A MULTIPHASIC PHYLOGENETIC AND GENOMIC CHARACTERIZATION OF THE FUNGAL ENDOPHYTE *STEMPHYLIUM* STRAIN PNW2016-02 AND A SEMI-AUTOMATED BIOINFORMATIC APPROACH TO INVESTIGATING THE STRAIN SECRETOME

by

Nathaniel Tobey

A Thesis

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Master of Science

Department of Biological Sciences

Hammond, Indiana

December 2021

THE PURDUE UNIVERSITY GRADUATE SCHOOL

STATEMENT OF COMMITTEE APPROVAL

Dr. Scott Bates, Chair

Department of Biological Sciences

Dr. Lindsay Gielda

Department of Biological Sciences

Dr. Dawit Gizachew

Department of Chemistry and Physics

Approved by:

Dr. Wei-Tsyi Evert Ting

ACKNOWLEDGMENTS

We thank the Purdue University Northwest College of Engineering and Sciences (CES) for their support through the Summer Research Grant Award and the PNW Interdisciplinary Award. In addition, we thank the PNW Biological Sciences Department for their backing of research within CES. In addition, we thank the Nils K. Nelson Endowment through CES for funding this research. We would also like to acknowledge various undergraduate and graduate researchers who have assisted in the project: Victoria Conn, Evan Curl, John Downing, Justin Golday, Rachel Kunnen, and Christopher Weliczko.

TABLE OF CONTENTS

LIST OF TABLES 6
LIST OF FIGURES 7
ABSTRACT 9
CHAPTER 1. INTRODUCTION 10
1.1 Endophytes and Microbial Interactions 10
1.2 Endophytic Fungal Isolate 12
1.3 Metabolites of Endophytic Fungi 19
1.4 Modification of Host-Endophyte Interactions Through Chemical Means 21
1.5 Bioinformatic Approach for Gaining Insights into Endophyte Metabolic Pathways 22
1.6 Research Questions 23
CHAPTER 2. METHODS 25
2.1 Bioinformatic Pipeline Overview 25
2.2 Extraction of Fungal DNA and Genomic Sequencing 25
2.3 Fungal Genome Assembly 25
2.4 Phylogenetic Analysis 26
2.5 HPLC-MS/MS of *Stemphylium* PNW2016-02 Secretome 27
2.6 Chemical Structure Cluster Analysis 28
2.7 Pathway Mapping 28
2.8 Bioinformatic Pipeline Code 29
CHAPTER 3. RESULTS 32
3.1 Genome Assembly 32
3.2 Phylogenetic and Comparative Genomic Analysis 35
3.3 Metabolite Chemical Cluster Analysis 39
3.4 Pathway Mapping 41
CHAPTER 4. DISCUSSION 46
4.1 Identity of *Stemphylium* PNW2016-02 46
4.2 Potential Activity of Evaluated Metabolites 48
4.3 Secretome Flavonoids and Related Compounds 50
4.4 Implications and Other Future Directions 58

REFERENCES 60
SUPPLEMENTAL DATA 84

LIST OF TABLES

Table 1. Gram-positive bacterial growth inhibition averages after 24 hours of direct contact with fungal endophyte isolates of *Stemphylium* (strain numbers are indicated), with standard deviation indicating variance between three individual trials. Agar plugs containing fungal isolates were placed directly onto bacterial lawns and incubated at 25°C for 24 hours..... 16

Table 2. Gram-negative bacterial growth inhibition averages after 24 hours of direct contact with fungal endophyte isolates of *Stemphylium* (strain numbers are indicated), with standard deviation indicating variance between three individual trials. Agar plugs containing fungal isolates were placed directly onto bacterial lawns and incubated at 25°C for 24 hours..... 17

LIST OF FIGURES

Figure 1. A relational map diagramming the dynamic nature of interactions between microbial endophytes and their plant hosts. 11

Figure 2. Stemphylium PNW2016-02, isolated from spinach and shown here growing on potato dextrose agar. A) view of culture from top of plate; B) view of culture from bottom of plate. ... 14

Figure 3. Stemphylium PNW2016-03, isolated from spinach and shown here growing on potato dextrose agar. A) view of culture from top of plate; B) view of culture from bottom of plate. ... 14

Figure 4. Filter sterilized fungal liquid growth media added to *E. coli* O157:H7 (strain EDL933), either untreated (solid lines) or heat stable after heating at 100°C for 15 minutes (dashed lines), inhibited bacterial growth over 8 hours as determined by optical density. Ethyl acetate fractionation determined antimicrobial activity is in both the organic and inorganic phases, suggesting multiple antimicrobial metabolites present in strain PNW2016-02 and PNW2016-03. 19

Figure 5. A simplified map of common metabolic pathways in fungi (after Thines et al., 2006).20

Figure 6. Process diagram of the semi-automated bioinformatic pipeline. Inputs are shown in yellow, graphic output with blue, and results with green. 31

Figure 7. Contig view ordered by size after processing through QUAST. The 1,126 contigs occupy the gray area in the top bar. These contigs are of a length shorter than 2503 bp compounding to a total length of 1,361,516 bp. 32

Figure 8. Pie chart showing the relative percentages of functional sequence predictions from the genome of PNW2016-02 in MG-RAST. 34

Figure 9. Pie chart showing the relative percentages of predicted sequence features of PNW2016-02 in MG-RAST. 34

Figure 10. Pie chart showing the relative percentages of functional gene product subsystems identified in the PNW2016-02 draft genome by MG-RAST. 35

Figure 11. Venn diagram of predicted gene features identified through Funannotate in three isolates of *Stemphylium*. *S. vesicarium* has GenBank accession number VHPU01002624 and *S. lycopersici* has accession number LC032102. 36

Figure 12. Molecular phylogeny of selected *Stemphylium* isolates based on partial 18S, ITS1, 5.8S, ITS2, and partial 28S sequence inferred using the UPGMA method in MEGA X. The optimal tree is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (10,000 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths gauged in units of number of base pair substitutions per site as determined by the Tamura-Nei method. The rate variation among sites was calculated with a gamma distribution of shape parameter equaling 1.00. This tree contains fifty-eight nucleotide sequences (10 *Pleospora*, eight scaffolds from PNW2016-02, 39 *Stemphylium*, 1 *Alternaria*). All ambiguous positions were

removed from the model for each sequence pair. There was a total of 505 positions in the final dataset. 37

Figure 13. Maximum-likelihood bootstrap consensus tree of ITS1 and ITS2 sequences of several similar species within *Stemphylium* and *Pleospora* with bootstrap values over 50% shown. (right) Graph showing the presence or absence of seven different genes within the genome 39

Figure 14. The similarity matrix generated from the metabolites of PNW2016-02 using the PubChem fingerprint syntax. The Tanimoto coefficient is used with a value of 1.00 (black) indicating identical structure and 0.00 (white) indicating dissimilar structure. 40

Figure 15. A dendrogram of the metabolites of PNW2016-02 with SMILES formulae found in the PubChem database. The branches are clustered using the Tanimoto coefficient and the PubChem chemical fingerprint syntax..... 41

Figure 16. A similarity matrix of the metabolites of PNW2016-02 using the pathway names obtained from MetaCyc. Higher similarity of the pathways in which two metabolites participate are denoted by darker tiles. 42

Figure 17. A dendrogram of the fungal metabolites of PNW2016-02 with identified metabolic pathways from the MetaCyc database. The branch lengths represent the relative dissimilarity in relevant pathways to the other branches determined by Jaccard index. 43

Figure 18. Tanglegram comparing the structural similarity and pathway similarity analyses. The lack of resolution in the pathway data is due to the limited pathway data in the MetaCyc database. 44

Figure 19. NMDS plot using the structural similarity of metabolites with identified pathways in the MetaCyc database. 45

Figure 20. Flavonoid chemical class groupings of secondary metabolites produced by the phenylpropanoid pathway (figure after Lacroix et al., 2018) 52

Figure 21. Structural similarity dendrogram (top) of the metabolome of PNW2016-02 with overlay of flavonoids (red), glycosides (blue), or both (violet). Glycosylated compounds were placed and named according to their aglycone but labelled as a glycoside. Expanded view of the subtree containing all of the flavonoids (bottom)..... 54

Figure 22. Section of metabolic pathways (map01100) in KEGG (Kanehisa et al., 2016) with pathways relevant to flavonoid synthesis highlighted in pink. Aglycones found in the metabolome of PNW2016-02 are labelled in black..... 56

ABSTRACT

As part of development of a human pathogen suppressing in plantae model system within *Spinacia oleracea*, a known pathogen transmitting produce plants, the secretome of the *Stemphylium* strain PNW2016-02, an endophytic fungal isolate of spinach plants, was studied. This strain was previously isolated in Purdue University Northwest laboratories in the Biological Sciences Department. As the secondary metabolite secretions of PNW2016-02 were shown to have antibiotic properties against a broad range of bacteria, we sought to improve our understanding of these properties and other characteristics of PNW2016-02 by sequencing and annotating the genome of this strain. Chemical compound characterization was achieved using HPLC-MS, which provided the ability to identify and quantify chemical compounds contained within the PNW2016-02 secretome. Through multi-gene phylogenetic analysis and comparative genomics, we found that PNW2016-02 clusters with a clade representing *Stemphylium vesicarium*; however, genome annotation also uncovered several genes unique to PNW2016-02. To assist in more fully understanding the secondary metabolites PNW2016-02, a semi-automated bioinformatic pipeline was developed in the R statistical environment in order to reconstruct and characterize the metabolic pathways, especially those pertinent to antibiotic production. Analysis using the bioinformatic pipeline revealed the presence of a number of antimicrobial metabolites, such as flavonoids, and suggests these compounds are produced and transformed in pathways, such as the phenylpropanoid synthesis pathway. These findings also suggest an important vital role of the shikimate pathway for antimicrobial metabolite production within PNW2016-02. Overall, the findings presented here have implications for understanding the antimicrobial strategies employed by *Stemphylium* PNW2016-02 and its potential for use in the above-mentioned model system.

CHAPTER 1. INTRODUCTION

1.1 Endophytes and Microbial Interactions

Microorganisms are ubiquitous and their presence impacts ecosystems in a variety of ways. For example, they can decompose dead plant tissue to release recalcitrant nutrients (Yarwood, 2018), serve as a food source to larger organisms (Niccolai et al., 2019), and harm other organisms as pathogens (West et al., 2012). Microbes do not exist independently in the environments they live in as they are typically engaged in interactions with other organisms, and sometimes the nature of these interactions are dynamic. Microbes may also engage in symbiotic relationships that include mutualism, commensalism, and parasitism, depending upon which microbe-host pairing is observed (Wani et al., 2015). The nature of these relationships is not static and may change depending on factors like microbial density, environmental conditions, or genetic changes (Kholkani et al., 2017; Keilhofer, 2018). For example, *Escherichia coli* is primarily found living symbiotically within the intestinal tract of around 90% of humans (Tenaillion et al., 2010) where this bacterium contributes to host health by providing vitamins that humans are unable to synthesize (Gao et al., 2014); however, some strains of this species, such as Enterohemorrhagic *E. coli* (EHEC) H7:O157, can act as a lethal pathogen (Anderson, Whitlock, & Harwood, 2005; Nakkarach et al., 2021). Every year EHEC is responsible for numerous hospitalizations and deaths, and case reports confirm that this pathogen is often transmitted through produce, especially leafy greens such as lettuce and spinach (Marder et al., 2014; Sharapov et al., 2016; Centers for Disease Control and Prevention, 2021). EHEC is known to exist on the leaf surfaces, thereby infecting humans that consume the produce, and recent work suggests that EHEC is capable moving inside the plant where it may live is an endophyte (*endo* - within, *phyte* - plant), thus increasing the risk of infection as the pathogen cannot simply be washed off the surface of the leaves (Dong et al., 2003; Shaw et al., 2008).

Microbes such as bacteria and fungi often battle one another for dominance within ecosystems; this is also true of endophytes that are competing for resources within the interior tissue of the plant in a specialized niche space known as the endophytome (Gómez-Lama Cabanás et al., 2021). Within the endophytome bacterial and fungal endophytes, therefore, compete with

one another as individual species for survival inside a plant (Akbaba & Ozaktan, 2018). As with the microbial interactions cited above, endophyte interactions with the host are likewise dynamic (Figure 1), and extending from these interactions, endophytes can also ultimately be involved in modulating health outcomes for the plant host. For example, *Diplodia sapinea* is a fungal pathogen of pine that causes shoot blight, and damage to the tree host can be indirectly minimized when healthy endophyte communities prevent pathogen colonization through competitive exclusion (Oliva et al., 2020). In another example, *Xanthomonas campestris* pv. *campestris* is a bacterial phytopathogen that causes black rot of crucifers (Vicente & Holub, 2012). Endophytic bacterial isolates from healthy canola tissue were able to reduce the bacterial load of *X. campestris* in this crop plant (Romero et al., 2018). In a comparable manner, bacterial endophytes isolated from healthy plant tissue can be used to directly prevent angular leaf spot, a disease caused by the bacterial pathogen *Pseudomonas syringae* pv. *lachrymans*, in cucumber through antagonistic interaction in nutrient competition (Akbaba & Ozaktan, 2018). Beyond pathogen suppression, endophytes can also influence other outcomes for the host, such as the promotion of plant growth (Romero et al., 2018; Poveda et al., 2021) or stimulation of plant defenses through the production of secondary metabolites (Romero et al., 2018; Fadiji & Babalola, 2020).

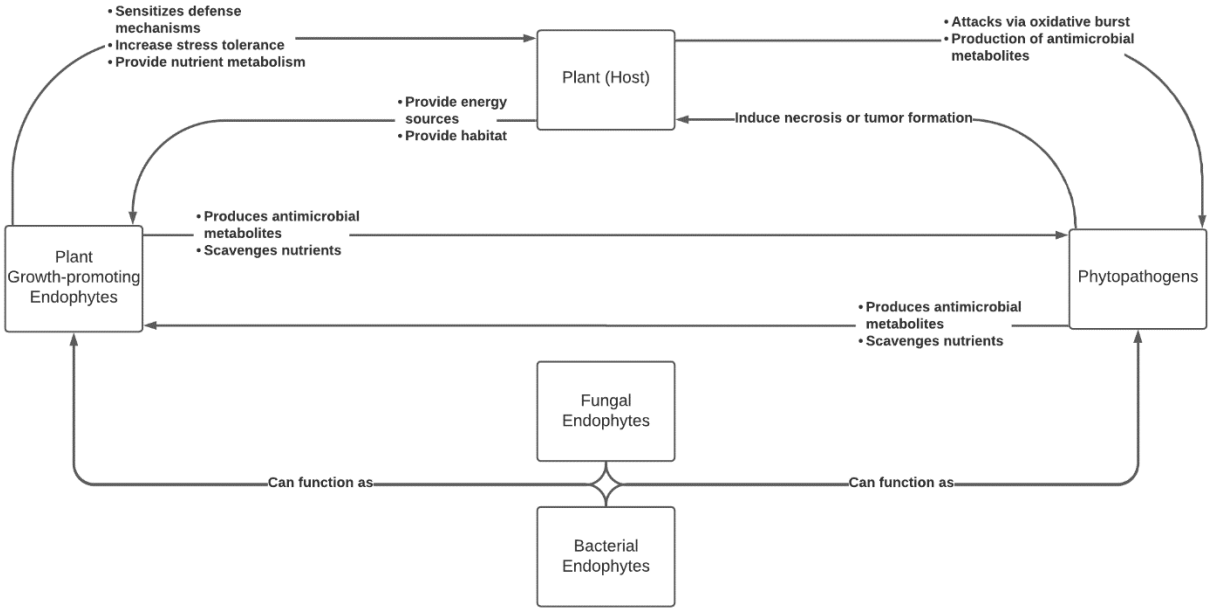


Figure 1. A relational map diagramming the dynamic nature of interactions between microbial endophytes and their plant hosts.

Considering these interactions, it is important to note that, like other microbes, endophytes are known to use chemical strategies to gain advantages over their competitors. Some of the compounds produced have the potential for use as biocontrol agents as they can serve to improve health outcomes for the host. For example, in one of the studies cited above the suppression of black rot caused by *Xanthomonas campestris* pv. *campestris* was thought to be achieved through the direct production of antimicrobial compounds by one of the 'healthy' bacterial endophytes (Romero et al., 2018). In another case, a fungal endophyte, *Gliocladium* sp., was isolated from tissue of the South American tree, *Eucryphia cordifolia,* and this fungus secreted a series of volatile organic compounds that were found to selectively inhibit the growth of phytopathogens of this tree (Stinson et al., 2003). Similarly, volatile secretions of the fungal endophyte *Daldinia* cf. *concentrica* isolated from olive trees, *Olea europaea,* inhibited the growth of *Aspergillus niger* infection in peanuts (Liarzi et al., 2016). The chemical tactics employed by each microbe in competitive interactions have coevolved with other endophytic microorganisms and the host itself, and this coevolution has led to the continuous generation of new metabolites with antimicrobial properties (Yan et al., 2018). Often, a new metabolite is a modification of existing structure, such as those found in short peptide-derived antimicrobials, which leads to the evolution of novel metabolic pathways (Ibrahim et al., 2020). Fungal endophytes, in particular, are known to be highly diverse and ubiquitous within plants (Arnold et al. 2002), thus they represent a considerable storehouse of unique bioactive products with immense potential (Strobel, 2003).

1.2 Endophytic Fungal Isolate

In previous work, an undergraduate student at Purdue University Northwest isolated endophytic fungal strains of *Stemphylium* (PNW2016-02 and PNW2016-03) from spinach (*Spinacia oleracea*) that had many interesting properties; these are described below. Species in the genus *Stemphylium* are known to cause plant disease on many major crops including onions, lentils, and spinach (Razzak et al., 2018; Spawton, McGrath, & du Toit, 2020; Wang et al., 2021), and species such as *S. beticola* are known to cause spinach leaf spot disease that appears as black lesions dotting the surface of leaves (Spawton, McGrath, & du Toit, 2020). *Stemphylium* species are also known to produce a number of interesting chemical compounds, some with antimicrobial activities and/or iron interactions (Manulis et al., 1984; Andersen et al., 1995; Debbab et al., 2009;

Medina et al., 2021). While our isolated strains had an affinity to *S. vesicarium* (Figure 1) and this species along with *S. beticola* are predominant leaf spot causing pathogens in spinach production areas in the United States (Liu et al. 2020), our strains were, however, isolated from non-symptomatic tissue. A well-known nineteenth century German botanist, Karl Friedrich Wallroth, established the genus *Stemphylium* in the nineteenth century survey of plants and fungi, Flora Cryptogamica Germaniae (Wallroth, Bluff & Fingerhut, 1833) with *S. botryosum* selected as the type species for the genus. Since the publication of Wallroth et al. (1833), many species have been proposed in *Stemphylium*; however, the need to reevaluate species circumscriptions in the genus using DNA-based methods was recognized following an outbreak of sugar beet leaf spot disease in the Netherlands (Woudenberg et al., 2017). Woudenberg et al. (2017) suggested combining members of a clade that included *S. alfalfae, S. cremanthodii*, *S. herbarum, S. mali*, *S. pomorum, S. sedicola*, and *S. tomatonis,* in synonymy under *S. vesicarium* on the basis of genetic and morphological evidence. While our isolates appeared to have a morphological affinity with the aforementioned clade, there was some uncertainty as to the actual species that our isolates represented. This thesis, however, will establish the true identity of our isolate.

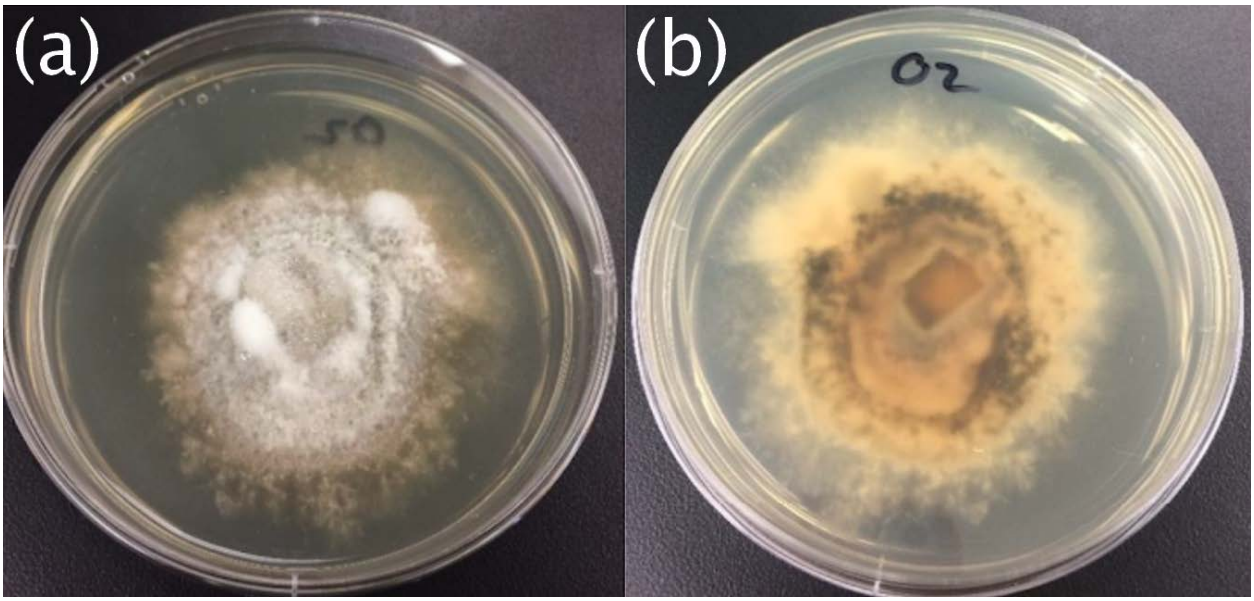


Figure 2. Stemphylium PNW2016-02, isolated from spinach and shown here growing on potato dextrose agar. A) view of culture from top of plate; B) view of culture from bottom of plate.

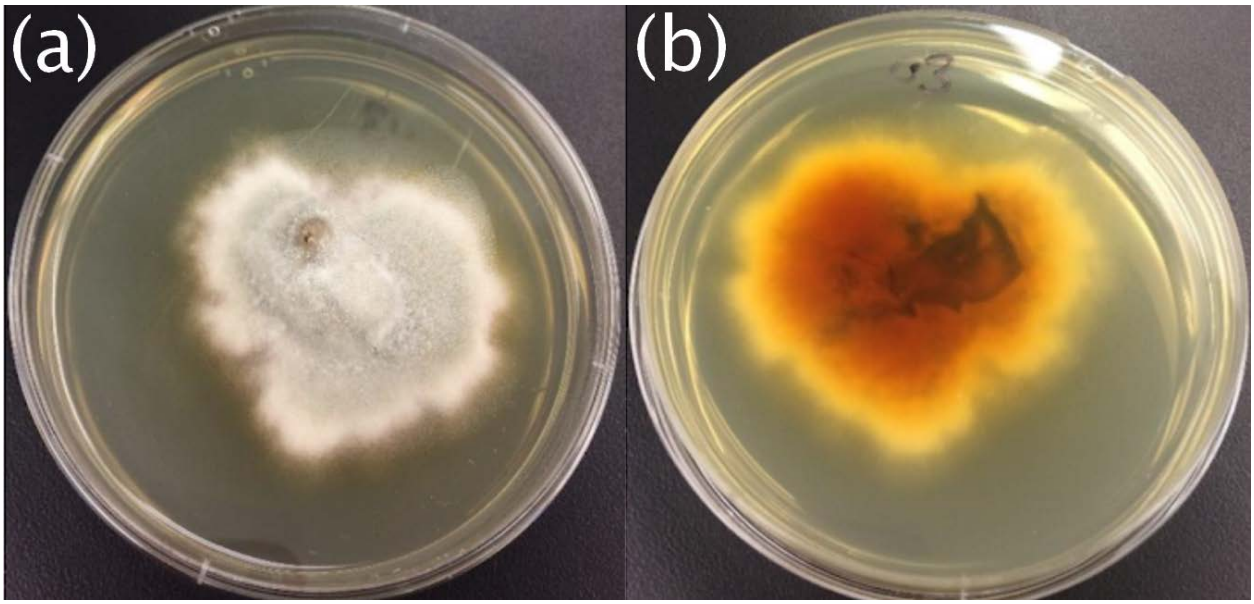


Figure 3. Stemphylium PNW2016-03, isolated from spinach and shown here growing on potato dextrose agar. A) view of culture from top of plate; B) view of culture from bottom of plate.

Following isolation, a battery of growth inhibition tests using PNW2016-02 and PNW2016-03 against numerous Gram-negative and Gram-positive bacteria revealed broad, yet variable, antimicrobial activity *in vitro* (Table 1-2). Most interestingly, neither of our *Stemphylium* strains had inhibitory effects on *Serratia marcescens*, suggesting this Gram-negative bacterium has genetic capabilities that allow it to counteract the antimicrobial activities of metabolites produced by PNW2016-02 and PNW2016-03 during growth. The mechanism of resistance by *S. marcescens* against exudates from our *Stemphylium* isolates were then studied extensively by a master's student in the Gielda lab (Sopovski, 2019). Additionally, our strains inhibited the growth of teaching strains of *E. coli*, which led to questions about whether the exudates from our *Stemphylium* strains could be effective against pathogenic strains of *E. coli*.

Table 1. Gram-positive bacterial growth inhibition averages after 24 hours of direct contact with fungal endophyte isolates of *Stemphylium* (strain numbers are indicated), with standard deviation indicating variance between three individual trials. Agar plugs containing fungal isolates were placed directly onto bacterial lawns and incubated at 25°C for 24 hours.

Gram-Positive Bacterium	**PNW2016-02**	**PNW2016-03**
Bacillus cereus	1.0 cm (+/- 0.4)	1.0 cm (+/- 0.5)
Bacillus megaterium	0.4 cm (+/- 0.3)	0.7 cm (+/- 0.2)
Enterococcus faecalis	1.0 cm (+/- 0.1)	0.8 cm (+/- 0.1)
Lactococcus lactis	0.5 cm (+/- 0.3)	0.7 cm (+/- 0.2)
Lactococcus plantarum	0.7 cm (+/- 0.2)	0.0 cm
Micrococcus luteus	1.1 cm (+/- 0.1)	0.9 cm (+/- 0.4)
Micrococcus roseus	1.1 cm (+/- 0.1)	1.0 cm (+/- 0.1)
Staphylococcus aureus	1.2 cm (+/- 0.3)	1.2 cm (+/- 0.5)
Staphylococcus epidermidis	0.9 cm (+/- 0.1)	0.9 cm (+/- 0.4)
Staphylococcus saprophyticus	1.1 cm (+/- 0.4)	0.9 cm (+/- 0.5)

Table 2. Gram-negative bacterial growth inhibition averages after 24 hours of direct contact with fungal endophyte isolates of *Stemphylium* (strain numbers are indicated), with standard deviation indicating variance between three individual trials. Agar plugs containing fungal isolates were placed directly onto bacterial lawns and incubated at 25°C for 24 hours.

Gram-Negative Bacterium	**PNW2016-02**	**PWN2016-03**
Proteus vulgaris	0.3 cm (+/- 0.1)	0.1 cm (+/- 0.2)
Salmonella typhi	0.6 cm (+/- 0.05)	0.5 cm (+/- 0.2)
Escherichia coli K12	0.4 cm (+/- 0.1)	0.4 cm (+/- 0.1)
Enterobacter aerogenes	1.0 cm (+/- 0.0)	0.4 cm (+/- 0.1)
Aeromonas hydrophila	0.7 cm (+/- 0.2)	0.7 cm (+/- 0.1)
Serratia liquefaciens	1. 1 cm (+/- 0.3)	1.1 cm (+/- 0.5)
Serratia marcescens	0.0 cm	0.0 cm

On the basis of this fundamental question regarding efficacy of chemical compounds produced by our strains against pathogenic *E. coli*, a series of growth inhibition assays were carried out on EHEC O157:H7 strain EDL933 using metabolite exudates of the PNW2016-02 and PNW2016-03 as well as those of other fungal endophytes isolated from spinach. Of the fungi tested, our endophytic *Stemphylium* strains showed considerable promise (Figure 2-3) in suppressing the growth of EHEC. Considering that PNW2016-02 and PNW2016-03 were endophytic fungal strains initially isolated from spinach plants, that these strains can suppress the growth of pathogenic *E. coli* known to be transmitted via produce plants, like spinach, and that such transmission events have been the cause of EHEC outbreaks in the past, we recognized that our strains showed potential to be used in a model system whereby suppression of a human pathogen (viz. EHEC) could be achieved *in plantae* through inoculating our endophytic *Stemphylium* strains in pathogen transmitting produce plants (viz. spinach). Thus, natural human pathogen suppressing spinach plants could be produced; therefore, another master's student in the Bates lab developed a system for inoculating PNW2016-02 and PNW2016-03 back into spinach plants (Golday, 2019). Golday (2019) was able to successfully inoculate both endophytic strains into spinach and developed a molecular assay for the detecting these strains in the plant tissue, and it was also found that spinach plants inoculated with PNW2016-02 had higher rates of inoculation success and germinating plant survival. Additionally, PNW2016-02 inoculated spinach plants retained the fungal endophyte for the full two weeks of the study post-inoculation and these plants also showed a significantly higher growth rate when compared to the controls (Golday 2019). Recognizing the potential of *Stemphylium* PNW2016-02 we set out to further characterize the genome and secretome of this strain, and to better understand the true identity of endophytic species.

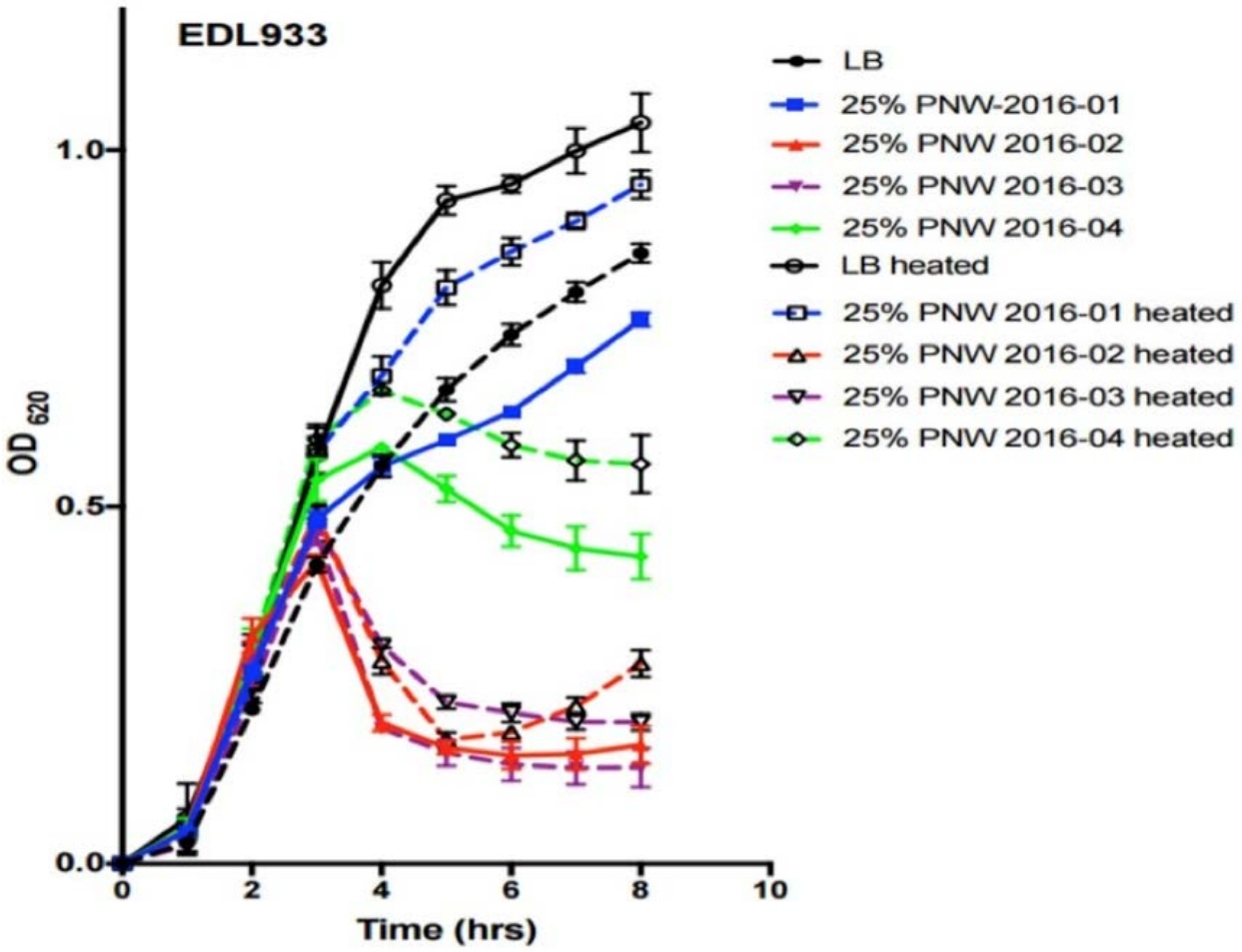


Figure 4. Filter sterilized fungal liquid growth media added to *E. coli* O157:H7 (strain EDL933), either untreated (solid lines) or heat stable after heating at 100°C for 15 minutes (dashed lines), inhibited bacterial growth over 8 hours as determined by optical density. Ethyl acetate fractionation determined antimicrobial activity is in both the organic and inorganic phases, suggesting multiple antimicrobial metabolites present in strain PNW2016-02 and PNW2016-03.

1.3 Metabolites of Endophytic Fungi

As heterotrophic fungal endophytes receive energy in the form of carbonous molecules (i.e., sugars) from the plant host, these compounds can be processed by the endophytes into a myriad of chemical metabolites (Behie et al., 2017; Dudeja et al., 2021). Many factors play a role in the quantity and structure of secondary metabolites produced by any particular microbial species, and one approach to the discovery of metabolites is the “one strain many compounds” strategy which varies the conditions at which microorganisms are cultured with the aim of increasing the diversity of the compounds produced. This strategy is based on evidence that a variable expression

of genes controlling metabolite production can be achieved from a given species dependent upon the condition of its environment (Pan et al., 2019). Currently, there are several well-established pathways that can be maintained at an elevated level of expression through which most metabolites are produced or modified (Sarsaiya, Shi & Chen, 2019; Ogbe, Finnie, & Van Staden, 2020). Originating from carbon sources provided by the host plants, fungal endophytes can produce enzymes that allow these carbon sources to be modified by being funneled into one of a few metabolite producing pathways as seen in Figure 5, where glucose represents the source of carbon (Sarsaiya, Shi & Chen, 2019).

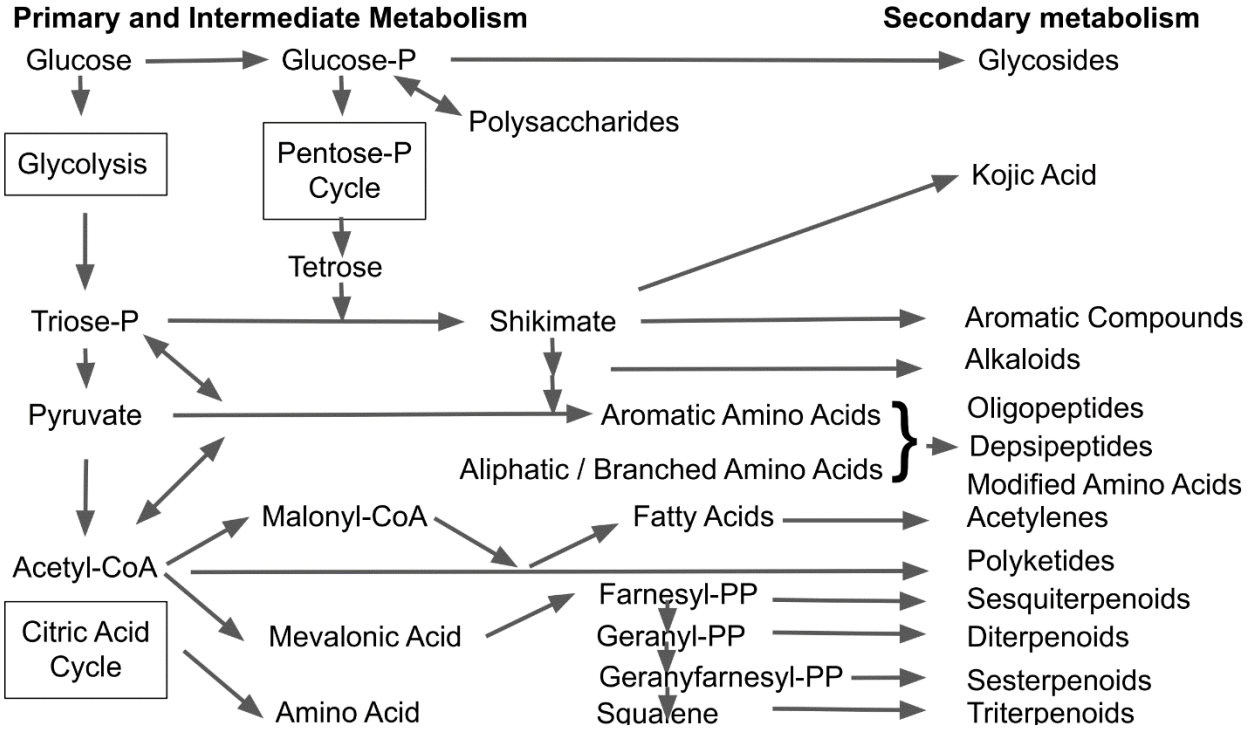


Figure 5. A simplified map of common metabolic pathways in fungi (after Thines et al., 2006).

As natural products, secondary metabolites represent an underexplored pool of potential antimicrobials and/or antimicrobial strategies, and endophytic fungi, including *Stemphylium* species, show potential to contribute to this area (Manulis et al., 1984; Andersen et al., 1995; Strobel, 2003, Barbosa et al., 2020; Akther, Ranjani, & Hemalatha, 2021). Recent research on the metabolome of *Stemphylium*, suggests that species in this genus produce a considerable number of

novel compounds that demonstrate antimicrobial activity (Debbab et al., 2009; Medina et al., 2021). Those compounds first identified in and named after the genus include stemphone, stemphol, stemphylin, stemphyperylenol, and stemphyltoxins (Arnone et al., 1986; Andersen et al., 1995). Other bioactive compounds have been found in *Stemphylium* isolates that were first identified in other species in the same taxonomic family (*Pleosporaceae*), as is the case with alterporriols and altersolanols that were first isolated from species within the genus *Alternaria* (Medina et al., 2021). Outside of the order *Pleosporales*, but within class *Dothideomycetes* that contains these taxa, a fungiculous (a fungus that grows parasitically on another fungus) isolate similar to *Cladosporium* produced a bioactive anthraquinone that was named altersolanol J after its highly similar structure to that of the antibacterial altersolanol A (Höller, Gloer, & Wicklow, 2002). Most of these compounds have aromatic carbon rings and may be polycyclic (e.g., anthraquinones) similar to flavonoids, another class of metabolites.

1.4 Modification of Host-Endophyte Interactions Through Chemical Means

Numerous interactions of endophytic fungi and their plant hosts occur regularly at a chemical level, and both participants in these exchanges can modify the activity of the other (Minerdi, Maggini, & Fani, 2021). Hosts plants can generally protect themselves against the internal growth of microbes, which might represent pathogenic infections, by producing phytoanticipins, which are antimicrobial compounds generated in 'healthy' plant tissue before infection events occur (Westrick, Smith, & Kabbage, 2021). Plants can also produce defensive compounds in response to infection events, and these are known as phytoalexins (Santra & Banerjee, 2020). The localization, availability and biological activity of plant defense compounds can also be chemically regulated through mechanisms such as glycosylation and glycoside hydrolysis (Le Roy et al., 2016). Conversely, endophytes can produce an array of compounds that modify plant growth or maturation, an example being the production of auxins also known as plant growth hormones (Suebrasri et al., 2020). Fungal endophytes employ a number of strategies for countering the effect of plant-produced defense compounds, including intracellular or secreted fungal enzymes that disarm plant-defense chemicals, and efflux pumps that can move these compounds out of fungal cell (Chen et al., 2019; Westrick, Smith, & Kabbage, 2021). An example in the plant *Arabidopsis thaliana* is the fungus *Sclerotinia sclerotiorum* that enzymatically cleaves

glycones from aglycones via fungal glycosyltransferase activity and subsequently detoxified the aglycones via catabolism (Chen et al., 2019).

1.5 Bioinformatic Approach for Gaining Insights into Endophyte Metabolic Pathways

Reconstructing the metabolic pathways of a microorganism has continued to prove difficult due its multi-omic nature (Wang, Dash, Ng, & Maranas, 2017). Standardized data collection practices and databases have evolved to amplify our understanding of the ecology of microorganisms on a global scale. As such, the scientific community leverages this widely available information to stimulate computational research (Keller et al., 2020). Over the last couple of decades, researchers have been compiling an ever-increasing number of databases that can be utilized in collecting new data to facilitate experimental research in areas like genomics, transcriptomics, and metabolomics. All-in-one multi-omics platforms like Reactome (Jassal et al., 2020) and MetaboAnalyst (Pang et al., 2021) cross-reference many of these databases to give a well-rounded interpretation of the metabolic pathways of a target species. In several cases, the analysis is limited to specific organisms like humans, *Homo sapiens*, or fungal yeast, *Saccharomyces cerevisiae*, limiting how applicable the analyses would be to other organisms such as the *Stemphylium* strain that is the focus of this study (Muthubharathi, Gowripriya, & Balamurugan, 2021). A benefit of having curated datasets with multiple sources of experimental data is the ability to maintain consistency across current and future experiments; however, this comes with the additional requirement of maintaining these databases. A way around this dilemma is by also having non-curated datasets available, such as EcoCyc (Karp et al., 2021).

The multiphasic characterization of *Stemphylium* strain PNW2016-02 prompts the use of tools that can harness the information contained within multiple databases and are able to dynamically respond to changes or updates within said databases. Use of these databases can be optimized and focused in this characterization effort through a dedicated pipeline that can access, parse, manipulate, and generate reports on data gathered from the source databases in a scalable manner (Wratten et al., 2021). Creating semi-automated bioinformatic pipelines can help future studies standardize their results and more rapidly incorporate previous results of other experiments (Wratten et al., 2021). Many of these pipelines rely on application programming interfaces (APIs) to access data stored in large repositories (Woody et al. 2020). A REST or RESTful API is an API

that abides by the conventions of representational state transfer (REST) architecture which are used to maintain consistency in data interpretability across the world wide web. With REST, users request information from a database through an API, and the database can return a representation of the requested information to the user (Richardson & Amundsen, 2013). Bioinformatics can capitalize on the availability of RESTful APIs in order to incorporate data from multiple sources into the same study using computers in a repeatable and reproducible protocol.

Three primary databases were targeted in our pipeline development that can be generally useful in identifying key components to the antimicrobial activity observed in PNW2016-02, and these are the National Institutes of Health's public chemistry database (PubChem), Stanford Research Institute (SRI) International's Metabolic Pathway Database (MetaCyc), and European Bioinformatics Institute's Chemical Entities of Biological Interest (ChEBI). PubChem, MetaCyc, and ChEBI were dynamically integrated into the pipeline we developed, with Kyoto University's Kyoto Encyclopedia of Genes and Genomes (KEGG) used in the visualization of metabolic pathways recovered in the pipeline. PubChem stores a gamut of published or calculated data on biologically relevant compounds (Kim et al., 2021). MetaCyc, a counterpart to the BioCyc database, handles information, such as ontology, metabolic pathway, and source organism, on identified metabolites (Caspi et al., 2020). ChEBI manages much of the same information as MetaCyc, but it also includes functional annotations for each of the compounds included (Hastings et al., 2015). Outside of the pipeline, European Bioinformatics Institute's protein families (PFAM) and Gene Ontology and Gene Ontology Annotations (QuickGO) databases were implemented to parse the annotated genome. PFAM contains data on proteins, protein families, and domains (Mistry et al., 2020). QuickGO is a gene ontology database which helps annotate the functions of PFAM terms (Binns et al., 2009). All databases used in this study were accessed via RESTful API services provided by the above-mentioned institutions for the databases cited.

1.6 Research Questions

Recognizing the potential for our fungal endophyte strain PNW2016-02 to be used in a model system of natural human pathogen suppressing spinach plants, in this study we carry out a multiphasic characterization of this *Stemphylium* strain using phylogenetic, genomic, and secretome data and bioinformatic approaches to ask the following questions: Is *Stemphylium*

PNW2016-002 conspecific with other known taxa within the genus, such as *S. vesicarium*, or does the strain represent a novel species? Do genomic comparisons between our strain and other known strains of *Stemphylium* reveal distinct genomic features for PNW2016-02? Can bioinformatics be used to gain insights into the PNW2016-02 secretome? Do metabolite clusters recovered in our bioinformatic analyses relate to unique functional features (e.g., antibiotic potential) or metabolic activity (e.g., metabolite modification)? Also, do the compounds in these clusters have shared or distinct metabolite biosynthetic pathways that can be discovered and/or reconstructed in our bioinformatic pipeline?

CHAPTER 2. METHODS

2.1 Bioinformatic Pipeline Overview

A bioinformatic pipeline (Figure 6) was developed taking two inputs of the raw genetic data in FASTA format as well as a table of the metabolites from an organism, for example a list of chemical compounds from the secretome in this study, and the outputs being the species identity and a reconstruction of metabolic pathways, those related to antimicrobial activities for this study, recognized within the particular source organism, the *Stemphylium* strain PNW2016-02 in this study. More information on the bioinformatic pipeline development is provided below.

2.2 Extraction of Fungal DNA and Genomic Sequencing

DNA was extracted from mycelial tissue of the fungal endophyte *Stemphylium* strain PNW2016-02 from cultures grown on potato dextrose agar plates using the DNeasy PowerSoil DNA Extraction Kit (Qiagen, Venlo, Netherlands). The manufacturer's protocol was followed with the addition of a preliminary 10-minute heating step at 75°C prior to beaded tube vortexing, which assisted in degradation of the mycelial tissue. Extracted DNA from PNW-2016-02 was then sent to the Purdue Genomics Core Facility for genomic sequencing (2X150 base S-4 run) on an Illumina NOVAseq after adaptive focused acoustic shearing.

2.3 Fungal Genome Assembly

Raw sequencing reads were processed, evaluated for quality, and contigs were assembled in the Galaxy open-source web-portal (Grüning et al., 2017) using Cutadapt (Martin 2011), Unicycler (Wick et al. 2016), and QUAST (Gurevich et al. 2013), respectively. Gene prediction and annotation was carried out using the Funannotate pipeline (Palmer 2016).

After obtaining the raw contigs, the Funannotate pipeline was used to perform the final assembly and annotation of the draft genome (Palmer & Stajich, 2020). The assembly begins with renaming, filtering, and sorting the contigs by size using the *clean* and *sort* functions which was then fed into the *mask* function. Only contigs of a length >150 bp were used. The contigs were run

through RepeatMasker using the *mask* function which returns contigs truncated based on species-dependent patterns. No species was specified for our data, so RepeatMasker used its default repeat library to trim the contigs. The annotation file was returned in FASTA format which was loaded into the prediction/annotation function.

The functional annotation of the draft genome was done using the *predict* function in Funannotate which uses several programs to parse, analyze, and organize the metadata. The function first primed GeneMark-ES (Ter-Hovhannisyan et al., 2008) using the draft genome. An augmented BUSCO2 (Manni et al. 2021) program identified orthologs within the sequences which was then passed as a training set to Augustus (Nachtweide & Stanke, 2019). Any high-quality predictions made by Augustus were recorded. The BUSCO2 data was also fed into *snap* (Korf, 2004) and GlimmerHMM (Majoros et al., 2004) to generate prediction models. The data from Augustus, *snap,* GlimmerHMM, and GeneMark was passed into the evidence modeler which filtered gene predictions by length as well as identified transposable elements. The predicted sequences were passed through tRNAscan-SE (Chan & Lowe, 2019) to predict transcripts.

