Biotechnological Communication

Biosci. Biotech. Res. Comm. 12(1): 124-133 (2019)

BBRC Bioscience Biotechnology Research Communications

Synteny analysis of *Glycine max* and *Phaseolus vulgaris* revealing conserved regions of NBS-LRR coding genes

Gaurav Singh* and Anil Kumar

School of Biotechnology, Devi Ahilya University, Khandwa Rd., Indore-452001, India

ABSTRACT

Soybean (*Glycine max*) and common bean (*Phaseolus vulgaris*) are the two important of the leguminous family, Phaseoleae. Synteny gives a framework in which preservation of genes and gene order is determine between genomes of various species. The syntenic relationship between *G. max* and *P. vulgaris* is important to determine the potential for comparative genomic analysis. Here, synteny analysis has been performed between *G. max* and *P. vulgaris* by using the tool 'SATSUMA'. Genome data of these two legume species were retrieved from NCBI database, gene synteny alignment was then performed and comparatively analysed. Result was visualized by developing custom script in BioPython software version 3.7. The soybean chromosome 13 was aligned with whole genome of common bean as it contains genes which code for nucleotide binding site and leucine rich repeats (NBS-LRR) protein. The NBS-LRR genes play a major role in defense against pathogens. On alignment, a set of genes linked with disease resistant proteins (NBS-LRR) in *G. max* showed synteny with the different chromosomes of *P. vulgaris*. The data supported the theory that in legumes, genes are highly conserved, as extensive regions of synteny exist between these two species. The present study will be helpful to use both genomic resources as well as genetic data for crucial agronomic traits for the improvement of these two species.

KEY WORDS: DISEASE RESISTANCE GENES, *GLYCINE MAX*, NBS-LRR, *PHASEOLUS VULGARIS*, SYNTENY

INTRODUCTION

Synteny word came from greek where 'syn' means together and 'taenia' means ribbon. When gene order is conserved between the organism is commonly refered as Synteny. It is of two different type's viz. macrosynteny when many genes or large chromosomal segments of different organisms are syntenic; and microsynteny when only few genes are conserved in different species (Zhu et al., 2005 and Passoupathy, 2016). Relationship between plants, pathogens and pests have been currently discussed in different models (Andolfo and Ercol-

Article Information:*Corresponding Author: singhgauravbiotech@gmail.com
Received 21/02/2019 Accepted after revision 30/03/2019
Published: 30th March 2019 Pp-124-133

Available at: https://bbrc.in/
Article DOI: http://dx.doi.org/10.21786/bbrc/12.1/16

ano, 2015). This models mainly include proteins which are coded by clustered disease-resistance (*R*) genes in plant genomes (Hulbert et. al., 2001). In earlier study, the *R* gene-encoded proteins were divided into eight major group on the basis of amino acid motif organization and localization in the cell (Gururani et. al., 2012). Among these group two main *R* gene proteins are (CC) NBS-LRR or CNL proteins and TNL proteins. NBS-LRR genes have been grouped into three classes (TNL, CNL and *R* NLs) (Shao et. al., 2016 Neupane et. al., 2018).

The legumes are extremely diverse and can be distributed into 3 subfamilies viz. *Caesalpinioideae, Mimosoideae* and *Papilionoideae* (Doyle and Luckow, 2003). The subfamily, Papilionoideae consists of almost all commercially important legumes like *Arachis hypogaea* (peanut), *Cicer arietinum* (chick pea), *Glycine max* (soybean), *Phaseolus vulgaris* (common bean), *Medicago sativa* (alfalfa), *Vigna radiata* (mungbean), *Lens culinaris* (lentil) and *Pisum sativum* (pea). Although these crop legumes have close phylogenetic relationship but differ extremely in their chromosome number, genome size and ploidy level (Table 1). However, previous studies on comparative genetic mapping suggested that the Papilionoideae subfamily members showed broad genome conservation (Weeden et al., 1992; Zhu et al., 2005). Young et al (2003) showed that legumes form a systematic taxonomic group with ubiquitous and prevailing macro- and micro- synteny. There are reports indicating that closely related species contain genomes having various genes with same map positions (Bennetzen, 2000; Devos and Gale, 2000; Paterson et al., 2000; Schmidt, 2000; Gualtieri et al., 2002).

