Application of Bioinformatics in Crop Improvement: Annotating the Putative Soybean Rust resistance gene Rpp3 for Enhancing Marker Assisted Selection

Okii D, Chilagane LA, Tukamuhabwa P and Maphosa M

Abstract

Despite the wide availability of DNA sequence information freely available online, the challenge is to convert this mass of data into knowledge that can be readily applied in crop improvement programs. The main objective of this study was to annotate the Rpp3 locus in soybean for enhancing the crop's marker assisted selection (MAS). The specific objectives were: (i) to do structural and functional annotation of the Rpp3 locus genetically mapped on Linkage groups (LG) - C2 and physical located on chromosome 6 and (ii) to generate novel markers linked to the rust resistance for MAS in soybean. The soybean query sequence of interest was downloaded from NCBI (www.ncbi. nlm.nih.gov/nuccore/NW_003722736.1) and subsequently analysed with an array of bioinformatics tools to capture information on the characteristics of the Rpp3 gene. The study found DNA transposons as the predominant repeats in the soybean genomic region analysed. 16 non-overlapping genes were predicted to be tightly linked to marker Satt460 and code for various functions from BLASTx analyses. Gene 1 and 12, both code for structural and enzymatic roles, while gene 13 suggests storage proteins mobilization in seeds. Genes 6, 7 and 8 codes for transcription activation, while gene 10 is a transcription deactivator. There was homology to model organisms; Arabidopsis thaliana (dicots) Chromosome 5 as best hit, with expected-value (E-value) of 3e-128 and 76% sequence identity to Oryza sativa japonica Chromosome 2, Oryza sativa, with E-value of 2e-21 and 84% sequence identity. 15 short random primer sequences with 18-24 base pairs were designed to amplify the Rpp3 gene, predicted genes and introns in soybean chromosome 6 though not validated in the study due to economic reasons. Similar studies are recommended on other genes conferring resistance to rust disease for effective gene pyramiding and shortening the soybean breeding cycle.