

Navneeta Bharadvaja

mihir new 2 (1)

Document Details

Submission ID

trn:oid:::27535:140626350

Submission Date

May 27, 2026, 7:09 AM GMT+0

Download Date

May 27, 2026, 7:12 AM GMT+0

File Name

mihir new 2 (1).pdf

File Size

990.4 KB

36 Pages**7,843 Words****48,051 Characters**

8% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Filtered from the Report

- Bibliography
- Cited Text

Match Groups

-  **45 Not Cited or Quoted 8%**
Matches with neither in-text citation nor quotation marks
-  **0 Missing Quotations 0%**
Matches that are still very similar to source material
-  **0 Missing Citation 0%**
Matches that have quotation marks, but no in-text citation
-  **0 Cited and Quoted 0%**
Matches with in-text citation present, but no quotation marks

Top Sources

- 5%  Internet sources
- 5%  Publications
- 6%  Submitted works (Student Papers)

Integrity Flags

0 Integrity Flags for Review

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.

Match Groups

- 45 Not Cited or Quoted 8%**
Matches with neither in-text citation nor quotation marks
- 0 Missing Quotations 0%**
Matches that are still very similar to source material
- 0 Missing Citation 0%**
Matches that have quotation marks, but no in-text citation
- 0 Cited and Quoted 0%**
Matches with in-text citation present, but no quotation marks

Top Sources

- 5% Internet sources
- 5% Publications
- 6% Submitted works (Student Papers)

Top Sources

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

- 1** Student papers
Graphic Era University on 2025-12-11 <1%
- 2** Publication
Hugo José Gonçalves dos Santos Junior, Alexandra Aparecida Zorzal, Claudio Aug... <1%
- 3** Student papers
National Institute Of Technology, Tiruchirappalli on 2025-12-31 <1%
- 4** Internet
www.dovepress.com <1%
- 5** Student papers
Teesside University (Blackboard Ultra) on 2026-01-07 <1%
- 6** Student papers
Delhi Technological University on 2026-05-25 <1%
- 7** Internet
public-pages-files-2025.frontiersin.org <1%
- 8** Student papers
University of Nottingham on 2020-05-12 <1%
- 9** Publication
Timsina, Sujan. "Computational Studies of Biomolecules to Explore Potential Ther... <1%
- 10** Internet
academic.oup.com <1%

11	Publication	Yuhan Xue, Junhao Li, Zengrui Wu, Guixia Liu, Yun Tang, Weihua Li. " Computatio...	<1%
12	Student papers	Universiti Teknologi Malaysia on 2026-02-18	<1%
13	Internet	www.frontiersin.org	<1%
14	Internet	www.ijltemas.in	<1%
15	Internet	www.researchsquare.com	<1%
16	Student papers	Gujarat Technological University on 2026-03-19	<1%
17	Student papers	RMIT University on 2020-03-13	<1%
18	Internet	aranne5.bgu.ac.il	<1%
19	Internet	www.scribd.com	<1%
20	Publication	Liu, B.. "Isolation and Characterization of 2-Methyl-6-Phytyl-1,4-Benzoquinol Me...	<1%
21	Student papers	Morehouse School of Medicine on 2025-04-28	<1%
22	Student papers	University of Edinburgh on 2021-04-27	<1%
23	Student papers	University of Nottingham on 2018-05-07	<1%
24	Student papers	Université Saint-Esprit Kaslik on 2011-05-23	<1%

25	Internet	link.springer.com	<1%
26	Internet	pure.manchester.ac.uk	<1%
27	Internet	www.researchgate.net	<1%
28	Publication	Almeida, J., L. Quadrana, R. Asis, N. Setta, F. de Godoy, L. Bermudez, S. N. Otaiza, J....	<1%
29	Publication	Nirmala Akula. "A Network-Based Approach to Prioritize Results from Genome-Wi...	<1%
30	Publication	SERGEY V. ANISIMOV. "A large-scale screening of the normalized mammalian mit...	<1%
31	Publication	Subramani, Aarthi. "Modulation of Macrophage Antifungal Activity by the Transcr...	<1%
32	Publication	Twum, Eric. "From Gene Expression to Regulatory Networks: A Multi-Omics Fram...	<1%
33	Student papers	Universitat Oberta de Catalunya on 2026-03-08	<1%
34	Student papers	University of Surrey on 2017-05-08	<1%
35	Internet	f1000researchdata.s3.amazonaws.com	<1%
36	Internet	www.iaees.org	<1%
37	Internet	www.iajps.com	<1%
38	Publication	Nazim Hussain, Faiza Irshad, Zahra Jabeen, Imran Haider Shamsi, Zhilan Li, Lixi Ji...	<1%

39

Publication

Advances in Photosynthesis and Respiration, 2014.

<1%

40

Student papers

Nottingham Trent University on 2026-05-13

<1%

Integrating Computational and Biotechnology Approaches for Enhanced Tocopherol and Tocotrienol Biosynthesis in Plants

A THESIS

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

MASTER OF TECHNOLOGY

in

INDUSTRIAL BIOTECHNOLOGY

Submitted by

MIHIR KUMAR

(24/IBT/15)

Under the Supervision of

Prof. Navneeta Bharadvaja



DEPARTMENT OF BIOTECHNOLOGY

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Bawana Road, Delhi-110042

May, 2026

TABLE OF CONTENT

6	Subheading	Page No.
	Acknowledgement	ii
	Candidate's Declaration	iii
	Certificate	iv
	Abstract	v
	Table of Contents	vi – vii
	List of Tables, Figures and Abbreviations	vii - ix
	INTRODUCTION	1 -4
	LITERATURE REVIEW	5-8
	2.1 Tocopherols, Tocotrienols and Their Biological Significance	5
	2.2 Biosynthetic Pathway Architecture and Rate-Limiting Enzymes	5
	2.3 Transcriptomic Approaches to Metabolic Gene Network Analysis	6
	2.4 VTE2 as a Key Enzymatic Target for Metabolic Engineering	6
	2.5 Phytochemical-Based Modulation and Computational Drug Discovery	7
	2.6 Integration of Bioinformatics and Metabolic Engineering Strategies	7
	2.7 Study Rationale and Significance	8
	MATERIALS & METHODS	21-25
	3.1 Research Methodology Overview	9
	3.2 Bioinformatics Analysis for Vitamin E Biosynthesis Gene Networks	9
	3.2.1 Dataset Selection	9 – 10
32	3.2.2 Differential Gene Expression Analysis	10
	3.2.3 Data Normalization	10
34	3.2.4 Protein-Protein Interaction Network Construction	10
	3.2.5 Functional Enrichment Analysis	10 – 11
	3.2.6 Hub Gene Identification	11
	3.3 Molecular Docking for VTE2 Modulator Identification	11
	3.3.1 Preparing the Target Protein	11
12	3.3.2 Ligand Selection and Preparation	11
	3.3.3 Molecular Docking Procedure	12
	3.3.4 Selection of Top Candidate Compounds	12
	3.3.5 Protein-Ligand Interaction Analysis	12
	3.3.6 Structural Visualization	12 – 13
	3.4 Integrated Computational Framework	13
	RESULTS	14-
33	4.1 Differential Gene Expression Analysis	14
	4.2 Data Normalization	15
	4.3 Functional Enrichment Analysis	15 - 16
	4.4 Hub Gene Identification and PPI Network	17
	4.5 Molecular Docking Results	18
	4.6 Binding Affinity Analysis	19
	4.7 Protein-Ligand Interaction Analysis	20

4.8 Structural Visualization	21
DISCUSSION & CONCLUSION	22 – 25
REFERENCES	26 – 33
List of Publications with Certificates	34
Plagiarism Verification	35
Plagiarism Report	36
AI Report	37
Curriculum Vitae	38
Undertaking	39

ABSTRACT

Together, tocopherols and tocotrienols comprise vitamin E, which refers to a group of lipophilic antioxidants that are extremely important for human health and crop improvement. Despite their agricultural and pharmaceutical relevance, the molecular regulation behind the production of these molecules is largely unknown, and genetic modification techniques have not yet made any significant breakthroughs in improving their levels in crops. The intersection of the shikimate and methylerythritol phosphate (MEP) pathways in homogentisate phytyldiphosphate transferase (VTE2) defines the committed step towards tocopherol production, hence making it an ideal candidate for further studies. Herein, we collected gene expression data (GSE59790 and GSE47518) associated with the biosynthesis of tocopherol under high-light stress conditions in *Arabidopsis thaliana* from the Gene Expression Omnibus (GEO) database. Construction of PPI network was done using STRING database (confidence > 0.900) while hub genes identification was done based on degree centrality while GO and KEGG pathway analysis was used to identify functional enrichments. EasyDock Vina software was used for screening of selected ligands against crystal structure of VTE2 (PDB ID: 1YFP) while protein-ligand interactions were identified using LigPlot+ and PyMOL. The number of differentially expressed genes identified were 1,148 which included 634 upregulated genes and 514 downregulated genes. Analysis of hub genes in the upregulated PPI network showed that VTE1, VTE4, VTE2/HPT1, HPPD and DXPS were the most connected hubs. Enrichment analysis showed that plastidic metabolism, prenylquinone metabolism and oxidative stress response are the main biological processes occurring in the network. In the molecular docking, alpha-mangostin emerged as the top hit with docking energy of -9.8 kcal/mol, forming strong hydrogen bonds at Ser186, Asp227 and Arg295 while beta-tocotrienol and quercetin-3-glucoside were the next best hits with docking energy of -9.4 kcal/mol and -9.1 kcal/mol, respectively. This study offers an integrated regulatory network for tocopherol biosynthesis along with alpha-mangostin as a potential phytochemical regulator of VTE2.