Earlier, in plants, synteny analyses were aimed only on the species belonging to the families *Poaceae* and *Brassicaceae*. In recent time, such studies are being extended on other plants especially legumes (Gualtieri et al., 2002). It has been shown that sequence based tools help to study the evolution, organization and syntenic relationships of genomes. Macro- and micro- synteny of various species can be compared using linkage maps (Mc Connell et al., 2010). The studies indicated that level of synteny is found high in closely related species, and it decreases with growing phylogenetic distance (Choi et al., 2004). Recently, a linkage map for *Apios Americana* (a tuberous perennial legume in phaseolae tribe) has been reported and it showed synteny with selected warm-seasoned legumes. It also revealed a translocation event in *Glycine max* and common bean against *Apios* and *Vigna* species (Singh et al., 2018).

The two most important members of legume family Phaseoleae are soybean and common bean. These two species are commercially valuable legumes, common bean as a nutritional crop for poor population and soybean for its various human and animal uses (Mc Clean et al., 2010). Common bean has been diverged from soybean 19 million years ago. Considering syntenic relationship between two species is crucial to determine the potential for comparative genomic analysis. Whole genome duplication is one of the most salient characteristics of the soybean genome (Cannon and Shoemaker, 2012).

The RFLP mapping (Shoemaker et al., 1996) and EST Ks analysis have provided substantial evidences that whole genome duplication (WGD) in soybean has occurred about 13 million years ago (Schmutz et al., 2010 and Schlueter et al., 2004). Boutin et al (1995) have reported a number of syntenic linkage blocks between common bean and soybean by exploiting shared RFLP markers. They used using low density RFLP mapping, they showed that these two species share a high degree of homology in sequence, however, synteny has been found over short blocks of genomes. A clear one to two relationship between common bean and soybean genomes was shown by Lee et al (2001). A few synteny analyses have been reported between soybean and common bean. Mc Clean et al., (2010) studied the synteny

Table 1. Genomic information of nine legume species.

Species name	Common name	Genome size (Mbp)	Chr. No.	Gene number	Remark	Reference
Medicago truncatula	Barrel medic	470	2n = 2x = 16	50,894	WGS	Young et al. 2011
Medicago sativa	Alfafa	830–860	2n = 4x = 32	NA	–	Bauchan and Hossain2001
Pisum sativum	Pea	4300	2n = 2x = 14	NA	–	Franssen et al. 2011
Cicer arietinum	Chickpea	864	2n = 2x = 16	28,269	WGS	Varshney et al. 2013
Lotus japonicas	Bird's-foot trefoil	471	2n = 2x = 12	39,735	WGS	Sato et al. 2008
Vigna radiata	Mung bean	333	2n = 2x = 22	22,368	WGS	Kang et al. 2014
Glycine max	Soybean	1115	2n = 2x = 40	56,044	WGS	Schmutz et al. 2010
Phaseolus vulgaris	Common bean	625	2n = 2x = 22	38,482	WGS	Schmutz et al. 2014
Arachis hypogaea	Pea nut, ground nut	2,800	2n = 4x = 40	50,324	WGS	Chen et al. 2016

*NA- Not available, WGS - Whole genome sequencing completed.

between soybean and common bean on the basis of 300 gene base loci. Synteny blocks of averaging 32 cM in common bean and 4.9 Mb in soybean were found for all 11 common bean linkage groups which were mapped to all soybean linkage groups. A total 55 syntenic blocks each having 7 loci were observed between the common bean and soybean. By analysing the location of these blocks, it was revealed that each locus of common bean genome mapped to two loci of soybean genome (Mc Clean et al., 2010).

In 2014, Schmutz and co-workers observed considerable synteny between *Phaseolus vulgaris* and *Glycine max*, except in pericentromeric regions, where due to genomic expansion in one or both genomes micro- colinearity was much extended and due to that was lesser dense. Conservation of genome macrostructure has been reported between the soybean and other legumes including *Phaseolus vulgaris* (Lee et al., 2017). Studies showed the correlation between *Phaseolus* linkage group Pv01 and *Glycine max* chromosomes Gm06 and Gm04 (Cannon and Shoemaker 2012). It is also reported that Gm5 and Gm8 chromosomes share synteny to common bean Pv2 (Mc Clean et al., 2010). In another study, nearly 63% of common bean unigenes were shown to have homology to soybean (Kalavacharla et al., 2011). Microsynteny analysis was also reported between common bean and soybean where 6 BAC clones were sequenced and analyzed for microsynteny (Yadegari, 2013). Recently the genomes of *G.max* and *P. vulgaris* were analyzed for investigating synteny. Syntenic blocks were found between chromosomes Gm03 and Pv10, Gm10 and Pv07 as well as Gm14 and Pv01. A large number of present research make use of whole genome sequences to study *R* genes in legumes (Anderson 2016, Christie et. al., 2016 and Neupane et. al., 2018).

