**Genome-wide analysis of cytochrome P450 genes of *Macrophomina phaseolina*: annotation and evolutionary relationships and structural analysis**

**Wadzani Dauda Palnam1, Daji Morumda2, Peter Abraham3, Charles Oluwaseun Adetunji 4, Shakira Ghazanfar5**  **Elkanah Glen6, Shittu, Emmanuel Abraham7, Peter, Grace Wabba8, Ogra Israel Ogwuche8, and Mawuli Kwamla Azameti9**

1Crop Science Unit, Department of Agronomy, Federal University, Gashua, Yobe State, Nigeria

2Department of Microbiology, Federal University Wukari, Taraba State, Nigeria

3Department of Horticulture, Federal College of Horticulture Dadin Kowa, Gombe, Nigeria

4Applied Microbiology, Biotechnology and Nanotechnology Laboratory, Department of Microbiology, Edo University Iyamho, PMB 04, Auchi, Edo State, Nigeria.

5National Agricultural Research Centre, National Institute of Genomics and Agriculture Biotechnology (NIGAB), Park Road, Islamabad 45500, Pakistan

6Department of Biochemistry, Federal University Lokoja

7Department of Agronomy, Bayero University Kano, Nigeria

8Department of Biochemistry, Faculty of Life sciences, Ahmadu Bello University, Zaria

9Division of Molecular Biology and Biotechnology, Indian Agricultural Research Institute, New Delhi, India

**\*Corresponding author’s email address**: wadzanidauda@gmail.com, ORCID I.D:0000-0002-5233-0058

**Novelty Statement**

*Macrophomina phaseolina* is an important soil-borne plant pathogen infecting over five hundred crops and causing diseases of economic importance including damping-off, colour rot, charcoal rot, stem rot, root rot, and seedling blight. Cytochrome P450 (CYPs) is one of the indispensable enzyme families that regulate myriads of metabolic processes in living systems. Several *in-silico* studies have been documented on the Cyps of some important fungal species but this is, however, lacking in *Macrophomina phaseolina.* The results of this present study elucidated relevant information on the cyps of *M. phaseolina* that could be harnessed for its effective management.

**Abstract**

Cytochrome P450 (CYPs) are a haem-containing monooxygenases family of enzymes involved in the metabolism of vast wide groups of compounds. Moreover, different aspects of fungal cytochrome p450 have been studied, such as their family diversity, catalytic versatility, P450 family enrichment, thermostable P450s, P450s as drug-target, and their values in biotechnology. The fungal organism *Macrophomina phaseolina* is an important plant pathogen with a wide host range capable of producing a vast array of enzymes that mediate its pathogenic lifestyle. In this study, CYP genes of *Macrophomina phaseolina* were retrieved, aligned, and phylogenetic analysis was carried out. Gene structure and motif elucidation were done, and the subcellular location of the CYPs was determined. The proteins were found to originate from a single ancestor and branched into three broad clades. 7 motifs were identified, and gene structure revealed both mono and poly-exonic genes belonging to three CYP clans. The majority of the cytochrome p450 proteins were found to be localized in the endomembrane system. This study gives an insight into the gene structure, phylogenetic, protein motif, functional classification, and subcellular localization of cytochrome p450 proteins in *Macrophomina phaseolina*.

**Keywords:** Cytochrome P450, Gene structure, *Macrophomina phaseolina,* Phylogeny, Subcellular localization

**Introduction**

Cytochrome P450 (CYPs) are haem-containing monooxygenases family of enzymes (Kelly *et al*., 2009) involved in the metabolism of extensive diversity of both exogenous and endogenous compounds. Their ability to catalyze the regio-chemo- and stereospecific oxidation of a huge number of substrates under mild reaction circumstances makes them an important set of actors in xenobiotic degradation. They are actively involved in both primary and secondary metabolism, thereby enabling them to accomplish chemical transformations, (Chadha *et al.,* 2018), exploiting compounds as main carbon and energy sources, cellular detoxification, etc., (Jossue *et al.,* 2020). All living organisms belonging to the biological kingdom possess CYPs, and among these kingdoms, fungal CYPs is the most studied, and the different aspects of these enzymes, including their family diversity, catalytic versatility, P450 family enrichment, thermostable P450s, their use as drug-target and provision of biotechnologically valuable products has been known (Qhanya *et al.,* 2015). The fitness and fecundity of fungi to various ecological niches can be attributed to the extensive participation of CYPs in a wide variety of physiological reactions (Chen *et al*., 2014). CYPs are made used by filamentous Fungi to produce a wide array of secondary metabolites that are important to biomedicine, agriculture, and industries. Extensions and functional variations of the fungal CYP families can be attributed to the evolution of fungal pathogenicity (Chen *et al.,* 2014). *Macrophomina phaseolina* is a fungus that has been implicated as an important plant pathogen. It is soil-borne and has a very wide host range of approximately 500 cultivated and wild plant species globally (Babu *et al.,* 2007). It causes damping-off, colour rot, charcoal rot, stem rot, root rot, and seedling blight in many important crops of economic significance. *Macrophomina phaseolina* secrets a wide array of enzymes that degrade the cell wall of plants by depolymerizing it, leading to the collapse of the overall plant structure (Marquez *et al.,* 2018). Among these enzymes, pectinases, xylanases, cellulases and proteases have been identified as the significant enzymes secreted by this organism (Khan *et al.,* 2017). Considering the importance of *Macrophomina phaseolina* to agriculture and the essential role CYPs play in the biology of fungi, it is therefore important to annotate and study the evolution of the CYPs in this organism which will help in both the commercial exploitation of these enzymes and in designing control measures. Although earlier studies on the CYPs of several important fungal species have been well documented (Moktali *et al.,* 2012; Chen *et al.,* 2014; Kgosiemang *et al.,* 2014; Sello *et al.,* 2015; Dauda *et al.,*2021a, 2022a, 2022b) this subject has not been covered in this important plant pathogen. Hence, this study intends to perform a genome-wide study of the structure, phylogenetic analysis, and functional annotation of CYPs genes of *Macrophomina phaseolina* coupled with the classification and subcellular localization analysis of the CYP proteins with a broader aim to facilitate the classification/nomenclature and to understand the functional diversity and evolution of the P450s in this important plant pathogenic fungus.

**Materials and Methods**

***Sequence retrieval***

The cytochrome p450 protein sequences of *Macrophomina phaseolina* were retrieved from the Joint Genome Institute mycocosm the fungal genome database (mycocosm.jgi.doe.gov/pages/search-for-genes.jsf).

***Structural feature analysis of CYP protein sequence***

This was done based on the analysis of the sequences with the CYP signature in the JGI database (mycocosm.jgi.doe.gov/pages/search-for-genes.jsf). The CYPs conserved domains were further identified using the Conserved Domain Database to confirm the presence of CYP sequences in the retrieved sequence.

***Phylogenetic analysis***

The downloaded protein sequences in FASTA format were aligned using the multiple sequence alignment program MUSCLE® as implemented in MEGA X software (Kumar *et al.* 2018). The evolutionary history was inferred by using the Maximum Likelihood method and JTT matrix-based model (Jones *et al*., 1992). Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the JTT model and then selecting the topology with a superior log-likelihood value. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. This analysis involved 318 amino acid sequences. There were a total of 2490 positions in the final dataset. Evolutionary analyses were conducted in MEGA X (Kumar *et al.,* 2018) as described by Dauda et al. (2021b).

***Gene structure analysis***

The exon-intron structures for cytochrome P450 genes of *Macrophomina phaseolina* were analyzed using Gene Structure Display Server (GSDS 2.0) ([http://gsds.gao-lab.org/)](http://gsds.gao-lab.org/)%20) (Bo *et al.,*2015) the server graphically displayed the numbers and positions of exons and introns after loading the coding and genomic FASTA sequences.

***Protein Motif Elucidation***

The discovery of the motifs of cytochrome P450 genes of *Macrophomina phaseolina* was performed 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) which was set for 7 motifs. The minimum width of the protein motif was 29 amino acids, while the maximum was 50 amino acids, respectively.

