FUDMA Journal of Sciences (FJS)
ISSN online: 2616-1370
ISSN print: 2645 - 2944
Vol. 7 No. 4, August, 2023, pp 246 - 264
DOI: https://doi.org/10.33003/fjs-2023-0704-1922

ELUCIDATING THE FUNCTIONAL ANNOTATION AND EVOLUTIONARY RELATIONSHIPS OF CYTOCHROME P450 GENES IN *XYLARIA SP*. FL1777 USING IN-SILICO APPROACHES

[*1]Wadzani Palnam Dauda, [2]Elkanah Glen, [3]Peter Abraham, [4]Charles Oluwaseun Adetunji, [5]Daji Morumda, [6]Shittu Emmanuel Abraham, [7]Grace Peter Wabba, [8]Israel Ogra Ogwuche

[1]Crop Science Unit, Department of Agronomy, Federal University, Gashua, Yobe State, Nigeria
[2]Department of Biochemistry, Federal University Lokoja, Kogi State, Nigeria
[3]Department of Horticulture, Federal College of Horticulture, Dadin Kowa, Gombe, Nigeria,
[4]Applied Microbiology, Biotechnology and Nanotechnology Laboratory, Department of Microbiology, Edo University Iyamho, PMB 04, Auchi, Edo State, Nigeria
[5] Department of Microbiology, Federal University Wukari, Taraba State, Nigeria
[6]Department of Agronomy, Bayero University Kano, Nigeria
[7]Department of Biochemistry, Ahmadu Bello University, Zaria
[8]UNESCO International Centre of Biotechnology Nsukka, Nigeria

*Corresponding authors' email: Wadzanidauda@gmail.com
ORCID: 0000-0002-5233-0058

ABSTRACT

The higher level of human activities has resulted in several forms of anthropogenic activities with diverse adverse effects on human and environmental sustainability. The traditional means of handling xenobiotics pollutants are no longer sustainable due to the high cost involved, complex procedures and demanding regulatory requirements. Bioremediation using fungi (mycoremediation) is now recognized as an efficient and workable biotechnological tool that effectively employ fungal enzymes via the process of absorption and mineralization to get rid of contaminants. Cytochrome P450s (CYPs) are diverse and unique gene families with varying degree of complexities in the eukaryotes. CYPs mainly utilize molecular oxygen to modify substrate conformation, thereby establishing a mechanism of action for achieving their important physiological and ecological processes. *Xylariaceae* belongs to the main and highly diversified families of filamentous Ascomycota; it plays an important role as saprotrophs of wood, soil, litter and dung. Genome-wide annotation analysis was carried out to explore the possibility of utilizing the CYPs of *Xylaria* sp. for achieving mycoremediation. The evolutionary analysis has divided the 214 *Xylaria* CYPs into fifteen (15) clades. The CYPs were categorized into forty-seven (47 clans) and eighty-six (86) families. MEME suite identified ten (10) conserved motifs. The gene structural investigation reveals high dynamic intron-exon organization. Most of the CYPs have been predicted to be localized in the endoplasmic reticulum. This study therefore calls for deeper exploration of the *Xylaria* sp and its high potential for application in bioremediation for the degradation of environmental contaminants.

Keywords: Bioremediation, Cytochrome P450, Genome, Pollutant, Xenobiotic, *Xylaria*

INTRODUCTION

Humankind is involved in various activities for survival, which consequently led to the liberation of different pollutants into the environment. Li *et al.* (2019) reported that pollutants from human activities constitute a major risk to the health of human and environmental sustainability. The agelong solution of eliminating xenobiotic pollutants includes the use of ultraviolet decomposition, incineration at high temperature, pit disposal and chemical degradation (Bhandari *et al.*, 2021). These techniques are gradually phasing out due to high-cost implications, complex procedures, burdensome regulatory requirements, inadequate provision for space and secondary pollutants arising from the processes (Bhandari *et al.*, 2021). Hence, there is an urgent need for environmentally friendly remediation techniques that could be applied for effective bioremediation of these contaminants mentioned above (Kevin *et al.*, 2019). Singh (2006) stated that fungi because of their ability to actively decompose various chemicals have been recognized as a potential workable biotechnological tool that could be applied in the bioremediation of heavily polluted environments.

Similarly, Buddolla *et al.* (2014) identified fungi as an essential organism for bioremediation due to their ability to exploit significantly minimal living conditions by producing enzymes capable of undertaking chemically difficult reactions. It has been discovered that fungi could effectively remove toxic and intractable products like waste from pharmaceuticals, polyaromatic and chlorinated hydrocarbons, pesticides and mineral oils from the soil have also been reported by Jasu *et al. (*2021) where they named cytochrome P450 monooxygenases as one of the intracellular enzymes to perform such task. Bioremediation using fungi (mycoremediation) is a method that utilizes enzymes in live fungi to clear up contaminants through mineralization or absorption (Kevin *et al.*, 2019). Similarly, shiyuki *et al.* (2013) reported using microorganisms in the activated sludge process as bioremediation techniques for industrial extract chemicals and the polychlorinated forms of dibenzo-p-dioxin and dibenzofuran (PCDD &PCDF).

Cytochrome P450s (CYPs) are diverse and unique gene families with varying degree of complexities in the eukaryotes. CYPs mainly utilize molecular oxygen to modify substrate conformation, thereby establishing a mechanism of action for achieving their important physiological, toxicological and ecological processes (Nelson *et al.*, 2013). The Cytochrome P450 enzymes (P450s) are largely disseminated across organisms and perform essential roles in the biosynthesis (of steroids and natural products), xenobiotics degradation, and metabolism drugs. P450s are generally regarded as the most adaptable natural biocatalysts

because of the wide range of substrate configuration and the kinds of reactions they catalyze (Li *et al*., 2010). Generally, ascomycetes inhabit a wider niche in soil than their basidiomycetes counterpart, yet they have not received attention for bioremediation studies when compared with basidiomycetes that have been well-studied (Li *et al.,* 2019). *Xylariaceae* belongs to the major and highly diversified families of filamentous Ascomycota. According to U'Ren *et al.,* (2016), *Xylariaceae* are active saprotrophs of litter, dung, wood, soil, and plant pathogens in a natural and agricultural system, as other facultative fungal organisims (Dauda et al., 2018; Palnam et al., 2019; Zarafi and Dauda, 2019)

Xylariaceous are progressively been recognized as a chief source of new products of metabolism for utilization in biofuel, environment, agriculture, medicine and industrial applications (Wu *et al*., 2017). Li *et al.,* (2019) reported that *Xylaria* sp. *BNL1* can degrade carbaryl in contaminated soil with a degradation rate of 59.0% in fifteen (15) days; this implies that *Xylaria* sp. *BNL1* can survive various attacks from indigenous microorganisms. The role played by P450s in economically important fungi such as *Aspergillus* spp., (Kelly et al., 2009; Dauda et al., 2022a *Alternaria* spp., (Dauda et al., 2022b) Candida tropicalis (Dauda et al., 2022c) , *Trichoderma* spp., (Chadha et al., 2018) have been well elucidated. Considering the diverse potential biotechnological applications of *Xylaria* and the impact of cytochrome P450 in the biological, physiological and biochemical activities of fungi this study intends to perform an evolutionary relationship and genome-wide analysis of cytochrome P50 genes in *Xylaria* sp. FL1777 to open room for commercial exploitation of these proteins, especially in bioremediation.

MATERIALS AND METHODS

Sequence Retrieval and Alignment

Protein, genomic and coding sequences of cytochrome P450 of *Xylaria* sp FL1777 were downloaded from the Joint Genome Institute (JGI) fungal genome database-MycoCosm (mycocosm.jgi.doe.gov/pages/search-for-genes.jsf). The protein sequences were aligned using the MUSCLE algorithm with all parameters set at default (gap open: -2.9, gap extend:0.00, hydrophobicity multiplier:1.20, maximum iterations:16, clustering method: UPGMA and min diag lenth:24).

Structural feature analysis of CYP protein sequences:

The conserved domain of cytochrome P450 in *Xylaria* sp. FL1777 were analyzed using the conserved domain database (CDD). The proteins sequenced were analysed for the presence of the CYP family signature domains viz; heme-binding and oxygen-binding motifs. The sequences used were only those with the two CYP signature domains (Matowane *et al*., 2018).

Evolutionary relationships of taxa

The evolutionary relationship was analysed using 214 amino acid sequences. The pairwise deletion was used to clear off all ambiguous positions on each sequence. The Phylogenetic tree was conducted in MEGA X (Kumar *et al.,* 2018) using the Neighbour-Joining method (Olszewska-Tomczyk *et al.,* 2016) and as described by Dauda et al., (2021). The optimal tree with the sum of branch length = 110.19813185 is shown. Poisson correction method was used to compute the evolutionary distances (Tomczyk *et al.,* 2016). The final dataset comprises a total of 2078 positions.

Identification of clans, families and putative functions

The putative CYP names for all P450 genes in *Xylaria sp* FL1777 were assigned by the logic in the FCDP pipeline (http://p450.riceblast.snu.ac.kr) following the nomenclature format as proposed by Nelson (2006), two CYPs with more than 40% sequence similarity belong to the same family. Therefore, each *Xylaria* CYPs was blasted against all known fungal cytochrome P450 available at "Fungal cytochrome P450 database" where blast result with a best hit (greater than 40% sequence similarity) to the query sequence is assigned to that family. Clans were identified by comparing families obtained against clans and families in the fungal cytochrome P450 database.

Identification of Motif and Analysis of Gene Structure:

The conserved motifs of cytochrome P450 gene of *Xylaria* sp. FL1777 were identified by an online server, Multiple Expectation Maximization for Motif Elicitation (MEME) Suite (http://meme-suite.org/tools/meme) using the genomic sequence (Bailey *et al*., 2009). A set of 214 protein sequences between 95 and 1153 in length with an average length of 491.6 have been submitted. The number of motif counts was set at 10, the minimum width of the motif was set at 6 amino acids, while the maximum was 100 amino acids. Similarly, structures of both intron and exon of cytochrome P450 gene in *Xylaria* sp. FL1777 were analysed using an online server called Gene Structure Display Server (GSDS 2.0) (http://gsds.gao-lab.org/) (Bo *et al*.,2015), the positions and numbers of both introns and exons were graphically displayed by the server after loading the coding and genomic FASTA sequences of *Xylaria* sp. FL1777.