The predictions as well as the annotation data were converted into an NCBI annotation table and then translated to the GenBank format using tbl2asn (Ouellette & Boguski, 1997). The results were separated into several files of which this study will focus only on the high-quality predictions (>90% fidelity) which serves as the default output in both gff3 and gbk formats. The genomes for a strain of *S. vesicarium* and *S. lycopersici* accessed through GenBank were also run through the Funannotate pipeline. The contigs were then also loaded into the MG-RAST metagenomic analysis server (Meyer et al., 2008) and run through the shotgun metagenome analysis tool for further functional annotation and characterization.

2.4 Phylogenetic Analysis

As described in the previous section, the raw genomic sequence data was extracted, assembled, and annotated, which made it possible to also isolate phylogenetic marker genes, such as the internal transcribed spacer (ITS) of the nrRNA gene (see below) for use in the phylogenetic analysis that included other strains of *S. vesicarium*; the anamorph, *Pleospora herbarum*; and other species within the genus, *S. lycopersici, S. mali, S. gracilariae, S. symphyti, S. eturmiuna, S. alfalfae,* and *S. lancipes*; as well as an outgroup species, *Alternaria alternata*, within *Pleosporales*.

All marker genes from the above-mentioned species were recovered from GenBank (Supplemental Table 1) and the alignment was performed in MEGA X (Kumar et al., 2018) using the MUSCLE algorithm (Edgar, 2004). A phylogenetic tree was generated using the UPGMA method with 10,000 bootstrap replicates generated. The Tamura-Nei nucleotide substitution model was used with gamma-distributed substitution rate with a shape parameter of 1.00.

Phylogenetic marker genes of interest (e.g., ITS1 and ITS2) were isolated within the PNW2016-02 genome using ITSx (Bengtsson-Palme et al., 2013). The two other *Stemphylium* genomes containing ITS1 and ITS2, *S. vesicarium* (VHPU01002624) and *S. lycopersici* (LC032102), were also run through ITSx to isolate these marker genes. The ITS sequences from PNW2016-02 were included in an analysis using the UNITE (Kõljalg et al., 2013) to clarify species predictions.

The annotated genome features of PNW2016-02, *S. vesicarium*, and *S. lycopersici* were parsed to extract the identified gene sequences, and these were then compiled in the R statistical environment (R Core Team, 2017), version 4.1.1, where resultant identified genes were compared using the *ggvenn* package (Yan, 2021). A multi-gene phylogeny was then constructed using the following extracted or downloaded marker: ITS1, ITS2, elongation factor 1α (EF1A), and RNA polymerase II large subunit (RPB1) genes. A simple bootstrap consensus tree was formed using the maximum-likelihood method following the General Time Reversible nucleotide substitution model with uniform rates assumed for transitions and transversions. The consensus tree was based on 1,000 bootstrap replicates.

2.5 HPLC-MS/MS of *Stemphylium* PNW2016-02 Secretome

Stemphylium PNW2016-02 was grown in a biphasic potato dextrose media for two weeks in a 20mL tissue culture flask. The broth was filter sterilized to remove any fungal cells. It was then fractionated using ethyl acetate to separate the organic and inorganic fractions. Three independent triplicate samples were collected from the isolate for each fraction. These were then sent to the Purdue University Metabolite Core and analyzed using HPLC-MS/MS.

2.6 Chemical Structure Cluster Analysis

All metabolic data was retrieved from HPLC-MS/MS analysis of the samples of exudates from *Stemphylium* PNW2016-02. The HPLC-MS/MS results (Supplemental Table 2) included the identity of the chemical compounds from the secretome as well as relative abundance (peak intensity) and confidence scores. Compounds that were significantly enriched against the broth only control samples were then referenced against the Human Metabolome Database (HMDB; Wishart et al., 2018) to identify the most similar spectra. Each spectral comparison was associated with a fidelity score; however, to increase the breadth of the metabolic analysis, these scores were not incorporated into the qualitative analysis of the metabolite data. Each metabolite identifier, as pulled from HMDB, was then queried in our pipeline to NCBI's PUG-REST, a RESTful API for PubChem, in order to retrieve the parsed canonical SMILES encoding for chemical structure; this was done with the aid of the *xml2* (Wickham et al., 2020) and *rvest* (Wickham et al., 2021) R packages. Using the *RxnSim* R package, a similarity matrix was constructed using PubChem's syntax for chemical fingerprinting and, with this a dendrogram of chemical similarity was built for the metabolites of PNW2016-02 using the hclust hierarchical clustering algorithm included with base R program. All graphical representations were generated using *ggplot2* (Wickham, 2016) and this figure was later manually annotated in order to highlight the presence of flavonoids and glycosides in the dendrogram.

2.7 Pathway Mapping

Metabolites that had corresponding entries in PubChem were also queried in our bioinformatic pipeline through the BioCyc database again using *xml2* and *rvest* and the API services to retrieve relevant metabolic pathways. The metabolites were organized into a comparative similarity matrix based on the Jaccard index of their included pathways. The resultant matrix was then used to produce a dendrogram in R with hclust to represent metabolic pathway similarity profiles. The dendrograms from the structural comparison and pathway comparison were rendered into a tanglegram using the *dendextend* (Galili, 2021) R package to highlight similarity and differences between the two clustering approaches.

Specific metabolites of interest, such as antimicrobial and iron-binding compounds, were identified in our pipeline via the ChEBI database based on the database's internal molecular ontology identifiers (Hastings et al., 2016). These data were then overlaid over the structural metabolite dendrogram to aid in identifying clusters of interest. The resulting dendrogram was cross-referenced with the pathway similarity dendrogram to identify overlapping pathways if present. All data pulled from the above-mentioned databases was current as of September 5, 2021.

2.8 Bioinformatic Pipeline Code

The entire bioinformatic pipeline (Figure 6) developed for the integration of genomic and secretomic analysis, including that used in our species-level characterization, used several packages in R as well as several external tools and databases to process the raw data. The codebase is included here in the Supplementary figures S1-S4. The pipeline started with raw whole genome shotgun sequencing data which was checked for quality in Galaxy (see above) and then moved to Funannotate (see above) for functional prediction and annotation. The genomic sequences were passed to ITSx (see above) with the output being sent to UNITE (see above) for species prediction. The sequences following the identified ITS are matched and aligned with those of similar species in MEGA X (see above). This alignment was shortened to include only ITS1 and sent for a maximum-likelihood bootstrap analysis. It is also processed through the UPGMA method for a multi-gene phylogeny. Comparative genomics starts with the named gene products of Funannotate predict which were collected and compared against other species in R. Three packages, *plyr* (Wickham, 2020), *readr* (Wickham et al., 2021), and *data.table* (Dowle & Srinivasan, 2021) were used to coerce the data into a usable format. The *stringr* package (Wickham, 2019) was used to filter out the named gene products and then *ggplot2* and *ggvenn* (Yan, 2021) were used to generate a Venn diagram showing overlapping predicted gene products between the species examined (see above).

The secretions from PNW were quantified using HPLC-MS. The results, essentially compound names, were processed in R using the *data.table* package to extract the compound names. Using *rvest* and *xml2* in R, each compound is queried through the PUG-REST API (Kim et al., 2019) to retrieve canonical SMILES information. The chemical fingerprint of the available SMILES structures was compared with the use of *RxnSim* (Giri, 2017). The data was passed

through *reshape* (Wickham, 2018) to change into a form usable by *ggplot2* for graphical output. *dplyr* (Wickham et al., 2021) and *rlist* (Ren, 2021) were used to reformat data syntax for further analysis by *ggdendro* (de Vries & Ripley, 2020) and *dendextend* (Galili, 2021) into dendrograms and a tanglegram. The package *berryFunctions* (Boessenkool, 2021) is used for error handling while accessing PUG-REST.

Accessing and analyzing data from MetaCyc followed a similar approach to PUG-REST. *reshape2* (Wickham, 2020) was used to coerce the data for graphical output; *stringr* (Wickham, 2019) was also used to reformat query strings. The package *gtools* (Warnes et al., 2021) was used in permutation generation to assign similarity values to a matrix for graphical display. The ChEBI database was obtained through their publicly-available full data spreadsheet allowing for access via *data.table*. Subsequent GO term querying required the use of *xml2*, *httr*, and *berryFunctions* for XML parsing, HTML parsing, and error handling, respectively.

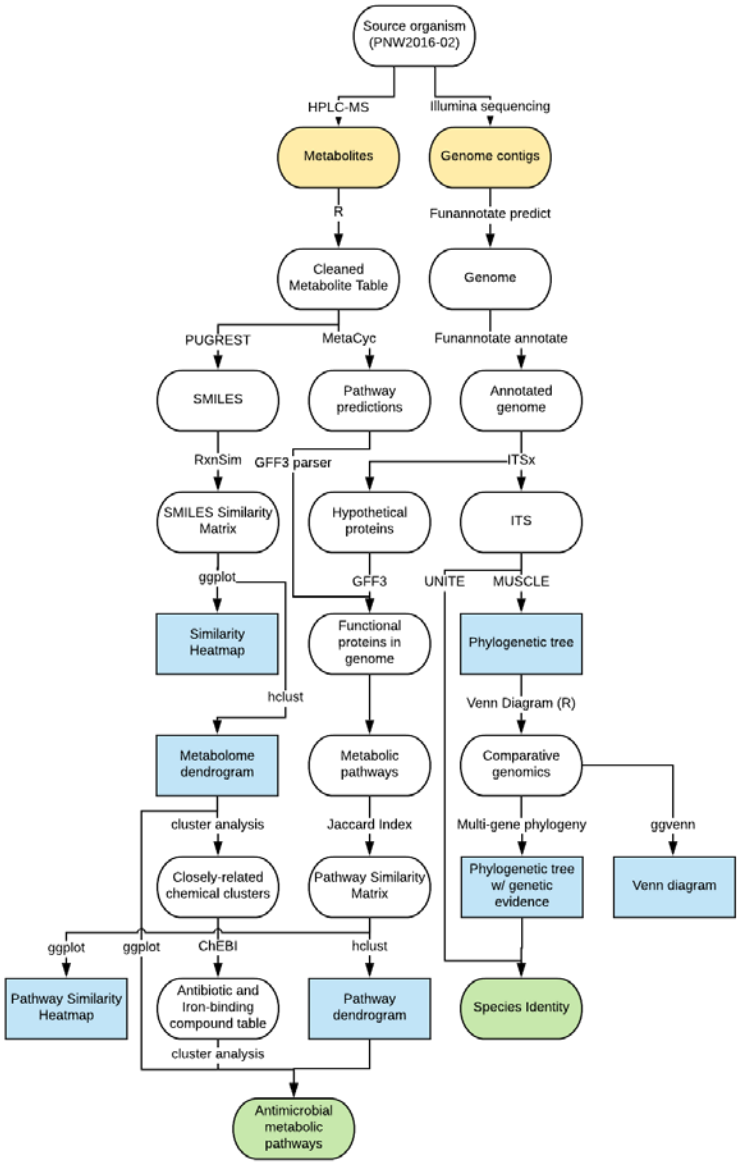


Figure 6. Process diagram of the semi-automated bioinformatic pipeline. Inputs are shown in yellow, graphic output with blue, and results with green.

CHAPTER 3. RESULTS

3.1 Genome Assembly

Whole genome shotgun sequencing of both paired-end reads resulted in over 10 million sequences of 151 ± 2 bp length on average for a total of 1,601,405,614 bp. In the Galaxy portal, the QUAST analysis of the draft genome found 2,126 contigs, with 1,620 being greater than or equal to 1,000 base pairs in length. The largest contig was 569,005 base pairs long lending to an L50 of 88 and an N50 of 119023. The total length of the contigs was 38,042,803 base pairs including contigs less than 1,000 base pairs in length (Figure 7).

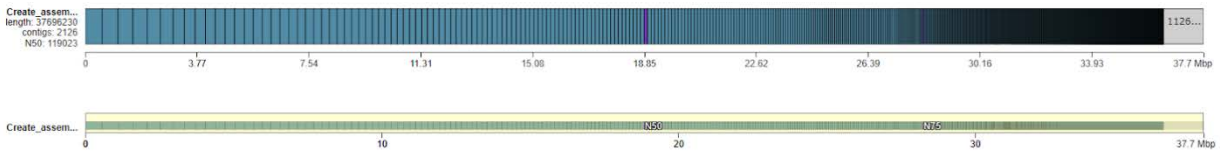


Figure 7. Contig view ordered by size after processing through QUAST. The 1,126 contigs occupy the gray area in the top bar. These contigs are of a length shorter than 2503 bp compounding to a total length of 1,361,516 bp.

Funannotate found that there were in excess of 10^7 sequences of 151 ± 2 base pairs in length in the total coverage. Genome assembly in the Funannotate pipeline also recovered 2126 contigs, and the resultant scaffolding of contigs was assembled into a 37.7 Mbp draft genome. The GC content of the draft genome was 50.63%. Funannotate identified 843 named proteins as well as 2,752 polypeptide domains, with 10,602 hypothetical proteins in addition.

MG-RAST identified 7,600,728 predicted sequence features with 1,751,648 failing in the quality control pipeline and 1,265,490 being of unknown function (Figure 8). Of the 7,600,728 features, 4,775,343 predictions were able to be functionally annotated, whereas 96,944 were classified as ribosomal RNA-related features, and the other 2,728,441 were unable to be annotated (Figure 9). The analysis also revealed some details about the metabolic subsystems present within PNW2016-02. A majority of the identified protein products that were classified into a subsystem

fell into carbohydrate metabolism (19.51%), amino acids and derivatives metabolism (15.01%), protein metabolism (10.13%), or clustering-based subsystems (8.42%; Figure 10).

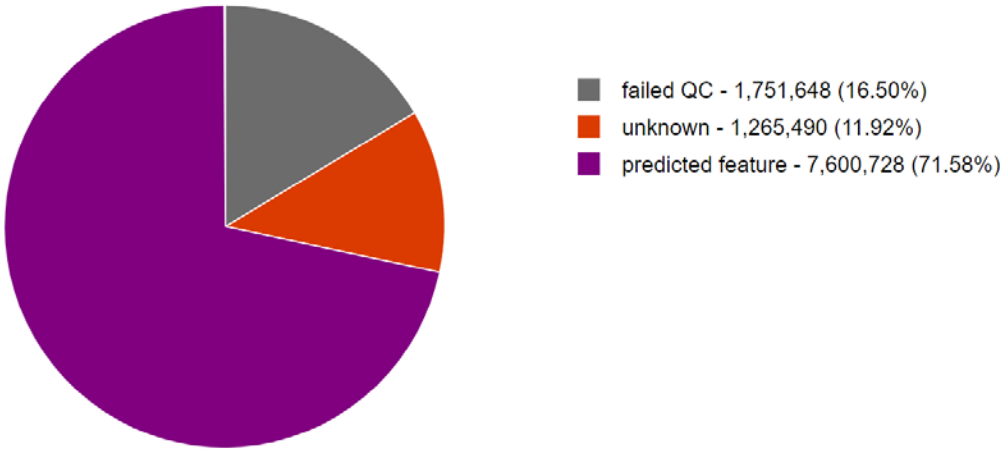


Figure 8. Pie chart showing the relative percentages of functional sequence predictions from the genome of PNW2016-02 in MG-RAST.

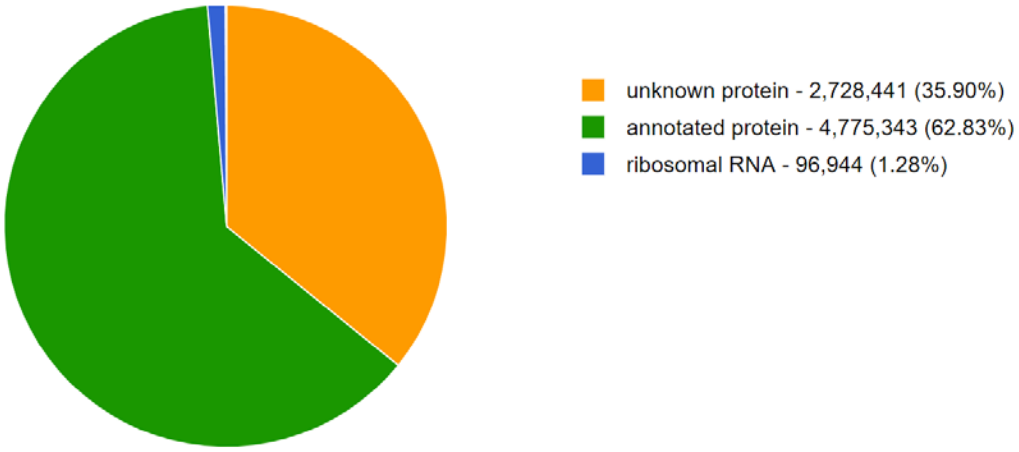


Figure 9. Pie chart showing the relative percentages of predicted sequence features of PNW2016-02 in MG-RAST.

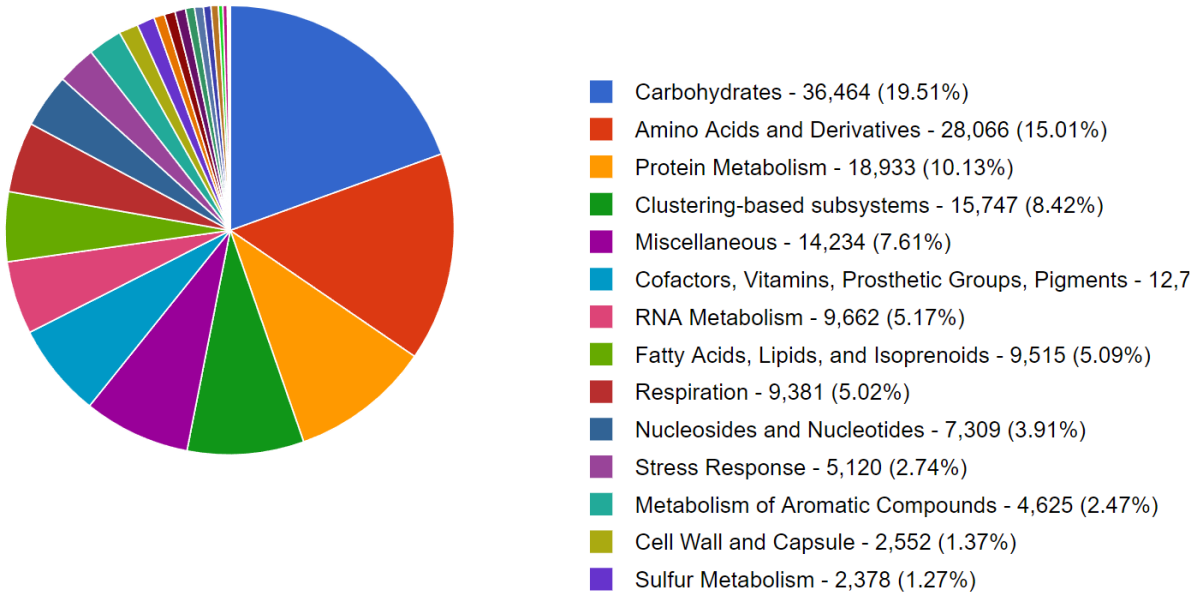


Figure 10. Pie chart showing the relative percentages of functional gene product subsystems identified in the PNW2016-02 draft genome by MG-RAST.

3.2 Phylogenetic and Comparative Genomic Analysis

Using default parameters in UNITE, the BLAST searches of the nine extracted ITS1 sequences returned best hits in *Stemphylium vesicarium* with BLAST E-values all less than $1 \cdot 10^{-70}$. The genome annotation of the three genomes through Funannotate yielded 10,358, 10,698, and 11,586 gene products in *Stemphylium lycopersici*, PNW2016-02, and *Stemphylium vesicarium*, respectively. The R analysis of the predicted functional gene features based on sequences found in the annotated genomes revealed that PNW2016-02 has 699 named gene features in common with *S. vesicarium*, of which fifty gene features were exclusive to the two. In contrast, the isolate shared 659 identified gene features with *S. lycopersici*, with only ten of these being exclusive to the pair (Figure 11).

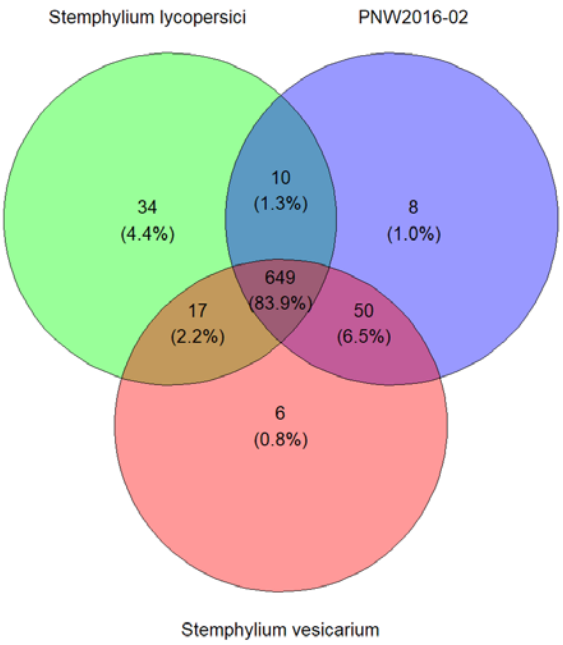
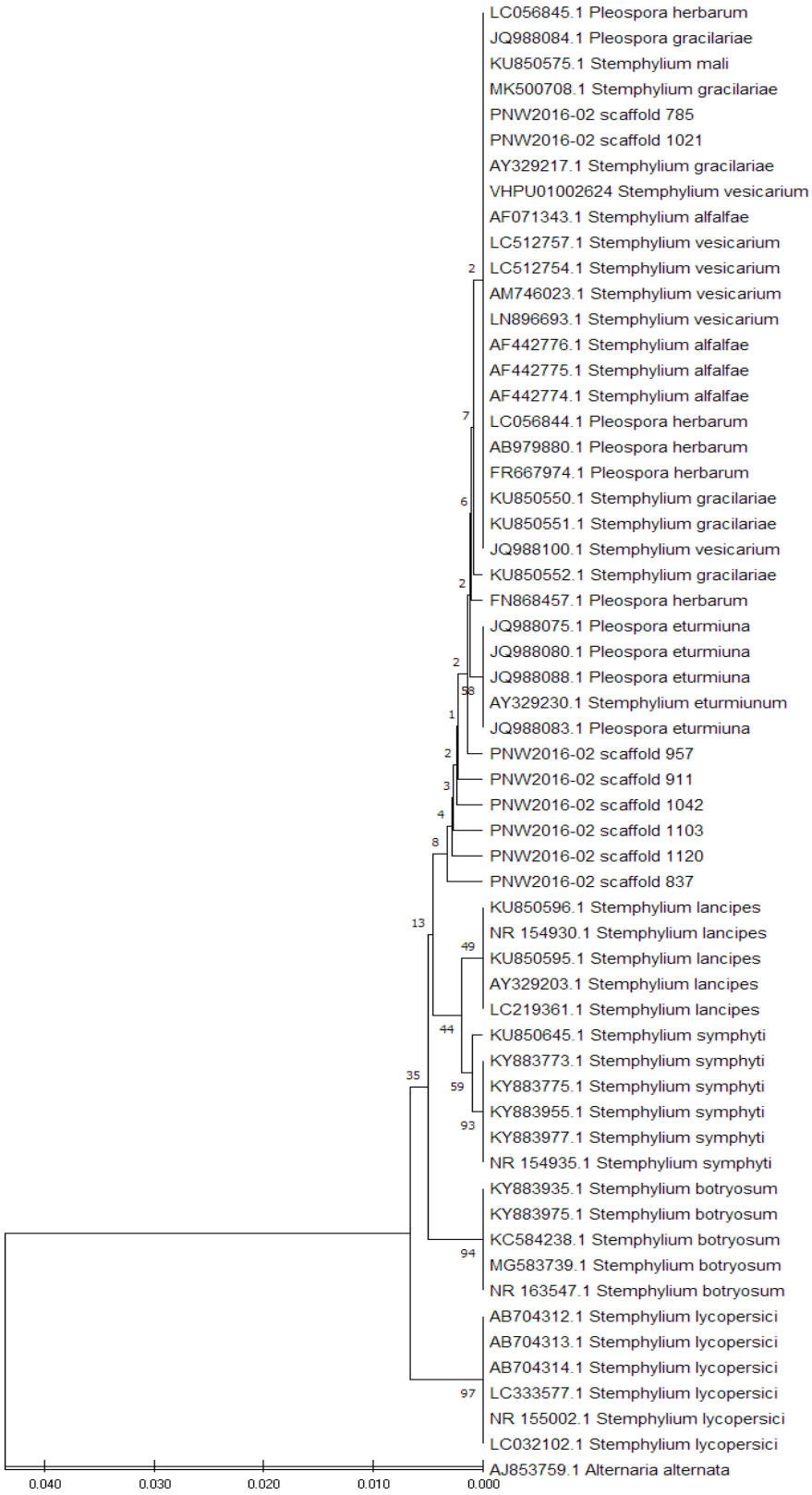


Figure 11. Venn diagram of predicted gene features identified through Funannotate in three isolates of *Stemphylium*. *S. vesicarium* has GenBank accession number VHPU01002624 and *S. lycopersici* has accession number LC032102.

Using ITS1 from several different species within *Stemphylium* or *Pleospora*, an anamorphic form of genus, the phylogenetic analysis included PNW2016-02 within a clade that contained various other isolates of *Stemphylium vesicarium* and other putative *S. vesicarium* synonyms. All ITS sequences recovered from the genome scaffolds extracted using ITSx belonged to one of two distinct clades in the phylogeny; ITS sequences from scaffolds 785 and 1021 fell within a large clade that included all isolates of *S. vesicarium*; whereas, ITS sequences from the other the six scaffolds were basal to that large clade, with this basal clade including ITS sequences representing isolates of *Pleospora herbarum, S. gracilariae, P. eturmiuna, S. alfalfae,* and *S. mali* (Figure 12). Bootstrap values showed only three clades were supported above the 60% level, with these clearly distinct clades representing *S. symphyti*, *S. botryosum*, and *S. lycopersici.*

Figure 12. Molecular phylogeny of selected *Stemphylium* isolates based on partial 18S, ITS1, 5.8S, ITS2, and partial 28S sequence inferred using the UPGMA method in MEGA X. The optimal tree is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (10,000 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths gauged in units of number of base pair substitutions per site as determined by the Tamura-Nei method. The rate variation among sites was calculated with a gamma distribution of shape parameter equaling 1.00. This tree contains fifty-eight nucleotide sequences (10 *Pleospora*, eight scaffolds from PNW2016-02, 39 *Stemphylium*, 1 *Alternaria*). All ambiguous positions were removed from the model for each sequence pair. There was a total of 505 positions in the final dataset.


LC056845.1 Pleospora herbarum
JQ988084.1 Pleospora gracilariae
KU850575.1 Stemphylium mali
MK500708.1 Stemphylium gracilariae
PNW2016-02 scaffold 785
PNW2016-02 scaffold 1021
AY329217.1 Stemphylium gracilariae
VHPU01002624 Stemphylium vesicarium
AF071343.1 Stemphylium alfalfae
LC512757.1 Stemphylium vesicarium
LC512754.1 Stemphylium vesicarium
AM746023.1 Stemphylium vesicarium
LN896693.1 Stemphylium vesicarium
AF442776.1 Stemphylium alfalfae
AF442775.1 Stemphylium alfalfae
AF442774.1 Stemphylium alfalfae
LC056844.1 Pleospora herbarum
AB979880.1 Pleospora herbarum
FR667974.1 Pleospora herbarum
KU850550.1 Stemphylium gracilariae
KU850551.1 Stemphylium gracilariae
JQ988100.1 Stemphylium vesicarium
KU850552.1 Stemphylium gracilariae
FN868457.1 Pleospora herbarum
JQ988075.1 Pleospora eturmiuna
JQ988080.1 Pleospora eturmiuna
JQ988088.1 Pleospora eturmiuna
AY329230.1 Stemphylium eturmiunum
JQ988083.1 Pleospora eturmiuna
PNW2016-02 scaffold 957
PNW2016-02 scaffold 911
PNW2016-02 scaffold 1042
PNW2016-02 scaffold 1103
PNW2016-02 scaffold 1120
PNW2016-02 scaffold 837
KU850596.1 Stemphylium lancipes
NR 154930.1 Stemphylium lancipes
KU850595.1 Stemphylium lancipes
AY329203.1 Stemphylium lancipes
LC219361.1 Stemphylium lancipes
KU850645.1 Stemphylium symphyti
KY883773.1 Stemphylium symphyti
KY883775.1 Stemphylium symphyti
KY883955.1 Stemphylium symphyti
KY883977.1 Stemphylium symphyti
NR 154935.1 Stemphylium symphyti
KY883935.1 Stemphylium botryosum
KY883975.1 Stemphylium botryosum
KC584238.1 Stemphylium botryosum
MG583739.1 Stemphylium botryosum
NR 163547.1 Stemphylium botryosum
AB704312.1 Stemphylium lycopersici
AB704313.1 Stemphylium lycopersici
AB704314.1 Stemphylium lycopersici
LC333577.1 Stemphylium lycopersici
NR 155002.1 Stemphylium lycopersici
LC032102.1 Stemphylium lycopersici
AJ853759.1 Alternaria alternata
0.040
0.030
0.020
0.010
0.000

Another simplified molecular phylogeny to highlight the comparative genomic analysis was undertaken solely using the ITS1 and ITS2 sequences (Figure 13). Fewer isolates were used for each species. The bootstrap consensus tree for the analysis yielded a strong grouping of PNW2016-02 within a clade containing *S. vesicarium* and synonyms, which was similar to the above-mentioned multi-gene phylogeny (Figure 12). Paired with the consensus tree is a graph showing the presence or absence of selected gene features within the genomes of PNW2016-02, *S. vesicarium*, and *S. lycopersici* performed using the named gene product dataset that generated Figure 11.

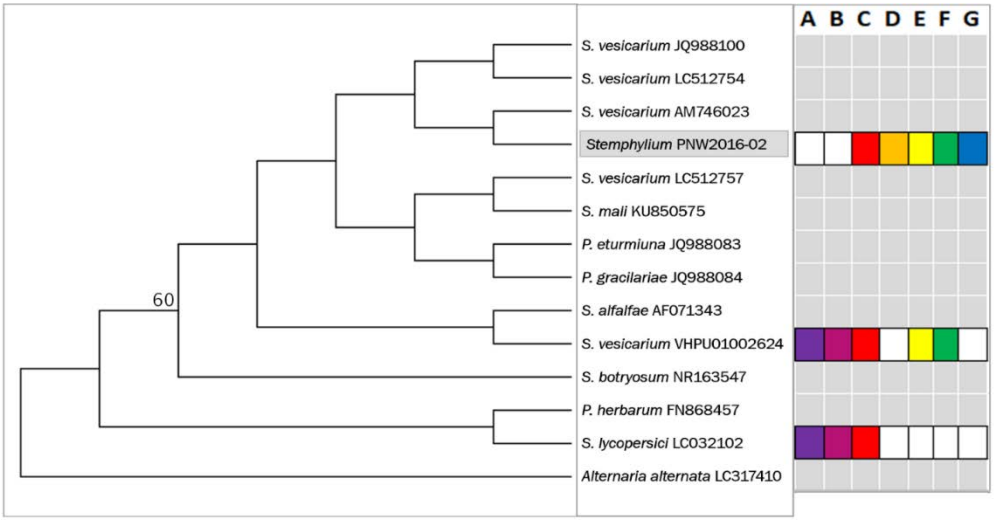


Figure 13. Maximum-likelihood bootstrap consensus tree of ITS1 and ITS2 sequences of several similar species within *Stemphylium* and *Pleospora* with bootstrap values over 50% shown. (right) Graph showing the presence or absence of seven different genes within the genome

3.3 Metabolite Chemical Cluster Analysis

Not all of the metabolites pulled from the HMDB had entries in PubChem; however, a similarity matrix (Figure 14) was produced for the cluster analysis that included the over 200 SMILES structures recovered. The dendrogram (Figure 15) showed the metabolite clustered into four distinct groups. One of these groups contained mostly phenylpropanoid pathway derivatives, including a variety of flavonoids. The similarity matrix (Figure 16) formed from the comparison of the chemical pathway fingerprint (see below) for each of the metabolites included 220 of the

298 metabolites that have canonical SMILES structures in the PubChem database. After hierarchical clustering, a dendrogram was produced (Figure 18) and this allowed for comparison with the cluster analysis of metabolites that had associated SMILES data with relative similarity distances used as branch lengths.

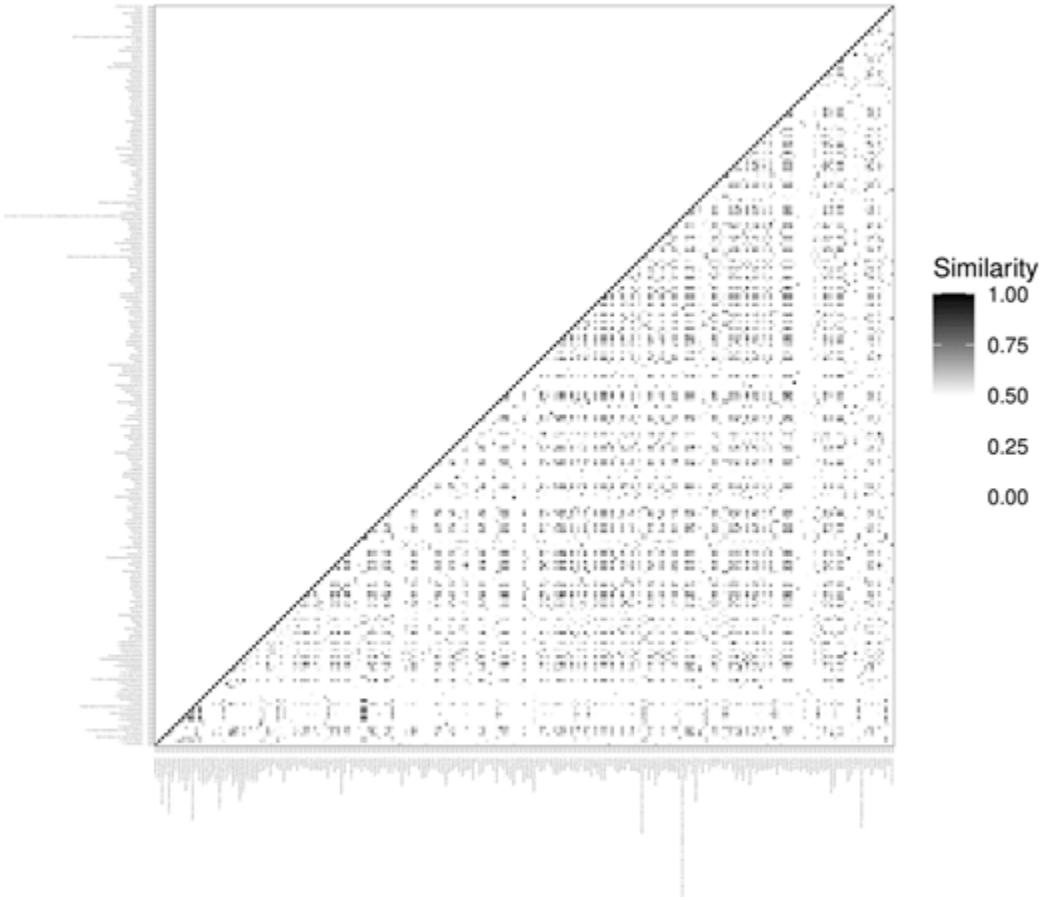


Figure 14. The similarity matrix generated from the metabolites of PNW2016-02 using the PubChem fingerprint syntax. The Tanimoto coefficient is used with a value of 1.00 (black) indicating identical structure and 0.00 (white) indicating dissimilar structure.

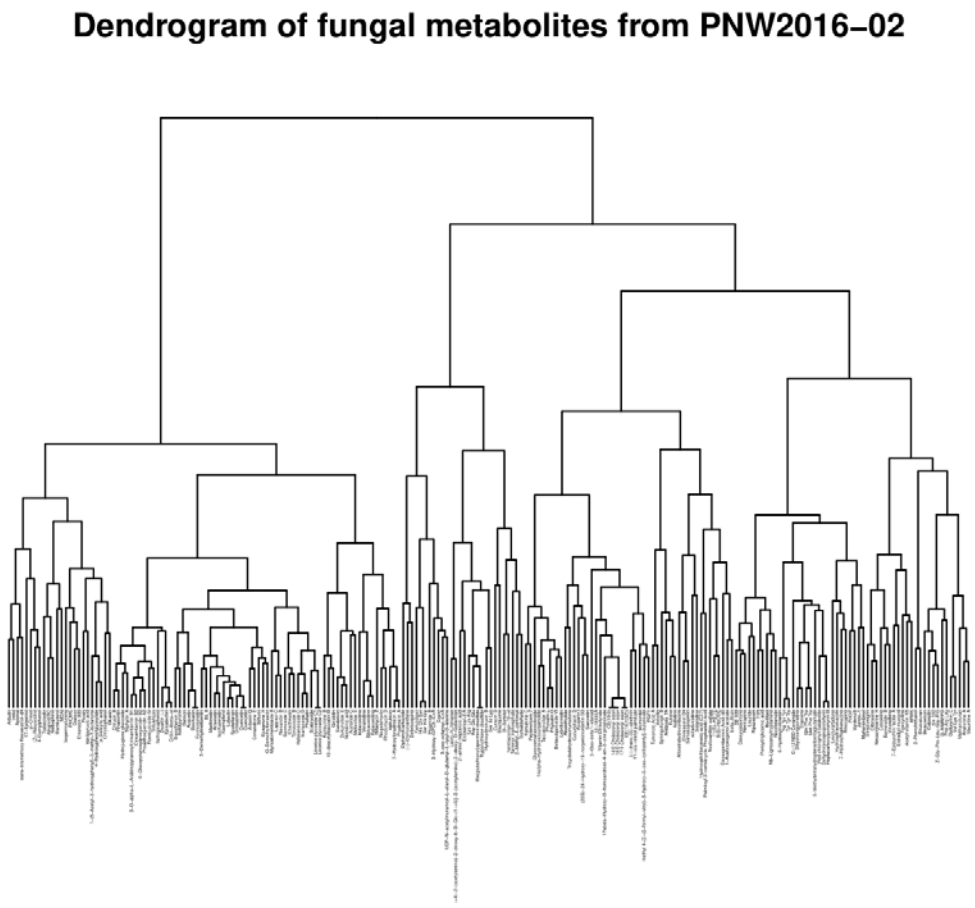


Figure 15. A dendrogram of the metabolites of PNW2016-02 with SMILES formulae found in the PubChem database. The branches are clustered using the Tanimoto coefficient and the PubChem chemical fingerprint syntax.

3.4 Pathway Mapping

Due to the limited number of metabolites with matching entries in the MetaCyc, along with other limitations to the database, and this compounded by the already reduced number of metabolites in our dataset overall for several reasons (see above), there were only few pathways to compare amongst the metabolites. Overall, only thirty metabolites had information in MetaCyc; however, only six had pathway data associated with their entry. A similarity matrix was generated from corresponding metabolites for their metabolic pathways (Figure 16) which was also converted to a dendrogram (Figure 17). The pathway similarity dendrogram (Figure 17) showed some relations among the metabolites that had pathway data but lacked any resolution with regard

to compounds with no such data. A tanglegram was used for the visualization of the comparison (Figure 18) but shows little consistency between the two trees.

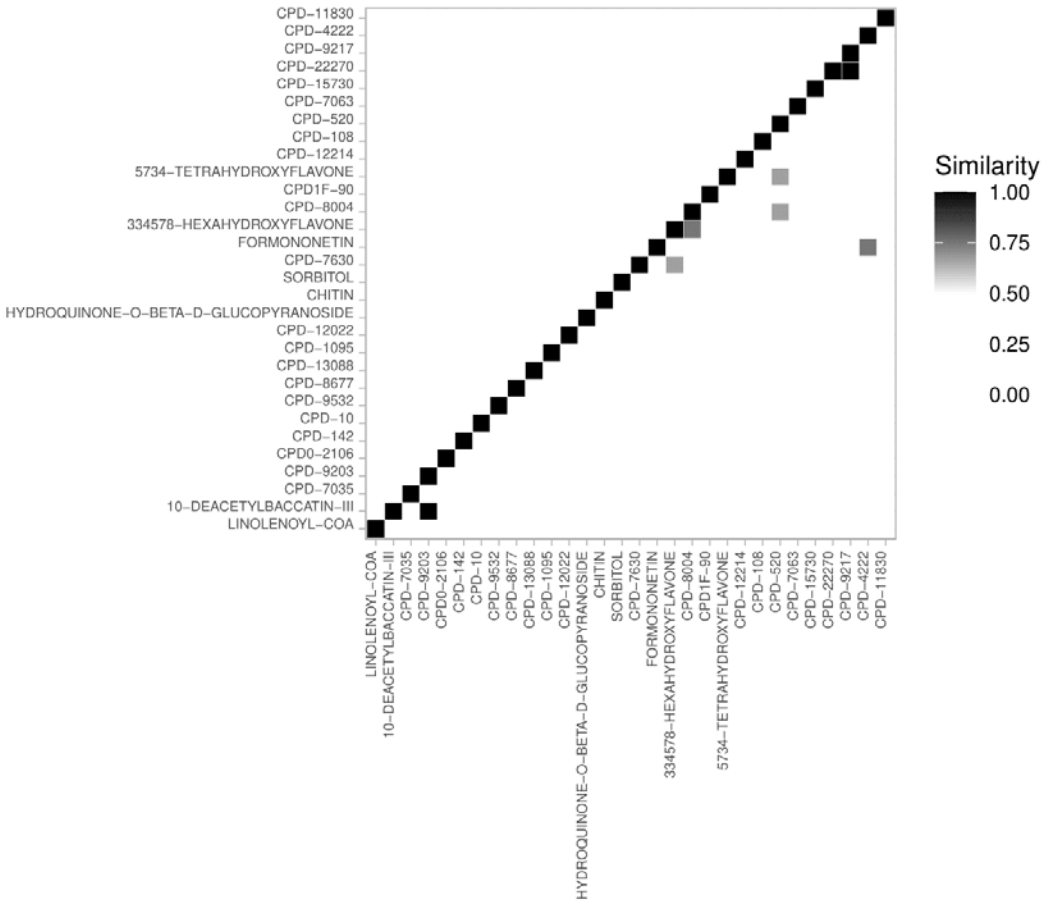


Figure 16. A similarity matrix of the metabolites of PNW2016-02 using the pathway names obtained from MetaCyc. Higher similarity of the pathways in which two metabolites participate are denoted by darker tiles.

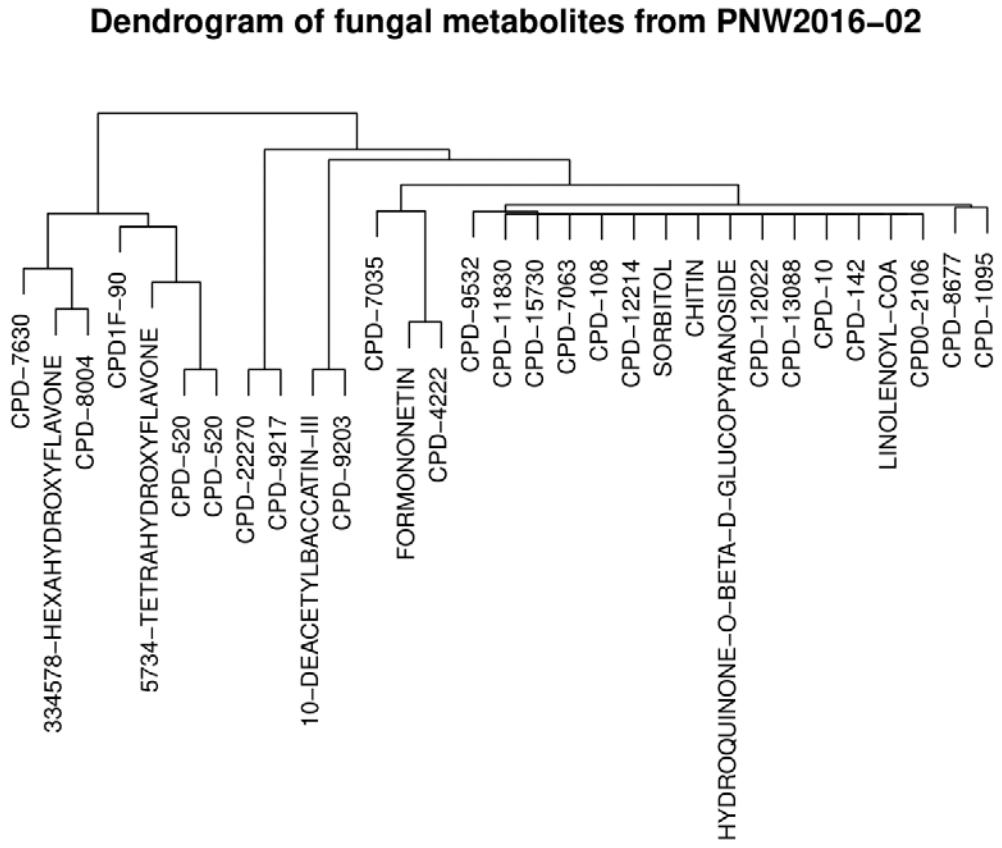


Figure 17. A dendrogram of the fungal metabolites of PNW2016-02 with identified metabolic pathways from the MetaCyc database. The branch lengths represent the relative dissimilarity in relevant pathways to the other branches determined by Jaccard index.

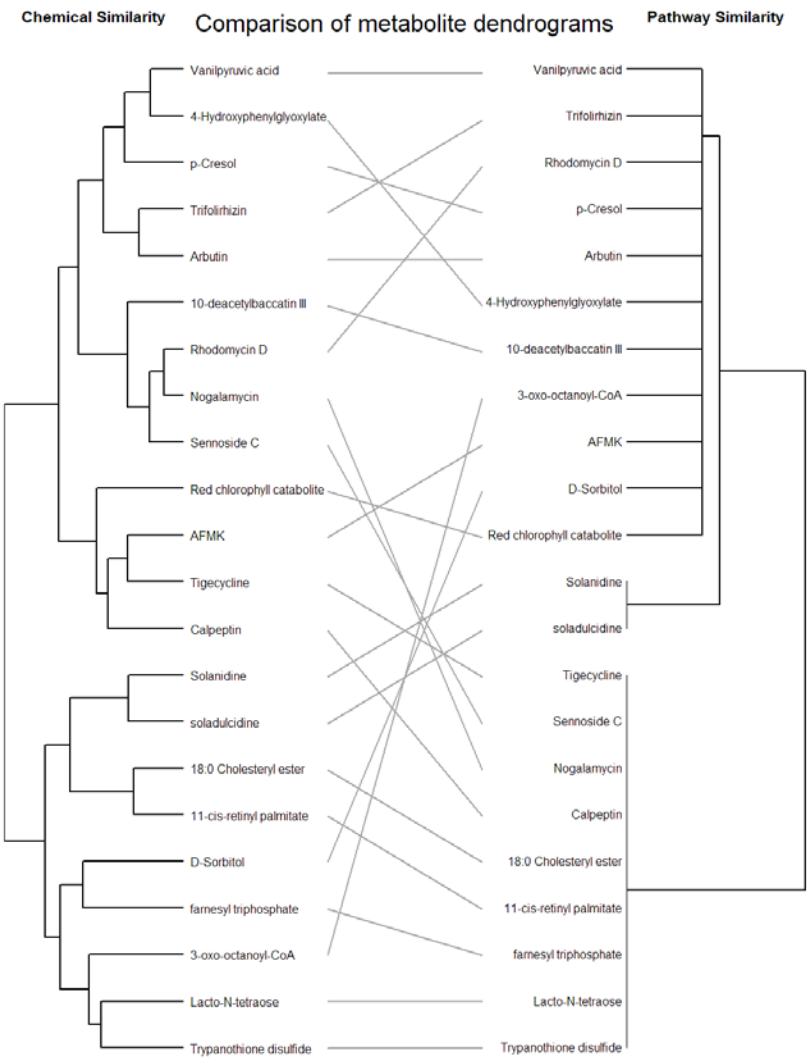


Figure 18. Tanglegram comparing the structural similarity and pathway similarity analyses. The lack of resolution in the pathway data is due to the limited pathway data in the MetaCyc database.

