Validated markers for sunflower (Helianthus annuus L.) breeding

Saeed Rauf1,*, Marilyn Warburton2, Amina Naeem1 and Wardah Kainat1

1 Department of Plant Breeding and Genetics, College of Agriculture, University of Sargodha-Pakistan, Sargodha, Pakistan
2 USDA ARS Corn Host Plant Research, Resistance Unit, Mississippi, Mississippi State, USA

Received 30 May 2020 – Accepted 13 August 2020

Abstract – Sunflower is native to North America and is now grown around the world for edible oil, seed roasting, confectionary products and bird food. Genetic diversity in cultivated and wild germplasm is characterized for use with various breeding objectives. Molecular markers have been developed to facilitate sunflower breeding. This review was undertaken to discuss molecular markers, which have been validated in different genetic backgrounds for traits of economic interest in sunflower. Markers found to be linked to monogenic traits in mapping populations may be used to select plants with those traits; review of the literature identified markers available for several monogenic traits including resistance against pests and pathogens. Markers linked to Quantitative Trait Loci (QTL) for many disease resistance and economically important traits that have also been identified in specific populations and target environments are also reported here. These identified linked markers should be validated in different genetic backgrounds and environments to ensure widespread utility. Publicly available inbred lines carrying traits of interest and validated markers related to them are summarized in this review, which also highlights traits for which these resources are still lacking, possibly due to lack of funding despite the importance of this hybrid crop. Genomic sequence data is now available for sunflower, which must now be exploited to develop new SNP based markers linked to genes of interest to mine allelic diversity related to economically important traits, especially traits well studied in other organisms, such as seed oil content and resistance genes.

Keywords: QTL / disease resistance / gene pyramiding / single nucleotide polymorphism / introgression

Résumé – Marqueurs validés pour la sélection du tournesol (Helianthus annuus L.). Le tournesol est originaire d’Amérique du Nord et est maintenant cultivé dans le monde entier comme source d’huile alimentaire, pour la torréfaction de ses graines, les tournesols de bouche et les aliments pour oiseaux. La diversité génétique du matériel végétal cultivé et sauvage a été caractérisée pour être utilisée à des fins de sélection diverses. Des marqueurs moléculaires ont été développés pour faciliter la sélection du tournesol. Cet article a pour objectif de discuter des marqueurs moléculaires qui ont été validés dans différents contextes génétiques pour les caractères d’intérêt économique du tournesol. Les marqueurs qui se sont avérés liés à des traits monogéniques dans les populations de cartographie peuvent être utilisés pour sélectionner des plantes présentant ces traits ; l’examen de la littérature a permis d’identifier les marqueurs disponibles pour plusieurs traits monogéniques, notamment la résistance aux parasites et aux agents pathogènes. Les marqueurs liés à des locus de caractères quantitatifs (QTL pour quantitative trait loci) pour de nombreuses résistances aux maladies et des caractères économiquement importants qui ont également été identifiés dans des populations spécifiques et des environnements cibles sont également mentionnés. Ces marqueurs liés identifiés doivent être validés dans différents contextes et environnements génétiques pour garantir une utilité généralisée. Les lignées consanguines disponibles de manière publique portant des traits d’intérêt et les marqueurs validés qui leur sont liés sont résumées dans cette revue, qui met également en évidence les traits pour lesquels ces ressources font encore défaut, peut-être en raison d’un manque de financement malgré l’importance de cette culture hybride. Les données de séquences génomiques sont maintenant disponibles pour le tournesol, qui doivent maintenant être exploitées pour développer de nouveaux marqueurs basés sur des SNP liés aux gènes d’intérêt pour exploiter la diversité allélique liée aux caractères économiquement importants, en particulier les caractères bien étudiés dans d’autres organismes, tels que la teneur en huile de graines et les gènes de résistance.

Mots clés : QTL / résistance aux maladies / pyramide des gènes / polymorphisme d’un seul nucléotide / introgression

*Corresponding author: saeedbreeder@hotmail.com
1 Introduction

Sunflower is an important oilseed or confectionary crop consumed by a large percentage of the world’s people. It is ranked as the fourth most popular oilseed crop after palm, soybean and rapeseed (FAOSTAT, 2016), and supplies 10% of the total edible oil produced in the world (FAOSTAT, 2018). Sunflower oil is a rich source of linoleic acid, which is excellent for human health. Its oil contains about 90% fatty acids, 9% phytosterols and 1% vitamin E, which is known to reduce low density lipids, improve immunity and helps to protect against cardiovascular disease (Staughton, 2019). The plant is native to eastern North America, with some evidence also supporting domestication in Mexico (Harter et al., 2004). Wild sunflower is spread across the entire United States, Northern Canada and the Trans-Mexican Volcanic Belt. It was used as a food source and for curing of diseases and wounds by Native Americans. Spanish travelers brought this crop to Europe where it was adapted as an ornamental plant.

