DECIPHERING BIOSYNTHETIC MACHINERY OF PODOPHYLLOTOXIN – ANTICANCER METABOLITE OF *PODOPHYLLUM HEXANDRUM* ROYLE

Thesis submitted in fulfillment for the requirement of the Degree of

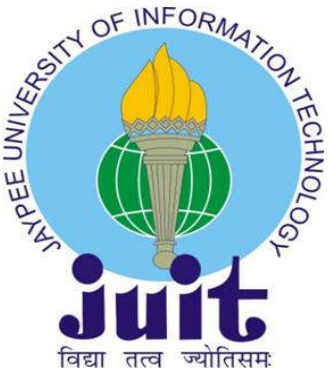
DOCTOR OF PHILOSOPHY

IN

BIOTECHNOLOGY

By

PAWAN KUMAR



Department of Biotechnology and Bioinformatics

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY

WAKNAGHAT, DISTRICT SOLAN, H.P., INDIA

DECEMBER 2017

DECLARATION BY THE SCHOLAR

I hereby declare that the work reported in the Ph.D. thesis entitled **"Deciphering Biosynthetic Machinery of Podophyllotoxin – Anticancer Metabolite of *Podophyllum hexandrum* Royle"** submitted at **Jaypee University of Information Technology, Waknaghat, India**, is an authentic record of my work carried out under the supervision of **Prof. (Dr.) Rajinder Singh Chauhan (Supervisor)** and **Dr. Hemant Sood (Co-Supervisor)**. I have not submitted this work elsewhere for any other degree or diploma. I am fully responsible for the contents of my Ph.D. Thesis.

(Pawan Kumar)

Department of Biotechnology and Bioinformatics

Jaypee University of Information Technology, Waknaghat, India

Date: 21/12/17

SUPERVISOR'S CERTIFICATE

This is to certify that the work reported in the Ph.D. thesis entitled "**Deciphering Biosynthetic Machinery of Podophyllotoxin – Anticancer Metabolite of *Podophyllum hexandrum* Royle**", submitted by **Pawan Kumar** at **Jaypee University of Information Technology, Waknaghat, India,** is a bonafide record of his original work carried out under my supervision. This work has not been submitted elsewhere for any other degree or diploma.

21/12/17

Prof. (Dr.) Rajinder Singh Chauhan

(Supervisor)

Former Dean (Biotechnology) and Head

Department of Biotechnology & Bioinformatics

Jaypee University of Information Technology,

Waknaghat, India

Date:

21/12/17

Dr. Hemant Sood

(Co-Supervisor)

Assistant Professor (Senior Grade)

Department of Biotechnology & Bioinformatics

Jaypee University of Information Technology,

Waknaghat, India

Date:

ACKNOWLEDGEMENT

Here is the acknowledgement page of my PhD thesis. Well, the list of the people I need to thank will not fit to a single Acknowledgement section. I just mention some people whose contribution is obvious. I owe my gratitude to all those people who have made this dissertation possible and because of whom my graduate experience has been one that I will cherish forever.

Completing the PhD degree was probably the most challenging activity of my life. It has been a great privilege to spend several years in the Department of Biotechnology and Bioinformatics at Jaypee University of Information Technology, Waknaghat, Solan, India, and its members will always remain dear to me.

*My first debt of gratitude must go to my supervisor**, Dr. Rajinder Singh Chauhan**. He patiently provided the vision, encouragement and advice necessary for me to proceed through the doctorial program and complete my dissertation. I want to thank **Dr. Chauhan** for his unflagging encouragement and serving as a role model to me. He has been a strong and supportive adviser to me throughout my doctorial career.*

*Special thanks to my co-supervisor, **Dr. (Mrs.) Hemant Sood**, for her support, guidance, motivation and helpful suggestions. She encouraged me throughout my PhD work and her guidance has served me well and I owe her my heartfelt appreciation.*

*I emphatically express my loyal and venerable thanks to JUIT administration, **Prof. (Dr.) Vinod Kumar (Vice Chancellor), Prof. (Dr.) S. C. Saxena (Former Vice Chancellor), JUIT; Brig. (Retd.) Balbir Singh, (Ex- Director) JUIT; Prof. (Dr.) Samir Dev Gupta (Director & Academic Head), Maj Gen Rakesh Bassi (Retd.) (Registrar and Dean of Students), Brig. (Retd.) K.K. Marwah (Ex- Director) JUIT; Prof. (Dr.) T. S. Lamba (Ex-Dean) (Academic & Research)** for providing opportunity to pursue a Doctorate Degree, teaching assistantship and advanced lab infrastructure to accomplish this scientific venture of my life.*

*I am thankful to the **Department of Biotechnology, Ministry of Science and Technology, Government of India, New Delhi** for providing financial support. I thank **Dr. Jagdish Singh** (Scientist), Himalayan Forest Research Institute, Shimla for providing valuable plant material.*

I am also thankful to all the members of technical and non-technical staff of the department, especially ***Mrs. Somlata Sharma****,* ***Mrs. Mamta Mishra****,* ***Mr. Baleshwar Shukla****,* ***Mr. Kamlesh Kumar and Mahinder Ji*** *for their assistance and valuable contributions.*

I am fortunate to have friends who have always stood beside me. I extend my heartfelt thanks to ***Kirti****,* ***Nikhil, Archit****,* ***Jibesh****,* ***Tarun****,* ***Ira****,* ***Neha****,* ***Pooja, Varun****,* ***Sampan****,* ***Prashant, Ashwani****,* ***Ankush, Shivam, Dr. Varun Jaiswal and Dr. Sree Krishna Chanumolu*** *for their sustained support and ever needed cooperation. It is my pleasure to express my gratitude to all research scholars of the Biotechnology & Bioinformatics Department for keeping me blessed with best wishes.*

I wish to thank my parents, ***Shri Hukam Chand*** *and* ***Mrs. Prem Lata****. Their love, support and motivation provided me inspiration and encouragement and was my driving force to successfully complete my work. I owe them everything and wish I could show them just how much I love and appreciate them.*

I would like to express my heartfelt gratitude to all those who have contributed directly or indirectly towards obtaining my doctorate degree and apologize if have missed out anyone. Last, but not the least, I thanks the one above all of us, omnipresent God, for answering my prayers, for giving me the strength to plod on during each and every phase of my life.

All may not be mentioned, but no one is forgotten.

Pawan Kumar

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

DAHPS	3-Deoxy-*D*-arabinoheptulosonate 7-phosphate synthase
4CL	4-Coumarate-CoA ligase
EPSPS	5-enolpyruvylshikimic acid-3-phosphate synthase
ADT	Arogenate/prephenate dehydratase
PPA-AT	Aromatic-amino-acid transaminase
CAM	Caffeic acid-3-*o*-methyltransferase
COMT	Caffeoyl-CoA-*O*-methyltransferase
CS	Chorismate synthase
CM	Chorosmate mutase
C4H	Cinnamate-4-hydroxylase
CCR	Cinnamoyl CoA reductase
CAD	Cinnamyl alcohol dehydrogenase
COG	Clusters of Orthologous Groups of proteins
DHQD	Dehydroquinate dehydratase
DHQS	Dehydroquinate synthase
SD	Dehydroshikimate dehydratase
E4P	Erythrose-4-phosphate
FPKM	Fragments Per Kilobase of Exon Per Million Fragments Mapped
GO	Gene ontology
HPLC	High-performance liquid chromatography
KAAS	KEGG Automatic Annotation Server

C3H	*p*-Coumarate 3-hydroxylase
PAL	Phenylalanine ammonia lyase
PA4H	Phenylalanine-4-hydroxylase
PEP	Phosphoenolpyruate
PLR	Pinoresinol lariciresinol reductase
PMRD	Plant microRNA Database
PS	Pluviatolide synthase
PD	Prephenate dehydrogenase
qRT-PCR	Quantitative real-time polymerase chain reaction
QSD	Quinate/shikimate dehydrogenase
SRD	Secoisolariciresinol dehydrogenase
SK	Shikimate kinase

ABSTRACT

Podophyllum hexandrum Royle (*P. hexandrum*) is known for its vast medicinal properties particularly, anticancer. It contains higher amount of podophyllotoxin (4.3%), compared to *P. peltatum* (0.025%) and other plant species, as a result has been used worldwide in the preparation of various drugs including, anticancer, antimalarial, antiviral, antioxidant, antifungal, and so on. Currently, Etoposide (VP-16-213), VUMON® (Teniposide; VM-26), Etopophos®, Pod-Ben- 25, Condofil, Verrusol and Warticon are available in the market. Podophyllotoxin is commercially isolated from roots and rhizomes of *Podophyllum* species. However, limited availability and finite amount of raw material, highly complex synthesis, and low cell culture yields of podophyllotoxin (0.3%), necessitate alternate sources/production routes for meeting increasing demands of podophyllotoxin. Further increase in podophyllotoxin content can only be achieved through genetic interventions in podophyllotoxin biosynthetic pathway. However, the knowledge on podophyllotoxin biosynthetic and regulatory machinery (pathway genes, transcription factors, ABC transporters and miRNAs) vis-à-vis podophyllotoxin content is lacking. Therefore, the current study was focused on i) identification of pathway components/genes contributing to podophyllotoxin biosynthesis ii) determination of crucial regulatory components such as transcription factors, ABC transporters and miRNAs implicated in podophyllotoxin biosynthesis in *Podophyllum hexandrum.*

The present work utilized the available transcriptomes resources of *Podophyllum hexandrum* to identify, characterize and validate the molecular components associated with the podophyllotoxin biosynthesis. The work led to the identification of candidate pathway genes *COMT, CM, PD, ADT, C4H, C3H, CCR, CAD* and *SRD* that showed increase in transcript abundance up to 1.4 to 23.05 folds, respectively vis-à-vis podophyllotoxin content in roots (1.37%) and rhizomes (3.05%) of *P. hexandrum.* Comparative transcriptome analysis of *Podophyllum* species (*Podophyllum hexandrum* and *Podophyllum peltatum*), quantification of podophyllotoxin and comparative expression analysis between high (2.51%) versus low (0.59%) podophyllotoxin content accessions revealed key transcription factors (bZIP, MYB, bHLH and WRKY) associated with the podophyllotoxin biosynthesis. bZIP TF showed the highest transcript abundance (19.60-folds) in *P. hexandrum* rhizomes (2.51%) compared to shoots (0.01%). Total 252 ABC transporter genes were identified as unigenes out of which 22 were further validated using qRT-PCR in different tissues of *P. hexandrum* varying for

podophyllotoxin content. Differential expression analysis and Pearson's correlation coefficient revealed two candidate genes, *PhABC6* and *PhABCIII* having positive correlation with the podophyllotoxin content. *PhABCIV* showed highest expression (20.53-folds) in rhizomes compared to shoots suggesting its possible role in the transport and accumulation of podophyllotoxin. Six potential miRNAs were identified in transcriptomes of *P. hexandrum.* Eight potential mRNA targets were identified for these six miRNAs in *P. hexandrum* transcriptomes that are possibly involved in the regulation of podophyllotoxin biosynthesis. The identified miRNAs and their mRNA targets (UDP glycosyltransferase, flavnol synthase, glyceraldehyde 3-phosphate dehydrogenase, peroxidise, malate dehydrogenase, phosphoenolpyruvate carboxylase, WRKY and MYB transcription factors) were validated through qRT-PCR . miRNAs vun-miR396b, mtr-miR2673a, vvi-miR828b and peu-miR2910 established negative correlation with their targets thereby, suggesting downregulation of mRNA targets linked to podophyllotoxin biosynthesis.

The current work led to the identification of crucial molecular components (pathway genes, key transcription factor, ABC transporters and miRNAs) involved in the biosynthetic machinery of podophyllotoxin. These molecular components, after their functional validation, can be utilized in designing a suitable genetic intervention strategy to enhance podophyllotoxin content in *Podophyllum hexandrum.*

CHAPTER 1

INTRODUCTION

Secondary metabolites are "natural products", mainly produced to aid in survival of plants under stress conditions. Secondary metabolites can be divided into groups consisting of 29,000 terpenoid, 12,000 alkaloid derivatives and 8,000 phenolics produced through different biosynthetic routes. Secondary metabolites are used as food additives, drugs, flavors, fragrances, dyes, colours, pesticides, pharmaceuticals, agrochemicals, biopesticides and more recently as nutraceuticals.

The North-Western Himalayan regions are a rich resource of biodiversity having 1748 medicinal and aromatic plant species [1]. The medicinal plants produce a wide array of phytochemicals/metabolites having high therapeutic value to mankind. Secondary metabolites like podophyllotoxin, vinblastine/vincristine, taxol (paclitaxel), artemisinin, shikonin, quinine, picrosides, hypericum, ajmalicine, serpentine, campothecin, etc. are being used by the industries for their large scale production because of their anticancer, antimalarial, hepatoprotective, anti-HIV, anti-pyretic, antifungal, anti-inflammatory properties and so on. These metabolites have high demand in international markets, but their supply is limited because of complexity in *in-vitro* synthesis, high cost of production and decreasing availability of raw material of medicinal plants. Another reason is that the exact biosynthetic pathways of these metabolites are not known completely or the regulatory steps/genes controlling the metabolites biosynthesis are not known clearly so as to plan genetic interventions for enhanced metabolite yields.

The higher cost of secondary metabolites production and geographical limitation of medicinal plants, has forced researchers to find out alternative methods or sources for the production of these metabolites. Various research groups have been working throughout the world towards the enhancement of secondary metabolite production in *in-vitro* cultures but only limited success has been achieved. Therefore, there exists an urgent need to not only decipher biosynthetic pathways but also to identify key points in the biosynthesis, transport, and accumulation of metabolites.

Podophyllum hexandrum Royle (*P. hexandrum*) is an important medicinal plant species native to alpine and sub-alpine regions of the North-Western Himalayas, North America and South West of China (2500-4200m) and commonly used as a source of podophyllotoxin and its derivatives- the anticancer bioactive compounds. Podophyllotoxin is present in different plant species *Podophyllum hexandrum* Royle (*syn. P. emodi* Wall), *Podophyllum peltatum, Linum album, Callitris drummondii, Juniperous chinensis, Polygala polygama*, etc., however *P. hexandrum* has been used as a primary source due to higher content (4.3%) [2,3]. Several anticancer drugs such as Etoposide (VP-16-213), VUMON® (Teniposide; VM-26), Etopophos®, Pod-Ben-25, Condofil, Verrusol and Warticon prepared from podophyllotoxin and its derivatives are available in the market. Many other drugs, GL331, Top 53, NK 611 and CPH 82 are in phase III clinical trials containing podophyllotoxin and its derivatives [4]. Owing to these medicinal properties, *P. hexandrum* has been overexploited and therefore, placed under the category of critically endangered plant species [5].

Podophyllotoxin is commercially isolated from roots and rhizomes of *Podophyllum* species. However, limited availability and finite amount of raw material necessitate alternate sources/production routes for meeting increasing demands of podophyllotoxin. Different cell cultures interventions have been tried for the production of podophyllotoxin such as elicitor's treatments, precursor feeding, modification in production media, immobilization of cells and scale-up through bioreactors, etc. [6]. However, podophyllotoxin yields have been low (0.3%) in callus/cell suspension cultures as compared to roots and rhizomes (4.3%) of plants grown in natural habitat [7, 8]. Further increase in podophyllotoxin content can only be achieved through genetic interventions in podophyllotoxin biosynthetic machinery. However, knowledge of podophyllotoxin biosynthetic pathway is fragmentary with missing links in the pathway as well as lack of relative contribution of pathway steps in contributing to the podophyllotoxin content.

Secondary metabolites are formed in different tissues/organs through the combinatorial role of multiple pathways (Mevalonic acid (MVA) pathway, non-mevalonate (MEP), phenylpropanoid/shikimic acid and iridoid pathways) in plant species. Podophyllotoxin and its derivatives are derived from the shikimic acid/phenylpropanoid pathway which has not yet been completely elucidated [1, 9]. Thirty one steps are known in the podophyllotoxin biosynthetic pathway up to the intermediate compound, pluviatolide; thereafter possible 36 steps were predicted but not validated either for the structures of intermediates or for genes/proteins/enzymes [10]. Total 26 genes encoding podophyllotoxin pathway enzymes are

known to catalyse 31 known steps, starting from *D*-erythrose-4-phosphate to intermediate compound pluviatolide. Out of 26 genes, only 12 genes *PAL, C4H, PCH, P4H, HCT, 4CL, CMT, CCMT, CCR, CAD, DPO, PLR* and *SD* have been characterized in *Podophyllum* spp. to correlate their role in podophyllotoxin biosynthesis through expression analysis [11, 12, 13]. Most of these 12 genes represent enzymatic steps, starting from phenylalanine to matairesinol. However, no information exists on remaining 14 genes, starting from *D*-erythrose-4-phosphate to cinnamic acid. The amount of *p*-coumaric acid biosynthesized is crucial in providing supply of further intermediates such as caffeic acid, ferulic acid, coniferyl aldehyde, coniferyl alcohol and so on in the pathway that eventually lead to the formation of podophyllotoxin. The crucial role of *p*-coumaric acid has been highlighted in the biosynthesis of secondary metabolites such as caffeic acid, *p*-hydroxybenzoic acid, etc. [14, 15]. Therefore, identification of genes contributing towards the production of *p*-coumaric acid is important in furthering the understanding of podophyllotoxin biosynthetic machinery by completing the biosynthetic pathway starting from the first compound *D*-erythrose-4-phosphate to the last known intermediate compound, pluviatolide.

No information exists as of today on transcription factors (TFs), ABC transporters and miRNAs controlling the biosynthesis and accumulation of podophyllotoxin content in *P. hexandrum*. Therefore, studies need to be undertaken to elucidate or decipher important molecular components controlling podophyllotoxin content in *P. hexandrum* so that appropriate stretigies for metabolic engineering can be undertaken to fulfil the escalating demand of podophyllotoxin.

The availability of genomics, transcriptomics and metabolomics resources can provide a powerful tool for identification, validation, characterization and functional analyses of molecular components in order to infer their role in regulation of metabolic processes in plant species. The assembled transcriptomes data of *Podophyllum* species (*P. hexandrum*—2, 27, 885 contigs, *P. peltatum*—1, 47, 960 contigs) were downloaded from the medicinal plants transcriptomics database (http://medplants.ncgr.org/) [9] and the transcriptomes were mined and analyzed for identification, characterisation and classification of molecular components by using in-house developed PERL scripts. This resource has substantially broadened our understanding of podophyllotoxin biosynthesis in *P. hexandrum*.

The current research work focussed on identification, characterisation, validation and expression analysis of molecular components such as biosynthetic pathway genes, transcription factors, ABC transporters and miRNAs to deduce the biosynthetic machinery of podophyllotoxin in *P. hexandrum.* To understand the molecular basis of podophyllotoxin biosynthesis the expression status of entire biosynthetic pathway genes was carried out in different tissues/organs of *P. hexandrum* varying for podophyllotoxin content. Further, the identified key regulatory components i.e. transcription factors, ABC transporters and miRNAs possibly associated with podophyllotoxin biosynthesis through extensive transcriptome analysis followed by wet lab experiments.

Keeping in view the immense medicinal importance of *P. hexandrum*, and lack of knowledge on molecular basis of podophyllotoxin biosynthetic machinery, the present work was carried with the following objectives:

Objective 1: Identification of pathway components/genes contributing to podophyllotoxin biosynthesis in *Podophyllum hexandrum*

Objective 2: Dissecting regulatory machinery/components of podophyllotoxin biosynthesis and accumulation in *Podophyllum hexandrum*

CHAPTER 2

REVIEW OF LITERATURE

2.1 Introduction

Podophyllum is a genus belonging to the family Berberidaceae, and represented by nine species of perennials from North America and parts of Asia. The term *Podophyllum* has been derived from the ancient Greek word "Foot leaf" and is also known as Mayapple, because fruits of these plants ripen at the time of spring. The two most representative and important species of this genus are *Podophyllum hexandrum* also known as the Himalayan Mayapple or the Indian Mayapple and *Podophyllum peltatum*, generally referred to as American Mayapple or American Mandrake (Figure 2.1). *Podophyllum hexandrum* Royle is a native of the lower elevation of North-Western Himalayas. Flowering in this plant occurs between May-August, the flowers have six petals and six stamens, which has inspired the name of its species-hexandrum, meaning six stamens. *P. peltatum* is a relative of *P. hexandrum*, and is commonly found in the deciduous forests of eastern North America. Extracts obtained from *Podophyllum* species have found usage as antidotes against poison, treatments for skin disorders or as cathartic, cholagogue, purgative, anti-helminthic, vesicant, and suicide agents [16].

Podophyllotoxin is a secondary metabolite naturally occurring in plants belonging to families Berberidaceae, Cupressaceace, Linaceae, Polygalaceae and Oleaceae such as *Podophyllum hexandrum, P. peltatum, Linum album, Callitris drummondii, Juniperous chinensis, Polygala polygama* and so on (Figure 2.2). These compounds are aryltetralin lignan type derivatives having various useful properties– anti-cancer, anti-viral, antioxidant, anti-fungal, anti-mitotic, immunostimumaltory [17, 18, 19]. Till date, podophyllotoxin has been used as the starting compound for the production of semi-synthetic anti-cancer drugs etoposide (VP-16-213), teniposide (VM-26) and ethophos, Pod-Ben- 25, Condofil, Verrusol, Warticon etc. which are used in the treatment of lung and testicular cancer, leukaemia and rheumatoid arthritis (Table 2.1). Several drugs which are under phase III clinical trials such as GL331 and Top 53 are also prepared from podophyllotoxin [20, 21, 22]. Indian *Podophyllum hexandrum* Royle has high level of podophyllotoxin content (4.3%) as compared to other podophyllotoxin

producing medicinal plants such as North American *Podophyllum peltatum* (0.25%)*, Linum album* (0.35%), *Juniperus chinensis* (0.0025%), *Linum nodiflorum* (1.7%) and *Callitris drummondii* (1.4%) [2, 3, 23, 24, 25]. Owing to these properties, genus *Podophyllum* is over-exploited and therefore, placed under the category of critically endangered species [5].

A scientific group of IIT Delhi, India under the guidance of Dr. Ashok Kumar Srivastava had worked on the *Podophyllum* species for almost 12 years towards production of secondary metabolites (podophyllotoxin, azadirachtin) through plant cell /hairy root culture, fermentation technology. Various physico-chemcial parameters/factors were studied by this group to enhance the production of podophyllotoxin including *in-vitro* production and its scale-up in bioreactor by optimizing various growth factors like nutrients levels, oxygen supplementation, inoculum density, low shear stress to cells, carbon source, elicitors/precursors, light, pH and substrate feeding rate [26, 27, 28, 29, 30]. A number of R&D groups belongs to CSIR-IIIM, CSIR-IICB, CSIR-IHBT, IIT-Delhi, Institute of Nuclear Medicine & Allied Sciences (INMAS), G.B. Pant institute of Himalayan Environment and Development Uttarakhand, Bharathidasan University, Tamil Nadu, etc., across India are working on *Podophyllum hexandrum* with the ultimate goal of *in-vitro* production, conservation, genetic variations, micropropagation, radio preventive studies and so on. Similarly, Dr. Sharmila Chattopadhyay's groups of CSIR- Indian Institute of Chemical Biology (IICB), Kolkata has worked on *in-vitro* regeneration, metabolite elicitation and expression analysis of proteins related to pathway of podophyllotoxin [31, 6].

In spite of above mention work done, several gaps still exist in complete understanding of biosynthetic pathway for podophyllotoxin. Cloning and expression of genes involved in podophyllotoxin biosynthetic pathway like *secoisolariciresinol dehydrogenase* (SDH), *Phenylalanine ammonia-lyase* (PAL), *Dirigent protein oxidase* (DPO), *Cinnamoyl-CoA reductase* (CCR), *Cinnamyl-alcohol dehydrogenase* (CAD), *Pinoresinol-lariciresinol reductase gene* (PLR), *Cinnamic acid 4-hydroxylase* (C4H) has been reported [11, 12, 13].

There are no reports published on the differential biosynthesis, accumulation, spatio-temporal behaviour, expression analysis of potential genes/proteins, and elucidation of pathway of podophyllotoxin. One report published by Marques et al. [10], showed possible steps in the pathway of podophyllotoxin with the identification of cytochrome P450s enzymes (CYP719A24, CYP719A23). They used parallel sequencing technologies to understand the biosynthetic processes in poorly understood (non-model) medicinal plants. One main

advantage is that such technologies can potentially lower the time frame for discovery of new genes and thus more rapidly improve our understanding of metabolism. Secondary metabolite production can be improved by a proper understanding of biosynthetic pathway thereby; ultimately goal of enhancement in the production of valuable metabolites in the medicinal plant, *P. hexandrum* can be achieved. Gene expression approaches such as differential display/subtractive hybridization and cDNA-AFLP offer a feasible alternative for capturing transcripts/genes associated with the secondary metabolite production. Moreover, these approaches give quantitative expression profiles and allow identification of novel genes [32]. The cDNA-AFLP technology coupled with analysis of target/ intermediate metabolites has been successfully utilized in identifying missing links in the biosynthetic pathways through genes-to-metabolite network analysis of biosynthesis of anticancer drugs vinblastine and vincristine in *Catharanthus roseus* [33].

This shows that various research groups have been working throughout the world towards the enhancement of podophyllotoxin production in *in-vitro* cultures but only limited success can be achieved. This may be due to poor understanding of its biosynthetic pathway. Therefore, there exists an urgent need to decipher the podophyllotoxin biosynthetic machinery by means of identification, characterisation and validation of missing links/steps, regulatory genes/enzymes involved in its biosynthesis through biotechnological approaches.

2.2 *Podophyllum hexandrum* as a potential source of podophyllotoxin

Podophyllum hexandrum Royle (*syn. P. emodi* Wall) is native to alpine and sub-alpine regions of Himalayan countries like India, Nepal, Pakistan, China, Afghanistan, and Bhutan (2500-4200m amsl) commonly known as "bankakri or Giriparpat" in Western Himalayan region. *P. hexandrum* Royle has high level of podophyllotoxin content (4.3%) as compared to other podophyllotoxin producing medicinal plants such as North American *Podophyllum peltatum* (0.25%), *Linum album* (0.35%), *Juniperus chinensis* (0.0025%), *Linum nodiflorum* (1.7%) and *Callitris drummondii* (1.4%). Currently, the roots and rhizome of *P. hexandrum* Royle are the main source of podophyllotoxin [34]. The rhizome of *P. hexandrum* is preferred more over *P. peltatum* because its resin is more copious and richer in podophyllotoxin, with a content of about 4.3% dry weight against 0.25% in *P. peltatum* [3]. Although it is possible to chemically synthesize podophyllotoxin but this process has not yet been optimized to be economically feasible. Consequentially, the rhizomes of *P. hexandrum* have been indiscriminately collected to meet the ever-increasing demand for the drug and its

derivatives. This over-collection has led to severe habitat destruction which has been followed by an acute depletion in the population of this herb. The cost of production, lack of organized cultivation, long juvenile phase, poor fruit setting ability, seed dormancy, over-exploitation and geographical limitation of *Podophyllum species,* forced the researchers to find other methods or sources to deal with the increasing demands of podophyllotoxin in the industries.

