*Current Agriculture* **9(2) 17-48 (April, 2025)**
**www.currentagriculture.com**

---

**REVIEW ARTICLE**

# Comprehensive review on application of marker assisted selection in Arabica coffee

**D. Geneti and D. Merga**

**Ethiopian Institute of Agricultural Research, Jimma Agricultural Research Center, P.O. Box 192, Jimma, Ethiopia**

**Corresponding authors email: dawitmerga@gmail.com and dulageneti983@gmail.com**

Manuscript received: April 1,2025; Decision on manuscript: April 8,2025; Manuscript accepted: April 15, 2025

---

## Abstract

Improving coffee through conventional breeding programs is challenging due to its perennial nature, the high cost of field trials, and dependence on environmental conditions. These factors have prompted researchers to apply marker-assisted selection (MAS) studies to further exploit the genetic diversity of Arabica coffee and improvement for desirable traits. This review summarizes the available information on the application of MAS in Arabica coffee and identifies the challenges and opportunities for future research to address issues related to its application. Research on the phylogenetic origin and evolution of Arabica coffee using various MAS techniques has shown Coffea species clustering based on their botanical types, geographical origins, and maternal progenitors, with a specific focus on Ethiopian-origin Arabica coffee. Studies on genetic diversity using MAS have demonstrated that Arabica coffee germplasm in Ethiopia has a broad genetic base, making it valuable for developing varieties capable of sustaining global coffee production. Arabica coffee germplasm gene bank were assessed using MAS, identifying 2,533 genotypes for core collections to assist breeders in selecting parent plants and reducing maintenance costs. Research on linkage and QTL mapping through MAS has identified mapping strategies for various desired traits, although progress in this area is still in its infancy, with limited success. Different candidate genes have been identified for major coffee diseases such as coffee berry dissease (T, R, K, and ckl-1 genes), coffee leaf rust (SH1–SH9), and coffee wilt disease (strains affecting only Arabica coffee). Furthermore, biochemical candidate genes have been identified for key compounds like caffeine (Caffeine-5), sucrose (Sucrose-3), trigonelline (Trigonelline-1), fatty acids (Fatty acid-1), and CGAs (CGAs-1). However, significant challenges remain in overcoming the limitations of MAS research on Arabica coffee, particularly in Ethiopia, where the highest diversity of Arabica coffee is recorded. To advance MAS applications in Arabica coffee research, particularly in developing countries like Ethiopia, adequate policies must be implemented to encourage innovation and investment in this field.

**Key words:** Arabica coffee, Disease resistance, Improvement, Marker-assisted selection, QTLs

## Introduction and concept

Coffee is a stimulant beverage crop, and it belongs to the family Rubiaceae and the genus Coffea. They are mostly grown in subtropical and tropical regions (Morris, 2018) and consist of 124 species (Zhou *et al.*, 2016; Vega *et al.*, 2008).

*Coffea canephora* Pierre and *Coffea arabica* L. are the only two economically important species and are widely cultivated worldwide. *Coffea liberica* is also cultivated on a small scale to satisfy local consumption (USDA, 2020). Almost all coffee species are diploid (2n=2x=22) and most are self-incompatible, except for *Coffea arabica* L., which is a self-fertile species and a natural allotetraploid (2n=4x = 44). Arabica coffee is a self-pollinating species with a common outcrossing rate of less than 10%, which is sufficient to induce some variation in offspring and free-pollinating cultivars (Morris, 2018). Ethiopia is both the center of origin and diversification of Arabica coffee (Bayetta, 2001). The crop is being grown in the country from the river bank of Gambella plain (550 2 m.a.s.l) to the northern highlands of the country with an altitude of 2600m (Bayetta, 1986). Within this range of altitudes and agro-ecological diversity, considerable genetic diversity exists among the cultivated and traditionally recognized landraces of Arabica coffee in Ethiopia, as confirmed by various studies (Ermias *et al.,* 2005; Dessalegn *et al.,* 2018; Zenebe *et al.,* 2019; Getachew *et al.,* 2019, Tadesse *et al.,* 2020). Despite the existence of high genetic diversity in the Arabica coffee population that provides immense opportunities for improvement programs, MAS research on Arabica coffee is limited. Until today, most researches conducted on Arabica coffee are conventional (traditional) methods, particularly in developing countries that produce Arabica coffee in large amounts (Adem, 2020). Conventional breeding offers limitations due to the long regeneration time of the coffee tree (three years), the high cost of field trials (Lashermes *et al.,* 2000b), and depend on environment. In cases backcrossing or breeding cycle is done over five generations, a minimum of 25-30 years after initial hybridization is required to ensure the improvement for desired traits (Lashermes *et al.,* 2000b). The tetraploid nature of Arabica coffee poses difficulties in breeding with other diploid *Coffea species* (Fazuoli *et al.,* 2000). Hence, to tackle the problem, efforts have been underway with a method called MAS breeding, which offer newer, easy, and more efficient practical alternatives to surmount the conventional breeding problems faced in coffee improvement. The MAS strategies in coffee research involve the use of DNA variations based on genetic markers, providing new hopes and possibilities for genetic improvement for difficult species like Arabica coffee (Lashermes *et al.,* 2000). The great potential of MAS-based technologies is now well demonstrated, and these are being utilized for crop improvement throughout the world for genotyping,varietal identification and claiming intellectual property rights, germplasm fingerprinting, construction of linkage maps, Quantitative Trait Loci (QTL) identification to finally develop genetically improved crops with desirable traits (Prasad and Ramesh, 2014). The MAS technologies have been also useful in helping decipher many of the evolutionary puzzles about crop species' origin, spread, and taxonomic relationships (Muhammad *et al.,* 2017). In the last 3 decades, conscious efforts have begun globally to integrate MAS-based technologies, which can provide impetus, dependability, and directionality to the genetic improvement. For instance, some valuable MAS studies have been conducted on Arabica coffee, such as research on the origin and evolution of Arabica coffee (Berthou et al., 1983; Virginie *et al.,* 2019), phylogenic studies (Mikiru et al, 2022), diversity studies (Tadesse *et al.,* 2020; Mikiru *et al.,* 2022), identification of candidate genes for *Arabica coffee* major diseases, and wide range of variations in biochemical compounds and identification of their candidate genes (caffeine, chlorogenic acids, sucrose) (Gichiru *et al.,* 2008; James *et al.,* 2021). Additionally, studies on introgressed varieties (Vose et al., 1995; Leshermen *et al.,* 2000; Christophe, *et al.,* 2020), and many other MAS studies have been used for the improvement programs.

Thus, coffee research is now experiencing the evolution and flourishing of many of these technologies, which are expected to be utilized in the future. However, there is still a long way to go before visible gains become a reality. Large coffee genomics programs are currently underway in many countries, including Brazil, France, Italy, Colombia, and more recently India and Kenya (Prasad and Ramesh, 2014), but not in Ethiopia. Information on coffee genomics is scant in Ethiopia, despite the high genetic diversity of Arabica coffee reported in the country. In this context, this article was initiated to conduct systematic investigations by focusing on coffee genetic studies assessed using MAS approaches with the objectives to compile concisely the available information on the application of MAS in Arabica coffee and to identify challenges and opportunity for the future research areas to address the problems related to the application of MAS in Arabica coffee.

**Development of MAS for genetic studies in Arabica coffee**

Molecular markers are DNA sequences that are found at the specific location of the genome that is usually inherited following the standard laws of heredity. Markers tightly linked, usually at less than 5 cM, to the gene of interest, serve as a chromosomal landmark for tracking the introgression of the desired gene in progenies in a cross (Mekonnen *et al.*, 2017, James *et al.,* 2021). DNA-based molecular markers are the best markers especially for closely related genotypes as they can be detected at all stages of an organism's development. These markers are not dependent on the stage of growth or the environment occupied by an individual, and they occur in unlimited numbers within the genome (Teressa *et al.,* 2010).These markers include non-PCR-based, PCR-based, and sequence-based markers. Non-PCR-based are also first-generation markers, including Restriction Fragment Length Polymorphism (RFLP), whose observed Polymorphism is based on the length generated by digestion with restriction enzymes (Gimase *et al.,* 2014). The PCR-based markers include Random Amplified Polymorphic DNA (RAPD), Amplified Fragment Length Polymorphism (AFLP), and Simple sequence repeat (SSR) or Microsatellites. The amplifications of AFLP and RAPD markers only indicate the presence or absence of alleles and therefore cannot differentiate homozygote from heterozygote variants (Dinesh *et al.,* 2011). The SSR markers are co-dominant and hence reproducible, in addition to being locus-specific (Barua *et al.,* 2003). The sequence-based markers include Single nucleotide polymorphism (SNP) markers (Ray, and Satya, 2014). Similar to SSR markers, the SNP markers are highly reproducible. This attribute makes the SSRs and SNPs markers the markers of choice in genetic studies and next-generation plant breeding (Ray and, Satya, 2014, James *et al.,* 2021). Concerning applications of DNA markers in coffee, the detection of genetic variation at the DNA level has been made possible by the advent of molecular markers started in the first 1990s. Lashermes *et al.,* (1996a) reported that genetic factors are more accurately tested by molecular markers in coffea species research. Promising efforts have been made in exploring and developing DNA markers for Arabica coffee; however, there is limited information on Arabica coffee compared to other species of the genus Coffea.

**Origin and Evolution of Arabica coffee studies using MAS**

The evolution of Arabica coffee has always been an intriguing puzzle for researchers because it is the only autogamous and tetraploid species in the genus Coffea. In addition to that, when we consider the evolution of Arabica coffee different set of research shows that the species evolve from a cross between two species of coffee genius (*Coffea canephora* and *Coffea eugenioides*).

The natural ranges of *Coffea canephora* and *Coffea eugenioides* overlap in East-Central Africa. However, the natural hybrids between these species in this area are not known so far. This can be explained by the absence of recent hybrids between *Coffea canephora* and *Coffea eugenioides*, for three reasons. First, although both species can be found in the same area, their habitat preference differs substantially. *Coffea eugenioides* is especially found near forest edges, while coffea canephora is mainly restricted to the forest interior (Noirot *et al.*, 2016). Second, the flowering time of both species does not coincide (Noirot *et al.*, 2016). The flowering time of Coffea species is highly species-specific and genetically controlled, hampering interspecific gene flow via pollination (Gomez *et al.*, 2016). Third, the success rate of induced cross-pollination between *Coffea canephora* and *Coffea eugenioides* is very low, suggesting the presence of additional reproductive barriers (Noirot *et al.*, 2016). However, changes in environmental conditions may have broken some of the reproductive barriers between the two specious in the past, enabling a successful interspecific hybridization between these species at the origin of Arabica coffee.

Nevertheless, a number of DNA marker approaches have provided evidence for the evolution of Arabica coffee. Evolutionary informative DNA signatures of organelle genomes (cp-DNA, mt-DNA) and ITS (Internally Transcribed Spacer) region of the nuclear 5S rDNA, and nuclear length polymorphism-based markers like RAPDs, AFLPs, and ISSRs, have been tried on a limited number of coffee species and also accessions with varying success (Table 1). Briefly, the results from these studies show that low polymorphism in organelle DNA provides poor resolution of species relationships, but provides support for monophyly of the Coffea genus. Additionally, these studies suggest Coffee eugioniodes as the possible maternal progenitor of Arabica coffee (Berthou *et al.*, 1983; Lashermes *et al.*, 1997; Cros et al., 1998; Raina et al., 1998; Lashermes et al., 1999; Maurin *et al.*, 2007; Tesfaye *et al.*, 2007; Hamon *et al.*, 2009). Some of such studies carried out are summarized in Table 1.

Conformity to that, the recent study of the origination of Arabica coffee by GBS (Genotyping by sequencing) provides a clear hypothesis regarding the evolutionary origin of Arabica Coffee. The GBS data proved to be more informative than the molecular data used in previous studies because a substantial amount of informative sites seems to be required to get reliable genetic distance estimates for coffee species and the result of this study showed that Coffee eugenioides species was the ovule donor in the Arabica coffee hybridization. Similarly, based on the similarity in plastid DNA markers, *Coffee eugenioides* is a close relative of *Arabica coffee* species and suggested that Coffee eugenioides is the ovule donor in the Arabica coffee hybridization event (Maurin et al., 2007; Tesfaye et al., 2007; Guyeux *et al.*, 2019). Besides, Bawin *et al.*, (2020), using GBS confirmed that *Coffee eugenioides* is genetically more similar to Arabica coffee and Coffee canephora was the putative pollen donor in the hybridization event prior to the emergence of Arabica coffee. When we consider the origins of *Coffea arabica*, at the present time, Arabica coffee is mainly found in the southwestern highlands of Ethiopia (Fig. 1), with some occurrence on the Boma plateau in southeastern South Sudan (Thomas, 1942), and on Mount Imantong in Sudan and Mount Marsabit in northern Kenya (Berthaud and Charrier, 1988). Arabica coffee is the main Coffea species that occurs in those regions and is geographically isolated from all diploid coffee species in the genus, which includes its two progenitor species *Coffee canephora* and *Coffee eugenioides*. Thus, in this point of view, precise localization in Africa of the cradle of *Arabica coffee*, based on the present distribution of its two progenitor species appears difficult.

Indeed, it has been suggested that plants with double genomes auto or allopolyploids have the potential to develop phenotypic novelties, increase their adaptability and obtain higher fitness features that would render them more tolerant towards changing conditions than their diploid counterparts (Amborella *et al.,* 2013). Virginie, *et al.,* (2019) using SNP markers and reported that, in fact, Coffee canephora was probably able to find suitable habitats in Ethiopia and the climatic changes could have reduced the diploid distributions to their current locations. As a consequence, the birthplace of Arabica coffee could possibly be not only in Ethiopia but also in the entire region (South Sudan, Uganda, North Kenya) followed either by migration to present-day Ethiopia or by survival in that region alone.

