Phylogeny, phytochemical analysis and relative expression of curcumin synthase genes from *Curcuma* spp. of South India

Thesis submitted to the
University of Calicut in partial fulfillment of the
requirement for the award of the degree of

DOCTOR OF PHILOSOPHY IN BOTANY

By

SANTHOSHKUMAR R.

BIOTECHNOLOGY DIVISION
DEPARTMENT OF BOTANY
UNIVERSITY OF CALICUT
KERALA- 673 635
JUNE 2019

UNIVERSITY OF CALICUT

DEPARTMENT OF BOTANY
Calicut University, Kerala

Dr. A. Yusuf
Assistant Professor
Biotechnology Division

CERTIFICATE

No corrections were mentioned by the adjudicators and resubmitted the thesis entitled **"Phylogeny, phytochemical analysis and relative expression of curcumin synthase genes from *Curcuma* spp. of South India".** The content included in the Thesis and CD are same.

30/12/2019
Dr. A. Yusuf

C.U. Campus
30/12/2019

UNIVERSITY OF CALICUT
DEPARTMENT OF BOTANY

Dr. A. Yusuf
Assistant professor
Biotechnology Division

Calicut University Kerala,
India. 673635
Mob : 9497192730

CERTIFICATE

Certified that the Ph.D. thesis entitled "**Phylogeny, phytochemical analysis and relative expression of curcumin synthase genes from *Curcuma* spp. of South India**" is an authentic record of the original research work accomplished by **Mr. Santhoshkumar R.** under my supervision at the Biotechnology Division in the Department of Botany, University of Calicut and that no part thereof has been presented earlier for the award of any other degree or diploma. Also certified that the contents in the thesis are subjected to **Plagiarism Check** using the software **Urkund** and that no text or data is reproduced from other's work.

C.U. Campus **Dr. A. Yusuf**

DECLARATION

I, **Santhoshkumar R.** do hereby declare that this thesis entitled, "**Phylogeny, phytochemical analysis and relative expression of curcumin synthase genes from *Curcuma* spp. of South India**" is the summary of the research work carried out by me under the supervision of **Dr. A.Yusuf,** Assistant Professor, Biotechnology Division, Department of Botany, University of Calicut, in partial fulfillment of the requirement for the award of Degree of Doctor of Philosophy in Botany of the University of Calicut and also declare that no part of this thesis has been submitted by me for the award of any other degree or diploma.

Calicut University **Santhoshkumar R.**
Date:

ACKNOWLEDGEMENTS

In my journey towards this degree, I have found a teacher, a friend, an inspiration and a pillar of support in my Guide, ***Dr. A. Yusuf****, Asst. Professor, Biotechnology Division (Interuniversity Centre for Plant Biotechnology), Department of Botany, University of Calicut, for the expert guidance, inspiring advices, criticisms, constant support, patience, and attentive efforts extended throughout the period of this research work leading to the successful completion of the venture.*

I convey my sincere gratitude to ***Prof. K. V. Mohanan****, former Director, Interuniversity Centre for Plant Biotechnology and former Head of the Department of Botany, for his support, encouragements and providing facilities during my work.*

I extend my sincere thanks to ***Prof. Santhosh Nampy****, Head, Department of Botany, University of Calicut for providing facilities to undertake research.*

I am thankful to ***Prof. John E. Thoppil****,* ***Prof. K. M. Jayaram****, former Heads, Department of Botany, University of Calicut for their whole hearted help offered throughout the period and for creating conducive environment in the department for the successful completion of my work.*

My deep felt gratitude to ***Prof. P. Manimohan****,* ***Prof. M. Sabu*** *and* ***Dr. A. K. Pradeep****, Department of Botany, University of Calicut, for their help and suggestions to complete my work. I am deeply thankful to* ***Prof. Sailas Benjamin*** *(late), Enzyme technology Division, Department of Botany, University of Calicut for providing access to use lab facilities. I owe my heartiest gratitude to* ***Dr. V. V. Radhakrishnan****, professor, Department of Botany, University of Calicut for his encouraging words.*

I gratefully extend my sincere thanks to all the other Teaching and Nonteaching Staffs, Department of Botany for providing needful help during the research period.

I also express my thanks to ***Mr. P. M. Prakashan****, Librarian, Department of Botany for the help rendered.*

I would like to thank the University Grants Commission for the award of RGNF Fellowship.

I am sincerely thankful to ***Mrs. Divya,*** *Apsara Innovations and* ***Sandeep Sen****, ATREE, Bangalore, Karnataka for providing me timely suggestions and advices regarding Phylogenetic analysis, comparative modeling and protein structure prediction.*

I also express my thanks to ***Mr. K. Rajesh,*** *Bina Photostat, Villunniyal for their tireless support in preparing this manuscript.*

My special words of thanks to all my friends from Department of Botany, Hostel and other Department those who directly or indirectly helped me during this period.

I extend my sincere thanks to ***Mr. Dipin Thacharakkal****, Department of Chemistry, University of Calicut for his help in doing FTIR analysis.*

I extend my sincere thanks to my dearest friends ***Dr. Deepa P., Dr, Showmy K. S., Mrs. Aparna M. B., Dr. Lamiya K. M., Mrs. Ambili P., Mrs. Nanditha N. R., Mrs. Savitha, N., Mrs. Sreeja., Mr. Nithin K. N., Ms. Anju V. V., Mr. Lins Simon., Ms. Maya R., Mrs. Raseena M., Mr. Habeeb K. C., Mr. Jishin Prakash T.S., Ms. Abhirami K. S., Mrs. Neethu M., Ms. Mahija P., Mrs. Irfana Jasmine., Ms. Amurtha T., Ms. Janeefa J., Ms. Risha Ravi*** *for all their useful suggestions and for being there to listen when I needed an ear.*

I express my heartfelt gratitude to ***Dr. Rathy M. C., Dr. Sajith., Dr. Snisha S., Dr. Karthika Menon*** *Environment Science Division, for their love and special care throughout my research period. Special thanks to all members of Enzymology Division for their kind support. I also thank* ***Mr. Arun T Ram.,***

Hareesh E. S., Janeeshma E., Abdul Faisal P., Pravisya P. and *Smisha* Research Scholars Department of Botany for the help extended. I am also thankful to all members of Genetics and Plant Breeding Division; *Dr. Shintu, Dr. Soorya, Ms. Thushara, Mrs. Athira, Ms. Surekha, Ms. Neethu, Mrs. Krshina Priya, Ms. Litty* for their friendly support during my work. The help rendered by *Mrs. Bindu T. V.,* Supporting staff, Interuniversity Centre for Plant Biotechnology is remembered with thanks and gratitude.

I am sincerely thankful to *Dr. Anusha T. S., Megha K. B.*, Department of Biotechnology, University of Calicut for their help and co-operation.

My special thanks to *Dr. Jayasree,* (Rubber Research Institute of India) for their timely support for UFLC analysis and valuable suggestions.

I deeply acknowledge *Dr. Rajesh Kumar T.*, Assistant professor, NSS College, Manjeri for his immense support during the field collection and plant identification.

Words are not enough to express my love and gratitude to my family, the greatest gift god has given me. It gives me great pleasure to thank my *Aunty, Sister, Brother In-law* and my cute and dearest *nephews* for their love, tremendous patience encouragement and sincere prayers. I cannot complete my acknowledgement without expressing how thankful I am to my parents *Mr. Raman* and *Mrs. Gowri* for all the sacrifices they have made, all their support, unconditional love and great inspiration throughout my life.

Finally, I would like to thank everybody who was important to the successful realization of my thesis, as well as expressing my apology that I could not mention all personally one by one.

Above all these, I owe to the Almighty for giving me the strength and health for the successful completion of the work.

Santhoshkumar R.

CONTENTS

ABBREVIATIONS

AIC	:	Akaike Information Criterion
ATP	:	Adenosine triphosphate
ATP1	:	ATP synthase subunit alpha
atpA	:	ATP synthase alpha subunit
atpB	:	ATP synthase beta subunit
atpF-H	:	ATP synthase subunits CFO I- CFO III
BDMC	:	Bisdemethoxycurcumin
BEAST	:	Bayesian Evolutionary Analysis by Sampling Trees
BI	:	Baysesian Inference
BIC	:	Bayesian Information Criterion
BLAST	:	Basic local alignment search tool
BOLD	:	Barcode of Life Data System
bp	:	Base Pair
BPB	:	Bromophenol blue
BS	:	Bootstrap
CBOL	:	Consortium for the Barcode of life
CDS	:	Coding Sequence
COX1	:	Cytochrome oxidase subunit 1
C_T	:	Threshold cycle
CTAB	:	Cetyl trimethyl ammonium bromide
CUR	:	Curcumin
CURS1	:	Curcumin synthase1
CURS2	:	Curcumin synthase2
CURS3	:	Curcumin synthase3
ddNTPs	:	Di-deoxy nucleoside tri phosphate
DDW	:	Double distilled water
DEPC	:	Di-ethyl pyrocarbonate /diethyl dicarbonate
DMC	:	Demethoxycurcumin
DMSO	:	Di-Methyl sulfoxide
DNA	:	Deoxyribonucleic acid
DNA	:	Deoxyribonucleic acid

dsDNA	:	Double stranded DNA
EDTA	:	Ethylenediamine-tetra-acetic acid
EMBOSS	:	European Molecular Biology Open Software Suite
et al.	:	et alia (and others)
EtBr	:	Ethidium bromide
etc.	:	et cetera (and other things)
E-value	:	Expect Value
FA buffer	:	Formaldehyde acetate buffer
FA-EtBr	:	Formaldehyde Ethidium bromide
FASTA	:	FAST Alignment
FdxH	:	Ferredoxin
Fe-EDTA	:	Iron- ethylenediaminetetraacetic acid
Fig.	:	Figure
FNR	:	Ferredoxin: NADP+ oxidoreductase
g	:	Gram
GC-MS	:	Gas Chromatography –Mass Spectrometry
GMQE	:	Global quality estimation
GRAVY	:	Grand average of hydropathicity
h	:	Hour
H_2O	:	Water
HCl	:	Hydrochloric acid
HKGs	:	Housekeeping genes
HKY	:	Hasegawa, Kishino and Yano Model
HPLC	:	High Performance Liquied Chromatography
iBOL	:	The International Barcode of Life project
ITS1	:	Internal transcribed spacer 1
ITS2	:	Internal transcribed spacer 2
K2P	:	Kimura -2- Parameter
Kb	:	Kilobyte
l	:	Litre
LRT	:	Likelihood-Ratio Test
M	:	Mole/litre
matK	:	Maturase K gene
MCL	:	Maximum Composite Likelihood

MCMC	:	Markov Chain Monte Carlo
MEGA6.0	:	Molecular evolutionary genetics analysis 6.0
mg	:	Milligram
mg/ml	:	Milligram per mil
$MgCl_2$	:	Magnesium chloride
min	:	Minute (s)
ML	:	Maximum Likelihood
mm	:	Millimetre
MOPS	:	3-(N-morpholino) propane sulfonic acid
MQ	:	Milli-Q Water
mRNA	:	Messenger ribonucleic acid
MULTALIN	:	Multiple Alignment
MUSCLE	:	Multiple Sequence Comparison by Log-Expectation
MW	:	Molecular weight
N	:	Normal
n	:	number
NaCl	:	Sodium Chloride
NaOH	:	Sodium hydroxide
NCBI	:	National center for biotechnology information
ng	:	Nanogram
NGS	:	Next generation sequencing
NJ	:	Neighbour Joining
NMR	:	Nuclear Magnetic Resonance
nrITS	:	Nuclear ribosomal internal transcribed spacer
O.D	:	Optical density
ORF	:	Open reading frame
PAGE	:	Poly acrylamide gel electrophoresis
PAUP	:	Phylogenetic Analysis Using Parsimony
PCR	:	Polymerase Chain Reaction
PDB	:	Protein data bank
pH	:	Potentia hydrogenii (Latin: hydrogen power)
PHYLIP	:	Phylogeny Inference Package
pI	:	isoelectric point
PIR	:	Protein Information Resource

PRF	:	Protein Research Foundation
PROCHECK	:	Proteincheck
psbA -trnH	:	photosystem II protein D1 spacer - tRNA for histidine
psbI	:	Photosystem II reaction center protein I
psbK	:	Photosystem II reaction center protein K
PSI-BLAST	:	Position-Specific Iterative Basic Local Alignment Search Tool
PSIPRED	:	PSI-blast based secondary structure Prediction
PVPP	:	Polyvinylpolypyrrolidone
QMEAN	:	Qualitative Model Energy Analysis
qRT-PCR	:	quantitative real time PCR
rbcL	:	Large subunit of the ribulose-bisphosphate carboxylase/oxygenase gene
RNA	:	Ribonucleic acid
RNase	:	Ribonuclease
rpm	:	Rotations per minute
RSDs	:	Relative standard deviations
RT-PCR	:	Reverse transcriptase polymerase chain reaction
RUBISCO	:	Ribulose bisphosphate carboxylase/oxygenase
RuBP	:	Ribulose 1,5-bisphosphate
SDS	:	Sodium dodecyl sulphate
SDS- PAGE	:	Sodium dodecyl sulphate-Poly acrylamide gel electrophoresis
Sec	:	Second
SNP	:	Single Nucleotide Polymorphism
SOPMA	:	Self Optimized Prediction Method with Alignment
St. dev.	:	Standard Deviation
TAE	:	Tris acetate-EDTA
TE	:	Tris EDTA
Thr	:	Threonine
Tm	:	Melting temperature
Ts	:	Transition
Tv	:	Transversion
UFLC	:	Ultra fast liquid chromatography
UPGMA	:	Unweighted pair group method with arithmetic mean

v/v	:	Volume/volume
w/v	:	Weight per volume
WGS	:	Whole genome shotgun
B	:	Beta
γ	:	Gamma
µg	:	Microgram
µl	:	Microlitre
%	:	Percent
~	:	Approximately
°C	:	Degree Celsius

CHAPTER I
INTRODUCTION

The pantropical family Zingiberaceae (gingers) belongs to a monophyletic order Zingiberales distributed in the Indo-Malayan region consists of approximately 53 genus and more than 1377 species (Kress *et al.,* 2002; Kong *et al.,* 2010), having economic and ornamental potential (Uma and Muthukumar, 2014). Species level identification is complicated in this family because of similarities for the morphological characters between the species, phenotype plasticity of plants and seasonal flowering cycles of short duration. In the purpose of classification, conservation, cultivar development and phylogenetic studies the proper identification and characterization of the taxa are essential (Silva *et al.,* 2018). The family Zingiberaceae has been a taxonomically neglected group mainly due to the inaccessible nature of the wet evergreen forests in which they grow and their short flowering period coinciding with the monsoon season. The delicate nature of flowers, loss of color and formation of a gummy mass soon after collection makes the floral morphology based taxonomic studies difficult.

The currently accepted classification of the family Zingiberaceae is based on the vegetative and floral characters (Petersen, 1889; Schumann, 1904; Holttum, 1950; Smith, 1981; Larsen *et al.,* 1998) and the family is divided into four tribes (Hedychieae , Alpinieae, Zingibereae and Globbeae). *Alpinia, Amomum*, *Curcuma* and *Zingiber* are the major genera belonging to this family and other genera such as *Boesenbergia, Kaempferia, Elettaria, Elettariopsis, Etlingera* and *Hedychium* are lesser extent (Jatoi *et al.,* 2007). Recently molecular sequence data have been used to explore the phylogenetic relationships of this family and the genera of the family Zingiberaceae is divided into four subfamilies, the Siphonochiloideae (the

genus *Siphonochilus* only), the Tamijioideae (the single genus *Tamijia*), the Alpinioideae (most of the former Alpinieae), and the Zingiberoideae (including the former tribes Hedychieae, Zingibereae, and Globbeae) (Kress *et al.,* 2002). *Curcuma* is one of the important genera from this family which is difficult for plant hunters, herbarium technicians as well as taxonomists. The correct identification is necessary for proper utilization and conservation of Zingiberaceous crops, hence molecular markers are used to solve the problem of delimitation of some taxa. However, strong molecular support is lacking in the diversity studies of Zingiberaceous members in India.

1.1 *Curcuma* L.

Curcuma L. is one of the third-largest genera in the family Zingiberaceae comprising 120 species; widely distributed throughout the tropical and sub-tropical Asia and few species extending to Australia and Pacific region (Skornickova *et al.,* 2004). Twenty one species including one variety have been reported in South India (Sabu, 2006). These plants are used as spices, in medicines, culinary and ornamental plants. The genus has confused species discrimination; hence the proper taxonomic identification of the specimen is very important for meaningful bioprospecting as well as for effective conservation. Kress *et al.* (2002) proposed a new classification for Zingiberaceae and placed *Curcuma* under the tribe Zingibereae and this genus is taxonomically quite confusing. Morphology based classification of the genus has its own drawback as some of the key taxonomic traits remain controversial (Santapau, 1945). Molecular characterization is used as an addition to the traditional methods of germplasm characterization in many crop plants (Semagn *et al.,* 2006). *C. longa* syn. *C. domestica* is one of the most commonly utilized species, other species such as *C. aromatica, C. amada, C. kwangsiensis, C. zedoaria, C. caesia, C. malabarica, C. angustifolia, C. montana, C. decipiens, C. alismatifolia, C. zanthorrhiza,*

C. aeruginosa and *C. pseudomontana* are some of the economically important species of *Curcuma*.

1.2 DNA Barcode in Angiosperms

DNA-barcoding is a method used to study the taxonomic characterization and phylogenetic analysis of different organisms by using specific DNA sequences. The defined region within the DNA of an organism undergoes evolutionary divergence, which is conserved between and within the species. For the rapid identification of biological specimens at the species level the short DNA segments of nuclear and or cytoplasmic genome was utilized. The term "Taxon Barcodes" was proposed first time by Hebert *et al.* (2003) for species level identification. In the assessment of biodiversity hotspots and international trade of rare species and in conservation biology the plant DNA barcoding plays an essential role.

The DNA barcoding did not attain much attention during earlier years due to the inability to use cytochrome oxidase (*cox1*) as a universal barcode (Cho *et al*., 2004). The plant mitochondrial genome evolves much more slowly than animal mitochondrial genome. The mitochondrial *cox1* (cytochrome oxidase 1) gene is not suitable for the distinction of plant species (Rubinoff, 2006). Many candidate genes have been recommended as potential barcodes for plants, however, no accepted universal barcodes were agreed for plants, though; nine intergenic spacers were used as plant DNA barcodes. Based on this a barcode criteria were proposed for most of the variable regions (*ycf6-psbM*, *trnV-atpE*, *trnC-ycf6*, *psbM-trnD, trnK-rps16*, *trnH-psbA*, *rp136-rps8*, *atpB-rbcL*, and *trnL-F*) that meets these criteria. The nuclear and chloroplast genome of ITS (internal transcribed spacer) and *trnH-psbA* spacer are used as DNA barcodes for flowering plants recommended (Kress *et al.,* 2005). The conservation of genetic diversity DNA barcodes are used for plant identification, genetic resource tags (Jeanson *et al.,* 2011),

inventorization and analysis of biodiversity and also minibarcodes such as smaller fragments of DNA probably using for conservation biology (Meusnier *et al.,* 2008). However, the DNA barcoding is used in wide variety of applications in forest biosecurity and biosurveillance of habitats by identifying invasive species from the native species (Armstrong and Ball, 2005).

As concluded by CBOL plant-working group (PWG), plant DNA barcodes are multi-locus, with one "anchor" (i.e. universal across the plant kingdom) and "identifiers" to distinguish closely related species. Still, there is no consensus on the best candidate marker for plant DNA barcoding. The combination of barcoding genes containing *trnH-psbA* noncoding intergenic spacers (Kress *et al*., 2005) and *mat*K coding sequences of plastid/ chloroplast (Chase *et al*., 2007) were also used. The ability and use of barcoding in plants permit the use of highly degraded samples (permafrost samples) and other samples such as processed food and medicinal plants (Taberlet, 2007). Since, the chloroplast *trnL* (UAA) intron or a shorter fragment of this intron (the P6 loop, 10-143 bp) is recommended while they have relatively low resolution could be improved with highly conserved primers.

Three locus barcode combinations of *mat*K+*rpoB*+*rpoC1* and *mat*K+*rpoC1*+*trnH-psbA* was proposed by Chase *et al.* (2007) and the combinations of *rbcL, mat*K, *rpoC1* and *trnH-psbA* regions act as a universal plant DNA barcode for land plants (Hollingsworth *et al.*, 2009). The plant's core barcode combination of *mat*K and *rbcL* was recommended by CBOL Plant Working Group (2009). Two gene core barcodes for ferns were used and proposed (Ming *et al.*, 2011).

1.3 Phylogenetic analysis

Phylogenetic analysis is a method used to study the evolutionary relationship or history among the group of taxa such as species from their ancestors with the order of branching and divergence over a time span. The term “Phylogeny” is borrowed from a grouping of Greek words Phylon stand for “tribe” or “clan” or “race” and genesis means “origin” or “source”. The term also can apply for the genealogy of genes derived from a common ancestral gene. Nowadays, DNA or proteins based molecular phylogenetic analysis become more prominent due to the advantages such as; (1) popularity of DNA sequencing method (2) establishment of methods for phylogenetic tree construction using gene or protein sequences (3) the results of a phylogenetic analysis being treated in a quantitative pattern (4) availability of many programs for constructing a phylogenetic tree. The basic information carried out from the phylogenetic analysis contributes to basic biology (e.g. evolutionary history of the species, the evolution of genes and identification of sampled species) and applied biology (*e.g.* examination of the method of the infection of pathogenic microorganisms). Phylogenetic trees are usually constructed based on the evolutionary relationship among species (Nei and Kumar, 2000). In recent years, molecular phylogeny is improved with techniques and analyses of nucleic acid and protein sequencing. The rRNA typing used to direct reverse transcriptase mediated sequencing of the portion of both the small and large subunits of ribosome (Sarich and Wilson, 1967). In the total cellular RNAs the rRNA is the major RNA and it is relatively easy to obtain enough RNA for sequencing.

In the past 20 years one major achievement in the field of phylogenetic studies (in particular TOL) is related to molecular data (especially, gene sequence data). For example, 159 released data from GenBank [a member of the International Nucleotide Sequence Database Collaboration (INSDC)]

contains 75 billion nucleotides in 72 million sequences plus another 93 billion nucleotides just in the WGS (whole genome shotgun) sequence division stemming from 787 registered genome projects (Benson *et al.,* 2011). Numerous tools and algorithms have been used in phylogenetics to further improve the capabilities of large data processing. However, there are still many technical problems to be solved, including data collection and screening of DNA sequences, automatic construction of large trees (Supertree) (Meng *et al.,* 2014).

1.3.1 Reconstruction of Phylogenetic Trees

The evolutionary history cannot be directly observed, it must be inferred by comparative morphological, physiological or molecular analyses. During the late 1800s reconstructing a tree of life by resolving the evolutionary and genealogical relationships among organisms has been an important focus of evolutionary biology (Futuyma, 2005). In addition to deciphering the evolutionary relationships among taxa, phylogenetic trees are useful for understanding the adaptive evolution and the evolution of multigene families. Several statistical methods are available for reconstructing the phylogenetic trees based on the molecular data mainly classified into three main groups: distance methods, parsimony methods and maximum likelihood methods (Nei and Kumar, 2000). Different algorithms are available for processing nucleotide or amino acid sequences to generate phylogenetic trees that reflect the estimated evolutionary relationships among taxonomic units. In a phylogenetic tree, the inferred evolutionary relationships are displayed by tree topology and branch lengths. Therefore, an approximation of these two parameters should be as accurate as possible. While branch length calculations are based on relatively simple statistical models, determination of the true topology is challenging due to a large number of possible alternative topologies.

1.3.2 Selection of the candidates for phylogeny analysis (genes or proteins)

The selection of datasets is the main criteria for a successful tree constriction used for phylogenetic analysis (such as DNA sequences, RNA sequences of functional RNA, or amino acid sequences of proteins). The molecules for phylogenetic analysis are selected based on two major important criteria. The genes should be present in all the given species and the genes should have proper evolutionary rates, however, the evolutionary rates of the proteins vary among the organisms (Miyata *et al.,* 1980). For constructing the evolutionary tree the distant related species having the same molecule with lesser evolutionary rate should be chosen. This is because nucleotide or amino acid substitution of gene/ proteins reaches to saturation between distant species when the evolutionary rate is high. The nucleotide sequence of a gene is easy to attain saturation than the amino acid sequence of the coded protein. In such cases the housekeeping gens with less evolutionary rates are suitable. Since, the accuracy reduces the substitution rates leading to lesser mutations, however, the tissue specific genes having high evolutionary rate are suitable as gene markers.

1.4 Chloroplast genome

Study of the chloroplast genome has been an important source of information for the phylogenetic analysis for the past several decades. For molecular evolution and systematics studies in plants, chloroplast is the most effectively analyzed genetic components. Many genes have been considered as potential plant barcodes due to the presence of important characters like uniparental mode of inheritance, non-recombination and structural stability in both the genic and intergenic regions of the chloroplast. The chloroplast DNA of higher plants is circular having a size between 120–160kb. The chloroplast genome harbours proteins synthesized by rRNA genes, tRNA genes and other

genes. An earlier phylogenetic study acts as milestone for the selection of potentially useful genic and intergenic chloroplast loci as potential candidates for barcoding of land plants. Some potential candidate barcode loci proposed in plants are described below (Ledford, 2008).

1.4.1 *rbcL* (Ribulose Bisphosphate Carboxylase/oxygenase large subunit)

Plant working group such as Consortium for the Barcode of Life (CBOL) has suggested standard barcode genes like ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (*rbcL*) and or maturase K (*mat*K) for all land plants having high recoverability, sequence quality and levels of species discrimination (CBOL Plant Working Group, 2009). Most of the higher plant chloroplast genome has a single copy of *rbcL* gene. It contains only exons and polypeptide with ~475 amino acids. The two promoters are situated 400 bp apart in opposite orientations because divergent transcription occurs. RNA polymerase interferes with binding at the *rbcL* promoter and transcription of the *atpB* promoter so they do not function independently (Hanley-Bowdoin and Chua, 1987). The substitutions flanked by the *rbcL* of the parent species are non-synonymous (Iida *et al*., 2007). However, the replacement of single amino acid in *rbcL* changes in the CO_2 and O_2 specificity of ribulose 1, 5-bisphosphate carboxylase/oxygenase (Rubisco) (Galmes *et al.*, 2005) and it provides all the catalytically important residues of Rubisco, an important enzyme for both the reductive and oxidative photosynthetic carbon cycles. Based on the conserved nature, the *rbcL* sequence has great phylogenetic importance, even though the substitutions occur in sites of recognized functional importance (Kellogg and Juliano, 1997). This potential candidate gene has importance next to the *mat*K in the identification efficiency.

1.4.2 *mat*K (Maturase K)

The *mat*K gene is one of the protein coding regions of plastid which is widely used to decipher the coding regions of DNA. The *mat*K gene was first identified from tobacco (*Nicotiana tabacum*) plant during the sequencing *trnK* gene encoding the $tRNA^{Lys}$ (UUU) of chloroplast. This chloroplast gene is 1500 bp long located within the intron of the *trnK* and codes for maturase like protein (Wolfe *et al.,* 1992). This gene has high substitution rates within the same species and is a potential candidate to study plant systematics and evolution. The polymorphism of chloroplast DNA regions comprising *trnK*, *mat*K and intergenic *trnL-trnF* have been used to study the phylogenetic evolution of various plant species (Wolfe *et al.,* 1987). Studies were conducted using *mat*K gene to resolve intergeneric or interspecific relations among different flowering plants such as, Poaceae (Liang and Hilu, 1996), Solanaceae (Aoki and Ito, 2000), Orchidaceae (Salazar *et al.,* 2003) and in many other angiosperms (Hilu *et al.,* 2003).

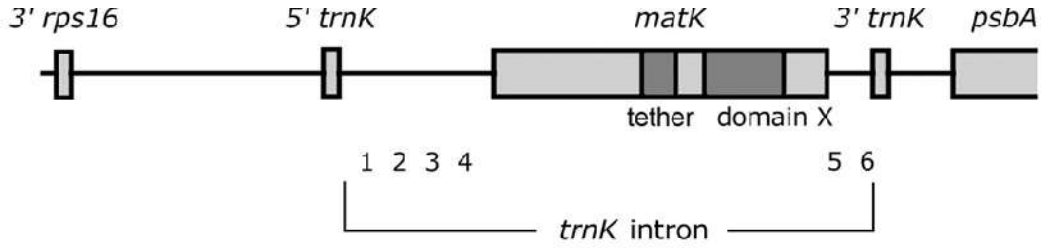


Fig. 1.1: The structure of the *mat*K gene (Duffy *et al.,* 2009).

1.4.3 *trnH-psbA* intergenic spacer (D1 Protein of Photosystem II)

In the population genetic studies of plants, the intergenic chloroplast DNA regions contain abundant information useful for predicting phylogenetic relationships (Provan *et al.,* 2001) and given more importance (Kress *et al.,* 2005). The unknown genes have non-coding regions, but much of the

distinction results from the spread of mutations that are not controlled by selection (Hamilton *et al.,* 2003). Intergenic spacer *psbA-trnH* occurs near to the *psbA* end spacer that has a highly conserved region and revealed large inversion at the end nearest to *trnH*, containing highly variable regions in the population genetic studies.

In the chloroplast of angiosperm the *trnH-psbA* intergenic spacer is one of the most variable genome segments. This non-coding gene is presented between the regions of the *psbA* gene and *trnH* gene, variation in gene size is observed among different species (Hollingsworth *et al.,* 2009; CBOL PWG, 2009). In angiosperms, gymnosperms, ferns, mosses and wild liverworts the *trnH–psbA* locus has been successfully amplified using PCR, however, failed in some taxa (Vijayan and Tsou, 2010).

Mononucleotide repeats are present in *trnH-psbA* genes so these are difficult for sequence accuracy and to estimate the levels of insertion (Wang *et al.,* 2008). However, in some groups, variation and differentiation among the closely related species was not observed using *trnH-psbA* gene (Sass *et al.,* 2007) and high intra-specific variation was observed in others (Edwards *et al.,* 2008).

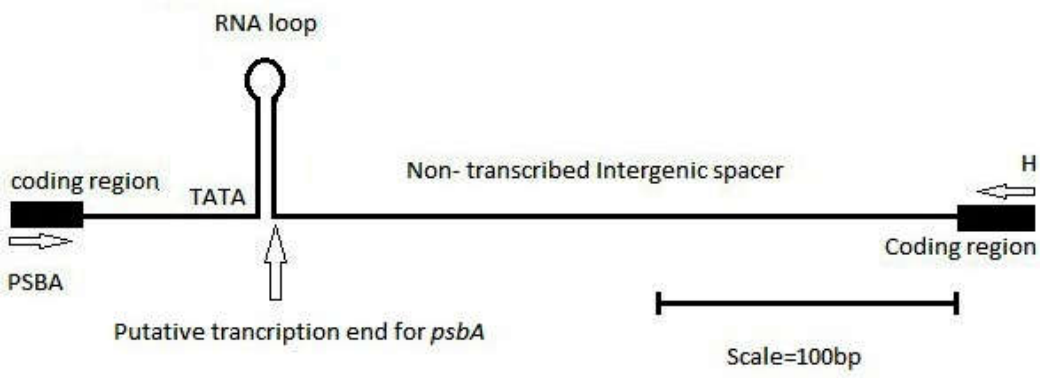


Fig. 1.2: The architecture of the chloroplast *psbA-trnH* (Storchova and Olson, 2007).

1.4.4 *atpF-atpH*

The *atpF-atpH* intergenic spacer is a potential plant barcode region recommended by second International Barcode of Life Conference (CBOL, 2009). These non-coding regions present between the *atpF* and *atpH* gene regions which encode the ATP synthase subunits CFO I and CFO III respectively (Drager and Hallick, 1993). The CBOL Plant Working Group pointed out that *atpF-atpH* has comparatively simple discriminatory power, intermediate sequence quality and universality and could be used as a plant DNA barcode. *atpF-atpH* is a noncoding spacer could serve as a universal DNA barcode marker for species-level identification confirmed by barcoding of duckweeds. The use of this noncoding region in the identification of new species is due to the termination of amplification, straightforward sequence alignment and rates of DNA variation (Wang *et al.,* 2010).

1.4.5 *psbK-psbI* intergenic spacer

The *psbK-psbI* intergenic spacer is non-coding regions encoded by two low molecular mass polypeptides (K and I respectively). These genes are present in between the genes *psbK* and *psbI* of the photosystem II (Meng *et al.,* 1991). These two conventional loci are present in the algae to land plants and even in parasitic plants (Vijayan and Tsou, 2010). However, due to the irregularity in nucleotide sequencing this has only been considered as a supplementary locus that can be used for better resolution if need arises (CBOL PWG, 2009).

1.4.6 *trnL* (UAA)–*trnF*(GAA) intergenic spacer

Taberlet *et al.,* (1991) was the first to use *trnL* intron gene for plant systematic studies. The non-coding locus of *trnL* (UAA)–*trnF* (GAA) contains *trnL*(UAA) gene, its intron and the intergenic region between *trnL* (UAA) and *trnF* (GAA). The highly conserved core structure of group I

introns are characterized and encoded by an active site that mediates self-splicing from the pre-tRNA. The *trnL* intron is the first group I intron described in chloroplast DNA that interrupt a tRNA gene (Bonnard *et al.,* 1984). Molecular phylogenetic studies of several taxa at various taxonomic levels were conducted using this gene sequences (Vijayan and Tsou, 2010). In the chloroplast DNA the *trnL* (UAA) – *trnF* (GAA) intergenic spacer is not the most variable non-coding region (Shaw *et al.,* 2005) but in this the non-coding region has an ease of amplification and good resolution power varying in different plant species.

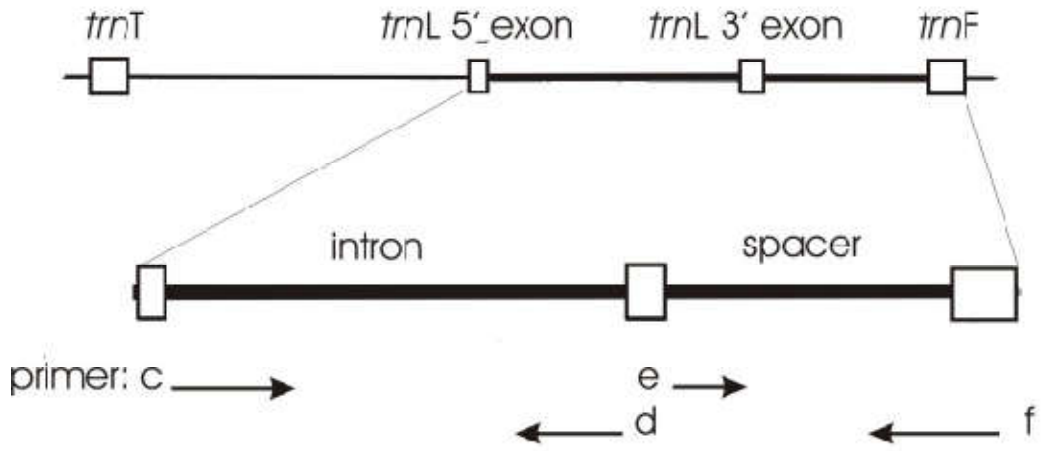


Fig. 1.3: The *trnL-F* region of chloroplast DNA. Primer annealing sites are marked (Adapted from Taberlet *et al.,* 1991)

1.5 Mitochondrial genes

Mitochondrial DNA (mtDNA) has a simple genetic structure containing a large number of foreign DNA and repeated sequences with a limited exposure to genetic recombination (Waugh, 2007). These mitochondrial DNA genes have suitable molecular markers for the study of closely related taxa (Hebert *et al.*, 2003). A mitochondrial gene encoding cytochrome c oxidase I (CO1-5′) and 648-bp region have been proposed as

the core of global bio-identification systems for eukaryotes (Hebert *et al.,* 2003; Waugh, 2007).

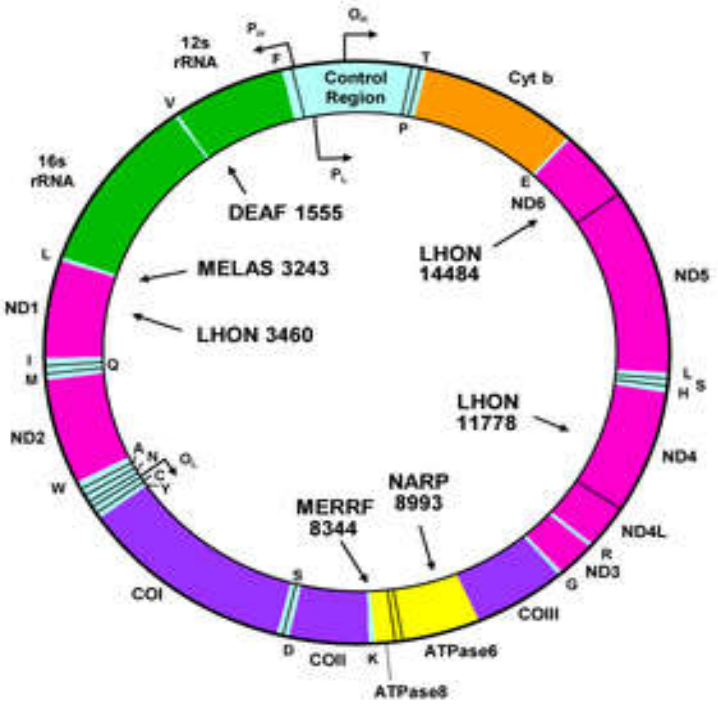


Fig. 1.4: Reprinted from Genetics Home: https://ghr.nlm.nih.gov/ mitochondrial-dna.pdf (2017)

1.5.1 Cytochrome c oxidase subunit 1 gene (*cox1)*

Cytochrome c oxidase subunit 1 gene (*cox1*) was the first gene used for DNA barcoding to assist the discrimination of closely related species (Hebert *et al.,* 2003). The mitochondrial protein cytochrome c oxidase is located in the inner membrane of mitochondria and it is a key electron transport chain enzyme. Therefore, it plays an important role in the aerobic metabolism of eukaryotic organisms and consists of several subunits. The catalytic activity of cytochrome c oxidase subunit 1 is encoded in the mitochondrial genome. The *cox1* gene was chosen for DNA barcoding due to its central role in major metabolic processes and therefore, its presence in almost all eukaryotes.

Moreover, compared to evolutionary rates of nuclear and chloroplast genome mitochondrial genes have higher evolutionary rates (*e.g.* small subunit ribosomal RNA) and should be a better candidate gene to discriminate between closely related taxa (Hebert *et al.,* 2003).

1.5.2 *ATP* 1 gene

The *ATP*ase complex present in the inner membrane of mitochondria consists of protein subunits synthesized in the mitochondrial matrix and cytoplasm (Dujon, 1981). As evident from the other sources, mitochondrial complexes are structurally and functionally conserved (Futai and Kanazawa, 1983). Analysis of genes encoding different subunits of the enzyme from sources as divergent as *Escherichia coli* and mammals indicates that the primary sequence of proteins constitute the catalytic core of the F_1- ATPase α, β and γ-subunits as well as subunits of the membrane sector F_o are homologous.

1.6 Nuclear ribosomal genes

In phylogenetic studies, the ribosomal RNA is considered as the best genomic region due to its ubiquitous nature and is composed of highly conserved as well as variable domains (Woese, 1987). Ribosomes consist of rRNA and proteins. The ribosome of all the organisms consists of two subunits (small and large) small ribosomal subunit (SSU) contains a single RNA species (the 18S rRNA in eukaryotes and the 16S rRNA in others). The large subunit (LSU) of bacteria and archaea contains two rRNA species (the 5S and 23S rRNAs) in most eukaryotes the large subunit contains three RNA species (the 5S, 5.8S and 25S/28S rRNAs). Moreover, rRNA genes are most important for the phylogenetic analysis of distantly related species because these genes are evolved more slowly than protein encoding genes (Moritz *et al.,* 1987).

1.6.1 ITS (Internal transcribed spacer)

The ribosomal DNA (rDNA) region of an internal transcribed spacer (ITS) is the only nuclear DNA that has been used in plant systematics (Vijayan and Tsou, 2010). The non-functional RNA sequence of ribosomal DNA (rDNA) is located between18S and 25S rRNA coding regions. Two regions are located in different regions, Internal transcribed spacer (ITS1) is located between 18S and 5.8S rRNA and ITS2 is located between 5.8S and 25S rRNA (Gerbi, 1985). Single to several thousand copies of tandem repeat units are present in the nuclear ribosomal RNA (rRNA) gene complex. For the phylogenetic analysis at interspecific and intraspecific levels, the whole ITS region was used (Kim *et al.,* 2015). However, ITS2 region is an equally efficient phylogenetic marker as the entire ITS region was suggested (Han *et al.,* 2013).

In the flowering plants (Angiosperms) length of ITS1 and ITS2 regions varies, and they range from 300 bp for ITS1 (187 to 298) and nearly 250 bp for ITS2 (187 to 252). Around 700 bp with the 5.8S rRNA region for total length of ITS region and which has a consistent length of 163 or 164 bp (Gonalez *et al.,* 1990). Based on the point mutation the regions have a different length.

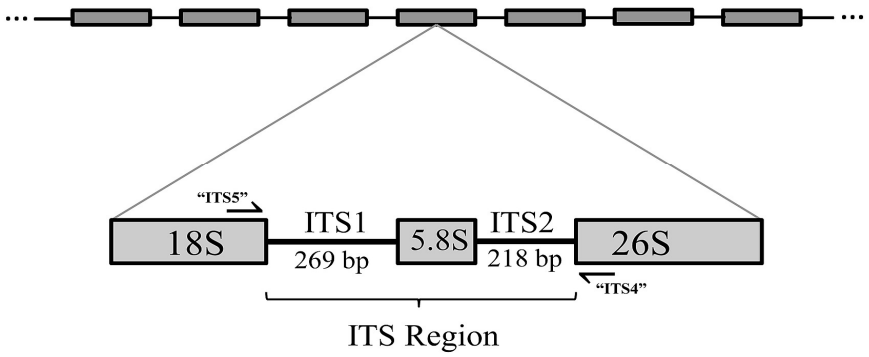


Fig. 1.5: The Structure of Intragenic Spacer Region ITS (Zhang *et al.,* 2015)

1.7 Phytochemistry

Phytochemistry deals with the extraction and analysis of chemical constituents from plants. Plants are better source of secondary metabolites with various biological activities. Phytochemical analysis strives to describe the structures of large number of secondary metabolites found in plants, the functions of these compounds in human and plant biology and the biosynthesis of these compounds. Such metabolic compounds (alkaloids, flavonoids, tannins, phenolics, amino acids, steroids, terpenoids, glycosides *etc.*) have higher antimicrobial and antioxidant properties (Westh, 2004). Secondary metabolites such as alkaloids, anthocyanins, flavonoids, quinines, lignins, steroids and terpenoids are used in commercial pharmaceutical and biochemical applications and are part of drugs, dyes, flavors, fragrances, and insecticides (Veerpoorte *et al.,* 2002).

1.7.1 Advanced techniques

1.7.1.1 Chromatography

Chromatography is a technique where the molecules are separated based on their size, shape and charge (Heftmann, 1992). During chromatography the solvent moves through a solid phase that acts as a sieving material. As the molecule proceeds further through the molecular sieve it gets separated. Paper and thin layer chromatography are the chromatographic techniques which readily provide qualitative information leading to generate the quantitative data (Ingle *et al.,* 2017).

1.7.1.2 GC-MS

Gas chromatography- mass spectrometry (GC-MS) is a technique comprising a gas chromatograph (GC) coupled to a mass spectrometer (MS) used to separate, identify and quantify the mixtures of biological samples

(Jayapriya and Shoba, 2015). It is often the analytical method of choice in toxicology, forensics, food science, and environmental research. It is mainly used for the determination of thermochemical constants, for the purification of compounds and qualitative and quantitative analysis of mixtures.

GC-MS technique is applied in various branches of biological research such as pharmaceutical industries for analytical research and development, quality control, quality assurance, production and pilot plants for the identification of active pharmaceutical ingredients (API), bulk drugs and formulations *etc*. It is used in the process and method development, identification of impurities in drug and pharmaceutical ingredients. It is an important part of research connected with medicinal chemistry (synthesis and characterization of compounds), pharmaceutical analysis (stability testing, impurity profiling), pharmacognosy, pharmaceutical process control, pharmaceutical biotechnology *etc.* (CDER, 2000). Gas chromatography mass spectrometry (GC-MS) has become conventional method for secondary metabolite profiling (Kell *et al.,* 2005).

1.7.1.3 High performance liquid chromatography (HPLC)

HPLC is a conventional technique used for the identification of natural products and it is a chromatographic technique that can be used to isolate single ingredients from a mixture of compounds. HPLC is used to study the phytochemical constitution to identify, quantify and purify the individual components in a sample (Piana *et al.,* 2012). Now-a-days this technique is applied to various analytical techniques like fingerprinting to analyze the quality of herbal extracts (Fan *et al.,* 2006). The resolving power of HPLC is ideally suited for the processing of such multi component samples in both analytical and preparative scale of plant extracts (Martin and Guiochon, 2005). Characterization and quantification of secondary metabolites in plant

extracts such as phenol compounds, steroids, flavonoids, alkaloids using HPLC were studied (Reis Ede *et al.,* 2014).

1.7.1.4 Ultra fast liquid chromatography (UFLC)

UFLC is ten times faster and three times better separation compared to other conventional liquid chromatography and its derivative Ultra-Performance Liquid Chromatography (UPLC). Therefore, state of being important UFLC provides excellent speed and low system back pressure. It is based on high analysis precision and reliability (Gangadasu *et al.,* 2015). Secondary metabolites are mostly methanol soluble and this solvent creates a back pressure exceeding 4000 psi in a conventional HPLC pump (Ramaswamy *et al.,* 2014) and UFLC withstand pressure up to 6800 psi (Gannu *et al.,* 2009).

1.7.1.5 Fourier-transform infrared spectroscopy (FTIR)

Fourier-transform infrared spectroscopy is a precious tool for the identification of functional ingredients present in the plant extract. It helps in the identification and structure prediction of the molecule. Different sample preparation methods are used for FTIR such as liquid samples, the easiest is to place one drop of sample between two plates of sodium chloride and the drop forms a thin film between the plates. Solid samples can be crushed with potassium bromide (KBr) and condensed into a thin pellet which can be analyzed or solid samples can be dissolved in methylene chloride and a few drops of the solution is placed on to a single High Attenuated Total Reflectance (HATR) plates and spectra can be recorded as percentage transmittance. The peaks at specific wave length were assigned deciphering the bonds and functional group as per the reference given in Varian FTIR instrument manual (Hazra *et al.,* 2007).

1.7.1.6 Nuclear Magnetic Resonance Spectroscopy (NMR)

Nuclear Magnetic Resonance Spectroscopy analyzes the physical, chemical and biological properties of a substance. One dimensional technique is normally used for the analysis but the complicated structure of the molecules could be analysed through two dimensional NMR techniques. Molecular structures of solids can be determined by using solid state NMR spectroscopy. The carbon structure present in the compounds are identified using radiolabelled C NMR and hydrogen atoms are identified in a compound using H-NMR (Cosa *et al.,* 2006).

1.7.1.7 Mass spectrometry (MS)

Mass spectrometry is a highly sensitive analytical technique for the identification of unknown compounds (Ingle *et al.,* 2017). It is working on ionizing chemical compounds to generate charged molecules or molecule fragments and measuring their mass to charge ratios. The ions are separated based on their mass-to-charge ratio in an analyzer by the electromagnetic fields and the ions are detected, usually by a quantitative method. The most recent applications of analytical chemistry are oriented towards biochemical activites such as proteome, metabolome, high throughput in drug discovery and metabolism. Other analytical applications are routinely applied in pollution control, forensic science, natural products or process monitoring and other applications including atomic physics, reaction physics, reaction kinetics, geochronology, inorganic chemical analysis, ion–molecule reactions (Hoffmann and Stroobant, 2007).

1.7.2 Conventional extraction

Various classical techniques are used to extract the bioactive compounds from plant parts and different solvents are used as extraction agents. The bioactive compounds are obtained from plants using conventional

techniques such as (1) soxhlet extraction, (2) maceration and (3) hydrodistillation (Azmir *et al.,* 2013). A soxhlet extractor is a conventional method first proposed by German chemist Franz Ritter Von Soxhlet (1879). Valuable bioactive compounds are obtained from different natural sources by using Soxhlet extraction. Maceration is a conventional technique used in homemade tonic preparation from a long time. It is a popular and inexpensive way of extracting the essential oils and bioactive compounds from different samples. For small scale extraction, maceration consists of several steps such as grinding of plant materials, maceration process, filtration and preparing the crude extract. Intermittent shaking during maceration facilitates the extraction by two ways; (a) increased diffusion (b) removal of the concentrated solution from the sample (Azmir *et al.,* 2013).

1.8 Curcumin

Curcumin is a hydrophobic polyphenolic, orange-yellow crystalline powder with molecular weight of 368.37g/mol and a melting point of 183°C almost insoluble in water but it is soluble in methanol, ethanol, DCM or DMSO (Tonnesen and Karlsen, 1985). Lampe and Milobedeska, (1913) first described the structure of curcumin ($C_{21}H_{20}O_6$) and shown to be diferuloylmethane. The main coloring agent of curcumin present in turmeric rhizome commonly known for its tradicinal medicinal values. Curcuminoids are a group of phenolic substances such as curcumin, demethoxycurcumin and bisdemethoxycurcumin which account for 60%-80%, 15%-30% and 2%-6% respectively (Rohman, 2012).

Curcuminoids have strong anti-oxidant (Choi, 2009), anti-inflammatory (Lantz *et al.,* 2005), antibacterial (Park *et al.,* 2005) and anti-carcinogenic activities (Das and Vinayak, 2012). In ancient medicine curcuminoids are used to treat dental diseases, digestive disorders such as dyspepsia and acidity, indigestion, flatulence, ulcers, as well to alleviate the

hallucinatory effects of *hashish* and other psychotropic drugs. Natural yellow colored, curcumin is currently used for making perfumes and as approved food preservative to flavor a variety of curries (Shishodia *et al.,* 2005).

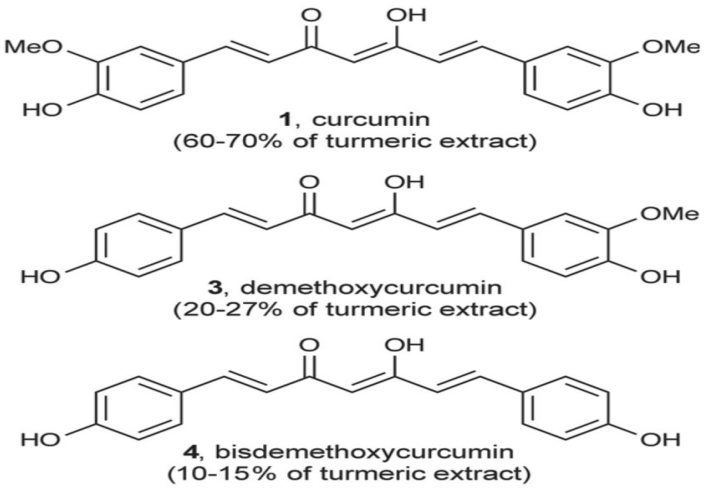


Fig. 1.6: The structures of different curcuminoids from *Curcuma* species (Nelson *et al.,* 2017)

1.9 Molecular characterization methods

1.9.1 DNA and RNA isolation and purification

The standardization of suitable extraction procedures for biomolecules such as DNA, RNA and protein is the most vital method used in molecular biology. The isolation procedure of DNA and RNA includes guanidium-thiocyanate phenol-chloroform method (Chomczynski and Sacchi, 1987), CTAB method (Carra *et al.,* 2006), purification of poly $(A)^+$ mRNA by Oligo (dT) cellulose chromatography (Konieczny and Legocki, 1981), EtBr-CsCl density gradient centrifugation (El-Gewely and Helling, 1980) *etc*. The unstable molecule of RNA has a very short half-life, it is extracted from the cell or tissues (Brooks, 1998). The RNA isolation is the first step to study of gene expression from the extracts of sufficient quantity and high quality total

RNA from the plant material is most important. There are different types of RNA occurring in nature such as ribosomal RNA (rRNA) (80%–90%), messenger RNA (mRNA) (2.5%-5%) and transfer RNA (tRNA). RNA can be easily degraded, hence special care and precautions must be taken during RNA isolation (Buckingham, 2011). RNA molecules are prone to degradation by ozone in the air whether the RNA sample is liquid or solid (Cataldo, 2005). In the field of plant molecular biology RNA quality and quantity plays a vital role for many studies including northern blotting, microarray hybridization, qRT-PCR and cDNA library construction.

Presence of secondary metabolites like phenolic substances, alkaloid, flavonoids and large quantities of polysaccharides causes major problem in isolating pure DNA and RNA. There are several methods employed in the isolation of pure DNA and RNA. Plants have large quantity of carbohydrates hence, phenol extraction is preferred are CTAB (Cetyl Trimethyl Ammonium Bromide or Cetrimonium bromide) method for isolation. CTAB is a cationic detergent used to isolate nucleic acids from the samples containing a high quantity of protein and polysaccharide. The CTAB method is used for the isolation of nucleic acids from bacterial cells, subsequently used for isolating DNA and RNA from plant tissues containing a high quantity phenolic substances and polysaccharides (Dahm, 2005). After the sequential extraction using chloroform and phenol helped in removing the CTAB-Polysaccharides-Protein complex, addition of ethanol helps to recover RNA and precipitated from the supernatant.

1.9.2 Quantification of nuclic acids

1.9.2.1 Agarose gel electrophoresis

Agarose gel electrophoresis is a generally used method for separating proteins, DNA or RNA (Kryndushkin *et al.,* 2003). DNA samples isolated

from different sources are mixed with ethidium bromide and marker dye are loaded in a well of agarose gel which kept in position in the electrophoresis chamber containing running buffer under electric current (typically 100 V for 30 min). Ethidium bromide (EtBr) is the frequently used as the intercalating dye for nucleic acid visualization. In the earlier protocol that describes the usage of ethidium bromide (2,7-diamino-10-ethyl-9- phenylphenan-thridiniumbromide-) for staining DNA and RNA in agarose gels dates as far back as 1970s (Sharp *et al.,* 1973). The movement of molecules are determined by the molecular weight where small weight molecules migrate faster than larger ones (Sambrook and Russel, 2001). Extension of the technique includes excising of the DNA "band" from a stained gel under a UV transilluminator (Sharp *et al.,* 1973).

1.9.2.2 Quantitative assessment of DNA

The concentration of DNA can be precisely measured by using UV absorbance spectrophotometer. The absorbed quantity of UV radiations in a solution of DNA is directly proportional to the quantity of DNA in the sample. Generally, the absorbance is measured at 260 nm and one OD corresponds to 50 μg of double strand DNA. The UV absorbance 260 nm and 280 nm also used to check the purity of DNA with the pure sample shows the ratio of absorbance at A260/A280 1.8. The ratio below 1.8 indicates that the samples are contaminated either with proteins or with phenol. A known concentration of standard DNA like salmon sperm DNA can be prepared and the readings at 260 nm can be measured using a UV spectrophotometer. The increase in concentration is directly proportional with the OD value (Psifidi, 2010).

1.9.3 Reverse transcriptases and cDNA synthesis

In polymerase chain reaction (PCR) a single copy of DNA template makes multiple copies with the help of DNA polymerase enzyme. The enzyme reverse transcriptase was discovered in 1970 (Temin and Mizutami, 1970). PCR is very specific and can be used to amplify or copy a specific DNA target from the whole DNA in three steps: denaturation, annealing and extension. Reverse transcription combined with polymerase chain reaction (RT-PCR) is used for the synthesis of cDNA using mRNA as the template. Reverse transcription PCR is a sensitive PCR technique used to qualitatively analyse the mRNA through the synthesis of cDNA by exponentially amplifying the target DNA sequences using traditional PCR. The components for the reaction include, RNA dependent DNA polymerase (reverse transcriptase), oligo-dT primer, dNTPs and Mg^{2+}. The capability of reverse transcriptase to synthesize cDNA from mRNA has been utilized in various laboratories. The oligo-dT primer or oligo hexamers are short repetitive sequences at the 3' poly (A+) tail of mRNA, that acts as a starting point for reverse transcriptases (Sangha *et al.,* 2010). The commonly used reverse transcriptases are Avian Myeloblastosis Virus (AMV) reverse transcriptase and Moloney Murine Leukemia Virus (M-MuLV, MMLV) reverse transcriptase (Kawasaki *et al*., 1988).

1.9.4 Gene expression

The central dogma of molecular biology describes the method by which information is carried forward from genes and used to create proteins. DNA is transcribed to RNA and the genetic information is translated to proteins. This process is called as gene expression and all life forms use it to create the building blocks of life from genetic information (Crick, 1970). There are numerous methodologies and strategies that can be used to study gene expression, the most important are complementary DNA (cDNA)

libraries; cDNA-amplified fragment length polymorphism (cDNA-AFLP) analysis; microarray analysis; suppressive subtractive hybridization (SSH); differential display (DD); RNA fingerprinting by arbitrary primed PCR (RAP-PCR); expressed sequence tags (EST) sequencing; serial analysis of gene expression (SAGE); representational difference analysis (RDA); and RNA sequencing (RNA-Seq) (Casassola *et al*., 2013).

1.9.5 Real time PCR

Real-time PCR (or qRT-PCR-quantitative real time PCR) allows the sensitive, specific and reproducible quantification of nucleic acids. It quantifies the relative expression of mRNA using very small quantity of the sample. Even though the quantitative PCR (qPCR) is most accurate method for quantifying gene expression it is mostly dependent on the use of a consistent, steadily expressed reference gene (Stanton *et al*., 2017). From the different samples real time PCR helps to determine the absolute transcript level and providing valuable quantitative information on gene expression (Nolan *et al.,* 2006).

Quantitative PCR quantifies and detect the quantity of DNA formed after each cycle using fluorescent dyes. During the amplification of the DNA template, the fluorescence exponentially increases. In SYBR green method the determination and quantification of the PCR products in real-time PCR reactions are highly sensitive (Nygard *et al.,* 2007). SYBR green chemistry is a DNA binding property of the dye or strand specific probe that are simplest and low cost method compared to TaqMan, Molecular Beacon (a hairpin probe) or hybridization probes (also called FRET) and it is easy to optimize regularly because SYBR green method is economical and beneficial when compared to strand specific probes (Fraga *et al.,* 2008). The reference genes or housekeeping genes are important for real time PCR reaction, should be stable as it is present in all types of tissues and cells either not or

constitutively expressed under all experimental conditions. There are number of specific housekeeping genes are used in expression studies (Vandesompele *et al.,* 2002), ubiquitin (UBQ), ribosomal protein 19 (RPL-19) and 18S ribosomal RNA (18S rRNA), gluteraldehyde-3-phosphate dehydrogenase (GAPDH), β-actin, β-tubulin and phosphoglycerate kinase (PGK) are commonly used references genes in real-time PCR. qRT-PCR is considered to be the most powerful, sensitive, quantitative method for mRNA quantification. Real-time PCR has much application in research, to analyze the expression of single or multiple genes, expression patterns for identifying infections and diseases, plant and animal pathogens *etc*.

1.9.6 Genome sequencing

Gene sequencing methods are used to determine the nucleotide sequence of a particular DNA to identify the organism and relationship of species by phylogenetic analysis as well as confirmation of proteins with the help of bioinformatics tools. It is also used in analyzing a newly cloned cDNA fragment to confirm the identity of a clone or mutant to check the fidelity of a newly created mutant, ligation junction or PCR product. In 1977 first two widely-known DNA sequencing methods were invented, the chemical cleavage sequencing method was developed by Maxam and Gilbert (Maxam *et al.,* 1977) and Sanger and collaborators developed the second method of chain terminator sequencing, dideoxy sequencing or Sanger sequencing, this method utilized ddNTPs (dideoxynucleoside triphosphate), a nucleoside analog which is substituted for deoxynucleotides at a random position during template directed copying of a DNA strand by a DNA polymerase (Sanger *et al.,* 1977).

Next-generation sequencing (NGS) is known as deep or massively parallel sequencing, developed using the technological advancement in DNA sequencing which can generate millions of sequence reads per run. The first

NGS machine 454 GS 20, was developed in 2005 capable of producing 2,00,000 constituting 100bp reads in a single run. Currently, Illumina Hiseq X Ten sequencer is used for sequencing technology it is capable of sequencing 18,000 genomes (Boonham *et al*., 2016). NGS is used in wide-range applications including (i) whole genome sequencing, (ii) pathogen discovery, (iii) metagenomic/microbiome analyses, (iv) transcriptome profiling and (vi) infectious disease diagnosis.

1.9.7 Comparative protein modeling

Based on one or more templates, a three dimensional model predition for an unknown protein (the target) can be constructed by comparative or homology protein structure modeling (Fiser and Sali, 2003). It helps to explicate the structure of an unknown protein where no crystal structure is available and leads to the identification of function if the 3-dimensional protein structure is available. 3D structures are developed using experimentally determined protein template that shows better homology with the target protein. Protein structure prediction helps to provide the biological function and mechanism of action of an unknown protein (Khan *et al*., 2016). The homology models are commonly used in virtual screening, designing site-directed mutagenesis experiments or in rationalizing the effects of sequence variations (Bordoli *et al.,* 2008). However, stable, reliable and accurate systems for automated homology modeling are required which are easy to use for both non-specialists and experts in structural bioinformatics. In protein modeling there are numerous automated computer programs and web servers used in the modeling process. Swiss-Model server (http://www.expasy./swissmodel.org) is the first web server for automated comparative modeling, followed by CPHModels (http://www .cbs.dtu.dk/services/CPHmodels/), SDSC1 (http://cl.sdsc.edu/hm), FAMS (http://physchem.pharm.kitasato u.ac.jp/FAMS/fams.html) and ModWeb

(http://guitar.rockefeller.edu/modweb). Similarly, MODLOOP, MOULDER, MODBASE *etc* are the various web servers used for homology modeling of proteins (Eswar *et al.,* 2003).

SWISS-MODEL is one of the most widely-used and longest standing servers used to build protein structure which uses its own comparative modeling functions. This server performs fully automated modeling, also allow the users to select templates for modeling. The server also help to find out complementarity to model ligands into structures also incorporate quaternary structure into modeling, if this is known from the template (Biasini *et al.,* 2014).

Objectives

Major objectives of the present study were:

1 Collection, identification and establishment of field germplasm for different species of *Curcuma.*

2 Standardization of DNA and RNA extraction procedures from leaves and rhizomes of all the species of *Curcuma.*

3 Standardization of PCR conditions for *mat*K, *rbcL, psbK-psbI, trnlC-trnlF, psbA-trnH, atpF-atpH, atp1* and *cox1* and ITS, further analysis of sequences using various softweares.

4 Phytochemical analysis of the species using GC-MS and UFLC

5 Cloning and relative expression studies of CURS1, CURS2 and CURS3 genes from the rhizomes of *Curcuma* species during different growth periods.

6 Comparative modeling of CURS1, CURS2 and CURS3 proteins using standard bioinformatic tools.

CHAPTER 2
REVIEW OF LITERATURE

2.1 Family Zingiberaceae

The members of Zingiberaceae are perennial rhizomatous herbs characterized by their type of inflorescence, colour and position (lateral or terminal), often covered by sheathing leaf bases. The word ‘gingers' refer to the members of the family Zingiberaceae, whereas "spiral gingers" to the members of the family Costaceae (Sabu, 2006). Fifty three genera and 1300 species are reported and the genera such as *Aframomum, Amomum, Curcuma, Elettaria, Kaempferia*, and *Zingiber* are the most important members of this family. The family Zingiberaceae is distributed mainly in the tropics, subtropics and center of distribution in the Indo-Malayan region, but extending through tropical Africa to Central and South America (Kress *et al*., 2002). Schumann (1904) reported 38 genera and 800 species, whereas Willis (1948) reported 45 genera and 800 species and Kong *et al.,* (2010) reported 53 genera with 1377 species.

Traditionally plant taxonomy is mainly dependent on the relative external morphological characters. Taxonomic confusion is reported to be prevailing in the family Zingiberaceae which is often difficult to discriminate based on the conventional taxonomic tools. A few studies on the morphological, anatomical and biochemical characterization of *Curcuma* species have been attempted earlier (Zhou *et al.,* 2007).

2.2 Significance of the Genus *Curcuma*

Curcuma is one of the important genera belonging to the family Zingiberaceae within the subfamily Zingiberoideae and tribe Zingiberaceae

(Kress *et al.,* 2002). This genus comprises of 120 species mainly distributed in tropical and subtropical Asia (Skornickova, 2007). Linnaeus first established this genus in 'Species Plantarum' in 1753. Baker (1890, 1892) classified the genus and reported 29 species from India. Roxburgh (1820) divided *Curcuma* into two subgenera based on the position of spikes (lateral/central). Horaninow (1862) distinguished three sections namely Exantha (Spikes always lateral), Mesantha (Spikes invariably terminal) and Amphiantha (Spikes both terminal and lateral) in this genus. Baker (1898) further subdivided the genus into three sections, Exantha, Mesantha, Hitcheniopsis and confirmed twenty-seven species of *Curcuma* and reported in the 'Flora of British India'. Fourteen species were included in the section Exantha with turmeric (*C. longa* L.) and also other economically important species such as *C. angustifolia* Roxb., *C. aromatica* Salisb., *C. zedoaria* Rosc. *etc*. (Velayudhan *et al.,* 1996).

A wide taxonomic study of the genus was carried out by Valeton (1918) and this genus was classified into two subgenera namely *Paracurcuma* and *Eucurcuma* based on the presence and absence of anther spur. Two species viz, *C. ecalcarata* and *C. aurantiaca* in the genus *Paracurcuma* were included based on the presence or absence of anther spur. Based on the presence or absence of tubers and stolons, *Eucurcuma* is further divided into three main sections namely tuberosa, nontuberosa and stolonifera. Each sections are again divided into subsections based on the aerial characters and underground features. Indian *Curcuma* species were divided into two subgenera based on the presence or absence of anther spurs (Velayudhan *et al.,* 1996).

Forty species of *Curcuma* was reported from India by Velayudhan *et al.* (1996) and many of them were recognized as duplicates and are now being recognized as synonyms as in the case of species *C. zedoaria* synonym of *C.*

zanthorrhiza (Sasikumar, 2005; Skornickova and Sabu, 2005). The existence of duplicate entities in the *Curcuma* species from China is also reported. *C. albicoma* and *C. chuanyujin* are synonyms of *C. sichuanensis* and *C. kwangsiensis*, respectively. *Curcuma wenyujin* is considered as a synonym of *C. aromatica*, *C. sichuanensis* as a synonym of *C. albicoma, C. sichuanensis* and *C. chuanyujin* are treated as varieties of *C. longa* and *C. phaeocaulis* has been misidentified in the past as *C. zedoaria, C. caesia* and *C. aeruginosa* in China (Nian and Wu, 1999). Twenty nine species have been reported from India, in almost all the states and many are cultivated and naturalized and distributed in South West India, North East India and Andaman and Nicobar Islands (Sabu, 2006). Twenty species and one variety have been reported from South India (Sabu, 2006).

2.3 Molecular characterization

2.3.1 Molecular Markers

Molecular markers are DNA sequences used to detect slight variations in gene sequences and their inheritance can be easily monitored in different individuals. The expression of morphological markers and biochemical markers including proteins markers are expressed through specific genes, molecular markers can be assessed based on DNA or on genetic characters (Choudhary *et al.,* 2008). Molecular markers have several advantages such as easy detection, rich throughout any genome including developed cultivars, independent of environmental conditions and can be distinguished practically during the developmental stages of the plant. Molecular markers are used in several areas *viz*; germplasm characterization, genetic diagnostics, characterization of transformants, the study of genome organization and phylogenetic analysis (Gupta and Roy, 2002).

Different DNA based molecular markers are utilized to evaluate DNA polymorphism, such as ISSR-inter simple sequence repeats (Zeitkiewicz *et al.,* 1994), RAPD-random amplified polymorphic DNA (Williams *et al.,* 1990), AFLP- amplified fragment length polymorphism (Vos *et al.,* 1995), RFLP-restriction fragment length polymorphism and VNTR-variable number of tandem repeats (Nakamura *et al.,* 1988) and SNPs-single nucleotide polymorphism (Sobrino *et al.,* 2005). It is necessary to know that different markers have different properties and will expose different characteristics in genetic variability (Karp and Edwards, 1995). Molecular markers are classified based on its capacity to decipher the DNA polymorphism and are commonly categorized as hybridization and polymerase chain reaction (PCR) based molecular markers. Some of the most important molecular markers used in the family Zingiberaceae for genetic diversity analysis are listed below.

2.3.2 RAPD (Random Amplified Polymorphic DNA)

In 1980 the Polymerase Chain Reaction (PCR) technology was used for genome identification (Saiki *et al.,* 1988). By using the advantages of PCR technology, enormous number of molecular markers were developed for rapid analysis of genetic polymorphisms and other related studies. RAPDs (Williams *et al.,* 1990) and AFLPs (Vos *et al.,* 1995) are widely used as PCR markers. However, each technique has its own advantages and disadvantages. RAPD markers are arbitrary sequence used as a primer based technique known to be rapid and easy to analyze the variability of species but this lacks reproducibility (Virk *et al.,* 2000).

Rajeshkumar *et al.* (2016) studied the genetic viability of 19 *Curcuma* species using RAPD. Twenty RAPD primers and generated total of 2226 scorable bands, identified in the 19 species of *Curcuma*, out of which 1025

were polymorphic. The polymorphism percentage ranged within the species from 56.7% to 36.5%.

Jan *et al.* (2011) reported molecular genetic fingerprints of indigenous turmeric (*Curcuma longa* L.) genotypes using RAPD marker to elucidate the genetic assortment among the genotypes. Ten random primers generated 95 fragments out of which 92 fragments were polymorphic with 96.84% of polymorphism.

Angel *et al.* (2008) assessed the genetic variability of *Curcuma* species using Random Amplified Polymorphic DNA (RAPD) technique. Twenty random primers were used to develop total 274 bands, of which 264 were polymorphic. The dendrogram was developed for all the species. *C. aromatica, C. leucorrhiza* and *C. brog* formed a cluster and *C. longa* and *C. zedoaria* formed a subgroup. *C. haritha* was genetically distinct from all the other *Curcuma* species.

Corcolon *et al.* (2015) studied the genotypic characterization of *Curcuma longa* L. using RAPD method. Ten random primers were used to generate 209 amplicons. Polymorphic information content was almost 0.5 for the six primers (OPD08, OPB07, OPA11, OPA12, OPC05 and OPN16) and except for OPA12 the multiplex ratio and marker index were also relatively high.

Donipati *et al.* (2015) examined the relationships among the six medicinal species of *Curcuma* using RAPD markers. Two polymorphic primers (OPC-4 and OPC-7) were used to evaluate the genetic diversity of *Curcuma* species. Between these two primers, the highest numbers of bands were observed in *Curcuma longa* (S 1) for primer OPC-7. The results showed the maximum similarity between C. *longa* and C. *amada*.

Theanphong *et al.* (2016) analyzed the phylogenetic relationship between 15 *Curcuma* species using 30 RAPD primers. Twenty-two to twenty-eight amplicons were observed, with an average of 24.5 bands for each primer. The cluster diagram was developed and divided into three major groups and the phylogenetic relationships were analyzed and correlated with the morphological characters. The results concluded that RAPD marker can be successfully used for differentiating 15 *Curcuma* species and provide a simple and rapid tool for differentiation.

Khan *et al.* (2013) studied the fingerprint of 18 local turmeric genotypes, using 22 RAPD primers. Prashanth *et al.* (2015) also used RAPD technique for analysing the genetic diversity of turmeric genotypes using four primers.

2.3.4 Amplified Fragment Length Polymorphism (AFLP)

AFLP is an appropriate molecular marker technique to detect a large number of variable loci distributed throughout the genome and without previous knowledge of genomic sequences (Vos *et al*., 1995). The intraspecific variation and hybrids were detected using this powerful method (Zaveska *et al*., 2011). Ghosh *et al.* (2011) identified three *Zingiber* (six accessions of *Z. officinale*, five accessions of *Z. montanum*, and five accessions of *Z. zerumbet*) species using AFLP methods to detect the variation and constructed a dendrogram for analysis of the phylogenetic relationship among these species. Techaprasan *et al.* (2008) studied the genetic relationships of 15 *Boesenbergia* species using AFLP with a total of 893 allelic fragments were generated, of which 99.78% were polymorphic using six primer combinations.

Kaewsri *et al.* (2007) assessed the genetic relationship among 45 zingiberaceous plants, *Amomum* and 5 outgroup taxa: *Alpinia, Etlingera* 1,

Etlingera 2, *Elettaria* and *Geostachys* using AFLP technique. Five primers were used to generate 122 polymorphic loci and cluster analysis was performed by using the UPGMA method. Kladmook *et al.* (2010) analyzed the genetic diversity in cassumunar ginger (*Zingiber cassumunar* Roxb.). 309 fragments were developed of which, 242 bands were polymorphic and UPGMA method was employed for cluster analysis. Wahyuni *et al.* (2003) studied the genetic relationships among ginger accessions based on AFLP marker, 28 accessions were used for the analysis.

Ghosh *et al.* (2011) identified three zinger species *Zingiber officinale*, *Z. montanum* and *Z. zerumbet* using species-specific AFLP markers with seven selected primers. A total of 837 fragments were generated using these primer pairs. Species-specific markers were identified for three Zingiber species (91 for *Z. officinale*, 82 for *Z. montanum*, and 55 for *Z. zerumbet*). Dendogram constructed using the bands showed that the species *Z. montanum* and *Z. zerumbet* are phylogenetically closer to each other than to *Z. officinale.*

Moon *et al.* (2013) investigated the genetic diversity of five *Curcuma* species viz, *C. longa, C. aromatica, C. zedoaria, C. phaeocaulis and C. kwangsiensis* using AFLP. Six primer combinations were used to reveal 643 fragments of which 349 were polymorphic with 54.3% polymorphism. Inter relationship between the different accessions of *Curcuma zanthorrhiza* was detected using AFLP method, 28 primer combinations were used to differentiate the accession from others (Damayanti 2012).

2.3.5 RFLP (Restriction Fragment Length Polymorphism)

Restriction fragment length polymorphism (RFLP) is one of the first technique extensively used for detecting the variation in the DNA sequences. DNA molecules are digested with different restriction enzymes, resulting in different sized DNA fragments from the two organisms, that differ in the

distance between sites of cleavage of a particular restriction endonuclease. Mutations encountered by the organism affect the DNA molecules in different ways due to this they develop fragments of variable lengths. These differences in the fragment length can be observed using appropriate methods such as gel electrophoresis, hybridization and visualization.

Ahmad *et al.* (2009) studied the genetic variability in the chloroplast DNA in 22 accessions of Zingiberaceae using restriction fragment length polymorphism. Specific combinations of restriction enzymes were selected to find out the relationship of species; suggested that RFLP method can be used to resolve interspecific variation in the highly conserved and variable regions of chloroplast genome.

2.3.6 Inter-Simple Sequence Repeat (ISSR)

The genetic relationships of *Curcuma alismatifolia* (Zingiberaceae) accessions were studied using ISSR markers. Sixteen primers were used to amplify 139 fragments; among these varieties, 77% showed high polymorphism (Taheri *et al.,* 2012). Saha *et al.* (2016) analyzed molecular fingerprints of four different species of *Curcuma,* viz*., C. amada, C. caesia, C. longa* and *C. zedoaria* using inter simple sequence repeats. Twenty ISSR primers were used to amplify 116 loci and cluster analysis was performed using the UPGMA method. Reena *et al.* (2017) evaluated the molecular characterization of the endangered medicinal plant *Hedychium coronarium* using 9 inter simple sequence repeats.

Siriluck *et al.* (2014) identified 24 species of Zingiberaceae using ISSR technique. Eleven primers were used to amplify 127 fragments which were 97.64% polymorphic. Hsieh *et al.* (2014) studied the genetic diversity of Zingiberaceae and assessed the genetic diversity and phylogenetic relationship by using inter-simple sequence repeat (ISSR) markers. Singh *et*

al. (2012) evaluated the genetic diversity in turmeric (*Curcuma longa* L.) using RAPD and ISSR markers. 6 ISSR markers were used to amplify 66 bands, the percentage polymorphism using ISSR primers ranged from 83 to 100 with an average of 95.4%.

Mohanty *et al.* (2014) analyzed the genetic diversity and gene differentiation among ten species of Zingiberaceae using 19 RAPD, 8 ISSR and 8 SSR primers. Prashanth *et al.* (2015) studied DNA isolation and PCR amplification of turmeric varieties using RAPD, ISSR, SSR markers. 20 ISSR primers were used to examine the relationships between different genotypes.

2.4 Biochemical Markers

2.4.1 Isozymes

Isozyme markers are considered as a quick and easy to use method because they do not require DNA extraction, sequence information, primers or probes. The biochemical marker variations are used for the detection of protein polymorphism. Approximately 90 isozyme systems have been used for plants and in many studies, the isozyme loci were mapped (Kumar, 2010). Isozyme markers have been successfully applied for Zingiberaceaea members like *Curcuma alismatifolia*, *Boesenbergia, Kaempferia, Scaphochlamys* and wild ginger (*Siphonochilus aethiopicus*). Using the markers, the natural populations of ginger demonstrated higher diversity compared to cultivated populations. Percentage of polymorphisms varied from 4.5 - 100% in cultivated populations, while in natural populations, they were found to be 50-100%. Isozyme markers are also been used to examine the relationship among ginger taxa and supported the taxonomical classification (Vanijajiva *et al.,* 2003).

Setyawan *et al.* (2014) analyzed the variation in the isozyme pattern of three ginger (*Zingiber officinale*) varieties using isozyme method. The

relationship among the ginger varieties was determined using the UPGMA method. Kar *et al.* (2014) studied the genetic diversity of ginger species (*Zingiber officinale* Rosc.) using different isozymes. Totally 108 proteins were determined, out of which 33 were unique. A total of 85 bands were detected for the four isozymes. Usefulness of the isozyme markers in the identification some of the early flowering *Curcuma* species have been reported (Apavatjrut *et al.,* 1999). Twenty-one isozymes were used; out of these, eight isozymes showed reliable polymorphism. Patterns from isozyme data were analyzed using cluster analysis and UPGMA to produce a dendrogram for depicting relationship among the species.

Vanijajiva *et al.* (2003) investigated the relationship among *Boesenbergia* (Zingiberaceae) using isozymes markers. Nine enzymes were used to analyze the relationship between 11 taxa of *Boesenbergia*. The results were analyzed using the UPGMA method. Tang *et al.* (2008) explored the genetic relationships of six *Curcuma* species using isozymes. Deng *et al.* (2011) studied the relationships among the six herbal species (*Curcuma*) using four isozymes. Totally 168 polymorphic isozymes bands were identified. These bands were used to construct the phylogenetic tree using the UPGMA method.

Shanmugapriya and Prabha, (2012) analyzed the genetic relationship between *Amomum* species using isoenzyme method. Genetic relationship among four *Amomum* species was analysed using four isoenzymes like peroxidase, esterase, polyphenol oxidase and acid phosphatase. Setyawan *et al.* (2014) studied the diversity of cardamom using isozyme method, ten accessions of local Java cardamom (*Amomum compactum*) and true cardamom (*Elettaria cardamomum*) was selected for the study. Two isozyme systems were assayed, namely esterase (EST) and peroxidase (PER, PRX)

with nine esterase and ten peroxidase isozymic bands and the phylogenetic relationship was determined using UPGMA method.

2.5 Phylogenetic analysis

Molecular plant systematics involves the analysis of variation in organelle and nuclear DNA to decipher plant phylogeny at all taxonomic levels (Soltis and Soltis, 1998). Phylogenetic reconstruction of the selected locus (DNA) mainly based on individual copy number, the capability to resolve relationships, character similarity (low homoplasy), a suitable number of parsimony-informative characters and evolution rate of the relative taxonomic group apart from the function of the gene. The constructed phylogenetic trees can be used to study the evolutionary history of the living organism and also provide the evidence of the climatic and geological history of the earth. A phylogenetic tree displays the relationship among the taxonomic groups in a hierarchical order (Futuyma, 2005).

Molecular phylogenetic methods are mainly used in the perspectives of biological systematics and are used for a wide variety of applications, such as community ecology, proteomics, as well as inference of protein-protein interactions (Pazos and Valencia, 2001). The gene must show sufficient variation (variable nucleotides) to mark the important cladogenic events. Although the variability with visible multiple changes in each site have erased the evolutionary information and give background noise without phylogenetic signal (Doyle and Luckow, 2003). From the phylogenetic point of view, a well-corroborated phylogeny would provide means for evaluating character evolution (Borsch *et al.,* 2008), molecular evolution (Grimm and Denk, 2007) and historical biogeography (Lohne *et al.,* 2008).The more recent studies of the genus *Curcuma* were done by Velayudhan *et al.,* (1999) and molecular taxonomy by Kress *et al.,* (2002).

2.5.1 The molecular phylogenetic analysis in family Zingiberaceae

Some of the most commonly used DNA region (which is also used in this study) for the phylogenetic analysis of the genus *Curcuma* and other genera of Zingiberaceae are discussed below.

Sabu (1991) studied the taxonomic and phylogenetic analysis of Zingiberaceae from South India. The gene should show adequate variation (variable nucleotides) to mark the important cladogenic events, without varyning multiple changes in each site probebly have erased the evolutionary information (Doyle and Luckow, 2003).

Dhivya *et al.* (2008) studied the phylogenetic relationship of Zingiberaceae using *mat*K gene region. Forty-seven genera were studied and a phylogenetic analysis was carried out using the PHYLIP package. The results showed that the Zingiberaceae members *Afromonum*, *Alpinia*, *Globba*, *Curcuma* and *Zingiber* showed a polyphyletic relationship. Shreth *et al.* (2012) analysed the phylogenetic relationship of Zingiberaceae using *mat*K gene sequences and the sequence was retrieved from GeneBank and further used to construct a phylogenetic tree. ClustalW program was used to align the sequences and phylogenetic analysis was carried out using PHYLIP package. The phylogenetic relationship showed five important genera, viz. *Alpinia, Boesenbergia, Curcuma, Hedychium* and *Kaempferia*.

Santhoshkumar and Yusuf (2017) studied the Phylogenetic relationship among the *Curcuma* species using chloroplast *mat*K gene. Twenty *Curcuma* species were used and constructed for phylogenetic tree. The phylogenetic analysis was carried out using MEGA 6.0 software. Based on the molecular character related to phylogenetic analysis, *C. aromatica* and *C. raktakanta* are close to each other. The results showed the paraphyletic origin of these two species is moderate with (100%) bootstrap value.

Santhoshkumar and Yusuf (2018) evaluated the efficiency of chloroplast markers for species discrimination among the genus *Curcuma*. Maximum parsimony trees were generated for *rbcL* and *mat*K loci. The two species like *C. aromatica* and *C. raktakanta* was supported with 100% bootstrap in a single clade. The *mat*K sequences showed considerable variations between *Curcuma* species, so these results recommended that *mat*K can be used as a promising candidate barcode for species identification and also for inferring relationship within the *Curcuma* species.

Julius *et al.* (2008) studied the molecular phylogeny of Bornean *Plagiostachys* (Zingiberaceae) using an internal transcribed spacer (ITS). Totally 111 taxa were used including 25 taxa of *Plagiostachys* for this analysis. Based on the ITS phylogenetic results, *Plagiostachys* comprised three subclades (A, B and C) and each subclade was moderately to strongly supported with relatively high bootstrap values.

Ngamriabsakul *et al.* (2003) examined the phylogenetic relationship of the tribe Zingibereae (Zingiberaceae) using nuclear ITS (nrDNA) and chloroplast DNA *trnL–F* (cpDNA) regions. Phylogenetic trees were generated using PAUP* Version 4.0b4. Zingiberaceae is monophyletic with two major clades, *Curcuma* clade and the *Hedychium* clade. *Bosenbergia* is not a monophyletic group; *Curcuma* clade composed of seven genera and appeared as the third group. The first group comprised of three species *C. alismatifolia, C. harmandii* and *C. parviflora*. The second group is comprised of *C. amada* and *C. rubescens.*

Ngamriabsakul *et al.* (2006) used *a pet*A-*psb*J spacer to clarify the relationship within the genus *Boesenbergia.* The cladogram showed *Boesenbergia* is monophyletic with moderate support (81% bootstrap value). Two subclades are recognized, *B. bambusetorum* is grouped within the clade

of the other two populations of *B. longiflora* with weak support (67% bootstrap value).

Smith *et al.* (1993) constructed the cladogram of Zingiberaceae based on *rbcL* sequences. The order can be divided into two sister groups, one containing Costaceae and Marantaceae, and the other, the remaining six families. All families are monophyletic with the exception of the Musaceae, which is paraphyletic with the family Cannaceae.

Cao *et al.* (2010) examined the interspecies phylogeny to identify the *Curcuma* species using 18S rRNA gene in the nuclear ribosomal DNA (rDNA) and *trnK* gene in chloroplast DNA (cpDNA). The results showed that *Curcuma* is a monophyletic group having two sister genera *Hedychium* and *Zingiber*. The *trnK* sequences showed significant variations between *Curcuma* species. They suggested the *trnK* gene as a promising candidate for barcoding of *Curcuma* species, which provide valuable information for inferring the relationship within species but are insufficient to resolve relationships among closely related taxa.

Zaveska *et al.* (2012) performed phylogenetic analysis of the genus *Curcuma* using plastid and nuclear sequences. A new infrageneric classification was proposed with a formal description of a new subgenus, by cloning the ITS regions and analysing the 'intra-individual ITS polymorphism'. The maximum parsimony and Bayesian analyses were chosen to predict the phylogenetic relationship and these results support a broad generic boundary for *Curcuma*, with the addition of *Laosanthus*, *Paracautleya*, *Stahlianthus*, *Smithatris* and some species of *Kaempferia* (*K. scaposa*, *K. candida*) and *Hitchenia* (*H. caulina*, *H. glauca*). They suggested that further studies are required to reveal the relationship of highly complex species in this subgenus.

Deng *et al.* (2015) analyzed the evolutionary relationship of six *Curcuma* species using three chloroplast regions *rbcL, psbA-trnH* and *petA-psbJ*. Maximum Parsimony tree was constructed using MEGA 4.0. The study reported that *Curcuma sichuanensis* is a mutated species of *C. longa*, hence cannot defined as a single species and *C. chuanhuangjiang* is an individual species.

Kress *et al.* (2005) studied the molecular phylogeny of *Alpinia* using plastid *mat*K region and nuclear internal transcribed spacer (ITS) regions. He proposed a new system of classification using molecular phylogeny. A parsimony analysis was chosen for both individual and combined data sets, recognized six polyphyletic clades with *Alpinia* species. The results suggest that many large genera including *Alpina, Amomum, Boesenbergia, Curcuma, Etlingera* and *Globba* are not monophyletic. Since, the genus *Curcuma* appeared in three groups, the first group having a sister to the genus *Hitchenia*, the second group is *Smithatris* and last group is *Stahlianthus.*

Williams *et al.* (2004) studied the phylogeny, evolution and classification of the genus *Globba* using nuclear internal transcribed spacer (ITS) and plastid *trnK-mat*K regions. He proposed a new infrageneric classification system for *Globba*, recognizing three subgenera and seven sections. The results showed a monophyletic group of *Mantisia* nested within *Globba* and four species of *Mantisia* are formally transferred into *Globba* but retained as a distinct section.

Theerakulpisut *et al.* (2012) investigated the relationship among 23 species of *Zingiber* using nuclear ribosomal DNA (ITS1, 5.8S and ITS*2*) sequences. Phylogenetic analysis of the ITS region has resolved the taxa under study into four clades, it is related to previously known sectional classification of the genus based on inflorescence type (sects *Zingiber*,

Dymczewiczia (Horan.) Benth., *Pleuranthesis* Benth. and *Cryptanthium* Horan.).

Xia *et al.* (2004) examined the evolutionary relationship between 53 accessions representing 13 genera of Zingiberaceae, 31 accessions of *Amomum* (Alpinioideae) was included, nuclear ribosomal internal transcribed spacer (ITS) and the chloroplast *mat*K coding and non-coding regions were used. Acordingly *Amomum* was classified into a polyphyletic group with three major groups, *Paramomum* as sister to *Elettariopsis*, which are both grouped within one group of *Amomum* and other two groups of *Amomum* share common ancestors with additional genera of the Alpinioideae.

Searle and Hedderson, (2000) studied the molecular phylogeny of tribe Hedychiaeae (Zingiberaceae) using internal transcribed spacer (ITS) of 18S-26S nuclear ribosomal DNA. It was observed that *Zingiber* is the sister group of *Cornukaempferia* and nested within tribe *Hedychiceae*. *Siphonochilus* is excluded from *Stahlianthus*. *Camptandra* is the sister group of these two genera with high support.

Kress *et al.* (2007) investigated the evolutionary relationship based on molecular phylogeny and proposed a new classification of the Zingiberaceae and various genera within the family, e.g., *Globba, Hedychium, Roscoea, Etlingera, Alpinia* and *Amomum*. The phylogenetic results obtained using ITS and *mat*K sequence data seventeen clades were recognized at the generic level even though some remain tentative and this study suggested the need of additional analysis before final taxonomic circumscriptions can be made within this family.

Pedersen, (2004) studied the phylogeny of the subfamily *Alpiniaceae*, mainly *Elingera* based on nuclear (ITS) and plastid DNA (*rps16*) regions.

Cladogram of that the genus *Elingeria* formed a monophyletic genus with four major clades and have *Hornstedtia* as a sister group. The inclusion of *Achasma, Geanthus* and *Nicolaia* is also strongly supported, but the genus *Geranthus* formed a monophyletic group and could be concluded as a different genus.

Wood *et al.* (2000) analysed the phylogeny of *Hedychium* and their related genera of the family Zingiberaceae used as ribosomal nuclear DNA sequence data. The cladistic analysis results strongly support that *Hedychium* is a monophyletic genus, but relationships with other genera are poorly supported. Four major clades are moderately supported within the genus *Hedychium*. These clades are also differentiated on the basis of the number of flowers per bract.

Prince and Kress (2006), studied the evolutionary distinctiveness and phylogenetic position of Marantaceae using three genomic regions such as chloroplast: *mat*K, *ndhF*, *rbcL, rps*16 intron and *trnL-trnF* intergenic spacer; mitochondrion: *coxI*; nucleus*:* ITS region. The phylogenetic tree was predicted using the parsimony method. These results identified four non-monophyletic genera like *Calathea, Marantochloa, Phrynium*, and *Schumannianthus*.

Molecular phylogenetic analysis of *A. galanga* showed close relation to *A. nigra* in sect. *Allughas* (with tubular bracteoles) using internal transcribed spacer (ITS) region assessed (Rangsiruji *et al.,* 2000). *Alpinia* is paraphyletic with respect to sect. *Alpinia* subsect. *Catimbium* and also strongly support the sister relationship of *A. rafflesiana* and *A. javanica* in section *Allughas* hence the whole section is paraphyletic.

Chen *et al.* (2015) studied the phylogenetic relationship of closely related species of *Curcuma* (Zingiberaceae) based on DNA barcodes. Four

chloroplast DNA regions (*mat*K, *rbc*L, *trn*H-*psb*A and *trn*L-F) and one nuclear region (ITS2) were used to analyse 44 *Curcuma* species. PCR amplification success rate, intra- and inter-specific genetic distance took an account for correct identification of barcode regions. PCR and sequence success rate was also noted and the highest rates such as *mat*K (89.7%), *rbcL* (100%), *trnH-psbA* (100%), *trnL-F* (95.7%) and ITS2 (82.6%) reported in all the regions. No barcode gaps were present using four candidate chloroplast barcoding regions (*mat*K, *rbcL, trnH-psbA* and *trnL-F*) so this results signify that genus *Curcuma* represents a challenging group for DNA barcoding. The blastclust method was also used to analyse these five regions.

Vinitha *et al.* (2014) analyzed the discrimination of Zingiberaceae species using DNA barcoding method. Nine plastids (*mat*K, *rbcL*, *rpoC*1, *rpoB*, *rpl36-rps8*, *ndhJ*, *trnL-F, trnH-psbA, accD*) and two nuclear (ITS and ITS2) barcode regions were used to analyze 60 accessions of 20 species. Forward and reverse sequences were recovered for all loci using direct sequencing of PCR products. But only 35 (58%) and 40 accessions (66%) yielded ITS and ITS2 sequences by direct sequencing. Different bioinformatics analyses were conducted using plastid and nuclear loci, 15 species grouped (75%) into monophyletic groups and five species into two paraphyletic groups using *mat*K and *rbcL*. Only 12 species (60%) was distinguished by using 173 ITS sequences, including 138 cloned sequences from 23 accessions and the remaining species were entered into three paraphyletic groups. The results showed that using the barcodes *mat*K and *rbcL* loci for the family Zingiberaceae can be analysed.

Shi *et al.* (2011) resolved the potential of DNA barcodes for species identification in family Zingiberaceae using chloroplast and nuclear locus to compare the genetic divergence and species identification efficiency of ITS2*, rbcL, mat*K*, psbK–psbI, trnH–psbA* and *rpoB* for many closely related

species. The highest interspecific divergence and significant differences between inter and intraspecific divergence was noticed using ITS2 marker, but *mat*K and *rbcL* produced much lower divergence values.The discrimination ability of the ITS2 locus was 99.5% at the genus level and 73.1% at the species level for the 260 species belonging to 30 genera in the family Zingiberaceae. So ITS2 is the ideal DNA barcode sequence for identifying Zingiberaceae members.

Zhang *et al.* (2014) studied the DNA barcode of the genus Roscoea (Zingiberaceae) using three chloroplast markers (*rbcL*, *trnH-psbA* and *mat*K) and one nrDNA marker (ITS). These regions (ITS and *trnH-psbA*) had a high success rate in PCR amplification and bidirectional sequencing. Based on the *rbcL* regions, sequences possessed no genetic variation PCR amplification and DNA sequencing in *the mat*K region was relatively difficult. The combination of ITS and *trnH-psbA* could effectively discriminate 90% species using the method of Neighbor-Joining (NJ) tree.

Wijayasiriwardene *et al.* (2017) identified the endemic species of *Curcuma albiflora* using DNA barcoding method. Two chloroplast genomic regions of *mat*K and *rbcL* were used. The amplified fragment of *rbcL* (600 bp) and *mat*K (850 bp) was identified. The interspecific distance for *rbcL* was 0.0010 and for *mat*K 0.05. Based on the interspecific distance the matK sequences were selected for identification of *C. albiflora, C. zedoaria* and *Z. zerumbet.* Neighbor-Joining method was used to construct a phylogenetic tree, *C. albiflora* appeared as a different group. The *mat*K gene and *rbcL* gene of *C. albiflora* and other species showed a total of 212 variable sites (*mat*K) and 2 variable sites in *rbcL*. Therefore, *mat*K gene is an efficient marker for identifying *C. albiflora* from other two species.

2.6 Curcumin synthease gene expression

The curcumin isolated from *Curcuma* species is an orange-yellow crystalline powder practically insoluble in water. Lampe and Milobedeska, (1913) was first described the structure of curcumin ($C_{21}H_{20}O_6$). Several scientific reports confirmed various pharmacological properties of curcumin and as a chemopreventive agent as well as a potential therapeutic agent against several chronic diseases (Priyadarsini, 2014).

Using UV-VIS spectrophotometric investigation, maximum light absorption of curcumin occurred around 400-430 nm depending on the organic solvent (Jasim and Ali, 1989). In sunlight, curcumin is decomposed, both in ethanolic and methanolic extracts and as a solid. Vanillin, vanillic acid, ferulic aldehyde and ferulic acid are identified as the degradation products (Wang *et al.,* 1997).

Gilanil *et al.* (2015) analysed the molecular genetic diversity of curcuminoid genes in *Curcuma amada*. Eight curcumin-containing accessions of *C. amada* and six curcumin free accessions were used in this analysis. Flow cytometry results revealed that the ploidy level was higher in curcumin-containing accessions than in curcumin free accessions. Phylogenetic analysis conducted based on CURS genes showed that CURS2 diverged from CURS1 and CURS3, whereas the DCS genes evolved independently of the CURS genes. DNA barcode analysis showed that the *rbcL, mat*K and *psbA-trnH* intergenic region were identical in all 14 accessions and thus confirmed that all accessions belonged to *C. amada*.

Minami *et al.* (2009) estimated the curcumin content in different *Curcuma* species using DNA polymorphisms. The chloroplast region of *trnS-trnfM* intergenic spacer was used. The *trnSfM* marker provided valuable information for the identification of four *Curcuma* species and for predicting

the curcumin content in the rhizome of *C. longa.* Curcumin was detectable in the rhizomes of *C. longa*, *C. aromatica* and *C. zanthorrhiza*, but was not detectable in that of *C. zedoaria*. The curcumin content of C. *longa* ranged from 0.19% to 3.08% *C. aromatica* ranged from 0.03% to 0.09% DW and *C. zanthorrhiza* was 1.12% dry weight.

Behar *et al.* (2016) studied the semi-quantitative expression of genes involved in the biosynthesis of curcuminoid in *C. caesia* Roxb using major genes (CURS, CURS2, CURS3, DCS & CHS1). RNA extracted from the rhizomes (5- month & 10-month-old) and leaves of the plant was used for cDNA preparation and semi-quantitative expression studies of genes. The genes showed higher expression in rhizome compared to leaves and among all the genes, CURS showed maximum expression followed by DCS, CHS1, CURS2 and least by CURS3. The 5-month-old rhizomes showed four times higher CURS expression compared to the 10-month-old rhizomes.

Sandeep *et al.* (2017) studied the CURS gene expression in *C. longa* during various growth periods and two different agroclimatic zones. Curcumin content was calculated spectrophotometrically and also evaluated the environmental factors. The curcumin content and relative expression were calculated and the highest curcumin content and relative expression were analysed during the sixth month with impact of soil and environment in the genetic improvement of turmeric.

Katsuyama *et al.* (2009) identified and characterized multiple curcumin synthases genes from *C. longa.* Two additional type III polyketide synthases, named CURS2 and CURS3, that are capable of curcuminoid synthesis was identified and characterized. The studies revealed that the CURS2 chose feruloyl-CoA as initial substrate and CURS3 chose both feruloyl-CoA and p-coumaroyl-CoA as starter substrates. These results suggested that CURS2 synthesizes curcumin or demethoxycurcumin and CURS3 synthesizes

curcumin, bisdemethoxycurcumin and demethoxycurcumin. These three enzyme expression levels contribute to curcuminoid synthesis with different substrates.

Hayakawa *et al.* (2011) developed a molecular marker to identify high curcumin content in *C. longa* L. Several sequences of chloroplast DNA were used to identify curcumin content. Proper outgroup was identified and used to conduct infraspecific analyses of *C. longa* and also molecular phylogenetic tree was constructed using allied species. Based on the results *C. aromatica* and *C. zedoaria* are closely related to *C. longa*. A unique haplotype was identified within *C. longa* corresponds to high curcumin content using this molecular marker. However, the chloroplast microsatellite regions were used as markers to determine the lines of this species with high curcumin content.

2.7 Phytochemical analysis

Chemical diversity, low abundance and variability within the same species are challenging tasks in the analysis of secondary metabolites. Gas chromatography (GC) and Mass spectrometry (MS) is a tool for detecting the volatile components according to their elution order on column (Steinmann and Ganaera, 2011). In human beings, several studies established a link between phytochemicals and the range of biological activities that impart on health. They were widely used in Ayurvedha and other traditional medicines (Moon *et al.,* 2010).

Several groups of phytochemicals such as polyphenols (anthocyanins, proanthocyanidins, flavanones, isoflavones and ellagic acid), non-nutrient chemical and dietary constituents are currently used in the pharmaceutical industry. Spices contain a large number of active phytochemicals mainly spices belonging to the genera *Curcuma* and *Zingiber* are well known for their

multiple uses as medicines, cosmetics, dyes, flavorings and neutraceuticals (Policegoudra *et al.,* 2011)

2.7.1 *Curcuma* GC-MS

The volatile oil of turmeric harbours an aroma and flavor and its composition is determined by GC-MS. Physicochemical property and the distillation process also various between individual samples (Krishnamurthy *et al.,* 1976). The differences can appear from a number of factors, including the origin and method of curing the spice, its age and distillation conditions and the stage of maturity of the rhizomes at harvest.

Santhoshkumar and Yusuf, (2019) studied the chemical composition of volatile methanolic rhizome extracts of twenty *Curcuma* species using gas chromatography–mass spectrometry (GC-MS) method. The major components identified in majority of the species were the beta-caryophyllene epoxide, camphor, retinal and alloaromadendrene oxide. Based on the presence of major chemical components, cluster analysis was performed. All the twenty species were grouped based on the presence or absence of a particular compound and a dendrogram was constructed, which grouped the 20 species into eight clusters.

The essential oils of two endemic species of *Curcuma viz C. haritha* and *C. raktakanta* were obtained using hydrodistillation method were analysed using GLC from fresh rhizomes. Eleven components were identified from *C. haritha,* Campher (21.24%) was the major component and ten components were identified from *C. raktakanta* of which ethyl p-methoxycinnamate (16.57%) was the main component. The common components identified in both the species were A-pinene, camphor, terpinyl acetate, tumerone and ethyl P-methoxy-cinnamate (Dan *et al.,* 2002).

Prinitha *et al.* (2012) studied the volatile oil components from the rhizome of *C. longa* isolated using distillation and characterized using Gas Chromatography Mass Spectroscopy. The essential oil yield from methanolic extract was 24.0%, 18.0% from acetone extract and 14.0% from hexane rhizome extract. Major constituents identified were, α-turmerone (38.24%), Camphor (32.3%) and β -turmerone (22.25). George and Britto, (2015), assessed the chemical composition (GC-MS) of essential oil from *C. aeruginosa* Roxb. Eighteen compounds were identified with four major components are Ethoxybenzene, Santolinatriene, Eucalyptol and Camphene.

Awasthi and Dixit, (2009) extracted essential oil from *Curcuma longa* using hydro distillation method and studied the chemical constitutions. GC and GC-MS analysis identified 73 constituents in rhizomes comprising 95.2% of the oil, of which the major ones were ar-turmerone (31.7%), *α*-turmerone (12.9%), *β*-turmerone (12.0%) and (Z) *β*-ocimene (5.5%). In the oil, 75 constituents comprised of *α*-phellantrene (9.1%), terpinolene (8.8%), 1, 8-cinceole (7.3%), undecanol (7.1) and p-cymene (5.5%).