***Identification of clans, families, and putative functions***

*Macrophomina phaseolina* CYP protein sequences were blasted in the NCBI CDD site for the identification of their characteristic functional class. The identified proteins were further compared with the already characterized CYPs in the Fungal cytochrome P450 database (http://p450.riceblast.snu.ac.kr) for the determination of their family and clan.

***Subcellular Localization analysis***

Sub-cellular localization analysis was carried out to obtain more understanding of the functional mechanism of P450 genes. This was done using the BUSCA integrative web server for predicting the subcellular localization of proteins in fungi http:busca.biocomp.unibo.it (Savojardo *et al.,* 2018). For this, 318 cytochrome P450 protein sequences were used, and the analysis was done at the default settings of the server by setting the taxonomic origin of the inputted sequence to eukarya fungi 9-compartment.

**Results**

***Sequence retrieval***

A total of 318 cytochrome P450 genes of *Macrophomina phaseolina* were retrieved from the joint genome institute*.*

***Phylogenetic analysis***

The evolutionary studies performed during this study showed that there are 3 phylogenetic groups of cytochrome P450 proteins of *Macrophomina phaseolina,* which are indicated by different colours as illustrated in Fiure 1. Group 1 (green bars) had the highest number of protein clusters consisting of a total of 140 protein families, this is followed by group 3 (red bars), which had 124 proteins, while group 2 (blue bars) had the least consisting of 54 protein family.

***Gene structure***

The exon-intron structure of the 318 genes presented in Figure 1 was analyzed by comparing their coding sequences with the corresponding genomic sequences using the online tool GSDS2.0 (<http://gsds.gao-lab.org/Gsds>). The highest number of introns (10) was in the gene MPH07555, MPH03440, MPH00935 immediately follow this, and MPH01839, having 9 introns, respectively.

**Motif analysis**

MEME analysis of the 318 cytochrome p450 genes of *Macrophomina phaseolina* revealed the result of 7 motifs as shown in Figures 2 a and b. Motif 6 had the highest frequency of 236 sites with the highest amino acid width of 50. followed by motif 7 with 211 sites and a width of 29 amino acids, while motifs 1, 2, 3, 4, and 5 had 128, 124, 106, 56, and 143 sites, respectively. Motif 4 had the highest E-value, followed by motif 1.

***Clans,* and *families of cytochrome P450 in Macrophomina phaseolina***

In *Macrophomina phaseolina*, three CYP clans were discovered, which were CYP533, CYP65 and CYP52, as presented in Table 1. All the CYP clans in this organism had one CYP family, where CYP533 had CYP64 family, CYP65 clan had CYP60, and CYP52 clan had CYP52 family, respectively. No clan was discovered for the CYP120 family. However, the distribution of the CYP proteins in these respective families differed, with CYP60 having the highest number of proteins (27), followed by CYP64, which had 16 proteins, while CYP52 had 5 proteins while the CYP120 family had the least number of proteins (2).

***Subcellular Localization***

The subcellular prediction of cytochrome P450 proteins was carried out using the BUSCA integrative web server, as presented in figure 4. For this, it was discovered that the 318 cytochrome P450 proteins of *Macrophomina phaseolina* were localized within eight (8) subcellular compartments, which are the cytoplasm, endomembrane system, extracellular space, mitochondrial membrane, mitochondrion, nucleus, organelle membrane and plasma membrane. The Endomembrane system was discovered to have the highest cytochrome P450 proteins in this organism (151). Cytoplasm (63) followed this, while 39 were localized in organelle membrane, 29 in Mitochondrion, 24 in Extracellular space and 10 in Plasma membrane.

**Discussion**

The 318 Cytochrome P450 genes recovered in this organism reveal a high amount of these genes in this fungus showing the importance of these proteins in the survival of this fungus as they play important biological and physiological roles (Chen *et al.,* 2014). This could be part of the factors that can enable this organism to overcome the host plant defence system to establish itself and mediate its damage which could eventually be the reason for its success as a pathogen of over 500 different plant species (Islam *et al*., 2012). *M. phaseolina* has a very large collection of hydrolytic enzymes that are able to degrade all the major components of the plant cell wall and cuticle, including cellulose, hemi-cellulose, pectin, lignin, and cutin. Also, research has shown that the cellulolytic activity of *M. Phaseolina* was found to be higher than that of other plant pathogenic fungi thereby making it an excellent pathogen adapted for its pathogenic lifestyle (Ramos *et al*., 2016).

The phylogenetic analysis of the Cytochrome P450 revealed 3 phylogenetic groups of the proteins in *Macrophomina phaseolina.* These proteins exhibit a monophyletic evolutionary trait with a common ancestor that has diverged over time into three broad clades. The clustering of this cytochrome p450 into three clades, each with many families indicates gene duplication and divergent evolution (Chen *et al.,* 2014). This expansion and functional diversification of these cytochromes proteins explain the reasons for the broad host range of *Macrophomina phaseolina*, which includes plants in many orders and families, which are distributed in varied climatic regions of the world and its advanced pathogenicity which is needed for (Chen *et al.,* 2014 and Sarr *et al*. 2014) as studies of cytochromes have well associated them with pathogenicity of fungi to plants (Chen *et al.,* 2014). Each of the three clades is seen to have an irregular topology and pattern of distribution. This topology is similar to the reports of Machodo *et al.* (2018) who carried out a phylogenetic analysis of the Translation elongation factor-1 (TEF1-α), β-tubulin, Actin and calmodulin in *Macrophomina phaseolina and* discovered that each of these characters was separated into three broad clades thereby confirming a monophyly pattern with irregular distribution of the different families of whose evolution can also be attributed to gene duplication which is closely associated with the fungal lifestyle in their respective habitats. This can also provide useful information on fungal evolutionary adaptation to various ecological niches (Chen *et al.,* 2014).

The exon-intron structure of the 318 genes presented in Figure 1 where the highest number of introns (10) was in the gene MPH07555, MPH03440, MPH00935 immediately follow this, and MPH01839, having 9 introns, respectively. Introns have been reported to regulate alternative splicing, a molecular mechanism that produces multiple variants of a protein from a single gene, enhances gene expression, regulates nonsense-mediated decay by recognizing premature mRNA targets, and participates in mRNA transport or chromatin assembly (Jo and Choi, 2015). Fifty (15%) of these genes are mono-exonic, while 75% are poly-exonic. There were 4 upstream/downstream. Upstream/downstream are functional elements involved in gene regulation by playing a significant role in post-transcriptional control particularly under stress conditions and at the translational level (Lawless *et al.,* 2009). This can confer adaptational advantages to this pathogen thereby enabling it to cope with stress it comes in contact with in the host enabling this pathogen to broad host spectrum capability it has.

The result of motif analysis as sown in Figures 2 a and b revealed that motifs 6 and 7 had the highest frequency with highest number of amino acids. This means that motifs 7 and 6 are more vastly distributed than the others as found in almost all genomic sequences. It also shows that they are more conserved than the others. This is closely The variation in the E-value, amino acid width and sites does not have a regular pattern as it can be seen that the motif with the highest E-value had 128 sites and a width of 50. This reveals a low level of sequence semblance, as reported by Yu *et al.* (2014), although most of the sequences contain almost all the seven motifs indicating that they are conserved.

