Whole Genome Re-sequencing of Soybean Accession EC241780 Providing Genomic Landscape of Candidate Genes Involved in Rust Resistance

Milind Balkrishna Ratnaparkhe1,*, Niharika Marmat1,2, Giriraj Kumawat1, Maranana Shivakumar1, Viraj Gangadhar Kamble3, Vennampally Nataraj1, Shunmugiah Veluchamy Ramesh1, Milind Panjabrao Deshmukh3, Ajay Kumar Singh4, Humira Sonah5, Rupesh Kailasrao Deshmukh5, Manoj Prasad6, Suresh Chand2 and Sanjay Gupta1

1ICAR-Indian Institute of Soybean Research, Indore-452001, MP, India; 2School of Life Sciences, Devi Ahilya University, Indore, MP, India; 3Agricultural Research Station, (MPKV), Kasbe Digraj, Sangli, MS, India; 4ICAR-National Institute of Abiotic Stress Management, Baramati, MS, India; 5National Agri-Food Biotechnology Institute, Mohali, India; 6National Institute of Plant Genome Research, New Delhi, India

Abstract: Background: In this study, whole genome re-sequencing of rust resistant soybean genotype EC241780 was performed to understand the genomic landscape involved in the resistance mechanism.

Methods: A total of 374 million raw reads were obtained with paired-end sequencing performed with Illumina HiSeq 2500 instrument, out of which 287.3 million high quality reads were mapped to Williams 82 reference genome. Comparative sequence analysis of EC241780 with rust susceptible cultivars Williams 82 and JS 335 was performed to identify sequence variation and to prioritise the candidate genes.

Results: Comparative analysis indicates that genotype EC241780 has high sequence similarity with rust resistant genotype PI 200492 and the resistance in EC241780 is conferred by the Rpp1 locus. Based on the sequence variations and functional annotations, three genes Glyma18G51715, Glyma18G51741 and Glyma18G51765 encoding for NBS-LRR family protein were identified as the most prominent candidate for Rpp1 locus.

Conclusion: The study provides insights of genome-wide sequence variation more particularly at Rpp1 loci which will help to develop rust resistant soybean cultivars through efficient exploration of the genomic resource.

Keywords: Soybean, whole genome re-sequencing, disease resistance, rust, Phakopsora pachyrhizi, single nucleotide polymorphisms (SNPs).

1. INTRODUCTION

Soybean, one of the most important oilseed crop worldwide and is severely affected by Soybean rust. Soybean rust is caused by a devastative pathogen Phakopsora pachyrhizi Syd. & P. Syd and spread of the disease occurs through wind disseminated urediniospores. Urediniospores can develop rapidly, causing loss of foliar area and ultimately leading to a severe reduction in seed yield. Soybean rust was reported for the first time in 1902 in Japan, and subsequently in other parts of Asia and Australia during 1934 [1]. Later on, by 2005, soybean rust spread to most of the soybean growing area worldwide, including north and south America [2-4]. Soybean rust is now a global problem that requires urgent attention to secure oil seed production under continuously increasing market demand.

Development of resistant cultivars is considered as most effective approach for the soybean rust control. Resistance to rust pathogen was found in wild species of the genera Glycine [5, 6] and resistance is mediated by a hypersensitive reaction (HR). This is a common type of response triggered when plant resistance genes (R-genes) are challenged by pathogen avirulence genes (Avr-genes) [7]. Resistant genotypes show red–brown (RB) lesions and a spectrum ranging from no to high levels of sporulation depending on the genotype. Susceptible genotypes are characterized mostly by light brown (TAN) lesions with profuse sporulation.

Host plant resistance mediated by Rpp (Resistance to P. pachyrhizi) genes has been found in numerous soybean accessions. About 10 Rpp genes along with different alleles have been mapped to seven genetic loci namely Rpp1, Rpp2, Rpp3, Rpp4, Rpp5, Rpp6 and Rpp7 [8-12]. Garcia et al. also discovered a recessive resistance allele (rpp5) in PI 200456 [13]. Monteros et al. mapped a resistance gene (designated Rpp2Hyuaga) at the Rpp3 locus in the Japanese cultivar Hyuaga (PI 506764) [14], and Kendrick et al. later found a second resistance gene in Hyuaga at the Rpp5 locus [10].
Chakraborty et al. identified Rpp1-b, a different dominant resistance allele at the Rpp1 locus in PI 594538A [15], Ray et al. discovered additional resistance alleles at the Rpp1 locus in PI 587886 and PI 587880A [9], and Hossain et al. also reported other alleles at the Rpp1 locus in PI 594767A and PI 587905 [16].

