Int.J.Curr.Microbiol.App.Sci (2019) *8*(10): 716-726

International Journal of Current Microbiology and Applied Sciences
***ISSN: 2319-7706* Volume 8 Number 10 (2019)**
Journal homepage: http://www.ijcmas.com

Original Research Article https://doi.org/10.20546/ijcmas.2019.810.083

LEA Genes Play Important Role in Seed and Pod Development in *Cajanus cajan*

Antara Das[1], Kuldeep Kumar[1], Kishor Tribhuvan[1], Rekha Joshi[2],
Kumar Durgesh[2] and Kishor Gaikwad[1]*

[1]ICAR-National Institute for Plant Biotechnology, New Delhi, 110012, India
[2]Division of genetics and plant breeding, Indian Agricultural Research Institute, New Delhi, 110012, India

**Corresponding author*

ABSTRACT

Keywords

LEA gene, *Cajanus cajan*, Paralogs, Seed and pod

Article Info

Accepted:
07 September 2019
Available Online:
10 October 2019

Seed and pod development is one of the important stages affecting the yield potential of a plant. Number of seeds per pod and pod length are some important yield attribute. *Cajanus cajan* has immense diversity in term of these traits. *LEA* (*late embryogenesis abundant*) gene family members are known to accumulate in seed and pod during several stresses and seed development. In this study we have tried to find out total number of LEA genes present in *Cajanus cajan*, and their phylogenetic analysis to search for the structural homologs. In this study also tried to build the heat map showing the expression level of these genes. It was found that some of the *LEA* gene family members *viz C.cajan_17192, C.cajan_21717, C.cajan_37355, C.cajan_43531, C.cajan_10424, C.cajan_10963, C.cajan_03928, C.cajan_20859, C.cajan_31323, C.cajan_06188, C.cajan_14597, C.cajan_35463, C.cajan_09914, C.cajan_09556, C.cajan_29204, C.cajan_04295, C.cajan_45355, C.cajan_00461* and *C.cajan_21796* were showing up regulation in reproductive mature seeds and pods. These *LEA* gene members may play important role in seed and pod development in pigeonpea.

Introduction

Pigeonpea (*Cajanus cajan* (L.) Millsp.), also known as red gram is a nutritionally rich and an important grain legume belonging to the *Cajaninae* sub-tribe of the tribe *Phaseoleae* under sub-family *Papilionoideae* of family *Leguminosae*. It is known to be originated from India (Van der Maesen, 1980). It is one of the high protein food legumes of rainfed tropic and sub-tropic environments. Pigeonpea is a hardy crop which shows tolerance toward heat and drought though it is having susceptibility toward extended cold and water logging. It has diploid genome with 11 pairs of chromosomes (2n =2x= 22) and the estimated genome size is 833.07 Mbp (Varshney *et al.,* 2012).

Many factors are accountable for stumpy productivity; like lack of superior cultivars, susceptibility toward the various pest and diseases. Proper agronomic practices have been equally important in this regard. Besides this there are many morphological characteristics which significantly hamper the yield.

Number of pods per plant, pod length and number of seeds per pod are some of them. Thus, optimization of all these factors will help us in attaining higher yield.

Seed development is one of the largely multifaceted genetically regulated as well as metabolically active process in the plant life cycle. Ultimately seeds are the final outcome of plants life cycle.

Many studies have been performed to understand the metabolic and hormonal involvement and changes during the seed development stages in legume which concerning about synthesis of carbohydrates, protein, lipids and other metabolites and their proper processing and partitioning as the form of assimilates. In case of pigeon pea it is observed that the seed development processes may require 25 to 35 days from the day of anthesis to maturation, this time period varies based on the genotype and the moisture content in the mature seeds. During seed development the proteins and soluble sugars play significant roles to getting hold of desiccation tolerance in the seeds.

During last stage of embryogenesis a group of hydrophilic proteins known as Late Embryogenesis Abundant (LEA) proteins accumulates. These proteins are also found in vegetative tissues during heat and drought condition. Due to its extensively wide range distribution from algae (Honjoh *et al.,* 1995) to angiosperms in the plant kingdom shows its significant role in the plants during different response.