Finally, non-metric multidimensional scaling (NMDS) was used to pare down the structural similarity to identify any additional clustering not represented in the similarity matrix. This revealed clustering involving several compounds including p-cresol, tigecycline, and calpeptin (Figure 19). D-sorbitol and farnesyl triphosphate lie furthest from this cluster on the plot.

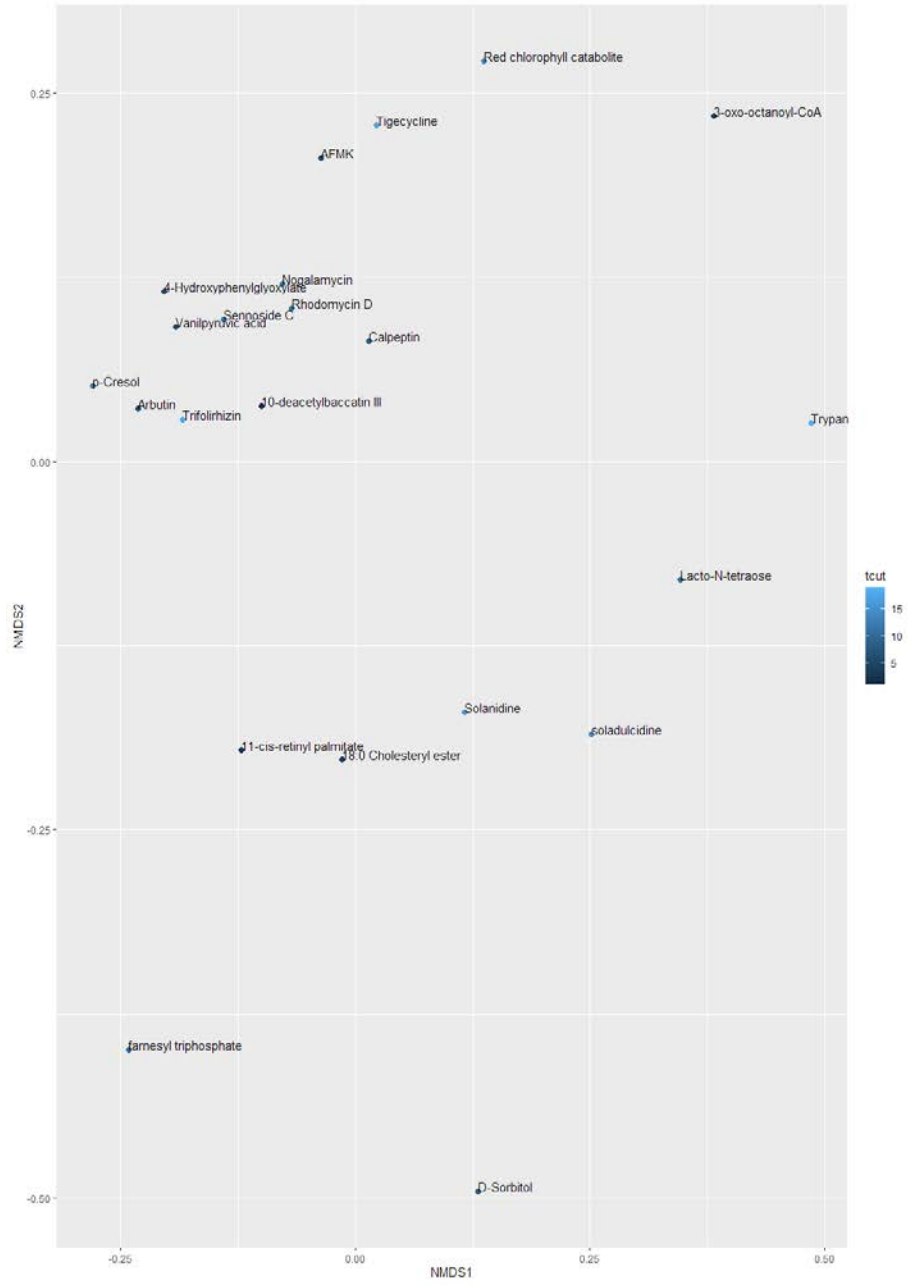


Figure 19. NMDS plot using the structural similarity of metabolites with identified pathways in the MetaCyc database.

CHAPTER 4. DISCUSSION

4.1 Identity of *Stemphylium* PNW2016-02

Stemphylium, a genus of plant pathogens and endophytes, contains a number of well-documented species (e.g., *S. vesicarium*); however, discrepancies in species identities of many isolates continue to be reassigned given new molecular evidence (Woudenberg et al., 2017). Reconstruction of the draft genome of PNW2016-02 and subsequent phylogenetic analysis using marker genes revealed an inconclusive node containing *S. vesicarium*, synonyms of *S. vesicarium*, and *S. gracilariae* (anamorphic *Pleospora gracilariae*) as well as several of the ITS sequences from PNW2016-02. Bootstrap values for the phylogenetic tree, however, show that the clade representing PNW2016-02, as well as other *S. vesicarium* isolates culled from public sequence repositories, was clearly distinct from clades representing both *S. botryosum* (bootstrap value 94) and *S. lycopersici* (bootstrap value 97).

The length of the assembled genome of PNW2016-02 fell within the expected range for *Stemphylium* at approximately 38 Mbp, as the genome of *Stemphylium vesicarium* has been shown to be around 38.66 Mbp (Gazzeti et al., 2019), 39.85 Mbp (Sharma et al., 2020), and 40.83 Mbp (Sharma et al., 2020). The median GC content is 50.7% compared to a GC content of 50.63% in PNW2016-02. *Stemphylium lycopersici* has a genome length near 35 Mbp derived from the average length of three draft genomes with lengths of 35.18 Mbp (Franco et al., 2015), 34.16 Mbp, and 35.00 Mbp (Medina et al., 2018). Genomes for *S. lycopersici* maintain a median GC content at 51.20%.

From the genome annotation, the three isolates had 84% of their gene products in common; however, *Stemphylium lycopersici* only shared 1-2% of its gene products with either of the other genomes, and the gene products of PNW2016-02 overlapped with those of *S. vesicarium* by 7% of the total gene products. The difference in correspondence of the gene products clearly provides genomic support for the functional grouping of PNW2016-02 with *S. vesicarium* rather than *S. lycopersici*. PNW2016-02 did have eight unique named gene products representing 1% of the products in the comparison that were found exclusively in its genome. The non-ribosomal genes exclusive to PNW2016-02 as identified in the Funannotate pipeline are alpha-ionylideneethane

synthase abl3, anthranilate synthase component 1, cysteine dioxygenase, endochitinase B1, endoplasmic oxidoreductin-1, oxidation- resistance protein 1, formate dehydrogenase (NAD+), suppressor of tub2 mutation, and ferrioxamine B transporter.

One of the eight aforementioned gene products, alpha-ionylideneethane synthase abl3 is part of the metabolic pathway in *Botrytis cinerea*, a fungal plant pathogen, which produces abscisic acid which is part of the pathogen's mechanism of establishing infection (Izquierdo-Bueno et al., 2018). In this process, abscisic acid is produced from the cyclization of farnesyl diphosphate (Inomata et al., 2004), a lower energy form of farnesyl triphosphate which can be found in the metabolome of PNW2016-02. Anthranilate synthase component 1 is part of a protein that catalyzes the formation of anthranilate from chorismate, a product of the Shikimate pathway, in the tryptophan synthesis (Herrmann & Weaver, 1999; Lin et al., 2009). Cysteine dioxygenase is part of the metabolism of cysteine where it oxidizes the thiol group in the residue to form cysteine sulfinic acid. The enzyme's active site uses an iron cation to perform the necessary electron transfers (Joseph & Maroney, 2007). Chitinase activity is the primary action of endochitinase B1 (Jaques et al., 2003); however, the enzyme also has some non-specific glycosyltransferase activity (Lü et al., 2009). Endoplasmic oxidoreductin-1 is involved in the formation of disulfide bonds in proteins within the endoplasmic reticulum. The enzyme works in conjunction with protein disulfide isomerases to redistribute disulfide bonds in unfolded proteins while producing reactive oxygen species (Tu & Weissman, 2002). Conversely, oxidation resistance protein 1 works to reduce the impact of reactive oxygen species (Volkert et al., 2000). Formate dehydrogenase functions as the last step in the detoxification of methanol, formaldehyde, and formate to carbon dioxide (Overkamp et al., 2002; Chistoserdova et al., 2004). Suppressor of tub2 mutation is a microtubule-binding protein that stabilizes microtubules during mitosis (Yin et al., 2002). Ferrioxamine B transporter or siderophore iron transporter 1 is a transmembrane transporter protein that actively shuttles the siderophore ferrioxamine B into the cell (Lesuisse et al., 1998). The genes solely identified in PNW2016-02 do not share any functional relationships nor do they appear to participate in any of the same metabolic pathways. Of note, PNW2016-02 carries genes encoding for enzymes that help establish fungal plant infections corresponding with the phytopathogenic effects associated with some *Stemphylium* isolates. The presence of an exclusive siderophore transporter does also implicate iron in one form of potential antimicrobial activity for PNW2016-02, for example, via direct iron-nutrient competition.

Genomic and phylogenetic analyses that included other representatives of *Stemphylium*, suggested a strong relationship between PNW2016-02 and *Stemphylium vesicarium*. Blast results from UNITE also suggested all nine ITS1 sequences recovered from the PNW2016-02 draft genome to be highly similar to *S. vesicarium*. UNITE currently only includes ITS sequences for the three closely related species: *S. vesicarium, S. loti,* and *S. solani*. According to the maximum-likelihood analysis performed by Woudenberg et al. (2017), *S. solani* fits into the phylogeny as a sister taxon to *S. lycopersici.* The maximum-likelihood bootstrap consensus tree presented here placed PNW2016-02 within a cluster of species including *S. vesicarium* and synonyms as well as *S. gracilariae*. Analysis using the UPGMA method corroborated the groupings around PNW2016-02; however, the branch structure of the phylogeny isolated the ITS sequences found in PNW2016-02 to two nodes: one within the aforementioned group and the node basal to the group. The separation of the sequences may be due to the lack of complete coverage of each scaffold within the length of 505 nucleotides included in the analysis. Only one of the sequences (scaffold 1042) of PNW2016-02 occupy all sites. The shortest sequence is scaffold 1021 at 234 base pairs in length.

Taken together, the similarity in genomic features as well as the congruence of gene products and ribosomal DNA patterns, confirm the species identity of PNW2016-02 to be *Stemphylium vesicarium*. The distinct genes in the PNW2016-02 genome that were distinguishable in the genomic analysis using other published genomes of *S. vesicarium* and *S. lycopersici* do raise interesting questions about the uniqueness or our strain, especially with respect to exclusive genes that appear to be associated with flavonoid production, the Shikimate pathway, and iron sequestration or transport; however, the majority of these genera appear to be generally common to fungal species, including plant pathogens. Understanding the unique characteristics of PNW2016-02, however, awaits further experimentation that might include, for example, monitoring expression of distinct identified genes related to iron sequestration and transport in transcription (e.g., using RNAseq).

4.2 Potential Activity of Evaluated Metabolites

Previous work with PNW2016-02 involving knockout mutants of *Serratia marcescens* found that mutants with knockouts in iron regulatory pathways were more susceptible to the

inhibitory activity of the isolate's secretions (Sopovski, 2019). Iron is an essential mineral for most life on Earth; this includes microorganisms like endophytes. It can determine the success of a pathogen or the resistance of a host (Greenshields, Liu & Wei, 2007). Iron complexes act as a critical ferry for electrons and small molecules as well as a reducing agent to reactive oxygen species present in many organisms like heme proteins in mammals and iron siderophores (*sidero*-, "iron;" -*phore*, "bearing") like triacetylfusarinine C and ferrichrome in fungi (Emery & Neilands, 1961; Adjimani & Emery, 1988; Chen, Yang, & Chung, 2014). In microorganisms, these small iron-chelating compounds are produced by many species in the absence of available iron to increase the uptake of the ion from the environment (Siebner-Friedbach et al., 2005). One mechanism by which PNW2016-02 may facilitate iron-dependent antimicrobial activity is through secreted siderophore nutrient scavenging. This action has been found in rice with the bacterial endophyte *Streptomyces sporocinereus* against the plant pathogen *Magnaporthe oryzae* (Zeng et al., 2018). Similarly, *Aspergillus fumigatus* uses the secreted iron-chelating compounds to resist the antifungal activity of *Pseudomonas aeruginosa* (Sass et al., 2019). Chelation involves organic compounds exchanging electrons with metal ions by complexing with one-another. This action of harvesting environmental metal ions is an effective strategy in microbial competition especially with microorganisms that can be sensitive to ion flux (Kousser et al., 2019). By cornering the market for available essential nutrients through chelation, microorganisms are able to starve and eliminate competition (Maheshwari, Bhutani, & Suneja, 2019).

Siderophores refer to any small compound that can transport iron ions, but most naturally derived siderophores occupy at least one of three classes—catecholates, hydroxamates, and carboxylates—but may have the moieties of more than one class (Miethke & Marahiel, 2007). Within the secretome of PNW2016-02, there are several compounds with one or more siderophore moieties from all three classes. Catecholate moieties occurred most often coinciding with the abundance of phenylpropanoid pathway derivatives including flavonoids, stilbenoids, and tannins. Compounds such as italidipyrone with similar lipophilic structures surrounding oxygen atoms would be able to coordinate with iron taking into consideration polar surface area and molecular geometry (Chaves et al., 2012). Overall, this suggests that PNW2016-02 uses a multifaceted approach to iron sequestration in a strategy that has been likely highly effective against other bacteria, including especially the iron regulatory pathways knockouts of *S. marcescens* mentioned above (Sopovski, 2019).

In addition to siderophores, the secretome of PNW2016-02 also contains several known antibiotics that can be utilized in microbial competition. Of these antimicrobials, some are bacterial metabolites with antibacterial properties. Streptonigrin, rifamycin Z chartreusin, and enterocin 900 fall under this category. There are several similar antibiotic compounds found in the secretome which are termed anthracyclines; these include nogalamycin, rhodomycin D, and pradimicin A. The presence of the gene product predictions of several different polyketide synthases, necessary enzymes in the synthesis of anthracyclines, in the genome of PNW2016-02 supports the existence of an anthracycline synthesis pathway (Metsä-Ketelä et al., 2007). A similar compound, tigecycline, was reported in the metabolome as a modified version of tetracyclines known as a glycylcycline (Greer, 2006). All of these compounds have broad-spectrum antibiotic activity or are further metabolized to have broad-spectrum activity (Bhuyan & Reusser, 1970; Oki et al., 1987 Dickens et al., 1997) with the exception of the limited activity of enterocin 900 on *Listeria monocytogenes* and *Enterococcus faecium* (Franz et al., 1996). Another compound with antibiotic properties, antibiotic TA has been observed in limiting the growth of *E. coli* (Xiao et al., 2012). The secretome of PNW2016-02 contains both siderophores and antimicrobial compounds that can be used in microbial competition *in vitro*. The variety of antimicrobials produced by PNW2016-02 suggests the isolate is able to compete with a broad range of other microbes within the endophytome and even repel other phytopathogens. These results are consistent with what was shown in the growth inhibition experiments, where compounds exuded in the secretome of PNW2016-02 were able to suppress the growth of a number of Gram-positive and Gram-negative bacteria, with the exception of *Serratia marcescens* which appears to have more aggressive iron nutrient scavenging strategy than PNW2016-02. Overall, these results also suggest that PNW2016-02 may have utility for a model system within spinach to suppress bacterial species *in plantae*, including, potentially, those that are pathogenic to humans, such as EHEC.

4.3 Secretome Flavonoids and Related Compounds

Flavonoids have been of interest in several areas of therapeutics including antimicrobial research for decades (Barber, McConnell, & DeCaux, 2000). These compounds occupy a multifunctional class of polyketide molecules derived, for example, from the phenylpropanoid synthesis pathway which itself derives chemical structures from the shikimate pathway (Figure 22)

that is found in virtually all fungi (Richards et al., 2020), and that are known to occur more frequently in the genera such as *Stemphylium* (Ge et al., 2019). A majority of the compounds in the flavonoid class have been identified in many different organisms that include plants, fungi, and algae (Zhang et al., 2017; Yonekura-Sakakibara, Higashi, & Nakabayashi, 2019; Mohanta, 2020; Richards et al., 2020). Flavonoids consist of several groups with closely related carbon backbones formed from a C6-C3-C6 configuration with two phenyl groups separated by a heterocyclic ring containing oxygen. They occupy a diverse range of functions where some have been found to nurture the formation of hyphae in arbuscular mycorrhizal fungi or inhibit spore germination in other fungal plant pathogens (Cushnie & Lamb, 2005; Salloum et al., 2018). Those of particular interest are highly hydroxylated flavanols, like kaempferol with four hydroxyl groups, quercetin with five hydroxyl groups, and gossypetin with six hydroxyl groups all of which have a high potential for glycosylation (Musialik et al., 2009). The abundance of hydroxyl groups in these compounds lends to their antioxidant properties including the reduction of metal ions and chelation of elemental metals from their oxidized forms (Escandar & Sala, 1991). Endophytic, and more specifically phytopathogenic, fungi use this antioxidant property of flavonoids to resist the host-response to infection characterized by the release of reactive oxygen species near the sites of infection (Zou, Wu, & Kuča, 2020).

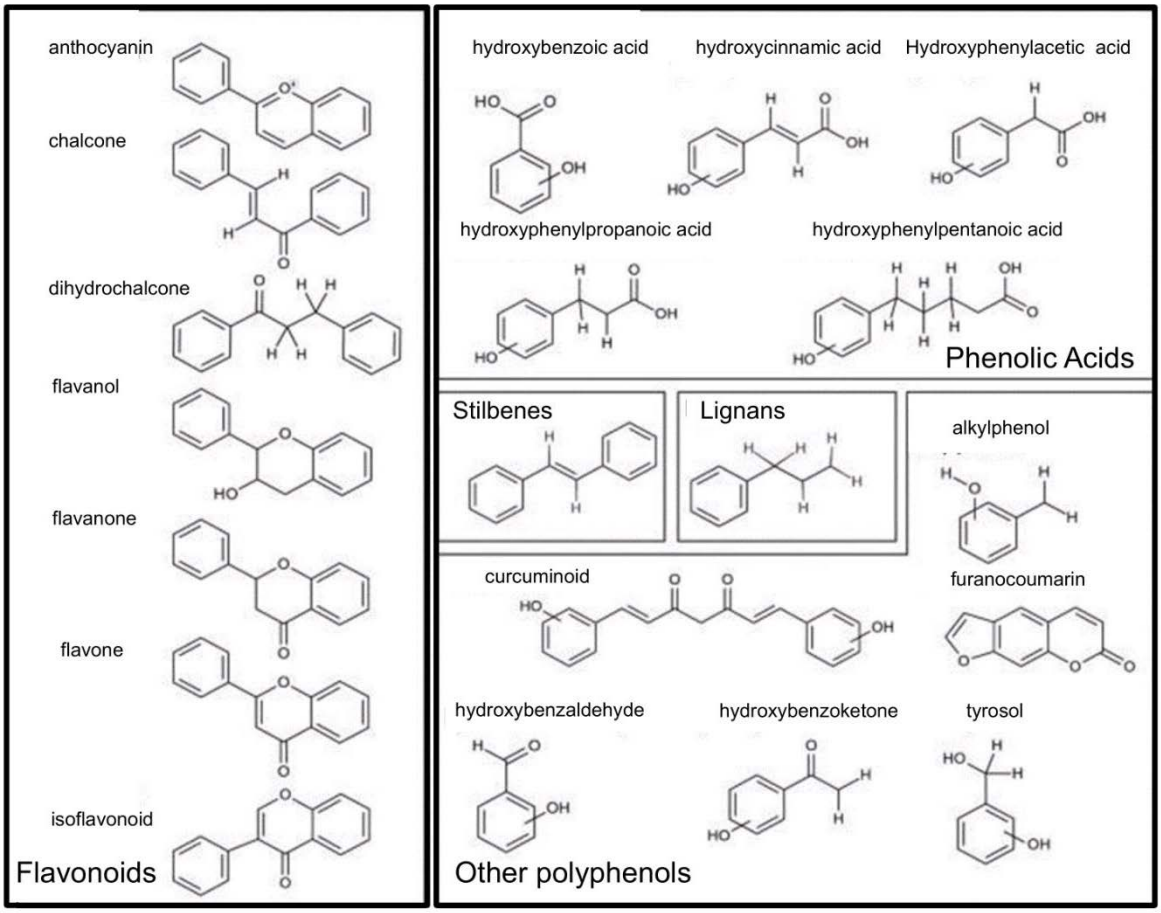


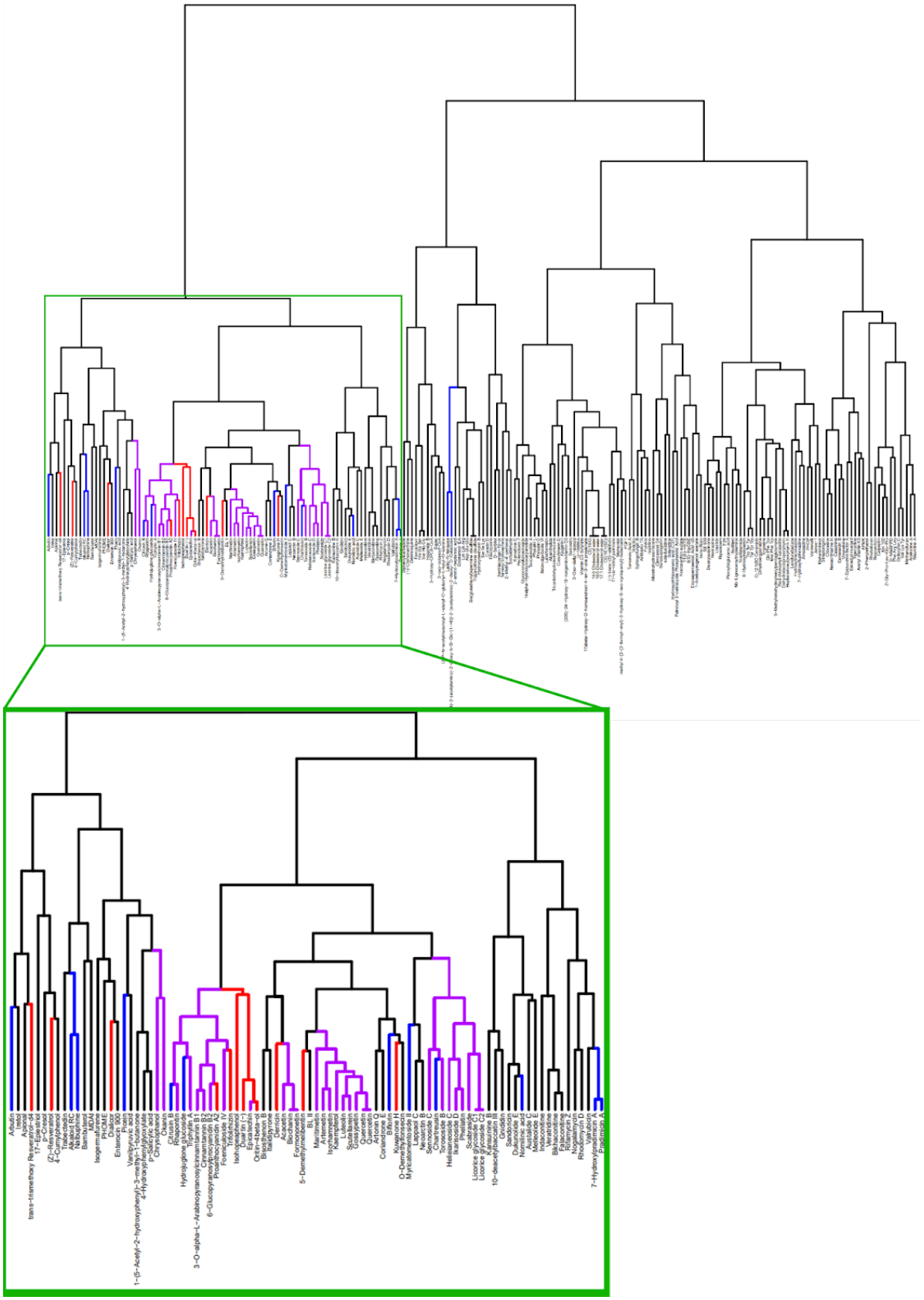
Figure 20. Flavonoid chemical class groupings of secondary metabolites produced by the phenylpropanoid pathway (figure after Lacroix et al., 2018)

Glycosylation is the modification of organic compounds by the addition of one or more sugar moieties. This is done by almost all organisms in order to facilitate changes in structure, activity (e.g., as a chemical switch), and localization of the original compounds (Vasudevan & Lee, 2020). Focusing on the glycosylation of small molecules, the chemical modification of the original compound, referred to as an aglycone, by the linkage of the carbohydrate group, known as a glycone, results in a glycoconjugate that can be further classified as a glycoside (O-linked) or glycosylamine (N-linked) (McNaught et al., 1996). In the general transformation of the chemical properties of a compound of origin, glycosylation tends to increase the aqueous solubility of some compounds it modifies due to the several hydrogen bond-capable functional groups on the glycone (Hollman et al., 1999). The increased solubility lends itself to increasing the membrane permeability of glycosides aiding in the diffusion of small compounds like phytoalexins; however,

it negatively affects the permeability of larger molecules like glycoproteins (Kren & Rezanka, 2008). The glycosylation of small molecules is achieved through the glycosyltransferase protein family which have moderate specificity toward groups of aglycones. Removing glycones is the primary activity of glycoside hydrolases (Henrissat & Davies, 1997). As previously mentioned in the discussion of predicted gene products exclusive to PNW2016-02, endochitinase B1 has some nonspecific glycosyltransferase activity; however, it is not the only enzyme in the genome of PNW2016-02 that is able to transfer glycoside moieties.

Flavonoids constitute a significant percentage (~21%) of the identified compounds in the metabolome of PNW2016-02. Related compounds such as stilbenoids (e.g., trans-trismethoxyresveratrol-d4) also co-located with flavonoids in the structural similarity dendrogram (Figure 21). A total of sixty-four flavonoids and other phenylpropanoid pathway derivatives were identified of which thirty-nine are shown on Figure 21. Interestingly, fifty of the sixty-four had at least one glycosyl moiety accounting for almost 80% of the flavonoids. An additional twenty-three compounds were glycosylated but were not flavonoids. The structural dendrogram was able to cluster the flavonoids together; however, it did not resolve clusters at the chemical subclass level. An example of this is evident in the clustering of acacetin, a flavone, with two isoflavonoids, formonetin and biochanin, instead of quercetin, another flavone. Other compounds (e.g., iretol) were interspersed in the flavonoid clusters using the shared phenol/catechol moiety. Steroids were established in a well-resolved cluster. Tripeptides did not resolve into one cluster in the dendrogram likely due to the variety of the residues; tripeptide sequences with two identical amino acids did cluster. Peptide sequences >3 amino acids held no strong affiliation with other polypeptides. The overall organization of the dendrogram, as computed from the SMILES annotation, in Figure 21 appears to be related to gross structure of the compounds themselves, for example with chemicals having higher oxygen-to-carbon ratio found toward the left, and this is to be expected considering the algorithm used here. However, the close proximity of some biologically active compounds, such as proanthocyanidin A2, with their glycosylated forms, including 6-glucopyranosylprocyanidin B2, at the finer clustering scale in the dendrogram, suggest we may also be detecting the regulating activity of these compounds, as for example, in a chemical switching process. Compounds clustered to the right of the dendrogram, conversely, tend to have less oxygen and more nitrogen such as in the several tripeptides or lisuride.

Figure 21. Structural similarity dendrogram (top) of the metabolome of PNW2016-02 with overlay of flavonoids (red), glycosides (blue), or both (violet). Glycosylated compounds were placed and named according to their aglycone but labelled as a glycoside. Expanded view of the subtree containing all of the flavonoids (bottom).



As a class of polyphenols, flavonoids can perform numerous roles. One of these roles is as a phytoalexin serving to reduce the bioburden or prevent the infection of a plant's tissues by harmful microbes (Nandakumar et al., 2021). One of the ways that flavonoids are able to protect host tissues is by infiltrating and destabilizing bacterial cell membranes including the cell membrane of *E. coli* (He et al., 2014). Another method involves the interference and induction of degradation of quorum-sensing receptors such as TraR in *Pseudomonas aeruginosa* which reduces the ability of the bacteria to form biofilms which leads to a decrease in bacterial propagation (Zeng et al., 2008). Another study found flavonoid compounds in a series of fungal endophytes which exhibited bacteriostatic and inhibitory activity when plated with *E. coli*, *P. aeruginosa*, *Staphylococcus aureus*, and *Bacillus subtilis* (Tang et al., 2020). Of the compounds found that had antimicrobial properties, glycosides of formonetin and epigallocatechin gallate are also found in the secretions of PNW2016-02 which likely contribute to the antimicrobial capabilities of the fungus. Flavonoids and other derivatives of the phenylpropanoid synthesis pathway do appear to constitute a large part of the secretome of PNW2016-02 and may play a key role, overall, in the antibacterial activity of the isolate.

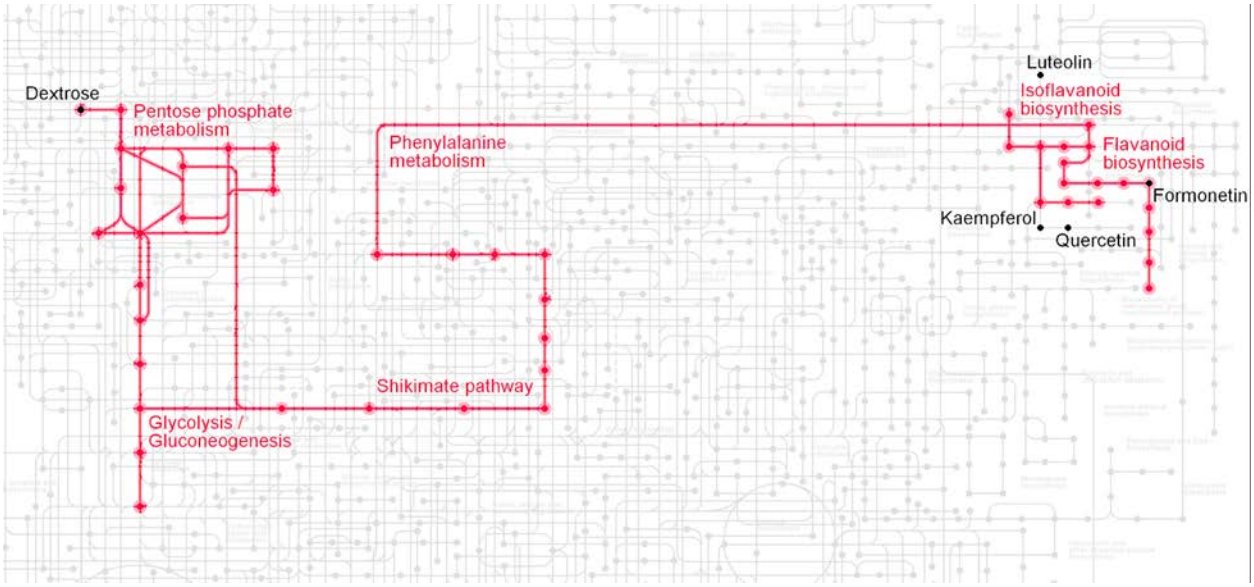


Figure 22. Section of metabolic pathways (map01100) in KEGG (Kanehisa et al., 2016) with pathways relevant to flavonoid synthesis highlighted in pink. Aglycones found in the metabolome of PNW2016-02 are labelled in black.

In the analysis using our bioinformatic pipeline, we encountered a number of compounds in the secretome, such as (+)-maackiain 3-O-glucoside and 2,3,4,4-tetrahydroxychalcone, that could not be located, and/or linked to metabolic pathways, using the publicly available databases, such as MetaCyc. Though there was less data in these databases than necessary to support a fully developed pathway reconstruction using our pipeline, evidence from metabolomic sources point toward the operation of some pathways that produce secondary metabolites within PNW2016-02. In the secretome, flavonoids and closely related polyphenolic compounds are very prevalent, as are many different configurations of steroid-containing molecules, indicating the presence of enzymes that catalyze the formation of these compounds. Figure 22 highlights a metabolic pathway that leads from common carbon source, D-glucose, through the pentose phosphate pathway, the shikimate pathway, and phenylalanine metabolism, then branches out into flavonoid biosynthesis (Kanehisa et al., 2016). Within the genome of PNW2016-02, genome annotation recovered domains, for example, associated with polyketide synthesis, chalcone-flavanone isomerization, chorismate-dependent anthranilate synthesis, and overall shikimate metabolism. Pathways leading to flavonoid biosynthesis have been found in other endophytic fungi namely *Aspergillus oryzae* (Seshime et al., 2005) as well as *Alternaria*, a close evolutionary relative to *Stemphylium* (Lu et al., 2019), and such compounds have been found abundantly in these genera, *Stemphylium* included (Ge et al., 2019). Work done involving *Alternaria* included supplementing its media and observing metabolic changes. Of note, the output of a metabolite falling under the flavonoid-like stilbenoid class called resveratrol increased to 25% above its original rate (Lu et al., 2019). This indicates a metabolic pathway linking sugar metabolism to phenylpropanoid synthesis exists in *Alternaria*. Given the ecological and phylogenetic proximity of *Alternaria* with *Stemphylium*, there likely exists a similar pathway in PNW2016-02.

While there are still gaps in the public database that prevent our bioinformatic pipeline from being able to fully reconstruct all metabolic pathway activity for our isolate and its secretome, our approach was successful overall in finding clear evidence for some pathway-associated activities, and this aspect of the analysis pointed to an important, perhaps central, role of the shikimate pathway and its products in the antimicrobial activities of PNW2016-02. As more data becomes available, more complete pathway reconstructions will no doubt be possible in the future. Further, our work here does show the excellent potential for using bioinformatic pipelines, such as the one developed here, in metabolic pathway discovery and reconstruction for gaining important

insights into the overall chemical functioning of microorganisms and their potential biological activities.

4.4 Implications and Other Future Directions

Pertaining to the questions answered in this thesis, the establishment of the species identity of PNW2016-02 as *Stemphylium vesicarium* will be critical in further studies of the isolate and its utility in model systems. In addition to the species identity, phylogenetic placement of the species revealed some overlap between the analyzed genes of *S. vesicarium* and *S. gracilariae*. Additional research would be required to establish genetic partitions between the species definitions of *S. vesicarium* and *S. gracilariae*. Some genes were present in the genome of PNW2016-02 that may differentiate it as a distinct form of *S. vesicarium*. The functionality and expression of these genes would need to be determined to establish them as being critical for differentiating our isolate. This could be achieved through the analysis of transcriptional magnitude through the use of reverse transcriptase polymerase chain reaction (RT-PCR). Of pertinence, measuring the activity of the ferrioxamine B transporter through both RT-PCR and spectrophotometric analysis as described previously (Kosman, 2003) would help support the iron sequestration hypothesis in that the ferrioxamine B transporter is able to transport multiple siderophores through fungal cells in PNW2016-02.

Implementation of a bioinformatic pipeline to integrate metabolomic and genomic, or other omics data can be useful in the discovery of natural products and exploration of the biosynthetic potential of understudied species like the endophytic fungi examined here. A further integrated automation and more fully developed databases will be a critical next step in the application of the outlined pipeline. Furthermore, genome annotation using Funannotate yielded several predicted gene products as well as hypothetical proteins carrying a variety of domains that may be used for further analysis of the metabolic capabilities of PNW2016-02, particularly in linking functional genes of the genome with their role in the production of compounds observed in the secretome via metabolic pathways that can be ascertained. Optimally, this would be achieved through the integration of additional bioinformatic analytical tools within the semi-automated pipeline established in this thesis. This would ultimately allow for a more robust understanding of the PNW2016-02 secretome, and metabolome more generally, the pathways by which the

secondary metabolites are produced, as well as how the gene-products discovered in the PNW2016-02 genome are integrated within those pathways.

Determination of the validity of the glycoside chemical switching hypothesis could be carried out by introducing glycoside hydrolases to the secretions of PNW2016-02 and measuring the presence of flavonoid aglycones and antimicrobial activity of the resultant product. Current databases appear to include less information on glycosides than on aglycones, so further bioinformatic analysis of these products would likely yield a more informative dendrogram both in terms of the robustness of the results returned from the databases as well as the removal of the carbohydrate moieties as a confounding factor in the clustering of the metabolites.

With regard to the antimicrobial activity of PNW2016-02; antibacterial compounds with previously established efficacy against a variety of bacteria, including enterohemorrhagic *E. coli*, exist in the secretome of this fungal endophyte isolate. Accompanying these known antibiotics are other natural products that have bactericidal and/or bacteriostatic effects that further increase resistance to bacterial growth. Iron sequestration is also suggested as a notable role in the potential antimicrobial activity of PNW2016-02 via direct iron-nutrient competition. Future work to confirm both the ability and activity of the secretome compounds of PNW2016-02 may be carried out, for example, in spectrophotometric fashion involving ferrozine as described in Chakraborty et al. (2021). Time-kill assays involving suspected iron-chelating compounds and a variety of bacteria similar to those listed in Tables 1 and 2 could be used to identify the chelation potentials of individual compounds. The same methodology can be applied to the flavonoids and other phenylpropanoids derived from the secretome of PNW2016-02. Understanding in these areas will also contribute to development of the human pathogen suppressing model system using our endophytic *Stemphylium* strains *in plantae* within spinach, a known pathogen transmitting produce plants.

REFERENCES

Adjimani, J. P., & Emery, T. (1988). Stereochemical aspects of iron transport in *Mycelia sterilia* EP-76. *Journal of Bacteriology, 170*(3), 1377–1379. https://doi.org/10.1128/jb.170.3.1377-1379.1988

Aguirre, L., Vidal, A., Saminati, C., Tello, M., Redondo, N., Darwich, L., & Martin, M. (2020). Antimicrobial resistance profile and prevalence of extended-spectrum beta-lactamases (ESBL), AmpC beta-lactamases and colistin resistance (mcr) genes in *Escherichia coli* from swine between 1999 and 2018. *Porcine Health Management*, *6*(8), 1–6. https://doi.org/doi.org/10.1186/s40813-020-00146-2

Akbaba, M., & Ozaktan, H. (2018). Biocontrol of angular leaf spot disease and colonization of cucumber (*Cucumis sativus* L.) by endophytic bacteria. *Egyptian Journal of Biological Pest Control*, *28*(1), 1–10. https://doi.org/10.1186/S41938-017-0020-1

Akther, T., Ranjani, S., & Hemalatha, S. (2021). Nanoparticles engineered from endophytic fungi (*Botryosphaeria rhodina*) against ESBL-producing pathogenic multidrug-resistant *E. coli*. *Environmental Sciences Europe, 33*(1), 1–8. https://doi.org/10.1186/S12302-021-00524-9

Andersen, B., Solfrizzo, M., & Visconti, A. (1995). Metabolite profiles of common *Stemphylium* species. *Mycological research*, *6*, 672–676. https://doi.org/10.1016/S0953-7562(09)80526-1

Arnone, A., Nasini, G., Merlini, L., & Assante, G. (1986). Secondary mould metabolites. Part 16. Stemphyltoxins, new reduced perylenequinone metabolites from *Stemphylium botryosum* var. *lactucum*. *Journal of the Chemical Society, Perkin Transactions 1*, *0*, 525–530. https://doi.org/10.1039/P19860000525

Barber, M. S., McConnell, V. S., & Decaux, B. S. (2000). Antimicrobial intermediates of the general phenylpropanoid and lignin specific pathways. *Phytochemistry*, *54*(1), 53–56. https://doi.org/10.1016/S0031-9422(00)00038-8

Barbosa, F., Pinto, E., Kijjoa, A., Pinto, M., & Sousa, E. (2020). Targeting antimicrobial drug resistance with marine natural products. *International Journal of Antimicrobial Agents*, *56*(1), 106005. https://doi.org/10.1016/J.IJANTIMICAG.2020.106005

Barthélemy, M., Elie, N., Genta-Jouve, G., Stien, D., Touboul, D., & Eparvier, V. (2021). Identification of antagonistic compounds between the palm tree *Xylariale* endophytic fungi and the phytopathogen *Fusarium oxysporum*. *Journal of Agricultural and Food Chemistry*, *69*(37), 10893–10906. https://doi.org/10.1021/ACS.JAFC.1C03141

Behie, S. W., Moreira, C. C., Sementchoukova, I., Barelli, L., Zelisko, P. M., & Bidochka, M. J. (2017). Carbon translocation from a plant to an insect-pathogenic endophytic fungus. *Nature Communications 2017 8:1*, *8*(1), 1–5. https://doi.org/10.1038/ncomms14245

Belisario, A., Vitale, S., Luongo, L., Nardi, S., Talevi, S., & Corvi, F. (2008). First report of *Stemphylium vesicarium* as causal agent of wilting and root rotting of radish sprouts in Italy. *Plant Disease*, *92*(4), 651. https://doi.org/10.1094/PDIS-92-4-0651C

Bengtsson-Palme, J., Ryberg, M., Hartmann, M., Branco, S., Wang, Z., Godhe, A., De Wit, P., Sánchez-García, M., Ebersberger, I., de Sousa, F., Amend, A., Jumpponen, A., Unterseher, M., Kristiansson, E., Abarenkov, K., Bertrand, Y. J. K., Sanli, K., Eriksson, K. M., Vik, U., … Nilsson, R. H. (2013). Improved software detection and extraction of ITS1 and ITS2 from ribosomal ITS sequences of fungi and other eukaryotes for analysis of environmental sequencing data. *Methods in Ecology and Evolution*, *4*(10), 914–919. https://doi.org/10.1111/2041-210X.12073

Bhuyan, B. K., & Reusser, F. (1970). Comparative biological activity of nogalamycin and its analogs. *Cancer Research*, *30*(4).

Binns, D., Dimmer, E., Huntley, R., Barrell, D., O'Donovan, C., & Apweiler, R. (2009). QuickGO: a web-based tool for Gene Ontology searching. *Bioinformatics*, *25*(22), 3045–3046. https://doi.org/10.1093/BIOINFORMATICS/BTP536

Boessenkool, B. (2021). CRAN - Package berryFunctions. https://cran.r-project.org/web/packages/berryFunctions/index.html

Botella, L., & Javier Diez, J. (2011). Phylogenic diversity of fungal endophytes in Spanish stands of *Pinus halepensis*. *Fungal Diversity*, *47*, 9–18. https://doi.org/10.1007/S13225-010-0061-1

Caspi, R., Billington, R., Keseler, I., Kothari, A., Krummenacker, M., Midford, P., Kit Ong, W., Paley, S., Subhraveti, P., & Karp, P. (2020). The MetaCyc database of metabolic pathways and enzymes - a 2019 update. *Nucleic Acids Research*. 48, 445–453. https://doi.org/10.1093/nar.gkz862

Chahtane, H., Füller, T. N., Allard, P. M., Marcourt, L., Queiroz, E. F., Shanmugabalaji, V., Falquet, J., Wolfender, J. L., & Lopez-Molina, L. (2018). The plant pathogen *Pseudomonas aeruginosa* triggers a DELLA-dependent seed germination arrest in Arabidopsis. *ELife*, *7*. https://doi.org/10.7554/ELIFE.37082

Chakraborty, A., Majumdar, S., & Bhowal, J. (2021). Phytochemical screening and antioxidant and antimicrobial activities of crude extracts of different filamentous fungi. *Archives of Microbiology*, *203*(10), 6091–6108. https://doi.org/10.1007/S00203-021-02572-4/FIGURES/2

Chan, P. P., & Lowe, T. M. (2019). tRNAscan-SE: Searching for tRNA genes in genomic sequences. *Methods in Molecular Biology (Clifton, N.J.)*, *1962*, 1–14. https://doi.org/10.1007/978-1-4939-9173-0_1

Chaves, S., Canário, S., Carrasco, M. P., Mira, L., & Santos, M. A. (2012). Hydroxy(thio)pyrone and hydroxy(thio)pyridinone iron chelators: Physico-chemical properties and anti-oxidant activity. *Journal of Inorganic Biochemistry*, *114*, 38–46. https://doi.org/10.1016/J.JINORGBIO.2012.04.019

Chen, J., Ullah, C., Reichelt, M., Gershenzon, J., & Hammerbacher, A. (2019). S*clerotinia sclerotiorum* circumvents flavonoid defenses by catabolizing flavonol glycosides and aglycones. *Plant Physiology*, *180*(4), 1975–1987. https://doi.org/10.1104/PP.19.00461

Chen, L.-H., Yang, S. L., & Chung, K.-R. (n.d.). Resistance to oxidative stress via regulating siderophore-mediated iron acquisition by the citrus fungal pathogen *Alternaria alternata*. *Microbiology (Reading, England)*, *160*(5), 970-979. .https://doi.org/10.1099/mic.0.076182-0

Chistoserdova, L., Laukel, M., Portais, J. C., Vorholt, J. A., & Lidstrom, M. E. (2004). Multiple formate dehydrogenase enzymes in the facultative methylotroph *Methylobacterium extorquens* AM1 are dispensable for growth on methanol. *Journal of Bacteriology*, *186*(1), 22–28. https://doi.org/10.1128/JB.186.1.22-28.2004

Cushnie, T. P. T., & Lamb, A. J. (2005). Antimicrobial activity of flavonoids. *International Journal of Antimicrobial Agents*, *26*(5), 343–356. https://doi.org/10.1016/J.IJANTIMICAG.2005.09.002

De Silva, N. I., Brooks, S., Lumyong, S., & Hyde, K. D. (2019). Use of endophytes as biocontrol agents. *Fungal Biology Reviews*, *33*(2), 133–148. https://doi.org/10.1016/J.FBR.2018.10.001

de Vries, A., & Ripley, B. (2020). Create Dendrograms and Tree Diagrams Using “ggplot2” [R package ggdendro version 0.1.22]. https://cran.r-project.org/package=ggdendro

Debbab, A., Aly, A. H., Edrada-Ebel, R. A., Wray, V., Müller, W. E. G., Totzke, F., Zirrgiebel, U., Schächtele, C., Kubbutat, M. H. G., Wen, H. L., Mosaddak, M., Hakiki, A., Proksch, P., & Ebel, R. (2009). Bioactive metabolites from the endophytic fungus *Stemphylium globuliferum* isolated from *Mentha pulegium*. *Journal of Natural Products*, *72*(4), 626–631. https://doi.org/10.1021/NP8004997

Dickens, M. L., Priestley, N. D., & Strohl, W. R. (1997). In vivo and in vitro bioconversion of epsilon-rhodomycinone glycoside to doxorubicin: functions of DauP, DauK, and DoxA. *Journal of Bacteriology*, *179*(8), 2641. https://doi.org/10.1128/JB.179.8.2641-2650.1997

Dong, Y., Iniguez, A. L., Ahmer, B. M. M., & Triplett, E. W. (2003). Kinetics and strain specificity of rhizosphere and endophytic colonization by enteric bacteria on seedlings of *Medicago sativa* and *Medicago truncatula*. *Applied and Environmental Microbiology*, *69*(3), 1783–1790. https://doi.org/10.1128/AEM.69.3.1783-1790.2003

Dowle, M., & Srinivasan, A. (2021). Extension of “data.frame” [R package data.table version 1.14.2]. https://cran.r-project.org/package=data.table

Dubey, A., Malla, M. A., Kumar, A., Dayanandan, S., & Khan, M. L. (2020). Plants endophytes: unveiling hidden agenda for bioprospecting toward sustainable agriculture. *Critical Reviews in Biotechnology*, *40*(8), 1210–1231. https://doi.org/10.1080/07388551.2020.1808584

Dudeja, S. S., Suneja-Madan, P., Paul, M., Maheswari, R., & Kothe, E. (2021). Bacterial endophytes: Molecular interactions with their hosts. *Journal of Basic Microbiology*, *61*(6), 475–505. https://doi.org/10.1002/JOBM.202000657

Edgar, R. C. (2004). MUSCLE: A multiple sequence alignment method with reduced time and space complexity. *BioMed Central Bioinformatics*, *5*(1), 1–19. https://doi.org/10.1186/1471-2105-5-113/FIGURES/16

Emery, T., Neilands, J. B., & 83, V. (n.d.). Structure of the Ferrichrome Compounds 1,3. Retrieved October 14, 2021, from https://pubs.acs.org/sharingguidelines

Escandar, G. M., & Sala, L. F. (1991). Complexing behavior of rutin and quercetin. *Canadian Journal of Chemistry*, *69*(12), 1994–2001. https://doi.org/10.1139/V91-288

Fang, K., Zhou, J., Chen, L., Li, Y.-X., Yang, A.-L., Dong, X.-F., & Zhang, H.-B. (2021). Virulence and community dynamics of fungal species with vertical and horizontal transmission on a plant with multiple infections. *PLOS Pathogens*, *17*(7), e1009769. https://doi.org/10.1371/JOURNAL.PPAT.1009769

Fiołka, M. J., Zagaja, M. P., Piersiak, T. D., Wróbel, M., & Pawelec, J. (2010). Gut bacterium of *Dendrobaena veneta* (*Annelida*: *Oligochaeta*) possesses antimycobacterial activity. *Journal of Invertebrate Pathology* (Vol. 105, Issue 1, pp. 63–73). https://doi.org/10.1016/j.jip.2010.05.001

Fraly Erbabley, N. Y. G., & Junianto, J. (2020). Chemical characteristics and phytochemicals of the brown alga *Sargassum filipendulla* from kelanit waters of southeast Maluku. *Egyptian Journal of Aquatic Biology and Fisheries*, *24*(4), 535–547. https://doi.org/10.21608/EJABF.2020.105542

Franco, M. E. E., López, S., Medina, R., Saparrat, M. C. N., & Balatti, P. (2015). Draft genome sequence and gene annotation of *Stemphylium lycopersici* strain CIDEFI-216. *Genome Announcements*, *3*(5). https://doi.org/10.1128/GENOMEA.01069-15

Franz, C. M. A. P., Schillinger, U., & Holzapfel, W. H. (1996). Production and characterization of enterocin 900, a bacteriocin produced by *Enterococcus faecium* BFE 900 from black olives. *International Journal of Food Microbiology*, *29*(2–3), 255–270. https://doi.org/10.1016/0168-1605(95)00036-4

Galili, T. (2015). dendextend: an R package for visualizing, adjusting and comparing trees of hierarchical clustering. *Bioinformatics (Oxford, England)*, *31*(22), 3718–3720. https://doi.org/10.1093/BIOINFORMATICS/BTV428

Gao, Y. D., Zhao, Y., & Huang, J. (2014). Metabolic modeling of common *Escherichia coli* strains in human gut microbiome. *BioMed Research International*, *2014*. https://doi.org/10.1155/2014/694967

Gazzetti, K., Diaconu, E. L., Nanni, I. M., Ciriani, A., & Collina, M. (2019). Genome sequence resource for *Stemphylium vesicarium*, causing brown spot disease of pear. *Molecular Plant-Microbe Interactions*, *32*(8), 935–938. https://doi.org/10.1094/MPMI-11-18-0299-A/ASSET/IMAGES/LARGE/MPMI-11-18-0299-A_T1.JPEG

Ge, X., Sun, C., Feng, Y., Wang, L., Peng, J., Che, Q., Gu, Q., Zhu, T., Li, D., & Zhang, G. (2019). Anthraquinone derivatives from a marine-derived fungus *Sporendonema casei* HDN16-802. *Marine Drugs* 17, 334. https://doi.org/10.3390/md17060334.

