Research Article **Molecular biology**

International Journal of Pharma and Bio Sciences **ISSN 0975-6299**

EVOLUTIONARY EVIDENCE FOR THE PREVALENCE OF TY3 GYPSY SUB-CLASS RETROTRANSPOSONS DURING ANNOTATION OF GERANIOL DEHYDROGENASE GENE IN JAVA CITRONELLA (*CYMBOPOGON WINTERIANUS* JOWITT)

AAKANKSHA WANY, VINOD KUMAR NIGAM AND DEV MANI PANDEY*

Department of Bio-Engineering, Birla Institute of Technology, Mesra, Ranchi, Jharkhand- 835215, India

ABSTRACT

When *GDH*- specific amplified product was annotated after three times of sequencing, it showed the similarity with Ty3 gypsy class retrotransposon protein of *O. sativa*, *Z. mays* and *S. bicolor*. Annotation was confirmed by the phylogenetic analysis performed between the new and LTR-RT sequences of the model plants. It was confirmed that *C. winterianus* have been evolved from *Z. mays* around 22.85 mya and thus, only a small subset of LTR sequence (6%) is showing similarity with *Z. mays* as compared to the 1083 bp *GDH* sequence. A new 1536 bp with gene coordinates <1…714, 730…909, 949…1407, 1426…1536> Ty3-gypsy class of RT containing mobile element sequence spanning from 1186-1536 bp with GenBank accession no. KF415292 in NCBI as *Cymbopogon winterianus* RT sequence was released. Thus, there may a prevalence of LTR-RT's across the *Citronella* genome as evidenced by the evolutionary history of *Citronella* being distantly related to wild *Z. mays*.

KEYWORDS: Annotation, GDH, LTR, Phylogenetic analysis, Retrotransposons, Sequencing

DEV MANI PANDEY
Department of Bio-Engineering, Birla Institute of Technology, Mesra, Ranchi, Jharkhand- 835215, India

***Corresponding author**

INTRODUCTION

Out of several aromatic grasses, *Cymbopogon* is one of the most important essential oil yielding genera belonging to the family Poaceae comprising 140 species worldwide, of which 45 species or more occur in India[18]. The name *Cymbopogon* is derived from the Greek words "kymbe" (boat) and "pogon" (beard), referring to the specific flower spike arrangement[25]. It is a native of Sri Lanka in the name of *Andropogon* with limited difference in some of the plant traits. It is reported that plants belonging to this genus might be a sub-genus of *Andropogon*[3]. The most important species of this sub-tribe (Andropogonae) is *Zea mays* (Teosinte grass-wild maize variety) and *Sorghum bicolor* which has evolved structures of vascular bundles and bundle sheath cells to exhibit C4 photosynthesis than C3 plants which grow in anaerobic conditions. *Cymbopogons* are tufted perennial C4 grasses with numerous stiff stems arising from a short, rhizomatous rootstock with distinct citronella flavour due to its essential oils which can be dried, powdered or even used fresh[23]. *Java Citronella* (*C. winterianus*; Figure 1), as compared to other species is the most common chemo cultivar displaying wide variation in morphological attributes and essential oil composition at inter and intra specific level[31]. Commercially, *Citronella* oil is classified into two chemotypes, Ceylon *Citronella* oil, obtained from *Cymbopogon nardus*, (inferior type), while Java type *Citronella* oil obtained from *Cymbopogon winterianus* (superior type). *C. winterianus* essential oil is rich in monoterpenes alcohols such as citronellal, geraniol, nerol and citronellol[10, 31] .

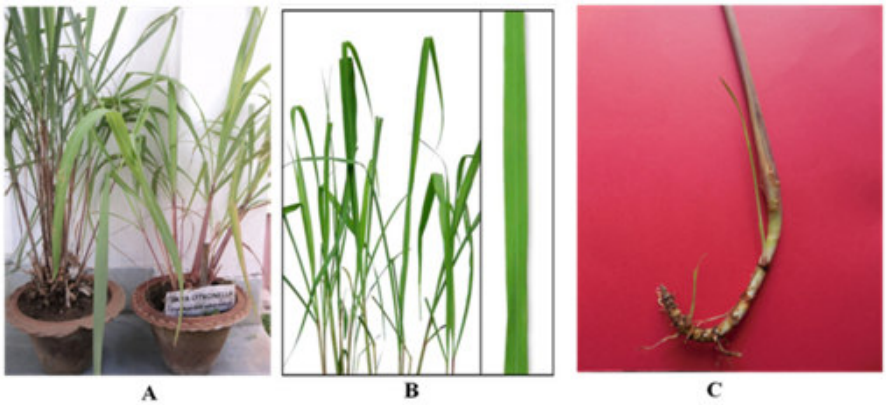


Figure 1
A. Potted plant of Cymbopogon winterianus B. Herbarium sheet C. Its vegetative slip

Irrespective of the advanced molecular biology techniques, very limited information about the plant is available, although, the different species have been well characterized on the basis of its essential oil components[33]. The major constrain in the thorough study of this plant is its least molecular characterization. Many literatures suggests that the variation in the percentage of essential oil components depends upon the age, length, shape (leaf) and physiological attributes of the chemo cultivar[10,30,31] .While searching through many literatures and databases, it was found that a single enzyme is responsible for the degradation of these essential oils in the geraniol degradation pathway- geraniol dehydrogenase (EC 1.1.1.183). The gene (*GDH*) responsible for this enzyme was aimed to be blocked so that there is an accumulation of essential oils through RNA*i* approaches and inhibitors[33]. The isolation of *GDH* gene and *in silico* characterization using modified DNA extraction protocol in *O. americanum* has been reported[32]. To verify the existence of *GDH* gene from ancient times and its evolutionary relationship with *ADH-1* gene of model plants, phylogenetic study was carried out. Also, the relationship among *Cymbopogon* and *Andropogon* genus members was checked. Therefore, present investigation deals with the isolation and characterization of the *GDH* gene from *Citronella* to block the gene, thereby enriching the essential oil components using modified

DNA extraction protocol. Also, the modified protocol for extracting the gDNA has been standardized with respect to seasons and age of the plant. Characterization of *GDH* gene via post-sequencing analysis revealed the presence of LTR-RT's. Hence, their phylogenetic analysis with known sequences of model plants has been conducted to confirm the evolutionary history.

2. MATERIALS AND METHODS

The leaves of Java Citronella were collected from the Pharmaceutical Indigenous Medicinal Plant Research Farm, BIT, Mesra. A herbarium was sent to the Botanical Survey of India, Kolkata.

2.1 Identification of GDH gene and primer designing

Identification and screening of the geraniol dehydrogenase gene (*GDH*) through a metabolic pathway available in KEGG database was done (Figure 2). Six sets of gene- specific primers and a check primer (ADH- 1) were designed (Primer3). Also, two sets of degenerate primers using Hyden v1.0 (http://acgt.cs.tau.ac.il/hyden/HYDEN.html) were designed. The third set of degenerate primer (DPGDH-3) was designed by Dr. A. Kohli, Senior Scientist, IRRI, Philippines. All primer sets were procured from Metabion International AG, Germany and summarized in Table 1.

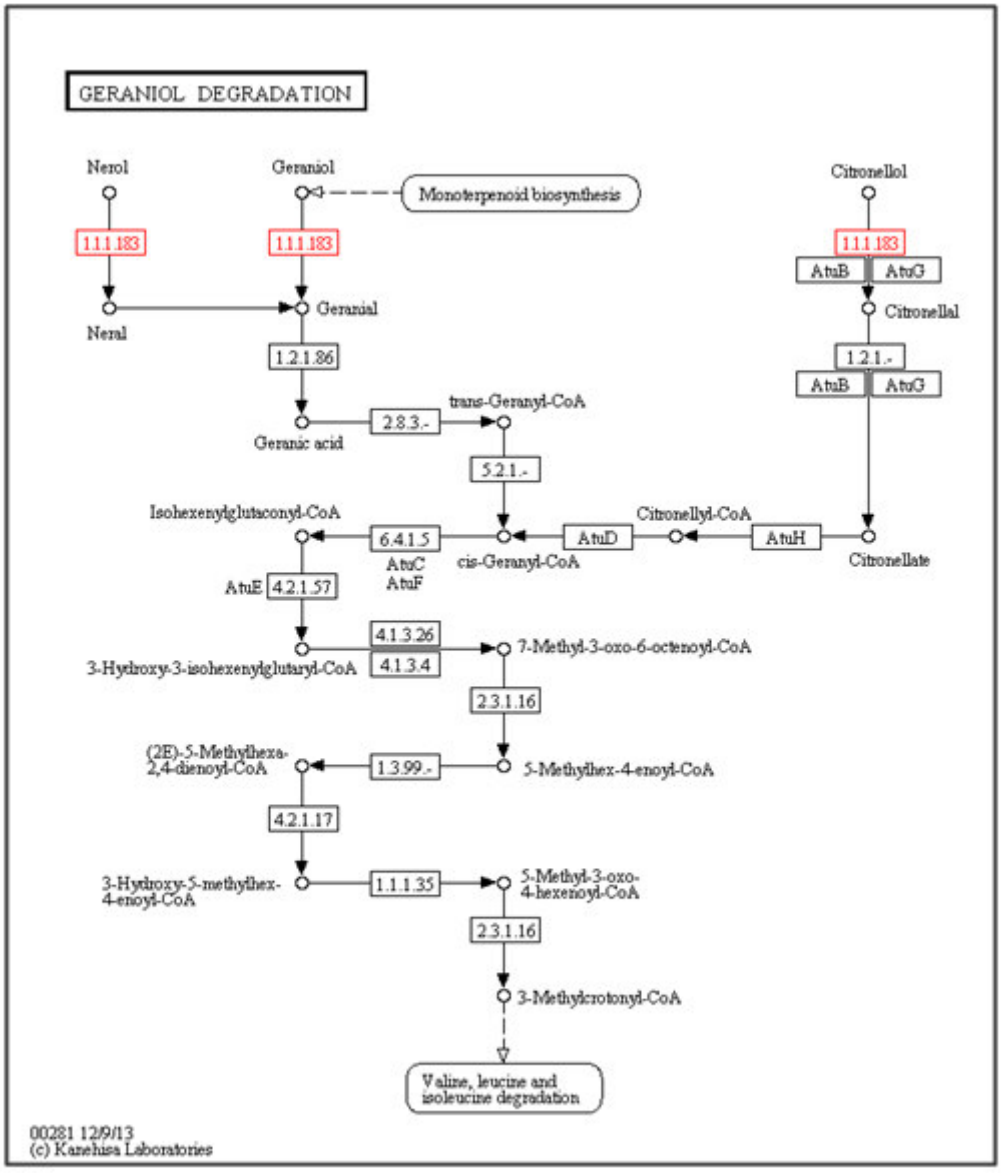


Figure 2
Geraniol degradation pathway

Table 1
Primers procured for identification of GDH gene as sequencing proceeded

Sequencing	Name	Primers	Annealing Temperature (°C)	Amplicon Size (bp)
First sequencing	ObGDH	GAAATCACCAGAAACAGAGC (Forward)	55	1083
		TCACTGAGGTTTCAAGGATT (Reverse)		
	ObGDH-1	GAAATCACCAGAAACAGAGC (Forward)	54	1047
		ATGACGAAACGGTACTTGAC (Reverse)		
	ObGDH-2	GCCTTGGGATGGGCAGCCAC (Forward)	64	892
		GCGTCTCCTTAAGCCCGCCG (Reverse)		
Second sequencing & Third sequencing	OaGDH-3	CGAGGATACCGTCATGGAAG (Forward)	45	705
		GTCCGTGAAGTGATCCCAAT (Reverse)		
	OaGDH-4	TGTCTCTTGCGTCCTTTGTG (Forward)	46	700
		CTCCGGGTGGAAATCTTTTT (Reverse)		
	OaGDH-5	ACGTCATAGCCCCAGAATTG (Forward)	45	702
		TGGCCAATCATCACATTGTT (Reverse)		
	DPGDH 1	CAYCCHRNDMAGGCHTTBGGATGGGC (Forward)	55	
		TTDGCAAAYTTBACDGCYAYATGDCC(Reverse)		
	DPGDH 2	GGHCATRTRGCHGTVAARTTTGCHAA (Forward)	52	
		GCYTTDGCAAAYTTBACDGCYAYATG (Reverse)		
	DPGDH 3	TAYTGTGGMGTTTGYCAYWCAG (Forward)	55	
		CATYTCTTGCGTCTCCTTMAKYC (Reverse)		
Check primer	ADH-1	CCAGTTCAGCAGGTACTTGC(Forward) CAGGATACACAGAAGAACCG(Reverse)	55	484

