ANALYSIS OF DE NOVO TRANSCRIPTOME SEQUENCE FOR IDENTIFICATION OF PUTATIVE CANDIDATE GENES RELATED TO β- ODAP BIOSYNTHESIS PATHWAY IN *LATHYRUS SATIVUS*

M. Sc. (Ag.) Thesis

By

Jajati Keshari Nayak

DEPARTMENT OF PLANT MOLECULAR BIOLOGY AND BIOTECHNOLOGY
COLLEGE OF AGRICULTURE
INDIRA GANDHI KRISHI VISHWAVIDYALAYA, RAIPUR (Chhattisgarh)

2019

ANALYSIS OF DE NOVO TRANSCRIPTOME SEQUENCE FOR IDENTIFICATION OF PUTATIVE CANDIDATE GENES RELATED TO β- ODAP BIOSYNTHESIS PATHWAY IN *LATHYRUS SATIVUS*

Thesis

Submitted to the

Indira Gandhi Krishi Vishwavidyalaya, Raipur

By

Jajati Keshari Nayak

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF

Master of Science
in
Agriculture

(Plant Molecular Biology and Biotechnology)

UE ID No. 20171827404 College ID No. 120117175

July, 2019

CERTIFICATE – I

This is to certify that the thesis entitle "Analysis of De novo transcriptome sequence for identification of putative candidate genes related to β ODAP biosynthesis pathway in *Lathyrus sativus*" submitted in partial fulfilment of the requirements for the degree of **Master of Science in Agriculture (Plant Molecular Biology and Biotechnology)** of the Indira Gandhi Krishi Vishwavidyalaya, Raipur, (Chhattisgarh) is a record of the bonafide research work carried out by **Jajati Keshari Nayak** under my/our guidance and supervision. The subject of the thesis has been approved by Student's Advisory Committee and the Director of Instructions.

No part of the thesis has been submitted for any other degree or diploma (certificate course). All the assistance and help received during the course of the investigations have been duly acknowledged.

Date: 27/7/19

Dr. Shubha Banerjee
Chairman

THESIS APPROVED BY THE STUDENT'S ADVISORY COMMITTEE

Chairman	Dr. Shubha Banerjee	
Member	Dr. S.B.Verulkar	27/7
Member	Dr. Sunil K Nair	
Member	Dr. R.R. Saxena	

CERTIFICATE – II

This is to certify that the thesis entitled "Analysis of De novo transcriptome sequence for identification of putative candidate genes related to β ODAP biosynthesis pathway in *Lathyrus sativus*" submitted by **Jajati Keshari Nayak** to the Indira Gandhi Krishi Vishwavidyalaya, Raipur in partial fulfillment of the requirements for the degree of **M.Sc. (Agriculture)** in the **Department of Plant Molecular Biology and Biotechnology** has been approved by external examiner and Student's Advisory Committee after oral examination.

Signature External Examiner

Date: 03.08.2019

Major Advisor ____________________

Head of the Department ____________________

Faculty Dean ____________________

Approved/Not approved

Director of Instructions ____________________

ACKNOWLEDGEMENTS

*I would like to express my heartful regards towards "**MY PARENTS**"who allowed me to stand in the world and hoe to tackle the struggling situation.*

*Life is meaningless and directionless without a monitor. I emphatically extend my loyal and vulnerable thanks towards the meticulous guidance extended by Chairperson of my Advisory Committee **Dr. Shubha Banerjee** for innovative ideas, motivation for hardwork, constructive criticism, methodological approach and invaluable guidance.*

I offer my sincere thanks to members of my advisory committee, Dr. S.B.Verulkar (Professor and head, Department of Plant Molecular Biology and Biotechnology) for his generous help and invaluable suggestion. I am very thankful to Dr. Dr. Sunil K Nair (Department of Genetics and Plant Breeding) and Dr. R. R. Saxena (Department of Agricultural Statistics & Social Science) for their guidance and timely help.

Furthermore, I want to extend my appreciation to the Dr. Archana Prasad, Dr. Girish Chandel, Dr. Zenu Jha and other technical staff of Department of Plant Molecular Biology and Biotechnology. .

I would like to mention my sincere thanks to my seniors viz, Miru mam, Helen mam, Arun sir, Shrikant sir, Datta sir, Sanjay bhariya sir,Mayur sir ,Abhiram sir for their help and support.

It is also pleasure to mention my good friends, Like Lohitha,Pooja,Ankit,Taruna,Shriraj,Ritesh,Kajal,Savita,Archana,Ritesh,Sunil, Shambhu for their support.

My thanks are also to office staff of the Department of PMBB for helping me during my studies. Thanks to one and all and to those whose names could not appear but who at one stage or the other has helped me.

Diction is not enough to express the heartful gratitude to my best friend Priyadarsini sahu and Debajani Nayak , whose filial affection, encouragement, selfless help, sincere prayer and love have always been the most vital source of inspiration and raising me to the present position. I am also very much thankful to my Roommates Sudipta sourav sahoo and Biswajit Sir for their support, affection and *encouragement.*

At last, I would like to convey my grateful thanks to all those unmentioned persions who helped me to see my dream comes true.

Jajati Keshari Nayak

Department of Plant molecular Biology and Biotechnology

College of Agriculture,IGKV,Raipur.(C.G) 492012

Date : 27/7/19

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

%	Per cent
^{o}C	Degree Celsius
β-ODAP	β-*N*-oxalyl-l-α,β-diaminopropionic acid
C.G.	Chhattisgarh
TPM	Transcripts per Million
SAT	serine acetyltransferase
RT-PCR	Real Time PCR
RNA	Ribonucleic acid
PLP	Pyridoxal 5'-phosphate
OPT	o-phthalaldehyde
NGS	Next Generation Sequencing
NCBI	National center for Biotechnological information
KEGG	Kyoto Encyclopedia of Genes and Genomes
IDT	Integreted DNA Technology
GO	Gene Ontology
CS	Cysteine synthase
CAS	Cyanoalanine synthase
BLAST	Basic Local Alignment Search Tool
DNA	Deoxyribo nucleic Acid
et al.	And others/Co-workers/ et alia
PCR	Polymerase chain reaction
Fig.	Figure
Ha	Hectare
i.e.	That is
Mm	Milimetre
max	Maximum

THESIS ABSTRACT

a) Title of the Thesis: "Analysis of De novo transcriptome sequence for identification of putative candidate genes related to β ODAP biosynthesis pathway in *Lathyrus sativus*"

b) Full Name of Student: Jajati Keshari Nayak

c) Major Subject: Plant Molecular Biology and Biotechnology

d) Name and Address of the Major Advisor: Dr. Shubha Banerjee (Asst. Professor) Dept. of Plant Molecular Biology and Biotechnology, College of Agriculture, IGKV, Raipur.

e) Degree to be Awarded: M.Sc.(Ag.) Plant Molecular Biology and Biotechnology

Signature of the Student

Signature of Major Advisor

Signature of Head of the Department

Date: 27.07-2019

Grass pea (*Lathyrus sativus)* is a valuable pulse crop of winter season in terms of low water requirement, tolerance to major abiotic stress, high protein content (~31.6%) and low input cost. It is grown as a fellow crop after rice in tropical region of India, Bangladesh, Nepal, Ethiopia etc. The association of Lathyrus with neurolathyrism, motor degenerative disease caused in human on consumption of L. sativus pulse has led to ignorance of this wonderful pulse crop in main-stream agriculture. The neurological disorder is caused because of presence of a non-protein amino acid β-ODAP in L. sativus seeds which is synthesized as a by-product of by O-acetyl serine metabolism pathway (Malathi *et al.,* 1970). Although the pathway of ODAP biosynthesis was reported long back but the genes which codes the proteins and key enzymes in the pathway are not elucidated yet. Therefore to identify the genes involved in biosynthesis of β-ODAP, RNA sequencing and *de-novo* transcriptome analysis was done.

Leaf tissue transcriptome data of two contrasting ODAP content variety (Mahateora and RLK-1950) was analysed using homology search, protein modelling and enzyme – ligand docking techniques. About 28,000 transcripts were obtained in both the genotypes which were initially annotated using *Phaseolus vulgaris* gene sequences (Banerjee et al., 2017 unpublished data).

NCBI BLAST and LOCAL BLAST was done to classify transcripts based on their function and localization in cell. ODAP is synthesized in mitochondrial and chloroplast of plants where amino acid transferase and synthase classes of enzymes play important role thus sequence alignment search of about 14000 transcripts for similarity with amino acid transferase or synthase enzymes with localization Mitochondria or chloroplast was done. Further Homology modelling through swiss model workspace (https://swissmodel.expasy.org/interactive) to assess protein structure identity and similariy and Modbase (https://modbase.compbio.ucsf.edu/modweb/) narrowed down to 11 putative candidate transcripts. Docking of the protein corresponding to the candidate transcripts with ligands indicated involvement of two isoforms of candidate gene L-alanine synthase enzymes in the process. Conserve domain search also showed that the candidate genes are involved in cysteine metabolism and have role in amino acid biosynthesis pathway. Pathway mapping was done by BLAST2GO and NCBI biosystem to deduce the involvement of putative candidate genes in various physiological and biochemical pathways. Sequence search and protein modelling led to identification of 11 candidate genes which play key role in ODAP synthesis.

Six primers were designed from these potential transcripts and used for expression profiling in leaf tissue of selected *L sativus* genotypes having different seed ODAP contents. Differential expression of the primers PLPD1 was observed in the high and low seed ODAP containing genotypes. The primers PLPD1 showed two amplicons of length 150 bp and 200 bp, again correlating to the two isoforms found in the sequence analysis. Of these the longer amplicos 250 nt was found to be correlated to high seed ODAP content and the smaller amplicon of 150 nt with low seed ODAP content in Mahateora and Pusa-24. So far, to the best of our knowledge this is the first report of identification of two putative candidate genes encoding L-alanine synthase isoforms in *L. sativus*. Further sequencing of the amplicons will be done to confirm their role in ODAP biosynthesis and isolate the candidate genes encoding two isoforms of L-alanine synthase in *L. sativus*.

शोध सारंश

शोध का शीर्षक	:	लाथिरस सैटिवस में बीटा ओडैप बायोसिंथेसिस पाथवे से संबंधित पुटीय उम्मीदवार जीन की पहचान के लिए डी नोवो प्रतिलेख अनुक्रम का विश्लेषण
विद्यार्थी का पूरा नाम	:	जजाती केसरी नायक
मुख्य विषय	:	पादप आणविक जीव विज्ञान एवं जैव प्रौद्योगिकी
मुख्य सलाहकार का नाम व पता	:	डॉ शुभा बनर्जी (सहायक प्राध्यापक) पादप आणविक जीव विज्ञान एवं जैव प्रौद्योगिकी विभाग, कृषि महाविद्यालय, इं. गा. कृ. वि., रायपुर (छ.ग.), 492012
डिग्री से सम्मानित किया जाना है	:	एम. एससी. (कृषि, पादप आणविक जीव विज्ञान एवं जैव प्रौद्योगिकी.)

विद्यार्थी के हस्ताक्षर

मुख्य सलाहकार के हस्ताक्षर

दिनांकः 27/7/15

विभागाध्यक्ष के हस्ताक्षर

सारंश

तिवरा (लैथिरस सैटियस) कम पानी की आव"यकता, प्रमुख अजैविक तनाव, उच्च प्रोटीन सामग्री (31–6 प्रति"ात) और कम इनपुट लागत के संदर्भ में सर्दियों के मौसम की की मूल्यवान दलहन फसल है। यह भारत, बांग्लादे"ा, नेपाल, इथियोपिया आदि के उष्णकटिबंधीय में चावल के बाद ,क साथी फसल के साथ में उगया जाता है। लैथिरस सैटियस सातिवस नाड़ी के सेवन पर मानव में होने वाली न्यूरोलैथिज्म, मोटर डिजनरेटिव बीमारी के साथ लैथिरस का संबंध इस अद्‌भुत पल्स की अज्ञानता का कारण बना है। मुख्य धारा की कृषि में फसल। न्यूरोलॉजिकल डिसऑर्डर β-ODAP बीज में L. sativus प्रोटीन अमीनो की उपस्थिति के कारण होता है, जिसे O-acetyl सेरीन मेटाबॉलिज्म मा॰ (Malathi *et al.*, 1970) }ारा उप–उत्पाद के :प में सं'ले"ित किया जाता है। हालांकि ओडैप बायोसिंथेसिस का लंबे समय पहले बताया गया था, लेकिन जिन जीनों में प्रोटीन और प्रमुख एनजाइम होते हैं, उन्हें अभी तक स्पस्ट नहीं किया गया है। इसलि, β-ODAP, RNA में 'ामिल जीनों की पहचान करने के लिये अनुक्रम और

डे–नोवो ट्रांस्क्रिप्टोम वि”ालेषण किया या था। दो वि’'ाम ओडी,पी सामग्री किस्म (Mahateora and RLK-1950) के लीफ टि”ाू ट्रांसक्रिप्टानल डेटा का वि”सले”ाण होमोलॉजी खोज, प्रोटीन मॉडलिंग और ;जाइम – तकनीकों का उपयोग करके किया गया था। दोनों जीनोटाइप्स में लगभग 28,000 टेप प्राप्त किये, थे, जिन्हें *Phaseolus vulgaris* जीन अनुक्रमों (Banerjee et al., 2017) का उपयोग किया गया था।

NCBI BLAST और LOCAL BLAST को सेल में उनके कार्य और स्थानीयकरण के आधार पर टेप का वर्गीकरण करने के लि, किया गया था। ODAP को पौधों के माइटोकॉन्ड्रियल और क्लोरोप्लास्ट में सं'लेषित किया जाता है, जहां एजाइमों के अमीनो ,सिड ट्रांसफरेज और सिंथेज महत्वपूर्ण भूमिका निभाती हैं, इस प्रकार स्थानीयकरण माइटोकोंड्रिया या क्लोरोप्लास्ट के साथ ,मिनो ,सिड ट्रांसफरेज या सिंथेज ;जाइमों के साथ समानता के लिये, लगभग 14000 टेपों के अनुक्रम संरेखा खोज किया गया था। प्रोटीन संरचना की पहचान और सिमिलियारी और मोडबेस (https://swissmodel.expasy.org/interactive) 11 सूचक तक सीमित करने के लि, स्विस मॉडल कार्य (https://modbase.compbio.ucsf.edu/modweb/) के माध्यम से होमोलॉजी मॉडलिंग। उम्मीदवार का संकेत। लिं•ड्स के साथ उम्मीदवार प्रतिलेख के अनुसार प्रोटीन का डॉकिंग प्रक्रिया में उम्मीदवार जीन ,ल–अलैनिन सिंथेस एजाइमों के दो आइसोफॉर्म की भागीदारी को इंगित करता है। संरा"ित डोमेन खोज से यह भी पता चला है कि उम्मीदवार जीन सिस्टीन चयापचय में शामिल हैं और अमीनो ,सिड बायोसिंथेन मार्ग में भूमिका है। विभिन्न शारीरिक और जैव रासायनिक रास्ते में पुट्टी उम्मीदवार जीन की भागीदारी को कम करने के लिये BLAST2GO और NCBI बायोसिस्टम धारा पाथवे मैपिं• की गई थी। अनुक्रम खोज और प्रोटीन मॉडलिंग ने 11 उम्मीदवार जीनों की पहचान की जो ओडैप सं”लेि"षत में महत्वपूर्ण भूमिका निभाते हैं।

छह प्राइमरों को इन संभावित टेपों से डिजाइन किया गया था और अलग–अलग बीज ODAP सामग्री वाले चयनित *L sativus* जीनोटाइप के पत्ती ऊतक में अभिव्यक्ति प्रोफाइलिंग के लिये उपयोग किया गया था। प्राइमरों की अलग–अलग अभिव्यक्ति जीनोटाइप वाले उच्च और निम्न बीज ओडैप में देखी गई थी। प्राइमर PLPD1 ने लंबाई के दो आयाम दिखा, 150 और 250, फिर से अनुक्रम वि'ले”ाष में पाया .दो आइसोफोर्मों से संबंधित। इनमें से लंबे समय तक ,म्पलीकॉस (250 nt) को उच्च बीज ओडैप सामग्री से संबंधित पाया गया और महाथोरा और पूसा –24 में कम बीज वाली ओडैप सामग्री के साथ (150 nt) के छोटे को अब तक हमारे ज्ञान का सबसे अच्छा करने के लिये ,ल। "ैतानी में ,ल–अलैनिन सिंथेस आइसोफॉर्म को ,न्कोडिं• करने वाले दो पु''ठीय उम्मीदवार जीन की पहचान की यह पहली रिपोर्ट है। इसके अलावा ,म्पलीकॉन्स की अनुक्रमाक ODAP में उनकी भूमिका की पुष्टि करने के लिये किया जाता और L-alanine synthase in *L. sativus* उम्मीदवार जीन अलग है।

CHAPTER – I

INTRODUCTION

Pulse remained a major food component as a supplier of protein source. It plays a vital role in minimizing the malnutrition and maintaining sustainable agariculture. Increasing pulse production is directly correlates with availability of enriched food, increases soil health, maintainins diet quality and ensures nutrition security. An investigation made by ICRISAT has revealed that there is no remarkable increase in pulse production from 2000 to 2014. Food security is the ensurement of high quality protein and high caloric energy. Source of protein mainly comes from animal sources and plant sources like pulse,legume and nuts.In a survey it has observed that 72% of total world population solely depend upon the pulse for proten content.Consumption of animal diet for protein is only 7%.

(*Lathyrus sativus L.*) is a diploid (2n=14), self-pollinated pulse Crop. According to Allkin *et al*. 1983 genus Lathyrus has 187 species and sub-species. Old World species have been originated from Asia Minor and the Mediterranean region (Zeven and de Wet 1982). However, only one species *Lathyrus sativus* is widely cultivated as a food crop (Jackson and Yunus 1984), while other species are cultivated to a lesser extent for both food and forage. Some species are valued as ornamental plants, especially the sweet pea (*L. odoratus*). It is cultivated mainly in Chhattisgarh, Bihar, Madhya Pradesh (Haque and Manan, 1989; Hanbury *et al*., 1999).Being hardy crop with low moisture requirement, Lathyrus is a crop of immense economic significance, especially in developing nations including India, Bangladesh, Pakistan, Nepal and Ethiopia (Dixit *et al.*, 2016). It is often broadcast-seeded into standing rice as fellow crop. Grass pea is the third most important cool-season pulse crop of India, occupying an area of 0.58 million ha with an annual production of 0.43 million tonnes (http://agricoop.nic.in/). The grass pea is endowed with many properties that combine to make it an attractive food crop in drought-stricken, rain-fed areas where soil quality is poor and extreme environmental conditions prevail (Palmer *et al*. 1989). It is called as insurance crop of farmer where water is not available for subsequent winter crops.

The grass pea is utilized in many areas of South East Asia as a ‘utera’ crop. The seeds are broadcast into a standing rice crop. When the water is drained to allow for harvesting of the rice, the seeds germinate and the crop utilizes the remaining moisture for growth. Normally the smaller-seeded type lines are utilized as the growers believe that they will remain dormant longer under flooded conditions and give better germination when the water is drained. In many areas of South East Asia and China where grass pea production occurs the crops are either grown under ‘utera’ conditions and utilize remnant water or are sown on rain-fed areas where they must exist on minimum moisture until harvest. The crops normally are considered to require low or zero inputs and therefore not only utilize remnant water but also must utilize remnant soil nutrients. As many of these soils are deficient in zinc this aspect requires further study to determine the effects on the crop.

In India, the seed of *lathyrus* (also called “teora” or khesari dal)are sometimes boiled whole, but are most often processed to split into *dal*. *Dal*, a soup-like dish, is the most common method of retailing the crop in the Indian subcontinent (Pandey, et al.,). The flour, made from grinding either the whole or split seed, is sold as *basan*. In many parts of Bangladesh, *roti* (unleavened bread) made out of grass pea flour is a staple for the landless labourers. More recently the *dal* or *basan* has been used to adulterate pigeonpea *dal* and chickpea *basan* as these crops demand a higher price on the market (Rahman, *et al*)

Lathyrus is an outstanding crop in terms of nutritional point of view. Seeds contain 18.2-34.6% protein, 0.6% fat, 58.2% carbohydrate (about 35% starch) (Duke, 1981; Williams *et al*., 1994).It also conatains sucrose 1.5%, phytin 3.6% ,pentosans 6.8% and albumin 6.69%((Duke,1981). But People and livestock consuming grasspea as the principal diet for months develop a paralytic disease known as "Lathyrism". Livestock that consume grasspea seeds (30-50% of the diet) for 3-6 months develop Neurolathyrism (Williams *et al*., 1994; Smartt *et al*., 1994). Amino acid derivatives from the seeds of other species of the genus *Lathyrus* and of some species of the genus *Vicia* produced similar effects on experimental animals (Duke, 1981; Williams *et al*., 1994). Beta-Noxalyl- L-alpha-beta-diamino-propionic acid (ODAP) also referred as Beta-N-oxalylamino-L-alanine (BOAA) occurs in grasspea and is a neurotoxin which causes paralysis (Smartt *et al*., 1994; Williams *et al*., 1994). ODAP concentrations vary widely (from 0.2 to greater than 1.01 mg per gram of seed) among a total of 1262 accessions collected from India and Ethiopia, and also the ODAP distribution in embryo was the greatest (400) followed by cotyledon (126 mg per gram). To decrease seed ODAP content, different approaches are going on.Some low varities were produced to with the help of conventional breeding approaches.But unfortunately these varieties had yield and contained unwanted compounds like α,β-diaminobutyric acid and γ-cyanoalaline,which are lathyrogenic in nature.(Bell,E *et al*.,1966.,Ressler,C.,1962). These varities had nrgetive correlations with the seed ODAP content and crude protein (Mishra,B.K.,1981). So till date no varities are developes which are devoid of neurotoxin.Development of lathyrisim free crop,is therefore, an important explotion of the potential of the crop.To achieve this goal different biotechnological approaches should be followed like gene knock down, gene knock out or genome editing.To approach these methods we should have the gene sequences of enzymes which are related in ODAP biosynthesis pathway.But unfortunately no genes were identified yet. To identify these genes,transcriptome data of *Lathyrus sativus* was taken from thesis Dr. Shubha Banerjee(Asst. Prof. IGKV,Raipur)

So transcritome data of two genotypes of *Lathyrus satives* having high and low ODAP content e.g RLK-1950 and Mahateora respectively were analysed and following objectives were formulated.

1. In silico analysis of transcriptome sequence for protein modelling and pathway mapping.
2. Biochemical estimation of seed ODAP content in diverse Lathyrus genotypes
3. Semi quantitative RT –PCR based validation of selected putative candidate genes in high and low ODAP containing genotypes

CHAPTER – II

REVIEW OF LITERATURE

This literature found in the heading of **“Analysis of De novo transcriptome sequence for identification of putative candidate genes related to β ODAP biosynthesis pathway in *Lathyrus sativus*”** is divided into following catagories.

2.1The cultivated area and consumption patterm

2.2 Why lathyrus sativus should be cultivated ?

2.2 Problem in Lathyrus consumption.

2.3 Role of β -ODAP in Plants

2.4 Lathyrus sativus breeding

2.5 ODAP biosynthesis.

2.6 Relation of β-ODAP with sulphur metabolism

2.1The cultivated area and consumption patterm

It is a most valuable crop in India.From stastical analysis it ranks 3rd among all cultivated rabi season crop.Produces .43MT from .52mha.(Kaul et al.,1986).Production ocuures in major states like MP,West Bengal,Maharashtra and Chhattisgarh(Khandare et al.,2014).

India has created name by holding a postion of rank 1st in terms of production and holds largest producer with 384,800t.Bangladesh and Ethipioa ranks 2nd and 3rd with a production of 232,500t, 202,126 respectivily.Country like Australia, and Canada has expontially increased its production as subsidiary crop between cereals due to its high nitrogen fixing ability capability.

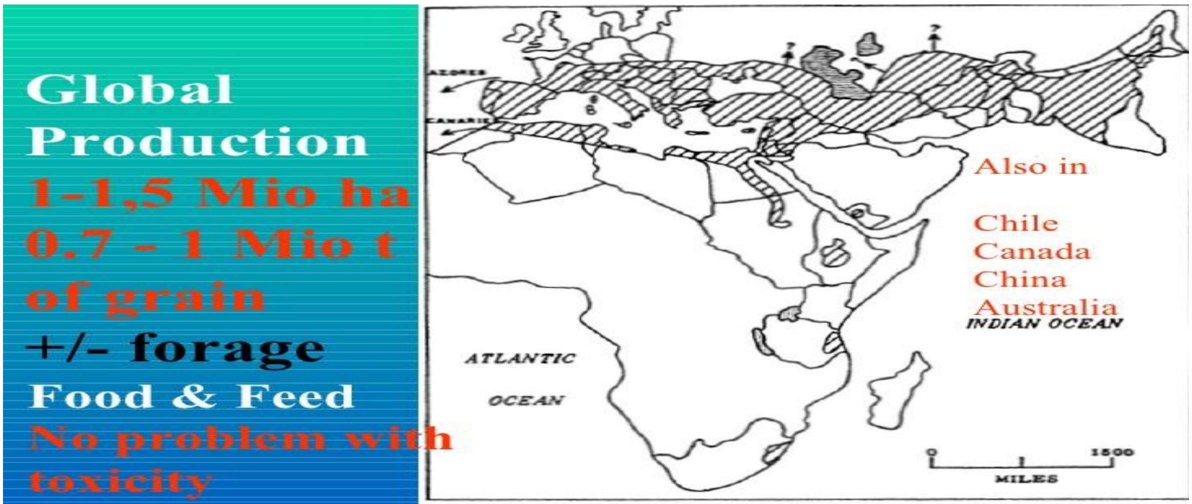


Figure 1. Maximum distribution of the grass pea as a food and fodder crop. (Adapted from V. S. Muratova)

If we talk about Chhattisgarh,then 60% of farmer in Raipur,bilaspur and Baster region adopt Grass pea cultivation as a insurance crop.They include this as a substancial crop in their cropping system.Bt the total food intake percentage is only 3%,in comparision to other pulse crop.Most of the farmer prefer to cultivate Chick pea(35%) over pigeon pea (25.3%), blackgram (17.5%), and grass pea (11.2%). About 60% residents in Gondia district in Maharashtra take grass pea as a part of their meal; but, the amount of grass pea consumed per capita was reported to be less than 25 g (Khandare *et al*., 2014).

2.2 Why lathyrus sativus should be cultivated ?

It has some sort of advantageous biological and agronomicall characters than other pulse crop.Most importantly it is very tolerant to drought and water-logging. It can be grown in soils of semiarid region which is poor in nutrient and organic matter.This crop shows considerable resistance to insects and pests. Nitrogen fixation character is very high-lighted in this crop.;It has outstanding production capacity and protein content.. (Kaul et al., 1986; Spencer, 1989; Campbell et al., 1994; Croft et al.)

1999).From the above statements we can get a clear idea that this legume can thus provide an good economic yield under hostile environmental conditions lie unequal rainfall and imbalance heat and gives high potential of production in the region of scarce raifall. Indeed, this has made it a popular crop in subsistence farming in certain developing countries that have extreme weather conditions (Rao et al., 1964; Ludolph et al; Tekele-Haimanot et al., 1990; Praveen et al., 1994).

If we look into the botanical structure of the plant,its is very hardy and has penetrating root system.This bulky root system helps in growth on a wide range of soil types, including very poor soil and heavy clays. As a legume, it improves to soil quality through the action of nitrogen-fixing symbiotic bacteria associated with the root system. Lu et al., 1990 demonstrate that L. sativus survives in the the extremely low rainfall (200–400 mm p.a.) in areas of northwestern China, it can also Gives an acceptable yield unde low humidityr conditions. (10% moisture) . In short, L. sativus is a hardy plant suited to dry climates, producing good seed crops on poor soils in some of the world_s most difficult farming areas. Itcan be used s good model and ideal plant for investigating the mechanism of drought resistance and can be identified genes associated with drought resistance.

During the famine it provides the required amount of protein and carbohydrate. (Spencer, 1989).

The nutritive amount of lathyrus is high with cpmparision to other crops like C. arietinum (18.0%), field pea (25.7%) and faba bean (26.9%), but lower than lupin (35.1%)

2.2 Problem in Lathyrus consumption

Though Lathyrus is an marvellous candidate crop for consumption and fodder purpose,but it has some sort of harmful effect on human consumption.It has an non standard,non protein neurotoxic amino acid,called as β-ODAP in its seed and other vegetative as well as reproductive plant parts. (Murti et al., 1964; Jiao et al., 2006.) It has been observed that prolonged consumption of lathyrus seed can cause neurodegenerative disease in human beings as well as animals.Some times over consumption may cause the same. (Spencer et al., 1986; Yan et al., 2006). β-ODAP is the main reason of ban in production and consumption. (Dixit et al., 2016).

