



In silico genome-wide discovery and characterization of SSRs and SNPs in powdery mildew disease resistant and susceptible cultivated and wild *Helianthus* species

M. Y. Dudhe¹  · Sujatha Mulpuri¹

Received: 11 August 2021 / Revised: 20 May 2022 / Accepted: 30 May 2022
© The Author(s) under exclusive licence to Society for Plant Research 2022

Abstract

In silico approach was applied for the discovery and characterization of SSRs and SNP markers. RNA sequencing was applied to generate large scale genome-wide transcripts associated with powdery mildew disease of sunflower (*Helianthus annuus* L.). The sunflower materials from which the transcripts were generated comprised one susceptible (PS 2023B), two cultivated resistant accessions (TX16R, ID 25) and three wild species (*H. praecox* [1823], *H. niveus* [1452], and *H. debilis* [DEB-689]). For the susceptible and resistant accessions, control (no infection) and pooled (with infection) samples were created based on which the transcripts were developed. A total of 123,058 SSRs were discovered from the assembled transcriptome of each accession. Out of this total obtained SSRs, 112,326 were classified as perfect and 10,732 as compound SSRs. The percent contribution of compound SSRs in susceptible and in all the tolerant accessions ranged from 8.51 to 9.15%. There was no prominent or significant difference in the number of SSR motifs, comparing the susceptible and resistant accessions tested. A significantly higher CG motif was observed in the susceptible accessions than in the resistant accessions. We, therefore, speculate that the CG motif might play a vital role in the susceptibility of sunflower to powdery mildew disease. Similarly, the percentage of GC motif was found to be significantly higher in the resistant accessions. It is likely that the GC motifs might be involved in the induction of resistance to powdery mildew in sunflower plants. Trinucleotide motif based predictions resulted in significantly higher levels of the following amino acids in the susceptible accession: Ile, Lys, Met, Phe, Pro, Thr and Val. It is presumed that these amino acids might be involved in the vulnerability to the powdery mildew disease in 2023B. In total 35,390 Single Nucleotide Polymorphisms (SNPs) loci were identified that were distributed on all 17 linkage groups with the average resolution of 443.2 Kb/SNP. Linkage group 10 had the highest localized SNPs (3631) with an average of 576 Kb/ SNP and 7 had least (781) with average of 441 Kb/ SNP. Most of the SNPs identified were observed to be transition mutations with mainly A/G type substitutions. The discovered markers will prove useful after wet laboratory studies for screening and marker assisted breeding for powdery mildew resistance in sunflower.

Keywords Powdery mildew · Sunflower · Genome-wide transcript · SSRs · SNPs · Linkage group

Introduction

Sunflower oil is mostly used for consumption in India. In 2017–18, a production level of 0.241 million tons at a productivity of 699 kg/ha was reported (Anonymous 2018; Choudhari et al. 2022). Among the most prominent fungal diseases in sunflower, powdery mildew is of very serious

concern due to its widespread distribution with high severity particularly in the tropics (Gulya et al. 1997).

In India powdery mildew was first reported in Mexican sunflower in 2008 (Baiswar et al. 2008) and later in cultivated sunflower in 2009. Around this period, powdery mildew infestation more commonly emerged during the spring season with the flowering and post flowering stages being the most vulnerable. Earlier researchers reported that the disease incidence and severity have become serious during all growing seasons and affect all growing stages and all plant parts (Zimmer and Hoes 1978; Gulya et al. 1997). Disease severity is observed mostly during the post anthesis stage. For this reason, disease control measures are generally

✉ M. Y. Dudhe
mangesh.dudhe@icar.gov.in

¹ ICAR-Indian Institute of Oilseeds Research, Rajendranagar, Hyderabad, Telangana 500 030, India