Chapter 1

Introduction

1.1 Background

Abrus precatorius is a plant known to produce valuable triterpenoid saponins such as abrusoside and glycyrrhizin, which supports their use in herbal medicine and potential as natural, low-calorie sweeteners. These compounds contain glycosyl (sugar) moieties that contribute to their bioactivity and stability, as well as their sweet taste which can be up to 150 times sweeter than sucrose (Dwivedi, 2022). Their unique properties make them attractive in the food and pharmaceutical industries as alternatives to synthetic sweeteners for both food and medications (Zhang et al., 2017). Beyond their commercial value, these secondary metabolites produced in the leaf and roots of *A. precatorius*, may also play roles in the plant's stress response (Okhale & Nwanosike, 2016).

The final step in the biosynthesis of these two triterpenoid saponins involve glycosylation, a reaction catalyzed by glycosyltransferase enzymes (GTs). This enzyme transfers sugar moieties to specific substrates, which affects its solubility, stability and biological activity (Yu et al., 2018). GTs have important roles in industrial applications, such as in modifying starch branching for improved food texture and solubility, and in improving the characteristics of these triterpenoid saponins. However, *A. precatorius* is genetically underexplored, making its important GT genes poorly characterized.

In this study, the effects of hydroponic cultivation and the plant phytohormone, methyl jasmonate (MeJA), as plant stressors were studied. Hydroponic systems, where the plant roots are submerged in water, represents a form of abiotic stress for soil-adapted plants as the aqueous environment causes oxidative stress where the roots do not get sufficient oxygen (Oetomo, 2024). On the other hand,

MeJA acts as a signaling molecule that mimics biotic stress, such as herbivore attacks or mechanical injury, and is widely used to trigger the expression of genes related to secondary metabolite synthesis (Yu et al., 2018). The combination of hydroponic and MeJA treatment may result in unique or amplified transcriptional responses compared to each treatment alone which creates a valuable setup to observe stress-responsive genes (Lowe et al., 2017). The transcriptomic data of these *A. precatorius* samples provided by Istiandari et al. (in press) resulted in the identification of candidate differentially expressed GT genes potentially involved in secondary metabolite synthesis.

Transcriptomics data such as the one provided by Istiandari et al. (in press), carry high risk of false positive results which causes the common employment of qPCR to corroborate in silico findings (Benjamini and Hochberg, 1995; Li et al., 2022; Kernfeld et al., 2024). In response to the need for efficient validation of transcriptomic findings, this study introduces a practical pipeline designed to guide the validation of gene expression from transcriptomic data through semi-quantitative or quantitative Polymerase Chain Reaction (qPCR) analysis of a specific gene of interest. The pipeline incorporates in silico analysis to effectively narrow down a strong candidate for subsequent laboratory validation from an exhaustive list of candidate genes available in the transcriptomic data. Instead of relying solely on qPCR, a more common and precise quantification method that can be costly and labor-intensive, the proposed workflow starts with the independent identification of differentially expressed genes (DEGs) from transcriptome datasets, followed by validation using semi-quantitative PCR. This method offers a more accessible and cost-effective alternative, suitable for narrowing down candidates before moving on to more sensitive quantification (Ferre, 1992). By cross-referencing previous analyses and validating gene relevance under stress conditions, the pipeline supports the confirmation of transcriptomic insights while contributing to our understanding of GT gene expression and molecular responses in underexplored plant species.

1.2 Objective

This study aims to validate transcriptome-derived candidate glycosyltransferase (GT) genes in *A. precatorius* that are potentially involved in the plant's stress response, using a combination of bioinformatics screening and semi-quantitative PCR analysis. Specifically,

- To identify GT gene candidates from *A. precatorius* transcriptome data that show differential expression under stress treatment.
- To narrow down the strong gene candidate for glycosyltransferase in *A. precatorius* using bioinformatic tools including motif analysis, phylogenetic inference, and expression clustering to find the potential glycosyltransferase candidate playing role in abrusoside or glychyrrhizin biosynthesis.
- To validate transcriptomic expression of the chosen glycosyltransferase gene candidate results experimentally through semi-quantitative PCR analysis.

1.3 Hypothesis

For identification of candidate genes from transcriptome data:

- Null Hypothesis (H_0): Glycosyltransferase (GT) candidate genes with differential expression under stress conditions cannot be identified from the *A. precatorius* transcriptome data and characterized using an in silico study.
- Alternative Hypothesis (H_1): Glycosyltransferase (GT) candidate genes with differential expression under stress conditions can be identified from the *A. precatorius* transcriptome data and characterized using an in silico study.

For gene expression analysis:

- Null Hypothesis (H_0): Semi-quantitative PCR does not support the expression trends of GT candidate genes predicted by the transcriptome analysis under stress conditions.
- Alternative Hypothesis (H_1): Semi-quantitative PCR supports the expression trends of GT candidate genes predicted by the transcriptome analysis under stress conditions.