Cause of neurolathyrisim is β- ODAP)(F. Wang, X. Chen et al.,2000 C. Ressler.,1962).Charectrization of this disease is spastic paraplegia.ODAP specifically act as promoting signal og glutamate receptor (S. M. Ross et al.,1989) and inhibit the activity of tyrosine aminotransferase.(S. Pearson et al.,1981 K. S. Vardhan et al.,1997). F. Lambein found that ODAP also inhibit the chelating metal activity like Zn.

2.3 Role of β -ODAP in Plants

β-ODAP has been found in 13 crotolaria sp.,17Acasia species and 21 Lathyrus species.

(Quereshi et al.,1977).Except this several non-leguminous plant also produce ODAP.(Long, Y.C.,1996 , Pan, M.D.,1997 Kuo, Y.H.,1998).The function of odap as Zn ion carrier (Lambein, F.,1994),scavenger of OH ion (Zhou, G.K.,2001) and as helps the chlorophyll from photo-oxidation(Zhang,J.,2003) at high light intensity during day time. As we know earlier that odap gives power to plant to cope up with drought and insect resistance.

2.4 Lathyrus sativus breeding

From above we knew about the importance and scope of lathyrus.So the objective of every breeder is to develop a lathyrus variety with low odap content.By evaluating the seed odap content in germplasm it is came to know that, in a vast range of odap from 0.02% to 2.59% of seed weight is present in germplasm.(Kumar, S.,2011)

Kumar,S et al. discovered that know that low odap plant very prone to disease and pest and also it is very vulenerable to drought.Then he conclude that ODAP seems to be play a great role in plant defence mechanisim.After that several line had been released through conventional hybridization, somaclonal variation and through mutant breeding in India as well as in Ethiopia.Through Multi location trial it is confirmed that these lines also give accountable yield.To justify this, one line produced from ICARDA (International Center for Agricultural Research in the Dry Areas) gives 1.67ton/ha and its odap content is low which is .08% of the seed weight.Kumar et al., also found that ODAP content gives a positive correlation with the adverse field condition like water stress,salinity and drought.

After crossing of two different diverse genotypes the Tripathy et al; observed that the gene resoponsible for production of neurotoxin is quantitative in nature. But problem is that this gene has not been discovered yet.

A study was undertaken for calculating body mass index(BMI) with sample size N=1070.It has observed that among all people,65% people have normal BMI,in which only 37% people consumesd pulses throughout the year.4.39% people have pulse in there diet only for a quarter of the year and 15% people have pulse 3 quarter per year. Remaining 35% are under nourished or over nourished.Above analysis and studied clearly indicates how pulse is very much important in our day to day life.

.,).Nepalis are used to eat lathyrus as dal by crushing the seeds with a help of a stone. The grains are also ground and made into flour for use in a pancake-like preparation of *badi* or *pakoda* (Yadov, pers. comm.). Grass pea flour increasingly is being used to adulterate the higher-priced legume flours such as chickpea and mungbean. In Ethiopia, particularly the northern regions, and in Eritrea tef, wheat, barley, maize and sorghum, either singly or in combination, are used to produce a fermented, sour pancake-like unleavened bread called *enjera*. *Lathyrus* grain is ground into *shiro* and is used in the preparation of *wott*, a sauce that is eaten together with the *enjera.* For snacks, cereals, legumes or their mixture are most often
consumed roasted or boiled. Boiled grass pea (*nifro*) is consumed in most areas. *Kitta*, an unleavened bread made from grass pea, is consumed to a more limited extent, mainly at times of acute food shortages (Tekele-Haimanot *et al.* 1993).

The first effort to systematically collect the grass pea germplasm in India was in 1967 **(Mehra *et al.* 1995**). Approximately 600 accessions were collected and analyzed for ODAP content. In 1969 subsequent collections were made from Bihar, Eastern U.P., West Bengal, Gujarat and Haryana. In 1975 IARI scientists collected from the tribal areas of Bihar. In 1976 some indigenous accessions were sent to the Lathyrus Improvement Program at Raipur M.P. from Jawahar Lal Nehru Krishi Vishwa Vidyalaya, Jabalpur. Later on the Directorate of Pulses Research, Kanpur also supplied some indigenous materials to this programme. Exotic accessions belonging to Italy, Canada, Germany, Bangladesh, France and Iowa were also received from NBPGR (National Bureau of Plant Genetic Resources), New Delhi. During 1989-90 and 1990-91, 1187 landraces of grass pea were collected from the growing regions of Madhya Pradesh. At the present time 2659 accessions are being maintained at Raipur. A set of landraces has been stored in the National Genebank at NBPGR, New Delhi. These have been coded as IC 142554 to IC 143565. Short-term storage facilities have been developed at the Indira Gandhi Agricultural University, Raipur

2.5 ODAP biosynthesis

ODAP biosynthesis biochemical pathway is not fully cleared yet. (Ikegami, F.,1991). It is hypothesized that O-acetylserine is the main precursor of the β-odap.O-acetylserine is formed from Serine on the presence of enzyme Serine acetyltransferase.Then O-acetylserine on the presence of β-cyanoalaline synthase(CAS) converted into β-(isoxazolin-5-on-2-yl)alanine (BIA). Then this BIA finally converted into short lived 2,3,-L-diaminopropanoic acid and this derivative of propionic acid with help oxlyl co-enzyme forms β-odap.(Malathi, K.,1967 Padmanaban, G.,1967)

Lambein, F et al.,2009 also gave their opinion that high concentration of β-(isoxazolin-5-on-2-yl)alanine found in young seedlings of lentil and grass pea but considerable amount of 2,3-L-diaminopropanoic acid has npt been found. O-acetyl serine also act as important substrate for cysteine. Cysteine synthase helps in production of cysteine with presencr of hydrogen sulphide.(Ikegami, F.,1991).Then cysteine is again used as substrate for CAS with hydrogen cyanide to produce β-cyanoalanine and this cyanoalanien gives isoxazolin-5-one with a intermediate amino acid asparagines.This conversion also detoxify the cyanide because HCN is used a substrate here. (Machingura, M.,2016)

Methionine and cysteine metabolism is related with neurotoxicity of ODAP. Getahun et al. suggested that we we took lathyrus vaegetable and sulphur containing vegetables simultaneously in our daily meal then definitely it will decrease the harmfull effect of lathyrus.

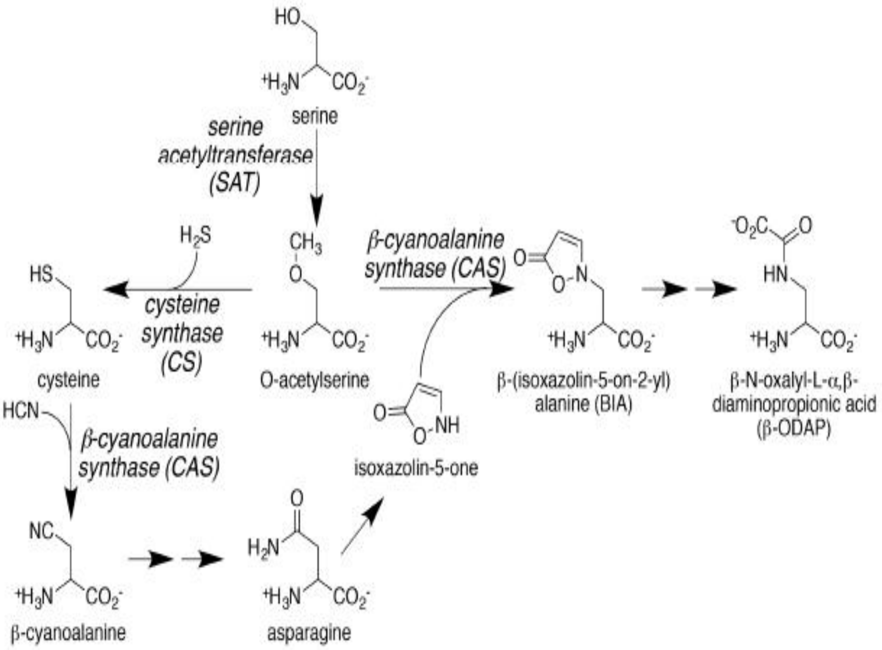


Fig 1-Biosynthetic pathway of β-ODAP

2.6 Relation of β-ODAP with sulphur metabolism

From the figure we have an clear idea that ODAP synthesis is related with sulphur amino acid cysteine.Precursor of β-ODAP is heterocyclic BIA (β-isoxazolin-5-on-2-yl)alaline.It is formed by CAS(β-cyanoalaline synthase).Here is an interestin incident that both CAS and CS act on the same substrate that is o-acetyl serine.Both the enzyme have pyridoxal dependent activity.Both enzyme comes under same family β-substituted alaline synthase(BSAS family) (Watanabe, M.,2008 Yi.,2012).

Ikegami et al. isolated the soforms of Cysteine synthase having 35kDa an 39kDa.It is an dimeric protein and both show the similar activity on o-acetyl serine.

Malathi has given the oxalylating steps of ODAP including two reactions catalyzed by two different enzymes

$$\text{Oxalate+ATP+Coenzyme A} \xleftrightarrow[\text{Oxlyl-CoA synthetase}]{Mg^{2+}} \text{Oxlyl-CoA+AMP+PPi}$$

ODAP
ssynthease

Oxalyl - CoA + L -1, 2 _ diaminopronic acid (DAPA) ⟶ ODAP +CoA

Recent studies have proven that biosynthesis of ODAP requires oxalic acid,which is a important factor for resistance in various plant.Physiologically and biochemically oxalic acid has similarity with ABA,synthesized by various drought induced plant.(D.W. Zhang et al.,2005). N. Bagni et al.,told about polyamines that these are ubiquitous nitrogenous compound and these are present in all plant cells upto milimolar conc. And plays a vital role in growth and development of the plant.These PA(polyamines) are indicator of environmental changes.(H.P. Liu et al).

Upto now the role of ABA and PA with relation to ODAP production was not known. But You-Cai Xiong et al., conducted an experiment and found the regulatory effect of ABA and free PAs metabolism in accumulation of ODAP in grass pea. In their study, they assessed the roles of ABA and PAs in regulating ODAP metabolism through application of exogenous α-difluoromethylarginine (DFMA), a polyamine biosynthesis inhibitor and exogenous Put. The study aims to find out whether ABA, as a result of water stress, regulates the biosynthesis of ODAP and polyamine in grass pea and what role polyamines play in the adaptation of crops to drought stress.They used 10% polyethanol glycol for this purpose.They also found that exogeneous application of ABA also increases the content of endogenous ABA in the leaf tissue.

Several marker analysis has been done in lathyrus to know the diversity analysis of the crop.((Noli *et al*., 1997; Paul *et al*., 1997; Yee *et al*., 1999).Both (Tanksley, 1983) and (Brown, 1978) did morphological and enzymatical analysis. A variety of DNA marker and isoenzymes are used.

R.L. Ambade, S.K. Verma et al, did an experiment in the field of IGAU,RAIPUR. They used 24 IsSR marker and among them 12 ISSR marker showed polymorphism. Aftrr that these marker are used for diversity array analysis.They used 48 geneotypes and among them RLK-466a nd RLK-637 are most similar with co-efficient .92. two variety BIO-L-203 and RLK-120 are less similar with co-efficient of .63.

Rabiul M. Haque et al. Conducted an experiment in by culturing the callus through tisue culture technique to know the relation of production of ODAP with environmental factors. He found the more the harsh environment more is the ODAP production. So he gave an conclusion that ODAP production is directly proportional to hostile environment.

Jaspreet kour *et al*., (2016) had conducted an experiment in IGKV ,RAIPUR on eighteen advanced breeding lines with their five parents,and they evaluated for their yield and yield attributing characters,protein and β-ODAP (β-*N*-oxalyl-L-α, β-diaminopropionic acid) at different growth stages. The correlations between traits were investigated.Positive significant correlations were found among seed yield,number of seeds/plant and number of pods/plant. β-ODAP content at different stage shows no and negative correlation with seed yield and protein content.Seed protein content shows negative or no correlation with seed yield and β-ODAP at different growth stage. Therefore, it can be said that development of higher yielding grasspea lines with low β-ODAP and high protein content is possible.

Fumio ikgemi *et al.,* (1999) gave an idea about enzymatic foramation of β-*N*-oxalyl-L-α, β-diaminopropionic acid, a neurotoxin present in *Lathyrus sativus,* by the break down of BIA(β-isoxzolin-5-on-2-yl)-L-alaline..

ODAP estimation by spectrometer

Rao *et al*., (1978) used a colouring reagent *o*-phthalaldehyde (OPT) to determine the presence of diamino propionic acid in the rat tissue. OPT givs a slight colour with that DAP. More the amount of DAP more is the depth of that colour.The reaction is used to determine specifically the DAP content in tissues of the rat injected with this amino acid. The method involved the alkaline hydrolysis of ODAP to yield α, β- diaminopropanoic acid, which in turn is complexed with ophthalaldehyde in the presence of ethanethiol to form a colored product that is quantified at 420 nm.Above procedure is also used to determine the ODAP content in Lathyrus sativus.. The procedure is rapid and would be of use in studies with DAP and the *L. Sativus* neurotoxin. However, the method does not differentiate between the two isomers of ODAP (α- ODAP and β-ODAP). α-ODAP is a nontoxic or much less toxic isomer, and easily formed from β-ODAP during heating.

Briggs *et al.*, (1983) demonstarated the extraction and Spectrophotometric assay procedure for analysis of ODAP in small samples (0.500mg) of powedered seed. This experiment uses o-phthalaldehyde (OPT) as a colouring reagent, followed by hydrolysis with 3N KOH solution , and this is suitable suitable for the derermination of ODAP content of 3-4 seeds of most *Lathyrus* species.Now a days this method is commonly used un breeding programme.

Husain *et al*., (1994) made various modifications of the in the Rao et al., described protocols.But when these protocol were followed in same sample with different time then a considerable variation was observed which signifies about the low reproducibility about above prescribed method. But the scientist of Agriculture Canada Research Station, Mordsen, Manitoba, Canada (Campbell method) developed the new method.In this method Cambell et al.,used 0.5M tetraborate buffer of pH9.9.This buffer helped in maintain the constant result gave high level of reproducibility.

Urga *et al.,* (2005) demonstrated the ODAP variation of by taking 100gm of seeds. He also analysed the variation of crude proteins in different Lathyrus seeds.

Bell and Odonovan (1966) worked in *Lathyrus latifolicus*.He found two oxlyl derivative of gamma butyric acid.Trnasfer of oxlyl group between alpha and beta group is demonstrated by the scientist.

Tarade *et al.,* (2006) A nonprotein neurotoxic amino acid, b-N-oxalyl-L-2,3-diaminopropionic acid (ODAP), found in *Lathyrus* sativus (grass pea or chickling vetch) seeds is known to be relatively heat stable. The present study aims at development of a kinetic model for degradation of ODAP in *Lathyrus sativus* subjected to a defined set of processing conditions. This study was carried out at pH 4.0 and 9.2. Isothermal condition experiments were carried out over a temperature range of 60–120°C. For non iso thermal conditions, three different cooking methods viz., open pan, pressure cooking and cooking in recently developed and patented fuel efficient „EcoCooker" were used. The degradation of ODAP was adequately modeled by Arrhenius type of equation. A mathematical model based on the time temperature data of the nonisothermal heat process and isothermal kinetic rate parameters has been developed to predict the degradation of ODAP in any nonisothermal heating process of known time temperature profiles.

Isildak *et al*., (2014) reported that a novel potentiometric sensor based on ionophore (Cd(NH2CH2CH2OCH2CH2OCH2CH2NH2)Ag3(CN)5) for the determination of β-N-oxalyl-l-α, β-di amino propionic acid (ODAP) is developed. The ODAP-selective membrane sensor demonstrates high sensitivity and short response time. The detection limit of the ODAP-selective membrane sensor is about 2 × 10−6 mol L−1 and the response time is shorter than 6s. The linear dynamic range of the ODAP-selective membrane sensor is between ODAP concentrations of 1.0 × 10−2 and 1 × 10−6 mol L−1.The ODAPselective membrane sensor exhibits good operational stability for at least one week in dry conditions at 4–6∘C. It has a reproducible and stable response 32 during continuous work for at least 10 h with a relative standard deviation of 0.28% (n = 18).

Woldesemayat *et al*., sequenced the leaf transcriptome using Illumina MiSeq platform and generated 2,590,652 paired-end reads that were assembled de novo into 45,450 high-quality transcripts. Annotation of these transcripts revealed best hits for homology to discover more than 17,800 functional genes and conserved domains. A total of 93 KEGG pathways and associated genes encoded by more than 90% of the coding transcripts are responsible for the biosynthesis of these life-saving metabolites. They validated and enriched the genes by GO annotation and linked this to enzymepowered pathways through interactive network map. Caffeine metabolism, flavonoid, phenylpropanoid and terpenoids biosynthesis and xenobiotics degradation were typical in tea quality and drug therapy. The relatedness of more than 80 gene families encoding key enzymes was shown using unrooted phylogenetic tree. The sequence of the *Lathyrus* transcriptome sequencing was recently determined by SAGE for investigation of *Ascochyta lathyri*. However, a whole genome sequence has not yet been reported and the only published genomewide study of *Lathyrus* is a transcriptome analysis of a genotype BGE015746 resistant to *A. lathyri* combined tissues of leaves (Almida *et al*., 2015).

Breyne *et al*., (2002) identified novel genes related in cell cycle mechanisi, and cell division process.He did this experiment by analysing the whole transcriptome data. He found some modulator genes related to cell cycle division process.

Xu *et al*., (2013) worked in *Setaria viridis*. He also worked in the ares of transcriptome analysis. He collected different plant material of *Setaria viridis* of different developmental stage of seed germination,reproduction and vegetative growth.With the help of next generation sequencing he analysed minimum 71 lakh transcriptome of 100bp reads.This analysis helped many researchers for development of different SSR markers.

Zhang *et al*., (2016) compared the transcriptomes of Kentucky bluegrass under cold treatment (-5°C) and a control treatment (at 20°C) by RNA-seq and *de novo* assembly. Totally 75,934 unigenes were generated, among which 53,762 were successfully annotated in public databases. Upon comparing the transcriptomes of the control and cold-treated plants, 3,896 unigenes were identified as differentially expressed. Among these genes, 2,410 were down-regulated and 1,486 were up-regulated in the cold-treated plants. A few previously reported cold-induced proteins, antioxidant enzymes, and osmoregulation proteins were identified, and their expression levels were estimated. Moreover, ten differentially expressed genes were selected for qRT-PCR verification. Their expression patterns were consistent with the results of the RNA-seq. Additionally, the transcription factor families, i.e., ethylene response factors, heat stress transcription factors, NAC proteins, WRKY domain containing proteins, and auxin response factors, were identified as differentially expressed genes.

Padder *et al*., (2011) conducted a comprehensive transcriptomeanalysis using Illumina sequencing of two near isogenic lines (NILs) differingfor the presence of the *Co-1* gene on chromosome Pv01 during a time coursefollowing infection with race 73 of *C. lindemuthianum*. From this, we identified 3,250 significantly differentially expressed genes (DEGs) withinand between the NILs over the time course of infection. During the biotrophicphase the majority of DEGs were up regulated in the susceptible NIL, whereas more DEGs were up-regulated in the resistant NIL during the necrotrophic phase. Various defense related genes, such as those encoding PR proteins, peroxidases, lipoxygenases were up regulated in the resistant NIL. Conversely, genes encoding sugar transporters were up-regulated in the susceptible NIL during the later stages of infection. Additionally, numerous transcription factors (TFs) and candidate genes within the vicinity of the *Co-1* locus were differentially expressed, suggesting a global reprogramming of gene expression in and around the *Co-1* locus.

Nakasugi *et al*., (2013) prepared RNA-seq libraries from 9 different tissues were deep sequenced and assembled, de novo, into a representation of the transcriptome. The assembly, of16GB of sequence, yielded 237,340 contigs, clustering into 119,014 transcripts (unigenes). Between 80 and 85% of reads from all tissues could be mapped back to the full transcriptome. Approximately 63% of the unigenes exhibited a match to the Solgenomics tomato predicted proteins database. Approximately 94% of the Solgenomics *N. benthamiana* unigene set (16,024 sequences) matched our unigene set (119,014 sequences). Using homology searches we identified 31 homologues that are involved in RNAi-associated pathways in *Arabidopsis thaliana*, and show that they possess the domains characteristic of these proteins. Of these genes, the RNA dependent RNA polymerase gene, Rdr1, is transcribed but has a 72 nt insertion in exon1 that would cause premature termination of translation. Dicer-like 3 (DCL3) appears to lack both the DEAD helicase motif and second dsRNA binding motif, and DCL2 and AGO4b have unexpectedly high levels of transcription.

Wang *et al*., (2016) studied that a transcriptome, including 76,014 unigenes, was assembled from dwarf polish wheat (DPW) roots, stems, and leaves using the software of Trinity. Among these unigenes, 61,748 (81.23%) unigenes were functionally annotated in public databases and classified intodifferentially functional types. Aligning this transcriptome against draft wheat genome released by the International Wheat Genome Sequencing Consortium (IWGSC), 57,331 (75.42%) unigenes, including 26,122 AB-specific and 2,622 D-specific unigenes, were mapped on A, B, and/or D genomes. Compared with the transcriptome of *T. turgidum*, 56,343 unigenes were matched with 103,327 unigenes of *T. turgidum*. Compared with the genomes of rice and barley, 14,404 and 7,007 unigenes were matched with 14,608 genes of barley and 7,708 genes of rice, respectively. On the other hand, 2,148, 1,611, and 2,707 unigenes were expressed specifically in roots, stems, and leaves, respectively. Finally, 5,531 SSR sequences were observed from4,531 unigenes, and 518 primer pairs were designed.

Sun *et al*., (2016) studied that identification of potential genes and regulatory networks related to the floral development in *C. faberi* by using transcriptome sequencing, de novo assembly and computational analyses. The vegetative and flower buds of *C. faberi* were sampled for such comparisons. The RNA-seq yielded about 189,300 contigs that were assembled into 172,959 unigenes. Furthermore, a total of 13,484 differentially expressed unigenes (DEGs) were identified between the vegetative and flower buds. There were 7683 down-regulated and 5801 up-regulated DEGs in the flower buds compared to those in the vegetative buds, among which 3430 and 6556 DEGs were specifically enriched in the flower or vegetative buds, respectively. A total of 173 DEGs orthologous to known genes associated with the floral organ development, floral symmetry and flowering time were identified, including 12 TCP transcription factors, 34 MADS-box genes and 28 flowering time related genes.

CHAPTER-III

MATERIALS AND METHODS

The present thesis entitled **"Analysis of De novo transcriptome sequence for identification of putative candidate genes related to β ODAP biosynthesis pathway in *Lathyrus sativus*"** was conducted during the sesson of 2018-19 at the department of Plant Molecular Biology and Biotechnology, Indira Gandhi Krishi Vishwavidyalaya, Raipur, India. This chapter contains the experimental materials used in the course of research and methods adopted for the same. The detail of the materials used, methods adopted were presented as under the individual headings and sub headings.

3.1. Experimental site

The experiment was conducted in the field and different laboratories of Department of Plant Molecular Biology & Biotechnology, College of Agriculture, Indira Gandhi Krishi Vishwavidyalaya, Raipur.

3.2 Materials used

Thirty *Lathyrus* genotypes were collected from Department of Genetics and Plant Breeding at IGKU, Raipur, Chhattisgarh for biochemical analysis.Out of them 11 genotypes were selected on the basis of their ODAP content for identification of candidate genes related to odap biosynthesis.

Sl no	**Variety**	**Sl no**	**Variety**
1	AKL-19	16	IC-28112
2	SEL-481	17	SEL-531
3	PRATEEK	18	BIOL-203
4	JRL-41	19	RLK-2059

5	PUSA-24	20	BIOL-202
6	SEL-519	21	SEL-516
7	SEL-522	22	RATAN
8	SEL-527	23	SEL-508
9	SEL-439	24	NO-2203
10	SEL-520	25	RLS-2
11	RLK-498	26	MAHATEORA
12	SEL-554	27	IL-143099
13	SEL-553	28	NO-2208
14	BIOR231	29	SEL-587
15	SEL-504	30	SEL-563

(Table 3.1 List of thirty diverse *Lathyrus* genotypes)

3.3 METHODOLOGY

3.3.1 Biochemical characterization of diverse *Lathyrus* genotypes

ODAP ESTIMATION:

According to the protocol of Rao *et al.*, 1964. β-N-Oxalyl-α, β diaminopropionic acid (ODAP) in dry seeds was estimated. An hydrolysing agent OPT((O-phthaladehyde) was used for this purpose.

The protocol as mentioned below :

- 500mg seed were grinded and powered sample was kept in 10ml 60% ethanol for overnight.

- Vortex well and after that 2ml supernatant was centrifuged at 4500rpm for 15 minutes at room temperature.
- Supernatant was hydrolysed with 4ml of 3N KOH solution and kept for 90^{0}C for 30 min.

- Blank solution containing 2ml 60 percent ethanol and 4 ml KOH (3 N) in place of extract was processed in parallel to hydrolyzed extract.
- After cooling 250µl supernatant of that hydrolyded extract was taken and added in 750µl dd water and 2ml opt reagent.
- The sample was incubated at room tempt for 2 hrs.
- A bright yellow colour was developed and absorbance was measured at 425nm in Visible spectrophotometer.

  The absorbance of blank solution was also taken. Final absorbance estimated by substrcting the blank absorbance from sample absorbance.

3.3.2 Preparation of OPT reagent

For ODAP estimation always freshly prepared OPT reagent is used. It is a hydrolysing agent. 100mg of OPT in 1 ml of 95% ethanol,.2ml β-mercaptoethanol is mixed properly and volume is make up is done to 100ml by adding potassium tetraborate buffer. pH of that buffer should be 9.9 and 0.5M concentration.

3. 3. 3 Standard Curve

For standard curve LDAP chemical was taken. This LDAP is a analogoys to ODAP. 1% LDAP was dissolved in 4ml of 60% ethanol and 2ml 3N KOH kept at 95◦C in waterbath.

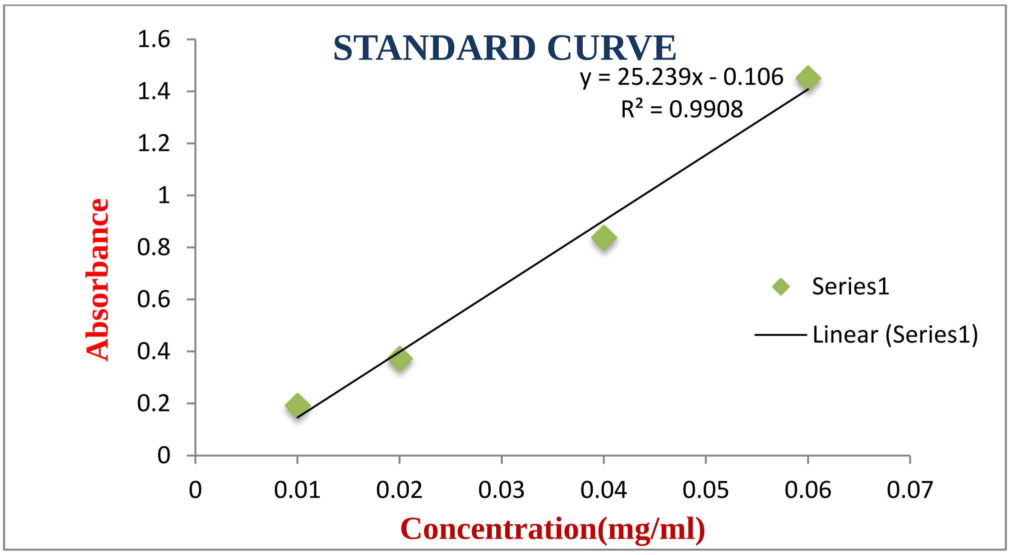


Fig-3.1 Standard Curve

Among 30 genotypes, 11 genotypes are selected according to their ODAP content

Sl no	Variety	Sl no	Variety
1	N0-2208	7	PUSA-24
2	SEL-519	8	BIO L-202
3	IL-143099	9	RLK-498
4	SEL-527	10	SEL-522
5	AKL-19	11	RLK-1950
6	MAHATEORA		

(Table3.2 Eleven selected genotypes talen for RNA extraction)

Seeds of these genotypes was germinated in petriplate and grown in field condition for 10 days.And RNA extraction was done 10 days old plant.

