REVIEW

Research advances and prospects of molecular markers in sesame: a review

Micheale Yifter Weldemichael[1] · Hailay Mehari Gebremedhn[1]

Abstract

Sesame (*Sesamum indicum* L, $2n = 2x = 26$) is one of the ancient and most important oilseed crops, extensively cultivated in tropical and sub-tropical countries. Due to its high oil quality and content, rich in vitamins and minerals, ingredients of medical products, antioxidant resistance as well as the main component in the production of soaps, cosmetics, perfumes, and insecticides, sesame is considered as the queen of oilseeds. In spite of its industrial, economical, medicinal and nutritional importance, sesame is considered as an orphan crop as very limited attention has been given from science, academia, industry and policy-makers. In recent years, the situation has substantially changed due to significant development in molecular markers. In this regard, numerous studies have been conducted on molecular markers and their importance in marker-assisted breeding, genetic conservation and genetic diversity. This paper, therefore, aims to review research advances and prospects of molecular markers employed so far for sesame genetic diversity and conservation studies. In this section, we present a review on different molecular techniques which were developed and used for sesame genetic improvement. We also pinpoint the progress of different molecular markers targeted for genetic diversity studies, germplasm conservation, cultivar identification, determination of genetic variation, construction of genetic linkage maps, identification of QTLs/genes and discovery of molecular markers using freely available online databases. Finally, this review identifies hot topics and upcoming directions for future sesame improvement programs.

Keywords Accessions · Countries · Cluster · Genetic diversity · Germplasm conservation · Molecular markers · Sesame

Introduction

Sesame (*Sesamum indicum* L.) is one of the most ancient cultivated plants in the world renowned for its high content and quality of oil (Bedigian 2003; Yu et al. 2019; Zhang et al. 2013b). Sesame, a diploid species ($2n = 26$) with an estimated genome size of 369 Mb (Zhang et al. 2013b), is one of the most seasoned crops widely produced in Asia (54.9%), Africa (40.8%), and South and Central America (4.3%) (Weldemichael et al. 2020). Sesame seeds are abundant in oil (50–55%), proteins (18–20%), carbohydrate (13.4–25.0%), and digestible fiber (9.8%) (Makinde and Akinoso 2013). Besides, sesame seeds are rich in various vitamins and minerals, such as pantothenic acid (0.60%),

calcium (1.35%), and potassium (0.67%), as well as phosphorous, iron, zinc, magnesium, and vitamin B1 (Yuan et al. 2018). Sesame lignans are nutritious sources owing to high contents of unsaturated fatty acids and natural antioxidants like sesamolin and sesamin as well as small amounts of sesaminol, piperitol, sesamolinnol, pinoresinol, episesaminone, hydroxymatairesinol, and larisiresinol (Shittu et al. 2008). Sesame is also used as a component for the manufacture of soaps, perfumes, cosmetics, insecticides, and pharmaceutical products (Zhang et al. 2019). Moreover, it promotes health benefits providing relief from hypertension, oxidative stress, and neurodegenerative diseases (Nakano et al. 2002). Currently, the crop is cultivated in 146 countries of the world. Several poor and middle-income farmers depend on this important crop for their livelihoods.

Traditional crop improvement approaches have been used in the emerging oilseed crop improvement for several years to produce new genotypes with desired traits, but are very slow and laborious. This can, hence, no longer sustain the global oilseed production demands of the rapidly increasing

✉ Micheale Yifter Weldemichael
y.mickye@gmail.com; micheale.yifter1@mu.edu.et

[1] Department of Biotechnology, College of Dryland Agriculture and Natural Resources, Mekelle University, P.O. Box 231, Mekelle, Tigrai, Ethiopia

population and declining agricultural resources. The production and quality of oilseeds need to be improved at an accelerated rate in order to achieve the current and future demands. Hence, the demand for vegetable oil is growing globally, where by 2050, it is projected to reach 240 million tons (Barcelos et al. 2015). In 2018, the largest harvested area of sesame (3,480,000 ha) was in Sudan with total production per year of 981,000 tons, while the lowest cultivated area (210,000 ha) was in Uganda with a total yield of 140,000 tons (Table1; Fig 1).

Despite its numerous applications in economic, medicinal, and nutritional importance, little attention was given in the area of genomics for sesame genetic improvement (Dossa et al. 2017a). In recent years, improving sesame yield and resistance levels of sesame to various biotic and abiotic stresses remains the primary objectives in sesame breeding. So far, development of new genotypes with better tolerance to abiotic stresses was given a priority by sesame researchers (Dossa et al. 2016a). Several research groups were engaged in developing genomic resources that assist breeding for high yielding sesame varieties during the last couples of years. As a result, different molecular techniques including Random Amplified Polymorphic DNA (RAPD) (Bhat et al. 1999), Amplified Fragment Length Polymorphism (AFLP) (Laurentin and Karlovsky 2006), Simple Sequence Repeat (SSR), such as Inter-Simple Sequence Repeat (ISSR) (Kim et al. 2002), Expressed Sequence Tags-SSR (EST-SSR) (Badri et al. 2014; Li-Bin et al. 2008; Sehr et al. 2016; Wei et al. 2011), Complementary DNA-SSR (cDNA-SSR) (López‑Villavicencio et al. 2005; Spandana et al. 2012; Surapaneni et al. 2014), Genome Sequence-SSR (Dossa 2016b; Yu et al. 2017), Chloroplast SSR (cpSSR) (Sehr et al. 2016), Restriction site-Associated DNA sequencing (RAD-seq) (Wang et al. 2016; Wu et al. 2014a), Specific Length Amplified Fragment Sequencing (SLAF-seq) (Zhang et al. 2013d), RNA-Seq (Wei et al. 2014a), Whole-Genome Sequencing (WGS) (Wang et al. 2015; Wei et al. 2015; Zhang et al. 2016), and Genotyping by sequencing (GBS) (Uncu et al. 2016), were developed for sesame improvement. But still, further applications of advanced genomics tools are required for the biotic and abiotic stresses as well as inherent genetic problems, such as seed shattering, indeterminate flowering, and other desirable morphological traits. This review, thus, explores the various findings of

Table 1 List of top nine sesame producing countries: total area, production, and yield in 2018

Top producing countries	Area (ha)	Yield (Kgha^{-1})	Production (tha^{-1})
China	311,203	13,926	433,386
Ethiopia	415,112	7258	301,302
India	1,730,000	4321	746,000
Myanmar	1,463,447	5254	768,858
Nigeria	539,086	10,625	572,761
South Sudan	617,908	3342	206,522
Sudan	3,480,000	2819	981,000
Uganda	210,000	6667	140,000
The United Republic of Tanzania	800,000	7014	561,103

Source: Food and Agricultural Statistical Databases (FAOSTAT 2020)

Fig. 1 List of top nine sesame producing countries
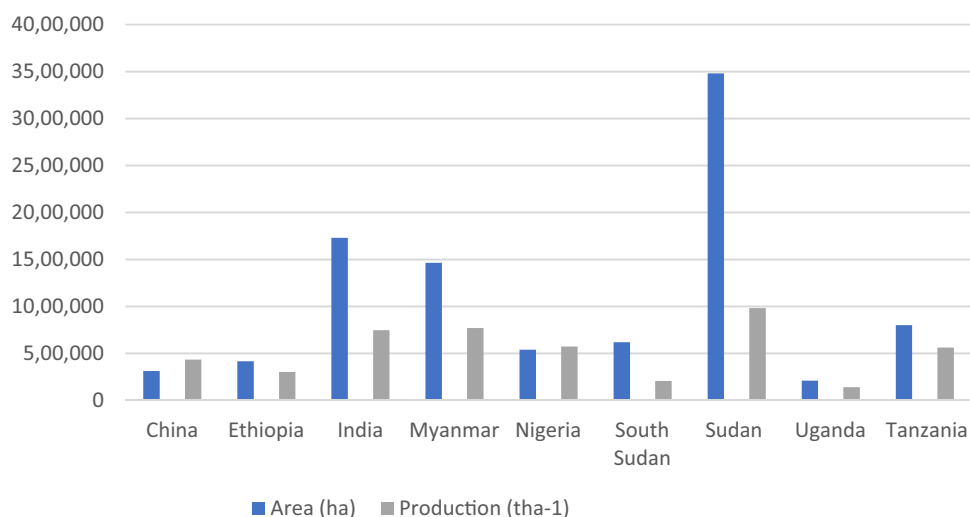
molecular markers for genetic diversity studies, construction of genetic linkage maps, identification QTLs/genes, and discovery of molecular markers using freely available online databases and finally forwards the possibilities of manipulating advanced molecular tools for future sesame improvement programs.

Genetic resources and core collection of sesame

Taxonomically, sesamum is related to *Utricularia gibba* as revealed by high-throughput genome sequencing and is estimated to be diverged from *U. gibba* approximately 98 million years ago. On the bases of chromosome number, 38 species of sesame have been classified into three groups, few of which are indicted below (Table 2). The five proposed centers for sesame genetic diversity include India, China, Central Asia, the Middle East, and Ethiopia (Zeven and Zhukovsky 1975). Great effort has been devoted to collect from and preserve sesame accessions in India, China, and Korea (Zhang et al. 2012a). The United States Department of Agriculture,

Agricultural Research Service, Plant Genetic Resources Conservation Unit (USDA, ARS, PGRCU), contains about 1226 accessions from Africa, Asia, Europe, North America, and South America (Morris 2009). About 6658 (Bisht et al. 1998) and 7698 (Park et al. 2015) sesame accessions were conserved in India and South Korea, respectively.

Moreover, several countries including China, Ethiopia, India, Israel, Kenya, Nigeria, South Korea, and the United States host gene banks for both wild and cultivated sesame (Table 2) via core collections, of which 453 genotypes are in China (Xiurong et al. 2000), 362 in India (Bisht et al. 1998), 287 in South Korea (Kang et al. 2006), and 870 in Ethiopia (Woldesenbet et al. 2015a). This is a huge germplasm, which is crucial for effective exploitation and utilization of genetic variation to facilitate future sesame improvement. However, utilization of this rich genetic resource for sesame improvement is unfortunately inadequate and the current diversity in the germplasm remains unknown. In this regard, the Asian sesame accessions have been well studied (Dossa et al. 2016b), but the sesame accessions from Africa, America, and Europe have been less preserved and characterized (Basak et al. 2019).

