# Molecular sexual determinants in Pistacia genus by KASP assay

**Zeyneb Nur Sahin**
Marmara University: Marmara Universitesi
**Ezgi Cabuk Sahin**
Marmara University: Marmara Universitesi
**Yildiz Aydin**
Marmara University: Marmara Universitesi
**Ahu Altinkut Uncuoglu** ( ahu.uncuoglu@marmara.edu.tr )
Marmara University: Marmara Universitesi

**Research Article**

**Keywords:** KASP, Pistacia, Sex discrimination, SNP

**Posted Date:** May 17th, 2021

**DOI:** https://doi.org/10.21203/rs.3.rs-499676/v1

## Abstract

# I. Background: Pistacia

is kind of unisexual plant species that can take 4–5 years to initiate bearing stage of trees in terms of economic value. The fruits of species have an extreme importance in food, health and baking industry as a raw material. So, identify the individual in early juvenile period for pollination and position of trees is very crucial for growers. The objective of this study is developing markers for each *Pistacia* species which helps the screening the sex of plant seedling quickly before they reach reproduction stage without waiting morphological data.

# II. Methods and Results:

Within the context, by using the power of KASP (Kompetitive Allele Specific PCR) assay technology as a marker screening system, we successfully discriminate 7 *Pistacia* species out of 8 which are *P. atlantica*, *P. integerrima*, *P. khinjuk*, *P. mutica*, *P. terebinthus*, *P. vera* and *P. lentiscus*. We have used a high-throughput DNA sequence read archive (SRA) to assemble a reference genome to use in our studies as de novo bioinformatics method. 4 genomic regions were sequenced with collected plant material from Antepfıstığı Research Institute Collection Garden predominantly and 8 species were aligned intraspecifically for Single Nucleotide Polymorphism (SNP) mining. 12 SNP markers are converted to KASP markers and 5 of them (SNP-PIS-133396, SNP-PIS-167992, P-ATL-91951-565, P-INT-91951-256, P-KHI-91951-115) showed obvious allelic discrimination between male and female species. SNP-PIS-167992 and P-ATL-91951-565 were identified as best marker assays due to the allelic frequency differences in graphs which are apparent for all individuals and homozygosis/heterozygosis observed clearly. These markers could be most comprehensive for whole genus because the discriminability power of several species in once.

# III. Conclusions:

As a conclusion, this study is reported first *Pistacia* gender discrimination by KASP, besides the precursor study for sex discrimination by KASP in plants.

## Introduction

*Pistacia* is the branch of the Anacardia (Cashew) [1] family and the genus has increasing economic value [2] on food, drug [3], healthcare [4, 5] and baking industry. The primarily producers of *Pistacia vera* are respectively USA, Iran, Turkey and China however the Turkey's yield value is two times more than worldwide yield [6]. Species have different character of leaves like membranous, leather, deciduous, evergreen, stipule and trifoliate. Besides, the most commonly known characteristics of the species was being dioecious, wind pollination (Anemophilous), homeochlamydic perianth [7] and have unisexual flowers.

One of the essential and common character for *Pistacia* genus is sex differences for individual. This Sexual Size Dimorphism (SSD) of species effects from fertilization season to individual body or fruit size [8]. These are the anatomical and structural traits of male and female individuals. Male (♂ staminate) and female (♀ pistillate) flowers occur on different individuals on unisexual plants [9], therefore fruit trees require the proximity of the opposite sex trees to produce fruit and identify the individual is very crucial for producers. Marker Assisted Selection (MAS) is a promising process and may help to screen the sex or many other traits of plant seedlings quickly before they reach reproduction stage or show morphological data in more than several years [10].

Molecular markers are nucleic acid (DNA/RNA) fragments that are any gene region or associated with genes in the genome [11]. These markers can be used for a variety of purposes ranging from creating of genetic and character maps of families to identifying and defining species [12]. Each molecular marker has advantages and disadvantages [13] so this feature gives the opportunity to choose the best fit research routes to researchers. There are also comprehensive studies which used different molecular markers together on the same study to make sense of genetic diversity between species [14].

Single Nucleotide Polymorphism (SNP) is a single point mutation in the genome of an individual. There are two main advantages for SNP usage; their abundance on genomes (e.g. 1/36 bp [1 SNP per 36 base pair] for Arabidopsis, 1/21 bp for potato and 1/78 bp for barley) and availability of high throughput SNP genotyping technologies [15]. SNP markers are often used on many plants species and many platforms like QTL (Quantitative Trait Locus) analysis on rice [16], Nested Association-Mapping (NAM) on maize for complex trait flowering time [17] or founding gene-specific diagnostic marker against fungal pathogens like leaf rust, stripe rust, and powdery mildew diseases for maize [18]. In addition to this, high-throughput SNP genotyping could be performed on Fixes-arrays or Genotyping by Sequencing (GBS) technologies. While fixed-arrays has been used on large number of samples, however, the higher cost it's an always limiting factor. As for that flexible technologies like TaqMan®, Douglas Array Type and KASP™ are preferable method of choice on certain fields like agriculture or personal medicine [19].

