 doi:10.20944/preprints202010.0501.v1

Whole-genome sequencing and phenotyping revealed structural variants and varied level of resistance against leaf crumple disease in diverse lines of snap bean (*Phaseolus vulgaris*)

Gaurav Agarwal[1], Saritha Raman Kavalappara[1], Saurabh Gautam[2], Andre da Silva[3], Alvin Simmons[4], Rajagopalbabu Srinivasan[2], Bhabesh Dutta[1*]

[1] Department of Plant Pathology, Coastal Plain Experimental Station, University of Georgia, Tifton, Georgia, United States of America

[2] Department of Entomology, 1109 Experiment Station, Griffin, University of Georgia, Griffin, Georgia, United States of America

[3] Department of Horticulture, Coastal Plain Experimental Station, University of Georgia, Tifton, Georgia, United States of America

[4]U.S. Vegetable Laboratory, Agricultural Research Service, United States Department of Agriculture, Charleston, SC

[*] Corresponding author: bhabesh@uga.edu

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 October 2020 doi:10.20944/preprints202010.0501.v1

Abstract

The production and quality of *Phaseolus vulgaris* (snap bean) have been negatively impacted by leaf crumple disease caused by two whitefly-transmitted begomoviruses; cucurbit leaf crumple virus (CuLCrV) and sida golden mosaic Florida virus (SiGMFV), which often appear as a mixed infection in Georgia. Host resistance is the most economical management strategy against whitefly-transmitted viruses. Currently, information is not available with respect to resistance to these two viruses in commercial cultivars. In two field seasons (2018 and 2019), we screened *Phaseolus* sp. genotypes (n = 84 in 2018; n = 80 in 2019; most of the genotypes were common in both years with a few exceptions) for resistance against CuLCrV and/or SiGMFV. Twenty genotypes with high- to moderate-levels of resistance (disease severity ranging from 5-50%) to CuLCrV and/or SiGMFV were identified. Twenty-one genotypes were found to be highly susceptible with a disease severity of ≥66%. Adult whitefly counts differed significantly among snap bean genotypes for both years. The whole genome of these *Phaseolus* sp. (n=82) genotypes was sequenced and genetic variability among them was identified. Over 900 giga-base (Gb) of filtered data were generated and >88% of the resulting data were mapped to the reference genome, and SNP and Indel variants in *Phaseolus* genotypes were obtained. A total of 645,729 SNPs and 68,713 Indels including 30,169 insertions and 38,543 deletions were identified, which were distributed in 11 chromosomes with chromosome 02 harboring the maximum number of variants. These phenotypic and genotypic information will be helpful in genome-wide association studies that will aid in identifying genetic basis of resistance to these begomoviruses in *Phaseolus* sp.

Keywords: Cucurbit leaf crumple virus, Sida golden mosaic Florida virus, Whitefly, Snap beans, Lima beans

1. Introduction

Among commonly grown *Phaseolus* sp., *Phaseolus vulgaris* L. (common bean, snap bean) is an annual legume crop with a diploid genome size of 521.1 Mb ($2n$ = 22) [1]. Snap bean is one of the most important affordable food legumes for humans [2], which is consumed by over 80 million poor people in regions of Latin America, the Caribbean, and Eastern and Southern Africa. In the U.S., snap bean is an important horticultural crop, especially for the state of Georgia where snap bean is grown in 9,979 acres and generates an annual revenue of $24 million dollars [3]. However, the production and quality of snap bean have been negatively impacted by two whitefly-transmitted begomoviruses namely, cucurbit leaf crumple virus (CuLCrV) and sida golden mosaic Florida virus (SiGMFV), which often appear as a mixed infection in Georgia.

CuLCrV is a bipartite begomovirus first identified in watermelon in the Imperial Valley of southern California in 1998 [4] and in Georgia in snap beans in 2009 [5]. In August 2018, snap beans with characteristic begomovirus infection symptoms (crumpled, curled, and thickened leaves) were found in Tifton, Georgia, and these plants were heavily infested with whiteflies. Subsequent analysis with degenerate and specific begomovirus primers revealed the presence of SiGMFV in infected plant tissues. In the Southeastern United States, SiGMFV (a bipartite virus) was first reported in Florida in 2006 on snap beans with infected plants displaying leaf mottling,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 October 2020 doi:10.20944/preprints202010.0501.v1

puckering, and severe curling symptoms [6]. Both the viruses are transmitted by the sweet-potato whitefly (*Bemisia tabaci* Gennadius). Currently, leaf crumple disease management is centered around vector control, which usually occur via insecticides. Disease management via vector control is unreliable and insufficient. On the contrary, host-resistance is the more economical and sustainable approach that can potentially minimize fields infestation, but there is considerable lack of information on host resistance against these two begomoviruses on snap bean in the United States. It is imperative to understand the genetics of host resistance, which involves identifying markers and genes for breeding highly resistant snap bean varieties.

Single nucleotide polymorphisms (SNPs) have high frequency of occurrence throughout the genome and are considered as preferred genetic marker in breeding for disease resistance. SNPs along with longer sequence variants; insertions and deletions (InDels), aided in discovery of quantitative trait loci (QTLs) and genes associated with disease resistance and agronomic traits in many cultivated crops. Prior genomic studies on *P. vulgaris* (dry beans) focused on agronomic and abiotic stress related traits (drought stress, salt tress) but none of them focused on identifying resistance to viral pathogens. Biparental QTL mapping and genome wide association studies (GWAS) have been used to discover such traits in common bean [7-9]. In the current study, beside evaluating the response of snap beans to natural infection of begomoviruses under field conditions for two consecutive years (seasons), sequence variants (SNPs and Indels), and their distribution in the *Phaseolus* cultivars were also identified using whole-genome sequencing (WGS).

2. Materials and Methods

2.1. Plant Materials

Eighty-four *Phaseolus* genotypes including 82 snap bean and two Lima bean (*P. lunatus*) genotypes were used in 2018. Two Lima bean genotypes that are close relatives of snap beans, Jackson wonder and Fordhook were also included. Eighty genotypes were tested in 2019 out of which seventy-six genotypes were the same as those tested in 2018. Seeds of BMN- RMR- 13, Bronco 2, Lakatte, SB4734, SB4735, SB4744, SB4679, SV1137 were not available for evaluation in 2019. Hence, four genotypes of snap beans, Achiever, Blue Lake 274, Coyote and Greenback were evaluated only in 2019 (**Table 1**). Seeds were collected from commercial seed companies and Germplasm Resource Information Network (GRIN) of the United States Department of Agriculture (USDA) (**Table 1**).

2.2. Experimental design, layout and environmental conditions

Genotypes mentioned above were evaluated for resistance to CuLCrV and SiGMFV under field conditions at the University of Georgia, Tifton. In both years (2018 and 2019), seeds were grown in 12 individual 138 m long-raised beds. Each raised bed was divided into plots with dimension of 3.04 m × 0.91 m. Each plot was comprised 20 plants planted in a in-row spacing of 7.62 cm, double rows spaced at 46 cm were used within each bed. Treatments (genotypes) were replicated (r = 3) using a randomized complete block design. Natural whitefly infestation was relied upon virus transmission and resultant disease. All cultural practices and disease management followed the UGA Cooperative Extension recommendations [10]. Insecticides were not sprayed to ensure survival of whiteflies for disease incidence and spread. Averages of maximum and minimum temperatures in 2018 during the growing period were 34.5 °C and 21.1 °C, respectively with an

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 October 2020 doi:10.20944/preprints202010.0501.v1

accumulated precipitation of 0.25 cm. In the 2019 growing period, averages of maximum and minimum temperatures were 32.5 °C and 22.5 °C and an accumulated precipitation was 0.23 cm.

2.3. Response of *Phaseolus* sp. (snap beans and Lima beans) genotypes to leaf crumple disease in the field

In 2018, evaluation of genotypes for virus resistance was conducted at 30 days after sowing (DAS). Since Hurricane Michael destroyed the crop in early October, 2018, a second evaluation of resistance was not possible. In 2019, leaf crumple disease evaluation was conducted twice, at 30 and 45 DAS. For each genotype, plants were evaluated visually for disease incidence and severity. Disease severity in five randomly selected plants per plot per genotype was evaluated using a severity scale of 0 to 100. A plant with no crumpling, mosaic and stunting was scored as 0 (**Fig 1A)**. A plant with severe leaf crumpling, mosaic and stunting was scored as 100 (**Fig 1B)**. Genotypes with disease severity ≤ 20% were rated as highly resistant, 21-50% as moderately resistant, 51-65% as susceptible and ≥65% as highly susceptible.

2.4. Whitefly count

Adult whiteflies were counted at 30 DAS in 2018 and 45 DAS in 2019 on each genotype. Counting was conducted on the lower side of leaves in the morning hours when whiteflies are not very active. Whiteflies adults were enumerated on top three, fully expanded leaves by gently turning the leaf over by tip. Whitefly counts were taken from 15 plants for each genotype, five from each replicate. Whitefly count data for 2018 and 2019 were analyzed using linear mixed techniques in software R version 3.4.2. Snap bean genotypes were considered as fixed effects and replicates were considered as random effects. To meet the assumption of ANOVA (normality and homoscedasticity of variance) prior to analysis data were log(X+1) transformed. Post-hoc analyses were performed using the 'emmeans' package with the default Tukey's honest significant difference test (p=0.05).