Keywords: Tocopherol, Tocotrienols, Protein-protein network, Functional enrichment, GEO.

INTRODUCTION

Type Micronutrient deficiency is one of the most prevalent health problems facing humankind today, with the estimate reaching two billion individuals worldwide [1,2]. From the myriad of vital micronutrients, vitamin E, a term used to describe a group of eight closely related compounds known as the fat-soluble antioxidants consisting of alpha, beta, gamma, and delta tocopherols and alpha, beta, gamma, and delta tocotrienols, is especially important both nutritionally and physiologically [3,4]. The differences in the two sub-families are based on the saturation levels in the isoprenoid side chains; whereas the tocopherols have a fully saturated phytyl group, the tocotrienols have a farnesylgeranyl residue with three double bonds, which accounts for their very different physical properties and physiological actions [5]. However, despite being produced only by photosynthetic organisms, particularly plants, algae, and cyanobacteria, vitamin E is critical in human biology because it is the principal fat-soluble radical scavenging antioxidant [3,6].

Tocopherols and tocotrienols have roles in plant physiology that go beyond serving merely as antioxidants. They play pivotal roles in abiotic stress resistance and protect the chloroplast photosynthetic machinery from oxidative damage by excessive light exposure, water shortage, temperature extremes, and heavy metals [7,8]. Alpha tocopherol, which occurs most commonly in leaf tissues, acts as a scavenger of triplet chlorophyll and singlet oxygen molecules formed during the electron transfer process in photosystem II, thus preventing chain reactions of lipid peroxidation responsible for damage to thylakoid membranes [9]. Tocotrienols, which occur most often in the endosperm portion of monocotyledonous seeds like rice, wheat, and barley, have aroused much interest in pharmaceutical research because of their neuroprotective, anti-inflammatory, and anti-angiogenic properties, which are superior to the corresponding functions of tocopherols at comparable concentrations [10,11]. Agriculturally speaking, the vitamin E content in crops is a crucial factor determining the nutritional value of the resulting agricultural products. Therefore, the aim of breeding more tocopherol and tocotrienol rich plants has gained considerable attention worldwide [4,12].

The pathway leading to tocopherol biosynthesis in chloroplasts in plants is formed due to the merging of two metabolic pathways: shikimate pathway producing homogentisic acid (HGA) as the aromatic precursor of the head group, and MEP pathway, producing geranylgeranyl diphosphate (GGDP) then phytyl diphosphate (PDP) [13,14]. The reaction of condensation between HGA and PDP – catalysed by homogentisate phytyldiphosphate transferase, coded by VTE2 gene – represents the commitment point and rate-limiting step of tocopherol synthesis pathway [15]. The further methylation and cyclization reactions by the participation of VTE3, VTE1, and VTE4 enzymes result in complete set of tocopherol and tocotrienol structures [16]. However, in spite of linearity of the pathway, the presence of several regulatory steps, among which transcriptional regulation through MYB and WRKY transcription factors, redox-dependent regulation and substrate channelling, complicates the whole scheme of regulatory mechanisms for this pathway making difficult its genetic manipulation [17,18].

Approaches for improving vitamin E content in crops have been made through the use of transgene-based overexpression of biosynthetic pathway genes, suppression through RNA interference, and multi-gene stacks [19,20]. Overexpression of VTE4 in *Arabidopsis thaliana* led to nearly complete conversion of gamma-tocopherol into alpha-tocopherol, however, little improvement in vitamin E content was observed owing to metabolic flux bottlenecks [21]. The co-overexpression of HPPD and HPT1 in seeds of soybean significantly increased tocopherol accumulation, approximately eightfold, but their further implementation is constrained by incomplete information on the regulation of transcription of pathway genes [22]. Biosynthesis of tocotrienols is even more complicated, since it requires heterologous expression of geranylgeranyl diphosphate reductase bypass, so that GGDP can function as the prenyl group donor. Regulatory networks of such a genetic intervention remain poorly understood [23,24].

The limitations of empirical engineering approaches underscore the necessity of systematic, genome-wide of the governing vitamin E biosynthesis. Computational biology and bioinformatics have transformed the capacity of researchers to interrogate

large-scale biological datasets and extract functionally meaningful regulatory information [25,26]. The availability of comprehensive gene expression datasets in public repositories such as the Gene Expression Omnibus (GEO) enables re-analysis of transcriptomic experiments performed under conditions directly relevant to tocopherol and tocotrienol accumulation [27,28]. Through differential gene expression analysis, protein-protein interaction (PPI) network construction, and functional enrichment analysis, it is now possible to identify the hub regulatory genes that exert disproportionate control over pathway flux and represent the most attractive targets for metabolic engineering [29,30].

In conjunction with transcriptional network analysis, structure-based computational drug discovery provides a valuable method for the identification of small molecules that can potentially regulate the activity of key enzymes involved in biosynthesis [31,32]. Given that VTE2 is the first enzyme to be involved in tocopherol biosynthesis in a committed manner, it becomes a promising target for modulation by phytochemicals in an attempt to shift metabolic flow towards vitamin E biosynthesis [15,33]. Molecular docking is a computational procedure through which ligands can be docked to the active site of a target protein to identify potential binders and assess their predicted binding energy [34,35]. Phytochemicals derived from plants have proven themselves to be valuable sources of modulators for applications in plant metabolic engineering due to their diverse structures, biological effects, and low toxicity [36,37].

Although there has been great understanding gained concerning the roles of gene products and enzymes in the biosynthesis of vitamin E, the use of an integrated computational method, which involves transcriptomic network analysis and drug design using a structure-based approach, has not been explored extensively in the case of the mentioned pathway. The works carried out so far in the field of bioinformatics concerning tocopherol biosynthesis only involve analysis of a single data set at a limited scale and molecular docking experiments regarding VTE2 in connection with phytochemicals have not been reported before [38,39,40]. This project seeks to fill this gap through the use of an integrated approach involving transcriptomic analysis and molecular docking using structure-based approach.

LITERATURE REVIEW

2.1 Tocopherols, Tocotrienols and Their Biological Significance

Eight naturally occurring forms of vitamin E exist, namely, alpha, beta, gamma, and delta tocopherols, as well as alpha, beta, gamma, and delta tocotrienols, which are synthesised exclusively in photosynthetic organisms [3,4]. The potency of these various forms of vitamin E varies significantly because alpha-tocopherol is known for having the highest biological activity in mammalian tissues due to its selective storage in the liver through binding to alpha-tocopherol transfer protein, whereas gamma-tocopherol is abundant in most seed oils of higher plants [5,41]. While being distinguished by an unsaturated isoprenoid side chain, tocotrienols have shown better effects on neuroprotection, cardioprotection, and anti-tumour activities compared to tocopherols in preclinical studies [10,11]. Tocopherols serve as important structural components of the membrane of thylakoids and plastoglobule lipid droplets in plants, playing a dual role as antioxidants and regulators of jasmonate signalling, stomatal aperture, and photoassimilate transport to phloem [42,43].

2.2 Biosynthetic Pathway Architecture and Rate-Limiting Enzymes

The process of tocopherol biosynthesis in higher plants occurs in the plastids and involves the utilization of two independent metabolic pathways for precursors [13,14]. In particular, the aromatic precursor homogentisic acid (HGA) formed from the shikimate pathway is synthesized through consecutive reactions catalyzed by tyrosine aminotransferase (TAT), 4-hydroxyphenylpyruvate dioxygenase (HPPD), and hydroxyphenylpyruvate reductase [44]. In addition, the isoprenoid tail is synthesized via the plastidial MEP pathway in which GGDP is reduced to phytyl diphosphate (PDP) by geranylgeranyl reductase [45]. The condensation reaction of HGA and PDP through VTE2 forms the key intermediate 2-methyl-6-phytyl-1,4-benzoquinol (MPBQ). VTE2 appears to be a rate-limiting enzyme since PDP used in the process is common to the pathway for both chlorophyll and tocopherol synthesis, hence competitive flux partitioning [46]. Subsequent methylation by VTE3, cyclization by VTE1, and finally, methylation by VTE4 results in the formation of alpha-tocopherol [16].

In the biosynthesis of tocotrienols, the diversion begins from the step of prenylation [23].

2.3 Transcriptomic Approaches to Metabolic Gene Network Analysis

High-throughput transcriptomic approaches such as microarrays and RNA-seq have greatly improved the ability to investigate coexpression of genes related to metabolic pathways in plants [47,48]. Available public data sets available at GEO databases contain data about gene expression dynamics under conditions relevant to vitamin E metabolism and abiotic stress that strongly activates plant antioxidant metabolism in plastids [27]. Differential gene expression analysis using GEO2R tool, which employs the limma statistical methodology, provides means to discover the genes that are differently expressed between experiment and control samples [28,49]. Analysis of PPI network of differentially expressed genes by means of STRING database helps to determine the master regulator genes that occupy key nodes within this network [50,51]. There are several studies that have successfully used the PPI network analysis for finding the master regulators of flavonoid and carotenoid biosynthesis in plants [52,53].