Kompetitive Allele-Specific PCR (KASP) is the PCR-based fluorescently-labelled marker scanning system which developed by KBioscience. The low-cost KASP assay platform is defined as breeder-friendly [20], time effective, novel and more reliable for routine screening than the other similar technologies like Taqman assay and Sanger sequencing [21]. The KASP is a flexible genotyping assay which able to use from high-throughput large scale to small research laboratories [22]. In addition, according to Semagn's review in 2014, KASP is 12.3–46.1% less expensive than GoldenGate or BeadXpress platforms. Via this uniplex SNP genotyping [23], one marker can be run and scan on real-time by fluorescent plate readers. The technology based on KASP assay is Fluorescence Resonance Energy Transfer (FRET) cassettes combination with PCR system. FRET cassettes originate from fluorescent tagged dyes FAM (6-carboxyfluorescein) and HEX (hexa-chloro-fluorescein) with specific oligo tails and their complementary quencher molecules to stop signal generation before matching. Preparation of two unlabeled allele-specific primer is the other key point of the reaction. Dye's sequences are the same with extension part of our primers one to one and quencher

have complementary of these extensions. At the end of the PCR cycles, used primer/primers are defined through signal frequency [24]. The KASP has a wide range of usage on different livings and it can be used easily for genotyping with cost advantage [22]. Cacao [25], apple [26, 27], soybean [28, 29], grain amaranth [30] ash tree [31], robus [32], dentrobium [33], watermelon [34], peanut [35; 36], maize [23], rice [37] and wheat [20, 38, 39, 40, 41] are some of the plants species which KASP assay has successfully applied for different purposes like Whole Genome Sequencing (WGS), GBS, breeding or identify to genetic diversity up to the present.

## Materials And Methods

### Plant material and genomic DNA extraction

Eight different male and female *Pistacia* species (*Pistacia atlantica, Pistacia integerrima, Pistacia khinjuk, Pistacia mutica, Pistacia palaestina, Pistacia terebinthus, Pistacia vera* and *Pistacia lentiscus*) were used as plant material that were collected from Antepfıstığı Research Institute Collection Garden (Gaziantep, Turkey) and from Çeşme Çiftlikköy region (*Pistacia lentiscus* L.) in İzmir, Turkey. CTAB method was used for DNA isolation [42]. Each the extracted genomic DNA sample was quantified with the Qubit 2.0 fluorometer according to the manufacturer's instructions and evaluated by loading of DNA based on Qubit measurements on a 0.8% agarose gel electrophoresis. Then genomic DNA samples were diluted to 2 ng $\mu L^{-1}$ and stock &diluted genomic DNA samples were stored separately at –20 °C. The plan of study was designed as shown in the chart (Fig. 1) and practiced step by step.

### SNP mining

In order to establish a broad range of molecular markers for sex discrimination in these *Pistacia* species, we focus the genomics regions shown to be definitive for this purpose [43]. Three SNP positions were selected (SNP-PIS-133396, SNP-PISTACIA-176863- SNP-PISTACIA-167992) (Table 1) according to their High Resolution Melting (HRM) positive results and were defined as suitable for KASP primer transformation. We have also utilized two set of WGS reads that were available in National Center for Biotechnology Information (NCBI's) high-throughput DNA and RNA SRA database with accession numbers *SRX2270199* and *SRX2270198*. These reads were assembled as a Denovo Assembly (Appendix 1) and used for the further examining the genomic areas of interest in order to develop further molecular markers for sex discrimination of other *Pistacia* species.

The data was operated for quality control by FastQC tool, used adapter contents on sequence and bad quality reads (Phred score <30) were removed. Short Oligonucleotide Analysis Package (SOAP) denovo Assembler (v2.01) were utilized to assemble raw WGS data without model organisms. The assembly tool is included these modules; Construction of Bruijn graph from k-mers which best lengths are estimated by KmerGenie tool, Paired-end (PE) 90 and PE150 reads mapping used for the assembly, contigs and scaffolds assembled and their fasta files created, scaffold fasta files are used for the final output of assemblies. Probable 4 hot regions were defined (Appendix 1) according to the previous studies [43] for Sanger sequencing.

### Sanger sequencing

The sequencing analysis has been performed by LGC Genomics GmbH Laboratory. We amplified the 4 target sites (Appendix 1) of each 16 *Pistacia* samples DNA template individually with specific primers (Table 1) to verify of the DNA sequence variants (substitutions or indels) detected by de novo sequencing at LGC Genomics GmbH Laboratory. Sequencing runs were performed on a 3730xl DNA Analyzer (Applied Biosystems™/ Thermo Fisher Scientific). PCR clean-up was performed with ExoSAP-IT™ PCR Product Cleanup Reagent (Thermo Fisher Scientific) according to the manufacturer's protocol.

### Multiple alignment and new SNP's

By alignment of four regions (REG-PIS-136404, REG-PIS-174431, REG-PIS-133396, REG-PIS-167992), we selected 9 SNP among polymorphisms. For each species female and male individuals were grouped and overlapped via Geneious Prime Program to identify gender-based SNP differences. For this program Pairvise to Geneious / Multiple Alignment and alignment tool was utilized as a kind of global alignment features with free end gaps, 65% was chosen as a cost matrix. Nine selected SNPs were transformed into KASP assay to test for new functional markers (Table 2).