Recent development in Next-generation Sequencing (NGS) technologies provides opportunities to perform whole genome re-sequencing in a cost effective manner [17, 18]. One of the many applications of whole genome re-sequencing is to understand the genetic basis of phenotypic differences by comparing the genomic sequences. In this regard, in the present investigation, whole genome re-sequencing of rust resistant soybean genotype EC241780 and large scale comparative sequence analysis with susceptible cultivars Williams 82 and JS 335 was performed. The sequences were analysed to discover a large number of DNA polymorphisms including SNPs and InDels in the two parental lines. Further, the functional relevance of SNPs and InDel were analysed based on their presence in intergenic regions, coding and non-coding regions. Functional impact of the SNP on the functionality of the protein was predicted using computational approaches. To our knowledge, this is the first report of whole genome re-sequencing of soybean rust resistant line with sequence depth of ~30X. The results provided here can be used in future studies on high-throughput genotyping and molecular breeding for resistance to rust in soybean.

2. MATERIALS AND METHODS

2.1. Plant Sample Preparation and Sequencing

For whole genome re-sequencing, genomic DNA from the leaves of 22 days old soybean genotype EC241780 was extracted using Qiagen DNeasy kit (Qiagen, GmbH, Hilden). The quality and quantity of the extracted DNA were determined using Bioanalyzer 2100 (Agilent Technologies, Singapore) and Qubit 2.0 Fluorometer (Invitrogen Life Technologies). The library preparation was performed according to the manufacturer’s instructions. Paired-end sequencing of the prepared libraries was performed using Illumina HiSeq 2500 platform (Illumina Technologies) in order to generate 100-bp paired-end reads. Only high quality sequence reads were employed for downstream analysis by trimming low quality reads.

2.2. Genomic Resources Used for Comparative Genomics

For comparative analysis, whole genome re-sequenced data of rust susceptible cultivar JS 335 reported by Yadav et al. was obtained from the National Institute of Plant Genome Research (NIPGR), New Delhi, India [19]. In addition, Williams 82 reference genome was also used for the comparative genomics.

2.3. Read Mapping

The high quality paired-end reads obtained from whole genome sequencing were aligned to the Williams 82 reference genome (V1.0.26) retrieved from the Ensembl database (http://plants.ensembl.org/) [20]. The sequence alignment was performed using the Burrows-Wheeler Alignment Tool (BWA software available at http://bio-bwa.sourceforge.net) [21] with default parameters [22]. The mapping output of BWA program was further filtered to retain only those reads that mapped to the reference genome at only one position. Base quality score recalibration and estimates on the coverage of the reference genome were performed using FastQC v0.10.1 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) [23] and SAMtools v0.1.18 (https://sourceforge.net/projects/samtools/files/samtools/) [24], respectively. The aligned reads were initially sorted using Picard tool and then the read duplicates were removed.

2.4. Identification and Functional Annotations of SNPs and InDels

The high-quality reads were used for the identification of SNPs and InDels by individually comparing the EC241780 genome with the Williams 82 reference genomes and JS 335. Variant calling was performed using Genome Analysis TKLite-2.3-9 (GATK) toolkit Unified Genotyper to identify SNPs and InDels. The identified variants were further filtered based on default parameters including depth, and variant score along with slight modifications to retain quality variants. The stringent criteria used for variant filtering were minimum variant frequency of ≥90%, average quality of the SNP base ≥30 and minimum read depth of 10. Genomic distribution of variant calling was performed by studying the frequency and position of SNPs and InDels for each chromosome.

2.5. Annotation of Genomic and Genic SNPs and InDels

Soybean annotation files were obtained from the Ensembl Database (http://plants.ensembl.org/) [20] in order to identify gene positions, functional descriptions, locations of introns, exons and UTRs. SNPs and InDels identified from EC241780 were annotated using data from the Ensemble database and using in-house bioinformatics pipeline (VARI-MAT – Variation and Mutation Annotation Toolkit). Similarly, functional annotation of the SNPs was performed to identify synonymous and non-synonymous substitutions.

2.6. Haplotype Analysis

Publicly available 50 thousand SNP data for 19,652 Glycine max accessions [25] were retrieved to perform haplotype analysis. Using the 50K SNP data, a unique haplotype was identified that spans the rust resistance locus identified in EC241780. Phylogenetic analysis of disease resistance genes was performed using MEGA7 software tool [26]. The evolutionary history was inferred using the UPGMA method.