2.2 Genomic DNA isolation and purification

In this study, a range of concentrations of CTAB, PVP and β- mercaptoethanol along with an extended RNase treatment have been studied. In this modified protocol, 0.8-1 g of fresh leaf samples were powdered using liquid nitrogen and homogenized in extraction buffer. The RNA contamination was removed by treating DNA samples with RNase Cocktail according to the protocol provided (Invitrogen, India). The rest of the protocol including the purification of genomic DNA was followed as described in our previous work[32]. The DNA isolation protocol was optimized (in triplicates) in different conditions such as collecting the leaf samples from different seasons and age of the plants (Y1-7 days, Y2- 21 days and M-more than 30 days old).

2.3 Total RNA isolation and cDNA synthesis

Total cellular RNA was extracted from the fresh leaves (1g) of *Citronella* (grounded with liquid nitrogen) using TRIzol reagent® (MRC Inc., USA) following the standard protocol. The DNA contamination was removed using RNase free DNase I (Sigma Aldrich, USA) following the manufacturer's instructions. RNA quality was checked on a 1% TAE-agarose gel with ethidium bromide and formaldehyde. Isolated total RNA was quantified with a biophotometer (Eppendorf, USA). Samples with over 1.9, 260/280 and 260/230 ratios were used. Subsequently, full length cDNA was synthesized using Omniscript RT kit (Qiagen, USA) from the purified RNA as per manufacturer's protocol.

2.4 PCR amplification and GDH annotation

PCR amplification of *GDH* was performed using TaKaRa™ mix, Japan. The PCR protocol for standard as well as optimized conditions was followed as per previous work [32]. Amplified products were resolved in 1.3% agarose gel and photographed in Gel documentation system (Syngene Bioimaging Pvt. Ltd. India). PCR optimization was done with respect to the number of cycles, annealing temperature and the concentration of template DNA. The best primer set which gave single band on amplification was sent for sequencing to Chromous Biotech, Bangalore, India. The sequencing was carried out three times after obtaining the desired amplicon size from all the primers. The structural and functional annotation of *GDH* gene was done using different online gene annotation workbenches.

2.5 Phylogenetic analysis

To establish whether there is any evolutionary relationship between the retrotransposon families of model plants and *Cymbopogon winterianus* retrotransposon Ty3-gypsy subclass, phylogenetic analysis using MEGA 5.2 was performed. Also, another set of gene family (*ADH-1*) of the above model plants was studied to have a phylogenetic relationship with *GDH* family of reported plants to confirm that whether geraniol dehydrogenase have been evolved from alcohol dehydrogenase superfamily. The sequences for this analysis were retrieved from their respective databases. Therefore, linearized trees were constructed by the neighbour-joining method[26].

3. RESULTS

The plant specimen was confirmed as *Cymbopogon winterianus* Jowitt (Specimen no- 12235) from Botanical Survey of India, Kolkata which is being cultivated in our Institute.

3.1 Identification of GDH gene

In the geraniol degradation pathway, only single enzyme geraniol dehydrogenase (EC 1.1.1.183) is responsible for the degradation of geraniol, nerol and citronellol (KEGG database: http://www.genome.jp/kegg) (Figure 2) and hence its encoding gene was further, aimed to be isolated and characterized using gene specific and degenerate primer sets.

3.2 Isolation of genomic DNA and its optimization studies

For the entire study, the plant material was collected randomly in triplicates from the three differently localized fields of *Citronella.* A range of 2% - 3.5% CTAB and 3% PVP was studied and reported in Table 2 (Figure 3). The same protocol was followed with three different varieties of *Ocimum* (*O. basilicum*, *O. americanum and O. sanctum*) which was taken as a control (Figure 3 and 4). The gDNA has been isolated from the three different stages Y1, Y2 and M plants (Figure 5 and 6; Table 3). The results obtained after collecting leaves from different seasons and its gDNA isolation is reported in Table 4, Figure 7.

Table 2
Isolation of genomic DNA with different concentrations of CTAB, PVP and β-ME from Citronella

Treatment	Band	CTAB (%)	PVP (%)	β-ME (%)	$A_{260/280}$	Colour of the pellet	Yield (µg/g)* FW
Without RNase treatment	1	2	2	0.2	1.37	Brown	2138
	2	2.5	2	0.2	1.56	Brown	2196
	3	3	3	0.3	1.65	Light Yellow	2233
	4	3.5	3	0.3	1.81	White	2978
With RNase treatment	5	2	2	0.2	1.71	Light Brown	1380
	6	2.5	2	0.2	1.79	Light Brown	1517
	7	3	3	0.3	1.80	White	1528
	8	3.5	3	0.3	1.89	Transparent	1535

** indicates average readings in triplicates of all the fields*

Figure 3
Genomic DNA isolated from Basil and Citronella leaves, C- C. winterianus, Ob- O. basilicum Oa- O. americanum and Os- O. sanctum (a) 2% CTAB + 2% PVP (b) 3.5% CTAB + 3% PVP {RNase treatment}

Figure 4
RNase treated gDNA isolated from O. americanum, O. basilicum and C. winterianus leaves

Table 3
Genomic DNA isolated from different stages of leaf till maturity in Citronella

Plant age	Fields	$A_{260/280}$	Yield (µg/g)* FW
First leaf stage (Y1)	Field number 1	1.58	2845
	Field number 2	1.56	
	Field number 3	1.57	
Two leaf stage (Y2)	Field number 1	1.66	1894
	Field number 2	1.68	
	Field number 3	1.69	
Mature leaf stage (M)	Field number 1	1.78	1597
	Field number 2	1.80	
	Field number 3	1.79	

**** indicates average readings in triplicates of all the fields***

Figure 5
Leaves sampled from different stages studied for gDNA isolation from different fields

Figure 6
RNase treated gDNA isolated from Citronella leaves where Y1- First leaf, Y2- second leaf and M- Mature leaf from different plants

Table 4
Genomic DNA isolated during different seasons with the modified protocol

Season	Treatment	Yield (µg/g)* FW	$A_{260/280}$	Colour of the pellet
March 2012 (summer)	Without RNase treatment	2318.12	1.65	Cream
	With RNase treatment	1315.48	1.71	Transparent
July 2012 (monsoon)	Without RNase treatment	2199.33	1.70	White
	With RNase treatment	1528.45	1.79	Transparent
December 2012 (winter)	Without RNase treatment	2985.68	1.72	White
	With RNase treatment	1632.14	1.81	Transparent

**** indicates average readings in triplicates***

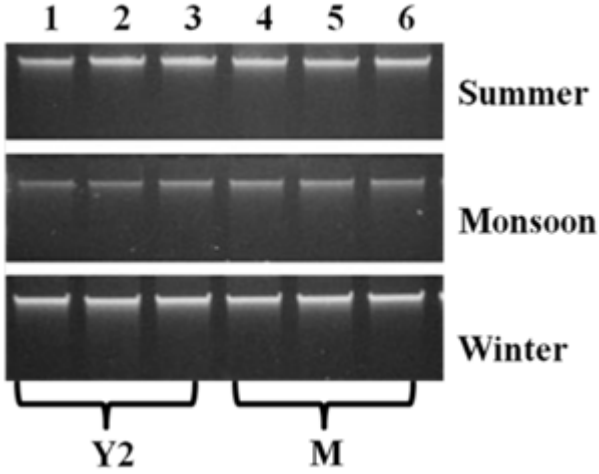


Figure 7
RNase treated gDNA isolated during different seasons with modified protocol where 1, 2, 3- Young leaves (Y2); 4, 5, 6- Mature leaves (M) from Citronella

3.3 Total RNA extraction

High-quality total cellular RNA was obtained by following the standard protocol using TRIzol® reagent[20]. The yields of total RNA (µg /g fresh weight [FW]) were 1179.88 and 1099.45 in *Ocimum* and *Citronella* respectively. This indicated that the RNA was of high purity and devoid of polyphenols and polysaccharide contamination (Table 5). For both the RNA samples, distinct rRNA bands without degradation were observed (Figure 8).

Table 5
Quantification of RNA using Eppendorf® BioPhotometer

Plant	Sample Type	$A_{260/230}$	$A_{260/280}$	Yield (µg/g)*
O.americanum	RNA	2.34	2.02	1179.88
C. winterianus	RNA	2.24	1.96	1099.45

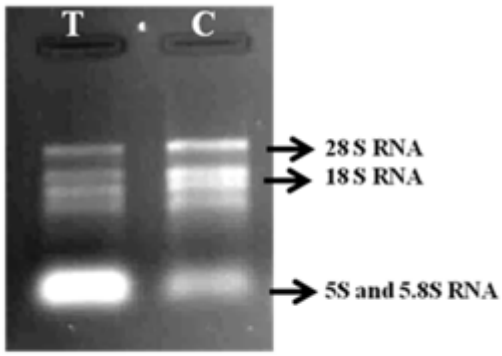


Figure 8
Denaturing gel of total cellular RNA isolated from Basil (T) and Citronella (C) leaves

3.4 PCR amplification and sequencing

Amplification was carried out with the gene specific primers and degenerate primers listed in Table 1. Out of nine sets of primers procured along with one set of check primer, only three sets of primers gave amplification at optimum conditions viz., ObGDH-1 (first sequencing), ObGDH-1 and 2 (second

sequencing) and OaGDH-3 (third sequencing) (Figure 9). The check primer ADH-1 gave amplification with the gDNA template and the cDNA template both. The annealing temperature in the PCR thermal profile was 54°C with ObGDH-1 (1000 bp); 58°C with ObGDH-2 (892 bp) and 45°C with OaGDH-3 (700 bp) primers. ObGDH-1 and OaGDH-3 gave single bands but ObGDH-2 gave multiple bands after number of optimization conditions, hence, all the multiple bands were sent to sequencing in the second time.