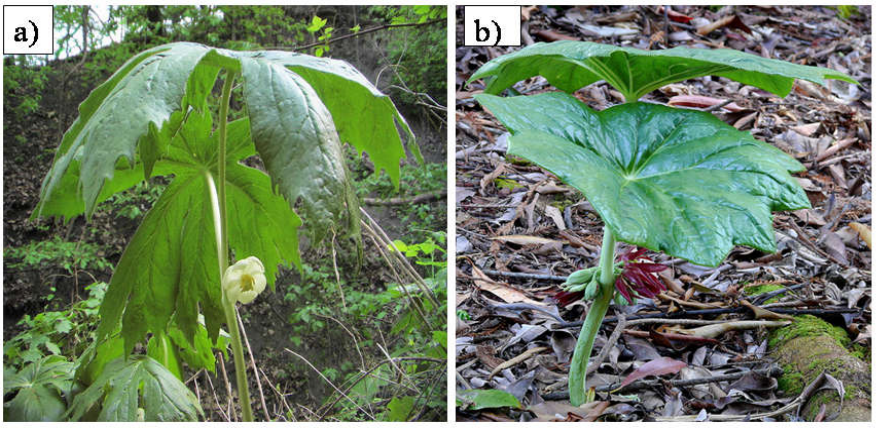
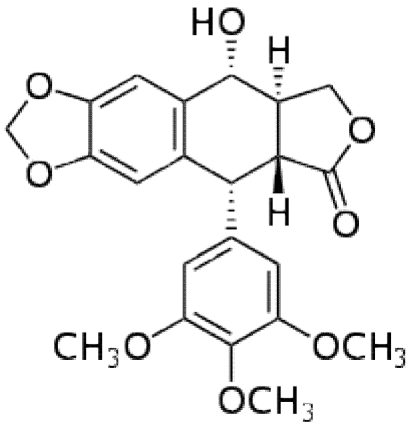
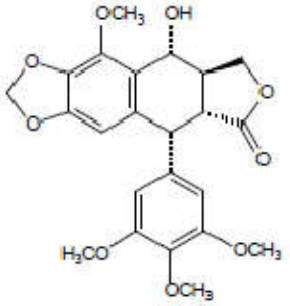


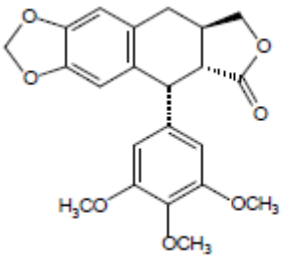
Figure 2.1 Species of *Podophyllum* in its natural habitat – a) *Podophyllum hexandrum* Royle (Indian/Himalayan May Apple)*;* b) *Podophyllum peltatum* (North American May Apple)



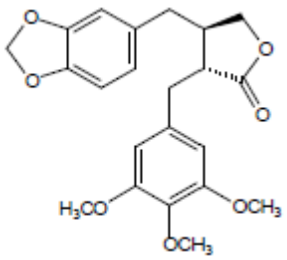
Podophyllotoxin



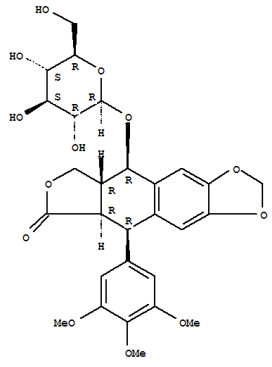
Deoxypodophyllotoxin



6-methoxypodophyllotoxin



Yatein



Podophyllotoxin 4-*O*-glucoside

Figure 2.2 Podophyllotoxin and its derivatives

Table 2.1 Anticancer drugs available in the market obtained from *Podophyllum* species

Podophyllotoxin (Main ingredient)	
Brand Name (Anticancer drugs)	**Country**
Etoposide (Vepesid)	India, USA
Teniposide	USA
Etopophos	USA
Pod-Ben- 25	USA
Condofil	Italy
Warticon	UK
Podofin	USA
Podofilm	Canada
Podophyllotoxin Derivatives	
Brand Name (Anticancer drugs)	**Country**
Boldolaxine	Australia
Canthacur-PS	Canada
Cantharone-Plus	USA, Canada
Opobyl	UK
Posalfilin	UK
Salicylin-P	Australia
Verban	Canada
Verrex	USA
Verrusol	USA
Wartkil	Australia
Wart-off	Australia

2.3 Genetic characterisation of variability in *P. hexandrum*

Molecular markers such as random amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP), microsatellites or simple sequence repeat (SSR), inter-simple sequence repeats (ISSR) have been used for genetic characterisation and population structure and their relationships with the metabolites content in medicinal plant species. Molecular markers provide valuable information of higher metabolites yielding genotypes as these were correlated with metabolites content, thereby have important role in controlling the quality raw materials [35]. Marker-assisted breeding has been utilised for developing high artimisinin content yielding genotypes of *Artemisia annua* 'CIM-arogya' [36].

Variation in podophyllotoxin content with different environmental cues and their correlation with the molecular markers has been done in *P. hexandrum* populations [37, 38, 39, 40]. The increasing availability of genome/transcriptome resources of medicinal plant species opens

the new doors for the identification and characterisation of PCR-based molecular markers that allowed the detailed analyses and evaluation of genetic diversity and also, the detection of genes associated with economically important traits [41].

Microsatellite markers were isolated from *P. hexandrum* and used to quantify the genetic structure within 28 populations [37]. Podophyllotoxin content varied with the altitude of different geographical regions. Genetic diversity has also been performed by using AFLP markers in 28 populations of *P. hexandrum* representing the whole Indian Himalayas [42]. The results of this study showed that regardless of geographical location, the populations from Indian Himalayas are intermixed and are composed broadly of two types of genetic populations. Different types of markers have been identified in *P. hexandrum* that could be utilised for spatial genetic structure and population diversity studies in order to conserve the *P. hexandrum* populations in North-Western Himalayan regions [43].

2.4 *In-vitro* production of podophyllotoxin

Biotechnological strategies have lead to the production of podophyllotoxin through *in-vitro* cell cultures but the yield was very low (0.3%). It was in 1982 that for the first time production of podophyllotoxin in callus cultures of *P. peltatum* was reported [44] since then several studies have confirmed the potential of plant cell and organ cultures to produce this compound. Podophyllotoxin content ranging from 0.008 to 0.30% DW has been obtained by *in-vitro* cultures of *P. hexandrum* and *P. peltatum*, but the growth of *in-vitro* cultures of this species was very slow with a tendency to browning [45]. Even after, over three decades of research in attempting to improve *in-vitro* podophyllotoxin production, the physio-biological factors such as slow growth, browning of cell/callus, culture conditions (medium, pH, oxygen, light and temperature), nutrients, including carbon source and plant growth regulators (PGRs) are needed to be optimised for efficient *in-vitro* production of podophyllotoxin.

In recent years, various plant species have been reported for the production of podophyllotoxin as listed in table 2.2. *In-vitro* cell culture systems have been established from various plant and fungus species as listed in Table no. 2.2. Systematic strategies, including identification and establishment of higher yielding strains/ genotypes and hairy root cultures, optimisation of *in-vitro* cultures systems i.e. shake flasks and large-scale

bioreactors, suitable elicitors/precursors, using two-phase cultures systems and metabolic engineering can lead to higher podophyllotoxin production.

Table 2.2 *In-vitro* podophyllotoxin content in different plants and fungus species

Plant Family	**Medicinal plants**	**Podophyllotoxin content (% of dry wt)**	**Source**	**Reference(s)**
Berbediaceae	*Podophyllum hexandrum*	4.3	Roots/Rhizomes	[2, 3, 34]
	Podophyllum peltatum	0.25	Shoots, Roots/Rhizomes	[2, 3, 34]
	Podophyllum versipelle	0.32	Roots/Rhizomes	[123]
	Diphylleia grayi	1.3	Roots	[123]
	Diphylleia cymosa	0.54	Leaves	[123]
Linaceae	*Linum album* Kotschy ex Boiss	0.35	Callus/suspension cultures	[13]
	Linum nodiflorum	1.7	Callus/suspension cultures	[24]
	Linum flavum	0.2	Callus/suspension cultures	[124]
	Linum flavum	0.16	Roots	[125]
Cupressaceae	*Juniperus chinensis*	0.0025	Leaves (needles)	[23]
	Juniperus chinensis	0.005	Callus/suspension cultures	[23]
	Juniperus scopulorum	0.17	Leaves (needles)	[126]
	Juniperus sabina	0.20	Leaves (needles)	[126]
	Juniperus lucayana	0.10	Leaves (needles)	[126]
	Juniperus virginiana	0.10	Leaves (needles)	[126, 129]
	Juniperus silicicola	0.04	Leaves (needles)	[126]

	Juniperus thurifera	1.3	Leaves (needles)	[128]
	Callitris drummondii	1.4	Leaves (needles)	[25]
Polygalaceae	*Polygala polygama*	0.08	Whole plant	[127]
Fungus species	*Trametes hirsute*	0.03	cultures	[45]
	Fusarium oxysporum	0.02	cultures	[45]

The growing demand for podophyllotoxin and its derivatives cannot be met by the collection of plant sources or commercially unviable chemical synthesis, therefore there is a need to develop alternative methods of production. In recent years genetic engineering of metabolic pathways has given good results in the successful manipulation of secondary metabolite production. Once the rate limiting reactions, regulatory components/steps or transportation components involved in a biosynthetic pathway have been identified metabolic engineering can be used to achieve enhanced levels of production of the desired metabolite from that particular pathway. Metabolic engineering of plant derived secondary metabolites holds high potential, and has been well studied extensively for anthocyanin and flavonoid biosynthetic pathways [46].

Engineering of Lignan biosynthesis pathway in podophyllotoxin for achieving enhanced levels of podophyllotoxin requires a thorough knowledge of the whole biosynthetic pathway and a detailed understanding of the regulatory mechanisms controlling the onset and the flux of the pathway. Unfortunately as mentioned before, the complete information related to podophyllotoxin biosynthesis is not yet available which explains why only limited success has been obtained by metabolic engineering in case of podophyllotoxin production. Nevertheless the potential is high for future as successful application of metabolic engineering has been achieved for other plants which are involved in lignan biosynthetic pathway [47, 48, 49]. *Forsythia koreana* is one such plant that produces matairesinol from pinoresinol via the catalysis of Pinoresinol/Lariciresinol reductase (PLR) and secoisolaricinesol dehydrogenase; both these enzymes are also involved in the biosynthetic pathway of podophyllotoxin. Application of RNA interference for the PLR gene (PLR-RNAi) in cell cultures of *Forsythia* increased the yield of pinoresinol and its glucoside by approximately 20-fold due to blocked biosynthesis of matairesinol [48].

2.5 Genome/transcriptome resources of *Podophyllum* species

Modern 'omics' resources such as genomics, transcriptomics, proteomics and metabolomics could reveal the molecular components associated with the biosynthesis and accumulations of metabolites in medicinal plant species. The genome size of *P. hexandrum* was estimated about 31442.70 Mbp through flow Cytometry by Nag et al [50]. The transcriptome resources of *Podophyllum* species (*Podophyllum hexandrum* and *Podophyllum peltatum*) have been generated and available at medicinal transcriptomes database (http://medplants.ncgr.org) [10]. This medicinal transcriptomes database is a part of NIH funded project "Transcriptome Characterization, Sequencing, and Assembly of Medicinal Plants Relevant to Human Health" operated by many institutes such as Institute of Biological Chemistry (IBC), Washington State University (WSU), Donald Danforth Plant Science Center (DDPSC) and National Center for Genome Resources (NCGR). This database contained the raw and assembled transcriptomes data from different tissues/organs not only of *Podophyllum* species but also for many other medicinal plant species which are quite useful for exploring and understanding the physio-biological processes in plant systems.

The availability of transcriptomes resources from different tissues/organs of two *Podophyllum* species differing in the sites of biosynthesis and accumulation of podophyllotoxin provided a robust dataset to mine for potential molecular components and correlate with podophyllotoxin content.

2.6 Molecular basis of podophyllotoxin biosynthesis

Podophyllotoxin belongs to lignan group of secondary metabolites. Lignans are a diverse and ubiquitously distributed group of compounds that are biosynthetically derived from the phenylpropanoid pathway/shikimic acid pathway. The shikimate pathway is the entry to the biosynthesis of phenylpropanoids, which is intra-cellular plastidial in origin. L-phenylalanine is one of the final products of shikimic acid pathway. Phenylpropanoids and their derivatives are compounds containing a C_6C_3- moiety derived from L-phenylalanine [51]. Phenylalanine one of the essential amino acids obtained from this pathway serves as the precursor for the biosynthesis of a range of phenylpropanoids which are further utilized for the production of a range of plant secondary metabolites (Figure 2.3).

The biosynthetic route leading to formation of podophyllotoxin remains to be elucidated completely, only twenty six genes are known which catalyse 31 steps in podophyllotoxin biosynthetic pathway up to the intermediate compound pluviatolide [2, 10].

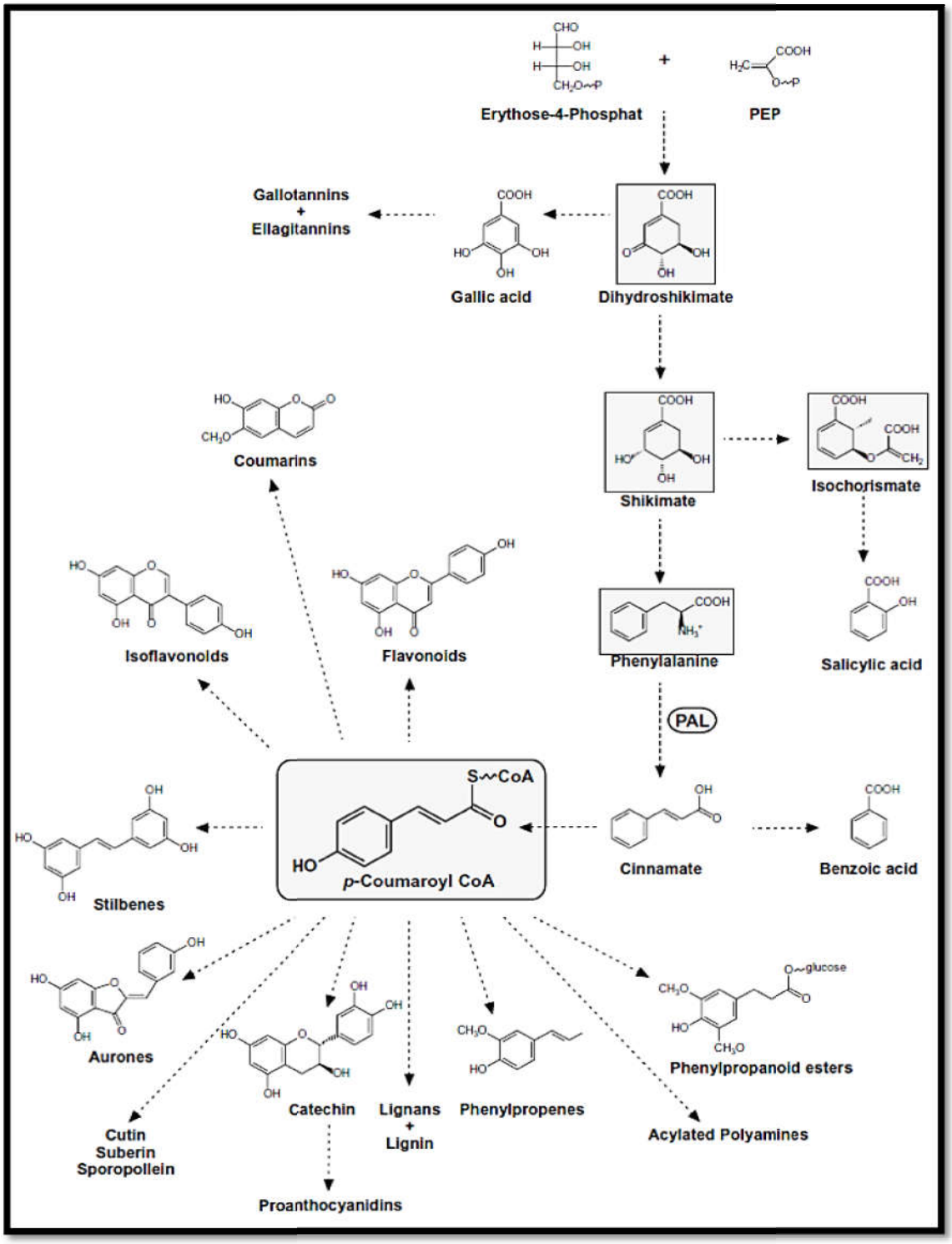


Figure 2.3 Diversification of phenylpropanoids based on the general phenylpropanoid pathway (Vogt [130])

The initial steps in podophyllotoxin biosynthesis are common to most of the phenolic compounds. Deamination of phenylalanine to yield cinnamic acid is catalysed by the enzyme *phenylalanine ammonia lyase* (PAL). *Cinnamic acid 4-hydroxylase* (C4H) then causes hydroxylation of cinnamic acid to form *p*-coumaric acid and after additional steps coniferyl

alcohol is formed by *cinnamyl alcohol-dehydrogenase* (CAD). Two molecules of coniferyl alcohol are coupled to pinoresinol. From pinoresinol, the biosynthetic pathway divides, leading to a great diversity of lignans. In case of podophyllotoxin, pinoresinol undergoes enantiospecific reduction, via action of *pinoresinol reductase* (PLR), to sequentially yield (+)-lariciresinol and (–)-secoisolariciresinol. Stereo-specific dehydrogenation by Secoisolariciresinol dehydrogenase next converts secoisolariciresinol into (–)-matairesinol, the last unequivocal known step in (–)-podophyllotoxin biosynthesis [10]. Figure 2.4 is an illustration of the steps involved in formation of podophyllotoxin from coniferyl alcohol.

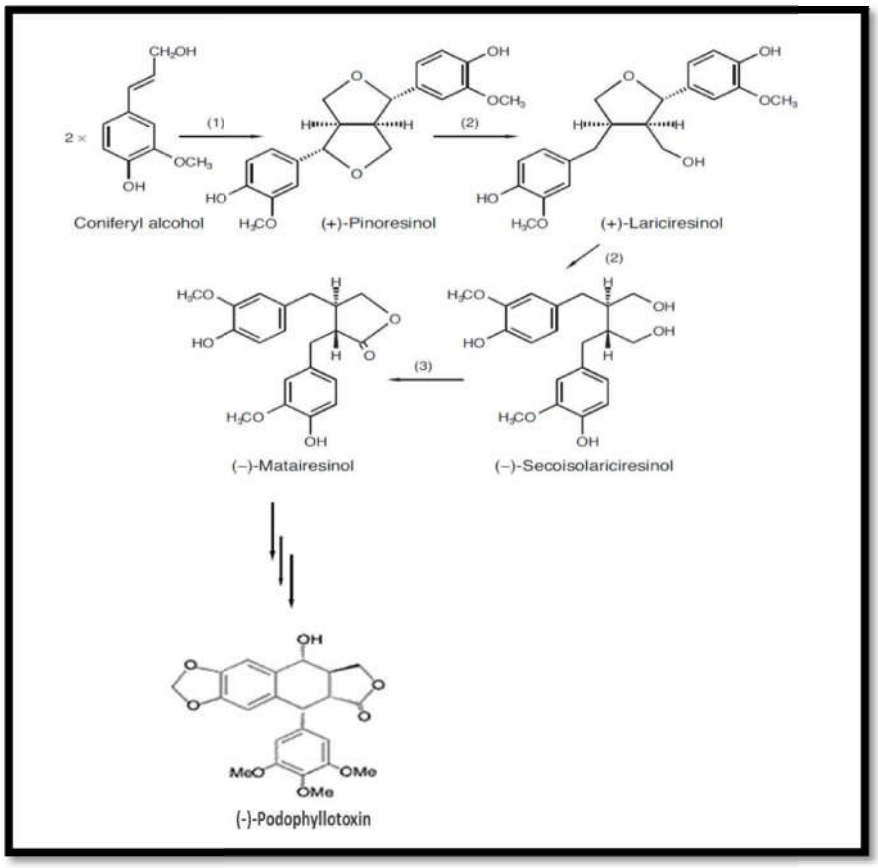


Figure 1.4 Biosynthesis of podophyllotoxin from coniferyl alcohol

Molecular cloning and characterisation was done only for 7 genes (*PS, PLR, SD, DPO, CAD*, *C4H* and *PCH*) implicated in the biosynthesis of podophyllotoxin [11, 12, 13]. Persual of literature showed that no study was done on identification of candidate genes in relation to podophyllotoxin content in *P. hexandrum.*

2.7 Regulators of podophyllotoxin biosynthesis

Regulatory components are known to regulate various physiological, biochemical and biological processes including growth, development, senescence, seed germination, environmental stress responses, etc. contributing towards secondary metabolites biosynthesis in plant species. The availability of genomics, transcriptomics and metabolomics resources can provide a powerful tool for identification, characterization and functional analyses of regulatory components in order to infer their role in regulation of metabolic processes in plant species.

2.7.1 Transcription factors

Transcription factors (TFs) are DNA-binding specific proteins that interact with *cis*-acting elements in promoter regions of target genes, thus regulating the gene expression [52]. Generally, TFs represent ~7% of all genes present in various plant genomes and presently, 108 TF families have been identified in the plant kingdom [53]. Over the last several years, TFs have emerged as a center point of scientific research on elucidation of regulatory mechanisms underlying the biosynthetic pathways of natural products/phytochemicals/metabolites. In plants, several TF families such as bZIP, WRKY, MYB, bHLH, AP2/ERF, mTERF, TFIIIA, DOF, NAC, Zinc finger and SPL have been reported to be involved in regulating the secondary metabolites biosynthesis [52, 54]. TFs can regulate the expression of multiple genes implicated in a metabolic pathway through the recognition of sequences in promoter regions [55].

The MYB TFs were the largest family and mainly involved in controlling biological processes including in biotic and abiotic responses, metabolism, defense, development, etc. A number of MYB transcription factors were known in various plant species, including maize, petunia, snapdragon, grape, apple and *Arabidopsis* that control phenylpropanoid pathway [56]. For example, R2R3MYB TF encoding genes have been identified in several plant species such as 121 genes in tomato (*Solanum lycopersicum* L.) (Zhao et al. 2014), >200 genes in *Zea mays*, 126 genes in *Arabidopsis thaliana*, 117 genes in *Vitis vinifera*, etc. signifying their major roles in plant response to stress conditions [57, 58, 59, 60]. Previous studies have demonstrated that bHLH TFs (GL3, eGL3, TT8) played important role in regulation of anthocyanin biosynthesis in *Arabidopsis thaliana* through interaction with MYB TFs [61]. MYC2 protein of bHLH class binds to *cis*-elements in the ORCA3gene promoter,

thus controlling the expression of several terpenoid indole alkaloids (TIA) biosynthesis genes in *Catharanthus roseus* [62]. Similarly, WRKY TFs such as GaWRKY1, AaWRKY1, WRKY3, WRKY6 and WRKY33 are also known to control the regulation of diverse biological processes in plants like cotton, *Artemisia annua*, tobacco and *Arabidopsis*, respectively, contributing to the biosynthesis of gossypol, artemisinin, and camalexin, respectively [63, 64, 65]. In plants, bZIP is a master regulator actively involved in physiological and biological processes, including growth and development, seed formation, abiotic and biotic stress responses, etc. [66]. The bZIP TFs are well known and characterized in grapevine (*Vitis vinifera* L.) for their role in controlling the flavonoid biosynthesis through the regulation of phenylpropanoid pathway [67]. However, no information is available as of today on transcription factors (TFs) contributing to podophyllotoxin biosynthesis in *Podophyllum* species.

2.7.2 MicroRNAs as regulators of secondary metabolites biosynthesis

MicroRNAs (miRNAs) are non-coding small RNAs, ranging from 18 to 22 nucleotides known to regulate various biological processes in plant systems. miRNAs have perfect or near-perfect complementarity to their mRNA targets, thereby, down-regulate their expression either through inhibition of translation or cleavage in plants [68]. miRNAs have been identified to regulate physio-biological processes such as signal transduction, growth and development, seed germination, stress responses, secondary metabolism in plant species *Picrorhiza kurroa*, *Arabidopsis thaliana*, *Lycopersicum esculentum, Oryza sativa, Solanum tuberosum, Panicum virgatum*, etc.[68, 69, 70, 71, 72].

In plants, most of the miRNAs are conserved across species and can bind to multiple targets either through coding or non-coding regions [73]. Previously, the publically available plant genomic resources including genomic survey sequences (GSS), high throughput genomic sequences (HTGS), expressed sequence tags (ESTs) and transcript sequence assemblies (TSA) were used for the identification of miRNAs in plant species [74]. Next generation sequencing (NGS) in combination with comparative genomics is a powerful strategy for identification of miRNAs from transcriptomes of plant species. miRNAs target genes have been identified on the basis of their sequence complementary with miRNAs. The functions associated with miRNAs target genes have been identified with NCBI BlastX, GO (gene ontology) and KAAS (KEGG automatic annotation server) tools [75]. However, there is no report on role of miRNAs in podophyllotoxin biosynthesis in *P. hexandrum*.

Research on miRNAs in plant species has opened the new window to enhance metabolites content through understanding the regulation mechanisms involved in the biosynthesis of metabolites. The biosynthetic pathways operating in plant systems have been altered through the coordination mechanism of miRNAs to obtain desired amount of natural products/metabolites. Previously, silencing of chalcone synthase (CHS) gene by siRNAs has altered the biosynthetic pathway of flavonoids in *Glycine max* [76]. Similarly, in *Arabidopsis thaliana*, miRNAs miR156, miR163, miR393 and miR828 have been found to regulate the biosynthesis of secondary metabolites such as anthocyanin, flavonols, etc. [77].

miRNA vun-miR396b has been reported for its role in regulation of metabolites through regulation of growth regulating factor (GRF) and bHLH (basic helix-loop-helix) family of transcription factors in *Arabidopsis thaliana* [78]. Moreover, miRNA vvi-miR828b was shown to control the MYB transcription factor regulating flavonoid and lignin biosynthesis in *Arabidopsis thaliana* [79]. MYB and WRKY TFs are known for their regulatory role in various physio-biological processes in different plant species [80]. WRKY TFs (GaWRKY1, AaWRKY1, WRKY3, WRKY6 and WRKY33) were found to be upregulated in gossypol, artemisinin and camalexin biosynthesis in cotton, *Artemisia annua*, tobacco and *Arabidopsis*, respectively [63, 64, 65].

2.8 Transporters of secondary metabolites vis-`a-vis podophyllotoxin

ATP-binding cassette (ABC) transporters are known to transport secondary metabolites into the vacuoles of plant cells [80]. ABC transporters constitute one of the largest protein families that are present ubiquitously from bacteria to humans [81]. The function of ABC proteins in plants was originally identified as transporters of phytotoxic xenobiotics into the vacuoles [82], however, since then several studies on their functionality in plants have been carried out and these proteins are now known to be involved in a range of diverse processes including plant growth, nutrition and development, response to abiotic stresses, pathogen resistance and interaction of plants with environment [83]. The role of ABC proteins in transport of secondary metabolites within plants is of particular interest for directing future efforts in enhancing the level of metabolites being produced in different production systems.

ABC transporters can be classified into subfamilies based on their domain architecture, similarities/structure comparison with human and microbial systems, functional and phyllogenetic information. The ABC transporter subfamilies such as ABCA, ABCB, ABCC,

ABCD, ABCE, ABCG, $ABCG_{WBC}$ group and $ABCG_{PDR}$ group, etc. have been identified in the plant systems [84, 85]. The ABCA, ABCB, ABCC and ABCG and ABCGWBC group of ABC transporters subfamilies were thought to be involved in the transport of various terpenoids, alkaloids and lignins compounds [86].

The biosynthesis of secondary metabolites is cell/tissue specific usually occurring in different tissues/organs such as leaf, root, rhizome, stems and flower, etc. For example, phytochemicals/metabolites like shikonin biosynthesis in vesicles of endoplasmic reticulum of *Lithospermum erythrorhizon,* picrosides in different tissues/organs of *Picrorhiza kurroa*, etc. and finally stored in vacuoles [87, 88]. Similarly, podophyllotoxin stored in vacuoles of *Podophyllum* species. However, no information exits on types of ABC transporters involved in the transport of podophyllotoxin in *P. hexandrum*.

The complete review of literature on podophyllotoxin, sources of podophyllotoxin and its derivatives, genetic diversity, *in-vitro* production strategies, genome/transcriptome resources, molecular basis of metabolites biosynthesis and regulatory components of biosynthetic pathways highlights the following major gaps in this highly valuable medicinal herb:

1) Knowledge of podophyllotoxin biosynthesis is fragmentary with missing links in the biosynthetic pathway
2) Lack of relative contribution of pathway steps in contributing to the podophyllotoxin content
3) Lack of knowledge on molecular components of podophyllotoxin biosynthesis
4) Lack of knowledge on regulatory control of podophyllotoxin biosynthesis
5) Lack of genetic intervention strategies to improve podophyllotoxin content *in- vitro*
6) No genetically defined accessions known for commercial use

CHAPTER 3

MATERIAL AND METHODS

3.1 Plant materials and growth conditions

Eleven accessions of *P. hexandrum* Royle belonging to different geographical locations of North-Western Himalayas were procured from the Himalayan Forest Research Institute at Jagatsukh (1,900 m altitude, $20^0$35.6′–$32^0$6.1′N and $78^0$57.8′–$77^0$33.7′E), Kullu, Himachal Pradesh, India. The plants were maintained in plastic pots (20 cm height x 20 cm top width x 12 cm bottom width) containing soil, sand and farm yard manure mixture in a ratio of 2:1:1 with controlled environmental conditions [light (intensity 300-1400 lx), temperature (25±2^{0}C), relative humidity (≈ 80%) and photoperiod (14 h day/10h night) at the green house of Jaypee University of Information Technology, Waknaghat, Solan, India. Shoot, root and rhizome tissues of *P. hexandrum* (3 year old plants) were harvested, frozen in liquid nitrogen and stored at -80^{0}C for metabolite quantification and expression analysis through qRT-PCR.