3. RESULTS

3.1. Mapping of Illumina Reads

Whole genome re-sequencing of rust resistant line EC241780 was performed using Illumina HiSeq 2500 platform and quality reads were obtained after FastQ analysis. A total of 374 million raw reads were obtained and adaptor trimming removed 12792 reads (0.003%). High quality reads were mapped to soybean reference genome Williams 82. Overall 287.3 million reads (76.64 %) of total reads were mapped to the reference genome. Out of the total aligned reads, 203.5 million reads (70.84 %) mapped uniquely to the
3.2. Detection and Distribution of Variations

Using minimum default parameters implemented in GATK, initially, we identified a total of 2,082,134 SNPs and 311,361 InDels between the reference Williams 82 genome and soybean accession EC241780. Stringent parameters were further used to minimize the detection of false positive SNPs and InDels. Finally, only those SNPs and InDels with minimum read depth of 10, average base quality of 30 and polymorphism call rate of ≥90% were filtered. Based on these parameters, a total of 2,024,915 polymorphic loci were detected out of which 1,799,343 SNPs and 225,572 InDels were obtained. The distribution of DNA polymorphisms detected for Williams 82/EC241780 across the twenty soybean chromosomes were analysed. The number of SNPs and InDels varied across the chromosome (Fig. 1). The total number of SNPs and InDels within a chromosome was found to be directly proportional to the chromosome length. The largest number of SNPs and InDels were detected in chromosome 18 and smallest number in chromosome 12 (Supplementary Table S1).

Table 1. Summary of sequence reads derived from rust resistance soybean genotypes EC241780 which were mapped onto the reference genome (Williams 82).

| S. No. | Mapping details of EC241780     | -       |
|--------|--------------------------------|---------|
| 1.     | Total number of quality reads  | 374,907,382 |
| 2.     | Number of aligned reads        | 287,318,932 |
| 3.     | Percentage of aligned reads    | 76.64   |
| 4.     | Number of uniquely aligned reads | 203,558,177 |
| 5.     | Percentage of uniquely aligned reads | 70.84   |

Furthermore, SNPs and InDels were identified by comparing sequence of EC241780 and soybean cultivar JS 335, which is most widely grown soybean cultivar in India and has been commonly used as one of the parent for the development of mapping populations. Initially, 484,704 SNPs and 73,124 InDels were identified between JS 335 and EC241780. Subsequently, with the stringent parameters 414,924 SNPs and 48,962 InDels distributed on 20 chromosomes were identified (Supplementary Table S2). Between JS 335 and EC241780 the largest number of SNPs (68,428) and InDels (5008) were detected in chromosome 14 and chromosome 18 respectively, whereas smallest numbers of SNPs (5106) and InDels (1006) were detected in chromosome 13 and in chromosome 4 respectively.

3.3. Gene Annotation

The annotation of soybean reference genome was used to reveal the distribution of SNPs and InDels within various genomic regions such as intergenic and intragenic. SNPs and InDels identified from EC241780 were annotated using gene model annotation information for Glycine max present at Ensembl Database. Genomic annotation of variants (SNPs and InDels) in EC241780 revealed that the majority of the variants (~64% of total variants) fall in the intergenic region based on the current gene model. Only 2.9 % of SNPs and 1.03% of the InDels were identified within the coding region of the genome. A total of 1,751,039 SNPs and 275,713 InDels were identified in the intergenic region. Genic region comprised of 311,779 SNPs and 65,756 InDels. Within coding region, 75,905 SNPs and 3,437 InDels were identified, whereas, 191,705 SNPs and 46,673 InDels were present in the intronic region. The 5’ UTRs and 3’ UTRs regions also showed the presence of SNPs (44,169) and InDels (14,736) (Supplementary Table S3). Functional annotation of SNPs in EC241780 revealed that mostly the variants resulted in missense (53% of total variants) and silent (40% of variants) changes in the coding regions (Supplementary Table S4). To validate the variants, over 69 representative SNPs were selected from different positions and were validated using mass array genotyping approach (Supplementary Table S5).

3.4. Soybean Accession EC241780 Shows High Sequence Similarity With PI 200492 at Rpp1 Locus on Chromosome 18

To identify rust resistance gene, comparative analysis was performed between accession EC241780 and other soybean accessions. Single nucleotide polymorphisms (SNPs) from EC241780 were compared with SoySNP50K data of various rust resistant and susceptible accessions [25]. Interestingly, SNPs from chromosome 18 showed high sequence similarity between accessions EC241780 and PI 200492 near Rpp1 locus (Supplementary Table S6). Accession PI 200492 has Rpp1 gene and shows resistance to soybean rust. Phylogenetic tree was further constructed using SNP data from 81 soybean accessions. These 81 accessions include various rust resistant soybean lines which are known source of Rpp1, Rpp2, Rpp3, Rpp4, and Rpp5. All the 81 genotypes grouped into seven different clades (Fig. 2). Soybean accession EC241780 grouped with PI 200492 in clade VI along with PI 547875, PI 594177 and PI 368039, which are known to have Rpp1 rust resistance gene. These results clearly indicate that
soybean genotype EC241780 has the same Rpp1 rust resistance gene as in PI 200492. To identify the putative Rpp1 rust resistant gene in EC241780, NBS-LRR genes were identified from over 12 Mb genomic regions from position 48753972 – 61504294 on chromosome 18 and were further characterized (Fig. 3).