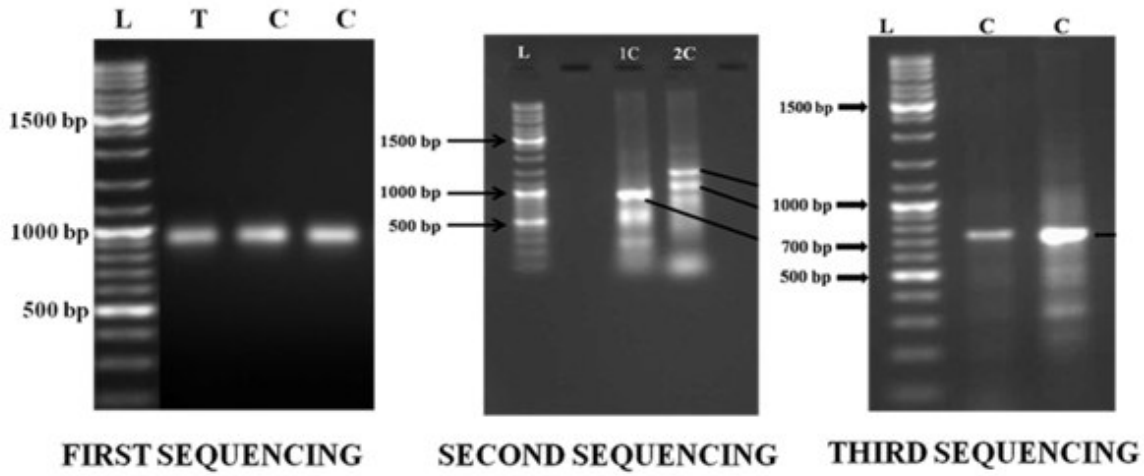


Figure 9
Gels sent for sequencing to Chromous Biotech Pvt. Ltd. Bangalore, India

Degenerate PCR was carried out with three sets of degenerate primers under standard and optimized conditions using cDNA template. Under standard amplifying conditions, lower annealing temperatures (40 -55°C) were tried at 35- 50 cycles. DPGDH-1 and 2 primers gave no amplification in any conditions. A faint band was observed at 95°C- 2 minutes, 95°C -15 seconds, 52°C - 30 seconds, 72°C - 1.30 minutes, 72°C -7 minutes with 42 cycles using DPGDH-3 primer but was unable to elute due to its less quantity. Under optimized conditions, where first four cycles were kept at lower gradients (42-52°C) are supposed to anneal with multiple mismatches with the chance of annealing to the target gene and then followed with 38-40 cycles of higher gradients (56-62°C) at same conditions. The result again showed feeble bands which are very difficult to elute and purify. The reason behind faint bands may be the degeneracy of the primers (512 mixes or more) or competitive inhibition leading to non-specificity and consequently no amplification. This means that the first PCR cycles are very inefficient, and even after 50 cycles a faint band of the target gene is seen[17]. Hence, it can be said that, these contain conserved dehydrogenase domains which can be easily detected by degenerate or mixed primers. Therefore, a feeble band with slight amplification was not sent for sequencing (Figure 10).

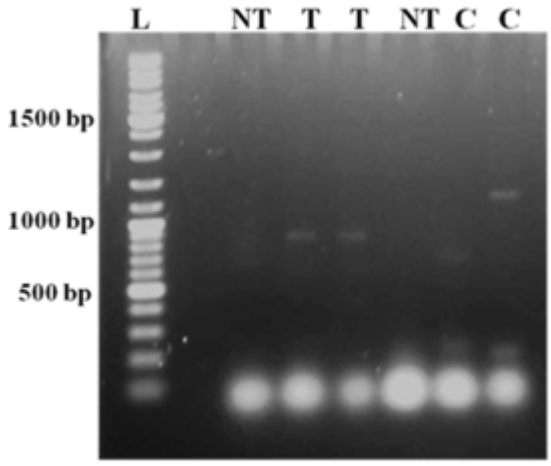


Figure 10
Slight amplification of degenerate primer (DPGDH-3) primer in Basil and Citronella with No template control (NTC)

3.5 Annotation of GDH gene

3.5.1. Structural annotation

The structural annotation of *GDH* gene started with similarity searches using BLAST and BLASTx. All the hits were of either hypothetical, uncharacterized polyprotein or retrotransposon protein Ty3 gypsy subclass *Oryza sativa* Japonica group (ABF96514.1). The assembled contig (DnaBaser v3.5.3) gave a single hit with *Sorghum bicolor* having 72% identity and only 27% query coverage. Simultaneously, BLASTx of contig was performed which showed the similarity with RT protein, putative, Ty3-gypsy subclass (*Oryza sativa* Japonica Group) with 55% identity and 82% query coverage. After sequence assembly and complete annotation, a new sequence of CwRTR was constructed and named on *Cymbopogon winterianus* as mobile element type retrotransposon nucleotide sequence. A 1536 bp length of new sequence was constructed with gene coordinates <1..714, 730..909, 949...1407, 1426..1536 > as validated by Wise2 program[2]. Also, when this newly constructed sequence was scanned for the presence of domains with InterProScan (http://www.ebi.ac.uk/Tools/pfa/iprscan/), it showed putative retrotransposon protein Ty3 gypsy subclass *Oryza sativa* Japonica group and also, contained reverse transcriptase, RNase-H like domains and integrase domains (Supplementary Figures 1-8). After three times of sequencing, it was confirmed that the genome of *Citronella* contains retrotransposon sequence sparsing or either flanking *GDH* gene and was submitted in NCBI (KF415292-Supplementary information) in the name of *Cymbopogon winterianus* Ty3-gypsy subclass retrotransposon, complete sequence.

3.5.2. Functional annotation

It confines mainly to assign the function to the gene or its protein or to find the regulatory motifs referring to molecular function and biological processes. This study was done using I-TASSER server which predicts the 3D structure of the target protein[21]. It predicted five best 3D models of the given protein. The C-Score for the best predicted model and TM-score was observed as (-0.23, 0.68±0.12) (Figure 11). As per Gene Ontology terms, the molecular function of the protein suggests its selective non-covalent interaction with a DNA or RNA sequence in order to modulate transcription. This may or may not also interact selectively with a protein or macromolecular complex (GO: 0001071) and is a cellular response to stimulus (GO: 005171614) as biological process. Thus, it was confirmed that by function, it is a retrotransposon protein which modulates the transcription of the gene/s by its selective interaction with DNA or RNA and arise in response to stimulus such as wound, pathogenic attack or biotic and abiotic stresses.

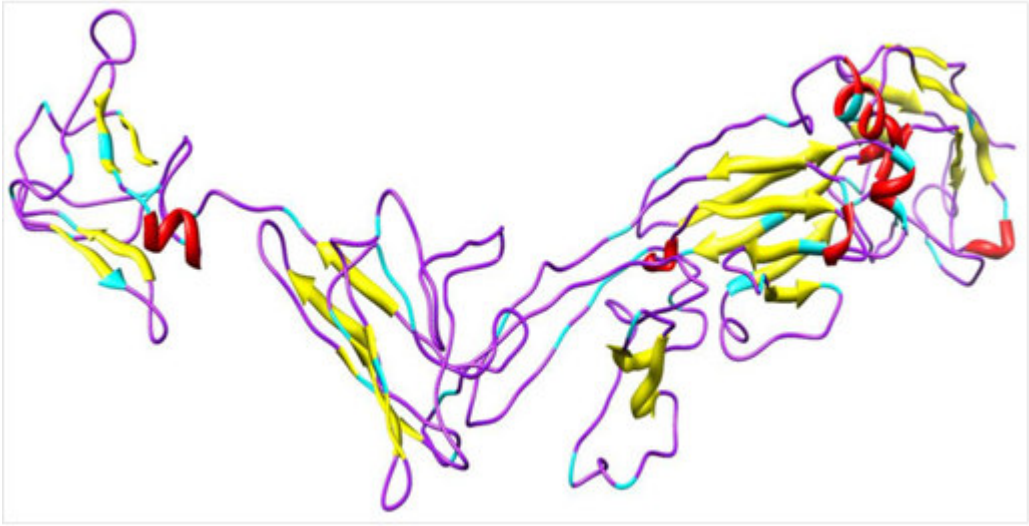


Figure 11
Predicted 3D model (I-TASSER) of CwRTR encoded protein where α-Helices are shown in red colour, β-Sheets are in yellow and coils are in purple

3.6 Phylogenetic analysis

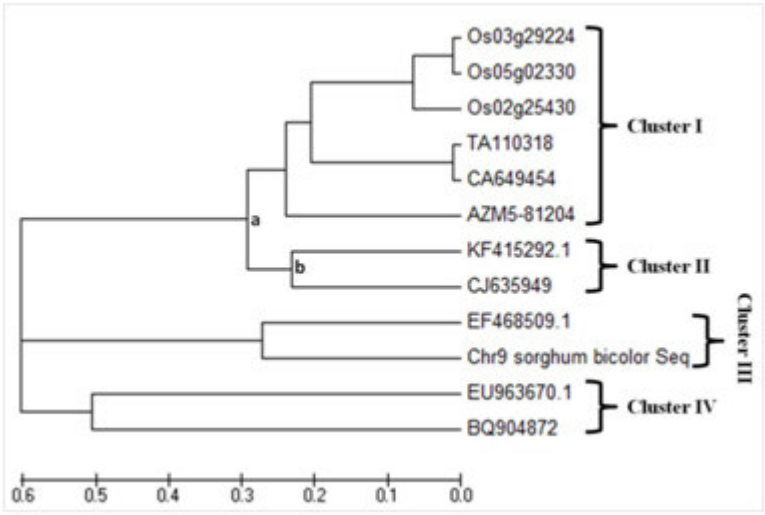
Evolutionary relationship analysis was conducted between retrotransposon families (Table6; Figure 12) and *ADH/GDH* gene families (Table 7: Figure 13) between corresponding sequences of model plants and submitted sequence of *GDH* gene i.e., JQ765502 and Ty3 gypsy subclass RT i.e., KF415292 constructing individual linearized trees using MEGA5.2 software[26].

Table 6
Description of the retrotransposon sequences of model plants represented in clusters of the phylogenetic tree

Clusters	GenBank Id/Locus Id/ Source	Accession no.	Description
Cluster I	Os03g29224		*Oryza sativa* ssp japonica cv. Nipponbare_retrotransposon protein, putative, Ty3-gypsy subclass, expressed
	Os05g02330		*Oryza sativa* ssp japonica cv. Nipponbare_retrotransposon protein, putative, Ty3-gypsy subclass, expressed
	Os02g25430		*Oryza sativa* ssp japonica cv. Nipponbare_retrotransposon protein, putative, Ty3-gypsy subclass, expressed
	TA110318		Putative GAG-POL [*Aegilops tauschii* (Tausch's goatgrass) (Aegilops squarrosa)]
	CA649454		Putative GAG-POL [*Aegilops tauschii* (Tausch's goatgrass) (Aegilops squarrosa)]
	AZM5_81204		*Zea mays*_rpt_family="grande" rpt_super_family="[RLG] Gypsy"
Cluster II	gi|557132934	KF415292.1	*Cymbopogon winterianus* retrotransposon Ty3-gypsy subclass, complete sequence
	CJ635949		Retrotransposon protein, putative, Ty3-gypsy sub-class [*Oryza sativa* (japonica cultivar-group)]
Cluster III	gi|149930441	EF468509.1	*Zea mays* clone pB3-202 retrotransposons GrandeB, complete sequence
	Sorghum bicolor v1.4*		Chromosome_9_*Sorghum bicolor*:21019436..21019812
Cluster IV	gi|195627915	EU963670.1	*Zea mays* clone 265470 retrotransposon protein Ty3-gypsy subclass mRNA, complete cds
	BQ904872		Hypothetical protein [*Solanum tuberosum* (Potato)]

Source of the sequence- http://www.phytozome.net/sorghum

Figure 12
listed in Linearized tree of the retrotransposon sequences from different model plants as Table-6.