Nayak *et al.* (2014) characterized the phytoconstituents of methanolic rhizome extracts from *Zingiber roseum, Curcuma angustifolia* and *Globba marantina* using GC-MS method. Seventeen compounds from *Z. roseum*, fifteen from *C. angustifolia*, and nine from *G. marantina* were identified. *Z. roseum* has Deoxynivalenol (0.43%), n-heptane (0.74%), Cyclopentane,1-methyl-2 (2-propyl) trans-(2.76%), Ethyl propionate (0.24%), Ethyl heptanoate (0.21%), Naphazoline (0.76%), Estazolam (3.20%), 2-hexane (20.00%), 3-Heptanone (29.29%), 1,4-butanediol (0.93%), Diehylene glycol methyl ether (0.20%), Iso-propyl benzene (Cumene) (4.47%), Isoprene (0.34%), Propofol (0.51%), Pentadecenol (18.08%), n-Amylmercaptan (8.65%), Methyl cyclopentane (9.20%), *C. angustifolia* has 3-hydroxyhexanoic acid(0.51%),α –thujene (1.41%), β- pinene (1.67%), Cis-p-

menth-2,8-dienol (2.14%), 3,5-dimethoxy (Tolune) (1.19%),Caryophyllene (1.39), Flurprimidol (2.61%), 2,7-napthalenediol (12.44%), Camphor (15.38%), Trans-nerolidol(12.02%), Humulen-6,7-epoxide (9.21%), Octadecanoic acid, butyl ester (11.56%), α-amorphene (23.02%), Aristoloshene (3.80%), Diflobenzuran (1.63%), *G. marantina* has Heptadecane (6.19%), Pinocarvone (54.27%), L-linalool (2.54%), Alloaromadendrene (4.58%), β–caryophyllene (13.65%), α- humulene (3.88%), Terpineol (6.31%), Lavandulol (2.63%),2,6-Dimethyl-1,5,7-Octatrien-3-ol (5.95%) .

Essential oil components of four *Curcuma* species, *C. manga, C. zanthorrhiza, C. aeruginosa* and *C. longa* analysed using GC-MS (Jantan *et al.,* 1999). Myrcene (81.4%) was the highest component of the essential oil of *C. manga*. The essential oil of *C. zanthorrhiza,* xanthorrhizol (45.5%) had 1, 8-Cineole (eucalyptol) (23.2%) and Curzerenone (28.4%) major constituent as the essential oil of *C. aeruginosa* the *α*-tumerone (45.3%) and *β*-tumerone (13.5%) were the major components.

Dung *et al.* (1995) studied the leaf essential oil components of *C. domestica* (*C.longa*) using GC and GC-MS, the essential oil was obtained using the steam distillation. More than 20 components were identified, of which the monoterpenes alpha-phellandrene (24.5%), 1, 8-cineole (eucalyptol) (15.9%), P-cymene (13.2%) and beta-pinene (8.9%) were the major ones. Kitamura *et al.* (2007) compared essential oil components in rhizomes of three *Curcuma* sp. using GC-MS technique. The total ion exchange chromatogram for essential oil fractions prepared from the rhizome of *Curcuma* sp.

Jarikasem *et al.* (2003) identified the hydro-distilled phytochemical components of essential oil *C. zanthorrhiza, C. aromatica* and *C. aeruginosa* (Zingiberaceae) analyzed by capillary GC and GC-MS. The major

components identified from *C. zanthorrhiza* were 1,8-cineol (37.58%) and curzerenone (13.70%) whereas camphor (26.94%), ar-curcumene (23.18%) and xanthorrhizol (18.70%) were found in the essential oil of *C. aromatica*. For *C. aeruginosa*, β-Pinene (7.71%), 1, 8-cineol (9.64%) and curzerenone (41.63%) were major constituents.

Zwaving and Bos, (1992) evaluated the chemical composition of the essential oils from the rhizomes of five *Curcuma* species by GC-MS. The major compounds identified were *C. zanthorrhiza* Roxb.: *ar*-curcumene (41.4%), xanthorrhizol (21.5%); *C. domestica* Val.: *ar*-turmerone (24.7%), turmerone (29.5%), turmerol (20.0%), α-atlantone (2.4%) and *C. aromatica* Salisb.: *ar*-curcumene (18.6%), β-curcumene (25.5%), xanthorrhizol (25.7%); *C. aeruginosa* Roxb.: isocurcumenol (8.5%), β-eudesmol (6.5%), curdione (3.6%), curcumenol (9.9%), curcumanolides A,B (11.4%), dehydrocurdione (9.4%), curcumenone (1.9%); *C. heyneana* Val.: 1,8-cineole/limonene (14.2%), isocurcumenol (7.4%), β-eudesmol (4.7%), curcumanolides A,B (13.1%), dehydrocurdione (10.2%), curcumenone (2.3%) respectively.

Raina *et al.* (2005) studied the essential oil composition of *C. longa* L. (Zingiberaceae) isolated from the rhizomes and leaves using hydrodistillation and oil was analyzed using GC and GC/MS. The major constituents obtained from rhizome oil was, *α*-turmerone (44.1%), *β*-turmerone (18.5%) and ar-turmerone (5.4%). 61 compounds were identified from the obtained 99.8% leaf oil and most important chemical compounds were identified as α-Phellandrene (53.4%), Terpinolene (11.5%) and 1, 8-Cineole (10.5%) respectively.

Uchegbu *et al.* (2014) determined the essential oil constituents from the ethanolic extract of *Curcuma longa* rhizome by GC- MS analysis. Five compounds were identified with an ester, 3-methyl-2-butenoic acid, 3-phenyl-

2-propenyl ester (45.64%), followed by a ketone, known ascurlone (26.84%) and a carboxylic acid, 9-octadecenoic acid (16.78%).

Garg *et al.* (2002) investigated the essential oil components from *C. longa* L. (Zingiberaceae), the oil obtained by hydrodistillation was used for GC and GC–MS analysis. Twenty compounds were identified, accounting for 72% of the contents. The oil of *C. longa* contain mainly of monoterpenoids; monoterpene hydrocarbons (57%), oxygenated monoterpenes (10%), sesquiterpene hydrocarbons (3.3%) and oxygenated sesquiterpenes (2.1%). The major constituents of the oil was p-cymene (25.4%) and 1, 8-cineole (18%), followed by cis-sabinol (7.4%) and α-pinene (6.3%).

2.7.2 HPLC (High Performance Liquid Chromatography) curcumin estimation

A large number of analytical methods have been used in curcumin quantification and recent versions of new methods are also reported (Revathy *et al.*, 2011; Jayaprakasha *et al.*, 2002). Syed *et al.* (2015) reported curcumin stability in the different formulation using the HPLC method.

Long *et al.* (2014) analyzed three active components of curcumin, demethoxycurcumin and bisdemethoxycurcumin in *C. longa* L. using HPLC coupled with electrochemical detection method. Acetonitrile and 10 mM Na_2HPO_4–H_3PO_4 (pH 5.0) (50:50, v/v) was used as mobile phase. Good linearity was observed in the methanolic extract, ranging from 0.208–41.6 for curcumin, 0.197–39.4 for demethoxycurcumin and 0.227–114μM for bisdemethoxycurcumin.

Nugroho *et al.* (2015) studied the quantitative analysis of curcumin in ethanolic extract of turmeric (*C. longa* L.) and *C. zanthorrhiza* (Zingiberaceae) using the HPLC method. Waters Xterra MS C_{18} column (4.6×250 mm, 5 μm) was used to separate the three curcuminoids such as

curcumin, demethoxycurcumin and bisdemethoxycurcumin. Aquabidestilata and acetonitrile (65:35 v/v) containing 1% (v/v) acetic acid was used to study the mobile phase and linearity, sensitivity, precision and accuracy were confirmed using the analytical method. Curcumin concentration ranged from 10-60 µg/mL with a correlation coefficient value of 0.999. The levels of curcumin in the rhizomes were in the range of 3.03±0.01-7.31±0.02 in *C. longa* and 1.69 ± 0.02-4.92 ± 0.01µg/g in *C. zanthorrhiza*.

Sharma *et al.* (2012) determined the curcumin concentration from drugs used in pharmaceuticals using methanol as a solvent. UV-visible spectrophotometry was used to quantify curcumin content and spectrophotometric detection was recorded at 421 nm. The detector response was linear over the concentration range of 1-7µg/ml with correlation coefficient 0.9995. The LOD and LOQ were reported to range from 0.05 and 0.172µg/ml respectively.

Revathy *et al.* (2011) analyzed the curcuminoid samples using HPLC with SPD-M20AuV detector in isocratic mode. The mobile phase was 40% THF and 60% water containing 1% citric acid. The isolated curcuminoids (CU, DMC, and BDMC) showed peaks at 10.81, 12.79 and 13.03 min interval respectively.

Ashraf *et al.* (2012) developed HPTLC method for the quantification of curcumin, toluene: chloroform: methanol (5:4:1, v/v/v) was used as mobile phase and silica gel 60 F254 used as stationary phase using a wavelength of 430nm. Linearity was observed in the concentration range between 200-1000ng/ml.

Hastati *et al.* (2015) identified the curcumin pigments from *Curcuma* species using HPLC method, acetonitrile: acetic acid: aquabides was used as mobile phase and detected at 425nm. Thus the analytical methods used for

curcumin were possible for quantitative analysis. The curcuminoid content from the ethanolic extract of *C. domestica* had the highest curcumin (16.1%) followed demetoxycurcumin (3.2%) and bisdemetoxycurcumin (2.8%).

Himesh *et al.* (2011) investigated the qualitative and quantitative determination of curcumin in the ethanolic extract of *C. longa*. The total phenolic content of the ethanolic extract in *C. longa* was found to be 11.24 as mg GAE/g. The pharmacologically important active curcuminoid components such as curcumin, demethoxycurcumin and bisdemethoxycurcumin were determined in *C. longa* using spectrophotometry and HPLC methods. The separation of samples were performed on a Cyber Lab C-18 column (250 x 4.0 mm, 5μ) using acetonitrile and 0.1% orthophosphoric acid solution in water in the ratio 60: 40 (v/v) at a flow rate of 0.5 mL/min. Detection of curcuminoids was performed at 425 nm.

Pandey *et al.* (2010) identified curcuminoid pigments from different turmeric genotypes (*C. domestica* Val) and quantified the levels of curcuminoids using HPLC method. Twenty two *Curcuma* species were selected and isolated curcumin from all the *C. domestica* genotypes using organic solvents like ether, benzene and ethanol by Soxhlet apparatus. The isolated curcumin was purified using column chromatography and the fractions were identified using thin layer chromatography. The total curcuminoid percentage ranged from 0.069 to 6.16. The maximum curcuminoid content was identified in genotype T22 (Wild local 2) (6.1%) followed by T21 (Wild local 1) (4.04%), T1 (TCP-1) (2.817%), T2 (TCP-2) (2.28%). The range of curcumin content varied from 1.87% to 0.015%. The genotype T22 (Wild local 2) showed the highest curcumin percentage.

Lee *et al.* (2013) identified the curcuminoids from Asian turmeric (*C. longa* L.) using RP-HPLC and LC-MS methods. Three major curcuminoids were extracted and identified using a fast and reliable HPLC-UV-MS method

and thin layer chromatography, including curcumin, demethoxycurcumin and bismethoxycurcumin. The analysis used C_{18} column and the UV wavelengths of 425 and 254 nm. The extraction time of 1 to 4 h is the best time for total extraction of curcumin and demethoxycurcumin. The efficient molecular weight information for the two major curcuminoids in the extracts were provided by the LC-MS using thin layer chromatography.

Jayaprakasha *et al.* (2002) determined the levels of curcumin, demethoxycurcumin, and bisdemethoxycurcumin using HPLC method. These three isoforms of curcumin were isolated and separated using column chromatography and identified using spectroscopic methods. The curcuminoids were purified and analysed by HPLC and separation was performed on a C_{18} column using three solvents, methanol, 2% AcOH, and acetonitrile at 425 nm. Four different samples were selected and analysed to detect the percentage of three curcuminoids (curcumin, demethoxycurcumin, and bisdemethoxycurcumin 2.34 ± 0.171 to 9.18 ± 0.232%).

Jadhav *et al.* (2007) used isocratic reverse phase-HPLC method to determine the curcuminoids. The mobile phase consist of acetonitrile: 0.1% trifluoroacetic acid (50:50 v/v), and the flow rate was 1.5 mL min^{-1} and the elution was monitored at 420 nm. The retention time of BDC, DC and C were reported at 7.2, 8.12 and 9.0 min, respectively. Linear range was from 100 to 600 ng/ mL.

The purity of the curcuminoids fractions was determined using HPLC (Naidu *et al*., 2009) in which exil-amino column with a 2-propanol: water (95:05, v/v) mixture as the mobile phase at a flow rate of 1 mL min^{-1} . Related studies were reported by Almeida *et al.* (2005) in which HPLC with NH_2 column and a mobile phase containing ethanol: water (85:15, v/v), the flow rate of 1 mL min^{-1} was used with diode array detector. The isolated curcuminoids showed peaks at retention time 4.71, 3.41, 2.90 min and

identified as curcumin, demethoxycurcumin, bisdemethoxycurcumin respectively.

Avula *et al.* (2012) quantified the curcuminoid content from *C. longa* and dietary supplements by UPLC-UV-MS Method. C-18 column was used and the mobile phase contained water/acetonitrile and formic acid at a temperature of 35˚C, the separation was obtained within 3.5min. The limits of detection and limits of quantification of curcuminoid were found to be 0.01µg/mL and 0.03µg/mL, respectively. The curcuminoids content of *C. longa* (curcumin, desmethoxycurcumin, bisdemethoxycurcumin) was found to be in the range from 1.16-4.92% and 0.83-35.37% respectively. The curcuminoid content was 0.004% in *C. zedoaria* and 0.0006% for *C. phaecaulis*.

Tejavathi *et al.* (2017) estimated the curcuminoid content of *C. karnatakensis* using ultra-high performance liquid chromatography-mass spectrometry/selected reaction monitoring method. The samples were collected from two different habitats and the stock solution was prepared using 5 mg of standard in 1 mL of methanol. C-18 column was used for curcuminoid separation with a flow rate of 0.2 mL/min^{-1}.

Quantitative analysis of curcuminoids in different accessions of *C. longa* used high-performance thin layer chromatography (Paramasivam *et al.,* 2009). TLC aluminum plates pre-coated with silica gel 60GF254 was used as stationary phase and solvent system contained chloroform: methanol (48:2, v/v). The compact spots were identified as curcumin, demethoxycurcumin and bisdemethoxycurcumin (RF value of 0.66 ± 0.02, 0.48 ± 0.02 and 0.30 ± 0.02), respectively. Densitometric analysis of curcuminoids was determined by absorption– reflection detection mode at 425 nm.

Zhang *et al.* (2009) quantified the curcuminoids using a simple HPLC fluorescence method. The separation of curcuminoids was achieved within 30 min using Cadenza CD-C_{18} column and a mixture of acetate buffer and CH_3CN. Good linearity was obtained in 0.993 as the correlation coefficient. The detection limits of CU, DMC and BDMC were 1.5, 0.9 and 0.09 ng mL^{-1} respectively. The results suggested that this method is more effective in determining curcuminoids in commercial turmeric products, such as turmeric powders, tablet, dressing, beverage, tea, and crude drugs.

Chao *et al.* (2018) investigated the bioactive compounds in the rhizome of *C. longa* L. (Turmeric), three curcuminoids; Curcumin, bisdemethoxy-curcumin, and demethoxycurcumin; and three volatile components; ar-turmerone, *β*-turmerone, and *a*-turmerone were quantified – using HPLC. Pressurized liquid extraction (PLE) method was employed for - sample extraction. In HPLC analysis Zorbax SB-C18 column was used and acetonitrile and 0.4% (v/v) aqueous acetic acid were used as the gradient mobile phase with a flow rate of 1.0 mL min^{-1}. The curcuminoids were detected at 430 nm and this method was successfully applied to detect increased curcuminoids content in different turmeric samples.

Osorio-Tobon *et al.* (2016) used HPLC method with a fused-core column for enabling faster and highly efficient separations of curcuminoids from *C. longa* L. A step-by-step procedure was used to optimize temperature between 40–55^{o}C and the flow rate of 1.0–2.5 mL min^{-1}. The combination of acidified water and acetonitrile was used as a gradient method with high column temperature (55^{o}C) and flow rate (2.5 mL min^{-1}). The separation of these three curcuminoids had taken 1.3 min with an average total analysis time of 7 mints.

2.8 Protein modeling

Comparative protein modeling predicts an unknown protein structure based on its sequence homology to one or more known homologous protein structures. Template based methods have strong impact in predicating the structure of a protein (Petrey *et al.,* 2015). The homology modeling can sometime provide a useful 3-D model for a protein that is related to at least one known protein structure, in the absence of an experimentally determined structure (Marti-Renom *et al*., 2000). Comparative modeling build a 3-D structure of a given input of protein sequence (target) based initially on its alignment to one or more proteins of known structure (Webb and Sali, 2017). Modeling of protein structures usually requires wide knowledge in structural biology and the use of extremely specialized computer programs for each of the individual steps in the modeling process (Tramontano *et al.,* 2001).

The comparative modeling of protein consists of four main steps (Marti-Renom *et al.,* 2000); (i) fold assignment, which identifies similarity between the target and at least one known template structure; (ii) target and template sequence alignment (iii) building a model using the aligned template(s) and (iv) predicting model errors (Fig. 2.1). In automate comparative modeling process numerous computer programs and web servers are used. The precision of the models are calculated by many of these servers and evaluated by CAMEO (Haas *et al.,* 2013) and the biannual CASP (Moult, 2005; Moult *et al.,* 2009).

SWISS-MODEL workspace (Arnold *et al.,* 2006) can be freely accessed by the biological community on the modeling of protein structures at web http://swissmodel.expasy.org/workspace/. It has been the first automated modeling server freely accessible for public (Peitsch, 1995). The workspace of SWISS-MODEL provides a summary showing a small ribbon representation, experimental details, information about bound molecules, as

well as links to PDB (Westbrook *et al.,* 2003), CATH (Pearl *et al.,* 2005), PDBsum (Laskowski *et al.,* 2005) and MSD (Velankar *et al.,* 2005).

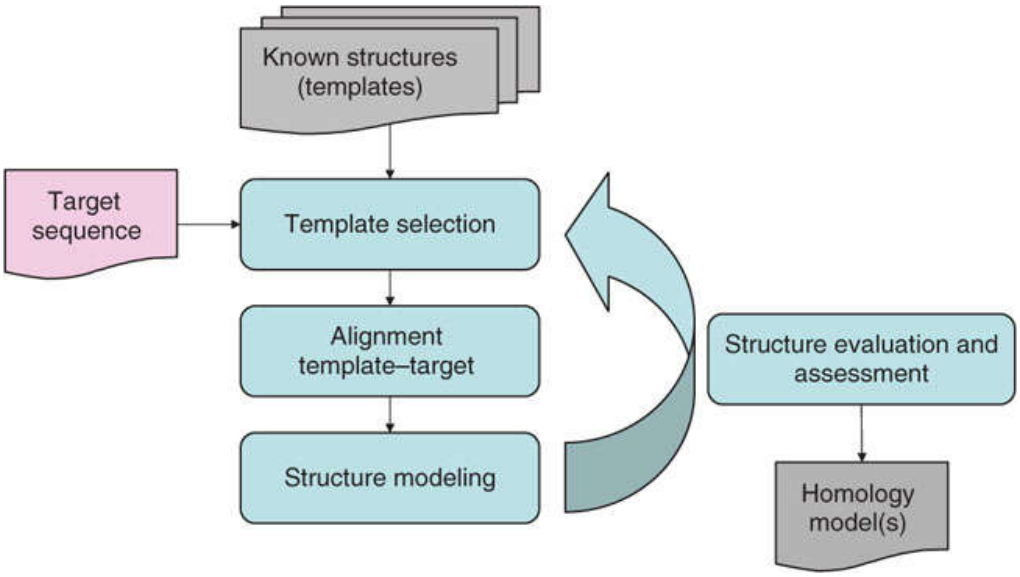


Fig.2.1. Steps involved in comparative protein structure modeling (Bordoli *et al.,* 2008).

CHAPTER 3
MATERIALS AND METHODS

3.1 Collection of plant material

Twenty *Curcuma* species and one variety were collected from different regions of South India. All the collected plants were maintained in pots in the premises of Department of Botany, University of Calicut and the leaf and rhizomes were used for the extraction of DNA, RNA and active ingredients. The location of collection and voucher numbers provided for different *Curcuma* species are presented in Table 3.1.

3.1.1 Plant material used for the study

C. aromatica Salisb

Known as aromatic turmeric, wild turmeric, or yellow zedoary, is the most widely used *Curcuma* species next to the common turmeric (*C. longa* L.). Rhizomes are orange yellow in colour, aromatic inside, with many sessile tubers. Flowers are brownish-yellow, pseudostem 90-120 cm high, inflorescence lateral, corolla funnel shaped, pinkish-white. The rhizomes are used for treating scabies and the eruption of small pox.

C. bhatii (R.M. Sm.) Skornickova and M. Sabu

Curcuma bhatii (*Paracautleya bhatii* R.M. Sm.) is an endemic, threatened plant. It is the smallest southern Indian Zingiber with a height of 12–15 cm. having small rhizomes. It grows in the crevices of laterite rocks. It is located as few spread populations in Udupi District, Karnataka state. Rhizomes small, leaves glabrous, tip acute, inflorescences central, flower yellow in each bract.

C. coriacea Mangaly and M. Sabu

This species has high ornamental value. Grows in open grasslands or exposed rocky regions at higher elevations and also along the margins of secondary forests. Rhizome small, non aromatic, cylindrical to conical, inflorescence lateral, lateral central, coma bracts deep pink to violet, flowers longer than bracts, and yellow in colour.

C. decipiens Dalzell

Rhizomatous herbs, rhizome coloured pale yellow-white, inflorescence both lateral and central, coma bracts many, lower portion green with purple tip, upper deep purple-pink. Calyx truncate, tip shortly 3-lobed, white with pink spots, corolla funnel-shaped, deep purple, labellum purple with a deep yellow band. This species can be distinguished from other species by a deep purple coma and it is closely related to *C. inodora* Balleter.

C. aurantiaca Zijp

Rhizome small, conical yellow inside, sessile tubers absent, roots many. The inflorescence is terminal in position on a leafy shoot. The bracts are yellow - green in colour and the comma bracts are red to pink or purple in colour. The petal colour is orange to yellow. Flowers longer than the bracts, yellow or orange-yellow, commonly grows in the clear zone of evergreen forests and rubber plantations in South India. It is also a well-known ornamental plant.

C. haritha Mangaly and M.Sabu

Rhizome pale yellow or white inside, non-aromatic branching with root tubers, leaves elliptic, acuminate, and pubescent beneath. Inflorescence lateral, produced early in the season, coma bracts bright pink, corolla white, labellum light yellow with a medium dark yellow band.

C. inodora Blatt.

C. inodora is common in barren laterite areas, plains, and roadsides and is known to occur in tropical moist deciduous forests, although, relatively few collections of the taxon are available. *C. inodora* is closely related to *C. decipiens*. It can be distinguished by its flowers equaling the bracts, three to four flowered cincinni, purple corolla and staminoides and labellum with a dark yellow band at the centre. Rhizome conical pale yellow in the centre, white towards the periphery, leaves distichous, tip acuminate, inflorescence both lateral and central, violet coma, flowers white, corolla tube deep purple, labellum purple with a median bright yellow band.

C. karnatakensis Amalraj, Velay. and Mural.

Plant short, root stock small, tapering, flesh cream colour with a slight mango ginger flavors, leaves broadly ovate-elliptic, inflorescence lateral or central, bracts pale green, ovate, acute, plant with pinkish petiole have pinkish green, flowers rose to white, corolla tube pale yellow with a rose segment. Labellum has a three lobed bright yellow band in the middle.

C. kudagensis Velay., V.S.Pillai and Amalraj

Rhizome small white inside, sessile tubers absent, leaves distichous, margin entire, tip acuminate, pubescent, inflorescence lateral, coma bracts light pink colour, flower long, corolla tube bright yellow, oblanceolate, labellum bright yellow and deeply lobed.

C. longa L.

Rhizome medium conical, deep orange-yellow inside, strongly aromatic, root tubers rare. Leaves distichous, inflorescence central, flowers long equal to the bracts, calyx truncate, white, corolla white, glabrous, labellum trilobed, median dark yellow band. This species is closely related to

C. amada Roxb. but can be distinguished from orange-yellow rhizomes and white coma.

C. montana Roxb.

Rhizome short, light yellow. Leaves elliptic, acute at either end. Inflorescence central, coma bracts white with pink in the distal half. Flowers almost equal or slightly smaller than bracts, pinkish yellow; calyx tube long, lobes truncate, glabrous; corolla tube 18 mm long, lobes unequal, oblong, stamen light yellow, labellum 3-lobed deep yellow.

C. mutabilis Škornick., M.Sabu and Prasanthk.

Rhizome ovoid without branches, cylindrical to conical, deeply buried in the ground, light brown externally, faintly aromatic, fleshy roots end in ovoid root tubers. Leaves petiolate. Pseudostem 5-30 cm long, deeply buried in the soil, inflorescence ventral in the beginning of the season, central in the later season. Coma inconspicuous, usually only uppermost 2-3 bracts, light green or with a red tinge(sometimes pretty deep), upper side sparsely hairy, lower side glabrous, tips rounded with a patch of deep violet mauve colour. Corolla tube white, yellowish or yellow at the base, towards lobes sometimes tinged with pink, red or violet, glabrous; dorsal corolla lobe varies in colour from white, yellowish or yellow, sometimes tinged pink red or deep violet-bluish, mucronate.

C. neilgherrensis Wight

Rhizome small, conical white inside, root tubers fusiform, white inside. Leaves distichous, ovate elliptic to obovate. Inflorescence both lateral or central, coma bracts light to dark pink or violet. Flowers longer than bracts, light yellow, corolla tube light yellow, lobes unequal, pubescent. Labellum with a median cleft yellow with a deep yellow median band.

C. oligantha Trimen.

Rhizome small, conical, non-aromatic, yellow inside, sessile tubers absent, roots many, ovate or fusiform, white inside. Pseudostem 7-10cm, leaves distichous, tip acuminate. Inflorescence lateral or central, fertile bracts green or with pinkish tinge. Flowers longer than the bracts, calyx white to light green, corolla tube equal to or longer than the bract, light yellow, yellow or pinkish tinged outside. Labellum shortly 3-lobed, deeply split at the apex, white with a yellow spot in the middle or orange-yellow throughout.

C. pseudomontana J.Graham

Rhizome small, conical or slightly cylindrical, yellow in the center, white towards periphery, pleasantly aromatic. Leaves distichous, petioles longer than lamina, green or purple, tip acuminate or acute. Inflorescence both lateral and central, lateral inflorescence smaller, coma bracts bright pink or with shades. Corolla tube funnel shaped, yellowish-white, labellum 3-lobed, bright yellow.

C. zedoaria (Christm.) Roscoe (*C. raktakanta* Mangaly and Sabu)

Rhizome medium conical, greyish to yellow inside, whitish towards the periphery, aromatic, sessile tubers finger-shaped, branched and elongated. Leaves 4-6, distichous, inflorescence lateral; peduncle 20-25 cm, covered by reddish-purple sheaths. Flowers are equal in size or slightly smaller than the bracts, corolla tube funnel-shaped; lobes unequal, light pink. Labellum broad, margin wavy with a medium cleft, light yellow with a medium dark yellow band.

C. vamana M.Sabu and Mangaly

Rhizome small, conical, orange within, sessile tubers absent, roots several. Leafy shoot upto 50 cm tall. Leaves distichous, tip acuminate, base

subequal, closely pinnately nerved, inflorescence central, 15-22 cm long peduncle, concealed within the leaf sheaths, light green, glabrous. Flowers shorter than the bracts, corolla tube 8 mm long; lobes almost equal, glabrous, yellowish-white. Labellum tip notched, lobe rounded, margin crisped, golden yellow, hairy at the middle.

C. zanthorrhiza Roxb.

Rhizomes large, with camphor smell, yellow to deep yellow inside, paler on younger parts, short; roots with large tubers. Leaves distichous, petiolate, Inflorescence lateral; peduncle 15-25 cm, coma bracts fused only at base, dark pink. Flowers almost equal to the bracts. Corolla tube funnel-shaped, white with a pinkish tinge. Labellum short 3-lobed, middle lobe emarginate, pale yellow with a deep yellow band.

C. aeruginosa Roxb.

Known as pink and blue ginger, is a perennial tropical plant with unbranched leafy stems up to 70-100 cm tall and slightly aromatic, rhizomes are blue in colour in the centre, inflorescence lateral, pink or violet coma bracts, corolla tube pink, purple or reddish-brown patch along the sides.

C. amada Roxb.

Rhizome light yellow inside, white towards the periphery with the smell of green mango. Leaves much smaller, Inflorescence lateral or central. Coma bracts fused only at the base, spreading, light violet. Corolla tube funnel-shaped, 3 cm long, pale yellow, minutely pubescent; lobes unequal, white. Labellum somewhat elliptic 3-lobed, midlobe emarginate, recurved, pale yellow with a median dark yellow band, glabrous.

Table 3.1. Site of collection and voucher numbers provided for *Curcuma* spp. collected from different regions of South India.

S. No	Name of the species	Locality	Voucher Number
1	*C. aeruginosa* Roxb.	Malappuram	6943 (CALI)
2	*C. amada* Roxb.	Malappuram	6946 (CALI)
3	*C. aromatica* Salisb	Idukki	6941 (CALI)
4	*C. bhatii* (R.M. Sm.) Skornickova and M. Sabu	Udupi, Karnataka	6955 (CALI)
5	*C. coriacea* Mangaly and M. Sabu	Idukki	6954 (CALI)
6	*C. decipiens* Dalzell	Alathur	6936 (CALI)
7	*C. aurantiaca* Zijp (*C. ecalcarata*)	Nilambur	6938 (CALI)
8	*C. haritha* Mangaly and M.Sabu	Calicut	6950 (CALI)
9	*C. inodora* Blatt.	Karwar, Karnataka	6952 (CALI)
10	*C. karnatakensis* Amalraj, Velay. and Mural.	Udupi, Karnataka	6951 (CALI)
11	*C. kudagensis* Velay., V.S.Pillai & Amalraj	Thalakkavari, Coorg	6948 (CALI)
12	*C. longa* L.	Malappuram	6949 (CALI)
13	*C. montana* Roxb.	Mysore	6942 (CALI)
14	*C. mutabilis* Škornick., M.Sabu and Prasanthk.	Nilambur	6937 (CALI)
15	*C. neilgherrensis* Wight	Wayanad	6953 (CALI)
16	*C. oligantha* Trimen.	Kannur	6940 (CALI)
17	*C. pseudomontana* J.Graham	Attapadi	6939 (CALI)
18	*C. zedoaria* (Christm.) Roscoe (*C. raktakanta* Mangaly and Sabu)	Alwaye	6944 (CALI)
19	*C. vamana* M.Sabu and Mangaly	Peechi, Thrissur	6945 (CALI)
20	*C. zanthorrhiza* Roxb.	Calicut	6947 (CALI)

3.2 Isolation of genomic DNA

3.2.1 Materials required

Mortar and pestle, measuring cylinders (5-250 ml), beakers, conical flasks, variable volume pipettes (1-1000 µl), electronic balance, cooling centrifuges, Eppendorf tubes, scissors, ice *etc*.

Genomic DNA was isolated from fresh *Curcuma* leaves using modified CTAB method (Doyle and Doyle, 1987). The procedure used for obtaining pure DNA is as follows;

1. Approximately 150 mg of fresh young leaf tissues without midrib were cut and grind to a fine powder using liquid nitrogen (LN_2) in a pre-chilled mortar and pestle.

2. The powder was immediately transferred to 2.0 ml centrifuge tube containing 1 ml of pre-warmed (65°C) extraction buffer containing 1M Tris-HCl (pH 8.0), 0.5M EDTA (pH 8.0), 5M NaCl, 2% (w/v) Cetyltrimethyl ammonium bromide (CTAB) and 0.2% (v/v) β-mercaptoethanol. Added 1 µl (10 µg/µl) proteinase K. (For preparation see Table -3.2 and 3.3)

3. The powder was emulsified gently and incubated at 65°C for one hour with intermittent mixing by gentle swirling.

4. Added equal volume of chloroform: isoamyl alcohol (24:1) to the supernatant and centrifuged at 4°C for 12 min at 12,000 rpm.

5. The supernatant was extracted twice with equal volume of phenol: chloroform: isoamyl alcohol (25:24:1) followed by extraction with

chloroform: isoamyl alcohol (24:1) and centrifuged at 12,000 rpm 4°C for 12 min.

6. Added 2/3 volume of isopropanol and mixed by gentle inversions and kept at -20°C for 1 hour. The precipitated DNA was washed twice with 70% (v/v) ethanol. Centrifuged at 12000 rpm for 10 minutes, discarded the supernatant and the pellet was air dried for 20 min. The pellet was dissolved in minimum TE buffer containing 10mM Tris, pH 8.0 and 1.0mM EDTA. (If sufficient quantity of DNA was not obtained, the DNA was pelletted as described below).

7. Added isopropanol and incubated overnight at 4°C.

8. Centrifuged at 12,000 rpm for 15 minutes at 4°C.

9. Supernatant was discarded and 70% (v/v) ethanol was added to wash the pellets.

10. Centrifuged at 12,000 rpm for 15 min at 4°C. Poured off the supernatant, to drain off excess alcohol and left the pellet to air dry.

11. The dried pellet was dissolved in minimum quantity of TE buffer.

12. Stored at -20°C until further use.

3.2.2 Purification of DNA

1. The obtained DNA was subjected for RNase treatment. Added 1 µl (10 µg/µl) of DNase free RNase and incubated at 37°C for 1 hour.

2. Added equal volume of phenol: chloroform: isoamyl alcohol (25:24:1) and mixed well.

3. Spin at 12,000 rpm for 15 min at 4ºC and the aqueous layer was transferred to fresh centrifuge tubes.

4. Added equal volume of chloroform: isoamyl alcohol (24:1), spin at 12,000 rpm for15 min at 4ºC and transferred the aqueous phase to fresh tubes.

5. Added 1/10 volume of 3 M sodium acetate (pH 5.2) and 0.8 volume of isopropanol and inverted the tubes and incubated at 4ºC overnight.

6. The tubes were centrifuged at 12,000 rpm for 12 minutes at 4ºC.

7. Decanted the supernatant and washed the pellet using 70% (v/v) ethanol.

8. Air dried the pellet and dissolved in TE buffer.

Table 3.2. Preparation of buffers for genomic DNA isolation (Sambrook *et al.*, 1989)

	Buffer	**Method of preparation**	**Comments**
1	CTAB extraction buffer: 1 litre 1MTrisHCl (pH 8.0) 0.5mM EDTA (pH 8.0) 5 M NaCl 2% (w/v)CTAB (Himedia) 0.2% (v/v)β mercaptoethanol (Himedia)	Measure 100 ml Tris (1 M), 280 ml of NaCl, 40 ml of EDTA (0.5 M). Mix with about 400 ml of hot distilled water, added part by part 20g of CTAB to this. Adjust final volume to 1 litre. Dispensed to reagent bottles and autoclaved. Just before use, added 0.2% (v/v) β-mercaptoethanol.	CTAB will take time to dissolve. Avoid foaming.
2	TE (10:1mM) buffer 100 ml, 100mM Tris-HCl (pH 8), 1 mM	Take 1 ml of TrisHCl (1 M), 20 ml of EDTA (0.5 M). Mixed with 99 ml of sterile	TE (10mM) is written since there is TE with

	EDTA (pH 8)	distilled water taken in a reagent bottle, mixed thoroughly and autoclaved.	1 mM EDTA.
3	TAE buffer 50X :1 litre	Weighed 242 gm of Tris base; added 100 ml of EDTA (0.5 M); 57.1 ml of glacial acetic acid and around 0.50 ml of sterile distilled water. Dissolved the salt and adjusted volume to 1 litre. Autoclaved.	It will dissolve much easily in 500 ml solution.
4	Gel loading buffer(6X)100 ml 0.25%(w/v) bromophenol blue(Himedia) 30% (v/v) glycerol (Himedia)	Dissolved 0.25 gm of BPB in 99 ml of 30% (v/v) glycerol. Kept on a magnetic stirrer for several hours to get the dye completely dissolved. Dispensed to reagent bottles and kept at $4^{o}C$.	Strong dyes handle carefully.
5	Proteinase K – storage buffer. Glycerol (50 ml), 1 M Tris-HCl, pH 7.5(1 ml), $CaCl_2$ (0.29 g) dd H_2O to 100 ml.	Kept 100 ml storage buffer in a screw cap tube, add 100 mg of proteinase K, mix well and aliquot to 1.5 ml eppendorf tubes.	Store at $-20^{o}C$
6	RNase A	Prepare 10 mg per ml stock solution in 10 mM sodium acetate buffer, pH 5.2. Heat to 100°C for 15 minutes; allowed to cool at room temperature, and adjust the pH to 7.4 using 0.1 volume of 1 M Tris-HCl, pH 7.5.	Aliquot and store at $-20^{o}C$.

Table 3.3. Preparation of stock solution (Sambrook *et al*., 1989)

	Solutions	Method of preparation	Comments
1	Tris (pH 8.0), 500 ml	Dissolved 60.55gm Tris base (Himedia) in 300 ml distilled water. Adjusted the pH to 8 by adding conc.HCI. Made the volume to 500 ml .Dispensed to reagent bottles and sterilized by autoclaving.	The salt will take time to dissolve.
2	0.5M EDTA (pH 8.0), 100ml	Dissolved 18.61 g of EDTA – disodium salt (Himedia) in 100ml of water. Adjusted pH to 8.0 by adding NaOH pellets. Made the volume to 100 ml. Dispensed into reagent bottles and autoclaved.	pH of EDTA solution is temperature dependent. EDTA will completely dissolve only when pH becomes 8.
3	5M NaCl 500ml	Weighed 146.1g NaCl (Himedia) added 200ml of water and mixed well. When the salts gets completely dissolved, adjust the final volume to 500 ml. Dispensed into reagent bottles and autoclaved.	The salt will take much time to dissolve.
4	3M sodium acetate (pH 5.2), 100 ml	Dissolved 24.609g of anhydrous sodium acetate (Merk) in 100ml of water and mixed well. When dissolved completely adjusted the pH of the solution to 5.2 with glacial acetic acid (99-100%). Dispensed to reagent bottles and autoclaved.	The salt will take much time to dissolve.
5	Ethidium Bromide (100 mg/ml),100ml.	Added 1g Ethidium bromide to 100 ml of distilled water. Kept on magnetic stirrer to ensure that the dye has dissolved completely. Dispensed to amber coloured reagent bottle and stored at 4°C.	Ethidium Bromide is a powerful mutagen and is moderately toxic. So handle carefully.

6	70% (v/v) ethanol, 500ml	Take 355ml of ethanol: mix with 145ml of distilled water. Dispensed to reagent bottle and stored at 4°C.	Stock ethanol is 99% (v/v) hence 355 ml is taken instead of 350ml.
7	Chloroform: isomyl alcohol (24:1),500ml.	Measured 480 ml of chloroform and 20 ml of isomyl alcohol. Mixed well and stored in reagent bottle in room temperature.	Chloroform will evaporate so close the cap tightly and keep in amber coloured bottles.
8	1M $MgCl_2$100 ml.	Weighed 20.33g $MgCl_2$ (Himedia), dissolved in double distilled water, made up to 100ml, dispensed into reagent bottles.	

3.2.3 Quantification of DNA

The purified DNA was quantified using UV scanning Thermo scientific Nano Drop^{TM} 2000 spectrophotometer. The amount of DNA was calculated based on its OD at 260/280 nm and 0.8% (w/v) and the agarose gel was used for DNA visualization for qualitative analysis. The DNA samples were mixed with a sample buffer containing bromophenol blue and 3µl of the sample were loaded to each well. The gel was incorporated with 1µl of (10µg/µl) ethidium bromide. The molecular weight for the DNA was compared by using 1kb ladder. The gel was run using a horizontal electrophoretic unit containing 1x TAE buffer until the tracking dye reaches the bottom edge of the gel. The OD of the samples were measured at 280 nm and the 260/280 value for the DNA samples were calculated to assure their quality. The DNA quantity was calculated based on the assumption that 1OD of DNA=50ng of double stranded DNA.

3.2.4 PCR amplification

PCR was performed to amplify the genes using specific DNA and respective gene specific primers. PCR was carried out in an Eppendorf mastercycler pro S, Germany. The desalted custom DNA primers were

obtained from sigma Aldrich, Bangalore. PCR reaction mixture consists of Taq DNA polymerase, 10 X reaction buffers, dNTPs and magnesium chloride procured from Genei, Banglore, India. The PCR reaction mixture composition is shown in the Table 3.4 and the list of specific PCR primers used and the cycler conditions are given in Table 3.5. 1kb ladder and low melting agarose for gel electrophoresis were obtained from Hi-Media, India.

After the amplification and agarose gel electrophoresis, the gel documentation system (Gel Doc 1000 System, Bio-Rad) and UV photography were used to record the banding patterns with standard DNA ladders for finding molecular size of the amplified DNA on the agarose gel. The amplified PCR products were sequenced in Scigenome Lab Pvt Ltd (Cochin, Kerala) and submitted to GenBank, NCBI.

Table 3.4: The reaction mixture for PCR

Reagents	**Reaction Volume (25µL)**
10 x PCR buffer	2 µL
$MgCl_2$- 25mM	2 µL
dNTPs- 100mM	3 µL
Forward primer	1 µL
Reverse primer	1 µL
DNA template (~50ng)	0.5 µL
Taq DNA polymerase (5U/µL)	0.25 µL
Sterile double distilled water to make the final volume upto 25 µL	15.25 µL

3.2.5 Elution of DNA from the agarose gel

The elution was performed using a Mini Gel Purification Kit (ODP208-02, Origin, India) according to the manufacturer's instructions. All the centrifugation steps were carried out at 12,000 rpm (~13,400 × g) in a conventional table-top centrifuge 5424 (Eppendorf) at room temperature.

3.3 Primer designing

The Chloroplast, mitochondria and nuclear gene specific primers were designed using NCBI primer blast software (http://www.ncbi.nlm.nih.gov). Both, the forward and the reverse primers were provided from Integrated DNA Technologies (IDT Vision Scientific Ltd, UK).

Table 3.5: PCR Primer sequences and their reaction conditions

S. No	Gene	Primer Sequences (5' to 3')	Reaction conditions
1	*rbcL*	F –5' -TCTGTTACTAACATGTTTACTTC-3' R-5' -TCCCTCATTACGAGCTTGTACACA-3'	94ºC 2 min, 94ºC 15 sec 55ºC & 55.5ºC 30 sec, 72ºC 1 min 30 cycles, 72ºC 10 min
2	*mat*K	F- 5' - GAAGATAGATCTCGGCAAC-3' R- 5' - TTACATAAAAATGTATTCGCTC-3'	94ºC 2 min, 94ºC 15 sec 55.6ºC 30 sec, 72ºC 1 min 30 cycles, 72ºC 10 min
3	*psbK-psbI*	F – 5' -TTAGCATTTGTTTGGCAAG-3' R - 5' -AAAGTTTGAGAGTAAGCAT-3'	94ºC 2 min, 94ºC 15 sec 53ºC 30 sec, 72ºC 1 min 30 cycles, 72ºC 10 min
4	*atpF-atpH*	F - 5' -ACTCGCACACACTCCCTTTCC-3' R -5' - GCTTTTATGGAAGCTTTAACAAT-3'	94ºC 2 min, 94ºC 15 sec 58ºC & 59ºC 30 sec, 72ºC 1 min 30 cycles, 72 ºC 10 min
5	*trnlC-trnlF*	F - 5' -CGAAATCGGTAGACGCTACG-3' R - 5' -ATTTGAACTGGTGACACGAG-3'	94ºC 2 min, 94ºC 15 sec 59.8ºC 30 sec, 72ºC 1 min 30 cycles, 72 ºC 10 min

6	*psbA-trnH*	F - 5' -CTTGGTATGGAAGTAATGCA-3' R - 5' -ATCCACTTGGCTACATCCG-3'	94ºC 2 min, 94ºC 15 sec 58.6ºC & 59ºC 30 sec, 72 ºC 1 min 30 cycles, 72 ºC 10 min
7	*atp1*	F -5' -AATTCTCACCCAGAGCTGC-3' R -5' - GCATCTGGTTCCATCTTTCTTTC-3'	94ºC 2 min, 94ºC 15 sec 56.4ºC 30 sec, 72ºC 1 min 30 cycles, 72 ºC 10 min
8	*cox1*	F- 5' -GACCCGGCGATCAAATTCTTG-3' R- 5' -ACCGAAGAACCGAAAGAGATG-3'	94ºC 2 min, 94ºC 15 sec 57.4ºC & 59ºC 30 sec, 72ºC 1 min 30 cycles, 72 ºC 10 min
9	ITS1	F-5' -GGAAGGAGAAGTCGTAACAAGG-3' R-5' -GGCACAACTTGCGTTCAAAG-3'	94ºC 2 min, 94ºC 15 sec 61ºC & 63ºC 30 sec, 72ºC 1 min 30 cycles, 72 ºC 10 min

3.3.1 Sequencing of PCR amplicons

The chloroplast, mitochondrion and nuclear PCR amplicons of all the *Curcuma* spp. were sequenced using ABI 3730XL DNA Analyzer Scigenom Lab Pvt, Ltd (Cochin, Kerala) according to the Sanger dideoxy sequencing method. Chromatograms of the reverse and forward DNA sequences were analyzed using Bio-Edit.v.7.1.3 software (Ibis Biosciences, Carlsbad, CA 92008). All sequences were deposited in the NCBI GenBank database (http://www.ncbi.nlm.nih.gov) and as accession numbers was obtained.

3.3.2 Sequence alignment

The DNA sequences for *mat*K, *rbcL, psbK-psbI, trnlC-trnlF, psbA-trnH, atpF-atpH* and mitochondrial regions (*atp1* and *cox1*) nuclear internal transcribed spacer (*nr*ITS) region were retrieved from the databases of

GenBank. Sequence alignment was performed using Clustal W multiple sequence alignment program. Aligned sequences were represented as rows within a matrix. Phylogenetic tree was constructed using these multiple aligned sequences based on the maximum likelihood and maximum parsimony methods and assigned bootstrap values.

Apart from this, the following online bioinformatic tools were used,

CLUSTAL W (http://www.genome.jp/tools/clustalw)

Molecular Evolutionary Genetic Analysis (MEGA 6.0)

BioEdit Sequence Alignment Editor

Phylogenetic Analysis Using Parsimony (PAUP 4.0)

3.3.3 Molecular phylogenetic analysis

Sequence characteristics of all the genomic regions were calculated using Molecular Evolutionary Genetic Analysis (MEGA) version 6.0 (Tamura *et al*., 2013). Phylogenetic analysis using the chloroplast, mitochondrion and nuclear gene sequences were performed for further confirmation of the degree of relationship among the members of South Indian *Curcuma* species. The phylogenetic tree was predicted by aligning all the sequence data executing Maximum likelihood methods with MEGA 6.0 (Tamura *et al.,* 2013) and parsimony and distance methods as implemented in PAUP* 4.0 (Swofford, 1998). Maximum parsimony (MP) analysis were performed using heuristic search with MULPARS, tree-bisection-reconnection (TBR) branch swapping, and RANDOM stepwise addition with 1000 replicates. A total of 21 species was aligned in the clustalW programme using the MEGA 6.0. The best-fit nucleotide-substitution model was determined based on Bayesian Information Criterion (BIC) and Akaike Information Criterion (AIC) with the same software. The pairwise distance of the estimated values was determined using

the Maximum Composite Likelihood (MCL) approach, and the topology was selected with superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among the sites [5 categories (+G, parameter= 0.4961)]. Positions containing gaps and missing data were eliminated from all the datasets (complete deletion option). To estimate the levels of assurance in monophyletic clades, the bootstrap method with 1000 replications were employed (Felsenstein, 1985) in all analyses.

3.3.4 Basic Local Alignment Search Tool (BLAST)

BLAST is a commonly used method for the identification and homology search of sequences (www.ncbi.nlm.nih.gov/BLAST/) based on the best sequence alignment to all or only to a part of the query sequence (Holder and Lewis, 2003). It is a popular tool and quickly identifies the regions of local similarity between two sequences. It works based on statistical algorithms that can determine the alignment between two sequences. BLAST results are conflicting and variable sometimes in certain real life situations (Little and Stevenson, 2007). In gymnosperms BLAST can correctly identify only up to the genus level but not at the species level, because top hits are often not the closest phylogenetic relatives (Koski and Golding, 2001). Simple measures of genetic distance have commonly been used in barcoding studies to infer identity, with thresholds used to distinguish non-specifics from hetero specifics (Hebert *et al.,* 2003).

3.4 Extraction of RNA

3.4.1 Materials required

Mortar and Pestle, Eppendorf tubes, variable volume pipettes (1-1000 µl), measuring cylinders, beakers, conical flasks, electronic balance, cooling centrifuge *etc.*

3.4.2 Solutions and reagents

Extraction buffer: 0.3% SDS, 100 mM Tris-HCl, pH 8, 25 mM EDTA - pH 8. PVP (0.3%) in added fresh and β-mercaptoethanol (0.3% v/v)

0.1 % (v/v) DEPC treated autoclaved water

Acid phenol: chloroform

Sodium acetate (0.5 M; adjusted to pH 5 with glacial acetic acid)

Cold isopropanol

Cold 70% ethanol

3.4.3 Procedure

Total RNA from the *Curcuma* rhizomes (3^{rd}, 6^{th} and 9^{th} month's old) was isolated and extracted using modified SDS method (Deepa *et al.,* 2014). The following procedure was used for isolating pure RNA.

1. Grind 100 mg of rhizome tissue to a fine powder along with 2% PVP (powder form) in a mortar and pestle using liquid nitrogen.
2. Added 2 ml of pre-warmed (65°C) extraction buffer (2% SDS, 100mM Tris-HCl – pH 8, 25mM EDTA- pH 8 and 1% (v/v) β-mercaptoethanol), grind further and transferred the contents equally to 2 ml centrifuge tubes (Table 3.6).
3. Added equal volume of acid phenol: chloroform and vortex for 10 Sec. Leave it for 10 min at room temperature for the complete dissociation of nucleoprotein complexes. Centrifuged at 15,000g for 10 min at 4°C

4. Transferred the supernatant to fresh tubes and added 0.3 volume of 5 M sodium acetate and 0.7 volume of acid phenol: chloroform. Mixed by inverting the tubes and incubated in ice for 10 min.

5. Centrifuged at 15,000g for 10 min at 4°C and transferred the supernatant to fresh 1.5 ml centrifuge tubes.

6. Added 0.1 volume of 3 M sodium acetate and equal volume of isopropanol. Incubated the tubes at -20°C for 1 h.

7. Centrifuged at 15,000g for 10 min at 4°C and discarded the supernatant.

8. Washed the pellet with 70% ethanol by centrifuging at 7,500g for 5 min at 4°C.

9. Air dried the pellet for 5–10 mins and dissolved in 25 µl RNase free water.

Table 3.6: Preparation of stock solution

Buffer and stock solution	**Preparation**
Extraction buffer: 3% (w/v) SDS, 100 mM Tris-HCl (pH 8.0), 25 mM EDTA - pH 8.0 and β-mercaptoethanol (0-3%)	Dissolved 1.21g of Tris HCl in 25ml DDW and adjusted pH to 8.0. 0.09g of EDTA dissolved in 25 ml DDW adjusted pH to 8.0. Mixed the above two solutions and added 3g of SDS and made the final volume to 100 ml
Acid phenol:Chloroform (24:1)	Phenol crystals were melt at 60^{o}C Equal volume of liquid phenol and 1M Tris-HCl (pH 8.0) were mixed well and kept above a magnetic stirrer for 15 to 20 minutes Kept the mixture at room temperature, transferred the upper layer to a fresh bottle and check the pH to be 8.0.

	Mixed equal volume of acid phenol and chloroform. Stored the mixture in an amber coloured reagent bottle.
3M Sodium acetate	12.31 g of sodium acetate dissolved in 25 ml DDW and final volume was made up to 50 ml adjusted the pH to 5.0 with glacial acetic acid.
5M Sodium acetate	20.51 g of sodium acetate dissolved in 25 ml DDW and final volume made up to 50 ml, adjusted pH to 5 with add glacial acetic acid.

3.4.4 Purification of RNA

The obtained total RNA was purified using 1U DNase 1 (Thermo Scientific, Japan) to reduce the DNA contamination. The components of the reaction mixture are;

Components for DNase treatment

Components	Volume	
RNA	5 µl	The reaction mixture was kept in ice
10X Reaction Buffer with $MgCl_2$	1 µl	
DNase1, RNase-free	1 µl	
Double distilled water	3 µl	

The following procedure was used to purify RNA using DNase:

1. Incubated the 10 µl reaction mixture at 37^{o}C for 30 min.
2. Added 1.0 µl of 50 mM EDTA in the reaction mixture and mixture was incubated at 65^{o}C for 10 min.
3. Extracted the RNA with an equal volume of Phenol (saturated): Chloroform (1:1).

4. Centrifuged at 15,000g for 10 min at 4°C and transferred the supernatant to fresh 1.5 ml centrifuge tubes.
5. Added 0.1 volume of 3M sodium acetate and equal volume of isopropanol. Incubated at -20°C for 1 h.
6. Centrifuged at 15,000 g for 10 min at 4°C and discarded the supernatant.
7. Washed the pellet with 70% ethanol by centrifuging at 7,500g for 5 min at 4°C.
8. Air dried the pellet for 5–10 mins and dissolved in 25 µl RNase free water.
9. The purified RNA was used for RT-PCR and cDNA synthesis.

3.4.5 Quantification of RNA

The purified total RNA was quantified using UV scanning NanoDrop^{TM} 2000 spectrophotometer (Thermo Scientific, Germany). The ratio of absorbance at 260/280 and 260/230nm were recorded to assess the purity of RNA. The 260/280 value of ~ 2.0 justifies the purity of RNA. The ratio of absorbance at 260/230nm are often higher than the respective 260/280 values. The quantity of RNA was calculated based on the theory that 1 OD. at 260 nm for RNA molecules = 40 ng/µl of RNA. 1.5% (w/v) formaldehyde agarose gel was used to check the qualitative of RNA and these samples were stored at -80^{o}C until further use.

3.4.6 Preparation of FA gel for running RNA

1.5 g agarose was added to 10 ml 10X formaldehyde /agarose (FA) gel buffer (pH 7.0) containing 200 mM 3N-Morphilino propane sulfonic acid (MOPS), 50 mM sodium acetate and 10 mM EDTA and made up to 30 ml by adding RNase-free water. The agarose was melt and cooled to 65^{o}C and added 1.8 ml of 37% (w/v) formaldehyde and 2 µl of 10 µg/µl ethidium bromide, mixed gently (Table 3.7).

Table 3.7: Buffers and stocks used for FA/agarose denaturing gel.

Buffers and stocks	Preparations
10X FA gel buffer: 200mM MOPS , 50mM sodium acetate and 10mM EDTA	Dissolved 4.19g MOPS, 0.41 g sodium acetate and 0.37g EDTA in 80ml sterile DDW and made up the volume to 100ml pH 7.0 adjusted to using 1N NaOH
1X FA gel running buffer	10ml of 10X FA gel buffer and 2.0 ml of 37% formaldehyde was added in to 88 ml RNase free water
Ethidium bromide	10 mg of ethidium bromide dissolved in 1 ml sterile DDW.
Gel loading buffer (6x) – 100 ml	Dissolved 0.25 g of BPB in 99 ml of 30% glycerol. Keep on magnetic stirrer for several hours to get the dye completely dissolved. Dispense to reagent bottles and store at 40^{o}C.

3.4.7 RNA sample preparation and loading to gel

RNA samples (<200µg/µl) were prepared for electrophoresis by adding 2 µl of 10X FA gel buffer, 4 µl of 37% (w/v) formaldehyde and 10 µl of formamide. Reaction mixture was incubated at 65^{o}C for 10-15 min followed by cooling for 10 min and centrifuged for 5 sec. to settle the components to the bottom of the tubes. 10X FA gel loading buffer (2µl) containing 50% (v/v) glycerol, 10mM EDTA (pH 8.0), 0.25% (w/v) Bromophenol blue and 0.25% (w/v) Xylene cyanol was added to the sample mixture, cooled for a seconds and loaded to each well. The gel was run in a horizontal electrophoretic apparatus containing 1X FA gel running buffer at 5-7 V/cm until the tracking dye reaches the bottom edge of the gel.

3.4.8 cDNA synthesis

cDNA was prepared from total RNA using Takara PrimeScript^{TM} RT Reagent Kit **(**Cat.#RR037A, Takara Bio Inc., Japan) following the manufacturer's instructions. One microgram of purified total RNA was used

for first-strand complementary DNA (cDNA) synthesis with oligo dTprimer and random primers using PrimeScript RT enzyme mix.

A total of 20µl reaction components were used to synthesize first strand cDNA, components of the reaction mixture (20µl) are summarized in Table 3.8.

1. Prepared the following reaction mixture in 200 µl microtubes and kept on ice.

Table 3.8: Components of first strand cDNA reaction mixture

Reagent	**Volume**
5X PrimeScript Buffer	2.0 µl
PrimeScript RT enzyme mix	0.5 µl
Oligo dT Primer (50 µM)	0.5 µl
Random 6 mers (100 µM)	2.0 µl
Template RNA	10 µl
The volume was adjusted to 20 µl by adding RNase-free water.	

2. The incubation condition for the reaction mixture used in PCR was:-

Reverse transcription at 37^{o}C for 15 min.
Inactivation of Reverse transcription at 85^{o}C for 5 sec.
Hold at 4^{o}C

3.4.9 Quantification of cDNA

The synthesized cDNA quality was assessed using 260/280 and 260/230 ratio using a Nanodrop2000 spectrophotometer (Thermo Scientific, Germany). The cDNA samples were run on a 1% (w/v) agarose gel to visualize the bands.

3.4.10 Gene cloning and primer designing

A thermal cycler with a 96 well with gradient block (Eppendorf Mastercycler Pro S, Germany) was used for RT-PCR reactions. The desalted custom primers for RT-PCR and 1kb DNA ladder were obtained from Sigma-Aldrich, USA. One step RT-PCR kit was procured from Takara, Japan.

Good quality RNA was used as the template for one step RT-PCR reactions to amplify the curcumin synthase genes using the designed CURS (CURS1, CURS2 and CURS3) primers. Primers for cloning CURS genes were designed using Multalin and Primer-BLAST software of NCBI (Table 3.9). CURS (CURS1, CURS2 and CURS3) genes were designed using *C. zedoaria, C. longa* and *C. amada* genes retrieved from GenBank (Accession Nos: CURS1: MF402846.1, KM880189.1 and AB495007.1; CURS2: LC064068.1, AB506762.1, KF980982.1 and KF980981.1; CURS3: KX154461.1, MF987835.1, AB506763.1 and KM880190.1) was used as the template for primer designing.

Table 3.9: Primers used for RT-PCR

Name	Sequence	Tm (°C)	PCR product length (bp)
CURS1-F	5'- ATGGTGAAGAAGCGGTACCTG-3'	65.1	900
CURS1-R	5'- TGTTGCCGTACTCTGTGAAGA-3'	63.5	
CURS2-F	5'- GCTAATCAGTCAATCCAGATGG-3'	62.4	1100
CURS2-R	5'- CGTCTATCGATTGATCGATCGT-3'	65.0	
CURS3-F	5'-GTCAACCGCCTCATGCTCTACA-3'	68.5	590
CURS3-R	5'-TCACCTCGTCCATCACGAAGTAC-3'	67.3	

Desalted primers were diluted using sterile distilled water to a final concentration of 100 µM, working solutions were prepared from these stocks and stored at -20^{o}C for future use. PCR were carried out in a Mastercycler Pro S (Eppendorf, Germany).

RT-PCR reaction containing a final volume of 25µl was used for the amplification of the specific gene of interest. Components of the reaction mixture (25µl) are summarized in Table 3.10.

Table 3.10: The reaction mixture for PCR contains

Reagents	**Reaction Volume (25µL)**
10 x PCR buffer	2 µL
$MgCl_2$- 25mM	2 µL
dNTPs- 100mM	3 µL
Forward primer	1 µL
Reverse primer	1 µL
cDNA template (~50ng)	0.5 µL
Taq DNA polymerase (5U/µL)	0.25 µL
The final volume was adjusted to 25 µl with sterile double distilled water	

PCR cycles for different sets of primers

1st Cycle: Denaturation at 95°C for 15 min.

2nd Cycle: Denaturation at 95°C for 20s

Annealing temperature at (51.5, 52.1, 53.4, 54.0, 55.4 and 55.9°C) for 40 s

35-40 Cycles

3rd Cycle: Extension at 72°C for 1 min

4th Cycle: Final extension at 72°C for 10 min.

3.4.11 Visualization of PCR products

PCR products were visualized using 1% (w/v) agarose gel and molecular weight was detected using 1kb DNA ladder. The amplified band

corresponding to the product size of the primer was eluted and purified using MiniGel DNA Purification Kit Spin Column (ODP208-02) (Origin, India).

3.4.12 Elution of DNA from agarose gel

The elution was performed using MinElute Gel Extraction Kit (Origin, USA) according to the manufacturer's instructions. All centrifugation steps were carried out at 12,000 rpm (~13,400 × g) in a conventional table-top Centrifuge 5424 (Eppendorf) at room temperature.

3.4.13 DNA sequencing

Purified PCR products were sequenced from Scigenom Labs Private Ltd, India on charge basis. Homology of the cloned sequences was checked using BLAST. The cloned sequences using CURS genes were submitted to NCBI.

3.5 Primer designing for Quantitative Real-Time PCR

3.5.1 Gene specific CURS primers

The gene specific primer pairs were designed from the cloned sequences submitted in NCBI. Primers, with the GC content (40%, 60%), length (18 to 25 bases) and Tm (60 ± 1°C) were selected. Primers were designed using Primer 3.0 (http://primer3.ut.ee/) by using the cloned CURS genes sequence (Accession No. CURS1: MK515083, CURS2:MG386670 and CURS3: MK511334), cloned different *Curcuma* species.

3.5.2 Actin (Internal control) primer

Actin was used as the internal control and the primers were designed using Multalin and Primer-BLAST software of NCBI. The *Arabidopsis thaliana* actin sequences are available at NCBI (NM_001338359.1, NM_001338358.1, XM_021012169.1, U41998.1 and NM_112764.4) (Table

3.11). This primer from the actin sequence was used to clone the actin gene from *Curcuma* species and the cloned gene fragment was sequenced and submitted to NCBI.

A novel actin primer was designed using the newly cloned *Curcuma* specific actin sequences (MG751913). The GC content (40%, 60%), length (18 to 25 bases) and Tm (60 ± 1°C) for the primers were selected. Primers were designed using Primer 3.0 (http://primer3.ut.ee/) by utilizing developed actin gene sequence (Accession No. MG751913). cDNA primers and actin primers were obtained from Sigma (Table 3.12) and diluted to 100 µM concentrations using sterile double distilled water and the aliquots were prepared from these stocks and stored at -20^{o}C.

Table 3.11: Actin primer

Name	Sequence	Tm (°C)	GC%	PCR product length (bp)
Actin-F	5'-CTGATTACCTCATGAAGATC-3'	53.5	40	450
Actin-R	5'-ATCCAGACACTGTACTTCCT-3'	55.1	45	

Table 3.12: Primers synthesized for qPCR analysis

Name	Sequence	Size (bp)	Tm (°C)	GC%	Reference
ACT-F	5'-GAGAAGCTGGCCTACATTGC-3'	158	63.8	55	This study
ACT-R	5'-ATGGATGGTTGGAAAAGCAC-3'		63.6	45	
CURS1-F	5' -AGATCATGAGGCGGTTGACG-3'	199	67.5	55	This study
CURS1-R	5' -GACATCGTGGTGGAGGAGAT-3'		66.1	55	
CURS2-F	5' -GGGATCAAGGACTGGAACAA -3'	184	63.9	50	This study
CURS2-R	5' -GTTTCCTGAGCTCGTCCATC-3'		63.8	55	
CURS3-F	5' -TGATCGCGGAAAACATCGAG-3'	189	64.4	50	This study
CURS3-R	5' -GGCTGGACCTTCCACTTCTT-3'		66.0	55	

3.6 Quantitative Real-time PCR/qRT-PCR assay

3.6.1 Reaction conditions for Sso Advanced Universal SYBR Green Supermix 172-5270

Quantitative RT-PCR (qRT-PCR) reactions were carried out in a total volume of 10µl reaction mixture; the reaction components are summarized in Table 3.13.

Table 3.13: Components of qRT-PCR reaction

Components	Volume per 10µl	Final concentration
Supermix	5 µl	1X
Forward primer Reverse primer	Variable	200-500nM each
Template	Variable	cDNA: 120ng
Nuclease free water	Variable	
Total reaction volume	Variable	

qRT-PCR conditions:-

The reaction condition of qRT-PCR was:

1st Step: Initial hold at 95°C for 30 s.

2nd Step: Denaturation at 95ºC for 15s

3rd Step: Annealing at 55.6ºC for 45s in CURS1, 54.5ºC for 45s in CURS2 and 55.1ºC for 45s in CURS3

4th Step: Extension at 72 ºC for 30s

} 39 Cycles

Go to Step 2

5th Step Melting at 95ºC for 10s

6th Step Melt curve at 65ºC to 95ºC, increment 0.5ºC for 5s

3.6.2 Quantitative analysis of putative CURS genes: Livak method

All the reactions were performed in triplicate in a 96-well reaction plate using CFX96 Touch™ Real-Time PCR Detection System (Bio-Rad, USA). A No-template control was included for each gene to detect any background signals arising from the amplification of DNA contamination or primer dimer formation during the reaction. The relative quantity of mRNA from three, six and nine months of *C. longa, C. aromatica, C. zedoaria* and *C. zanthorrhiza* was recorded based on the method of Livak and Schmittgen (2001). The housekeeping gene, actin from *Curcuma* was used as internal control (reference gene).The threshold cycle (C_T) value of the internal control was subtracted from that of the gene of interest (CURS) to obtain a ΔC_T value. The mean C_T value of the samples was calculated from the triplicate reactions. The expression level of CURS genes was quantified using the following method.

$$\Delta C_{T\,(Sample)} = C_{T\,(Sample)} - C_{T\,(Actin)}$$

$$\Delta C_{T\,(Control)} = C_{T\,(Control)} - C_{T\,(Actin)}$$

$$\Delta\Delta C_{T\,(Sample)} = \Delta\Delta C_{T\,(Sample)} - \Delta\Delta C_{T\,(Control)}$$

$$\text{Relative expression level of target gene} = 2^{-\Delta\Delta C_T}$$

3.6.3 Amplification and melting curve analysis

Amplification plot and melting curve of each sample were obtained from the software used in qRT-PCR reaction. In the amplification plot, C_T values indicate the concentration of the target gene in the samples. The melting curve showed the plotting of -dF/dT against the temperature. The single peak of the melting curve indicates the pure and specific PCR product in the reaction mixture. Any genomic DNA contaminations, primer-dimer formations and the presence of splice variations in the reaction can be detected form the melting curve.

3.7. Extraction of active ingredients from rhizome for GC-MS and UFLC

3.7.1 Plant materials

The rhizomes were harvested after 10 months for GC-MS and 3rd, 6th, 9th months for UFLC of cultivation and used for extraction. All chemicals used in this study were of analytical grade purchased from Hi-Media laboratories, Mumbai, India and from *Sigma*- Aldrich, USA.

3.7.2 Preparation of the rhizome extract

The fresh rhizomes were thoroughly washed using distilled water and cut into small pieces and shade dried and pulverized into a powder using a mortar and pestle. The powder was tightly packed in polythene bags and stored at 4°C temperature. 30 g of the powder were dissolved in 50 ml of methanol and kept for 48 hours and filtered through Whatman No1 filter paper. The filtered extract was concentrated using rotary evaporator and the extract and solvent were separated. The crude extract was kept in a refrigerator at 4°C for further phytochemical analysis. The methanolic fractions of each species were used for GC-MS analysis.

3.7.3 GC-MS analysis of methanolic extracts

The analysis of the rhizome extracts were carried out using a gas chromatography- mass spectrometry (GC–MS) instrument (Shimadzu GC-MS Model Number: QP2010SE) equipped with a Rxi-5Sil MS capillary column (0.25 mm×30m, 0.25μm film thickness) (Shimadzu, Kyoto Prefecture, Japan). The operating conditions were as follows; the GC temperature program was held at 80°C for 2 min, then increased at 4.25°C/min from 80°C to 260°C and held at 280°C for 2 min. The injector temperature was 260°C; the sample injection volume was 1 μl; and the split ratio was 100:1. Mass spectrometry conditions were as follows; electron impact as the ion source,

ionizing voltage 70 eV, source temperature 150^{o}C, electron multiplier at 2000 eV, scan speed 1000amu/s and scan range 50–500 amu. The constituents of methanolic extracts were identified by comparing retention indices and matching the recorded mass spectra of each compound with NIST Chemistry web book and literature published (Gounder *et al.,* 2012).

3.7.4 Cluster analysis

The fractions obtained from the rhizome extracts were scored for the presence or absence of specific components and the data were analyzed using NTSYS-pc version 2.1 software. Cluster analysis was carried out with UPGMA (Unweighted Pair Group Method with Arithmetic averages) for the composition of 20 methanolic extracts to determine the pattern of identical and diverse chemical compounds between different species.

3.8 UFLC

3.8.1 Preparation of the rhizome extract

A modified method was used for the extraction of active ingredients from the rhizomes. The fresh rhizomes were thoroughly washed using distilled water and cut into small pieces and shade dried for 10 days and pulverized into a fine powder using a mortar and pestle. The powder was tightly packed in polythene bags and stored at 4°C. 30 g of the powder was dissolved in 50 ml methanol in an amber coloured flask, kept for 48 hours on a rotary shaker (198 rpm) and filtered initially using muslin cloth followed by Whatman filter paper. The filtered extract was concentrated keeping in a water bath at 64^{o}C and the free methanol was evaporated. The crude extract was kept in a refrigerator at 4^{o}C until UFLC analysis.

3.8.2 UFLC condition

UFLC analysis was performed using UFLC system (Shimadzu, Kyoto, Japan) containing LC-20AD prominence liquid chromatography pump, PDA-M20A (200-800 nm) detector. The elution was carried out in C_{18} column with gradient solvent systems having a flow rate of 1.0 ml/min at ambient temperature. The mobile phase consisted of water +0.2% phosphoric acid (A) and acetonitrile (B), levels of curcuminoids were determined using the above solvents programmed linearly from 40 to 60% acetonitrile for 0-15 min. The compounds were quantified using LC-solution software. The analyte detection was done using UV-VIS detector at λ 421 nm. The injection volume was 20 µL and running time was 30 minutes.

3.8.3 Standard preparation

The curcumin analytical standard (≥98.0% HPLC) was obtained from Sigma-Aldrich. The methanolic stock solutions of curcumin were prepared at a concentration of 1 mg/mL.

3.8.4 Method Validation

Validation of the method for the determination of curcumin (CUR), demethoxycurcumin (DMC) and bisdemethoxycurcumin (BDMC) were performed by following the guidelines of single-laboratory validation from the Association of Official Analytical Chemists (AOAC, 2016). The method validation parameters were the stability of the standard and sample, specificity, the linearity of the calibration curves, precision, accuracy, limit of detection (LOD) and limit of quantification (LOQ).

3.8.5 Solution stability

Standard and sample solutions were kept at room temperature for 24 hours. Assay percentage was calculated at zero hour and after every 4 hour

intervals, the assay percentage was compared with the initial time period. The change in the assay percentage was evaluated and percentage RSD was calculated to ascertain the stability of the test solution.

3.8.6 Specificity

Specificity is another analytical method used to measure accurately and specifically the presence of the component that is expected to be present in the sample matrix. Chromatograms of standard and sample solutions of the curcuminoids were compared and the peak purity spectra obtained using photodiode array detector (PDA) was recorded in order to ascertain the specificity of the method.

3.8.7 Linearity

Curcumin stock solution was diluted to five different concentrations using HPLC grade methanol to obtain 1, 10, 100, 200 and 300µg/ml concentrations. Each concentration was performed in triplicate and the mean value was calculated. A 0.22mm membrane filter was used to filter the standard solution and injected into the UFLC instrument. The linear equation and R^2 were calculated.

3.8.9 Precision

The intra-day and inter-day precisions of the proposed method was evaluated by measuring the corresponding responses six times on the same day and subsequently three times for three different days for the test mixture of three different concentrations of curcuminiods (1,3,5µg/ml). The results were evaluated to obtain relative standard deviation.

3.8.10 Accuracy (% Recovery)

The accuracy of the method was determined by calculating the recovery rate of CUR, DMC and BDMC by the standard addition method. Known quantity of standard CUR, DMC and BDMC (1, 2, 3µg/ml) were added to a pre-quantified sample solution of CUR (2µg/ml), DMC (4µg/ml) and BDMC (5µg/ml). Each solution was injected in triplicate and the percentage recovery was calculated by measuring the peak areas and fitting these values into the regression equation of the respective calibration curves.

3.8.11 Limit of detection and limit of quantification

The limit of detection (LOD) and the limit of quantification (LOQ) were calculated using the standard deviation of y-intercept of the calibration curve from linear lower concentrations (N) and slope (S) of the calibration curve. LOD and LOQ values are calculated using the equation;

$LOD = 3.3 \times N/S$, $LOQ = 10 \times N/S$

3.9 *In silico* modeling and structure prediction of curcumin synthase (CURS1, CURS2 and CURS3) using bioinformatics tools

3.9.1 Materials

The following materials have been used to achieve this objective.

3.9.2 Databases

- NCBI
- PDB

3.9.3 Tools

- BLAST
- PSI-BLAST

- PSIPRED
- SOPMA
- Protparam
- SWISS MODEL

3.9.4 DATABASES USED FOR PROTEIN SECONDARY STRUCTURE PREDICTION

3.9.4.1 NCBI (www.ncbi.nlm.nih.gov/)

The NCBI (National Center for Biotechnology Information) is an organization for collection of data of nucleotide sequences and it's a part of the U.S National Library of Medicine (NLM) and the branch of National Institutes of Health. NCBI houses of genome sequencing data in GenBank and an index of biomedical research articles such as PubMed Central, PubMed and other information relevant to biotechnology. All the information of the databases in NCBI is available online through the Entrez computer programme.

3.9.4.2 Protein Data Bank (PDB) (http://www.pdb.org/)

Protein databank is a repository of three dimensional structures of biological macromolecules like DNA, RNA, Proteins and Protein-Protein complexes. Any three dimensional structure can be downloaded from PDB using a PDB ID which is a four letter code of alpha numeric in world wide. The stored protein structure data are typically obtained by X-ray crystallography, NMR spectroscopy, or, increasingly, cryo-electron microscopy methods (Rose *et al.,* 2016). The stored structure was submitted by biologists and biochemists from around the world are freely accessible on

the Internet through the websites of its member organizations (PDBe, PDBj, RCSB and BMRB).

3.9.4.3 PSIPRED (Psi-blast based secondary structure prediction)

PSIPRED is a technique used to examine the secondary structure of particular protein and it employs to a network for machine learning methods in its algorithm. It is a server-side program, featuring a website serving as a front-end interface, which can predict a protein's secondary structure (beta sheets, alpha helices and coils), from the primary sequence (McGuffin *et al.,* 2000).

3.9.4.4 SOPMA (Self Optimized Prediction Method from Alignment)

SOPMA is a tool used to predict secondary structure of proteins and these methods are based on the homology method developed by Levin *et al.,* (1993). The method utilizes the information from an alignment of sequences belonging to the same family. If there are no homologous sequences, the SOPMA prediction is the SOPM one. The parallel method called the self-optimized prediction method (SOPM) developed to improve the success rate in the prediction of the secondary structure of proteins. It predicts all the sequences of a set of aligned proteins belonging to the same family. This improved SOPM method (SOPMA), correctly predicts 69.5% of amino acids for a three-state description of the secondary structure (alpha-helix, beta-sheet and coil) in a whole database containing 126 chains of non-homologous (less than 25% identity) proteins.

3.9.4.5 PROTPARAM

PROTPARAM is a tool working with ExPasy server to analyze the protein sequences of various physical and chemical parameters for a given protein stored in Swiss-Prot or TrEMBL or for a user entered sequence. The

physicochemical properties such as molecular weight, theoretical isoelectric point, aliphatic index and grand average hydropathy (GRAVY) were calculated (Gasteiger *et al.,* 2005).

3.9.4.6 PSI BLAST

BLAST (Basic Local Alignment Search Tool) is a sequence similarity search method, in which a query protein or nucleotide sequence is compared to nucleotide or protein sequences in a target database to identify regions of local alignment and report those alignments that score above a given score threshold. Position-Specific Iterative (PSI)-BLAST is a protein sequence profile search method that builds off the alignments generated by a run of the BLASTp program. The first iteration of a PSI-BLAST search is identical to a run of BLASTp program (Altschul *et al.,* 1997).

3.9.4.7 SWISS MODEL

SWISS-MODEL (http://swissmodel.expasy.org) is a server for automated comparative modeling of three-dimensional (3D) protein structures. SWISS-MODEL repository is integrated with several external resources, such as UniProt, InterPro, STRING, and Nature PSI SBKB (Schwede *et al.,* 2003).

SWISS-MODEL channel work in four main steps that are involved in building a homology model of a given a particular protein structure;

1. Identification of structural template(s). BLAST and HHblits are used to identify templates. The templates are stored in the SWISS-MODEL Template Library (SMTL), which is derived from PDB.
2. Alignment of target sequence and template structure(s).
3. Model building and energy minimization. SWISS-MODEL tools are rigid fragment assembly approach for modelling.

4. Assessment of the model's quality using QMEAN, a statistical potential of mean force.

3.9.5 CURS protein sequence analysis

The cloned CURS nucleotide sequences were translated using ORF finder program (https://www.ncbi.nlm.nih.gov/orffinder/) predicted corresponding amino acid sequences. The homology was analyzed using PSI-BLAST and SMART BLAST. The primary structure of CURS1, CURS2 and CURS3 proteins was studied using ProtParam tool of Expasy Proteomic Server (http://expasy.org/cgibin/protparam).

3.9.6 Comparative Modeling for structure prediction

The derived CURS proteins sequence was used as a query sequence for comparative modeling. SWISS-MODEL (http://swissmodel.expasy.org) was used for the 3D structure prediction of the CURS1, CURS2 and CURS3 and its integrated with several external resources, such as UniProt, InterPro, STRING, and Nature PSI SBKB. SWISS–MODEL server by a combination of methods comprising RCSB , PDBe, PDBj, PDBsum , CATH and PLIP. CATH was used to understand the class and homology of an unknown protein samples (Orengo *et al.,* 1997).

3.9.7 Model Validation

Different tools were used to evaluate the internal consistency and reliability of the modeled structure of the CURS1, CURS2 and CURS3. MolProbity was used to assess the stereochemical quality of the model by quantifying the residues in the allowed zones of Ramachandran plot (Chen *et al.,* 2010). The obtained protein structure was re-assessed for its reliability and model quality using QMEAN-Z scores from QMEAN server, http://swissmodel.expasy.org/docs/structure_assessment (Arnold *et al.,* 2006).

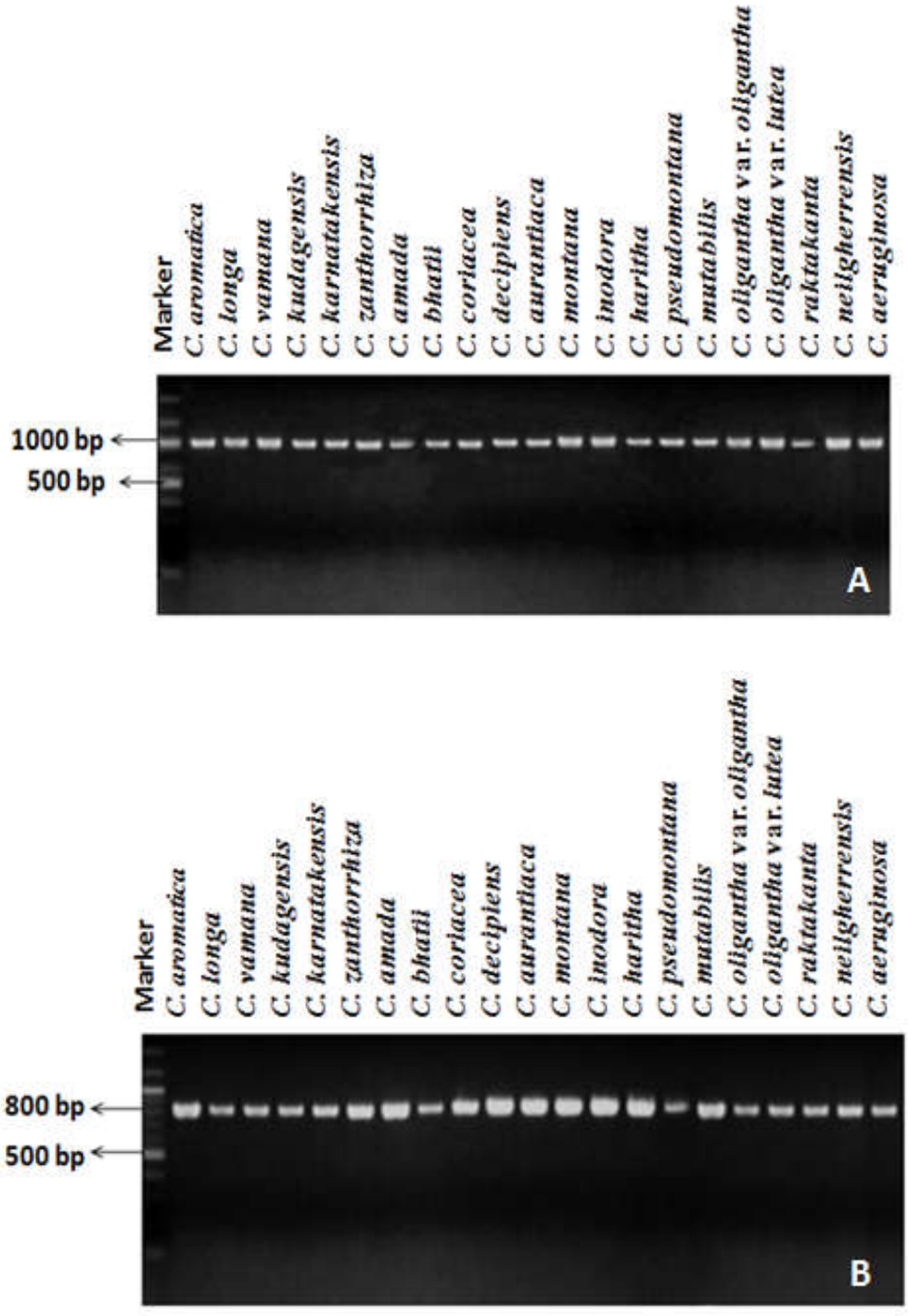
CHAPTER 4

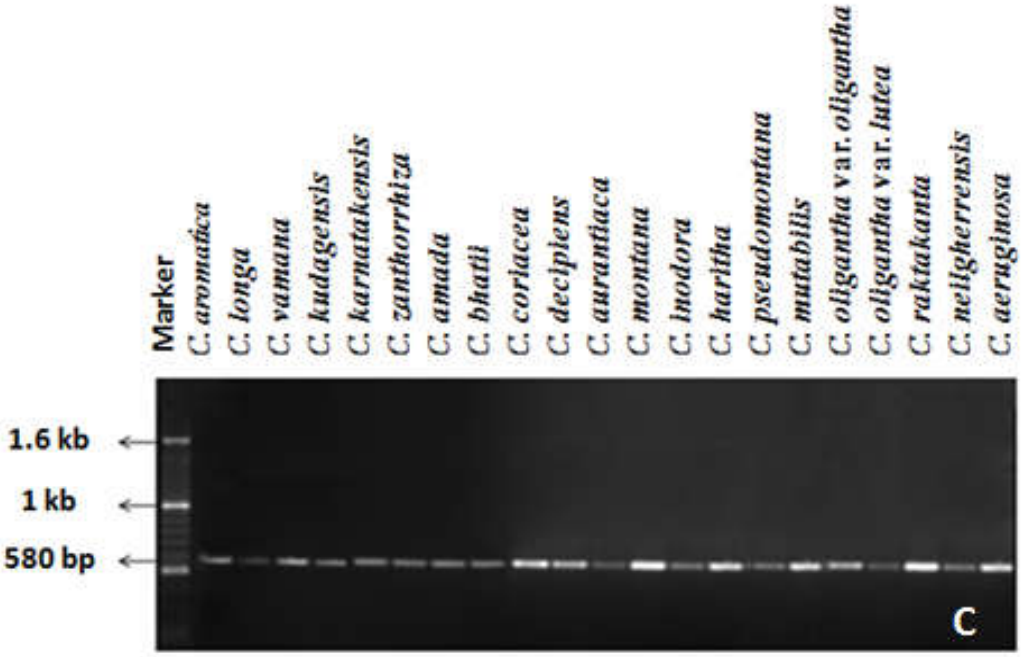
PHYLOGENETIC ANALYSIS USING MOLECULAR MARKERS

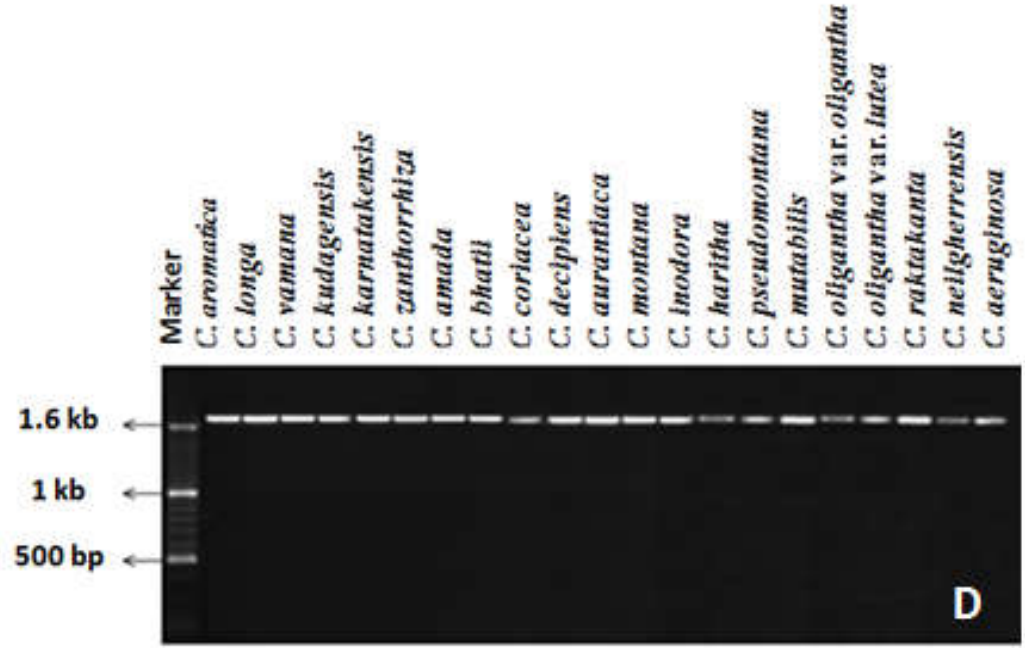
4.1 Molecular phylogenetic analysis of *Curcuma* based on chloroplast *mat*K, *rbcL*, *psbK-psbI*, *trnlC-trnlF*, *psbA-trnH*, *atpF-atpH* sequences.

4.1.1 Sequence characteristics of chloroplast regions

Agarose gel electrophoresis of the PCR products of the *Curcuma* species using chloroplast sequences showed distinct bands of *rbcL* ~1000 bp, *mat*K ~800 bp, *psbK-psbI* ~580 bp, *trnlC-trnlF*~ 1600 bp, *psbA-trnH*~800 bp and *atpF-atpH* ~750 bp respectively with a reliable amplicon represented in Fig. 4.1A-F**.** Amplified PCR products were sequenced by automated DNA sequencing. The obtained chromatograms of the sequences were analyzed by using Bio-Edit.v.7.1.3 software (Ibis Biosciences, Carlsbad, CA 92008). The obtained forward and reverse strand sequences were aligned and edited. Sequence homology was detected using BLAST homology search tools. All the cloned sequences were deposited in the NCBI GenBank database (http://www.ncbi.nim.nih.gov) and the provided accession numbers (Table 4.1).


Marker
C. aromatica
C. longa
C. vamana
C. kudagensis
C. karnatakensis
C. zanthorrhiza
C. amada
C. bhatii
C. coriacea
C. decipiens
C. aurantiaca
C. montana
C. inodora
C. haritha
C. pseudomontana
C. mutabilis
C. oligantha var. oligantha
C. oligantha var. lutea
C. raktakanta
C. neilgherrensis
C. aeruginosa
1000 bp
500 bp
A
Marker
C. aromatica
C. longa
C. vamana
C. kudagensis
C. karnatakensis
C. zanthorrhiza
C. amada
C. bhatii
C. coriacea
C. decipiens
C. aurantiaca
C. montana
C. inodora
C. haritha
C. pseudomontana
C. mutabilis
C. oligantha var. oligantha
C. oligantha var. lutea
C. raktakanta
C. neilgherrensis
C. aeruginosa
800 bp
500 bp
B


Marker
C. aromatica
C. longa
C. vamana
C. kudagensis
C. karnatakensis
C. zanthorrhiza
C. amada
C. bhatii
C. coriacea
C. decipiens
C. aurantiaca
C. montana
C. inodora
C. haritha
C. pseudomontana
C. mutabilis
C. oligantha var. oligantha
C. oligantha var. lutea
C. raktakanta
C. neilgherrensis
C. aeruginosa
1.6 kb
1 kb
580 bp
C


Marker
C. aromatica
C. longa
C. vamana
C. kudagensis
C. karnatakensis
C. zanthorrhiza
C. amada
C. bhatii
C. coriacea
C. decipiens
C. aurantiaca
C. montana
C. inodora
C. haritha
C. pseudomontana
C. mutabilis
C. oligantha var. oligantha
C. oligantha var. lutea
C. raktakanta
C. neilgherrensis
C. aeruginosa
1.6 kb
1 kb
500 bp
D

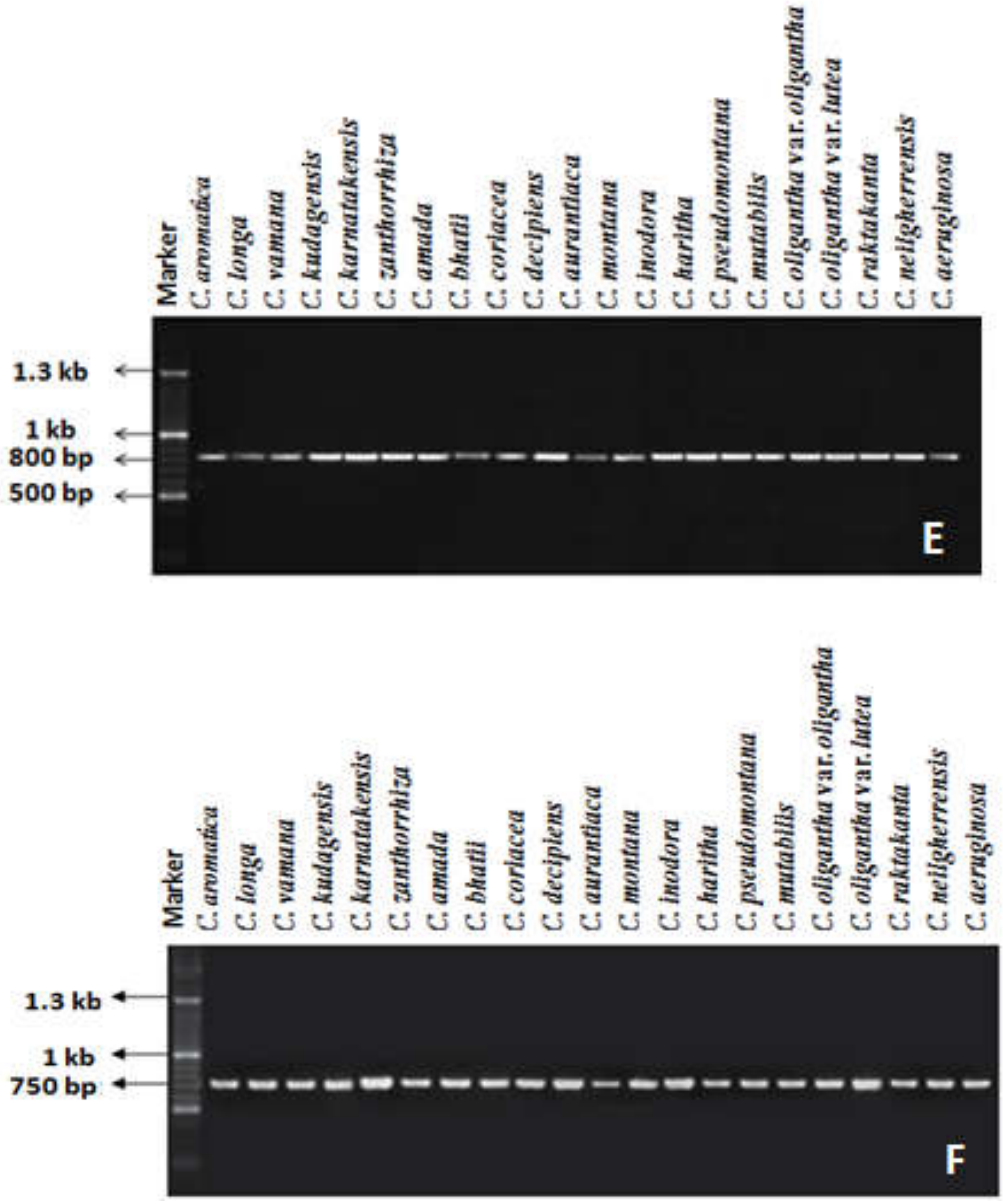


Fig. 4.1. Gel-electrophoresis of PCR product of *Curcuma* species shows a band of **A-** *rbcL* ~ 1000 bases, **B-** *mat*K~ 800 bases, **C-** *psbK-psbI*~ 580 bases, **D-** *trnlC-trnlF*~1600 bases, **E-** psbA-trnH~ 800 bases and **F-** *atpF-atpH* ~ 750with 750 bases.

Homology alignment of the six genomic regions resulted in a combined dataset with the following base pairs *rbcL*~966 bp*, trnLF*~1592 bp, *atpF*~761 bp, *psbK*~569 bp*, psbA-trnH*~921 bp and *mat*K~800 bp. Out of

5619 character sites, 2773 were constant, 1080 variable but parsimony uninformative and 1766 parsimony informative. Maximum parsimony analysis of the combined dataset yielded >300.0 maximally parsimonious trees with 3832 steps, 0.79 consistency index (CI), and 0.94 retention index (RI). Agreeability between all the data sets was assessed using the incongruence length difference (ILD) test and found high level of congruence between the six data sets (p=0.09) and the data was deemed combinable. The maximum likelihood (ML) tree and strict consensus most maximum parsimonious (MP) tree represented in Fig. 4.2 & 4.3.

Table 4.1. GenBank accession numbers for *mat*K, *rbcL, psbK-psbI*, *trnlC-trnlF, psbA-trnH, atpF-atpH* sequences of the twenty one species of *Curcuma*.

Name of the Species	*rbcL*	*mat*K	*trnlC-F*	*AtpF-atpH*	*psbK-psbI*	*psbA-trnH*
C. coriacea	KX608612	KU934093	KY303751	KY430376	KY818318	KY978409
C. karnatakensis	KX608613	KU736742	KY303765	KY463993	KY784120	KY963941
C. bhatii	KU697332	KX170829	KY303746	KY463991	KY784118	KY963942
C. zedoaria	KX608610	KX455852	KY303760	KY485935	KY614254	KY978411
C. oligantha var *oligantha*	KX650824	KX455853	KY303762	KY485940	KY818320	KY963940
C.aeruginosa	KX608611	KX455854	KY303752	KY303942	KY784121	KY847871
C. vamana	KX608615	KX455855	KY303761	KY464000	KY594917	KY851762
C. haritha	KX608606	KX148521	KY303747	KY485939	KY784115	KY963938
C. oligantha var *lutea*	KX608609	KX418654	KY200904	KY225981	KY818321	KY978410
C. amada	KX608605	KX650813	KX893883	KY430374	KY303943	KY430377
C.decipiens	KX608618	KX650814	KY303763	KY463999	KY818319	KY978412
C. kudagensis	KU886554	KX650815	KY303757	KY463997	KY784116	KY963939
C. aromatica	KX650825	KX650816	KY303748	KY430375	KY496973	KY847872
C. mutabilis	KX608607	KX650817	KY303756	KY485937	KY784119	KY978413
C. pseudomontana	KX608616	KX650818	KY303753	KY463995	KY614253	KY614253
C. neilgherrensis	KX608608	KX650819	KY303759	KY485934	KY614252	KY851761
C. longa	KX608614	KX650820	KY303754	KY485938	KY614251	KY851760
C. zanthorrhiza	KX650821	KX650811	KY303758	KY485936	KY594916	KY978404
C. inodora	KX650822	KX650809	KY303750	KY463996	KY496974	KY851763
C. montana	KX650823	KX650810	KY303749	KY463998	KY784117	KY851764
C. aurantiaca	KX608617	KX650812	KY303755	KY463992	KY614250	KY978414

4.1.2 Nucleotide- substitution model selection

Bayesian Information Criterion (BIC) and Akaike Information Criterion (AIC) are the best-fit nucleotide-substitution models determined using MEGA 6.0; it was found that T92 (Tamura 3-parameter model), with the lowest BIC score (17172.057), and lowest AIC score (16760.754). Models with the lowest BIC scores (Bayesian Information Criterion) are depicts the best substitution pattern. Non-uniformity of evolutionary rates among sites were also modeled by using a discrete Gamma distribution (+*G*) with 5 rate categories and by assuming that a certain fraction of sites are evolutionarily invariable (+*I*). Estimates of gamma shape parameter and or the estimated fraction of invariant sites were shown in Fig. 4.4. Estimated values of transition/transversion bias (*R*) are shown for each model. They were followed by nucleotide frequencies (*f*) and rates of base substitutions (*r*) for each nucleotide pair. Relative values of instantaneous *r* were considered and for simplicity, sum of the *r* values is made equal to 1 for each model **(**Fig. 4.4).

4.1.3 Maximum likelihood phylogenetic analysis within the genus *Curcuma* on the basis of chloroplast sequences

The phylogenetic tree obtained from the maximum likelihood (ML) analysis based on six chloroplast sequences of 24 species comprising 3 related outgroup taxa (APG III 2009) provided four groups. The bootstrap support for each clade is shown below the branches. The rooted ML tree was deciphered using *Zingiber officinale, Z. nimmonii* and Z. *zerumbet* as outgroup. The group one clustered with 9 different species *Curcuma* with monophyletic origin, all the members shared a common ancestor. On the other hand, the species *C. inodora* and *C. kudagensis* clustered with a single clade with 100% bootstrap value (Fig. 4.2). The species *C. aromatica* and *C. zedoaria* are grouped together with 100% bootstrap value in to a single group. The

remaining species (*C. montana*, *C. zanthorrhiza, C. oligantha* var. *oligantha, C. neilgherrensis* and *C. mutabilis*) constitute a single group with >50% bootstrap value. The present study observed that on the basis of chloroplast nucleotide sequences, the species *C. bhatii* could be grouped with other species but exists as a monoclade with least bootstrap value (Fig. 4.2).

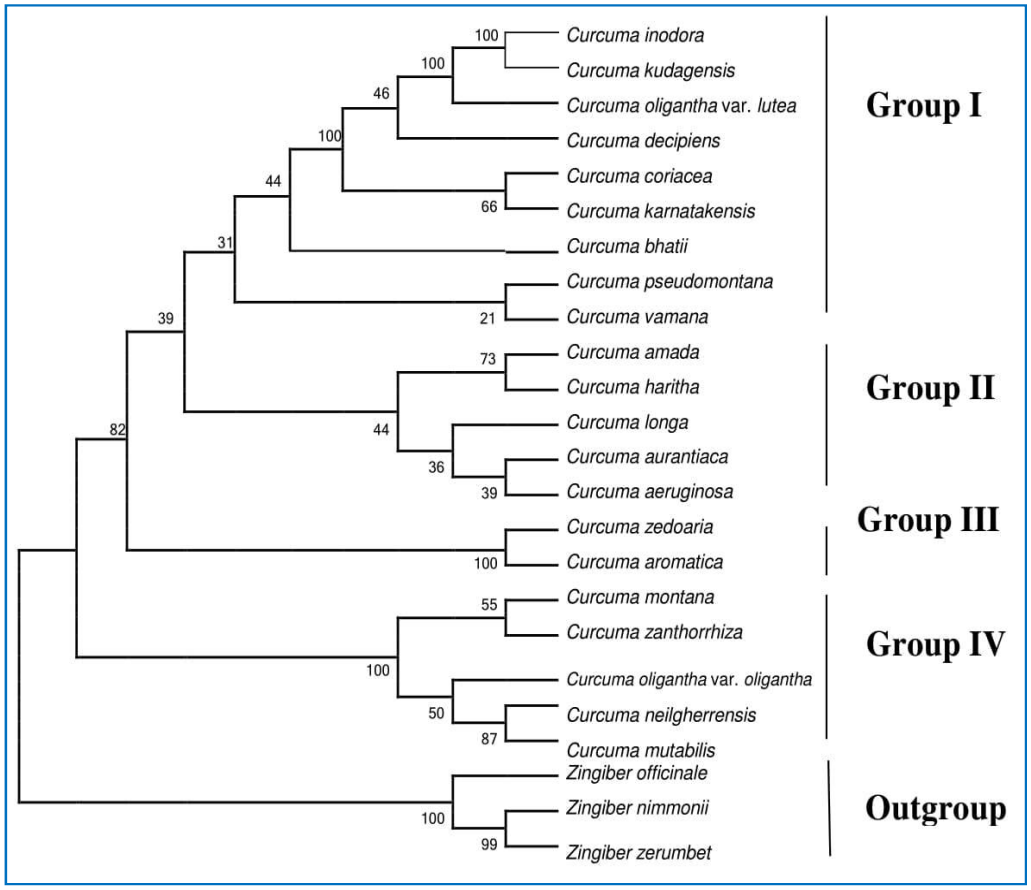


Fig. 4.2: Maximum likelihood phylogeny of the 21 species of the genus *Curcuma* based on chloroplast sequences data. Numbers beneath nodes are Bootstrap support (BS) indices. The three *Zingiber* species was selected as outgroup.