The real potential of sunflower as a crop was not realized until it reached Russia (the former Soviet Union). V.S. Pustovoi used half seed reserve selection methods for the improvement of oil content in achenes. He was able to improve oil content from 30% to 40% from 1935 to 1944 with further improvement to 54% in the next decade (Kaya et al., 2012). High oil content lines from Russia were used as source of breeding material to develop new oilseed varieties in many parts of the world. Another breakthrough in sunflower breeding came with the discovery of cytoplasmic male sterility by Leclercq in Helianthus petiolaris (Leclercq, 1969; 1979), followed by the discovery of fertility restoration by Kinman (1970). This opened a new window for hybrid development, which increased the achene yield potential without sacrificing oil content. Sunflower is now cultivated in more than 70 countries with total oil production of 15.85 million tons. World average production stands at 1700 kg ha⁻¹. Ukraine and Russia are the major growers of sunflower, and collectively produced 47% of the world production (FAOSTAT, 2016).

Sunflower breeding has been facilitated by the use of markers (Dimitrijevic and Horn, 2018). Good markers are highly polymorphic, generally co-dominant, have a strong linkage with the trait of interest, measurable at all growth stages, and phenotypically neutral. Many studies in sunflower have documented the importance of markers in marker-assisted selection (MAS), estimation of genetic diversity (Mandel et al., 2011), identification of inbred lines in heterosis breeding (Yue et al., 2007) and understanding heterotic patterns (Iqbal et al., 2010). Markers potentially suitable for MAS have been identified via Quantitative Trait Loci (QTL) mapping of economically important traits including yield and oil content (Abdi et al., 2012; Bert et al., 2004; Mokrani et al., 2002; Léon et al., 2003; Yu et al., 2003); disease resistance including head and stalk rot (Micic et al., 2005a), downy mildew (Brahm and Friedt, 2000; Liu et al., 2012a) and rust (Talukder et al., 2014); nutritional quality such as high oleic acid level (Schuppert et al., 2006; Dimitrijevic et al., 2017), β-tocopherol content (Vera-Ruiz et al., 2006), and γ-tocopherol content (Garcia-Moreno et al., 2006); herbicide resistance (Bulos et al., 2013b); Orobanche resistance (Lu et al., 2000; Iuorás et al., 2004); and selection of cytoplasmic male sterility (Schnabel et al., 2008), nuclear male sterility (Chen et al., 2006) or fertility restorer sources (Horn et al., 2003; Yue et al., 2010; Liu et al., 2012b). Many of these traits are discussed below.

Mapped markers included amplified fragment length polymorphism (AFLP), single nucleotide polymorphism (SNP), cleaved amplified polymorphism (CAP), simple sequence repeat (SSR), and sequence characterized amplified region (SCAR). Markers mapped within or very tightly linked to monogenic traits may be directly useful for MAS in multiple genetic backgrounds. However, markers identified linked to quantitative traits in specific populations (i.e., mapping populations of F₂:₃, RILs or doubled haploids), should be validated in other genetic backgrounds and possibly other environments before use. Moreover, markers should only be used to move resistance genes from the same donor source they were originally mapped, or from any line derived from that same source (Lande and Thompson, 1990). The linkage between the good allele of a marker and the good allele of a gene only works in the mapped or derived lines, except for perfect markers, which are completely linked to the actual causal mutation. Genotyping by sequence (GBS) has allowed large-scale identification of SNP based diversity within genomes, which may be used to mine resistance genes, and identified SNP markers may be validated in other populations (Celik et al., 2016). This review will consolidate information about mapped and validated sunflower markers in the sections below, which will be a useful resource in the genetic improvement of sunflower.

2 Biotic stress resistance breeding

2.1 Rust (Puccinia helianthi)

Rust diseases are caused by the fungus Puccinia helianthi (Fig. 1A), and are prevalent in many parts of the world including the USA, Australia and South Africa. There are high levels of diversity within rust races; 29 races are endemic to its center of origin in the USA and races 300, 304 and 324 are the most prevalent in this region (Friskop et al., 2015). However, race 777 has been the most virulent, able to infect more than nine differentials (which are sunflower genotypes used to determine the various races of the pathogen, including inbred lines 7350, MC-90, MC-29, P-386, HA-R1, HA-R2, HA-R3, HA-R4, HA-R5; Qi et al., 2011). These differentials have been accepted by the ad hoc committee of the International Sunflower Association in 1987 (and reviewed by Gulya, 2006) as containing diverse rust resistance genes. Development of resistant hybrids is the only economical solution to control the disease without use of fungicides, which are expensive, dangerous for farmers to apply, and potentially hazardous to the environment (Bulos et al., 2013b). Phenotyping of resistance levels to determine presence of the resistant allele of rust resistance genes is difficult and laborious, making the development of molecular markers for selection highly desirable.