Earlier many research were done on NBS-LRR genes in *Glycine max* and other legumes by using bioinformatics approach (Benson 2014, Shao et. al., 2016 and Neupane et. al., 2018). Nepal and Benson (2015) showed the complete evolutionary relationship of the CNL- *R* genes in soybean and common bean. In the present study, synteny analysis between chromosome 13 of soybean having NBS-LRR enriched sequences and whole genome of common bean have been carried out since common bean has been reported to have NBS-LRR enriched regions which help in disease resistance. The results from our study will be helpful in understanding evolutionary relationships of NBS-LRR genes with potential implication in crop improvement.

MATERIALS AND METHODS

The gene synteny alignment, and genomic linkage analyses between the chromosome 13 of *G. max* and whole genome of *P. vulgaris* was performed by using 'SATSUMA version 3.1.0 freely available on the website, http://satsuma.sourceforge.net/. All the data in FASTA and GFF files required for genome sequence and annotation of soybean and common bean were downloaded from the NCBI database (https://www.ncbi.nlm.nih.gov/genome/). For synteny analysis, only sequence alignments of more than 200 bp and percent identity of more than 85% were considered and rest were filtered out. The results were visualized by developing custom script in Biopython module of BioPython, version 3.7 freely available on the website, http://biopython.org/DIST/docs/install/Installation.html.

RESULTS AND DISCUSSION

Comparison of orthologous regions of soybean and common bean

The syntenic analysis between soybean and common bean showed many conserved regions between the two as shown in Fig. 1. Twenty one genes mentioned in Table 2, which were screened on the basis of few parameters of our interest as highest, lowest percent similarity, protein kinase, diseases resistance protein, NBS-LRR proteins represent less than 1/4th of the genes we have obtained after synteny mapping. These genes were mentioned with their gene id, locus tag along with the product formed as shown in table 2. Soybean chromosome 13 region having start from 22449055 bp and up to 22449570 bp showed gene homology with the sequence starting from 31929662 to 31930177 of *Phaseolus vulgaris*. This genomic region in both the species codes for guanine nucleotide binding protein subunit gamma 3-like isoform X1. The guanine nucleotide binding protein is important for signalling pathway. It is to be noted that the Rpg1b genes in soybean mainly codes for NB and LRR proteins. Various NB-LRR protein coding genes which are present on chromosome 13 of *G. max* are also showing synteny with *P. vulgaris*. For instance, F-box/LRR-repeat protein 4-like (LOC100816140) is located at 27041793 bp to 27042286 bp in *Phaseolus vulgaris*. A mitogen-activated protein kinase2 (LOC100789734) located on chromosome 13 of *Glycine max* is showing homology with *Phaseolus vulgaris*.

A G type lectin S-receptors like serine/threonine-protein kinase coding gene located at chromosome 13 in *Glycine max* is also present in *Phaseolus vulgaris* genomic region having starts from 30144768 bp to 30146245 bp. It is having several molecular function like ATP binding, Calmodulin binding and protein serine/threonine kinase activity and few biological functions like protein phosphorylation and recognition of pollen. Glycogen phosphorylase 1-like isoform X1 coding gene which is present

