ISSN: 0975-8585

Research Journal of Pharmaceutical, Biological and Chemical Sciences

Bioinformatics Overview of *lantana camara*, an Environmental Weed.

Anita S Goswami-Giri* And Ritesh Oza.

Chemistry Research Laboratory, Department of Chemistry, B. N. Bandodkar College of Science,Thane - 400 601, Maharashtra. India.

ABSTRACT

Chemical and biological principles of medicinal plants have lead to the discovery of newer and cost-effective drugs. *Lantana camara* has been used in various folk medicinal preparations and its chemical composition and pharmacological activities have been elaborated recently. However there remains still a huge scope for use of modern scientific methods - genomics, proteomics, metabolomics and bioinformatics in this plant. Bioinformatics shall facilitate analysis and integration of information from these related fields to enable the identification of genes and gene products and elucidate the functional relationships between genotype and observed phenotype. This review provides a state-of-the-art overview of bioinformatics study of *Lantana camara* with emphasis on the current progress and future directions, which shall provide tools and resources necessary to understand and promote advances in this important field.

Keywords: Phytochemistry, Bioinformatics, *Lantana camara*, Poly (3-hydroxyalkanoates) (PHAs)

Corresponding author

ISSN: 0975-8585

INTRODUCTION

Phytochemistry reports thousands of newer organic compounds and natural products from various plants every year. Pharmacological testing, modifying, derivatising and research on these natural products from such natural biomass sources, represent a major strategy for discovering and developing new drugs [1]. Discovering newer molecular structures from these sources from nature remains a major challenge and source of novel opportunity in combinatorial chemistry, drug designing, bioinformatics, genomics and bioenergetics etc [2]. *Lantana camara* is a rich source of many bioactive molecules like triterpenes, steroids, flavonoids [3,4] and Poly (3-hydroxyalkanoates) (PHAs) [5]**.** The recent progress in bioinformatics techniques has made it possible to obtain information about sequences of protein and nucleotides of *Lantana camara*, which may play an important role in regulation of these bioactive molecules. Yet, further research is still required to fully understand the mechanism of this regulation, which may play an important role in its use as therapeutic agent or as source of important bioactive molecules. Plants have been used as a source of medicine since historic times and several commercially important drugs are of plant-based origin [6]. *Lantana camara* is a storehouse of numerous bioactive molecules. It is a rambling perennial shrub found growing up to 2000 m altitude in tropical, sub tropical and temperate parts of the world including India with a number of flower colors viz. red, pink, white, yellow and violet [7]. It is an important medicinal plant with several medicinal uses in traditional medication system. It is been used to cure many health problems in different parts of the world. Leaves are used to treat cuts, rheumatisms, ulcers, catarrhal infection, tetanus, rheumatism, malaria, cancer, chicken pox, asthma, ulcer, swelling, eczema, tumor, highblood pressure, bilious fever, ataxy of abdominal viscera, sores, measles, fevers, cold and high blood pressure. In Ghana, infusion of the whole plant is used to cure bronchitis and the powdered root in milk was given to children for stomach-ache and as a vermifuge. Lantana oil is used in the treatment of skin, itches, as an antiseptic for wounds. In leprosy and scabies decoctions were applied externally [8].

However the traditional approach towards discovery of plant-based drugs often times involves significant amount of time and expenditure. These labor-intensive approaches have struggled to keep pace with the rapid development of high-throughput technologies. In the era of high volume, high-throughput data generation across the biosciences, bioinformatics plays a crucial role. This has generally been the case in the context of drug designing and discovery. However, there has been limited attention to date to the potential application of bioinformatics approaches that can leverage plant-based knowledge [9]. Here, we review bioinformatics studies that have contributed to *Lantana camara* research. In particular, we highlight areas in *Lantana camara* research where the application of bioinformatics methodologies may result in quicker and potentially cost-effective leads toward finding plant-based remedies or mass production of bioactive molecules.