Table 2 List of some important *sesame* species with their chromosome number variation

Chromosome number	Species	Sources
2n = 26	*Sesamum indicum*	IPGRI (2004)
	S. latum	
	S. capense	
	S. malabaricum	
	S. mulayanum	
	S. schenckii	
	S. africanum	
2n = 32	S. *angustifolium*	
	S. laciniatum	
	S. latifolium	
	S. angolense	
	S. protratum	
2n = 64	*S. radiatum*	
	S. occidentale	
	S. schinzianum	
2n = Indeterminate	*S. abbreviatum Merxm*	
	S. calycinum Welw. ssp. calycinum	
	S. calycinum Welw. ssp. baumii (Stapf) Seidenst. ex. Ihlenf	
	S. calycinum Welw.ssp. pseudoangolense Seidenst ex.Ihlent	
	S. marlothii Engl	
	S. parviflorum Grabow-Seidenst	
	S. pedalioides Heirn	
	S. rigidum Peyr. ssp. rigidum	
	S. rigidum ssp. merenksyanum Ihlenf. & Seidenst	
	S. schinzianum Aschers. ex. Schinz	
	S. triphyllum Welw. ex. Aschers	
	S. triphyllum Welw. ex. Aschers. var. grandiflorum (Schinz) Merxm	

Concerning sesame core collection, it has been established by assessing the genetic diversities collected worldwide and fairly represent pre-selected core collection. These collections have been done to efficiently explore and utilize novel genetic variations (Hodgkin et al. 1995) and resulted in 362 Indian accessions (Bisht et al. 1998), 453 Chinese accessions (Zhang et al. 2012b), and 278 Korean accessions (Park et al. 2015). These core collections serve as valuable genetic resources for present and future sesame genetic improvement. However, sesame accessions from Africa have not been well characterized and conserved as compared to the Asian germplasm (Dossa et al. 2016b). Hence, extra efforts are required to collect indigenous accessions and wild relatives of sesame from Africa and constitute an extensive gene bank for their efficient conservation and exploitation in future improvement programs.

Molecular markers

Molecular markers are essential to explore highly diverse genotypes and relationships among populations so as to speed up breeding programs and identify target genes in several crops. They are used for increasing the genetic gain, minimizing the breeding cycles, and useful for molecular breeding (Varshney et al. 2014). Up to now, a number of different molecular markers have been used for genetic diversity study, genetic linkage map construction, molecular characterization, functional online database development, and identification of genomic regions (QTLs) for different traits associated with various biotic and abiotic factors. This review, hence, presents the detail description of these markers as indicated below:

Random amplified polymorphic DNA (RAPD) marker

RAPD marker has been employed for genetic diversity studies of different sesame accessions across many countries and resulted a significant level of polymorphism. So far, a study conducted on 58 accessions of sesame from 18 different states of India and four adjoining countries using 24 RAPD markers grouped 48 of the total accessions into six clusters, the remaining 10 accessions being highly diverse (Bhat et al. 1999). In another study, 35 out of 38 accessions of sesame from four different areas of Turkey were clustered into six groups using seven RAPD markers (Ercan et al. 2004). Genetic diversity of two commercial cultivars and seven lines collected from Venezuela were grouped into two major clusters using 12 RAPD markers (Salazar et al. 2006). Similarly, a set of 10 sesame germplasm obtained from diverse regions of Sudan were clustered into two main groups suing 10 RAPD markers (Abdellatef et al. 2008). In a comparative study of genetic diversity, 20 sesame genotypes from divers' geographical zones of India were clustered based on their geographical origin using 30 RAPD markers (Kumar and Sharma 2009). Similarly, 16 Indian genotypes of sesame collected from various geographical locations were grouped into two major clusters using 26 RAPD markers (Sharma et al. 2009). These findings revealed a significant discrimination power of the marker for the characterization of Indian sesame genotypes.

In another study, accessions of sesame from Cambodia and Vietnam were assessed and grouped into four clusters using 10 RAPD markers (Pham et al. 2009, 2011). In this study, a considerable amount of genetic diversity was reported in both the Vietnamese and Cambodian populations with a relatively high genetic diversity in populations from South Vietnam. Besides, two major clusters and six sub-groups were obtained from 20 Pakistani sesame accessions using 10 RAPD markers (Akbar et al. 2011). The authors reported some geographically distant accessions clustered together suggesting to be a result of large movement of Pakistani farmers to various regions moving sesame seeds for farming into their new geographical locations. Genetic diversity of 27 sesame accessions (13 Iranian genotypes from different locations and 14 exotic genotypes) with extensive geographical coverage was analyzed using 15 RADP markers and clustered into five groups (Tabatabaei et al. 2011). Compared with other studies conducted with RAPD markers, this finding reported higher variability than that of Indian, Southeast Asian, and Sudanese sesame genotypes (Abdellatef et al. 2008; Kumar and Sharma 2009; Pham et al. 2009). Besides, 47 accessions of sesame representing diverse agro-ecological zones of India were assessed and grouped into two major clusters using 22 RAPD markers (Dar et al. 2017). In their study, brown- and white-seeded varieties were reported to be clustered together, while black-seeded varieties were outside the cluster. Another study comprising nine sesame cultivars from Tamil Nadu (India) where clustered into two major groups using 10 RAPD markers (Anandan et al. 2018). In this study, the diversity among the available cultivars was suggested to be governed by origin of the cultivar and geographical condition of growing regions.

Various studies are done on RAPD-based techniques and significant levels of variability for accession of sesame from various countries and climatic conditions have been achieved, even with the use of limited set of primers. RAPD technique is, thus, a robust as well as time- and cost-effective technique valuable for germplasm conservation, cultivar identification, and determination of genetic variation and effective selection of parents in sesame-breeding programs.

Amplified fragment length polymorphism (AFLP)

AFLP, a PCR-based technique, is multi-locus fingerprinting approach that selectively amplifies a subset of restricted DNA fragments to create and compare unique fingerprints for desire genomes. Thus, the marker has been employed for the analysis of genetic diversity in several sesame accessions collected from different countries and agro-ecological zones (Table 3). Genetic relationship and diversity study of 32 accessions of sesame collected from the Venezuelan germplasm collection using 457 AFLP markers revealed a large genetic diversity within the germplasm collection, 25 of which were clustered in two major groups but without any association between genotypes and geographical origin (Laurentin and Karlovsky 2006). Moreover, 20 commercial cultivars of sesame collected from Venezuela were analyzed and grouped into four clusters using 339 AFLP markers (Laurentin and Karlovsky 2007). Based on previously studied AFLPs, genome diversity of 10 sesame accessions representing the majority of genome diversity in India, Sudan, Venezuela, and western Asia was analyzed using AFLP markers (Laurentin et al. 2008). In their finding, genetic analysis based on neutral markers was reported to be inaccurate predictor of the potential parental lines in sesame-breeding program. Similarly, 96 accessions of sesame from various parts of the glob were divided into two major clusters (eastern Asian origin and South Asian origin) using AFLP techniques (Ali et al. 2007). The similar pattern of AFLP in their finding reflected geographical origin and morphological characteristics of the accessions. In addition, 40 accession of *S. radiatum* (Table 2) gathered from diverse regions of Benin were analyzed using four AFLP markers and grouped into three clusters (Adéoti et al. 2011). Their findings indicated that the accessions from both northern and southern regions of the country clustered together indicated independent clustering of the accessions from their geographical origin, which was similar to the finding of (Laurentin and Karlovsky 2006) but in contradiction to the finding of (Adéoti et al. 2011).

Simple sequence repeat (SSR)

Simple sequence repeats also known as microsatellites are highly versatile, operationally flexible, and low-cost markers applied in an extensive range of important and appropriate field of sesame improvement (Table 3), such as genetic variability, gene/QTL mapping, markers-assisted breeding (MAB), genotype identification, and evolution study (Dossa et al. 2016b; Shi et al. 2014; Surapaneni et al. 2014). A study isolated 50 SSRs from improved library of sesame, 10 of which were polymorphic and used to determine genetic diversity of 16 sesame accessions (Dixit et al. 2005). In another study, 20 genotypes of sesame, commercially cultivated in diverse geographical regions of India, were grouped into four major clusters using 10 SSR markers with the average similarity coefficient of 0.67 as a cutoff value, providing a basic tool for variety characterization and establishment of genetic association among genotypes of sesame (Kumar and Sharma 2011). Similarly, genetic diversity, population structure, and phylogenetic relationship among 150 sesame accessions collected from 22 countries were studied using 16 SSR markers and grouped into three clusters (Cho et al. 2011). Cluster analyses of this study reported a multifaceted distribution pattern of accession with mean genetic similarity coefficient of 0.52 by population and 0.45 by accession. Molecular diversity among 60 accessions of sesame germplasm originating from diverse sources (8 states of India) was studied using five polymorphic SSR markers and revealed high level of polymorphism with only some primer pairs, while the majority of them were monomorphic in nature (Spandana et al. 2012). In a different study, 545 sesame accessions were systematically analyzed and clustered into three groups with 42 SSR primer pairs (Yue et al. 2012). Their study showed that the genetic bases of Chinese varieties were relatively narrow while those of introduced varieties were abundant, suggesting the importance of introducing foreign sesame genotypes.

Genetic diversity study of 130 Chinese accessions of sesame including four wild germplasm accessions, 44 landraces, and 82 cultivars using 88 polymorphic SSR markers showed the existence of five subgroups belonging to two main clusters (Wu et al. 2014b). This finding suggested that sesame landraces have broader genetic bases than improved cultivars, indicating the consequence of genetic drift or selection during breeding process. Using selective hybridization strategy and 95 mining ESTs of the NCBI database, 25 microsatellite markers were developed as new set of microsatellite markers, which were analyzed and screened for genetic diversity in a collection of 16 accessions from sesame germplasm (Badri et al. 2014). This finding showed five clusters with a varied number of accessions: cluster I (two accessions), cluster II (five accessions), cluster III (five accessions), cluster IV (two accessions), and cluster V (two accessions). Similarly, 31 accessions of sesame obtained from 12 countries were screened and grouped into two clusters using 218 polymorphic SSRs (Wei et al. 2014b). The first group included 16 accessions (five from China, eight from East and South Asia, and three from Africa), while the second group included 15 accessions (12 of which were from China and the other three accessions all from the United States). In this study, most samples from China and other countries were clustered into different groups, indicating Chinese sesame accessions were distantly related to those of other countries. Genetic diversity of 96 sesame accessions, collected from 22 countries scattered over six different geographic regions in Africa and Asia,

Table 3 Summary of molecular marker used for population structure and genetic diversity analysis in sesame

Source of collection	Number of accessions	Marker type	Marker size	Detected alleles	Number of alleles per locus	Genetic diversity	PIC	Sub populations identified	References
12 countries	31	SSR	218	NA	5.46	0.685	0.63	2	Wei et al. (2014b)
155 outside China and 390 from China	545	SSR	42	106	NA	0.645	0.41	3	Yue et al. (2012)
19 countries	129	SSR	23	157	8.5	0.41	0.65	8	Asekova et al. (2018
22 countries	150	SSR	16	121	7.6	0.46	0.42	3	Cho et al. (2011)
22 countries	277	SSR	14	158	11.3	NA	0.568	4	Park et al. (2014)
25 countries	94	ISSR	34	64.5	7.59	0.97	0.80	5	Kumar et al. (2012)
Africa and Asia (22 countries)	96	SSR	33	137	4.15	0.508	0.45	5	Dossa et al. (2016b)
Brazil	36	SSR	10	22	2	0.34	0.27	9	ARAÚJO et al. (2019)
China	96	SSR and InDel	44	113	2.6	0.37	0.31	2	Wei et al. (2016)
China	24	EST-SSR	50	NA	6.55	0.76	0.70	NA	Wei et al. (2011)
China	36	C DNA-SSR	59	127	2.15	0.76	NA	5	Wang et al. (2012a)
China	216	SSR, SRAP and AFLP	79	338	2	0.2092	0.25	2	Wei et al. (2013)
China	404	SRAP and SSR	14	126	9	0.24	0.39	2	Zhang et al. (2010)
China	453	SRAP and SSR	14	126	9	0.2218	0.347	9	Zhang et al. (2012c)
China	67	SRAP	21	NA	NA	0.910	0.071	3	Zhang et al. (2011)
China	130	SSR and InDel	88	325	3.69	0.432	0.36	2	Wu et al. (2014b)
China	151	SNP	47	94	0.690	0.55	0.316	3	Wei et al. (2017a)
		InDel	47	52	0.710	0.55	0.296	3	
		SSR	46	52	2.67	0.55	0.360	3	
East African (Kenya, Tanzania and Uganda)	46	ISSR	51	NA	NA	0.34	0.67	4	Nyongesa et al. (2013)
Ethiopia	100	SSR	27	NA	NA	0.30	0.25	2	Teklu et al. (2021)
Ethiopia	6	ISSR	4	36	9	0.34	NA	4	Admas Alemu and Tesfaye (2013)
Ethiopian	120	ISSR	6	44	7.33	0.24	0.37	2	Woldesenbet et al. (2015b)
Ethiopian	128	ISSR	7	89	12.6	0.39	0.43	3	Abate et al. (2015)
Ghana	25	SSR	38	410	19.52	0.91	0.91	5	Adu-Gyamfi et al. (2019)