More than a dozen rust resistance genes have been identified in various accessions of sunflower, including R₁–R₉, R₁₁–R₁₃, R₁₃α, R₁₃β, R₁₄, R₉adv, R₉₁₁, Ps₇₆, Rs₅₃, and R₉D₆ (Zhang et al., 2016). The chromosomal locations of all R genes
have been reported in a review by Jan et al. (2014) and markers have been developed for all mapped rust resistance genes. The genes for which linked markers have been validated in other populations \((R_1, R_2, R_4, R_{4a}, R_5, R_{11}, R_{12}, R_{13a}, R_{13b}, \text{ and } R_{HAR6})\) are shown in Table 1, which can be used as a reference of donor lines for resistance breeding programs. Studies have shown that the resistance provided by race specific genes can be overcome by the evolution of the pathogen into new pathotypes. This process can be slowed by the use of breeding lines containing multiple resistance genes to achieve durable resistance. The creation of new inbred parental lines with multiple resistance genes is only possible through gene pyramiding with the help of molecular markers, since once one resistance gene is fixed for the resistant allele, phenotypic selection cannot differentiate lines fixed for a single gene from lines with additional resistance alleles from other genes (Qi and Ma, 2020).

New lines have been developed having multiple resistance genes identified originally in one line each, such as MC29, HA-R2, and HA-R6. These multiply resistant lines (HA-R12 and HA-R13) were developed by the pedigree method and molecular marker assisted selection. HA-R12 contains genes \(R_2\) from MC29 (AUS) and \(R_{13a}\) from HA-R6. HA-R13 contains \(R_5\) from HA-R2 and \(R_{13a}\) from HA-R6. Both HA-12 and HA-13 showed high levels of resistance to rust races 336 and 777 (Ma et al., 2016). Sunflower confection lines HA-DM2 (carrying the genes \(R_{12}\) and \(P_{ARG}\)) and HA-DM3 (carrying the genes \(R_{13a}\) and \(P_{l_{12}}\)) carry rust resistance and downy mildew genes and were developed through pedigree and marker assisted selection (Ma et al., 2019). Gene stacking has been carried out by crossing parents carrying multiple rust resistance genes such as RO18 × MC29 carrying gene pair \(R_{18}\) and \(R_2\), and linked markers were used in the segregating generations to select lines with multiple rust resistance genes (Lawson et al., 2003). Similarly, an \(F_2\) population originating from HAR2 × MC29 carrying rust resistance genes \(R_5\) and \(R_2\) were genotyped using molecular markers ORS-316, ORS-630, ORS-333, SFW-00211 and SFW-01272 for the selection of multiple rust resistance genes (according to Tab. 1) Recently, a study showed the successful pyramiding of five rust resistance genes in homozygous condition and in various combinations \((P_{ARG} \times R_5/R_{12} \times P_{l_{15}}, R_5/R_{12} \times P_{l_{15}}, R_{13a}/R_{12} \times P_{l_{15}}, R_{13a}/R_{12} \times P_{l_{15}}, \text{ and } R_{13a}/R_{15})\) selected via SSR and SNP markers (Qi and Ma, 2020). Three combinations included rust and downy mildew resistance genes.

### 2.2 Downy mildew \((Plasmopara halstedii)\)

Downy mildew (DM) is one of the most destructive diseases of sunflower, caused by an obligate biotroph, which induces damage to sunflower leaves (Fig. 1B). It was first reported in late 1890s in Northeastern US and is now present in all sunflower-producing countries of the world with the exception of Australia (Spring, 2019). It has been known to cause 50% yield losses under artificial inoculation trials (Spring, 2019). A total of 36 \(Pl\) resistance genes \((R\) genes), \(P_{l_{1}}-P_{l_{35}}\) and \(P_{l_{ARG}}\) have been reported in sunflower to date (Qi et al., 2019). However, in the past four decades, a total of 44 \(P.\) halstedii races have been identified globally, with the highest diversity of races present in North America and France (Trojanová et al., 2017; Viranyi et al., 2015). Resistance genes induce complete resistance against downy mildew; however, markers have been validated for only a few genes \((P_{1ARG}, P_{l_{1}}, P_{l_{2}}, P_{l_{5}}, P_{l_{8}}, P_{l_{16}}-P_{l_{20}}, P_{l_{14}}, \text{ and } P_{l_{15}}; \text{ Tab. 2})\). Resistance genes \(P_{l_{8}}, P_{l_{15}}\) and \(P_{l_{ARG}}\) originated from the wild species \(H.\) argophyllus. \(P_{l_{ARG}}\) confers resistance to all races of downy mildew, and \(P_{l_{15}}\) has been introgressed into cultivated germplasm.

The mapped position of resistance genes were reported by Dimitrijevic and Horn (2018) and Talukder et al. (2019) and are presented in Table 2. MAS and backcrossing were used to develop two resistant confection sunflower lines. HA-DMS

![Fig. 1](image-url)
Table 1. List of validated markers for rust (*Puccinia helianthi*) resistance breeding.