These proteins are not only coupled to water deficit caused by environmental changes but also to water constraint created during plant development under optimal growth conditions, such as during development of seeds and pollen grains, or some stages of shoot and root development (Colmenero-Flores *et al.,* 1999; Vicient *et al.,* 2000; Sheoran *et al.,* 2006).

Biotic stresses such as drought, salinity, osmotic, cold, and freezing temperatures construct cellular water deficient condition, which escort to the gathering of a collection of exceedingly hydrophilic LEA proteins (Battaglia *et al.,* 2008; Bies-Etheve *et al.,* 2008); Hundertmark and Hincha, 2008). Some of the LEA proteins are also involved seed germination to advancement into seedling growth.

The majority of the LEA proteins recognized till now belong to hydrophilins. It is well and extensively distributed protein group containing high level of charged amino acid residues *viz.,* glycine, alanine, serine, or threonine and lack of tryptophanes and cysteines (Garay-Arroyo *et al.,* 2000).

In pulses investigation of LEA protein was done based on conserved amino acid sequences and seven groups named as LEA1 to LEA7 were identified. Though LEA protein was studied in pluses including *Cajanus cajan, Phaseolus vulgaris, Vigna* sp. but individually extensive study in *Cajanus cajan* LEA genes was lacking. In this study we tried to identify and characterized all LEA genes present in *C. cajan* and also to analyse their expression pattern.

Materials and Methods

Genome wide identification of LEA genes in *C. cajan*

The protein sequence of LEA genes were downloaded from LIS (Legume information data base) database. We searched the LIS database using ‘LEA’ as key words and mRNAs as well as polypeptide sequences were downloaded in fasta format.

Phylogenetic analysis to search homologs of LEA gene

Multiple sequence alignment was performed to the protein sequences of all LEA gene via MEGA10 software to build the phylogenetic tree for all LEA protein in *Cajanus cajan*.

Expression pattern analysis of LEA family genes using gene expression atlas

Expression atlas of *C. cajan* developed by Pazhamala *et al.,* (2017) from 10 tissues of a *C. cajan* cv. Asha and was used to visualize the expression profile of all LEA genes in different tissues. The gene expression data in the form of FPKM values of each selected genes was filtered from the gene expression atlas and used for the preparation of heatmap using ‘R’ script.

Result and Discussion

A set of 82 LEA proteins were identified through the search option from LIS database in *C. cajan*. The information such as chromosome name, start and end position, domains present, their function are provided in table 1. Chromosome CcLG02 and CcLG03 contains most number of LEA gene i.e. both of these contains 8 LEA gene members. In terms of numbers LEA14 is the most abundant LEA protein in *C. cajan*, as it have 60 members. Whole phylogeny was classified into 4 major clades. Clade I contains 2 members, clade II contains 21 members, clade III contains 3 members while clade IV contains 56 members.

Both clade I members are not assigned to any further subgroup. Clade II contains all LEA3, LEA5 and some LEA 14 members. Two members of clade III i.e. cajca.C.cajan_02499.1 and cajca.C.cajan_ 21796.1 are not well characterized but they falls very close to cajca.C.cajan_10424.1, a LEA18 protein. This concludes that both of these proteins viz., cajca.C.cajan_02499.1 and cajca.C.cajan_21796.1 may be having LEA18 like function. Clade IV comprises solely of LEA14 members.

The heat map developed from expression atlas data developed by Pazhmahla *et al.,* (2012) revealed the expression pattern of all these LEA genes. FPKM values of these particular genes in reproductive mature pod, reproductive mature seed, reproductive stamen, reproductive pistil, reproductive petal, reproductive sepal, reproductive immature pod, reproductive immature seed, reproductive bud, reproductive shoot apical meristem, reproductive petiole and reproductive leaf were used. A total of 19 LEA gene family member *i.e. C.cajan_17192, C.cajan_21717, C.cajan_37355, C.cajan_43531, C.cajan_ 10424, C.cajan_10963, C.cajan_03928, C.cajan_20859, C.cajan_31323, C.cajan_ 06188, C.cajan_14597, C.cajan_35463, C.cajan_09914, C.cajan_09556, C.cajan_ 29204, C.cajan_04295, C.cajan_45355, C.cajan_00461* and *C.cajan_21796* were found to be upregulated in reproductive mature bud and reproductive mature seed as compared to the reproductive immature bud and reproductive immature seed.