**Table 1: DNA markers-based studies for deciphering species relationship among *Coffea taxa***

| **Technique** | **Species assessed** | **Groupings and relationship** | **Study source/ Reference** |
|---|---|---|---|
| Chloroplast (cp) and mitochondrial (mt) DNA RFLP polymorphism using Hpa II enzyme and Sal I enzyme respectively | Arabica (2 acc.), robusta (2 acc.), C. eugenioides (2 acc.), C. congensis (2 acc.), 'nana' taxon, C. excelsa, C. liberica, *Para coffea ebracteolata*, *C. arabusta* | **cp−DNA Group A:** *C. arabica,* *C. eugenioides,* *C. congensis* Group A': *C. canephora*, 'nana' taxon mt –DNA **Group 1:** C. arabica, C. eugenioides **Group 2:'**nana' *C. arabica* has an ancestor similar to *C. eugenioides*. | Berthou *et al.,* (1983) |
| Chloroplast DNA RFLP in the atpB-rbcL intergenic region | 52 trees from 25 *Coffea taxa* Total = 52 | – Coffea is monophyletic and recent in origin | Lashermes *et al.,* (1996c) |
| Chloroplast PCR-RFLP (trnT-L, trnL, trnL-F) | Arabica (4 acc.), robusta (3 acc.), *C. eugenioides* (2 acc.), *C. liberica* (2 acc.), *C. stenophylla*, *C. racemosa*, *C. humilis*, *C.pseudozanguebariae*, *C. congensis*, *C. sessi flora*, *C. breviceps* (1 acc. each); Total = 18 | **I =** *C. arabica,* *C. eugenioides,* *C. humilis,* C. stenophylla **II =** *C. canephora,* *C. liberica,* *C. breviceps,* C. congensis **III =** *C. pseudozanguebariae,* *C. sessi flora,* *C. racemosa* | Orozco-Castillo *et al.,* (1996) |
| Mitochondria PCR-RFLP (V7 rDNA) ITS 1 and 2 regions of nuclear ribosomal DNA | 26 Coffea taxa, 3 *Psilanthus taxa* Total = 37 | Strong geographical correspondence among the clusters. | Lashermes *et al.,* (1997) sequence variants were observed. |
| TrnL-trnF intergenic spacer of chloroplast DNA | 23 Coffea taxa and 2 *Psilanthus taxa* | Strong geographical correspondence among the clusters. Supports radial mode and recent origin of Coffea taxa in Africa arabica, and C. sp. Moloundou. | Cros *et al.,* (1998) |

| Technique | Species assessed | Groupings and relationship | Study source/ Reference |
|---|---|---|---|
| RAPD (20 primers), ISSR (10 primers) | 15 *Coffea spp.*, 4 Psilanthus spp. Individual as well as pooled samples | Well-resolved species clusters showing correspondence to their botanical types, as well as geographical origin; *C. kapakata* was indicated to be a *Coffea spp.* rather than belonging to Psilanthus or Psilanthopsis genra. Four endemic paracoffea species appeared as a distinct cluster under related genus Psilanthus | Prasad and Ramesh, 2014 |
| SSR (three primers) | 15 *Coffea spp*., specifically focus on four the most known spp. | C. arabica has an ancestor smilar to C. eugenioides | Pearl *et al.,* (2004) |
| | Three coffea spp | Nuclear genomes have remained essentially unaltered since the formation of the hybrid. | Yu *et al.,* (2011) |
| 8.5K SNP array | Three *Coffee spp*. | Two progenitor species of coffee Arabica *C. canephora* and *C. eugenioides*<br>The birthplace of *C. arabica* could possibly be not only in Ethiopia but also in the entire region (South Sudan, Uganda, North Kenya) | Virginie *et al.,* 2019 |

**Fig 1. Coffea arabica and its progenitor species. (a) Range distribution of the three related species. Dotted lines represent their schematic distribution limit, whereas names with colour labels correspond to sampled sites (*C. canephora* in blue, *C. eugenioides* in gold, and *C. arabica* in red), Map data ©2018 Google, ORION-ME;**

**Phylogenetic studies of *Arabica coffee* species using MAS**

There is some set of research conducted on phylogenetic studies of arabica coffee by using different marker-assisted selection in the 1990s and clustered into different distinct groups. This set of the research reported well-resolved species clusters showing correspondence to their botanical types, as well as geographical origin (Lashermes *et al.*, 1997; Cros *et al.*, 1998;). Depending on their botanical types, they clustered coffee specious into different distinct classes such that, Cluster-I, *C. arabica*, *C. eugenioides*, *C. humilis*, *C. stenophylla*, cluster-II, *C. canephora*, *C. liberica*, *C. breviceps*, and *C. congensis*, cluster-III, *C. pseudozanguebariae*, *C. sessiflora*, and *C. racemos*. They further reported the origin of arabica coffee is southwestern Ethiopia (Table 2).

**Fig 2, A. Phylogeny of Arabica coffee accessions using SNPs Cluster I (Black), - NW and SW, Cluster II (Purple) - SW and N. Cluster III (Blue) SW, N, S, and SE parts. Clustered in group IV (Green) - S and SE accession. Source - mikru *et al.*, 2022. B. Dendrogram depicting relationships among 128 coffee samples based on Jaccard's similarity coefficients of RAPD data Source Aga *et al.*, 2005**

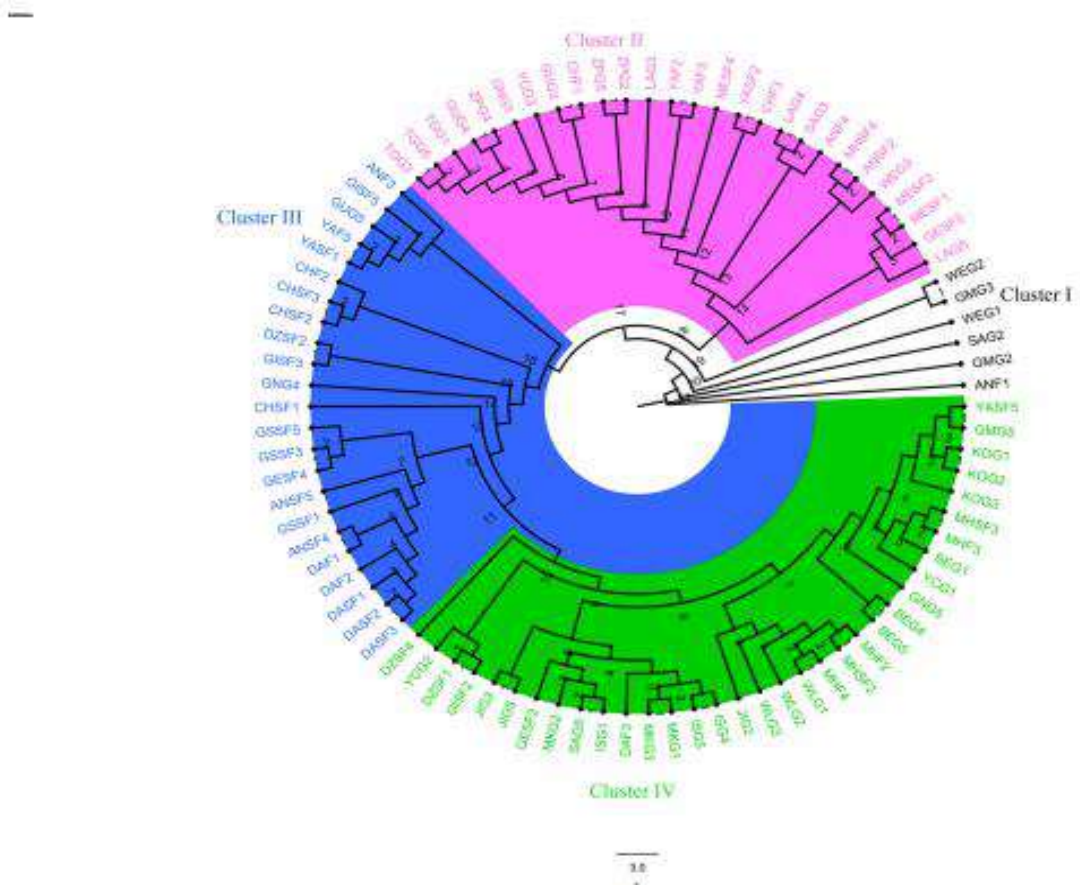
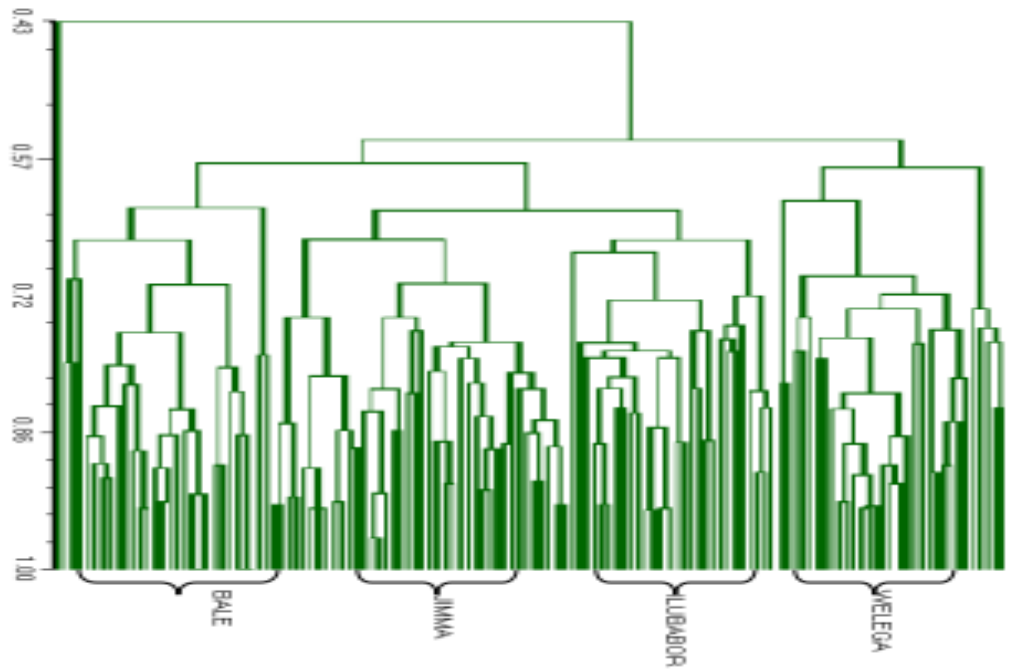


**Fig 2 B**




Hence, when we consider the phylogenetic study of Arabica coffee, a recent study using SNP and SSR markers showed that coffee accessions originating from the same localities had the potential to separate genetically due to the domestication process and human selection activity. For instance, Castilillo *et al.,* (2020) reported the SNP markers generated by DAr T-seq technology separated the 87 accessions of Ethiopian originated arabica coffee into five distinct groups. Similarly, more recently Mikiru et al., (2022) studied the phylogenetic relationship of the 90 coffee accessions using whole-genome SNP markers and, clustered into four major groups, comprising several sub-clades (Fig .2 A)1. Clusters I, II, III, and, IV mainly contained accessions from the southwest.

In conformity to that Montagnon, and Bouharmont, (1996) using RAPD markers reported the similarity between southeast and southwest coffee samples and suggested that the southeastern coffee trees could have been introduced from the southwest by man. The clustering of southwest coffee accessions with the south and southeast population was also previously reported (Anthony *et al.,* 2001; Mekbib *et al.,* 2020; Aga *et al.,* 2005, Mishra *et al.,* 2014). Similarly, Benti *et al.,* (2021) studied the Ethiopian commercial arabica coffee varieties by using SSR markers and also found the grouping of varieties into different clusters regardless of their geographic origin.

These authors suggested that the clustering pattern found in their study could indicate the presence of southwest Ethiopia coffee accessions in different clusters could explain coffee accessions that originated in that region had a broad genetic base and high level of genetic diversity within coffee accessions sampled from the same geographic origin.

**Applications of MAS for genetic diversity studies in Arabica coffee**

A number of studies carried out using molecular markers *viz.*, isozymes, RFLP, RAPD, AFLP, and SSR to detect variability in Arabica and Robusta coffee germplasm (Lashermes et al., 1993, Prakash *et al.,* 2005, Aga *et al.,* 2005). As a result, the information from these markers revealed that a very low level of variability in cultivated Arabica compared to wild Arabica accessions as well as a very low variability of Arabica compared to diploid coffee germplasm. Even though the result is exact in all markers but the size of variability in Arabica coffee was different in a different set of markers used, and this is in turn suggested by the suitability of the markers of Arabica coffee. In the 1990s, the suitability of markers for Arabica coffee was tried to identify by approaches like RAPD and AFLP that provided relatively high multiplex ratios to resolve the subtle variation in Arabica germplasm compared to approaches like isozymes and RFLP that generally failed to detect the same (Lashermes *et al.,* 1996a; Lashermes *et al.,* 1999). Similarly, some authors studied the effectiveness of some PCR-based molecular markers (RAPD, AFLP and SSR) (Aga *et al.,* 2005; Sousa *et al.,* 2017; Macedo *et al.,* 2020). According to their studies, they reported that all three PCR-based DNA markers used in their study proved to be useful for the characterization of Arabica coffee germplasm. Therefore, differences in marker systems used (Vuylsteke *et al.,* 1999), and played a significant role in the variations between the results our investigations. Thus, this paper focuses on these three genetic diversity studies of PCR-based markers and the recent sequence-based markers including Single nucleotide polymorphism (SNP) markers.

