




Article

In Silico Analysis of P450s and Their Role in Secondary Metabolism in the Bacterial Class *Gammaproteobacteria*

Ntombizethu Nokuphiwa Msomi ¹, Tiara Padayachee ¹, Nomfundo Nzuzwa ¹, Puleng Rosinah Syed ², Justyna Dorota Kryś ³, Wanping Chen ⁴, Dominik Gront ^{3,*}, David R. Nelson ^{5,*} and Khajamohiddin Syed ^{1,*}

¹ Department of Biochemistry and Microbiology, Faculty of Science and Agriculture, University of Zululand, KwaDlangezwa 3886, South Africa; msomizethu085@gmail.com (N.N.M.); teez07padayachee@gmail.com (T.P.); nomfundonzuzwa11@gmail.com (N.N.)

² Department of Pharmaceutical Chemistry, College of Health Sciences, University of KwaZulu-Natal, Durban 4000, South Africa; prosinah@gmail.com

³ Biological and Chemical Research Center, Faculty of Chemistry, University of Warsaw, Pasteura 1, 02-093 Warsaw, Poland; juchxd@gmail.com

⁴ Department of Molecular Microbiology and Genetics, University of Göttingen, 37077 Göttingen, Germany; chenwanping1@foxmail.com

⁵ Department of Microbiology, Immunology and Biochemistry, University of Tennessee Health Science Center, Memphis, TN 38163, USA

* Correspondence: dgront@gmail.com (D.G.); dnelson@uthsc.edu (D.R.N.); khajamohiddinsyed@gmail.com (K.S.); Tel.: +48-225526367 (D.G.); +1-901-448-8303 (D.R.N.); +27-035-902-6857 (K.S.)



Citation: Msomi, N.N.; Padayachee, T.; Nzuzwa, N.; Syed, P.R.; Kryś, J.D.; Chen, W.; Gront, D.; Nelson, D.R.; Syed, K. In Silico Analysis of P450s and Their Role in Secondary Metabolism in the Bacterial Class *Gammaproteobacteria*. *Molecules* **2021**, *26*, 1538. <https://doi.org/10.3390/molecules26061538>

Academic Editor: Daniel Krug

Received: 12 February 2021

Accepted: 8 March 2021

Published: 11 March 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: The impact of lifestyle on shaping the genome content of an organism is a well-known phenomenon and cytochrome P450 enzymes (CYPs/P450s), heme-thiolate proteins that are ubiquitously present in organisms, are no exception. Recent studies focusing on a few bacterial species such as *Streptomyces*, *Mycobacterium*, *Cyanobacteria* and *Firmicutes* revealed that the impact of lifestyle affected the P450 repertoire in these species. However, this phenomenon needs to be understood in other bacterial species. We therefore performed genome data mining, annotation, phylogenetic analysis of P450s and their role in secondary metabolism in the bacterial class *Gammaproteobacteria*. Genome-wide data mining for P450s in 1261 *Gammaproteobacteria* species belonging to 161 genera revealed that only 169 species belonging to 41 genera have P450s. A total of 277 P450s found in 169 species grouped into 84 P450 families and 105 P450 subfamilies, where 38 new P450 families were found. Only 18% of P450s were found to be involved in secondary metabolism in *Gammaproteobacteria* species, as observed in *Firmicutes* as well. The pathogenic or commensal lifestyle of *Gammaproteobacteria* species influences them to such an extent that they have the lowest number of P450s compared to other bacterial species, indicating the impact of lifestyle on shaping the P450 repertoire. This study is the first report on comprehensive analysis of P450s in *Gammaproteobacteria*.

Keywords: cytochrome P450 monooxygenases; *Gammaproteobacteria*; genome data mining; lifestyle; secondary metabolism; biosynthetic gene clusters; *Firmicutes*; *Streptomyces*; *Mycobacterium*; *Cyanobacteria*

1. Introduction

The phylum *Proteobacteria* is the most metabolically diverse bacterial phylum [1,2], constituting most of the known bacteria of medicinal, industrial, and agricultural significance [1,2]. Based on the 16S rRNA gene sequence and multiprotein phylogenetic analysis, *Proteobacteria* are divided into seven classes: *Alphaproteobacteria*, *Betaproteobacteria*, *Gammaproteobacteria*, *Deltaproteobacteria*, *Epsilonproteobacteria*, *Zetaproteobacteria* and *Acidithiobacillia* [3]. *Proteobacteria* members are well-known human pathogens; especially the class *Gammaproteobacteria* contain bacterial species that are medically, ecologically and scientifically important [1,2]. *Gammaproteobacteria* contain species with different physiology such as anaerobic, microaerophilic, and facultative aerobic, as well as different shapes such as rods, cocci or curve shapes [1]. Most of the bacterial species in this class are commensals

with the ability to cause diseases and some of them are strictly pathogenic. Below we provide brief information on well-known pathogens of this class (Table 1).

Table 1. Information on well-known pathogens of the bacterial class *Gammaproteobacteria*.

Species Name	General Information	Reference(s)
<i>Citrobacter freundii</i>	It is a commensal resident of the intestinal tracts of both humans and animals. It causes diarrhea and other infections in humans.	[4]
<i>Citrobacter koseri</i>	It is present as part of human flora. It causes meningitis in neonates, infants and immune-compromised people and in adults it causes urinary tract infections.	[5]
<i>Enterobacter cloacae</i>	It is a common nosocomial pathogen capable of producing a wide variety of infections, such as pneumonia, urinary tract infections, and septicemia.	[6]
<i>Enterobacter aerogenes</i>	It is a commensal of humans and an opportunistic pathogen. It causes respiratory, urinary, blood, or gastrointestinal tract infections.	[7]
<i>Enterobacter sakasaki</i>	It is a food-borne pathogen causing a severe form of sepsis, necrotizing enterocolitis and meningitis in neonatal infants.	[8]
<i>Escherichia coli</i>	It is commonly found in human flora. Pathogenic <i>E. coli</i> can be divided into two groups: extra-intestinal pathogens associated mainly with neonatal meningitis and urinary tract infections in adults and intestinal pathogens causes diarrhea.	[9]
<i>Escherichia albertii</i>	It is an emerging human enteropathogen and avian pathogen that causes diarrhea.	[10,11]
<i>Klebsiella pneumoniae</i>	It used to be an opportunistic pathogen causing a wide range of infections in immunocompromised patients, but has recently emerged as a pathogen of healthy and immunocompetent people. It causes pneumonias, urinary tract infections, bacteremia, and liver abscesses.	[12,13]
<i>Klebsiella oxytoca</i>	It is an emerging pathogen causing a wide range of infections similar to <i>K. pneumoniae</i> . <i>K. oxytoca</i> causes neonatal infections of the bloodstream, urinary tract, central nervous system, lung, skin, and soft tissues.	[14]
<i>Proteus mirabilis</i>	It is a human opportunistic pathogen. It causes infections of the urinary tract, including cystitis and pyelonephritis. It is also found in cases of asymptomatic bacteriuria, particularly in the elderly and patients with type 2 diabetes.	[15–17]
<i>Proteus vulgaris</i>	It is also an opportunistic pathogen causing infections such as <i>P. mirabilis</i> .	[18]
<i>Salmonella enterica</i>	This food-borne pathogen causes typhoid and paratyphoid fever in humans.	[19]
<i>Serratia marcescens</i>	It is an opportunistic pathogen causing bacteremia/sepsis.	[20]
<i>Shigella dysenteriae</i> , <i>S. flexneri</i> , <i>S. sonnei</i> and <i>S. boydii</i>	<i>Shigella</i> spp. are responsible for acute diarrhea and are major contributors to the global diarrheal disease burden.	[21,22]
<i>Yersinia pestis</i>	It causes bubonic plague and pneumonia. It is well-known to cause one of the most devastating diseases of human history, the black death.	[23]
<i>Yersinia enterocolitica</i>	It causes acute diarrhea, mesenteric adenitis, terminal ileitis, and pseudo-appendicitis.	[24]
<i>Yersinia pseudotuberculosis</i>	It causes mesenteric lymphadenitis, diarrhea, and septicemia in humans.	[25]

Because of the pathogenic nature of these species, quite a large number of Gammaproteobacterial species' genomes have been sequenced and are available for public use at different genome databases, including the Kyoto Encyclopedia of Genes and Genomes (KEGG) [26], the National Center for Biotechnology Information (NCBI) (<https://www.>

ncbi.nlm.nih.gov/genome/microbes/) (accessed on 15 January 2021) and Joint Genome Institute Integrated Microbial Genomes and Microbiomes [27]. Most of these genome-sequencing articles focused on finding traits that play a role in pathogenesis to come up with novel drug targets. However, key enzymes such as cytochrome P450 monooxygenases (CYPs/P450s) have been neglected or not reported in the genome sequencing studies owing to their typical nomenclature criteria that involve identification of key motifs and assigning the family and subfamily based on International P450 Nomenclature rules [28–30]. Only a handful of studies reported the presence of a number of P450s in certain groups of *Gammaproteobacteria* [31,32]. P450s are heme-thiolate proteins known to be involved in organisms' primary and secondary metabolism and are found in species belonging to different biological kingdoms [33], including in non-living entities such as viruses [34]. P450s perform enzymatic reactions in regio- and stereo-specific manners [35,36] and because of this inherent property they have been under the spotlight for the last 57 years to harness their potential [37]. P450s' enzymatic reactions are essential in assessing the drug toxicity of prodrugs and xenobiotic compounds have been well-established [38]. P450s serve as drug targets against pathogenic microbes [39–41]. Well-known examples of P450s targeted are the CYP51 [39] and CYP53 [40] families. P450s have been investigated for various biotechnological, environmental and pharmaceutical applications and recently, their role in the generation of secondary metabolites valuable to humans has been thoroughly reviewed [42,43]. One of the best examples of the role of P450s is antibiotic production, such as in the biosynthesis of the macrolide antibiotic erythromycin and glycopeptide antibiotics [44,45] and in the production of the anticancer drugs taxol and perillyl alcohol [46,47]. Gammaproteobacterial species, especially marine species, are well-known producers of a variety of secondary metabolites [48]. These metabolites include antibacterial, antiviral and anti-tumor compounds (violacein, pentabromopseudilin, alterochromides and thiomarinol) and a variety of siderophores (vibrioferrin, anguibactin, vanchrobactin, vibriobactin, avaroferrin, putrebactin, and bisucaberin) [48]. The synthesis of secondary metabolites such as thiomarinol, the alterochromides, pentabromopseudilin, and violacein compounds has been studied in detail and their synthesis has been linked to the biosynthetic gene clusters [48]. Information on Gammaproteobacterial species producing different secondary metabolites, secondary metabolites' biological properties and their genetic information has been reviewed [48].