Note: The scale bar corresponds to 0.1 estimated amino acid substitutions per site)

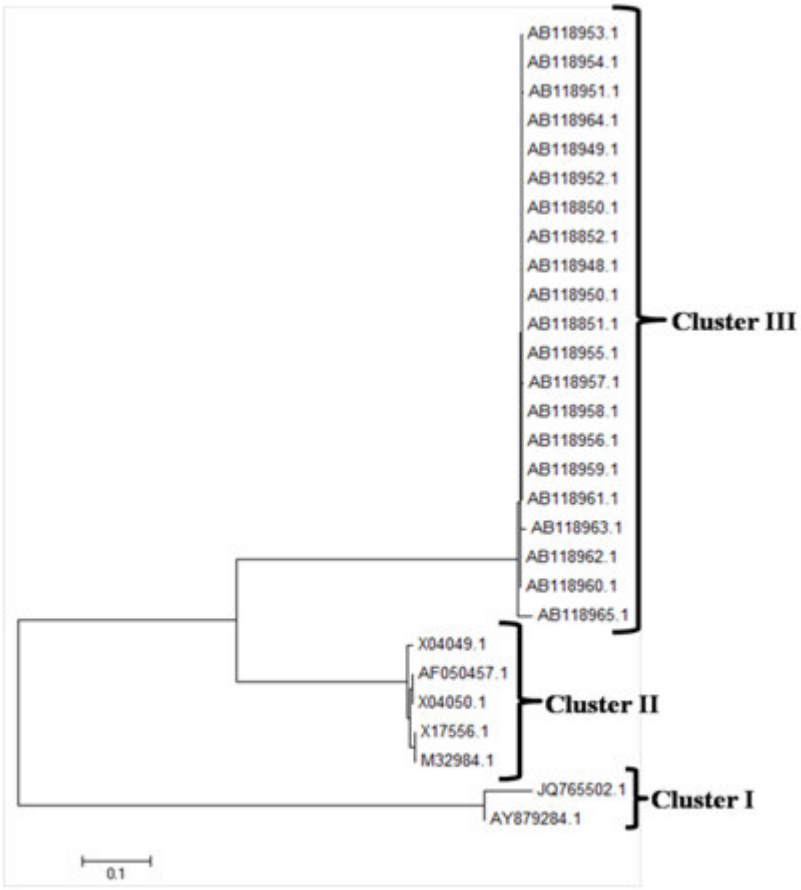
The branch length estimation of linearized tree in Figure 12 depicted that the branch leading to cluster IV is significantly longer than the branches for the other clusters which signifies faster evolution than the other clusters (Table 6). Though, substitution rates among clusters do not have considerable difference. Cluster II consists of *Cymbopogon winterianus* retrotransposon Ty3-gypsy subclass sequence (KF415292) which diverged (node b) faster from *O. sativa* retrotransposon protein putative Ty3-gypsy sub-class (CJ635949) with less substitution rate having split time of 22.85 million years ago (Mya). Again, divergence from the cluster I with splitting time of 29.86 Mya (node a) suggesting slower evolution was observed. Thus, it can be concluded that this class of RT have been diverged from *O. sativa* retrotransposon protein and evolved from *Z. mays* in minimum evolutionary time of 23 mya.

Table 7
Description of the ADH/ GDH gene sequences of model plants represented in clusters of the phylogenetic tree

Clusters	GenBank Id	Accession no.	Description
Cluster I	gi|389889214	JQ765502.1	*Ocimum* americanum geraniol dehydrogenase 1 (GDH) gene, complete cds
	gi|62461967	AY879284.1	*Ocimum* basilicum geraniol dehydrogenase (GEDH) mRNA, complete cds
Cluster II	gi|168406	M32984.1	*Z.mays* alcohol dehydrogenase (ADH-1 C-m allele) gene, complete cds
	gi|22118	X17556.1	*Z.mays* DNA for Adh1-Cm allele
	gi|22121	X04050.1	Maize alcohol dehydrogenase 1 gene (Adh1-1F)
	gi|22123	X04049.1	Maize alcohol dehydrogenase 1 gene (Adh1-1S)
	gi|3420020	AF050457.1	*Zea mays* alcohol dehydrogenase 1 (adh1) gene, adh1-F allele, complete cds
Cluster III	gi|34787328	AB118959.1	*Oryza rufipogon* Adh1 gene for alcohol dehydrogenase I, complete cds, strain:VT51
	gi|34787322	AB118956.1	*Oryza rufipogon* Adh1 gene for alcohol dehydrogenase I, complete cds, strain:W1956
	gi|34787332	AB118961.1	*Oryza rufipogon* Adh1 gene for alcohol dehydrogenase I, complete cds, strain:CB27
	gi|34787334	AB118962.1	*Oryza sativa* Indica Group Adh1 gene for alcohol dehydrogenase I, complete cds, strain:435
	gi|34787320	AB118955.1	*Oryza rufipogon* Adh1 gene for alcohol dehydrogenase I, complete cds, strain:SN
	gi|34787330	AB118960.1	*Oryza rufipogon* Adh1 gene for alcohol dehydrogenase I, complete cds, strain:CT56
	gi|34787316	AB118953.1	*Oryza rufipogon* Adh1 gene for alcohol dehydrogenase I, complete cds, strain:W593
	gi|34787314	AB118952.1	*Oryza rufipogon* Adh1 gene for alcohol dehydrogenase I, complete cds, strain:CB22
	gi|34787318	AB118954.1	*Oryza rufipogon* Adh1 gene for alcohol dehydrogenase I, complete cds, strain:YG2A
	gi|34787324	AB118957.1	*Oryza rufipogon* Adh1 gene for alcohol dehydrogenase I, complete cds, strain:KA
	gi|34787310	AB118950.1	*Oryza rufipogon* Adh1 gene for alcohol dehydrogenase I, complete cds, strain:PT1A
	gi|34787304	AB118852.1	*Oryza rufipogon* Adh1 gene for alcohol dehydrogenase I, complete cds, strain:W1972

	gi|34787308	AB118949.1	*Oryza rufipogon* Adh1 gene for alcohol dehydrogenase I, complete cds, strain:W120
	gi|34787312	AB118951.1	*Oryza rufipogon* Adh1 gene for alcohol dehydrogenase I, complete cds, strain:W108
	gi|34787306	AB118948.1	*Oryza rufipogon* Adh1 gene for alcohol dehydrogenase I, complete cds, strain:W1976
	gi|34787300	AB118850.1	*Oryza rufipogon* Adh1 gene for alcohol dehydrogenase I, complete cds, strain:LV61
	gi|34787326	AB118958.1	*Oryza rufipogon* Adh1 gene for alcohol dehydrogenase I, complete cds, strain:W630
	gi|34787302	AB118851.1	*Oryza rufipogon* Adh1 gene for alcohol dehydrogenase I, complete cds, strain:W1965
	gi|34787336	AB118963.1	*Oryza barthii* Adh1 gene for alcohol dehydrogenase I, complete cds, strain:W607
	gi|34787338	AB118964.1	*Oryza glumipatula* Adh1 gene for alcohol dehydrogenase I, complete cds, strain:W1167
	gi|34787340	AB118965.1	*Oryza meridionalis* Adh1 gene for alcohol dehydrogenase I, complete cds, strain:W1627

Figure 13
Linearized tree of the ADH and GDH gene sequences corresponding to different model plants as listed in Table-7



(Note: The scale bar corresponds to 0.1 estimated amino acid substitutions per site)

In the second phylogenetic tree (Figure 13), there are three clusters of orthologous genes. The branch lengths were re-estimated under the assumption of rate constancy. It was seen that the branch leading to cluster I was significantly longer than the branches for the other two clusters (Table 7). Cluster I consists of *Ocimum americanum* geraniol dehydrogenase gene (JQ765502) and *Ocimum basilicum* geraniol dehydrogenase sequence (AY879284), whereas, cluster II and III have five maize and 21 rice orthologous *ADH* sequences, respectively. The longer branch length of cluster I implies that *GDH* has evolved faster than *ADH* (cluster II and III) with higher substitution rate of divergence time 70 Mya. The sequences in cluster II and III have very less significant substitution rate.

Supplementary Information

Sequence Assembly and newly constructed sequence of *Cymbopogon winterianus* retrotransposon Ty3-gypsy subclass (*CwRTR*), complete sequence (1536 bp)

>gi|557132934|gb|KF415292.1| *Cymbopogon winterianus* retrotransposon Ty3-gypsy subclass, complete sequence

TCGGCTAAGATCCCTCCTGCCGACCAGGTAGTAGCACCACTGGACGAGGATACCATCATGGA
AGACGCCGACCCTCCAGACCAGCAGGGGCCAAACCACTGGGAACCTGAGGACCAGCCCCC
AAACCACGCTCCGTCAAACAGCCTAAACCCCCCAAAGCCTGACTGGGAGGAACCTTTCATTG
AGTACCTGGTCTCCACCTCTTTCCAAAGGACAAATCAAAAGCAAAACGTTAATGGCGCACGCA
CGCAACTACGTCATAATCAACAATCATTTGATGCGCAGAACTCCAAACCAAATTTACTAAAAAA
TGGCTTTACTCGGAAAGCTGGTCAAAATTTTGGTTGAAATCCCCCAAGGCATGGGGGGTAAC
CACGCCGCATCCAAAACGCTGGTCGGCAGGGTTTTCGAAGCAGGTTTCTACGGGCCTACAG
CATAACCAAACCCCAAGAACATGGTTCGCCGTTGTCGGAATGGCCAAATGTTCCCCAAACAA
ATCCTGGTCCGGCTCACAACTCCAAACCATACCCGCTTCAGGGCCGTCTCTCAGGCTGGGAA
CTGACATGATGGGCCGTTCAACCAACCACCAGAAGAATTCAAGTACGTCAAGGAACTAATAA
CAAGTTCCCAAGTGAATAGAATACAAGCCTTAAAGGTTGCTACGGCACAAAAGCAGCAGATTT
TCCCGATGAGTCATCCACAGTCGATACCCATAGTATCATCACTGATTGGGATCACTTCACGGG
ACCACTCTGGACCTTGCAACAACTCAGCATCAAGTCAGACGCCTCGCCGTCATAAACGGTCC
GAGGATACCCTCATGGAAGGCGCCGACCTTCCAGACCAGCAGGGCCAGACCACGTGGCAG
CTGAGGACCAGCCTCCAGACCACGCTCCATCAGAGCAGCCTAAACCCCCAGAGCCTGACTG
GAGGGAACCTTTCATTGAGTACTGGTCAATCACCTCGTTCCAGAGGACAAATCAGAAGCAGA
ACGCTTAATGCGGCGCGCACGCAACTATGTCATAGTAGACAACAGTTGATGCGCAAGAACTC
CAGAGCAGATATACTATAGAAGTGCATTACTCGGGAAGCTGGTCAAGAAATTTTGGTTGAAAT
CCACCGAGGCATGTGTGGTAACCACGCCGCATCCAGAACGCTGGTCGGCAAGGCTTTCCGA
GCAGGTTTCTACTGGCCTACAGCAGTAGCAGACGCCGAGGACATGGTTCGTCGTTGTCTGGA
GTGCCAGATGTCACCAAACAGATCCATGTTCCGGCTCACGAGCTCCAGACCATACCCGCTTC
ATGGCCGTTCTCATGCTGGGGACTGGACATGATTGGCCCGTTCAAACGAGCACCAGGAGGA
TTCGAGTACGTCTATGTACTAATGACAAGTTCACCAAGTGGATAGAATACAAGCCATTGAAGG
TTGCTACCGCACAAAAGGCAGCAGATTTTCTCGATGAAGTCATCCACAGGTTCGGAGTACCC
AACAGTATCATCACTGATTTGGGATCCACTTCACCGACACCACTT

The red colored sequences (322 bp) are the highlighted features showing mobile element retrotransposon Ty3-gypsy subclass in *Citronella*.