<table>
<thead>
<tr>
<th>Gene (chromosome locations)</th>
<th>Markers</th>
<th>Original mapping population</th>
<th>Validated population</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1 (LG8)</td>
<td>SCT06 (950 b)</td>
<td>RHA 279</td>
<td>MC 90, RHA 279, MC 29, HA-R8, RHA 379, RHA 464 Qi et al. (2011)</td>
</tr>
<tr>
<td>R2 (LG9)</td>
<td>ORS-333</td>
<td>MC 29</td>
<td>MC 90, MC 29a, PhRR3 Qi et al. (2011)</td>
</tr>
<tr>
<td>R2 (LG9)</td>
<td>SFW-00211 and SFW-01272</td>
<td>117 F2 individual HA-89 × MC 29</td>
<td>46 sunflower breeding lines Qi et al. (2015b)</td>
</tr>
<tr>
<td>R4 (LG13)</td>
<td>ORS-316</td>
<td>HA-R3</td>
<td>HA-R13 Qi et al. (2011)</td>
</tr>
<tr>
<td>R4u (LG13)</td>
<td>ORS-799</td>
<td>Suncross 53</td>
<td>Suncross 53a Hysun 47 Qi et al. (2011)</td>
</tr>
<tr>
<td>R4u (LG13)</td>
<td>ORS-45</td>
<td>Suncross 53</td>
<td>Suncross 53a Hysun 47 Qi et al. (2011)</td>
</tr>
<tr>
<td>R5 (LG2)</td>
<td>ORS-316</td>
<td>HA-R2</td>
<td>MC 29a Qi et al. (2011)</td>
</tr>
<tr>
<td>R5 (LG2)</td>
<td>ORS-630</td>
<td>HA-R2</td>
<td>RHA 279a, P-386, RHA 397, Hysun 36, PH4 Qi et al. (2011)</td>
</tr>
<tr>
<td>R11 (LG13)</td>
<td>ORS-728 and ORS-45</td>
<td>Rf ANN-1742</td>
<td>F2 population Qi et al. (2012)</td>
</tr>
<tr>
<td>R12 (LG11)</td>
<td>CRT-275 and ZVG-53</td>
<td>RHA 464</td>
<td>R12 allele very rare Talukder et al. (2014) NSA_003426 and NSA_004155 SNP Gong et al. (2013)</td>
</tr>
<tr>
<td>R13a and R13b (LG13)</td>
<td>ORS-316</td>
<td>HA-R6 and RHA 397</td>
<td>Sunflower germplasm Qi et al. (2011)</td>
</tr>
<tr>
<td>RHAR6 (LG13)</td>
<td>ZVG-61 and ORS-581</td>
<td>HAR6</td>
<td>Susceptible line into a rust resistant isoline Bulos et al. (2013a)</td>
</tr>
<tr>
<td>R4, R12, PLARG (LG13)</td>
<td>ORS-316, NSA-001392, NSA-002798, linked to genes R4, R12, and PlARG</td>
<td>HA-R3, HA-R2, HA-R8, RHA-397</td>
<td>Segregating population Qi and Ma (2020)</td>
</tr>
<tr>
<td>R15 (LG8)</td>
<td>SFW01920, SFW00128, SFW05824 NSA_008457</td>
<td>HA-R8</td>
<td>F2 Ma et al. (2018)</td>
</tr>
</tbody>
</table>

2.3 Sclerotinia White Mold (*Sclerotinia sclerotiorum*)

Sclerotinia white mold is a major disease of sunflower in temperate areas, causing damage to all parts of the plant including the head (Fig. 1C–D). Resistance to the disease has been found in related wild species but is known to be of polygenic inheritance and was absent in cultivated germplasm until recently (Micie et al., 2005a). However, breeding lines such as HA-BSR1 have been registered, which have high tolerance to *Sclerotinia* basal stalk mold originating from parents HA 441/RHA 439 (Talukder et al., 2017). Narrow sense heritability ranged between 2–60% for *Sclerotinia* resistance (Zubrzycki and Maringolo, 2017), indicating the possibility for good gain from selection, and the possibility of identifying genes of major effect on resistance. Gene identification studies using different methods have been successful in this effort. Forty-three candidate genes have been identified through transcript profiling of sunflower and contained Pl16 and HA-DM6 contained Pl35 (Qi et al., 2020). Durable resistance could be achieved by pyramiding two or more genes into a single line using the markers below. SNPs markers were used to develop line with three resistance genes i.e. Pl16, PlARG and R12 in F2 segregating populations from a cross between RHA 340 × RHA 464 (Qi et al., 2017). Dominant markers ORS-1008 and HT-636 were recommended for selection of Pl16 and Pl13 genes, respectively, in their mapping lines as well as new sunflower germplasm (Liu et al., 2012a). Utilization of genomic data and development of SNP markers for DM resistance genes has also been successful following recent studies utilizing the new SNP markers developed for sunflower for GWAS (Pecrix et al., 2018). For instance, resistance genes of two wild and eight domesticated ecotypes of sunflower were mapped on to the sunflower reference genome creating a genotyping array of 49 449 SNPs; this was used to identify 10 resistance genes, (Pecrix et al., 2018).
Brassica napus (Fusari et al., 2012). Association mapping showed that candidate gene HaRIC_B caused a 20% reduction in the head rot incidence. Several mapping studies were carried out to identify QTLs linked to Sclerotinia resistance (Davar et al., 2010; Zubrzycki and Maringolo, 2017). However, markers proposed by transcript profiling, association, or linkage mapping have not been validated in independent populations. Polygenic traits are generally affected by genotype × environment interactions so it is also important to validate lines containing proposed markers in different environments as well as different genetic backgrounds (Dimitrijevic and Horn, 2018).