**RAPD Marker**

The first-ever study by Lashermes *et al.,* (1993) using PCR-based 20 RAPD markers, revealed that there were sufficient inter-specific variations between Arabica and Canephora, there was almost no detectable variation within the few Arabica accessions analyzed in their study. They further reported that genetic variability in Arabica populations is expected to have reduced further owing to its autogamous behavior, leading to genetic uniformity. Similarly, Kathurima *et al.,* (2012) studied the diversity of 24 coffee genotypes in Kenya originated from different countries (Kenya, Puerto Rico, Tanzania, Reunion, Portugal, Yemen, Guatemala, and Colombia) using 10 RAPD markers, and the narrow genetic base was reported in Arabica coffee. Thus, emphasized the need to widen the existing genetic diversity through interspecific hybridization. These results are in agreement with the work of Agwanda *et al.,* (1997) and Hue (2005) which revealed high genetic similarity between Kenyan traditional commercial varieties by using the RAPD marker. Comparatively, when we consider materials specifically targeting Ethiopia using the RAPD marker higher genetic diversity has been reported among wild coffee populations than within cultivated genotypes (Anthony *et al.,* 2000; Aga *et al.,* 2003; Masumbuko *et al.,* 2003; Masumbuko and Bryngelsson 2006; Maluf *et al.,* 2005). Similarly, Aga *et al.,*, 2003 studied Genetic diversity within the forest Arabica coffee gene pool in Ethiopia using 12 RAPD markers and reported the presence of genetic variability among forest Arabica coffee populations in Ethiopia. The information from the *Arabica coffee* diversity studies using RAPD markers clearly showed the presence of high genetic diversity in Ethiopia than in other parts of the world.

**AFLP marker**

Most scholars illustrated that the use of AFLP markers showed that a genetic alteration has been limited in Arabica coffee somatic embryogenesis (Landey *et al.,* 2013). But some of them didn't agree with the conclusion.

For instance, Sousa *et al.,* (2017) applied AFLP marker for the same sample of cultivars/progenies from the Brazil National Coffee Trial that analyzed by SSR markers. Comparatively, they found much higher variation both within and between cultivars using AFLP marker; hence, this revealed that AFLPs are much more efficient to explore the genetic variability that still exists in Arabica coffee.

Similarly, Macedo *et al.,* (2020) used four AFLP markers to assess the genetic distinctiveness of 32 Arabica coffee genotypes belonging to the National Coffee Trial of Brazil, which included cultivars released by different research centers. They reported the presence of variability among research center cultivars but not within the centers in Brazil. Regarding Ethiopian-origin materials, Dessalegn *et al.,*. (2009) studied 28 *Arabica coffee* genotypes using six AFLP markers and reported high genetic variation among the materials. The information from AFLP markers also showed the presence of higher genetic variability in Ethiopia compared to other countries.

**SSR markers**

A different set of research on SSR markers showed that the low level of genetic diversity and the narrow genetic base of the commercial cultivars of Arabica coffee. For instance, Anthony et al. (2002) conducted genetic diversity study among 15 commercial varieties compared to wild coffee accessions using six SSR markers and reported the presence of a low percent polymorphism. In line with this, a low level of genetic diversity was reported by Moncada and McCouch (2004)and Maluf *et al.,* (2005) in 12 and 26 Colombian and Brazilian cultivated Arabica coffee varieties using 34 SSR and 23 SSR markers, respectively. Additionally, Tornincasa *et al.,* (2006) reported a low level of genetic diversity in 45 commercial Arabica coffee varieties obtained from Brazil, Guatemala, India, and Africa using 23 SSR; in agreement, low genetic diversity demonstrated among 55 commercial Arabica coffee varieties of France using 32 SSR markers ( Teressa *et al.,* 2010). A low level of genetic variation was also reported by Al-Murish *et al.,* (2013) using 58 SSR markers with 17 Arabica coffee cultivars grown in Yemen. Similarly, Geleta *et al.,* (2012) using 12 SSR markers and reported the presence of a low level of genetic diversity among eight Nicaraguan commercial Arabica coffee varieties.

When we consider, the material specifically targeting Ethiopian origins Arabica coffee, a wide genetic variability was reported among 96 Ethiopian accessions using 12 SSR markers (Tornincasa *et al.,* 2006). Similarly, Tadesse *et al.,* (2020) conducted genetic diversity study on 42 Ethiopian commercial Arabica coffee cultivars using 14 SSR markers and reported that the number of rare alleles across the 28 coffee varieties ranged from one to six. The presence of rare alleles in their studied coffee varieties reflected their rich genetic diversity. This result clearly suggests the presence of a high level of genetic diversity among commercial Arabica coffee varieties currently grown in Ethiopia and also the presence of a high level of genetic diversity in Ethiopia rather than in other countries. The information from SSR markers still showed the presence of high genetic variability in Ethiopia rather than in other countries.

**SNP markers**

SNPs markers are more efficient for genetic studies in perennial crops; this is great technology for germplasm management which has not been fully utilized in coffee. However, some studies involving the detection of SNPs in coffee have been carried out (De Kochko *et al.,* 2010; Vidal *et al.,* 2010; Combes, *et al.,* 2013; Yuyama *et al.,* 2016).

Recently, Virginie *et al.,* (2019) using 8.5K SNPs array that contains 8580 unique and informative SNPs, covering the whole Arabica coffee genome, reported the largest proportion of genetic variation in Arabica coffee genotypes. Likewise, Gustavo et al., (2018) studied 107 Arabica Coffee accessions including wild genotypes from the historical FAO collection from Ethiopia analysis using SNP markers revealed that the collection of Arabica coffee used in their study has a higher genetic diversity than traditional cultivars. In this context, Ethiopian germplasm collection has been shown to be a valuable source of novel favorable bio-chemical characteristic-related alleles, which can be explored by breeding programs. Similarly, Mikiru *et al.*, 2022 studied the diversity of Ethiopian-origin Arabica coffee genotypes by using SNP markers and reported the presence of genetic variability between genotypes. As mentioned earlier, these results clearly suggest that the Arabica coffee germplasm found in Ethiopia has a broad genetic base, and is valuable in developing varieties that could sustain global coffee production.

**Applications of DNA markers for core collection identification in coffee gene banks**

The molecular characterization of coffee accessions using MAS is an accurate tool for the conservation and more efficient use of genetic resources by breeders through evaluating the redundancies and deficiencies of the germplasm that generates information on the efficiency of the collection, maintenance, and expansion of a germplasm bank (Ferrão *et al.*, 2015, Sousa *et al.*, 2017). In addition, this provides fundamental information to help breeders choose parents to integrate into cross-breeding schemes, as well as in directing the improvement of the genetic base during the course of a breeding program. Thus, different the world-leading significant germplasm resources and conservation of the Coffea genus gene bank assessed are summarized in Table 2. Source and compiled from Juliano *et al.*, (2020).

**Table 2: International coffee gene bank**

| Name of research institute | Country | Year of establishment | No of accessions | Coffea specious | Techniques used |
|---|---|---|---|---|---|
| Centre National de Recherche Agronomique (CNRA) | Ivory Coast | 1998 | 8003 | Arabica and others | Using all molecular markers and detailed use of SSR, SNP, AFLP markers |
| United States Department of Agriculture (USDA) | USA | - | 800 | Arabica an others | |
| CATIE botanical garden and germplasm | Costa Rica | 1942 | 1987 | Arabica | |
| Centro de Cooperación Internacional de Investigación Agricola para el Desarrollo (CIRAD) | Birazil | 1960s | 3800 | Arabica and others | |
| Ethiopian Institute of Agricultural Reseatrch EIAR/Jimma Agricultural Research center (JARC) | Ethiopia | 1967 | 5853 | Arabica | Not assessed yet |
| Institute of Biodiversity Conservation | Ethiopia | - | 5196 | Arabica | Not assessed yet |
| Instituto Agronômico de Campinas (IAC) | Brazil | - | 5451 | Arabica and others | Detailed use of SSR, AFLP and SNP markers and others |
| IAPAR, (EPAMIG), (UFV), (INCAPER | Brazil | - | 13,856 | Arabica and others | |
| **Total** | 44,946 | | | | |

when we consider the Arabica coffee number excluding the Ethiopians gene bank, the gene banks around the world have a collection of Arabica coffee which stands out with the most significant number of accessions (11,415), immediately succeeded by *C. canephora* (625), *C. liberica* (94), *C. eugenioides* (81) and other *Coffea species* (more than 700) (Bramel *et al.,* 2017).

**Table 3: Number of a selected core collection for the different categories of coffee genetic material that conserved under World Coffee Research**

| **Categories** | **Number of the f selected coffee core collection** |
|---|---|
| Typica/Bourbon | 458 |
| East African varieties | 132 |
| Kivu varieties | 129 |
| Ethiopian landraces | 406 |
| Sudanese landraces | 24 |
| Introgressed varieties | 1150 |
| $F_1$ hybrid varieties or experimental crosses | 234 |
| **Total** | **2533** |

Among international coffee gene banks assessed for coffee diversity around the world by using different marker-assisted selections, about 2533 genotypes were identified for coffee core collection (Pruvot *et al.,* 2020). These genotypes correspond to the core collection of the germplasm of the Tropical Agricultural Research and Higher Education Center, accessions from Southern Sudan, and cultivars/germplasm from North, Central, and South America as well as Africa and Asia. Currently, these accessions are conserved under World Coffee Research. The categories of these accessions under World Coffee Research are summarized in Table 3. When we consider Ethiopian coffee gene banks the information in this regard is scanty. Therefore to avoid the redundancy of genotypes that increase the cost of maintenance and conservation Ethiopia should develop strategies by using marker-assisted selection. In addition to that, this strategy is paramount important to generate information for coffee breeders on the efficiency of the collection, maintenance, and expansion of a germplasm bank.

**Applications of MAS for Molecular Linkage Maps of Arabica coffee**

Virginie *et al.,* (2019) reported that the generation of a high-density Arabica map is still severely hampered by its allotetraploid nature and the narrow genetic diversity among Arabica coffee accessions. A preliminary linkage map was constructed using AFLP markers on a pseudo-$F_2$ population derived from a cross between the cultivars 'Tall Mokka' and 'Catimor' (Pearl et a1. 2004).

Similarly, Nagai *et al.,* (2007) used the $F_2$ population of the same parental cultivars and identified more linkage groups. Later, an interspecific $F_2$ population of Arabica coffee and *Coffee canephora* was used for genetic map construction. As mentioned earlier, despite the situation of Arabica coffee, some mapping efforts have been developed more recently with both SSR and SNP markers (Pestana *et al.,* 2015; Moncada *et al.,* 2016; Diola *et al.,* 2011; Maria *et al.,* 2019; James *et al.,* 2021) and summarized in table 4.

**Table 4: Genetic linkage maps of coffee Arabica using MAS applications**

| Markers | Population used | Length (cM) | No of linkage groups | Traits/ purpose | Authors |
|---|---|---|---|---|---|
| AFLP | pseudo-$F_2$ population (Mokka hybrid x Catimor) | 1,802.8 | 31 | source-sink traits | Pearl *et al.*, (2004) |
| AFLP | $F_2$ of Tall Mokka and Catimor | 1,042.4 | 40 | Cupping quality and morphology | Nagai *et al.*, (2007) |
| AFLP and SSR | $F_2$ of *C. arabica* x *C. canephora* | 1,011 | 37 | Quality and productivity | Priolli *et al.*, (2009) |
| RAPD | $F_2$ of (Mundo Nov x Hybrido de Timor) | 540.6 | 8 | Partial linkage map | Teixeira Cabral *et al.*, (2004) |
| SSR | $F_2$ and $F_3$ of Caturra x CCC1046 (Wil type Ethiopian origin accession) | 3800 | 22 | Yield, plant height, and fruit size | Moncada *et al.*,(2014) |
| SSR | | 976.8 | 12 | CBD Resistant | Pestana *et al.*, (2015) |
| SSR and SNP | $F_2$ offspring from a cross between C. arabica var. Caturra and a wild C. arabica accession from Ethiopia | 3840 | 22 | CBD Resistant | Moncada *et al.*, (2016) |
| SNP | Varieties of Rume Sudan | 5525.39 | 11 | CBD Resistant | Jmes *et al.*, 2021 |
| SNP and SSR | A cross between *C. arabica var*. Caturra and a wild *C. arabica* accession from Ethiopia, (CCC1146) | 3840 | 21 | Yield, plant height, and bean size | Maria *et al.*, 2019 |
| SCAR | Natural crossed Hybrido de Timor accessions | | 9 | Rust resistance | Diola *et al.*, 2011 |

Similarly, Nagai *et al.,* (2007) used the $F_2$ population of the same parental cultivars and identified more linkage groups. Later, an interspecific $F_2$ population of Arabica coffee and Coffee canephora was used for genetic map construction. As mentioned earlier, despite the situation of Arabica coffee, some mapping efforts have been developed more recently with both SSR and SNP markers (Pestana *et al.,* 2015; Moncada *et al.,* 2016; Diola *et al.,* 2011; Maria *et al.,* 2019; James *et al.*, 2021) and summarized in Table 4.

**Applications of MAS for QTL mapping in Arabica coffee**

It is now well established that most of the plant traits of agronomical interest are quantitative, controlled by multiple genes that need to be judiciously manipulated to develop genetically improved germplasm, but hitherto were not accessible through conventional breeding approaches. With the development and availability of DNA marker-based linkage maps for many animal and plant species, identification, mapping, and selection of Quantitative Trait Loci (QTLs), have become practically feasible and achievable.

The QTLs can be identified by monitoring mapped or unmapped DNA markers in conjunction with the target trait using a segregating mapping population (Hackett, 2002), and/or diverse genotypes employing dis-equilibrium mapping (Gupta *et al.*, 2005). Though having a molecular linkage map is not a prerequisite for linking a QTL, it is essential to map them for efficient selection and subsequent transfer of the trait. In the case of *Arabica coffee*, QTL mapping is in its infancy with few successful efforts in recent years.

**Table 5: Quantitative Trait Loci and Linkage Groups in coffee using applications of MAS**

| Traits | No of QTLs and LGs | Markers | Authors |
|---|---|---|---|
| CBD resistance | 3 QTLs, qCBD 1-1, qCBD 2-1, qCBD 2-2 | 699 SNP | James *et al.*, 2021 |
| Yield, | 2 QTLs, B0013, 08832 | 338 SSR and SNP | Maria *et al.*, 2019 |
| Plant height | 1 QTLs, 10639 | 338 SSR and SNP | Maria *et al.*, 2019 |
| Lipids and CAF and KAH | and identified 21 SNP/ trait associations | 2,587 SNP | Gustavo *et al.*, 2018 |
| Rust resistance | $Q_{CLR\_4}$ | 4SSR | Glady *et al.*, 2013 |

In the first such study in coffee, few RAPD markers were linked with resistance to coffee berry disease (CBD) caused by Colletotrichum kahawae (Agwanda *et al.*, 1997), using the indirect approach of diverse genotypes (5 susceptible and 8 resistant Arabica cultivars/selections) and unmapped RAPD markers. To date, QTL analyses relating to quality compounds and cups have been largely performed on Coffee canephora and other species, but none has been reported for Arabica coffee. Some of the QTL mappings identified in Coffea arabica are summarized in Table 5.