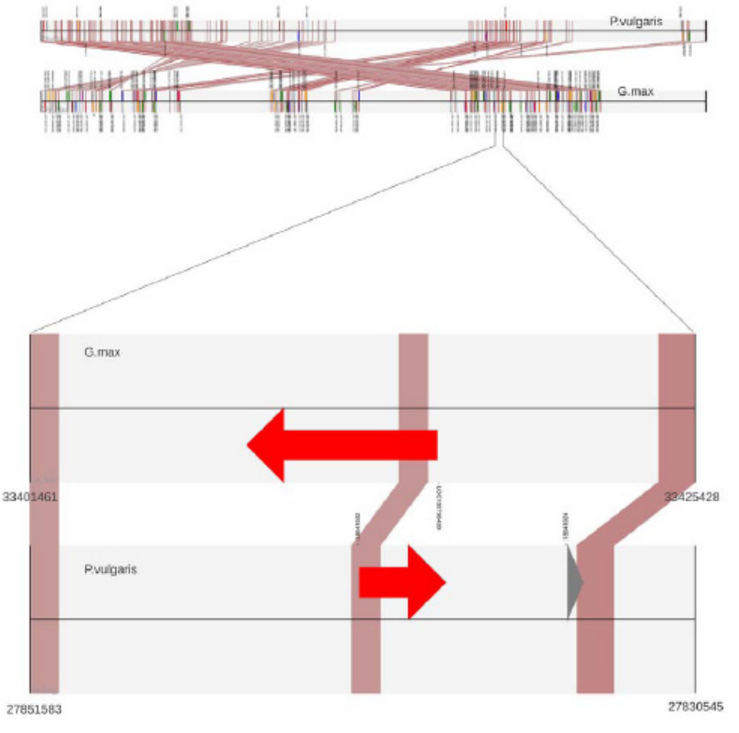

FIGURE 1. Syntenic regions between *Glycine max* and *Phaseolus vulgaris*. A synteny region has been amplified where synteny length has been indicated using colors. The colors of the links are based on the magnitude of synteny length (Light grey color indicates minimum synteny length whereas red brick color maximum synteny length).

at locus LOC100779066 on *Glycine max* chromosome 13 is present in *Phaseolus vulgaris*. Uncharacterized ATP-dependent helicase C23E6.02-like (LOC100792901) and B3 domain-containing protein Os07g0563300-like both were conserved in these legumes namely *Glycine max* and *Phaseolus vulgaris*. Highest synteny was found to be 91.794 % similarity between *Glycine max* and *Phaseolus vulgaris* for a region which codes for an uncharacterized protein LOC100775655 in both the legume. Similarly, lowest synteny was found to be 85.049 % similarity at loci LOC102661544 of *Glycine max*.

Various syntenic regions were found between the two legumes, which codes for different proteins like MAP Kinase kinase2, Guanine nucleotide binding protein, 1-phosphatidylinositol-3-Phosphate 5-kinase and several Serine threonine kinases. All the proteins have different molecular and biological functions like ATP and ADP binding, Kinase activity, Metal and Calcium Ion binding, signal transduction, defense responses respectively. Different gene functions and their position in soybean and common bean are described with details in Table 3.

In the present study, a different approach has been used for comparison of soybean genome with common bean. Advantage of synteny analysis is based on the concept that the species which are evolutionary related are diverged from their common ancestor and conserved genome synteny can be efficiently interpreted from a well-studied species to another less characterized genomes. Synteny is found to be higher between the closely related species (Lee et. al., 2017) and this concept is confirmed in this study. The outcome from the analysis identifies the best match to the *Glycine max* genome. Here focus was done only on chromosome 13 of *Glycine max*. We selected this chromosome because in *Glycine max,* diseases resistance coding genes are known to be found mostly on this region (Chr. 13). By the comparison of both the legumes sequences we get NB-LRR protein coding genes, which are responsible for disease resistance in *Glycine max* and *Phaseolus vulgaris*.

Identification of conserved synteny has noticeable advantage in understanding the legumes genetics. Large region of synteny occurs between *Glycine max* and *Phaseolus vulgaris* (Mc Clean et. al., 2010), our results also

Table 2. Comparison of Soybean (*Glycine_max* cultivar_Williams_82 chromosome_13, *Glycine_max*_v2.0, whole_genome shotgun sequence) with *Phaseolus vulgaris*.

