientific community, since the early industrial revolution, in the mid-18th century (Kamadjeu, 2019).

1) Europe PMC data vs Elsevier data

Considering the interest of this study in obtain the greatest amount of relevant information from different sources, the ScienceDirect platform was also considered and investigated as a potential platform to extract data for further analysis within this system study. Nonetheless, while Europe PMC is an open access repository containing millions of research papers that allows querying documents with specific field attributes (or filtering), therefore, it is possible to make search query match to specific needs, which is determining on how many documents are there with any of the relevant terms in either the title or the abstract. On the other hand, Elsevier does not seem to handle well multifield query specification for article results on SEARCH API. Scarce information on the website and delay in online support were equally evident. That is why in the present study there were only made comparisons of queries against appearances of any of the relevant terms in the title and abstract. In a first search by the Europe PMC website, the executed queries were the following ones, unlike Elsevier's website, that was impossible to determine the following queries to obtain compliant results.

Europe PMC specific results:

TITLE:fungi OR ABSTRACT:fungi OR TITLE:fungus OR ABSTRACT:fungus OR TITLE:fungal OR ABSTRACT:fungal → 251 330 publications

TITLE:fungi OR ABSTRACT:fungi → 108 081 publications

TITLE:fungus OR ABSTRACT:fungus → 71 724 publications

TITLE:fungal OR ABSTRACT:fungal → 137 248 publications

Table 3.6 presents in more detail a comparison between the Europe PMC and the Elsevier platforms, namely in terms of the queries obtained and, how as it is possible to observe, a difference between the number of publications between the two platforms is visible, reaching a difference of around 80% in the title and 95% in the abstract using the keywords of interest “fungi”, “fungal”, and “fungus”, highlighting the Europe PMC platform as a practical repository of scientific papers. Moreover, the Europe PMC platform is equipped with the most available and organized information regarding research help on the website, as well as faster and more objective researcher support.

For a better understanding of the results obtained by the Europe PMC platform, and the reason of the sum of articles from, for example, the "fungi" query in the title added to the "fungus" query in the title is higher (by 140 publications) than the single "OR" query is that there are (140) articles

which contain both words in the title. These 140 publications will appear in the results list of each separate query but are not duplicated in the "OR" query. Reason why the following Venn diagram (Figure 3.45) show article titles containing "fungi" (26 362) on one side, "fungus" (20 093) on the other, and an intersection (140) of the two, and so on. The potential equivalent query to get the keywords intersection could be "Title:(fungi AND fungus)", nevertheless, even in this intersection challenges were faced, since the "AND" query aggregate the selected words, that is, only the articles that has both 2 words in the title will be visible. If only one of the keywords ("fungi" or "fungus") were included in the title, it would not be considered a publication of interest.

Table 3.6. Europe PMC and Elsevier platforms comparison regarding to the different queries between 1786 and 2020.

	Platform	
Queries	**Europe PMC (number of publications)**	**Elsevier (number of publications)**
Title (fungi)	26 362	9 430
Title (fungus)	20 093	4 098
Title (fungal)	34 095	8 693
Title (fungi OR fungus)	46 315	11 382
Title (fungal OR fungus)	53 940	10 907
Title (fungi OR fungal)	60 016	14 397
Title (fungi OR fungus OR fungal)	79 722	15 861
Abstract (fungi)	97 482	4 509
Abstract (fungus)	63 354	1 901
Abstract (fungal)	125 762	4 297
Abstract (fungi OR fungal OR fungus)	228 275	7 229

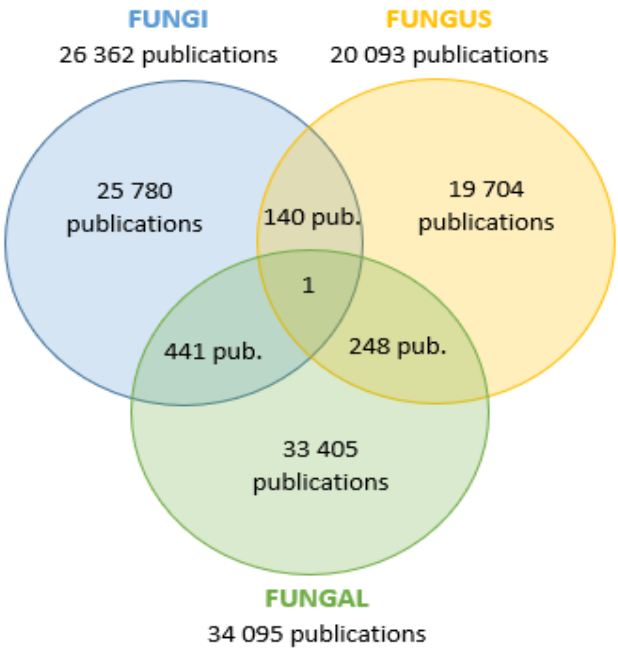

Figure 3.45. Venn diagram representing the logical relation between Europe PMC sets in the Title.

2) Elsevier data extraction using R programming language

Considering the difficulty in extracting data using the C# programming language, an alternative programming language was tested, namely R programming language. Therefore, to compile relevant information between 2010 and 2020, the 10 most significant words (Top10) were graphically represented, aiming for an easily visual illustration and consequent perception, which emphasizes that since 2010 the major topics that have been extensively studied are the following ones: “marine”; “fungi”; “compounds”; “terrestrial”, “activity”; “bioactive”; “isolated”, “potential”, “fungal”, and “natural” as it is demonstrated in Figure 3.46 (see Appendix III). In turn, for a better visualization of the word frequency, a word cloud was developed (Figure 3.47), representing the set of words most commons, being essentially based on the most frequent words that were treated in the extracted abstracts.

Regarding to the analysis of the count words per year, the raw data obtained results demonstrate that, and regardless of the year, the most frequent themes remained to be associated to the topics “marine” and “fungi”, with the research about these concepts increasing exponentially over time (see Appendix III, Figures 3.48).

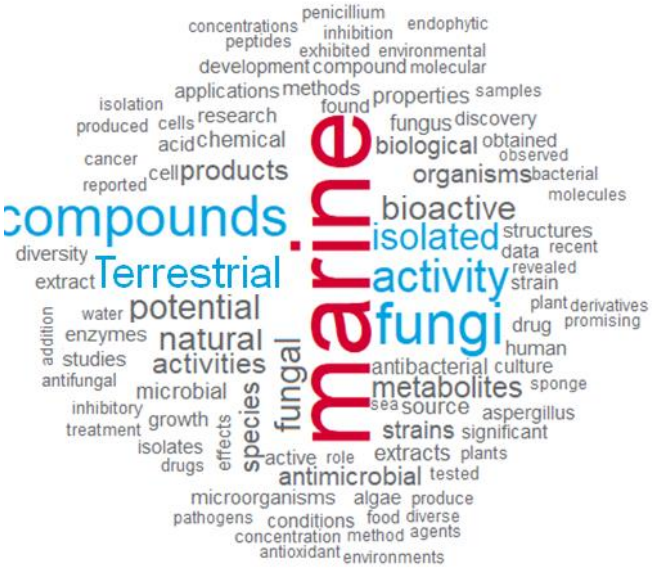

Figure 3.47. Word Cloud representation between 2010 and 2020 (the hundred most frequent words used in in scientific articles).

Furthermore, another item of interest would be to know which are the top 10 journals and analyse which words are the most highlighted and, consequently searched as well. By observing the obtained results, for the most relevant journals collected and summarized, the most investigated themes continue to be the topics regarding “marine” and “fungi”, even though a greater word variability is more visible (see Appendix III, Figure 3.49).

It should be borne in mind that given the very low number of data available, the analysis of extracted scientific articles and consequent treatment of results may be influenced. After realizing the number of scientific articles that were extracted mentioning keywords, it was observed that many times when the extraction process fails, it is impossible to develop a correct analysis regarding how many studies were already published using a specific topic. Nevertheless, the few numbers of collected articles highlighted the scarcity of the scientific information available online in the selected database. Thus, as expected of an innovative research work, the results demonstrated that there is still little investigation regarding the present thesis’ subject.

3.4 Conclusions

The main goal of this study was to assess the current landscape of fungi-derived compounds research, both marine and terrestrial, by using a data mining tool. Hence, the aim was to understand if there are several studies within this specific area of expertise or if, on the other hand, there is a lack of information and data. Data mining tool, more specifically text mining tool, using ETL process,

was the selected technique due to its efficiency when extracting patterns, relevant information, or knowledge from several sources.

Despite posing some challenges, the ETL method was beneficial to the present investigation, generating insightful results that comprehend a vast timeline (1786-2020). As a matter of fact, this tool allowed the characterization of the research that has been conducted throughout the years, namely in terms of its global geographic distribution, the most relevant terms, the number of publications per year, the languages of the publications, the publications' format, their authors, cities, countries, and document type, as well as the journal where they were published, and the main used keywords (ultimately represented in a word cloud).

Overall, the results demonstrate that most countries have been developing research about fungi compounds, both marine and terrestrial, throughout the years, even though there are some countries that show truly little research about the topic. During the past 10 years, the number of publications about marine and terrestrial bioactive compounds has been increasing substantially, even though there are more publications about marine bioactive compounds in comparison to those about terrestrial bioactive compounds. Still, it is quite outstanding the fact that every single country presents research about this subject. The terms that seem to be the most relevant in the developed research are directly related to the top keywords, including a list of 35 concepts in total and comprehending those same top keywords.

Considering the number of publications per year, it is possible to conclude that it started quite low, with only 2 publications in 1786. However, throughout the years publications started to increase, reaching a total of 190 in 1946. The peak point was definitely in 2019, when there were around 33 330 publications published during that year. Most of these publications was written in English, even though there were some written in other languages, such as Mandarin, Russian, German, French, and Spanish, and now the most common publication format is printed-electronic, which demonstrates an evolution from previous years, when the most common format was printed. This might be due to the constant and increasing implementation of technology within academic fields, with technology providing a crucial resource in several investigations.

In terms of the countries with more publications, these are almost the same throughout the years, with a clear emphasis on the United States of America, China, and United Kingdom. The existing research is mostly composed by journal articles, being frequently published in well-known scientific journals, such as the *PloS one*, *Scientific Reports, The Journal of Biological Chemistry*, and *Frontiers in Microbiology*. Within these specific publications, the most commonly used keywords are the following ones: "fungi", "candida albicans", "inflammation", "bacteria", "infection", "yeast", "antimicrobial activity", "microbiome", "aspergillus", and "antifungal activity". Even though the number of publications has been increasing since 1786, when the first publication about the subject was published, the obtained results demonstrate that more research is needed in this specific field of expertise, since there is truly little research based on marine and terrestrial fungi with the bioactivity of interest in the present investigation.

In addition, it was noted that there are several areas where the extraction algorithm can still be improved to extract higher quality metadata or extended to capture more information hidden in unstructured documents. The algorithm currently processes only born-digital document, in which the text is present as PDF stream rather than the images of scanned pages. The documents containing scanned pages are not properly processed, which concerns older resources in particular.

In the future, plan to successfully extend the range of information the algorithm is able to extract, it would be attractive, that is, not only to collect keywords from the title and abstract, but also from the body of the scientific article.

Last, but not least, in documents of various domains there are a lot of useful and important information present in the text of the document, but not expressed directly, such as the reasons the document cites other documents, the methods used in the paper, the problem the paper addresses, or the experiment results. In the future, it would be interesting to experiment with machine learning and natural language processing techniques in order to acquire a deeper understanding of the text of the input scientific publication.

Equally interesting to address would be to understand the impact of the use of these programming tools on the advancement of the scientific community. Today's advanced scientific knowledge is the cumulative effect of previous research work, each research paper reference or citation to the earlier research article. In another way, each current scientific publication is an extension of the previous research topic. Thus, the intellectual links are established between past and current research topics. Knowledge is interdisciplinary, multifaceted, multidimensional and multidirectional in nature, growing at a faster pace. It is essential to classify the knowledge to understand the growth and evolution in a better way. There are several potential methods developed over time with the purpose of classifying knowledge. Data mining analysis is one of the most effective and efficient methods of applying knowledge. Many attempts have been made in the past manually to understand the structure of knowledge in a limited extent. Advances in computing and information technology and the development of well-structured databases have largely paved the way for new tools applications.

3.5 References

Abilhoa, W. D., & Castro, L. N. (2014). A keyword extraction method from twitter messages represented as graphs. *Applied Mathematics and Computation,* 240, 308-325.

Allahyari, M., Trippe, E. D., & Gutierrez, J. B. (2017). A Brief Survey of Text Mining: Classification, Clustering and Extraction Techniques. ArXiv:1707.02268 [Cs].

Ananiadou, S., Kell, D. B., & Tsujii, J.-I. (2006). Text mining and its potential applications in systems biology. *Trends in Biotechnology*, 24, 571–579.

Baba, S. W., & Kumar, R. S. (2016). Data Mining: Text Classification System for Classifying Abstracts of Research Papers. *International Journal of Advance Research, Ideas and Innovations in Technology*, 2, 1-6.

Barateiro, J., & Galhardas, H. (2005). A survey of data quality tools. *Datenbank Spektrum,* 14, 15-21.

Beliga, S., Meštrović, A., & Martinčić-Ipšić, S. (2015). An overview of graph-based keyword extraction methods and approaches. *Journal of Information and Organizational Sciences,* 39, 1-20.

Berg, J. M., Bhalla, N., Bourne, P. E., Chalfie, M., Drubin, D. G., Fraser, J. S., Greider, C. W., Hendricks, M., Jones, C., Kiley, R., King, S., Kirschner, M. W., Krumholz, H. M., Lehmann, R., Leptin, M., Pulverer, B., Rosenzweig, B., Spiro, J. E., Stebbins, M., Strasser, C., Swaminathan, S., Turner, P., Vale, R. D., VijayRaghavan, K., & Wolberger, C. (2016). Preprints for the life sciences. The time is right for biologists to post their research findings onto preprint servers. *Science*, 352, 899-901.

Berry, M. W., & Kogan, J. (2010). *Text mining: Applications and theory*. UK: Wiley.

Boulis, C., & Ostendorf, M. (2005). Text Classification by Augmenting the Bag-of-Words Representation with Redundancy-Compensated Bigrams. In *Proceedings of the SIAM International Conference on Data Mining at the Workshop on Feature Selection in Data Mining*.

Buse, R. P. L., & Weimer, W. (2012). Synthesizing API Usage Examples. IEEE Xplore: 34th International Conference on Software Engineering (ICSE), 782-792.

Carroll, A. R., Copp, B. R., Davis, R. A., Keyzers, R. A., & Prinsep, M. R. (2021). Marine natural products. *Natural Product Reports*, 38, 362-413.

Cui, W., Wu, Y., Liu, S., Wei, F., Zhou, M., & Qu, H. (2010). Context-preserving, dynamic word cloud visualization. *IEEE Computer Graphics & Applications,* 30, 42-53.

Dang, S., & Ahmad, P. H. (2015). A review of text mining techniques associated with various application areas. *International Journal of Science and Research,* 4, 2461-2466.

DePaolo, C. A., & Wilkinson, K. (2014). Get your head into the clouds: using word clouds for analyzing qualitative assessment data. *TechTrends*, 58, 38–44.

Elasticsearch (2021a). *Elastic.* Available at https://www.elastic.co/pt/. Consulted on February, 2021.

Elasticsearch (2021b). Text field type. *Elastic.* Available at https://www.elastic.co/guide/en/elasticsearch/reference/current/text.html. Consulted on February, 2021.

Elasticsearch (2021c). Filter aggregation usage. *Elastic.* Available at https://www.elastic.co/guide/en/elasticsearch/client/net-api/current/filter-aggregation-usage.html. Consulted on February, 2021.

Elasticsearch (2021d). Terms aggregation usage. *Elastic.* Available at https://www.elastic.co/guide/en/elasticsearch/client/net-api/current/terms-aggregation-usage.html. Consulted on February, 2021.

Elasticsearch (2021e). Testing. *Elastic.* Available at https://www.elastic.co/guide/en/elasticsearch/client/javascript-api/7.x/client-testing.html. Consulted on February, 2021.

Elasticsearch (2021f). Stop token filter. Elastic. Available a https://www.elastic.co/guide/en/elasticsearch/reference/current/analysis-stop-tokenfilter.html#analysis-stop-tokenfilter-stop-words-by-lang. Consulted on February, 2021.

Europe PMC (2020). *EBI Europe PMC SOAP Web Service 6.5.0: Reference guide.* Available at http://europepmc.org/docs/EBI_Europe_PMC_Web_Service_Reference.pdf. Consulted on February, 2021.

Europe PMC (2021). *Articles RESTful API*. Available at http://europepmc.org/RestfulWebService. Consulted on February, 2021.

Fan, W., Wallace, L., Rich, S., & Zhang, Z. (2006). Tapping the power of text mining. *Communications of the ACM*, 49, 77-82.

Farooqui, N. A., & Mehra, R. (2018). Design of a data warehouse for medical information system using data mining techniques. *5th IEEE International Conference on Parallel, Distributed and Grid Computing*, 199-203.

Feldman, R., & Sanger, J. (2007). The text mining handbook: Advanced approaches in analyzing unstructured data. New York: Cambridge University Press.

Feldman, R., & Sanger, J. (2006). The Text Mining Handbook. Cambridge University Press.

Feinerer, I., Hornik, K., & Meyer, D. (2008). Text mining infrastructure in R. *Journal of Statistical Software*, 25, 1-54.

Gaikwad, S. V., Chaugule, A., & Patil, P. (2014). Text mining methods and techniques. *International Journal of Computer Applications*, 85.

Gilbert, K., Izquierdo, J., Holmes, G., Athanasiadis, I., Comas, J., & Sànchez-Marrè, M. (2008). On the role of pre and post-processing in environmental data mining. *iEMSs 2008: International Congress on Environmental Modelling and Software Integrating Sciences and Information Technology for Environmental Assessment and Decision Making*.

Ginsparg, P. (2016). Preprint déjà vu. *The EMBO Journal*, 35, 2620-2625.

Girija, N., & Srivatsa, S. K. (2006). A research study: Using data mining in knowledge base business strategies. *Information Technology Journal,* 5, 590-600.

Gladfelter, A. S., James, T. Y., & Amend, A. S. (2019). Marine fungi. *Current Biology*, 29, R191-R195.

González-Albo, B., & Bordons, M. (2011). Articles vs. proceedings papers: Do they differ in research relevance and impact? A case study in the Library and Information Science field. *Journal of Informetrics*, 5, 369-381.

Gupta, V., & Lehal, G. S. (2009). A survey of text mining techniques and applications. *Journal of Emerging Technologies in Web Intelligence*, 1, 60–76.

Gupta, S., Kaiser, G. E., Grimm, P., Chiang, M. F., & Starren, J. (2005). Automating content extraction of html documents. *World Wide Web*, 8, 179–224.

Hamm, S. E. (2011). Using word clouds for reflection and discussion in an online class. *Teaching Theology & Religion,* 14, 156-156.

Han, J., Kamber, M., & Pei, J. (2011). Data mining: Concepts and techniques (3rd ed.). Amsterdam: Elsevier.

Hassani, H., Huang, X., Silva, E. S., & Ghodsi, M. (2016). A review of data mining applications in crime. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, 9, 139–154.

Hoy, M. B. (2020). Rise of the Rxivs: How preprint servers are changing the publishing process. *Medical Reference Services Quaternaly*, 39: 84-89.

Huang, A. (2008). Similarity measures for text document clustering. In: *Proceedings of the sixth New Zealand Computer Science Research Student Conference (NZCSRSC2008)*, Christchurch, New Zealand, pp. 49–56.

Jayashankar, S., & Sridaran, R. (2017). Superlative model using word cloud for short answers evaluation in eLearning. *Education and Information Technologies*, 22, 2383-2402.

Jiang, J. (2013). Information extraction from text. In C. C. Aggarwal & C. X. Zhai (Eds.), *Mining text data*, London: Springer, pp. 11-42.

Jivani, A. G. (2011). A comparative study of stemming algorithms. *International Journal of Computer Applications in Technology,* 2, 1930-1938.

Jo, T. (2019). Text mining: Concepts, implementation, and big data challenge. Switzerland: Springer.

Jones, E. B. G., Pang, K.-L., Abdel-Wahab, M. A., Scholz, B., Hyde, K. D., Boekhout, T., Ebel, R., Rateb, M. E., Henderson, L., Sakayaroj, J., Suetrong, S., Dayarathne, M. C., Kumar, V., Raghukumar, S., Sridhar, K. R., Bahkali, A. H. A., Gleason, F. H., & Norphanphoun, C. (2019). An online resource for marine fungi. *Fungal Diversity*, 96, 347–433.

Jones, K. S. (1972). A statistical interpretation of term specificity and its application in retrieval. *Journal of Documentation*, 28, 11-21.

Jonnalagadda, S., & Topham, P. (2010). NEMO: Extraction and normalization of organization names from PubMed affiliation strings. *Journal of Biomedical Discovery and Collaboration*, 5, 50-75.

Justicia de la Torre, C., Sánchez, D., Blanco, I., & Martín-Bautista, M. J. (2018). Text mining: Techniques, applications, and challenges. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems,* 26, 553-582.

Kamadjeu, R. (2019). English: the lingua franca of scientific research. Correspondence, 7, E1174.

Kimball, R., & Caserta, J. (2004). *The data warehouse ETL toolkit: Practical techniques for extracting, cleaning, conforming, and delivering data.* New Jersey: Wiley.

Khattab, A. R., & Farag, M. A. (2021). Marine and terrestrial endophytic fungi: a mine of bioactive xanthone compounds, recent progress, limitations, and novel applications. Critical Reviews in Biotechnology, 1-28.

Kobayashi, V. B., Mol, S. T., Berkers, H. A., Kismihók, G., & Den Hartog, D. N. (2018). Text mining in organizational research. *Organizational Research Methods,* 21(3), 733-765.

Kononenko, I., & Kukar, M. (2007). *Machine learning and data mining.* Amsterdam: Elsevier.

Krassmann, A. L., Herpich, F., Bercht, M., & Cazella, S. C. (2017). Analyzing trends in academic papers about ubiquitous virtual worlds in education using text mining. *International Journal for Innovation Education and Research*, 5, 157-10.

Kroeze, J. H., Matthee, M. C., & Bothma, T. J. D. (2003). Differentiating Data- and Text-Mining Terminology. *Proceedings of the 2003 annual research conference of the South African institute of computer scientists and information technologists on Enablement through technology*, 93-101.

Maalej, W., & Robillard, M. P. (2013). Patterns of Knowledge in API Reference Documentation. In *IEEE Transactions on Software Engineering*, 39, 1264-1282.

MacCallum, C. J. (2006). ONE for all: the next step for PLoS. *PLoS biology*, 4, 1875-1876.

Manku, G. S., Jain, A., & Sarma, A. D. (2007). Detecting near-duplicates for web crawling. In WWW '07: Proceedings of the 16th international conference on World Wide Web, 141-150.

Markov, Z., & Larose, D. T. (2007). Data mining the web: Uncovering patterns in web content, structure, and usage. John Wiley &Sons, Inc., Hoboken, New Jersey.

Maslove, D. M. (2018). Medical preprints: a debate worth having. Journal of the American Medical Association (JAMA), 319, 443-444.

Mathen, M. P. (2010). Data warehouse testing. *Infosys Developer IQ Magazine,* 1-8.

Mayer, A. M. S., Guerrero, A. J., Rodríguez, A. D., Taglialatela-Scafati, O., Nakamura, F., & Fusetani, N. (2021). Marine Pharmacology in 2016-2017: Marine Compounds with Antibacterial, Antidiabetic, Antifungal, Anti-Inflammatory, Antiprotozoal, Antituberculosis and Antiviral Activities; Affecting the Immune and Nervous Systems, and Other Miscellaneous Mechanisms of Action. *Marine Drugs*, 19, 1-75.

McLachlan, G. J., & Krishnan, T. (2007). The EM Algorithm and Extensions. John Wiley & Sons, Inc.

Microsoft (2021). Text Analytics API documentation. *Microsoft*. Available at https://docs.microsoft.com/en-us/azure/cognitive-services/text-analytics/. Consulted on February, 2021.

Miner, G., Elder, J., Hill, T., Nisbet, R., Delen, D., & Fast, A. (2012). Basic Text Mining Principles. *Practical Text Mining and Statistical Analysis for Non-structured Text Data Applications* (pp.3-121). Elsevier Inc.

Mukhopadhyay, D. (Ed.) (2019). Web searching and mining. Singapore: Springer.

Narayana, B. L., & Kumar, S. P. (2015). A new clustering technique on text in sentence for text mining. *International Journal of Science Engineering and Advance Technology*, 3, 69–71.

Navathe, S. B., & Elmasri, R. (2000). Data warehousing and data mining. In Fundamentals of Database Systems, *Pearson Education pvt Inc*. Singapore, 841-872.

Nenkova, A., & McKeown, K. (2012). A survey of text summarization techniques. In C. C. Aggarwal & C. X. Zhai (Eds.), *Mining text data* (pp.43-76). London: Springer.

Nie, B., & Sun, S. (2017). Using text mining techniques to identify research trends: A case study of design research. *Applied Sciences*, 7, 1-21.

Padhy, N., Mishra, D. P., & Panigrahi, R. (2012). The survey of data mining applications and feature scope. *International Journal of Computer Science, Engineering and Information Technology* (IJCSEIT), 2, 43-58.

Peyrat, L.-A., Tsafantakis, N., Georgousaki, K., Ouazzani, J., Genilloud, O., Trougakos, I. P., & Fokialakis, N. (2019). Terrestrial microorganisms: Cell factories of bioactive molecules with skin protecting applications. *Molecules*, 24, 1-35.

Ramasubramanian, C., & Ramya, R. (2013). Effective pre-processing activities in text mining using improved Porter's stemming algorithm. *International Journal of Advanced Research in Computer and Communication Engineering,* 2, 4536-4538.

Rai, M., Gade, A., Zimowska, B., Ingle, A. P., & Ingle, P. (2018). Marine-derived *Phoma*—the gold mine of bioactive compounds. *Applied Microbiology and Biotechnology*, 102, 9053–9066.

Reddy, C. L., & Venkatadri, M. (2011). A Review on data mining from past to the future. *International Journal of Computer Applications*, 15, 19-22.

Salloum, S. A., Al-Emran, M., Monem, A. A., & Shaalan, K. (2018). Using text mining techniques for extracting information from research articles. In K. Shaalan, A. E. Hassanien, & F. Tolba (Eds.), *Intelligent natural language processing: Trends and applications. Studies in Computational Intelligence*. New York: Springer International Publishing, pp. 373-397.

Salloum, S. A., Al-Emran, M., Monem, A. A., & Shaalan, K. (2017). A survey of text mining in social media: facebook and twitter perspectives. *Advances in Science, Technology and Engineering Systems Journal*, 2, 127-133.

Sarica, S. & Luo, J. (2021). Stopwords in technical language processing. *PLoS ONE*, 16, 1-13.

ScienceDirect (2021). Faceted Search. *ScienceDirect.* Available at https://www.sciencedirect.com/topics/computer-science/faceted-search. Consulted on February, 2021.

Sebastiani, F. (2002). Machine Learning in Automated Text Categorization, 34, 1–47.

Sharma, D. (2012). Stemming algorithms: A comparative study and their analysis. *International Journal of Applied Information Systems*, 4, 7-12.

Siguenza-Guzman, L. Saquicela, V., Avila-Ordóñez, E., Vandewalle, J., & Cattrysse, D. (2015). Literature review of data mining applications in academic libraries. *The Journal of Academic Librarianship*, 41, 499-510.

Sinclair, J., & Cardew-Hall, M. (2008). The folksonomy tag cloud: when is it useful? *Journal of Information Science*, 34, 15–29.

Sikos, L. F. (2015). Mastering structured data on the semantic web: From HTML5 microdata to linked open data. Apress, New York.

Song, M., & Kim, S. Y. (2013). Detecting the knowledge structure of bioinformatics by mining full-text collections. *Scientometrics,* 96, 183-201.

Sulova, S., & Nacheva, R. (2017). Using Text Mining to Classify Research Papers. *17th International Multidisciplinary Scientific GeoConference SGEM 2017*, 17, 647-654.

Sullivan, D. (2001). Document warehousing and text mining. Techniques for improving business operations, marketing, and sales. John Wiley & Sons, Inc. New York, NY, USA ©2001

Sumathi, S., & Sivanandam, S. N. (2006). Introduction to data mining and its applications. Berlin: Springer.

Sun, S., Luo, C., Chen, J. (2017). A review of natural language processing techniques for opinion mining systems, *Information Fusion*, 36, 10-25.

Talib, R., Hanif, M. K., Ayesha, S., & Fatima, F. (2016). Text Mining: Techniques, applications and issues. *International Journal of Advanced Computer Science and Applications* (IJACSA), 7, 414-418.

Tang, J., Alelyani, S., & Liu, H. (2014). Feature selection for classification: A Review. In *Data Classification: Algorithms and Applications* (pp.1-25).

Tromp, E. & Pechenizkiy, M. (2011). Graph-based N-gram language Identification on short texts. *Proceedings of the 20th Machine Learning conference of Belgium and The Netherlands.*

Vale, R. D. (2015). Accelerating scientific publication in biology. *Proceedings of the National Academy of Sciences* (PNAS), 112, 13439-13446.

Viegas, F. B., Wattenberg, M., Van Ham, F., Kriss, J., & McKeon, M. (2007). Manyeyes: a site for visualization at internet scale. *IEEE Transactions on Visualization and Computer Graphics*, 13, 1121–1128.

Wagstaf, K., Cardie, C., Rogers, S., & Schroedl, S. (2001). Constrained K-means clustering with background knowledge. In *Eighteenth International Conference on Machine Learning* (pp.1-8).

Web of Science (2021). Using the proceedings paper document type. Retrieve February 20, 2021. In: http://isiwebofknowledge.com/productstools/multidisciplinary/webofscience/cpci/usingproceedings/

Chapter 4

BIOACTIVE EXTRACTS FROM FUNGI

4.1 Introduction

4.1.1 Environmentally friendly extraction techniques, such as High-Pressure Extraction (HPE) to obtain bioactive extracts

High pressure processing is an emerging non-thermal technique that has been showing great promise in food and pharmaceutical industries and in biotechnological research. It is an environmentally friendly technique, and it was recently developed for extracting bioactive compounds from natural products (Aktas & Yildiz, 2011; Jun, 2013; Lebovka et al., 2011). Usually, the extraction process is based on two main phases. First, the target material for extraction is immersed in a solvent to be swelled and hydrated, and then the soluble elements move to the extraction solvent through mass transfer by diffusion and permeation (Mustafa & Turner, 2011; Sadus, 2012). Compared to traditionally applied methods, such as heat or solvent extraction, the HPE method is faster, generates greater extraction gains, lower quantities of impurities and, as it happens at room temperature, prevents thermal degradation and also bioactivity loss of the extracted elements.

HPE began to be explored initially in 1914 as a developing non-thermal technique that could achieve the same degree of food safety as heat pasteurization, as well as inactivate pathogens and enzymes with insignificant structural changes with no impact on nutritional characteristics and food quality (Pereira & Vicente, 2010; Rastogi et al., 2007). Nevertheless, HPE was only considered as a true extraction technique in 2004 and, it was found to be effectively short regarding to the extract time consuming and a very efficient technique (Shouqin et al., 2004). Shouqin et al. (2004) described some preliminary studies to demonstrate the HPE application capacity. The authors observed that diverse HPE conditions produced herbal extracts with different colors, and the one extracted with the lowest quantity of impurities was reached at 600 MPa for five minutes, using water as solvent (Shouqin et al., 2004).

HPE can increase the rate of mass transfer by adjusting the level of concentration and diffusion capacity, consequently causing damage to cell membranes, and intensifying the solvent permeability into the cells, leading to a shorter time process, improving process safety, reducing costs, and achieving better yield with respect to compounds (Xi et al., 2011a). To obtain a gain in cell membrane permeability, the process of choosing the appropriate extraction methodology, which has a high capacity to increase the mass transfer during the extraction technique, whether of animal origin or vegetable is too relevant. Furthermore, the chosen methodology also needs to be environmentally safe and energy efficient (Corrales et al., 2008). Therefore, HPE can be considered as a new technology that uses high pressure conditions to extract active elements from various raw material and effectively improve the low extraction levels by increasing the quality and concentration of active elements achieved through conventional extraction methods (Wang et al., 2016).

Over the years, HPE has been applied to extract a several bioactive compounds with unique molecules, from a variety of food products, such as vegetables, fruits and plants (Table 4.1) (Huang et al., 2013a; Moreira et al., 2019). HPE has already been successfully used to extract flavonoids

from propolis (Shouqin et al., 2005); lycopene from tomato paste (Briones-Labarca et al., 2019; Xi, 2006a, 2006b); phenolics from sour cherry pomace (Adil et al., 2008) and fermented fig (Alexandre et al., 2017); caffeine and polyphenols from green tea leaf (Jun, 2009; Jun et al., 2009; Xi, 2009a, 2009b; Xi et al., 2010; Xi et al., 2011b; Xi et al., 2013); corilagin and phenolics from longan fruits pericarp (Prasad et al., 2009a, 2009c, 2009d, 2010a, 2010b); ginsenosides from Panax ginseng (Chen et al., 2009; Lee et al., 2011; Shin et al., 2010; Souqin et al., 2006, 2007b); pectin from orange peel (Guo et al., 2012); polysaccharides from the Ligusticum rhizome (Liu et al., 2015); rutin from Amaranth leaf (Kraujalis et al., 2015) and anthocyanins from black chokeberry (Grunovaitė et al., 2016).

In addition, HPE shows a higher level of efficiency to extract active compounds, which allows the safe use of solvents even if they have low selectivity (Bi et al., 2009; Huang et al., 2013a; Pais et al., 2019). Research, such as those developed by Corrales et al. (2008, 2009) and Liao et al. (2012), have demonstrated numerous gains from HPE technology, such as shorter process times, lower energy consumption, higher extraction yield, and lower volume of impurities present in the liquid extracted compared to other methodologies for extraction (Corrales et al., 2008, 2009; Liao et al., 2012).

The HPE technology covers the most important phases of the extraction process. They are the pressure boost phase, which includes mixing materials and solvent; the very short time space until reaching the desired pressure inside the pot and the resulting balance between the interior and exterior of the cells; the pressure maintaining phase, which consists of high pressure treatment for a certain period of time; and finally, the pressure relief phase with fast release of target pressure to atmospheric pressure in just a few seconds, causing the cell to expand and the fluid to circulate, resulting in significant damage to the cells and their membranes, generating greater permeability and consequently, the substances of interest for extraction are concentrated and purified (Huang et al., 2013b; Shouqin et al., 2005). The extraction process during the HPE is greatly impacted by numerous parameters, such as solvent type and volume, temperature, pressure, time to carry out the process, number of cycles, among others (Chen et al., 2009). Thus, the HPE extraction method works at low temperatures (generally not exceeding 60 ºC) and high pressures (generally 100-1 000 MPa), to extract the compounds in an agile and fast way, which requires reduced volumes of organic solvents and, at the same time, offers a good process performance (Shouqin et al., 2005).

In addition to the isostatic principle, HPE follows the *principle of Le Chattelier*, who stated that when a system is disturbed, it can respond by modifying the temperature or concentration restoring its equilibrium (Khan et al, 2018). In other words, high pressure is used in order to trigger numerous events, such as the phase transition from one configuration to another, modification in the dynamics of the reaction, structural alteration in the molecule, among others, which promotes a response, resulting in greater efficiency in extraction (Xi et al., 2011b). Thus, the high pressure intensifies the breaking of the ionic connections due to the electrostriction forces that act on water. As a result, high pressure can cause changes in the structure of the cell molecule (Yang et al., 2009), such as lipid proteins, enzymes and outer cell membranes, which cause damage to the wall of cells and cellular

structure at the internal level, decreasing the potential for resistance to mass transfer within the cell, therefore, the bioactive members are excited, leaving another elements such as, for example, vitamins, flavonoids, saponin, alkaloids, pigments, fragrance components, among other compounds not impacted by the method (Hall, 2015; Knorr et al., 2011; Linton & Patterson, 2000).