### KASP (Kompetitive Allele Specific PCR)

The KASP assay was performed in 20 µL reaction volume consisting of 10 µL KASP master mix, 0.28 µL KASP primer assay mix and 10 µL of 2 ng $\mu L^{-1}$ DNA template. Thermocycling was performed in a ABI 9700 real-time with the following LGC standard 61–55 °C touchdown cycling conditions: 94 °C 15 min, followed by 10 cycles of 94 °C 20 s, 61 °C (decreasing by 0.6 °C per cycle) 60 s and finally 26 cycles of 94 °C 20 s, 55 °C 60 s in triplicate. Fluorescence was tracked in real-time with plate reads at the end of every amplification cycle and allelic discrimination was performed (Applied Biosystems' StepOneTM Software version 2.2.2).

## Results

### Sanger sequencing and SNP sex association

Sequence analyzes were performed with 75% success and quality sequences between a minimum of 250 bp and a maximum of 1142 bp were selected. After alignment of male and female samples for each species, SNP's were determined between genders (Fig. 2) for use in the KASP analysis. Consensus sequences were supplied with Appendix 2.

### Verification of sex association

In this study 12 SNP markers were converted to KASP markers (Table 1) to use in the assays. Five of them (SNP-PIS-133396, SNP-PIS-167992, P-ATL-91951-565, P-INT-91951-256, P-KHIN-91951-115) were discriminative for gender in 7 *Pistacia* species out of 8 namely *P. atlantica, P. integerrima, P. khinjuk, P. mutica, P. terebinthus, P. vera* and *P. lentiscus* (Fig. 3).

According to intraspecific consideration, SNP-PIS-167992 (A/T) marker showed female/male distinction for 3 species which are *P. vera*, *P.atlantica* and *P. terebinthus*. While the male samples showed homozygous allele 2-character (T/T) for this assay and species, the signal frequencies of female samples indicated the heterozygote allele1/ allele 2 (A/T) feature as shown in Fig. 3a, e. The marker P-INT-91951-256 had only determinative result for *P. integerrima* as shown in Fig. 3b. When examined the whole genus, one of the best assay results was P-ATL-91951-565. The assay separated the male and female in the 4 species which are (*P. atlantica, P. khinjuk, P. mutica, and P. terebinthus*) in *Pistacia* genus intraspecifically (Fig. 3c, d). SNP-PIS-133396 marker had only discriminative result for *Pistacia vera* among the genus (Fig. 1f). The 5$^{th}$ marker which discriminated more than one species was P-KHIN-91951-115. The assay showed distinction for 3 species and they are *P. khinjuk*, *P.atlantica* and *P. lentiscus* (Fig. 3g).

Separations of the male & female samples of *P. atlantica* was provided with 3 different SNP assays which are SNP-PIS-167992, P-ATL-91951-565 and P-KHIN-91951-115. The best result among these reactions was SNP-PIS-167992 (Fig. 3a). According to this assay means value of male samples showed homozygous (T/T) and female samples showed heterozygous (AT) characters (M= A1:0.4584, A2: 0.4886; F= A1:0.2076, A2:0.5052) as shown in Fig. 3a. The same assay reaction was also best for *P. terebinthus* (Fig. 3e). The allelic frequencies of the male & female species are respectively A1:0.0395, A2:0.3476 & A1:0.2412, A2:0.1977 as shown in Fig. 3e (M:Male, F:Female; A1:Allelle 1, A2: Allelle 2).

*P. integerrima* was discriminated only with P-INT-91951-256 assay. The allelic frequencies are respectively A1:0.7416 & A2: 0.2464 for mean value of *P. integerrima* males and A2:0.3387 & A2:0.1154 for mean value of *P. integerrima* females shown in Fig. 3b.

The most obvious results for *P. khinjuk* and *P. mutica* were obtained from P-ATL-91951-565 T/C base-pair substitution assay (Fig. 3c, d). On this assay while both *P. khinjuk* and *P. mutica* males showed homozygous (T/T) allele 1 genotype, their females have heterozygous (T/C) feature. The *P. khinjuk* and *P. mutica*'s mean values of allelic frequencies are respectively A1M: 1.3269, A2M: 0.575; A1F: 0.7439, A2F: 0.4111 & A1M: 1.3157, A2M: 0.1580; A1F: 0.7473, A2F: 0.4208 for male and female samples.

Analysis with SNP-PIS-133396 containing the C/G base-pair substitution revealed that *P. vera* female was heterozygous respect to mean value of allele frequencies (A1: 0.9239, A2: 0.5016) while all the individuals in the genus including *P. vera* male (A1: 0.0933; A2: 0,4846) were exhibited homozygous allele 2 (G/G) character as shown in Fig. 3f.

*P. lentiscus* is only discriminated with P-KHIN-91951-115 assay in terms of sex. While the demonstration points of genders shown close in plot (Fig. 3g) and in allele 2 frequencies are similar values, according to allele 2 the differences are more than 3 times (A1M: 0.1258, A2M: 0.4786; A1F: 0.3937, A2F: 0.5203)

The result of P-INT-91951-120 assay indicated that while male and females of *P. integerrima* was showing heterozygous (T/A) genotype, all other male and female samples of species in genus were homozygous allele 2 (A/A). This marker can be accepted as a species-specific for *Pistacia* genus.