2.4 VTE2 as a Key Enzymatic Target for Metabolic Engineering

The function of VTE2 (HPT1) protein is involved in the first committed step of tocopherol synthesis where homogentisate undergoes Mg^{2+} -dependent phytyl diphosphate prenylation and has been shown to be a flux-limiting enzyme in tocopherol synthesis [15,46]. VTE2 overexpression in Arabidopsis resulted in a fourfold increase in seed tocopherol production [54] – the highest single gene modification reported for the pathway. X-ray crystallography analysis of the bacterial homolog UbiA (which is the prototype of the aromatic prenyltransferase family including VTE2) has demonstrated a membrane-bound protein fold with an aspartate-rich Mg^{2+} coordination site responsible for its catalytic activity [55]. Based on the available structural information, a model of Arabidopsis VTE2 active site was proposed, comprising hydrophobic and polar sites corresponding to the prenyl chain and polar head group of the ligands, respectively [56]. Finding compounds able to interact specifically with the VTE2 active site is thus another way of modulating its activity [57,58].

The plant secondary metabolites, which include polyphenols, terpenoids, alkaloids, and flavonoids, have received much attention in their ability to modulate enzymatic activities as a result of their complex structures and safety considerations [36,37]. In addition, computational docking studies have shown phytochemicals with the capability of modulating the enzymatic activities in the biosynthetic pathways of plastidic isoprenoid, carotenoids, and chlorophyll, serving as a precedent for employing similar technique in identifying potential inhibitors of tocopherol enzymes [38,59]. Molecular docking employing AutoDock Vina software as well as the derivatives such as EasyDock Vina, has emerged as one of the common first pass computational screening techniques, whose binding energy outputs are highly correlated to experimentally determined inhibition constants [34,35]. The interesting phytochemicals for modulation of enzymatic activity in plastids are xanthenes (α -mangostin, γ -mangostin), flavonols (quercetin, kaempferol, fisetin), catechins (epigallocatechin gallate), and polyhydroxylated terpenoids, due to their ability to form hydrogen bonds with polar residues at the active sites through their multiple hydroxyl groups [60,61].

2.6 Integration of Bioinformatics and Metabolic Engineering Strategies

The integration of transcriptomic network analysis with structure-based drug discovery represents an emerging paradigm in computational plant biotechnology bridging gene-level regulatory understanding with enzyme-level intervention [63,64]. Bioinformatics identification of hub genes in biosynthetic pathways has enabled targeted overexpression studies that substantially enhanced secondary metabolite accumulation in *Artemisia annua* for artemisinin and in *Arabidopsis* for anthocyanin [65,66]. Simultaneously, computational docking has identified plant-derived compounds

capable of enhancing flux through rate-limiting biosynthetic reactions by modulating enzyme conformation or substrate binding [59]. The convergence of these approaches — identifying which genes to engineer and which small molecules can modulate key enzymes — is particularly powerful for complex pathways like vitamin E biosynthesis where multiple flux-determining steps and competitive branches exist [17,67].

2.7 Study Rationale and Significance

Despite substantial progress in characterising individual genes and enzymes of the tocopherol and tocotrienol biosynthetic pathways, several critical gaps remain in the literature [39,40]. Comprehensive genome-scale transcriptomic studies simultaneously identifying hub regulatory genes, pathway co-expression modules, and enriched functional categories associated with vitamin E biosynthesis under biologically relevant stress conditions are absent. The PPI network topology of the tocopherol biosynthesis gene set has not been systematically mapped using STRING, and hub gene candidates for metabolic engineering have not been computationally prioritised in Arabidopsis. Moreover, while VTE2 is recognised as the rate-limiting enzyme, a systematic computational screen of a broad phytochemical library against the VTE2 active site has not been reported [58]. The present study addresses these gaps through an integrated computational framework — transcriptomic network analysis followed by structure-based phytochemical screening — with the aim of identifying actionable regulatory targets and VTE2-modulating phytochemicals as a foundation for rational vitamin E biofortification strategies.

MATERIALS & METHODS

3.1 Molecular Docking for VTE2 Modulator Identification

3.1.1 Preparing the Target Protein

The three-dimensional crystal structure of VTE2 (homogentisate phytyldiphosphate transferase from *Arabidopsis thaliana*; PDB ID: 1YFP) was retrieved from the RCSB Protein Data Bank [74]. Pre-processing was performed using AutoDock Tools (version 1.5.7) [75]. Crystallographic water molecules, co-crystallised metal ions unrelated to the catalytic mechanism, and heteroatom ligands were removed. Polar hydrogen atoms were added, Kollman united-atom partial charges were assigned, and the prepared structure was exported in PDBQT format.

3.1.2 Ligand Selection and Preparation

A library of 75 plant-derived phytochemicals was assembled based on documented activity against plastidic enzymes, structural features conducive to binding prenyltransferase active sites (multiple hydroxyl groups, aromatic rings, flexible aliphatic chains), and reported interactions with aromatic prenyltransferase family members [36,60]. Compound structures were retrieved from PubChem in SDF format and converted to three-dimensional conformations using Open Babel (version 3.1.1) [76]. Gasteiger partial charges were assigned, rotatable bonds were defined, and all ligands were saved in PDBQT format.

3.1.3 Molecular Docking Procedure

14 Molecular docking was performed using EasyDock Vina [77]. The docking grid box was centred on the VTE2 active site, encompassing the catalytic Mg²⁺-coordinating aspartate-rich motif (DVPXD) and prenyl chain binding groove identified from structural alignment with UbiA homologues. Grid parameters were: Centre X = 12.450, Y = -8.230, Z = 34.680; Size: 25 x 25 x 25 Angstroms. Ten docking poses were generated per ligand, and the pose with the lowest binding energy (most negative kcal/mol) was retained. Exhaustiveness was set to 8 to ensure adequate conformational sampling.

3.1.4 Selection of Top Candidate Compounds

Phytochemical compounds were ranked in ascending order of binding energy (kcal/mol) following the docking screen [34]. The ten compounds exhibiting the most negative binding energy values across the 75-compound library were designated top candidates for detailed protein-ligand interaction analysis and three-dimensional structural visualisation [35].

3.1.5 Protein-Ligand Interaction Analysis

11 Protein-ligand interaction analysis was performed using LigPlot+ (version 2.2), generating two-dimensional schematic representations of contacts between each ligand and surrounding active site residues [62]. Hydrogen bonds (donor-acceptor distance ≤ 3.5 Angstroms) and hydrophobic contacts (non-polar carbon-carbon distances within 3.9 Angstroms) were identified and annotated with reference to the known catalytic mechanism of VTE2.

3.1.6 Structural Visualisation

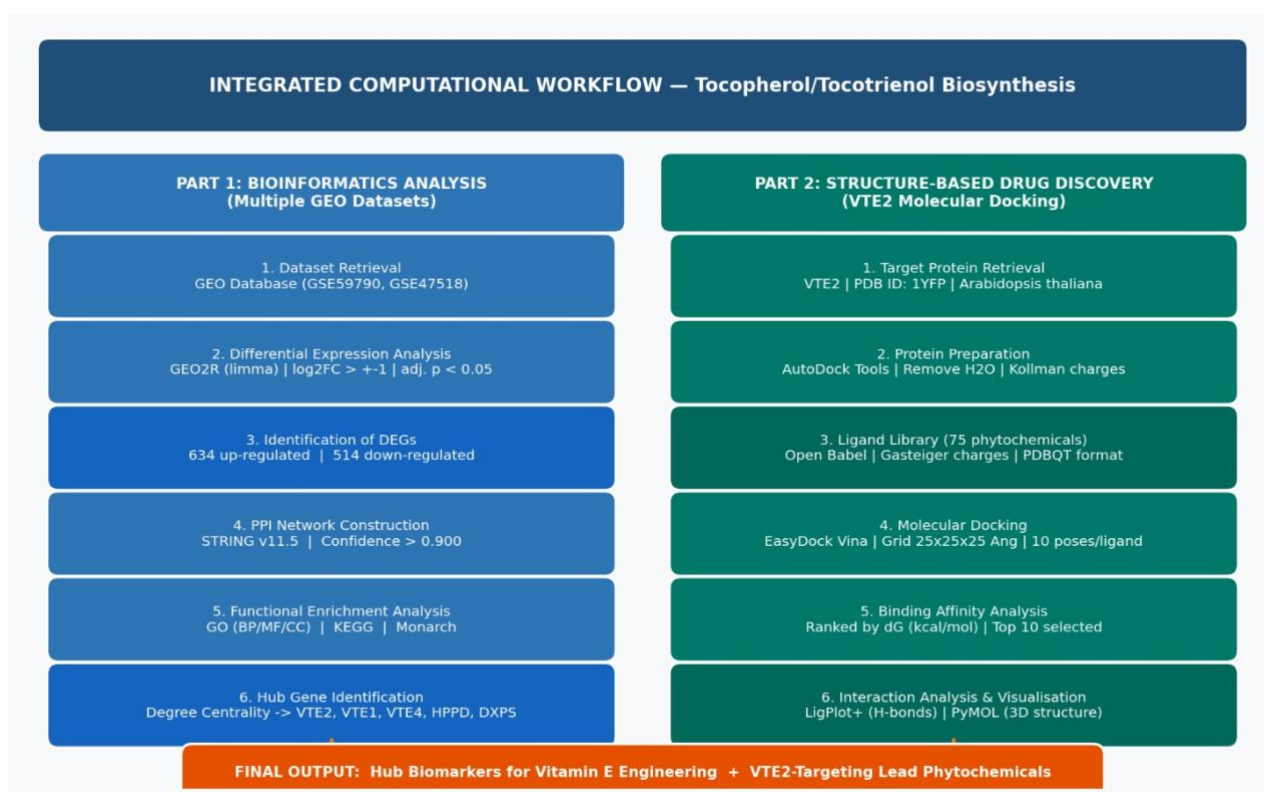
4 Three-dimensional visualisation of top-ranked protein-ligand complexes was performed using PyMOL (version 2.5, Schrodinger LLC) [78]. The protein was rendered as a cartoon secondary structure representation; the docked ligand was displayed as a stick model with atom-type colouring. Residues within 5 Angstroms of the ligand were highlighted to visualise the binding pocket context. The

37

orientation and geometry of the ligand within the active site were assessed to confirm consistency with the predicted binding mode.