Sub-cellular localization analysis:

The localization of the *Xylaria* CYPs was predicted using an online web server for predicting the subcellular localization of eukaryotic proteins, including those with multiple sites in a different organism known as Euk-mPLoc 2.0 (Cheng *et al.,* 2018), which is accessible at http://www.csbio.sjtu.edu.cn/bioinf/Cell-PLoc/. The protein FASTA sequence of the organism was used.

Identification of *Xylaria* CYPs Involved in secondary metabolism-related gene clusters

Secondary metabolism-related gene clusters of *Xylaria sp.* FL1777 were identified from the joint genome institute mycoCosm using the annotations on the homepage and a search for all cluster types, namely; dimethylallyltryptophan (DMAT), PKS/NRPS (HYBRID), Non-ribosomal Peptide Synthase (NRPS), NRPS-like, polyketide synthase (PKS), PKS-like and terpene cyclase (TC).

RESULTS AND DISCUSSION

Phylogenetic Analysis:

The result obtained from the phylogenetic analysis shown in figure 1 revealed that 214 protein sequences were divided into thirteen (13) clades. About half of these proteins (95) were clustered in four clades (I, VIII, IX, and XI), having 24, 24, 18 and 31 proteins, respectively. In contrast, clades with the least cluster of proteins were IV, VI, and X with 3, 2 and 3 proteins respectively. Clades II, V and XIII were having relatively equal distribution of proteins consisting of 14, 15, and 14 proteins respectively.

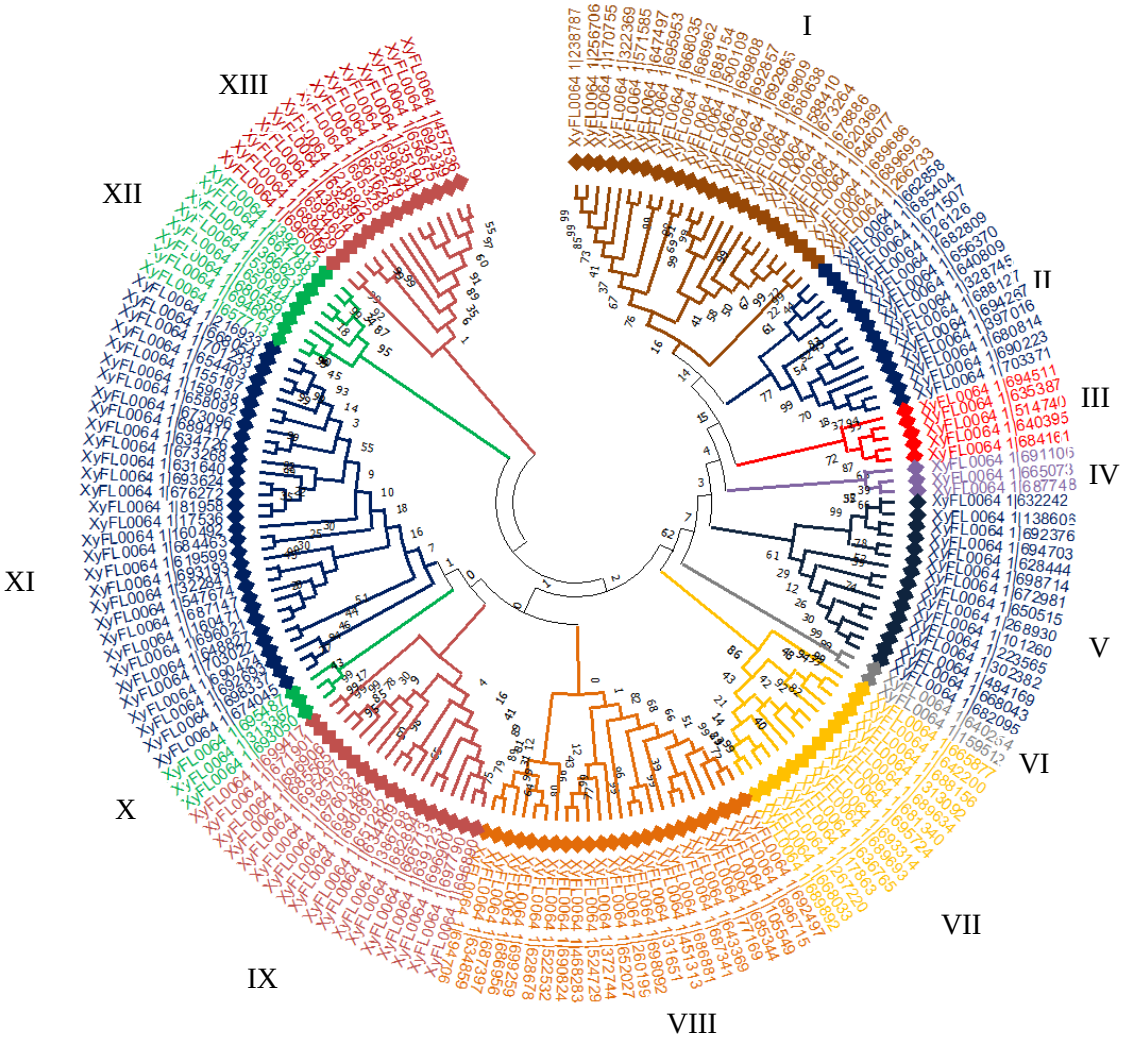


Figure 1: Evolutionary relationship of Cytochrome P450 proteins in *Xylaria* Sp. FL1777. Neighbor-Joining method was used to infer the evolutionary history using MEGA-X software. Thirteen (13) clades have been labelled (I-XIII) and separated by colour demarcation.

Identification of Clan and Family

Moreover, the 214 genes obtained from Joint Genome Institute Mycocosm used in this study were putatively distributed into 47 clans and 86 families, as shown in Table 1. Fifty-eight (58) CYPs have no matches in the fungal cytochrome P50 database. CYP531 clan has the highest number of families (8), consisting of families CYP531, CYP532, CYP536, CYP629, CYP631, CYP675, CYP5080 and CYP5104. The entire clan has 11 proteins that are involved in xenobiotic metabolism. Clan CYP58 is the second-highest in family size consisting of six (6) families which are; CYP58, CYP682, CYP5104, CYP5112, CYP5094 and CYP551. The clan has the highest number of proteins (12) which were participate in xenobiotic and secondary metabolism. Clan CYP54 has four families with 11 proteins that are involved in secondary metabolism. Nineteen clans are orphans with only a single protein each. Twenty-one (21) clans have no corresponding putative function in the fungal cytochrome P450 database. Generally, five (5) clans are involved in primary metabolism, seven (7) in secondary metabolism while thirteen (13) in xenobiotic metabolism.

Spread of Conserved Motifs in *Xylaria* sp. FL1777

Furthermore, the spreading of the ten conserved motifs across the 214 cytochrome P450 genes was established during this study, as shown in Figure 2. The study revealed that thirty-six (36) genes have all the ten conserved motifs. The result also revealed that Thirteen (13) CYPs have only one conserved motif each. Motifs 1(FXXGXXXCXG), motif 2 (EXXR), motif 3 (PERW), motif 5 (LXXPXXXLXE) and motif 7 (HXGXRXP) appeared the most, occurring at 154, 154, 126, 128 and 123 sites, respectively. On the other hand, motif 10 (HXXXRXFSXXR) is the widest, while motif 3 is the shortest. The other motifs (2,4.5,6,7 and 8) have relatively equal width. Similarly, motif 6 was the least conserved as it appeared at 48 sites only.

Exon-intron Analysis

The result of exon-intron structures of cytochrome P450 gene in *Xylaria sp*. FL1777 was shown in figure 3. All the genes have a minimum of one and a maximum of nine introns except for twelve (12) genes (XYFL 801046, 789729,787010, 783228, 781286, 761684, 643781, 164544, 350436, 324633, 437355, 783920 which have none. XYFL437355 exists as the longest single exon with about 2,500bp. XYFL799538 has the highest number of introns (9), while twelve others have eight introns each in their sequences. All the genes have no untranslated regions (UTR).

Sub-cellular localization of CYPs of *Xylaria* sp FL1777

The *Xylaria* CYPs were established in this study (Table 2) to be majorly localised in the endoplasmic reticulum (160 out of the 214 CYPs), representing 74.77% 49 CYPs representing 22.9% were found to be localized in the cytoplasm. Three genes were found each in the plasma membrane, chloroplast, peroxisome, nucleus and microsome. The extracellular compartments and mitochondrion were each shown to contain 10 CYPs. Twenty-one (21) CYPs are localized in at least two organelles with XYFL 763710, 413182 and 382656 occurring in 6, 5 and 4 locations respectively. The three CYPs mentioned above were all present in mitochondria, cytoplasm, and plasma membrane.

Secondary metabolism-related gene clusters

The annotations on the mycoCosm homepage of joint genome institute for *Xylaria sp*. FL1777 revealed 64 genes of the 214 (which represent 29.9%) cytochrome P450 are linked to secondary metabolism-related gene clusters (figure 4), specifically HYBRID (3), NRPS (6), NRPS-like (12), PKS (30), PKS-like (3) and TC (10).

Discussion

The phylogenetic analysis performed during this study revealed an unequal distribution of cytochrome P450 cluster sizes in *Xylaria* sp. FL1777 and is in line with Chadha *et al.* (2018), who stated that there are high expansions and contractions of certain CYP families in the course of evolution. Expansion of cytochrome P450 across different clades in *Xylaria* sp FL1777 could be very instrumental in their survival in respective habitats. The observed numerous branches in the tree imply their highly evolved divergence (Chen *et al*., 2014). The high evolutionary diversity observed in the *Xylaria* CYPs may not only be due to significant sequence variation but also incredible functional diversification as earlier reported in a similar study by Sezutsu *et al.* (2013). Most of the cytochrome P450 genes in *Xylaria* sp FL1777 have demonstrated a close relationship in phylogeny, hence inferring a common ancestral lineage which agrees with the earlier report of Chen *et al.* (2014) on fungal cytochrome P450. The observed variation in *Xylaria* cytochrome P450 might be linked to gene duplication; more so, the resemblance in protein sequence identity of cytochromes P450 in *Xylaria* as seen in clustering of about half of the genes in just four clades is an indication of recent duplication in that specie. This also agrees with the findings of Chen *et al.*, (2014).