Table 3.1 *Podophyllum hexandrum* accessions from different geographical locations used for quantification of podophyllotoxin

Accession	Location	District	Altitude (amsl)	Latitude	Longitude
Ph-1	Jammu Dug	Chamba	3620	33006'27.1"N	76029'171.1"E
Ph-2	Kulri	Kinnaur	3330	31030'01.3"N	78056'34.5"E
Ph-3	Kallo Noi	Kullu	3597	31045.412″N	77027.680″E
Ph-4	Jalori Pass	Chamba	3590	32°31′28″N	76°31′01″E
Ph-5	Chamba	Chamba	3354	31018'23.6"N	77045'02.1"E
Ph-6	Parot	Kullu	3144	32°0'19"N	77°14'43"E
Ph-7	Panertu	Kinnaur	3645	31020'42.2"N	78026'70.2"E
Ph-8	Thothi Dhar	Kullu	3200	31032.448″N	77024.605″ E
Ph-9	Thang Bo	Kinnaur	3435	31040'6.2"N	78001'11.9"E
Ph-10	Gulaba	chamba	2703	31°04′32″N	77°38′36″E
Ph-11	Homkhani	Kullu	3979	32°22′17″N	77°14′47″E

3.2 Quantification of podophyllotoxin content

The quantification of podophyllotoxin content (expressed as % of fresh wt) was carried by reverse phase high-performance liquid chromatography (HPLC) in shoots, roots and rhizomes tissues of eleven accessions of *P. hexandrum* (Figure 3.1). The quantification of podophyllotoxin was performed on Waters HPLC system (HPLC Waters 515) equipped with 515 HPLC Pump, Waters 717 autosampler, Waters 2996 photodiode-array detector and Waters Spherisorb® C18 (4.6×250 mm, 5 µm) reverse phase column using Empower-2 software at 210 nm UV wavelength [89]. The podophyllotoxin was identified on the basis of retention time and comparison of UV spectra with the authentic podophyllotoxin standard (Chroma Dex, CA, USA).

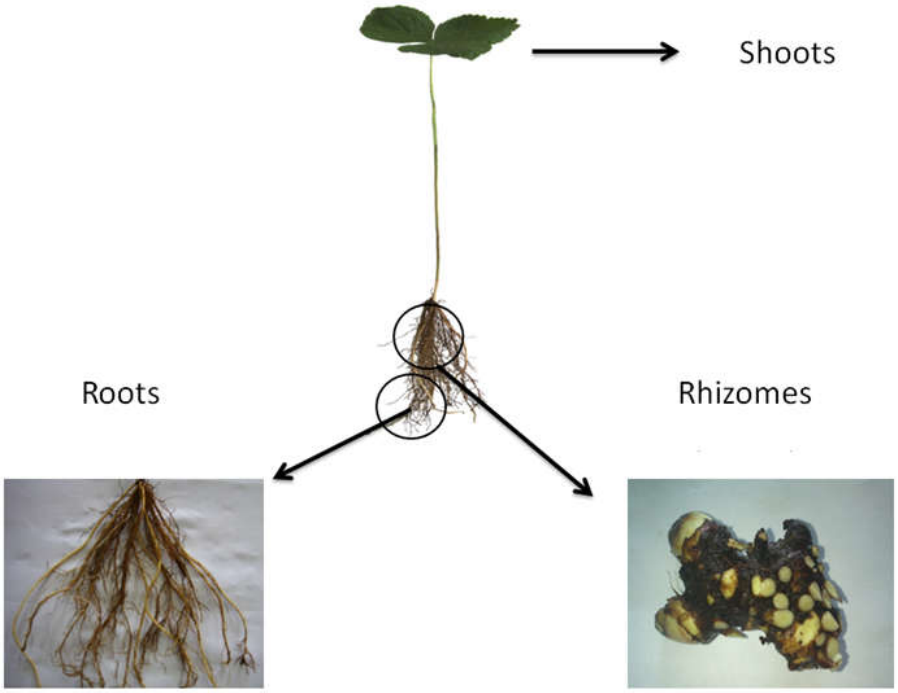


Figure 3.1 Different parts of *P. hexandrum* plant used for HPLC quantification

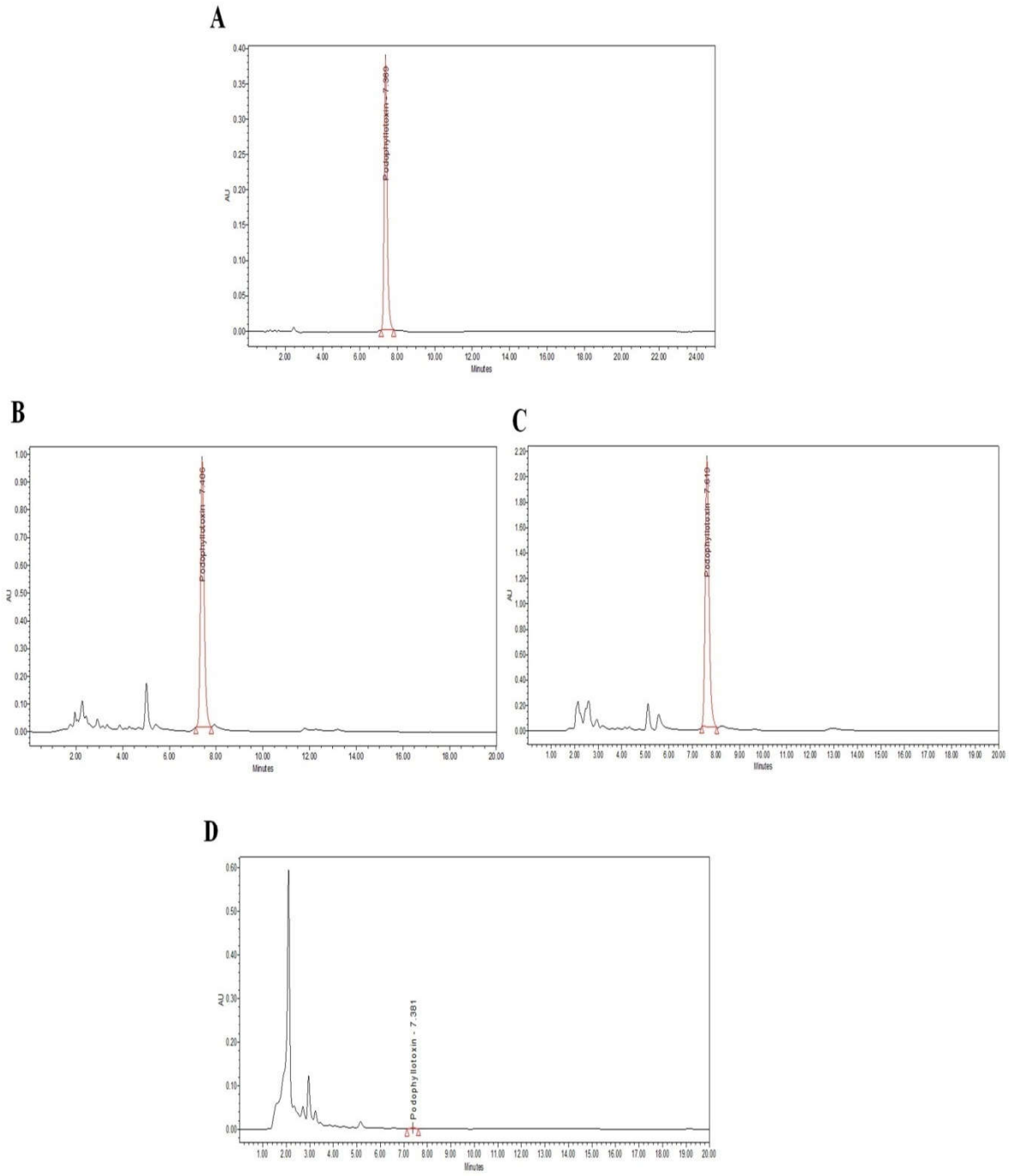


Figure 3.2 HPLC chromatogram of podophyllotoxin: A) standard compound; B) roots, C) rhizomes D) shoots

3.3 Identification of high and low metabolite content accessions of *P. hexandrum*

On the basis of podophyllotoxin quantification through HPLC, high and low content accessions of *P. hexandrum* were identified and selected for further experiments to ascertain the role of molecular components implicated in podophyllotoxin biosynthesis.

3.4 Analysis of transcriptomes of *Podophyllum* species

Assembled transcriptomes data of *Podophyllum* species (*P. hexandrum* - 2, 27, 885 contigs, *P. peltatum*- 1, 47, 960 contigs) were downloaded from medicinal plants transcriptomics database (http://medplants.ncgr.org/) [10] and the transcriptomes were mined and analyzed for molecular components such as biosynthetic pathway genes, transcription factors (TFs), ABC transporters and miRNAs by using in-house developed PERL scripts. Workflow depicting mining, identification and experimental validation of molecular components from *Podophyllum* species transcriptomes is given in Figure 3.3.

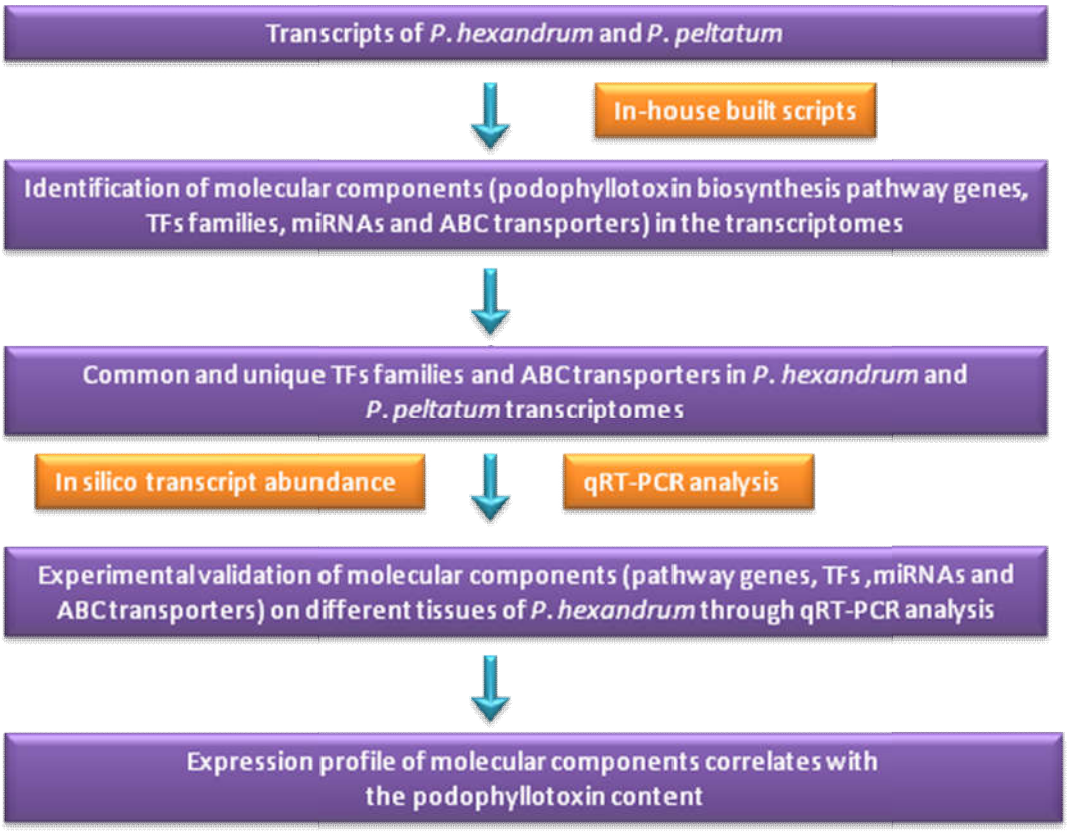


Figure 3.3 Workflow depicting transcriptomes mining, identification and experimental validation of molecular components from *Podophyllum* species

3.5 Identification of pathway genes contributing to podophyllotoxin biosynthesis

The functional annotation of *P. hexandrum* transcriptome was done by comparing transcripts against the NR protein sequences database available at NCBI by using BLASTX algorithm with E-value threshold of 10^{-05}. Using in-house developed perl scripts, fourteen genes (*DAHPS, DHQS, DHQD, COMT, EPSPS, CS, CM, PPA-AT, PD, ADT, PAL, 4CL, COMT, CCR*) related to podophyllotoxin biosynthesis were identified based on high homology (identity value ≥ 30%) (Figure 3. 4). Besides this, nucleotide sequences of 7 genes, namely *CYP719A23, PLR, SRD, PS, CAD, C4H,* and *C3H* were retrieved from *P. hexandrum* CDSs available at GenBank Database (http://www.ncbi.nlm.nih.gov/genbank/).

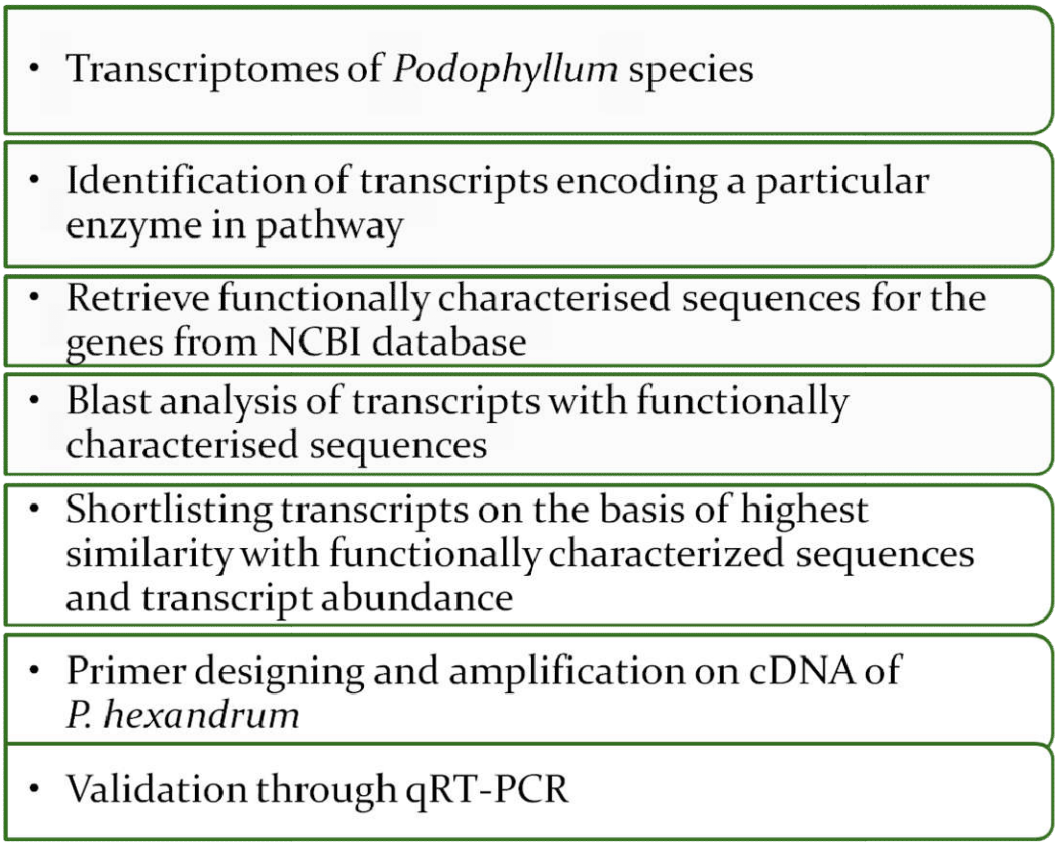


Figure 3.4 Workflow for identification of podophyllotoxin pathway genes

3.6 Computational mining of TFs families in transcriptomes of *Podophyllum* species

Plant transcription factor database (PlantTFcat) (http://plantgrn.noble.org/PlantTFcat/familylist.do) was used for identification and classification of TFs families involved in secondary metabolites biosynthesis. Transcripts coding for TFs families in transcriptomes of *Podophyllum* species (*P. hexandrum* and *P. peltatum*) were identified by searching against all transcription factor protein sequences at PlantTFcat using BLASTX with an E-value cutoff $1e^{-6}$ and according to the presence of InterPro domains. All transcripts were annotated to proteins (according to possible open reading frames) and searched through InterProScan. PlantTFcat classified these transcripts into 108 known TF families. The conserved domains in TFs were identified by using conserved domain database available at NCBI (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). Workflow is given in Figure 3.5 depicting the identification of transcription factors from transcriptome resources of *Podophyllum* species.

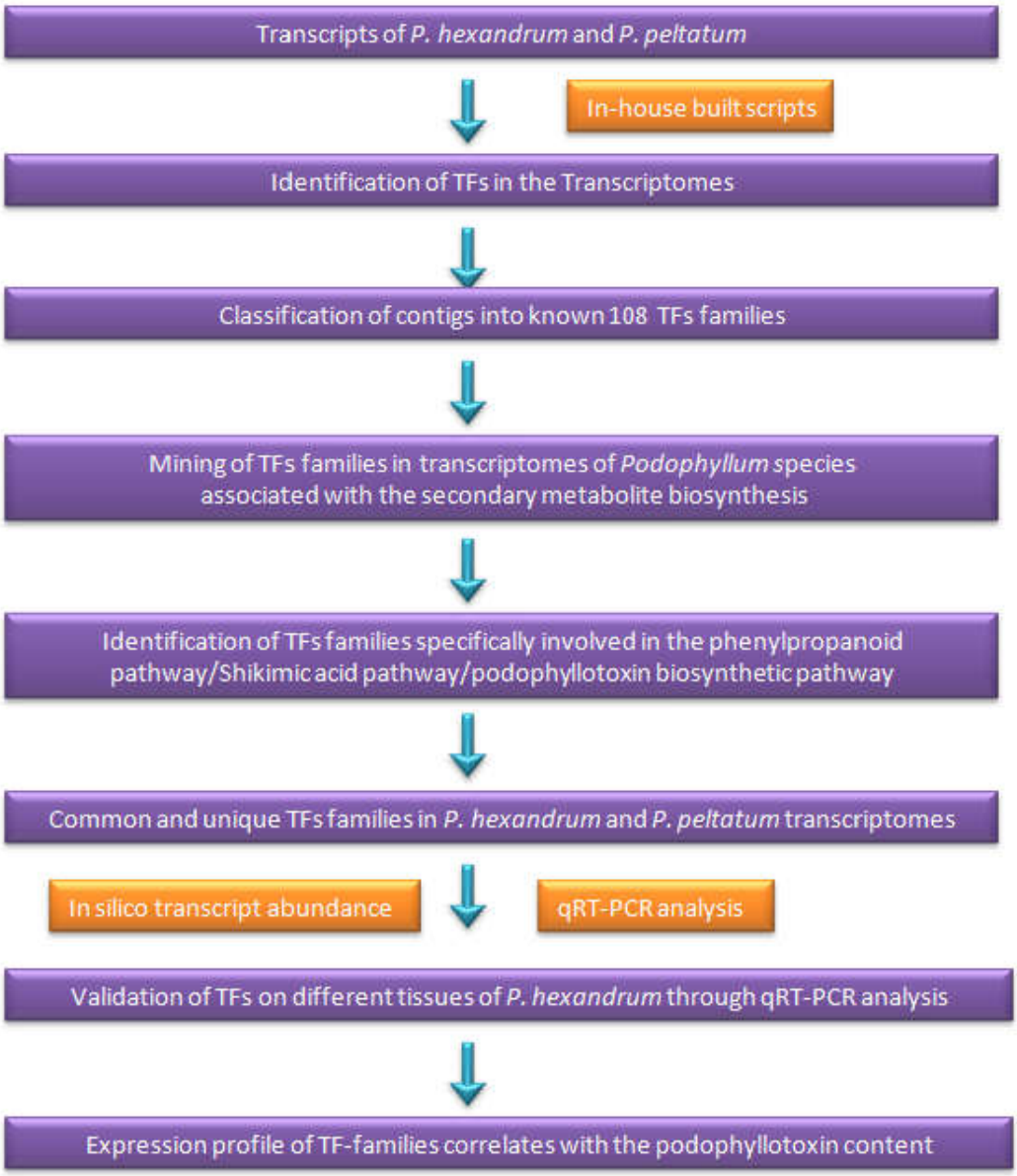


Figure 3.5 Workflow for identification and analysis of TFs in transcriptomes of *Podophyllum* species

3.7 Mining ABC transporters in transcriptomes of *P. hexandrum* and their phyllogenetic analysis

P. hexandrum assembled transcriptomes data (2, 27, 885 contigs) were downloaded from medicinal plants transcriptomics database (http://medplants.ncgr.org/) [10] and the transcriptomes were mined and analyzed for ABC transporters by using in-house developed PERL scripts. Protein family database (Pfam) (pfam.xfam.org) was used for identification of ABC transporters in the transcriptomes of *P. hexandrum*. Transcripts encoding ABC transporters in *P. hexandrum* were mined by searching the transcriptomes against all

available ABC transporters sequences at Pfam database using in-house developed perl scripts. For phylogenetic reconstruction of identified ABC transporters, protein sequences of known ABC transporters characterized in model plant species were downloaded from NCBI Genbank database. The alignment of sequences was carried out by CLUSTALX and final tree was constructed in MEGA 6 (Molecular Evolutionary Genetics Analysis 6.0) software using Maximum Likelihood method [90].

3.8 Mining miRNAs and their mRNAs targets in transcriptomes of *P. hexandrum*

miRNAs were identified in transcriptomes of *P. hexandrum* using in-house built perl-scripts (Figure 3.6). PMRD (plant microRNA database) (http://bioinformatics.cau.edu.cn/PMRD/) was used as a reference database for mining conserved miRNAs in different tissues/organs of *P. hexandrum*. psRNATarget (A Plant Small RNA Target Analysis Server) (http://plantgrn.noble.org/psRNATarget/) target prediction tool was used to identify miRNAs targets in P. *hexandrum* transcriptomes [91]. NCBI-NR database was used for identification of functions governed by miRNAs by BLASTX program with default parameters. The identified miRNAs targets were involved in various physio-biological processes in *P. hexandrum*. In order to better understand the functions of miRNAs, their possible targets were identified through GO, COG and KAAS using default parameters.

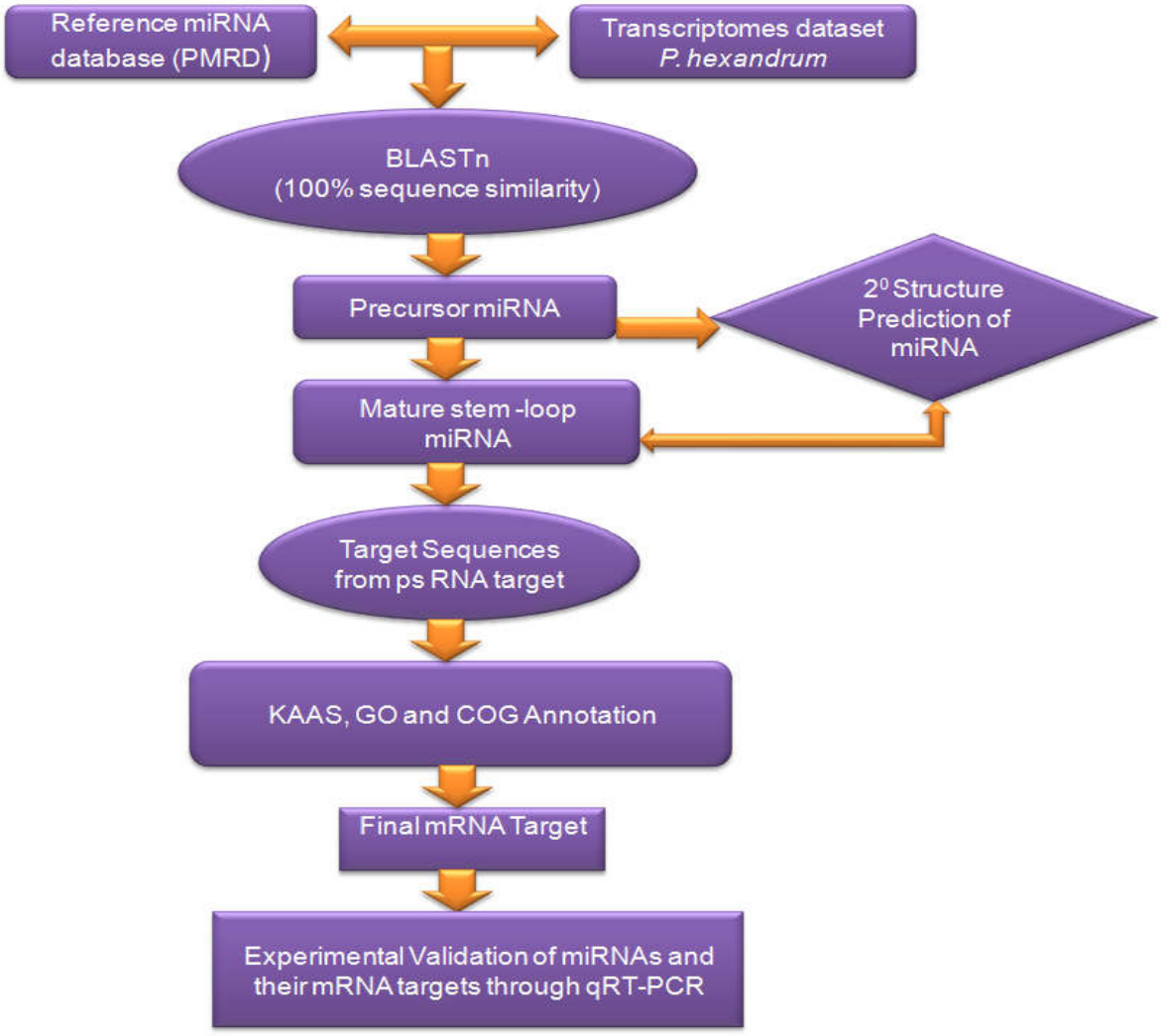


Figure 3.6 Workflow for identification of miRNAs and their potential targets

3.9 *In-silico* transcript abundance for pathway genes, ABC transporters and transcription factors through RSEM

RNA-Seq by Expectation Maximization (RSEM) was used for quantification of transcript abundance from the RNA-Seq high quality data. RSEM measures the expression profiles in terms of FPKM (Fragments per kilobase of transcript per million mapped reads) and RPKM (Reads per kilobase per million) [92]. The transcript abundance of the transcripts encoding pathway genes, ABC transporters and transcription factors in transcriptomes of *P. hexandrum* was calculated using FPKM parameter of RSEM package. All default parameters were used in the query option.

3.10 Identification of putative *cis*-regulatory elements

The *cis*-regulatory elements in the promoter regions of biosynthetic pathway genes were identified through *in-silico* analysis using Plant Genome Database (http://www.plantgdb.org/) and Plant *cis*-acting Regulatory DNA Elements Database (PLACE) (https://sogo.dna.affrc.go.jp/cgiin/sogo.cgi?lang=en&pj=640&action=page&page=newplace) [93]. The complete coding sequences of pathway genes (*CAD, C4H, PS, C3H*, *PAL, 4CL, PLR* and *SRD*) were retrieved from NCBI Genbank and BlastN analysis was performed in Plant Genome Database (http://www.plantgdb.org/) against the genome sequences of *Arabidopsis thaliana*, *Solanum tubersoum*, *Vitis vinifera* and *Glycine max*. The *cis*-regulatory elements were mined through PLACE database by taking upstream region up to 1 kb from translation start site of each gene.

3.11 Isolation of RNA and cDNA synthesis

Total RNA was isolated from 100 mg of frozen tissues of shoots and rhizomes of *P. hexandrum* by using QIAGEN RNeasy mini kit and digested with DNaseI (Invitrogen). First-strand complementary DNA synthesis (cDNA) was done using Verso cDNA synthesis kit (Thermo Scientific, USA) from total RNA (2 µg) template as per manufacturer's instructions. The primers for qRT-PCR analysis were designed using Primer 3 v.0.4.0 software.