Giri, V. (2017). Functions to compute chemical reaction similarity [R package RxnSim version 1.0.3]. https://cran.r-project.org/package=RxnSim

Golday, J. S. (2019). Progress toward a novel model system to investigate fungal endophytic suppression of human pathogens in spinach. https://doi.org/10.25394/PGS.8055995.V1

Greenshields, D. L., Liu, G., & Wei, Y. (2007). Roles of iron in plant defence and fungal virulence. *Plant Signaling & Behavior*, *2*(4), 300. https://doi.org/10.4161/PSB.2.4.4042

Greer, N. D. (2006). Tigecycline (Tygacil): the first in the glycylcycline class of antibiotics. *Proceedings (Baylor University. Medical Center)*, *19*(2), 155. https://doi.org/10.1080/08998280.2006.11928154

Grüning, B. A., Rasche, E., Rebolledo-Jaramillo, B., Eberhard, C., Houwaart, T., Chilton, J., Coraor, N., Backofen, R., Taylor, J., & Nekrutenko, A. (2017). Jupyter and Galaxy: Easing entry barriers into complex data analyses for biomedical researchers. *PLOS Computational Biology*, *13*(5), e1005425. https://doi.org/10.1371/JOURNAL.PCBI.1005425

Gurevich, A., Saveliev, V., Vyahhi, N., & Tesler, G. (2013). QUAST: quality assessment tool for genome assemblies. *Bioinformatics*, *29*(8), 1072–1075. https://doi.org/10.1093/BIOINFORMATICS/BTT086

Siebner-Freibach, H., Yariv, S., Lapides, Y., Hadar, Y. & Chen, Y.. (2005). Thermo-FTIR spectroscopic study of the siderophore ferrioxamine B: Spectral analysis and stereochemical implications of iron chelation, pH, and temperature. *Journal of Agricultural and Food Chemistry*, *53*(9), 3434–3443. https://doi.org/10.1021/JF048359K

Hastings, J., Owen, G., Dekker, A., Ennis, M., Kale, N., Muthukrishnan, V., Turner, S., Swainston, N., Mendes, P., & Steinbeck, C. (2016). ChEBI in 2016: Improved services and an expanding collection of metabolites. *Nucleic Acids Research*, *44*(D1), D1214–D1219. https://doi.org/10.1093/NAR/GKV1031

He, M., Wu, T., Pan, S., & Xu, X. (2014). Antimicrobial mechanism of flavonoids against *Escherichia coli* ATCC 25922 by model membrane study. *Applied Surface Science*, *305*, 515–521. https://doi.org/10.1016/J.APSUSC.2014.03.125

Henrissat, B., & Davies, G. (1997). Structural and sequence-based classification of glycoside hydrolases. *Current Opinion in Structural Biology*, *7*(5), 637–644. https://doi.org/10.1016/S0959-440X(97)80072-3

Herrmann, K. M., & Weaver, L. M. (1999). The shikimate pathway. *Annual Review of Plant Physiology and Plant Molecular Biology*, *50*, 473–503. https://doi.org/10.1146/ANNUREV.ARPLANT.50.1.473

Höller, U., Gloer, J. B., & Wicklow, D. T. (2002). Biologically active polyketide metabolites from an undetermined fungicolous hyphomycete resembling *Cladosporium*. *Journal of Natural Products*, *65*, 39. https://doi.org/10.1021/np020097y

Hollman, P. C. H., Bijsman, M. N. C. P., Gameren, Y. van, Cnossen, E. P. J., Vries, J. H. M. de, & Katan, M. B. (1999). The sugar moiety is a major determinant of the absorption of dietary flavonoid glycosides in man. *Free Radical Research*, *31*(6), 569–573. https://doi.org/10.1080/10715769900301141

Ibrahim, A., Tanney, J. B., Fei, F., Seifert, K. A., Cutler, G. C., Capretta, A., Miller, J. D., & Sumarah, M. W. (2020). Metabolomic-guided discovery of cyclic nonribosomal peptides from *Xylaria ellisii* sp. nov., a leaf and stem endophyte of *Vaccinium angustifolium*. *Scientific Reports*, *10*(1), 1–17. https://doi.org/10.1038/s41598-020-61088-x

Inomata, M., Hirai, N., Yoshida, R., & Ohigashi, H. (2004). The biosynthetic pathway to abscisic acid via ionylideneethane in the fungus *Botrytis cinerea*. *Phytochemistry*, *65*(19), 2667–2678. https://doi.org/10.1016/J.PHYTOCHEM.2004.08.025

Ito, J., Cox, E. C., & Yanofsky, C. (1969). Anthranilate synthetase, an enzyme specified by the tryptophan operon of *Escherichia coli*: purification and characterization of component I. *Journal of Bacteriology*, *97*(2), 725–733. https://doi.org/10.1128/JB.97.2.725-733.1969

Izquierdo-Bueno, I., González-Rodríguez, V. E., Simon, A., Dalmais, B., Pradier, J. M., Le Pêcheur, P., Mercier, A., Walker, A. S., Garrido, C., Collado, I. G., & Viaud, M. (2018). Biosynthesis of abscisic acid in fungi: identification of a sesquiterpene cyclase as the key enzyme in *Botrytis cinerea*. *Environmental Microbiology*, *20*(7), 2469–2482. https://doi.org/10.1111/1462-2920.14258

Jaques, A. K., Fukamizo, T., Hall, D., Barton, R. C., Escott, G. M., Parkinson, T., Hitchcock, C. A., & Adams, D. J. (2003). Disruption of the gene encoding the ChiB1 chitinase of *Aspergillus fumigatus* and characterization of a recombinant gene product. *Microbiology (Reading, England)*, *149*(Pt 10), 2931–2939. https://doi.org/10.1099/MIC.0.26476-0

Jassal, B., Matthews, L., Viteri, G., Gong, C., Lorente, P., Fabregat, A., Sidiropoulos, K., Cook, J., Gillespie, M., Haw, R., Loney, F., May, B., Milacic, M., Rothfels, K., Sevilla, C., Shamovsky, V., Shorser, S., Varusai, T., Weiser, J., … D'Eustachio, P. (2020). The reactome pathway knowledgebase. *Nucleic Acids Research*, *48*(D1), D498–D503. https://doi.org/10.1093/NAR/GKZ1031

Joseph, C. A., & Maroney, M. J. (2007). Cysteine dioxygenase: structure and mechanism. *Chemical Communications*, *32*, 3338–3349. https://doi.org/10.1039/B702158E

Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M., & Tanabe, M. (2016). KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Research*, *44*(D1), D457–D462. https://doi.org/10.1093/NAR/GKV1070

Karp, P. D., Ong, W. K., Paley, S., Billington, R., Caspi, R., Fulcher, C., Kothari, A., Krummenacker, M., Latendresse, M., Midford, P. E., Subhraveti, P., Gama-Castro, S., Muñiz-Rascado, L., Bonavides-Martinez, C., Santos-Zavaleta, A., Mackie, A., Collado-Vides, J., Keseler, I. M., & Paulsen, I. (2018). The EcoCyc Database. *EcoSal Plus*, *8*(1). https://doi.org/10.1128/ECOSALPLUS.ESP-0006-2018

Keilhofer, N., Nachtigall, J., Kulik, A., Ecke, M., Hampp, R., Süssmuth, R. D., Fiedler, H.-P., & Schrey, S. D. (2018). *Streptomyces* AcH 505 triggers production of a salicylic acid analogue in the fungal pathogen *Heterobasidion abietinum* that enhances infection of Norway spruce seedlings. *Antonie van Leeuwenhoek*, *111*(5), 691–704. https://doi.org/10.1007/S10482-018-1017-9

Keller, A., Hohlfeld, S., Kolter, A., Rg Schultz, J., Gemeinholzer, B., & Ankenbrand, M. J. (2020). BCdatabaser: on-the-fly reference database creation for (meta-)barcoding. *Bioinformatics*, *36*(8), 2630-2631. https://doi.org/10.1093/bioinformatics/btz960

Khokhani, D., Lowe-Power, T. M., Tran, T. M., & Allen, C. (2017). A single regulator mediates strategic switching between attachment/spread and growth/virulence in the plant pathogen *Ralstonia solanacearum*. *MBio*, *8*(5). https://doi.org/10.1128/MBIO.00895-17

Kim, S., Chen, J., Cheng, T., Gindulyte, A., He, J., He, S., Li, Q., Shoemaker, B. A., Thiessen, P. A., Yu, B., Zaslavsky, L., Zhang, J., & Bolton, E. E. (2021). PubChem in 2021: new data content and improved web interfaces. *Nucleic Acids Research*, *49*(D1), D1388–D1395. https://doi.org/10.1093/NAR/GKAA971

Kim, S., Thiessen, P. A., Cheng, T., Zhang, J., Gindulyte, A., & Bolton, E. E. (2019). Pug-View: Programmatic access to chemical annotations integrated in PubChem. *Journal of Cheminformatics*, *11*(1), 1–11. https://doi.org/10.1186/S13321-019-0375-2/FIGURES/5

KL, A., JE, W., & VJ, H. (2005). Persistence and differential survival of fecal indicator bacteria in subtropical waters and sediments. *Applied and Environmental Microbiology*, *71*(6), 3041–3048. https://doi.org/10.1128/AEM.71.6.3041-3048.2005

Kosman D. J. (2003). Molecular mechanisms of iron uptake in fungi. *Molecular Microbiology*, *47*(5), 1185–1197. https://doi.org/10.1046/j.1365-2958.2003.03368.x

Kõljalg, U., Nilsson, R. H., Abarenkov, K., Tedersoo, L., Taylor, A. F. S., Bahram, M., Bates, S. T., Bruns, T. D., Bengtsson-Palme, J., Callaghan, T. M., Douglas, B., Drenkhan, T., Eberhardt, U., Dueñas, M., Grebenc, T., Griffith, G. W., Hartmann, M., Kirk, P. M., Kohout, P., … Larsson, K. H. (2013). Towards a unified paradigm for sequence-based identification of fungi. *Molecular Ecology*, *22*(21), 5271–5277. https://doi.org/10.1111/MEC.12481

Korf, I. (2004). Gene finding in novel genomes. *BioMed Central Bioinformatics*, *5*(1), 1–9. https://doi.org/10.1186/1471-2105-5-59/TABLES/4

Kousser, C., Clark, C., Sherrington, S., Voelz, K., & Hall, R. A. (2019). *Pseudomonas aeruginosa* inhibits *Rhizopus microsporus* germination through sequestration of free environmental iron. *Scientific Reports*, *9*(1), 1–14. https://doi.org/10.1038/s41598-019-42175-0

Křen, V., & Řezanka, T. (2008). Sweet antibiotics - The role of glycosidic residues in antibiotic and antitumor activity and their randomization. *FEMS Microbiology Reviews*, *32*(5), 858–889. https://doi.org/10.1111/j.1574-6976.2008.00124.x

Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution*, *35*(6), 1547–1549. https://doi.org/10.1093/MOLBEV/MSY096

Künzler, M. (2018). How fungi defend themselves against microbial competitors and animal predators. *PLoS Pathogens*, *14*(9). https://doi.org/10.1371/journal.ppat.1007184

Lacroix, S., Klicic Badoux, J., Scott-Boyer, M. P., Parolo, S., Matone, A., Priami, C., Morine, M. J., Kaput, J., & Moco, S. (2018). A computationally driven analysis of the polyphenol-protein interactome. *Scientific Reports*, *8*(1). https://doi.org/10.1038/s41598-018-20625-5

Le Roy, J., Huss, B., Creach, A., Hawkins, S., & Neutelings, G. (2016). Glycosylation is a major regulator of phenylpropanoid availability and biological activity in plants. *Frontiers in Plant Science*, *7*(5), 735. https://doi.org/10.3389/FPLS.2016.00735/BIBTEX

Lesuisse, E., Simon-Casteras, M., & Labbe, P. (1998). Siderophore-mediated iron uptake in *Saccharomyces cerevisiae*: the SIT1 gene encodes a ferrioxamine B permease that belongs to the major facilitator superfamily. *Microbiology (Reading, England)*, *144*(12), 3455–3462. https://doi.org/10.1099/00221287-144-12-3455

Liarzi, O., Bar, E., Lewinsohn, E., & Ezra, D. (2016). Use of the endophytic fungus *Daldinia* cf. *concentrica* and its volatiles as bio-control agents. *PLOS ONE*, *11*(12). https://doi.org/10.1371/JOURNAL.PONE.0168242

Lin, X., Xu, S., Yang, Y., Wu, J., Wang, H., Shen, H., & Wang, H. (2009). Purification and characterization of anthranilate synthase component I (TrpE) from *Mycobacterium tuberculosis* H37Rv. *Protein Expression and Purification*, *64*(1), 8–15. https://doi.org/10.1016/J.PEP.2008.09.020

Liu, B., Stein, L., Cochran, K., du Toit, L., Feng, C., Dhillon, N., & Correll, J. (2020). Characterization of leaf spot pathogens from several spinach production areas in the United States. *Plant Disease*, *104*(7), 1994–2004. https://doi.org/10.1094/PDIS-11-19-2450-RE

Lou, J., Fu, L., Peng, Y., & Zhou, L. (2013). Metabolites from *Alternaria* fungi and their bioactivities. *Molecules*, *18*, 5891–5935. https://doi.org/10.3390/molecules18055891

Lu, Y., Ye, C., Che, J., Xu, X., Shao, D., Jiang, C., Liu, Y., & Shi, J. (2019). Genomic sequencing, genome-scale metabolic network reconstruction, and in silico flux analysis of the grape endophytic fungus *Alternaria* sp. MG1. *Microbial Cell Factories*, *18*(1), 1–16. https://doi.org/10.1186/S12934-019-1063-7/TABLES/4

Luo, J., Tao, Q., Jupa, R., Liu, Y., Wu, K., Song, Y., Li, J., Huang, Y., Zou, L., Liang, Y., & Li, T. (2019). Role of vertical transmission of shoot endophytes in root-associated microbiome assembly and heavy metal hyperaccumulation in *Sedum alfredii*. *Environmental Science & Technology*, *53*(12), 6954–6963. https://doi.org/10.1021/ACS.EST.9B01093

M, P., L, R., E, M., C, M., & I, L. (2015). Combined morphological and molecular approach for identification of *Stemphylium vesicarium* inoculum in pear orchards. *Fungal Biology*, *119*(2–3), 136–144. https://doi.org/10.1016/J.FUNBIO.2014.11.006

Maheshwari, R., Bhutani, N., & Suneja, P. (2019). Screening and characterization of siderophore producing endophytic bacteria from *Cicer arietinum* and *Pisum sativum. Journal of Applied Biology & Biotechnology*, *7*(05), 7-14. https://doi.org/10.7324/JABB.2019.70502

Majoros, W. H., Pertea, M., & Salzberg, S. L. (2004). TigrScan and GlimmerHMM: two open source ab initio eukaryotic gene-finders. *Bioinformatics (Oxford, England)*, *20*(16), 2878–2879. https://doi.org/10.1093/BIOINFORMATICS/BTH315

Manias, D., Verma, A., & Soni, D. K. (2020). Isolation and characterization of endophytes: Biochemical and molecular approach. *Microbial Endophytes: Prospects for Sustainable Agriculture*, 1–14. https://doi.org/10.1016/B978-0-12-818734-0.00001-2

Manulis, S., Kashman, Y, Netzer, D. & Barash, L (1984). Phytotoxins from *Stemphylium botryosum*: structural determination of stemphyloxin II, productIon in culture and interaction with iron. *Phytochemistry* 23, 2193–2198.

Mattoo, A. J., & Nonzom, S. (2021). Endophytic fungi: understanding complex cross-talks. *Symbiosis*, *83*(3), 237–264. https://doi.org/10.1007/S13199-020-00744-2

McNaught, A. D. (1996). Nomenclature of carbohydrates (IUPAC Recommendations 1996). *Pure and Applied Chemistry*, *68*(10), 1919–2008. http://dx.doi.org/10.1351/pac199668101919*8*.

Medina, R., Franco, M. E. E., da Cruz Cabral, L., Bahima, J. V., Patriarca, A., Balatti, P. A., & Saparrat, M. C. N. (2021). The secondary metabolites profile of *Stemphylium lycopersici*, the causal agent of tomato grey leaf spot, is complex and includes host and non-host specific toxins. *Australasian Plant Pathology*, *50*(1), 105–115. https://doi.org/10.1007/S13313-020-00753-1/TABLES/3

Medina, R., Franco, M. E. E., Lucentini, G., Saparrat, M. C. N., & Balatti, P. A. (2018). Draft Genome Sequences of Sporulating (CIDEFI-213) and Nonsporulating (CIDEFI-212) Strains of *Stemphylium lycopersici*. *Microbiology Resource Announcements*, *7*(6). https://doi.org/10.1128/MRA.00960-18

Metsä-Ketelä, M., Niemi, J., Mäntsälä, P., & Schneider, G. (2007). Anthracycline Biosynthesis: Genes, Enzymes and Mechanisms. *Topics in Current Chemistry*, *282*, 101–140. https://doi.org/10.1007/128_2007_14

Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E. M., Kubal, M., Paczian, T., Rodriguez, A., Stevens, R., Wilke, A., Wilkening, J., & Edwards, R. A. (2008). The metagenomics RAST server - A public resource for the automatic phylogenetic and functional analysis of metagenomes. *BioMed Central Bioinformatics*, *9*(1), 1–8. https://doi.org/10.1186/1471-2105-9-386/FIGURES/4

Miethke, M., & Marahiel, M. A. (2007). Siderophore-based iron acquisition and pathogen control. *Microbiology and Molecular Biology Reviews : MMBR*, *71*(3), 413. https://doi.org/10.1128/MMBR.00012-07

Minerdi, D., Maggini, V., & Fani, R. (2021). Volatile organic compounds: from figurants to leading actors in fungal symbiosis. *FEMS Microbiology Ecology*, *97*(5). https://doi.org/10.1093/FEMSEC/FIAB067

Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G. A., Sonnhammer, E. L. L., Tosatto, S. C. E., Paladin, L., Raj, S., Richardson, L. J., Finn, R. D., & Bateman, A. (2021). Pfam: The protein families database in 2021. *Nucleic Acids Research*, *49*(D1), D412–D419. https://doi.org/10.1093/NAR/GKAA913

Mohanta, T. K. (2020). Fungi contain genes associated with flavonoid biosynthesis pathway. *Journal of Functional Foods*, *68*, 103910. https://doi.org/10.1016/J.JFF.2020.103910

Mrudulakumari Vasudevan, U., & Lee, E. Y. (2020). Flavonoids, terpenoids, and polyketide antibiotics: Role of glycosylation and biocatalytic tactics in engineering glycosylation. *Biotechnology Advances*, *41*, 107550. https://doi.org/10.1016/J.BIOTECHADV.2020.107550

Musialik, M., Kuzmicz, R., Pawłowski, T. S., & Litwinienko, G. (2009). Acidity of hydroxyl groups: An overlooked influence on antiradical properties of flavonoids. *Journal of Organic Chemistry*, *74*(7), 2699–2709. https://doi.org/10.1021/JO802716V

Muthubharathi, B. C., Gowripriya, T., & Balamurugan, K. (2021). Metabolomics: small molecules that matter more. *Molecular Omics*, *17*(2), 210–229. https://doi.org/10.1039/D0MO00176G

Nachtweide, S., & Stanke, M. (2019). Multi-genome annotation with AUGUSTUS. *Methods in Molecular Biology, 1962*, 139–160. https://doi.org/10.1007/978-1-4939-9173-0_8

Nakkarach, A., Foo, H. L., Song, A. A.-L., Mutalib, N. E. A., Nitisinprasert, S., & Withayagiat, U. (2021). Anti-cancer and anti-inflammatory effects elicited by short chain fatty acids produced by *Escherichia coli* isolated from healthy human gut microbiota. *Microbial Cell Factories*, *20*(1), 1–17. https://doi.org/10.1186/S12934-020-01477-Z

Nandakumar, M., Malathi, P., Sundar, A. R., & Viswanathan, R. (2021). Host-pathogen interaction in sugarcane and red rot pathogen: exploring expression of phytoalexin biosynthesis pathway genes. *Indian Phytopathology*, *74*(2), 529–535. https://doi.org/10.1007/S42360-020-00306-Y/FIGURES/3

Niccolai, A., Chini Zittelli, G., Rodolfi, L., Biondi, N., & Tredici, M. R. (2019). Microalgae of interest as food source: Biochemical composition and digestibility. *Algal Research 42*(9). https://doi.org/10.1016/j.algal.2019.101617

Nozawa, S., Horii, N., Ishizaki, T., & Watanabe, K. (2016). Occurrence of *Stemphylium* leaf blight caused by *Stemphylium lycopersici* and gray mold caused by *Botrytis cinerea* on *Nemesia caerulea* in Japan. *Japanese Journal of Phytopathology*, *82*(1), 19–23. https://doi.org/10.3186/JJPHYTOPATH.82.19

Ogbe, A. A., Finnie, J. F., & Van Staden, J. (2020). The role of endophytes in secondary metabolites accumulation in medicinal plants under abiotic stress. *South African Journal of Botany*, *134*, 126–134. https://doi.org/10.1016/J.SAJB.2020.06.023

Ouellette, B. F. F., & Boguski, M. S. (1997). Database divisions and homology search files: A guide for the perplexed. *Genome Research*, *7*(10), 952–955. https://doi.org/10.1101/GR.7.10.952

Overkamp, K. M., Ktter, P., Van Hoek, R. Der, Schoondermark-Stolk, S., Luttik, M. A. H., Van Dijken, J. P., & Pronk, J. T. (2002). Functional analysis of structural genes for NAD(+)-dependent formate dehydrogenase in *Saccharomyces cerevisiae*. *Yeast (Chichester, England)*, *19*(6), 509–520. https://doi.org/10.1002/YEA.856

Marder, E., Garman, K., Ingram, L., & Dunn, J. (2014). Multistate outbreak of *Escherichia coli* O157:H7 associated with bagged salad. *Foodborne Pathogens and Disease*, *11*(8), 593–595. https://doi.org/10.1089/FPD.2013.1726

Palmer, J. M., & Stajich, J. (2020). Funannotate v1.8.1: Eukaryotic genome annotation. https://doi.org/10.5281/ZENODO.4054262

Pang, Z., Chong, J., Zhou, G., de Lima Morais, D. A., Chang, L., Barrette, M., Gauthier, C., Jacques, P.-É., Li, S., & Xia, J. (2021). MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. *Nucleic Acids Research*, *49*(W1), W388–W396. https://doi.org/10.1093/NAR/GKAB382

Pedras, S., & Ahiahonu, P. (2005). Metabolism and detoxification of phytoalexins and analogs by phytopathogenic fungi. *Phytochemistry*, *66*, 391–411. https://doi.org/10.1016/j.phytochem.2004.12.032

Prevention, C. for D. C. and. (2021). Reports of Selected *E. coli* Outbreak Investigations. 2021. https://www.cdc.gov/ecoli/outbreaks.html

R Core Team. (2017). *R:* A language and environment for statistical computing. R Foundation for Statistical Computing.

Razzak, M. A., Mohammad, H., Islam, M. A., Rahman, M., Sathi, M. A., Hasan, M. M., Mudassir, E., & Nasir, A. (2018). Screening Of lentil germplasm against *Stemphylium* blight by observing disease reaction in three different stages. *Malaysian Journal of Halal Research (MJHR)*, *1*(2), 15–18. https://doi.org/10.26480/mjhr.02.2018.15.18

Ren, K. (2021, September 3). CRAN - Package rlist. https://cran.r-project.org/web/packages/rlist/index.html

Richards, T. A., Dacks, J. B., Campbell, S. A., Blanchard, J. L., Foster, P. G., McLeod, R., & Roberts, C. W. (2006). Evolutionary origins of the eukaryotic shikimate pathway: Gene fusions, horizontal gene transfer, and endosymbiotic replacements. *Eukaryotic Cell*, *5*(9), 1517–1531. https://doi.org/10.1128/EC.00106-06

Richardson, L., & Amundsen, M. (2013). RESTful Web APIs.

Romero, F. M., Rossi, F. R., Gárriz, A., Carrasco, P., & Ruíz, O. A. (2019). A bacterial endophyte from apoplast fluids protects canola plants from different phytopathogens via antibiosis and induction of host resistance. *Phytopathology*, *109*(3), 375–383. https://doi.org/10.1094/PHYTO-07-18-0262-R

Salloum, M. S., Menduni, M. F., Benavides, M. P., Larrauri, M., Luna, C. M., & Silvente, S. (2018). Polyamines and flavonoids: key compounds in mycorrhizal colonization of improved and unimproved soybean genotypes. *Symbiosis*, *76*(3), 265–275. https://doi.org/10.1007/S13199-018-0558-Z

Santra, H. K., & Banerjee, D. (2020). Natural products as fungicide and their role in crop protection. *Natural Bioactive Products in Sustainable Agriculture*, 131. https://doi.org/10.1007/978-981-15-3024-1_9

Sarsaiya, S., Shi, J., & Chen, J. (2019). A comprehensive review on fungal endophytes and its dynamics on *Orchidaceae* plants: current research, challenges, and future possibilities. *Bioengineered*, *10*(1), 316–334. https://doi.org/10.1080/21655979.2019.1644854

Sass, G., Ansari, S. R., Dietl, A. M., Déziel, E., Haas, H., & Stevens, D. A. (2019). Intermicrobial interaction: *Aspergillus fumigatus* siderophores protect against competition by *Pseudomonas aeruginosa*. *PLOS ONE*, *14*(5), e0216085. https://doi.org/10.1371/JOURNAL.PONE.0216085

Seshime, Y., Praveen, R. J., Fujii, I., & Kitamoto, K. (2005). Genomic evidences for the existence of a phenylpropanoid metabolic pathway in *Aspergillus oryzae*. *Biochemical and Biophysical Research Communications*, *337*(3), 747–751. https://doi.org/10.1016/J.BBRC.2005.08.233

Sharapov, U. M., Wendel, A. M., Davis, J. P., Keene, W. E., Farrar, J., Sodha, S., Hyytia-Trees, E., Leeper, M., Gerner-Smidt, P., Griffin, P. M., BRADEN, C. (2016). Multistate outbreak of *Escherichia coli* O157:H7 infections associated with consumption of fresh spinach: United States, 2006. *Journal of Food Protection*, *79*(12), 2024–2030. https://doi.org/10.4315/0362-028X.JFP-15-556

Sharma, S., Hay, F. S., & Pethybridge, S. J. (2020). Genome resource for two *Stemphylium vesicarium* isolates causing *Stemphylium* leaf blight of onion in New York. *Molecular Plant-Microbe Interactions*, *33*(4), 562–564. https://doi.org/10.1094/MPMI-08-19-0244-A/ASSET/IMAGES/LARGE/MPMI-08-19-0244-A_T1.JPEG

Shaw, R. K., Berger, C. N., Feys, B., Knutton, S., Pallen, M. J., & Frankel, G. (2008). Enterohemorrhagic *Escherichia coli* exploits EspA filaments for attachment to salad leaves. *Applied and Environmental Microbiology*, *74*(9), 2908–2914. https://doi.org/10.1128/AEM.02704-07

Siedler, S., Balti, R., & Neves, A. R. (2019). Bioprotective mechanisms of lactic acid bacteria against fungal spoilage of food. *Current Opinion in Biotechnology*, *56*, 138–146. https://doi.org/10.1016/J.COPBIO.2018.11.015

Smith, G. (1938). Cryptogamic Botany Vol I. https://archive.org/details/cryptogamicbotan031880mbp/page/n7/mode/2up

Sopovski, D. (2019). Antimicrobial Resistance in *Serratia marcescens*. https://doi.org/10.25394/PGS.8059304.V1

Sordon, S., Popłoński, J., Tronina, T., & Huszcza, E. (2019). Regioselective O-glycosylation of flavonoids by fungi *Beauveria bassiana*, *Absidia coerulea* and *Absidia glauca*. *Bioorganic Chemistry*, *93*. https://doi.org/10.1016/j.bioorg.2019.01.046

Spawton, K. A., McGrath, M., & Toit, L. J. du. (2020). First report of *Stemphylium* leaf spot of spinach (*Spinacia oleracea*) caused by *Stemphylium beticola* in New York State. *Plant Disease*, *104*(11), 3068. https://doi.org/10.1094/PDIS-02-20-0343-PDN

Stinson, M., Ezra, D., Hess, W. M., Sears, J., & Strobel, G. (2003). An endophytic *Gliocladium* sp. of *Eucryphia cordifolia* producing selective volatile antimicrobial compounds. *Plant Science*, *165*(4), 913–922. https://doi.org/10.1016/S0168-9452(03)00299-1

Strobel, G. A. (2003). Endophytes as sources of bioactive products. *Microbes and Infection*, *5*(6), 535–544. https://doi.org/10.1016/S1286-4579(03)00073-X

Suebrasri, T., Harada, H., Jogloy, S., Ekprasert, J., & Boonlue, S. (2020). Auxin-producing fungal endophytes promote growth of sunchoke. *Rhizosphere*, *16*, 100271. https://doi.org/10.1016/J.RHISPH.2020.100271

Tang, Z., Wang, Y., Yang, J., Xiao, Y., Cai, Y., Wan, Y., Chen, H., Yao, H., Shan, Z., Li, C., & Wang, G. (2020). Isolation and identification of flavonoid-producing endophytic fungi from medicinal plant *Conyza blinii* H.Lév that exhibit higher antioxidant and antibacterial activities. *PeerJ*, *2020*(4), e8978. https://doi.org/10.7717/PEERJ.8978/SUPP-3

Tenaillon, O., Skurnik, D., Picard, B., & Denamur, E. (2010). The population genetics of commensal *Escherichia coli*. *Nature Reviews. Microbiology*, *8*(3), 207–217. https://doi.org/10.1038/NRMICRO2298

Ter-Hovhannisyan, V., Lomsadze, A., Chernoff, Y. O., & Borodovsky, M. (2008). Gene prediction in novel fungal genomes using an ab initio algorithm with unsupervised training. *Genome Research*, *18*(12), 1979–1990. https://doi.org/10.1101/GR.081612.108

Thines, E., Aguirre, J., Foster, A. J., & Deising, H. B. (2006). Genetics of phytopathology: Secondary metabolites as virulence determinants of fungal plant pathogens. *Botany*, *67*, 134–161. https://doi.org/10.1007/3-540-27998-9_6

Tu, B. P., & Weissman, J. S. (2002). The FAD- and O2-dependent reaction cycle of Ero1-mediated oxidative protein folding in the endoplasmic reticulum. *Molecular Cell*, *10*(5), 983–994. https://doi.org/10.1016/S1097-2765(02)00696-2

Ul Haq, I., & Ijaz, S. (2020). History and Recent Trends in Plant Disease Control: An Overview. 1–13. https://doi.org/10.1007/978-3-030-35955-3_1

Křen, V., & Řezanka, T. (2008). Sweet antibiotics - the role of glycosidic residues in antibiotic and antitumor activity and their randomization. *FEMS Microbiology Reviews*, *32*(5), 858–889. https://doi.org/10.1111/J.1574-6976.2008.00124.X

van Overbeek, L. S., & Saikkonen, K. (2016). Impact of bacterial–fungal interactions on the colonization of the endosphere. *Trends in Plant Science*, *21*(3), 230–242. https://doi.org/10.1016/J.TPLANTS.2016.01.003

Vasconcelos, F. N. C., Li, J., Pang, Z., Vincent, C., & Wang, N. (2021). The total population size of '*Candidatus Liberibacter asiaticus*' inside the phloem of citrus trees and the corresponding metabolic burden related to Huanglongbing disease development. *Phytopathology*, *111*(7), 1122-1128. https://doi.org/10.1094/PHYTO-09-20-0388-R

Vicente, J. G., & Holub, E. B. (2013). *Xanthomonas campestris* pv. *campestris* (cause of black rot of crucifers) in the genomic era is still a worldwide threat to brassica crops. *Molecular Plant Pathology*, *14*(1), 2–18. https://doi.org/10.1111/J.1364-3703.2012.00833.X

Volkert, M. R., Elliott, N. A., & Housman, D. E. (2000). Functional genomics reveals a family of eukaryotic oxidation protection genes. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(26), 14530–14535. https://doi.org/10.1073/PNAS.260495897

Vu, D., Groenewald, M., de Vries, M., Gehrmann, T., Stielow, B., Eberhardt, U., Al-Hatmi, A., Groenewald, J. Z., Cardinali, G., Houbraken, J., Boekhout, T., Crous, P. W., Robert, V., & Verkley, G. J. M. (2019). Large-scale generation and analysis of filamentous fungal DNA barcodes boosts coverage for kingdom fungi and reveals thresholds for fungal species and higher taxon delimitation. *Studies in Mycology*, *92*, 135–154. https://doi.org/10.1016/J.SIMYCO.2018.05.001

Wallroth, K. F., Bluff, M. J., & Fingerhuth, C. (1833). Flora cryptogamica Germaniae: Algas et fungos. In *Compendium Florae Germanicae*. https://books.google.com/books?hl=en&lr=&id=22s-AAAAcAAJ&oi=fnd&pg=PR2&dq=Flora+cryptogamica+Germaniae+&ots=aOwLJ86Rwk&sig=xXlinC25LVwaePsuJ12x4FJVpnY#v=onepage&q=Flora cryptogamica Germaniae&f=false

Wang, J.-F., Liu, S.-S., Song, Z.-Q., Xu, T.-C., Liu, C.-S., Hou, Y.-G., Huang, R., & Wu, S.-H. (2020). Naturally occurring flavonoids and isoflavonoids and their microbial transformation: A review. *Molecules*, *25*(21), 5112. https://doi.org/10.3390/molecules25215112

Wang, L., Dash, S., Ng, C. Y., & Maranas, C. D. (2017). A review of computational tools for design and reconstruction of metabolic pathways. *Synthetic and Systems Biotechnology*, *2*(4), 243–252. https://doi.org/10.1016/J.SYNBIO.2017.11.002

Wang, M. C.-H., Tsai, M. Y. C., Tsai, M. I., Chung, D. C.-L., Lin, M. Y.-C., Hsieh, M. W. T., Hung, D. T.-H., Suwannarach, D. N., Cheewangkoon, D. R., Elgorban, D. A., & Ariyawansa, P. H. A. (2021). Stemphylium leaf blight of Welsh onion (Allium fistulosum): An emerging disease in Sanxing, Taiwan. Plant Disease. https://doi.org/10.1094/PDIS-11-20-2329-RE

Wang, S.-S., Liu, J.-M., Sun, J., Sun, Y.-F., Liu, J.-N., Jia, N., Fan, B., & Dai, X.-F. (2019). Diversity of culture-independent bacteria and antimicrobial activity of culturable endophytic bacteria isolated from different *Dendrobium* stems. *Scientific Reports*, *9*(1), 1–12. https://doi.org/10.1038/s41598-019-46863-9

Wani, Z. A., Ashraf, N., Mohiuddin, T., & Riyaz-Ul-Hassan, S. (2015). Plant-endophyte symbiosis, an ecological perspective. *Applied Microbiology and Biotechnology*, *99*(7), 2955–2965. https://doi.org/10.1007/S00253-015-6487-3

Warnes, G., Bolker, B., & Lumley, T. (2021). CRAN - Package gtools. https://cran.r-project.org/web/packages/gtools/index.html

West, J. S., Townsend, J. A., Stevens, M., & Fitt, B. D. L. (2012). Comparative biology of different plant pathogens to estimate effects of climate change on crop diseases in Europe. *European Journal of Plant Pathology*, *133*(1), 315–331. https://doi.org/10.1007/S10658-011-9932-X

Westrick, N. M., Smith, D. L., & Kabbage, M. (2021). Disarming the Host: Detoxification of Plant Defense Compounds During Fungal Necrotrophy. *Frontiers in Plant Science*, *12*. https://doi.org/10.3389/fpls.2021.651716

Wickham, H. (2020). Tools for working with URLs and HTTP [R package httr version 1.4.2]. https://cran.r-project.org/package=httr

Wickham, H. (2020). CRAN - Package reshape2. https://cran.r-project.org/web/packages/reshape2/index.html

Wickham, H. (2018). Flexibly reshape data [R package reshape version 0.8.8]. https://cran.r-project.org/package=reshape

Wickham, H. (2019). Simple, consistent wrappers for common string operations [R package stringr version 1.4.0]. https://cran.r-project.org/package=stringr

Wickham, H. (2020). Tools for splitting, applying and combining data [R package plyr version 1.8.6]. https://cran.r-project.org/package=plyr

Wickham, H. (2016). ggplot2: Elegant graphics for data analysis. *Springer-Verlag (New York, New York)*.

Wickham, H. (2021). Easily harvest (scrape) web pages [R package rvest version 1.0.2]. Comprehensive R Archive Network (CRAN). https://cran.r-project.org/package=rvest

Wickham, H., François, R., Henry, L., & Müller, K. (2021). A grammar of data manipulation [R package dplyr version 1.0.7]. https://cran.r-project.org/package=dplyr

Wickham, H., Hester, J., & François, R. (2021). Read rectangular text data [R package readr version 2.1.0]. https://cran.r-project.org/package=readr

Wickham, H., Hester, J., & Ooms, J. (2020). Parse XML [R package xml2 version 1.3.2]. Comprehensive R Archive Network (CRAN). https://cran.r-project.org/package=xml2

Wishart, D. S., Feunang, Y. D., Marcu, A., Guo, A. C., Liang, K., Vázquez-Fresno, R., Sajed, T., Johnson, D., Li, C., Karu, N., Sayeeda, Z., Lo, E., Assempour, N., Berjanskii, M., Singhal, S., Arndt, D., Liang, Y., Badran, H., Grant, J., … Scalbert, A. (2018). HMDB 4.0: the human metabolome database for 2018. *Nucleic Acids Research*, *46*(D1), D608–D617. https://doi.org/10.1093/NAR/GKX1089

Woody, S. K., Burdick, D., Lapp, H., & Huang, E. S. (2020). Application programming interfaces for knowledge transfer and generation in the life sciences and healthcare. *Nature Portfolio Journal Digital Medicine*, *3*(1), 1–5. https://doi.org/10.1038/s41746-020-0235-5

Woudenberg, J. H. C., Hanse, B., van Leeuwen, G. C. M., Groenewald, J. Z., & Crous, P. W. (2017). *Stemphylium* revisited. *Studies in Mycology*, *87*, 77. https://doi.org/10.1016/J.SIMYCO.2017.06.001

Wratten, L., Wilm, A., & Göke, J. (2021). Reproducible, scalable, and shareable analysis pipelines with bioinformatics workflow managers. *Nature Methods*, *18*(10), 1161–1168. https://doi.org/10.1038/s41592-021-01254-9

Xiao, Y., Gerth, K., Müller, R., & Wall, D. (2012). *Myxobacterium*-produced antibiotic TA (myxovirescin) inhibits type II signal peptidase. *Antimicrobial Agents and Chemotherapy*, *56*(4), 2014–2021. https://doi.org/10.1128/AAC.06148-11

Xie, L., Zhang, L., Wang, C., Wang, X., Xu, Y. ming, Yu, H., Wu, P., Li, S., Han, L., Gunatilaka, A. A. L., Wei, X., Lin, M., Molnár, I., & Xu, Y. (2018). Methylglucosylation of aromatic amino and phenolic moieties of drug-like biosynthons by combinatorial biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(22), E4980–E4989. https://doi.org/10.1073/pnas.1716046115

Yan, L. (2021). Draw Venn diagram by "ggplot2" [R package ggvenn version 0.1.9]. https://cran.r-project.org/package=ggvenn

Yan, W., Cao, L.-L., Zhang, Y.-Y., Zhao, R., Zhao, S.-S., Khan, B., & Ye, Y.-H. (2018). New metabolites from endophytic fungus *Chaetomium globosum* CDW7. *Molecules*, *23*(11), 2873. https://doi.org/10.3390/MOLECULES23112873

Yarwood, S. A. (2018). The role of wetland microorganisms in plant-litter decomposition and soil organic matter formation: a critical review. *FEMS Microbiology Ecology*, *94*, 175. https://doi.org/10.1093/femsec/fiy175

Yin, H., You, L., Pasqualone, D., Kopski, K. M., & Huffaker, T. C. (2002). Stu1p is physically associated with beta-tubulin and is required for structural integrity of the mitotic spindle. *Molecular Biology of the Cell*, *13*(6), 1881–1892. https://doi.org/10.1091/MBC.01-09-0458

Yonekura-Sakakibara, K., Higashi, Y., & Nakabayashi, R. (2019). The origin and evolution of plant flavonoid metabolism. *Frontiers in Plant Science*, *10*, 943. https://doi.org/10.3389/FPLS.2019.00943

Yuan, M. M., Kakouridis, A., Starr, E., Nguyen, N., Shi, S., Zhou, J., Firestone, M., Pett-Ridge, J., & Nuccio, E. (2021). Fungal-bacterial cooccurrence patterns differ between arbuscular mycorrhizal fungi and nonmycorrhizal fungi across soil niches. *MBio*, *12*(2). https://doi.org/10.1128/MBIO.03509-20

Zeng, J., Xu, T., Cao, L., Tong, C., Zhang, X., Luo, D., Han, S., Pang, P., Fu, W., Yan, J., Liu, X., & Zhu, Y. (2018). The role of iron competition in the antagonistic action of the rice endophyte *Streptomyces sporocinereus* OsiSh-2 against the pathogen *Magnaporthe oryzae*. *Microbial Ecology*, *76*(4), 1021–1029. https://doi.org/10.1007/S00248-018-1189-X/FIGURES/5

Zeng, Z., Qian, L., Cao, L., Tan, H., Huang, Y., Xue, X., Shen, Y., & Zhou, S. (2008). Virtual screening for novel quorum sensing inhibitors to eradicate biofilm formation of *Pseudomonas aeruginosa*. *Applied Microbiology and Biotechnology*, *79*(1), 119–126. https://doi.org/10.1007/S00253-008-1406-5/FIGURES/8

Zhang, Q., Xue, X.-Z., Miao, S.-M., Cui, J.-L., & Qin, X.-M. (n.d.). Differential relationship of fungal endophytic communities and metabolic profiling in the stems and roots of *Ephedra sinica* based on metagenomics and metabolomics. *Symbiosis*, *81*(2), 115-125. https://doi.org/10.1007/s13199-020-00685-w

Zhang, X., Abugri, D. A., & Witola, W. H. (2017). Activity of green algae extracts against *Toxoplasma gondii*: mushrooms as novel sources for new parasitic drug development. *Medicinal & Aromatic Plants*, *6*(3). https://doi.org/10.4172/2167-0412.1000293

Zhang, Y., Yu, X., Zhang, W., Lang, D., Zhang, X., Cui, G., & Zhang, X. (2019). Interactions between endophytes and plants: Beneficial effect of endophytes to ameliorate biotic and abiotic stresses in plants. *Journal of Plant Biology*, *62*, 1–13. https://doi.org/10.1007/s12374-018-0274-5

Zhou, X.-M., Zheng, C.-J., Chen, G.-Y., Song, X.-P., Han, C.-R., Tang, X.-Z., Liu, R.-J., & Ren, L.-L. (2015). Two new stemphol sulfates from the mangrove endophytic fungus *Stemphylium* sp. 33231. *The Journal of Antibiotics*, *68*, 501–503. https://doi.org/10.1038/ja.2015.16

Zou, Y.-N., Wu, Q.-S., & Kuča, K. (2021). Unravelling the role of arbuscular mycorrhizal fungi in mitigating the oxidative burst of plants under drought stress. *Plant Biology*, *23*(S1), 50–57. https://doi.org/10.1111/PLB.13161