Table 3 (continued)

Source of collection	Number of accessions	Marker type	Marker size	Detected alleles	Number of alleles per locus	Genetic diversity	PIC	Sub populations identified	References
Global	36	EST-SSR	155	108	2.45	0.47	0.39	4	Li-Bin et al. (2008)
Greece, Bulgaria, Italy and several Asian countries	37	EST-SSR	7	NA	NA	NA	0.82	4	Stavridou et al. (2021)
India	18	SSR	30	134	4.5	0.75	0.63	5	Bhattacharjee et al. (2020)
India	49	SSR	20	NA	3	NA	0.718	2	Yepuri et al. (2013)
India	60	EST-SSR	8	27	3.37	0.57	NA	2	Pandey et al. (2015)
India	20	ISSR	25	NA	12.45	0.78	0.78	5	Kumar and Sharma 2011)
		SSR	10	5	3.6	0.67	0.67	4	
India	20	RAPD	30	309 (24)	10.3	0. 619	0.16	NA	Kumar and Sharma (2009)
India	16	RAPD	26	194	7.5	0.22	0.78	2	Sharma et al. (2009)
		ISSR	17	163	9.5	0.29	0.71	2	
India	7	SSR	30	15	3.75	0.5	NA	2	Das et al. (2013)
India	60	SSR	5	13	2.6	NA	0.77	NA	Spandana et al. (2012)
India	47	RAPD	22	191	11.6	0.09	0.186	2	Dar et al. (2017)
		SSR	18	64	3.55	0.07	0.194	2	
India	9	RAPD	10	102	10.2	0.45	NA	3	Anandan et al. (2018)
		SSR	10	19	1.9	0.42	NA	2	
India	18	SSR	30	134	4.5	0.967	0.63	5	Bhattacharjee et al. (2020)
India	30	SSR	21	93	4.42	0.83	0.43	5	Bhattacharjee et al. (2019)
Iran and other Asian countries	24	ISSR	13	170	7–19			7	Parsaeian et al. (2011)
Korea	16	SSR	10	46	4.6	0.69	0.66	NA	Dixit et al. (2005)
Korea and different parts of the world	139	SSR	23	157	6.5	0.15	0.59	2	Asekova et al. (2018)
Mediterranean sesame core collections	32	InDel	16	NA	NA	0.34	0.33	2	Kizil et al. (2020)
Morocco	31	ISSR	24	NA	NA	0.870	0.18	2	El Harfi et al. (2021)
Nigeria	22	SSR	12	47	3.92	0.62	0.56	2	Oduoye et al. (2020)
Ugandan	121	CpSSR	44	9	3.12	0.75	0.41	4	Sehr et al. (2016)

PIC polymorphic information content, *NA* data not available

was studied using 33 polymorphic SSR markers and revealed more diversity from the Asian accessions compared to the African accessions (Dossa et al. 2016b). Molecular variance analysis of this study showed more than 44% of genetic variation due to diversity among the geographic regions. In addition, another study constructed a comprehensive physical map from 151 published genomic sequences along with more than 83,135 non-redundant SSRs supplied with their physical positions and characteristics (Dossa 2016). Besides to its provision of the genomic distribution of 151 genomic sequences including functional genes, QTLs, and markers linked to some important sesame traits, this map sheds light on certain functional genomic regions that need future in-depth research to effectively locate the functional genes. This would assist future sesame research in fine mapping of more functional genes based on the available functional markers and QTLs.

Genetic diversity of 47 different sesame accessions assessed with 18 SSR markers (Dar et al. 2017) and nine sesame cultivars assessed with 10 SSR markers (Anandan et al. 2018) was grouped into two and three major clusters, respectively. The genetic diversity study among 129 sesame cultivars and landraces from Korea and various parts of the glob were studied using 23 informative SSR markers and revealed two major and six minor clusters (Asekova et al. 2018). In this study, clustering of genotypes was independent to their geographic origin, showing substantial gene flow among the genotypes across the selected geographic regions. Moreover, a high degree of variation was reported among individuals within (85.84%) than among the populations (14.16%). Similarly, a high variance was reported among the individuals within (90.45%) than between the countries of origins.

Recently, 38 SSR marker-based dendrogram analysis clustered 25 sesame landraces collected from five districts of northern Ghana into two main clusters: cluster A sub-divided into four sub-clusters and cluster B sub-divided into three sub-clusters (Adu-Gyamfi et al. 2019). Other genetic diversity studies reported five clusters from 30 diverse genotypes from different parts of West Bengal and 18 elite Indian sesame varieties analyzed using 21 polymorphic and 30 sesame-specific SSR markers, respectively (Bhattacharjee et al. 2019, 2020). Several authors reported grouping of genotypes belonging to different geographical origin in the same cluster, showing independent clustering of genotypes to their geographical origin (Asekova et al. 2018; Bhattacharjee et al. 2019, 2020; Kim et al. 2002; Laurentin and Karlovsky 2006; Pandey et al. 2015; Parsaeian et al. 2011). This contradicts to the usual assumption that selection of genotypes from different geographical origin maximizes the chance of producing segregation leading to diversity. Analysis of genetic diversity of 36 sesame accessions using 10 SSR markers revealed nine groups with the dissimilarity measures varying from 2.49 to 9.0, suggesting a narrow genetic basis among the evaluated 36 accessions (ARAÚJO et al. 2019). In a very recent study, 22 sesame germplasm accessions collected from two ecological zones of Nigeria were studied using 12 SSR markers (ODUOYE et al. 2020) (Table 4). This study revealed three major clusters and one outlier, where Savana was a more diverse ecological zone with 13 accessions and the humid forest was less diverse with nine accessions. More recently, 100 sesame accessions collected from Tigray, Afar, Amhara, Gambela, and Oromiya regions of Ethiopia were studied using 27 SSR markers, of which 63 accessions were allocated to the four populations, whereas 37 accessions were admixtures with no specific membership (Teklu et al. 2021). Their findings have reported various population clusters: population I (24 accessions), population II (13 accessions), population III (nine accessions), and population IV (17 accessions).

Table 4 List of functional online databases for sesame genomics

Database name	Website	Authors
GinMicrosatDb	http://backwin.cabgrid.res.in:8080/Gingelly7	Purru et al. (2018)
NCBI*	http://www.ncbi.nlm.nih.gov/genome/?term=sesame	–
ocsESTdb	http://www.ocri-genomics.org/ocsESTdb/index.html	Ke et al. (2015)
PMDBase	http://www.sesame-bioinfo.org/PMDBase	Yu et al. (2017)
PSMD	http://big.cdu.edu.cn/psmd/	Du et al. (2020)
PTGBase	http://www.ocri-genomics.org/PTGBase/index.html	Yu et al. (2015)
Sesame Germplasm Resource Information Database	http://www.sesame-bioinfo.org/phenotype/index.html	–
SesameFG	http://www.ncgr.ac.cn/SesameFG	Wei et al. (2017b)
SesameHapMap	http://202.127.18.228/SesameHapMap/	Wei et al. (2015)
Sinbase	http://www.ocri-genomics.org/Sinbase/index.html	Wang et al. (2015)
SisatBase	http://www.sesame-bioinfo.org/SisatBase/	Dossa et al. (2017b)
The Sesame Genome Project	http://www.sesamegenome.org	Zhang et al. (2013b)

Besides to the well-established application of SSR markers for the study of genetic diversity, they represent a suitable type marker for the construction of genetic map and development of databases to search for sesame SSR information at the whole-genome level and perform functional analysis (Tables 4 and 5). Thus, SSR markers have been employed for the development of freely available and accessible web-based databases, such as Sinbase that provides a useful search function and data mining tools (Wang et al. 2015), SisatBase that gives free access to microsatellite marker data and a unified platform for functional analysis (Dossa et al. 2017b), PMDBase that facilitates researchers to identify microsatellite (Du et al. 2020). These functional databases would assist the beneficial resources for genomic studies, genetic assessment, and breeding advancement in sesame. Moreover, based on an F_2 segregating population of an intra-specific cross between two cultivars, the first sesame linkage map was constructed comprising 220 molecular markers with a total genetic length of 936.72 cM, and average distance of 4.93 cM between markers on 30 LGs (Wei et al. 2009). This has opened a door for the construction of other linkage maps and increase its density using different types of markers. As a result, a high-density linkage map containing 653 (30 SSR, SNP, 50 AFLP and 573 RSAMPL) marker loci distributed on 14 LGs covering a total length of 1,216 cM and 1.86 cM between adjacent markers was constructed and mapped a QTL for seed coat color (Zhang et al. 2013c). These findings provided a foundation for future improvement of sesame MAB.

Inter-simple sequence repeat (ISSR)

ISSR markers have been established and employed mainly for genetic diversity studies of sesame during the last couples of decades. About 75 *Sesamum indicum* accessions obtained from divers areas in Korea (26 cultivars, 17 breeding lines and landraces, and 32 introductions from different countries across the world) were characterized using ISSR primers and clustered into seven groups based on unweighted pair-group method with arithmetic averages (UPGMA) cluster analysis (Kim et al. 2002). Another ISSR-based dendrogram analysis classified 16 sesame genotypes collected from different geographical locations of India into two clusters, showing maximum discrimination power of the marker for the characterization of sesame genotypes (Sharma et al. 2009). In another study, 20 commercially grown sesame genotypes from various agro-climatic regions of India were grouped into five major clusters using 25 ISSR primers with an average similarity of 0.78 as a cutoff value, suggesting these markers provide a powerful tool for genotype identification and establishment of genetic relationships among sesame genotypes (Kumar and Sharma 2011). In a different study, a set of 34 confirmed polymorphic ISSR primers were reported

to cluster 94 accessions of sesame from 25 countries into six major groups based on their geographical origin, viz., India sub-continent (50), Central Asia (11), America (12), East Asia (eight), Europe (two), and Africa (11) (Kumar et al. 2012). Besides, genetic diversity study of 18 sesame genotypes obtained from various regions of Iran and six accessions of sesame from other Asian countries revealed a cluster of seven groups using ISSR primers (Parsaeian et al. 2011). According to this study, genetic diversity-based selection of parental lines for hybridization is better than their geographical distribution.