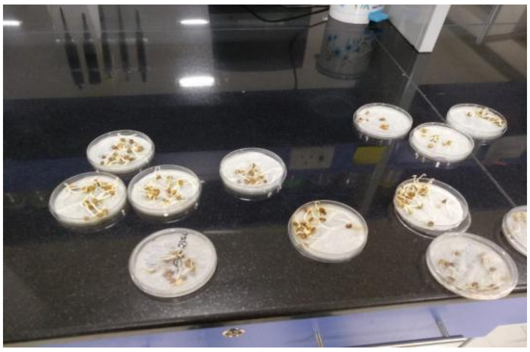


Fig-3.2 Germinated seed in petriplate Fig3.3 Transplanting of germinated seed

3.4 Semi quantitative expression analysis of candidate genes responsible for ODAP biosynthesis among different *Lathyrus sativus* genotypes.

3.4.1 Collection of leaf tissue and isolation of RNA:

Leaf samples of 11 *Lathyrus* genotypes were collected from 10 days old plant because ODAP Synthesis is higger in 10 days old plant. Healthy and tender leaf were taken for this. Damaged and diseased leaves were discarded.These leaves were crushed in liquid N_2.

3.4.2 RNA extraction:

RNA extraction was done using TRIZOL Reagent.

MATERIALS REQUIRED:

- Trizol
- Chloroform:Isoamyl Alcohol mixture(24:1,v:v)
- Isopropanol
- DEPC H_2O
- 75% ethanol (prepared with DEPC water)

METHODS

Preparation of DEPC water

1. 0.1% DEPC water is prepared by dissolving 0.1g DEPC(diethyl pyro carbonate) to 1L double distilled water. Mix well by stirring for 2 hours.(**DEPC IS VERY HIGHLY TOXIC COMPOUND).** So everything should be done in proper care and protection.
2. DEPC water is double autoclaved. DEPC get hydrolysed and degraded during the autoclaved process. After that it can be stored for several days at normal room tempt.
3. Incubate all the mortor and pestles ,spatula overnight in DEPC treated water.

Preparation of Utensils For RNA extraction

1.Soak mortar pestles and other glass wares used for RNA extraction in 30% H_2O_2 for 15 minutes.After that rinse with DEPC treated water and keep for over night and dry in hot air oven.

2. Always use RNase free microcentrifuge 1.5 ml tube and RNase free filter tips every time.

RNA EXTRACTION PROCEDURE

1. 10 days old leaves were collected and crushed to fine poweder with liquid $N_{2,}$ with the help of mortar and pestle. The Powered is then transferred to 2 ml tube.
2. 500µl of Trizol reagent is added and vortex well.Incubated for 10 minures.
3. After that add 200µl of Chloroform:Isoamyl alcohol mixture(24:1) and vortex well.Tubes are again incubated for 10 mins. at room tempt then spin it at 4^0c fpr 15 minutes at maximum speed at bench top microcentrifuge.
4. After centrifugation pipette out the top aqueous phase into clean 1.5 ml microcentrifuge tube.
5. Add 500µl Isopropanoland mixed gently. Leave that tube intact for 10 minutes.
6. Again centrifuge at maximum speed in 4^0c for 15 minutes. See the pellate.
7. Decant isopropanol.
8. Add 500µl 70% ethanol(DEPC treated) and resuspended the pellate by gentely tapping. Again centrifuge for 5 minutes and decant the ethanol.
9. Invert the tubes on paper towel for air dry until no droplets is seen in the tube.
10. Re-suspend pellatein 20µl of RNAse free DEPC treated water.

3.4.3 cDNA synthesis:

RNA isolated and quantified was used for cDNA synthesis using Bio-RAD iScriptTM cDNA Synthesis kit as per manufacturer's instructions. Reaction mix and temperature profile for cDNA synthesis are presented in Table 3.2 and 3.3. The kit consisted of following components and is suitable for 1pg to 1µg of RNA. Kit components include following items:

1. 5X iScript Reaction Mix
2. iScript Reverse Transcriptase
3. Nuclease free water

Procedure for cDNA synthesis:

1. Thaw the template RNA samples on ice. Calculate the amount/ volume of RNA required reaching 1 µg concentration and adjusting the volume of water accordingly. Dispense the required quantities of RNA and water in each tube (0.2 ml capacity).
2. Prepare the master mix on ice by mixing the mentioned amounts of 5x iScript reaction mix, iScript Reverse transcriptase, nuclease free water.
3. Gently Mix by pipetting.
4. Dispense 10 µl of master mix into each tube containing RNA samples and water. Mix by gentle tapping and spin down the contents.
5. Incubate the tubes in PCR machine on reverse transcription cycling program described in table 3.4.

Table 3.3: cDNA synthesis (reverse-transcription) reaction components.

Component	Volume reaction	Final concentration
5x iScript reaction mix	4 µl	1x
iScriptReverseTranscriptase	1 µl	-
Nuclease free water	Variable	-
RNA template	1-51 µl	1 µg
Total volume	20 µl	

Table 3.4: Reverse transcription cycling program

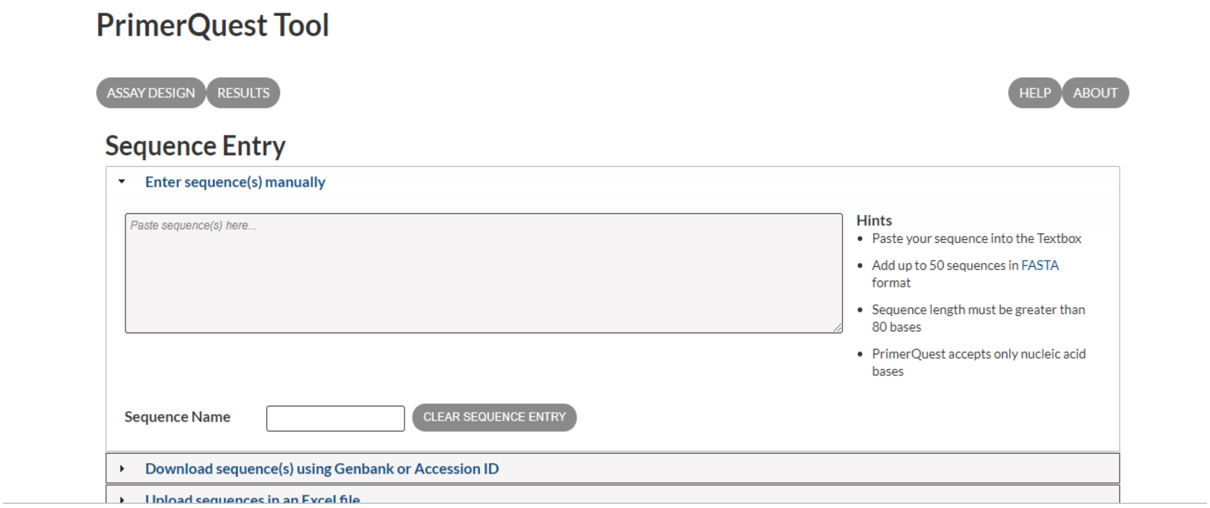
Step	Temperature	Time	No.of cycles
Priming	25˚C	5min	1
Reverse transcription	46˚C	20min	1
RT inactivation	95˚C	1min	1
Optional step	4˚C	Hold	

Quantification and dilution of cDNA:

Quality and quantity of cDNA were estimated by measuring the absorbance using Nanodrop spectrophotometer (ND1000). 1µl of DNA was placed over tip of Nanodrop to record absorbance at 260nm. The absorption ratio (A260/ A280) and (A260/ A280) was recorded for each sample to estimate quantity and purity of cDNA.

3.5 In silico designing of primers:

Candidate genes were screened on the basis of synthase family and transferase family which are related to ODAP biosynthesis from the transcriptome data. After screening all the sequences were subjected for Primer designing through primerquest IDT(Integrated DNA Technology)(https://eu.idtdna.com/PrimerQuest/Home/Index)



(Fig-3.4 Home page of primer Quest)

Precautions for Primer designing

Most imp factor for primer designing are Melting tempt, Annealing temperature (T_a) and total GC content.

- Ideal melting tempt is between 60-62^{0}C. Ideally, the melting temperatures of the 2 primers should not differ by more than 2°C in order for both primers to bind simultaneously and efficiently amplify the product.
- The annealing temperature chosen for PCR relies directly on length and composition of the primers. This temperature should be no more than 5°C below the T_m of your primers. One consequence of having T_a too low is that one or both primers will anneal to sequences other than the intended target because internal single-base mismatches or partial annealing may be tolerated. This can lead to nonspecific PCR amplification and will consequently reduce the yield of the desired product. Conversely, if T_a is too high, reaction efficiency may be reduced because the likelihood of primer annealing is reduced significantly. Optimal annealing temperatures will result in the highest product yield with the correct amplicon.
- GC content og the primer should be in between 35-65%. Ideal content is 50%,which permits complexity during maintain a unique sequence.Primer should not contain 4 or more consecutive G residues.
- Sequence length must be >80b

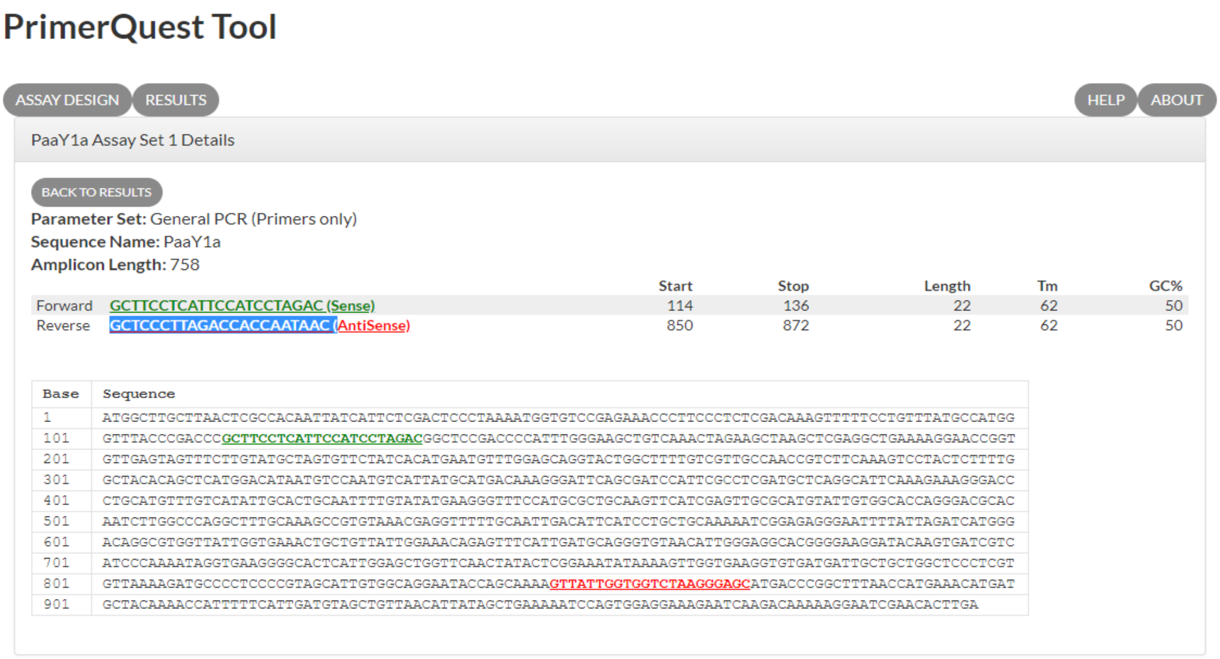


Fig-3.5 Designing of primer

LIST OF PRIMERS

SL NO	PRIMER NAME	FORWARD PRIMER	REVERSE PRIMER	T_m	Amplicon length
1	**PaaY1**	**TAACACCGCAATCCTCAACA**	**TTAAGCCATGAGGCGACTAAG**	**62**	**112**
2	**PLPD1**	**CACAACCATTCCTTCTCTACCA**	**CAGATGATGAAGAGAACACATGAAC**	**62**	**129**
3	**PLN1**	TCGGCTTCCAAAGCAAAGA	**GGACCTGCACTTAACCAGTATC**	**62**	**125**
4	**PLN2**	**GGAAGGAGATGGAACTGATACC**	**ATTGAAGGCGAACCAGACA**	**62**	**142**
5	**TRP1**	**CAAGAAGACGGGCATCTGAA**	**GAGTGAGTTATGGTCGTGTCTG**	**62**	**122**

6	**PLN3**	**GGTCTTTGTGGGAAG AGTAGAC**	**TCGCGAACCAT TCTCTGAATAA**	**62**	**146**
7	**UNK1**	**GCTAAGATTGCAAAC AACACTCTC**	**AGCGTCTTTGA TCAGGCATATT**	**62**	**149**
8	**UNK2**	**TGTAACACCGGATGG AGATACT**	**CTGTGACCTGC CAGGATTTATT**	**63**	**100**
9	**AMIF**	**CACGTAACAGGGATG GCTTAT**	**ACTCGTATCGT GCAGTCATTC**	**62**	**99**
10	**AMIFa**	**CTACAGCACTCTCTG ACAAACA**	**GTTGTTGGGA GGGCAATAGA**	**62**	**130**
11	**DegQ**	**CCAATGCTAACACCG CAATC**	**GGCCGGAGTT GAAGACTATT**	**62**	**97**

(Table 3.5 List of 11 Primers)

3.6 Semi quantitative RT- PCR based gene expression analysis:

Semi quantitative reverse transcriptase PCR was carried out to study the expression of ODAP biosynthesis related genes. The cDNA generated from the total RNA isolated from leaf tissues of 11 different lathyrus genotype having differential ODAP content were subjected to semi-quantitative expression profiling in 10 µl reaction using candidate gene primers designed from transcrioptome data of RLK-1950. The resultant PCR product was then resolved on 1.5 % Agarose gel at 70V. The presence of amplicons and their respective intensity were recorded under gel documentation system. The expression was analyzed by comparing the relative fluorescent intensities of cDNA amplicons under gel documentation system. *LsActin* (primer designed from *Lathyrus sativus*)was used as an internal control for normalization of RNA concentration. Details of PCR components and temperature profile are depicted below in table 3.6 and table 3.7.

Table 3.6. PCR components their quantity used for semi-quantitative PCR.

Components	Concentration	Quantity
cDNA	1000 ng/µl	1 µl
PCR Buffer	10X	1 µl
dNTP mix	10Mm	1 µl
Primer Forward	10 mM	0.5 µl
Primer Reverse	10 mM	0.5 µl
Taq Polymerase	5U/ µl	0.1 µl
$MgCl_2$	50mM	0.25 µl
Nanopure water	-	5.65 µl
Total		10 µl

Table 3.7 Temperature profiles used for semi quantitative RT-PCR.

Activity	Temperature	Duration	Cycles
Initial Denaturation	94 ˚C	5 min	1
Denaturation	94 ˚C	45 sec	
Annealing	55^0C	30 sec	35
Extension	72˚C	1 min	
Final Extension	72˚C	7 min	1
Store	4˚C	∞	1

3.7 Agarose Gel Electrophoresis

1.5 percent Agarose gels (horizontal) were used for vizualising larger size PCR amplified products. Gels were casted in Borad electrophoresis unit . Casting tray was thoroughly cleaned before making the gel solution.

(I) Pouring the Agarose gel

(a) To prepare the agrosose gel, 1.5g of agrose was weighted and put in 100ml of 1X TBE Buffer and heated for 2 minutes. After cooling down of agrose, 2.5µl of EtBr stock solution was added.

(b) Fixed the comb in casting tray and agarose gel was put in the casting tray and kept for about 30 minutes to get solidified.

(c) After sodification of gel, agrose gel was placed in the gel casting tank which was filled with 1X TBE Buffer.

(d) Added the 3µl loading dye in 20µl PCR products. Finally, 20µl of each samples were loaded into wells facilitating the sizing of the various alleles Ladder (100 bp and 1kb) was loaded in the first well.

(e) Electrophoresis was done at 70 to 80 volts for about 30-45 minutes

(II) Visualization of amplified PCR products through agarose gel

Agarose gel was placed in the Gel Documentation system (Gel Doc) for visualizing the amplified PCR products.

3.8 GEL ELUTION

3.8.1 Materials

Razor blade or scalper, gel image, Gel Doc, microcentrifuge tube, hand gloves, Thermofisher Gel Extraction kit, Centrifuge machine .

3.8.2 Preapartion

- First Run the sample in gel electrophoresis then check the it in Gel Doc and take print the gel image.
- Turn on the illuminator using the switch, and slide the gel viewer at the top of the cover open.
- Using the scalper or razore blade, make quick and accurate cuts around the each band. Cut as close to the DNA as possible to minimized the gel volume.

- Place each gel piece into an empty pre-weighed microcentrifuge tube and labelled. Record the weight of the gel slice.

3.8.3 DNA extraction from the gel using centrifuge

Procedure

- Add 1:1 volume of Binding Buffer to gel slice.
- Incubate the gel mixture at 50-60℃ for 10 min or until the gel slice is completely dissolve. Mix the tube by inversion every few minutes to smooth the progress of the melting process. Ensure that the gel is completely dissolved. Vortex the gel mixture briefly before loading the column.
- Transfer up to 800 µl to the solubilised gel solution to the GeneJET purification column. Centrifuge on 12000 rpm for 1 min, Discard the flow through and place column back into the same collection tube. Repeat this step until solubilized gel solution will finish.
- Add 100 µl of Binding buffer to the GeneJET purification columm. Centrifuge 12000 rpm for 1 min. Discard the flow-through and take place the columm back into the same collection tube.
- Add 700 µl of wash buffer(diluted with ethanol) to the GeneJET purification column. Centrifuge on 12000 rpm for 1 min. Discard the flow through and place the column back into the same collection tube
- Centrifuge the empty column tube for an additional 1 min to completely remove residual wash buffer.
- Transfer the GeneJET purification column into a clean 1.5 ml microcentrifuge tube. Add 35 µl Elution buffer to the centre of the purification column membrane and centrifuge on 12000 rpm for 1 min.

3.9 HOMOLOGY MODELLING

Prediction of the three dimensional structure of a target protein from the amino acid sequence of a homologous (template) protein for which an X-ray or NMR structure is available.

Also called Comparative modelling & Knowledge-based modelling.

Method.

Figure below illustrates the mayor steps of obtaining structure from sequence.

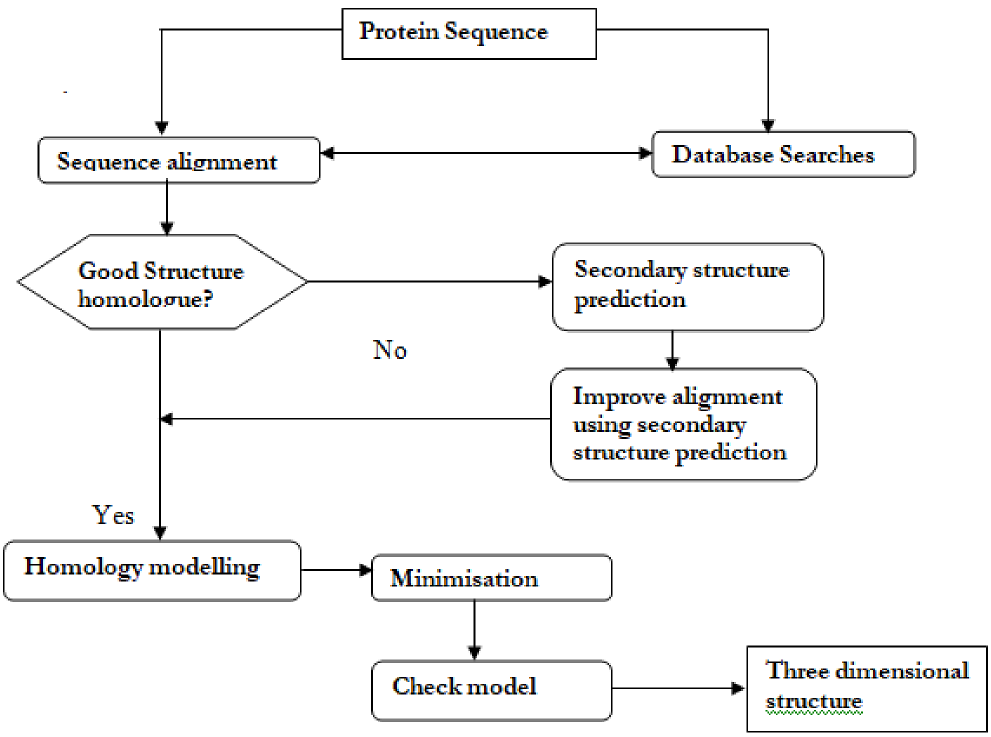


Fig-3.6 Steps of Homology modelling

Steps in molecular modelling:

1. Identification of structures that will form the template for the target structure (model).
2. Alignment – **the most important** step. Alignment of low homology sequences can be improved using secondary structure prediction (align-model-realign-remodel).
3. Transfer of coordinates from the template(s) to the target of structurally conserved regions (SCR's) - many fragment method
   - single structure.
4. Modelling variable regions
   - Loops
     - Insertions: Search of a high resolution fragment database
     - Deletions: Local minimisation often sufficient.
5. Modelling side chains (practically a virtual step)
6. Minimisation:

- Local – especially loop-hinge regions
- Global.

7. Molecular Dynamics: To study regional flexibility.
8. Checking the correctness of the model.
    - Correctness of the overall fold by:
        - **Bad**: Non-polar side chains exposed to the solvent.
        - **Bad**: Buried ionizable groups.
        - Conformational energy calculations – Incorrect folds have high solvation energy.
        - Luthy's method.
    - Stereochemical properties: ***PROCHECK***
        - Bond angles
        - Bond Length

Modelling through Swiss-Model(http://www.expasy.ch/swissmod/)

- Swiss_model first does a database search for homologous proteins. Then it Superposes all the structures it finds.

- It generates a multiple alignment with the sequence to be modelled and all the homologous structure

- Generates 3D framework for the target protein sequence.
- Atoms that occupy a similar spatial area and are aligned to the target sequence and are used to compute the averaged atomic position of the framework from which the target will be build.
- Side chains with incorrect geometries are removed

.

- Building of insertions or loops.SWISS_Model uses two techniques: The first method is the same as I described earlier. It also uses first principles, in other words it searches conformational space to build loops where: is uses 7 allowed ϕ, ψ combinations adequate space allocation for the loop space allocation for each α-carbon Both methods exclude loops in conflict with structure

- Side chain building

It also uses a library of allowed side-chain rotamers.First the distorted but otherwise complete side chains are corrected

Then the incomplete side chains are built with a probabilistic approach using the rotamers. A van der Waals exclusion test and dihedral angle constraints can be used to select the "best" side chain conformation

- ➢ Minimization

- ➢ The correctness of the structure is checked by analysing the conformational space of each residue energetically. The correctness of the structure is also checked by looking at the packing density of the model which is compared to what is expected.

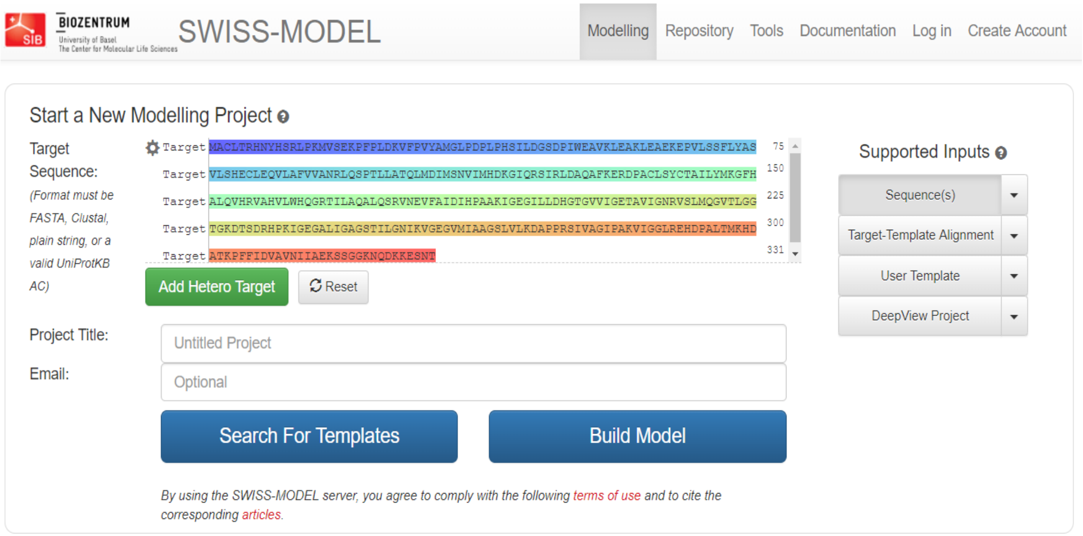


FIG-3.7 Interface of swiss model

.

3.10 DOCKING:

Step I – Building the Receptor This step the 3D structure of the receptor should be considered which can be downloaded from PDB; later the available structure should be processed. This should include removal of the water Exp - 12 molecules from the cavity, stabilizing the charges, filling the missing residues, generation the side chains etc according to the parameters available. The receptor should be biological active and stable state. Step II – Identification of the Active Site After the receptor is built, the active site within the receptor should be identified. The receptor may have many active sites but the one of the interest should be selected. Most of the water molecules and heteroatom if present should be removed. Step III – Ligand Preparation Ligands can be obtained from various databases like ZINC, PubChem or can be sketched using tools Chemsketch. While selecting the ligand, the LIPINSKY'S RULE OF 5 should be applied. The rule is important for drug development where a pharmacologically active lead structure is optimized stepwise for increased activity and selectivity, as well as drug-like properties as described. For selection of a ligand according to the LIPINSKY'S RULE: • Not more than 5 –H bond donors. • Molecular Weight NOT more than 500 Da. • Log P not over 5. • NOT more than 10 H bond acceptors. Step IV- Docking This is the last step, where the ligand is docked onto the receptor and the interactions are checked. The scoring function generates score depending on which the best fit ligand is selected.

CHAPTER – IV

RESULTS AND DISCUSSION

The present thesis entitled **"Analysis of *De novo* transcriptome sequence for identification of putative candidate genes related to β ODAP biosynthesis pathway in *Lathyrus sativus*"** was undertaken to understand the differential expression of different ODAP containing varieties of *Lathyrus sativus*. Further the correlation of genes expression levels with and biochemical traits associated with the ODAP response gene in Lathyrus sativus was also assessed.

4.1 Biochemical estimation of seed ODAP content in diverse Lathyrus genotypes

Biochemical estimation of seed ODAP estimation in 30 genotypes was done following the protocol of Rao et al.,2011..Aborbance was taken in Visible spectrophotometer at 425 nm wavelength.Abosorbance was taken in three biological replication then mean was calculated. Eleven genotypes including 6 high ODAP and 5 low ODAP) containing ines were choosen for total RNA isolation to see the differential gene expression pattern of high and low ODAP containg lines.

From 30 lines we have choosen 10 lines viz; AKL-19,IL-143099,SEL-519,SEL-522and SEL-527 are high ODAP containing line shown in red colour; & MAHATEORA,,NO-2208,PUSA-24,BIO L-202 are low ODAP containg lines shown in green colour and RLK-498 is a intermediate ODAP containg line.Apart from this we also have added RLK-1950 which is a high ODAP containg line.

TABLE 4.1-absorbance data of 30 genotypes

Variety	**R1**	**R2**	**R3**	**R4**	**MEAN**
AKL-19	0.858	0.96	0.98	0.922	0.93
BIOL-202	0.328	0.266	0.299	0.303	0.299
BIOL-203	0.791	0.859	0.829	0.822	0.825
BIOR231	0.737	0.767	0.634	0.661	0.700
IC-28112	0.413	0.445	0.425	0.379	0.416
IL-143099	0.836	0.864	0.833	0.815	0.837
JRL-41	0.72	0.827	0.759	0.861	0.792
MAHATEORA	0.202	0.192	0.2	0.161	0.189
NO-2203	0.663	0.699	0.672	0.714	0.687
NO-2208	0.355	0.347	0.386	0.32	0.352
PRATEEK	0.605	0.682	0.633	0.611	0.512
PUSA-24	0.533	0.523	0.508	0.482	0.512
RATAN	0.315	0.329	0.311	0.296	0.313
RLK-2059	0.729	0.775	0.751	0.754	0.752
RLK-498	0.639	0.616	0.584	0.54	0.595
RLS-2	0.784	0.798	0.765	0.744	0.773
SEL-439	0.521	0.539	0.544	0.483	0.522
SEL-481	0.687	0.725	0.624	0.648	0.671
SEL-504	0.656	0.542	0.554	0.57	0.581
SEL-508	0.687	0.672	0.713	0.639	0.678
SEL-516	0.713	0.716	0.686	0.654	0.692
SEL-519	0.856	0.921	0.91	0.895	0.896
SEL-520	0.789	0.808	0.776	0.751	0.781
SEL-522	0.825	0.907	0.928	0.766	0.857
SEL-527	0.879	0.85	0.943	0.973	0.911
SEL-531	0.463	0.481	0.453	0.444	0.460
SEL-553	0.676	0.71	0.69	0.631	0.677
SEL-554	0.744	0.72	0.671	0.73	0.716
SEL-563	0.787	0.86	0.802	0.81	0.815
SEL-587	0.432	0.491	0.454	0.434	0.453

4.2 In-silico analysis of transcriptome for identification of candidate genes related to ODAP biosynthesis.

Transcriptome is the sum total RNA content of an oraganisim , expreesed in a particular stage and particular tissue.