## Discussion

The first sex determination study on *Pistacia* with molecular markers was done by Hormaza [10]. One thousand Random Amplified Polymorphic DNA (RAPD) markers had used for discrimination in that study. Bulk *P. vera* L. male and female samples had amplified with 700 decamer primers and as a result OPO08945 (OPO-08) had defined as a sex marker. 945 bp size band was observed on female bulk sample and this band was absent on male samples. Yakubov et al. reported to verify the primer and develop OPO-08 RAPD primer with touchdown PCR and they determined PVF1-2 Sequence Characterized Amplified Region (SCAR) primers based on OPO-08. In that study 4-Point Base Deletion (PBD) for female samples and 1 PBD for males were defined as a mutation (again with *P. vera*) [44]. In other completed gender discrimination study was associated RAPD markers in wild *Pistacia* which conducted by Kafkas et al. [45]. BC156 (1300bp) & BC360 (500bp) markers for *P. eurocarpa* and OPAK-09 (850bp) & BC346 (700bp) markers for *P. atlantica* was indicated as a marker. In the articles published by Esfendiyari et al. [46, 47], the first twenty 10-mer RAPD primers than thirty different 10-mer RAPD primer were tried on 3 different species; *P. atlantica*, Desf subsp. *mutica*, *P. khinjuk* and *P. vera*. As a result, they reported that BC1200 primer is a discriminator marker and designed the primers PVF1 (forward) and PVF2 (reverse) from the PCR amplified region sequence (convert RAPD to SCAR). The SCAR primers which has 300 bp amplifying region, while generate a band in all female samples, there wasn't any band in male samples.

When studies are examined, it is seen that many of them are based on *P. vera* genus [48, 49, 43, 50, 44] due to the world wide trade income. On this basis, in 2008, *P. vera* L. cultivars were generated with Inter Simple Sequence Repeat (ISSR) primers and 2 of them were able to distinguish genders. These markers are (AC) 8GC and (AC) 8TA 10-mer repeats and band sizes are 2400 [48]. Following year the same research group was experienced 32 arbitrary primers for RAPD analysis and mentioned FPK 106 and FPK 105 but there was no data related to primer sequence. Vedramin et al. carried out the SSR study covering 4 species [51]. Their analysis contains 50 *P. vera*, 4 *P. integerrima*, 12 *P. mutica* and 16 *P. terebinthus* for a total of 82 accessions but the tested markers were insufficient for gender base discrimination.

Up to the present, the most reliable study about sex linked marker on Pistacia genus was published by Kafkas et al. [43]. In that study 38 putative sex-associated SNP markers were identified as heterozygous in female individuals and homozygous in male individuals using RAD sEq. suggests a ZW/ZZ sex determination system in *P. vera*.

As a result, some of the 8 species in the genus (*P. vera*, *P atlantica*, *P integerrima*, *P. terebinthus*, *P. mutica*, *P. khinjuk*) were studied separately before, while about the remaining part of the genus (*P. palaestina*, *P. lentiscus*) there is no gender discrimination study. However, it has not been previously collected in a single study, which indicates the comprehensiveness of our study.

There are sex discrimination marker studies already done for the *Pistacia* species and different molecular marker methods have been used in these studies like RAPD [47] or ISSR [48]. When the previous studies were compared with the results of KASP, there are extra steps in previous techniques like preparing gel, loading samples or screening the gel. These extra steps can cause disadvantageous in terms of time, effort and cost. While each extra step increases the risk of making mistakes, KASP is very suitable for scale-up as it does not contain any of these steps.

It's obviously known that, this new and popular technique KASP assay has a wide range of usage on different species from crop improvements; maize [23], wheat [20], tomato [52] to detection of disease in human [21]. Nonetheless, sex discrimination studies using KASP technique is not seen except for a few fish like salmon [53] or halibut [54] and one insect species [55]. These results show that, this study is a not just first for *Pistacia* in terms of gender discrimination but also the precursor study for sex discrimination by KASP in plants.

In our study, 12 SNP regions were screened for 8 species and 5 SNP (SNP-PIS-133396, SNP-PIS-167992, P-ATL-91951-565, P-INT-91951-256, P-KHI-91951-115) showed clear allelic discrimination between sexes in species. As a result of this allelic distinction, 5 KASP primers identified as a sex-linked markers for 7 of 8 species. In the sex discrimination marker study conducted by Kafkas et al. [43] for *P. vera* species, 8 of 13 SNP had differentiated between *P. vera* genders. When these markers were tested with some other species (*P. atlantica, P. terebinthus, P. eurycarpa, P. integerrima*) in the genus, a distinction could not be observed in these wild types.

In terms of allelic frequency, the best marker assays were SNP-PIS-167992 and P-ATL-91951-565 (Fig. 4a, c). In allelic graphs of these markers, frequency differences are apparent for all individuals whose distinction was observed, and the distinction is exact in terms of homozygosis and heterozygosis (Fig. 4). At the same time, they can be considered more comprehensive because it differentiates several species at once for whole genus (Fig. 4a, b, c). Each SNP markers show discrimination for 3 *Pistacia* species; SNP-PIS-167992 for *P. atlantica, P. terebinthus, P. vera* (Fig. 4a); P-KHIN-91951-115 for *P. atlantica, P. khinjuck, P. lentiscus* (Fig. 4b); P-ATL-91951-565 for *P. atlantica, P. khinjuck, P. mutica* (Fig. 4c).