Recent studies from our laboratory focusing on bacterial species revealed that the lifestyle of an organism plays a key role in shaping P450 content vis à vis P450s helping microbes to adapt to ecological niches. However, these studies were limited to few groups of bacterial species belonging to the genera *Streptomyces* [49,50] and *Mycobacterium* [51–53] and from the phyla *Cyanobacteria* [54] and *Firmicutes* [55,56]. Furthermore, to date, a comparative analysis of P450s in the Gammaproteobacterial species has not been reported. Thus, this study is aimed at addressing these research gaps by not only performing genome data mining, annotation and phylogenetic analysis of P450s in Gammaproteobacterial species, but also performing comparative analysis with other bacterial species to see the impact of lifestyle shaping the P450 content in this group of bacteria, among others with respect to P450s' involvement in secondary metabolism.

2. Results and Discussion

2.1. Only Few Gammaproteobacterial Species Have P450s

Genome-wide data mining for P450s in 1261 Gammaproteobacterial species belonging to 161 genera revealed the presence of P450s only in 169 Gammaproteobacterial species, indicating that most of these species do not have P450s in their genomes (Figure 1 and Supplementary Dataset 1; Sheet 1). Analysis of the *Gammaproteobacteria* genera disclosed that of the 161 genera, species belonging to 41 genera have P450s in their genomes (Supplementary Dataset 1; Sheet 1). A point to be noted is that the number of species genomes available in the 120 genera is very low. Sometimes only a single species genome is available and thus future availability of a higher number of species genomes will provide more accurate infor-

mation on P450s in these genera (Supplementary Dataset 1: Sheet 1). However, a significant number of species belonging to genera such as *Shewanella*, *Aeromonas* and *Haemophilus* were analyzed in this study and no P450s were found, suggesting that species in these genera probably do not have P450s (Supplementary Dataset 1: Sheet 1). Genera-level comparison revealed that at least 50% of *Pseudomonas* and *Xanthomonas* species have P450s. Analysis of P450s in 169 Gammaproteobacterial species revealed the presence of 277 P450s in their genomes (Supplementary Dataset 1: Sheet 2). The P450 count in the Gammaproteobacterial species ranged from a single P450 to six P450s; *Gammaproteobacterium* HdN1 had the highest number of P450s in its genome (Supplementary Dataset 1: Sheet 2). Apart from the complete P450 sequences (277 P450s), one P450 fragment and 10 P450 false-positive hits were also found in some Gammaproteobacterial species (Supplementary Dataset 1: Sheet 2). Comparative analysis of the average number of P450s with *Firmicutes* species, *Streptomyces*, mycobacterial species and cyanobacterial species revealed that Gammaproteobacterial species have the lowest average number of P450s (0.2%) in their genomes (Table 2). Gammaproteobacterial species P450 sequences along with P450 fragment and false positive hit proteins were presented in Supplementary Dataset 2.

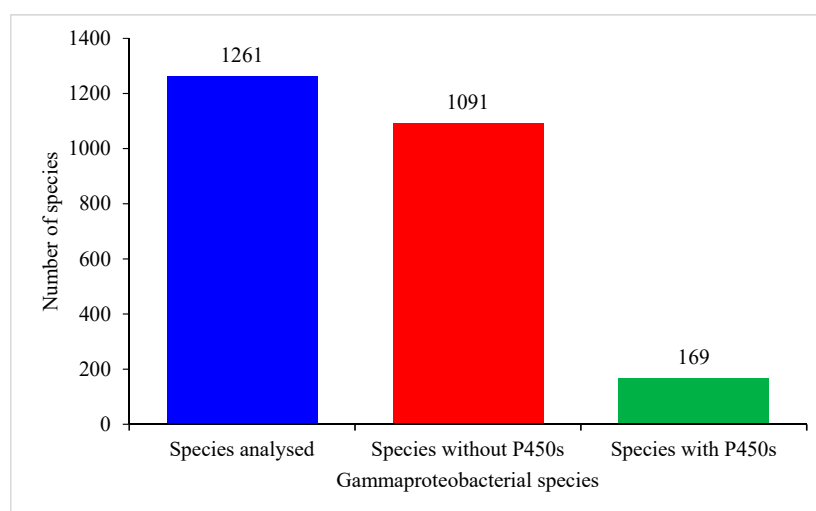


Figure 1. Analysis of P450s in Gammaproteobacterial species. Detailed analysis of the species and P450s is presented in Supplementary Dataset 1.

2.2. A Few P450 Families Are Expanded in Gammaproteobacterial Species

Based on the P450 nomenclature criteria [29,30], 277 P450s of Gammaproteobacterial species can be classified into 84 P450 families and 105 P450 subfamilies (Table 3). Annotation of P450s in this study was further verified by performing phylogenetic analysis (Figure 2). As shown in Figure 2, P450s belonging to the same family grouped together, indicating the correct annotation of P450s. Gammaproteobacterial species P450s identified in this study, along with their protein sequences and species, are presented in Supplementary Dataset 2. The number of P450 families in Gammaproteobacterial species is found to be highest compared to *Firmicutes* species, mycobacterial species and cyanobacterial species, indicating the highest diversity of P450s in Gammaproteobacterial species (Table 2). However, the number of P450 families in Gammaproteobacterial species was found to be lowest compared to *Streptomyces* species (Table 2). In order to understand the P450 diversity in Gammaproteobacterial species, the P450 diversity percentage was calculated by modifying the previously proposed formulae [54,57] to reflect the actual number of species with P450s. The new formula is presented below:

$$\text{P450 diversity percentage} = 100 \times \frac{\text{Total number of P450 families}}{(\text{Total number of P450s} \times \text{Number of species with P450s})} \quad (1)$$

Table 2. Comparative analysis of key features of P450s and their association with secondary metabolism between Gammaproteobacterial species and different bacterial species. BGCs: biosynthetic gene clusters.

	Gammaproteobacterial Species	Firmicutes Species	Streptomyces Species	Mycobacterial Species	Cyanobacterial Species
Total no. of species analyzed	1261	972	203	60	114
No of Species have P450s	169	229	203	60	114
No. of P450s	277	712	5460	1784	341
No. of families	84	14	253	77	36
No. of subfamilies	105	53	698	132	79
Dominant P450 family	CYP133	CYP107	CYP107	CYP125	CYP110
Average no. of P450s	0.2	1	27	30	3
P450 diversity percentage	0.18	0.008	0.02	0.07	0.09
No. of P450s part of BGCs	49	126	1231	204	27
Percentage of P450s part of BGCs	18	18	23	11	8
Reference(s)	This work	[55,56]	[49,50]	[49,51]	[54]

Table 3. Comparative analysis of P450 families and subfamilies in Gammaproteobacterial species.

Family	Subfamily	P450 Count	Percentage Contribution
CYP101	M	1	0.36
CYP102	J	1	0.36
CYP1043	C	1	0.36
CYP1049	A	1	0.36
CYP105	BQ	4	1.44
CYP107	DJ	3	7.94
	E	4	
	HT	1	
	S	14	
CYP108	R	1	0.36
CYP1097	B	1	0.36
CYP1104	B	3	1.08
CYP111	A	1	0.36
CYP112	A	3	1.08
CYP1138	A	1	0.36
CYP1142	A	1	0.36
CYP114	A	4	1.44
CYP1157	C	1	0.36
CYP115	A	3	1.08
CYP1164	A	1	0.36
CYP1165	A	1	0.72
	B	1	
CYP116	B	2	0.72
CYP1172	A	1	0.36
CYP1179	A	1	0.36

Table 3. Cont.

