



HYDROPATHICITY, PHYLOGENETICS AND POSSIBLE CODONS ANALYSIS FOR ASPARAGINE, GLUTAMINE, ARGININE, AND PROLINE PROTEINS IN VICILIN GENES OF *ABROMA ANGUSTA*, *SOLANUM LYCOPERSICUM*, *THEOBROMA CACAO*, *HERRANIA NYCTERODENDRON*, *Z. FURFURACEA*, *PISUM SATIVUM*, *VICIA FABA*, *THEOBROMA CACAO*

LAYDONG LEPCHA*, BASISTHA B.C. AND SUBBA K.B.

Bioinformatics Centre (Sub-DISC), Sikkim State Council of Science & Technology, Department of Science & Technology and Climate Change, Vigyan Bhawan, Deorali-737 102, Gangtok, Sikkim, India.

*Corresponding Author: Email-laysimick@gmail.com

Received: March 03, 2016; Revised: April 11, 2016; Accepted: April 12, 2016; Published: April 28, 2016

Abstract- The study was conducted by analyzing and retrieving DNA *Vicilin genes* of various plant species whose partial genomic information or database are so far been experimented and submitted in gene bank. The DNA sequences of *Vicilin genes* were retrieved from plant species such as *Abroma angusta*, *Solanum lycopersicum*, *Theobroma cacao*, *Herrania nycteroendron*, *Z. furfuracea*, *Pisum sativum*, *Vicia faba*, *Theobroma cacao*. Specific Bioinformatics software tools were applied to analyze the necessary genomic information from studied *Vicilin genes*. The codons ratio of proteins amino acids such as *Asparagine* (AAC,AAT,GAC,GAT), *Glutamine* (CAA,CAG), *Arginine* (AGA,AGG,CGA,CGC,CGG,CGT), and *Proline* (CCA,CCC,CCG,CCT), were analyzed from within the retrieved DNA *Vicilin genes*. The study found that the major storage protein amino acids *Asparagine* (Asn), *Arginine* (Arg) and *Proline* (Pro) is ingredient more in *Vicilin genes* of *Abroma angusta*. It also studied the hydrophobicity nature of specific amino acids from various *vicilin genes* of studied plant species. It helps to understand the protein amino acids interaction behaviour with the hydrogen bond of H₂O. According to the Phylogenetic study, two of the *Vicilin genes* possessing plant species; *Pisum sativum* and *Vicia faba* are close relatives, belonging to family *Fabaceae*. This signifies that family *Fabaceae* is possibly being a dominant plant family in exhibiting secretion of *Vicilin genes*, an anti-decay gene. The main objective of the study was to analyze and understand the efficient *Vicilin gene* from various studied *Vicilin genes* of species respectively regarding their constituents of major proteins amino acids such as *Asparagine*, *Glutamine*, *Arginine*, and *Proline*. And also to find out the plant family that exhibits the most dominant status for the ingredients of *Vicilin genes*.

Keywords- Codons, Proteins, Bioinformatics, Genomic, Phylogenetic, Hydrophobicity, Amino acids

Citation: Laydong Lepcha, et al., (2016) Hydrophobicity, Phylogenetics and Possible Codons Analysis for *Asparagine*, *Glutamine*, *Arginine*, and *Proline* Proteins in *Vicilin genes* of *Abroma angusta*, *Solanum lycopersicum*, *Theobroma cacao*, *Herrania nycteroendron*, *Z. furfuracea*, *Pisum sativum*, *Vicia faba*, *Theobroma cacao*. International Journal of Bioinformatics Research, ISSN: 0975-3087 & E-ISSN: 0975-9115, Volume 7, Issue 1, pp.-329-345.

Copyright: Copyright©2016 Laydong Lepcha, et al., This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution and reproduction in any medium, provided the original author and source are credited.

Introduction

Vicilin genes are used as seed storage proteins for conserving valuable seeds of fruits and foods. It is an anti-decay proteins. *Glutamine* (Gln), *Asparagine* (Asx), and *Arginine* (Arg) Protein's amino acids are considered to be Nitrogen rich proteins [1a]. The healthy growth or development of plant has positive impacts towards its ingredient quality and quantity fruits. Seasonal fruits have a high potential quality to pay healthy and sustainability life cycle but it needs early development of plants. Nitrogen along with other amino acids play an important role for the germination and early development of plants [1b]. The proteins such as *Globulins*, *albumins* and *prolamins* are major seed storage protein groups [2]. Among these protein group *Globulin* has two genes; *Vicilin* (7S) and *Legumin* (11S). The *Vicilin* and *Legumin* gene are regarded as a major source for the supply of Nitrogen and amino acids for the germination and early plant development. The longevity and stability of gene expression depends upon the nature of sequences of nucleotides. The sequence has less chances of becoming short with the fewer inceptions of stop codons. In DNA the stop codons are TAG, TAA and TGA.

Various proteins based studies such as designing method, understanding the strategies for isolated homology identification, detect the protein folding characters and its comparative study rely upon the interaction nature of side chains [3-5]. The importance of interaction between amino acids and H₂O is well realized [6-9]. The hydrophobicity is a physicochemical study of proteins, leading towards the

development of hydrophobicity and hydrophilic atmospheres of proteins. The hydrophobicity are positive (+) in nature and tend not to be in an aqueous environment [10], which intrinsically means that interaction with hydrogen bond of H₂O is completely divided. While the hydrophilic are negative (-) in nature, more the negative numbers greater the affinity or interaction for water [11]. *Arginine* (Arg) is a basic and polar essential amino acid. The formation of polar amino acids enhances the interaction between H₂O hydrogen bond and proteins, leading to a potential dimensional structure of proteins. Medicinally *Arginine* is also a valuable protein, when added to the diet it allow to participate in hepatic clearance of ammonia from the body by means of urea cycle, which reduces the risk of hyperammonemia, one of the clinical signs of *Arginine* deficiency [12]. As a cation, *Arginine*, as well as lysine, plays a role in maintaining the overall charge balance of a protein (www.biology.arizona.edu). *Arginine* also plays an important role in nitrogen metabolism. *Asparagine* (N) is a significant component of seed storage proteins in legumes and a major transport compound in cereals [13]. *Glutamine* (Q) acts as a sensor to sense its reduced nitrogen status and regulate the uptake and reduction of nitrate. Several lines of evidence have suggested that *Glutamine* may play an important role [14]. The *Proline* (P) amino acid plays an important role for maintaining tolerance ability of plant during the adverse environmental conditions [15-16].

This comparative analysis between Genomic *Vicilin* DNA nucleotides of various plant species taken from various plant family will reveal the blueprint *Genomic and*

Proteomic distribution of Asparagine, Glutamine, Arginine and Proline Protein amino acids in *Vicilin* genes including all other content protein amino acids. It will also promote analysis and understanding the close relative families between plant species that produces *Vicilin* genes and a better prediction of durable proteins that encodes Nitrogen for better seed germination and early development of plant.

Materials and Methods

Database sources

The Gene bank *Vicilin* gene sequences of *Abroma angusta* (Sterculiaceae) Accession No. AF113031.1, *Solanum lycopersicum* (Solanaceae) Accession No. AM932874.1, *Theobroma cacao* (Malvaceae) Accession No. X62625.1, *Pisum sativum* (Fabaceae) Accession No. AJ276876.1, *Vicia faba* (Leguminosae) Accession No. Y00506.1, *Z furfuracea* (Zamiaceae) Z54365.1 and *Herrania nycterodendron*, (Byttneriaceae) Accession No. AF113059.1, were retrieved from National Centre of Biotechnology Information (NCBI) web database source. The sequences were duly analyzed in other biological web sources such as Protein Data Bank (PDB), The European Molecular Biology Laboratory-European Bioinformatics Institute (EMBL-EBI).

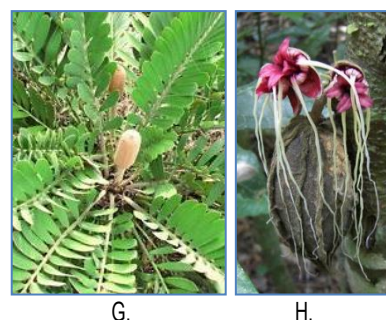
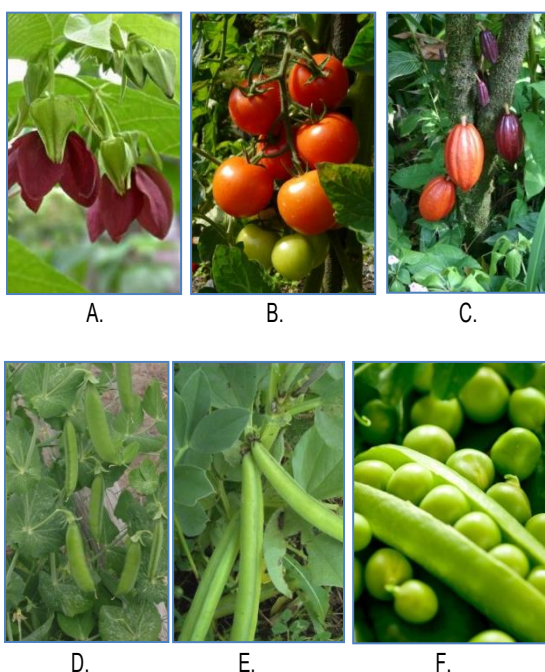


Fig-1. (A). *Abroma angusta* (B). *Solanum lycopersicum* (C). *Theobroma cacao* (D). *Pisum sativum* (E). *Vicia faba* (F). Pea (G). *Z. furfuracea* (H). *Herrania nycterodendron* (source: <https://www.google.co.in>)

Bioinformatics Software

Genomic Evaluation of retrieved *Vicilin* DNA sequences

The study was done in Bioinformatics Sub-Distributed Information Centre (Sub-DISC), Sikkim State Council of Science & Technology, with the support of major facilities available at the center from the Department of Biotechnology, Government of India. The retrieved *Vicilin* DNA sequences were evaluated, authenticated and calibrated by referring various gene sequencing parameters, from various Bioinformatics software sources such as NCBI Blast methodology, ACUA software (1.0 version) and Geneious Pro (Trial version).

Phylogenetic study

The Phylogenetic analysis was done by using MAFFT PHYLOGENY Bioinformatics tool service of European Bioinformatics Institute.

Genomic Sequence Transcription and Translation

The retrieved DNA *Vicilin* genes were transcribed to mRNA (Messenger Ribonucleic Acid) and translated into Protein sequences by using Attotron Biosensor Corporation's Bioinformatics Software (<http://www.attotron.com/cybertory/analysis/trans.htm>).

Amino acids composition or occurrence ratio and Hydropathicity analysis

To understand the composition or occurrence nature of all the amino acids, composition of negative and positive charge residues, contents in all the *Vicilin* genes, and their Hydropathicity, Bioinformatics software ProtParam tool (ExPASy) was used.

Results

The transcribed and translated *Vicilin* genes sequences detail are as follows:

Table-1 Transcribed and translated of *Vicilin* genes sequences.

Gene Ascension No.	DNA Sequence	mRNA (Codons encoded)	Protein
[6469858 gb AF113031.1 <i>Abroma angusta</i> <i>Vicilin</i> gene, partial cds	GAAGAARGCANRKYCAAGKTCCTTCAGAGGTTTGGCTCAGAGT CCCGTCTACTTAAGGTCATTGATAAC TACCGTTGGCCATGTT CGAAGCAATCCCAAGACTTTTGTCTCCACACCCTGTGAT TCCGAGTCAATTTACTTTTGTGACAAACGGTAAATCTCTTTTTC TCTTGACAAACATTTGGCCGCTTTATCACCTAAGGCCAAACA AACCATTTACTAAAACATGGTAACTTCTTTCTGTGTGAGGAA GGGACAAATCACTTTGTGAAATCATGAAACAAAGAATGTCCTAT AATTTAGAGCGTGGAACTGTTGTGACAAATCTCTGCTGGAAGCA CTGTGACTTGGTCAACCAAGGCAACAAAGAAAGCTGACCAT AGCCGTGCTAGCACTGCCTGTTAATACCCCTGGCAATCTGA GGTTGATTTCTACTTAACATTAAITTCCAAACATTTTCTATCTT AATTTAGAGCTTAGTCTCTTCGTCTCGCTATTAAACCTGGGC TTTATAATGTGCTCAGAAATCTTCCCTGCTGGAAGTGAAT CCTCAGTCATATTTCTGTCTTCAAGTTSTGAAATCTTTGAGAG TGTATTCAATGTATAGTTCTTTCACATGAAATATGAAATCAATTT TGTTACTATCTCTTGTCTTTTATTATCTATTTATTTCTTTC TATCTATTATTTTCTCTTATTAGACACCAATAGAGAGTTAGA GAAGCTCTTTGAGGCGCGAGCGGGCAGCAGGCGAGAAGA GTCCAGAGGGTGTATTCGGCAGAGCCAGCAGATTA GGGCAATGAGCACAAGTCTACTTCGCAAGGCAGAGGCG	UGAAGAAGCACAAGUCCUUAGAGGCUUU GCUAGAGUCCCGUCUACUUAAGGCA UUGAUAAUJACCGUUGGCCAUGUU UAGCAAAUCCCAACACUUUUUGUUCUCCAC ACCACUGUGAUUCGAGUCAAUUUACUU UGUGACAACGGUAAAUUCUCUUUUCCU CUUGAACAAACUUUGGCAGCUUUUAUCA CCUAAGGCCAAACAACCAUUUACUAAA ACAUGGUAGUUCUUCUGUGUCAGGAA AGGGGACAUAUCAGUUGUGAAUCAUGA AAACAAGAAGUCCUUAUUUUAGAGCGUG GAAUGUUGUGACAUAUCCUGCGGGAAG CACUGUGUACUUGGUCAACCAAGGCAC AAAGAAAGCUGACCAUAGCCGUGCUAG CACUGCCUGUAAUACCCUUGGCAAAUC UAGGUGUUAUUUCUACUUAAGAUUU UUCCAACAUAUUUCUUAUUUUUUA GCUUAGUCUCUUUGUCUGCUAAUUC CUGGGCUUUAAUUUGCUGCAGAAUUC UUCUGCUGGAAGUGAAAUCUAGU	RSTSPSEVCSEVPST.GH..LP LGHVRSKSHFCSPPTL.FV NLLCDKILFSS.TNFR.LYH LFPNKPFKNMVTSFVCKG DNHVCES.KRVL.FAWNC CDNSCWKHCVLGPRCK ADHSASTAC.YPW.I.G.FLL NINFPNIFYLNLEA.SLR LAINP GLYNCCRNSSLLEVEILSHS LSSV.NSDICIVL.P.NMN FCLLSLLFFYLSIYSFSLYYFS LIRHRS.SSLRGRAGSRG RRVRRVYSGEPSRSRGLA NKLLRGR.EAGNLFYSIY.AN HLSTPTKDAFLRVP.NSS NFR.T.LS.LL.H.N.TRYLKPND FSII..ISFMNLC.SILF.GSIFVP