**Applications of MAS for major Arabica coffee disease studies**

The availability of markers linked to genes of interest allows for identifying sources of resistance, even when the pathogen is absent. According to Ortega and Lopez-Vizcon (2012), when used at the appropriate stage of the breeding process, molecular markers closely related to resistance genes enable the early selection of resistant individuals. Moreover, compared to artificial inoculations, molecular markers are more rapid, inexpensive, and reliable for screening individuals with resistance genes. Thus, marker-assisted selection (MAS) is a powerful tool for increasing the efficiency of breeding programs, reducing the time required for selection, and allowing the search for durable, broad-spectrum resistance through the pyramiding of genes of interest (Gartner *et al.*, 2013; Romero *et al.*, 2014). Therefore, the progress made in MAS research on major coffee diseases is summarized.

**Coffee Berry Disease (CBD)**

Efforts using conventional breeding methods to develop resistance to CBD began in 1971, with the primary goal of creating cultivars that combine resistance to CBD, high production, good beverage quality, and desirable growth habit for high-density planting (Van Der Vossen and Walyaro, 1980). Using conventional approaches, genes for resistance to CBD were introduced to susceptible Arabica coffee varieties by crossing them with donor varieties and backcrossing to standard varieties to restore desirable attributes (Walyaro, 1983). However, this approach takes a long time to develop a coffee variety due to the long juvenile nature of the Coffea genus (Moncada *et al.*, 2016).

Within this context, some progress is made to identify resistance genes for major coffee diseases using molecular markers. Based on inheritance studies, Van Der Vossen and Walyaro (1980) reported the existence of a locus (T) for resistance to CBD using Híbrido de Timor. After a couple of years, Agwanda *et al.,* (1997) and Silva *et al.,* (2006), also reported that resistance to CBD is controlled by at least three loci (T, R, and K) present in Híbrido de Timor. They further reported the presence of those genes in different varieties. Catimor varieties contain (gene T), Rume Sudan (genes R and K), and K7 (gene K). In a study carried out by Gichuru *et al.,* (2008), the locus identified as being responsible for resistance to *Colletotrichum kahawae* was termed Ck1. Although these authors suggest that this locus is similar to the T locus described by Van Der Vossen and Walyaro (1980), they do not discard the possibility of the existence of another locus conferring resistance. Using the cultivars Catimor 88 and Catimor 127 as a source of resistance, Gichuru *et al.,* (2008) also identified eight AFLP and two SSR markers linked to the gene for resistance to CBD. The gene termed Ck-1 was found to be located in a segment of 11 cM. The CBD-resistant cultivar R11, and Batain is an $F_1$ hybrid, after several generations of selfing to fix the CBD resistant genes (Omondi *et al.,* 2001, Gichimu *et al.,* 2014). SL 28 is a Bourbon-type single-tree selection that combines high yield, high quality, and drought tolerance but is highly susceptible to CBD (Walyaro, 1983). James *et al.,* (2021) further evaluated the genetic relationship and the occurrence of multiple gene resistance to coffee berry disease within selected Arabica coffee varieties by previous researchers mentioned above in Kenya and reported that all the genotypes within the CBD-resistant varieties R11 and Batian carry the T gene, while R11 carries additionally, the R gene and therefore with multiple gene resistance to CBD. Hence coffee varieties with multiple gene resistance to CBD is a reality. The genotypes confirmed carrying the two genes for resistance to CBD are recommended for further distribution to growers since resistance will not break easily to new disease races.

**Coffee Leaf Rust (CLR)**

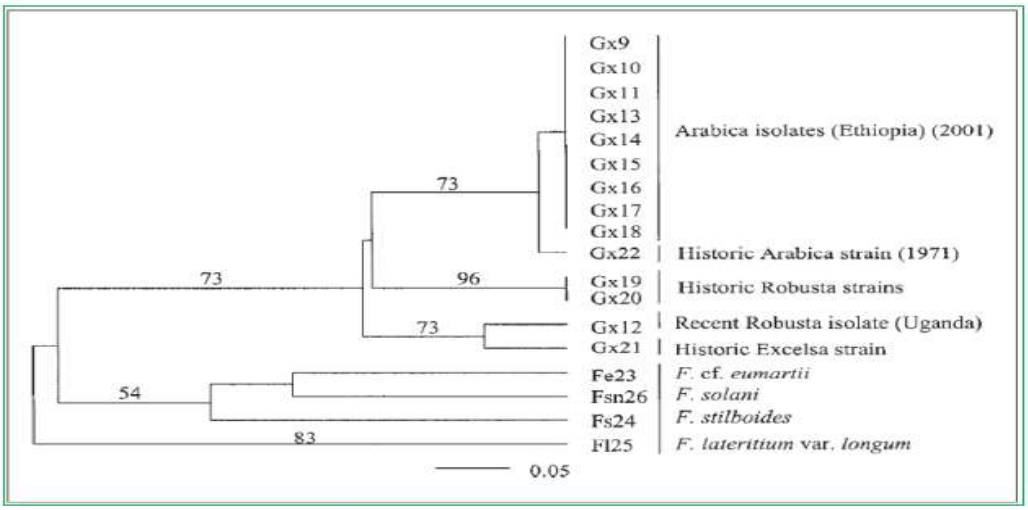
Resistance to coffee rust is conferred by at least nine dominant genes (SH1 to SH9), either singly or in combination. The resistance genes identified in Arabica coffee (SH1, SH2, SH4, and SH5) and SH5 to SH9 genes have been identified in Híbrido de Timor, a coffee plant resulting from the natural cross between Arabica coffee and Coffea canephora (Cabral *et al.,* 2009; Fernandez *et al.,* 2012; Maia *et al.,* 2013). In contrast, the SH3 gene was identified in Indian selections, which are derived from natural crosses between Arabica coffee and Coffea liberica, Coorg (Bettencourt and Rodrigues 1988; Prakash *et al.,* 2004; Ram 2006; Prakash *et al.,* 2011). Híbrido de Timor and Indian selections, which are tetraploid materials, are used to the facilitate introgression of genes of interest from species with ploidy levels (*Coffea canephora* and *Coffea liberica*) that differ from that of Arabica coffee. Indeed, genes from Coffea canephora (SH6 and SH9) and *Coffea liberica* (SH3) have provided durable resistance in coffee plants evaluated in the field (Bettencourt and Rodrigues 1988; de Brito *et al.,* 2010; Diola *et al.,* 2011). In an attempt to assist introgression of genes from other species into Arabica coffee, Prakash *et al.,* (2004) identified 21 amplified fragment length polymorphism (AFLP) markers linked to the SH3 gene derived from introgression of *Coffea liberica* into *Coffea arabica.* The locus SH3 originated from the introgression of Coffea liberica into Arabica coffee mapped by Prakash *et al.,* (2004) and characterized by Mahé *et al.,* (2008). However, the physiological races of harmerale vastratix in Brazil have already overcome the SH3 resistance gene, making the available markers not useful for breeding purposes in Brazil.

De Brito *et al.*, (2010) characterized one of the resistance genes found in Híbrido de Timor using accession UFV 427-15, which presents dominant monogenic resistance, and named the gene SH?. According to the authors, this gene corresponds to one of the SH genes (SH7, SH8, or SH9) previously identified in genotypes derived from Híbrido de Timor or another unknown gene. Later, De Brito *et al.*, (2010) and Diola *et al.,* (2011) identified molecular markers linked to the SH? gene. Diola *et al.,* (2011) also developed a high-density genetic map with six SCAR markers, delimiting a chromosomal region of 9.45 cM and flanking the SH? gene within 0.7 and 0.9 cM (Table 6).

**Coffee Wilt Disease (CWD)**

It is obvious that there is no resistance Arabica coffee varieties to CWD. However, some research had conducted to identify *Gibberella xylarioides* fussarium specious in different coffee specious including Arabica coffee. RAPD was one of the first molecular techniques used to look for differences between organisms. The technique analyses randomly chosen parts of the genome for differences between genomes of two or more organisms, and can successfully distinguish between species but is now often regarded as too unreliable when it comes to determining between strains of the same species. Out of 12 oligonucleotide primers, RAPD-PCR analysis with five oligonucleotide primers produced clear informative and reproducible polymorphic DNA banding patterns (Girma *et al.,* 2005). The analysis showed that all the recent Arabica isolates including those isolates derived from the same ascus had monomorphic RAPD amplification and clustered into a single group indicating homogeneity of the population.

**Fig 3: Clustering of Gibberella xylarioides strains from Coffea arabica, C. canephorae, C. excels and other Fusarium species (Girma et al., 2005)**




However, clear DNA polymorphism among *Gibberella xylarioides* strains from *Coffea arabica*, *Coffea canephora*, and Coffea excels with varying fragment lengths. Gibberella xylarioides were distinctly polymorphic to Fusarium stilboides, *F. solani*, F. cf. eumartii, and *F. lateritium* var. Longum (Fig. 3) conforms to the taxonomic classification of these species. Even though they originated from diverse environments like host cultivars, agro-ecological zones as well as production systems, and varied significantly in aggressiveness in the pathogenicity test.

The historic Arabica strain of 1971, however, seems to be slightly different from the recent Arabica collections which may implicate little genetic changes in the pathogen populations over the last 30 years (1971–2001). The results of RAPD-PCR markers corroborated the existence of host specialization into at least two pathogenic forms within *Gibberella xylarioides* populations. These are *Gibberella xylarioides* f. sp. Abyssiniae (anamorph: *Fusarium xylarioides* f. sp. abyssiniae) for the fungal strains attacking only Arabica coffee and *Gibberella xylarioides* f. sp. Canephorae (anamorph: *F. xylarioides* f. sp. canephorae) pathogenic to *C. canephorae* and *C. excels*.

**Applications of MAS for quality studies in coffee**

The presence of a number of species with special features in relation to quality such as low or no caffeine, high trigonelline, or low CGAs could be an important breeding focus (Thi *et al.,* 2017). However, the success of interspecific hybridization may vary due to genetic barriers between species; so genetic improvement focusing on using materials within the same species of Arabica coffee is therefore the priority. As for other crops, the determination of molecular predictors for coffee quality traits would help reduce the length of breeding selection cycles and phenotypic evaluation cost. However, the use of DNA technology in coffee quality improvement is still in its infancy. Pot *et al.*, (2007) used polymorphisms generated from SNPs, INDELs, and SSRs (simple sequence repeats) to identify the nucleotide diversity of four sucrose metabolism enzymes in Coffea canephora genotypes using direct sequencing. The variation of these genes was also analyzed between different Coffea species to allow the identification of more polymorphic sites using parallel *in silico* analysis of expressed sequence tag (EST) resources (Pot *et al.*, 2007; Thi *et al.,* 2017). AFLP (amplified fragment length polymorphism) and SSR markers were used to construct a genetic map of an $F_2$ population between *C. arabica* and *C. canephora* (artificial tetraploid) (Table 5). The number of markers associated with quality traits identified was 19 for sugar content, 8 for caffeine and chlorogenic acids (CGAs), and one for caffeine and CGAs (Priolli *et al.,* 2009). These markers need to be validated in other genotypes for consistency before they can be used in marker-assisted selection (MAS) in coffee breeding. Recently, a total of 33,239 SNPs specific to Arabica coffee and 87,271 SNPs specific to *Coffea canephora* were developed using targeted genome capture strategies and next-generation sequencing and were evaluated on 72 samples from *Coffea canephora* and 72 from Arabica coffee. These genomic resources will support genome assemblies, accelerate the breeding of interesting traits, and manage genetic diversity in coffee species (Resende *et al.,* 2016). When we consider, the identification of genes related to quality a number of coffee candidate genes have been identified and some of them have been cloned and characterized. These results are useful to the coffee genetics community, especially those on genes encoding the enzymes of key metabolic processes. These are candidate genes that may control the variability of coffee quality (Leroy *et al.,* 2006). Genes regulating the main chemical components that are thought to be involved in the flavor and sensory quality of *Coffea arabica* are listed in table 6. Although several genes encoding the biosynthesis of biochemical compounds in coffee have been identified in *Arabica coffee* and *Coffea canephora*, there are no genes identified for trigonelline synthesis and no studies on allelic variation associated with low and high levels of the key biochemical compounds which can be utilized in MAS (Thi *et al.,* 2017). Sequences from the known genes can serve as useful references in re-sequencing to detect polymorphisms for genetic mapping of candidate genes contributing to genetic variation in biochemical compounds.