Shouqin et al. (2005) verified that the impact of high pressure on the structure of the lycopene and flavonoid molecule is irrelevant, since it has no effect on small molecules of plants. Thus, it was noted that pressures above 100 MPa are sufficient to disrupt plant cell membranes (Shouqin et al, 2005), as for example intracellular vacuoles in onions (Butz et al., 1994). As a consequence of the degradation process, the compounds of a chemical nature existing inside the cell are quickly released in the extraction solvent until the equilibrium is reached. The results of the extraction of bioactive compounds increase as the pressure increases, which is achieved in a process of multi-stages. Generally, most natural phenomena of bioactive compounds demonstrate an increase in solubility potential due to increased pressure employed, or with a higher density solvent (Khan et al., 2018; Xi., 2009a). Usually, the pressure used in HPE, as already mentioned, is the range between 100 to 1,000 MPa, while the highest pressure achieved in other methods, like the SFE method is approximately 100 MPa and in supercritical carbon dioxide, it is approximately 10 MPa. Thus, the higher pressure employed by the HPE method is significantly higher compared to other extraction methods. Therefore, more essential components can be extracted by the HPE technique, as noted due to its higher achievable fraction, which is understood as the mass fraction of essential components in solution over the mass of raw material (Khan et al., 2018).

Like pressure, temperature plays an important role in determining the extraction efficiency, as higher temperature values lead to greater diffusivity of the solvent towards components within the matrix, thus raising the extraction level. Furthermore, there is an increase in the rate of the release process of active local components in the materials (Prasad et al., 2009e). Thus, the general temperature range needed to extract bioactive compounds is between 20 ºC and 60 ºC in the HPE method (Shin et al., 2010) and the higher temperature range results in greater extraction efficiency due to the increase in the rate of diffusion and solubility of analytes in solvents (Ju & Howard, 2003). Nevertheless, there are substances sensitive to the temperature that need to be efficiently extracted. An increase in temperature can break phenolic matrix connections and structurally impact the plant cell membrane, making it less selective and coagulating lipoproteins (Prasad et al., 2009e). In the HPE technique, the total phenolic extraction with greater efficiency of Pinot Noir, a type of grape skin, was achieved in the research by Prasad et al. (2009b), in the higher temperature range when increased from 30 ºC to 50 ºC under a higher-pressure impact (500 MPa). Furthermore, the rise in temperature is followed by a reduction in the viscosity of the solvents in liquid form, which facilitates their entry into the matrix and results in more efficient extraction. Thus, at higher temperatures, the surface tension that acts on solutes, solvent and matrix is also reduced, which allows a more efficient infusion of the solvent into the matrix. However, each compound degrades at a certain temperature, thus restricting the maximum rise in temperature (Khan et al., 2018). The performance with respect to the results of the HPE technique is also much defined according to the retention time of the

extraction and the number of cycles involved in it, which can be considered as associated parameters. Thus, a longer retention time ensures that the solute and solvent are in total contact for a sufficient period of time, generating a better extraction gain (Xi, 2009a). Regarding to the matrix, when this comes into contact with the solvent for a longer period of time, it will swell up and, as a consequence, more solvent will permeate the sample pores (Chen et al., 2009). Hence, the number of times that the HPE technique was performed, or the matrix was mixed with the solvent is the number of cycles of the method (Khan et al., 2018).

In order to maximize HPE's results, numerous algorithmic projects have been used, such as a Factorial Central Composition project, and an Orthogonal Matrix's, among others (Chen et al., 2009). The data that is needed for the parameters and requirements regarding the process have been described in several compounds, such as to extract proteins from cedrus male cones pollen (Altuner et al., 2012), pectin from honey pomelo (Guo et al., 2014), phenolic compounds in orange peel (M'hiri et al., 2015) and pomegranate peel (Alexandre et al., 2017b, 2019), carotenoids and betalains compounds from yellow prickly pear peel (Castro et al., 2019), among others. The most relevant parameters that impact the HPE technique, as previously highlighted are the pressure, temperature, time period for extraction, numerical quantitative extraction cycles, as well as the nature, quantity and concentration of the solvent (Xi et al., 2011b). Choosing the appropriate extraction solvent is one of the first steps in the process to optimize the HPE method. The selection of the solvent and its concentration is closely associated with the compounds to be extracted; this must be non-toxic and easily evaporated from the extract reached at the end of the process (Ji et al., 2010; Xi, 2009a). The extraction solvent should have the characteristic of being selectively soluble, that is, only the active target compound should be soluble, while the others should not be simultaneously extracted from the matrix (Ji et al., 2010). Thus, the fraction between solvent and raw material is another relevant indicator to be considered, since the dilution of bioactive elements in the solvent is a physical phenomenon, and if the amount of solvent is sufficiently, there will be a much-increased chance that it makes contact with those compounds targeted for extraction, which will lead to greater gains for the extraction process (Xi, 2009a). The HPE method is robust, ecological, safe and, energy efficient, thus requiring a small volume of solvent (Khan et al., 2018).

According to *Pascal's theory,* during the HPE method, the passage of pressure to the entire material occurs uniformly and instantaneously (Chen et al., 2005). As a result, there is a very short time to reach an equilibrium pressure between the inside and outside of the cells. In such conditions, the greatest extraction gain is reached in a too short time because of the fast solvent diffusion speed, which makes this method much faster than traditional ones (Nieto et al., 2008). Thus, often one to ten minutes of extraction time and one to two cycles are enough to extract bioactive compounds from biological materials (Chen et al., 2005). Guo et al. (2012) noted that the HPE technique was the best and most rapid methodology (10 minutes) to extract pectin from the navel of orange peel compared to traditional heat extraction techniques (60 min) and the micro-waves assisted extraction technique (21 min) (Guo et al., 2012). The reduced treatment time involves more efficiency in the process, as the process time is reduced and, therefore, the production capacity in the process increases. Thus,

the ideal pressure treatment time period of five minutes was considered the most appropriate (Li et al, 2016).

Parameters are defined in advance for target structures in order to maintain the support that guides the equipment for the HPE technique. Thus, precise information regarding the parameters in use, like pressure, temperature and time, should be fixed to optimize the extraction process (Shoqin et al., 2004). The demands for food or pharmaceutical products, having bioactive compounds with nutritional and health gains, are growing, which has led industries to create extraction techniques that demonstrate efficiency. The HPE technique has a prominent place among modern extraction techniques, where high pressure being used for a variable time from some second to few minutes (Khan et al., 2018). Thus, by choosing the best technique among the existing ones, HPE makes a significant difference in the value gain to the food by-product and, as a consequence, will reduce the costs of the product prepared for sale. Considerations in using organic waste as a source of bioactive compounds by the HPE method, will make it easy to improve the financial condition of farmers and business holders and, in particular, an option to manage remains (Khan et al., 2018). One of the viable ways to reduce waste associated with by-products is through their use to extract bioactive compounds. Thus, numerous diverse and high-value chemical compounds, such as polysaccharides, proteins, fibers, among others with flavoring properties, can be extracted from food waste and used later as functional ingredients and nutraceutical-type supplements. Such bioactive compounds have antioxidant properties and can help eliminate radicals; thus, it can contribute to the inhibition or deceleration of oxidative processes with regard to lipids, proteins and DNA (Baiano, 2014; Khan et al., 2018). The HPE technique is applicable for use in a wide variety of food or biological products, including fruit and vegetable, oilseed, cereal bran and husk, seafood, some herb and medicinal plant remains. However, most of the research for HPE has been carried out in the laboratory and attempts should be made to expand the method in order to satisfy the marketing requirements for high quality bioactive substances or extracts (Khan et al., 2018).

Table 4.1. Studies describing bioactive compounds extracted through high-pressure extraction (HPE) from multiple sources and determined bioactivities.

			HPE conditions						Bioactivities and methodologies of analysis				
Source	**Bioactive compounds**	**Pre-treatment**	**Pressure (MPa)**	**Time (min)**	**T (ºC)**	**Ratio Mass to volume (g/mL)**	**Solvent**	**Other extraction methods**	**Antioxidant**	**Antimicrobial**	**Structural changes**	**Other analyses (methods)**	**References**
Tomato puree	Carotenoids (lycopene, β-carotene, total carotenoids)	Crude	400	15	25	-	Water	-	DPPH; Lipid oxidation	-	-	pH; Titratable acidity; Soluble solids; Total solids; Colour; HPLC	Sánchez-Moreno et al. (2004)
Propolis	Flavonoids	Crude	500	1	RT	1/35	75% Ethanol	RT; Heat reflux	-	-	-	Total flavonoids (colorimetry)	Shouqin et al. (2005)
American Ginseng	Ginsenosides	Dried in vacuum	200	2	25	1/50	60% Ethanol	US; MW; SC-CO2; Soxhlet; Heat reflux	-	-	-	HPLC	Shouquin et al. (2006)
Tomato paste	Lycopene	Dried in air drier	500	1	RT	1/5	50% Ethanol	Solid-liquid extraction with sonication	-	-	-	HPLC	Xi (2006a)
Tomato paste	Lycopene	Dried in air drier	500	1	RT	1/6	75% Ethanol	Solid-liquid extraction with sonication	-	-	-	HPLC	Xi (2006b)
Propolis	Flavonoids	Crude	500	1	RT	1/35	75% Ethanol	Leaching at RT; Heat reflux	β-carotene bleaching; DPPH	-	-	Phenolic compounds (FolinCiocalteu)	Xi (2006c)
Ginseng	Ginsenosides	Dried	500	2	RT	1/75	50% Ethanol	RT; US; SC-CO2	-	-	-	-	Shouqin et al. (2007)
Rhodiola sachalinensis	Flavonoids; salidroside	Dried in vacuum	500	3	RT	1/70	60% Ethanol	US; Leaching; Soxhlet; Heat reflux	DPPH	-	-	HPLC	Zhang et al. (2007)
Propolis	Phenolic compounds;	Crude	500	1	RT	1/35	75% Ethanol	Leaching at RT;	β-carotene bleaching;	-	-	Phenolic	Xi & Shouqin

			HPE conditions						Bioactivities and methodologies of analysis				
Source	**Bioactive compounds**	**Pre-treatment**	**Pressure (MPa)**	**Time (min)**	**T (ºC)**	**Ratio Mass to volume (g/mL)**	**Solvent**	**Other extraction methods**	**Antioxidant**	**Antimicrobial**	**Structural changes**	**Other analyses (methods)**	**References**
	Flavonoids							Heat reflux	DPPH			compounds (FolinCiocalteu); Total flavonoids (colorimetry)	(2007)
Sour cherry pomace	Phenolic compounds	Crashed, heated and pressed	176-193	25	60	1/15	Ethanol	SC-CO2; Solid-liquid extraction	DPPH	-	-	Phenolic compounds (FolinCiocalteu)	Adil et al. (2008)
Grape waste	Anthocyanins	Whole skins	600	60	70	1/4.5	50% Ethanol	PEF; US	ABTS	-	-	LC-DAD/ESI-MS; Phenolic compounds (FolinCiocalteu)	Corrales et al. (2008)
Grape waste	Anthocyanins	Whole skins	600	30	70	1/4.5	50% Ethanol	-	ABTS	-	-	HPLC-DAD/ESI-MS	Corrales et al. (2009)
Ginseng	Ginsenosides	Dried in oven	200	5	60	1/50	70% Ethanol	MW; US; Soxhlet; Heat reflux	DPPH	-	Scanning electron microscope (SEM)	Colorimetry	Chen et al. (2009)
Rhodiola sachalinensis	Salidroside	Dried in vacuum	300	3	25	1/50	60% Ethanol	Cellulase; US; Leaching; Heat reflux; Soxhlet	-	-	-	-	Bi et al. (2009)
Green tea	Phenolic compounds	Dried in vacuum	500	1	RT	1/20	50% Ethanol	RT; US; Heat reflux	-	-	-	Phenolic compounds (FolinCiocalteu)	Xi et al. (2009b)
Green tea	Caffeine	Dried in vacuum	500	1	RT	1/20	50% Ethanol	RT; US; Heat reflux	-	-	-	-	Xi (2009a)
Litchi fruit pericarp	Phenolic compounds	Dried in hot air	500	2.5	70	1/50	85% Ethanol	US; RT	DPPH; Superoxide	-	-	Phenolic	Prasad et al.

			HPE conditions						Bioactivities and methodologies of analysis				
Source	**Bioactive compounds**	**Pre-treatment**	**Pressure (MPa)**	**Time (min)**	**T (ºC)**	**Ratio Mass to volume (g/mL)**	**Solvent**	**Other extraction methods**	**Antioxidant**	**Antimicrobial**	**Structural changes**	**Other analyses (methods)**	**References**
		oven							anion			compounds (FolinCiocalteu)	(2009e)
Longan fruit pericarp	Phenolic compounds	Dried in hot air oven	500	2.5	50	1/50	50% Ethanol	RT	DPPH; Superoxide anion	-	-	Gallic acid calibration	Prasad et al. (2009d)
Longan fruit pericarp	Phenolic compounds	Dried in hot air oven	500	2.5	30	1/50	50% Ethanol	RT	Phosphomolyb denum; lipid peroxidation; DPPH; superoxide anion	-	-	Gallic acid calibration; Identification (HPLC)	Prasad et al. (2009a)
Longan fruit pericarp	Corilagin	Dried in hot air oven	500	2.5	30	1/50	50% Ethanol	US; Solid-liquid extractio	-	-	-	Identification (HPLC)	Prasad et al. (2009c)
Litchi fruit pericarp	Flavonoids	Dried in hot air oven	400	30	25	1/40	Ethanol: HCl (85:15)	US; Solid-liquid extraction	DPPH; Superoxide anion	-	-	Phenolic compounds (FolinCiocalteu); Identification (HPLC)	Prasad et al. (2009b)
Longan fruit pericarp	Polysaccharides; Lignins; Cellulose	Dried in freeze dryer	500	30	25	1/15	Distilled water	Control (0.1 MPa, 25 ºC, 30 min)	-	-	-	Isolation of polysaccharides and cellulose; Acid hydrolysis of cellulose	Yang et al. (2009)
Schisandra chinensis	Deoxyschisandrin; yschisandrin	Dried in vacuum	400	5	RT	1/90	90% Ethanol	Heat reflux; US	DPPH	-	-	Identification (HPLC)	Liu et al. (2009)
Berberis koreana	Phenolic compounds	Crude	500		RT	1/10	Water	US; Solid-liquid extraction	DPPH; Xanthine oxidase	-	-	HPLC; Phenolic	Qadir et al. (2009)

			HPE conditions						Bioactivities and methodologies of analysis				
Source	**Bioactive compounds**	**Pre-treatment**	**Pressure (MPa)**	**Time (min)**	**T (ºC)**	**Ratio Mass to volume (g/mL)**	**Solvent**	**Other extraction methods**	**Antioxidant**	**Antimicrobial**	**Structural changes**	**Other analyses (methods)**	**References**
												compounds (FolinCiocalteu)	
Korean barberry	Phenolic compounds	Dried stem brought	500	30	30	1/90	Distilled water	Solid-liquid extraction	-	Probiotic activity; MIC	-	pH; Phenolic compounds (FolinCiocalteu)	Lee et al. (2010)
Deodeok roots	Phenolic compounds and Flavonoids	Dried in cabinet-type convective drier, and grinded	500	30	50	-	70% Ethanol	Solid-liquid extraction	DPPH; Ferric reducing power	Probiotic activity; MIC and MBC	-	pH; Phenolic compounds (FolinCiocalteu); Identification (HPLC)	He et al. (2010)
Longan fruit pericarp	Phenolic compounds	Dried in hot-air oven	500	30	30	1/50	50% Ethanol	US; Solid-liquid extraction	DPPH; Reducing power; Total antioxidant activity; Superoxide anion radical; Lipid peroxidation	-	-	Identification (HPLC)	Prasad et al. (2010b)
Bitter melon	Momordicosides	Dried in hot air oven	423.1	7	30	1/45.3	70% Ethanol	Heat reflux	-	-	-	Total momordicosides (UV/Vis); HPLC	Ji et al. (2010)
Ginseng	Ginsenosides	Dry powder (purchased)	600	5	RT	-	Water	RT	-	-	-	Identification (HPLC)	Shin et al. (2010)

			HPE conditions						Bioactivities and methodologies of analysis				
Source	**Bioactive compounds**	**Pre-treatment**	**Pressure (MPa)**	**Time (min)**	**T (ºC)**	**Ratio Mass to volume (g/mL)**	**Solvent**	**Other extraction methods**	**Antioxidant**	**Antimicrobial**	**Structural changes**	**Other analyses (methods)**	**References**
Longan fruit pericarp	Phenolic compounds	Dried in hot air oven	500	30	30	1/50	50% Ethanol	RT	DPPH; Superoxide anion; Phosphomolybdenum	-	-	Identification (HPLC)	Prasad et al. (2010a)
Green tea	Catechins and caffeine	Dried in vacuum	400	15	RT	1/20	50% Ethanol	Solid-liquid extraction	-	-	-	HPLC	Xi et al. (2010)
Ginseng	Ginsenosides	Fresh roots versus Dried roots	80	12h	30	1/20	Water	Heat extraction	-	-	-	Ginsenosides (HPLC); Phenolic compounds (FolinCiocalteu); Total sugars (phenolH2SO4); Volatile compounds (GC-MS)	Lee et al. (2011)
Deodeok	Phenolic compounds; Flavonoids	Inlet air temperature	300	20	30	1/5	Water	Heat extraction	DPPH	-	-	Phenolic compounds (FolinCiocalteu); Identification (HPLC)	He et al. (2011)
Green tea	Phenolic compounds	Fresh leaves pulverized	400	15	RT	1/20	50% Ethanol	-	-	-	SEM; TEM	-	Xi et al. (2011a)
Epimedium koreanum Naka	Flavonoids	Dry powder (purchased)	350	5	-	-	50% Ethanol	Ultrasounds, Heat reflux, SC-CO2	-	-	-	Total flavonoids (colorimetry)	Hou et al. (2011)
Green tea	Phenolic compounds	Dried in hot air oven	450	5	RT	1/20	50% Ethanol	Solid-liquid extraction	DPPH; Phosphomolybdenum	-	-	Phenolic	Xi et al. (2011b)

			HPE conditions						Bioactivities and methodologies of analysis				
Source	**Bioactive compounds**	**Pre-treatment**	**Pressure (MPa)**	**Time (min)**	**T (ºC)**	**Ratio Mass to volume (g/mL)**	**Solvent**	**Other extraction methods**	**Antioxidant**	**Antimicrobial**	**Structural changes**	**Other analyses (methods)**	**References**
												-compounds (FolinCiocalteu)	
Cedrus male cones	Pollen protein	Air dried	330	30	RT	1/5	0.2M PBS	-	-	-	Light microscopy; SEM	Bradford	Altuner et al. 2012)
Orange peel	Pectin	Vacuum drying oven	500	10	55	1/50	Water	Heat extraction; MW	-	-	Rheology (viscosity)	Gelling properties; Activation energy; Degree of esterification	Guo et al. (2012)
Dendrobium candidum	Polysaccharides	Fresh flowers	445.3	6.7	-	1/237.9	-	Heat reflux	-	-	-	-	Tao et al. (2012)
Dysosma versipellis	Podophyllotoxin; 4'-demethylpodophyllotoxin	Crude	200	1	-	1/12	80% Methanol	Heat reflux	-	-	-	HPLC; ESI-MS; NMR	Zhu et al. (2012)
Green tea	Phenolic compounds	Dried in hot air oven	500	15	RT	1/20	50% Ethanol	-	-	-	-	Phenolic compounds (FolinCiocalteu)	Xi et al. (2013)
Beer wort	Xanthohumol	Boiled	250	5	25	-	-	Boiling	-	-	-	HPLC-UV/Vis	Santos et al. (2013)
Honey pomelo	Pectin	Dried in a vacuum freeze dyer	500	10	55	1/50	Distilled water + 0.5M hydrochloric acid	High-speed shearing; Thermal extraction	-	-	Viscosity; Light microscopy; Emulsion stability	Galacturonic acid; Degree of esterification; Protein content; Molecular weight	Guo et al. (2014)

			HPE conditions						Bioactivities and methodologies of analysis				
Source	**Bioactive compounds**	**Pre-treatment**	**Pressure (MPa)**	**Time (min)**	**T (ºC)**	**Ratio Mass to volume (g/mL)**	**Solvent**	**Other extraction methods**	**Antioxidant**	**Antimicrobial**	**Structural changes**	**Other analyses (methods)**	**References**
Mango peel	Mangiferin; Lupeol	Freeze dried	150	20	25	1/10	80% Ethanol; Hexane	Maceration; Soxhlet; US; MW	-	-	-	HPLC	Ruiz-Montanez et al. (2014)
Lemon peels	Phenolic compounds	Crude	500	3	10	-	-	Control (0.1 MPa, 3-10 min)	DPPH	-	-	Phenolic compounds (FolinCiocalteu)	Casquete et al. (2014)
Citrus peels	Phenolic compounds	Crude	300	3	10	-	-	Control (0.1 MPa, 3-10 min)	DPPH; ABTS	Growth inhibition (Halo formation)	-	Phenolic compounds (FolinCiocalteu)	Casquete et al. (2015)
Tomato waste	Total carotenoids; Lycopene	Air dried and crushed	700	10	25	1/4	Ethyl lactate	Control (0.1 MPa, 25 ºC, 30 min)	-	-	-	Total carotenoids (colorimetry); Lycopene (HPLC)	Strati et al. (2015)
Chilean papaya seeds	Antioxidants; Sulforaphane; Fatty acids	Air died in dark	500	15 pulses of 1 min	RT	-	80% Methanol	Solid-liquid extraction; US	DPPH; FRAP	-	-	Moisture; Protein content; Lipid content; Fiber; Ash; Phenolic compounds; Total flavonoids; Sulforaphane; Oil extraction	Briones-Labarca et al. (2015)
Orange peel	Phenolic compounds; Flavonoids	Freeze dried	50	30	35	1/10	80% Ethanol	Solid-liquid extraction; US; MW; SC-CO2	ABTS	-	-	Phenolic compounds (FolinCiocalteu); Total flavonoids (colorimetry); Identification	M'hiri et al. (2015)

			HPE conditions						Bioactivities and methodologies of analysis				
Source	**Bioactive compounds**	**Pre-treatment**	**Pressure (MPa)**	**Time (min)**	**T (ºC)**	**Ratio Mass to volume (g/mL)**	**Solvent**	**Other extraction methods**	**Antioxidant**	**Antimicrobial**	**Structural changes**	**Other analyses (methods)**	**References**
												(HPLC)	
Moringa seeds	Essential oil	Purchased and cleaned	19.63	27.17	85.57	-	Water	-	-	-	-	Moisture content and yield	Fakayode and Ajav (2016)
Shrimp waste	Astaxanthin	Shells separated from flesh and vacuum dried	200	5	-	1/20	Ethanol	Solid-liquid extraction	-	-	SEM	-	Li et al. (2016)
Passion fruit peel	Pectin	Grounded and dried in oven with air circulation	300	20	50	1/30	Nitric acid, pH 1.0	Pressure pretreatment + high temperature / Heat extraction	-	-	-	Pectin purification; Galacturonic acid; Degree of esterification; Apparent viscosity	Oliveira et al. (2016)
Egg yolk	5-methyltetrahydrofolate	Fresh white shelled eggs	400	5	RT	1% solids	Mili-Q water	-	Total reducing capacity (FolinCiocalteau); DPPH	-	-	Total nitrogen content; RP-HPLC; Electrophoresis	Naderi et al. (2017)
Garden pansy	Phenolic compounds	Freeze-dried and grounded	384	15	RT	1/30	35% Ethanol	-	DPPH; ABTS; FRAP	-	-	Flavonoids (colorimetry); Hydrolysable tannins (colorimetry); Total monomeric anthocyanins (pH differential)	Fernandes et al. (2017)

			HPE conditions						Bioactivities and methodologies of analysis				
Source	**Bioactive compounds**	**Pre-treatment**	**Pressure (MPa)**	**Time (min)**	**T (ºC)**	**Ratio Mass to volume (g/mL)**	**Solvent**	**Other extraction methods**	**Antioxidant**	**Antimicrobial**	**Structural changes**	**Other analyses (methods)**	**References**
Pomegranate peel	Phenolic compounds	Dried in a laboratory incubator with air circulation	356-600	30	RT	1/15	32-56% Ethanol	-	DPPH; ABTS; FRAP	-	-	Phenolic compounds (FolinCiocalteau); Total condensed tannins (vanillin method); Total flavonoids (Dowd method); Total anthocyanin (pH differential); Identification (uHPLC and LCDAD/ESI-MS)	Alexandre et al. (2017b)
Fermented fig	Phenolic compounds	Dried and grounded	600	5-30	RT	1/15	<15% Ethanol	-	DPPH; ABTS; FRAP	-	-	Phenolic compounds (FolinCiocalteau); Total condensed tannins (Vanillin method); Total flavonoids (Dowd method); Identification (LCDAD/ESI-MS, and HPLC-DAD)	Alexandre et al. (2017a)

			HPE conditions						Bioactivities and methodologies of analysis				
Source	**Bioactive compounds**	**Pre-treatment**	**Pressure (MPa)**	**Time (min)**	**T (ºC)**	**Ratio Mass to volume (g/mL)**	**Solvent**	**Other extraction methods**	**Antioxidant**	**Antimicrobial**	**Structural changes**	**Other analyses (methods)**	**References**
Japanese raisin tree	Total phenolic acids	-	400	60	-	-	-	Hot water extraction	-	-	-	Enzymes: alcohol dehydrogenase, aldehyde dehydrogenase, glutathione-Stransferase	Lee (2017)
Blue honeysuckle berries	Anthocyanins	-	426	7	-	1/14.7	-	Ultrasound assisted extraction; Solid-liquid extraction	DPPH; ABTS; FRAP	-	-	Identification (HPLC-DAD-MS)	Li et al. (2018)
Xinjiang jujube leaves	Flavonoids	Cleaned, dried in lyophilizer	342.39	11.56	50	1/43.95	70% Methanol	Ultrasound assisted extraction	DPPH; ABTS	-	-	Total Flavonoids (colorimetry); Identification (UPLC-ESI-MS)	Zhang et al. (2019)
Pomegranate peel	Phenolic compounds	Dried in a laboratory incubator with air circulation	300-600	15	RT	1/62	Water	Enzymatic extraction prior to HPE	DPPH	Well diffusion; MIC; MBC	-	Phenolic compounds (FolinCiocalteau); Identification (uHPLC)	Alexandre et al. (2019)
Ecliptae herba	Wedelolactone and isodemethylwedelolactone	Grounded	200	3	RT	1/20	80% Methanol	Heat reflux	-	-	-	Identification (HPLC)	Zhao et al. (2019)

			HPE conditions						Bioactivities and methodologies of analysis				
Source	**Bioactive compounds**	**Pre-treatment**	**Pressure (MPa)**	**Time (min)**	**T (ºC)**	**Ratio Mass to volume (g/mL)**	**Solvent**	**Other extraction methods**	**Antioxidant**	**Antimicrobial**	**Structural changes**	**Other analyses (methods)**	**References**
Yellow prickly pear peel	Bioactive compounds	Dried in a laboratory incubator with air circulation	300-600	5-30	RT	1/40	0-80% Ethanol	Soxhlet	DPPH; ABTS; FRAP	Well diffusion	-	Phenolic compounds (FolinCiocalteau); Total condensed tannins (Vanillin method); Total flavonoids (Dowd method); Total betalains (colorimetry); Total carotenoids (colorimetry)	Castro et al. (2019)
Tomato pulp	Flavonoids; Lycopene	Washed and blended	450	10	20	1/2	60% Hexane	Solid-liquid extraction	DPPH; FRAP	-	-	Phenolic compounds (FolinCiocalteau); Total Flavonoids (colorimetry); Identification (HPLC); Simulated gastrointestinal tract model	Briones-Labarca et al. (2019)

4.1.2 Bioactivity screening

Biological extracts may exert several actions at the biological level, such as antifungal, antimicrobial, antioxidant, anti-proliferative, anti-inflammatory, anticholinesterase, among other properties (Cetkovic et al., 2007; Elgndi et al., 2017; Silva et al., 2019). Moreover, the research carried out by Moreira et al. (2020a) aimed to evaluate the impact of winter savory leaves extraction using HPE technology optimized by the response surface methodology to achieve winter temperate extract with a high level of bioactive compounds and high antioxidant action. Therefore, the study showed that the methods predicted an ideal condition between 200–500 MPa, extraction time around 1–20 minutes and optimal ethanol concentration of 0–70% (v/v). In this way, compared to Atmospheric Pressure Extraction, HPE proved to be more efficient, which allowed a growth of approximately 40% for all compounds, and a growth of 29, 48 and 70% for the antioxidant activity by the ferric reducing antioxidant power (FRAP), 2,2-diphenyl-1-picrylhydrazy (DPPH) and 2,2'-azinobis (3-ethylbenzothiazoline-6-sulfonic acid (ABTS) tests, in that order (Moreira et al., 2020a, 2020b).

According to Thaipong et al. (2006), there are some essays available to predict the antioxidant action of herbal extracts, among which stand out as ABTS and DPPH radical scavenging assays (Thaipong et al., 2006). The DPPH radical has greater stability than ABTS, which allows for greater reproduction capacity, being also quick and simple, but it is better diluted in organic solvents, and only demonstrate the antioxidant action of lipophilic substances (Moreira et al., 2020b). Commonly, the antioxidant action varies according to the total concentration of phenolic compounds in extracts (Prasad et al., 2009a, 2009d, 2010a, 2010b; Xi, 2006c; Xi & Shouqin, 2007). Casquete et al. (2014, 2015) researched the impact of the HPE technique on antioxidant and antimicrobial actions of phenolic substances contained in extracts of various citrus fruit peels (Casquete et al., 2014, 2015). Thus, it was noted that the antioxidant action, in experiments ABTS and DPPH increases with the highest concentration of phenolic substances, being the maximum values obtained after HPE at 300 MPa, for 3 min (Casquete et al., 2015).

Xi (2006c) drew comparisons between the antioxidant action of propolis flavonoids using HPE at 500 MPa, 1 minute, at room temperature and 75% ethanol as solvent and extraction through the heat reflux method. The authors described that there was a greater antioxidant action (~70%, DPPH assay) of the HPE extracts compared to the method that aims to extract through heat reflux (~60%, DPPH assay) (Xi, 2006c). Thus, in addition to demonstrating better performance compared to other extraction methods, HPE resulted in extracts with greater antioxidant action compared to synthetic ones, such as ascorbic acid (Prasad et al., 2009b, 2009e; Xi, & Zhang, 2007) and butylated hydroxytoluene (Prasad et al., 2010b). Shouqin et al. (2007a) researched ethanol extracts of *Rhodiola sachalinensis* obtained after HPE and described an antioxidant action of approximately 92% (DPPH assay). In addition, Chen et al. (2009) achieved ginseng extracts with 55% and 58% action of DPPH radical after HPE and to extract through heating reflux, in that order. The biggest advantage was that the HPE took only five minutes for the extraction to complete, while the heat reflux lasted for more than four hours (Chen et al., 2009).

Another interesting example displaying promising bioactive properties is the pansies flower, which can be eaten and used for seasoning and garnishing in salad, soup, dessert, and beverage. Endowed with natural antioxidants due to the presence of carotenoids, they have huge interest in industry, specifically as dietary supplementation or as additives to improve flavors and colors in food and cosmetic products (Gamsjaeger et al., 2011; Jun, 2013; Skowyra et al., 2014; Vukics et al., 2008). Extraction of carotenoids, flavonoids, and other unknown antioxidant elements from pansies, were only carried out in the laboratory under traditional conditions using solvents at atmospheric pressure (Rop et al., 2012; Skowyra et al., 2014, Vukics et al., 2008; Zhang et al., 2012), usually for a long period (24h), under agitation or vortexing (Rop et al., 2012; Skowyra et al., 2014, Zhang et al., 2012). Thus, the research objective of Fernandes et al. (2017) was for the first time to evaluate the implementation of HPE to extract bioactive compounds from pansy flowers with the goal to obtain antioxidant bioactive substances, such as flavonoids and phenols by Total Reducing Capacity (TRC) and antioxidant action by DPPH to eliminate free radicals (Fernandes et al., 2017).

The industry is more adapted with the modern technologies, such as those used to manufacture artificial synthetic food. Nevertheless, the accessibility to new biological sources came to counter these trends (Sun et al., 2020). Thus, mushrooms came to show that it is not only foods that kill hunger, but also, they have a pleasant taste, as well as important nutrients with value for medicine, once it can prevent pathologies, being a source of compounds that have bioactivity (Chaudhary & Tripathy, 2015).

More than 3 000 mushrooms, according to Wasser (2011), can be used as food and around 100 of these mushrooms can be used for commercial purposes (Wasser, 2011). Mushrooms, in addition to not being scarce, have numerous compounds, like polysaccharides, proteins, vitamins and fibers, with a wide spectrum of bioactivities such as antifungal, anti-inflammatory, antitumor, antiviral, antibacterial and immunomodulatory actions (Ha et al., 2000; Ishara et al., 2018). Among the main species of edible fungi produced in the world, *Agaricus bisporus*, *Agrocyte cylindraceaare*, *Auricularia auricula, Flammulina velutipes, Grifola frondosa, Hericium erinaceus, Lentinus edodes, Pleurotus eryngii, Pleurotus ostreatus, Tremella fuciformis, Volvariella vovlaca* stand out (Hua & Zhang, 2018; Wang et al., 2017).

Mushrooms can easily be used as a functional food (Gu et al., 2016). Thus, the increased consumption of mushrooms in a regular diet is considered useful to prevent numerous pathologies caused by oxidative stress, representing an option for treatment based on natural products (Janjušević et al., 2017). Branen (1975, as cited in Seedevi et al., 2019), related that an antioxidant is a substance capable of reacting with reactive oxygen species (ROS), reducing the threat of chronic pathologies in humans and other living beings, containing antioxidants that can protect them from damage caused by oxidation; however, such antioxidants are not enough to prevent the resulting damage (Branen, 1975; Seedevi et al., 2019). Synthetic antioxidants, such as butylated hydroxyanisole (BHA), butylated hydroxytoluene (BHT) and propyl gallate (PG) contain intense radical scavenging activity (RSA); however, the use of such antioxidants is limited because of their adverse impacts, namely skin allergies, gastrointestinal tract problems and in some cases increased

the risk of cancer. As a result, huge interest is being turned to researching natural antioxidants to replace synthetic ones. Currently, the progress of natural antioxidants provided with effectively and safely is one of the most important tasks in the field of antioxidant research. In this view, the antioxidants found in fungus are of enormous interest as powerful agents that protect against damage from oxidation (Lourenço et al., 2019; Mau et al., 2002).

A polysaccharide called PNP from *Pholiota nameko* was extracted and purified. *Pholiota nameko* is a fungus that can be eaten and is widely cultivated in China. Traces of the PNP structure were researched by combining chemical and instrumental level analyses. According to the results of the chemical and structural analysis carried out by UV-Visible spectra, FTIR and ^{1}H-NMR spectroscopy, showed that PNP is essentially formed from glucose, mannose, polysaccharides and polyphenols, (Zhu et al., 2019). Furthermore, based on additionally results carried out using FTIR-ATR (attenuated total reflectance) and ^{1}H-NMR, it was confirmed that PNP is formed by *α*- and *β*-glycosidic structures, as well as α-glucans, *β*-glucans and glucan-protein complexes. Thus, *in vitro* antioxidant results indicated that the elimination actions for radical scavenging activity of the polysaccharide, superoxide, hydroxyl and the reducing potential increased with increasing polysaccharide concentration. Nevertheless, each of the four antioxidant potentials was lower than Vc and the hydroxyl radical scavenging action was the most potent with 70.7% of RSA (Fan et al., 2020; Rodrigues et al., 2017; Zhu et al., 2019). The *in vitro* antioxidant results showed that PNP has antioxidant potential and may play a role in protecting the biofilm and acting against aging. Thus, *P. nameko* can be used for healthy eating and medication creation (Fan et al. 2020; Zhu et al., 2019).

According to Janjušević and its collaborators (2017), *Trametes versicolor (L.),* often referred to as turkey tail, is a widespread white rot ligneous fungus that proliferates in several deciduous trees (oak, *Prunus*) and other conifers (spruce and pine) with most basidium's appearing on stumps and trunks over the course of a year (Janjušević et al., 2017). It is a species that cannot be eaten; however, it is the world's most famous medicinal mushroom (Kamiyama et al., 2013). Current *in vitro* and *in vivo* research, like those of the next authors mentioned, has shown that *T. versicolor* has numerous properties with medical applications, such as activity against diabetes (Liu et al., 2012), tumors (Luo et al., 2014), microbes (Helba, 2014), immune system stimulant (Trovato et al., 2016) and antioxidant action (Karaman et al., 2010; Kozarski et al., 2012; Sun et al., 2014). According to Wan (2013), the main compounds with bioactivity identified in this species integrate, for the most part, the set of proteins and polysaccharides, with the polysaccharide krestin and the polysaccharopeptide PSP being the most researched (Wan, 2013).