In *Xylaria* sp. FL1777, clan CYP52 comprises four families (CYP 52, CYP538, CYP539 and CYP 655) with five (5) proteins. Werner *et al.,* (2017) reported that this clan is known to catalyze alpha-omega-dicarboxylic acids from alkanes and fatty acids. CYP 51, CYP61 and CYP505 proteins observed in this study are linked to primary metabolism in Xylaria and consist of only five proteins. CYP61 has been reported by Venegas *et al.,* (2020) to be responsible for the coding of sterol 22 desaturase, which plays a significant role in the advanced phase of the ergosterol pathway in metabolizing Ergosta-5,7,24(28)-trienol to Ergosta5,7,22,24(28)-tetraenol by introducing a C-22(23) double bond in the sterol side chain. Clan CYP51 has been reported to be involved in sterol biosynthesis in basidiomycetes and ascomycetes and is known as housekeeping CYP. This has made them target most antifungal control of fungal human diseases (Shin *et al.,* 2018). Seven clans out of the forty-seven clans discovered in this study in *Xylaria* (CYP54, CYP65, CYP526, CYP547, CYP550, CYP559 and CYP574) have been linked to secondary metabolism. CYP65 has been reported to catalyze the epoxidation reaction during the synthesis of trichothecenes biosynthesis in *F. graninearumn* (Gao *et al.,* 2020) and radicicol (Chedha *et al.,* 2018). Similarly, thirteen clans (CYP613, CYP548, CYP537, CYP533, CYP531, CYP530, CYP528, CYP507, CYP504, CYP62, CYP59, CYP53 and CYP52) comprising of 56 proteins in *Xylaria* have been linked to Xenobiotic metabolism. This finding has agreed with an earlier study by Chedha *et al.,* (2018) where they reported the involvement of CYP507, CYP530, CYP531, CYP532 and CYP548 to be involved in Xenobiotic metabolism. The Copiousness of these proteins in *Xylaria* may be responsible for the exceptional ability of this fungus to degrade a diverse range of xenobiotics, including fungicides.

The study established Motif 2 (EXXR) in *Xylaria sp.* which has arginine and glutamic acid residues to be highly conserved. This signature motif was earlier reported by Deng *et al.* (2007) to be actively involved in stabilizing the main structure of CYP proteins. Motif 4 (AGXDTT) was

reported to constitute the domain for binding and activation of oxygen (Chen *et al.,* 2014). This shows that these motifs were widely distributed and have the strongest conservatism in the gene sequence of *Xylaria sp.* The evaluation of the conserved motifs is one way to predict the functions of *Xylaria* CYP genes (Jiu *et al.,* 2020). Many signature motifs have been reported to be conserved in the CYP protein of fungi (Chadha *et al.,* 2018). Generally, the sequence similarities are minimal looking at the characteristics differences in the motif, however, the motifs are highly conserved across the cytochrome P450 genes. This finding agrees with the earlier submission of Yu *et al.* (2014), who reported highly conserved characteristics motif of fungal cytochrome P450 with very low overall sequence resemblance. The FXXGXXXCXG conserved motif (also known as CXG) is reported to have a domain that binds heme which contains a consistent cysteine residue that binds to the iron in the heme. Also, it was reported by Deng *et al.* (2007) and Moktali *et al.* (2012) that the few conserved domains in fungal cytochrome P450 are responsible for their major characteristics, which tallies with the conservation of enzymatic functions and the tertiary structure.

The result of exon-intron structures of cytochrome P450 gene in *Xylaria sp. during this study* revealed twelve (12) intron-less genes (mono-exonic) hence can easily be translated. One gene has 9 introns. The longest single exon has about 2,500bp. All the genes have no untranslated regions (UTR). This observed variation in the length of both introns and exons of *Xylaria CYP* also agrees with the findings of Raghavendra *et al.* (2012) where they reported about the highly dynamic nature of the intron-exon structure of the cytochrome P450 superfamily.

The localization of over 70% of *Xylaria CYPs* in the endoplasmic reticulum and about 20% in the cytoplasm as shown in this study also validates the claim by Kelly *et al.* (2009) where they stated that cytochrome P450 of Eukaryotes generally attached themselves to the endoplasmic reticulum via its cytoplasmic surface. The functions of cytochrome P450 in *Xylaria sp.* depend on these enzymes' ability to relate with their oxidation-reduction partners, NADPH-cytochrome P450 reductase and cytochrome b5 in the endoplasmic reticulum (Park *et al.,* 2014). These interactions with redox partners in the endoplasmic reticulum might be responsible for the intercellular catalysis of xenobiotics and other environmental pollutants.

Fungi are among the most prolific producers of secondary metabolites, which are both beneficial (as antibiotics and pharmaceuticals) and harmful (toxic and carcinogenic) to mankind in particular and the universe in general (Keller *et al.,* 2005). The 29.9% of cytochrome P450 that are involved in secondary metabolism-related gene clusters in *Xylaria* sp. FL1777 shows the abundance of secondary metabolites in the organism. Therefore, there is the need for systematic studies that will lead to a discovery of a new pathway, intermediate or metabolite, that can be harnessed for an improved bioremediation application.

CONCLUSION

This study revealed the distribution of 214 protein sequences into fifteen (15) clades, with more than half of them (125) clustering in just four clades. The result obtained implies a close relationship in phylogeny, hence inferring a common ancestral lineage. Moreover, the 214 genes were putatively distributed into 47 clans and 86 families. The majority of these CYPs have been implicated in xenobiotic metabolism. Furthermore, ten conserved motifs have been predicted in the study. The signature motifs; EXXR, AGXDTT and FXXGXXXCXG have been linked with stabilizing CYPs structures, binding and activation of oxygen and heme-binding domains, respectively. The study also showed consistency in organizations of exon-intron structures. More so, 97.67% of *Xylaria* CYPs are localized in the endoplasmic reticulum and cytoplasm. *Xylaria* CYPs have not been characterized nor classified before now; this work has laid a foundation for further characterization and systematic studies that will fully annotate the functions of these genes. Therefore, there is a need to identify a gene or set of genes that can be effectively harnessed for application in bioremediation.

Declaration of Interest:

The authors hereby declare that there is no conflict of interest whatsoever.

REFERENCES

Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., *et al*. (2009). MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.* 37(Web Server): 202-8.

Bhandari, S., Poudel, D. K., Marahatha, R., Dawadi, S., Khadayat, K., Phuyal, S., ... & Parajuli, N. (2021). "Microbial Enzymes Used in Bioremediation", *Journal of Chemistry.* https://doi.org/10.1155/2021/8849512.

Bo H., Jinpu J., An-Yuan G., He Z, Jingchu L. and Ge G. (2015). GSDS 2.0: an upgraded gene features visualization server. *Bioinformatics*, 31(8):1296- 1297.

Budddolla, V., Bandi R., Avilala J., Arthala P.K. and Golla N. (2014). Fungal Laccases and the Application in Bioremediation. *Enzyme Research* 201:1-21

Chadha S., Mahetre S.T., Bansal R., Kuo A., Aerts A., Grigoriev I.V., Druzhinina S.I., Mukherjee PK. (2018). Genome-wide analysis of cytochrome P450 of *Trichoderma* spp.: annotation and evolutionary relationships. *Fungal Biol Biotechnol* 5(12): 11-12

Chen W., Lee M., Jefcoate C., Kim S, Chen F., Yu J.H., (2014). Fungal Cytochrome P450 *Monooxygenases*: Their Distribution, Structure, Functions, Family Expansion, and Evolutionary Origin. *Genome Biol. Evol*. 6(7):1620-1634

Cheng, X., Xiao, X., & Chou, K. C. (2018). pLoc-mEuk: Predict subcellular localization of multi-label eukaryotic proteins by extracting the key GO information into general PseAAC. *Genomics*, *110* (1), 50-58.

Dauda WP, Peter GW, Abraham P, Adetunji CO, Glen E, Daji M, Ogra IO, Shittu EA., Azameti MK, Ghazanfar S, Osemwegie OO, Olaniyan OT, and Anyakudo MMC (2022c). Bioinformatics Based Structural Analysis of Cytochrome P450 genes in Candida tropicalis. Nigerian Journal of Parasitology 43(2): 379-390.

Dauda WP, Abraham P, Fasogbon IV, Adetunji CO, Banwo OO, Kashina BD, Alegbejo MD (2021). Cassava mosaic virus in Africa: Functional analysis of virus coat proteins based on evolutionary processes and protein structure. Gene Reports, p.101239.

Dauda WP, Abraham P, Glen E, Adetunji CO, Ghazanfar S, Ali S, Al-Zahrani M, Azameti MK, Alao SEL, Zarafi AB, et

al. (2022a). Robust Profiling of Cytochrome P450s (P450ome) in Notable Aspergillus spp. Life.12, 451.https://doi.org/10.3390/life12030451.

Dauda WP, Morumda D, Abraham P, Adetunji CO, Ghazanfar S, Glen E, Abraham SE, Peter GW, Ogra IO, Ifeanyi UJ, Musa H, Azameti MK, Paray BA, Gulnaz A (2022b). Genome-Wide Analysis of Cytochrome P450s of Alternaria Species: Evolutionary Origin, Family Expansion and Putative Functions. J. Fungi (Basel) 8 (4): 324. doi: 10.3390/jof8040324.

Dauda, W.P., Alao, S.E.L., Zarafi, A.B. and Alabi, O., 2018. First Report of die-back disease of onion (Allium cepa L.) induced by Fusarium equiseti (Mart) Sacc in Nigeria. Inter J Plant Soil Sci, 21, pp.1-8.