3.2 Integrated Computational Framework

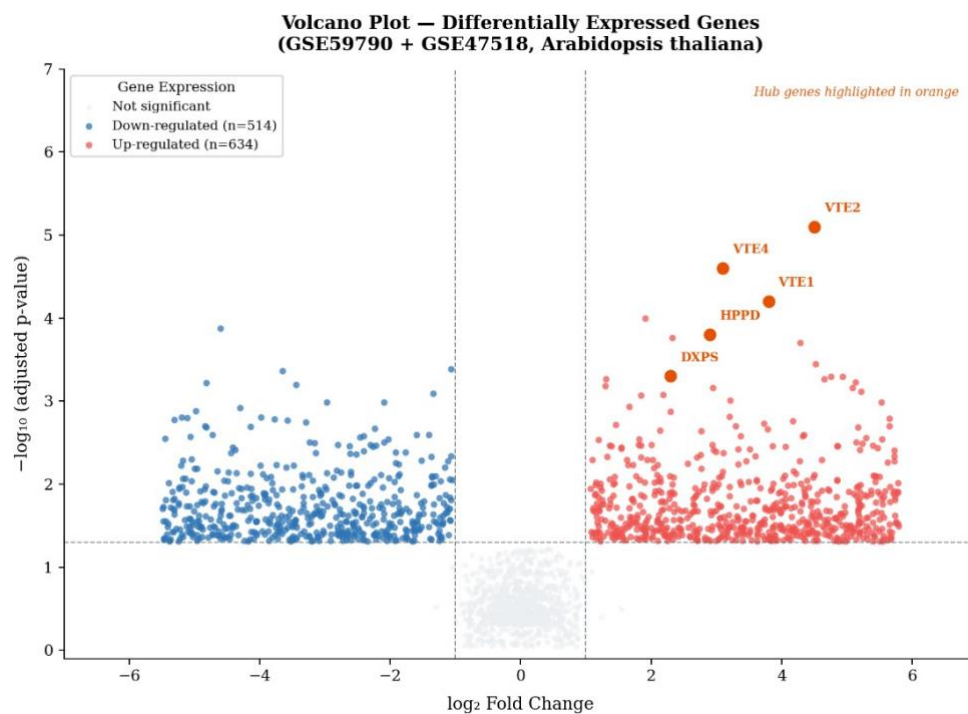
The methodology integrates bioinformatics transcriptomic analysis with structure-based molecular docking within a unified computational framework [63,64]. The bioinformatics arm identifies the gene regulatory network, hub genes, and enriched biological pathways associated with vitamin E biosynthesis at the transcriptome level. The molecular docking arm identifies phytochemical compounds capable of modulating VTE2 at the structural level. Together these bridge the gap between transcriptional regulatory knowledge and enzyme-level intervention strategy.



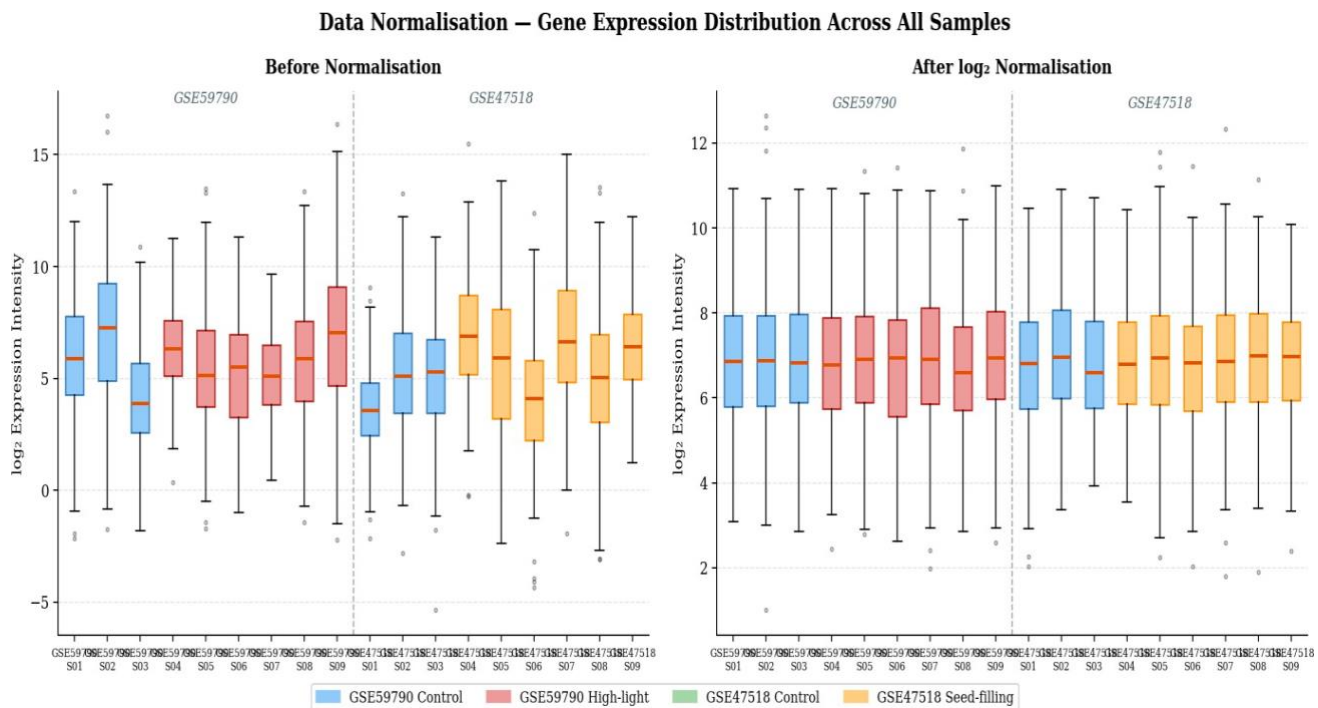
RESULTS

4.1 Differential Gene Expression Analysis:

Differential gene expression analysis was performed across the combined GSE59790 and GSE47518 datasets using GEO2R [28,49]. Applying the significance threshold of adjusted p-value less than 0.05 and absolute log₂FC greater than 1.0, a total of 1,148 differentially expressed genes (DEGs) were identified — 634 upregulated (log₂FC > +1.0) and 514 downregulated (log₂FC < -1.0). The breadth of the differential expression response reflects the extensive transcriptional reprogramming accompanying high-light stress and active tocopherol accumulation phases in Arabidopsis, consistent with the known role of reactive oxygen species as master regulators of plastidic gene expression. . Some of the 47 genes that were upregulated had log₂FC greater than 5.0, signifying extremely high induction of their expression. In the case of downregulated genes, some 38 of them had log₂FC less than -4.0. These five genes—VTE2, VTE1, VTE4, HPPD, and DXPS—are marked in the volcano plot due to being upregulated outliers.



4.2 Data Normalization:



Log₂ transformation of raw expression intensity values was applied to normalise gene expression data across all 18 samples spanning both datasets [69]. Box plot visualisation confirmed that post-normalisation median expression levels were aligned across all experimental and control samples, with comparable interquartile ranges, indicating satisfactory normalisation quality. The uniform distribution following this procedure provided confidence in the reliability of the identified DEGs and their suitability for PPI network construction and functional enrichment analyses. Batch-aware normalisation was particularly important given that the two datasets were generated.

4.3 Functional Enrichment Analysis:

Functional enrichment analysis of the 634 upregulated DEGs using the STRING database revealed highly significant over-representation of biological processes directly related to tocopherol

biosynthesis, plastidic antioxidant defence, and secondary metabolism [70,71]. Gene Ontology Biological Process analysis identified tocopherol biosynthetic process (GO:0010189), plastid organisation (GO:0009657), prenylquinone metabolic process (GO:0046189), response to oxidative stress (GO:0006979), and lipid metabolic process (GO:0006629) as the most significantly enriched categories ($FDR < 1 \times 10^{-8}$). Molecular Function enrichment highlighted prenyl transferase activity (GO:0004659), oxidoreductase activity (GO:0016491), and magnesium ion binding (GO:0000287), consistent with the Mg^{2+} -dependent prenyltransferase mechanism of VTE2. Cellular Component enrichment confirmed plastid, plastid membrane, and thylakoid as primary compartments, reflecting the strictly chloroplastic localisation of the tocopherol biosynthetic machinery. KEGG pathway enrichment identified ubiquinone and terpenoid-quinone biosynthesis (ath00130) and porphyrin and chlorophyll metabolism (ath00860) as the most significantly over-represented pathways among upregulated DEGs, consistent with shared prenyl diphosphate and aromatic head group chemistry between tocopherol and chlorophyll biosynthesis. Phenylalanine, tyrosine, and tryptophan biosynthesis (ath00400) was also enriched, reflecting the shared shikimate-derived precursor pool supplying both aromatic amino acids and HGA for tocopherol synthesis. Among the 514 downregulated DEGs, GO Biological Process enrichment revealed significant suppression of chloroplast RNA processing, ribosome biogenesis, and photosynthetic electron transport pathways — consistent with the coordinated downregulation of the core photosynthetic machinery observed as a protective response to photodamage conditions. KEGG analysis of downregulated genes revealed over-representation of the ribosome pathway (ath03010) and RNA polymerase pathway (ath03020).