These LEA genes may be of primary important for transformation of reproductive immature bud and reproductive immature seed to reproductive mature bud and reproductive mature seed (Fig. 1 and 2).

Table.1

Name	Chr. no.	Start	End	Domain	LEA-Family	Description
C.cajan_ 00461	CcLG11	4337141	4337546	IPR005513	LEA-25/ LEA-D113	seed maturation protein; IPR005513 (Late embryogenesis abundant protein, LEA-25/LEA-D113); GO:0009790 (embryo development)
C.cajan_ 00500	CcLG11	4735764	4736484	IPR004864	LEA-14	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_ 02499	CcLG11	27354459	27356340	IPR025423		Late embryogenesis abundant protein (LEA) family protein; IPR025423 (Domain of unknown function DUF4149)
C.cajan_ 03928	CcLG11	43136957	43138475	IPR004238	LEA-3	late embryogenesis abundant protein, putative / LEA protein, putative; IPR004238 (Late embryogenesis abundant protein, LEA-3)
C.cajan_ 04295	CcLG11	46876402	46876624	IPR004926	LEA-5	Late embryogenesis abundant 3 (LEA3) family protein; IPR004926 (Late embryogenesis abundant protein, LEA-5); GO:0006950 (response to stress)
C.cajan_ 05658	CcLG02	12473333	12474143	IPR004864	LEA-14	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_ 05699	CcLG02	12813034	12814144	IPR004864, IPR004864	LEA-14	late embryogenesis abundant protein; IPR004864 (Late embryogenesis abundant protein, LEA-14), IPR004864 (Immunoglobulin-like fold); GO:0009269 (response to desiccation)
C.cajan_ 05978	CcLG02	15418451	15419144	IPR004864	LEA-14	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_ 06049	CcLG02	16405988	16406904	NA		Late embryogenesis abundant protein (LEA) family protein
C.cajan_ 06188	CcLG02	17870000	17871698	IPR004238	LEA-3	late embryogenesis abundant protein, putative / LEA protein, putative; IPR004238 (Late embryogenesis abundant protein, LEA-3)
C.cajan_ 06305	CcLG02	19270566	19271535	IPR004864	LEA-14	protein YLS9 [Glycine max]; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_ 06725	CcLG02	24130143	24130788	IPR004926	LEA-5	Late embryogenesis abundant 3 (LEA3) family protein; IPR004926 (Late embryogenesis abundant protein, LEA-5); GO:0006950 (response to stress)
C.cajan_ 07993	CcLG02	36538747	36539626	IPR004864	LEA-14	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_	CcLG03	12841276	12842044	IPR004864	LEA-14	Late embryogenesis abundant (LEA)

09281						hydroxyproline-rich glycoprotein family; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_ 09556	CcLG03	15642569	15643139	IPR005513	LEA-25/ LEA-D113	seed maturation protein; IPR005513 (Late embryogenesis abundant protein, LEA-25/LEA-D113); GO:0009790 (embryo development)
C.cajan_ 09914	CcLG03	19363346	19363997	IPR004238	LEA-3	35 kDa seed maturation protein [Glycine max]; IPR004238 (Late embryogenesis abundant protein, LEA-3)
C.cajan_ 10424	CcLG03	23987705	23987948	IPR018930	LEA-18	Late embryogenesis abundant protein, group 6; IPR018930 (Late embryogenesis abundant protein, LEA-18)
C.cajan_ 10532	CcLG03	25145030	25145813	IPR004864	LEA-14	protein YLS9 [Glycine max]; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_ 10818	CcLG03	27827288	27827672	IPR004926	LEA-5	late embryogenesis abundant protein; IPR004926 (Late embryogenesis abundant protein, LEA-5); GO:0006950 (response to stress)
C.cajan_ 10963	CcLG03	28997247	28997841	IPR004864	LEA-14	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_ 10997	CcLG03	29303412	29303877	IPR004864	LEA-14	Late embryogenesis abundant hydroxyproline-rich glycoprotein family, putative n=1 Tax=Theobroma cacao RepID=UPI00042B1EF8; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_ 11442	CcLG06	4615050	4616742	IPR004864	LEA-14	uncharacterized protein LOC100811519 [Glycine max]; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_ 12887	CcLG06	19537893	19539944	IPR004864	LEA-14	late embryogenesis abundant protein; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_ 13277	CcLG06	23254692	23255256	IPR004864	LEA-14	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_ 13655	CcLG10	3270951	3271617	IPR004864, IPR013783	LEA-14	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family; IPR004864 (Late embryogenesis abundant protein, LEA-14), IPR013783 (Immunoglobulin-like fold)
C.cajan_ 13744	CcLG10	4153772	4154141	IPR004864	LEA-14	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_ 13745	CcLG10	4165613	4166177	IPR004864	LEA-14	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_	CcLG10	14360576	14361272	NA		Late embryogenesis abundant protein (LEA)