Three CYP clans were discovered in this analysis which was CYP533, CYP65 and CYP52, as presented in Table 1. All the CYP clans in this organism had one CYP family, where CYP533 had CYP64 family, CYP65 clan had CYP60, and CYP52 clan had CYP52 family, respectively. No clan was discovered for the CYP120 family. Cytochrome P450s from the CYP52 family have been reported to participate in the assimilation of alkanes and fatty acids (Jossue *et al.,* 2020). They catalyze a rate-limiting step in the hydroxylation of n-alkanes and fatty acids—which are then further metabolized via the β-oxidation pathway. The CYP60 family have been thought to be one member of the CYPs clusters that mediate in the pathway of the synthesis of aflatoxin in *Aspergillus flavus* and *A. parasiticus.* CYP102A1, also known as P450 BM3, is a fatty acid hydroxylase that is believed to be formed by the fusion of a cytoplasmic P450 (N-terminus) to a soluble CPR through a flexible peptide linker region, is an evolutionary adaptation of the eukaryotic class II system and has a much higher catalytic rate than the eukaryotic class II enzymes. BM3 and related enzymes catalyze rapid oxidation of fatty acids near the ω-methyl group (typically hydroxylation at ω-1 to ω-3 positions). However, notwithstanding being expansively studied to comprehend its molecular properties, BM3’s physiological role remains unclear. However, propositions have been made about its participation in the metabolism of toxic unsaturated fatty acids derived from plants (Paul, 2015) which could be why this pathogen is successful over a wide host range of plants.

As presented in figure 4, it was discovered that the 318 cytochrome P450 proteins of *Macrophomina phaseolina* were localized within eight (8) subcellular compartments, which are the cytoplasm, endomembrane system, extracellular space, mitochondrial membrane, mitochondrion, nucleus, organelle membrane and plasma membrane with the endomembrane system having the highest concentration of Cytochrome P450 in this organism. This could be because an endomembrane system is a group of membranes that are connected directly or through vesicular transport, enabling them to exchange materials, thereby forming a single functional and developmental unit (Smith, 1997), and the organelles of the endomembrane system in eukaryotes consist of the nuclear membrane, endoplasmic reticulum, Golgi apparatus, lysosomes, vesicles, endosomes and plasma membrane.

The concentration of these proteins in this system could also possibly be because cytochrome P450 (CYPs) are a family of enzymes (Kelly *et al*., 2009) that are involved in the metabolism of extensive diversity of both exogenous and endogenous compounds, and this system is the predominant location where active synthesis, packaging and transportation of substances occur in the cell. The mitochondrial membrane and Nucleus had only 1 cytochrome P450 protein localized in them, respectively. This pattern of cytochrome p450 proteins distribution is further proof indicating how this plant pathogen is well adapted for its parasitic lifestyle as most of this enzyme system goes into either synthesis, packaging or transportation of substances that are either endogenous or exogenous.

**Conclusion**

*Macrophomina phaseolina* is an important plant pathogen estimated to affect over 500 species of cultivated and wild plants. This study has revealed an enormous number of cytochrome p450 genes (318) in *Macrophomina phaseolina* originating from a single ancestor into three phylogenetic groups with 140, 123 and 54 members in each family. Their gene structure and protein motif analysis revealed an irregular pattern showing the varied diversity of these proteins, which could offer this organism such an ability. Three CYP clans discovered are CYP533, CYP52 and CYP64, each having one family, respectively. Also, CYP120 with an unidentified clan was discovered. It was found that 151 CYP proteins in this organism were concentrated in the endomembrane system. The findings from this study will provide a basis for the biotechnological exploitation of these enzymes and provide the basis for further functional molecular research on the management of this plant pathogen.

**Author Contributions**: Conceptualization, W.P.D.; writing—original draft preparation, M.D.; methodology, P.A.; writing—review and editing, W.P.D., P.A., C.O.A.; formal analysis, E.G.; validation, G.P.W.; investigation, I.O.O., data curation S.E.A.; visualization, M.K.A.; project administration, S.G., and supervision, W.P.D. and P.A. All authors have read and agreed to the published version of the manuscript.

**Conflicts of Interest**

The authors hereby confirm that we have no conflict of interest.

**Data Availability**

Data supporting the findings of this study are available in this article.

**Ethics Approval**

This article does not contain any studies with human participants or animals. The collection materials of the plants, complies with the relevant institutional, national, and international guidelines and legislation.

**Funding Source**

No funding was received from any organization for this study.

**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.

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.

Chen Wanping, Mi-Kyung Lee, Colin Jefcoate, Sun-Chang Kim, Fusheng Chen, and Jae-Hyuk Yu 2014. Fungal Cytochrome P450 Monooxygenases: Their Distribution, Structure, Functions, Family Expansion, and Evolutionary Origin. *Genome Biol. Evol*. 6(7):1620–1634. doi:10.1093/gbe/evu132

Dauda, W.P., Abraham, P., Glen, E., Adetunji, C.O., Ghazanfar, S., Ali, S., Al-Zahrani, M., Azameti, M.K., Alao, S.E.L., Zarafi, A.B. and Abraham, M.P., 2022a. Robust Profiling of Cytochrome P450s (P450ome) in Notable Aspergillus spp. Life, 12(3), p.451.

Dauda, W.P., Morumda, D., Abraham, P., Adetunji, C.O., Ghazanfar, S., Glen, E., Abraham, S.E., Peter, G.W., Ogra, I.O., Ifeanyi, U.J. and Musa, H., 2022b. Genome-Wide Analysis of Cytochrome P450s of Alternaria Species: Evolutionary Origin, Family Expansion and Putative Functions. *Journal of Fungi*, 8(4), p.324.

Dauda, W.P.; Glen, E.; Abraham, P.; Adetunji, C.O.; Morumda, D.; Abraham, S.E.; Wabba, G.P.; Ogwuche, I.O.; Azameti, M.K. 2021a. Comparative Phylogenomic Analysis of Cytochrome P450 Monooxygenases from *Fusarium* Species. Search. Sq. 2021. Preprint from Research Square, 02 December 2021a.

Dauda, W.P.; Abraham, P.; Fasogbon, I.V.; Adetunji, C.O.; Banwo, O.O.; Kashina, B.D.; Alegbejo, M.D. 2021b. Cassava mosaic virus in Africa: Functional analysis of virus coat proteins based on evolutionary processes and protein structure. Gene Rep. 24, 101239.

Diane E. Kelly, Nada Kraševec, Jonathan Mullins, David R. Nelson 2009. The CYPome (Cytochrome P450 complement) of Aspergillus nidulans. *Fungal Genetics and Biology* 46:S53–S61

In vitro growth and cell wall degrading enzyme production by Argentinean isolates of *Macrophomina phaseolina*, the causative agent of charcoal rot in corn. *Revista Argentina de Microbiología*, Volume 48, (4)267-273, <https://doi.org/10.1016/j.ram.2016.06.002>.

Jo, B. S., & Choi, S. S. 2015. Introns: The Functional Benefits of Introns in Genomes. *Genomics & informatics*, *13*(4), 112–118. https://doi.org/10.5808/GI.2015.13.4.112

Jones D.T., Taylor W.R., and Thornton J.M. 1992. The rapid generation of mutation data matrices from protein sequences. *Computer Applications in the Biosciences*8: 275-282.

Jossue Ortiz-Álvarez, Arturo Becerra-Bracho, Alfonso Méndez-Tenorio3, Jazmin Murcia-Garzón, Lourdes Villa-Tanaca & César Hernández-Rodríguez 2020. Phylogeny, evolution, and potential ecological relationship of cytochrome CYP52 enzymes in Saccharomycetales yeasts. *Scientific Report nature research. 10:10269* <https://doi.org/10.1038/s41598-020-67200-5>

Jossue ortiz-Álvarez, Arturo Becerra-Bracho, Alfonso Méndez-tenorio, Jazmin MurciaGarzón, Lourdes Villa-tanaca & césar Hernández-Rodríguez 2020. Phylogeny, evolution, and potential ecological relationship of cytochrome CYP52 enzymes in Saccharomycetales yeasts. *Sci Rep* **10,**10269 <https://doi.org/10.1038/s41598-020-67200-5>

Kelly, D.E.; Kraševec, N.; Mullins, J.; Nelson, D.R. 2009. The CYPome (cytochrome P450 complement) of Aspergillus nidulans. Fungal Genet. Biol. 46, S53–S61.