When all results are evaluated, the less discriminative result of allelic signals between genders showed itself in separation of *P. lentiscus* species with P-KHIN-91951-115 assay. While allele 2 (HEX) radiations were similar in this result and indicative dots seems close to each other, allele 1 (FAM) had more than 3 times stronger signaling difference in the female individuals. Some of the assays were irradiated after reaction but they did not display a difference on allelic discrimination plot. In one marker (P-VERA-10733-132), no signal was received, which is actually in fact an indication of the KASP system accuracy and the specificity of primer sequences.

The assay P-INT-91951-120 were discriminated only *P. integerrima* species apart from genus and actually it was not surprising, because the assay was designed special and in the direction of *P. integerrima* sequence. When we assess the SNP-PIS-167992 assay result, while the Kafkas et al. [43] research discriminate only *P. vera*, our assay discriminated 2 more species which are *P. atlantica* and *P. terebinthus* in addition.

## Conclusion

This analysis provides insight into the sex discrimination of *Pistacia* genus using sequencing-based assays. SNP assay described in this paper can be used to determine and confirm the sex of *Pistacia* varieties so this study underlines the usefulness of using SNP marker technology in future breeding programs by cost and time effectively. The KASP technique could be a reliable genotyping assay for routine to identify the sex of *Pistacia* genus. Understanding sex/gender in plant reproduction is of immense practical importance for biotechnology, the conservation of biodiversity and the control of invasive species.

## Declarations

### Acknowledgments

We acknowledge to Antepfıstığı Research Institute Collection Garden (Gaziantep, Turkey) for providing plant material. This research has been supported by Marmara University Research Foundation (Project No: FEN-C-YLP-090217-0049).

## References

1. UniProt Consortium (2017) Taxonomy - *Pistacia vera* (Pistachio) (Species). http://www.uniprot.org/taxonomy/55513. Accessed 20 November 2017.
2. Al-Saghir MG, Porter DM (2012) Taxonomic revision of the genus *Pistacia* L. (Anacardiaceae). Am J Plant Sci 3(1):12-32. https://doi.org/10.4236/ajps.2012.31002
3. Mosbah R, Mokrani N, Mosbahi I, Rouabhi S, Mantovani A (2016) Effect of methomyl on the biochemical and reproductive parameters in pregnancy rats: the protective role of *Pistacia lentiscus* oil. 18th European Congress of Endocrinology BioScientifica 41. https://doi.org/10.1530/endoabs.41.EP671
4. Mohagheghzadeh A, Faridi P, Ghasemi Y (2010) Analysis of Mount Atlas mastic smoke: a potential food preservative. Fitoterapia 81(6):577-580. https://doi.org/10.1016/j.fitote.2010.01.022
5. Orhan I, Küpeli E, Aslan M, Kartal M, Yesilada E (2006) Bioassay-guided evaluation of anti-inflammatory and antinociceptive activities of pistachio, *Pistacia vera* L. J Ethnopharmacol 105(1):235-240. https://doi.org/10.1016/j.jep.2005.10.023
6. FAO, Food and Agriculture Organizations of the United Nations (2017) Production of pistachios in worldwide top 10 producers in 2016, Crops http://www.fao.org/faostat/en/#data/QC/visualize. Accessed 15 December 2017.
7. Mabberley DJ (1997) The plant book. Cambridge University Press, Cambridge, UK.
8. Fairbairn DJ (1997) Allometry for sexual size dimorphism: pattern and process in the coevolution of body size in males and females. Annu Rev Ecol Syst 28(1):659-687.