Translated retrotransposon Protein sequence (*EMBOSS workbench)*

>*Cymbopogon winterianus* Ty3 gypsy subclass retrotransposon protein

SAKIPPADQVVAPLDEDTIMEDADPPDQQGPNHWEPEDQPPNHAPSNSLNPPKPDWEEPFIEYL
VSTSFQRTNQKQNVNGARTQLRHNQQSFDAQNSKPNLLKNGFTRKAGQNFGNPPRHGGPRRIQ
NAGRQGFRSRFLRAYSITKPQEHGSPLSEWPNVPQTNPGPAHNSKPYPLQGRLSGWELTWAVQ
PTTRRIQVRQGTNNKFPSENTSLKGCYGTKAADFPDESSTVEYPYHHLGSLHGTTLDLATTQHQV
RRLAVINGPRIPSWKAPTFQTSRARPRGSGPASRPRSIRAATPRALEGTFHVLVNHLVPEDKSEA
ERLMRRARNYVIVDNQLMRKNSRADILKCITREAGQEILVEIHRGMCGNHAASRTLVGKAFRAGF
YWPTAVADAEDMVRRCLECQMFTKQIHVPAHELQTIPASWPFSCWGLDMIGPFKRAPGGFEYVY
VLMTSSPSGNTSHRLLPHKRQQIFSMKSSTGSEYPTVSSLIWDPLHRTPL

Figure 1
Snapshot of BLAST result of sequenced forward primer of Citronella

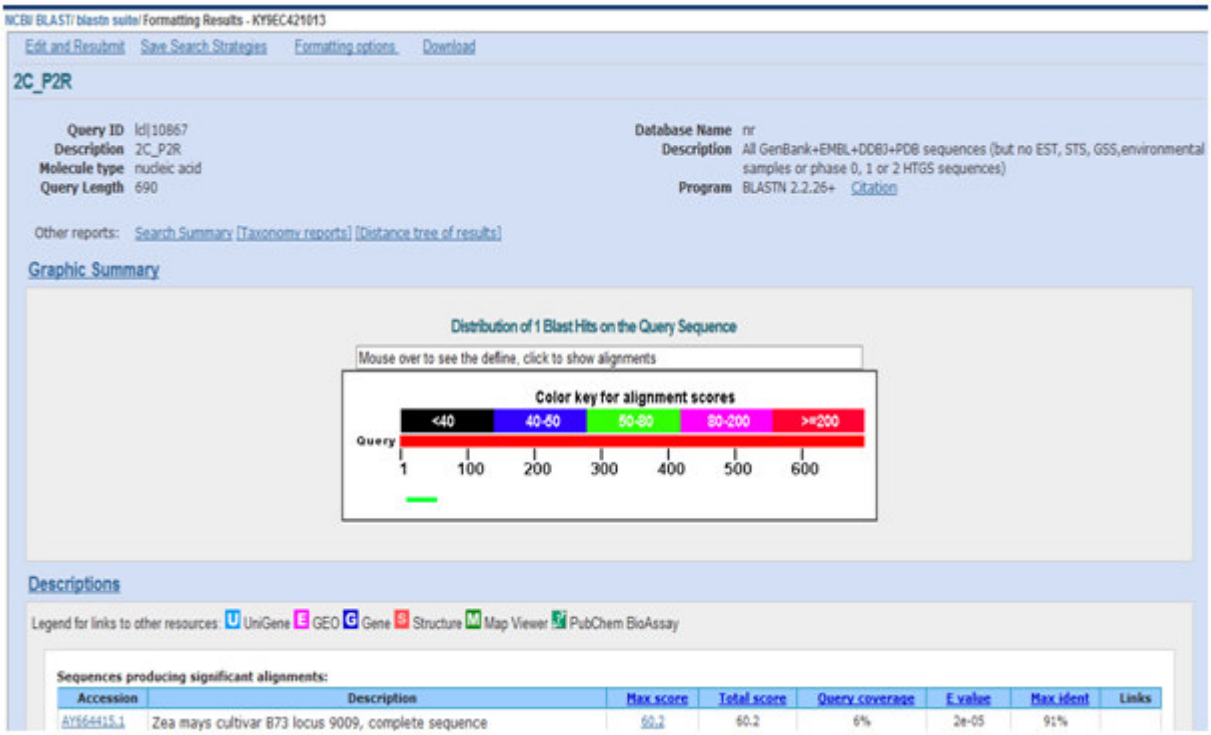


Figure 2
Snapshot of BLAST result of sequenced reverse primer of Citronella showing similarity with the Z. mays cultivar