Family	Subfamily	P450 Count	Percentage Contribution
CYP117	A	3	1.08
CYP1199	A	2	0.72
CYP1200	B	1	0.36
CYP1201	A	1	0.72
	B	1	
CYP1202	A	1	0.36
CYP1225	A	1	0.36
CYP1234	A	5	2.17
	B	1	
CYP1247	A	9	3.25
CYP1261	B	1	0.36
CYP126	C	4	1.44
CYP1278	A	1	0.36
CYP1311	A	1	0.36
CYP133	B	22	8.30
	D	1	
CYP134	B	2	0.72
CYP136	F	2	0.72
CYP1413	A	3	1.44
	B	1	
CYP1414	A	1	0.72
	B	1	
CYP1415	A	3	1.08
CYP1464	A	2	0.72
CYP1465	A	4	1.44
CYP1466	A	1	0.36
CYP1467	A	1	0.36
CYP1468	A	1	0.36
CYP1469	A	2	0.72
CYP1470	A	3	1.08
CYP1471	A	1	0.36
CYP1472	A	1	0.36
CYP1473	A	1	0.36
CYP1474	A	3	2.53
	B	2	
	C	2	
CYP1475	A	1	0.36
CYP1476	A	1	0.36
CYP1477	A	1	0.36
CYP1478	A	1	0.36

Table 3. Cont.

Family	Subfamily	P450 Count	Percentage Contribution
CYP1479	A	1	0.36
CYP1480	A	1	0.36
CYP1481	A	2	0.72
CYP1482	A	1	0.36
CYP1483	A	1	0.36
CYP151	A	2	0.72
CYP152	G	3	3.25
	R	2	
	W	2	
	X	1	
	Y	1	
CYP153	A	16	6.86
	E	3	
CYP159	B	3	1.08
CYP168	A	18	7.22
	B	1	
	C	1	
CYP169	A	13	4.69
CYP172	B	1	0.36
CYP177	E	2	0.72
CYP198	A	3	1.08
CYP221	A	2	0.72
CYP226	A	5	1.81
CYP229	A	7	6.50
	D	7	
	E	4	
CYP234	A	1	0.36
CYP236	A	3	1.08
CYP238	A	1	0.36
CYP261	D	1	0.36
CYP289	A	6	2.53
	D	1	
CYP51	B	2	0.72
CYP107	S	1	0.36
CYP1229	A	1	0.36
CYP1414	A	1	0.36
CYP2242	A	1	0.36
CYP159	B	1	0.36
CYP163	K	1	0.36
CYP1779	A	1	0.36

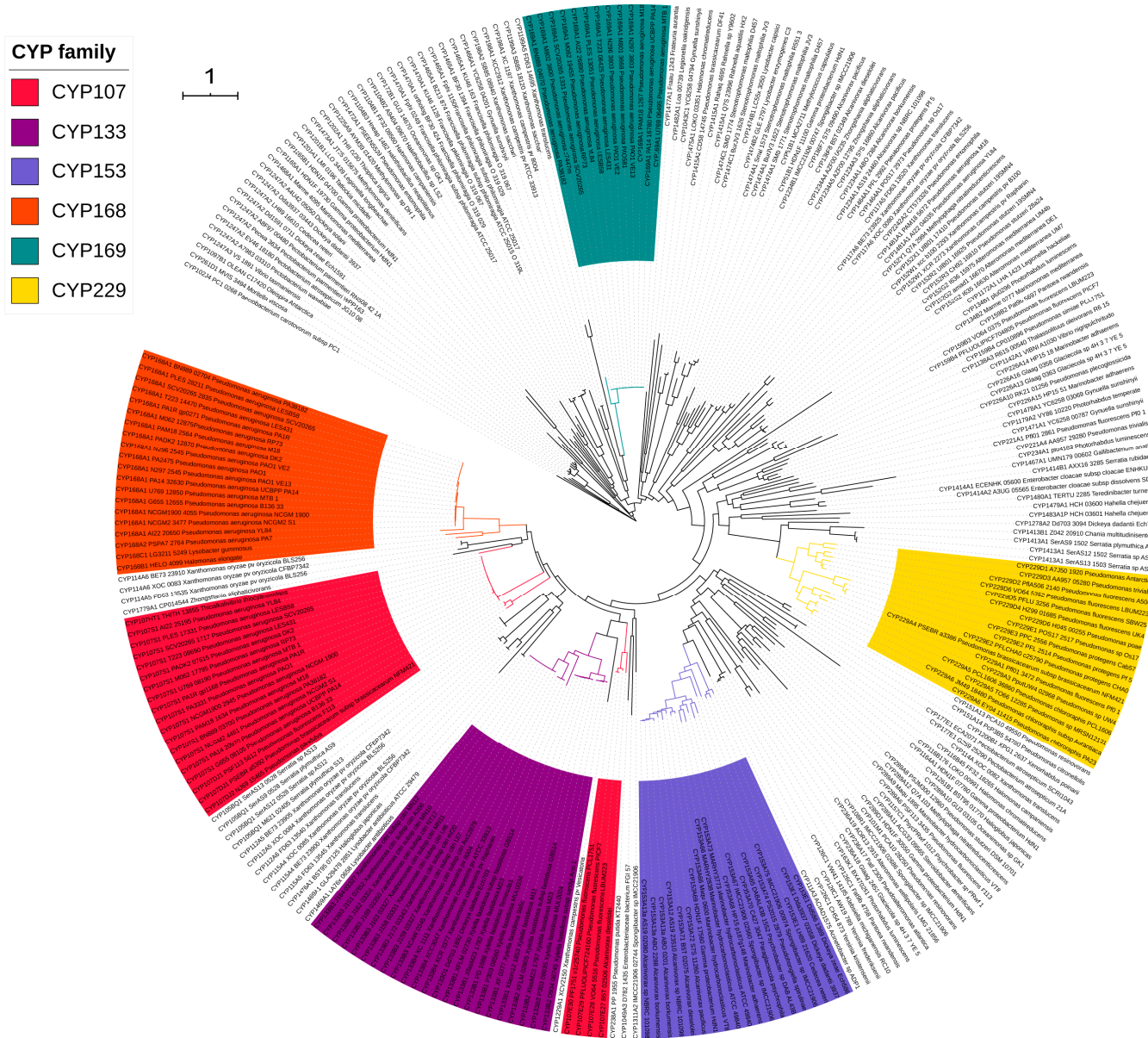


Figure 2. Phylogenetic analysis of Gammaproteobacterial species P450s. P450 families that are expanded in these species were highlighted in different colors and indicated in the figure.

Based on the above formulae, Gammaproteobacterial species were found to have the highest P450 diversity percentage compared to *Firmicutes* species, mycobacterial species, cyanobacterial species, and *Streptomyces* species (Table 2).

Among P450 families, the CYP133 has the highest number of member P450s with two-digit members: (23 P450s) contributing to 8.3% of total P450s in Gammaproteobacterial species (Table 3), followed by CYP107 (22 P450s contributing 7.9%), CYP168 (20 P450s contributing 7.2%), CYP153 (19 P450s contributing 6.8%), CYP229 (18 P450s contributing 6.5%) and CYP169 (13 P450s contributing 4.7%) (Table 3). The number of member P450s in the remaining 78 P450 families ranged from one to nine members (Table 3). The CYP133 family was found to be dominant in Gammaproteobacterial species, whereas in other bacterial species P450 families such as CYP107, CYP125 and CYP110 are dominantly present (Table 3). Considering the number of member P450s and the large number of species analyzed in this study, it can safely be said that these P450 families are expanded and not bloomed (a few P450 families with many genes) in Gammaproteobacterial species, which is contrary to other bacterial species where some P450 family blooming was observed [49–51,54]. The

number of subfamilies in a family ranged from one to five subfamilies; CYP152 has five subfamilies, followed by CYP107 with four subfamilies (Table 3). Interestingly, some kind of P450 subfamily level expansion was observed in Gammaproteobacterial species where subfamily S in the CYP07 family, subfamily B in the CYP133 family, subfamily A in the CYP153 family and subfamily A in the CYP168 family have the highest number of members (Table 3). Contrary to the other bacterial species, analysis of the conservation of P450 families in Gammaproteobacterial species revealed that none of the 84 P450 families were conserved in these species and that none of the P450 families were co-present (Figure 3). The comparative analysis of P450 families in different Gammaproteobacterial species is presented in Supplementary Dataset 1: Sheet 4.