Deng J.X., Carbone I., Dean R.A. (2007). The Evolutionary History of Cytochrome P450 Genes in Four Filamentous Ascomycetes. *BMC Evol Biol.* 7:30

Felsenstein J. (1985). Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* 39:783-791.

Gao, S., Zeng, R., Xu, L., Song, Z., Gao, P., & Dai, F. (2020). Genome sequence and spore germination-associated transcriptome analysis of Corynespora cassiicola from cucumber. *BMC microbiology*, 20 (1), 1-20.

Jasu, A., Lahiri, D., Nag, M., & Ray, R. R. (2021). Fungi in bioremediation of soil organic pollutants. In *Fungi Bio-Prospects in Sustainable Agriculture, Environment and Nano technology* (pp. 381-405). Academic Press.

Jiu, S., Xu, Y., Wang, J., Wang, L., Liu, X., Sun, W., ... & Zhang, C. (2020). The Cytochrome P450 Monooxygenase Inventory of Grapevine (Vitis vinifera L.): Genome-Wide Identification, Evolutionary Characterization and Expression Analysis. *Frontiers in genetics*, *11*, 44.

Keller, N. P., Turner, G., & Bennett, J. W. (2005). Fungal secondary metabolism—from biochemistry to genomics. *Nature Reviews Microbiology*, *3*(12), 937-947.

Kelly D.E., Krasevec N., Mullins J., Nelson D.R. (2009). The Cypone (Cytochrome P450 complement) of *Aspergillus nidulans*. *Fungal Genetics and Biology* 46:53-61.

Kevin, D. H., Jianchu X., Sylvie, R., Rajesh, J., Saisamorn, L., Allen G. T. N., Pranami, D.A.… & Marc, S. (2019). The amazing potential of fungi: 50 ways we can exploit fungi industrially. *Fungal Diversity* 97:1–136

Kumar S., Stecher G., Li M., Knyaz C., and Tamura K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Molecular Biology and Evolution* 35:1547-1549.

Kuo-Chen C. and Hong-Bin S. (2010). Cell-PLoc 2.0: an improved package of web-servers for predicting subcellular localization of proteins in various organisms, *Natural Science*, 2: 1090-1103.

Li Z., Jlang Y., Guengerich F.P., Ma L., Shengying L. and Zhang W. (2020). Engineering cytochrome P450 enzymes systems for biomedical and biotechnological applications.

Li, F., Di, L., Liu, Y., Xiao, Q., Zhang, X., Ma, F., & Yu, H. (2019). Carbaryl biodegradation by Xylaria sp. BNL1 and its metabolic pathway. *Ecotoxicology and environmental safety*, *167*, 331-337.

Matowane, R. G., Wieteska, L., Bamal, H. D., Kgosiemang, I. K. R., Van Wyk, M., Manume, N. A., ... & Syed, K. (2018). In silico analysis of cytochrome P450 monooxygenases in chronic granulomatous infectious fungus Sporothrix schenckii: Special focus on CYP51. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, *1866*(1), 166-177.

Mockali V., Park J., Fedorova N.D., Park B., Choi J., Lee Y.H., and Kang S., (2012). Systematic and searchable classification of cytochrome P450 proteins encoded by fungal and oomycete genomes. *BMC Genomics* 13,525

Nelson D.R. (2006). Cytochrome P450 nomenclature, 2004. Methods in Molecular Biology (Clifton, N.J.) 320:1-10

Nelson, D. R., Goldstone, J. V., & Stegeman, J. J. (2013). The cytochrome P450 genesis locus: the origin and evolution of animal cytochrome 450s. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 368(1612), 20120474.

Olszewska-Tomczyk, M., Świętoń, E., Minta, Z., & Śmietanka, K. (2016). Occurrence and phylogenetic studies of chicken anaemia virus from polish broiler flocks. *Avian diseases*, *60*(1), 70-74.

Palnam, W.D., Laykay, A.S.E. and Hanatu, M., 2019. Effect of Plant Extracts on Sporulation of Aspergillus niger and Penicillium chrysogenum from Sunflower (Helianthus annuus L.) Seeds. Journal of Food Stability, 2(1), pp.43-48.

Park, J. W., Reed, J. R., Brignac-Huber, L. M., & Backes, W. L. (2014). Cytochrome P450 system proteins reside in different regions of the endoplasmic reticulum. *Biochemical Journal*, 464(2), 241-249.

Raghavendra, K.N., Reddy, B.P., Prasad, G.B.K.S. (2012) Anopheles gambiae (Diptera: Culicidae) Cytochrome P450 (P450) Supergene Family: Phylogenetic Analyses and Exon-Intron Organization. *Entomol Ornithol Herpeto* 1(1):1-10

Saitou N. and Nei M. (1987). The neighbour-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* 4:406-425.

Sezutsu H., Le Goff G., Feyereisen R. (2013). Origins of P450 diversity. Philos Trans R Soc B. 368:20120428

Shin, J., Kim, J. E., Lee, Y. W., & Son, H. (2018). Fungal cytochrome P450s and the P450 complement (CYPome) of *Fusarium graminearum*. *Toxins* 10 (3), 112.

Toshiyuki, S., Keiko, Y. Shinichi, I. (2013). Possibility of Application of Cytochrome P450 to bioremediation of dioxins. Wiley Online Library DOI: 1 0. 1 0 0 2 / b a b. 1 0 6 7

Venegas, M., Barahona, S., González, A. M., Sepúlveda, D., Zúñiga, G. E., Baeza, M., ... & Alcaíno, J. (2020). Phenotypic

analysis of mutants of ergosterol biosynthesis genes (ERG3 and ERG4) in the red yeast Xanthophyllomyces dendrorhous. *Frontiers in Microbiology*, *11*, 1312.

Werner, N., Dreyer, M., Wagner, W., Papon, N., Rupp, S., & Zibek, S. (2017). Candida guilliermondii as a potential biocatalyst for the production of long-chain α, ω-dicarboxylic acids. *Biotechnology letters*, *39*(3), 429-438.

Wu, W., Davis R.W., Tran-Gyamfi, M.B., Kuo A., LaButti K., Mihalcheva S., Hundley H., Chovatia M., Lindquist E., Berry K., Grigoriev I.V., Henrissat B., Gladden J.M. (2017). Characterization of four endophytic fungi as potential consolidated bioprocessing hosts for conversion of lignocellulose into advanced biofuels. *Bioenergy Biofuels* 101: 2603-2618.

.

Yu, J.H., Wanping C., Mi-Kyung L., Colin J., Sun-Chang K and Fusheng C. (2014). Fungal Cytochrome P450 Monooxygenases: their Distribution, Structure, Functions Family Expansion, and Evolutionary Origin. *Genome of Biology Evolution* 6(7): 1620-1634

Zarafi, A.B. and Dauda, W.P., 2019. Exploring the importance of fungi in agricultural biotechnology. Int J Agric Sci Vet Med, 7, pp.1-12.

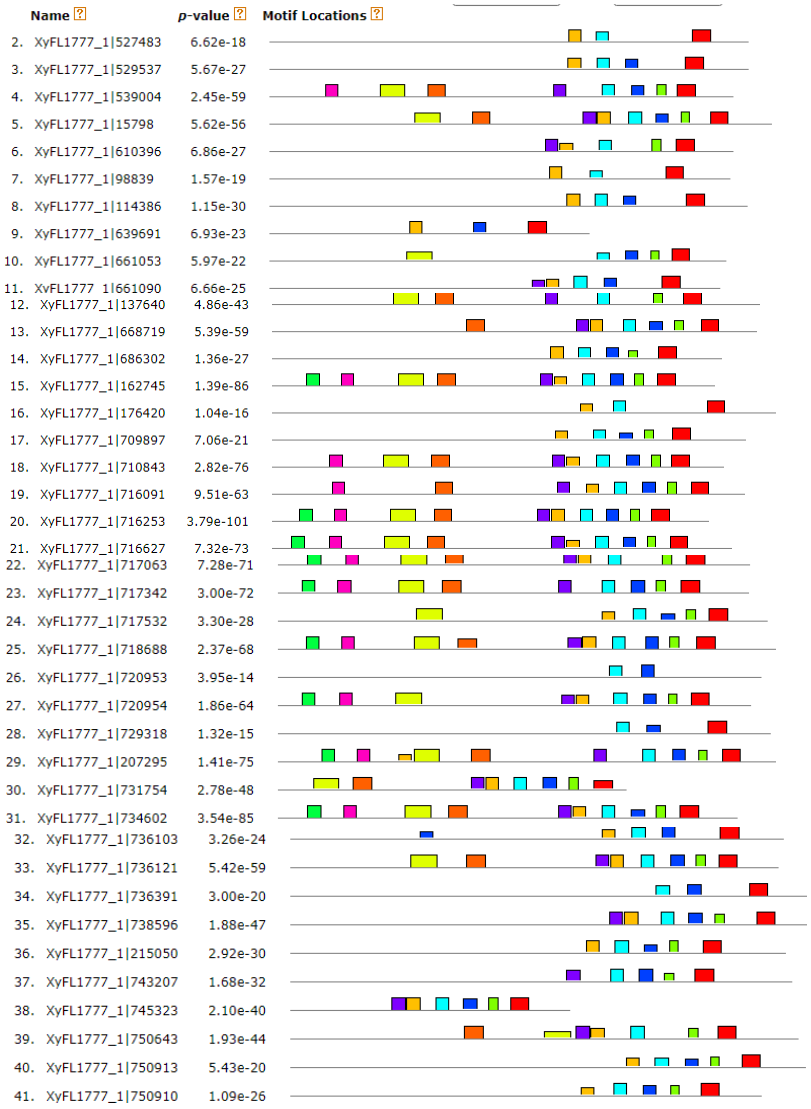
Zuckerkandl E. and Pauling L. (1965). Evolutionary divergence and convergence in proteins. Edited in *Evolving Genes and Proteins* by V. Bryson and H.J. Vogel, pp. 97-166. Academic Press, New York