	<p>GGGGAGAAGCTGTTTTCATTTGAGCCAAACACCTGTCT CACTC CAACCAAAA CGGACCGCTTTTGGAGCTTGTCCGAG GAA TTCAGCAAGCTT CAGAACATGACTGTCTCAGTTACAGCAT TAAAATG AACCCGGTATTTGAAACCAATGATTTTCAAATAA TTGATAGATAGCTTCATGAAATTTGTGCTGATCAAACCTTTTTC AGGGATCAAAATTTGTGCCA</p>	<p>CAUAUUUCUCUGUCUU CAGUUUGAAAUU CUUGACUGUAUUCAAUGUAUAGUUCU UUGACCAUGAAAUUGAAUCAAUUUUUGUU UACUAUCUCUCUUUGUUUUUUUUUUUAUC UAUUUUUUUUUUUUUUUUUUUUUUUUUUU CUCUUUUUACACCAAUACGAAGAUUAC GAAGCUCUUUUGAGCGGCGAGCGGCA GCAGAGGCGAGACAGUCAGAGGGUGUA UUCGGGAGGCCAAGCAGGAGUA GCGCAAUGAGCCAAAGCUACUUCGCC AAGGCGAGAGGCGGGGCAACUUGUU UUCAUUCUUAUGAGCCAAUACCCUGUCCA CUCCAACCAAACGGACGCUUUUUUUGAG GCUUGUCCCGAGGAUUCAAAGCAAGUUC AGAACUGACUGUCUCAGUUAACCAUUA AAUUUGAACCGGUUUUGAACCAAUG AUUUUUCAAUUUUUGAUUAUUAGCUUC AUGAAUUUGUCUGAUCAAUCUUUUUUC AGGGAUCAAUUUUUGGCCA</p>	
<p>166053039 emb A M932874.1 <i>Solanum lycopersicum</i> Vicilin gene for vicilin, isolated from seeds</p>	<p>ATCGGCCGAAATGATCATCACTAAAATACAAATAAAGTTTTTC ACTCACTATGGCAAITTTCAAAAACCAAGCTTTTATTTATCT TCTTCTTCTCTCTCTTTGGTCTTGTATCTCAAGCTATGAT CAAAACCTAGGGTTTACCAAGACCTCAGGAGAAATACGA GAGTGC CAACAACGTTGTGAGACAACAACAGGACAACAG AAACAGTTGTGCAACAACGTTGTGAAACAACAGTATAGAAAG GCAACAACAACAATGAGAGGGGAGACTGGTGAAGATGATC TAGGCAATCGTGGGCTGATAGA GCTACAAAGATTTGCAAG AATGC CAACGTTAGTGCAGAGTGAAACAACAGGCGCAACGAC TACAAGAGTGTCAACAACGTTGTCAACAAGAGTACCAAAGAGA GAAAGGACAACACCAAGGTGAAACTAAGCCAAGTGGGAACA ACAAGAAATCAAACAATCCATACTTATTCGAGTCTCAGCGAT TCAGGTCTCGATTCAGAGTGTGATGATTTCCGAACTCCT CGAGAGATTCAATCAAAGATCTCAGCTTCTCAAAGGAACTCGAA AATACCGTGTCCGATCCTCGAATGGAACTCAGTCTTTCG GTTGGCTCATCTGTGATGGTGAAGCATTATGTGCTGGT TAAAGGACAAGGAGTAAATATATAGCAGAACAGATAACA AAGTCTTCAACTTGCAGAAAGGAGATGTAATCAGATTGTTG CTGGTCAAATGCTACTTGTCAACAAGATACAACGAAAA GCTCTTTGTTAAGTCTGCTAAGTCCGTCAAAGCCCTGGGA AATTTGCAGAAATACCTTAGTGCCGGAGGCCAAAATCCTGAT CCTTCTACAGAGCATTAGCAGCGATATCCTGGAGTCTGCTTT CAATAACCCAAAGGATAGTTAGAGAGGCTATTTGACAAACAC AAGGAGGATATAATAAAGCTAGTGAAGCAAAATTCGA GCTTAAGCCGAACAACCGCTCCGCTCAACTCAGCAAACTAGA GGTAGAACACAGGAGCTTTCAATTTGATGAAGGAAACGCCG GTATTGCAAAAGTATTTGAGCAGTCTTTGAAGCACTCCCG AAAGATACGAGCAATGAGGGAGTTGATGCTGCTGTTGTTT CATGAACATTAAACCAAGTGGAAATGGTACTACCATACTACAAC ACGAAATCTACAAATGTTATGTCATAGAGGAAACCGCTC GATTTGAAATGGCATGCTCTATCTTGGTAGACAAGCCAGAG CCCATGGTCTCGTGCACAAGGAGAGAGCAAGAGCGAGAG AAGAACAAGACAAGAAAGAGAGATGTCATTACCAGAAAT CCGCGGTAACTAAAAGTTGGTATGTTGGTAAATCCCTGCT GGCCATCAAATACCTTTGTAGCAACCGGAACCTCGAATCTGA GGATCGTTGGTTCCGAGTCTGATGCTGAGAACAAACAAAGA AGTTCTTGCAGGTAAACAAGATATGGAGAACATAGACAG GGAAGCGAAAGAACTGCTTTAGCATGCCGGGAAGGAGGT AGAAGAGATATTCCAGAGGCAAGACAGTCTACTTTGTGGCT GGACCGGAGCATCGCCAGCAGAGGAGAGGGGTGAAGAAGG AAGAAGGGACAAGATCAAATTTGTCTTCAAATTTGACTTTG TTTTCTAAAAGATGTTAAGTATGCACAAAGCAGATATGCGT GTGCCAAAGTAAAAGCTTTGTGAGTAAATAAGAAATGAGGCT CCAGAAGAAATATGTAATACTACTAATGTAACCTCTTAAAAGTA ATGTGGTCTAAAGATTTTATAGAGGTTTTTATGACTCTTGT AGTATGATGATGATTTTGGTGTGCAATGTAACTTTTGCATCAT A</p>	<p>AUCGGCGAAUUGAUCAUCACUAAAUAAC AAUAUAAAGUUUUUACACUCACUAUGGCAA UUUUCAAAAACCAAGCUUUUUUUUUUAUC UUCUUCUUGAUCCUCUCUUUGGCUUGU UAUCUCAAUGCUAUGAUCAAACCCUAGG GGUUAACCAAGCCCUAGGAGAAAGUA GAGUGGCAACAACGUUGUGAGGAC ACAACGAGACCAACAGAACAGUUGUCA AACAACGUUGUAACAACAGUAUAGGAA GCAACAACAACAACAUUGGAGGAG CUGGUAAGUUAUCUAGGCAAUUGG GCCUGUAUAGGCUACAAGUUGCAA AAUUCCAACGUAGGUGCAGGUGAACAA ACAGGCGCAACGACUACAAGUUGUCA CAACGUUGUCACCAAGUACCAAGAGG GAACGCAACAACCAAGGUGAAAGUAGC CACAGUUGGAAACAACAAAGAAUCAAAC AAUCAUACUUUUUGAGUCUCAGCGAU UCAGGUCUCGUAUCAGGCUAGUCAUG UGAUUUCGAAUCCUGGAGAUUCAU CAAACAUUCAGCUUCUCAAAGGAUUCGA AAAUJACCGUGCGCAUCCUCAAUUG GAACCUAGUCUUUCGUUGUCCUCAUC ACUGUAUGUGAAGCAAUUUUGUUGU GGUUUAAGCAAGAGUAAUUAUUAUG CAGAACAAAUAAACAAACUCCUUCAC UUGCAGAGGAGUGUAUUCAGUUGU UUGCUGUCAAUUGUCUACUUGCUCAA CAAACAUAAACAACGAAAGCUCUUUUU ACGUCUCGUAAGUCGUCAAUUGCCC UGGAAAUUUGCAGGAUACUUUAGUGCG GAGGCCAAAACUCCUUAUCUAC GAGCAUUCAGCAGCGAUUCCUGGAGUC UGCUUUCAAUAAACCAAGGUAUAGUU GAGGCUAUUUUGGACAAACAAAGGAGG GGAUAAUUAUCAAAGCUAGUGAGGACAA AUUCAGCUUAUAGCGAACACCGUCGCG GCUCAACUCAGCAACUACAGGUAGAAC CAGGACCUUUCAAUUUGAUGAAGGAAAC GCCCGUAUUUGAAAGUUAUUUGGACA GUUCUUUGAAGCAGUCCCGAAAGUAC GACAAUUGAGGACUUGGAGUUGGUCUG UUGUUUCAUGAACAUUACCAAGGUGG AAUGGUACUACCAUACUACAAACGAAAU CUACAAAUUGGUUAUGGUCAUAGAGG GAAGGCUUGAUUUUGAAUUGGCAUCCU CAUCUUGGUAGCAAGCAGGACCCAU GGUCUCGUGGACAAAGGAGGACAAAG CGCGAGCAAACAAACAAACAAAGGAG GAGUUGUCCAUUACAGAAAUCCGCGG UAAACUAAACGUUGGUAUGUUGGUA AUCCUGCUGGCAUCAAUACCUUUG UAGCAACCGAAAGUCGAAUCUGAGGAU CGUUGUUUCGAGUCGAGUUGGAGAAC AACAAACAAAGUCCUUGCAGUAAACA AACAAUAGGAGGAAGAUACAGGGAAG CGAAAGAACUGUCCUUUAGCAUGCCGG AAGGAGGUAGAGGAUUAUUCAGAGG CAAGACAGUCCUACUUUGGCGUGGAC</p>	<p>IGRIDH.NTISFSLTMAIFTK PKLLFIFLILSLVLVS CYD NPGYDPKLEKRECDRC ERCPGCKLCKRCE GYRKEQKHGGETGEDDL GNRGPKSYKRLTECFRC SEGGRLECDRC EYREKGHGETNPWE GLEKSNPPLFESRFRRF RASHGDFLRFNRSLL KGIEKRVVAILELPSFVLP HHCDEGAIYVVVKGVINIA EDNKNSFNLKGDVIRLFA GSNVYLLNKDNNEKLFVYV AKSVNAPGNLEYFSAGG NPESFYAFSDIGETSNFP RDKLERLFGHKEGIIKASE ERRAISEHASSTTRGR TGPFNLMKERPVFESRF FFEARPERYERDLDAV GFMNINGGMVLPPYNTKST KLMVIEGNAFEMACPHLG RASLSPWSRGCREERE EEIEEGDVHYKIR NVGDVLPVIPHGPITVATGN SNLRIVFGVD AENKKNFL AGKNIVNIDREAKELSFS MPGEVEEIFRDRSYFVA GPEHRGCRGEEGRG DYLSLIDFVFKSVRYAC RYAFVPSEKSFVSK.E.GSREI CTTNTVPT.K.MWSKEYFRG LDSLIVCMYVLVCNVTAS</p>