4.1.4 Maximum parsimony phylogenetic analysis of genus *Curcuma* based on the chloroplast sequences.

The maximum parsimony (MP) phylogenetic tree was constructed using the chloroplast sequences of 24 species including 3 outgroup taxa (Fig. 4.3). The similar topology was used that of the ML phylogenetic tree. Three groups were resolved from this MP phylogenetic tree with respective bootstrap support for each clade shown below the branches (Fig. 4.3). Similar to the ML tree, the MP tree was rooted using the related outgroup species such as *Z. officinale, Z. nimmonii* and *Z. zerumbet*. This MP phylogenetic tree also revealed the monophyletic status of the *Curcuma* species.

From the MP phylogenetic tree it was also observed that, the members of the *Curcuma* clustered into three groups based on the characteristics of the chloroplast region (Fig. 4.3). The group one has >50% bootstrap value comprised of nine *Curcuma* species. In the MP tree the species *C. inodora* and *C. kudagensis* are clustered into a single clade with 100% bootstrap value. At the same time another species namely *C. oligantha* var. *lutea* grouped together with *C. inodora* clade with 100% bootstrap values. In MP analysis the species *C. bhatii* grouped with other species but exists as a monoclade with least bootstrap value. The second group consists of seven species, the species *C. aromatica* and *C. zedoaria* clustered in a single clade with 100% bootstrap value. *C. oligantha* var. *oligantha* showed the relationship with other members of this group like *C. mutabilis* and *C. montana* clade with above 60% of bootstrap value in Fig 4.3.

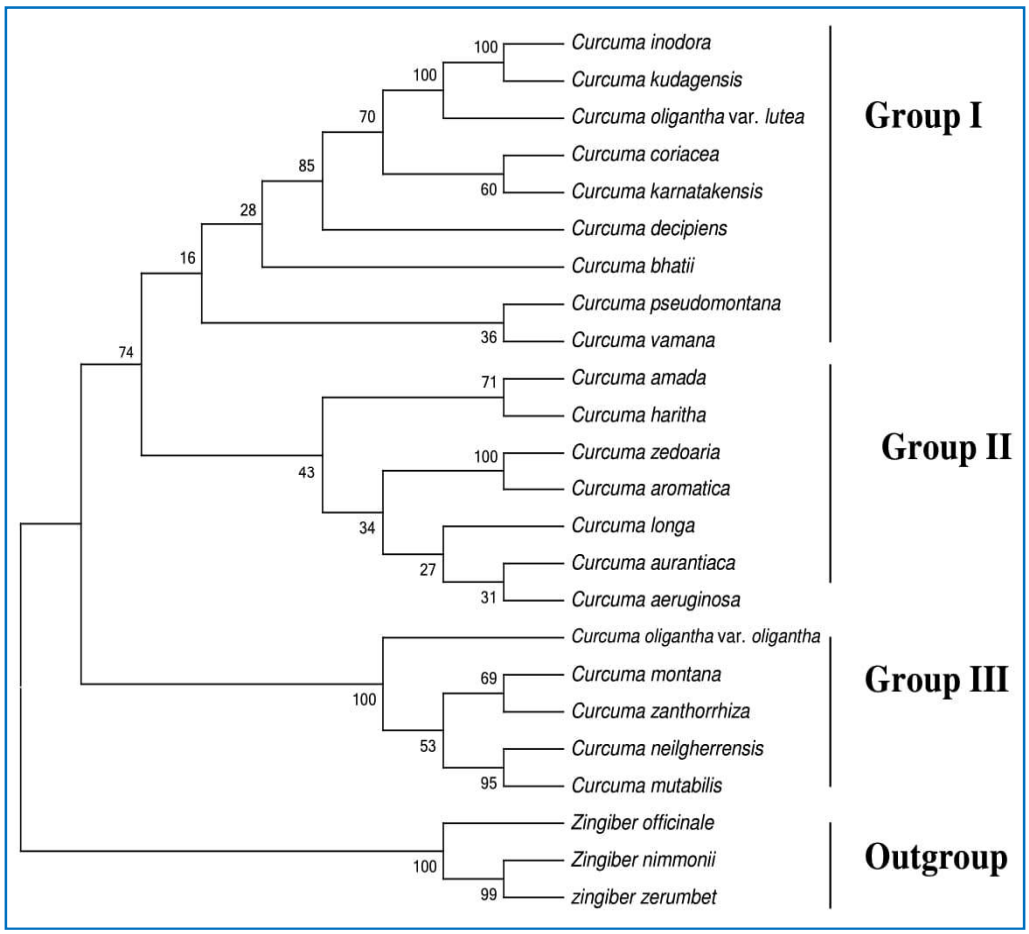


Fig. 4.3: Strict consensus of three equally most parsimonious trees of 21 *Curcuma* species generated from sequences of six genes of chloroplast DNA. Numbers beneath nodes are Bootstrap support (BS) indices.

Model	Parameters	BIC	AICc	*lnL*	(+*I*)	(+*G*)	*R*	*f*(A)	*f*(T)	*f*(C)	*f*(G)	*r*(AT)	*r*(AC)	*r*(AG)	*r*(TA)	*r*(TC)	*r*(TG)	*r*(CA)	*r*(CT)	*r*(CG)	*r*(GA)	*r*(GT)	*r*(GC)
T92	47	17172.057	16760.754	-8333.329	n/a	n/a	0.94	0.328	0.328	0.172	0.172	0.080	0.042	0.088	0.080	0.088	0.042	0.080	0.168	0.042	0.168	0.080	0.042
T92+I	48	17182.810	16762.758	-8333.329	0.00	n/a	0.94	0.328	0.328	0.172	0.172	0.080	0.042	0.088	0.080	0.088	0.042	0.080	0.168	0.042	0.168	0.080	0.042
T92+G	48	17183.218	16763.165	-8333.532	n/a	200.00	0.94	0.328	0.328	0.172	0.172	0.080	0.042	0.088	0.080	0.088	0.042	0.080	0.168	0.042	0.168	0.080	0.042
T92+G+I	49	17193.971	16765.169	-8333.532	0.00	200.00	0.94	0.328	0.328	0.172	0.172	0.080	0.042	0.088	0.080	0.088	0.042	0.080	0.168	0.042	0.168	0.080	0.042
HKY	49	17201.363	16772.561	-8337.228	n/a	n/a	0.94	0.322	0.335	0.157	0.187	0.082	0.038	0.095	0.079	0.080	0.046	0.079	0.171	0.046	0.165	0.082	0.038
TN93	50	17209.917	16772.366	-8336.129	n/a	n/a	0.94	0.322	0.335	0.157	0.187	0.082	0.038	0.090	0.079	0.086	0.045	0.079	0.183	0.045	0.155	0.082	0.038
HKY+I	50	17212.116	16774.566	-8337.228	0.00	n/a	0.94	0.322	0.335	0.157	0.187	0.082	0.038	0.095	0.079	0.080	0.046	0.079	0.171	0.046	0.165	0.082	0.038
HKY+G	50	17212.504	16774.953	-8337.422	n/a	200.00	0.94	0.322	0.335	0.157	0.187	0.082	0.038	0.095	0.079	0.080	0.046	0.079	0.171	0.046	0.165	0.082	0.038
TN93+I	51	17220.670	16774.371	-8336.129	0.00	n/a	0.94	0.322	0.335	0.157	0.187	0.082	0.038	0.090	0.079	0.086	0.045	0.079	0.183	0.045	0.155	0.082	0.038
TN93+G	51	17221.068	16774.769	-8336.328	n/a	200.00	0.94	0.322	0.335	0.157	0.187	0.082	0.038	0.090	0.079	0.086	0.045	0.079	0.183	0.045	0.155	0.082	0.038
HKY+G+I	51	17223.257	16776.957	-8337.422	0.00	200.00	0.94	0.322	0.335	0.157	0.187	0.082	0.038	0.095	0.079	0.080	0.046	0.079	0.171	0.046	0.165	0.082	0.038
TN93+G+I	52	17231.821	16776.773	-8336.328	0.00	200.00	0.94	0.322	0.335	0.157	0.187	0.082	0.038	0.090	0.079	0.086	0.045	0.079	0.183	0.045	0.155	0.082	0.038
GTR	53	17233.299	16769.503	-8331.690	n/a	n/a	0.94	0.322	0.335	0.157	0.187	0.071	0.047	0.089	0.068	0.084	0.046	0.096	0.181	0.046	0.153	0.082	0.038
GTR+I	54	17244.053	16771.508	-8331.690	0.00	n/a	0.94	0.322	0.335	0.157	0.187	0.071	0.047	0.089	0.068	0.084	0.046	0.096	0.181	0.046	0.153	0.082	0.038
GTR+G	54	17244.435	16771.890	-8331.881	n/a	200.00	0.94	0.322	0.335	0.157	0.187	0.071	0.047	0.089	0.068	0.084	0.046	0.096	0.181	0.046	0.153	0.082	0.038
GTR+G+I	55	17255.188	16773.895	-8331.881	0.00	200.00	0.94	0.322	0.335	0.157	0.187	0.071	0.047	0.089	0.068	0.084	0.046	0.096	0.181	0.046	0.153	0.082	0.038
K2	46	17463.855	17061.301	-8484.604	n/a	n/a	1.11	0.250	0.250	0.250	0.250	0.059	0.059	0.132	0.059	0.132	0.059	0.059	0.132	0.059	0.132	0.059	0.059
K2+I	47	17474.608	17063.305	-8484.604	0.00	n/a	1.11	0.250	0.250	0.250	0.250	0.059	0.059	0.132	0.059	0.132	0.059	0.059	0.132	0.059	0.132	0.059	0.059
K2+G	47	17482.476	17071.172	-8488.538	n/a	200.00	1.19	0.250	0.250	0.250	0.250	0.057	0.057	0.136	0.057	0.136	0.057	0.057	0.136	0.057	0.136	0.057	0.057
K2+G+I	48	17493.229	17073.176	-8488.538	0.00	200.00	1.19	0.250	0.250	0.250	0.250	0.057	0.057	0.136	0.057	0.136	0.057	0.057	0.136	0.057	0.136	0.057	0.057
JC	45	17545.689	17151.884	-8530.898	n/a	n/a	0.50	0.250	0.250	0.250	0.250	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083
JC+I	46	17556.442	17153.888	-8530.898	0.00	n/a	0.50	0.250	0.250	0.250	0.250	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083
JC+G	46	17556.783	17154.229	-8531.068	n/a	200.00	0.50	0.250	0.250	0.250	0.250	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083
JC+G+I	47	17567.536	17156.233	-8531.068	0.00	200.00	0.50	0.250	0.250	0.250	0.250	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083

Fig. 4.4: Nucleotide-substitution model selection for the chloroplast sequences of the studied taxa of the genus *Curcuma*; [GTR: General Time Reversible; HKY: Hasegawa-Kishino-Yano; TN93; Tamura-Nei; T92: Tamura 3-parameter; K2: Kimura 2-parameter; JC: Jukes-Cantor]. [(+G): Gamma distribution; (+I): Evolutionarily invariable site; (R): Transition/ transversion bias; (f): Nucleotide frequencies; (r): Rates of base substitutions]

4.2 Molecular phylogenetic analysis based on mitochondrial sequences

4.2.1 Sequence characteristics

To evaluate the phylogenetic relationship within the members of *Curcuma* species using mitochondrial (*atp1* and *cox1*) regions was also taken into consideration. Amplification of PCR product of *Curcuma* species using mitochondrial sequences showed distinct bands of *atp1*~1300 and *cox1*~600 sizes respectively, with high intense amplicon (Fig. 4.5A and B). The PCR products were sequenced using automated DNA sequencing. The obtained chromatograms were analyzed using Bio-Edit.v.7.1.3 software (Ibis Biosciences, Carlsbad, CA 92008) and the forward and reverse strands of each sequences were aligned and edited using this software. Sequences were compared with other sequences from BLAST. All the obtained sequence data have been deposited in the NCBI GenBank database (http://www.ncbi. nlm.nih.gov) under the accession numbers listed in the Table 4.2.

Alignment of the two genomic regions resulted in a combined dataset with the following: *atp1* and *cox1*. Out of the 2968 character sites, 779 were constant, 1601 were variable but parsimony uninformative, and 588 parsimony informative. Maximum parsimony analyses of the combined dataset yielded >200.0 maximally parsimonious trees of 960 steps, consistency index (CI) of 0.80, and retention index (RI) of 0.86. Incongruence between the two data sets was assessed using the incongruence length difference (ILD) test and found the high level of congruence between the two data set (p=0.08), and the data was combined. The maximum likelihood (ML) tree is shown is Fig 4.6 and strict consensus most maximum parsimonious (MP) tree is shown in Fig.4.7.

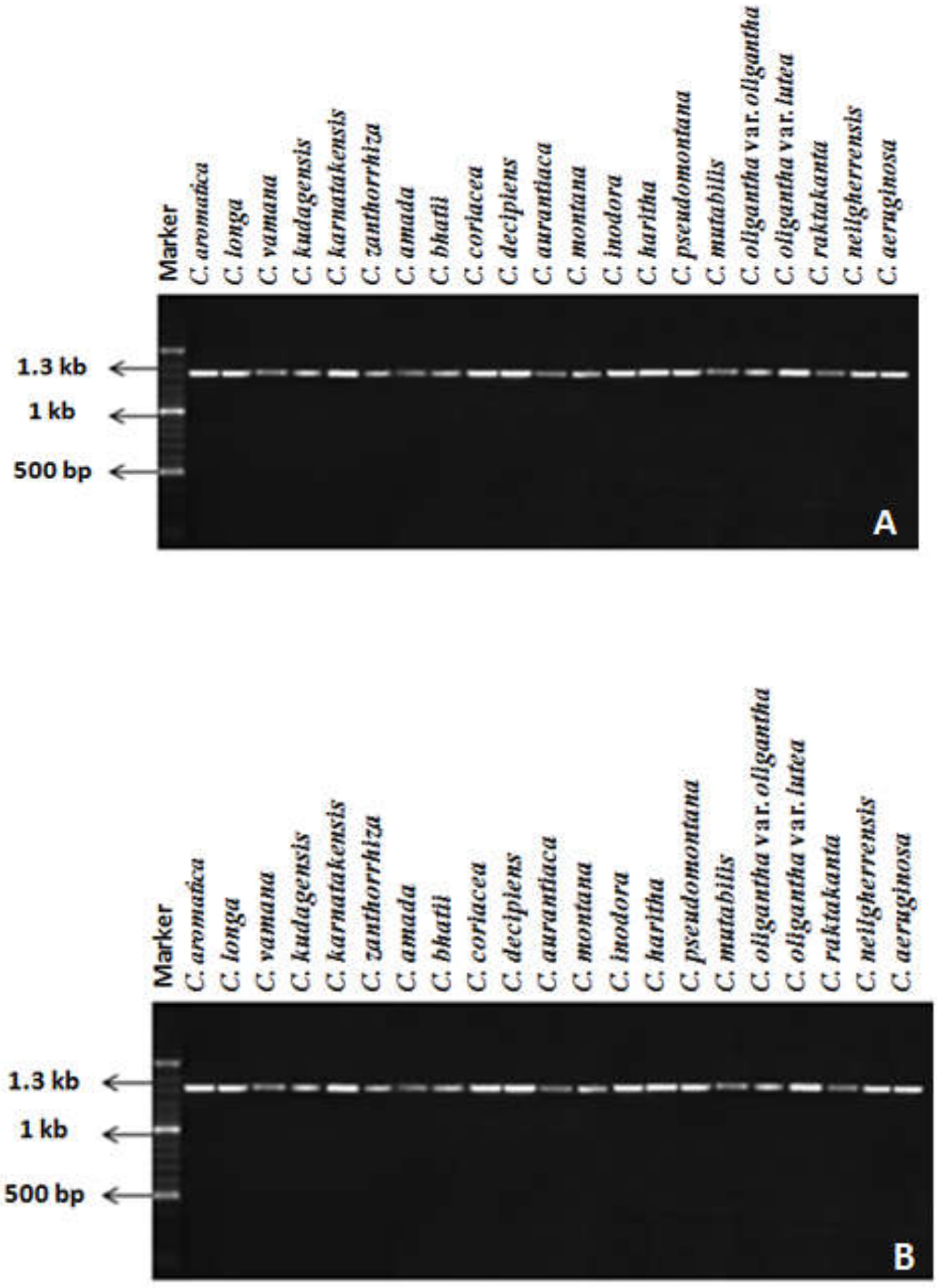


Fig. 4.5. Gel-electrophoresis of PCR product of *Curcuma* species shows a band of **A**-*atp1*~ 1300 bases and **B**-*cox1*~ 600 bases.

Table 4.2: GenBank accession numbers for *atp1* and *cox1* sequences of twenty one species of the *Curcuma*

Name of the Species	*atp1*	*cox1*
C. coriacea	KY200896	KY225972
C. karnatakensis	KY200910	KY225987
C. bhatii	KY200897	KY225971
C. zedoaria	KY200908	KY225982
C. oligantha var *oligantha*	KY200907	KY225983
C. aeruginosa	KY200906	KY225968
C. vamana	KY200905	KY225985
C. haritha	KY200899	KY225975
C. oligantha var *lutea*	KY200904	KY225981
C. amada	KY170860	KY225969
C. decipiens	KY200898	KY225973
C. kudagensis	KY200902	KY225977
C. aromatica	KY170861	KY225970
C. mutabilis	KY200903	KY225980
C. pseudomontana	KY864244	KY225984
C. neilgherrensis	KY200901	KY225988
C. longa	KY030908	KY225978
C. zanthorrhiza	KY864243	KY225986
C. inodora	KY200900	KY225976
C. montana	KY864242	KY225979
C. aurantiaca	KY200909	KY225974

4.2.2 Nucleotide- substitution model selection

Bayesian Information Criterion (BIC) and Akaike Information Criterion (AIC) are the best-fit nucleotide-substitution models. Both were assessed by using MEGA 6.0; with similarity to the chloroplast model T92 (Tamura 3-parameter model), having a lowest BIC score (7705.106) and lowest AIC score (7373.129). The lowest BIC (Bayesian Information Criterion) scores are considered to represent the best substitution model. Non-uniformity of evolutionary rates among sites were also modeled using a

discrete Gamma distribution (+*G*) with 5 rate categories and by assuming that a certain fraction of sites are evolutionarily invariable (+*I*). Estimates of gamma shape parameter and/or the estimated fraction of invariant sites are represented in Fig 4.8. Estimated values of transition/ transversion bias (*R*) are shown for each model with nucleotide frequencies (*f*) and rates of base substitutions (*r*) for each nucleotide pair. Relative values of instantaneous *r* were considered and for simplicity, sum of the *r* values is made equal to 1 in each model (Fig. 4.8).

4.2.3 Maximum likelihood and Maximum parsimony phylogenetic analysis within the *Curcuma* species on the basis of mitochondrial sequence

Both ML and MP phylogenetic analysis were conducted within the species of *Curcuma* by using two mitochondrial sequences. The ML and MP phylogenetic tree obtained from this analysis indicated the formation of two groups (Fig. 4.6 & 4.7). The species *C. montana*, *C. zanthorrhiza, C. pseudomontana* and *C. aurantiaca* grouped together and formed a single group, the remaining species were grouped together and shared a common ancestor. The mitochondrial sequences of *Maranta leuconeura* (AY299801.1, AJ223432.1) were used as outgroup to construct both ML and MP. The two groups were resolved from this ML and MP phylogenetic tree based on the mitochondrial sequences with respective bootstrap support for each clade shown below the branches (Fig.4.6 & 4.7). The mitochondrial sequence based phylogenetic analysis, was slightly different compared to chloroplast sequence based phylogenies (Fig.4.6 &4.7), the member of the genus *Curcuma bhatii* exist as a monoclade in the ML analysis with >50% bootstrap value (Fig.4.6).

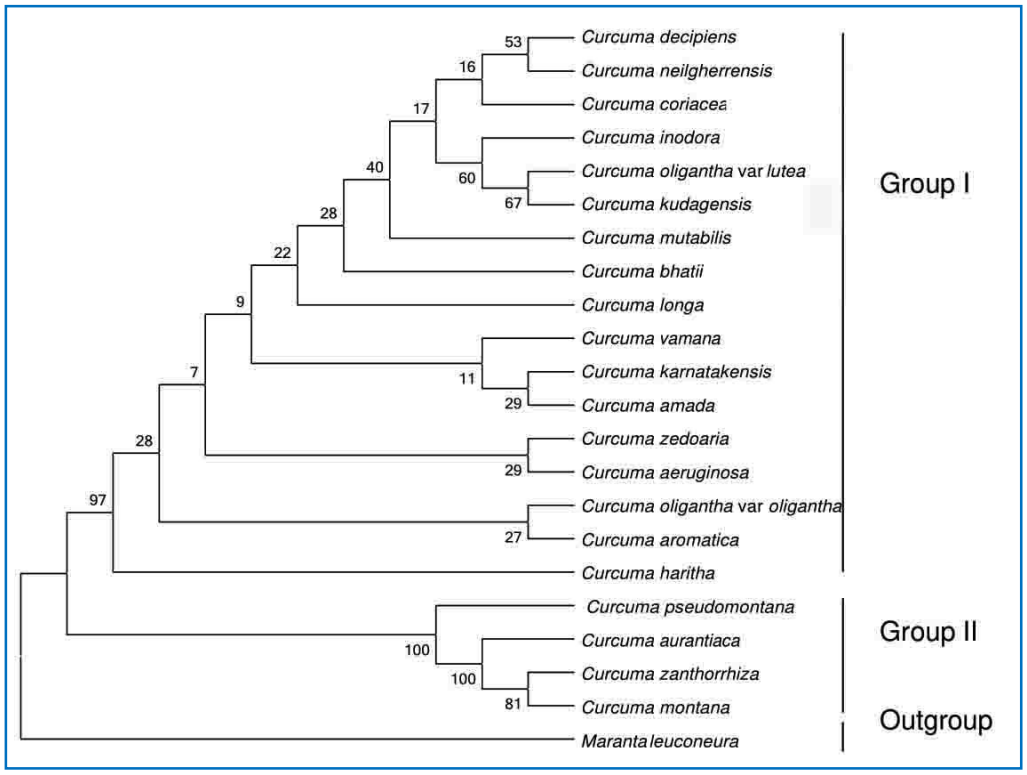


Fig. 4.6: Maximum likelihood phylogeny of the studied members of the genus *Curcuma* based on mitochondrial sequences data. Numbers beneath nodes are Bootstrap support (BS) indices.

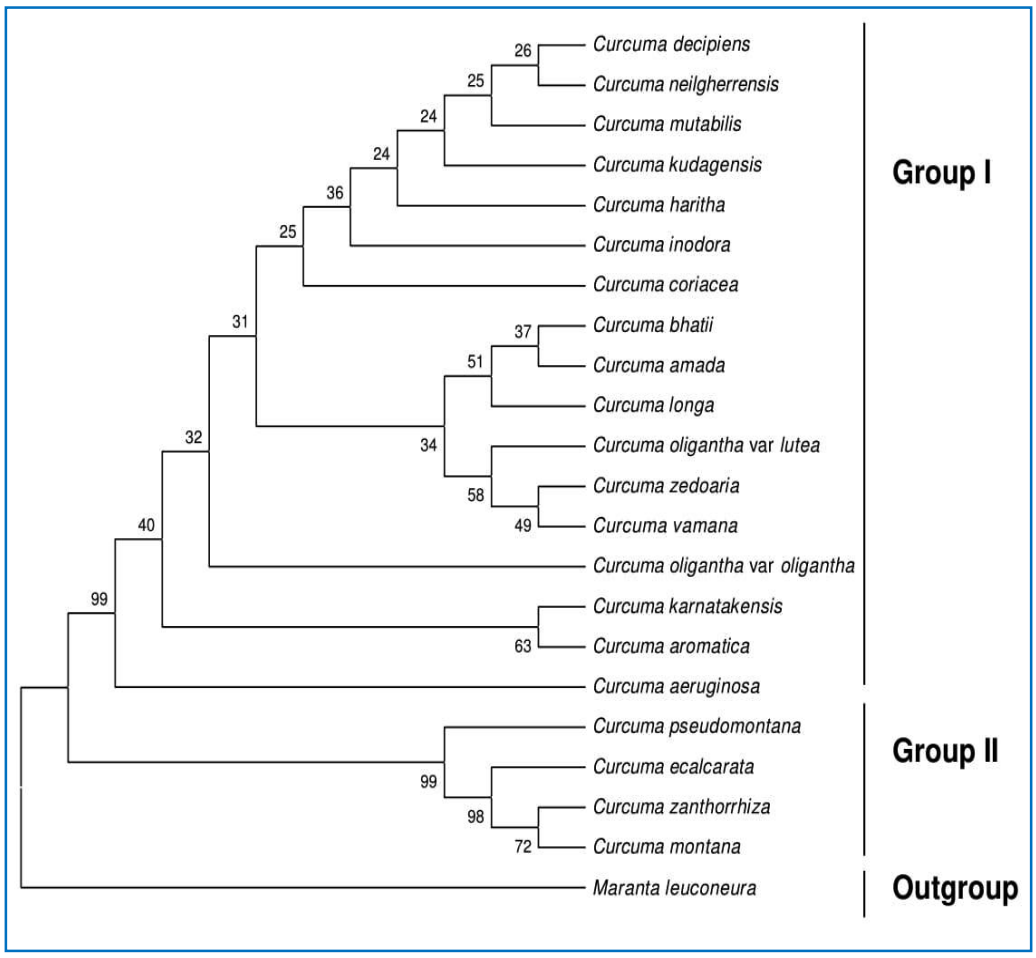


Fig. 4.7: Strict consensus of three equally most parsimonious trees of 21 *Curcuma* species generated from sequences of two genes of mitochondrial DNA. Numbers beneath nodes are Bootstrap support (BS) indices.

Model	Parameters	BIC	AICc	lnL	(+I)	(+G)	R	f(A)	f(T)	f(C)	f(G)	r(AT)	r(AC)	r(AG)	r(TA)	r(TC)	r(TG)	r(CA)	r(CT)	r(CG)	r(GA)	r(GT)	r(GC)
T92	43	7705.106	7373.129	-3643.451	n/a	n/a	0.99	0.282	0.282	0.218	0.218	0.070	0.054	0.109	0.070	0.109	0.054	0.070	0.142	0.054	0.142	0.070	0.054
K2	42	7713.507	7389.245	-3652.514	n/a	n/a	0.99	0.250	0.250	0.250	0.250	0.063	0.063	0.124	0.063	0.124	0.063	0.063	0.124	0.063	0.124	0.063	0.063
T92+G	44	7714.341	7374.649	-3643.206	n/a	18.86	1.00	0.282	0.282	0.218	0.218	0.070	0.054	0.110	0.070	0.110	0.054	0.070	0.142	0.054	0.142	0.070	0.054
T92+I	44	7714.832	7375.139	-3643.451	0.00	n/a	0.99	0.282	0.282	0.218	0.218	0.070	0.054	0.109	0.070	0.109	0.054	0.070	0.142	0.054	0.142	0.070	0.054
K2+G	43	7722.581	7390.604	-3652.189	n/a	16.27	0.99	0.250	0.250	0.250	0.250	0.063	0.063	0.125	0.063	0.125	0.063	0.063	0.125	0.063	0.125	0.063	0.063
K2+I	43	7723.233	7391.256	-3652.514	0.00	n/a	0.99	0.250	0.250	0.250	0.250	0.063	0.063	0.124	0.063	0.124	0.063	0.063	0.124	0.063	0.124	0.063	0.063
T92+G+I	45	7724.067	7376.660	-3643.206	0.00	18.86	1.00	0.282	0.282	0.218	0.218	0.070	0.054	0.110	0.070	0.110	0.054	0.070	0.142	0.054	0.142	0.070	0.054
HKY	45	7725.340	7377.933	-3643.842	n/a	n/a	0.99	0.259	0.306	0.242	0.194	0.077	0.061	0.096	0.065	0.120	0.049	0.065	0.152	0.049	0.129	0.077	0.061
TN93	46	7728.763	7373.641	-3640.691	n/a	n/a	0.99	0.259	0.306	0.242	0.194	0.076	0.060	0.115	0.064	0.104	0.048	0.064	0.131	0.048	0.153	0.076	0.060
K2+G+I	44	7732.307	7392.614	-3652.189	0.00	16.27	0.99	0.250	0.250	0.250	0.250	0.063	0.063	0.125	0.063	0.125	0.063	0.063	0.125	0.063	0.125	0.063	0.063
HKY+G	46	7734.535	7379.413	-3643.577	n/a	18.07	1.00	0.259	0.306	0.242	0.194	0.077	0.061	0.096	0.065	0.121	0.049	0.065	0.152	0.049	0.129	0.077	0.061
HKY+I	46	7735.066	7379.944	-3643.842	0.00	n/a	0.99	0.259	0.306	0.242	0.194	0.077	0.061	0.096	0.065	0.120	0.049	0.065	0.152	0.049	0.129	0.077	0.061
TN93+G	47	7737.740	7374.903	-3640.317	n/a	16.57	1.00	0.259	0.306	0.242	0.194	0.076	0.060	0.115	0.064	0.103	0.048	0.064	0.131	0.048	0.154	0.076	0.060
TN93+I	47	7738.361	7375.524	-3640.627	0.00	n/a	0.99	0.259	0.306	0.242	0.194	0.076	0.060	0.115	0.064	0.104	0.048	0.064	0.131	0.048	0.153	0.076	0.060
HKY+G+I	47	7744.261	7381.425	-3643.577	0.00	18.07	1.00	0.259	0.306	0.242	0.194	0.077	0.061	0.096	0.065	0.121	0.049	0.065	0.152	0.049	0.129	0.077	0.061
TN93+G+I	48	7747.466	7376.915	-3640.317	0.00	16.57	1.00	0.259	0.306	0.242	0.194	0.076	0.060	0.115	0.064	0.103	0.048	0.064	0.131	0.048	0.154	0.076	0.060
GTR	49	7751.720	7373.456	-3637.581	n/a	n/a	0.99	0.259	0.306	0.242	0.194	0.066	0.075	0.114	0.056	0.103	0.042	0.081	0.131	0.050	0.153	0.067	0.062
JC	41	7757.156	7440.610	-3679.202	n/a	n/a	0.50	0.250	0.250	0.250	0.250	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083
GTR+G	50	7760.855	7374.877	-3637.286	n/a	17.08	1.00	0.259	0.306	0.242	0.194	0.065	0.075	0.115	0.055	0.103	0.042	0.081	0.131	0.050	0.154	0.067	0.062
GTR+I	50	7762.085	7376.107	-3637.901	0.00	n/a	0.99	0.259	0.306	0.242	0.194	0.066	0.075	0.114	0.056	0.103	0.042	0.081	0.131	0.050	0.153	0.067	0.062
JC+G	42	7766.472	7442.210	-3678.997	n/a	20.78	0.50	0.250	0.250	0.250	0.250	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083
JC+I	42	7766.881	7442.619	-3679.201	0.00	n/a	0.50	0.250	0.250	0.250	0.250	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083
GTR+G+I	51	7770.581	7376.889	-3637.286	0.00	17.08	1.00	0.259	0.306	0.242	0.194	0.065	0.075	0.115	0.055	0.103	0.042	0.081	0.131	0.050	0.154	0.067	0.062
JC+G+I	43	7776.198	7444.220	-3678.997	0.00	20.78	0.50	0.250	0.250	0.250	0.250	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083

Fig. 4.8: Nucleotide- substitution model selection for the mitochondrial sequences of studied taxa of the genus *Curcuma*; [GTR: General Time Reversible; HKY: Hasegawa-Kishino-Yano; TN93: Tamura-Nei; T92: Tamura 3-parameter; K2: Kimura 2-parameter; JC: Jukes-Cantor]. [(+G): Gamma distribution; (+I): Evolutionarily invariable site; (R): Transition/ transversion bias; (f): Nucleotide frequencies; (r): Rates of base substitutions

4.3 Molecular phylogenetic analysis based on internal transcribed spacer (ITS1) sequences.

4.3.1 Sequence characteristics

The genomic DNA was isolated from all the species using modified CTAB method. The quantity and quality of the isolated DNA was checked

using a nanodrop spectrophotometer. To assess the phylogenetic relationship between the members of *Curcuma* species internal transcribed spacer (ITS1) sequence was used. The PCR amplification was carried out using gene specific primers. Agarose gel electrophoresis of the PCR products generated using internal transcribed spacer (ITS1) sequences showed distinct band of molecular size~450 each species (Fig. 4.9). Amplicons generated were sequenced using automated DNA sequencer. The chromatograms were analyzed using Bio-Edit.v.7.1.3 software (Ibis Biosciences, Carlsbad, CA 92008). The forward and reverse strands of all the sequences were manually aligned and edited used this software. Sequence homology with other sequences were detected using BLAST search tool of NCBI. All the obtained sequence data were deposited in the NCBI GenBank database (http://www.ncbi.nim.nih.gov) under the accession numbers listed in the Table 4.3.

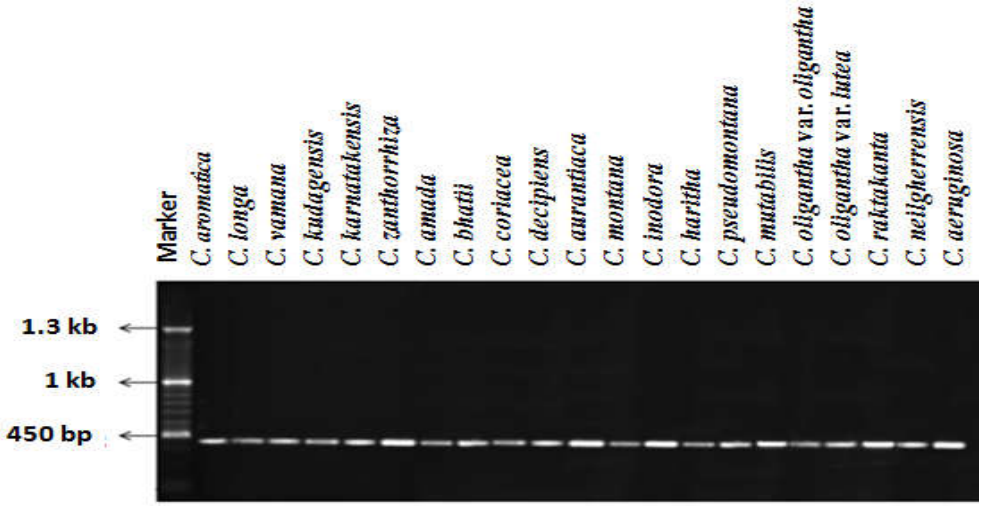


Fig. 4.9. Gel-electrophoresis of PCR product of *Curcuma* species shows a band of ITS1 with 450 base pair.

Alignment of the genomic region resulted in a dataset with a total of 722 character sites, 193 were constant, 542 were variable but parsimony

uninformative, and 347 were parsimony informative. Maximum parsimony analysis yielded a maximally parsimonious tree of 245 steps, with consistency index (CI) of 0.75 and retention index (RI) of 0.80. The maximum likelihood (ML) tree (Fig 4.10) and strict consensus most maximum parsimonious (MP) tree was deciphered (Fig. 4.11).

Table 4.3 List of NCBI GenBank accession numbers for ITS1 sequences of twenty one collected species of the genus *Curcuma*

Name of the Species	**ITS1**
C. coriacea	MF576506
C. karnatakensis	MF595875
C. bhatii	MF576505
C. zedoaria	MF571833
C. oligantha var *oligantha*	MF611628
C. aeruginosa	MF628251
C. vamana	MF611631
C. haritha	MF628254
C. oligantha var *lutea*	MF611629
C. amada	MF628252
C. decipiens	MF595872
C. kudagensis	MF595676
C. aromatica	MF628253
C. mutabilis	MF628256
C. pseudomontana	MF611630
C. neilgherrensis	MF611627
C. longa	MF628255
C. zanthorrhiza	MF611632
C. inodora	MF595874
C. montana	MF611626
C. aurantiaca	MF595873

4.3.2 Model selection

The accepted nucleotide substitution models were Bayesian Information Criterion (BIC) and Akaike Information Criterion (AIC) as the best-fit nucleotide-substitution models determined by using MEGA 6.0; it was found to be model T92+G (Tamura 3-parameter model), with the lowest BIC score (10757.816), and lowest AIC score (10405.460) (Fig. 4.12).

Model	Parameters	BIC	AICc	lnL	(+I)	(+G)	R	f(A)	f(T)	f(C)	f(G)	r(AT)	r(AC)	r(AG)	r(TA)	r(TC)	r(TG)	r(CA)	r(CT)	r(CG)	r(GA)	r(GT)	r(GC)
T92+G	48	10757.816	10405.460	-5154.524	n/a	2.45	1.17	0.225	0.225	0.275	0.275	0.051	0.063	0.149	0.051	0.149	0.063	0.051	0.122	0.063	0.122	0.051	0.063
T92+G+I	49	10767.165	10407.477	-5154.524	0.00	2.45	1.17	0.225	0.225	0.275	0.275	0.051	0.063	0.149	0.051	0.149	0.063	0.051	0.122	0.063	0.122	0.051	0.063
K2+G	47	10775.283	10430.259	-5167.932	n/a	2.26	1.46	0.250	0.250	0.250	0.250	0.051	0.051	0.149	0.051	0.149	0.051	0.051	0.149	0.051	0.149	0.051	0.051
HKY+G	50	10778.025	10411.005	-5155.279	n/a	2.44	1.17	0.229	0.220	0.284	0.267	0.050	0.065	0.145	0.053	0.154	0.061	0.053	0.119	0.061	0.124	0.050	0.065
T92+I	48	10781.153	10428.796	-5166.193	0.08	n/a	1.13	0.225	0.225	0.275	0.275	0.052	0.064	0.147	0.052	0.147	0.064	0.052	0.120	0.064	0.120	0.052	0.064
K2+G+I	48	10784.632	10432.276	-5167.932	0.00	2.26	1.46	0.250	0.250	0.250	0.250	0.051	0.051	0.149	0.051	0.149	0.051	0.051	0.149	0.051	0.149	0.051	0.051
TN93+G	51	10787.214	10412.862	-5155.199	n/a	2.44	1.17	0.229	0.220	0.284	0.267	0.050	0.065	0.148	0.053	0.150	0.061	0.053	0.117	0.061	0.127	0.050	0.065
HKY+G+I	51	10787.374	10413.023	-5155.279	0.00	2.44	1.17	0.229	0.220	0.284	0.267	0.050	0.065	0.145	0.053	0.154	0.061	0.053	0.119	0.061	0.124	0.050	0.065
K2+I	47	10790.175	10445.151	-5175.379	0.08	n/a	1.13	0.250	0.250	0.250	0.250	0.059	0.059	0.132	0.059	0.132	0.059	0.059	0.132	0.059	0.132	0.059	0.059
T92	47	10795.910	10450.886	-5178.246	n/a	n/a	1.11	0.225	0.225	0.275	0.275	0.053	0.065	0.145	0.053	0.145	0.065	0.053	0.119	0.065	0.119	0.053	0.065
TN93+G+I	52	10796.563	10414.880	-5155.199	0.00	2.44	1.17	0.229	0.220	0.284	0.267	0.050	0.065	0.148	0.053	0.150	0.061	0.053	0.117	0.061	0.127	0.050	0.065
HKY+I	50	10801.336	10434.316	-5166.935	0.08	n/a	1.13	0.229	0.220	0.284	0.267	0.051	0.066	0.143	0.053	0.152	0.062	0.053	0.118	0.062	0.123	0.051	0.066
TN93+I	51	10810.638	10436.287	-5166.912	0.09	n/a	1.14	0.229	0.220	0.284	0.267	0.051	0.066	0.141	0.053	0.153	0.062	0.053	0.119	0.062	0.121	0.051	0.066
K2	46	10811.327	10473.636	-5190.629	n/a	n/a	1.34	0.250	0.250	0.250	0.250	0.054	0.054	0.143	0.054	0.143	0.054	0.054	0.143	0.054	0.143	0.054	0.054
GTR+G	54	10815.010	10418.666	-5155.073	n/a	2.44	1.17	0.229	0.220	0.284	0.267	0.052	0.062	0.148	0.054	0.150	0.063	0.050	0.117	0.061	0.127	0.052	0.065
HKY	49	10816.344	10456.656	-5179.114	n/a	n/a	1.10	0.229	0.220	0.284	0.267	0.052	0.067	0.141	0.054	0.150	0.063	0.054	0.116	0.063	0.121	0.052	0.067
GTR+G+I	55	10824.359	10420.686	-5155.073	0.00	2.44	1.17	0.229	0.220	0.284	0.267	0.052	0.062	0.148	0.054	0.150	0.063	0.050	0.117	0.061	0.127	0.052	0.065
TN93	50	10825.464	10458.444	-5178.999	n/a	n/a	1.10	0.229	0.220	0.284	0.267	0.052	0.067	0.144	0.054	0.146	0.063	0.054	0.113	0.063	0.124	0.052	0.067
GTR+I	54	10838.228	10441.884	-5166.683	0.09	n/a	1.14	0.229	0.220	0.284	0.267	0.054	0.062	0.141	0.056	0.153	0.064	0.050	0.119	0.062	0.121	0.053	0.065
GTR	53	10853.328	10464.315	-5178.907	n/a	n/a	1.10	0.229	0.220	0.284	0.267	0.054	0.065	0.144	0.056	0.146	0.064	0.052	0.113	0.063	0.124	0.052	0.067
JC+G	46	10878.392	10540.701	-5224.162	n/a	2.70	0.50	0.250	0.250	0.250	0.250	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083
JC+G+I	47	10887.741	10542.717	-5224.162	0.00	2.70	0.50	0.250	0.250	0.250	0.250	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083
JC+I	46	10898.798	10561.107	-5234.365	0.08	n/a	0.50	0.250	0.250	0.250	0.250	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083
JC	45	10910.058	10579.700	-5244.669	n/a	n/a	0.50	0.250	0.250	0.250	0.250	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083

Fig. 4.12: Nucleotide- substitution model selection for the rDNA ITS1 sequences of studied taxa of the genus *Curcuma*; [GTR: General Time Reversible; HKY: Hasegawa-Kishino-Yano; TN93: Tamura-Nei; T92: Tamura 3-parameter; K2: Kimura 2-parameter; JC: Jukes-Cantor]. [(+G): Gamma distribution; (+I): Evolutionarily invariable site; (R): Transition/ transversion bias; (f): Nucleotide frequencies; (r): Rates of base substitutions]

4.3.3 Maximum likelihood and Maximum parsimony phylogenetic analysis of genus *Curcuma* on the basis of Internal transcribed spacer (ITS1) sequence

The maximum likelihood (ML) and Maximum parsimony (MP) phylogenetic tree was constructed for all the *Curcuma* species using ITS1 sequences. The ML and MP phylogenetic tree obtained from this analysis indicated the formation of two groups (Fig. 4.10 & 4.11). The ITS1 sequences of *Zingiber officinale* (KR816713.1), *Z. nimmonii* (KJ872237.1) and *Z. montnum* (KJ872211.1) retrived from NCBI were used as outgroups. The two trees constructed using the same topology, ML and MP (Fig.4.19 and 4.20), among the member of the genus, 7 species were grouped together with single group in ML tree and the remaining species form another group, the species *C. bhatii* exist as a monoclade with <50% bootstrap value in ML tree and this species is clustered in *C. decipiens* with <50% bootstrap value in the MP tree (Fig.4.10). The species *C. aurantiaca* and *C. vamana* cluster together with single clade to form a single group with 93% of bootstrap value.

The MP tree divided two groups, 14 species are clustered in a single group such as *C. coriacea, C. oligantha* var. *oligantha, C. kudagensis, C. oligantha* var. *lutea, C. neilgherrensis, C. pdeudomontana, C. decipiens, C. bhatii, C. zanthorrhiza, C. aeruginosa, C. montana, C. aurantiaca, C. vamana* and *C. aromatica* and remaining 7 species (*C. amada, C. zedoaria, C. longa, C. haritha, C. karnatakensis, C. inodora* and *C. mutabilis*) were clustered in another group.

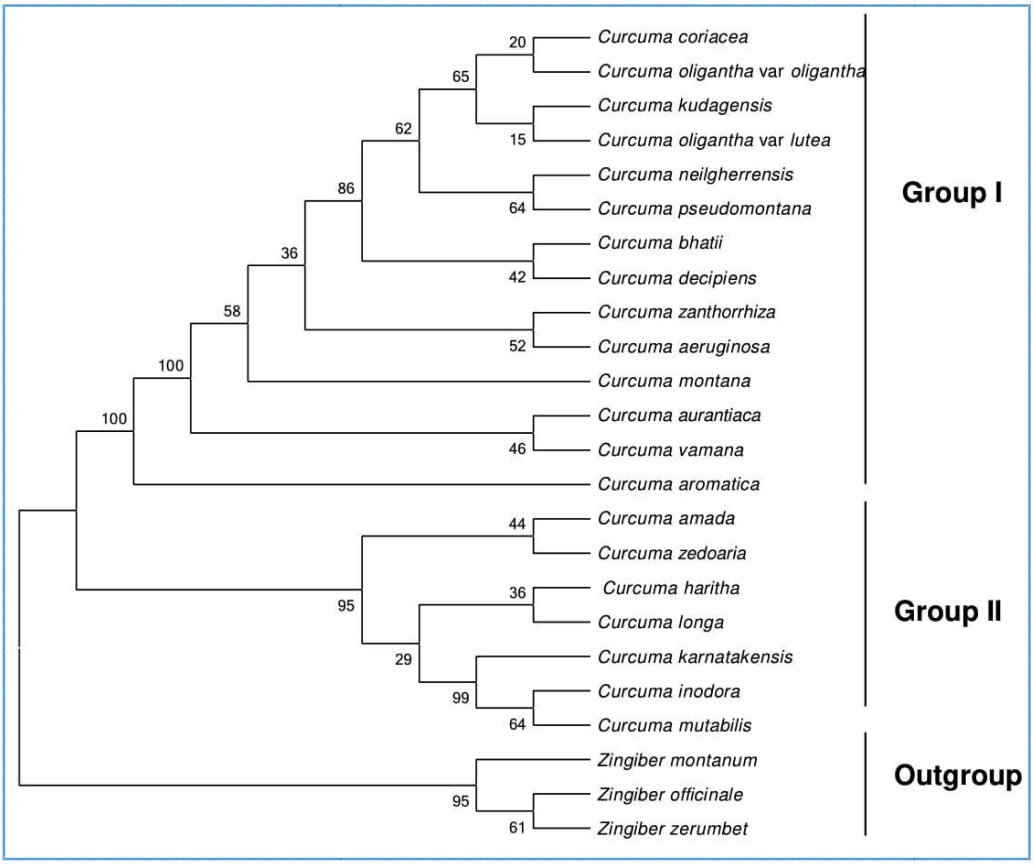


Fig. 4.10: Maximum likelihood phylogeny of the studied members of the genus *Curcuma* based on nuclear ITS data. Numbers beneath nodes are Bootstrap support (BS) indices.

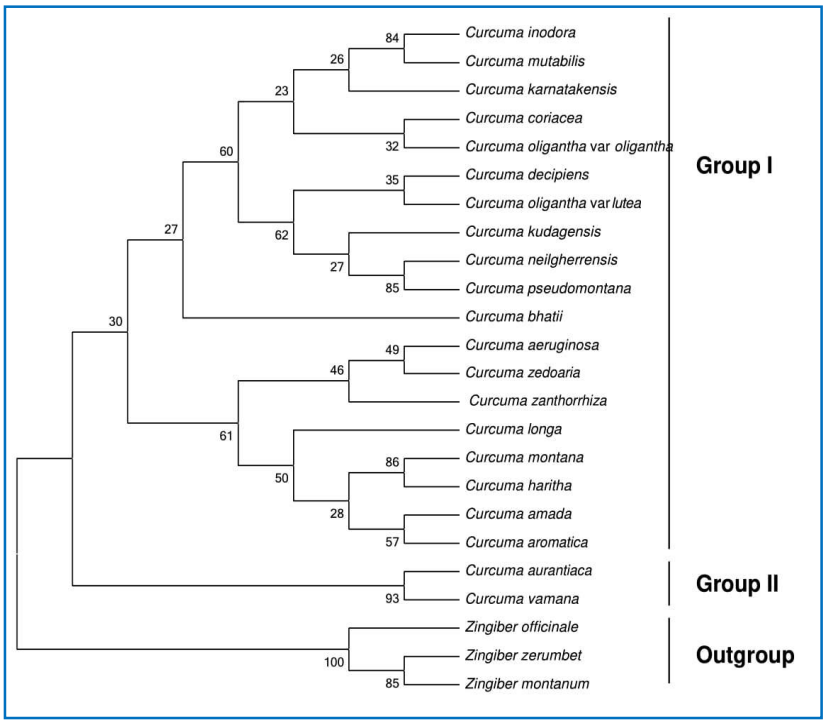


Fig. 4.11: Strict consensus of three equally most parsimonious trees of 21 *Curcuma* species generated from sequence of nuclear ITS. Numbers beneath nodes are Bootstrap support (BS)

4.4 Molecular phylogenetic analysis based on the combined data of chloroplast and Internal transcribed spacer (ITS1) sequences.

Alignment of genomic region of chloroplast and nuclear genes resulted in a dataset with a total of 6210 character sites, among these 868 were constant, 4727 were variable but parsimony uninformative, and 4184 were parsimony informative. The analysis of the combined ITS and chloroplast (*mat*K, *rbcL, psbK-psbI*, *trnlC-trnlF, psbA-trnH, atpF-atpH*) sequence data resulted in equally parsimonious trees of 2606 steps (number of parsimony-informative characters 4184; CI 0.82; RI 0.90; RC 0.75). *Z. officinale* and *Z. zerumbet* are used as outgroup and their sequences retrieved from GenBank.

The maximum likelihood (ML) tree is shown in Fig. 4.13 and strict consensus most maximum parsimonious (MP) tree is shown in Fig.4.14.

Both the ML and MP analysis the species were grouped in two groups, the group one of ML tree included three clades, with *C. inodora* and *C. kudagensis* clustered together with 100% bootstrap support. On the other hand the species *C. oligantha* var *lutea* showed strong bootstrap support with this clade (Fig. 4.13). In the ML and MP analysis, the species *C. aromatica* and *C. zedoaria* were clustered as a single clade with 100% bootstrap value. The same result was obtained using six chloroplast regions. Group two clustered with three species (*C. aeruginosa, C. pseudomontana* and *C. vamana*) in both ML and MP tree with strong bootstrap support value. The species *C. bhatii* exist as monoclade in both ML and MP analysis (Fig. 4.13 & 4.14).

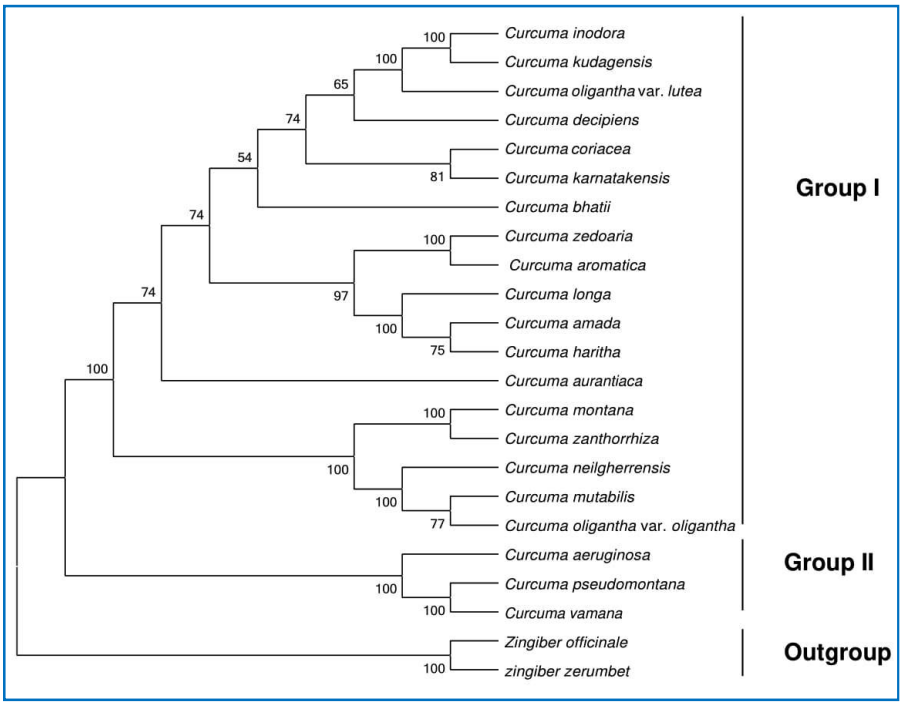


Fig. 4.13: Maximum likelihood phylogeny of the studied members of the genus *Curcuma* based on combined chloroplast and nuclear rDNA sequences data. Numbers beneath nodes are Bootstrap support (BS) indices.

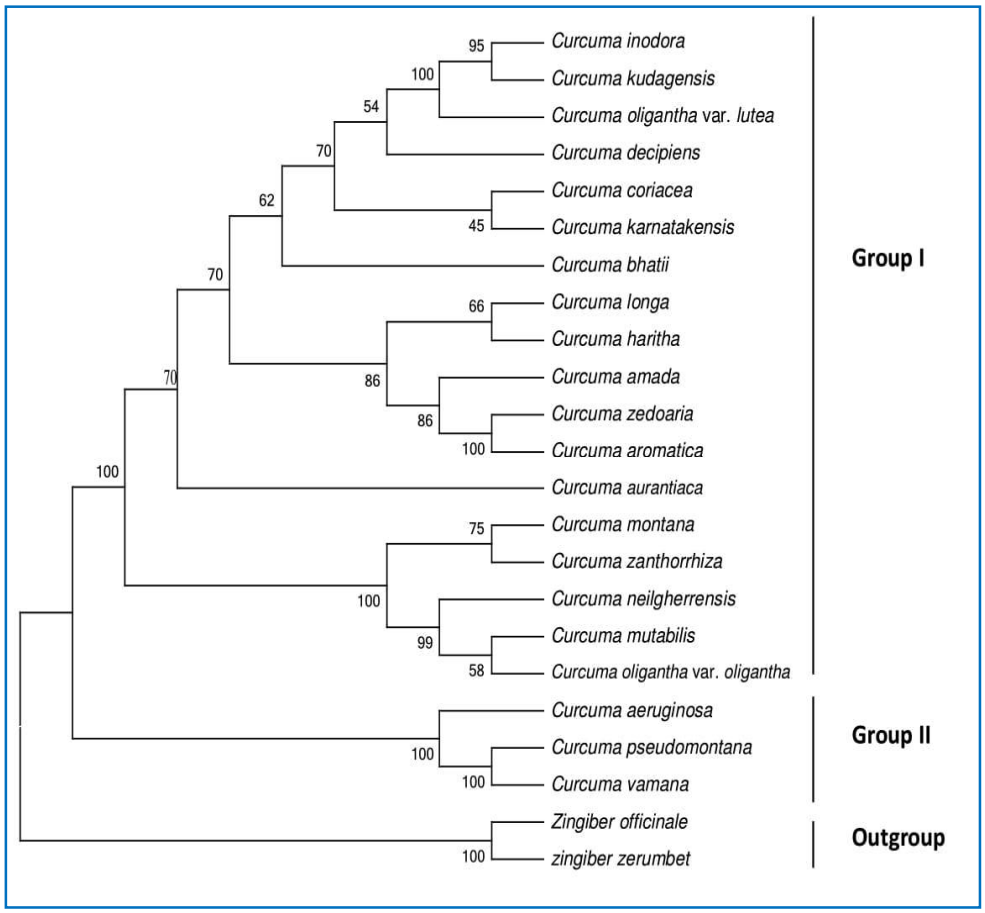


Fig. 4.14: Strict consensus of equally most parsimonious tree of 21 *Curcuma* species generated from sequences of combined chloroplast and nuclear rDNA. Numbers beneath nodes are Bootstrap support (BS) indices

4.5 Discussion

4.5.1 Circumscription of *Curcuma*

Earlier infrageneric classifications of *Curcuma* were based on morphological characters using position of inflorescence that is confusing in many species (Skornickova, 2007). Molecular phylogenetic studies showed that *Curcuma* is a paraphyletic genus (Kress *et al.,* 2002; Ngamriabsakul *et al.,* 2003; Zaveska *et al.,* 2012). Ngamriabsakul *et al.* (2003) suggested that the genus *Curcuma* is paraphyletic in origin based on the hybridization and

polyploidization of the species and all the species were well nested in the *Curcuma* complex and also reported that representatives of the *Curcuma*-like genera are derived from *Curcuma* species and formal transfer of these species into *Curcuma*, which should be recognized in a broad sense (*Curcuma* s .l.).

Several studies grouped *Curcuma* as infrageneric groups in *Curcuma* (Roxburgh, 1810; Horaninow, 1862; Baker, 1890; Valeton, 1918; Velayudhan *et al.,* 1996, 1999) with two subgenera, subg. *Curcuma* and subg. *Hitcheniopsis* (Baker) K. Schum. (Schumann, 1904) which was currently accepted. The current status of information about *Curcuma* species is based on broad geographical sampling but none of the earlier attempts were suitable for the classification of *Curcuma* species. Most of the classifications proposed by the earlier workers were based only on materials collected from the western part of the distribution range of the genus, whereas omitting several morphological character types from Indochina (Zaveska *et al.,* 2012). In addition, the morphological characters employed for the classifications such as position of the inflorescence and presence or absence of anther spurs is confusing in many species (Skornickova, 2007). Polyploidy was reported in the Indian representatives of *Curcuma* subg *Curcuma* (Leong-Skornickova *et al.,* 2007). These polyploidy accounts for creating long-standing taxonomic problems in this group (Skornickova, 2007).

Morphological delimitation of the genus *Curcuma* is more challenging because this large genus contains many extremely specialized and morphologically advanced species. No character has been found that is both exclusive to *Curcuma* and present in all the species. However, there is a set of characters that is common to majority of the species of *Curcuma* and includes basally connate refluxed bracts, well-developed lateral staminodes, versatile anthers and a cincinnus of two or more flowers. The previously reported phylogenetic studies using *mat*K gene, the species *C. aromatica* and *C.*

zedoaria are clustered together are congruent with our results (Santhoshkumar and Yusuf, 2017, 2018). Mainly molecular studies on phylogenetic relationships in *Curcuma* species were studied based on ITS data alone (Sirirugsa *et al.,* 2007). In plants, the chloroplast genes for maturase K (*mat*K), large subunit of the ribulose 1,5-bisphosphate carboxylase (*rbcL*), and *trnH-psbA* intergenic spacer are often used for molecular phylogenetic analysis (Hilu and Liang, 1997).

The species *C. inodora* is closely related with *C. decipiens* based on the morphological classifications (Sabu, 2006), but our studies focused on the molecular phylogenies using chloroplast and nuclear genes showed that *C .inodora* clustered with *C. kudagensis* and *C. vamana* closely resembled *C. burttii* but differs in leaf base, condensed spike and bracts, absence of anther-crest and presence of spurs on fertile anther (Sabu and Mangaly, 1988). These morphological results are dissimilar with molecular phylogeny and the *C. vamana* showed a close relationship with *C. pseudomontana* as these two species are clustered in a single clade. The co-habiting species, *C. decipiens* and *C. ecalcarata* (*C. aurantiaca*) showed much similarity in floral, vegetative and rhizome characters and anther spur length variation also cohabitate at high altitudes in rocky patches (Sabu, 1991). Our results showed these two species clustered in different clades.

Morphology based classification of the species *C. haritha* showed close relationship with *C. aromatica* and shares a few characters with *C. zedoaria,* however, *C. zedoaria* closely resembled *C. aeruginosa* in morphological character (Sabu, 2006). Based on the molecular phylogenetic analysis, using chloroplast and nuclear genes *C. aromatica* and *C. zedoaria* are close to each other. The present study showed the paraphyletic origin of these two species is moderate with (100%) bootstrap value in single as well as combined chloroplast and nuclear genes; however, *C. zedoaria* did not show

any relationship with *C. aeruginosa.* Morphologically close relationship was established between *C. karnatakensis* and *C. oligantha* (Amalraj *et al.,* 1999), where as these two species are grouped with different clade based on the chloroplast, mitochondrial and nuclear genes sequences. However, based on the phylogenetic tree using chloroplast, mitochondrial, nuclear sequences alone and the combination of chloroplast and nuclear genes showed both the species are placed in different clades.

The endemic and monotypic genus of *Paracautleya* was established by R.M. Smith in 1977. The type species of *Paracautleya bhatii* R.M.Sm. was described based on a collection by K.G. Bhat from the surroundings of Udupi, South India. Several characters were used to elucidate the similarity of the species with genus *Paracautleya* but unfortunately most of the characters did not support this species in the genus *Paracautleya* (Skornickova and Sabu, 2005). Hence the genus *Paracautleya* merged with *Curcuma* based on the similarity of morphological characters, now the species *Paracautleya bhatii* renamed as *Curcuma bhatii*, but the molecular phylogenetic results showed the species exist as a monoclade with lowest bootstrap support using chloroplast and ITS genes and same result were obtained by using combined tree chloroplast and nuclear ITS sequences.

The phylogenetic analysis were conducted using chloroplast and nuclear genes the chloroplast genes poorly resolved relationship with the plant species compared to nuclear ITS data (Zaveska *et al.,* 2012). The current studies, inconsistent with phylogenetic signals between datasets are often explained either by hereditary changes or by ancient hybridization (Okuyama *et al.,* 2005). The observed topology of the mDNA tree, despite its low support, reflects the geographical distribution of the species in a group and lesser mutation in the region (Fig. 4.6, 4.7).

Finally, the ITS and cpDNA phylogenetic relationships among twenty one species of *Curcuma* are congruent with our results. The most notable example in the case of *C. aromatica* and *C. zedoaria* that are grouped together in chloroplast phylogenies but these two species were clustered as a different clade in the ITS tree. The assortment of hybridization and the multiple-copy character of the ITS region means that several ITS repeat types can be maintained in the genome and or further undergo complete or incomplete homogenization (Alvarez and Wendel, 2003). Relationship of species within this group is very difficult to distinguish because polyploidization and hybridization is important for speciation (Leong-Skornickova *et al.,* 2007; Zaveska *et al.,* 2011). The hierarchical structure within the group dependant on the variation in genome size (Leong-Skornickova *et al.,* 2007).

4.5.2 Methods in tree construction

Several methods were used to study the phylogenetic analysis, the most common and important among them are the Maximum parsimony and Maximum likelihood methods. MP is one of the correct method that can easily determine the insertions or deletions of nucleotides and provide important phylogenetic information. Although MP is quite competent in obtaining the correct topology (Nei, 1996 & 1991), it may also give incorrect trees when the rate of nucleotide substitution is fairly constant among the taxa (Takezaki *et al.,* 1994). Cavalli-Sforza & Edwards (1967) was first developed the idea of using an ML method for phylogenetic analysis for gene frequency data, but they encountered a number of problems in implementing the method.

This ML method refined as a separate one compared to parsimony method and it was specifically developed by using molecular data called maximum likelihood (Felsenstein, 1981). It seeks to find the trees that yield the maximum likelihood value to an observed data set on the explicit model

of evolution. This method is considered to be more versatile than parsimony to analyze the sequences because it evaluates all characters and can accommodate other assumptions, such as different rates and patterns of substitution. However, it can give inconsistent results (Siddall, 1998) and is computationally demanding. Because of these reasons we chose these two methods were constricting phylogenetic tree and compare the topologies of each separate tree.

4.5.3 Genomes for phylogenetic anlysis

Phylogenetic trees are important tools for organizing knowledge of biological diversity, and they hypothesized the evolutionary relationships among nested groups of taxa (monophyletic groups) that are supported by shared traits known as *synapomorphies*. The predicted phylogenetic trees are a type of graphic diagram that illustrates theoretical concepts rather than appearances of objects (iconic diagrams) or quantitative relationships (Lee, 2010).

The plant cells have multiple copies of chloroplast genomes (Pyke, 1999), analysis of total genome sequencing data are useful for both nuclear marker achievement as well as for the assembly of chloroplast genomes (Nock *et al.,* 2011). Generally, in plants the mitochondrial genome evolves at a slowest rate, the chloroplast genome at a slightly faster rate and the nuclear genome at the fastest rate (Wolfe *et al.,* 1987). The conventional evolution of the chloroplast genome provides different advantages in phylogenetic studies compared to mitochondria (Palmer, 1985) and nuclear (Flavell, 1980) genomes of plants and which is much more dynamic and diverse. Researchers in the ancient times have used numerous noncoding cpDNA regions to obtain adequate characters for phylogenetic resolution (Schonenberger *et al.,* 2003). The chloroplast DNA (cpDNA) gene sequences (e.g. *rbcL, atpB, mat*K and *ndhF*) have been widely used at the family level and above (Soltis *et al.,*

1999), whereas non-coding sequences such as introns (e.g. *rpL16, rpoC1, rpS16, trnL, trnK*) and intergenic spacers (e.g. *trnT–trnL, trnL–trnF, atpB–rbcL, psbA–trnH*) are used more frequently at lower taxonomic levels (Taberlet *et al.,* 1991).

The cpDNA genomes harbors such limitations, as well as to obtain additional and independent estimates of phylogeny, nuclear rDNA has been generally accepted as a tool for studying plant systematics. Slowly evolving rRNA genes are used for higher taxonomic levels (Soltis and Soltis 1998), although lower taxonomic levels internal and external intergenic spacers are more commonly employed (Bailey *et al.,* 2003).

Mitochondrial genome is not compatible for plant phylogenetic analysis because of its high frequency of rearrangements, gene order and content of the plant mitochondrial genome is poorly conserved across the plants (Knoop *et al.,* 2011). In addition, horizontal gene transfer is one of the characteristic features for molecular genomes, plant mitochondrial genomes are more commonly compared to plastid and nuclear genomes (Xi *et al.,* 2013). On the other hand RNA editing in mitochondria is not limited to compare protein coding sequences, but is also evident for tRNAs, introns, and 5' and 3' untranslated sequences. For this reason, even if working with mitochondrial sequences, comparison at the cDNA or peptide level genomes is likely to produce more accurate results. However, mitochondrial genes including cytochrome oxidase subunits (Sha *et al.,* 2014), NADH dehydrogenase subunits (Sanjur *et al.,* 2002), atpA (alpha subunit of mitochondrial ATP synthase) (Seberg *et al.,* 2012), matR (intron encoded maturase R), small subunit (19S) ribosomal DNA (Duff and Nickrent, 1999) and rps genes that encode ribosomal proteins, have been used for resolving plant phylogenies. In addition, researchers reported the employment of huge sets of plant mitochondrial genes in evolutionary studies (Liu *et al.,* 2014).

Based on the previous study we used two mitochondrial regions such as *cox1* and *atpA* for studying the evolutionary relationship of *Curcuma* species. Here also the species *C. bhatii* previously reported *Paracautleya bhatii* exist as a monoclade with >50% bootstrap value (Fig. 4.6). The same results obtained from when we use chloroplast and nuclear sequences.

In plant systematics the impact of molecular data can hardly be overstated. In combination with the unambiguous methods for phylogenetic analysis, molecular data have reshaped concepts of relationships and circumscriptions at all levels of the taxonomic hierarchy (Crawford, 2000). Molecular markers used in character based identification of species might draw from the chloroplast, mitochondrion, and nucleus (Savolainen and Chase, 2003). Estimating the phylogenetic relationship of particular organisms, the combined sequence data from two or more loci is used as a powerful method. Compared to single gene analysis combining sequences provide improved resolution and internal support of phylogenies (Soltis and Soltis, 1998). The two locus combination of ribulose-1, 5-bisphosphate carboxylase oxygenase large subunit (*rbcL*) and maturase K (*mat*K) genes are recommended by the consortium for the Barcode of Life (CBOL) plant working group (CBOL Plant Working Group, 2009) for the identification of the species. Seven chloroplast genomic regions were evaluated by the Consortium for the Barcode of Life Plant Working Group and suggested a combination of *mat*K+*rbcl* for the identification and authentication of angiosperm. Many recent studies have proved combined *rbcL* and *mat*K sequence data to resolve problems at the subfamily and tribe levels in angiosperms (Soltis *et al.,* 1996). Many fragments of coding and non-coding regions, introns, and intergenic spacers, including *atpB, atpB-rbcL, mat*K*, ndhF, rbcL, rpl16, rps4-trnS, rps16, trnH-psbA, trnL-F, trnS-G*, *etc*., have been used for studying the phylogenetic reconstructions at various taxonomic levels (Peterson *et al.,* 2010). Several regions of the chloroplast genome, for

example, *atpF-H, mat*K*, psbK-I, rbcL, rpoB, rpoC1* and *trnH-psbA* have been relied upon heavily for the development of candidate markers for plant DNA barcoding (Kress *et al.,* 2005; Hollingsworth *et al.,* 2011).

For this study we used the nuclear *ITS* marker since it had previously used effectively for resolving phylogenetic relationships in the family Zingiberaceae (Kress *et al.,* 2005). In addition, we analyzed six coding and non-coding regions of cpDNA (*rbcL , atpF-H, trnL-trnF, psbA-trnH, mat*K and *psbA-psbK*). A number of *Curcuma* species are projected to be polyploid or hybrid in origin (Leong-Skornickova *et al.,* 2007), the combination of a biparentally inherited rDNA region with maternally inherited cpDNA was selected for its potential to distinguish ancient species and hybrid speciation, as well as ancestry of polyploids (Alvarez and Wendel, 2003). In the present study we used chloroplast, mitochondrial and nuclear regions to compare the evolutionary relationship of *Curcuma* species collected from different region of South India based on previously reported markers genes. On the other hand we found some discordance between plastid and nrDNA gene trees. For example, in the chloroplast genome analysis (Fig. 4.2, 4.3), *C. inodora* grouped with *C. kudagensis* and *C. oligantha* var *lutea* (BS = 100), whereas in the nuclear analysis (Fig. 4.10 & 4.11**)** it was closely related with *C. mutabilis* and *C. decipiens*. This type of incongruence can be explained by reticulate evolution, which has been confirmed in earlier studies (Peng and Wang, 2008).

CHAPTER 5
PHYTOCHEMICAL ANALYSIS

5.1 GC-MS based chemical profiling

The *Curcuma* species used in the present study were collected from different regions of South India and cultivated under identical conditions so as to rule out any variability in the chemical profile of methanolic extracts due to environmental factors. Unpeeled rhizomes were used to make rhizome extracts. The colour of most of the *Curcuma* rhizomes was light blue, yellow or gray, while color of methanolic extracts was brownish or yellow. The color of the rhizomes extracts did not correlate with the color of rhizome and little information is available about the constituents of rhizome extracts of *Curcuma* species.

The composition of the methanolic extracts was determined using GC–MS and individual components were identified by comparing the relative retention indices of the peaks with the NIST library of standard extracts. The number of chemical constituents of individual species ranged from 6 to 35. The maximum number of compounds was detected in *C. aeruginosa, C. aromatica C. decipiens, C. longa, C. zanthorrhiza* and *C. zedoaria* rhizome extracts. The identified major components of the rhizome extracts of *Curcuma* species are given below:

C. aeruginosa

The GC-MS profile of the rhizome extract from the species *C. aeruginosa* is represented in (Fig. 5 A**)**. A total of 28 major components were identified viz, butyl 9, 12-octadecadienoate (16.17%), (-)-isolongifolene (4.8%), camphor (2.3%), eucalyptol (1.8%), isovelleral (5.4%), (E)-.beta.-

famesene (5.3%), (+)-isocurcumenol (4.3%) beta-elemen-(2) (2.2%), 4-[(3,4-dimethylphenyl)sulfanyl]aniline (11.3%), 1,5,9-trimethyl-1,5,9-cyclododecatriene (23%) and isopulegol 2 (1.2%).

C. amada

Seventeen compounds were identified form the GC-MS chromatogram of rhizome extract of *C. amada.* The major components are beta-caryophyllene oxide (75%), beta myrcene (9.4%), acetic acid, 1-[2-(2,2,6-trimethyl-bicyclo[4.1.0]hept-1-yl)-ethyl]-vinyl ester (2.6%), alpha.-springene (1.5%) and card-20(22)-enolide, 3,5,14,19-tetrahydroxy-8-methyl-, (3β,5β,9β) (1.8%) respectively (Fig. 5 B).

C. aromatica

The GC-MS analysis of the methonalic extract of *C. aromatica* rhizome detected 27 compounds represented in Fig. 5 C. curdione (20.5%), isovelleral (15.3%), trimethyl-1, 5, 9-cyclododecatriene (13.9%), camphor (7.9%), germacrone (3.1%), borneol (4.6), 9-cis retinal (2.4%) and isolongifolene (3.3%) are major components.

C. bhatii

Thirteen chemical compounds were identified from the rhizome extracts of *C. bhatii*; the major ones were beta-caryophyllene oxide (87%), kauren-18-ol, acetate, (4beta) (2.4%), retinal (1.8%) and duvatriendiol (1.5%) identified in the GC-MS profile (Fig. 5 D).

C. coriacea

Analysis of *C. coriacea* rhizomes detected 17 compounds; the major components were 1-heptatriacotanol (39.9%), (E)-.beta.-famesene (19.9%), (-)-β-caryophyllene epoxide (12.3%), 1S,2E,4S,5R,7E,11E)-cembra-2,7,11-trien-4,5-diol (10%), corymbolone (3.3%), eucalyptol (3.3%) and

tricyclo[20.8.0.0(7,16)]triacontane, 1(22),7(16)-diepoxy- (2.5%) the GC-MS profile is represented in Fig. 5 E.

C. decipiens

The GC-MS analysis of *C. decipiens* rhizome yielded 28 compounds; kauren-18-ol, acetate, (4.beta.)- (53.8%), (-)-β-caryophyllene epoxide (14.4%), 9-cis retinal (11.4%), kauren-19-ol-acetate (2.6%) and acetic acid, 1-[2-(2, 2, 6-trimethyl-bicyclo [4.1.0] hept-1-yl)-ethyl]-vinyl ester (1.7%) are the major components. The GC-MS profile is represented in Fig. 5 F.

C. aurantiaca

Twenty four compounds were detected from the rhizome extract of *C. aurantiaca*; the major components are 3-Terpinolene (30.7%), dihydrochrysin (18.3%), 3,3-dimethyl-4-phenyl-4-penten-2-one (15%), eucalyptol (8.3%), trans-β-caryophyllene (6.5%), E-nerolidol (4.1%) and camphor (3.5%) as represented in the GC-MS profile (Fig. 5 G).

C. haritha

The major chemical compounds identified from the rhizome extracts of *C. haritha* were isoaromadendrene epoxide (27.8%), (-)-β-caryophyllene epoxide (26.6%), isovelleral (12.3%), 2,4-di-tert-butylphen (7.3%), camphor (4.5%), 2-bornanol (3.5%), 2-camphanol, exo (2.7%) and diazoprogesterone (2.4%). A total of 20 compounds were identified using GC-MS and is represented in (Fig. 5 H).

C. inodora

C. inodora rhizome extract were subjected to GC-MS analysis; a total of 22 compounds were detected, alloaromadendrene oxide-(1) (53.3%), 9-cis retinal (11.3%), (+,-)-β bisabolene (7.7%), andrographolid

(5.7%),2,4a,8,8tetramethyldecahydrocyclopropa[d]naphthalene (2.9%) and (+)-alpha-santalene (1.4%) are the major components. The peaks and the obtained chromatogram are represented in Fig. 5 I.

C. karnatakensis

Rhizome of *C. karnatakensis* detected 19 compounds; the major components were alloaromadendrene oxide (50.8%), 9-cis retinal (11.7%), (-)-β-caryophyllene epoxide (8.3%), 3-bromocholest-5-ene (6.1%), 10-chlorotricyclo [4.2.1.1(2, 5)] deca-3, 7-dien-9-ol (4.5%), kauren-18-ol, acetate, (4.beta.)- (4.5%), 1-ethyl-4,4-dimethyl-cyclohex-2-en-1-ol (3.0%) and ergost-25-ene-6,12-dione, 3,5-dihydroxy-, (3.beta.,5.alpha.)- (2%). The GC-MS profile depicted in Fig. 5 J.

C. kudagensis

Fifteen compounds were detected using the GC-MS analysis of the species *C. kudagensis*; the following compounds being the major components as, kauren-18-ol, acetate, (4.beta.) (38.3%), 9-cis retinal (18%), (-)-β-caryophyllene epoxide (17.3%), isovelleral (4.7%), acetic acid, 1-[2-(2, 2, 6-trimethyl-bicyclo [4.1.0] hept-1-yl)-ethyl]-vinyl ester (3.9%), 2-24-nitro butanoyl 7 cyclohexanone (3.4%) and kauren-19-ol-Acetate (1.3%) the chromatogram are repersented in Fig. 5 K.

C. longa

Rhizome extract of *C. longa* yielded 31 compounds, tumerone (19.1%), (-)-zingiberene (11.6%), beta-sesquiphellandrene (8.8%), (-)-.alpha.-Santalene (8.6%), Curlone (6.4%), ar-tumerone (6.1%), β-bisabolene (5.5%), benzenemethanol, 4-methyl-.alpha.-(1-methyl-2-propenyl)-, (R*,R*) (4.7%), (E)-.beta.-famesene (4%) and butane-1,1-dicarbonitrile, 1-cyclohexyl-3-

methyl-(2.6%) as the major compounds from the extract as shown in the profile Fig. 5 L.

C. montana

Totally 6 components were identified in the GC-MS analysis of the rhizomes of *C. montana;* the major components are alloaromadendrene oxide (60.7%), 2-tert-butyl-4,6-bis(3,5-di-tert-butyl-4-hydroxybenzyl)phenol (12.8%), violanthrene a (11.4%) and (-)-β-caryophyllene epoxide (6.6%) respectively (Fig. 5 M).

C. mutabilis

Fifteen components were identified from the methanolic extract of *C. mutabilis*; the major components were Kauren-18-ol, acetate, (4.beta.)- (59.1%), (-)-β-caryophyllene epoxide (17.2%), retinal (8.1%), trispiro [4.2.4.2.4.2.]heneicosane (3.0%), kauren-19-ol-Acetate (3.0%), 1-ethyl-4,4-dimethylcyclohex-2-en-1-ol (2.4%) and aromadendrene oxide-(2) (1.3%). The GC-MS profile is represented in Fig. 5N.

C. neilgherrensis

C. neilgherrensis rhizome extract subjected to GC-MS analysis identified a total of 15 components; the major components were 1-heptatriacotanol (29.9%), 9 cis retinal (23.3%), aromadendrene oxide-(1) (17.6%), methyl (4,4-dimethyl-2,4,5,6-tetrahydro-1H-inden-2-yl)acetate (5.0%),(-)-β-caryophyllene epoxide (3.7%) andrographolide (3.1%) and 1,5-epoxy-salvial(4)14-en (2.4%) (Fig .5 O).

C. oligantha* var. *oligantha

The rhizome extract of *C. oligantha* var. *oilgantha* were subjected to GC-MS analysis and 26 components were identified retinal (34.1%), 3-

bromocholest-5-ene (21.2%), (-)-β-caryophyllene epoxide (8.0%), kauren-18-ol, acetate, (4.beta.)- (7.9%), 4-hexen-1-ol, 6-(2,6,6-trimethyl-1-cyclohexenyl)-4-methyl-, (E)-(3.2%), (+)-carotol (2.2%) and camphor (1.4%) are the major components. The obtained chromatogram represented in Fig. 5 P.

C. pseudomontana

Ten components were detected in GC-MS analysis of *C. pseudomontana* rhizome extracts; the major components are kauren-18-ol, acetate, (4.beta.)- (54.0%), 1-heptatriacotanol (17.4%), (-)-β-caryophyllene epoxide (10.9%) and 9 cis retinal (9.5%). The GC-MS profile is represented in Fig. 5 Q.

C. zedoaria

The rhizome extract of *C. zedoaria* produced 35 components such as, isovelleral (23.7%), germacrone (10.7%), isogermafurene (6.8%), aromadendrene oxide-(1) (6.5%), camphor (5.3%), 1-heptatriacotanol (4.3%), beta-elemene (3.6%), (-)-beta-caryophyllene (3.3%), germacrene B (2.6%) and 6.beta.-hydroxymethandienone (2.1%). The major components are represented in Fig. 5 R.

C. vamana

GC-MS analysis of *C. vamana* rhizome detected the following compounds as represented in the chromatogram (Fig 5S). Totally 26 components were identified; the major compounds are stahlianthusone (27.7%), copaene (15.5%), 9,10-dimethyl-1,2,3,4,5,6,7,8-octahydroanthracene (11.4%), Camphor (11.2%), aromadendrene (7.3%), coumarine, 8-allyl-7-hydroxy-6-ethyl-4-methyl (4.0%) and (-)-β-caryophyllene epoxide (2.9%).

C. zanthorrhiza

Thirty components were identified using GC-MS analysis from the rhizome extract of *C. zanthorrhiza* (Fig. 5T); the major components were 2-methyl-5-(1,2,2-trimethyl cyclopentyl)phenol (25.8%), beta.-curcumene (21.0%), alpha-curcumene (17.6%), N-(4-hydroxyphenyl)-N,N',N'-trimethylsulfamide (4.8%), germacrone (4.1%) and camphor (3.8%).

5.1.1 Cluster analysis

The major components were selected for the relationship analysis between the species. The components selected were above ≥5% concentration. The major and common components identified by GC–MS analyses are presented in Table 5.1. The fractions obtained from the rhizome extracts were scored for the presence or absences of specific components and the data were analyzed using NTSYS-pc version 2.1 software. Cluster analysis was carried out using UPGMA (Unweighted Pair Group Method with Arithmetic averages) for the compositions of 20 rhizome methanolic extracts to determine the pattern of identical and diverse chemical compounds between different species (Fig. 5.1). The data also revealed that camphor, germacrene D, aromadendrene oxide, β -caryophyllene epoxide, 1-heptatriacotanol, germacrone, retinal, Isovelleral, kauran-18-al, 17-(acetyloxy)-, (4.beta.) and β-caryophyllene were the common compounds in most of the species. The identified unique components from all the species are listed in Table 5.1.

Table 5.1: Major and Common Components identified from Twenty *Curcuma* species using GC-MS analysis

Components Name	Species Name Percentage quantity	Components	Species Name (Quantity)
β-caryophyllene	*C. decipiens* (3.8%) *C. aurantiaca* (6.5%) *C. inodora* (1.3%) *C. zedoaria* (1.1%)	1-heptatriacotanol	*C. coriacea* (39.9%) *C. aurantiaca* (1.0%) *C. neilgherrensis* (29.5%) *C. pseudomontana* (17.4%) *C. zedoaria* (4.3%)
β-caryophyllene epoxide	*C. amada* (75.7%) *C. bhatii* (87.0%) *C. coriacea* (12.3%) *C. decipiens* (14.4%) *C. aurantiaca* (1.4%) *C. harith* (26.6%) *C. karnatakensis*(8.3%) *C. kudagensis* (17.3%) *C. longa* (0.8%) *C. montana* (6.6%) *C. mutabilis* (17.2%) *C. neilgherrensis* (3.7%) *C. oligantha* (8.0%) *C. pseudomontana* (10.9%) *C. zedoaria* (0.4%) *C. vamana* (2.9%)	beta-bisabolene	*C. coriacea* (5.1%) *C. amada* (1.4%) *C. decipiens*(1.5%) *C. inodora* (7.7%) *C. karnatakensis*(0.3%) *C. longa* (5.5%) *C. oligantha* (1.6%)
β-myrcene	*C. amada* (9.4%)	kauren-18-ol, acetate, (4.beta.)-	*C. bhatii* (2.4%) *C. decipiens* (53.8%) *C. karnatakensis* (4.5%) *C. kudagensis* (38.3%) *C. mutabilis* (59.1%) *C. oilgantha* (7.9%) *C. pseudomontana* (54.0%) *C. vamana* (0.9%)
curdione	*C. aromatica* (20.5%) *C. zedoaria* (1.6%)	retinal	*C. bhatii* (1.8%) *C. decipiens* (11.4%) *C. karnatakensis* (11.7%) *C. kudagensis* (18.0%) *C. mutabilis* (18.1%) *C. oilgantha* (34.1%) *C. pseudomontana* (9.5%) *C. amada* (0.4%) *C. aromatica* (2.4%)

			C. coriacea (1.7%) *C. inodora* (11.3%) *C. neilgherrensis* (23.3%)
isovelleral	*C. aeruginosa* (5.4%) *C. aromatica* (15.3%) *C. haritha* (12.3%) *C. zedoaria* (23.7%) *C. zanthorrhiza* (0.7%)	juniper camphor	*C. decipiens* (0.6%) *C. inodora* (1.7%) *C. zedoaria* (0.7%)
andrographolide	*C. aromatica* (1.0%) *C. inodora* (5.7%) *C. neilgherrensis* (3.1%)	alloaromadendrene oxide-1	*C. aeruginosa* (0.6%) *C. amada* (0.9%) *C. coriacea* (0.5%) *C. decipiens* (1.1%) *C. inodora* (53.3%) *C. karnatakensis*(50.8%) *C. montana* (60.7%) *C. mutabilis* (1.3%) *C. neilgherrensis* (17.6%) *C. zedoaria* (6.5%) *C. vamana* (0.7%)
tumerone	*C. decipiens* (0.3%) *C. longa* (19.1%)	zingiberene	*C. zanthorrhiza* (1.8%) *C. longa* (11.6%) *C. decipiens* (0.2%)
(E)-.beta.-farnesene	*C. aeruginosa* (11.6%) *C. coriacea* (19.9%) *C. aurantiaca* (0.7%) *C. longa* (4.0%) *C. oilgantha* (0.4%) *C. zanthorrhiza* (2.9%)	camphor	*C. aeruginosa* (2.3%) *C. aromatica* (7.9%) *C. aurantiaca* (3.5%) *C. decipiens* (0.4%) *C. haritha* (4.5%) *C. inodora* (0.7%) *C. karnatakensis* (0.5%) *C. oilgantha* (1.4%) *C. zedoaria* (5.3%) *C. vamana* (11.2%) *C. zanthorrhiza* (3.8%)
germacrene B	*C. aeruginosa* (0.8%) *C. zedoaria* (2.6%) *C. zanthorrhiza* (2.2%)	beta-elemene	*C. aeruginosa* (2.2%) *C. aromatica* (2.2%) *C. zedoaria* (3.6%) *C. zanthorrhiza* (0.5%) *C. haritha* (0.6%)
beta-curcumene	*C. longa* (0.3%) *C. zanthorrhiza* (21.0%) *C. oligantha* (0.8%)	germacrene D	*C. aeruginosa* (1.6%) *C. aromatica* (0.6%) *C. zedoaria* (0.7%) *C. zanthorrhiza* (0.6%)
germacrone	*C. aeruginosa* (1.7%)	(+)-delta-cadinene	*C. aurantiaca* (1.2%)

	C. aromatica (3.1%) *C. zedoaria* (10.7%) *C. zanthorrhiza* (4.1%)		*C. aromatica* (0.8%)
Ar- tumerone	*C. decipiens* (0.4%) *C. longa* (6.1%)	eucalyptol	*C. aeruginosa* (1.8%) *C. coriacea* (3.3%) *C. aurantiaca* (8.3%) *C. longa* (0.5%)

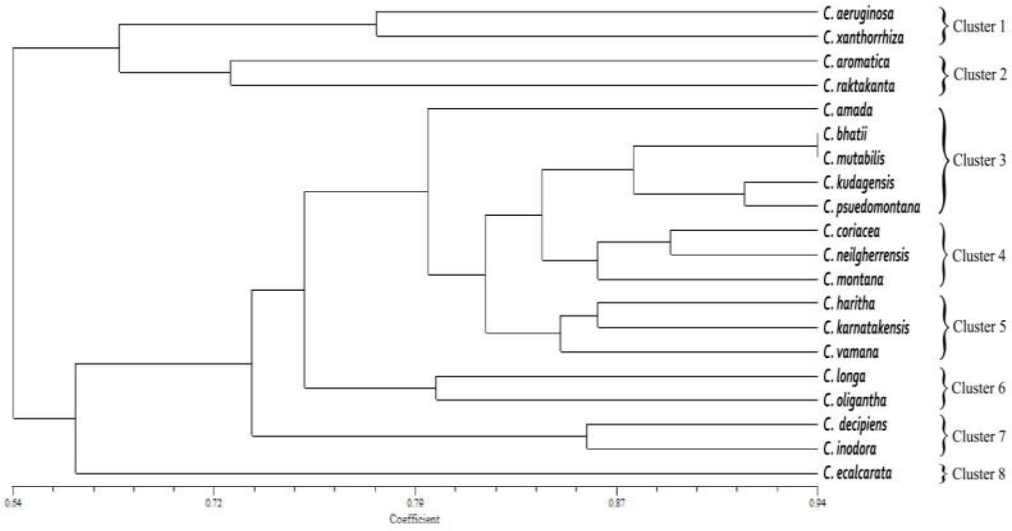
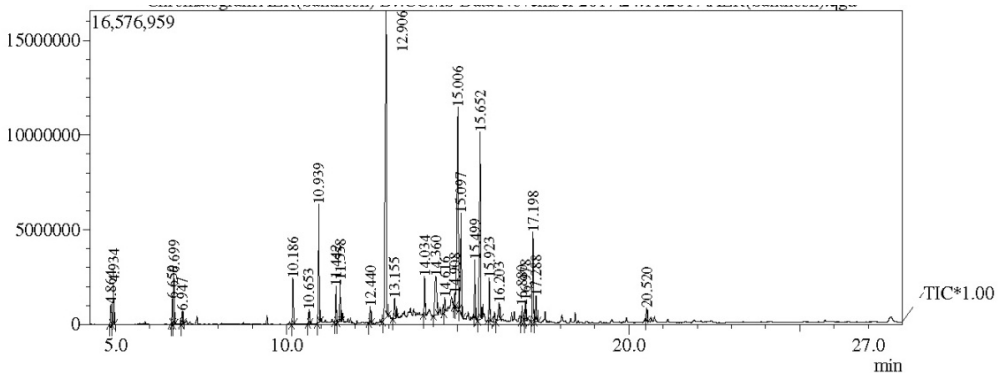


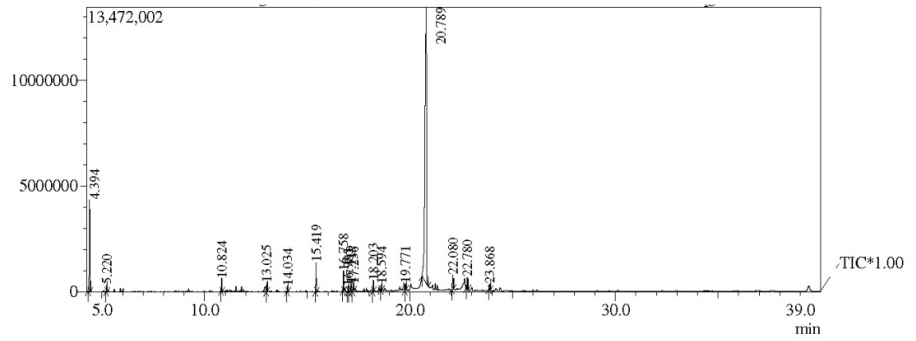
Fig. 5.1. UPGMA derived dendrogram of *Curcuma* species based on GC/MS analysis of methanolic rhizome extracts. The numerical scale indicates genetic similarity. Species were scored for the presence (1) or absence (0) of individual essential oil components with ≥5% concentration. Data analyzed using NTSYS – pc version 2.1 software and dendrogram generated based on the UPGMA cluster. The numerical scale indicates genetic similarity and the members of specific cluster/groups are genetically related with similar/near coefficient value



Peak Report TIC

Peak#	R.Time	Area	Area%	Height	Height%	Name	Base m/z
1	4.864	2192111	1.17	1040469	1.21	ISOPULEGOL 2	81.10
2	4.934	3541646	1.89	2210943	2.58	Eucalyptol	81.10
3	6.659	2360209	1.26	1327026	1.55	Spirobicyclo[2.2.1]heptane-2,2'-[1',3'-dioxa-2'-oxocyclohex-5'-ene)], 1,6',7,7-tetramethyl-	95.10
4	6.699	4346692	2.32	2666250	3.11	Camphor	95.10
5	6.947	1115576	0.60	670722	0.78	Isoborneol	95.10
6	10.186	4275327	2.29	2433676	2.84	BETA-ELEMEN-{2}	81.10
7	10.653	1315652	0.70	749282	0.87	Caryophyllene	93.10
8	10.939	10058682	5.38	6228293	7.26	(E)-.beta.-Famesene	69.05
9	11.442	3046362	1.63	1894433	2.21	(-)-GERMACRENE D	161.10
10	11.558	3544805	1.90	2220050	2.59	Isogermafurene	108.05
11	12.440	1536928	0.82	879068	1.03	GERMACRENE B	121.10
12	12.906	43001822	23.00	16238577	18.94	Phenol, 4-[[(dimethylamino)sulfonyl]methylamino]-	122.05
13	13.155	2585849	1.38	1039261	1.21	Acetic acid, (1,2,3,4,5,6,7,8-octahydro-3,8,8-trimethylnaphth-2-yl)methyl ester	105.10
14	14.034	3272356	1.75	2020200	2.36	(E,E)-Germacrone	107.10
15	14.360	8061399	4.31	2028864	2.37	(+)-ISOCURCUMENOL	105.10
16	14.616	1194549	0.64	779496	0.91	2,3,3-Trimethyl-2-(3-methylbuta-1,3-dienyl)-6-methylenecyclohexanone	163.05
17	14.908	1524400	0.82	673274	0.79	Spathulenol	162.10
18	15.006	30237191	16.17	10643234	12.41	Butyl 9,12-octadecadienoate	67.05
19	15.097	10237709	5.48	5323845	6.21	Isovelleral	232.10
20	15.499	4296385	2.30	3030660	3.53	ELEMENE	68.05
21	15.652	21874003	11.70	9909274	11.56	1,5,9-Cyclododecatriene, 1,5,9-trimethyl-	68.05
22	15.923	3776188	2.02	2261243	2.64	5,8-Dihydroxy-4a-methyl-4,4a,4b,5,6,7,8,8a,9,10-decahydro-2(3H)-phenanthrenone	174.10
23	16.203	2604482	1.39	924216	1.08	VALERENAL	228.05
24	16.880	1934368	1.03	814183	0.95	2-Isopropenyl-4a,8-dimethyl-1,2,3,4,4a,5,6,7-octahydronaphthalene	109.05
25	16.978	1918683	1.03	1075591	1.25	S-OCTAHYDROANTHRACENE	158.05
26	17.198	9131719	4.88	4687837	5.47	(-)-ISOLONGIFOLENE	175.05
27	17.288	2775674	1.48	1347219	1.57	3-Oxomanoyl oxide	81.10
28	20.520	1228920	0.66	616920	0.72	Alloaromadendrene oxide-(1)	81.10
		186989687	100.00	85734106	100.00		

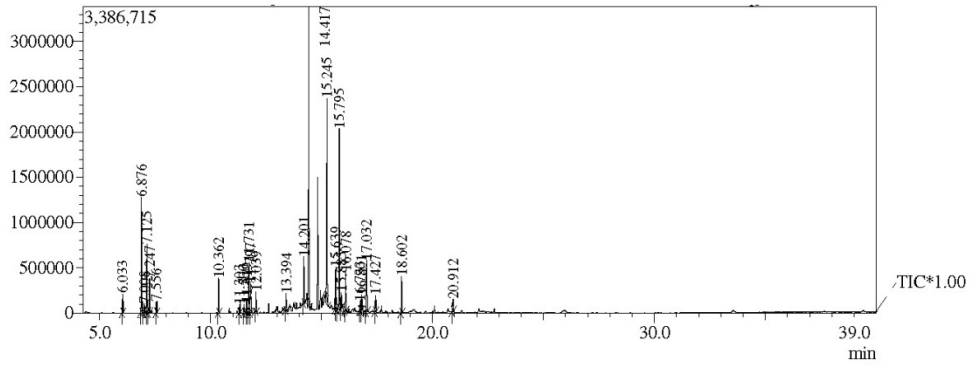
Fig. 5 A. GC-MS Chromatograms of methanolic extract of *C. aeruginosa* rhizome.



Peak Report TIC

Peak#	R.Time	Area	Area%	Height	Height%	Name	Base m/z
1	4.394	8389799	9.43	4331450	17.42	.beta.-Myrcene	93.10
2	5.220	536917	0.60	320673	1.29	.beta.-Ocimene	93.10
3	10.824	907640	1.02	606368	2.44	Caryophyllene	93.10
4	13.025	639764	0.72	423699	1.70	Phenol, 4-[[(dimethylamino)sulfonyl]methylamino]-	122.05
5	14.034	404088	0.45	273925	1.10	Spiro[4.5]dec-6-en-8-one, 1,7-dimethyl-4-(1-methylethyl)-	109.05
6	15.419	2369007	2.66	1388702	5.58	Acetic acid, 1-[2-(2,2,6-trimethyl-bicyclo[4.1.0]hept-1-yl)-ethyl]-vinyl ester	137.10
7	16.758	1343275	1.51	948761	3.82	(E,E,E)-.alpha.-Springene	69.10
8	16.985	429078	0.48	261132	1.05	.ALPHA.-BISABOLENE	137.10
9	17.119	510385	0.57	352098	1.42	BETA-BISABOLENE	69.10
10	17.236	653010	0.73	337559	1.36	4,8,13-Duvatriene-1,3-Diol	81.10
11	18.203	863571	0.97	496796	2.00	Alloaromadendrene oxide-(1)	81.10
12	18.594	438261	0.49	272422	1.10	RETINAL	147.10
13	19.771	575663	0.65	315004	1.27	4-Hexen-1-ol, 6-(2,6,6-trimethyl-1-cyclohexenyl)-4-methyl-, (E)-	137.10
14	20.789	67399474	75.72	13066225	52.54	(-)-.BETA.-CARYOPHYLLENE EPOXIDE	81.05
15	22.080	1650244	1.85	668142	2.69	Card-20(22)-enolide, 3,5,14,19-tetrahydroxy-, (3.beta.,5.beta.)-	55.00
16	22.780	1199860	1.35	532564	2.14	PREGNANE DIOL DIMETHYLSILYL ETHER	103.05
17	23.868	702221	0.79	273510	1.10	12,12-DICHLOROBICYCLO[8.2.0]DODECAN-11-ONE	98.05
		89012257	100.00	24869030	100.00		

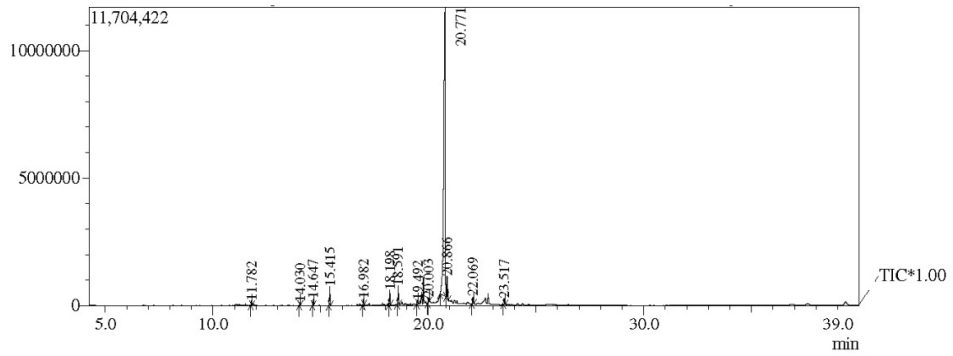
Fig. 5 B. GC-MS Chromatograms of *C. amada* rhizome methanolic extract.



Peak Report TIC

Peak#	R.Time	Area	Area%	Height	Height%	Name	Base m/z
1	6.033	303886	1.16	210266	1.36	.BETA.-LINALOOL	71.05
2	6.876	2092548	7.96	1277705	8.28	Camphor	95.10
3	7.008	131282	0.50	97068	0.63	EXO-METHYL-CAMPHENILOL	71.05
4	7.125	1217823	4.63	754243	4.89	Borneol, exo-	95.10
5	7.247	592875	2.26	369358	2.39	endo-Borneol	95.10
6	7.556	191168	0.73	124821	0.81	(+)-ALPHA-TERPINEOL (P-MENTH-1-EN-8-OL)	59.05
7	10.362	594723	2.26	385034	2.50	BETA-ELEMEN-(2)	93.10
8	11.303	145811	0.55	94525	0.61	ALPHA.-CARYOPHYLLENE	93.10
9	11.500	162357	0.62	92377	0.60	(-)-.ALPHA.-AMORPHENE	161.15
10	11.619	166952	0.64	112752	0.73	(-)-GERMACRENE D	161.10
11	11.731	884889	3.37	559053	3.62	Isogermafurene	108.05
12	11.797	695307	2.65	321702	2.09	PHENOL, 3,5-BIS(1,1-DIMETHYLETHYL)-	191.10
13	12.039	320967	1.22	219507	1.42	(+)-DELTA-CADINENE	161.10
14	13.394	280777	1.07	186329	1.21	(-)-SPATHULENOL	119.10
15	14.201	824947	3.14	520302	3.37	(E,E)-GERMACRONE	107.10
16	14.417	5406678	20.58	3270994	21.20	Curdione	69.05
17	15.245	4036619	15.36	2212631	14.34	Isovelleral	232.10
18	15.639	807509	3.07	464531	3.01	ELEMENE	68.05
19	15.795	3670364	13.97	1992454	12.91	TRIMETHYLCYCLODODECATRIENE	68.05
20	15.885	263570	1.00	181133	1.17	Andrographolide	105.10
21	16.078	732020	2.79	433141	2.81	5,8-Dihydroxy-4a-methyl-4,4a,4b,5,6,7,8,8a,9,10-decahydro-2(3H)-phenanthrenone	174.10
22	16.733	202239	0.77	93405	0.61	Diazoprogesterone	159.10
23	16.801	280218	1.07	172411	1.12	6R,9R-3-Oxo-.alpha.-ionol	108.10
24	17.032	1000819	3.81	563020	3.65	(3E,5E,7E)-6-Methyl-8-(2,6,6-trimethyl-1-cyclohexenyl)-3,5,7-octatrien-2-one	109.05
25	17.427	323925	1.23	182316	1.18	2-Butenal, 2-methyl-4-(2,6,6-trimethyl-1-cyclohexen-1-yl)-	81.10
26	18.602	641108	2.44	404933	2.62	Retinal, 9-cis-	147.10
27	20.912	305994	1.16	132562	0.86	Cholest-6-ene-3.beta.,5.beta.,8.beta.-triol	125.10
		26277375	100.00	15428573	100.00		

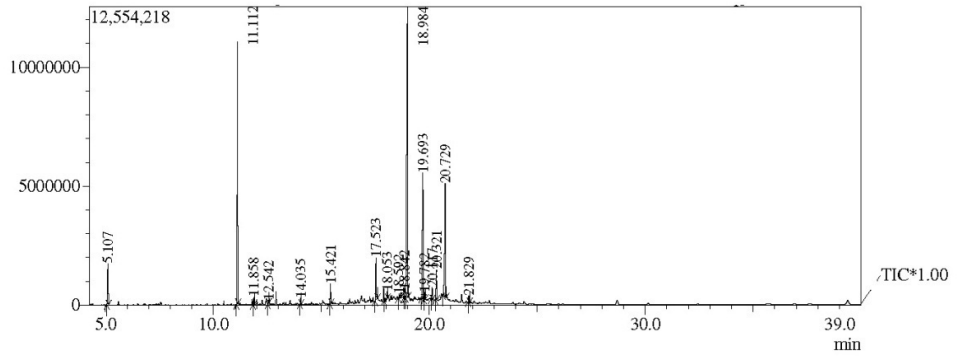
Fig. 5 C. GC-MS Chromatograms of *C. aromatica* rhizome.