3.5. Comparative Analysis Identified Three Candidate NBS-LRR Disease Resistance Genes at Rpp1 Locus

To characterize the rust resistance genes at Rpp1 locus, we identified nine NBS-LRR genes (Glyma18G46875, Glyma18G51533, Glyma18G51546, Glyma18G51715, Gly-
Comparative analysis identified several SNPs and InDels in 60,910,083 which are potential candidates for chromosome 18 between genomic region 59,283,411 to ma18G51930, Gm18:58722971 and PI 587886 was mapped to the region of Gm18:60518978. Allele Xiao Jing Huang, Himeshirazu (Huang Dou (PI 587905), Zhao Ping Hei Dou (PI 594767A), Huang Dou (PI 587880A), Bai Dou (PI 587886), Xiao Jing Huang, Himeshirazu (PI 587905) and PI 587905). The other soybean lines carrying the allele includes Huahuang Dou, Bai Dou (PI 587880A), Xiao Huang Dou (PI 587905), Zhao Ping Hei Dou (PI 594767A), Xiao Jing Huang, Himeshirazu (PI 594777) and PI 561356. Allele Rpp1? in PI 561356 was mapped to the region of Gm18:60518978-60613377 [28]; and Rpp1? in PI 587880A and PI 587886 was mapped to the region of Gm18:58722971-60612672 [15]. We identified five NBS-LRR genes viz. Glyma18G51533, Glyma18G51546, Glyma18G51930, Glyma18G51950 and Glyma18G51960 on chromosome 18 between genomic region 59,283,411 to 60,910,083 which are potential candidates for Rpp1? alleles. Comparative analysis identified several SNPs and InDels in the putative NBS-LRR genes near Rpp1? locus. The largest numbers of 69 SNPs were found in the gene Glyma18G51950 in EC241780 with a total of 18 synonymous and 51 non-synonymous mutations (Supplementary Table S7). Gene Glyma18g51546 has 3 InDels with two in-frame mutation and one frame-shift mutation. Large number of SNPs and Indels in the NBS-LRR genes near Rpp1? locus caused significant changes in the NBS-LRR gene sequences in EC241780 as compared to Williams 82 (Supplementary Table S8). The functional effect for the amino acid changes located in the candidate genes predicted using PROVEAN tool identified two variations in Glyma18G51741 and five in Glyma18G51765 with deleterious effect (Table 2, Supplementary Table S9). Whereas, none of the variations in gene Glyma18G51715 showed a deleterious effect.

4. DISCUSSION

This study was initiated with a broad objective of identifying genomic variations between the rust resistant line EC241780 and susceptible cultivar JS 335, which are used as parental lines in soybean breeding programme. Whole genome sequence information of parental lines provides an opportunity to identify informative markers useful for subsequent breeding applications, to prioritise candidate genes at known loci, and most importantly to get allelic information of the candidate genes. Over a million of markers identified with whole genome re-sequencing facilitates high resolution mapping and provides tightly linked or gene specific markers. Apart from the obvious applications, whole genome sequence analysis of resistance and susceptible soybean genotypes performed in the present study helped us to understand the genomic landscape involved in the rust resistance mechanism. Earlier, similar efforts have been made in chickpea, where whole genome re-sequencing of the parental lines have been successfully used to identify candidate genes and to develop genomic resources for the breeding applications [29].
Required sequence depth for whole genome re-sequencing applications largely depends on the biological question. Another major concern is higher cost involved in the high depth sequencing. Earlier, whole genome re-sequencing effort by Zhou et al. achieved an average 11x depth in soybean genotypes [30]. Similarly, Lam et al. performed whole genome re-sequencing of 32 diverse soybean genotypes with about 5x sequence depth [31]. Another significant study by Valliyodan et al. has provided whole genome re-sequencing of 106 soybean lines at 17x depth [32]. All these studies have great importance and provide valuable resource to the soybean community. However, because of low depth the data has limitations in identifying reliable sequence variation in high copy genes and gene family members sharing high level of sequence similarity. In this study, we generated high quality re-sequencing data at 30x depth which is the gold standard as per the present genomics standard. The 30x sequence data helped to identify sequence variation in NBS-LRR gene family with more confidence. We have identified all possible polymorphisms between EC241780 and JS 335 and selected highly informative variations. Our data allowed us to develop genetic markers covering the whole genome which will be highly informative for genetic mapping of rust resistance and may be utilized in a breeding scheme to develop a rust resistant cultivar.