		<p>CGGAGCAUCGCAGCAGAGGGAGAGGGG UGAGAGAGGAAGAGGGGACAAAGUCAA UAUUUGUCUUCAAUUUUGGACUUUUGUUU UCUAAGAGUGUAUCGUUAUGCAAAUG CAGUAUGCGUUGUGCCAAUGAAGAG GCUUUGUGAGUAAAUAAGAAUGGCU CCAGGAAAUUGUACUACUACUAAUGU ACUCCUAAAAGUAAAUGUGGUCUAAAG GUUUUAUAGGGUUUUUAGCUCUCUU GUAGUAUGUAUGUAUGUUUGGUGCA AUGUAUUUUUGCAUCAUA</p>	
<p>21910 emb X62625 .1 <i>Theobroma cacao</i> csv gene for seed vicilin</p>	<p>TCACITTTATCCAGAGATATTTTTTCAAAITTTCTCCATTTAA CGAGGAAATAAAAAAATATCGTCAAATTTGTTTCATATCAT ATCCTTGCAGCTCATCGCATGCAAGCAGCAGGTGTCAAC ATGAGCGGTAGATCTGCGAGAGTGCGAGATCATTATCAACTCAA CTTAACTCGTGTGTCCTCAATCCAACTCAAAGAAACGCTTTC CTAACTCGCTTAAACACACAAATCCCACTCTTTAACTCAACT AAGTTCAAGCAAGAAACAAATTAATCGACGAGCTCTGCTGC CAAACACCATCCTTCCGTACATCTAGTTGCGAGACCGACCAA CAAATTCATGGTACAAAGCTTGTCCGGACCTGCGCCGTACCA CGTGAAGGATGTTGCTCTGAGTTTCCCATCTTAAGAA AAGGATCAAAGTATTTGCTGTTAAATGACACTCGTCACT TTGCATGTCAAAGTCTTACAGTAAAGGAGATTTGCATGG AAAGCTTAGCCTCGCTTCTATAAATAGCTTGCCTCTCTTTG CTCTTATCACAAGAAAGAACACAGATCAAAGCATAGCAA ATGTTGATCAGTAAGTCTCTTTCATAGTTTATCTTCTCT TTCTCTTCTTTTCTGTTGCTTGTCTGTTGCTGAGCGCTAT GGAGAAACAAATGAGCGTGTCTCTGACAGCAAACGAG CAAAGCCAGAGCGATGCGAGTCCGGAAGCGACAGAAAG GGAGCAAGCGAGTGTGAACAAGCGTGTGAAGGAGTAA GGAGCAGCAGAGCAGCAAGCAAGAAAGAGCTTCAAAGCAA CCAGCAAAGTCAAGGCGTGTGAAGCAACAACAGGGCC GAGAGAGCAGCAGCAGTCCAGAGAAAGTCTGGGAGCAA TAAGGAAACAAGAGAGGCGAGCAGCAAGATACCATAA CAAAAAATAGGAGCGAAGCAAGAAAGAAAGGCAACAA CAACTTACTATTTTCTAAAAGAAGATCATTCAAAGTGA TCAGGATGAGAGAGGCAAGTTCAGATCCTCCAGAGTTT CTGAGACTCTCCTCACTCAAGGCAATCAAGATACCGCTT GGCAGTTCGAGCAAAACCAACACTTTTATTTCTCCGCAC CACTGTGATGCTGAGGCAATTTACTTCTGACAAACGGTAA TCTCTTCCCTTTCGAAACAAATTTTGGGCTTTTATCACAA CACCATTCATGAAAGATAGTAAATATTTCTGTGTGAGGAA GGCACAAATAGTTGTGACTCATGAAACAAGAGTCTCT AATGTACAGCGTGGAAAGTAGTCAAGCTTCTGCAAGG ACTGTTTACCTGTTAGCAGAGCAACAAGAGAGCTAACCA TAGCTGTCTCGCCCTGCCTGTTAAATCTCTGGCAAATG GGTTTTCTACTAACATTTATTTAGAAGTTTCTACATACCT TBAATGCTTGGTTAGTCTATTCAACCTGTACTTAAATGCTG AGTTATTTCTCCCGCTGGAATAAATAAAGCTCAATCATTTAC GGAGCCTCAGCTATGAAGTCTTGAAGCCGCTTCAAATGAT ATGGTCTTTGACTAGCTATGAAATTTTTGCTTTTACCTT TCTTCTGCTTTATTTATTTATTTATTTATTTATTTATCTGT TTGTTAGACACAAAGAGAGAGCTGGAGGAGATCTTGGAGGA ACAGAGAGGGCAGAGAGGAGCAGGAGGAGGAGGATATGT TCCGGAAGGCACAAACAGAGCAGATAGAGCAAAGCCAAAC AAGCTACTTCTCAGAGCAGAGGAGGAGGAGGAGGACTTGCCA TCAAATCTATTGAGCAAACCGCTGTCTACTCACAACAAACGG ACGCTTCTTTAGGCTTGTCTGAGGAGTTCAAGTCAAATTCAG AACATGATGCTCGCTGTTTCAAGCTTCAAAGTCAATCAGGTAC TAAATCAATGATTTTTTTTCAAATTTCTATTAATAATTTCTC TTGTACTACTACAGCACTACTAGATAGCTTCTAGTAAATAAGC TCATGAACTTTTGCAGGAGCCATATTTGTCCACACTACA ATCTAAGGCTACATTCGTGTTGTTGTACCGGACGGATATGG GTAACCTCAAAGGCTTCCCTCTCCAGACAGAGCCAA GGATCCCAAAGTGGAGGCAAGACAGAAAGAAACAAGAA AGAGTCAAGAGGAGACATTTGAGAAATTCAGCAGGTCAA AGCCCCATTGTCACTGGTCAAGCTTCTTGTAGCCCCGGCAGG CATGCAGTTACATCTTTGCTCAAGAGCAGCCCTGAAAG CAGTCTCGTTGAGTCAAACCGCCAGAAACAAACAGAGAA TTTCTTGCAGGTAGGCTTTTTTTCTGAAACAAAGCAAAACA AATGTCATCAAATTCATGTAAGGCTTCAAGTAAATTTGAT CAAATGGTCTTGTGTTTCTATAGGAAAAAGAAATGTTGTC GACAAATGGATAGGAGGCAAGAGGATTTATCTTTGGGTAC CATCGAATGGTAAATATATTCAACAACCGGATGAGT GATTTTCATGCTTCTCTCAACAGAGGAGCGTGGAGATGAA AGGAGGGCAAATCCCTTGGCCTCAAATCTGAGTTTGTCCCG</p>	<p>UCACUUUAUCAGGAUUUUUUUUCACAA UUUUCUCAAUUUUAAUAGGAAAUAAA AAAAAUUUAGCUCAAAAUUUGUUAUCAU AUUCCUUGCAGCUCUAUCCGCAUGCAGC CCACAGGUGUACCAACAUAGAGCGU UUCUGCAGGUGCAGUCAUUAUCACU CAUUCUUAAUGUCUUCGCAACCAUCCA ACUCAUGAACGCAAUUCCUAAUUUCGCU UAACACACAGAAUUCACACUUAUUAUC AACUAGUUAACGCAAAACAAAAUUA AUCGAGCAGCUCUGGCGCAAGCACCAU CACUUCGCUACAUUAGUUGCAGUCCG ACCAAGAAUUAUCAGGUACCAUUCUUGU CCGGACCCUGCCGUAACCAUGAAGG AUUGUGUCUAGUAGUUUCCCAUCU UAAGAAAGAAUUAUUAAGUUAUUGCCUG UUAUUUGCACACUCUUCACCUUGCAUG UCAUUGUCUUCUACAGUAGUUGCAU UUGCAUGCAAAGCUUAGCCUCGCUU CUUAUAAUAGUCUUCGCUUCUUCUUCU UAUACCAAAGAAACACAGUCAAAA GCAUAGCAAAUAGGUGAUAGUAGUC UCCUUUAUUAUUUUAUCUUCUUCU CCUUUCUUAUUUGCGUUCUUAUUUGU GUGUCAGCGCUAUGGCAAGAAUA UGAGCGUUAUCUCGACAGCAUACGAG CAUUCGCAAGGCGAUGCGAUGCGAAG CGACAGAGAAAGGAGCAAAGCAGUG UGAACAACGCUUGUAAAGGAGUACAG GAGCAGCAGGACAGCAAGAAAGAGC UUCAAGGCAAUACAGCAUUGCAAGG GCGUUGUCAAGCAACAACAGGCGAG GCGCAGCAGCAGCAGUCCAGAGAAAU GCUGGAGCAUUAUAGGAACAAGAG AGGCGAGCAGCAAUUACCAUUAUCACA AAAAAAUAGGAGCAAGAAAGAGGCG CAACAAGCAAAUCCUUAUUAUUUCC UAAAAGAAUCAUUCAAACUCGAUUC CGGAUGAGAGGCAACUUCAAAGUCCU CAGAGGUAUUGCUGAAGUCUCCUCA CUCAGGCAUCACAGAUUACCGUUGG CCAUGUUCAAGCAAAUCCCAACUUAU AUUCUUCGCAACACUGUGAUGCUGAGG CAUUUAUCUCGACAAACGGUAAUUAU UCUCUCCUUUCGAAACAAAAUUUGCGC UUUUAUCACCAAAUACCAUUAUUAU ACAUAGAAAUUAUUCUGUGUAGGAA GGGGCAAAUUCGUUGUGACUCAUGAA AACAAAGGUCCUUAUUAUUAACAGCGU GAAACAGUAGUCAGCGUUCUGCAGGAG CACUGUUUAUGUUGUAGCCAAACAAC CAAGGAAGCUAACCAUAGCUGGUCUG CCGGCGUUAUUAUUCUCCGCAAAUUA UAGGUUUUCUACUUAACAUUUUAUUA UAGUUAUCUAUUAACUUGAUGUCU UGUUAUGUCUUAUAGCCUGUACUUAU AUUGCUGCAGUUAUUCUCCCGGUGGA AAUAAUUAACUGAAUUAUUAUUCGAGC CUUCAGCUAUGAAGUUCUUGCGCGUC UUCAUUAUUAUUGGUUUUUAUUAUUA AUGAAUUAUUUUUUAUUAUUAUUAU AUUCAUUUAUUAUUAUUAUUAUUAU AAAGAGAGAGCUGGAGGAGUUCUUGA GGAACAGAGGCGCAGAGAGGCGAGCAG</p>	<p>SLYPEIFHNFHLNKGK.KKIT SKFVHIISLIAMHANCCTT. AVDSAECSLSTSLVLP LNETHS.FALTHIPTL.LN.V HAEIN.STSSAAKHHFAT SSCSDCFHGTNACP RTT.RDVASRFPHLKRCSI CPVNLHTRHLACCLLHVVG DLHAKA.PPSINTLPLFALITK KKTIKSIANMVKISPFVILF SLLSFALLCSGVSAYGRKI YERDPRUYECCRCESE ATEERECECECEREYK EDRDCEELRKYDCC GPCREKCFR KWEYKEERGEHENYHN HKNPSEEEEGRRNPPYY FPKRRSFRDEEGNFKI LRFAENSPPLKINDYLA MFEANPTFLPHHCKDEAYI FVTNGKFSSLSNKILLLSPN HH.LKT.IISVSGKTVFVTHE NKESYVRRGTVVSPAGST VYVYVNDREKLTIAVLALPV NSPGKYEVFYLTFFY.KFLHNL DVLVSLLLTYL.LLFFPAGN NKPESSYGFAYSVELVTFVN VYKSLTYESIFHCFLPFCSLL IYLFHFYLVCTREKLEILE ERGRKRGGMGMFKA KPEIRAISSATSPYHRGG ERLAINLLSPVSNRGR FFEACDFSNMVAVS AFKLNVLKIF.FSISLKFSS CLLRLD.LHEYKLMNPFAGS HICATL.F.GYIRGCHGRW VRNGLPASLTPGIPKWK ARKTRRRVRCDIWRIPA GSPVITWRLCSPGRHAVT FFSKDP.LNAVAFGLNAN NRIFLAGPFFLNHNNTN VIKFTVKASAY.LINWFLFSI GKKNLVRMDSEAKELSFV PSKLVNIFNPNDESYFMSF SORRGRGDERRGNPLASLD FARLF.AAASFVSDMRHV M.ISWPMRGESLLSCLSNL VFELVNF.CNESYLSKLMN KRPGLNFPNTSSPIKHLS NVRWLVSCLAKVPV.IA.FL. SFLYELEFPFLSFT.TKMSYIM HELHCFNRKAY.TIISFG .ES.SIIKC.PFNCKESKVYV NVSS</p>

