BSA-seq approach identifies genomic regions associated with altered plant growth in cucumber (*Cucumis sativus* L.)

R. Słomnicka[1], K. Kaźmińska[1], S. Mużacz[1], M. Grad[1], K. Zembrzucki[1], D. Stokowiec[1], G. Bartoszewski[1]

[1]Department of Plant Genetics Breeding and Biotechnology, Warsaw University of Life Sciences – SGGW
Nowoursynowska 159, 02-787 Warsaw, Poland

Introduction

Plant architecture plays a crucial role in crop management and yield. Identification of genes that control plant growth habit is essential for modern plant breeding and can significantly accelerate breeding progress. In cucumber, only several genes responsible for altered plant growth at the molecular level have been identified (Liu et al. 2021).
The study aimed to identify genomic regions carrying *cp-2*, *scp* and *sp-2* genes that control growth habit in cucumber.

Materials & Methods

We applied BSA-seq to three segregating $F_{2:3}$ populations derived from the cross between maternal line (L500) with normal growth and paternal lines (L504, L505 and L511) characterized by altered growth due to *cp-2*, *scp* and *sp-2* genes (Fig. 1). For each population, two bulked samples were prepared, consisting of equimolar DNA isolated from dominant homozygotes (normal growth) and recessive homozygotes (altered growth). Paired-end libraries for each pool were sequenced using Illumina NovaSeq technology. The re-sequencing data were mapped on the cucumber reference genome (9930 v3, Li et al. 2019). Single nucleotide polymorphisms (SNPs) and insertions/deletions (INDELs) were identified and annotated. The ΔAll-index parameter was used to define genomic regions associated with altered plant growth (Zhou et al. 2022).

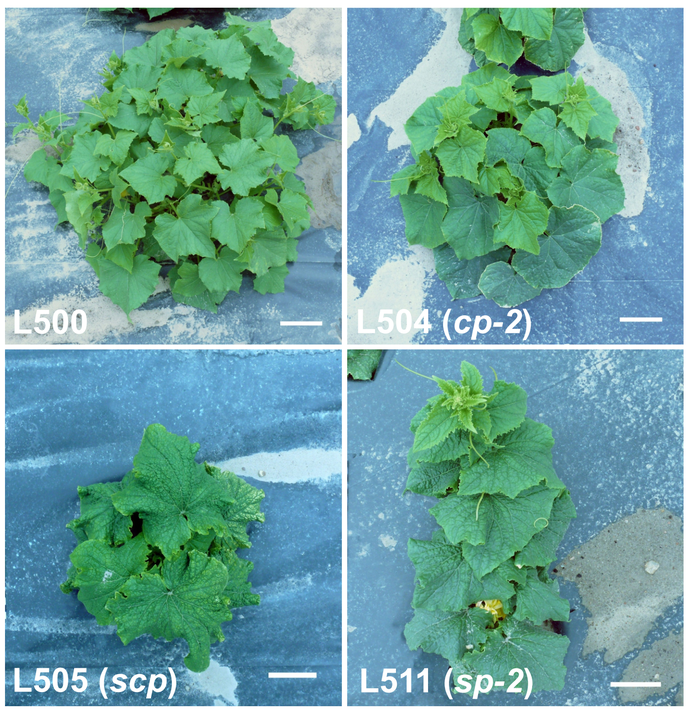

Figure 1. Control inbred line L500 with normal growth and inbred lines characterized by altered growth grown under field conditions. Scale bars, 10 cm.

Results

The re-sequencing of bulked samples resulted in high-quality reads that mapped to the reference genome with average mapping rate of 96% and genome coverage exceeding 30× (Tab. 1). The average number of SNPs was 100k, including 34k exonic SNPs, while 25k INDELs, including 2.2k exonic INDELs were identified for each pooled samples. The details for exonic SNPs and INDELs are shown in Fig. 2.

Table 1. Summary of the re-sequencing and mapping of the reads for six bulked samples representing dominant homozygotes (normal growth) and recessive homozygotes (altered growth) in segregating $F_{2:3}$ populations L500×L504, L500×L505 and L500×L511.

#	Bulked sample	Total lenght of clean reads (Gb)	Number of clean reads	Q20	Number of mapped reads	Mapping ratio (%)	Average depth
			L500×L504				
1	***Cp-2/Cp-2***	20.3	135,577,514	97.0	129,120,146	95.2	67×
2	***cp-2/cp-2***	18.9	92,747,506	96.9	89,705,616	96.7	48×
			L500×L505				
3	***Scp/Scp***	18.2	121,411,346	96.3	116,193,316	95.7	63×
4	***scp/scp***	25.1	167,416,596	96.5	160,880,425	96.1	81×
			L500×L511				
5	***Sp-2/Sp-2***	19.1	127,373,994	96.6	120,738,688	94.8	62×
6	***sp-2/sp-2***	15.5	103,179,152	96.9	97,950,663	97.0	52×

