Academic Sciences International Journal of Pharmacy and Pharmaceutical Sciences

ISSN- 0975-1491 Vol 3, Suppl 5, 2011

Research Article

COMPARATIVE PHYLOGENETIC APPROACH FOR DISOVERING ALTERNATIVE SOURCES OF CAMPTOTHECIN – AN ANTICANCER DRUG

NUPUR SONI[1] AND ALOK MITTAL[2]

[1]Department of Bioinformatics,MANIT,Bhopal,India, [2]Department of Chemistry, MANIT ,Bhopal ,India.

Received: 6 July 2011, Revised and Accepted: 2 Nov 2011

ABSTRACT

Camptothecin is an essential precursor of semisynthetic chemotherapeutic agents for cancers throughout the world. In spite of the rapid growth of market demand, camptothecin raw material is still harvested by extraction from *Camptothecin acuminata* and *Nothapodytes foetida* .But still its total synthesis is not cost-effective. However, these natural resources are being threatened day by day due to the destructive collection of wight bark and Chinese happy tree seeds for camptothecin production Thus, for the stable supply of camptothecin and for the development of more potent anticancer compounds, alternative sources for the biosynthesis of camptothecin are urgently anticipated. The computational approaches for prediction of the alternatives are fast and economical and therefore can be used to complement the existing wet lab techniques. In this paper an attempt has been made to find alternative sources for camptothecin by using Comparative Phylogenetics approach based on molecular, structural and physicochemical level so as to lessen the threat on the existing species. The analysis was performed using enzymes physiochemical data, domain organization and sequence data. Dendogram was generated through physiochemical data whereas phylogenetic trees (based on sequence alignment and structural alignment) were generated through sequence data. Identification of domain families of homologous protein was done to analyze their structural homology relationship. Consequently after comparison ,three different organisms, *Arabidopsis thaliana, Ricinus communis* (castor oil plant) and *Oryza sativa*(Rice) were observed to be related to Camptothecin plant(Chinese Happy Tree) sequentially, physiochemically and structurally, thus providing alternatives for Camptothecin production.

Keywords: Camptothecin, Comparative phylogenetics, Structural alignment, Physiochemical properties, Dendogram, Phylogenetic tree

INTRODUCTION

Plants have been used for treating various diseases of human beings and animals since time immemorial. They maintain the health and vitality of individuals, and also cure diseases, including cancer without causing toxicity[1]. More than 50% of all modern drugs in clinical use are natural products, many of which have the ability to control cancer cells. A recent survey shows that more than 60% of cancer patients use vitamins or herbs as therapy [2,3] .Therefore, natural protection against cancer has recently been receiving a great deal of attention not only from cancer patients but, surprisingly, from physicians as well.Camptothecin; a pentacyclic quinoline alkaloid, has been discovered as a potential anticancer agent through an extensive screening by the U.S. National Cancer Institute [4]. It is a potent broad spectrum antitumor agent which is of interest due to a variety of camptothecin derivatives currently being used in a variety of clinical protocols. Camptothecin functions by effectively converting DNA topoisomerase I into a DNA damaging poison, causing it to irreversibly bind DNA midway through its catalytic action, leading to a double-stranded break when encountered by a replication fork[5] Camptothecin and its derivatives are active against carcinomas of the stomach, colon, head, neck and bladder.

At present, semisynthetic water-soluble camptothecin analogs, topotecan and irinotecan, are prescribed as clinical antitumor drugs throughout the world. CPT derivatives that exhibit remarkable antitumour and antileukaemia activity, are in clinical trials to treat breast, small-cell lung cancer and leukemia etc[6].The worldwide market size of irinotecan/topotecan in 2003 was estimated at about US$1 billion[7] .Despite the rapid growth of the market, camptothecin is still obtained by the extraction from natural resources such as the seeds of *Camptothecin acuminata* (Chinese Happy Tree) and the bark of *Nothapodytes foetida*(Wight). A number of studies on biotechnological production of camptothecin have also been reported [7,8,9]. These are mostly focused on cell/tissue cultures including transformed hairy roots or adventitious roots [10] of the plants which accumulate camptothecin. The investigation of organogenesis of the explants derived from the camptothecin-producing plants has been conducted [7,8,10,11] However, these natural resources are being threatened day by day due to the destructive collection of wight bark and Chinese happy tree seeds for camptothecin production .Thus, for the stable supply of camptothecin and for the development of more potent anticancer compounds, alternative sources for the biosynthesis of camptothecin are urgently anticipated [8]