Sunflower wild species H. tuberosus and H. maximiliani are known to carry resistance (Rönicke et al., 2004; Fusari et al., 2012). Breeding line TUB-5-3234 carrying partial resistance was developed from introgression and phenotypic selection. This line was then used to develop molecular markers related to Sclerotinia white mold resistance component traits such as decreased stem lesions, decreased leaf lesions and speed of fungal growth (Micie et al., 2005a). Markers such as ORS-337, ORS-1129 and ORS-588 were mapped to linkage group (LG) 4, LG10 and LG17 and found to be associated with traits related to white mold resistance (Tab. 3). QTLs were also validated in another population originating from the cross NDBLOSsel (partial resistance)/C2 CM625; although these markers may be useful for the selection of white mold resistance, they must be validated in each genetic background the resistance will be moved into. Genome wide association studies developed 20 522 high quality bi-allelic SNPs suitable for use in diverse genetic backgrounds such as inbred lines, hybrids, open pollinated varieties, and landraces. The developed array was found superior to improve the Sclerotinia mid stalk rot resistance when compared with currently available tools for determining genetic diversity for disease resistance (Livaja et al., 2016).

### Broomrape resistance (Orobanche spp)

Broomrape is a parasitic plant, which causes significant economic damage to sunflower growers around the world, but specifically in Eastern Europe (Kaya, 2014; Fig. 1E). Unlike weeds, parasitic plants directly feed on the sunflower for food and shelter. This causes reduction in leaf area and head diameter, as many of the sunflower’s resources are being taken...
by the parasite. Various parasitic broomrape races denoted A through H have been characterized, and they differ on the basis of their virulence with “H” being the most virulent pathotype (Kaya, 2014). Races F–H are the most prevalent in sunflower growing countries. Resistance is monogenic and race specific, and thus could be defeated by the origin of new races (Kaya, 2014). Multiple resistance genes have been identified and are denoted in the Or series (Tab. 4), and each gene induces vertical (race specific) resistance.

Several resistant hybrids have been developed to reduce yield losses due to the parasitic plants through conventional breeding methods of phenotypic screening and selection (Čvejić et al., 2020). The genomic tools may be further exploited to identify resistance genes and subsequent introduction in elite inbred lines, particularly to pyramid resistance and discourage evolution of new virulence types (Čvejić et al., 2020). SSR markers from LG3 showed strong association with various resistance genes i.e. Or2, Or5, and Or6 (Imerovski et al., 2013). The Or5 gene was identified on the telomeric end of LG3 and it was not possible to find flanking markers for this gene. Therefore, MAS was limited to one side of this gene, closer to the centromere, and four primers i.e. CRT-392, CRT-314, ORS-1034 and ORS-1040 were located within a 6.2 to 11.2 cM range (Tang et al., 2003). Primer ORS-683 was closely linked to gene orab-vl-8 with genetic distance of 1.5 cM on LG3 (Molinero-Ruiz et al., 2015). Primers sets such as C12Q1/6895 and C12Q1/6881 were patented by Gao, (2019) to select sunflower genotypes with increased Orobanche resistance (US Patent, No. 15/946,105; Gao et al., 2019). The primers allow selection of the resistant alleles of the OrDEB2 gene. Gene BRS11 was located on linkage group 4 (Hoef et al., 2011), and the nomenclature of this gene has been changed to ORS11 for the purpose of consistency with identified genes (Martin-Sanz et al., 2020). The gene was linked to SNP marker M-4_30.40 on linkage group 4 (Martin-Sanz et al., 2020).

Markers linked to QTL for Orobanche resistance have also been identified, which explain less of the phenotypic variation for resistance but may be used to increase tolerance. These have been poorly validated in sunflower as compared to disease resistance genes for rust and downy mildew. Genotype-by-sequencing was used to map QTLs related to Orobanche resistance. Two major QTLs or3.1 and or3.2 were identified; or3.1 was mapped near the resistance gene Or5 while Or3.2 was identified for the first time. QTL Or3.1 region later fine-mapped with CAPS markers (Primer sequences presented in supplementary file S1; Imerovski et al., 2019). Exploration of 6.5 MB (31.9 to 38.48 Mb) of genomic sequence on LG3 where resistance QTLs in parental lines were overlapping identified 123 genes including two resistance genes, HanXRQCht-030065701 and HanXRQCht-030065841. In exploration of a 3.72 MB genomic region (97.13 and 100.85 Mb), another putative resistance gene HanXRQCht-030076321 was identified (Imerovski et al., 2019).