	<p>TTGTTCTAAGCAGCTGCTTCCACTTTTGTATCAGACATGCAGA GGCATGTAAGCAAATAAAGTTGGCCTATGTAAAGAGGAGA GAGTTTGTCTTTGTCTTTGCTTAAGCTTTGTTTAACTAGTAA AGTTTCAAAGTATGAGAGTTGTTATCTTTCTAAGTTAATGAAT AAAAGACAGGGAACTCCTTTTCTTAAACAAATTCTC TATAAAACATCTTTTGTCCACGTTTCGATGGCTAGTGTCAAA GTTTAGCCAAAGTACCCTTTAGATGCTTAAATTTTGTAACT CTTTCTTTATGAAATTTCTTTTCTTTGTTCACTTGAAC AAAGATGTCATACATAAATGCATGACAACTACATCAGCAATCA ATAGATGAAAGGCCTATTAGACCGCAATCATTTCATTTGGATG AGAAGTTAATCAAATATTAATAAATAATGCCTTTCAAATGTA AAGAATCAAAGTTGTGATTTAATGTATCCTCTGC</p>	<p>GGCAGCAGGGUAUGUUCGGAAAGCCAA AACCAGAGCAGUAAGCAAAUAGCCAA CAAGCUACUUCUCAAAGGCAAGAGGCG GGGACUUGCCAUCAAUCUUAUUGAG CCAAUCGCGUGUCUACUCCAAACAAAGG ACCGUUCUUUGAGGCUUGUCUGAGGA CUUCAGUCAAUUUAGAAACAUUGAUGUC GGUGUUUCAGCCUUCAAAGUGAAUCAGG UACUUAAAUCAAAUAGUUUUUUCAAUA AUUUCAAUAAAUUUUUUCUUCUUGUCUACU CAUAAGCUACUAGAUAGCUUCAUGAAU AUAAGCUCAUGAAUCUUUUUCAGCGAG CCAAUUUUGGCCACACUCAAUUCUAAG ACUACAUUCGUGUGUUUGUCAAGGAGCG GAUUGGGUAAGCUCAAUUGGCUUGGCC GCAUCUCUCAGACAGAGCAGGGAUCC CAAAGUGGAAGGCAAGCAGAGGAAACA AAGAAGAGGUCAGAGAGGAGCAUUU GGCAAUUCAGCAGGUCAAAGCCCAU UGUCACCUGGUGAGCUUUUGUAGCCCG GGCAGCAUGCAGUUACAUUUUUGCAU CCAAAGCAGCCCCGAAUAGCAGUUGCG GUUUGGACUACAGCCAGAACACCCAG GAUUUUUCUUGCAGGUAGGCCUUUUU UUCUGAAUCAAGCAAACCAAUUGUC AUCAAUUCACUGUAAAAGCCUCGGUUA UUAAUUGAUCAAUGGUUCUUGUUUGAU UCUAUAGGAAAAAGAACUUGGUAGCA AAUGGAUAGCAGGCAAGAGAGUUUCA UUUGGGUACCAUCAAUUUGGUAGUA AUAAUUCACAAACCCGGAUGAGUCGU UUCAGUCUUUCUCUACACAGAGGAGCG GUGGAGUAGAAAGGAGGGCAAUCCCU GGCCUCAAUCUGGACUUUGGCCCGUUG UUCUAAGCAGCUGCUUCCACUUUGUUA CAGCAUGCAGAGCAGUUAUUGCAAUA AUAAUUGGCCUUAUGUAAAGGAGAGGA GUUUGCUUUUGUCUUGUUCUACCUUGU UUUUGAACUAGUAAACUUUCAUUAUUG AGAUUGUUAUCUUUCUAAAGUAAUAGAA UAAAAGCAGGGAAUCUCGCUUUUCU AAUAAGGUUCUCCAAUUUAAACAUUCU UUUGUCACAGUUCGAGUUGGCUAGUGCA CAAUGUUUAGCCAAAGUACCAGUUUCG UUGCUUAAUUUUUGUAAUCCUUUCUUA UGAAUUCAAUUUCCUUUUUCUUGUUC ACUUGAACAAAGUUGUCAUCAAUUGCA UGAACAAUCACUACAGCAAUCAAUAGAU GAAAGGCCUUAUAGCGCAAUCUUAUUA UUUGGAUGAGAAAGUUAAUCAAUUAUAA AUGAAAUGCCUUUCAAUUGUAAAGAAU CAAAGUUGUUAUUUAAUUAUCCUCU GC</p>	<p>DLGVKHNHY.Y.INISFKCFY. KIIVCLSLLLHLFTNNALRIRI YLDMTNITVLIYHLCGDELGN YIFFN.GL...K.IIYIIFLF.YIK.SI VYVK.TDK.INN.MTYIYNYFI Y.VYYSYIVDG.IDR.LNDVY VELHFSHDKYKPMHF.VVY GVICMS.SWSLCCR.HLLKF SS.FINSYTHRR...SILLFH KNV.IH.TKHLHPSNVLFIL PL.FCTF.HTSICMAHG.LL VHVNYIVFFISINTLISLYFIHP ERP.KISFN.HGCYYN</p>
<p>7339552 emb AJ27 6876.1 Pisum sativum vclc gene for vicilin, promoter region</p>	<p>GATCTTGGTGTGATAAAAGATATAATACTAATATTAATTAAT ATTAGTTTTAATGTTTTTATTAATAATAATGTATGTTATCCT TATTACTCCATCTTCTTTACTATAAAGCCTGAGAAACGG ATATATTTGACATGACAAATATACCGTTTATAAATTTATCAT CTTTGTGATGGAGAAATATGTAACTATATTTTTTAAATGAGG ATTTGTAAATAAAGTAGATATATAATATATATTTATTTTTCTTTT TAAATATTAATAAAGTATAGTATATGTAAGTAACCGATA ATAATAGATAATAATAATAATGATGACATATATCTCAATACATTTT TATATATTAAGTATATATAGTATATCAAGTACAGCGATATA AATAGATAGATAATATAATGACGATATGTAGAAATACATTTTTTC ACATGACAAAGTACAACCTATGCCTTCTAAGTGCAAGTTTAT GGAGTTATTTGCATGTCTTGAAGCTGGAGCTTGTGTTGAGGA ACAAACTTGTAAATTTCTCTAGTCAATCATTAAATCATATA CACATGGCCGAGACAAATAAAGCATCCTCCTTTCCATA GAATGTCAAATTCATCAAAATCAAACAAAATCCACCACC GTAAATGTTCTTTTCTTTTGCCTTCAATTTGTACATTTTAAAC ACACGTCATATGCATGGCAACATGGCCAACTGTTGGTGC ATGTTAATATATAGTTTTATTTTTATATCTATAAATACACTCAT CTCACTGTACTTTATTCATCAGACGCAACAAAGTGAATATT AGTTTCAAACAATGAGCTGCTACTACAAATGA</p>	<p>GAUCUUGGUGUAGUAAACAAUAAUUA CUAAUUAUAAUUAUUAUUAUUAUUAU GUUUUUAUAAAUAUUAUUAUUAUUAU UUCUUUAUACUCCAUUCUUCUUAUUA UAAUGCCUGGAAUACGGAAUUAUUAU GACAUGACAAAUUAACCGUUUAUUAU UUAUCACUUCUUGUGAUGGCAAUUAUG AACUAAUUAUUUUUUUAUUAUUAUUA AUAAAAGUAGUUAUUAUUAUUAUUAU UUUUUCUUUUUAUUAUUAUUAUUAU AUAGUUAUUAAGUAAACCGAAUAAU AUAGUUAUUAUUAUUAUUAUUAUUAU ACAAUUAUUAUUAUUAUUAUUAUUAU UAUAGUUAUUAUUAUUAUUAUUAUUAU AAUAGAUUAUUAUUAUUAUUAUUAU UJAGAUUAUUAUUAUUAUUAUUAUUAU AAACUUAUUAUUAUUAUUAUUAUUAU UGGAGUUAUUAUUAUUAUUAUUAUUAU AGCUUGUUAUUAUUAUUAUUAUUAUUAU UAAAUUAUUAUUAUUAUUAUUAUUAU UAUACACAUUAUUAUUAUUAUUAUUAU CAUCCUUAUUAUUAUUAUUAUUAUUAU UUAUUAUUAUUAUUAUUAUUAUUAU CCAAUUAUUAUUAUUAUUAUUAUUAU UUAUUAUUAUUAUUAUUAUUAUUAU UUAUUAUUAUUAUUAUUAUUAUUAU</p>	<p>DLGVKHNHY.Y.INISFKCFY. KIIVCLSLLLHLFTNNALRIRI YLDMTNITVLIYHLCGDELGN YIFFN.GL...K.IIYIIFLF.YIK.SI VYVK.TDK.INN.MTYIYNYFI Y.VYYSYIVDG.IDR.LNDVY VELHFSHDKYKPMHF.VVY GVICMS.SWSLCCR.HLLKF SS.FINSYTHRR...SILLFH KNV.IH.TKHLHPSNVLFIL PL.FCTF.HTSICMAHG.LL VHVNYIVFFISINTLISLYFIHP ERP.KISFN.HGCYYN</p>

<p>TTCTGCTGGTCAAITGGTTATTTAGTTAACCGAGATGACGA GGAGGACTTCTAGAGTACTAGATCTTGTATCCCGTAAATAGA CCAGGTGACCTCAGGTAAATATATCAATTAATCTTTCTAATAT CAACTAAATGATGCATGCAATGTCATAAATAGATTTTGGAGAT CTTTTCATACATCAACAACAATATTACTATGTTTATCCACA TTTTTCTCAAITTTTTCTATTGCAITCTTATAAITTTTTTCCCT TCAACAGTCTTTCTATTGTCTGGTAAACAAAACCAACCGTCCA TCTTATCTGCAITCAGCAAGAATTTTTAGAGGCTTCTCTCAA GTAAGCATAAACAAACATTTTCTTCTTTTTCATGTATTAGTTT CATTACTCTTTGTAAATATTCACTTTGTCTTGTATATATTGTA AATATTGTAGACCGATATAAAGAGATAGAGAGGTTCTCTTA GAAGAATGATGGAAAGAGATATCACAGGAGAGGCCTTAA GATAGGAGGCAAGCGGGCCCAAGAAGAAATGTAATAGTCAA ATATCAAGGAACAATTTAGGAAATGAACAAAATGCAAAGT CTAGCTCAAAGAATGATCATCTTCTGAACTGAAACCGTTCAA TTGAGAGTCTGAGCCATCTATTTCAACAAGTTTGGTAA TTTTTGAATCACCCCAAGAATATCCAACTTCAAGATTT GAAATATTTTGTCAATATGTGGAAATTAAGCAGGTATATCAA ATATCTTATAACATAAACAAATAAATCCCAACAAGTTCTATA TAAATGATTCTTAAATAAATAAAGTATAGTCAAATACAAATG TTCTAAAGGATTAATAATTTTGTCCATTGAAGTTATTAATGCAA CTTTTATTAATCACTCATTAATAAATAATTTATATACCAGATAA CATGTTTTAAATAGAACATATTTGCTAAATAATGATATTTGTTA TTCTCTAAATTTAACTTATAGGATCTCTTTTGTGCCACACTA CAATTCAGGCGCCATAGTAAATGTAACAGTCAATGAAGGAAA GGGATTTGAACTTTGTGGTCAAAGAAATGAAACCAACAG GGCTGAGAGAAATATGACGAGGAAAGGAAACAAGGAGAA GAAGAGATAAGGAAACAATGTCAAATATCAAAGCTAAGTTGT CTCCAGGTCAATGTTTTAGTAAATCCAGCGGGTTATCCAGTTGC CATAAAGCTTCAATCTCAAATTTGGTGGATTTGGCATCA ATGCTGAGAACCAACAGAGATACTCTTGTGAGGTATAAATATA TGAAATTAAGTCAAGTGTTTAAGTCACTTATTTTATTAATC AAAGAAGAAAATAAATAAGATGTTTACTTTTATGATTTGA AAATTAAGGTGAGGAAGCAATGTATAAGTCAACATACATA AACCAAGTGAAGAACTTGCATTCCTGATCAGCTCAAGAGGT AGATACCGTACTAGAAATCAGAAACAATCTCACTTTGCAAT GCTCAACTCTGAGAAAGGAAAGAGGAAGCCAGAAATAAAG GATCATCTGACTCAATTTGGGTTCTTTTTAAATGATGATCAA ATGTTATGTATGTATGCTTAAAGTACTATAGCTCATAGTGAGC AAATAAATAACGACTCTCTCTTGTAACTATTAATTAATCCCA CCTTTCTACTAAGAAATAAATAAAGAAATTTGTGTGTCTCC TCTTAACCTACTATTTTTTTTGTAACTTACATAATAAATAA ATTTTTTTGTTAACTTAAATTTGTTGTAATAATATAAATAA ATGTTTAAATATCTGCTTATACCGTCACTAAAGATAGATAACTTT TTTTCAAAAATAAATAACAACAGTTTGAAGTATAAATATCATG AGACAAATTTGTAATTTGTTGAATAAACCTTCTATAACAATA GTCAAATAAAGAAGATGATTTGTTCAAACGTCACCATTACTAGT GAGGTTGTCTAGACCGTGCACCTATATATTACTCGAGATTTGGG GCCAAAGATTTTTCAAAGGTGGGTTAGAAATAGGCTTCTGACAC ACAGATTTCCCAAATTTGTTGTTCAAATTTGTTGCAACCCAAATAA GCGATAATAGAACCGCACAACTCATAAAAACCTCTCAC AAAA</p>	<p>UUUUUAUUUAUUAGAACAGUGCUUGAAU AAUUAAUUCACAACAUAUUAAUCUUUAU AGUUAGUGUAAAAAGUUUUUUUUUUUAU UAUAAAAUUUGAAUUUUUUUUUUUUUU GAUUCUUUUUUUUUUUUUUUUUUUUUU GCUAACAACACUGCAAAGCAUGAAG CUGCCAAAUUCAUUUUUUUUUUUUUU UGCAUUUUUUUUUUUUUUUUUUUUUU UAGAGGUAGUUUUUUUUUUUUUUUU AGAAUUUUUUUUUUUUUUUUUUUUUU AGUGAUGAACAACAAUUUUUUUUUUUU CUAUUCUUUUUUUUUUUUUUUUUUUU AAUUCAAACAACUCCACCACCACUUUU AUGUUUUUUUUUUUUUUUUUUUUUUUU UUGUACUUUUUUUUUUUUUUUUUUUU UGGAGAUCAUGACCACAAUUGUUGGUGCA UGUUCAUUACAUUUUUUUUUUUUUUU UAUAAAAUUUUUUUUUUUUUUUUUUUU CAUUCAUCCACAUUGAAAGUUUUUUUU UUAAUUCAAACAAAGUUGCAGCUACCAC AUUGAAAUUUUUUUUUUUUUUUUUUU UUGUUGGAAUUCGUUUUUUUUUUUUU UGUGUCUUUUUUUUUUUUUUUUUUUU UAAACCUUUUUUUUUUUUUUUUUUUUU UUUCAACUCUUUUUUUUUUUUUUUUUU UCACAUUCGUUUUUUUUUUUUUUUUU AACAUUCACAAUACUUGGAAUUCUCCAA AAUUACCGUUUUUUUUUUUUUUUUUUU ACUCACACCAUAAUUUUUUUUUUUUUU GUGACGCGGAAUUUCAUUUUUUUUUU UAUUUAAAACAUAUUUUUUUUUUUUUU UAAAAACAAGGUUUUUUUUUUUUUUU UGUUUUUUUUUUUUUUUUUUUUUUUU CAUUUUUUUUUUUUUUUUUUUUUUUU UUUGCAGAAAAGCUAUACUCAAGUGU UAUUACCAACAGCAGAACUCCUUUUGU CUUGACCGAGGAAUACCAUAAAACUUU CUGCUGGUACAAUUGUUUUUUUUUUUU CCGAAUGACGAGGAGGAAUCUUUUUU CUUGAUUUUUUUUUUUUUUUUUUUUU CAGGUGAACCUUAGGAAUUUUUUUUUU AUACUUUUUUUUUUUUUUUUUUUUUU UGCAAUUGCAUAAAACAAUUUUUUUU CUUUUCAUAAAUACAAACAACAAUUUU ACUAUUGUUUUUUUUUUUUUUUUUUUU UUUUUUUUUUUUUUUUUUUUUUUUUU UCCUUUCAAACAGUCUUUUUUUUUUUU GUAAUCAAACCAACCGUCCAUCUUUU GGAAUCAGCAAACAACUUUUUUUUUU CCUUCAAUGAAGCAUAAACAACUAAU UCUUCUUUUUUUUUUUUUUUUUUUUUU UACUCUUUUUUUUUUUUUUUUUUUUUU UGUAUUUUUUUUUUUUUUUUUUUUUU UAUAAAGAUAGGAGGUUUUUUUUUUU AGAGCAUUGGAAAGGAAAUUAUCAAGG CAGGCCUUAAAGAUAGGAGGAGCGGGG CCAAAGAGAAAUUUUUUUUUUUUUUU CAAGGAAACAACUUUUUUUUUUUUUU AAUGCAAAGUCUAGCUCCAAACAAAGUAC AUCUUUCUGAAUCUGAACCGUCCAUUUU AGAGUGGUGAGCCCAUCUUAUCCAAAC GUUUUGUAAAUUUUUUUUUUUUUUUU AGGAAAUCCACAAUUCAAAGCUUGAA UAUUUUUUUUUUUUUUUUUUUUUUUU AGGUAAUCAAACAAUUUUUUUUUUUU ACAAAUAAACUCCACAAAGUUUUUUUU AAUUAUUUUUUUUUUUUUUUUUUUUUU UCAAAUUCAAUUUUUUUUUUUUUUUU AAUUUUUUUUUUUUUUUUUUUUUUUU ACUUUUUUUUUUUUUUUUUUUUUUUU UUUUUUUUUUUUUUUUUUUUUUUUUU AUUUAACAAUUAUGCUAAAAGUAAUUUU UUUUUUUUUUUUUUUUUUUUUUUUUU CUCUUUUUUUUUUUUUUUUUUUUUUUU CGCCAUAGUGAUAGUAAACAGUCAAUG GAAAAGGGGAUUUUUUUUUUUUUUUU</p>	<p>CSTSRRKRRKPKRNGKSSVLN SANNKTLSCNYYNNTSTFLF GFFLMSDMLCMYALKYYS RINKLMFVCSLLFFVNT INYYFFVNIKPVKYYKNCLIS AYTVTKDRLFFKLDQFEL YHETICNCNKPS.TCKLKE DDCYNRHHY.WLSLPVHY LLEIGAKDFSKVGNRLLTR FRNCCSNCCNPAAITATTLL NSH</p>
--	---	---