Peak Report TIC

Peak#	R.Time	Area	Area%	Height	Height%	Name	Base m/z
1	11.782	271222	0.43	173709	1.10	PHENOL, 3,5-BIS(1,1-DIMETHYLETHYL)-	191.10
2	14.030	187105	0.30	121277	0.77	Spiro[4.5]dec-6-en-8-one, 1,7-dimethyl-4-(1-methylethyl)-	109.10
3	14.647	415365	0.66	252464	1.59	CUPAROPHENOL	136.10
4	15.415	1225458	1.93	746063	4.71	Acetic acid, 1-[2-(2,2,6-trimethyl-bicyclo[4.1.0]hept-1-yl)-ethyl]-vinyl ester	137.10
5	16.982	380165	0.60	235782	1.49	4-Hexen-1-ol, 6-(2,6,6-trimethyl-1-cyclohexenyl)-4-methyl-, (E)-	137.10
6	18.198	1000471	1.58	592071	3.74	DUVATRIENDIOL	81.05
7	18.591	1191858	1.88	729917	4.61	RETINAL	147.05
8	19.492	309249	0.49	154872	0.98	Methyl octadec-6,9-dien-12-ynoate	149.05
9	20.003	432944	0.68	173570	1.10	2,4a,8,8-Tetramethyldecahydrocyclopropa[d]naphthalene	97.00
10	20.771	55173325	87.06	11334176	71.57	(-)-.BETA.-CARYOPHYLLENE EPOXIDE	81.05
11	20.866	1571381	2.48	840863	5.31	Kauren-18-ol, acetate, (4.beta.)-	81.05
12	22.069	657114	1.04	270816	1.71	KAUREN-19-YL-ACETATE	55.00
13	23.517	555664	0.88	211426	1.34	Verrucarol	81.05
		63371321	100.00	15837006	100.00		

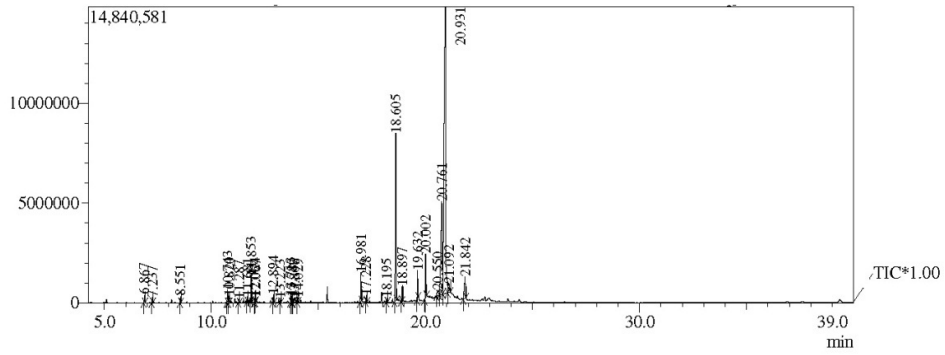
Fig. 5 D. GC-MS Chromatograms of *C. bhatii*.



Peak Report TIC

Peak#	R.Time	Area	Area%	Height	Height%	Name	Base m/z
1	5.107	3334669	3.31	1729915	4.17	Eucalyptol	81.10
2	11.112	20113971	19.97	11022691	26.58	(E)-.beta.-Famesene	69.05
3	11.858	512941	0.51	313190	0.76	.beta.-Bisabolene	69.05
4	12.542	173095	0.17	138053	0.33	LIMONENE DIOXIDE 2	79.05
5	14.035	489046	0.49	261148	0.63	Spiro[4.5]dec-6-en-8-one, 1,7-dimethyl-4-(1-methylethyl)-	109.05
6	15.421	1510622	1.50	836832	2.02	Acetic acid, 1-[2-(2,2,6-trimethyl-bicyclo[4.1.0]hept-1-yl)-ethyl]-vinyl ester	137.10
7	17.523	3387717	3.36	1792411	4.32	Corymbolone	121.10
8	18.053	429636	0.43	302686	0.73	Succinic acid, dodec-9-yn-1-yl nonyl ester	101.05
9	18.592	1758450	1.75	175359	0.42	RETINAL	147.10
10	18.842	1197671	1.19	321136	0.77	LIMONENE DIOXIDE 3	107.10
11	18.984	40275923	39.98	12234162	29.50	1-Heptatriacotanol	69.05
12	19.693	10688636	10.61	5285027	12.74	(1S,2E,4S,5R,7E,11E)-Cembra-2,7,11-trien-4,5-diol	95.05
13	19.782	534572	0.53	289329	0.70	ALLOAROMADENDRENOXID-(1)	137.10
14	20.117	791962	0.79	488089	1.18	Calusterone	135.10
15	20.321	2597854	2.58	1262239	3.04	Tricyclo[20.8.0.0(7,16)]triacontane, 1(22),7(16)-diepoxy-	95.05
16	20.729	12407709	12.32	4766904	11.49	(-)-.BETA.-CARYOPHYLLENE EPOXIDE	81.10
17	21.829	532207	0.53	254381	0.61	Pregnane-3,11,20-triol, (3.alpha., 11.beta., 20.beta.)-	93.10
		100736681	100.00	41473552	100.00		

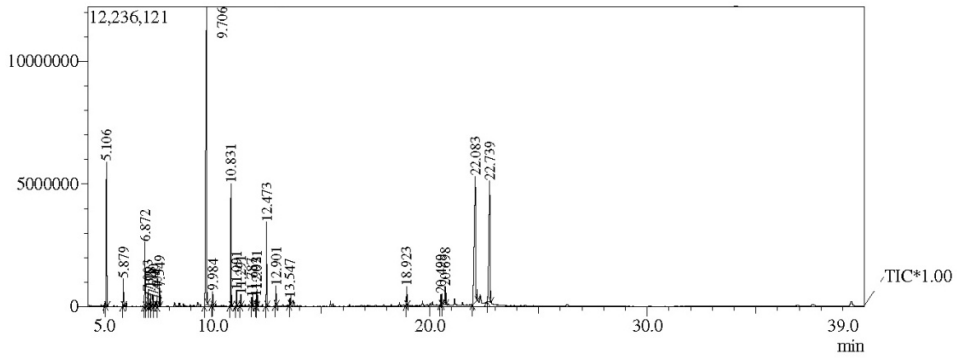
Fig. 5 E. *C. coriacea* GC-MS Chromatograms profile.



Peak Report TIC

Peak#	R.Time	Area	Area%	Height	Height%	Name	Base m/z
1	6.867	619546	0.43	390894	0.94	Camphor	95.10
2	7.237	447514	0.31	255490	0.62	endo-Borneol	95.10
3	8.551	469572	0.33	313550	0.76	.alpha.-Citral	69.10
4	10.743	982582	0.68	671848	1.62	(-)-.alpha.-Santalene	94.10
5	10.820	546599	0.38	351943	0.85	(-)-BETA-CARYOPHYLLEN	93.10
6	11.287	311055	0.22	188713	0.45	.ALPHA.-CARYOPHYLLENE	93.10
7	11.681	353611	0.24	266457	0.64	(-)-ZINGIBERENE	119.15
8	11.853	2246060	1.55	1322455	3.19	.beta.-Bisabolene	69.10
9	12.024	396532	0.27	265687	0.64	(+-)-LEPIDOZENE	121.15
10	12.067	294876	0.20	204715	0.49	.BETA.-SESQUIPHELLANDRENE	69.10
11	12.894	703629	0.49	389704	0.94	cis-sesquisabinene hydrate	83.05
12	13.223	390210	0.27	230546	0.56	alpha.-Humulene epoxide II	67.05
13	13.713	578335	0.40	315841	0.76	Ar-tumerone	83.10
14	13.780	482662	0.33	274841	0.66	Tumerone	83.05
15	13.867	913995	0.63	200925	0.48	JUNIPERCAMPHOR	81.10
16	14.029	434935	0.30	248936	0.60	Spiro[4.5]dec-6-en-8-one, 1,7-dimethyl-4-(1-methylethyl)-	109.10
17	16.981	2612363	1.81	1465037	3.53	4-Hexen-1-ol, 6-(2,6,6-trimethyl-1-cyclohexenyl)-4-methyl-, (E)-	137.15
18	17.228	576056	0.40	320421	0.77	Dehydrosaussurea lactone	81.10
19	18.195	381353	0.26	216933	0.52	4,8,13-DUVATRIENE-1,3-DIOL	81.05
20	18.605	16539144	11.45	8417698	20.30	Retinal, 9-cis-	147.10
21	18.897	1383750	0.96	747055	1.80	1-Ethyl-4,4-dimethyl-cyclohex-2-en-1-ol	125.10
22	19.632	2498936	1.73	1411178	3.40	Acetic acid, 1-[2-(2,2,6-trimethyl-bicyclo[4.1.0]hept-1-yl)-ethyl]-vinyl ester	97.05
23	20.002	4411786	3.05	2210855	5.33	TRISPIRO[4.2.4.2.4.2.]HENEICOSANE	97.05
24	20.550	1621950	1.12	274268	0.66	2-Hexynyl aldehyde diethyl acetal	125.10
25	20.761	20915669	14.48	4657655	11.23	(-)-.BETA.-CARYOPHYLLENE EPOXIDE	137.15
26	20.931	77857219	53.89	14364777	34.63	Kauren-18-ol, acetate, (4.beta.)-	81.10
27	21.092	1713679	1.19	423164	1.02	ALLOAROMADENDRENOXID-(1)	81.10
28	21.842	3794199	2.63	1073923	2.59	KAUREN-19-YL-ACETATE	69.05
		144477817	100.00	41475509	100.00		

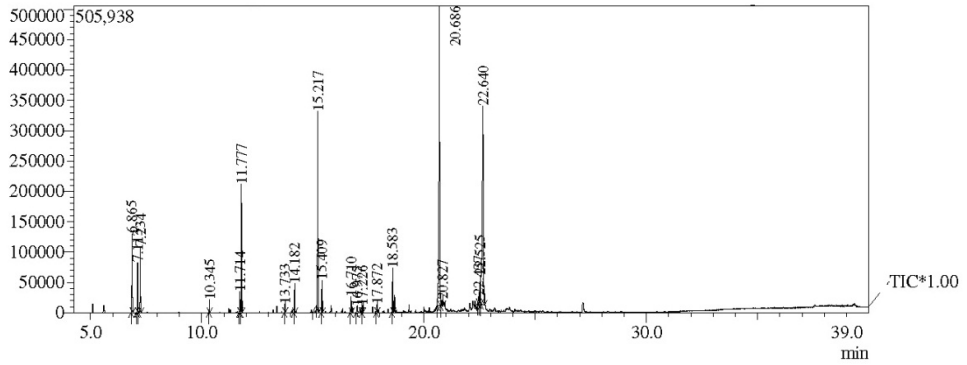
Fig. 5 F. GC-MS Chromatograms of *C. decipiens* rhizome.



Peak Report TIC

Peak#	R.Time	Area	Area%	Height	Height%	Name	Base m/z
1	5.106	10469878	8.36	5880108	12.08	Eucalyptol	81.10
2	5.879	1938010	1.55	1136309	2.33	.ALPHA.-TERPINOLEN	93.10
3	6.872	4496346	3.59	2626635	5.39	Camphor	95.10
4	7.003	1051853	0.84	586407	1.20	EXO-METHYL-CAMPHENILOL	71.05
5	7.118	521155	0.42	315028	0.65	Isoborneol	95.10
6	7.240	796484	0.64	417374	0.86	endo-Borneol	95.10
7	7.404	748419	0.60	207129	0.43	CYMEN-8-OL, P-	135.10
8	7.549	1130681	0.90	757243	1.56	(+)-ALPHA-TERPINEOL (P-MENTH-1-EN-8-OL)	59.05
9	9.706	38512386	30.75	12166563	24.98	3-Terpinolenone	150.10
10	9.984	872965	0.70	572243	1.18	5,7-DIMETHYLOCTAHYDROCOUMARIN 1	95.05
11	10.831	8181166	6.53	5003939	10.28	TRANS(.BETA.)-CARYOPHYLLENE	93.10
12	11.091	926471	0.74	630889	1.30	(E)-.beta.-Famesene	69.10
13	11.291	729509	0.58	462691	0.95	.ALPHA.-CARYOPHYLLENE	93.10
14	11.787	531505	0.42	334173	0.69	PHENOL, 3,5-BIS(1,1-DIMETHYLETHYL)-	191.10
15	11.992	565652	0.45	357447	0.73	.GAMMA.-MUUROLEN	161.15
16	12.031	1000312	0.80	676097	1.39	(+)-DELTA-CADINENE	161.15
17	12.473	5244467	4.19	3425541	7.03	E-Nerolidol	69.05
18	12.901	1312423	1.05	801676	1.65	Isoaromadendrene epoxide	79.05
19	13.547	507269	0.41	288734	0.59	.tau.-Cadinol	161.15
20	18.923	1300560	1.04	784788	1.61	1-Heptatriacotanol	83.10
21	20.499	830928	0.66	439021	0.90	3-OXO-5-PHENYL-PENTANOIC ACID	91.05
22	20.698	1837492	1.47	733043	1.51	(-)-.BETA.-CARYOPHYLLENE EPOXIDE	81.10
23	22.083	22943053	18.32	5156384	10.59	Dihydrochrysin	256.10
24	22.739	18782886	15.00	4936099	10.14	3,3-Dimethyl-4-phenyl-4-penten-2-one	131.10
		125231870	100.00	48695561	100.00		

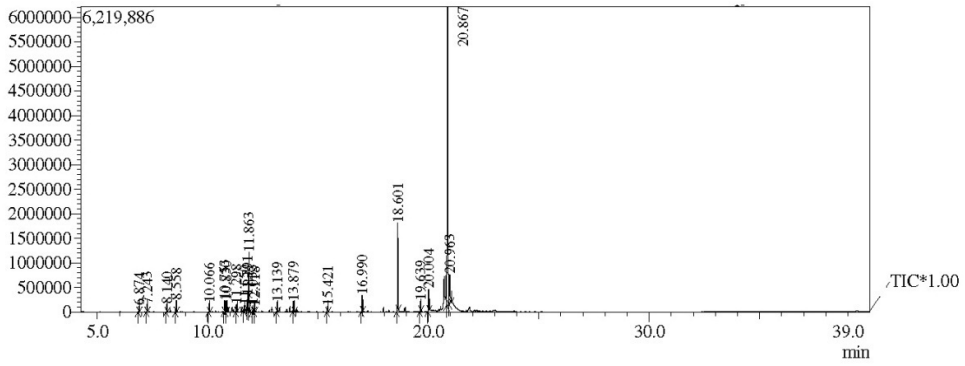
Fig. 5 G. *C. aurantiaca* GC-MS Chromatograms.



Peak Report TIC

Peak#	R.Time	Area	Area%	Height	Height%	Name	Base m/z
1	6.865	199872	4.55	129934	6.31	(-)-CAMPHOR	95.10
2	7.113	122643	2.79	82720	4.02	2-CAMPHANOL, EXO-	95.10
3	7.234	157334	3.58	108476	5.27	2-BORNANOL	95.10
4	10.345	27021	0.62	21415	1.04	BETA. ELEMENE	81.05
5	11.714	49179	1.12	34919	1.70	Isosericenin	108.10
6	11.777	321366	7.31	212023	10.30	2,4-DITERT-BUTYLPHENOL	191.15
7	13.733	15087	0.34	13901	0.68	ISONARDOSINON	59.05
8	14.182	67413	1.53	48450	2.35	Germacron	107.10
9	15.217	541399	12.32	326323	15.86	Isovelleral	232.15
10	15.409	81134	1.85	53401	2.60	LABDA-8(17),13Z-DIEN-15-OL	137.10
11	16.710	44915	1.02	26612	1.29	6-HEPTENSAEURE, 4-ISOPROPENYL-6-METHYL-2-OXO-, ETHYLESTER	55.10
12	16.975	38230	0.87	10167	0.49	(E)-3-METHYL-5-(2',6',6'-TRIMETHYLCYCLOHEX-1'-ENYL)PENT-2-ENYL BROMIDE	137.20
13	17.226	15831	0.36	12009	0.58	(Z)-4-METHYL-6-(2',6',6'-TRIMETHYL-2'-CYCLOHEXENYL)HEX-3-EN-1-OL	81.10
14	17.872	21879	0.50	13550	0.66	2-(2-Hydroxyethylamino)pyrimidine	108.05
15	18.583	107065	2.44	73865	3.59	Diazoprogesterone	147.10
16	20.686	1170777	26.65	494222	24.02	(-)-.BETA.-CARYOPHYLLENE EPOXIDE	81.10
17	20.827	29389	0.67	9582	0.47	1-Oxaspiro[2.5]octane, 4,4-dimethyl-8-methylene-	137.20
18	22.437	23559	0.54	11985	0.58	trans-2-Caren-4-ol	109.10
19	22.525	136857	3.12	46964	2.28	DRIMENOL	109.10
20	22.640	1222298	27.82	327241	15.90	Isoaromadendrene epoxide	55.05
		4393248	100.00	2057759	100.00		

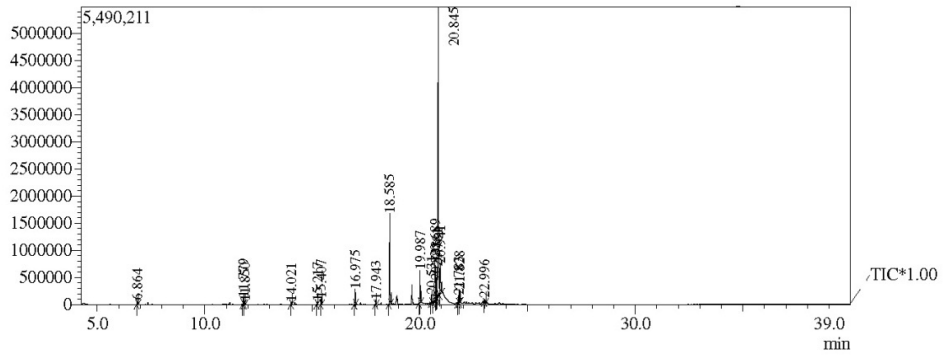
Fig. 5 H. GC-MS Chromatograms of *C. haritha* rhizome.



Peak Report TIC

Peak#	R.Time	Area	Area%	Height	Height%	Name	Base m/z
1	6.874	187264	0.71	120312	0.91	Camphor	95.10
2	7.243	221875	0.85	153652	1.16	endo-Borneol	95.10
3	8.140	239801	0.91	162012	1.22	.BETA.-CITRAL	69.10
4	8.558	345969	1.32	231544	1.74	.alpha.-Citral	69.05
5	10.066	249569	0.95	178632	1.34	NERYL ACETATE	69.10
6	10.753	367542	1.40	239628	1.80	(-)-ALPHA-SANTALENE	94.05
7	10.830	364339	1.39	230351	1.73	(-)-.BETA.-CARYOPHYLLENE	93.10
8	11.298	245429	0.94	156926	1.18	.ALPHA.-CARYOPHYLLENE	93.10
9	11.650	157962	0.60	123132	0.93	.ALPHA.-SELINENE	189.15
10	11.791	344540	1.31	279039	2.10	PHENOL, 2,4-BIS(1,1-DIMETHYLETHYL)-	191.10
11	11.863	2026833	7.72	1179893	8.88	(+,-)-.BETA.-BISABOLENE	69.10
12	12.033	209388	0.80	127442	0.96	.DELTA.-CADINENE	161.10
13	12.118	156140	0.59	111903	0.84	(-)-.alpha.-Panasinsen	122.10
14	13.139	302385	1.15	203067	1.53	Carotol	161.15
15	13.879	451997	1.72	211268	1.59	Juniper camphor	81.05
16	15.421	189497	0.72	117899	0.89	Acetic acid, 1-[2-(2,2,6-trimethyl-bicyclo[4.1.0]hept-1-yl)-ethyl]-vinyl ester	137.10
17	16.990	574545	2.19	339640	2.56	4-Hexen-1-ol, 6-(2,6,6-trimethyl-1-cyclohexenyl)-4-methyl-, (E)-	137.10
18	18.601	2980764	11.36	1794425	13.51	Retinal, 9-cis-	147.10
19	19.639	361977	1.38	210640	1.59	Longipinocarveol, trans-	97.00
20	20.004	770035	2.93	440092	3.31	2,4a,8,8-Tetramethyldecahydrocyclopropa[d]naphthalene	97.00
21	20.867	14003209	53.35	6074768	45.74	ALLOAROMADENDRENOXID-(1)	81.10
22	20.963	1494862	5.70	595078	4.48	Andrographolide	147.10
		26245922	100.00	13281343	100.00		

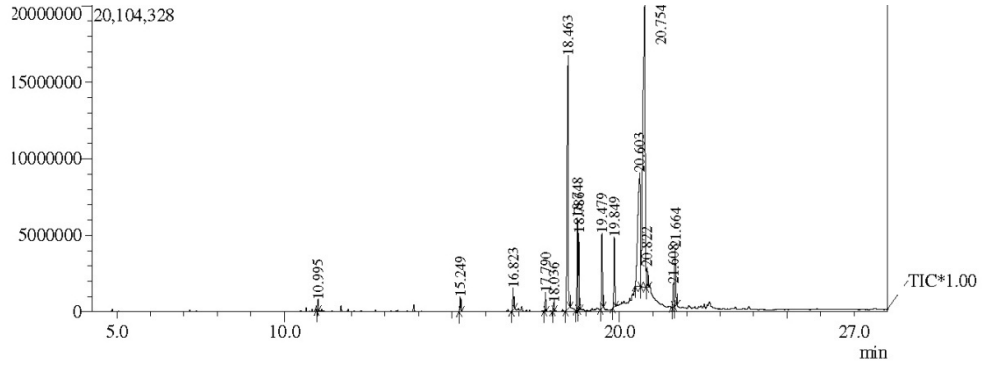
Fig. 5 I. GC-MS Chromatograms of *C. inodora* rhizome methanolic extract.



Peak Report TIC

Peak#	R.Time	Area	Area%	Height	Height%	Name	Base m/z
1	6.864	121384	0.51	77724	0.69	Camphor	95.10
2	11.779	200304	0.84	129176	1.14	2,4-DITERT-BUTYLPHENOL	191.20
3	11.850	83617	0.35	58277	0.52	(-)-.BETA.-BISABOLENE	69.10
4	14.021	81575	0.34	47546	0.42	Isocyclocitral	109.10
5	15.217	62206	0.26	51880	0.46	4,8,11,11-TETRAMETHYL-8-TRICYCLO[7.2.0.0(2,5).]UNDECEN-4-OL	232.15
6	15.407	180681	0.76	115763	1.03	3-Cyclohexene-1-carboxaldehyde, 1,3,4-trimethyl-	137.15
7	16.975	471271	1.97	284303	2.52	4-Hexen-1-ol, 6-(2,6,6-trimethyl-1-cyclohexenyl)-4-methyl-, (E)-	137.20
8	17.943	129017	0.54	86859	0.77	Diazoprogesterone	146.10
9	18.585	2818918	11.79	1674674	14.84	RETINAL	147.10
10	19.987	1076803	4.50	606823	5.38	Kauren-18-ol, acetate, (4.beta.)-	81.10
11	20.531	249057	1.04	145408	1.29	Bicyclo[3.1.1]heptan-3-one, 6,6-dimethyl-2-(2-methylpropyl)-	125.10
12	20.689	2007735	8.39	774748	6.87	((-)-.BETA.-CARYOPHYLLENE EPOXIDE	81.10
13	20.738	720260	3.01	334448	2.96	1-Ethyl-4,4-dimethyl-cyclohex-2-en-1-ol	125.15
14	20.781	1092847	4.57	567923	5.03	10-Chlorotricyclo[4.2.1.1(2,5)]deca-3,7-dien-9-ol	147.10
15	20.845	12157291	50.83	5341322	47.33	ALLOAROMADENDRENOXID-(1)	81.10
16	20.941	1469143	6.14	550764	4.88	3-BROMOCHOLEST-5-ENE #	147.10
17	21.783	368783	1.54	140226	1.24	Longipinocarveol, trans-	69.05
18	21.828	480940	2.01	232906	2.06	Ergost-25-ene-6,12-dione, 3,5-dihydroxy-, (3.beta.,5.alpha.)-	69.10
19	22.996	147524	0.62	64159	0.57	DUVATRIENDIOL	81.10
		23919356	100.00	11284929	100.00		

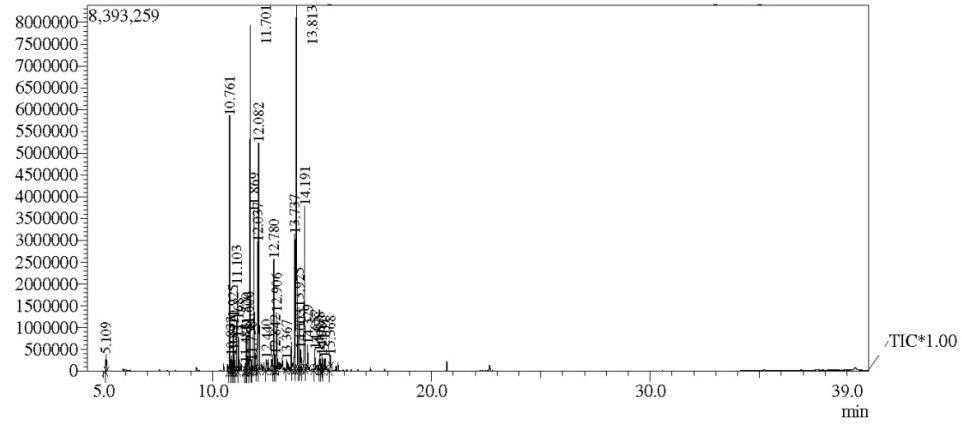
Fig. 5 J. GC-MS profile of *C. karnatakensis* rhizome.



Peak Report TIC

Peak#	R.Time	Area	Area%	Height	Height%	Name	Base m/z
1	10.995	1037748	0.45	694474	0.94	.BETA.-SESQUIPHELLANDRENE	69.05
2	15.249	1623339	0.71	981621	1.32	Acetic acid, 1-[2-(2,2,6-trimethyl-bicyclo[4.1.0]hept-1-yl)-ethyl]-vinyl ester	137.10
3	16.823	2653153	1.16	1430187	1.93	4-Hexen-1-ol, 6-(2,6,6-trimethyl-1-cyclohexenyl)-4-methyl-, (E)-	137.10
4	17.790	1935255	0.85	1195435	1.61	2(1H)-Naphthalenone, 7-ethynyl-4a,5,6,7,8,8a-hexahydro-1,4a-dimethyl-, (1.alpha.,4a.beta.,7.beta.,8a.alpha.)-	146.05
5	18.036	866597	0.38	509854	0.69	SPINACENE	81.10
6	18.463	41349349	18.08	16545444	22.29	Retinal, 9-cis-	147.10
7	18.748	10906865	4.77	6014391	8.10	1-Ethyl-4,4-dimethyl-cyclohex-2-en-1-ol	125.10
8	18.786	7827533	3.42	5005552	6.74	2-(4-NITROBUTANOYL)CYCLOHEXANONE	125.10
9	19.479	9017118	3.94	4868294	6.56	Acetic acid, 1-[2-(2,2,6-trimethyl-bicyclo[4.1.0]hept-1-yl)-ethyl]-vinyl ester	97.05
10	19.849	8573085	3.75	4559743	6.14	Trispiro[4.2.4.2.4.2.]heneicosane	97.05
11	20.603	39589876	17.31	7411793	9.98	(-)-.BETA.-CARYOPHYLLENE EPOXIDE	81.10
12	20.754	87715902	38.35	18473138	24.88	Kauren-18-ol, acetate, (4.beta.)-	81.10
13	20.822	3021396	1.32	1289976	1.74	22,23-Dibromostigmasterol acetate	147.10
14	21.608	4176592	1.83	1493779	2.01	2,6,10,14,18,22-TETRACOSAHEXAENE, 2,6,10,15,19,23-HEXAMETHYL-	69.05
15	21.664	8416158	3.68	3766726	5.07	Ergost-25-ene-6,12-dione, 3,5-dihydroxy-, (3.beta.,5.alpha.)-	69.05
		228709966	100.00	74240407	100.00		

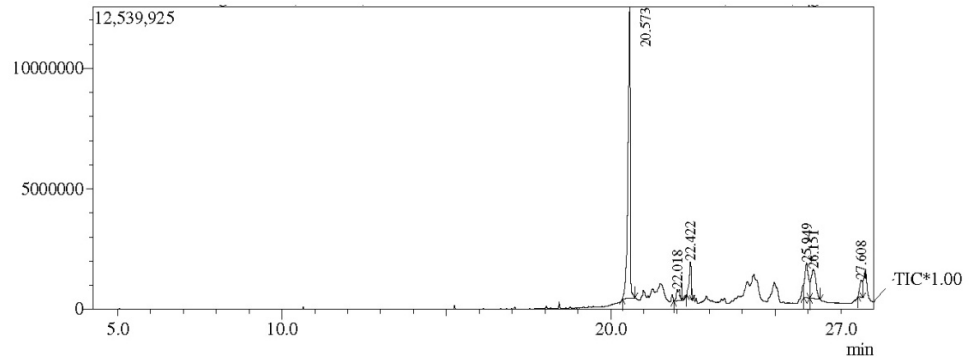
Fig. 5 K. *C. kudagensis* rhizome GC-MS Chromatograms.



Peak Report TIC

Peak#	R.Time	Area	Area%	Height	Height%	Name	Base m/z
1	5.109	621760	0.57	390552	0.69	Eucalyptol	81.05
2	10.761	9421022	8.67	5851231	10.29	(-)-.alpha.-Santalene	94.10
3	10.833	559563	0.51	350975	0.62	Caryophyllene	93.10
4	10.925	1604064	1.48	1058799	1.86	trans-.alpha.-Bergamotene	119.10
5	10.971	769487	0.71	477749	0.84	Teresantalol	121.15
6	11.103	4368702	4.02	1973416	3.47	(E)-beta.-Famesene	69.05
7	11.168	1084419	1.00	760629	1.34	TRANS-BETA.-FARNESENE	69.10
8	11.483	357823	0.33	188529	0.33	.GAMMA.-CURCUMENE	121.15
9	11.520	1344829	1.24	889450	1.56	ALPHA-CURCUMEN	132.15
10	11.606	1421935	1.31	916286	1.61	cis-.beta.-Farnesene	69.10
11	11.701	12666973	11.66	7914967	13.92	(-)-ZINGIBERENE	119.10
12	11.758	391248	0.36	222471	0.39	ALPHA.-BISABOLENE	93.10
13	11.869	6056464	5.57	3630198	6.38	(+,-)-.BETA.-BISABOLENE	69.10
14	12.037	5164575	4.75	2928844	5.15	Benzenemethanol, 4-methyl-.alpha.-(1-methyl-2-propenyl)-, (R*,R*)-	121.15
15	12.082	9624192	8.86	5197383	9.14	.BETA.-SESQUIPHELLANDRENE	69.05
16	12.440	890214	0.82	251307	0.44	(-)-.BETA.-CARYOPHYLLENE EPOXIDE	93.10
17	12.780	4633125	4.26	2537889	4.46	6-Isopropyl-1,2-dimethyl-4-oxo-bicyclo[3.3.1]non-2-ene-9-carboxaldehyde	83.10
18	12.842	560450	0.52	351020	0.62	.alpha.-Santalol	85.10
19	12.906	2850311	2.62	1325435	2.33	Butane-1,1-dicarbonitrile, 1-cyclohexyl-3-methyl-	83.05
20	13.367	736030	0.68	223277	0.39	7-epi-cis-sesquisabinene hydrate	69.05
21	13.737	6728366	6.19	3102169	5.46	Ar-tumerone	83.10
22	13.813	20813727	19.15	8345766	14.68	Tumerone	83.05
23	13.925	3374144	3.11	1416464	2.49	Bicyclo[3.3.1]nonan-9-one, 1,2,4-trimethyl-3-nitro-, (2-endo,3-exo,4-exo)-(.+-.)-	95.10
24	14.003	1291864	1.19	465447	0.82	7-epi-trans-sesquisabinene hydrate	93.10
25	14.191	6974244	6.42	3739480	6.58	Curlone	120.10
26	14.329	1257183	1.16	573416	1.01	ALPHA. TERPINEOL	59.05
27	14.652	762273	0.70	461148	0.81	Naphthalene, 1,1'-methylenebis[decahydro-	137.10
28	14.878	768198	0.71	414534	0.73	Cyclohexane, (2-nitro-2-propenyl)-	83.05
29	14.980	502220	0.46	333602	0.59	1H-Indene, 2,3,3a,4,7,7a-hexahydro-2,2,4,4,7,7-hexamethyl-	83.10
30	15.107	460987	0.42	266547	0.47	Cyclohexene, 1-methyl-3-(1-methylethyl)-	95.10
31	15.368	606294	0.56	304830	0.54	Dehydrozingerone	145.00
		108666686	100.00	56863810	100.00		

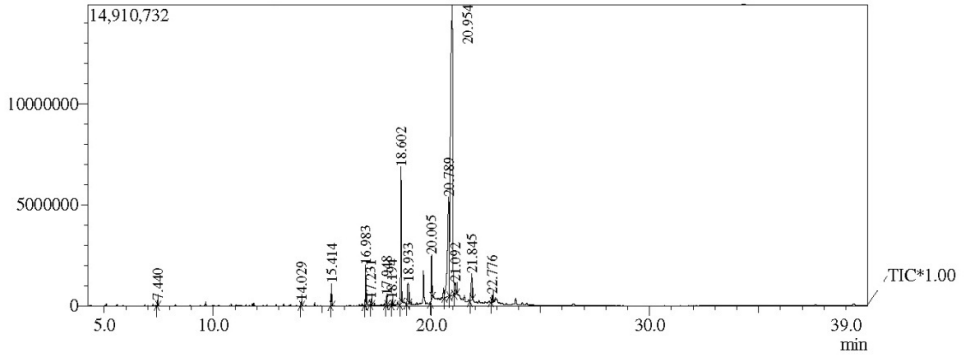
Fig. 5 L. GC-MS Chromatograms of methanolic extract form *C. longa* rhizome.



Peak Report TIC

Peak#	R.Time	Area	Area%	Height	Height%	Name	Base m/z
1	20.573	58415947	60.78	12078649	69.29	Alloaromadendrene oxide-(1)	81.10
2	22.018	3410457	3.55	475836	2.73	ANDROSTAN-3-ONE, 17-HYDROXY-, (5.ALPHA.,17.BETA.)-	181.10
3	22.422	6367682	6.63	1521445	8.73	(-)-.BETA.-CARYOPHYLLENE EPOXIDE	55.10
4	25.949	11035593	11.48	1445468	8.29	VIOLANTHRENE A	426.30
5	26.151	12351065	12.85	1222715	7.01	2-tert-Butyl-4,6-bis(3,5-di-tert-butyl-4-hydroxybenzyl)phenol	57.10
6	27.608	4523502	4.71	688902	3.95	o[3,5,6-metheno-1-silapentalene-1(2H),1'-[1,3]disilacyclobutane-3',1''(2''H)-[3,5,6]metheno[1]silapentalene], 2,2	57.10
		96104246	100.00	17433015	100.00		

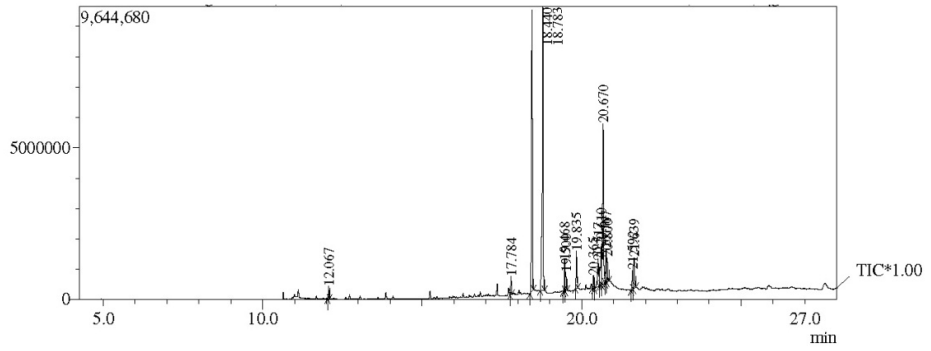
Fig. 5 M. GC-MS Chromatograms profile of *C. montana* rhizome.



Peak Report TIC

Peak#	R.Time	Area	Area%	Height	Height%	Name	Base m/z
1	7.440	441799	0.28	267600	0.74	Cryptone, L-	96.05
2	14.029	383029	0.24	219812	0.61	Spiro[4.5]dec-6-en-8-one, 1,7-dimethyl-4-(1-methylethyl)-	109.10
3	15.414	1912235	1.20	1107116	3.05	Acetic acid, 1-[2-(2,2,6-trimethyl-bicyclo[4.1.0]hept-1-yl)-ethyl]-vinyl ester	137.15
4	16.983	3431272	2.15	1977843	5.45	4-Hexen-1-ol, 6-(2,6,6-trimethyl-1-cyclohexenyl)-4-methyl-, (E)-	137.15
5	17.231	600959	0.38	304555	0.84	4,8,13-Duvatriene-1,3-Diol	81.10
6	17.948	750613	0.47	458026	1.26	7-Ethynyl-1,4a-dimethyl-4a,5,6,7,8,8a-hexahydro-2(1H)-naphthalenone	146.10
7	18.194	485945	0.30	279965	0.77	Alloaromadendrene oxide-(1)	81.10
8	18.602	13015867	8.14	6789279	18.71	RETINAL	147.10
9	18.933	3969350	2.48	1057661	2.92	1-Ethyl-4,4-dimethyl-cyclohex-2-en-1-ol	125.10
10	20.005	4840360	3.03	2259406	6.23	TRISPIRO[4.2.4.2.4.2]HENEICOSANE	97.05
11	20.789	27525402	17.21	4924944	13.57	(-)-.BETA.-CARYOPHYLLENE EPOXIDE	137.15
12	20.954	94639145	59.18	14385986	39.65	Kauren-18-ol, acetate, (4.beta.)-	81.10
13	21.092	2199794	1.38	576686	1.59	Aromadendrene oxide-(2)	81.05
14	21.845	4834286	3.02	1322574	3.65	KAUREN-19-YL-ACETATE	69.05
15	22.776	900154	0.56	350176	0.97	Diazoprogesterone	137.15
		159930210	100.00	36281629	100.00		

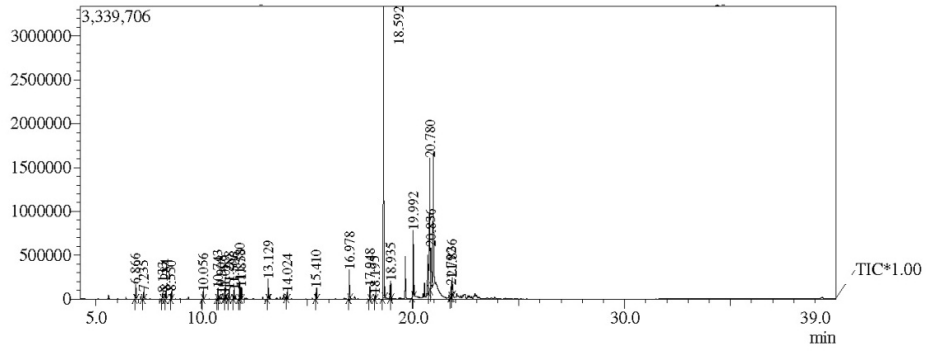
Fig. 5 N. GC-MS Chromatograms of methanolic rhizome extract of *C. mutabilis*.



Peak Report TIC

Peak#	R.Time	Area	Area%	Height	Height%	Name	Base m/z
1	12.067	636320	0.89	417464	1.23	1,3,3-TRIMETHYL-2-OXABICYCLO[2.2.2]OCT-6-YL ACETATE	108.10
2	17.784	915116	1.29	600768	1.77	.BETA.-COPAEN-4 .ALPHA.-OL	146.05
3	18.440	16631989	23.36	9273132	27.36	Retinal, 9-cis-	147.10
4	18.783	21064869	29.59	9345645	27.57	1-Heptatriacotanol	125.10
5	19.468	1774568	2.49	1016150	3.00	1,5-EPOXYSALVIAL-4(14)-ENE	97.10
6	19.500	1028074	1.44	613506	1.81	2,6,10,10-Tetramethyl-1-oxaspiro[4.5]decan-6-ol	85.15
7	19.835	2101183	2.95	1221329	3.60	1,5-EPOXYSALVIAL-4(14)-ENE	97.10
8	20.365	688739	0.97	416708	1.23	1-Ethyl-4,4-dimethyl-cyclohex-2-en-1-ol	125.10
9	20.517	2637073	3.70	860384	2.54	(-)-.BETA.-CARYOPHYLLENE EPOXIDE	81.10
10	20.610	3617812	5.08	1276308	3.77	(4,4-Dimethyl-2,4,5,6-tetrahydro-1H-inden-2-yl)acetic acid	147.10
11	20.670	12595639	17.69	5290146	15.61	ALLOAROMADENDRENOXID-(1)	81.10
12	20.767	2219791	3.12	1105412	3.26	Andrographolide	147.10
13	20.800	1440393	2.02	802009	2.37	Cycloheptanol, 1-allyl-2-methylene-	125.10
14	21.592	1602470	2.25	633436	1.87	Ergost-25-ene-6,12-dione, 3,5-dihydroxy-, (3.beta.,5.alpha.)-	69.05
15	21.639	2244009	3.15	1022513	3.02	Acetic acid, 1-[2-(2,2,6-trimethyl-bicyclo[4.1.0]hept-1-yl)-ethyl]-vinyl ester	69.10
		71198045	100.00	33894910	100.00		

Fig. 5 O. *C. neilgherrensis* GC-MS Chromatograms profile.

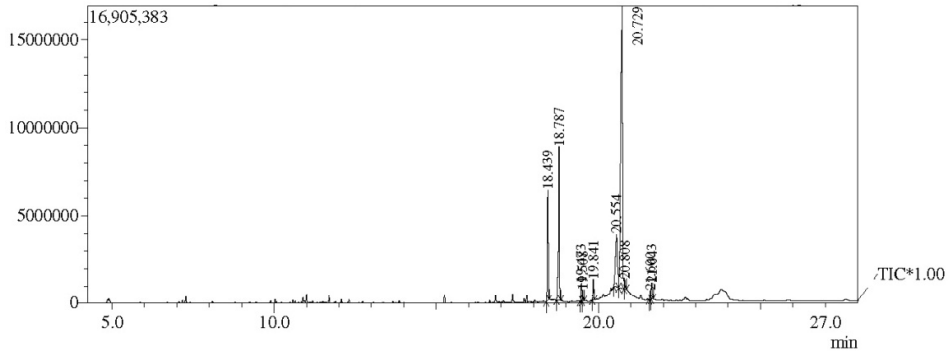


Peak Report TIC

Peak#	R. Time	Area	Area%	Height	Height%	Name	Base m/z
1	6.866	242458	1.47	162298	1.83	Camphor	95.05
2	7.235	111298	0.68	81427	0.92	2-BORNANOL	95.10
3	8.133	88456	0.54	64618	0.73	.BETA.-CITRAL	69.05
4	8.284	135140	0.82	89571	1.01	.BETA.-GERANIOL	69.05
5	8.550	120418	0.73	84090	0.95	.ALPHA.-CITRAL	69.10
6	10.056	121083	0.74	84430	0.95	2,6-OCTADIEN-1-OL, 3,7-DIMETHYL-, ACETATE	69.05
7	10.743	170064	1.03	122074	1.37	.ALPHA.-SANTALENE	94.05
8	10.908	196927	1.20	54518	0.61	(-)-.ALPHA.-TRANS-BERGAMOTENE	93.10
9	11.089	74803	0.45	53862	0.61	(Z)-.BETA.-FARNESENE	69.05
10	11.288	171567	1.04	108366	1.22	ALPHA.-CARYOPHYLLENE	93.10
11	11.506	137611	0.84	96672	1.09	ALPHA-CURCUMEN	132.15
12	11.780	301716	1.84	197379	2.22	2,4-DITERT-BUTYLPHENOL	191.10
13	11.853	274748	1.67	130800	1.47	(-)-.BETA.-BISABOLENE	69.05
14	13.129	375975	2.29	227750	2.56	(+)-CAROTOL	161.10
15	14.024	107915	0.66	71953	0.81	(Z)-iso-Geraniol	109.10
16	15.410	188753	1.15	126407	1.42	1-CYCLOHEXENE-1-PROPANOL, .ALPHA.,.GAMMA.,2,6,6-PENTAMETHYL-	137.10
17	16.978	535014	3.25	328838	3.70	4-Hexen-1-ol, 6-(2,6,6-trimethyl-1-cyclohexenyl)-4-methyl-, (E)-	137.15
18	17.948	209625	1.28	140764	1.58	Androst-5-ene-3,19-diol, 3-acetate, (3.beta.)-	146.05
19	18.195	75770	0.46	47865	0.54	LABDA-8(17),13Z-DIEN-15-OL	81.05
20	18.592	5615073	34.15	3332504	37.48	RETINAL	147.05
21	18.935	301489	1.83	201780	2.27	1-Ethyl-4,4-dimethyl-cyclohex-2-en-1-ol	125.10
22	19.992	1312704	7.98	757119	8.52	Kauren-18-ol, acetate, (4.beta.)-	97.05
23	20.780	3500172	21.29	1514501	17.03	3-BROMOCHOLEST-5-ENE #	147.05
24	20.836	1319682	8.03	474579	5.34	(-)-.BETA.-CARYOPHYLLENE EPOXIDE	81.05
25	21.792	251976	1.53	120053	1.35	Ergost-25-ene-6,12-dione, 3,5-dihydroxy-, (3.beta.,5.alpha.)-	69.05
26	21.836	499792	3.04	216550	2.44	4,8,13-DUVATRIENE-1,3-DIOL	69.05
		16440229	100.00	8890768	100.00		

Fig. 5 P. GC-MS Chromatograms of compounds from *C. oligantha* var. *oligantha* rhizome.

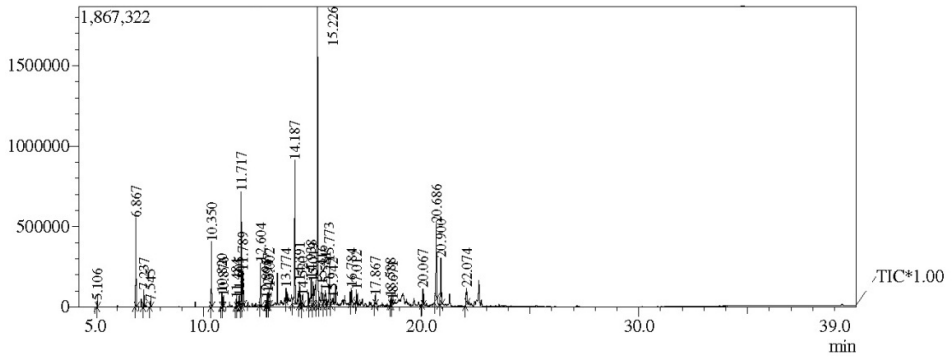
.



Peak Report TIC

Peak#	R.Time	Area	Area%	Height	Height%	Name	Base m/z
1	18.439	11217290	9.56	6317754	16.20	Retinal, 9-cis-	147.10
2	18.787	20440104	17.42	8755450	22.45	1-Heptatriacotanol	83.05
3	19.473	2057630	1.75	1067777	2.74	(ALBICANOL) DECAHYDRO-2-METHYLENE-5,5,8A-TRIMETHYL-1-NAPHTHALENEMETHANOL	97.05
4	19.508	1064834	0.91	639747	1.64	NEROLIDOL-EPOXYACETATE	85.10
5	19.841	2130707	1.82	1200762	3.08	2,4a,8,8-Tetramethyldecahydrocyclopropa[d]naphthalene	97.05
6	20.554	12853972	10.96	2996513	7.68	(-)-.BETA.-CARYOPHYLLENE EPOXIDE	81.10
7	20.729	63407841	54.05	16032212	41.11	Kauren-18-ol, acetate, (4.beta.)-	81.05
8	20.808	1000730	0.85	532935	1.37	1-Oxaspiro[4.5]decan-2-one, 6-isopropenyl-9-methyl-	125.10
9	21.600	1393282	1.19	516751	1.33	SPINACENE	69.10
10	21.643	1739962	1.48	938457	2.41	Ergost-25-ene-6,12-dione, 3,5-dihydroxy-, (3.beta.,5.alpha.)-	69.05
		117306352	100.00	38998358	100.00		

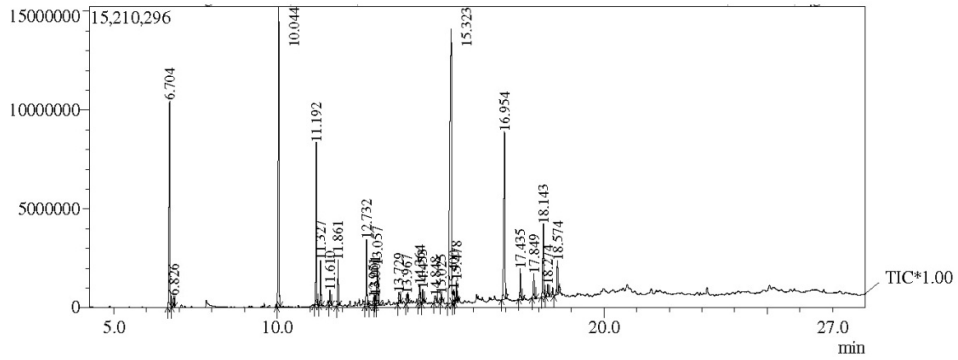
Fig. 5 Q. *C. pseudomontana* rhizome components GC-MS Chromatograms.



Peak Report TIC

Peak#	R.Time	Area	Area%	Height	Height%	Name	Base m/z
1	5.106	55058	0.34	39839	0.50	1, 8-CINEOL	81.10
2	6.867	865397	5.30	553364	6.94	Camphor	95.10
3	7.237	162601	1.00	104228	1.31	2-BORNANOL	95.10
4	7.545	51805	0.32	35491	0.44	(-)-.ALPHA.-TERPINEOL	59.05
5	10.350	597286	3.66	407200	5.10	BETA-ELEMEN-(2)	81.05
6	10.820	131547	0.81	87414	1.10	(-)-BETA-CARYOPHYLLENE	93.10
7	10.876	82981	0.51	63432	0.80	.gamma.-Elemene	121.10
8	11.484	82441	0.51	50436	0.63	5.BETA.,10.ALPHA.-EUDESMA-4(14),11-DIENE	105.10
9	11.605	126663	0.78	66179	0.83	(-)-GERMACRENE D	161.10
10	11.717	1109092	6.80	713183	8.94	Isogermafurene	108.05
11	11.789	562695	3.45	216661	2.72	1-(4-ISOPROPYLPHENYL)-2-METHYLPROPYL ACETATE	191.10
12	12.604	436298	2.67	267599	3.35	GERMACRENE B	121.10
13	12.895	75481	0.46	45455	0.57	(-)-.BETA.-CARYOPHYLLENE EPOXIDE	69.05
14	12.947	134185	0.82	79336	0.99	(+)-BETA-GUAJEN	161.10
15	13.002	237369	1.45	111627	1.40	.beta.-Elemenone	107.10
16	13.774	124415	0.76	82477	1.03	Juniper camphor	81.10
17	14.187	1757245	10.77	893465	11.20	(E,E)-GERMACRONE	107.10
18	14.391	266419	1.63	148331	1.86	Curdione	69.05
19	14.520	149552	0.92	63406	0.79	Isoaromadendrene epoxide	93.10
20	14.938	585800	3.59	165018	2.07	5,8-Dihydroxy-4a-methyl-4,4a,4b,5,6,7,8,8a,9,10-decahydro-2(3H)-phenanthrenone	159.05
21	15.019	440425	2.70	146117	1.83	(-)-Spathulenol	119.10
22	15.226	3873570	23.74	1856848	23.28	Isovelleral	232.10
23	15.416	549877	3.37	124267	1.56	(-)-BETA-CARYOPHYLLENE	109.05
24	15.581	347564	2.13	87652	1.10	6.beta.-Hydroxymethandienone	232.10
25	15.773	547256	3.35	265666	3.33	ELEMENE	68.05
26	15.942	163272	1.00	37543	0.47	1(10)-CIS-COSTUNOLIDE	68.00
27	16.784	154810	0.95	88756	1.11	6R,9R-3-Oxo-.alpha.-ionol	108.05
28	17.012	158195	0.97	85312	1.07	(3E,5E,7E)-6-Methyl-8-(2,6,6-trimethyl-1-cyclohexenyl)-3,5,7-octatrien-2-one	109.00
29	17.867	110991	0.68	53324	0.67	Longipinocarveol, trans-	137.15
30	18.588	76310	0.47	47307	0.59	Diazoprogesterone	147.05
31	18.671	51534	0.32	30467	0.38	Isocyclocitral	109.10
32	20.067	192823	1.18	100151	1.26	GAILLARDIN	188.00
33	20.686	1069185	6.55	471207	5.91	ALLOAROMADENDRENOXID-(1)	81.05
34	20.900	704473	4.32	285823	3.58	1-Heptatriacotanol	81.05
35	22.074	280599	1.72	102828	1.29	Card-20(22)-enolide, 2,3,14-trihydroxy-, (2.alpha.,3.beta.,5.alpha.)-	124.10
		16315214	100.00	7977409	100.00		

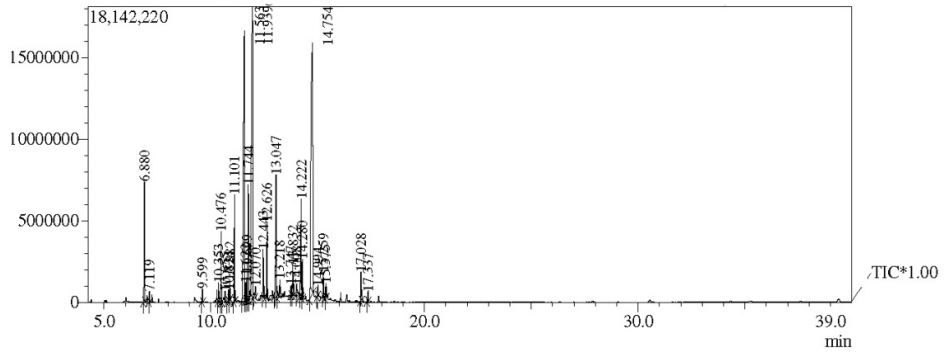
Fig. 5 R. GC-MS Chromatograms of methanolic extraction form *C. zedoaria* rhizome.



Peak Report TIC

Peak#	R.Time	Area	Area%	Height	Height%	Name	Base m/z
1	6.704	21503287	11.24	10429834	12.67	Camphor	95.15
2	6.826	1022833	0.53	573617	0.70	Camphene hydrate	71.10
3	10.044	29771291	15.56	15155814	18.41	Copaene	119.15
4	11.192	14092480	7.37	8202714	9.96	Aromandendrene	91.10
5	11.327	3520497	1.84	2278650	2.77	GAMMA-MUUROLEN	161.20
6	11.610	1385358	0.72	738218	0.90	cubedol	161.20
7	11.861	3845611	2.01	2305732	2.80	4-epi-cubedol	161.20
8	12.732	5720464	2.99	3356849	4.08	(-)-.BETA.-CARYOPHYLLENE EPOXIDE	79.10
9	12.951	1119561	0.59	403246	0.49	Diepicedrene-1-oxide	81.10
10	13.001	931895	0.49	504858	0.61	Globulol	122.15
11	13.057	3954276	2.07	1996807	2.43	Humulene epoxide 2	109.15
12	13.729	1427090	0.75	561439	0.68	ALLOAROMADENDRENOXID-(1)	91.10
13	13.967	873187	0.46	453830	0.55	(-)-SPATHULENOL	159.15
14	14.364	2064708	1.08	949762	1.15	1,4A,7,7-TETRAMETHYLDECAHYDROCYCLOPROPA[7,8]AZULENO[3A,4-B]OXIRENE	132.15
15	14.453	1291006	0.67	692687	0.84	7-Acetyl-2-hydroxy-2-methyl-5-isopropylbicyclo[4.3.0]nonane	153.15
16	14.848	1260168	0.66	338248	0.41	.BETA.-COPAEN-4.ALPHA.-OL	203.15
17	15.025	837125	0.44	444540	0.54	4-(4-METHOXY-2,3,6-TRIMETHYLPHENYL)-3-BUTEN-2-ONE	203.15
18	15.323	53099402	27.75	13883045	16.86	STAHLIANTHUSONE	228.15
19	15.400	1881980	0.98	673502	0.82	KAURAN-18-AL, 17-(ACETYLOXY)-, (4.BETA.)-	55.10
20	15.478	2124391	1.11	1116813	1.36	6-Isopropenyl-4,8a-dimethyl-1,2,3,5,6,7,8,8a-octahydro-naphthalen-2-ol	159.15
21	16.954	21831912	11.41	8480889	10.30	Anthracene, 1,2,3,4,5,6,7,8-octahydro-9,10-dimethyl-	199.15
22	17.435	2762396	1.44	1553891	1.89	6-Acetyl-5-methoxy-2,7-dimethyl-1,4-naphthoguinone	243.15
23	17.849	2497806	1.31	1255245	1.52	2-[5-(2,2-Dimethyl-6-methylene-cyclohexyl)-3-methyl-pent-2-enyl]-[1,4]benzoquinone	161.15
24	18.143	7679643	4.01	3682104	4.47	Coumarine, 8-allyl-7-hydroxy-6-ethyl-4-methyl-	229.10
25	18.274	1103889	0.58	590686	0.72	4-(6,6-Dimethyl-2-methylene-cyclohex-3-enylidene)-pentan-2-one	161.10
26	18.574	3722600	1.95	1709208	2.08	6-METHYL-5-(1-METHYLETHYLIDENE)-3,6,9-DECATRIEN-2-ONE	161.15
		191324856	100.00	82332228	100.00		

Fig. 5 S. Rhizome extract GC-MS profile of *C. vamana*.



Peak Report TIC

Peak#	R.Time	Area	Area%	Height	Height%	Name	Base m/z
1	6.880	13482951	3.85	7358916	6.30	Camphor	95.10
2	7.119	1016716	0.29	625873	0.54	Isoborneol	95.10
3	9.599	1174082	0.33	795546	0.68	DELTA-ELEMENE	121.10
4	10.353	1804703	0.51	1173546	1.00	BETA-ELEMEN	93.10
5	10.476	6587520	1.88	4314109	3.69	(-)-ZINGIBERENE	119.10
6	10.655	1090806	0.31	690695	0.59	trans-.alpha.-Bergamotene	93.10
7	10.824	1179143	0.34	731876	0.63	Caryophyllene	93.10
8	10.882	1769906	0.50	1205561	1.03	.gamma.-Elemene	121.10
9	11.101	10318809	2.94	6546647	5.60	(E)-.beta.-Famesene	69.05
10	11.563	61786529	17.62	16548517	14.16	ALPHA-CURCUMEN	132.15
11	11.622	2212940	0.63	1097979	0.94	(-)-GERMACRENE D	161.15
12	11.699	3138240	0.90	1612036	1.38	Alpha-Cedrene	119.10
13	11.744	11830191	3.37	7054561	6.04	Curzerene	108.05
14	11.939	73820586	21.05	17915798	15.33	.beta.-curcumene	119.10
15	12.070	1659151	0.47	719168	0.62	(+)-.BETA.-FUNEBRENE	69.05
16	12.443	4412459	1.26	2858216	2.45	7-epi-cis-sesquisabinene hydrate	69.05
17	12.626	7836700	2.24	4710009	4.03	GERMACRENE B	121.15
18	13.047	16903760	4.82	7499597	6.42	N-(4-HYDROXYPHENYL)-N,N',N'-TRIMETHYLSULFAMIDE	122.05
19	13.218	2892265	0.82	1101075	0.94	Cedren-13-ol, 8-	119.10
20	13.747	1018292	0.29	605917	0.52	(+)-.BETA.-EUDESMOL	59.05
21	13.832	3816277	1.09	1883865	1.61	.beta.-bisabolol	82.10
22	14.008	2974774	0.85	607504	0.52	7-epi-trans-sesquisabinene hydrate	119.10
23	14.222	14370310	4.10	6019596	5.15	GERMACRON	107.10
24	14.280	3689904	1.05	2372986	2.03	2-Heptanone, 6-methyl-6-[3-methyl-3-(1-methylethenyl)-1-cyclopropen-1-yl]-	135.10
25	14.754	90585245	25.83	15422794	13.20	Phenol, 2-methyl-5-(1,2,2-trimethylcyclopentyl)-, (S)-	136.10
26	14.994	1009889	0.29	588202	0.50	5,8-Dihydroxy-4a-methyl-4,4a,4b,5,6,7,8,8a,9,10-decahydro-2(3H)-phenanthrenone	159.10
27	15.259	2498660	0.71	1514231	1.30	Isovelleral	232.10
28	15.375	1591705	0.45	888073	0.76	Dehydrozingerone	145.00
29	17.028	3104241	0.89	1783371	1.53	Cyclohexyldimethylsilyloxybenzene	151.05
30	17.337	1056308	0.30	608050	0.52	(-)-ISOLONGIFOLINE	175.05
		350633062	100.00	116854314	100.00		

Fig. 5 T. Methanolic rhizome extract GC-MS Chromatograms of *C. zanthorrhiza*.

5.2 UFLC analysis

5.2.1 Method Development and Optimization

5.2.1.1 Selection of wavelength for determination

The standard solution of curcumin (1mg/mL) was scanned in the wavelength range of 300-600nm. The responses of the standard solution measured with PDA detector showed a good resolution at 421 nm using the UFLC method.

5.2.1.2 Selection of Columns

For conducting UFLC analysis, various columns are available but our main objective was to determine the curcumin content from different *Curcuma* species. the C18 (150mm x 4.6mm i.d., 5µm) column was selected for curcumin analysis and this column was chosen to give good peak and high resolution, besides high peak symmetry, good retention in curcuminoid separation from all the species without any interference within a short run time.

5.2.1.3 Selection of mobile phase

The chromatographic conditions were optimized for the method, which can separate curcuminoids from the crude extract with a good resolution. The mobile phase was chosen after several trials with water, methanol, phosphoric acid and acetonitrile (ACN). Finally, the mobile phase consists of water+0.2% (v/v) phosphoric acid (A), and acetonitrile (B) in an isocratic mode was selected to achieve maximum separation. Flow rates were selected between 0.5 and 2.5 ml/min. A flow rate of 1 ml/min gave an optimum signal/noise ratio with a reasonable separation time (30 min) and the best separation efficiency was obtained using C_{18} column.

5.2.1.4 Optimized chromatographic condition

Column	**:**	Kromasil C18 (150mm x 4.6mm i.d., 5µm)
Mobile phase	**:**	Acetonitrile and Water+0.2% (v/v) phosphoric acid, pH was adjusted at 5.0)
Detection wavelength	**:**	421nm
Flow rate	**:**	1 ml/min
Injection volume	**:**	20µl
Column Oven Temperature	**:**	40 ^{o} C
Sample Cooler Temperature	**:**	25 ° C

5.2.2 Method validation

5.2.2.1 Solution stability

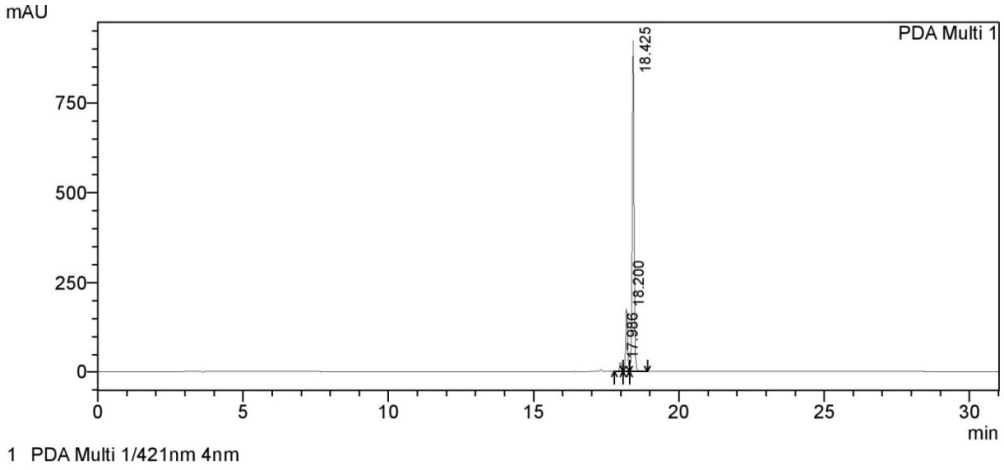
Stability of the analytes in the sample solution was evaluated by analyzing the sample solution of 5µg/ml at 4, 8, 12, 16, 20 and 24 hr after the preparation of sample at room temperature. The analytes were stable in the sample solution with RSDs values between 0.25% and 1.76% for all targeted compounds. The stability is supported by the low%RSDs, as shown in Table 5.2 and demonstrates that this method had good stability.

Table 5.2: the standard solution stability of various curcuminoids obtained from UFLC method.

	CUR	DMC	BDMC
%RSD	0.25	1.70	1.76

5.2.2.2 Specificity

The present analytical method observed was specific with no interference in curcuminoids separation. The analytical result showed that the curcuminoid peaks were free from any impurities (Fig. 5.1).



PeakTable

PDA Ch1 421nm 4nm

Peak#	Ret. Time	Area	Height	Area %	Height %
1	17.986	124497	27410	2.416	2.434
2	18.200	813072	177264	15.780	15.738
3	18.425	4215066	921651	81.804	81.828
Total		5152635	1126326	100.000	100.000

Fig. 5.1. UFLC chromatogram of curcumin standard solution (1mg/ml) at wavelength using C18 column

5.2.2.3 Linearity and range, limits of detection and quantification

The curcuminoid contents were determined to prepare calibration graphs of two *Curcuma* species. The calibration curves were obtained from three injections of five different concentrations of curcuminoids versus the peak area. Linearity was determined in the concentration range between 1µg and 300µg, with high reproducibility and accuracy (Fig. 5.2A-C). Regression analysis of the experimental data points showed a linear relationship with excellent correlation coefficients (r^2) of curcumin, demethoxycurcumin, and

bisdemethoxycurcumin of 0.998, 0.994, and 0.993 respectively. The linear regression equations for the curves for curcumin, demethoxycurcumin, and bisdemethoxycurcumin were $y = 30805x + 889.6$, $y = 832.6x - 30839$, and $y = 7042x - 20553$, respectively.

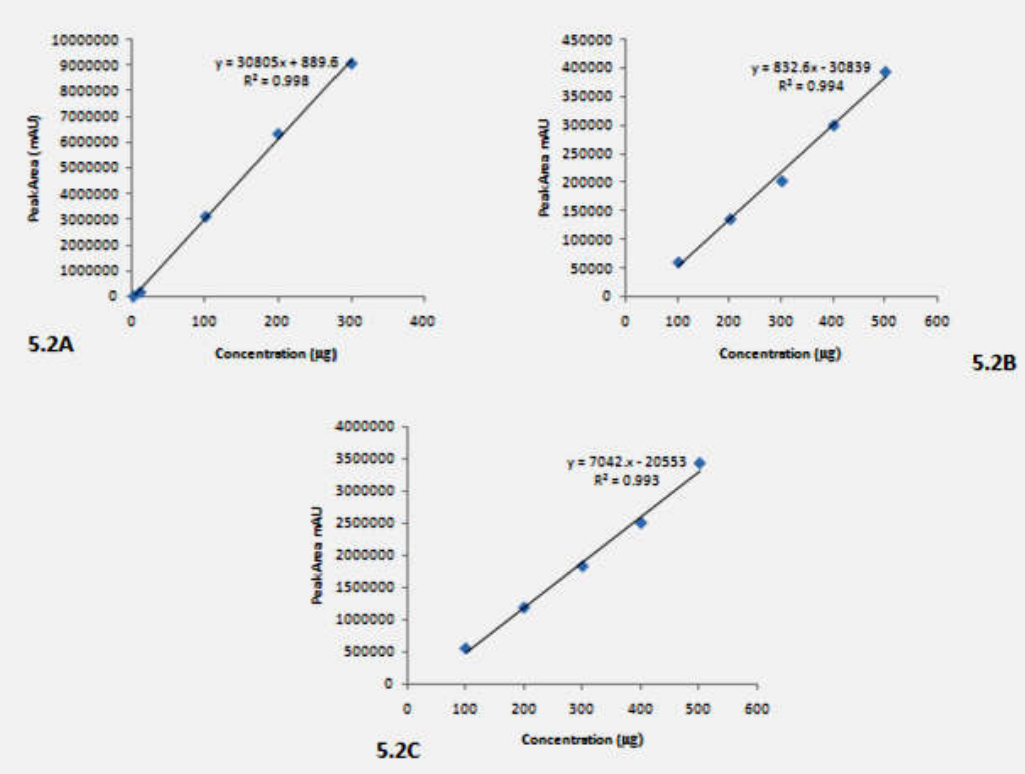


Fig. 5.2. Linear relationship between peak area and concentration of **A-** curcumin, **B-** demethoxycurcumin and **C-** bisdemethoxycurcumin**.**

5.2.2.4 Precision

The intra-day and inter-day precision was estimated by analyzing six independently prepared samples using the same method, operator, instruments, and equipment. Intraday assay variation was evaluated by injecting curcumin (1µg/ml) six times during the same day. Inter-day assay variation was assessed by injecting three curcumin concentrations (1, 3, and

5µg/ml) in replicates of three during three different days. The relative standard deviation was determined for evaluating the precision. RSDs for intra-day and inter-day precision were below 4.5% for all of the compounds. The% RSD for intra-day precision was found to be in the range of 1.00-0.27, 2.81-1.14 and 2.92-1.38 while inter-day precision was found to be in the range of 0.01-1.61, 0.02-1.71 and 0.64-0.85 for CUR, DMC and BDMC, respectively (Table 5.3).

Table 5.3: Intra-day and Inter-day precision parameters for curcumin, demethoxycurcumin and bisdemethoxycurcumin by using different concentrations of the sample. The experiments were repeated as Intra-day (n= 6) and Inter-day (n=3).

Analyte Concentration (1µg, 3µg and 5µg/mL)		Intra-day (n= 6)		Inter-day (n=3)	
		Detected (µg/mL)	%RSD	Detected (µg/mL)	%RSD
CUR	1µg	1.4328±0.01	1.00	1.8023±0.12	1.61
	3µg	2.4841±0.01	0.70	2.4313±0.38	0.03
	5µg	4.7475±0.01	0.27	4.0605±0.56	0.01
DMC	1µg	1.2391±0.03	2.81	1.5648±0.01	1.02
	3µg	1.7163±0.05	3.37	2.9037±0.01	1.71
	5µg	3.9308±0.04	1.14	4.4844±0.02	0.58
BDMC	1µg	1.47±0.04	2.92	1.6715±0.01	0.79
	3µg	2.6208±0.03	1.38	2.99±0.02	0.85
	5µg	4.8625±0.07	1.44	3.4515±0.02	0.64

5.2.2.5 Accuracy

Three different quantities (low, medium, and high) of the standards were added into the sample to evaluate the percentage recovery of the

developed analytical method. The amount of each component was subsequently calculated from the calibration curves. The method had an acceptable accuracy with the overall recovery from 99.15 to 105.50% and percentage RSDs below 3.25% for the three marker compounds. The value of overall recovery and low percentage RSDs showed that the methods used in the analysis had a good accuracy. A good recovery indicates the accuracy of the method. The results are summarized in Table 5.4.

Table 5.4: Results of recovery studies of curcuminoids extracts of *C. longa, C. aromatica, C. zanthorrhiza* and *C. zedoaria* (n=3).

Samples	Amt. of Sample (µg/mL)	Amt. of Std Added (µg/mL)	Amt. found (µg/mL)	%Recovery± S.D. (n=3)	%RSD
CUR	5	4	8.99	99.90±1.22	1.22
	5	7	11.89	99.16±2.21	2.23
	5	9	14.77	105.50±1.46	1.39
DMC	4	4	7.97	99.68±1.77	1.78
	4	6	9.92	99.29±3.18	3.20
	4	8	11.83	98.58±3.20	3.24
BDMC	5	1	5.90	98.33±3.15	3.20
	5	2	6.91	98.77±1.26	1.27
	5	3	7.93	99.15±2.59	2.61

5.2.2.6 Limit of detection and limit of quantification:

The estimated LOD and LOQ in this study were found to be 0.72-0.92 and 0.80-1.42 respectively (Table 5.5).

Table 5.5: Result of limit of detection and limit of quantification

Curcuminoids	LOD (µg/mL)	LOQ (µg/mL)
Curcumin	0.72	1.42
Demethoxycurcumin	0.92	0.80
Bisdemethoxycurcumin	0.86	1.04

5.2.3 Determination of curcuminoids in *C. zanthorrhiza, C. aromatica, C. longa* and *C. zedoaria.*

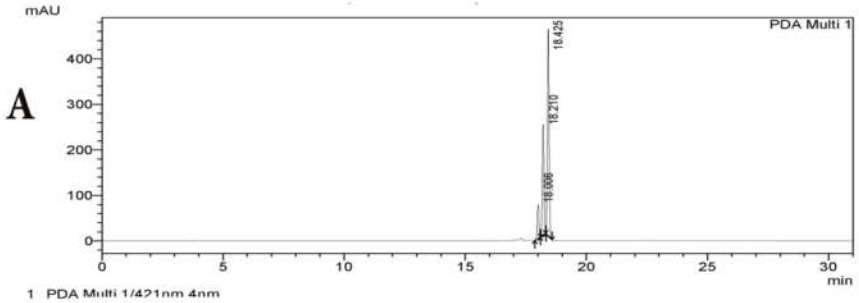
UFLC method was successfully utilized for the determination of CUR, DMC and BDMC from economically important *Curcuma* species like *C. longa*, *C. aromatica*, *C. zanthorrhiza* and *C. zedoaria*. Using the calibration curve of each compound, the three analytes in the four *Curcuma* species were determined (Table 5.6) and it was observed that the levels of the three individual analytes present in the samples varied considerably. The highest curcumin content was determined in the nine months old *Curcuma* species (15.0-1705.7µg/g) followed by six (14.5-352.6µg/g) and three (3.7-111.6µg/g) months old rhizomes. The highest curcumin content was detected in *C. longa* (1918.3µg/g) followed by *C. zanthorrhiza* (1240.4µg/g), *C. aromatica* (91.8µg/g) and *C. zedoaria* (33.2µg/g).

However, the other curcuminoids, DMC (8.55-54106.8µg/g) and BDMC (1.4-6683.3µg/g) were quantitatively different among different species. It is well known that the quality of herbal products can be influenced by many factors, such as the cultivating site, harvesting time, and post-harvest

handling. The percentage of curcuminoids was found to be satisfactory (Table 5.6). The compounds identified were curcumin, demethoxycurcumin and bisdemethoxycurcumin and retention time and peak area of all the species shown in Fig. 5.6-5.9.

Table 5.6: Curcuminoids concentration of different aged plants of *C. zanthorrhiza*, *C. aromatica, C. longa* and *C. zedoaria.*

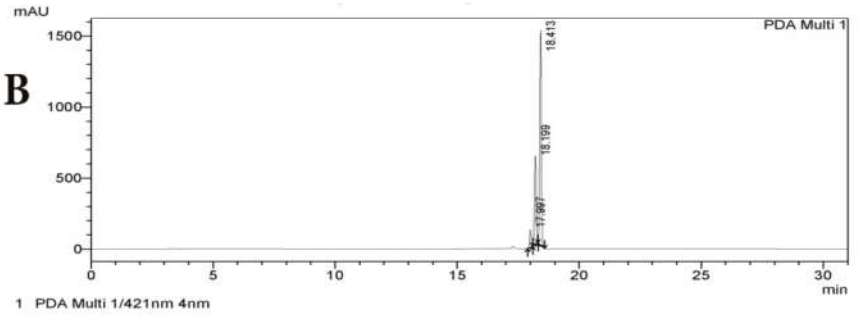
Curcuminoids	Months	*C. zanthorrhiza*	*C. aromatica*	*C. longa*	*C. zedoaria*
CUR	3	111.6µg/g	15.5µg/g	72.5µg/g	3.7µg/g
	6	352.6µg/g	30.3µg/g	140.1µg/g	14.5µg/g
	9	776.2µg/g	46.0µg/g	1705.7µg/g	15.0µg/g
DMC	3	1070 µg/g	153.1µg/g	3812.1µg/g	8.5µg/g
	6	615.4 µg/g	100.0µg/g	1759.7µg/g	203.8µg/g
	9	1557.9 µg/g	104.0µg/g	5410.8µg/g	163.6µg/g
BMDC	3	692.5 µg/g	92.6 µg/g	666.0µg/g	1.4µg/g
	6	266.4 µg/g	59.5 µg/g	270.3µg/g	6.9µg/g
	9	2135.4 µg/g	17.0 µg/g	6683.3µg/g	3.6µg/g



PeakTable

PDA Ch1 421nm 4nm

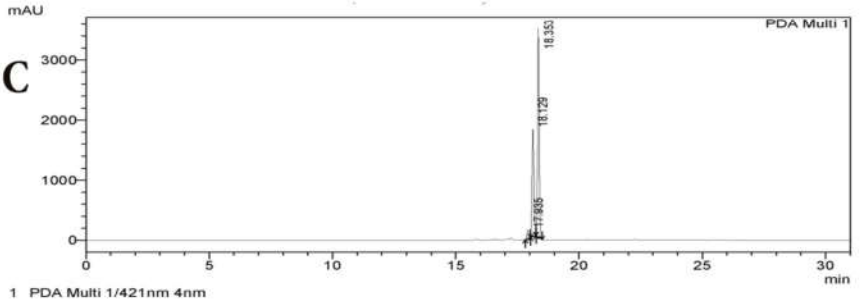
Peak#	Ret. Time	Area	Height	Area %	Height %
1	18.006	338304	74361	9.533	9.641
2	18.210	1146264	243415	32.301	31.561
3	18.425	2064162	453485	58.166	58.798
Total		3548730	771261	100.000	100.000



PeakTable

PDA Ch1 421nm 4nm

Peak#	Ret. Time	Area	Height	Area %	Height %
1	17.997	565827	125183	5.640	5.528
2	18.199	2946844	625398	29.376	27.618
3	18.413	6518936	1513860	64.984	66.854
Total		10031606	2264441	100.000	100.000

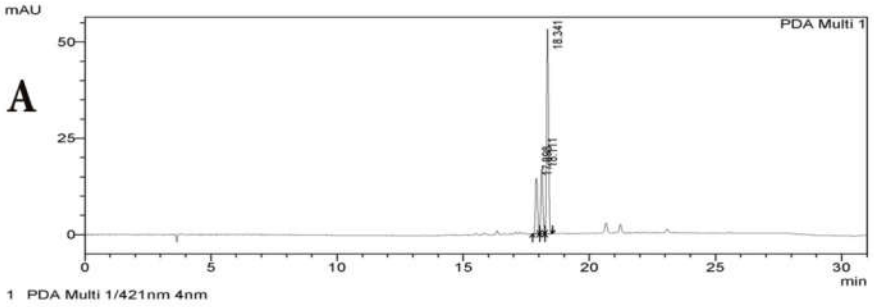


PeakTable

PDA Ch1 421nm 4nm

Peak#	Ret. Time	Area	Height	Area %	Height %
1	17.935	809149	149462	3.343	2.763
2	18.129	9043198	1786040	37.367	33.017
3	18.353	14348880	3473898	59.290	64.220
Total		24201228	5409400	100.000	100.000

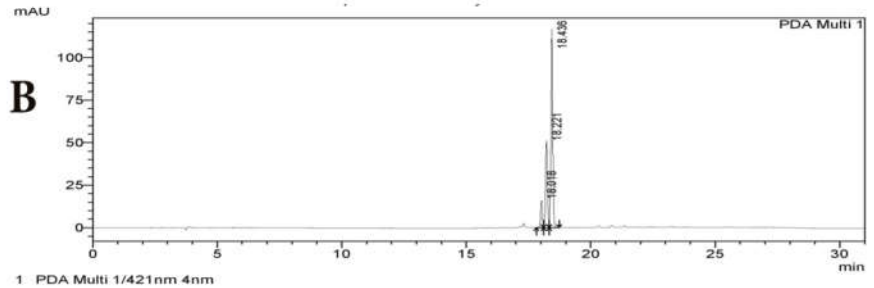
Fig. 5.6. *C. zanthorrhiza* UFLC chromatograms of curcuminoids-**A-**3rd, **B-** 6th and **C**-9thmonths.



PeakTable

PDA Ch1 421nm 4nm

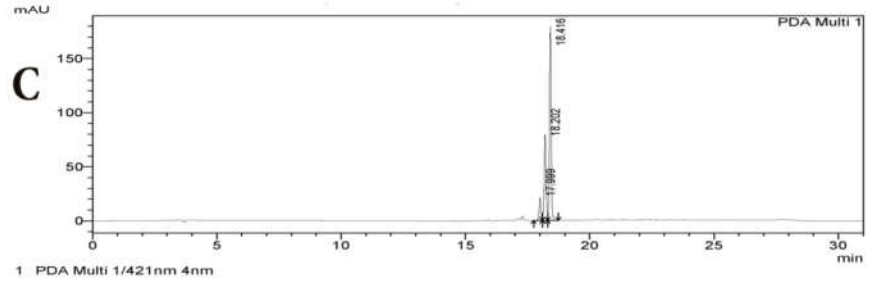
Peak#	Ret. Time	Area	Height	Area %	Height %
1	17.898	83137	14388	17.924	17.106
2	18.111	92411	16593	19.924	19.727
3	18.341	288275	53132	62.152	63.167
Total		463824	84113	100.000	100.000



PeakTable

PDA Ch1 421nm 4nm

Peak#	Ret. Time	Area	Height	Area %	Height %
1	18.018	81172	15946	8.870	8.685
2	18.221	272289	51042	29.754	27.801
3	18.436	561677	116610	61.376	63.514
Total		915137	183598	100.000	100.000

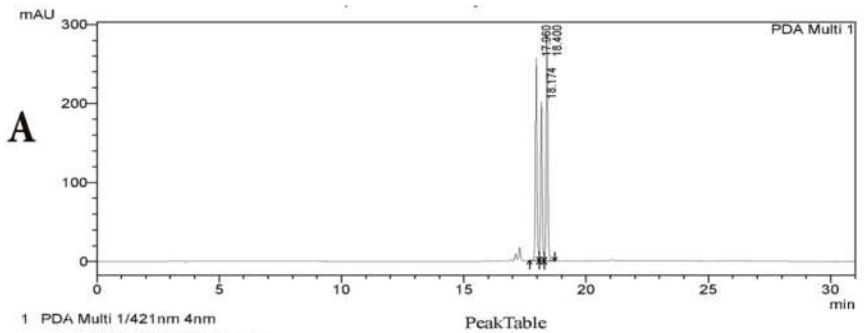


PeakTable

PDA Ch1 421nm 4nm

Peak#	Ret. Time	Area	Height	Area %	Height %
1	17.999	107368	21156	7.831	7.565
2	18.202	412058	79469	30.054	28.417
3	18.416	851612	179026	62.114	64.018
Total		1371039	279651	100.000	100.000

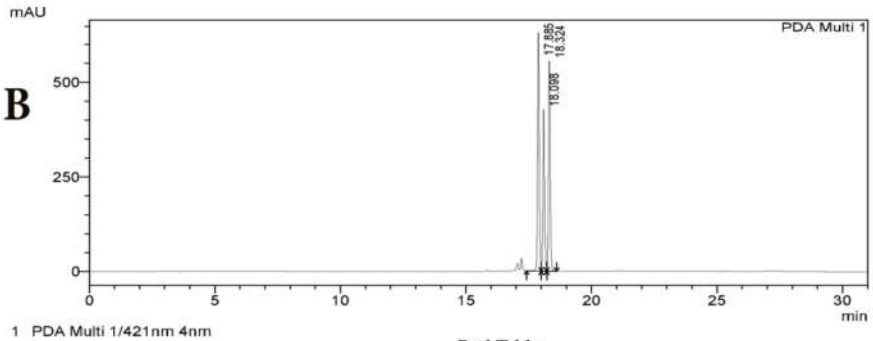
Fig. 5.7. *C. aromatica* UFLC chromatograms of curcuminoids-**A**-3rd, **B**-6th and **C**-9th months.



PeakTable

PDA Ch1 421nm 4nm

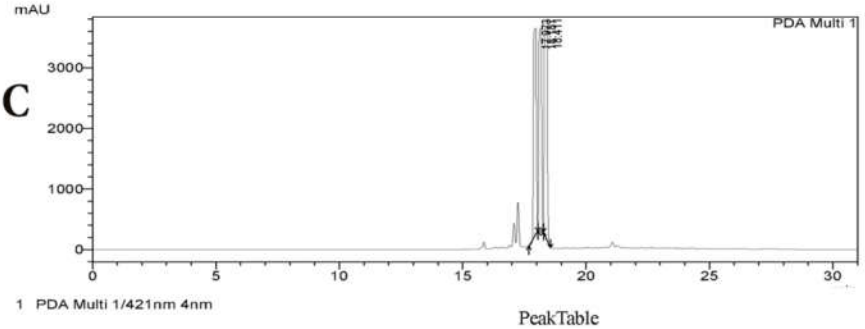
Peak#	Ret. Time	Area	Height	Area %	Height %
1	17.960	1162916	256333	34.059	34.486
2	18.174	909944	200919	26.650	27.030
3	18.400	1341534	286053	39.291	38.484
Total		3414394	743305	100.000	100.000



PeakTable

PDA Ch1 421nm 4nm

Peak#	Ret. Time	Area	Height	Area %	Height %
1	17.885	2834874	630557	38.513	39.077
2	18.098	1935217	427180	26.290	26.473
3	18.324	2590817	555906	35.197	34.450
Total		7360909	1613643	100.000	100.000

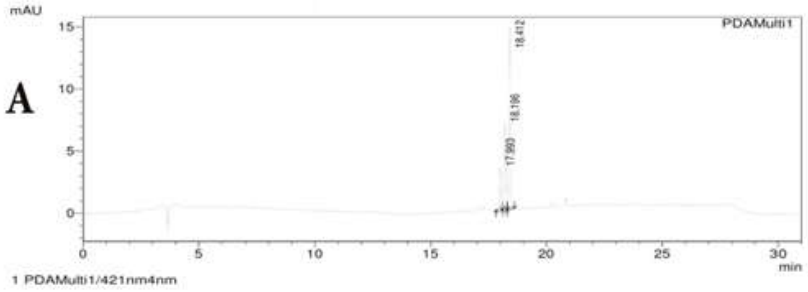


PeakTable

PDA Ch1 421nm 4nm

Peak#	Ret. Time	Area	Height	Area %	Height %
1	17.973	28259188	3396863	32.538	33.191
2	18.187	27060403	3353966	31.158	32.772
3	18.411	31528898	3483513	36.303	34.037
Total		86848489	10234342	100.000	100.000

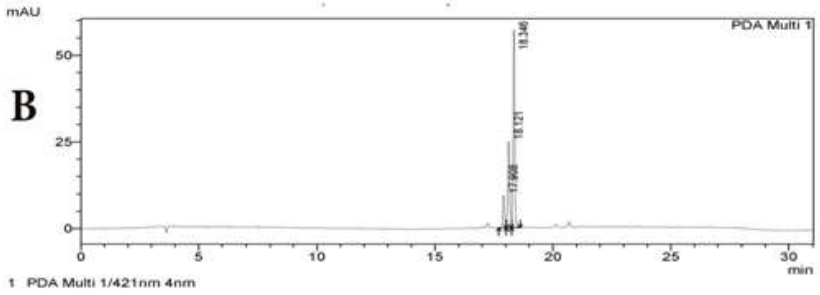
Fig. 5.8. *C. longa* UFLC chromatograms of curcuminoids-**A-**3rd, **B-**6th and **C-**9th months.



PeakTable

PDACh1421nm4nm

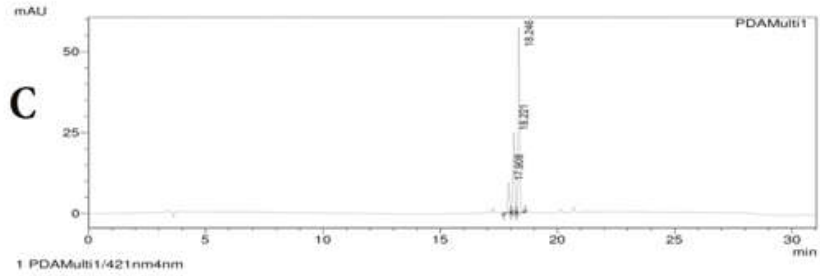
Peak#	Ret.Time	Area	Height	Area %	Height%
1	17.993	26835	3291	13.793	13.344
2	18.196	35113	6850	28.768	27.780
3	18.412	70107	14518	57.439	58.876
Total		132055	24659	100.000	100.000



PeakTable

PDA Ch1 421nm 4nm

Peak#	Ret. Time	Area	Height	Area %	Height %
1	17.908	49930	9324	11.042	10.220
2	18.121	132690	24822	29.345	27.208
3	18.346	269554	57086	59.613	62.572
Total		452175	91232	100.000	100.000



PeakTable

PDACh1421nm4nm

Peak#	Ret.Time	Area	Height	Area %	Height%
1	17.908	35930	8314	12.042	11.120
2	18.221	112590	21522	28.945	28.408
3	18.246	278594	58036	59.013	60.472
Total		427116	87872	100.000	100.000

Fig. 5.9. *C. zedoaria* UFLC chromatograms of curcuminoids-**A**-3rd, **B**-6th and **C**-9th months**.**

Table 5.7: Summary of Validation Parameters

Parameter		CUR	DMC	BDMC
Linearity Range (µg/ml)		1-300µg	1-300µg	1-300µg
Correlation coefficient (R^2)		0.998	0.994	0.993
Regression equation		y=30805x + 889.6	y =832.6x - 30839	y = 7042x - 20553
Precision (% R.S.D)	**Intraday precision (n=6)**	1.00-0.27	2.81-1.14	2.92-1.44
	Interday precision (n=3)	1.61-0.01	1.02-0.58	0.85-0.64
Accuracy (%recovery) (n=3)		99.16-105.50	98.58-99.68	98.33-99.15
Limit of Detection (µg/ml)		0.72	0.92	0.86
Limit of Quantitation (µg/ml)		1.42	0.80	1.04

5.3 Discussion

5.3.1 GC-MS

The rhizome extracts of *Curcuma* species has wide range pharmacological properties, such as anti-inflammatory, anticancerous, antiproliferative, hypocholesterolemic, antidiabetic, antihepatotoxic, antidiarrheal, carminative, diuretic, antirheumatic, hypotensive, antioxidant, antimicrobial, antiviral, insecticidal, larvicidal, antivenomous, antithrombotic, antityrosinase and cyclooxygenase-1 (COX-1) inhibitory activities, *etc.* (Dosoky and Setzer, 2018).

Previous studies on the important constituents were carried out using the essential oil extracted from the rhizome, while the present work was carried out using methanolic extracts. The results indicated that methanolic

extracts from these species are a good source of 1, 8 cineole, camphor, camphene, ar-curcumene and beta-elemene, which are known to have great application in pharmaceutical and flavoring industries. These compounds obtained from the essential oils of *Curcuma* species possess analgesic, antimicrobial, anti-inflammatory, insecticidal and anticancer properties (Chen *et al.,* 2013). The present study identified the same compounds from the rhizomes extracts of *Curcuma* species investigated. Fifteen components were identified from the methanolic extracts of *C. amada* (George and Britto, 2016). *C. amada* is a perennial herb called as "mango ginger" and "manga manjal" because of its raw mango flavor that is mainly attributed to the presence of -3-carene, myrcene, and (Z)-ocimene (Varadarajan *et al.,* 2018). It is used in culinary preparations, medicines, and as a source of starch (Policegoudra *et al.,* 2011). Myrcene is another dominant compound obtained from the rhizome essential oil of *C. amada* (Padalia *et al.,* 2013). Other components present in *C. amada* were (E)-hydroocimene, (Z)-hydroocimene, myrcene, and linalool, ar-curcumene, camphor, curzerenone, and 1, 8-cineole (Srivastava *et al.,* 2001). The major chemical constituents in *C. aromatica* are camphor, camphene, borneol, 1, 8-cineole, β-elemene, 8, 9-dehydro-9-formyl-cycloisolongifolene, germacrone, ar-turmerone, turmerone, curdione (Xiang *et al.,* 2018), camphor (Herath *et al.,* 2017; Angel *et al.,* 2014), xanthorrhizol, ar-curcumene, di-epi-cedrene (Nampoothiri *et al.,* 2015).

Essential oils exhibit a number of biological activities and hence have important applications in the pharmaceutical industry. Beta-elemene, with tumouricidal effects, was detected in the rhizomes of *C. aromatica, C. brog, C. zedoaria* and essential oils (Fu *et al.,* 1984). These components were identified in the rhizome methanolic extracts of *C. aromatica*, *C. zedoaria, C. zanthorrhiza* and *C. aeruginosa.* Curcumenol with analgesic properties was detected in the *C. aeruginosa* oil (De Fatima *et al.,* 2002). Camphor is one of the main compound responsible for developing the cholagogic effect of

essential oils (Ozaki and Liang, 1988), which is a common constituent in the essential oil of *Curcuma* species and also is detected in the methanolic extracts of *Curcuma* species.

The major components present in the *C. longa* essential oil were ar-turmerone, α-tumerone, β- turmerone, β-caryophyllene and eucalyptol (Naz *et al.,* 2010). The major components reported in the essential oil of *C. inodora* were caryophyllene and alpha-caryophyllene (Rajesh Kumar and Yusuf, 2017) and curzerenone, germacrone, curdione, and 1,8-cineole (Malek *et al.,* 2006). The chemical composition of essential oils from eight underutilized starchy *Curcuma* species was studied (Angel *et al.,* 2014), and the major compounds identified were, 1, 8 cineole, camphor, camphene, α -pinene, and β-pinene. The components of essential oil were identified form 12 *Curcuma* species using GC-MS and the major chemical components were curdione, ar-turmerone, germacrone, β-elemenone and velleral (Zhang *et al.,* 2017). The essential oil of *C. aeruginosa* rhizome contain 8, 9-dehydro-9-formyl-cycloisolongifolene, dihydrocostunolide, germacrone, curzerenone (Theanphong *et al.,* 2015), dehydrocurdione, curcumenol, 1, 8-cineole, germacrone (Srivilai *et al.,* 2018), camphor, germacrone, curcumenol, α-pinene (Angel *et al.,* 2014).

The essential oil of *C. aurantiaca* contain1,8-cineole, camphor, germacrone, β-elemene, curzerene, and β-elemenone (Liu *et al.,* 2012). *C. mangga* rhizome essential oil have two chemotypes, caryophyllene oxide and caryophyllene-rich chemotype and myrcene-dominated chemotype (Wahab *et al.,* 2011). The major compounds detected in the essential oil of *C. oligantha* rhizome are caryophyllene, phytol, humulene, elemene, and caryophyllene respectively (Herath *et al.,* 2017). *C. pseudomontana* rhizome from India contain chemical compounds like β-elemenone, pseudocumenol, germacrone,

2-(4-methoxyphenyl) N, N-trimethyl-1-pyrrolamine and (1,5-dimethyl-4-hexenyl)-4-methylbenzene respectively (Muniyappan *et al.,* 2014).

5.3.2 Relationship between *Curcuma* species

The dendrogram generated using the statistical analysis of the identified compounds (≥ 5% concentration) of each species is shown in Fig. 5.1. Cluster I consists of *C. aeruginosa* and *C. zanthorrhiza*. Cluster II consists of *C. aromatica* and *C. zedoaria.* Cluster III includes *C. amada*, *C. bhatii, C. mutabilis, C. kudagensis* and *C. pseudomontana*. Cluster IV includes *C. coriacea, C. neilgherrensis* and *C. montana*. Cluster V contained *C. haritha*, *C. karnatakensis* and *C. vamana* respectively. Cluster VI harbours the species like *C. longa* and *C. oligantha*. Cluster VII included the species like *C. decipiens* and *C. inodora*. *C. aurantiaca* exist a monoclade in Cluster VIII. The evolutionary relationship of *Curcuma* species using *mat*K and *mat*K+*rbc*L sequence analysis of leaf tissues (Santhoshkumar and Yusuf 2017 & 2018) showed slight similarity with the present dendrogram constructed on the basis of methanolic extracts of rhizome, indicating the usefulness of rhizome extracts data in chemo taxonomical applications.

Curcuma is one of the most important plants that have been used as food and pharmaceutical ingredients worldwide. The identification of major compounds in the rhizome extracts from twenty *Curcuma* species provided a chemical blueprint assisting in the identification of these species. The chemical composition obtained from the rhizome extracts were used to construct the dendrogram divided into eight clusters for the twenty species. In this study, we included some of the wild species that are not been exploited pharmacologically, and thus this study will assist in their species identification and utilization.

5.3.3 UFLC

Ultra fast liquid chromatography (UFLC) is one of the most commonly used methods for the determination of different compounds present in plant material. The UFLC analysis of methanol extract of *Curcuma* plants showed different peaks. Curcumin is one of the major active ingredients of dietary spice found in rhizomes of turmeric (*C. longa*) (Prasad *et al*., 2014). Curcuminoids are polyphenolic substances containing a mixture of curcumin (60%–80%), demethoxycurcumin (15%–30%) and bisdemethoxycurcumin (2%–6%) (Wichitnithad *et al.,* 2009) and it is soluble in methanol, ethanol, DCM or DMSO and insoluble in water (Tonnesen and Karlsen, 1985). Curcuminoids have medicinal properties such as anti-inflammatory, antimutagenic, anti-diabetic, anti-bacterial, hepatoprotective activities (Krup *et al.,* 2013). It has antioxidant activity as well as free-radical scavenging properties (Kalpravidh *et al.,* 2010), healing of dermal wound (Gopinath *et al.,* 2004), and prevention of Alzheimer's disease (Shen and Ji, 2012). Most importantly curcumin inhibit the cell growth of various cancer cell lines, induces apoptosis of cancer cells and is also effective on the cell-cycle regulation of cancer cells (Liu *et al.,* 2007). Curcumin is sensitive to light, hence it is suggested that their biological samples containing curcumin should be protected from light (Prasad *et al.,* 2014).

5.3.4 Validation of UFLC method

The pure curcuminoids stock solution (1mg/ml) was prepared. Standard solutions were also prepared using the stock solution of individual curcumin and derived the calibration curve using the linearity range 1-300 µg/ ml (Fig. 5.1A-C).

Generally ethanol is used as the mobile phase in HPLC analysis of curcuminoids. The three curcuminoids (curcumin, demethoxycurcumin and

bisdemethoxycurcumin) was first isolated by Tonnesen and Karlsen (1983) using HPLC in a Nucleosil NH_2 column stationary phase, flow rate of 1.2 m1 as mobile phase of ethanol. Other than ethanol the various solvents such as methanol, acetonitrile and tetrahydrofuran are used as stationary phase (C_{18} and NH_2) and different methods (isocratic and gradient elution) are used for the analysis of curcuminoids (Cooper *et al.,* 1994). Curcuminoids were analyzed using HPLC with acetonitrile: water containing (0.2% phosphoric acid), 60:40 (v/v). Individual pure peak could be seen in the HPLC chromatogram. The HPLC analysis of compounds showed single peaks with retention times of 18.4, 18.2 and 17.9 min, respectively. Spectroscopic data confirmed that the single peak at 18.4, 18.2 and 17.9 min represents curcumin, demethoxycurcumin and bisdemethoxycurcumin respectively. Jayaprakasha *et al.* (2002) reported that HPLC chromatogram showed three peaks, followed by Goel *et al.* (2008) identified these peaks with the major one as curcumin and two minor peaks were identified as demethoxycurcumin and bisdemethoxycurcumin determined by co-injection of curcumin standards and confirmed elusions of peaks by ^{1}H and ^{13}C NMR analysis. The peaks for curcumin, demethoxycurcumin and bisdemethoxycurcumin within the two species were observed and identified in the chromatogram.

Jayaprakasha *et al.* (2002) determined curcuminoids using HPLC separation with C_{18} column, methanol, 2% AcOH, and acetonitrile was used as mobile phase, while, Wichitnithad *et al.* (2009) reported a simple isocratic method with the elution of acetonitrile and 2% acetic acid (40:60, v/v) at a flow rate of 2.0 mL min for separation of the three curcuminoids. The isocratic method served as the best method for the separation of all the three curcuminoids. Another study Jadhav *et al.,* (2007) reported that RP C-18 column and acetonitrile: water (50:50 v/v) and 0.1% trifluroacetic acid was used as mobile phase. The retention times of bisdemethoxycurcumin, demethoxycurcumin and curcumin were recorded at retention time of 7.2,

8.12 and 9.0 min interval, respectively. Revathy *et al*. (2011) reported HPLC method for the analysis of curcuminoids usingC18 column and mobile phase comprising of tetrahydrofuran: water containing 1% citric acid (40:60) were used to separate curcumin, demethoxycurcumin and bisdemethoxycurcumin at retention time of 10.81, 12.79, 13.03 min, respectively. Erpina *et al.,* (2017) reported the HPLC analysis of curcuminoids using the Phenomenex C_{18} column with acetonitrile-0.001% formic acid used as a mobile phase to separate curcuminoids at retention time of 9.11, 9.96 and 10.82. The separation of curcuminoids using HPLC with fluorescence detector reported by Tonnesen and Karlsen (1986). Rafi *et al.,* (2015) reported the HPLC analysis of curcuminoids in *C. zanthorrhiza* and *C. longa* samples using C_{18} column with acetonitrile and 0.5% acetic acid in water used as a mobile phase to separate of curcuminoids at retention time of 9.07, 9.84 and 10.65 respectively. Separation curcuminoids (CUR, DMC and BDMC) in the HPLC system is known to be influenced by the mobile phase (Jayaprakasha *et al.,* 2002).