2.5. DNA isolation, library preparation, sequencing and quality filtering of raw data

A total of 82 *Phaseolus* genotypes (80 snap beans and two Lima beans) were sequenced. Total DNA was isolated from a single plant of each genotype collected arbitrarily from the field using DNEasy plant mini kit (Qiagen) following manufacturer's instructions. A 50ng/ µl of DNA per sample was used for library preparation as per the standard protocol**.** Genomic DNA of each sample was randomly sheared into short fragments of about 300-500 bp. The obtained fragments were subjected to library construction using the NEBNext® DNA Library Prep Kit following manufacturer's instructions. Library was subsequently sequenced and the raw FASTQ reads obtained were quality filtered. We discarded the paired reads: when either read contained adapter contamination; when uncertain nucleotides (N) constitute more than 10 percent of either read; when low quality nucleotides (base quality less than 5, Q ≤ 5) constitute more than 50 percent of either read.

2.6. Mapping of filtered read data on the reference genome and variant calling

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 October 2020 doi:10.20944/preprints202010.0501.v1

The filtered sequencing data was aligned on *Phaseolus vulgaris* reference genome (Pvulgaris_442_v2.0_softmasked) available at legume information system (LIS). BWA software (parameters: mem -t 4 -k 32 -M) was used for alignment and the mapping rate and coverage were counted according to the alignment results. The duplicates were removed by SAMtools. Individual SNP variations were detected using GATK. SNPs and InDels were further filtered based quality and depth. All variants with Qual <30, SOR > 3.0, DP <6, heterozygous and multi-allelic calls were filtered out. Filtered SNPs and InDels were annotated using Annovar [11]. SNP and InDel densities per kb were calculated in 100 kb bins all throughout the 11 chromosomes of *P. vulgaris*.

2.7. Confirmation of begomoviruses (CuLCrV and SiGMFV) infection associated with leaf crumple symptoms in *Phaseolus* sp.

In order to ensure if the symptoms observed were associated with begomoviruses, we tested *Phaseolus* sp. leaf samples from forty randomly collected genotypes from the field. A real time PCR (qRTPCR) using primers and protocol developed by Gautam et al. [12] was used. DNA samples tested positive for CulCrV and SiGMFV earlier were included as positive controls. Water was added in place of DNA in negative controls.

3 Results

3.1. Response of *Phaseolus* sp. (snap beans and Lima beans) genotypes to leaf crumple disease in the field

In both years, typical symptoms of virus infection included yellow mosaic, leaf crumpling, and shortening in varying degrees in different genotypes (**Figure 1, Supplementary figure S1**). In 2018, each plant in the field was examined for visual symptoms and 100% of the genotypes had at least one symptomatic plant per plot. In 2019, data from only five plants were recorded individually for disease incidence. One hundred percent of the plants visually screened for each genotype had leaf crumple incidence; however, disease severity among genotypes varied considerably (**Figure 1, Supplementary figure S1**). None of the genotypes were symptomless or immune in both years tested.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 October 2020 doi:10.20944/preprints202010.0501.v1

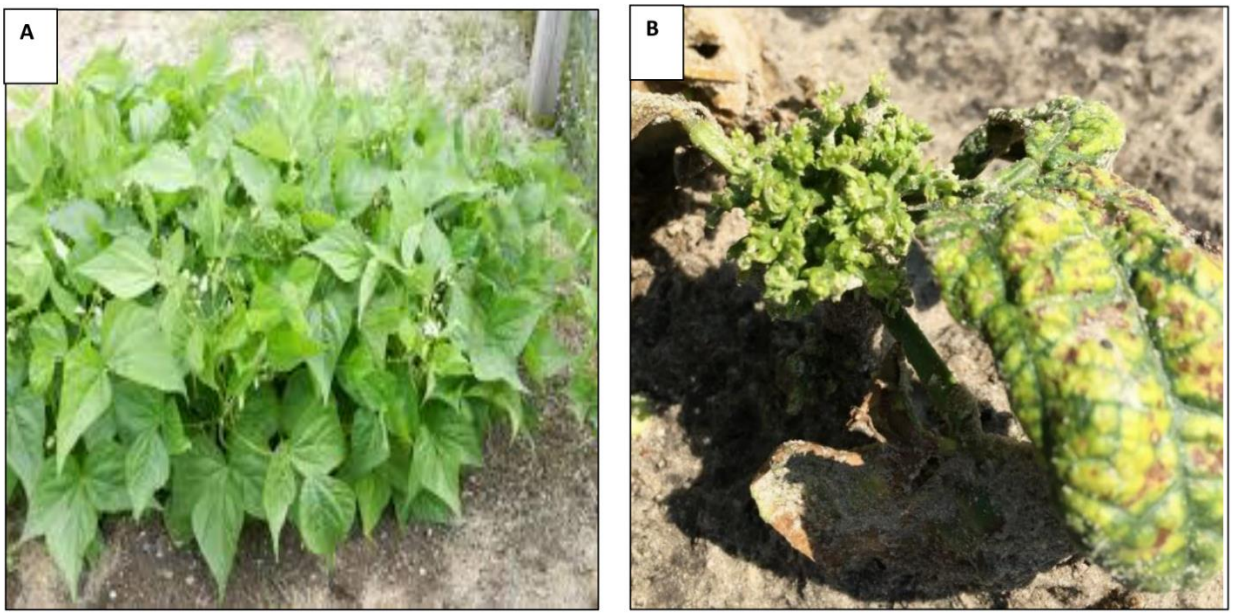


Figure 1. Healthy snap bean plants (A) and snap bean plant severely affected by leaf crumple disease with severe yellow mosaic and clumping of stunted leaves at the top (B).

In 2018, out of 84 genotypes, 19 genotypes showed a high level of resistance to leaf crumple, 25 genotypes were moderately resistant, 11 genotypes were susceptible, and 29 genotypes were found to be highly susceptible (**Table 1)**.

In general, disease severities were higher in most of the genotypes in 2019 compared to 2018. Many genotypes that were resistant in 2018 were susceptible in 2019. At 30 DAS, disease severity was higher in most genotypes compared to 45 DAS. Sixteen snap bean genotypes classified as highly resistant in 2018 showed higher disease severities in 2019. In 2019, twenty-four snap bean genotypes were moderately resistant, 24 were susceptible and 31 genotypes were highly susceptible (**Table 1**).

The two Lima bean genotypes (*P. lunatus*), Jackson Wonder and Fordhook had low disease severity in 2018 and 2019 (**Table 1**). Eight *Phaseolus* genotypes were moderately resistant in 2018 and 2019 with disease severities ranging from 21% to 50% for Barron, Carson, Cedric Larson, Fordhook, Furano, Hastings white cornfield, Hmx 175724 and Wyatt.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 October 2020 doi:10.20944/preprints202010.0501.v1

Table 1. Response of *Phaseolus* genotypes to leaf crumple disease in the field during 2018 and 2019 cropping seasons.

	Genotype	Source	Disease Severity (%)					
S. No.			2018 30 DAS	HR[a]/MR[b] /S[c]/HS[d]	2019 30 DAS[e]	2019 45 DAS	2019 Mean (%) ± SE	HR/MR /S/HS
1	Abunda*	GRIN	77[f]	HS	100	100	100±0	HS
2	Affirmed*	Seminis	17	HR	53	57	55±2	S
3	Amethyst*	Johhny's seed	20	HR	33	33	33±0	MR
4	Apollo*	GRIN	33	MR	60	60	60±0	S
5	BA0958*	Jenna	53	S	70	74	72±2	HS
6	BA1006*	Jenna	77	HS	65	65	65±0	S
7	Barron*	Harris Moran	30	MR	73	27	50±23	MR
8	Belmidak-Rust Resistant-1*	GRIN	73	HS	60	60	60±0	S
9	Belmidak-Rust Resistant-2*	GRIN	67	HS	40	37	38.5±1.5	MR
10	BLUSH*	GRIN	17	HR	47	50	48.5±1.5	MR
11	BMN- RMR- 13*	GRIN	52	S	NS[g]	NS	_	_
12	BMN-RMR-10*	GRIN	53	S	63	67	65±2	S
13	BMN-RMR-11*	GRIN	73	HS	70	77	73.5±3.5	HS
14	BMN-RMR-12*	GRIN	80	HS	83	86	84.5±1.5	HS
15	BMN-RMR-8*	GRIN	50	MR	72	80	76±4	HS
16	BMN-RMR-9*	GRIN	60	S	70	70	70±0	HS
17	Bronco 1*	Seminis	80	HS	53	57	55±2	S
18	Bronco 2*	Seminis	90	HS	NS	NS	_	_
19	Bush Blue Lake 283*	Asgrow Seed Co	80	HS	100	100	100±0	HS
20	Capitole Snap*	GRIN	80	HS	70	73	71.5±1.5	HS
21	Caprice*	Harris Moran	87	HS	100	100	100±0	HS
22	Carson*	Syngenta	22	MR	43	47	45±2	MR
23	Cascade*	GRIN	42	MR	57	59	58±1	S
24	Cedric Larson*	GRIN	27	MR	37	43	40±3	MR