3.12 Experimental validation and expression analysis using quantitative real-time PCR analysis

Quantitative real-time PCR (qRT-PCR) based analysis for biosynthetic pathway genes, transcription factors, ABC transporters and miRNAs were performed on shoots, roots and rhizome tissues of *P. hexandrum*. PCR reaction was performed in 12.5 µl mixture containing diluted cDNA sample (100 ng) as a template, SYBR Green PCR master mix (Bio-Rad Laboratories) and 10µM each of forward and reverse gene-specific primers (Table 3.2, Table 3.3, Table 3.4, Table 3.5). The reactions were performed using CFX96 system (Bio-Rad Laboratories; Hercules CA, USA) with the following programme: 94°C (3 min), 94°C (30 s), corresponding annealing temperature (30 s), 72°C (20 s) at 40 cycles. To normalize the variance in cDNA input, *cyclophilin*, *Elongation factor-1 alpha* (*EF-1 alpha*) and 5.8sRNA genes were used as an internal control. The relative expression of each gene was calculated using the comparative Ct value method. The significant differences between transcripts were statistically evaluated by standard deviation.

Table 3.2 Primers used in qRT-PCR expression analysis of podophyllotoxin biosynthesis pathway genes

Sr. No.	Primer Name	Sequence (5′–3′)	Annealing Temperatures (°C)
1.	Ph-SRDF	CATCAAGCTGCAAACCTCAA	54
2.	Ph-SRDR	TCAAGCCAATCCATGTTTCA	
3.	Ph-CADF	GGCATCTCCAGAGCAAGAAC	55
4.	Ph-CADR	TGACCATGTGGAGGTCAGAA	
5.	Ph-PLRF	ATTCTTGCCATCGGAGTTTG	50
6.	Ph-PLRR	CCCCAAAAAGTATCCAGCAA	
7.	Ph-PSF	TCAAGCAACGCATACCATGT	57
8.	Ph-PSR	TCGTCGAACACAACCACATT	
9.	Ph-C4HF	GGCCAACGTAATCTCATCGT	53
10.	Ph-C4HR	CCGCATTTTTCTCCAGTGTT	
11.	Ph-C3HF	GCTCGTAGGGATCGTCTCAC	55
12.	Ph-C3HR	AGCAGTGATCATGTCCCACA	
13.	Ph- DAHPS F	CGGAGGTTGTTGTGGAAGAT	50
14.	Ph- DAHPS R	CCTCTCGTCCAAATCCATGT	
15.	Ph-3DHQDF	GCCTAACGAATCCAACGAAA	53
16.	Ph-3DHDR	ATCTCCCAATCCAAGCTCCT	
17.	Ph- EPSPS F	CCACAGTGGCAGAGTAGTGC	56
18.	Ph- EPSPS R	GTTCGGATGTGCCTGTAACC	
19.	Ph-CSF	AGGGCTCTTGTAGCATTCCA	50
20.	Ph-CSR	TCTCATAGCACGTGGTCGTC	
21.	Ph-CMF	CTGAATCCCTCCATGGAAAA	52
22.	Ph-CMR	CCATGTTTCTGCCGTTTCTT	
23.	Ph-PALF	CACACATGGCTGTGAAAACC	52
24.	Ph-PALR	ATACCCCGGCAATTCCTAAC	
25.	Ph-PDF	GATTCTGGATTCCTCGTCCA	51
26.	Ph-PDR	CTCTTCGTCGACGTCCTTTC	
27.	Ph-ADTF	GCTCAGCATAAGCTGGTTCC	50
28.	Ph-ADTR	CGACGAGGAATTGTCGATTT	
29.	Ph-4CLF	TGGAGAGAATCCCAATTTGC	52
30.	Ph-4CLR	CCGAAACCTTGTACCGTTGT	
31.	Ph-COMTF	ATGGAAAGATCTTGGCGATG	50
32.	Ph-COMTR	CTTGTCAGCGTCCACGAATA	
33.	Ph-CCRF	CACTTCATCCGAGCACTTCA	51
34.	Ph-CCRR	AGGCTTATGTGGACGTGAGG	
35.	Ph-CMTF	CATCGGCTTTTGGAACACTT	50
36.	Ph-CMTR	ATGTTGGTGGTGGAACTGGT	
37.	Ph-PPA-ATF	GCATCAAGGCTTCCACTAGC	52
38.	Ph-PPA-ATR	GCACCAGATGGTTCTTTCGT	
39.	Ph-CYP719A23F	TAGGGGACTTACGCATGGTC	52
40.	Ph-CYP719A23R	CAAGGATTGCGAGGAATGAT	
41.	Ph- DHQS F	TCATGGAGAAGCTGTTGCAG	52
42.	Ph- DHQS R	GTCAACCGCCATGTAGGACT	

Table 3.3 Primers used for expression analysis of key transcription factors through qRT-PCR

Sr. no.	TFs	Forward Primer	Reverse Primer	Size (bp)	Tm
1.	*Phex*_MYB	CGACCAGGCCTCAAAGAAAC	CGTGCTTTGCTTGAAACTGC	165	60^0C
2.	*Phex*_WRKY	GCGCGACGGAACTCCGCCAT	GTGGGGTCTCGGCTTCTCTC	147	59 ^{0}C
3.	*Phex*_bHLH	AAAGCCCTCAACAAACACCG	GGGACACAGGACTTTGTTGG	133	60^0C
4.	*Phex*_bZIP	TTTCAATCTTCTGCAGCCGG	AGACGCTGCCATGTTTGTTT	156	52^0C

Table 3.4 Primers used for expression analysis of ABC transporters through qRT-PCR

Name	Forward Primer 5'-3'	Reverse Primer 5'-3'	Tm (^{0}C)	Product Size (bp)
PhABC1	GTGCACTTGACGCTGAATCA	ATCAAATCCGCATGCCTTCC	55	172
PhABC2	CACGACAAACTTCAGCCTTCA	GGTCCCTCTCCATCCTCATC	52	160
PhABC3	CCCCGTATCGACAAGAAGGA	CACCTCTCTTCCTCCTTGCA	59	148
PhABC4	GGAAGCTTGTGTCTGGTTGG	CCAGGTCTGCACATACTGGT	55	169
PhABC5	CCAGGAGGTCTTACTGCACT	TAGGCAGCAGTTCCAAAGGA	58	155
PhABC6	GATGAATGACACTCCGCCAG	CCTGGTTAACGCGTTGGTTA	56	152
PhABC7	CCTTAAACAGATGCGTCCCC	TGGCTTTTGGGGATCAGACA	60	168
PhABC8	TATGGCTGTGTTGGGTGCTA	TTCAACGGTCAGCATAGGGT	61	190
PhABC9	TATGGGTCCTTTGTCAGCGT	CCCTCTCTCATCTGCTTGGT	57	173
PhABC10	TTCTCCAAAGCAGCCCAAAC	AGGACGGAGAGCAAGTCAAA	61	159
PhABC11	TTCGTATGCCAGTGACACCT	CTGTGTGTGCGCTCTATGTC	54	150
PhABC12	ATCCCACAAGACCCAACCAT	TTGACCCATGCTCCAGTTCT	54	186
PhABC13	TCATTGATCCTGGCATTGCG	ACTGCCTGGCCTATGTTCTT	58	185
PhABC14	CCGCGAGACATTGGACTTTT	GGTGATGAGTTGGCTTGCAA	54	182

PhABC15	CAGAGAGTCACGGTCCAAGA	CGTCATCTCCAAGCACTTCG	58	178
PhABC16	ATGGTTCTGTTGTCGCACAC	GTCCGAGGGTCAGAAGTCTC	51	150
PhABC17	CGAGTAGCAGGATCTTGGGG	TCCCTGCCTGATGGTTTCAA	54	100
PhABCI	TTGGCGCTTATGGGATTGTG	GTCCATCAAGATTCACCGCC	59	181
PhABCII	AGAGCAGGCCAGTTCATTCT	CCAATTTCTCGGCGCATGTA	58	183
PhABCIII	TCTGGTCGAGATGGGTAAGC	CAGGGGAGGTGAAGCTGTTA	59	150
PhABCIV	GTCTCGGCTGCTACTTTTGG	TCTGGTGCCTCGAGAAACTT	58	188
PhABCV	ATTTGAGTGGACGGTAGCCA	CACTGGCATTCTCGGTCTTG	57	196

Table 3.5 Primers used for expression analysis of miRNAs targets through qRT-PCR

Name	**Forward Primer 5'-3'**	**Reverse Primer 5'-3'**	**Tm (^{0}C)**	**Product Size (bp)**
UDP-Gly	AGCATTTGGATGCAACTCAG	TGAGGAAGAGAGCTGGAGAGA	53	200
FS	GGGACAAGAATGGTGAGTGC	GCTTGGAGTTGAAGGCAATG	55	156
GAPDH	AGGTGGTGTCCAACGACTTC	CTAGTTGTCGACCGTGCTCA	60	197
WRKY	GAAGGAGGAGGAGGATTTGG	GCTCGTCTCTCCGCAATACT	58	155
PEP	CAGTTGGAAGAGGTGGTGGT	CTGGAGCGTTCTGAAACACA	55.5	146
MD	GCCGAGGTCTCGTTCGTTA	GGGGAAACTTACCAGGTCCA	56	141
MYB	TGCACAAGAGGAGGATCTGA	GGCCTCAGGAAGGTGTGAAT	52	153
PRO	ACCCAACGAAGACCAATCTG	CATTGGGCTTTCTTCCTCAA	55	188

CHAPTER 4

RESULTS

4.1 Podophyllotoxin content

Highest amount of podophyllotoxin was detected in rhizomes (3.05%) followed by roots (1.37%) and shoots (0.01%) of *P. hexandrum* (Figure 4.1). The podophyllotoxin content differed in different tissues/organs, thereby providing differential conditions to correlate the transcript abundance of molecular components (pathway genes, TFs, ABC transporters and miRNAs) with podophyllotoxin content in *P. hexandrum*. To add one more filter to further confirm the correlation between transcripts profile and metabolite content, we also quantified podophyllotoxin content in rhizomes of 11 different accessions of *P. hexandrum* collected from different geographical locations. Podophyllotoxin content ranged from 0.59% to 3.05% among different accessions (Fig. 4.2). The accession Ph5 of *P. hexandrum* showed the highest amount of podophyllotoxin (3.05%) whereas accession Ph6 contained lowest podophyllotoxin content (0.59%). The high and low podophyllotoxin content aceesions of *P. hexandrm* were selected for expression analysis through qRT-PCR. The podophylltoxin conent was not detected in shoots of these accessions.

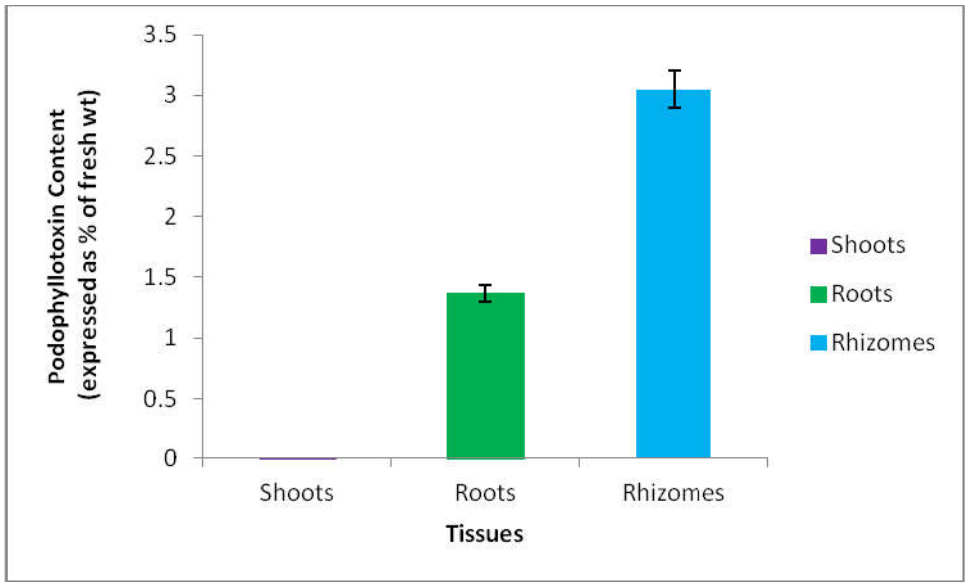


Figure 4.1 Podophyllotoxin content in different tissues of *P. hexandrum* Error bars show standard deviation (n=3)

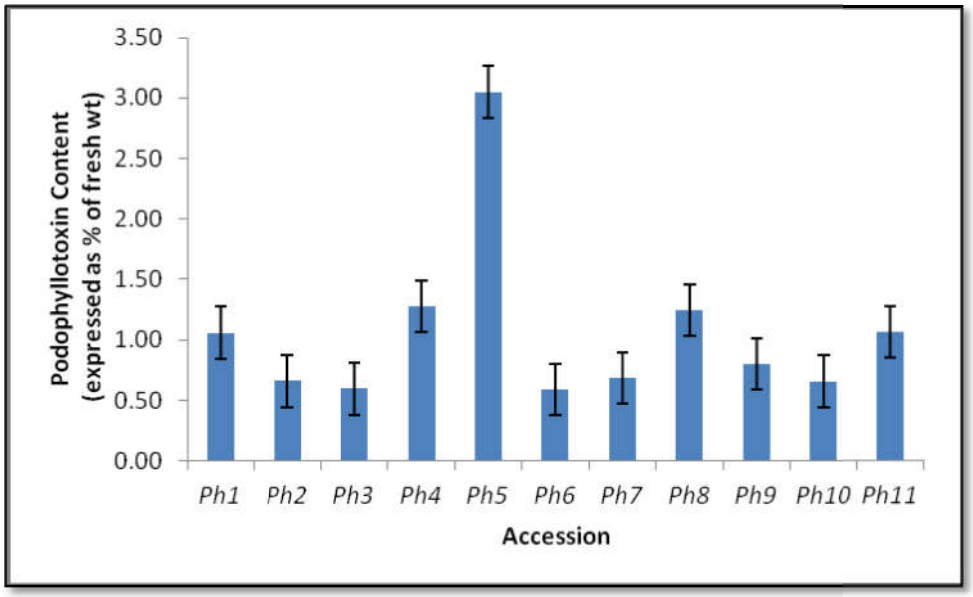


Figure 4.2 Podophyllotoxin content in rhizomes of different accessions of *Podophyllum hexandrum.* Error bars show standard deviation (n=3)

4.2 Identification of podophyllotoxin biosynthetic pathway components/genes through transcriptome mining

The whole picture of podophyllotoxin biosynthetic pathway has not yet been completely elucidated. Till date, 31 steps were known up to intermediate compound pluviatolide in podophyllotoxin biosynthesis; thereafter it was unknown until the end product podophyllotoxin. Fourteen genes namely, *DAHPS, DHQS, DHQD, COMT, EPSPS, CS, CM, PPA-AT, PD, ADT, PAL, 4CL, COMT, CCR* involved in the podophyllotoxin biosynthetic pathway were identified in the transcriptomes of *P. hexandrum* on the basis of high sequence identity (≥ 30% and E-value threshold of 10^{-05}) with the genes in other plant species. Four genes, *3DSD, QD, SK* and *PA4H* could not be ascertained due to their low sequence similarity (Figure 4.3).

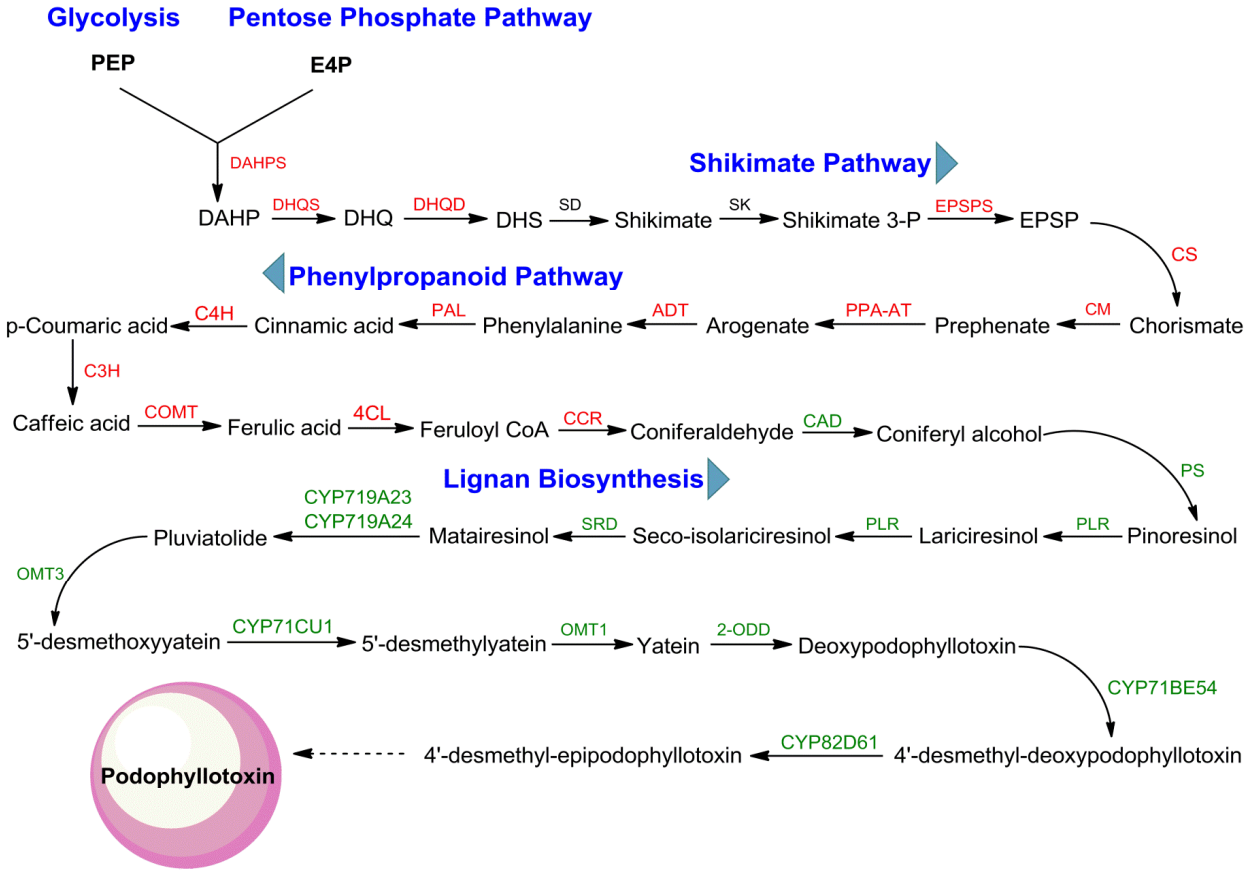


Figure 4.3 Biosynthetic pathway for podophyllotoxin in *Podophyllum hexandrum* [2, 10]. The pathway genes identified in the present work are highlighted in red color (Kumar et al [89]

4.3 Expression analysis of podophyllotoxin biosynthetic pathway genes

qRT-PCR analysis of 21 pathway genes involved in podophyllotoxin biosynthesis revealed that the transcript abundance ranged from 0.004 to 23.06 folds in shoots, roots and rhizomes of *P. hexandrum* (Figure 4.4). Multiple genes of podophyllotoxin biosynthetic pathway showed elevated levels of expression vis-à-vis podophyllotoxin content. The *SRD* gene revealed the highest expression in roots (19.46 folds) and rhizomes (23.05 folds) compared to shoots. The fold increase in transcript amount of pathway genes was in consistence with podophyllotoxin content i.e. 1.37% and 3.05% in roots and rhizomes, respectively.

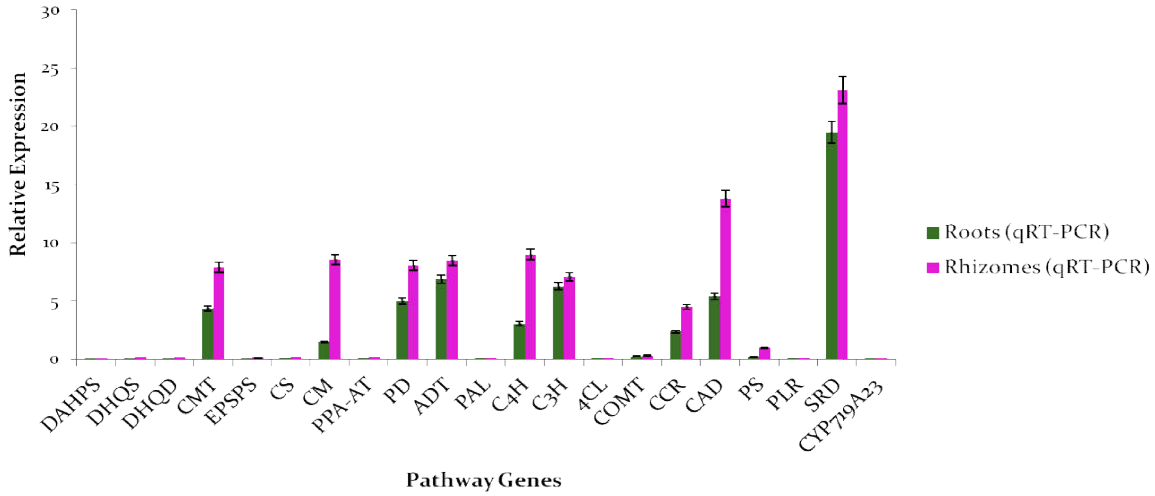


Figure 4.4 Expression status of pathway genes in roots and rhizomes w.r.t. shoots of *P. hexandrum.* Error bars show standard deviation (n=3)

4.4 Identification of TFs families in transcriptomes of *Podophyllum* species

The mining of *P. hexandrum* and *P. peltatum* transcriptomes derived from shoots and rhizomes resulted in the identification of 3477, 3517 and 3383, 3824 TFs encoding transcripts belonging to 97 and 95 TF families. The identified TFs encoding transcripts were suggested to regulate different biological processes such as photosynthesis, seed germination, disease resistance, secondary metabolites biosynthesis, etc (https://sites.google.com/site/combiogroup/data-access). We further identified TFs encoding transcripts contributing to the biosynthesis of secondary metabolites in *Podophyllum* species. Literature survey showed that eleven TFs families namely, WRKY, MYB, bZIP, bHLH, NAC, DOF, Zing finger, SPL, AP2/ERF, mTERF and TFIIIA have been implicated in the biosynthesis of secondary metabolites in various plant species [94, 95].

Out of eleven, four TFs families, WRKY, MYB, bZIP and bHLH have been suggested to contribute in lignin and terpenoid biosynthesis. Therefore, we mined transcriptomes of *Podophyllum* species for these four TFs families which revealed 82, 278, 70 and 90 transcripts in shoots and 89, 273, 72 and 91 transcripts in rhizomes of *P. hexandrum* transcriptomes and similarly, 70, 268, 48 and 92 transcripts in shoots and 58, 245, 41 and 85

transcripts in rhizomes of *P. peltatum* transcriptomes corresponding to bZIP, MYB, WRKY and bHLH TF families (Figure 4.5).

The multiple sequence alignment of TF encoding transcripts was performed using clustalW and corresponding conserved domains were identified through conserved domain database available at NCBI (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) (Figure 4.6).

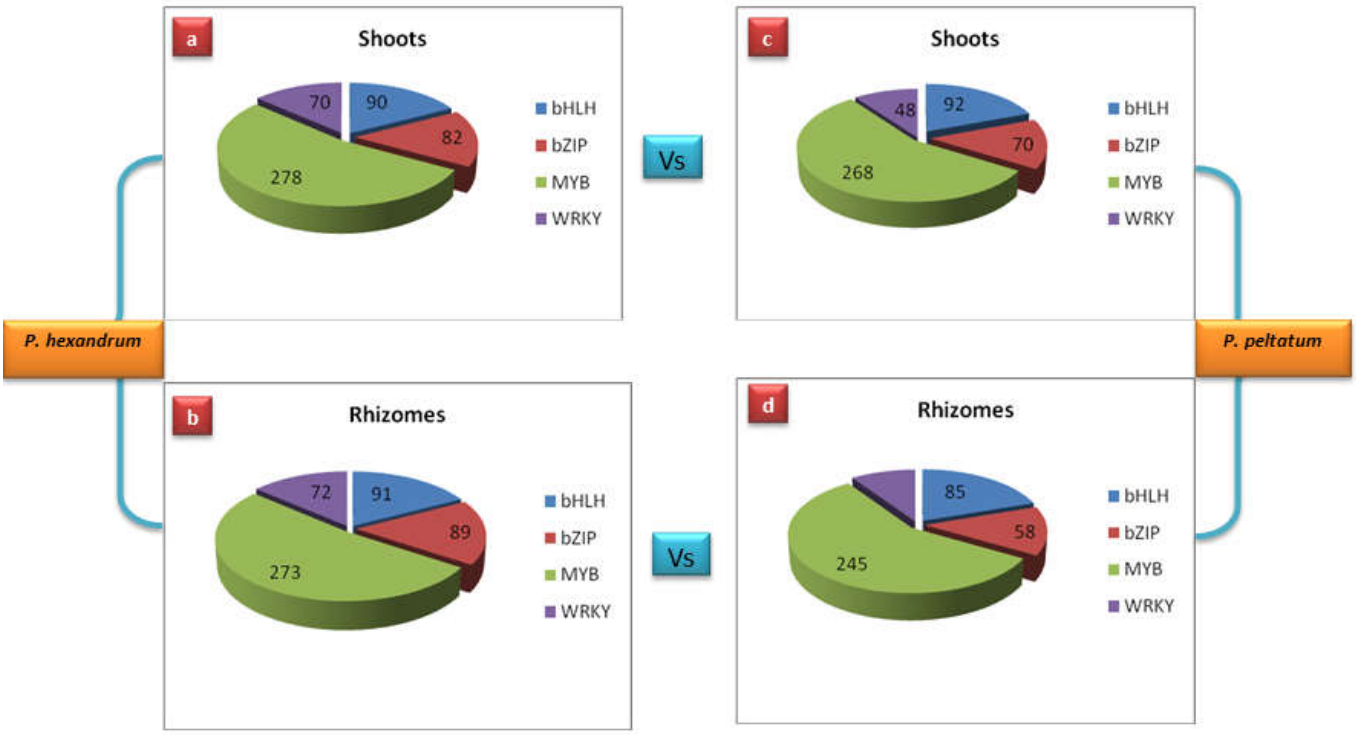


Figure 4.5 Distribution of four TFs families in shoots and rhizomes transcriptomes of *Podophyllum* species a) & b) *P. hexandrum;* c) & d) *P. Peltatum*

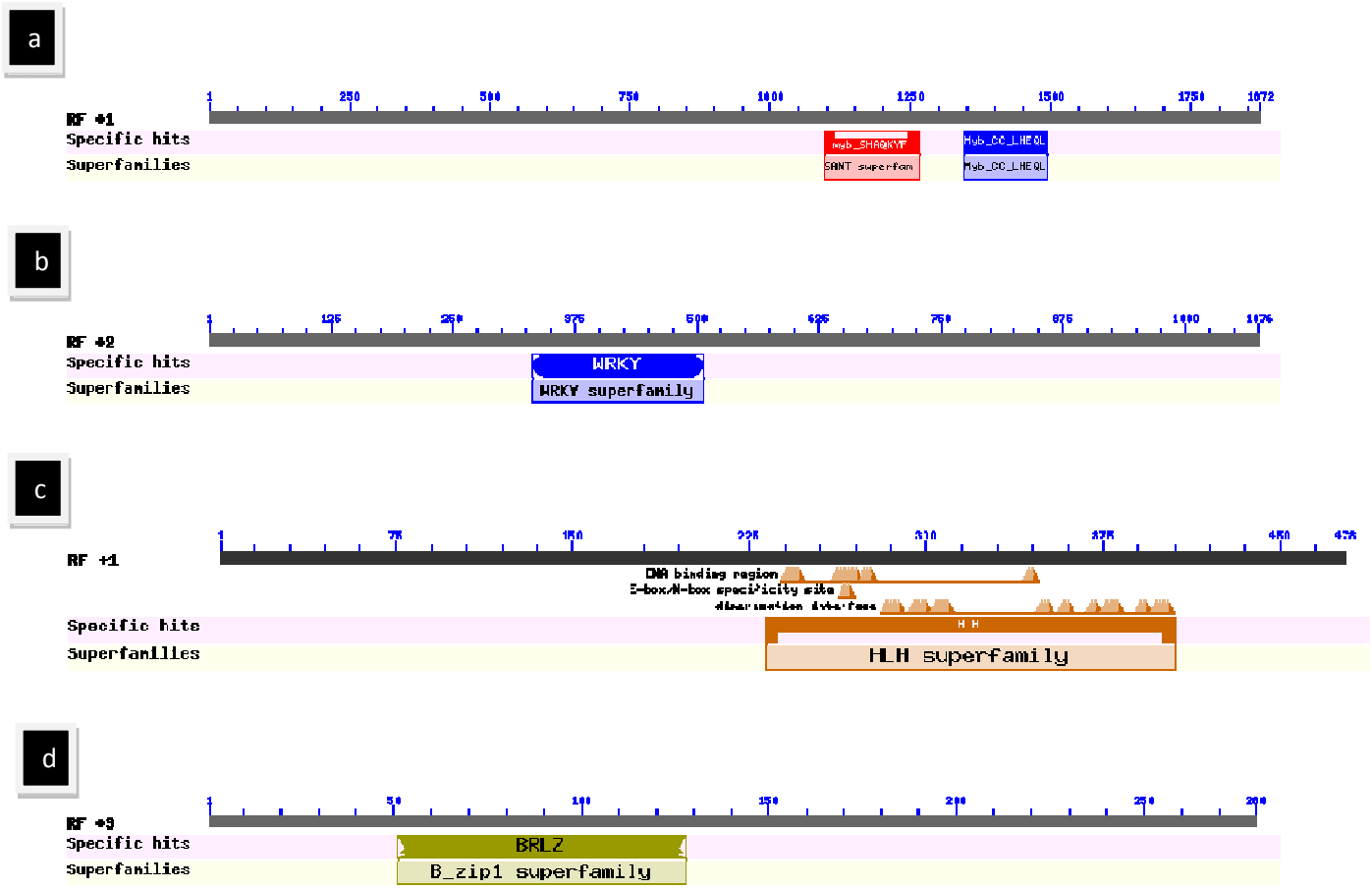


Figure 4.6 Conserved domains in different classes of TFs a) MYB b) WRKY c) bHLH d) bZIP

4.5 Identification of TFs families uniquely associated with podophyllotoxin content in *P. hexandrum* and *P. peltatum*

The podophyllotoxin content is negligible in leaves of *P. hexandrum* (0.01%) as compared to the leaves of *P. peltatum* (2.05%) whereas roots/rhizomes of *P. hexandrum* contain 2.51% podophyllotoxin compared to 0.25% in roots/rhizomes of *P. peltatum* [39]. The transcriptomes of these differential podophyllotoxin content tissues thus provided a good resource to mine for TFs families associated with podophyllotoxin biosynthesis. Two unique transcripts, medp_podpe_41091 and medp_podpe_2547 coding for bHLH and MYB/SANT TFs were detected in the leaves of *P. peltatum*, but absent in leaves of *P. hexandrum*. Similarly, two unique transcripts coding for bZIP and MYB TFs were identified only in the rhizomes of *P. hexandrum* as compared to *P. peltatum.* Moreover, 218 transcripts encoding WRKY, MYB, bZIP and bHLH TFs were common in the shoots and rhizomes transcriptomes of *P. hexandrum* and *P. peltatum*, respectively.