Transcriptome data of Mahateora and RLK-1950 was taken from Dr. Shubha Banerjee(Asst. Prof,IGKV,Raipur) which was published in thesis of Bhariya,sanjay.,2019.Based on available litreture on ODAP estimation data,it has been reported that Mahateora is low ODAP content line and RLK-1950 is high ODAP content line..Transcriptome sequencing was done by outsourcing through illumina sequencing. Two data sets 201M were generated (Mahateora transcriptome) and 202R (RLK-1950 transcriptome), which consists of 40,068,526 sequences in pairs and 39,165,312 sequences in pairs respectively. The total sequences and the total nucleotides in both data sets are 79,233,838 sequences and 11,964,309,538 nucleotides respectively.

	MAHATEORA		**RLK-1950**	
Reads mapped in pairs	16,893,882	42.16%	16,205,062	41.38%
Reads mapped in broken pairs	5,515,964	13.11%	5,740,956	14.66%
Reads not mapped	17,658,680	44.07%	17,219,294	43.97%
Total	40,068,526	100%	39,165,312	100.00

(Table 4.2- Graphical representation of Sequencing data)

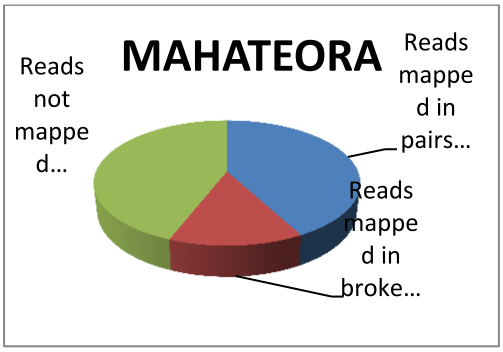


Fig-4.1

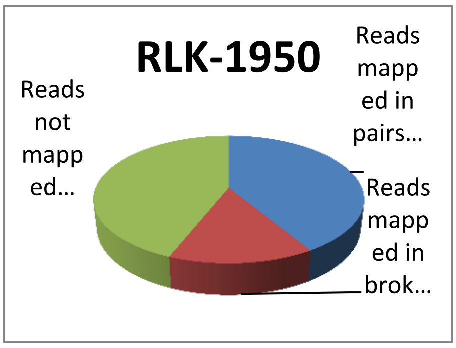


Fig-4.2

In the above fig-4.1 and Fig-4.2, transcriptome data of Mahateora and RLK-1950 are shown in pie chart

QC report provides length distribution, GC content, ambiguous base content, quality distribution, coverage, nucleotide distribution 5 mers and sequence duplication level, which are shown below :

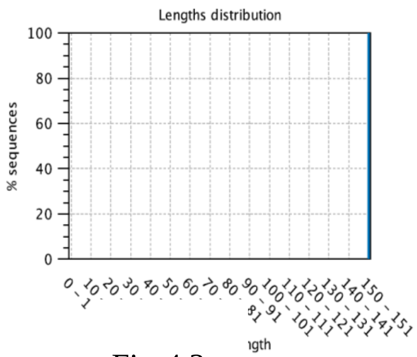


Fig-4.3

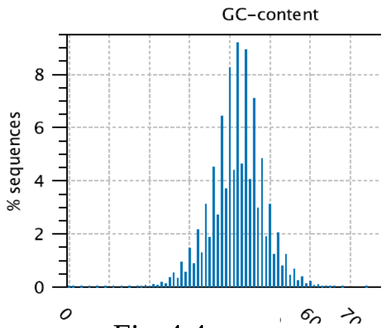


Fig-4.4

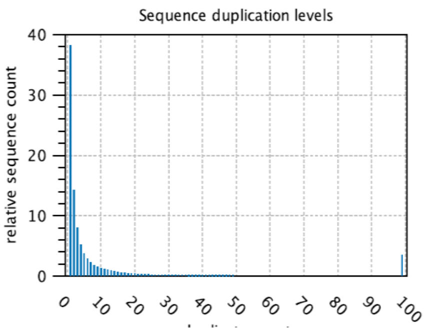


Fig-4.5

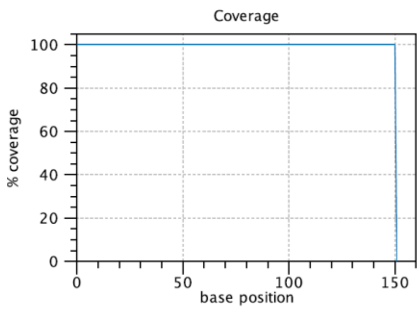


Fig-4.6

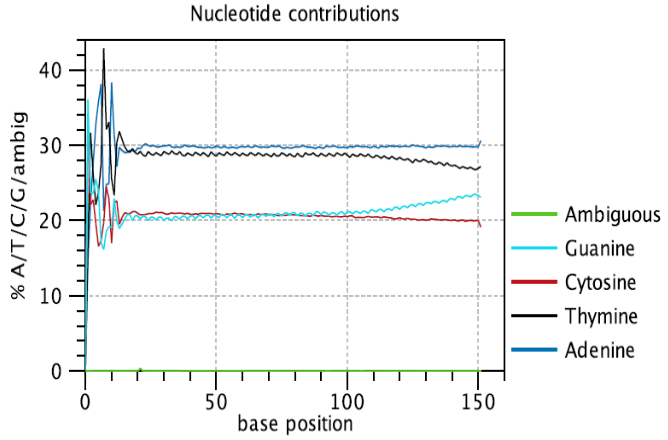


Fig-4.7

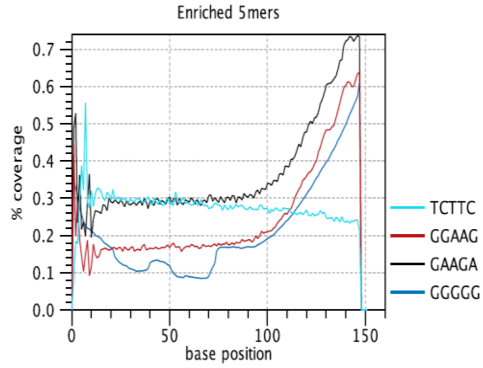


Fig-4.8

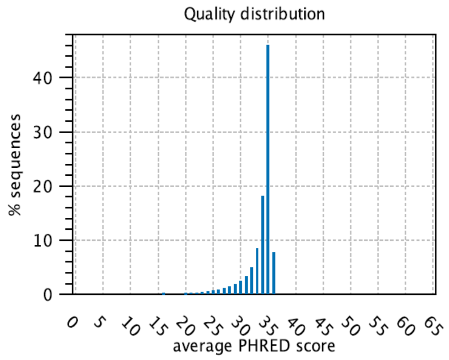


Fig-4.9

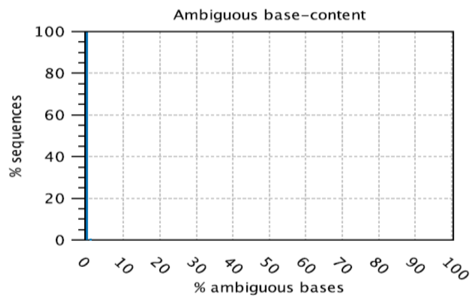


Fig-4.10

The above figures from 4.3 to 4.10 depicted the quality of transcriptome which was generated through QC report.Modern high-throughtput sequencers generates million no of fragments in a single run.So before analysing these sequence to draw biological conclusions we should always check the quality report of sequencing to ensure wheather raw data is good or bad.

Here Fig 4.3 shows the length of nucleotide distribution.If all the fragments are evenly distributed then it will be a good result.We can see the all the fragments are in a range 150-151.

Fig-4.4 shows the total GC content.A good quality sequence results should have normally distributed GC ontent where central peak corresponds to the overall GC content of the underlying genome. An unusually shaped distribution could indicate a contaminated library or some other kinds of biased subset

Fig-4.5 shows the duplication level of sequencing result.In a diverse libreray most of the sequences occurs only once in the set.A low level of duplication,which showed in the range of 0-20, depicts that it is very good.

Fig-4.6 shows the total coverage of the sequence. Low level of duplication shows high level of coverage.Fif-4.6 shows the coverage of 100.

Fig 4.7- shows the % of nucleotide contribution.% of nucleotide contribution should be in the range of 5 % to 20%.

Fig-4.8 shows 5 mer enriched graph. The Kmer module starts from the assumption that any small fragment of sequence should not have a positional bias in its apearance within a diverse library. This module measures the number of each 5-mer at each position in above library and then uses a binomial test to look for significant deviations from an even coverage at all positions.

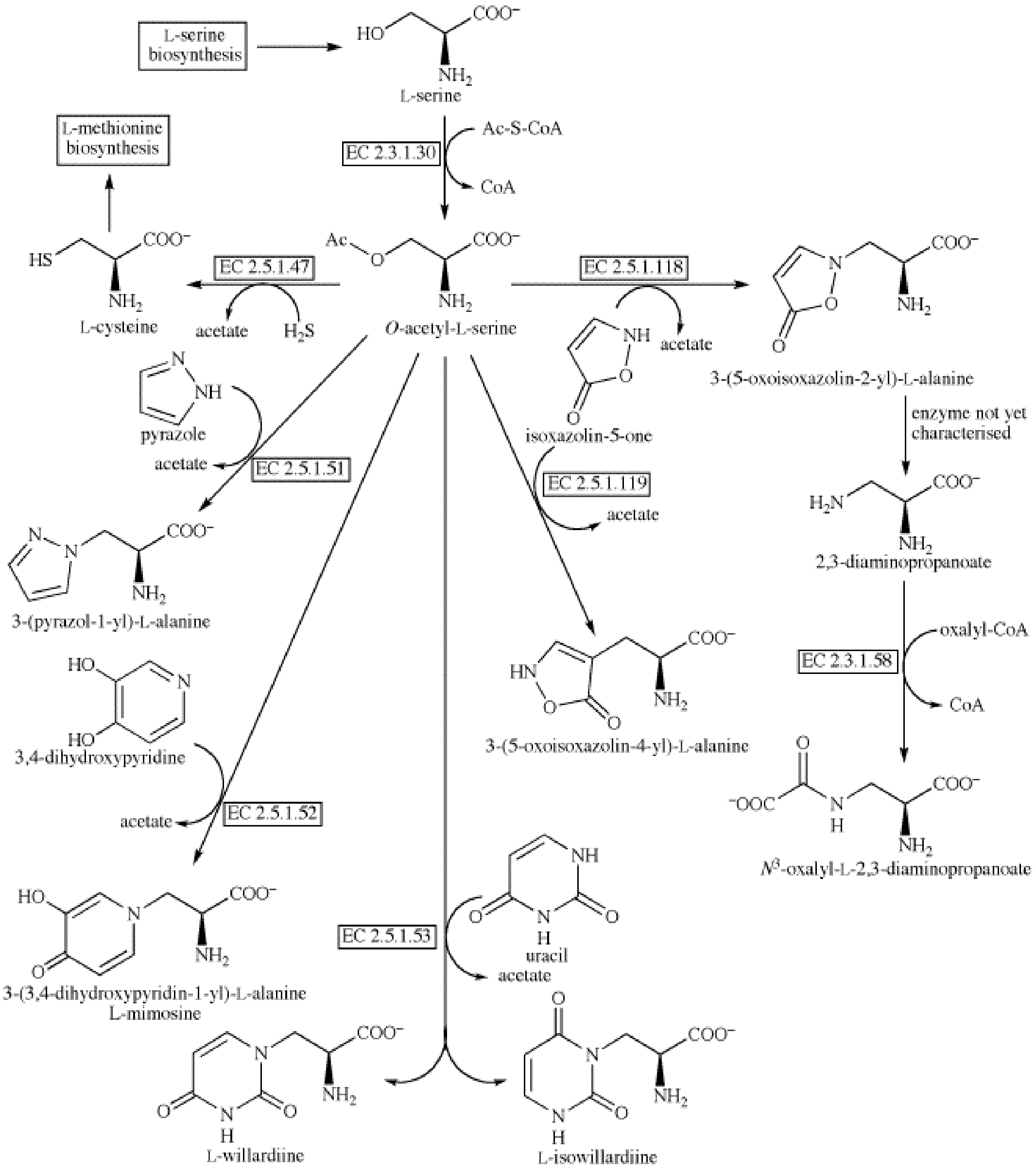
Fig-4.9 shows the quality distribution.Observed Phred mean should be greater than 27,this equated 0.2% error rate. In our graph mean phred score is 35.

Fig-4.10 shows the ambiguous base content.

All this explanation are referred from barbraham bioinformatics.(https://www.bioinformatics.babraham.ac.uk/projects/fastqc/Help/3%20Analysis%20Modules/)

From above graphs and tables we have an clear idea that our sequencing result was very good.Our objective was to screen the sequences related to ODAP biosynthesis pathway.If we look upon into the biosynthetic pathway(Fig-4.11) (Malathi *et al*., 1970 and Kuo *et al*., 1998),then we can know that EC 2.3.1.30 serine *O*acetyltransferase ,EC 2.3.1.582,3diaminopropionate *N*-oxalyltransferase ;EC 2.5.1.47 cysteine synthase ;EC 2.5.1.51 βpyrazolylalanine Synthase ;EC 2.5.1.52 Lmimosine synthase ;EC 2.5.1.53 uracilylalanine synthase EC 2.5.1.118β(isoxazolin5on2yl) L-alaline synthase EC 2.5.1.119β(isoxazolin5on4yl)Lalanine synthase, are participated in ODAP biosynthesis pathway.

Amomg these enzyme we worked upon to screen two enzymes viz; serine *O*-acetyltransferase(EC 2.3.1.30) and L-alanine synthase.For that reason we followed two approaches viz;NCBI BLAST method and LOCAL BLAST method for scanning of our transcriptome data.



(Fig-4.11 β- ODAP Biosynthesis in Grass pea)

EC 2.3.1.30 serine *O*-acetyltransferase
EC 2.3.1.58 2,3-diaminopropionate *N*-oxalyl transferase
EC 2.5.1.47 cysteine synthase
EC 2.5.1.51 β-pyrazolylalanine synthase
EC 2.5.1.52 L-mimosine synthase
EC 2.5.1.53 uracilylalanine synthase
EC 2.5.1.118 β-(isoxazolin-5-on-2-yl)-L-alanine synthase
EC 2.5.1.119 β-(isoxazolin-5-on-4-yl)-L-alanine synthase

4.3 NCBI BLAST METHOD

NCBI(National Center for Biotechnology Information) (https://www.ncbi.nlm.nih.gov/)

Is a strong online server data base and depository of all the nucleotide and protein sequenc which has been sequence. It gives comaparitive results with our query sequences.

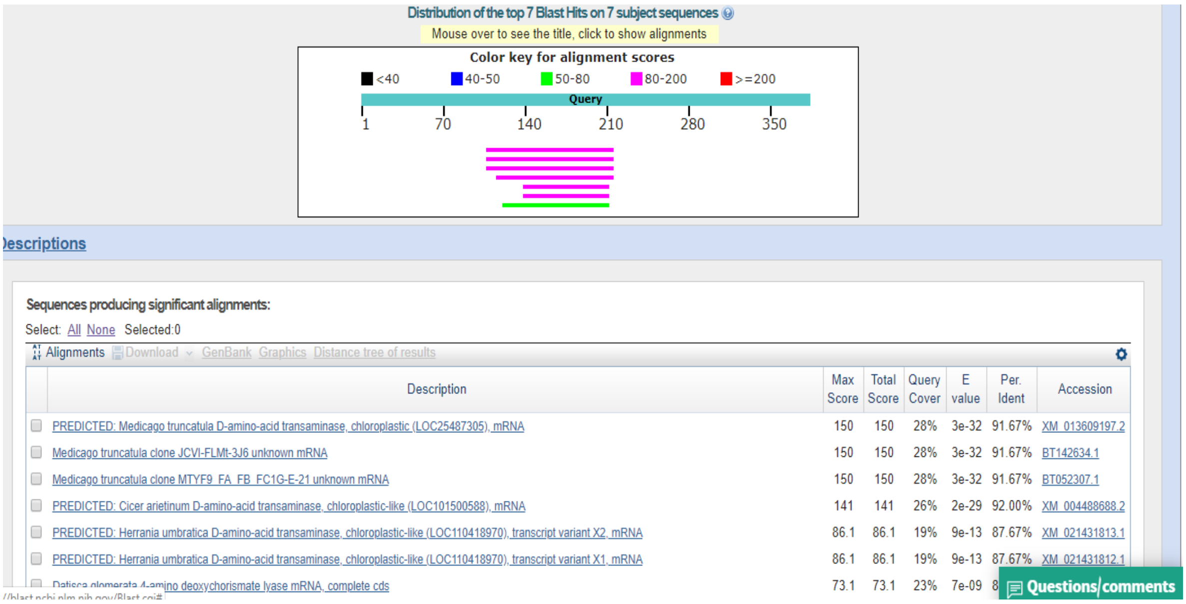
To screen the gene of our interest, nucleotide blast was conducted. Default parameters like cost to open a gap [integer], cost to extend a gap [integer], reward for a match in the BLAST portion of run [integer], penalty for a mismatch in the blast portion of run [integer],expection value(E), word size, max score, max alignment was kept as 5, 2, 1, 3, 10.0, 11, 25 and 15 respectively.

Among 28,135 annoted transcripts, we blasted around 14000 transcripts. Blast was carried out through direct NCBI blast by putting our query sequence in online server and through BLAST2GO software. From the BLASTn result we screened 500 sequence like synthase,transferase.ligase,transporters and enzymes related in amino acid metabolism.

Ikegami, F.; Horiuch et al has told that ODAP biosynthetic pathway is connected with the nitrogen and sulphur metabolism,which is alsp related to amino acid metabolism pathway. So we screened all the transcripts which are related to nitrogen,sulphur and amino acid metabolism.

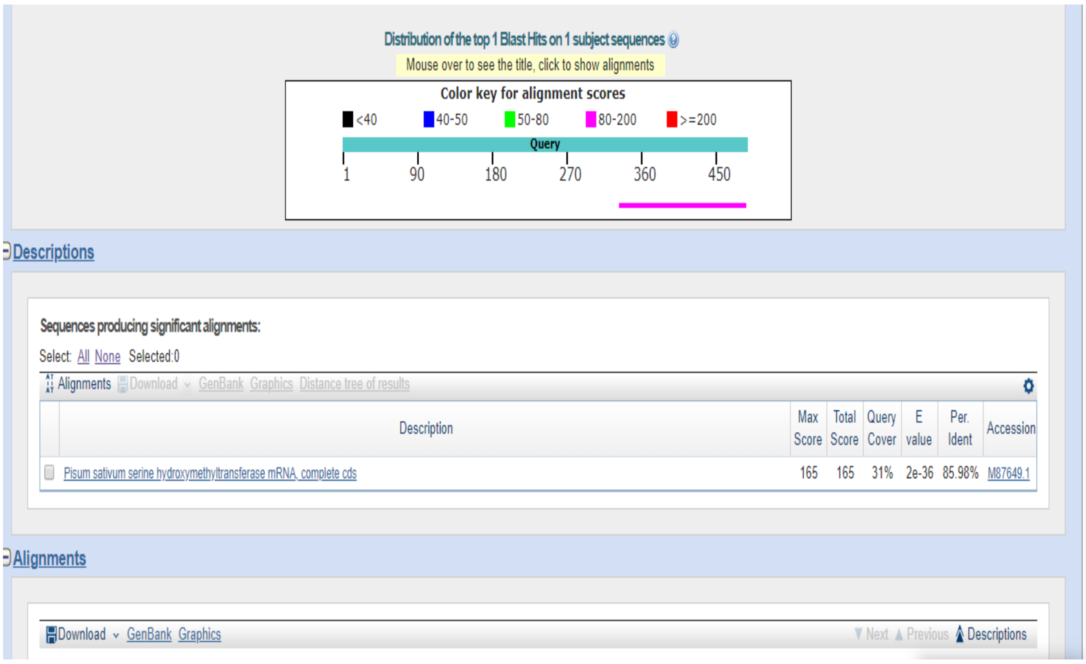
From the BLASTn reult we identified 6 transcripts viz; D-amino acid transferase(gene id-13637,13636); serine hydroxyl methyl transferase(ID-13825); β-alaline synthase(ID-1676); Cysteine Synthase(ID-13847); Branched chain Amino acid transferase(ID-12053).

Blast result of above transcripts:



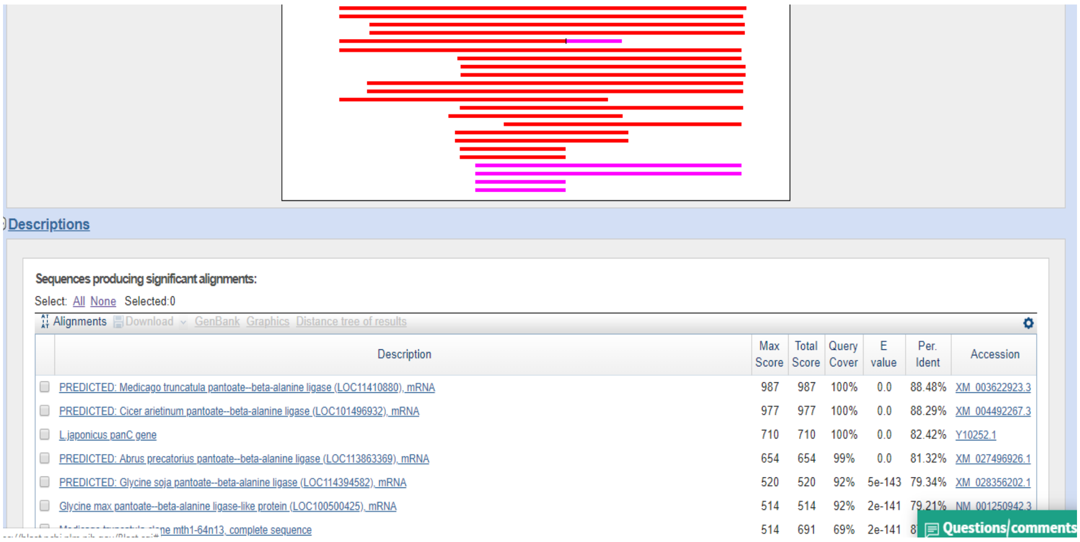
	Description	Max Score	Total Score	Query Cover	E value	Per. Ident	Accession
	PREDICTED: Medicago truncatula D-amino-acid transaminase, chloroplastic (LOC25487305), mRNA	150	150	28%	3e-32	91.67%	XM_013609197.2
	Medicago truncatula clone JCVI-FLMt-3J6 unknown mRNA	150	150	28%	3e-32	91.67%	BT142634.1
	Medicago truncatula clone MTYF9_FA_FB_FC1G-E-21 unknown mRNA	150	150	28%	3e-32	91.67%	BT052307.1
	PREDICTED: Cicer arietinum D-amino-acid transaminase, chloroplastic-like (LOC101500588), mRNA	141	141	26%	2e-29	92.00%	XM_004488688.2
	PREDICTED: Herrania umbratica D-amino-acid transaminase, chloroplastic-like (LOC110418970), transcript variant X2, mRNA	86.1	86.1	19%	9e-13	87.67%	XM_021431813.1
	PREDICTED: Herrania umbratica D-amino-acid transaminase, chloroplastic-like (LOC110418970), transcript variant X1, mRNA	86.1	86.1	19%	9e-13	87.67%	XM_021431812.1
	Datisca glomerata 4-amino deoxychorismate lyase mRNA, complete cds	73.1	73.1	23%	7e-09	8[illegible]	[illegible]

(Fig-4.12 blast result of ID-13637)



	Description	Max Score	Total Score	Query Cover	E value	Per. Ident	Accession
	Pisum sativum serine hydroxymethyltransferase mRNA, complete cds	165	165	31%	2e-36	85.98%	M87649.1

(Fig-4.13 Blast result of ID13825)



	Description	Max Score	Total Score	Query Cover	E value	Per. Ident	Accession
	PREDICTED: Medicago truncatula pantoate--beta-alanine ligase (LOC11410880), mRNA	987	987	100%	0.0	88.48%	XM_003622923.3
	PREDICTED: Cicer arietinum pantoate--beta-alanine ligase (LOC101496932), mRNA	977	977	100%	0.0	88.29%	XM_004492267.3
	L.japonicus panC gene	710	710	100%	0.0	82.42%	Y10252.1
	PREDICTED: Abrus precatorius pantoate--beta-alanine ligase (LOC113863369), mRNA	654	654	99%	0.0	81.32%	XM_027496926.1
	PREDICTED: Glycine soja pantoate--beta-alanine ligase (LOC114394582), mRNA	520	520	92%	5e-143	79.34%	XM_028356202.1
	Glycine max pantoate--beta-alanine ligase-like protein (LOC100500425), mRNA	514	514	92%	2e-141	79.21%	NM_001250942.3
	...ne mth1-64n13, complete sequence	514	691	69%	2e-141	8	

Fig-4.14 Blast result of ID 1676,β-alaline ligase

4.4 LOCAL BLAST METHOD:

We conducted local blast through Blast2go software.We used RLK-1950 transcriptome data as a domain of data base then we formatted our data base to .nhd,nhr,.nin.,nog. Format. After that we downloaded the sequence of SAT(Serine acetyltransferase) from *Staphylococcus aureus (strain MRSA252). Uniprot id-* Q6GJE0 (CYSE_STAAR). We used this sequence as our query sequence.