Kgosiemang IKR, Mashele SS, and Syed K 2014. Comparative genomics and evolutionary analysis of cytochrome P450 monooxygenases in fungal subphylum Saccharomycotina. *J Pure Appl Microbiol.* 8: 291–302

Khan AN, Shair F, Malik K, Hayat Z, Khan MA, Hafeez FY and Hassan MN 2017. Molecular Identification and Genetic Characterization *of Macrophomina phaseolina* Strains Causing Pathogenicity on Sunflower and Chickpea*. Front. Microbiol.* 8:1309. doi: 10.3389/fmicb.2017.01309

Kishore Babu, B., Saxena, A. K., and Arora, D. K 2007. Identification and detection of *Macrophomina phaseolina* by using specie-specific oligonucleotide primers and probes. *Mycologia* 99: 797-803

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.

Lawless, C., Pearson, R. D., Selley, J. N., Smirnova, J. B., Grant, C. M., Ashe, M. P., Pavitt, G. D., & Hubbard, S. J. 2009. Upstream sequence elements direct post-transcriptional regulation of gene expression under stress conditions in yeast. *BMC genomics*, *10*, 7. https://doi.org/10.1186/1471-2164-10-7

Lehlohonolo Benedict Qhanya, Godfrey Matowane, Wanping Chen, Yuxin Sun, Elizabeth Mpholoseng Letsimo, Mohammad Parvez, Jae-Hyuk Yu, Samson Sitheni Mashele, Khajamohiddin Syed 2015. Genome-Wide Annotation and Comparative Analysis of Cytochrome P450 Monooxygenases in Basidiomycete Biotrophic Plant Pathogens. *PLOS ONE* DOI: 10.1371/journal.pone.0142100

Machado A. R., Pinho D. B., Soares D. J., Gomes A. A. M., and Pereira O. L. 2018. Bayesian analyses of five gene regions reveal a new phylogenetic species of *Macrophomina* associated with charcoal rot on oilseed crops in Brazil. *Eur J Plant Pathol.* <https://doi.org/10.1007/s10658-018-1545-1>

Md Shahidul Islam, Md Samiul Haque, Mohammad Moinul Islam, Emdadul Mannan Emdad, Abdul Halim, Quazi Md Mosaddeque Hossen, Md Zakir Hossain, Borhan Ahmed, Sifatur Rahim, Md Sharifur Rahman, Md Monjurul Alam, Shaobin Hou, Xuehua Wan, Jennifer A Saito and Maqsudul Alam 2012. Tools to kill: Genome of one of the most destructive plant pathogenic fungi *Macrophomina phaseolina. BMC Genomics* 13:493<http://www.biomedcentral.com/1471-2164/13/493>

Moktali V, Park J, Fedorova-Abrams ND, Park B, Choi J, Lee YH, 2012. Systematic and searchable classification of cytochrome P450 proteins encoded by fungal and oomycete genomes. *BMC Genomics* 13: 525. doi:10.1186/1471-2164-13-525PMID:23033934

Nathalie Marquez, Maria L. Giachero, Adrien Gallou, Humberto J. Debat, Sylvie Cranenbrouck, Julio A. Di Rienzo, Maria J. Pozo, Daniel A. Ducasse, and Stephane Declerck 2018. Transcriptional changes in Mycorrhizal and non-Mycorrhizal Soybean upon infection with the fungal pathogen *Macrophomina phaseolina. Molecular Plant Interaction* Vol 31 No. 8 pp 842-845

Paul R. Ortiz de Montellano 2015. Cytochrome P450: Structure, Mechanism, and Biochemistry 4th edition. *Springer* International Publishing Switzerland

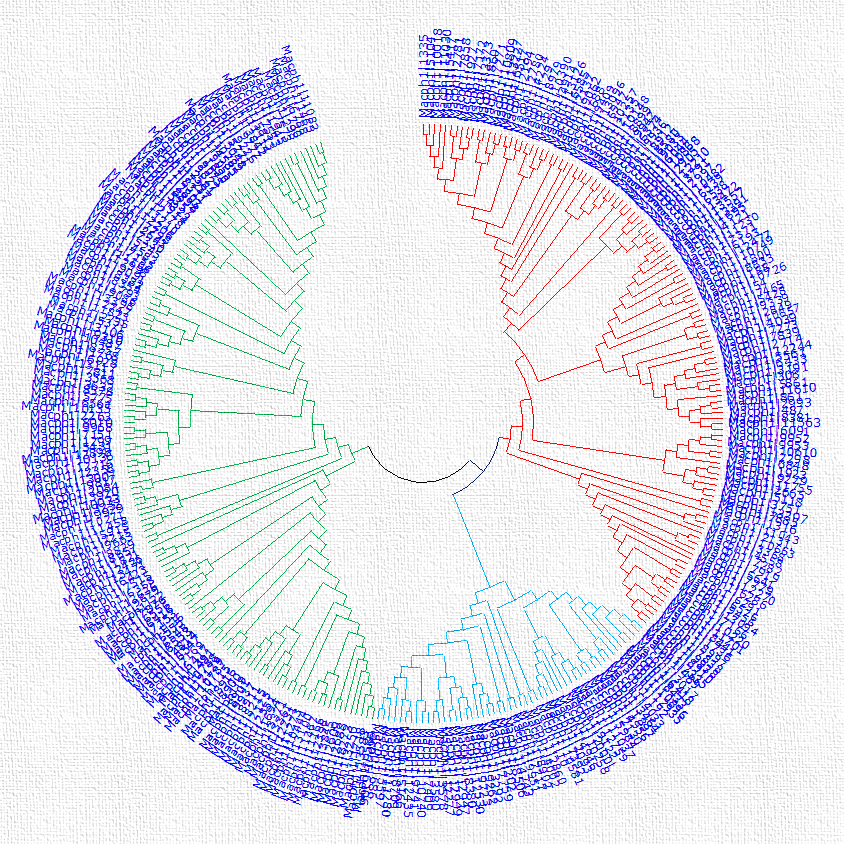
Sarr, M. P., Diaye, M. N., Groenewald, J. Z., & Crous, P. 2014. Genetic diversity in *Macrophomina* *phaseolina*, the causal agent of charcoal rot. *Phytopathologia Mediterranea*, 53,250–268.

Sello MM, Jafta M, Nelson DR, Chen W, Yu J-H, Parvez M, 2015. Diversity and evolution of cytochrome P450 monooxygenases in Oomycetes. *Sci Rep*. 5: 11572. doi:10.1038/srep11572PMID: 26129850