4.6 *In-silico* transcript abundance of TFs

The transcript abundance for common and unique TFs was checked in shoots and rhizomes transcriptomes of *P. hexandrum* and *P. peltatum* through *in silico*-based RSEM approach. RSEM calculates the transcript abundance of RNA-Seq data in terms of FPKM value. The FPKM values corresponding to respective TFs ranged from 0.0014-12.01 and 0.014-1162.01 in shoots and rhizomes of *P. hexandrum* and *P. peltatum*, respectively. The RSEM analysis revealed that transcripts medp_podhe_24032 and medp_podhe_111874 corresponding to bHLH and bZIP TFs have the highest FPKM value of 8.62 and 12.01 corresponding to 0.01% and 2.51% podophyllotoxin content in shoots and rhizomes of *P. hexandrum*. In case of *P. peltatum,* transcripts medp_podpe_6198 and medp_podpe_4999 showed highest transcript abundance with 1162.01 and 285.73 FPKM values corresponding to bZIP TF family in the shoots and rhizomes. On analyzing the common transcription factors, the transcript abundance ranged from 0.29 to 285.73 in shoots and 0.12 to 1162.01 in rhizomes of *P. hexandrum*. Similarly, transcript abundance ranged from 0.16 to 184.09 in shoots and 0.26 to 173.66 in rhizomes of *P. peltatum.* On comparative tissue profiling of *P. hexandrum* and *P. peltatum*, it was found that the transcripts medp_podhe_7127 in shoots coding for WRKY TF in *P. hexandrum* and medp_podpe_6198 in shoots coding for bZIP TF in *P. peltatum* showed the highest transcript abundance with FPKM value of 184.09 and 285.73. Similarly, on comparing the rhizomes of both *Podophyllum* species, it was found that the transcripts medp_podhe_8048 coding for MYB TF in rhizomes of *P. hexandrum* showed the highest transcript abundance with FPKM value of 173.66 and medp_podpe_4999 coding for bZIP in rhizomes of *P. peltatum* showed highest transcript abundance with FPKM value of 1162.0. It was observed that the transcript abundance of common TFs was higher in rhizomes as compared to shoots, which was in conjunction with higher metabolite content in rhizomes (2.51%) compared to shoots (0.01%).

4.7 Mining of *cis*-regulatory elements

The putative *cis*-regulatory elements were identified in the promoter regions of podophyllotoxin pathway genes in order to deduce the possible binding sites of transcription factors in *P. hexandrum*. The *in-silico* analysis revealed that MYB and WRKY transcription

factor binding sites were commonly present in the promoter regions of podophyllotoxin pathway genes i.e. *CAD, C4H, DPO, PCH, PAL, 4CCL, PLR* and *SD*. The identified transcription factors are, therefore, suggested to be regulating podophyllotoxin biosynthesis through binding with the sequence elements in the promoters of target pathway genes in *P. hexandrum*.

4.8 Experimental validation and differential expression analysis of TFs using qRT-PCR in *P. hexandrum*

The transcripts of four TFs were validated through qRT-PCR in *P. hexandrum* tissues differing for podophyllotoxin content. The partial cloned fragments of TFs matched transcriptome contigs sequences perfectly (100%), therefore, again validating the transcriptomes mining procedures employed in the present study. The partial sequences of TFs (212-434bp) have been submitted to the GenBank under accession numbers KP299959 to KP299961. These sequences were utilized for designing primers for gene expression analysis. *In-silico* transcript abundance values of respective TFs ranged between 0.0014-12.01 in shoots and rhizomes of *P. hexandrum* (Figure 4.7). In order to ascertain the relevance of FPKM-based expression values obtained through RSEM analysis, qRT-PCR was carried out in shoots and rhizomes of *P. hexandrum.* Increase in transcript abundance with 1.09-folds for MYB, 0.62-folds for WRKY and 0.26-folds for bHLH were observed in rhizomes with respect to shoots of *P. hexandrum.* The bZIP class of TFs showed higher transcript abundance with 19.60-folds in rhizomes compared to shoots of *P. hexandrum* (Figure 4.8). The 3 TFs belonging to bZIP, MYB and WRKY families showed down-regulation in transcript amounts in shoots compared to rhizomes, except for bHLH TF with higher amount of transcript in shoots compared to rhizomes.

Furthermore, the transcript abundance of identified transcription factors was further assessed between high (Ph5) (3.05%) versus low (Ph6) (0.59%) podophyllotoxin content accessions of *P. hexandrum*. This comparative expression analysis further confirmed the expression status of transcription factors in rhizomes of *P. hexandrum* (Figure 4.9). Transcript abundance of TFs obtained through two different platforms i.e. RSEM and qRT-PCR resulted in the identification of potential TFs possibly regulating podophyllotoxin biosynthesis in *P. hexandrum*. The Pearson's correlation coefficient was performed to find out relationship between podophyllotoxin content and transcript abundance of TFs. The transcript abundance

pattern of TFs (bZIP, MYB, WRKY and bHLH) showed positive correlation with podophyllotoxin content which suggested their role in regulating and controlling the podophyllotoxin biosynthesis in *P. hexandrum*. The bZIP TFs showed highest correlation (≥ 0.99) with podophyllotoxin biosynthesis.

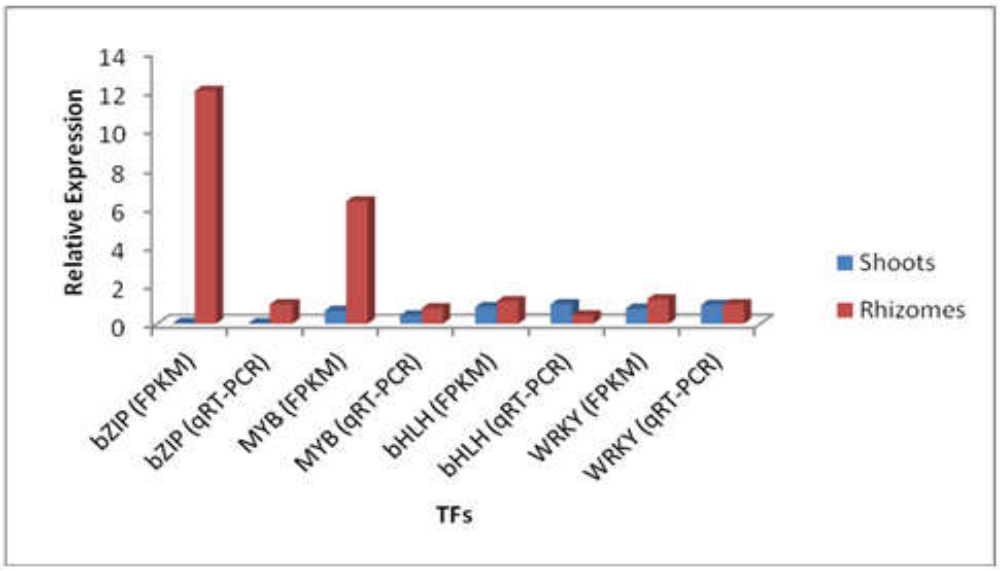


Figure 4.7 Relative expression analysis of TF genes in shoots and rhizomes of *P. hexandrum* based upon the data obtained by fragments per kilobase of exon per million fragments mapped (FPKM) values and validated by quantitative real time-polymerase chain reaction (qRT-PCR)

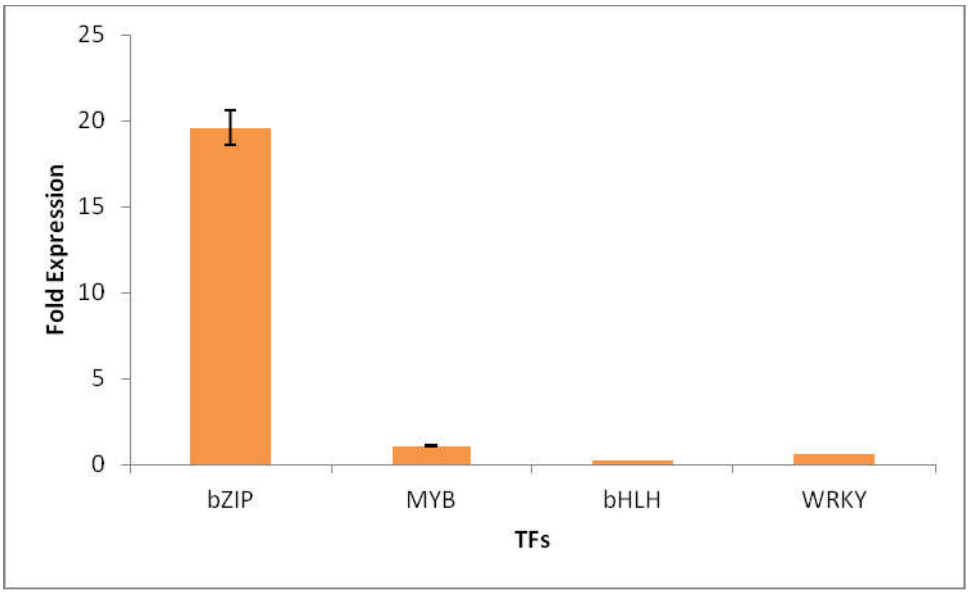


Figure 4.8 Fold increase in transcript abundance of TF genes representing different TF families in rhizomes w.r.t. shoots of *P. hexandrum*

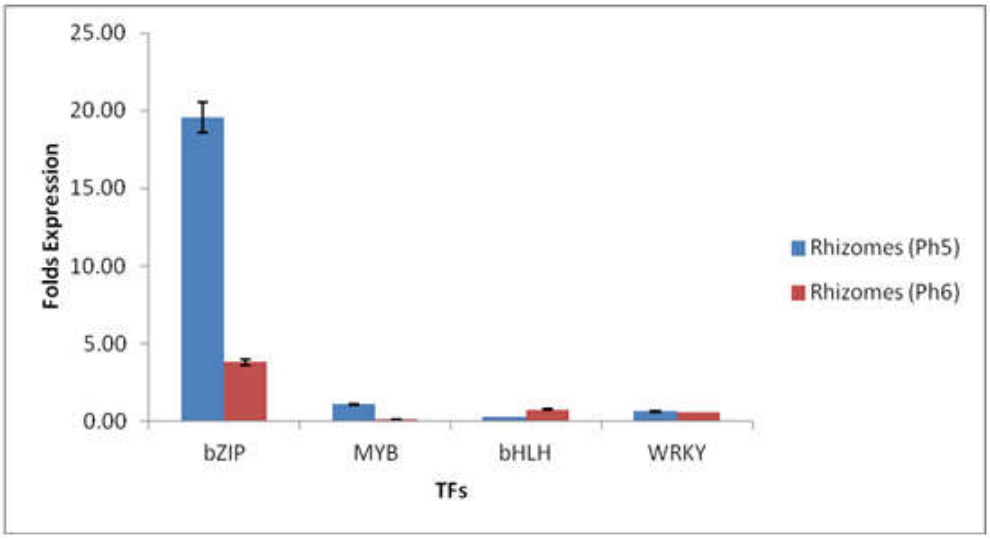


Figure 4.9 Expression status of transcription factors in rhizomes of high (Ph5) versus low (Ph6) podophyllotoxin content accessions of *P. hexandrum*

4.9 Mining ABC transporters in transcriptomes of *P. hexandrum* and their phyllogenetic analysis

Total 252 putative ABC transporters were identified and classified from the transcriptome of *P. hexandrum* by searching protein sequences corresponding to full complement of ABC transporters through Pfam database (Figure 4.10). The identified transporters genes were suggested to be involved in the transport of primary and secondary metabolites in *P. hexandrum.* We further identified orthologous of five functionally validated transporters genes, three from *Arabidopsis thaliana* (*PhABCI, PhABCII, PhABCIII*), and one each from *Vitis vinifera* (*PhABCIV*) and *Nicotiana plumbaginifolia* (*PhABCV*) in the transcriptomes of *P. hexandrum*. These transporter genes are assumed to play a role in the transport of secondary metabolites in *P. hexandrum* as has been demonstrated in other plant species. Phylogenetic analysis of identified ABC transporters was performed through MEGA6 software using functionally characterised ABC transporters retrieved from model plant species such as *Arabidopsis thaliana*, *Nicotiana plumbaginifolia, Zea mays* and *Medicago truncatula.* The ABC transporters genes showed high homology with well characterised

genes in model plants that further confirms the biological functions of identified ABC transporters in *P. hexandrum*.

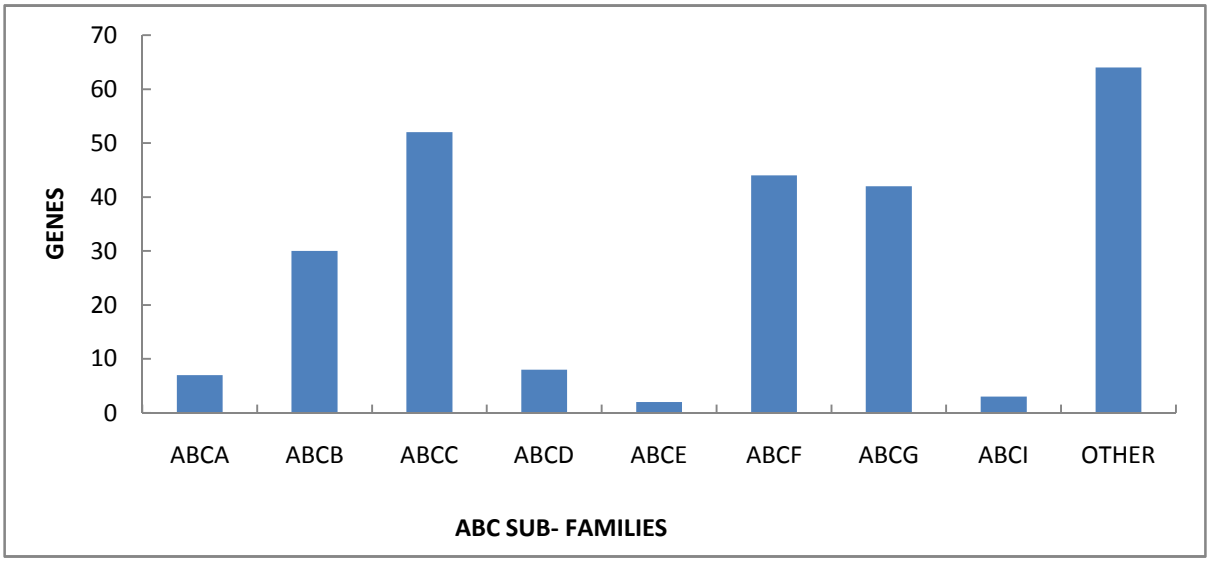


Figure 4.10 Distribution of ABC transporter family into 9 sub-families ABCA, ABCB, ABCC, ABCD, ABCE, ABCF, ABCG, ABCI and others in transcriptomes of *P. hexandrum*

4.7 Identification of common and unique ABC transporters

Comparative analysis was done for identification of common and unique ABC transporters genes in transcriptomes from shoots and rhizomes of *P. hexandrum.* 249 ABC transporters genes were found commonly present in shoots and rhizomes transcriptomes while 16 transporters genes were uniquely present. Out of 16, 12 genes were present only in shoots and 4 genes in rhizomes of *P. hexandrum*. The common ABC transporters genes in shoots and rhizomes transcriptomes of *P. hexandrum* may be responsible for the transport of metabolites existing in both the tissues.

4.8 Transcript abundance of ABC transporters in transcriptomes

Transcript abundance was determined for 252 ABC transporters through RSEM approach using in-house developed perl scripts. Transcript abundance for commonly present ABC transporters genes ranged from 0.41 to 252.71 in shoots and 0.06 to 402.88 FPKM value in rhizomes. The transcript abundance for uniquely present genes in shoots and rhizomes of *P. hexandrum* was also analysed, where transcript abundance ranged from 0.00 to 11.96 in rhizomes and from 0.00 to 0.71 in shoots. Transcripts medp_podhe_20101112|14082 showed highest expression in shoots (252.71 FPKM value) and rhizomes (402.88 FPKM value) of *P.*

hexandrum. The ABC transporters genes with higher transcript abundance were further selected for validation through qRT-PCR to assess their association with podophyllotoxin accumulation (Figure 4.11).

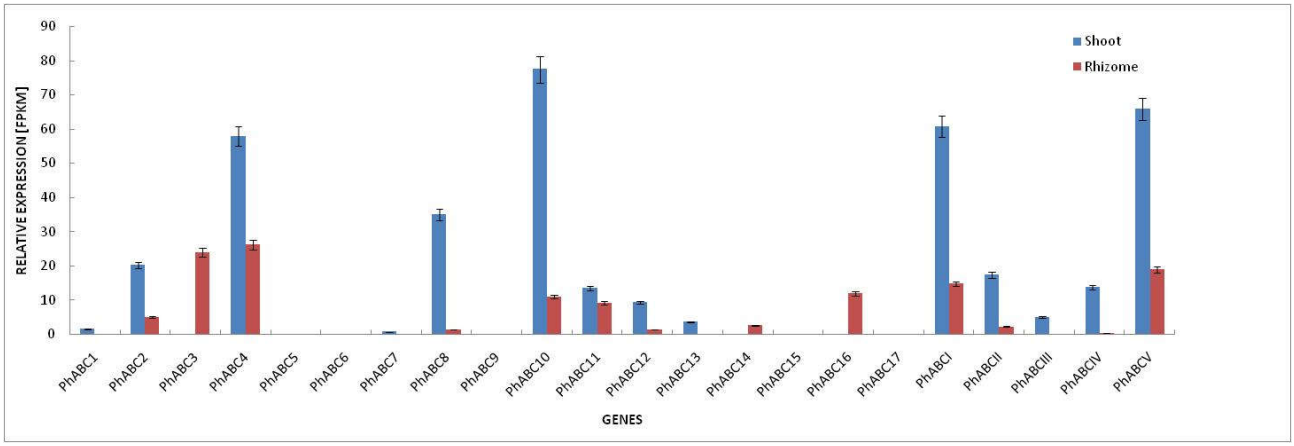


Figure 4.11 Relative expression of ABC transporters genes in shoots versus rhizomes of *P. hexandrum* analyzed by FPKM based approach. *Error bars* indicate percentage error (5%)

4.9 Experimental validation and differential expression analysis of ABC transporters using qRT-PCR

Seventeen ABC transporters genes were selected on the basis of higher transcript abundance value for validation through qRT-PCR in shoots, roots and rhizome tissues varying for podophyllotoxin contents. The qRT-PCR-based relative transcript abundance of transporters genes ranged from 0.002 to 1.673 in shoots, roots and rhizomes of *P. hexandrum* (Figure 4.12). Gene expression pattern observed through two different approaches was in agreement with each other. The *PhABC3*, *PhABC6* and *PhABCIV* transporter genes showed highest transcript abundance with 1.673, 1.25 and 1.0 and in shoots, roots and rhizomes of *P. hexandrum*. Most of the ABC transporters genes showed higher transcript abundance in shoots suggesting their role in transporting macro and micro molecules in shoots. Moreover, the expression of five functionally characterized transporters genes (*PhABCI*, *PhABCII*, *PhABCIII*, *PhABCIV* and *PhABCV*) from *Arabidopsis thaliana, Vitis vinifera* and *Nicotiana plumbaginifolia* were also evaluated through qRT-PCR across different tissues/organs of *P. hexandrum.* Expression analysis revealed that *PhABCII*, *PhABCV* showed higher transcript abundance in shoots and *PhABCIII* in rhizomes of *P. hexandrum* (Figure 4.13).

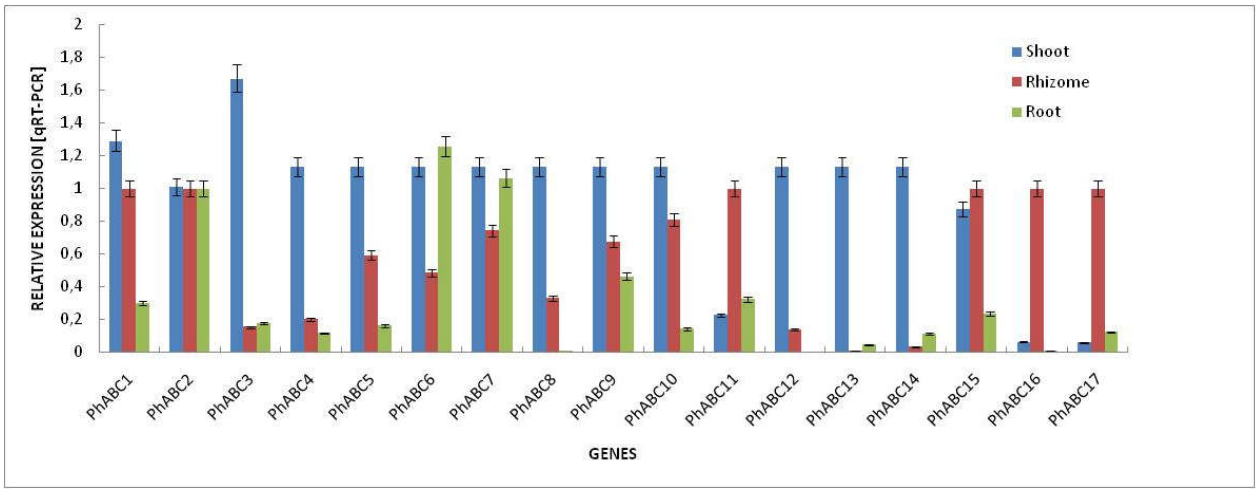


Figure 4.12 Relative expression of ABC transporters genes in shoots, roots and rhizomes of *P. hexandrum* analyzed by qRT-PCR. *Error bars* indicate percentage error (5%)

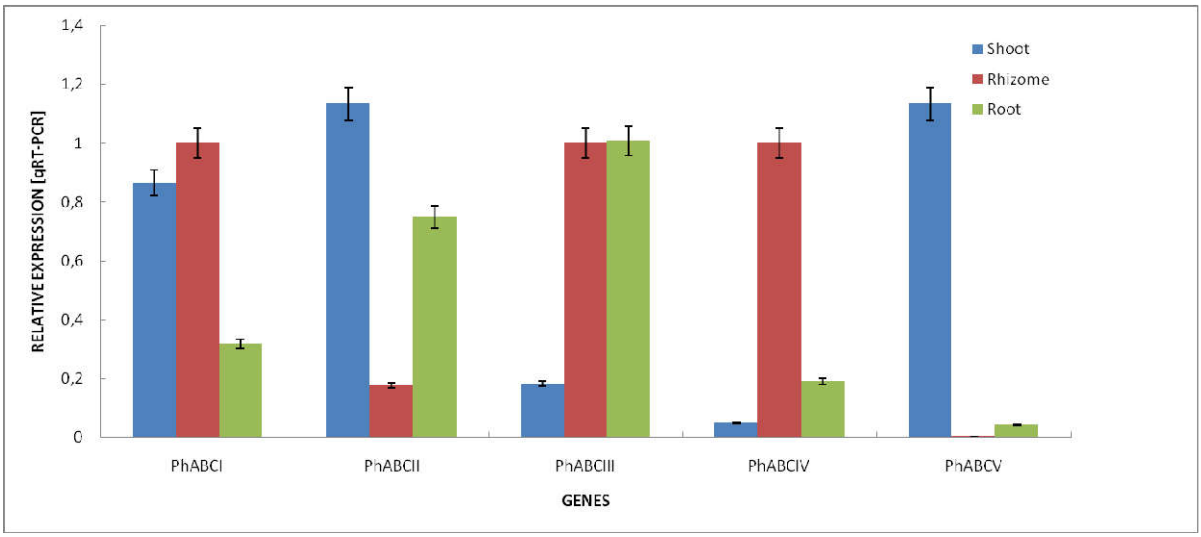


Figure 4.13 Relative expression of orthologous ABC transporters in shoots, roots and rhizomes of *P. hexandrum* analyzed by qRT-PCR; *PhABCI*, *PhABCII*, *PhABCIII* from *Arabidopsis thaliana*; *PhABCIV* from *Vitis vinifera* and *PhABCV* from *Nicotiana plumbaginifolia*

4.10 Identification of miRNAs and their mRNAs targets in transcriptomes

PMRD (plant microRNA database) (http://bioinformatics.cau.edu.cn/PMRD/) was used to identify miRNAs from transcriptomes of *P. hexandrum* [96]. Precursor miRNAs sequences were mined and used for identifying potential miRNAs using in-house perl scripts and the criteria are mentioned in the materials and method section. Total 279 putative miRNAs were

identified from transcriptomes that were conserved across various plant species ranging in size from 18 to 22 nucleotides with perfect complementarity (100%) to homologous sequences. Out of 279, 6 miRNAs were selected for further validation on the basis of their possible role in regulation of podophyllotoxin biosynthesis (Figure 4.14).