Table 1: Identification of clans, families and putative functions of cytochrome P450 in *Xylaria* sp. FL1777

S/no.	Clan	Families	Number of entries	Putative functions
1	CYP51	CYP51	1	Primary metabolism
2	CYP52	CYP52, CYP538, CYP539, CYP655	5	Xenobiotic metabolism
3	CYP53	CYP53	2	Xenobiotic metabolism
4	CYP54	CYP602, CYP649, CYP503, CYP560	11	Secondary metabolism
5	CYP 58	CYP58, CYP682, CYP5104, CYP5112, CYP5094, CYP551	12	Secondary/Xenobiotic Metabolism
6	CYP 59	CYP59, CYP586, CYP587	7	Xenobiotic metabolism
7	CYP61	CYP61	1	Primary metabolism
8	CYP62	CYP84	2	Xenobiotic metabolism
9	CYP65	CYP65, CYP563, CYP567,	7	Secondary metabolism
10	CYP68	CYP68, CYP596	9	-
11	CYP504	CYP504	1	Xenobiotic metabolism
12	CYP505	CYP505, CYP541	3	Primary metabolism
13	CYP507	CYP527, CYP535, CYP570	6	Xenobiotic metabolism
14	CYP526	CYP526, CYP591, CYP644	5	Secondary metabolism
15	CYP528	CYP528	1	Xenobiotic metabolism
16	CYP529	CYP529, CYP543, CYP545	4	-
17	CYP530	CYP530, CYP5093	4	Xenobiotic metabolism
18	CYP531	CYP531, CYP532, CYP536, CYP629, CYP631, CYP675, CYP5080, CYP5104	11	Xenobiotic metabolism
19	CYP533	CYP620, CYP621	5	Xenobiotic metabolism
20	CYP 537	CYP537	2	Xenobiotic metabolism
21	CYP540	CYP540	2	Primary metabolism
22	CYP544	CYP544	1	-
23	CYP546	CYP5053	1	-
24	CYP547	CYP547, CYP617, CYP618,	5	Secondary metabolism
25	CYP548	CYP548	9	Xenobiotic metabolism
26	CYP549	CYP549	1	-
27	CYP550	CYP611, CYP634, CYP636, CYP660	5	Secondary metabolism
28	CYP559	CYP559, CYP623	3	Secondary metabolism
29	CYP572	CYP5109	2	-
30	CYP574	CYP5076	2	Secondary metabolism
31	CYP578	CYP578	1	-
32	CYP589	CYP5075 CYP614	6	-
33	CYP603	CYP603	1	-

34	CYP605	CYP605	1	-
35	CYP607	CYP607	2	-
36	CYP608	CYP608	1	-
37	CYP609	CYP609	2	-
38	CYP613	CYP613	1	Xenobiotic metabolism
39	CYP627	CYP628	1	-
40	CYP630	CYP 630	2	Primary metabolism
41	CYP639	CYP5100	2	-
42	CYP648	CYP648	1	-
43	CYP653	CYP654	1	-
44	CYP678	CYP679	1	-
45	CYP5014	CYP5014	1	-
46	CYP5084	CYP5125	1	-
47	CYP5042	CYP5042	1	-
	Total		156	
	No. of CYPs without matches in FCPD		58	

CYP= cytochrome P450, FCPD= fungal cytochrome P450 database



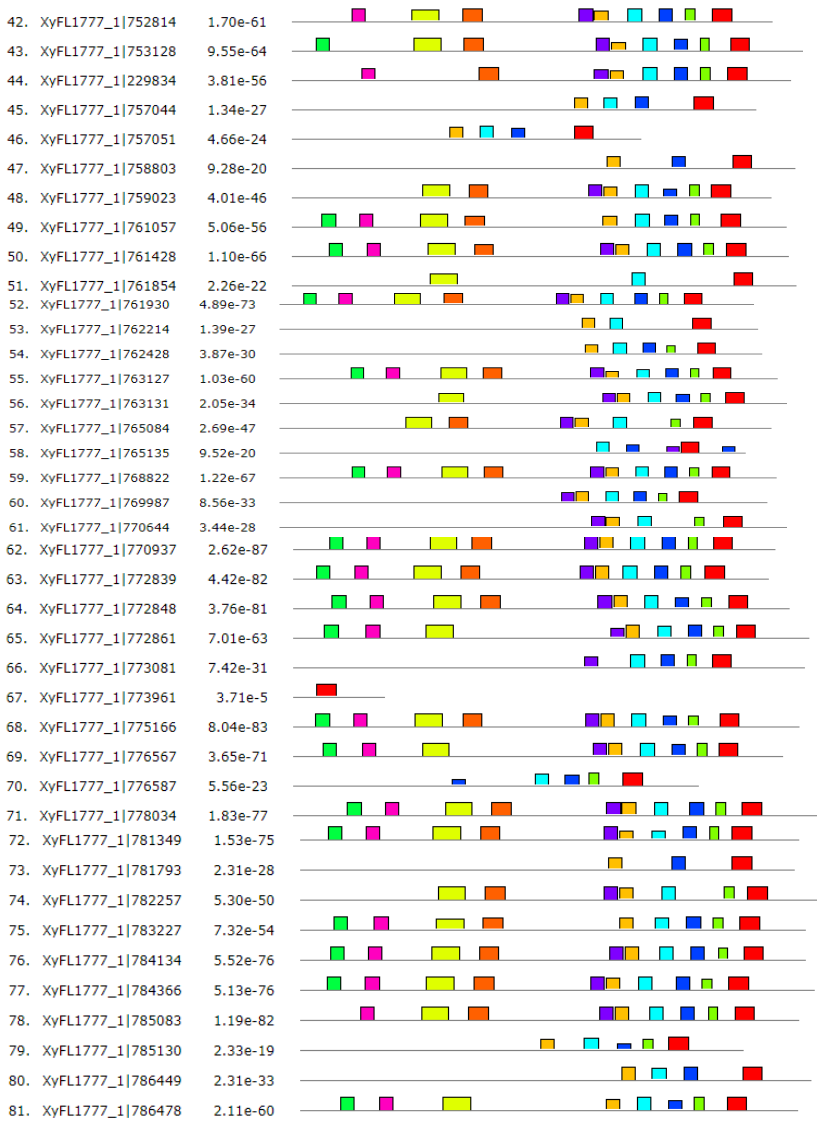


Figure 2a: predicted motifs of cytochrome P450 genes in *Xylaria* sp. FL1777

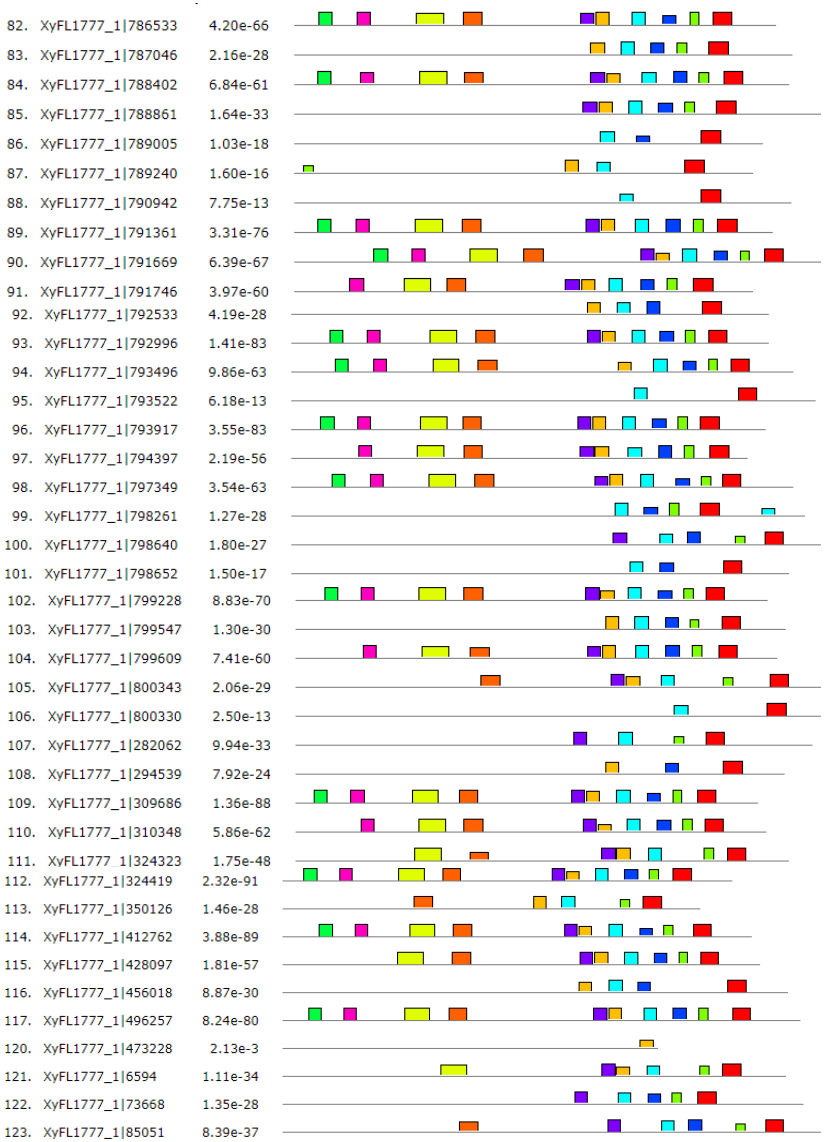
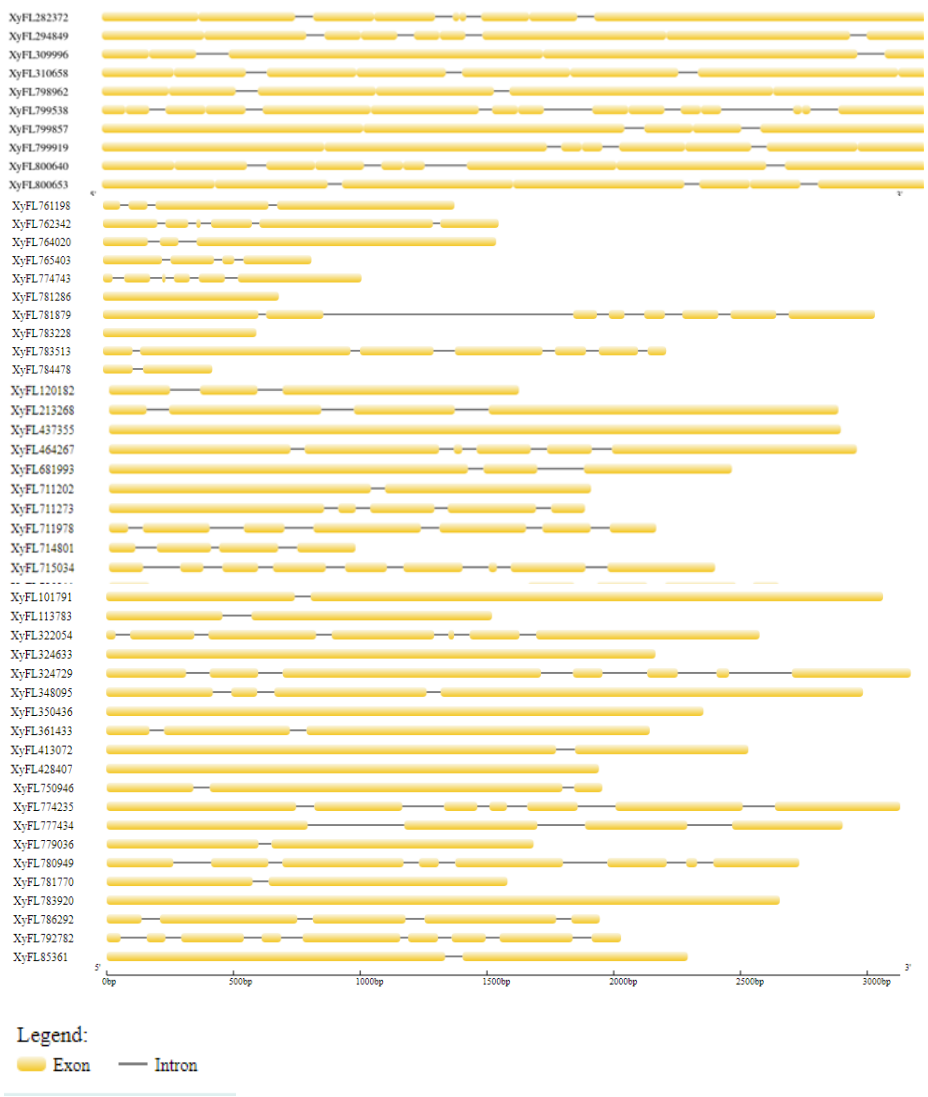