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 October 2020 doi:10.20944/preprints202010.0501.v1

25	Champagne*	GRIN	55	S	47	59	53±6	S
26	Coloma*	GRIN	97	HS	93	100	96.5±3.5	HS
27	Colter*	Harris Moran	27	MR	57	60	58.5±1.5	S
28	Cosmos*	Johnny's seed	60	S	53	63	58±5	S
29	Desoto*	Harris Moran	20	HR	40	50	45±5	MR
30	Desperado*	Burpee	20	HR	47	47	47±0	MR
31	Early Harvest*	GRIN	80	HS	60	100	80±20	HS
32	Executive Bush Snap*	GRIN	87	HS	100	100	100±0	HS
33	E-Z pick*	Johhny's seed	82	HS	97	99	98±1	HS
34	Fordhook*¶	Seedway	23	MR	25	18	21.5±3.5	MR
35	Furano*	Syngenta	22	MR	32	40	36±4	MR
36	Gardengreen*	GRIN	33	MR	67	83	75±8	HS
37	Gold Mine*	Seminis	87	HS	100	100	100±0	HS
38	Goldcoast*	GRIN	67	HS	100	100	100±0	HS
39	Goldcrop*	GRIN	17	HR	47	47	47±0	MR
40	Greencrop*	Seedway	80	HS	88	67	77.5±10.5	HS
41	Hastings White Cornfield*	GRIN	35	MR	45	45	45±0	MR
42	Hmx175724*	Harris Moran	27	MR	50	47	48.5±1.5	MR
43	Hmx5106*	Harris Moran	12	HR	47	46	46.5±0.5	MR
44	Horticultural*	Seedway	93	HS	81	90	85.5±4.5	HS
45	Jackson Wonder*¶	GRIN	5	HR	23	12	17.5±5.5	HR
46	Jade II*	Harris Moran	40	MR	57	57	57±0	S
47	Kentucky Blue*	Sieger	50	MR	70	72	71±1	HS
48	Kentucky Wonder*	Seedway	35	MR	67	68	67.5±0.5	HS
49	King Horticultural*	GRIN	40	MR	60	65	62.5±2.5	S
50	Lakatte*	GRIN	93	HS	NS	NS	_	_
51	Lasalle*	Harris Moran	80	HS	95	100	97.5±2.5	HS
52	London Horticultural*	GRIN	33	MR	50	57	53.5±3.5	S
53	Longval*	GRIN	87	HS	77	95	86±9	HS
54	Lows Champion*	GRIN	17	HR	67	43	55±12	S
55	Maxibel*	Johhny's seed	57	S	53	57	55±2	S

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 October 2020 doi:10.20944/preprints202010.0501.v1

56	Missouri Wonder*	GRIN	57	S	60	80	70±10	HS
57	Momentum*	Syngenta	20	HR	37	40	38.5±1.5	MR
58	Morses Pole No 191*	GRIN	53	S	53	53	53±0	S
59	Outlaw*	Stokes seeds	73	HS	47	63	55±8	S
60	Polaris*	GRIN	40	MR	66	70	68±2	HS
61	Prevail*	Syngenta	13	HR	45	47	46±1	MR
62	Provider*	Seedway	93	HS	95	97	96±1	HS
63	PV-857*	Seedway	20	HR	35	37	36±1	MR
64	PV-905*	PopVriend	27	MR	53	53	53±0	S
65	Roma II*	Seedway	93	HS	80	100	90±10	HS
66	Roundup*	GRIN	80	HS	80	87	83.5±3.5	HS
67	Royal Burgundy*	Johhny's seed	23	MR	60	43	51.5±8.5	S
68	SB4679*	GRIN	17	HR	NS	NS	_	_
69	SB4734*	GRIN	20	HR	NS	NS	_	_
70	SB4735*	GRIN	50	MR	NS	NS	_	_
71	SB4744*	GRIN	37	MR	NS	NS	_	_
72	Spartan Half Runner*	GRIN	53	S	40	40	40±0	MR
73	Striped Half Runner*	GRIN	33	MR	79	39	59±20	S
74	SV1003GF*	Stokes seed	20	HR	70	40	55±15	S
75	SV1137*	GRIN	63	S	NS	NS	_	_
76	Sybaris*	Seminis	13	HR	35	37	36±1	MR
77	Tavera*	Johhny seed	53	S	43	53	48±5	MR
78	Tema*	Semins	5	HR	47	50	48.5±1.5	MR
79	Topcrop*	Seedway	100	HS	77	83	80±3	HS
80	Valentino*	Stokes seed	17	HR	66	45	55.5±10.5	S
81	Wyatt*	Harris Moran	37	MR	37	40	38.5±1.5	MR
82	Yakima*	GRIN	20	HR	53	57	55±2	S
83	Achiever	Dave's garden	NS	_	53	57	55±2	S
84	Bluelake 274	Ferry Morse	NS	_	75	75	75±0	HS
85	Coyote	Syngenta	NS	_	45	47	46±1	MR
86	Golden Rod	Seminis	77	HS	100	100	100±0	HS

87	Greenback	Seedway	NS	_	40	40	40±0	MR
88	K Bush Bean	GRIN	83	HS	97	100	98.5±1.5	HS

[¶]Lima beans (*Phaseolus lunatus*)
[*]Genotypes sequenced
[a] Highly resistant
[b] Moderately resistant
[c] Susceptible
[d] Highly susceptible
[e]DAS: days after sowing
[f] Mean disease severity from 15 plants, five each from three replicated plots.
[g]NS: not evaluated

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 October 2020 doi:10.20944/preprints202010.0501.v1

3.2. Confirmation of begomoviruses (CuLCrV and/or SiGMFV) infection in *Phaseolus* sp.

In 2018 and 2019 both the viruses were detected in the field and were widely prevalent. In 2018 out of the 38 samples tested, at least one virus was detected in 11 samples and both viruses were detected in 27 samples. The CuLCrV was detected in 37 samples while SiGMFV was detected in 28 samples. In 2019, out of 40 samples tested, at least one virus was detected in 21 samples while both viruses were detected in 17 samples. The CuLCrV was detected in 18 samples while SiGMFV was detected in 36 samples.

3.3. Whitefly count

Overall, the whitefly counts were lower in 2018 compared to 2019. In 2018, whitefly counts differ significantly between snap bean genotypes ($F_{(86,1164)}$=7.12, p<0.001), the lowest whitefly count was recorded on the Gold Mine (Fig. 2A). Similarly, in 2019, whitefly counts differ significantly between snap bean genotypes ($F_{(76,1078)}$=9.13,p<0.001), but the lowest whitefly adult count was recorded on the Executive bush bean (Fig. 2B).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 October 2020 doi:10.20944/preprints202010.0501.v1