Total 375 targets were found for 6 conserved miRNAs in *P. hexandrum* having perfectly or near-perfect complementarity to their mRNA targets using psRNATarget tool. The identified potential targets were involved in various biological processes including, plant defense response such as biotic and abiotic stresses, secondary metabolisms, growth and development, signal transduction, host–pathogen interaction, etc. The miRNA miR2673a/b has maximum number of targets (186 targets). The functional analysis of identified putative targets of miRNAs was performed through KAAS (KEGG automatic annotation server) (http://www.genome.jp/tools/kaas/) using default settings [75]. KAAS provides information pertaining to metabolism, genetic information processing and signalling and cellular processes. Total 26 targets were found for secondary metabolites biosynthesis through KAAS. Out of 26, 8 targets were selected for further validation and expression profiling on the basis of their association with podophyllotoxin biosynthesis.

4.11 Functional analysis

The functional analysis of identified putative targets of miRNAs was performed through KAAS (KEGG automatic annotation server) (http://www.genome.jp/tools/kaas/) [75] using default settings. KAAS provides information pertaining to metabolism, genetic information processing and signalling and cellular processes. Total 26 targets were found for secondary metabolites biosynthesis through KAAS. Out of 26, 8 targets were selected for further validation and expression profiling on the basis of their association with podophyllotoxin biosynthesis.

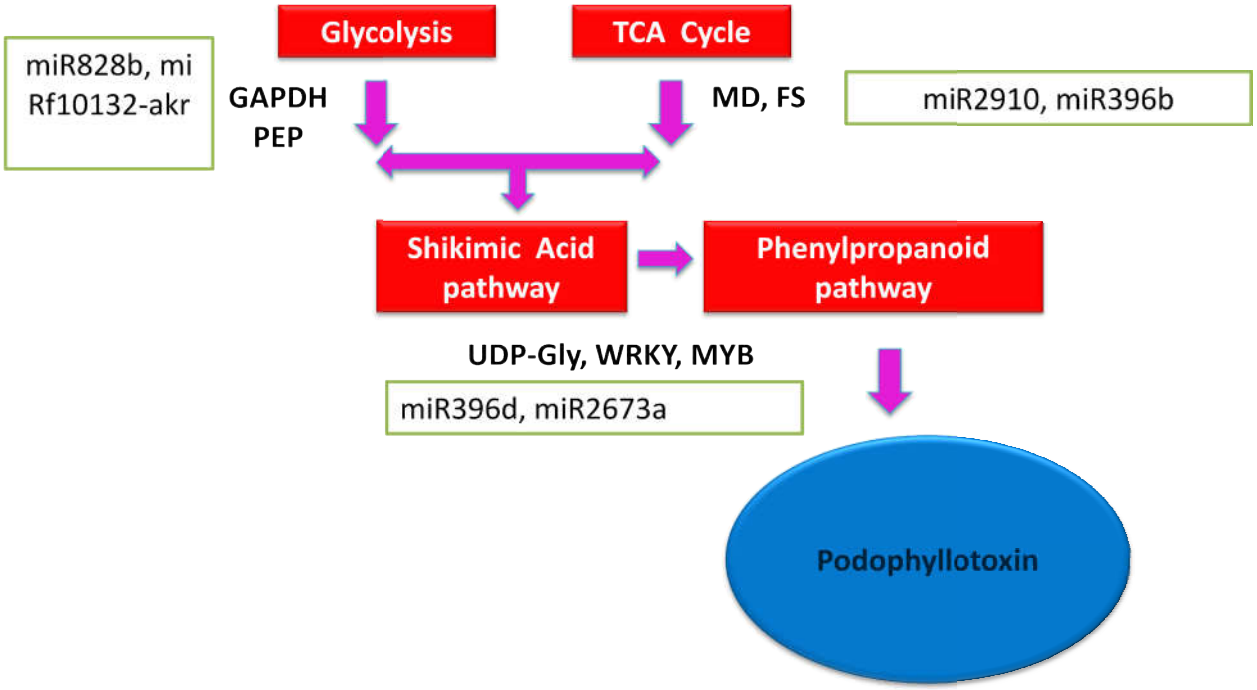


Figure 4.14 An outline of podophyllotoxin biosynthetic pathway showing miRNAs and their mRNA targets

4.12 Experimental validation and differential expression profiling of miRNAs using stem-loop qRT-PCR

The identified 6 conserved miRNAs from transcriptomes were selected for validation and expression analysis through stem-loop qRT-PCR on the basis of their association with secondary metabolites biosynthesis in *P. hexandrum.* The transcript abundance of miRNAs ranged from 0.03 to 747-folds in roots and rhizomes of *P. hexandrum*. miR396b showed highest expression in roots with 2-fold and 747-fold in rhizomes of *P. hexandrum*. miR828b showed least expression in roots with 0.03-fold and miRf10132-akr 0.27-fold in rhizomes (Figure 4.15).

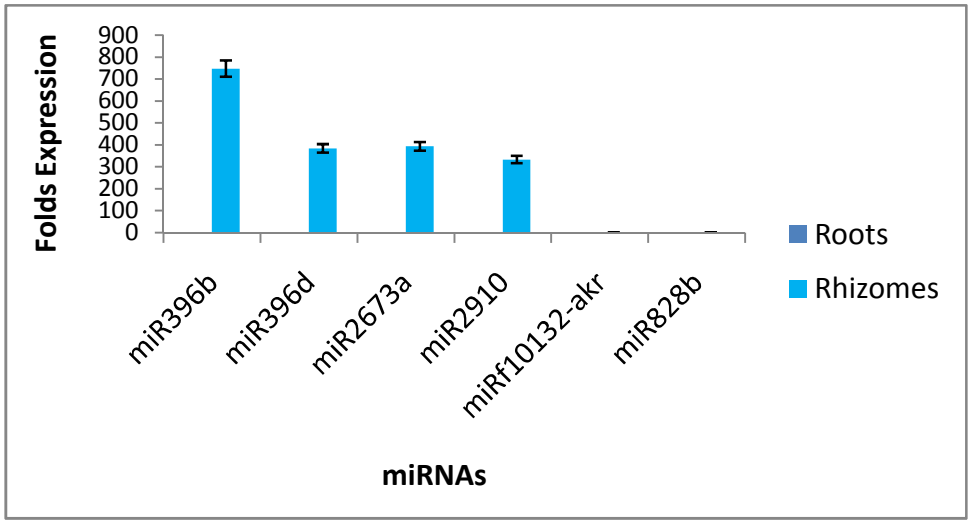


Figure 4.15 Compartartive expression profile of miRNAs in roots and rhizomes w.r.t. shoots of *P. hexandrum*

4.13 Validation and expression analysis of miRNAs targets by qRT-PCR

Out of 375 predicted targets, eight targets (UDP glycosyltransferase (*UDP-Gly*), flavnol synthase (*FS*), glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*), peroxidise (*PRO*), malate dehydrogenase (*MD*), phosphoenolpyruvate carboxylase (*PEP*), WRKY and MYB transcription factors were selected for validation through qRT-PCR on the basis of their association with secondary metabolites biosynthesis. The expression profiles of target genes and their corresponding miRNAs could provide valuable information pertaining to regulation of secondary metabolites biosynthesis. The relative expression of target genes mediated by miRNAs ranged from 0.4 to 26 in roots and 0.1 to 28 in rhizomes of *P. hexandrum*. The target gene UDP-Gly showed higher transcript abundance with 305-fold in roots and WRKY gene with 130-fold in rhizomes of *P. hexandrum*. The PEP gene showed lowest expression with 0.4 fold and 0.1-fold in roots and rhizomes of *P. hexandrum* (Figure 4.16).

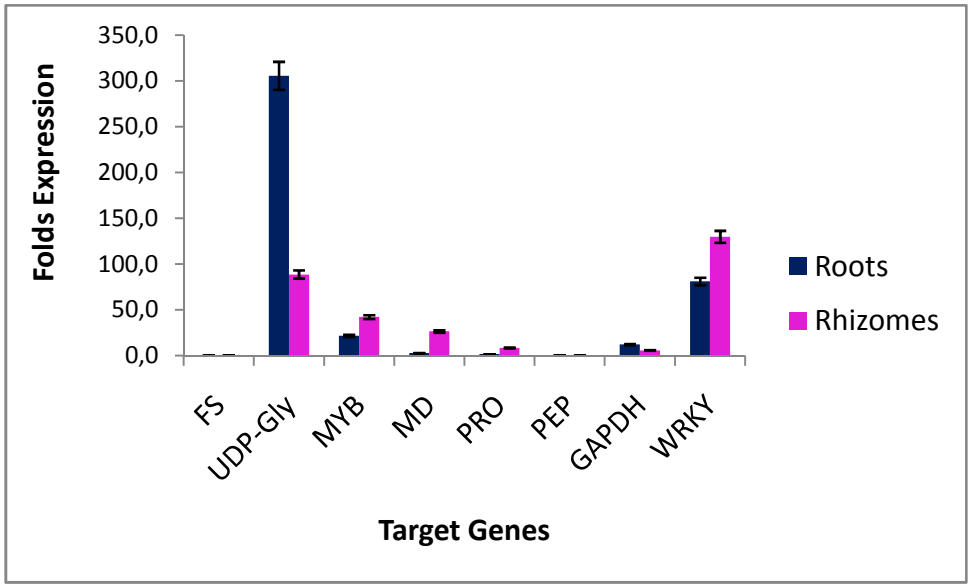


Figure 4.16 Compartartive expression profile of miRNAs mediated target genes in roots and rhizomes w.r.t. shoots of *P. hexandrum*

CHAPTER 5

DISCUSSION

The demand for podophyllotoxin has been increasing day-by-day due to its usage in the manufacture of anti-cancer drugs. The limited raw material and low (0.3%) *in-vitro* production of podophyllotoxin in cell cultures necessitates the development of alternate production routes/sources to meet the industrial demands. Research on molecular components (pathway genes, TFs, ABC transporters and miRNAs) in plant species has off late opened a new window to enhance metabolites content.

5.1 Validation and expression analysis of podophyllotoxin biosynthesis pathway genes

In order to gain better understanding on the molecular basis of podophyllotoxin biosynthesis, quantitative expression of 21 pathway genes was studied in different tissues of *P. hexandrum* by using transcript abundance analysis through FPKM and further validation through qRT-PCR. The FPKM values of pathway genes ranged from 1.15 to 52.09 in rhizome transcriptome of *P. hexandrum*. The qRT-PCR based relative transcript abundancy of pathway genes ranged from 0.004 to 23.06 in shoots, roots and rhizomes of *P. hexandrum* (Figure 4.4). Out of 21 genes retrieved from GenBank Database (7 genes) and *P. hexandrum* transcriptomes (14 genes), 5 genes *SRD, C3H, CAD, CCR* and *C4H* showed higher expression in roots and rhizomes of *P. hexandrum* compared to shoots which was in agreement with the previous reports [11,13]. In addition, 4 genes *PD, CM*, *COMT* and *ADT* also showed higher expression in roots and rhizomes corresponding to podophyllotoxin content (Figure 4.4) The fold increase in transcript abundance was 4.98 and 8.04 for *PD*, 6.24 and 7.05 for *C3H*, 1.49 and 8.53 for *CM*, 4.35 and 7.90 for *COMT*, 2.37 and 4.48 for *CCR*, 3.05 and 8.98 for *C4H*, 5.38 and 13.76 for *CAD*, 6.88 and 8.44 for *ADT*, respectively in roots and rhizomes w.r.t. shoots of *P. hexandrum* (Figure 4.4). The expression of *SRD* gene was highest among all genes in the roots (19.46 folds) and rhizomes (23.05 folds) compared to shoots. The fold increase in transcript amount of pathway genes analysed by qRT-PCR was in consistence with the podophyllotoxin content i.e. 1.37% and 3.05% in roots and rhizomes, respectively.

These observations indicated that podophyllotoxin biosynthesis occurs in roots and rhizomes of *P. hexandrum*. Our results were in concurrence with the previous reports where metabolite content was positively correlated with the expression pattern of biosynthetic pathway genes studied for picrosides in *P. kurroa*, catechins in *C. sinensis*, shikonin in *A. euchroma* and aconites in *A. heterophyllum* [88, 97].

The expression of *SRD* gene was highest among all genes in the roots (19.46 folds) and rhizomes (23.05 folds) compared to shoots. This gene is well known to catalyze secoisolariciresinol into matairesinol, a central precursor of podophyllotoxin biosynthesis [9]. Lan et al. [12] reported that *SRD* gene had maximum expression in root and rhizome tissues contributing to podophyllotoxin production in *Dysosma tsayuensis*. The transcript level of pathway genes such as *COMT, CM, PD, ADT, C4H, C3H, CCR, CAD* and *SRD* correlated positively with the podophyllotoxin content as observed through Pearson's correlation coefficient values (≥0.99). Such values suggested a significant agreement between the transcript level and podophyllotoxin content. The higher expression of multiple genes in podophyllotoxin biosynthetic pathway may be due to the common regulation through regulatory components such as transcription factors, microRNAs (miRNAs) and *cis*-regulatory elements. The expression of multiple genes of MIA pathway in *C. roseus* was attributed to a common regulator methyl jasmonate-responsive AP2 domain transcription factor, ORCA-3. Furthermore, relatively higher expressions of *COMT, C4H, CCR, CAD* genes was also reported in *Arabidopsis thaliana* for the production of lignin [25]. Previous studies had suggested that *C4H, CAD, CCR* play crucial role in the regulation of phenylpropanoid pathway leading to podophyllotoxin biosynthesis. Multiple genes of flavonoid pathway such as *C4H, CHS, CHI* and *FLS* showed increased transcript in relation to rutin content in *Fagopyrum* spp. [98]. Similarly, multiple genes such as *DXPS, ISPD, PMK* and *MECPS* showed elevated expression w.r.t. picrosides content in different tissues of *P. kurroa* [23]. Higher expression of multiple genes was also observed in rhizomes of ginger (*Zingiber officinale Rosc*.), turmeric (*Curcuma longa* L.), bamboo (*Phyllostachys praecox*), red rice (*Oryza longistaminata*) and *Sorghum propinquum* contributing to the biosynthesis of terpenoids [99, 100].

5.2 Expression analysis of transcription factors vis-à-vis podophyllotoxin content

Podophyllotoxin biosynthesis occurs in different tissues of *Podophyllum* species through the contribution of phenylpropanoid/shikimic acid pathway which is controlled by regulatory components, particularly TFs. TFs are fascinating vital elements controlling the diverse physiological, biochemical and biological processes including growth, development, senescence, seed germination, environmental stress responses, etc. by modulating gene expression. TFs are DNA-binding specific proteins that interact with *cis*-acting elements in promoter regions of target genes, thus regulating the gene expression [52]. It's very crucial to know the status of TFs for better understanding the biosynthesis of podophyllotoxin. Availability of transcriptomes of *Podophyllum* species provided a platform to identify and characterize candidate TFs families in shoots and rhizomes of *Podophyllum* species to infer their role in regulation of podophyllotoxin biosynthesis [10]. An extensive literature survey revealed that four TF families namely, bZIP, MYB, WRKY and bHLH regulate the phenylpropanoid/shikimic acid pathway in various plant species (Table 5.1). These TF families might also be associated with the regulation of podophyllotoxin biosynthesis in *Podophyllum* species. Therefore, we focused on identifying these four TFs in shoots and rhizomes of *P. hexandrum* in reference to podophyllotoxin biosynthesis. We mined transcriptomes of shoots and rhizomes of *P. hexandrum* in order to identify, scrutinize, validate and evaluate the transcript abundance of transcripts/genes that encode bZIP, MYB, WRKY and bHLH TF families. Transcript abundance of respective TF families was evaluated by *in silico* FPKM and qRT-PCR-based transcript abundance methods. The FPKM values of TFs ranged between 0.0014-12.01 and 0.014-1162.01 in transcriptomes corresponding to shoots and rhizomes of *P. hexandrum* and *P. peltatum*, respectively. The qRT-PCR based relative transcript abundance of TF encoding transcripts ranged from 0.55 to 19.60-folds in rhizomes w.r.t. shoots of *P. hexandrum* (Figure 4.8). Gene expression pattern observed through two different platforms was found to be in conjunction with each other (Figure 4.7).

To ascertain whether the elevated levels of transcripts of TF families is only affecting the biosynthesis of podophyllotoxin in rhizomes/roots compared to shoots where podophyllotoxin content is almost negligible. We applied two-step refinement process where, the fold increase in transcript abundance of particular TFs was evaluated by comparing the transcripts abundance in rhizomes w.r.t shoots of *P. hexandrum* accessions varying for

podophyllotoxin content. High (Ph5; 2.51%) and low (Ph6; 0.59%) podophyllotoxin content accession of *P. hexandrum* were selected for comparative expression analysis. The transcripts corresponding to TFs families (MYB, WRKY, bHLH and bZIP) showed increase in transcript abundance with 1.09, 0.62, 0.26 and 19.60-folds in rhizomes compared to shoots of *P. hexandrum* (Figure 4.8). To further validate the expression status in relation to podophyllotoxin content, we compared the transcript abundance of TFs in rhizomes of high vs low podophyllotoxin content accessions of *P. hexandrum*. The expression of bZIP, MYB and WRKY decreased to 0.04-15.78-folds between rhizomes of high versus low podophyllotoxin content (Figure 4.9). There was overall decrease in the expression level of transcription factors when compared between rhizomes of high versus low podophyllotoxin content accessions (Fig. 4.9). The transcript abundance of bZIP decreased from 19.60-fold to 3.82-fold. Similarly, the expression of MYB and WRKY decreased from 1.09-fold to 1-fold and 0.26-fold to 0.04-fold, respectively in rhizomes of high versus low podophyllotoxin content accessions. The decrease in transcript abundance was due to the fact that comparison has been made for the same tissue/organ rather between two different organs (rhizomes versus shoots) and, therefore, the expression status of genes is influenced not only due to variation in contents of metabolites but also due to development pattern of tissues/organs. The differences in transcript abundance for identified transcription factors between high versus low podophyllotoxin content accessions reflects their association as per inherent genetic differences between accessions solely for podophyllotoxin contents. The transcript abundance of transcription factors was in conjunction with the metabolites content as revealed through comparative expression analysis between high versus low podophyllotoxin content accessions. Furthermore, a strong positive correlation was obtained between transcript abundance of TF families and podophyllotoxin content as observed by Pearson's correlation coefficient value (≥0.99). Expression profiling of genes in relation to metabolites content in high vs low metabolite content accessions of various plant species such as *Aconitum heterophyllum*, *Picrorhiza kurroa*, *Jatropha curcus*, etc. was studied previously, where metabolites contents were correlated with the expression profiles of biosynthetic pathway genes [101, 102, 103].

The MYB TFs were the largest and mainly involved in controlling biological processes, including in biotic and abiotic responses, metabolism, defense, development, etc. A number of MYB transcription factors were known in various plant species, including maize, petunia, snapdragon, grape, apple and *Arabidopsis* that control phenylpropanoid pathway [56]. For

example, R2R3MYB TF encoding genes have been identified in several plant species such as 121 genes in tomato (*Solanum lycopersicum* L.), >200 genes in *Zea mays*, 126 genes in *Arabidopsis thaliana*, 117 genes in *Vitis vinifera* signifying their major roles in plant response to stress conditions [57, 58, 59, 60]. Previous studies had demonstrated that bHLH TFs (GL3, eGL3, TT8) certainly play an important role in regulation of anthocyanin biosynthesis in *Arabidopsis thaliana* through interaction with MYB TFs [61]. MYC2 protein of bHLH class binds to *cis*-elements in the ORCA3gene promoter, thus controlling the expression of several terpenoid indole alkaloids (TIA) biosynthesis genes in *Catharanthus roseus* [62]. Similarly, WRKY TFs such as GaWRKY1, AaWRKY1, WRKY3, WRKY6 and WRKY33 are also known to control the regulation of diverse biological processes in plants like cotton, *Artemisia annua*, tobacco and *Arabidopsis*, respectively, contributing to the biosynthesis of gossypol, artemisinin, and camalexin, respectively [63, 64, 65]. In plants, bZIP is a master regulator being actively involved in physiological and biological processes, including growth and development, seed formation, abiotic and biotic stress responses, etc. [66]. The bZIP TFs are well known and characterized in grapevine (*Vitis vinifera* L.) for their role in controlling the flavonoid biosynthesis through regulation of the phenylpropanoid pathway [67]. Similarly, bZIP TFs might be involved in controlling the podophyllotoxin biosynthesis via regulating the biosynthetic pathways genes as suggested via higher transcript abundance in rhizomes of *P. hexandrum* (Figure 4.9).

In addition, we have also been identified unique and common TFs between *P. hexandrum* and *P. peltatum,* respectively. Podophyllotoxin content differed significantly in two *Podophyllum* species wherein shoots of *P. peltatum* contain 2.05% podophyllotoxin content whereas shoots of *P. hexandrum* had almost negligible amount of podophyllotoxin content (0.01%). Roots/rhizomes of *P. hexandrum* contained highest podophyllotoxin content of 3.05% as compared to *P. peltatum* (0.25%). These contrasting conditions of podophyllotoxin biosynthesis provided an opportunity for identification, characterization and comparative analysis of TFs commonly or uniquely present in similar tissue/types and their association with the biosynthesis of podophyllotoxin and its derivatives. Two unique transcripts (medp_podpe_41091 and medp_podpe_2547) encoded bHLH and MYB/SANT TFs in shoots of *P. peltatum*. Similarly, medp_podhe_9316 and medp_podhe_111874 transcripts encoded MYB and bZIP TFs in rhizomes of *P. hexandrum.* On analyzing the common TFs between *P. hexandrum* and *P. peltatum*, 218 TFs were found common in the shoots and rhizomes of both the plant species. The identified common transcription factors were further analysed on the

basis of transcript abundance obtained in terms of FPKM value through RSEM. The transcript abundance ranged from 0.29 to 285.73 in shoots and 0.12 to 1162.01 in rhizomes of *P. hexandrum*. Similarly, for *P. peltatum*, transcript abundance ranged from 0.16 to 184.09 in shoots and 0.26 to 173.66 in rhizomes**.** *In silico* transcript abundance has been correlated with metabolite contents in plant species like *P. kurroa*, *C. sinensis*, *Solanum tuberosum* L and *Malus x domestica* [88, 104, 105, 106]. The common TFs identified from both the *Podophyllum* species suggested that these are conserved in nature and may be responsible for providing conserved defense responses against various environmental cues [107]. Moreover, conserved domains were also identified in transcripts that encoded TF families in transcriptomes of *P. hexandrum* (Figure 4.6). Conserved domains provide valuable information pertaining to cellular processes, including DNA binding, protein interactions, enzyme activity, etc., within plant systems.

Table 5.1 Transcription factor families associated with regulation of secondary metabolite biosynthesis in different plant species

Sr. no.	**TFs**	**Function**	**Plant species**
1.	MYB	Regulates the anthocyanin biosynthesis	*A. thaliana*
2.	WRKY	Regulates the artemisinin, camaleaxin and TIA biosynthesis	*C. roseus, A. annua, A. thaliana*
3.	bHLH	Anthocyanin, terpene, nicotine biosynthesis	*N. tabacum, C. roseus, A. annua*
4.	bZIP	Control and regulates proanthocyanidin, terpenoids biosynthesis	*Diospyros kaki*, *A. thaliana*
5.	NAC	Regulates the camaleaxin biosynthesis	*A. thaliana*
6.	DOF	Regulates the flavonoid biosynthesis	*A. thaliana*
7.	Zing Finger	Regulates terpenoid indole alkaloids biosynthesis (TIA)	*C. roseus*
8.	SPL	Regulates anthocyanin biosynthesis	*A. thaliana*
9.	AP2/ERF	Regulates artemisinin and TIA biosynthesis	*A. annua*
10.	mTERF	Participates in the growth and development, adaptive mechanism against harsh environmental cues	*A. thaliana, Z. mays*
11.	TFIIIA	Participates in the adaptive mechanism in response to salt and drought stress	*A. thaliana, Medicago truncatula, Oryza sativa* L.

The identified putative *cis*-regulatory elements in pathway genes are suggested to be involved in the regulation of podophyllotoxin biosynthesis pathway. The *in-silico* analysis revealed that binding sites for two transcription factors MYB and WRKY are commonly present in the promoter regions of pathway genes such as *CAD, C4H, PS, C3H, PAL, 4CL, PLR* and *SRD*. These transcription factors are suggested to regulate the podophyllotoxin biosynthesis by complementary binding with the sequence elements in the promoters of target pathway genes. However, the validation of binding pattern of *cis*-regulatory elements with transcription factors can be taken up in future.

5.3 Experimental validation and expression profiling of ABC transporters

ABC transporters are known as the largest families of transporters in plant systems having role in transport of various macro and micro molecules across biological membranes [108, 109]. ABC transporters have broad substrate specificity and a number of ABC proteins have been implicated in transport of secondary metabolites in different plant species. Most of the identified transporters belonging to ABC family included ABCB/MDR (multi drug resistance), ABCC/MRP (multidrug and resistance protein) and ABCG/PDR (pleiotropic drug resistance). The transporters had significant role in various physiological and biochemical processes in plant systems. The role of transporters is not fully understood due to limited availability of desired radiolabeled compounds requisite to decipher the transport mechanisms in plant systems. There is no report available on identification of ABC transporters genes and dissecting their role in secondary metabolites biosynthesis in *P. hexandrum*. Therefore, transcriptomes of *P. hexandrum* were mined to identify and validate ABC transporters genes.

Total 252 ABC transporter genes were identified as unigenes in transcriptomes of *P. hexandrum*, which were further classified into 9 sub-families such as ABCA, ABCB, ABCC, ABCD, ABCE, ABCF, ABCG, ABCI, and others (Figure 4.10). Phylogenetic analysis of ABC transporters showed high homology with the functionally characterised genes of *Nicotiana plumbaginifolia, Medicago truncatula, Arabidopsis thaliana*, *Zea mays,* etc. Clustering of ABC transporters genes with the well characterised genes of model plant species further confirmed the function governed by the identified ABC transporters in *P. hexandrum*. Out of 252, 22 genes were selected for validation on the basis of transcript abundance to gain insights into regulation of podophyllotoxin biosynthesis. Transcript

abundance for 22 ABC transporters genes was checked by two approaches- FPKM and qRT-PCR analysis. Expression analysis of 22 ABC transporters genes revealed 14 genes with higher expression in shoots as compared to roots and rhizomes of *P. hexandrum* suggesting their major role in the transport of various macro and micromolecules in shoots (Figure 4.11). Shoots are the primary organs for synthesising carbohydrates, starch and glucose in the photosynthesis required for various physio-biological and biochemical processes including secondary metabolites biosynthesis in plant systems [110]. The expression of ABC transporters in shoots might be due to the their involvement in transporting sugars and other carbohydrates, amino acids, peptides, polyamines, metal ions, etc., to rhizomes- the storage organ of *P. hexandrum*. The expression pattern observed by two approaches was found in concurrence with each other. Further in order to find out suitable candidate genes correlating with the podophyllotoxin content in *P. hexandrum,* Pearson's correlation coefficient was determined. Two genes, namely *PhABC6* and *PhABCIII* showed positive correlation with podophyllotoxin content. These genes were suggested to be the potential candidates for designing possible genetic intervention strategies aimed at enhancing podophyllotoxin content.

Furthermore, common and unique ABC transporters genes were also identified in shoot and rhizome transcriptomes of *P. hexandrum*. 239 ABC transporters genes were found common in shoots and rhizomes of *P. hexandrum.* Similarly, 16 genes (3 in rhizomes and 12 in shoots) were uniquely present in *P. hexandrum*. Transcript abundance analysis for unique genes revealed transcript medp_podhe_95366 with highest abundance with 11.96 FPKM value in *P. hexandrum*. Similarly, transcript medp_podhe_20101112|14082 showed highest abundance (402.88 FPKM value) among commonly present ABC transporters genes in shoots and rhizomes of *P. hexandrum.* Previously, the unique and common genes have been identified in plant species such as *Jatropha curcas*, *Ricinus communis*, *Arabidopsis thaliana*, etc. for their role in plant defense systems [107]. The genes commonly present between shoots and rhizomes of *P. hexandrum* may be involved in the transport of common metabolites. The uniquely present genes in rhizomes might be of much more interest as they were suggested to be involved in the transport of specific compounds in *P. hexandrum*.