**Applications of MAS for Arabica coffee introgression lines**

Introgressed Arabica genotypes derived from the Timor Hybrid (interspecific hybridization between *Coffea arabica* and *Coffea canephora*) were analyzed for the presence of coffee canephora genetic material using the amplified fragment length polymorphism (AFLP) approach (Vos *et al.*, 1995). In order to gain insights into the mechanism of introgression in Arabica coffee, Lashermes *et al.*, (2000) estimated the amount of introgression percent in such material. The Timor Hybrid-derived genotypes were evaluated using 42 different AFLP primer combinations, and compared to 23 accessions of Arabica coffee and 8 accessions of *Coffeacanephora* and reported 8% to 27% of the Coffea canephora genome in Hybrido de Timor varieties. Some findings showed the introgression of different coffee species for desirable traits, such as disease resistance, and genetic diversity using MAS (Table 7). Different sets of the research reported the genetic diversity observed in the Timor Hybrid-derived genotypes appeared approximately double that in Arabica coffee. Although representing only a small proportion of the genetic diversity available in *Coffea canephora*, and the Timor Hybrid obviously constitutes a considerable source of genetic diversity for Arabica breeding (Leshermase et al., 2000; J.C Hererra *et al.*, 2002; Prakash *et al.*, 2004; Tesfaye *et al.*, 2009; Gichimu *et al.*, 2016 ). Recent research conducted by Christophe *et al.*, (2020) studied on introgression of coffee Arabica maintained by world coffee research (WCR) by a DNA fingerprinting database composed of 2533 Arabica samples gathered since 2014 using SSR molecular markers with relevant genetic authentication methods for Arabica coffee. The team reported that the precise authentication of varieties depends on the degree of fixation (selfing generation). Those varieties are located at the bottom of the tree, below the orange dot line as Marsellesa, CR95, or Lempira are fixed and homogeneous lines that are easy to authenticate (Fig 4). On the other hand, Small deviations from the reference can be clearly identified as residual segregation that always occurs even in well-fixed varieties. The Cat129 fingerprint is more a single branch of the tree than a single reference. Still, this branch is individualized from other varieties. In contrast, high deviation from the reference can be clearly identified as that always occurs even in not well-fixed introgressed varieties.

**Fig 4: Neighbor-Joining tree from the single allelic data (0/1) of 2533 C. arabica samples from the WCR DNA fingerprinting database. The area of some genetic categories or positions of some varieties is indicated. One point might represent more than one sample**.

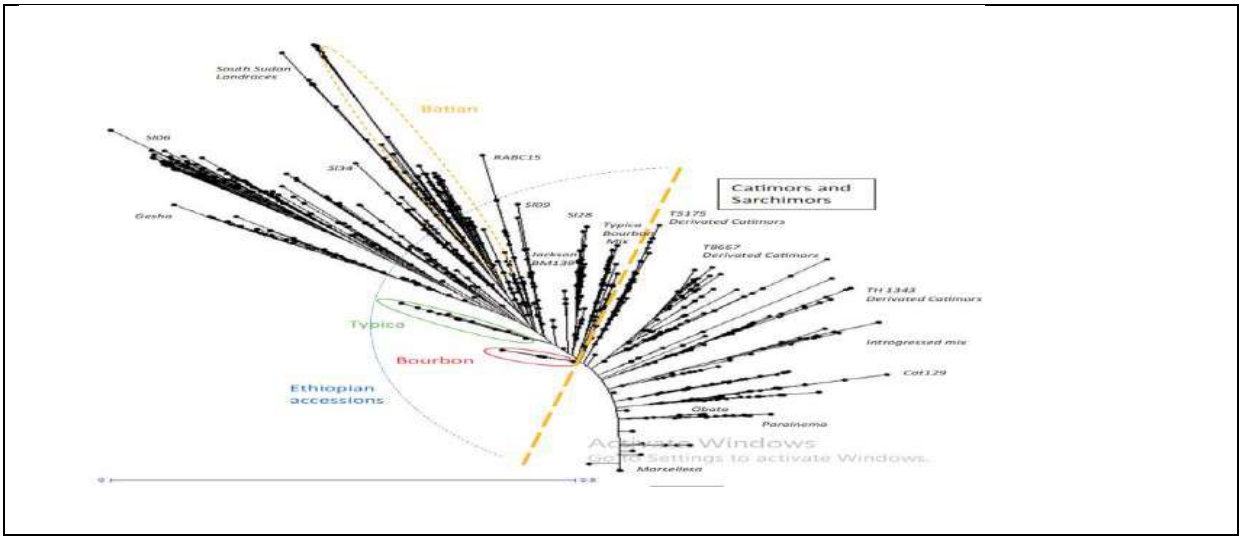


**Table 7: Some studies of introgressed Arabica coffee populations using MAS applications**

| Markers | Population used | Results | Traits/ purpose | Authors |
|---|---|---|---|---|
| AFLP | An $F_2$ Matari, an arabica accession x liberica-introgressed line S.288 | The progeny segregated for the $S_H3$ gene in a 3:1 ratio, as expected for a single dominant gene | Rust resistance | Prakash *et al*, 2004 |
| SSR | Ruiru 11 is a composite of sixty-six (66) $F_1$ hybrid sibs | Ruiru 11 sibs have canephora genome range from 8.7 to 24.14% and R11-11, R11-22, R11-107, and R11-121 had also good resistance to CBD | CBD resistance | Gichimu *et al*., 2016 |
| AFLP, RAPD, and SSR | Hybrido de Timor population (101 cultivars included) | 1 8.7 % of the Hybrido de Timor Arabica coffee genome introgressed from canephora and the high diversity of the Hybrido de Timor population | Genetic diversity study | Tesfaye *et al*., 2009 |
| AFLP | 19 Arabica coffee introgression lines ($BC_1F_4$) and two accessions derived from a spontaneous interspecific cross (i.e. Timor Hybrid) | Introgressed genotypes were estimated to represent from 9% to 29% of the *C. canephora* genome | Genetic diversity study | Phillipe *et al*., 2000 |
| SSR and AFLP | Interspecific triploid hybrid plants between the tetraploid species *Coffea arabica* L. and the diploid species *C. canephora* P. were backcrossed to *C. Arabica* | High resistance to leaf rust was obtained | Leaf rust resistance | J.C Hererra *et al*., 2002 |
| AFLP | Timor Hybrid | The presence of coffee canephora genetic material in Timor hybrid was reported for the first time | Diversity study of *Arabica coffee* materials | (Vos *et al.*, 1995) |
| AFLP | Timor Hybrid (832-1, 832-2, and 1343) and commercial arabica cultivars (progenies 832-1 and 1343), and 19 introgression arabica lines (BC1F4) | Estimates represent 8% to 27% of the coffee canephora genome in Hybrido De Tiomor | Genetic diversity study of *Arabica coffee* | Lashermes *et al*. 2000 |
| SSR | 2533 Arabica coffee conserved in world coffee research (WCR) | The precise authentication of varieties depends on the degree of fixation (selfing generation) Marsellesa, CR95 or Lempira are fixed and homogeneous lines that are easy to authenticate | Genetic authentication method for *Arabica coffee* | Christophe *et al*, 2020 |

**Current status of MAS application in Arabica coffee, achievements and limitations**

When we consider the current status of the application of MAS technologies in Arabica coffee, it is best to consider the achievements made and limitations with the developed marker for coffee genomic study. Currently, there are achievements that present opportunities to develop coffee research through marker-assisted selection (MAS). Over the past thirty years, MAS research on coffee has pursued major objectives, such as identifying important genes that code for desirable traits through a functional genomics approach and developing high-throughput sequencing (HTS) technologies that allow the rapid acquisition of significant amounts of sequence data. These advancements have also increased our understanding of the genomics of particular species. As mentioned earlier, various molecular markers, such as restriction fragment length polymorphism (RFLP), random amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP), inter-simple sequence repeat (ISSR), simple sequence repeats (SSR), and expressed sequence tag derived simple sequence repeats (EST-SSR), have been used in coffee genetic studies (Davis *et al.*, 2006; Volk *et al.*, 2020). More recently, a new type of molecular marker known as a sequence-related amplified polymorphism (SRAP) has been used in the genetic diversity analysis of coffee cultivars and species (Benti, 2017). SRAP markers were also successfully used to discriminate between parents in hybrid identification (Teferi, 2019) and therefore have great potential in coffee breeding programs. In addition to the above, single-nucleotide polymorphisms (SNPs) and PCR-RFLP markers have been used in coffee genome analysis, This revealed that in Arabica coffee, polymorphisms are created by paralogous chromosomes, whereas the homozygosity of many genes is maintained by the self-fertile nature of the species (Teferi, 2019). However, not all these resources are published or available, which limits their use. A few achievements of the current contributions of marker-assisted selection for coffee that can be considered as opportunities are summarized in table 8. This summary will be very important for designing strategies and making decisions in breeding programs, as well as in sequencing projects and identifying candidates' genes for desired traits. On the other hand, when we consider the limitation, there are currently significant challenges to overcome in Arabica coffee MAS research. Among them, there are some other limitations that hamper the development of MAS research in Arabica coffee, particularly in Ethiopia, a place known for its high genetic diversity of Arabica coffee. Unfortunately, not all of these limitations can be covered in this review paper.

**Absence of urgent institute measures**

The research on genetics and genomics for Arabica coffee is relatively very limited compared to other crops (Thi *et al.,* 2017) and thus belies its potential and economic contribution. The reasons could possibly be due to the lack of funding. Most coffee-growing regions are developing countries. The complexity of the quality traits and the limitations of the technology used have impact on the advancement of the crop improvement. This clearly shows the absence of strong institute measures in most coffee-growing countries particularly in Ethiopia because they are developing countries. However, there are some institutes established for MAS coffee research in some countries like Brazil, USA, Italy, Kenya, etc. but there is not strong enough to change MAS Arabica coffee because they have low Arabic coffee diversity in these countries. Most researchers analyzed a small number of genotypes derived from the two botanical varieties ("Typica" and "Bourbon") and Hybrido de Timor, which is a spontaneous interspecific hybrid between *Arabica coffee* and *coffee canephora*.

**Table 8: Some achievements of molecular markers in coffee**

<table>
<tr><th>Titles</th><th>Achievements</th></tr>
<tr><td>Coffee diversity</td><td>A large number of commercial and non-commercial coffee samples of American, Indian, and African origin were also analyzed using highly polymorphic PCR-based and sequence-based markers revealed that Indian cultivars were genetically diverse from the American and African cultivars, except those from Ethiopia. A recent study showed that Ethiopian-origin commercial and accessions genotypes are more diverse than those from other parts of the world (Donkor et. al., 2020).</td></tr>
<tr><td>Construction of large insert genomic libraries</td><td>C. arabica cultivar 'IAPAR 59' was used for HAC library construction for its resistance to leaf rust and root-knot nematode<br>The HAC library consists of 88,813 clones with an average insert size of 130 kilobases (kb) In parallel with the development of the EST database, Biotechnology Research International 5 BAC libraries of coffee species, C. Arabica, and C. canephora were established</td></tr>
<tr><td>Coffee-expressed sequence tags</td><td>A total of 246,500 ESTs have been reported with 69,801 for C. canephora, 166,133 for C. Arabica, and 10,566 for C. racemosa<br>Very recently, the Italian group has generated an additional 161 660 ESTs which will be publicly available on the website (http://www.coffeedna.net/)<br>the Brazilian government funded an ambitious coffee genome program, and this has resulted in the establishment of 200 000 ESTs which led to the identification of 30000 genes</td></tr>
<tr><td>Candidate genes</td><td>Different candidate genes for desired traits were identified
<table>
<tr><th>Trait</th><th>Gene</th><th>Varieties/species</th></tr>
<tr><td>Yield</td><td></td><td>ccc1146</td></tr>
<tr><td>Dwarf</td><td>Ct</td><td>Catura</td></tr>
<tr><td>CBD resistance</td><td>T</td><td>R11-123, R11-195, CR30-809</td></tr>
<tr><td>CR30-809</td><td>R</td><td>R11-123, R11-195,</td></tr>
<tr><td></td><td>Ckl</td><td></td></tr>
<tr><td>Rust resistance</td><td>SH11 – SH4</td><td>Coffee Arabica</td></tr>
<tr><td></td><td>SH – 3</td><td>Coffee liberica</td></tr>
<tr><td></td><td>SH5 – SH9</td><td>Coffee canephora</td></tr>
<tr><td>Quality</td><td>K</td><td>K-7</td></tr>
</table></td></tr>
<tr><td>Core collection</td><td>About 21,00 germplasm assessed – 2253 core collection identified<br>Anyone can witness as many million dollars was saved per annum to maintain the materials</td></tr>
</table>

There is no strong research that specifically targets Ethiopian-origin coffee the place where high Arabica coffee genetic diversity is present. There is a compelling urgency to institute measures that ensure funding of coffee molecular research, especially in most coffee-producing countries particularly Ethiopia the place where Arabica coffee genetic diversity is present. To ensure this strategy, the major hindrances to weak institutions include inadequate investment; sub-optimal human resources; inability to innovate as evidenced in the prevailing inadequate deployment of appropriate science and technology; inadequate infrastructure; and poor policy regimes. Coffee improvement, by fostering genetic gains that aid production through enhanced productiveness, is a very critical component of the development of MAS research in coffee. Therefore, to achieve this, MAS in coffee must be re-oriented in a number of very critical ways.

**Limit application of transfer candidate genes through available biotechnological tools**

It was evident from the earlier research in coffee that, to date, several candidate genes have been identified for disease resistance and other desired traits and many more will be known in future. For instance, the genes identified from coffee canephora provide the main source of disease and pest resistance traits not found in arabica coffee, including coffee leaf rust (*H. vastatrix*), coffee berry disease (*C. kahawae*), and root-knot nematode (*Meloidogyne spp*.) (Table 8). Likewise, other coffee species are of considerable interest in this respect. Besides, *C. liberica* has been used as a source of resistance to leaf rust, while *C. racemosa* constitutes a promising source of resistance to the coffee leaf miner. The exploitation of such genetic resources has so far relied on conventional procedures, in which a hybrid is produced between an outstanding variety and a donor genotype carrying the trait of interest. The progeny is then backcrossed to the recurrent parent in most coffee-producing countries. Undesirable genes from the donor parent are gradually eliminated through selection. However, conventional coffee breeding methodologies face considerable difficulties. In particular, these include the long generation time of a coffee tree (5 years), the high cost of field trials, and the lack of accuracy in the current strategy. A minimum of 25 years after hybridization is required to restore the genetic background of the recipient cultivar, ensuring the good quality of the improved variety. It is evident that the development of advanced Arabica coffee varieties using MAS (Marker-Assisted Selection) and other biotechnological tools in the international coffee sector particularly in most coffee-growing countries lags significantly behind relative to other crops (Merga *et al.*, 2020).