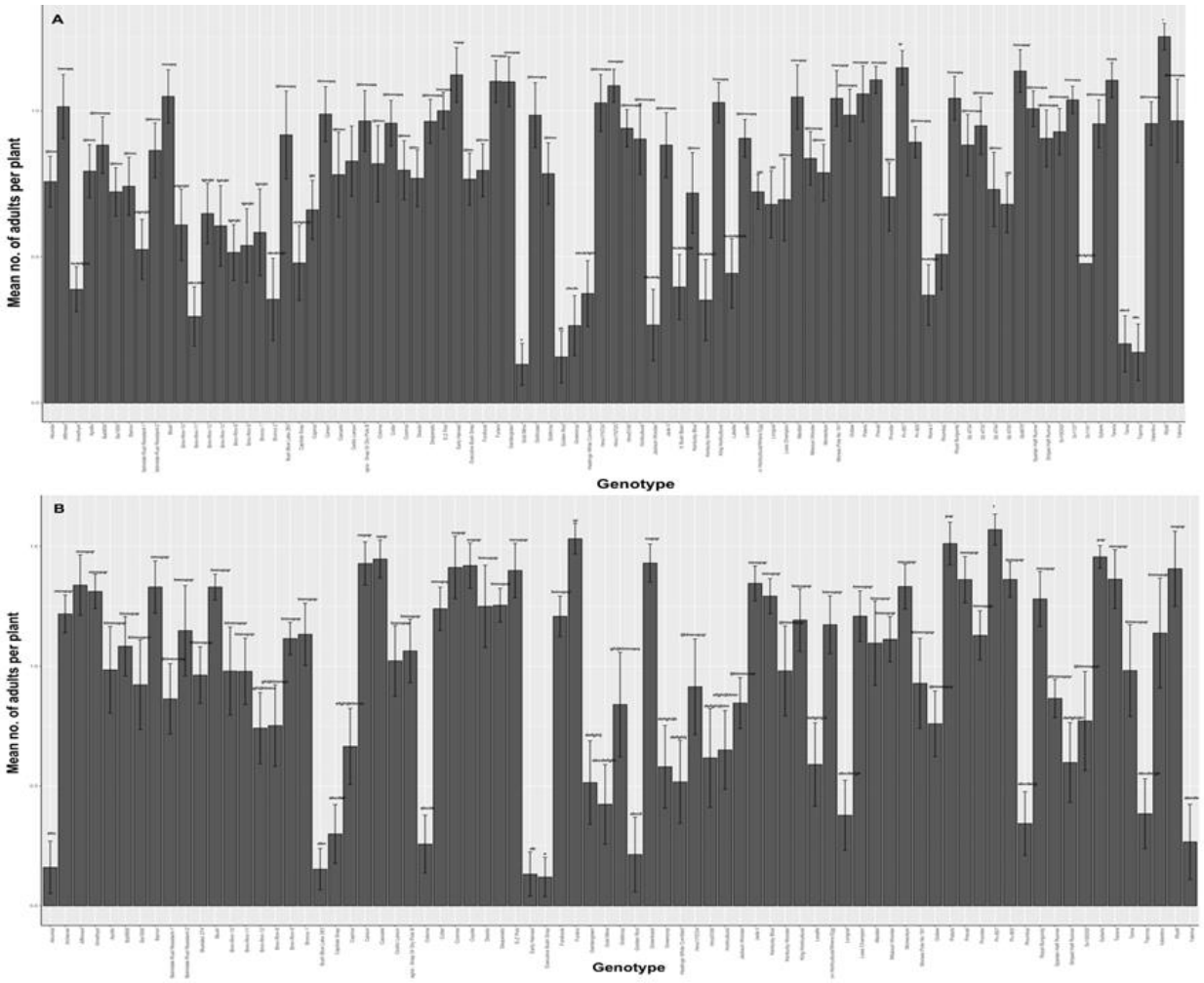


Figure 2. Mean number of whitefly adults counted on different snap bean genotypes in 2018 (A) and 2019 (B). Bars with standard errors represent average number of whiteflies present on each plant of the genotype. Whiteflies adults were enumerated on three top leaves of the plant. Y-axis is shown in a logarithmic scale. Bars with different letters are significantly different from one another ($p < 0.05$).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 October 2020 doi:10.20944/preprints202010.0501.v1

3.4. Data filtering, mapping and variants identification

A total of over six billion raw-read data were generated. Per sample the raw data generated ranged from a minimum of 29.8 million to a maximum of 73.1 million paired-end reads (**Supplementary Table S1, Supplementary figure S2**). The raw read data were quality filtered. Overall, more than 97% data (903.6 Gb) were retained (**Table 2, Supplementary figure S3**). The filtered read data were mapped on to the reference genome of *P. vulgaris*. A total of over five billion reads were mapped on the reference genome amounting to 88.6 % mapping rate. Seventy-one *Phaseolus* genotypes showed that more than 75% of filtered reads were successfully mapped (75% - 97%); however, 11 genotypes showed less than 75% mapping (**Supplementary figure S3**). Details of total number of filtered reads mapped for each sample is listed in the supplementary table S2. Average depth (X) of mapped reads at each site ranged from 11.8 to 22.3 calculated based on the total number of bases in the mapped reads divided by size of the assembled genome. Percentage of genome coverage with more than one read mapped (at least 1X) ranged from 46.5% to 96.9%. Percentage of genome coverage with 4X ranged from 39.5% to 94.98%. Overall, Fordhook (55.8%) and Jackson wonder (60. 8%) displayed low mapping because of low genome coverage of 39.5% and 39.19% respectively (**Supplementary table S2**). A total of 2,10,42,255 raw SNPs and 41,56,878 raw InDels were identified on eleven chromosomes and 467 scaffolds of *P. vulgaris*. Out of these variants initially identified, 2,01,17,468 biallelic SNPs and 37,32,869 biallelic InDels were retained further. After removing the non-variant sites, 1,15,72,528 SNPs and 24,65,936 SNPs were retained. Further, applying missing variant site, minor allele frequency (0.10) filters and excluding the variants present on scaffolds, a total of 6,45,729 SNPs and 68,713 InDels were identified on eleven chromosomes of *P. vulgaris*.

Table 2. Overview of raw data generated, data retained after quality filtering and reads mapped on the reference genome.

Total raw reads	Filtered clean reads	Filtered data (Gb)	Total reads mapped	Av reads mapped (%)
6,033,783,354	6,026,076,892	903.6	5,204,929,327	88.59

3.5. Analysis and annotation of SNPs and InDels

SNP and Indel densities varied among the chromosomes. The maximum SNP density of 6.39 SNPs/kb was identified on chromosome 1 (19.2 Mb to 19.3 Mb bin) followed by chromosomes 7 (5.7 Mb to 5.8 Mb bin) and 5 (2 Mb to 2.1 Mb bin) with SNP densities of 5.91/kb and 5.28/kb respectively (**Supplementary tables S3-S5**). Maximum insertion and deletion densities of 0.37/kb and 0.54/kb were identified in the same region of chromosome 5 (3 Mb to 3.1 Mb bin). Overall, five 100 kb bins on chromosomes 1, 2, 5, 6 and 11 were found with density >500 SNPs/100 kb, seven such bins with insertions > 30/100 kb were identified on chromosomes 2, 3, 5 (two bins), 9 (two bins) and 10. Similarly, six bins with deletions > 40/100 kb on chromosomes 2, 4, 5 (two regions), 7 and 11 were identified (**Figure 3, Supplementary tables S3-S5**). A maximum of

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 October 2020 doi:10.20944/preprints202010.0501.v1

73,326 and a minimum of 46,028 SNPs were identified on chromosomes 02 and 07, respectively (**Table 3, Figure 4A**). Length of a chromosome and number of variants are generally directly correlated i.e. the longest chromosome is expected to possess greatest number of SNPs and InDels. However, in the current study we identified the maximum number of SNPs (73,326) on chromosome 02 (the fifth longest chromosome with length 49.67 Mb). On the contrary, Chromosome 08 (the largest with length 63.05 Mb) contained second largest number of SNPs (69,823). Similarly, the greatest number of insertions (3,522) and deletions (4,429) were found on chromosome 02. Minimum number of Indels were identified on chromosome 10 (**Table 3**, **Figure 4B, C**). Length of insertions ranged from 1 to 181 bp and that of deletions ranged from -1 to -109 bp. The number of deletions (38,550) were higher than the number of insertions (30,165) with a maximum frequency of 1 bp insertions and deletions (**Table 3**, **Supplementary tables S6 and S7**). Investigation on the nucleotide substitution type of SNPs indicated higher frequency of transitions (C/T and G/A; Ts = 407,325) than transversions (C/A, G/T, C/G and T/A; Tv = 238,404) and the ratio of Ts/Tv was 1.71 (**Table 4, Figure 5**).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 October 2020 doi:10.20944/preprints202010.0501.v1