		<p>CAUACUCAAUAAUGUUUUUUUUUUGA ACAUUAUCUAAUGUAAUUUUUAAUUAU UCUUUCAAUUUAAUUUUUAGGGAUCUUUA UUGUUGCCACACUACAAUUCAGGGGCCA UAGUAAUAGUAAAGUUAAAGGAAGAAA GGAGUUUUUAAUUUGGGUCAAAGAA AUGAAACCACAAAGCAGGAAAAGAA GAUGACGAGGAAAGGAAACAGGAGAA AGGAGUAAAUAACAAUGGCAAAUUAU AAAGCUAAAUAUCUUACAGGAGUUGUUU UGUGAUUCAGCAGCCAUUCAGUUGCC GUAAAAGCUUCCUCAAUCUUGAUUUGC UUGGUUUUGUAUUAAUGCUGAGAACAA UCAGAGGAAUUUCUUGCAGGUUAUUA UAUUAUCACCAGUCUCUGACUUAUUUA UUCAUUUUAAGUGUGUUAUUUUAAAGU GACUUCUUAUUAAUCAAAGGAAAUAU UAAGUUAUGCUUUAUUUUUGGUAUUA AAAAUUGAAGGCGAUGAGAAUUAUG AUUAGUCAGUACAGCGACAGUGAAAG GCUUGCAUUCUUGGAAUCAGCUCAAAG GUUGAUAGGAUACUAGGAAUAGAACAA AUCCACUUUGCAGUUCUACCCUCAAG AAAGGAGGAGGAAAGUCUAAAACAA AGAUAGCUUAUCUUAGUUUAAAUGU UCUUAAUGAGUGGACAAAUAUUAUUA GUAUGCUUAACAAGUAUAUCUCAAGG GGAGCAAUGAAUAAACAUAUUAUCUUA UAAUUAUUUAUUAUCCACUUUUUUAU UAUGAAUUUUAAUCAAAGUUAUUUGG UGUUCCUUUAAGUAUUUAUUUUUA GAAGUCAAUAAUAAAAGUAACAACAC UCAUUAAAUAUAAUUAUUAUUAUUA AAUAUCUAUCCAUAGAUUAUUAUUA CAUUAUCAUCAAUUAUUUAUUAUUA</p>	
<p>1019832 emb Z543 65.1 Z. furfuracea gene for vicilin</p>	<p>GGCAGATTGCTTTATAAGTCAAGGAGTTTGCATTGTACA GTAAAGATGCCTGCTTGTGATGATGAAATCGATTGTATG TGATTGTGAGTGTACATTCCACAGGAAGAGGAGGGTTGCG TTTGTTTCATGAGAGAGAGTGGTAGAAGGCAGCTGCGGGAC GGAGATGTGTACCGAATTGCTGACAGGTATACCGTTTTATTC TCAACAGGGATGACAGTCGGCGCTTTTCATTCACTGTCTCCT GCCACCGAGTGTCTACTACCGGACTCTATGAGGTACTCATT CACACTCACTCACTAGTGTGCTGATGATCTGAACTAAC CGTGTACCGTTGGGCATGACAGTCGTTTTACCGTGGTGGAG GCCCGAAACCGCAGATGTTTTGTCCCAATCAGCGAGGACG TTCTGCAAGCTGCATTCAACGTAACCTTGTCAATTTTCA TTGATATTATTGCAACCGCGGCTTCGTTAATTTACTCTAAC CTAAGTCAAAGCTTGAATGAGCAGAGTTCGAAAGGCGTCCCTG GATCCTATGCTGTAAGTGGGTTAACAGAGGGGCCATTAATA CAGTGTCCAGAGCAACAATGGAAAGGCTGAGCAGGGGTAGAA TTAGGGATTCGGAGGCTGAGGAGCCACAGCCCTTCAAC TGCTCTACAGGAACCCGACTTTTCCCAACAACAACGCGTGA TTTCAACAGCAGACCGCCGAGATCATCGCGTTTTTACCGCGGCT GAATGTTGGAGTGCAGCTTCTCAACTTAAAGCCGGTAAGCTCA AACACAGCCAGCCAGAGTTCTAATTAATTAATAGGACCGTTAC AATTAGCATTACAGTATCGTCTGTTTGGAGGAGATGATTATA TTGCATTGGGAGCGCTCAATGACGGCACCGCATTACGATAC AAGGTGCAGCAAGAAATGGCATCGTTAGGAACGGGAGGGAA CCTCGAATTAGTGCAGCCCGCAGAAACAAGAAACAACAACA ACAACAAGGTCACATACCAGAAAGTACGGGCCAACTCAA CCCCAGGACTGTTTCTTACCCCGCCCGGCTACCTTCCAC TGTAAATGCCTCTGGCAAAGGCAATTGAGATATTGTACTT GACATTAAATCCCAAGGCAATCGCAGGCAATCTCTCGCAGGT GAGATGTTCTCCCGGAAATCAGTCTTTTCAAATTCGTTAGT ATAAATAGTAAACCAAATACCCAAAGCAACGCCATCGCAGGAAG GAGCAAATGTGTGAGATATCTGCC</p>	<p>GGCAGUUUUGCUUUAUACGUCACAGGA GGUUUGCAUUGUAAGUAAAGUAGCCUG CUUGUAAUUGAAUUAUUAUUAUUAUUA UUGAUUUGACUGACUGACAUUCCACAGG AAGAGGAGGUGUUGCUUUGUUAUUAUUA GAGGACUGGUAGAGGACAGCUGCGGG AGCCAGUUGUUAUUAUUAUUAUUAUUA UAUACGUGUUUUAUUAUUAUUAUUAUUA GACAGUCCGCGCGCUUUUAUUAUUAUUA UCCUGCGCAGCAGUUCUUAUUAUUAUUA ACUCUAUAGCUGACUUAUUAUUAUUAUUA CACUGCACUAGUGGAAUUAUUAUUAUUA GUAAACCGUGUACCGUUGGGCAUGACAG UCGUUUAUUAUUAUUAUUAUUAUUAUUA CGCAGAAUUGUUUUGCCCAAUUAUUAUUA GGACGUCUGCAAGCUGCAUUAUUAUUA ACUUGUUAUUAUUAUUAUUAUUAUUAUUA UUAUUAUUAUUAUUAUUAUUAUUAUUA UACUCUAUUAUUAUUAUUAUUAUUAUUA GACAGUUCGAAAGGCGUCCUGGAUUAUUA AUGCGUUAUUAUUAUUAUUAUUAUUAUUA CCAUAAUUAUUAUUAUUAUUAUUAUUAUUA GAAAGCGUGAGCAGCGGUUAUUAUUAUUA GAUUCCGAGGUGUCAGGAGCCACAGCC CUUCAACUGCUCUAUUAUUAUUAUUAUUA UUUUUAUUAUUAUUAUUAUUAUUAUUAUUA AGCAGCGCCCGCAGUCAUUAUUAUUAUUA CGCCCGCUGAAUUGUUGGAGUGCAGCUUC UCAAGUUAAGCAGGUAAGCUAAUUAUUA CAGCCAGGUUCUUAUUAUUAUUAUUAUUA ACGCUUAUUAUUAUUAUUAUUAUUAUUA CUGUUUAGGAGGAGUUAUUAUUAUUAUUA AUUGGCGAGCGCUAAUGAAGGACCGC AUUAAGUAACAAGGUCGACAAUUAUUAUUA CAUCGUUAGGAAACGGAGGGGAUUCUUA GAUUUAGUGCCCGCAAGAACAAUUAUUA ACAACAACAACAACAAGUCCCAUUAUUA AGAAACUACCGGCCAACCUAAUUAUUAUUA CACUGUUUUCUUAUUAUUAUUAUUAUUAUUA CCUUCACUUAUUAUUAUUAUUAUUAUUAUUA AGGCAUUAUUAUUAUUAUUAUUAUUAUUA</p>	<p>GRFGFIRRRFALYSKHAFL MNEICDCDCDIPEEEG LRLFMKRDW.KGSCGTEMCT LLVYRFIFSTMTVGAFSF TVSCARSALLPDSMRYSF THCTSAE.ITNRVPFGHDSRF TWWEAARRMFCPNSAHTF TNGKAE.G.N.GIRRV.GATAL PALPEPLFDGR.NFHSR RRRSSRFTPAECWSAASL KAGKLNAS.SSN.LIRTLI.H SVSSV.EEMINIALGSA.RHR ITIGRELASLGTGESSN. CAKKNKNNNNNNKVPHTRN YGPTSTPALFS.PAPATPLPL PLAMRHCYCTSTLIPKAIAG NSSMRCSSRNSVFSNFVSD NSNTKYPMNAIAGSNVLR L</p>

		UAAUUCCAAAGCAAUCCAGGCAAUUC CUCCAGGUGCAUGUUUUUCGGAAU UCAGUUUUUUCAAUUUCGUUAGUGAU AUAGUAAUACCAAUACCCAAUAGCGC AUCGAGGAAGGAGCAAUGUGUUUGAU AUCGCC	
6469914 gb AF113 059.1 <i>Herrania nycterodendron</i> vicilin (vic) gene, partial cds	TGAAGAGGGCAACTTCAAGATCCTCCAGAGGTTTGTGAGAA CTCTAGTCCACTCAAGGCGATCAACGATACCGCTTCGCGCATT TTCGAAGCAAACCCAACTTTTGTCTTCCACACCACTCTGA TGCTGAGGCAAATTACTTTCGTGACAAAGGTAATCTCTTCC CTTTTGAACAAATTTTGGGCTTTTTTTCACAAATCATCATTG ATTGAAATCATGTAATATTTCTGTGTGAGGAAAGGGGACAA ATACGTTTGTGACTGATGAAACAAGGAGTCTATATGTAC AGCGTGGAACCGTAGTCAGCGTTCCTGCAGGATGCACGTGTT ACGTTGTTAACCAAGCAACAAGAGAAGCTAACCATAGCCGT GCTCGCCCTGCCTGTTAATACTCCTGGCAATATGAGGTTTTT TACTTTAATTTATTTATATAAGTTTCTACATATCCTTGTGT CTTGGTAGTTTATTGACCACTACTTATATATGCTGAGGTGT TCTTCCCCTGGTGAATAATAAACCTGAGTCATATTACAGAGC CTTCAAGCTGTGAAGTTCTTGTAGGCGCTTCAAATGTAATGTT TCTCTGACTAGCTATGAAATCAATTTTTGTTTTTACTTTCTCTG TCTTTACTTTATTTATTTATTTATTTATTTATCTCGTTGTTAG ACACAAAGAGAGAGCTGGAGAAGATCTTTGAGGAACAGAGA GGGCAGGAGGGGAGCAGGATGTTTCCGGAGAGCCAAACC AGAGCAGATAGAGCAATGAGCCAAACGCTACTTCGCCAGG GCAGAGAGGCGGGGAGATGTTGCCGTCAACTATTTAGCCA ATCGCCTGTCTACTCCAAACGAAACGGACCGCTTCTTTGAGTT TGTCCTACTGAGGCCTTCACTCAAATTCAGAACATGATGTC CTGTTGCAGCCATCAAACTGATCCGGTACTTAAATCAATG ATTTTTTCAAATAATTCATTAATAATCTTTCTTGTGACTCAT ACTACTGCCAGATTAGCTTCATGATGCATGCTCATGATCCT TTTGACGGGAGCAATATTGTGCCA	UGAAGAGGGCAACUUCAASAJCCUCAG AGGUUUGCUGAGCAUCUAGUCCACUCA AGGCAUCACAGAUUACCGCUUCGCAU UUUCGAGCAAUCCCAACACUUUUGUUC UUCACACCACUCUGAUGCUGAGGCAAU UUACUUUGUACAAAGGUAUUUUCUCU UCCUUUUGAACAAUUUUGCGCUU UUUCACAAAUCAUUAUUAUUAUUA UGUAUUUUUUUUCUGUGUCAGGAAGG GGCAAUUAUGUUGACUGAUGAAA CAAGAGUCCUAUUUAUGUACAGCGUGGA ACCGUAGUCAGGUGUCCUGAGUGCA CUGUUUAUGUUAACCAAGACAACA GAGGAGUAACCAUAGCCGUGCUCGCC UGUUUUUUAUACUCCUGGCAAAUUA GGUUUUUUAUUUAUUUUUUUUUU AAGUUUUUAUUUAUUUUUUUUUUUU UUAGUUUUUUAUUAUUUUUUUUUU GCUGCAGUGUUUUUCCCGUGGAAU AAUUAACUGAGUCAUUUAAGGCUU CAGCUGUUAAGUUCUUGAGCUGCUU AAUGUAAUUGUUCUCUGACUAGCUAUG AAUCAAUUUUUUUUUUUUUUUUUU GUCUUUUUUUUUUUUUUUUUUUUUU UUUUUUUUUUUUUUUUUUUUUUUU GAGAGCUGGAGAAUUCUUUAGGAACA GAGAGGAGGAGGAGGAGGAGGUAU GUUCGGAGGCCAAACAGGAGUA AGAGCAAUAGCCAAACAGCUACUUCG CAGGAGAGGAGGCGGUGUCUUG CGUCAAUCUUAUUUAGCCAAUCGCUUG UACUCCAACCGAAACGAGCGUUCUUU AGGUUUUUGUCCUACUGAGGCUUCAGUCA AUUUUAGAAACUAGGAUGCGUUGUCA GCAUCAAACUGAAUCGGUACUUUUUU CAAUUGAUUUUUUUCAAUUUUUUUU AAUCUUUCUCUUGUACUACUACUACU GCAGUUAGCUUUAUUAUUAUUAUUA AUGAAUCCUUUUGAGGAGCAAUUUU GUGCCA	RGLLDPPEVC.EL.STIGH RLPLRHFBSKSHFCSSTP AAFTKSSLIENMNVYFCV L.C.GNLLFDK.GILFPFENF AAFFTKSSLIENMNVYFCV GDNVYCD.KGVLCTAWN SRSRMRHCLRGPRRTREA NHSRARPAC.YSWLIGLLH LFIIISFYLSMSWL.VY.PRTYN CCRCSSPLEIINLSHITEPSAV KFLRPPSSM.MVL.LAMNFFV FTFSCLLLIYLFHIFYLVC.TR EKLEKIFEEFRGEGCGMF RRAKPEIRAMSQATSPG RRGGESLAVNLFSPVYSN RNGRFFEVCPTEAFSFIN MDVAVAAIKLNPLVKIFFSII SLKFSFVLLILLPLHECMMLN NPFAGSNICA