The previous studies reported the analysis of curcuminoids from different *Curcuma* species, including *C. longa*, *C. zanthorrhiza, C. mangga, C. heyneana, C. aeruginosa, C. comosa, C. longa, C. malabarica, C. montana, C. zedoaria* and *C. soloensis* using HPLC (Syamkumar, 2008 thesis). In the present study we analyzed curcuminoids from four *Curcuma* species *viz*, *C. longa, C. aromatica, C. zedoaria* and *C. zanthorrhiza* during different growth periods using the developed UFLC method. The highest percentage of curcuminoids was detected in the nine months old *C. longa* and the lowest was in *C. zedoaria*. However, three and six month old rhizomes of *C. zanthorrhiza* showed higher curcumin content compared with *C. longa*. Higher amount of DMC content showed in *C. zedoaria* compared to CUR and BDMC.

5.3.5 Validation of analytical methods

The developed method was evaluated for the calculation of Limit of detection (LOD) and Limit of quantification (LOQ) of the three curcuminoids (CUR, DMC and BDMC) employing two different approaches. The primary approach, i.e. signal to noise ratio technique is generally used in HPLC method development. Second approach used the dispersion characteristics of the regression line with the help of peak area and concentration of analyte (Roughley *et al.*, 1973). Based on this approach, standard error of peak areas for known concentrations was used for the detection and quantification of curcuminodis. Using signal to noise ratio of 3, LOD were calculated as 0.72, 0.92, 0.86 µg/ml for curcumin, demethoxycurcumin and bisdemethoxy-curcumin respectively. LOQ for curcumin were 1.42µg/ml, demethoxy-curcumin were 0.80µg/mL and bisdemethoxycurcumin were 1.04 µg/ml. Our results are in agreement with the previously reported studies. However, important variations were reported in the LOD and LOQ of these compounds. Syed *et al.* (2015) reported 0.3 µg/ml LOD of curcumin, whereas Wichitnithad *et al.* (2009) reported 0.9 µg/ml and lover LOD was repoted by Rafi *et al.,* (2015) 1.53ng/ml. LOQ of curcumin, demethoxycurcumin, bisdemethoxycurcumin were reported to be found 2.73. 2.53 and 0.23 µg/ml respectively (Wichitnithad *et al.*, 2009). Interestingly, much lower LOD values of curcumin, demethoxycurcumin, bisdemethoxycurcumin as 27.9, 31.9 and 21.8 ng/ml (Jadhav *et al.*, 2007) and 5.11, 5.47 and 4.21 ng/ml reported (Rafi *et al.,* 2015) attributed to different detectors used in the analysis.

Accuracy and precision was calculated for the individual curcuminoids after intra-day and inter-day runs at different concentration level of 1, 3 and 5µg/ml (Table 5.3 and 5.4). The intra-day accuracy ranged between 0.85 to 1.61% and Intra-day 1.00 to 2.82% so this result indicated efficiency of the method. All the data are within the standard criteria i.e. below 5% level.

CHAPTER 5
PHYTOCHEMICAL ANALYSIS

5.1 GC-MS based chemical profiling

The *Curcuma* species used in the present study were collected from different regions of South India and cultivated under identical conditions so as to rule out any variability in the chemical profile of methanolic extracts due to environmental factors. Unpeeled rhizomes were used to make rhizome extracts. The colour of most of the *Curcuma* rhizomes was light blue, yellow or gray, while color of methanolic extracts was brownish or yellow. The color of the rhizomes extracts did not correlate with the color of rhizome and little information is available about the constituents of rhizome extracts of *Curcuma* species.

The composition of the methanolic extracts was determined using GC–MS and individual components were identified by comparing the relative retention indices of the peaks with the NIST library of standard extracts. The number of chemical constituents of individual species ranged from 6 to 35. The maximum number of compounds was detected in *C. aeruginosa, C. aromatica C. decipiens, C. longa, C. zanthorrhiza* and *C. zedoaria* rhizome extracts. The identified major components of the rhizome extracts of *Curcuma* species are given below:

C. aeruginosa

The GC-MS profile of the rhizome extract from the species *C. aeruginosa* is represented in (Fig. 5 A**)**. A total of 28 major components were identified viz, butyl 9, 12-octadecadienoate (16.17%), (-)-isolongifolene (4.8%), camphor (2.3%), eucalyptol (1.8%), isovelleral (5.4%), (E)-.beta.-

famesene (5.3%), (+)-isocurcumenol (4.3%) beta-elemen-(2) (2.2%), 4-[(3,4-dimethylphenyl)sulfanyl]aniline (11.3%), 1,5,9-trimethyl-1,5,9-cyclododecatriene (23%) and isopulegol 2 (1.2%).

C. amada

Seventeen compounds were identified form the GC-MS chromatogram of rhizome extract of *C. amada.* The major components are beta-caryophyllene oxide (75%), beta myrcene (9.4%), acetic acid, 1-[2-(2,2,6-trimethyl-bicyclo[4.1.0]hept-1-yl)-ethyl]-vinyl ester (2.6%), alpha.-springene (1.5%) and card-20(22)-enolide, 3,5,14,19-tetrahydroxy-8-methyl-, (3β,5β,9β) (1.8%) respectively (Fig. 5 B).

C. aromatica

The GC-MS analysis of the methonalic extract of *C. aromatica* rhizome detected 27 compounds represented in Fig. 5 C. curdione (20.5%), isovelleral (15.3%), trimethyl-1, 5, 9-cyclododecatriene (13.9%), camphor (7.9%), germacrone (3.1%), borneol (4.6), 9-cis retinal (2.4%) and isolongifolene (3.3%) are major components.

C. bhatii

Thirteen chemical compounds were identified from the rhizome extracts of *C. bhatii*; the major ones were beta-caryophyllene oxide (87%), kauren-18-ol, acetate, (4beta) (2.4%), retinal (1.8%) and duvatriendiol (1.5%) identified in the GC-MS profile (Fig. 5 D).

C. coriacea

Analysis of *C. coriacea* rhizomes detected 17 compounds; the major components were 1-heptatriacotanol (39.9%), (E)-.beta.-famesene (19.9%), (-)-β-caryophyllene epoxide (12.3%), 1S,2E,4S,5R,7E,11E)-cembra-2,7,11-trien-4,5-diol (10%), corymbolone (3.3%), eucalyptol (3.3%) and

tricyclo[20.8.0.0(7,16)]triacontane, 1(22),7(16)-diepoxy- (2.5%) the GC-MS profile is represented in Fig. 5 E.

C. decipiens

The GC-MS analysis of *C. decipiens* rhizome yielded 28 compounds; kauren-18-ol, acetate, (4.beta.)- (53.8%), (-)-β-caryophyllene epoxide (14.4%), 9-cis retinal (11.4%), kauren-19-ol-acetate (2.6%) and acetic acid, 1-[2-(2, 2, 6-trimethyl-bicyclo [4.1.0] hept-1-yl)-ethyl]-vinyl ester (1.7%) are the major components. The GC-MS profile is represented in Fig. 5 F.

C. aurantiaca

Twenty four compounds were detected from the rhizome extract of *C. aurantiaca*; the major components are 3-Terpinolene (30.7%), dihydrochrysin (18.3%), 3,3-dimethyl-4-phenyl-4-penten-2-one (15%), eucalyptol (8.3%), trans-β-caryophyllene (6.5%), E-nerolidol (4.1%) and camphor (3.5%) as represented in the GC-MS profile (Fig. 5 G).

C. haritha

The major chemical compounds identified from the rhizome extracts of *C. haritha* were isoaromadendrene epoxide (27.8%), (-)-β-caryophyllene epoxide (26.6%), isovelleral (12.3%), 2,4-di-tert-butylphen (7.3%), camphor (4.5%), 2-bornanol (3.5%), 2-camphanol, exo (2.7%) and diazoprogesterone (2.4%). A total of 20 compounds were identified using GC-MS and is represented in (Fig. 5 H).

C. inodora

C. inodora rhizome extract were subjected to GC-MS analysis; a total of 22 compounds were detected, alloaromadendrene oxide-(1) (53.3%), 9-cis retinal (11.3%), (+,-)-β bisabolene (7.7%), andrographolid

(5.7%),2,4a,8,8tetramethyldecahydrocyclopropa[d]naphthalene (2.9%) and (+)-alpha-santalene (1.4%) are the major components. The peaks and the obtained chromatogram are represented in Fig. 5 I.

C. karnatakensis

Rhizome of *C. karnatakensis* detected 19 compounds; the major components were alloaromadendrene oxide (50.8%), 9-cis retinal (11.7%), (-)-β-caryophyllene epoxide (8.3%), 3-bromocholest-5-ene (6.1%), 10-chlorotricyclo [4.2.1.1(2, 5)] deca-3, 7-dien-9-ol (4.5%), kauren-18-ol, acetate, (4.beta.)- (4.5%), 1-ethyl-4,4-dimethyl-cyclohex-2-en-1-ol (3.0%) and ergost-25-ene-6,12-dione, 3,5-dihydroxy-, (3.beta.,5.alpha.)- (2%). The GC-MS profile depicted in Fig. 5 J.

C. kudagensis

Fifteen compounds were detected using the GC-MS analysis of the species *C. kudagensis*; the following compounds being the major components as, kauren-18-ol, acetate, (4.beta.) (38.3%), 9-cis retinal (18%), (-)-β-caryophyllene epoxide (17.3%), isovelleral (4.7%), acetic acid, 1-[2-(2, 2, 6-trimethyl-bicyclo [4.1.0] hept-1-yl)-ethyl]-vinyl ester (3.9%), 2-24-nitro butanoyl 7 cyclohexanone (3.4%) and kauren-19-ol-Acetate (1.3%) the chromatogram are repersented in Fig. 5 K.

C. longa

Rhizome extract of *C. longa* yielded 31 compounds, tumerone (19.1%), (-)-zingiberene (11.6%), beta-sesquiphellandrene (8.8%), (-)-.alpha.-Santalene (8.6%), Curlone (6.4%), ar-tumerone (6.1%), β-bisabolene (5.5%), benzenemethanol, 4-methyl-.alpha.-(1-methyl-2-propenyl)-, (R*,R*) (4.7%), (E)-.beta.-famesene (4%) and butane-1,1-dicarbonitrile, 1-cyclohexyl-3-

methyl-(2.6%) as the major compounds from the extract as shown in the profile Fig. 5 L.

C. montana

Totally 6 components were identified in the GC-MS analysis of the rhizomes of *C. montana;* the major components are alloaromadendrene oxide (60.7%), 2-tert-butyl-4,6-bis(3,5-di-tert-butyl-4-hydroxybenzyl)phenol (12.8%), violanthrene a (11.4%) and (-)-β-caryophyllene epoxide (6.6%) respectively (Fig. 5 M).

C. mutabilis

Fifteen components were identified from the methanolic extract of *C. mutabilis*; the major components were Kauren-18-ol, acetate, (4.beta.)- (59.1%), (-)-β-caryophyllene epoxide (17.2%), retinal (8.1%), trispiro [4.2.4.2.4.2.]heneicosane (3.0%), kauren-19-ol-Acetate (3.0%), 1-ethyl-4,4-dimethylcyclohex-2-en-1-ol (2.4%) and aromadendrene oxide-(2) (1.3%). The GC-MS profile is represented in Fig. 5N.

C. neilgherrensis

C. neilgherrensis rhizome extract subjected to GC-MS analysis identified a total of 15 components; the major components were 1-heptatriacotanol (29.9%), 9 cis retinal (23.3%), aromadendrene oxide-(1) (17.6%), methyl (4,4-dimethyl-2,4,5,6-tetrahydro-1H-inden-2-yl)acetate (5.0%),(-)-β-caryophyllene epoxide (3.7%) andrographolide (3.1%) and 1,5-epoxy-salvial(4)14-en (2.4%) (Fig .5 O).

C. oligantha* var. *oligantha

The rhizome extract of *C. oligantha* var. *oilgantha* were subjected to GC-MS analysis and 26 components were identified retinal (34.1%), 3-

bromocholest-5-ene (21.2%), (-)-β-caryophyllene epoxide (8.0%), kauren-18-ol, acetate, (4.beta.)- (7.9%), 4-hexen-1-ol, 6-(2,6,6-trimethyl-1-cyclohexenyl)-4-methyl-, (E)-(3.2%), (+)-carotol (2.2%) and camphor (1.4%) are the major components. The obtained chromatogram represented in Fig. 5 P.

C. pseudomontana

Ten components were detected in GC-MS analysis of *C. pseudomontana* rhizome extracts; the major components are kauren-18-ol, acetate, (4.beta.)- (54.0%), 1-heptatriacotanol (17.4%), (-)-β-caryophyllene epoxide (10.9%) and 9 cis retinal (9.5%). The GC-MS profile is represented in Fig. 5 Q.

C. zedoaria

The rhizome extract of *C. zedoaria* produced 35 components such as, isovelleral (23.7%), germacrone (10.7%), isogermafurene (6.8%), aromadendrene oxide-(1) (6.5%), camphor (5.3%), 1-heptatriacotanol (4.3%), beta-elemene (3.6%), (-)-beta-caryophyllene (3.3%), germacrene B (2.6%) and 6.beta.-hydroxymethandienone (2.1%). The major components are represented in Fig. 5 R.

C. vamana

GC-MS analysis of *C. vamana* rhizome detected the following compounds as represented in the chromatogram (Fig 5S). Totally 26 components were identified; the major compounds are stahlianthusone (27.7%), copaene (15.5%), 9,10-dimethyl-1,2,3,4,5,6,7,8-octahydroanthracene (11.4%), Camphor (11.2%), aromadendrene (7.3%), coumarine, 8-allyl-7-hydroxy-6-ethyl-4-methyl (4.0%) and (-)-β-caryophyllene epoxide (2.9%).

C. zanthorrhiza

Thirty components were identified using GC-MS analysis from the rhizome extract of *C. zanthorrhiza* (Fig. 5T); the major components were 2-methyl-5-(1,2,2-trimethyl cyclopentyl)phenol (25.8%), beta.-curcumene (21.0%), alpha-curcumene (17.6%), N-(4-hydroxyphenyl)-N,N',N'-trimethylsulfamide (4.8%), germacrone (4.1%) and camphor (3.8%).

5.1.1 Cluster analysis

The major components were selected for the relationship analysis between the species. The components selected were above ≥5% concentration. The major and common components identified by GC–MS analyses are presented in Table 5.1. The fractions obtained from the rhizome extracts were scored for the presence or absences of specific components and the data were analyzed using NTSYS-pc version 2.1 software. Cluster analysis was carried out using UPGMA (Unweighted Pair Group Method with Arithmetic averages) for the compositions of 20 rhizome methanolic extracts to determine the pattern of identical and diverse chemical compounds between different species (Fig. 5.1). The data also revealed that camphor, germacrene D, aromadendrene oxide, β -caryophyllene epoxide, 1-heptatriacotanol, germacrone, retinal, Isovelleral, kauran-18-al, 17-(acetyloxy)-, (4.beta.) and β-caryophyllene were the common compounds in most of the species. The identified unique components from all the species are listed in Table 5.1.

Table 5.1: Major and Common Components identified from Twenty *Curcuma* species using GC-MS analysis

Components Name	Species Name Percentage quantity	Components	Species Name (Quantity)
β-caryophyllene	*C. decipiens* (3.8%) *C. aurantiaca* (6.5%) *C. inodora* (1.3%) *C. zedoaria* (1.1%)	1-heptatriacotanol	*C. coriacea* (39.9%) *C. aurantiaca* (1.0%) *C. neilgherrensis* (29.5%) *C. pseudomontana* (17.4%) *C. zedoaria* (4.3%)
β-caryophyllene epoxide	*C. amada* (75.7%) *C. bhatii* (87.0%) *C. coriacea* (12.3%) *C. decipiens* (14.4%) *C. aurantiaca* (1.4%) *C. harith* (26.6%) *C. karnatakensis*(8.3%) *C. kudagensis* (17.3%) *C. longa* (0.8%) *C. montana* (6.6%) *C. mutabilis* (17.2%) *C. neilgherrensis* (3.7%) *C. oligantha* (8.0%) *C. pseudomontana* (10.9%) *C. zedoaria* (0.4%) *C. vamana* (2.9%)	beta-bisabolene	*C. coriacea* (5.1%) *C. amada* (1.4%) *C. decipiens*(1.5%) *C. inodora* (7.7%) *C. karnatakensis*(0.3%) *C. longa* (5.5%) *C. oligantha* (1.6%)
β-myrcene	*C. amada* (9.4%)	kauren-18-ol, acetate, (4.beta.)-	*C. bhatii* (2.4%) *C. decipiens* (53.8%) *C. karnatakensis* (4.5%) *C. kudagensis* (38.3%) *C. mutabilis* (59.1%) *C. oilgantha* (7.9%) *C. pseudomontana* (54.0%) *C. vamana* (0.9%)
curdione	*C. aromatica* (20.5%) *C. zedoaria* (1.6%)	retinal	*C. bhatii* (1.8%) *C. decipiens* (11.4%) *C. karnatakensis* (11.7%) *C. kudagensis* (18.0%) *C. mutabilis* (18.1%) *C. oilgantha* (34.1%) *C. pseudomontana* (9.5%) *C. amada* (0.4%) *C. aromatica* (2.4%)

			C. coriacea (1.7%) *C. inodora* (11.3%) *C. neilgherrensis* (23.3%)
isovelleral	*C. aeruginosa* (5.4%) *C. aromatica* (15.3%) *C. haritha* (12.3%) *C. zedoaria* (23.7%) *C. zanthorrhiza* (0.7%)	juniper camphor	*C. decipiens* (0.6%) *C. inodora* (1.7%) *C. zedoaria* (0.7%)
andrographolide	*C. aromatica* (1.0%) *C. inodora* (5.7%) *C. neilgherrensis* (3.1%)	alloaromadendrene oxide-1	*C. aeruginosa* (0.6%) *C. amada* (0.9%) *C. coriacea* (0.5%) *C. decipiens* (1.1%) *C. inodora* (53.3%) *C. karnatakensis*(50.8%) *C. montana* (60.7%) *C. mutabilis* (1.3%) *C. neilgherrensis* (17.6%) *C. zedoaria* (6.5%) *C. vamana* (0.7%)
tumerone	*C. decipiens* (0.3%) *C. longa* (19.1%)	zingiberene	*C. zanthorrhiza* (1.8%) *C. longa* (11.6%) *C. decipiens* (0.2%)
(E)-.beta.-farnesene	*C. aeruginosa* (11.6%) *C. coriacea* (19.9%) *C. aurantiaca* (0.7%) *C. longa* (4.0%) *C. oilgantha* (0.4%) *C. zanthorrhiza* (2.9%)	camphor	*C. aeruginosa* (2.3%) *C. aromatica* (7.9%) *C. aurantiaca* (3.5%) *C. decipiens* (0.4%) *C. haritha* (4.5%) *C. inodora* (0.7%) *C. karnatakensis* (0.5%) *C. oilgantha* (1.4%) *C. zedoaria* (5.3%) *C. vamana* (11.2%) *C. zanthorrhiza* (3.8%)
germacrene B	*C. aeruginosa* (0.8%) *C. zedoaria* (2.6%) *C. zanthorrhiza* (2.2%)	beta-elemene	*C. aeruginosa* (2.2%) *C. aromatica* (2.2%) *C. zedoaria* (3.6%) *C. zanthorrhiza* (0.5%) *C. haritha* (0.6%)
beta-curcumene	*C. longa* (0.3%) *C. zanthorrhiza* (21.0%) *C. oligantha* (0.8%)	germacrene D	*C. aeruginosa* (1.6%) *C. aromatica* (0.6%) *C. zedoaria* (0.7%) *C. zanthorrhiza* (0.6%)
germacrone	*C. aeruginosa* (1.7%)	(+)-delta-cadinene	*C. aurantiaca* (1.2%)

	C. aromatica (3.1%) *C. zedoaria* (10.7%) *C. zanthorrhiza* (4.1%)		*C. aromatica* (0.8%)
Ar- tumerone	*C. decipiens* (0.4%) *C. longa* (6.1%)	eucalyptol	*C. aeruginosa* (1.8%) *C. coriacea* (3.3%) *C. aurantiaca* (8.3%) *C. longa* (0.5%)

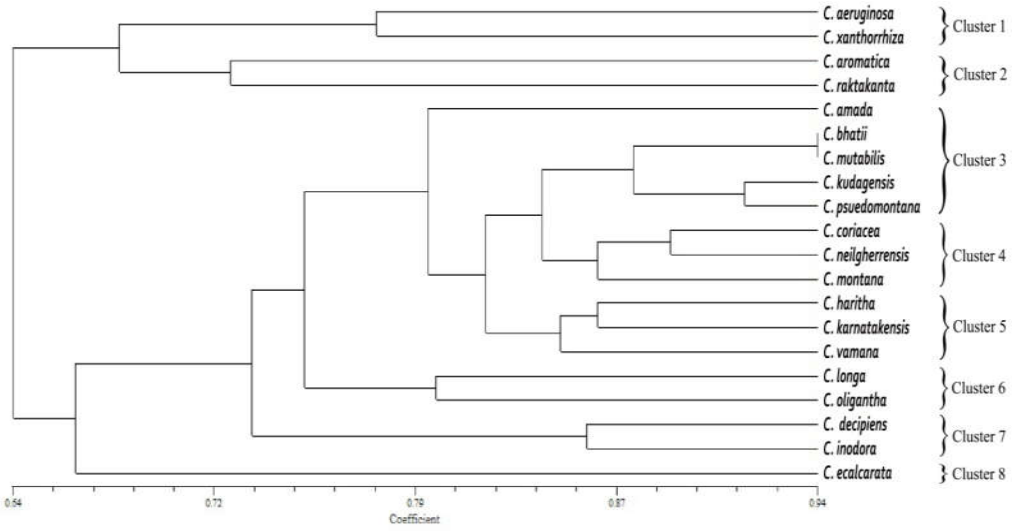
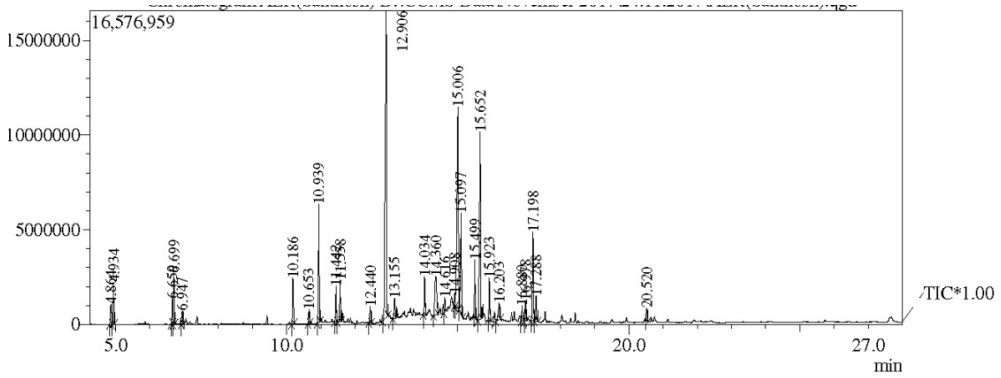


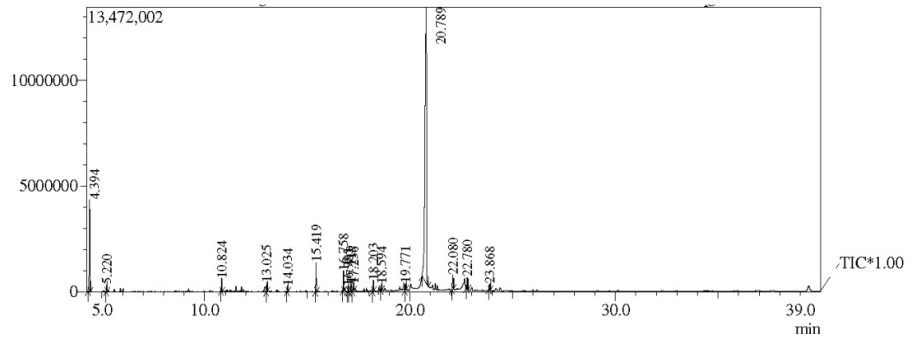
Fig. 5.1. UPGMA derived dendrogram of *Curcuma* species based on GC/MS analysis of methanolic rhizome extracts. The numerical scale indicates genetic similarity. Species were scored for the presence (1) or absence (0) of individual essential oil components with ≥5% concentration. Data analyzed using NTSYS – pc version 2.1 software and dendrogram generated based on the UPGMA cluster. The numerical scale indicates genetic similarity and the members of specific cluster/groups are genetically related with similar/near coefficient value



Peak Report TIC

Peak#	R.Time	Area	Area%	Height	Height%	Name	Base m/z
1	4.864	2192111	1.17	1040469	1.21	ISOPULEGOL 2	81.10
2	4.934	3541646	1.89	2210943	2.58	Eucalyptol	81.10
3	6.659	2360209	1.26	1327026	1.55	Spirobicyclo[2.2.1]heptane-2,2'-[1',3'-dioxa-2'-oxocyclohex-5'-ene)], 1,6',7,7-tetramethyl-	95.10
4	6.699	4346692	2.32	2666250	3.11	Camphor	95.10
5	6.947	1115576	0.60	670722	0.78	Isoborneol	95.10
6	10.186	4275327	2.29	2433676	2.84	BETA-ELEMEN-{2}	81.10
7	10.653	1315652	0.70	749282	0.87	Caryophyllene	93.10
8	10.939	10058682	5.38	6228293	7.26	(E)-.beta.-Famesene	69.05
9	11.442	3046362	1.63	1894433	2.21	(-)-GERMACRENE D	161.10
10	11.558	3544805	1.90	2220050	2.59	Isogermafurene	108.05
11	12.440	1536928	0.82	879068	1.03	GERMACRENE B	121.10
12	12.906	43001822	23.00	16238577	18.94	Phenol, 4-[[(dimethylamino)sulfonyl]methylamino]-	122.05
13	13.155	2585849	1.38	1039261	1.21	Acetic acid, (1,2,3,4,5,6,7,8-octahydro-3,8,8-trimethylnaphth-2-yl)methyl ester	105.10
14	14.034	3272356	1.75	2020200	2.36	(E,E)-Germacrone	107.10
15	14.360	8061399	4.31	2028864	2.37	(+)-ISOCURCUMENOL	105.10
16	14.616	1194549	0.64	779496	0.91	2,3,3-Trimethyl-2-(3-methylbuta-1,3-dienyl)-6-methylenecyclohexanone	163.05
17	14.908	1524400	0.82	673274	0.79	Spathulenol	162.10
18	15.006	30237191	16.17	10643234	12.41	Butyl 9,12-octadecadienoate	67.05
19	15.097	10237709	5.48	5323845	6.21	Isovelleral	232.10
20	15.499	4296385	2.30	3030660	3.53	ELEMENE	68.05
21	15.652	21874003	11.70	9909274	11.56	1,5,9-Cyclododecatriene, 1,5,9-trimethyl-	68.05
22	15.923	3776188	2.02	2261243	2.64	5,8-Dihydroxy-4a-methyl-4,4a,4b,5,6,7,8,8a,9,10-decahydro-2(3H)-phenanthrenone	174.10
23	16.203	2604482	1.39	924216	1.08	VALERENAL	228.05
24	16.880	1934368	1.03	814183	0.95	2-Isopropenyl-4a,8-dimethyl-1,2,3,4,4a,5,6,7-octahydronaphthalene	109.05
25	16.978	1918683	1.03	1075591	1.25	S-OCTAHYDROANTHRACENE	158.05
26	17.198	9131719	4.88	4687837	5.47	(-)-ISOLONGIFOLENE	175.05
27	17.288	2775674	1.48	1347219	1.57	3-Oxomanoyl oxide	81.10
28	20.520	1228920	0.66	616920	0.72	Alloaromadendrene oxide-(1)	81.10
		186989687	100.00	85734106	100.00		

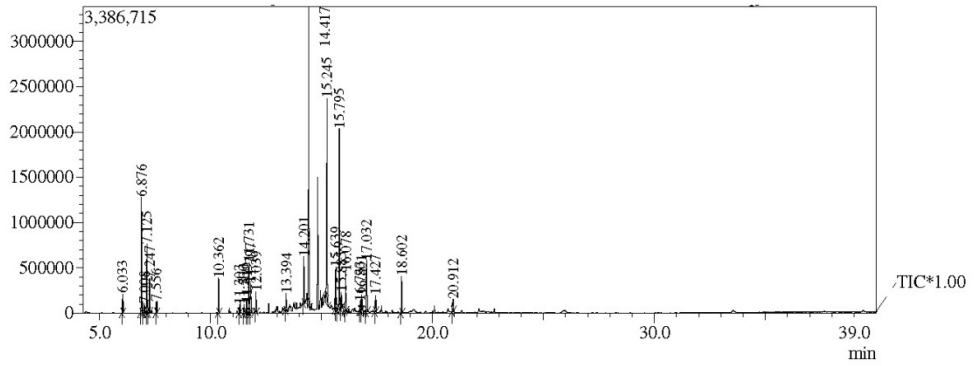
Fig. 5 A. GC-MS Chromatograms of methanolic extract of *C. aeruginosa* rhizome.



Peak Report TIC

Peak#	R.Time	Area	Area%	Height	Height%	Name	Base m/z
1	4.394	8389799	9.43	4331450	17.42	.beta.-Myrcene	93.10
2	5.220	536917	0.60	320673	1.29	.beta.-Ocimene	93.10
3	10.824	907640	1.02	606368	2.44	Caryophyllene	93.10
4	13.025	639764	0.72	423699	1.70	Phenol, 4-[[(dimethylamino)sulfonyl]methylamino]-	122.05
5	14.034	404088	0.45	273925	1.10	Spiro[4.5]dec-6-en-8-one, 1,7-dimethyl-4-(1-methylethyl)-	109.05
6	15.419	2369007	2.66	1388702	5.58	Acetic acid, 1-[2-(2,2,6-trimethyl-bicyclo[4.1.0]hept-1-yl)-ethyl]-vinyl ester	137.10
7	16.758	1343275	1.51	948761	3.82	(E,E,E)-.alpha.-Springene	69.10
8	16.985	429078	0.48	261132	1.05	.ALPHA.-BISABOLENE	137.10
9	17.119	510385	0.57	352098	1.42	BETA-BISABOLENE	69.10
10	17.236	653010	0.73	337559	1.36	4,8,13-Duvatriene-1,3-Diol	81.10
11	18.203	863571	0.97	496796	2.00	Alloaromadendrene oxide-(1)	81.10
12	18.594	438261	0.49	272422	1.10	RETINAL	147.10
13	19.771	575663	0.65	315004	1.27	4-Hexen-1-ol, 6-(2,6,6-trimethyl-1-cyclohexenyl)-4-methyl-, (E)-	137.10
14	20.789	67399474	75.72	13066225	52.54	(-)-.BETA.-CARYOPHYLLENE EPOXIDE	81.05
15	22.080	1650244	1.85	668142	2.69	Card-20(22)-enolide, 3,5,14,19-tetrahydroxy-, (3.beta.,5.beta.)-	55.00
16	22.780	1199860	1.35	532564	2.14	PREGNANE DIOL DIMETHYLSILYL ETHER	103.05
17	23.868	702221	0.79	273510	1.10	12,12-DICHLOROBICYCLO[8.2.0]DODECAN-11-ONE	98.05
		89012257	100.00	24869030	100.00		

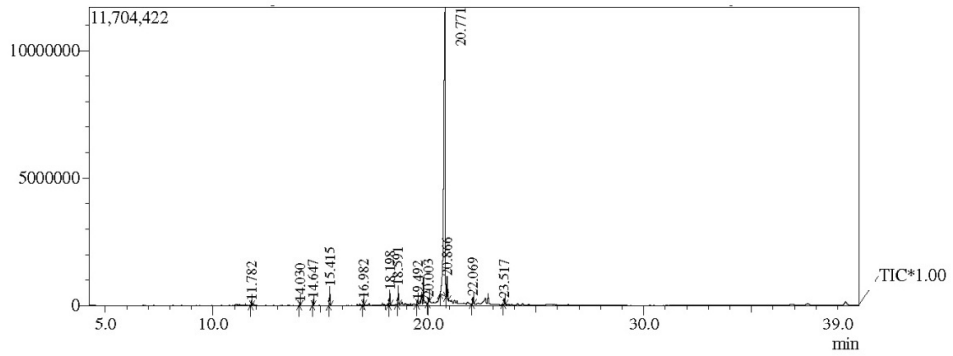
Fig. 5 B. GC-MS Chromatograms of *C. amada* rhizome methanolic extract.



Peak Report TIC

Peak#	R.Time	Area	Area%	Height	Height%	Name	Base m/z
1	6.033	303886	1.16	210266	1.36	.BETA.-LINALOOL	71.05
2	6.876	2092548	7.96	1277705	8.28	Camphor	95.10
3	7.008	131282	0.50	97068	0.63	EXO-METHYL-CAMPHENILOL	71.05
4	7.125	1217823	4.63	754243	4.89	Borneol, exo-	95.10
5	7.247	592875	2.26	369358	2.39	endo-Borneol	95.10
6	7.556	191168	0.73	124821	0.81	(+)-ALPHA-TERPINEOL (P-MENTH-1-EN-8-OL)	59.05
7	10.362	594723	2.26	385034	2.50	BETA-ELEMEN-(2)	93.10
8	11.303	145811	0.55	94525	0.61	ALPHA.-CARYOPHYLLENE	93.10
9	11.500	162357	0.62	92377	0.60	(-)-.ALPHA.-AMORPHENE	161.15
10	11.619	166952	0.64	112752	0.73	(-)-GERMACRENE D	161.10
11	11.731	884889	3.37	559053	3.62	Isogermafurene	108.05
12	11.797	695307	2.65	321702	2.09	PHENOL, 3,5-BIS(1,1-DIMETHYLETHYL)-	191.10
13	12.039	320967	1.22	219507	1.42	(+)-DELTA-CADINENE	161.10
14	13.394	280777	1.07	186329	1.21	(-)-SPATHULENOL	119.10
15	14.201	824947	3.14	520302	3.37	(E,E)-GERMACRONE	107.10
16	14.417	5406678	20.58	3270994	21.20	Curdione	69.05
17	15.245	4036619	15.36	2212631	14.34	Isovelleral	232.10
18	15.639	807509	3.07	464531	3.01	ELEMENE	68.05
19	15.795	3670364	13.97	1992454	12.91	TRIMETHYLCYCLODODECATRIENE	68.05
20	15.885	263570	1.00	181133	1.17	Andrographolide	105.10
21	16.078	732020	2.79	433141	2.81	5,8-Dihydroxy-4a-methyl-4,4a,4b,5,6,7,8,8a,9,10-decahydro-2(3H)-phenanthrenone	174.10
22	16.733	202239	0.77	93405	0.61	Diazoprogesterone	159.10
23	16.801	280218	1.07	172411	1.12	6R,9R-3-Oxo-.alpha.-ionol	108.10
24	17.032	1000819	3.81	563020	3.65	(3E,5E,7E)-6-Methyl-8-(2,6,6-trimethyl-1-cyclohexenyl)-3,5,7-octatrien-2-one	109.05
25	17.427	323925	1.23	182316	1.18	2-Butenal, 2-methyl-4-(2,6,6-trimethyl-1-cyclohexen-1-yl)-	81.10
26	18.602	641108	2.44	404933	2.62	Retinal, 9-cis-	147.10
27	20.912	305994	1.16	132562	0.86	Cholest-6-ene-3.beta.,5.beta.,8.beta.-triol	125.10
		26277375	100.00	15428573	100.00		

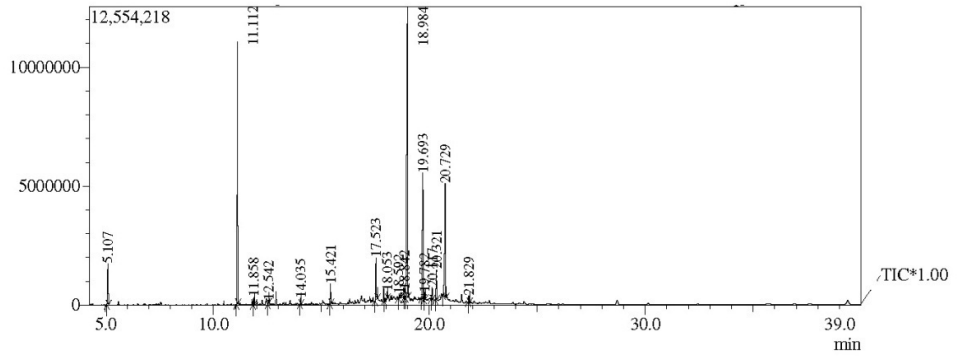
Fig. 5 C. GC-MS Chromatograms of *C. aromatica* rhizome.



Peak Report TIC

Peak#	R.Time	Area	Area%	Height	Height%	Name	Base m/z
1	11.782	271222	0.43	173709	1.10	PHENOL, 3,5-BIS(1,1-DIMETHYLETHYL)-	191.10
2	14.030	187105	0.30	121277	0.77	Spiro[4.5]dec-6-en-8-one, 1,7-dimethyl-4-(1-methylethyl)-	109.10
3	14.647	415365	0.66	252464	1.59	CUPAROPHENOL	136.10
4	15.415	1225458	1.93	746063	4.71	Acetic acid, 1-[2-(2,2,6-trimethyl-bicyclo[4.1.0]hept-1-yl)-ethyl]-vinyl ester	137.10
5	16.982	380165	0.60	235782	1.49	4-Hexen-1-ol, 6-(2,6,6-trimethyl-1-cyclohexenyl)-4-methyl-, (E)-	137.10
6	18.198	1000471	1.58	592071	3.74	DUVATRIENDIOL	81.05
7	18.591	1191858	1.88	729917	4.61	RETINAL	147.05
8	19.492	309249	0.49	154872	0.98	Methyl octadec-6,9-dien-12-ynoate	149.05
9	20.003	432944	0.68	173570	1.10	2,4a,8,8-Tetramethyldecahydrocyclopropa[d]naphthalene	97.00
10	20.771	55173325	87.06	11334176	71.57	(-)-.BETA.-CARYOPHYLLENE EPOXIDE	81.05
11	20.866	1571381	2.48	840863	5.31	Kauren-18-ol, acetate, (4.beta.)-	81.05
12	22.069	657114	1.04	270816	1.71	KAUREN-19-YL-ACETATE	55.00
13	23.517	555664	0.88	211426	1.34	Verrucarol	81.05
		63371321	100.00	15837006	100.00		

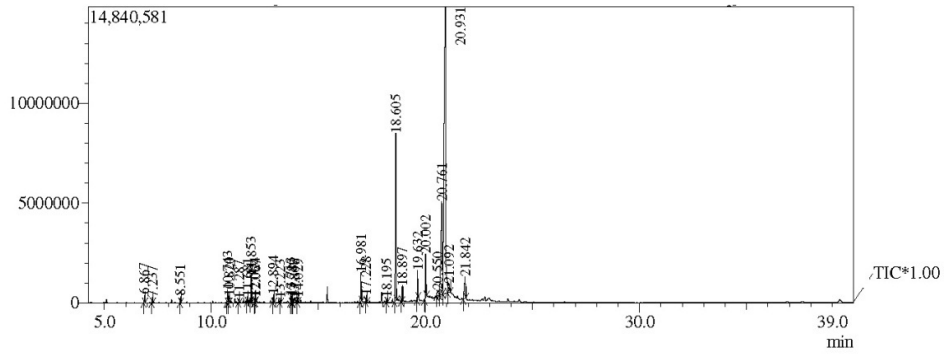
Fig. 5 D. GC-MS Chromatograms of *C. bhatii*.



Peak Report TIC

Peak#	R.Time	Area	Area%	Height	Height%	Name	Base m/z
1	5.107	3334669	3.31	1729915	4.17	Eucalyptol	81.10
2	11.112	20113971	19.97	11022691	26.58	(E)-.beta.-Famesene	69.05
3	11.858	512941	0.51	313190	0.76	.beta.-Bisabolene	69.05
4	12.542	173095	0.17	138053	0.33	LIMONENE DIOXIDE 2	79.05
5	14.035	489046	0.49	261148	0.63	Spiro[4.5]dec-6-en-8-one, 1,7-dimethyl-4-(1-methylethyl)-	109.05
6	15.421	1510622	1.50	836832	2.02	Acetic acid, 1-[2-(2,2,6-trimethyl-bicyclo[4.1.0]hept-1-yl)-ethyl]-vinyl ester	137.10
7	17.523	3387717	3.36	1792411	4.32	Corymbolone	121.10
8	18.053	429636	0.43	302686	0.73	Succinic acid, dodec-9-yn-1-yl nonyl ester	101.05
9	18.592	1758450	1.75	175359	0.42	RETINAL	147.10
10	18.842	1197671	1.19	321136	0.77	LIMONENE DIOXIDE 3	107.10
11	18.984	40275923	39.98	12234162	29.50	1-Heptatriacotanol	69.05
12	19.693	10688636	10.61	5285027	12.74	(1S,2E,4S,5R,7E,11E)-Cembra-2,7,11-trien-4,5-diol	95.05
13	19.782	534572	0.53	289329	0.70	ALLOAROMADENDRENOXID-(1)	137.10
14	20.117	791962	0.79	488089	1.18	Calusterone	135.10
15	20.321	2597854	2.58	1262239	3.04	Tricyclo[20.8.0.0(7,16)]triacontane, 1(22),7(16)-diepoxy-	95.05
16	20.729	12407709	12.32	4766904	11.49	(-)-.BETA.-CARYOPHYLLENE EPOXIDE	81.10
17	21.829	532207	0.53	254381	0.61	Pregnane-3,11,20-triol, (3.alpha., 11.beta., 20.beta.)-	93.10
		100736681	100.00	41473552	100.00		

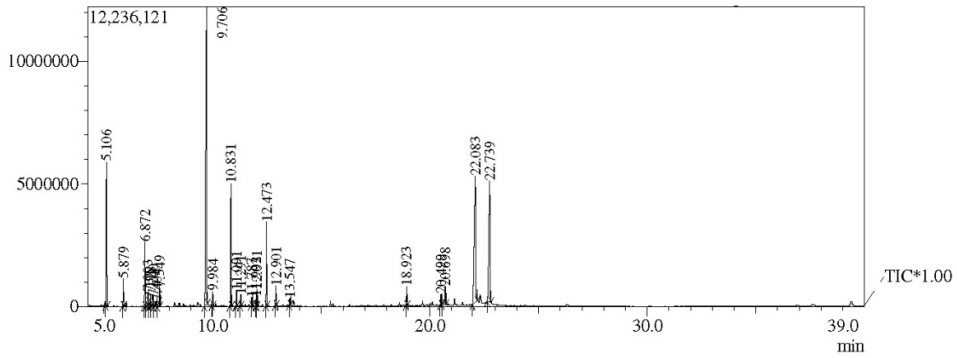
Fig. 5 E. *C. coriacea* GC-MS Chromatograms profile.



Peak Report TIC

Peak#	R.Time	Area	Area%	Height	Height%	Name	Base m/z
1	6.867	619546	0.43	390894	0.94	Camphor	95.10
2	7.237	447514	0.31	255490	0.62	endo-Borneol	95.10
3	8.551	469572	0.33	313550	0.76	.alpha.-Citral	69.10
4	10.743	982582	0.68	671848	1.62	(-)-.alpha.-Santalene	94.10
5	10.820	546599	0.38	351943	0.85	(-)-BETA-CARYOPHYLLEN	93.10
6	11.287	311055	0.22	188713	0.45	.ALPHA.-CARYOPHYLLENE	93.10
7	11.681	353611	0.24	266457	0.64	(-)-ZINGIBERENE	119.15
8	11.853	2246060	1.55	1322455	3.19	.beta.-Bisabolene	69.10
9	12.024	396532	0.27	265687	0.64	(+-)-LEPIDOZENE	121.15
10	12.067	294876	0.20	204715	0.49	.BETA.-SESQUIPHELLANDRENE	69.10
11	12.894	703629	0.49	389704	0.94	cis-sesquisabinene hydrate	83.05
12	13.223	390210	0.27	230546	0.56	alpha.-Humulene epoxide II	67.05
13	13.713	578335	0.40	315841	0.76	Ar-tumerone	83.10
14	13.780	482662	0.33	274841	0.66	Tumerone	83.05
15	13.867	913995	0.63	200925	0.48	JUNIPERCAMPHOR	81.10
16	14.029	434935	0.30	248936	0.60	Spiro[4.5]dec-6-en-8-one, 1,7-dimethyl-4-(1-methylethyl)-	109.10
17	16.981	2612363	1.81	1465037	3.53	4-Hexen-1-ol, 6-(2,6,6-trimethyl-1-cyclohexenyl)-4-methyl-, (E)-	137.15
18	17.228	576056	0.40	320421	0.77	Dehydrosaussurea lactone	81.10
19	18.195	381353	0.26	216933	0.52	4,8,13-DUVATRIENE-1,3-DIOL	81.05
20	18.605	16539144	11.45	8417698	20.30	Retinal, 9-cis-	147.10
21	18.897	1383750	0.96	747055	1.80	1-Ethyl-4,4-dimethyl-cyclohex-2-en-1-ol	125.10
22	19.632	2498936	1.73	1411178	3.40	Acetic acid, 1-[2-(2,2,6-trimethyl-bicyclo[4.1.0]hept-1-yl)-ethyl]-vinyl ester	97.05
23	20.002	4411786	3.05	2210855	5.33	TRISPIRO[4.2.4.2.4.2.]HENEICOSANE	97.05
24	20.550	1621950	1.12	274268	0.66	2-Hexynyl aldehyde diethyl acetal	125.10
25	20.761	20915669	14.48	4657655	11.23	(-)-.BETA.-CARYOPHYLLENE EPOXIDE	137.15
26	20.931	77857219	53.89	14364777	34.63	Kauren-18-ol, acetate, (4.beta.)-	81.10
27	21.092	1713679	1.19	423164	1.02	ALLOAROMADENDRENOXID-(1)	81.10
28	21.842	3794199	2.63	1073923	2.59	KAUREN-19-YL-ACETATE	69.05
		144477817	100.00	41475509	100.00		

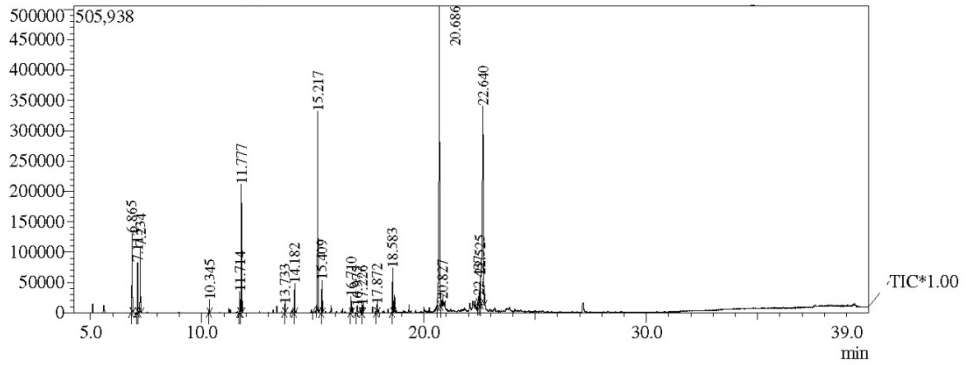
Fig. 5 F. GC-MS Chromatograms of *C. decipiens* rhizome.



Peak Report TIC

Peak#	R.Time	Area	Area%	Height	Height%	Name	Base m/z
1	5.106	10469878	8.36	5880108	12.08	Eucalyptol	81.10
2	5.879	1938010	1.55	1136309	2.33	.ALPHA.-TERPINOLEN	93.10
3	6.872	4496346	3.59	2626635	5.39	Camphor	95.10
4	7.003	1051853	0.84	586407	1.20	EXO-METHYL-CAMPHENILOL	71.05
5	7.118	521155	0.42	315028	0.65	Isoborneol	95.10
6	7.240	796484	0.64	417374	0.86	endo-Borneol	95.10
7	7.404	748419	0.60	207129	0.43	CYMEN-8-OL, P-	135.10
8	7.549	1130681	0.90	757243	1.56	(+)-ALPHA-TERPINEOL (P-MENTH-1-EN-8-OL)	59.05
9	9.706	38512386	30.75	12166563	24.98	3-Terpinolenone	150.10
10	9.984	872965	0.70	572243	1.18	5,7-DIMETHYLOCTAHYDROCOUMARIN 1	95.05
11	10.831	8181166	6.53	5003939	10.28	TRANS(.BETA.)-CARYOPHYLLENE	93.10
12	11.091	926471	0.74	630889	1.30	(E)-.beta.-Famesene	69.10
13	11.291	729509	0.58	462691	0.95	.ALPHA.-CARYOPHYLLENE	93.10
14	11.787	531505	0.42	334173	0.69	PHENOL, 3,5-BIS(1,1-DIMETHYLETHYL)-	191.10
15	11.992	565652	0.45	357447	0.73	.GAMMA.-MUUROLEN	161.15
16	12.031	1000312	0.80	676097	1.39	(+)-DELTA-CADINENE	161.15
17	12.473	5244467	4.19	3425541	7.03	E-Nerolidol	69.05
18	12.901	1312423	1.05	801676	1.65	Isoaromadendrene epoxide	79.05
19	13.547	507269	0.41	288734	0.59	.tau.-Cadinol	161.15
20	18.923	1300560	1.04	784788	1.61	1-Heptatriacotanol	83.10
21	20.499	830928	0.66	439021	0.90	3-OXO-5-PHENYL-PENTANOIC ACID	91.05
22	20.698	1837492	1.47	733043	1.51	(-)-.BETA.-CARYOPHYLLENE EPOXIDE	81.10
23	22.083	22943053	18.32	5156384	10.59	Dihydrochrysin	256.10
24	22.739	18782886	15.00	4936099	10.14	3,3-Dimethyl-4-phenyl-4-penten-2-one	131.10
		125231870	100.00	48695561	100.00		

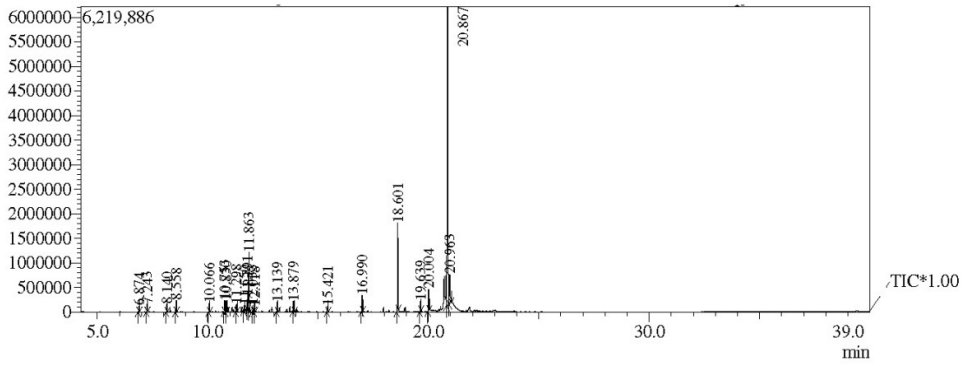
Fig. 5 G. *C. aurantiaca* GC-MS Chromatograms.



Peak Report TIC

Peak#	R.Time	Area	Area%	Height	Height%	Name	Base m/z
1	6.865	199872	4.55	129934	6.31	(-)-CAMPHOR	95.10
2	7.113	122643	2.79	82720	4.02	2-CAMPHANOL, EXO-	95.10
3	7.234	157334	3.58	108476	5.27	2-BORNANOL	95.10
4	10.345	27021	0.62	21415	1.04	BETA. ELEMENE	81.05
5	11.714	49179	1.12	34919	1.70	Isosericenin	108.10
6	11.777	321366	7.31	212023	10.30	2,4-DITERT-BUTYLPHENOL	191.15
7	13.733	15087	0.34	13901	0.68	ISONARDOSINON	59.05
8	14.182	67413	1.53	48450	2.35	Germacron	107.10
9	15.217	541399	12.32	326323	15.86	Isovelleral	232.15
10	15.409	81134	1.85	53401	2.60	LABDA-8(17),13Z-DIEN-15-OL	137.10
11	16.710	44915	1.02	26612	1.29	6-HEPTENSAEURE, 4-ISOPROPENYL-6-METHYL-2-OXO-, ETHYLESTER	55.10
12	16.975	38230	0.87	10167	0.49	(E)-3-METHYL-5-(2',6',6'-TRIMETHYLCYCLOHEX-1'-ENYL)PENT-2-ENYL BROMIDE	137.20
13	17.226	15831	0.36	12009	0.58	(Z)-4-METHYL-6-(2',6',6'-TRIMETHYL-2'-CYCLOHEXENYL)HEX-3-EN-1-OL	81.10
14	17.872	21879	0.50	13550	0.66	2-(2-Hydroxyethylamino)pyrimidine	108.05
15	18.583	107065	2.44	73865	3.59	Diazoprogesterone	147.10
16	20.686	1170777	26.65	494222	24.02	(-)-.BETA.-CARYOPHYLLENE EPOXIDE	81.10
17	20.827	29389	0.67	9582	0.47	1-Oxaspiro[2.5]octane, 4,4-dimethyl-8-methylene-	137.20
18	22.437	23559	0.54	11985	0.58	trans-2-Caren-4-ol	109.10
19	22.525	136857	3.12	46964	2.28	DRIMENOL	109.10
20	22.640	1222298	27.82	327241	15.90	Isoaromadendrene epoxide	55.05
		4393248	100.00	2057759	100.00		

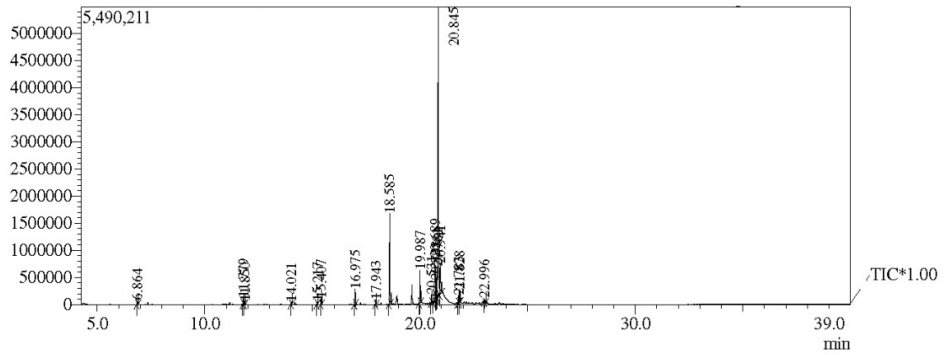
Fig. 5 H. GC-MS Chromatograms of *C. haritha* rhizome.



Peak Report TIC

Peak#	R.Time	Area	Area%	Height	Height%	Name	Base m/z
1	6.874	187264	0.71	120312	0.91	Camphor	95.10
2	7.243	221875	0.85	153652	1.16	endo-Borneol	95.10
3	8.140	239801	0.91	162012	1.22	.BETA.-CITRAL	69.10
4	8.558	345969	1.32	231544	1.74	.alpha.-Citral	69.05
5	10.066	249569	0.95	178632	1.34	NERYL ACETATE	69.10
6	10.753	367542	1.40	239628	1.80	(-)-ALPHA-SANTALENE	94.05
7	10.830	364339	1.39	230351	1.73	(-)-.BETA.-CARYOPHYLLENE	93.10
8	11.298	245429	0.94	156926	1.18	.ALPHA.-CARYOPHYLLENE	93.10
9	11.650	157962	0.60	123132	0.93	.ALPHA.-SELINENE	189.15
10	11.791	344540	1.31	279039	2.10	PHENOL, 2,4-BIS(1,1-DIMETHYLETHYL)-	191.10
11	11.863	2026833	7.72	1179893	8.88	(+,-)-.BETA.-BISABOLENE	69.10
12	12.033	209388	0.80	127442	0.96	.DELTA.-CADINENE	161.10
13	12.118	156140	0.59	111903	0.84	(-)-.alpha.-Panasinsen	122.10
14	13.139	302385	1.15	203067	1.53	Carotol	161.15
15	13.879	451997	1.72	211268	1.59	Juniper camphor	81.05
16	15.421	189497	0.72	117899	0.89	Acetic acid, 1-[2-(2,2,6-trimethyl-bicyclo[4.1.0]hept-1-yl)-ethyl]-vinyl ester	137.10
17	16.990	574545	2.19	339640	2.56	4-Hexen-1-ol, 6-(2,6,6-trimethyl-1-cyclohexenyl)-4-methyl-, (E)-	137.10
18	18.601	2980764	11.36	1794425	13.51	Retinal, 9-cis-	147.10
19	19.639	361977	1.38	210640	1.59	Longipinocarveol, trans-	97.00
20	20.004	770035	2.93	440092	3.31	2,4a,8,8-Tetramethyldecahydrocyclopropa[d]naphthalene	97.00
21	20.867	14003209	53.35	6074768	45.74	ALLOAROMADENDRENOXID-(1)	81.10
22	20.963	1494862	5.70	595078	4.48	Andrographolide	147.10
		26245922	100.00	13281343	100.00		

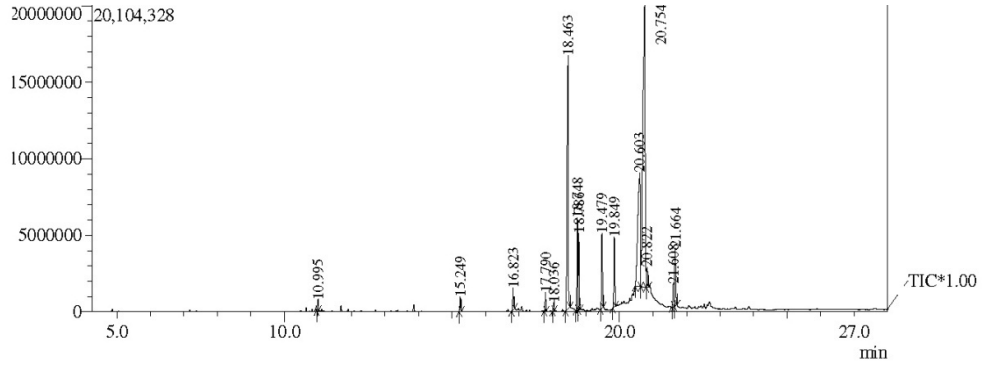
Fig. 5 I. GC-MS Chromatograms of *C. inodora* rhizome methanolic extract.



Peak Report TIC

Peak#	R.Time	Area	Area%	Height	Height%	Name	Base m/z
1	6.864	121384	0.51	77724	0.69	Camphor	95.10
2	11.779	200304	0.84	129176	1.14	2,4-DITERT-BUTYLPHENOL	191.20
3	11.850	83617	0.35	58277	0.52	(-)-.BETA.-BISABOLENE	69.10
4	14.021	81575	0.34	47546	0.42	Isocyclocitral	109.10
5	15.217	62206	0.26	51880	0.46	4,8,11,11-TETRAMETHYL-8-TRICYCLO[7.2.0.0(2,5).]UNDECEN-4-OL	232.15
6	15.407	180681	0.76	115763	1.03	3-Cyclohexene-1-carboxaldehyde, 1,3,4-trimethyl-	137.15
7	16.975	471271	1.97	284303	2.52	4-Hexen-1-ol, 6-(2,6,6-trimethyl-1-cyclohexenyl)-4-methyl-, (E)-	137.20
8	17.943	129017	0.54	86859	0.77	Diazoprogesterone	146.10
9	18.585	2818918	11.79	1674674	14.84	RETINAL	147.10
10	19.987	1076803	4.50	606823	5.38	Kauren-18-ol, acetate, (4.beta.)-	81.10
11	20.531	249057	1.04	145408	1.29	Bicyclo[3.1.1]heptan-3-one, 6,6-dimethyl-2-(2-methylpropyl)-	125.10
12	20.689	2007735	8.39	774748	6.87	((-)-.BETA.-CARYOPHYLLENE EPOXIDE	81.10
13	20.738	720260	3.01	334448	2.96	1-Ethyl-4,4-dimethyl-cyclohex-2-en-1-ol	125.15
14	20.781	1092847	4.57	567923	5.03	10-Chlorotricyclo[4.2.1.1(2,5)]deca-3,7-dien-9-ol	147.10
15	20.845	12157291	50.83	5341322	47.33	ALLOAROMADENDRENOXID-(1)	81.10
16	20.941	1469143	6.14	550764	4.88	3-BROMOCHOLEST-5-ENE #	147.10
17	21.783	368783	1.54	140226	1.24	Longipinocarveol, trans-	69.05
18	21.828	480940	2.01	232906	2.06	Ergost-25-ene-6,12-dione, 3,5-dihydroxy-, (3.beta.,5.alpha.)-	69.10
19	22.996	147524	0.62	64159	0.57	DUVATRIENDIOL	81.10
		23919356	100.00	11284929	100.00		

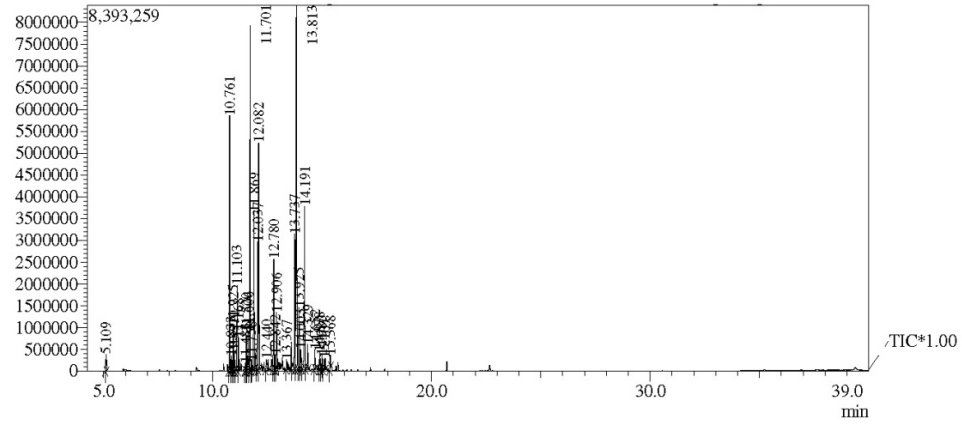
Fig. 5 J. GC-MS profile of *C. karnatakensis* rhizome.



Peak Report TIC

Peak#	R.Time	Area	Area%	Height	Height%	Name	Base m/z
1	10.995	1037748	0.45	694474	0.94	.BETA.-SESQUIPHELLANDRENE	69.05
2	15.249	1623339	0.71	981621	1.32	Acetic acid, 1-[2-(2,2,6-trimethyl-bicyclo[4.1.0]hept-1-yl)-ethyl]-vinyl ester	137.10
3	16.823	2653153	1.16	1430187	1.93	4-Hexen-1-ol, 6-(2,6,6-trimethyl-1-cyclohexenyl)-4-methyl-, (E)-	137.10
4	17.790	1935255	0.85	1195435	1.61	2(1H)-Naphthalenone, 7-ethynyl-4a,5,6,7,8,8a-hexahydro-1,4a-dimethyl-, (1.alpha.,4a.beta.,7.beta.,8a.alpha.)-	146.05
5	18.036	866597	0.38	509854	0.69	SPINACENE	81.10
6	18.463	41349349	18.08	16545444	22.29	Retinal, 9-cis-	147.10
7	18.748	10906865	4.77	6014391	8.10	1-Ethyl-4,4-dimethyl-cyclohex-2-en-1-ol	125.10
8	18.786	7827533	3.42	5005552	6.74	2-(4-NITROBUTANOYL)CYCLOHEXANONE	125.10
9	19.479	9017118	3.94	4868294	6.56	Acetic acid, 1-[2-(2,2,6-trimethyl-bicyclo[4.1.0]hept-1-yl)-ethyl]-vinyl ester	97.05
10	19.849	8573085	3.75	4559743	6.14	Trispiro[4.2.4.2.4.2.]heneicosane	97.05
11	20.603	39589876	17.31	7411793	9.98	(-)-.BETA.-CARYOPHYLLENE EPOXIDE	81.10
12	20.754	87715902	38.35	18473138	24.88	Kauren-18-ol, acetate, (4.beta.)-	81.10
13	20.822	3021396	1.32	1289976	1.74	22,23-Dibromostigmasterol acetate	147.10
14	21.608	4176592	1.83	1493779	2.01	2,6,10,14,18,22-TETRACOSAHEXAENE, 2,6,10,15,19,23-HEXAMETHYL-	69.05
15	21.664	8416158	3.68	3766726	5.07	Ergost-25-ene-6,12-dione, 3,5-dihydroxy-, (3.beta.,5.alpha.)-	69.05
		228709966	100.00	74240407	100.00		

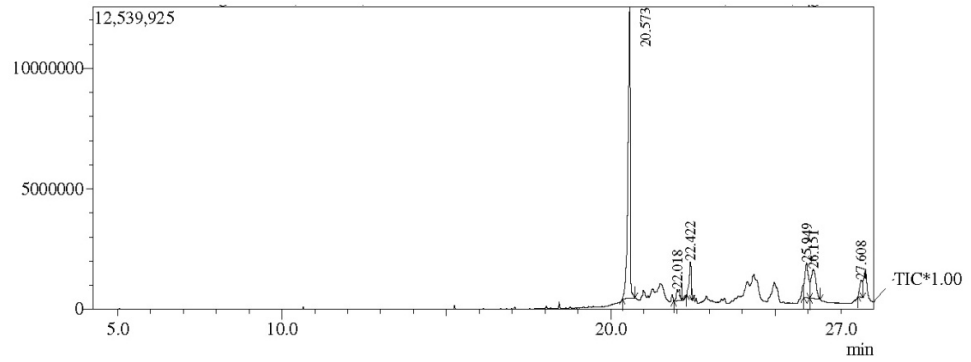
Fig. 5 K. *C. kudagensis* rhizome GC-MS Chromatograms.



Peak Report TIC

Peak#	R.Time	Area	Area%	Height	Height%	Name	Base m/z
1	5.109	621760	0.57	390552	0.69	Eucalyptol	81.05
2	10.761	9421022	8.67	5851231	10.29	(-)-.alpha.-Santalene	94.10
3	10.833	559563	0.51	350975	0.62	Caryophyllene	93.10
4	10.925	1604064	1.48	1058799	1.86	trans-.alpha.-Bergamotene	119.10
5	10.971	769487	0.71	477749	0.84	Teresantalol	121.15
6	11.103	4368702	4.02	1973416	3.47	(E)-beta.-Famesene	69.05
7	11.168	1084419	1.00	760629	1.34	TRANS-BETA.-FARNESENE	69.10
8	11.483	357823	0.33	188529	0.33	.GAMMA.-CURCUMENE	121.15
9	11.520	1344829	1.24	889450	1.56	ALPHA-CURCUMEN	132.15
10	11.606	1421935	1.31	916286	1.61	cis-.beta.-Farnesene	69.10
11	11.701	12666973	11.66	7914967	13.92	(-)-ZINGIBERENE	119.10
12	11.758	391248	0.36	222471	0.39	ALPHA.-BISABOLENE	93.10
13	11.869	6056464	5.57	3630198	6.38	(+,-)-.BETA.-BISABOLENE	69.10
14	12.037	5164575	4.75	2928844	5.15	Benzenemethanol, 4-methyl-.alpha.-(1-methyl-2-propenyl)-, (R*,R*)-	121.15
15	12.082	9624192	8.86	5197383	9.14	.BETA.-SESQUIPHELLANDRENE	69.05
16	12.440	890214	0.82	251307	0.44	(-)-.BETA.-CARYOPHYLLENE EPOXIDE	93.10
17	12.780	4633125	4.26	2537889	4.46	6-Isopropyl-1,2-dimethyl-4-oxo-bicyclo[3.3.1]non-2-ene-9-carboxaldehyde	83.10
18	12.842	560450	0.52	351020	0.62	.alpha.-Santalol	85.10
19	12.906	2850311	2.62	1325435	2.33	Butane-1,1-dicarbonitrile, 1-cyclohexyl-3-methyl-	83.05
20	13.367	736030	0.68	223277	0.39	7-epi-cis-sesquisabinene hydrate	69.05
21	13.737	6728366	6.19	3102169	5.46	Ar-tumerone	83.10
22	13.813	20813727	19.15	8345766	14.68	Tumerone	83.05
23	13.925	3374144	3.11	1416464	2.49	Bicyclo[3.3.1]nonan-9-one, 1,2,4-trimethyl-3-nitro-, (2-endo,3-exo,4-exo)-(.+-.)-	95.10
24	14.003	1291864	1.19	465447	0.82	7-epi-trans-sesquisabinene hydrate	93.10
25	14.191	6974244	6.42	3739480	6.58	Curlone	120.10
26	14.329	1257183	1.16	573416	1.01	ALPHA. TERPINEOL	59.05
27	14.652	762273	0.70	461148	0.81	Naphthalene, 1,1'-methylenebis[decahydro-	137.10
28	14.878	768198	0.71	414534	0.73	Cyclohexane, (2-nitro-2-propenyl)-	83.05
29	14.980	502220	0.46	333602	0.59	1H-Indene, 2,3,3a,4,7,7a-hexahydro-2,2,4,4,7,7-hexamethyl-	83.10
30	15.107	460987	0.42	266547	0.47	Cyclohexene, 1-methyl-3-(1-methylethyl)-	95.10
31	15.368	606294	0.56	304830	0.54	Dehydrozingerone	145.00
		108666686	100.00	56863810	100.00		

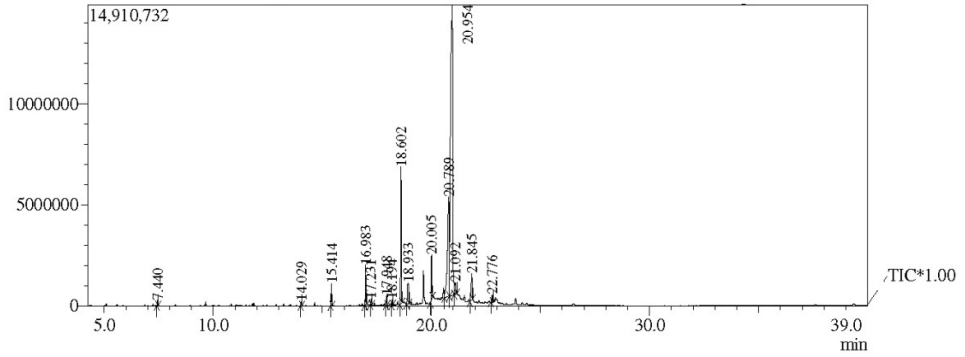
Fig. 5 L. GC-MS Chromatograms of methanolic extract form *C. longa* rhizome.



Peak Report TIC

Peak#	R.Time	Area	Area%	Height	Height%	Name	Base m/z
1	20.573	58415947	60.78	12078649	69.29	Alloaromadendrene oxide-(1)	81.10
2	22.018	3410457	3.55	475836	2.73	ANDROSTAN-3-ONE, 17-HYDROXY-, (5.ALPHA.,17.BETA.)-	181.10
3	22.422	6367682	6.63	1521445	8.73	(-)-.BETA.-CARYOPHYLLENE EPOXIDE	55.10
4	25.949	11035593	11.48	1445468	8.29	VIOLANTHRENE A	426.30
5	26.151	12351065	12.85	1222715	7.01	2-tert-Butyl-4,6-bis(3,5-di-tert-butyl-4-hydroxybenzyl)phenol	57.10
6	27.608	4523502	4.71	688902	3.95	o[3,5,6-metheno-1-silapentalene-1(2H),1'-[1,3]disilacyclobutane-3',1''(2''H)-[3,5,6]metheno[1]silapentalene], 2,2	57.10
		96104246	100.00	17433015	100.00		

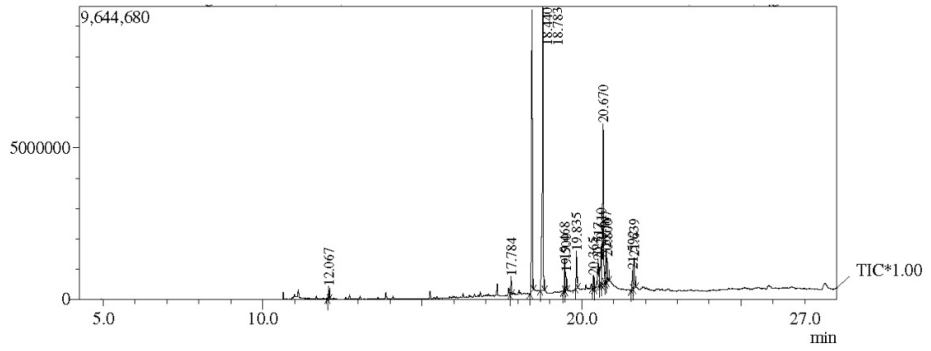
Fig. 5 M. GC-MS Chromatograms profile of *C. montana* rhizome.



Peak Report TIC

Peak#	R.Time	Area	Area%	Height	Height%	Name	Base m/z
1	7.440	441799	0.28	267600	0.74	Cryptone, L-	96.05
2	14.029	383029	0.24	219812	0.61	Spiro[4.5]dec-6-en-8-one, 1,7-dimethyl-4-(1-methylethyl)-	109.10
3	15.414	1912235	1.20	1107116	3.05	Acetic acid, 1-[2-(2,2,6-trimethyl-bicyclo[4.1.0]hept-1-yl)-ethyl]-vinyl ester	137.15
4	16.983	3431272	2.15	1977843	5.45	4-Hexen-1-ol, 6-(2,6,6-trimethyl-1-cyclohexenyl)-4-methyl-, (E)-	137.15
5	17.231	600959	0.38	304555	0.84	4,8,13-Duvatriene-1,3-Diol	81.10
6	17.948	750613	0.47	458026	1.26	7-Ethynyl-1,4a-dimethyl-4a,5,6,7,8,8a-hexahydro-2(1H)-naphthalenone	146.10
7	18.194	485945	0.30	279965	0.77	Alloaromadendrene oxide-(1)	81.10
8	18.602	13015867	8.14	6789279	18.71	RETINAL	147.10
9	18.933	3969350	2.48	1057661	2.92	1-Ethyl-4,4-dimethyl-cyclohex-2-en-1-ol	125.10
10	20.005	4840360	3.03	2259406	6.23	TRISPIRO[4.2.4.2.4.2]HENEICOSANE	97.05
11	20.789	27525402	17.21	4924944	13.57	(-)-.BETA.-CARYOPHYLLENE EPOXIDE	137.15
12	20.954	94639145	59.18	14385986	39.65	Kauren-18-ol, acetate, (4.beta.)-	81.10
13	21.092	2199794	1.38	576686	1.59	Aromadendrene oxide-(2)	81.05
14	21.845	4834286	3.02	1322574	3.65	KAUREN-19-YL-ACETATE	69.05
15	22.776	900154	0.56	350176	0.97	Diazoprogesterone	137.15
		159930210	100.00	36281629	100.00		

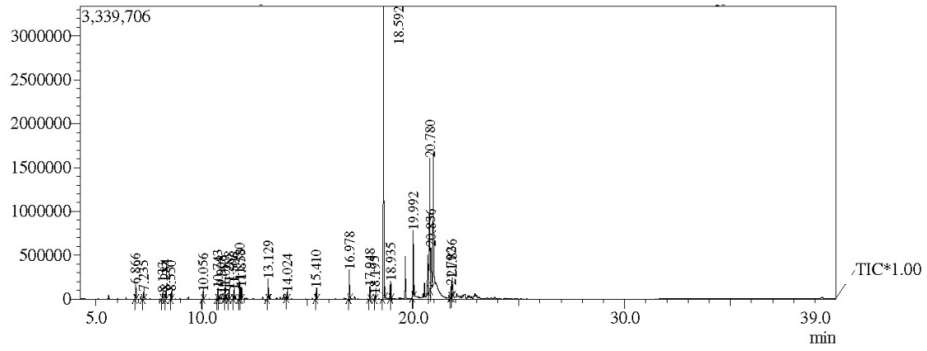
Fig. 5 N. GC-MS Chromatograms of methanolic rhizome extract of *C. mutabilis*.



Peak Report TIC

Peak#	R.Time	Area	Area%	Height	Height%	Name	Base m/z
1	12.067	636320	0.89	417464	1.23	1,3,3-TRIMETHYL-2-OXABICYCLO[2.2.2]OCT-6-YL ACETATE	108.10
2	17.784	915116	1.29	600768	1.77	.BETA.-COPAEN-4 .ALPHA.-OL	146.05
3	18.440	16631989	23.36	9273132	27.36	Retinal, 9-cis-	147.10
4	18.783	21064869	29.59	9345645	27.57	1-Heptatriacotanol	125.10
5	19.468	1774568	2.49	1016150	3.00	1,5-EPOXYSALVIAL-4(14)-ENE	97.10
6	19.500	1028074	1.44	613506	1.81	2,6,10,10-Tetramethyl-1-oxaspiro[4.5]decan-6-ol	85.15
7	19.835	2101183	2.95	1221329	3.60	1,5-EPOXYSALVIAL-4(14)-ENE	97.10
8	20.365	688739	0.97	416708	1.23	1-Ethyl-4,4-dimethyl-cyclohex-2-en-1-ol	125.10
9	20.517	2637073	3.70	860384	2.54	(-)-.BETA.-CARYOPHYLLENE EPOXIDE	81.10
10	20.610	3617812	5.08	1276308	3.77	(4,4-Dimethyl-2,4,5,6-tetrahydro-1H-inden-2-yl)acetic acid	147.10
11	20.670	12595639	17.69	5290146	15.61	ALLOAROMADENDRENOXID-(1)	81.10
12	20.767	2219791	3.12	1105412	3.26	Andrographolide	147.10
13	20.800	1440393	2.02	802009	2.37	Cycloheptanol, 1-allyl-2-methylene-	125.10
14	21.592	1602470	2.25	633436	1.87	Ergost-25-ene-6,12-dione, 3,5-dihydroxy-, (3.beta.,5.alpha.)-	69.05
15	21.639	2244009	3.15	1022513	3.02	Acetic acid, 1-[2-(2,2,6-trimethyl-bicyclo[4.1.0]hept-1-yl)-ethyl]-vinyl ester	69.10
		71198045	100.00	33894910	100.00		

Fig. 5 O. *C. neilgherrensis* GC-MS Chromatograms profile.

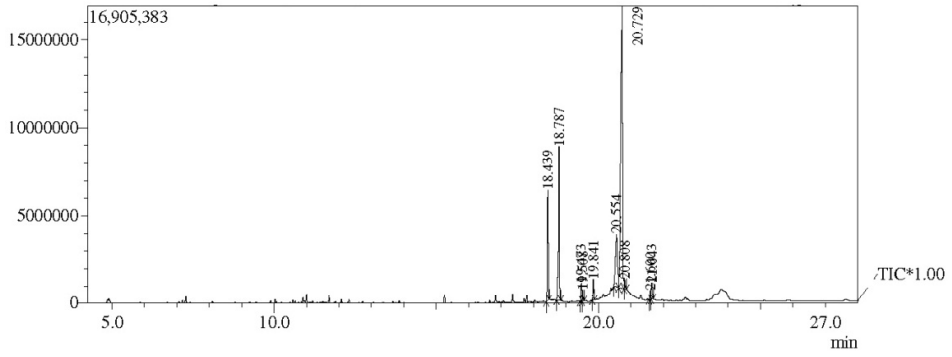


Peak Report TIC

Peak#	R. Time	Area	Area%	Height	Height%	Name	Base m/z
1	6.866	242458	1.47	162298	1.83	Camphor	95.05
2	7.235	111298	0.68	81427	0.92	2-BORNANOL	95.10
3	8.133	88456	0.54	64618	0.73	.BETA.-CITRAL	69.05
4	8.284	135140	0.82	89571	1.01	.BETA.-GERANIOL	69.05
5	8.550	120418	0.73	84090	0.95	.ALPHA.-CITRAL	69.10
6	10.056	121083	0.74	84430	0.95	2,6-OCTADIEN-1-OL, 3,7-DIMETHYL-, ACETATE	69.05
7	10.743	170064	1.03	122074	1.37	.ALPHA.-SANTALENE	94.05
8	10.908	196927	1.20	54518	0.61	(-)-.ALPHA.-TRANS-BERGAMOTENE	93.10
9	11.089	74803	0.45	53862	0.61	(Z)-.BETA.-FARNESENE	69.05
10	11.288	171567	1.04	108366	1.22	ALPHA.-CARYOPHYLLENE	93.10
11	11.506	137611	0.84	96672	1.09	ALPHA-CURCUMEN	132.15
12	11.780	301716	1.84	197379	2.22	2,4-DITERT-BUTYLPHENOL	191.10
13	11.853	274748	1.67	130800	1.47	(-)-.BETA.-BISABOLENE	69.05
14	13.129	375975	2.29	227750	2.56	(+)-CAROTOL	161.10
15	14.024	107915	0.66	71953	0.81	(Z)-iso-Geraniol	109.10
16	15.410	188753	1.15	126407	1.42	1-CYCLOHEXENE-1-PROPANOL, .ALPHA.,.GAMMA.,2,6,6-PENTAMETHYL-	137.10
17	16.978	535014	3.25	328838	3.70	4-Hexen-1-ol, 6-(2,6,6-trimethyl-1-cyclohexenyl)-4-methyl-, (E)-	137.15
18	17.948	209625	1.28	140764	1.58	Androst-5-ene-3,19-diol, 3-acetate, (3.beta.)-	146.05
19	18.195	75770	0.46	47865	0.54	LABDA-8(17),13Z-DIEN-15-OL	81.05
20	18.592	5615073	34.15	3332504	37.48	RETINAL	147.05
21	18.935	301489	1.83	201780	2.27	1-Ethyl-4,4-dimethyl-cyclohex-2-en-1-ol	125.10
22	19.992	1312704	7.98	757119	8.52	Kauren-18-ol, acetate, (4.beta.)-	97.05
23	20.780	3500172	21.29	1514501	17.03	3-BROMOCHOLEST-5-ENE #	147.05
24	20.836	1319682	8.03	474579	5.34	(-)-.BETA.-CARYOPHYLLENE EPOXIDE	81.05
25	21.792	251976	1.53	120053	1.35	Ergost-25-ene-6,12-dione, 3,5-dihydroxy-, (3.beta.,5.alpha.)-	69.05
26	21.836	499792	3.04	216550	2.44	4,8,13-DUVATRIENE-1,3-DIOL	69.05
		16440229	100.00	8890768	100.00		

Fig. 5 P. GC-MS Chromatograms of compounds from *C. oligantha* var. *oligantha* rhizome.

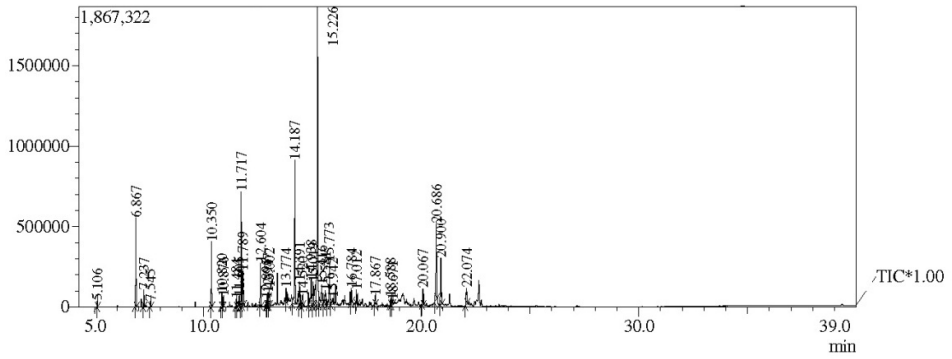
.



Peak Report TIC

Peak#	R.Time	Area	Area%	Height	Height%	Name	Base m/z
1	18.439	11217290	9.56	6317754	16.20	Retinal, 9-cis-	147.10
2	18.787	20440104	17.42	8755450	22.45	1-Heptatriacotanol	83.05
3	19.473	2057630	1.75	1067777	2.74	(ALBICANOL) DECAHYDRO-2-METHYLENE-5,5,8A-TRIMETHYL-1-NAPHTHALENEMETHANOL	97.05
4	19.508	1064834	0.91	639747	1.64	NEROLIDOL-EPOXYACETATE	85.10
5	19.841	2130707	1.82	1200762	3.08	2,4a,8,8-Tetramethyldecahydrocyclopropa[d]naphthalene	97.05
6	20.554	12853972	10.96	2996513	7.68	(-)-.BETA.-CARYOPHYLLENE EPOXIDE	81.10
7	20.729	63407841	54.05	16032212	41.11	Kauren-18-ol, acetate, (4.beta.)-	81.05
8	20.808	1000730	0.85	532935	1.37	1-Oxaspiro[4.5]decan-2-one, 6-isopropenyl-9-methyl-	125.10
9	21.600	1393282	1.19	516751	1.33	SPINACENE	69.10
10	21.643	1739962	1.48	938457	2.41	Ergost-25-ene-6,12-dione, 3,5-dihydroxy-, (3.beta.,5.alpha.)-	69.05
		117306352	100.00	38998358	100.00		

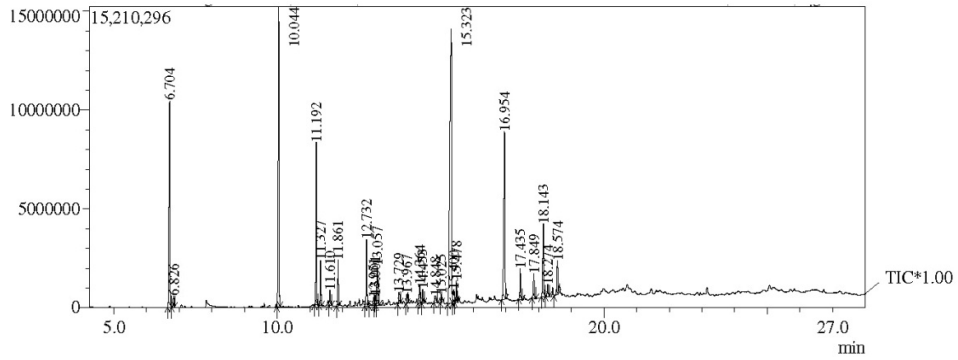
Fig. 5 Q. *C. pseudomontana* rhizome components GC-MS Chromatograms.



Peak Report TIC

Peak#	R.Time	Area	Area%	Height	Height%	Name	Base m/z
1	5.106	55058	0.34	39839	0.50	1, 8-CINEOL	81.10
2	6.867	865397	5.30	553364	6.94	Camphor	95.10
3	7.237	162601	1.00	104228	1.31	2-BORNANOL	95.10
4	7.545	51805	0.32	35491	0.44	(-)-.ALPHA.-TERPINEOL	59.05
5	10.350	597286	3.66	407200	5.10	BETA-ELEMEN-(2)	81.05
6	10.820	131547	0.81	87414	1.10	(-)-BETA-CARYOPHYLLENE	93.10
7	10.876	82981	0.51	63432	0.80	.gamma.-Elemene	121.10
8	11.484	82441	0.51	50436	0.63	5.BETA.,10.ALPHA.-EUDESMA-4(14),11-DIENE	105.10
9	11.605	126663	0.78	66179	0.83	(-)-GERMACRENE D	161.10
10	11.717	1109092	6.80	713183	8.94	Isogermafurene	108.05
11	11.789	562695	3.45	216661	2.72	1-(4-ISOPROPYLPHENYL)-2-METHYLPROPYL ACETATE	191.10
12	12.604	436298	2.67	267599	3.35	GERMACRENE B	121.10
13	12.895	75481	0.46	45455	0.57	(-)-.BETA.-CARYOPHYLLENE EPOXIDE	69.05
14	12.947	134185	0.82	79336	0.99	(+)-BETA-GUAJEN	161.10
15	13.002	237369	1.45	111627	1.40	.beta.-Elemenone	107.10
16	13.774	124415	0.76	82477	1.03	Juniper camphor	81.10
17	14.187	1757245	10.77	893465	11.20	(E,E)-GERMACRONE	107.10
18	14.391	266419	1.63	148331	1.86	Curdione	69.05
19	14.520	149552	0.92	63406	0.79	Isoaromadendrene epoxide	93.10
20	14.938	585800	3.59	165018	2.07	5,8-Dihydroxy-4a-methyl-4,4a,4b,5,6,7,8,8a,9,10-decahydro-2(3H)-phenanthrenone	159.05
21	15.019	440425	2.70	146117	1.83	(-)-Spathulenol	119.10
22	15.226	3873570	23.74	1856848	23.28	Isovelleral	232.10
23	15.416	549877	3.37	124267	1.56	(-)-BETA-CARYOPHYLLENE	109.05
24	15.581	347564	2.13	87652	1.10	6.beta.-Hydroxymethandienone	232.10
25	15.773	547256	3.35	265666	3.33	ELEMENE	68.05
26	15.942	163272	1.00	37543	0.47	1(10)-CIS-COSTUNOLIDE	68.00
27	16.784	154810	0.95	88756	1.11	6R,9R-3-Oxo-.alpha.-ionol	108.05
28	17.012	158195	0.97	85312	1.07	(3E,5E,7E)-6-Methyl-8-(2,6,6-trimethyl-1-cyclohexenyl)-3,5,7-octatrien-2-one	109.00
29	17.867	110991	0.68	53324	0.67	Longipinocarveol, trans-	137.15
30	18.588	76310	0.47	47307	0.59	Diazoprogesterone	147.05
31	18.671	51534	0.32	30467	0.38	Isocyclocitral	109.10
32	20.067	192823	1.18	100151	1.26	GAILLARDIN	188.00
33	20.686	1069185	6.55	471207	5.91	ALLOAROMADENDRENOXID-(1)	81.05
34	20.900	704473	4.32	285823	3.58	1-Heptatriacotanol	81.05
35	22.074	280599	1.72	102828	1.29	Card-20(22)-enolide, 2,3,14-trihydroxy-, (2.alpha.,3.beta.,5.alpha.)-	124.10
		16315214	100.00	7977409	100.00		

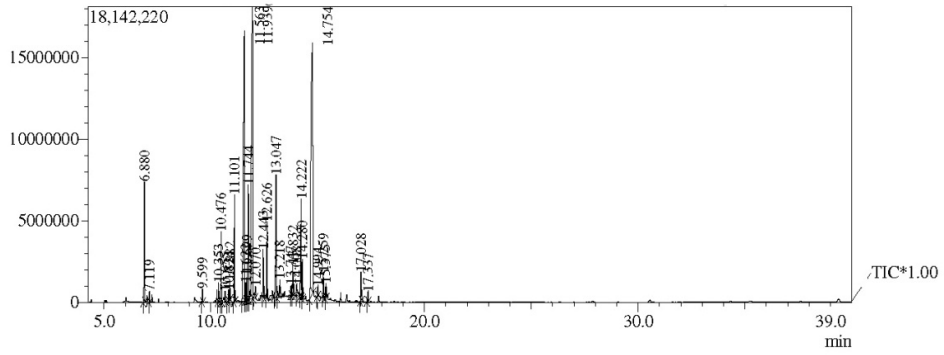
Fig. 5 R. GC-MS Chromatograms of methanolic extraction form *C. zedoaria* rhizome.



Peak Report TIC

Peak#	R.Time	Area	Area%	Height	Height%	Name	Base m/z
1	6.704	21503287	11.24	10429834	12.67	Camphor	95.15
2	6.826	1022833	0.53	573617	0.70	Camphene hydrate	71.10
3	10.044	29771291	15.56	15155814	18.41	Copaene	119.15
4	11.192	14092480	7.37	8202714	9.96	Aromandendrene	91.10
5	11.327	3520497	1.84	2278650	2.77	GAMMA-MUUROLEN	161.20
6	11.610	1385358	0.72	738218	0.90	cubedol	161.20
7	11.861	3845611	2.01	2305732	2.80	4-epi-cubedol	161.20
8	12.732	5720464	2.99	3356849	4.08	(-)-.BETA.-CARYOPHYLLENE EPOXIDE	79.10
9	12.951	1119561	0.59	403246	0.49	Diepicedrene-1-oxide	81.10
10	13.001	931895	0.49	504858	0.61	Globulol	122.15
11	13.057	3954276	2.07	1996807	2.43	Humulene epoxide 2	109.15
12	13.729	1427090	0.75	561439	0.68	ALLOAROMADENDRENOXID-(1)	91.10
13	13.967	873187	0.46	453830	0.55	(-)-SPATHULENOL	159.15
14	14.364	2064708	1.08	949762	1.15	1,4A,7,7-TETRAMETHYLDECAHYDROCYCLOPROPA[7,8]AZULENO[3A,4-B]OXIRENE	132.15
15	14.453	1291006	0.67	692687	0.84	7-Acetyl-2-hydroxy-2-methyl-5-isopropylbicyclo[4.3.0]nonane	153.15
16	14.848	1260168	0.66	338248	0.41	.BETA.-COPAEN-4.ALPHA.-OL	203.15
17	15.025	837125	0.44	444540	0.54	4-(4-METHOXY-2,3,6-TRIMETHYLPHENYL)-3-BUTEN-2-ONE	203.15
18	15.323	53099402	27.75	13883045	16.86	STAHLIANTHUSONE	228.15
19	15.400	1881980	0.98	673502	0.82	KAURAN-18-AL, 17-(ACETYLOXY)-, (4.BETA.)-	55.10
20	15.478	2124391	1.11	1116813	1.36	6-Isopropenyl-4,8a-dimethyl-1,2,3,5,6,7,8,8a-octahydro-naphthalen-2-ol	159.15
21	16.954	21831912	11.41	8480889	10.30	Anthracene, 1,2,3,4,5,6,7,8-octahydro-9,10-dimethyl-	199.15
22	17.435	2762396	1.44	1553891	1.89	6-Acetyl-5-methoxy-2,7-dimethyl-1,4-naphthoguinone	243.15
23	17.849	2497806	1.31	1255245	1.52	2-[5-(2,2-Dimethyl-6-methylene-cyclohexyl)-3-methyl-pent-2-enyl]-[1,4]benzoquinone	161.15
24	18.143	7679643	4.01	3682104	4.47	Coumarine, 8-allyl-7-hydroxy-6-ethyl-4-methyl-	229.10
25	18.274	1103889	0.58	590686	0.72	4-(6,6-Dimethyl-2-methylene-cyclohex-3-enylidene)-pentan-2-one	161.10
26	18.574	3722600	1.95	1709208	2.08	6-METHYL-5-(1-METHYLETHYLIDENE)-3,6,9-DECATRIEN-2-ONE	161.15
		191324856	100.00	82332228	100.00		

Fig. 5 S. Rhizome extract GC-MS profile of *C. vamana*.



Peak Report TIC

Peak#	R.Time	Area	Area%	Height	Height%	Name	Base m/z
1	6.880	13482951	3.85	7358916	6.30	Camphor	95.10
2	7.119	1016716	0.29	625873	0.54	Isoborneol	95.10
3	9.599	1174082	0.33	795546	0.68	DELTA-ELEMENE	121.10
4	10.353	1804703	0.51	1173546	1.00	BETA-ELEMEN	93.10
5	10.476	6587520	1.88	4314109	3.69	(-)-ZINGIBERENE	119.10
6	10.655	1090806	0.31	690695	0.59	trans-.alpha.-Bergamotene	93.10
7	10.824	1179143	0.34	731876	0.63	Caryophyllene	93.10
8	10.882	1769906	0.50	1205561	1.03	.gamma.-Elemene	121.10
9	11.101	10318809	2.94	6546647	5.60	(E)-.beta.-Famesene	69.05
10	11.563	61786529	17.62	16548517	14.16	ALPHA-CURCUMEN	132.15
11	11.622	2212940	0.63	1097979	0.94	(-)-GERMACRENE D	161.15
12	11.699	3138240	0.90	1612036	1.38	Alpha-Cedrene	119.10
13	11.744	11830191	3.37	7054561	6.04	Curzerene	108.05
14	11.939	73820586	21.05	17915798	15.33	.beta.-curcumene	119.10
15	12.070	1659151	0.47	719168	0.62	(+)-.BETA.-FUNEBRENE	69.05
16	12.443	4412459	1.26	2858216	2.45	7-epi-cis-sesquisabinene hydrate	69.05
17	12.626	7836700	2.24	4710009	4.03	GERMACRENE B	121.15
18	13.047	16903760	4.82	7499597	6.42	N-(4-HYDROXYPHENYL)-N,N',N'-TRIMETHYLSULFAMIDE	122.05
19	13.218	2892265	0.82	1101075	0.94	Cedren-13-ol, 8-	119.10
20	13.747	1018292	0.29	605917	0.52	(+)-.BETA.-EUDESMOL	59.05
21	13.832	3816277	1.09	1883865	1.61	.beta.-bisabolol	82.10
22	14.008	2974774	0.85	607504	0.52	7-epi-trans-sesquisabinene hydrate	119.10
23	14.222	14370310	4.10	6019596	5.15	GERMACRON	107.10
24	14.280	3689904	1.05	2372986	2.03	2-Heptanone, 6-methyl-6-[3-methyl-3-(1-methylethenyl)-1-cyclopropen-1-yl]-	135.10
25	14.754	90585245	25.83	15422794	13.20	Phenol, 2-methyl-5-(1,2,2-trimethylcyclopentyl)-, (S)-	136.10
26	14.994	1009889	0.29	588202	0.50	5,8-Dihydroxy-4a-methyl-4,4a,4b,5,6,7,8,8a,9,10-decahydro-2(3H)-phenanthrenone	159.10
27	15.259	2498660	0.71	1514231	1.30	Isovelleral	232.10
28	15.375	1591705	0.45	888073	0.76	Dehydrozingerone	145.00
29	17.028	3104241	0.89	1783371	1.53	Cyclohexyldimethylsilyloxybenzene	151.05
30	17.337	1056308	0.30	608050	0.52	(-)-ISOLONGIFOLINE	175.05
		350633062	100.00	116854314	100.00		

Fig. 5 T. Methanolic rhizome extract GC-MS Chromatograms of *C. zanthorrhiza*.

5.2 UFLC analysis

5.2.1 Method Development and Optimization

5.2.1.1 Selection of wavelength for determination

The standard solution of curcumin (1mg/mL) was scanned in the wavelength range of 300-600nm. The responses of the standard solution measured with PDA detector showed a good resolution at 421 nm using the UFLC method.

5.2.1.2 Selection of Columns

For conducting UFLC analysis, various columns are available but our main objective was to determine the curcumin content from different *Curcuma* species. the C18 (150mm x 4.6mm i.d., 5µm) column was selected for curcumin analysis and this column was chosen to give good peak and high resolution, besides high peak symmetry, good retention in curcuminoid separation from all the species without any interference within a short run time.

5.2.1.3 Selection of mobile phase

The chromatographic conditions were optimized for the method, which can separate curcuminoids from the crude extract with a good resolution. The mobile phase was chosen after several trials with water, methanol, phosphoric acid and acetonitrile (ACN). Finally, the mobile phase consists of water+0.2% (v/v) phosphoric acid (A), and acetonitrile (B) in an isocratic mode was selected to achieve maximum separation. Flow rates were selected between 0.5 and 2.5 ml/min. A flow rate of 1 ml/min gave an optimum signal/noise ratio with a reasonable separation time (30 min) and the best separation efficiency was obtained using C_{18} column.

5.2.1.4 Optimized chromatographic condition

Column	**:**	Kromasil C18 (150mm x 4.6mm i.d., 5µm)
Mobile phase	**:**	Acetonitrile and Water+0.2% (v/v) phosphoric acid, pH was adjusted at 5.0)
Detection wavelength	**:**	421nm
Flow rate	**:**	1 ml/min
Injection volume	**:**	20µl
Column Oven Temperature	**:**	40 ^{o} C
Sample Cooler Temperature	**:**	25 ° C

5.2.2 Method validation

5.2.2.1 Solution stability

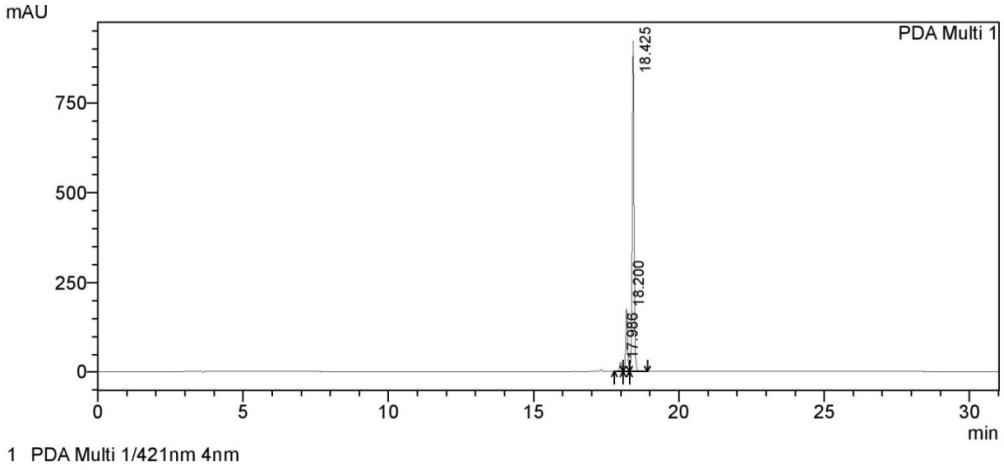
Stability of the analytes in the sample solution was evaluated by analyzing the sample solution of 5µg/ml at 4, 8, 12, 16, 20 and 24 hr after the preparation of sample at room temperature. The analytes were stable in the sample solution with RSDs values between 0.25% and 1.76% for all targeted compounds. The stability is supported by the low%RSDs, as shown in Table 5.2 and demonstrates that this method had good stability.

Table 5.2: the standard solution stability of various curcuminoids obtained from UFLC method.

	CUR	DMC	BDMC
%RSD	0.25	1.70	1.76

5.2.2.2 Specificity

The present analytical method observed was specific with no interference in curcuminoids separation. The analytical result showed that the curcuminoid peaks were free from any impurities (Fig. 5.1).