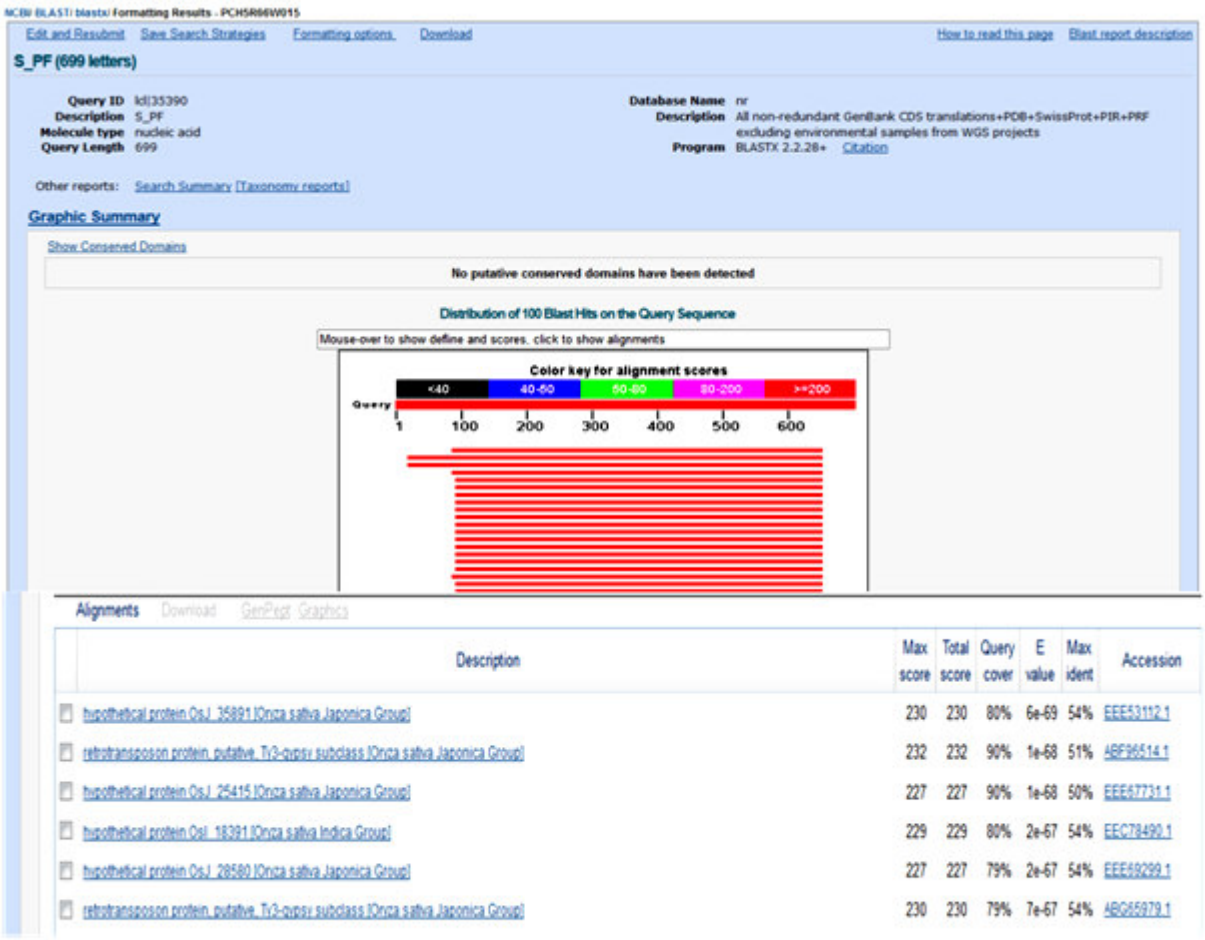


Figure 3
Snapshot of BLASTx result of sequenced forward primer of Citronella

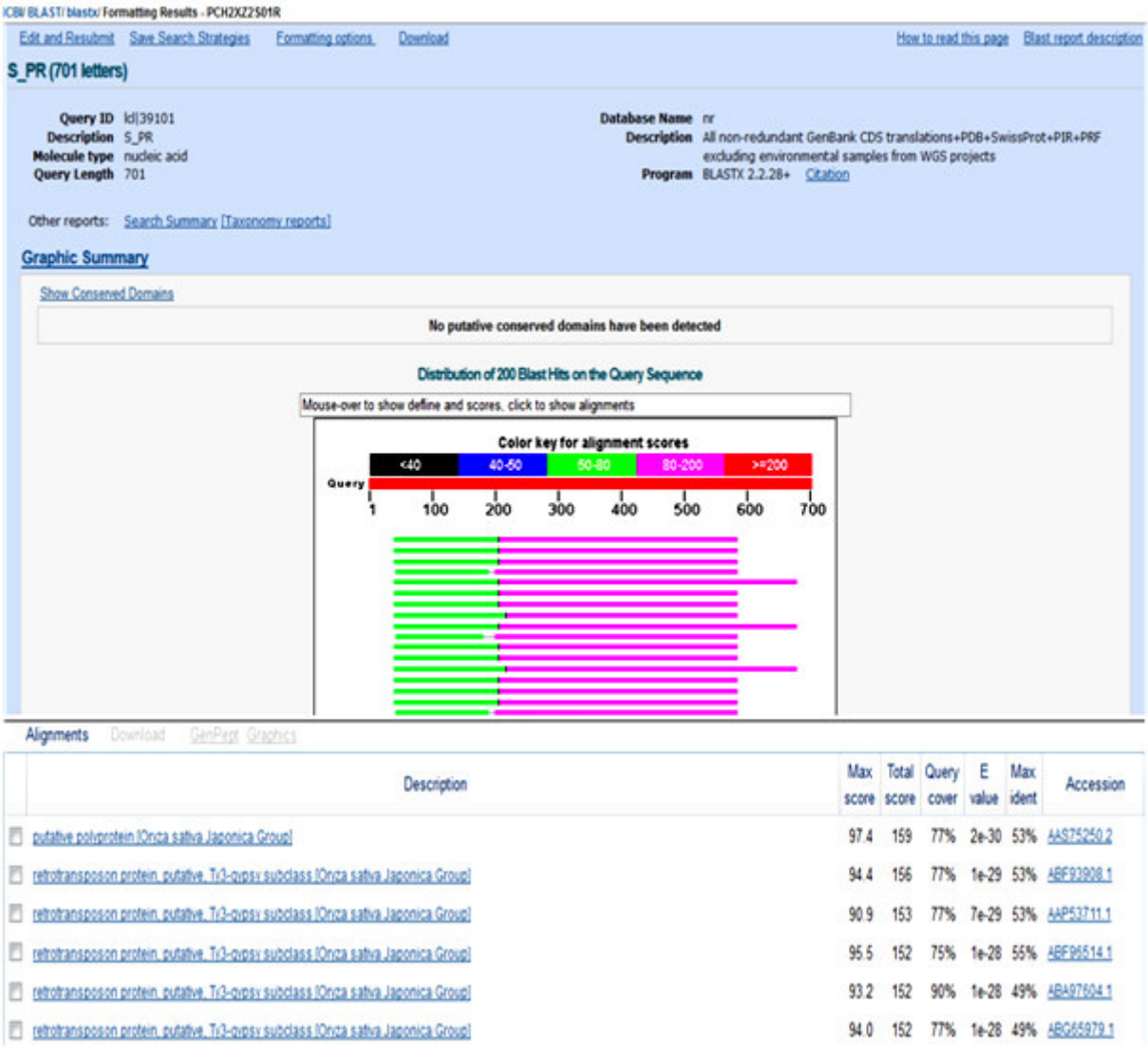


Figure 4
Snapshot of BLASTx result of sequenced reverse primer of Citronella

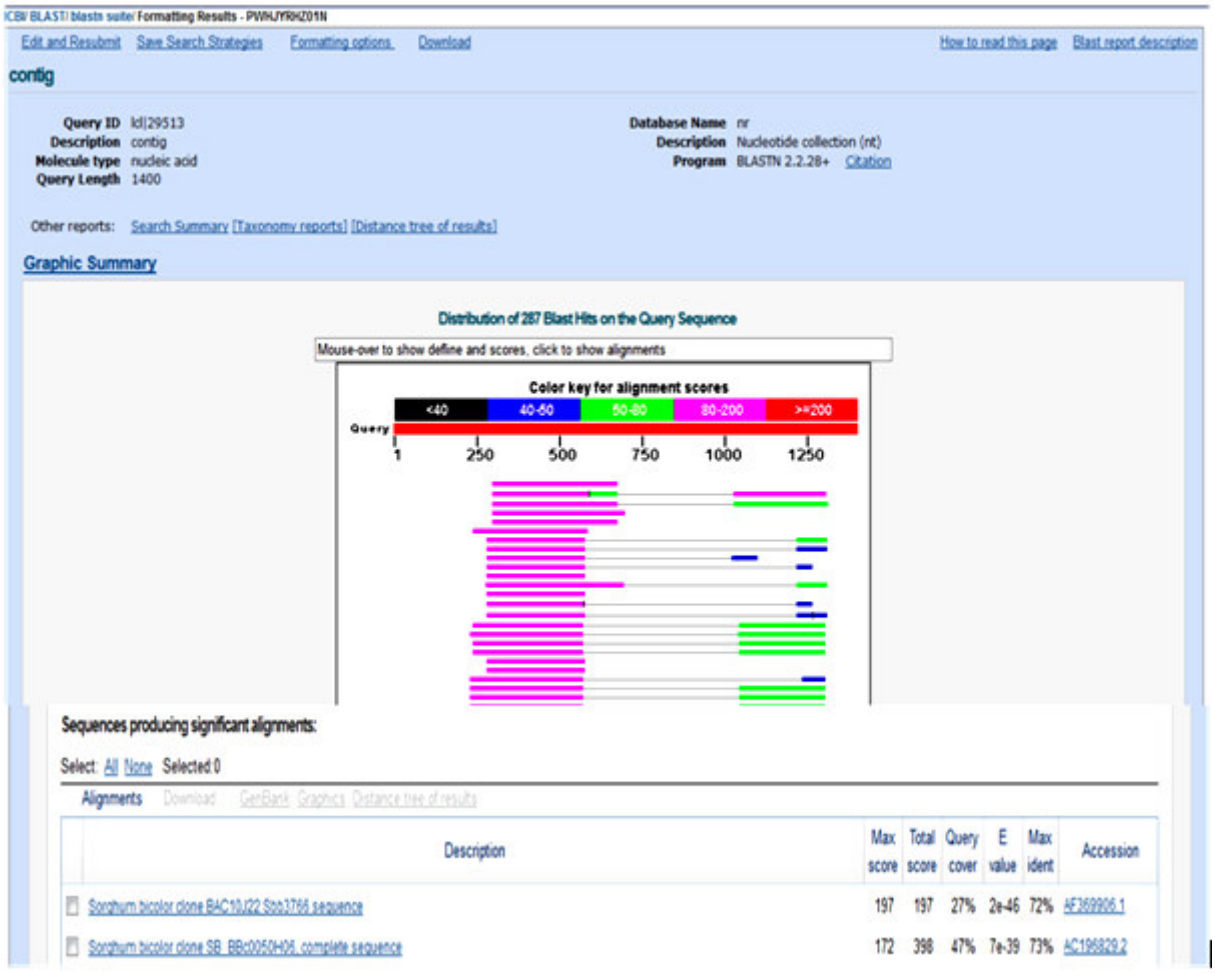


Figure 5
Snapshot of BLASTn of assembled contig showing the similarity with Sorghum bicolor

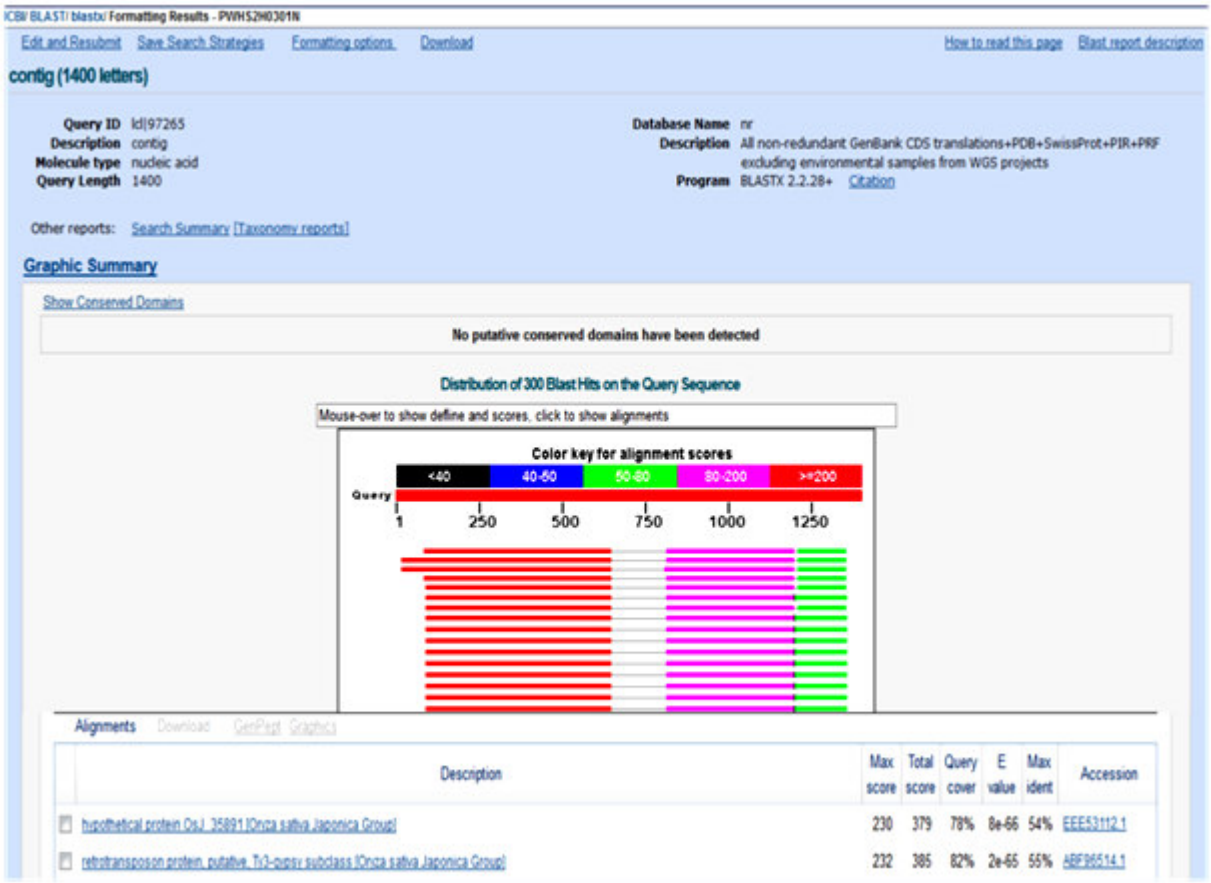


Figure 6
Snapshot of BLASTx of the assembled contig showing the similarity with retrotransposon protein