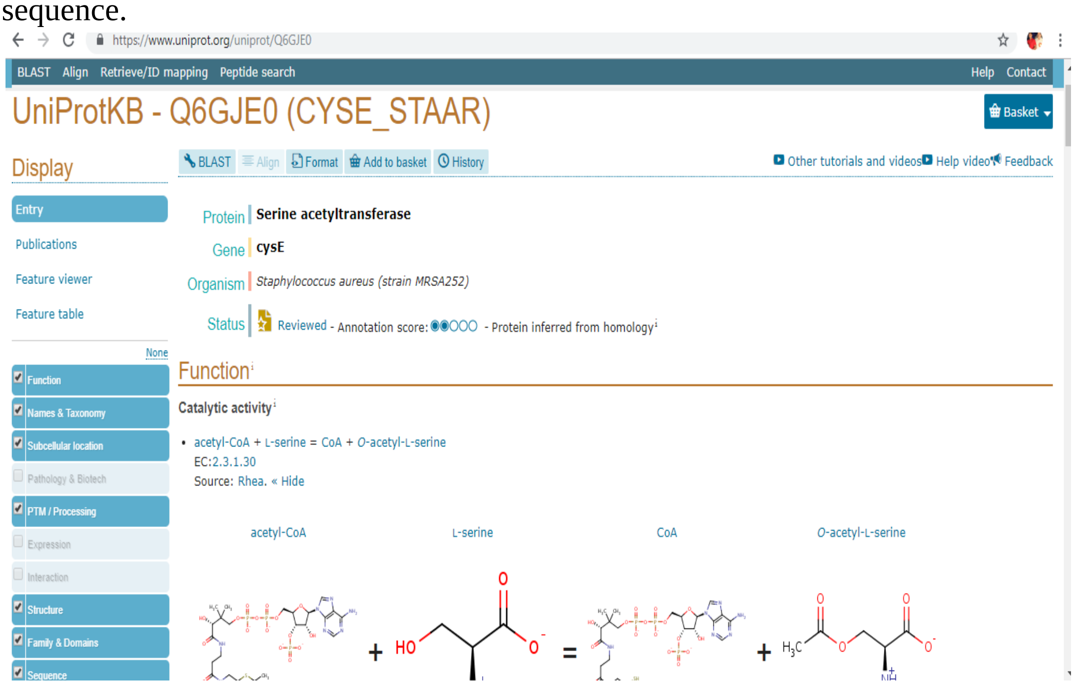


Fig-4.15- SAT gene of *Staphylococcus aureus*

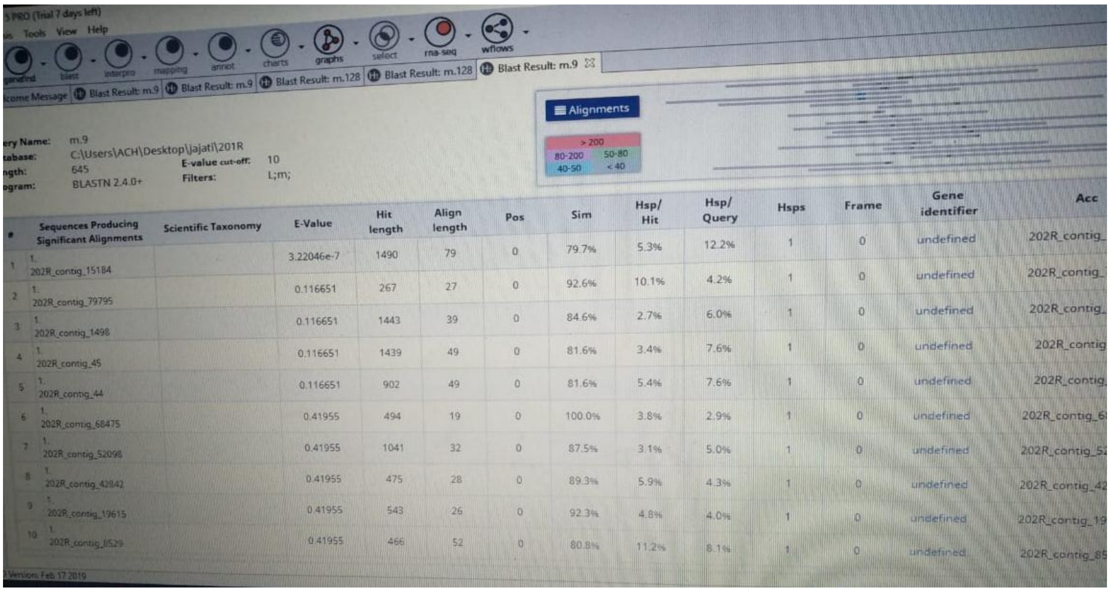


Fig-4.16 Local blast result of SAT

From the above Blast2go result we pick up 2 sequence having high align length,similairity, and high hit length viz; contig no 15184 and gene ID 15478.Then we again did the NCBI blast to know the exact name of the above sequence. We found that gene ID 15478 is resembles with serine acetyl-transferase.

(Fig-4.17 Blast result of candidate gene contig No-15184)

4.5 Amplification of cDNA with candidate gene prime

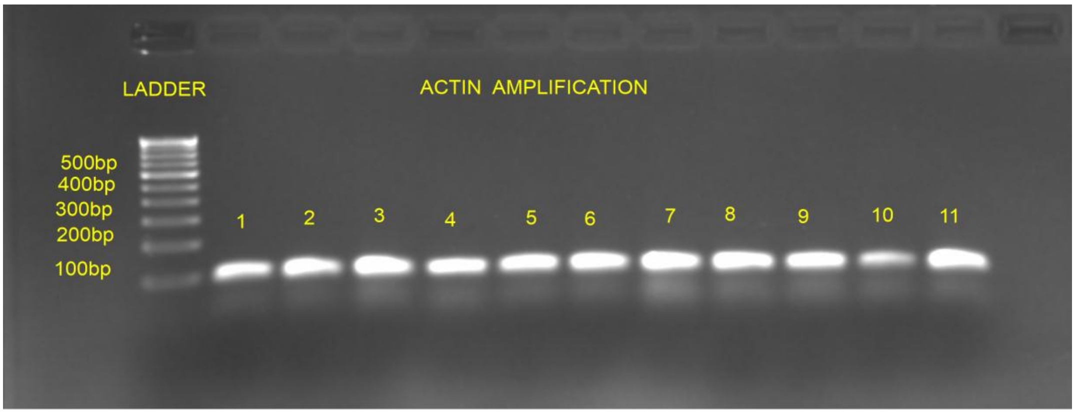
For expression analysis ,RNA was extracted from 10 days old leaves,because ODAP synthesis is known to occure in early seedling satge.

From the above results we got the candidate genes of serine acetyltransferase and L-alaline synthase.However it was not clear wheather these sequence are active in ODAP biosyntehsis pathway.Therefore we have designed primer from the 11 candidate sequences from primequest IDT.(https://eu.idtdna.com/PrimerQuest/Home/Index) .The primers were than used for amplification of our cDNA,isolated from 10 days old plant. For expression analysis ,RNA was extracted from 10 days old leaves,because ODAP synthesis is known to occure in early seedling satge. (Jiao,C.J;Xu,Q.L *et al.*,).In contrast to it all mature leaves,stems and roots show a low level of ODAP accumulation. From their experiment it was evident that ODAP content is inversely proportional to the age of plant.

(Table 4.3 primer list of 11 transcripts)

Sl no	PRIMER NAME	Candidate genes
1	PaaY	Serine acetyl transferase(ID-15478)
2	PLN1	D-Amino acid transferase(ID-13637)
3	PLPD1	D-Amino acid transferase(ID-13637)
4	TRP1	Cysteine syntahse(ID-13847)
5	PLN2	Serine hydroxyl methyl transferase(ID-13825)
6	PLN3	β-alaline syntahse(ID-13847)
7	UNK1	Uknown sequence(contig-68475)
8	UNK2	Uknown sequence(contig-19615)
9	AMIF	Uredio propionase(ID-11699)
10	AMIFa	Uredio propionase(ID-5965)
11	DeGQ	Chloroplastic, mRNA(ID-13179)

To normalize the expression profiling,expression of actin was done.Actin primers are also designed from the *Lathyrus* transcriptome sequencing for amplification.Expression of actin gene was similar in all the genotypes as expected,being a homologous gene.



(FIG-4.18 Amplification of house keeping gene LsActin)

1-No-2208 ;**2**-SEL-519 ;**3**-IL-143099; **4**-IL-527 **5**-AKL-19;**6**-MAHATEORA(.32); **7**-PUSA-24 ;**8**-BIO L-202 **9**-RLK-498 **10**-SEL-522 **11**-RLK-1950

Gel image of 11 primes are listed below.

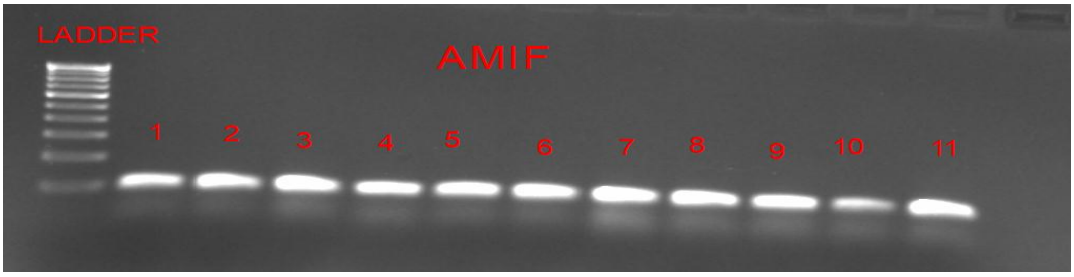


Fig-4.19 primer AMIF- Uredio propionase(ID-11699

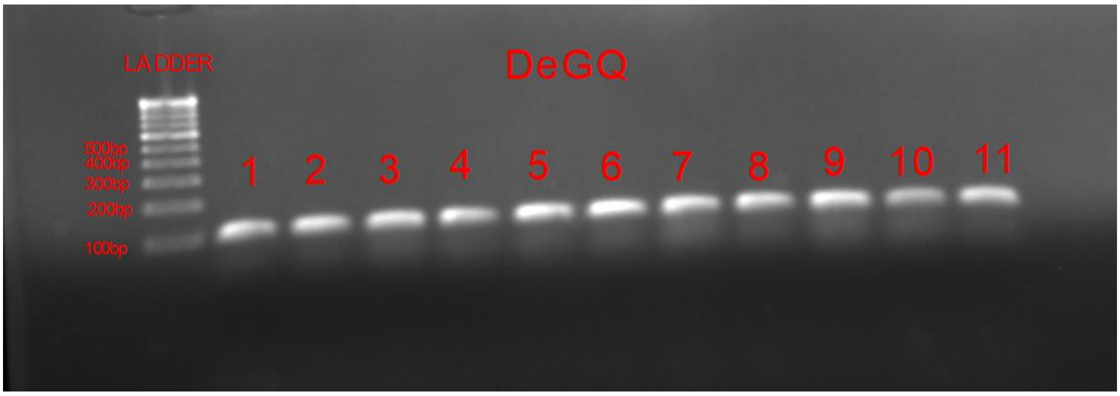


Fig-4.20 Primer DeGQ- Chloroplastic, mRNA(ID-13179)

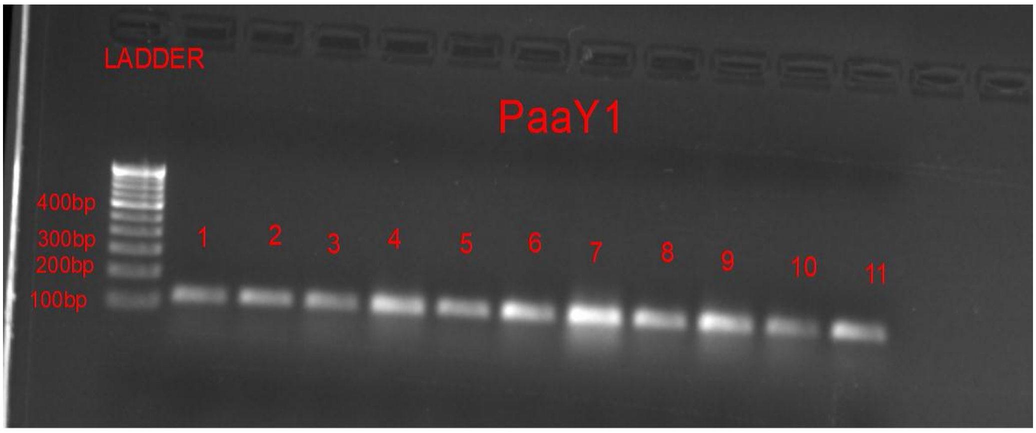


Fig-4.21 Primer PaaY- Serine acetyl transferase(ID-15478)

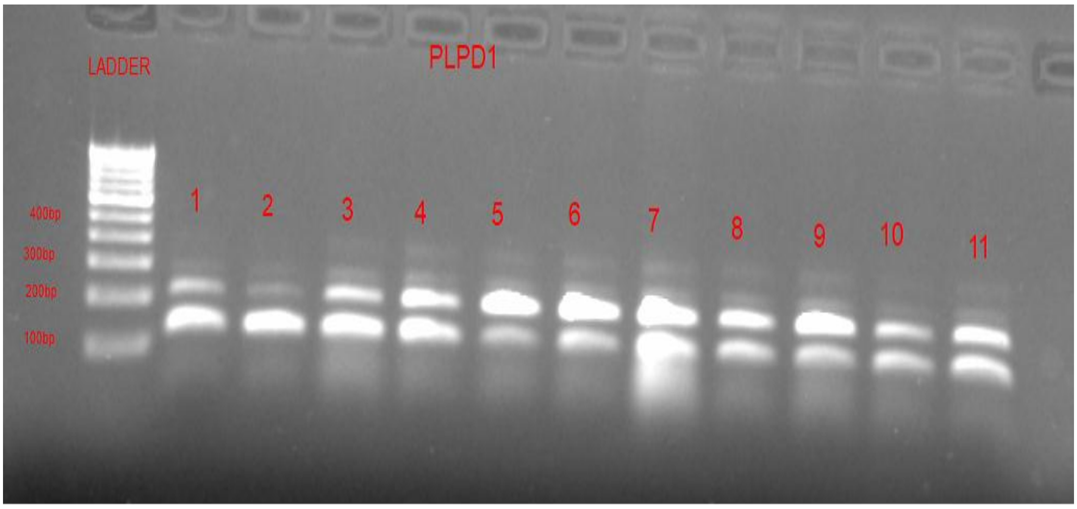


Fig-4.22 Primer PLPD1- D-Amino acid transferase(ID-13637)

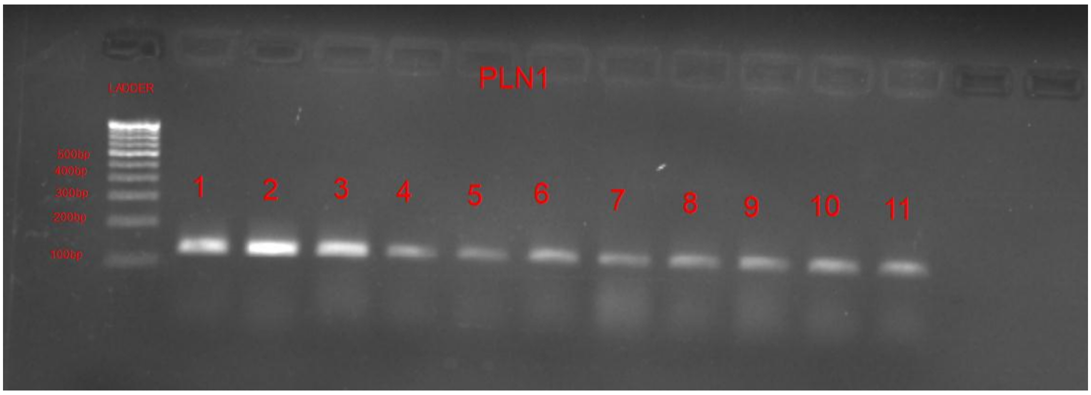


Fig-4.23 Primer PLN1- D-Amino acid transferase(ID-13637)

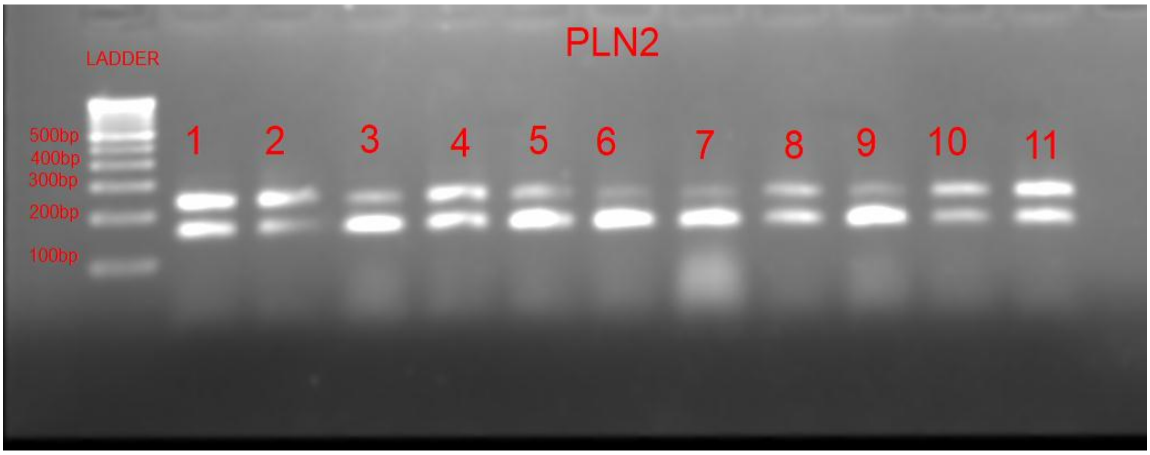


Fig-4.24 Primer PLN2- Serine hydroxyl methyl transferase(ID-13825)

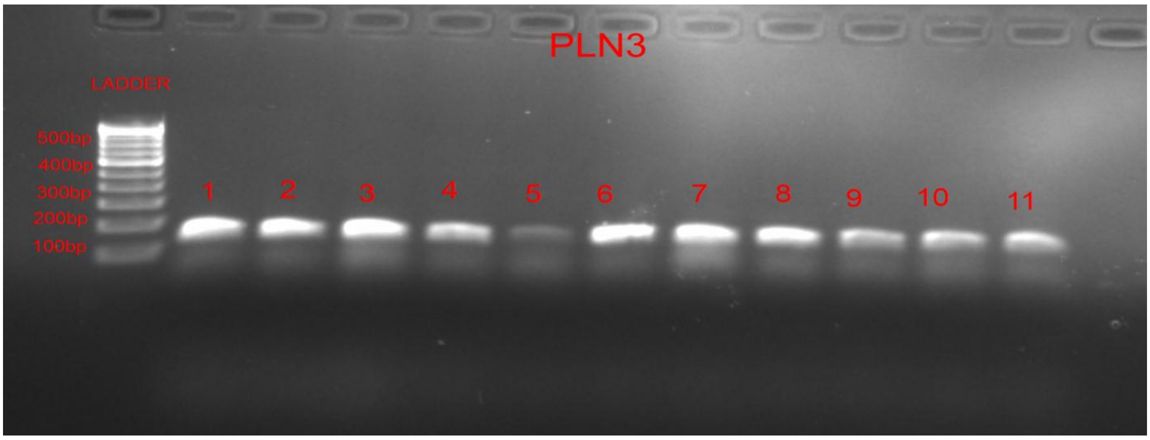


Fig-4.25 Primer PLN3- β-alaline syntahse(ID-13847)

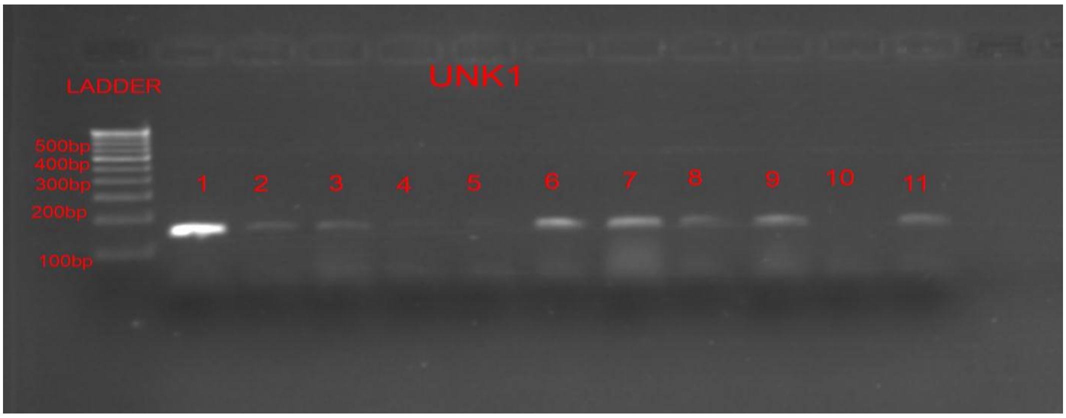


Fig-4.26 Primer UNK1- Uknown sequence(contig-68475)

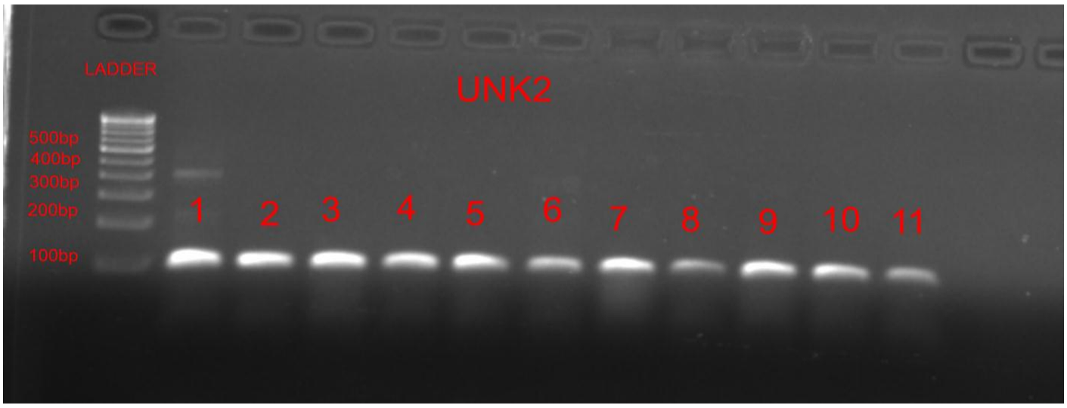


Fig-4.27 Primer UNK2- Uknown sequence(contig-19615)

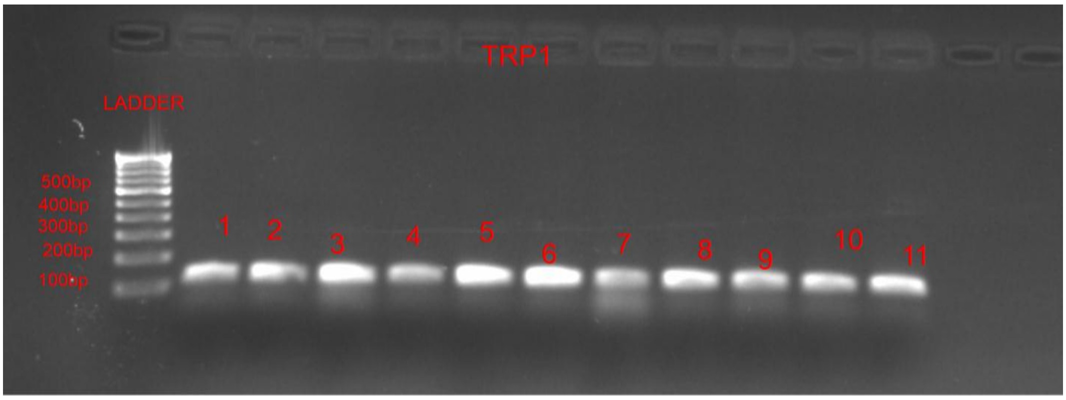


Fig-4.28 Primer TRP1- Cysteine syntahse(ID-13847)

No of genotypes shown sequentially.

1-No-2208 ;**2**-SEL-519 ;**3**-IL-143099; **4**-IL-527 **5**-AKL-19;**6**-MAHATEORA(.32); **7**-PUSA-24 ;**8**-BIO L-202 **9**-RLK-498 **10**-SEL-522 **11**-RLK-195

From the above gel image we can understand that primer(PLPD) designed from candidate gene ID-13637 showed a two bands of 150 and 250 bp with differential expression to the seed ODAP content was observed in the 11 genotypes for both the amplicon. PaaY shows similar expression through out 11 genotypes which was in accordance with to its role in ODAP biosynthesis.

The gel result of primer PLPD can be correlated with the pathway which was proposed by Malathi et al., K., Padmanaban, G. and Sarma(1970). This Pathway clearly shows that l-alaline has two isoforns viz; EC 2.5.1.118 β(isoxazolin-5on-2yl) L-alanine synthase ;EC 2.5.1.119 β(isoxazolin-5-on-4yl) Lalanine synthase.

Another interesting observation is that when the expression of β(isoxazolin-5-on-4yl is high then the expression of β(isoxazolin-5on-2yl decreases. So expression of both the isoforms are inversely proportional to each other. This might be depicted from the gel result that when the enzyme β(isoxazolin-5-on-4yl) L-alanine synthase act more on the substrate o-acetyl L-serine then 3-(5-isoxazolin-5-on-4yl) L-alaline is formed and when β(isoxazolin-5-on-2yl) L-alanine synthase act upon the same substare then β(isoxazolin-5-on-2yl) L-alanine is formed which eventually leads to ODAP fornation. So o-acetyl l serine is a common substrate for both the isoforms.

(TABLE 4.4- correlation of gel image with ODAP biosynthesis pathway)

Ligand	enzyme	Product	Remark
So o-acetyl l serine	β(isoxazolin-5on-2yl) L-alanine synthase(EC 2.5.1.118)	β(isoxazolin-5-on-2yl) L-alanine	Corealated with the upper band of primer PLPD1
So o-acetyl l serine	EC 2.5.1.119 β(isoxazolin-5-on-4yl) Lalanine synthase	3-(5-isoxazolin-5-on-4yl) L-alaline	Corealated with the lower band of primer PLPD1

Gel image of primer PaaY1(serine acetyl transferase) showed a similar level of expression in all the genotypes. If we again correlate this result with the pathway of Malathi *et al;(*1970) then we can say that equal expression of the gene in both high as well as low ODAP containing genotypes is obvious.It act upon the substrate L-serine and gives O-acetyl l-serine.

4.6 DOCKING OF THE PUTATIVE PROTEIN

Molecular docking is used to determine the structure of the intermolecular complex formed between two molecules. Smaller molecule called ligand which interact with the protein's binding site. Binding sites are the proteins active site. There are several possible mutual conformations in which binding may occur. These are commonly called binding modes. It also predicts the strength of the binding, the energy of the complex; the types of signal produced and calculate the binding affinity between two molecules using scoring functions. Lower the activation energy more is the stability of ligand-protein complex. (Inbar,Yuval *et al.*,2005)

We used the pathcdock server for docking purpose.

4.6.1 Patch dock results of candidate gene L-alaline synthase with O-acetyl serine.

PatchDock is an online server and it does geometry-based molecular docking algorithm. (Duhovny,D., Nussinov,R. *et al.*, 2002) It is aimed at finding docking transformations that yield good molecular shape complementarity. Such transformations, when applied, induce both wide interface areas
and small amounts of steric clashes

In this docking process a template of candidate gene L-alaline synthase as protein and O-acetyl serine as ligand was taken . XML format of liagnd o-acetyl serine from ligandbook(https://ligandbook.org) was doenloaded. Ligandbook is a online depository of different ligands and small drug like molecules .It enables the rapid set up of reproducible molecular dynamics simulations of ligands and protein-ligand complexes. Then the the ligand, from XML format to .pdb format was converted through **openbible2.4** software. The Homology modelling results indicated a a template 5e25 which was downloaded from PDB database.(https://www.rcsb.org/structure/5E25). Then simulation process was carried out through patchdock online server.(https://bioinfo3d.cs.tau.ac.il/PatchDock/).ALL The structure are viewed in PYMOL software.