To determine the antioxidant activity, total phenolic compounds and total flavonoids content of two mushroom species, namely *Tramestes versicolor* and *Trametes gibbosa*, harvested from north-west part of Romania, mushroom extracts of water, acetone and methanol were analyzed using Ultraviolet-visible (UV-Vis) spectroscopy, FTIR spectroscopy and Liquid chromatography-Mass Spectrometry (LC-MS) techniques. In total, 28 bioactive compounds were characterized as phenolic acids (11 compounds), flavonols (6 compounds), flavones (6 compounds), coumarins (2 compounds), flavanols, isoflavonoids and biflavonoids (1 compound). The most potent antioxidant

action was attributed to the methanol extract, while the highest quantities of total polyphenols and flavonoids were specifically for the water extract. *Trametes* species have been presented as a rich source of nutritional compounds with important pharmacological properties, such as antioxidant, anticancer and anti-inflammatory actions (Pop et al., 2018).

Shiitake (*Lentinula edodes*) is the most famous consumed edible mushroom on the world market (Royse et al., 2017), of great value in Eastern and currently Western cuisine because of its typical flavor. Such mushroom has molecules that produce excellent impacts on human health, such as phenolic compounds and ergothioneine with antioxidant action, ergosterol, *β*-glucans and eritadenine with hypocholesterolemic action, antihypertensive peptides, lenthionine with antithrombotic action and among others, however, it is necessary to give special attention to lentinan (Morales et al., 2018). It is a well-known glucan based on the sequence *β-D*-glucopyranose units *(1 → 3)* linked *β-D*-glucopyranose. This polysaccharide attracted the medical curiosity due to its intense activity against tumors *in vitro* and *in vivo*, as well as immunomodulatory and antiviral potentials (Xu et al., 2014; Zhang et al., 2011).

According to research by Morales et al. (2020), a polysaccharide extract obtained from *Lentinula edodes* was subjected to numerous stages of purification to proceed with the separation of the three distinct fraction of *D*-glucans having especially a linear *(1 → 3) -α- D*-glucan, a linear *(1 → 6) -β-D*-glucan and one branched *(1 → 3), (1 → 6) -β-D*-glucan, being characterized by Gas chromatography-mass (GC-MS), FTIR, NMR technologies. Furthermore, the antioxidant, anti-tumor, anti-inflammatory and hypocholesterolemic action *in vitro* were also studied. Glucans showed interesting ability to inhibit the coenzyme A reductase 3-hydroxy-3-methylglutaryl (HMGCR), however, only the factions of *β- (1 → 6) and β- (1 → 3), (1 → 6)* showed potential for inhibit DPPH. Glucans also showed potential to inhibit Interleukin 1 beta (IL-1β) and IL-6 secretion by LPS-activated THP-1/M cells and showed cytotoxic potential in a breast cancer cell line, which was not noted in healthy breast cells. Such *in vitro* studies indicated relevant results for further *in vivo* research, evidencing distinct impacts of each chemical structure of glucans under isolation of *shiitake* mushrooms (Morales et al., 2020).

Seedevi et al. (2019) purified a polysaccharide taken from *Pleurotus sajor-caju*, a misnamed for *Lentinus sajor-caju*, by anion exchange column chromatography and subjected to purification using gel permeation column chromatography. The polysaccharide chemically characterized revealed that it has in its entirety, 90.16% of carbohydrates, 12.7% of ash, 5.2% of moisture and 0% of protein; however, the quantity of carbon, hydrogen and nitrogen were 31.53%, 4.28% and 3.01%, respectively. The polysaccharide characterized with the chemical structure *→ 6) α-D-Glciv (1 → 6) α-D-Glciii (1 → 6) β-D-Glcii (1 → 6) α-D-Glci (1 →* units, showed 21.67-68.35% of DPPH, 16.01-70.09% of ABTS, 24.31-73.64% of superoxide radical scavenging activity, 16.64-63.51% of hydroxyl radical activity and 18.61-63.21% of anticancer activity against AGS human gastric carcinoma cell line. The active principle of the polysaccharide may be industrially used in the food and pharmacy sector in the future (Seedeve et al., 2019).

The basidiomycete *Phanerochaete chrysosporium* is a type of *ligninolytic* fungus that has been researched. Studies carried out by Liu and its collaborators (2014) showed that extracts of *P. chrysosporium* had a notorious antioxidant action, as well as a high potential to eliminate reactive oxygen radicals ($O_2^{\cdot-}$ radical, ·OH radical and H_2O_2). The extracts of *P. chrysosporium* showed an evident increase in the antioxidant action with the increase of the fungi concentrations. Furthermore, in the research aforementioned, it was determined that several antioxidant compounds, such as the catalase, superoxide dismutase (SOD) and glutathione peroxidase (GPx) enzymes, and phenolic, flavonoids and glutathione were permanently present in *P. chrysosporium* extracts. Very positive correlations between antioxidant action and antioxidants were discovered by analyzing the relationship between the enzymes action and the action of the compounds. In addition, antioxidant stimuli were observed in *P. chrysosporium* extracts exposed to Cd, which generated another dimension of systemic impact of *P. chrysosporium* antioxidants for immune defense. An increase of 2.78-fold and 2.35-fold was found in superoxide dismutase and total phenolic, respectively. In short, these data are relevant to clarify the functions of such antioxidants in the physiological level of antioxidant action, which offers a deepening of recent knowledge about the defense processes of *P. chrysosporium* adjusting to reactive oxygen stress (Liu et al., 2014). The fact that a huge number of pathologies and clinical conditions are related to oxidative stress has stimulated the emergence of much research on other antioxidant sources in nature, particularly coming from microorganisms and mushrooms (Khatua et al., 2013; Mau et al., 2002).

Equally interesting is the broad spectrum of fungi with antibacterial activity (Figure 4.1). In the research carried out by Janeš et al. (2007), among the fungi extracts tested, three demonstrated action against different bacteria. Thus, the mushroom extract *Amanita virosa* (Fr.) Bertill. (Amanitaceae) showed action against *Pseudomonas aeruginosa*; *Cortinarius praestans* Cordier (Cortinariaceae) against *Staphylococcus aureus* and endophytic fungus extract *Trucatella hartigii* (Tubeuf) Steyaert (Amphisphaeriaceae) against *S. aureus*; and *Enterococcus faecalis* (Janeš et al., 2007).

In association with ongoing screening actions for biologically active secondary metabolites of fungi, there is the research carried out by Kumar et al., (2010), which intended to screen new microorganisms and to select strains with antibacterial and anti-*Candida* action from unexplored areas of Kaziranga National Park, Assam, India, which has very moist soil and excellent conditions for fungal growth. The screening result showed that 42 out of 130 fungi isolated showed antimicrobial action, 15 with antibacterial action, 20 with antibacterial and anti-*Candida* action, and 7 with exclusively anti-*Candida albicans* action. Thirteen among those submitted to isolation with antibacterial action showed action against gram-positive bacteria of the type *Bacillus subtilis* and *Staphylococcus aureus*; and only 2 among the subjects to the isolation demonstrated activity against gram-negative bacteria (Kumar et al., 2010).

Santos et al. (2015), described pioneering research on the antimicrobial action of endophytic fungi found on the leaf of *Indigofera suffruticosa* Miller. Among the 65 endophytic fungi submitted to isolation and, subsequently, submitted to an antimicrobial screening, 18 fungi showed action against

Bacillus subtilis, but the most effective result was noticed against *Staphylococcus aureus* (Santos et al., 2015). In a study, carried out by Manganyi et al., (2018), *Palargonium sidoides* fungi were evaluated for activity against *Escherichia coli* bacteria using the standard Kirby-Bauer agar disk diffusion technique. A total of 133 fungi from 32 genres were effectively isolated and identified based on morphological analysis. Thus, according to the aforementioned research, fungi of the *Penicillium* and *Aspergillus* genus showed dominance antimicrobial activity (Manganyi et al., 2018).

The natural products developed from microorganisms, according to Demain (2014), have the potential to be useful in several areas that include agriculture, industry and, the medical field. However, according to Ventola (2015), there is a significant reduction of antibiotics found in recent times, which was due to a series of factors, including the lack of interest from the pharmaceutical industrial sector and the competition of semi-synthetic substances as potential medicines. In this context, the scenario impacts the research aimed at the screening of antibiotics, which can be perceived as a first phase in the map of the pharmaceutical sector sales stages, which is not something positive, considering that there are micro-organisms resistant to multiple drugs and disease-causing agents detected today (Ventola, 2015).

According to Kaul et al. (2012), fungi are known to produce a wide variety of bioactive metabolites. As part of a research program based on the variety of fungi, Synytsya and this collaborators in 2016 tried to carry out a previous screening of certain fungal strains that presented the potential to produce metabolites with activity against disease-causing microbes. *In vitro* studies pointed out that the chosen fungal strains generated metabolites with bioactive activity against a wide spectrum of bacteria (*Bacillus cereus, Enterococcus faecalis, Micrococcus luteus, Staphylococcus aureus, Staphylococcus epidermidis, Enterobacter cloaconae* and *Pseudomonas aeruginosa* (Synytsya et al., 2017).

The action of *shiitake* or *Lentinula edodes* extract against bacteria was also studied once mushrooms are considered as a source to develop natural antibiotics. Thus, numerous mushroom by-products have been applied against disease-causing agents in humans, to activate the immune system and improve human health because of their antioxidant and antitumor activity (Kitzberger et al., 2007). The S*hiitake* ethanol extract presented bactericidal action against *Micrococcus luteus* and *Bacillus cereus* and chloroform extract against *Streptococcus mutans*, which generates dental caries and *Prevotella intermedia,* which causes periodontal pathology (Kitzberger et al., 2007).

Despite the fungi present bioactive compounds with interesting potential for application in the pharmaceutical industry, what can provide a research-related opening for new sources of a biological nature for possible drugs, the widespread use of antibiotics has led to the emergence of agents that cause antibiotic-resistant pathologies, which include strains with multiple drug resistances and an increasing number of agents capable of causing infections that do not respond to antibiotics in patients all over the planet (Kumar & Schweizer, 2005; Levy, 2005). Therefore, the emergence of drug resistance in bacteria that cause disease in humans, among microorganisms such as *Mycobacterium tuberculosis*, *Staphylococcus sp.* and *Streptococcus sp*., induced the search for other and more effective antibiotics (Strobel, 2003; Strobel et al., 2004). That's why there is an urgent

demand to control antimicrobial resistance through more efficient use of antibiotics and the reduction of infections in hospitals (French, 2005). Nevertheless, the discovery and development of new potential antibiotics must be maintained, as they are important relevance in controlling and maintaining good effectiveness of antimicrobial treatments (Janeš et al., 2007; van der Waaij & Nord, 2000). Thus, with high nutritional value and recognized worldwide as a food and medicinal source for millennia, mushrooms have proven to be a strong ally in the fight against various diseases, with numerous articles from the beginning of the 21st century describing and highlighting their action against bacteria and fungi (Dulger, 2004; Hatvani, 2001; Hur et al., 2004; Ngai & Ng, 2004).

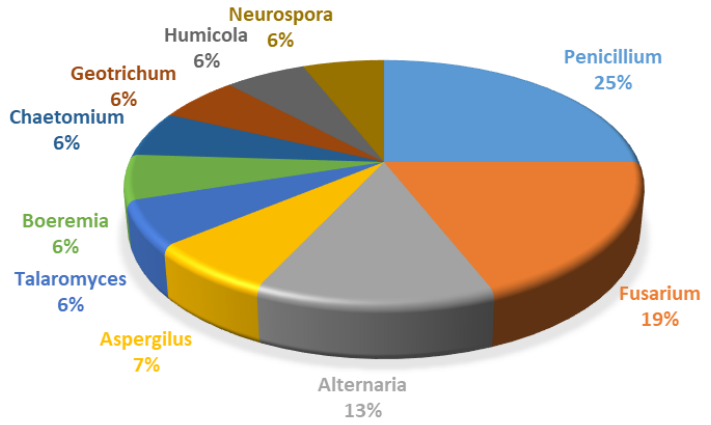

Figure 4.1. Most predominant fungi sources with antibacterial activity. Adapted from Manganyia et al., 2019.

4.1.3 Structural characterization and determination of bioactive extracts

The examination of biomedical, biotechnological or bioenergetics correlations and processes has been the focus of several interdisciplinary work related to the life sciences since the beginning of this century. Methods based on spectroscopy technologies can provide detailed information about complex substances, their dynamics, and reactions (Wolpert & Hellwing, 2006). Nuclear Magnetic Resonance spectroscopy is a powerful technology to help elucidate structures, since the characteristics it shows may be associated with the structure of the molecule. The chemical environment of a certain nucleus is related to the chemical shift (δ, ppm), and the resonance area associated with the number of nuclei that will originate the NMR signal. Thus, the interaction between each individual nucleus, mediated by electrons in a chemical bond defines the coupling constant (J, Hz) (Santos & Silva, 2014).

Carbon-13 NMR spectroscopy or ^{13}C-NMR spectroscopy is the application of nuclear magnetic resonance spectroscopy to carbon. It is analogous to proton NMR (^{1}H-NMR) and enables the detection of carbon atoms in organic molecules, as well as ^{1}H-NMR detects hydrogen atoms, also strongly supported by other 1D (DEPT) and 2D (COSY, TOCSY, HSQC/HMQC, HMBC) NMR

techniques. 13C-NMR spectroscopy shows some limitations that are not found in ^{1}H-NMR spectroscopy, justified by the lower sensitivity to carbon comparing to the sensitive of ^{1}H-NMR to hydrogen, since the main isotope of carbon, the ^{12}C isotope, has a spin quantum number of zero and, thus, it is not magnetically active and, consequently not detected by NMR technique. Only the ^{13}C isotope, much less common, naturally existing at 1.1% of natural abundance, is magnetically active and so detected by NMR. Thus, as mentioned, only the few ^{13}C nucleus present in a sample will vibrate in the magnetic field, although this can be overcome by isotopic enrichment of e.g., protein samples (Balci, 2005; Caytan et al., 2007).

According to Yadav (2013), a possible problem of these spectra arises from the presence of large one bond J-coupling constants between carbon and hydrogen, usually 100 to 250 Hz. Although, only informative, such couplings can hamper spectra and decrease sensitivity. For these reasons, ^{13}C-NMR spectra are generally registered with proton NMR decoupling. Couplings between carbons can be neglected because of 13C their reduced presence in nature. That said, in contrast to common ^{1}H-NMR spectra that demonstrate multiplets for each proton location, carbon NMR spectra show a single peak for each chemically nonmatching carbon atom. Another potential difficult is that the signal intensity is often not proportional to the number of corresponding ^{13}C atoms. Instead, it is heavily impacted, being proportional to the adjacent quantitative spins, often ^{1}H (Yadav, 2013).

The most frequent ^{13}C spectra recording is proton-noise decoupling, which intend to handle the high J values for ^{13}C-H (110–320 Hz), ^{13}C-C-H (5–60 Hz) and ^{13}C-C-C-H (5–25 Hz) that otherwise produce ^{13}C spectra fully coupled to protons that are difficult to interpret. With proton-noise decoupling, where most of the spectra are carried out, a noise decoupler will intensely irradiate the sample with a wide range (around 1000 Hz) of radio frequencies that cover the range where the protons will change their rotations at the nuclear level (100 MHz) (Yadav, 2013). The quick modifications in the proton spin generates an effective heteronuclear decoupling, which increases the intensity of the carbon signal due to the Nuclear Overhauser effect (NOE) and makes the spectrum simpler so that each non matched carbon generates a singlet peak. Thus, the relative intensity is unreliable as some carbons have a higher spin lattice relaxation time period and, others have a weaker NOE highlight (Rahman et al., 2016).

The first attempt at assigning complex ^{13}C-NMR spectral signals is through Distortionless Enhancement by Polarization Transfer (DEPT) experiments, which provide information on the atomic nature of carbon, determining the presence of primary, secondary and tertiary carbon atoms. The DEPT experiment differentiates ^{13}C signals between methyl (CH_3), methylene (CH_2), and methine (CH) carbons by variation of the selection angle parameter (Primasova et al., 2017; Santos & Silva, 2014).

Three DEPT spectra are usually indispensable for integral analysis and use a complex chain of pulses in ^{13}C spaces, a 45° pulse (DEPT-45), where all H-bearing carbon peaks are positive, a 90° pulse (DEPT-90) where only the methine (CH) carbons are visualized, and a 135° pulse (DEPT-135) when the methyl (CH_3) and methine (CH) carbons are positive signals and the methylene carbons (CH_2) have negative signs. Quaternary carbons and other carbons with no attached protons are not

visualized in DEPT spectra as the method is based on transfer polarization, which concerns transferring magnetization from protons to directly bound carbon atoms (Primasova et al, 2017; Rahman et al., 2016; Santos & Silva, 2014).

It has been described in the literature that numerous sterols and triterpenoids isolated from mushroom have inhibitory actions on inflammation generated by *12-O-tetradecanoyl-phorbol-13-acetate* (TPA), a very known tumor promoter (Yaoita et al., 2015). Over the past two decades, the research by Yaoita and its collaborators (2015) have been directed towards phytochemical studies of sterols and triterpenoids on mushroom. Thus, among 18 species, namely, *Amanita pantherina, Amanita virgineoides, Daedaleopsis tricolor, Flammulina velutipes, Grifola frondosa, Hypsizigus marmoreus, Lentinula edodes, Lyophyllum shimeji, Naematoloma sublateritium, Omphalia lapidescens, Panellus serotinus, Pholiota nameko, Pleurotus eryngii, Plellillus ostreatus, Polyporus umbellatus, Sarcodon aspratus, Tricholoma matustake* and *Tricholoma portentosum,* 28 recent sterols and 3 recent triterpenoids were isolated (Yaoita et al., 2015).

According to the aforementioned authors, the structure of the original compounds was determined by spectroscopic techniques, especially 2D-COSY (two-dimensional correlation spectroscopy), NMR and Mass Spectrometry (MS) analysis. Unique and interesting structural traces show a sterol with Δ5.8 structure, 6 sterols with 5α,6α-epoxy group, 5 sterols with 5α,9α-epidioxy group, 6 sterols with enone, diene and ketone, 4 sterols with 3,5,6,9-tetrol structure and 3,5,6,7-tetrol structure, 1 sterol with 1,2,3,4,5,10,19-heptanor structure, 5 sterols of the 23-methylergostane-type and 3 triterpenoids of the lanostane-type were characterized (Yaoita et al., 2015).

Nowadays, the methods used to determine and characterize the structures of diverse bioactive compounds from natural sources are often simple and fast, due to the extraordinary advances in spectroscopic technology. The progress in spectroscopic techniques changed the process of attributing structures, which previously was basically based on deterioration at the chemical level followed by a partial or total synthesis. Due to the common method, named *Fourier Transform Infrared Spectroscopy* (FTIR), fast and precise measurements are possible, since this method is used to map cell constituents, such as lipids and proteins (Levin & Bhargava, 2005; Petibois & Deleris, 2006). This method, according to Berthomieu and Hienerwade (2009), offers complementary information to data on three-dimensional structures obtained through X-ray diffraction or Nuclear Magnetic Resonance (NMR) (Berthomieu & Hienerwade, 2009).

FTIR has been demonstrated as a valuable tool for characterize and detect compounds or chemical connections groups existing in unknown mixtures of extracts (Eberhardt et al., 2007; Hazra et al., 2007). Furthermore, the FTIR spectra of pure compounds are often so unique that they operate like molecular fingerprints. However, for most common plant compounds, the spectra of unknown compounds can be detected by comparisons with a range of known compounds (Sasidharan et al., 2011). Thus, FTIR can be used to collect high resolution data from a wide wavelength, in general between 5000 and 400 cm^{-1} (mid-infrared region). In FTIR, the light of several frequencies can be measured at the same time and this process can be carried out again several times, which offers the FTIR something advantageous as it will result in a higher signal-to-noise relationship for a particular

scan period (Song, 2017). Consequently, given that FTIR spectroscopy determining the presence of fundamental molecular vibrations that are characteristic of a chemical compound or class of compounds, it may be used in the qualitative elucidation of the composition of biomass. The fingerprint region between 1800 and 650 cm^{-1} is generally of particular interest because it contains the most spectral information regarding to the chemical composition of a material. In the literature, several bands have been associated with carbohydrates due to their related functional groups. The resulting peaks due to the polysaccharides include 897 cm^{-1} and 1030 cm^{-1} from the C-H and C-O stretch, respectively. Band in 1157 cm^{-1} correspond to C-O-C vibration, 1239 cm^{-1} from C-O stretch and O-H in plane in polysaccharides, 1465 cm−1 from C-H deformation, and 1740 cm−1 from the C=O stretching of unconjugated ketones. The peak at 1122 cm−1 occurs due to aromatic skeletal and C-O stretch. C-O stretching causes a peak to rise at 1270 cm−1 and syringyl ring breathing creates the peak at 1365 cm−1. The peak at 1505 cm−1 is attributed to the C=C stretch characteristic of aromatic skeletal compounds. The peaks occurring at 2935 cm−1 have been associated with the bending and stretching of C-H, and 3345 cm−1 has been assigned to bonded O-H (Acquah et al., 2016). For example, in 2015, Parmar and Kumar reported that, using FTIR, GC-MS spectrometry and X-ray fluorescence (XRF) were carried out to evaluate the constitution at the chemical level of *Pleurotus cornucopiae* (Paulet), collected from Kalpa on Himachal Pradesh, thirteen peaks were recorded by FTIR spectrophotometry which were equivalent to nine distinct functional groupings. The 3260 cm^{-1} peak corresponding to the ≡C-H stretch; 2925 cm^{-1} and 2858 cm^{-1} matching the C-H stretch; the peak 2117 cm^{-1} which is equivalent to the C≡C stretch; 1633 cm^{-1} corresponding to the C=O stretch; 1566 cm^{-1} and 1551 cm^{-1} matching the N=O stretch; 1380 cm^{-1} which equals the N=O bend; 1153 cm^{-1} and 1020 cm^{-1} corresponding to the C-O stretch; 937 cm^{-1} consistent with the =C-H bend; and 687 cm^{-1} and 651 cm^{-1} corresponding to the C-Cl stretch (Jin-Zhe et al., 2013; Parmar & Kumar, 2015).

Similarly, in a research work carried out by Sarangi et al. (2006), three neutral proteoglycans from *Pleurotus ostreatus* mycelia mushrooms were characterized by the FTIR method. The attribution of the most important absorption traits of the glycosidic structure is associated with the stretch O-H (3000-3500 cm^{-1}), stretch C-O (1078 cm^{-1}) and stretch C-O-C (1165 cm^{-1}) (Sarangi et al., 2006). In the same way, protein patterns were also noted with absorption at 1664 cm^{-1} and 1527 cm^{-1} (Gonzaga et al., 2005). Thus, because of the existence of bands in the spectra at 3430, 2920, 2850, 1407–1475, 1160 and 1078 cm^{-1}, such fractions were detected as possessing a pyranose ring. The existence of protein in the samples was explained by the presence of 1664 cm^{-1} (amide I) and 1527 cm^{-1} (amide II). Furthermore, FTIR spectra (1024 cm^{-1} and 867 cm^{-1}) indicated that a beta-glycosidic connection appeared in the three fractions (Sarangi et al., 2006).

This introductory chapter describes the theoretical information currently available in the literature regarding the use of an efficient, and environmentally friendly technique (HPE) for extracting bioactive extracts and compounds, as well as to presenting several examples of fungi exhibiting antibacterial and antioxidant activities, characterized using well-known techniques, such as NMR and FTIR.

4.2 Materials and Methods

4.2.1 Fungi Samples

Samples of 6 fungi species, including 5 understudied mushrooms, from pre-existing cultures kept at the Department of Chemistry, were used. These were *Lentinula edodes*, *Lentinus sajor-caju, Phanerochaete chrysosporium*, *Pholiota nameko*, and *Trametes versicolor*. One more understudied mushroom, named *Pleurotus ostreatus* was obtained from Faculdade de Ciências Agrárias of Universidade Estadual Paulista – UNESP, São Paulo Brasil. All fungi samples were stored at -20 ºC.

4.2.2 Growth and isolation of fungi samples in laboratorial conditions

A batch reactor methodology was applied for growth and proliferation of six microorganisms. Thus, in order to cultivate and reproduce in laboratory fungi species using this methodology, *Pholiota nameko*, *Trametes versicolor*, *Lentinula edodes*, *Lentinus sajor-caju* and *Phanerochaete chrysosporium* were grown at 25°C in a growth medium containing 20 g/L of glucose, 20 g/L of malt extract (Oxoid®, UK) and 1 g/L of peptone (Rao, 2012), as well as 35 g/L of sea salts (Sigma-Aldrich, 2016), under stirring for 5 days prior to removal assays. *Pleurotus ostreatus* was sub-cultured and maintained on Potato Dextrose Agar (PDA; Himédia®, Índia) at 4 ºC. Growth was performed in sterilized medium containing 2 g/L glucose (Riedel-Haen®, Germany), 2 g/L starch (Absolve®, Portugal), 0.1 g/L peptone (Himédia®, India) and fragments of wheat straw (1%) for 7 days at 25 ºC.

Batch reactors were incubated (HWY-200D, Lan Technics, USA) in the dark and stirring was maintained at 120±10 rpm for a maximum of 28 days. Temperature was kept at a constant 25°C. At the end of the experiment, mycelia were freeze-dried using a bench-top freeze dryer (ScanVac, 55-4). After growth, mycelia were collected by filtration with sterilized gaze and kept in sterile plastic containers at 4 ºC, for a maximum of 24 hours, until further use.

4.2.3 DNA isolation, amplification, and identification of fungi samples

The nucleotide sequences for the six fungi samples mentioned in the previous section were determined and obtained from the genomic DNA used for molecular typing of all isolates. The identification of the strain was confirmed through a phylogenetic analysis of sequences of the rDNA internal transcribed spacer (ITS) region. Representative isolates of each group were selected and subjected to Polymerase Chain Reaction (PCR) amplification. The PCR amplification conditions of the ITS region of the rDNA was as described by Gonçalves et al. (2019). Basic Local Alignment Search Tool (BLAST) search for bioinformatics analysis, against the nucleotide collection database using the ITS sequences was carried out to determine the homologous nucleotide sequences, which

were added to the sequence alignment (https://blast.ncbi.nlm.nih.gov/Blast.cgi) (Chaudhary & Tripathy, 2015). On the basis of this information, species status was assigned.

4.2.4 Extraction of Bioactive Extracts

For the extraction of bioactive extracts from fungi samples, the process was used the High-Pressure Extraction equipment (HPE, Model 55, Hyperbaric, Burgos, Spain), with a pressure vessel of 55 L, connected to a refrigeration unit (RMA KH 40 LT, Ferroli, San Bonifacio, Italy) in order to control the temperature of the input water used as pressure-transmitting fluid.

The process for extraction of fungi consisted in the re-suspension of 2 g of lyophilized mycelium in 160 mL of ethanol at 22% (v/v) in a flexible bag that will be later vacuum-sealed. Then, the resulting mixture was subjected to a pressure of 200 MPa during 1 min. After, the extracts were filtered through glass fiber filters (GFF, Whatman), and then evaporated in a rotary evaporator (Laborata 4000 WB, Heidolph) under vacuum at 40 ºC until complete desiccation with a final volume of approximately 20 mL. The extraction yield was calculated as the weight percentage of lyophilized extract (M) per dried mycelium sample (W) as given in Equation (1):

$$Extraction\ yield\ (\%) = (M/W) \times 100\% \quad (1)$$

4.2.5 Bioassays for Bioactivity Screening

4.2.5.1 Antibacterial activity

The disk diffusion method was performed using square plates (20x20 cm) containing Mueller-Hinton agar (BD, Heidelberg, Germany) and inoculated with a bacteria suspension (*Escherichia coli* ATCC® 25992*, Klebsiella pneumoniae* ATCC® 13883*, Pseudomonas aeruginosa* ATCC® 9027*, Enterococcus faecalis* ATCC® 19433*, Kocuria rhizophila* ATCC® 9341 and *Staphylococcus aureus* ATCC® 6538). For each inoculum was prepared an individual plate (Figure 4.2). Inoculated plates were allowed to dry before paper disks were applied to the surface of the agar containing 15µL fungal extracts. Prepared plates were incubated at 35 ºC – 37 ºC for 18h. After the incubation period, plates were examined to verify the inhibition. A positive result was defined as an inhibition zone (halo size) around the holes, therefore indicating the presence of antibacterial substance in the extracts tested. Amoxicillin (AML) (for *L. edodes*, *Ph. nameko* and *P. ostereatus*) and tigecycline (TGC) (for *T. versicolor*, *Ph. chrysosporium* and *L. sajor-caju*) were used as standard. An additional study was performed using a standard mixture of amoxicillin plus clavulanic acid in the proportion of 30/10 µg applied to all extracts to observe inhibition in *E. Coli*.

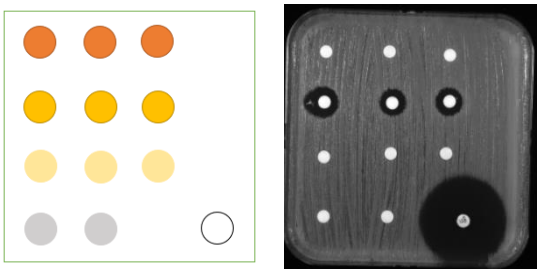

Figure 4.2. Disk diffusion method: (a) scheme-example of each plate; (b) real example.

4.2.5.2 Antioxidant activity

The following antioxidant procedures were adapted from Vamanu et al. 2012 and Microlab protocol.

a) DPPH-free radical scavenging assay

A total of 22 µL of mycelial extract solution and 200 µL of 120 µM DPPH solution dissolved in 80% ethanol were put into a 96-well microplate and then left in the darkness at room temperature for 30 min. The absorbance was measured at 517 nm (A1). The absorbance of solvent instead of mycelium extract solution was measured as A0. Each sample was determined in quintuplicate, and each absorbance was compared to a control without extract. Ascorbic acid as standard (1 mg/mL) were used to calculate calibration curve. The percentage of scavenging activity was determined using the following formula (2):

$$\text{Scavening } \% = \left(\frac{\text{A0}-\text{A1}}{\text{A0}}\right) \times 100\% \qquad (2)$$

A1 sample

A0 control

b) Phenolic compounds quantification

A total of 20 µL of mycelial extract solution, 90 µL of distilled water and 10 µL of Folin-Ciocalteau reagent solution were put into a 96-well microplate and left in the darkness at room temperature for 6 min. Then, 80 µL of 7 % sodium carbonate solution were added to each well and it was incubated again in the dark at room temperature for 2 hours. The absorbance was measured at 750 nm. Each sample was determined in quintuplicate. Gallic acid as standard (1 mg/mL) were used to calculate calibration curve.

c) Ortho-phenols quantification

A total of 160 µL of mycelial extract solution and 40 µL of 5 % sodium molybdate solution were put into a 96-well microplate and then left in the darkness at room temperature for 15 min. The absorbance was measured at 370 nm. Each sample was determined in quintuplicate. Gallic acid as standard (1 mg/mL) were used to calculate calibration curve.

d) Flavonoids quantification

A total of 60 µL of mycelial extract solution (20 mg/mL of 80% ethanol) and 28 µL of 5% sodium nitrite solution were put into a 96-well microplate and then left in the darkness at room temperature for 6 min. Then, 28 µL of 10% aluminum chloride solution were added to each well and incubated again for 6 min. After that, 120 µL of 4% sodium hydroxide solution were added and gently shook. The absorbance was measured at 370 nm. Each sample was determined in quintuplicate. Catechin/Quercetin as standard (1 mg/mL) were used to calculate calibration curve (twelve concentrations determined in triplicate).

4.2.6 Characterization of bioactive extracts via elemental analysis (CNHS), and nuclear magnetic resonance (NMR) and Fourier transform infrared (FTIR) spectroscopies

Organic elements C, Hydrogen (H), Sulfur (S) and N in fungi samples and respective lyophilized extracts were quantified using a Truspec 630-200-200 Elemental Analyser (Mönchengladbach, Germany). Samples of up to 3 mg for each extract were placed under combustion at 1075 ºC. Carbon, H and S were detected by infrared absorption whereas N was detected by thermal conductivity.

FTIR and ^{13}C-NMR samples were prepared by lyophilizing the fungal biomass and using solid mycelium for analysis.

FTIR spectra were recorded on a Perkin Elmer Spectrum BX FTIR device (Perkin Elmer, USA), using a Golden Gate single reflection diamond ATR system (Specac Lda, USA). All spectra were obtained using 64 scans each and a resolution of 4 cm^{-1}, within the 4000–550 cm^{-1} range. Air was used for the background spectrum.

The ^{13}C-NMR spectra were obtained using a Bruker Avance III 300 MHz spectrometer (Karlsruhe, Germany) with an operating frequency of 300.13 MHz. Spectra were acquired with a spinning rate of 20 Hz, a contact time of 4.75s, and with the pulse program, ZG30. The recycle delay was 1s and the length of the proton 90 pulses was 9.00 µs. About 56 scans were collected for each spectrum. A 0.3 Hz line broadening weighting function and a baseline correction were applied. The identification of functional groups in the NMR spectra was based on their chemical shift (δH) relative to that of the water (4.7 ppm).

4.2.7 Statistical analysis

The analysis of experimental samples results was expressed as the mean value ± standard deviation (SD) for bioactive tests (n=5). Statistical analysis was performed using One-way Analysis of Variance (ANOVA) using the Minitab Statistical Software v.18. The Pearson correlations were evaluated by the Pearson's correlation coefficient (R) and the statistical significance of the coefficient (*p*-value) using Minitab Statistical Software v.18. Differences at $p<0.05$ were considered significant.

4.3 Results and Discussion

4.3.1 Identification of fungi samples

The BLAST finds regions of local similarity between sequences. The program compares nucleotide or protein sequences to sequence databases and calculates the statistical significance of matches. BLAST can be used to infer functional and evolutionary relationships between sequences, as well as help identify members of gene families. Then, it can be found the nucleotide sequences of the field isolated fungi samples.