PeakTable

PDA Ch1 421nm 4nm

Peak#	Ret. Time	Area	Height	Area %	Height %
1	17.986	124497	27410	2.416	2.434
2	18.200	813072	177264	15.780	15.738
3	18.425	4215066	921651	81.804	81.828
Total		5152635	1126326	100.000	100.000

Fig. 5.1. UFLC chromatogram of curcumin standard solution (1mg/ml) at wavelength using C18 column

5.2.2.3 Linearity and range, limits of detection and quantification

The curcuminoid contents were determined to prepare calibration graphs of two *Curcuma* species. The calibration curves were obtained from three injections of five different concentrations of curcuminoids versus the peak area. Linearity was determined in the concentration range between 1µg and 300µg, with high reproducibility and accuracy (Fig. 5.2A-C). Regression analysis of the experimental data points showed a linear relationship with excellent correlation coefficients (r^2) of curcumin, demethoxycurcumin, and

bisdemethoxycurcumin of 0.998, 0.994, and 0.993 respectively. The linear regression equations for the curves for curcumin, demethoxycurcumin, and bisdemethoxycurcumin were $y = 30805x + 889.6$, $y = 832.6x - 30839$, and $y = 7042x - 20553$, respectively.

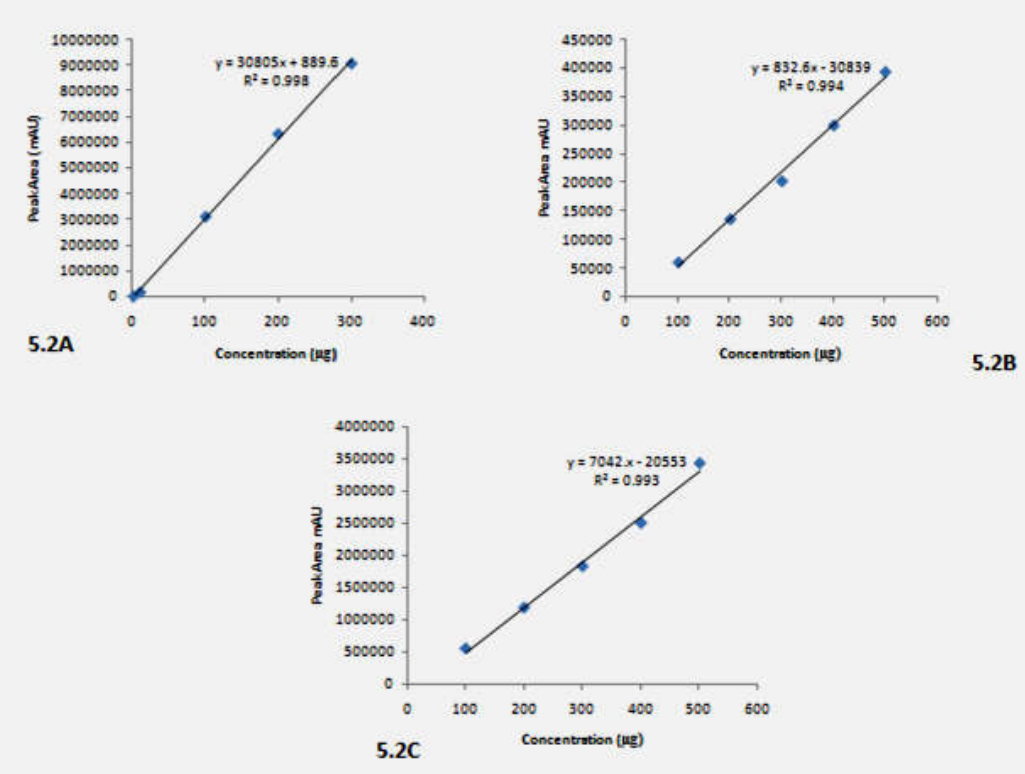


Fig. 5.2. Linear relationship between peak area and concentration of **A-** curcumin, **B-** demethoxycurcumin and **C-** bisdemethoxycurcumin**.**

5.2.2.4 Precision

The intra-day and inter-day precision was estimated by analyzing six independently prepared samples using the same method, operator, instruments, and equipment. Intraday assay variation was evaluated by injecting curcumin (1µg/ml) six times during the same day. Inter-day assay variation was assessed by injecting three curcumin concentrations (1, 3, and

5µg/ml) in replicates of three during three different days. The relative standard deviation was determined for evaluating the precision. RSDs for intra-day and inter-day precision were below 4.5% for all of the compounds. The% RSD for intra-day precision was found to be in the range of 1.00-0.27, 2.81-1.14 and 2.92-1.38 while inter-day precision was found to be in the range of 0.01-1.61, 0.02-1.71 and 0.64-0.85 for CUR, DMC and BDMC, respectively (Table 5.3).

Table 5.3: Intra-day and Inter-day precision parameters for curcumin, demethoxycurcumin and bisdemethoxycurcumin by using different concentrations of the sample. The experiments were repeated as Intra-day (n= 6) and Inter-day (n=3).

Analyte Concentration (1µg, 3µg and 5µg/mL)		Intra-day (n= 6)		Inter-day (n=3)	
		Detected (µg/mL)	%RSD	Detected (µg/mL)	%RSD
CUR	1µg	1.4328±0.01	1.00	1.8023±0.12	1.61
	3µg	2.4841±0.01	0.70	2.4313±0.38	0.03
	5µg	4.7475±0.01	0.27	4.0605±0.56	0.01
DMC	1µg	1.2391±0.03	2.81	1.5648±0.01	1.02
	3µg	1.7163±0.05	3.37	2.9037±0.01	1.71
	5µg	3.9308±0.04	1.14	4.4844±0.02	0.58
BDMC	1µg	1.47±0.04	2.92	1.6715±0.01	0.79
	3µg	2.6208±0.03	1.38	2.99±0.02	0.85
	5µg	4.8625±0.07	1.44	3.4515±0.02	0.64

5.2.2.5 Accuracy

Three different quantities (low, medium, and high) of the standards were added into the sample to evaluate the percentage recovery of the

developed analytical method. The amount of each component was subsequently calculated from the calibration curves. The method had an acceptable accuracy with the overall recovery from 99.15 to 105.50% and percentage RSDs below 3.25% for the three marker compounds. The value of overall recovery and low percentage RSDs showed that the methods used in the analysis had a good accuracy. A good recovery indicates the accuracy of the method. The results are summarized in Table 5.4.

Table 5.4: Results of recovery studies of curcuminoids extracts of *C. longa, C. aromatica, C. zanthorrhiza* and *C. zedoaria* (n=3).

Samples	Amt. of Sample (µg/mL)	Amt. of Std Added (µg/mL)	Amt. found (µg/mL)	%Recovery± S.D. (n=3)	%RSD
CUR	5	4	8.99	99.90±1.22	1.22
	5	7	11.89	99.16±2.21	2.23
	5	9	14.77	105.50±1.46	1.39
DMC	4	4	7.97	99.68±1.77	1.78
	4	6	9.92	99.29±3.18	3.20
	4	8	11.83	98.58±3.20	3.24
BDMC	5	1	5.90	98.33±3.15	3.20
	5	2	6.91	98.77±1.26	1.27
	5	3	7.93	99.15±2.59	2.61

5.2.2.6 Limit of detection and limit of quantification:

The estimated LOD and LOQ in this study were found to be 0.72-0.92 and 0.80-1.42 respectively (Table 5.5).

Table 5.5: Result of limit of detection and limit of quantification

Curcuminoids	LOD (µg/mL)	LOQ (µg/mL)
Curcumin	0.72	1.42
Demethoxycurcumin	0.92	0.80
Bisdemethoxycurcumin	0.86	1.04

5.2.3 Determination of curcuminoids in *C. zanthorrhiza, C. aromatica, C. longa* and *C. zedoaria.*

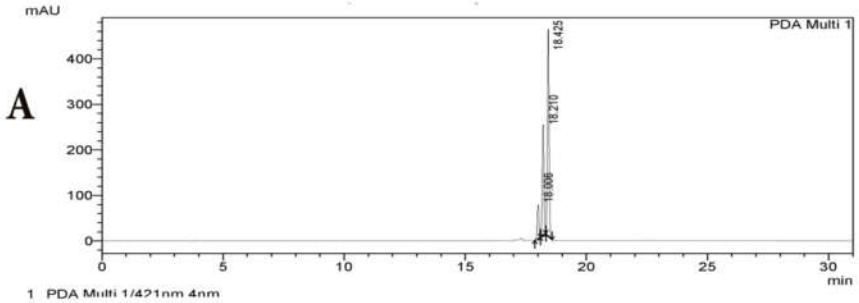
UFLC method was successfully utilized for the determination of CUR, DMC and BDMC from economically important *Curcuma* species like *C. longa*, *C. aromatica*, *C. zanthorrhiza* and *C. zedoaria*. Using the calibration curve of each compound, the three analytes in the four *Curcuma* species were determined (Table 5.6) and it was observed that the levels of the three individual analytes present in the samples varied considerably. The highest curcumin content was determined in the nine months old *Curcuma* species (15.0-1705.7µg/g) followed by six (14.5-352.6µg/g) and three (3.7-111.6µg/g) months old rhizomes. The highest curcumin content was detected in *C. longa* (1918.3µg/g) followed by *C. zanthorrhiza* (1240.4µg/g), *C. aromatica* (91.8µg/g) and *C. zedoaria* (33.2µg/g).

However, the other curcuminoids, DMC (8.55-54106.8µg/g) and BDMC (1.4-6683.3µg/g) were quantitatively different among different species. It is well known that the quality of herbal products can be influenced by many factors, such as the cultivating site, harvesting time, and post-harvest

handling. The percentage of curcuminoids was found to be satisfactory (Table 5.6). The compounds identified were curcumin, demethoxycurcumin and bisdemethoxycurcumin and retention time and peak area of all the species shown in Fig. 5.6-5.9.

Table 5.6: Curcuminoids concentration of different aged plants of *C. zanthorrhiza*, *C. aromatica, C. longa* and *C. zedoaria.*

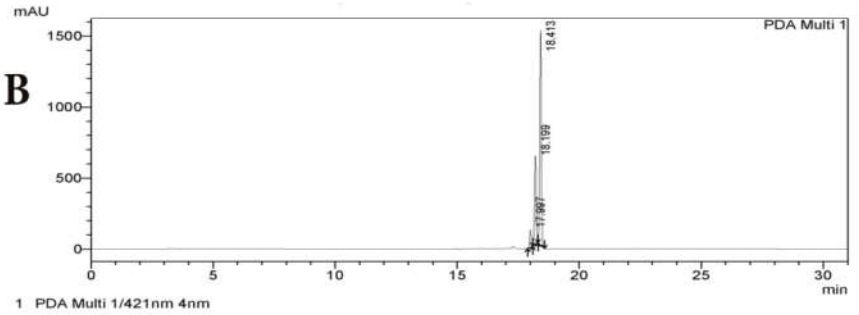
Curcuminoids	Months	*C. zanthorrhiza*	*C. aromatica*	*C. longa*	*C. zedoaria*
CUR	3	111.6µg/g	15.5µg/g	72.5µg/g	3.7µg/g
	6	352.6µg/g	30.3µg/g	140.1µg/g	14.5µg/g
	9	776.2µg/g	46.0µg/g	1705.7µg/g	15.0µg/g
DMC	3	1070 µg/g	153.1µg/g	3812.1µg/g	8.5µg/g
	6	615.4 µg/g	100.0µg/g	1759.7µg/g	203.8µg/g
	9	1557.9 µg/g	104.0µg/g	5410.8µg/g	163.6µg/g
BMDC	3	692.5 µg/g	92.6 µg/g	666.0µg/g	1.4µg/g
	6	266.4 µg/g	59.5 µg/g	270.3µg/g	6.9µg/g
	9	2135.4 µg/g	17.0 µg/g	6683.3µg/g	3.6µg/g



PeakTable

PDA Ch1 421nm 4nm

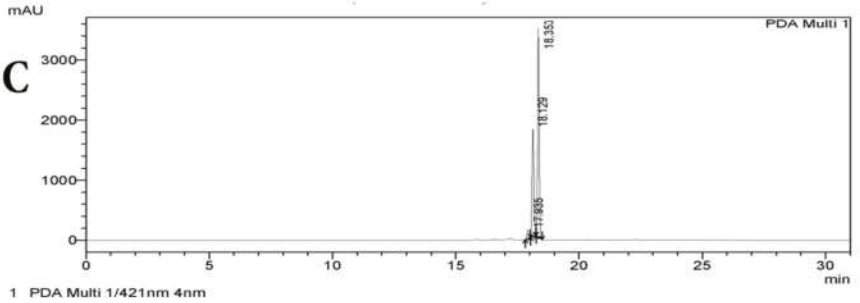
Peak#	Ret. Time	Area	Height	Area %	Height %
1	18.006	338304	74361	9.533	9.641
2	18.210	1146264	243415	32.301	31.561
3	18.425	2064162	453485	58.166	58.798
Total		3548730	771261	100.000	100.000



PeakTable

PDA Ch1 421nm 4nm

Peak#	Ret. Time	Area	Height	Area %	Height %
1	17.997	565827	125183	5.640	5.528
2	18.199	2946844	625398	29.376	27.618
3	18.413	6518936	1513860	64.984	66.854
Total		10031606	2264441	100.000	100.000

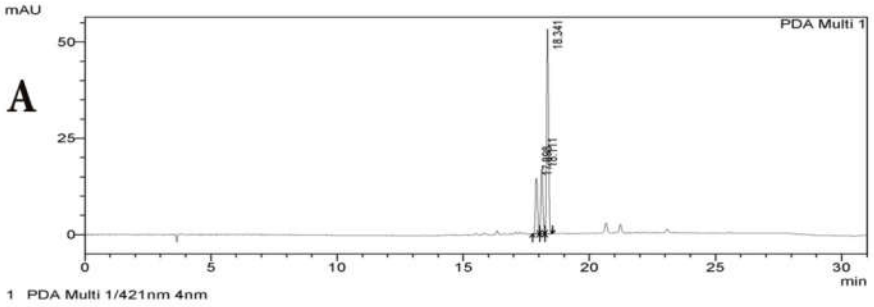


PeakTable

PDA Ch1 421nm 4nm

Peak#	Ret. Time	Area	Height	Area %	Height %
1	17.935	809149	149462	3.343	2.763
2	18.129	9043198	1786040	37.367	33.017
3	18.353	14348880	3473898	59.290	64.220
Total		24201228	5409400	100.000	100.000

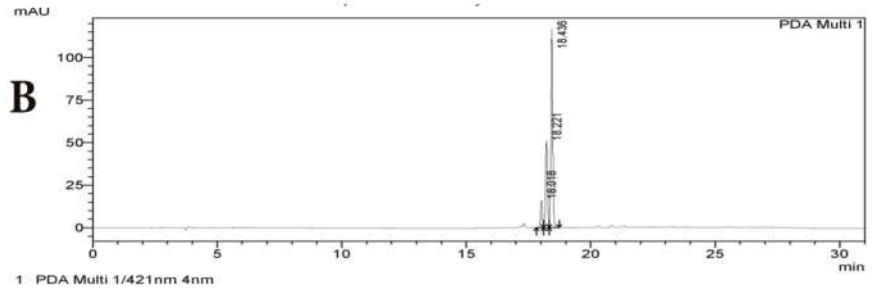
Fig. 5.6. *C. zanthorrhiza* UFLC chromatograms of curcuminoids-**A-**3rd, **B-** 6th and **C**-9thmonths.



PeakTable

PDA Ch1 421nm 4nm

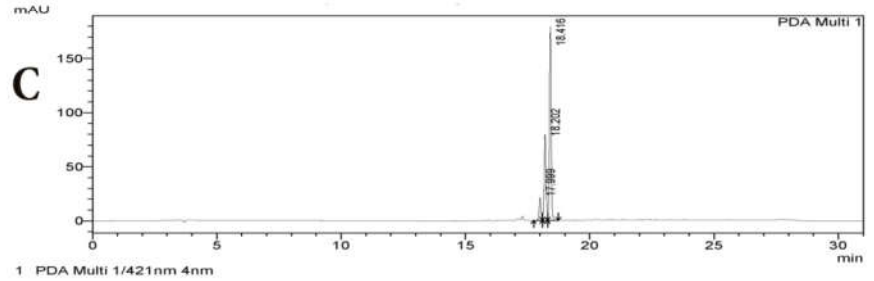
Peak#	Ret. Time	Area	Height	Area %	Height %
1	17.898	83137	14388	17.924	17.106
2	18.111	92411	16593	19.924	19.727
3	18.341	288275	53132	62.152	63.167
Total		463824	84113	100.000	100.000



PeakTable

PDA Ch1 421nm 4nm

Peak#	Ret. Time	Area	Height	Area %	Height %
1	18.018	81172	15946	8.870	8.685
2	18.221	272289	51042	29.754	27.801
3	18.436	561677	116610	61.376	63.514
Total		915137	183598	100.000	100.000

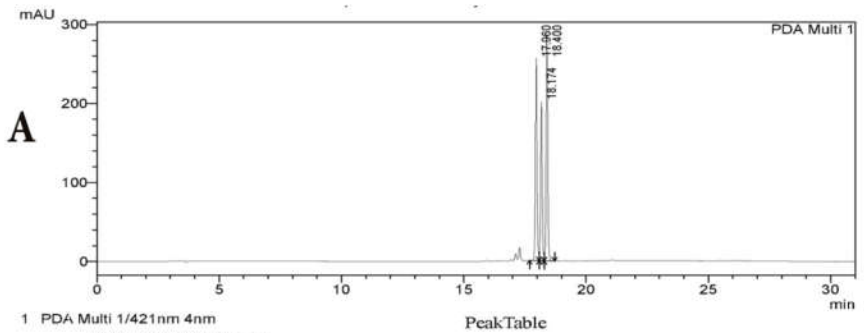


PeakTable

PDA Ch1 421nm 4nm

Peak#	Ret. Time	Area	Height	Area %	Height %
1	17.999	107368	21156	7.831	7.565
2	18.202	412058	79469	30.054	28.417
3	18.416	851612	179026	62.114	64.018
Total		1371039	279651	100.000	100.000

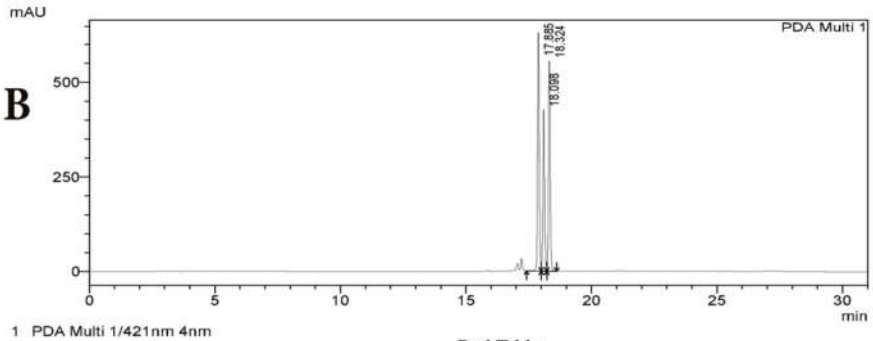
Fig. 5.7. *C. aromatica* UFLC chromatograms of curcuminoids-**A**-3rd, **B**-6th and **C**-9th months.



PeakTable

PDA Ch1 421nm 4nm

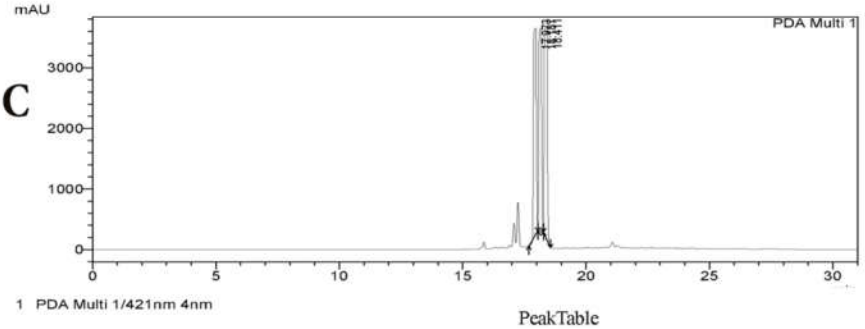
Peak#	Ret. Time	Area	Height	Area %	Height %
1	17.960	1162916	256333	34.059	34.486
2	18.174	909944	200919	26.650	27.030
3	18.400	1341534	286053	39.291	38.484
Total		3414394	743305	100.000	100.000



PeakTable

PDA Ch1 421nm 4nm

Peak#	Ret. Time	Area	Height	Area %	Height %
1	17.885	2834874	630557	38.513	39.077
2	18.098	1935217	427180	26.290	26.473
3	18.324	2590817	555906	35.197	34.450
Total		7360909	1613643	100.000	100.000

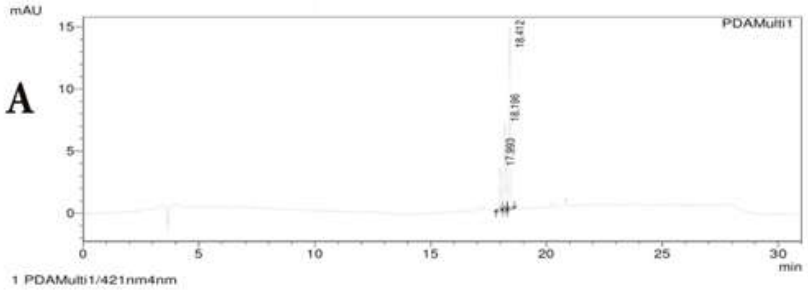


PeakTable

PDA Ch1 421nm 4nm

Peak#	Ret. Time	Area	Height	Area %	Height %
1	17.973	28259188	3396863	32.538	33.191
2	18.187	27060403	3353966	31.158	32.772
3	18.411	31528898	3483513	36.303	34.037
Total		86848489	10234342	100.000	100.000

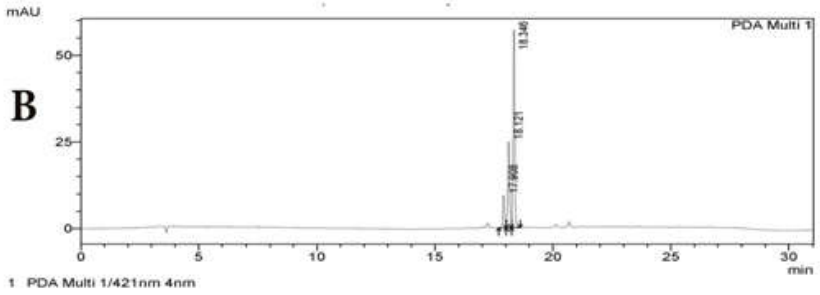
Fig. 5.8. *C. longa* UFLC chromatograms of curcuminoids-**A-**3rd, **B-**6th and **C-**9th months.



PeakTable

PDACh1421nm4nm

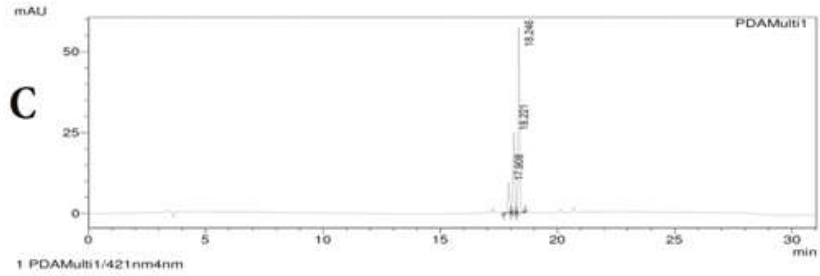
Peak#	Ret.Time	Area	Height	Area %	Height%
1	17.993	26835	3291	13.793	13.344
2	18.196	35113	6850	28.768	27.780
3	18.412	70107	14518	57.439	58.876
Total		132055	24659	100.000	100.000



PeakTable

PDA Ch1 421nm 4nm

Peak#	Ret. Time	Area	Height	Area %	Height %
1	17.908	49930	9324	11.042	10.220
2	18.121	132690	24822	29.345	27.208
3	18.346	269554	57086	59.613	62.572
Total		452175	91232	100.000	100.000



PeakTable

PDACh1421nm4nm

Peak#	Ret.Time	Area	Height	Area %	Height%
1	17.908	35930	8314	12.042	11.120
2	18.221	112590	21522	28.945	28.408
3	18.246	278594	58036	59.013	60.472
Total		427116	87872	100.000	100.000

Fig. 5.9. *C. zedoaria* UFLC chromatograms of curcuminoids-**A**-3rd, **B**-6th and **C**-9th months**.**

Table 5.7: Summary of Validation Parameters

Parameter		CUR	DMC	BDMC
Linearity Range (µg/ml)		1-300µg	1-300µg	1-300µg
Correlation coefficient (R^2)		0.998	0.994	0.993
Regression equation		y=30805x + 889.6	y =832.6x - 30839	y = 7042x - 20553
Precision (% R.S.D)	**Intraday precision (n=6)**	1.00-0.27	2.81-1.14	2.92-1.44
	Interday precision (n=3)	1.61-0.01	1.02-0.58	0.85-0.64
Accuracy (%recovery) (n=3)		99.16-105.50	98.58-99.68	98.33-99.15
Limit of Detection (µg/ml)		0.72	0.92	0.86
Limit of Quantitation (µg/ml)		1.42	0.80	1.04

5.3 Discussion

5.3.1 GC-MS

The rhizome extracts of *Curcuma* species has wide range pharmacological properties, such as anti-inflammatory, anticancerous, antiproliferative, hypocholesterolemic, antidiabetic, antihepatotoxic, antidiarrheal, carminative, diuretic, antirheumatic, hypotensive, antioxidant, antimicrobial, antiviral, insecticidal, larvicidal, antivenomous, antithrombotic, antityrosinase and cyclooxygenase-1 (COX-1) inhibitory activities, *etc.* (Dosoky and Setzer, 2018).

Previous studies on the important constituents were carried out using the essential oil extracted from the rhizome, while the present work was carried out using methanolic extracts. The results indicated that methanolic

extracts from these species are a good source of 1, 8 cineole, camphor, camphene, ar-curcumene and beta-elemene, which are known to have great application in pharmaceutical and flavoring industries. These compounds obtained from the essential oils of *Curcuma* species possess analgesic, antimicrobial, anti-inflammatory, insecticidal and anticancer properties (Chen *et al.,* 2013). The present study identified the same compounds from the rhizomes extracts of *Curcuma* species investigated. Fifteen components were identified from the methanolic extracts of *C. amada* (George and Britto, 2016). *C. amada* is a perennial herb called as "mango ginger" and "manga manjal" because of its raw mango flavor that is mainly attributed to the presence of -3-carene, myrcene, and (Z)-ocimene (Varadarajan *et al.,* 2018). It is used in culinary preparations, medicines, and as a source of starch (Policegoudra *et al.,* 2011). Myrcene is another dominant compound obtained from the rhizome essential oil of *C. amada* (Padalia *et al.,* 2013). Other components present in *C. amada* were (E)-hydroocimene, (Z)-hydroocimene, myrcene, and linalool, ar-curcumene, camphor, curzerenone, and 1, 8-cineole (Srivastava *et al.,* 2001). The major chemical constituents in *C. aromatica* are camphor, camphene, borneol, 1, 8-cineole, β-elemene, 8, 9-dehydro-9-formyl-cycloisolongifolene, germacrone, ar-turmerone, turmerone, curdione (Xiang *et al.,* 2018), camphor (Herath *et al.,* 2017; Angel *et al.,* 2014), xanthorrhizol, ar-curcumene, di-epi-cedrene (Nampoothiri *et al.,* 2015).

Essential oils exhibit a number of biological activities and hence have important applications in the pharmaceutical industry. Beta-elemene, with tumouricidal effects, was detected in the rhizomes of *C. aromatica, C. brog, C. zedoaria* and essential oils (Fu *et al.,* 1984). These components were identified in the rhizome methanolic extracts of *C. aromatica*, *C. zedoaria, C. zanthorrhiza* and *C. aeruginosa.* Curcumenol with analgesic properties was detected in the *C. aeruginosa* oil (De Fatima *et al.,* 2002). Camphor is one of the main compound responsible for developing the cholagogic effect of

essential oils (Ozaki and Liang, 1988), which is a common constituent in the essential oil of *Curcuma* species and also is detected in the methanolic extracts of *Curcuma* species.

The major components present in the *C. longa* essential oil were ar-turmerone, α-tumerone, β- turmerone, β-caryophyllene and eucalyptol (Naz *et al.,* 2010). The major components reported in the essential oil of *C. inodora* were caryophyllene and alpha-caryophyllene (Rajesh Kumar and Yusuf, 2017) and curzerenone, germacrone, curdione, and 1,8-cineole (Malek *et al.,* 2006). The chemical composition of essential oils from eight underutilized starchy *Curcuma* species was studied (Angel *et al.,* 2014), and the major compounds identified were, 1, 8 cineole, camphor, camphene, α -pinene, and β-pinene. The components of essential oil were identified form 12 *Curcuma* species using GC-MS and the major chemical components were curdione, ar-turmerone, germacrone, β-elemenone and velleral (Zhang *et al.,* 2017). The essential oil of *C. aeruginosa* rhizome contain 8, 9-dehydro-9-formyl-cycloisolongifolene, dihydrocostunolide, germacrone, curzerenone (Theanphong *et al.,* 2015), dehydrocurdione, curcumenol, 1, 8-cineole, germacrone (Srivilai *et al.,* 2018), camphor, germacrone, curcumenol, α-pinene (Angel *et al.,* 2014).

The essential oil of *C. aurantiaca* contain1,8-cineole, camphor, germacrone, β-elemene, curzerene, and β-elemenone (Liu *et al.,* 2012). *C. mangga* rhizome essential oil have two chemotypes, caryophyllene oxide and caryophyllene-rich chemotype and myrcene-dominated chemotype (Wahab *et al.,* 2011). The major compounds detected in the essential oil of *C. oligantha* rhizome are caryophyllene, phytol, humulene, elemene, and caryophyllene respectively (Herath *et al.,* 2017). *C. pseudomontana* rhizome from India contain chemical compounds like β-elemenone, pseudocumenol, germacrone,

2-(4-methoxyphenyl) N, N-trimethyl-1-pyrrolamine and (1,5-dimethyl-4-hexenyl)-4-methylbenzene respectively (Muniyappan *et al.,* 2014).

5.3.2 Relationship between *Curcuma* species

The dendrogram generated using the statistical analysis of the identified compounds (≥ 5% concentration) of each species is shown in Fig. 5.1. Cluster I consists of *C. aeruginosa* and *C. zanthorrhiza*. Cluster II consists of *C. aromatica* and *C. zedoaria.* Cluster III includes *C. amada*, *C. bhatii, C. mutabilis, C. kudagensis* and *C. pseudomontana*. Cluster IV includes *C. coriacea, C. neilgherrensis* and *C. montana*. Cluster V contained *C. haritha*, *C. karnatakensis* and *C. vamana* respectively. Cluster VI harbours the species like *C. longa* and *C. oligantha*. Cluster VII included the species like *C. decipiens* and *C. inodora*. *C. aurantiaca* exist a monoclade in Cluster VIII. The evolutionary relationship of *Curcuma* species using *mat*K and *mat*K+*rbc*L sequence analysis of leaf tissues (Santhoshkumar and Yusuf 2017 & 2018) showed slight similarity with the present dendrogram constructed on the basis of methanolic extracts of rhizome, indicating the usefulness of rhizome extracts data in chemo taxonomical applications.

Curcuma is one of the most important plants that have been used as food and pharmaceutical ingredients worldwide. The identification of major compounds in the rhizome extracts from twenty *Curcuma* species provided a chemical blueprint assisting in the identification of these species. The chemical composition obtained from the rhizome extracts were used to construct the dendrogram divided into eight clusters for the twenty species. In this study, we included some of the wild species that are not been exploited pharmacologically, and thus this study will assist in their species identification and utilization.

5.3.3 UFLC

Ultra fast liquid chromatography (UFLC) is one of the most commonly used methods for the determination of different compounds present in plant material. The UFLC analysis of methanol extract of *Curcuma* plants showed different peaks. Curcumin is one of the major active ingredients of dietary spice found in rhizomes of turmeric (*C. longa*) (Prasad *et al*., 2014). Curcuminoids are polyphenolic substances containing a mixture of curcumin (60%–80%), demethoxycurcumin (15%–30%) and bisdemethoxycurcumin (2%–6%) (Wichitnithad *et al.,* 2009) and it is soluble in methanol, ethanol, DCM or DMSO and insoluble in water (Tonnesen and Karlsen, 1985). Curcuminoids have medicinal properties such as anti-inflammatory, antimutagenic, anti-diabetic, anti-bacterial, hepatoprotective activities (Krup *et al.,* 2013). It has antioxidant activity as well as free-radical scavenging properties (Kalpravidh *et al.,* 2010), healing of dermal wound (Gopinath *et al.,* 2004), and prevention of Alzheimer's disease (Shen and Ji, 2012). Most importantly curcumin inhibit the cell growth of various cancer cell lines, induces apoptosis of cancer cells and is also effective on the cell-cycle regulation of cancer cells (Liu *et al.,* 2007). Curcumin is sensitive to light, hence it is suggested that their biological samples containing curcumin should be protected from light (Prasad *et al.,* 2014).

5.3.4 Validation of UFLC method

The pure curcuminoids stock solution (1mg/ml) was prepared. Standard solutions were also prepared using the stock solution of individual curcumin and derived the calibration curve using the linearity range 1-300 µg/ ml (Fig. 5.1A-C).

Generally ethanol is used as the mobile phase in HPLC analysis of curcuminoids. The three curcuminoids (curcumin, demethoxycurcumin and

bisdemethoxycurcumin) was first isolated by Tonnesen and Karlsen (1983) using HPLC in a Nucleosil NH_2 column stationary phase, flow rate of 1.2 m1 as mobile phase of ethanol. Other than ethanol the various solvents such as methanol, acetonitrile and tetrahydrofuran are used as stationary phase (C_{18} and NH_2) and different methods (isocratic and gradient elution) are used for the analysis of curcuminoids (Cooper *et al.,* 1994). Curcuminoids were analyzed using HPLC with acetonitrile: water containing (0.2% phosphoric acid), 60:40 (v/v). Individual pure peak could be seen in the HPLC chromatogram. The HPLC analysis of compounds showed single peaks with retention times of 18.4, 18.2 and 17.9 min, respectively. Spectroscopic data confirmed that the single peak at 18.4, 18.2 and 17.9 min represents curcumin, demethoxycurcumin and bisdemethoxycurcumin respectively. Jayaprakasha *et al.* (2002) reported that HPLC chromatogram showed three peaks, followed by Goel *et al.* (2008) identified these peaks with the major one as curcumin and two minor peaks were identified as demethoxycurcumin and bisdemethoxycurcumin determined by co-injection of curcumin standards and confirmed elusions of peaks by ^{1}H and ^{13}C NMR analysis. The peaks for curcumin, demethoxycurcumin and bisdemethoxycurcumin within the two species were observed and identified in the chromatogram.

Jayaprakasha *et al.* (2002) determined curcuminoids using HPLC separation with C_{18} column, methanol, 2% AcOH, and acetonitrile was used as mobile phase, while, Wichitnithad *et al.* (2009) reported a simple isocratic method with the elution of acetonitrile and 2% acetic acid (40:60, v/v) at a flow rate of 2.0 mL min for separation of the three curcuminoids. The isocratic method served as the best method for the separation of all the three curcuminoids. Another study Jadhav *et al.,* (2007) reported that RP C-18 column and acetonitrile: water (50:50 v/v) and 0.1% trifluroacetic acid was used as mobile phase. The retention times of bisdemethoxycurcumin, demethoxycurcumin and curcumin were recorded at retention time of 7.2,

8.12 and 9.0 min interval, respectively. Revathy *et al*. (2011) reported HPLC method for the analysis of curcuminoids usingC18 column and mobile phase comprising of tetrahydrofuran: water containing 1% citric acid (40:60) were used to separate curcumin, demethoxycurcumin and bisdemethoxycurcumin at retention time of 10.81, 12.79, 13.03 min, respectively. Erpina *et al.,* (2017) reported the HPLC analysis of curcuminoids using the Phenomenex C_{18} column with acetonitrile-0.001% formic acid used as a mobile phase to separate curcuminoids at retention time of 9.11, 9.96 and 10.82. The separation of curcuminoids using HPLC with fluorescence detector reported by Tonnesen and Karlsen (1986). Rafi *et al.,* (2015) reported the HPLC analysis of curcuminoids in *C. zanthorrhiza* and *C. longa* samples using C_{18} column with acetonitrile and 0.5% acetic acid in water used as a mobile phase to separate of curcuminoids at retention time of 9.07, 9.84 and 10.65 respectively. Separation curcuminoids (CUR, DMC and BDMC) in the HPLC system is known to be influenced by the mobile phase (Jayaprakasha *et al.,* 2002).

The previous studies reported the analysis of curcuminoids from different *Curcuma* species, including *C. longa*, *C. zanthorrhiza, C. mangga, C. heyneana, C. aeruginosa, C. comosa, C. longa, C. malabarica, C. montana, C. zedoaria* and *C. soloensis* using HPLC (Syamkumar, 2008 thesis). In the present study we analyzed curcuminoids from four *Curcuma* species *viz*, *C. longa, C. aromatica, C. zedoaria* and *C. zanthorrhiza* during different growth periods using the developed UFLC method. The highest percentage of curcuminoids was detected in the nine months old *C. longa* and the lowest was in *C. zedoaria*. However, three and six month old rhizomes of *C. zanthorrhiza* showed higher curcumin content compared with *C. longa*. Higher amount of DMC content showed in *C. zedoaria* compared to CUR and BDMC.

5.3.5 Validation of analytical methods

The developed method was evaluated for the calculation of Limit of detection (LOD) and Limit of quantification (LOQ) of the three curcuminoids (CUR, DMC and BDMC) employing two different approaches. The primary approach, i.e. signal to noise ratio technique is generally used in HPLC method development. Second approach used the dispersion characteristics of the regression line with the help of peak area and concentration of analyte (Roughley *et al.*, 1973). Based on this approach, standard error of peak areas for known concentrations was used for the detection and quantification of curcuminodis. Using signal to noise ratio of 3, LOD were calculated as 0.72, 0.92, 0.86 µg/ml for curcumin, demethoxycurcumin and bisdemethoxy-curcumin respectively. LOQ for curcumin were 1.42µg/ml, demethoxy-curcumin were 0.80µg/mL and bisdemethoxycurcumin were 1.04 µg/ml. Our results are in agreement with the previously reported studies. However, important variations were reported in the LOD and LOQ of these compounds. Syed *et al.* (2015) reported 0.3 µg/ml LOD of curcumin, whereas Wichitnithad *et al.* (2009) reported 0.9 µg/ml and lover LOD was repoted by Rafi *et al.,* (2015) 1.53ng/ml. LOQ of curcumin, demethoxycurcumin, bisdemethoxycurcumin were reported to be found 2.73. 2.53 and 0.23 µg/ml respectively (Wichitnithad *et al.*, 2009). Interestingly, much lower LOD values of curcumin, demethoxycurcumin, bisdemethoxycurcumin as 27.9, 31.9 and 21.8 ng/ml (Jadhav *et al.*, 2007) and 5.11, 5.47 and 4.21 ng/ml reported (Rafi *et al.,* 2015) attributed to different detectors used in the analysis.

Accuracy and precision was calculated for the individual curcuminoids after intra-day and inter-day runs at different concentration level of 1, 3 and 5µg/ml (Table 5.3 and 5.4). The intra-day accuracy ranged between 0.85 to 1.61% and Intra-day 1.00 to 2.82% so this result indicated efficiency of the method. All the data are within the standard criteria i.e. below 5% level.

CHAPTER 6

RELATIVE QUANTITATIVE EXPRESSION OF CURS GENES FROM DIFFERENT GROWTH PERIODS OF *CURCUMA* SPECIES

The age dependent relative expression of CURS genes was studied in *Curcuma* species (*C. longa, C. zanthorrhiza, C. aromatica a*nd *C. zedoaria* (*C. raktakanta*). The rhizome of all the species was harvested after 3rd, 6th and 9th months of growth and used for molecular analysis.

6.1 Total RNA isolation

Total RNA was isolated from the rhizomes of the *Curcuma* species during different growth periods (3rd, 6th and 9th months) (Fig. 6.1-6.3). The secondary metabolite (Polyphenols and alkaloids) contents were higher in the rhizomes of *Curcuma* species, hence, CTAB method failed to yield good quality RNA from the rhizomes. High-quality total RNA from *Curcuma* rhizomes was isolated using a modified SDS-Acid Phenol method and 1U DNase1 was used to remove the co-precipitated DNA. The quantity and quality of isolated RNA were assessed using a Nanodrop spectrophotometer. A 260/280 values was between 1.8 and 2.0 confirmed good quality RNA, free from proteins and phenolic compounds. EtBr-Formaldehyde agarose gel (1.2% w/v) was used to visualize 28S and 18S rRNA bands without any degradation (Fig. 6.1-6.3).

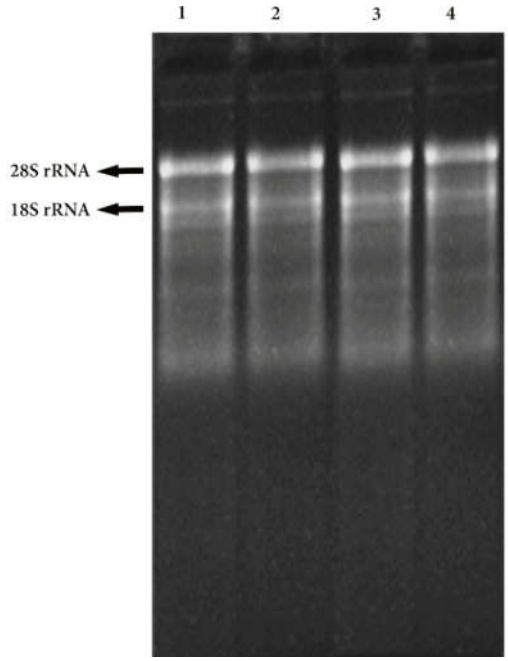


Fig. 6.1. Total RNA from three month old rhizomes of **1.** *C. longa*, **2.** *C. zanthorrhiza*, **3.** *C. aromatica* and **4.** *C. zedoaria* (*C. raktakanta*).

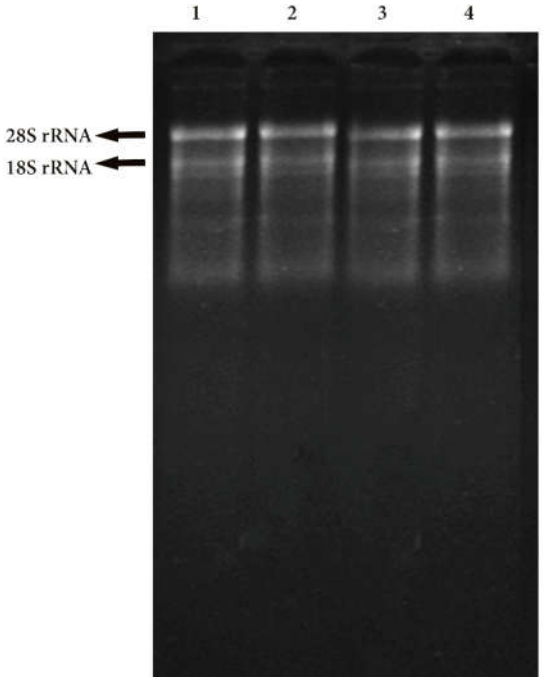


Fig. 6.2. Total RNA from six month old rhizomes of **1.** *C. longa*, **2.** *C. zanthorrhiza*, **3.** *C. aromatica* and **4.** *C. zedoaria* (*C. raktakanta*).

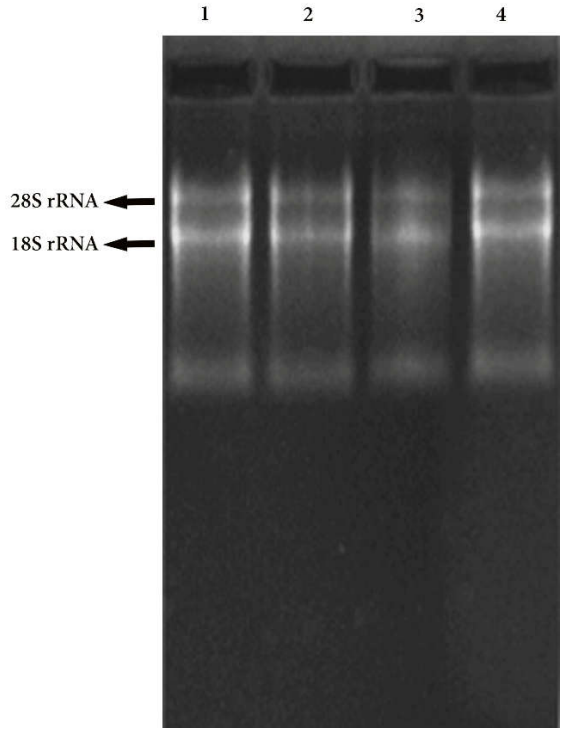


Fig. 6.3. Total RNA from nine month old rhizomes of **1.** *C. longa*, **2.** *C. zanthorrhiza*, **3**. *C. aromatica* and **4.** *C. zedoaria* (*C. raktakanta*).

6.1.1 cDNA synthesis

cDNA was synthesized using total RNA from the different aged rhizomes of (3^{rd}, 6^{th} and 9^{th} months) *Curcuma* species (Fig. 6.4-6.6). One microgram of purified total RNA was used for first-strand complementary DNA (cDNA) synthesis with oligo dTprimer and random primers using PrimeScript RT enzyme mix.

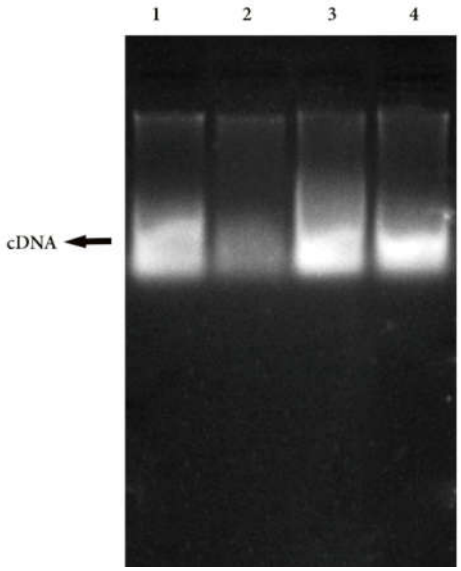


Fig. 6.4. Gel electrophoresis of the reverse transcription product of three months old rhizomes showing a smear of cDNA on 1% TAE-agarose gel. **1.** *C. longa*, **2.** *C. zanthorrhiza*, **3.** *C. aromatica* and **4.** *C. zedoaria* (*C. raktakanta*).

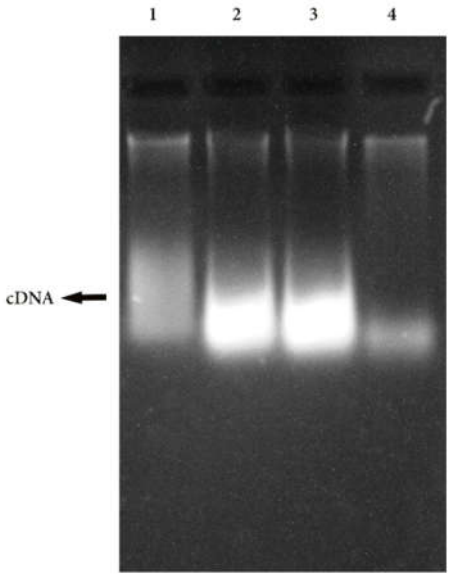


Fig. 6.5. Gel electrophoresis of reverse transcription product of six months old rhizomes showing a smear of cDNA on 1% TAE-agarose gel. **1.** *C. longa*, **2.** *C. zanthorrhiza*, **3.** *C. aromatica* and **4.** *C. zedoaria* (*C. raktakanta*).

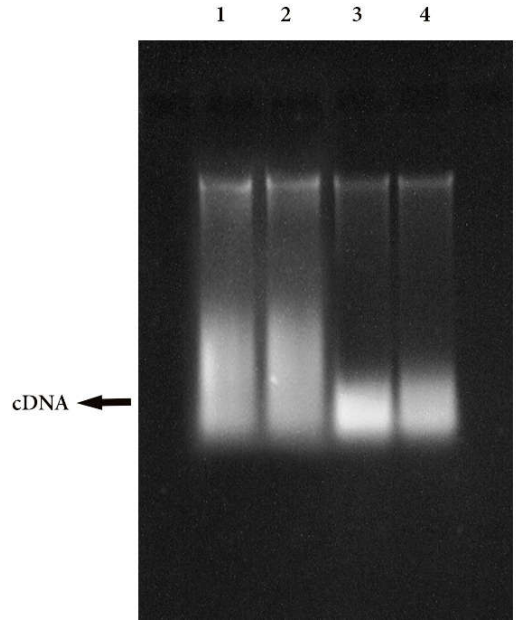


Fig. 6.6. Gel electrophoresis of reverse transcription product of nine months old rhizomes showing a smear of cDNA on 1% TAE-agarose gel. **1.** *C. longa*, **2.** *C. zanthorrhiza*, **3.** *C. aromatica* and **4.** *C. zedoaria* (*C. raktakanta*).

6.1.2 Multiple sequence alignment and Primer designing for CURS genes

CURS (CURS1, CURS2 and CURS3) genes were designed using *C. zedoaria, C. longa* and *C. amada* genes (Accession Nos: CURS1: MF402846.1, KM880189.1 and AB495007.1; CURS2: LC064068.1, AB506762.1, KF980982.1 and KF980981.1; CURS3: KX154461.1, MF987835.1, AB506763.1 and KM880190.1) retrieved from the NCBI database (http://www.ncbi.nlm.nih.gov) and used to obtain consensus sequence using Multalin with a hierarchical clustering of multiple sequence alignment software (http://www.multalin.toulouse,inra.fr/multalin/).The primers were CURS1 forward (5'- ATGGTGAAGAAGCGGTACCTG-3'); reverse (5' TGTTGCCGTACTCTGTGA AGA-3'); CURS2 forward (5'-GCTAATCAGTCAATCCAGATGG-3'); reverse (5' CGTCTATCGATTGAT CGATCGT-3') and CURS3 forward (5'-GTCAACCGCCTCAT GCTCTACA-3'); reverse (5' -TCACCTCGTCCATCACGAAGTAC-3'). The

primers did not show any hairpin loops, self-dimers or primer dimers during analysis using IDT Oligo Analyzer Tool (eu.idtdna.com/calc/analyzer).

6.1.3 Cloning of CURS genes

Curcuma cDNA having CURS genes from the selected species was cloned using Emerald AmpR GT PCR Master Mix (Cat. # RR310A Takara, Japan). The cDNA was amplified using an annealing temperature of 60.3 ºC for CURS1, 53.1 ºC for CURS2 and 57.6 ºC for CURS3. The PCR product was run using 1% (w/v) TAE-EtBr agarose gel showed a clear band of high intensity with a molecular mass of 900bp (CURS1), 1100bp (CURS2) and 590bp (CURS3) (Fig 6.7-6.9). The obtained band was eluted and sequenced using Sanger Sequencing method.

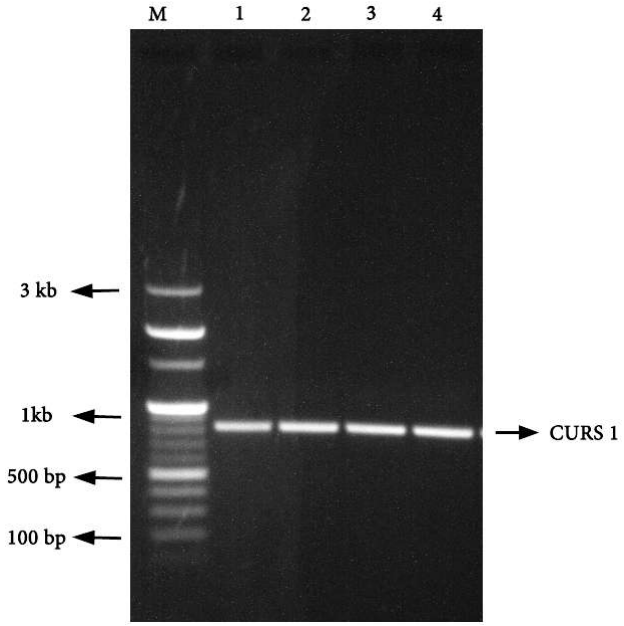


Fig. 6.7. Gel electrophoresis of RT-PCR product of rhizomes showing a band of CURS1. **M-** Marker, **1.** *C. longa*, **2**. *C. zanthorrhiza***, 3.** *C. aromatica* **4.** *C. zedoaria* (*C. raktakanta*).

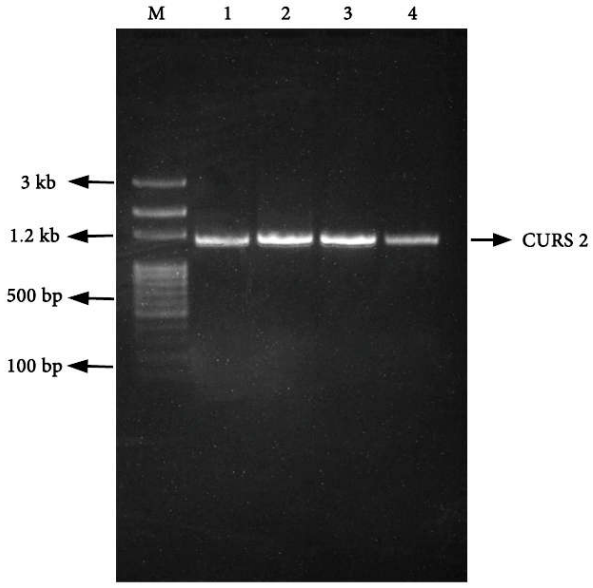


Fig. 6.8. Gel electrophoresis of RT-PCR product of rhizomes showing a band of CURS2. **M-** Marker **1.** *C. longa***, 2.** *C. zanthorrhiza***, 3.** *C. aromatica* **4.** *C. zedoaria* (*C. raktakanta*).

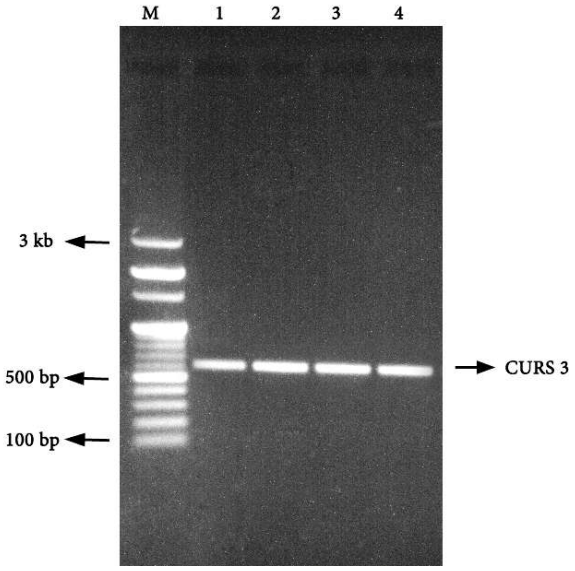


Fig. 6.9. Gel electrophoresis of RT-PCR product of rhizomes showing a band of CURS3. **M-** Marker, **1.** *C. longa*, **2.** *C. zanthorrhiza*, **3.** *C. aromatica* **4.** *C. zedoaria* (*C. raktakanta*).

The cloned gene fragment of CURS genes, both forward and reverse were merged using BioEdit software, yield CURS (CURS1, CURS2 and CURS3) genes of 400 to 990 bp developing upon the species. The sequences of partial cDNA showed homology with other CURS genes confirmed by BLAST analysis. CURS1 showed similarity with *C. zedoaria* (Acc. No. MF402846.1), CURS2 was homologous with *C. longa* (Acc. No.AB506762.1) and CURS3with *C. longa* (KM880190.1) showed 100% identity, 100% Query coverage, and an E-value of 0.0 (Table 6.1). The sequences was confirmed for homology and submitted in NCBI and accession numbers (CURS1: MK515082, MK515083, MK515084 and MK515085; CURS2: MG386668.1, MG386669.1, MG386670.1 and MG386671.1; CURS3: MK511333, MK511334, MK511335and MK511336) were provided and designated CURS genes of *Curcuma* species.

> **MK515085 *C. aromatica* curcumin synthase 1 (CURS1) mRNA, partial cds**

ATGCCCGGCGCCGACTACCGCCTCGCCACGCTCCTCGGCCTCCCTCTCACC
GTCAACCGCCTCATGATCTACAGCCAGGCCTGCCACATGGGCGCCGCCAT
GCTCCGCATCGCCAAGGACCTTGCCGAGAACAACAGGGGCGCGCGCGTGC
TGGTGGTCGCCTGCGAGATCACCGTGCTCAGCTTCCGCGGCCCGAACGAG
GGCGACTTCGAGGCGCTCGCGGGGCAGGCCGGCTTCGGCGACGGCGCGGG
GGCCGTCGTCGTCGGGGCCGACCCGCTGGAAGGAATTGAAAAGCCCATCT
ACGAGATCGCGGCGGCGATGCAGGAGACGGTGGCGGAGAGCCAGGGGGC
GGTGGGCGGCCACCTGCGGGCGTTCGGCTGGACGTTCTACTTCCTGAACCA
GCTGCCGGCGATCATCGCCGACAACCTCGGGAGGAGCCTGGAGCGGGCGT
TGGCGCCGCTGGGGGGAGGGAGTGGAACGACGTCTTCTGGGTGGCGCACC
CGGGCAACTGGGCCATCATGGACGCCATCGAAGCCAAGCTGCAGCTGAGC
CCGGACAAGCTCAGCACCGCCCGCCACGTCTTCACA

> **MK515083 *C. longa* curcumin synthase 1 (CURS1) mRNA, partial cds**

ATGCCCGGCGCCGACTACCGCCTCGCCACGCTCCTCGGCCTCCCTCTCACC
GTCAACCGCCTCATGATCTACAGCCAGGCCTGCCACATGGGCGCCGCCATG
CTCCGCATCGCCAAGGACCTCGCCGAGAACAACAGGGGCGCGCGCGTGCT
GGTGGTCGCCTGCGAGATCACCGTGCTCAGCTTCCGCGGCCCGAACGAGGG
CGACTTCGAGGCGCTCGCGGGGCAGGCCGGCTTCGGCGACGGCGCGGGGG
CCGTCGTCGTCGGGGCCGACCCGCTGGAAGGAATTGAAAAACCCATCTAC
GAGATCGCGGCGGCGATGCAGGAGACGGTGGCGGAGAGCCAGGGGGCGG
TGGGCGGCCACCTGCGGGCCTTCGGCTGGACGTTCTACTTCCTGAACCAGC
TGCCGGCGATCATCGCCGACAACCTCGGGAGGAGCCTGGAGCGGGCGTTG
GCGCCGCTGGGGGTGACGGAGTGGAACGACGTCTTCTGGGTGGCGCACCC
GGGCAACTGGGCCATCATGGACGCCATCGAAGCCAAGCTGCAGCTGAGCC
CGGACAAGCTCAGCACCGCCCGCCACGTCTTCACAGAGTAA

> **MK515082 *C. zanthorrhiza* curcumin synthase 1 (CURS1) mRNA, partial cds**

ATGCCCGGCGCCGACTACCGCCTCGCCACGCTCCTCGGCCTCCCTCTCACC
GTCAACCGCCTCATGATCTACAGCCAGGCCTGCCACATGGGCGCCGCCATG
CTCCGCATCGCCAAGGACCTCGCCGAGAACAACAGGGGCGCGCGCGTGCT
GGTGGTCGCCTGCGAGATCACCGTGCTCAGCTTCCGCGGCCCGAACGAGGG
CGACTTCGAGGCGCTCGCGGGGCAGGCCGGCTTCGGCGACGGCGCGGGGG
CCGTCGTCGTCGGGGCCGACCCGCTGGAAGGAATTGAAAAACCCATCTAC
GAGATCGCGGCGGCGATGCAGGAGACGGTGGCGGACAGCCAGGGGGCGG
TGGGCGGCCACCTGCGGGCGTTCGGCTGGACGTTCTACTTCCTGAACCAGC
TGCCGGCGATCATCGCCGACAACCTCGGGAGGAGCCTGGAGCGGGCGTTG
GCGCCGCTGGGGGTGACGGAGTGGAACGACGTCTTCTGGGTGGCGCACCC
GGGCAACTGGGCCATCATGGACGCCATCGAAGCCAAGCTGCAGCTGAGCC
CGGACAAGCTCAGCACCGCCCGCCACGTCTTCACA

> **MK515084 *C. zedoaria* curcumin synthase 1 (CURS1) mRNA, partial cds**

ATGCCCGGCGCCGACTACCGCCTCGCCACGCTCCTCGGCCTCCCTCTCACCGT
CAACCGCCTCATGATCTACAGCCAGGCCTGCCACATGGGCGCCGCCATGCTC
CGCATCGCCAAGGACCTCGCCGAGAACAACAGGGGCGCGCGCGTGCTGGTG
GTCGCCTGCGAGATCACCGTGCTCAGCTTCCGCGGCCCGAACGAGGGCGACT
TCGAGGCGCTCGCGGGGCAGGCCGGCTTCGGCGACGGCGCGGGGGCCGTCGT
CGTCGGGGCCGACCCGATGGAAGGAATTGAAAAACCCATCTACGAGATCGCG
GCGGCGATGCAGGAGACGGTGGCGGAGAGCCAGGGGGCGGTGGGCGGCCAC
CTGCGGGCGTTCGGCTGGACGTTCTACTTCCTGAACCAGCTGCCGGCGATCAT
CGCCGACAACCTCGGGAGGAGCCTGGAGCGGGCGTTGGCGCCGCTGGGGGT
GACGGAGTGGAACGACGTCTTCTGGGTGGCGCACCCGGGCAACTGGGCCATC
ATGGACGCCATCGAAGCCAAGCTGCAGCTGAGCCCGGACAAGCTCAGCACCG
CCCGCCACGTCTTCACAGAG

>**MG386669.1*C. aromatica* curcumin synthase 2 (CURS2) mRNA, partial cds**

ATGCTCTACAGCCAGGCCTGCCACATGGGCGCCGCCATGCTGCGCATAGCCA
AGGACATCGCCGAGAACAACCGCTCCGCTCGCGTCCTCGTCGTCGCCTGCGA
GATCACCGTGCTCAGCTTCCGCGGCCCGGACGAGCGCGACTTCCAGGCGCTG
GCCGGCCAGGCCGGCTTCGGGGACGGCGCCGGCGCGATGATCGTCGGGGCCG
ACCCCGTCCTCGGCGTCGAGCGGCCGCTCTACCACATCATGTCGGCGACTCA
GACGACGGTGCCGGAGAGCGAGAAGGCGGTGGGGGGCCACCTCCGCGAGGT
GGGGCTGACCTTCCACTTCTTCAACCAGCTGCCGGCGATCATCGCCGACAAC
GTGGGGAACAGCCTGGCGGAGGCGTTCGAACCGATCGGGATCAAGGACTGG
AACAACATCTTCTGGGTGGCGCACCCGGGCAACTGGGCCATCATGGACGCCA
TCGAGACCAAGCTGGGCCTGGAACAGAGCAAG

>MG386668.1 *C. longa*) curcumin synthase 2 (CURS2) mRNA, partial cds

ATGCTCTACAGCCAGGCCTGCCACATGGGCGCCGCCATGCTGCGCATAGCCA
AGGACATCGCCGAGAACAACCGCTCCGCGCGCGTCCTCGTCGTCGCCTGCGA
GATCACCGTGCTCAGCTTCCGCGGCCCGGACGAGCGCGACTTCCAGGCGCTG
GCCGGCCAGGCCGGCTTCGGGGACGGCGCCGGCGCGATGATCGTCGGGGCCG
ACCCCGTCCTCGGCGTCGAGCGGCCGCTCTACCACATCATGTCGGCGACTCA
GACGACGGTACCGGAGAGCGAGAAGGCGGTGGGGGGCCACCTCCGCGAGGT
GGGGCTGACCTTCCACTTCTTCAACCAGCTGCCGGCGATCATCGCCGACAAC
GTGGGGAACAGCCTGGCGGAGGCGTTCGAACCGATCGGGATCAAGGACTGG
AACAACATCTTCTGGGTGGCGCACCCGGGCAACTGGGCCATCATGGACGCCA
TCGAGACCAAGCTGGGCCTGGAACAGAGCAAGCTGGCCACCGCACGCCACGT
CTTCTCCGAGTTCGGCAACATGCAGAGCGCCACCGTCTACTTCGTGATGGACG
AGCTCAGGAAACGGTCGGCGGCGGAGAACCGGGCGACCACCGGCGACGGGC
TCCGGTGGGGCGTGCTCTTCGGCTTCGGCCCGGGCATCAGCATCGAAACCGT
CGTGCTCCAAAGCGTGCCGCTTTAG

> MG386670.1 *C. zanthorrhiza* curcumin synthase 2 (CURS2) mRNA, partial cds

ATGGGCGCCGCCATGCTGCGCATAGCCAAGGACATCGCCGAGAACAACCGCTCC
GCTCGCGTCCTCGTCGTCGCCTGCGAGATCACCGTGCTCAGCTTCCGCGGCCCGG
ACGAGCGCGACTTCCAGGCGCTGGCCGGCCAGGCCGGCTTCGGGGACGGCGCCG
GCGCGATGATCGTCGGGGCCGACCCCGTCCTCGGCGTCGAGCGGCCGCTCTACC
ACATCATGTCGGCGACTCAGACGACGGTGCCGGAGAGCGAGAAGGCGGTGGGG
GGCCACCTCCGCGAGGTGGGGCTGACCTTCCACTTCTTCAACCAGCTGCCGGCG
ATCATCGCCGACAACGTGGGGAACAGCCTGGCGGAGGCGTTCGAACCGATCGG
GATCAAGGACTGGAACAACATCTTCTGGGTGGCGCACCCGGCAACTGGGCCATC
ATGGACGCCATCGAGACCAAGCTGGGCCTGGAACAGAGCAAGCTGGCCACCGC
ACGCCACGTCTTCTCCGAGTTCGGCAACATGCAGGCGCCACCGTCTACTTCGTGA
TGGACGAGCTCAGGAAACGGTCGGCGGCGGGAACCGGGCGACCACCGGCGACG
GGCTCCGGTGGGGCGTGCTCTTCGGCTTCGCCCCGGCATCAGCATCGAAACCGT
CGTGCTCCAAAGCGTGCCGCTTTAG

>MG386671.1 *C. raktakanta* (*C. zedoaria*) curcumin synthase 2 (CURS2) mRNA, partial cds

ATGCTCTACAGCCAGGCCTGCCACATGGGCGCCGCCATGCTGCGCATAGCCAAG
GACATCGCCGAGAACAACCGCTCCGCTCGCGTCCTCGTCGTCGCCTGCGAGATC
ACCGTGCTCAGCTTCCGCGGCCCGGACGAGCGCGACTTCCAGGCGCTGGCCGGC
CAGGCCGGCTTCGGGGACGGCGCCGGCGCGATGATCGTCGGGGCCGACCCCGTC
CTCGGCGTCGAGCGGCCGCTCTACCACATCATGTCGGCGACTCAGACGACGGTG
CCGGAGAGCGAGAAGGCGGTGGGGGGCCACCTCCGCGAGGTGGGGCTGACCTT
CCACTTCTTCAACCAGCTGCCGGCGATCATCGCCGACAACGTGGGGAACAGCCT
GGCGGAGGCGT

> **MK511334 *C. longa* curcumin synthase 3 (CURS3) mRNA, partial cds**

ATGCTCTACAGCCAGGCCTGCCACATGGGCGCGCAGATGCTGCGCATAGCC
AAGGACCTCGCGGAGAACAACCGGGGCGCGCGCGTCCTGGCCGTCTCCTGC
GAAATCACCGTGCTCAGCTTCCGCGGTCCCGACGCGGGCGACTTCGAGGCC
CTCGCGTGTCAGGCCGGCTTCGGCGACGGTGCCGCTGCCGTCGTCGTCGGG
GCCGACCCCCTCCCGGGCGTCGAGAGGCCCATCTACGAGATCGCGGCGGCG
ATGCAGGAAACCGTGCCGGAGAGCGAGAGGGCGGTGGGGGGCCACCTGAG
GGAAATCGGCTGGACCTTCCACTTCTTCAACCAGCTGCCGAAGCTGATCGC
GGAAAACATCGAGGGCAGCCTGGCGCGGGCGTTCAAGCCGCTGGGGATAA
GCGAGTGGAACGACGTGTTCTGGGTGGCGCACCCGGGGAACTGGGGCATC
ATGGACGCCATCGAGACCAAGCTGGGGCTGGAACAGGGGAAGCTCGCCAC
GGCGCGCCACGTCTTCAGCGAGTACGGAAACATGCAGAGCGCCACCGTGT
ACTTCGTGATGGAC

> **MK511333 *C. aromatica* curcumin synthase 3 (CURS3) mRNA, partial cds**

ATGCTCTACAGCCAGGCCTGCCACATGGGCGCGCAGATGCTGCGCATAGCC
AAGGACCTCGCGGAGAACAACCGGGGCGCGCGCGTCCTGGCCGTCTCCTGC
GAAATCACCGTGCTCAGCTTCCGCGGTCCCGACGCGGGCGACTTCGAGGCC
CTGCGTGTCAGGCCGGCTTCGGCGACGGTGCCGCTGCCGTCGTCGTCGGGG
CCGACCCCCTCCCGGGCGTCGAGAGGCCCATCTACGAGATCGCGGCGGCGA
TGCAGGAAACGGTGCCGGAGAGCGAGAGGGCGGTGGGGGGCCACCTGAGG
GAAATCGGCTGGACCTTCCACTTCTTCAACCAGCTGCCGAAGCTGATCGCG
GAAAACATCGAGGGCAGCCTGGCGCGGGCGTTCAAGCCGCTGGGGATCAG
CGAGTGGAACGACGTGTTCTGGGTGGCGCACCCGGGGAACTGGGGCATCA
TGGACGCCATCGAGACCAAGCTGGGGCTGGAACAGGGGAAGCTCGCCACG
GCGCGCCACGTCTTCAGCGAGTACGGAAACATGCAGAGCGCCACCGTGTAC
TTCGTGATGGACCGAGGTGAA

> **MK511335 *C. zanthorrhiza* curcumin synthase 3 (CURS3) mRNA, partial cds**

ATGCTCTACAGCCAGGCTTGCCACATGGGCGCGCAGATGCTGCGCATA
GCCAAGGACCTCGCGGAGAACAACCGGGGCGCGCGCGTCCTGGCCGT
CTCCTGCGAAATCACCGTGCTCAGCTTCCGCGGTCCCGACGCGGGCGA
CTTCGAGGCCCTCGCGTGTCAGGCCGGCTTCGGCGACGGTGCCGCTGC
CGTCGTCGTCGGGGCCGACCCCCTCCCGGGCGTCGAGAGGCCCATCTA
CGAGATCGCGGCGGCGATGCAGGAAACCGTGCCGGAGAGCGAGAGGG
CGGTGGGGGGCCACCTGAGGGAAATCGGCTGGACCTTCCACTTCTTCA
ACCAGCTGCCGAAGCTGATCGCGGAAAACATCGAGGGCAGCCTGGCG
CGGGCGTTCAAGCCGCTGGGGATCAGCGAGTGGAACGACGTGTTCTGG
GTGGCGCACCCGGGGAACTGGGGCATCATGGACGCCATCGAGACCAA
GCTGGGGCTGGAACAGGGGAAGCTCGCCACGGCGCGCCACGTCTTCA
GCGAGTACGGAAACATGCAGAGCGCCACCGTGTACTTCGTGATGGACC
GAGGTGAA

> **MK511336 *C. zedoaria* curcumin synthase 3 (CURS3) mRNA, partial cds**

ATGCTCTACAGCCAGGCCTGCCACATGGGCGCGCAGATGCTGCGCATA
GCCAAGGACCTCGCGGAGAACAACCGGGGCGCGCGCGTCCTGGCCGT
CTCCTGCGAAATCACCGTGCTCAGCTTCCGCGGTCCCGACGCGGGCGA
CTTCGAGGCCCTCGCGTGTCAGGCCGGCTTCGGCGACGGTGCCGCTGC
CGTCGTCGTCGGGGCCGACCCCCTCCCGGGCGTCGAGAGGCCCATCTA
CGAGATCGCGGCGGCGATGCAGGAAACGGTGCCGGAGAGCGAGAGGG
CGGTGGGGGGCCACCTGAGGGAAATCGGCTGGACCTTCCACTTCTTCA
ACCAGCTGCCGAAGCTGATCGCGGAAAACATCGAGGGCAGCCTGGCG
CGGGCGTTCAAGCCGCTGGGGATCAGCGAGTGGAACGACGTGTTCTGG
GTGGCGCACCCGGGGAACTGGGGCATCATGGACGCCATCGAGACCAA
GCTCGGGCTGGAACAGGGGAAGCTCGCCACGGCGCGCCACGTCTTCA
GCGAGTACGGAAACATGCAGAGCGCCACCGTGTACTTCGTGATGGACC
GAGGT

Table 6.1: BLAST result of CURS gene sequences showing homology with other CURS gene sequences.

Plant species	CURS genes	E value	Identity (%)	GenBank Accession No.
Curcuma zedoaria	CURS1	0.0	99.75	MF402846.1
Curcuma longa	CURS1	0.0	99.25	KM880189.1
Curcuma longa	CURS1	0.0	99.25	AB495007.1
Curcuma longa	CURS2	0.0	100	AB506762.1
Curcuma zedoaria	CURS2	0.0	99	MF402847.1
Curcuma longa	CURS2	0.0	99	LO064048.1
Curcuma longa	CURS3	0.0	99.65	KM880190.1
Curcuma zedoaria	CURS3	0.0	99.30	MF987835.1
Curcuma amada	CURS3	0.0	99.30	KX154461.1

6.1.4 CURS gene specific primer designing

The cloned CURS genes sequences of all the *Curcuma* species submitted to NCBI was used to design the primers using Primer 3.0 software (http://primer3.ut.ee/). A primer pair with CURS1 forward primer (5'-AGATCATGAGGCGGTTGACG-3') and reverse primer (5'-GACATCGTGGTGGAGGAGAT-3') amplifying 199 bases were selected for the amplification CURS1 gene. The primers were 20bp in length having 67.5^{o}C and 66.1 ^{o}C Tm and the GC content of 55% produced an amplification product of 199 bases at an annealing temperature of 55.6ºC (Fig. 6.10).

CURS2 forward primer (5'-GGGATCAAGGACTGGAACAA-3') and reverse primer (5'- GTTTCCTGAGCTCGTCCATC-3') with 184 bases amplification was selected for the cloning of CURS2 gene. The forward primer was 21 bases in length having 63.9ºC Tm, 50% GC content and reverse primer 20 bases in length having a Tm of 63.8ºC, 55% GC produced an amplification product at an annealing temperature of 55.4ºC (Fig. 6.11). The primers of CURS3 (F: 5'-TGATCGCGGAAAACATCGAG-3': R: 5'-GGCTGGACCT TCCACTTCTT-3') amplified 189 bases with the forward primer 20 bases in length, Tm 64.4 ºC and 50% GC content for forward primer and 66.0 ºC Tm, 55% GC content for reverse primer and produced an amplification product at an annealing temperature of 55.1ºC (Fig. 6.12). This designed CURS forward and reverse primers were used for the relative quantification of CURS genes in all the selected *Curcuma* species.

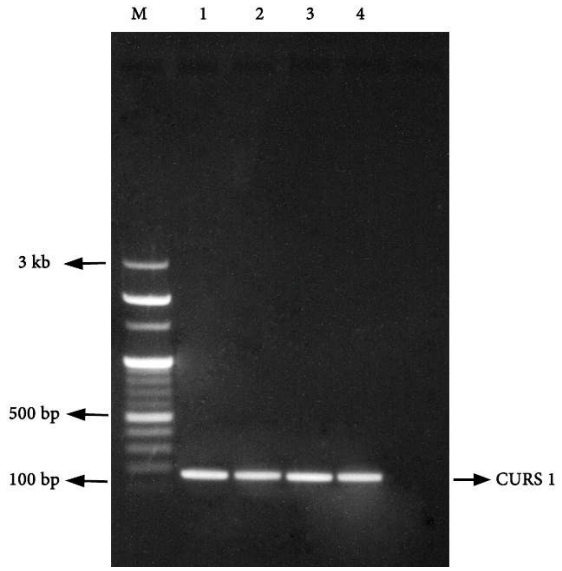


Fig. 6.10. Gel electrophoresis of the PCR product using the mRNA from the rhizomes of CURS1 with 199 bp. **M-**Marker**, 1.** *C. longa*, **2.** *C. zanthorrhiza*, **3.** *C. aromatica* **4.** *C. zedoaria* (*C. raktakanta*).

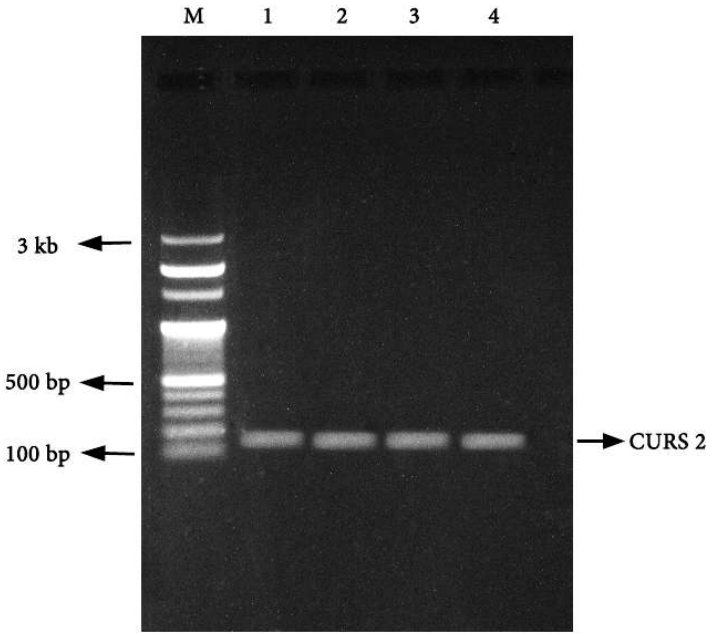


Fig. 6.11. Gel electrophoresis of the PCR product using the mRNA from the rhizomes of CURS2 with 184 bp. **M-**Marker, **1.** *C. longa*, **2.** *C. zanthorrhiza*, **3.** *C. aromatica* **4.** *C. zedoaria* (*C. raktakanta*).

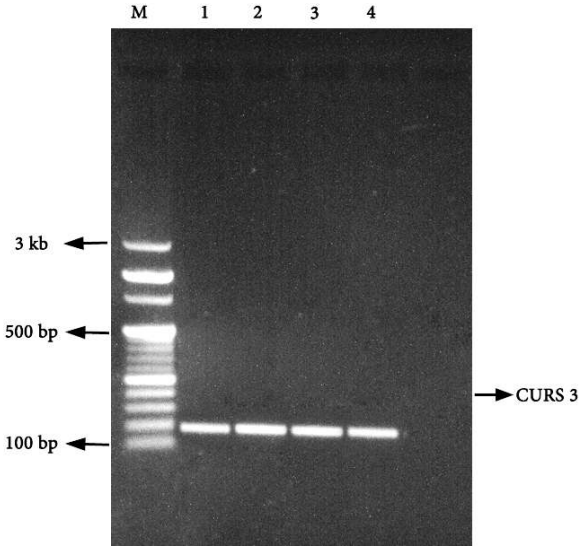


Fig. 6.12. Gel electrophoresis of the PCR product using the mRNA from the rhizomes of CURS3 with 189 bp. **M-**Marker, **1.** *C. longa*, **2.** *C. zanthorrhiza*, **3.** *C. aromatica* **4.** *C. zedoaria* (*C. raktakanta*).

6.1.5 Internal control for Real-Time PCR

Actin was used as internal control/housekeeping gene in the qRT-PCR assay, showed better expressions in all the growth conditions. The primers were designed using Multalin and Primer-BLAST softwares from NCBI. *Arabidopsis thaliana* actin sequences were selected for designing the actin primers, using the sequences available at NCBI (NM_001338359.1, NM_001338358.1, XM_021012169.1, U41998.1 and NM_112764.4). The forward primer 5'- CTGATTACCTCATGAAGATC-3'and reverse primer 5'-ATCCAGACACTGTACTTCCT-3' was designed from the consensus sequences. The forward primer was 20 bases in length with 53.5ºC Tm, 40% GC content and reverse primer was 20 bases in length, with 55.1ºC Tm, 45% GC amplified 450 bases at annealing temperature of 46.5ºC (Fig. 6.13). The sequence was confirmed and submitted to NCBI and an accession number (MG751913.1) was provided and designated actin gene of *Curcuma* species

> **MG751913.1 *Curcuma* sp. actin 1 mRNA, partial cds**

ATGAAGATCCTGACGGAGCGTGGTTACTCCTTCACCACCTCGGCTGA
GCGTGAAATCGTGCGTGACATCAAGGAGAAGCTGGCGTACATTGCCC
TGGACTTCCAGCAGGAGCTGGAGACTGCCAAGACCAGCTCCTCCATA
GAGAAGAACTACGAGCTGCCCGACGGCCAAGTCATCACCATCGGCTC
GGAGCGCTTCCGGTGCCCGGAGGTGCTTTTCCAACCATCCATGATAG
GCATGGAAGCTGCCGGCATCCACGAGACCACATACAACTCCATCATG
AAGTGCGACGTGGACATAAGGAAGGATCTGTATGGAAATATTGTGCT
CAGCGGCGGGACGACCATGTTC

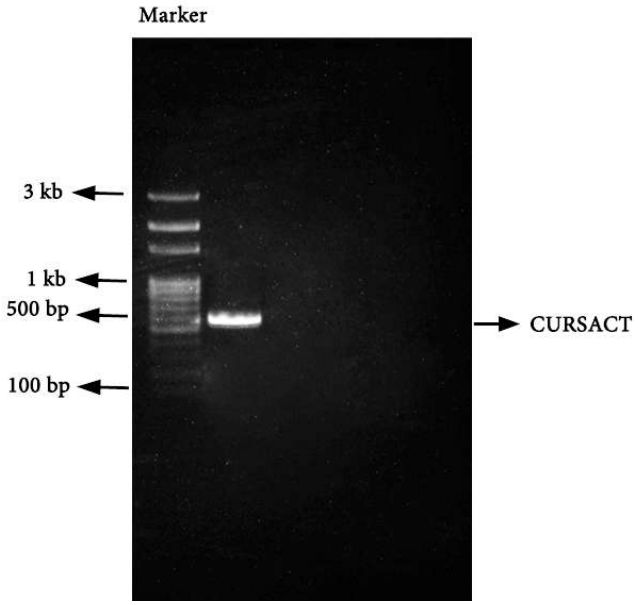


Fig. 6.13. Gel electrophoresis of the PCR product obtained rhizome of CURSACT with 450 bp**.**

6.1.6 Designing of specific actin primer

The cloned actin gene sequence from the *Curcuma* species submitted to NCBI (MG751913) was used to design specific actin primers using Primer 3.0 software (http://primer3.ut.ee/). A primer pair with forward primer (5'-GAGAAGCTGGCCTACATTGC-3') and reverse primer (5'-ATGGATGGTTGGAAAAGCAC-3') flanking a region between 73rd and 230th bases containing 157 bases was selected for the amplification of actin gene. The forward primer was 20 bases in length having 63.8ºC Tm, 55% GC content and reverse primer 20 bases in length having a Tm of 63.6ºC, 45% GC produced an amplification product at an annealing temperature of 55.4ºC (Fig. 6.14). This designed actin forward and reverse primers were used internal control for relative quantification of CURS genes relative in the selected *Curcuma* species using real-time PCR.

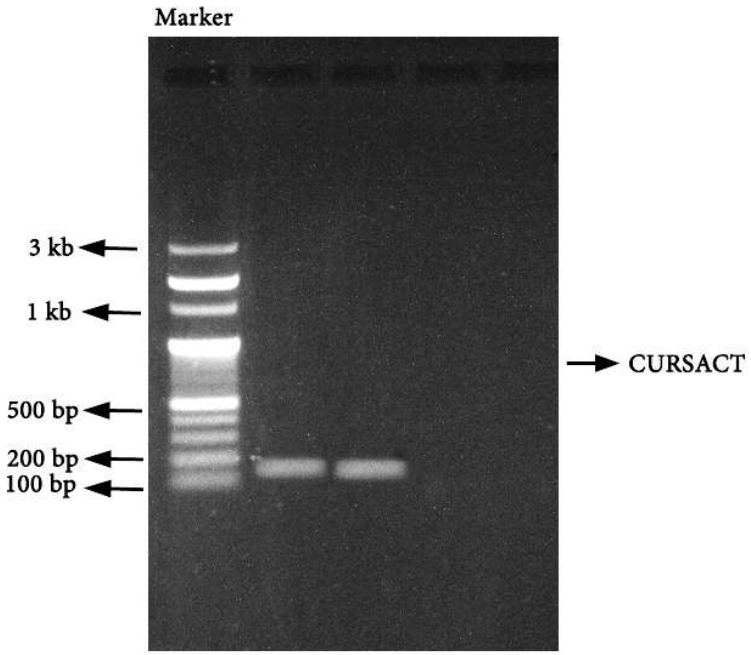


Fig. 6.14. Gel electrophoresis of the PCR product obtained rhizome of CURSACT with 157 bp.

6.1.7 Relative expression of CURS genes analysis by Real-Time PCR

Transcript levels of CURS genes from the rhizomes of *Curcuma* species were quantified using Real-Time PCR and the data obtained was analyzed by Livak's method ($2^{-\Delta\Delta C_T}$). Actin was used asinternal control and its expression in all the samples was studied. cDNA isolated from different growth periods served as a template for qRT-PCR.

6.1.8 Melting/dissociation curve

The melting/dissociation curve of the CURS genes and actin gene showed single peak indicating pure single specific amplicon in SYBR-Green assay of qRT-PCR (Fig. 6.15-6.17). The qRT-PCR was programmed to produce the melting curve. The melting curve in qRT-PCR was programmed having a cycle of 95^{o}C for 30 sec, 95^{o}C for 15 sec and after initial

amplification, the increase in temperature from 65ºC to 95ºC with a 0.5ºC increment at every 5sec led to the denaturation of double stranded DNA to single stranded DNA resulting in the decrease in fluorescence by the dissociation of SYBR Green. Each double standard DNA has a melting temperature (T_m) at which 50% of the DNA is denatured to single stranded DNA. T_m value of the actin and CURS amplicons is specific depending on its sequence length and GC content. The single melting curve peak of CURS and the internal control actin confirmed the presence of a single PCR product in the real-time analysis with T_m of 89.8^{o}C and 86.4^{o}C for CURS1, T_m of 89^{o}C and 87.5^{o}C for CURS2 and 82^{o}C and 86.2^{o}C for CURS3 respectively.

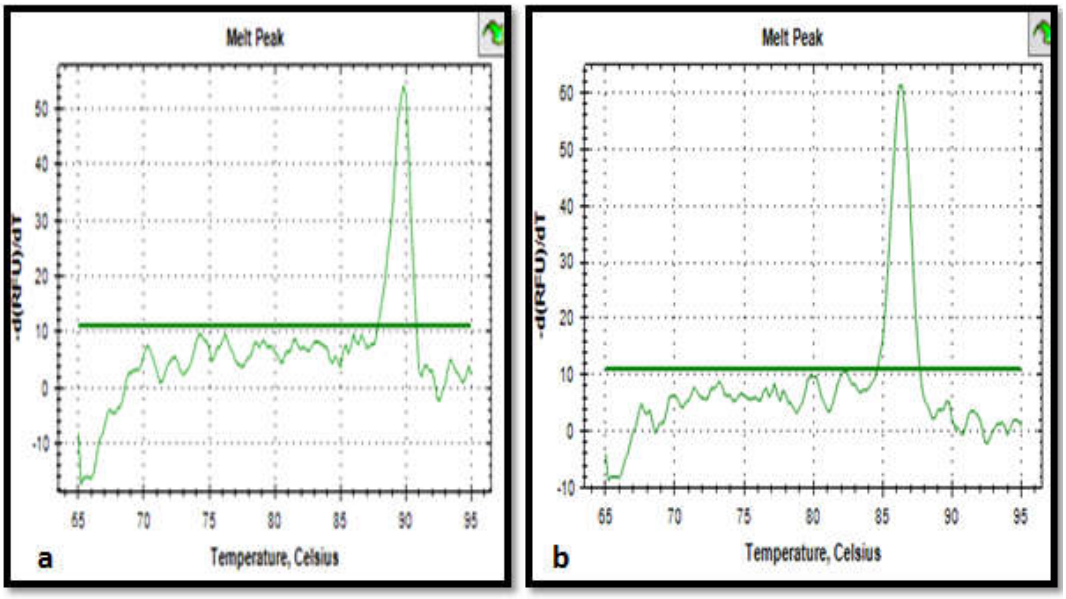


Fig. 6.15. Melting/dissociation curves of CURS1 **(a)** and CURSACT **(b)** amplicons with single peak represents amplicon specificity with a T_m of 89.8^{o}C and 86.4^{o}C respectively.

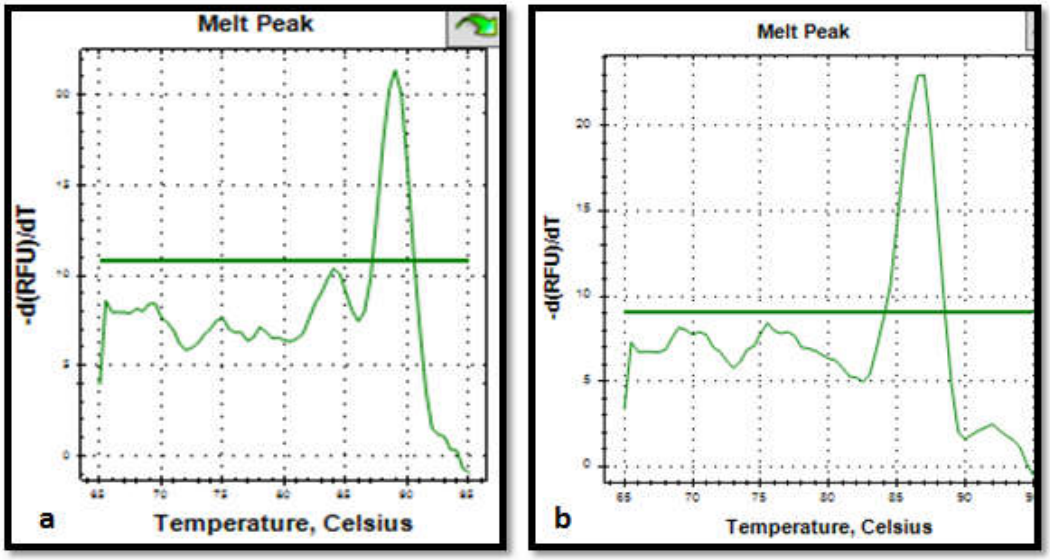


Fig. 6.16. Melting/dissociation curves of CURS2 **(a)** and CURSACT **(b)** amplicons with single peak represents amplicon specificity with a T_m of 89^oC and 87.5^oC respectively.

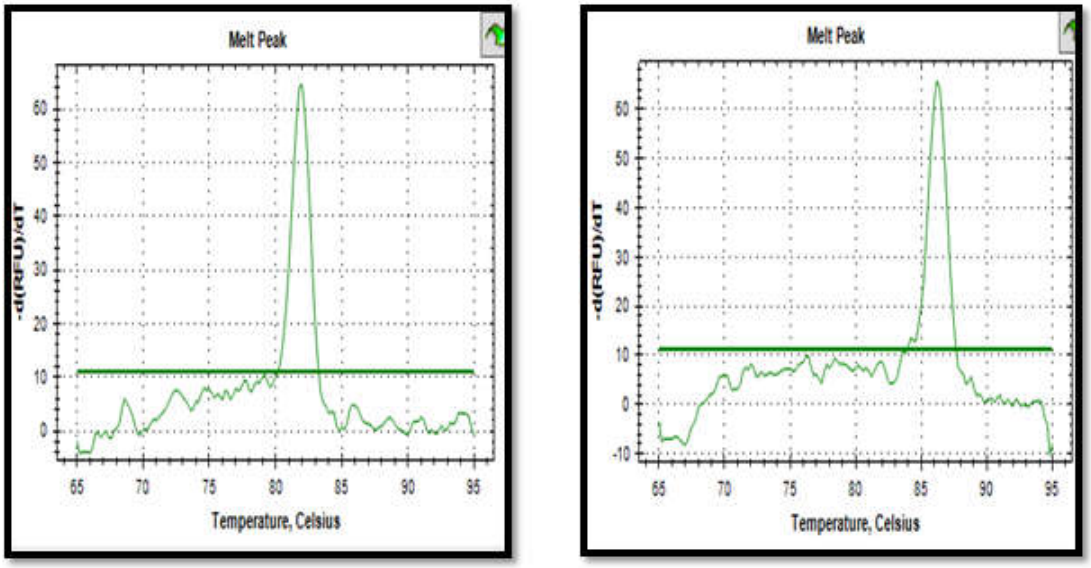


Fig. 6.17. Melting/dissociation curves of CURS3 **(a)** and CURSACT **(b)** amplicons with single peak represents amplicon specificity with a T_m of 82^oC and 86.2^oC respectively.

6.1.9 Amplification of CURS genes using qRT-PCR assay

Amplification plot of CURS genes was generated by plotting RFU (Relative Fluorescence Units- ΔRn (Fluorescent signals subtracted from the background noise level) against C_T values (Threshold cycle, the fractional cycle number at which the fluorescence passes the fixed threshold), correlated with initial concentration of cDNA for the period of exponential phase of PCR and help to distinguish the amplification signals (Fig. 6.18).The baseline of the plot indicate the initial cycles of PCR in which there is little change in fluorescence signal. The threshold level of the signals reflected a statistically significant increase over the baseline signal. The cycle number in which the fluorescent signal crosses the threshold level is represented as the Threshold cycle (C_T). The C_T value is inversely proportional to the initial levels of cDNA; lowest C_T value indicated a higher concentration of initial CURS cDNA and the samples with lowest initial cDNA concentration showed higher C_T value.

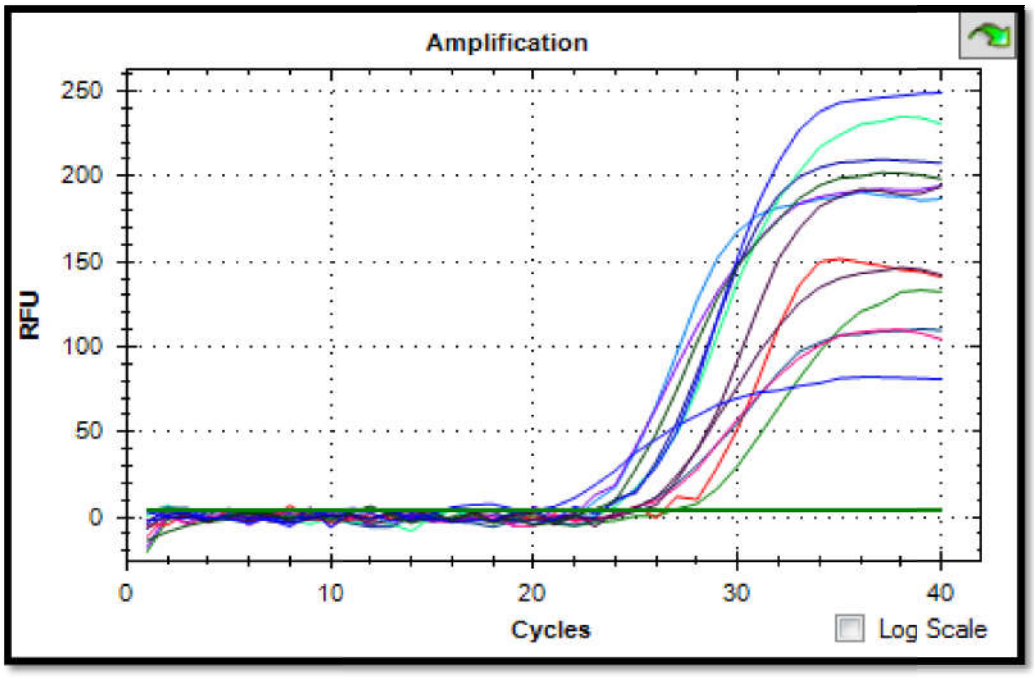


Fig. 6.18. qRT-PCR amplification plot of CURS amplicons using the cDNA from the rhizomes of different growth periods of *Curcuma* species.

6.1.10 Age dependent differential expression of CURS genes

Transcript levels of CURS (CURS1, CURS2 and CURS3) genes from the 3rd, 6th and 9 month old rhizomes were quantified using qRT-PCR and $2^{-\Delta\Delta C_T}$ method (Table 6.2-6.4). The results indicated the age-dependent differential expression of CURS1 genes from the rhizomes. The upregulation of CURS1 transcript level was observed in the six months old rhizome of *C. longa* (10.46 ± 0.11 fold) followed by six month old *C. zanthorrhiza* (6.54 ± 0.06 fold), *C aromatica* (4.01 ± 0.09 fold) and *C. zedoaria* (1.90 ± 0.06 fold). The down regulation was observed in the nine month old rhizome of *C. zedoaria* (1.13 ± 0.19 fold) followed by *C. aromatica* (1.75 ± 0.19 fold), *C. zanthorrhiza* (2.35 ± 0.04 fold) and *C longa* (4.16 ± 0.13 fold). The transcription level was lowest in the three months old rhizome of *C. zedoaria* (1.02 ± 0.14 fold) followed by *C. aromatica* (1.59 ± 0.16 fold), *C. zanthorrhiza* (2.39 ± 0.03 fold) and *C. longa* (3.58 ± 0.13 fold).

In the case of CURS2 gene the highest upregulation of CURS2 transcript level was observed in the six months old rhizome of *C. longa* (4.92 ± 0.10 fold) followed by the ninth month old rhizome (2.82 ± 0.17 fold). The expression level was lowest in the three month old rhizomes (1.51 ± 0.06 fold). In the case of *C. zanthorrhiza* the 6th month old rhizome showed higher expression (3.24 ± 0.18 fold) followed by 9th month rhizome(1.86 ± 0.05 fold). The transcription level was lowest in the three months old rhizome (1.74 ± 0.01 fold). The higher expression level was showed in 6th month old rhizomes of *C. aromatica* (2.00 ± 0.25fold) and *C. zedoaria* (*C. raktakanta*) (1.23 ± 0.10 fold) followed by 9th month (1.14 ± 0.09 and 1.14± 0.09 fold) and the lowest expression level was obtained from 3rd month of the rhizome in both cases (0.75 ± 0.18 and 0.93 ± 0.13) respectively (Table 6.3).

The least upregulation of CURS3 transcript level was observed in the six months old rhizome of *C. longa* (7.01 ± 0.14 fold) followed by three months (1.82 ± 0.05 fold). The expression level was lowest in the nine month

old rhizome (1.80 ± 0.03 fold). In *C. zanthorrhiza* the six month old rhizome showed higher expression (3.48 ± 0.13 fold) followed by 9th month (1.84 ± 0.18 fold). The transcription level was lowest in the three months old rhizome (1.61 ± 0.23 fold). The higher expression was recorded in 6th month old rhizomes of *C. aromatica* (2.17 ± 0.13 fold) and *C. zedoaria* (1.70± 0.19 fold) followed by 9th month (1.67 ± 0.19 and 1.37± 0.13 fold) and the lowest expression level was obtained from three month old rhizome in both cases (1.07 ± 0.08 and 1.20 ± 0.07) respectively (Table 6.4). The transcript analysis of the CURS genes from the rhizomes at different growth periods revealed that the highest transcript levels are observed in the vegetative growth phase during the 6th month, lesser CURS genes expression was observed during the 3rd and 9th month, compared to the 6th month (Fig 6.19-6.21). Hence, CURS genes expression CURS1 showed higher expression followed by CURS3 and CURS2.

Table 6.2: Relative expression level of CURS1 in different aged rhizomes of *Curcuma* species

Species Name	Months	C_T Mean	Δ C_T Mean	ΔΔ C_T Mean	$2^{-\Delta\Delta C_T}$ ± SE (N=3)
***C. zedoaria* (*C. raktakanta*)**	3rd	19.4	-0.13	-0.01	1.02 ± 0.14
	6th	20.0	-1.05	-0.93	1.90± 0.06
	9th	19.2	-0.24	-0.12	1.13± 0.19
C. aromatica	3rd	19.7	-0.70	-0.58	1.59 ± 0.16
	6th	16.4	-2.12	-2.0	4.01 ± 0.09
	9th	20.1	-0.93	-0.81	1.75 ± 0.19
C. zanthorrhiza	3rd	17.6	-1.38	-1.26	2.39 ± 0.03
	6th	19.7	-2.83	-2.71	6.54 ± 0.06
	9th	21.7	-1.35	-1.23	2.35 ± 0.04
C. longa	3rd	17.8	-1.95	-1.83	3.58 ± 0.13
	6th	15.9	-3.50	-3.38	10.46± 0.11
	9th	19.5	-2.17	-2.05	4.16 ± 0.13

Table 6.3: Relative expression level of CURS2 in different aged rhizomes of *Curcuma* species

Species Name	Months	C_T Mean	Δ C_T Mean	ΔΔ C_T Mean	$2^{-\Delta\Delta C_T} \pm$ SE (N=3)
***C. zedoaria* (*C. raktakanta*)**	3rd	25.5	0.4	0.1	0.93 ± 0.13
	6th	22.8	0	-0.3	1.23± 0.10
	9th	25.4	0.1	-0.2	1.14± 0.09
C. aromatica	3rd	25.7	0.7	-0.4	0.75 ± 0.18
	6th	25.4	-0.7	-1.0	2.00 ± 0.25
	9th	25.9	0.1	-0.2	1.14 ± 0.09
C. zanthorrhiza	3rd	20.9	-0.5	-0.8	1.74 ± 0.01
	6th	20.4	-1.4	-1.7	3.24 ± 0.18
	9th	23.8	-0.6	-0.9	1.86 ± 0.05
C. longa	3rd	22.9	-0.3	-0.6	1.51 ± 0.06
	6th	24.2	-2.0	-2.3	4.92 ± 0.10
	9th	23.9	-1.2	-1.5	2.82 ± 0.17

Table 6.4: Relative expression level of CURS3 in different aged rhizomes of *Curcuma* species

Species Name	Months	C_T Mean	Δ C_T Mean	ΔΔ C_T Mean	$2^{-\Delta\Delta C_T} \pm$ SE (N=3)
***C. zedoaria* (*C. raktakanta*)**	3rd	19.8	-0.29	-0.26	1.20 ± 0.07
	6th	19.6	-0.77	-0.74	1.70± 0.19
	9th	20.1	-0.48	-0.45	1.37± 0.13
C. aromatica	3rd	16.6	-0.13	-0.10	1.07 ± 0.08
	6th	18.6	-1.14	-1.11	2.17 ± 0.13
	9th	19.2	-0.75	-0.72	1.67 ± 0.19
C. zanthorrhiza	3rd	19.5	-0.69	-0.66	1.61 ± 0.23
	6th	19.8	-1.83	-1.8	3.48 ± 0.13
	9th	19.9	-0.9	-0.87	1.84 ± 0.18
C. longa	3rd	18.5	-0.89	-0.86	1.82 ± 0.05
	6th	15.4	-2.84	-2.81	7.01 ± 0.14
	9th	17.1	-0.88	-0.85	1.80 ± 0.03

Fig. 6.19. CURS1 transcript levels in the different aged rhizomes of *Curcuma* species *C. zedoaria* (*C. raktakanta*), *C. aromatica*, *C. zanthorrhiza*, *C. longa*. The relative expression level is mean ± SE, N=3.