In Ethiopia, UPGMA cluster analysis of 6 varieties and farmers' landraces from north-western Ethiopia grouped each individual of the respective populations in separate groups using four ISSR primers (Admas Alemu and Tesfaye 2013). Similarly, the genetic diversity study of 128 genotypes of sesame (comprising 119 land races and nine commercial varieties) collected from 20 Administrative Regions and Zones of Ethiopia resulted three major groups sub-grouped into nine clusters using 17 ISSR primers (Abate et al. 2015). In this study, most of the landraces were grouped in separate clusters irrespective of their geographic origins, while the cultivars were grouped in one cluster, suggesting less variability within the released varieties than the landraces. Another study also revealed 120 (38 exotic and 82 Ethiopian) sesame accessions grouped into two main clusters and three outliers using six ISSR primers (Woldesenbet et al. 2015b). This study showed weak association between ecological region and genetic variation of sesame genotypes, and less diversity of exotic accessions than Ethiopian ones. In addition, 10 genotypes of sesame obtained from various regions of the country were characterized using 11 ISSR primers and clustered into four groups (Gebru et al. 2019). More recently, 31 genotypes of sesame collected from various locations of Tadla area in Morocco were characterized using 24 ISSR primers and grouped into three large groups and seven independent branches (El Harfi et al. 2021). This result indicated that Moroccan sesame populations are characterized by a low genetic diversity. This indicates the need to either create a new germplasm through mutagenesis breeding or introduce from diverse geographical origins in order to broaden the existing genetic variability. More recently, genetic diversity of various cultivated sesame genotypes was studied in Saudi Arabia using karyological, physiological, biochemical and molecular markers and were grouped into two main clusters (ALshamrani et al. 2022).

Expressed sequence tags-SSR (EST-SSR)

Advancement of EST-SSR markers using publicly available sesame EST data has speeded up the application of molecular markers in sesame improvement (Table 3). In this regard, Illumina paired-end sequencing identified 148

Table 5 Summary of genetic and association mapping studies for different traits of sesame

Traits	Marker system	Parents of the cross	Marker number	Population	Major results	References
Seed coat color, seed size and thousands of seed weight	SLAF-seq	Ganzhi 6 (black-seeded) × Gaoyou 8 (white-seeded)	2159	122 F2	14 QTLs, 14 QTLs, and 3 QTLs linked to seed coat color, seed size and TSW respectively	Du et al. (2019)
Determinate growth habit	SNP	Zhongzhi No. 13 (indeterminate genotype) × Det-4247 (determinate genotype)	13,679	232 F2	Determinacy gene located on the up end of LG09 out of 13	Zhang et al. (2018)
Number of flowers per axial and branching habit	SLAF-seq	Yuzhi 4 × Bengal Small-seed	9378	150 BC1	Identified SiFA and SiBH for number of flowers per axil and basal branching habit, respectively	Mei et al. (2017)
Charcoal rot disease	SSR	ZZM2748 × Zhongzhi No. 13	498	548 RIL	14 novel QTLs linked to charcoal rot disease resistance mapped distributed into 13LG	Wang et al. (2017)
Plant height and seed coat color	SNP	Zhongzhi No.13 × ZZM2748 (semidwarf)	11,924	430 RIL	41 QTLs linked to plant height and 9 QTLs linked to seed coat color were mapped with 13 LGs	Wang et al. (2016)
Seed coat color	SSR	Zhongzhi No.13 (white seed) × Mishuozhima (Black seed)	400	550 F6	1 major QTL with SIN_1016759, a homologous of PPO gene related to black seed coat color	Wei et al. (2016)
Determinant growth habit	SNP	Yuzhi 11 (indeterminate growth) × YuzhiDS899 (determinate growth)	30,193	120 F2	Determinacy gene SiDt identified	Zhang et al. (2016)
56 agronomic traits	SNP	Worldwide sesame accessions	1,800,000	705	Identified 549 associated loci and 46 causative genes linked to key agronomic traits	Wei et al. (2015)
Oil and protein content	SSR	Worldwide sesame accessions	112	369	Detected 19 SSR for oil content and 24 SSR for protein content	Li et al. (2014)
Grain yield traits	SNP	Miaoqianzhima × Zhongzhi 14	1230	224 RIL	Detected 13 QTLs on 7 LGs and 17 QTLs on 10 LGs for seven grain yield-related traits	Wu et al. (2014a)
Seed quality traits	SRAP, SSR and AFLP	Chinese collection	79	216	Identified 10 markers associated with oil, oleic acid, linoleic acid and protein	Wei et al. (2013)
Seed coat color	SSR, AFLP and SAMPL	COI1134 (white-seeded) × RXBS (black-seeded)	653	260 F2	Detected 4 QTLs for seed coat color (QTL1-1, QTL11-1, QTL11-2 and QTL13-1)	Zhang et al. (2013a)

SSR sequences consisting of 155 EST-SSRs from a total of 1785 non-redundant EST sets assembled among 3328 identified sesame ESTs (Li-Bin et al. 2008). This study revealed 27 (61.4%) out of 44 successfully amplified EST-SSRs to be polymorphic in 36 globally collected sesame accessions and classified them into four groups based on UPGMA dendrogram analysis. Similarly, Illumina paired-end sequenced transcriptomes from five sesame tissues converted a total of 7702 uni-genes into SSR markers (EST-SSR) dominant by dinucleotide SSR repeat motifs (67.07%) (Wei et al. 2011). In this study, 40 primer pairs out of randomly selected 50 EST-SSRs amplified DNA fragments successfully and detected high level of polymorphism among 24 sesame accessions. In another study, 20 EST-derived SSRs were used to study the diversity of 49 sesame accessions and grouped them into five separate clusters exhibiting a genetic similarity coefficient ranging from 0.59 to 1.0 (Yepuri et al. 2013). These identified EST-derived SSR markers could be valuable in assessing the diversity of sesame accessions and in identifying various parents for sesame improvement programs. Moreover, 60 globally collected genotypes were evaluated in India and clustered into two major groups using eight polymorphic EST-SSR markers, where genotypes belonging to the same geographical area did not always occupy the same cluster (Pandey et al. 2015). More recently, 35 sesame landraces collected from Bulgaria, Greece, Italy and several Asian countries were studied using seven EST-SSR markers and clustered into three clusters: cluster I (12 cultivars), cluster II (9 cultivars) and cluster III (16 cultivars) (Stavridou et al. 2021). The molecular and morphological analysis in this study clearly showed that sesame cultivars collected from similar geographical origin were not always clustered in the same group, making heterotic clusters: nevertheless, grouping patters were detected for the Greek landraces. These EST-derived SSRs could be valuable to study the relationship among sesame accessions and help identify distinct parents for further sesame improvement programs.

Complementary DNA-SSR (cDNA-SSR)

Complementary DNA-SSRs are markers directly associated with transcribed genes and which give some fundamental advantages over genomic SSRs in lowering cost of development and significant level of transferability to related species (Wang et al. 2010). Hence, 59 cDNA-SSR markers were developed based on unigenes to genotypes to investigate the genetic variation among 36 sesame accessions (Wang et al. 2012a). Such characterized and identified polymorphic cDNA-SSR markers will greatly facilitate genetic structure, conservation and identification of sesame.

Chloroplast SSR (cpSSR)

Chloroplast-SSRs are newly developed markers, which have become a marker of choice for the study of genetic diversity in several crops including sesame. In sesame, very few genetic studies have been reported using cpSSRs (Table 3). Seven large repeated loci and 28 different SSR loci were detected in the chloroplast genome of *S. indicum*, of which 21 loci were homopolymers, four di-polymers, and three tri-polymers (Yi and Kim 2012). This study also identified seven large repeated loci in the S*esamum indicum* chloroplast genome which are useful for developing sesame-specific chloroplast genome vectors. From 121 Ugandan accessions of sesame, four haplotypes were developed, of which 96% of the individuals belonged to haplotype B (Sehr et al. 2016). In addition, their findings reported 117 sesame landraces grouped in 13 districts based on nine SSR markers.

Recently developed molecular markers

Insertions and deletions (InDels)

InDels have been developed and used in sesame genetic analysis for their high abundance, high polymorphism, and reproducibility in the sesame genome. Accordingly, a comprehensive characterization of molecular and genetic variability of 130 sesame accessions from Chinese was evaluated using 88 polymorphic markers and revealed two main groups with five subgroups (Wu et al. 2014b). The rate of polymorphism (48.0%) for the InDels in this study exhibited better capacity to discern genetic diversity that contradicts to the work of (Wang et al. 2012a), 3.8%. The authors also developed 79 InDels from the cultivar Zhongzhi 14 using reduced-representation sequencing method. A different study reported the detection of 17,961 small InDels using a parallel re-sequencing of 29 strains of sesame obtained from 12 different countries and found consistent phylogenetic relationships among the Indian sesame accessions (Wang et al. 2014a). This study provides an insight into genetic dissimilarity of the sesame germplasm and facilitates a broad range of evolutionary and functional studies, such as population genetics, genome evolution, gene identification, and marker-assisted breeding. In addition, 45 InDels in combination with 259 SNPs and 124 SSR markers constructed a new genetic map spanning 1790.08 cM using 107 F_2 lines derived from Zhongzhi No. 13 (ZZM2289) population (Wang et al. 2014b). These authors suggested an association of high genetic variation between lipid-related genes and oil content from the re-sequenced data of 29 sesame accessions collected from 12 countries. Recently, the first successful 7477 InDel sites were discovered using ddRADSeq data from 95 accessions belonging to Mediterranean sesame core collection (Kizil et al. 2020). The InDel frequency (1 per 37.74 kb)

in this study was much higher than that of (Wei et al. 2014a), 1 per 137 kb, who used transcriptome assembly for InDel detection. Moreover, more InDels (7477) were found in this study as compared to the work of (Wu et al. 2014b), 75 genomic InDel markers, who used restriction site-associated DNA (RAD) sequencing of 'Miaoqianzhima' and 'Zhongzhi 14'. This study also revealed no clear pattern with respect to the geographical origin, which is in agreement with the work of (Laurentin and Karlovsky 2006), where no association was reported between accession origin and genetic differentiation in sesame. These differences, hence, could be a consequence of different factors, such as the number of genotypes used for genotyping, the sequencing method, and the bioinformatics tools for the exploration of variants.