Current status

Alongside the increase in numbers of sequencing projects, based on numerous crops in different laboratories, there has been a concomitant growth in plant databases [9]. Table 1 list some of major bioinformatics sites used for this purpose viz Swiss Prot, GenBank,

EMBL, DNA Data Bank of Japan etc. One of the more significant changes to plant genome databases has been the move towards graphical user interfaces that provide a more user-friendly search environment. Although graphical user interfaces have been developed, more recent crop databases have used the Ensemble database scheme, which has a strong emphasis on graphical user interaction. A recent advance in crop database interfaces is the use of a standardized client-based scalable vector graphics viewer, which enables data views to be manipulated without the need for constant web page updating.

The diversity of database formats and interfaces reflects the needs of various research groups, but such diversity creates a challenge for bioinformatics because it reduces the scope for data integration and interrogation across databases. With the maturation of genomics has come a move towards the adoption of standard data formats and schemata for crop genome information, and hence current databases has been designed with cross connectivity capabilities as a priority. Complex biological data integration is also driven by developments in grid computing.

Table 1: Major Bioinformatics site.

Topic	Name of database	URLs
Bibliographic databases	Pub Med	http://www.ncbi.nlm.nih.gov/sites/entrez?db=PubMed
	Pub Med Central	http://www.ncbi.nlm.nih.gov/sites/entrez?db=Pmc
	OMIM	http://www.ncbi.nlm.nih.gov/sites/entrez?db=OMIM
	Books	http://www.ncbi.nlm.nih.gov/sites/entrez?db=Books
Sequence databases	EntreZ Nucleotide Sequences	http://www.ncbi.nlm.nih.gov/sites/entrez?db=Nucleotide
	EntreZ Protein Sequences	http://www.ncbi.nlm.nih.gov/sites/entrez?db=Protein
	Gen Bank	http://www.ncbi.nlm.nih.gov/Genbank
	EMBL	http://www.ebi.ac.uk/embl
Protein Database	PDB	http://www.rcsb.org/pdb/home/home.do
	SWISS PROT	http://www.expasy.org/sprot/
	UNI PROT	http://www.uniprot.org/uniprot/
Enzyme database	Enzyme Structure database	http://www.biochem.ucl.ac.uk/bsm/enzymes
	BRENDA	http://www.brenda-enzymes.info

Tools

Primary tools for sequence comparison and assembly have grown in line with an expansion of the datasets that they analyze. Without basic local alignment search tool (BLAST) and related sequence comparison tools, much of the data coming from the many high-throughput sequencing laboratories would be nothing more than strings of letters. BLAST remains the fastest means by which to identify specific sequences in large datasets andenables the rapid annotation of novel sequences. Although BLAST is the standard tool for identifying sequence similarities in large datasets, there are several options for assembling sequence datasets, the choice of which depends on hardware availability, dataset size, data format, structure and the genetic structure of the organism.

ISSN: 0975-8585

Sequence similarity search and assembly tools are the foundation of many software applications for analyzing plant genomic information. The ability to rapidly identify similarities to previously characterized sequences greatly enhances the sequence annotation process. This has led to the development of comparative sequence databases, whereas sequence assembly packages both reduce the high level of redundancy in datasets and enable variations in sequence to be identified.

Proteomics

Proteomics currently encompasses databases of protein sequences, databases of predicted protein structures and, more recently, databases of protein expression analysis, and the field is expanding with emerging technologies.

The principal protein sequence database remains SwissProt (Table 1), which was established in 1986 as a repository for predicted protein sequences and now contains multi-level protein data. In the bid to link the genome and the proteome with associated phenotypes, there has been a push towards the prediction of protein structures in relation to their sequence. This drive has come mainly from the pharmaceutical industry, although structure prediction has applications in plant biotechnology research [8].