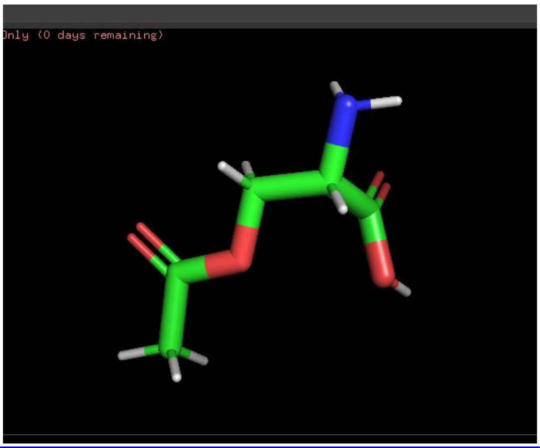
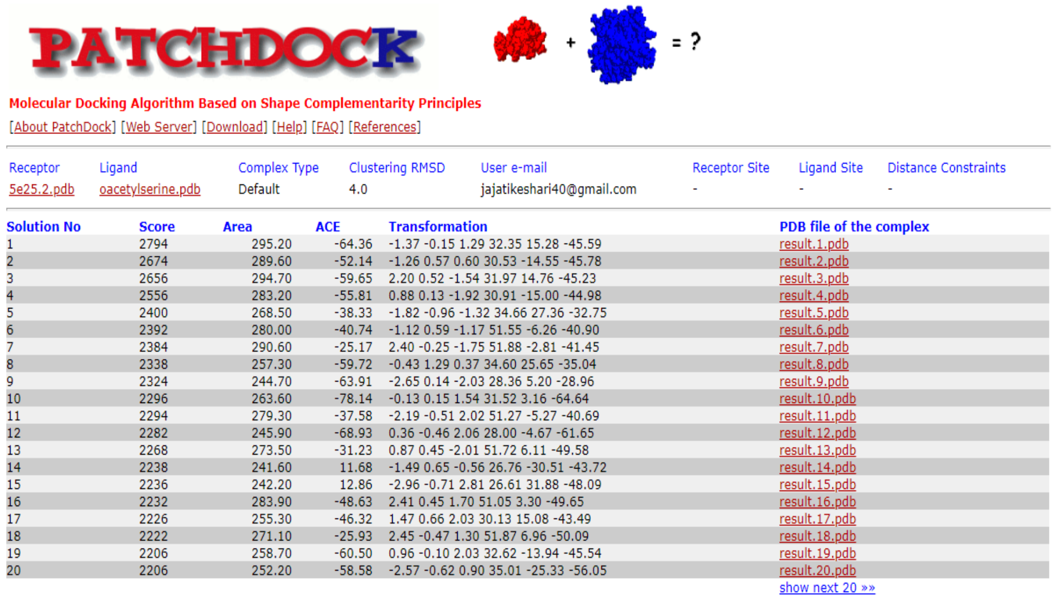


Fig-4.29 pdb format of ligand O-acetyl serine)

Fig-4.30 3D view of candidate gene L-alaline synthase


PATCHDOCK

Molecular Docking Algorithm Based on Shape Complementarity Principles

[About PatchDock] [Web Server] [Download] [Help] [FAQ] [References]

Receptor	Ligand	Complex Type	Clustering RMSD	User e-mail	Receptor Site	Ligand Site	Distance Constraints
5e25.2.pdb	oacetylserine.pdb	Default	4.0	jajatikeshari40@gmail.com	-	-	-

Solution No	Score	Area	ACE	Transformation	PDB file of the complex
1	2794	295.20	-64.36	-1.37 -0.15 1.29 32.35 15.28 -45.59	result.1.pdb
2	2674	289.60	-52.14	-1.26 0.57 0.60 30.53 -14.55 -45.78	result.2.pdb
3	2656	294.70	-59.65	2.20 0.52 -1.54 31.97 14.76 -45.23	result.3.pdb
4	2556	283.20	-55.81	0.88 0.13 -1.92 30.91 -15.00 -44.98	result.4.pdb
5	2400	268.50	-38.33	-1.82 -0.96 -1.32 34.66 27.36 -32.75	result.5.pdb
6	2392	280.00	-40.74	-1.12 0.59 -1.17 51.55 -6.26 -40.90	result.6.pdb
7	2384	290.60	-25.17	2.40 -0.25 -1.75 51.88 -2.81 -41.45	result.7.pdb
8	2338	257.30	-59.72	-0.43 1.29 0.37 34.60 25.65 -35.04	result.8.pdb
9	2324	244.70	-63.91	-2.65 0.14 -2.03 28.36 5.20 -28.96	result.9.pdb
10	2296	263.60	-78.14	-0.13 0.15 1.54 31.52 3.16 -64.64	result.10.pdb
11	2294	279.30	-37.58	-2.19 -0.51 2.02 51.27 -5.27 -40.69	result.11.pdb
12	2282	245.90	-68.93	0.36 -0.46 2.06 28.00 -4.67 -61.65	result.12.pdb
13	2268	273.50	-31.23	0.87 0.45 -2.01 51.72 6.11 -49.58	result.13.pdb
14	2238	241.60	11.68	-1.49 0.65 -0.56 26.76 -30.51 -43.72	result.14.pdb
15	2236	242.20	12.86	-2.96 -0.71 2.81 26.61 31.88 -48.09	result.15.pdb
16	2232	283.90	-48.63	2.41 0.45 1.70 51.05 3.30 -49.65	result.16.pdb
17	2226	255.30	-46.32	1.47 0.66 2.03 30.13 15.08 -43.49	result.17.pdb
18	2222	271.10	-25.93	2.45 -0.47 1.30 51.87 6.96 -50.09	result.18.pdb
19	2206	258.70	-60.50	0.96 -0.10 2.03 32.62 -13.94 -45.54	result.19.pdb
20	2206	252.20	-58.58	-2.57 -0.62 0.90 35.01 -25.33 -56.05	result.20.pdb

show next 20 »»

(Fig4.31- Results of ligand and protein interaction obtained from patchdock)

By default Patchdock provides 20 most probable ligand-protein complex having highest score. Among best 20 complexes we choose 1st complex having lowest activation energy and highest scoring function.

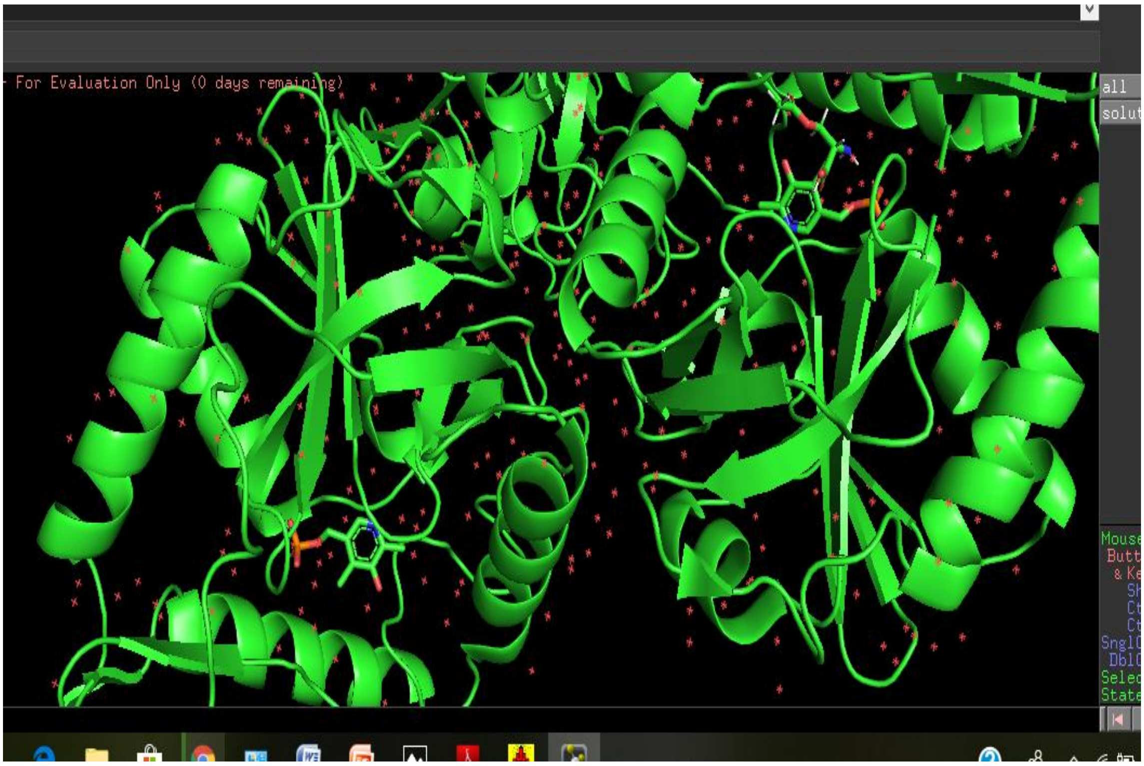


Fig-4.32 PYMOL view of ligand-protein complex.

Malathi et al., had proposed that enzyme l-alaline synthase has two isoforms viz. β(isoxazolin-5on-2yl) L-alanine synthase(EC 2.5.1.118) and EC 2.5.1.119 β(isoxazolin-5-on-4yl) Lalanine synthase and worked on same substrate o-acetyl-l-serine.

In figure 4.32 it is clearly observed that the candidate protein l-alaline synthase has two ligand binding site or ligand binding pocket.This results of docking correlates with the gel image of primer PLPD1.So two binding pocket of candidate genes binds with ligand and gives rise to two different products viz : β(osoxazolin5on2yl)Lalanine and β(osoxazolin5on4yl) L-alanine (Malathi et al;1970)

4.6.2 Patch dock results of candidate gene serine o- acetyl transferase with ligand L-serine

Like above ligand L-serine structure was downloaded from ligandbook and converted into .pdb format through openbable and viewed through PYMOL.These two ligand and protein was docked in patchdock sever and viewed in PYMOL.A template of protein serine o-acetyl transferase(SAT) was downloaded from pdb data bank. PDB data bank is repository of proteins 3D structure which id determined by NMR spectroscopy and X-RAY crystallography.

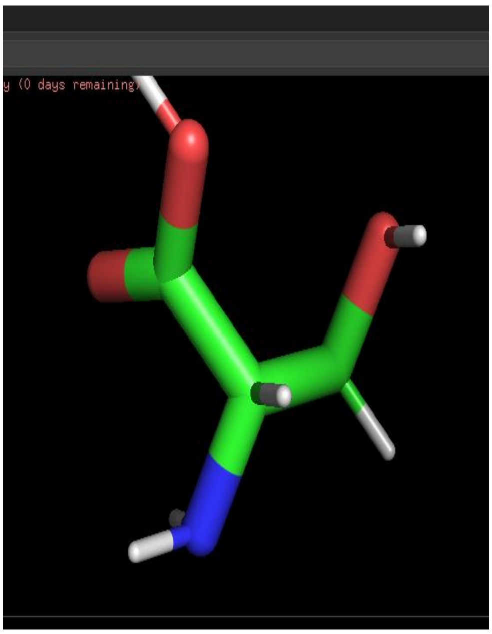


Fig-4.33 Ligand L-serine structure

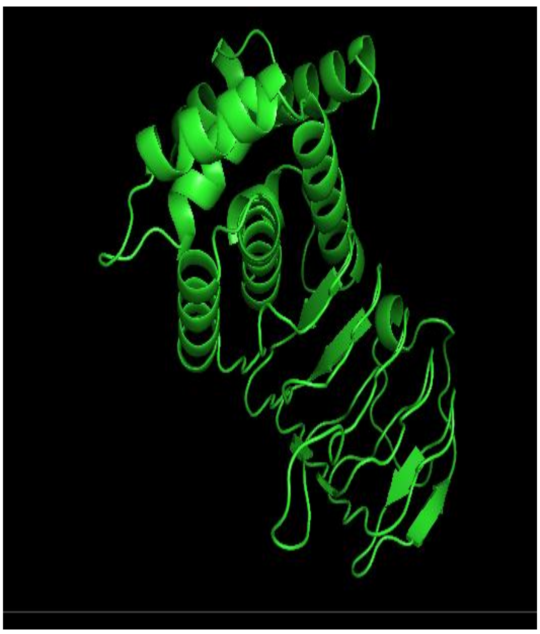
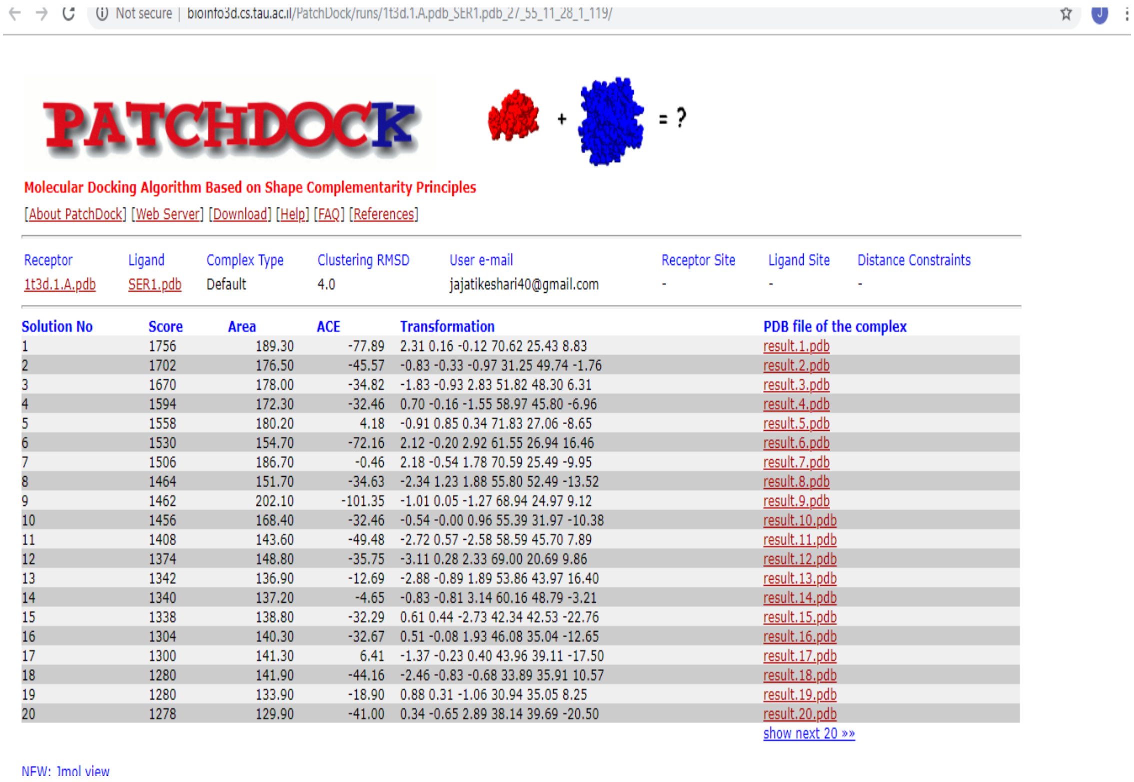


Fig-4.34 SAT Protein structure



Not secure | bioinfo3d.cs.tau.ac.il/PatchDock/runs/1t3d.1.A.pdb_SER1.pdb_27_55_11_28_1_119/

PATCHDOCK

Molecular Docking Algorithm Based on Shape Complementarity Principles

[About PatchDock] [Web Server] [Download] [Help] [FAQ] [References]

Receptor	Ligand	Complex Type	Clustering RMSD	User e-mail	Receptor Site	Ligand Site	Distance Constraints
1t3d.1.A.pdb	SER1.pdb	Default	4.0	jajatikeshari40@gmail.com	-	-	-

Solution No	Score	Area	ACE	Transformation	PDB file of the complex
1	1756	189.30	-77.89	2.31 0.16 -0.12 70.62 25.43 8.83	result.1.pdb
2	1702	176.50	-45.57	-0.83 -0.33 -0.97 31.25 49.74 -1.76	result.2.pdb
3	1670	178.00	-34.82	-1.83 -0.93 2.83 51.82 48.30 6.31	result.3.pdb
4	1594	172.30	-32.46	0.70 -0.16 -1.55 58.97 45.80 -6.96	result.4.pdb
5	1558	180.20	4.18	-0.91 0.85 0.34 71.83 27.06 -8.65	result.5.pdb
6	1530	154.70	-72.16	2.12 -0.20 2.92 61.55 26.94 16.46	result.6.pdb
7	1506	186.70	-0.46	2.18 -0.54 1.78 70.59 25.49 -9.95	result.7.pdb
8	1464	151.70	-34.63	-2.34 1.23 1.88 55.80 52.49 -13.52	result.8.pdb
9	1462	202.10	-101.35	-1.01 0.05 -1.27 68.94 24.97 9.12	result.9.pdb
10	1456	168.40	-32.46	-0.54 -0.00 0.96 55.39 31.97 -10.38	result.10.pdb
11	1408	143.60	-49.48	-2.72 0.57 -2.58 58.59 45.70 7.89	result.11.pdb
12	1374	148.80	-35.75	-3.11 0.28 2.33 69.00 20.69 9.86	result.12.pdb
13	1342	136.90	-12.69	-2.88 -0.89 1.89 53.86 43.97 16.40	result.13.pdb
14	1340	137.20	-4.65	-0.83 -0.81 3.14 60.16 48.79 -3.21	result.14.pdb
15	1338	138.80	-32.29	0.61 0.44 -2.73 42.34 42.53 -22.76	result.15.pdb
16	1304	140.30	-32.67	0.51 -0.08 1.93 46.08 35.04 -12.65	result.16.pdb
17	1300	141.30	6.41	-1.37 -0.23 0.40 43.96 39.11 -17.50	result.17.pdb
18	1280	141.90	-44.16	-2.46 -0.83 -0.68 33.89 35.91 10.57	result.18.pdb
19	1280	133.90	-18.90	0.88 0.31 -1.06 30.94 35.05 8.25	result.19.pdb
20	1278	129.90	-41.00	0.34 -0.65 2.89 38.14 39.69 -20.50	result.20.pdb

show next 20 »»

NEW: Jmol view

(Fig 4.35 - Results of SAT and L-serine obtained from Patchdock)

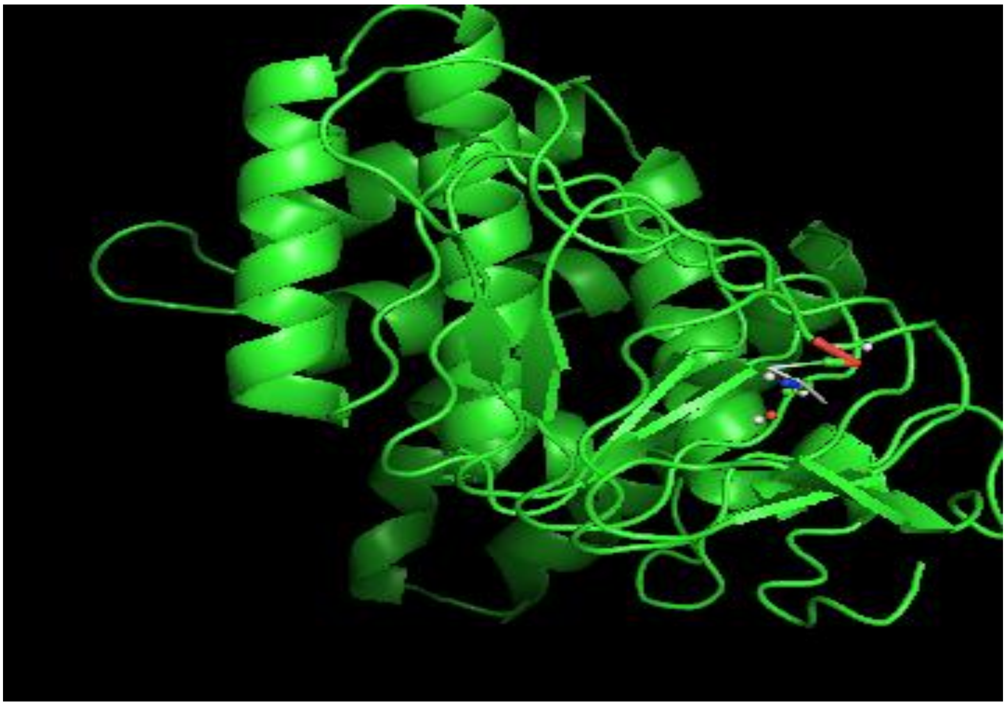


Fig-4.36 Liagnd and protein interaction

This docking results showed that enzyme serine –o scetyl transferase has a single binding site.This results is strongly correlates with our gel image which has single band. This docking results also correlated with the ODAP biosynthesis pathway (Malathi et al., 1970); sulphur and nitrogen metabolism pathway (Ikegami, F et al.,1991) and cyanide detoxification pathway (Machingura, M.,2016)

4.7 HOMOLOGY MODELLING

Homology modelling or comparative modelling or knowledge based modelling used for construction of 3D structure. It uses the pre determined 3D structure of different proteins of same family as a reference.On the basis of sequence alignment, output comes. It is not possible to determine the 3D structure of every protein experimentally. So homology modelling is a best way for this purpose.

3D structure gives important information about biochemical function and interaction with different molecules at a molecular level.If we calculate the ratio of known structure protein to total available protein, then it will be surely <0.01.Homology modelling consists 4 main steps

i) The proteins which are evolutionary related to the experimental related proteins were identified and used as template.

ii) The candidate proteins were mapped with template protein by means of sequence alignment and manual adjustment.

iii) 3D structure of candidate proteins werw built with help of sequence alignment.

iv) Quality of final outcome result was evaluated.(Kopp, J. & Schwede 2004)

For the homology modelling SWISS-MODEL workspace (https://swissmodel.expasy.org/interactive) and modebase online data server were used. (https://modbase.compbio.ucsf.edu/modweb/) . The protein structure of approx. 400 candidate genes were built.. From the homology modelling reults of 400 sequences we screened two most appropriate transcripts of candidate alalline synthase and serine o-acetyl transferase.For the modelling purpose we convert our transcripts into amino acid through Expassy translational tool. (https://web.expasy.org/translate/). After getting the amino acid we built the 3D structure of protein.

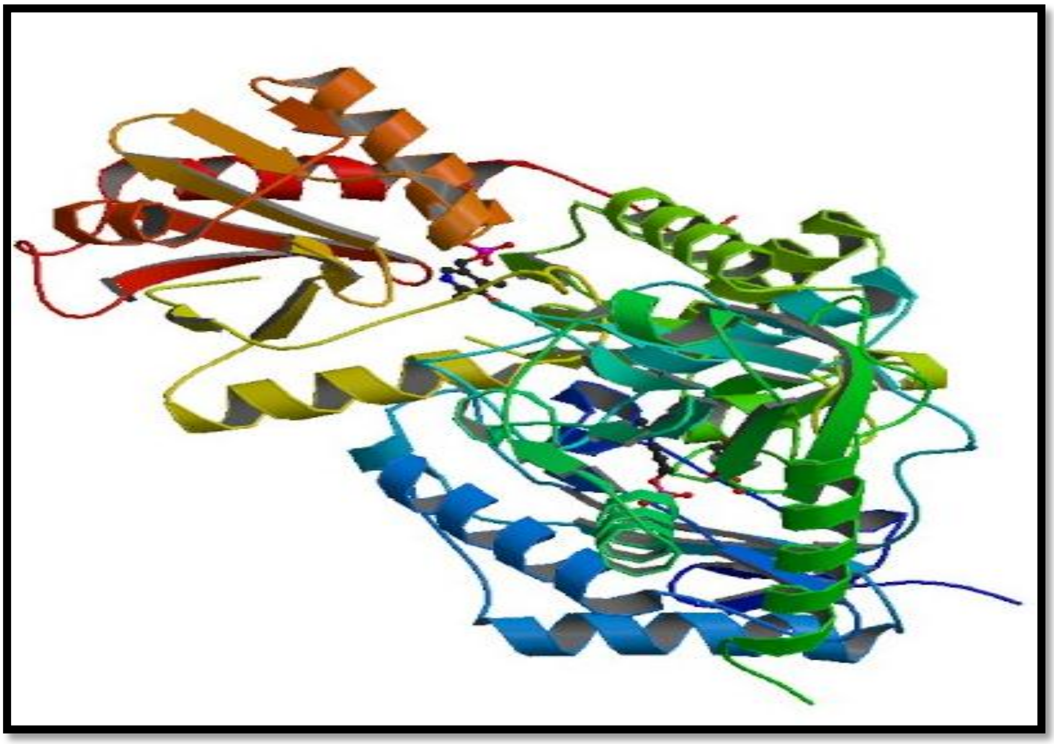
4.7.1 Results of candidate gene l-alaline synthase

Template	Seq Identity	Oligo-state	QSQE	Found by	Method	Resolution	Seq Similarity	Coverage	Description
5e25.2.A	23.68	homo-dimer	0.55	HHblits	X-ray	2.20Å	0.32	0.92	branched-chain aminotransferase
1a0g.1.A	23.32	homo-dimer	0.56	HHblits	X-ray	2.00Å	0.30	0.90	D-AMINO ACID AMINOTRANSFERASE
5e25.1.A	23.68	homo-dimer	0.54	HHblits	X-ray	2.20Å	0.32	0.92	branched-chain aminotransferase
2daa.1.A	24.22	homo-dimer	0.56	HHblits	X-ray	2.10Å	0.31	0.90	D-AMINO ACID AMINOTRANSFERASE
3lqs.1.A	23.98	homo-dimer	0.55	HHblits	X-ray	1.90Å	0.31	0.89	D-alanine aminotransferase
6fte.1.A	24.89	homo-dimer	0.50	HHblits	X-ray	1.52Å	0.32	0.94	Amine transaminase (fold IV)
1g2w.1.A	23.32	homo-dimer	0.53	HHblits	X-ray	2.00Å	0.30	0.90	D-ALANINE AMINOTRANSFERASE
3wwi.1.A	21.03	homo-dimer	0.51	HHblits	X-ray	2.27Å	0.31	0.94	(R)-amine transaminase
4cmd.1.A	21.03	homo-dimer	0.51	HHblits	X-ray	1.68Å	0.31	0.94	AMINOTRANSFERASE
5daa.1.A	24.20	homo-dimer	0.51	HHblits	X-ray	2.90Å	0.31	0.88	D-AMINO ACID AMINOTRANSFERASE
4uug.1.A	21.46	homo-dimer	0.48	HHblits	X-ray	1.60Å	0.32	0.94	AMINE TRANSAMINASE
4uug.1.B	21.46	homo-dimer	0.48	HHblits	X-ray	1.60Å	0.32	0.94	AMINE TRANSAMINASE

(Fig-4.37 Reults of Swiss-model workspace)

From the swiss-model results different template proteins weres found which showed similarity with our candidate proteins. Among different template, we chose template 5e25(**PDB ID-5E25**),because it has higher similarity and sequence identity with our candidate genes.

For docking simulation(discussed in section 4.3) This template gave two binding site in docking interaction with our provided ligand.In the above result template 5e25 gave more similarity, sequence identity and more coverage than other template.(refer fig 4.40)



(Fig 4. 38 3D view of template 5e25 source Boyko, K.M., Nikolaeva, A.Y. *et al*)

4.7.2 Results of candidate gene SAT

From swiss model results, we found different template viz; 1t3d,3gvd,4h7o,4hzo,1ssq,4hzc,4hzd,3mc4. We took template 1t3d(PDB ID-1T3D) for docking analysis with ligand L-Serine.1t3d has higher sequence similarity, sequence identity and more coverage with our candidate genes.

File Edit View Document Tools Window Help

4/20/2019 Untitled Project | Report

Template	Seq Identity	Oligo-state	QSQE	Found by	Method	Resolution	Seq Similarity	Coverage	Description
1t3d.1.A	48.83	homo-hexamer	0.47	HHblits	X-ray	2.20Å	0.43	0.77	Serine acetyltransferase
3gvd.1.A	48.03	homo-hexamer	0.44	HHblits	X-ray	2.40Å	0.43	0.77	Serine acetyltransferase
4h7o.1.A	45.35	homo-hexamer	0.44	HHblits	X-ray	2.17Å	0.43	0.78	Serine acetyltransferase
4hzc.1.A	47.64	homo-hexamer	0.46	BLAST	X-ray	1.97Å	0.43	0.77	CysE, serine acetyltransferase
4h7o.1.A	46.80	homo-hexamer	0.44	BLAST	X-ray	2.17Å	0.43	0.76	Serine acetyltransferase
1ssq.1.A	45.28	homo-hexamer	0.46	HHblits	X-ray	1.85Å	0.42	0.77	Serine acetyltransferase
4hzc.1.A	46.54	homo-hexamer	0.46	HHblits	X-ray	1.97Å	0.42	0.79	CysE, serine acetyltransferase
4hzd.1.A	47.64	homo-hexamer	0.44	BLAST	X-ray	1.87Å	0.43	0.77	CysE, serine acetyltransferase
3mc4.1.A	47.64	homo-trimer	0.45	BLAST	X-ray	1.95Å	0.43	0.77	WW/Rsp5/WWP domain:Bacterial transferase hexapeptide repeat:Serine O-acetyltransferase
4hzd.1.A	46.54	homo-hexamer	0.44	HHblits	X-ray	1.87Å	0.42	0.79	CysE, serine acetyltransferase
1ssm.1.A	46.86	homo-hexamer	0.44	HHblits	X-ray	2.15Å	0.43	0.72	Serine acetyltransferase

Fig4.39 Swiss-model result of SAT

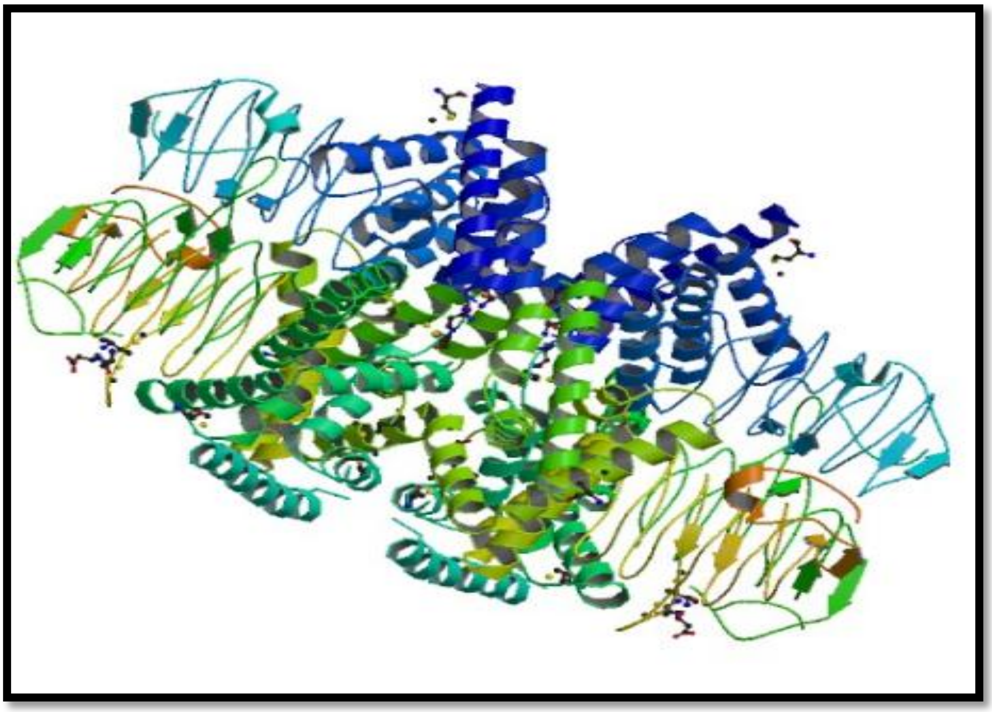


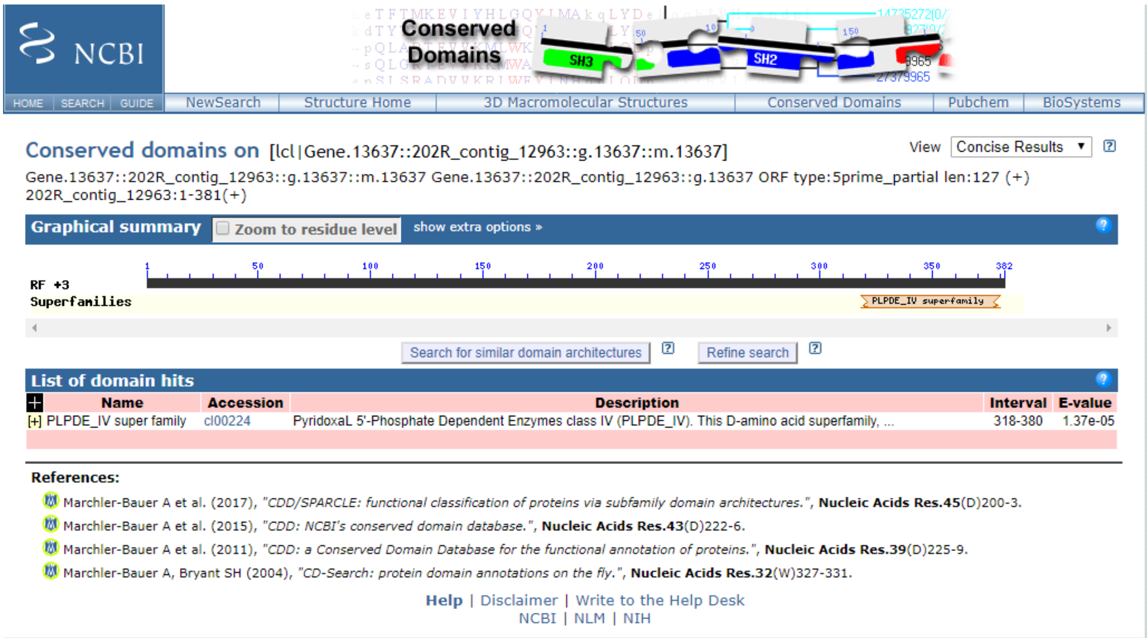
Fig-4.40 3D View of 1T3D source author Pye, V.E., Tingey

4.8 CD(CONSERVED DOMAIN) SEARCH

Conserve domain data base is protein annotation repository. They utilises position specific score matrix for rapid identification of related protein family via RPS-BLAST(Reverse Position-Specific BLAST) and CDD content.It gives a related idea about similar protein family according to querry sequence.