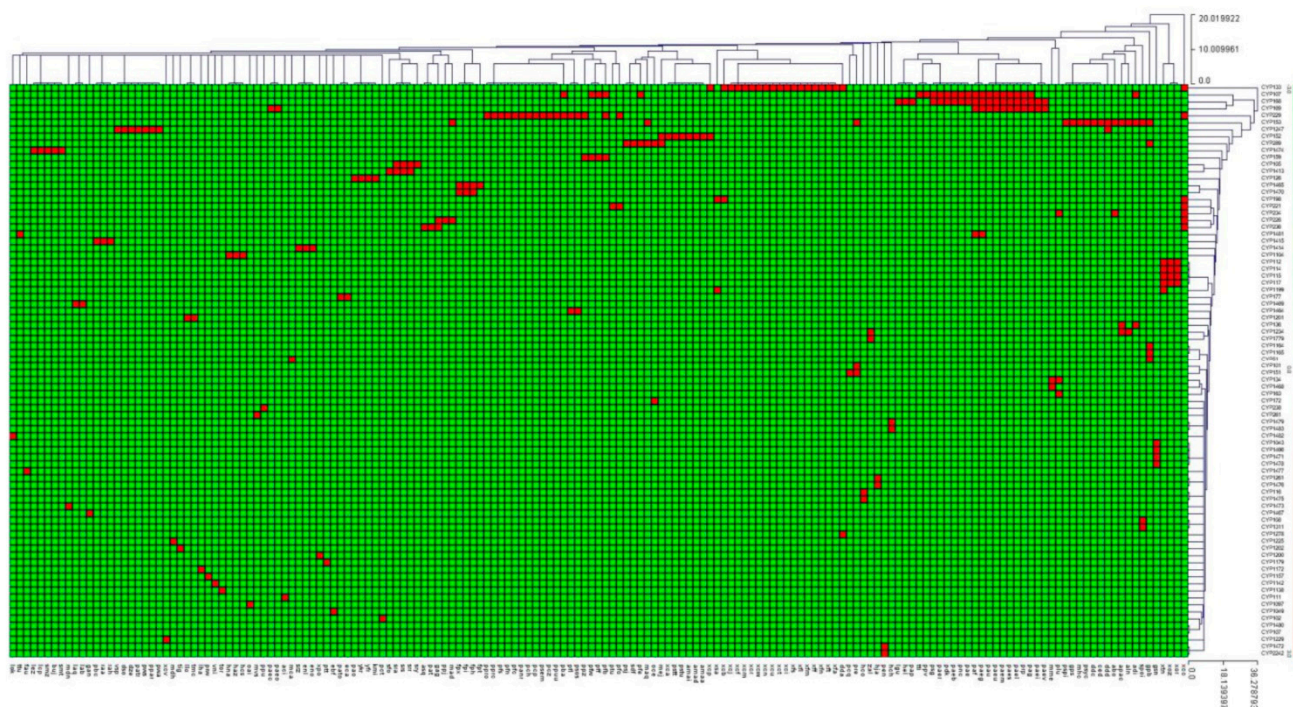


Figure 3. Heat map figure representing the presence or absence of cytochrome P450 families in 169 species of Gammaproteobacterial species. The data have been represented as -3 for family absence (green) and 3 for family presence (red). One-hundred and sixty-nine Gammaproteobacterial species form the horizontal axis and 84 P450 families form the vertical axis.

Analysis of P450 families revealed the presence of 38 new P450 families in Gammaproteobacterial species. Eighty-one P450s were found to be part of these 38 new P450 families. The list of new P450 families found in these species is presented in Supplementary Dataset 1: Sheet 5.

2.3. A Few P450s Are Involved in Secondary Metabolism in Gammaproteobacterial Species

Analysis of the P450s part of secondary metabolite biosynthetic gene clusters (BGCs) revealed that only a few P450s (18%) are part of these clusters, indicating their involvement in secondary metabolism in Gammaproteobacterial species (Figure 4). The percentage of P450s involved in Gammaproteobacterial species was found to be the same as in *Firmicutes* species (Table 2). Among 277 P450s, only 49 P450s belonging to 22 P450 families were found to be part of a secondary metabolite gene cluster (Figure 4). Among these families, CYP107 P450 family members were dominantly present (37%) in secondary metabolite BGCs (Figure 4). The same phenomenon was observed in *Streptomyces* species [49,50] and *Firmicutes* species [55], where CYP107 P450 family members were dominantly present as part of BGCs. Interestingly, among six P450 families that are expanded in these species, only two P450 family members were found to be part of the clusters (Figure 4). Among

these two P450 families, the CYP107 family has 18 members and the CYP153 family has two members involved in secondary metabolism. This suggests that P450 family expansion has no correlation with P450s' involvement in secondary metabolism in Gammaproteobacterial species, with the exception of the CYP107 family. The analysis of P450 secondary metabolite gene clusters revealed the presence of 12 types of BGCs, where the nonribosomal peptides (NRPS) were dominant (Supplementary Dataset 1: Sheet 2). One interesting point to be noted is that CYP107 family members are associated with the NRPS cluster (Supplementary Dataset 1: Sheet 2). The P450 secondary metabolite gene clusters showed the lowest percentage similarity to known secondary metabolite clusters, suggesting that these gene clusters probably produce novel compounds. Detailed information on P450s, their particular type of cluster and percentage identity to known secondary metabolite gene clusters is presented in Supplementary Dataset 1: Sheet 2).

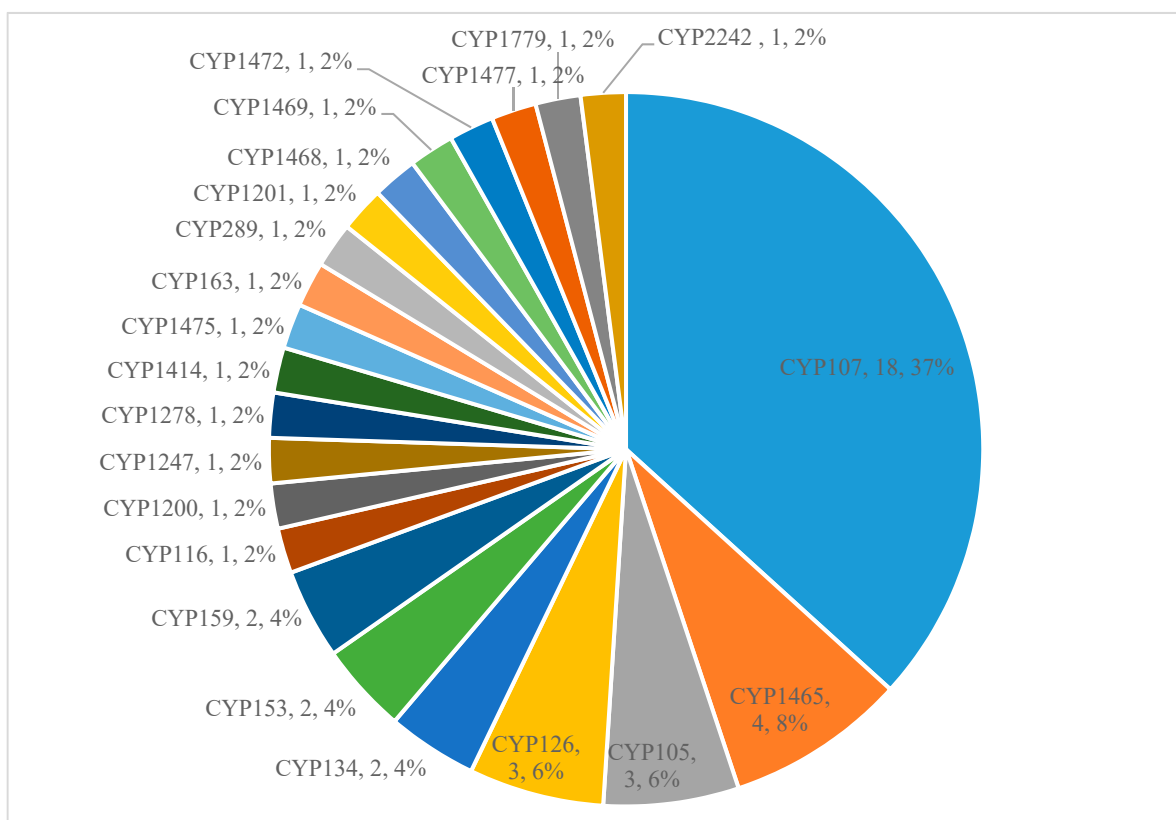


Figure 4. Comparative analysis of P450s involved in secondary metabolism in Gammaproteobacterial species. The P450 family name, number of P450s and the percentage of the total number of P450s that are part of BGCs are presented in the figure. Detailed information on secondary metabolite clusters, species and P450s is presented in Supplementary Dataset 1: Sheet 2.

2.4. Functional Prediction of Gammaproteobacterial P450s

Most of the Gammaproteobacterial P450s are orphans without an assigned biological function. Based on homolog P450s from other organisms, some of the Gammaproteobacterial P450s functions can be predicted. CYP51, also known as sterol 14 α -demethylase, is a highly conserved P450 across the phyla and stimulates a key enzymatic reaction that involves the stereo-selective three-step oxidative removal of the 14 α -methyl group from the sterol during the synthesis of membrane sterol [39]. CYP102 family members are involved in the hydroxylation of fatty acids [58–60]. CYP101 family members are involved in camphor hydroxylation [61]. CYP105, CYP107, and CYP109 have been found to display highly diverse functions [43,62,63]. CYP107 and CYP105 family members are involved in degradation and biotransformation of a large number of xenobiotic and secondary

metabolites [63,64]. P450s belonging to the CYP107, CYP109, and CYP134 families were found to hydroxylate different steroids, albeit with different substrate specificities [22]. CYP152 family members were found to be peroxygenases catalyzing the hydroxylation and decarboxylation of fatty acids [65–67]. CYP153 members are well-known alkane hydroxylases [68]. CYP226 family members are involved in the catabolism of dehydroabietic acid and abietic acid [69,70]. CYP261 family members were found to have a high affinity to fatty acids [71]. Based on the biosynthetic gene cluster analysis, in this study, we conclude that 49 P450s are involved in the synthesis of secondary metabolites (Supplementary Dataset 1: Sheet 2). However, the functions of these P450s in the synthesis of different secondary metabolites need to be determined.