SUPPLEMENTAL DATA

Supplementary Figure 1: R Code for Metabolome Chemical Structure Analysis

```
# clear global environment
rm(list = ls())

# install packages
{ install.packages("xml2")
  install.packages("RxnSim")
  install.packages("reshape")
  install.packages("ggplot2")
  install.packages("berryFunctions")
  install.packages("data.table")
  install.packages("dendextend")
  install.packages("ggdendro")
  install.packages("dplyr")
  install.packages("rlist")
}
{ library(xml2)
  library(RxnSim)
  library(reshape)
  library(ggplot2)
  library(berryFunctions)
  library(data.table)
  library(dendextend)
  library(ggdendro)
  library(dplyr)
  library(rlist)
}

# choose the file containing the chemical names, load file, and determine
# number of rows
```

```
pth_names <- file.choose("")
names_table <- fread(pth_names, header = F, sep = "\t",
                                          quote = "", encoding = "UTF-8")
table_rows <- dim(names_table)[1]

# assign API URL
apiUrl <- "https://pubchem.ncbi.nlm.nih.gov/rest/pug/compound/name/"

runQuery <- function(x){
  # declare query start
  message(paste("Sending get request"))
  # format the api url call
  queryUrl <- paste(apiUrl, x, "/XML", sep="")
  # Encode for compatibility
  queryUrl <- URLencode(queryUrl)
  # query the api and handle errors
  if(is.error(read_html(queryUrl)) == TRUE){
    queryResult <- NA
    return(c(x,queryResult))
    try(closeAllConnections())
  }else{
  # parse the xml results
  querySpec <- xml_text(xml_find_all(read_html(queryUrl),
                                       "//pc-infodata_value_sval")[12])
  # create a blank vector for data
  singleOut <- c(x,querySpec)
  try(closeAllConnections())
  return(singleOut)
  }
}

# search down the list for the number of rows in table
cpdNames <- as.character(unlist(names_table))
```

```
resultList <- mapply(runQuery, cpdNames)
resultDf <- t(as.data.frame(resultList))
row.names(resultDf) <- seq_along(cpdNames)
colnames(resultDf) <- c("Compound Name","SMILES")

# extract viable SMILES
cleanResultDf <- resultDf[is.na(resultDf[,2]) == FALSE, ]
row.names(cleanResultDf) <- c(1:nrow(cleanResultDf))

# trim names absent from BioCyc
resultDframe <- as.data.frame(resultId)
bioCycInt <- intersect(resultDframe$BioCyc, names(pwyListFinal))
bioCycFiltered <- resultDframe$BioCyc[lapply(resultDframe$BioCyc, function(x)
which(x %in% bioCycInt)) == 1]
bioCycDerivative <- bioCycFiltered[is.na(bioCycFiltered) == F]

# create a similarity matrix
#Sim_matrix <- ms.compute.sim.matrix(cleanResultDf[,2], format = 'smiles',
#                                    sim.method = 'tanimoto', fp.type = "pubchem")
Sim_matrix <- ms.compute.sim.matrix(overlapResults[,2], format = 'smiles',
                                      sim.method   =   'tanimoto',   fp.type   =
"pubchem")
colnames(Sim_matrix) <- overlapResults[,1]
rownames(Sim_matrix) <- overlapResults[,1]
# create similarity matrix plot
get_lower_tri<-function(Sim_matrix){
  Sim_matrix[upper.tri(Sim_matrix)] <- NA
  return(Sim_matrix)
}
low_Sim <- get_lower_tri(Sim_matrix)
melt_sim <- melt(low_Sim)
head(melt_sim)
mat_plot <- ggplot(data = melt_sim, aes(x=Var1, y=Var2, fill=value))+
  geom_raster() +   geom_tile(color = "white") +
```

```
  scale_fill_gradient2(low = "white", high = "Black", midpoint = 0.5,
                       limit = c(0,1), space = "Lab",  na.value = "white",
                       name="Similarity") + theme_light() +
  theme(axis.text.x = element_text(angle = 90, vjust = 0, size = 1, hjust =
1),axis.text.y = element_text(angle = 0, vjust = 0, size = 1, hjust = 1)) +
  coord_fixed()



plot(mat_plot)
# generate cluster dendrogram
#colnames(Sim_matrix) <- cleanResultDf[,1]
#rownames(Sim_matrix) <- cleanResultDf[,1]
colnames(Sim_matrix) <- overlapResults[,1]
rownames(Sim_matrix) <- overlapResults[,1]
colnames(Sim_matrix) <- names(pwyListRmsp)
rownames(Sim_matrix) <- names(pwyListRmsp)
Sim_dist <- dist(Sim_matrix, method = "euclidean")
Sim_h <- hclust(Sim_dist)
Sim_dendro <- as.dendrogram(Sim_h)
labels_cex(Sim_dendro) <- 1
Sim_plot <- plot(Sim_dendro,
                 main  =  "Dendrogram  of  fungal  metabolites  from  PNW2016-
02",#"Dendrogram of fungal metabolites also found in BioCyc",
                 sub = NA, xlab = NA, ylab = NA,  xaxt='n',
                 yaxt='n')


# Compare the dendrograms
simDend <- as.dendrogram(Sim_h)
compDend <- as.dendrogram(compH)
tanglegram(untangle(dendlist(simDend, compDend), method = "step2"),
                    sort = F,
                    main_left = "Chemical Similarity",
                    main_right = "Pathway Similarity",
                    lwd = 2,
```

```
                edge.lwd = 2,
                left_dendo_mar = c(2.5,1,2.5,15),
                right_dendo_mar = c(2.5,15,2.5,1),
                columns_width = c(15, 7, 15),
                intersecting = T,
                dLeaf = -0.05,
                axes = F,
                type = "r",
                lab.cex = 1.5,
                remove_nodePar = T,
                hang = F,
                main = "Comparison of metabolite dendrograms",
                highlight_distinct_edges = F,
                common_subtrees_color_lines = F,
                common_subtrees_color_lines_default_single_leaf_color    =
"grey",
                common_subtrees_color_branches = FALSE,
                highlight_branches_col = FALSE,
                highlight_branches_lwd                =                F)
```

Supplementary Figure 2: R Code for Metabolic Pathway Screening

```
# clear global environment
rm(list = ls())

# install packages
{ install.packages("rvest")
  install.packages("xml2")
  install.packages("reshape2")
  install.packages("rlist")
  install.packages("stringr")
  install.packages("gtools")
  install.packages("ggplot2")
  install.packages("vegan")
}
{ library(rvest)
  library(xml2)
  library(reshape2)
  library(rlist)
  library(stringr)
  library(gtools)
  library(ggplot2)
  library(vegan)
}

# load metabolite names
metaboliteDir <- pth_names
metaboliteTable <- as.vector(data.table::fread(metaboliteDir, header = F, sep
= "\t",
                                                quote = "", encoding = "UTF-
8"))

compoundId <- as.matrix(metaboliteTable)
compoundId <- as.matrix(overlapResults[,1])


# query for BioCyc ID.
colnames(compoundId) <- "Name"
findId <- function(compoundName){
  idList <- matrix(nrow = 0, ncol = 14)
  for(i in seq_along(compoundName)){
    compoundUrlName <- toupper(URLencode(compoundName[i,1]))
    compoundUrlName <- gsub("\\[","%5B",compoundUrlName)
    compoundUrlName <- gsub("\\]", "%5D", compoundUrlName)
    bioCycUrl <- paste("https://metacyc.org/META/metabolite-translation-
service?file=",
                       compoundUrlName,sep = "")
    pageResult <- read_html(bioCycUrl)
    pageTest <- grep("unknown", html_text(pageResult), value = F)
    if(length(pageTest) == 0){
      pageText <- html_text(xml_find_all(pageResult, "//p"))
      textSplitN <- strsplit(pageText, "\n")
      textSplitNt <- lapply(textSplitN,strsplit,"\t")
      htmlTextDf <- t(as.data.frame(textSplitNt))
      colnames(htmlTextDf) <- htmlTextDf[1,]
      rownames(htmlTextDf) <- 1:(nrow(htmlTextDf))
```

```
      htmlTextFinal <- as.vector(htmlTextDf[-1,])
      idList <- rbind(idList, htmlTextFinal)
      message(paste("BioCyc ID", htmlTextFinal[4], "for", compoundName[i,1],
                    "has been added to the results. \n"))
    }else{
      message(paste(compoundName[i,1],"is not found in the BioCyc
database.\n",
                    sep = " "))
    }
  }
  colnames(idList) <- colnames(htmlTextDf)
  rownames(idList) <- NULL
  return(idList)
}

# Pull BioCyc IDs. This may take a bit
resultId <- findId(compoundId)

findPathway <- function(bioCycId){
  pwyResult <- list()
  bioCycIdUrl <- URLencode(bioCycId)
  pathUrl <- paste("https://websvc.biocyc.org/apixml?fn=pathways-of-
compound&id=META:",
                   bioCycIdUrl,"&detail=low", sep = "")
  pageResult <- read_html(pathUrl)
  pwyName <- html_text(xml_find_all(pageResult, "//parent/pathway/@frameid"))
  if(length(pwyName) > 0){
    pwyResult[bioCycId] <- list(pwyName)
    message(pwyResult)
  }else{
    pwyResult[bioCycId] <- list()
  }
  message(length(unlist(pwyResult, recursive = F)))
  return(pwyResult)
}

# Apply findPathway to all metabolites. This may take a bit
pwyList <- lapply(resultId[,4], findPathway)
pwyListMod <- unlist(pwyList, recursive = F)
pwyListMod <- lapply(pwyListMod, as.list)
pwyListClean <- lapply(pwyListMod, unique)

# Remove Super-Pathways from pathwayLists
rmSp <- function(pathwayListElement){
  elementSpIndex <- lapply(pathwayListElement, str_detect, pattern = "Super-
Pathways")
  pathwayListElement <- pathwayListElement[elementSpIndex == FALSE]
}

pwyListFinal <- lapply(pwyListClean, rmSp)

pwyListRmsp <- lapply(pwyListMod, rmSp)

# Generates Jaccard index by comparing the elements of two vectors
jIndex <- function(vectorA, vectorB){
  andAB <- intersect(vectorA, vectorB)
  orAB <- union(vectorA, vectorB)
```

```
  j <- length(andAB) / length(orAB)
  return(j)
}

compareVector <- pwyListFinal

loadSquareMatrix <- function(compareVector){
  sideLength <- length(compareVector)
  compMat <- matrix(nrow = sideLength, ncol = sideLength)
  matIndex <- permutations(sideLength, 2, repeats.allowed = TRUE)
  for(i in 1:nrow(matIndex)){
    compA <- matIndex[i,1]
    compB <- matIndex[i,2]
    compJ <- jIndex(compareVector[[compA]], compareVector[[compB]])
    compMat[compA,compB] <- compJ
    message(paste("Set value", compJ, "at index",
                  paste("[", compA, ",", compB, "].", sep = ""), "Finished",
                  i, "of", sideLength^2))
  }
  return(compMat)
}

# assign data to be plotted
pwyData <- pwyListFinal

pathSim <- loadSquareMatrix(pwyData)
colnames(pathSim) <- names(pwyListRmsp)
rownames(pathSim) <- names(pwyListRmsp)

# Plot graphs
compMat <- pathSim
get_lower_tri<-function(compMat){
  compMat[upper.tri(compMat)] <- NA
  return(compMat)
}
low_Sim <- get_lower_tri(compMat)
melt_sim <- melt(low_Sim)
head(melt_sim)
compPlot <- ggplot(data = melt_sim, aes(x=Var1, y=Var2, fill=value))+
  geom_raster() +   geom_tile(color = "white") +
  scale_fill_gradient2(low = "white", high = "Black", midpoint = 0.5,
                       limit = c(0,1), space = "Lab",  na.value = "white",
                       name="Similarity") + theme_light() +
  theme(axis.text.x = element_text(angle = 90, vjust = 0, size = 6, hjust =
1), axis.text.y = element_text(angle = 0, vjust = 0, size = 6, hjust = 1)) +
  coord_fixed()

plot(compPlot)
# Generate cluster dendrogram
#colnames(compMat) <- cleanResultDf[,1]
#rownames(compMat) <- cleanResultDf[,1]
colnames(compMat) <- names(pwyListRmsp)
rownames(compMat) <- names(pwyListRmsp)
compDist <- dist(compMat, method = "euclidean")
compH <- hclust(compDist)
compDPlot <- plot(compH,
                  main = "Dendrogram of fungal metabolites from PNW2016-02",
```

```
                sub = NA,  cex = 0.8, xlab = NA, ylab = NA,  xaxt='n',
                yaxt='n')

# Construct data frame containing all comparisons of pathway data
i <- 1

intersectDf <- function(compareVector){
  sideLength <- length(compareVector)
  compDf <- data.frame()
  matIndex <- permutations(sideLength, 2, repeats.allowed = TRUE)
  for(i in 1:nrow(matIndex)){
    compA <- matIndex[i,1]
    compB <- matIndex[i,2]
    compJ <- list(intersect(compareVector[[compA]], compareVector[[compB]]))
    if(length(compJ) == 0){
      compDf[compA,compB][[1]] <- NULL
    }else{
      compDf[compA,compB][[1]] <- compJ
    }
  }
  return(compDf)
}

intDf <- intersectDf(pwyListMod)
row.names(intDf) <- names(pwyListMod)
colnames(intDf) <- names(pwyListMod)

# identify metabolites with common pathways
featList <- pwyListFinal
pwyName <- names(pwyListFinal)
uniquePwy <- unique(as.character(unlist(featList)))
featDf <- data.frame()

overlapper <- function(listNamedPathways){
  allPwy <- unique(as.character(unlist(listNamedPathways)))
  for(i in 1:length(listNamedPathways)){
    colReturn <- lapply(allPwy, grepl, listNamedPathways[i])
    featDf[i,1:length(allPwy)] <- colReturn
  }
  rownames(featDf) <- names(listNamedPathways)
  colnames(featDf) <- allPwy
  return(featDf)
}

overlapList <- overlapper(featList)
overlapBinary <- mapply(as.numeric, overlapList)
rownames(overlapBinary) <- pwyName

# invert values to generate a dissimilarity matrix
invert <- function(valueReg){
  valueInv <- 1 - valueReg
  return(valueInv)
}
Dissim_matrix <- apply(Sim_matrix, 1, invert)
compDissim <- apply(compMat, 1, invert)

# generate NMDS with dissimilarity matrix
```

```
anosimV <- c()
for(i in 1:NROW(breakMat)){
  anosimCut <- anosim(Dissim_matrix, breakMat[i,], 9999)
  anosimV <- c(anosimV, anosimCut)
}

metabMDS <- metaMDS(Dissim_matrix)
metabScores <- as.data.frame(scores(metabMDS))
metabScores$tcut = breakMat[2,]
metabScores$name = overlapResults[,1]
metabPlot = ggplot(metabScores, aes(x = NMDS1, y = NMDS2)) +
  geom_point(size = 2, aes(colour = tcut)) +
  geom_text(aes(label=colnames(Dissim_matrix),hjust=0, vjust=0))
plot(metabPlot)
```

Supplementary Figure 3: R Code for Parsing gff3 Files

```
# clear global environment
rm(list = ls())

# install packages
{ install.packages("ape")
  install.packages("plyr")
  install.packages("stringr")
  install.packages("rvest")
  install.packages("xml2")
  install.packages("jsonlite")
}
{ library(ape)
  library(plyr)
  library(stringr)
  library(rvest)
  library(xml2)
  library(jsonlite)
}

# select GFF3 file and reformat
gffFile <- file.choose("")
gffDf <- read.gff(gffFile)
gffDfSplit <- mapply(strsplit, gffDf$attributes,";")
gffDfF <- ldply(gffDfSplit,rbind)
gffDf <- cbind(gffDf[1:8],gffDfF)
rm(gffDfF, gffDfSplit)

# extract all PFAM IDs
pfamMatches <- grepl("PFAM", gffDf)
pfamRaw <- gffDf[pfamMatches == TRUE]
pfamIds <- mapply(strsplit, pfamRaw, "PFAM:")
pfamIds <- gsub("DBxref=", "", pfamIds)
pfamIds <- gsub(",", "", pfamIds)
pfamIdIndex <- grep("PF",pfamIds)
pfIds <- as.list(pfamIds[pfamIdIndex])

allPfIds <- str_match(pfIds, "PF[0-9][0-9][0-9][0-9][0-9]")
allPfIds <- unique(allPfIds)

queryBase <- "http://pfam.xfam.org/family/"
pfamNames <- list()
for(i in 1:length(allPfIds)){
  queryUrl <- paste(queryBase, allPfIds[i], sep = "")
  urlContent <- read_html(queryUrl)
  domainSummary <- html_text(xml_find_all(urlContent, "//body/div/div/h1"))
  domainFunction <- html_text(xml_find_all(urlContent,
"//div/div/div/div/div/div/div/table/tbody/tr/td/a[contains(text(),'GO')]"))
  pfamListElement <- list(domainFunction)
  names(pfamListElement) <- domainSummary
  panaxes[i] <- list(pfamListElement)
  writeLines(paste(domainSummary, "\n", i, " of ", length(allPfIds), " (",
round(100*(i/length(allPfIds)), digits = 2), "%)", sep = "" ))
}
```

```
rm(i, pfamListElement, domainFunction, domainSummary, queryUrl)
```

Supplementary Figure 4: R Code for Comparing Gene Products

```
# clear global environment
rm(list = ls())

# install packagesinstall.packages("plyr")
{ install.packages("readr")
  install.packages("data.table")
  install.packages("stringr")
  install.packages("ggplot2")
  install.packages("ggvenn")
}
{ library(plyr)
  library(readr)
  library(data.table)
  library(stringr)
  library(ggplot2)
  library(ggvenn)
}

# select GFF3 files
importPath <-choose.dir("")
importFiles <- list.files(importPath,".gff3", full.names = T)

# pull gene products
grabProducts <- function(x){
  importTable <- fread(x, header = T)
  dbXrefColumn <- importTable[,9]
  productIndex <- grep("product=", unlist(dbXrefColumn))
  allProducts <- unlist(dbXrefColumn)[productIndex]
  products <- str_extract_all(allProducts,"product=*(.*?)*;")
  products <- gsub("product=", "", products)
  products <- gsub(";", "", products)
  return(products)
}

# compare gene products and generate Venn diagram
productsToAnalyze <- lapply(importFiles, grabProducts)
uniqueProductsToAnalyze <- lapply(productsToAnalyze, unique)
names(uniqueProductsToAnalyze) <- c("Stemphylium lycopersici", "PNW2016-02",
"Stemphylium vesicarium")
uniqueProductsToAnalyze <- lapply(uniqueProductsToAnalyze,
function(x){x[2:length(x)]})

compVend <- ggvenn(uniqueProductsToAnalyze, stroke_size = 0, set_name_size =
7, text_size = 7, fill_color = c("green", "blue", "red"), fill_alpha = 0.4)
plot(compVend)
```

Supplementary Table 1: GenBank Accession Numbers for Species Included in Phylogenetic Analysis

GenBank Accession No.	Species	Sample Location	Sequence Source
AJ853759.1	*Alternaria alternata*	Unknown	Gene sequence
JQ988075.1	*Pleospora eturmiuna*	Spain	Gene sequence
JQ988080.1	*Pleospora eturmiuna*	Spain	Gene sequence
JQ988088.1	*Pleospora eturmiuna*	Spain	Gene sequence
JQ988083.1	*Pleospora eturmiuna*	Spain	Gene sequence
JQ988084.1	*Pleospora gracilariae*	Spain	Gene sequence
AB979880.1	*Pleospora herbarum*	Japan	Gene sequence
FR667974.1	*Pleospora herbarum*	Spain	Gene sequence
LC056844.1	*Pleospora herbarum*	Japan	Gene sequence
LC056845.1	*Pleospora herbarum*	Japan	Gene sequence
FN868457.1	*Pleospora herbarum*	Spain	Gene sequence
AF442774.1	*Stemphylium alfalfae*	USA	Gene sequence
AF442775.1	*Stemphylium alfalfae*	USA	Gene sequence
AF442776.1	*Stemphylium alfalfae*	USA	Gene sequence
AF071343.1	*Stemphylium alfalfae*	Unknown	Gene sequence
KY883935.1	*Stemphylium botryosum*	Denmark	Gene sequence
KY883975.1	*Stemphylium botryosum*	Unknown	Gene sequence
KC584238.1	*Stemphylium botryosum*	Canada	Gene sequence
MG583739.1	*Stemphylium botryosum*	Iran	Gene sequence
NR 163547.1	*Stemphylium botryosum*	Canada	Gene sequence
AY329230.1	*Stemphylium eturmiunum*	New Zealand	Gene sequence
KU850550.1	*Stemphylium gracilariae*	Spain	Gene sequence
KU850551.1	*Stemphylium gracilariae*	Spain	Gene sequence
AY329217.1	*Stemphylium gracilariae*	Israel	Gene sequence
KU850552.1	*Stemphylium gracilariae*	China	Gene sequence
MK500708.1	*Stemphylium gracilariae*	Italy	Gene sequence
MH862230.1	*Stemphylium gracilariae*	Israel	Gene sequence
KU850595.1	*Stemphylium lancipes*	New Zealand	Gene sequence
KU850596.1	*Stemphylium lancipes*	USA	Gene sequence
AY329203.1	*Stemphylium lancipes*	New Zealand	Gene sequence
NR 154930.1	*Stemphylium lancipes*	USA	Gene sequence
LC219361.1	*Stemphylium lancipes*	Japan	Gene sequence
AB704312.1	*Stemphylium lycopersici*	Japan	Gene sequence
AB704313.1	*Stemphylium lycopersici*	Japan	Gene sequence
AB704314.1	*Stemphylium lycopersici*	Japan	Gene sequence
LC333577.1	*Stemphylium lycopersici*	Japan	Gene sequence
NR 155002.1	*Stemphylium lycopersici*	China	Gene sequence
LC032102.1	*Stemphylium lycopersici*	Japan	Gene sequence
KU850575.1	*Stemphylium mali*	China	Gene sequence
KU850645.1	*Stemphylium symphyti*	New Zealand	Gene sequence
KY883773.1	*Stemphylium symphyti*	Unknown	Gene sequence
KY883775.1	*Stemphylium symphyti*	Unknown	Gene sequence
KY883955.1	*Stemphylium symphyti*	New Zealand	Gene sequence
KY883977.1	*Stemphylium symphyti*	Unknown	Gene sequence
NR 154935.1	*Stemphylium symphyti*	USA	Gene sequence

LN896693.1	*Stemphylium vesicarium*	Italy	Gene sequence
AM746023.1	*Stemphylium vesicarium*	Italy	Gene sequence
JQ988100.1	*Stemphylium vesicarium*	Spain	Gene sequence
LC512754.1	*Stemphylium vesicarium*	South Korea	Gene sequence
LC512757.1	*Stemphylium vesicarium*	South Korea	Gene sequence
VHPU01002624	*Stemphylium vesicarium*	USA	Genome

Supplementary Table 2: Key of Database Identification Tags for Metabolites of PNW2016-02

Compound Name	IUPAC Name	Canonical SMILES	MetaCyc ID
(-)-Chimonanthine	(3aS,8bS)-8b-[(3aS,8bS)-3-methyl-1,2,3a,4-tetrahydropyrrolo[2,3-b]indol-8b-yl]-3-methyl-1,2,3a,4-tetrahydropyrrolo[2,3-b]indole	CN1CCC2(C1NC3=CC=CC=C32)C45CCN(C4NC6=CC=CC=C56)C	
(-)-trans-(S)-allethrin	[(1S)-2-methyl-4-oxo-3-prop-2-enylcyclopent-2-en-1-yl] (1S,3S)-2,2-dimethyl-3-(2-methylprop-1-enyl)cyclopropane-1-carboxylate	CC1=C(C(=O)CC1OC(=O)C2C(C2(C)C)C=C(C)C)CC=C	
(+)-12-Isocopalene-15,16-dial	(4*bS*,8*aS*,10*aR*)-4*b*,8,8,10*a*-tetramethyl-4,4*a*,5,6,7,8*a*,9,10-octahydro-1*H*-phenanthrene-1,2-dicarbaldehyde	CC1(CCCC2(C1CCC3(C2CC=C(C3C=O)C=O)C)C)C	
(13Z,16Z)-Docosadi-13,16-enoyl-CoA	S-[2-[3-[[4-[[[(2*R*,3*S*,4*R*,5*R*)-5-(6-aminopurin-9-yl)-4-hydroxy-3-phosphonooxyoxolan-2-yl]methoxy-hydroxyphosphoryl]oxy-hydroxyphosphoryl]oxy-2-hydroxy-3,3-dimethylbutanoyl]amino]propanoylamino]ethyl] (13*Z*,16*Z*)-docosa-13,16-dienethioate	CCCCCC=CCC=CCCCCCCCCCCCC(=O)SCCNC(=O)CCNC(=O)C(C(C)(C)COP(=O)(O)OP(=O)(O)OCC1C(C(C(O1)N2C=NC3=C(N=CN=C32)N)O)OP(=O)(O)O)O	
(17Z)-1,25-dihydroxy-17,20-didehydro-21-norvitamin D3			
(20S)-24-Hydroxy-19-norgeminivitamin D3	(3*R*,6*S*)-6-[(1*R*,3*aS*,4*E*,7*aR*)-4-[2-[(3*R*,5*R*)-3,5-dihydroxycyclohexylidene]ethylidene]-7*a*-methyl-2,3,3*a*,5,6,7-hexahydro-1*H*-inden-1-yl]-2,10-dimethylundecane-2,3,10-triol	CC12CCCC(=CC=C3CC(CC(C3)O)O)C1CCC2C(CCCC(C)(C)O)CCC(C(C)(C)O)O	
(22Z)-1-hydroxy-22,23-didehydrovitamin D3			
(25R)-26,26,26-trifluoro-1,25-dihydroxyvitamin D3			
(2E,5E)-Undecadienal			
(6R)-vitamin D3 6,19-sulfur dioxide adduct	(3*R*,5*S*)-3-[(*E*)-[(1*R*,3*aS*,7*aR*)-7*a*-methyl-1-[(2*R*)-6-methylheptan-2-yl]-2,3,3*a*,5,6,7-hexahydro-1*H*-inden-4-ylidene]methyl]-2,2-dioxo-1,3,4,5,6,7-hexahydro-2-benzothiophen-5-ol	CC(C)CCCC(C)C1CCC2C1(CCCC2=CC3C4=C(CCC(C4)O)CS3(=O)=O)C	
(E,E)-3,7,11-Trimethyl-2,6,10-dodecatrienyl octanoate	[(2*E*,6*E*)-3,7,11-trimethyldodeca-2,6,10-trienyl] octanoate	CCCCCCCC(=O)OCC=C(C)CCC=C(C)CCC=C(C)C	
(S)-2,3-Dihydro-7-hydroxy-2-methyl-4-oxo-4H-1-benzopyran-5-acetic acid	2-(7-hydroxy-2-methyl-4-oxo-2,3-dihydrochromen-5-yl)acetic acid	CC1CC(=O)C2=C(C=C(C=C2O1)O)CC(=O)O	
(Z)-Resveratrol	5-[(*Z*)-2-(4-hydroxyphenyl)ethenyl]benzene-1,3-diol	C1=CC(=CC=C1C=CC2=CC(=CC(=C2)O)O)O	
Linolenoyl-CoA	[(2*R*,3*S*,4*R*,5*R*)-5-(6-aminopurin-9-yl)-4-hydroxy-2-[[[[(3*R*)-3-hydroxy-2,2-dimethyl-4-[[3-[2-[(9*Z*,12*Z*,15*Z*)-octadeca-9,12,15-trienoyl]sulfanylethylamino]-3-oxopropyl]amino]-4-oxobutoxy]-oxidophosphoryl]oxy-oxidophosphoryl]oxymethyl]oxolan-3-yl] phosphate	CCC=CCC=CCC=CCCCCCCCC(=O)SCCNC(=O)CCNC(=O)C(C(C)(C)COP(=O)([O-])OP(=O)([O-])OCC1C(C(C(O1)N2C=NC3=C(N=CN=C32)N)O)OP(=O)([O-])[O-])O	LINOLENOYL-COA
1(15),8(19)-Trinervitadiene-3alpha,5alpha,18-triol	(1*R*,6*S*,11*S*,12*S*,13*R*,15*S*)-12-(hydroxymethyl)-6,16-dimethyl-2-methylidenetricyclo[7.5.2.012,15]hexadec-9(16)-ene-11,13-diol	CC1CCCC(=C)C2CC(C3(C2C(=C(CC1)CC3O)C)CO)O	
1-(5-Acetyl-2-hydroxyphenyl)-3-methyl-1-butanone	1-(5-acetyl-2-hydroxyphenyl)-3-methylbutan-1-one	CC(C)CC(=O)C1=C(C=CC(=C1)C(=O)C)O	
1,2-Dioleoyl-sn-Glycero-3-Phosphate	[2-[(*E*)-octadec-9-enoyl]oxy-3-phosphonooxypropyl] (*E*)-octadec-9-enoate	CCCCCCCCC=CCCCCCCCC(=O)OCC(COP(=O)(O)O)OC(=O)CCCCCCC=CCCCCCCCC	
1,4-Dideoxy-1,4-imino-D-ribitol	(2*R*,3*R*,4*S*)-2-(hydroxymethyl)pyrrolidine-3,4-diol	C1C(C(C(N1)CO)O)O	
100-1			

10-deacetylbaccatin III	[(1*S*,2*S*,3*R*,4*S*,7*R*,9*S*,10*S*,12*R*,15*S*)-4-acetyloxy-1,9,12,15-tetrahydroxy-10,14,17,17-tetramethyl-11-oxo-6-oxatetracyclo[11.3.1.03,10.04,7]heptadec-13-en-2-yl] benzoate	CC1=C2C(C(=O)C3(C(CC4C(C3C(C(C2(C)C)(CC1O)O)OC(=O)C5=CC=CC=C5)(CO4)OC(=O)C)O)C)O	10-DEACETYLBACCATIN-III
11(S)-HETE	(5*Z*,8*Z*,11*S*,12*E*,14*Z*)-11-hydroxyicosa-5,8,12,14-tetraenoic acid	CCCCCC=CC=CC(CC=CCC=CCCCC(=O)O)O	
11-cis-retinyl palmitate	[(2*E*,4*Z*,6*E*,8*E*)-3,7-dimethyl-9-(2,6,6-trimethylcyclohexen-1-yl)nona-2,4,6,8-tetraenyl] hexadecanoate	CCCCCCCCCCCCCCCC(=O)OCC=C(C)C=CC=C(C)C=CC1=C(CCCC1(C)C)C	
13-Hydroxy-2-(hydroxymethylene)-3-oxo-13,17-seco-5alpha-androstan-17-oic acid, delta-lactone	(4*aS*,4*bR*,6*aS*,9*Z*,10*aS*,10*bS*,12*aS*)-9-(hydroxymethylidene)-10*a*,12*a*-dimethyl-3,4,4*a*,4*b*,5,6,6*a*,7,10,10*b*,11,12-dodecahydronaphtho[2,1-f]chromene-2,8-dione	CC12CCC3C(C1CCC(=O)O2)CCC4C3(CC(=CO)C(=O)C4)C	
14:0 Cholesteryl ester	[(3*S*,8*S*,9*S*,10*R*,13*R*,14*S*,17*R*)-10,13-dimethyl-17-[(2*R*)-6-methylheptan-2-yl]-2,3,4,7,8,9,11,12,14,15,16,17-dodecahydro-1*H*-cyclopenta[a]phenanthren-3-yl] tetradecanoate	CCCCCCCCCCCCCC(=O)OC1CCC2(C3CCC4(C(C3CC=C2C1)CCC4C(C)CCCC(C)C)C)C	
14alpha-Hydroxyixocarpanolide	15-[1-(4,5-dimethyl-6-oxooxan-2-yl)-1-hydroxyethyl]-5,18-dihydroxy-10,14-dimethyl-3-oxapentacyclo[9.7.0.02,4.05,10.014,18]octadec-7-en-9-one	CC1CC(OC(=O)C1C)C(C)(C2CCC3(C2(CCC4C3C5C(O5)C6(C4(C(=O)C=CC6)C)O)C)O)O	
16beta-Hydroxy-3,11-dioxopregna-4,17(20)-dien-21-oic acid, gamma-lactone	(1*S*,2*S*,4*S*,9*S*,12*S*,13*R*)-9,13-dimethyl-5-oxapentacyclo[10.8.0.02,9.04,8.013,18]icosa-7,17-diene-6,11,16-trione	CC12CCC(=O)C=C1CCC3C2C(=O)CC4(C3CC5C4=CC(=O)O5)C	
17:1 Cholesteryl ester	[(3*S*,8*S*,9*S*,10*R*,13*R*,14*S*,17*R*)-10,13-dimethyl-17-[(2*R*)-6-methylheptan-2-yl]-2,3,4,7,8,9,11,12,14,15,16,17-dodecahydro-1*H*-cyclopenta[a]phenanthren-3-yl] (*Z*)-heptadec-9-enoate	CCCCCCCC=CCCCCCCCC(=O)OC1CCC2(C3CCC4(C(C3CC=C2C1)CCC4C(C)CCCC(C)C)C)C	
17abeta-Hydroxy-D-homoandrost-4-en-3-one propionate	[(1*S*,4*aS*,4*bR*,10*aR*,10*bS*,12*aS*)-10*a*,12*a*-dimethyl-8-oxo-2,3,4,4*a*,4*b*,5,6,9,10,10*b*,11,12-dodecahydro-1*H*-chrysen-1-yl] propanoate	CCC(=O)OC1CCCC2C1(CCC3C2CCC4=CC(=O)CCC34C)C	
17-Epiestriol	(8*R*,9*S*,13*S*,14*S*,16*R*,17*S*)-13-methyl-6,7,8,9,11,12,14,15,16,17-decahydrocyclopenta[a]phenanthrene-3,16,17-triol	CC12CCC3C(C1CC(C2O)O)CCC4=C3C=CC(=C4)O	
18:0 Cholesteryl ester	[(3*S*,8*S*,9*S*,10*R*,13*R*,14*S*,17*R*)-10,13-dimethyl-17-[(2*R*)-6-methylheptan-2-yl]-2,3,4,7,8,9,11,12,14,15,16,17-dodecahydro-1*H*-cyclopenta[a]phenanthren-3-yl] octadecanoate	CCCCCCCCCCCCCCCCCC(=O)OC1CCC2(C3CCC4(C(C3CC=C2C1)CCC4C(C)CCCC(C)C)C)C	
1-Acetoxyeugenol acetate	(6-acetyloxy-6-methoxy-4-prop-2-enylcyclohexa-2,4-dien-1-yl) acetate	CC(=O)OC1C=CC(=CC1(OC)OC(=O)C)CC=C	
1-hexadecanyl-2-((2-alpha-glucosyl)-beta-glucosyl)-3-beta-xylosyl-sn-glycerol	(2*S*,5*S*,6*R*)-2-[(2*R*,5*S*,6*R*)-2-[(2*R*)-1-[(2*S*,3*R*,5*S*)-3,4-dihydroxy-5-(hydroxymethyl)oxolan-2-yl]oxy-3-hexadecoxypropan-2-yl]oxy-4,5-dihydroxy-6-(hydroxymethyl)oxan-3-yl]oxy-6-(hydroxymethyl)oxane-3,4,5-triol	CCCCCCCCCCCCCCCCOCC(COC1C(C(C(O1)CO)O)O)OC2C(C(C(C(O2)CO)O)O)OC3C(C(C(C(O3)CO)O)O)O	
2-(1,3-Benzodioxol-5-yl)-7-hydroxy-3,5,6,8-tetramethoxy-4H-1-benzopyran-4-one	2-(1,3-benzodioxol-5-yl)-7-hydroxy-3,5,6,8-tetramethoxychromen-4-one	COC1=C(C(=C(C2=C1C(=O)C(=C(O2)C3=CC4=C(C=C3)OCO4)OC)OC)O)OC	
2-(4-Hydroxy-3,5-dimethoxyphenyl)ethanol	2-[4-(2-hydroxyethyl)-2,6-dimethoxyphenoxy]-6-(hydroxymethyl)oxane-3,4,5-triol	COC1=CC(=CC(=C1OC2C(C(C(C(O2)CO)O)O)O)OC)CCO	
2,2,4,4-Tetramethyl-6-(1-oxobutyl)-1,3,5-cyclohexanetrione	6-butanoyl-2,2,4,4-tetramethylcyclohexane-1,3,5-trione	CCCC(=O)C1C(=O)C(C(=O)C(C1=O)(C)C)(C)C	
2,2-Dibutyl-3-(4-methoxyphenyl)-4-methyl-2H-1-benzopyran-7-ol acetate	[2,2-dibutyl-3-(4-methoxyphenyl)-4-methylchromen-7-yl] acetate	CCCCC1(C(=C(C2=C(O1)C=C(C=C2)OC(=O)C)C)C3=CC=C(C=C3)OC)CCCC	
2,6-Diamino-4-hydroxy-5-N-methylformamidopyrimidine	N-(2,4-diamino-6-oxo-1*H*-pyrimidin-5-yl)-*N*-methylformamide	CN(C=O)C1=C(N=C(NC1=O)N)N	
24-Nor-5-cholane-3,7,12-triol			
28-Glucosyl-30-methyl-3b,23-dihydroxy-12-oleanene-28,30-dioate	3,5-dihydroxy-6-[[4-(hydroxymethyl)-11-methoxycarbonyl-4,6a,6b,11,14b-pentamethyl-8a-[3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxycarbonyl-1,2,3,4a,5,6,7,8,9,10,12,12a,14,14a-tetradecahydropicen-3-	CC1(CCC2(CCC3(C(=CCC4C3(CCC5C4(CCC(C5(C)CO)OC6C(C(C(C(O6)C(=O)O)O)O)OC7C(C(C(CO7)O)O)O)O)C)C)C2C1)C)C(=O)	

	yl]oxy]-4-(3,4,5-trihydroxyoxan-2-yl)oxyoxane-2-carboxylic acid	OC8C(C(C(C(O8)CO)O)O)O)C(=O)OC	
2-amino-heptanoic acid	2-aminoheptanoic acid	CCCCCC(C(=O)O)N	
2E,4Z-Heptadienal	(2*E*,4*Z*)-hepta-2,4-dienal	CCC=CC=CC=O	
2E,6E-Octadienal	(2*E*,6*E*)-octa-2,6-dienal	CC=CCCC=CC=O	
2-Methyl-4-pentenal	2-methylpent-4-enal	CC(CC=C)C=O	
2-N-Acetyl-6-deamino-6-hydroxyneomycin C	N-[(2R,3R,4R,5S,6R)-2-[(2R,3S,4R,5S)-5-[(1R,2R,3S,5R,6S)-3,5-diamino-2-[(2R,3R,4R,5S,6R)-3-amino-6-(aminomethyl)-4,5-dihydroxyoxan-2-yl]oxy-6-hydroxycyclohexyl]oxy-4-hydroxy-2-(hydroxymethyl)oxolan-3-yl]oxy-4,5-dihydroxy-6-(hydroxymethyl)oxan-3-yl]acetamide	CC(=O)NC1C(C(C(OC1OC2C(OC(C2O)OC3C(C(CC(C3OC4C(C(C(C(O4)CN)O)O)N)N)N)O)CO)CO)O)O	
2-O-p-Hydroxybenzoylorientin	[(2*S*,4*S*)-2-[2-(3,4-dihydroxyphenyl)-5,7-dihydroxy-4-oxochromen-8-yl]-4,5-dihydroxy-6-(hydroxymethyl)oxan-3-yl] 4-hydroxybenzoate	C1=CC(=CC=C1C(=O)OC2C(C(C(OC2C3=C(C=C(C4=C3OC(=CC4=O)C5=CC(=C(C=C5)O)O)O)O)C(O)O)O)O)O	
2-Phenylethanol	2-phenylethanol	C1=CC=C(C=C1)CCO	CPD-7035
3,4,5-Trimethoxycinnamyl alcohol acetate	[(*Z*)-3-(3,4,5-trimethoxyphenyl)prop-2-enyl] acetate	CC(=O)OCC=CC1=CC(=C(C(=C1)OC)OC)OC	
3-Hydroxy-OPC8-CoA	S-[2-[3-[[4-[[[(2R,3S,4R,5R)-5-(6-aminopurin-9-yl)-4-hydroxy-3-phosphonooxyoxolan-2-yl]methoxy-hydroxyphosphoryl]oxy-hydroxyphosphoryl]oxy-2-hydroxy-3,3-dimethylbutanoyl]amino]propanoylamino]ethyl] 3-hydroxy-8-[(1S,2S)-3-oxo-2-[(Z)-pent-2-enyl]cyclopentyl]octanethioate	CCC=CCC1C(CCC1=O)CCCCCC(CC(=O)SCCNC(=O)CCNC(=O)C(C(C)(C)COP(=O)(O)OP(=O)(O)OCC2C(C(C(O2)N3C=NC4=C(N=CN=C43)N)O)OP(=O)(O)O)O)O	
3-N-Debenzoyltaxol	[(1*S*,2*S*,3*R*,4*S*,7*R*,9*S*,10*S*,12*R*,15*S*)-4,12-diacetyloxy-15-[(2*R*,3*S*)-3-amino-2-hydroxy-3-phenylpropanoyl]oxy-1,9-dihydroxy-10,14,17,17-tetramethyl-11-oxo-6-oxatetracyclo[11.3.1.03,10.04,7]heptadec-13-en-2-yl] benzoate	CC1=C2C(C(=O)C3(C(CC4C(C3C(C(C2(C)C)(CC1OC(=O)C(C(C5=CC=CC=C5)N)O)O)OC(=O)C6=CC=CC=C6)(CO4)OC(=O)C)O)C)OC(=O)C	CPD-9203
3-O-alpha-L-Arabinopyranosylcinnamtannin B1	5,13-bis(3,4-dihydroxyphenyl)-7-[2-(3,4-dihydroxyphenyl)-3,5,7-trihydroxy-3,4-dihydro-2*H*-chromen-8-yl]-21-(3,4,5-trihydroxyoxan-2-yl)oxy-4,12,14-trioxapentacyclo[11.7.1.02,11.03,8.015,20]henicosa-2(11),3(8),9,15,17,19-hexaene-6,9,17,19-tetrol	C1C(C(OC2=C1C(=CC(=C2C3C(C(OC4=C3C(=CC5=C4C6C(C(O5)(OC7=CC(=CC(=C67)O)O)C8=CC(=C(C=C8)O)O)OC9C(C(C(CO9)O)O)O)O)C1=CC(=C(C=C1)O)O)O)O)O)C1=CC(=C(C=C1)O)O)O	
3-Oxo-delta1-steroid	10,13-dimethyl-4,5,6,7,8,9,11,12,14,15,16,17-dodecahydrocyclopenta[a]phenanthren-3-one	CC12CCCC1C3CCC4CC(=O)C=CC4(C3CC2)C	
3-oxo-octanoyl-CoA	S-[2-[3-[[(2*R*)-4-[[[(2*R*,3*R*,5*R*)-5-(6-aminopurin-9-yl)-4-hydroxy-3-phosphonooxyoxolan-2-yl]methoxy-hydroxyphosphoryl]oxy-hydroxyphosphoryl]oxy-2-hydroxy-3,3-dimethylbutanoyl]amino]propanoylamino]ethyl] 3-oxohexadecanethioate	CCCCCCCCCCCCCC(=O)CC(=O)SCCNC(=O)CCNC(=O)C(C(C)(C)COP(=O)(O)OP(=O)(O)OCC1C(C(C(O1)N2C=NC3=C(N=CN=C32)N)O)OP(=O)(O)O)O	CPD0-2106
4-(2,6,6-Trimethylcyclohex-1-enyl)but-2-en-4-one	(*E*)-1-(2,6,6-trimethylcyclohexen-1-yl)but-2-en-1-one	CC=CC(=O)C1=C(CCCC1(C)C)C	
4,2,3,4-Tetrahydroxychalcone	[(2*S*,4*S*,5*S*)-2-[2,3-dihydroxy-4-[(*E*)-3-(4-hydroxyphenyl)prop-2-enoyl]phenoxy]-4,5-dihydroxy-6-(hydroxymethyl)oxan-3-yl] (*E*)-3-(4-hydroxyphenyl)prop-2-enoate	C1=CC(=CC=C1C=CC(=O)C2=C(C(=C(C=C2)OC3C(C(C(C(O3)CO)O)O)OC(=O)C=CC4=CC=C(C=C4)O)O)O)O	
4,2,4,6-Tetrahydroxy-3-methoxychalcone	(*E*)-3-(4-hydroxy-3-methoxyphenyl)-1-(2,4,6-trihydroxyphenyl)prop-2-en-1-one	COC1=C(C=CC(=C1)C=CC(=O)C2=C(C=C(C=C2O)O)O)O	
4-Cumylphenol	4-(2-phenylpropan-2-yl)phenol	CC(C)(C1=CC=CC=C1)C2=CC=C(C=C2)O	
4-Hydroxyphenylglyoxylate	2-(4-hydroxyphenyl)-2-oxoacetic acid	C1=CC(=CC=C1C(=O)C(=O)O)O	CPD-142
4-O-Demethyl-13-dihydroadriamycinone			
5,10-Methylene-H4SPT	(2S)-2-[[2-[[(2R,3S,4R,5S)-5-[(2R,3S,4S)-5-[4-[(6S,6aR,7R)-3-amino-6,7-dimethyl-1-oxo-2,5,6,6a,7,9-hexahydroimidazo[1,5-f]pteridin-8-yl]phenyl]-2,3,4-	CC1C2C(N(CN2C3=C(N1)N=C(NC3=O)N)C4=CC=C(C=C4)CC(C(C(COC5C(C(C(O5)COP(=O)(O)O	CPD-10

	trihydroxypentoxy]-3,4-dihydroxyoxolan-2-yl]methoxy-hydroxyphosphoryl]oxy-4-carboxybutanoyl]amino]pentanedioic acid	C(CCC(=O)O)C(=O)NC(CCC(=O)O)C(=O)O)O)O)O)O)O)C	
5,2,5-Trihydroxy-3,7,8-trimethoxyflavone 2-acetate	2-[4-hydroxy-2-(5-hydroxy-3,7,8-trimethoxy-4-oxochromen-2-yl)phenyl]acetic acid	COC1=C(C2=C(C(=C1)O)C(=O)C(=C(O2)C3=C(C=CC(=C3)O)CC(=O)O)OC)OC	
5,4-Dihydroxy-3,3-dimethoxy-6,7-methylenedioxyflavone	9-hydroxy-6-(4-hydroxy-3-methoxyphenyl)-7-methoxy-[1,3]dioxolo[4,5-g]chromen-8-one	COC1=C(C=CC(=C1)C2=C(C(=O)C3=C(C4=C(C=C3O2)OCO4)O)OC)O	
5,7,11,14-Eicosatetraenoic acid, 9-hydroxy-, (E,Z,Z,Z)-; 9-HETE	(5*E*,7*Z*,11*Z*,14*Z*)-9-hydroxyicosa-5,7,11,14-tetraenoic acid	CCCCCC=CCC=CCC(C=CC=CCCCC(=O)O)O	
5,7,11,14-Eicosatetraenoic acid, 9-oxo-, (E,Z,Z,Z)-	(5*E*,7*Z*,11*Z*,14*Z*)-9-oxoicosa-5,7,11,14-tetraenoic acid	CCCCCC=CCC=CCC(=O)C=CC=CCCCC(=O)O	
5,7,3-Trihydroxy-3,6,8,4,5-pentamethoxyflavone	5,7-dihydroxy-2-(3-hydroxy-4,5-dimethoxyphenyl)-3,6,8-trimethoxychromen-4-one	COC1=CC(=CC(=C1OC)O)C2=C(C(=O)C3=C(C(=C(C(=C3O2)OC)O)OC)O)OC	
5?-Cholane-3,7,12,24-tetrol			
5alpha-Ethoxy-6beta-hydroxy-5,6-dihydrophysalin B	14-ethoxy-5,15-dihydroxy-2,9,26-trimethyl-3,19,23,28-tetraoxaoctacyclo[16.9.1.118,27.01,5.02,24.08,17.09,14.021,26]nonacos-11-ene-4,10,22,29-tetrone	CCOC12CC=CC(=O)C1(C3CCC4(C(=O)OC5(C46C7C(=O)C(C3CC2O)(O6)OCC8C7(CC5OC8=O)C)C)O)C	
5-Demethylmelibentin	2-(1,3-benzodioxol-5-yl)-5-hydroxy-3,6,7,8-tetramethoxychromen-4-one	COC1=C(C(=C2C(=C1O)C(=O)C(=C(O2)C3=CC4=C(C=C3)OCO4)OC)OC)OC	
5-Hydroxy-3,3,7,8-tetramethoxy-4,5-methylenedioxyflavone	3-hydroxy-5,7,8-trimethoxy-2-(7-methoxy-1,3-benzodioxol-5-yl)chromen-4-one	COC1=CC(=CC2=C1OCO2)C3=C(C(=O)C4=C(O3)C(=C(C=C4OC)OC)OC)O	
5-Methyltetrahydropteroylpentaglutamate	(2*S*,6*R*,11*S*)-2,11-diamino-6-[[(4*S*)-4-amino-4-carboxybutanoyl]-[4-[(2-amino-5-methyl-4-oxo-3,6,7,8-tetrahydropteridin-6-yl)methylamino]benzoyl]amino]-6-[(4*S*)-4-amino-4-carboxybutanoyl]oxycarbonyl-7-(carboxymethyl)-5,8-dioxododecanedioic acid	CN1C(CNC2=C1C(=O)NC(=N2)N)CNC3=CC=C(C=C3)C(=O)N(C(=O)CCC(C(=O)O)N)C(C(CC(=O)O)C(=O)CCC(C(=O)O)N)C(=O)CCC(C(=O)O)N)C(=O)OC(=O)CCC(C(=O)O)N	
6,8-nonadienal	(6*E*)-nona-6,8-dienal	C=CC=CCCCCC=O	
6-Glucopyranosylprocyanidin B2	(2*R*,3*R*,4*R*)-2-(3,4-dihydroxyphenyl)-4-[(2*R*,3*R*)-2-(3,4-dihydroxyphenyl)-3,5,7-trihydroxy-3,4-dihydro-2*H*-chromen-8-yl]-6-[(3*R*,4*R*,5*S*,6*R*)-3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]-3,4-dihydro-2*H*-chromene-3,5,7-triol	C1C(C(OC2=C1C(=CC(=C2C3C(C(OC4=C3C(=C(C(=C4)O)C5C(C(C(C(O5)CO)O)O)O)O)O)C6=CC(=C(C=C6)O)O)O)O)O)C7=CC(=C(C=C7)O)O)O	
6-hydroxydoxazosin	[4-(4-amino-6,7-dimethoxyquinazolin-2-yl)piperazin-1-yl]-(6-hydroxy-2,3-dihydro-1,4-benzodioxin-3-yl)methanone	COC1=C(C=C2C(=C1)C(=NC(=N2)N3CCN(CC3)C(=O)C4COC5=C(O4)C=C(C=C5)O)N)OC	
6-Hydroxyindoramin	N-[1-[2-(6-hydroxy-1*H*-indol-3-yl)ethyl]piperidin-4-yl]benzamide	C1CN(CCC1NC(=O)C2=CC=CC=C2)CCC3=CNC4=C3C=CC(=C4)O	
6-Keto-PGF1?			
7,12-Dimethylbenz[a]anthracene 5,6-oxide	12,19-dimethyl-3-oxapentacyclo[9.8.0.02,4.05,10.013,18]nonadeca-1(19),5,7,9,11,13,15,17-octaene	CC1=C2C3C(O3)C4=CC=CC=C4C2=C(C5=CC=CC=C15)C	
7-Epizucchini factor A	[11-(benzoyloxymethyl)-6-hydroxy-4,4,6*a*,6*b*,8*a*,11,14*b*-heptamethyl-1,2,3,4*a*,5,6,7,8,9,10,12,12*a*,13,14-tetradecahydropicen-3-yl] 4-aminobenzoate	CC1(C(CCC2(C1CC(C3=C2CCC4(C3(CCC5(C4CC(CC5)(C)COC(=O)C6=CC=CC=C6)C)C)C)O)C)OC(=O)C7=CC=C(C=C7)N)C	
7-Hydroxy-2,4,5-trimethoxyisoflavone	7-hydroxy-3-(2,4,5-trimethoxyphenyl)chromen-4-one	COC1=CC(=C(C=C1C2=COC3=C(C2=O)C=CC(=C3)O)OC)OC	CPD-9532
7-Hydroxy-2,5,6-trimethoxy-4,5-methylenedioxyisoflavone	[2-[5,6-dimethoxy-3-(6-methoxy-1,3-benzodioxol-5-yl)-4-oxochromen-7-yl]oxy-4,5-dihydroxy-6-(hydroxymethyl)oxan-3-yl] (*Z*)-3-(4-hydroxyphenyl)prop-2-enoate	COC1=CC2=C(C=C1C3=COC4=CC(=C(C(=C4C3=O)OC)OC)OC5C(C(C(C(O5)CO)O)O)OC(=O)C=CC6=CC=C(C=C6)O)OCO2	