>Lentinula edodes

CATTATTGAATTTTTTGGTGGTGGATTGTTGCTGGCCTTTGGGTATGTGCACATCCTCCTCCGA
TTTCTATTCATCCACCTGTGCACTTTTTGTAGGAGTTCTTTCATCGGGTTTTTGAAGGTGCTCAT
TATGAGTTACTTGAAAAGACTAGTTGACAAGGCTTCTATGTTCTTATAAACCATTGAAGTATGTT
ATAGAATGATCTTGTTATTGGGACTTTATTGACCCTTTAAACTTAATACAACTTTCAGCAACGGA
TCTCTTGGCTCTCCCATCGATGAAGAACGCAGCGAAATGCGATAAGTAATGTGAATTGCAGAA
TTCAGTGAATCATCGAATCTTTGAACGCACCTTGCGCCCTCTGGTATTCCGGAGGGCATGCCT
GTTTGAGTGTCATTAAATTCTCAACTTTATAAGTTTTTACTTATTAAAGCTTGGATGGTGGAGGC
TTGCAGGCGTTTGTCAGCTCCTCTTAAATTTATTAATGGGAACCCTGGTTTGGTAGTTCTAACC
TTGGTGTGATAATTATCTACATTTTGGTGGAACCTTACAATAATAAAGCTCTATTGGTTTGGGTT
GGTGCATTTAGTTTGCTCAATCTGGTCTATTCATTGGAGAAAAAGGGAAGTTCCGCTTTCTAAC
TGGCTTGATTGACTATATATAACTTATTTGCTTGACCTCAAATCAGGTAGGATTACCCGCTGAA
CTTAA

Identification sample name result: Lentinula edodes

Percentage of identity: 99.02%

> Lentinus sajor-caju (=Pleurotus sojar caju)

CATTAATGAATTCACTATGGAGTTGTTGCTGGCCTCTAGGGGCATGTGCACGCTTCACTAGTC
TTTCAACCACCTGTGAACTTTTGATAGATCTGTGAAGTCGTCCTTCAAGTCGTCAGACTTGGTT
TGCTGGGATTTAAACGTCTCGGTGTGACAACGCAGTCTATTTACTTAACACACCCCAAATGTAT
GTCTACGAATGTCATTTAATGGGCCTTGTGCCTATAAACCATAATACAACTTTCAACAACGGAT
CTCTTGGCTCTCGCATCGATGAAGAACGCAGCGAAATGCGATAAGTAATGTGAATTGCAGAAT
TCAGTGAATCATCGAATCTTTGAACGCACCTTGCGCCCCTTGGTATTCCGAGGGGCATGCCTG
TTTGAGTGTCATTAAATTCTCAAACTCACATTTATTTGTGATGTTTGGATGTTTGGGGGGTTGCT
GGCTGTAACAAGTCGGCTCCTCTTAAATGCATTAGCAGGACTTCTCATTGCCTCTGCGCATGA

TGTGATAATTATCACTCATCAATAGCACGCATGAATAGAGTCCAGCTCTCTAATCGTCCGCAAG
GACAATTTGACAATTTGACCTCAAATCAGGTAGGACTACCCGCTGAACTTAA

Identification sample name result: Pleurotus pulmonarius or Lentinus sajor

Percentage of identity: 99.68%

>Phanerochaete chrysosporium

CATTAACGAGTAACTGAACAGGTTGTAGCTGGCCTCTCGGGGCATGTGCACGCCTGGCTCAT
CCACTCTTCAACCTCTGTGCACTTGTTGTAGGTCGGTAGAAGAGCGAGCATCCTCTGATGCTT
TGCTTGGAAGCCTTCCTATGTTTTACTACAAACGCTTCAGTTTAAGAATGTCTACCTGCGTATA
ACGCATCTATATACAACTTTCAGCAACGGATCTCTTGGCTCTCGCATCGATGAAGAACGCAGC
GAAATGCGATAAGTAATGTGAATTGCAGAATTCAGTGAATCATCGAATCTTTGAACGCACCTTG
CGCTCCCTGGTATTCCGGGGAGCATGCCTGTTTGAGTGTCATGGTATCCTCAACCTTCATAAC
TTTTTGTTATCGAAGGCTTGGACTTGGAGGTTGTGCTGGCTTCTAGTCGAGTCGGCTCCTCTT
AAATGTATTAGCGTGAGTGTAACGGATCGCTTCGGTGTGATAATTATCTGCGCCGTGGTCGTG
AAGTAACATAAGCTTGCGCTTCTAACCGTCCTTCAGTTGGACAACTTACTTTGACATCTGACCT
CAAATCAGGTAGGACTACCCGCTGAACTTAA

Identification sample name result: Phanerochaete chrysosporium

Percentage of identity: 100.00%

>Pholiota nameko

CATTATTGAATAAAACTTGGTTGGATTGTTGCTGGCCTGAATGAGGGCATGTGCACATCTGCC
ATCTTTTATCTTTCCACCTGTGCACACTTTGTAGGTCTGGGATTAAACTTTCTGAGGTCAACTC
AGTTTTGAGGACTGCTGTTAGCAATAATGGCTTTCCTTGTCTTTCCAGATCTATGTTTTCATATA
CACCATAAAAATGTAATAGAATGTGTTAATAAGCCTTGTGCTTATAAACTATATACAACTTTCAG
CAACGGATCTCTTGGCTCTCGCATCGATGAAGAACGCAGCGAAATGCGATAAGTAATGTGAAT
TGCAGAATTCAGTGAATCATCGAATCTTTGAACGCACCTTGCGCTCCTTGGTATTCCGAGGAG
CATGCCTGTTTGAGTGTCATTAAATTATCAATCTTTGCAGCTTTTGTTGTTAAAGACTTGGAGGT
GGGGGTTTTAATTTGGAGGCTTTTCGGAGTGTCCTCCCCTAAAATGTATTAGCTAGTTGCTCG
TGCGGACTTGTCTATTGGTGTGATAATTATCTACGCCATGGACAGACTGCCATTAAGTAGCAC
TGCTTCTAATCGTCCTTTACTGGACAACTTATGACAATTTGACCTCAAATCAGGTAGGACTACC
CGCTGAACTTAA

Identification sample name result: Pholiota microspora (=Pholiota nameko)

Percentage of identity: 99.54%

>Pleurotus ostreatus

ACTGCGGAAGGACATTAATGAATTCACTATGGAGTTGTTGCTGGCCTCTAGGGGCATGTGCAC
GCTTCACTAGTCTTTCAACCACCTGTGAACTTTTGATAGATCTGTGAAGTCGTCTCTCAAGTCG
TCAGACTTGGTTGGTGGGATTTTAACGGCTCGGTGGGGCTACCCAATCCAATTA

Identification sample name result: Pleurotus ostreatus

Percentage of identity: 96.45%

>Trametes versicolor

ACTGCGGAAGGACATTAACGAGTTTTGAAACGAGTTGTAGCTGGCCTTCCGAGGCATGTGCA
CGCTCTGCTCATCCACTCTACCCCTGTGCACTTACTGTAGGTTGGCGTGGGCTCCTTAACGG
GAGCATTCTGCCGGCCTATGTATACTACAAACACTTTAAAGTATCAGAATGTAAACGCGTCTAA
CGCATCTATAATACAACTTTTAGCAACGGATCTCTTGGCTCTCGCATCGATGAAGAACGCAGC
GAAATGCGATAAGTAATGTGAATTGCAGAATTCAGTGAATCATCGAATCTTTGAACGCACCTTG
CGCTCCTTGGTATTCCGAGGAGCATGCCTGTTTGAGTGTCATGGAATTCTCAACTTATAAATCC
TTGTGATCTATAAGCTTGGACTTGGAGGCTTGCTGGCCCTCGTTGGTCGGCTCCTCTTGAATG
CATTAGCTCGATTCCGTATGGATCGGCTCTCAGTGTGATAATTGTCTACGCTGTGACCGTGAA
GTGTTTTGGCGAGCTTCTAACCGTCCATTAGGACAACTTTTTAACATCTGACCTCAAATCAGGT
AGGACTACCCGCTGAACTTAA

Identification sample name result: Trametes versicolor

Percentage of identity: 99.66%

All BLAST results of the six fungi samples under study showed at least 96% match with the respective and expected fungi.

4.3.2 Extraction yield

Extraction is the first crucial step in the process to obtain cellular bioactive compounds from natural sources (Sasidharan et al., 2011). Thus, it is expected that HPE technique allying the high hydrostatic pressure (e.g. in our study 200 MPa) with to the easier breakage of the cell wall and consequent higher volume of solvent inside the cell membrane increase extraction yield, while the extraction time (1 min for our study), and the energy required for the process to occur is reduced (He et al., 2011; Moreira et al., 2020a). The process efficiency is quantitatively related to extraction yield (Kitzberger et al., 2007).

Mushrooms in general are of important nutritional value, constituting a great source of vitamins (B1, B2, B12, C, D, and E), minerals (selenium and potassium), carbohydrates (37-48%), proteins (20-25%), dietary fiber (13-24%) and some secondary metabolites, including phenolic compounds, polyketides, terpenes, steroids, carotenoids and flavonoids (Sabaratnam et al., 2011; Vaz et al., 2010). Nevertheless, among all the bioactive compounds of mushroom origin, polysaccharides are the most extensively researched, followed by proteins (Roupas et al., 2012; Xu et al., 2011). In contrast, they are low in fats (4-5%) (Sabaratnam et al., 2011).

The choice of solvent is considered the utmost parameter for any extraction process, being dependent on the solubility of the compounds of interest, as well as its interaction with the sample matrix. Ethanol is one of the most used solvents for the extraction of some bioactive compounds, such as phenolic compounds, especially due to its moderate polarity (ε= 25.5 at 20 °C), easy removal from the final extract by evaporation, non-toxic and low cost (Moreira et al., 2020a; Shouqin et al., 2005).

Despite all fungal samples extracted in this study have undergone the same experimental variables (pressure, temperature, solvent and extraction time), some differences in the extraction yield were observed. As shown in Figure 4.3, it can be observed that the fungus in which the HPE technique was more effective, allowing a higher percentage of extraction was *Lentinus sajor-caju* (97%), followed by *Pleurotus ostreatus* (93%), *Pholiota nameko* (86%), *Phanerochaete chrysoporium* (84%), *Lentinula edodes* (67%) and finally with the lowest extraction percentage *Trametes versicolor* (55%). Although HPE extracts presented some differences in extraction yield values, this technique, in addition to increasing the quantity of bioactive compounds in the extract when compared to other traditional techniques, also gets to preserve its level in terms of natural quality and operationalization, due to its application at lower temperatures.

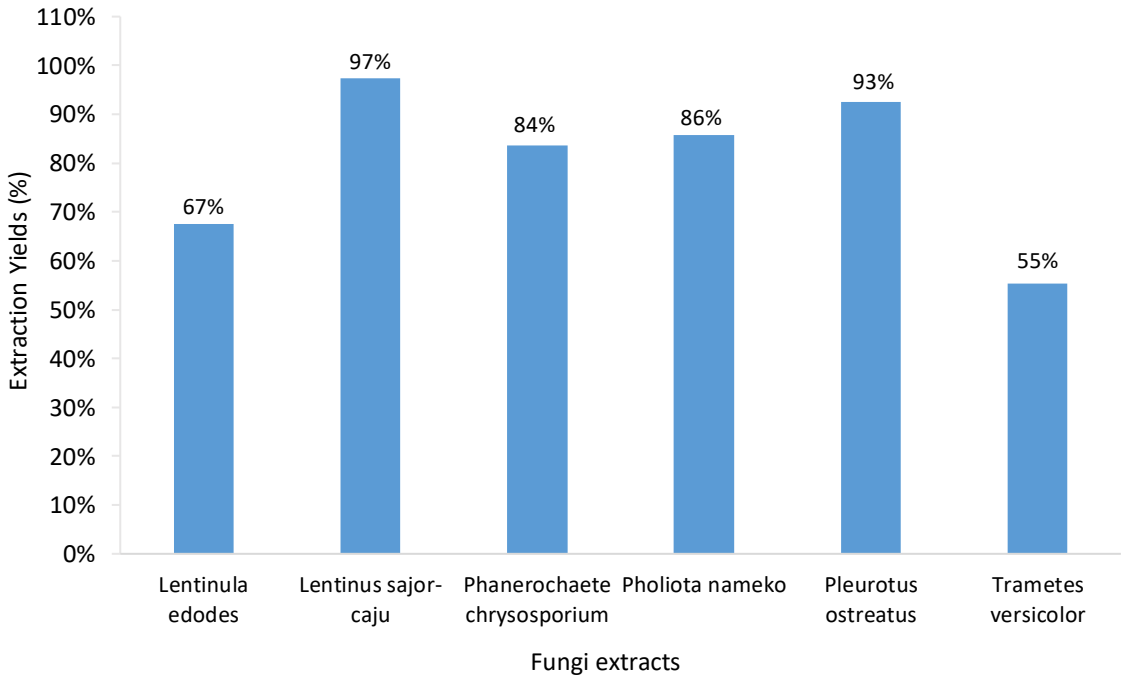

Figure 4.3. Extraction yields of fungi extracts obtained by high-pressure extraction (HPE), (n=1).

The fact that there is any moisture in samples can reduce the extraction efficiency, however, our fungi samples were submitted to the freeze-drying process and the solvent used was polar (Khan et al., 2018). Thus, one of the explanations for this lower extraction yield in *Trametes versicolor* fungus may be the difference in solubility of the active compounds (Miller et al, 2000). However, the replicates are not significantly representative and, the standard deviation associated with the experiment was considerably high (SD equal to 16%) and, therefore, the differences were not statistically significant. Shouqin et al., (2004) showed that the presence of any moisture in the extract may lower extraction efficiency. Depending upon the product the samples can be oven-dried or

freeze-dried. Wet samples can be effectively extracted using polar solvents such as methanol, ethanol, acetonitrile, ethyl acetate, etc. or from a mixture of solvent like hexane/acetonitrile, hexane/acetone, among others, making the drying step less important. Despite using a suitable drying agent, sometimes, water may be co-extracted and can later be used while cleaning up, for preparing extract concentration or in the direct analysis (Shouqin et al., 2004).

Each of the fungal extracts under study was then subjected to bioactivity screening and kept for further structural characterization.

4.3.3 Antibacterial activity

The diffusion methods indicate the active compounds, in this case, which extracts are susceptible to the microorganism under study. Despite the disk diffusion method is sensitive to detect microbial growth, it has a qualitative character and should not be recommended to quantify the antimicrobial activity of a substance based on the size of the inhibition zone formed during the analyses (Kitzberger et al., 2007). Anyhow, six fungi extracts were tested against several gram-positive and gram-negative bacteria (see section 3.2.5.1), using the disk diffusion method in order to provide indication for detection of minimum inhibition concentration.

Table 4.2 shows the results of disk diffusion essays in terms of size of inhibition zone (mm) for the extracts tested against the studied microorganisms. The only fungi extract that showed inhibition action was *Ph. nameko* against *K. rhizophila*, *S. aureus* and *E. coli* (presenting statistical significance ($p<0.05$)), with most positive response against *K. rhizophila*. Despite the *E. coli* was the less resistant microorganism for *Ph. nameko* extract, studies performed by Janeš et al. (2007) revealed that among a wide range of wood-colonizing mushrooms, *Trametes versicolor* was the only one that showed an antibacterial effect, although weak in *P. aeruginosa*. Most of the fungal extracts were not active in the disk diffusion test. This can be the result of a low diffusion of active compound into the agar due to low solubility or high molecular mass (Janeš et al., 2007). Further studies performed using a standard mixture of amoxicillin plus clavulanic acid as an antibiotic, revealed that none of the extracts showed antibacterial activity against the bacteria *E.coli*.

Table 4.2. Bioactive tests for *Pholiota nameko* against gram-positive and negative bacteria's.

	AML (mm)	TGC (mm)	*Pholiota nameko*			
			Replicate 1 (mm)	Replicate 2 (mm)	Replicate 3 (mm)	Average ± Standard deviation
Escherichia coli	20	21	9	9	9	9 ± 0
Klebsiella pneumoniae	0	20	0	8	0	3 ± 5
Pseudomonas aeruginosa	0	21	0	0	0	NAP
Enterococcus faecalis [(a)]	31	22	0	0	0	NAP
Kocuria rhizophila [(b)]	42	34	14	12	13	13 ± 1
Staphylococcus aureus [(c)]	23	22	10	12	13	12 ± 2

***a), b) c)** Different letters indicate significant differences (p<0.05) between species.

NAP – Not applicable

4.3.4 Antioxidant activity

Numerous studies have been investigating the mechanisms of action of free radicals, as well as to discover effective substances towards preventing and even reversing the occurrence of oxidative damages (Andreu et al 2018; Diaz et al., 2012; Luna-Guevara et al., 2018; Tungmunnithum et al., 2018). Not only synthetic but also natural antioxidants have proved to be highly effective to control the magnitude of free radicals' production, to prevent its undesirable effects, as well as to support the organism antioxidant and detoxifying mechanisms (Holst & Williamson, 2008, Kapravelou et al., 2015, Valko et al., 2007, Yeh & Yen, 2006). The study of the antioxidant potential of phenolic extracts derived from plant species is one of the hot topics among the scientific community; being the *in vitro* studies the most common (Dai & Mumper, 2010, Larrosa et al., 2010, Rubió et al., 2013). Thus, in this study, the antioxidant activity of six fungi extracts were measured through the DPPH assay, which evaluates the hydrogen-donating or RSC of extracts, following the procedure reported before in section 4.2.5.2. This method was choice due to its simplicity, sensitivity, and reproducibility. The antioxidant activity was expressed as a percentage of DPPH radical reduction.

DPPH is a free radical, stable at room temperature, which produces a violet solution in ethanol. In presence of antioxidant compounds, the DPPH is reduced producing a non-color ethanolic solution (Kitzberger et al, 2007). The extracts that showed higher antioxidant activity were *L. sajor-caju* (77.6%), *T. versicolor* (76.8%), *Ph. chrysosporium* (74.6%) and *P. ostereatus* (72.7%), in this order, being the *Ph. nameko* fungus the one with the least activity showing only 16.7% (Figure 4.4). High yields are probably due to the presence of polar substances in the extracts responsible for the cited activity. On the other hand, low yields are possibly caused by the non-polar characteristic of the solvent, resulting in the extraction of mainly non-polar components, with low antioxidant activity. The

efficiency of the extraction method is an important factor in determining antioxidant activity (Kitzberger et al, 2007). Results performed by Zhu et al. (2019) indicated that the scavenging activities towards DPPH enhanced with the increasing concentration of polysaccharides. The *in vitro* antioxidant results using extracts from *Ph. nameko* and *L. edodes* suggest that polysaccharides, such as PNP and linear (1→6)-*β*-*D*-glucan, respectively, have antioxidant capacity and it can play a role in protecting biofilm and anti-ageing (Fan et al., 2020; Morales et al., 2020; Zhu et al., 2019). Reduction of phenolic amounts may occur due to isomerization of certain compounds caused by the high temperatures (Minatel et al., 2017). In plants, the phenolics can be found linked to the cell membranes/walls or can be free, and the food processing methods, such as the use of high temperatures or freezing, can cause the release of these compounds, which is implied by an increase of its bioavailability in the human body. Some reports show that heating affects the content of some polyphenols, including flavonoids, due to the extractability alteration by the rupture of the cell wall. In this way, polyphenols linked to the wall could be released more easily on cooking than from the raw material (Dewanto et al., 2002; Jeong et al., 2004). Thus, the extracted total phenolic compounds increased obviously with increasing extraction temperature. Despite the *Ph. nameko* fungus was the one with the lowest activity in our study, potential antioxidant effects of the polysaccharide extract on this fungus have been however reported by Zhang et al. (2015). Recently, a ~43 kDa antioxidant protein was isolated from *Ph. nameko* exhibiting potential antioxidant activity (Qian et al., 2016).

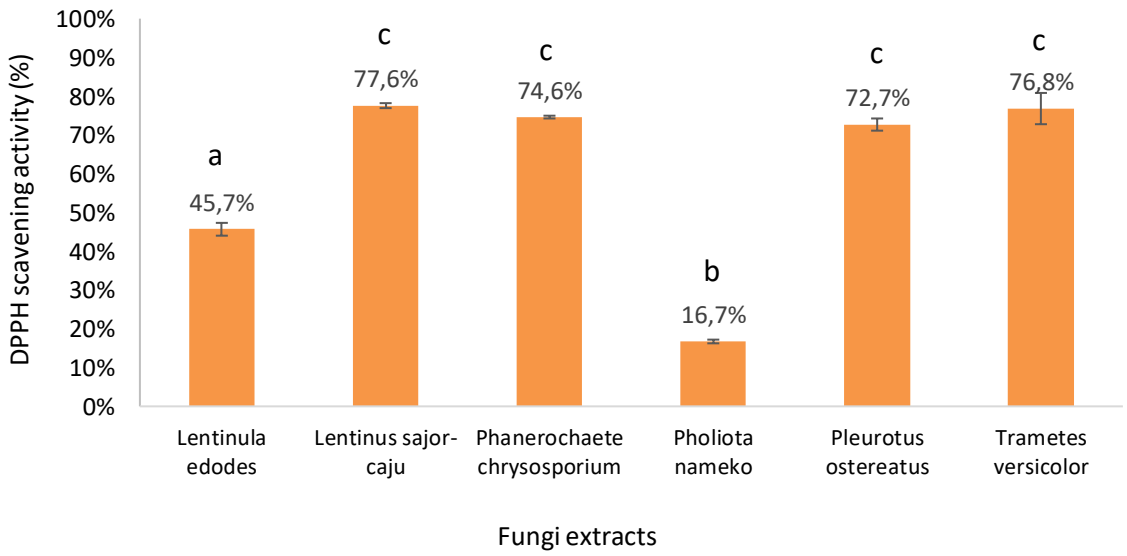

Figure 4.4. DPPH scavenging activity of fungi extracts. Different letters indicate significant differences (p<0.05) between species (n=5).

The inhibitory concentration of the extract required to decrease the initial DPPH radical concentration by 50% (IC_{50}) was determined from the graph of DPPH reduction percentage in

function of extracts concentration by the regression linear of gallic acid as standard with the equation: y=-0.0534x+0.7007, R^2=0.8661. Based on the R^2 value, the linear relationship can indicate some limitations. The *L. sajor-caju* and *T. versicolor* extracts that has a lower IC_{50} and, therefore, has a greater antioxidant capacity. This difference between the extracts may be due to the extraction yield, it may have been possible to extract a greater quantity of bioactive compounds for the extracts *L. sajor-caju* and *T. versicolor*, respectively. It should be noted that, since the antioxidant capacity is greater the lower the IC_{50} value, if there is any relationship, a negative correlation is expected between the IC_{50} and the content of total phenolics and flavonoids.

The antioxidant activity of fungi extracts has been correlated to their content of phenolic components, ortho-phenols and flavonoids due to their property of scavenging free radicals (Fernandes et al., 2017; Tungmunnithum et al., 2018). Therefore, it is important to consider the effect of the total phenolic and flavonoids quantity in the antioxidant activity of the fungi extracts. The results obtained in our study are in accordance with the principle that a higher phenols and flavonoids content is associated with a higher antioxidant power, with greater prominence for *Ph. chrysosporium*, followed by *T. versicolor*, *L. sajor-caju*, and *P. ostereatus* (Figure 4.5), showing statistical significances (p<0.05). High correlations between total phenolic content (TPC), total flavonoids content (TFC) and DPPH for ethanol extracts were observed by Fernandes et al. (2017). These correlations point out the important roles of flavonoids and polysaccharides in the ethanol extracts. This behavior is probably due to the ethyl acetate capacity to solubilize flavonoid components from the fungi samples, substances detected by the Folin-Ciocalteu method (Fernandes et al., 2017). Dependent upon flavonoids content of glycosides, isoprenoids, and aliphatic ethers, the flavonoids have almost any polarity and, thus, are soluble in a range of solvents (Avello et al., 2013, Havsteen, 2002, Taârit et al., 2012, Trabelsi et al., 2012).

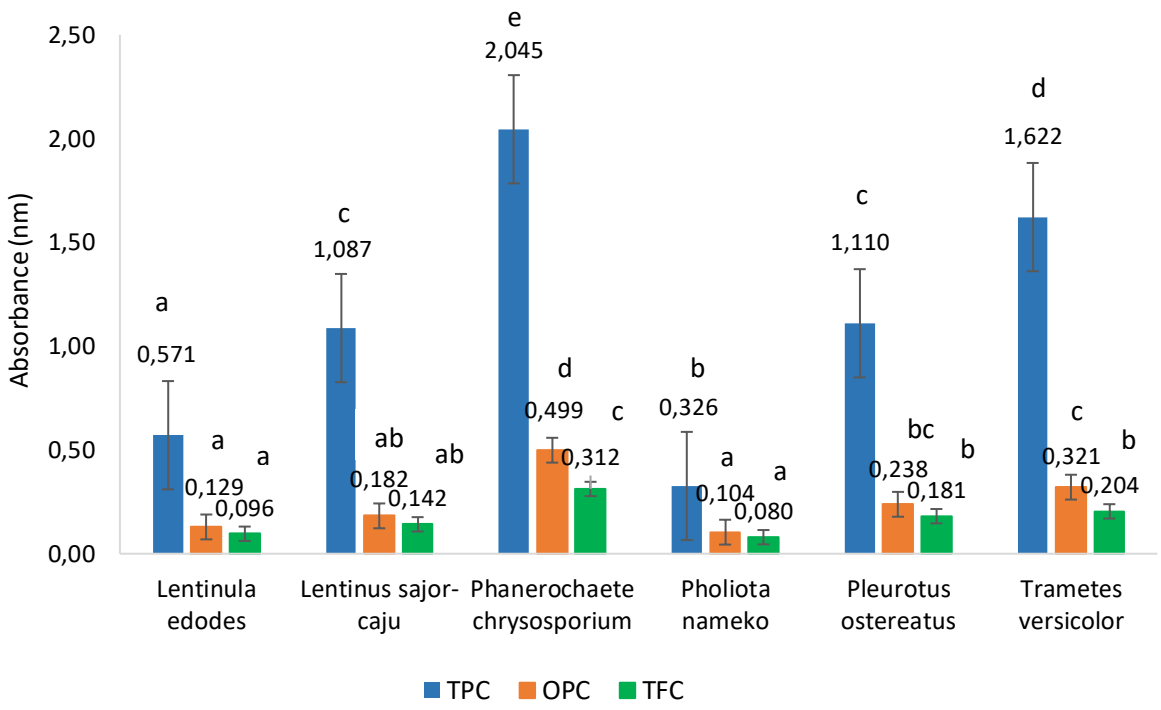

Figure 4.5. Antioxidant activity of different fungi extracts based on total phenolic content (TPC), total flavonoids content (TFC) and ortho-phenolic content (OPC). Different letters indicate significant differences (p<0.05) between species (n=5).

The antioxidant activity of phenolic compounds is attributed to the capacity of scavenging free radicals, donating hydrogen atoms, electrons, or chelate metal cations. Molecular structures, particularly the number and positions of the hydroxyl groups, and the nature of substitutions on the aromatic rings, confers to phenolic compounds the capacity of inactivating free radicals, which is referred to as structure-activity relationship. Among the several classes of phenolic compounds, the phenolic acids, flavonoids, and tannins are regarded as the main phenolic compounds, nevertheless, anthocyanins, anthraquinones, catechins, chalcones, xanthones, stilbenes, and naphthoquinones are also found (Minatel et al., 2017; Soković & Liaras, 2021). Flavonoid and phenolic compounds are known to act as antioxidants, radical scavengers, and metal chelators (Mira et al., 2002). Thus, these fungal extracts under study seem to present appreciable amounts of flavonoid and phenolic compounds that may contribute to an antioxidant ability, corroborating the theory. Studies performed by Diaz et al. (2012) evidenced that the water extracts from several herbs that had the highest phenolic content also had the highest flavonoid content, showing a positive correlation between the content of phenolic compounds and flavonoids, which is equally observed in our results. Phenolics and flavonoids neutralize free radicals by donating a hydrogen atom or an electron and chelate metal ions (Petti & Scully, 2009). The ethanol extracts obtained by Diaz et al. (2012) contained a higher proportion of flavonoids than phenolics (Diaz et al., 2012). On the other hand, Janjušević et al. (2017) verified that all three crude extracts along with polysaccharide (PSH) fraction of *T. versicolor* showed

better results with water than ethanol. The difference may have been caused by the choice of the solvent concentration used during extraction (Janjušević et al., 2017).

As stated in literature till date by several researches, the RSC of fungal extracts on DPPH increases with increasing concentrations of extracts (Karaman et al., 2010; Kozarski et al., 2012; Zhu et al., 2019). It has been well established that, phenolic substances are strongly associated with antiradical activity (Karaman et al., 2010; Jeong et al., 2009; Kim et al., 2008; Ferreira et al., 2009). Actually, the antiradical activity depends on the basic structural organization of the phenolic compounds (Mathew et al., 2015; Zhang et al., 2014). Phenolic compounds comprise one (phenolic acids) or more (polyphenols) aromatic rings with attached hydroxyl groups in their structures (Minatel et al., 2017). Quercetin and rutin, derivatives of quinic acid, showed high scavenging DPPH activity in studies performed by Zhang et al. (2014). Both the substituents on the phenyl ring and conjugated carbon skeleton were previously found to be of importance for this antiradical activity (Mathew et al., 2015; Sroka, 2014). Consequently, quercetin, kaempferol, and myrcetin (flavonols with a free 3-hydroxy group) exhibited better antiradical activity compared to flavonoids with glycosylated OH group (Mathew et al., 2015; Sroka, 2014; Zhang et al., 2014). Sroka (2014) emphasized potent anti-DPPH activity of gallic and caffeic acids, compounds that were identified in *Trametes versicolor* water extract (Sroka, 2014).

4.3.5 Characterization by elemental analysis (CNHS), nuclear magnetic resonance (NMR) and Fourier transform infrared (FTIR) spectroscopies

Considering that plant extracts are usually a combination of several types of bioactive compounds or phytochemicals with different polarities, their separation still remains a big challenge for the identification and characterization process of these compounds (Sasidharan et al., 2011).

a) Elemental analysis (CNHS)

The elemental concentrations of fungal biomass varied from 36.87% to 49.68% for %C, 0.62% to 4.82 % for %N, and 5.53% to 6.85% for %H. Regarding to the extracts, the values varied from 35.50% to 40.88%, 0.76% to 6.71%, and 5.14% to 7.38%, respectively.

The organic elemental contents are very similar in fungi and their respective extracts (Figure 4.6), although a slightly higher content in C, N and H is observed for the *Phanerochaete chrysosporium* fungus (49.68%, 4.82% and 6.85%, respectively); on the other hand, in the extracts it was registered a higher C content in *Pleurotus ostreatus* (40.88%), N in *Trametes versicolor* (6.71%) and H in *Lentinula edodes* (7.38%). In summary, it is possible to observe that the element carbon is the one that clearly dominates in each of the fungi and respective extracts analyzed, and in contrast, nitrogen is the one that is less represented. Differences in fungal physiology may have

important influences on large scale C and N cycling (Waring et al., 2013). However, relatively little is known about fungal stoichiometry (Gadd, 2004, 2007). Most plant species form symbioses with soil fungi, and up to 80% of plant N is provided by mycorrhizal fungi, since the mycorrhizal fungi explore the soil volume searching nutrients (Behie & Bidochka, 2014; van der Heijden et al., 2015).

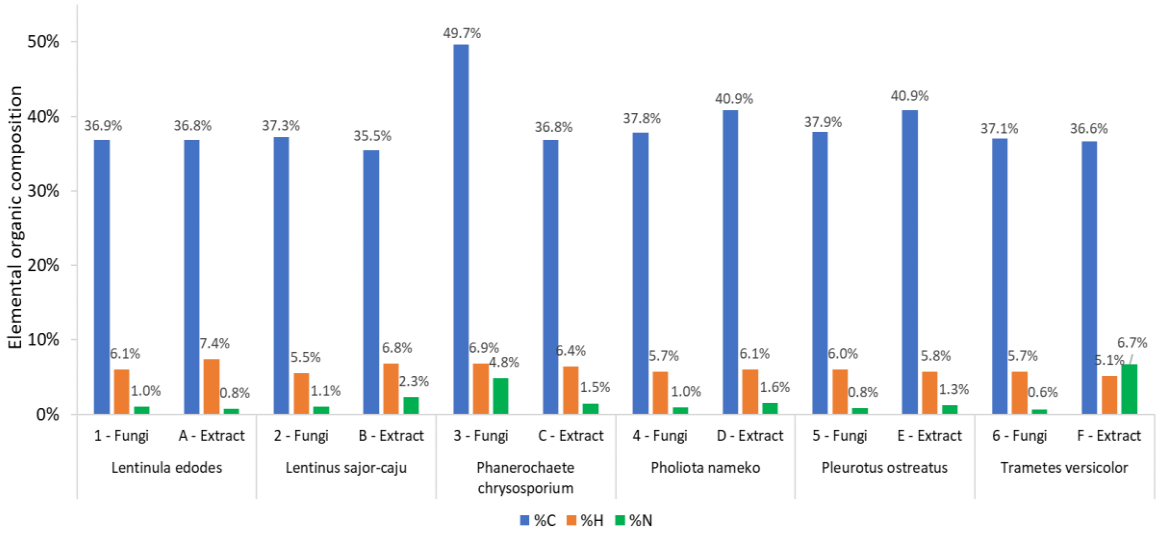

Figure 4.6. Elemental organic composition (Carbon, Hydrogen and Nitrogen) of different fungi extracts under study (n=1).

b) Nuclear magnetic resonance (NMR) spectroscopy

The 13C-NMR spectra of fungi samples and the respective extracts which had been extracted from the HPE methodology were recorded and shown in Figure 4.7. The fungi raw material is composed of many diverse constituents, so the assignment of the peaks obtained is often difficult. Nevertheless, from the spectra of the fungi samples and extracts shown in Figure 4.7, the pattern of carbohydrate structure is clearly defined. According to Chen et al. (2000) and Gonzaga et al. (2005) the fungi spectra were dominated by the carbohydrate peaks (100 and 70 ppm).

The general representative peaks of carbohydrate were assigned into the following regions: 0-40 ppm corresponds to aliphatic C and, according to Gonzaga et al. (2005), the presence of additional peaks in the range 20–40 ppm may suggest the consideration of the presence of glucan–protein structure. Peaks in these regions were more noticeable in only one fungi sample *Ph. chrysosporium* and three fungi extracts, named *L. edodes*, *Ph. nameko*, and *P. ostreatus*. At 40-60 ppm to methoxyl lignin C, 60-90 ppm to hydroxyl carbohydrate C (very defined range in all fungi spectra showing two peaks and only one peak in extracts spectra), 90-120 ppm to aromatic C (more evident in fungal samples than in extracts), 120-140 ppm to C substituted aromatic C (not so highlight in our spectra), 140-160 ppm to phenolic aromatic C (no relevant peaks), and 160-220 ppm mainly

to C in carboxyl groups, and to a limited extent to aldehyde C (Chen et al., 2000; Santos et al., 2021). Chemical shifts at 160–180 ppm in Figure 4.7 (a, b, c and d) showed a signal, suggesting the presence of uronic acid (Zhu et al., 2019), and it is consistent with the results obtained in FTIR spectra. Peaks between 100 and 104 ppm, suggest detections of anomeric carbon sign of both *α* and *β* configurations, respectively (Gonzaga et al., 2005; Zhu et al., 2019). Comparison of the spectra showed *α* and *β* glucan signals, being more significant in fungal samples than in extracts, and being more abundant the *α*-glucan, which is consistent with the results obtained by FTIR.

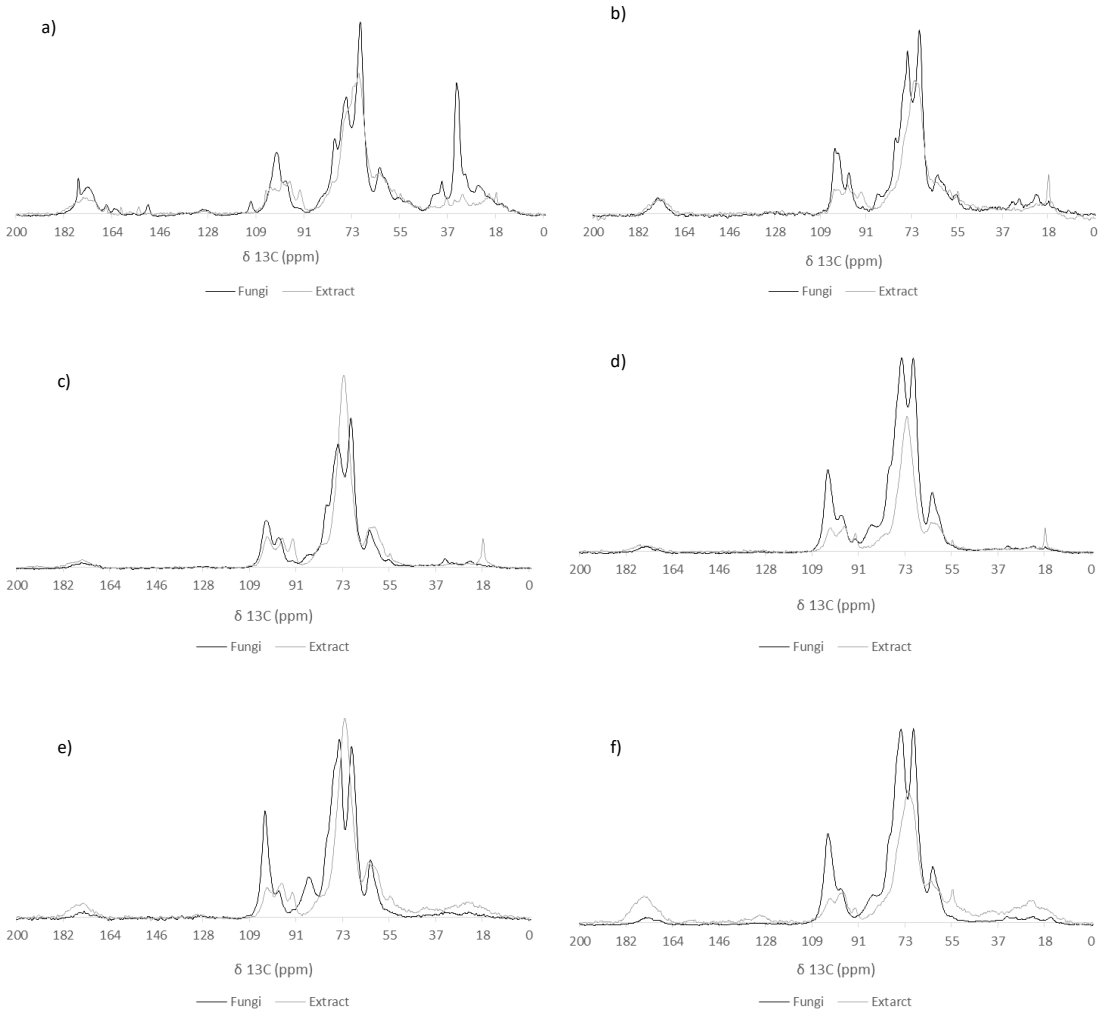

Figure 4.7. The 13C-RMN spectra of extracts and fungo of a) *Phanerochaete chrysosporium*, b) *Lentinula edodes*, c) *Pholiota nameko*, d) *Pleurotus ostreatus*, e) *Lentinus sajor-caju* and f) *Trametes versicolor*.

c) Fourier transform infrared (FTIR)

Mushrooms are rich in polysaccharides, followed by protein content and with low fat content (Alam et al., 2008; Sabaratnam et al., 2011), however, a more specific characterization of these compounds is an additional resource for establishing structure-bioactivity relationships.