Figure 2a: predicted motifs of cytochrome P450 genes in *Xylaria* sp.

Figure 2b: Sequence Logo of the ten Motifs predicted in the cytochrome P450 of *Xylaria* sp. The type and position of amino acids are shown on the x-axis. The y-axis shows the amino acid stacks. The total height of the amino acid stacks specifies the sequence conservation at a given position, while the individual height of each symbol within a stack specifies the relative frequency of a nucleotide base at that position.



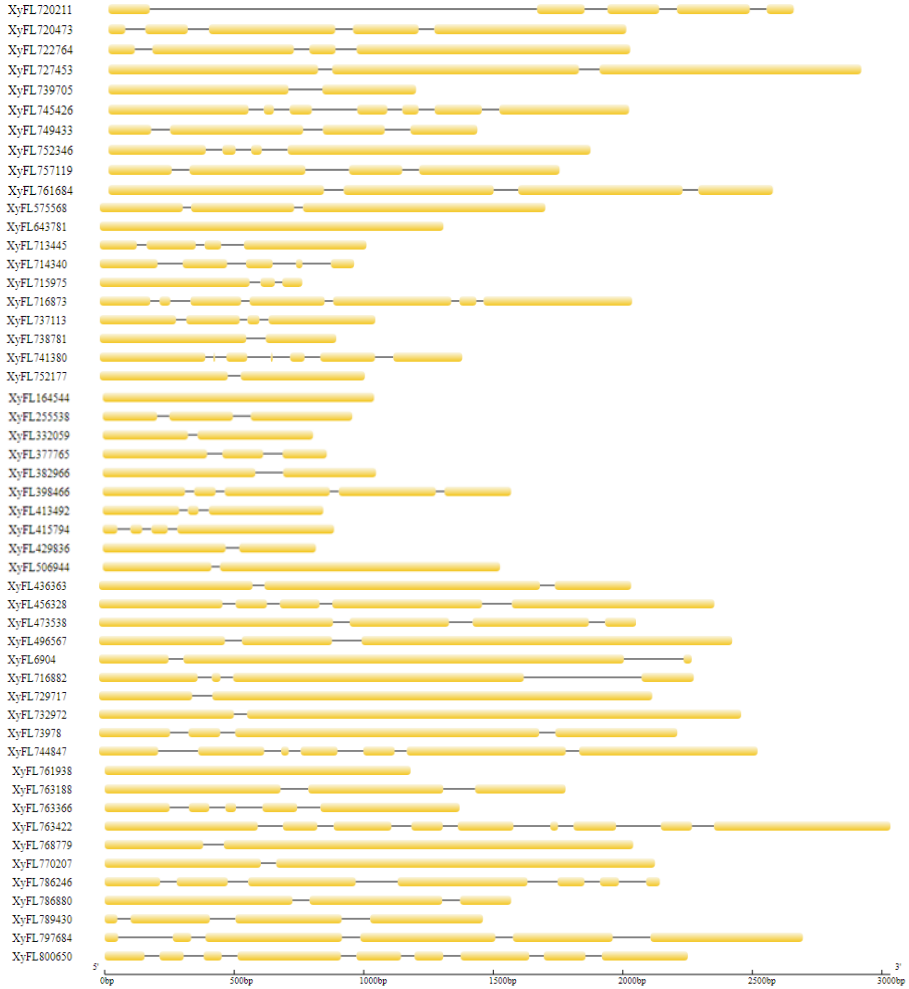


Figure 3**:** exon-intron structures of cytochrome P450 genes in *Xylaria* sp. using GSDS. The scale indicates the relative length and position of the exons and introns

Figure 3 (cont): exon-intron structures of cytochrome P450 genes in *Xylaria sp.* using GSDS. The scale indicates the relative length and position of the exons and introns.

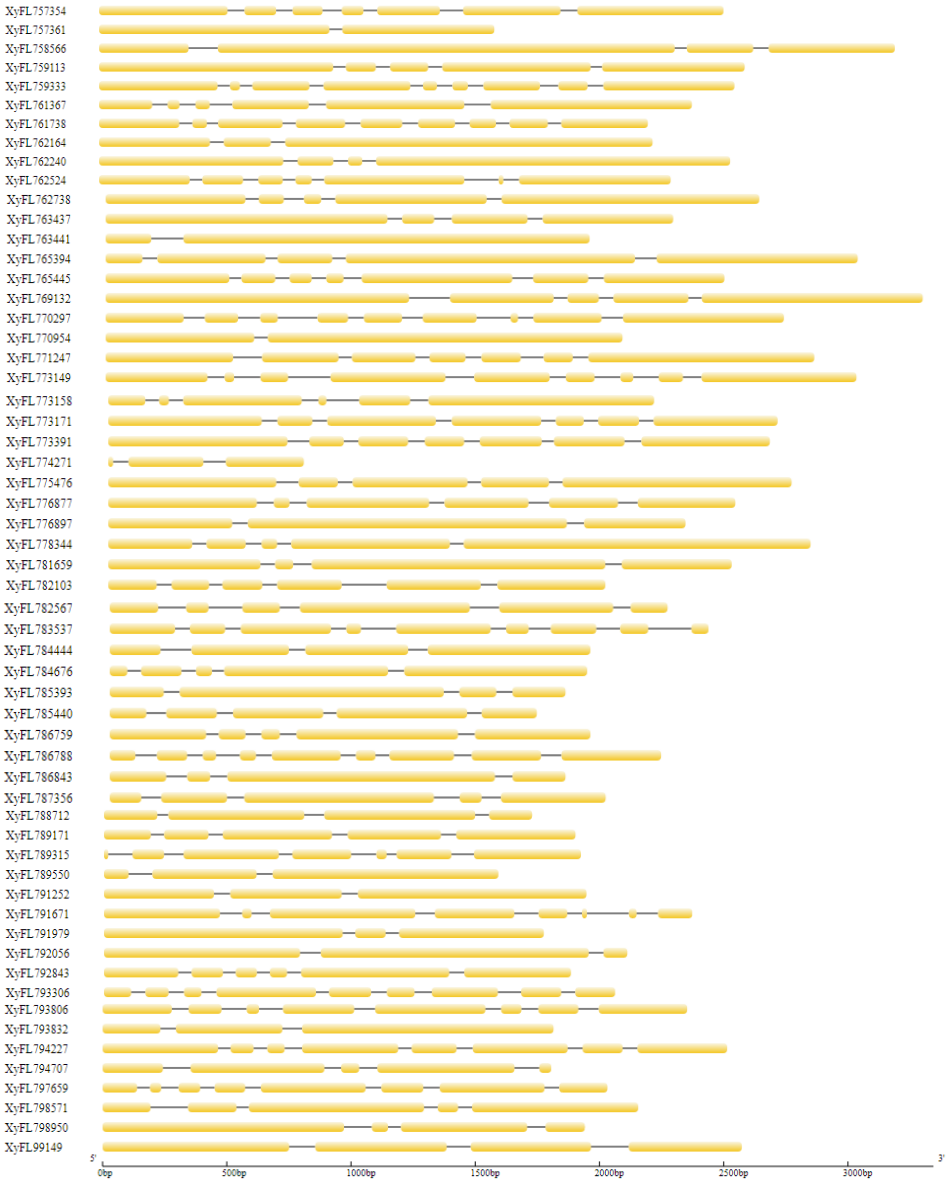
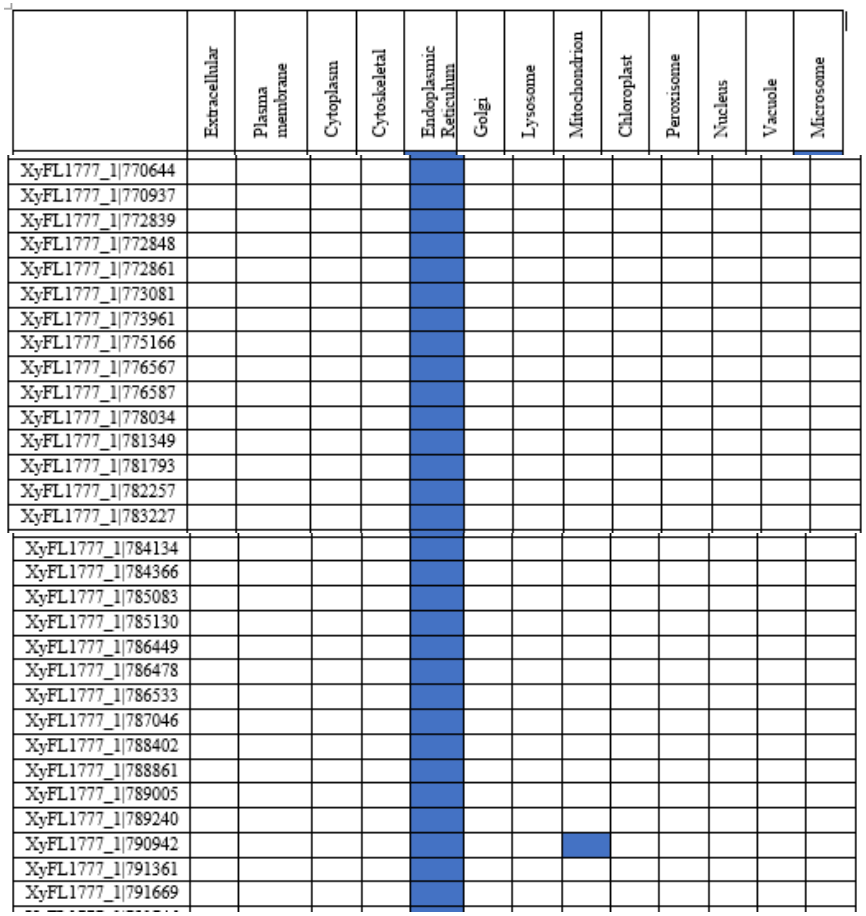