The computational approaches for prediction of the alternatives are fast and economical and therefore can be used to complement the existing wet lab techniques. They not only shorten the R&D time cycle but also reduce the ever increasing cost involved in the drug discovery process[12]. Bioprospecting through molecular phylogenetics analysis provides new horizons for this type of study. Enzymes involved in the biosynthetic pathway of the compounds act as mapping units in bioprospecting,. Questions concerning enzyme function and performance remain unanswered if only molecular data is used. However these questions can be answered by analyzing structural and physiochemical data of these enzymes. Physiochemical properties and their similarity level can very well explore the qualitative measurement of the enzymes , whereas the structure-based homology annotations can very accurately record functional homology relationships. Because, the structural similarity is preserved well beyond sequence similarity [13],protein structures are often the gold standard for annotating homology relationships. Studies in this direction focusing on individual pathways [14] or on the entire metabolic repertoire [15] have been attempted. But so such work has been reported for camptothecin biosynthesis pathway. In the present work, an attempt has been made to find alternate source for camptothecin through molecular, physiochemical and structure based Comparative Phylogenetics Approach by analyzing its biosynthesis pathway. The comparative phylogenetic analysis is performed using protein sequence and physiochemical data of all the 5 enzymes involved in the biosynthetic pathway of camptothecin. Enzymes of the source that gave same results with physiochemical data, sequence data and structural data were proposed to be the most prominent alternate source for camptothecin biosynthesis.

MATERIALS AND METHODS

Enzymes regulating camptothecin biosynthesis pathway and their BLAST search.

Except for a few undefined steps , the complete camptothecin biosynthetic pathway was elucidated and many genes encoding certain enzymes, which regulate its biosynthesis pathway, have been cloned and characterized in *Ophiorrhiza pumila* Though

camptothecin is structurally classified as a quinoline alkaloid, it belongs to a family of monoterpenoid indole alkaloids from the viewpoint of biosynthesis[4].

Table 1: **List of 5 reported enzymes of camptothecin biosynthetic pathways**

Enzyme	Accession no.	Reference
strictosidine synthase (STR)	OpSTR; EC 4.3.3.2	[16]
tryptophan decarboxylase (TDC)	OpTDC; EC 4.1.1.28	[16]
NADPH:cytochrome P450 reductase (CPR)	OpCPR; EC 1.6.2.4	[16]
secologanin synthase (SLS)	SLS; EC 1.3.3.9	[17,18]
geraniol 10-hydroxylase (G10H)	G10H; EC 1.14.14.1	[16]

Nucleotide and protein sequences of each of the 5 enzymes were retrieved from the Genebank accession Numbers, which were then subjected to a BLASTp[19] search against the non-redundant database with the *e*-value inclusion threshold, set to 0.005. By using an option available in the BLAST program to select a particular organism, the search was restricted to plants. In this way potential homolog of enzymes were selected.

Data preparation and characterization

Sequence data, physiochemical data and structural data of enzymes was collected, processed and analyzed for phylogenetic analysis.

A) Physiochemical data

The protein sequence of each enzyme was subjected to ExPASy tool, ProtParam to obtained its physiochemical property such as No. of AAs, No. of Atoms, Molecular weight, Iso-electric point, Positive charge residues(Asp + Glu), Negatively charge residues(Arg + Lys), Aliphatic index, Instability index [20] Hydropathocity[21], and Extinction coefficient.

B) Sequence based data

In order to prepare data for phylogenetic tree analysis, entire protein sequence of each set of homologus enzymes was subjected to site based alignment. MEGA4.0 software [22] was used to analyse the Constant sites(C), Variable sites (V) and Singleton sites(S). Then overall similarity of the alignments was calculated by the Feature similarity score (FSS).

Feature similarity score (FSS) = Constant site+ Singleton site/ Total sites

C) Structure based Data

To study the phylogenetic relationship, the protein sequences of each set of homologus enzymes was subjected to structure based alignment through Expresso (3D- Coffee) – Regular(http://tcoffee.vital-it.ch/)[23] which aligns the templates using structural aligners like sap or TMalign .Thereafter ,these structure based alignments are used as a template for realigning the original sequences . Moreover, family identification of sequence homolouges was done through FISH Server (Family identification with Structure Anchored HMM) (http://babel.ucmp.umu.se/fish/.) [24]

Phylogenetics Tree and Dendogram generation

A) Dendogram generation

The physiochemical properties of each enzyme and their homologous is collected and calculated by ExPASy tool, ProtParam. Thereafter, distance measurement and linkage method are chosen for clustering of the physiochemical data. Finally, the statistical software Minitab is use to compute and generate dendogram.

B) Tree generation on basis of sequence alignment and structural alignment

The sequence alignment obtained from the MEGA software is subjected to phylogenetic tree generation. There are different methods of tree generation. Which method should be adopted is a great question. Thus, depending upon the nature of data (built alignment), the following categories were made:

-If data has high similarity (high FSS) i.e. greater than 75% - character based method (MP).

-If data has moderate similarity (medium FSS) i.e. less than 75% and more than 50% - distance based method (NJ, UPGMA)

-If data has lower similarity (low FSS) i.e. less than 50% - ML method

After observation of data, Cut off (%) is planned.