According to the previous studies, soybean lines PI 200492, PI 368039, PI 587886, Xiao Jing Huang, and Himeshirazu may carry the original Rpp1 as identified by Hartwig & Bromfield [33]. Bhor et al. had indicated the presence of Rpp1 gene in genotype EC241780 using SSR markers [34]. In this study, three NBS-LRR genes were identified within the Rpp1 genomic region in soybean accession EC241780. The SNPs, and InDels identified in the putative rust resistance genes will be valuable for genotyping the mapping populations originating from a combination of EC241780 and JS 335 and for marker assisted selection. Further, to facilitate genomics assisted resistance breeding efforts in soybean, chromosome-wise list of useful SNPs have been identified. Thus, we predict that our analyses will be valuable, not only for the fundamental analysis of EC241780/JS 335 derived populations, but also for enhancing our general knowledge of variation in the soybean genome. Identification of molecular markers at the genome wide scale would not only identify markers linked to currently prevalent rust races but also would result in the characterization of novel genomic variants.

The functional annotation of SNPs in NBS-LRR candidate genes showed higher non-synonymous SNPs resulting into amino acid changes. Previous reports in soybean have suggested two models of NBS-LRR gene evolution, first where the genes were tandemly duplicated and then undergone through selective pressures and then copied to different locations [35]. The second model suggest duplication of a single NBS-LRR gene first and then the tandemly duplications as suggested in the first model. Disease resistance genes are known to have a positive selection and more particularly NBS-LRR gene family [36-38]. In Arabidopsis, Mondragón-Palomino et al. have reported strong positive selection in many members of the NBS-LRR gene family [39].

SSR and SNP markers play a great role in the construction of high density linkage map in crop species. The availability of a large number of SSRs, SNPs and InDels from parental lines increases the utility of such markers in genomics assisted breeding. Hence, the development of novel genomic resources for mapping rust-resistant loci in soybean is imperative. This study offers high-density coverage of DNA polymorphisms across the entire soybean genome, which could be utilized for high-resolution genotyping and in molecular breeding programs. In this study, we obtained high quality homozygous SNPs from JS 335 and EC241780. Such genomic variation will greatly help in enhancing molecular markers repository, fine mapping investigations, and molecular genetics research for soybean rust.
Recently, Pedley et al. have identified genes governing rust resistance through extensive VIGS mediated screening [40]. This study identified Rpp1 and Rpp1b on the same chromosomal region suggesting the complexity of the locus with possible multiple resistance genes with high allelic variation. Almost a decade before Pedley et al. [40], Chakraborty et al. have shown the co-localization of Rpp1 and Rpp1b using QTL mapping approach [15]. Several similar examples where reported candidate genes does not have high confidence or miss identified due to the localization of different resistance genes have been reported in soybean [41]. A high number of alleles, the drastic difference between allelic responses, and co-localization of resistance genes makes it difficult to select the appropriate candidate. The identification of the correct candidate is indeed highly important for molecular understanding but does not limit the exploration of loci in the breeding program.

CONCLUSION

The sequence variations identified in the present study not only provide the promising candidates but also facilitate marker development for breeding application. The information generated in the present study pave the path for the exploration of Rpp1 locus for the development of rust resistance soybean cultivar.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

HUMAN AND ANIMAL RIGHTS

No animals/humans were used for studies that are the basis of this research.

CONSENT FOR PUBLICATION

Not applicable.

AVAILABILITY OF DATA AND MATERIALS

The raw sequence data supporting the conclusion of this manuscript were made available under the BioProjectPRJNA486454 at NCBI SRA database. The data can be accessed with the link: https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA486454.

FUNDING

This study was supported by the Indian Council of Agricultural Research (ICAR), Grant No: IISR 1.34/17.

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

Declared none.

SUPPLEMENTARY MATERIAL

Supplementary material is available on the publisher’s website along with the published article.

REFERENCES

[1] Kochman, J.K. Soybean rust in Australia. Rust of soybean—the problem and research needs; Ford, R.E.; Sinclair, J.B., Eds.; International Agricultural Publications: Manila, 1977, pp. 44-48.