**Lack of resources, training, and capabilities**

The key limitations of Arabica coffee MAS (Marker-Assisted Selection) research include a lack of resources, training, and capabilities in most of the world's Arabica coffee improvement programs through molecular breeding, particularly in Ethiopia. Therefore, it is important to expand the scope and access to new marker platforms to provide efficient and cost-effective screening services to Arabica coffee breeders in most developing coffee-producing countries. Communication and mechanisms for the delivery of materials to these breeders must also be developed. There is an urgent need to enhance the capacity of coffee breeding programs to adopt new strategies. The clearly documented high rate of return on such investments in the past should be considered for coffee and other crops globally (J. M. Alston *et al.,* 2000). When evaluating MAS investments in other crops, concerns about food security and the likely impact of environmental change on food production have added urgency to accelerating the rates of genetic gain in breeding programs. Further technological developments are essential, and a major challenge will be ensuring that the technological advances already achieved are effectively deployed.

Arabica coffee is a stimulant beverage crop that belongs to the family Rubiaceae. It is a self-pollinating species with a common outcrossing rate of less than 10%, which is sufficient to induce some variation in offspring and free-pollinating cultivars. Ethiopia is both the center of origin and diversification for Arabica coffee. Despite the high genetic diversity in the Arabica coffee population, which provides immense opportunities for improvement programs, research on Arabica coffee MAS (Marker-Assisted Selection) is limited. To date, most Arabica coffee research has been conducted through conventional (traditional) methods, particularly in developing countries that produce large quantities of Arabica coffee.

In conclusion, in the last three decades, conscious efforts have begun globally to integrate MAS-based technologies, which provide impetus, dependability, and directionality to genetic improvement efforts for coffee. Several DNA marker studies have been conducted on Arabica coffee, covering different topics. Studies on the origin and evolution of Arabica coffee suggest that *Coffea eugenioides* is the possible maternal progenitor of Arabica coffee, while *Coffea canephora* was the putative pollen donor in the hybridization event that preceded the emergence of Arabica coffee. Almost all MAS studies of Arabica coffee provide evidence that the birthplace of Arabica coffee is likely Ethiopia. Research on the phylogenetic studies of Arabica coffee, using various marker-assisted selections, suggests that southwestern Ethiopia is its origin. Numerous studies on Arabica coffee diversity conducted via MAS have reported a high level of genetic diversity in Ethiopia compared to other countries. This diversity is valuable for developing varieties that could sustain global coffee production. However, information on this topic remains scarce in Ethiopia. Similarly, about 21,000 Arabica coffee germplasm accessions were assessed among world-leading gene banks using MAS, and 2,533 genotypes were identified for core collections. MAS linkage and QTL mapping have been developed for different desirable traits, though they are in their infancy, with few successful examples. Various candidate genes have been identified for major coffee diseases: for CBD (T, R, K, and ckl-1 genes), for CLR (SH1-SH9), and for CWD, which is caused by *Gibberella xylarioides* f. sp. abyssiniae (anamorph: *Fusarium xylarioides* f. sp. abyssiniae). Candidate genes that may control the variability of coffee quality have also been identified through MAS. While there have been some achievements in the application of MAS in coffee research, there are significant challenges to overcome, especially in Ethiopia, where a high diversity of Arabica coffee is recorded. Unfortunately, the specific challenges were not specified in this seminar.

**Recommendation and future prospective**

Coffee breeding in developing countries, particularly Ethiopia, must be supported by adequate policies, including those that promote innovation and investment. To address and reverse the worrisome trend of declining capacities in the application of MAS (Marker-Assisted Selection) for Arabica coffee improvement, a new generation of Arabica coffee breeders must be trained. Equally important are successful partnerships, including public-private sector synergies, to ensure that 21st-century coffee breeding efforts bear fruit. To achieve these strategies, MAS in coffee must be re-oriented in several critical ways. These include the establishment of strong institutions equipped with adequate infrastructure, the development of robust policy frameworks, the application of new biotechnological tools, and other supportive measures. The following research areas are crucial for solving the challenges in applying MAS to Arabica coffee, particularly in Ethiopia, where there is significant genetic diversity: Analysis of Arabica coffee genetic diversity, identification of genes responsible for desired traits, application of identified genes using modern biotechnological tools in improvement programs and development of large genome libraries for Arabica coffee genome studies through MAS.

**References**

1. Aga, E., Bryngelsson, T., Bekele, E and Salomon B. 2003. Genetic diversity of forest arabica coffee (*Coffea arabica* L.) in Ethiopia revealed by random amplified polymorphic DNA (RAPD) analysis. Hereditas, 138(1):36–46.

2. Agwanda, C., Lashermes, P., Trouslot, P., Combes, M.C and Charrier A. 1997. Identification of RAPD markers for resistance to coffee berry disease, *Colletotrichum kahawae*, in arabica coffee. Euphytica, 97: 241–248.
3. Alkimim, E.R., Caixeta, E.T., Sousa, T.V., Da Silva, F.L., Sakiyama, N.S and Zambolim L.2018. High-throughput targeted genotyping using next-generation sequencing applied in *Coffea canephora* breeding. Euphytica, 214:1-8.
4. Alkimim, E.R., Caixeta, E.T., Sousa, T.V., Pereira, A.V., de Oliveira, A.C.B., Zambolim, L and Sakiyama N.S. 2017. Marker-assisted selection provides Arabica coffee with genes from other Coffea species targeting multiple resistance to rust and coffee berry disease. Mol. Breed., 37(6): 1-10.
5. Al-Murish, T. M., Elshafei, A. A., Al-Doss, A. A and Barakat M. N. 2013. Genetic diversity of coffee (*Coffea Arabica* L.) in Yemen via SRAP, TRAP and SSR markers. J. Food, Agril. Environ., 11 (2): 411–416.
6. Anthony, F., Combes, M. C., Astorga, C., Bertrand, B., Graziosi, G and Lashermes P. 2002. The origin of cultivated *Coffea Arabica* L. varieties revealed by AFLP and SSR markers. Theor. Applied Genet., 104 (5): 894–900.
7. Baruah, A., Naik, V., Hendre, P., Rajkumar, P and Aggarwal R. K. 2003. Isolation and characterization of nine microsatellite markers from *Coffea arabica* L. showing wide cross-species amplification. Mol.Eco. Notes, 3(4): 647–650.
8. Bawin, Y.T. R., Ariane, S., Annelies, H., Piet, S., Jean, C., Ithe, M.M., Isabel, R.-R., Olivier , H and Steven B. J. 2020. The phylogenomic analysis clarifies the evolutionary origin of *Coffea arabica* L. J. Systematics Evol., 59(5): 953-963.
9. Bayetta, B. 1986. Exploration and collection of coffee germplasm from Gambella plain. IAR newsletter. Addis Ababa, 1 (2): 3-5.
10. Bayetta, B. 2001. Arabica coffee breeding for yield and resistance to coffee berry disease (*Colletotrichum kahawae* Sp.nov.). A Ph.D. degree thesis submitted to the University of London.
11. Benti, T., Gebre, E., Tesfaye, K., Berecha, G., Lashermes, P., Kyallo, M and Kouadio Y. N. 2021. Genetic diversity among commercial arabica coffee (*Coffea arabica* L.) varieties in Ethiopia using simple sequence repeat markers. J. Crop Imp., 35(2), 147-168.
12. Benti, T. 2017. Progress in Arabica Coffee Breeding in Ethiopia: Achievements, Challenges and Prospects. Int. J. Sci. Basic Appl. Res., 33(2):15-25.
13. Berthaud, J and Charrier A. 1988. Genetic resources of Coffea. In Coffee Agronomy (Clarke, R.J. and Macrae, R. eds),. Londres: Elsevier Applied Sci., 4: 1–42.
14. Berthou, F., Mathieu, C and Vedel, F. 1983. Chloroplast and mitochondrial DNA variation as an indicator of phylogenetic relationships in the genus Coffea L. Theor Appl Genet., 65:77–84.
15. Bettencourt, A and Rodrigues, C.1988. Principles and practice of coffee breeding for resistance to rust and other diseases. In: Clarke R, Macrae R (eds) Coffee agronomy. Elsevier Applied Science, London, p. 199–234.
16. Bobadilla, L. R., Cenci, A., Georget, F., Bertrand, B., Camayo, G., Dechamp, E., Herrera, J.C., Santoni, S., Lashermes, P., Simpson, J and Etienne H. 2013. High genetic and epigenetic stability in Coffea arabica plants derived from embryogenic suspensions and secondary embryogenesis as revealed by AFLP, MSAP and the phenotypic variation rate. PLOS one, 8(2), p.e56372.
17. Bramel, P., Krishnan, S., Horna, D., Lainoff, B. and Montagnon C. 2017. Global conservation strategy for coffee genetic resources. 1st ed. Bonn: Crop Trust, p72 .
18. Macedo,C.R., de Godoy, S.M., Ruas, E.A., Góes, B.D., Chaves, C.L., Ruas, C.F., Sera, T., Sera, G.H and Ruas P.M. 2020. Genetic diversity based on AFLP

markers in germplasm of the Brazilian national *Coffea arabica* trial. DOI http://dx.doi.org/10.4238/gmr18772.

19. Cabral, P.G.C., Zambolim, E.M., Zambolim, L., Lelis, T.P., Capucho, A.S and Caixeta E.T. 2009. Identification of a new race of *Hemileia vastatrix* in Brazil. Australas Plant Dis. Notes, 4:129–130.
20. Campa, C., Noirot, M., Bourgeois, M., Pervent, M., Ky, C., Chrestin, H., Hamon, S. and de Kochko A. 2003. Genetic mapping of a caffeoyl-coenzyme A 3-0-methyltransferase gene in coffee trees. Impact on chlorogenic acid content. Theor. Appl. Genet., 107: 751–756.
21. Christophe, M., Sarada, K., William, S.,Tim, S., Lucile ,T., Benoit, B and Sole`ne P.W.2020. Authentication of Coffea arabica Varieties through DNA fingerprinting and its significance for the Coffee Sector, 2020. J. Aoac Int., 103(2), 325–334.
22. Combes, M.C., Dereeper, A., Severac, D., Bertrand, B and Lashermes P. 2013. Contribution of subgenomes to the transcriptome and their intertwined regulation in the allopolyploid *Coffea arabica* grown at contrasting temperatures. New Phytol., 200:251-260.
23. Cros, J., Combes, M.C., Trouslot, P., Anthony, F., Hamon, S., Charrier, A and Lashermes P. 1998. Phylogenetic analysis of chloroplast DNA variation in Coffea L. Mol. Phyl. Evol., 9:109–117.
24. Alemayehu, D. 2017. Review on Genetic Diversity of Coffee (*Coffea Arabica*. L) in Ethiopia. Intern. J. Forestry Horti., 3(2): 1-10.
25. Davis, A.P., Govaerts, R., Bridson, D.M. and Stoffelen P. 2006. An annotated taxonomic conspectus of the genus Coffea (Rubiaceae). Bot. J. Linn Soc., 152:465-512.
26. de Brito, G.G., Caixeta, E.T., Gallina, A.P., Zambolim, E.M., Zambolim, L., Diola, V and Loureiro M.E. 2010. Inheritance of coffee leaf rust resistance and identification of AFLP markers linked to the resistance gene. Euphytica, 173: 255–264.
27. De Kochko, A., Akaffou, S and Andrade A.C. 2010. Advances in Coffea genomics. Adv. Bot. Res., 53: 23-63.
28. Denoeud, F., Carretero-Paulet, L., Dereeper, A and Droc G. 2014. The coffee genome provides insight into the convergent evolution of caffeine biosynthesis. Sci., 345: 1181-1184.
29. Dessalegn, Y., Herselman, L and Labuschagne M. T. 2009. Comparison of SSR and AFLP analysis for genetic diversity assessment of ethiopian Arabica coffee genotypes. South African J. Plant Soil, 62 (2): 119–125.
30. Dinesh, K. P., Shivanna, M. B and Santa R. A. 2011. Identification of RAPD (Random Amplified Polymorphic DNA) markers for Ethiopian wild *Coffea arabica* L. genetic resources Conserved in India. The IIOAB J., 2(4): 1–7.
31. Donkor, E.F. Ohene-Asare, D. and Adjei, R.R. 2020. Association and variability studies for yield and yield components of robusta coffee hybrids (*Coffea canephora*) J. Genet. Geno. Plant Breed. 2020; 4:103-111
32. Diola, V., de Brito, .G.G., Caixeta, E.T., Maciel-Zambolim, E., Sakiyama, N.S and Loureiro, M.E. 2011. High-density genetic mapping for coffee leaf rust resistance. Tree Genet Genom., 7:1199–1208.
33. Fazuoli, L.C., Gallo, P.B., Martins, A.L.M., Guerreiro-Filho, O., Medina-Filho, H.P., Bordignon, R. and Gonçalves W. 2001. Efficiency of early selection in the Icatu coffee. 19th conference of ASIC, Trieste, Italy
34. Fernandez, D., Tisserant, E., Talhinhas, P., Azinheira, H., Vieira, A., Petitot, A.S., Loureiro, A., Poulain, J., Da Silva, C., Silva Mdo, C and Duplessis S. 2012. 454-pyrosequencing of *Coffea arabica* leaves infected by the rust fungus *Hemileia vastatrix* reveals in planta-expressed pathogen-secreted proteins and plant functions in a late compatible plant-rust interaction. Mol. Plant Pathol., 13: 17–37.
35. Ferrão, L.F. V., Caixeta, E.T., Pena, G., Zambolim, E.M., Cruz, C.D., Zambolim

L, *et al.,* 2015. New EST–SSR markers of *Coffea arabica*: transferability and application to studies of molecular characterization and genetic mapping. Mol. Breed. 35:31.