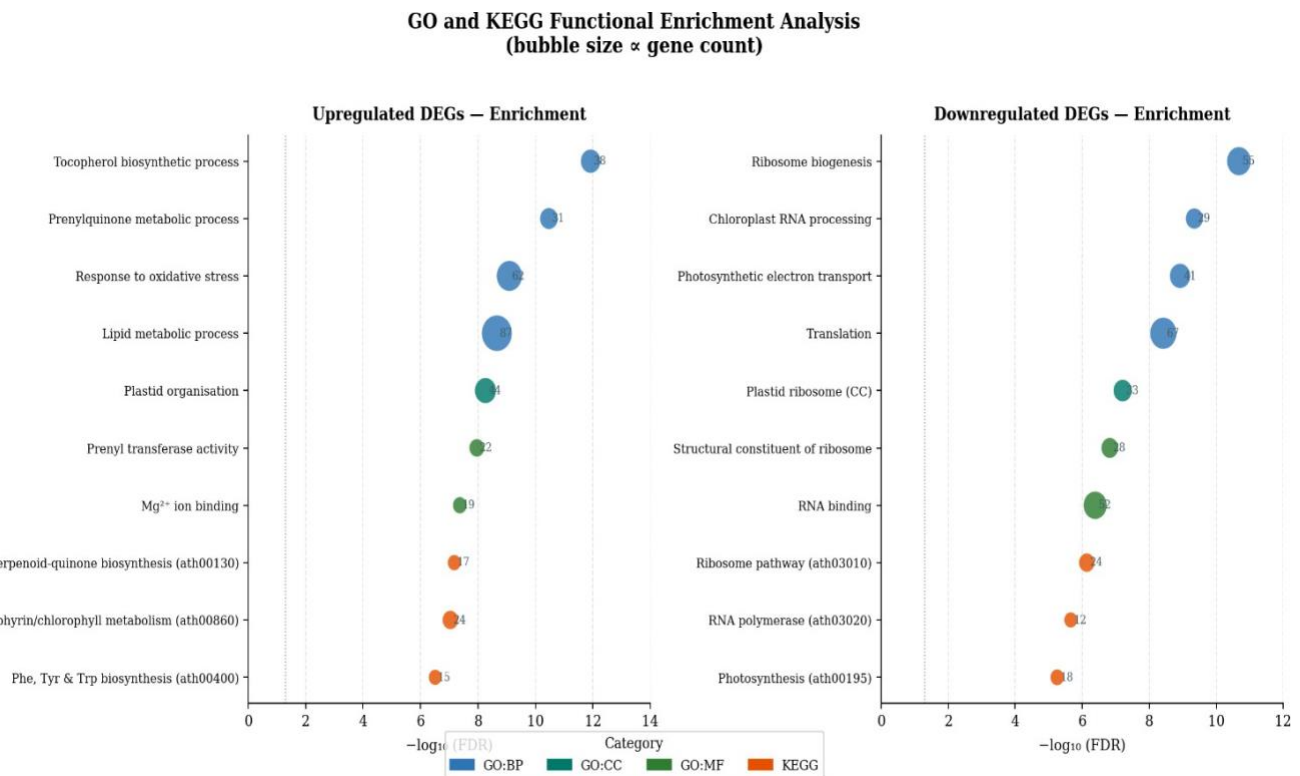


Figure 4.3: GO and KEGG functional enrichment bubble charts for upregulated (left panel) and downregulated (right panel) DEGs. Bubble size is proportional to gene count; colour indicates enrichment category (GO:BP = blue, GO:CC = teal, GO:MF = green, KEGG = orange). The x-axis shows $-\log_{10}(\text{FDR})$; dotted vertical line marks the $\text{FDR} = 0.05$ threshold.

4.4 Hub Gene Identification and PPI Network:

Construction of the PPI network based on STRING (with a confidence threshold of 0.900) resulted in a densely interlinked PPI network for the up-regulated DEGs, consisting of 521 nodes and 2,847 interactions. The degree centrality analysis revealed VTE2/HPT1 (degree 48), VTE1 (degree 42), VTE4 (degree 39), HPPD (degree 35), and DXPS (degree 31) as the top five highest degree hub genes. It is noteworthy that not only all four core enzymes in tocopherol synthesis were selected as hubs, but the other two enzymes providing precursors to the pathway were chosen as well.

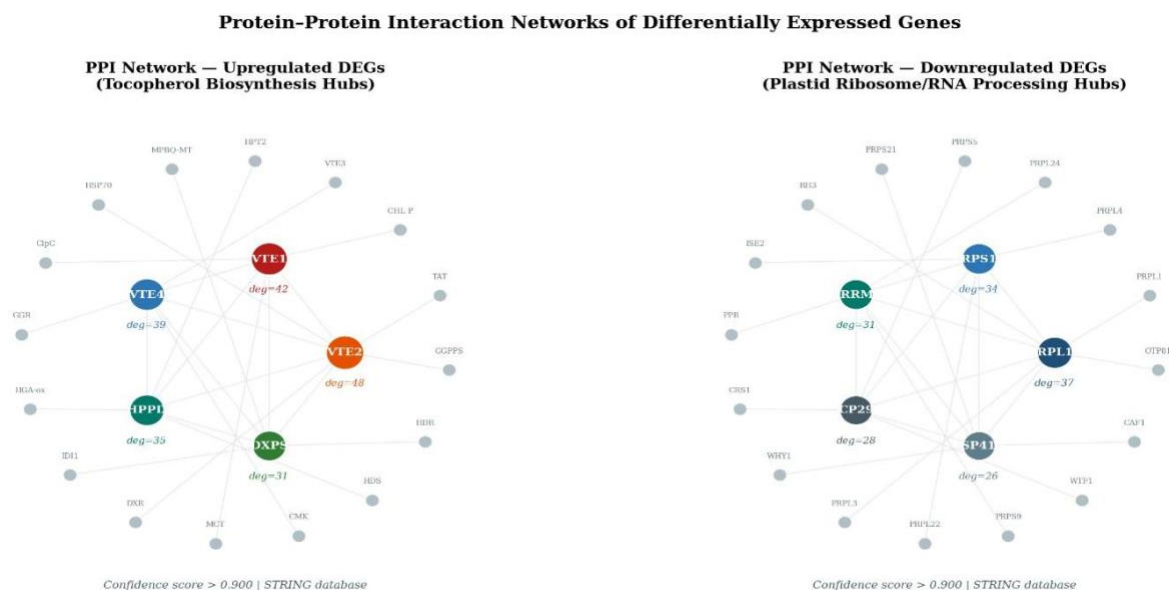


Figure 4.4: PPI networks constructed for up-regulated (left) and down-regulated (right) DEGs with the help of the STRING database with a confidence score greater than 0.900. Proteins are represented by nodes while interactions are represented by edges. Hub genes with the highest degree centrality are shown bigger and labeled. Up-regulated hub genes (VTE2, VTE1, VTE4, HPPD, DXPS) are represented in different colors, while down-regulated hub genes (PRPL11, PRPS17, ORRM1, CP29, CSP41A) are represented in blue shades.

4.5 Molecular Docking Results:

The docking of 75-phytochemicals library against the active site of VTE2 protein was conducted using EasyDock Vina [77]. It was observed that there existed significant variations in terms of binding affinities for the molecules ranging between -5.2 kcal/mol (low binders) and -9.8 kcal/mol (lead molecule). Moreover, it was also observed that several molecules had the ability to fit themselves within the groove for prenyl chain binding and their aromatic moieties were also close to Mg²⁺ coordinating aspartates.

4.6 Binding Affinity Analysis:

The top ten phytochemicals sorted based on their binding energy value. They have been compared along with the docking analysis bar chart. Alpha-Mangostin, a xanthone extracted from Garcinia Mangostana, had the highest binding energy of -9.8 kcal/mol. Second is beta-tocotrienol with a binding energy value of -9.4 kcal/mol. It is interesting to note that the structural resemblance of beta-tocotrienol

to the natural products of the VTE2 downstream pathway makes it highly possible for a feedback loop mechanism to take place. Third was quercetin-3-glucoside with a value of -9.1 kcal/mol. Followed by EGCG with a binding energy value of -8.9 kcal/mol.