Till date, only 22 ABC transporters genes were known which have been functionally characterised in different tissues/organs of *Arabidopsis thaliana* having diverse role such as cutin and pollen exine formation, plastid lipid formation, auxin transport, fatty acyl-CoA

import to peroxisome, cuticle formation, stomata regulation, abscisic acid import, metal/metalloid tolerance, folate transport, chlorophyll catabolite transport, kanamycin tolerance, biotic and abiotic stress tolerance [108]. Out of these 22 ABC transporters genes from Arabidopsis, 3 genes were found to have orthologous in the transcriptomes of *P. hexandrum*, and orthologous of two ABC transporter genes from *Vitis vinifera* and *Nicotiana plumbaginifolia* were also identified. The identified ABC transporters genes were further validated and their expression pattern was evaluated across different tissues/organs of *P. hexandrum* using qRT-PCR in order to correlate their expression with podophyllotoxin content (Figures 4.12 and 4.13). *PhABCII* and *PhABCV* showed higher transcript abundance in shoots and *PhABCIII* in rhizomes of *P. hexandrum* (Figure 4.13). The expression pattern of *PhABCV* was consistent with previous reports, where this gene was highly expressed in shoots of *Nicotiana plumbaginifolia* having role in terpenoids secretion [111]. Similarly, *PhABCII* showed higher expression in shoots, where it was involved in transportation of metabolites, like xenobiotics, folates, conjugates of chlorophyll catabolite, phytohelatins etc. in *Arabidopsis thaliana* [112, 113].

The podophyllotoxin biosynthesis occurs in roots/rhizomes of *P. hexandrum*; however the transport mechanisms of podophyllotoxin remain to be elucidated. Podophyllotoxin is mainly stored in the vacuole in rhizomes of *P. hexandrum*. ABC transporters were thought to be involved in transporting podophyllotoxin and its derivatives from different compartments to vacuoles of rhizomes. Several questions need to be answered such as how many transporters are involved in the transportation of podophyllotoxin. An attempt has been made by providing preliminary information of ABC transporters genes through identification and validation in different tissues/organs of *P. hexandrum.* ABC transporters were well known for transporting various secondary metabolites including lignins, flavonoids, terpenoids, etc. across plasma membranes and from cytosol to vacuole in the plant systems [114]. To gain insights into podophyllotoxin biosynthesis and accumulation, we also evaluated the expression status of ABC transporters genes in roots and rhizomes with respect to shoots of *P. hexandrum*. Five genes namely *PhABC11*, *PhABC16*, *PhABC17*, *PhABCIII* and *PhABCIV* showed higher expression ranging from 4.45-folds to 20.53-folds in rhizomes and roots as compared to shoots (Figures 4.12 and 4.13). *PhABCIV* showed highest expression (20.53-folds) in rhizomes suggesting their major role in accumulation and transport of podophyllotoxin content in rhizomes of *P. hexandrum*. *PhABCIV* is well characterised for the transport of anthocyanidin 3-*O*-glucosides into the vacuole in *Vitis vinifera* [115].

The identified ABC transporters genes are thought to be involved in podophyllotoxin accumulation as their transcript abundance correlated with podophyllotoxin content. However, the functional validation of these highly correlated genes remains to be deciphered. The present study has set-up a new platform by providing ABC transporters gene sequences to scientific community for further research. This study would help in elucidating the physiological and biological functions of ABC transporters in *P. hexandrum*.

5.4 Role of miRNAs in regulation of podophyllotoxin biosynthesis

Our previous study revealed that the presence of multiple pathway genes contributed towards podophyllotoxin biosynthesis [89]. In order to enhance the podophyllotoxin content in *in-vitro* the metabolic engineering requires at multiple steps of pathways or there is urgent need to find out the crucial regulatory components such as transcription factors and miRNAs involved in the regulation of podophyllotoxin. However, no information was available on molecular basis of regulation of podophyllotoxin biosynthesis in *P. hexandrum*.

The biosynthesis of primary and secondary metabolites is controlled through miRNA-mediated regulatory networks in plant systems. The NGS transcriptomes can provide valuable information about the role of miRNAs in various physio-biological processes in plants species. The present study identified 279 conserved miRNAs for the first time from transcriptomes known to regulate different biological processes in *P. hexandrum*. The putative mRNAs targets were identified for 279 conserved miRNAs using psRNATarget tool. This analysis revealed 896 mRNAs targets in *P. hexandrum* transcriptomes. To check the functions of identified 896 mRNA targets, KAAS analysis was performed which provided functional annotation for mRNAs targets. 26 mRNAs targets have been implicated in different secondary metabolites biosynthetic pathways in plant species.

Out of 26 mRNAs targets, 8 targets *UDP-Gly, FS, GAPDH, PRO, MD, PEP, WRKY* and *MYB* transcription factors have been suggested to regulate the primary and secondary metabolites biosynthesis, including podophyllotoxin biosynthesis (Kumar et al., 2015 (Unpublished) [116, 117] (Table 5.2). Six miRNAs (vun-miR396b, gma-miR396d, osa-miRf10132-akr, mtr-miR2673a, peu-miR2910 and vvi-miR828b) were corresponded to these eight mRNAs targets (Figure 4.14). miRNAs mtr-miR2673a/b had maximum number of targets (3 targets) i.e. peroxidise, MYB and WRKY transcription factors known to have their

crucial role in controlling the biological processes as confirmed by psRNATarget and KAAS analysis.

Six miRNAs were further experimentally validated on the basis of their predicted mRNAs targets implicated in regulation of podophyllotoxin biosynthesis. qRT-PCR validation and expression profiling revealed candidate miRNAs from *P. hexandrum* that can act as suitable targets for designing a genetic intervention strategy. Out of several mRNA targets predicted, an important target is UDP glycosyltransferase governed by vun-miR396b. The expression profiles of vun-miR396b (747-fold) and its target UDP-Gly (~89-folds) was in negative correlation with each other in rhizomes of *P. hexandrum*. UDP-Gly played crucial role in the biosynthesis of secondary metabolites in various plant species through glucosylation of intermediates in biosynthetic pathways [118]. UDP-Gly is known for its role in the formation of key intermediates of monoterpenoid and alkaloids biosynthesis pathways [119]. The involvement of UDP glycosyltransferase in podophyllotoxin biosynthesis had been shown previously, where it was suggested to be involved in the translocation of intermediates coumaryl and coniferyl alcohol in the form of glucosides from cytosol to cell wall [116, 120].

In this study, we have also found one miRNA (mtr-miR2673a/b) targeting transcription factors (TF) MYB and WRKY (Table 2). MYB and WRKY TFs are known for their regulatory role in various physio-biological processes such as plant growth and development, responses to environmental cues, primary and secondary metabolites biosynthesis, etc. [121]. WRKY TFs (GaWRKY1, AaWRKY1, WRKY3, WRKY6 and WRKY33) have been implicated in the regulation of gossypol, artemisinin and camalexin biosynthesis in cotton, *Artemisia annua*, tobacco and *Arabidopsis*, respectively [63, 64, 65]. Similarly, MYB TFs were involved in the regulation of physio-biological processes such as growth and development, metabolites biosynthesis, abiotic and biotic stress responses, seed formation and phenylpropanoid biosynthesis, etc. [56]. These TFs have also been suggested in the formation and regulation of podophyllotoxin biosynthesis [121]. The expression analysis showed negative correlation of MYB and WRKY TFs with their corresponding miRNAs in rhizomes of *P. hexandrum.*

The selected miRNAs (vun-miR396b, gma-miR396d, osa-miRf10132-akr, mtr-miR2673a, peu-miR2910 and vvi-miR828b) showed possible role in the regulation of podophyllotoxin biosynthesis as their mRNA targets were implicated in different pathway modules such as glycolysis, TCA cycle, shikimic acid and phenylpropanoid pathways contributing towards

podophyllotoxin biosynthesis (Figure 4.14). The mRNA targets belonged to the biosynthesis of primary metabolites such as *GAPDH, PRO, MD* and *PEP* (Table 5.2). The biosynthesis of primary metabolites is prerequisite for the formation of secondary metabolites as the products of primary metabolic pathways act as precursors for the secondary metabolites biosynthesis, including podophyllotoxin biosynthesis. The expression analysis of *GAPDH, PRO, MD* and *PEP* and their miRNAs revealed negative correlation thereby, suggesting the downregulation of these mRNA targets by complementary miRNA:mRNA base pairing in *P. hexandrum* (Figures 4.15 and 4.16). Due to downregulation of these mRNA targets the biosynthesis of podophyllotoxin got affected and ultimately led to lesser podophyllotoxin production. The identified miRNAs could be knocked-down to increase podophyllotoxin content in *P. hexandrum.* The selected six miRNAs are known in many plant species for their role in regulation of metabolites such as vun-miR396b was regulating the growth regulating factor (GRF) and (basic helix-loop-helix) bHLH family of transcription factors in *Arabidopsis thaliana* [122]. Similarly, vvi-miR828b was shown to regulate the MYB transcription factor involved in the biosynthesis of flavonoid and lignin in *Arabidopsis thaliana* [79].

Table 5.2 Potential targets of miRNAs from *P. hexandrum* transcriptomes

miRNA family	**Sequence**	**Target description**	**Function**
vun-miR396b	GCTCAAGAAAGCTGTGGGAG	Flavnol synthase	Phenylpropanoid biosynthesis
		UDP glycosyltransferase	Phenylpropanoid biosynthesis, Transferase activity
gma-miR396d	AAGAAAGCTGTGGGAGAATATGG	Pyruvate decarboxylase	Krebs cycle
osa-miRf10132-akr	GCGAGCTTCTCGAAGATGTCGTTGA	Glyceraldehydes 3-phosphate dehydrogenase	Glycolysis pathway
peu-miR2910	TAGTTGGTGGAGCGATTTGTC	Malate dehydrogenase	Krebs cycle
vvi-miR828b	TCTTGCTCAAATGAGTGTTCCA	Phosphoenolpyruvate carboxylase	Glycolysis pathway
mtr-miR2673a	CCTCTTCCTCTTCCTCTTCCA	WRKY	Phenylpropanoid biosynthesis
		MYB	Phenylpropanoid biosynthesis
		Peroxidise	Response to oxidative stress

CONCLUSION AND FUTURE PROSPECTS

The outcome of this work results in the identification and validation of crucial molecular components (pathway genes, TFs, ABC transporters and miRNAs) contributing to podophyllotoxin biosynthesis machinery in *P. hexandrum*. The information on molecular basis of these molecular components will provide further insights into the existing knowledge of podophyllotoxin biosynthesis. Expression analysis revealed nine genes (*SRD, PD, C3H, CM, COMT, CAD, CCR, C4H* and *ADT*) of podophyllotoxin biosynthetic pathway in root and rhizome tissues of *P. hexandrum* that were positively correlated with the podophyllotoxin content. Further, the work on regulatory components i.e. TFs and miRNAs revealed four key transcription factors (MYB, WRKY, bHLH and bZIP) and six miRNAs (miR396b, miR396d, miRf10132-akr, miR2673a, miR2910 and miR828b) targeted UDP glycosyltransferase, flavnol synthase, glyceraldehyde 3-phosphate dehydrogenase, peroxidise, malate dehydrogenase, phosphoenolpyruvate carboxylase, WRKY and MYB transcription factors showed possible role in regulation of podophyllotoxin biosynthesis in *P. hexandrum*. These regulatory components would facilitate the research on understanding the regulatory machinery of podophyllotoxin biosynthesis. Moreover, the work on ABC transporters would lay a basis for understanding the transportation of podophyllotoxin by revealing the candidate transporters genes *PhABC6* and *PhABCIII* that showed positive correlation with podophyllotoxin content. The research work deciphered molecular components involved in the biosynthetic machinery of podophyllotoxin that could be the key targets for designing a genetic intervention strategy to enhance podophyllotoxin content.

The research work laid the foundation for future research aimed at escalates the production of bioactive compounds in *Podophyllum* species and eventually beneficial for the industries and society. Exploration on molecular components associated with the podophyllotoxin biosynthetic machinery open new ways to design the novel metabolic engineering strategies. The present research work explain and extend information on whole pathway genes, transcription factors, ABC transporters and miRNAs involved in the metabolic pathways related to podophyllotoxin synthesis.

Overall, this work provides a foundational base for future studies aimed at deciphering the regulatory networks of podophyllotoxin biosynthesis in *Podophyllum* species. Functional validation of key genes, either through silencing or over-expression can help in metabolic engineering of *P. hexandrum* towards enhanced production of podophyllotoxin. Moreover, it further builds a platform for upcoming researchers and established scientific communities who wish to do work on related field to fulfil their future endeavours.

SUMMARY

Deciphering the biosynthetic machinery of podophyllotoxin has provided deep insights into understanding the biosynthesis and accumulation of podophyllotoxin in *P. hexandrum*. In the present study, the available transcriptome resources of *Podophyllum* species- *Podophyllum hexandrum* and *Podophyllum peltatum* from medicinal plants transcriptomics database (http://medplants.ncgr.org/) were utilized to mine molecular components implicated in podophyllotoxin biosynthetic machinery. The present study led to the identification of crucial molecular components associated with podophyllotoxin biosynthesis. The identified candidate pathway genes in podophyllotoxin biosynthesis would enhance our current knowledge on biosynthetic machinery. Expression pattern of nine genes - *SRD, C3H, CAD, CCR, C4H*, *PD, CM*, *COMT* and *ADT* of podophyllotoxin biosynthetic pathway in root and rhizome tissues of *P. hexandrum* was positively correlated with the podophyllotoxin content. Roots and rhizomes of *P. hexandrum* are active in the biosynthesis and accumulation of podophyllotoxin. To further gain insights into podophyllotoxin biosynthetic machinery, regulatory components such as transcription factors, *cis*-regulatory elements, ABC transporters and miRNAs were also identified, characterised, validated and dissected for their possibly role in podophyllotoxin biosynthesis. *In silico* analysis of putative *cis*-regulatory elements revealed that Skn-1 motif and MBS elements were common among the up-stream regions of higher expressed genes CMT, CAD, CCR, C4H and ADH, thereby, suggesting the common regulation of podophyllotoxin biosynthesis.

Furthermore, the study provided first glimpse on regulation of podophyllotoxin biosynthesis by identifying four key transcription factors- MYB, WRKY, bHLH and bZIP, two candidate ABC transporters- *PhABC6* and *PhABCIII* and four miRNAs- miR396d, miR2673a, miR2910 and miR396b involved in biosynthesis and accumulation of podophyllotoxin in *P. hexandrum*.

The present study is the first attempt to decipher podophyllotoxin biosynthetic machinery and discern the possible role of molecular components associated with podophyllotoxin biosynthesis and accumulation in *P. hexandrum.* The identified candidate molecular components (pathway genes, transcription factors, ABC transporters and miRNAs)

will help to unravel the molecular basis of biosynthesis and regulation of secondary metabolites in *P. hexandrum.* These molecular components could serve as important genetic intervention points for metabolic engineering to enhance the podophyllotoxin content in *P. hexandrum.*

Further, the study also provided a repertoire of transcripts corresponding to pathway genes, transcription factors and ABC transporters in *P. hexandrum* for future endeavours at following link: https://sites.google.com/site/combiogroup/data-access.

REFERENCES

[1] S.S. Samant, Dhar U., and Palni L.M S., "Medicinal Plants of Indian Himalaya", *Himavikas Publ*., GBPIHED, Kosi Katarmal, Almora, 1998.

[2] D.E. Jackson, and P.M. Dewick, "Aryltetralin lignans from *Podophyllum hexandrum* and *Podophyllum peltatum*", *Phytochemistry*, vol. 23, pp.1147-1152, 1984.

[3] A. Nag, P. Bhardwaj, P.S. Ahuja, and R.K. Sharma, "Identification and characterization of novel UniGene-derived microsatellite markers in *Podophyllum hexandrum* (Berberidaceae)", *J. Genet,* vol. 92, pp. e4–e7, 2013

[4] O. Pagani, M. Zucchetti, C. Sessa, J. Jong, M.D. Incalci, M.D. Fusco, A. Kaeser-Fröhlich, A. Hanauske, and F. Cavalli, "Clinical and pharmacokinetic study of oral NK 611, a new podophyllotoxin derivative", *Cancer Chemotherapy Pharmacol*, vol. 38, pp. 541–547, 1996.

[5] R. Gupta, and K.L. Sethi, "Conservation of medicinal plant resources in the Himalayan region. In: Conservation of tropical plant resources", Eds. SK Jain and KL Mehra. Botanical survey of India. *Howarh*, pp. 101-107, 1983.

[6] D. Bhattacharyya , R. Sinha, S. Ghanta , A. Chakraborty , S. Hazra, and S. Chattopadhyay, "Proteins differentially expressed in elicited cell suspension culture of *Podophyllum hexandrum wi*th enhanced podophyllotoxin content", *Proteome Sci,* vol. 10, pp.34, doi: 10.1186/1477-5956-10-34, 2012.

[7] A. Giri, and L.M. Narasu, "Production of podophyllotoxin from *Podophyllum hexandrum*: a potential natural product for clinically useful anticancer drugs", *Cytotechnology*, vol. 34, pp. 17–26, 2000.

[8] S. Chattopadhyay, R.S. Mehra, A.K. Srivastava, S.S. Bhojwani, and V.S. Bisaria, "Effect of major nutrients on podophyllotoxin production in *Podophyllum hexandrum* suspension cultures", *Applied Microbiology and Biotechnology,* vol. 60, pp. 541-546, 2003.

[9] Z.Q. Xia, M.A. Costa, J. Proctor, L.B. Davin, and N.G. Lewis, "Dirigent-mediated podophyllotoxin biosynthesis in *Linum flavum* and *Podophyllum peltatum*", *Phytochemistry* vol. 55, pp. 537–549, 2000.

[10] J.V. Marques, K.W. Kim, C. Lee, M.A. Costa, G.D. May, J.A. Crow, L.B. Davin, and N.G. Lewis, "Next generation sequencing in predicting gene function in podophyllotoxin biosynthesis", *Journal of Biological Chemistry*, vol. 288, pp. 466-479, 2013.

[11] D.P. Wankhede, D.K. Biswas, S. Rajkumar, and A.K. Sinha, "Expressed sequence tags and molecular cloning and characterization of gene encoding pinoresinol/lariciresinol reductase from *Podophyllum hexandrum*", *Protoplasma*, vol. 250, pp. 1239-1249, 2013.

[12] X. Lan, S. Ren, Y. Yang, M. Chen, C. Yang, H. Quan, G. Zhong, and Z. Liao, "Molecular cloning and characterization of the secoisolariciresinol dehydrogenase gene involved in podophyllotoxin biosynthetic pathway from Tibet *Dysosma*", *Journal of Medicinal Plants Research*, vol. 4, pp. 484-489, 2010.

[13] M. Yousefzadi, M. Sharifi, M. Behmanesh, E. Moyano, and J. Palazon, "Salicylic acid improves podophyllotoxin production in cell cultures of *Linum album* by increasing the expression of genes related with its biosynthesis", *Biotechnol Lett*, vol.32, pp. 1739–1743, 2010.

[14] S.Y. Kang, O. Choi, J. K. Lee, B. Y. Hwang, T.B. Uhm, and Y.S. Hong, "Artificial biosynthesis of phenylpropanoic acids in a tyrosine overproducing *Escherichia coli* strain", *Microbial Cell Factories*, vol. 11, pp. 153, 2012.

[15] S. Sachan,, S.K. Ghosh, Sen, and A. Mitra, "Co-production of caffeic acid and *p*-hydroxybenzoic acid from *p*-coumaric acid by *Streptomyces caeruleus* MTCC 6638", *Appl Microbiol Biotechnol*, vol.7, pp.720–727, 2006.

[16] K. Oksman-Caldentey, and D. Inzé, "Plant cell factories in the post-genomic era: new ways to produce designer secondary metabolites", *Trends in Plant Science*, vol. 9, no. 9, pp. 433-440, 2004.

[17] M. Li, L. Zhou, D. Yang, T. Li, and W. Li, "Biochemical composition and antioxidant capacity of extracts from *Podophyllum hexandrum* rhizome", *BMC Complementary and Alternative Medicine*, vol. 12, pp. 263, 2012.

[18] H.F. Stahelin, and A.V. Wartburg, "The chemical and biological route from podophyllotoxin glucoside to etoposide: ninth Cain memorial Award lecture", *Cancer Res* 51: 5–15, 1991.

[19] T. Imbert, "Discovery of podophyllotoxins", *Biochimie*, vol. 80, pp. 207-222, 1998.

[20] T. Huang, C. Shu, Y. Shih, H. Huang, Y. Su, Y. Chao, W. Yang, and J. Whang-Peng, "Protein tyrosine phosphatase activities are involved in apoptotic cancer cell death induced by GL331, a new homolog of etoposide", *Cancer Letters*, vol. 110, no. 1-2, pp. 77-85, 1996.

[21] O. Pagani, M. Zucchetti, C. Sessa, J. de Jong, M. D'Incalci, M. Fusco, A. Kaeser-Fröhlich, A. Hanauske, and F. Cavalli, "Clinical and pharmacokinetic study of oral NK611,

a new podophyllotoxin derivative", *Cancer Chemotherapy and Pharmacology*, vol. 38, no. 6, pp. 541-547, 1996.

[22] T. Utsugi, J. Shibata, Y. Sugimoto, K. Aoyagi, K. Wierzba, T. Kobunai, T. Terada, T. Oh-hara, T. Tsuruo, and Y. Yamada, "Antitumor activity of a novel podophyllotoxin derivative (Top-53) against lung cancer and lung metastic cancer", *Cancer Res*, vol. 56, pp. 2809–2814, 1996.

[23] T. Muranaka, M. Miyata, K. Ito, and S. Tachibana, "Production of podophyllotoxin in *Juniperus Chinensis* Callus cultures treated with oligosaccharides and a biogenetic precursor in honour of professor G. H. Neil Towers 75th birthday", *Phytochemistry*, vol. 49, no. 2, pp. 491-496, 1998.

[24] S. Kuhlmann, K. Kranz, B. Lücking, A. Alfermann, and M. Petersen, "Aspects of cytotoxic lignan biosynthesis in suspension cultures of *Linum nodiflorum*", *Phytochemistry Reviews*, vol. 1, no. 1, pp. 37-43, 2002.

[25] W. Van Uden, N. Pras, E. Vossebeld, J. Mol and T. Malingré, "Production of 5-methoxypodophyllotoxin in cell suspension cultures of *Linum flavum* L.", *Plant Cell Tiss Organ Cult*, vol. 20, no. 2, pp. 81-87, 1990.

[26] S. Chattopadhyay, R. Mehra, A. Srivastava, S. Bhojwani and V. Bisaria, "Effect of major nutrients on podophyllotoxin production in *Podophyllum hexandrum* suspension cultures", *Appl Microbiol Biotechnol*, vol. 60, no. 5, pp. 541-546, 2003.

[27] S. Chattopadyay, A. Srivastav, and V. Bisaria, "Optimization of culture parameters for production of podophyllotoxin in suspension culture of *Podophyllum hexandrum*", *Applied Biochemistry and Biotechnology*, vol. 102-103, no. 1-6, pp. 381-394, 2002.

[28] S. Chattopadhyay, A.K. Srivastava, S.S. Bhojwani, and V.S. Bisaria, "Development of suspension culture of *Podophyllum hexandrum* for the production of podophyllotoxin", *Biotechnology Letters*, vol. 23, pp. 2063-2066, 2001.

[29] A. Baldi, Jain N., Gupta A. Srivastava, and V. Bisaria, "Co-culture of arbuscular mycorrhiza-like fungi (*Piriformospora indica* and *Sebacina vermifera*) with plant cells of *Linum album* for enhanced production of podophyllotoxins: a first report", *Biotechnology Letters*, vol. 30, no. 9, pp. 1671-1677, 2008.

[30] A. Baldi S., Farkya A., Jain, N. Gupta, R. Mehra, V. Datta, A. Srivastava, and V. Bisaria, "Enhanced production of podophyllotoxins by co-culture of transformed *Linum album* cells with plant growth-promoting fungi", *Pure and Applied Chemistry*, vol. 82, no. 1, 2010.

[31] A. Chakraborty, D. Bhattacharya, S. Ghanta, and S. Chattopadhyay, "An efficient protocol for in vitro regeneration of *Podophyllum hexandrum*, a critically endangered medicinal plant", *Indian Journal of Biotechnology*, vol. 9, pp. 217-220, 2010.

[32] A. Goossens, S. Hakkinen, I. Laakso, T. Seppanen-Laakso, S. Biondi, V. De Sutter, F. Lammertyn, A. Nuutila, H. Soderlund, M. Zabeau, D. Inze, and K. Oksman-Caldentey, "A functional genomics approach toward the understanding of secondary metabolism in plant cells", *Proceedings of the National Academy of Sciences*, vol. 100, no. 14, pp. 8595-8600, 2003.

[33] H. Rischer, M. Oresic, T. Seppanen-Laakso, M. Katajamaa, F. Lammertyn, W. Ardiles-Diaz, M. Van Montagu, D. Inze, K. Oksman-Caldentey, and A. Goossens, "Gene-to-metabolite networks for terpenoid indole alkaloid biosynthesis in *Catharanthus roseus* cells", *Proceedings of the National Academy of Sciences*, vol. 103, no. 14, pp. 5614-5619, 2006.

[34] P.H. Qazi, A. Rashid, and S.A. Shawl, "*Podophyllum hexandrum* -a versatile medicinal plant", vol. 3, no. 5, pp. 261-268, 2011.

[35] M.K. Dhar, B. Friebe, S. Kaul, and B.S. Gill, "Characterization and physical mapping of 18S - 5.8S - 25S and 5S ribosomal RNA gene families in *Plantago* species", *Annals of Botany*, vol. 97, pp. 541–548, 2006.

[36] S.P.S. Khanuja, S. Paul, A.K. Shasany, A.K. Gupta, M.P. Darokar, M.M. Gupta, R.K. Verma, G. Ram, A. Kumar, R.K. Lal, and R.P. Bansal, "Genetically tagged improved variety 'CIM-Arogya' of *Artemisia annua* for high artemisinin yield", *Journal of Medicinal and Aromatic Plant Sciences*, vol. 27, pp. 520-524, 2005.

[37] Md A. Alam, P. Gulati, A.K. Gulati, G.P. Mishra, and P.K. Naik, "Assessment of genetic diversity among *Podophyllum hexandrum* genotypes of the North-western Himalayan region for podophyllotoxin production", *Indian Journal of Biotechnology*, vol. 8, pp. 391-399, 2009.

[38] M. Anrini, and S. Jha, "Characterization of podophyllotoxin yielding cell lines of *Podophyllum hexandrum*", *Caryologia*, vol. 62, no. 3, pp. 220-235, 2009.

[39] P. Sultan, A. Shawl, A. Jan, B. Hamid, and H. Irshad, "Germplasm conservation and quantitative estimation of podophyllotoxin and related glycosides of *Podophyllum hexandrum*", *Journal of Crop Science and Biotechnology*, vol. 12, no. 4, pp. 267-274, 2009.

[40] W. Liu, D. Yin, J. Liu, and N. Li, "Genetic diversity and structure of *Sinopodophyllum hexandrum* (Royle) Ying in the Qinling mountains, China", *PLoS ONE*, vol. 9, pp. e110500, 2014.

[41] P. Poczai, I. Varga, N.E. Bell, and J. Hyvonen, "Genomics meets biodiversity: Advances in molecular marker development and their applications in plant genetic diversity assessment", *The Molecular Basis of Plant Genetic Diversity*, Mahmut Caliskan (ed.), InTech, doi: 10.5772/33614, 2012,

[42] A. Nag, P. Ahuja, and R. Sharma, "Genetic diversity of high-elevation populations of an endangered medicinal plant", *AoB PLANTS*, vol. 7, no. 7, pp. plu076-plu076, 2015.

[43] P. Sultan, A.S. Shawl, P.W. Ramteke, A. Kour, and P.H. Qazi, "Assessment of diversity in *Podophyllum hexandrum* by genetic and phytochemical markers", *Scientia horticulturae*, vol. 115, no. 4, 398-408, 2008.

[44] O.P. Chaurasia, B. Ballabh, A. Tayade, R. Kumar, G.P. Kumar, and S.B. Singh, "Podophyllum L.: An endangered and anti-cancerous medicinal plant- An overview", *Indian Journal of Traditional Knowledge*, vol. 11, no. 2, pp. 234-241, 2012.

[45] M. Yousefzadi, M. Sharifi, M. Behmanesh, E. Moyano, M. Bonfill, R. Cusido, and J. Palazon, "Podophyllotoxin: Current approaches to its biotechnological production and future challenges", *Eng. Life Sci.*, vol. 10, no. 4, pp. 281-292, 2010.

