Victoria Dixon

ABT780 SPR25 SEC01

2025-02-16

*A Simple Phylogenetic Analysis of a Tomato Detoxification Gene*

**Introduction:**

Tomato (*Solanum lycopersicum*) is an important global crop that is often exposed to pesticides during production. To manage these chemical stresses, tomato plants rely on detoxification enzymes—especially those from the cytochrome P450 superfamily—to break down and neutralize harmful compounds (Hamberger & Bak, 2013; Schuler & Werck-Reichhart, 2003). These enzymes are a diverse group of heme-thiolate proteins that perform monooxygenation reactions, which increase the water solubility of hydrophobic compounds, making them easier to detoxify (Mizutani & Ohta, 2010).

This diversity results from extensive gene duplication and sequence divergence, allowing plants to develop specialized enzymes that respond to various xenobiotics, including pesticides (Hamberger & Bak, 2013). Variations in these detoxification genes among different tomato cultivars and related Solanaceae species may lead to differences in pesticide tolerance by altering conserved motifs or catalytic efficiency (Li et al., 2010). By analyzing the evolutionary relationships among these genes through phylogenetic methods, we can identify patterns of gene duplication, divergence, and conservation that contribute to enhanced detoxification (Tang et al., 2024).

Understanding these evolutionary patterns is essential for breeders working to develop tomato varieties with improved resilience to chemical stress. This research will focus on one specific cytochrome P450 gene from tomato, comparing its sequences across several cultivars and closely related species, building on studies by Vasav et al. (2019), Li et al. (2010), and Hamberger & Bak (2013).

**Research Objectives:**

To create a sequence collection by retrieving nucleotide or protein sequences for a selected tomato cytochrome P450 detoxification gene from various tomato cultivars and related species. Perform multiple sequence alignments to identify conserved regions and sequence variations among these sequences. Conduct phylogenetic analysis by constructing a phylogenetic tree to visualize evolutionary relationships and detect potential gene duplication events. Interpret the data by assessing whether the observed sequence differences correlate with variations in pesticide detoxification capacity among tomato cultivars.

This investigation will provide a foundational analysis on how sequence variation in a key detoxification gene may contribute to pesticide tolerance in tomatoes, with potential applications in crop improvement.

**Programming Languages and Code Description:**

My research aims to build a comprehensive collection of sequences by retrieving nucleotide or protein data for a selected tomato cytochrome P450 detoxification gene from various tomato cultivars and related species. I will perform multiple sequence alignments on these sequences to identify conserved regions and variations that may correlate with differences in pesticide detoxification capacity. Next, I will construct a phylogenetic tree to visualize evolutionary relationships and detect potential gene duplication events, providing insights into the evolutionary dynamics underlying differential pesticide tolerance.

To achieve these objectives, I will primarily use R for data analysis, sequence alignment, phylogenetic tree construction, and visualization, with Python as an optional tool for initial data handling and preprocessing.

The workflow begins with sequence retrieval, where I will programmatically download the target cytochrome P450 gene sequences from public databases (e.g., NCBI, Ensembl Plants) using the R package *rentrez*. The retrieved sequences will then be aligned using the R package *msa* (or external tools such as MAFFT or ClustalW via system calls), and the resulting alignment will be saved for further analysis. For phylogenetic tree construction, I will employ the R packages *ape* and/or *phangorn*, using methods like Neighbor-Joining (NJ) or Maximum Likelihood (ML) with optional bootstrap analysis to assess branch support and annotate the tree with relevant sequence features. Finally, I will generate annotated tree plots using *ggtree* or base R plotting functions to clearly present the key findings.

**Expected Outcome:**

This integrated program will automate the retrieval, alignment, and phylogenetic analysis of tomato detoxification gene sequences, creating a clear, annotated phylogenetic tree that illustrates the relationship among the selected detoxification gene sequences. Thereby, it will address the research question by revealing the evolutionary dynamics that may underlie differential pesticide tolerance.**Bottom of Form**

**References:**

Li, Z., Hao, Y., Yang, Y. *et al.* Molecular cloning and expression analysis of a cytochrome P450 gene in tomato. *Plant Growth Regul* **61**, 297–304 (2010). https://doi.org/10.1007/s10725-010-9477-6

Vasav AP, Barvkar VT. Phylogenomic analysis of cytochrome P450 multigene family and their differential expression analysis in Solanum lycopersicum L. suggested tissue specific promoters. BMC Genomics. 2019 Feb 7;20(1):116. doi: 10.1186/s12864-019-5483-x. PMID: 30732561; PMCID: PMC6367802.

Tang M, Zhang W, Lin R, Li L, He L, Yu J, Zhou Y. Genome-wide characterization of cytochrome P450 genes reveals the potential roles in fruit ripening and response to cold stress in tomato. Physiol Plant. 2024 May-Jun;176(3):e14332. doi: 10.1111/ppl.14332. PMID: 38710502.

Hamberger, B. & Bak, S. (2013) Plant P450s as versatile drivers for evolution of species specific chemical diversity. *Philosophical Transactions of The Royal Society B Biological Sciences*, **368**, 20120426.

Schuler, M. A., & Werck-Reichhart, D. (2003). Functional genomics of plant P450s. *Annual Review of Plant Biology, 54*, 629–667.

Nelson, D. and Werck-Reichhart, D. (2011), A P450-centric view of plant evolution. The Plant Journal, 66: 194-211. https://doi.org/10.1111/j.1365-313X.2011.04529.x

Mizutani, M., & Ohta, D. (2010). Diversification of P450 genes during land plant evolution. *Annual Review of Plant Biology, 61*, 291–315.