Restriction site-associated DNA sequencing (RAD-Seq)

Next-generation sequencing (NGS) techniques have recently been used for genome sequencing due to fast and cost-efficient development of sesame-specific DNA markers. RAD-seq is used for sequencing genomic regions flanking specific restriction enzyme sites and develop a reduced representation of genome (Wang et al. 2012b). Besides, RAD-seq is used for genotyping and discovery of SNPs in sesame (Wang et al. 2016; Wu et al. 2014b). With this regard, RAD-seq has developed 3,804 pairs of new DNA markers including SNPs (3769) and InDels (35), which were used to construct the most high-density genetic map in combination with SSR markers using a RIL population derived from a cross between 'Zhongzhi 14' and 'Miaoqianzhima' sesame genotypes (Wu et al. 2014b). This map was about 844.46 cM in length containing 1,230 markers (18 InDels, 22 SSRs and 1190 SNPs) distributed on 14 LGs with 0.69 cM average distance between adjacent markers. This study also identified a total of 30 QTLs for seven grain yield-related traits and was the first QTL map for yield-related traits using a RIL population. By adding a new RAD-seq, the first bin map was generated containing 1522 bins covering a genetic distance of 1090.99 cM with 0.72 cM between adjacent bins on 13 LGs and identified 50 QTLs: 9 for seed coat color and 41 for sesame plant height (Wang et al. 2016). This was the first genetic map reported in sesame to have the same number of linkage groups as the number of chromosomes in sesame. The bin distance (0.72 cM) reported in this study was in line to that of (Wu et al. 2014b), 0.69 cM generated by RAD-seq, but shorter than the maps constructed using randomly selective amplification markers (1.86 cM) (Zhang et al. 2013c) or the SLAF-seq technology (1.20 cM) (Zhang et al. 2013d). Moreover, more than 99% of the interval distances between adjacent bins were shorter than 6 cM, representing high density of the new sesame genetic map. The aforementioned findings reported high-density genetic map that would be an ideal for comparative genomic analysis,

map-based cloning for yield-related traits, QTL/gene mapping and applying marker-assisted selection (MAS) for sesame genetic improvement.

Specific length amplified fragment sequencing (SLAF-Seq)

SLAF-seq is a recently developed high-resolution strategy used for large-scale genotyping and discovery of new SNPs (Table 5). SLAF-seq was employed to detect a total of 71,793 high-quality SLAFs, of which 1272 of the 3673 polymorphic SLAFs were used to construct sesame genetic map (Zhang et al. 2013d). This study developed a dense genetic linkage map of sesame containing 1,233 markers on 15 LGs spanning a distance of 1474.87 cM with 1.20 cM between adjacent markers using SLAF-seq, of which 1079 were SNP_only, 86 were InDels_only, and 68 were SNP & InDel markers. In addition, the study reported segregation distortion of 205 markers including 171 'SNP_only'-type markers, 20 'Indels_only'-type markers, and 14 'SNP & Indels' markers with percentages of 83.41%, 9.76%, and 6.83%, respectively. Another study generated a sequence data of 81.2 Gb containing 902.36 million per end reads using SLAF-seq on 366 germplasm of sesame accessions collected from China (329) and 11 other countries (37) (Cui et al. 2017). Similarly, a high-resolution genetic map containing 9,378 SLAF on 13 LGs was constructed, spanning 1974.23 cM in length with mean length of 151.86 cM and distance of 0.22 cM between adjacent markers using SLAF-seq (Mei et al. 2017). The authors reported mapping of genes for basal branching habit (*SiBH*) and flowers per leaf axil (*SiFA*) on LG 5 and LG11, respectively and suggested their finding to serve as a reference linkage map to anchor sequence scaffolds of the physical map.

More recently, a SLAF-seq constructed a high-density SNP linkage map of 2128.51 cM in length, comprising 2159 SNPs distributed on 13 LGs with an average distance of 0.99 cM between adjacent markers (Du et al. 2019). This genetic map contained the same number of linkage groups as the number sesame chromosomes which agrees to (Wang et al. 2016) using RAD-seq. The authors reported that this was the first QTL map for seed-related traits in sesame using an F_2 population, revealing 17 important seed-related traits and 19 major effect QTLs (eight QTLs for seed coat color, nine QTLs for seed size and two QTLs for thousand seed weight).

RNA- sequencing (RNA-Seq)

The advancement of largescale RNA-seq has led to generate large expression datasets in non-model species due to its high speed, cost-effectiveness and reliability (Marioni et al. 2008). In a comprehensive transcriptome study of sesame samples from different developmental and growth

stages, RNA-seq identified 2164 genic-SSR markers in 4440 SSR-containing sequences (Zhang et al. 2012a). Out of the identified SSR markers in this study, primer pairs of 300 SSRs were selected randomly and validated in 25 accessions of sesame, 92.0% (276) of which successfully yielded PCR amplification. Besides, the genetic similarity coefficients of this study showed low correlation between geographical source and genotype as well as narrow genetic basis of sesame in China, which agrees with other findings reported using SSR markers (Wu et al. 2014b; Yue et al. 2012). Similarly, screening 29 sesame accessions of variable oil content from 12 countries generated a new high-density genetic map with more than 120 Gb clean data and identified 2,348,008 SNPs (Wang et al. 2014b). In a different study, a total of 7450 SNPs and 362 InDels were identified from a total of 33.47 Gbp from three sesame transcriptome datasets using RNA-seq (Wei et al. 2014a). The newly developed polymorphic markers (21 SNPs and 16 InDels) were used to evaluate the genetic diversity of 39 sesame cultivars and classify them into four groups with a similarity value of 0.636. These new genic-SSR and SNP/InDel markers will provide a useful tool for sesame genetic research and comparative genome analysis so as to increase the identification of candidate functional genes and the efficiency of MAS.

Whole-genome sequencing (WGS)

Advanced genetic and genomic research improvment in sesame is possible with understating of whole-genome sequencing technique. WGS has been widely applied in SNP discovery and genotypic studies in sesame (Table 4). Accordingly, potential candidate genes were identified for different desirable traits including oil quality and content, yield components, fatty acid composition, morphological characteristics, growth cycle, and coloration and disease resistance putative candidate genes (Wang et al. 2015; Wei et al. 2015; Zhang et al. 2016). These will help to undertake further research activities on functional genomics, such as genome editing using RNA interference (RNAi), clustered regularly interspaced short palindromic repeats)/Cas9 technology (CRISPR/Cas9), transcription activator-like effector nuclease (TALEN), and zinc finger nuclease (ZFNs), as well as genetic transformation. WGS of 29 sesame strains generated more than 120 Gb filtered data and identified an average of 127,347 SNPs (Wang et al. 2014b). This study resulted two major groups based on monoflower (13) and triple-flower (16), where 695 genes with equivalent SNPs between the two pools were identified. Besides, the authors expected the number of flowers per leaf axil to be determined by adenylate iso-pentenyl-transferase (ITP) genes in sesame, suggesting a novel role of the gene. Moreover, WGS has been used to develop various comprehensive and functional online user friendly databases, such as SesameFG

(Wei et al. 2017b), SesameHapMap (Wei et al. 2015), Sinbase (Wang et al. 2014a), and the Sesame Genome Project (Zhang et al. 2013b), which could play vital roles to study functional and comparative genetic studies of desirable traits in sesame. Similarly, another freely online available database, 'GinMicrosatDb', has been developed using WGS data of 'Swetha' sesame variety consisting of microsatellites localized on both LGs and scaffold with their genomic co-ordinates (Purru et al. 2018). These databases facilitate selection of unlinked markers across linkage groups and ensured the random distribution by reducing the biases in estimation of genetic diversity and relatedness of genotypes in sesame. Overall, such freely online available databases can be easily accessed by researcher and prove useful search function and mining tools including genome browser and local BLAST service.

With regard to SNP-based genetic diversity analysis in sesame, a large number of high-throughput SNP markers have been discovered using recently developed molecular markers (Table 6) including 362 InDels and 7450 SNPs using RNA-seq (Wei et al. 2014a), 15,693 SNPs using RAD-seq (Wang et al. 2016; Wu et al. 2014a), more than 15,000 SNPs using GBS (Uncu et al. 2016), and 93,053 SNPs using SLAF-seq (Cui et al. 2017; Du et al. 2019; Mei et al. 2017) from various datasets and varieties. Markers based on high-density SNPs are highly favored in the study of genetic diversity and population structure, high-density genetic linkage map construction and genome-wide association study of sesame accessions (Table 5). SNP genotyping is a useful tool for mapping QTLs/genes, map-based cloning and MAS in several crops (Hayashi et al. 2006). The genetic structure of 366 sesame germplasm accessions collected from 18 provinces of China and other 11 countries of the world were clustered into three sub-groups (Pop1, Pop2 and mixed) using 89,924 SNPs (Cui et al. 2017). This clustering was primarily along the bases of their geographic collection, which is in contradiction to majority of other findings (Asekova et al. 2018; Bhattacharjee et al. 2019, 2020; Kim et al. 2002; Laurentin and Karlovsky 2006; Pandey et al. 2015; Parsaeian et al. 2011). In another study, combinations of different markers were reported to provide a useful protocol for the identification and authentication of various sesame cultivars (Wei et al. 2017a). In their study, a combination of 140 polymorphic markers (47 SNPs, 47 InDels and 46 SSRs) were employed for DNA fingerprinting of 151 Chinese sesame cultivars and classified them into three main groups. This revealed a narrow range of genetic variation among the cultivars and were consistent with their geographical distribution (Wei et al. 2017a), which agrees with other findings based on using SSRs (Wu et al. 2014b; Yue et al. 2012) and RNA-seq (Zhang et al. 2012a). Recently, 198 sesame accessions from 21 geographical regions spread over four different continents were analyzed using 5292 high-quality SNPs

Table 6 Summary of SNP marker-based population structure and genetic diversity analysis in sesame

Number of accessions	Sources of accessions	Number of effective SNPs	SNPs detection approach	Average marker density/SNP	Average nucleotide diversity of the panel	Genetic distance	Average gene diversity	Number of subgroups identified	References
29	China gene bank	127,347	Whole-genome sequencing	NA	1.5×10^{-4}	NA	NA	NA	Wang et al. (2014b)
705	China gene bank	254,781	Whole-genome sequencing	1SNP/50 bp	2.4×10^{-4}	0.02	NA	2	Wei et al. (2015)
336	HSRC-HAAS	89,924	SLAF-seq	1SNP/2.6 kb	1.1×10^{-3}	0.01 to 0.42	0.17	3	Cui et al. (2017)
95	Mediterranean sesame core collection	5,292	ddRAD	46SNP/kb	NA	0.023 to 0.524	0.28	3	Basak et al. (2019)

and were clustered into three sub-groups (Basak et al. 2019). This study showed about 5.53% genetic variability among continental groups of the total variation, indicating highly diversification of the Mediterranean sesame core collections.

In a high-throughput SNP identification and genotypic study, a total of 15,521 SNPs were identified from 91 sesame RILs derived from a cross Acc. No. 95-223 (Africa) × Acc. No. 92-3091 (Korea), 14,786 SNPs (95.26%) of which were located along sesame genome assembly pseudomolecules (Uncu et al. 2016). Consequently, 770 SNPs selected through a high stringency filtering protocol in conjunction with 50 SSRs were used in linkage analysis, and resulted in 13 LGs encompassing a total genetic distance of 914 cM with 432 markers (420 SNPs and 12 SSRs). Besides, an ultra-dense SNP map comprising 30,193 SNPs in 13 LGs with an average marker density of 0.10 cM was constructed by resequencing 120 progeny of an F_2 population derived from the cross Yuzhi 11 (indeterminate) × Yuzhi DS899 (determinate) (Zhang et al. 2016). This study identified the target SNP (SiDt27-1) and the determinacy gene (*DS899s00170.023*, *SiDt*) in Scaffold 00170 of the Yuzhi 11 reference genome. Another study also reported a high-density bin map with 1086.403 cM in length, containing 13,679 SNPs with an average distance of 0.918 cM between adjacent bins using GBS in an F_2 sesame population (Zhang et al. 2018). This study reported a novel sesame determinacy gene mapped on LG09 in a genome region of 41 kb which could be a base for developing of sesame cultivars suited for mechanization. More recently, 300 genotypes comprising 75 exotic collections from different Asian and African countries, 209 Ethiopian landraces, and 16 varieties were studied using Diversity Arrays Technology (DArT) and SNP markers (Tesfaye et al. 2022). In this study, a total of 6474 SNP markers and 6115 silicoDArT markers were discovered,

of which 5821 and 5065 markers of SNP and DArT markers, respectively were aligned with the reference sesame genome.