9. Westergaard M (1958) The mechanism of sex determination in dioecious flowering plants. Adv Genet 9:217-281. https://doi.org/10.1016/S0065-2660(08)60163-7
10. Hormaza JI, Dollo L, Polito VS (1994) Identification of a RAPD marker linked to sex determination in *Pistacia vera* using bulked segregant analysis. Theor Appl Genet 89(1):9-13.
11. Devran Z (2003) Moleküler işaretleyicilerin dayanıklılık ıslahında kullanılması. Derim 20(1):1-6.
12. Rafalski JA (2002) Novel genetic mapping tools in plants: SNPs and LD-based approaches. Plant Sci 162(3):329-333. https://doi.org/10.1016/S0168-9452(01)00587-8
13. Ahn S, Tanksley SD (1993) Comparative linkage maps of the rice and maize genomes. PNAS 90(17):7980-7984. https://doi.org/10.1073/pnas.90.17.7980
14. Abuduli A, Aydin Y, Sakiroglu M, Onay A, Ercisli S, Uncuoglu AA (2016) Molecular evaluation of genetic diversity in wild-type mastic tree (*Pistacia lentiscus* L.). Biochem Genet 54(5):619-635. https://doi.org/10.1007/s10528-016-9742-0
15. Henry RJ (2008) SNP discovery in plants. Plant genotyping II: SNP technology. CAB International, USA.
16. Yu H, Xie W, Wang J, Xing Y, Xu C, Li X, Xiao J, Zhang Q (2011) Gains in QTL detection using an ultra-high-density SNP map based on population sequencing relative to traditional RFLP/SSR markers. PloS One 6(3):e17595. https://doi.org/10.1371/annotation/f2eb75fb-ae22-4a57-b828-1506aa506c6d
17. Buckler ES, Holland JB, Bradbury PJ, Acharya CB, Brown PJ, Browne C, Ersoz E, Flint-Garcia S et al (2009) The genetic architecture of maize flowering time. Science 325(5941):714-718. https://doi.org/10.1126/science.1174276
18. Lagudah ES, Krattinger SG, Herrera-Foessel S, Singh RP, Huerta-Espino J et al (2009) Gene-specific markers for the wheat gene Lr34/Yr18/Pm38 which confers resistance to multiple fungal pathogens. Theor Appl Genet 119(5):889-898. https://doi.org/10.1007/s00122-009-1097-z
19. Thomson MJ (2014) High-throughput SNP genotyping to accelerate crop improvement. Plant Breed Biotechnol 2(3):195-212. https://doi.org/10.9787/PBB.2014.2.3.195
20. Tan CT, Assanga S, Zhang G, Rudd JC, Haley SD, Xue Q, Ibrahim A, Bai G, Zhang X et al (2017) Development and validation of KASP markers for wheat streak mosaic virus resistance gene Wsm2. Crop Sci 57(1):340-349. https://doi.org/10.2135/cropsci2016.04.0234
21. Landoulsi Z, Benromdhan S, Djebara MB, Damak M, Dallali H, Kefi R, Abdelhak S et al (2017) Using KASP technique to screen LRRK2 G2019S mutation in a large Tunisian cohort. BMC Genet 18(1):70. https://doi.org/10.1186/s12881-017-0432-5
22. He C, Holme J, Anthony J (2014) SNP genotyping: the KASP assay. In: Crop breeding Humana Press, New York, NY, pp 75-86. https://doi.org/0.1007/978-1-4939-0446-4_7
23. Semagn K, Babu R, Hearne S, Olsen M (2014) Single nucleotide polymorphism genotyping using Kompetitive Allele Specific PCR (KASP): overview of the technology and its application in crop improvement. Mol Breed 33(1):1-14. https://doi.org/10.1007/s11032-013-9917-x
24. LGC Limited, Laboratory of the Government Chemist (2017) Assay search tool. https://www.lgcgroup.com/assays/#.WeCXlY-0PIUw. Accessed 02 November 2018.
25. Gopaulchan D, Motilal LA, Bekele FL, Clause S, Ariko JO, Ejang HP, Umaharan P (2019) Morphological and genetic diversity of cacao (*Theobroma cacao* L.) in Uganda. Physiol Mol Biol Plants 25(2):361-375. https://doi.org/10.1007/s12298-018-0632-2
26. Baumgartner IO, Kellerhals M, Costa F, Dondini L, Pagliarani G, Gregori R, Tartarini S, Leumann L, Laurens F, Patocchi A (2016) Development of SNP-based assays for disease resistance and fruit quality traits in apple (Malus×domestica Borkh.) and validation in breeding pilot studies. Tree Genet. Genomes12(3):35. https://doi.org/10.1007/s11295-016-0994-y
27. Wang H, Zhao S, Mao K, Dong Q, Liang B, Li C, Wei Z, Li M, Ma F (2018) Mapping QTLs for water-use efficiency reveals the potential candidate genes involved in regulating the trait in apple under drought stress. BMC Plant Biol 18(1):136. https://doi.org/10.1186/s12870-018-1308-3
28. Shi Z, Liu S, Noe J, Arelli P, Meksem K, Li Z (2015) SNP identification and marker assay development for high-throughput selection of soybean cyst nematode resistance. BMC Genomics 16(1):314. https://doi.org/10.1186/s12864-015-1531-3
29. Yuan J, Wen Z, Gu C, Wang D (2014) Introduction of high throughput and cost effective SNP genotyping platforms in soybean. Plant Genet Genomics Biotechnol 2(1):90-94. https://doi.org/10.5147/pggb.v2i1.155
30. Maughan PJ, Smith SM, Fairbanks DJ, Jellen EN (2011) Development, characterization, and linkage mapping of single nucleotide polymorphisms in the grain amaranths (*Amaranthus* sp.). The Plant Genome 4(1):92-101. https://doi.org/10.3835/plantgenome2010.12.0027
31. Sollars ES, Harper AL, Kelly LJ, Sambles CM, Ramirez-Gonzalez RH, Swarbreck D et al (2017) Genome sequence and genetic diversity of European ash trees. Nature541(7636):212-216. https://doi.org/10.1038/nature20786
32. Ryu J, Kim WJ, Im J, Kim SH, Lee KS, Jo HJ, Kim EY, Kang SY, Lee JH, Ha BK (2018) Genotyping-by-sequencing based single nucleotide polymorphisms enabled Kompetitive Allele Specific PCR marker development in mutant Rubus genotypes. EJB Electron J Biotechnol35:57-62. https://doi.org/10.1016/j.ejbt.2018.08.001
33. Ryu J, Kim WJ, Im J, Kang KW, Kim SH, Jo YD, Kang SY, Lee JH, Ha BK (2019) Single nucleotide polymorphism (SNP) discovery through genotyping-by-sequencing (GBS) and genetic characterization of Dendrobium mutants and cultivars. Sci Hortic 244:225-233. https://doi.org/10.1016/j.scienta.2018.09.053
34. Fall LA, Clevenger J, McGregor C (2018) Assay development and marker validation for marker assisted selection of *Fusarium oxysporum* f. sp. niveum race 1 in watermelon. Mol Breed38(11):130. https://doi.org/10.1007/s11032-018-0890-2
35. Khera P, Upadhyaya HD, Pandey MK, Roorkiwal M, Sriswathi M, Janila P, Guo Y et al (2013) Single nucleotide polymorphism–based genetic diversity in the reference set of peanut (*Arachis* spp.) by developing and applying cost-effective kompetitive allele specific polymerase chain reaction genotyping