The structural alignments obtained from the EXPRESSO (3D Coffee) are then used for phylogenetic tree generation using PHY.FI (cgi-www.cs.au.dk/cgi-chili/phyfi/go) [25]. This program produces colored phylogenetc tree from the parenthesized Newick format phylogeny obtained from EXPRESSO (3D Coffee) RESULTS

Similarity search of enzymes involved in the biosynthesis pathway helps in the prediction of other biological sources that contains the same enzymes and thus provides an useful approach for discovering alternative and cost effective source for the camptothecin biosynthesis.

Output from physiochemical data

Physiochemical data of strictosidine synthase (STR) and its homologous which is calculated by ExPASy tool, ProtParam is shown in table-**2**. STR is one of the 5 known enzymes involved in camptothecin biosynthetic pathway.

Table 2: **Physiochemical data of STR enzyme and its homologs**

Name of protein enzyme	N	M	MW	PI	NG	PS	I I	EC	AI	H
AB060341.1strictosidine synthase*[Ophiorrhiza pumila]*	358	5413	38890.6	5.15	45	30	34.11	44600	83.02	-0.19
ACF21007.1strictosidine synthase*[Ophiorrhiza japonica]*	353	5482	39285.3	5.7	41	33	31.82	4584	82.55	-0.2
ABZ79473.1strictosidine synthase*[Mitragyna speciosa]*	352	5407	38751.5	5.65	38	31	27.14	47330	81.93	-0.19
P68175.1strictosidine synthase*[Rauvolfia serpentina]*	344	5332	38162	5.19	41	28	35.61	47330	84.68	-0.18
P68174.1strictosidine synthase*[Rauvolfia mannii]*	342	5305	37959.7	5.19	41	28	35.4	47330	84.88	-0.19
P18417.2strictosidine synthase[*Catharanthus roseus*]	352	5412	39093.7	5.14	43	29	47.42	48820	73.04	-0.29
CAN78856.1hypothetical protein *[Vitis vinifera]*	600	6410	64700.1	5.23	63	52	30.94	52830	93.27	0.048
ACU19372.1unknown *[Glycine max]*	336	5105	36360.2	8.32	27	29	38.23	25900	867.9	-0.09
NP_177540.3strictosidine synthase *[Arabidopsis thaliana]*	329	4931	34666.3	9.66	23	39	17.95	24410	83.22	0.024
AAB40593.1strictosidine synthase *[Arabidopsis thaliana]*	335	4978	35293.1	5.6	31	29	26.49	29910	86.12	0.07
XP_002532597.1strictosidine synthase *[Ricinus communis]*	375	5946	42292.2	6.26	47	45	34.71	62340	91.28	-0.21
AAR87254.1strictosidine synthase *[Oryza sativa Japonica]*	480	7507	53706.2	8.75	59	65	41.85	62340	79.65	-0.29

(Abbreviations- N=No. AAs, M= No. of atoms, MW= Molecular Weight, PI= Isoelectric Point, NG= (Asp+ Glu) PS= (Arg+Lys), II= Instability Index, EC= Extension Coefficient, AI=Aliphatic Index, H=Hydropathicity)

Similarly, physiochemical data was collected for other 4 enzymes and dendrogram was generated from these data for each enzyme by using statistical software Minitab.

Output from Sequence based data

Site based alignment was done to find similarity between all the 5 enzymes involved in the pathway.

From the table 3 it is clear that, out of 5 enzymes of biosynthetic pathway, enzymes with higher FSS are *zero*, moderate FSS is *one* and low FSS are *four*, subsequently distance based and maximum liklihood tree are constructed by MEGA and Phylip software. The phylogenetic analysis reveals that the trees are unrooted.

Output from Structure based data

The structural alignment of the protein sequences through EXPRESSO (3D Coffee) (figure1) and Domain superfamily identification from FISH revealed the presence of following homologous domains as in table 4

Table 3: **Characterization of all 5 known enzymes of camptothecin biosynthetic pathway through MEGA 4.0**

S. No	Name o Enzyme	Coserved/ Constant site	Variable sites	Singleton sites	Total sites	FSS index (%)
1	Strictosidine synthase	54	452	100	668	23
2	Tryptophan decarboxylase	40	394	82	543	40.8
3	NADPH-cytochrome P-450 reductase	297	412	155	801	56
4	Secologanin synthase	3	658	197	970	20.6
5	Geraniol10-hydroxylase	82	449	68	553	27.14