We did the CD serach(https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) of our candidate gene to determine wheather our candidate gene family has similarity with genes of our interest related to ODAP biosynthesis family. From CD search we found that our candidate gene l-alaline synthase has PLPD_IV family .PLPD protein family are pyridoxal-5'-phosphate dependent enzymes. Pyridoxal phosphate (PLP; a vitamin B6 derivative) act as important cofactor in many biological and biochemical pathways.(John, 1995; Jansonius, 1998; Mehta & Christen, 2000; Schneider *et al.*, 2000; Christen & Mehta, 200). Strisovsky *et al.* (2003) recently has proposed that deamination of serine in serine racimization traction is PLP dependant. Han *et al.* (2001 also described that aminotransferse reaction is also PLP dependant. cysteine-*S*-conjugate β-lyase (EC 4.4.1.13) related in amino acid metabolism is also a PLP dependent.(Cooper *et al.*, 2002). (Graveley, 2001 and (Graveley, 2001 also proved that some splice variant of certain enzyme dependence upon PLP domain for their catalytic activity.

4.8.1 CD result of l-alaline synthase



(Fig-4.41 CD result of candidate gene l-alaline synthase)

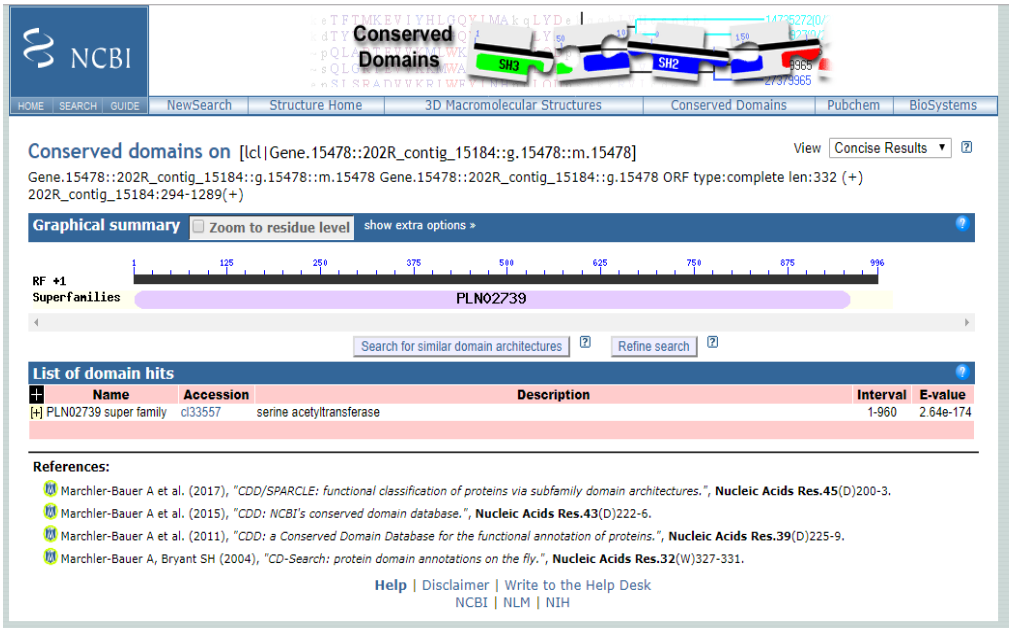
From the above result we saw that l-alaine synthase has PLP domain. From above section we knew that PLP domain has transaminase, taminotransferase and amino acid metabolism activity. So most probably our candidate gene may have some role in ODAP biosynthesis pathway.

4.8.2 CD result of candidate gene SAT

From CD search result we found that our candidate gene SAT comes under superfamily PLN-02739. This superfamily contains serine-o-acetyl transferase.(https://www.ncbi.nlm.nih.gov/proteinclusters?term=PLN02739).From CD search we can sure that our candidate gene is our gene of our interest.

(Table 4.5 KEGG result of candidate gene SAT)

EC Number:	**2.3.1.30**
KEGG KO:	**K00640**
InterPro:	**IPR005881; IPR010493; IPR011004; IPR018357; IPR001451**



(Fig-4.42 CD search of candidate gene SAT)

4.9 Sub cellular localization of our candidate genes

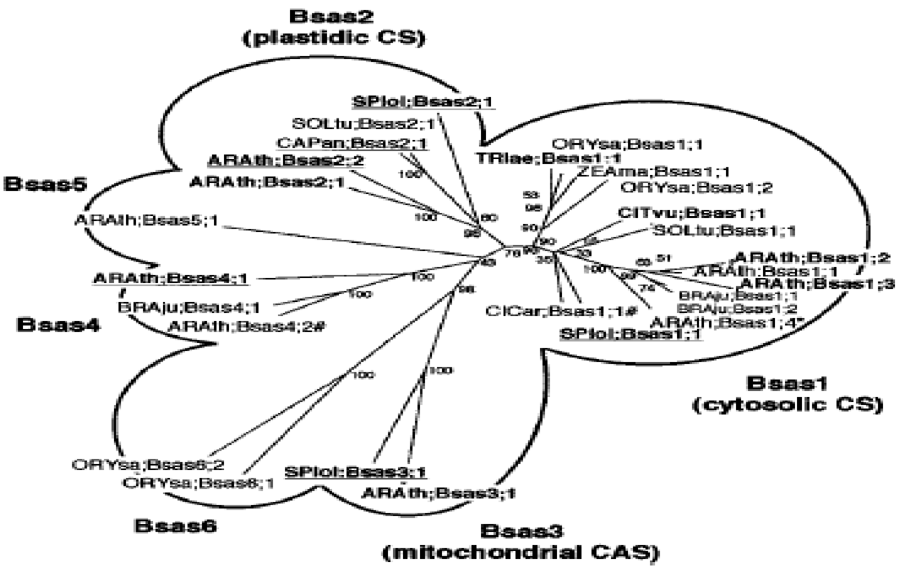
Blumenthal et al., 1968 had proved that b-Cyano-Ala synthase (CAS; EC 4.4.1.9) helps in the formation of cyanide,3-cyano-L-alaline and sulphide.

$$HSCH_2CH(NH_2)COOH + HCN \longrightarrow CNCH_2CH(NH_2)COOH + H_2S$$

L-Cys Cyanide 3-cyano-L-alaline Sulphide

Dunnill and Fowden, 1965 *et al.*, had detected this type of mechanism in different species of bacteria. Floss *et al.*, 1965; Hendrickson proved that above mechanism also take part in plants. Meyers and Ahmad, 1991 found that b-Cyano-Ala synthase activity found in mitochondrial matrix,main target for cyanide toxicity. Masada *et al.*, 1975; Droux *et al.*, 1992 proved that CAS and CS are homodimer and both are depend upon pyridoxal 5-phosphate. Ikegami and Murakoshi, 1994 gave an idea that both CS an CAS comes under same family infact CAS shows some extent of similar activity with CS.(Hendrickson and Conn, 1969; Ikegami *et al.*, 1988a, 1988b; Maruyama *et al.*, 1998). Warrilow and

Hawkesford, 1998 found that cyanoalaline synthase present both in chloroplast and mitochondria.



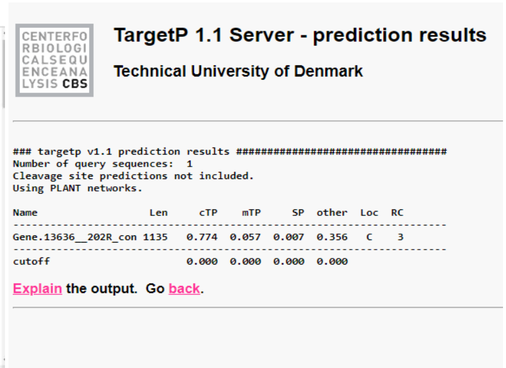
(4.43 phylogenetic tree was constructed by Yves Hatzfeld, Akiko Maruyama *et al.,)*

From the above demonstration of different scientist and phylogentic trees we got to know that cyanoalaline synthase(CAS; EC 4.4.1.9) found in Mitochondria and chloroplast and has a great role cyanide detoxification and Cysteine synthase activity.

From sub-cellular localization search of our Cndidate gene we also got that our candidate gene also found in mitochondria and chloroplast.From this result we can say that our candidate gene has some activity in ODAP biosynthesis pathway.

4.9.1 Sub-cellular search of candidate gene L-alaline synthase

(Fig-4.44 WolFpsort result of Alaline synthase)



(Fig-4.45 Target P results L-alaline synthase)

Sub cellular search was done through two online server. TargetP 1.1 Server(http://www.cbs.dtu.dk/services/TargetP/) and WoLF PSORT: Advanced Protein Subcellular Localization Prediction Tool(https://www.genscript.com/wolf-psort.html)

4.9.2 Sub-cellular search of candidate gene serine o-acetyl transferase

From subcellular search of our candidate gene we also knew that like l-alaline synthase, SAT also found in mitochondria and chloroplast.So from both result we can conclude that both candidate gene SAT and L-alaline synthase found in chloroplast and mitochondria and takes part in ODAP biosynthetic path

15513342172009 WoLFPSORT output

15513342172009 WoLFPSORT prediction chlo_mito: 6, chlo: 5.5, mito: 5.5, cyto: 1, vacu: 1

traditional PSORTII prediction

14 Nearest Neighbors

id	site	distance	identity	comments
DAPA_SOYBN	chlo	185.78	10.6825%	[Uniprot] SWISS-PROT45:Chloroplast.
DHE2_ARATH	mito	194.212	15.8151%	[Uniprot] SWISS-PROT45:Mitochondrial.
ARGD_ALNGL	mito	222.674	17.0354%	[Uniprot] SWISS-PROT45:Mitochondrial.
MDHG_CITLA	pero	226.728	14.4847%	[Uniprot] SWISS-PROT45:Glyoxysomal.
TIPA_PHAVU	vacu	229.631	14.5015%	[Uniprot] SWISS-PROT45:Integral membrane protein; tonoplast membrane.
SODM_ARATH	mito	230.444	14.1994%	[Uniprot] SWISS-PROT45:Mitochondrial matrix.
LEU3_SOLTU	chlo	233.476	14.2458%	[Uniprot] SWISS-PROT45:Chloroplast.
FDH_SOLTU	mito	237.334	13.6483%	[Uniprot] SWISS-PROT45:Mitochondrial.
ENO_LYCES	cyto	239.638	13.964%	[Uniprot] SWISS-PROT45:Cytoplasmic.
G3PA_CHLRE	chlo	245.209	12.5668%	[Uniprot] SWISS-PROT45:Chloroplast.
MDHP_MESCR	chlo	249.397	14.2857%	[Uniprot] SWISS-PROT45:Chloroplast.
At1g04640.1	mito	249.581	14.8036%	[Arath]
At5g14780.1	chlo_mito	249.939	12.2396%	[Arath]
DAP2_WHEAT	chlo	252.409	14.8541%	[Uniprot] SWISS-PROT45:Chloroplast

Fig 4.46-WolF-Psort result of Candidate gene SAT)

4.10 Pathway mapping through KEGG and BLAST2GO software

It is the last bioinformatocal analysis for our candidaye genes to make sure wheatther our candidate genes are related in ODAP biosynthesis pathway.We did pathwat mapping through NCBI biosystem data base (https://www.ncbi.nlm.nih.gov/Structure/biosystems/docs/biosystems_about.html) and KEGG data base(https://www.genome.jp/kegg/). KEGG (Kyoto Encyclopedia of Genes and Genomes) is a online repository of different biochemical and physiological reaction which are determined by genomic sequencing and different high throughput technologies.It gives a better idea about inter connection between different molecules and chemical compounds.(Kiyoko F. Aoki-Kinoshita *et al*.)

Through NCBI biosystem data base, we can directly retrieve the genes and their related function by using gene name or gene ID.Moreover this data base helps in gene annotation.

We did the pathway mapping of our two candidate genes. Reults are shown below.

4.10.1 Pathway mapping of candidate gene serine-o-acetyl transferase

From the pathway mapping of our candidate gene we got to know that our candidate gene involves in 3 different pathway.

i) Carbon metabolism
ii) Biosynthesis of different amino acid
iii) Cysteine biosynthesis pathway

.

i) Carbon metabolism:

Carbon metabolism is the most basisc and fundamental part of our life.This map consists of overall central carbon metabolism.Main pathways involved in this map are glycolysis,citric acid cycle and pentose phosphate pathway.This pathway also contains 1) reductive pentose phosphate cycle (Calvin cycle) in plants and cyanobacteria that perform oxygenic photosynthesis, (2) reductive citrate cycle in photosynthetic green sulfur bacteria and some chemolithoautotrophs, (3) 3-hydroxypropionate bi-cycle in photosynthetic green nonsulfur bacteria, two variants of 4-hydroxybutyrate pathways in Crenarchaeota called (4) hydroxypropionate-hydroxybutyrate cycle and (5) dicarboxylate-hydroxybutyrate cycle, and (6) reductive acetyl-CoA pathway in methanogenic bacteria.

We can also view this result through https://www.ncbi.nlm.nih.gov/biosystems/816666?Sel=geneid:11409367#show=genes

ii)Biosynthesis different amino acid

This pathway consists of arichitecture of 20 different amino acids which are essential as well non essential.Core part of this pathway is the conversion of glyceraldehydes-3-phosphate to pyruvate.Glycine and serine biosynthesis also play amajor role in this pathway.Details are provided through this link.(https://www.kegg.jp/pathway/mtr01230). mtr01230 is KEGG ID of this result

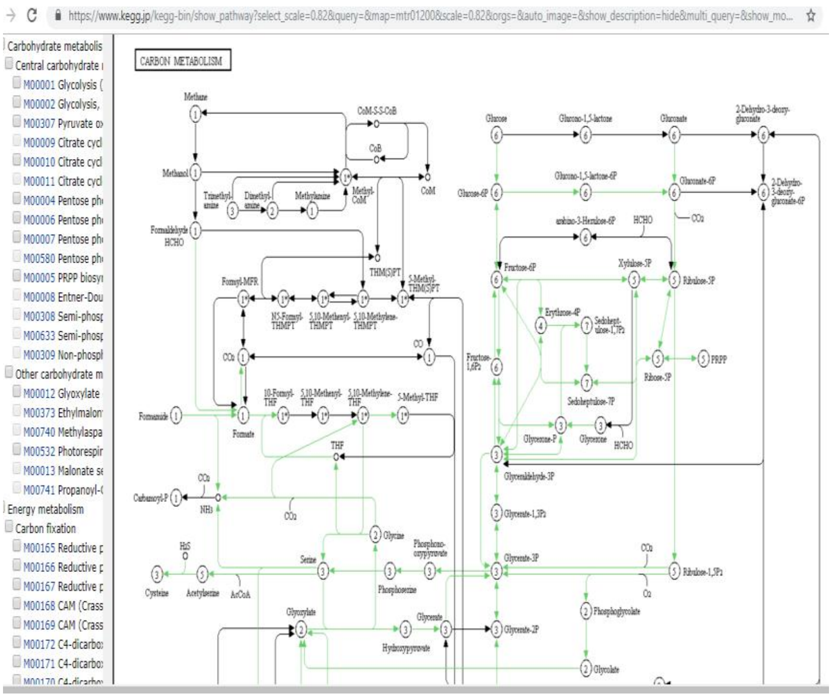


Fig-4.47 Pathway mapping of Carbon metabolisim

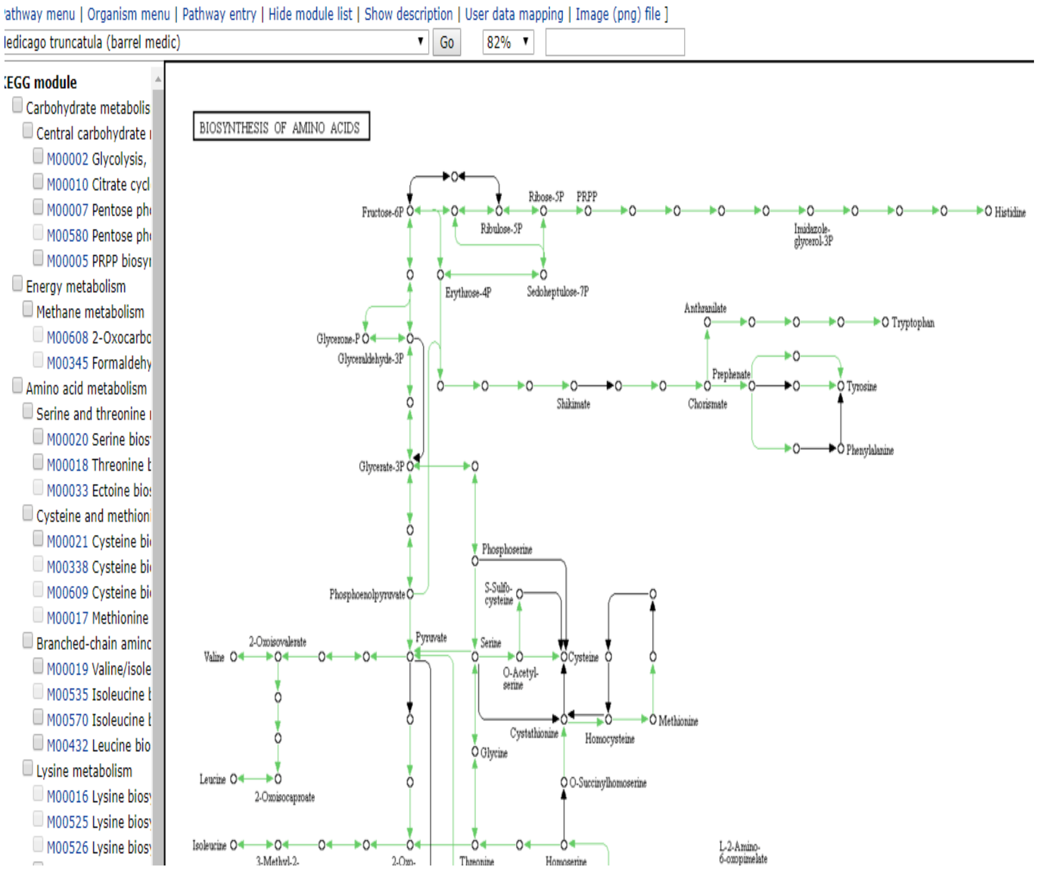


Fig-4.48 Pathway mapping og Amino acid biosynthesis.

iii) Cysteine Biosynthetic pathway:

Cysteine and methionine are sulfur-containing amino acids. Cysteine is synthesized from serine through different pathways in different organism groups. In bacteria and plants, cysteine is converted from serine (via acetylserine) by transfer of hydrogen sulfide [MD:M00021]. In animals, methionine-derived homocysteine is used as sulfur source and its condensation product with serine (cystathionine) is converted to cysteine [MD:M00338]. Cysteine is metabolized to pyruvate in multiple routes. Methionine is an essential amino acid, which animals cannot synthesize. In bacteria and plants, methionine is synthesized from aspartate [MD:M00017]. S-Adenosylmethionine (SAM), synthesized from methionine and ATP, is a methyl group donor in many important transfer reactions including DNA methylation for regulation of gene expression. SAM may also be used to regenerate methionine in the methionine salvage pathway.

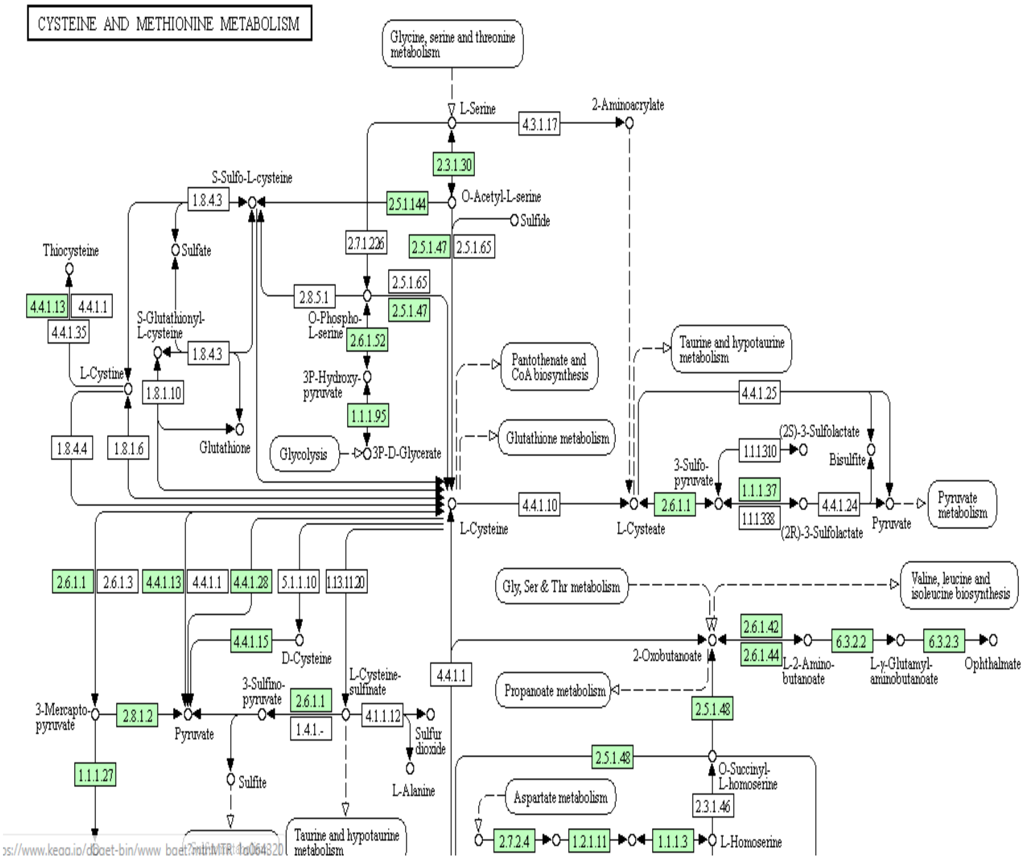


Fig-4.52 Pathway mapping of Cysteine and Methionine Metabolism

From the above results we got a breakthrough. As our candidate gene involved in amino acid metabolism and cysteine metabolism, here it is clearly proved that our candidate gene serine o-acetyl transferase also involved in ODAP biosynthesis pathway. Because ODAP biosynthesis also linked with cysteine metabolism and amino acid metabolism.(Ikegami, F.; Itagaki, S.; Kamiya 1996; Machingura, M.; Salomon.,et al; Malathi, K., Padmanaban et al.)

4.7.2 Pathway mapping of candidate gene L-alaline synthase

Pathway mapping of this candidate gene shows that it takes part in folate metabolism pathway.It has from KEGG source record: mtr00790

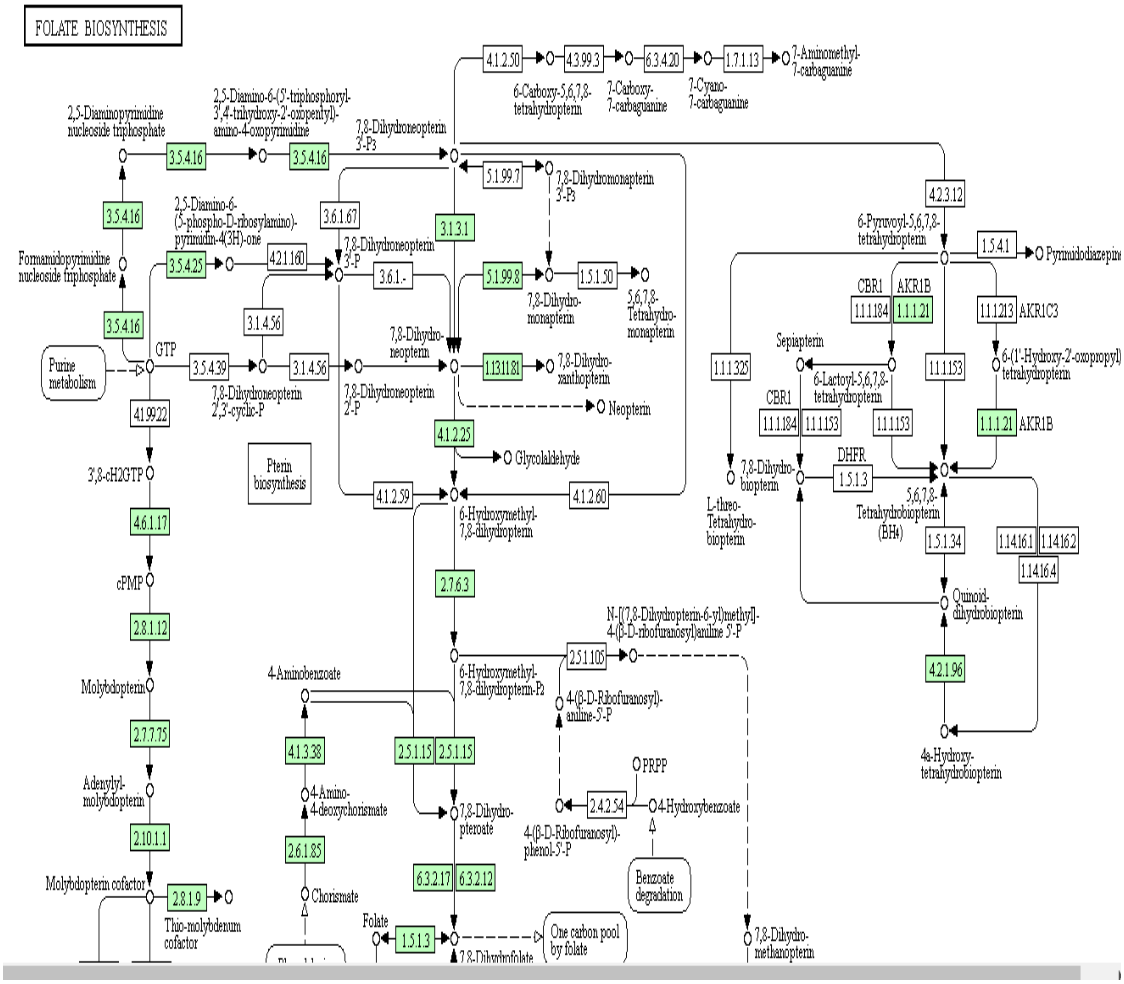


Fig-4.53 Pathway mapping of folate biosynthesis.