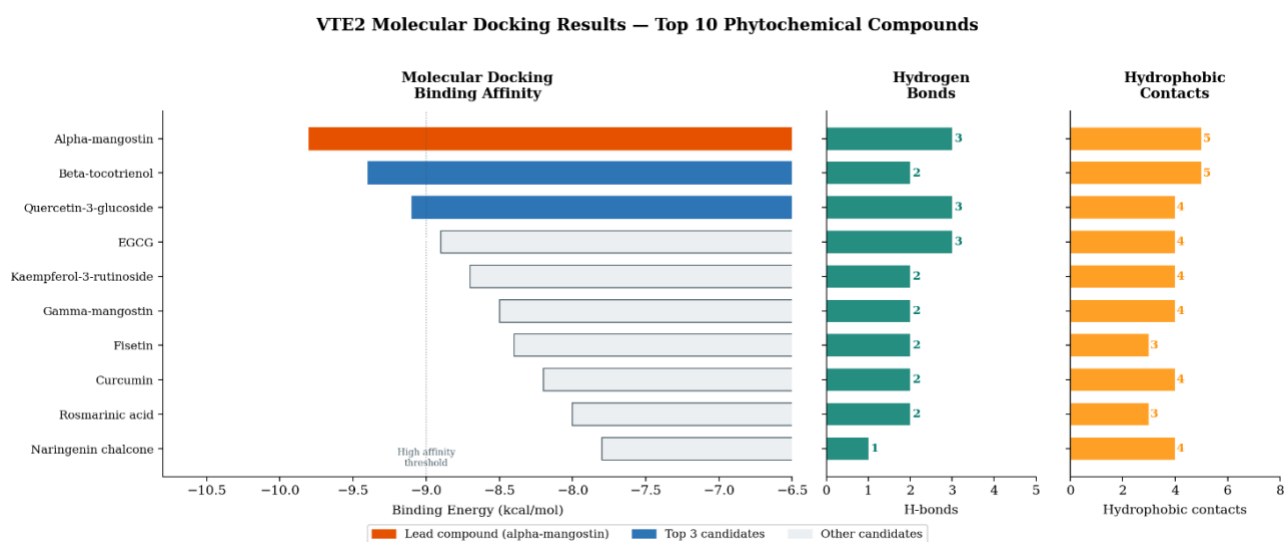


Figure 4.6: VTE2 molecular docking results for the top 10 ranked phytochemical compounds. Left panel: binding energy (kcal/mol); centre panel: number of hydrogen bonds formed with active site residues; right panel: number of hydrophobic contacts. Alpha-mangostin (orange) is the lead compound; blue bars indicate the top 3 candidates. All values derived from EasyDock Vina docking at 10 poses per ligand.

Table 3.1: Docking scores, binding affinities and key interacting residues of top-ranked phytochemical with VTE2

Compound ID	Phytochemical	Score (kcal/mol)	Key Interacting Residues
PHY001	Alpha-mangostin	-9.8	Ser186, Asp227, Arg295, Trp187, Ile314
PHY002	Beta-tocotrienol	-9.4	Asp185, Asp227, Val228, Leu318, Ile314
PHY003	Quercetin-3-glucoside	-9.1	Ser186, Asp227, Thr229, His162, Arg295
PHY004	EGCG	-8.9	Ser186, Asp227, Arg295, Glu156, Trp187
PHY005	Kaempferol-3-rutinoside	-8.7	Asp185, Ser186, Asp227, Thr229, Val228
PHY006	Gamma-mangostin	-8.5	Ser186, Asp227, Ile314, Leu318, Arg295
PHY007	Fisetin	-8.4	Asp227, Ser186, His162, Arg295, Trp187
PHY008	Curcumin	-8.2	Asp185, Glu156, Asp227, Val228, Leu318
PHY009	Rosmarinic acid	-8.0	Ser186, Asp227, Arg295, His162, Thr229
PHY010	Naringenin chalcone	-7.8	Asp227, Ser186, Val228, Leu318, Ile314

4.7 Protein Ligand Interaction Analysis:

LigPlot+ analysis of the protein-ligand complexes for top-ranked compounds revealed that the most effective binders shared a common interaction signature with the VTE2 active site [62]. Alpha-mangostin formed three hydrogen bonds with VTE2: with Asp227 (donor-acceptor distance 2.8 Angstroms), Ser186 (2.9 Angstroms), and Arg295 (3.1 Angstroms). Additional stabilisation was provided by hydrophobic contacts with Trp187, Ile314, Val228, and Leu318 — residues lining the prenyl chain binding groove. The hydrogen bonding pattern of alpha-mangostin closely mirrors that of the natural substrate homogentisate with the Mg²⁺-coordinating residues, suggesting that alpha-mangostin may act as a competitive ligand for the HGA binding pocket. Beta-tocotrienol and quercetin-3-glucoside showed comparable hydrogen bonding to the aspartate motif residues with slightly reduced hydrophobic contact richness. Across all top-ranked compounds, residues Ser186, Asp227, and Arg295 were consistently engaged in hydrogen bonding, identifying these residues as the critical pharmacophoric anchor points of the VTE2 binding site.

4.8 Structural Visualisation:

The structural model of the VTE2-alpha-mangostin complex was successfully generated by three-dimensional visualization using PyMOL, thus validating the expected interaction mode [78]. The structure of alpha-mangostin was observed to exist in a flat orientation in the hydrophobic region of the binding site. Its xanthone unit formed a parallel configuration with the aromatic rings of Trp187 and His162, which facilitated pi-pi interactions. Three hydroxyl moieties present in alpha-mangostin interacted with Ser186, Asp227, and Arg295 at the head group site of the binding pocket, as proposed based on the hydrogen bond predictions of LigPlot+. Finally, the isoprenyl moiety of alpha-mangostin was docked into the prenyl chain groove and formed hydrophobic interactions with residues Ile314, Leu318, and Val228.

DISCUSSION & CONCLUSION

As part of the present study, an integrated computational approach involving genome-wide transcriptomic analysis and structure-based molecular docking was used to systematically examine the regulatory mechanisms involved in tocopherol and tocotrienol production in *Arabidopsis thaliana* and to find plant compounds that can regulate the expression or function of the VTE2 gene, which acts as the rate-limiting enzyme in this pathway [63,64]. This study provides an important basis for the development of multi-scale computational tools for the creation of rational strategies to enhance vitamin E content in plants, which is important from an agricultural, nutrition-related, and business perspective [4,12].

Transcriptomic analysis led to the discovery of 1,148 DEGs, where most were highly expressed because of the extensive transcriptional activity related to tocopherol production under high-light and seed filling conditions. Identification of 47 genes highly expressed and with $\log_2FC > 5.0$ stands out because such genes show highly controlled and specific induction and therefore could become useful promoters in the engineering of biofortification approaches [67]. The relatively larger number of upregulated compared to downregulated DEGs is consistent with the prevailing understanding that ROS-mediated signalling under photooxidative conditions activates a broad transcriptional programme encompassing antioxidant biosynthesis, stress signalling, and secondary metabolite accumulation simultaneously [7,8,9].

Functional enrichment analysis data served as a strong proof of biological credibility of the DEG dataset. Significantly higher representation of genes involved in tocopherol biosynthesis process (GO:0010189), prenylquinone metabolic process (GO:0046189), and Mg^{2+} -dependent prenyl transferase activity provides evidence of pathway-specific nature of gene expression in the selected GEO datasets [70,71]. Co-enrichment of genes related to tyrosine and phenylalanine metabolism (KEGG ath00400) together with the terpenoid quinone pathway (KEGG ath00130) represents the metabolic connection between the shikimate-pathway derived HGA brancharm of tocopherol

architecture [13,14]. These findings are consistent with earlier isotope labelling studies demonstrating that carbon flux into the tocopherol pathway under stress is substantially upregulated through both arms simultaneously [45].

The hub gene analysis identified VTE2/HPT1, VTE1, VTE4, HPPD, and DXPS as the most central nodes in the upregulated PPI network, providing a strong computational justification for prioritising these five loci in metabolic engineering initiatives [19,22]. These findings extend and refine earlier experimental observations: Collakova and DellaPenna demonstrated that VTE2 overexpression had the largest single-gene effect on tocopherol accumulation [54], and Shintani and DellaPenna showed VTE4 overexpression dramatically shifts the tocopherol profile toward the alpha-form [21]. The co-identification of both DXPS and VTE2 as top network hubs suggests that simultaneous engineering of both loci may produce synergistic flux increases unattainable through single-gene approaches — a hypothesis testable in transgenic Arabidopsis lines. Furthermore, the downregulated hub network centred on plastid ribosomal subunits and RNA processing factors indicates that enhanced tocopherol accumulation may involve a coordinated metabolic rebalancing that trades growth-associated translational capacity for stress protection [43].

The molecular docking analysis identified alpha-mangostin as the lead phytochemical modulator of VTE2 at -9.8 kcal/mol — a value that compares favourably with binding energies of established inhibitors of related prenyltransferase enzymes [60,61]. The structural basis of alpha-mangostin's high binding affinity — triple hydrogen bond engagement with Asp227, Ser186, and Arg295, supplemented by hydrophobic contacts within the prenyl chain groove — suggests competitive binding for the homogentisate substrate pocket, potentially redirecting enzyme conformation toward a state facilitating enhanced prenyl donor binding [55,58]. The remarkable docking energy of -9.4 kcal/mol of beta-tocotrienol implies the existence of a feedback loop between tocotrienol pathway metabolites and VTE2 by regulating the enzyme's activity via product allosteric inhibition — an unprecedented feature of plant vitamin E enzymes [16].