Concluding remarks and future perspectives

Sesame is one of the oldest oilseed crops widely cultivated in the tropics and subtropics of the world. The production and productivity of sesame are, however, highly affected by both biotic and abiotic stresses. So far, traditional crop improvement approaches have been employed in sesame improvement for many years to produce new genotypes with desired traits, but are very slow and laborious. To minimize the problems of traditional breeding, several advanced molecular markers, such as RAPD, AFLP, SSR, ISSR, EST-SSR, cDNA-SSR, cpDNA, InDel, RAD-seq, SLAF-seq, RNA-seq, and WGS, have been developed and made available genetic improvement of sesame. Such molecular markers are valuable and provide a foundation for cultivar identification and conservation of the species, genetic diversity studies, linkage mapping, qualitative and quantitative trait mapping, whole-genome sequence data, transcriptomes-based research, marker-assisted breeding and transferability studies among sesame accessions. This review presented the application of molecular markers on population structure and genetic diversity studies as well as construction of genetic linkage maps and the discovery of molecular markers for freely available online databases using different approaches. Understanding the population structure and diversity of sesame collection is essential for efficient management of genetic resources to improve sesame breeding. Availability of a dense genetic map would also be useful for fine mapping of genes/QTLs, map-based gene identification, and advanced molecular breeding for sesame and would act

as a reference for positioning sequence scaffolds on a physical map, thereby assisting the assembly of sesame genome sequences. Moreover, the freely online available systematic data bases and integrative analytical web services help to enhance molecular markers application for genome evolution and structure variants, gene regulation, and future advanced sesame breeding. Hence, the following are future prospective for further sesame improvement:

1. Africa is the second largest producer of sesame next to Asia but the genetic diversities of the African sesame accessions are not well studied. There is, hence, a need to conduct further genetic diversity studies using recently developed molecular markers in the major sesame producing African countries including Ethiopia, Nigeria, South Sudan, Sudan, Uganda, and the United Republic of Tanzania.

2. Collecting cultivars in various growing areas of African countries with the intention of establishing sesame gene banks are urgently needed.

3. Locally adapted and wild species of sesame are rich in genomic resources which play a significant role for developing resistant genes to different stresses and give high genetic variation. Wild species are, henceforth, important to contribute required genes when crossed with the available elite sesame genotypes in order to improve their resistance to the different biotic and abiotic stresses and their adaptability to various environmental conditions. Therefore, there is an urgent need to undertake an intensive marker-assisted breeding to unravel potential candidate genes associated to abiotic and biotic stresses in sesame.

4. In this comprehensive review, the progresses of various molecular markers including RAPD, AFLP, SSR, ISSR, EST-SSR, cDNA-SSR, and cpSSR have been comprehensively reviewed. Besides, recently developed markers, such as InDels, RAD-seq, SLAF-seq, RNA-seq, WGS, and SNPs, have been comprehensively reviewed. Therefore, special attentions should be given on these research findings for sesame improvement through high-density genetic map construction, comparative genomic analysis, map-based cloning for yield-related traits, QTL/gene mapping and marker-assisted selection (MAS). Moreover, the use of different molecular markers related to different traits, such as yield and yield-related, branching habit, determinate growth, non-shattering, oil content and quality, tolerance to waterlogging and drought stresses, and resistance to various diseases and insect/pests, should be applied to get high specificity in crop improvement programs.

5. Production and productivity of African sesame especially Ethiopian sesame genotypes are very low due to low genetic potential, biotic and abiotic factors. To improve sesame production it is, therefore, very important to know the ideal environmental conditions, sesame phenotype, and exploiting sesame diversity. In addition, an intensive research work is needed on phenotype interaction with environment to improve sesame production.

6. Sesame genetic resources characterization is considered as a prerequisite of breeding activities. Therefore, sesame genetic resources collection, characterization, and conservation is urgently demanding to discover new variations for farther breeding programs.

7. Ethiopian sesame accessions are less characterized and preserved as compared to other top sesame producing countries. As a result, the country is getting low yield which is about 5 Qt/ha. Hence, special attention should be given to omics technologies to enhance production and productivity of Ethiopian sesame accessions. In addition, shattering is the major yield reducing factor in Ethiopian sesame. Therefore, applying genome editing and genetic transformation to solve this problem is of a paramount important.

8. Root system architecture (RSA) of sesame is very important as it plays a significant role in adapting various climatic conditions. However, no study has been conducted in RSA of sesame so an intensive research on RSA is urgently demanding.

Declarations

Conflict of interest The authors have declared that no competing interests exist.

Ethics approval and consent to participate Not Applicable.

Consent for publication Not applicable.

References

Abate M, Mekbib F, Ayana A, Nigussie M (2015) Assessment of genetic diversity in Ethiopian sesame (*Sesamum indicum* L.) germplasm using ISSR markers. Biotechnol J Int 8:1–13

Abdellatef E, Sirelkhatem R, Ahmed MM, Radwan K, Khalafalla M (2008) Study of genetic diversity in Sudanese sesame (*Sesamum indicum* L.) germplasm using random amplified polymorphic DNA (RAPD) markers. Afr J Biotechnol 7

Adéoti K, Rival A, Dansi A, Santoni S, Brown S, Beule T, Nato A, Henry Y, Vodouhe R, Loko L (2011) Genetic characterization of two traditional leafy vegetables (*Sesamum radiatum* Thonn. ex Hornem and *Ceratotheca sesamoides* Endl.) of Benin, using flow cytometry and amplified fragment length polymorphism (AFLP) markers. Afr J Biotech 10:14264–14275

Admas Alemu YP, Tesfaye K (2013) Genetic distance of sesame (*Sesamum indicum* L.) cultivars and varieties from Northwestern Ethiopia using inter simple sequence repeat markers. East Afr J Sci 7:31–40

Adu-Gyamfi R, Prempeh R, Zakaria I (2019) Diversity assessment of some sesame (*Sesamum indicum* L.) genotypes cultivated in northern ghana using morphological and simple sequence repeat (SSR) markers. Adv Agric

Akbar F, Rabbani MA, Masood MS, Shinwari ZK (2011) Genetic diversity of sesame (*Sesamum indicum* L.) germplasm from Pakistan using RAPD markers. Pak J Bot 43:2153–2160

Ali GM, Yasumoto S, Seki-Katsuta M (2007) Assessment of genetic diversity in sesame (*Sesamum indicum* L.) detected by amplified fragment length polymorphism markers. Electron J Biotechnol 10:12–23

ALshamrani SM, Safhi FA, Alshaya DS, Ibrahim AA, Mansour H, Abd El Moneim D (2022) Genetic diversity using biochemical, physiological, karyological and molecular markers of *Sesamum indicum* L. Front Genet 13

Anandan R, Deenathayalan T, Prakash M, Sunilkumar B, Narayanan G (2018) Assessment of genetic diversity among sesame (*Sesamum indicum* L.) germplasm as revealed by RAPD and SSR markers

Araújo EdS, Arriel NHC, dos Santos RC, de Lima L (2019) Assessment of genetic variability in sesame accessions using SSR markers and morpho-agronomic traits. Embrapa Algodão-Artigo em periódico indexado (ALICE)

Asekova S, Kulkarni KP, Oh KW, Lee M-H, Oh E, Kim J-I, Yeo U-S, Pae S-B, Ha TJ, Kim SU (2018) Analysis of molecular variance and population structure of sesame (*Sesamum indicum* L.) genotypes using simple sequence repeat markers. Plant Breed Biotechnol 6:321–336

Badri J, Yepuri V, Ghanta A, Siva S, Siddiq EA (2014) Development of microsatellite markers in sesame (*Sesamum indicum* L.). Turk J Agric for 38:603–614

Barcelos E, Rios SDA, Cunha RN, Lopes R, Motoike SY, Babiychuk E, Skirycz A, Kushnir S (2015) Oil palm natural diversity and the potential for yield improvement. Front Plant Sci 6:190

Basak M, Uzun B, Yol E (2019) Genetic diversity and population structure of the Mediterranean sesame core collection with use of genome-wide SNPs developed by double digest RAD-Seq. PLoS ONE 14:0223757 https://doi.org/10.1371/journal.pone.0223757

Bedigian D (2003) Evolution of sesame revisited: domestication, diversity and prospects. Genet Resour Crop Evol 50:779–787

Bhat KV, Babrekar PP, Lakhanpaul S (1999) Study of genetic diversity in Indian and exotic sesame (*Sesamum indicum* L.) germplasm using random amplified polymorphic DNA (RAPD) markers. Euphytica 110:21–34

Bhattacharjee M, Iqbal A, Singha S, Nath D, Prakash S, Dasgupta T (2019) Genetic diversity in *Sesamum indicum* L. Bangladesh J Bot 48:497–506

Bhattacharjee M, Prakash S, Roy S, Soumen S, Begum T, Dasgupta T (2020) SSR-based DNA fingerprinting of 18 elite Indian varieties of sesame (*Sesamum indicum* L.). Nucleus 63:67–73

Bisht I, Mahajan R, Loknathan T, Agrawal R (1998) Diversity in Indian sesame collection and stratification of germplasm accessions in different diversity groups. Genet Resour Crop Evol 45:325–335

Cho Y-I, Park J-H, Lee C-W, Ra W-H, Chung J-W, Lee J-R, Ma K-H, Lee S-Y, Lee K-S, Lee M-C (2011) Evaluation of the genetic diversity and population structure of sesame (*Sesamum indicum* L.) using microsatellite markers. Genes Genomics 33:187–195

Cui C, Mei H, Liu Y, Zhang H, Zheng Y (2017) Genetic diversity, population structure, and linkage disequilibrium of an association-mapping panel revealed by genome-wide SNP markers in sesame. Front Plant Sci 8:1189

Dar AA, Mudigunda S, Mittal PK, Arumugam N (2017) Comparative assessment of genetic diversity in *Sesamum indicum* L. using RAPD and SSR markers. 3 Biotech 7:10

Das A, Pandey S, Dasgupta T (2013) Association of heterosis with combining ability and genetic divergence in sesame (*Sesamum indicum* L.). Int J Sci Tech Res 2:307–314

Dixit A, Jin MH, Chung JW, Yu JW, Chung HK, Ma KH, Park YJ, Cho EG (2005) Development of polymorphic microsatellite markers in sesame (*Sesamum indicum* L.). Mol Ecol Notes 5:736–738

Dossa K (2016) A physical map of important QTLs, functional markers and genes available for sesame breeding programs. Physiol Mol Biol Plants 22:613–619

Dossa K, Diouf D, Cissé N (2016a) Genome-wide investigation of Hsf genes in sesame reveals their segmental duplication expansion and their active role in drought stress response. Front Plant Sci 7:1522