assays. The Plant Genome 6(3). https://doi.org/10.3835/plantgenome2013.06.0019

36. Zhao S, Li A, Li C, Xia H, Zhao C, Zhang Y, Hou L, Wang X (2017) Development and application of KASP marker for high throughput detection of AhFAD2 mutation in peanut. EJB Electron J Biotechnol 25:9-12. https://doi.org/10.1016/j.ejbt.2016.10.010
37. Steele KA, Quinton-Tulloch MJ, Amgai RB, Dhakal R, Khatiwada SP, Vyas D, Heine M, Witcombe JR (2018) Accelerating public sector rice breeding with high-density KASP markers derived from whole genome sequencing of indica rice. Mol Breed 38(4):38. https://doi.org/10.1007/s11032-018-0777-2
38. Eagles HA, Hyles J, Wilson J, Cane K, Forrest KL, Hayden MJ, Ramm K, Trevaskis B (2018) A linked SNP marker to genotype Fr-B2 in wheat. Crop Pasture Sci 69(9):859-863. https://doi.org/10.1071/CP18248
39. Liu J, Xu Z, Fan X, Zhou Q, Cao J, Wang F, Ji G, Yang L, Feng B, Wang T (2018) A genome-wide association study of wheat spike related traits in China. Front Plant Sci 9:1584. https://doi.org/10.3389/fpls.2018.01584
40. Rasheed A, Jin H, Xiao Y, Zhang Y, Hao Y, Zhang Y, Hickey LT, Morgounov AI, Xia X, He Z (2019) Allelic effects and variations for key bread-making quality genes in bread wheat using high-throughput molecular markers. J Cereal Sci 85:305-309. https://doi.org/10.1016/j.jcs.2018.12.004
41. Yang Y, Basnet BR, Ibrahim AM, Rudd JC, Chen X, Bowden RL, Xue Q, Wang S et al (2019) Developing KASP Markers on a Major Stripe Rust Resistance QTL in a Popular Wheat TAM 111 Using 90K Array and Genotyping-by-Sequencing SNPs. Crop Sci 59(1):165-175. https://doi.org/10.2135/cropsci2018.05.0349
42. Doyle JJJL, Doyle JL (1987) Genomic plant DNA preparation from fresh tissue-CTAB method. Phytochem Bull 19(11):11-15.
43. Kafkas S, Khodaeiaminjan M, Güney M, Kafkas E (2015) Identification of sex-linked SNP markers using RAD sequencing suggests ZW/ZZ sex determination in *Pistacia vera*L. BMC Genomics 16(1):98. https://doi.org/10.1007/BF00226975
44. Yakubov B, Barazani O, Golan-Goldhirsh A (2005) Combination of SCAR primers and Touchdown-PCR for sex identification in *Pistacia vera*L. Sci Hortic 103(4):473-478. https://doi.org/10.1016/j.scienta.2004.06.008
45. Kafkas SK, Cetiner GD, Perl-Treves R, Nissim-Levi AN (2001) Development of sex-associated RAPD markers in wild Pistacia species. J Hortic Sci Biotechnol 76(2):242-246. https://doi.org/10.1080/14620316.2001.11511358
46. Esfandiyari B, Nejad GD, Kiani FSM (2010) Sex determination in *Pistacia*species using molecular markers. J Appl Hortic 12(2):122-124. https://doi.org/10.1007/s10681-011-0527-6
47. Esfandiyari B, Davarynejad GH, Shahriari F, Kiani M, Mathe A (2012) Data to the sex determination in *Pistacia*species using molecular markers. Euphytica185(2):227-231. https://doi.org/10.1007/s10681-011-0527-6
48. Ehsanpour AA, Tavassoli M, Arab L (2008) Sex determination of *Pistacia vera*L. using ISSR markers. Malaysian Appl Biol 37(2):25-28. https://doi.org/10.1007/s10681-011-0527-6
49. Ehsanpour A, Arab L (2010) Application of RAPD (Random Amplified Polymorphic DNA) marker for sex determination of *Pistacia vera* L. J Cell Mol Res 1(2):68-71. https://doi.org/10.22067/jcmr.v1i2.3221
50. Kamiab F, Ebadi A, Panahi B, Tajabadi A (2014) RAPD Analysis for Sex Determination in *Pistacia vera*L. J Nuts 5(1):51-55. https://doi.org/10.22034/JON.2014.515697
51. Vendramin E, Dettori MT, Verde I, Micali S, Giovinazzi J, Mardi M, Avanzato D, Quarta R (2007) Molecular characterization of Pistacia genus by microsatellite markers. I Balkan Symposium on Fruit Growing 825:55-62. https://doi.org/10.17660/ActaHortic.2009.825.5
52. Devran Z, Göknur A, Mesci L (2016) Development of molecular markers for the *Mi-1* gene in tomato using the KASP genotyping assay. HEB 57(2):156-160. https://doi.org/10.1007/s13580-016-0028-6
53. Carmichael SN, Bekaert M, Taggart JB, Christie HR, Bassett DI, Bron JE, Skuce PJ et al (2013) Identification of a sex-linked SNP marker in the salmon louse (*Lepeophtheirus salmonis*) using RAD sequencing. Plos One 8(10). https://doi.org/10.1371/annotation/9a5a8f89-38b6-4ba3-aceb-65aa256f6157
54. Palaiokostas C, Bekaert M, Davie A, Cowan ME, Oral M, Taggart JB, Gharbi K et al (2013) Mapping the sex determination locus in the Atlantic halibut (*Hippoglossus hippoglossus*) using RAD sequencing. BMC Genomics14(1):566. https://doi.org/10.1186/1471-2164-14-566
55. De Boer JG, Groenen MA, Pannebakker BA, Beukeboom LW, Kraus RH (2015) Population-level consequences of complementary sex determination in a solitary parasitoid. BMC Evol Biol 15(1):98. https://doi.org/10.1186/s12862-015-0340-2