2.5. Impact of Lifestyle on P450 Repertoire Is Also Evident in Gammaproteobacteria

It is now well known that lifestyle plays a key role in shaping the genome content of organisms, and P450s are no exception. The impact of lifestyle on P450 profiles in organisms or P450s' role in the adaptation of organisms to different ecological niches reported in animals [72], plants [73], oomycetes [57] and fungi [28,40,74–79] was scrutinized. A number of P450s or a number of members of a P450 family have been found to be present in large numbers in species compared to the pathogenic species or species adapted to a simpler lifestyle where simpler/abundant carbon sources are available for survival. This phenomenon of the impact of lifestyle shaping P450 content has also been observed in bacterial species belonging to the genera *Streptomyces* [49,50], *Mycobacterium* [51,52], and the phyla *Cyanobacteria* [54] and *Firmicutes* [55]. Especially in *Firmicutes*, it has been reported that the pathogenic/commensal lifestyle was influenced to such an extent that some of the species belonging to quite a large number of genera have no P450s in their genomes [55]. The same phenomenon of a low number or no P450s was also observed in Gammaproteobacterial species' genomes. Most of the bacterial species in the class *Gammaproteobacteria* are commensals with the ability to cause diseases and some of them are strictly pathogenic [1,2] which possibly led to have fewer P450s, as observed for *Firmicutes* [55]. The presence of the lowest number of P450s compared to the large number of species analyzed in the study (Table 2) clearly indicates that these bacterial species adapted to mostly living on simple carbon sources available in the host environment, which led to having fewer P450s in their genomes as observed in other species [55,77]. However, a noticeable difference was that the Gammaproteobacterial species had the highest number of P450 families and subfamilies compared to *Firmicutes* species, and thus the highest P450 diversity percentage compared to other bacterial species (Table 2). Nevertheless, the percentage of P450s involved in BGCs is the same as in *Firmicutes* species, indicating that most of the P450s play a role in primary metabolism. In summary, based on the number of P450s present in Gammaproteobacterial species and in comparison with other bacterial species, it is safe to say that the lifestyle of Gammaproteobacterial species profoundly affect their P450 repertoire, which is the same as what is observed in other organisms, as mentioned earlier.

3. Materials and Methods

3.1. Species and Their Genome Database Information

In this study, 1261 Gammaproteobacterial species belonging to 161 genera that are available for public use at the KEGG [80] database were used. Information on genera, species names, species codes and their genome IDs is presented in Supplementary Dataset 1: sheet 1.

3.2. Genome Data Mining and Annotation of P450s

P450 data mining and annotation were carried out following the standard procedure described previously by our laboratory [50,54,55]. Briefly, proteomes of each bacterial species were downloaded from the KEGG and subjected to the NCBI Batch Web CD-Search Tool [81]. The result was analyzed and proteins that belong to the P450 superfamily were

selected and searched for the presence of characteristic P450 motifs, EXXR and CXG [82,83]. Proteins that were short in amino acid length and lacking both motifs were regarded as P450 fragments and these P450 fragments were not considered for further analysis. Proteins having both motifs were selected and subjected to BLAST analysis against annotated P450s at the P450 website (<http://www.p450.unizulu.ac.za/>) (accessed on 15 January 2021). Based on the percentage identity to the named homolog P450s at the P450 website, selected P450s were assigned to specific P450 families and P450 subfamilies, following the International P450 Nomenclature Committee rule [28–30] that proteins with >40% identity and >55% identity will be grouped under the same family and subfamily, respectively. P450s with less than 40% identity were assigned to a new P450 family.

3.3. Phylogenetic Analysis of P450s

Phylogenetic analysis of P450s was carried out following the protocol previously described by our laboratory [54,55]. Briefly, the P450 protein sequences were aligned using the MAFFT v6.864 program [84], which is available at the Trex web server [85]. For tree inferring and optimization, the Trex web server [85] was used. Then, the server inferred the tree with different algorithms, including maximum likelihood, for the best phylogenetic tree in the least-squares sense. Finally, the best-inferred tree was visualized, colored and generated by the Interactive Tree of Life (iTOL) [86].

3.4. Generation of P450 Profile Heat Maps

The generation of the heat map profile was carried out according to the method previously reported by our laboratory [28,54,55]. Gammaproteobacterial P450 family data information was used to show the presence and absence of P450s in this class. The data were presented as positive 3 with the presence of P450 indicated in red and negative 3 indicated in green for the absence of P450. A tab-delimited file containing P450 presence and absence data was loaded into the multi-experiment viewer using a two-color array [87]. Hierarchical clustering using a Euclidean distance metric was used to cluster the data. P450 families formed the vertical axis and Gammaproteobacterial species formed the horizontal axis.

3.5. Identification of P450 Part of Secondary Metabolite BGCs

The identification of secondary metabolite BGCs and P450s forming part of the BGCs was carried out following the procedure previously reported by our laboratory [54–56]. Briefly, genome ID of Gammaproteobacterial species (Supplementary Dataset 1: sheet 1) having P450s was submitted to anti-SMASH (antibiotics & Secondary Metabolite Analysis Shell) [88] for the identification of secondary metabolite BGCs. Anti-SMASH results were downloaded both in the form of gene cluster sequences and Excel spreadsheets representing species-wise cluster information. P450s that formed part of a specific gene cluster were identified by manual data mining of gene cluster sequences. Standard gene cluster abbreviation terminology available at the anti-SMASH database [88] was maintained in this study.

3.6. Data Analysis

All calculations were carried out following the procedure reported previously by our laboratory [54]. The average number of P450s was calculated using the formula: Average number of P450s = Number of P450s/Number of species. The P450 diversity percentage was calculated using the formula presented in the main text of this article. The percentage of P450s that formed part of BGCs was calculated using the formula: Percentage of P450s part of BGCs = $100 \times \text{Number of P450s part of BGCs} / \text{Total number of P450s present in species}$.

3.7. Comparative Analysis of P450s and Gene Cluster Data

P450s and their BGCs data for *Firmicute* species [55], *Streptomyces* species [49,50], mycobacterial species [49,51] and cyanobacterial species [54] were retrieved from published articles and used for comparative analysis.

4. Conclusions

Sustained research into genome sequencing resulted in the sequencing of a large number of species genomes. This gives us an opportunity to look at specific protein families, their distribution and their role in organisms' physiology. This study is an example of such work, where cytochrome P450 monooxygenases (CYPs/P450s) were analyzed in the bacterial class *Gammaproteobacteria*. The study revealed that species in this class have few P450s in their genomes compared to other bacterial species. Interestingly, despite having a low number of P450s, Gammaproteobacterial species have the highest P450 diversity percentage compared to *Firmicutes* species, *Streptomyces* species, mycobacterial species and cyanobacterial species. The lifestyle of Gammaproteobacterial species, pathogenic or commensalism, had a profound impact on the P450 count in their genomes, leading to having a low number of P450s, as observed in *Firmicutes* species, indicating that the lifestyle of an organism plays a key role in shaping P450 content. Since P450 enzymes follow typical nomenclature criteria, this study will serve as a reference for P450 annotation in the class *Gammaproteobacteria*. Studies are in progress to determine the impact of lifestyle in other bacterial species.

Another interesting result from the study is the identification of secondary metabolite biosynthetic gene clusters that have P450s in the Gammaproteobacterial species. In this post-genomic era, before entering the laboratory one can analyze large numbers of genomes and select the best candidates for the production of unique/novel secondary metabolites. This approach is now gaining momentum, as it saves a tremendous amount of wet-laboratory work, money and laborious practice. In this study, we followed this approach and identified secondary metabolite biosynthetic gene clusters that seem to be novel, as they have the lowest percentage identity to known or characterized secondary metabolite gene clusters, indicating that the compounds synthesized by these gene clusters are novel and possibly have biotechnological values. Future studies include cloning of these novel clusters to enable the synthesis of metabolites and assessment of the biological properties of the synthesized metabolite.

Supplementary Materials: The following are available online, Supplementary Dataset 1: Sheet1: Information on species used in the study; Sheet 2: Secondary metabolite biosynthetic gene clusters analysis in Gammaproteobacterial species; Sheet 3: P450 profile heat map information; Sheet 4: Comparative analysis of P450 families in Gammaproteobacterial species; Sheet 5: New P450 families and their count in the bacterial class *Gammaproteobacteria*; Supplementary Dataset 2: Gammaproteobacterial species P450 sequences identified in this study. Each P450 sequence is presented with its assigned name, protein code (in parenthesis) and species name. Full-length P450 sequences along with P450-fragment and false positive hit proteins were presented.

Author Contributions: Conceptualization, K.S.; data curation, N.N.M., T.P., N.N., P.R.S., J.D.K., W.C., D.G., D.R.N. and K.S.; formal analysis, N.N.M., T.P., N.N., P.R.S., J.D.K., W.C., D.G., D.R.N. and K.S.; funding acquisition, D.G. and K.S.; investigation, N.N.M., T.P., N.N., P.R.S., J.D.K., W.C., D.G., D.R.N. and K.S.; methodology, N.N.M., T.P., N.N., P.R.S., J.D.K., W.C., D.G., D.R.N. and K.S.; project administration, K.S.; resources, K.S.; supervision, K.S.; validation, N.N.M., T.P., N.N., P.R.S., J.D.K., W.C., D.G., D.R.N. and K.S.; visualization, N.N.M., T.P., N.N., P.R.S., J.D.K., W.C., D.G., D.R.N. and K.S.; writing—original draft, N.N.M., T.P., N.N., P.R.S., J.D.K., W.C., D.G., D.R.N. and K.S.; writing—review and editing, N.N.M., T.P., N.N., P.R.S., J.D.K., W.C., D.G., D.R.N. and K.S. All authors have read and agreed to the published version of the manuscript.