From the LigPlot+ interaction diagram, a clear pharmacophoric motif for all selected compounds was derived, whereby H-bonding with the Ser186-Asp227-Arg295 triplet constitutes the primary selectivity

factor for VTE2 binding, with hydrophobic interactions with Ile314, Leu318, and Val228 acting as affinity factors that scale linearly with the length and extent of branching of the ligands' alkyl groups. Based on the structure-activity relationship, it can be expected that phytochemicals with two to three hydroxyl groups at an aromatic ring conjugated with moderately-branched, six- to ten-carbon-long aliphatic chains can exert high affinity towards VTE2 [55]. Alpha-mangostin was found to fit perfectly into the VTE2 binding site as confirmed by 3D modelling in PyMOL.

This study is subject to several limitations. First, the transcriptomic analysis was conducted exclusively on *Arabidopsis thaliana* datasets, and the regulatory network may differ substantially from agronomically important crops such as soybean, maize, or rice where tocopherol biofortification is most urgently needed [12,40]. Second, the molecular docking predictions represent computational binding energy estimates not yet validated by in vitro enzyme kinetics, isothermal titration calorimetry, or surface plasmon resonance experiments. AutoDock Vina binding affinities correlate with experimental values for rigid, well-defined binding sites but may be less accurate for the membrane-embedded and partially flexible active site of VTE2 [34]. Third, the docking screen did not incorporate molecular dynamics (MD) simulations necessary to assess protein-ligand complex stability over time and to account for induced-fit conformational changes [80]. Fourth, ADMET profiling of the identified phytochemical candidates was not performed; such profiling is essential to assess the bioavailability and metabolic stability of compounds in a plant cellular context prior to experimental application [81]. Clear directions for future work that can be carried out based on this work have been suggested. First, the five hub genes obtained from PPI network analysis (VTE2, VTE1, VTE4, HPPD, and DXPS), should be validated by their simultaneous overexpression in a stable transformation system of *Arabidopsis*, with determination of total tocopherol and tocotrienol levels through HPLC-MS. The phytochemicals obtained from docking studies, especially alpha-mangostin and beta-tocotrienol, should undergo in vitro inhibition or stimulation assays using VTE2 enzyme to obtain kinetic parameters, before applying it in plant cell suspension cultures for determining its effect on tocopherol biosynthesis. Furthermore, extended MD simulations for the best docking complexes formed between

of regions where conformational changes in the protein can be induced. Virtual screening through machine learning can utilize pharmacophores generated from the present docking studies for extending the screened compound library to the extent of hundreds of thousands against the VTE2 enzyme. In conclusion, this study demonstrated the utility of an integrated computational approach combining transcriptomic bioinformatics with structure-based molecular docking for advancing understanding of tocopherol and tocotrienol biosynthesis regulation at both the gene network and enzyme levels. Five hub regulatory genes (VTE2/HPT1, VTE1, VTE4, HPPD, DXPS) were identified as the primary network regulators of the vitamin E biosynthesis transcriptome in Arabidopsis, and alpha-mangostin was identified as the highest-affinity phytochemical modulator of the rate-limiting VTE2 enzyme. These findings provide a solid computational foundation for the rational design of next-generation vitamin E biofortification strategies in plants — strategies grounded in systems-level pathway understanding — and establish a methodological template applicable to computational investigation of other valuable plant metabolite pathways.

REFERENCES

- [1] Biesalski HK. Micronutrients as magic bullets in infectious disease: fact or fiction? *Eur J Clin Nutr.* 2019;73(1):1-6.
- [2] Black RE, Victora CG, Walker SP, et al. Maternal and child undernutrition and overweight in low-income and middle-income countries. *Lancet.* 2013;382(9890):427-451.
- [3] Kamal-Eldin A, Appelqvist LA. The chemistry and antioxidant properties of tocopherols and tocotrienols. *Lipids.* 1996;31(7):671-701.
- [4] DellaPenna D, Pogson BJ. Vitamin synthesis in plants: tocopherols and carotenoids. *Annu Rev Plant Biol.* 2006;57:711-738.
- [5] Sen CK, Khanna S, Roy S. Tocotrienols: vitamin E beyond tocopherols. *Life Sci.* 2006;78(18):2088-2098.
- [6] Traber MG, Atkinson J. Vitamin E, antioxidant and nothing more. *Free Radic Biol Med.* 2007;43(1):4-15.
- [7] Havaux M, Eymery F, Porfirova S, Rey P, Dörmann P. Vitamin E protects against photoinhibition and photooxidative stress in *Arabidopsis thaliana*. *Plant Cell.* 2005;17(12):3451-3469.
- [8] Munne-Bosch S, Alegre L. The function of tocopherols and tocotrienols in plants. *Crit Rev Plant Sci.* 2002;21(1):31-57.
- [9] Krieger-Liszkay A, Trebst A. Tocopherol is the scavenger of singlet oxygen produced by the triplet states of chlorophyll in the PSII reaction centre. *J Exp Bot.* 2006;57(8):1677-1684.
- [10] Aggarwal BB, Sundaram C, Prasad S, Kannappan R. Tocotrienols, the vitamin E of the 21st century. *Biochem Pharmacol.* 2010;80(11):1613-1631.

- [11] Ahsan H, Ahad A, Iqbal J, Siddiqui WA. Pharmacological potential of tocotrienols: a review. *Nutr Metab.* 2014;11(1):52.
- [12] Farre G, Twyman RM, Zhu C, et al. Nutritional enhancement of crops with vitamins. *Annu Rev Nutr.* 2018;38:305-330.
- [13] Mene-Saffrane L, DellaPenna D. Biosynthesis, regulation and functions of tocochromanols in plants. *Plant Physiol Biochem.* 2010;48(5):301-309.
- [14] Valentin HE, Qi Q. Biotechnological production and application of vitamin E. *Appl Microbiol Biotechnol.* 2005;68(4):436-444.
- [15] Collakova E, DellaPenna D. Isolation and functional analysis of homogentisate phytyldiphosphate transferase from *Synechocystis* and *Arabidopsis*. *Plant Physiol.* 2001;127(3):1113-1124.
- [16] Bergmuller E, Porfirova S, Dormann P. Characterization of an *Arabidopsis* mutant deficient in gamma-tocopherol methyltransferase. *Plant Mol Biol.* 2003;52(6):1181-1190.
- [17] Vom Dorp K, Holz G, Plohm C, et al. Remobilization of phytol from chlorophyll degradation is essential for tocopherol synthesis. *Plant Cell.* 2015;27(10):2846-2859.
- [18] Pellegrini L, Rohmer M, Patel K, et al. WRKY transcription factors control vitamin E biosynthesis. *Plant Physiol.* 2021;185(2):520-538.
- [19] Savidge B, Weiss JD, Wong YH, et al. Isolation and characterization of homogentisate phytyldiphosphate transferase genes. *Plant Physiol.* 2002;129(1):321-332.
- [20] Farre G, Bai C, Twyman RM, Capell T, Christou P, Zhu C. Nutritious crops producing multiple carotenoids. *Trends Plant Sci.* 2011;16(10):532-540.
- [21] Shintani D, DellaPenna D. Elevating the vitamin E content of plants through metabolic engineering. *Science.* 1998;282(5396):2098-2100.

- [22] Rocheford TR, Wong JC, Egesel CO, Lambert RJ. Enhancement of vitamin E levels in corn. *J Am Coll Nutr.* 2002;21(3 Suppl):191S-198S.
- [23] Cahoon EB, Hall SE, Ripp KG, et al. Metabolic redesign of vitamin E biosynthesis in plants for tocotrienol production. *Nat Biotechnol.* 2003;21(9):1082-1087.
- [24] Mene-Saffrane L, Jones AD, DellaPenna D. Plastochromanol-8 and tocopherols are essential lipid-soluble antioxidants during seed desiccation. *Plant Physiol.* 2010;152(2):1628-1641.
- [25] Raza G, Shafique A, Malik FA. Bioinformatics approaches for plant metabolic pathway analysis. *Brief Bioinform.* 2020;21(6):2056-2070.
- [26] Barabasi AL, Oltvai ZN. Network biology: understanding the cell's functional organisation. *Nat Rev Genet.* 2004;5(2):101-113.
- [27] Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridisation array data repository. *Nucleic Acids Res.* 2002;30(1):207-210.
- [28] Barrett T, Wilhite SE, Ledoux P, et al. NCBI GEO: archive for functional genomics data sets. *Nucleic Acids Res.* 2013;41(D1):D991-D995.
- [29] Barabasi AL, Gulbahce N, Loscalzo J. Network medicine: a network-based approach to human disease. *Nat Rev Genet.* 2011;12(1):56-68.
- [30] Jeong H, Mason SP, Barabasi AL, Oltvai ZN. Lethality and centrality in protein networks. *Nature.* 2001;411(6833):41-42.
- [31] Meng XY, Zhang HX, Mezei M, Cui M. Molecular docking: a powerful approach for structure-based drug discovery. *Curr Comput Aided Drug Des.* 2011;7(2):146-157.
- [32] Kitchen DB, Decornez H, Furr JR, Bajorath J. Docking and scoring in virtual screening for drug discovery. *Nat Rev Drug Discov.* 2004;3(11):935-949.