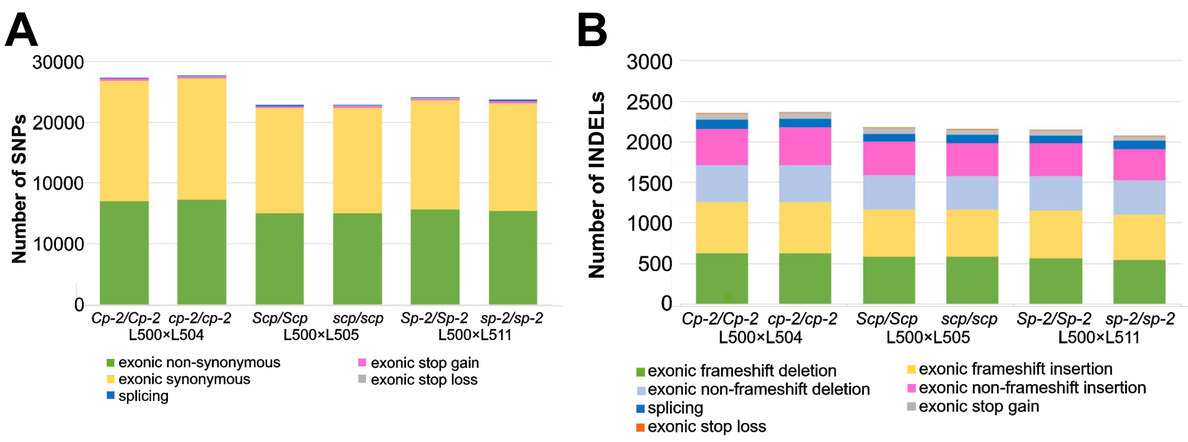

Figure 2. Histograms of the number of exonic SNPs (A) and INDELs (B) obtained by comparison each bulked sample with reference genome 9930 v3.

Acknowledgements

The study was supported by a research program of the Polish Ministry of Agriculture and Rural Development ”Basic research for biological progress in crop production 2021-2027”, task 33 “Identification of genes associated with altered growth type in cucumber plants (*Cucumis sativus* L.)”.

Results

Re-sequencing data of bulked samples and parental lines were analyzed. BSA-seq analysis provided information on the distribution of polymorphisms across the cucumber genome. For populations L500×L505 and L500×L511, several genomic regions were found for which only single polymorphisms were identified. The genomic regions containing *scp* and *cp-2* genes were identified at the upper and lower arm of chr 4, respectively. A relatively large region containing *sp-2* gene was identified on chr 3 (Fig. 3).

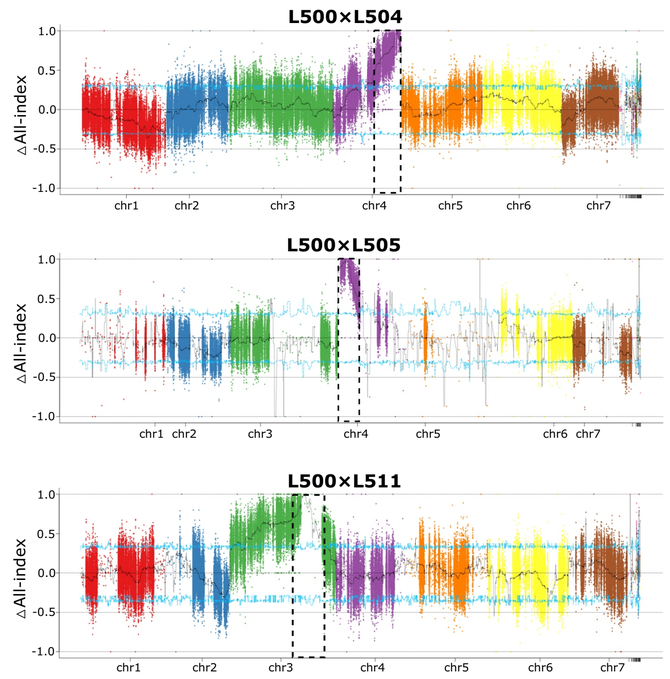

Figure 3. Identification of candidate genomic regions associated with altered growth habit in segregating $F_{2:3}$ populations L500×L504, L500×L505 and L500×L511.
ΔAll index value close to 1 indicates that a particular region of the genome is linked to altered growth habit. The genomic regions associated with altered growth are marked with a black dashed line.

Conclusions

1. SNPs and INDELs analysis allowed to identify *scp* and *cp-2* genes at the upper and lower arm of chr 4 respectively, and *sp-2* on the chr 3.
2. BSA-seq revealed an uneven distribution of polymorphisms for the L500×L505 and L500×L511 populations, which could be related to similar genetic background of the parental lines.
3. This study provides new insights into the genomic regions controlling altered growth habit in cucumber.

Literature

1. Li Q et al. (2019) A chromosome-scale genome assembly of cucumber (*Cucumis sativus* L.). Gigascience 8:giz072
2. Liu X et al. (2021) Genetic regulation of shoot architecture in cucumber. Hortic Res 8:143
3. Zhou et al. (2022) Conjunctive analyses of BSA-Seq and BSR-Seq unveil the Msβ-GAL and MsJMT as key candidate genes for cytoplasmic male sterility in alfalfa (*Medicago sativa* L.). Int J Mol Sci 23:7172