14597						family protein
C.cajan_ 14840	CcLG10	16958021	16958600	IPR004864, IPR013783	LEA-14	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family; IPR004864 (Late embryogenesis abundant protein, LEA-14), IPR013783 (Immunoglobulin-like fold)
C.cajan_ 15135	CcLG10	20124814	20125555	IPR004864	LEA-14	protein YLS9 [Glycine max]; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_ 15504	CcLG08	1770036	1770594	IPR004864	LEA-14	protein YLS9-like [Glycine max]; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_ 15522	CcLG08	1887897	1888992	IPR004864	LEA-14	uncharacterized protein LOC100787767 [Glycine max]; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_ 16206	CcLG08	9717806	9718436	IPR004864	LEA-14	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_ 17192	CcLG08	19359194	19361713	IPR004864	LEA-14	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_ 20707	CcLG01	16077098	16077791	IPR004864	LEA-14	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_ 20854	CcLG01	17614454	17615117	IPR004864	LEA-14	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_ 20859	CcLG01	17636185	17636797	IPR004864, IPR013783	LEA-14	Late embryogenesis abundant protein; IPR004864 (Late embryogenesis abundant protein, LEA-14), IPR013783 (Immunoglobulin-like fold); GO:0009269 (response to desiccation)
C.cajan_ 21068	CcLG04	1997430	1999176	IPR004864	LEA-14	late embryogenesis abundant protein; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_ 21674	CcLG04	7986489	7987122	NA		late embryogenesis abundant protein
C.cajan_ 21676	CcLG04	8014879	8015677	IPR004864	LEA-14	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_ 21717	CcLG04	8355887	8356442	IPR004864	LEA-14	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family; IPR004864 (Late embryogenesis abundant protein, LEA-14); GO:0009269 (response to desiccation)
C.cajan_ 21796	CcLG04	8989147	8990184	IPR025423		Late embryogenesis abundant protein (LEA) family protein; IPR025423 (Domain of unknown function DUF4149)

C.cajan_ 22736	CcLG09	6775857	6776442	IPR004864	LEA-14	uncharacterized protein LOC100797168 [Glycine max]; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_ 22769	CcLG09	7226278	7227061	IPR004864	LEA-14	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_ 23280	CcLG05	2045435	2045903	IPR004864	LEA-14	protein YLS9-like [Glycine max]; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_ 24412	Scaffold 000046	908804	909962	IPR005513	LEA-25/ LEA-D113	late embryogenesis abundant protein; IPR005513 (Late embryogenesis abundant protein, LEA-25/LEA-D113); GO:0009790 (embryo development)
C.cajan_ 25044	Scaffold 127746	500414	500783	IPR004926	LEA-5	late embryogenesis abundant protein; IPR004926 (Late embryogenesis abundant protein, LEA-5); GO:0006950 (response to stress)
C.cajan_ 25170	Scaffold 000321	151017	151650	IPR004864	LEA-14	protein YLS9-like [Glycine max]; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_ 25699	Scaffold 000332	214072	214786	IPR004864	LEA-14	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_ 26931	Scaffold 128870	108738	109659	IPR004864	LEA-14	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_ 26932	Scaffold 128870	124905	125535	IPR004864	LEA-14	protein YLS9-like [Glycine max]; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_ 26934	Scaffold 128870	162312	162978	IPR004864	LEA-14	protein YLS9 [Glycine max]; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_ 26943	Scaffold 128870	216816	217317	IPR004864	LEA-14	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_ 27095	Scaffold 000144	144731	145714	IPR004864	LEA-14	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_ 27530	Scaffold 000159	249736	250420	IPR004864	LEA-14	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_ 27598	Scaffold 132776	453667	456495	IPR004864	LEA-14	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_ 27619	Scaffold 133584	141240	141717	IPR004864	LEA-14	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family; IPR004864 (Late embryogenesis abundant