Codons analysis

Targeted amino acid's codons:

- Asparagine or Aspartic acid (Asn) D : AAC, AAT, GAC, GAT
- Glutamine (Gln) Q : CAA, CAG
- Arginine (Arg) R : AGA, AGG, CGA, CGC, CGG, CGT
- Proline (Pro) P : CCA, CCC, CCG, CCT

Following equation was used to find the Occurrence ration of Codons Amino acids:

$$\text{Codon Amino acid occurrence ratio} = \frac{\text{No. of Codon Amino acid}}{\text{Total no. of nucleotides}} \times 100$$

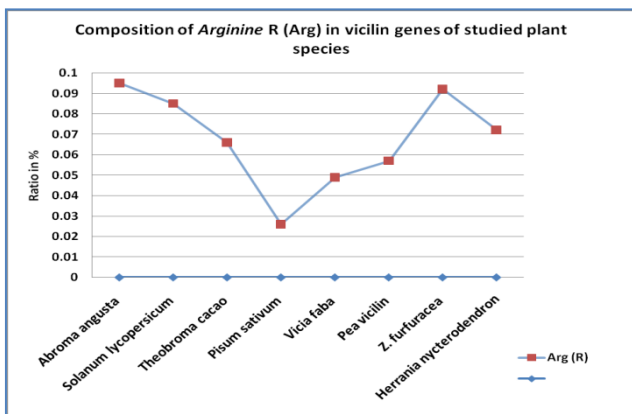
Table-2 Occurrence nature of gene codons in DNA sequence of Vicilin genes of studied plant species

Gene	Codon	No. of codons	Occurrence ratio (in %)	No. of nucleotide	Plant species
Asparagine or Aspartic acid (Asn)	AAT	31	2.81%	1102	<i>Abroma angusta</i> Vicilin gene, partial cds
		46	2.33%	1973	<i>Solanum lycopersicum</i> Vicilin gene
		99	2.98%	3317	<i>Theobroma cacao</i> vicilin
		45	5.14%	875	<i>Pisum sativum</i> vclc gene vicilin
		260	4.88%	5327	<i>Vicia faba</i> Vicilin gene
		176	4.15%	4239	<i>Pea</i> Vicilin gene
		40	3.06%	1304	<i>Z. furfuracea</i> gene for vicilin
		30	2.71%	1105	<i>Herrania nycterodendron</i> vicilin
		24	2.17%	1102	<i>Abroma angusta</i> Vicilin gene, partial cds
	AAC	64	3.24%	1973	<i>Solanum lycopersicum</i> Vicilin gene
		59	1.77%	3317	<i>Theobroma cacao</i> vicilin
		12	1.37%	875	<i>Pisum sativum</i> vclc gene vicilin
		110	2.06%	5327	<i>Vicia faba</i> Vicilin gene
		82	1.94%	4239	<i>Pea</i> Vicilin gene
		28	2.14%	1304	<i>Z. furfuracea</i> gene for vicilin
		20	1.80%	1105	<i>Herrania nycterodendron</i> vicilin
		GAC	9	0.81%	1102
	19		0.96%	1973	<i>Solanum lycopersicum</i> Vicilin gene

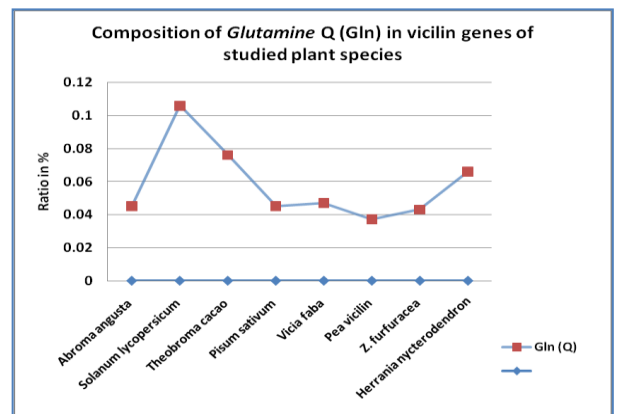
Hydropathicity, Phylogenetics and Possible Codons Analysis for Asparagine, Glutamine, Arginine, and Proline Proteins in Vicilin genes of *Abroma angusta*, *Solanum lycopersicum*, *Theobroma cacao*, *Herrania nycterodendron*, *Z. furfuracea*, *Pisum sativum*, *Vicia faba*, *Theobroma cacao*

		30	0.90%	3317	<i>Theobroma cacao vicilin</i>		
		8	0.91%	875	<i>Pisum sativum vclc gene vicilin</i>		
		32	0.60%	5327	<i>Vicia faba Vicilin gene</i>		
		30	0.70%	4239	<i>Pea Vicilin gene</i>		
		14	1.07%	1304	<i>Z. furfuracea gene for vicilin</i>		
		8	0.72%	1105	<i>Herrania nycterodendron vicilin</i>		
		GAT	9	0.81%	1102	<i>Abroma angusta Vicilin gene, partial cds</i>	
			32	1.62%	1973	<i>Solanum lycopersicum Vicilin gene</i>	
			40	1.20%	3317	<i>Theobroma cacao vicilin</i>	
			13	1.48%	875	<i>Pisum sativum vclc gene vicilin</i>	
			81	1.52%	5327	<i>Vicia faba Vicilin gene</i>	
			66	1.55%	4239	<i>Pea Vicilin gene</i>	
			17	1.30%	1304	<i>Z. furfuracea gene for vicilin</i>	
12	1.08%		1105	<i>Herrania nycterodendron vicilin</i>			
CAA	35	3.17%	1102	<i>Abroma angusta Vicilin gene, partial cds</i>			
	81	4.10%	1973	<i>Solanum lycopersicum Vicilin gene</i>			
	112	3.37%	3317	<i>Theobroma cacao vicilin</i>			
	21	2.4%	875	<i>Pisum sativum vclc gene vicilin</i>			
	162	3.04%	5327	<i>Vicia faba Vicilin gene</i>			
	112	2.64%	4239	<i>Pea Vicilin gene</i>			
	37	2.83%	1304	<i>Z. furfuracea gene for vicilin</i>			
	31	2.80%	1105	<i>Herrania nycterodendron vicilin</i>			
	CAG	17	1.54%	1102	<i>Abroma angusta Vicilin gene, partial cds</i>		
		33	1.67%	1973	<i>Solanum lycopersicum Vicilin gene</i>		
		69	2.08%	3317	<i>Theobroma cacao vicilin</i>		
		1	0.11%	875	<i>Pisum sativum vclc gene vicilin</i>		
		42	0.78%	5327	<i>Vicia faba Vicilin gene</i>		
		35	0.82%	4239	<i>Pea Vicilin gene</i>		
		33	2.53%	1304	<i>Z. furfuracea gene for vicilin</i>		
		21	1.90%	1105	<i>Herrania nycterodendron vicilin</i>		
	AGA	23	2.08%	1102	<i>Abroma angusta Vicilin gene, partial cds</i>		
		65	3.29%	1973	<i>Solanum lycopersicum Vicilin gene</i>		
90		2.71%	3317	<i>Theobroma cacao vicilin</i>			
13		1.48%	875	<i>Pisum sativum vclc gene vicilin</i>			
110		2.06%	5327	<i>Vicia faba Vicilin gene</i>			
84		1.98%	4239	<i>Pea Vicilin gene</i>			
22		1.68%	1304	<i>Z. furfuracea gene for vicilin</i>			
21		1.90%	1105	<i>Herrania nycterodendron vicilin</i>			
AGG		16	1.45%	1102	<i>Abroma angusta Vicilin gene, partial cds</i>		
		43	2.17%	1973	<i>Solanum lycopersicum Vicilin gene</i>		
		54	1.62%	3317	<i>Theobroma cacao vicilin</i>		
		2	0.22%	875	<i>Pisum sativum vclc gene vicilin</i>		
		48	0.90%	5327	<i>Vicia faba Vicilin gene</i>		
		43	1.01%	4239	<i>Pea Vicilin gene</i>		
		31	2.37%	1304	<i>Z. furfuracea gene for vicilin</i>		
		22	1.99%	1105	<i>Herrania nycterodendron vicilin</i>		
		CGA	4	0.36%	1102	<i>Abroma angusta Vicilin gene, partial cds</i>	
			22	1.11%	1973	<i>Solanum lycopersicum Vicilin gene</i>	
21			0.63%	3317	<i>Theobroma cacao vicilin</i>		
2			0.22%	875	<i>Pisum sativum vclc gene vicilin</i>		
26			0.48%	5327	<i>Vicia faba Vicilin gene</i>		
16			0.37%	4239	<i>Pea Vicilin gene</i>		
8			0.61%	1304	<i>Z. furfuracea gene for vicilin</i>		
3			0.27%	1105	<i>Herrania nycterodendron vicilin</i>		
CGC		5	0.45%	1102	<i>Abroma angusta Vicilin gene, partial cds</i>		
		8	0.40%	1973	<i>Solanum lycopersicum Vicilin gene</i>		
		22	0.66	3317	<i>Theobroma cacao vicilin</i>		
		-	-	875	<i>Pisum sativum vclc gene vicilin</i>		
		12	0.22%	5327	<i>Vicia faba Vicilin gene</i>		
		6	0.14%	4239	<i>Pea Vicilin gene</i>		
		22	1.68%	1304	<i>Z. furfuracea gene for vicilin</i>		
		8	0.72%	1105	<i>Herrania nycterodendron vicilin</i>		
		CGG	6	0.54%	1102	<i>Abroma angusta Vicilin gene, partial cds</i>	
			8	0.40%	1973	<i>Solanum lycopersicum Vicilin gene</i>	
13			0.39%	3317	<i>Theobroma cacao vicilin</i>		
3			0.34%	875	<i>Pisum sativum vclc gene vicilin</i>		
8			0.15%	5327	<i>Vicia faba Vicilin gene</i>		
8			0.18%	4239	<i>Pea Vicilin gene</i>		
17			1.30	1304	<i>Z. furfuracea gene for vicilin</i>		
5			0.45%	1105	<i>Herrania nycterodendron vicilin</i>		
CGT		5	0.45%	1102	<i>Abroma angusta Vicilin gene, partial cds</i>		
		16	0.81%	1973	<i>Solanum lycopersicum Vicilin gene</i>		
	26	0.78%	3317	<i>Theobroma cacao vicilin</i>			
	3	0.34%	875	<i>Pisum sativum vclc gene vicilin</i>			
Arginine (Arg)		23	2.08%	1102	<i>Abroma angusta Vicilin gene, partial cds</i>		
		65	3.29%	1973	<i>Solanum lycopersicum Vicilin gene</i>		
		90	2.71%	3317	<i>Theobroma cacao vicilin</i>		
		13	1.48%	875	<i>Pisum sativum vclc gene vicilin</i>		
		110	2.06%	5327	<i>Vicia faba Vicilin gene</i>		
		84	1.98%	4239	<i>Pea Vicilin gene</i>		
		22	1.68%	1304	<i>Z. furfuracea gene for vicilin</i>		
		21	1.90%	1105	<i>Herrania nycterodendron vicilin</i>		
		AGG	16	1.45%	1102	<i>Abroma angusta Vicilin gene, partial cds</i>	
			43	2.17%	1973	<i>Solanum lycopersicum Vicilin gene</i>	
			54	1.62%	3317	<i>Theobroma cacao vicilin</i>	
			2	0.22%	875	<i>Pisum sativum vclc gene vicilin</i>	
			48	0.90%	5327	<i>Vicia faba Vicilin gene</i>	
			43	1.01%	4239	<i>Pea Vicilin gene</i>	
			31	2.37%	1304	<i>Z. furfuracea gene for vicilin</i>	
			22	1.99%	1105	<i>Herrania nycterodendron vicilin</i>	
			CGA	4	0.36%	1102	<i>Abroma angusta Vicilin gene, partial cds</i>
				22	1.11%	1973	<i>Solanum lycopersicum Vicilin gene</i>
21	0.63%	3317		<i>Theobroma cacao vicilin</i>			
2	0.22%	875		<i>Pisum sativum vclc gene vicilin</i>			
26	0.48%	5327		<i>Vicia faba Vicilin gene</i>			
16	0.37%	4239		<i>Pea Vicilin gene</i>			
8	0.61%	1304		<i>Z. furfuracea gene for vicilin</i>			
3	0.27%	1105		<i>Herrania nycterodendron vicilin</i>			
CGC	5	0.45%	1102	<i>Abroma angusta Vicilin gene, partial cds</i>			
	8	0.40%	1973	<i>Solanum lycopersicum Vicilin gene</i>			
	22	0.66	3317	<i>Theobroma cacao vicilin</i>			
	-	-	875	<i>Pisum sativum vclc gene vicilin</i>			
	12	0.22%	5327	<i>Vicia faba Vicilin gene</i>			
	6	0.14%	4239	<i>Pea Vicilin gene</i>			
	22	1.68%	1304	<i>Z. furfuracea gene for vicilin</i>			
	8	0.72%	1105	<i>Herrania nycterodendron vicilin</i>			
	CGG	6	0.54%	1102	<i>Abroma angusta Vicilin gene, partial cds</i>		
		8	0.40%	1973	<i>Solanum lycopersicum Vicilin gene</i>		
13		0.39%	3317	<i>Theobroma cacao vicilin</i>			
3		0.34%	875	<i>Pisum sativum vclc gene vicilin</i>			
8		0.15%	5327	<i>Vicia faba Vicilin gene</i>			
8		0.18%	4239	<i>Pea Vicilin gene</i>			
17		1.30	1304	<i>Z. furfuracea gene for vicilin</i>			
5		0.45%	1105	<i>Herrania nycterodendron vicilin</i>			
CGT	5	0.45%	1102	<i>Abroma angusta Vicilin gene, partial cds</i>			
	16	0.81%	1973	<i>Solanum lycopersicum Vicilin gene</i>			
	26	0.78%	3317	<i>Theobroma cacao vicilin</i>			
	3	0.34%	875	<i>Pisum sativum vclc gene vicilin</i>			