Funding: Khajamohiddin Syed expresses sincere gratitude to the University of Zululand Research Committee for funding (Grant No. C686) and to the National Research Foundation (NRF), South Africa for a research grant (Grant No. 114159). Master's students Ntombizethu Nokuphiwa Msomi, Tiara Padayachee, Nomfundo Nzuza and Mbalenhle Sizamile Mfeka thank the NRF, South Africa for master's scholarships (Grant numbers SFH180530337436, MND190619448759, MND190626451135 and SFH170716253893, respectively). Puleng Rosinah Syed also thanks the NRF, South Africa for a doctoral scholarship (Grant number MND190606443406). Dominik Gront and Justyna D. Kryś were supported by the Polish National Science Centre (NCN) 2018/29/B/ST6/01989.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors want to thank Barbara Bradley, Pretoria, South Africa for English language editing.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Sample Availability: Samples of the compounds are not available from the authors.

References

1. Berman, J.J. *Taxonomic Guide to Infectious Diseases: Understanding the Biologic Classes of Pathogenic Organisms*; Academic Press: Cambridge, MA, USA, 2019.
2. Kersters, K.; De Vos, P.; Gillis, M.; Swings, J.; Vandamme, P.; Stackebrandt, E. Introduction to the *Proteobacteria*. In *The Prokaryotes: A Handbook on the Biology of Bacteria*; Springer: New York, NY, USA, 2006; Volume 5, pp. 3–37.
3. Williams, K.P.; Kelly, D.P. Proposal for a new class within the phylum *Proteobacteria*, *Acidithiobacillia* classis nov., with the type order *Acidithiobacillales*, and emended description of the class *Gammaproteobacteria*. *Int. J. Syst. Evol. Microbiol.* **2013**, *63*, 2901–2906. [[CrossRef](#)]
4. Liu, L.; Chen, D.; Liu, L.; Lan, R.; Hao, S.; Jin, W.; Sun, H.; Wang, Y.; Liang, Y.; Xu, J. Genetic diversity, multidrug resistance, and virulence of *Citrobacter freundii* from diarrheal patients and healthy individuals. *Front. Cell. Infect. Microbiol.* **2018**, *8*, 233. [[CrossRef](#)] [[PubMed](#)]
5. Stewart, Z.E.; Shaker, M.; Baxter, J.D. Urinary tract infection caused by *Citrobacter koseri* in a patient with spina bifida, an ileal conduit and renal caluli progressing to peri-nephric abscess and empyema. *Urol. Case Rep.* **2017**, *11*, 22–24. [[CrossRef](#)]
6. Annavajhala, M.K.; Gomez-Simmonds, A.; Uhlemann, A.-C. Multidrug-resistant *Enterobacter cloacae* complex emerging as a global, diversifying threat. *Front. Microbiol.* **2019**, *10*, 44. [[CrossRef](#)]
7. Davin-Regli, A. *Enterobacter aerogenes* and *Enterobacter cloacae*, versatile bacterial pathogens confronting antibiotic treatment. *Front. Microbiol.* **2015**, *6*, 392. [[CrossRef](#)]
8. Abdesselam, K.; Pagotto, F. Bacteria: *Cronobacter* (*Enterobacter*) *sakazakii* and Other *Cronobacter* spp. *Encycl. Food Saf.* **2014**, *1*, 424–432. [[CrossRef](#)]
9. Rojas-Lopez, M.; Monterio, R.; Pizza, M.; Desvaux, M.; Rosini, R. Intestinal pathogenic *Escherichia coli*: Insights for vaccine development. *Front. Microbiol.* **2018**, *9*, 440. [[CrossRef](#)]
10. Oaks, J.L.; Besser, T.E.; Walk, S.T.; Gordon, D.M.; Beckmen, K.B.; Burek, K.A.; Haldorson, G.J.; Bradway, D.S.; Ouellette, L.; Rurangirwa, F.R. *Escherichia albertii* in wild and domestic birds. *Emerg. Infect. Dis.* **2010**, *16*, 638. [[CrossRef](#)] [[PubMed](#)]
11. Ooka, T.; Tokuoka, E.; Furukawa, M.; Nagamura, T.; Ogura, Y.; Arisawa, K.; Harada, S.; Hayashi, T. Human gastroenteritis outbreak associated with *Escherichia albertii*, Japan. *Emerg. Infect. Dis.* **2013**, *19*, 144–146. [[CrossRef](#)]
12. Paczosa, M.K.; Meccas, J. *Klebsiella pneumoniae*: Going on the offense with a strong defense. *Microbiol. Mol. Biol. Rev.* **2016**, *80*, 629–661. [[CrossRef](#)]
13. Effah, C.Y.; Sun, T.; Liu, S.; Wu, Y. *Klebsiella pneumoniae*: An increasing threat to public health. *Ann. Clin. Microbiol. Antimicrob.* **2020**, *19*, 1–9. [[CrossRef](#)] [[PubMed](#)]
14. Nizet, V.; Klein, J. Bacterial sepsis and meningitis. *Infect. Dis. Fetus Newborn Infant* **2011**, *8*, 217–271.
15. Schaffer, J.N.; Pearson, M.M. *Proteus mirabilis* and urinary tract infections. In *Urinary Tract Infections: Molecular Pathogenesis and Clinical Management*; John Wiley & Sons: New York, NY, USA, 2017; pp. 383–433.
16. Matthews, S.J.; Lancaster, J.W. Urinary tract infections in the elderly population. *Am. J. Geriatr. Pharmacother.* **2011**, *9*, 286–309. [[CrossRef](#)]
17. Papazafropoulou, A.; Daniil, I.; Sotiropoulos, A.; Balampani, E.; Kokolaki, A.; Bousboulas, S.; Konstantopoulou, S.; Skliros, E.; Petropoulou, D.; Pappas, S. Prevalence of asymptomatic bacteriuria in type 2 diabetic subjects with and without microalbuminuria. *BMC Res. Notes* **2010**, *3*, 169. [[CrossRef](#)]