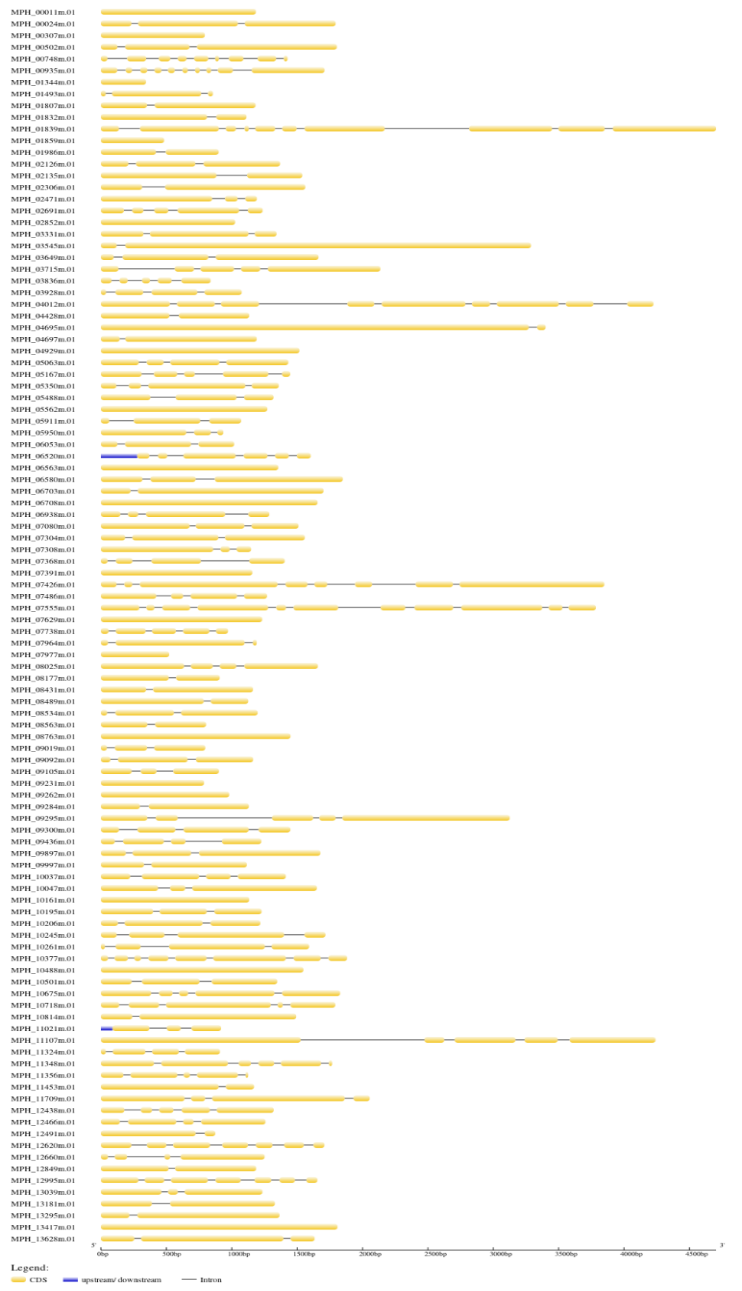
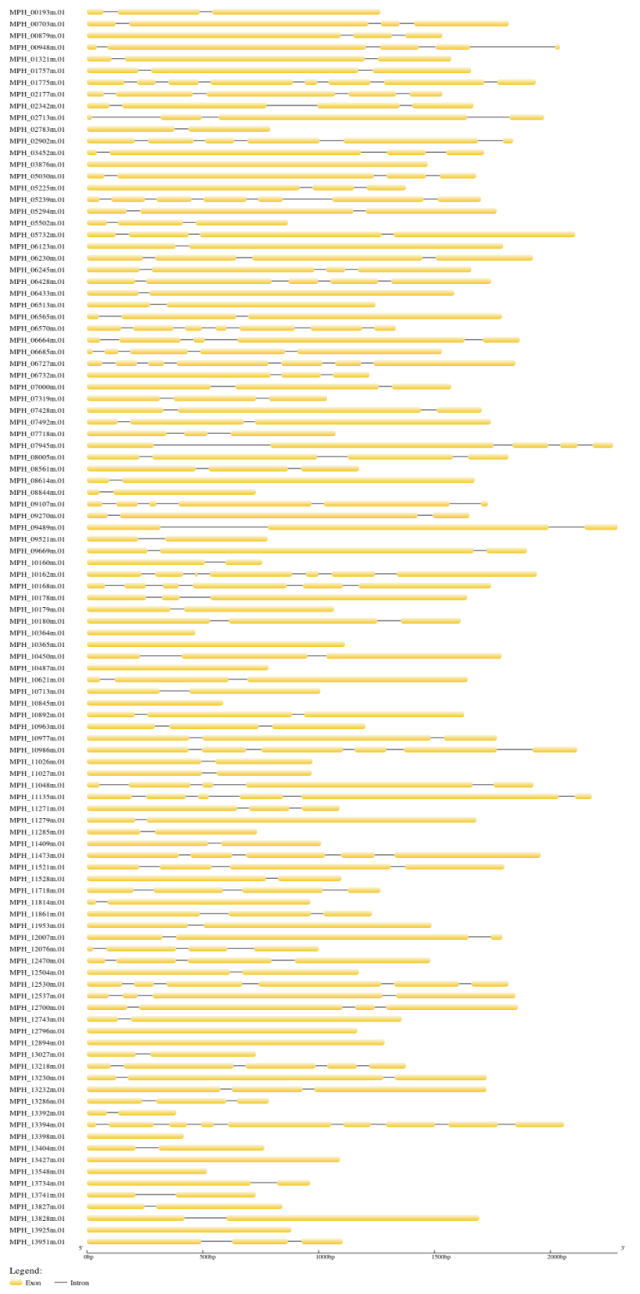
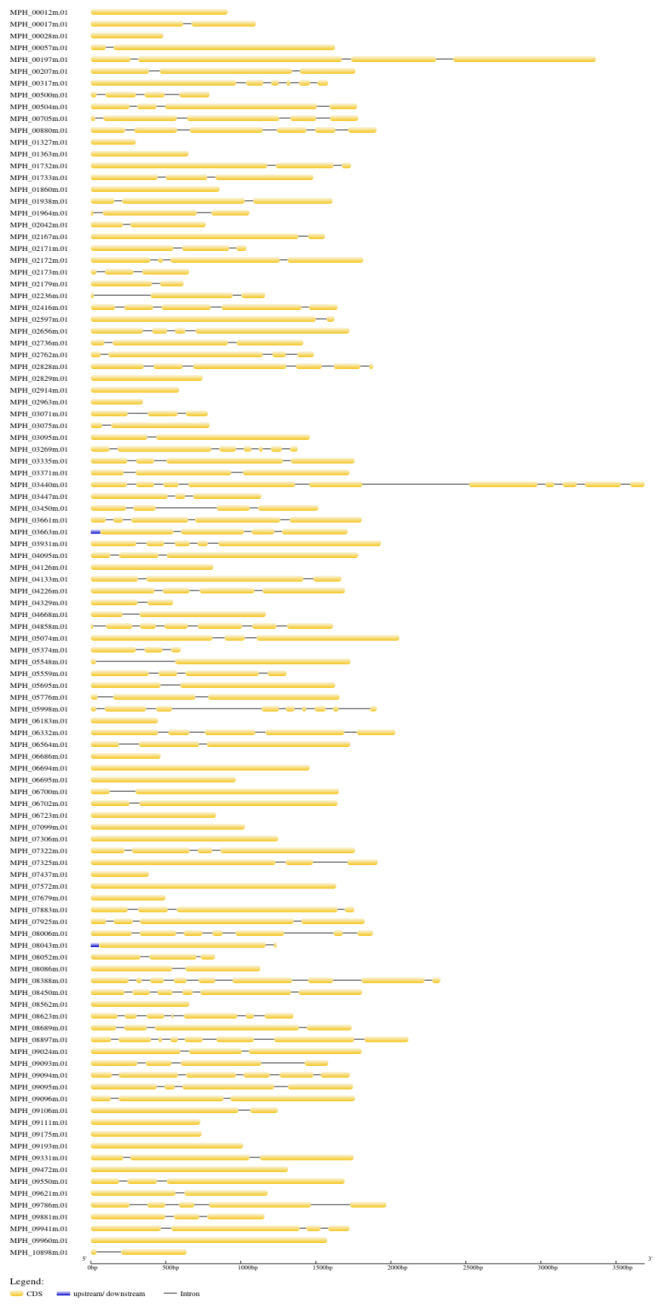
Smith A L. 1997. Oxford dictionary of biochemistry and molecular biology. Oxford: Oxford University press. Pp.206. ISBN 978-0-19-854768-6

Sonia Chadha, Sayaji T. Mehetre, Ravindra Bansal, Alan Kuo2, Andrea Aerts, Igor V. Grigoriev, Irina S. Druzhinina3 and Prasun K. Mukherjee 2018. Genome‑wide analysis of cytochrome P450s of *Trichoderma* spp.: annotation and evolutionary relationships. *Fungal Biol Biotechnol* 5:12 <https://doi.org/10.1186/s40694-018-0056-3>