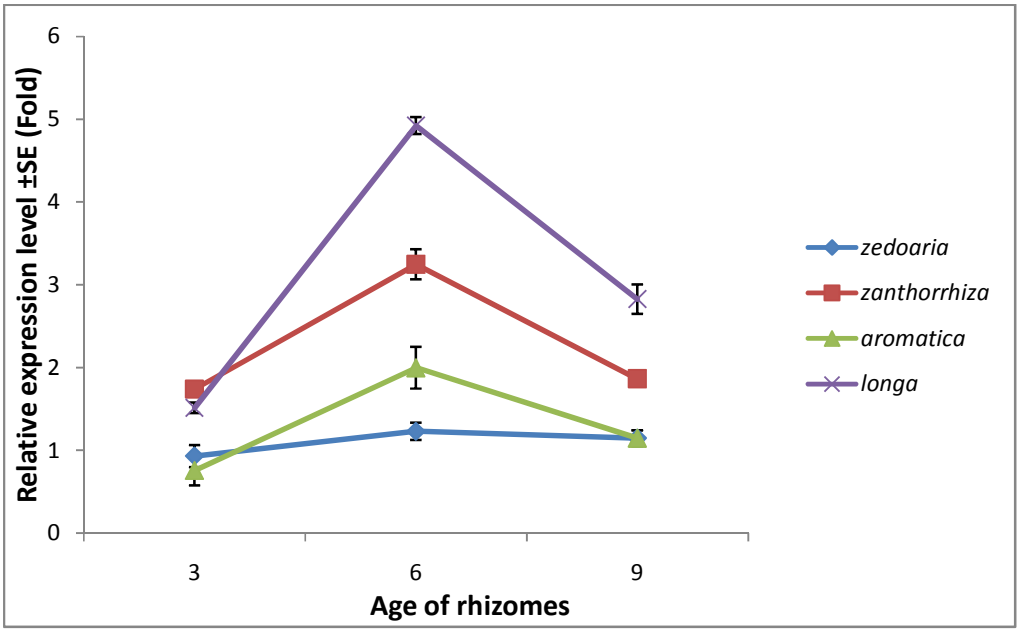


Fig. 6.20. CURS2 transcript levels in the different aged rhizomes of *Curcuma* species *C. zedoaria* (*C. raktakanta*), *C. aromatica*, *C. zanthorrhiza*, *C. longa*. The relative expression level is mean ± SE, N=3.

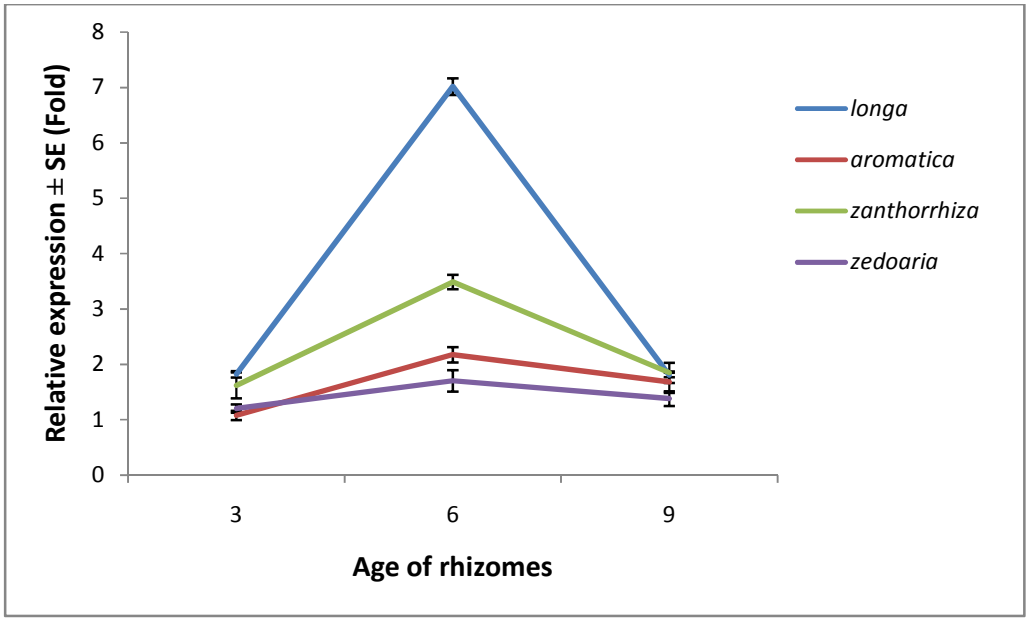


Fig. 6.21. CURS3 transcript levels in the different aged rhizomes of *Curcuma* species *C. zedoaria* (*C. raktakanta*), *C. aromatica*, *C. zanthorrhiza*, *C. longa*. The relative expression level is mean ± SE, N=3.

6.2 Discussion

6.2.1 RNA isolation

The RNA isolation from the *Curcuma* rhizomes is one of the major hurdle as the rhizomes contain polyphenols, polysaccharides and alkaloids which interfere with RNA extraction and complicate downstream applications. The Trizol method, Plant RNeasy kit (Qiagen) yielded RNA of low yield and purity from rhizomes (Deepa *et al.,* 2014). Hou *et al.* (2011) selected and modified a simple, rapid protocol and also changed the sample extraction buffer compositions to Tris-HCl, sodium chloride, SDS and β-mercaptoethanol yielded good quality and quantity of RNA. The presence of sodium chloride in the extraction buffer decreased the RNA yield but the addition of PVP and EDTA increased the yield (Deepa *et al.,* 2014).

Sodium chloride and sodium acetate play an important role in the nucleic acid extraction from polysaccharide containing tissues (Rubio-Pina and Zapata-Perez, 2011). β-mercaptoethanol inhibit RNase activity and to prevent sample oxidation (Gonzalez-Mendoza *et al.,* 2008) and SDS is widely used as a cell disrupting agent in RNA isolation (Hou *et al.,* 2011). Phenol: chloroform extraction aided in the protein precipitation (Perry and Kelley, 1972). Re-extraction of the supernatant along with sodium acetate and acid phenol: chloroform decreased the polysaccharide and protein contamination (Deepa *et al.,* 2014).

6.2.2 Relative expression of CURS genes

Secondary metabolites play a fundamental role in the development of plants grown in different environments. Several factors influence the production of secondary metabolites (SMs) in plants such as temperature, rainfall, humidity and minerals; these indirectly affect the secondary metabolite production (Paul and Kumar, 2015). The biosynthesis of SMs in the plants are dependent on environmental factors, any change in one factor, alter the content of SMs, though the other factors remain stable (Verma and Shukla, 2015) as reported in the production of curcumin in *C. longa* (Sandeep *et al.,* 2016). In the present study, the expression of CURS genes, a major polyketide family member involved in curcumin synthesis pathway in *Curcuma* species are correlated with at different growth periods.

The environmental factors always affect the gene expression in plant secondary metabolites. *Curcuma* plants from different environmental conditions were collected and planted in same environmental condition to avoid the interaction of environment in the expression of CURS genes. The results showed a variation in the quantity of curcumin content during the 3rd, 6th and 9th month of growth in the rhizomes. *Curcuma* plants exhibited better vegetative growth during the 4th and 8th months and the higher quantity of

curcumin was produced during the 6th to 8th month, without significant difference during the 8th and 10th months (Asghari *et al.,* 2009). On the contrary, our results showed the expression of higher curcumin content during the 9th month. These results indicate that 9th month is the most appropriate time to harvest the rhizome of *Curcuma* species in order to achieve maximum yield of curcumin. There are several reports that the concentrations (increase or decrease) of phytochemicals such as alkaloids (Sporer *et al.,* 1993), sesquiterpenes (Wang *et al.,* 1990) varied seasonally. Curcuminoid contents are affected by both intrinsic and extrinsic factors including seasonal variation, environmental condition, post-harvest handling, storage, microbial contamination, *etc*. (Monton *et al.,* 2016).

Upregulation of CURS genes (CURS1, CURS2 and CURS3) were observed in the 6th month-old rhizome compared to 9th month-old rhizome. It has been reported that CURS gene showed maximum expression during the 5th month compared to 10th month old rhizome (Behar *et al.,* 2016). Curcumin synthase genes are actively involved in curcuminoid synthesis during the mid period of growth, when the plants achieve peak vegetative growth and the expression level of CURS genes down regulate in the fully matured plant. The CURS genes expression is upregulated during the initial to mid vegetative stages and after the formation of curcuminoids, the mRNA synthesis declines, signifying that curcuminoid compounds synthesized in the leaves are post-translationally modified, translocated and stored in the rhizomes. CURS from *C. longa* have different functions in plants, with potential function in chalcone synthesis and curcuminoid biosynthetic pathways, respectively (Katsuyama *et al.,* 2009b). The relative expression of CURS showed maximum expression at 120 dap (Deepa *et al.,* 2017) and reports suggest that the level of curcumin is more are less stable after 180 days (Neema, 2005).

Curcumin yield increased from M0 to M6 in *curcuma* species with a parallel increase in expression of CURS gene. Our results are in agreement with the secondary metabolites production during different growth stages in *Artemisia annua* (Sandeep *et al.,* 2017; Nair *et al.,* 2013). Two genes namely, DCS (Diketide synthase) and CURS reported to have a crucial role in curcumin biosynthetic pathway in turmeric. Three curcumin synthase isoforms (CURS1, CURS2 and CURS3) are involved in the synthesis of three major curcuminoids (Katsuyama *et al.,* 2009a). The cucumin synthase genes mainly, the CURS showed maximum expression in the rhizome compared to the leaves of *Curcuma* species (Behar *et al.,* 2016).

It is evident from the above results that the expressions of CURS genes are not only tissue specific but also time specific. According to Behar *et al.,* (2016) CURS genes except CURS3 showed high level of relative expression in the rhizome, while CURS3 showed equal expression in rhizome as well as leaves. CURS3 is the only gene involved in synthesis of bis-demethoxycurcumin and the lowest expression of this gene in the rhizome and in leaves suggests either this compound is not present or present in very low amount (Katsuyama *et al*., 2009).

CHAPTER 7

IN SILICO STRUCTURE ELUCIDATION AND HOMOLOGY MODELING OF CURS (CURS1, CURS2 AND CURS3) PROTEINS

7.1 Sequence analysis of CURS proteins

The cloned CURS nucleotide sequences of *C. longa* were retrieved from NCBI with the accession numbers (MK515083, MG386668 and MK511334) and the sequences were analyzed to find out the proteins. The ORF Finder showed a putative 588bp, 675bp and 570bp open reading frame (ORF) for the cloned CURS nucleotide sequences which code the proteins with 195, 224 and 189 amino acid residues with ATG as a initiation codon (Fig. 7.1-7.3) used for the Homology modeling of CURS proteins.

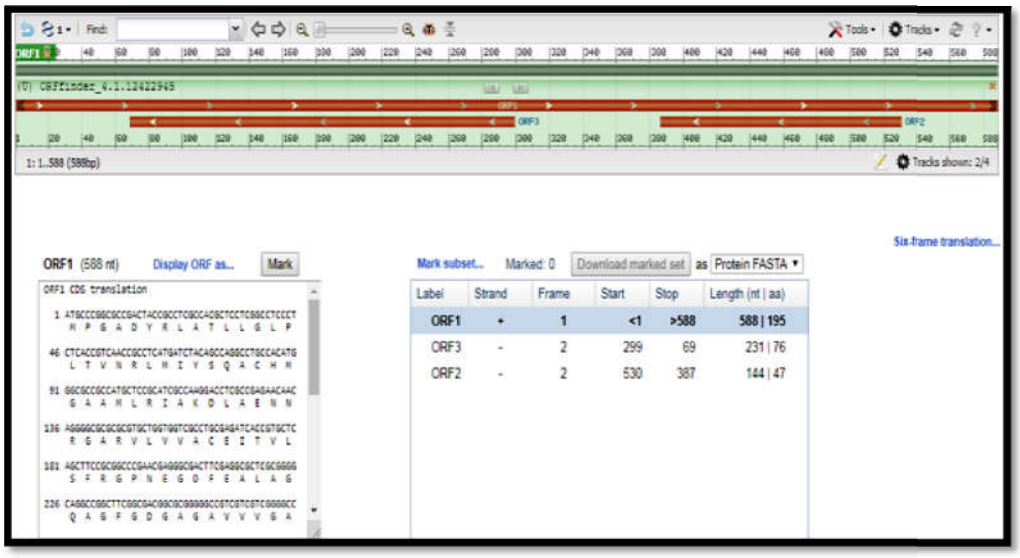


Fig. 7.1. Open Reading Frame of the cloned CURS1 gene sequence of *Curcuma longa* identified by using ORF program with 195 aminoacids

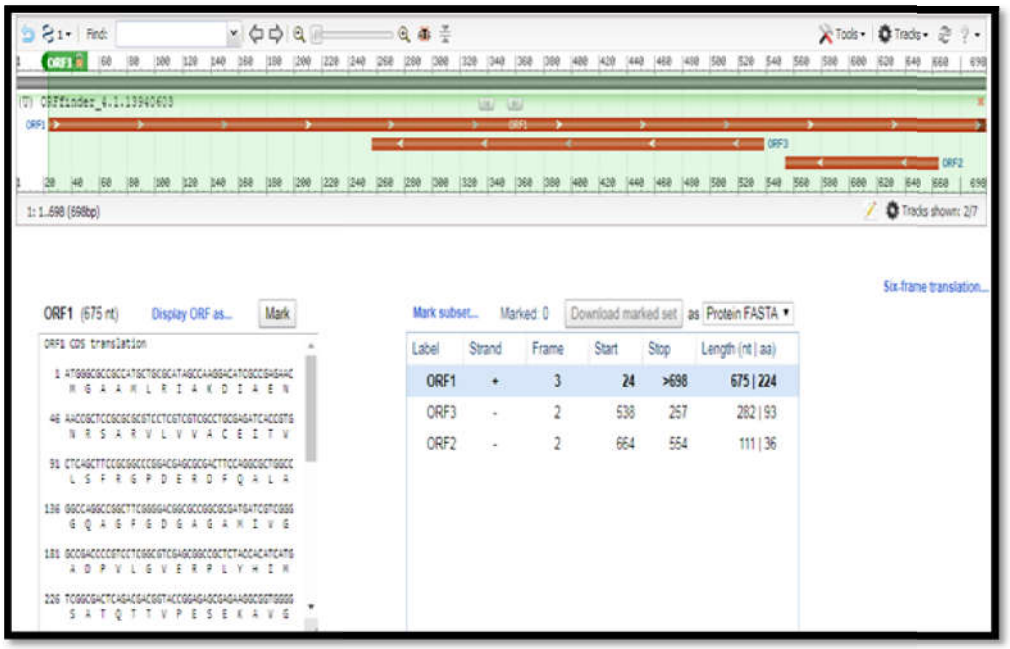


Fig. 7.2. Open Reading Frame of the cloned CURS2 gene sequence of *Curcuma longa* identified by using ORF program with 224 aminoacids

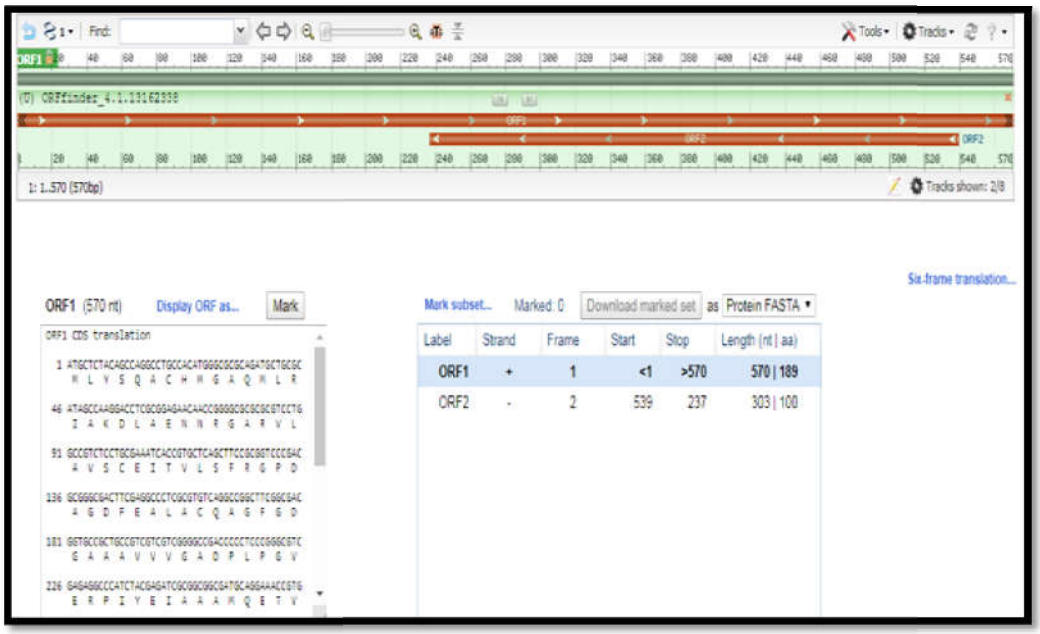


Fig. 7.3. Open Reading Frame of the cloned CURS3 gene sequence of *Curcuma longa* identified by using ORF program with 189 aminoacids

The CURS (CURS1, CURS2 and CURS3) mRNA sequences translated into protein sequences used for bioinformatics studies are:

>CURS1

MPGADYRLATLLGLPLTVNRLMIYSQACHMGAAMLRIAKDLAENNR
GARVLVVACEITVLSFRGPNEGDFEALAGQAGFGDGAGAVVVGADP
LEGIEKPIYEIAAAMQETVAESQGAVGGHLRAFGWTFYFLNQLPAIIA
DNLGRSLERALAPLGVTEWNDVFWVAHPGNWAIMDAIEAKLQLSPD
KLSTARHVFTE

>CURS2

MLYSQACHMGAAMLRIAKDIAENNRSARVLVVACEITVLSFRGPDER
DFQALAGQAGFGDGAGAMIVGADPVLGVERPLYHIMSATQTTVPESE
KAVGGHLREVGLTFHFFNQLPAIIADNVGNSLAEAFEPIGIKDWNNIF
WVAHPGNWAIMDAIETKLGLEQSKLATARHVFSEFGNMQSATVYFV
MDELRKRSAAENRATTGDGLRWGVLFGFGPGISIETVVLQSVPL

>CURS3

MLYSQACHMGAQMLRIAKDLAENNRGARVLAVSCEITVLSFRGPDA
GDFEALACQAGFGDGAAAVVVGADPLPGVERPIYEIAAAMQETVPES
ERAVGGHLREIGWTFHFFNQLPKLIAENIEGSLARAFKPLGISEWNDVF
WVAHPGNWGIMDAIETKLGLEQGKLATARHVFSEYGNMQSATVYFV
MD

7.1.1 Homology modeling of CURS proteins

The putative CURS protein sequences were analyzed using PSI-BLAST and confirmed its homology with other curcumin synthase protein sequences in the database. The search generated homologous templates were ranked according to the E-value (Fig. 7.4-7.6). The database consisted of all non-redundant GenBank CDS translations, PDB, SwissProt, PIR and PRF excluding environmental samples from WGS projects database. The CURS protein sequences were further analyzed using SMART BLAST and the output results showed homology alignment with other plant polyketides family domains (Fig. 7.7-7.9)

Fig. 7.4. PSI-BLAST analysis of CURS1 protein sequence showing homology with other CURS1 protein sequences

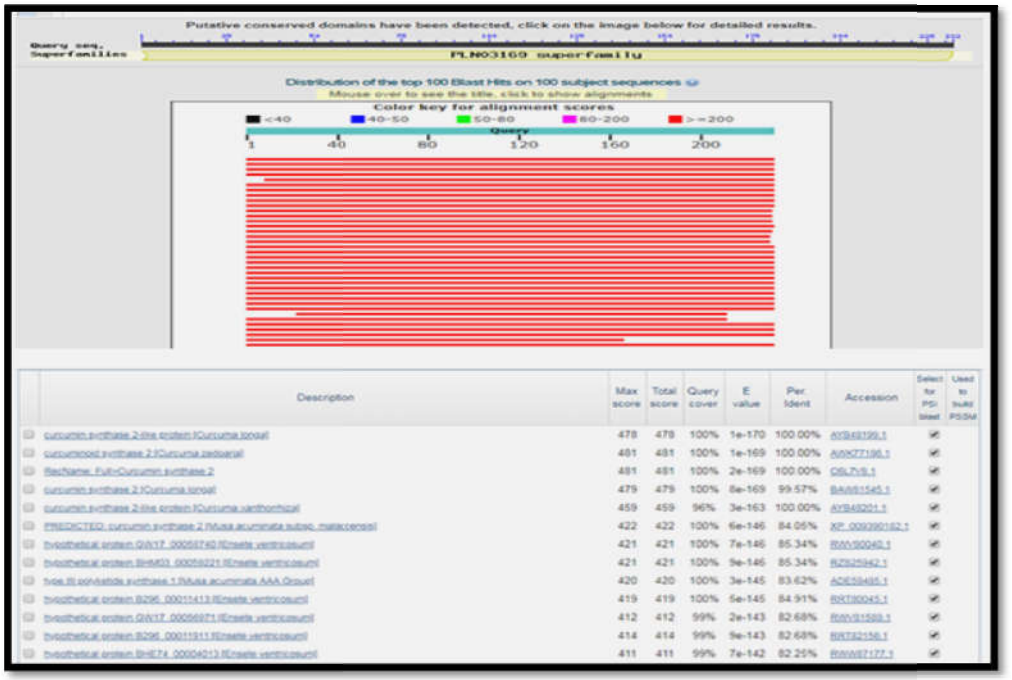


Fig. 7.5. PSI-BLAST analysis of CURS2 protein sequence showing homology with other CURS1 protein sequences

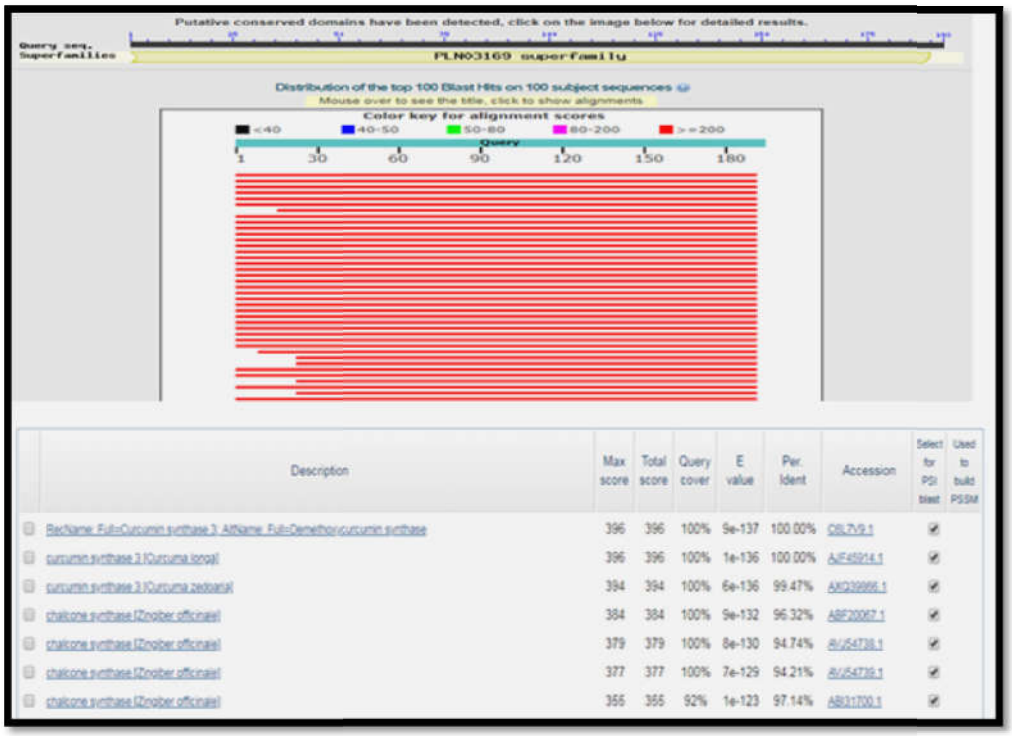


Fig. 7.6: PSI-BLAST analysis of CURS3 protein sequence showing homology with other CURS3 protein sequences

Fig. 7.7. SMART BLAST output showing alignment of CURS1 protein sequences with polyketides family DOMAIN

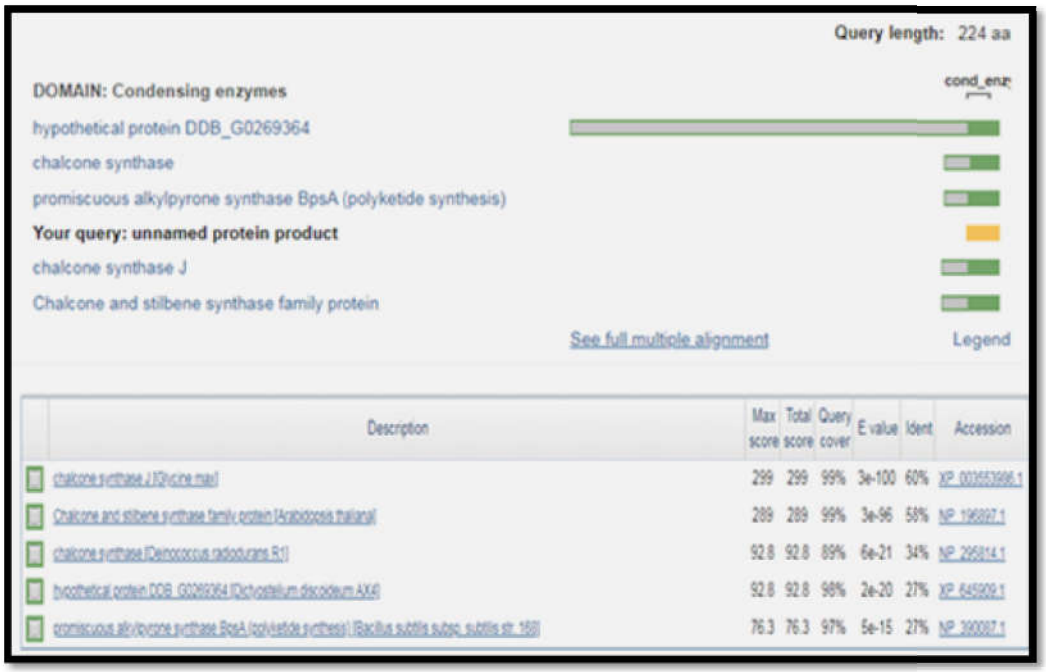


Fig. 7.8. SMART BLAST output showing alignment of CURS2 protein sequences with polyketides family DOMAIN

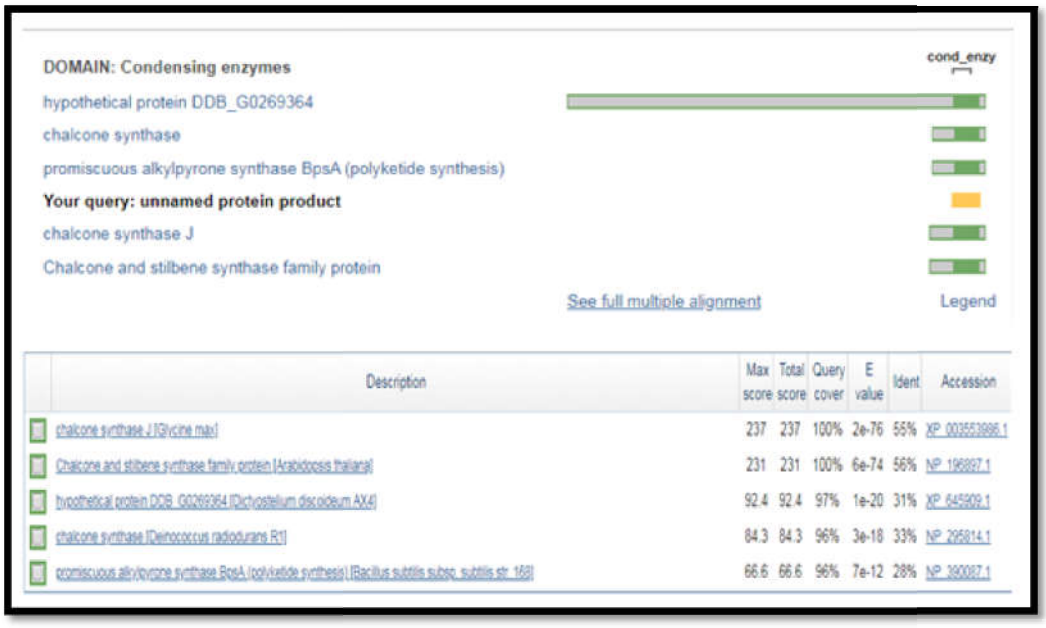
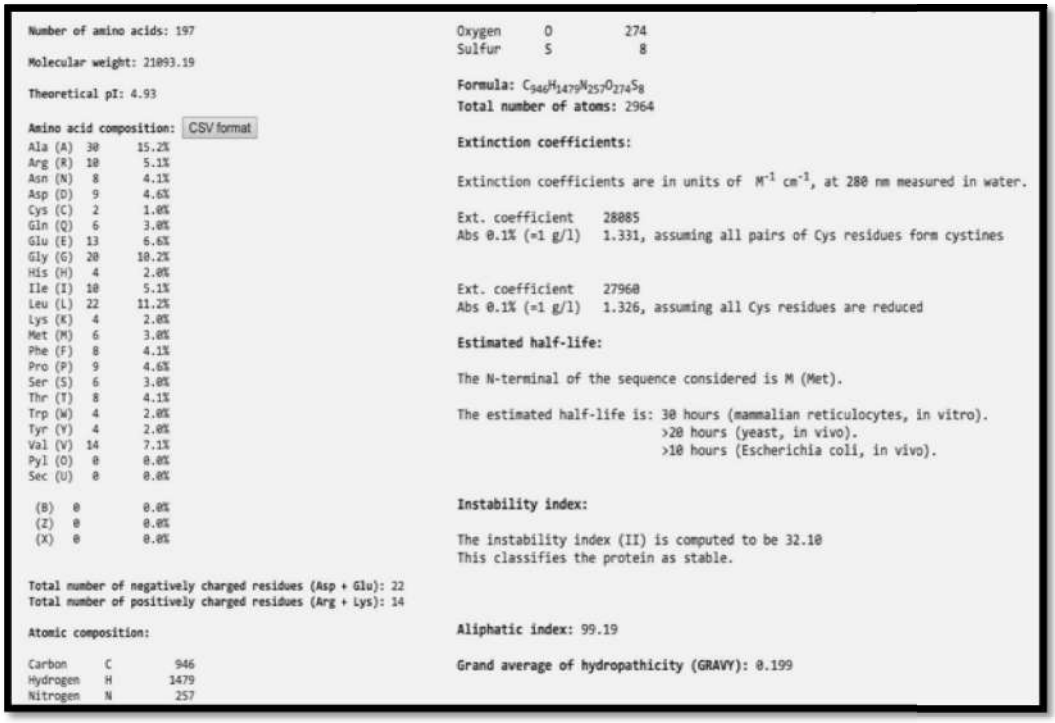


Fig. 7.9. SMART BLAST output showing alignment of CURS3 protein sequences with polyketides family DOMAIN

Physicochemical properties of CURS1, CURS2 and CURS3 proteins were examined using ExPASy ProtParam tool. The molecular weight of CURS1 is 21093.19, theoretical pI as 4.93 and an aliphatic index of 99.19. The Instability index was 32.10 and GRAVY was 0.199 (Fig. 7.10). Molecular weight of CURS2 and CURS3 proteins are 20266.13, 20629.52, theoretical pI as 5.28 and 4.96 and an aliphatic index of 89.30 and 86.37 respectively. The Instability index was 37.84 and 31.33, GRAVY was 0.118 and 0.058 for CURS2 and CURS3 (Fig. 7.11 & 7.12). Phosphorylation sites were predicted using NetPhos 2.0 server. The CURS1 protein containing 3 Ser, 2 Thr and 1 Tyr (Fig. 7.13), CURS2 having 4 Ser, 5 Thr and 2 Tyr (Fig. 7.14) and CURS3 showed 4 Ser, 1 Thr and 3 Tyr (Fig. 7.15).



Number of amino acids: 197

Molecular weight: 21093.19

Theoretical pI: 4.93

Amino acid composition: CSV format

Ala (A)	30	15.2%
Arg (R)	10	5.1%
Asn (N)	8	4.1%
Asp (D)	9	4.6%
Cys (C)	2	1.0%
Gln (Q)	6	3.0%
Glu (E)	13	6.6%
Gly (G)	20	10.2%
His (H)	4	2.0%
Ile (I)	10	5.1%
Leu (L)	22	11.2%
Lys (K)	4	2.0%
Met (M)	6	3.0%
Phe (F)	8	4.1%
Pro (P)	9	4.6%
Ser (S)	6	3.0%
Thr (T)	8	4.1%
Trp (W)	4	2.0%
Tyr (Y)	4	2.0%
Val (V)	14	7.1%
Pyl (O)	0	0.0%
Sec (U)	0	0.0%
(B)	0	0.0%
(Z)	0	0.0%
(X)	0	0.0%

Total number of negatively charged residues (Asp + Glu): 22
Total number of positively charged residues (Arg + Lys): 14

Atomic composition:

Carbon	C	946
Hydrogen	H	1479
Nitrogen	N	257
Oxygen	O	274
Sulfur	S	8

Formula: $C_{946}H_{1479}N_{257}O_{274}S_8$
Total number of atoms: 2964

Extinction coefficients:

Extinction coefficients are in units of M^{-1} cm^{-1}, at 280 nm measured in water.

Ext. coefficient 28085
Abs 0.1% (=1 g/l) 1.331, assuming all pairs of Cys residues form cystines

Ext. coefficient 27960
Abs 0.1% (=1 g/l) 1.326, assuming all Cys residues are reduced

Estimated half-life:

The N-terminal of the sequence considered is M (Met).

The estimated half-life is: 30 hours (mammalian reticulocytes, in vitro).
>20 hours (yeast, in vivo).
>10 hours (Escherichia coli, in vivo).

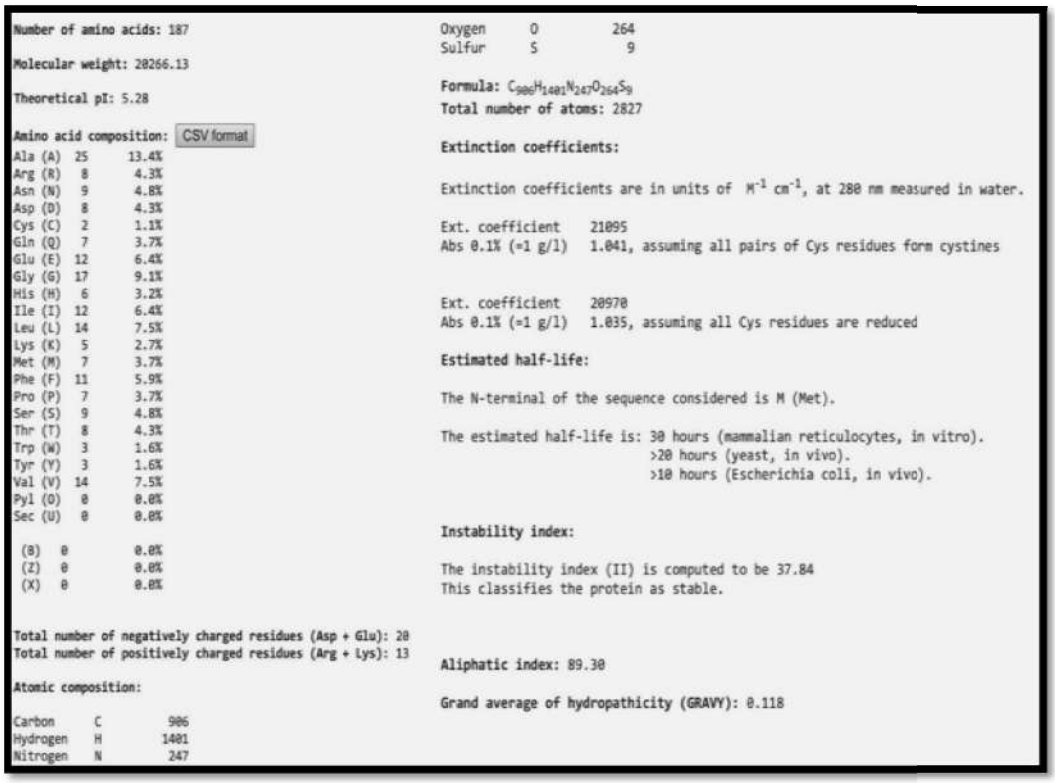
Instability index:

The instability index (II) is computed to be 32.10
This classifies the protein as stable.

Aliphatic index: 99.19

Grand average of hydropathicity (GRAVY): 0.199

Fig. 7.10. ExPASy ProtParam result of CURS1 protein



Number of amino acids: 187

Molecular weight: 20266.13

Theoretical pI: 5.28

Amino acid composition: CSV format

Ala (A)	25	13.4%
Arg (R)	8	4.3%
Asn (N)	9	4.8%
Asp (D)	8	4.3%
Cys (C)	2	1.1%
Gln (Q)	7	3.7%
Glu (E)	12	6.4%
Gly (G)	17	9.1%
His (H)	6	3.2%
Ile (I)	12	6.4%
Leu (L)	14	7.5%
Lys (K)	5	2.7%
Met (M)	7	3.7%
Phe (F)	11	5.9%
Pro (P)	7	3.7%
Ser (S)	9	4.8%
Thr (T)	8	4.3%
Trp (W)	3	1.6%
Tyr (Y)	3	1.6%
Val (V)	14	7.5%
Pyl (O)	0	0.0%
Sec (U)	0	0.0%
(B)	0	0.0%
(Z)	0	0.0%
(X)	0	0.0%

Total number of negatively charged residues (Asp + Glu): 20
Total number of positively charged residues (Arg + Lys): 13

Atomic composition:

Carbon	C	906
Hydrogen	H	1401
Nitrogen	N	247
Oxygen	O	264
Sulfur	S	9

Formula: $C_{906}H_{1401}N_{247}O_{264}S_9$
Total number of atoms: 2827

Extinction coefficients:

Extinction coefficients are in units of M^{-1} cm^{-1}, at 280 nm measured in water.

Ext. coefficient 21095
Abs 0.1% (=1 g/l) 1.041, assuming all pairs of Cys residues form cystines

Ext. coefficient 20970
Abs 0.1% (=1 g/l) 1.035, assuming all Cys residues are reduced

Estimated half-life:

The N-terminal of the sequence considered is M (Met).

The estimated half-life is: 30 hours (mammalian reticulocytes, in vitro).
>20 hours (yeast, in vivo).
>10 hours (Escherichia coli, in vivo).

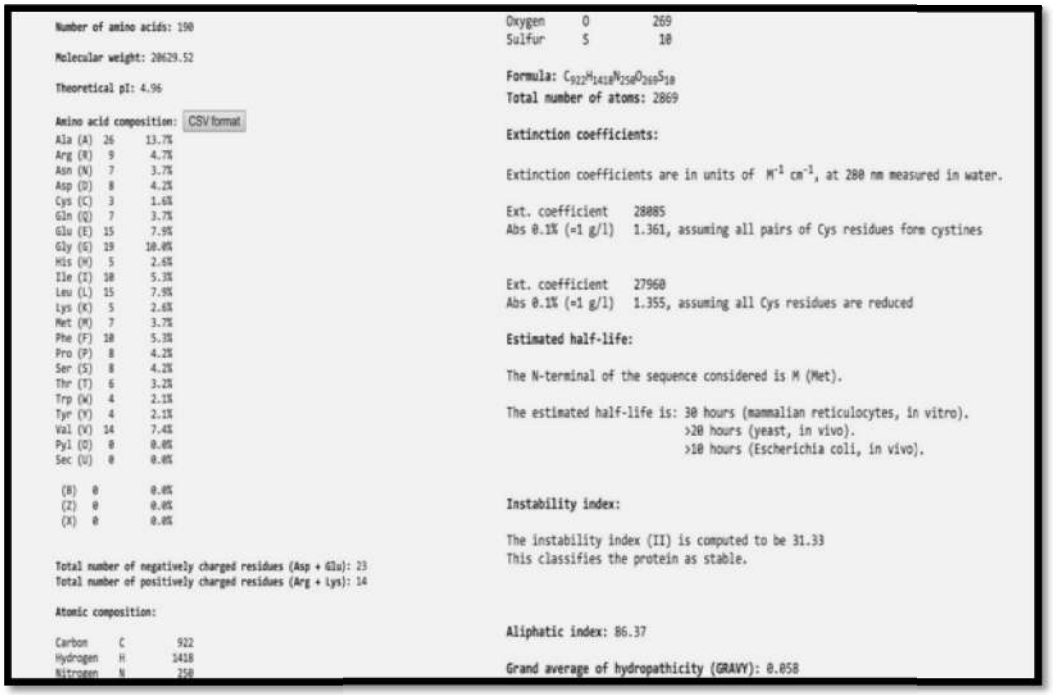
Instability index:

The instability index (II) is computed to be 37.84
This classifies the protein as stable.

Aliphatic index: 89.30

Grand average of hydropathicity (GRAVY): 0.118

Fig. 7.11. ExPASy ProtParam result of CURS2 protein



Number of amino acids: 190

Molecular weight: 20629.52

Theoretical pI: 4.96

Amino acid composition: CSV format

Ala (A)	26	13.7%
Arg (R)	9	4.7%
Asn (N)	7	3.7%
Asp (D)	8	4.2%
Cys (C)	3	1.6%
Gln (Q)	7	3.7%
Glu (E)	15	7.9%
Gly (G)	19	10.0%
His (H)	5	2.6%
Ile (I)	10	5.3%
Leu (L)	15	7.9%
Lys (K)	5	2.6%
Met (M)	7	3.7%
Phe (F)	10	5.3%
Pro (P)	8	4.2%
Ser (S)	8	4.2%
Thr (T)	6	3.2%
Trp (W)	4	2.1%
Tyr (Y)	4	2.1%
Val (V)	14	7.4%
Pyl (O)	0	0.0%
Sec (U)	0	0.0%
(B)	0	0.0%
(Z)	0	0.0%
(X)	0	0.0%

Total number of negatively charged residues (Asp + Glu): 23
Total number of positively charged residues (Arg + Lys): 14

Atomic composition:

Carbon	C	922
Hydrogen	H	1418
Nitrogen	N	250
Oxygen	O	269
Sulfur	S	10

Formula: $C_{922}H_{1418}N_{250}O_{269}S_{10}$
Total number of atoms: 2869

Extinction coefficients:

Extinction coefficients are in units of M^{-1} cm^{-1}, at 280 nm measured in water.

Ext. coefficient 28085
Abs 0.1% (=1 g/l) 1.361, assuming all pairs of Cys residues form cystines

Ext. coefficient 27960
Abs 0.1% (=1 g/l) 1.355, assuming all Cys residues are reduced

Estimated half-life:

The N-terminal of the sequence considered is M (Met).

The estimated half-life is: 30 hours (mammalian reticulocytes, in vitro).
>20 hours (yeast, in vivo).
>10 hours (Escherichia coli, in vivo).

Instability index:

The instability index (II) is computed to be 31.33
This classifies the protein as stable.

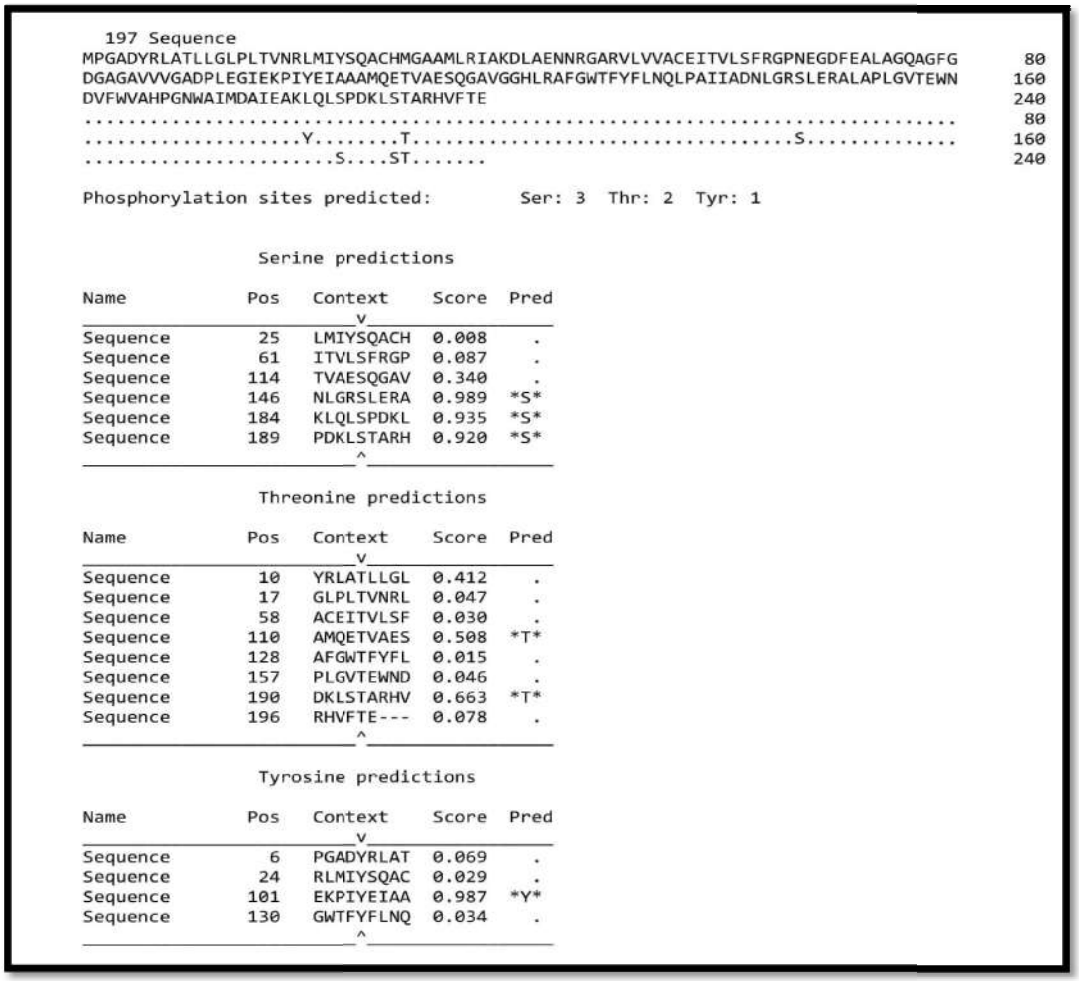
Aliphatic index: 86.37

Grand average of hydropathicity (GRAVY): 0.058

Fig. 7.12. ExPASy ProtParam result of CURS3 protein

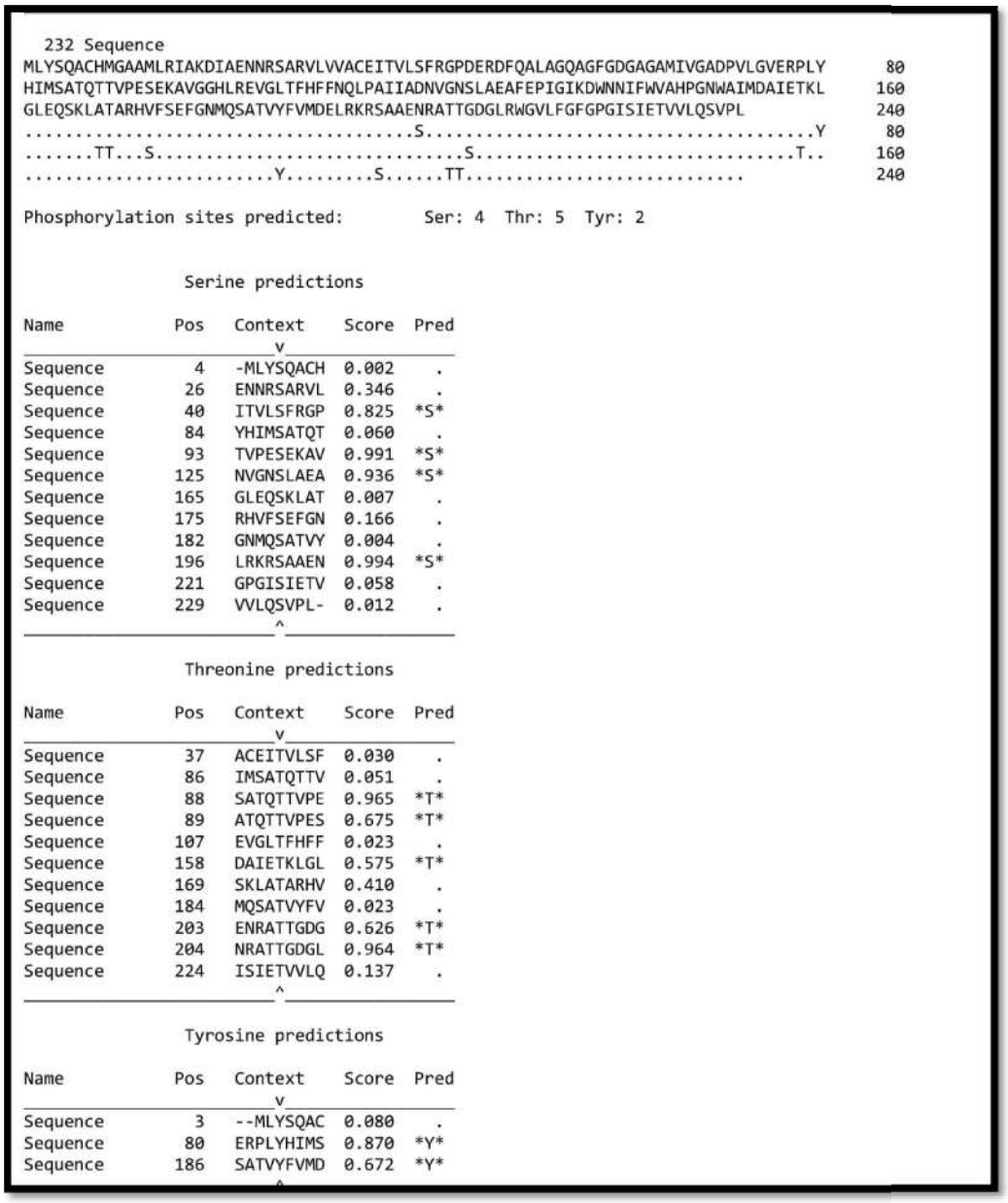
7.1.2 Secondary structure prediction

Consensus secondary structure was predicted using SOPMA server revealed that CURS1 protein has 47.72% alpha helix, 24.87% random coil, 16.24% extended strand and 0% of ambiguous state (Fig. 7.16). CURS2 has 41.38% alpha helix, 31.03% random coil, 19.40% extended strand and 0% of ambiguous state (Fig.7.17). While CURS3 protein has 44.74% alpha helix, 27.89% random coil, 17.89% extended strand and 0% of ambiguous state (Fig.7.18). Further the protein secondary structure was analyzed by using PSIPRED server (Fig.7.20-7.22).



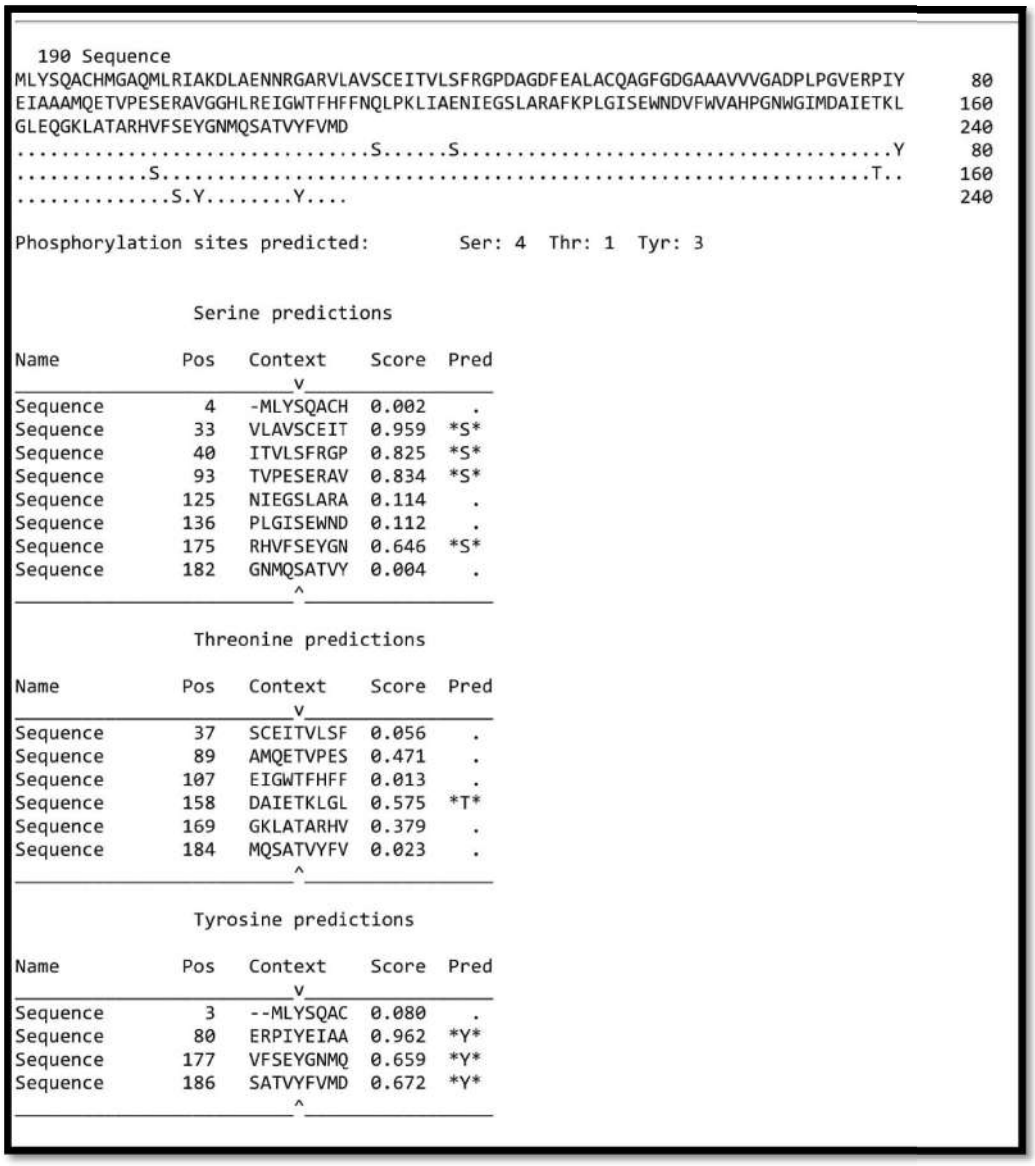
```
  197 Sequence
MPGADYRLATLLGLPLTVNRLMIYSQACHMGAAMLRIAKDLAENNRGARVLVVACEITVLSFRGPNEGDFEALAGQAGFG      80
DGAGAVVVGADPLEGIEKPIYEIAAAMQETVAESQGAVGGHLRAFGWTFYFLNQLPAIIADNLGRSLERALAPLGVTEWN     160
DVFWVAHPGNWAIMDAIEAKLQLSPDKLSTARHVFTE                                                  240
.....                                                                                   80
.....Y.....T.....S.....                                                                160
.....S....ST.....                                                                      240

Phosphorylation sites predicted:        Ser: 3  Thr: 2  Tyr: 1


                Serine predictions

Name            Pos   Context    Score  Pred
                           v
Sequence         25   LMIYSQACH  0.008    .
Sequence         61   ITVLSFRGP  0.087    .
Sequence        114   TVAESQGAV  0.340    .
Sequence        146   NLGRSLERA  0.989  *S*
Sequence        184   KLQLSPDKL  0.935  *S*
Sequence        189   PDKLSTARH  0.920  *S*
                           ^

                Threonine predictions

Name            Pos   Context    Score  Pred
                           v
Sequence         10   YRLATLLGL  0.412    .
Sequence         17   GLPLTVNRL  0.047    .
Sequence         58   ACEITVLSF  0.030    .
Sequence        110   AMQETVAES  0.508  *T*
Sequence        128   AFGWTFYFL  0.015    .
Sequence        157   PLGVTEWND  0.046    .
Sequence        190   DKLSTARHV  0.663  *T*
Sequence        196   RHVFTE---  0.078    .
                           ^

                Tyrosine predictions

Name            Pos   Context    Score  Pred
                           v
Sequence          6   PGADYRLAT  0.069    .
Sequence         24   RLMIYSQAC  0.029    .
Sequence        101   EKPIYEIAA  0.987  *Y*
Sequence        130   GWTFYFLNQ  0.034    .
                           ^
```

Fig. 7.13. NetPhos 2.0 result of Phosphorylation sites of CURS1 protein



```
  232 Sequence
MLYSQACHMGAAMLRIAKDIAENNRSARVLVVACEITVLSFRGPDERDFQALAGQAGFGDGAGAMIVGADPVLGVERPLY      80
HIMSATQTTVPESEKAVGGHLREVGLTFHFFNQLPAIIADNVGNSLAEAFEPIGIKDWNNIFWVAHPGNWAIMDAIETKL     160
GLEQSKLATARHVFSEFGNMQSATVYFVMDELRKRSAAENRATTGDGLRWGVLFGFGPGISIETVVLQSVPL              240
.........................................S.......................................Y      80
.......TT...S..............................S..............................T..     160
.........................Y.........S......TT............................              240

Phosphorylation sites predicted:        Ser: 4  Thr: 5  Tyr: 2


                Serine predictions

Name            Pos    Context    Score  Pred
                             v
Sequence          4    -MLYSQACH  0.002    .
Sequence         26    ENNRSARVL  0.346    .
Sequence         40    ITVLSFRGP  0.825  *S*
Sequence         84    YHIMSATQT  0.060    .
Sequence         93    TVPESEKAV  0.991  *S*
Sequence        125    NVGNSLAEA  0.936  *S*
Sequence        165    GLEQSKLAT  0.007    .
Sequence        175    RHVFSEFGN  0.166    .
Sequence        182    GNMQSATVY  0.004    .
Sequence        196    LRKRSAAEN  0.994  *S*
Sequence        221    GPGISIETV  0.058    .
Sequence        229    VVLQSVPL-  0.012    .
                             ^

                Threonine predictions

Name            Pos    Context    Score  Pred
                             v
Sequence         37    ACEITVLSF  0.030    .
Sequence         86    IMSATQTTV  0.051    .
Sequence         88    SATQTTVPE  0.965  *T*
Sequence         89    ATQTTVPES  0.675  *T*
Sequence        107    EVGLTFHFF  0.023    .
Sequence        158    DAIETKLGL  0.575  *T*
Sequence        169    SKLATARHV  0.410    .
Sequence        184    MQSATVYFV  0.023    .
Sequence        203    ENRATTGDG  0.626  *T*
Sequence        204    NRATTGDGL  0.964  *T*
Sequence        224    ISIETVVLQ  0.137    .
                             ^

                Tyrosine predictions

Name            Pos    Context    Score  Pred
                             v
Sequence          3    --MLYSQAC  0.080    .
Sequence         80    ERPLYHIMS  0.870  *Y*
Sequence        186    SATVYFVMD  0.672  *Y*
                             ^
```

Fig. 7.14. NetPhos 2.0 result of Phosphorylation sites of CURS2 protein



```
  190 Sequence
MLYSQACHMGAQMLRIAKDLAENNRGARVLAVSCEITVLSFRGPDAGDFEALACQAGFGDGAAAVVVGADPLPGVERPIY      80
EIAAAMQETVPESERAVGGHLREIGWTFHFFNQLPKLIAENIEGSLARAFKPLGISEWNDVFWVAHPGNWGIMDAIETKL     160
GLEQGKLATARHVFSEYGNMQSATVYFVMD                                                        240
.............................S......S.....................................Y      80
............S.............................................................T..     160
..............S.Y........Y....                                                        240

Phosphorylation sites predicted:        Ser: 4  Thr: 1  Tyr: 3


                Serine predictions

Name            Pos   Context    Score  Pred
                           v
Sequence          4   -MLYSQACH  0.002    .
Sequence         33   VLAVSCEIT  0.959  *S*
Sequence         40   ITVLSFRGP  0.825  *S*
Sequence         93   TVPESERAV  0.834  *S*
Sequence        125   NIEGSLARA  0.114    .
Sequence        136   PLGISEWND  0.112    .
Sequence        175   RHVFSEYGN  0.646  *S*
Sequence        182   GNMQSATVY  0.004    .
                           ^

                Threonine predictions

Name            Pos   Context    Score  Pred
                           v
Sequence         37   SCEITVLSF  0.056    .
Sequence         89   AMQETVPES  0.471    .
Sequence        107   EIGWTFHFF  0.013    .
Sequence        158   DAIETKLGL  0.575  *T*
Sequence        169   GKLATARHV  0.379    .
Sequence        184   MQSATVYFV  0.023    .
                           ^

                Tyrosine predictions

Name            Pos   Context    Score  Pred
                           v
Sequence          3   --MLYSQAC  0.080    .
Sequence         80   ERPIYEIAA  0.962  *Y*
Sequence        177   VFSEYGNMQ  0.659  *Y*
Sequence        186   SATVYFVMD  0.672  *Y*
                           ^
```

Fig. 7.15. NetPhos 2.0 result of Phosphorylation sites of CURS3 protein

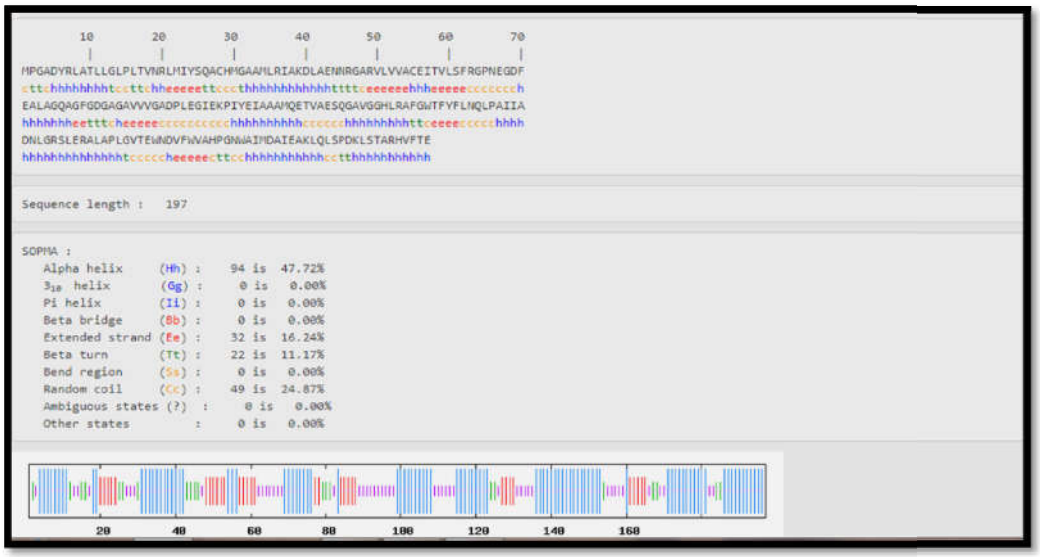


Fig. 7.16. SOPMA result for CURS1 protein

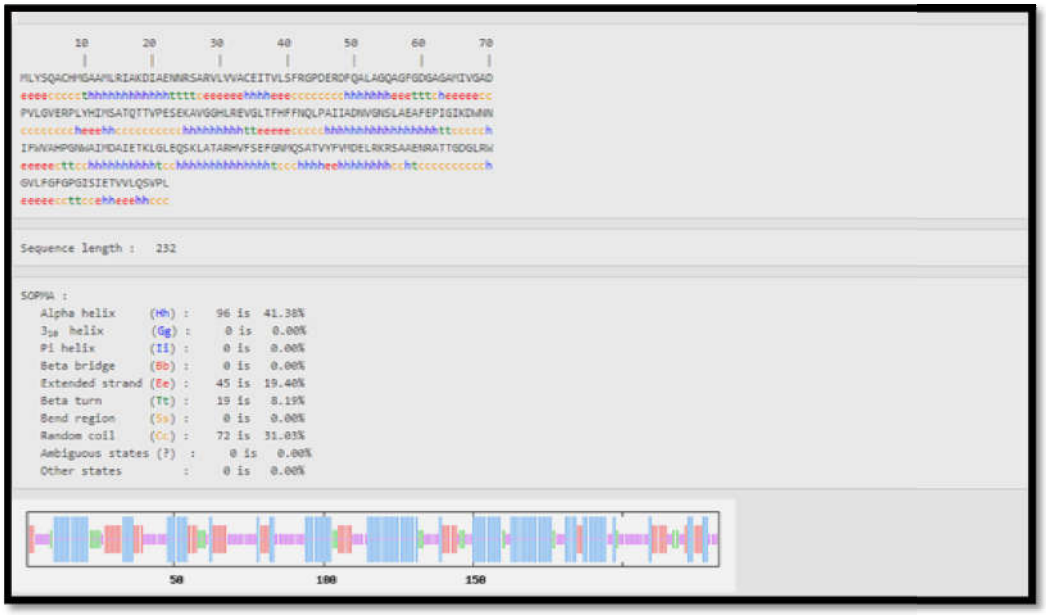


Fig. 7.17. SOPMA result for CURS2 protein

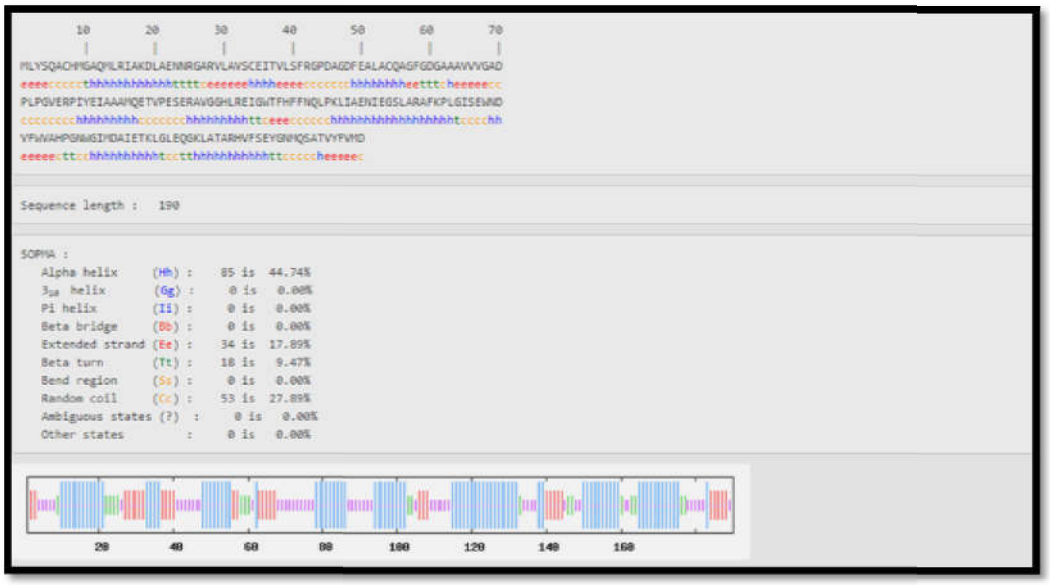


Fig. 7.18. SOPMA result for CURS3 protein

InterPro tool predicted the CURS1, CURS2 and CURS3 proteins has better homology with Polyketide synthase, type III protein family having Chalcone synthases (CHS) and stilbene synthases (STS) domain (Fig. 7.19). CHS is a key enzyme in flavanoid biosynthesis and STS is an important enzyme in stilbene-type phytoalexin biosynthesis. These enzymes catalyse the addition of three molecules of malonyl-CoA to a starter CoA ester (a typical example is 4-coumaroyl-CoA) to produce either a chalcone (with CHS) or stilbene (with STS). These enzymes have a conserved cysteine residue, located in the central core of the protein sequence, which is essential for the catalytic activity of both enzymes and probably represents the binding site for the 4-coumaryl-CoA group.

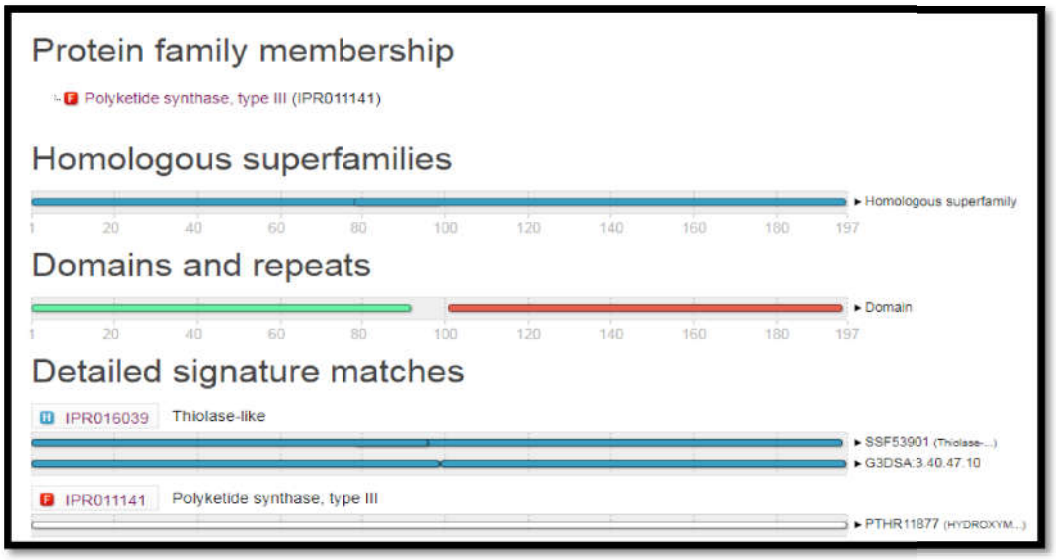


Fig. 7.19. InterPro analysis of CURS proteins

7.1.3 Model building of CURS1, CURS2 and CURS3 proteins

Homology modeling of the CURS proteins was performed using automated homology protein modeling server of SWISS MODEL, which relies on ProMod3, an in-house comparative modeling engine based on Open Structure (Fig. 7.23-7.25). Homology modeling is based on four main steps: (i) identification of structural template(s), (ii) alignment of target sequence and template structure(s), (iii) model-building and (iv) model quality evaluation. Identification of the best templates was performed using PSI-BLAST search against PDB database. The higher E-value template was selected for modeling. The potential energy was adjusted manually for predicting 3D models of CURS proteins. The best templates were identified according to the E- value and sequence similarity. The best five PDB structures (3ov2.1.A, 3ov3.1.A, 1bi5.1.A, 6co0.1.B and 6co0.1.A) were selected for homology modeling of CURS proteins.

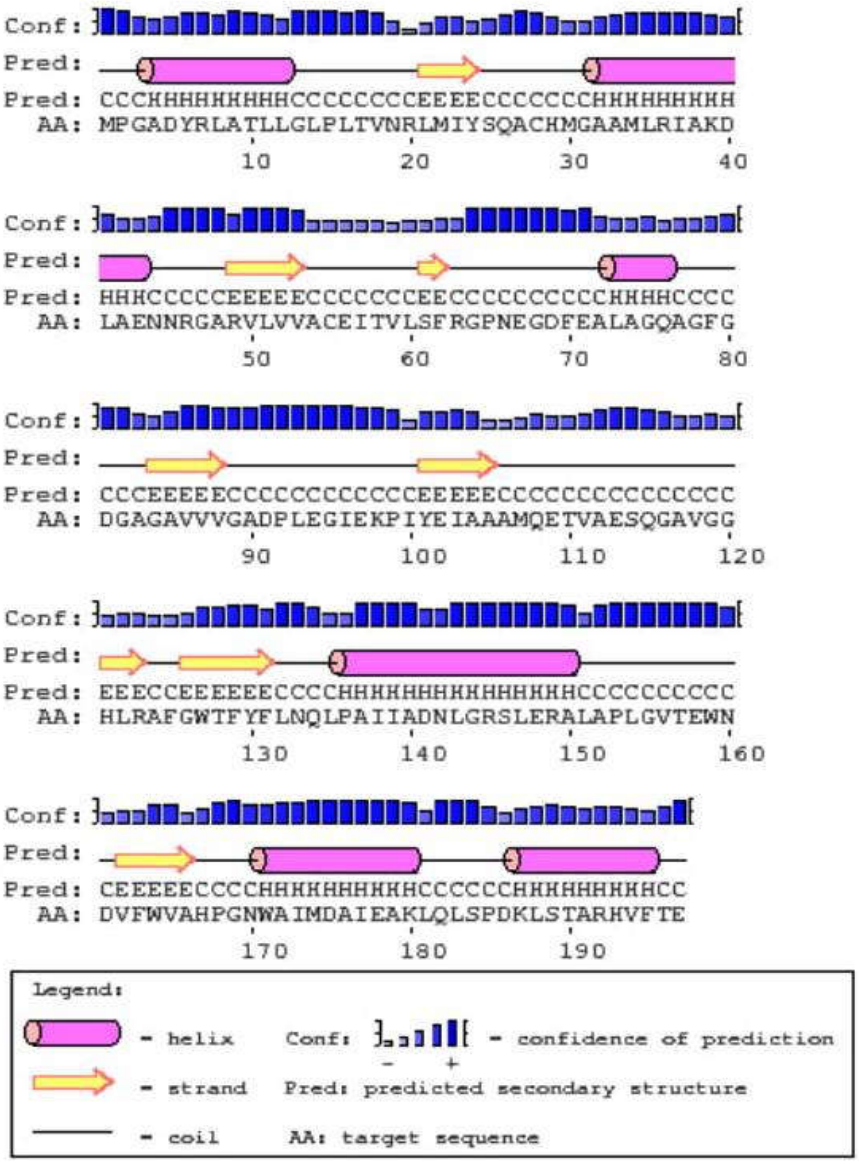


Fig. 7. 20. PSIPRED output of two dimensional graphical representation of CURS1 protein

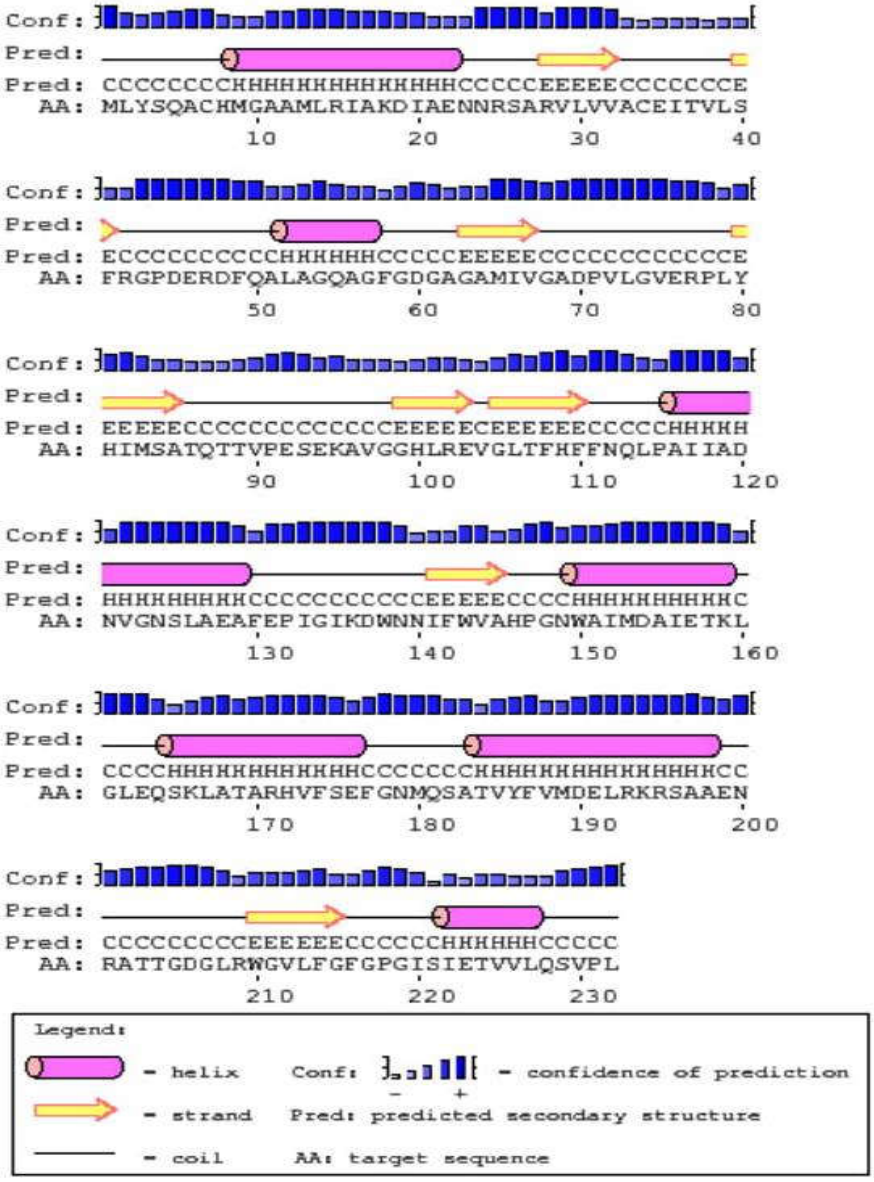


Fig. 7. 21. PSIPRED output of two dimensional graphical representation of CURS2 protein

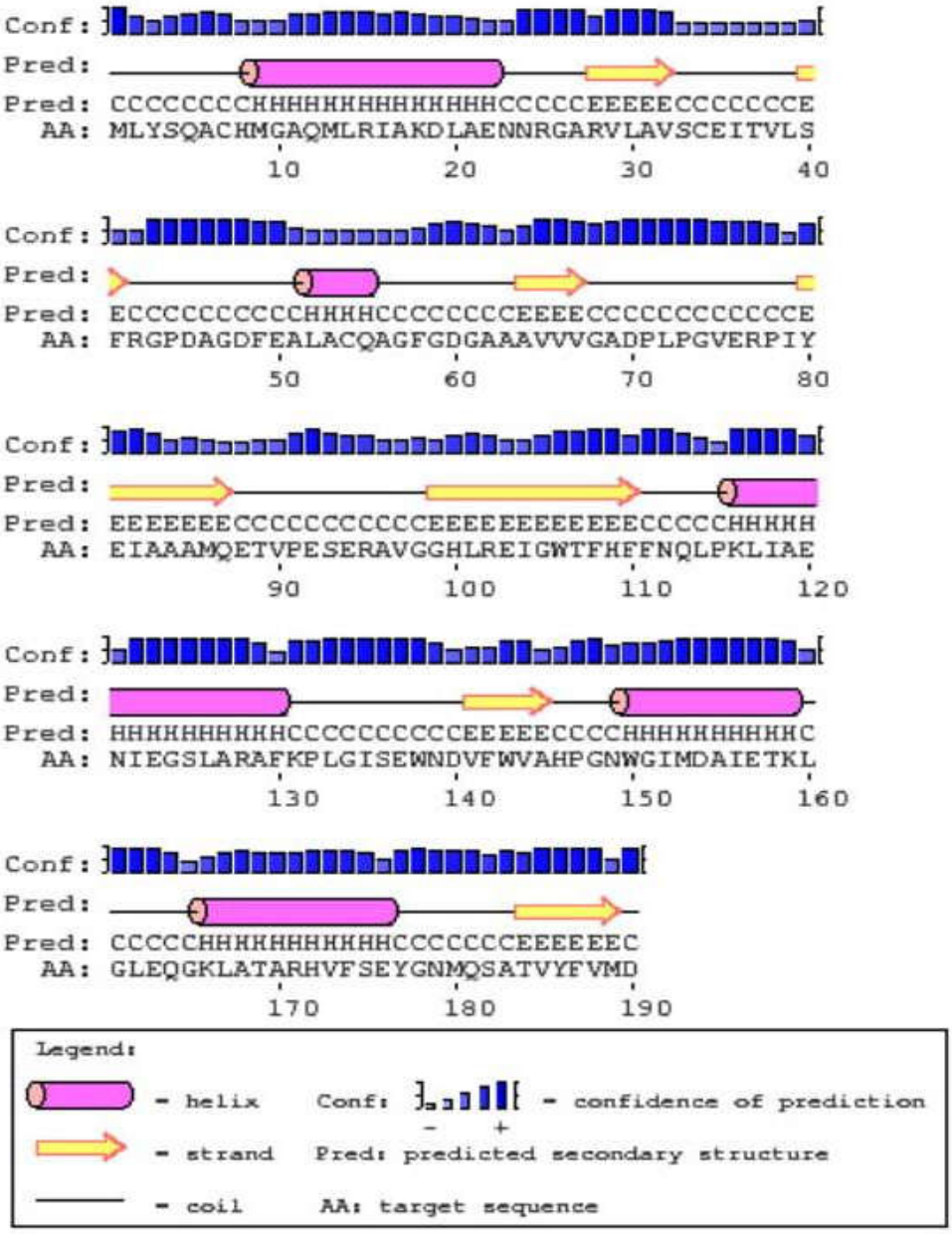


Fig. 7. 22. PSIPRED output of two dimensional graphical representation of CURS3 protein

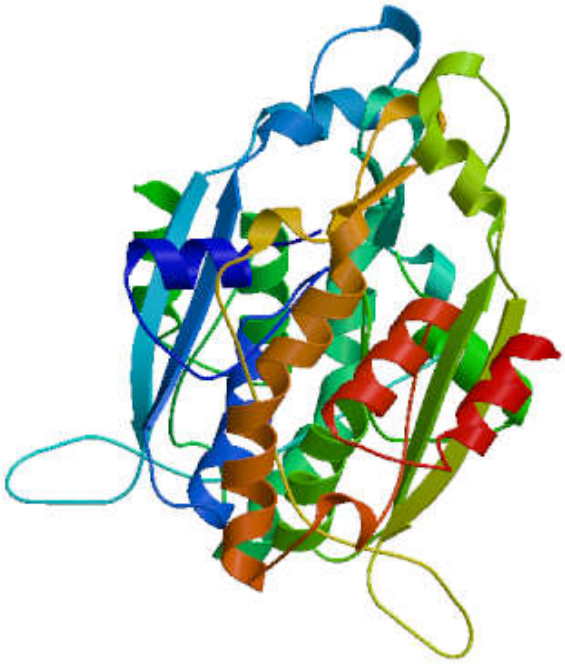


Fig. 7.23. Model 3D structure of CURS1 protein from *Curcuma longa*

Fig. 7.24. Model 3D structure of CURS2 protein from *Curcuma longa*

Fig. 7.25. Model 3D structure of CURS3 protein from *Curcuma longa*

7.1.4 Model validation

The CURS1, CURS2 and CURS3 protein models were verified using Ramachandran plot available at the MolProbity program and validated all the amino acid residues of the modeled protein fit in the allowed regions of the Ramachandran plot. The CURS1 showed 1.3% MolProbity score, 97.67% residues were in the favored regions, 0% is outliers regions and the Clash score was 0.68%. The MolProbity score is 1.6%, favored residues were 95.45%, outliers regions with 0.22% and Clash Score was 1.85% for CURS2 protein. In CURS3 protein 1.33% of MolProbity score, 96.01% of favored residues, 0% outliers regions and 0.52% Clash Score (Fig 7.26-7.28). The main and side chain parameters are summarized in and discussed in the context of accuracy of the structure prediction. The parameters of main chain such as omega angle standard deviation, bad contacts/100 residues, zeta angle

standard deviation, hydrogen bond energy standard deviation and overall G factor, which underscores to structural accuracy of a structure as determined by PROCHECK, also showed the SWISS-MODEL generated structure having better structural attributes (Tables 7.1A-7.1C).

The parameters of side chains showed as chi1-chi2 side chain torsion angle combinations for all residues with side chains long enough to have both these angles. The chi1-chi2 plots which were presented as side chain parameters in Tables 7.2A-7.2C indicate SWISS-MODEL generated structure exhibited higher score in terms of overall structural accuracy. In conclusion, the results from PROCHECK analysis clearly suggested that the modeled structure of CURS1, CURS2 and CURS3 generated by SWISS-MODEL has superior quality.

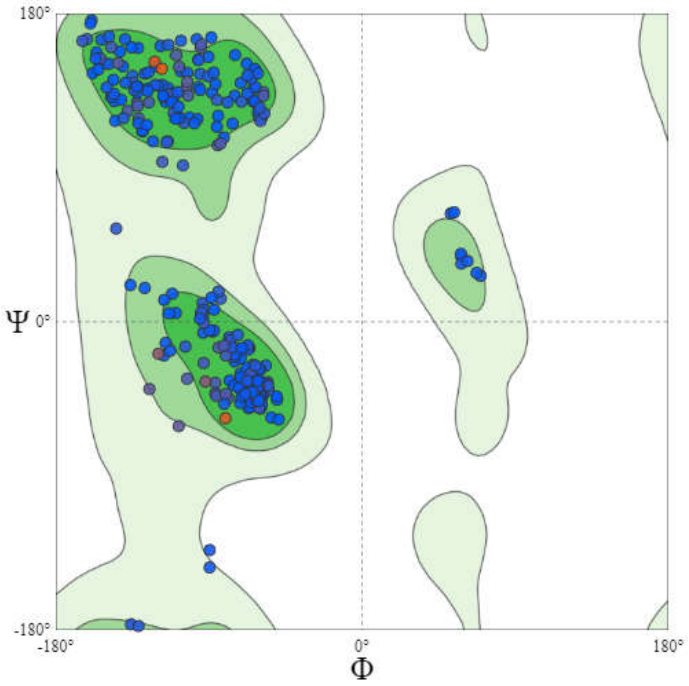


Fig. 7.26. The stereo chemical validation of the hypothetical model using Ramachandran plot of CURS1 protein

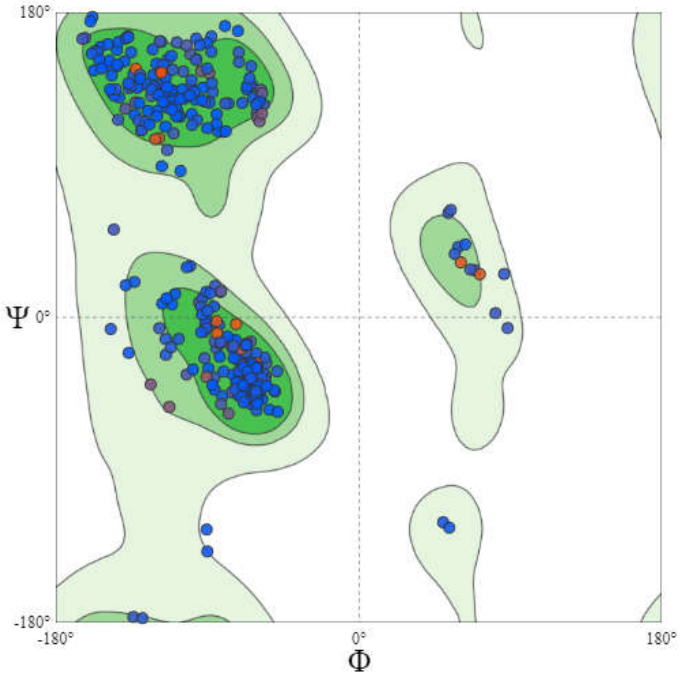


Fig. 7.27. The stereo chemical validation of the hypothetical model using Ramachandran plot of CURS2 protein

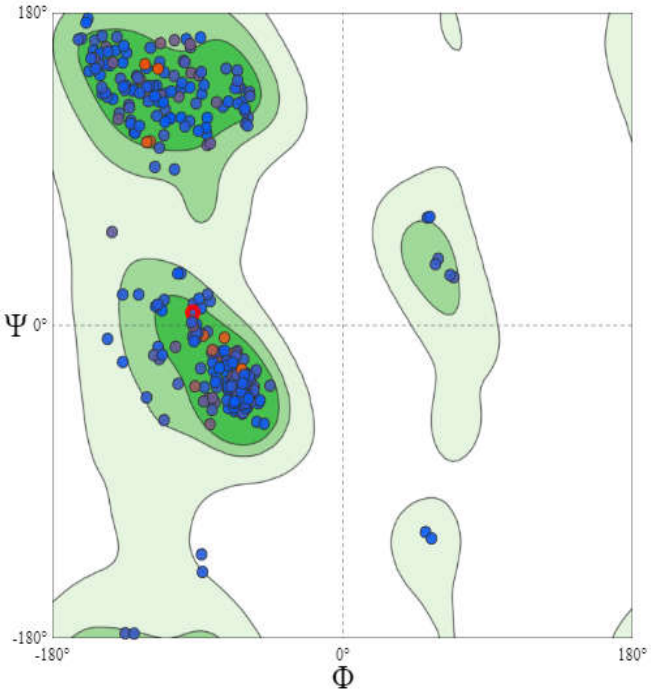


Fig. 7.28. The stereo chemical validation of the hypothetical model using Ramachandran plot of CURS3 protein

The homology model of the 3D structure of CURS (CURS1, CURS2 and CURS3) of *Curcuma longa* was submitted in PMDB (http://mi.caspur.it/PMDB/) and an accession numbers (ID: PMDB: PM0082212, PM0082213 and PM0082214) was assigned.

Table 7.1A: Summary of main chain parameters of CURS1 for the structure generated by SWISS-MODEL

Stereochemical quality	No. of data points	Parameter value	Comparison typical value	Value band width	No. of band widths from mean	Interpretation
%-tage residues in A,B,L	332	94.0	82.2	10.0	0.6	Inside
Omega angle	390	6.4	6.0	3.0	0.1	Inside
Bad Contact/100 Residue	0	0.0	1.0	10.0	-0.1	Inside
Zeta angle	354	1.8	3.1	1.6	-0.8	Inside
H-bond Energy	206	0.7	0.7	0.2	0.2	Inside
Overall G-Factor	394	-0.1	-0.2	0.3	0.4	Inside

Table 7.1B: Summary of main chain parameters of CURS2 for the structure generated by SWISS-MODEL

Stereochemical quality	No. of data points	Parameter value	Comparison typical value	Value band width	No. of band widths from mean	Interpretation
%-tage residues in A,B,L	396	93.2	82.2	10.0	0.5	Inside
Omega angle	460	6.4	6.0	3.0	0.1	Inside
Bad Contact/100 Residue	0	0.0	1.0	10.0	-0.1	Inside
Zeta angle	418	1.8	3.1	1.6	-0.8	Inside
H-bond Energy	286	0.8	0.7	0.2	0.4	Inside
Overall G-Factor	464	-0.1	-0.2	0.3	0.3	Inside

Table 7.1C: Summary of main chain parameters of CURS3 for the structure generated by SWISS-MODEL

Stereochemical quality	No. of data points	Parameter value	Comparison typical value	Value band width	No. of band widths from mean	Interpretation
%-tage residues in A,B,L	322	93.5	82.2	10.0	0.5	Inside
Omega angle	378	6.5	6.0	3.0	0.2	Inside
Bad Contact/100 Residue	0	0.0	1.0	10.0	-0.1	Inside
Zeta angle	342	1.9	3.1	1.6	-0.8	Inside
H-bond Energy	209	0.8	0.7	0.2	0.3	Inside
Overall G-Factor	380	-0.1	-0.2	0.3	0.2	Inside

The accuracy of structure is depicted in the order BETTER > INSIDE > WORSE each parameter**.**

Table 7.2A: Summary of side chain parameters of CURS1 for the structure generated by SWISS-MODEL

Stereochemical quality	No. of data points	Parameter value	Comparison typical value	Value band width	No. of band widths from mean	Interpretation
Chi-1 gauche minus	43	9.0	13.6	6.5	-0.7	Inside
Chi-1 trans	83	11.4	15.3	5.3	-0.7	Inside
Chi-1 gauche plus	151	12.2	13.8	4.9	-0.3	Inside
Chi-1 pooled	276	11.8	14.3	4.8	-0.5	Inside
Chi-2 trans	94	14.2	17.7	5.0	-0.7	Inside

Table 7.2B: Summary of side chain parameters of CURS1 for the structure generated by SWISS-MODEL

Stereochemical quality	No. of data points	Parameter value	Comparison typical value	Value band width	No. of band widths from mean	Interpretation
Chi-1gauche minus	46	14.5	13.6	6.5	0.1	Inside
Chi-1trans	104	11.0	15.3	5.3	-0.8	Inside
Chi-1gauche plus	194	10.8	13.8	4.9	-0.6	Inside
Chi-1 pooled	344	11.5	14.3	4.8	-0.6	Inside
Chi-2 trans	103	10.7	17.7	5.0	-1.4	BETTER

Table 7.2C. Summary of main chain parameters of CURS3 for the structure generated by SWISS-MODEL

Stereochemical quality	No. of data points	Parameter value	Comparison typical value	Value band width	No. of band widths from mean	Interpretation
Chi-1 gauche minus	31	12.2	13.6	6.5	-0.2	Inside
Chi-1 trans	86	13.2	15.3	5.3	-0.4	Inside
Chi-1 gauche plus	156	11.4	13.8	4.9	-0.5	Inside
Chi-1 pooled	274	12.3	14.3	4.8	-0.4	Inside
Chi-2 trans	90	13.4	17.7	5.0	-0.9	Inside

The accuracy of structure is depicted in the order BETTER > INSIDE > WORSE each parameter.

The GMQE score is 0.98 in CURS1, 0.91 in CURS2 and 0.93 in CURS3 proteins. QMEAN Z-score is -0.83, -0.89 and -1.09 for CURS1, CURS2 and CURS3 respectively. The individual Z-scores compare the interaction potential between Cβ atoms only. All atoms with the resolution potential and the torsion angle potential are shown in Figure (7.29-7.31). The "Local Quality" was estimated, for each residue of the model (reported on the x-axis), the expected similarity to the native structure (y-axis). Usually, residues showing a score below 0.6 are expected to be of low quality. In the "Comparison" plot (Fig. 7.29-7.31), model quality scores of individual models are related to scores obtained for experimental structures of similar size. The x-axis shows protein length (number of residues) and the y-axis is the normalized QMEAN score. Every dot represents one experimental protein structure. Black dots are experimental structures with a normalized QMEAN

score within 1 standard deviation of the mean (|Z-score| between 0 and 1), experimental structures with a Z-score between 1 and 2 are grey. Experimental structures that are even farther from the mean are light grey. The proposed protein model is represented as a red star.

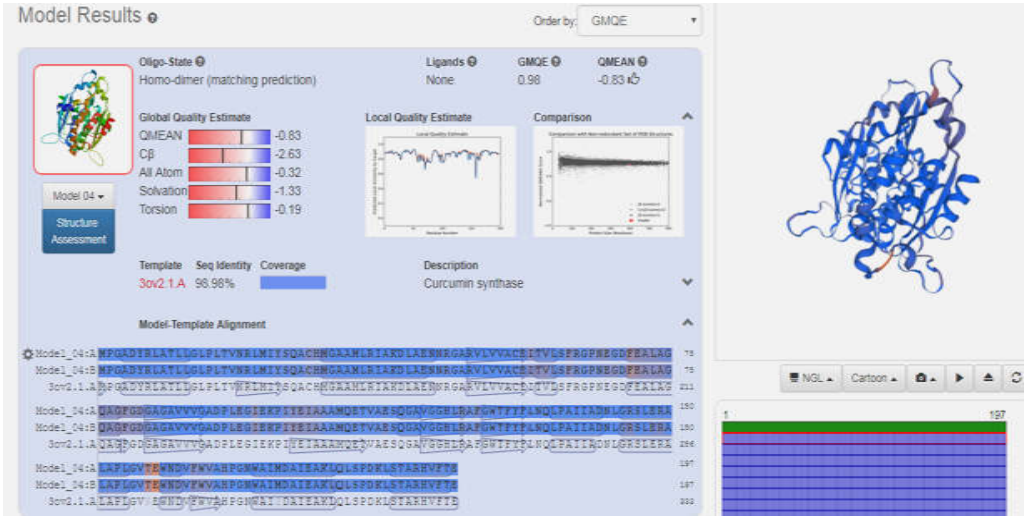


Fig. 7.29. Quality estimation (GMQE, QMEAN, Local Quality Estimate and Comparison plot) of CURS1 protein

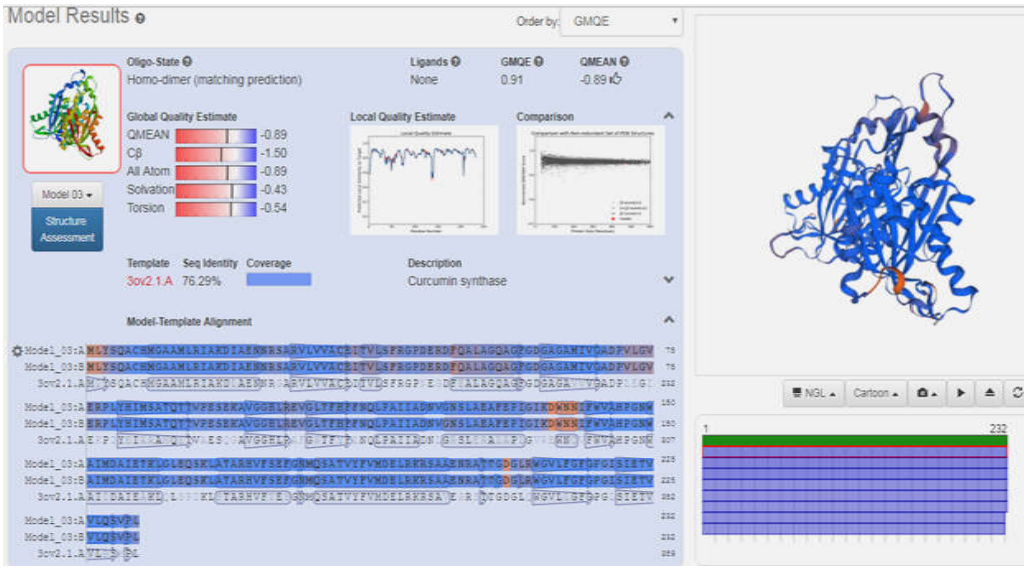


Fig. 7.30. Quality estimation (GMQE, QMEAN, Local Quality Estimate and Comparison plot) of CURS2 protein

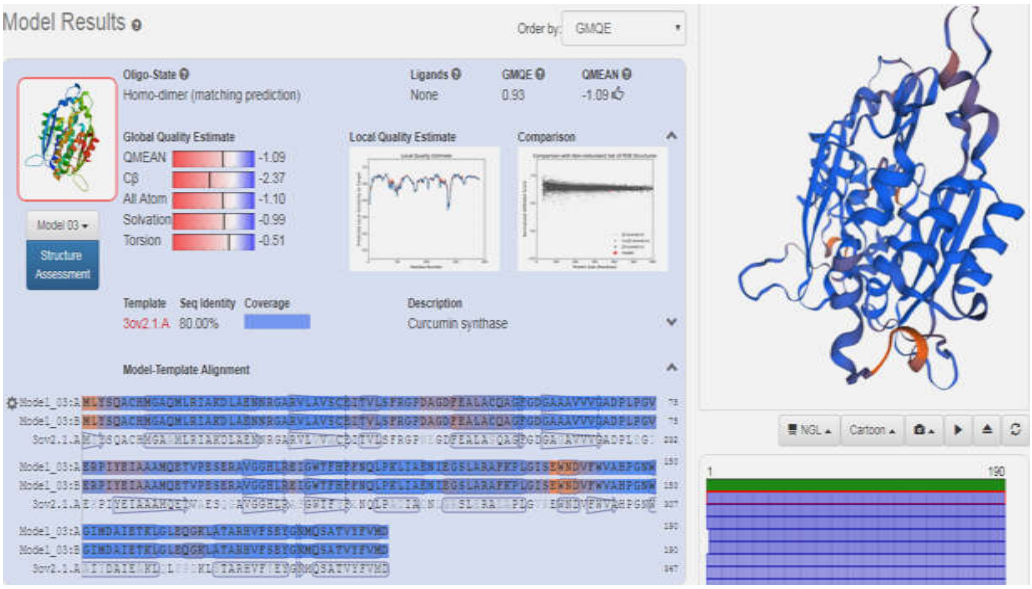


Fig. 7.30. Quality estimation (GMQE, QMEAN, Local Quality Estimate and Comparison plot) of CURS3 protein

7.2 Discussion

Homology modeling plays a crucial role in determination of protein structure in the structural genomics studies. However, a protein sequence with over 30% identity to a known structure can often be predicted with an accuracy equivalent to a low-resolution X-ray structure. The protein structures are developed by computational modeling, which is time-consuming and often complicated, will continue to spur the development of a host of new computational methods useful in filling the gaps and further contribute to understanding the relationship between protein structure and function.

Perceptive mechanism of the protein function usually required knowledge of protein three-dimensional structure, which is ultimately determined by protein sequence (Gething and Sambrook, 1992). X-ray crystallography or NMR spectroscopy methods used in protein structure determination but it is time consuming and not successful with all proteins, particularly in membrane proteins (Johnson *et al.,* 1994). The enormous gap among the number of available sequences and experimentally solved protein

structures could possibly be resolved by computational methods. Hypothetical structure prediction of protein can be divided into two different extreme methods: ab-initio method and homology modeling (Kolinski *et al.,* 1999). The aims of first approach are fold prediction from physical chemistry principles. 3D structure of a given protein sequence can be predicted based on the primarily sequence similarity to one or more proteins of known structures determined by second method.