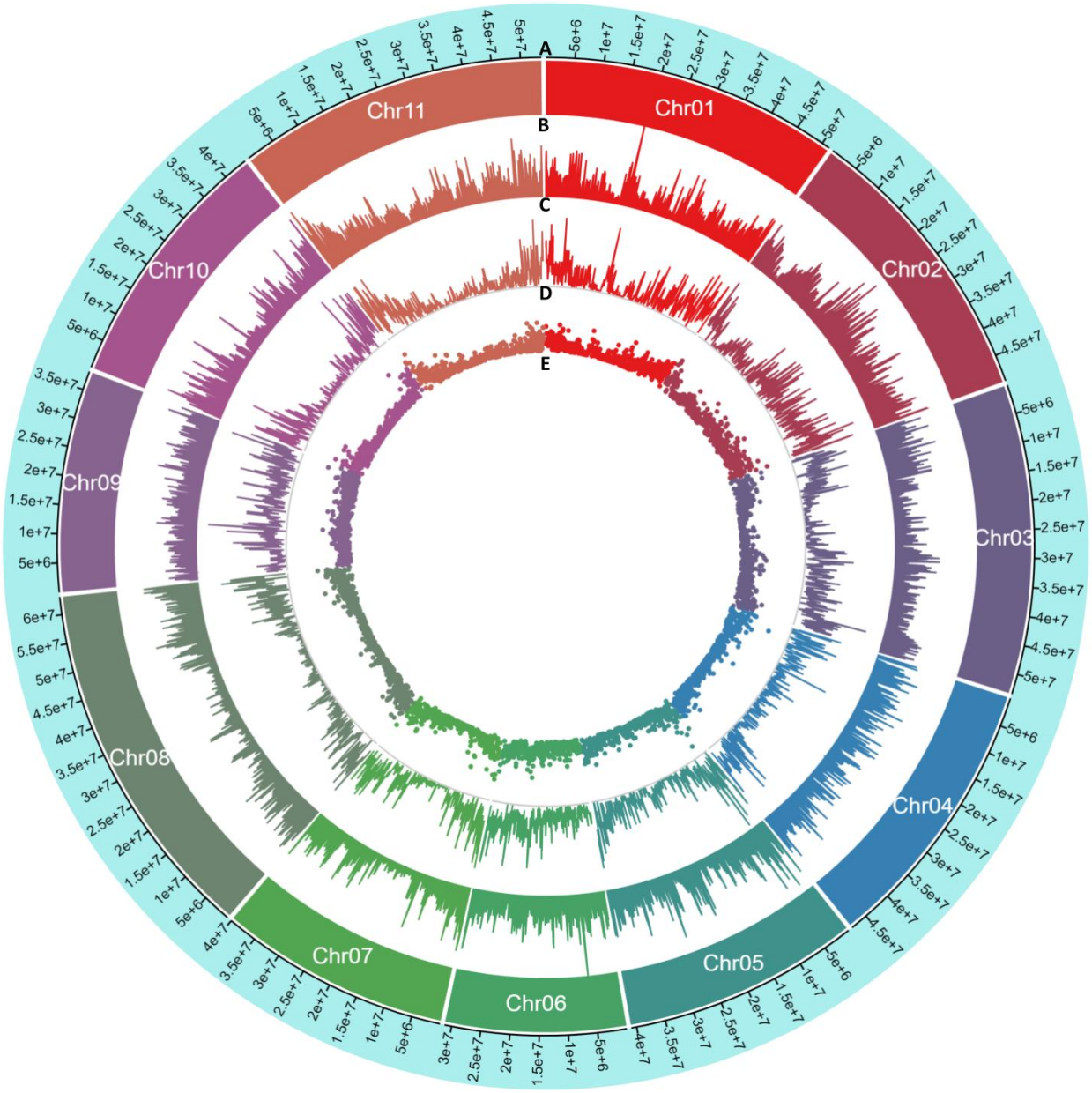
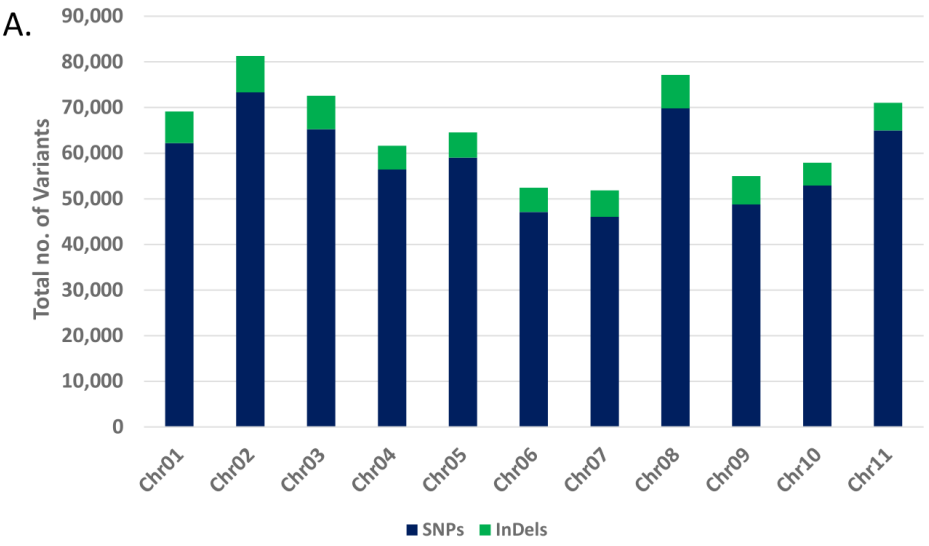


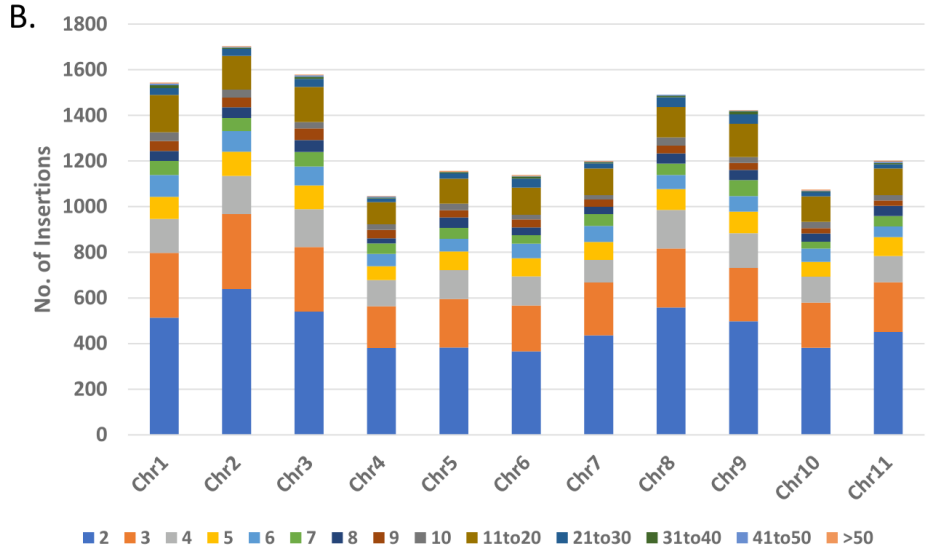
Figure 3. Circos plot to show the density (no. of variants/kb) of SNPs, insertions and deletions in bins of 100 Kb on 11 chromosomes of *Phaseolus vulgaris*. The outermost track (A) denotes the physical distance in scientific notation on each of the eleven chromosomes at 5 Mb break-point. Track B denotes the chromosome numbers. Track C shows the area plot of SNP density. Track D represent the line plot of insertion density and Track E represent the scatter plot of deletion density.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 October 2020 doi:10.20944/preprints202010.0501.v1

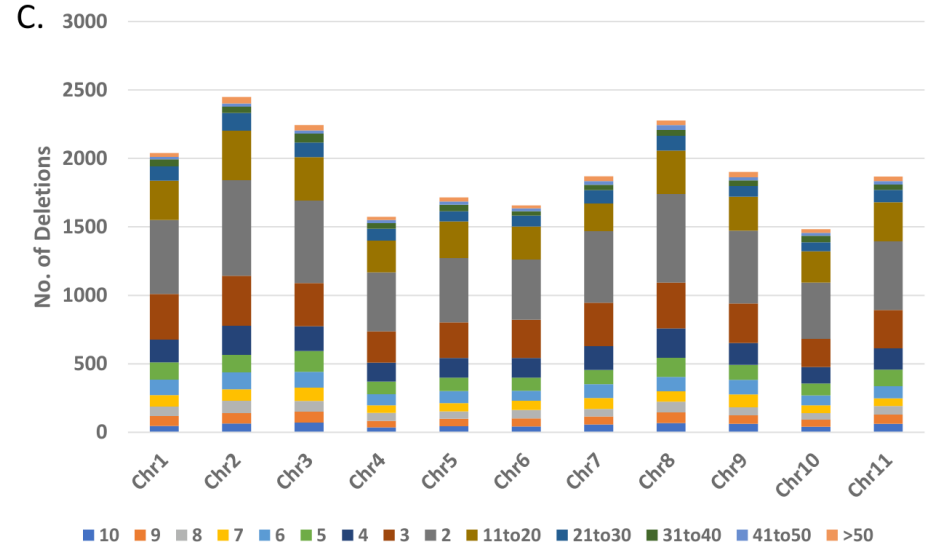
A.



B.



C.



Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 October 2020 doi:10.20944/preprints202010.0501.v1

Figure 4. Genome-wide distribution of SNPs and InDels. Chromosome-wise distribution of SNPs and InDels (A). Chromosome-wise distribution of indels ranging from 2bp to >50 bp (B). Chromosome-wise distribution of indels ranging from 2bp to >50 bp (C).

Table 3. Distribution of SNPs and InDels on eleven chromosomes of *Phaseolus vulgaris*.

Chromosome no.	Size (Mb)	No. of SNPs	No. of Insertions	No. of Deletions	No. of InDels
Chr01	51.43	62,199	3,156	3,766	6,922
Chr02	49.67	73,326	3,522	4,429	7,951
Chr03	53.44	65,222	3,241	4,094	7,335
Chr04	48.05	56,422	2,249	2,928	5,177
Chr05	40.92	59,007	2,383	3,146	5,529
Chr06	31.24	47,079	2,339	3,014	5,353
Chr07	40.04	46,028	2,500	3,301	5,801
Chr08	63.05	69,823	3,144	4,177	7,321
Chr09	38.25	48,746	2,830	3,414	6,244
Chr10	44.30	52,900	2,226	2,793	5,019
Chr11	53.58	64,977	2,575	3,488	6,063
Total	**513.97**	**645,729**	**30,165**	**38,550**	**68,715**

Table 4. Number of transitions and transversions based on the SNPs identified in 82 lines of *Phaseolus* species based on reference genome of *Phaseolus vulgaris*.