36. Gartner, G.A.L., McCouch, S.R and Moncada M.D.P. 2013. A genetic map of an interspecific diploid pseudo testcross population of coffee. Euphytica, 192:305–323.
37. Geleta, M., Herrera, I., Monzón, A and Bryngelsson T. 2012. Genetic diversity of arabica coffee (*Coffea arabica* L.) in Nicaragua as estimated by simple sequence repeat markers. The Sci. World J., 2012: 1–11.
38. Geromel, C., Ferreira, L. P., Guerreiro, S. M. C., Cavalari, A. A., Pot, D., Pereira, L. F. P., Leroy, T., Vieira, L. G. E and Marraccini P. M. 2006. Biochemical and genomic analysis of sucrose metabolism during coffee (*Coffea arabica*) fruit development. J. Exper. Bot., 57: 3243–3258.
39. Getachew,W-M., Sentayehu, A., Leta, T and Gezahegn B. 2020. Genetic variability of Ethiopian Coffee (*Coffea arabica* L.) accessions collected from East Wollega Zone for bean biochemical constituents. Ethiop. J. Agric. Sci., 30(3): 77-97.
40. Gichimu, B. M., Gichuru, E. K., Mamati, G. E and Nyende A. B. 2014. Occurrence of Ck-1 gene conferring resistance to Coffee Berry Disease in *Coffea arabica* cv. Ruiru 11 and its parental genotypes. J. Agri. Crop Res., 2(3): 51-61.
41. Gichuru, E. K., Agwanda, C. O., Combes, M. C., Mutitu, E. W., Ngugi, E. C. K., Bertrand, B and Lashermes P. 2008. Identification of molecular markers linked to a gene conferring resistance to Coffee berry disease (*Colletotrichum kahawae*) in *Coffea arabica.* Plant Patho., 57:1117–1124.
42. Gimase, J.M. 2014. Genetic diversity of Arabusta coffee (*C. arabica* L. x *C. canephora* P.) and their parental genotypes. M.Sc. Thesis, Kenyatta University, Kenya.
43. Girma, A., Hindorf, H., Steiner, U., Nirenberg, H. I., Dehne, H. W and Schellander K. 2005. Genetic diversity in the coffee wilt pathogen (*Gibberella xylarioides*) populations: differentiation by host specialization and RAPD analysis. J. Plant Dis. Pro., 112 (2): 134-145.
44. Romero, G., Vásquez, L.M., Lashermes, P. and Herrera, J.C., 2014. Identification of a major QTL for adult plant resistance to coffee leaf rust (Hemileia vastatrix) in the natural Timor hybrid (*Coffea arabica* x *C. canephora*). Plant Breed., 133(1):121-129.
45. Gomez, C., Desinoy, M., Hamon, S., Hamon, P., Salmon, D., Akaffou, D.S., Legnate, H., de Kochko, A., Mangeas, M and Poncet V. 2016. Shift in precipitation regime promotes interspecific hybridization 20 of introduced *Coffea species*. Ecology Evol., 6: 3240–3255.
46. Gustavo, C., Sant'Ana, Luiz, F. P., David, P., Suzana, T. I., Douglas, S. D., Rafaelle, V. F., Natalia, F. P., Bruna, S. R., Lívia, M. N., Cintia, S. G. K., Maria, B. S. S., Fernanda, F. O., Gustavo, H. S., Lilian, P., Jean-Pierre, L., Romain, G., Pierre, C and Thierry L. 2017. Genome-wide association study reveals candidate genes infuencing lipids and diterpenes contents in *Cofea arabica* L. Scienti. Reports, 8(1):465.
47. Guyeux, C., Charr, J.C., Tran, H.T.M., Furtado, A., Henry, R.J., Crouzillat, D., Guyot, R and Hamon P. 2019. Evaluation of chloroplast genome annotation tools and application to the analysis of the evolution of coffee species. PLoS One 14: e0216347.
48. Hamon, P., Siljak-Yakovlev, S., Srisuwan, S., Robin, O., Poncet, V., Hamon, S and De Kochko A. 2009. Physical mapping of rDNA and heterochromatin in chromosomes of 16 Coffea species: a revised view of species differentiation. Chromosome Res., 17(3): .291-304.
49. Herrera, J.C., Combes, M.C., Anthony, F., Charrier, A. and Lashermes, P., 2002.

Introgression into the allotetraploid coffee (*Coffea arabica* L.): segregation and recombination of the C. canephora genome in the tetraploid interspecific hybrid (*C. arabica*× *C. canephora*). Theore. Appl. Genet., 104:661-668.

50. James, M. G., Wilson, M. T and Chrispine O. 2021. Pyramiding of genes conferring resistance to coffee berry disease using marker assisted selection. An M.Sc. thesis submitted to the school of graduate studies of Kenya University, p149.
51. Joët, T., Bertrand, B and Dussert S. 2014b. Environmental effects on coffee seed biochemical composition and quality attribute a genomic perspective. In "The 25th International conference on coffee and science ASIC 2014", Colombia.
52. Joët, T., Laffargue, A., Descroix, F., Doulbeau, S., Bertrand, B., kochko, A. d and Dussert S. 2010. Influence of environmental factors, wet processing and their interactions on the biochemical composition of green Arabica coffee beans. Food Chem., 118: 693-701.
53. Joët, T., Laffargue, A., Salmona, J., Doulbeau, S., Descroix, F., Bertrand, B., Kochko, A. d and Dussert S. 2009. Metabolic pathways in tropical dicotyledonous albuminous seeds: *Coffea arabica* as a case study. New Phytologi., 182: 146–162.
54. Joët, T., Laffargue, A., Salmona, J., Doulbeau, S., Descroix, F., Bertrand, B., Lashermes, P and Dussert S. 2014a. Regulation of galactomannan biosynthesis in coffee seeds. J. Exp.Bot., 65:323–337.
55. Joët, T., Salmona, J., Laffargue, A., Descroix, F and Dussert S. 2010. Use of the growing environment as a source of variation to identify the quantitative trait transcripts and modules of co- expressed genes that determine chlorogenic acid accumulation. Plant, Cell Enviro., 33: 1220- 1233.
56. Juliano, L.F., Eveline, T.C., Fernanda, F.C., Tesfahun, S., Gustavo, C.S and Leila M. F. 2020. Genetic Diversity of *Coffea arabica*. IntechOpen. Available at: http://dx.doi.org/10.5772/intechopen.94744.
57. Kathurima, C.W., Kenji, G.M., Muhoho, S.M., Boulanger, R., Gichimu, B.M and Gichuru E.K. 2012. Genetic diversity among commercial coffee varieties, advanced selections and museum collections in Kenya using molecular markers. Intl. J. Biodiver Conserv., 4 (2): 39-46.
58. Kato, M and Mizuno K. 2004. Caffeine synthase and related methyltransferases in plants. Front. Biosci., 1: 1833-1842.
59. Koshiro, Y., Zheng, X. Q., Wang, M.L., Nagai, C and Ashihara H. 2006. Changes in content and biosynthetic activity of caffeine and trigonelline during growth and ripening of *Coffea arabica* and *Coffea canephora* fruits. Plant Sci., 171: 242-250.
60. Lashermes, P., Combes, M.C., Trouslot, P and Charrier A. 1997. Phylogenetic relationship of coffee-tree species (*Coffea arabica* L.) as inferred from ITS sequences of nuclear ribosomal DNA. Theor Appl. Genet., 94:947–955.
61. Lashermes P., Cros, J,, Combes, M.C., Trouslot, P., Anthony, F., Hamon, S and Charrier A. 1996. Inheritance and restriction fragment length polymorphism of chloroplast DNA in the genus Coffea arabica L. Theor Appl. Genet., 93:626–632.
62. Lashermes,P., Paczek, V., Trouslot, P., Combes, M.C., Couturon, E and Charrier A. 2000. Single-locus inheritance in the allotetraploid *Coffea arabica* L. and interspecific hybrid *C. arabica* × *C. canephora*. J. Hered., 91:81–85.
63. Lashermes, P., Cros, J., Marmey, P and Charrier A. 1993. Use of Random amplified DNA markers to analyze genetic variability and relationships of Goflea species. Genet. Resour. Crop Evol., 40: 91–99.
64. Lashermes, P., Combes, M.C., Robert, J., Trouslot, P.,Hont, A. D., Anthony, F and Charrier A. 1999. Molecular characterization and origin of the *Coffea*

*arabica* L. genome. Mol. General Genet., 261:259-266.

65. Lepelley, M., Mahesh, V., McCarthy, J., Rigoreau, M., Crouzillat, D., Chabrillange, N., de Kochko, A and Campa C. 2012. Characterization, high-resolution mapping and differential expression of three homologous PAL genes in *Coffea canephora* Pierre (Rubiaceae). Planta, 236: 313-26.
66. Leroy, T., Marraccini, P., Dufour, M., Montagnon, C., Lashermes, P., Sabau, X., Ferreira, L. P., Jourdan, I., Pot, D., Andrade, A. C., Glaszmann, J. C., Vieira, L. G and Piffanelli P. 2005. Construction and characterization of a *Coffea canephora* BAC library to study the organization of sucrose biosynthesis genes. Theor Appl. Genet.,111: 1032-41.
67. Leroy, T., Ribeyre, F., Bertrand, B., Charmetant, P., Dufour, M., Montagnon, C., Marraccini, P and Pot D. 2006. Genetics of coffee quality. Brazilian J. Plant Physio., 18, 229-242.
68. Merga, D., Mohammed, H., and Ayano A. 2020. Studies on the genetic variability among wollega coffee (*Coffea arabica* L.) landrace in western Ethiopia. J. Genet. Genom. Plant Breed., 4 (3): 112-124.
69. Mahé, L., Combes, M.C., Várzea, V.M.P., Guilhaumon, C and Lashermes P. 2008. Development of sequence characterized DNA markers linked to leaf rust (*Hemileia vastatrix*) resistance in coffee (*Coffea arabica* L.). Mol. Breed., 21:105–113.
70. Mahesh, V., Rakotomalala, J. J., Le Gal, L., Vigne, H., De Kochko, A., Hamon, S., Noirot, M and Campa, C.2006. Isolation and genetic mapping of a *Coffea canephora* phenylalanine ammonialyase gene (CcPAL1) and its involvement in the accumulation of caffeoyl quinic acids. Plant Cell Reports, 25:986-992.
71. Maia,T.A., Maciel-Zambolim, E., Caixeta, E.T., Mizubuti, E.S.G and Zambolim L. 2013. The population structure of *Hemileia vastatrix* in Brazil inferred from AFLP. Australas Plant Pathol., 42:533–542.
72. Maluf, M. P., Silvestrini, M., Ruggiero, L. M., de Guerreiro-filho, C. O and Colombo, C. A. 2005. Genetic diversity of cultivated *Coffea arabica* inbred lines assessed by RAPD, AFLP and SSR Marker systems. Scientia Agricola, 62: 366–373.
73. Masumbuko, L.I and Bryngelsson T. 2006. Inter simple sequence repeat (ISSR) analysis of diploid coffee species and cultivated *Coffea arabica* L. from Tanzania. Genet. Res. Crop Evol., 53:357–166.
74. Masumbuko, L. I., Bryngelsson, T., Mneney, E. E and Salomon B. 2003. Genetic diversity in Tanzanian Arabica coffee using Random Amplified Polymorphic DNA (RAPD) markers. Hereditas, 139: 56–63.
75. Maurin, O., Davis, A.P., Chester, M., Mvungi, E.F., Jaufeerally-Fakim, Y and Fay M.F. 2007. Towards a phylogeny for coffea (Rubiaceae): Identifying well-supported lineages based on nuclear and 23 Plastid DNA sequences. Annals Bot., 100: 1565–1583.
76. Mekbib, Y., Saina, J.K., Tesfaye, K., Eshetu, G and Hu G. Chloroplast genome sequence variations and development of polymorphic markers in *Coffea arabica*. Plant Mol. Biol. Rep., 38:491–502.
77. Mekonnen, T., Haileselassie, T and Tesfaye K. 2017. Identification, Mapping and Pyramiding of Genes/Quantitative Trait Loci (QTLs) for Durable Resistance of Crops to Biotic Stresses. J. Plant Patho. Microbi., 8 (6): 412.
78. Mishra, M.K., Nishani, S., Gowda, M., Padmajyothi, D., Suresh, N., Sreenath, H. and Raghuramulu, Y., 2014. Genetic diversity among Ethiopian coffee (*Coffea arabica* L.) collections available in Indian gene bank using sequence related amplified polymorphism markers. Plant Breed. Seed Sci., 70(1):29-40.
79. Mizuno, K., Matsuzaki, M., Kanazawa, S., Tokiwano, T., Yoshizawa, Y and Kato M. 2014. Conversion of nicotinic acid to trigonelline is catalyzed by N-methyltransferase belonged to motif B'

methyltransferase family in *Coffea arabica*. Biochem. Biophysical Res. Communications, 452:1060–1066.