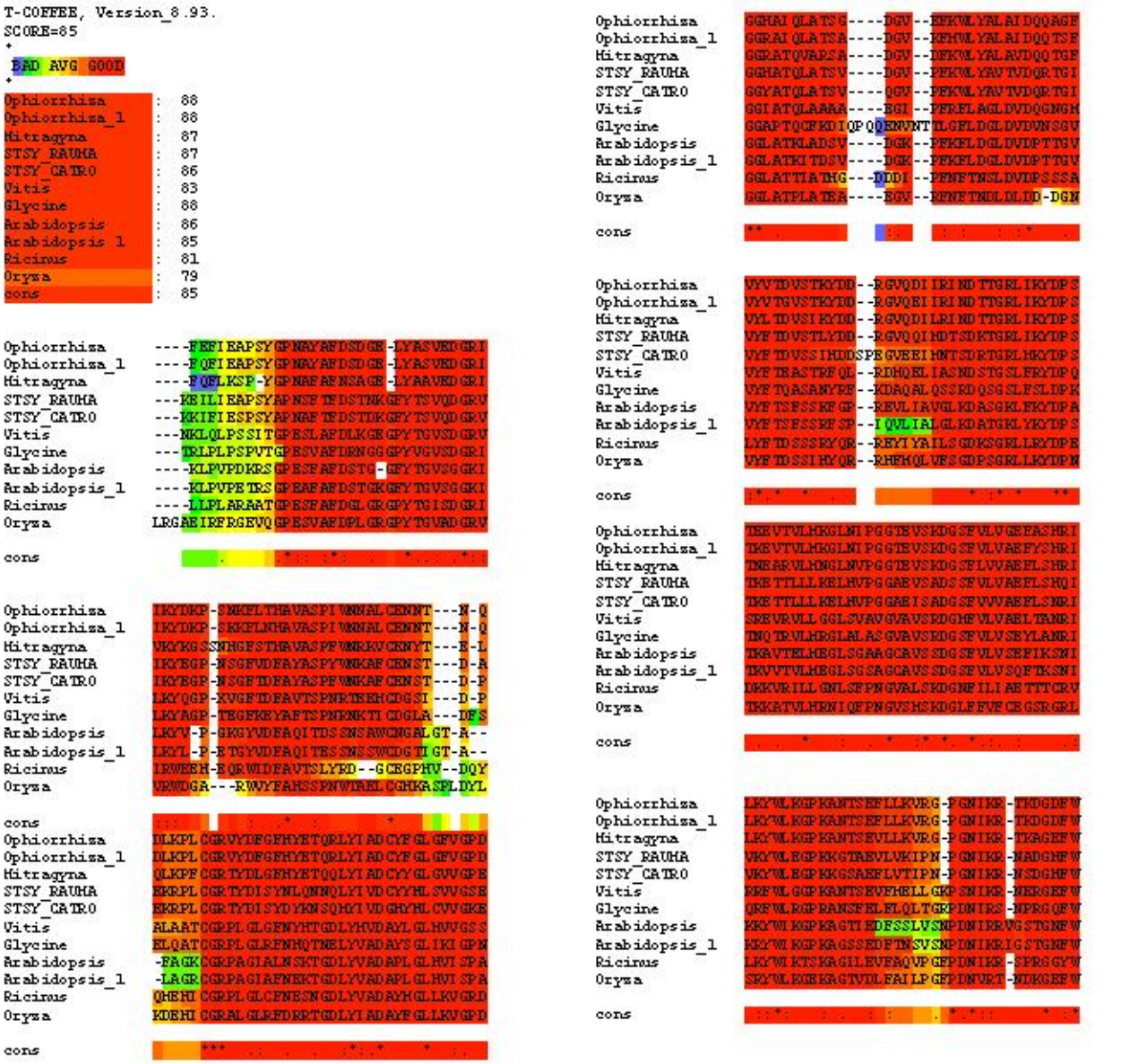


Fig. 2: **shows phylogenetic trees of Enzymes- Strictosidine synthase (STR) from physiochemical data, molecular data and structural data**

Two phylogenetic trees and dendogram belonging to each group are shown in (figure-2, 3 and 4).

Table 4: Domain Family Information for homologues of enzyme Strictosidine synthase (STR)

Enzyme	No. of Domains	Class	Fold	Superfamily	Family
STR Ophiorrhiza pumila BAB47180.1	1	All beta proteins	6-bladed beta-propeller	Calcium-dependent phosphotriesterase	SGL-like
STR Ophiorrhiza japonica ACF21007.1	1	All beta proteins	6-bladed beta-propeller	Calcium-dependent phosphotriesterase	SGL-like
STR Mitragyna speciosa ABZ79473.1	1	All beta proteins	6-bladed beta-propeller	Calcium-dependent phosphotriesterase	SGL-like
STSY_RAUMA P68174.1	1	All beta proteins	6-bladed beta-propeller	Calcium-dependent phosphotriesterase	SGL-like
STSY_CATRO P18417.2	1	All beta proteins	6-bladed beta-propeller	Calcium-dependent phosphotriesterase	SGL-like
STR Vitis vinifera CAN78856.1	2	All beta proteins	6-bladed beta-propeller	1.Calcium-dependent phosphotriesterase 2.NHL repeat	1.SGL-like 2. NHL repeat
STR Glycine max ACU19372.1	3	1,2. All beta proteins 3.Alpha and beta proteins (a+b)	1,2. 6-bladed beta-propeller 3.TBP-like	1.Calcium-dependent phosphotriesterase 2.NHL repeat 3.Bet v1-like	1.SGL-like 2. NHL repeat 3.STAR domain
STR Arabidopsis thaliana NP_177541.1	2	All beta proteins	6-bladed beta-propeller	1.Calcium-dependent phosphotriesterase 2.NHL repeat	1.SGL-like 2. NHL repeat
STR Arabidopsis thaliana AAB40593.1	2	All beta proteins	6-bladed beta-propeller	1.Calcium-dependent phosphotriesterase 2.NHL repeat	1.SGL-like 2. NHL repeat
STR Ricinus communis XP_002532597.1	2	All beta proteins	6-bladed beta-propeller	1.Calcium-dependent phosphotriesterase 2.NHL repeat	1.SGL-like 2. NHL repeat
STR Oryza sativa Japonica Group AAR87254.1	1	All beta proteins	6-bladed beta-propeller	Calcium-dependent phosphotriesterase	SGL-like