7-Hydroxyfluphenazine	10-[3-[4-(2-hydroxyethyl)piperazin-1-yl]propyl]-8-(trifluoromethyl)phenothiazin-3-ol	C1CN(CCN1CCCN2C3=C(C=C(C=C3)O)SC4=C2C=C(C=C4)C(F)(F)F)CCO	
7-Hydroxylpradimicin A	2-[[(5*S*,6*S*)-1,6,7,9,14-pentahydroxy-5-[(2*S*,3*R*,4*S*,5*S*,6*R*)-3-hydroxy-6-methyl-5-(methylamino)-4-[(2*S*,3*R*,4*S*,5*R*)-3,4,5-trihydroxyoxan-2-yl]oxyoxan-2-yl]oxy-11-methoxy-3-methyl-8,13-dioxo-5,6-dihydrobenzo[a]tetracene-2-carbonyl]amino]propanoic acid	CC1C(C(C(C(O1)OC2C(C3=C(C4=C2C=C(C(=C4O)C(=O)NC(C)C(=O)O)O)C)C(=C5C(=C3O)C(=O)C6=C(C5=O)C=C(C=C6O)OC)O)O)O)OC7C(C(C(CO7)O)O)O)NC	
8,9-dihydroxy-5,11,14-eicosatrienoic acid	(5*E*,11*E*,14*E*)-8,9-dihydroxyicosa-5,11,14-trienoic acid	CCCCCC=CCC=CCC(C(CC=CCCCC(=O)O)O)O	
8-Hydroxy-perphenazine glucuronide			
9(S)-HpODE	(9*S*,10*E*,12*Z*)-9-hydroperoxyoctadeca-10,12-dienoic acid	CCCCCC=CC=CC(CCCCCCCC(=O)O)OO	CPD-8677
9,10-12,13-Diepoxyoctadecanoate	8-[3-[(3-pentyloxiran-2-yl)methyl]oxiran-2-yl]octanoic acid	CCCCCC1C(O1)CC2C(O2)CCCCCCCC(=O)O	CPD-13088
Absintholide	(1*S*,2*R*,4*S*,7*S*,8*S*,11*S*,13*S*,14*R*,15*R*,17*S*,20*S*,21*S*,24*S*)-11,13,24-trihydroxy-2,7,11,15,20,24-hexamethyl-5,18-dioxaheptacyclo[13.10.1.02,14.03,12.04,8.016,25.017,21]hexacosa-3(12),16(25)-diene-6,19,26-trione	CC1C2CCC(C3=C(C2OC1=O)C4(C(C3O)C5(C6=C(C4C5=O)C(CCC7C6OC(=O)C7C)(C)O)C)C)(C)O	
Acacetin	5,7-dihydroxy-2-(4-methoxyphenyl)chromen-4-one	COC1=CC=C(C=C1)C2=CC(=O)C3=C(C=C(C=C3O2)O)O	CPD-1095
Acarbose (Glucobay)	(2*R*,3*R*,4*R*,5*S*,6*R*)-5-[(2*R*,3*R*,4*R*,5*S*,6*R*)-5-[(2*R*,3*R*,4*S*,5*S*,6*R*)-3,4-dihydroxy-6-methyl-5-[[(1*S*,4*R*,5*S*,6*S*)-4,5,6-trihydroxy-3-(hydroxymethyl)cyclohex-2-en-1-yl]amino]oxan-2-yl]oxy-3,4-dihydroxy-6-(hydroxymethyl)oxan-2-yl]oxy-6-(hydroxymethyl)oxane-2,3,4-triol	CC1C(C(C(C(O1)OC2C(OC(C(C2O)O)OC3C(OC(C(C3O)O)O)CO)CO)O)O)NC4C=C(C(C(C4O)O)O)CO	
Acetyl vitamin K5	N-(4-hydroxy-3-methylnaphthalen-1-yl)acetamide	CC1=C(C2=CC=CC=C2C(=C1)NC(=O)C)O	
AFMK	N-[3-(2-formamido-5-methoxyphenyl)-3-oxopropyl]acetamide	CC(=O)NCCC(=O)C1=C(C=CC(=C1)OC)NC=O	CPD-12022
Ajugarin I	[(4*R*,4*aR*,5*S*,7*R*,8*S*,8*aR*)-5-acetyloxy-7,8-dimethyl-8-[2-(5-oxo-2*H*-furan-3-yl)ethyl]spiro[2,3,5,6,7,8*a*-hexahydro-1*H*-naphthalene-4,2'-oxirane]-4*a*-yl]methyl acetate	CC1CC(C2(C(C1(C)CCC3=CC(=O)OC3)CCCC24CO4)COC(=O)C)OC(=O)C	
Alkaloid RC	2-(hydroxymethyl)-6-[(13-methyl-5,7,19,21,25-pentaoxa-13-azahexacyclo[12.11.0.02,10.04,8.015,23.018,22]pentacosa-2,4(8),9,15(23),16,18(22)-hexaen-24-yl)oxy]oxane-3,4,5-triol	CN1CCC2=CC3=C(C=C2C4C1C5=C(C(O4)OC6C(C(C(CO6)CO)O)O)O)C7=C(C=C5)OCO7)OCO3	
Allotetrahydrocortisone	(3*R*,5*S*,8*S*,9*S*,10*S*,13*S*,14*S*,17*R*)-3,17-dihydroxy-17-(2-hydroxyacetyl)-10,13-dimethyl-2,3,4,5,6,7,8,9,12,14,15,16-dodecahydro-1*H*-cyclopenta[a]phenanthren-11-one	CC12CCC(CC1CCC3C2C(=O)CC4(C3CCC4(C(=O)CO)O)C)O	
Amoxapine	8-chloro-6-piperazin-1-ylbenzo[b][1,4]benzoxazepine	C1CN(CCN1)C2=NC3=CC=CC=C3OC4=C2C=C(C=C4)Cl	
Antibiotic TA	(12*Z*,14*E*)-16-ethyl-6,8,9-trihydroxy-12-(methoxymethyl)-25,27-dimethyl-2-propyl-1-oxa-4-azacyclooctacosa-12,14-diene-3,20,28-trione	CCCC1C(=O)NCC(CC(C(CCC(=CC=CC(CCCC(=O)CCCCC(CC(C(=O)O1)C)C)CC)COC)O)O)O	
APC			
Apional	(*E*)-3-(4,7-dimethoxy-1,3-benzodioxol-5-yl)prop-2-enal	COC1=C2C(=C(C(=C1)C=CC=O)OC)OCO2	
Aralionine A	(2*S*,3*S*)-*N*-[(3*R*,4*S*,7*S*,10*Z*)-7-benzoyl-5,8-dioxo-3-phenyl-2-oxa-6,9-diazabicyclo[10.2.2]hexadeca-1(14),10,12,15-tetraen-4-yl]-2-(dimethylamino)-3-methylpentanamide	CCC(C)C(C(=O)NC1C(OC2=CC=C(C=C2)C=CNC(=O)C(NC1=O)C(=O)C3=CC=CC=C3)C4=CC=CC=C4)N(C)C	
Arbutin	(2*R*,3*S*,4*S*,5*R*,6*S*)-2-(hydroxymethyl)-6-(4-hydroxyphenoxy)oxane-3,4,5-triol	C1=CC(=CC=C1O)OC2C(C(C(C(O2)CO)O)O)O	HYDROQUINONE-O-BETA-D-GLUCOPYRANOSIDE
Artemetin	2-(3,4-dimethoxyphenyl)-5-hydroxy-3,6,7-trimethoxychromen-4-one	COC1=C(C=C(C=C1)C2=C(C(=O)C3=C(C(=C(C=C3O2)OC)OC)O)OC)OC	
Artonin Q	methyl 7,14-dihydroxy-18,18-dimethyl-21-(3-methylbut-2-enyl)-5,12-dioxo-9-prop-1-en-2-yl-2,19-	CC(=CCC1=C2C(=C(C3=C1OC4=C5C(=O)CC(C5=C(C=C4C3=O)C(	

	dioxapentacyclo[11.8.0.03,11.04,8.015,20]henicosa-1(13),3,8,10,14,16,20-heptaene-7-carboxylate	=C)C)(C(=O)OC)O)O)C=CC(O2)(C)C)C	
Asp Glu Glu	(2*S*)-2-[[(2*S*)-2-[[(2*S*)-2-amino-3-carboxypropanoyl]amino]-4-carboxybutanoyl]amino]pentanedioic acid	C(CC(=O)O)C(C(=O)NC(CCC(=O)O)C(=O)O)NC(=O)C(CC(=O)O)N	
Austalide C	[(1*R*,2*R*,4*S*,16*R*,17*S*,18*R*,20*S*)-18-acetyloxy-13,20-dimethoxy-4,7,17,22,22-pentamethyl-11-oxo-5,10,21,23-tetraoxahexacyclo[18.2.1.01,17.04,16.06,14.08,12]tricosa-6(14),7,12-trien-2-yl] acetate	CC1=C2COC(=O)C2=C(C3=C1OC4(CC(C56C(OC(O5)(CC(C6(C4C3)C)OC(=O)C)OC)(C)C)OC(=O)C)C)OC	
Bambuterol	[3-[2-(*tert*-butylamino)-1-hydroxyethyl]-5-(dimethylcarbamoyloxy)phenyl] *N*,*N*-dimethylcarbamate	CC(C)(C)NCC(C1=CC(=CC(=C1)OC(=O)N(C)C)OC(=O)N(C)C)O	
Beauvericin	(3*S*,6*R*,9*S*,12*R*,15*S*,18*R*)-3,9,15-tribenzyl-4,10,16-trimethyl-6,12,18-tri(propan-2-yl)-1,7,13-trioxa-4,10,16-triazacyclooctadecane-2,5,8,11,14,17-hexone	CC(C)C1C(=O)N(C(C(=O)OC(C(=O)N(C(C(=O)OC(C(=O)N(C(C(=O)O1)CC2=CC=CC=C2)C)C(C)CC3=CC=CC=C3)C)C(C)C)CC4=CC=CC=C4)C	
beta-Bixin	(2*E*,4*E*,6*E*,8*E*,10*E*,12*E*,14*E*,16*E*,18*E*)-20-methoxy-4,8,13,17-tetramethyl-20-oxoicosa-2,4,6,8,10,12,14,16,18-nonaenoic acid	CC(=CC=CC=C(C)C=CC=C(C)C=CC(=O)OC)C=CC=C(C)C=CC(=O)O	
Betavulgaroside VI	4-[2-carboxy-1-(carboxymethoxy)-2-hydroxyethoxy]-3,5-dihydroxy-6-[[4-(hydroxymethyl)-4,6a,6b,11,11,14b-hexamethyl-8a-[3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxycarbonyl-1,2,3,4a,5,6,7,8,9,10,12,12a,14,14a-tetradecahydropicen-3-yl]oxy]oxane-2-carboxylic acid	CC1(CCC2(CCC3(C(=CCC4C3(CCC5C4(CCC(C5(C)CO)OC6C(C(C(C(O6)C(=O)O)O)OC(C(C(=O)O)O)OCC(=O)O)O)C)C)C2C1)C)C(=O)OC7C(C(C(C(O7)CO)O)O)O)C	
Biflorin	5,7-dihydroxy-2-methyl-6-[(2*S*,3*R*,4*R*,5*S*,6*R*)-3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]chromen-4-one	CC1=CC(=O)C2=C(O1)C=C(C(=C2O)C3C(C(C(C(O3)CO)O)O)O)O	
Bikhaconitine	[(1*S*,2*R*,3*R*,4*R*,5*S*,6*S*,8*R*,9*R*,13*S*,16*S*,17*R*,18*R*)-8-acetyloxy-11-ethyl-5-hydroxy-6,16,18-trimethoxy-13-(methoxymethyl)-11-azahexacyclo[7.7.2.12,5.01,10.03,8.013,17]nonadecan-4-yl] 3,4-dimethoxybenzoate	CCN1CC2(CCC(C34C2C(C(C31)C5(CC(C6(CC4C5C6OC(=O)C7=CC(=C(C=C7)OC)OC)O)OC)OC(=O)C)OC)OC)COC	
Biochanin	5,7-dihydroxy-3-(4-methoxyphenyl)chromen-4-one	COC1=CC=C(C=C1)C2=COC3=CC(=CC(=C3C2=O)O)O	
Bis(glutathionyl)spermine disulfide	(2*S*)-2-amino-5-[[(4*R*,27*R*)-27-[[(4*S*)-4-amino-4-carboxybutanoyl]amino]-5,8,23,26-tetraoxo-1,2-dithia-6,9,13,18,22,25-hexazacyclooctacos-4-yl]amino]-5-oxopentanoic acid	C1CCNCCCNC(=O)CNC(=O)C(CSSCC(C(=O)NCC(=O)NCCCNC1)NC(=O)CCC(C(=O)O)N)NC(=O)CCC(C(=O)O)N	
Bisosthenon B	8-[2,3-diacetyl-4-(7-methoxy-2-oxochromen-8-yl)cyclobutyl]-7-methoxychromen-2-one	CC(=O)C1C(C(C1C(=O)C)C2=C(C=CC3=C2OC(=O)C=C3)OC)C4=C(C=CC5=C4OC(=O)C=C5)OC	
BL II	[3,6,8,9-tetraacetyloxy-7-(4-hydroxyphenyl)dibenzofuran-2-yl] acetate	CC(=O)OC1=C(C=C2C(=C1)C3=C(O2)C(=C(C(=C3OC(=O)C)OC(=O)C)C4=CC=C(C=C4)O)OC(=O)C)OC(=O)C	
C-12 NBD Ceramide	N-[(*E*,2*S*,3*R*)-1,3-dihydroxyoctadec-4-en-2-yl]-12-[(4-nitro-2,1,3-benzoxadiazol-7-yl)amino]dodecanamide	CCCCCCCCCCCCCCC=CC(C(CO)NC(=O)CCCCCCCCCCCNC1=CC=C(C2=NON=C12)[N+](=O)[O-])O	
Calpeptin	benzyl *N*-[(2*S*)-4-methyl-1-oxo-1-[[(2*S*)-1-oxohexan-2-yl]amino]pentan-2-yl]carbamate	CCCCC(C=O)NC(=O)C(CC(C)C)NC(=O)OCC1=CC=CC=C1	
Cassinine	[(1*S*,3*R*,13*R*,18*R*,19*S*,20*R*,21*R*,22*S*,23*R*,24*R*,25*R*,26*R*)-19,20,22,23-tetraacetyloxy-21-(acetyloxymethyl)-13-ethyl-3,26-dimethyl-6,16-dioxo-2,5,17-trioxa-11-azapentacyclo[16.7.1.01,21.03,24.07,12]hexacosa-7(12),8,10-trien-25-yl] benzoate	CCC1CCC(=O)OC2C(C34C(C(C(C(C3(C(C2OC(=O)C)OC(=O)C)COC(=O)C)OC(=O)C)OC(=O)C)C(O4)(COC(=O)C5=C1N=CC=C5)C)OC(=O)C6=CC=CC=C6)C	
CAY10500			
CAY10598			
CE(16:1(9Z))	[(3*S*,8*S*,9*S*,10*R*,13*R*,14*S*,17*R*)-10,13-dimethyl-17-[(2*R*)-6-methylheptan-2-yl]-2,3,4,7,8,9,11,12,14,15,16,17-dodecahydro-1*H*-cyclopenta[a]phenanthren-3-yl] (*Z*)-hexadec-9-enoate	CCCCCCC=CCCCCCCCC(=O)OC1CCC2(C3CCC4(C(C3CC=C2C1)CCC4C(C)CCCC(C)C)C)C	
CE(19:0)	[(3*S*,8*S*,9*S*,10*R*,13*R*,14*S*,17*R*)-10,13-dimethyl-17-[(2*R*)-6-methylheptan-2-yl]-2,3,4,7,8,9,11,12,14,15,16,17-dodecahydro-1*H*-cyclopenta[a]phenanthren-3-yl] nonadecanoate	CCCCCCCCCCCCCCCCCCCC(=O)OC1CCC2(C3CCC4(C(C3CC=C2C1)CCC4C(C)CCCC(C)C)C)C	
Chartreusin	3-[(2*S*,3*R*,4*S*,5*R*,6*R*)-3-[(2*R*,3*R*,4*S*,5*S*,6*R*)-3,5-dihydroxy-4-methoxy-6-methyloxan-2-yl]oxy-4,5-dihydroxy-6-methyloxan-2-yl]oxy-8-hydroxy-15-methyl-11,18-	CC1C(C(C(C(O1)OC2=CC=CC3=C2C4=C5C6=C(C=CC(=C6C(=O)	

	dioxapentacyclo[10.6.2.02,7.09,19.016,20]icosa-1(19),2(7),3,5,8,12(20),13,15-octaene-10,17-dione	O4)C)OC(=O)C5=C3O)OC7C(C(C(C(O7)C)O)OC)O)O)O	
Chitin			CHITIN
Chrysophanol	1,8-dihydroxy-3-methylanthracene-9,10-dione	CC1=CC2=C(C(=C1)O)C(=O)C3=C(C2=O)C=CC=C3O	
Cinnamtannin B2	(1R,5R,6R,7S,13R,21R)-5,13-bis(3,4-dihydroxyphenyl)-7-[(2R,3R)-2-(3,4-dihydroxyphenyl)-3,5,7-trihydroxy-3,4-dihydro-2H-chromen-8-yl]-16-[(2R,3R,4R)-2-(3,4-dihydroxyphenyl)-3,5,7-trihydroxy-3,4-dihydro-2H-chromen-4-yl]-4,12,14-trioxapentacyclo[11.7.1.02,11.03,8.015,20]henicosa-2(11),3(8),9,15,17,19-hexaene-6,9,17,19,21-pentol	C1C(C(OC2=C1C(=CC(=C2C3C(C(OC4=C3C(=CC5=C4C6C(C(O5)(OC7=C(C(=CC(=C67)O)O)C8C(C(OC9=CC(=CC(=C89)O)O)C1=CC(=C(C=C1)O)O)O)C1=CC(=C(C=C1)O)O)O)O)C1=CC(=C(C=C1)O)O)O)O)C1=CC(=C(C=C1)O)O)O	
Cinncassiol D4	(1*R*,2*S*,3*S*,4*S*,6*R*,7*R*,10*S*,11*S*,13*S*,14*S*)-3,7,10-trimethyl-11-propan-2-yl-15-oxapentacyclo[7.5.1.02,6.07,13.010,14]pentadecane-4,6,9,14-tetrol	CC1C(CC2(C1C3C4(C5C2(CC(C4(C(C5)C(C)C)C)(O3)O)C)O)O)O	
Citbrasine	1,5-dihydroxy-2,3,4-trimethoxy-10-methylacridin-9-one	CN1C2=C(C=CC=C2O)C(=O)C3=C1C(=C(C(=C3O)OC)OC)OC	
Citrusin B	(2*S*,3*R*,4*S*,5*S*,6*R*)-2-[4-[1,3-dihydroxy-2-[4-[(*E*)-3-hydroxyprop-1-enyl]-2,6-dimethoxyphenoxy]propyl]-2-methoxyphenoxy]-6-(hydroxymethyl)oxane-3,4,5-triol	COC1=CC(=CC(=C1OC(CO)C(C2=CC(=C(C=C2)OC3C(C(C(C(O3)CO)O)O)O)O)OC)O)OC)C=CCO	
Corchorusoside E	3-[(3S,5S,8R,9S,10R,11R,13R,14S,17R)-5,11,14-trihydroxy-3-[(2R,4S,5S,6R)-4-hydroxy-6-methyl-5-[(2S,3R,4S,5S,6R)-3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxyoxan-2-yl]oxy-10,13-dimethyl-2,3,4,6,7,8,9,11,12,15,16,17-dodecahydro-1H-cyclopenta[a]phenanthren-17-yl]-2H-furan-5-one	CC1C(C(CC(O1)OC2CCC3(C4C(CCC3(C2)O)C5(CCC(C5(CC4O)C)C6=CC(=O)OC6)O)C)O)OC7C(C(C(C(O7)CO)O)O)O	
Coriandrone E	(7*S*,8*R*)-8-hydroxy-4-methoxy-7-methyl-7,8-dihydrofuro[2,3-g]isochromen-5-one	CC1C(C2=CC3=C(C=CO3)C(=C2C(=O)O1)OC)O	
Cucurbitacin D	(2*S*,8*S*,9*R*,10*R*,13*R*,14*S*,16*R*,17*R*)-17-[(*E*,2*R*)-2,6-dihydroxy-6-methyl-3-oxohept-4-en-2-yl]-2,16-dihydroxy-4,4,9,13,14-pentamethyl-2,7,8,10,12,15,16,17-octahydro-1*H*-cyclopenta[a]phenanthrene-3,11-dione	CC1(C2=CCC3C4(CC(C(C4(CC(=O)C3(C2CC(C1=O)O)C)C)C(C)(C(=O)C=CC(C)(C)O)O)O)C)C	
Curacin D	4-[(1*Z*,5*E*,7*E*)-11-methoxytetradeca-1,5,7,13-tetraenyl]-2-[(1*R*,2*S*)-2-methylcyclopropyl]-4,5-dihydro-1,3-thiazole	CC1CC1C2=NC(CS2)C=CCCC=CC=CCCC(CC=C)OC	
Debrisoquine	3,4-dihydro-1*H*-isoquinoline-2-carboximidamide	C1CN(CC2=CC=CC=C21)C(=N)N	
Dehydroisocoproporphyrin	3-[13,18-bis(2-carboxyethyl)-17-(carboxymethyl)-8-ethenyl-3,7,12-trimethyl-22,24-dihydroporphyrin-2-yl]propanoic acid	CC1=C(C2=CC3=NC(=CC4=C(C(=C(N4)C=C5C(=C(C(=N5)C=C1N2)C)CCC(=O)O)CCC(=O)O)CC(=O)O)C(=C3C)CCC(=O)O)C=C	
Deltorphin C	(3*S*)-3-[[(2*S*)-2-[[(2*R*)-2-[[(2*S*)-2-amino-3-(4-hydroxyphenyl)propanoyl]amino]propanoyl]amino]-3-phenylpropanoyl]amino]-4-[[(2*S*)-1-[[(2*S*)-1-[(2-amino-2-oxoethyl)amino]-3-methyl-1-oxobutan-2-yl]amino]-3-methyl-1-oxobutan-2-yl]amino]-4-oxobutanoic acid	CC(C)C(C(=O)NC(C(C)C)C(=O)NCC(=O)N)NC(=O)C(CC(=O)O)NC(=O)C(CC1=CC=CC=C1)NC(=O)C(C)NC(=O)C(CC2=CC=C(C=C2)O)N	
Deoxyvasicinone	2,3-dihydro-1*H*-pyrrolo[2,1-b]quinazolin-9-one	C1CC2=NC3=CC=CC=C3C(=O)N2C1	
Derricin	(*E*)-1-[2-hydroxy-4-methoxy-3-(3-methylbut-2-enyl)phenyl]-3-phenylprop-2-en-1-one	CC(=CCC1=C(C=CC(=C1O)C(=O)C=CC2=CC=CC=C2)OC)C	
DHOAA(AcO)-Val-Phe(NMe)-Pro-Phe(NMe)-Gly-OMe			
Dialicor	1-[2-[2-(diethylamino)ethoxy]phenyl]-3-phenylpropan-1-one;hydrochloride	CCN(CC)CCOC1=CC=CC=C1C(=O)CCC2=CC=CC=C2.Cl	
Diaziquone	ethyl *N*-[2,5-bis(aziridin-1-yl)-4-(ethoxycarbonylamino)-3,6-dioxocyclohexa-1,4-dien-1-yl]carbamate	CCOC(=O)NC1=C(C(=O)C(=C(C1=O)N2CC2)NC(=O)OCC)N3CC3	
Dihydroergocristine	(6*aR*,9*R*,10*aR*)-*N*-[(1*S*,2*S*,4*R*,7*S*)-7-benzyl-2-hydroxy-5,8-dioxo-4-propan-2-yl-3-oxa-6,9-diazatricyclo[7.3.0.02,6]dodecan-4-yl]-7-methyl-6,6*a*,8,9,10,10*a*-hexahydro-4*H*-indolo[4,3-fg]quinoline-9-carboxamide	CC(C)C1(C(=O)N2C(C(=O)N3CCCC3C2(O1)O)CC4=CC=CC=C4)NC(=O)C5CC6C(CC7=CNC8=CC=CC6=C78)N(C5)C	
Diphenylcarbazide	1,3-dianilinourea	C1=CC=C(C=C1)NNC(=O)NNC2=CC=CC=C2	
D-NMAPPD	N-[(1*R*,2*R*)-1,3-dihydroxy-1-(4-nitrophenyl)propan-2-yl]tetradecanamide	CCCCCCCCCCCCCC(=O)NC(CO)C(C1=CC=C(C=C1)[N+](=O)[O-])O	

D-Sorbitol	(2*R*,3*R*,4*R*,5*S*)-hexane-1,2,3,4,5,6-hexol	C(C(C(C(C(CO)O)O)O)O)O	SORBITOL
Duartin (-)	(3*S*)-3-(3-hydroxy-2,4-dimethoxyphenyl)-8-methoxy-3,4-dihydro-2*H*-chromen-7-ol	COC1=C(C(=C(C=C1)C2CC3=C(C(=C(C=C3)O)OC)OC2)OC)O	
Dukunolide E	15-(furan-3-yl)-3,11-dihydroxy-5,5,10,16-tetramethyl-9,14,20-trioxahexacyclo[10.8.1.01,19.03,11.06,10.016,21]henicos-12(21)-ene-4,8,13-trione	CC1(C2CC(=O)OC2(C3(C4=C5C(CCC6C5(O6)CC3(C1=O)O)(C(OC4=O)C7=COC=C7)C)O)C)C	
Eicosanoyl-EA	N-(2-hydroxyethyl)icosanamide	CCCCCCCCCCCCCCCCCCCC(=O)NCCO	
Eicosapentaenoic Acid-d5	(5*Z*,8*Z*,11*Z*,14*Z*,17*Z*)-19,19,20,20,20-pentadeuterioicosa-5,8,11,14,17-pentaenoic acid	CCC=CCC=CCC=CCC=CCC=CCCCC(=O)O	
Enkephaline, (D-Ala)2-Leu			
Enterocin 900	1-[2-[4-(2-phenyl-6,7,8,9-tetrahydro-5*H*-cyclohepta[f][1]benzofuran-3-yl)phenoxy]ethyl]pyrrolidine	C1CCC2=CC3=C(C=C2CC1)OC(=C3C4=CC=C(C=C4)OCCN5CCCC5)C6=CC=CC=C6	
Epicatechin	(2*R*,3*R*)-2-(3,4-dihydroxyphenyl)-3,4-dihydro-2*H*-chromene-3,5,7-triol	C1C(C(OC2=CC(=CC(=C21)O)O)C3=CC(=C(C=C3)O)O)O	CPD-7630
Ergosine	(6*aR*,9*R*)-*N*-[(1*S*,2*S*,4*R*,7*S*)-2-hydroxy-4-methyl-7-(2-methylpropyl)-5,8-dioxo-3-oxa-6,9-diazatricyclo[7.3.0.02,6]dodecan-4-yl]-7-methyl-6,6*a*,8,9-tetrahydro-4*H*-indolo[4,3-fg]quinoline-9-carboxamide	CC(C)CC1C(=O)N2CCCC2C3(N1C(=O)C(O3)(C)NC(=O)C4CN(C5CC6=CNC7=CC=CC(=C67)C5=C4)C)O	
Esmeraldin B	11-[1-(2-hydroxy-6-methylbenzoyl)oxyethyl]-16-methyl-1,6,13,22-tetrazaheptacyclo[15.11.1.02,15.05,14.07,12.021,29.023,28]nonacosa-2(15),3,5,7,9,11,13,17(29),18,20,23,25,27-tridecaene-4,24-dicarboxylic acid	CC1C2=C3C(=CC=C2)NC4=C(C=CC=C4N3C5=C1C6=NC7=C(C=CC=C7N=C6C(=C5)C(=O)O)C(C)OC(=O)C8=C(C=CC=C8O)C)C(=O)O	
Estradiol mustard	[(8*R*,9*S*,13*S*,14*S*,17*S*)-3-[2-[4-[bis(2-chloroethyl)amino]phenyl]acetyl]oxy-13-methyl-6,7,8,9,11,12,14,15,16,17-decahydrocyclopenta[a]phenanthren-17-yl] 2-[4-[bis(2-chloroethyl)amino]phenyl]acetate	CC12CCC3C(C1CCC2OC(=O)CC4=CC=C(C=C4)N(CCCl)CCCl)CCC5=C3C=CC(=C5)OC(=O)CC6=CC=C(C=C6)N(CCCl)CCCl	
Falaconitine	[(1*S*,2*R*,3*S*,4*R*,5*S*,6*S*,9*R*,13*R*,14*R*,16*S*,17*S*,18*R*)-11-ethyl-5,14-dihydroxy-6,16,18-trimethoxy-13-(methoxymethyl)-11-azahexacyclo[7.7.2.12,5.01,10.03,8.013,17]nonadec-7-en-4-yl] 3,4-dimethoxybenzoate	CCN1CC2(C(CC(C34C2C(C(C31)C5=CC(C6(CC4C5C6OC(=O)C7=CC(=C(C=C7)OC)OC)O)OC)OC)OC)OC)O)COC	
Famciclovir	[2-(acetyloxymethyl)-4-(2-aminopurin-9-yl)butyl] acetate	CC(=O)OCC(CCN1C=NC2=CN=C(N=C21)N)COC(=O)C	
farnesyl triphosphate	[hydroxy(phosphonooxy)phosphoryl] [(2*E*,6*E*)-3,7,11-trimethyldodeca-2,6,10-trienyl] hydrogen phosphate	CC(=CCCC(=CCCC(=CCOP(=O)(O)OP(=O)(O)OP(=O)(O)O)C)C)C	
FIPI	5-fluoro-*N*-[2-[4-(2-oxo-3*H*-benzimidazol-1-yl)piperidin-1-yl]ethyl]-1*H*-indole-2-carboxamide	C1CN(CCC1N2C3=CC=CC=C3NC2=O)CCNC(=O)C4=CC5=C(N4)C=CC(=C5)F	
Foeniculoside IV	(2S,3R,4S,5S,6R)-2-[3-hydroxy-5-[(2R,3R)-2-(4-hydroxyphenyl)-4-[(2S,3S)-2-(4-hydroxyphenyl)-4-[(E)-2-(4-hydroxyphenyl)ethenyl]-6-[(2S,3R,4S,5S,6R)-3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxy-2,3-dihydro-1-benzofuran-3-yl]-6-[(2S,3R,4S,5S,6R)-3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxy-2,3-dihydro-1-benzofuran-3-yl]phenoxy]-6-(hydroxymethyl)oxane-3,4,5-triol	C1=CC(=CC=C1C=CC2=C3C(C(OC3=CC(=C2)OC4C(C(C(C(O4)CO)O)O)O)O)C5=CC=C(C=C5)O)C6=C7C(C(OC7=CC(=C6)OC8C(C(C(C(O8)CO)O)O)O)O)C9=CC=C(C=C9)O)C1=CC(=CC(=C1)OC1C(C(C(C(O1)CO)O)O)O)O)O	
Formononetin	7-hydroxy-3-(4-methoxyphenyl)chromen-4-one	COC1=CC=C(C=C1)C2=COC3=C(C2=O)C=CC(=C3)O	FORMONONETIN
Fulvine	(1*R*,4*R*,5*S*,6*S*,16*R*)-5-hydroxy-4,5,6-trimethyl-2,8-dioxa-13-azatricyclo[8.5.1.013,16]hexadec-10-ene-3,7-dione	CC1C(=O)OCC2=CCN3C2C(CC3)OC(=O)C(C1(C)O)C	
Gal?(1-3)[Fuc?(1-2)]Gal?(1-4)Glc?-Sp			
GalNAc?1-3[Fuc?1-2]Gal?1-3GlcNAc?-Sp			
Gazer	2-[2-[5-[2-[(E)-4-[[9-[3,4-dihydroxy-5-(hydroxymethyl)oxolan-2-yl]purin-6-yl]amino]-2-methylbut-2-enoxy]-3-hydroxy-6-(hydroxymethyl)-5-(3,4,5-trihydroxyoxan-2-yl)oxyoxan-4-yl]oxy-4-hydroxy-2-(hydroxymethyl)oxolan-3-yl]oxy-4,5-dihydroxy-6-(hydroxymethyl)oxan-3-yl]oxy-6-(hydroxymethyl)oxane-3,4,5-triol	CC(=CCNC1=C2C(=NC=N1)N(C=N2)C3C(C(C(O3)CO)O)O)COC4C(C(C(C(O4)CO)OC5C(C(C(CO5)O)O)O)OC6C(C(C(O6)CO)OC7C(C(C(C(O7)CO)O)O)OC8C(C(C(C(O8)CO)O)O)O)O)O	
Gln Phe Trp	(2*S*)-2-[[(2*S*)-2-[[(2*S*)-2,5-diamino-5-oxopentanoyl]amino]-3-phenylpropanoyl]amino]-3-(1*H*-indol-3-yl)propanoic acid	C1=CC=C(C=C1)CC(C(=O)NC(CC2=CNC3=CC=CC=C32)C(=O)O)NC(=O)C(CCC(=O)N)N	

Glucoconvallatoxoloside	3-[3-[5-[3,4-dihydroxy-6-(hydroxymethyl)-5-[3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxyoxan-2-yl]oxy-3,4-dihydroxy-6-methyloxan-2-yl]oxy-5,14-dihydroxy-10-(hydroxymethyl)-13-methyl-2,3,4,6,7,8,9,11,12,15,16,17-dodecahydro-1H-cyclopenta[a]phenanthren-17-yl]-2H-furan-5-one	CC1C(C(C(C(O1)OC2CCC3(C4CCC5(C(CCC5(C4CCC3(C2)O)O)C6=CC(=O)OC6)C)CO)O)O)OC7C(C(C(C(O7)CO)OC8C(C(C(C(O8)CO)O)O)O)O)O)O	
Gnididin	[(1*R*,2*R*,6*S*,7*S*,8*R*,10*S*,11*S*,12*R*,16*S*,17*R*,18*R*)-6,7-dihydroxy-8-(hydroxymethyl)-4,18-dimethyl-5-oxo-14-phenyl-16-prop-1-en-2-yl-9,13,15,19-tetraoxahexacyclo[12.4.1.01,11.02,6.08,10.012,16]nonadec-3-en-17-yl] (2*E*,4*E*)-deca-2,4-dienoate	CCCCCC=CC=CC(=O)OC1C(C23C4C=C(C(=O)C4(C(C5(C(C2C6C1(OC(O6)(O3)C7=CC=CC=C7)C(=C)C)O5)CO)O)O)C)C	
Gossypetin	2-(3,4-dihydroxyphenyl)-3,5,7,8-tetrahydroxychromen-4-one	C1=CC(=C(C=C1C2=C(C(=O)C3=C(O2)C(=C(C=C3O)O)O)O)O)O	334578-HEXAHYDROXYFLAVONE
Helieianeoside C	2-(3,4-dihydroxyphenyl)-5,7-dihydroxy-3-[(2*S*,3*S*,5*R*)-4-hydroxy-6-methyl-5-[(2*S*,4*S*,5*R*)-3,4,5-trihydroxyoxan-2-yl]oxy-3-[(2*S*,4*S*,5*S*)-3,4,5-trihydroxy-6-[[(2*S*,4*S*,5*R*)-3,4,5-trihydroxyoxan-2-yl]oxymethyl]oxan-2-yl]oxyoxan-2-yl]oxychromen-4-one	CC1C(C(C(C(O1)OC2=C(OC3=CC(=CC(=C3C2=O)O)O)C4=CC(=C(C=C4)O)O)OC5C(C(C(C(O5)COC6C(C(C(CO6)O)O)O)O)O)O)OC7C(C(C(CO7)O)O)O	
hentriacontan-16-ol	hentriacontan-16-ol	CCCCCCCCCCCCCCCC(CCCCCCCCCCCCCCC)O	
Heptacarboxylporphyrin III	3-[7,8,12,17-tetrakis(2-carboxyethyl)-13,18-bis(carboxymethyl)-3-methyl-22,23-dihydroporphyrin-2-yl]propanoic acid	CC1=C(C2=CC3=NC(=CC4=C(C(=C(N4)C=C5C(=C(C(=CC1=N2)N5)CCC(=O)O)CCC(=O)O)CCC(=O)O)O)CC(=O)O)C(=C3CC(=O)O)CCC(=O)O)CCC(=O)O	
Hexandraside C	[(2R,3R,4S,5R,6S)-6-[(2S,3R,4R,5S,6S)-3,5-dihydroxy-2-[5-hydroxy-2-(4-methoxyphenyl)-8-(3-methylbut-2-enyl)-4-oxochromen-7-yl]oxy-6-methyloxan-4-yl]oxy-3,4,5-trihydroxyoxan-2-yl]methyl acetate	CC1C(C(C(C(O1)OC2=C(OC3=C(C2=O)C(=CC(=C3CC=C(C)C)OC4C(C(C(C(O4)CO)O)O)O)O)C5=CC=C(C=C5)OC)O)OC6C(C(C(C(O6)COC(=O)C)O)O)O)O	
His Gly Ser			
Hoiamide B	(5R,9S,12S,13S,14R,17S,20S,23R,24S,25S)-12,17-bis[(2S)-butan-2-yl]-13-hydroxy-20-[(1R)-1-hydroxyethyl]-23-[(2S,3S,4R)-3-hydroxy-4-methylheptan-2-yl]-25-methoxy-5,9,14,24-tetramethyl-16,22-dioxa-3,7,28-trithia-11,19,30,31,32-pentazatetracyclo[25.2.1.12,5.16,9]dotriaconta-1(29),2(32),6(31),27(30)-tetraene-10,15,18,21-tetrone	CCCC(C)C(C(C)C1C(C(CC2=NC(=CS2)C3=NC(CS3)(C4=NC(CS4)(C(=O)NC(C(C(C(=O)OC(C(=O)NC(C(=O)O1)C(C)O)C(C)CC)C)O)C(C)CC)C)C)OC)C)O	
Hydrojuglone glucoside	2-(5,8-dihydroxynaphthalen-1-yl)oxy-6-(hydroxymethyl)oxane-3,4,5-triol	C1=CC2=C(C=CC(=C2C(=C1)OC3C(C(C(C(O3)CO)O)O)O)O)O	
Hydroxy-3-O-methyl-6?-naltrexol			
Hydroxyanastrozole			
Hydroxydestruxin B	16-butan-2-yl-3-(2-hydroxy-2-methylpropyl)-10,11,14-trimethyl-13-propan-2-yl-4-oxa-1,8,11,14,17-pentazabicyclo[17.3.0]docosane-2,5,9,12,15,18-hexone	CCC(C)C1C(=O)N(C(C(=O)N(C(C(=O)NCCC(=O)OC(C(=O)N2CCCC2C(=O)N1)CC(C)(C)O)C)C)C(C)C)C	
Hydroxyitraconazole	4-[4-[4-[4-[[2-(2,4-dichlorophenyl)-2-(1,2,4-triazol-1-ylmethyl)-1,3-dioxolan-4-yl]methoxy]phenyl]piperazin-1-yl]phenyl]-2-(3-hydroxybutan-2-yl)-1,2,4-triazol-3-one	CC(C(C)O)N1C(=O)N(C=N1)C2=CC=C(C=C2)N3CCN(CC3)C4=CC=C(C=C4)OCC5COC(O5)(CN6C=NC=N6)C7=C(C=C(C=C7)Cl)Cl	
Hydroxyphthiocerani c acid (C37)	(2*S*,4*S*,6*S*,8*S*,10*R*,12*R*,14*R*)-15-hydroxy-2,4,6,8,10,12,14-heptamethyltriacontanoic acid	CCCCCCCCCCCCCCCCC(C(C)CC(C)CC(C)CC(C)CC(C)CC(C)CC(C)C(=O)O)O	
Ibudilast	2-methyl-1-(2-propan-2-ylpyrazolo[1,5-a]pyridin-3-yl)propan-1-one	CC(C)C1=NN2C=CC=CC2=C1C(=O)C(C)C	
Ikarisoside D	[(3*R*,5*S*,6*S*)-6-[5,7-dihydroxy-2-(4-hydroxyphenyl)-8-(3-methylbut-2-enyl)-4-oxochromen-3-yl]oxy-4,5-dihydroxy-2-methyloxan-3-yl] acetate	CC1C(C(C(C(O1)OC2=C(OC3=C(C(=CC(=C3C2=O)O)O)CC=C(C)C)C4=CC=C(C=C4)O)O)O)OC(=O)C	
Indaconitine	[(1*S*,2*R*,3*R*,4*R*,5*S*,6*S*,8*R*,9*R*,13*R*,14*R*,16*S*,17*S*,18*R*)-8-acetyloxy-11-ethyl-5,14-dihydroxy-6,16,18-trimethoxy-13-(methoxymethyl)-11-	CCN1CC2(C(CC(C34C2C(C(C31)C5(CC(C6(CC4C5C6OC(=O)C7=CC=C	

	azahexacyclo[7.7.2.12,5.01,10.03,8.013,17]nonadecan-4-yl] benzoate	C=C7)O)OC)OC(=O)C)OC)OC)O)COC	
Iretol	2-methoxybenzene-1,3,5-triol	COC1=C(C=C(C=C1O)O)O	
Isatidine	(1*R*,4*Z*,6*R*,7*S*,17*R*)-4-ethylidene-7-hydroxy-7-(hydroxymethyl)-6-methyl-14-oxido-2,9-dioxa-14-azoniatricyclo[9.5.1.014,17]heptadec-11-ene-3,8-dione	CC=C1CC(C(C(=O)OCC2=CC[N+]3(C2C(CC3)OC1=O)[O-])(CO)O)C	
Isodonal			
Isogermafurene	6-ethenyl-3,6-dimethyl-5-prop-1-en-2-yl-5,7-dihydro-4*H*-1-benzofuran	CC1=COC2=C1CC(C(C2)(C)C=C)C(=C)C	
Isohopeaphenol	(1R,8S,9R,16R)-8,16-bis(4-hydroxyphenyl)-9-[(1R,8S,9R,16R)-4,6,12-trihydroxy-8,16-bis(4-hydroxyphenyl)-15-oxatetracyclo[8.6.1.02,7.014,17]heptadeca-2(7),3,5,10(17),11,13-hexaen-9-yl]-15-oxatetracyclo[8.6.1.02,7.014,17]heptadeca-2(7),3,5,10(17),11,13-hexaene-4,6,12-triol	C1=CC(=CC=C1C2C(C3=C4C(C(OC4=CC(=C3)O)C5=CC=C(C=C5)O)C6=C2C(=CC(=C6)O)O)C7C(C8=C(C=C(C=C8O)O)C9C(OC1=CC(=CC7=C91)O)C1=CC=C(C=C1)O)C1=CC=C(C=C1)O)O	
Isorhamnetin	3,5,7-trihydroxy-2-(4-hydroxy-3-methoxyphenyl)chromen-4-one	COC1=C(C=CC(=C1)C2=C(C(=O)C3=C(C=C(C=C3O2)O)O)O)O	CPD-8004
Isoscoparin	5,7-dihydroxy-2-(4-hydroxy-3-methoxyphenyl)-6-[(2*S*,3*R*,4*R*,5*S*,6*R*)-3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]chromen-4-one	COC1=C(C=CC(=C1)C2=CC(=O)C3=C(O2)C=C(C(=C3O)C4C(C(C(C(O4)CO)O)O)O)O)O	
Isoswertisin 3-O-(2-methylbutyrate)	[(2*S*,4*R*,5*R*)-3,5-dihydroxy-2-[5-hydroxy-2-(4-hydroxyphenyl)-7-methoxy-4-oxochromen-8-yl]-6-(hydroxymethyl)oxan-4-yl] 2-methylbutanoate	CCC(C)C(=O)OC1C(C(OC(C1O)C2=C(C=C(C3=C2OC(=CC3=O)C4=CC=C(C=C4)O)O)OC)CO)O	
Italidipyrone	6-ethyl-3-[[3-[(6-ethyl-4-hydroxy-5-methyl-2-oxopyran-3-yl)methyl]-2,4,6-trihydroxy-5-(2-methylbutanoyl)phenyl]methyl]-4-hydroxy-5-methylpyran-2-one	CCC1=C(C(=C(C(=O)O1)CC2=C(C(=C(C(=C2O)C(=O)C(C)CC)O)CC3=C(C(=C(OC3=O)CC)C)O)O)O)C	
Janex-1	4-[(6,7-dimethoxyquinazolin-4-yl)amino]phenol	COC1=C(C=C2C(=C1)C(=NC=N2)NC3=CC=C(C=C3)O)OC	
Kaempferol	3,5,7-trihydroxy-2-(4-hydroxyphenyl)chromen-4-one	C1=CC(=CC=C1C2=C(C(=O)C3=C(C=C(C=C3O2)O)O)O)O	CPD1F-90
Kamahine C	1,5'-dihydroxy-2,4',8,8-tetramethylspiro[6-oxabicyclo[3.2.1]oct-2-ene-7,2'-oxolane]-4-one	CC1CC2(C3(C(=CC(=O)C(C3(C)C)O2)C)O)OC1O	
Kansuinine B	[(1*S*,3*S*,4*S*,6*S*,9*R*,10*S*,12*R*,13*S*,14*R*,15*R*,16*S*)-1,14-diacetyloxy-10-benzoyloxy-9,15,16-trihydroxy-3,7,7,15-tetramethyl-11-methylidene-2,8-dioxo-5-oxatricyclo[11.3.0.04,6]hexadecan-12-yl] benzoate	CC1C2C(O2)C(C(=O)C(C(C(=C)C(C3C(C(C(C3(C1=O)OC(=O)C)O)(C)O)OC(=O)C)OC(=O)C4=CC=CC=C4)OC(=O)C5=CC=CC=C5)O)(C)C	
Kni 102	benzyl *N*-[(2*S*)-4-amino-1-[[(2*S*,3*S*)-4-[(2*S*)-2-(*tert*-butylcarbamoyl)pyrrolidin-1-yl]-3-hydroxy-4-oxo-1-phenylbutan-2-yl]amino]-1,4-dioxobutan-2-yl]carbamate	CC(C)(C)NC(=O)C1CCCN1C(=O)C(C(CC2=CC=CC=C2)NC(=O)C(CC(=O)N)NC(=O)OCC3=CC=CC=C3)O	
Kuwanone H	8-[(1*S*,5*R*,6*S*)-6-[2,4-dihydroxy-3-(3-methylbut-2-enyl)benzoyl]-5-(2,4-dihydroxyphenyl)-3-methylcyclohex-2-en-1-yl]-2-(2,4-dihydroxyphenyl)-5,7-dihydroxy-3-(3-methylbut-2-enyl)chromen-4-one	CC1=CC(C(C(C1)C2=C(C=C(C=C2)O)O)C(=O)C3=C(C(=C(C=C3)O)CC=C(C)C)O)C4=C(C=C(C5=C4OC(=C(C5=O)CC=C(C)C)C6=C(C=C(C=C6)O)O)O)O	
L-4-Hydroxy-3-methoxy-a-methylphenylalanine			
Labriformidin	(1*S*,2*S*,3*S*,5*R*,6*R*,9*R*,10*R*,12*S*,14*R*,16*R*,18*S*,20*R*,23*R*,25*R*)-3,9,23-trihydroxy-1,5,20-trimethyl-6-(5-oxo-2*H*-furan-3-yl)-11,17,19,24-tetraoxaheptacyclo[12.12.0.02,10.05,9.010,12.016,25.018,23]hexacosane-4,22-dione	CC1CC(=O)C2(C(O1)OC3CC4CC5C6(O5)C(C4(CC3O2)C)C(C(=O)C7(C6(CCC7C8=CC(=O)OC8)O)C)O)O	
Lacto-N-tetraose	N-[(2S,3R,4R,5S,6R)-2-[(2R,3S,4S,5R,6S)-3,5-dihydroxy-2-(hydroxymethyl)-6-[(2R,3S,4R,5R)-4,5,6-trihydroxy-2-(hydroxymethyl)oxan-3-yl]oxyoxan-4-yl]oxy-5-hydroxy-6-(hydroxymethyl)-4-[(2R,3R,4S,5R,6R)-3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxyoxan-3-yl]acetamide	CC(=O)NC1C(C(C(OC1OC2C(C(OC(C2O)OC3C(OC(C(C3O)O)O)CO)CO)O)CO)O)OC4C(C(C(C(O4)CO)O)O)O	
L-alpha-Acetyl-N,N-dinormethadol	[(3*S*,6*S*)-6-amino-4,4-diphenylheptan-3-yl] acetate	CCC(C(CC(C)N)(C1=CC=CC=C1)C2=CC=CC=C2)OC(=O)C	
Lappaol C	4-[[3-[1,3-dihydroxy-1-(4-hydroxy-3-methoxyphenyl)propan-2-yl]-4-hydroxy-5-methoxyphenyl]methyl]-3-[(4-hydroxy-3-methoxyphenyl)methyl]oxolan-2-one	COC1=CC(=CC(=C1O)C(CO)C(C2=CC(=C(C=C2)O)OC)O)CC3COC(=O)C3CC4=CC(=C(C=C4)O)OC	