80. Mizuno, K., Okuda, A., Kato, M., Yoneyama, N., Tanaka, H., Ashihara, H and Fujimura T. 2003. Isolation of a new dual-functional caffeine synthase gene encoding an enzyme for the conversion of 7-methylxanthine to caffeine from coffee (*Coffea arabica* L.). FEBS Letters 534, 75- 81.
81. Moncada, M.P., Tovar, E., Montoya, J.C., Gonzalez, A., Spindel, J and McCouch S. 2016. A genetic linkage map of coffee (*Coffea arabica* L.) and QTL for yield, plant height, and bean size. Tree Genet. Genom., 12:1- 7
82. Moncada, P., and S. McCouch. 2004. Simple sequence repeat diversity in diploid and tetraploid coffea species. Genome 47: 501–509. doi:10.1139/g03-129.
83. Moncada, P., Tovar, E., Montoya, J. C., Gonzalez, A., Spindel, J., and McCouch, S. (2014). A genetic of linkage map of coffee (*Coffea arabica* L) and QTL for yield, plant height and bean size. In "The 25th International conference on coffee and science ASIC 2014", Colombia.
84. Montagnon, C and Bouharmont P. 1996. Multivariate analysis of phenotypic diversity of *Coffea arabica*. Genetic Resour. Crop Evol., 43: 221–227.
85. Morris, J. 2018. Coffee: A Global History. 1st ed. London: Reaktions Books, p176.
86. Nadeem, M.A., Nawaz, M.A., Shahid, M.Q., Doğan, Y., Comertpay, G., Yıldız, M., Hatipoğlu, R., Ahmad, F., Alsaleh, A., Labhane, N and Özkan, H.2018. DNA molecular markers in plant breeding: current status and recent advancements in genomic selection and genome editing. Biotech. Biotechnol. Equipment, 32(2):261-285.
87. Nagai, C., Jones, M. R., Byers, A. E., Adamski, D. J and Ming R. 2007. Development and characterization of a true F2 population for genetic and QTL mapping in Arabica. In 21st International Conference on Coffee Science , Montpellier, France, p 771-777.
88. Noirot, M., Charrier, A., Stoffelen, P and Anthony F. 2016. Reproductive isolation, gene flow and speciation in the former Coffea subgenus: a review. Trees, 30:597–608.
89. Ogawa, M., Herai, Y., Koizumi, N., Kusano, T and Sano H. 2001. 7-Methylxanthine Methyltransferase of coffee plants: gene isolation and enzymatic properties. J. Biolo. Chem., 276: 8213-8218.
90. Ogita, S., Uefuji, H., Morimoto, M and Sano H. 2004. Application of RNAi to confirm theobromine as the major intermediate for caffeine biosynthesis in coffee plants with potential for construction of decaffeinated varieties. Plant Mol. Bio., 54: 931-941.
91. Omondi, C.O., Ayiecho, P. O., Mwang'ombe, A.W and Hindorf H. 2001. Resistance of *Coffea arabica* cv. Ruiru 11 tested with different isolates of *Colletotrichum kahawae*, the causal agent of coffee berry disease. Euphytica, 121: 19–24.
92. Orozco-Castillo, C., Chalmers, K.J., Powell, W and Waugh R.1996. RAPD and organellar specific PCR re-affirms taxonomic relationship within the genus Coffea. Plant Cell Rep., 15:337–341.
93. Ortega, F and Lopez-Vizcon C. 2012. Application of molecular marker-assisted selection (MAS) for disease resistance in a practical potato breeding program. Potato Res., 55:1–13.
94. Pearl, H.M. 2004. Construction of a genetic map for Arabica coffee. Theor. Appl. Genet., 108: 829-835.
95. Pearl, H., Nagai, C., Moore, P., Steiger, D., Osgood, R and Ming R. 2004. Construction of a genetic map for arabica coffee. Theor. Appl. Genet., 108: 829-835.
96. Pestana, K.N., Capucho, A.S., Caixeta, E.T., de Almeida, D.P., Zambolim, E.M., Cruz, C.D., Zambolim, L., Pereira, A.A., de Oliveira, A.C.B and Sakiyama N.S. 2015. Inheritance study and linkage

mapping of resistance loci to *Hemileia vastatrix* in Híbrido de Timor UFV 443-03. Tree Genet. Genom., 11:1-13.

97. Pot, D., Bouchet, S., Marraccini, P., De Bellis, F., Cubry, P., Jourdan, I., Pereira, L.F.P., Vieira, L.G.E., Musoli, C.P and Legnate H. 2019. Nucleotide diversity of genes involved in sucrose metabolism. Towards the identification of candidates genes controlling sucrose variability in Coffea sp. Conilon Coffee; Capixaba Institute for Research, Technical Assistance and Rural Extension: Linhares, Brazil, p 679-686.
98. Prakash, N.S., Marques, D.V., Varzea, V.M., Silva, M.C., Combes, M.C and Lashermes P. 2004. Introgression molecular analysis of a leaf rust resistance gene from *Coffea liberica* into *C. Arabica* L. Theor Appl Genet., 109:1311–1317.
99. Prakash, N.S., Muniswamy, B., Hanumantha, B.T., Sreenath, H.L., Sundaresha, K.D., Suresh, N., Santhosh, P., Soumya, P.R., Asha, B.M and Bhat S.S. 2011. Marker-assisted selection and breeding for leaf rust resistance in coffee (*Coffea arabica* L.) some recent leads. Indian J. Genet. Plant Breed., 71:185–189.
100. Prakash, N.S., Combes, M.C., Dussert, S., Naveen, S. and Lashermes P. 2005. Analysis of genetic diversity in Indian robusta coffee gene pool (*Coffea canephora*) in comparison with a representative core collection using SSRs and AFLPs. Genet. Resour. Crop Evol., 52: 333-343.
101. Hendre, P.S and Ramesh K. A. 2007.DNA markers: development and application for genetic improvement of coffee." In Genomics-Assisted Crop Improvement: Vol. 2: Genomics Applications in Crops,. Dordrecht: Springer Netherlands, p 399-434200.
102. Priolli, R. H. M., Paulo, Siqueira, W. J., Möller, M., Zucchi, M. I., Ramos, L. C. S., Gallo, P. B and Colombo C. A. 2008. Caffeine inheritance in interspecific hybrids of *Coffea arabica* x *Coffea canephora* (Gentianales, Rubiaceae). Genet. Mol. Biol., 31: 498-504.
103. Privat, I., Bardil, A., Gomez, A. B and Severac D. 2011. The 'PUCE CAFE' Project: the first 15K coffee microarray, a new tool for discovering candidate genes correlated to agronomic and quality traits. BMC Genomi., 12: 1-14
104. Privat, I., Foucrier, S., Prins, A., Epalle, T., Eychenne, M., Kandalaft, L., Caillet, V., Lin, C., Tanksley, S and Foyer C. 2008. Differential regulation of grain sucrose accumulation and metabolism in *Coffea arabica* (Arabica) and *Coffea canephora* (Robusta) was revealed through gene expression and enzyme activity analysis. New Phytologist, 178: 781-797.
105. Pruvot-Woehl, S., Krishnan, S., Solano, W., Schilling, T., Toniutti, L and Bertrand B. 2020. Authentication of *Coffea arabica* varieties through DNA fingerprinting and its significance for the Coffee Sector. J. AOAC Int.,103:325-334.
106. Raina, S.N., Mukai, Y and Yamamoto M. 1998. In situ hybridization identifies the diploid progenitor species of *Coffea arabica* (Rubiaceae). Theor. App. Genet., 97: 1204–1209.
107. Ram, A.S. 2006. Popular Indian coffee selections. Indian Coffee, 70:12–18.
108. Ray,S and Satya P .2014. Next-generation sequencing technologies for next-generation plant breeding. Front. Plant Sci., 5:367.
109. Resende, M., Caixeta, E and Alkimim E.R. 2016. High-throughput targeted genotyping of *Coffea arabica* and *Coffea canephora* using next generation sequencing. San Diego, CA, 1.
110. Romero, G., Vásquez, L.M., Lashermes, P and Herrera J.C. 2014. Identification of a major QTL for adult plant resistance to coffee leaf rust (*Hemileia vastatrix*) in the natural Timor hybrid (*Coffea arabica* x *C. canephora*). Plant Breed., 133: 121–129.
111. Rutherford, M. A. 2005. Epidemiology and variability of *Gibberella xylarioides*, the coffee wilt pathogen. DFID – Crop

Protection Program, Final Technical Report CAB International, UK, p 24.

112. Santana, G.C., Pereira, L.F., Pot, D., Ivamoto, S.T., Domingues, D.S., Ferreira, R.V., Pagiatto, N.F., da Silva, B.S., Nogueira, L.M., Kitzberger, C.S and Scholz M.B. 2018. Genome-wide association study reveals candidate genes influencing lipids and diterpenes contents *in Coffea arabica* L. Scient. Reports, 8(1):465.
113. Silva, M.C., Várzea, V., Guerra-Guimarães, L., Azinheira, H.G., Fernandez, D., Petitot, A., Bertrand, B., Lashermes, P and Nicole, M. 2006. Coffee resistance to the main diseases: leaf rust and coffee berry disease. Braz. J. Plant Physiol.18:119–147.
114. Simkin, A. J., Qian, T., Caillet, V., Michoux, F., Ben Amor, M., Lin, C., Tanksley, S and McCarthy J. 2006. Oleosin gene family of *Coffea canephora*: Quantitative expression analysis of five oleosin genes in developing and germinating coffee grain. J. Plant Physi., 163: 691-708.
115. Sousa, T.V., Caixeta, E.T., Alkimin, E.R and Oliveira A.C.B. 2013. Molecular markers are useful to discriminate Coffea arabica cultivars with high genetic similarity. Euphytica, p1-15.
116. Spinoso-Castillo, J.L., Escamilla-Prado, E., Aguilar-Rincón, V.H., Morales Ramos, V., De Los Santos, G.G., Pérez-Rodríguez, P and Corona-Torres T. 2020. Genetic diversity of coffee (Coffea spp.) in Mexico evaluated by using DArTseq and SNP markers. Genet. Resou. Crop Evol., 67:1795-1806.
117. Teferi, D. 2019.Achievements and Prospects of Coffee Research in Ethiopia: A Review. Int. J. Res. Stud. Agric. Sci., 5:41-51.
118. Teixeira-Cabral, T. A., Sakiyama, N. S., Zambolim, L., Pereira, A. A and Schuster I. 2004. Single-locus inheritance and partial linkage map of *Coffea arabica* L. Crop Breed. Appl. Biotech., 4:416-421.
119. Teressa, A., Crouzillat, D., Petiard, V and Brouhan,P. 2010. Genetic diversity of Arabica coffee (*Coffea arabica* L.) collections. Ethiopian J. Applied Sci. Tech., 1(1):63-79.
120. Tesfaye, K., Borsch, T., Govers, K and Bekele E. 2007. Characterization of Coffea chloroplast microsatellites and evidence for the recent divergence of *C. arabica* and *C. eugenioides* chloroplast genomes. Genome, 50: 1112–1129.
121. Tesfaye, N., Sussumu, S., Evliene, T., Cosme, D and Antonio C. 2009. Genetic diversity and Genome introgression in coffee. . A Ph.D. thesis submitted to the school of graduate studies of Minas Gerais Brasil University, p73.
122. Thomas, A.S. 1942. The wild arabica coffee of the Boma Plateau, Anglo-Egyptian Sudan. Empir. J. Exp. Agric. 10, 207–212.
123. Tornincasa, P., Dreos, R., De Nardi, B., Asquini, E., Devasia, J., Mishra, M.K., Del Terra, L., Crisafulli, P., Pallavicini, A and Graziosi G. 2006. Genetic diversity of commercial coffee (C. arabica L) from America, India and Africa assessed by simple sequence repeats (SSRs). Proceedings of the 21st International Association for Coffee Science (ASIC'06), p778-785.
124. Uefuji, H., Ogita, S., Yamaguchi, Y., Koizumi, N and Sano H. 2003. Molecular cloning and functional characterization of three distinct N-methyltransferases involved in the caffeine biosynthetic pathway in coffee plants. Plant physic., 132: 372-380..
125. USDA.2020. Coffee: World Markets and Trade. Coffee: World Markets and Trade. Available at: http://apps.fas.usda. gov/psdonline/circulars/coffee.pdf
126. Van der Vossen, H.A.M and Walyaro D. J. 1980. Breeding for resistance to coffee berry disease in Coffea arabica L. II. Inheritance of the resistance. Euphytica, 29: 777-791.
127. Vega, F.E., Ebert, A.W and Ming R. 2008. Coffee germplasm resources, genomics, and breeding. Plant Breed. Rev., 30:415-447.

128. Vidal, R.O., Mondego, J.M.C and Pot D. 2010. A high-throughput data mining of single nucleotide polymorphisms in *Coffea species* expressed sequence tags suggests differential homologous gene expression in the allotetraploid *Coffea arabica*. Plant Physio., 154:1053-1066.
129. Merot-Anthoene, V., Tournebize, R., Darracq, O., Rattina, V., Lepelley, M., Bellanger, L., Tranchant-Dubreuil, C., Coulée, M., Pégard, M., Metairon, S and Fournier C. 2019. Development and evaluation of a genome-wide Coffee 8.5 K SNP array and its application for high-density genetic mapping and for investigating the origin of *Coffea arabica* L. Plant Biotech. J., 17(7):1418-1430.
130. Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Plot, J., Peleman, J., Kuiper, M and Zabeau M. 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res., 23: 4407–4414.
131. Vuylsteke, M., Mank, R., Antonise, R., Bastiaans, E., Senior, M.L., Stuber, C.W., Melchinger, A.E., Lübberstedt, T., Xia, X.C., Stam, P and Zabeau M. 1999. Two high-density AFLP® linkage maps of Zea mays L.: analysis of distribution of AFLP markers. Theor Appl Genet., 99: 921-935.
132. Walyaro, D.J.A., 1983. Consideration in Breeding for Improved Yield on Quality Arabica Coffee (Coffea arabica L.) PhD thesis. Wageningen, The Netherlands: Agricultural University.
133. Yeshitila, M., Kassahun, T., Xiang, D., Josphat, K. S., Guang Wan, H and Qing Feng W. 2022. Whole-genome resequencing of *Coffea arabica* L. (Rubiaceae) genotypes identifies SNP and unravels distinct groups showing a strong geographical pattern. BMC Plant Bio., 22(1): 69.
134. You, Q., Yang, X., Peng, Z., Xu, L and Wang J. 2018. Development and applications of a high throughput genotyping tool for polyploid crops: single nucleotide polymorphism (SNP) array. Front. Plant Sci., 9:104.
135. Yu, Q., Guyot, R., de Kochko, A., Byers, A., Navajas-pérez, R and Langston, B.J., Dubreuil-Tranchant, C., Paterson, A.H., Poncet, V and Nagai C et al. 2011. Micro-collinearity and genome evolution in the vicinity of an ethylene receptor gene of cultivated diploid and allotetraploid coffee species (Coffea). The Plant J., 67: 305–317.
136. Yuyama, P.M., Reis Júnior, O., Ivamoto, S.T., Domingues, D.S., Carazzolle, M.F., Pereira, G.A.G and Leroy T. 2016. Transcriptome analysis in Coffea eugenioides, an Arabica coffee ancestor, reveals differentially expressed genes in leaves and fruits. Mol. Genet. Genom., 291: 323–336.
137. Zhou, L., Vega, F.E., Tan, H., Lluch, A.E.R., Meinhardt, L.W. and Fang W. *et al.*, 2016. Developing single nucleotide polymorphism (SNP) markers for the identification of coffee germplasm. Trop PlantBiol.,9:82-95.