18. Drzewiecka, D. Significance and roles of *Proteus* spp. bacteria in natural environments. *Microb. Ecol.* **2016**, *72*, 741–758. [[CrossRef](#)] [[PubMed](#)]
19. Andino, A.; Hanning, I. *Salmonella enterica*: Survival, colonization, and virulence differences among serovars. *Sci. World J.* **2015**, *2015*, 1–16. [[CrossRef](#)] [[PubMed](#)]
20. Khanna, A.; Khanna, M.; Aggarwal, A. *Serratia marcescens*—A rare opportunistic nosocomial pathogen and measures to limit its spread in hospitalized patients. *J. Clin. Diagn. Res. JCDR* **2013**, *7*, 243.
21. Drancourt, M. Acute diarrhea. *Infect. Dis.* **2017**, *1*, 335–340.e2.
22. Baker, S.; The, H.C. Recent insights into *Shigella*: A major contributor to the global diarrhoeal disease burden. *Curr. Opin. Infect. Dis.* **2018**, *31*, 449. [[CrossRef](#)]
23. Ari, T.B.; Neerincx, S.; Gage, K.L.; Kreppel, K.; Laudisoit, A.; Leirs, H.; Stenseth, N.C. Plague and climate: Scales matter. *PLoS Pathog.* **2011**, *7*, e1002160.
24. Sabina, Y.; Rahman, A.; Ray, R.C.; Montet, D. *Yersinia enterocolitica*: Mode of transmission, molecular insights of virulence, and pathogenesis of infection. *J. Pathog.* **2011**, *2011*, 1–10. [[CrossRef](#)]
25. Fukushima, H.; Shimizu, S.; Inatsu, Y. *Yersinia enterocolitica* and *Yersinia pseudotuberculosis* detection in foods. *J. Pathog.* **2011**, *2011*, 735308. [[CrossRef](#)] [[PubMed](#)]
26. Kanehisa, M.; Sato, Y.; Kawashima, M.; Furumichi, M.; Tanabe, M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* **2016**, *44*, D457–D462. [[CrossRef](#)] [[PubMed](#)]
27. Chen, I.-M.A.; Chu, K.; Palaniappan, K.; Pillay, M.; Ratner, A.; Huang, J.; Huntemann, M.; Varghese, N.; White, J.R.; Seshadri, R.; et al. IMG/M v.5.0: An integrated data management and comparative analysis system for microbial genomes and microbiomes. *Nucleic Acids Res.* **2019**, *47*, D666–D677. [[CrossRef](#)]
28. Akapo, O.O.; Padayachee, T.; Chen, W.; Kappo, A.P.; Yu, J.H.; Nelson, D.R.; Syed, K. Distribution and diversity of cytochrome P450 monooxygenases in the fungal class *Tremellomycetes*. *Int. J. Mol. Sci.* **2019**, *20*, 2889. [[CrossRef](#)]
29. Nelson, D.R. Cytochrome P450 nomenclature. *Methods Mol. Biol.* **1998**, *107*, 15–24. [[PubMed](#)]
30. Nelson, D.R. Cytochrome P450 Nomenclature, 2004. *Methods Mol. Biol.* **2006**, *320*, 1–10. [[PubMed](#)]
31. Spring, S.; Scheuner, C.; Goker, M.; Klenk, H.P. A taxonomic framework for emerging groups of ecologically important marine *Gammaproteobacteria* based on the reconstruction of evolutionary relationships using genome-scale data. *Front. Microbiol.* **2015**, *6*, 281. [[CrossRef](#)]
32. Korzhenkov, A.A.; Toshchakov, S.V.; Golyshina, O.V.; Ferrer, M.; Chernikova, T.N.; Jaeger, K.-E.; Yakimov, M.M.; Golyshin, P.N. Aerobic hydrocarbon-degrading *Gammaproteobacteria*: *Oleiphilaceae* and relatives. *Taxon. Genom. Ecophysiol. Hydrocarb. Degrad. Microbes* **2019**, 153–166. [[CrossRef](#)]
33. Nelson, D.R. Cytochrome P450 diversity in the tree of life. *Biochim. Biophys. Acta (BBA) Proteins Proteom.* **2018**, *1866*, 141–154. [[CrossRef](#)]
34. Lamb, D.C.; Follmer, A.H.; Goldstone, J.V.; Nelson, D.R.; Warrilow, A.G.; Price, C.L.; True, M.Y.; Kelly, S.L.; Poulos, T.L.; Stegeman, J.J. On the occurrence of cytochrome P450 in viruses. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 12343–12352. [[CrossRef](#)]
35. Bernhardt, R. Cytochromes P450 as versatile biocatalysts. *J. Biotechnol.* **2006**, *124*, 128–145. [[CrossRef](#)]
36. Sono, M.; Roach, M.P.; Coulter, E.D.; Dawson, J.H. Heme-containing oxygenases. *Chem. Rev.* **1996**, *96*, 2841–2888. [[CrossRef](#)] [[PubMed](#)]
37. Yamazaki, H. *Fifty Years of Cytochrome P450 Research*; Springer: Tokyo, Japan, 2014; pp. IX, 409.
38. Guengerich, F.P. A history of the roles of cytochrome P450 enzymes in the toxicity of drugs. *Toxicol. Res.* **2021**, *37*, 1–23. [[CrossRef](#)]
39. Lepesheva, G.I.; Friggeri, L.; Waterman, M.R. CYP51 as drug targets for fungi and protozoan parasites: Past, present and future. *Parasitology* **2018**, *145*, 1820–1836. [[CrossRef](#)]
40. Jawallapersand, P.; Mashele, S.S.; Kovačič, L.; Stojan, J.; Komel, R.; Pakala, S.B.; Kraševc, N.; Syed, K. Cytochrome P450 monooxygenase CYP53 family in fungi: Comparative structural and evolutionary analysis and its role as a common alternative anti-fungal drug target. *PLoS ONE* **2014**, *9*, e107209. [[CrossRef](#)] [[PubMed](#)]
41. Ortiz de Montellano, P.R. Potential drug targets in the *Mycobacterium tuberculosis* cytochrome P450 system. *J. Inorg. Biochem.* **2018**, *180*, 235–245. [[CrossRef](#)] [[PubMed](#)]
42. Podust, L.M.; Sherman, D.H. Diversity of P450 enzymes in the biosynthesis of natural products. *Nat. Prod. Rep.* **2012**, *29*, 1251–1266. [[CrossRef](#)]
43. Greule, A.; Stok, J.E.; De Voss, J.J.; Cryle, M.J. Unrivalled diversity: The many roles and reactions of bacterial cytochromes P450 in secondary metabolism. *Nat. Prod. Rep.* **2018**, *35*, 757–791. [[CrossRef](#)]
44. Andersen, J.F.; Tatsuta, K.; Gunji, H.; Ishiyama, T.; Hutchinson, C.R. Substrate specificity of 6-deoxyerythronolide B hydroxylase, a bacterial cytochrome P450 of erythromycin A biosynthesis. *Biochemistry* **1993**, *32*, 1905–1913. [[CrossRef](#)]
45. Bischoff, D.; Bister, B.; Bertazzo, M.; Pfeifer, V.; Stegmann, E.; Nicholson, G.J.; Keller, S.; Pelzer, S.; Wohlleben, W.; Süßmuth, R.D. The biosynthesis of vancomycin-type glycopeptide antibiotics—A model for oxidative side-chain cross-linking by oxygenases coupled to the action of peptide synthetases. *ChemBioChem* **2005**, *6*, 267–272. [[CrossRef](#)] [[PubMed](#)]
46. Jennewein, S.; Park, H.; DeJong, J.M.; Long, R.M.; Bollon, A.P.; Croteau, R.B. Coexpression in yeast of *Taxus* cytochrome P450 reductase with cytochrome P450 oxygenases involved in Taxol biosynthesis. *Biotechnol. Bioeng.* **2005**, *89*, 588–598. [[CrossRef](#)] [[PubMed](#)]

47. van Beilen, J.B.; Holtackers, R.; Lüscher, D.; Bauer, U.; Witholt, B.; Duetz, W.A. Biocatalytic production of perillyl alcohol from limonene by using a novel *Mycobacterium* sp. cytochrome P450 alkane hydroxylase expressed in *Pseudomonas putida*. *Appl. Environ. Microbiol.* **2005**, *71*, 1737–1744. [[CrossRef](#)]
48. Timmermans, M.L.; Paudel, Y.P.; Ross, A.C. Investigating the biosynthesis of natural products from marine proteobacteria: A survey of molecules and strategies. *Mar. Drugs* **2017**, *15*, 235. [[CrossRef](#)]
49. Senate, L.M.; Tjatji, M.P.; Pillay, K.; Chen, W.; Zondo, N.M.; Syed, P.R.; Mnguni, F.C.; Chiliza, Z.E.; Bamal, H.D.; Karpoomath, R.; et al. Similarities, variations, and evolution of cytochrome P450s in *Streptomyces* versus *Mycobacterium*. *Sci. Rep.* **2019**, *9*, 1–12. [[CrossRef](#)]
50. Mnguni, F.C.; Padayachee, T.; Chen, W.; Gront, D.; Yu, J.-H.; Nelson, D.R.; Syed, K. More P450s are involved in secondary metabolite biosynthesis in *Streptomyces* compared to *Bacillus*, *Cyanobacteria* and *Mycobacterium*. *Int. J. Mol. Sci.* **2020**, *21*, 4814. [[CrossRef](#)]
51. Parvez, M.; Qhanya, L.B.; Mthakathi, N.T.; Kgosiemang, I.K.R.; Bamal, H.D.; Pagadala, N.S.; Xie, T.; Yang, H.; Chen, H.; Theron, C.W.; et al. Molecular evolutionary dynamics of cytochrome P450 monooxygenases across kingdoms: Special focus on mycobacterial P450s. *Sci. Rep.* **2016**, *6*, 33099. [[CrossRef](#)]
52. Syed, P.R.; Chen, W.; Nelson, D.R.; Kappo, A.P.; Yu, J.-H.; Karpoomath, R.; Syed, K. Cytochrome P450 monooxygenase CYP139 family involved in the synthesis of secondary metabolites in 824 *Mycobacterial* Species. *Int. J. Mol. Sci.* **2019**, *20*, 2690. [[CrossRef](#)] [[PubMed](#)]
53. Ngcobo, N.S.; Chiliza, Z.E.; Chen, W.; Yu, J.-H.; Nelson, D.R.; Tuszynski, J.A.; Preto, J.; Syed, K. Comparative analysis, structural insights, and substrate/drug interaction of CYP128A1 in *Mycobacterium tuberculosis*. *Int. J. Mol. Sci.* **2020**, *21*, 4816. [[CrossRef](#)]
54. Khumalo, M.J.; Nzuzza, N.; Padayachee, T.; Chen, W.; Yu, J.H.; Nelson, D.R.; Syed, K. Comprehensive analyses of cytochrome P450 monooxygenases and secondary metabolite biosynthetic gene clusters in *Cyanobacteria*. *Int. J. Mol. Sci.* **2020**, *21*, 656. [[CrossRef](#)]
55. Padayachee, T.; Nzuzza, N.; Chen, W.; Nelson, D.R.; Syed, K. Impact of lifestyle on cytochrome P450 monooxygenase repertoire is clearly evident in the bacterial phylum *Firmicutes*. *Sci. Rep.* **2020**, *10*, 1–12. [[CrossRef](#)]
56. Mthethwa, B.C.; Chen, W.; Ngwenya, M.L.; Kappo, A.P.; Syed, P.R.; Karpoomath, R.; Yu, J.-H.; Nelson, D.R.; Syed, K. Comparative analyses of cytochrome P450s and those associated with secondary metabolism in *Bacillus* Species. *Int. J. Mol. Sci.* **2018**, *19*, 3623. [[CrossRef](#)]
57. Sello, M.M.; Jafta, N.; Nelson, D.R.; Chen, W.; Yu, J.-H.; Parvez, M.; Kgosiemang, I.K.R.; Monyaki, R.; Raseleman, S.C.; Qhanya, L.B.; et al. Diversity and evolution of cytochrome P450 monooxygenases in *Oomycetes*. *Sci. Rep.* **2015**, *5*, 11572. [[CrossRef](#)] [[PubMed](#)]
58. Munro, A.W.; Leys, D.G.; McLean, K.J.; Marshall, K.R.; Ost, T.W.; Daff, S.; Miles, C.S.; Chapman, S.K.; Lysek, D.A.; Moser, C.C.; et al. P450 BM3: The very model of a modern flavocytochrome. *Trends Biochem. Sci.* **2002**, *27*, 250–257. [[CrossRef](#)]
59. Li, H.; Poulos, T.L. The structure of the cytochrome p450BM-3 haem domain complexed with the fatty acid substrate, palmitoleic acid. *Nat. Struct. Biol.* **1997**, *4*, 140–146. [[CrossRef](#)] [[PubMed](#)]
60. Noble, M.; Miles, C.; Reid, G.; Chapman, S.; Munro, A. Catalytic properties of key active site mutants of flavocytochrome P-450 BM. *Biochem. Soc. Trans.* **1999**, *27*, A44. [[CrossRef](#)]
61. Poulos, T.L.; Finzel, B.C.; Howard, A.J. Crystal structure of substrate-free *Pseudomonas putida* cytochrome P-450. *Biochemistry* **1986**, *25*, 5314–5322. [[CrossRef](#)] [[PubMed](#)]
62. McLean, K.J.; Leys, D.; Munro, A.W. Microbial cytochrome P450s. In *Cytochrome P450: Structure, Mechanism, and Biochemistry*, 4th ed.; Ortiz de Montellano, P.R., Ed.; Springer International Publishing: New York, NY, USA, 2015; Chapter 6; pp. 261–407.
63. Moody, S.C.; Loveridge, E.J. CYP105-diverse structures, functions and roles in an intriguing family of enzymes in *Streptomyces*. *J. Appl. Microbiol.* **2014**, *117*, 1549–1563. [[CrossRef](#)]
64. Li, Z.-Z.; Li, X.-F.; Yang, W.; Dong, X.; Yu, J.; Zhu, S.-L.; Li, M.; Xie, L.; Tong, W.-Y. Identification and functional analysis of cytochrome P450 complement in *Streptomyces virginiae* IBL₁₄. *BMC Genom.* **2013**, *14*, 130. [[CrossRef](#)] [[PubMed](#)]
65. Lee, D.-S.; Yamada, A.; Matsunaga, I.; Ichihara, K.; Adachi, S.-I.; Park, S.-Y.; Shiro, Y. Crystallization and preliminary X-ray diffraction analysis of fatty-acid hydroxylase cytochrome P450BS β from *Bacillus subtilis*. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2002**, *58*, 687–689. [[CrossRef](#)]
66. Lee, D.-S.; Yamada, A.; Sugimoto, H.; Matsunaga, I.; Ogura, H.; Ichihara, K.; Adachi, S.-I.; Park, S.-Y.; Shiro, Y. Substrate recognition and molecular mechanism of fatty acid hydroxylation by cytochrome P450 from *Bacillus subtilis* crystallographic, spectroscopic, and mutational studies. *J. Biol. Chem.* **2003**, *278*, 9761–9767. [[CrossRef](#)] [[PubMed](#)]
67. Girvan, H.M.; Poddar, H.; McLean, K.J.; Nelson, D.R.; Hollywood, K.A.; Levy, C.W.; Leys, D.; Munro, A.W. Structural and catalytic properties of the peroxygenase P450 enzyme CYP152K6 from *Bacillus methanolicus*. *J. Inorg. Biochem.* **2018**, *188*, 18–28. [[CrossRef](#)] [[PubMed](#)]
68. Funhoff, E.G.; Bauer, U.; García-Rubio, I.; Witholt, B.; van Beilen, J.B. CYP153A6, a soluble P450 oxygenase catalyzing terminal-alkane hydroxylation. *J. Bacteriol.* **2006**, *188*, 5220–5227. [[CrossRef](#)] [[PubMed](#)]
69. Smith, D.J.; Patrauchan, M.A.; Florizone, C.; Eltis, L.D.; Mohn, W.W. Distinct roles for two CYP226 family cytochromes P450 in abietane diterpenoid catabolism by *Burkholderia xenovorans* LB400. *J. Bacteriol.* **2007**, *190*, 1575–1583. [[CrossRef](#)] [[PubMed](#)]
70. Morgan, C.; Wyndham, R. Characterization of tdt genes for the degradation of tricyclic diterpenes by *Pseudomonas diterpeniphila* A19-6a. *Can. J. Microbiol.* **2002**, *48*, 49–59. [[CrossRef](#)] [[PubMed](#)]