Levofuraltadone	(5*S*)-5-(morpholin-4-ylmethyl)-3-[(*E*)-(5-nitrofuran-2-yl)methylideneamino]-1,3-oxazolidin-2-one	C1COCCN1CC2CN(C(=O)O2)N=CC3=CC=C(O3)[N+](=O)[O-]	
Licorice glycoside C1	[(4*S*,5*S*)-5-[(2*S*,4*S*,5*S*)-4,5-dihydroxy-6-(hydroxymethyl)-2-[4-[(2*R*)-7-hydroxy-4-oxo-2,3-dihydrochromen-2-yl]phenoxy]oxan-3-yl]oxy-3,4-dihydroxyoxolan-3-yl]methyl (*E*)-3-(4-hydroxy-3-methoxyphenyl)prop-2-enoate	COC1=C(C=CC(=C1)C=CC(=O)OCC2(COC(C2O)OC3C(C(C(OC3OC4=CC=C(C=C4)C5CC(=O)C6=C(O5)C=C(C=C6)O)CO)O)O)O)O	
Licorice glycoside C2	[(3S,4R,5S)-5-[(2S,3R,4S,5S,6R)-4,5-dihydroxy-6-(hydroxymethyl)-2-[4-[(2S)-7-hydroxy-4-oxo-2,3-dihydrochromen-2-yl]phenoxy]oxan-3-yl]oxy-3,4-dihydroxyoxolan-3-yl]methyl (E)-3-(4-hydroxy-3-methoxyphenyl)prop-2-enoate	COC1=C(C=CC(=C1)C=CC(=O)OCC2(COC(C2O)OC3C(C(C(OC3OC4=CC=C(C=C4)C5CC(=O)C6=C(O5)C=C(C=C6)O)CO)O)O)O)O	
Lisuride	3-[(6*aR*,9*S*)-7-methyl-6,6*a*,8,9-tetrahydro-4*H*-indolo[4,3-fg]quinolin-9-yl]-1,1-diethylurea	CCN(CC)C(=O)NC1CN(C2CC3=CNC4=CC=CC(=C34)C2=C1)C	
Luteolin	2-(3,4-dihydroxyphenyl)-5,7-dihydroxychromen-4-one	C1=CC(=C(C=C1C2=CC(=O)C3=C(C=C(C=C3O2)O)O)O)O	5734-TETRAHYDROXYFLAVONE
Lys Lys Arg	(2*S*)-2-[[(2*S*)-6-amino-2-[[(2*S*)-2,6-diaminohexanoyl]amino]hexanoyl]amino]-5-(diaminomethylideneamino)pentanoic acid	C(CCN)CC(C(=O)NC(CCCCN)C(=O)NC(CCCN=C(N)N)C(=O)O)N	
Mahanimbinol	1-[(2*E*)-3,7-dimethylocta-2,6-dienyl]-3-methyl-9*H*-carbazol-2-ol	CC1=CC2=C(C(=C1O)CC=C(C)CCC=C(C)C)NC3=CC=CC=C32	
Maritimetin	(2*Z*)-2-[(3,4-dihydroxyphenyl)methylidene]-6,7-dihydroxy-1-benzofuran-3-one	C1=CC(=C(C=C1C=C2C(=O)C3=C(O2)C(=C(C=C3)O)O)O)O	
Mauritine A	(2*S*)-*N*-[(2*S*)-1-[(3*S*,7*S*,10*S*,13*Z*)-10-benzyl-8,11-dioxo-2-oxa-6,9,12-triazatricyclo[13.2.2.03,7]nonadeca-1(18),13,15(19),16-tetraen-6-yl]-3-methyl-1-oxobutan-2-yl]-2-(dimethylamino)propanamide	CC(C)C(C(=O)N1CCC2C1C(=O)NC(C(=O)NC=CC3=CC=C(O2)C=C3)CC4=CC=CC=C4)NC(=O)C(C)N(C)C	
MDAI	6,7-dihydro-5*H*-cyclopenta[f][1,3]benzodioxol-6-amine	C1C(CC2=CC3=C(C=C21)OCO3)N	
Melleolide E	[2*a*,7-dihydroxy-3-(hydroxymethyl)-6,6,7*b*-trimethyl-1,2,4*a*,5,7,7*a*-hexahydrocyclobuta[e]inden-2-yl] 2,4-dihydroxy-6-methylbenzoate	CC1=CC(=CC(=C1C(=O)OC2CC3(C2(C(=CC4C3C(C(C4)(C)C)O)CO)O)C)O)O	
Met Phe Trp	(2*S*)-2-[[(2*S*)-2-[[(2*S*)-2-amino-4-methylsulfanylbutanoyl]amino]-3-phenylpropanoyl]amino]-3-(1*H*-indol-3-yl)propanoic acid	CSCCC(C(=O)NC(CC1=CC=CC=C1)C(=O)NC(CC2=CNC3=CC=CC=C32)C(=O)O)N	
Methanofuran	(3*R*,4*S*)-7-[[(1*S*)-4-[[(1*S*)-4-[2-[4-[[5-(aminomethyl)furan-3-yl]methoxy]phenyl]ethylamino]-1-carboxy-4-oxobutyl]amino]-1-carboxy-4-oxobutyl]amino]-7-oxoheptane-1,3,4-tricarboxylic acid	C1=CC(=CC=C1CCNC(=O)CCC(C(=O)O)NC(=O)CCC(C(=O)O)NC(=O)CCC(C(CCC(=O)O)C(=O)O)C(=O)O)OCC2=COC(=C2)CN	
Methyl 3,4,5-trimethoxycinnamate	methyl (*E*)-3-(3,4,5-trimethoxyphenyl)prop-2-enoate	COC1=CC(=CC(=C1OC)OC)C=CC(=O)OC	
methyl 4-[2-(2-formyl-vinyl)-3-hydroxy-5-oxo-cyclopentyl]-butanoate	methyl 4-[3-hydroxy-5-oxo-2-[(*E*)-3-oxoprop-1-enyl]cyclopentyl]butanoate	COC(=O)CCCC1C(C(CC1=O)O)C=CC=O	
Misoprostol (free acid)-d5			
Momordol	5-[8,10-dihydroxy-11-(hydroxymethyl)-4,7-dimethyltridecyl]-6-ethyl-4-hydroxy-4,5-dimethylcyclohex-2-en-1-one	CCC1C(=O)C=CC(C1(C)CCCC(C)CCC(C)C(CC(C(CC)CO)O)O)O)(C)O	
Myricatomentoside II	3,8-dihydroxy-16,17-dimethoxy-15-[(2*S*,3*R*,4*S*,5*S*,6*R*)-3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxytricyclo[12.3.1.12,6]nonadeca-1(17),2,4,6(19),14(18),15-hexaen-9-one	COC1=C2C=C(CCCCC(=O)C(CC3=CC2=C(C=C3)O)O)C(=C1OC)OC4C(C(C(C(O4)CO)O)O)O	
N-(2R-methyl-3-hydroxy-ethyl)-16,16-dimethyl-5Z,8Z,11Z,14Z-docosatetraenoyl amine	(5*Z*,8*Z*,11*Z*,14*Z*)-*N*-[(2*R*)-1-hydroxypropan-2-yl]-16,16-dimethyldocosa-5,8,11,14-tetraenamide	CCCCCCC(C)(C)C=CCC=CCC=CCC=CCCCC(=O)NC(C)CO	
N1,N5,N10-Triferuloyl spermidine	(*E*)-3-(4-hydroxy-3-methoxyphenyl)-*N*-[4-[[(*E*)-3-(4-hydroxy-3-methoxyphenyl)prop-2-enoyl]-[3-[[(*E*)-3-(4-hydroxy-3-methoxyphenyl)prop-2-enoyl]amino]propyl]amino]butyl]prop-2-enamide	COC1=C(C=CC(=C1)C=CC(=O)NCCCCN(CCCNC(=O)C=CC2=CC(=C(C=C2)O)OC)C(=O)C=CC3=CC(=C(C=C3)O)OC)O	CPD-12214

Nalbuphine	(4*R*,4*aS*,7*S*,7*aR*,12*bS*)-3-(cyclobutylmethyl)-1,2,4,5,6,7,7*a*,13-octahydro-4,12-methanobenzofuro[3,2-e]isoquinoline-4*a*,7,9-triol	C1CC(C1)CN2CCC34C5C(CCC3(C2CC6=C4C(=C(C=C6)O)O5)O)O	
Nap-Trp-OH			
Nb-Lignoceroyltryptamine	N-[2-(1*H*-indol-3-yl)ethyl]tetracosanamide	CCCCCCCCCCCCCCCCCCCCCCCC(=O)NCCC1=CNC2=CC=CC=C21	
N-demethyl-6-O-methylerythromycin	(3*R*,4*S*,5*S*,6*R*,7*R*,9*R*,11*R*,12*R*,13*R*,14*S*)-14-ethyl-12,13-dihydroxy-4-(5-hydroxy-4-methoxy-4,6-dimethyloxan-2-yl)oxy-6-[3-hydroxy-6-methyl-4-(methylamino)oxan-2-yl]oxy-7-methoxy-3,5,7,9,11,13-hexamethyl-oxacyclotetradecane-2,10-dione	CCC1C(C(C(C(=O)C(CC(C(C(C(C(C(=O)O1)C)OC2CC(C(C(O2)C)O)(C)OC)C)OC3C(C(CC(O3)C)NC)O)(C)OC)C)C)O)(C)O	
N-dodecanoyl-L-Homoserine lactone-3-hydrazone-fluorescein			
Neoacrimarine H	6-hydroxy-11-[(9-hydroxy-8,8-dimethyl-2-oxo-9,10-dihydropyrano[2,3-f]chromen-10-yl)oxy]-3,3,12-trimethylpyrano[2,3-c]acridin-7-one	CC1(C=CC2=C(O1)C=C(C3=C2N(C4=C(C3=O)C=CC=C4OC5C(C(OC6=C5C7=C(C=C6)C=CC(=O)O7)(C)C)O)C)O)C	
Neoarctin B	3-[(4-hydroxy-3,5-dimethoxyphenyl)methyl]-4-[[4-[4-[[4-[(4-hydroxy-3,5-dimethoxyphenyl)methyl]-5-oxooxolan-3-yl]methyl]-2-methoxyphenyl]-3-methoxyphenyl]methyl]oxolan-2-one	COC1=CC(=CC(=C1O)OC)CC2C(COC2=O)CC3=CC(=C(C=C3)C4=C(C=C(C=C4)CC5COC(=O)C5CC6=CC(=C(C(=C6)OC)O)OC)OC)OC	
Nevirapine	2-cyclopropyl-7-methyl-2,4,9,15-tetrazatricyclo[9.4.0.03,8]pentadeca-1(11),3,5,7,12,14-hexaen-10-one	CC1=C2C(=NC=C1)N(C3=C(C=CC=N3)C(=O)N2)C4CC4	
Nicofetamide	N-(1,2-diphenylethyl)pyridine-3-carboxamide	C1=CC=C(C=C1)CC(C2=CC=CC=C2)NC(=O)C3=CN=CC=C3	
Nogalamycin	methyl (1R,10S,12S,13R,21R,22S,23R,24R)-23-(dimethylamino)-4,8,12,22,24-pentahydroxy-1,12-dimethyl-6,17-dioxo-10-[(2R,3R,4R,5S,6S)-3,4,5-trimethoxy-4,6-dimethyloxan-2-yl]oxy-20,25-dioxahexacyclo[19.3.1.02,19.05,18.07,16.09,14]pentacosa-2,4,7(16),8,14,18-hexaene-13-carboxylate	CC1C(C(C(C(O1)OC2CC(C(C3=CC4=C(C(=C23)O)C(=O)C5=C(C=C6C(=C5C4=O)OC7C(C(C(C6(O7)C)O)N(C)C)O)O)C(=O)OC)(C)O)OC)(C)OC)OC	
Nomilinic acid	(3*R*)-3-acetyloxy-3-[(1*R*,2*R*,5*R*,6*R*,7*R*,10*S*,11*R*,14*S*)-11-(furan-3-yl)-5-(2-hydroxypropan-2-yl)-2,6,10-trimethyl-3,13-dioxo-12,15-dioxatetracyclo[8.5.0.01,14.02,7]pentadecan-6-yl]propanoic acid	CC(=O)OC(CC(=O)O)C1(C2CCC3(C(OC(=O)C4C3(C2(C(=O)CC1C(C)(C)O)C)O4)C5=COC=C5)C)C	
Nostoxanthin sulfate	sodium;[(1*R*,6*R*)-4-[(1*E*,3*E*,5*E*,7*E*,9*E*,11*E*,13*E*,15*E*,17*E*)-18-[(4*R*,5*R*)-4,5-dihydroxy-2,6,6-trimethylcyclohexen-1-yl]-3,7,12,16-tetramethyloctadeca-1,3,5,7,9,11,13,15,17-nonaenyl]-6-hydroxy-3,5,5-trimethylcyclohex-3-en-1-yl] sulfate	CC1=C(C(C(C(C1)O)O)(C)C)C=CC(=CC=CC(=CC=CC=C(C)C=CC=C(C)C=CC2=C(CC(C(C2(C)C)O)OS(=O)(=O)[O-])C)C)C.[Na+]	
N-Valerylglycine methyl ester	methyl 2-(pentanoylamino)acetate	CCCCC(=O)NCC(=O)OC	
O-b-D-Gal-(1->3)-O-[O-b-D-Gal-(1->4)-2-(acetylamino)-2-deoxy-b-D-Glc-(1->6)]-2-(acetylamino)-2-deoxy- D-Galactose	N-[(2*R*,3*R*,4*R*,6*R*)-2-[(2*R*,3*S*,4*R*)-5-acetamido-2,3-dihydroxy-6-oxo-4-[(2*R*,3*R*,4*S*,5*R*,6*R*)-3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxyhexoxy]-4-hydroxy-6-(hydroxymethyl)-5-[(2*S*,3*R*,6*R*)-3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxyoxan-3-yl]acetamide	CC(=O)NC1C(C(C(OC1OCC(C(C(C(=O)NC(=O)C)OC2C(C(C(C(O2)CO)O)O)O)O)O)CO)OC3C(C(C(C(O3)CO)O)O)O)O	
O-Demethylfonsecin	2,5,6,8-tetrahydroxy-2-methyl-3*H*-benzo[g]chromen-4-one	CC1(CC(=O)C2=C(C3=C(C=C(C=C3C=C2O1)O)O)O)O	
Okanin	(*E*)-3-(3,4-dihydroxyphenyl)-1-(2,3,4-trihydroxyphenyl)prop-2-en-1-one	C1=CC(=C(C=C1C=CC(=O)C2=C(C(=C(C=C2)O)O)O)O)O	
Oritin-4beta-ol	(2*R*,3*S*,4*S*)-2-(4-hydroxyphenyl)-3,4-dihydro-2*H*-chromene-3,4,7,8-tetrol	C1=CC(=CC=C1C2C(C(C3=C(O2)C(=C(C=C3)O)O)O)O)O	
Palmitoyl 3-carbacyclic Phosphatidic Acid	(2-hydroxy-2-oxo-1,2λ5-oxaphospholan-5-yl)methyl hexadecanoate	CCCCCCCCCCCCCCCC(=O)OCC1CCP(=O)(O1)O	
p-Cresol	4-methylphenol	CC1=CC=C(C=C1)O	CPD-108
Pectenotoxin 7			
PGF1	7-[(1*R*,2*R*,3*R*,5*S*)-3,5-dihydroxy-2-[(*E*,3*S*)-3-hydroxyoct-1-enyl]cyclopentyl]heptanoic acid	CCCCCC(C=CC1C(CC(C1CCCCCC(=O)O)O)O)O	

Phe Arg Leu	(2*S*)-2-[[(2*S*)-2-[[(2*S*)-2-amino-3-phenylpropanoyl]amino]-5-(diaminomethylideneamino)pentanoyl]amino]-4-methylpentanoic acid	CC(C)CC(C(=O)O)NC(=O)C(CCCN=C(N)N)NC(=O)C(CC1=CC=CC=C1)N	
Phellatin	3,5-dihydroxy-6-(3-hydroxy-3-methylbutyl)-2-(4-hydroxyphenyl)-7-[(2*S*,4*S*,5*S*)-3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxychromen-4-one	CC(C)(CCC1=C(C=C2C(=C1O)C(=O)C(=C(O2)C3=CC=C(C=C3)O)O)OC4C(C(C(C(O4)CO)O)O)O)O	
PHOME	[cyano-(6-methoxynaphthalen-2-yl)methyl] 2-(3-phenyloxiran-2-yl)acetate	COC1=CC2=C(C=C1)C=C(C=C2)C(C#N)OC(=O)CC3C(O3)C4=CC=CC=C4	
PHOP	1-([1,3]oxazolo[4,5-b]pyridin-2-yl)-6-phenylhexan-1-one	C1=CC=C(C=C1)CCCCCC(=O)C2=NC3=C(O2)C=CC=N3	
Physalin L	(1*S*,2*R*,3*R*,5*R*,6*S*,7*R*,14*R*,15*S*,18*S*,21*S*,22*R*,25*S*)-5,7,18-trihydroxy-1,14,21,25-tetramethyl-4,20,23-trioxaheptacyclo[20.3.1.12,5.03,18.03,21.06,15.09,14]heptacosa-8,10-diene-13,19,24,27-tetrone	CC1C(=O)OC2CC1(C3C(=O)C4(C5C(CCC6(C3(C2(OC6=O)C)O4)O)C7(C(=O)CC=CC7=CC5O)C)O)C	
Physalin O	5,7,18-trihydroxy-1,14,21,25-tetramethyl-4,20,23-trioxaheptacyclo[20.3.1.12,5.03,18.03,21.06,15.09,14]heptacosa-8,11-diene-13,19,24,27-tetrone	CC1C(=O)OC2CC1(C3C(=O)C4(C5C(CCC6(C3(C2(OC6=O)C)O4)O)C7(C(=CC5O)CC=CC7=O)C)O)C	
Picein	1-[4-[(2*S*,3*R*,4*S*,5*S*,6*R*)-3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxyphenyl]ethanone	CC(=O)C1=CC=C(C=C1)OC2C(C(C(C(O2)CO)O)O)O	
Pradimicin A	(2*R*)-2-[[(5*S*,6*S*)-1,6,9,14-tetrahydroxy-5-[(2*S*,3*R*,4*S*,5*S*,6*R*)-3-hydroxy-6-methyl-5-(methylamino)-4-[(2*S*,3*R*,4*S*,5*R*)-3,4,5-trihydroxyoxan-2-yl]oxyoxan-2-yl]oxy-11-methoxy-3-methyl-8,13-dioxo-5,6-dihydrobenzo[a]tetracene-2-carbonyl]amino]propanoic acid	CC1C(C(C(C(O1)OC2C(C3=CC4=C(C(=C3C5=C2C=C(C(=C5O)C(=O)NC(C)C(=O)O)C)O)C(=O)C6=C(C4=O)C(=CC(=C6)OC)O)O)O)OC7C(C(C(CO7)O)O)O)NC	
Proanthocyanidin A2	(1*R*,5*R*,6*R*,13*S*,21*R*)-5,13-bis(3,4-dihydroxyphenyl)-4,12,14-trioxapentacyclo[11.7.1.02,11.03,8.015,20]henicosa-2(11),3(8),9,15,17,19-hexaene-6,9,17,19,21-pentol	C1C(C(OC2=C1C(=CC3=C2C4C(C(O3)(OC5=CC(=CC(=C45)O)O)C6=CC(=C(C=C6)O)O)O)O)C7=CC(=C(C=C7)O)O)O	
Propanoic acid, 2-hydroxy-3-[(4-hydroxy-1-naphthalenyl)oxy]-	2-hydroxy-3-(4-hydroxynaphthalen-1-yl)oxypropanoic acid	C1=CC=C2C(=C1)C(=CC=C2OCC(C(=O)O)O)O	
p-Salicylic acid	4-hydroxybenzoic acid	C1=CC(=CC=C1C(=O)O)O	
PtdIns (4)-P1-fluorescein			
PtdIns-(4)-P1 (1,2-dioctanoyl) biotin			
Pteroyltriglutamic acid	2-[[4-[[4-[[4-[(2-amino-4-oxo-3*H*-pteridin-6-yl)methylamino]benzoyl]amino]-4-carboxybutanoyl]amino]-4-carboxybutanoyl]amino]pentanedioic acid	C1=CC(=CC=C1C(=O)NC(CCC(=O)NC(CCC(=O)NC(CCC(=O)O)C(=O)O)C(=O)O)C(=O)O)NCC2=CN=C3C(=N2)C(=O)NC(=N3)N	
Pyridoxal (Vitamin B6)	(4-formyl-5-hydroxy-6-methylpyridin-3-yl)methyl dihydrogen phosphate	CC1=NC=C(C(=C1O)C=O)COP(=O)(O)O	
Quercetin	2-(3,4-dihydroxyphenyl)-3,5,7-trihydroxychromen-4-one	C1=CC(=C(C=C1C2=C(C(=O)C3=C(C=C(C=C3O2)O)O)O)O)O	CPD-520
Red chlorophyll catabolite	3-[(2*Z*,3*S*,4*S*)-5-[(*Z*)-(4-ethenyl-3-methyl-5-oxopyrrol-2-ylidene)methyl]-2-[(5*R*)-2-[(3-ethyl-5-formyl-4-methyl-1*H*-pyrrol-2-yl)methyl]-5-methoxycarbonyl-3-methyl-4-oxo-1*H*-cyclopenta[b]pyrrol-6-ylidene]-4-methyl-3,4-dihydropyrrol-3-yl]propanoic acid	CCC1=C(NC(=C1C)C=O)CC2=C(C3=C(N2)C(=C4C(C(C(=N4)C=C5C(=C(C(=O)N5)C=C)C)C)CCC(=O)O)C(C3=O)C(=O)OC)C	CPD-7063
Rhapontin	(2*S*,3*R*,4*S*,5*S*,6*R*)-2-[3-hydroxy-5-[(*E*)-2-(3-hydroxy-4-methoxyphenyl)ethenyl]phenoxy]-6-(hydroxymethyl)oxane-3,4,5-triol	COC1=C(C=C(C=C1)C=CC2=CC(=CC(=C2)OC3C(C(C(C(O3)CO)O)O)O)O)O	
Rhodomycin D	methyl (1*R*,2*R*,4*S*)-4-[(2*R*,4*S*,5*S*,6*S*)-4-amino-5-hydroxy-6-methyloxan-2-yl]oxy-2-ethyl-2,5,7,12-tetrahydroxy-6,11-dioxo-3,4-dihydro-1*H*-tetracene-1-carboxylate	CCC1(CC(C2=C(C1C(=O)OC)C(=C3C(=C2O)C(=O)C4=C(C3=O)C=CC=C4O)O)OC5CC(C(C(O5)C)O)N)O	CPD-15730
Rhubafuran	2,4-dimethyl-4-phenyloxolane	CC1CC(CO1)(C)C2=CC=CC=C2	
Rifamycin Z	(1*S*,2*E*,16*Z*,18*E*,20*S*,21*S*,22*R*,23*R*,24*R*,25*S*,28*R*,29*R*)-6,8,21,23,28-pentahydroxy-3,7,16,20,22,24,29-heptamethyl-26-oxa-14-azatetracyclo[23.2.2.19,13.05,10]triaconta-2,5(10),6,8,12,16,18-heptaene-4,11,15,27,30-pentone	CC1C=CC=C(C(=O)NC2=CC(=O)C3=C(C(=C(C(=C3C2=O)O)O)C)O)C(=O)C(=CC4C(C(C(C(C(C1O)C)O)C)OC4=O)C)O)C)C	
Ripazepam	1-ethyl-3-methyl-8-phenyl-4,6-dihydropyrazolo[4,3-e][1,4]diazepin-5-one	CCN1C2=C(C(=N1)C)NC(=O)CN=C2C3=CC=CC=C3	

Sandoricin	methyl 2-[(1*S*,3*S*,5*S*,7*S*,8*S*,9*R*,11*S*,12*S*,13*R*,16*S*)-5,11-diacetyloxy-13-(furan-3-yl)-16-hydroxy-6,6,8,12-tetramethyl-17-methylidene-15-oxo-2,14-dioxatetracyclo[7.7.1.01,12.03,8]heptadecan-7-yl]acetate	CC(=O)OC1CC2C(=C)C3(C1(C(OC(=O)C3O)C4=COC=C4)C)OC5C2(C(C(C(C5)OC(=O)C)(C)C)CC(=O)OC)C	
Sarcostin	(3*S*,8*S*,9*R*,10*R*,12*R*,13*R*,14*R*,17*S*)-17-[(1*S*)-1-hydroxyethyl]-10,13-dimethyl-1,2,3,4,7,9,11,12,15,16-decahydrocyclopenta[a]phenanthrene-3,8,12,14,17-pentol	CC(C1(CCC2(C1(C(CC3C2(CC=C4C3(CCC(C4)O)C)O)O)C)O)O)O	
Sarsasapogenin	(1*R*,2*S*,4*S*,5'*S*,6*R*,7*S*,8*R*,9*S*,12*S*,13*S*,16*S*,18*R*)-5',7,9,13-tetramethylspiro[5-oxapentacyclo[10.8.0.02,9.04,8.013,18]icosane-6,2'-oxane]-16-ol	CC1CCC2(C(C3C(O2)CC4C3(CCC5C4CCC6C5(CCC(C6)O)C)C)C)OC1	
SB 939	(*E*)-3-[2-butyl-1-[2-(diethylamino)ethyl]benzimidazol-5-yl]-*N*-hydroxyprop-2-enamide	CCCCC1=NC2=C(N1CCN(CC)CC)C=CC(=C2)C=CC(=O)NO	
Scabraside	[(2*R*,3*R*,4*S*,5*R*,6*S*)-4-acetyloxy-2-(acetyloxymethyl)-6-[[(3*S*,4*R*,4a*S*)-4-ethenyl-8-oxo-4,4*a*,5,6-tetrahydro-3*H*-pyrano[3,4-c]pyran-3-yl]oxy]-5-benzoyloxyoxan-3-yl] 2-hydroxy-3-[(2*S*,3*R*,4*S*,5*S*,6*R*)-3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxybenzoate	CC(=O)OCC1C(C(C(C(O1)OC2C(C3CCOC(=O)C3=CO2)C=C)OC(=O)C4=CC=CC=C4)OC(=O)C)OC(=O)C5=C(C(=CC=C5)OC6C(C(C(C(O6)CO)O)O)O)O	
Scutellarein	5,6,7-trihydroxy-2-(4-hydroxyphenyl)chromen-4-one	C1=CC(=CC=C1C2=CC(=O)C3=C(O2)C=C(C(=C3O)O)O)O	
Sennoside C	(9*R*)-4-hydroxy-9-[(9*R*)-4-hydroxy-2-(hydroxymethyl)-10-oxo-5-[(2*S*,3*R*,4*S*,5*S*,6*R*)-3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxy-9*H*-anthracen-9-yl]-10-oxo-5-[(2*S*,3*R*,4*S*,5*S*,6*R*)-3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxy-9*H*-anthracene-2-carboxylic acid	C1=CC2=C(C(=C1)OC3C(C(C(C(O3)CO)O)O)O)C(=O)C4=C(C2C5C6=C(C(=CC=C6)OC7C(C(C(C(O7)CO)O)O)O)O)C(=O)C8=C5C=C(C=C8O)C(=O)O)C=C(C=C4O)CO	
Ser His Gly			
Ser Ile Lys	(2*S*)-6-amino-2-[[(2*S*,3*S*)-2-[[(2*S*)-2-amino-3-hydroxypropanoyl]amino]-3-methylpentanoyl]amino]hexanoic acid	CCC(C)C(C(=O)NC(CCCCN)C(=O)O)NC(=O)C(CO)N	
soladulcidine	(1*R*,2*S*,4*S*,5'*R*,6*R*,7*S*,8*R*,9*S*,12*S*,13*S*,16*S*,18*S*)-5',7,9,13-tetramethylspiro[5-oxapentacyclo[10.8.0.02,9.04,8.013,18]icosane-6,2'-piperidine]-16-ol	CC1CCC2(C(C3C(O2)CC4C3(CCC5C4CCC6C5(CCC(C6)O)C)C)C)NC1	CPD-22270
Solanidine	(1*S*,2*S*,7*S*,10*R*,11*S*,14*S*,15*R*,16*S*,17*R*,20*S*,23*S*)-10,14,16,20-tetramethyl-22-azahexacyclo[12.10.0.02,11.05,10.015,23.017,22]tetracos-4-en-7-ol	CC1CCC2C(C3C(N2C1)CC4C3(CCC5C4CC=C6C5(CCC(C6)O)C)C)C	CPD-9217
Sorbaldehyde	(2*E*,4*E*)-hexa-2,4-dienal	CC=CC=CC=O	
Sphingofungin B	5-amino-6-(7-amino-6-methoxy-5,8-dioxoquinolin-2-yl)-4-(2-hydroxy-3,4-dimethoxyphenyl)-3-methylpyridine-2-carboxylic acid	CC1=C(C(=C(N=C1C(=O)O)C2=NC3=C(C=C2)C(=O)C(=C(C3=O)N)OC)N)C4=C(C(=C(C=C4)OC)OC)O	
Streptonigrin	5-amino-6-(7-amino-6-methoxy-5,8-dioxoquinolin-2-yl)-4-(2-hydroxy-3,4-dimethoxyphenyl)-3-methylpyridine-2-carboxylic acid	CC1=C(C(=C(N=C1C(=O)O)C2=NC3=C(C=C2)C(=O)C(=C(C3=O)N)OC)N)C4=C(C(=C(C=C4)OC)OC)O	
Taccalonolide A	[(1*S*,2*S*,3*R*,5*S*,7*S*,9*S*,10*R*,11*R*,12*S*,13*S*,14*R*,15*R*,16*S*,17*S*,22*S*,23*S*,24*R*,25*R*)-10,14,25-triacetyloxy-3,22-dihydroxy-11,15,17,22,23-pentamethyl-4,21-dioxo-8,20-dioxaheptacyclo[13.10.0.02,12.05,11.07,9.016,24.019,23]pentacos-18-en-13-yl] acetate	CC1C=C2C(C3C1C4(C(C3OC(=O)C)C5C(C(C4OC(=O)C)OC(=O)C)C6(C(CC7C(C6OC(=O)C)O7)C(=O)C5O)C)C)(C(C(=O)O2)(C)O)C	
Tigecycline	(4*S*,4a*S*,5a*R*,12a*R*)-9-[[2-(*tert*-butylamino)acetyl]amino]-4,7-bis(dimethylamino)-1,10,11,12*a*-tetrahydroxy-3,12-dioxo-4*a*,5,5*a*,6-tetrahydro-4*H*-tetracene-2-carboxamide	CC(C)(C)NCC(=O)NC1=CC(=C2CC3CC4C(C(=O)C(=C(C4(C(=O)C3=C(C2=C1O)O)O)O)C(=O)N)N(C)C)N(C)C	
Torososide B	1-[(2S,3R,4S,5S,6R)-6-[[(2R,3R,4S,5R,6R)-3,5-dihydroxy-6-(hydroxymethyl)-4-[(2S,3R,4S,5S,6R)-3,4,5-trihydroxy-6-[[(2R,3R,4S,5S,6R)-3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxymethyl]oxan-2-yl]oxyoxan-2-yl]oxymethyl]-3,4,5-trihydroxyoxan-2-yl]oxy-8-hydroxy-3-methoxy-6-methylanthracene-9,10-dione	CC1=CC2=C(C(=C1)O)C(=O)C3=C(C2=O)C=C(C=C3OC4C(C(C(C(O4)COC5C(C(C(C(O5)CO)O)O)OC6C(C(C(C(O6)COC7C(C(C(C(O7)CO)O)O)O)O)O)O)O)O)O)O)OC	
Trabectedin	[(1R,2R,3R,11S,12S,14R,26R)-5,6',12-trihydroxy-6,7'-dimethoxy-7,21,30-trimethyl-27-oxospiro[17,19,28-trioxa-24-thia-13,30-diazaheptacyclo[12.9.6.13,11.02,13.04,9.015,23.016,20]triaconta-4(9),5,7,15,20,22-hexaene-26,1'-3,4-dihydro-2H-isoquinoline]-22-yl] acetate	CC1=CC2=C(C3C4C5C6=C(C(=C7C(=C6C(N4C(C(C2)N3C)O)COC(=O)C8(CS5)C9=CC(=C(C=C9CCN8)O)OC)OCO7)C)OC(=O)C)C(=C1OC)O	

trans-4,5-Dihydroxy-4,5-dihydropyrene	(4*S*,5*S*)-4,5-dihydropyrene-4,5-diol	C1=CC2=C3C(=C1)C(C(C4=CC=CC(=C43)C=C2)O)O	
trans-trismethoxy Resveratrol-d4	1,2,4,5-tetradeuterio-3-[(*E*)-2-(3,5-dimethoxyphenyl)ethenyl]-6-methoxybenzene	COC1=CC=C(C=C1)C=CC2=CC(=CC(=C2)OC)OC	
Tricyclodehydroisohumulone	(1*S*,5*S*,8*R*,10*R*)-2,5-dihydroxy-9,9-dimethyl-3-(3-methylbutanoyl)-10-prop-1-en-2-yltricyclo[6.3.0.01,5]undec-2-ene-4,6-dione	CC(C)CC(=O)C1=C(C23CC(C(C2CC(=O)C3(C1=O)O)(C)C)C(=C)C)O	
Trifolirhizin	(2*R*,3*S*,4*S*,5*R*,6*S*)-2-(hydroxymethyl)-6-[[(1*R*,12*R*)-5,7,11,19-tetraoxapentacyclo[10.8.0.02,10.04,8.013,18]icosa-2,4(8),9,13(18),14,16-hexaen-16-yl]oxy]oxane-3,4,5-triol	C1C2C(C3=C(O1)C=C(C=C3)OC4C(C(C(C(O4)CO)O)O)O)OC5=CC6=C(C=C25)OCO6	CPD-4222
Triphyllin A	(2*S*,4*S*,5*S*)-2-[[(2*R*,4*S*)-4-hydroxy-6-(hydroxymethyl)-2-(4-methoxyphenyl)-8-methyl-5-[(2*S*,4*S*,5*S*)-3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxy-3,4-dihydro-2*H*-chromen-7-yl]oxy]-6-(hydroxymethyl)oxane-3,4,5-triol	CC1=C2C(=C(C(=C1OC3C(C(C(C(O3)CO)O)O)O)CO)OC4C(C(C(C(O4)CO)O)O)O)C(CC(O2)C5=CC=C(C=C5)OC)O	
tritriacontane-16,18-dione	tritriacontane-16,18-dione	CCCCCCCCCCCCCCCC(=O)CC(=O)CCCCCCCCCCCCCCC	
Trp Tyr Tyr	(2*S*)-2-[[(2*S*)-2-[[(2*S*)-2-amino-3-(1*H*-indol-3-yl)propanoyl]amino]-3-(4-hydroxyphenyl)propanoyl]amino]-3-(4-hydroxyphenyl)propanoic acid	C1=CC=C2C(=C1)C(=CN2)CC(C(=O)NC(CC3=CC=C(C=C3)O)C(=O)NC(CC4=CC=C(C=C4)O)C(=O)O)N	
Trypanothione disulfide	(2*S*)-2-amino-5-[[(4*R*,23*R*)-4-[[(4*S*)-4-amino-4-carboxybutanoyl]amino]-5,8,19,22-tetraoxo-1,2-dithia-6,9,13,18,21-pentazacyclotetracos-23-yl]amino]-5-oxopentanoic acid	C1CCNC(=O)CNC(=O)C(CSSCC(C(=O)NCC(=O)NCCCNC1)NC(=O)CCC(C(=O)O)N)NC(=O)CCC(C(=O)O)N	
Tumonoic Acid A	(2*S*)-1-[(*E*,2*R*,3*S*)-3-hydroxy-2,4-dimethyldodec-4-enoyl]pyrrolidine-2-carboxylic acid	CCCCCCCC=C(C)C(C(C)C(=O)N1CCCC1C(=O)O)O	
Tyr Tyr Trp	(2*S*)-2-[[(2*S*)-2-[[(2*S*)-2-amino-3-(4-hydroxyphenyl)propanoyl]amino]-3-(4-hydroxyphenyl)propanoyl]amino]-3-(1*H*-indol-3-yl)propanoic acid	C1=CC=C2C(=C1)C(=CN2)CC(C(=O)O)NC(=O)C(CC3=CC=C(C=C3)O)NC(=O)C(CC4=CC=C(C=C4)O)N	
U-18666A	(3*S*,8*R*,9*S*,10*R*,13*S*,14*S*)-3-[2-(diethylamino)ethoxy]-10,13-dimethyl-1,2,3,4,7,8,9,11,12,14,15,16-dodecahydrocyclopenta[a]phenanthren-17-one;hydrochloride	CCN(CC)CCOC1CCC2(C3CCC4(C(C3CC=C2C1)CCC4=O)C)C.Cl	
UDP-3-(3R-hydroxy-tetradecanoyl)-D-glucosamine			
UDP-N-acetylmuramoyl-L-alanyl-D-glutamyl-L-lysyl-D-alanine	(2R)-2-[[(2S)-2-[[(2R)-2-[(2R,3R,4R,5S,6R)-3-acetamido-2-[[[(2R,3S,4R,5R)-5-(2,4-dioxopyrimidin-1-yl)-3,4-dihydroxyoxolan-2-yl]methoxy-hydroxyphosphoryl]oxy-hydroxyphosphoryl]oxy-5-hydroxy-6-(hydroxymethyl)oxan-4-yl]oxypropanoyl]amino]propanoyl]amino]-5-[[(2S)-6-amino-1-[[(1R)-1-carboxyethyl]amino]-1-oxohexan-2-yl]amino]-5-oxopentanoic acid	CC(C(=O)NC(CCC(=O)NC(CCCCN)C(=O)NC(C)C(=O)O)C(=O)O)NC(=O)C(C)OC1C(C(OC(C1O)CO)OP(=O)(O)OP(=O)(O)OCC2C(C(C(O2)N3C=CC(=O)NC3=O)O)O)NC(=O)C	
Val Lys Tyr	(2*S*)-2-[[(2*S*)-6-amino-2-[[(2*S*)-2-amino-3-methylbutanoyl]amino]hexanoyl]amino]-3-(4-hydroxyphenyl)propanoic acid	CC(C)C(C(=O)NC(CCCCN)C(=O)NC(CC1=CC=C(C=C1)O)C(=O)O)N	
Valrubicin	[2-oxo-2-[(2*S*,4*S*)-2,5,12-trihydroxy-4-[(2*R*,4*S*,5*S*,6*S*)-5-hydroxy-6-methyl-4-[(2,2,2-trifluoroacetyl)amino]oxan-2-yl]oxy-7-methoxy-6,11-dioxo-3,4-dihydro-1*H*-tetracen-2-yl]ethyl] pentanoate	CCCCC(=O)OCC(=O)C1(CC(C2=C(C1)C(=C3C(=C2O)C(=O)C4=C(C3=O)C=CC=C4OC)O)OC5CC(C(C(O5)C)O)NC(=O)C(F)(F)F)O	
Vanillin 1,2-butylene glycol acetal	4-(4,5-dimethyl-1,3-dioxolan-2-yl)-2-methoxyphenol	CC1C(OC(O1)C2=CC(=C(C=C2)O)OC)C	
Vanilpyruvic acid	3-(4-hydroxy-3-methoxyphenyl)-2-oxopropanoic acid	COC1=C(C=CC(=C1)CC(=O)C(=O)O)O	CPD-11830
Vardenafil	2-[2-ethoxy-5-(4-ethylpiperazin-1-yl)sulfonylphenyl]-5-methyl-7-propyl-3*H*-imidazo[5,1-f][1,2,4]triazin-4-one	CCCC1=NC(=C2N1N=C(NC2=O)C3=C(C=CC(=C3)S(=O)(=O)N4CCN(CC4)CC)OCC)C	
Veratridine	[(1*R*,2*S*,6*S*,9*S*,10*R*,11*S*,12*S*,14*R*,15*S*,18*S*,19*S*,22*S*,23*S*,25*R*)-1,10,11,12,14,23-hexahydroxy-6,10,19-trimethyl-24-oxa-4-azaheptacyclo[12.12.0.02,11.04,9.015,25.018,23.019,25]hexacosan-22-yl] 3,4-dimethoxybenzoate	CC1CCC2C(C3(C(CC4(C5CCC6C7(C5(CC4(C3CN2C1)O)OC6(CCC7)OC(=O)C8=CC(=C(C=C8)OC)OC)O)C)O)O)O)(C)O	
Vinorelbine	methyl (1R,9R,10S,11R,12R,19R)-11-acetyloxy-12-ethyl-4-[(12S,14R)-16-ethyl-12-methoxycarbonyl-1,10-diazatetracyclo[12.3.1.03,11.04,9]octadeca-3(11),4,6,8,15-pentaen-12-yl]-10-hydroxy-5-methoxy-8-methyl-8,16-diazapentacyclo[10.6.1.01,9.02,7.016,19]nonadeca-2,4,6,13-tetraene-10-carboxylate	CCC1=CC2CC(C3=C(CN(C2)C1)C4=CC=CC=C4N3)(C5=C(C=C6C(=C5)C78CCN9C7C(C=CC9)(C(C(C8N6C)(C(=O)OC)O)OC(=O)C)CC)OC)C(=O)OC	

Vitamin D3 butyrate	[(1S,3Z)-3-[(2E)-2-[(1R,3aS,7aR)-7a-methyl-1-[(2R)-6-methylheptan-2-yl]-2,3,3a,5,6,7-hexahydro-1H-inden-4-ylidene]ethylidene]-4-methylidenecyclohexyl] butanoate	CCCC(=O)OC1CCC(=C)C(=CC=C2CCCC3(C2CCC3C(C)CCCC(C)C)C)C1	
Volicitin	(2*S*)-5-amino-2-[[(9*Z*,12*Z*,15*Z*,17*S*)-17-hydroxyoctadeca-9,12,15-trienoyl]amino]-5-oxopentanoic acid	CC(C=CCC=CCC=CCCCCCCCC(=O)NC(CCC(=O)N)C(=O)O)O	
zeleplon			
Z-Gly-Pro-Leu-Gly-Pro	(2*S*)-1-[2-[[(2*S*)-4-methyl-2-[[(2*S*)-1-[2-(phenylmethoxycarbonylamino)acetyl]pyrrolidine-2-carbonyl]amino]pentanoyl]amino]acetyl]pyrrolidine-2-carboxylic acid	CC(C)CC(C(=O)NCC(=O)N1CCCC1C(=O)O)NC(=O)C2CCCN2C(=O)CNC(=O)OCC3=CC=CC=C3	