Table 2: Sub-cellular Localization of CYPs in *Xylaria* sp.

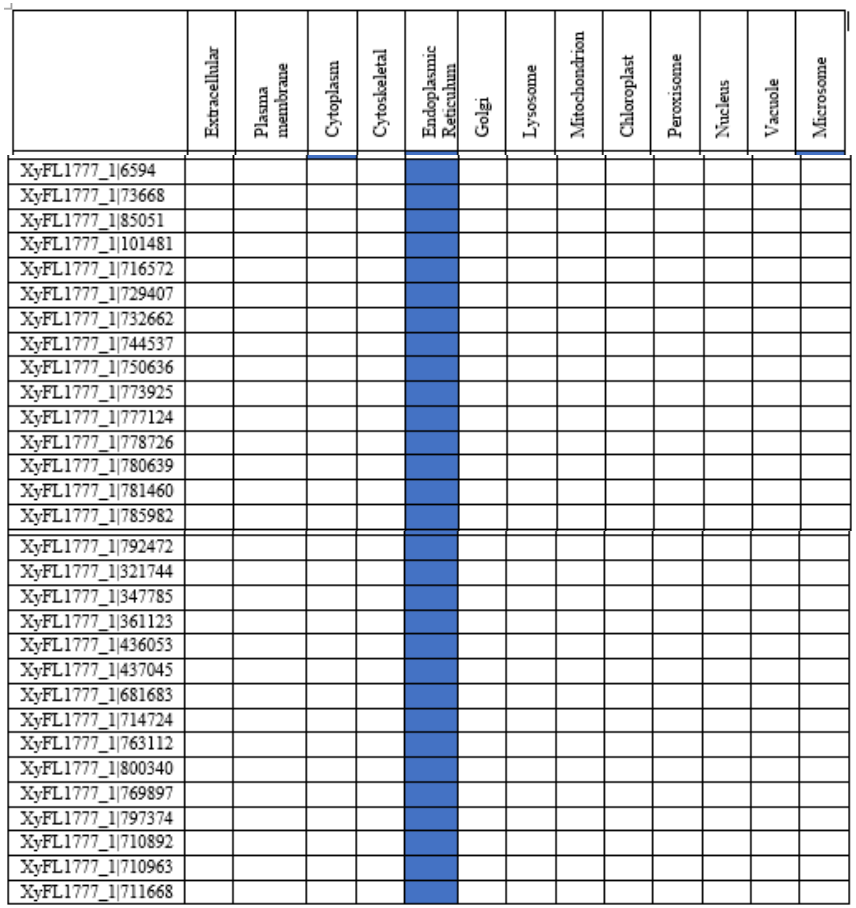
	Extracellular	Plasma membrane	Cytoplasm	Cytoskeletal	Endoplasmic Reticulum	Golgi	Lysosome	Mitochondrion	Chloroplast	Peroxisome	Nucleus	Vacuole	Microsome
XyFL1777_1|758256					■								■
XyFL1777_1|527483					■								
XyFL1777_1|529537					■								
XyFL1777_1|539004					■								
XyFL1777_1|15798					■								
XyFL1777_1|610396					■								
XyFL1777_1|98839					■								
XyFL1777_1|114386					■								
XyFL1777_1|639691					■								
XyFL1777_1|661053					■								
XyFL1777_1|661090					■								
XyFL1777_1|137640					■								
XyFL1777_1|668719					■								
XyFL1777_1|686302					■								
XyFL1777_1|162745					■								
XyFL1777_1|176420					■								
XyFL1777_1|709897					■								
XyFL1777_1|710843													■
XyFL1777_1|716091					■								
XyFL1777_1|716253					■								
XyFL1777_1|716627					■								
XyFL1777_1|717063					■								
XyFL1777_1|717342					■								
XyFL1777_1|717532					■								
XyFL1777_1|718688					■								
XyFL1777_1|720953					■								
XyFL1777_1|720954					■								
XyFL1777_1|729318					■								
XyFL1777_1|207295					■								
XyFL1777_1|731754					■								
XyFL1777_1|734602					■								
XyFL1777_1|736103					■								
XyFL1777_1|736121					■								
XyFL1777_1|736391					■								
XyFL1777_1|738596					■								
XyFL1777_1|215050					■								
XyFL1777_1|743207					■								
XyFL1777_1|745323					■								
XyFL1777_1|750643					■								
XyFL1777_1|750913					■								
XyFL1777_1|750910					■								
XyFL1777_1|752814					■								
XyFL1777_1|753128					■								
XyFL1777_1|229834					■								
XyFL1777_1|757044					■								
XyFL1777_1|757051					■								
XyFL1777_1|758803					■								
XyFL1777_1|759023					■								
XyFL1777_1|761057					■								
XyFL1777_1|761428					■								
XyFL1777_1|761854					■								
XyFL1777_1|761930					■								
XyFL1777_1|762214					■								
XyFL1777_1|762428					■								
XyFL1777_1|763127					■								
XyFL1777_1|763131					■								
XyFL1777_1|765084					■								
XyFL1777_1|765135					■								
XyFL1777_1|768822													
XyFL1777_1|769987					■								



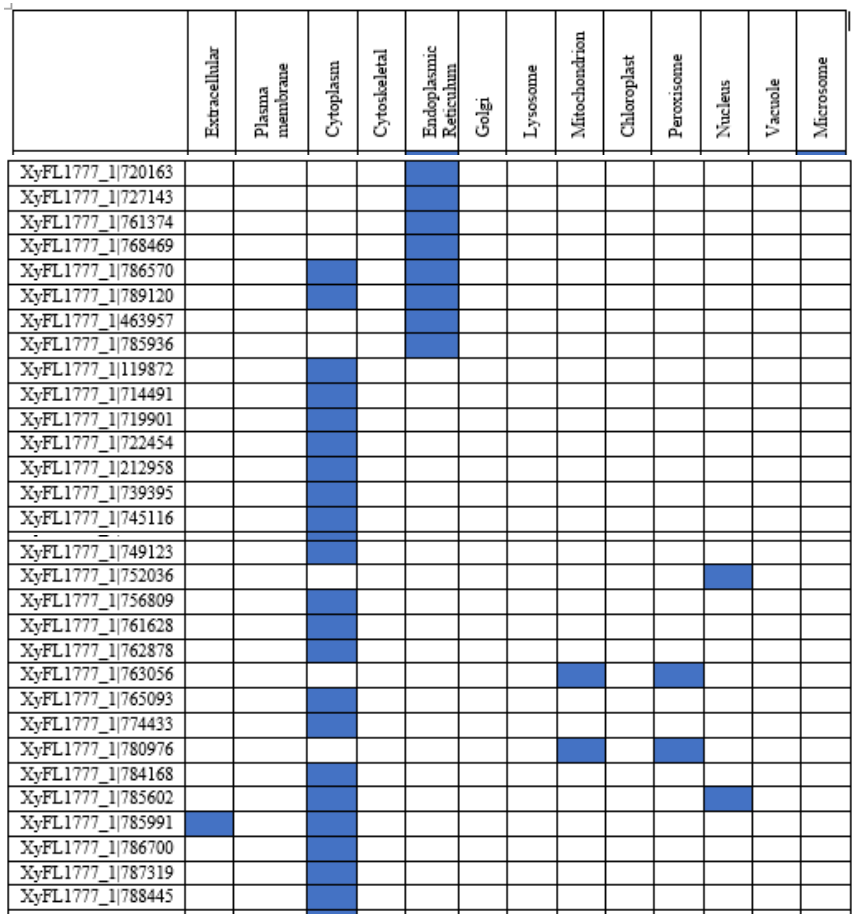
	Extracellular	Plasma membrane	Cytoplasm	Cytoskeletal	Endoplasmic Reticulum	Golgi	Lysosome	Mitochondrion	Chloroplast	Peroxisome	Nucleus	Vacuole	Microsome
XyFL1777_1|770644					■								
XyFL1777_1|770937					■								
XyFL1777_1|772839					■								
XyFL1777_1|772848					■								
XyFL1777_1|772861					■								
XyFL1777_1|773081					■								
XyFL1777_1|773961					■								
XyFL1777_1|775166					■								
XyFL1777_1|776567					■								
XyFL1777_1|776587					■								
XyFL1777_1|778034					■								
XyFL1777_1|781349					■								
XyFL1777_1|781793					■								
XyFL1777_1|782257					■								
XyFL1777_1|783227					■								
XyFL1777_1|784134					■								
XyFL1777_1|784366					■								
XyFL1777_1|785083					■								
XyFL1777_1|785130					■								
XyFL1777_1|786449					■								
XyFL1777_1|786478					■								
XyFL1777_1|786533					■								
XyFL1777_1|787046					■								
XyFL1777_1|788402					■								
XyFL1777_1|788861					■								
XyFL1777_1|789005					■								
XyFL1777_1|789240					■								
XyFL1777_1|790942					■			■					
XyFL1777_1|791361					■								
XyFL1777_1|791669					■								