[46] R. Dixon, and C. Steele, "Flavonoids and isoflavonoids – a gold mine for metabolic engineering", *Trends in Plant Science*, vol. 4, no. 10, pp. 394-400, 1999.

[47] G. Forkmann, and S. Martens, "Metabolic engineering and applications of flavonoids", *Current Opinion in Biotechnology*, vol. 12, no. 2, pp. 155-160, 2001.

[48] H. Kim, E. Ono, K. Morimoto, T. Yamagaki, A. Okazawa, A. Kobayashi and H. Satake, "Metabolic engineering of lignan biosynthesis in *forsythia* cell culture", *Plant and Cell Physiology*, vol. 50, no. 12, pp. 2200-2209, 2009.

[49] H. Satake, E. Ono, and J. Murata, "Recent advances in the metabolic engineering of lignan biosynthesis pathways for the production of transgenic plant-based foods and supplements", *J. Agric. Food Chem.*, vol. 61, no. 48, pp. 11721-11729, 2013.

[50] A. Nag, S. Chanda, and S. Rajkumar, "Estimation of nuclear genome size of important medicinal plant species from Western Himalaya using flow cytometry", *Journal of Cell and Plant Sciences*, vol. 2, pp. 19–23, 2013.

[51] M. Wink, "Biochemistry of plant secondary metabolism 2nd edition", *Annual Plant Review*, vol. 40, 2010.

[52] Yang, X. Fang, X. Wu, Y. Mao, L. Wang and X. Chen, "Transcriptional regulation of plant secondary metabolism", *Journal of Integrative Plant Biology*, vol. 54, no. 10, pp. 703-712, 2012.

[53] X. Dai, S. Sinharoy, M. Udvardi, and P. Zhao, “PlantTFcat: an online plant transcription factor and transcriptional regulator categorization and analysis tool”, *BMC Bioinformatics*, vol. 14, 321, 2013, doi: 10.1186/1471-2105-14-321.

[54] Xu, J. Huang, S. Guo, X. Yang, Y. Bao, H. Tang, and H. Zhang, “Overexpression of a TFIIIA-type zinc finger protein gene ZFP252 enhances drought and salt tolerance in rice (*Oryza sativa* L.)”, *FEBS Letters*, vol. 582, no. 7, pp. 1037-1043, 2008.

[55] Y. Liao, H. Zou, W. Wei, Y. Hao, A. Tian, J. Huang, Y. Liu, J. Zhang, and S. Chen, “Soybean GmbZIP44, GmbZIP62 and GmbZIP78 genes function as negative regulator of ABA signaling and confer salt and freezing tolerance in transgenic Arabidopsis", *Planta*, vol. 228, no. 2, pp. 225-240, 2008.

[56] Hichri, F. Barrieu, J. Bogs, C. Kappel, S. Delrot, and V. Lauvergeat, “Recent advances in the transcriptional regulation of the flavonoid biosynthetic pathway”, *Journal of Experimental Botany*, vol. 62, no. 8, pp. 2465-2483, 2011.

[57] P. Zhao, Q. Li, J. Li, L. Wang, and Z. Ren, “Genome-wide identification and characterization of R2R3MYB family in *Solanum lycopersicum*”, *Molecular Genetics and Genomics*, vol. 289, no. 6, pp. 1183-1207, 2014.

[58] A.P. Dias, E.L. Braun, M.D. McMullen, and E. Grotewold, “Recently duplicated maize R2R3 Myb genes provide evidence for distinct mechanisms of evolutionary divergence after duplication”, *Plant Physiology*, vol. 131, no. 2, pp. 610-620, 2003.

[59] R. Stracke, M. Werber, and B. Weisshaar, “The R2R3-MYB gene family in *Arabidopsis thaliana*”, *Current Opinion in Plant Biology*, vol. 4, no. 5, pp. 447-456, 2001.

[60] J.T. Matus, F. Aquea, and P. Arce-Johnson, “Analysis of the grape MYB R2R3 subfamily reveals expanded wine quality-related clades and conserved gene structure organization across Vitis and Arabidopsis genomes”, *BMC Plant Biol*, vol. 8, no. 1, pp. 83, 2008.

[61] C. Dubos, R. Stracke, E. Grotewold, B. Weisshaar, C. Martin, and L. Lepiniec, “MYB transcription factors in Arabidopsis”, *Trends Plant Sci*, vol. 15, pp. 573–581, 2010.

[62] H. Zhang, S. Hedhili, G. Montiel, Y. Zhang, G. Chatel, M. Pré, P. Gantet, and J. Memelink, “The basic helix-loop-helix transcription factor CrMYC2 controls the jasmonate-responsive expression of the ORCA genes that regulate alkaloid biosynthesis in *Catharanthus roseus*”, *The Plant Journal*, vol. 67, no. 1, pp. 61-71, 2011.

[63] Y.H. Xu., J.W. Wang, S. Wang, J.Y. Wang, and X.Y. Chen, “Characterization of GaWRKY1, a cotton transcription factor that regulates the sesquiterpene synthase gene (+)-delta-cadinene synthase A”, *Plant Physiol*, vol. 135, no. 1, pp. 507-515, 2004.

[64] M. Skibbe, N. Qu, I. Galis, and I.T. Baldwin IT "Induced plant defenses in the natural environment: *Nicotiana attenuata* WRKY3 and WRKY6 coordinate responses to herbivory. *Plant Cell*, vol. 20, pp. 1984–2000, 2008.

[65] Qiu, B. Fiil, K. Petersen, H. Nielsen, C. Botanga, S. Thorgrimsen, K. Palma, M. Suarez-Rodriguez, S. Sandbech-Clausen, J. Lichota, P. Brodersen, K. Grasser, O. Mattsson, J. Glazebrook, J. Mundy, and M. Petersen, "*Arabidopsi*s MAP kinase 4 regulates gene expression through transcription factor release in the nucleus", *EMBO. J*, vol. 2, pp. 2214–2221, 2008.

[66] M.S. Alves, S.P. Dadalto, A.B. Gonçalves, G.B.D. Souza, V.A. Barros, and L.G. Fietto, "Plant bZIP transcription factors responsive to pathogens: a review", *Int. J. Mol. Sci, vol. 14, pp.* 7815-7828, 2013.

[67] G. Malacarne, E.V. Coller, U. Heppel, S. Czemmel, S. Bogs, and J.C. Moser, "Characterization of a bZIP transcription factor that regulates the phenylpropanoid pathway in grapevine", In: XXVIth International Conference on Polyphenols, Florence, Italy, 22-26 July 2012: 2. handle: http://hdl.handle.net/10449/21284

[68] I. Vashisht, P. Mishra, T. Pal, S. Chanumolu, T. Singh, and R.S. Chauhan, "Mining NGS transcriptomes for miRNAs and dissecting their role in regulating growth, development, and secondary metabolites production in different organs of a medicinal herb, *Picrorhiza kurroa*", *Planta*, vol. 241, no. 5, pp. 1255-1268, 2015.

[69] C. Fu, R. Sunkar, C. Zhou, H. Shen, J. Zhang, J. Matts, J. Wolf, D. Mann, C. Stewart, Y. Tang, and Z. Wang, "Overexpression of miR156 in switchgrass (*Panicum virgatum* L.) results in various morphological alterations and leads to improved biomass production", *Plant Biotechnology Journal*, vol. 10, no. 4, pp. 443-452, 2012.

[70] A. Martin, H. Adam, M. Diaz-Mendoza, M. Zurczak, N. Gonzalez-Schain, and P. Suarez-Lopez, "Graft-transmissible induction of potato tuberization by the microRNA miR172", *Development*, vol. 136, no. 17, pp. 2873-2881, 2009.

[71] G. Davuluri, A. van Tuinen, P. Fraser, A. Manfredonia, R. Newman, D. Burgess, D. Brummell, S. King, J. Palys, J. Uhlig, P. Bramley, H. Pennings, and C. Bowler, "Fruit-specific RNAi-mediated suppression of DET1 enhances carotenoid and flavonoid content in tomatoes", *Nat Biotechnol*, vol. 23, no. 7, pp. 890-895, 2005.

[72] R. Rodriguez, M. Mecchia, J. Debernardi, C. Schommer, D. Weigel, and J. Palatnik, "Control of cell proliferation in *Arabidopsis thaliana* by microRNA miR396", *Development*, vol. 137, no. 1, pp. 103-112, 2009.

[73] A. Krek, D. Grun, M.N. Poy, R. Wolf, L. Rosenberg, E.J. Epstein, P. MacMenamin, I. da Piedade, K.C. Gunsalus, M. Stoffel, and N. Rajewsky, "Combinatorial microRNA target predictions", *Nat Genet*, vol. 37(5), pp. 495–500, 2005.

[74] A. Das, S. Chaudhury, M. Kalita, and T. Mondal, "In silico identification, characterization and expression analysis of miRNAs in *Cannabis sativa* L.", *Plant Gene*, vol. 2, pp. 17-24, 2015.

[75] Y. Moriya, M. Itoh, S. Okuda, A. Yoshizawa, and M. Kanehisa, "KAAS: an automatic genome annotation and pathway reconstruction server", Nucleic Acids Res, vol. 35, W182-W185, 2007.

[76] J.H. Tuteja, G. Zabala, K. Varala, M. Hudson, and L. O. Vodkin, "Endogenous, tissue-specific short interfering RNAs silence the chalcone synthase gene family in *Glycine max* seed coats", *The Plant Cell*, vol. 21, no. 10, pp. 3063-3077, 2009.

[77] J.Y. Gou, F. Felippes, C. Liu, D. Weigel, and J.W. Wang, "Negative regulation of anthocyanin biosynthesis in Arabidopsis by a miR156-targeted SPL transcription factor", *The Plant Cell*, vol. 23, no. 4, pp. 1512-1522, 2011.

[78] J.M. Debernardi, R.E. Rodriguez, M.A. Mecchia, and J.F. Palatnik, "Functional specialization of the plant miR396 regulatory network through distinct microRNA-target interactions", PLoS Genet, vol. 8, pp. e1002419, doi:10.1371/journal.pgen.1002419, 2012.

[79] R. Rajagopalan, H. Vaucheret, J. Trejo, and D.P. Bartel, "A diverse and evolutionarily fluid set of microRNAs in *Arabidopsis thaliana*", Genes Dev. Vol. 20, pp. 3407-25, 2006.

[80] K. Yazaki, "Transporters of secondary metabolites", *Current Opinion in Plant Biology*, vol. 8, pp. 301–307, 2005.

[81] E. Martinoia, M. Klein, M. Geisler, L. Bovet, C. Forestier, Ü. Kolukisaoglu, B. Müller-Röber, and B. Schulz, "Multifunctionality of plant ABC transporters – more than just detoxifiers", *Planta*, vol. 214, no. 3, pp. 345-355, 2001.

[82] E. Martinoia, E. Grill, R. Tommasini, K. Kreuz, and N. Amrhein, "ATP-dependent glutathione S-conjugate 'export' pump in the vacuolar membrane of plants", *Nature*, vol. 364, pp. 247-249, 1993.

[83] K. Yazaki, A. Sugiyama, M. Morita, and N. Shitan, "Secondary transport as an efficient membrane transport mechanism for plant secondary metabolites", *Phytochemistry Reviews*. vol. 7, pp. 513–524, 2008.

[84] G. Andolfo, M. Ruocco, A. Di Donato, L. Frusciante, M. Lorito, F. Scala, and M. Ercolano, "Genetic variability and evolutionary diversification of membrane ABC

transporters in plants", *BMC Plant Biol*, vol. 15, no. 1, pp. 51, doi: 10.1186/s12870-014-0323-2, 2015.

[85] R. Sanchez-Fernandez, T. Davies, J. Coleman, and P. Rea, "The *Arabidopsis thaliana* ABC protein superfamily, a complete inventory", *Journal of Biological Chemistry*, vol. 276, no. 32, pp. 30231-30244, 2001.

[86] J. Kang, J. Park, H. Choi, B. Burla, T. Kretzschmar, Y. Lee, and E. Martinoia, "Plant ABC tansporters" *The Arabidopsis Book/American Society of Plant Biologists*, vol. 9: e0153, doi:10.1199/tab.0153, 2011.

[87] M. Tabata, "The mechanism of shikonin biosynthesis in *Lithospermum* cell cultures", *Plant Tissue Cultures Letters*, vol.13, pp.117-125, 1996.

[88] S. Pandit, K. Shitiz, P.K. Naik, H. Sood, and R.S. Chauhan, "Expression pattern of fifteen genes of non-mevalonate (MEP) and mevalonate (MVA) pathways in different tissues of endangered medicinal herb *Picrorhiza kurroa* with respect to picrosides content", *Molecular Biology Reports*, vol. 40, pp. 1053-63, 2013.

[89] P. Kumar, T. Pal, N. Sharma, V. Kumar, H. Sood, and R.S. Chauhan, "Expression analysis of biosynthetic pathway genes vis-à-vis podophyllotoxin content in *Podophyllum hexandrum* Royle", *Protoplasma,* vol. 252 no 5, pp. 1252-1262, 2015.

[90] K. Tamura, G. Stecher, D. Peterson, A. Filipski, and S. Kumar, "MEGA6: Molecular evolutionary genetics analysis version 6.0", *Molecular Biology and Evolution*, vol. 30, pp. 2725-2729, 2013.

[91] X. Dai, and P.X. Zhao, "PsRNATarget: a plant small RNA target analysis server", *Nucleic Acids Reserach*, vol. 39, pp. 155 –159, 2011.

[92] B. Li, and C.N. Dewey, "RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome", *BMC Bioinformatics,* vol. 12, pp. 323, 2011.

[93] K. Higo, Y. Ugawa, M. Iwamoto, and T. Korenaga, "Plant *cis*-acting regulatory DNA elements (PLACE) database", *Nucleic Acids Research*, vol. 27, pp. 297-300, 1999.

[94] C.Q. Yang, X. Fang, X.M. Wu, Y.B. Mao, L.J. Wang, and X.Y. Chen, "Transcriptional regulation of plant secondary metabolism", *J Integr Plant Biol*, vol. 54, pp. 703-712, 2012.

[95] D.Q. Xu, J. Huang , S.Q. Guo, X.Yang, Y.M. Bao, H.J. Tang, and H.S. Zhang, "Over expression of a TFIIIA-type zinc finger protein gene ZFP252 enhances drought and salt tolerance in rice (*Oryza sativa* L.)", *FEBS Letters*, vol. 582, pp. 1037–1043, 2008.

[96] Z. Zhang, J. Yu, D. Li, Z. Zhang, F. Liu, X. Zhou, T. Wang, Y. Ling and Z. Su, "PMRD: plant microRNA database", *Nucleic Acids Research*, vol. 3, pp. D806-D813, 2010.

[97] N. Malhotra, V. Kumar, H. Sood, T.R Singh, and R.S. Chauhan, "Multiple genes of mevalonate and non-mevalonate pathways contribute to high aconites content in an endangered medicinal herb, *Aconitum heterophyllum* Wall, *Phytochemistry*, vol. 108, pp. 26-34, 2014.

[98] N. Gupta, S.K. Sharma, J.C. Rana, and R.S. Chauhan "Expression of flavonoid biosynthesis genes vis-à-vis rutin content variation in different growth stages of *Fagopyrum* species", *Journal of Plant Physiology*, vol. 168, pp. 2117-2123, 2011.

[99] K.H. Wang, H.Z. Peng, E.P. Lin, Q.Y. Jin, X.Q. Hua, S. Yao, H.W. Bian, N. Han, and J.W. Pan, "Identification of genes related to the development of bamboo rhizome bud", *J Exp Bot,* vol. *6*1, pp. 551–561, 2010.

[100] H.J. Koo, E. McDowell, X. Ma, K. Greer, J. Kapteyn, Z. Xie, A. Descour, H. Kim, Y. Yu, D. Kudrna, R. Wing, C. Soderlund and D. Gang, "Ginger and turmeric expressed sequence tags identify signature genes for rhizome identity and development and the biosynthesis of curcuminoids, gingerols and terpenoids", *BMC Plant Biol*, vol. 13, no. 1, pp. 27, 2013

[101] N. Malhotra, V. Kumar, H. Sood, T.R Singh, and R.S. Chauhan, "Multiple genes of mevalonate and non-mevalonate pathways contribute to high aconites content in an endangered medicinal herb, *Aconitum heterophyllum* Wall, *Phytochemistry*, vol. 108, pp. 26-34, 2014.

[102] K. Shitiz, N. Sharma, .T Pal, H. Sood, and R.S. Chauhan, "NGS transcriptomes and enzyme inhibitors unravel complexity of picrosides biosynthesis in *Picrorhiza kurroa* Royle ex. Benth", *Plos One*, 10(12): e0144546, 2015.

[103] A. Sood, and R.S. Chauhan, "Regulation of FA and TAG biosynthesis pathway genes in endosperms and embryos of high and low oil content genotypes of *Jatropha curcas* L", *Plant Physiology and Biochemistry,* vol. 94, pp. 253-267, 2015.

[104] A. Paul, S. Jha, S. Bhardwaj, S. Singh, R. Shankar, and S. Kumar, "RNA-seq-mediated transcriptome analysis of actively growing and winter dormant shoots identifies non-deciduous habit of evergreen tree tea during winters", *Scientific Reports*, vol. 4, pp. 5932, doi: 10.1038/srep05932, 2014.

[105] A.K. Singh, V. Sharma, A.K. Pal, V. Acharya, and P.S. Ahuja, "Genome-wide organization and expression profiling of NAC transcription factor family in potato (*Solanum tuberosum* L)", *DNA research*, vol. 20, pp. 403-423, 2013.

[106] P. Arya, G. Kumar, V. Acharya, and A. Singh, "Genome-wide identification and expression analysis of NBS-encoding genes in *Malus x domestica* and expansion of NBS genes family in Rosaceae", *PLoS ONE*, vol. 9, pp. e107987, 2014.

[107] V. Haake, "Transcription factor CBF4 is a regulator of drought adaptation in Arabidopsis", *Plant Physiology*, vol. 130, no. 2, pp. 639-648, 2002.

[108] F.L. Theodolou, "Plant ABC transporters", *Biochimica et Biophysica Acta*, Vol. 1465, pp. 79-103, 2000.

[109] I.B. Holland, S.P.C. Cole, K. Kuchler, and C.F. Higgins, "ABC proteins: From bacteria to man", *San Diego, CA.- Academic Press*, 2003.

[110] S.C. Huber, "Biochemical mechanism for regulation of sucrose accumulation in leaves during photosynthesis", *Plant Physiology*, vol. 91, pp. 656-662, 1989.

[111] S. Grec, D. Vanham, J. De Ribaucourt, B. Purnelle and M. Boutry, "Identification of regulatory sequence elements within the transcription promoter region of NpABC1, a gene encoding a plant ABC transporter induced by diterpenes", *The Plant Journal*, vol. 35, no. 2, pp. 237-250, 2003.

[112] W. Song, J. Park, D. Mendoza-Cozatl, M. Suter-Grotemeyer, D. Shim, S. Hortensteiner, M. Geisler, B. Weder, P. Rea, D. Rentsch, J. Schroeder, Y. Lee and E. Martinoia, "Arsenic tolerance in Arabidopsis is mediated by two ABCC-type phytochelatin transporters", *Proceedings of the National Academy of Sciences*, vol. 107, no. 49, pp. 21187-21192, 2010.

[113] A. Raichaudhuri, M. Peng, V. Naponelli, S. Chen, R. Sanchez-Fernandez, H. Gu, J. Gregory, A. Hanson, and P. Rea, "Plant vacuolar ATP-binding cassette transporters that translocate folates and antifolates in vitro and contribute to antifolate tolerance in vivo", *Journal of Biological Chemistry*, vol. 284, no. 13, pp. 8449-8460, 2009.

[114] R. Sibout, and H. Höfte, "Plant cell biology: The ABC of monolignol transport", *Current Biology*, vol. 22, pp. 533-535, 2012.

[115] R. Francisco, A. Regalado, A. Ageorges, B. Burla, B. Bassin, C. Eisenach, O. Zarrouk, S. Vialet, T. Marlin, M. Chaves, E. Martinoia and R. Nagy, "ABCC1, an ATP Binding Cassette Protein from Grape Berry, Transports Anthocyanidin 3- O –Glucosides", *The Plant Cell*, vol. 25, no. 5, pp. 1840-1854, 2013.

[116] A. Kumari, H.R. Singh, A. Jha, M.K. Swarnkar, R. Shankar, and S. Kumar, "Transcriptome sequencing of rhizome tissue of *Sinopodophyllum hexandrum* at two temperatures", *BMC Genomics*, vol. 7 pp. 871, 2014.

[117] D. Bhattacharyya, R. Sinha, R. Datta, S. Hazra, and S. Chattopadhyay, "De novo transcriptome analysis using 454 pyrosequencing of the Himalayan May apple, *Podophyllum hexandrum*", BMC Genomics, vol.1, pp. 748, 2013.

[118] M. Nagatoshi, K. Terasaka, A. Nagatsu, and H. Mizukami, "Iridoid-specific glucosyltransferase from *Gardenia jasminoides*", *Journal of Biological Chemistry*, vol. 286, pp. 32866-32874, 2011.

[119] A. Pani, and R.K. Mahapatra, "Computational identification of microRNAs and their targets in *Catharanthus roseus* expressed sequence tags", *Genomics Data,* vol. 1, pp. 2-6, 2013.

[120] J. Wang, and B. Hou, "Glycosyltransferases: Key players involved in the modification of plant secondary metabolites", *Front Biol China*, vol. 4, pp. 39–46, 2009.

[121] P. Kumar, V. Jaiswal, T. Pal, J. Singh, and R.S. Chauhan, "Comparative whole-transcriptome analysis in *Podophyllum* species identifies key transcription factors contributing to biosynthesis of podophyllotoxin in *P. hexandrum*", *Protoplasma,* vol. 254, pp. 217–228, 2017.

[122] J.M. Debernardi, R.E. Rodriguez, M.A. Mecchia, and J.F. Palatnik, "Functional specialization of the plant miR396 regulatory network through distinct microRNA-target interactions", *PLoS Genet*, vol. 8, pp. e1002419, doi:10.1371/journal.pgen.1002419, 2012.

[123] A.J. Broomhead, and P.M. Dewick, "Tumor inhibitory aryltralin lignans in *Podophyllum versipelle*, *Diphyllcia cymosa* and *Diphyllicia grayi*", *Phytochemistry*, vol. 2, pp. 3831-3837, 1990.

[124] T. Smollny, H. Wichers, S. Kalenberg, A. Shasavari, M. Petersen, and A.W. Alfermann, "Accumulation of podophyllotoxin and related lignans in cell suspension cultures of *Linum albu*m", *Phytochemistry*, vol. 48, pp. 975–979, 1998.

[125] A.J. Broomhead, and P.M. Dewick, "Aryltetralin lignans from *Linum flavum* and *Linum capitatum", Phytochemistry*, vol. 29, pp. 3839–3844, 1990.

[126] J.L. Hertwell, J.M. Johnson, D.B. Fitzgerald, and M. Belkin, "Podophyllotoxin form Jiniperus species: Savinin", *J. Am. Chem. Soc*, vol. 75, pp. 235-236, 1953.

[127] G.C. Hokanson, "Podophylltoxion and 4-demethylpodophyllotoxin from *Polygala polygama* (Polygalaceae)", *Lioyodia*, vol. 41, pp. 497-498, 1979.

[128] A. San Feliciano, J.M. Miguel del Corral, M. Gordaliza, and M.A. Castro, "Lignans from *Juniperus thruifera*", *Phytochemistry*, vol. 28, pp. 2863-2866, 1989.

[129] T. Vogt, "Phenylpropanoid biosynthesis", *Molecular Plant*, vol. 3, pp. 2–20, 2010.

List of Publications

Papers in Peer Reviewed Journals

1. **Pawan Kumar**, Jibesh Kumar Padhan, Ashwani Kumar, Rajinder S. Chauhan (2017) Transcriptomes of *Podophyllum hexandrum* unravel candidate miRNAs and their association with the biosynthesis of secondary metabolites. Journal of Plant Biochemistry and Biotechnology. doi.org/10.1007/s13562-017-0414-x **(IF- 1.35)**

2. **Pawan Kumar**, Varun Jaiswal, Tarun Pal, Jagdish Singh, Rajinder S. Chauhan (2017) Comparative whole-transcriptome analysis in *Podophyllum* species identifies key transcription factors contributing to biosynthesis of podophyllotoxin in *P. hexandrum*. Protoplasma 254 (1): 217-228 **(IF-2.87)**

3. **Pawan Kumar,** Tarun Pal, Neha Sharma, Varun Kumar, Hemant Sood, Rajinder S. Chauhan (2015) Expression analysis of biosynthetic pathway genes vis-à-vis podophyllotoxin content in *Podophyllum hexandrum* Royle. Protoplasma 252(5): 1253-1262 **(IF- 3.2)**

4. **Pawan Kumar**, Ritika Sharma, Varun Jaiswal, Rajinder S. Chauhan (2016) Identification, validation and expression of ABC transporters in *Podophyllum hexandrum* Royle and their role in podophyllotoxin biosynthesis. Biologia Plantarum 60:3, 452-458 (In Print) **(IF-1.9)**

5. Tarun Pal*, Jibesh Kumar Padhan*, **Pawan Kumar**, Hemant Sood, Rajinder S. Chauhan (2017) Comparative transcriptomics uncovers differences in physiological/biochemical processes between photoautotrophic and photoheterotrophic modes of nutrition vis-à-vis secondary metabolites biosynthesis in *Swertia chirayita*. Molecular Biology Reports **(IF-1.82)** (Accepted)

6. Jibesh Kumar Padhan, **Pawan Kumar**, Hemant Sood, Rajinder S. Chauhan (2016) Prospecting NGS-Transcriptomes to assess regulation of miRNA-inspected secondary metabolites biosynthesis in *Swertia chirayita*, a medicinal herb of North-Western Himalayas. Medicinal Plants - International Journal of Phytomedicines and Related Industries 8(3): 219-228 **(Scopus Indexed)**

Conference Papers/Abstracts

1. **Pawan Kumar**, Tarun Pal, Varun Jaiswal, Hemant Sood, Rajinder S. Chauhan (2014). Computational mining of transcriptomes for transcription factors controlling various biological processes in *Podophyllum* spp. Agrotechnology; 2:4, http://dx.doi.org/10.4172/2168-9881.S1.008

2. **Pawan Kumar**, Neha Sharma, Tarun Pal, Hemant Sood, Rajinder S. Chauhan (2014). Differential expression analysis of phenylpropanoid pathway genes controlling podophyllotoxin biosynthesis in *Podophyllum hexandrum*: A medicinal herb of North-Western Himalayas. Proceedings of the National Symposium on Advances in Biotechnology for Crop Improvement , Baru Sahib, Himachal Pradesh, India July 12, 2014

3. **Pawan Kumar**, Neha Sharma, Tarun Pal, Hemant Sood, Rajinder S. Chauhan (2014). Molecular components control podophyllotoxin biosynthesis in *Podophyllum hexandrum*: A medicinal herb of North-Western Himalayas. Proceedings of International Conference on Emerging Trends in Biotechnology, page no-14, Jawaharlal Nehru University, New Delhi, India, 6th Nov-9th Nov. 2014

4. Neha Sharma, **Pawan Kumar**, Tarun Pal, Hemant Sood, Rajinder S. Chauhan (2014). In-*vitro* conservation and SSR fingerprinting in *Podophyllum hexandrum* Royle accession from North-Western Himalayas National Symposium on Advances in Biotechnology for Crop Improvement, Eternal University, Baru Sahib, Himachal Pradesh, India July 12, 2014

Best Poster Presentation Award

- **Pawan Kumar**, Tarun Pal, Varun Jaiswal, Hemant Sood and Rajinder S. Chauhan (2014) Computational mining of transcriptomes for transcription factors controlling various biological processes in *Podophyllum* spp. Poster presented in 2nd International Conference on Agricultural and Horticultural Sciences (Omics Group) Radisson Blu Plaza Hotel, Hyderabad, India, February 03-05, 2014

NCBI GenBank Submissions

Kumar P, Sharma N, Sood A, Sood H, and Chauhan RS (2015):

- Partial cloned sequence of MYB Family transcription factor (*PhMYB*) mRNA, partial cds KP299959
- Partial cloned sequence of bZIP Family transcription factor (*PhbZIP*) mRNA, partial cds KP299960
- Partial cloned sequence of WRKY Family transcription factor (*PhWRKY*) mRNA, partial cds KP299961