						protein, LEA-14)
C.cajan_ 29121	Scaffold 127472	83282	83848	IPR004864, IPR013783	LEA-14	Late embryogenesis abundant protein; IPR004864 (Late embryogenesis abundant protein, LEA-14), IPR013783 (Immunoglobulin-like fold); GO:0009269 (response to desiccation)
C.cajan_ 29142	Scaffold 127472	285988	287854	NA		Late embryogenesis abundant protein (LEA) family protein
C.cajan_ 29209	Scaffold 127411	160686	161074	IPR000389	LEA-B19.1A	late embryogenesis abundant protein B19.1A; IPR000389 (Stress induced protein)
C.cajan_ 31323	Scaffold 000286	24310	25219	NA		Late embryogenesis abundant protein (LEA) family protein
C.cajan_ 33267	Scaffold 130593	192997	193747	IPR004864	LEA-14	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_ 34938	Scaffold 133177	30950	31586	IPR004864	LEA-14	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_ 34939	Scaffold 133177	34525	34951	IPR004864	LEA-14	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_ 35463	Scaffold 131636	152855	153339	IPR000167		late embryogenesis abundant protein-like [Glycine max]; IPR000167 (Dehydrin); GO:0006950 (response to stress), GO:0009415 (response to water)
C.cajan_ 35879	Scaffold 133195	129374	130544	IPR004864	LEA-14	uncharacterized protein LOC100798888 [Glycine max]; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_ 36841	Scaffold 135508	32862	33528	IPR004864	LEA-14	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_ 36842	Scaffold 135508	44449	44989	IPR004864	LEA-14	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_ 37355	Scaffold 133269	88667	89231	IPR004864	LEA-14	Late embryogenesis abundant hydroxyproline-rich glycofamily protein n=1 Tax=Theobroma cacao RepID=UPI00042B23A2; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_ 37356	Scaffold 133269	95199	95730	IPR004864	LEA-14	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_ 39866	Scaffold 132067	130160	130562	IPR004864	LEA-14	protein YLS9-like [Glycine max]; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_	Scaffold	43522	44077	IPR004864	LEA-14	Late embryogenesis abundant (LEA)

41026	133482					hydroxyproline-rich glycoprotein family; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_ 41555	Scaffold 126477	15837	17915	IPR004864	LEA-14	late embryogenesis abundant protein; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_ 41962	Scaffold 133864	2357	3138	IPR004926	LEA-5	Late embryogenesis abundant 3 (LEA3) family protein; IPR004926 (Late embryogenesis abundant protein, LEA-5); GO:0006950 (response to stress)
C.cajan_ 43531	Scaffold 132354	17752	18265	IPR004864	LEA-14	uncharacterized protein [Glycine max]; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_ 43533	Scaffold 132354	77454	78051	IPR004864	LEA-14	late embryogenesis abundant protein; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_ 43535	Scaffold 132354	101430	102024	IPR004864	LEA-14	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_ 43908	Scaffold 134929	16099	16759	IPR004864	LEA-14	protein YLS9-like [Glycine max]; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_ 45355	Scaffold 137131	22392	23343	IPR005513	LEA-25/ LEA-D113	seed maturation protein; IPR005513 (Late embryogenesis abundant protein, LEA-25/LEA-D113); GO:0009790 (embryo development)
C.cajan_ 47118	Scaffold 117591	23	707	IPR004864	LEA-14	protein YLS9-like [Glycine max]; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_ 47458	Scaffold 132160	20344	21025	IPR004864	LEA-14	protein YLS9 [Glycine max]; IPR004864 (Late embryogenesis abundant protein, LEA-14)
C.cajan_ 48607	Scaffold 135722	24542	25490	IPR004864, IPR013783	LEA-14	late embryogenesis abundant protein; IPR004864 (Late embryogenesis abundant protein, LEA-14), IPR013783 (Immunoglobulin-like fold); GO:0009269 (response to desiccation)

Fig.1 Heat map showing expression pattern of all 82 LEA genes in different tissues