7.2.1 Homologue detection and alignment

The homology modeling starts with selection of homologous templates with known structures from the PDB data base. The query sequence has high sequence identity (>30%) with the structure, the homology detection is quite easy which is usually done by comparing the query sequence with all the sequences of the structures in the PDB data base. This can be generally achieved simply with dynamic programming method and its derivatives. The most popular software is BLAST (http://www.ncbi.nlm.nih.gov/blast/) that searches sequence databases for optimal local alignments to the query sequences with high similarities (Xiang, 2006). Homology hits from BLAST results are not reliable when the sequence identity was below 30%. But in this study the CURS (CURS1, CURS2 and CURS3) protein sequences showed higher similarity with other CURS sequences in NCBI data base. A number of alternative strategies have also been developed such as, template consensus sequences and profile analysis (Lolkema and Slotboom, 2008). All these approaches, based on either multiple sequence or structure alignments, are more sensitive because the consensus sequences are better representative of the sequence family, and the profile reflects the conserved structural or functional preferences. Position-specific profile search methods such as PSI-BLAST, BLASTP and BLASTN have vastly improved the precision of sequence alignments and detectable sequence similarity (Hu and Kurgan, 2019). In PSI-BLAST analysis all three CURS proteins showed higher similarity with other CURS proteins (Fig. 7.4-7.6).

7.2.2 Method validation

The proteins were modelled using SWISS-MODEL (Arnold *et al.,* 2006; Kiefer *et al.,* 2009). Three different models CURS1, CURS2 and CURS3 were developed (7.23-7.25) and submitted PMDB. The stereochemical quality and accuracy of the models were tested using PROCHECK (Laskowski *et al.,* 1996) acceptable values were obtained. Ramachandran plot is one of the oldest methods for evaluation of protein 3D model yet this most reliable method of determination of the quality of protein structure (Xiao *et al.,* 2016). A structure with ≥90% of its residues in the most favored regions A, B and L of Ramachandran plot is considered to be as accurate as a 2Å-resolution crystal structure (Laskowski *et al.,* 1996). The allowed areas of the Ramachandran plot differ very significantly between glycines and the other amino acids, and to a lesser extent between different amino acids (Read *et al.,* 2011). The φ/ψ angles have a strong validation power because their values are usually not restrained in the refinement, unless a special torsion angle refinement method is used (Brünger *et al.,* 1998). It was originally suggested that more than 90% of the φ/ψ pairs should be found in the most favored areas of the plots (Pramanik *et al.,* 2017). This type of in silico homology modeling was also done by several workers (Pathak *et al.,* 2014) to predict a variety of 3D protein models of interest. The developed CURS protein 3D models are showed above 95% of Ramachandran favored residues.

A newer approach to validate macromolecular structures use of various versions of the program suite MolProbity (Chen *et al.,* 2010) used as a web server, in a stand-alone mode, or as part of other program systems, such as Phenix (Adams *et al.,* 2010). In addition to analyzing the geometrical parameters discussed above, MolProbity relies very heavily on the analysis of interatomic clashes. For that purpose the program calculates the positions of the hydrogen atoms and adds them to the coordinate files (sometimes replacing the riding H atoms that might already be present there). Another

function of MolProbity, which is now based on a much larger database compared to PROCHECK, involves the analysis of the side-chain χ torsion angles. The preferred rotamers of the side chains are contoured by excluding 1% of high-quality data, and these definitions are periodically updated (Chen *et al.,* 2010).

7.2.3 GMQE

GMQE (Global Model Quality Estimation) is a quality estimation which combines properties from the target–template alignment and the template search method. The resulting GMQE score is expressed as a number between 0 and 1, reflecting the expected accuracy of a model built with that alignment and template and the coverage of the target. Higher numbers indicate higher reliability. Once a model is built, the GMQE gets updated for this specific case by taking into account the QMEAN score of the obtained model in order to increase reliability of the quality estimation. In the present homology study the CURS1, CURS2 and CURS3 proteins showed below 1% of GMQE score, 0.96, 0.93 and 0.91 respectively.

7.2.4 QMEAN

The QMEAN Z-score provides an estimate of the "degree of nativeness" of the structural features observed in the model on a global scale. It indicates whether the QMEAN score of the model is comparable to what one would expect from experimental structures of similar size. QMEAN Z-scores of approximately zero specify superior quality between the modelled structure and experimental structures. The obtained scores of -4.0 or below indicate the models with low quality. The QMEAN Z- scores of the CURS1, CURS2 and CURS3 proteins showed -0.83, -0.89 and -1.09 respectively and these results indicate that proposed homology model was reliable and acceptable.

QMEAN quality estimators is based on different geometrical properties and provides both global (i.e. for the entire structure) and local (i.e.

per residue) absolute quality estimates on the basis of one single model and its scoring function consisting of a linear combination of six structural descriptors (Benkert *et al.,* 2010). QMEAN Z-scores are applied for the experimental structures from the PDB database (Berman *et al.,* 2000).

The Isoelectric point is the condition of a solution where the amino acid maintains the same levels of positive and negative charges and the ultimate charge will be zero. Isoelectric point (pI) of CURS1, CURS2 and CURS3 was 4.93, 5.28 and 4.96 suggesting a moderately acidic protein. Approximately neutral pH is required in in vivo condition in contrast to in vitro for the optimum activity of the alkaline phosphatase enzyme (Aminfar and Tohidfar, 2018). The total number of positively charged and negatively charged residues refers to the total number of Lysine (K), Arginine(R) and Aspartate (D), Glutamate (E) respectively (Filiz and Koç, 2014). The instability indices were between 32.10, 37.84 and 31.33. While the values of instability index for CURS1, CURS2 and CURS3 were lower than 40, suggesting the stability of the proteins (Kim *et al.,* 2011). The proteins showed higher aliphatic index which suggested that the protein is thermostable. GRAVY is used for the computational analysis of various physical-chemical parameters for a given amino acid sequence (Pramanik *et al.,* 2017). Low range GRAVY value of 0.199, 1.118 and 0.058 indicates its high affinity with that improves the solubility of a protein (Pramanik *et al.,* 2017).

7.2.5 Structural and functional characterization of CURS proteins

The secondary structures of protein chains were analysed by SOPMA tool that predicted different form like the alpha helix, extended strand, beta turn and random coil (Fig 7.16-7.18). Alpha helical structure composed of Methionine (M), Alanine (A), Leucine (L), Glutamate (E), Lysine (K) amino acids; whereas the beta strand composed of Tryptophan (W), Tyrosine

(Y), Phenylalanine (F), Valine (V), Isoleucine (I), Threonine (T) furthermore, Glycine (G), Proline (P) amino acids help to build the relevant turns (Filiz and Koç, 2014). Such findings suggested that the quantities of amino acids are solely responsible for constructing respective secondary structure of protein sequences. The percentage score of amino acid distribution infers that alpha helix is dominated over other secondary structures followed by the random coil, extended strand and beta turn, Fig 7.16-7.18 represents the secondary structure of CURS proteins where alpha helix is maximum than other structures. The unstructured regions within the cellular protein molecules are essential for intracellular signalling pathways (Muthukumaran *et al.,* 2011).

7.2.6 Post-translational modifications

The process of post-translational modification mainly includes phosphorylation, glycosylation, uniquitination, nitrosylation, methylation, acetylation, lipidation and proteolysis. Phosphorylation regulates innate inflammatory responses through the activation, cellular translation and interaction of innate receptors, adaptors and downstream signaling of molecules in response to infectious and dangerous signals (Liu *et al.,* 2016). The CURS1 protein has 3 Ser, 2 Thr and 1Tyr sites. S-146 has a score of 0.989indicates its candidacy for a phosphorylation site than the other. CURS2 has 4 Ser, 5 Thr and 2Tyr sites and S-196 has a score of 0.994 and CURS3 has 4 Ser, 1 Thr and 3Tyr sites and S-33 has a score of 0.959. The higher scores reflect the confidence of the prediction and similarity to one or more of the phosphorylation sites used in utilizing the method (Ahamd *et al.,* 2006).

SUMMARY AND CONCLUSION

The genus *Curcuma* of the family Zingiberaceae is difficult to differentiate using traditional morphological characters because the morphological differences are attributed to wide species hybridization. The present investigation was to elucidate the interspecific phylogenetic relationships among twenty species and one variety of *Curcuma* from southern India. This is the first comprehensive study on the phytochemical and molecular analysis of South Indian *Curcuma* species. The collected 20 species and one variety were documented using the herbarium sheets for further reference. A live germplasm of the species is maintained in the Calicut University Botanical Garden (CUBG). The molecular phylogenetic studies using the chloroplast, mitochondria and nuclear genes indicated that genotypic characters are not in agreement with the morphological character based phenotypic characters particularly for closely related species within the genus *Curcuma*, the nucleotide sequences of combined chloroplast and nuclear genes could provide a partial evidence for the identification of medicinal and economically important *Curcuma* species. Indeed, each group of *Curcuma* species was found to have a unique sequence pattern in all the gene regions, so that they could be easily distinguished at the DNA level.

Phylogenetic analysis using chloroplast and nuclear genes showed that *Curcuma* species require a more intensive and integrative approach of taxonomy for its species delimitation as these loci resolve, although cases of congruency between cpDNA and ITS support strong taxonomic conclusions. Another interesting case the inability of the ML method to bring more resolution than MP although insufficient sampling and rigorous analysis cannot be ruled out. Most of the species position can be attributed to the degree of hybridization observed in *Curcuma* as it is an easily hybridized

species even in nature. Use of mitochondrion locus as phylogenetic or barcoding marker did not provide conclusive evidence in this study.

The molecular diversity studies of the South Indian *Curcuma* species using chloroplast *(mat*K, *rbcL, psbK-psbI, trnlC-trnlF, psbA-trnH* and *atpF-atpH*), mitochondrial (*atp1* and *cox1*) and nuclear (ITS1) sequences were analyzed and the generated sequences from the 21 samples were deposited in NCBI. 189 sequences have been submitted in GenBank and accession numbers were assigned. The phylogenetic analysis using the different methods generated maximum likelihood and maximum parsimony trees with species grouping.

GC-MS analysis was used as another tool for the study of chemical components from the rhizome methanolic extracts of the 20 species. The GC-MS technique detected a number of compounds that were identified from different *Curcuma* species that can be used as a marker compounds to distinguish between the different species. The identification of major compounds in the rhizome extracts of twenty *Curcuma* species provided a chemical blueprint assisting in the identification of these species. The major components identified in majority of the species were; beta-caryophyllene epoxide, camphor, retinal and alloaromadendrene oxide. The components identified using GC-MS from the rhizome extracts were scored for the presence and absence of specific components and the data were analyzed using the software NTSYS-pc version 2.1. The chemical composition obtained from the rhizome extracts provided an eight clustered dendrogram for the twenty species. In this study, we included some of the wild species also, as most of these species are not exploited pharmacologically, their active ingredient identification will help the utilization of such underutilized species. By conventional morphological methods, the identification of *Curcuma* up to species level is difficult; hence the chemical composition of the rhizome methanolic extracts can be used as an important tool for species identification and chemotaxonomic studies.

Curcuminoids from *Curcuma* spp. have world-wide acceptance due to their strong anti-oxidant, anti-inflammatory, antibacterial and anti-carcinogenic activities. The most important natural products isolated to date from *C. longa* and its cultivation all over the world is affected by rhizome rot of *Pythium aphanidermatum* and leaf spot diseases caused by various pathogens. The yield is seriously affected resulting in the use of synthetic curcumin in place of natural curcumin and it would be of immense significance to explore an alternative source for curcumin other than *C. longa*. In the present study determination of curcuminoid from the different aged rhizomes of *C. zanthorrhiza, C. longa, C. zedoaria* (*C. raktakanta*) and *C. aromatica* were conducted using UFLC method. A modified method was developed for the preparation of methanolic extraction and there are very few reports available on the curcuminoid content in *C. zedoaria* and *C. aromatica.* To the best of our knowledge, this is the first report as the quantification of curcuminoids (curcumin, demethoxycurcumin, and bisdemethoxycurcumin) from the different aged rhizomes *Curcuma* species by UFLC. The UFLC method for simultaneous determination of curcuminoids seems to be accurate, precise, reproducible and repeatable. The validation data makes this method more reliable than the other analytical methods.

Curcumin is one of the most important natural products to date isolated from *C. longa.* Three classes of curcumin synthase (CURS) genes were identified, curcumin synthase 1 (CURS1, the first identified CURS), rest are type III polyketide synthases(PKSs), named CURS2 and CURS3, which shows CURS-like activity with the substrate specificity slightly different from that of CURS1 and all these three types of CURS genes are involved in curcumin synthesis pathway. The present study correlates the age related synthesis of curcumin and the expression of putative CURS (CURS1, CURS2 and CURS3) genes from the rhizomes of *C. aromatica, C. zedoaria* (*C. raktakanta*), *C. zanthorrhiza* and *C. longa* rhizomes. The CURS gene primers were developed using available CURS sequences retrieved from NCBI. The

cloned CURS sequences were submitted in GenBank and assigned accession numbers (CURS1: MK515082, MK515083, MK515084 and MK515085; CURS2: MG386668.1, MG386669.1, MG386670.1 and MG386668.1; CURS3: MK511333, MK511334, MK511335and MK511336). The transcript analysis of the CURS (CURS1, CURS2 and CURS3) genes from the rhizomes during different growth periods revealed that the highest transcript levels are observed in the vegetative growth phase during the 6th month, lesser CURS genes expression was observed during the 3rd and 9th month, compared to the 6th month. *C. longa* showed higher expression of all three CURS genes followed by *C. zanthorrhiza*, *C. aromatica* and *C. zedoaria*. Hence, CURS genes expression CURS1 showed higher expression followed by CURS3 and CURS2. A total of 189 gene sequences, specific putative CURS expression sequences (13) were submitted in NCBI and accessions number were provided.

The Identification of the 3D structure of a protein using the analytical techniques, X-ray crystallography or NMR (Nuclear Magnetic Resonance) are time consuming and expensive. In this regard, a viable alternative approach was developed to predict the in silico 3D structure of proteins based on homology modeling that serves the purpose with better validation. The primary, secondary and tertiary structure of CURS proteins was predicted by using multiple bioinformatics tools such as ExPasy ProtParam tools, NetPhos 2.0 server, SOPMA and PSIPRED. Literature shows that the homology modeling of CURS (CURS1, CURS2 and CURS3) proteins has not been explored yet, suggesting that development of a model for better understanding of CURS structure. In the predicted 3D structure, the CURS proteins were analyzed by Swiss model and validated accordingly which can be considered as an acceptable model. The modeled structure was deposited in PMDB (ID: PM0082212, PM0082213 and PM0082214).

REFERENCES

Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., & McCoy, A. J. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D: Biological Crystallography*, 66(2), 213-221.

Ahmad, D., Kikuchi, A., Jatoi, S. A., Mimura, M., & Watanabe, K. N. (2009). Genetic variation of chloroplast DNA in Zingiberaceae taxa from Myanmar assessed by PCR-restriction fragment length polymorphism analysis. *Annals of Applied Biology*, 155(1), 91-101.

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, 25(17), 3389-3402.

Alvarez, I., & Wendel, J. F. (2003). Ribosomal ITS sequences and plant phylogenetic inference. *Molecular Phylogenetics and Evolution*, 29(3), 417-434.

Amalraj, V. A., Velayudhan, K. C., & Muralidharan, V. K. (1999). *Curcuma karnatakensis* sp. nov (Zingiberaceae), a new species from Uttar Kannad district of Karnataka state. . *Journal of Economic and Taxonomic Botany,*15(2), 490-492.

Aminfar, Z., & Tohidfar, M. (2018). In silico analysis of squalene synthase in Fabaceae family using bioinformatics tools. *Journal of Genetic Engineering and Biotechnology*, 16(2), 739-747.

Angel, G. R., Makeshkumar, T., Mohan, C., Vimala, B., & Nambisan, B. (2008). Genetic diversity analysis of starchy *Curcuma* species using RAPD markers. *Journal of Plant Biochemistry and Biotechnology*, 17(2), 173-176.

Angel, G. R., Menon, N., Vimala, B., & Nambisan, B. (2014). Essential oil composition of eight starchy *Curcuma* species. *Industrial Crops and Products*, 60, 233-238.

AOAC (2016) Official methods of analysis (20th ed.). Washington, DC: Association of Official Agricultural Chemists.

Aoki, S., & Ito, M. (2000). Molecular phylogeny of *Nicotiana* (Solanaceae) based on the nucleotide sequence of the matK gene. *Plant Biology*, 2(03), 316-324.

Apavatjrut, P., Anuntalabhochai, S., Sirirugsa, P., & Alisi, C. (1999). Molecular markers in the identification of some early flowering *Curcuma* L.(Zingiberaceae) species. *Annals of Botany*, 84(4), 529-534.

Armstrong, K. F., & Ball, S. L. (2005). DNA barcodes for biosecurity: invasive species identification. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1462), 1813-1823.

Arnold, K., Bordoli, L., Kopp, J., & Schwede, T. (2006). The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics*, 22(2), 195-201.

Asghari, G., Mostajeran, A., & Shebli, M. (2009). Curcuminoid and essential oil components of turmeric at different stages of growth cultivated in Iran. *Research in Pharmaceutical Sciences*, 4(1), 55-61.

Ashraf, K., Mujeeb, M., Ahmad, A., Amir, M., Mallick, M. N., & Sharma, D. (2012). Validated HPTLC analysis method for quantification of variability in content of curcumin in *Curcuma longa* L (turmeric) collected from different geographical region of India. *Asian Pacific Journal of Tropical Biomedicine*, 2(2), S584-S588.

Avula, B., Wang, Y. H., & Khan, I. A. (2012). Quantitative determination of curcuminoids from the roots of *Curcuma longa*, *Curcuma* species and dietary supplements using an UPLC-UV-MS method. *Journal Chromatography Separate Technique*, 3(120), 2.

Awasthi, P. K., & Dixit, S. C. (2009). Chemical composition of *Curcuma longa* leaves and rhizome oil from the plains of Northern India. *Journal of Young Pharmacists*, 1(4), 312.

Azmir, J., Zaidul, I. S. M., Rahman, M. M., Sharif, K. M., Mohamed, A., Sahena, F., Jahurul, M. H .A., Ghafoor, K., Norulaini, N. A. N., & Omar, A. K. M. (2013). Techniques for extraction of bioactive compounds from plant materials: A review. *Journal of Food Engineering*, 117(4), 426-436.

Bailey, C. D., Carr, T. G., Harris, S. A., & Hughes, C. E. (2003). Characterization of angiosperm nrDNA polymorphism, paralogy, and pseudogenes. *Molecular Phylogenetics and Evolution*, 29(3), 435-455.

Baker, J. G. (1892). Zingiber. *Flora of British India*, 6, 243-249.

Baker, J.G. (1890). *Curcuma*. Pp. 209–216 in: Hooker, J.D. (ed.), *The flora of British India*, vol. 6. London: L. Reeve.

Baker, J.G. 1898. Scitamineae in dyer. *Flora Trpoical Africa*, 7,311.

Behar, N., Tiwari, K. L., & Jadhav, S. K. (2016). Semi-quantitative expression studies of genes involved in biosynthesis of curcuminoid in *Curcuma caesia* Roxb. *Indian Journal of Biotechnology*, 15*: 491-494.*

Benkert, P., Biasini, M., & Schwede, T. (2010). Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics,* 27(3), 343-350.

Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., & Sayers, E. W. (2011) GenBank. *Nucleic Acids Research,* 39, D32–D37.

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., & Bourne, P. E. (2000). The protein data bank. *Nucleic Acids Research*, 28(1), 235-242.

Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., Kiefer, F., Cassarino, T. G., Bertoni, M., Bordoli, L., & Schwede, T. (2014). SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Research*, 42(W1), W252-W258.

Bonnard, G., Weil, J. H., & Steinmetz, A. (1984). The intergenic region between the Vicia faba chloroplast tRNA CAA Leu and tRNA UAA Leu genes contains a partial copy of the split tRNA UAA Leu gene. *Current Genetics*, 9(5), 417-422.

Boonham, N., Tomlinson, J., & Mumford, R. (Eds.). (2016). *Molecular methods in plant disease diagnostics: Principles and Protocols*. Wallingford, CABI.

Bordoli, L., Kiefer, F., Arnold, K., Benkert, P., Battey, J., & Schwede, T. (2008). Protein structure homology modeling using SWISS-MODEL workspace. *Nature Protocols*, 4(1), 1.

Borsch, T., Lohne, C., & Wiersema, J. H. (2008). Phylogeny and evolutionary patterns in Nymphaeales: integrating genes, genomes and morphology. *Taxon,* 57, 1052-1081.

Brooks, G. (1998). *Biotechnology in Healthcare: An Introduction to Biopharmaceuticals*. The Pharmaceutical Press, London, UK, pp228.

Buckingham, L. (2011). *Molecular diagnostics: Fundamentals, Methods and Clinical Applications*. FA Davis. Philadelphia, Pa, USA.

Cao, H., Sasaki, Y., Fushimi, H., & Komatsu, K. (2010). Authentication of *Curcuma* species (Zingiberaceae) based on nuclear 18S rDNA and plastid trnK sequences. *Yao xue xue bao=Acta Pharmaceutica Sinica*, 45(7), 926-933.

Carra, A. M., Gambino, G., & Schubert, A. (2006). A cetyltrimethylammonium bromide-based method to extract low-molecular-weight RNA from polysaccharide-rich plant tissues. *Analytical Biochemistry***.** 360:318–320.

Casassola, A., Brammer, S. P., Chaves, M. S., Martinelli, J. A., Grando, M. F., & Denardin, N. D. (2013). Gene expression: a review on methods for the study of defense-related gene differential expression in plants. *American Journal of Plant Sciences*, 4, 64-73.

Cataldo, F. (2005). Ozone degradation of ribonucleic acid (RNA). *Polymer Degradation and Stability*, 89(2), 274-281.

Cavalli-Sforza, L. L., & Edwards, A. W. (1967). Phylogenetic analysis: models and estimation procedures. *Evolution*, 21(3), 550-570.

CBOL Plant Working Group. (2009). A DNA barcode for land plants. *Proceedings USA*, 106, 12794-12797.

CDER, (2000). Guidance for Industry Analytical Procedures and Methods Validation. FDA.

Chao, I. C., Wang, C. M., Li, S. P., Lin, L. G., Ye, W. C., & Zhang, Q. W. (2018). Simultaneous Quantification of three curcuminoids and three volatile components of *Curcuma longa* Using Pressurized Liquid Extraction and High-Performance Liquid Chromatography *Molecules*, 23(7), 1568.

Chase, M. W., Cowan, R. S., Hollingsworth, P. M., van den Berg, C., Madriñán, S., Petersen, G., Seberg, O., Jørgsensen, T., Cameron, K. M., Carine, M., & Pedersen, N. (2007). A proposal for a standardised protocol to barcode all land plants. *Taxon*, 56(2), 295-299.

Chen, J., Zhao, J., Erickson, D. L., Xia, N., & Kress, W. J. (2015). Testing DNA barcodes in closely related species of *Curcuma* (Zingiberaceae) from Myanmar and China. *Molecular Ecology Resources*, 15(2), 337-348.

Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L.W., Richardson, J.S., & Richardson, D. C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D: Biological Crystallography*, 66(1), 12-21.

Chen, W., Vermaak, I., & Viljoen, A. (2013). Camphor-a fumigant during the black death and a coveted fragrant wood in ancient Egypt and Babylon-a review. *Molecules*, 18(5), 5434-5454.

Cho, Y., Mower, J. P., Qiu, Y. L., & Palmer, J. D. (2004). Mitochondrial substitution rates are extraordinarily elevated and variable in a genus of

flowering plants. *Proceedings of the National Academy of Sciences*, 101(51), 17741-17746.

Choi, H. Y. (2009). Antioxidant Activity-of *Curcuma longa* L. Novel Foodstuff. *Molecular and Cellular Toxicology*, 5(3), 237-242.

Chomczynski, P., & Sacchi, N. (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Analytical Biochemistry*, 162(1), 156-159.

Cooper, T.H., Clark, J.G., Guzinski, J.A. (1994). Analysis of curcuminoids by high-performance liquid chromatography. In Food Phytochemicals for Cancer PreVention. II. Teas, Spices and Herbs; *ACS Symposium Series,* 547.

Corcolon, E. A., Laurena, A. C., & Dionisio-Sese, M. L. (2015). Genotypic characterization of turmeric (*Curcuma longa* L.) accessions from Mindanao, Philippines using RAPD markers. *Procedia Chemistry*, 14, 157-163.

Cosa, G., Zeng, Y., Liu, H. W., Landes, C. F., Makarov, D. E., Musier-Forsyth, K., & Barbara, P. F. (2006). Evidence for non-two-state kinetics in the nucleocapsid protein chaperoned opening of DNA hairpins. *The Journal of Physical Chemistry B*, 110(5), 2419-2426.

Crawford, D. J. (2000). Plant macromolecular systematics in the past 50 years: one view. *Taxon*, 479-501.

Crick, F. (1970). Central dogma of molecular biology. *Nature*, 227(5258), 561.

Dahm, R. (2005). Friedrich Miescher and the discovery of DNA. *Developmental Biology*, 278(2), 274-288.

Damayanti, D. (2012). Genetic Variation Analysis of *Curcuma xanthorrhiza* Roxb. by Using Amplified Fragment Length Polymorphism (AFLP) Marker. Thesis, *Sekolah Pascasarjana IPB* Graduate School, Bogor.

Dan, M., George, V., & Pushpangadan, P. (2002). Studies on the essential oils of *Curcuma haritha* Mangaly & Sabu and *C. raktakanta* Mangaly & Sabu. *Journal of Spices and Aromatic Crops*, 11(1), 78-79.

De Fatima Navarro, D., De Souza, M. M., Neto, R. A., Golin, V., Niero, R., Yunes, R. A., Delle Monache, F., & Cechinel Filho, V. (2002). Phytochemical analysis and analgesic properties of *Curcuma zedoaria* grown in Brazil. *Phytomedicine*, 9(5), 427-432.

Deepa, K., Sheeja, T. E., Rosana, O. B., Srinivasan, V., Krishnamurthy, K. S., & Sasikumar, B. (2017). Highly conserved sequence of ClPKS11 encodes a novel polyketide synthase involved in curcumin biosynthesis in turmeric (*Curcuma longa* L.). *Industrial Crops and Products*, 97, 229-241.

Deepa, K., Sheeja, T. E., Santhi, R., Sasikumar, B., Cyriac, A., Deepesh, P. V., & Prasath, D. (2014). A simple and efficient protocol for isolation of high quality functional RNA from different tissues of turmeric (*Curcuma longa* L.). *Physiology and Molecular Biology of Plants*, 20(2), 263-271.

Deng, J. B., Ding, C. B., Zhang, L., Zhou, Y. H., & Yang, R. W. (2011). Relationships among six herbal species (*Curcuma*) assessed by four isozymes. *Phyton-Revista Internacional de Botanica Experimental*, 181.

Deng, J., Liu, J., Ahmad, K. S., Ding, C., Zhang, L., Zhou, Y., & Yang, R. (2015). Relationships evaluation on six herbal species (*Curcuma*) by DNA barcoding. *Pakistan Journal of Botany***,** 47(3), 1103-1109.

Donipati, P., & Sreeramulu, S. H. (2015). Relationships among six medicinal species of *Curcuma* assessed by RAPD markers. *International Journal of Recent Scientific Research*, 6(8), 5909-5912.

Dosoky, N., & Setzer, W. (2018). Chemical Composition and Biological Activities of Essential Oils of *Curcuma* Species. *Nutrients*, 10(9), 1196.

Doyle, J. J., & Doyle, J. L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin*, 19,11–15.

Doyle, J. J., & Luckow, M. A. (2003). The rest of the iceberg. Legume diversity and evolution in a phylogenetic context. *Plant Physiology*, 131(3), 900-910.

Drager, R. G., & Hallick, R. B. (1993). A novel Euglena gracilis chloroplast operon encoding four ATP synthase subunits and two ribosomal proteins contains 17 introns. *Current Genetics*, 23(3), 271-280.

Duff, R. J., & Nickrent, D. L. (1999). Phylogenetic relationships of land plants using mitochondrial small subunit rDNA sequences. *American Journal of Botany*, 86(3), 372-386.

Duffy, A. M., Kelchner, S. A., & Wolf, P. G. (2009). Conservation of selection on matK following an ancient loss of its flanking intron. *Gene*, 438(1-2), 17-25.

Dujon, B. (1981). Mitochondrial genetics and functions. *Cold Spring Harbor Monograph Archive*, 11, 505-635.

Dung, N. X., Tuyecirc; t, N. T. B., & Leclercq, P. A. (1995). Constituents of the leaf oil of *Curcuma domestica* L. from Vietnam. *Journal of essential oil Research*, 7(6), 701-703.

Edwards, D., Horn, A., Taylor, D., Savolainen, V., & Hawkins, J. A. (2008). DNA barcoding of a large genus, *Aspalathus* L.(Fabaceae). *Taxon*, 57, 1317–1327.

El-Gewely, M. R., & Helling, R. B. (1980). Preparative separation of DNA-ethidium bromide complexes by zonal density gradient centrifugation. *Analytical biochemistry*, 102(2), 423-428.

Erpina, E., Rafi, M., Darusman, L. K., Vitasari, A., Putra, B. R., & Rohaeti, E. (2017). Simultaneous quantification of curcuminoids and xanthorrhizol in *Curcuma zanthorrhiza* by high-performance liquid chromatography. *Journal of Liquid Chromatography and Related Technologies*, 40(12), 635-639.

Fan, X. H., Cheng, Y. Y., Ye, Z. L., Lin, R. C., & Qian, Z. Z. (2006). Multiple chromatographic fingerprinting and its application to the quality control of herbal medicines. *Analytica Chimica Acta*, 555(2), 217-224.

Felsenstein, J. (1981). Evolutionary trees from DNA sequences: a maximum likelihood approach. *Journal of Molecular Evolution*, 17(6), 368-376.

Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, 39(4), 783-791.

Filiz, E., & Koc, I. (2014). In silico sequence analysis and homology modeling of predicted beta-amylase 7-like protein in *Brachypodium distachyon*. L. *Journal of BioScience and Biotechnology*, 3(1), 61-67.

Fiser, A., & Sali, A. (2003). ModLoop: automated modeling of loops in protein structures. *Bioinformatics*, 19(18), 2500-2501.

Flavell, R. (1980). The molecular characterization and organization of plant chromosomal DNA sequences. *Annual Review of Plant Physiology*, 31(1), 569-596.

Fraga, D., Meulia, T., & Fenster, S. (2008). Real-time PCR. *Current Protocols Essential Laboratory Techniques*, (1), 10-3.

Futai, M. A. S. A. M. I. T. S. U., & Kanazawa, H. I. R. O. S. H. I. (1983). Structure and function of proton-translocating adenosine triphosphatase (F0F1): biochemical and molecular biological approaches. *Microbiological Reviews*, 47(3), 285.

Futuyma, D.J. (2005). Evolution. Sinauer & Associates. *Inc., Sunderland, Massachusetts*. 226-243

Galmes, J., Flexas, J., Keys, A. J., Cifre, J., Mitchell, R. A., Madgwick, P. J., Haslam, R. P., Medrano, H., & Parry, M. A. (2005). Rubisco specificity factor tends to be larger in plant species from drier habitats and in species with persistent leaves. *Plant, Cell and Environment*, 28(5), 571-579.

Gannu, R., Yamsani, V. V., Yamsani, S. K., Palem, C. R., Voruganti, S., & Yamsani, M. R. (2009). Development of ultra fast liquid chromatographic method for simultaneous determination of nitrendipine and carvone in skin diffusate samples. *Journal of Pharmaceutical and Biomedical Analysis*, 50(5), 1080-1084.

Gangadasu, B. R., Nagarjuna, R. G., & Dhanalakshmi, K. (2015). Comparison of UPLC with UFLC: Liquid Chromatography. *International Journal of Pharmaceutical Science Review and Research*, 31(1), 97-101.

Garg, S. N., Mengi, N., Patra, N. K., Charles, R., & Kumar, S. (2002). Chemical examination of the leaf essential oil of *Curcuma longa* L. from the north Indian plains. *Flavour and Fragrance Journal*, 17(2), 103-104.

Gasteiger, E., Hoogland, C., Gattiker, A., Wilkins, M. R., Appel, R. D., & Bairoch, A. (2005). Protein identification and analysis tools on the ExPASy server. In *The Proteomics Protocols Handbook,* Humana press, 571-607.

George, M., & Britto, S. J. (2015). Phytochemicaland antioxidant studies on the essential oil of the rhizome of *Curcuma aeruginosa* Roxb. *International Research Journal of Pharmacy*, 6(8), 573-579.

George, M., & Britto, S. J. (2016). Phytochemical, antioxidant and antibacterial studies on the leaf extracts of *Curcuma amada* Roxb. *International Journal of Current Pharmaceutical Research*, 8, 32-8.

Gerbi, S. A. (1985). Evolution of ribosomal DNA. *Molecular Evolutionary Genetics*, 419-517.

Gething, M. J., & Sambrook, J. (1992). Protein folding in the cell. *Nature*, 355(6355), 33.

Ghosh, S., Majumder, P. B., & Mandi, S. S. (2011). Species-specific AFLP markers for identification of *Zingiber officinale, Z. montanumand Z.*

zerumbet (Zingiberaceae). *Genetics and Molecular Research*, 10(1), 218-229.

Gilani, S. A., Kikuchi, A., Shimazaki, T., Wicaksana, N., & Watanabe, K. N. (2015). Molecular genetic diversity of curcuminoid genes in *Curcuma amada*: Curcuminoid variation, consideration on species boundary and polyploidy. *Biochemical Systematics and Ecology*, 61, 186-195.

Goel, A., Kunnumakkara, A. B., & Aggarwal, B. B. (2008). Curcumin as "Curecumin": from kitchen to clinic. *Biochemical Pharmacology*, 75 (4), 787-809.

Gonzalez-Mendoza, D., Moreno, A. Q., & Zapata-Perez, O. (2008). An improved method for the isolation of total RNA from *Avicennia germinans* leaves. *Zeitschrift für Naturforschung C*, 63(1-2), 124-126.

Gopinath, D., Ahmed, M. R., Gomathi, K., Chitra, K., Sehgal, P. K., & Jayakumar, R. (2004). Dermal wound healing processes with curcumin incorporated collagen films. *Biomaterials*, 25(10), 1911-1917.

Gounder, D. K., & Lingamallu, J. (2012). Comparison of chemical composition and antioxidant potential of volatile oil from fresh, dried and cured turmeric (*Curcuma longa*) rhizomes. *Industrial Crops and Products*, 38, 124-131.

Grimm, G. W., & Denk, T. (2007). ITS evolution in Platanus (Platanaceae): homoeologues, pseudogenes and ancient hybridization. *Annals of Botany*, 101(3), 403-419.

Gupta, P. K., & Roy, J. K. (2002). Molecular markers in crop improvement: Present status and future needs in India. *Plant Cell, Tissue and Organ Culture*, 70(3), 229-234.

Haas, J., Roth, S., Arnold, K., Kiefer, F., Schmidt, T., Bordoli, L., & Schwede, T. (2013). The Protein Model Portal-a comprehensive resource for protein structure and model information. *Database*, 2013.

Hamilton, M. B., Braverman, J. M., & Soria-Hernanz, D. F. (2003). Patterns and relative rates of nucleotide and insertion/deletion evolution at six chloroplast intergenic regions in New World species of the Lecythidaceae. *Molecular Biology and Evolution*, 20(10), 1710-1721.

Han, J., Zhu, Y., Chen, X., Liao, B., Yao, H., Song, J., Chen, S., & Meng, F. (2013). The short ITS2 sequence serves as an efficient taxonomic sequence tag in comparison with the full-length ITS. *BioMed Research International, 2013*.

Hanley-Bowdoin, L., & Chua, N. H. (1987). Chloroplast promoters. *Trends in Biochemical Sciences*, 12, 67-70.

Hastati, S., Hadju, V., & Alam, G. (2015). Determination of the Curcumin Pigment in Extract *Curcuma domestica* Val from South Sulawesi, Indonesia. *High Performance Liquid Chromatography*, 4, 95-98.

Hayakawa, H., Kobayashi, T., Minaniya, Y., Ito, K., Miyazaki, A., Fukuda, T., & Yamamoto, Y. (2011). Development of a molecular marker to identify a candidate line of turmeric (*Curcuma longa* L.) with a high curcumin content. *American Journal of Plant Sciences*, 2(01), 15.

Hazra, K. M., Roy, R. N., Sen, S. K., & Laskar, S. (2007). Isolation of antibacterial pentahydroxy flavones from the seeds of *Mimusops elengi* Linn. *African Journal of Biotechnology*, 6(12): 1446–1449.

Hebert, P. D., Cywinska, A., Ball, S. L., & Dewaard, J. R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 270(1512), 313-321.

Heftmann, F. (1992). Chromatography: Fundamentals and Application of Chromatographic and Electrophoretic Techniques. 5th edn., Elsevier, Amsterdam, The Netherlands.

Herath, I. C., Wijayasiriwardene, T. D. C. M. K., & Premakumara, G. A. S. (2017). Comparative GC-MS analysis of all *Curcuma* species grown in Sri Lanka by multivariate test. *Ruhuna Journal of Science*, 8(2).

Hilu, K. W., & Liang, G. (1997). The *mat*K gene: sequence variation and application in plant systematics. *American journal of botany*, 84(6), 830-839.

Hilu, K. W., Borsch, T., Müller, K., Soltis, D. E., Soltis, P. S., Savolainen, V., Chase, M. W., Powell, M. P., Alice, L. A., Evans, R., & Sauquet, H. (2003). Angiosperm phylogeny based on< 011> matK sequence information. *American journal of botany*, 90(12), 1758-1776.

Himesh, S., Sharan, P. S., Mishra, K., Govind, N., & Singhai, A. K. (2011). Qualitative and quantitative profile of curcumin from ethanolic extract of *Curcuma longa*. *International Research Journal of Pharmacy*, 2(4), 180-184.

Hoffman, E. D., & Stroobant, V. (2007). Mass spectrometry: principles and applications. *West Sussex: John Wiley & Sons, Bruxellas, Bélgica*, *1* (2), 85.

Holder, M., & Lewis, P. O. (2003). Phylogeny estimation: traditional and Bayesian approaches. *Nature Reviews Genetics*, 4(4), 275.

Hollingsworth, M. L., Andra Clark, A. L. E. X., Forrest, L. L., Richardson, J., Pennington, R. T., Long, D. G., Cowan, R., Chase, M. W., Gaudeul, M., & Hollingsworth, P. M. (2009). Selecting barcoding loci for plants: evaluation of seven candidate loci with species-level sampling in three

divergent groups of land plants. *Molecular Ecology Resources*, 9(2), 439-457.

Hollingsworth, P. M., Graham, S. W., & Little, D. P. (2011). Choosing and using a plant DNA barcode. *PloS one*, 6(5), e19254.

Holttum, R. E. (1950). The zingiberaceae of the Malay Peninsula. *Gardens' Bulletin. Singapore*, 13(1), 1-249.

Horaninow, P.F. (1862) *Prodromus Monographiae Scitaminearum*. Typis Academiae Caesareae Scientiarum, Petropoli [St. Petersburg], 22–24.

Hou, P., Xie, Z., Zhang, L., Song, Z., Mi, J., He, Y., & Li, Y. (2011). Comparison of three different methods for total RNA extraction from *Fritillaria unibracteata*, a rare Chinese medicinal plant. *Journal of Medicinal Plants Research*, 5(13), 2835-2839.

Hsieh, K., Lin, S., & Wu, Y. (2014). Genetic diversity of Zingiberaceae crop. *Journal of Taiwan Agricultural Research*, 63(3), 235-248.

Hu, G., & Kurgan, L. (2019). Sequence Similarity Searching. *Current Protocols in Protein Science*, 95(1), e71.

Iida, S., Yamada, A., Amano, M., Ishii, J., Kadono, Y., & Kosuge, K. (2007). Inherited maternal effects on the drought tolerance of a natural hybrid aquatic plant, Potamogeton anguillanus. *Journal of Plant Research*, 120(4), 473-481.

Ingle, K. P., Deshmukh, A. G., Padole, D. A., Dudhare, M. S., Moharil, M. P., & Khelurkar, V. C. (2017). Phytochemicals: Extraction methods, identification and detection of bioactive compounds from plant extracts. *Journal of Pharmacognosy and Phytochemistry*, 6, 32-36.

Jadhav, B. K., Mahadik, K. R., & Paradkar, A. R. (2007). Development and validation of improved reversed phase-HPLC method for simultaneous determination of curcumin, demethoxycurcumin and bis-demethoxycurcumin. *Chromatographia*, 65(7-8), 483-488.

Jan, H. U., Rabbani, M. A., & Shinwari, Z. K. (2011). Assessment of genetic diversity of indigenous turmeric (*Curcuma longa* L.) germplasm from Pakistan using RAPD markers. *Journal of Medicinal Plants Research*, 5(5), 823-830.

Jantan, I. B., Ahmad, A. S., Ali, N. A. M., Ahmad, A. R., & Ibrahim, H. (1999). Chemical composition of the rhizome oils of four *Curcuma species* from Malaysia. *Journal of Essential Oil Research*, *11(*6), 719-723.

Jarikasem, S., Thubthimthed, S., Chawananoraseth, K., & Suntorntanasat, T. (2003). Essential oils from three *Curcuma* species collected in Thailand. In *III WOCMAP Congress on Medicinal and Aromatic Plants-Volume 3: Perspectives in Natural Product Chemistry,* 677, 37-41.

Jasim, F., & Ali, F. (1989). Measurements of some spectrophotometric parameters of curcumin in 12 polar and nonpolar organic solvents. *Microchemical journal*, 39(2), 156-159.

Jatoi, S. A., Kikuchi, A., & Watanabe, K. N. (2007). Genetic diversity, cytology, and systematic and phylogenetic studies in Zingiberaceae. *Genes, Genomes and Genomics*, 1(1), 56-62.

Jayaprakasha, G. K., Jagan Mohan Rao, L., & Sakariah, K. K. (2002). Improved HPLC method for the determination of curcumin,

demethoxycurcumin, and bisdemethoxycurcumin. *Journal of Agricultural and Food Chemistry*, 50(13), 3668-3672.

Jayapriya, G., & Shoba, F. G. (2015). GC-MS analysis of bio-active compounds in methanolic leaf extracts of *Justicia adhatoda* (Linn.). *Journal of Pharmacognosy and Phytochemistry*, 4(1), 113-117.

Jeanson, M. L., Labat, J. N., & Little, D. P. (2011). DNA barcoding: a new tool for palm taxonomists?. *Annals of Botany*, 108(8), 1445-1451.

Johnson, M. S., Srinivasan, N., Sowdhamini, R., & Blundell, T. L. (1994). Knowledge-based protein modeling. *Critical Reviews in Biochemistry and Molecular Biology*, 29(1), 1-68.

Julius, A., & Suleiman, M. (2008). Preliminary molecular phylogeny of *Bornean plagiostachys* (Zingiberaceae) based on DNA sequence data of internal transcribed spacer (ITS). *Journal of Tropical Biology & Conservation (JTBC)*, 4 (1): 67 – 80.

Kaewsri, W., Paisooksantivatana, Y., Veesommai, U., Eiadthong, W., & Vajrodaya, S. (2007). Phylogenetic analysis of Thai *Amomum* (Alpinioideae: Zingiberaceae) using AFLP markers. *Kasetsart Journal of Natural Science*, 41, 213-226.

Kalpravidh, R.W., Siritanaratkul, N., Insain, P., Charoensakdi, R., Panichkul, N., Hatairaktham, S., Srichairatanakool, S., Phisalaphong, C., Rachmilewitz, E., & Fucharoen, S. (2010). Improvement in oxidative stress and antioxidant parameters in β-thalassemia/Hb E patients treated with curcuminoids. *Clinical Biochemistry*. 43(4-5):424-429

Kar, D., Pattanaik, P. K., Acharya, L., Panda, M. K., Sathapathy, K., Kuanar, A., & Mishra, B. (2014). Assessment of genetic diversity among some

elite cultivars of ginger (*Zingiber officinale* Rosc.) using isozyme and protein markers. *Brazilian Journal of Botany*, 37(4), 469-479.

Karp, A., & Edwards, K. J. (1995). Molecular techniques in the analysis of the extent and distribution of genetic diversity. In: *Molecular Genetic Techniques for Plant Genetic Resources. In: IPGRI workshop*.

Katsuyama, Y., Kita, T., & Horinouchi, S. (2009b). Identification and characterization of multiple curcumin synthases from the herb *Curcuma longa*. *FEBS letters*, 583(17), 2799-2803.

Katsuyama, Y., Kita, T., Funa, N., & Horinouchi, S. (2009a). Curcuminoid biosynthesis by two type III polyketide synthases in the herb *Curcuma longa*. *Journal of Biological Chemistry*, 284(17), 11160-11170.

Kawasaki, E. S., Clark, S. S., Coyne, M. Y., Smith, S. D., Champlin, R., Witte, O. N., & McCormick, F. P. (1988). Diagnosis of chronic myeloid and acute lymphocytic leukemias by detection of leukemia-specific mRNA sequences amplified in vitro. *Proceedings of the National Academy of Sciences*, 85(15), 5698-5702.

Kell, D. B., Brown, M., Davey, H. M., Dunn, W. B., Spasic, I., & Oliver, S. G. (2005). Metabolic footprinting and systems biology: the medium is the message. *Nature Reviews Microbiology*, 3(7), 557.

Kellogg, E. A., & Juliano, N. D. (1997). The structure and function of RuBisCO and their implications for systematic studies. *American Journal of Botany*, 84(3), 413-428.

Khan, F. I., Wei, D. Q., Gu, K. R., Hassan, M. I., & Tabrez, S. (2016). Current updates on computer aided protein modeling and designing. *International Journal of Biological Macromolecules*, 85, 48-62.

Khan, S., Shama Naz, R., & Naeem, R. (2013). Genetic fingerprinting of local turmeric genotypes using RAPDS. *Pakistan Journal of Botany***,** 45(S1), 339-346.

Kiefer, F., Arnold, K., Künzli, M., Bordoli, L., & Schwede, T. (2008). The SWISS-MODEL Repository and associated resources. *Nucleic acids research*, 37(suppl_1), D387-D392.

Kim, S. J., Cho, K. S., Yoo, K. O., Lim, K. B., Hwang, Y. J., Chang, D. C., & Kim, K. S. (2015). Sequence analysis of the internal transcribed spacer (ITS) region of the nuclear ribosomal DNA (nrDNA) Chrysanthemum species in Korea. *Horticulture, Environment, and Biotechnology*, 56 (1), 44-53.

Kim, Y. S., Cho, J. H., Park, S., Han, J. Y., Back, K., & Choi, Y. E. (2011). Gene regulation patterns in triterpene biosynthetic pathway driven by overexpression of squalene synthase and methyl jasmonate elicitation in *Bupleurum falcatum*. *Planta*, 233(2), 343-355.

Kitamura, C., Nagoe, T., Prana, M. S., Agusta, A., Ohashi, K., & Shibuya, H. (2007). Comparison of *Curcuma* sp. in Yakushima with *C. aeruginosa* and *C. zedoaria* in Java by trnK gene sequence, RAPD pattern and essential oil component. *Journal of Natural Medicines*, 61(3), 239-243.

Kladmook, M., Chidchenchey, S., & Keeratinijakal, V. (2010). Assessment of genetic diversity in cassumunar ginger (*Zingiber cassumunar* Roxb.) in Thailand using AFLP markers. *Breeding Science*, 60(4), 412-418.

Knoop, V. (2011). When you can't trust the DNA: RNA editing changes transcript sequences. *Cellular and Molecular Life Sciences*, 68(4), 567-586.

Kolinski, A., Rotkiewicz, P., Ilkowski, B., & Skolnick, J. (1999). A method for the improvement of threading based protein models. *Proteins: Structure, Function, and Bioinformatics*, 37(4), 592-610.

Kong, J. J., Xia, Y. M., & Li, Q. J. (2010). Inflorescence and flower development in the Hedychieae (Zingiberaceae): *Hedychium coccineum* Smith. *Protoplasma*, 247(1-2), 83-90.

Konieczny, A., & Legocki, A. B. (1981). Translation and characterization of poly (A)-lacking RNA from lupin root nodules. *Acta Biochimica Biochimica Polonica*, 28, 83.

Koski, L. B., & Golding, G. B. (2001). The closest BLAST hit is often not the nearest neighbour. *Journal of Molecular Evolution*, 52(6), 540-542.

Kress, W. J., Prince, L. M., & Williams, K. J. (2002). The phylogeny and a new classification of the gingers (Zingiberaceae): evidence from molecular data. *American Journal of Botany*, 89(10), 1682-1696.

Kress, W. J., Wurdack, K. J., Zimmer, E. A., Weigt, L. A., & Janzen, D. H. (2005). Use of DNA barcodes to identify flowering plants. *Proceedings of the National Academy of Sciences*, 102(23), 8369-8374.

Kress, W.J., Newman, M.F., Poulsen, A.D., Specht, C.D. (2007). An analysis of generic circumscriptions in tribe Alpinieae (Alpinioideae: Zingiberaceae). *Gardens' Bulletin Singapore*, 59: 113–128.

Krishnamurthy, N., Mathew, A. G., Nambudiri, E. S., Shivashankar, S., Lewis, Y. S., & Natarajan, C. P. (1976). Oil and oleoresin of turmeric. *Tropical Science*, 18(1), 37.

Krup, V., Prakash, L. H., & Harini, A. (2013). Pharmacological activities of turmeric (*Curcuma longa* linn): a review. *Journal of Homeopathy and Ayurvedic Medicine*, 2(133), 2167-1206.

Kryndushkin, D. S., Alexandrov, I. M., Ter-Avanesyan, M. D., & Kushnirov, V. V. (2003). Yeast [PSI+] prion aggregates are formed by small Sup35 polymers fragmented by Hsp104. *Journal of Biological Chemistry*, 278(49), 49636-49643.

Kumar, S. (2010). Development of molecular markers for genetic analysis of plants. In: Kumar A (ed) Plant Genetic Transformation and Molecular Markers. *Pointer Publisher*, *Jaipur*, 173–203.

Lampe, V., & Milobedzka, J. (1913). Studien uber curcumin. *Berichte der deutschen chemischen Gesellschaft*, 46(2), 2235-2240.

Lantz, R. C., Chen, G. J., Solyom, A. M., Jolad, S. D., & Timmermann, B. N. (2005). The effect of turmeric extracts on inflammatory mediator production. *Phytomedicine*, 12(6-7), 445-452.

Larsen, K., Lock, J. M., Maas, H., & Maas, P. J. M. (1998). Zingiberaceae. In *Flowering Plants· Monocotyledons* , 474-495. Springer, Berlin, Heidelberg.

Laskowski, R. A., Chistyakov, V. V., & Thornton, J. M. (2005). PDBsum more: new summaries and analyses of the known 3D structures of proteins and nucleic acids. *Nucleic Acids Research*, 33(suppl_1), D266-D268.

Laskowski, R. A., Rullmann, J. A. C., MacArthur, M. W., Kaptein, R., & Thornton, J. M. (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *Journal of Biomolecular NMR*, 8(4), 477-486.

Ledford, H. (2008). Botanical identities: DNA barcoding for plants comes a step closer. *Nature*, 451,616.

Lee, K. J., Kim, Y. S., & Ma, J. Y. (2013). Separation and identification of curcuminoids from Asian turmeric (*Curcuma longa* L.) using RP-HPLC and LC-MS. *Asian Journal of Chemistry*, 25(2), 909.

Lee, V. R. (2010). Adaptations and continuities in the use and design of visual representations in US middle school science textbooks. *International Journal of Science Education*, 32(8), 1099-1126.

Leong-Skornickova, J., Sída, O., Jarolimova, V., Sabu, M., Fer, T., Travnicek, P., & Suda, J. (2007). Chromosome numbers and genome size variation in Indian species of *Curcuma* (Zingiberaceae). *Annals of Botany*, 100(3), 505-526.

Levin, J.M., Pascarella, S., Argosand, P., and Garnier, J. (1993). Quantification of secondary structure prediction Improvement Using multiple alignment. *Protein Engineering,* 6, 849-854.

Liang, H., & Hilu, K. W. (1996). Application of the *mat*K gene sequences to grass systematics. *Canadian Journal of Botany*, 74(1), 125-134.

Little, D. P., & Stevenson, D. W. (2007). A comparison of algorithms for the identification of specimens using DNA barcodes: examples from gymnosperms. *Cladistics*, 23(1), 1-21.

Liu, E., Wu, J., Cao, W., Zhang, J., Liu, W., Jiang, X., & Zhang, X. (2007). Curcumin induces G2/M cell cycle arrest in a p53-dependent manner and upregulates ING4 expression in human glioma. *Journal of Neuro-oncology*, 85(3), 263-270.

Liu, H., Carvalhais, L. C., Kazan, K., & Schenk, P. M. (2016). Development of marker genes for jasmonic acid signaling in shoots and roots of wheat. *Plant Signaling & Behavior*, *11*(5), e1176654.

Liu, Y., Cox, C. J., Wang, W., & Goffinet, B. (2014). Mitochondrial phylogenomics of early land plants: mitigating the effects of saturation, compositional heterogeneity, and codon-usage bias. *Systematic Biology*, 63(6), 862-878.

Liu, Z., Zhao, N., Liu, C., Zhou, L., & Du, S. (2012). Identification of insecticidal constituents of the essential oil of *Curcuma wenyujin* rhizomes active against *Liposcelis bostrychophila* Badonnel. *Molecules*, 17(10), 12049-12060.

Livak, K. J., & Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C_T}$ method. *Methods*, 25(4), 402-408.

Lohne, C., Yoo, M. J., Borsch, T., Wiersema, J., Wilde, V., Bell, C. D., Barthlott, W., Soltis, D.E., & Soltis, P. S. (2008). Biogeography of Nymphaeales: extant patterns and historical events. *Taxon*, 57(4), 1123-19E.

Lolkema, J. S., & Slotboom, D. J. (2008). The major amino acid transporter superfamily has a similar core structure as Na+-galactose and Na+-leucine transporters. *Molecular Membrane Biology*, 25(6-7), 567-570.

Long, Y., Zhang, W., Wang, F., & Chen, Z. (2014). Simultaneous determination of three curcuminoids in *Curcuma longa* L. by high performance liquid chromatography coupled with electrochemical detection. *Journal of Pharmaceutical Analysis*, 4(5), 325-330.

Malek, S. N., Seng, C. K., Zakaria, Z., Ali, N. A., Ibrahim, H., & Jalil, M. N. (2006). The essential oil of *Curcuma inodora* aff. Blatter from Malaysia. *Journal of Essential Oil Research*, 18(3), 281-283.

Martin, M., & Guiochon, G. (2005). Effects of high pressure in liquid chromatography. *Journal of Chromatography A*, 1090(1-2), 16-38.

Marti-Renom, M. A., Stuart, A. C., Fiser, A., Sánchez, R., Melo, F., & Šali, A. (2000). Comparative protein structure modeling of genes and genomes. *Annual Review of Biophysics and Biomolecular Structure*, 29(1), 291-325.

Maxam, A. M., & Gilbert, W. (1977). A new method for sequencing DNA. *Proceedings of the National Academy of Sciences*, 74(2), 560-564.

McGuffin, L. J., Bryson, K., & Jones, D. T. (2000). The PSIPRED protein structure prediction server. *Bioinformatics*, 16(4), 404-405.

Meng, B. Y., Wakasugi, T., & Sugiura, M. (1991). Two promoters within the psbK-psbI-trnG gene cluster in tobacco chloroplast DNA. *Current Genetics*, 20(3), 259-264.

Meng, Z., Shao, J., Cao, W., Li, J., Zhou, Y., & Wang, X. (2014). RapidTree: a solution to rapid reconstruction phylogenetic tree. In *2014 11th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)* (pp. 513-517). IEEE.

Meusnier, I., Singer, G. A., Landry, J. F., Hickey, D. A., Hebert, P. D., & Hajibabaei, M. (2008). A universal DNA mini-barcode for biodiversity analysis. *BMC Genomics*, 9(1), 214.

Minami, M., Nishio, K., Ajioka, Y., Kyushima, H., Shigeki, K., Kinjo, K., Yamada, K., Nagai, M., Satoh, K., & Sakurai, Y. (2009). Identification of *Curcuma* plants and curcumin content level by DNA polymorphisms in the trnS-trnfM intergenic spacer in chloroplast DNA. *Journal of Natural Medicines*, 63(1), 75-79.

Ming, L. I., Hui, C. A. O., But, P. P. H., & Pang-Chui, S. H. A. W. (2011). Identification of herbal medicinal materials using DNA barcodes. *Journal of Systematics and Evolution*, 49(3), 271-283.

Miyata, T., Yasunaga, T., & Nishida, T. (1980). Nucleotide sequence divergence and functional constraint in mRNA evolution. *Proceedings of the National Academy of Sciences*, 77(12), 7328-7332.

Mohanty, S., Panda, M. K., Acharya, L., & Nayak, S. (2014). Genetic diversity and gene differentiation among ten species of Zingiberaceae from Eastern India. *3 Biotech*, 4(4), 383-390.

Monton, C., Charoenchai, L., Suksaeree, J., & Sueree, L. (2016). Quantitation of curcuminoid contents, dissolution profile, and volatile oil content of turmeric capsules produced at some secondary government hospitals. *Journal of Food and Drug Analysis*, 24(3), 493-499.

Moon, B. C., Kim, W. J., Ji, Y., Lee, Y. M., & Kim, H. K. (2013). Genetic diversity of *Curcuma* genus collected germplasm using analysis of AFLP. *Korean Journal of Medicinal Crop Science*, 21(6), 455-460.

Moon, K., Khadabadi, S. S., Deokate, U. A., & Deore, S. L. (2010). *Caesalpinia bonducella* F-an overview. *Report and Opinion*, 2(3), 83-90.

Moritz, C. T. E. D., Dowling, T. E., & Brown, W. M. (1987). Evolution of animal mitochondrial DNA: relevance for population biology and

systematics. *Annual Review of Ecology and systematics*, 18(1), 269-292.

Moult, J. (2005). A decade of CASP: progress, bottlenecks and prognosis in protein structure prediction. *Current Opinion in Structural Biology*, 15(3), 285-289.

Moult, J., Fidelis, K., Kryshtafovych, A., Rost, B., & Tramontano, A. (2009). Critical assessment of methods of protein structure prediction—Round VIII. *Proteins: Structure, Function, and Bioinformatics*, 77(S9), 1-4.

Muniyappan, N., & Nagarajan, N. S. (2014). Green synthesis of gold nanoparticles using *Curcuma pseudomontana* essential oil, its biological activity and cytotoxicity against human ductal breast carcinoma cells T47D. *Journal of Environmental Chemical Engineering*, 2(4), 2037-2044.

Muthukumaran, J., Manivel, P., Kannan, M., Jeyakanthan, J., & Krishna, R. (2011). A framework for classification of antifreeze proteins in over wintering plants based on their sequence and structural features. *Journal of Bioinformatics and Sequence Analysis*, 3(4), 70-88.

Naidu, M. M., Shyamala, B. N., Manjunatha, J. R., Sulochanamma, G., & Srinivas, P. (2009). Simple HPLC method for resolution of curcuminoids with antioxidant potential. *Journal of Food Science*, 74(4), C312-C318.

Nair, P., Misra, A., Singh, A., Shukla, A. K., Gupta, M. M., Gupta, A. K., Gupta, V., Khanuja, S. P., & Shasany, A. K. (2013). Differentially expressed genes during contrasting growth stages of *Artemisia annua* for artemisinin content. *PLoS One*, 8(4), e60375.

Nakamura, Y., Carlson, M., Krapcho, K., Kanamori, M., & White, R. (1988). New approach for isolation of VNTR markers. *American journal of Human Genetics*, 43(6), 854.

Nampoothiri, S.V., Philip, R.M., Kankangi, S., Kiran, C.R., & Menon, A.N. (2015). Essential oil composition, α-amylase inhibition and antiglycation potential of *Curcuma aromatica* Salisb. *Journal of Essential Oil Bearing Plants*. 18, 1051–1058.

Nayak, S., Jena, A. K., Mittal, D. K., & Joshi, D. (2014). GC–MS analysis of phytoconstituents of some wild Zingiberaceae plants methanolic rhizome extracts. *Research in Plant Sciences*, 2(1), 1-5.

Naz, S., Ilyas, S., Parveen, Z., & Javed, S. (2010). Chemical analysis of essential oils from turmeric (*Curcuma longa*) rhizome through GC-MS. *Asian Journal of Chemistry*, 22(4), 3153.

Neema, A. (2005). Investigations on the Biosynthesis of Curcumin in Turmeric (*Curcuma Longa* L.). University of Calicut at Kerala, India (Ph.D dissertation).

Nei, M. (1991). Relative efficiencies of different tree making methods for molecular data. In Recent Advances in Phylogenetic Studies of DNA Sequences. *Oxford University Press.* 133-47.

Nei, M. (1996). Phylogenetic analysis in molecular evolutionary genetics. *Annual Review of Genetics*, 30(1), 371-403.

Nei, M., & Kumar, S. (2000). Molecular evolution and phylogenetics. Oxford university press.

Nelson, K. M., Dahlin, J. L., Bisson, J., Graham, J., Pauli, G. F., & Walters, M. A. (2017). The essential medicinal chemistry of curcumin: miniperspective. *Journal of Medicinal Chemistry*, 60(5), 1620-1637.

Ngamriabsakul, C., & Techaprasan, J. (2006). The phylogeny of Thai *Boesenbergia* (Zingiberaceae) based on petA-psbJ spacer (chloroplast DNA). *Journal* of *Science and Technology*, 28, 49-57.

Ngamriabsakul, C., Newman, M. F., & Cronk, Q. C. B. (2003). The phylogeny of tribe Zingibereae (Zingiberaceae) based on its (nrDNA) and trnl–f (cpDNA) sequences. *Edinburgh Journal of Botany*, 60(3), 483-507.

Nian, L., & Deling, W. (1999). Notes on *Curcuma* in China. *Journal of Tropical and Subtropical Botany*, 7(2), 146-150.

Nock, C. J., Waters, D. L., Edwards, M. A., Bowen, S. G., Rice, N., Cordeiro, G. M., & Henry, R. J. (2011). Chloroplast genome sequences from total DNA for plant identification. *Plant Biotechnology Journal*, 9(3), 328-333.

Nolan, T., Hands, R. E., & Bustin, S. A. (2006). Quantification of mRNA using real-time RT-PCR. *Nature Protocols*, 1(3), 1559.

Nugroho, A., Rohman, A., Lukitaningsih, E., Rakhmawati, N., & Sudjadi, S. (2015). Analysis of curcumin in ethanolic extract of *Curcuma longa* Linn. and *Curcuma xanthorrhiza* Roxb. Using high performance liquid chromatography with UV-Detection. *Research Journal of Phytochemistry,* 9: 188-194.

Nygard, A. B., Jorgensen, C. B., Cirera, S., & Fredholm, M. (2007). Selection of reference genes for gene expression studies in pig tissues using SYBR green qPCR. *BMC Molecular Biology*, 8(1), 67.

Okuyama, Y., Fujii, N., Wakabayashi, M., Kawakita, A., Ito, M., Watanabe, M., Murakami, N., & Makoto, K. (2005). Nonuniform concerted evolution and chloroplast capture: Heterogeneity of observed introgression patterns in three molecular data partition phylogenies of Asian *Mitella* (Saxifragaceae). *Molecular Biology and Evolution.* 22: 285–296.

Orengo, C. A., Michie, A. D., Jones, S., Jones, D. T., Swindells, M. B., & Thornton, J. M. (1997). CATH–a hierarchic classification of protein domain structures. *Structure*, 5(8), 1093-1109.

Osorio-Tobon, J. F., Carvalho, P. I., Barbero, G. F., Nogueira, G. C., Rostagno, M. A., & de Almeida Meireles, M. A. (2016). Fast analysis of curcuminoids from turmeric (*Curcuma longa* L.) by high-performance liquid chromatography using a fused-core column. *Food Chemistry*, 200, 167-174.

Ozaki, Y., & Liang, O.B. (1988). Cholagogic action of the essential oil obtained from *Curcuma xanthorrhiza* Roxb. *Shoyakugaku Zasshi*. 42, 257–263.

Padalia, R. C., Verma, R. S., Sundaresan, V., Chauhan, A., Chanotiya, C. S., & Yadav, A. (2013). Volatile terpenoid compositions of leaf and rhizome of *Curcuma amada* Roxb. from Northern India. *Journal of Essential Oil Research*, 25(1), 17-22.

Palmer, J. D. (1985). Evolution of chloroplast and mitochemdrial DNA in plants and algae. *Molecular Evolutionary Biology*, 131-240.

Pandey, A. D. I. T. I., & Katiyar, S. K. (2010). Determination and comparison of the curcuminoid pigments in turmeric genotypes (*Curcuma domestica* Val) by high-performance liquid

chromatography. *International Journal of Pharmacy and Pharmaceutical Sciences*, 2, 125-27.

Paramasivam, M., Poi, R., Banerjee, H., & Bandyopadhyay, A. (2009). High-performance thin layer chromatographic method for quantitative determination of curcuminoids in *Curcuma longa* germplasm. *Food Chemistry*, 113(2), 640-644.

Park, B. S., Kim, J. G., Kim, M. R., Lee, S. E., Takeoka, G. R., Oh, K. B., & Kim, J. H. (2005). *Curcuma longa* L. constituents inhibit sortase A and *Staphylococcus aureus* cell adhesion to fibronectin. *Journal of Agricultural and Food Chemistry*, 53(23), 9005-9009.

Pathak, R., Narang, P., Chandra, M., Kumar, R., Sharma, P. K., & Gautam, H. K. (2014). Homology Modeling and Comparative Profiling of Superoxide Dismutase Among Extremophiles: Exiguobacterium as a Model Organism. *Indian Journal of Microbiology*, 54(4), 450-458.

Paul, A., & Kumar, S. (2015). An A20/AN1-zinc-finger domain containing protein gene in tea is differentially expressed during winter dormancy and in response to abiotic stress and plant growth regulators. *Plant Gene*, 1, 1-7.

Pazos, F., & Valencia, A. (2001). Similarity of phylogenetic trees as indicator of protein–protein interaction. *Protein Engineering*, 14(9), 609-614.

Pearl, F., Todd, A., Sillitoe, I., Dibley, M., Redfern, O., Lewis, T., Bennett, C., Marsden, R., Grant, A., Lee, D. & Akpor, A. (2005). The CATH Domain Structure Database and related resources Gene3D and DHS provide comprehensive domain family information for genome analysis. *Nucleic Acids Research*, 33(suppl_1), D247-D251.

Pedersen, L.B. (2004). Phylogenetic analysis of the subfamily Alpinioideae (Zingiberaceae), particularly *Etlingera giseke*, based on nuclear and plastid DNA. *Plant Systematics and Evolution, 245,* 239-258.

Peitsch, M. C. (1995). Protein modelling by E-mail. *BioTechnology*, 13, 658–660

Peng, D., & Wang, X. Q. (2008). Reticulate evolution in Thuja inferred from multiple gene sequences: implications for the study of biogeographical disjunction between eastern Asia and North America. *Molecular Phylogenetics and Evolution*, 47(3), 1190-1202.

Peret-Almeida, L., Cherubino, A. P. F., Alves, R. J., Dufosse, L., & Gloria, M. B. A. (2005). Separation and determination of the physico-chemical characteristics of curcumin, demethoxycurcumin and bisdemethoxycurcumin. *Food Research International*, 38(8-9), 1039-1044.

Perry, R. P., & Kelley, D. (1972). The production of ribosomal RNA from high molecular weight precursors: III. Hydrolysis of pre-ribosomal and ribosomal RNA by a 3′-OH specific exoribonuclease. *Journal of Molecular Biology*, 70(2), 265-279.

Petersen, O. G. (1889). Musaceae, Zingiberaceae, Cannaceae. Marantaceae. *In. Engler, A. & Prantl, K.(Eds.) Die natürlichen Pflanzenfamilien,*, 2, 1-43.

Peterson, P. M., Romaschenko, K., & Johnson, G. (2010). A classification of the Chloridoideae (Poaceae) based on multi-gene phylogenetic trees. *Molecular Phylogenetics and Evolution*, 55(2), 580-598.

Petrey, D., Chen, T. S., Deng, L., Garzon, J. I., Hwang, H., Lasso, G., Lee, H., Silkov, A., & Honig, B. (2015). Template-based prediction of protein function. *Current Opinion in Structural Biology*, 32, 33-38.

Piana, M., Zadra, M., de Brum, T. F., Boligon, A. A., Gonçalves, A. F. K., da Cruz, R. C., de Freitas, R. B., Canto, G.S.D., & Athayde, M. L. (2012). Analysis of rutin in the extract and gel of Viola tricolor. *Journal of Chromatographic Science*, 51(5), 406-411.

Policegoudra, R. S., Aradhya, S. M., & Singh, L. (2011). Mango ginger (*Curcuma amada* Roxb.)–A promising spice for phytochemicals and biological activities. *Journal of Biosciences*, 36(4), 739-748.

Pramanik, K., Ghosh, P. K., Ray, S., Sarkar, A., Mitra, S., & Maiti, T. K. (2017). An in silico structural, functional and phylogenetic analysis with three dimensional protein modeling of alkaline phosphatase enzyme of Pseudomonas aeruginosa. *Journal of Genetic Engineering and Biotechnology*, 15(2), 527-537.

Prasad, S., Tyagi, A. K., & Aggarwal, B. B. (2014). Recent developments in delivery, bioavailability, absorption and metabolism of curcumin: the golden pigment from golden spice. *Cancer Research and Treatment: Official Journal of Korean Cancer Association*, 46(1), 2.

Prashanth, N., Yugander, A., & Bhavani, N. L. (2015). DNA Isolation and PCR Amplification of Turmeric Varieties from Telangana State. *International Journal of Current Microbiology and Applied Sciences,* 4(5), 485-490.

Prince, L. M., & Kress, W. J. (2006). Phylogenetic relationships and classification in Marantaceae: insights from plastid DNA sequence data. *Taxon*, 55(2), 281-296.

Prinitha, M. H. P., Madoen, J. S. S., & Jacod, A. (2012). Phytochemical characterization and antimicrobial activity of oil and solvent extracts of *Curcuma longa*. *Research Jornal of Pharmaceutical, Biological and Chemical Sciences*, 3, 49-59.

Priyadarsini, K. I. (2014). The chemistry of curcumin: from extraction to therapeutic agent. *Molecules*, 19(12), 20091-20112.

Provan, J., Powell, W., & Hollingsworth, P. M. (2001). Chloroplast microsatellites: new tools for studies in plant ecology and evolution. *Trends in Ecology and Evolution*, 16(3), 142-147.

Psifidi, A., Dovas, C. I., & Banos, G. (2010). A comparison of six methods for genomic DNA extraction suitable for PCR-based genotyping applications using ovine milk samples. *Molecular and Cellular Probes*, 24(2), 93-98.

Pyke, K.A. (1999). Plastid division and development. *The Plant Cell*, 11(4), 549-556.

Ramaswamy, S., Kuppuswamy, G., Dwarampudi, P., Kadiyala, M., Menta, L., & Kannan, E. (2014). Development and validation of simultaneous estimation method for curcumin and piperine by RP-UFLC. *Pakistan Journal of Pharmaceutical Sciences*, 27(4), 901-6.

Rafi, M., Wulansari, L., Heryanto, R., Darusman, L. K., Lim, L. W., & Takeuchi, T. (2015). Curcuminoid's Content and Fingerprint Analysis for Authentication and Discrimination of *Curcuma zanthorrhiza* from *Curcuma longa* by High-Performance Liquid Chromatography-Diode Array Detector. *Food Analytical Methods* 8, 2185–2193.

Raina, V. K., Srivastava, S. K., & Syamsundar, K. V. (2005). Rhizome and leaf oil composition of *Curcuma longa* from the lower Himalayan

region of northern India. *Journal of Essential Oil Research*, 17(5), 556-559.

Rajesh Kumar, T., Santhoshkumar, R., & Yusuf, A. (2016). Morphological characters and random amplified polymorphic DNA based genetic diversity analysis of *Curcuma* species (Zingiberaceae) from India. *International Journal* of *Plant*, *Animal and Environmental Sciences*; 6:4.

Rajesh Kumar, T., & Yusuf, A. (2017). Analysis by gas chromatography-mass spectrometry of the essential oil of *Curcuma inodora* Blatt. (Zingiberaceae) from Southern India. *Journal of Pharmacognacy and Phytochemistry*, 6(4), 1629-1634.

Rangsiruji, A., Newman, M. F., & Cronk, Q. C. B. (2000). A study of the infrageneric classification of *Alpinia* (Zingiberaceae) based on the ITS region of nuclear rDNA and the trnL-F spacer of chloroplast DNA. *Monocots: systematics and evolution. Collingwood, Australia: CSIRO*, 695-709.

Read, R. J., Adams, P. D., Arendall III, W. B., Brunger, A. T., Emsley, P., Joosten, R. P., Kleywegt, G.J., Krissinel, E.B., Lütteke, T., Otwinowski, Z., & Perrakis, A. (2011). A new generation of crystallographic validation tools for the protein data bank. *Structure*, 19(10), 1395-1412.

Reena, P., Sujata, M., & Sanghamitra, N. (2017). Molecular characterization of endangered medicinal plant species *Hedychium coronarium* from Eastern India. *International Journal of Pharmacy and Pharmaceutical Sciences*, 9(1), 173-178.

Reis, E. D. M., Neto, F. W. S., Cattani, V. B., Peroza, L. R., Busanello, A., Leal, C. Q., Boligon, A.A., Lehmen, T.F., Libardoni, M., Athayde, M.L., & Fachinetto, R. (2014). Antidepressant-like effect of Ilex paraguariensis in rats. *BioMed Research International*, 958209.

Revathy, S., Elumalai, S., & Antony, M. B. (2011). Isolation, purification and identification of curcuminoids from turmeric (*Curcuma longa* L.) by column chromatography. *Journal of Experimental sciences*, 2(7).

Rohman, A. (2012). Analysis of curcuminoids in food and pharmaceutical products. *International Food Research Journal*, 19(1), 19-27.

Rose, P. W., Prlić, A., Altunkaya, A., Bi, C., Bradley, A. R., Christie, C. H., Costanzo, L.D., Duarte, J.M., Dutta, S., Feng, Z. & Green, R. K. (2016). The RCSB protein data bank: integrative view of protein, gene and 3D structural information. *Nucleic Acids Research*, 4, 45(D1), D271-D281.

Roughley, P. J., & Whiting, D. A. (1973). Experiments in the biosynthesis of curcumin. *Journal of the Chemical Society, Perkin Transactions* 1, 2379-2388.

Roxburgh, W. (1810). Descriptions of several of the monandrous plants of India. *Asiatic Researches,* 1, 318–362.

Roxburgh, W. (1820). *Curcuma* Linn. *Flora Indica*, 1, 20-37.

Rubinoff, D., Cameron, S., & Will, K. (2006). A genomic perspective on the shortcomings of mitochondrial DNA for "barcoding" identification. *Journal of Heredity*, 97(6), 581-594.

Rubio-Pina, J. A., & Zapata-Perez, O. (2011). Isolation of total RNA from tissues rich in polyphenols and polysaccharides of mangrove plants. *Electronic Journal of Biotechnology*, 14(5), 11-11.

Sabu, M. (1991). A taxonomic and phylogenetic study of South Indian Zingiberaceae Ph.D Thesis, Department of Botany, University of Calicut. 201–243.

Sabu, M. (2006). Zingiberaceae and Costaceae of South India. *Association for Angiosperm Taxonomy*. University of Calicut, Kerala, 126-186

Sabu, M., & Mangaly, J.K. (1988). *Curcuma vamana* (Zingiberaceae): A new species from South India. *Journal of Economic and Taxonomic Botany*,12: 307-309.

Saha, K., Sinha, R. K., Basak, S., & Sinha, S. (2016). ISSR Fingerprinting to Ascertain the Genetic Relationship of *Curcuma* sp. of Tripura. *American Journal of Plant Sciences*, 7(02), 259.

Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K.B., & Erlich, H. A. (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science*, 239(4839), 487-491.

Salazar, G. A., Chase, M. W., Soto Arenas, M. A., & Ingrouille, M. (2003). Phylogenetics of Cranichideae with emphasis on Spiranthinae (Orchidaceae, Orchidoideae): evidence from plastid and nuclear DNA sequences. *American Journal of Botany*, 90(5), 777-795.

Sambrook, J., & Russell, D. W. (2001). Molecular Cloning: A Laboratory Manual, 3rd edn, *Cold Spring Harbor*.

Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989). Molecular cloning: a laboratory manual (No. Ed. 2). *Cold Spring Harbor Laboratory Press*.

Sandeep, I. S., Das, S., Nasim, N., Mishra, A., Acharya, L., Joshi, R. K., Nayak, S., & Mohanty, S. (2017). Differential expression of CURS gene during various growth stages, climatic condition and soil nutrients in turmeric (*Curcuma longa*): Towards site specific cultivation for high curcumin yield. *Plant Physiology and Biochemistry*, 118, 348-355.

Sandeep, I. S., Kuanar, A., Akbar, A., Kar, B., Das, S., Mishra, A., Sial, P., Naik, P. K., Nayak, S., & Mohanty, S. (2016). Agroclimatic zone based metabolic profiling of turmeric (*Curcuma longa* L.) for phytochemical yield optimization. *Industrial Crops and Products*, 85, 229-240.

Sanger, F., Nicklen, S., & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences*, 74(12), 5463-5467.

Sangha, J. S., Gu, K., Kaur, J., & Yin, Z. (2010). An improved method for RNA isolation and cDNA library construction from immature seeds of *Jatropha curcas* L. *BioMed Central Research Notes*, 3(1), 126.

Sanjur, O. I., Piperno, D. R., Andres, T. C., & Wessel-Beaver, L. (2002). Phylogenetic relationships among domesticated and wild species of *Cucurbita* (Cucurbitaceae) inferred from a mitochondrial gene: Implications for crop plant evolution and areas of origin. *Proceedings of the National Academy of Sciences*, 99(1), 535-540.

Santapau, H. (1945). *Curcuma pseudomontana* Grah. *Journal of the Bombay Natural History Society*, 45, 618-624.

Santhoshkumar, R., & Yusuf, A. (2017). DNA Barcoding and Phylogenetic analysis of South Indian *Curcuma* species using chloroplast *mat*K gene. *International Journal of Advanced Research,* 5(2), 615–621.

Santhoshkumar, R., & Yusuf, A. (2018). Molecular evaluation using two Chloroplast genes of South Indian *Curcuma* species: Insight in to Phylogenetic relationship. *International Journal of Research and Analytical Reviews*, 5(3), 498-504.

Santhoshkumar, R., & Yusuf, A. (2019). Chemotaxonomic studies on rhizome extract compositions of twenty Curcuma species from South India. *Biochemical Systematics and Ecology*, 84, 21-25.

Sarich, V. M., & Wilson, A. C. (1967). Immunological time scale for hominid evolution. *Science*, 158(3805), 1200-1203.

Sasikumar, B. (2005). Genetic resources of *Curcuma*: diversity, characterization and utilization. *Plant Genetic Resources*, 3(2), 230-251.

Sass, C., Little, D. P., Stevenson, D. W., & Specht, C. D. (2007). DNA barcoding in the cycadales: testing the potential of proposed barcoding markers for species identification of cycads. *PloS one*, 2(11), e1154.

Savolainen, V., & Chase, M. W. (2003). A decade of progress in plant molecular phylogenetics. *TRENDS in Genetics*, 19(12), 717-724.

Schonenberger, J., & Conti, E. (2003). Molecular phylogeny and floral evolution of Penaeaceae, Oliniaceae, Rhynchocalycaceae, and Alzateaceae (Myrtales). *American Journal of Botany*, 90(2), 293-309.

Schumann, K. (1904). Zingberaceae. *Das Pflanzenreich*, *4*(46), 1-458.

Schwede, T., Kopp, J., Guex, N., & Peitsch, M. C. (2003). SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Research*, 31(13), 3381-3385.

Searle, R. J., & Hedderson, T. A. J. (2000). A preliminary phylogeny of the Hedychieae tribe (Zingiberaceae) based on ITS sequences of the nuclear rRNA cistron. *Monocots: Systematics and Evolution. Collingwood, Australia: CSIRO*, 710-718.

Seberg, O., Petersen, G., Davis, J. I., Pires, J. C., Stevenson, D. W., Chase, M. W., Fay, M.F., Devey, D.S., Jørgensen, T., Sytsma, K.J. & Pillon, Y. (2012). Phylogeny of the Asparagales based on three plastid and two mitochondrial genes. *American Journal of Botany*, 99(5), 875-889.

Selvaraj, D., Sarma, R. K., & Sathishkumar, R. (2008). Phylogenetic analysis of chloroplast matK gene from Zingiberaceae for plant DNA barcoding. *Bioinformation*, 3(1), 24.

Semagn, K., Bjørnstad, A., & Ndjiondjop, M. N. (2006). An overview of molecular marker methods for plants. *African Journal of Biotechnology*, 5(25), 2540-2568.

Setyawan, A. D., Wiryanto, S., & Bermawie, N. (2014). Variation in isozymic pattern of germplasm from three ginger (*Zingiber officinale*) varieties. *Bioscience*, 6, 86-93.

Setyawan, A. D., Wiryanto, W., Suranto, S., Bermawie, N., & Sudarmono, S. (2014). Comparisons of isozyme diversity in local Java cardamom (*Amomum compactum*) and true cardamom (*Elettaria cardamomum*). *Nusantara Bioscience*, 6(1): 94-101.

Sha, L.N., Fan, X., Zhang, H.Q., Kang, H.Y., Wang, Y., Wang, X.L., Zhang, L., Ding, C. B., Yang, R.W., and Zhou, Y.H. (2014). Phylogenetic

relationships in *Leymus* (Triticeae; Poaceae): Evidence from chloroplast trnH-psbA and mitochondrial coxII intron sequences *Journal of Systematics and Evolution*. 52: 722–734.

Shanmugapriya, M., & Prabha, M. L. (2012). Isoenzyme analysis of *Amomum* species. *International Journal of Pharmaceutical and Biological Archive*, 3(5), 1091-1094.

Sharma, K., Agrawal, S. S., & Gupta, M. (2012). Development and validation of UV spectrophotometric method for the estimation of curcumin in bulk drug and pharmaceutical dosage forms. *International Journal of Drug Development and Research*, 4(2), 375-380.

Sharp, P. A., Sugden, B., & Sambrook, J. (1973). Detection of two restriction endonuclease activities in *Haemophilus parainfluenzae* using analytical agarose-ethidium bromide electrophoresis. *Biochemistry*, 12(16), 3055-3063.

Shaw, J., Lickey, E. B., Beck, J. T., Farmer, S. B., Liu, W., Miller, J., Siripun, K. C., Winder, C. T., Schilling, E. E., & Small, R. L. (2005). The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *American Journal of Botany*, 92(1), 142-166.

Shen, L., & Ji, H. F. (2012). The pharmacology of curcumin: is it the degradation products?. *Trends in Molecular Medicine*, 18(3), 138-144.

Shi, L. C., Zhang, J., Han, J. P., Song, J. Y., Yao, H., Zhu, Y. J., Li, J.C., Wang, Z.Z., Xiao, W., Lin, Y.L. & Xie, C.X. (2011). Testing the potential of proposed DNA barcodes for species identification of Zingiberaceae. *Journal of Systematics and Evolution*, 49(3), 261-266.

Shishodia, S., Sethi, G., & Aggarwal, B. B. (2005). Curcumin: getting back to the roots. *Annals of the New York Academy of Sciences*, 1056(1), 206-217.

Shreth, C. N., William, S. J., & Thorat, S. S. (2012). In-silico construction of phylogenetic tree to establish the relationship amongst family Zingiberaceae found in Manipur, North-east India, based on chloroplast *mat*K gene. *International Journal of Novel Trends in Pharmaceutical Sciences,* 2(5), 174-177.

Siddall, M. E. (1998). Success of parsimony in the four-taxon case: long-branch repulsion by likelihood in the Farris zone. *Cladistics*, 14(3), 209-220.

Silva, A. B. W. R., Herath, H., Senanayake, S. P., & Swarnathilaka, D. B. R. (2018). Phenetic and genetic characterization of selected economically important species in the family Zingiberaceae. *Sri Lankan Journal of Biology*, 3(1), 34-43.

Singh, S., Panda, M. K., & Nayak, S. (2012). Evaluation of genetic diversity in turmeric (*Curcuma longa* L.) using RAPD and ISSR markers. *Industrial Crops and Products*, 37(1), 284-291.

Siriluck, I., Ratchanok, T., Worakij, H., Thitamin, K., Saothongnoi, V., Amkha, S., Inubushi, K., Smakgahn, K., Sanjaya, Y., Ocampo, V.R., & Caoili, B. L. (2014). 1. Identification of 24 species of Zingiberaceae in Thailand using ISSR technique. *Thai Journal of Agriculatural Science*, 47(1), 1-6.

Sirirugsa, P., Larsen, K., & Maknoi, C. (2007). The genus *Curcuma* L.(zingiberaceae): distribution and classification with reference to

species diversity in Thailand. *Gardens' bulletin Singapore*, 59(1-2), 203-220.

Skornickova, J. (2007).Taxonomic studies in Indian *Curcuma* L. Dissertation, Charles University, Prague, Czech Republic.

Skornickova, J., & Sabu, M. (2005). The recircumscription of *Curcuma* L. to include the genus *Paracautleya* RM Sm. *Gardens' Bulletin Singapore*, 57, 37-46.

Skornickova, J., Sabu, M., & Prasanthkumar, M. G. (2004). *Curcuma mutabilis* (Zingiberaceae): a new species from South India. *Gardens' Bulletin, Singapore*, 56, 43-54.

Smith, J. F., Kress, W. J., & Zimmer, E. A. (1993). Phylogenetic analysis of the Zingiberales based on rbcL sequences. *Annals of the Missouri Botanical Garden*, 620-630.

Smith, R. M. (1977). new genus of Zingiberaceae from South India. *Notes from the Royal Botanic Garden, Edinburgh*, 35 (3), 365-368.

Smith, R. M. (1981). Synoptic keys to the genera of Zingiberaceae pro parte. *Royal Botanic Garden*, 2,1–28.

Sobrino, B., Bríon, M., & Carracedo, A. (2005). SNPs in forensic genetics: a review on SNP typing methodologies. *Forensic Science International*, 154(2-3), 181-194.

Soltis, D. E., & Soltis, P. S. (1998). Choosing an approach and an appropriate gene for phylogenetic analysis. In: *Molecular systematics of plants II* (pp. 1-42). Springer, Boston, MA.

Soltis, D. E., Kuzoff, R. K., Conti, E., Gornall, R., & Ferguson, K. (1996). matK and rbcL gene sequence data indicate that *Saxifraga*

(Saxifragaceae) is polyphyletic. *American Journal of Botany*, 83(3), 371-382.

Soltis, P. S., Soltis, D. E., & Chase, M. W. (1999). Angiosperm phylogeny inferred from multiple genes as a tool for comparative biology. *Nature*, 402(6760), 402.

Sporer, F., Sauerwein, M., & Wink, M. (1993). Diurnal and developmental variation of alkaloid accumulation in *Atropa belladonna*. In *WOCMAP I-Medicinal and Aromatic Plants Conference:* part 3 of 4 331 (381-386).

Srivastava, A. K., Srivastava, S. K., & Shah, N. C. (2001). Constituents of the rhizome essential oil of *Curcuma amada* Roxb. from India. *Journal of Essential Oil Research*, 13(1), 63-64.

Srivilai, J., Waranuch, N., Tangsumranjit, A., Khorana, N., & Ingkaninan, K. (2018). Germacrone and sesquiterpene-enriched extracts from *Curcuma aeruginosa* Roxb. increase skin penetration of minoxidil, a hair growth promoter. *Drug Delivery and Translational Research*, 8(1), 140-149.

Stanton, K. A., Edger, P. P., Puzey, J. R., Kinser, T., Cheng, P., Vernon, D. M., Forsthoefel, N. R., & Cooley, A. M. (2017). A whole-transcriptome approach to evaluating reference genes for quantitative gene expression studies: a case study in *Mimulus*. *G3: Genes, Genomes, Genetics*, 7(4), 1085-1095.

Steinmann, D., & Ganzera, M. (2011). Recent advances on HPLC/MS in medicinal plant analysis. *Journal of pharmaceutical and Biomedical Analysis*, 55(4), 744-757.

Storchova, H., & Olson, M. S. (2007). The architecture of the chloroplast psbA-trnH non-coding region in angiosperms. *Plant Systematics and Evolution*, 268(1-4), 235-256.

Swofford, D. L. (1998). *Paup 4.0 beta version for windows: phylogenetic analysis using parsimony*. Sunderland, Massachusetts: Sinauer Associates.

Syamkumar, S. (2008). Molecular biochemical and morphological characterization of selected *Curcuma* accessions. Ph D Thesis Calicut University, Calicut, India.

Syed, H. K., Liew, K. B., Loh, G. O. K., & Peh, K. K. (2015). Stability indicating HPLC-UV method for detection of curcumin in *Curcuma longa* extract and emulsion formulation. *Food chemistry*, 170, 321-326.

Taberlet, P., Coissac, E., Pompanon, F., Gielly, L., Miquel, C., Valentini, A., Vermat, T., Corthier, G., Brochmann, C., & Willerslev, E. (2007). Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding. *Nucleic Acids Research*, 35, e14

Taberlet, P., Gielly, L., Pautou, G., & Bouvet, J. (1991). Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology*, 17(5), 1105-1109.

Taheri, S., Abdullah, T. L., Abdullah, N. A. P., & Ahmad, Z. (2012). Genetic relationships among five varieties of *Curcuma alismatifolia* (Zingiberaceae) based on ISSR markers. *Genetics and Molecular Research*, 11(3), 3069-3076.

Takezaki, N., & Nei, M. (1994). Inconsistency of the maximum parsimony method when the rate of nucleotide substitution is constant. *Journal of Molecular Evolution*, 39(2), 210-218.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., & Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution*, 30(12), 2725-2729.

Tang, J. Y., Q. M. Li, R. W. Yang, J.Q. Liao & Y. H. Zhou (2008). Study on isozymes in six species of *Curcuma*. *China Journal of Chinese Materia Medica,* 33, 1381-1386.

Techaprasan, J., Klinbunga, S., & Jenjittikul, T. (2008). Genetic relationships and species authentication of *Boesenbergia* (Zingiberaceae) in Thailand based on AFLP and SSCP analyses. *Biochemical Systematics and Ecology*, 36(5-6), 408-416.

Tejavathi, D. H., Sujatha, B. S., & Kannan, R. (2017). Estimation of Curcuminoids in *Curcuma Karnatakensis* (White Turmeric)-an Endemic Taxon. *Asian Journal of Pharmaceutical and Clinical Research*, 10, 360-363.

Temin, H. M., & Mizutami, S. (1970). RNA-dependent DNA polymerase in virions of Rous sarcoma virus. *Nature*, 226, 1211-1213.

Theanphong, O., Mingvanish, W., & Kirdmanee, C. (2015). Chemical constituents and biological activities of essential oil from *Curcuma aeruginosa* roxb. Rhizome. *Bulletin of Health Science and Technology*, 13(1), 16.

Theanphong, O., Thanakijcharoenpath, W., Palanuvej, C., Ruangrungsi, N., & Rungsihirunrat, K. (2016). RAPD marker for determination of

phylogenetic relationships of 15 *curcuma* species from thailand. *Bulletin of Health, Science and Technology*, 14 (1) : 45-56.

Theerakulpisut, P., Triboun, P., Mahakham, W., Maensiri, D., Khampila, J., & Chantaranothai, P. (2012). Phylogeny of the genus *Zingiber* (Zingiberaceae) based on nuclear ITS sequence data. *Kew Bulletin*, 67(3), 389-395.

Tonnesen, H. H., & Karlsen, J. (1983). High-performance liquid chromatography of curcumin and related compounds. *Journal of Chromatography A,* 259, 367-371.

Tonnesen, H. H., & Karlsen, J. (1985). Studies on curcumin and curcuminoids. *A, European Journal of Food Research and Technology,* 180(5), 402-404.

Tonnesen, H. H., & Karlsen, J. (1986). Studies on curcumin and curcuminoids VII. Chromatographic separation and quantitative analysis of curcumin and related compounds. *European Food Research and Technology*, 182(3), 215-218.

Tramontano, A., Leplae, R., & Morea, V. (2001). Analysis and assessment of comparative modeling predictions in CASP4. *Proteins: Structure, Function, and Bioinformatics*, 45(S5), 22-38.

Uchegbu, R.I., Ngozi-Olehi, L. C., & Ogbuneke, R. U. (2014). Essential oils composition of *Curcuma longa* rhizome from Nigeria. *American Journal of Chemistry and Applications*, 1(1), 1.

Uma, E., & Muthukumar, T. (2014). Comparative root morphological anatomy of Zingiberaceae. *Systematics and Biodiversity*, 12(2), 195-209.

Valeton, T. (1918). New notes on the Zingiberaceae of Java and Malaya n archipelago. *Bulletin du Jardin Botanique de Buitenzorg.* 2, 27, 1-166.

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., & Speleman, F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology*, 3(7), research0034-1.

Vanijajiva, O., Suvachittanont, W., & Sirirugsa, P. (2003). Isozyme analysis of relationships among *Boesenbergia* (Zingiberaceae) and related genera in Southern Thailand. *Biochemical Systematics and Ecology*, 31(5), 499-511.

Varadarajan, R., Mathew, M. C. R., & Souprayan, S. (2018). Hepatoprotective efficacy of ethanolic extracts of rhizome *Curcuma amada* Roxb. in experimental rats. *Annals of Plant Sciences*, 71, 1966-1972.

Velankar, S., McNeil, P., Mittard-Runte, V., Suarez, A., Barrell, D., Apweiler, R., & Henrick, K. (2005). E-MSD: an integrated data resource for bioinformatics. *Nucleic Acids Research*, 33(suppl_1), D262-D265.

Velayudhan, K. C., Muralidharan, V. K., Amalraj, V. A., Gautam, P. L., Mandal, S., & Kumar, D. (1999). *Curcuma* genetic resources. *Scientific Monograph*, (4), 149.

Velayudhan, K.C., Amalraj, V.A. & Muralidharan, V.K. (1996). The conspectus of the genus *Curcuma* in India. *Journal of Economic and Taxonomic Botany* 20: 375–382.

Verma, N., & Shukla, S. (2015). Impact of various factors responsible for fluctuation in plant secondary metabolites. *Journal of Applied Research on Medicinal and Aromatic Plants*, 2(4), 105-113.

Vijayan, K., & Tsou, C. H. (2010). DNA barcoding in plants: taxonomy in a new perspective. *Current Science*, 1530-1541.

Vinitha, M. R., Kumar, U. S., Aishwarya, K., Sabu, M., & Thomas, G. (2014). Prospects for discriminating Zingiberaceae species in India using DNA barcodes. *Journal of Integrative Plant Biology*, 56(8), 760-773.

Virk, P. S., Zhu, J., Newbury, H. J., Bryan, G. J., Jackson, M. T., & Ford-Lloyd, B. V. (2000). Effectiveness of different classes of molecular marker for classifying and revealing variation in rice (*Oryza sativa*) germplasm. *Euphytica*, 112(3), 275-284.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., Lee, T. V. D., Hornes, M., Friters, A., Pot, J., Paleman, J., Kuiper, M., & Zabeau, M. (1995). AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research*, 23(21), 4407-4414.

Wahab, I. R. A., Blagojevic, P. D., Radulovic, N. S., & Boylan, F. (2011). Volatiles of *Curcuma mangga* Val. & Zijp (Zingiberaceae) from Malaysia. *Chemistry and Biodiversity*, 8(11), 2005-2014.

Wahyuni, S., Xu, D. H., Bermawie, N., Tsunematsu, H., & Ban, T. (2003). Genetic relationships among ginger accessions based on AFLP marker. *Jurnal Bioteknologi Pertanian*, 8(2), 60-68.

Wang, J. L., & Langenheim, J. H. (1990). Seasonal and diurnal variations in leaf sesquiterpenes of greenhouse-grown saplings of *Hymenaea* and *Copaifera*. *Acta Botanica Yunnanica*, 12(1), 85-91.

Wang, L., & Stegemann, J. P. (2010). Extraction of high quality RNA from polysaccharide matrices using cetyltrimethylammonium bromide. *Biomaterials*, 31(7), 1612-1618.

Wang, R. J., Cheng, C. L., Chang, C. C., Wu, C. L., Su, T. M., & Chaw, S. M. (2008). Dynamics and evolution of the inverted repeat-large single copy junctions in the chloroplast genomes of monocots. *BMC Evolutionary Biology*, 8(1), 36.

Wang, W., Wu, Y., Yan, Y., Ermakova, M., Kerstetter, R., & Messing, J. (2010). DNA barcoding of the Lemnaceae, a family of aquatic monocots. *BMC Plant Biology*, 10(1), 205.

Wang, Y. J., Pan, M. H., Cheng, A. L., Lin, L. I., Ho, Y. S., Hsieh, C. Y., & Lin, J. K. (1997). Stability of curcumin in buffer solutions and characterization of its degradation products. *Journal of Pharmaceutical and Biomedical Analysis*, 15(12), 1867-1876.

Waugh, J. (2007). DNA barcoding in animal species: progress, potential and pitfalls. *BioEssays*, 29(2), 188-197.

Webb, B., & Sali, A. (2017). Protein Structure Modeling with MODELLER. In *Functional Genomics*. Humana Press, New York, NY. 39-54.

Westbrook, J., Feng, Z., Chen, L., Yang, H., & Berman, H. M. (2003). The protein data bank and structural genomics. *Nucleic Acids Research*, 31(1), 489-491.

Westh, H., Zinn, C. S., Rosdahl, V. T., & Sarisa Study Group. (2004). An international multicenter study of antimicrobial consumption and resistance in Staphylococcus aureus isolates from 15 hospitals in 14 countries. *Microbial Drug Resistance*, 10(2), 169-176.

Wichitnithad, W., Jongaroonngamsang, N., Pummangura, S., & Rojsitthisak, P. (2009). A simple isocratic HPLC method for the simultaneous determination of curcuminoids in commercial turmeric extracts. *Phytochemical Analysis*, 20(4), 314-319.

Wijayasiriwardene, T. D. C. M. K., Herath, H. M. I. C., & Premakumara, G. A. S. (2017). Identification of endemic *Curcuma albiflora* Thw. by DNA barcoding method. *Sri Lankan Journal of Biology*, 2(1).

Williams, J. G., Kubelik, A. R., Livak, K. J., Rafalski, J. A., & Tingey, S. V. (1990). DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research*, 18(22), 6531-6535.

Williams, K. J., Kress, W. J., & Manos, P. S. (2004). The phylogeny, evolution, and classification of the genus *Globba* and tribe Globbeae (Zingiberaceae): appendages do matter. *American Journal of Botany*, 91(1), 100-114.

Willis, J. C. (1948). *A dictionary of the flowering plants and ferns*. Cambridge University Press.

Woese, C. R. (1987). Bacterial evolution. *Microbiological Reviews*, 51(2), 221.

Wolfe, K. H., Li, W. H., & Sharp, P. M. (1987). Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proceedings of the National Academy of Sciences,* 84(24), 9054-9058.

Wolfe, K. H., Morden, C. W., & Palmer, J. D. (1992). Function and evolution of a minimal plastid genome from a nonphotosynthetic parasitic plant. *Proceedings of the National Academy of Sciences*, 89(22), 10648-10652.

Wood, T. H., Whitten, W. M., & Williams, N. H. (2000). Phylogeny of *Hedychium* and related genera (Zingiberaceae) based on ITS sequence data. *Edinburgh Journal of Botany*, 57(2), 261-270.

Xi, Z., Wang, Y., Bradley, R. K., Sugumaran, M., Marx, C. J., Rest, J. S., & Davis, C. C. (2013). Massive mitochondrial gene transfer in a parasitic flowering plant clade. *PLoS Genetics*, 9(2), e1003265.

Xia, Y. M., Kress, W. J., & Prince, L. M. (2004). Phylogenetic analyses of *Amomum* (Alpinioideae: Zingiberaceae) using ITS and matK DNA sequence data. *Systematic Botany*, 29(2), 334-344.

Xiang, H., Zhang, L., Xi, L., Yang, Y., Wang, X., Lei, D., Zheng, X., & Liu, X. (2018). Phytochemical profiles and bioactivities of essential oils extracted from seven *Curcuma* herbs. *Industrial Crops and Products*, 111, 298-305.

Xiang, Z. (2006). Advances in homology protein structure modeling. *Current Protein and Peptide Science*, 7(3), 217-227.

Xiao, S., Ellena, J. F., Armstrong, G. S., & Capelluto, D. G. (2016). Structure of the GAT domain of the endosomal adapter protein Tom1. *Data in Brief*, 7, 344-348.

Zaveska, E., Fer, T., Sida, O., Krak, K., Marhold, K., & Leong-Skornickova, J. (2012). Phylogeny of *Curcuma* (Zingiberaceae) based on plastid and nuclear sequences: Proposal of the new subgenus *Ecomata*. *Taxon*, 61(4), 747-763.

Zaveska, E., Fer, T., Sida, O., Leong-skornickova, J. A. N. A., Sabu, M., & Marhold, K. (2011). Genetic diversity patterns in *Curcuma* reflect differences in genome size. *Botanical Journal of the Linnean Society*, 165(4), 388-401.

Zhang, D., Duan, L., & Zhou, N. (2014). Application of DNA barcoding in *Roscoea* (Zingiberaceae) and a primary discussion on taxonomic status of *Roscoea cautleoides* var. *pubescens*. *Biochemical Systematics and Ecology*, 52, 14-19.

Zhang, J., Jinnai, S., Ikeda, R., Wada, M., Hayashida, S., & Nakashima, K. (2009). A simple HPLC-fluorescence method for quantitation of curcuminoids and its application to turmeric products. *Analytical Sciences*, 25(3), 385-388.

Zhang, L., Yang, Z., Wei, J., Su, P., Pan, W., Zheng, X., Lin, L., Tang, J., & Du, Z. (2017). Essential oil composition and bioactivity variation in wild-growing populations of *Curcuma phaeocaulis* Valeton collected from China. *Industrial Crops and Products.* 103, 274-282.

Zhang, W., Yuan, Y., Yang, S., Huang, J., & Huang, L. (2015). ITS2 secondary structure improves discrimination between medicinal "Mu Tong" species when using DNA barcoding. *PloS one*, 10(7), e0131185.

Zhou, X., Li, Z., Liang, G., Zhu, J., Wang, D., & Cai, Z. (2007). Analysis of volatile components of *Curcuma sichuanensis* XX Chen by gas chromatography–mass spectrometry. *Journal of Pharmaceutical and Biomedical Analysis*, 43(2), 440-444.

Zietkiewicz, E., Rafalski, A., & Labuda, D. (1994). Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification. *Genomics*, 20(2), 176-183.

Zwaving, J. H., & Bos, R. (1992). Analysis of the essential oils of five *Curcuma* species. *Flavour and Fragrance Journal*, 7(1), 19-22.