[2] Schneider, R.W.; Hollier, C.A.; Whitem, H.K.; Paltt, M.E.; McKenny, J.M.; Hernández, J.R.; Levy, L.; DeVries-Paterson, R. First report of soybean rust caused by Phakopsora pachyrhizi in the continental United States. Plant Dis., 2005, 89(7), 774. http://dx.doi.org/10.1094/PD-89-0774A PMID: 30791253

[3] Yorinori, T.; Paiva, W.M.; Frederick, R.D.; Costamalian, L.M.; Bernagolli, P.F.; Hartyman, G.E.; Godoy, C.V.; Nunes, J., Jr. Epidemics of soybean rust (Phakopsora pachyrhizi) in Brazil and Paraguay from 2001 to 2003. Plant Dis., 2005, 89(6), 675-677. http://dx.doi.org/10.1094/PD-89-0675 PMID: 30795398

[4] Ivanovich, A. Soybean Rust in Argentina. Facing the Challenge of Soybean Rust in South America; Kudo H, Suenaga K, Soares RM, Toledo A, Eds.: Japan International Research Centre for Agricultural Science (JIRCAS): Tsukuba, Ibaraki, Japan; 2008, pp. 14-17.

[5] Burdon, J.J.; Marshall, D.R. Evaluation of Australian native species of Glycine for resistance to soybean rust. Plant Dis., 1981, 65, 44-45. http://dx.doi.org/10.1094/PD-65-44

[6] Burdon, J.J. Major gene resistance to Phakopsora pachyrhizi in Glycine canescens, a wild relative of soybean. Theor. Appl. Genet., 1988, 75, 923-928. http://dx.doi.org/10.1007/BF00258055

[7] McDowell, J.M.; Simon, S.A. Recent insights into R gene evolution. Mol. Plant Pathol., 2006, 7(5), 437-448. http://dx.doi.org/10.1111/j.1364-3703.2006.00342.x PMID: 20507459

[8] Hyten, D.L.; Smith, J.R.; Frederick, R.D.; Tucker, M.L.; Song, Q.; Cregan, P.B. Bulk segregate analysis using the Golden Gate assay to locate the Rpp3 locus that confers resistance to Phakopsora pachyrhizi (soybean rust) in soybean. Crop Sci., 2009, 49, 265-271. http://dx.doi.org/10.2135/cropsci2008.08.0511

[9] Ray, J.D.; Morel, W.; Smith, J.R.; Frederick, R.D.; Miles, M.R. Genetics and mapping of adult plant rust resistance in soybean PI 587886 and PI 587880A. Theor. Appl. Genet., 2009, 119(2), 271-280. http://dx.doi.org/10.1007/s00122-009-1036-z PMID: 19396573

[10] Kendrick, M.D.; Harris, D.K.; Ha, B.K.; Hyten, D.L.; Cregan, P.B.; Frederick, R.D.; Boerma, H.R.; Pedley, K.F.; Identification of a new Asian soybean rust resistance gene in Hyuuga soybean. Phytopathology, 2011, 101(5), 535-543. http://dx.doi.org/10.1094/PHYTO-09-10-0257 PMID: 21244223

[11] Li, S.; Smith, J.R.; Ray, J.D.; Frederick, R.D. Identification of a new soybean rust resistance gene in PI 567102B. Theor. Appl. Genet., 2012, 125(1), 133-142. http://dx.doi.org/10.1007/s00122-012-1821-y PMID: 22374138

[12] Childs, S.P.; King, Z.R.; Walker, D.R.; Harris, D.K.; Pedley, K.F.; Buck, J.W.; Boerma, H.R.; Li, Z. Discovery of a seventh Rpp soybean rust resistance locus in soybean accession PI 605823. Theor. Appl. Genet., 2018, 131(1), 27-41. http://dx.doi.org/10.1007/s00122-017-2993-4 PMID: 2890046

[13] Garcia, A.; Calvo, E.S.; de Souza Kiihl, R.A.; Harada, A.; Hiramoto, D.M.; Vicrara, I.G.E. Molecular mapping of soybean rust (Phakopsora pachyrhizi) resistance genes: discovery of a novel locus and alleles. Theor. Appl. Genet., 2008, 117(4), 545-553. http://dx.doi.org/10.1007/s00122-008-0798-2 PMID: 18506417

[14] Monteros, M.J.; Missaoui, A.M.; Phillips, D.V.; Walker, D.R.; Boerma, H.R. Mapping and confirmation of the ‘Hyuga’ red-brown lesion resistance gene for Asian soybean rust. Crop Sci., 2007, 47, 829-836. http://dx.doi.org/10.2135/cropsci06.07.0462