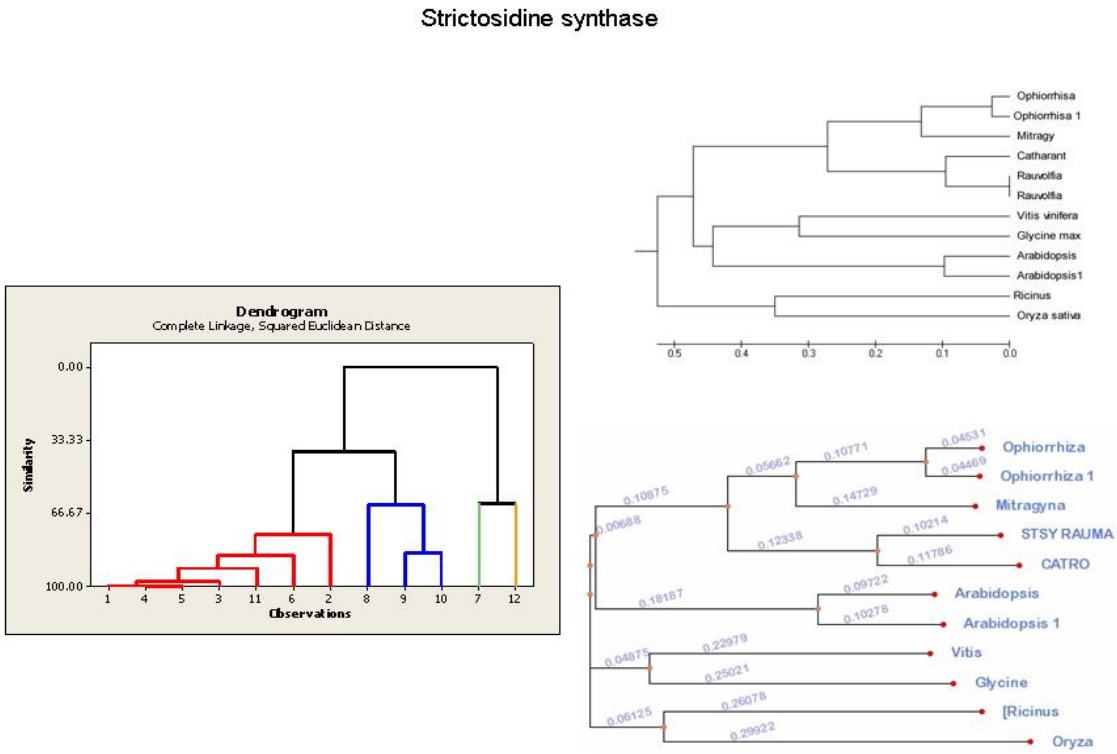


Fig. 3: shows phylogenetic trees of Enzymes- NADPH:cytochrome P450 reductase (CPR) from physiochemical data ,molecular data and structural data

(1-Ophiorrhiza pumila,2-Ophiorrhiza japonica,3-Mitragyna speciosa,4-Rauvolfia serpentine,5-[Rauvolfia mannii,6-Catharanthus roseus,7-Vitis vinifera,8-Glycine max, 9-Arabidopsis thaliana,10-Arabidopsis thaliana,11-Ricinus communis,12-Oryza sativa)