Table 2 (contd): Sub-cellular Localization of CYPs in *Xylaria* sp.

	Extracellular	Plasma membrane	Cytoplasm	Cytoskeletal	Endoplasmic Reticulum	Golgi	Lysosome	Mitochondrion	Chloroplast	Peroxisome	Nucleus	Vacuole	Microsome
XyFL1777_1|791746					■								
XyFL1777_1|792533					■								
XyFL1777_1|792996					■								
XyFL1777_1|793496					■								
XyFL1777_1|793522					■								
XyFL1777_1|793917					■								
XyFL1777_1|794397					■								
XyFL1777_1|797349					■								
XyFL1777_1|798261					■								
XyFL1777_1|798640					■								
XyFL1777_1|798652					■								
XyFL1777_1|799228					■								
XyFL1777_1|799547					■								
XyFL1777_1|799609					■								
XyFL1777_1|800343					■								
XyFL1777_1|800330					■			■					
XyFL1777_1|282062					■								
XyFL1777_1|294539					■								
XyFL1777_1|309686					■								
XyFL1777_1|310348					■								
XyFL1777_1|324323					■								
XyFL1777_1|324419					■								
XyFL1777_1|350126					■								
XyFL1777_1|412762					■								
XyFL1777_1|428097					■								
XyFL1777_1|456018					■								
XyFL1777_1|496257					■								
XyFL1777_1|113473					■								■
XyFL1777_1|783610			■						■				
XyFL1777_1|473228			■										



	Extracellular	Plasma membrane	Cytoplasm	Cytoskeletal	Endoplasmic Reticulum	Golgi	Lysosome	Mitochondrion	Chloroplast	Peroxisome	Nucleus	Vacuole	Microsome
XyFL1777_1|6594					■								
XyFL1777_1|73668					■								
XyFL1777_1|85051					■								
XyFL1777_1|101481					■								
XyFL1777_1|716572					■								
XyFL1777_1|729407					■								
XyFL1777_1|732662					■								
XyFL1777_1|744537					■								
XyFL1777_1|750636					■								
XyFL1777_1|773925					■								
XyFL1777_1|777124					■								
XyFL1777_1|778726					■								
XyFL1777_1|780639					■								
XyFL1777_1|781460					■								
XyFL1777_1|785982					■								
XyFL1777_1|792472					■								
XyFL1777_1|321744					■								
XyFL1777_1|347785					■								
XyFL1777_1|361123					■								
XyFL1777_1|436053					■								
XyFL1777_1|437045					■								
XyFL1777_1|681683					■								
XyFL1777_1|714724					■								
XyFL1777_1|763112					■								
XyFL1777_1|800340					■								
XyFL1777_1|769897					■								
XyFL1777_1|797374					■								
XyFL1777_1|710892					■								
XyFL1777_1|710963					■								
XyFL1777_1|711668					■								



	Extracellular	Plasma membrane	Cytoplasm	Cytoskeletal	Endoplasmic Reticulum	Golgi	Lysosome	Mitochondrion	Chloroplast	Peroxisome	Nucleus	Vacuole	Microsome
XyFL1777_1|720163					■								
XyFL1777_1|727143					■								
XyFL1777_1|761374					■								
XyFL1777_1|768469					■								
XyFL1777_1|786570			■		■								
XyFL1777_1|789120			■		■								
XyFL1777_1|463957					■								
XyFL1777_1|785936					■								
XyFL1777_1|119872			■										
XyFL1777_1|714491			■										
XyFL1777_1|719901			■										
XyFL1777_1|722454			■										
XyFL1777_1|212958			■										
XyFL1777_1|739395			■										
XyFL1777_1|745116			■										
XyFL1777_1|749123			■										
XyFL1777_1|752036											■		
XyFL1777_1|756809			■										
XyFL1777_1|761628			■										
XyFL1777_1|762878			■										
XyFL1777_1|763056								■		■			
XyFL1777_1|765093			■										
XyFL1777_1|774433			■										
XyFL1777_1|780976								■		■			
XyFL1777_1|784168			■										
XyFL1777_1|785602			■								■		
XyFL1777_1|785991	■		■										
XyFL1777_1|786700			■										
XyFL1777_1|787319			■										
XyFL1777_1|788445			■										

Table 2 (contd): Sub-cellular Localization of CYPs in *Xylaria* sp.

	Extracellular	Plasma membrane	Cytoplasm	Cytoskeletal	Endoplasmic Reticulum	Golgi	Lysosome	Mitochondrion	Chloroplast	Peroxisome	Nucleus	Vacuole	Microsome
XyFL1777_1|415484			■										
XyFL1777_1|429526			■										
XyFL1777_1|164234			■										
XyFL1777_1|255228			■										
XyFL1777_1|782918			■										
XyFL1777_1|794830			■										
XyFL1777_1|797204			■										
XyFL1777_1|800736	■		■		■								
XyFL1777_1|716563					■								
XyFL1777_1|781569							■						
XyFL1777_1|736803			■										
XyFL1777_1|741070			■										
XyFL1777_1|751867			■										
XyFL1777_1|796647					■								
XyFL1777_1|575258	■		■										
XyFL1777_1|713135			■										
XyFL1777_1|714030									■				
XyFL1777_1|715665	■		■					■					
XyFL1777_1|738471										■			
XyFL1777_1|760888			■					■					
XyFL1777_1|763710	■	■	■		■			■			■		
XyFL1777_1|787223	■								■				
XyFL1777_1|382656	■	■	■					■					
XyFL1777_1|398156			■										
XyFL1777_1|413182	■	■	■				■	■					
XyFL1777_1|506634			■										
XyFL1777_1|783203			■		■								
XyFL1777_1|643471	■		■					■					
XyFL1777_1|762032	■				■								
XyFL1777_1|789419			■										
XyFL1777_1|799475			■										

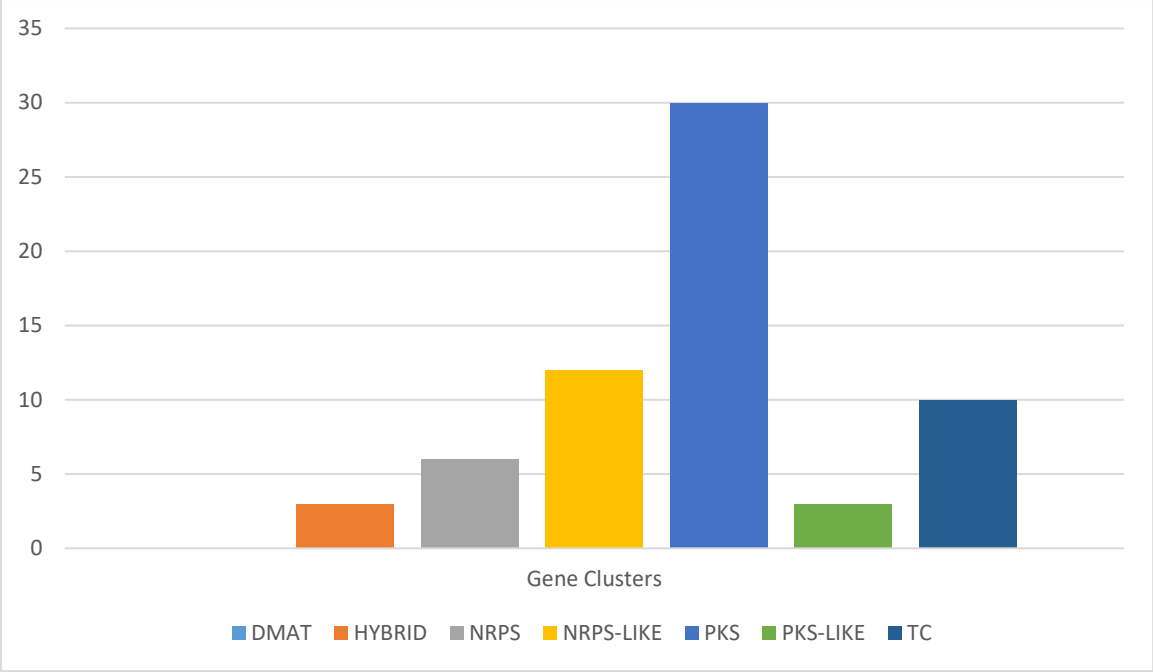


Figure 4: Distribution of secondary metabolic related gene clusters in *Xylaria* sp. FL1777
DMAT: dimethylallyltryptophan synthase
NRPS-Non ribosomal peptide synthetase
NRPS-like- Non ribosomal peptide synthetase-like
PKS- Polyketide synthase
PKS- polyketide synthase-Like
TC-Toxin complex