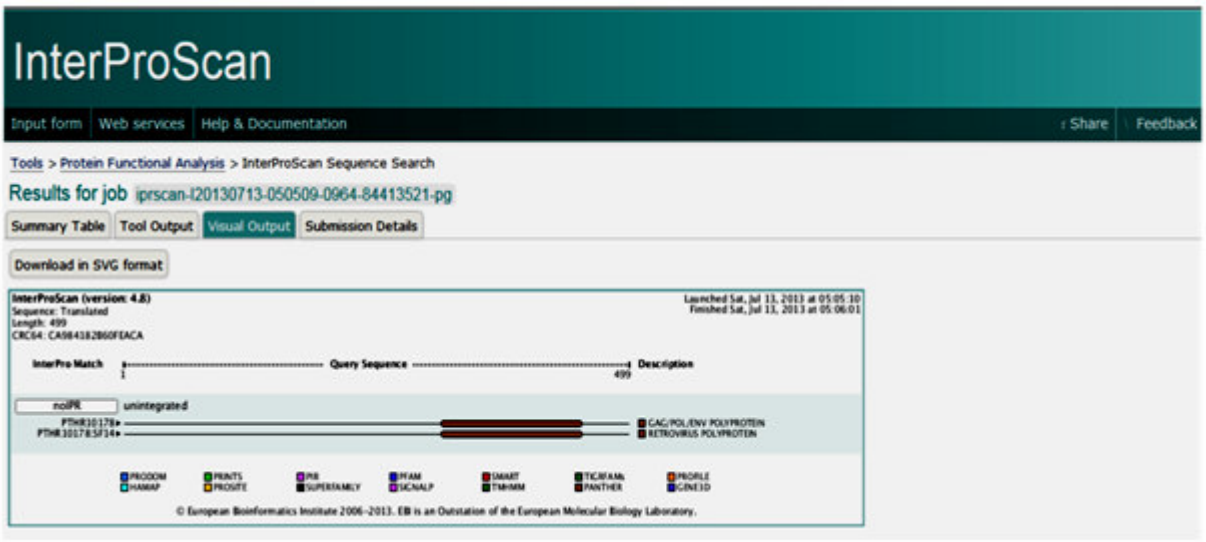


Figure 7
Snapshot of InterProScan of CwRTR newly constructed sequence

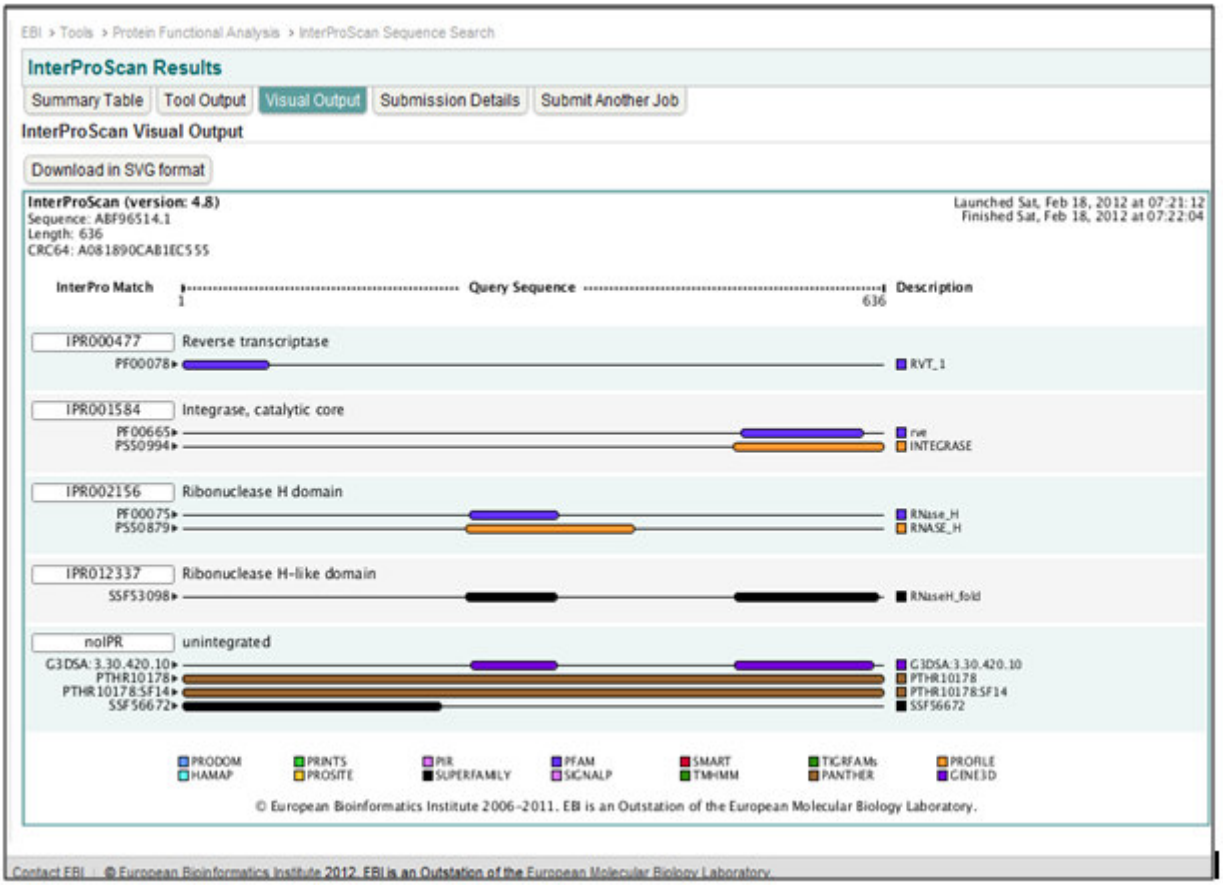


Figure 8

Snapshot of InterProScan of Citronella showed homology with putative retrotransposon protein Ty3 gypsy subclass Oryza sativa Japonica group (Accession no. - ABF96514.1)

4. DISCUSSION

As Java Citronella is least studied molecularly and the interference of secondary metabolites are concerned, an attempt was taken to isolate genomic DNA through modified protocol using 3.5% CTAB and 3% PVP with an extended RNase treatment. Our aim to identify and isolate *GDH* gene from the *Citronella* genome provided an impetus on the geraniol degradation pathway. In this regard, the key enzyme (geraniol dehydrogenase) responsible for the degradation of essential oil components (geraniol, citronellol and nerol) has to be blocked. After three times of sequencing, the results were validated by phylogenetic analysis to ensure the relationship between maize and *Citronella*.

4.1 DNA extraction optimization studies

Genomic DNA isolation has been a challenging task for monocot plants than dicots especially which possess numerous secondary metabolites. Therefore, an optimization study was undertaken aiming to standardize the isolation protocol of genomic DNA from *Citronella* by modified protocols. Many medicinal plants, crops, shrubs and ornamental plants containing polyphenols and polysaccharides have been attempted to use modified DNA extraction protocol using different compositions of extraction buffer[1, 7, 8,19,28,32]. A significantly high amount of gDNA (2978 µg/g) is obtained from the modified protocol involving 3.5% CTAB and 3% PVP (Table 2). Further, the yield of purified DNA was found to be 1535µg/g (Figure 3). In this study, the amount of purified DNA and quality ($A_{260/280}$ ratio) was 20 times (1535µg/g; 1.89) higher as compared to the protocol using leaf tissues (80.13µg/g; 1.80) using benzyl chloride reagent in the extraction buffer[7]. The result obtained in this study is in contrast with the findings which utilized two different protocols of DNA isolation from *C. winterianus* and *C. flexuosus*[7]. Although similar work has

been carried out, but contrasting results were found in the protocol used in terms of the yield of gDNA[1]. The same modified protocol was followed with three different varieties of *Ocimum* (*O. basilicum, O. americanum, O. sanctum*) which was taken as a control (Figure 3). *O. americanum* has not been studied till date as a popular cultivar as far as other *Ocimum* varieties are concerned. A significantly less amount of gDNA (250µg/ml) has been isolated from the species earlier[28] as compared to 1057µg/ml which is still four times higher in our findings[32]. Thus, the modified protocol using 3.5% CTAB and 3% PVP was standardized after obtaining reproducible results for both, *Ocimum* and *Citronella*. Isolated gDNA was found to be stable, reproducible and PCR amplifiable before and after extended storage at -20°C. far as age of plants in concerned, it was observed that, though the yield of gDNA was higher in Y1 and Y2 plants but the $A_{260/280}$ ratio was quite less as compared to M plants. During the growing stage of the leaves, the DNA is still in the preparatory phase and there is presence of other contaminants along with the DNA such as RNA and protein, shown as smeared appearance of the DNA bands in Y1 and Y2. But, the discrete gDNA band from the mature leaves shows the presence of pure and high quality DNA with the absence of other contaminants (Figure 6). Seasonal variation has always been an influential and decisive factor as far as regulation of some genes is concerned which, is involved in the complex metabolic pathways. Any climatic change or seasonal alteration allows the plant species of the same or related genera to show enormous variability in the complexity of pathways of dispensable functions as well as its biochemical composition[9]. *Citronella*, being cultivated throughout the year, faces these traumatic changes in the essential oil yield and thus, changes in the products of the metabolic pathways regulated by some specific genes. The study of the essential oil yield and its variations has been reported in many literatures but the yield of the gDNA of *Citronella* during these climatic changes remains untouched. Therefore, gDNA was isolated in three different seasons i.e. during March (summer), July (monsoon) and December 2012 (winter). The leaves were harvested for the gDNA isolation as well as for essential oil extraction[31] and its analysis. It was observed that, no significant difference in the gDNA yield and the purity due to seasonal collection was seen (Table 3, Figure 7).

4.3 Total cellular RNA extraction

As reported in Table 5, highly pure total cellular RNA with an average $A_{260/280}$ ratios range of 1.90 has been observed. This first successful attempt of isolating RNA from *Citronella* and *Ocimum* using TRIzol® in very less time has proven other protocols less yield oriented in maximum extraction period using expensive protocols such as Nucleon PhytoPure™ system (25-900µg/g; 1.89; 5h)[12] and modified extraction buffers. In this study, the time taken for isolating total cellular RNA using TRIzol® was 2 hours which is very less compared to other protocols.

4.4 PCR, sequencing and gene annotation studies

GDH gene specific primers and degenerate primers based on dehydrogenase domains were amplified against gDNA and cDNA templates of *Ocimum* and *Citronella*. The identification and isolation of *GDH* gene has been earlier studied in *O. americanum*[32] and is first time reported here in *Citronella*. *Cymbopogon* genus members have been identified by various genetic diversity analysis using RAPD[22, 24] and EST-SSR markers through PCR[14]. It suggests that, no studies have been conducted regarding the identification of *GDH* gene in *Citronella* rather diversity among various accessions of *Cymbopogon* species and their phylogenetic relationships have been established and discussed[11]. The use of degenerate primers has also been reported for the first time for the amplification of *GDH* gene in *Citronella*. *GDH* gene annotation studies during similarity searches revealed that most of the hits corresponded to Ty3 gypsy sub-class retrotransposon protein belonging to *O. sativa* japonica group with 50-54% identity and 75-90% query coverage after three times of *GDH* gene sequencing. Also, it showed similarity with the *Z. mays* with 97% identity and 6% query coverage. Functionally, as per protein prediction tool (I-TASSER), it possesses reverse transcriptase, RNase-H like domains and integrase domains. In a review, it has been explained that these mobile elements

have been characterized by PCR amplification of gDNA or cDNA[6].