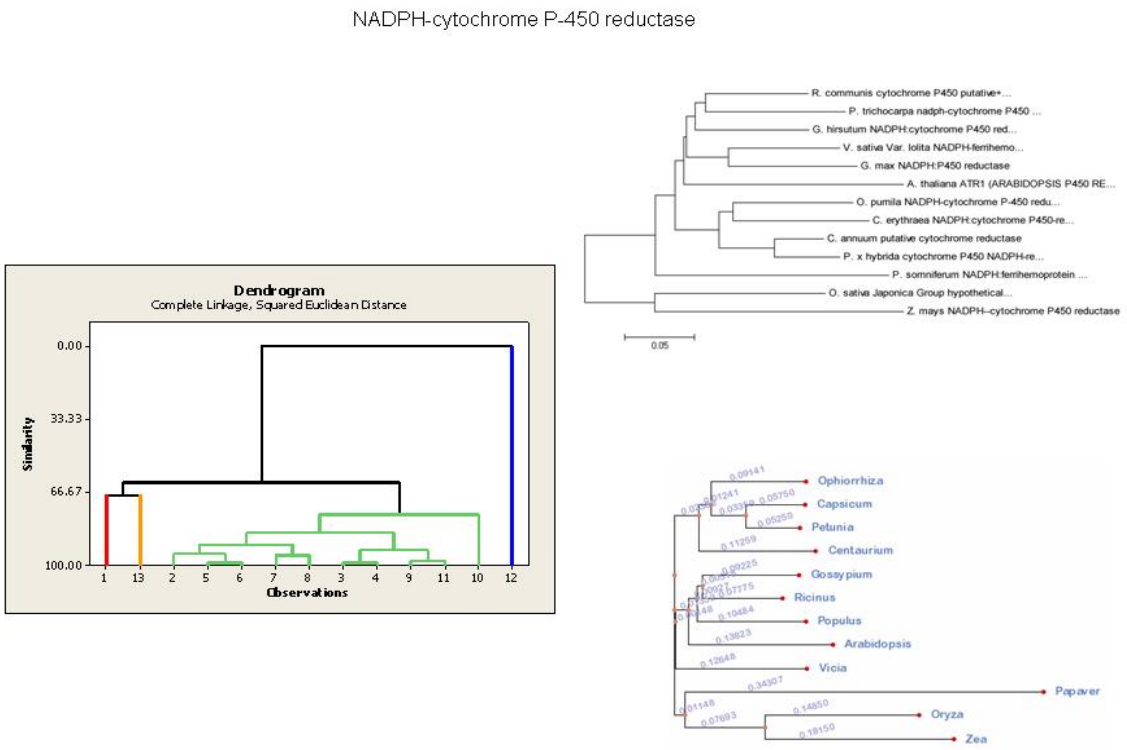
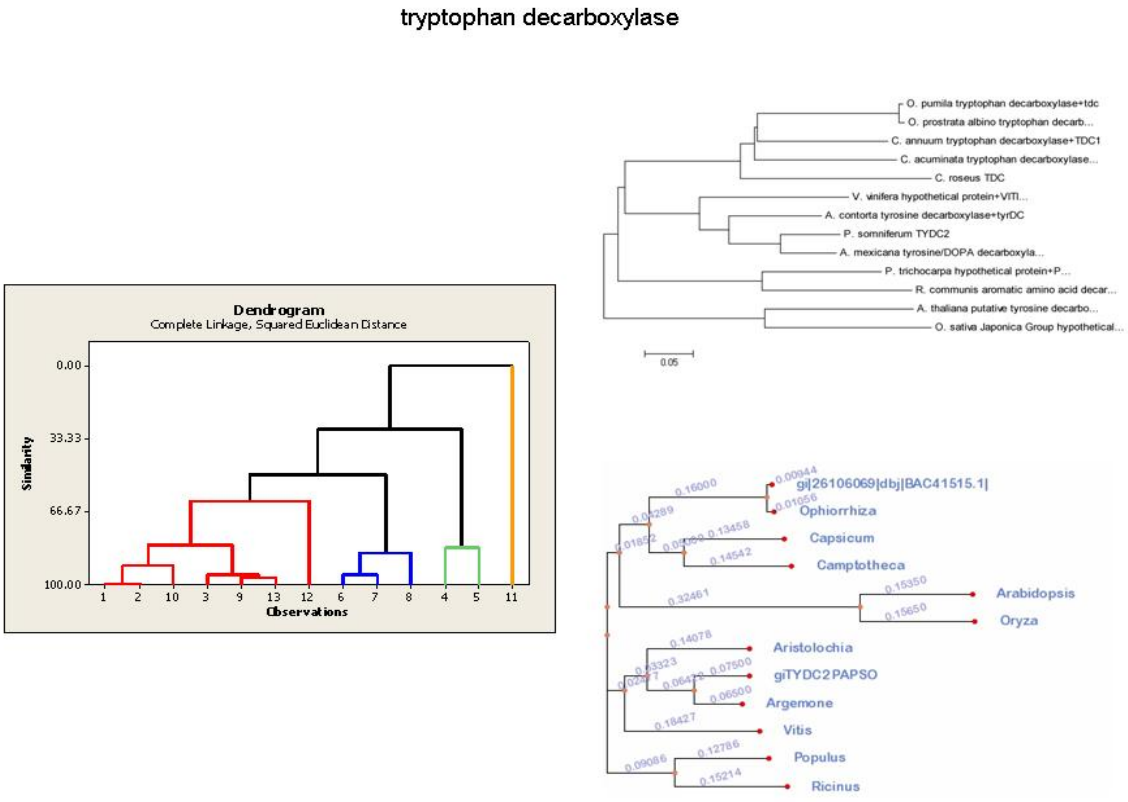


Fig. 4: shows phylogenetic trees of Enzymes- Tryptophan decarboxylase (TDS) from physiochemical data, molecular data and structural data