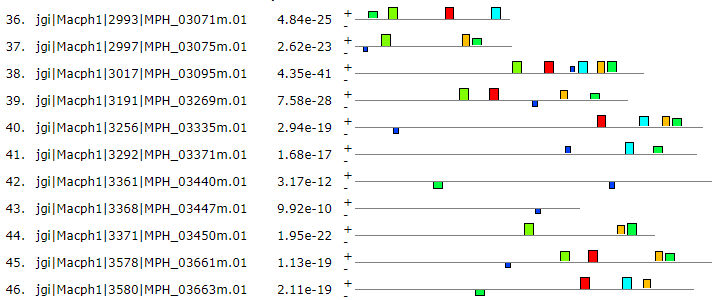
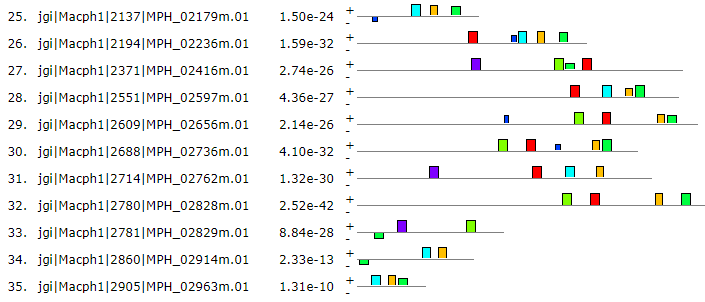
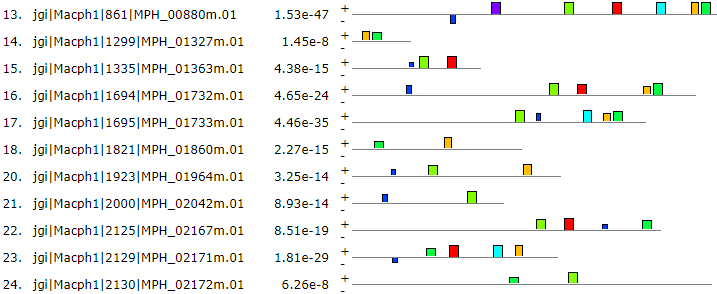
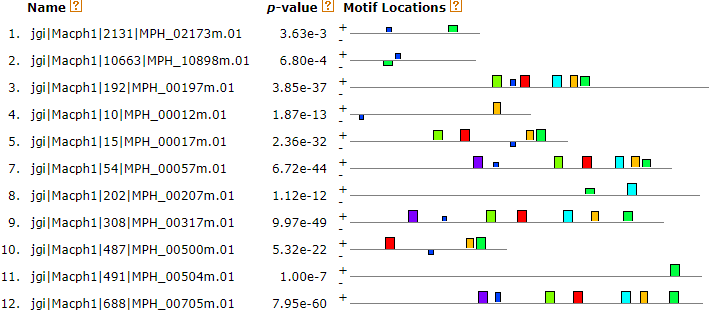
Yu J.H., Wanping C., Mi-Kyung L., Colin J., Sun-Chang K and Fusheng C. 2014. Fungal Cytochrome P450 Monooxygenases: their Distribution, Structure, Functions, Family Expansion, and Evolutionary Origin. *Genome of Biology Evolution*. 6(7):1620-1634.

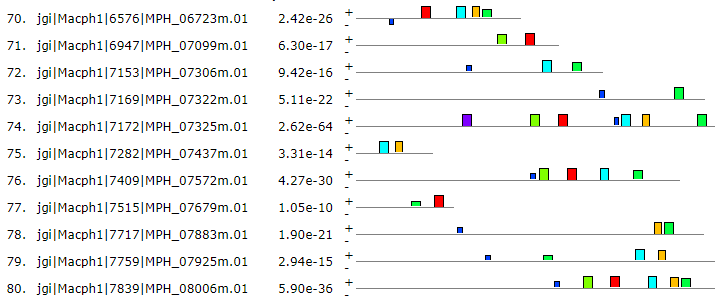
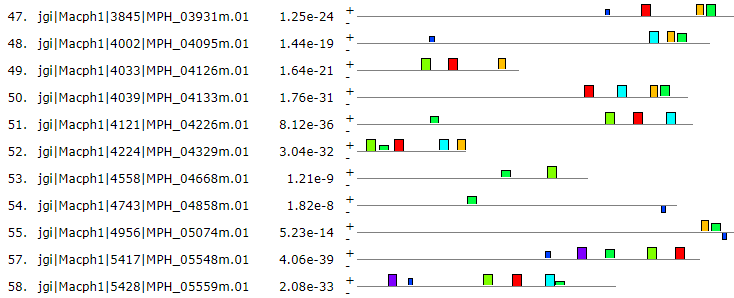


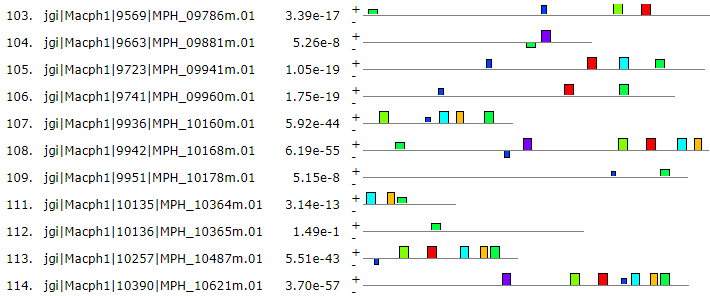
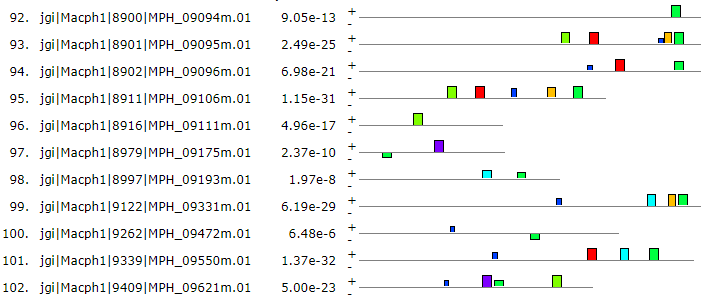
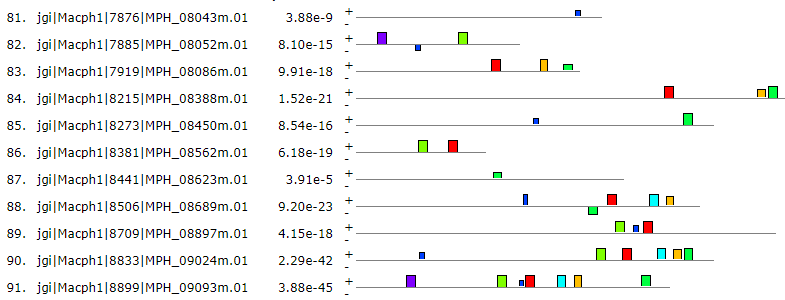
**Figure 1: Neighbor-joining unrooted phylogenetic tree of the 318 cytochrome p450 protein sequences from *Macrophomina phaseolina.* Pseudogene products are excluded due to large gaps in the sequences. Phylogenetic and molecular evolutionary analyses were conducted using MEGA X (Kumar *et al.,* 2018)**