Table 2 gives information about different proteins from *Lantana camara*, which have been studied using bioinformatics database UniProt [10]**.**

Table 2

No	Entry	Entry Name	Status	Protein names	Gene Names	Organism	Length
1	F1BDZ3	F1BDZ3_LANCA	unreviewed	30S ribosomal protein S3, chloroplastic (Fragment)	rps3	*Lantana camara* (Lantana)	195
2	Q95D38	Q95D38_LANCA	unreviewed	ATP synthase beta subunit (Fragment)	atpB	*Lantana camara* (Lantana)	283
3	F1BDM4	F1BDM4_LANCA	unreviewed	Cytochrome c heme attachment protein (Fragment)	ccsA	*Lantana camara* (Lantana)	294
4	C9DTU8	C9DTU8_LANCA	unreviewed	DNA-directed RNA polymerase (EC 2.7.7.6) (Fragment)	rpoC1	*Lantana camara* (Lantana)	238
5	K0A6L0	K0A6L0_LANCA	unreviewed	Flowering locus T (Fragment)	FT	*Lantana camara* (Lantana)	50
6	G3EIV1	G3EIV1_LANCA	unreviewed	MADS box transcription factor (Fragment)	GLO1	*Lantana camara* (Lantana)	187
7	G3EIQ2	G3EIQ2_LANCA	unreviewed	MADS box transcription factor (Fragment)	DEFB1	*Lantana camara* (Lantana)	213
8	G3EIR1	G3EIR1_LANCA	unreviewed	MADS box transcription factor (Fragment)	DEFA	*Lantana camara* (Lantana)	225
9	G3EIQ3	G3EIQ3_LANCA	unreviewed	MADS box transcription factor (Fragment)	DEFB2	*Lantana camara* (Lantana)	217

10	I1UHG8	I1UHG8_LANCA	unreviewed	Maturase K (Fragment)	matK	*Lantana camara* (Lantana)	259
11	I1UHH2	I1UHH2_LANCA	unreviewed	Maturase K (Fragment)	matK	*Lantana camara* (Lantana)	249
12	Q8HVW8	Q8HVW8_LANCA	unreviewed	Maturase K (Fragment)	matK	*Lantana camara* (Lantana)	444
13	G0X197	G0X197_LANCA	unreviewed	Maturase K (Fragment)	matK	*Lantana camara* (Lantana)	272
14	G8DT59	G8DT59_LANCA	unreviewed	Maturase K (Fragment)	matK	*Lantana camara* (Lantana)	257
15	C9DTS4	C9DTS4_LANCA	unreviewed	Maturase K (Fragment)	matK	*Lantana camara* (Lantana)	277
16	D2JX66	D2JX66_LANCA	unreviewed	Maturase K (Fragment)	matK	*Lantana camara* (Lantana)	275
17	F1BDQ7	F1BDQ7_LANCA	unreviewed	Maturase K (Fragment)	matK	*Lantana camara* (Lantana)	386
18	I1UIF5	I1UIF5_LANCA	unreviewed	Maturase K (Fragment)	matK	*Lantana camara* (Lantana)	258
19	I1UHG9	I1UHG9_LANCA	unreviewed	Maturase K (Fragment)	matK	*Lantana camara* (Lantana)	255
20	I1UHH0	I1UHH0_LANCA	unreviewed	Maturase K (Fragment)	matK	*Lantana camara* (Lantana)	248
21	I1UHH3	I1UHH3_LANCA	unreviewed	Maturase K (Fragment)	matK	*Lantana camara* (Lantana)	241
22	I1UHH1	I1UHH1_LANCA	unreviewed	Maturase K (Fragment)	matK	*Lantana camara* (Lantana)	253
23	E7CL85	E7CL85_LANCA	unreviewed	NADH dehydrogenase subunit F (Fragment)	ndhF	*Lantana camara* (Lantana)	695
24	K4PL96	K4PL96_LANCA	unreviewed	Pentatricopeptide repeat-containing protein 11 (Fragment)	Nil	*Lantana camara* (Lantana)	431
25	K4PAT0	K4PAT0_LANCA	unreviewed	Pentatricopeptide repeat-containing protein 123 (Fragment)	Nil	*Lantana camara* (Lantana)	381
26	K4PAP4	K4PAP4_LANCA	unreviewed	Pentatricopeptide repeat-containing protein 81 (Fragment)	Nil	*Lantana camara* (Lantana)	373
27	K4P8X8	K4P8X8_LANCA	unreviewed	Ribosomal protein L32 (Fragment)	rpl32	*Lantana camara* (Lantana)	52
28	Q9THN3	Q9THN3_LANCA	unreviewed	Ribulose 1,5-bisphosphate carboxylase large subunit (Fragment)	rbcL	*Lantana camara* (Lantana)	478
29	G8DSD4	G8DSD4_LANCA	unreviewed	Ribulose bisphosphate carboxylase large chain (EC 4.1.1.39) (Fragment)	rbcL	*Lantana camara* (Lantana)	184
30	M1GF68	M1GF68_LANCA	unreviewed	Ribulose bisphosphate carboxylase large chain (EC 4.1.1.39) (Fragment)	rbcL	*Lantana camara*	242