Some differences, in terms of qualitative and absorbance intensity are observed between spectra of the fungi and respective extracts, with greater prominence for the *Phanerochaete chrysosporium*, *Pholiota nameko* and *Trametes versicolor* fungi (Figures 4.8a, 4.8c and 4.8d). Four specific regions, 4000-1800 cm^{-1}, 1800-1500 cm^{-1}, 1500-750 cm^{-1} and 950-750 cm^{-1} well known are described in literature and observed in all our spectra (Zhao et al., 2006a, Zhao et al., 2006b). However, there are several authors who describe the chemical characteristic of specific regions of the FTIR spectra.

Considering the region between 4000-1800 cm^{-1}, with focus in 3500 cm^{-1}-3000 cm^{-1} stretching, a pronounced broad band centered around 3280 cm^{-1} is detected in all spectra, that could be assigned to O-H stretching vibrations of glyosidic structures (Pop et al., 2018; Sarangi et al., 2006; Zhao et al., 2006a, 2006b; Zhu et al., 2019) and C-H stretching vibrations at 2980 cm^{-1} (Seedevi et al., 2019) and at 2932–2922 cm^{-1} (Klaus et al., 2015). Seedevi et al. (2019) obtained a polysaccharide from *P. sajor-caju* which displayed a broad stretching peak at around 3410 cm^{-1}, indicating the existence of the OH group in the molecular structure. Moreover, according to Klaus et al. (2015), characteristic N-H vibration around 3400 cm^{-1} could be overlapped by O-H stretch vibration around 3500-3000 cm^{-1} of inter- and intra-hydrogen bonds that are present in polysaccharides (Klaus et al., 2015; Sadhana & Malini, 2018). An important structural polysaccharide of the cell wall of mushrooms is chitin which is the major component of crude fiber (Klaus et al., 2015).

Bands around 2920-2850 cm^{-1} assigned to CH_2 and CH_3 stretching of fatty acids from the cell wall (Pop et al., 2018; Zhao et al., 2006a, 2006b; Zheng et al., 2014) are more defined in the fungi spectrum than in the respective extracts spectra (Figure 4.8), which could result from loss of fatty acids as consequence of the extraction process.

In the region between 1800-1500 cm^{-1} composed of the vibrational mode of carbonyl (C=O) and the C=C double bond are usually found in mushrooms spectra. In addition, a band around 1740-1730 cm^{-1} that could correspond to carbonyl stretching vibration of alkyl-esters may indicate the presence of oil in fungi, however, it is not perceived in the spectra obtained (Gupta et al., 2015; Pop et al., 2018; Zhao et al., 2006a, 2006b; Zhu et al., 2019). Two major bands around 1646 and 1560 cm^{-1} assigned to amide I and amide II of proteins (Klaus et al., 2015; Sadhana & Malini, 2018; Zhao et al., 2006a, 2006b) are visible in *Phanerochaete chrysosporium* spectra, but not in the respective extracts. Likewise, protein patterns were also observed with absorption around 1660 cm^{-1} (amide I) and 1520 cm^{-1} (amide II) by Sarangi et al. (2006) and Gupta et al. (2015). A single band between these regions are observed for the respective spectra in this study.

Absorbance bands at 1500-750 cm^{-1} region are associated with vibrations of proteins, lipids and carbohydrates (Zhao et al., 2006a, 2006b; Mohacek-Grosev et al., 2001). Several main absorption peaks are observed in this stretching vibration in our spectra, not forgetting that sugars of natural products are also absorbed in this region (Zhao et al., 2006a, 2006b). Similarly, protein patterns have been associated with characteristic absorption at 1654, 1544 and 1409 cm^{-1}; a similar pattern was found in the 6 cultivated fungi species studied which according to Gonzaga et al. (2005) may be evidence of the presence of a glucan-protein complex.

Absorption bands between 1410 and 1310 cm^{-1} may indicate OH groups of phenolic compounds, and a specific band around 1405 cm^{-1} could be related to the aliphatic groups belonging to the phenolic pigments (Klaus et al., 2015), which is notable in our fungi extracts spectra. Other authors, in turn, indicate that absorption bands around 1400 cm^{-1} could also suggest the occurrence of uronic acids (O-C=O bending) (Zheng et al., 2014); and peaks at 1242 cm^{-1} could be associated to protein structures, which may be due to the presence of glucan-protein complex (Gonzaga et al., 2005).

The bands between 1165 cm^{-1} and 1150 cm^{-1} have been assigned to C-O-C stretching of glycosidic structures, 1074 cm^{-1} to anomeric C_1H group vibration and, 1042 cm^{-1} and 1020 cm^{-1} has been assigned to C-O stretching (Gonzaga et al., 2005; Klaus et al., 2015; Sarangi et al., 2006; Zhao et al., 2006a). The vibrational spectra indicate that the main compositions of these fungi are polysaccharide and protein (Klaus et al., 2015). According to Zheng et al. (2014) and reinforced by Seedevi et al. (2019), the strong characteristic absorption at 1200–1000 cm^{-1} is ascribed to sugar ring vibrations overlapping with stretching vibrations of C-OH side groups and the C-O-C glycosidic bonds vibration. Wang et al. (2014) reinforced the idea that each polysaccharide fraction from *Phellinus nigricans* mycelia had a specific band from 1200 to 1000 cm^{-1}. Moreover, according to Liu et al. (2006) the region between 1200-750 cm^{-1} could serve as fingerprints to discriminate mushrooms, whereas according to Mohacek-Grosev et al. (2001) the spectral region between 1200-1000 cm^{-1} could serve as an indicator of mushroom genus.

In the 950-750 cm^{-1} region that has been associated with identification of anomeric configuration of polysaccharides, can be deduced to contain pyran glycosidic bonds (Barbosa et al., 2003; Mohacek-Grosev et al., 2001; Zhao et al., 2006a, 2006b; Zhu et al., 2019). In particular, the 890 cm^{-1} band has been assigned to *β*-glycosides and 860-810 cm-1 region for *α*-glycosides (Klaus et al., 2015; Rodrigues et al., 2017; Zheng et al., 2014; Zhao et al., 2006a, 2006b; Zhu et al., 2019). This information is equally supported by Mohacek-Grosev et al. (2001) studies who used vibrational spectroscopy to characterize several wild growing mushroom species. According to Zhao et al. (2006a, 2006b) chitosan and b-glucan standards with b-glycosidic linkage presented bands at 897 and 889 cm^{-1}, respectively, whereas a-glycosidic linkage typical in a standard starch presented a characteristic band at 858 cm^{-1}. On the other hand, Barbosa et al. (2003) reported bands at 1370 and 890 cm^{-1} that are typical of b-glucans and of *(1 3)-β*-glucans, respectively. Sarangi et al. (2006) also reported bands at 1024 cm^{-1} and 867 cm^{-1} suggesting beta-glycosidic bonds. Small peaks at 860-750 cm^{-1} are observable in all our spectra (Figure 4.8), such could indicate that a-glycosidic

linkages could be present in these fungi species studied, as well as bands at 687-651 cm^{-1} corresponds to C-Cl stretch (Parmar & Kumar, 2015).

Overall, as observed by previous researchers mentioned through the chapter, it can be concluded that there is strong evidence of presence of polysaccharides and glycoproteins that are covalently linked together in a dynamic process that could occur extracellularly or intracellularly. The probing depth of FTIR spectroscope was enough to sense bond-vibration signal of these active groups.

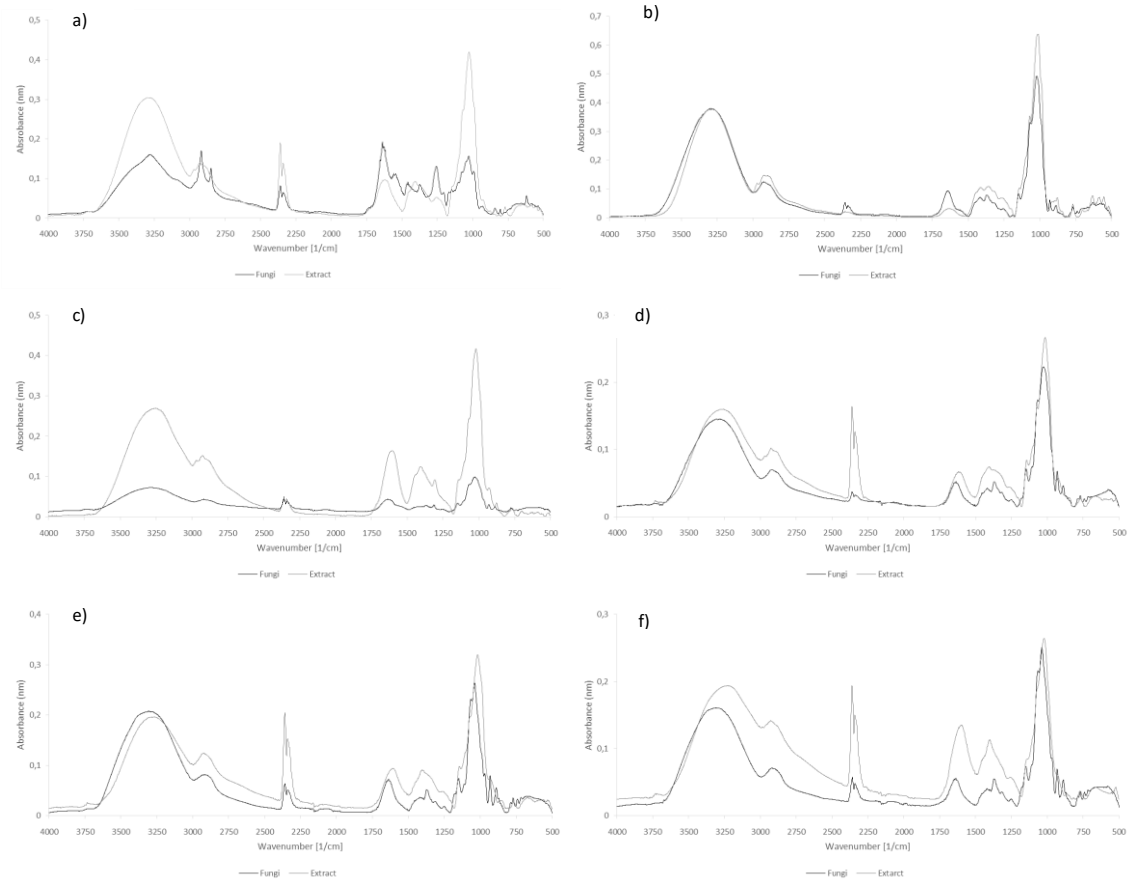

Figure 4.8. FTIR spectra of extracts and fungi of a) *Phanerochaete chrysosporium*, b) *Lentinula edodes*, c) *Pholiota nameko*, d) *Pleurotus ostreatus*, e) *Lentinus sajor-caju* and f) *Trametes versicolor*.

4.4 Conclusions

As our standard of living is improving, the requirement for higher quality foods is becoming more significant. In addition, the amount of health-related and anti-aging products is fast increasing, demonstrating the increased attention attributed to health nowadays. Therefore, to make use of the wide variety of compounds present within natural resources that have potential health and nutritional benefits, it is essential that currently used extraction technology be improved. The development of a highly efficient methodology that is rapid, convenient, safe, and sustainable would facilitate the separation of such compounds, increasing the purity of health care components whilst meeting the requirements for saving energy and protecting the environment, in addition to allowing scale up of the technology to an industrial level of production. HPE may provide such a platform, not only for use in the traditional natural product industries, but also for scientific research, product development, and technical changes in biology, medicine. This technology allows for the extraction efficiency to be greatly increased, with high yields obtained in only a few minutes. As it can be operated at room temperature, the bioactivity of compounds with low thermal stability can be protected, and as it is carried out in a closed environment, there is no volatilization of solvent, preventing environmental pollution. Although, the antioxidant activity results clearly demonstrate the potential of HPE in this context, the HPE can, therefore, be defined as an environmentally friendly, rapid, and highly efficient extraction method that is easily operated and highly mechanized. It is suitable for extension from a laboratory scale to an industrial production level for producing natural, safe, and healthy foods as required by consumers. Therefore, it becomes clear that this emerging HPE methodology will provide extensive benefits for the food and pharmaceutical industries.

The selected fungi could be a rich source of antioxidants and free radical scavenging compounds. The levels of phenolic and flavonoid compounds were positively correlated with the antioxidant activity of the fungi extracts. On the other hand, the results of antibacterial activity of this study were not affected by either total phenolic content or flavonoids content.

Regarding to compounds characterization, the NMR and FTIR spectra give information about chemical constituents of the crude fungi samples, which indicated that the main compositions of the fungi are polysaccharide and protein. This study illustrated that the FTIR spectroscopic method is a valuable tool for rapid and nondestructive analysis of fungi. Thus, in conclusion, this study showed that fungi could be promising sources for antioxidant compounds.

4.5 References

Acquah, G. E., Via, B. K., Fasina, O. O., & Eckhardt, L. G. (2016). Rapid Quantitative Analysis of Forest Biomass Using Fourier Transform Infrared Spectroscopy and Partial Least Squares Regression, *Journal of Analytical Methods in Chemistry*, 2016, 1-10.

Adil, İ., Yener, M., Bayındırlı, A. (2008) - Extraction of total phenolics of sour cherry pomace by high pressure solvent and subcritical fluid and determination of the antioxidant activities of the extracts. *Separation Science and Technology*, 43, 1091-1110.

Aktas, E. T., & Yildiz, H. (2011). Effects of electroplasmolysis treatment on chlorophyll and carotenoid extraction yield from spinach and tomato. *Journal of Food Engineering*, 106, 339–346.

Alam, N., Amin, R., Khan, A., Ara, I., Shim, M. J., Lee, M. W., & Lee, T. S. (2008). Nutritional analysis of cultivated mushrooms in Bangladesh - *Pleurotus ostreatus*, *Pleurotus sajor-caju*, *Pleurotus florida* and *Calocybe indica*. *Mycobiology*, 36, 228–232.

Alexandre, E., Araújo, P., Duarte, M., de Freitas, V., Pintado, M., & Saraiva, J. (2017a). High-pressure assisted extraction of bioactive compounds from industrial fermented fig by-product. *Journal of Food Science and Technology*, 54, 2519-2531.

Alexandre, E., Araújo, P., Duarte, M., de Freitas, V., Pintado, M., & Saraiva, J. (2017b). Experimental design, modeling, and optimization of high-pressure-assisted extraction of bioactive compounds from pomegranate peel. *Food and Bioprocess Technology*, 10, 886-900.

Alexandre, E., Silva, S., Santos, S., Silvestre, A., Duarte, M., Saraiva, J., & Pintado, M. (2019). Antimicrobial activity of pomegranate peel extracts performed by high pressure and enzymatic assisted extraction. *Food Research International*, 115, 167–176.

Altuner, E., Çeter, T., & Alpas, H. (2012). High hydrostatic pressure processing: a method having high success potential in pollen protein extraction. *High Pressure Research*, 32, 291-298.

Andreu L., Nuncio-Jáuregui, N., Carbonell-Barrachina, Á. A., Legua, P., & Hernández, F. (2018). Antioxidant properties and chemical characterization of Spanish Opuntia ficus-indica Mill. cladodes and fruits. *Journal of the Science of Food and Agriculture*, 98: 1566-1573.

Avello, M. A., Pastene, E. R., Bustos, E. D., Bittner, M. L., Becerra, J. A. (2013). Variation in phenolic compounds of *Ugni molinae* populations and their potential use as antioxidant supplement. Revista Brasileira de Farmacognosia, 23: 44-50.

Baiano, A. (2014). Recovery of biomolecules from food wastes - A review. *Molecules*, 19, 14821–14842.

Balci, M. (2005). Basic 1H- and 13C-NMR Spectroscopy, 1st Edition, Elsevier Science.

Barbosa, A. M., Steluti, R. M., Dekker, R. F., Cardoso, M. S., & da Silva M. C. (2003). Structural characterization of Botryosphaeran: a (1→3; 1→6)-beta-D-glucan produced by the ascomyceteous fungus, Botryosphaeria sp. *Carbohydrate Research*, 338, 1691-1698.

Brown, K.H., Rivera, J.A., Bhutta, Z., Gibson, R.S., King, J.C., Lönnerdal, B., Ruel, M.T., Sa

Behie, S. W., & Bidochka, M. J. (2014). Nutrient transfer in plant–fungal symbioses. Trends in Plant Science, 19, 734–740.

Berthomieu, C., & Hienerwade, R. (2009). Fourier transform infrared (FTIR) spectroscopy. *Photosynthesis Research*, 101, 157–170

Bi, H. M., Zhang, S. Q., Liu, C. J., & Wang, C. Z. (2009). High hydrostatic pressure extraction of salidroside from Rhodiola Sachalinensis. *Journal of Food Process Engineering*, 32, 53–63.

Branen, A., & Am, J. (1975). Toxicology and biochemistry of butylated hydroxyanisole and butylated hydroxytoluene. *Journal of the American Oil Chemist's Society,* 52, 59.

Briones-Labarca, V., Giovagnoli-Vicuña, C., & Cañas-Sarazúa, R., (2019). Optimization of extraction yield, flavonoids and lycopene from tomato pulp by high hydrostatic pressure-assisted extraction. *Food Chemistry*, 278, 751-759.

Butz, P., Koller, W. D., Tauscher, B., & Wolf, S. (1994). Ultra-high pressure processing of onions: Chemical and sensory changes. LWT-*Food Science and Technology*, 27(5), 463–467.

Casquete, R., Castro, S., Martin, A., Ruiz-Moyano, S., Saraiva, J., Córdoba, M., & Teixeira, P. (2015). Evaluation of the effect of high pressure on total phenolic content, antioxidant and antimicrobial activity of citrus peels. *Innovative Food Science and Emerging Technologies*, 31, 37-44.

Casquete, R., Castro, S., Villalobos, M., Serradilla, M., Queirós, R., Saraiva, J., Córdoba, M., & Teixeira, P. (2014). High pressure extraction of phenolic compounds from citrus peels. *High Pressure Research*, 34, 447-451.

Castro, L., Alexandre, E., Pintado, M., & Saraiva, J. (2019). Bioactive compounds, pigments, antioxidant activity and antimicrobial activity of yellow prickly pear peels. *International Journal of Food Science and Technology*, 54, 1225-1231.

Caytan, E., Remaud, G. S., Tenailleau, E., Akoka, S. (2007). Precise and accurate quantitative 13C NMR with reduced experimental time. *Talanta*, 71, 1016–1021.

Ćetković, S. G., Čanadanović-Brunet, M. J., Djilas, M. S., Tumbas, V., Markov, S., & Cvetkovic, D. (2007). Antioxidant potential, lipid peroxidation inhibition and antimicrobial activities of Satureja montana L. subsp. kitaibelii extracts. *International Journal of Molecular Sciences,* 8, 1013-1027.

Chaudhary, R., & Tripathy, A. (2015). Isolation and Identification of Bioactive Compounds from *Irpex Lacteus* Wild Fleshy Fungi. *Journal of Pharmaceutical Sciences and Research*, 7, 424-434.

Chen, R., Shouqin, Z., & Wang, C. (2005). High pressure extraction of total ginsenoside at room temperature. *Journal of Chemical Industry and Engineering (China)*, 56, 911.

Chen, R.., Meng, F., Zhang, S., & Liu, Z. (2009). Effects of ultrahigh pressure extraction conditions on yields and antioxidant activity of ginsenoside from ginseng. *Separation and Purification Technology*, 66, 340–346.

Chen, Y., Chefetz, B., Rosario, R., van Heemst, J. D H., Romaine, C. P., & Hatcher, P. G. (2000). Chemical nature and composition of compost during mushroom growth, *Compost Science & Utilization*, 8, 347-359.

Corrales, M., García, A. F., Butz, P., & Tauscher, B. (2009). Extraction of anthocyanins from grape skins assisted by high hydrostatic pressure. *Journal of Food Engineering,* 90, 415–421.

Corrales, M., Toepfl, S., Butz, P., Knorr, D., & Tauscher, B. (2008). Extraction of anthocyanins from grape by-products assisted by ultrasonics, high hydrostatic pressure or pulsed electric fields: A comparison. *Innovative Food Science & Emerging Technologies*, 9, 85–91.

Dai, J., & Mumper, R. J. (2010). Plant phenolics: extraction, analysis and their antioxidant and anticancer properties. *Molecules*, 15, 7313-7352.

Demain, A. (2014). Importance of microbial natural products and the need to revitalize their discovery. *Journal of Industrial Microbiology Biotechnology*, 41, 185-201.

Dewanto, V., Wu, X., & Liu, R. H. (2002). Processed sweet corn has higher antioxidant activity. *Journal of Agricultural and Food Chemistry*, 50: 4959–4964.

Diaz, P., Jeong, S. C., Lee, S., Khoo, C., & Koyyalamudi, S. R. (2012). Antioxidant and anti-inflammatory activities of selected medicinal plants and fungi containing phenolic and flavonoid compounds. *Chinese Medicine*, 7, 1-9.

Dulger B (2004). Antimicrobial activity of the macrofungus Pholiota adiposa. *Fitoterapia,* 75, 395–397.

Eberhardt, T. L., Li, X., Shupe, T. F., & Hse, C.Y. (2007). Chinese tallow tree (*Sapium Sebiferum*) utilization: Characterization of extractives and cell-wall chemistry. *Wood and Fiber Science: Journal of the Society of Wood Science and Technology*, 39, 319-324.

Elgndi, M., Filip, S., Pavlić, B., Vladic, J., Stanojkovic, T., Zizak, Z., & Zekovic, Z. (2017). Antioxidative and cytotoxic activity of essential oils and extracts of *Satureja montana L.*, *Coriandrum sativum L.* and *Ocimum basilicum L*. obtained by supercritical fluid extraction. *Journal of Supercritical Fluids*, 128, 128–137.

Fakayode, O., & Ajav, E. (2016). Process optimization of mechanical oil expression from Moringa (Moringa oleifera) seeds. *Industrial Crops and Products*, 90, 142-151.

Fan, Y., Chun, Z., Wang, G., Pu, S., Pan, Y., Ma, J., Miao, R., Luo, A. (2020). Isolation, structural characteristics, and *in vitro* and *in vivo* antioxidant activity of the acid polysaccharide isolated from *Pholiota nameko*. *Pharmacognosy Magazine*, 16, 738-44.

Fernandes, L., Casal, S., Pereira, J., Ramalhosa, E., & Saraiva, J. (2017). Optimization of high-pressure bioactive compounds extraction from pansies (Viola×wittrockiana) by response surface methodology, *High Pressure Research*, 37, 415-429.

Ferreira, I. C. F. R., Barros, L., & Abreu, R. M. V. (2009). Antioxidants in wild mushrooms. *Current Medicinal Chemistry*, 16: 1543–1560.

French, G. L. (2005). Clinical impact and relevance of antibiotic resistance. *Advanced Drug Delivery Reviews,* 57, 1514– 1527.

Gadd, G. M. (2004). Mycotransformation of organic and inorganic substrates. *Mycologist*, 18, 60–70.

Gamsjaeger, S., Baranska, M., Schulz, H., Heiselmayer, P., & Musso, M. (2011). Discrimination of carotenoid and flavonoid content in petals of pansy cultivars (Viola×wittrockiana) by FT-Raman spectroscopy. *Journal of Raman Spectroscopy*, 42, 1240–1247.

Gonçalves, M. F., Santos, L., Silva, B. M., Abreu, A. C., Vicente, T. F., Esteves, A. C., & Alves, A. (2019). Biodiversity of *Penicillium* species from marine environments in Portugal and description of *Penicillium lusitanum* sp. nov., a novel species isolated from sea water. *International Journal of Systematic and Evolutionary Microbiology*, 69, 3014–3021.

Gonzaga, M., Ricardo, N., Heatley, F., & Soares, S., (2005). Isolation and characterization of polysaccharides from *Agaricus blazei* Murill. *Carbohydrate Polymers*, 60, 43–9.

Grunovaitė, L., Pukalskienė, M., Pukalskas, A., & Venskutonis, P. R. (2016). Fractionation of black chokeberry pomace into functional ingredients using high pressure extraction methods and evaluation of their antioxidant capacity and chemical composition. *Journal of Functional Foods*, 24, 85–96.

Gu, K., Zhou, C.-Y., & Shao, Y. (2016). Advances research and utilization on active constituents of edible fungus. *Edible fungus of china, 35*, 1–9.

Guo, X., Han, D., Xi, H., Rao, L., Liao, X., Hu, X., & Wu, J. (2012). Extraction of pectin from navel orange peel assisted by ultra-high pressure, microwave or traditional heating: A comparison. *Carbohydrate Polymers*, 88, 441–448.

Guo, X., Zhao, W., Pang, X., Liao, X., Hu, X., & Wu, J. (2014). Emulsion stabilizing properties of pectins extracted by high hydrostatic pressure, high-speed shearing homogenization and traditional thermal methods: A comparative study. *Food Hydrocolloids*, 35, 217-225.

Gupta, B. S., Jelle, B. P., & Gao, T. (2015). Application of ATR-FTIR spectroscopy to compare the cell materials of wood decay fungi with wood mould fungi. International Journal of Spectroscopy, 1-7.

Ha, T. B., Gerhauser, C., Zhang, W. D., Ho-Chongline, N., & Fouraste, I. (2000). New lanostanoids from Ganoderma lucidium that include NADPH: Quinone oxidoreductase in cultured hepalcic 7 mureine hepatoma cell. *Planta Medica*, 66, 681-684.

Hall, J. E. (2015). Guyton and hall textbook of medical physiology e-book. Philadelphia, PA, Elsevier Health Sciences.

Hatvani, N. (2001). Antibacterial effect of the culture fluid of *Lentinus edodes* mycelium grown in submerged liquid culture. *International Journal of Antimicrobial Agents,* 17, 71–74.

Havsteen, B. (2002). The biochemistry and medical significance of the flavonoids. *Pharmacology & Therapeutics*, 96: 67-202.

Hazra, K. M., Roy, R. N., Sen, S. K., & Laska, S. (2007). Isolation of antibacterial pentahydroxy flavones from the seeds of Mimusops elengi Linn. *African Journal of Biotechnology*, 6, 1446-1449.

Helba, L. (2014). Antimicrobial activity of crude methanolic extracts from *Ganoderma lucidum* and *Trametes versicolor*. *Animal Science and Biotechnologies*, 47, 89–93.

ou, L., Zhang, S., Dou, J., Zhu, J., & Liang, Q. (2011). Ultrahigh hydrostatic pressure extraction of flavonoids from *Epimedium koreanum* Nakai. *Proceedings of SPIE*, 7752, 77520Z-1.

Holst, B., & Williamson, G. (2008). Nutrients and phytochemicals: from bioavailability to bioefficacy beyond antioxidants. *Current Opinion in Biotechnology*, 19: 73-82.

Hua, R., & Zhang, W. (2018). Main species and medicinal value of wild edible (medicinal) fungi in Vunnan Province. *Medicinal Plants*, 9, 1–4.

Huang, H.-W., Hsu, C.-P., Yang, B. B., & Wang, C.-Y. (2013a). Advances in the extraction of natural ingredients by high pressure extraction technology. *Trends in Food Science & Technology*, 33, 54–62.

Huang, H.-W., Hsu, C-P; Yang, B & Wang, C-Y (2013b). Advances in the extraction of natural ingredients by high pressure extraction technology. *Trends in Food Science and Technology*, 33, 54-62.

Hur, J.-M., Yang, C.-H., Han, S.-H., Lee, S.-H., You, Y.-O., Park, J.-C., & Kim, K.-J. (2004). Antibacterial effect of *Phellinus linteus* against methicillin-resistant *Staphylococcus aureus*. *Fitoterapia*, 75, 603–605.

Ishara, J., Sila, D., Kenji, G., Buzera, A., Mushagalusa, G. (2018). Nutritional and Physical Attributes of Maize-mushroom Complementary Porridges as Influenced by Mushroom Species and Ratio. *American Journal of Food and Nutrition*, 6, 17-27.

Janeš, D., Kreft, S., Jurc, M., Seme, K., & Štrukelj, B. (2007). Antibacterial Activity in higher fungi (mushrooms) and endophytic fungi from Slovenia, *Pharmaceutical Biology*, 45, 700-706.

Janjušević, L., Karaman, M., Šibul, F., Tommonaro, G., Iodice, C., Jakovljević, D., Pejin, B. (2017). The lignicolous fungusTrametes versicolor (L.) Lloyd (1920): a promising natural source of antiradical and AChE inhibitory agents. *Journal of Enzyme Inhibition and Medicinal Chemistry*, 32, 355-362.

Jeong, S.-M., Kim, S.-Y., Kim, D.-R., Jo, S.-C., Nam, K. C., Ahn, D. U., & Lee, S.-C. (2004). Effect of heat treatment on the antioxidant activity of extracts from citrus peels. Journal of Agricultural and Food Chemistry, 52, 3389–3393.

Jeong, J. B., Park, J. H., Lee, H. K., Ju, S. Y., Hong, S. C., Lee, J. R., Chung, G. Y., Lim, J. H., Jeong, H. J. (2009). Protective effect of the extracts from Cnidium officinale against oxidative damage induced by hydrogen peroxide via antioxidant effect. Food and Chemical Toxicology, 47: 525–529.

Ji, H., Zhang, L., Li, J., Yang, M., & Liu, X. (2010). Optimization of ultrahigh pressure extraction of momordicosides from bitter melon. *International Journal of Food Engineering*, 6.

Jin-Zhe, H., Qiao-Mei, R., Dan-Dan, D. & Pei-Long, S. (2013). Chemical characteristics and antioxidant properties of crude water soluble polysaccharides from four edible mushrooms. *Molecules*, 17, 4373-4387.

Ju, Z. Y., & Howard, L. R. (2003). Effects of solvent and temperature on pressurized liquid extraction of anthocyanins and total phenolics from dried red grape skin. *Journal of Agricultural and Food Chemistry,* 51, 5207–5213.

Jun X. (2013). High-pressure processing as emergent technology for the extraction of bioactive ingredients from plant materials. *Critical Reviews in Food Science and Nutrition*, 53, 837–852.

Jun, X. (2009). Caffeine extraction from green tea leaves assisted by high pressure processing. *Journal of Food Engineering*, 94, 105–109.

Jun, X., Deji, S., Shou, Z., Bingbing, L., Li, Y., & Zhang, R. (2009). Characterization of polyphenols from green tea leaves using a high hydrostatic pressure extraction. *International Journal of Pharmaceutics*, 382, 139–143.

Kamiyama, M., Horiuchi, M., Umano, K., Shibamoto, T., Kondo, K., & Otsuka, Y. (2013). Antioxidant/antiinflammatory activities and chemical composition of extracts from the mushroom *Trametes versicolor*. *International Journal of Nutritional and Food Science*, 2, 85–91.

Kapravelou, G., Martínez, R., Andrade, A. M., Chaves, C. L., López-Jurado, M., Aranda, P., Arrebola, F., Cañizares, F. J., Galisteo, M., & Porres J. M. (2015). Improvement of the antioxidant and hypolipidaemic effects of cowpea flours (*Vigna unguiculata*) by fermentation: results of *in vitro* and *in vivo* experiments. *Journal of the Science of Food and Agriculture*, 95: 1207-1216.

Karaman, M., Jovin, E., Malbasa, R., Matavuly, M., & Popovic, M. (2010). Medicinal and edible lignicolous fungi as natural sources of antioxidative and antibacterial agents. *Phytotherapy Research*, 24, 1473–81.

Khan, S., Aslam, R., & Makroo, H. (2018). High pressure extraction and its application in the extraction of bio-active compounds: A review. *Journal Food Process Engineering*, e12896.

Khatua, S., Paul, S., & Acharya, K. (2013). Mushroom as the potential source of new generation of antioxidant: a review. *Research Journal Pharmacy and Technology*, 6, 496–505.

Kim, M. Y., Seguin, P., Ahn, J. K., Kim, J.-J., Chun, S.-C., Kim, E.-H., Seo, S.-H., Kang, E.-Y., Kim, S.-L., Park, Y.-J., Ro‖, H.-M., & Chung, I.-M. (2008). Phenolic compound concentration and antioxidant activities of edible and medicinal mushrooms from Korea. Journal Agricultural and Food Chemistry, 56: 7265–7270.

Kitzberger, C., Smânia, A., Pedrosa, R., & Ferreira, S. (2007). Antioxidant and antimicrobial activities of *shiitake* (*Lentinula edodes*) extracts obtained by organic solvents and supercritical fluids. *Journal of Food Engineering,* 80, 631-638.

Klaus, A., Kozarski, M., Vunduk, J., Todorovic, N., Jakovljevic, D., Zizak, Z., Pavlovic, V., Levic, S., Niksic, M. & Van Griensven, L.J.L.D. (2015). Biological potential of extracts of the wild edible Basidiomycete mushroom *Grifola frondosa*. *Food Research International*, 67, 272-283.

Knorr, D., Froehling, A., Jaeger, H., Reineke, K., Schlueter, O., & Schoessler, K. (2011). Emerging technologies in food processing. *Annual Review* of *Food Science and Technology,* 2, 203-235.

Kozarski, M., Klaus, A., Niksic, M., Kozarki, M., Klaus, A., Niksic, M., Vrvic, M., Todorovic, N., Jakovljevic, D., & Van Griensen, L. (2012). Antioxidative activities and chemical characterization of polysaccharide extracts from the widely used mushrooms *Ganoderma*

applanatum, *Ganoderma lucidum*, *Lentinus edodes* and *Trametes versicolor*. *Journal of Food Composition and Analysis*, 26, 144–153.

Kraujalis, P., Venskutonis, P. R., Ibanez, E., & Herrero, M. (2015). Optimization of rutin isolation from *Amaranthus paniculatus* leaves by high pressure extraction and fractionation techniques. *The Journal of Supercritical Fluids*, 104, 234–242.

Kumar, A., & Schweizer, H. P. (2005). Bacterial resistance to antibiotics: Active efflux and reduced uptake*. Advanced Drug Delivery Reviews,* 57, 1486–1513.

Kumar, C., Mongolla, P., Joseph, J., Nageswar, Y., & Kamal, A. (2010). Antimicrobial activity from the extracts of fungal isolates of soil and dung samples from Kaziranga National Park, Assam, India. *Journal de Mycologie Médicale*, 20, 283-289.

Larrosa, M., García-Conesa, M. T., Espín, J. C., & Tomás-Barberán, F. A. (2010). Ellagitannins, ellagic acid and vascular health. *Molecular Aspects of Medicine*, 31: 513-539.

Lebovka, N., Vorobiev, E., & Chemat, F. (2011). High pressure-assisted extraction: method, technique, and application. In *Enhacing extraction processes in the food industry,* CRC Press., pp.303-322.

Lee, H. (2017). Alcohol hangover relieving effect of Hovenia dulcis thunb associated with antioxidant activities through ultra-high-pressure extraction process. *Research Journal of Biotechnology*, 12, 72-77.

Lee, H. Y., He, X., & Ahn, J. (2010). Enhancement of antimicrobial and antimutagenic activities of Korean barberry (Berberis koreana Palib.) by the combined process of high-pressure extraction with probiotic fermentation. *Journal of the Science of Food and Agriculture*, 90, 2399-404.

Lee, H.-S., Lee, H., Yu, H., Ju, D., Kim, Y., Kim, C.-T., Kim, C.-J., Cho, Y.-J., Kim N., Choi, S.-Y., & Suh, H. (2011). A comparison between high hydrostatic pressure extraction and heat extraction of ginsenosides from ginseng *(Panax ginseng CA* Meyer). *Journal of the Science of Food and Agriculture*, 91, 1466-1473.

Levin, I. W., & Bhargava, R. (2005). Fourier transform infrared vibrational spectroscopic imaging: integrating microscopy and molecular recognition. *Annual Review Physical Chemistry,* 56, 29–474.