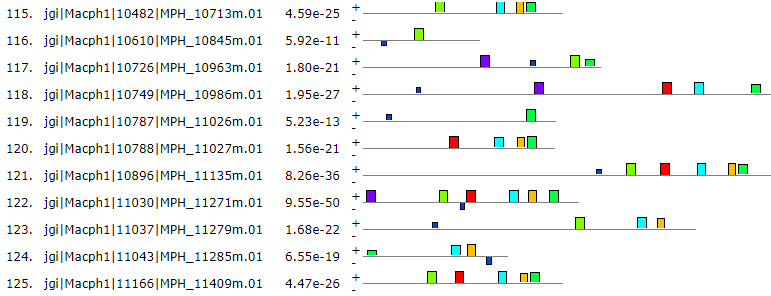


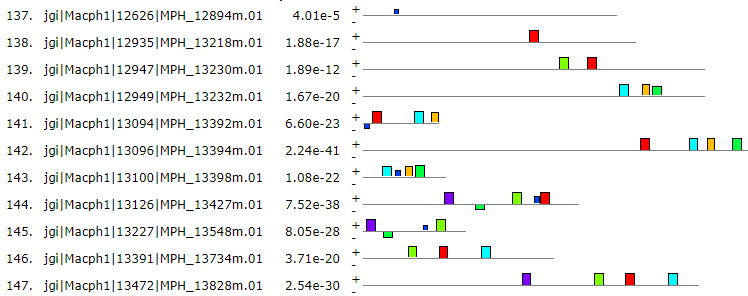
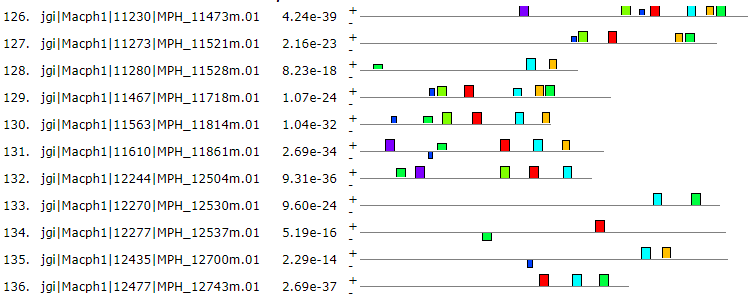
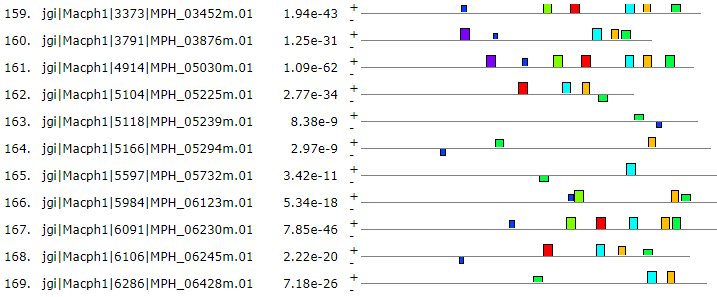
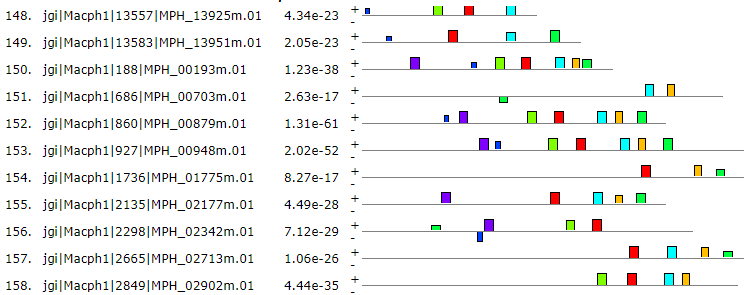
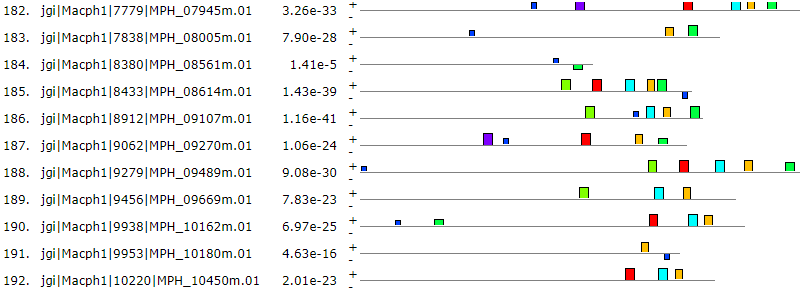
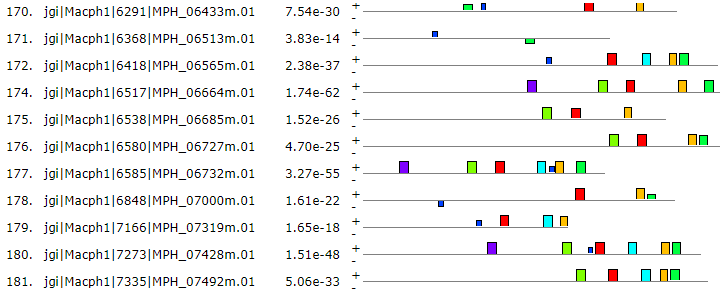
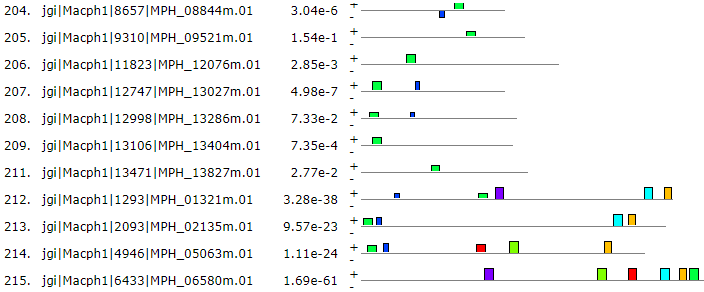
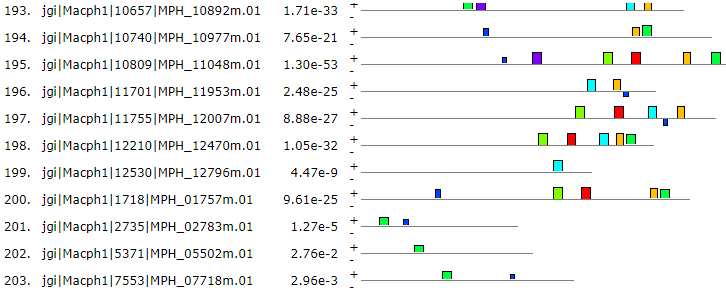
**Figure 2:** **Gene structure as obtained using the gene structure display server (GSDS**).

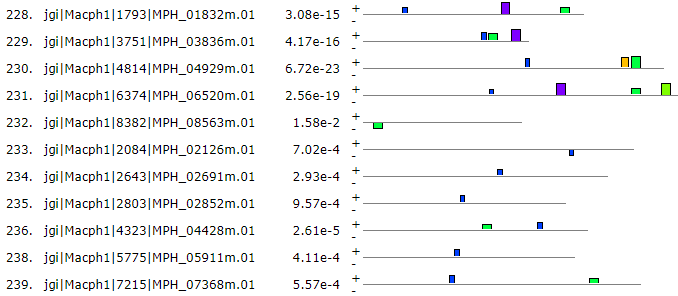
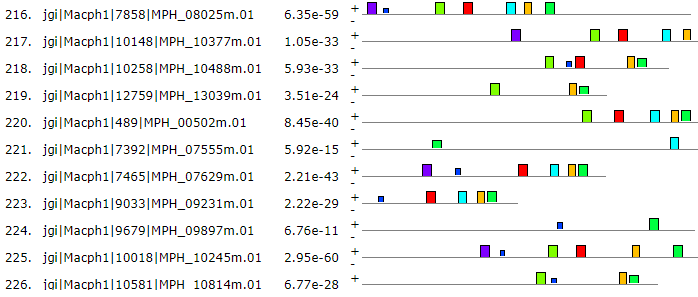
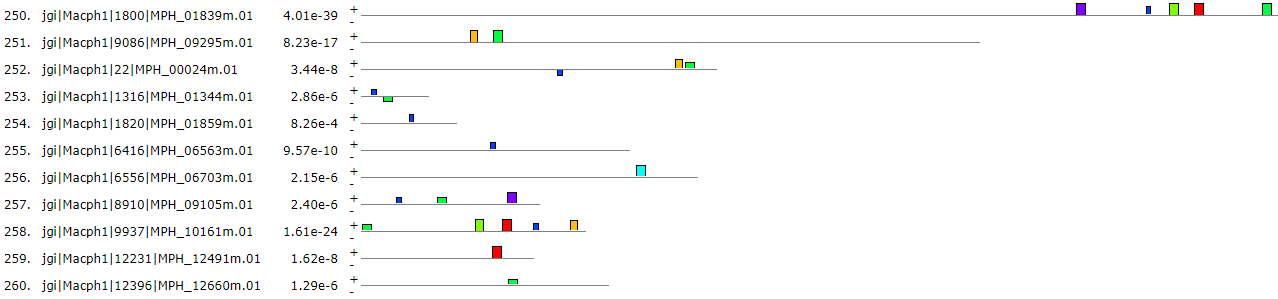
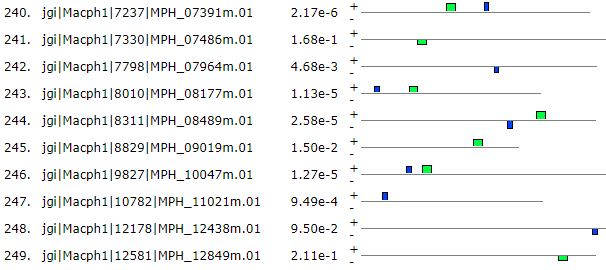