Taregt_ID	Target Start	Target End	Query ID	Query Start	Query End	Identity	Strand	Synteny Length	Gene Start	Gene End	Gene and Locus tag	Product
Glycine max and Phaseolus vulgaris												
NC_016100.2_ Glycine_max cultivar_ Williams_82 chromosome_13, Glycine_max_v2.0, whole_genome shotgun_sequence	22120404	22121394	PhaVulg_2	32502275	32503260	86.3636	-	986	22121113	22121337	LOC100789734 GLYMA_13G106900	mitogen-activated protein kinase kinase 2
	22449055	22449570	PhaVulg_2	31929662	31930177	85.2427	-	516	22449124	22449517	LOC100787066 GLYMA_13G110900	guanine nucleotide-binding protein subunit gamma 3-like isoform X1
	22847546	22849062	PhaVulg_2	30908984	30910502	87.7309	-	1519	22848615	22848943	LOC100811146 GLYMA_13G114800	1-phosphatidylinositol-3-phosphate 5-kinase FAB1B-like%2C transcript variant X1
	22847546	22849062	PhaVulg_2	30908984	30910502	87.7309	-	1519	22847654	22848520	LOC100811146 GLYMA_13G114800	1-phosphatidylinositol-3-phosphate 5-kinase FAB1B-like isoform X1
	23222659	23223739	PhaVulg_2	30060722	30061802	87.037	-	1081	23222716	23223648	LOC100778006 GLYMA_13G119600	probable GTP diphosphokinase CRSH%2C chloroplastic
	22110388	22111570	PhaVulg_2	32514048	32515233	91.794	-	1186	22110445	22111163	LOC100775655 GLYMA_13G106700	uncharacterized protein LOC100775655
	24110505	24111213	PhaVulg_6	37793082	37793792	87.0056	+	711	24110547	24110936	LOC100784243 GLYMA_13G128200	receptor-like serine/threonine-protein kinase ALE2%2C transcript variant X1
	24110505	24111213	PhaVulg_6	37793082	37793792	87.0056	+	711	24110547	24110936	LOC100784243 GLYMA_13G128200	receptor-like serine/threonine-protein kinase ALE2 isoform X1
	24933885	24934843	PhaVulg_6	39244442	39245400	85.2818	+	959	24934259	24934787	LOC100798007 GLYMA_13G136700	probable LRR receptor-like serine/threonine-protein kinase At1g67720
	27466724	27467253	PhaVulg_2	37761025	37761551	88.6578	+	527	27466879	27467198	LOC100305441 GLYMA_13G159400	protein kinase

	NC_016100.2_ Glycine_max cultivar_ Williams_82 chromosome_13, Glycine_max_ v2.0, whole_genome shotgun_ sequence	32416131	32416981	PhaVulg_5	26826933	26827782	85.7647	+	850	32416429	32416584	LOC100814013 GLYMA_13G210300	mitogen-activated protein kinase 15-like
		32523009	32523502	PhaVulg_5	27041793	27042286	86.4097	+	494	32523037	32523443	LOC100816140 GLYMA_13G211600	F-box/LRR-repeat protein 4-like%2C transcript variant X1
		32779883	32782745	PhaVulg_5	27246593	27249457	89.4479	+	2865	32779922	32781577	LOC100777288 GLYMA_13G214400	serine/threonine-protein kinase RUNKEL-like
		32848947	32850530	PhaVulg_5	27294165	27295744	88.0606	+	1580	32850020	32850505	LOC100778885 GLYMA_13G215000	inactive beta-amylase 9
		34551754	34552841	PhaVulg_5	29104549	29105630	87.3965	+	1082	34551796	34552806	LOC100792540 GLYMA_13G234800	serine/threonine-protein kinase-like protein ACR4
		34706927	34708223	PhaVulg_5	29239944	29241240	86.6512	+	1297	34707002	34708057	LOC100798896 GLYMA_13G236700	putative leucine-rich repeat receptor-like serine/threonine-protein kinase At2g14440%2C transcript variant X2
		34706927	34708223	PhaVulg_5	29239944	29241240	86.6512	+	1297	34707002	34708057	LOC100798896 GLYMA_13G236700	putative leucine-rich repeat receptor-like serine/threonine-protein kinase At2g 14440 isoform X1
		35582629	35584126	PhaVulg_5	30027808	30029305	90.314	+	1498	35582677	35584018	LOC100785128 GLYMA_13G247300	probable serine/ threonine-protein kinase abkC isoform X1
		35582629	35584126	PhaVulg_5	30027808	30029305	90.314	+	1498	35582677	35584018	LOC100785128 GLYMA_13G247300	probable serine/ threonine-protein kinase abkC%2C transcript variant X4
		35712009	35713485	PhaVulg_5	30144768	30146245	89.0244	+	1478	35712083	35713370	LOC100804024 GLYMA_13G249300	G-type lectin S-receptor-like serine/ threonine-protein kinase At4g27290 isoform X1
		31895480	31897607	PhaVulg_5	26318633	26320755	85.049	+	2123	31895602	31895959	LOC102661544 GLYMA_13G204900	uncharacterized protein LOC102661544

Table 3. Different functions of gene products of *Glycine max* and *Phaseolus vulgaris*.