4.5 Phylogenetic analysis

The major reason for the evolution of *Cymbopogon* from *Andropogon* genus (*Z. mays*) may be the presence of these retrotransposons. The phylogenetic analysis studies confirmed through the linearized trees that, *Citronella* has been evolved from *maize* and diverged from Ty3 gypsy subclass of RT from *O. sativa* japonica group around 23 mya. The results obtained directly reveals the evolutionary status of *C. winterianus* to a distant relative of *Z. mays*[15] is in accordance with them. RT's are generally more transcriptionally active in the monocot grasses than in other groups of plants, although transcription occurs in all groups[29]. Three major groups of RTs have been identified in plants, viz., copia-like, gypsy-like and non-long terminal repeat RTs[5, 13]. This study might emphasize on providing valuable insights in the prevalence of RTs in the Java *Citronella* genome as evidenced during *GDH* gene annotation. As *Cymbopogon* plants are able to survive extreme stress conditions, the reason may be the benefit from the stress conditions i.e, RTs, the plant may not silence the transcription of the retrotransposon genes and thus, they express themselves at different positions in host genomes which lead to genetic diversity and evolution[16]. As a matter of fact, transposons and RTs are first reported in *Z. mays* which evolved from the teosinte grass (Andropogon *sps*), and is also a distant relative of *Cymbopogon* due to environmental challenges. In fact, active RTs are switched on by stress and their promoter elements are similar to those of plant defence genes, and it may bind to similar defence-induced transcription factors[4]. Many cellular stress signals activate the RTs[5], but the molecular mechanisms controlling the activation of RTs remain unclear. Another phylogenetic study was carried out among the *GDH* and *ADH-1* gene sequences of the model plants. A total of 28 *ADH* sequences were aligned for the presence of dehydrogenase domains and evolutionary relationship has been proven by dividing them in three clusters. As per the branch length and substitution rate constancy, Cluster I (*GDH*) evolved faster than the five corresponding *ADH-1* maize and 21 rice orthologous genes in Cluster II and III respectively (Figure 13). *ADH-1* gene is an ideal gene due to its convenient size (1000bp-coding region) and present in low copy number in maize genome[27]. Differences in the sequences of *ADH* in different species have been used to create phylogenies justifying the closely related different species of plants. The *ADH-1* gene is used a check gene (check primer) in this study for confirming the presence of dehydrogenase domains in *Ocimum* and *Citronella* genome. The amplified bands in *Citronella* gave another conclusion that, *GDH* is evolved from the *ADH-1* gene, again proving the indirect and distant relationship of *Citronella* with maize. The estimates of the size and copy number of RTs in a 200-250 kb region flanking the *ADH-1* gene suggests that 33-62% of the maize genome is composed of RTs. Therefore, it can be concluded that RTs have increased its size two to five fold since the divergence of maize is from a common ancestor about 16 mya. Also, *Sorghum* genome has evolved at the similar time from the common ancestor as maize and is still three folds smaller than the maize genome. This proves that the similarity of *GDH* gene which was observed with the sorghum genome is due to these RTs and thus, it can be considered that just like maize; sorghum is also distantly related to *Citronella*. Hence, this study might be considered as the preliminary evolutionary evidence for the prevalence of Ty3 gypsy sub-class RT which were found in the *Citronella* genome during GDH gene annotation studies. Furthermore, RT-based specific probes can be designed and can be amplified across the *Citronella* genome, to know the exact copy number and heterogeneous distribution of this class of RT.

CONCLUSION

With the aim to identify, isolate and characterize *GDH* gene, the present study progressively revealed another facet of *Citronella* genome with the prevalence of Ty3 gypsy sub-class RT after three times of *GDH* gene sequencing. The existence of transposons/RTs and substantial increase in genome size are validated by many scientists since ancient times in maize. Another added fact is *Citronella* belongs to the same sub-tribe

Andropogonae as *Z. mays*. Phylogenetic analysis also, confirmed that *C. winterianus* is evolved from *Z. mays* and the Ty3 gypsy sub-class RT protein belonged to *O. sativa* japonica group which suggests its early divergence. The *GDH* gene belongs to the *ADH* superfamily of maize, again interpreting the indirect and distant relationship of *Citronella* with maize and other RT contain model plant genome such as *S. bicolor*, and *O. sativa*. This is a first and preliminary inadvertent investigation of *Citronella* genome towards isolating a specific *GDH* gene responsible for the degradation of essential oil components. The occurrence of RT in this genome is by far a conclusion revealing the fact which needs specific attention to confirm its existence with the help of RT based probes using PCR or RAPD.

ACKNOWLEDGEMENT

Dr. A. Kohli, Senior Scientist, Plant Molecular Biology Laboratory, IRRI, Philippines is duly acknowledged for kindly suggesting and designing degenerate primers for *GDH* gene amplification. The authors are grateful to Dr. Ashutosh Kumar, Research Associate, National Institute of Plant Genome Research, Aruna Asaf Ali Marg, New Delhi, India for constructing phylogenetic trees and for calculating evolutionary time in the present study. University Grants Commission, New Delhi, India is gratefully acknowledged for providing financial support to Ms. Aakanksha Wany (F1-17.1/2012-13/MANF-2012-13-CHR-CHH-9539) as Senior Research Fellow.

REFERENCES

1. Bhattacharya S., Bandopadhyay T K., Ghosh P D. Efficiency of RAPD and ISSR markers in assessment of molecular diversity in elite germplasms of *Cymbopogon winterianus* across West Bengal, India. Emir. J. Food Agric, 22(1): 13-24, (2010).
2. Birney E., Clamp M., Durbin R. GeneWise and Genomewise. Genome Res. 14(5): 988–995 (2004).
3. C M Bertea., M.E. Maffei.,Ed. The Genus Cymbopogon Botany, Including Anatomy, Physiology, Biochemistry,and Molecular Biology. In: A. Akhila (eds.), Essential oil-bearing grasses: The genus *Cymbopogon*, CRC press, Taylor and Francis group, Boca Raton, Florida, pp. 1-20, 2010.
4. Casacuberta J M., Santiago N. Plant LTR retrotransposons and MITEs: control of transposition and impact on the evolution of plant genes and genomes. Gene 31(1):1-11 (2003).
5. Cavrak V V., Lettner N., Jamge S., Kosarewicz A., Bayer L M., Scheid O M. How a retrotransposon exploits the plant's heat stress response for its activation. PLoS Genet. 10(1), e1004115. doi:10.1371/journal.pgen.1004115 (2014).
6. Cordaux R., Batzer M A. The impact of retrotransposons on human genome evolution. Nat. Rev. Genet. 10(10): 691–703 (2009).
7. Dey T., Saha S., Dhar T N., Adhikary S, Ghosh P. Optimization and comparison of efficiency between two DNA isolation protocols in *Cymbopogon* species, Gen. Appl. Plant Physiol. 36(3–4): 232–238 (2010).
8. Ibrahim R I H. A modified CTAB protocol for DNA extraction from young flower petals of some medicinal plant species. Gene conserve 10(40): 165-182 (2011).
9. K Weishing H Nybom, K Wolff, W Meyer, DNA isolation and purification. In: DNA fingerprinting in plants and fungi, CRC Press, Boca Raton, Florida, pp. 44-59 (1995).
10. Katiyar R, Gupta S, Yadav K R, *Cymbopogon winterianus*: an important species for essential Java *Citronella* oil and medicinal values. In: National conference on forest biodiversity: earth's living treasure FRI Kanpur (2011).
11. Khanuja S P S, Shasany A K, Pawar A, Lal R K, Darokar M P, Naqvi A A, Rajkumar S, Sundaresan V, Lal N, Kumar S, Essential oil constituents and RAPD markers to establish species relationship in *Cymbopogon* Spreng. (Poaceae). Biochem. Syst. Ecol. 33: 171–186. (2005).
12. Kiefer E, Heller W, Ernst A D, Simple and efficient protocol for isolation of functional RNA from plant tissues rich in secondary metabolites. Plant Mol. Biol. Rep. 18: 33-39. (2000).

13. Kumar A, Bennetzen J L, Plant Retrotransposons, Annu. Rev. Genet. 33: 479–532 (1999).
14. Kumar J, Verma V, Goyal A, Shahi A K, Sparoo R, Sangwan R S, Qazi G N, Genetic diversity analysis in *Cymbopogon* species using DNA markers. Plant Omics Journal, 2(1): 20-29 (2009).
15. Langdon T, Seago C, Mende M, Leggett M, Thomas H, Forster J W, Jones R N, Jenkins G,. Retrotransposon Evolution in Diverse Plant Genomes. Genetics 156: 313–325 (2000).
16. Mansour A, Epigenetic activation of genomic retrotransposons. J. Cell Mol. Biol. 6(2): 99-107 (2007).
17. Najafabadi H S, Torabi N, Chamankhah M, Designing multiple degenerate primers via consecutive pairwise alignments. BMC Bioinformatics 9: 55 (2008).
18. Padalia R C, Verma R S, Chanotiya C S, Yadav A, Chemical fingerprinting of the fragrant volatiles of nineteen Indian cultivars of *Cymbopogon* Spreng. Rec. Nat. Prod. 5: 290–299 (2011).
19. Quereshi S, Purwar P, Singh R, Khan N A, Mani A, Patel J, Studies on Essential Oils and DNA Extraction from *Ocimum* species. J. Phytol. 3(8): 23-27 (2011).
20. Rio D C., Ares M Jr., Hannon G J., Nilsen T W. Purification of RNA using TRIzol (TRI reagent). Cold Spring Harb. Protoc. (6), pdb.prot5439 (2010).
21. Roy A, Yang J, Zhang, COFACTOR: An accurate comparative algorithm for structure-based protein function annotation. Nucleic Acids Res. 40: W471-W477 (2012).
22. Sangwan N S, Yadav U, Sangwan R S, Genetic diversity among elite varieties of the aromatic grasses, *Cymbopogon martinii*. Euphytica 130: 117-130 (2003).
23. Shah G, Shri R, Panchal V, Sharma N, Singh, B, Mann A S, Scientific basis for the therapeutic use of *Cymbopogon citratus*, stapf (Lemon grass). J. Adv. Pharm. Tech. Res. 2: 3-8 (2011).
24. Shasany A K, Lal R K, Darokar M P, Patra N K, Garg A, Kumar S, Khanuja S P S, Phenotypic and RAPD diversity among *Cymbopogon winterianus* Jowitt accessions in relation to *Cymbopogon nardus* Rendle. Genet. Resour. Crop Evol. 47: 553–559 (2000).
25. Stapf O. The oil grass of India & Ceylon, Kew Bull. 8:297-267 (1906).
26. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S, MEGA5: Molecular Evolutionary Genetics Analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28: 2731-2739 (2011).
27. Thompson C, Fernandes C, De Souza O, De Freitas L, Salzano F, Evaluation of the impact of functional diversification on Poaceae, Brassicaceae, Fabaceae, and Pinaceae alcohol dehydrogenase enzymes. J. Mol. Model 16(5): 919–928 (2010).
28. Tiwari K L, Jadhav S K, Gupta S, Modified CTAB technique for isolation of DNA from some medicinal plants. Res. J. Med. Plant 6(1): 65-73 (2012).
29. Vicient C M, Jaaskelainen M J, Kalendar R, Schulman A H, Active retrotransposons are a common feature of grass genomes. Plant Physiol. 125: 1283–1292 (2001).
30. Wany A, Jha S, Nigam V K, Pandey D M, Chemical analysis and therapeutic uses of *Citronella* oil from *Cymbopogon winterianus*: A short review. I. J. A. R. 1(6): 504-521 (2013a).
31. Wany A, Kumar A, Nallapeta S, Jha S, Nigam V K, Pandey D M, Extraction and characterization of essential oil components based on geraniol and citronellol from Java *Citronella* (*Cymbopogon winterianus* Jowitt). Plant Growth Regul. 73:133–145 (2014b).
32. Wany A, Kumar A, Nigam V K, Pandey D M, Isolation and *in silico* characterization of geraniol dehydrogenase encoding gene using modified DNA extraction protocol from American Basil (*Ocimum americanum*). Int. J. Pharm. Bio. Sci. 4(3): 365 – 379 (2013b).
33. Wany A, Nigam V K, Pandey D M, Understanding new paradigm of plant metabolomes through secondary metabolite enrichment. Res. J. of Biotech. 9(4): 94-102 (2014a).