[15] Chakraborty, N.; Curley, J.; Frederick, R.D.; Hyten, D.L.; Nelson, R.L.; Hartman, G.L.; Diers, B.W. Mapping and confirmation of a new allele at Rpp1 from soybean PI 594538A conferring RB lesion type resistance to soybean rust. Crop Sci., 2009, 49(5), 785-790. http://dx.doi.org/10.2135/cropsci2008.06.0335

[16] Hossain, M.M.; Akamatsu, H.; Morishita, M.; Mori, T.; Yamaoka, Y.; Suenaga, K.; Soares, R.M.; Bogado, A.N.; Ivanovich, A.J.G.; Yamanaka, N. Molecular mapping of Asian soybean rust resistance in soybean landraces PI 594767A, PI 587905 and PI 416764. Plant Pathol., 2015, 64, 147-156. http://dx.doi.org/10.1111/ppa.12226
Whole Genome Re-sequencing of Soybean Accession EC241780

[Dong, W.; Wu, D.; Li, G.; Wu, D.; Wang, Z. Next-generation sequencing from bulked segregant analysis identifies a dwarfism gene in watermelon. Sci. Rep., 2018, 8(1), 2908. http://dx.doi.org/10.1038/s41598-018-21293-1 PMID: 29440685]

[Sun, Q.; Qiao, J.; Zhang, S.; He, S.; Shi, Y.; Yuan, Y.; Zhang, X.; Cai, Y. Changes in DNA methylation assessed by genomic bisulfite sequencing suggest a role for DNA methylation in cotton fruiting branch development. Peer J, 2018, 6(10), e4945. http://dx.doi.org/10.7717/peerj.4945 PMID: 29915693]

[Yadav, C.B.; Bharate, P.; Muthamilarasan, M.; Mukherjee, M.; Khan, Y.; Ratih, P.; Prasad, M. Genome-Wide SNP identification and characterization in two soybean cultivars with contrasting mungbean yellow mosaic India virus disease resistance traits. Plos One, 2015, 10(4), 0123897. http://dx.doi.org/10.1371/journal.pone.0123897]

[Ensembl Plants. http://plants.ensembl.org]

[Sourceforge.net. http://bio-bwa.sourceforge.net]

[Li, H.; Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics, 2009, 25(14), 1754-1760. http://dx.doi.org/10.1093/bioinformatics/btp324 PMID: 19451168]

[Babraham Bioinformatics. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/]

[Song, Q.; Hyten, D.L.; Jia, G.; Quigley, C.V.; Fickus, E.W.; Nelson, R.L.; Cregan, P.B. Development and evaluation of SoySNP50K, a high-density genotyping array for soybean. PLoS One, 2013, 8(1), e54985. http://dx.doi.org/10.1371/journal.pone.0054985]

[http://plants.ensembl.org]

[Song, Q.; Hyten, D.L.; Jia, G.; Quigley, C.V.; Fickus, E.W.; Nelson, R.L.; Cregan, P.B. Development and evaluation of SoySNP50K, a high-density genotyping array for soybean. PLoS One, 2013, 8(1), e54985.]

[Ding, Z.; Gao, Y.; Xiang, H.; Zhu, B.; Lee, S.H.; Wang, W.; Tian, Q.; Wu, M.; Wang, M.; Wu, Y.; Dong, Y.; Wang, X.; Ding, Z.; Gao, Y.; Xiang, H.; Zhu, B.; Lee, S.H.; Wang, W.; Tian, Q. Resequencing 302 wild and cultivated soybean genomes identifies patterns of genetic diversity and selection. Nat. Genet., 2010, 42(12), 1053-1059. http://dx.doi.org/10.1038/ng.175 PMID: 21076406]

[Valiyodan, B.; Dan Qiu, ; Patil, G.; Zeng, P.; Huang, J.; Dai, L.; Chen, C.; Li, Y.; Joshi, T.; Song, L.; Vuong, T.D.; Musket, T.A.; Xu, D.; Shannon, J.G.; Shifeng, C.; Liu, X.; Nguyen, H.T. Landscapes of genomic diversity and trait discovery in soybean. Sci. Rep., 2016, 6, 23598. http://dx.doi.org/10.1038/srep23598 PMID: 27029319]

[Hartwig, E.E.; Bromfield, K.R. Relationship among three genes conferring specific resistance to rust in soybeans. Crop Sci., 1983, 23, 237-239. http://dx.doi.org/10.2135/cropsci1983.0011183X002300020012x]