Substitution type	Substitution	Count
Transversions (Tv)	C/G	46,390
	G/T	59,644
	A/C	59,095
	A/T	73,275
Transitions (Ts)	A/G	204,568
	C/T	202,757
	Ts	407,325
	Tv	238,404
Ratio	**Ts/Tv**	**1.71**

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 October 2020 doi:10.20944/preprints202010.0501.v1

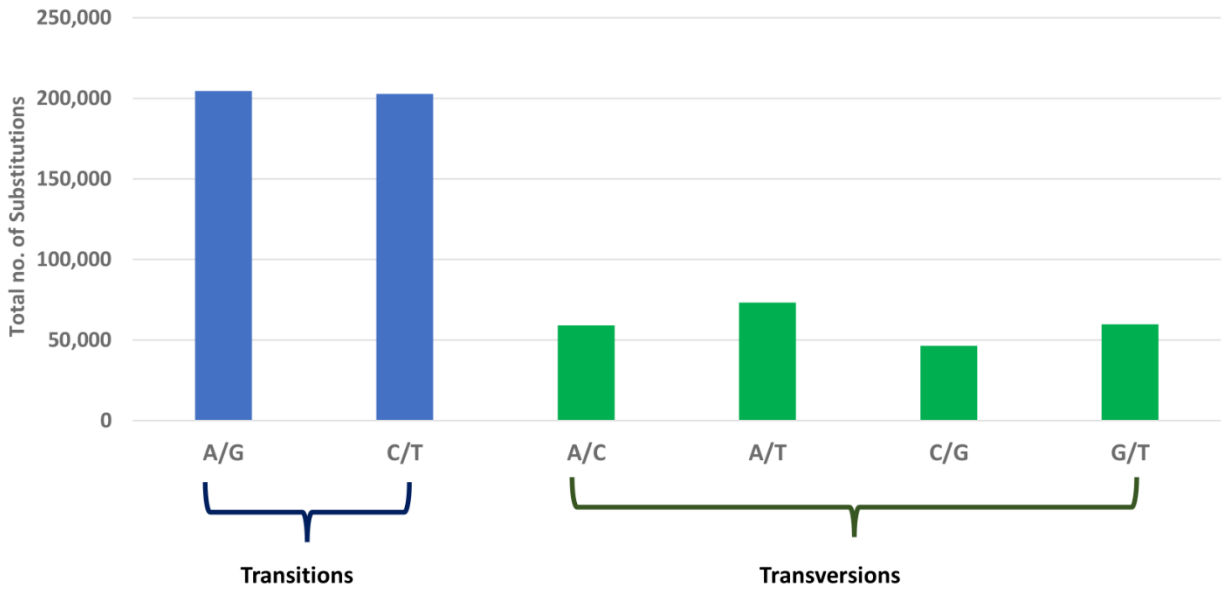


Figure 5. Total number of transitions and transversions observed.

Only 71,544 (11%) SNPs were identified in the exonic regions of chromosomes (**Figure 6A**). The exonic SNPs were further annotated into nonsynonymous (28,373, 39.65%), synonymous (42,906, 59.97%), stop gain (221, 0.003%) and stop loss (44, 0.0006%) SNPs (**Figure 6B**). Out of 68,715 InDels, only 1535 (2.2%) InDels were identified within the exons on chromosomes. Maximum number of exonic InDels were annotated as non-frameshift deletions (512), followed by non-frameshift insertion (420), frameshift deletion (341), frameshift insertion (236), stop gain (21) and the least, stop loss (5) (**Figure 6B**). Overall, SNPs and InDels showed similar distribution patterns in the genome. These variants were found in intergenic, intronic, splicing, UTR downstream and upstream region of genes (**Figure 6A, B**).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 October 2020 doi:10.20944/preprints202010.0501.v1

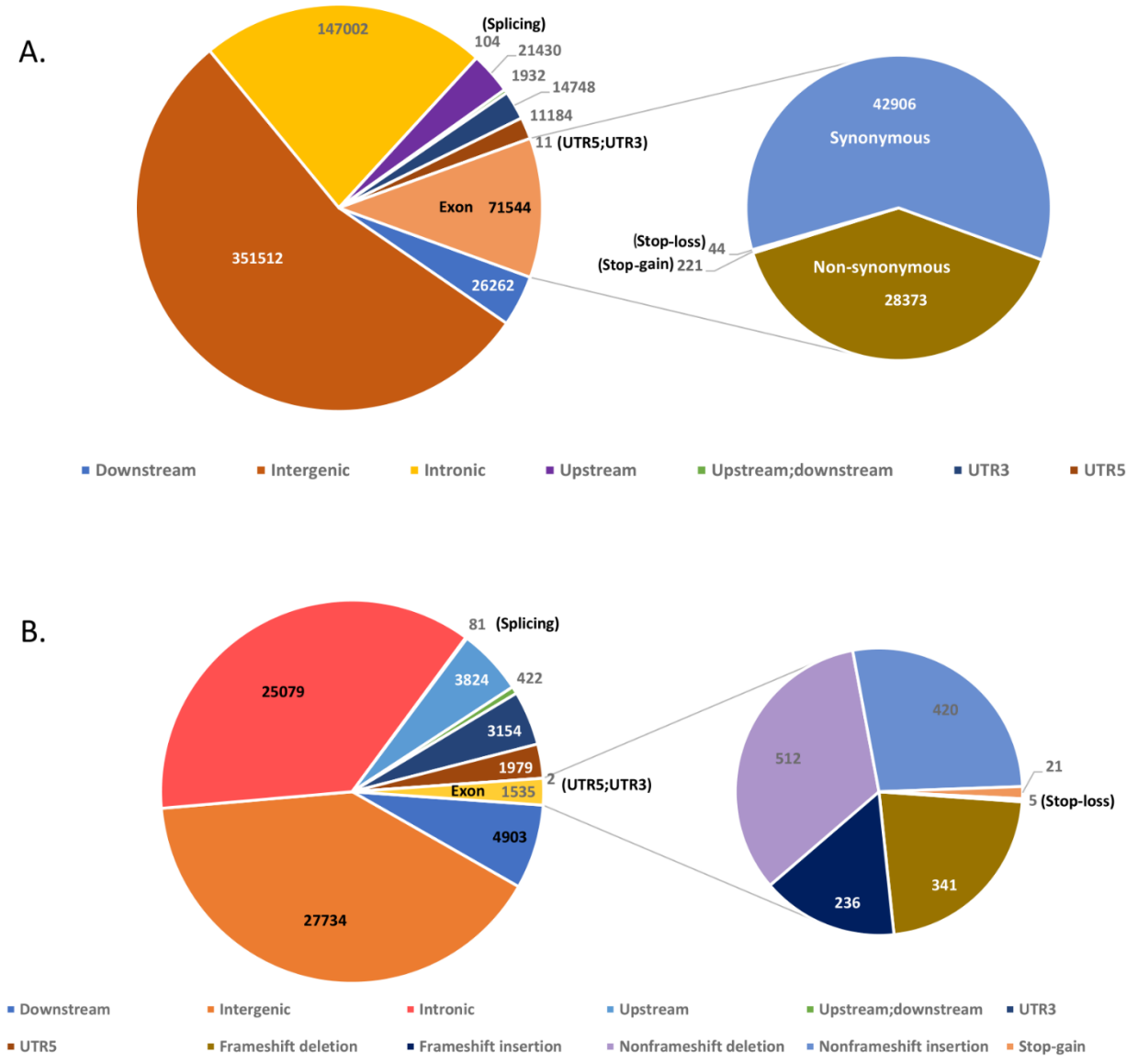


Figure 6. Annotation of single nucleotide polymorphisms (SNPs) and InDels. Distribution of SNPs (A) and InDels (B) in different regions of genome. SNPs in exonic regions are classified into synonymous, non-synonymous, stop-loss and stop-gain categories (A). Indels in exonic regions are classified into frameshift deletion, frameshift insertion, nonframeshift deletion, stop-gain and stop-loss categories (B).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 October 2020 doi:10.20944/preprints202010.0501.v1

4. Discussion

A total of 88 different *Phaseolus* genotypes were evaluated for natural resistance to CuLCrV and SiGMFV, with 84 evaluated in 2018 and 80 evaluated in 2019. There were 76 genotypes common in both the years. Further, the 82 genotypes were sequenced, and the SNP and InDel variants were identified. Overall, the aim was to identify both phenotypic variability (symptom severity to CuLCrV and SiGMFV) and diversity within the genomes of these genotypes. All genotypes displayed begomovirus-associated symptoms in the field suggesting that none of the genotype was immune. The disease severity ranged from 5 to 100%, indicating a considerable difference in disease resistance among the genotypes. However, we observed some inconsistencies for the phenotypic response of genotypes in 2018 and 2019. For example, the genotypes; Affirmed, Blush, Royal burgundy, Prevail and Tema showed high-to-moderate level of resistance against the leaf crumple disease in 2018 (severity: 5-23%). However, in 2019 the symptom severity for these genotypes ranged from 46-55%. This could be due to comparatively higher level of infestation with whiteflies in 2019 vs. 2018 resulting in presumably higher inoculation events with one and/or both begomoviruses. Moreover, percentage of genotypes that were mixed infected with both begomoviruses were higher in 2019 (38.5%) vs. 2018 (27.5%) and as per the previous observations these plants can display severe symptoms compared with when they are infected with either of the viral pathogens [13]. It is possible that more genotypes were mixed infected in 2019 than in 2018 resulting in severe symptoms as observed for the same genotypes earlier. Interestingly, the two Lima bean genotypes, Fordhook and Jackson wonder were resistant and highly-resistant in both years. Adult whitefly counts differed among the genotypes during both the years indicating a potential difference in preference to these genotypes or host-related factors that that repel whiteflies, which needs to be investigated further. Carefully planned extensive preference and biology experiments are required to fully comprehend the level of resistance of snap bean genotypes to whiteflies.