Levy, S., B. (2005). Antibiotic resistance - the problem intensifies. *Advanced Drug Delivery Review,* 57, 1446–1450.

Li, X., Meng, X., Tan, H., Ren, Q., Li, D., & Li, B. (2018). Ultra-high pressure extraction of anthocyanins from *Lonicera caerulea* and its antioxidant activity compared with ultrasound-assisted extraction. *International Journal of Agriculture and Biology*, 20, 2257-2264.

Li, Z., Smith, K. H., & Stevens, G. W. (2016). The use of environmentally sustainable bio-derived solvents in solvent extraction applications—A review. *Chinese Journal of Chemical Engineering,* 24, 215–220.

Liao, G., Liu, J., He, S., Chen, J., & Zhang, Z. (2012). Optimization for ultrahigh pressure extraction of berberine from cortex Phellodendri by central composite design-response surface methodology. *Journal of medicinal plants research*, 6, 3963–3970.

Linton, M., & Patterson, M. F. (2000). High pressure processing of foods for microbiological safety and quality (a short review). *Acta Microbiologica et Immunologica Hungarica*, 47, 175-182.

Liu, C., Zhang, S., & Wu, H. (2009). Non-thermal extraction of effective ingredients from *Schisandra chinensis* Baill and the antioxidant activity of its extract. *Natural Product Research*, 23, 1390-13401.

Liu, G., Song, D., Zhao, D., Liu, J., Zhou, Y., Ou, J., & Sun, S. (2006). A study of the mushrooms of boletes by Fourier transform infrared spectroscopy. In: von Bally, G., Luo, Q. (Eds.), *SPIE Proceedings*., (pp. 124-129). ICO20: Biomedical Optics.

Liu, J. L., Zheng, S. L., Fan, Q. J., Yuan, J. C., Yang, S. M., & Kong, F. L. (2015). Optimization of high-pressure ultrasonic-assisted extraction and antioxidant capacity of polysaccharides from the rhizome of *Ligusticum chuanxiong*. *International Journal of Biological Macromolecules*, 76, 80–85.

Liu, V., Xu, P., Zeng, G., Huang, D., Zhao, M., Lai, C., Chen, M., Li, N., Huang, C., Wang, C., Cheng, M., He, X., Lai, M., & He, Y. (2014). Inherent antioxidant activity and high yield production of antioxidants in *Phanerochaete chrysosporium*. *Biochemical Engineering Journal*, 90, 245–254.

Liu, Y., Sun, J., Luo, Z., Rao, S.-O., Su, Y.-J., Xu, R.-R., & Yang, Y.-J. (2012). Chemical composition of five wild edible mushrooms collected from Southwest China and their antihyperglycemic and antioxidant activity. *Food and Chemical Toxicology*, 50, 1238–1244.

Lourenço, S. C., Moldão-Martins, M., & Alves, V. D. (2019). Antioxidants of natural plant origins: from sources to food industry applications. *Molecules*, 24, 4132.

Luna-Guevara, L., Luna-Guevara, J., J., Hernández-Carranza, P., Ruíz-Espinosa, H., Ochoa-Velasco, C. E. (2018). Phenolic compounds: A good choice against chronic degenerative diseases. *Studies in Natural Products Chemistry*, 59, 79-108.

Luo, K., Yue, G., Ko, C., Lee, J., Gao, S., Li, L.-F., Li, G., Fung, K.-P., Leung, P.-C., & Lau, C. (2014). *In vivo* and *in vitro* antitumor and anti-metastasis effects of *Coriolus versicolor* aqueous extract on mouse mammary 4T1 carcinoma. *Phytomedicine*, 21, 1078–1087.

M'hiri, N., Ioannou, I., Boudhrioua, N., & Ghoul, M. (2015). Effect of different operating conditions on the extraction of phenolic compounds in orange peel. *Food and Bioproducts Processing*, 96, 161-170.

Manganyi, M., Regnier, T., Kumar, A., Bezuidenhout, C., Ateba, C. (2018). Biodiversity and antibacterial screening of endophytic fungi isolated from *Pelargonium sidoides*. *South African Journal of Botany,* 116, 192–199.

Manganyia, M. C., Tchatchouanga, C.-D. K., Regnierb, T., Bezuidenhoutc, C. C., & Ateba, C. N. (2019). Bioactive compound produced by Endophytic Fungi isolated from *Pelargonium sidoides* against selected bacteria of clinical importance. *Microbiology*, 47, 335-339.

Mathew, S., Abraham, T. E., & Zakaira, Z. A. (2015). Reactivity of phenolic compounds towards free radicals under in vitro conditions. *Journal of Food Science and Technology*, 52: 5790–5798.

Mau, J. L., Lin, H. C., & Chen, C. (2002). Antioxidant Properties of Several Medicinal Mushrooms, J*ournal of Agricultural and Food Chemistry*, 50, 6072–6077.

Mau, J. L., Lin, H. C., Song, S. F. (2002). Antioxidant properties of several specialty mushrooms. *Food Research International*, 35, 519–526.

Miller, H. E., Rigelhof, F., Marquart, L., Prakash, A., & Kanter, M. (2000). Antioxidant Content of Whole Grain

Breakfast Cereals, Fruits and Vegetables. *Journal of the American College of Nutrition*, 19, 312–319.

Minatel, I. O., Borges, C. V., Gomez, H. A. G., Ferreira, M. I., Chen, C., & Lima, G. P. P. (2017). Phenolic Compounds: Functional Properties, Impact of Processing and Bioavailability. In M. Soto-Hernandez, M. Palma-Tenango & M. R. Garcia-Mateos (Eds.), *Phenolic Compounds - Biological Activity* (pp. 1-24). IntechOpen.

Mira, L., Fernandez, M. T., Santos, M., Rocha, R., Florencio, M. H., & Jennings, K. R. (2002). Interactions of flavonoids with iron and copper ions: a mechanism for their antioxidant activity. *Free Radical Research*, 36: 1199–1208.

Mohacek-Grosev, V., Bozac, R., & Puppels, G. J. (2001). Vibrational spectroscopic characterization of wild growing mushrooms and toadstools. *Spectrochimica Acta. Part A, Molecular and Biomolecular Spectroscopy*, 57, 2815-2829.

Morales, D., Piris, A. J., Ruiz-Rodriguez, A., Prodanov, M., & Soler-Rivas, C. (2018). Extraction of bioactive compounds against cardiovascular diseases from *Lentinula edodes* using a sequential extraction method. *Biotechnology Progress*, 34, 746–755.

Morales, D., Rutekeviski, R., Villalva, M., Abreu, H., Sler-Rivas, C., Santoyo, S., Iacomini, M., & Smiderle, F. (2020). Isolation and comparison of *α*- and *β-D-glucans* from *shiitake* mushrooms (*Lentinula edodes*) with different biological activities. *Carbohydrate Polymers*, 229, 115521.

Moreira, S. A., Alexandre, E. M. C., Pintado, M. E., & Saraiva, J. A. (2019). Effect of emergent non-thermal extraction technologies on bioactive individual compounds profile from different plant materials. *Food Research International*, 115, 177–190.

Moreira, S. A., Pintado, M. E., & Saraiva, J. A. (2020a). Optimization of antioxidant activity and bioactive compounds extraction of winter savory leaves by high hydrostatic pressure. *High Pressure Research*, 40, 543-560.

Moreira, S. A., Silva, S., Costa, E., Pinto, S., Sarmento, B., Saraiva, J. A., & Pintado, M. (2020b). Effect of high hydrostatic pressure extraction on biological activities and phenolics composition of Winter Savory leaf extracts, Antioxidants, 9, 841.

Mustafa, A., & Tumer, C. (2011). Pressurized liquid extraction as a green approach in food and herbal plants extraction: a review. *Analytica Chimica Acta*, 703, 8-18.

Naderi, N., Pouliot, Y., House, J., & Doyen, A. (2017). High hydrostatic pressure effect in extraction of 5-methyltetrahydrofolate (5-MTHF) from egg yolk and granule fractions. *Innovative Food Science and Emerging Technologies*, 43, 191-200.

Ngai, P. H. K., Ng, T. B. (2004): A ribonuclease with antimicrobial, antimitogenic and antiproliferative activities from the edible mushroom *Pleurotus sajor-caju*. *Peptides,* 25, 11–17.

Nieto, A., Borrull, F., Marcé, R., & Pocurull, E. (2008). Pressurized Liquid Extraction of Contamination from Environmental samples. *Current Analytical Chemistry*, 4, 157-167.

Oliveira, C., Gurak, P., Cladera-Olivera, F., Marczak, L., & Karwe, M. (2016). Combined effect of High-Pressure and conventional heating on pectin extraction from passion fruit peel. *Food and Bioprocess Technology*, 1-10.

Pais, A. C. S., Pinto, C. A., Ramos, P. A. B., Pinto, R. J. B., Rosa, D., Duarte, M. F., Abreu, M. H., Rocha, S. M., Saraiva, J. A., Silvestre, A. J. D., & Santos, S. A. O. (2019). High pressure extraction of bioactive diterpenes from the macroalgae *Bifurcaria bifurcata*: an efficient and environmentally friendly approach. *Royal Society of Chemistry Advances*, 9, 39893–39903.

Parmar, R., & Kumar, D. (2015). Study of chemical composition in wild edible *mushroom Pleurotus cornucopiae (Paulet)* from Himachal Pradesh, India by using Fourier transforms infrared spectrometry (FTIR), Gas chromatography-mass spectrometry (GCMS) and X-ray fluorescence (XRF). *Biological Forum – An International Journal,* 7, 1057-1066.

Pereira, R. N., & Vicente, A. A (2010). Environmental impact of novel thermal and non-thermal technologies in food processing. *Food Research International*, 43, 1936-1943.

Petti, S., & Scully, C. (2009). Polyphenols, oral health and disease: a review. *Journal of Dentistry*, 37: 413–423.

Petibois C., & Déléris, G. (2006). Chemical mapping of tumor progression by FT-IR imaging: towards molecular histopathology. *Trends Biotechnology,* 24, 455–462.

Pop, R., Puia, I., Puia, A., Chedea, V., Leopold, N., Bocsan, I., & Buzoianu, A. (2018). Characterization of *Trametes versicolor*: medicinal mushroom with important health benefits. *Notulae Botanicae Horti Agrobotanici Cluj-Napoca,* 46, 343-349.

Prasad, K. N., Yang, B., Zhao, M., Wei, X., Jiang, Y., & Chen, F. (2009c). High pressure extraction of corilagin from longan (*Dimocarpus longan Lour.*) fruit pericarp. *Separation and Purification Technology*, 70, 41–45.

Prasad, K., Hao, J., Shi, J., Liu, T., Li, J., Wei, X., Qiu, S., Xue, S., & Jiang, Y. (2009a). Antioxidant and anticancer activities of high pressure-assisted extract of longan (*Dimocarpus longan* Lour.) fruit pericarp. *Innovative Food Science and Emerging Technologies*, 10, 413-419.

Prasad, K., Yang, B., Shi, J., Yu, C., Zhao, M., Xue, S., & Jiang, Y. (2010a). Enhanced antioxidant and antityrosinase activities of longan fruit pericarp by ultra-high-pressure-assisted extraction. *Journal of Pharmaceutical and Biomedical Analysis*, 51, 471-747.

Prasad, K., Yang, B., Zhao, M., & Sun, J. (2010b). Effects of high pressure or ultrasonic treatment on extraction yield and antioxidant activity of pericarp tissues of longan fruit. *Journal of Food Biochemistry*, 34, 838- 855.

Prasad, K., Yang, B., Zhao, M., Ruenroengklin, N., & Jiang, Y. (2009b). Application of ultrasonication or high-pressure extraction of flavonoids from litchi fruit pericarp. *Journal of Food Process Engineering,* 32, 828–843

Prasad, K., Yang, B., Zhao, M., Wang, B; Chen, F., & Jiang, Y. (2009e). Effects of high-pressure treatment on the extraction yield, phenolic content and antioxidant activity of litchi (*Litchi chinensis Sonn.*) fruit pericarp. *International Journal of Food Science and Technology*, 44, 960-966.

Primasova, H., Bigler, P., & Furrer, J. (2017). The DEPT Experiment and Some of Its Useful Variants. *Annual Reports on NMR Spectroscopy*, 92, 1-82.

Qadir, S., Kwon, M., Han, J., Ha, J., Chung, H., Ahn, J., & Lee, H. (2009). Effect of different extraction protocols on anticancer and antioxidant activities of Berberis koreana bark extracts. *Journal of Bioscience and Bioengineering*, 107, 331-8.

Qian, L., Zhang, Y., & Liu, F. (2016). Purification and characterization of a ~43 kDa antioxidant protein with antitumor activity from *Pholiota nameko*. *Journal of the Science of Food and agriculture*, 96, 1044-1052.

Rahman, A.-U., Choudhary, M., & Wahab, A.-T. (2016). Spin-Echo and Polarization Transfer. In: *Solving Problems with NMR Spectroscopy*, Elsevier Academic Press, pp. 133-190.

Rastogi, N., Raghavarao, K., Balasubramaniam, V., Niranjan, K., & Knorr, D. (2007). Opportunities and challenges in high pressure processing of foods. *Critical Reviews in Food Science and Nutrition*, 47, 69-112.

Rodrigues, D., Freitas, A., Sousa, S., Amorim, M., Vasconcelos, M., da Costa, J., Silva, A., Rocha-Santos, T., Duarte, A. & Gomes, A. (2017). Chemical and structural characterization of *Pholiota nameko* extracts with biological properties. *Food Chem*istry, 216, 176–185.

Rop, O., Mlcek, J., Jurikova, T., Neugibauerova, J., & Vabkova, J. (2012). Edible flowers – A new promising source of mineral elements in human nutrition. *Molecules*, 17, 6672–6683.

Roupas, P., Keogh, J., Noakes, M., Margetts, C., & Taylor, P. (2012). The role of edible mushrooms in health: Evaluation of the evidence. *Journal of Functional Foods*, 4, 687-709.

Royse, D. J., Baars, J., & Tan, Q. (2017). Current overview of mushroom production in the world. In D. C. Zield, & A. Pardo-Gimenez (Eds.). *Edible and medicinal mushrooms: Technology and applications*. Hoboken: John Wiley & Sons LtD, pp. 5–13.

Rubió, L., Motilva, M.-J., & Romero, M.-P. (2013). Recent advances in biologically active compounds in herbs and spices: a review of the most effective antioxidant and anti-inflammatory active principles. *Critical Reviews in Food Science and Nutrition*, 53: 943-953.

Ruiz-Montanez, G., Ragazzo-Sánchez, J., Calderón-Santoyo, M., Velázquez-de la Cruz, G., Ramírez de León, J., & Navarro-Ocaña, A. (2014). Evaluation of extraction methods for preparative scale obtention of mangiferin and lupeol from mango peels (*Mangifera indica* L.). *Food Chemistry*, 159, 267-272.

S. M., Saraiva, J. A., Silvestre, A. J. D., & Santos, S. A. O. (2019). High pressure extraction of bioactive diterpenes from the macroalgae *Bifurcaria bifurcata*: an efficient and environmentally friendly approach, *Royal Society of Chemistry*, 9, 39893.

Sadus, R. J. (2012). *High pressure phase behaviour of multicomponent fluid mixtures*. Amsterdam, Elsevier Science.

Sabaratnam, V., Kah-Hui, W., Naidu, M., & David, P.R. (2011). Neuronal Health – Can Culinary and Medicinal Mushrooms Help?. *Journal of traditional and complementary Medicine*, 3, 62-68.

Sadhana, B., & Malini R. H. (2018). FTIR analysis for harvested Oyster mushroom. *TEJAS Thiagarajar College Journal*, 3, 49-61.

Sánchez-Moreno, C., Plaza, L., de Ancos, B., & Cano, M. P. (2004). Effect of combined treatments of high-pressure and natural additives on carotenoid extractability and antioxidant activity of tomato puree (*Lycopersicum esculentum mill.*). *European Food Research and Technology*, 219, 151–160.

Santos, C., & Silva, A. (2014). Nuclear Magnetic Resonance Spectroscopy for Structural Characterization of Bioactive Compounds. In: *Analysis of Marine Samples in Search of Bioactive Compounds*, Elsevier, B.V..

Santos, I., Silva, L., Silva, M., Araújo, J., Cavalcanti, M., & Lima, V. (2015). Antibacterial activity of endophytic fungi from leaves of *Indigofera suffruticosa Miller* (*Fabaceae*). *Frontiers in Microbiology,* 6, 350.

Santos, T. L., Tavares, O. C. H., Lopes, S. A., Elias, S. S., Berbara, R. L. L., & García, A. C. (2021). Environmental implications of the organic matter structure for white-rot fungus *Pleurotus eryngii* growth in a tropical climate. *Fungal Biology*, 125, 845-859.

Sarangi, I., Ghosh, D., Bhutia, S., Malleck, S., & Maiti, T. (2006). Anti-tumor and immunomodulating effects of *Pleurotus ostreatus* mycelia-derived proteoglycans. *International Immunopharmacolo,* 6, 1287–1297.

Sasidharan, S., Chen, Y., Saravanan, K., Sundram, L., & Latha, Y. (2011). Extraction, Isolation and Characterization of Bioactive Compounds from plants extracts. *African Journal of Traditional, Complementary and Alternative Medicines,* 8, 1-10.

Seedevi, P., Ganesan, A., Mohan, K., Raguraman, V., Sivakumar, M., Sivasankar, P., Loganathan, S., Rajamalar, P., Vaviamani, S., & Shanmugam, A. (2019). Chemical structure and biological properties of a polysaccharide isolated from *Pleurotus sajor-caju*. Royal Society of Chemistry Advances, 9, 20472.

Shin, J.-S., Ahn, S., Choi, S., Lee, D., Kim, B., & Baik, M. (2010). Ultra high pressure extraction (UHPE) of ginsenosides from Korean Panax ginseng powder. *Food Science and Biotechnology*. 19, 743-748.

Shouqin, Z., Bi, H.-M., & Liu, C.-J. (2007a). Extraction of bio-active components from *Rhodiola sachalinensis* under ultrahigh hydrostatic pressure. *Separation and Purification Technology*, 57, 277–282.

Shouqin, Z., Jun, X., & Changzheng, W. (2005). High hydrostatic pressure extraction of flavonoids from propolis. *Journal of Chemical Technology and Biotechnology*, 80, 50–54.

Shouqin, Z., Junjie, Z., & Changzhein, W. (2004). Novel high pressure extraction technology. *International Journal of Pharmaceutics*, 278, 471-474.

Shouqin, Z., Ruizhan, C., & Changzheng, W. (2006). Ginsenoside extraction from *Panax quinquefolium* L. (American ginseng) root by using ultrahigh pressure. *Journal of Pharmaceutical and Biomedical Analysis*, 41, 57-63.

Shouqin, Z., Ruizhan, C., & Changzheng, W. (2007b). Experiment study on ultrahigh pressure extraction of ginsenosides. Journal of Food Engineering. 79, 1-5.

Silva, F. V. M., Martins, A., Salta, J., Neng, N., Nogueira, J., Gaspar, N., Mira, D., Justino, J., Grosso, C., Urieta, J., Palavra, A., & Rauter, A. (2009). Phytochemical profile and anticholinesterase and antimicrobial activities of supercritical versus conventional extracts of *Satureja montana*. *Journal of agricultural and Food Chemistry*, 57, 11557–11563.

Skowyra, M., Calvo, M., Gallego, M., Azman, N., & Almajano, M. (2014). Characterization of phytochemicals in petals of different colours from Viola × wittrockiana Gams. and their correlation with antioxidant activity. *Journal of Agricultural Science,* 6, 93–105.

Soković, M., & Liaras, K. (2021). Natural products as antifungals. In M. Soković & K. Liaras (Eds.), *Antifungal Compounds Discovery* (pp. 67-165). Elsevier.

Song, K. (2017). Interphase characterization in rubber nanocomposites. In: *Progress in Rubber Nanocomposites,* Elsevier, Ltd..

Souza, A. D., Maia, A. I. V., Rodrigues, T. H. S., Canuto, K. M., Ribeiro, P. R. V., Pereira, R. D. C. A., … de Brito, E. S. (2016). Ultrasound-assisted and pressurized liquid extraction of phenolic compounds from Phyllanthus amarus and its composition evaluation by UPLC-QTOF. *Industrial Crops and Products*, 79, 91–103.

Strati, I., Gogou, E., & Oreopoulou, V. (2015). Enzyme and high pressure assisted extraction of carotenoids from tomato waste. *Food and Bioproducts Processing*, 94, 668-674.

Strobel, G. A. (2003). Endophytes as sources of bioactive products. *Microbes and Infection Journal,* 5, 535–544.

Strobel, G., Daisy, B., Castillo, U., Harper, J. (2004). Natural products from endophytic microorganisms*. Journal Natural Product,* 67, 257–268.

Sun, X., Sun, Y., Zhang, Q., Zhang, H., Yang, B., Wang, Z. & Kuang, H. (2014). Screening and comparison of antioxidant activities of polysaccharides from *Coriolus versicolor*. *International Journal of Biological Macromolecules*, 69, 12–9.

Sun, Y., Zhang, M., & Fang, Z. (2020). Efficient physical extraction of active constituents from edible fungi and their potential bioactivities: A review. *Trend in Food Science & Technology*, 105, 468-482.

Synytsya, A., Monkai, J., Bleha, R., Macurkova, A., Ruml, T., Ahn, J., Chukeatirote, E. (2017). Antimicrobial activity of crude extracts prepared from fungal. *Asian Pacific Journal of Tropical Biomedicine*, 7, 257–261.

Taârit, M. B., Msaada, K., Hosni, K., Marzouk, B. (2012). Physiological changes, phenolic content and antioxidant activity of *Salvia officinalis* L. grown under saline conditions. *Journal of the Science of Food and Agriculture*, 92: 1614-1619.

Tao, X., Zhan, Y., Zhou, Q., Feng, F., & Yu, Y. (2012). Optimization of ultra-high pressure extraction process of polysaccharides from Dendrobium candidum by response surface method. In Liu, Z., Peng, F. & Liu, X. Advances in Chemical Engineering, 550-553.

Thaipong, K., Boonprakob, U., Crosby, K., Zevallos-Cisneros, L., & Byrne, D. (2006). Comparison of ABTS, DPPH, FRAP, and ORAC assays for estimating antioxidant activity from guava fruit extracts. *Journal of Food Composition and Analysis*, 19, 669–675.

Trabelsi, N., Waffo-Téguo, P., Snoussi, M., Ksouri, R., Mérillon, J. M., Smaoui, A., Abdelly, C. (2012). Variability of phenolic composition and biological activities of two Tunisian halophyte species from contrasted regions. *Acta Physiol Plant*, 35: 749-761.

Trovato, A., Siracusa, R., Di Paola, R., Suito, M., Fronte, V., Kovirech, G., Luca, M., Serra, A., Toscano, M., Petralia, A., Cuzzocrea, S., & Calabrese, V. (2016). Redox modulation of cellular stress response and lipoxin A4 expression by *Coriolus versicolor* in rat brain: relevance to Alzheimer's disease pathogenesis. *Neurotoxicology,* 53, 1–358.

Tungmunnithum, D., Thongboonyou, A., Pholboon, A., & Yangsabai, A. (2018). Flavonoids and other phenolic compounds from medicinal plants for pharmaceutical and medical aspects: an overview. *Medicines*, 5: 93

Valko, M., Leibfritz, D., Moncol, J., Cronin, M. T.D., Mazur, M., Telser, J. (2007). Free radicals and antioxidants in normal physiological functions and human disease. *The International Journal of Biochemistry & Cell Biology*, 39: 44-84.

van der Heijden, M. G. A., Martin, F. M., Selosse, M., & Sanders, I. R. (2015). Mycorrhizal ecology and evolution: the past, the present, and the future. The New Phytologist. 205, 1406–1423.

van der Waaij, D., Nord, C. (2000). Development and persistence of multi-resistance to antibiotics in bacteria; An analysis and a new approach to this urgent problem. *International Journal of Antimicrobial Agents,* 16, 191–197.

Vaz, J.A., Heleno, S.A., Martins, A., Almeida, G.M., Vasconcelos, M.H., & Ferreira, I.C.F.R. (2010). Wild mushrooms *Clitocybe alexandri* and *Lepista inversa*: *In vitro* antioxidant activity and growth inhibition of human tumor cell lines. *Food Chemistry and Toxicology*, 48, 2881-2884.

Ventola, C. L. (2015). The antibiotic resistance crisis: part 1: causes and threats. *Pharmacy and Therapeutics,* 40, 277-283.

Vukics, V., Kery, A., & Guttman, A. (2008). Analysis of polar antioxidants in heartsease (*Viola tricolor L.*) and garden pansy (Viola × wittrockiana Gams.). *Journal of Chromatographic Science,* 46, 823–827.

Wan, J. (2013). Polysaccharide Krestin (PSK) and Polysaccharopeptide (PSP). In: Kastin AJ, ed. *Handbook of biologically active peptides: fungal peptides*. USA: Elsevier Inc., pp. 180–184.

Wang, C.-Y, Huang, H.-W, Hsu, C.-P., & Yang, B. B. (2016). Recent advances in food processing using high hydrostatic pressure technology. *Critical Reviews in Food Science and Nutrition*, 56, 527-540.

Wang, X., Liu, H., Zhang, J., Li, T., & Wang, Y.-Z. (2017). Evaluation of heavy metal concentrations of edible wild-grown mushrooms from China. *Journal of Environmental Science & Health*, 52, 178–183.

Wang, Z., Zhou, F., & Quan, Y. (2014). Antioxidant and immunological activity in vitro of Polysaccharides from *Phellinus nigricans* mycelia. *International Journal of Biological Macromolecules*, 64, 139-143.

Waring, B. G., Averill, C., & Hawkes, C. V. (2013). Differences in fungal and bacterial physiology alter soil carbon and nitrogen cycling: insights from metaanalysis and theoretical models. *Ecology Letter*. 16, 887–894.

Wasser, S. P. (2011). New dietary supplements from medicinal mushrooms. *International. Journal of Medicinal Mushrooms*, 13, 307-313.

Wolpert, M., & Hellwing, P. (2006). Infrared spectra and molar absorption coefficients of the 20 alpha amino acids in aqueous solutions in the spectral range from 1800 to 500 cm^{-1}. *Spectrochimica Acta,* 64, 987–1001.

Xi, J. (2006a). Application of high hydrostatic pressure processing of food to extracting lycopene from tomato paste waste. *High Pressure Research*, 26, 33-41.

Xi, J. (2006b). Effect of high-pressure processing on the extraction of lycopene in tomato paste waste. *Chemical Engineering and Technology*. 29, 736-739.

Xi, J. (2006c). Comparison of antioxidant activity of ethanolic extracts of propolis obtained by different extraction methods. *The Canadian Journal of Chemical Engineering*, 84, 447-451.

Xi, J. (2009a). Caffeine extraction from green tea leaves assisted by high pressure processing. *Journal of Food Engineering*, 94,105-109.

Xi, J., & Shouqin, Z. (2007). Antioxidant activity of ethanolic extracts of propolis by high hydrostatic pressure extraction. *International Journal of Food Science and Technology*. 42, 1350- 1356.

Xi, J., et al. (2013). Artificial neural network modeling and optimization of ultrahigh pressure extraction of green tea polyphenols. *Food Chemistry*, 141, 320-326.

Xi, J., Shen, D., Li, Y. & Zhang, R. (2011a). Micromechanism of ultrahigh pressure extraction of active ingredients from green tea leaves. *Food Control*, 22, 1473-1476.

Xi, J., Shen, D., Li, Y., & Zhang, R. (2011b). Ultra-high pressure extraction as a tool to improve the antioxidant activities of green tea extracts. *Food Research International*, 44, 2783–2787.

Xi, J., Shen, D., Zhao, S., Lu, B., Li, Y., & Zhang, R. (2009b). Characterization of polyphenols from green tea leaves using a high hydrostatic pressure extraction. *International Journal of Pharmaceutics*, 382, 139–143.

Xi, J., Zhao, S., Lu, B., Zhang, R., Li, Y., Shen, D., & Zhou, G. (2010). Separation of major catechins from green tea by ultrahigh pressure extraction. *International Journal of Pharmaceutics*, 386, 229-231.

Xu, X., Yan, H., Chen, J., & Zhang, X. (2011). Bioactive proteins from mushrooms. *Biotechnology Advances*, 29, 667-674.

Xu, X.-F., Yan, H.-D., Tang, J., Chen, J., & Zhang, X.-W. (2014). Polysaccharides in *Lentinus edodes*: Isolation, structure, immunomodulating activity and future prospective. *Critical Reviews in Food Science and Nutrition*, 54, 474–487.

Yadav, L. D. S. (2013). Organic Spectroscopy. Springer. pp. 197–199.

Yang, Y., Shao, B., Zhang, J., Wu, Y., & Duan, H. (2009). Determination of the residues of 50 anabolic hormones in muscle, milk and liver by very-high-pressure liquid chromatography–electrospray ionization tandem mass spectrometry. *Journal of Chromatography B*, 877, 489–496.

Yaoita, Y., Kikuchi, M., & Machida, K. (2015). *Terpenoids and Sterols from Mushrooms.* In: *Studies in Natural Products Chemistry*, Elsevier B.V.

Yeh, C.-T., & Yen, G.-C. (2006). Modulation of hepatic phase II phenol sulfotransferase and antioxidant status by phenolic acids in rats. The Journal of Nutritional Biochemistry, 17: 561-569.

Zhang, J., Wang, L.-S., Gao, J.-M., Xu, Y-J; Li, L-F & Li, C-H (2012). Rapid separation and identification of anthocyanins from flowers of *Viola yedoensis* and *V. prionantha* by high-performance liquid chromatography-photodiode array detection-electrospray ionisation mass spectrometry. *Phytochemical Analysis*, 23, 16–22.

Zhang, J., Xu, N., Wang, G., Zhao, H., Lin, I., & Jia, I. (2015). *In vitro* and *in vivo* antioxidant effects of polysaccharides from Nameko medicinal mushroom, *Pholiota nameko* SW-01 (Higher Basidiomycetes). *International Journal of Medicinal Mushrooms*, 17, 671-680.

Zhang, L., Liu, P., Li, L., Huang, Y., Pu, Y., Hou, X., & Song, L. (2019). Identification and antioxidant activity of flavonoids extracted from xinjiang jujube (*Ziziphus jujube* Mill.) leaves with ultra-high pressure extraction technology. *Molecules*, 24, 122.

Zhang, Y., Li, S., Wang, X., Zhang, L., & Cheung, P. C. K. (2011). Advances in lentinan: Isolation, structure, chain conformation and bioactivities. *Food Hydrocolloids*, 25, 196–206.

Zhang, X. D., Liu, X. Q., Kim, Y. H., & Whang, W. K. (2014). Chemical constituents and their acetyl cholinesterase inhibitory and antioxidant activities from leaves of Acanthopanax henryi: potential complementary source against Alzheimer's disease. *Archives of Pharmaceutical Research*, 37: 606–616.

Zhao, D., Liu, G., Song, D., Liu, J., Zhou, Y., Ou, J. & Sun, S. (2006a). Fourier transform infrared spectroscopic study of truffles. Proc. SPIE 6026, *ICO20: Biomedical Optics*, 60260H, China.

Zhao, D., Liu, G., Song, D., Liu, J., Zhou, Y., Ou, J. & Sun, S. (2006b). Identification of Amanita mushrooms by Fourier transform infrared spectroscopy. Proc. SPIE 6047, *Fourth International Conference on Photonics and Imaging in Biology and Medicine*, 60471V, China.

Zheng, L., Zhai, G., Zhang, J., Wang, L., Ma, Z., Jia, M. & Lia, L. (2014). Antihyperlipidemic and hepatoprotective activities of mycelia zinc polysaccharide from *Pholiota nameko*. *International Journal of Biological Macromolecules*, 70, 523-529.

Zhu, Z.-Y., Pan, L.-C., Han, D., Sun, H.-Q., & Chen, L.-J. (2019) Structural properties and antioxidant activities of polysaccharide from fruit bodies of *Pholiota nameko. Natural Product Research*, 33, 1563-1569.

Chapter 5

REMARKS AND FINAL CONCLUSIONS

Across academia and industry, data and text mining has become a popular strategy for keeping up with the rapid growth of the scientific literature. Text mining in the scientific literature has mostly been carried out on collections of abstracts, due to their availability, though vital data is often only available in the full-text, which is not always available and/or accessible. Here we present an analysis of more than 500 000 scientific abstracts of articles published during the period 1786–2020. We describe the development in article length and publication sub-topics during nearly 230 years. We showcase the potential of text mining by extracting published fungi, fungus, fungal, antibacterial, antioxidant, marine and terrestrial associations using an entity recognition system, and quantitatively report on their accuracy using standard benchmark data sets to subsequently compare the findings to corresponding results obtained, showing that text mining is a remarkable and reliable tool. Therefore, it becomes interesting and possible to use this type of tools as a preparatory work before starting a research study, to become aware of the state of development in the literature of a particular topic of study.

Considering that and regarding to the study of the discovering of new extracts and bioactive compounds, this is a hot topic that has attracted the attention of numerous scientists. In fact, the use of botanical preparations dates to the pre-historic era. Prior to the use of synthetic drugs, botanical preparations were used for a multitude of health conditions, including as a tool for curbing the aging process. Then, with the increasing use of synthetic molecules, natural therapies passed to a second stage. But nowadays, the focus on natural matrices properties has gained attention again. Therefore, in this thesis, the search, selection, and characterization of six different mushrooms for the bioactivity screening was pursued. From these six mushrooms, although only one demonstrated potential antibacterial activity, several showed interesting antioxidant activity. The *In vitro* studies have been increasing in an exponential manner, and numerous bioactive properties were confirmed; then, *in vivo* studies mainly in animal models have also increased, but in smaller proportions, thus, this topic can be a potential focus of study.

Despite those advances, for most terrestrial and marine species no extensive knowledge is available, namely their bioactivity, their mechanisms of action, therapeutic and prophylactic doses, synergism, antagonisms and other inter-relations between them. Clinical trials are very important, in order to develop future and effective alternatives to improve the health and well-being of individuals. In most, if not all cases, new drugs are designed with activity against various ailments. Thus, optimization of the various active compounds for maximum product quality is also necessary through which continued advancement in science can be achieved. In addition, it is of the major importance to prove the effective therapeutic potential, together with the increasing number of publications related with the discovery of marine and terrestrial species with bioactive potential, as well as to evaluate other variables, such as bioavailability and bio-efficacy interactions. The theoretical investigations should take lead in determining the interaction of active compounds with microorganisms to produce drugs with better activity. The development in predicting the growing improvement of these bioactive constituents, such as those described in this dissertation, is given a lead to groundbreaking research.

For instance, marine and terrestrial bioactive compounds in general can improve the appearance, texture, stability, and quality of finished food products, despite being incredibly attractive and suitable to the food industry in general, given their relatively cost-effective extraction methods, their natural availability, and their biological activities, the latter promoting health and reducing the burden of several diseases among individuals.

With the rapid spread of diseases, many still without a known cure, the opportunity to discover new compounds from plants, animals, and microbial sources, including a new group of compounds to combat diseases, is therefore necessary. Hence, the discovery of new drugs through isolation from medicinal natural organisms, including the ability to determine and improve the quality of bioactive compounds from the microbial sources, should remain an active field of research in the modern scientific world.