Gene Location in *Glycine max*	Gene products	Functions	
		Molecular	Biological
LOC100789734	mitogen-activated protein kinase kinase 2	ATP binding, MAP kinase kinase activity and protein serine/threonine kinase activity	Activation of MAPK activity, activation of protein kinase activity, auxin transport, cold acclimation, defense response, incompatible interaction, response to cold, response to salt stress, signal transduction by protein phosphorylation, stress-activated protein kinase signalling cascade and xylem and phloem pattern formation
LOC100787066	guanine nucleotide-binding protein subunit gamma 3-like isoform X1	-	G protein-coupled receptor signalling pathway
LOC100811146	1-phosphatidylinositol-3-phosphate 5-kinase FAB1B-like%2C transcript variant X1	ATP binding, metal ion binding and phosphatidylinositol phosphate kinase activity	-
LOC100778006	probable GTP diphospho kinase CRSH%2C chloroplastic	Calcium ion binding and kinase activity	Guanosine tetra phosphate metabolic process
LOC100775655	uncharacterized protein LOC100775655		
LOC100784243	receptor-like serine/threonine-protein kinase ALE2%2C transcript variant X1	ATP binding and protein kinase activity	-
LOC100798007	probable LRR receptor-like serine/threonine-protein kinase At1g67720	ATP binding and protein kinase activity	-
LOC100814013	mitogen-activated protein kinase 15-like	ATP binding and protein kinase activity	-

LOC100816140	F-box/LRR-repeat protein 4-like%2C transcript variant X1	Peptidyl-prolyl cis-trans isomerase activity and unfolded protein binding	-
LOC100777288	serine/threonine-protein kinase RUNKEL-like	ATP binding and protein kinase activity	-
LOC100778885	inactive beta-amylase 9	Beta-amylase activity	Polysaccharide catabolic process
LOC100792540	serine/threonine-protein kinase-like protein ACR4	ATP binding and protein kinase activity	-
LOC100798896	putative leucine-rich repeat receptor-like serine/threonine-protein kinase At2g14440%2C transcript variant X2	Kinase activity	-
LOC100785128	probable serine/threonine-protein kinase abkC%2C transcript variant X4	Kinase activity	-
LOC100809556	homeobox-leucine zipper protein HDG2	DNA-binding transcription factor activity, lipid binding and sequence-specific DNA binding	Maintenance of floral organ identity and trichome-morphogenesis
LOC100801911	putative E3 ubiquitin-protein ligase RING1a	Metal ion binding and transferase activity	Cell fate determination, chromatin organization, maintenance of floral meristem identity, maintenance of inflorescence meristem identity, maintenance of shoot apical meristem identity, negative regulation of gene expression, epigenetic, negative regulation of transcription, DNA-templated and protein ubiquitination
LOC100776037	V-type proton ATPase subunit a1-like isoform X2	ATPase binding, proton-transporting ATPase activity and rotational mechanism	ATP hydrolysis coupled proton transport, proton-transporting V-type ATPase complex assembly and vacuolar acidification
LOC100792901	uncharacterized ATP-dependent helicase C23E6.02-like%2C transcript variant X2	ATP binding, helicase activity and metal ion binding	-
LOC100782114	B3 domain-containing protein Os07g0563300-like	DNA binding and zinc ion binding	Regulation of transcription, DNA-templated transcription, and DNA-templated
Source: InterPro, UniProtKB-UniRule,UniProtKB-KW, GO_Central, TAIR.			

showed the gene conservation between both the legumes. Studies of the NBS-LRR gene family in legume plants can provide knowledge on the genomic and molecular mechanism that form the basis of gene regulation and protein functions. In this study, various NBS-LRR proteins (as mentioned in table 3 showing different functions of gene products) were found to be conserved in *Glycine max* and *Phaseolus vulgaris*. Now a day's gene cloning approach is advancing to make disease resistance varieties of legumes as in *Medicago sativa* cloning of a disease resistance gene was done by utilizing *Medicago truncatula* gene information through synteny (Yang et. at., 2008). In the same way this synteny analysis result will be useful for evolutionary studies that help in long term planning, breeding and developing disease resistance varieties of legumes.

ACKNOWLEDGEMENTS

Authors acknowledge the facilities of the Department of Biotechnology, Ministry of Science and Technology, Government of India, New Delhi (DBT) under the Bioinformatics Sub Centre as well as the golden jubilee research fellowship (GJRF) provided by School of Biotechnology, Devi Ahilya University Indore. M.P. India.

Conflict of Interest

The authors confirm that they have no conflict of interest.

REFERENCES

Andersen, E. J., Ali, S., Reese, R. N., Yen, Y., Neupane, S., & Nepal, M. P. (2016). Diversity and evolution of disease resistance genes in barley (*Hordeum vulgare* L.). Evolutionary Bioinformatics, 12, EBO-S38085.