Dossa K, Wei X, Zhang Y, Fonceka D, Yang W, Diouf D, Liao B, Cissé N, Zhang X (2016b) Analysis of genetic diversity and population structure of sesame accessions from Africa and Asia as major centers of its cultivation. Genes 7:14

Dossa K, Diouf D, Wang L, Wei X, Zhang Y, Niang M, Fonceka D, Yu J, Mmadi MA, Yehouessi LW (2017a) The emerging oilseed crop Sesamum indicum enters the "Omics" era. Front Plant Sci 8:1154

Dossa K, Yu J, Liao B, Cisse N, Zhang X (2017b) Development of highly informative genome-wide single sequence repeat markers for breeding applications in sesame and construction of a web resource: SisatBase. Front Plant Sci 8:1470

Du H, Zhang H, Wei L, Li C, Duan Y, Wang H (2019) A high-density genetic map constructed using specific length amplified fragment (SLAF) sequencing and QTL mapping of seed-related traits in sesame (*Sesamum indicum* L.). BMC Plant Biol 19:588

Du L, Liu Q, Zhao K, Tang J, Zhang X, Yue B, Fan Z (2020) PSMD: an extensive database for pan-species microsatellite investigation and marker development. Mol Ecol Resour 20:283–291

El Harfi M, Charafi J, Houmanat K, Hanine H, Nabloussi A (2021) Assessment of genetic diversity in Moroccan sesame (*Sesamum indicum*) using ISSR molecular markers. OCL 28:3

Ercan AG, Taskin M, Turgut K (2004) Analysis of genetic diversity in Turkish sesame (*Sesamum indicum* L.) populations using RAPD markers★. Genet Resour Crop Evol 51:599–607

FAOSTAT (2020) FAOSTAT provides free access to food and agriculture data for Over 245 countries and territories and covers all FAO regional groupings. http://faostat.fao.org/. Accessed 19 May 2020

Gebru MM, Yifter M, Kiros M, Mesfin M, Tadesse K (2019) Genetic characterization of sesame (*Sesamum indicum* L.) cultivars in Ethiopia using issr markers. J Drylands 9:969–976

Hayashi K, Yoshida H, Ashikawa I (2006) Development of PCR-based allele-specific and InDel marker sets for nine rice blast resistance genes. Theor Appl Genet 113:251–260

Hodgkin T, Brown AH, Van Hintum TJ, Morales EAV (1995) Core collections of plant genetic resources. John Wiley & Sons, Hoboken

IPGRI N (2004) Descriptors for sesame, Sesamum

Kang CW, Kim SY, Lee SW, Mathur PN, Hodgkin T, De Zhou M, Lee JR (2006) Selection of a core collection of Korean sesame germplasm by a stepwise clustering method. Breed Sci 56:85–91

Ke T, Yu J, Dong C, Mao H, Hua W, Liu S (2015) ocsESTdb: a database of oil crop seed EST sequences for comparative analysis and investigation of a global metabolic network and oil accumulation metabolism. BMC Plant Biol 15:19

Kim D, Zur G, Danin-Poleg Y, Lee S, Shim K, Kang C, Kashi Y (2002) Genetic relationships of sesame germplasm collection as revealed by inter-simple sequence repeats. Plant Breeding 121:259–262

Kizil S, Basak M, Guden B, Tosun HS, Uzun B, Yol E (2020) Genome-wide discovery of InDel markers in sesame (*Sesamum indicum* L.) using ddRADSeq. Plants 9:1262

Kumar V, Sharma S (2009) Assessment of genetic diversity of sesame (*Sesamum indicum* L.) genotypes using morphological and RAPD markers. Indian J Genet Plant Breed 69:209–218

Kumar V, Sharma SN (2011) Comparative potential of phenotypic, ISSR and SSR markers for characterization of sesame (*Sesamum indicum* L.) varieties from India. J Crop Sci Biotechnol 14:163–171

Kumar H, Kaur G, Banga S (2012) Molecular characterization and assessment of genetic diversity in sesame (*Sesamum indicum* L.) germplasm collection using ISSR markers. J Crop Improv 26:540–557

Laurentin HE, Karlovsky P (2006) Genetic relationship and diversity in a sesame (*Sesamum indicum* L.) germplasm collection using amplified fragment length polymorphism (AFLP). BMC Genet 7:10

Laurentin H, Karlovsky P (2007) AFLP fingerprinting of sesame (*Sesamum indicum* L.) cultivars: identification, genetic relationship and comparison of AFLP informativeness parameters. Genet Resour Crop Evol 54:1437–1446

Laurentin H, Ratzinger A, Karlovsky P (2008) Relationship between metabolic and genomic diversity in sesame (*Sesamum indicum* L.). BMC Genomics 9:250

Li C, Miao H, Wei L, Zhang T, Han X, Zhang H (2014) Association mapping of seed oil and protein content in *Sesamum indicum* L. using SSR markers. PLoS One 9

Li-Bin W, Zhang H-Y, Zheng Y-Z, Wang-Zhen G, Zhang T-Z (2008) Developing EST-derived microsatellites in sesame (*Sesamum indicum* L.). Acta Agron Sin 34:2077–2084

López-Villavicencio M, Branca A, Giraud T, Shykoff JA (2005) Sex-specific effect of *Microbotryum violaceum* (Uredinales) spores on healthy plants of the gynodioecious *Gypsophila repens* (Caryophyllaceae). Am J Bot 92:896–900

Makinde F, Akinoso R (2013) Nutrient composition and effect of processing treatments on anti nutritional factors of Nigerian sesame (*Sesamum indicum* Linn.) cultivars. Int Food Res J 20:2293

Marioni JC, Mason CE, Mane SM, Stephens M, Gilad Y (2008) RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. Genome Res 18:1509–1517

Mei H, Liu Y, Du Z, Wu K, Cui C, Jiang X, Zhang H, Zheng Y (2017) High-density genetic map construction and gene mapping of basal branching habit and flowers per leaf axil in sesame. Front Plant Sci 8:636

Morris J (2009) Characterization of sesame (*Sesamum indicum* L.) germplasm regenerated in Georgia, USA. Genet Resour Crop Evol 56:925

Nakano D, Itoh C, Takaoka M, Kiso Y, Tanaka T, Matsumura Y (2002) Antihypertensive effect of sesamin. IV. Inhibition of vascular superoxide production by sesamin. Biol Pharm Bull 25:1247–1249

Nyongesa BO, Were BA, Gudu S, Dangasuk OG, Onkware AO (2013) Genetic diversity in cultivated sesame (*Sesamum indicum* L.) and related wild species in East Africa. J Crop Sci Biotechnol 16:9–15

Oduoye OT, Oluwasanya OA, Arikawe OO, Sunday A, Ayekun OA, Oyenpemi OS, Jandong E, Uguru M, Okechukwu E (2020) Genetic variation via SSR polymorphic information content and ecological distribution of Nigerian sesame. Afr J Biotech 19:165–170

Pandey SK, Das A, Rai P, Dasgupta T (2015) Morphological and genetic diversity assessment of sesame (*Sesamum indicum*

L.) accessions differing in origin. Physiol Mol Biol Plants 21:519–529

Park J-H, Suresh S, Cho G-T, Choi N-G, Baek H-J, Lee C-W, Chung J-W (2014) Assessment of molecular genetic diversity and population structure of sesame (*Sesamum indicum* L.) core collection accessions using simple sequence repeat markers. Plant Genet Resour 12:112–119

Park J-H, Suresh S, Raveendar S, Baek H-J, Kim C-K, Lee S, Cho G-T, Ma K-H, Lee C-W, Chung J-W (2015) Development and evaluation of core collection using qualitative and quantitative trait descriptor in sesame (*Sesamum indicum* L.) germplasm. Korean J Crop Sci 60:75–84

Parsaeian M, Mirlohi A, Saeidi G (2011) Study of genetic variation in sesame (*Sesamum indicum* L.) using agro-morphological traits and ISSR markers. Russ J Genet 47:314

Pham TD, Bui TM, Werlemark G, Bui TC, Merker A, Carlsson AS (2009) A study of genetic diversity of sesame (*Sesamum indicum* L.) in Vietnam and Cambodia estimated by RAPD markers. Genet Resour Crop Evol 56:679–690

Pham TD, Geleta M, Bui TM, Bui TC, Merker A, Carlsson AS (2011) Comparative analysis of genetic diversity of sesame (*Sesamum indicum* L.) from Vietnam and Cambodia using agro-morphological and molecular markers. Hereditas 148:28–35

Purru S, Sahu S, Rai S, Rao A, Bhat K (2018) GinMicrosatDb: a genome-wide microsatellite markers database for sesame (*Sesamum indicum* L.). Physiol Mol Biol Plants 24:929–937

Salazar B, Laurentín H, Dávila M, Castillo MA (2006) Reliability of the RAPD technique for germplasm analysis of sesame (*Sesamum indicum* L.) from Venezuela. Interciencia 31:456–460

Sehr EM, Okello-Anyanga W, Hasel-Hohl K, Burg A, Gaubitzer S, Rubaihayo PR, Okori P, Vollmann J, Gibson P, Fluch S (2016) Assessment of genetic diversity amongst Ugandan sesame (*Sesamum indicum* L.) landraces based on agromorphological traits and genetic markers. J Crop Sci Biotechnol 19:117–124

Sharma S, Kumar V, Mathur S (2009) Comparative analysis of RAPD and ISSR markers for characterization of sesame (*Sesamum indicum* L) genotypes. J Plant Biochem Biotechnol 18:37–43

Shi Y-Z, Forneris N, Rajora OP (2014) Highly informative single-copy nuclear microsatellite DNA markers developed using an AFLP-SSR approach in black spruce (*Picea mariana*) and red spruce (*P. rubens*). PLoS ONE 9:e103789

Shittu L, Bankole M, Ahmed T, Bankole M, Shittu R, Saalu C, Ashiru O (2008) Antibacterial and antifungal activities of essential oils of crude extracts of *Sesame radiatum* against some common pathogenic micro-organisms. Iran J Pharmacol Ther 6:165–170

Spandana B, Reddy VP, Prasanna GJ, Anuradha G, Sivaramakrishnan S (2012) Development and characterization of microsatellite markers (SSR) in Sesamum (*Sesamum indicum* L.) species. Appl Biochem Biotechnol 168:1594–1607

Stavridou E, Lagiotis G, Kalaitzidou P, Grigoriadis I, Bosmali I, Tsaliki E, Tsiotsiou S, Kalivas A, Ganopoulos I, Madesis P (2021) Characterization of the genetic diversity present in a diverse sesame landrace collection based on phenotypic traits and EST-SSR markers coupled with an HRM analysis. Plants 10:656

Surapaneni M, Yepuri V, Vemireddy LR, Ghanta A, Siddiq E (2014) Development and characterization of microsatellite markers in Indian sesame (*Sesamum indicum* L.). Mol Breed 34:1185–1200

Tabatabaei I, Pazouki L, Bihamta MR, Mansoori S, Javaran MJ, Niinemets U (2011) Genetic variation among iranian sesame ('*Sesamum indicum*'L.) Accessions vis-a-vis exotic genotypes on the basis of morpho-physiological traits and rapd markers. Aust J Crop Sci 5:1396