3 Economically important traits

3.1 Herbicide tolerance

The sunflower crop is protected from weeds generally by spraying pre-emergence and a few post emergence broad leaf herbicides to reduce infestation in the field (Kaya et al., 2012). Therefore, herbicide tolerance is a desirable trait in sunflower breeding to expand the range of herbicides including imidazolinones (IMI) and sulfonylureas (SU) applied to sunflower fields at various growth stages (Sala et al., 2018). Herbicide tolerance was observed in wild populations of sunflower, which survived in commercial fields of corn, soybean and other crops in Kansas and South Dakota, US (Al-Khatib et al., 1998; White et al., 2002). The trait was spontaneously induced in wild sunflower populations repeatedly treated with herbicides, which induced a partial dominant mutation at locus Ahas1l (acetohydroxyacid synthase). The mutant allele of ahas1l, was subsequently introgressed into elite breeding lines through hybridization and selection.

Subsequently, three alleles were discovered, including AhasIl-1, which confers moderate levels of tolerance against IMI and SU herbicides and contained a C to T mutation in codon 205 (Kolkmann et al., 2004); AhasIl-2, which conferred high levels of tolerance to SU and contained a C to T mutation in codon 197; and AhasIl-3, which contained a G to A mutation in codon 122 and showed high levels of resistance to IMI (Kolkmann et al., 2004). Herbicide resistance was also induced through treatment of mutagen ethyl methane sulfonate to seeds and subsequent selection with imazapyr herbicide in an M2 population (Sala et al., 2008). A homozygous resistant line (CLHA-PLUS) was developed from the selection in a treated population (Sala et al., 2008). Allele AhasIl-3 or allelic combination of AhasIl-1plus AhasIl-3 was sufficient to exploit the herbicide tolerance in hybrids for IMI resistance in sunflower. Marker assisted selection may help to introgress...
various variants of the AhasII gene into elite breeding lines to exploit the trait in hybrid breeding (Bulos et al., 2013b). The SSR primers p-AHAS18 and p-AHAS19, which were initially developed by Kolkman et al. (2004) were validated and found to be polymorphic in lines carrying the alleles AhasIl-1 and AhasIl-3 (Bulos et al., 2013b). Co-dominant inheritance of these SSRs allowed identification of heterozygous hybrids. Two SNP primers p-AHASNidF and pAHAS122TWT were polymorphic in lines carrying the ahasil, AhasIl-1, and AhasIl-2 alleles while two other primers p-AHASNidF with pAHAS122TMU may be used to amplify AhasIl-3 (Bulos et al., 2013b; Jacob et al., 2017). SSR primer pair AHAS16 and AHAS17 was originally reported by Kolkman et al. (2004) and were validated to study the size variation for allele AhasIl-1 in wild germplasm (Jacob et al., 2017).

### 3.2 Oleic acid

Development of high oleic acid lines is an important breeding objective for sunflower. Historically, sunflower contains about 18–25% oleic acid (Rauf et al., 2017). Because oleic acid is beneficial to human health, the high oleic acid trait was created through mutation breeding and then introgressed into new hybrids. “Pervenent” breeding lines contain a dominant mutation, which increases oleic acid content to more than 89% in the sunflower oil. Commercial varieties with high oleic acid content are available and now account for up to 4% of the total sunflower oil production, and generally enjoy a premium in price (Rauf et al., 2017). Selection for high oleic acid is expensive and slow due to the laborious gas chromatography and nuclear infrared resonance protocols. The use of molecular markers could greatly facilitate selection in early segregating generations, and only desirable plants would be carried forward into the next generation. Primer sets such as NI-3F/N2-IR have been used successfully to select for the defective version of the A12-desaturase or FAD2-1D gene which causes the accumulation of high levels of high oleic acid in the sunflower seeds. The primers are perfect markers for this trait, because they are completely linked to the causal mutation. This allows genotypes to be selected in all segregating material for which the trait exists and simplifies the use of this trait (Tab. 5).

### 3.3 Fertility restorer genes

Sunflower is the second biggest crop after maize cultivated through hybrid seed (Dimitrijevic and Horn, 2018). Commercial hybrid seed production is dependent on cytoplasmic male sterility (CMS) and male fertility restorer (RF) lines. Fertility restorer genes are nuclear based and tend to overcome the cytoplasmic male sterility in the F1 generation. A satisfactory restoration of fertility is necessary for high grain filling percentage. There are more than 70 cytoplasmic male sterility sources developed for sunflower, and some of them have been reviewed by Rauf (2019). These CMS sources can only be exploited with suitable restorer genes. Generally, hybrid sunflower breeding is dependent on the RFgene obtained from line T66006-2-1-B (Kinman, 1970). Diversification of cytoplasmic and fertility restorer genes is major breeding goal to reduce genetic vulnerability to diseases and pathogens. A range of fertility restorer genes (RF1–RF7, RF2-P_{RF}) have