Next generation sequencing (NGS) technology particularly the WGS with downstream computational analyses have provided a quick and accurate method to discover genome-wide variations and identify marker-trait associations as exemplified in several other studies [14-19]. We therefore generated WGS data of 82 *Phaseolus* genotypes and aligned it on the *P. vulgaris* reference genome [1]. A wide variation in the total number of sequenced reads was observed (59.6 million to 146.1 million) with a mapping rate ranging from 55.19% to 97.19%. The low mapping rate of genotypes Fordhook and Jackson wonder is due to the fact that these two genotypes belong to *P. lunatus*; however, were mapped on to the *P. vulgaris* reference genome. The reason for low mapping could lie in the breeding history of these cultivars, which might have resulted in allelic admixture events in these nine *P. vulgaris* genotypes. With an average density of 125 SNPs, five insertions and seven deletions/100kb, variants were differentially distributed throughout the genome. There were several 100 kb bins on each of the 11 chromosomes that did not contain such variants. Despite having uniform genome coverage of mapped reads several empty bins were identified because of the stringent variant calling parameters used as indicated in the methods section. Such significant differential distribution of DNA polymorphisms has also been reported in Arabidopsis and rice [20-22].

Varshney et al. [23] reported SNP and InDel densities (per100 kb) of 63.3 SNPs and 38 InDels in cultivated chickpea, and 103.4 SNPs and 67.4 InDels in wild chickpea using 412 cultivated and seven wild chickpea genotypes. We observed a higher SNP but lower InDel density in our 82

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 October 2020 doi:10.20944/preprints202010.0501.v1

genotypes when compared to a cool-season legume crop (chickpea). The variant density is expected to increase even further if we consider a larger set of genotypes for genotyping. This clearly indicate that *Phaseolus* has more genetic diversity than its cool season counterpart that can be deployed for breeding for disease resistance. The identified SNP density in this study (125/100 kb) is also comparable to a warm-season legume, soybean (~100 SNP/100 Kb) [24]. In our study, the ratio of non-synonymous to synonymous SNPs was found to be 0.66, which is less than the ratio observed in pigeon pea (*Cajanus cajan*; 1.18) [25], soybean (*Glycine max*.; 1.36) [24], rice (*Oryza sativa*; 1.18) [26], sorghum (*Sorghum bicolor*; 1.0) [27] and chickpea (*Cicer arietinum*; 1.20) [23]. Lesser ratio on our study indicate that synonymous substitutions in the studied *Phaseolus* genotypes are tolerated, but the non-synonymous substitutions are removed by purifying selection. It suggests that functionally constrained regions of genes evolve at a slower rate than regions that are not functionally constrained.

The Ts/Tv ratio is often used as a quality indicator of variation data produced from NGS experiments. A higher ratio is an indicator of good quality SNPs as sequencing errors and false positive variants have a ratio closer to one [28].We found SNP transitions (A/G and C/T) are the most common substitution in the genome, which is consistent with other crop species like foxtail millet (*Setaria italica*) [29], tea (*Camellia sinensis*) [30], soybean [31], rice [22]. We observed a Ts/Tv ratio of 1.71 is however less than the ratios reported in crops like rice [22], maize (*Zea mays*) [32] and tea [30]. The higher Ts/Tv could be because of more synonymous mutations resulting due to transitions than transversions, which brings out the change in protein structure and function.

Overall, we identified 20 genotypes (18 snap bean and 2 Lima bean) that consistently displayed high-to-moderate level of resistance to begomoviruses under field conditions. Further characterization and confirmation of resistance response should be conducted under controlled greenhouse conditions with standard parameters (exposure to standard or equal number of viruliferous whiteflies). Earlier studies deployed genotyping by sequencing (GBS) which resulted in reduced representation of genome and captured less genomic variants [18] or used much less frequently present simple sequence repeats (SSRs) [33]. The WGS approach captured an unprecedented number of genomic variants that can be used to identify the genetic basis of disease tolerance against the begomoviruses in *Phaseolus* species.

Conclusions

Our study reports the occurrence of CuLCrV and/or SiGMFV-induced symptoms in *Phaseolus* genotypes including 80 snap beans and two Lima bean genotypes. Based on our phenotyping experiments in field and genomics assisted studies we conclude that the tested genotypes depict significant variations in susceptibility against one and/or both viruses. Future comprehensive studies will be carried out with larger set of *Phaseolus* germplasm materials, which will aid in associating genetic diversity with diverse disease response against both the begomoviruses.

Acknowledgments

This study was supported in part by resources and technical expertise from the Georgia Advanced Computing Resource Center, a partnership between the University of Georgia Office of the Vice President for Research and Office of the Vice President for Information Technology. This work is partially supported by the Specialty Crop Block Grant AWD00011228 and the USDA Award AWD00011012. This work is also partially supported by the Specialty Crops Research

Initiative Award 2018-51181-28420 from the USDA National Institute of Food and Agriculture. The University of Georgia is an equal opportunity provider and employer

Author Contributions

BD conceived the project, GA performed the bioinformatics analyses. SRK and AS performed the phenotyping and DNA isolation of plant samples. SRK and SG analyzed the field phenotyping and whitefly data. GA and SRK complied the manuscript. BD, GA and RS contributed in planning and designing the experiment and manuscript revision. BD GA and RS designed and finalized the manuscript. BD planned the project, secured extramural funds, and revised and submitted manuscript.

Legends to supplementary figures and tables

Supplementary Figure S1. Response of bean genotypes to leaf crumple disease in the field in 2019. Genotype names from left to right- A: Jackson Wonder (5); B: Fordhook (18); C: Furano (40); D: Spartan half runner (40); E: Cascade and F: Caprice (100). Figures in parenthesis are mean disease severity values at 45 days after sowing.

Supplementary Figure S2. Overview of raw data generated and data retained for after quality filtering of 82 lines of *Phaseolus* species for mapping and downstream analyses. Overall >97% of data was retained after quality filtering of raw data.

Supplementary Figure S3. Read mapping statistics of filtered data of 82 lines of *Phaseolus* species on to the reference genome of *Phaseolus vulgaris*. Fourteen out of the 82 Phaseolus lines showed <75% mapping.

Supplementary table S1: Summary of raw read data generated and amount of clean data obtained after filtering the raw data.

Supplementary table S2: Total number of filtered reads obtained the number of reads mapped on to the reference genome with mapping percent, average depth and coverage distribution.

Supplementary table S3: SNP density (no. of SNPs/Kb) calculated in bins of 100 Kb throughout the genome on all eleven chromosomes.

Supplementary table S4: Insertion density (no. of SNPs/Kb) calculated in bins of 100 Kb throughout the genome on all eleven chromosomes.

Supplementary table S5: Deletion density (no. of SNPs/Kb) calculated in bins of 100 Kb throughout the genome on all eleven chromosomes.