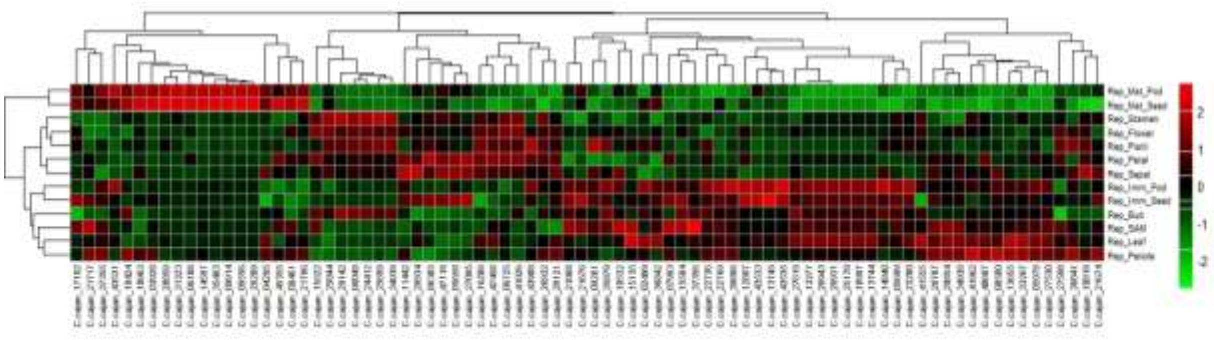
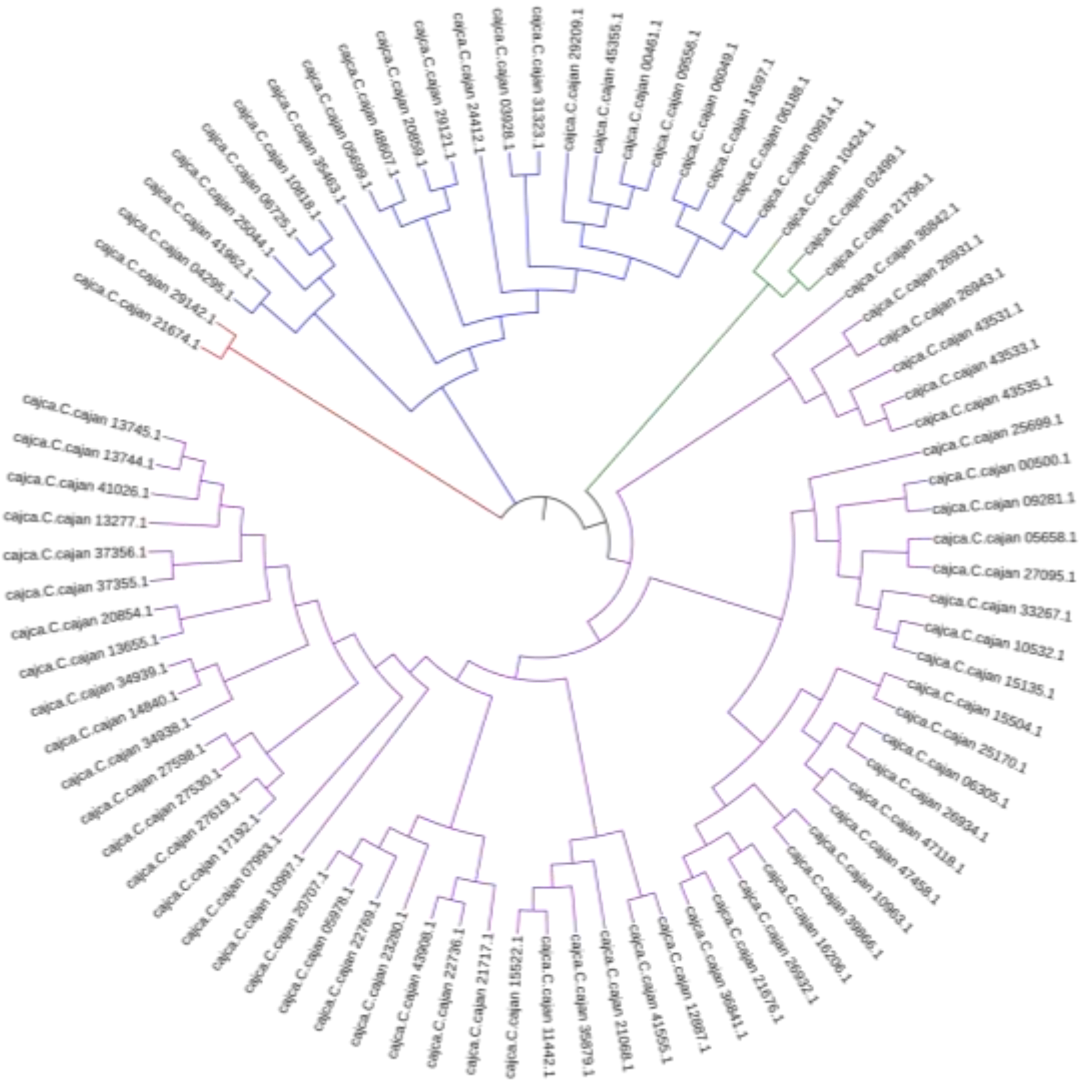