ANNEXES

Appendix I – Data mining code sample

```
public async Task<ISearchResponse<Document>> Filter(string search,
DocumentFilter filter)
{
    Func<QueryContainerDescriptor<Document>, QueryContainer> match;
    var filters = new List<Func<QueryContainerDescriptor<Document>,
QueryContainer>>();
    if(!string.IsNullOrEmpty(search))
    {
        match = q => q
            .MultiMatch(mm => mm
                .Fields(f => f.Field(p => p.Title).Field(p =>
p.Abstract))
                .Query(search)
                .MinimumShouldMatch(1)
                .AutoGenerateSynonymsPhraseQuery(false)
                .Type(TextQueryType.CrossFields)
            );
    }
    else
    {
        match = q => q
            .MatchAll();
    }

    if(!string.IsNullOrEmpty(filter.ResearchField))
    {
        filters.Add(fq => fq.Match(m => m
            .Field(df => df.ResearchField.Suffix("keyword"))
            .Query(filter.ResearchField)));
    }

    if(filter.Languages != null && filter.Languages.Any())
    {
        filters.Add(fq => fq.Terms(m => m
            .Field(df => df.Language.Suffix("keyword"))
            .Terms(filter.Languages)
            .Boost(1)
        ));
    }

    if (filter.StartYear > 0)
    {
        var startDate = new DateTime(filter.StartYear, 1, 1);
        if (filter.EndYear > 0)
        {
            var endDate = new DateTime(filter.EndYear, 12, 31);
            filters.Add(fq => fq.DateRange(m => m
                .Field(df => df.Date)
                .GreaterThanOrEquals(startDate)
                .LessThanOrEquals(endDate)
            ));
        }
        else
        {
            filters.Add(fq => fq.DateRange(m => m
```

```
                .Field(df => df.Date)
                .GreaterThanOrEquals(startDate)
                .LessThanOrEquals(DateMath.Now)
            ));
        }
    }
    else
    {
        if (filter.EndYear > 0)
        {
            var endDate = new DateTime(filter.EndYear, 12, 31);
            filters.Add(fq => fq.DateRange(m => m
                .Field(df => df.Date)
                .LessThanOrEquals(endDate)
            ));
        }
    }

    if(filter.Authors != null && filter.Authors.Any())
    {
        filters.Add(fq => fq.Terms(m => m
            .Field(df => df.Authors.Suffix("keyword"))
            .Terms(filter.Authors)
            .Boost(1)
        ));
    }

    if(filter.Countries != null && filter.Countries.Any())
    {
        filters.Add(fq => fq.Terms(m => m
            .Field(df => df.Affiliations.Suffix("country.keyword"))
            .Terms(filter.Countries)
            .Boost(1)
        ));
    }
    if(filter.Journals != null && filter.Journals.Any())
    {
        filters.Add(fq => fq.Terms(m => m
            .Field(df => df.Journal.Suffix("name.keyword"))
            .Terms(filter.Journals)
            .Boost(1)
        ));
    }
    if(filter.PublishFormats != null && filter.PublishFormats.Any())
    {
        filters.Add(fq => fq.Terms(m => m
            .Field(df => df.PublishFormat.Suffix("keyword"))
            .Terms(filter.PublishFormats)
            .Boost(1)
        ));
    }

    if(filter.DocumentTypes != null && filter.DocumentTypes.Any())
    {
        filters.Add(fq => fq.Terms(m => m
            .Field(df => df.PublishedAs.Suffix("keyword"))
            .Terms(filter.DocumentTypes)
            .Boost(1)
        ));
```

```
    }

    if(filter.Institutions != null && filter.Institutions.Any())
    {
        filters.Add(fq => fq.Terms(m => m
            .Field(df => df.Affiliations.Suffix("institution.keyword"))
            .Terms(filter.Institutions)
            .Boost(1.1)
        ));
    }
    if(filter.Keywords != null && filter.Keywords.Any())
    {
        filters.Add(fq => fq.Terms(m => m
            .Field(df => df.Keywords.Suffix("keyword"))
            .Terms(filter.Keywords)
            .Boost(1.1)
        ));
    }

    if(filter.Cities != null && filter.Cities.Any())
    {
        filters.Add(fq => fq.Terms(m => m
            .Field(df => df.Affiliations.Suffix("city.keyword"))
            .Terms(filter.Cities)
            .Boost(1.1)
        ));
    }

    var query = await _elasticClient.SearchAsync<Document>(s => s
        .Query(q => q
            .Bool(b => b
                .Must(match)
                .Filter(filters)
            )
        )
```

Figure 3.8. Code for calculation of the relevant aggregations for the desired fields in study (authors, countries, journals, publisher formats, document types, institutions, keywords, cities).

```
.DateHistogram("hist_pubs_per_year", date => date
     .Field(p => p.Date)
     .CalendarInterval(DateInterval.Year)
     .MinimumDocumentCount(50)
     .Order(HistogramOrder.KeyAscending)
)
```

Figure 3.9. Code of date histograms of publications during the “two past centuries”.

```
Document>> GetAggregations(string query = null)
{
    Func<QueryContainerDescriptor<Document>, QueryContainer> qd = null;
    if (string.IsNullOrEmpty(query))
    {
        qd = q => q.MatchAll();
    }
    else
    {
```

```
        qd = q => q.MultiMatch(mm => mm
            .Fields(ff => ff
                .Field(f => f.Abstract)
                .Field(f => f.Title)
            )
            .Query(query)
        );
    }
    return await _elasticClient.SearchAsync<Document>(s =>
        s.Query(qd)
        .Aggregations(ag => ag
            .DateHistogram("hist_pubs_per_year", date => date
                .Field(p => p.Date)
                .CalendarInterval(DateInterval.Year)
                .MinimumDocumentCount(50)
                .Order(HistogramOrder.KeyAscending)
            )
            .SignificantTerms("SignificantTerms_Most-Significant-Keywords", st =>
st
                .Field(p => p.Keywords.Suffix("keyword"))
                .MinimumDocumentCount(20)
                .MutualInformation(mi => mi
                    .BackgroundIsSuperSet()
                    .IncludeNegatives()
                )
            )
            .Terms("distrib_Cities", t => t
                .Field(f => f.Affiliations.Suffix("city.keyword"))
                .ExecutionHint(TermsAggregationExecutionHint.Map)
                .Order(d => d.CountDescending())
                .Aggregations(agg => agg
                    .Terms("distrib_departments", tt => tt
                        .Field(ff =>
ff.Affiliations.Suffix("department.keyword"))
                        .ExecutionHint(TermsAggregationExecutionHint.Map)
                        .Order(d => d.CountDescending())
                        .Size(5)
                    )
                    .Terms("distrib_institutions", tt => tt
                        .Field(ff =>
ff.Affiliations.Suffix("institutionName.keyword"))
                        .ExecutionHint(TermsAggregationExecutionHint.Map)
                        .Order(d => d.CountDescending())
                        .Size(5)
                    )
                )
            )
            .Terms("distrib_pubs_Countries", f => f
                .Field(ff => ff.Affiliations.Suffix("country.keyword"))
                .ExecutionHint(TermsAggregationExecutionHint.Map)
                .Order(d => d.CountDescending())
                .Size(300)
            )
            .Terms("TopN_Countries", f => f
                .Field(ff => ff.Affiliations.Suffix("country.keyword"))
                .ExecutionHint(TermsAggregationExecutionHint.Map)
                .Order(d => d.CountDescending())
                .Size(10)
            )
            .Terms("TopN_Journals", f => f
                .Field(ff => ff.Journal.Name.Suffix("keyword"))
                .ExecutionHint(TermsAggregationExecutionHint.Map)
                .Order(d => d.CountDescending())
                .Size(10)
            )
```

```
            .Terms("distrib_Journals", f => f
                .Field(ff => ff.Journal.Name.Suffix("keyword"))
                .ExecutionHint(TermsAggregationExecutionHint.Map)
                .Order(d => d.CountDescending())
            )
            .Terms("TopN_pub_Document-Types", f => f
                .Field(ff => ff.PublishedAs.Suffix("keyword"))
                .ExecutionHint(TermsAggregationExecutionHint.Map)
                .Order(d => d.CountDescending())
                .Size(10)
            )
            .Terms("distrib_pub_Document-Types", f => f
                .Field(ff => ff.PublishedAs.Suffix("keyword"))
                .ExecutionHint(TermsAggregationExecutionHint.Map)
                .Order(d => d.CountDescending())
            )
            .Terms("distrib_pub_Publication-Format", f => f
                .Field(ff => ff.PublishFormat.Suffix("keyword"))
                .ExecutionHint(TermsAggregationExecutionHint.Map)
                .Order(d => d.CountDescending())
            )
            .Terms("TopN_pub_Publication-Format", f => f
                .Field(ff => ff.PublishFormat.Suffix("keyword"))
                .ExecutionHint(TermsAggregationExecutionHint.Map)
                .Order(d => d.CountDescending())
                .Size(10)
            )
            .Terms("TopN_Languages", f => f
                .Field(ff => ff.Language.Suffix("keyword"))
                .ExecutionHint(TermsAggregationExecutionHint.Map)
                .Order(d => d.CountDescending())
                .Size(5)
            )
            .Terms("distrib_Languages", f => f
                .Field(ff => ff.Language.Suffix("keyword"))
                .ExecutionHint(TermsAggregationExecutionHint.Map)
                .Order(d => d.CountDescending())
            )
            .Terms("TopN_Authors", f => f
                .Field(ff => ff.Authors.Suffix("keyword"))
                .ExecutionHint(TermsAggregationExecutionHint.Map)
                .Order(d => d.CountDescending())
                .Size(10)
            )
            .Terms("distrib_Authors", f => f
                .Field(ff => ff.Authors.Suffix("keyword"))
                .ExecutionHint(TermsAggregationExecutionHint.Map)
                .Order(d => d.CountDescending())
            )
            .Terms("distrib_Keywords", f => f
                .Field(ff => ff.Keywords.Suffix("keyword"))
                .ExecutionHint(TermsAggregationExecutionHint.Map)
                .Order(d => d.CountDescending())
                .Size(100)
            )
        )
        .Sort(d => d.Descending(document => document.Date))
        .Size(0)
    );
}
```

Figure 3.10. Code sample for search specific and relevant aggregations (“TopN”) in study (countries, journals, document types, publication formats, language, authors, keywords).

```
.Terms("TopN_Countries", t => t
    .Field(f => f.Affiliations.Suffix("country.keyword"))
    .ExecutionHint(TermsAggregationExecutionHint.Map)
    .Order(d => d.CountDescending())
    .Size(5)
    .Aggregations(agg => agg
        .Terms("TopN_Journals", tt => tt
            .Field(ff => ff.Journal.Suffix("name.keyword"))
            .ExecutionHint(TermsAggregationExecutionHint.Map)
            .Order(d => d.CountAscending())
            .Size(5)
            .Aggregations(agg2 => agg2
                .Terms("TopN_Authors", ta => ta
                    .Field(ff => ff.Authors.Suffix("keyword"))
                    .ExecutionHint(TermsAggregationExecutionHint.Map)
                    .Order(d => d.CountDescending())
                    .Size(5)
                )
            )
        )
    )
)


.Terms("TopN_Cities", t => t
    .Field(f => f.Affiliations.Suffix("city.keyword"))
    .ExecutionHint(TermsAggregationExecutionHint.Map)
    .Order(d => d.CountDescending())
    .Size(5)
    .Aggregations(agg => agg
        .Terms("TopN_Journals", tt => tt
            .Field(ff => ff.Journal.Suffix("name.keyword"))
            .ExecutionHint(TermsAggregationExecutionHint.Map)
            .Order(d => d.CountAscending())
            .Size(5)
            .Aggregations(agg2 => agg2
                .Terms("TopN_Authors", ta => ta
                    .Field(ff => ff.Authors.Suffix("keyword"))
                    .ExecutionHint(TermsAggregationExecutionHint.Map)
                    .Order(d => d.CountDescending())
                    .Size(5)
                )
            )
        )
    )
)
```

Figure 3.11. Analysis of the top 5 countries and top 5 cities within the top 5 journals and the corresponding top 5 author (aggregations).

Appendix II - Elsevier data extraction using R programming language

```
# keywords <- '"marine fungi"

keywords <- 'fungi OR fungus OR fungal AND antibacterial AND anti-bacterial AND antioxidant AND anti-oxidant AND marine AND terrestrial'

#Define Time period & Number of papers

year_min <- 2010 # last five years should be enough
year_max <- 2020

# skip_step_large <- seq(0, 400, 100) # e.g.: seq(0, 800, 100) will collect 900 articles, in increments of 100
skip_step_small <- seq(0, 800, 100)

get_abstract <- function(doi_i){
```

Figure 3.12. Code for definition of the keywords and time interval to be applied in the study.

```
#CountWords Plot
publications_tokens %>%
  count(word, sort = TRUE) %>%
  top_n(10, n) %>%
  ggplot(aes(x = fct_reorder(word, n), y = n,fill=word)) +
  geom_bar(stat="identity",show.legend = FALSE) +
  coord_flip() +theme(axis.text.x = element_text(angle = 90, hjust = 1))+theme_minimal()

# 1. Retrieve and pre-process data ####
# _1.1 Extend search keywords ####
 keywords_1 <- 'Fungi OR Fungus OR Fungal'
 keywords_2 <- 'Fungi OR Fungus OR Fungal AND marine'
 keywords_3 <- 'Fungi OR Fungus OR Fungal AND terrestrial'
 keywords_4 <- 'Fungi OR Fungus OR Fungal AND antibacterial'
 keywords_5 <- 'Fungi OR Fungus OR Fungal AND anti-bacterial'
 keywords_6 <- 'Fungi OR Fungus OR Fungal AND antioxidant'
 keywords_7 <- 'Fungi OR Fungus OR Fungal AND anti-oxidant'

# _1.2 Search relevant papers by set of keywords ####
 publications_1 <- pmap_dfr(crossing("keywords" = keywords_1, year_min, year_max, "skip_step" = skip_step_small), pub_data)
 publications_2 <- pmap_dfr(crossing("keywords" = keywords_2, year_min, year_max, "skip_step" = skip_step_large), pub_data)
 publications_3 <- pmap_dfr(crossing("keywords" = keywords_3, year_min, year_max, "skip_step" = skip_step_large), pub_data)
 publications_4 <- pmap_dfr(crossing("keywords" = keywords_4, year_min, year_max, "skip_step" = skip_step_large), pub_data)
 publications_5 <- pmap_dfr(crossing("keywords" = keywords_5, year_min, year_max, "skip_step" = skip_step_small), pub_data)
 publications_6 <- pmap_dfr(crossing("keywords" = keywords_6, year_min, year_max, "skip_step" = skip_step_large), pub_data)
 publications_7 <- pmap_dfr(crossing("keywords" = keywords_7, year_min, year_max, "skip_step" = skip_step_small), pub_data)


# _1.3 Bind all search results  ####
 publications_all <- bind_rows(publications_1, publications_2, publications_3,
                               publications_4, publications_5) %>%
   distinct()


# _1.4 Collect title and abstracts for the publications of interest  ####
 abstracts_all <- publications_all %>%
   mutate(abstract = map_chr(doi, possibly(get_abstract, otherwise = NA_character_)))

 save(abstracts_all, file = "Abstracts_2010_2020_papers.RData")

 titles_all <- publications_all %>%
   mutate(titles = map_chr(doi, possibly(get_abstract, otherwise = NA_character_)))

 save(titles_all, file = "Abstracts_2010_2020_papers.RData")
```

Figure 3.13. Code for extract relevant scientific articles by using a set of keywords.

	keywords	doi	author	date	publication	title	abstract
2728	Fungi OR fungus OR fungal AND	10.1016/j.phytochem.2019.05.001	Ma José Iglesias	2019-08-...	Phytochemistry	NMR characterization and evaluation of ...	Abstract The chemical composition of
2729	Fungi OR fungus OR fungal AND	10.1016/j.ecoleng.2016.02.007	Hristina Bodin	2016-06-...	Ecological Engineering	Effects of biopellets composed of microa...	Abstract Removal of seven pharmace
2730	Fungi OR fungus OR fungal AND	10.1016/j.scitotenv.2009.01.054	A. C. Freitas	2009-05-...	Science of The Total Environment	Biological treatment of the effluent from...	Abstract Three white-rot fungi (Pleuro
2731	Fungi OR fungus OR fungal AND	10.1016/j.funeco.2014.12.001	Xiaohong Guo	2015-04-...	Fungal Ecology	Marine fungal communities in water and...	Abstract The planktonic and benthic f
2732	Fungi OR fungus OR fungal AND	10.1016/j.bjp.2018.01.005	Gabriele Andressa Zatelli	2018-04-...	Revista Brasileira de Farmacognosia	An overview of odoriferous marine seaw...	Abstract Since the middle of the twen
2733	Fungi OR fungus OR fungal AND	10.1016/j.conbuildmat.2012.12.022	P. Hughes	2013-04-...	Construction and Building Materials	Microscopic examination of a new mech...	Abstract This research presents an exp
2734	Fungi OR fungus OR fungal AND	10.1016/bs.mie.2018.02.018	Marco N. Allemann	2018-12-...	Methods in Enzymology	Chapter One: Characterization and Appli...	Abstract The long-chain omega-3 pol
2735	Fungi OR fungus OR fungal AND	10.1016/j.copbio.2018.02.001	Joseph H Collins	2018-10-...	Current Opinion in Biotechnology	Genetic engineering of host organisms f...	Pharmaceutical production hosts may
2736	Fungi OR fungus OR fungal AND	10.1016/j.sedgeo.2017.09.011	Eva De Boever	2017-11-...	Sedimentary Geology	What do we really know about early diag...	Abstract Non-marine carbonate rocks
2737	Fungi OR fungus OR fungal AND	10.1016/S1473-3099(16)30323-1	Usama Ramadan Abdelm...	2017-02-...	The Lancet Infectious Diseases	Potential of marine natural products agai...	Summary Antibiotics have revolutioni
2738	Fungi OR fungus OR fungal AND	10.1016/j.peptides.2013.04.004	Ming-Ching Lin	2013-06-...	Peptides	Truncated antimicrobial peptides from m...	Abstract Antimicrobial peptides (AMP
2739	Fungi OR fungus OR fungal AND	10.1016/j.bcab.2019.01.016	G. Ramachandran	2019-01-...	Biocatalysis and Agricultural Biotec...	Extraction and partial purification of sec...	Abstract The aim of this study is to id
2740	Fungi OR fungus OR fungal AND	10.1016/j.fitote.2017.08.010	Dan Xu	2017-10-...	Fitoterapia	New alkenylated tetrahydropyran derivat...	Abstract Six new alkenylated tetrahyd
2741	Fungi OR fungus OR fungal AND	10.1016/j.phytol.2015.05.014	Ibrahim A. M. Asiri	2015-09-...	Phytochemistry Letters	Penicillivinacine, antimigratory diketopip...	Abstract In the course of our ongoing
2742	Fungi OR fungus OR fungal AND	10.1016/j.micres.2013.07.014	Panchanathan Manivasag...	2014-04-...	Microbiological Research	Pharmaceutically active secondary meta...	Abstract Marine actinobacteria are or
2743	Fungi OR fungus OR fungal AND	10.1016/j.jaad.2009.01.046	Vidal Haddad	2009-11-...	Journal of the American Academy ...	Tropical dermatology: Marine and aquati...	Dermatoses caused by marine organi
2744	Fungi OR fungus OR fungal AND	10.1016/j.ibiod.2014.08.010	Carlotta Sacco Perasso	2015-04-...	International Biodeterioration & Bi...	Marine bioerosion of lapideous archaeol...	Abstract This research has been cond
2745	Fungi OR fungus OR fungal AND	10.1016/j.bmcl.2015.06.010	Bahaa El-Dien M. El-Gendy	2015-08-...	Bioorganic & Medicinal Chemistry ...	Antibacterial activity of diketopiperazine...	Abstract Nine diketopiperazines were
2746	Fungi OR fungus OR fungal AND	10.1016/j.scitotenv.2017.02.017	Ana Paço	2017-05-...	Science of The Total Environment	Biodegradation of polyethylene micropla...	Abstract Plastic yearly production has
2747	Fungi OR fungus OR fungal AND	10.1016/j.margen.2014.07.003	Jean-François Bloch	2014-10-...	Marine Genomics	Marine biotechnologies and synthetic bi...	Abstract The sea will be a source of e
2748	Fungi OR fungus OR fungal AND	10.1016/j.jenvrad.2015.01.012	N. S. Kudryasheva	2015-04-...	Journal of Environmental Radioacti...	Effect of low-dose ionizing radiation on l...	Abstract The paper summarizes studi
2749	Fungi OR fungus OR fungal AND	10.1016/j.scitotenv.2014.12.002	Ying Li	2015-03-...	Science of The Total Environment	Enhanced biodegradation of phthalate a...	Abstract Cylindrotheca closterium, a r
2750	Fungi OR fungus OR fungal AND	10.1016/B978-0-12-416003-3.00005-6	Anupam Giri	2012-12-...	Advances in Food and Nutrition Re...	Chapter 5: Bioactive Marine Peptides: Nu...	Abstract Marine organisms represent

Figure 3.14. Compilation of number of abstracts extracted.

```
# _1.5 Tokeninze abstracts ####
abstracts_tokens <- abstracts_all %>%
  unnest_tokens(token = "words", input = "abstract", output = "word", drop = FALSE) %>%
  # remove stop-words
  anti_join(stop_words, by = "word") %>%
  # remove tokens containing numbers
  filter(!str_detect(word, "[:digit:]")) %>%
  # remove unwanted words
  filter(!word %in% c("abstract","study","sp","ml","including", "analysis","results","anti","based","identified","review",
                      "sources","proportions","environment","production","derived","bacteria","secondary","oil","obtained","acid","liquid","content",
                      "longan","solid","ii","compared","model","conditions","low","assisted","ratio","samples","total","ps",
                      "treatment","dry","matter","uae","pop","ip","key","approach","scope","conclusions","background","findings"))
```

Figure 3.15. Code sample used for data cleaning.

```r
# 2. Analysis ####
# _2.1 Plots (word frequency count, wordcloud) ####

# CountWords Plot

r<-abstracts_tokens %>%
  count(word, sort = TRUE) %>%
  top_n(10, n) %>%
  ggplot(aes(x = fct_reorder(word, n), y = n,fill=word)) +
  geom_bar(stat="identity",show.legend = FALSE) +
  coord_flip() +theme(axis.text.x = element_text(angle = 90, hjust = 1))+theme_minimal()+
  labs(x = "Number of words", y = "Keyword",title = "Total Number of Words")

 r

 ggsave(plot = r, width = 15, height = 15, dpi = 300, filename = "WordTotal.pdf")


# we add some more metadata columns to the data frame
abstracts_tokens$year <- substr(abstracts_tokens$date, 0, 4)


# CountWords Plot by Year
q<-abstracts_tokens %>% group_by(year) %>%
  count(word, sort = FALSE) %>%
  top_n(10) %>%
  drop_na(word) %>%
  ggplot(aes(reorder(word, n), n, fill = year)) +
  geom_col() + scale_y_continuous(breaks=c(0,200,400,600,800,1000,1200,1400))+
  coord_flip() +
  facet_wrap(~year, scales = "free_y") +
  labs(x = NULL) +
  guides(fill = FALSE) +
  scale_fill_brewer(palette = "Set3")

q

ggsave(plot = q, width = 15, height = 15, dpi = 300, filename = "WordbyYear.pdf")
```

Figure 3.16. Code sample for count of words most frequent (Top 10) between 2010 and 2020 and the count of words most frequent by year (Top 10).

```r
# Find 10 top journals

abstracts_tokens1 <- abstracts_tokens %>%
  group_by(publication) %>%
  mutate(n= n()) %>%
  distinct(publication, .keep_all=TRUE) %>%
  arrange (desc(n)) %>%
  ungroup() %>%
  top_n(10)

# CountWords Plot by Journal

 p<-abstracts_tokens %>% group_by(publication) %>%
   count(word, sort = TRUE) %>%
   filter(publication %in% c("Process Biochemistry","Asian Pacific Journal of Tropical Biomedicine",
                             "Advances in Food and Nutrition Research","Biocatalysis and Agricultural Biotechnology",
                             "International Journal of Biological Macromolecules","Marine Pollution Bulletin",
                             "Phytochemistry Letters","Fungal Ecology","Microbiological Research")) %>%
   top_n(10) %>%
   ggplot(aes(reorder(word,n), n, fill = publication)) +
   geom_col() +
   coord_flip() +
   facet_wrap(~ publication, scales = "free_y") +
   labs(x = NULL) +
   guides(fill = FALSE) +
   scale_fill_brewer(palette = "Set3")

 p

 ggsave(plot = p, width = 15, height = 15, dpi = 300, filename = "WordbyJournal.pdf")
```

Figure 3.17. Code for determine the Top 10 journals using R programming language.

Appendix II – ETL graphics and tables representations

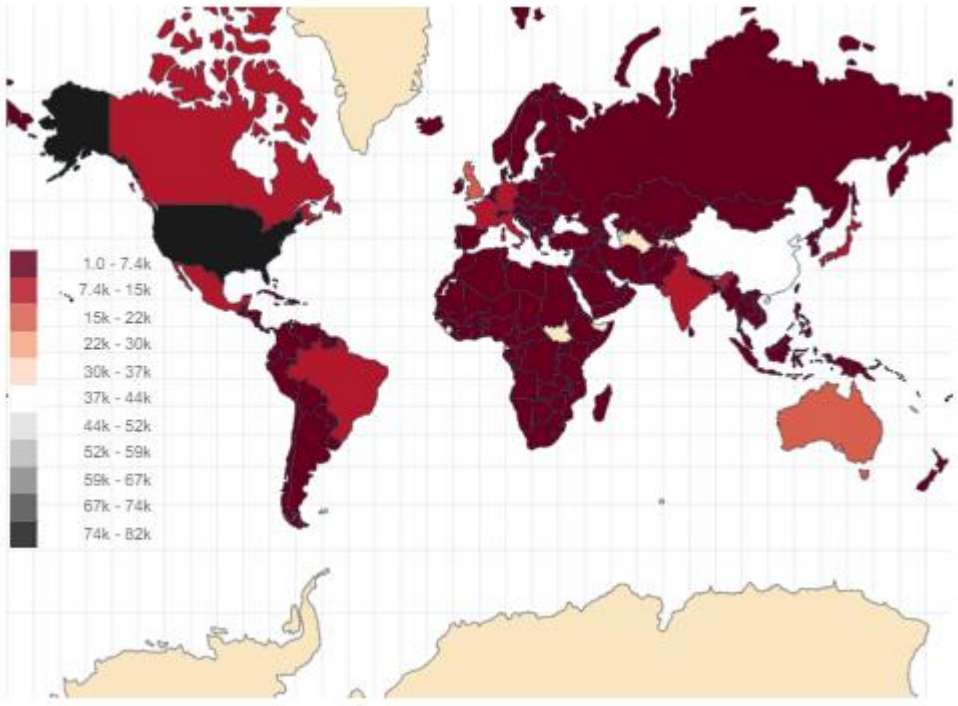

Figure 3.20. Global geographic distribution of scientific publications between 2010 and 2020. Number of scientific papers expressed in k (thousand) units with a colour gradient.

Table 3.2. Publications per year extracted from Europe PMC containing “fungi, fungus, or fungal” keywords in the Titles and Abstracts between 1786 and 2020.

Date	Obtained Data (number of publications)	Date	Obtained Data (number of publications)	Date	Obtained Data (number of publications)
1786	2	**1808**	4	**1821**	2
1787	1	**1809**	1	**1822**	3
1789	2	**1810**	4	**1823**	4
1792	1	**1811**	3	**1824**	2
1793	1	**1812**	7	**1825**	3
1797	2	**1813**	3	**1826**	5
1800	5	**1814**	6	**1827**	8
1801	2	**1815**	1	**1828**	7
1803	2	**1816**	5	**1829**	11
1804	4	**1817**	4	**1830**	7
1805	3	**1818**	7	**1831**	10
1806	3	**1819**	6	**1832**	7

Date	Obtained Data (number of publications)	Date	Obtained Data (number of publications)	Date	Obtained Data (number of publications)
1807	6	**1820**	4	**1833**	1
1834	2	**1869**	8	**1907**	7
1835	3	**1870**	8	**1908**	23
1836	2	**1871**	12	**1909**	25
1837	3	**1872**	15	**1910**	20
1838	2	**1873**	5	**1911**	33
1839	2	**1874**	1	**1912**	18
1840	1	**1876**	3	**1913**	24
1841	8	**1879**	1	**1914**	29
1842	6	**1880**	2	**1915**	26
1843	5	**1881**	2	**1916**	25
1844	11	**1882**	2	**1917**	30
1845	10	**1883**	2	**1918**	29
1846	4	**1884**	6	**1919**	20
1847	5	**1885**	4	**1920**	36
1848	12	**1886**	3	**1921**	25
1849	8	**1887**	5	**1922**	29
1850	6	**1888**	5	**1923**	36
1851	10	**1889**	5	**1924**	40
1852	6	**1890**	5	**1925**	37
1853	20	**1891**	2	**1926**	32
1854	14	**1892**	14	**1927**	46
1855	11	**1893**	20	**1928**	40
1856	13	**1894**	18	**1929**	52
1857	9	**1895**	14	**1930**	33
1858	15	**1896**	18	**1931**	65
1859	12	**1897**	21	**1932**	63
1860	3	**1898**	23	**1933**	57
1861	7	**1899**	16	**1934**	42
1862	11	**1900**	14	**1935**	46
1863	12	**1901**	10	**1936**	44
1864	8	**1902**	13	**1937**	45
1865	8	**1903**	9	**1938**	49
1866	8	**1904**	16	**1939**	69
1867	10	**1905**	23	**1940**	61

Date	Obtained Data (number of publications)	Date	Obtained Data (number of publications)	Date	Obtained Data (number of publications)
1868	6	**1906**	15	**1941**	61
1942	71	**1969**	1 223	**1996**	6 497
1943	59	**1970**	1 184	**1997**	7 045
1944	84	**1971**	1 373	**1998**	7 367
1945	86	**1972**	1 456	**1999**	7 765
1946	190	**1973**	1 743	**2000**	8 603
1947	224	**1974**	1 723	**2001**	8 893
1948	205	**1975**	1 928	**2002**	9 257
1949	212	**1976**	1 849	**2003**	9 286
1950	306	**1977**	1 801	**2004**	10 307
1951	349	**1978**	1 736	**2005**	11093
1952	413	**1979**	1 977	**2006**	11 622
1953	513	**1980**	1 912	**2007**	12 197
1954	496	**1981**	1 958	**2008**	13 694
1955	459	**1982**	2 060	**2009**	14 962
1956	494	**1983**	2 323	**2010**	16 884
1957	481	**1984**	2 693	**2011**	18 983
1958	503	**1985**	2 806	**2012**	21 101
1959	519	**1986**	2 886	**2013**	22 472
1960	512	**1987**	3 276	**2014**	24 960
1961	558	**1988**	3 491	**2015**	25 978
1962	555	**1989**	3 908	**2016**	26 855
1963	679	**1990**	4 463	**2017**	29 000
1964	816	**1991**	4 827	**2018**	31 568
1965	869	**1992**	5 212	**2019**	33 330
1966	851	**1993**	5 609	**2020**	18 156
1967	970	**1994**	6 179		
1968	1 184	**1995**	6 234		

Table 3.3. Publications per year containing “fungi, fungus and fungal” keywords in the Titles and Abstracts between Europe PMC and Web of Science collections.

Date	Obtained Data (number of publications)	Europe PMC Data (number of publications)	Web of Science Data (number of publications)
2010	7686	8905	8328
2011	8544	9161	9283
2012	8976	9134	9746
2013	9298	9451	10009
2014	10240	10406	10286
2015	10469	10642	11716
2016	10667	10982	12561
2017	11279	11665	12661
2018	12473	13146	13768
2019	13669	14715	15384
2020	7616	17005	16737

Table 3.4. Number of publications by relevant terms searched specifically in the Title, Abstracts or just one of them.

Terms	#documents with term in Title (number of publications)	#documents with term in Abstract (number of publications)	#Documents with term in Title OR Abstract (number of publications)
Fungi	22240	82033	90862
Fungus	17344	54713	61778
Fungal	29242	109536	119218
Antibacterial	2285	8657	9257
Anti-bacterial	76	480	526
Antioxidant	2253	6867	7288
Anti-oxidant	42	346	391
Marine	3027	5417	5960
Terrestrial	351	2607	2714

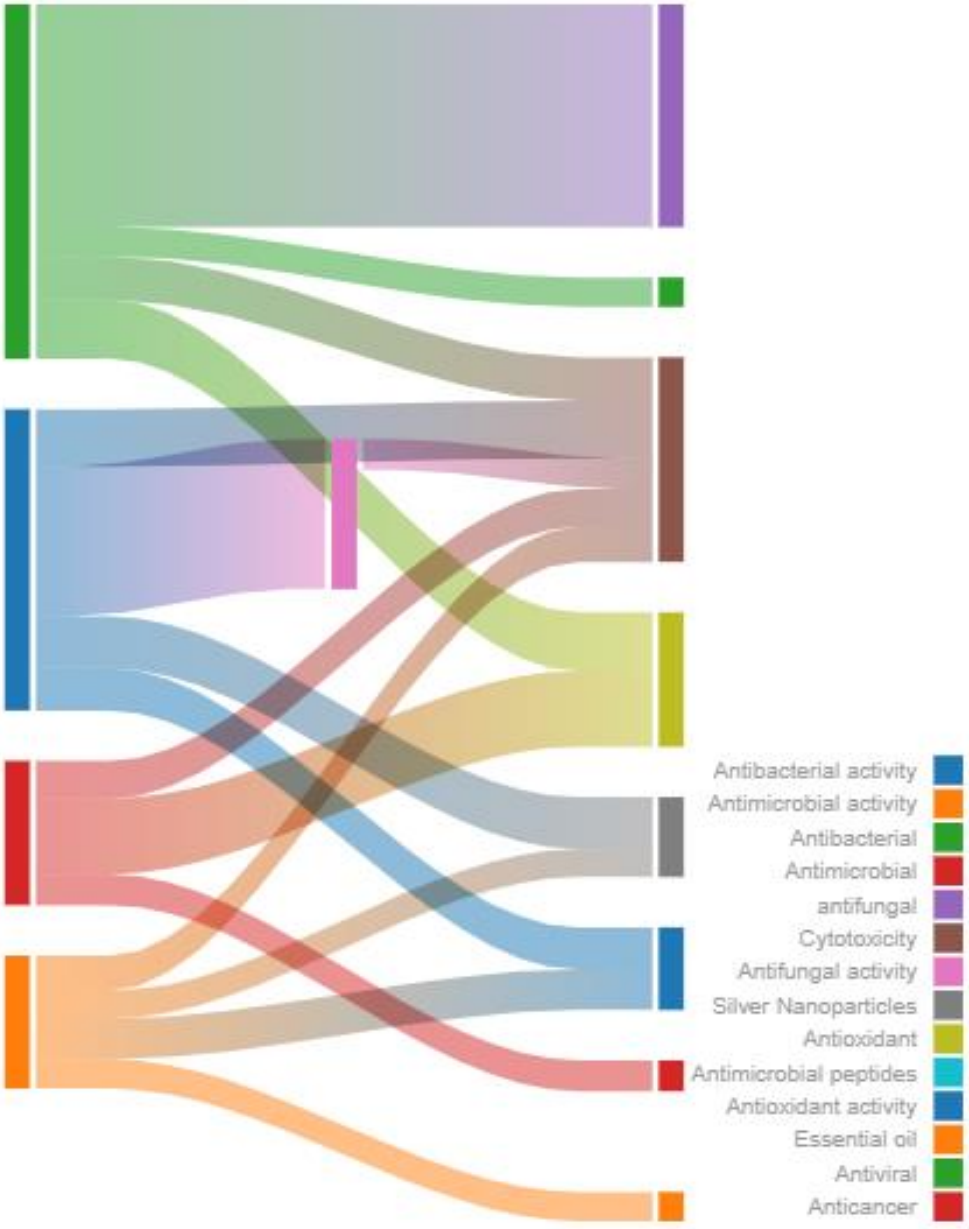

Figure 3.25. Main fourteen activities of marine and terrestrial bioactive compounds fungi related published between 1932 and 2020.

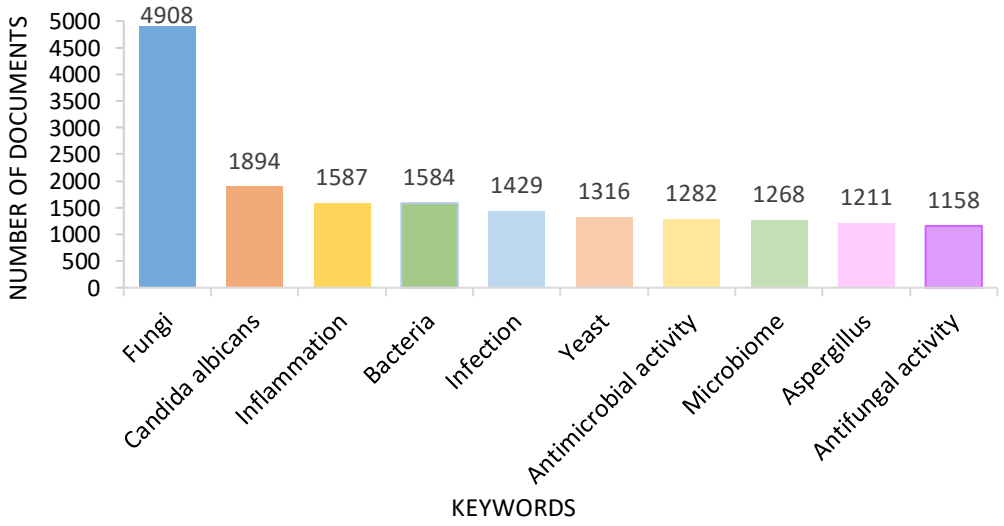

Figure 3.26. Top 10 keywords between 1786 and 2020.