Andolfo, G., & Ercolano, M. R. (2015). Plant innate immunity multicomponent model. Frontiers in plant science, 6, 987.

Bauchan, G. R., & Hossain, M. A. (2001). Distribution and characterization of heterochromatic DNA in the tetraploid African population alfalfa genome. Crop science, 41(6), 1921-1926.

Bennetzen, J. L. (2000). Comparative sequence analysis of plant nuclear genomes: microcolinearity and its many exceptions. The Plant Cell, 12(7), 1021-1029.

Benson, B. V. (2014). Disease resistance genes and their evolutionary history in six plant species.

Boutin, S. R., Young, N. D., Olson, T., Yu, Z. H., Vallejos, C. E., & Shoemaker, R. C. (1995). Genome conservation among three legume genera detected with DNA markers. Genome, 38(5), 928-937.

Cannon, S. B., & Shoemaker, R. C. (2012). Evolutionary and comparative analyses of the soybean genome. Breeding science, 61(5), 437-444.

Chen, X., Li, H., Pandey, M. K., Yang, Q., Wang, X., Garg, V., ... & Upadhyaya, H. (2016). Draft genome of the peanut A-genome progenitor (*Arachis duranensis*) provides insights into geocarpy, oil biosynthesis, and allergens. Proceedings of the National Academy of Sciences, 113(24), 6785-6790.

Choi, H. K., Mun, J. H., Kim, D. J., Zhu, H., Baek, J. M., Mudge, J., ... & Young, N. D. (2004). Estimating genome conservation between crop and model legume species. Proceedings of the National Academy of Sciences, 101(43), 15289-15294.

Christie, N., Tobias, P. A., Naidoo, S., & Külheim, C. (2016). The *Eucalyptus grandis* NBS-LRR gene family: physical clustering and expression hotspots. Frontiers in plant science, 6, 1238.

Devos, K. M., & Gale, M. D. (2000). Genome relationships: the grass model in current research. The Plant Cell, 12(5), 637-646.

Doyle, J. J., & Luckow, M. A. (2003). The rest of the iceberg. Legume diversity and evolution in a phylogenetic context. Plant Physiology, 131(3), 900-910.

Franssen, S. U., Shrestha, R. P., Bräutigam, A., Bornberg-Bauer, E., & Weber, A. P. (2011). Comprehensive transcriptome analysis of the highly complex *Pisum sativum* genome using next generation sequencing. BMC genomics, 12(1), 227.

Gualtieri, G., Kulikova, O., Limpens, E., Kim, D. J., Cook, D. R., Bisseling, T., & Geurts, R. (2002). Microsynteny between pea and *Medicago truncatula* in the SYM2 region. Plant molecular biology, 50(2), 225-235.

Gururani, M. A., Venkatesh, J., Upadhyaya, C. P., Nookaraju, A., Pandey, S. K., & Park, S. W. (2012). Plant disease resistance genes: current status and future directions. Physiological and molecular plant pathology, 78, 51-65.

Hulbert, S. H., Webb, C. A., Smith, S. M., & Sun, Q. (2001). Resistance gene complexes: evolution and utilization. Annual review of phytopathology, 39(1), 285-312.

Kalavacharla, V., Liu, Z., Meyers, B. C., Thimmapuram, J., & Melmaiee, K. (2011). Identification and analysis of common bean (*Phaseolus vulgaris* L.) transcriptomes by massively parallel pyrosequencing. BMC plant biology, 11(1), 135.

Kang, Y. J., Kim, S. K., Kim, M. Y., Lestari, P., Kim, K. H., Ha, B. K., ... & Shim, S. (2014). Genome sequence of mungbean and insights into evolution within *Vigna* species. Nature communications, 5, ncomms6443.

Lee, C., Yu, D., Choi, H. K., & Kim, R. W. (2017). Reconstruction of a composite comparative map composed of ten legume genomes. Genes & Genomics, 39(1), 111-119.

Lee, J. M., Grant, D., Vallejos, C. E., & Shoemaker, R. C. (2001). Genome organization in dicots. II. Arabidopsis as a'bridging species' to resolve genome evolution events among legumes. Theoretical and Applied Genetics, 103(5), 765-773.

McClean, P. E., Mamidi, S., McConnell, M., Chikara, S., & Lee, R. (2010). Synteny mapping between common bean and soybean reveals extensive blocks of shared loci. BMC genomics, 11(1), 184.