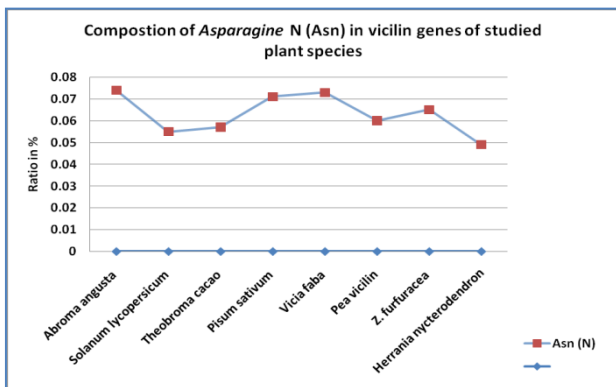
Proline (Pro)	CCA	24	0.45%	5327	<i>Vicia faba Vicilin gene</i>
		10	0.23%	4239	<i>Pea Vicilin gene</i>
		14	1.07%	1304	<i>Z. furfuracea gene for vicilin</i>
		11	0.99%	1105	<i>Herrania nycterodendron vicilin</i>
		21	1.90%	1102	<i>Abroma angusta Vicilin gene, partial cds</i>
		28	1.41%	1973	<i>Solanum lycopersicum Vicilin gene</i>
		54	1.62%	3317	<i>Theobroma cacao vicilin</i>
		11	1.25%	875	<i>Pisum sativum vclc gene vicilin</i>
		63	1.18%	5327	<i>Vicia faba Vicilin gene</i>
		41	0.96%	4239	<i>Pea Vicilin gene</i>
		16	1.22%	1304	<i>Z. furfuracea gene for vicilin</i>
		19	1.71%	1105	<i>Herrania nycterodendron vicilin</i>
	CCC	8	0.72%	1102	<i>Abroma angusta Vicilin gene, partial cds</i>
		9	0.45%	1973	<i>Solanum lycopersicum Vicilin gene</i>
		20	0.60%	3317	<i>Theobroma cacao vicilin</i>
		2	0.22%	875	<i>Pisum sativum vclc gene vicilin</i>
		15	0.28%	5327	<i>Vicia faba Vicilin gene</i>
		11	0.25%	4239	<i>Pea Vicilin gene</i>
		11	0.84%	1304	<i>Z. furfuracea gene for vicilin</i>
		4	0.36%	1105	<i>Herrania nycterodendron vicilin</i>
	CCG	9	0.81%	1102	<i>Abroma angusta Vicilin gene, partial cds</i>
		10	0.50%	1973	<i>Solanum lycopersicum Vicilin gene</i>
		16	0.48%	3317	<i>Theobroma cacao vicilin</i>
		2	0.22%	875	<i>Pisum sativum vclc gene vicilin</i>
		14	0.26%	5327	<i>Vicia faba Vicilin gene</i>
		8	0.18%	4239	<i>Pea Vicilin gene</i>
		15	1.15%	1304	<i>Z. furfuracea gene for vicilin</i>
		9	0.81%	1105	<i>Herrania nycterodendron vicilin</i>
	CCT	13	1.17%	1102	<i>Abroma angusta Vicilin gene, partial cds</i>
		24	1.21%	1973	<i>Solanum lycopersicum Vicilin gene</i>
		47	1.41%	3317	<i>Theobroma cacao vicilin</i>
		5	0.57%	875	<i>Pisum sativum vclc gene vicilin</i>
		43	0.80%	5327	<i>Vicia faba Vicilin gene</i>
		31	0.73%	4239	<i>Pea Vicilin gene</i>
		15	1.15%	1304	<i>Z. furfuracea gene for vicilin</i>
		14	1.26%	1105	<i>Herrania nycterodendron vicilin</i>



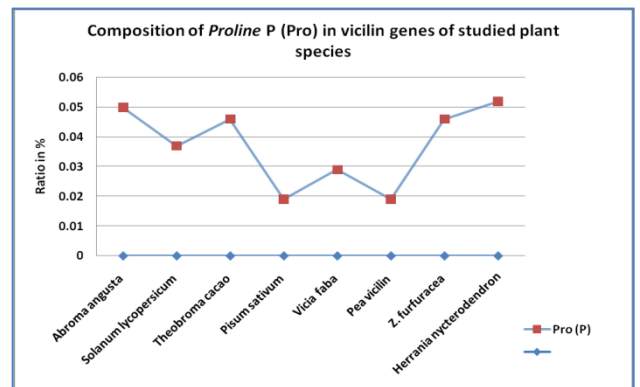
A



C



B



D

Fig-2 Composition detail of Protein amino acids: (A). Agrinine (Arg), (B). Asparagine (Asn), (C). Glutamine (Gln), (D). Proline (Pro) in Vicilin genes of studied plant species.

Table-3 Physicochemical composition detail of studied *Vicilin gene* sequences of various plant species

Plant <i>Vicilin gene</i>	No. of (+) charge (K,R,H)	No. of (-) charge (D,E)	Grand average of hydropathicity	No. of Stop codons	Total No. of Amino acid (excluding stop codons)
<i>Abroma angusta</i>	43(Arg + Lys)	15 (Asp + Glu)	-0.243	28	337
<i>Solanum lycopersicum</i>	89 (Arg + Lys)	82 (Asp + Glu)	-0.826	7	650
<i>Theobroma cacao</i>	132 (Arg + Lys)	90 (Asp + Glu)	-0.391	33	1072
<i>Pisum sativum</i>	20 (Arg + Lys)	12 (Asp + Glu)	0.397	22	269
<i>Vicia faba</i>	204 (Arg + Lys)	103 (Asp + Glu)	0.004	109	1666
Pea	170 (Arg + Lys)	67 (Asp + Glu)	0.052	100	1313
<i>Z. furfuracea</i>	52 (Arg + Lys)	26 (Asp + Glu)	-0.433	20	414
<i>Herrania nycterodendron</i>	38 (Arg + Lys)	23 (Asp + Glu)	-0.055	19	349

Table-4 Properties detail in studied *Vicilin gene* sequences

Plant <i>Vicilin gene</i>	Aliphatic (l) residues (I, V, L)	Aromatic (@) residues (Y,H,W,F)	Hydrophobic (h) residues (W, F, Y, M, L, I, V, A, C, T, H)	Alcohol (o) residues (S, T)	Polar (p) residues (D, E, H, K, N, Q, R, S, T)	Tiny (t) residues (A, G, C, S)	Small (s) residues (A, G, C, S, V, N, D, T, P)	Bulky (b) residues (E, F, I, K, L, M, Q, R, W, Y)	Charged (C) (D,E,K,R,H)
<i>Abroma angusta</i>	20.27%	12.87%	38.35%	15.34%	45.20%	23.56%	44.65%	44.65%	18.90%
<i>Solanum lycopersicum</i>	17.04%	11.87%	39.42%	9.13%	130.59%	20.39%	41.24%	87.51%	33.18%
<i>Theobroma cacao</i>	20.09%	13.66%	47.33%	12.48%	48.68%	20.45%	41.71%	52.03%	23.34%
<i>Pisum sativum</i>	29.55%	21.64%	60.48%	9.27%	38.48%	12.71%	33.33%	89.69%	16.83%
<i>Vicia faba</i>	25.52%	15.54%	52.39%	13.01%	44.16%	15.04%	36.61%	54.64%	19.88%
Pea	24.55%	18.89%	54.49%	12.88%	42.32%	15.64%	30.04%	52.20%	20.45%
<i>Z. furfuracea</i>	14.74%	11.75%	47.00%	16.12%	51.84%	27.64%	53.91%	42.39%	25.34%
<i>Herrania nycterodendron</i>	19.83%	13.58%	49.18%	11.95%	41.84%	23.09%	41.84%	50.54%	19.02%