Fig.2 Phylogenetic analysis result depicted the presence of paralogs of LEA gene family members.



Abbreviations

LEA (late embryogenesis abundant)

References

Van der Maesen LJG (1980) India is the native home of pigeonpea. In: Arends JC, Boelema G, de Groot CT, Leeuwenberg AJM, Veenman H, Zonen BV(Eds) Libergratulatorius in honorem H.C.D. de Witlandbouwhoge school, Miscellaneous paper no. 19, Wageningen, Netherlands, pp 257–262

Varshney, R.K. *et al.,* Draft genome sequence of pigeonpea (*Cajanus cajan*), an orphan legume crop of resource-poor farmers. *Nat. Biotechnol.* 30, 83–89 (2012)

Honjoh, K., Yoshimoto, M., Joh, T., Kajiwara, T., Miyamoto, T., and Hatano, S. (1995). Isolation and characterization of hardening-induced proteins in *Chlorella vulgaris* C-27:

identification of late embryogenesis abundant proteins. *Plant Cell Physiol*. 36, 1421–1430

Colmenero-Flores JM, Moreno LP, Smith CE, Covarrubias AA. Pvlea-18, a member of a new late-embryogenesis-abundant protein family that accumulates during water stress and in the growing regions of well-irrigated bean seedlings. Plant Physiol. 1999;120(1):93–104. doi:10.1104/pp.120.1.93

Vicient CM, Hull G, Guilleminot J, Devic M, Delseny M (2000) Differential expression of the *Arabidopsis* genes coding for Em-like proteins. J Exp Bot 51 1211–1220

Sheoran, I. S., Sproule, K. A., Olson, D. J. H., Ross, A. R. S., and Sawhney, V. K. (2006). Proteome profile and functional classification of proteins in *Arabidopsis thaliana* (Landsberg erecta) mature pollen. *Sex. Plant Reprod*. 19, 185–196.

Battaglia M and Covarrubias AA (2013) Late Embryogenesis Abundant (LEA) proteins in legumes. Front. Plant Sci. 4:190. doi: 10.3389/fpls.2013.00190

Bies-Ethève N, Gaubier-Comella P, Debures A, Lasserre E, Jobet E, Raynal M, Cooke R, Delseny M (2008) Inventory, evolution and expression profiling diversity of the LEA (late embryogenesis abundant) protein gene family in *Arabidopsis thaliana*. Plant Mol Biol 67 107–124

Hundertmark, M., and Hincha, D. K. (2008). LEA (late embryogenesis abundant) proteins and their encoding genes in *Arabidopsis thaliana*. *BMC Genom*. 9:118.

Garay-Arroyo, A., Colmenero-Flores, J. M., Garciarrubio, A., and Covarrubias, A. A. (2000). Highly hydrophilic proteins in prokaryotes and eukaryotes are common during conditions of water deficit. *J. Biol. Chem*. 275, 5668–5674.

Pazhamala LT, Purohit S, Saxena RK, *et al.*, (2017) Gene expression atlas of pigeonpea and its application to gain insights into genes associated with pollen fertility implicated in seed formation. J Exp Bot 68:2037–2054.

How to cite this article:

Antara Das, Kuldeep Kumar, Kishor Tribhuvan, Rekha Joshi, Kumar Durgesh and Kishor Gaikwad 2019. LEA Genes Play Important Role in Seed and Pod Development in *Cajanus cajan*. *Int.J.Curr.Microbiol.App.Sci.* 8(10): 716-726.
doi: https://doi.org/10.20546/ijcmas.2019.810.083