---

**Table 4.** List of validated markers for broomrape resistance.

<table>
<thead>
<tr>
<th>Gene (chromosome location)</th>
<th>Markers</th>
<th>Original mapping population</th>
<th>Validated population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Or5 (LG3)</td>
<td>Markers linkage group 3</td>
<td>262 recombinant inbred lines (RILs) (PHC × PHD)</td>
<td>25 inbred lines carrying genes Or6, Or2 and Or4 (Tang et al., 2003; Imorovski et al., 2013)</td>
</tr>
<tr>
<td>Or3 (LG3)</td>
<td>CRT392, CRT314 ORS1036, ORS1040</td>
<td>Bulk segregant analysis (PHC × PHD)</td>
<td>262 F5 RIL population (Tang et al., 2003)</td>
</tr>
<tr>
<td>OrDEB2 (LG4)</td>
<td>C12Q1/6895 and C12Q1/6881</td>
<td></td>
<td>Gao et al. (2019)</td>
</tr>
</tbody>
</table>

**Table 5.** List of validated markers for the selection of high oleic acid contents.

<table>
<thead>
<tr>
<th>Gene (chromosome location)</th>
<th>Markers</th>
<th>Original mapping population</th>
<th>Validated population</th>
</tr>
</thead>
<tbody>
<tr>
<td>A12-oleate desaturase gene (LG14)</td>
<td>NI-3F/N2-IR</td>
<td>Pervenent</td>
<td>PAC-3973, VSFH-2042 and RSFH Nagarathna et al. (2011)</td>
</tr>
<tr>
<td>A12-oleate desaturase gene (LG14)</td>
<td>NI-3F/N2-IR</td>
<td>Pervenent</td>
<td>A1, A2, A4 and A7 line in CMS, R4, R5, R6 and R7 lines in restorer, and A1 × R4, A1 × R5, A2 × R2, A2 × R4, A2 × R7, A4 × R2, A4 × R 6 lines (Tilak et al., 2018)</td>
</tr>
<tr>
<td>FAD2-1D alleles (LG14)</td>
<td>F4-R1, N1-3F/N2-1R</td>
<td>Pervenent</td>
<td>F2 population (Dimitrijević et al., 2017)</td>
</tr>
<tr>
<td>O1 (LG14)</td>
<td>HO_Fsp_b</td>
<td>HO 5–13</td>
<td>Germplasm differing for oil quality (Premnath et al., 2016)</td>
</tr>
</tbody>
</table>
been identified from various sources which are compatible with different CMS sources (Talukder et al., 2019). Some of the validated markers to select various RF genes and their assigned linkage group are shown in Table 6.

3.4 Tocopherol content

Tocopherols are vitamin E active compounds with antioxidant activity and are highly abundant in sunflower oil (Rauf et al., 2017). There are four derivatives (α, β, γ, δ) of fat-soluble tocopherol with vitamin E activity. These tocopherols protect cells as well as oil from oxidative damage, thus prolonging the shelf life of both seed and oil (Rauf et al., 2017). Sunflower seed is predominantly comprised of α-tocopherol; however, substituting the α-tocopherol with γ-tocopherol has improved the shelf life of the oil (García-Moreno et al., 2012).

The selection for oil with modified tocopherol content is expensive and laborious due to the laboratory protocols and the complicated inheritance of interacting loci which encode various tocopherol derivatives (García-Moreno et al., 2012). DNA based markers will provide a cost effective way to select desired oil ratios and levels (Vera-Ruiz et al., 2006). High β-tocopherol content is controlled by recessive alleles at the Tph1 gene. DNA marker based studies showed that three SSR markers (ORS-1093, ORS-222 and ORS-598) on LG1 are tightly linked with Tph1. ORS-716 successfully differentiated the low β-tocopherol genotype (CAS-12) from high β-tocopherol genotypes (i.e., T-589; Vera-Ruiz et al., 2006). Four inbred lines containing high γ-tocopherol content (85%) were developed and reported by García-Moreno et al. (2006). γ-tocopherol content is controlled by recessive alleles at Tph2, which is tightly linked to SSR markers ORS-312 and ORS-599 on LG8 (García-Moreno et al., 2006). The heterozygous recessive genotype tph1tph2 was superior to the homozygous genotype at this locus due to ability to produce both types of β- and γ-tocopherols. A primer combination of γ-TMT-F1/F2/R24 showed polymorphism between two high γ-tocopherol parents i.e. IAST-1 and nmsT2100, concluding that both parents had maximized γ-tocopherols at different loci and were polymorphic in an F2 population created by crossing the two lines. A list of polymorphic validated primer is presented in Table 7.