(1-Ophiorrhiza pumila, 2-Centaurium erythraea, 3-Capsicum annuum,4-Petunia x hybrida,5-Gossypium hirsutum , 6-Ricinus communis, 7-Vicia sativa. 8-Glycine max, 9-Populus trichocarpa, 10-Arabidopsis thaliana, 11-Papaver somniferum, 12-Oryza sativa , 13-Zea mays)



(1-Ophiorrhiza pumila,2-Ophiorrhiza prostrata3-Capsicum annuum,4-Camptotheca acuminate,5-Catharanthus roseus,6-Papaver somniferum,7-Argemone mexicana,8-Populus trichocarpa,9-Ricinus communis10-Arabidopsis thaliana,11-Oryza sativa)

DISCUSSIONS

Drugs of plants origin , maintain the health and vitality of individual and also cure various diseases including cancer without causing toxicity. Due to the adverse side effect of synthetic anticancer drugs, plant origin drug like Camptothecin is currently prescribed in the clinical field for cancer treatment. Camptothecin is still obtained by the extensive extraction from its natural resources which may possess a threat to existing species.

In the present study, an attempt has been made to determine the alternative sources of anticancer drug- Camptothecin through comparative phylogenetic analysis of the five known enzymes involved in its biosynthetic pathway. Feature similarity score (FSS) was calculated for molecular data, structural alignment was performed for phylogenetic relationship of domain superfamily, whereas distance measurement and linkage method was studied for dendogram generation. It was observed that-

1. *Ricinus communis, Oryza sativa* and *Arabidopsis thaliana* contain all the five enzymes (Enzymes-STR, TDC, CPR, SLS, G10H) of Camptothecin biosynthesis.
2. *Vitis vinifera* (STR, TDC, SLS, G10H) and *Populus trichocarpa* (TDC, CPR, SLS, G10H) comprises of four enzymes involved in Camptothecin biosynthesis.
3. *Glycine max* possess three enzymes (STR,CPR,G10H).Whereas, *Capsicum annuum* (TDC,CPR) and *Catharanthus roseus* (TDC,G10H) both possess two-two enzymes each of the Camptothecin biosynthetic pathway.
4. TDC is found in *Capsicum annuum* and *Papaver somniferum* which are distantly related to *Camptothecin acuminata* and *Ophiorrhiza pumila* plants.

On the basis of sequence based phylogenetics tree, physiochemical dendogram and structure based phylogenetic tree, two groups were identified:

A) Less supporting (Enzymes- STR,TDC,G10H), and

b) Well supporting (Enzymes-SLS,CPR) (as shown in table-5)

Table 5: **Comparative result of sequence, structural and physiochemical data based on phylogeny as well as quality of camptothecin enzymes with other plants**

Enzyme	Species	Comparative Result	Supported by
Strictosidine synthase(STR)	*Arabidopsis thaliana* *Ricinus communis* *Oryza sativa*	Less supported	Sequence and structural phylogenetic tree
tryptophan decarboxylase (TDC)	*Arabidopsis thaliana* *Ricinus communis* *Oryza sativa*	Less supported	Structural phylogenetic tree and Physiochemical dendogram
NADPH:cytochrome P450 reductase (CPR)	*Arabidopsis thaliana* *Ricinus communis* *Oryza sativa*	Well supported	Sequence and structural phylogenetic tree and Physiochemical dendogram
Secologanin synthase (SLS)	*Arabidopsis thaliana* *Ricinus communis* *Oryza sativa*	Well supported	Sequence and structural phylogenetic tree and Physiochemical dendogram
Geraniol10-hydroxylase (G10H)	*Arabidopsis thaliana* *Ricinus communis* *Oryza sativa*	Less supported	Sequence and structural phylogenetic tree

Three different organisms, *Arabidopsis thaliana, Ricinus communis* and *Oryza sativa,* have been observed to be related to Camptothecin producing plants sequentially, structurally and physiochemically.

CONCLUSION

This result data is of great significant value as it provides highly authentic information about the alternative source of a drug, by considering not only sequence similarity but also physiochemical relatedness and structural homology. Our work provides a new direction to the research in the area of bioprospecting for discovering alternative sources for the valuable compounds thereby improving the quality of products and their cost effectiveness.

ACKNOWLEDGEMENT

We are thankful to Department of Bioinformatics, MANIT, and Bhopal, India for their support and cooperation in the production of manuscript.