						(Lantana)	
31	M1GIR4	M1GIR4_LANCA	unreviewed	Ribulose bisphosphate carboxylase large chain (EC 4.1.1.39) (Fragment)	rbcL	*Lantana camara* (Lantana)	245
32	M1GCK2	M1GCK2_LANCA	unreviewed	Ribulose bisphosphate carboxylase large chain (EC 4.1.1.39) (Fragment)	rbcL	*Lantana camara* (Lantana)	240
33	M1GFG7	M1GFG7_LANCA	unreviewed	Ribulose bisphosphate carboxylase large chain (EC 4.1.1.39) (Fragment)	rbcL	*Lantana camara* (Lantana)	240
34	D2JXM9	D2JXM9_LANCA	unreviewed	Ribulose bisphosphate carboxylase large chain (EC 4.1.1.39) (Fragment)	rbcL	*Lantana camara* (Lantana)	189
35	F1BDT6	F1BDT6_LANCA	unreviewed	Ribulose-1,5-bisphosphate carboxylase large subunit (Fragment)	rbcL	*Lantana camara* (Lantana)	163
36	I1UWL6	I1UWL6_LANCA	unreviewed	Ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (Fragment)	rbcL	*Lantana camara* (Lantana)	123
37	G0WYS9	G0WYS9_LANCA	unreviewed	Ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (Fragment)	rbcL	*Lantana camara* (Lantana)	453
38	F1BDW6	F1BDW6_LANCA	unreviewed	RNA polymerase beta II subunit (Fragment)	rpoC2	*Lantana camara* (Lantana)	701
39	E6Y8N6	E6Y8N6_LANCA	unreviewed	Superoxide dismutase (EC 1.15.1.1)	Nil	*Lantana camara* (Lantana)	224
40	E6Y8N5	E6Y8N5_LANCA	unreviewed	Superoxide dismutase [Cu-Zn] (EC 1.15.1.1)	Nil	*Lantana camara* (Lantana)	152
41	K7U8U0	K7U8U0_LANCA	unreviewed	Superoxide dismutase [Cu-Zn] (EC 1.15.1.1) (Fragment)	SOD	*Lantana camara* (Lantana)	129