Supplemental Table 3: HPLC-MS Results of Ethyl Acetate Fraction of PNW2016-02 Secretome

Name	Formula	Score	Mass	Mass (DB)	Height	METLIN
(+)-Maackiain 3-O-glucoside	C22 H22 O10	98.33	446.122	446.1213	856274	48120
Isoscoparin 2"-O-ferulate	C32 H30 O14	99.24	638.1644	638.1636	819053	49191
Kaempferol 3-(4"-p-coumaroylglucoside)	C30 H26 O13	98.99	594.1381	594.1373	789244	50356
7-Hydroxy-2',4',5'-trimethoxyisoflavone 7-O-glucoside	C24 H26 O11	98.88	490.1483	490.1475	728012	47584
Kaempferol 3-gentiobioside-7-glucuronide	C33 H38 O22	92.48	786.1806	786.1855	674741	50179
Isoscoparin 2"-O-ferulate	C32 H30 O14	99.44	638.1643	638.1636	313942	49191
Antibiotic TA	C35 H61 N O8	93.43	623.4446	623.4397	487994	71087
Okanin 4-methyl ether 3'-glucoside	C22 H24 O11	99.72	464.1325	464.1319	264082	51971
PE-Cer(d16:1(4E)/18:0)	C36 H73 N2 O6 P	90.24	660.5185	660.5206	479702	103059
7-Hydroxy-2',4',5'-trimethoxyisoflavone 7-O-glucoside	C24 H26 O11	99.57	490.1482	490.1475	255114	47584
5-Demethylmelibentin	C20 H18 O9	99.49	402.0958	402.0951	398087	51761
(+)-12-Isocopalene-15,16-dial	C20 H30 O2	99.82	302.225	302.2246	301970	53624
Kaempferol 3-(4"-p-coumaroylglucoside)	C30 H26 O13	99.65	594.138	594.1373	164821	50356
Sennoside C	C42 H40 O19	99.58	848.2168	848.2164	110532	71349
7-Hydroxy-2',4',5'-trimethoxyisoflavone 7-O-glucoside	C24 H26 O11	99.24	490.148	490.1475	104147	47584
C-12 NBD Ceramide	C36 H61 N5 O6	97.55	659.4656	659.4622	127320	45025
Biochanin A-_-D-glucoside	C22 H22 O10	99.62	446.122	446.1213	213098	64300
11(S)-HETE	C20 H32 O3	99.66	320.2356	320.2351	405508	45056
Artonin Q	C31 H30 O8	63.28	530.196	530.1941	112617	95554
Hydroxyitraconazole	C35 H38 Cl2 N8 O5	59.1	720.2426	720.2342	59544	878
Esmeraldin B	C38 H28 N4 O7	90.71	652.1913	652.1958	129072	69427
Gossypetin 3-sophoroside-8-glucoside	C33 H40 O23	93.57	804.1905	804.196	81622	51594
Austalide C	C30 H38 O11	99.25	574.2419	574.2414	115886	89697
2-(1,3-Benzodioxol-5-yl)-7-hydroxy-3,5,6,8-tetramethoxy-4H-1-benzopyran-4-one	C20 H18 O9	99.5	402.0958	402.0951	131753	51767
2-(1,3-Benzodioxol-5-yl)-7-hydroxy-3,5,6,8-tetramethoxy-4H-1-benzopyran-4-one	C20 H18 O9	83.14	402.0958	402.0951	131753	51767
Quercetin 7-[xylosyl-(1->2)-rhamnosyl-(1->2)-rhamnosyl]-(1->6)-glucoside	C38 H48 O24	56.63	888.2638	888.2536	53227	50670
Kamahine C	C14 H20 O5	99.86	268.1313	268.1311	53837	93546
Helieianeoside C	C37 H46 O24	93.09	874.2318	874.2379	40025	50666
7-Hydroxy-2',4',5'-trimethoxyisoflavone 7-O-glucoside	C24 H26 O11	99.6	490.1475	490.1475	49512	47584
Kaempferol 7-(6"-p-coumarylglucoside)	C30 H26 O13	99.25	594.1377	594.1373	77507	50312
Okanin 4'-(4"-acetyl-6"-p-coumarylglucoside)	C32 H30 O14	99.74	638.1638	638.1636	55485	51963
Fulvine	C16 H23 N O5	95.65	309.1584	309.1576	46491	68233
Kaempferol 7-(6"-p-coumarylglucoside)	C30 H26 O13	98.68	594.1377	594.1373	69133	50312
Triphyllin A	C30 H40 O16	74.04	656.2269	656.2316	42268	47385
Isoscoparin 2"-O-ferulate	C32 H30 O14	99.39	638.1637	638.1636	53589	49191
Okanin 4'-(4"-acetyl-6"-p-coumarylglucoside)	C32 H30 O14	99.24	638.1638	638.1636	55485	51963
Misoprostol (free acid)-d5	C21 H31 D5 O5	64.28	373.2869	373.2877	35675	96463
Apional	C12 H12 O5	86.68	236.0688	236.0685	102022	88988
Physalin L	C28 H32 O10	99.57	528.1996	528.1995	23726	89861
Biochanin A-_-D-glucoside	C22 H22 O10	99.47	446.1217	446.1213	80884	64300
Hydroxyitraconazole	C35 H38 Cl2 N8 O5	58.87	720.2425	720.2342	43049	878
CAY10500	C32 H32 F3 N3 O2	73.05	547.2413	547.2447	38708	44995
Helieianeoside C	C37 H46 O24	96.07	874.2337	874.2379	26550	50666
Physalin L	C28 H32 O10	99.43	528.1994	528.1995	23670	89861
Deoxyvasicinone	C11 H10 N2 O	99.35	186.0787	186.0793	85163	68472
Artemetin 5-glucosyl-(1->4)-rhamnoside	C32 H40 O17	72.45	696.2193	696.2266	33695	51353
Alkaloid RC	C26 H29 N O11	95.34	531.1761	531.1741	43613	86222
10-deacetylbaccatin III	C29 H36 O10	95.33	544.2316	544.2308	47155	84946
Scutellarein 6,7,4'-trimethyl ether 5-glucoside	C24 H26 O11	99.74	490.1476	490.1475	50156	49556
Kaempferol 3-(2"-(Z)-p-coumaroylglucoside)	C30 H26 O13	99.48	594.1375	594.1373	42502	50355
Helieianeoside C	C37 H46 O24	91.11	874.2317	874.2379	19157	50666
Cinnamtannin B2	C60 H48 O24	98.85	1152.253	1152.2536	27338	71827
Gal_(1-3)[Fuc_(1-2)]Gal_(1-4)Glc_-Sp	C26 H45 N3 O20	92.35	719.258	719.2596	38178	3658
8,9-dihydroxy-5,11,14-eicosatrienoic acid	C20 H34 O4	99.92	338.246	338.2457	81426	36274
N-dodecanoyl-L-Homoserine lactone-3-hydrazone-fluorescein	C37 H40 N4 O8 S	85.92	700.2529	700.2567	32649	64901
Phellatin	C26 H30 O12	99.47	534.1737	534.1737	34325	50125

Methyl 3,4,5-trimethoxycinnamate [arabinosyl-(1->3)-[glucosyl-(1->6)]-glucosyl] ester	C29 H42 O19	87.56	694.2259	694.232	21783	90182
Cinnamtannin B2	C60 H48 O24	99.13	1152.2527	1152.2536	27507	71827
5,4'-Dihydroxy-3,3'-dimethoxy-6,7-methylenedioxyflavone	C18 H14 O8	99.87	358.0691	358.0689	50497	51433
2"-O-p-Hydroxybenzoylorientin	C28 H24 O13	99.5	568.1218	568.1217	35146	48947
Dialicor	C21 H27 N O2	99.4	325.2041	325.2042	26300	69504
Sennoside C	C42 H40 O19	99.83	848.2158	848.2164	21595	71349
5,4'-Dihydroxy-3,3'-dimethoxy-6,7-methylenedioxyflavone	C18 H14 O8	99.91	358.0691	358.0689	51440	51433
Citbrasine	C17 H17 N O6	99.69	331.1052	331.1056	23207	92611
Cucurbitacin D	C30 H44 O7	78.08	516.3089	516.3087	26325	41669
Momordol	C26 H48 O5	99.59	440.3502	440.3502	30929	86511
2-(4-Hydroxy-3,5-dimethoxyphenyl)ethanol 4'-glucoside	C16 H24 O9	84.12	360.1423	360.142	27436	93089
Hydroxyitraconazole	C35 H38 Cl2 N8 O5	59.52	720.2421	720.2342	13668	878
Helieianeoside C	C37 H46 O24	88.74	874.2319	874.2379	12400	50666
Kansuinine B	C38 H42 O14	62.98	722.2555	722.2575	18331	67445
Vanilpyruvic acid	C10 H10 O5	86.84	210.0527	210.0528	51728	62448
Labriformidin	C29 H36 O11	99.39	560.2256	560.2258	29170	67235
Hydroxyitraconazole	C35 H38 Cl2 N8 O5	58.49	720.2429	720.2342	18916	878
Chitin	C28 H49 N3 O16	72.21	683.3146	683.3113	15950	6903
2-amino-heptanoic acid	C7 H15 N O2	87.05	145.1109	145.1103	81861	35934
Physalin L	C28 H32 O10	99.63	528.1996	528.1995	14117	89861
Nomilinic acid 17-glucoside	C34 H48 O16	64.77	712.2914	712.2942	17299	92698
Formononetin 7-O-(2"-p-hydroxybenzoylglucoside)	C29 H26 O11	99.22	550.1474	550.1475	19155	47549
(S)-2,3-Dihydro-7-hydroxy-2-methyl-4-oxo-4H-1-benzopyran-5-acetic acid	C12 H12 O5	86.63	236.0685	236.0685	41707	94916
3-oxo-octanoyl-CoA	C37 H64 N7 O18 P3 S	78.64	1019.3295	1019.3241	19637	36648
Okanin 4-methyl ether 3'-glucoside	C22 H24 O11	99.78	464.1318	464.1319	24544	51971
Neoarctin B	C42 H46 O12	99.7	742.2987	742.2989	19156	71860
Acacetin 7-(6"-acetylglucoside)	C24 H24 O11	99.41	488.1319	488.1319	16091	48888
4,2',3',4'-Tetrahydroxychalcone	C15 H12 O5	99.81	272.0687	272.0685	106416	51955
Nostoxanthin sulfate	C40 H55 Na O7 S	88.68	702.3617	702.3566	12869	41363
Scabraside	C40 H44 O20	71.56	844.2357	844.2426	14661	71943
Kaempferol 3-isorhamninoside-7-rhamnoside	C39 H50 O23	90.11	886.2683	886.2743	18022	50102
Formononetin 7-(2-p-hydroxybenzoylglucoside)	C29 H26 O11	99.25	550.1475	550.1475	11101	93419
100-1	C31 H34 O10	65.2	566.216	566.2152	22517	63788
Enkephaline, (D-Ala)2-Leu	C29 H39 N5 O7	89.75	569.2844	569.2849	23451	44752
Met Phe Trp	C25 H30 N4 O4 S	58.81	482.1979	482.1988	18021	21978
Phellatin	C26 H30 O12	99.44	534.174	534.1737	14520	50125
2-(1,3-Benzodioxol-5-yl)-7-hydroxy-3,5,6,8-tetramethoxy-4H-1-benzopyran-4-one	C20 H18 O9	99.86	402.0951	402.0951	22074	51767
Valrubicin	C34 H36 F3 N O13	70.89	723.2208	723.2139	10587	43368
Physalin L	C28 H32 O10	99.14	528.1996	528.1995	9146	89861
7,12-Dimethylbenz[a]anthracene 5,6-oxide	C20 H16 O	72.99	272.12	272.1201	2729	73259
Enkephaline, (D-Ala)2-Leu	C29 H39 N5 O7	95.28	569.285	569.2849	16041	44752
Taccalonolide A	C36 H46 O14	72.33	702.2907	702.2888	13477	67087
3-Oxo-delta1-steroid	C19 H28 O	86.34	272.214	272.214	22289	65869
5alpha-Ethoxy-6beta-hydroxy-5,6-dihydrophysalin B	C30 H36 O11	99.48	572.2256	572.2258	8148	86386
5,10-Methylene-H4SPT	C36 H52 N7 O19 P	89.5	917.31	917.3056	12160	72566
Methanofuran	C34 H44 N4 O15	83.29	748.272	748.2803	13459	63295
Isoscoparin 2"-O-ferulate	C32 H30 O14	98.99	638.1635	638.1636	8523	49191
3-Hydroxy-OPC8-CoA	C39 H64 N7 O19 P3 S	88.93	1059.3161	1059.3191	22859	63357
(Z)-Resveratrol 3,4'-diglucoside	C26 H32 O13	99.57	552.184	552.1843	10143	94453
6-Keto-PGF1_	C20 H34 O6	81.95	370.2347	370.2355	10612	2558
Eicosanoyl-EA	C22 H45 N O2	99.31	355.3453	355.345	90679	3723
Arbutin	C12 H16 O7	86.08	272.0892	272.0896	14379	43989
Physalin O	C28 H32 O10	95.05	528.1984	528.1995	14207	93687
Cassinine	C44 H51 N O17	96.09	865.3174	865.3157	10552	68023
Kaempferol 3-isorhamninoside-7-rhamnoside	C39 H50 O23	88.43	886.2679	886.2743	14457	50102
Apional	C12 H12 O5	86.75	236.0686	236.0685	23429	88988
5,7,3'-Trihydroxy-3,6,8,4',5'-pentamethoxyflavone	C20 H20 O10	99.86	420.1057	420.1056	14645	51775
2-(1,3-Benzodioxol-5-yl)-7-hydroxy-3,5,6,8-tetramethoxy-4H-1-benzopyran-4-one	C20 H18 O9	99.67	402.0951	402.0951	15990	51767
(Z)-Resveratrol 3,4'-diglucoside	C26 H32 O13	99.14	552.1848	552.1843	9713	94453
4-O-Demethyl-13-dihydroadriamycinone	C20 H18 O9	99.86	402.0952	402.0951	19044	664

Dukunolide E	C26 H28 O9	99.69	484.1735	484.1733	10676	90988
His Gly Ser	C11 H17 N5 O5	56.48	299.123	299.123	17836	23116
Trabectedin	C39 H43 N3 O11 S	82.95	761.2534	761.2618	9641	71400
3-Oxo-5alpha-steroid	C19 H30 O	86.64	274.2296	274.2297	17088	65868
Enkephaline, (D-Ala)2-Leu	C29 H39 N5 O7	92.84	569.2843	569.2849	7944	44752
Esmeraldin B	C38 H28 N4 O7	93.05	652.1923	652.1958	9873	69427
Falaconitine	C34 H47 N O10	99.27	629.3196	629.32	11888	67117
Lisuride	C20 H26 N4 O	78.96	338.2088	338.2107	7158	1010
Kamahine C	C14 H20 O5	85.89	268.1311	268.1311	10252	93546
Diphenylcarbazide	C13 H14 N4 O	85.99	242.1155	242.1168	5665	68930
1'-Acetoxyeugenol acetate	C14 H16 O5	86.41	264.0997	264.0998	13446	68319
Physalin L	C28 H32 O10	77.48	528.1997	528.1995	7210	89861
Trifolirhizin	C22 H22 O10	99.84	446.1214	446.1213	18840	64288
Sphingofungin B	C20 H39 N O6	84.05	389.2777	389.2777	8482	53962
Maritimetin 6-(6"-p-coumarylglucoside)	C30 H26 O13	99.41	594.1371	594.1373	11505	52426
Isohopeaphenol	C56 H42 O12	77.73	906.2745	906.2676	6225	88905
5,7,11,14-Eicosatetraenoic acid, 9-hydroxy-, (E,Z,Z,Z)-; 9-HETE	C20 H32 O3	84.13	320.2351	320.2351	8937	75007
Beauvericin	C45 H57 N3 O9	80.18	783.4031	783.4095	7785	69114
PtdIns-(4)-P1 (1,2-dioctanoyl) biotin	C35 H63 N3 O18 P2 S	82.88	907.3284	907.3303	8513	64897
Licorice glycoside C2	C36 H38 O16	98.25	726.2166	726.216	5529	52521
Trabectedin	C39 H43 N3 O11 S	83.84	761.2539	761.2618	5314	71400
SB 939	C20 H30 N4 O2	47.62	358.2368	358.2369	684	96423
GalNAc_1-3[Fuc_1-2]Gal_1-3GlcNAc_-Sp	C30 H51 N5 O20	57.47	801.3169	801.3127	7596	3660
PtdIns (4)-P1-fluorescein	C46 H59 N O22 P2	75.54	1039.3117	1039.3004	6054	64896
3-oxo-octanoyl-CoA	C37 H64 N7 O18 P3 S	77.16	1019.3291	1019.3241	6248	36648
Rhodomycin D	C28 H31 N O11	94.91	557.191	557.1897	11034	63790
5alpha-Ethoxy-6beta-hydroxy-5,6-dihydrophysalin B	C30 H36 O11	75.35	572.2256	572.2258	3192	86386
Vanilpyruvic acid	C10 H10 O5	87.33	210.0529	210.0528	16176	62448
Hydroxyanastrozole	C17 H19 N5 O	83.56	309.1585	309.159	6953	679
Licorice glycoside C2	C36 H38 O16	98.54	726.2155	726.216	6400	52521
Ripazepam	C15 H16 N4 O	84.09	268.1309	268.1324	2846	73201
3-oxo-octanoyl-CoA	C37 H64 N7 O18 P3 S	75.26	1019.3287	1019.3241	7523	36648
Chartreusin	C32 H32 O14	73.65	640.1794	640.1792	3910	67494
2,2,4,4-Tetramethyl-6-(1-oxobutyl)-1,3,5-cyclohexanetrione	C14 H20 O4	85.71	252.1364	252.1362	5659	89070
Kaempferol 3-isorhamninoside-7-rhamnoside	C39 H50 O23	87.27	886.2666	886.2743	6700	50102
(13Z,16Z)-Docosadi-13,16-enoyl-CoA	C43 H74 N7 O17 P3 S	83.05	1085.4121	1085.4075	7891	63369
Kaempferol 3-rhamnoside-7-[6'''-ferulyglucosyl-(1->3)-rhamnoside]	C43 H48 O22	87.55	916.2551	916.2637	6370	50320
3',4',5'-Trimethoxycinnamyl alcohol acetate	C14 H18 O5	86.71	266.1154	266.1154	20559	95344
GalNAc_1-3[Fuc_1-2]Gal_1-3GlcNAc_-Sp	C30 H51 N5 O20	57.77	801.3168	801.3127	5351	3660
Gazer	C43 H67 N5 O28	78.93	1101.4067	1101.3973	6955	86213
Hoiamide B	C45 H73 N5 O10 S3	71.1	939.4616	939.452	5265	65382
Melleolide E	C23 H30 O7	66.22	418.1966	418.1992	11979	93428
Lappaol C	C30 H34 O10	74.42	554.2156	554.2152	2627	71863
Kaempferol 3-(2"-(Z)-p-coumaroylglucoside)	C30 H26 O13	75.26	594.1374	594.1373	7678	50355
beta-Bixin	C25 H30 O4	78.4	394.2125	394.2144	11735	90424
Okanin 4'-(4"-acetyl-6"-p-coumarylglucoside)	C32 H30 O14	99.27	638.1632	638.1636	8808	51963
2-Methyl-4-pentenal	C6 H10 O	87.64	98.0732	98.0732	9976	86895
Quercetin 3-xylosyl-(1->3)-rhanosyl-(1->6)-[apiosyl-(1->2)-galactoside]	C37 H46 O24	87.57	874.2312	874.2379	6325	50667
Sandoricin	C31 H40 O11	94.35	588.2573	588.2571	7412	92412
Ikarisoside D	C28 H30 O11	77.34	542.1786	542.1788	3957	50140
5-Methyltetrahydropteroylpentaglutamate	C40 H53 N11 O18	81.42	975.3493	975.357	4922	6511
4-Hydroxyphenylglyoxylate	C8 H6 O4	87.36	166.0269	166.0266	17356	63853
Ripazepam	C15 H16 N4 O	83.06	268.131	268.1324	3817	73201
PA(15:0/22:6(4Z,7Z,10Z,13Z,16Z,19Z))	C40 H67 O8 P	98.24	706.4592	706.4574	5918	81324
Beauvericin	C45 H57 N3 O9	84.87	783.4041	783.4095	7235	69114
Isorhamnetin 3-rhamnosyl-(1->6)-[rhamnosyl-(1->2)-(3'''-(E)-p-coumaroylgalactoside]-7-rhamnoside	C49 H58 O26	84.05	1062.3143	1062.3216	4314	50806
Kaempferol 3-(4"-p-coumaroylglucoside)	C30 H26 O13	63.29	594.1372	594.1373	5917	50356
4-Cumylphenol	C15 H16 O	85.1	212.1189	212.1201	8444	69918
Rhodomycin D	C28 H31 N O11	96.43	557.1904	557.1897	6353	63790
Kaempferol 3-gentiobioside-7,4'-diglucoside	C39 H50 O26	86.83	934.2653	934.259	4838	50198
Asp Glu Glu	C14 H21 N3 O10	66.83	391.121	391.1227	5711	18340
Torososide B	C40 H52 O25	57.91	932.2903	932.2798	4157	90247

1,2-Dioleoyl-sn-Glycero-3-Phosphate	C39 H73 O8 P	91.6	700.5088	700.5043	5711	34486
Methanofuran	C34 H44 N4 O15	85.41	748.2723	748.2803	4519	63295
UDP-N-acetylmuramoyl-L-alanyl-D-glutamyl-L-lysyl-D-alanine	C37 H60 N8 O25 P2	83.21	1078.3099	1078.3145	5506	66400
Trypanothione disulfide	C27 H47 N9 O10 S2	73.23	721.295	721.2887	4448	63641
4-Hydroxyphenylglyoxylate	C8 H6 O4	87.52	166.0265	166.0266	13643	63853
7-Hydroxy-2',5,6-trimethoxy-4',5'-methylenedioxyisoflavone 7-(2-p-coumaroylglucoside)	C34 H32 O15	93.08	680.1686	680.1741	4865	95579
4,2',4',6'-Tetrahydroxy-3-methoxychalcone	C16 H14 O6	99.65	302.0791	302.079	23786	52077
5alpha-Ethoxy-6beta-hydroxy-5,6-dihydrophysalin B	C30 H36 O11	47.62	572.2256	572.2258	3168	86386
2,2-Dibutyl-3-(4-methoxyphenyl)-4-methyl-2H-1-benzopyran-7-ol acetate	C27 H34 O4	80.01	422.2451	422.2457	9844	70555
Tricyclodehydroisohumulone	C21 H28 O5	47.54	360.1933	360.1937	2670	68527
Mahanimbinol	C23 H27 N O	46.89	333.2082	333.2093	940	95431
Kuwanone H	C45 H44 O11	69.39	760.2893	760.2884	5332	68131
Hexandraside C	C43 H56 O24	85.55	956.3092	956.3162	4759	50137
17-Epiestriol	C18 H24 O3	73.29	288.1683	288.1725	6887	5345
9,10-12,13-Diepoxyoctadecanoate	C18 H32 O4	85.1	312.2303	312.2301	5295	70341
DHOAA(AcO)-Val-Phe(NMe)-Pro-Phe(NMe)-Gly-OMe	C43 H61 N5 O9	98.12	791.4459	791.4469	2910	65478
Debrisoquine	C10 H13 N3	79.14	175.1096	175.1109	8180	58453
UDP-3-(3R-hydroxy-tetradecanoyl)-_D-glucosamine	C29 H51 N3 O18 P2	78.29	791.2683	791.2643	4701	41548
2E,4Z-Heptadienal	C7 H10 O	87.92	110.0731	110.0732	8969	75386
Nomilinic acid 17-glucoside	C34 H48 O16	77.05	712.2878	712.2942	4691	92698
Okanin 4'-(4"-acetyl-6"-p-coumarylglucoside)	C32 H30 O14	95.94	638.1631	638.1636	4313	51963
Beauvericin	C45 H57 N3 O9	86.21	783.4056	783.4095	6593	69114
Nalbuphine-3-glucuronide	C27 H35 N O10	78.33	533.226	533.2261	7532	1402
Nicofetamide	C20 H18 N2 O	85.15	302.1424	302.1419	3871	71885
5alpha-Ethoxy-6beta-hydroxy-5,6-dihydrophysalin B	C30 H36 O11	74.93	572.2261	572.2258	3358	86386
Lacto-N-tetraose	C26 H45 N O21	68.44	707.2562	707.2484	3764	58460
2,6-Diamino-4-hydroxy-5-N-methylformamidopyrimidine	C6 H9 N5 O2	46.47	183.0747	183.0756	3911	62402
Italidipyrone	C29 H34 O10	80.48	542.2182	542.2152	5622	95721
Indaconitine	C34 H47 N O10	68.75	629.319	629.32	4704	67125
1,4-Dideoxy-1,4-imino-D-ribitol	C5 H11 N O3	46.22	133.0731	133.0739	5232	94068
5alpha-Ethoxy-6beta-hydroxy-5,6-dihydrophysalin B	C30 H36 O11	71.49	572.2261	572.2258	2553	86386
Gazer	C43 H67 N5 O28	73.41	1101.4062	1101.3973	3674	86213
Licorice glycoside C2	C36 H38 O16	98.44	726.2153	726.216	6996	52521
Epicatechin 3-O-gallate-(4beta->6)-epigallocatechin 3-O-gallate	C44 H34 O21	84.39	898.1625	898.1593	4612	93887
farnesyl triphosphate	C15 H29 O10 P3	76.94	462.095	462.0974	8715	53847
2,2,4,4-Tetramethyl-6-(1-oxobutyl)-1,3,5-cyclohexanetrione	C14 H20 O4	84.51	252.1363	252.1362	4574	89070
1'-Acetoxyeugenol acetate	C14 H16 O5	86.23	264.1	264.0998	4155	68319
Physalin L	C28 H32 O10	78.29	528.2	528.1995	4628	89861
Isatidine	C18 H25 N O7	82.34	367.1634	367.1631	6119	68253
1-(5-Acetyl-2-hydroxyphenyl)-3-methyl-1-butanone	C13 H16 O3	82.13	220.1104	220.1099	8157	88547
N-Valerylglycine methyl ester	C8 H15 N O3	81.61	173.1038	173.1052	13862	5650
Kni 102	C31 H41 N5 O7	75.76	595.3016	595.3006	4122	70999
MGDG(18:5(3Z,6Z,9Z,12Z,15Z)/18:5(3Z,6Z,9Z,12Z,15Z))	C45 H66 O10	70.44	766.4681	766.4656	3602	46648
Picein	C14 H18 O7	85.85	298.1053	298.1053	8247	68525
AFMK	C13 H16 N2 O4	79.9	264.1106	264.111	5391	44788
Ergosine	C30 H37 N5 O5	76.19	547.278	547.2795	4788	67481
zeleplon	C17 H15 N5 O	76.56	305.1285	305.1277	3780	85038
4-Hydroxyphenylglyoxylate	C8 H6 O4	87.35	166.0269	166.0266	13397	63853
Diaziquone	C18 H22 N2 O6	99.91	362.1477	362.1478	147818	44225
Pradimicin A	C40 H44 N2 O18	79.12	840.2642	840.2589	3118	66544
Duartin (-)	C18 H20 O6	84.87	332.126	332.126	3721	43647
Isoswertisin 3"-O-(2'"-methylbutyrate)	C27 H30 O11	78.67	530.1787	530.1788	1979	49429
BL II	C28 H22 O12	76.9	550.111	550.1111	6530	94124
Citrusin B	C27 H36 O13	53.84	568.2095	568.2156	3344	93822
(20S)-24-Hydroxy-19-norgeminivitamin D3	C31 H54 O5	76.49	506.3947	506.3971	3557	42558
Hydrojuglone glucoside	C16 H18 O8	80.64	338.0978	338.1002	102400	95613
zeleplon	C17 H15 N5 O	77.78	305.128	305.1277	6618	85038
Trabectedin	C39 H43 N3 O11 S	50.64	761.2525	761.2618	2804	71400
Kaempferol 3-(4"-p-coumaroylglucoside)	C30 H26 O13	63.57	594.1371	594.1373	5475	50356

Amoxapine	C17 H16 Cl N3 O	47.48	313.0977	313.0982	30279	982
3-O-alpha-L-Arabinopyranosylcinnamtannin B1	C50 H44 O22	92.39	996.2312	996.2324	3621	94409
PS(22:4(7Z,10Z,13Z,16Z)/18:4(6Z,9Z,12Z,15Z))	C46 H74 N O10 P	57.17	831.5079	831.505	3501	78467
5alpha-Ethoxy-6beta-hydroxy-5,6-dihydrophysalin B	C30 H36 O11	73.33	572.2253	572.2258	2452	86386
Isodonal	C22 H28 O7	83.35	404.1838	404.1835	3534	67441
(-)-trans-(S)-allethrin	C19 H26 O3	80.09	302.1877	302.1882	7046	53293
Pteroyltriglutamic acid	C29 H33 N9 O12	63.62	699.2174	699.2249	2663	6381
3',4',5'-Trimethoxycinnamyl alcohol acetate	C14 H18 O5	86.47	266.1155	266.1154	6500	95344
7-Hydroxyfluphenazine glucuronide	C28 H34 F3 N3 O8 S	51.31	629.2108	629.2019	2761	2779
Heptacarboxylporphyrin III	C39 H38 N4 O14	83.11	786.2492	786.2385	3523	6400
Famciclovir	C14 H19 N5 O4	78.81	321.1427	321.1437	6965	2650
Vardenafil	C23 H32 N6 O4 S	55.64	488.2277	488.2206	4114	85438
5,4'-Dihydroxy-3,3'-dimethoxy-6,7-methylenedioxyflavone	C18 H14 O8	84.56	358.0687	358.0689	7154	51433
Dehydroisocoproporphyrin	C36 H36 N4 O8	70.63	652.2514	652.2533	3663	6570
Hydroxydestruxin B	C30 H51 N5 O8	87.01	609.3726	609.3738	4125	94675
CAY10598	C21 H29 N5 O2	76.35	383.23	383.2321	3023	45515
Bisosthenon B	C28 H24 O8	72.75	488.147	488.1471	2958	87952
N1,N5,N10-Triferuloyl spermidine	C37 H43 N3 O9	58.03	673.2952	673.2999	2518	64187
PE-Cer(d14:2(4E,6E)/21:0)	C37 H73 N2 O6 P	68.33	672.5197	672.5206	3210	103079
Eicosapentaenoic Acid-d5	C20 H25 D5 O2	95.99	307.2546	307.256	12788	96379
Kaempferol 3-rutinoside-7-sophoroside	C39 H50 O25	59.11	918.2586	918.2641	3245	50197
Pectenotoxin 7	C47 H68 O16	59.68	888.4495	888.4507	2912	88641
Z-Gly-Pro-Leu-Gly-Pro	C28 H39 N5 O8	76.13	573.2807	573.2799	3437	65925
Acarbose (Glucobay)	C25 H43 N O18	68.86	645.2445	645.248	3751	756
2E,6E-Octadienal	C8 H12 O	82.44	124.0888	124.0888	4759	46424
Vinorelbine	C45 H54 N4 O8	64.33	778.39	778.3942	3499	43370
Sarsasapogenin 3-[4"-glucosyl-6"-arabinosylglucoside]	C44 H72 O17	60	872.4705	872.477	232	86820
Gossypetin 3-sophoroside-8-glucoside	C33 H40 O23	61.94	804.1898	804.196	3285	51594
1(15),8(19)-Trinervitadiene-3alpha,5alpha,18-triol	C20 H32 O3	81.41	320.2351	320.2351	5626	53697
trans-trismethoxy Resveratrol-d4	C17 H14 D4 O3	64.75	274.1506	274.1507	3480	96499
Levofuraltadone	C13 H16 N4 O6	81.72	324.1079	324.107	3729	72985
N1,N5,N10-Triferuloyl spermidine	C37 H43 N3 O9	58.95	673.2959	673.2999	2262	64187
Gnididin	C37 H44 O10	63.9	648.296	648.2934	2583	67424
Gln Phe Trp	C25 H29 N5 O5	79.96	479.2178	479.2169	4088	20886
24-Nor-5_-cholane-3_,7_,12_-triol	C23 H40 O3	82.2	364.2966	364.2977	3689	4124
Kaempferol 3-gentiobioside-7-glucuronide	C33 H38 O22	58.83	786.1792	786.1855	2867	50179
Kaempferol 3-rutinoside-7-sophoroside	C39 H50 O25	89.32	918.2572	918.2641	3683	50197
Rhapontin	C21 H24 O9	80.36	420.1423	420.142	3958	44320
N1,N5,N10-Triferuloyl spermidine	C37 H43 N3 O9	59.27	673.2962	673.2999	3401	64187
7-Hydroxylpradimicin A	C40 H44 N2 O19	60.56	856.2566	856.2538	2355	66542
Ibudilast	C14 H18 N2 O	86.05	230.1416	230.1419	23741	85552
Aralionine A	C34 H38 N4 O5	59.27	582.2869	582.2842	2889	68078
PtdIns-(4)-P1 (1,2-dioctanoyl) biotin	C35 H63 N3 O18 P2 S	85.48	907.3278	907.3303	2452	64897
Physalin L	C28 H32 O10	78.65	528.1996	528.1995	5604	89861
Deltorphin C	C37 H52 N8 O10	62.39	768.3866	768.3806	2325	72025
Hydroxyitraconazole	C35 H38 Cl2 N8 O5	54.3	720.2381	720.2342	1921	878
N-demethyl-6-O-methylerythromycin	C37 H67 N O13	71.19	733.4552	733.4612	2047	1861
2-Phenylethanol glucuronide	C14 H18 O7	47.62	298.1053	298.1053	3121	61666
(E,E)-3,7,11-Trimethyl-2,6,10-dodecatrienyl octanoate	C23 H40 O2	71.46	348.3031	348.3028	6543	97253
Rifamycin Z	C35 H41 N O11	66.88	651.267	651.268	2315	63722
16beta-Hydroxy-3,11-dioxopregna-4,17(20)-dien-21-oic acid, gamma-lactone	C21 H24 O4	97.65	340.1687	340.1675	117561	70224
Kaempferol 3-isorhamninoside-7-rhamnoside	C39 H50 O23	56.05	886.2673	886.2743	2963	50102
PGF1_	C20 H36 O5	47.58	356.2565	356.2563	89879	3503
p-Cresol	C7 H8 O	84.43	108.0576	108.0575	7263	4236
Derricin	C21 H22 O3	74.64	322.1539	322.1569	3775	51815
FIPI	C23 H24 F N5 O2	73.28	421.1941	421.1914	2386	45560
1-hexadecanyl-2-((2'-alpha-glucosyl)-beta-glucosyl)-3-beta-xylosyl-sn-glycerol	C36 H68 O17	60.31	772.4444	772.4457	2077	46631
2'"-N-Acetyl-6'"-deamino-6'"-hydroxyneomycin C	C25 H47 N5 O15	53.97	657.3013	657.3069	2102	71796
Luteolin 7-neohesperidoside-4'-sophoroside	C39 H50 O25	58.7	918.2584	918.2641	2076	49083
Nap-Trp-OH	C29 H23 N3 O6	70.24	509.1534	509.1587	2553	65037
(2E,5E)-Undecadienal	C11 H18 O	41.47	166.1377	166.1358	3561	75346
5,2',5'-Trihydroxy-3,7,8-trimethoxyflavone 2'-acetate	C20 H18 O9	83.33	402.0949	402.0951	2946	51499
6-hydroxydoxazosin	C23 H25 N5 O6	73.23	467.1845	467.1805	4023	2457

UDP-N-acetylmuramoyl-L-alanyl-D-glutamyl-L-lysyl-D-alanine	C37 H60 N8 O25 P2	83.39	1078.3098	1078.3145	3028	66400
_-Linolenoyl-CoA	C39 H64 N7 O17 P3 S	73.88	1027.3368	1027.3292	3679	63380
Z-Gly-Pro-Leu-Gly-Pro	C28 H39 N5 O8	77.82	573.2805	573.2799	4463	65925
Bikhaconitine	C36 H51 N O11	70.21	673.3459	673.3462	2762	67099
L-4-Hydroxy-3-methoxy-a-methylphenylalanine	C11 H15 N O4	47.61	225.1002	225.1001	4206	1219
Licorice glycoside C1	C36 H38 O16	59.85	726.215	726.216	2139	52520
Neoacrimarine H	C33 H29 N O8	39.75	567.195	567.1893	1923	88923
Chrysophanol 8-(6-galloylglucoside)	C28 H24 O13	76.83	568.121	568.1217	3874	92063
14alpha-Hydroxyixocarpanolide	C28 H40 O7	74.21	488.2776	488.2774	3267	89855
5_-Cholane-3_,7_,12_,24-tetrol	C24 H42 O4	78.71	394.308	394.3083	3882	42885
Tigecycline	C29 H39 N5 O8	76.65	585.2802	585.2799	2450	69251
Sorbaldehyde	C6 H8 O	47.62	96.0575	96.0575	3844	36528
3'-N-Debenzoyltaxol	C40 H47 N O13	67.29	749.3039	749.3047	2148	64037
Red chlorophyll catabolite	C35 H38 N4 O7	71.77	626.2723	626.274	2314	63971
Curacin D	C22 H33 N O S	47.48	359.2288	359.2283	1620	65384
PE(P-16:0/15:1(9Z))	C36 H70 N O7 P	68.77	659.4839	659.489	3724	77590
Kaempferol 3-neohesperidoside-7-(6"-malonylglucoside)	C37 H44 O24	58.55	872.2151	872.2223	2250	50380
Hexandraside C	C43 H56 O24	56.89	956.3087	956.3162	2242	50137
Kaempferol 3-rutinoside-7-sophoroside	C39 H50 O25	56.99	918.2572	918.2641	2488	50197
Glucoconvallatoxoloside	C41 H64 O20	57.78	876.3925	876.3991	2443	86661
8-Hydroxy-perphenazine glucuronide	C27 H34 Cl N3 O8 S	49.48	595.1688	595.1755	1903	1806
Bambuterol	C18 H29 N3 O5	47.53	367.2103	367.2107	2137	85510
5-Methyltetrahydropteroylpentaglutamate	C40 H53 N11 O18	55.95	975.3501	975.357	1699	6511
trans-4,5-Dihydroxy-4,5-dihydropyrene	C16 H12 O2	81.79	236.085	236.0837	4453	72099
Myricatomentoside II	C27 H34 O11	74.93	534.2096	534.2101	3411	87786
PS(20:5(5Z,8Z,11Z,14Z,17Z)/18:3(6Z,9Z,12Z))	C44 H70 N O10 P	56	803.4773	803.4737	1905	78325
Corchorusoside E	C41 H64 O19	57.31	860.3975	860.4042	2321	88749
Absintholide	C30 H38 O8	62.55	526.2557	526.2567	2381	92908
Cinncassiol D4	C20 H32 O5	66.15	352.2216	352.225	2336	71844
BL II	C28 H22 O12	77.62	550.1106	550.1111	3341	94124
Apional	C12 H12 O5	85.88	236.0687	236.0685	4073	88988
Val Lys Tyr	C20 H32 N4 O5	61.82	408.2361	408.2373	2134	19434
Isogermafurene	C15 H20 O	82.58	216.1513	216.1514	3516	92888
Nogalamycin	C39 H49 N O16	66.2	787.3043	787.3051	2230	72425
Rhubafuran	C12 H16 O	76.55	176.1209	176.1201	5291	91199
5,10-Methylene-H4SPT	C36 H52 N7 O19 P	59.44	917.309	917.3056	1620	72566
Iretol	C7 H8 O4	86.22	156.0423	156.0423	3205	43632
Ajugarin I	C24 H34 O7	67.66	434.2321	434.2305	2764	67391
3-O-alpha-L-Arabinopyranosylcinnamtannin B1	C50 H44 O22	60.96	996.2316	996.2324	2022	94409
5-Methyltetrahydropteroylpentaglutamate	C40 H53 N11 O18	57.79	975.3509	975.357	1709	6511
U-18666A	C25 H41 N O2	70.69	387.3136	387.3137	2983	45165
Mauritine A	C32 H41 N5 O5	71.38	575.3113	575.3108	2424	86200
APC	C33 H38 N4 O8	73.37	618.2668	618.269	2688	71230
Pyridoxal (Vitamin B6)	C8 H9 N O3	47.57	167.0581	167.0582	3991	2203
5,7,11,14-Eicosatetraenoic acid, 9-oxo-, (E,Z,Z,Z)-	C20 H30 O3	84.31	318.219	318.2195	4703	75009
Hydroxyitraconazole	C35 H38 Cl2 N8 O5	51.49	720.2413	720.2342	1927	878
Kaempferol 3-neohesperidoside-7-(6"-malonylglucoside)	C37 H44 O24	58.68	872.2153	872.2223	2308	50380
CAY10598	C21 H29 N5 O2	66.68	383.2299	383.2321	2324	45515
5-Hydroxy-3,3',7,8-tetramethoxy-4',5'-methylenedioxyflavone	C20 H18 O9	47.62	402.0951	402.0951	3079	95674
Lisuride	C20 H26 N4 O	76.61	338.2085	338.2107	2255	1010
Foeniculoside IV	C60 H62 O24	59.72	1166.3611	1166.3631	2046	94669
Ser Ile Lys	C15 H30 N4 O5	55.91	346.2185	346.2216	3410	20274
p-Salicylic acid	C7 H6 O3	47.62	138.0317	138.0317	4041	3263
Hexandraside C	C43 H56 O24	59.65	956.3105	956.3162	2240	50137
13-Hydroxy-2-(hydroxymethylene)-3-oxo-13,17-seco-5alpha-androstan-17-oic acid, delta-lactone	C20 H28 O4	33.94	332.2038	332.1988	1878	70923
Veratridine	C36 H51 N O11	67.14	673.3468	673.3462	2156	3688
6,8-nonadienal	C9 H14 O	47.55	138.1043	138.1045	4643	75328
4-(2,6,6-Trimethylcyclohex-1-enyl)but-2-en-4-one	C13 H20 O	67.77	192.1499	192.1514	3143	88514
Estradiol mustard	C42 H50 Cl4 N2 O4	57.98	786.2493	786.2525	2104	73186
Okanin 4'-(4"-acetyl-6"-p-coumarylglucoside)	C32 H30 O14	71.99	638.1634	638.1636	2276	51963
Tumonoic Acid A	C19 H33 N O4	64.51	339.2413	339.241	1419	65396
6-Glucopyranosylprocyanidin B2	C36 H36 O17	68.04	740.1962	740.1952	1771	92320

Hydroxy-3-O-methyl-6_-naltrexol	C21 H27 N O5	80.65	373.1891	373.1889	8922	1431
5-Methyltetrahydropteroylpentaglutamate	C40 H53 N11 O18	57.53	975.3512	975.357	1558	6511
Kaempferol 3-(2",4"-di-(Z)-p-coumaroylrhamnoside)	C39 H32 O14	65.69	724.1791	724.1792	1709	50359
(25R)-26,26,26-trifluoro-1_,25-dihydroxyvitamin D3 / (25R)-26,26,26-trifluoro-1_,25-dihydroxycholecalciferol	C27 H41 F3 O3	99.36	470.2999	470.3008		42071
Calpeptin	C20 H30 N2 O4	99.64	362.22	362.2206		68942
Coriandrone E	C13 H12 O5	99.98	248.0686	248.0685		86628
Enterocin 900	C31 H33 N O2	76.75	451.2568	451.2511		93009
CE(16:1(9Z))	C43 H74 O2	86.06	622.5761	622.5689	22076	5629
methyl 4-[2-(2-formyl-vinyl)-3-hydroxy-5-oxo-cyclopentyl]-butanoate	C13 H18 O5	39.49	254.1185	254.1154	378	74530
Vitamin D3 butyrate	C31 H50 O2	91.41	454.3772	454.3811	13718	42580
Acetyl vitamin K5	C13 H13 N O2	99.62	215.0941	215.0946	305414	88800
DG(20:4(8Z,11Z,14Z,17Z)/20:1(11Z)/0:0)	C43 H74 O5	90.04	670.5602	670.5536	8018	59031
PHOME	C23 H19 N O4	83.54	373.1315	373.1314	12515	64666
DG(20:4(8Z,11Z,14Z,17Z)/20:1(11Z)/0:0)	C43 H74 O5	89.95	670.56	670.5536	6643	59031
18:0 Cholesteryl ester	C45 H80 O2	64.21	652.6108	652.6158	5681	83955
hentriacontan-16-ol	C31 H64 O	89.07	452.4935	452.4957	6801	46046
PHOP	C18 H18 N2 O2	79.99	294.1392	294.1368	7968	63082
(6R)-vitamin D3 6,19-sulfur dioxide adduct / (6R)-cholecalciferol 6,19-sulfur dioxide adduct	C27 H44 O3 S	63.72	448.3049	448.3011	6589	42187
Betavulgaroside VI	C47 H72 O21	81.24	972.4481	972.4566	3165	89261
17:1 Cholesteryl ester	C44 H76 O2	59.63	636.5917	636.5845	4291	41718
28-Glucosyl-30-methyl-3b,23-dihydroxy-12-oleanene-28,30-dioate 3-[arabinosyl-(1->3)-glucuronide]	C48 H74 O21	79.67	986.4637	986.4723	2933	93714
(22Z)-1_-hydroxy-22,23-didehydrovitamin D3 / (22Z)-1_-hydroxy-22,23-didehydrocholecalciferol	C27 H42 O2	74.32	398.3141	398.3185	4812	42090
CE(19:0)	C46 H82 O2	63.1	666.6263	666.6315	3532	58537
Nb-Lignoceroyltryptamine	C34 H58 N2 O	66.14	510.4504	510.4549	3644	95281
(22Z)-1_-hydroxy-22,23-didehydrovitamin D3 / (22Z)-1_-hydroxy-22,23-didehydrocholecalciferol	C27 H42 O2	73.76	398.3143	398.3185	4680	42090
Solanidine	C27 H43 N O	75.82	397.3348	397.3345	4366	3517
hentriacontan-16-ol	C31 H64 O	55.11	452.4924	452.4957	4007	46046
tritriacontane-16,18-dione	C33 H64 O2	54.1	492.4845	492.4906	2475	46575
9(S)-HpODE	C18 H32 O4	84.32	312.2303	312.2301	3616	36019
Bis(glutathionyl)spermine disulfide	C30 H54 N10 O10 S2	53.02	778.3395	778.3466	2396	63646
Palmitoyl 3-carbacyclic Phosphatidic Acid	C20 H39 O5 P	66.69	390.2521	390.2535	4912	44867
N-(2R-methyl-3-hydroxy-ethyl)-16,16-dimethyl-5Z,8Z,11Z,14Z-docosatetraenoyl amine	C27 H47 N O2	82.65	417.3603	417.3607	6584	36712
DG(18:4(6Z,9Z,12Z,15Z)/18:1(11Z)/0:0)	C39 H66 O5	63.82	614.4973	614.491	2443	58872
Janex-1	C16 H15 N3 O3	74.25	297.1098	297.1113	4294	45316
Phe Arg Leu	C21 H34 N6 O4	70.81	434.2618	434.2642	4442	21813
14:0 Cholesteryl ester	C41 H72 O2	66.51	596.5483	596.5532	2550	41700
Hydroxyphthioceranic acid (C37)	C37 H74 O3	61.29	566.5605	566.5638	4454	73736
soladulcidine	C27 H45 N O2	99.18	415.3452	415.345	62643	41803
11-cis-retinyl palmitate	C36 H60 O2	71.42	524.4552	524.4593	2648	53832