McConnell, M., Mamidi, S., Lee, R., Chikara, S., Rossi, M., Papa, R., & McClean, P. (2010). Syntenic relationships among

legumes revealed using a gene-based genetic linkage map of common bean (*Phaseolus vulgaris* L.). Theoretical and Applied Genetics, 121(6), 1103-1116.

Nepal, M. P., & Benson, B. V. (2015). CNL disease resistance genes in soybean and their evolutionary divergence. Evolutionary Bioinformatics, 11, EBO-S21782.

Neupane, S., Ma, Q., Mathew, F. M., Varenhorst, A. J., Andersen, E. J., & Nepal, M. P. (2018). Evolutionary Divergence of TNL Disease-Resistant Proteins in Soybean (*Glycine max*) and Common Bean (*Phaseolus vulgaris*). Biochemical genetics, 1-26.

Passoupathy, R. (2016). Synteny with Allied and Model Genomes. 10.1007/978-3-319-47789-3_6.

Paterson, A. H., Bowers, J. E., Burow, M. D., Draye, X., Elsik, C. G., Jiang, C. X., ... & Wright, R. J. (2000). Comparative genomics of plant chromosomes. The Plant Cell, 12(9), 1523-1539.

Sato, S., Nakamura, Y., Kaneko, T., Asamizu, E., Kato, T., Nakao, M., & Fujishiro, T. (2008). Genome structure of the legume, Lotus japonicus. DNA research, 15(4), 227-239.

Schlueter, J. A., Dixon, P., Granger, C., & Shoemaker, R. C. (2005). Mining the EST databases to determine evolutionary events in the legumes and grasses. In Genome Exploitation (pp. 163-181). Springer, Boston, MA.

Schmidt, R. (2000). Synteny: recent advances and future prospects. Current opinion in plant biology, 3(2), 97-102.

Schmutz, J., Cannon, S. B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., ... & Xu, D. (2010). Genome sequence of the palaeopolyploid soybean. Nature, 463(7278), 178.

Schmutz, J., McClean, P. E., Mamidi, S., Wu, G. A., Cannon, S. B., Grimwood, J., ... & Torres-Torres, M. (2014). A reference genome for common bean and genome-wide analysis of dual domestications. Nature genetics, 46(7), 707.

Shao, Z. Q., Xue, J. Y., Wu, P., Zhang, Y. M., Wu, Y., Hang, Y. Y., ... & Chen, J. Q. (2016). Large-scale analyses of angiosperm nucleotide-binding site-leucine-rich repeat (NBS-LRR) genes reveal three anciently diverged classes with distinct evolutionary patterns. Plant Physiology, pp-01487.

Shoemaker, R. C., Polzin, K., Labate, J., Specht, J., Brummer, E. C., Olson, T., ... & Kochert, G. (1996). Genome duplication in soybean (*Glycine subgenus soja*). Genetics, 144(1), 329-338.

Singh, J., Kalberer, S. R., Belamkar, V., Assefa, T., Nelson, M. N., Farmer, A. D., ... & Cannon, S. B. (2018). A transcriptome-SNP-derived linkage map of *Apios americana* (potato bean) provides insights about genome re-organization and synteny conservation in the phaseoloid legumes. Theoretical and Applied Genetics, 131(2), 333-351.

Varshney, R. K., Song, C., Saxena, R. K., Azam, S., Yu, S., Sharpe, A. G., ... & Millan, T. (2013). Draft genome sequence of chickpea *(Cicer arietinum*) provides a resource for trait improvement. Nature biotechnology, 31(3), 240.

Weeden, N. F., Muehlbauer, F. J., & Ladizinsky, G. (1992). Extensive conservation of linkage relationships between pea and lentil genetic maps. Journal of Heredity, 83(2), 123-129.

Yadegari, Z. (2013). Molecular mapping and characterization of phenylpropanoid pathway genes in common bean (*Phaseolus vulgaris* L.) (Doctoral dissertation).

Young, N. D., Debellé, F., Oldroyd, G. E., Geurts, R., Cannon, S. B., Udvardi, M. K., ... & Van de Peer, Y. (2011). The Medicago genome provides insight into the evolution of rhizobial symbioses. Nature, 480(7378), 520.

Young, N. D., Mudge, J., & Ellis, T. N. (2003). Legume genomes: more than peas in a pod. Current opinion in plant biology, 6(2), 199-204.

Zhu, H., Choi, H. K., Cook, D. R., & Shoemaker, R. C. (2005). Bridging model and crop legumes through comparative genomics. Plant physiology, 137(4), 1189-1196.