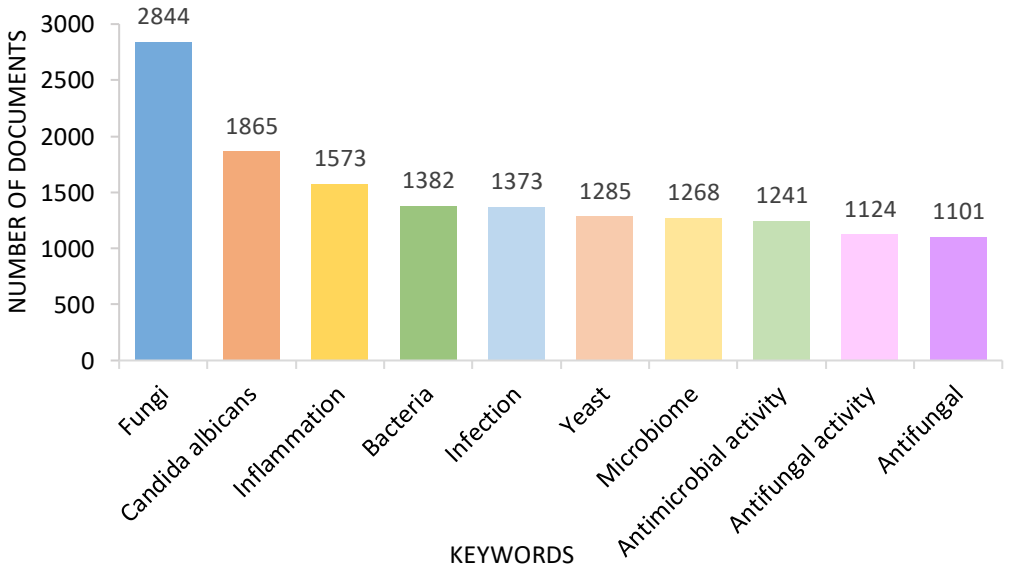

Figure 3.28. Top 10 keywords between 2010 and 2020.

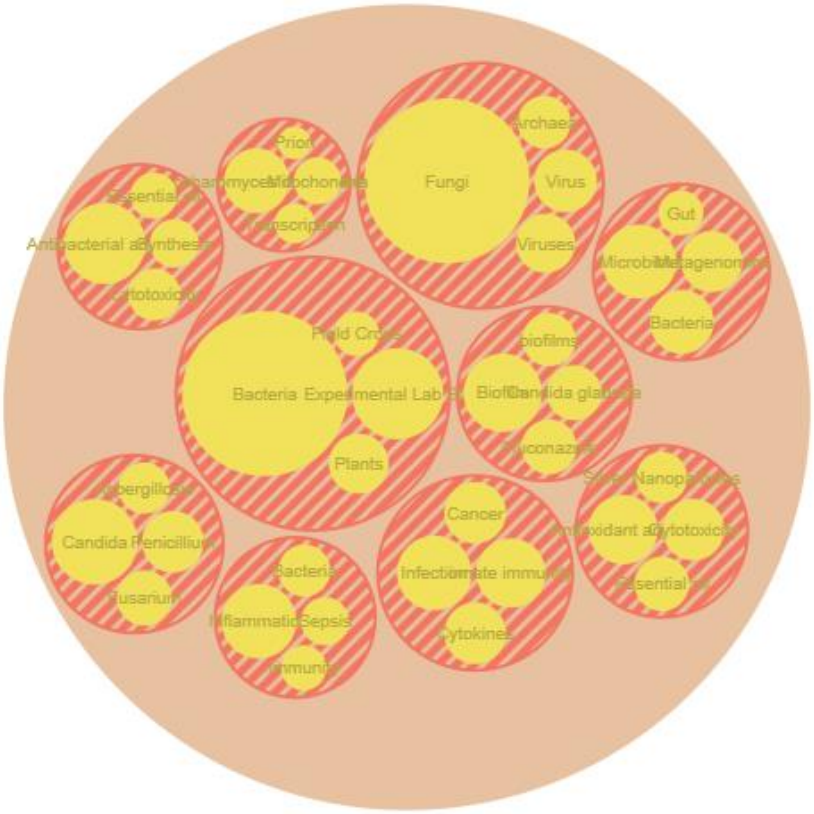

Figure 3.30. A different way to represent the Top 10 keywords.

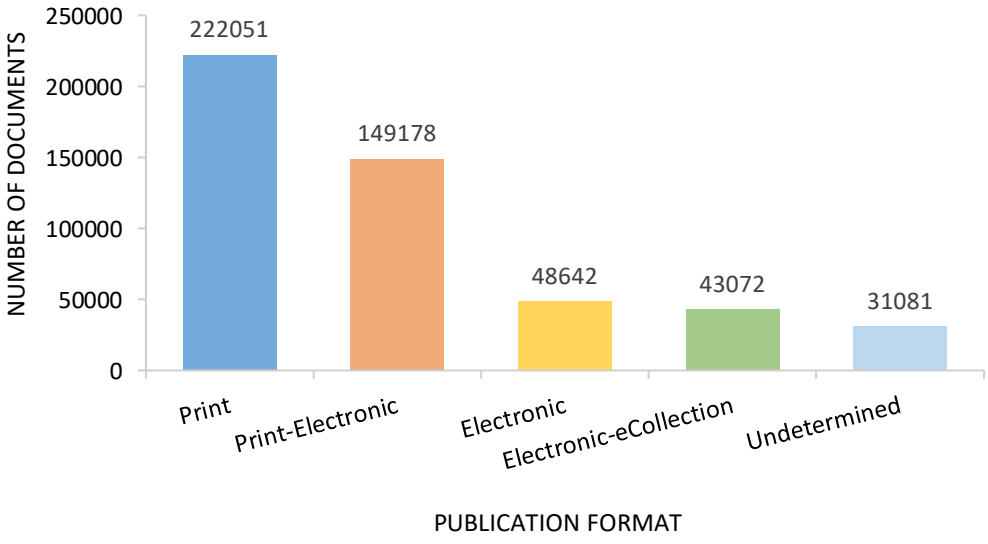

Figure 3.31. Top 5 publication formats between 1786 and 2020.

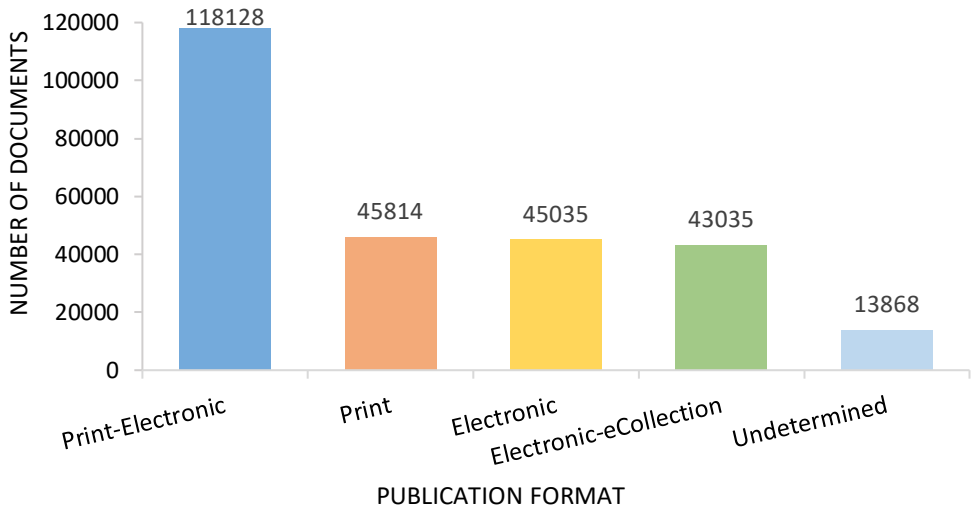

Figure 3.32. Top 5 publication formats between 2010 and 2020.

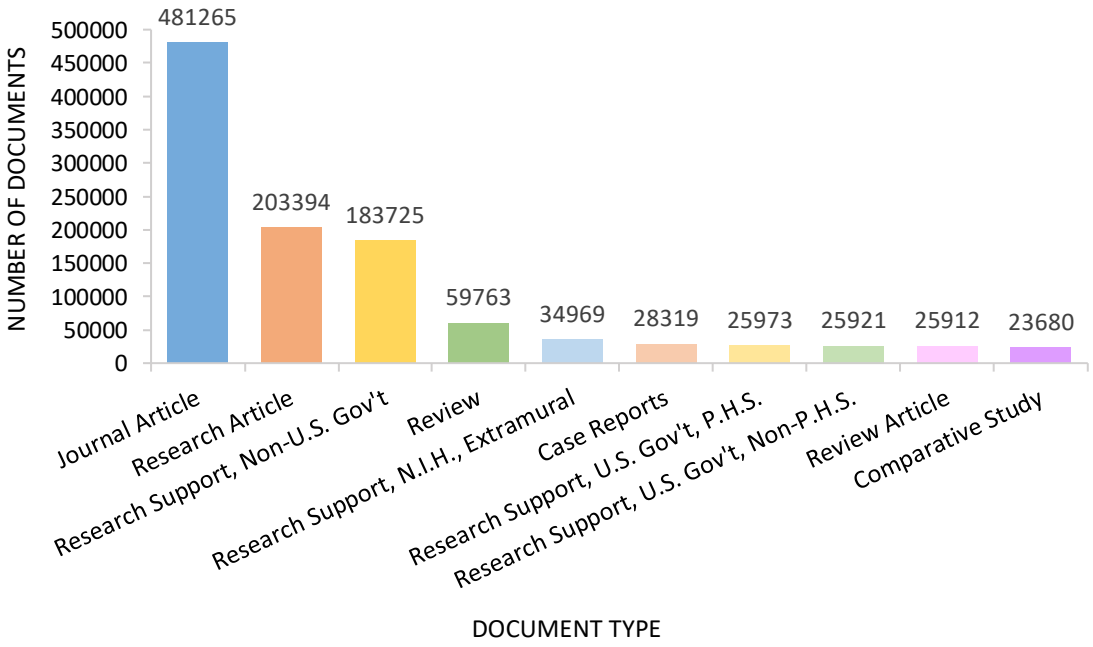

Figure 3.33. Top 10 document types between 1786 and 2020.

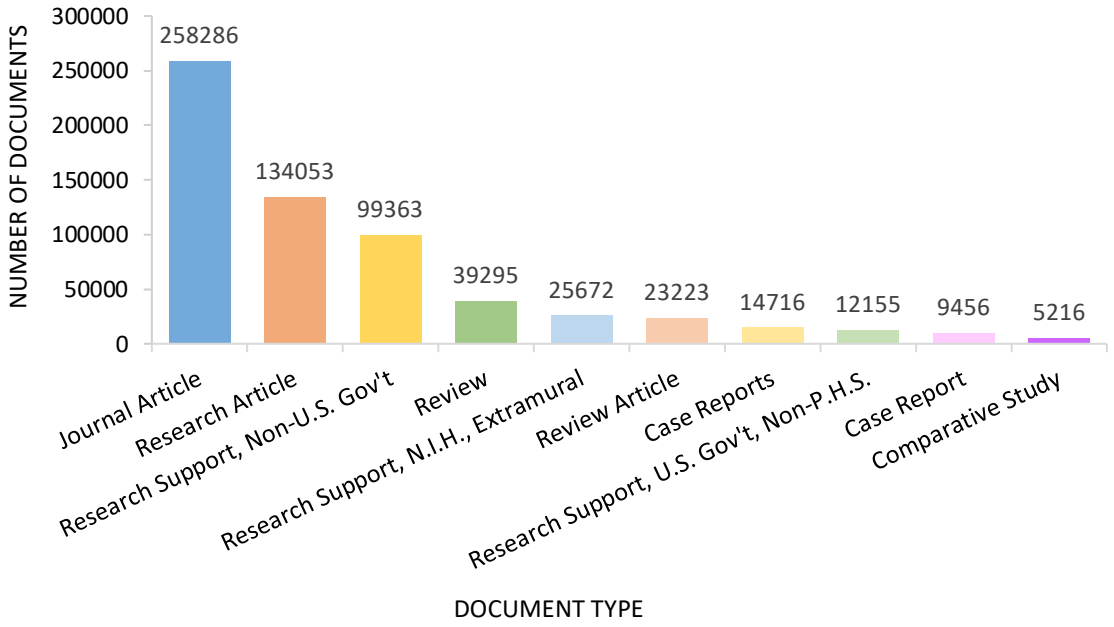

Figure 3.34. Top 10 document type between 2010 and 2020.

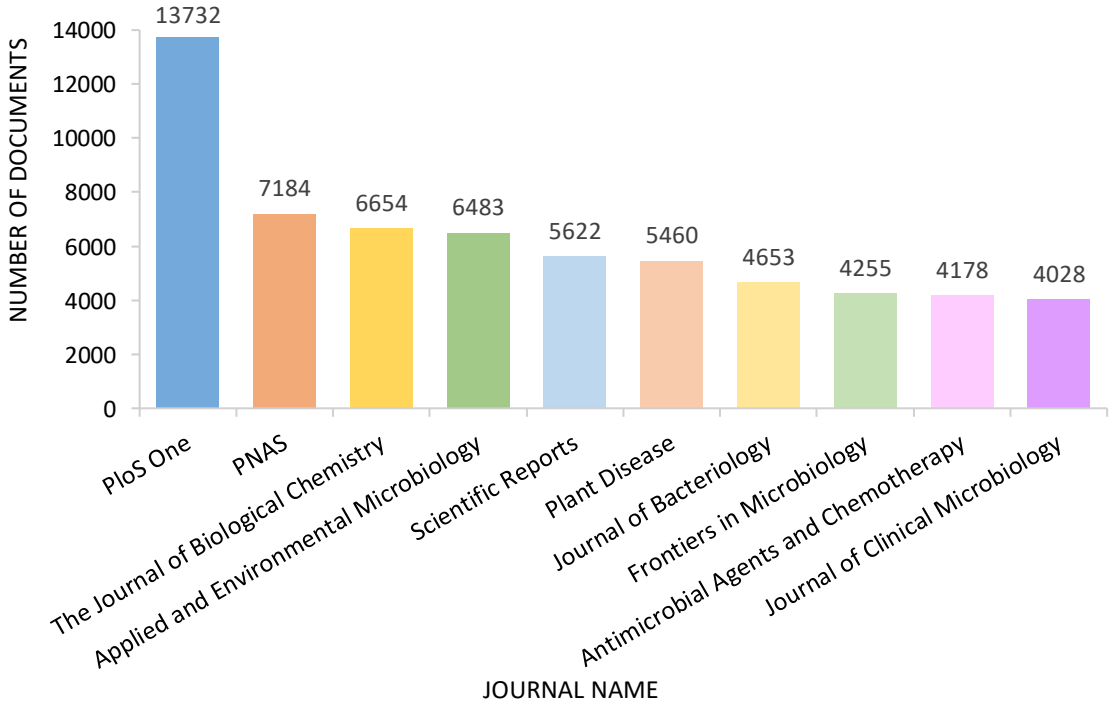

Figure 3.35. Top 10 journals between 1786 and 2020.

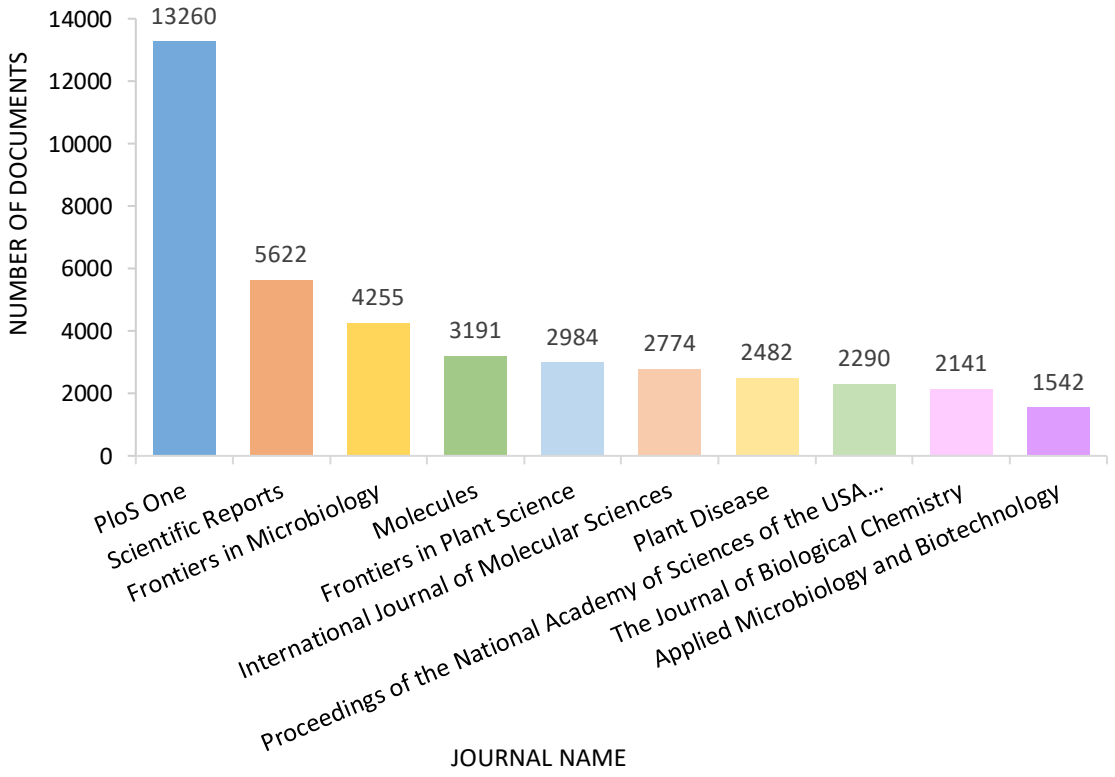

Figure 3.36. Top 10 journals between 2010 and 2020.

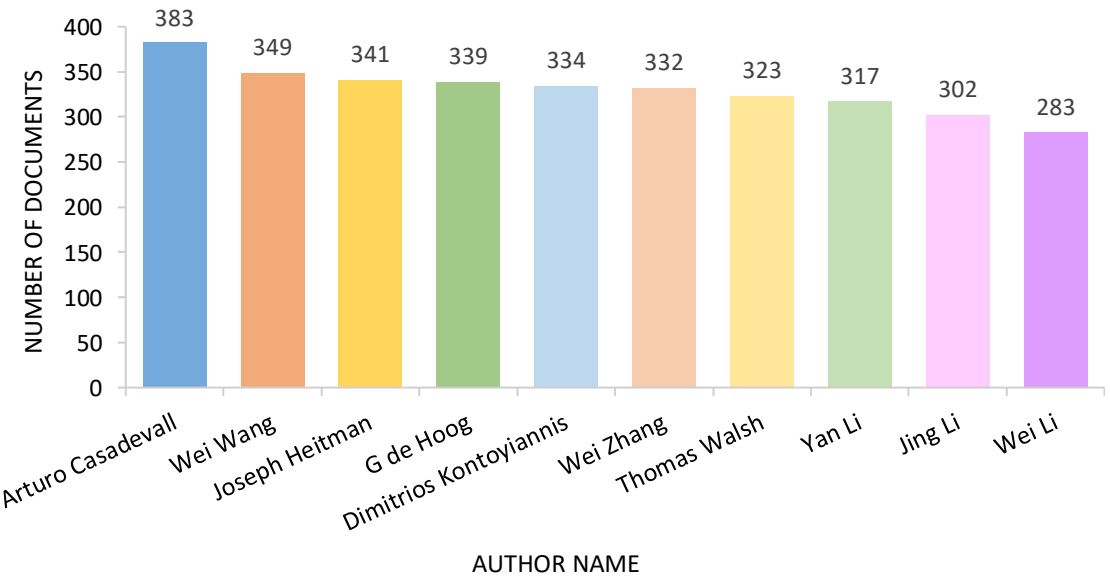

Figure 3.37. Top 10 authors between 1786 and 2020.

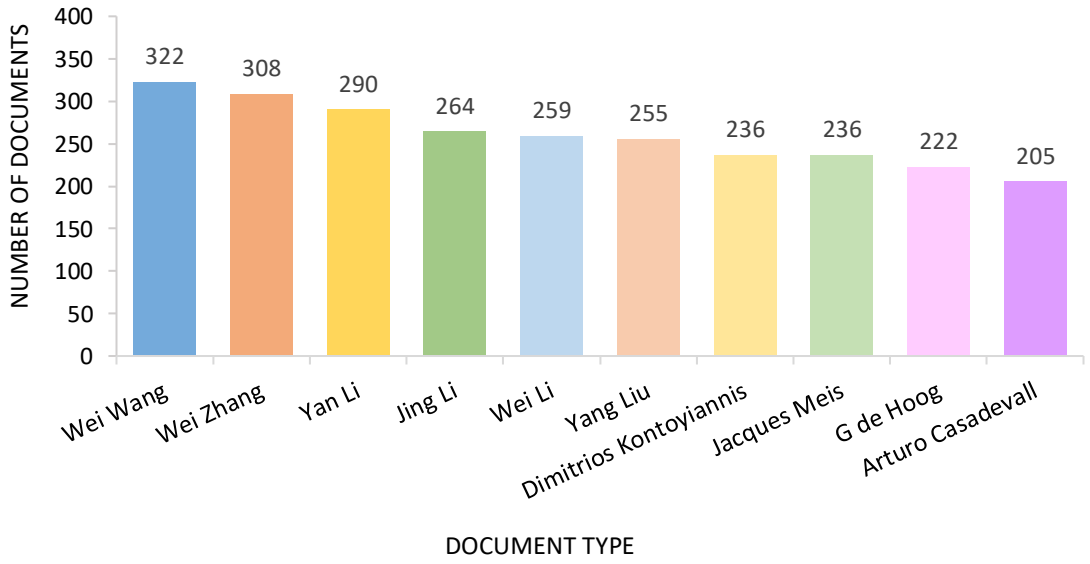

Figure 3.38. Top 10 authors between 2010 and 2020.

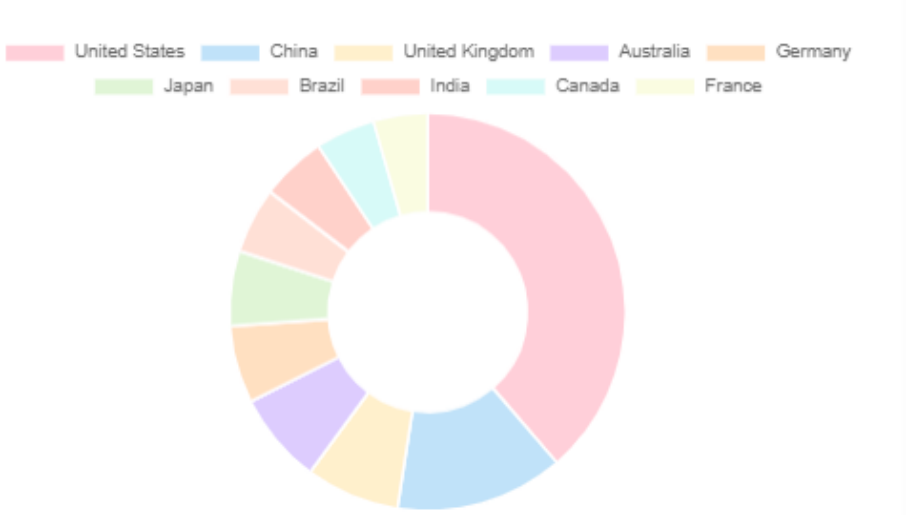

Figure 3.39. Top 10 countries between 1786 and 2020.

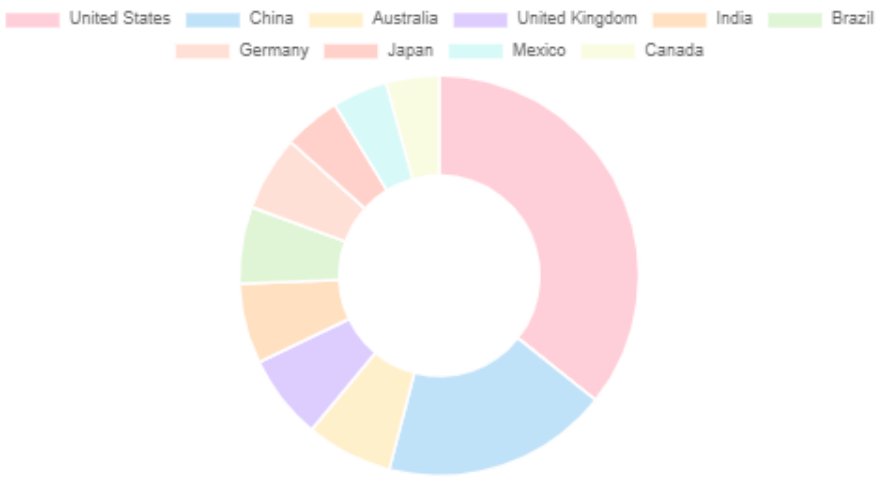

Figure 3.40. Top 10 countries between 2010 and 2020.

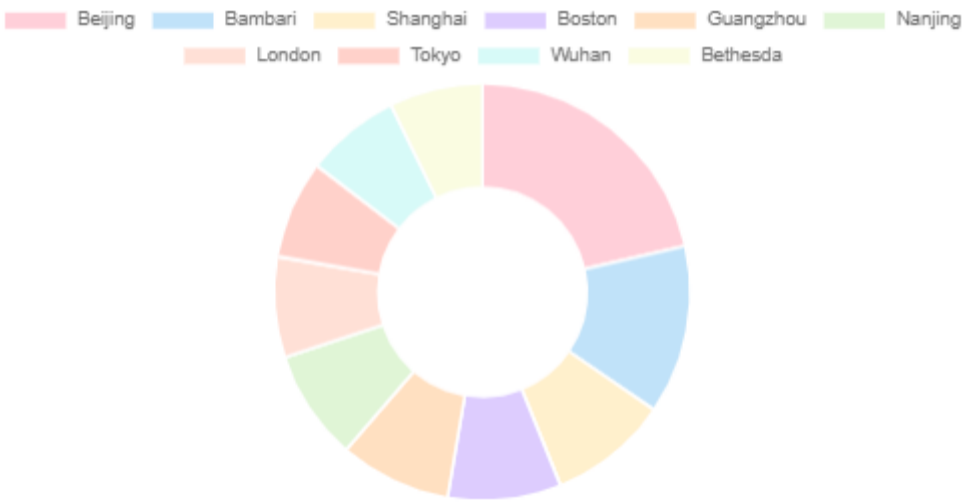

Figure 3.41. Top 10 cities between 1786 and 2020.

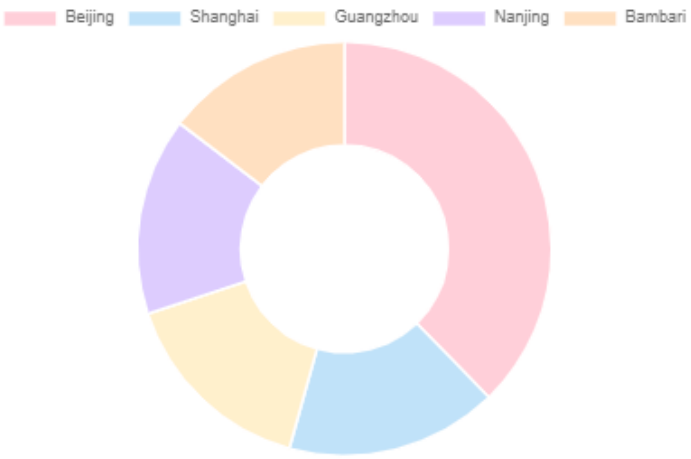

Figure 3.42. Top 5 cities between 2010 and 2020.

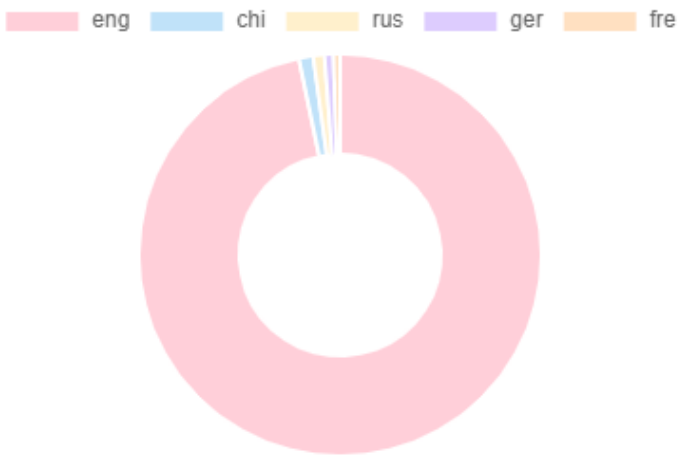

Figure 3.43. Top 5 languages of publications between 1786 and 2020.

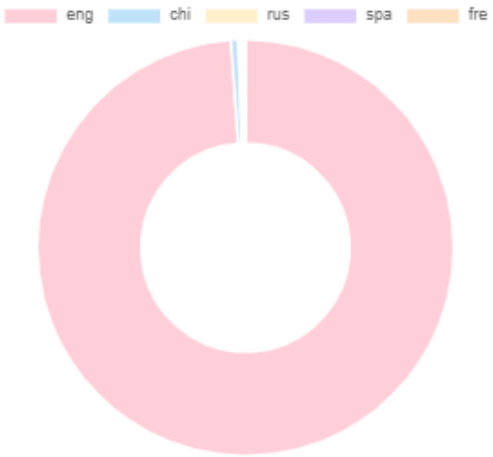

Figure 3.44. Top 5 languages of publications between 2010 and 2020.

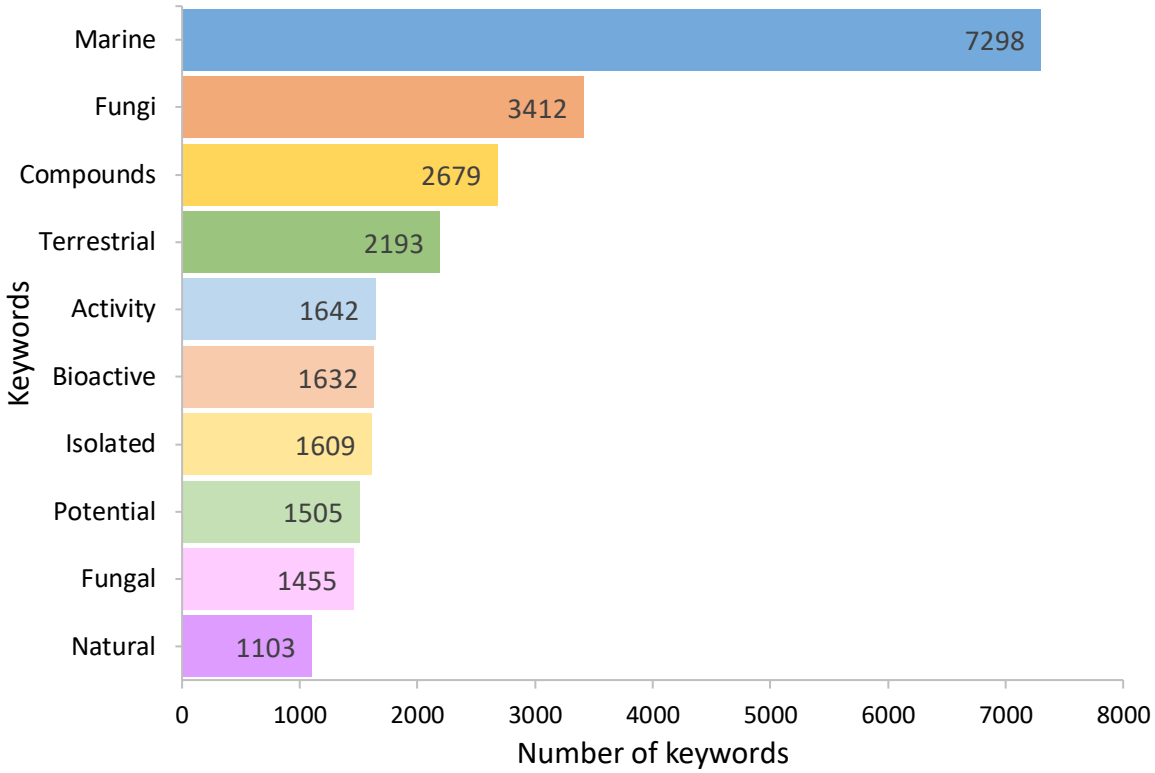

Figure 3.46. Top 10 words most frequent obtained using R programming language in Elsevier papers.

Figure 3.48. Annual count of the ten most frequent words analysed.

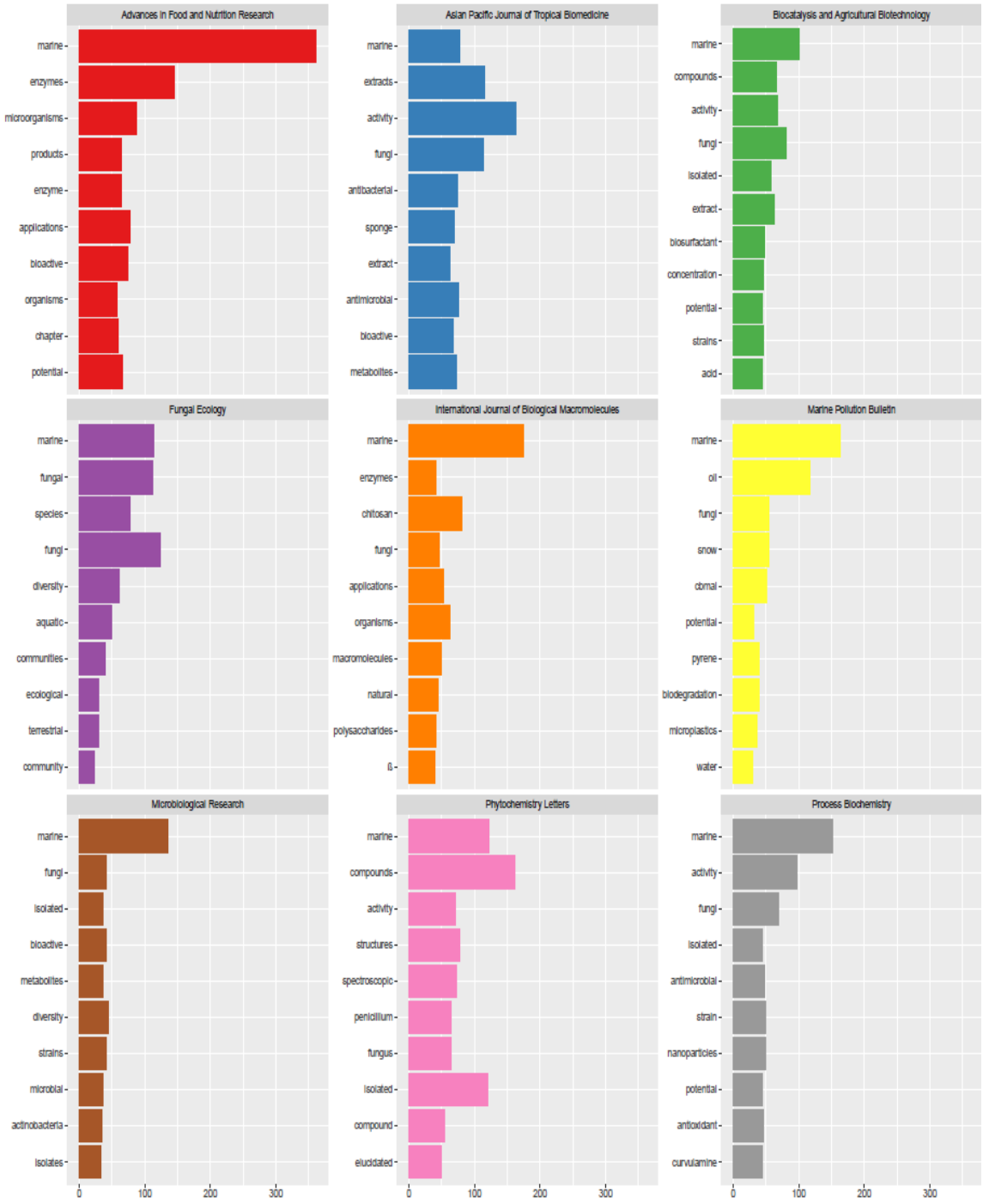

Figure 3.49. The top 10 journals obtained using R programming language.