Teklu DH, Shimelis H, Tesfaye A, Mashilo J, Zhang X, Zhang Y, Dossa K, Shayanowako A (2021) Genetic variability and population structure of Ethiopian Sesame (*Sesamum Indicum*

L.) germplasm assessed through phenotypic traits and simple sequence repeats markers

Tesfaye T, Tesfaye K, Keneni G, Ziyomo C, Alemu T (2022) Genetic diversity of Sesame (*Sesamum indicum* L) using high throughput diversity array technology. J Crop Sci Biotechnol 25:359–371

Uncu AO, Frary A, Karlovsky P, Doganlar S (2016) High-through-put single nucleotide polymorphism (SNP) identification and mapping in the sesame (*Sesamum indicum* L.) genome with genotyping by sequencing (GBS) analysis. Mol Breed 36:173

Varshney RK, Terauchi R, McCouch SR (2014) Harvesting the promising fruits of genomics: applying genome sequencing technologies to crop breeding. PLoS Biol 12

Wang Z, Fang B, Chen J, Zhang X, Luo Z, Huang L, Chen X, Li Y (2010) De novo assembly and characterization of root transcriptome using Illumina paired-end sequencing and development of cSSR markers in sweetpotato (*Ipomoea batatas*). BMC Genomics 11:726

Wang L, Zhang Y, Qi X, Gao Y, Zhang X (2012a) Development and characterization of 59 polymorphic cDNA-SSR markers for the edible oil crop *Sesamum indicum* (Pedaliaceae). Am J Bot 99:e394–e398

Wang N, Fang L, Xin H, Wang L, Li S (2012b) Construction of a high-density genetic map for grape using next generation restriction-site associated DNA sequencing. BMC Plant Biol 12:148

Wang L, Han X, Zhang Y, Li D, Wei X, Ding X, Zhang X (2014a) Deep resequencing reveals allelic variation in *Sesamum indicum*. BMC Plant Biol 14:225

Wang L, Yu S, Tong C, Zhao Y, Liu Y, Song C, Zhang Y, Zhang X, Wang Y, Hua W (2014b) Genome sequencing of the high oil crop sesame provides insight into oil biosynthesis. Genome Biol 15:R39

Wang L, Yu J, Li D, Zhang X (2015) Sinbase: an integrated database to study genomics, genetics and comparative genomics in *Sesamum indicum*. Plant Cell Physiol 56:e2–e2

Wang L, Xia Q, Zhang Y, Zhu X, Zhu X, Li D, Ni X, Gao Y, Xiang H, Wei X (2016) Updated sesame genome assembly and fine mapping of plant height and seed coat color QTLs using a new high-density genetic map. BMC Genomics 17:31

Wang L, Zhang Y, Zhu X, Zhu X, Li D, Zhang X, Gao Y, Xiao G, Wei X (2017) Development of an SSR-based genetic map in sesame and identification of quantitative trait loci associated with charcoal rot resistance. Sci Rep 7:1–8

Wei L-B, Zhang H-Y, Zheng Y-Z, Miao H-M, Zhang T-Z, Guo W-Z (2009) A genetic linkage map construction for sesame (*Sesamum indicum* L.). Genes Genomics 31:199–208

Wei W, Qi X, Wang L, Zhang Y, Hua W, Li D, Lv H, Zhang X (2011) Characterization of the sesame (*Sesamum indicum* L.) global transcriptome using Illumina paired-end sequencing and development of EST-SSR markers. BMC Genomics 12:451

Wei W, Zhang Y, Lü H, Li D, Wang L, Zhang X (2013) Association analysis for quality traits in a diverse panel of Chinese Sesame (*Sesamum indicum* L.) germplasm. J Integr Plant Biol 55:745–758

Wei L, Miao H, Li C, Duan Y, Niu J, Zhang T, Zhao Q, Zhang H (2014a) Development of SNP and InDel markers via de novo transcriptome assembly in *Sesamum indicum* L. Mol Breed 34:2205–2217

Wei X, Wang L, Zhang Y, Qi X, Wang X, Ding X, Zhang J, Zhang X (2014b) Development of simple sequence repeat (SSR) markers of sesame (*Sesamum indicum*) from a genome survey. Molecules 19:5150–5162. https://doi.org/10.3390/molecules19045150

Wei X, Liu K, Zhang Y, Feng Q, Wang L, Zhao Y, Li D, Zhao Q, Zhu X, Zhu X (2015) Genetic discovery for oil production and quality in sesame. Nat Commun 6:8609

Wei X, Zhu X, Yu J, Wang L, Zhang Y, Li D, Zhou R, Zhang X (2016) Identification of sesame genomic variations from genome comparison of landrace and variety. Front Plant Sci 7:1169

Wei L, Miao H, Xu F, Kong J, Zhang H (2017a) Chinese sesame cultivars, DNA fingerprinting, and two-dimensional barcodes using single nucleotide polymorphisms, insertions or deletions, and simple sequence repeat markers. Crop Sci 57:1941–1947

Wei X, Gong H, Yu J, Liu P, Wang L, Zhang Y, Zhang X (2017b) SesameFG: an integrated database for the functional genomics of sesame. Sci Rep 7:1–10

Weldemichael MY, Bayratsion YT, Sbhatu DB, Abraha GG, Juhar HM, Birhan Kassa A, Sibhatu FB, Berhe BD, Sahle TE, Mossa MM (2020) In vitro shoot regeneration of oil seed crop *Sesamum indicum* L. from seedling cotyledon explant to lay ground for genetic transformation in Ethiopia. Int J Agron

Woldesenbet DT, Tesfaye K, Bekele E (2015a) Genetic diversity of sesame germplasm collection (*Sesamum indicum* L.): implication for conservation, improvement and use. Int J Biotechnol Mol Biol Res 6:7–18

Woldesenbet DT, Tesfaye K, Bekele E (2015b) Genetic diversity of sesame germplasm collection (*Sesamum indicum* L.): implication for conservation, improvement and use. Int J Biotechnol Mol Biol Res 6:7–18

Wu K, Liu H, Yang M, Tao Y, Ma H, Wu W, Zuo Y, Zhao Y (2014a) High-density genetic map construction and QTLs analysis of grain yield-related traits in Sesame (*Sesamum indicum* L.) based on RAD-Seq techonology. BMC Plant Biol 14:274

Wu K, Yang M, Liu H, Tao Y, Mei J, Zhao Y (2014b) Genetic analysis and molecular characterization of Chinese sesame (*Sesamum indicum* L.) cultivars using Insertion-Deletion (InDel) and Simple Sequence Repeat (SSR) markers. BMC Genet 15:35

Xiurong Z, Yingzhong Z, Yong C, Xiangyun F, Qingyuan G, Mingde Z, Hodgkin T (2000) Establishment of sesame germplasm core collection in China. Genet Resour Crop Evol 47:273–279

Yepuri V, Surapaneni M, Kola VSR, Vemireddy LR, Jyothi B, Dinesh-kumar V, Anuradha G, Siddiq EA (2013) Assessment of genetic diversity in sesame (*Sesamum indicum* L.) genotypes, using EST-derived SSR markers. J Crop Sci Biotechnol 16:93–103. https://doi.org/10.1007/s12892-012-0116-9

Yi D-K, Kim K-J (2012) Complete chloroplast genome sequences of important oilseed crop *Sesamum indicum* L. PLoS One 7

Yu J, Ke T, Tehrim S, Sun F, Liao B, Hua W (2015) PTGBase: an integrated database to study tandem duplicated genes in plants. Database 2015

Yu J, Dossa K, Wang L, Zhang Y, Wei X, Liao B, Zhang X (2017) PMDBase: a database for studying microsatellite DNA and marker development in plants. Nucleic Acids Res 45:D1046–D1053

Yu J, Golicz AA, Lu K, Dossa K, Zhang Y, Chen J, Wang L, You J, Fan D, Edwards D (2019) Insight into the evolution and functional characteristics of the pan-genome assembly from sesame landraces and modern cultivars. Plant Biotechnol J 17:881–892

Yuan Q, Zhang H, Miao H, Duan Y, Wei Q, Wang X (2018) Effects of water logging stress on the quality of sesame seed and oil product. Acta Agri Bor-Sin 33:202–208

Yue W, Wei L, Zhang T, Li C, Miao H, Zhang H (2012) Genetic diversity and population structure of germplasm resources in sesame (*Sesamum indicum* L.) by SSR markers. Acta Agron Sin 38:2286–2296

Zeven AC, Zhukovsky PM (1975) Dictionary of cultivated plants and their centres of diversity: excluding ornamentals, forest trees and lower plants Pudoc

Zhang Y-X, Zhang X-R, Hua W, Wang L-H, Che Z (2010) Analysis of genetic diversity among indigenous landraces from sesame (*Sesamum indicum* L.) core collection in China as revealed by SRAP and SSR markers. Genes Genomics 32:207–215

Zhang Y-X, Jian S, Zhang X-R, Wang L-H, Zhuo C (2011) Analysis on genetic diversity and genetic basis of the main sesame cultivars released in China. Agric Sci China 10:509–518

Zhang H, Wei L, Miao H, Zhang T, Wang C (2012a) Development and validation of genic-SSR markers in sesame by RNA-seq. BMC Genomics 13:316

Zhang Y, Zhang X, Che Z, Wang L, Wei W, Li D (2012b) Genetic diversity assessment of sesame core collection in China by phenotype and molecular markers and extraction of a mini-core collection. BMC Genet 13:102

Zhang Y, Zhang X, Che Z, Wang L, Wei W, Li D (2012c) Genetic diversity assessment of sesame core collection in China by phenotype and molecular markers and extraction of a mini-core collection. BMC Genet 13:1–14

Zhang H, Li C, Miao H, Xiong S (2013a) Insights from the complete chloroplast genome into the evolution of *Sesamum indicum* L. PLoS One 8

Zhang H, Miao H, Wang L, Qu L, Liu H, Wang Q, Yue M (2013b) Genome sequencing of the important oilseed crop *Sesamum indicum* L. Genome Biol 14:401

Zhang H, Miao H, Wei L, Li C, Zhao R, Wang C (2013c) Genetic analysis and QTL mapping of seed coat color in sesame (*Sesamum indicum* L). PLoS One 8

Zhang Y, Wang L, Xin H, Li D, Ma C, Ding X, Hong W, Zhang X (2013d) Construction of a high-density genetic map for sesame based on large scale marker development by specific length amplified fragment (SLAF) sequencing. BMC Plant Biol 13:141

Zhang H, Miao H, Li C, Wei L, Duan Y, Ma Q, Kong J, Xu F, Chang S (2016) Ultra-dense SNP genetic map construction and identification of SiDt gene controlling the determinate growth habit in *Sesamum indicum* L. Sci Rep 6:31556

Zhang Y, Wang L, Gao Y, Li D, Yu J, Zhou R, Zhang X (2018) Genetic dissection and fine mapping of a novel dt gene associated with determinate growth habit in sesame. BMC Genet 19:38

Zhang H, Miao H, Ju M (2019) Potential for adaptation to climate change through genomic breeding in sesame, Genomic designing of climate-smart oilseed crops. Springer, New York, pp 371–440