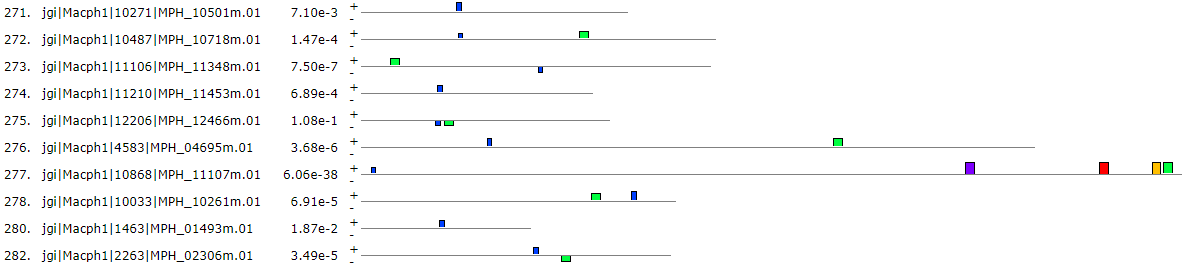
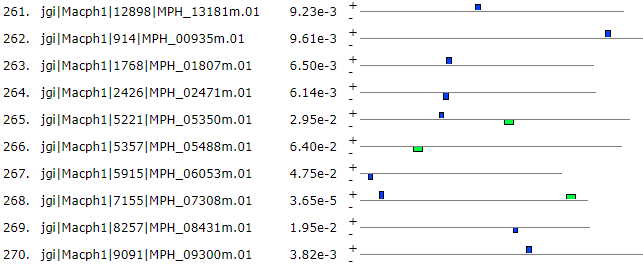
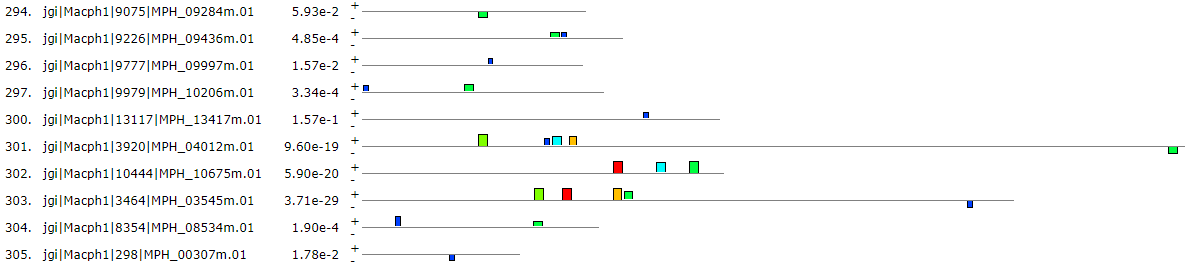
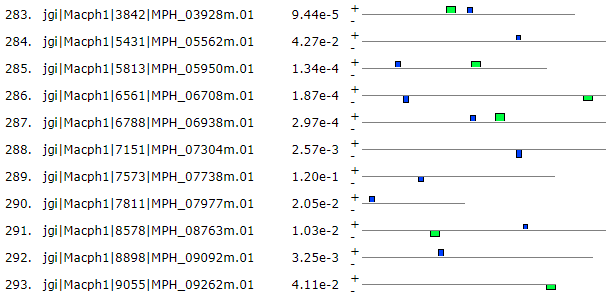
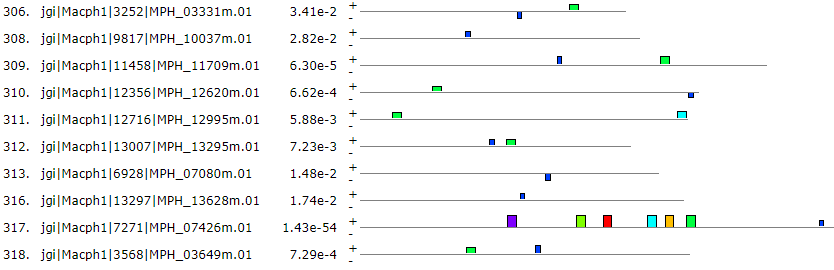


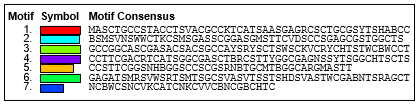




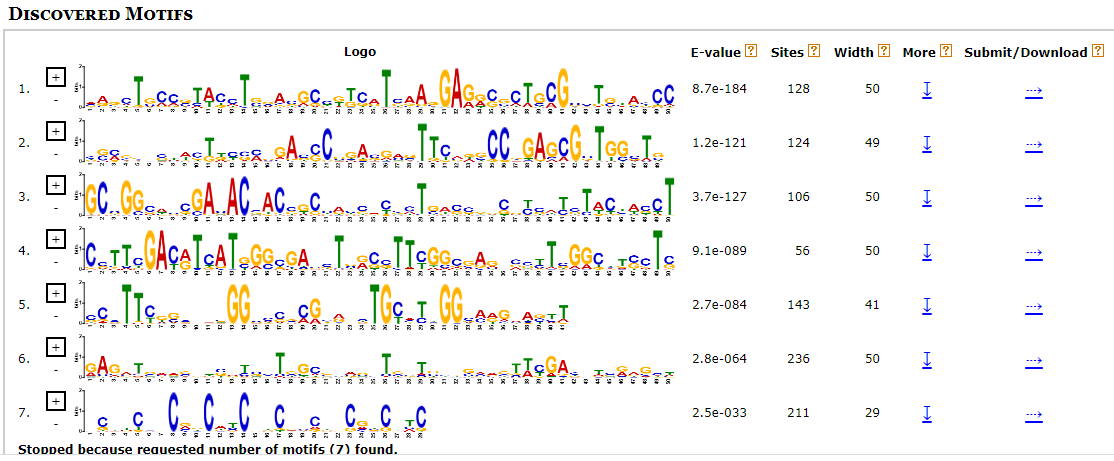
   



**Figure 3b: Predicted conserved motifs of cytochrome P450 genes of *Macrophomina phaseolina***



**Figure 3a: Sequence Logo of the Seven Motifs predicted in the cytochrome P450 of *Macrophomina phaseolina.* The amino acid type and position are shown on the x-axis. The overall height of the amino acid stacks, plotted on the y-axis, indicates the sequence conservation at a given position, while the height of individual symbols within a stack indicates the relative frequency of a nucleotide base at that position.**

**Table 1: Identified clans, families and putative functions of cytochrome P450 in *Macrophomina phaseolina***

|  |  |  |  |
| --- | --- | --- | --- |
| Clan | Family | Putative Function | Number of P450s |
| CYP533 | CYP64 | Xenobiotic Metabolism | 16 |
| CYP65 | CYP60 | Secondary Metabolism | 27 |
| CYP52 | CYP52 | Xenobiotic Metabolism, Alkane inducible P450 | 5 |
|  | CYP120 | Oxidation of fatty acids | 2 |

**Figure 4:** **Prediction of** **subcellular localization of Cytochrome P450 in *Macrophomina phaseolina*. Bars represent standard error at a 5 % probability level.**