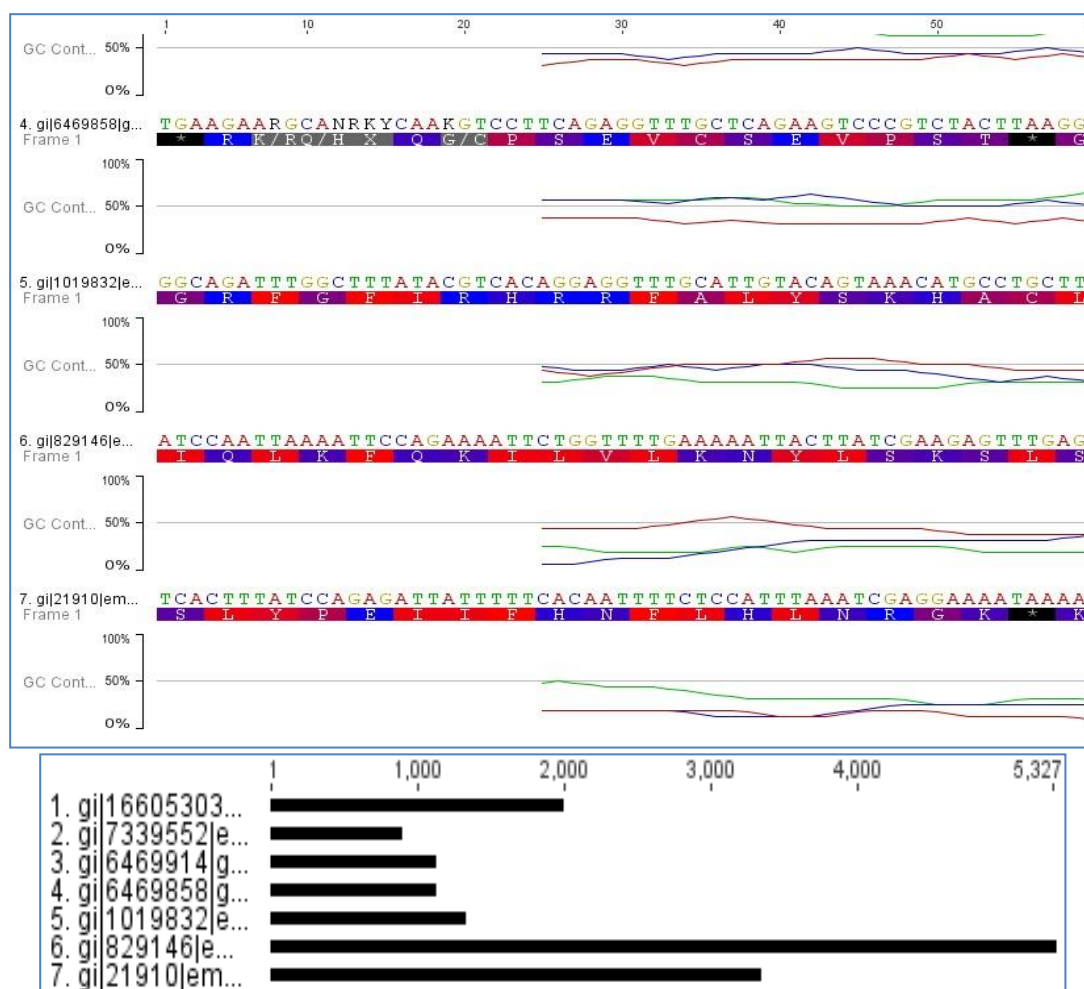


Fig-3 GC contents partial detail of *Vicilin genes* of various plant species

Phylogenetic analysis

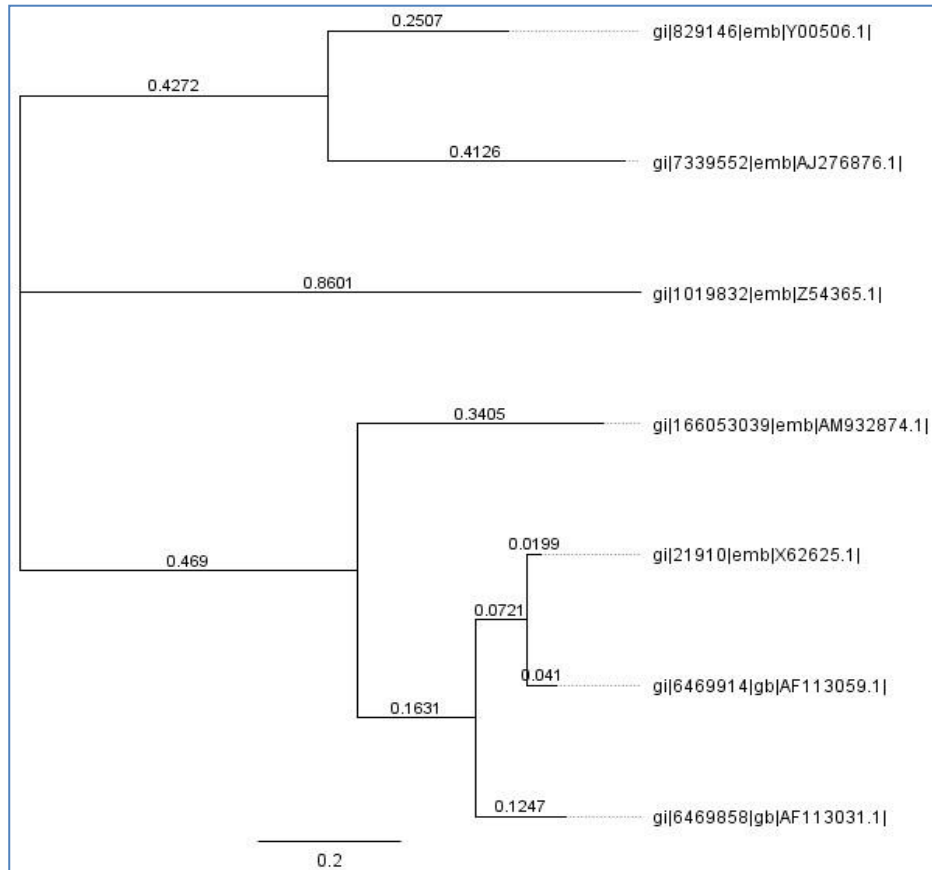


Fig-4 Phylogenetic tree of studied Vicilin genes of studied plant species

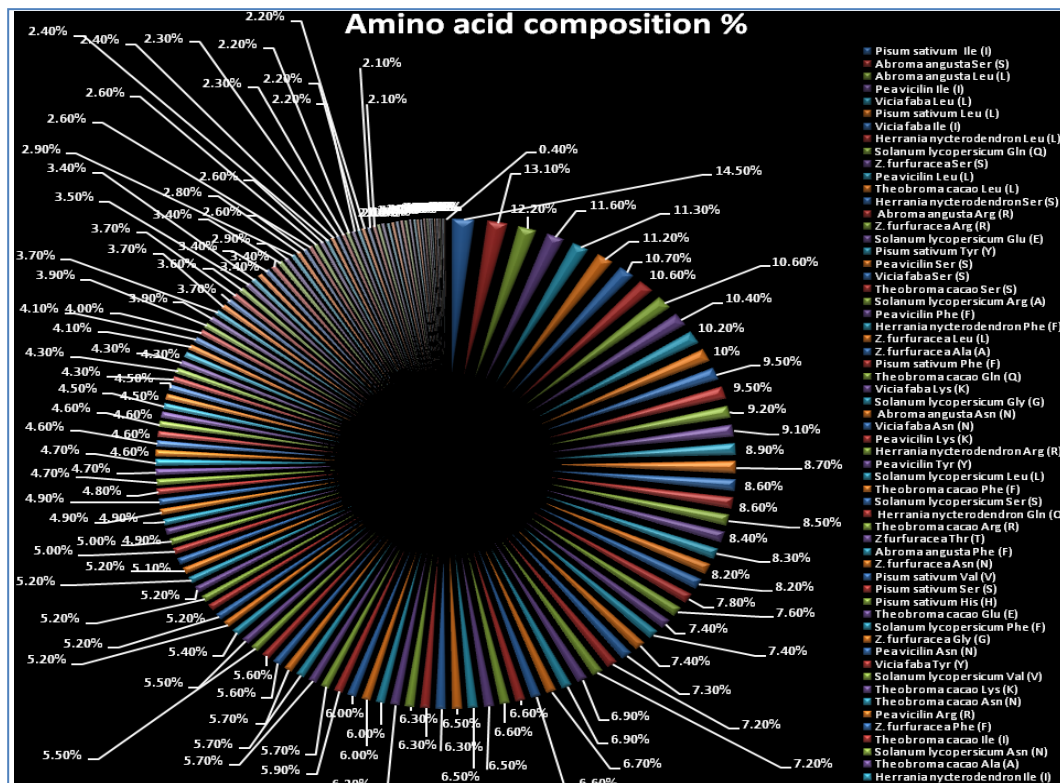


Fig-5 Amino acid composition chart

Discussion

The *Vicilin genes* are responsible for letting anti-decay activities in plant especially in seeds. But the amino acids developing this nature is not clearly specified till

time. *Vicilin gense* are major source for the supply of Nitrogen (N). The conducted study will put a possible assumption upon the codon amino acids responsible for creating anti-decay conditions. Nitrogen is an indispensable element requires by

all living organisms and the element is ingredient into amino acids [17-18]. *Asparagine* (N) promotes nitrogen cycle in plant vegetative organs during the stress conditions. *Asparagine* along with *Glutamine* Q (Gln) plays an important role to transport and storage of nitrogen element from the source [19]. Because of high ingredient ratios of nitrogen/carbon and stability, the *Asparagine* is considered as an optimal nitrogen transport and reserve compound [20]. The *Vicilin gene* are anti-decay in nature, the reason behind this is the presence of amide source amino acids such as *Asparagine*. It helps plant to rejuvenate even during the stress conditions by generating nitrogen cycle leading more towards the biosynthesis activities. This intrinsically benefits the plant seeds to remain raw and alive. Though the retrieved nucleotides in *Vicilin genes* of *Pisum sativum* from the gene bank has less nucleotides of 875 but the Bioinformatics analysis has hinted upon a fact, that massive sequence of the *Pisum sativum* devoid of stop codons can be durable source to generate anti-decay *Vicilin gene*.

According to the studied results of hydropathicity [Table-3], *Vicilin genes* in *Pisum sativum* has formed maximum grand average hydropathicity with 0.397, as compare to the other studied *Vicilin genes* of respective plant species. However, in *Pisum sativum* the hydrophobic (h) amino acids ratio is highest among all the *Vicilin genes* of species with 60% [Table-4], which could lead towards the formation of less interactive nature of amino acids with hydrogen bond and thus form non-binding protein. Since the hydrophobic amino acids are considered to be non interactive with H₂O hydrogen bond. The interactive natures of polar or hydrophilic residues are found to be 38.48%. The *Glutamine* Q (Gln) was analyzed more in *Solanum lycopersicum* with 10.6%. followed by 7.6% *Theobroma cacao*, 6.6% *Herrania nycterodendron*, 4.7% *Vicia faba*, 4.5% *Pisum sativum*, *Abroma angusta*, 4.3% *Z. furfuracea* and 3.7% in *Pea Vicilin*. It is very interesting to notice that lowest grand average of hydropathicity is found in *Solanum lycopersicum* with -0.826% [Table-3] but the composition of polar or hydrophilic residues are high [Table-4] with 130.59%. Only a docking methodology may figure out whether the protein amino acids are binding in nature or not.

Being anti-decay in nature, the *Vicilin gene* must go by an adverse situation during the storage pattern, to cope such condition; the gene must contain a specific protein, which can promote tolerance characteristics. *Proline* P (Pro) is such a protein amino acid which will give tolerance ability from certain adverse environmental conditions. The study also held an analysis for an understanding a distributions patterns of *Proline* in various *Vicilin genes*. It was found that *Vicilin gene* of *Herrania nycterodendron* contains more *Proline* proteins with 5.2%. In *abroma angusta* 5.0%, 4.6% in *Z. furfuracea* and *Theobroma cacao*, 3.7% *Solanum lycopersicum*, 2.9% *Vicia faba* and 1.9% in *Pisum sativum* and *Pea Vicilin*.

The major storage protein *Asparagine* (N) found to be content more in *Abroma angusta* with 7.4% followed by 7.3% *Vicia faba*, 6.5% *Z. furfuracea*, 6.0% *Pea vicilin*, 5.7% *Theobroma cacao*, 5.5% *Solanum lycopersicum*, 4.9% *Herrania nycterodendron*. *Asparagine* regulates biosynthesis activities, as it contains α -amino acids that promotes biosynthesis. The study indicated that the occurrence ratio of *Asparagine* in *Abroma angusta* and *Vicia faba* belonging to Fabaceae family, are somewhat similar. This shows that the *Asparagine* protein which promotes maximum rates of biosynthesis by transporting and storing important Nitrogen elements, happen to be found durably more in plants belonging to Fabaceae family. To adhere process of biosynthesis an active mechanism of Nitrogen is necessary which is performed by another important protein *Arginine* R (Arg). The *Arginine* is a polar or positively charged (cation) amino acid. It is located in the active centric portion of protein. It helps to maintain overall charge particles in balance form for protein. While studying the occurrence nature of *Arginine* protein within the studied *Vicilin gene* sequences of respective plant species, it was noted that *Vicilin gene sequence* of *Abroma angusta* was found to be dominated with 9.5%. *Z. furfuracea Vicilin gene* has 9.2%, followed by 8.5% *Solanum lycopersicum*, 7.2% *Theobroma cacao*, 7.2% *Herrania nycterodendron*, 6.6% *Theobroma cacao*, 5.7% *Pea vicilin*, 4.9% *Vicia faba* and 2.6% *Pisum sativum*. The efficiency of protein depends upon the binding property of proteins. Which is related with the interaction nature with hydrogen bond and proteins. Understanding the hydropathicity of specific amino acids it is found that amino

acids with high polar ratio has maximum chances to exhibits interaction with hydrogen bond. According to the report of property detail in studied *Vicilin genes* [Table-4], *Solanum lycopersicum Vicilin genes* isolated from seed with 103.59% composition ratio formed a highest ratio of polar amino acids as compared to other species *Vicilin genes*. With composition ratio of 51.84% *Vicilin gene* of *Z. furfuracea* was found to be second highest polar amino acids *Vicilin genes*. This reflected upon the potential binding nature of proteins within certain species. Understanding the results of hydropathicity, it was found that *Vicilin gene* of *Abroma angusta* and *Z. furfuraceae* with 9.5% and 9.2% composition ratio of polar (-) residues respectively, the perception of protein-protein interaction is high in these two genes.

The phylogenetic study shows that the four plant species *Vicilin gene* are close relative to each other i.e. *Pisum sativum* and *Vicia faba* with a distance of 0.4126 (*Pisum sativum*) to 0.2507 (*Vicia faba*), but they originates from a common ancestor which is originated at distance of 0.4272. It is very interesting to notice that both species belongs to same Fabaceae family. The other plant species *Vicilin genes* close to each other were *Theobroma cacao* with a distance 0.0199 to *Herrania nycterodendron Vicilin gene* with a distance of 0.04. Its common ancestor is originated at 0.0721. It is also found that *Vicia faba* and *Abroma angusta* is not close relative. The only plant species *Vicilin gene* with no relative to other *Vicilin genes* species is *Z. furfuracea gene*.

With findings of two close plant species *Vicilin genes* of *Pisum sativum* and *Vicia faba*, belonging to common family, the study created an assumption of probability of ingredients source of *Vicilin genes* specifically in plants belonging to family Fabaceae. In case of other two relative plant gene species the assumption cannot be pertained due to their respective families. As *Theobroma cacao* belongs to Malvaceae family while *Herrania nycterodendron* belongs to family Bytneriaceae.

Conclusion

The study found that *Vicilin genes* in *Pisum sativum* is more durable to form anti-decay nature genes than other plant species but at the same time the durable nature depend upon the stop codons diversity in gene. Among the studied plants secreting *Vicilin genes*; *Pisum sativum* and *Vicia faba* are very close relative and it was also found that both of these species belongs to Fabaceae family. It also provides an assumption and possibility of *Vicilin gene* ingredients in species of Fabaceae family. However, the highest composition of hydrophobic residues of amino acids in *Solanum lycopersicum Vicilin gene* has also put on a light that in this protein the water hydrogen binding capacity is less. The polar amino acid *Arginine* R was recorded more in *Vicilin genes* of *Abroma angusta* and *Z. furfuraceae* with 9.5% and 9.2% respectively. The *Abroma angusta Vicilin genes* is found to be the most efficient plant species in respect to the constituent of major storage protein amino acids *Asparagine* D (Asn), *Arginine* R (Arg) and *Proline* P (Pro), while *Glutamine* Q (Gln) was analyzed more in *Vicilin gene* of *Solanum lycopersicum*.

Acknowledgement

The study acknowledges the web lab of Bioinformatics Sub-Distributed Information Centre (DISC), Sikkim State Council of Science & Technology and also Department of Biotechnology, Government of India, for providing all the necessary facilities to complete this research.

Conflict of Interest: None declared

References

- [1] Xie Z., Neigel J. and Chlan C. (2012) *Biochemical Genetics*, 50(5-6), pp. 372-388.
- [2] Osborne T.B. (1924) *The vegetable proteins*, 2nd Edn. Longmans, Green and Co., London, pp.154.
- [3] Bhaduri A., Ravishankar R. and Sowdhamini R. (2004) *Proteins*, 54, 657-670.
- [4] Gromiha Michael M. and Selvaraj S. (2004) *Progress in Biophysics & Molecular Biology*, 86, 235-277.

- [5] Shao X. and Grishin Nick V. (2000) *Nucleic Acids Research*, 28(14), 2643–2650.
- [6] Burley S.K. and Petsko G.A. (1985) *Science*, 229(4708), pp.23–28.
- [7] Reid K.S.C., Lindley P.F. and Thornton J.M. (1985) *FEBS Letters*, 190(2), pp.209–213.
- [8] Satyapriya R. and Vishveshwara S. (2004) *Nucleic Acids Research*, 32, 4109–4118.
- [9] Barlow D.J. and Thornton J.M. (1983) *Journal of Molecular Biology*, 168, 867–885.
- [10] Biro J.C. (2006) *Theoretical Biology and Medical Modelling*, 3,15.
- [11] Eisenberg D., Weiss R.M. Terwilliger T.C. and Wilcox W. (1982) *Faraday Symposia of Chemical Society*, 17,109-120.
- [12] Czarnecki Gail L. and Baker David H. (1984) *The Journal of Nutrition*, 114, pp.581-590.
- [13] Mifflin Ben J. and Habash Dimah Z. (2002) *Journal of Experimental Botany*, 53(370),pp. 979–987.
- [14] Glass Anthony D.M., Britto Dev T., Kaiser Brent N., Kinghorn James R., Kronzucker Herbert J., Kumar A., Okamoto M., Rawat S., Siddiqi M.Y., Unkles Shiela E., Vidmar Joseph J. (2002) *Journal of Experimental Botany*, 53(370), pp.855–864.
- [15] Hong Z., Lakkineni K., Zhang Z. and Verma Des Pal S. (2000) *Plant Physiology*, 122(4), pp.1129–1136.
- [16] Miller G., Honig A., Stein H., Suzuki N., Mittler R. and Zilberstein A. (2009) *The Journal of Biological Chemistry*, 284(39), pp. 26482–26492.
- [17] Sungdae P. (2002) Structural studies on *Sinorhizobium meliloti* Dctd related to ATP binding and activation. Ph.D. Thesis, Pennsylvania state university, USA.
- [18] Tabatabaei S.J., Yusefi M. and Hajiloo J. (2008) *Scientia. Hort.*, 116,113-122.
- [19] Ireland R.J. and Lea P.J. (1999) The enzymes of *Glutamine*, glutamate, *Asparagine* and aspartate metabolism, in: B.K. Singh (Ed.), *Plants Amino Acids, Biochemistry and Biotechnology*, Dekker, M. Inc., New York/Basel/Hong Kong, pp. 49–109.
- [20] Gaufichon L., Reisdorf-Crena M. Rothstein S.J. Chardona F. and Suzuki A. (2010) *Plant Science*, 179, pp.141–153.