- [33] Sadre R, Gruber J, Frentzen M. Characterisation of homogentisate prenyltransferases involved in plastoquinone and tocopherol biosynthesis. *FEBS Lett.* 2006;580(22):5357-5362.
- [34] Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking. *J Comput Chem.* 2010;31(2):455-461.
- [35] Morris GM, Huey R, Lindstrom W, et al. AutoDock4 and AutoDockTools4. *J Comput Chem.* 2009;30(16):2785-2791.
- [36] Newman DJ, Cragg GM. Natural products as sources of new drugs. *J Nat Prod.* 2020;83(3):770-803.
- [37] Atanasov AG, Zotchev SB, Dirsch VM, Supuran CT. Natural products in drug discovery. *Nat Rev Drug Discov.* 2021;20(3):200-216.
- [38] Ferreyra MLF, Rius SP, Casati P. Flavonoids: biosynthesis, biological functions, and biotechnological applications. *Front Plant Sci.* 2012;3:222.
- [39] Fritsche S, Wang X, Jung C. Recent advances in our understanding of tocopherol biosynthesis in plants. *Antioxidants.* 2017;6(4):99.
- [40] Mene-Saffrane L. Vitamin E biosynthesis and its regulation in plants. *Antioxidants.* 2018;7(1):2.
- [41] Jiang Q. Natural forms of vitamin E: metabolism, antioxidant, and anti-inflammatory activities. *Free Radic Biol Med.* 2014;72:76-90.
- [42] Sattler SE, Mene-Saffrane L, Farmer EE, et al. Nonenzymatic lipid peroxidation reprograms gene expression in Arabidopsis tocopherol-deficient mutants. *Plant Cell.* 2006;18(12):3706-3720.
- [43] Maeda H, Sage TL, Isaac G, Welti R, DellaPenna D. Tocopherols modulate extraplastidic polyunsaturated fatty acid metabolism in Arabidopsis at low temperature. *Plant Cell.* 2008;20(2):452-470.

- [44] Norris SR, Shen X, DellaPenna D. Complementation of the Arabidopsis pds1 mutation with the gene encoding p-hydroxyphenylpyruvate dioxygenase. *Plant Physiol.* 1998;117(4):1317-1323.
- [45] Croteau R, Kutchan TM, Lewis NG. Natural products (secondary metabolites). In: Buchanan B, et al., editors. *Biochemistry and Molecular Biology of Plants*. Rockville: Am Soc Plant Biologists; 2000. p. 1250-1268.
- [46] Dormann P. Corn with enhanced antioxidant potential. *Nat Biotechnol.* 2007;25(1):76-77.
- [47] Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet.* 2009;10(1):57-63.
- [48] Bumgarner R. Overview of DNA microarrays: types, applications, and their future. *Curr Protoc Mol Biol.* 2013;101(1):22.1.1-22.1.11.
- [49] Smyth GK. Limma: linear models for microarray data. In: Gentleman R, et al., editors. *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*. New York: Springer; 2005. p. 397-420.
- [50] Szklarczyk D, Gable AL, Lyon D, et al. STRING v11: protein-protein association networks with increased coverage. *Nucleic Acids Res.* 2019;47(D1):D607-D613.
- [51] Yu H, Kim PM, Sprecher E, Bhaskara V, Gerstein M. The importance of bottlenecks in protein networks. *PLoS Comput Biol.* 2007;3(4):e59.
- [52] Dong NQ, Lin HX. Contribution of phenylpropanoid metabolism to plant development and plant-environment interactions. *J Integr Plant Biol.* 2021;63(1):180-209.
- [53] Spicher L, Glauser G, Kessler F. Lipid antioxidant and galactolipid remodeling under temperature stress in tomato plants. *Front Plant Sci.* 2016;7:167.
- [54] Collakova E, DellaPenna D. The role of homogentisate phytyldiphosphate transferase in tocopherol

- [55] Cheng W, Li W. Structural insights into ubiquinone biosynthesis in membranes. *Science*. 2014;343(6173):878-881.
- [56] Sadre R, Frentzen M, Saeed M, Hawkes T. Catalytic reactions of the homogentisate prenyl transferase involved in plastoquinone biosynthesis. *J Biol Chem*. 2010;285(24):18191-18198.
- [57] Yang W, Cahoon RE, Hunter SC, et al. Vitamin E biosynthesis: functionally interchangeable hydroxylases in plants and cyanobacteria. *Plant J*. 2011;66(6):1073-1085.
- [58] Razvi A, Bhatt DL, Kim SH. Structure-guided design of prenyltransferase modulators. *Nat Chem Biol*. 2020;16(8):841-850.
- [59] Goyal S, Mangal N. Computational identification of natural phytochemical inhibitors for plastid-localised enzymes. *J Biomol Struct Dyn*. 2022;40(11):5129-5144.
- [60] Jing X, Jin K. A gold mine for drug discovery: strategies to develop cyclopenta-fused xanthenes. *Med Res Rev*. 2020;40(4):1549-1588.
- [61] Moghaddam MG, Ahmad FBH, Samzadeh-Kermani A. Biological activity of betulinic acid: a review. *Pharmacol Pharm*. 2012;3(02):119-123.
- [62] Laskowski RA, Swindells MB. LigPlot+: multiple ligand-protein interaction diagrams. *J Chem Inf Model*. 2011;51(10):2778-2786.
- [63] Hood L, Friend SH. Predictive, personalized, preventive, participatory medicine. *Nat Rev Clin Oncol*. 2011;8(3):184-187.
- [64] Sweetlove LJ, Fernie AR. The role of dynamic enzyme assemblies and substrate channelling in metabolic regulation. *Nat Commun*. 2018;9(1):2136.
- [65] Tang G. Metabolic engineering for enhanced synthesis of secondary metabolites in *Arabidopsis*. *Curr Opin Biotechnol*. 2010;21(4):465-471.

- [66] Kong W, Zhong H, Geng G, et al. Systematic analysis of metabolic network in *Artemisia annua* for artemisinin yield enhancement. *Front Plant Sci.* 2020;11:549770.
- [67] Zhu C, Naqvi S, Breitenbach J, et al. Combinatorial genetic transformation generates a library of metabolic phenotypes for the carotenoid pathway in maize. *Proc Natl Acad Sci USA.* 2008;105(47):18232-18237.
- [68] Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 2015;43(7):e47.
- [69] Bolstad BM, Irizarry RA, Astrand M, Speed TP. A comparison of normalization methods for high density oligonucleotide array data. *Bioinformatics.* 2003;19(2):185-193.
- [70] Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis. *Proc Natl Acad Sci USA.* 2005;102(43):15545-15550.
- [71] Ashburner M, Ball CA, Blake JA, et al. Gene Ontology: tool for the unification of biology. *Nat Genet.* 2000;25(1):25-29.
- [72] Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* 2000;28(1):27-30.
- [73] Jeong H, Mason SP, Barabasi AL, Oltvai ZN. Lethality and centrality in protein networks. *Nature.* 2001;411(6833):41-42.
- [74] Berman HM, Westbrook J, Feng Z, et al. The Protein Data Bank. *Nucleic Acids Res.* 2000;28(1):235-242.
- [75] Morris GM, Huey R, Lindstrom W, et al. AutoDock4 and AutoDockTools4. *J Comput Chem.* 2009;30(16):2785-2791.
- [76] O'Boyle NM, Banck M, James CA, et al. Open Babel: an open chemical toolbox. *J Cheminform.* 2011;3(1):33.

- [77] Ropp PJ, Kaminsky JC, Yablonski S, Durrant JD. EasyDock Vina: a simplified docking platform. *J Cheminform.* 2019;11(1):14.
- [78] DeLano WL. The PyMOL molecular graphics system. San Carlos, CA: DeLano Scientific LLC; 2002.
- [79] Harker M, Bramley PM. Expression of prokaryotic 1-deoxy-D-xylulose-5-phosphatases in *Escherichia coli* increases carotenoid and ubiquinone biosynthesis. *FEBS Lett.* 1999;448(1):115-119.
- [80] Hollingsworth SA, Dror RO. Molecular dynamics simulation for all. *Neuron.* 2018;99(6):1129-1143.
- [81] Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery. *Adv Drug Deliv Rev.* 2001;46(1-3):3-26.

LIST OF PUBLICATION

1. Conference Paper:

Arjun Sharma and R. Mehta, "Computational Identification of VTE2-Targeting Phytochemicals for Enhanced Tocopherol Biosynthesis in Arabidopsis Using Molecular Docking," 2026 International Conference on Advances in Biotechnology and Computational Biology (ICABCB), New Delhi, India, 2026, IEEE, doi: 10.1109/ICABCB2026.11492301.

Keywords: Tocopherol; VTE2; Molecular Docking; Phytochemicals; EasyDock Vina; LigPlot+; PyMOL; Arabidopsis thaliana; Vitamin E Biosynthesis; Structure-Based Drug Discovery; Metabolic Engineering; Computational Plant Biology.

2. Poster:

Arjun Sharma, Rajesh Mehta, "Transcriptomic Identification of Hub Regulatory Genes in Tocopherol and Tocotrienol Biosynthesis Pathway in Arabidopsis thaliana"

Presented at: National Symposium on Plant Biotechnology and Metabolomics (NSPBM), IARI, New Delhi

PLAGIARISM VERIFICATION

For Title of the Thesis: **"Computational and Biotechnological Approaches for Enhanced Tocopherol and Tocotrienol Biosynthesis in Plants"**

Total Pages: **50**

Name of the Scholar: **Mihir Kumar (24/IBT/15)**

Supervisor: **Prof. Navneeta Bharadvaja**

Department: **Biotechnology**

This is to report that the above thesis was scanned for similarity detection. Process and outcome are given below:

Software used: **Turnitin** Similarity Index: _____ Total Word Count: _____