4 Future work

Sunflower belongs to a highly diverse genus with species that include extremophiles, which may be useful donors of genes to fulfill various sunflower-breeding objectives (Warburton et al., 2017). However, most of the diversity within the germplasm pool is as yet unexplored due to lack activity in the characterization and transfer of valuable genes from related species. Marker assisted selection can overcome drawbacks of phenotypic selection, and MAS for monogenic traits is particularly straightforward. Introgression via MAS can be carefully targeted and include lower linkage drag and allow gene pyramiding, which usually cannot be done through conventional breeding procedures. Today’s breeders are selecting multiple disease resistance genes in segregating populations with markers. Despite these successes, large parts of the genomes and of the collections of cultivated and wild species are as yet uncharacterized and under-utilized due to lack of structural and functional genomic information. The international consortium on sunflower genomics has been able to create a genomic database of 3.6 GB of sequence data available to the public (https://www.ncbi.nlm.nih.gov/assembl y/GCF_002127325.1/). This database may help speed the scanning of the sunflower genome to mine resistance genes in sunflower, using information from related species, including model species for which considerably more genetic information is available.

Genomic data has aided in the elucidation of the evolutionary history of sunflower, and the genetic architecture of at least two important traits (flowering time and the metabolism of oil content) is now better understood (Badouin et al., 2017; Bonnafous et al., 2018). Genome sequences of several of breeding lines showed that the cultivated pan genome is comprised of 61 205 genes, and 27% of these genes
vary between breeding lines (Hübner et al., 2019). A small percentage (1.5%) of the genes are introgressed from wild species, and majority of these genes induce biotic resistance in sunflower (Hübner et al., 2019). A genetic analysis of male and female lines used in development of sunflower hybrids and compared to open pollinated varieties (OPV) showed male lines had a higher percentage of introgressed genes from wild species than did the females or OPVs. Genetic analysis of male and female lines also revealed differentiation for biotic resistance, which was complementary to provide better resistance in hybrids (Owens et al., 2019).

Oil content is a highly economical but polygenic trait in sunflower. Genome wide association selection was used to identify several genes related to the metabolism of oil in sunflower. Utilization of newly developed SNP marker resources allowed identification of SNP markers for traits of interest, which can speed selection via marker assisted selection or targeted interventions of specific genes using gene editing. Genomic selection has been considered a useful way to increase breeding efficiency of uncharacterized parental lines and could be used as alternative to classical general combining ability analyses and phenotypic population selection, and may be helpful to reduce the labor and cost of phenotypic analysis for economical traits such as oil content (Mangin et al., 2019).

The Sunflower Genome Database (https://sunflowergenome.org/) and the XRO genome (https://www.heliagene.org/HanXRO-SUNRISE/) are available on the INRA Sunflower Bioinformatics Resources (https://www.heliagene.org/) and may be used to compare genomic regions of sunflower. Genomic resources regarding sunflower pest such as Orobanche cumana are also available which may provide a genomic insight to this pest, and helps to uncover genomic based diversity within various virulent races of this pest. Genomic resources may be used for expression analysis of genes under stress conditions, pathway and metabolism analysis of economically important traits, and gene sequences that may help finding SNPs linked to genes of interest.

5 Conclusions

Literature was reviewed and validated markers were sought for various traits in sunflower, and these are consolidated and presented here. Validated markers that are available for diverse monogenic traits including diseases resistance, Orobanche resistance, herbicide tolerance, high oleic acid, high tocopherol content, and fertility restorers were identified. Validated markers were also found to be available for the quantitative traits, Sclerotinia white mold and Orobanche resistance. Lists of all these markers are provided in Tables 1–7. These tables provide information on validated markers that will enable sunflower breeding to characterize, diversify and transfer genes between sunflower inbred lines within cultivated germplasm and from wild species without excessive linkage drag. Highly resistant germplasm is now available containing multiple resistance genes including rust, downy mildew and Sclerotinia white rot; these can be used as donor lines and the markers in Tables 1–7 can be used for marker assisted introgression of the beneficial alleles. Validated molecular markers are also available to modify tocopherol levels and fatty acids, which would help develop specialty sunflower lines at lower cost, and validated markers for fertility restorer genes may help to diversify the fertility restoration sources and may improve the performance of hybrids with better grain filling under various environments. The authors hope that the markers, available in one consolidated review article, will aid sunflower breeders around the world in the improvement of selection gain efficiency for traits of interest. Genome sequence resources are made available online which help to understand the evolution history of sunflower, sequence diversity within germplasm, and mine new genes of economic value. Genome sequencing may also be used to develop new SNP markers tightly linked with targeted traits or genes which could be validated and applied for MAS for improvement of sunflower populations.
Supplementary material

The supplementary material is available at https://www.ocl-journal.org/10.1051/ocl/2020042/olm.

**Supplementary Table S1.** Supplementary sequence data of primers 5′–3′.

**Acknowledgments.** The authors would like to thank Drs. Gerald Seiler and Ziv Attia for their thoughtful review of the manuscript. Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the USDA that is an equal opportunity employer.

**Conflicts of interest.** The authors declare that they have no conflicts of interest in relation to this article.

**Funding**

This manuscript has not received any specific funding.

**References**


Cite this article as: Rauf S, Warburton M, Naeem A, Kainat W. 2020. Validated markers for sunflower (Helianthus annuus L.) breeding. OCL 27: 47.