CHAPTER – V
SUMMARY AND CONCLUSION

5.1 SUMMARY

Pulse has an important role in idian agriculture.In India, pulse is grown in 23.8mha with a production of 18.6mt. Unlike other pulses, Lathyrus doesn't require much care and maintenance. It is comparatively tolerant towards drought, heat, cold, insect and pest than other pulse crops. Being a short duration hardy pulse, it is an alternative crop for mixed farming and intercropping.

Although it has many advantages, prolonged consumption Lathyrus is known to cause neurolathyrisim, a motor degenerative disease which is caused by β-ODAP metabolism in human system.This prevents cultivation of Lathyrus as major crop.Therefore,it is inevitable to reduce β-ODAP conrent in Lathyrus, that could be achieved by manipulating genes involved in the β-ODAP biosynthesis pathway, which is interlinked with sulphur metabolism, cysteine metabolisim and other amino acid metabolisim.

In our current investigation we are trying to find out the genes involved in the β-ODAP biosynthesis pathway. In the process of finding genes, the approaches such as BLAST and LOCAL BLAST have been followed. The primers viz. PaaY1, PLPD1, UNK1, UNK2, PLN1, PLN2, DeQ and AMIF were designed from the putative transcripts of Mahateora and RLK-1950 transcriptomes as their translated protein sequences are similar to putative genes such as SAT and Alanine synthase, which involved in ODAP biosynthesis pathway. After amplification of these primers with the cDNA of 11 Lathyrus genotypes( ), it was found that amplification of PaaY1a and PLPD1 correlated to ODAP biosynthesis pathway, as per Malathi et al.,1970, where PaaY1 amplified the putative genes SAT and PLPD1 amplified the L-alaline synthase.The amplified products of PLPD1,validated as L-alaline synthase has two bands, where upper band (250bp approx.) is positively correlated to high ODAP content Lathyrus genotypes while lower band (150bp approx.) is negatively correlated, showing that L-alaline synthase exists in two isoforms viz; β(isoxazolin-5on-2yl) L-alanine synthase(EC

2.5.1.118) and EC 2.5.1.119 β(isoxazolin-5-on-4yl) Lalanine synthase. And the amplified products of PaaY1, validated as SAT has single bands with aprrox. equall expression in all the 11 genotypes of *Larhyrus.* This proves that SAT is involved in ODAP biosynthesis pathway but not differentially expressed in high and low ODAP containing genotypes as its product O-acetyl-L-serine is common precursor of ODAP biosynthesis pathway,cysteine biosynthesis pathway and cyanide detoxification pathway. Therefore, our putative genes, amplified with the primers PaaY1 and PLPD1 were validated as SAT and L-alanine synthase.

5.2 CONCLUSION

- Esimation of β-ODAP content of 30 genotypes was done with the help of modified Rao *et al.*,2002 protocol. Among 30 genetoypes, 11 genotypes were screened on the basis of differential ODAP content, which was correlated with cDNA amplification.
- Two candidate genes, amplified with the primers PaaY1 and PLPD1 were validated as SAT and L-alanine synthase , which involved in the ODAP biosynthasis pathway, where the later was found to be correlated with the ODAP content. with respect to ODAP content.
- The candidate genes were confirmed as SAT and L-alanine synthase using bioinformatic tools such as conserve domain search, homology modelling of protein, molecular docking.

5.3 SUGGESTION FOR FUTURE WORK

- Full length gene of this candidate gene could be formed using RACE-PCR
- Then, the full length coyld be used for cloning and stored as gene library for future use.
- Further these genes could be used foe validation with population.

REFERENCES

Aziz, A., and Larher, F. 1996. Changes in polyamine titers associated with the proline response and osmotic adjustment of rape leaf discs submitted to osmotic stress.Plant Sci., 112: 175–186.

Kaul, A.K., Islam, M.Q., and A, Hamid. 1986. Screening of Lathyrus germplasm of Bangladesh for BOAA content and some agronomic characters, in: Kaul, A.K., D. Combes (Eds.), Lathyrus and Lathyrism. Third World Medical Research Foundation., New York , pp. 130–141.

Galston, A.W., Kaur-Sawhney, R.1995 Polyamines as endogenous growth regulators, in: P.J. Davies (Ed.), Plant hormones: physiology, biochemistry and molecular biology, Kluwer Academic Publishers., Dordrecht, pp. 158–178.

Allkin, R., T.D, Macfarlance., R.J, White., F.A, Bisby and M.E. Adey. 1983. Names and synonyms of species and subspecies in the Vicieae. Issue 2: Vicieae Database Project Publication No. 2, Southampton.

Basaran, U., Acar, Z., Karancan, M., and Onar, A.N., 2013. Variation and Correlation of morpho-agronomic traits and biochemical contents (Protein and β-ODAP) in Turkis grass pea (Lathyrus sativus L.) landraces, 166-173.

Bell,E and O'Donovan, J.P.,Phytochemistry.1966.5: 1211-1219

Ressler, C. 1962. "Isolation and identification from common vetch of the neurotoxin β- cyano-l-alanine, a possible factor in neurolathyrism," TheJournal of Biological Chemistry, vol. 237: pp. 733– 735,

C.Z, Li., Wang, G.X. 2004 Interactions between reactive oxygen species, ethylene and polyamines in leaves of Glycyrrhiza inflata seedlings under root osmotic stress, Plant Growth Regul. 42 : 55–60.

Campbell, C.G. 1988. Improvement of Lathyrus sativus, in: P.S. Spencer (Ed.), The Grasspeas Threat and Promise. Proceedings of the International Network for Improvement of Lathyrus sativus and Erradication of Lathyrism (INILSEL) workshop, London, May 1988, Third World Medical Research Foundation, New York, pp. 139-146.

Christen, P., & Mehta, P.K., 2001 From cofactor to enzymes. The molecular evolution of pyridoxal-5′-phosphate-dependent enzymes. Chem. Rec., 1: 436–447

Christmas, P., *et al*. 2001. Alternative splicing determines the function of CYP4F3 by switching substrate specificity. J. Biol. Chem., 276, 38166–38172. [PubMed] [Google Scholar

Connolly,M.L.1983 Analytical molecular surface calculation. J. Appl. Crystallogr., 16, 548–558.

Connolly, M.L. 1983. Solvent-accessible surfaces of proteins and nucleic acids. Science., 221, 27 :709–713.

Cooper, A.J., Bruschi, S.A. & Anders, M.W. 2002. Toxic, halogenated cysteine S-conjugates and targeting of mitochondrial enzymes of energy metabolism. Biochem. Pharmacol., 64: 553–564.

Zhang, D.W., Xing, G.M., H, Xu., Yan, Z.Y., Wang, C.Y., Wang, Y.F., Z.X. Li.2005. Relationship between oxalic acid and the metabolism of β-N-oxalyl- L-a,β-diaminopropionic acid (ODAP) in grass pea (Lathyrus sativus L.). Israel, J. Plant Sci., 53.

Dixit, G. P. , Parihar, A.K. , Bohra, A. , and Singh, N. P., (2016) Achievements and prospects of grass pea (Lathyrus sativus L.) improvement for sustainable food production., ICAR-Indian Institute of Pulses Research (IIPR), Kanpur 208024, India., The crop journal A L 4: 407 – 416.

Duhovny,D., Nussinov,R. and Wolfson,H.J. (2002) Efficient unbound docking of rigid molecules. In Guigo,R. and Gusfield,D. (eds), Proceedings of the Fourth International Workshop on Algorithms in Bioinformatics. Springer-Verlag GmbH Rome, Italy, September 17–21, 2002, Vol. 2452, pp. 185–200

Duhovny,D., Nussinov,R. and Wolfson,H.J. 2002 Efficient unbound docking of rigid molecules. In Guigo,R. and Gusfield,D. (eds), Proceedings of the Fourth International Workshop on Algorithms in Bioinformatics. Springer-Verlag GmbH Rome, Italy, September 17–21. Vol. 2452: pp. 185–200.

Duke, J.A. 1981. Handbook of legumes of world economic importance. p. 199-265. Plenum Press, New York.

F. Lambein, R. Haque, J. K. Khan, N. Kebede, and Y.-H. Kuo, “From soil to brain: zinc deficiency increases the neurotoxicity of Lathyrus sativus and may affect the susceptibility for the motorneurone disease neurolathyrism,” Toxicon, vol. 32: no. 4, pp. 461–466.

Wang, F., Chen, X., Q. Chen, X. Qin, and Z. Li,1994. “Determination of neurotoxin 3-N-oxalyl-2,3-diaminopropionic acid and nonprotein amino acids in Lathyrus sativus by precolumn derivatization with 1-fluoro-2,4-

dinitrobenzene," Journal of Chromatography A, vol. 883, no. 1-2, pp. 113–118.

Anonymous. 1999.Genotype-environmentinteractionforseedyieldandODAP concentrationof Lathyrussativus L. and L. cicera L. in Mediterranean-typeenvironments,Euphytica110 :45–60.

Getahun, H., Lambei, F., Vanhoorne, F., Stuyft, M., van der, P. 2003. Food-aid cereals to reduce neurolathyrism related to grasspea preparations during famine. Lancet 362: 1808–1810.

Getahun, H., Lambein, F., Vanhoorne, F., Stuyft, M., Van der, P. 2005 Neurolathyrism risk depends on type of grass pea preparation and on mixing with cereals and antioxidants. Trop. Med. Int. Health., 10:169–178.

Graveley, B.R. 2001. Alternative splicing: increasing diversity in the proteomic world. Trends Genet., 17: 100–107

Liu, H.P., Dong, B.H., Zhang, Y.Y., Liu, Z.P., Liu, Y.L. 2004. Relationship between osmotic stress and the levels of free, conjugated and bound polyamines in leaves of wheat seedlings. Plant Sci., 166: 1261–1267.

Hammer K, Knupfer H, Laghetti G, Perrino P. 1992. Seeds from the past. A catalogue of crop germplasm in South Italy and Sicily. Germplasm Institute of CNR, Bari, p 173

Han, Q., Fang, J., & Li J. 2001. Kynurenine aminotransferase and glutamine transaminase K of Escherichia coli: identity with aspartate aminotransferase. Biochem. J., 360: 617–623.

Hanbury C, White C, Mullan BP, Siddique KHM. 2000. A review of the potential of Lathyrus sativus L. and L. cicera L. grain for use as animal feed. Lathyrus Lathyrism Newsletter 1, 34.

Hristen P. & Mehta P.K. (2001) From cofactor to enzymes. The molecular evolution of pyridoxal-5′-phosphate-dependent enzymes. Chem. Rec., 1, 436–447. [PubMed] [Google Scholar]

Ahlawat, I.P.S., Singh, A., and Saraf, C.S. 1981. Effects of winter legumes on the nitrogen economy and productivity of succeeding cereals. Exp. Agric. 17 :57–62.

Ikegami, F., Horiuchi, S., Kobori, M., Morishige, I., Murakoshi, I. 1992. Biosynthesis of neuroactive amino acids by cysteine synthases in Lathyrus latifolius. Phytochemistry . 31: 1991–1996.

Ikegami, F., Itagaki, S., Ishikawa, T., Ongena, G., Kuo, Y.H., Lambein, F., Murakoshi, I.1991. Biosynthesis of beta-(isoxazolin-5-on-2-yl)alanine, the precursor of the neurotoxic amino acid β-N-oxalyl-L-α,β_-diaminopropionic acid. Chem. Parm. Bull. 39: 3376–3377.

Ikegami, F., Itagaki, S., Kamiya, M., Kuo, Y.H., Lambein, F., Murakoshi, I.1996 Enzymatic synthesis of two isoxazolylalanine isomers by cysteine synthases in Lathyrus species. Biol. Pharm. Bull. 19: 1214–1215.

Ikegami, F., Ongena, G., Sakai, R., Itagaki, S., Kobori, M., Ishikawa, T., Kuo, Y.H., Lambein, F., Murakoshi, I.1993. Biosynthesis of β-(isoxazolin-5-on-2-yl)-alanine, the precursor of the neurotoxin β-N-oxalyl-L-α, β diaminopropionic acid, by cysteine synthase in Lathyrus sativus. Phytochemistry . 33: 93–98.

Ikegemi, F., Yamamoto, A., Kuo,Yu., and Lambei,F.1999. enzymatic foramation of β-N-oxalyl-L-α,βdiaminopropionic acid, the direct precursor, neurotoxin present in Lathyrus sativu. Biol pharm. Bull(22)7:770-771

Vardhan, K.S., Rudra, M.P.P., and Rao, S.L.N. 1997 .Inhibition of tyrosine aminotransferase by β-N-oxalyl-l-α, β- diaminopropionic acid, the Lathyrus sativus neurotoxin. Journal of Neurochemistry., vol. 68, no. 6, pp. 2477–2484, 1997.

Kaul, A.K., M.Q. Islam, and A. Hamid. 1986. Screening of Lathyrus germplasm of Bangladesh for BOAA content and some agronomic characters. Pp. 130-141 in Lathyrus and Lathyrism (A.K. Kaul and D. Combes, eds.). Third World Medical Research Foundation, New York.

Kour, J.2016. Correlation analysis in advanced breeding lines of grasspea for yield and its attributing character with neurotoxin ODAP content. Asian Science11(1): 60-64

Kumar, S., Bejija, G., Ahmed, S., Nakkoul, H., Sarker, A.2011 Genetic improvement of grasspea for low neurotoxin(β_-ODAP) content. Food Chem. Toxicol. 49: 589–600.

Kuo, Y.H., Khan, J.K., Lambein F. 1994. Biosynthesis of the neurotoxin β-ODAP in developing seeds of Lathyrus sativus. Phytochem., 35: 911-913.

Kuo, Y.H.; Ikegami, F.; Lambein, F.1998. Metabolic routes of β-(isoxazolin-5-on-2-yl)-l-alanine (BIA), the precursor of the neurotoxin ODAP (_β-N-oxalyl-L-α,_β-diaminopropionic acid), in different legume seedlings. Phytochemistry., 49: 43–48.

L.Lioi,F.Sparvoli,G.Sonnante,G.Laghetti,F.Lupo,M.Zaccardelli.2011.Characterization of Italiangrasspea(LathyrussativusL.) germplasm using agronomic traits, biochemical and molecular markers,Genet.Resour.Crop.Evol.58: 425–437.

Lambein, F., Haque, R., Khan, J.K., Kebede, N., Kuo, Y.H.1994. From soil to brain: Zinc deficiency increases the neurotoxicity of Lathyrus sativus and may affect the susceptibility for the motorneurone disease neurolathyrism. Toxicon., 32: 461–466.

Long, Y.C., Ye, Y.H., Xing, Q.Y.1996. Studies on the neuroexcitotoxin β-N-oxalyl-L-α,β-diaminopropionic acid and its isomer β-N-oxalo-L-α,β-diaminopropionic acid from the root of Panax species. Int. J. Peptide Protein Res., 47: 42–46.

Lu, F.H., Bao, X.G., Liu, S.Z., 1990. A study of fivevein vetchling (Lathyrus L.) germplasm resources. Crop Genet. Res. (China) 3: 17–19.

Ludolph, A.C., Hugon, J., Dwivedi, M.P., Schaumburg, H.H., Spencer, P.S., 1987. Studies on the aetiology and pathogenesis of motor neuron disease. 1. Lathyrism: clinical findings in established cases. Brain 110, 149–165.

Ludolph, A.C., Spencer, P.S., 1996. Toxic models of upper motor neuron disease. J. Neurol. Sci. 139, 53–59.

M.C. PattoVaz,D.Rubiales.2014 Lathyrus diversity: available resources with relevan cetocrop improvement — L. sativus and L. cicera as casestudies,Ann.Bot.London113: 895–908. C.D. Hanbury,K.H.M.Siddique,N.W.Galwey,P.S.Cocks,

Machingura, M., Salomon, E., Jez, J.M., Ebbs, S.D.2016 The _β-cyanoalanine synthase pathway: Beyond cyanide detoxification. Plant Cell Environ. 39: 2329–2341

Malathi, K., Padmanaban, G., Rao, S.L.N., Sarma.1967 P.S. Studies on the biosynthesis of β-N-oxalyl-L-α,_β diaminopropionic acid, the Lathyrus sativus neurotoxin. Biochim. Biophys. Acta . 141: 71–78.

Malathi, K., Padmanaban, G.; Sarma, P.S. 1970. Biosynthesis of β_-N-oxalyl-L-α_,β_-diamino-propionic acid, the Lathyrus sativus neurotoxin. Phytochemistry ., 9:1603–1610

Meenakshi, N. H., (2016) Identification of DNA markers (EST SSR's, SSR's and ISSR's) associated to grain nutritive traits in Lathyrus sativus, IGKV Raipur (C.G.).

Mehra, R.B., D.B. Raju and K. Himabindu. 1995. Evaluation and utilization of Lathyrus sativus collection in India. Pp. 37-43 in Lathyrus Genetic Resources in Asia. Proceedings of a Regional Workshop, 27-29 December 1995, Indira Gandhi Agricultural University, Raipur, India (R.K. Arora, P.N. Mathur, K.W. Riley and Y. Adham, eds.). IPGRI Office for South Asia, New Delhi, India.

Mishra,B.K., Panda, P.C. and Barat,G.K.1979. Plant biochem physiology.Pflanzen.

N. Bagni, P. Torrigiani.1992. Polyamines: a new class of growth substances, in: C.M. Karssen, L.C. van Loon, D. Vreugdenhil (Eds.), Progress in plant growth regulation, Kluwer Academic Publishers, Dordrecht., pp. 264–275.

Palmer, V.S., A.K. Kaul and P.S. Spencer. 1989. International Network for the Improvement of Lathyrus sativus and the Eradication of Lathyrism (INILSEL): A TWMRF initiative. Pp. 219-223 in The Grass Pea: Threat and Promise. Proc. of the International Network for the Improvement of Lathyrus sativus and the Eradication of Lathyrism (P. Spencer, ed.). Third World Medical Research Foundation, New York.

Pan, M.D., Mabry, T.J., Cao, P., Moini, M.1997. Identification of nonprotein amino acids from cycad seeds as N-ethoxycarbonyl ethyl ester derivatives by positive chemical-ionization gas chromatography-mass spectrometry. J. Chromatogr. A 787: 288–294.

Quereshi, M.Y., Pilbeam, D.J., Evans, C.S., Bell, E.A.1977. The neurolathyrogen, _-amino-_-oxalylaminopropionic acid in legume seeds. Phytochemistry., 16: 477–479.

R.L. Ambade, S.K. Verma, H.C. Nanda, S.K. Nair and S.B. Verulkar.2015. Genetic diversity based on molecular markers in Grasspea (Lathyrus sativus L.) Agricultural research communication centre ., 43-46

R.L. Pandey, P. Shrivastava, A.K. Geda, R.N. Sharma.2000. Relative contribution of yield components and their relationship with neurotoxin content in grass pea (Lathyrus sativus L.). Ann. Agric. Res. 21: 11–16.

Ressler,C.,J. 1962. Biol chem.,237:733-735

Ross, S.M., Roy, D.N., and Spencer.1989 P.S. "β-N-oxalylamino-lalanine action on glutamate receptors. Journal of Neurochemistry., vol. 53, no. 3, pp. 710–715,.

Pearson, S., and Nunn,P.B.1981 "The neurolathyrogen, β-N-oxalyll- α,β-diaminopropionic acid, is a potent agonist at glutamate preferring receptors in the frog spinal cord," Brain Research, vol. 206, no. 1, pp. 178–182.

Rao,S.L.N., Adiga,P.R., Sharma,P.S.1964. Isolation and characterization of b-Noxalyl-L-a,b-diaminopropionic acid, a neurotoxin from seeds of Lathyrus sativus. Biochemistry., 3: 432–436.

Siddique, K.H.M. and Hanbury, C.D. 1998. The introduction, evaluation and utilization of Lathyrus germplasm in Australia. In: Mathur PN, Ramanatha RV and Arora RK (editors). Lathyrus genetic resources network. IPGRI, New Delhi, India. pp. 42-45.

Smartt, J. 1990. Grain Legumes: Evolution and genetic resources. Cambridge University Press, Cambridge, UK. p. 200.

Smartt, J., 1990. Pulses of the classical world. In: Summerfield, R.J., Ellis, E.H. (Eds.), Grain Legumes: Evaluation and Genetic Resources. Cambridge Univ. Press, Cambridge, pp. 190–200.

Smartt, J., A. Kaul, Wolde Amlak Araya, M.M. Rahman and J. Kearney. 1994. Grasspea (Lathyrus sativus L.) as a potentially safe food legume crop. p. 144-155. In: F.J. Muehlbauer and W.J. Kaiser (eds.), Expanding the Production and Use of Cool Season Food Legumes. Kluwer Academic Publishers. Dordrecht, Netherlands

Smartt, J., Kaul, A.K., Araya, W.A., Rahman, M.M., Kearney, J., 1994. Grasspea (Lathyrus sativus L.) as a potential safe legume food crop. In: Muehlbauer, F.J., Kaiser, W.J. (Eds.), Expanding the Production and Use of Cool Season Food Legumes. Kluwer Academic Publishers, Dordrecht, Netherlands, pp. 144–155.

Spencer, P.S., 1989. The Grass Pea: Threat and Promise. Third World Medical Research Foundation, New York, pp. 244.

Spencer, P.S., 1995. Lathyrism. In: De Wolff, F.A. (Eds.) Handbook of Clinical Neurology, vol. 21 (65): Intoxications of the Nervous System, Part II. Elsevier, New York, pp. 1–20.

Spencer, P.S., Ludolph, A.C., Kisby, G.E., 1993. Neurologic diseases associated with use of plant components with toxic potential. Environ. Res. 62, 106–113.

Spencer, P.S., Roy, D.N., Ludolph, A., Hugon, J., Schaumburg, H.H., 1986. Lathyrism: evidence for role of the neuroexcitatory amino acid BOOA. Lancet 2, 1066–1067.

Spencer, P.S., Roy, D.N., Ludolph, A., Hugon, J., Schaumburg, H.H., 1988. Primate model lathyrism – a human pyramidal disorder. In: Nappi, G., Hornykiewicz, O., Fariello, R.G., Agnoli, A., Klawans, H. (Eds.), Neurodegenerative Disorders: The Role Played by Endotoxins and Xenobiotics. Raven Press, New York, pp. 233–238.

Strisovsky K., Jiraskova J., Barinka C., Majer P., Rojas C., Slusher B.S. & Konvalinka J. (2003. Mouse brain serine racemase catalyzes specific elimination of L-serine to pyruvate. FEBS Lett., 535, 44–48.

Tekele-Haimanot, R., B. Abegaz, E. Wuhib, A. Kassina, Y. Kidane, N. Kebe, T. Alemu and P.S. Spencer. 1993. Nutritional and neuro-toxicological surveys of Lathyrus sativus consumption in northern Ethiopia. Pp. 41-45 in Lathyrus sativus and Human Lathyrism: Progress and Prospects (H.K.M. Yusuf and F. Lambein, eds.) Proc. 2nd Int. Colloq. Lathyrus/Lathyrism, Dhaka, 10-12 December 1993. University of Dhaka.

Tripathy, S.K.; Ranjan, R.; Dash, S.; Bharti, R.; Lenka, D.; Sethy, Y.D.; Mishra, D.R.; Mohapatra, B.R.; Pal, S. Genetic analysis of BOAA content in grasspea (Lathyrus sativus L.). Legume Res. 2015, 38, 465–468.

Watanabe, M.; Kusano, M.; Oikawa, A.; Fukushima, A.; Noji, M.; Saito, K. Physiological roles of the β_-substituted alanine synthase gene family in Arabidopsis. Plant Physiol. 2008, 146, 310–320.

Williams, P.C., R.S. Bhatty, S.S. Deshpande, L.A. Hussein and G.P. Savage. 1994. Improving nutritional quality of cool season food legumes. p. 113-129. In: F.J. Muehlbauer and W.J. Kaiser (eds.), Expandingthe Production and Use of Cool Season Food Legumes. Kluwer Academic Publishers, Dordrecht Netherlands.

Yadov, C.R. 1995. Genetic evaluation and varietal improvement of grasspea in Nepal. Pp. 21-27 in Lathyrus Genetic Resources in Asia. Proceedings of a Regional Workshop, 27- 29 December 1995, Indira Gandhi Agricultural University, Raipur, India (R.K. Arora, P.N. Mathur, K.W. Riley and Y. Adham, eds.). IPGRI Office for South Asia, New Delhi, India.

Yan ZY, Spencer PS, Li ZX, Liang YM, Wang YF, Wang CY, Li FM (2006) Lathyrus sativus (grasspea) and its neurotoxin ODAP. Phytochem 67:107–121. doi:10.1016/ j.phytochem..10.022

Yi, H.; Jez, J.M. Assessing functional diversity in the soybean _-substituted alanine synthase enzyme family. Phytochemistry 2012, 83, 15–24.

Yunus, A.G. 1982. A taxonomic study of the grass pea, Lathyrus sativus L. MSc. Thesis, University of Birmingham.

Yunus, A.G., M.T. Jackson and J.P. Catty. 1991. Phenotypic polymorphism of six enzymesin the grass pea (Lathyrus sativus L). Euphytica 55:33-42.

Zeven, A.C. and J.M.J. de Wet. 1982. Dictionary of cultivated plants and their regions of diversity. Pudoc, Wageningen.

Zhang, J.; Xing, G.M.; Yan, Z.Y.; Li, Z.X. _-N-oxalyl-L-_,_-diaminopropionic acid protects the activity of glycolate oxidase in Lathyrus sativus seedlings under high light. Russ. J. Plant Physiol. 2003, 50, 618 622.

Zhang,C., Vasmatzis,G., Cornette,J.L. and DeLisi,C. (1997) Determination of atomic desolvation energies from the structures of crystallized proteins. J. Mol. Biol., 267, 707–726.

Zhou, G.K.; Kong, Y.Z.; Cui, K.R.; Li, Z.X.; Wang, Y.F. 2001. Hydroxyl radical scavenging activity of β-N-oxalyl-L-α,β-diaminopropionic acid. Phytochemistry . 58: 759–762.

RESUME

Name	:	**Jajati Keshri Nayak**
Date of Birth	:	11-11-1995
Present address	:	Room No-214, Sundaram Boys hostel, IGKV, Krishak Nagar, Raipur 492012
Mobile no.	:	7008749898
E-mail	:	jajatikeshrinayak00@gmail.com
Permanent address	:	At Bhadrak post Bhadrak, Odisha (756125)

Academic Qualification:

Degree/ Class	**Year**	**Board**
10th	2010	KCVM,AGARPADA
12th	2012	BJB AUTONOMOUS COLLEGE
B.Sc. (Ag.)	2017	OUAT,BHUBANESWAR
M.Sc. (Ag)	Appearing	IGKV, RAIPUR

K Nayak

Signature

Gmail **Jajati keshari Nayak <jajatikesharinayak00@gmail.com>**

Manuscript Accepted: (Ref: Chemi-7-4-732)...

1 message

Chemistry Journal <chemi.journal@gmail.com> Mon, Aug 13, 2019 at 12:45 PM
To: jajatikesharinayak00@gmail.com

Dear **Author,**

Ref: **Chemi-7-4-732**

A manuscript titled "**Analysis of phenotypic characters in Lathyrus sativus related to ODAP**" is very well written and has been accepted for publication. Note: After pay fee kindly send receipt to our mail id. Please send complete filled copyright form.

You are required to pay publication fees of Rs 2500 + 18% GST = Rs. 2950. Kindly inform us after payment.

Pay the fee in following account.
Bank Name: IDBI Bank
A/C Holder Name: Linear Publication
A/C Number: 0163102000031064
A/C type: Current
IFS Code: IBKL0000163
Branch: Sector 3, Rohini, NEW DELHI

Best Regards
Office
International Journal of Chemical Studies
www.chemijournal.com