From the above sequences, further in-silico genomic analysis on maturase K (matK) gene has lead to the proposal of *Lantana camara* as an inexpensive and abundant source for biological production of Poly (3-hydroxyalkanoates) (PHAs) like poly-β-hydroxybutyrate (PHB). PHBs are a class of microbial polyesters that have potential applications as conventional plastics, specifically thermoplastic elastomers and this polymer can accumulate up to 90% of the cellular dry weight of some bacteria. Maturase K gene (matK) of chloroplast is highly conserved in plant, which is involved in intron splicing. These simulations focused on the maturase K (matK) gene – length of 1331 base pairs (bp) – of the chloroplasts, which affects the intron splicing. The computational approach considered a reconstructed chloroplast genome of *Lantana camara*, the trnL–trnFintergenic spacer (IGS), and trnH–psbA IGS in the chloroplast genome. The Blastn result indicated that the Verbenaceae genus Lantana displays 86% sequence identity with that of Solanaceae genus Nicotiana. Phylogenetic tree was constructed to identify an ideal comparison of *Lantana camara* matK gene and other plant species used for the insertion of phbA, phbB, and phbC enzymes, which regulate PHB synthesis. These computational results also proposed that the chloroplast transformation may be useful in biological production of PHB [5].

In addition to the above list, it has been also proposed that, bioinformatics tools can be used to compare different inhibitors of alkaline proteases of*Lantana camara*. These alkaline proteases hold a great potential for application in the detergent, leather industries

and their easy extraction from the senesced leaves can make *Lantana camara*, an economical choice for their large scale production [11]. This increases scope of the use of *Lantana camara* in large scale of production of such bioactive molecules along with its therapeutic uses.

Future prospects

Proteomics has significant prospects for advancing our understanding in plant biotechnology owing to its direct relationship with gene and transcript data. They also have a strong influence on the measured phenotype of the plant, either directly through protein content or function or indirectly through the relationship of a protein with the metabolome. The potential for bioinformatics to structure and integrate-omic data, therefore, relies on an ability to model both the proteome and its interactions [9].

CONCLUSION

Lantana camara is a waste land weed having various pharmacological properties and is a source of beneficial secondary metabolites. Computational studies of its key biosynthetic enzymes may provide valuable insights into the mechanism of action of the enzymes aiding in the ultimate aim of improving quality and quantity of the reaction products. Consequently, the use of appropriate computing and bioinformatics tools to allow automated data storage and efficient non-labor intensive data analysis may further help in metabolomics analysis of *Lantana camara*. This may then be used either for 'fingerprinting' samples to perform comparative analyses to detect differences for 'profiling' where individual differential secondary metabolites like alkaloids, phenolic components, are identified for further analysis.

With this background, the integration and structured interrogation of metabolome and transcriptome datasets yields result, that provides the basis for the integration of genome and phenome data. Linking gene expression, protein sequence and protein structure data with genetic and physical map data integrates genetics, genomics, transcriptomics and proteomics. The further incorporation of metabolomic and phenotypic data closes the loop and create the foundation for advanced knowledge bases – in otherwords, meta-integrated databases that facilitate queries across whole-systems biology[9].

REFERENCES

[1] Kayser O, Kidderlen AF and Croft SL. ActaTropica 2000;77:307 – 314.
[2] DK Sharma. J Sci Ind Res2006;65:477- 484.
[3] S Begum, SQ Zehra and BS Siddiqui. Chem Pharm Bull 2008;56:1317-1320.
[4] DK Verma, SK Singh and V Tripathi. Indian Drugs 1997;34:32-35.
[5] Protiti Bose, Rahul Shrivastava. J Plastic Film Sheeting 2012;28:228-243.
[6] Indra Neil Sarkar, Vivekanand Sharma. Brief Bioinform 2013;14(2):238-250.
[7] Dinesh Kumar, Ajay Kumar and Om Prakash. Asian Pacific J Trop Dis 2012;S42-S45.
[8] Sanjeeb Kalita, Gaurav Kumar, Loganathan Karthik, Kokati Venkata, Bhaskara Rao. Res J Pharm Tech 2012;5(6):711-715.

[9] David Edwards and Jacqueline Batley. Trends Biotechnol 2004;22(5):232 – 237.
[10] http://www.uniprot.org/uniprot/?query=lantana+camara&sort=score
[11] Smriti Gaur and Neeraj Wadhwa. African J Biotechnol 2008;7(24):4602 – 4608.