REFERENCES

1. Somya Prakash Rout,K.A Choudary, D.m Kar, Lopamudra Das,Avijeet Jain. Plants in traditional medicinal system- Future sources of new drugs .International Journal of Pharmacy and Pharmaceutical Sciences. 2009; 1(1):1-23
2. Madhuri, S. and Govind Pandey Some dietary agricultural plants with anticancer properties. Pl. Arch., 2008;8(1) : 13-16.
3. S Sivalokanathan; M Ilayara. Efficacy of Terminalia arjuna (Roxb.) on N-nitrosodiethylamine induced hepatocellular carcinoma in rats. Indian Journal of Experimental Biology. 2005; 43: 264-267
4. Yamazaki Mami; Asano, Takashi; Yamazaki, Yasuyo; Sirikantaramas, Supaart; Sudo, Hiroshi; Saito, Kaz .Biosynthetic system of camptothecin: an anticancer plant product. Pure and Applied Chemistry , 2010; report
5. Shicheng Liu, Yiming Yuan, Yutaka Okumura, Norihiro Shinkai and Hitoshi Yamauchi,.Camptothecin disrupts androgen receptor signaling and suppresses prostate cancer cell growth. Biochemical and Biophysical Research Communications,2010; 394(2) : 297-30
6. http://www.greatvistachemicals.com
7. Lorence, C. L. Nessler. Camptothecin, over four decades of surprising findings Phytochemistry ,2004; 65: 2735
8. S. Sirikantaramas, T. Asano, H. Sudo, M. Yamazaki, K. Saito. Camptothecin: therapeutic potential and biotechnology.Curr. Pharm. Biotechnology, 2007; 8 : 196
9. Sharma, S. Chhikara, S. N. Ghodeker, R. Gawande, A. G. Namdeo, K. R. Mahadik, D. P. Fulzele. Phytochemical and

Pharmacological investigations on *Boswellia serrata.* Pharmacognosy Rev.2008; 2: 219

10. K. P. Martin, C.-L. Zhang, M. E. Hembrom, A. Slater, J. Madassery .Adventitious root induction in *Ophiorrhiza prostrata* : a tool for the production of camptothecin (an anticancer drug) and rapid propagation. Plant Biotechnol. Rep.2008; 2: 163-169
11. S. Beegum, K. P. Martin, C.-L. Zhang, I. K. Nishitha, Ligimol, A. Slater, P. V. Madhusoodanan. Organogenesis from leaf and internode explants of Ophiorrhiza prostrata, an anticancer drug (camptothecin) producing plant .Elec. J. Biotechnol.2007; 10: 114
12. E.N Bharath, S.N Manjula,A.Vijaychan .In Silico drug design-tool for overcoming the innovation deficit in the drug discovery process. International Journal of Pharmacy and Pharmaceutical Sciences.2011; 3(2) :8-12
13. Gibrat JF, MadejT, Bryant SH.Surprising similarities in structure comparison. Curr. Opin. Struct. Biol . 1996; 6:377-385.
14. Forst C. and Schulten K.Phylogenetic analysis of metabolic pathways. Journal of Molecular Evolution,2001; 52:471-489.
15. Liao L., Kim S., and Tomb J. F. Genome comparisons based on profiles of metabolic pathways. In Sixth International Conference on Knowledge-Based Intelligent Information and Engineering Systems (KES), Crema, Italy ;2002.
16. Yasuyo Yamazaki, Hiroshi Sudo, Mami Yamazaki, Norio Aimi and Kazuki Saito. Camptothecin Biosynthetic Genes in Hairy Roots of *Ophiorrhiza pumila*: Cloning, Characterization and Differential Expression in Tissues and by Stress Compounds. Plant and Cell Physiology,2003; 44(4): 395-403
17. Irmler, S., Schröder, G., St-Pierre, B., Crouch, N.P., Hotze, M., Schmidt, J., Strack, D., Matern, U. and Schröder, J. Indole alkaloid biosynthesis in *Catharanthus roseus*: new enzyme activities and identification of cytochrome P450 CYP72A1 as secologanin synthase. Plant J.2000; 24: 797–804
18. Yamamoto, H., Katano, N., Ooi, A. and Inoue, K Secologanin synthase which catalyzes the oxidative cleavage of loganin into secologanin is a cytochrome P450. Phytochemistry .2000; 53: 7–12
19. 19.Altschul S. F., Thomas L.M., Alejandro A.S., Jinghui Z., Zheng Z.,Webb M. and David J.L Gapped BLAST and PSI BLAST: a new generation of protein database search programs. Nucleic Acids Res. 1997; 25:3389-3402..
20. Guruprasad, K., Reddy, B.V.B. and Pandit, M.W. Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence. Protein Eng. 1990;4: 155-161.
21. Kyte, J. and Doolittle, R.F.. A simple method for displaying the hydropathic character of a protein.J. Mol. Biol. 1982; 157:105-132
22. Hall B. G. Comparison of the Accuracies of Several Phylogenetic Methods Using Protein and DNA Sequences. Mol. Biol. Evol. 2005; 22: 792-802.
23. Poirot O, Suhre K, Abergel C, O'Toole E, Notredame C. 3DCoffee@igs: a web server for combining sequences and structures into a multiple sequence alignment. Nucleic Acids Res. 2004; 32(Web Server issue):W37-40.
24. Jeanette Tångrot, Lixiao Wang, Bo Kågström and Uwe H. Sauer. FISH—family identification of sequence homologues using structure anchored hidden Markov models. Nucleic Acids Res. 2006; 34(Web Server issue): W10–W14
25. Jakob Fredslund**.** PHY·FI: fast and easy online creation and manipulation of phylogeny color figures .BMC Bioinformatics2006; **7:** 315.