[Bhor, T.J.; Chimote, V.P.; Deshmukh, M.P. Molecular tagging of asiacitic soybean rust resistance in exotic genotype EC241780 reveals complementation of two genes. Plant Breed., 2015, 134(1), 70-77. http://dx.doi.org/10.1111/phb.12240]

[Kang, Y.J.; Kim, K.H.; Shim, S.; Yoon, M.Y.; Sun, S.; Kim, M.Y.; Van, K.; Lee, S.H. Genome-wide mapping of NBS-LRR genes and their association with disease resistance in soybean. BMC Plant Biol., 2012, 12(1), 139. http://dx.doi.org/10.1186/1471-2229-12-139 PMID: 22877146]

[Innes, R.W.; Ameline-Torregrosa, C.; Ashfield, T.; Cannon, E.; Cannon, S.B.; Chacko, B.; Chen, N.W.; Couloux, A.; Dalwani, A.; Denny, R.; Deshpande, S.; Egan, A.N.; Glover, N.; Hans, C.S.; Howell, S.; Iltut, D.; Jackson, S.; Lai, H.; Mammodov, J.; Del Campo, S.M.; Metcalf, M.; Nguyen, A.; O’Brien, M.; Pfeil, B.E.; Pedichi, R.; Ratnaparkhe, M.B.; Samain, S.; Sanders, I.; Ségurens, B.; Sévignac, M.; Sherman-Broyles, S.; Thareau, V.; Tucker, D.M.; Walling, J.; Wawrzynski, A.; Yi, J.; Doyle, J.J.; Geffroy, V.; Roe, B.A.; Maroof, M.A.; Young, N.D. Differential accumulation of retroelements and diversification of NB-LRR disease resistance genes in duplicated regions following polyploidy in the ancestor of soybean. Plant Physiol., 2008, 148(4), 1740-1759. http://dx.doi.org/10.1104/pp.108.127902 PMID: 18842825]

[Ratnaparkhe, M.B.; Wang, X.; Li, J.; Compton, R.A.; Rainville, L.K.; Lemke, C.; Kim, C.; Tang, H.; Paterson, A.H. Comparative analysis of peanut NBS-LRR gene clusters suggests evolutionary innovation among duplicated domains and erosion of gene microsynteny. New Phytol., 2011, 192(1), 164-178. http://dx.doi.org/10.1111/j.1469-8137.2010.03800.x PMID: 21707619]

[Van de Weyer, A.L.; Monteiro, F.; Furzer, O.J.; Nishimura, M.T.; Cevik, V.; Wiek, K.; Jones, J.D.G.; Dangl, J.L.; Weigel, D.; Bemm, F. A species-wide inventory of NLR genes and alleles in Arabidopsis thaliana. Cell, 2019, 178(5), 1260-1272. http://dx.doi.org/10.1016/j.cell.2019.07.038 PMID: 31442410]

[Mondragon-Palomino, M.; Meyers, B.C.; Michelmore, R.W.; Gau, B.S. Patterns of positive selection in the complete NBS-LRR gene family of Arabidopsis thaliana. Genome Res., 2002, 12(9), 1305-1315. http://dx.doi.org/10.1101/gr.159402 PMID: 12213767]

[Pedley, K.F.; Pandey, A.K.; Ruck, A.; Lincoln, L.M.; Whitham, S.A.; Graham, M.A. Rppl encodes a ULPI-NBS-LRR protein that controls immunity to Phakopsora pachyrhizi in soybean. Mol. Plant Microbe Interact., 2019, 32(1), 120-133. http://dx.doi.org/10.1094/MPMI-07-18-0198-FI PMID: 30307365]

[Vuong, T.D.; Walker, D.R.; Nguyen, B.T.; Nguyen, T.T.; Dinh, H.X.; Hyten, D.L.; Cregan, P.B.; Sleper, D.A.; Lee, J.D.; Shannon, J.G.; Nguyen, H.T. Molecular characterization of resistance to soybean rust (Phakopsora pachyrhizi Syd.) in soybean cultivar DT 2000 (PI 635999). Plos One, 2016, 11(12), e0164493.]

[http://dx.doi.org/10.1111/phy.12240

http://dx.doi.org/10.1104/pp.108.127902

http://dx.doi.org/10.1111/j.1469-8137.2010.03800.x

http://dx.doi.org/10.1094/MPMI-07-18-0198-FI

http://www.bioinformatics.babraham.ac.uk/projects/fastqc/]

[http://www.bioinformatics.babraham.ac.uk/projects/fastqc/]

[http://dx.doi.org/10.1093/bioinformatics/btp324 PMID: 19451168]

[http://dx.doi.org/10.1093/bioinformatics/btp324 PMID: 19451168]