71. Davydov, D.R.; Sineva, E.V.; Davydova, N.Y.; Bartlett, U.H.; Halpert, J.R. CYP261 enzymes from deep sea bacteria: A clue to conformational heterogeneity in cytochromes P450. *Biotechnol. Appl. Biochem.* **2013**, *60*, 30–40. [[CrossRef](#)] [[PubMed](#)]
72. Feyereisen, R. Insect CYP genes and P450 enzymes. In *Insect Molecular Biology and Biochemistry*; Academic Press: Cambridge, MA, USA, 2012; pp. 236–316.
73. Hamberger, B.; Bak, S. Plant P450s as versatile drivers for evolution of species-specific chemical diversity. *Philos. Trans. R. Soc. B Biol. Sci.* **2013**, *368*, 20120426. [[CrossRef](#)]
74. Ngwenya, M.L.; Chen, W.; Basson, A.K.; Shandu, J.S.; Yu, J.-H.; Nelson, D.R.; Syed, K. Blooming of unusual cytochrome P450s by tandem duplication in the pathogenic fungus *Conidiobolus coronatus*. *Int. J. Mol. Sci.* **2018**, *19*, 1711. [[CrossRef](#)]
75. Matowane, R.G.; Wieteska, L.; Bamal, H.D.; Kgosiemang, I.K.R.; Van Wyk, M.; Manume, N.A.; Abdalla, S.M.H.; Mashele, S.S.; Gront, D.; Syed, K. *In silico* analysis of cytochrome P450 monooxygenases in chronic granulomatous infectious fungus *Sporothrix schenckii*: Special focus on CYP51. *Biochim. Biophys. Acta (BBA) Proteins Proteom.* **2018**, *1866*, 166–177. [[CrossRef](#)]
76. Qhanya, L.B.; Matowane, G.; Chen, W.; Sun, Y.; Letsimo, E.M.; Parvez, M.; Yu, J.-H.; Mashele, S.S.; Syed, K. Genome-wide annotation and comparative analysis of cytochrome P450 monooxygenases in basidiomycete biotrophic plant pathogens. *PLoS ONE* **2015**, *10*, e0142100. [[CrossRef](#)]
77. Kgosiemang, I.K.R.; Syed, K.; Mashele, S.S. Comparative genomics and evolutionary analysis of cytochrome P450 monooxygenases in fungal subphylum *Saccharomycotina*. *J. Pure Appl. Microbiol.* **2014**, *8*, 12.
78. Syed, K.; Shale, K.; Pagadala, N.S.; Tuszynski, J. Systematic identification and evolutionary analysis of catalytically versatile cytochrome P450 monooxygenase families enriched in model basidiomycete fungi. *PLoS ONE* **2014**, *9*, e86683. [[CrossRef](#)]
79. Suzuki, H.; Macdonald, J.; Syed, K.; Salamov, A.; Hori, C.; Aerts, A.; Henrissat, B.; Wiebenga, A.; A Vankuyk, P.; Barry, K.; et al. Comparative genomics of the white-rot fungi, *Phanerochaete carnosae* and *P. chrysosporium*, to elucidate the genetic basis of the distinct wood types they colonize. *BMC Genom.* **2012**, *13*, 444. [[CrossRef](#)]
80. Kanehisa, M.; Sato, Y.; Furumichi, M.; Morishima, K.; Tanabe, M. New approach for understanding genome variations in KEGG. *Nucleic Acids Res.* **2019**, *47*, D590–D595. [[CrossRef](#)]
81. Lu, S.; Wang, J.; Chitsaz, F.; Derbyshire, M.K.; Geer, R.C.; Gonzales, N.R.; Gwadz, M.; I Hurwitz, D.; Marchler, G.H.; Song, J.S.; et al. CDD/SPARCLE: The conserved domain database in 2020. *Nucleic Acids Res.* **2020**, *48*, D265–D268. [[CrossRef](#)]
82. Graham, S.E.; Peterson, J.A. How similar are P450s and what can their differences Teach Us? *Arch. Biochem. Biophys.* **1999**, *369*, 24–29. [[CrossRef](#)] [[PubMed](#)]
83. Syed, K.; Mashele, S.S. Comparative analysis of P450 signature motifs EXXR and CXG in the large and diverse kingdom of fungi: Identification of evolutionarily conserved amino acid patterns characteristic of P450 Family. *PLoS ONE* **2014**, *9*, e95616. [[CrossRef](#)]
84. Katoh, K.; Kuma, K.; Toh, H.; Miyata, T. MAFFT version 5: Improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res.* **2005**, *33*, 511–518. [[CrossRef](#)]
85. Boc, A.; Diallo, A.B.; Makarenkov, V. T-REX: A web server for inferring, validating and visualizing phylogenetic trees and networks. *Nucleic Acids Res.* **2012**, *40*, W573–W579. [[CrossRef](#)]
86. Letunic, I.; Bork, P. Interactive Tree of Life (iTOL) v4: Recent updates and new developments. *Nucleic Acids Res.* **2019**, *47*, W256–W259. [[CrossRef](#)] [[PubMed](#)]
87. Saeed, A.; Sharov, V.; White, J.; Li, J.; Liang, W.; Bhagabati, N.; Braisted, J.; Klapa, M.; Currier, T.; Thiagarajan, M.; et al. TM4: A Free, Open-Source System for Microarray Data Management and Analysis. *Biotechniques* **2003**, *34*, 374–378. [[CrossRef](#)] [[PubMed](#)]
88. Blin, K.; Andreu, V.P.; Santos, E.L.C.D.L.; Del Carratore, F.; Lee, S.Y.; Medema, M.H.; Weber, T. The antiSMASH database version 2: A comprehensive resource on secondary metabolite biosynthetic gene clusters. *Nucleic Acids Res.* **2019**, *47*, D625–D630. [[CrossRef](#)] [[PubMed](#)]