Supplementary table S6: Frequency of length distribution of insertions among the 11 chromosomes

Supplementary table S7: Frequency of length distribution of deletions among the 11 chromosomes

References

1. Schmutz, J.; McClean, P.E.; Mamidi, S.; Wu, G.A.; Cannon, S.B.; Grimwood, J.; Jenkins, J.; Shu, S.; Song, Q.; Chavarro, C. A reference genome for common bean and genome-wide analysis of dual domestications. *Nature genetics* **2014**, *46*, 707-713.
2. United States Dry Bean Council. https://www.usdrybeans.com/industry/production-facts/. Availabe online: (accessed on
3. Georgia Agricultural Commodity Rankings. https://caed.uga.edu/content/dam/caes-subsite/caed/publications/annual-reports-farm-gate-value-reports/2018%20Farm%20Gate.pdf. **2018**.
4. Guzman, P.; Sudarshana, M.; Seo, Y.-S.; Rojas, M.; Natwick, E.; Turini, T.; Mayberry, K.; Gilbertson, R. A new bipartite geminivirus (begomovirus) causing leaf curl and crumpling in cucurbits in the Imperial Valley of California. *Plant Disease* **2000**, *84*, 488-488.
5. Larsen, R. First report of Cucurbit leaf crumple virus in snap bean in Georgia. In Proceedings of Phytopathology; pp. S68-S68.
6. Durham, T.; Baker, C.; Jones, L.; Snyder, L.U. First report of Sida golden mosaic virus infecting snap bean (Phaseolus vulgaris) in Florida. *Plant disease* **2010**, *94*, 487-487.
7. Mukeshimana, G.; Butare, L.; Cregan, P.B.; Blair, M.W.; Kelly, J.D. Quantitative trait loci associated with drought tolerance in common bean. *Crop Science* **2014**, *54*, 923-938.
8. Villordo-Pineda, E.; González-Chavira, M.M.; Giraldo-Carbajo, P.; Acosta-Gallegos, J.A.; Caballero-Pérez, J. Identification of novel drought-tolerant-associated SNPs in common bean (Phaseolus vulgaris). *Frontiers in Plant Science* **2015**, *6*, 546.
9. Moghaddam, S.M.; Mamidi, S.; Osorno, J.M.; Lee, R.; Brick, M.; Kelly, J.; Miklas, P.; Urrea, C.; Song, Q.; Cregan, P. Genome-wide association study identifies candidate loci underlying agronomic traits in a Middle American diversity panel of common bean. *The Plant Genome* **2016**, *9*, 1-21.
10. Extension bulletin. https://secure.caes.uga.edu/extension/publications/files/pdf/B%201369_4.PDF. **2007**.
11. Yang, H.; Wang, K. Genomic variant annotation and prioritization with ANNOVAR and wANNOVAR. *Nature protocols* **2015**, *10*, 1556-1566.
12. Gautam, S. The role of Bemisia tabaci in the transmission of vegetable viruses in the farmscape of Georgia (PhD thesis). Athens, GA, Dec, 2019.
13. Gautam, S.; Gadhave, K.R.; Buck, J.W.; Dutta, B.; Coolong, T.; Adkins, S.; Srinivasan, R. Virus-virus interactions in a plant host and in a hemipteran vector: Implications for vector fitness and virus epidemics. *Virus research* **2020**, *286*, 198069.
14. Agarwal, G.; Jhanwar, S.; Priya, P.; Singh, V.K.; Saxena, M.S.; Parida, S.K.; Garg, R.; Tyagi, A.K.; Jain, M. Comparative analysis of kabuli chickpea transcriptome with desi and wild chickpea provides a rich resource for development of functional markers. *PloS one* **2012**, *7*, e52443.
15. Agarwal, G.; Clevenger, J.; Pandey, M.K.; Wang, H.; Shasidhar, Y.; Chu, Y.; Fountain, J.C.; Choudhary, D.; Culbreath, A.K.; Liu, X. High-density genetic map using whole-genome resequencing for fine mapping and candidate gene discovery for disease resistance in peanut. *Plant biotechnology journal* **2018**, *16*, 1954-1967.
16. Agarwal, G.; Clevenger, J.; Kale, S.M.; Wang, H.; Pandey, M.K.; Choudhary, D.; Yuan, M.; Wang, X.; Culbreath, A.K.; Holbrook, C.C. A recombination bin-map identified a major QTL for resistance to Tomato Spotted Wilt Virus in peanut (Arachis hypogaea). *Scientific reports* **2019**, *9*, 1-13.
17. Clevenger, J.; Chu, Y.; Chavarro, C.; Agarwal, G.; Bertioli, D.J.; Leal-Bertioli, S.C.; Pandey, M.K.; Vaughn, J.; Abernathy, B.; Barkley, N.A. Genome-wide SNP genotyping resolves signatures of selection and tetrasomic recombination in peanut. *Molecular plant* **2017**, *10*, 309-322.

18. Oladzad, A.; Porch, T.; Rosas, J.C.; Moghaddam, S.M.; Beaver, J.; Beebe, S.E.; Burridge, J.; Jochua, C.N.; Miguel, M.A.; Miklas, P.N. Single and multi-trait GWAS identify genetic factors associated with production traits in common bean under abiotic stress environments. *G3: Genes, Genomes, Genetics* **2019**, *9*, 1881-1892.
19. Lobaton, J.D.; Miller, T.; Gil, J.; Ariza, D.; de la Hoz, J.F.; Soler, A.; Beebe, S.; Duitama, J.; Gepts, P.; Raatz, B. Resequencing of common bean identifies regions of inter–gene pool introgression and provides comprehensive resources for molecular breeding. *The plant genome* **2018**, *11*, 1-21.
20. Zhang, W.; Sun, X.; Yuan, H.; Araki, H.; Wang, J.; Tian, D. The pattern of insertion/deletion polymorphism in Arabidopsis thaliana. *Molecular Genetics and Genomics* **2008**, *280*, 351-361.
21. Jain, M.; Moharana, K.C.; Shankar, R.; Kumari, R.; Garg, R. Genomewide discovery of DNA polymorphisms in rice cultivars with contrasting drought and salinity stress response and their functional relevance. *Plant Biotechnology Journal* **2014**, *12*, 253-264.
22. Chai, C.; Shankar, R.; Jain, M.; Subudhi, P.K. Genome-wide discovery of DNA polymorphisms by whole genome sequencing differentiates weedy and cultivated rice. *Scientific reports* **2018**, *8*, 1-12.
23. Varshney, R.K.; Thudi, M.; Roorkiwal, M.; He, W.; Upadhyaya, H.D.; Yang, W.; Bajaj, P.; Cubry, P.; Rathore, A.; Jian, J. Resequencing of 429 chickpea accessions from 45 countries provides insights into genome diversity, domestication and agronomic traits. *Nature genetics* **2019**, *51*, 857-864.
24. Choi, I.; Hyten, D.; Specht, J.; Matukumalli, L.; Song, Q.; Quigley, C.; Lee, M.; Chase, K.; Lark, K.; Reiter, R. A soybean transcript map: Discovery and mapping of single nucleotide polymorphisms in soybean genes. *Genetics* **2007**, *176*, 685-696.
25. Varshney, R.K.; Saxena, R.K.; Upadhyaya, H.D.; Khan, A.W.; Yu, Y.; Kim, C.; Rathore, A.; Kim, D.; Kim, J.; An, S. Whole-genome resequencing of 292 pigeonpea accessions identifies genomic regions associated with domestication and agronomic traits. *Nature Genetics* **2017**, *49*, 1082.
26. Xu, X.; Liu, X.; Ge, S.; Jensen, J.D.; Hu, F.; Li, X.; Dong, Y.; Gutenkunst, R.N.; Fang, L.; Huang, L. Resequencing 50 accessions of cultivated and wild rice yields markers for identifying agronomically important genes. *Nature biotechnology* **2012**, *30*, 105-111.
27. Mace, E.S.; Tai, S.; Gilding, E.K.; Li, Y.; Prentis, P.J.; Bian, L.; Campbell, B.C.; Hu, W.; Innes, D.J.; Han, X. Whole-genome sequencing reveals untapped genetic potential in Africa's indigenous cereal crop sorghum. *Nature communications* **2013**, *4*, 1-9.
28. Wang, J.; Raskin, L.; Samuels, D.C.; Shyr, Y.; Guo, Y. Genome measures used for quality control are dependent on gene function and ancestry. *Bioinformatics* **2015**, *31*, 318-323.
29. Bai, H.; Cao, Y.; Quan, J.; Dong, L.; Li, Z.; Zhu, Y.; Zhu, L.; Dong, Z.; Li, D. Identifying the genome-wide sequence variations and developing new molecular markers for genetics research by re-sequencing a landrace cultivar of foxtail millet. *PloS one* **2013**, *8*, e73514.
30. Liu, S.; An, Y.; Tong, W.; Qin, X.; Samarina, L.; Guo, R.; Xia, X.; Wei, C. Characterization of genome-wide genetic variations between two varieties of tea plant (Camellia sinensis) and development of InDel markers for genetic research. *BMC genomics* **2019**, *20*, 935.
31. Ramakrishna, G.; Kaur, P.; Nigam, D.; Chaduvula, P.K.; Yadav, S.; Talukdar, A.; Singh, N.K.; Gaikwad, K. Genome-wide identification and characterization of InDels and SNPs in Glycine max and Glycine soja for contrasting seed permeability traits. *BMC plant biology* **2018**, *18*, 141.
32. Batley, J.; Barker, G.; O'Sullivan, H.; Edwards, K.J.; Edwards, D. Mining for single nucleotide polymorphisms and insertions/deletions in maize expressed sequence tag data. *Plant physiology* **2003**, *132*, 84-91.
33. Perseguini, J.M.K.C.; Oblessuc, P.R.; Rosa, J.R.B.F.; Gomes, K.A.; Chiorato, A.F.; Carbonell, S.A.M.; Garcia, A.A.F.; Vianello, R.P.; Benchimol-Reis, L.L. Genome-wide association studies of anthracnose and angular leaf spot resistance in common bean (Phaseolus vulgaris L.). *PLoS One* **2016**, *11*, e0150506.