## Tables

**Table 1.** Sanger sequence of primer pairs for four target regions.

| Region | Primer forward | Primer reverse |
|---|---|---|
| **REG-PIS-136404 (exp#: 91951)** | TTGAGGAAGAGGATGAAACC | TGAAATAGGATGAAAATCGTG |
| **REG-PIS-174431 (exp#:107303)** | TCCTTTTCTCTTTCCCTAAGC | GCTACCCTAGCGAAAGATCC |
| **REG-PIS-133396** | GATGGCCAGTGAGAGGATTTAT | GCATGGGGACACAGAGTTAG |
| **REG-PIS-167992** | CGTGACTTAAAGGACAACAGC | GTGCTATGCGCATTTCATTAAAATC |

**Table 2.** KASP primer sequences.

| NO | SNP ID | FAM | HEX | Sequence | KASP Primer Seq Allele FAM |
|---|---|---|---|---|---|
| 1 | SNP-PIS-133396 | C | G | AAGAAGATTAAAGTA[C/G]CTGATCCATGATCTT | GCAAACCGCAAAGAAGATTAAAGTAC |
| 2 | SNP-PIS-167992 | A | T | TCATAGCGTGAGCTC[A/T]TTGTTCCACCTSTAA | GAAAAATAACTTCATAGCGTGAGCTCA |
| 3 | SNP-PIS-176863 | A | G | TGAGCTTGGAAGGTA[A/G]TATCATACTTTGCAGAT | ATCGACCTGTCTGCAAAGTATGATAT |
| 4 | P-ATL-91951-565 | T | C | TTAATGCAGTTGAGG[T/C]TAGCGGTAAGAATGT | CAGAGAAAGTTATTAATGCAGTTGAGGT |
| 5 | P-INT-91951-120 | T | A | ACATGAGCAGCAAGT[T/A]CAGAGCATAACAAAT | GAATGAACAATAACATGAGCAGCAAGTT |
| 6 | P-KHIN-91951-115 | C | T | AATAGTACTTCTTTG[C/T]GCATTTTTTATTCAT | GATGGTTTCTTTGAATAGTACTTCTTTG |
| 7 | P-VERA 10733-132 | A | T | GTCACGTTCCCGCTC[A/T]CTTCTGTCTTTGTCA | CACCGTCACGTTCCCGCTCA |
| 8 | P-VERA-91951-460 | T | G | TATATTTTACTTATG[T/G]AGTTATAGTATGTGC | TAATCAGCTTTGAAGTATATTTTACTTATC |
| 9 | P-INT-91951-256 | A | G | TTATGTTGTTAGATT[A/G]AGATTAAAGATATAC | TGTTTCAGAGAGGTATATCTTTAATCTT |
| 10 | P-PAL-107303-INDEL | DEL | INS | TCTTTCACCGTCACG[TTCCCGCTCTCTTCTGTCTTTGTCATGATCACA/-]TTCCCGCTCTCTTCT | CTCTTCTGTCTTTGTCATGATCACA |
| 11 | P-KHIN-133396-479 | C | A | ATTTATATTTGCTGC[C/A]TCTGTTATATTTTAT | GTCAAAACCATGCAATAAAATATAACAGA |
| 12 | P-TRE-167992-63 | T | C | GTATGTATATGGCTT[T/C]ACAGTATCACCTGTA | CAAGCCCTTACAGGTGATACTGTA |

## Figures

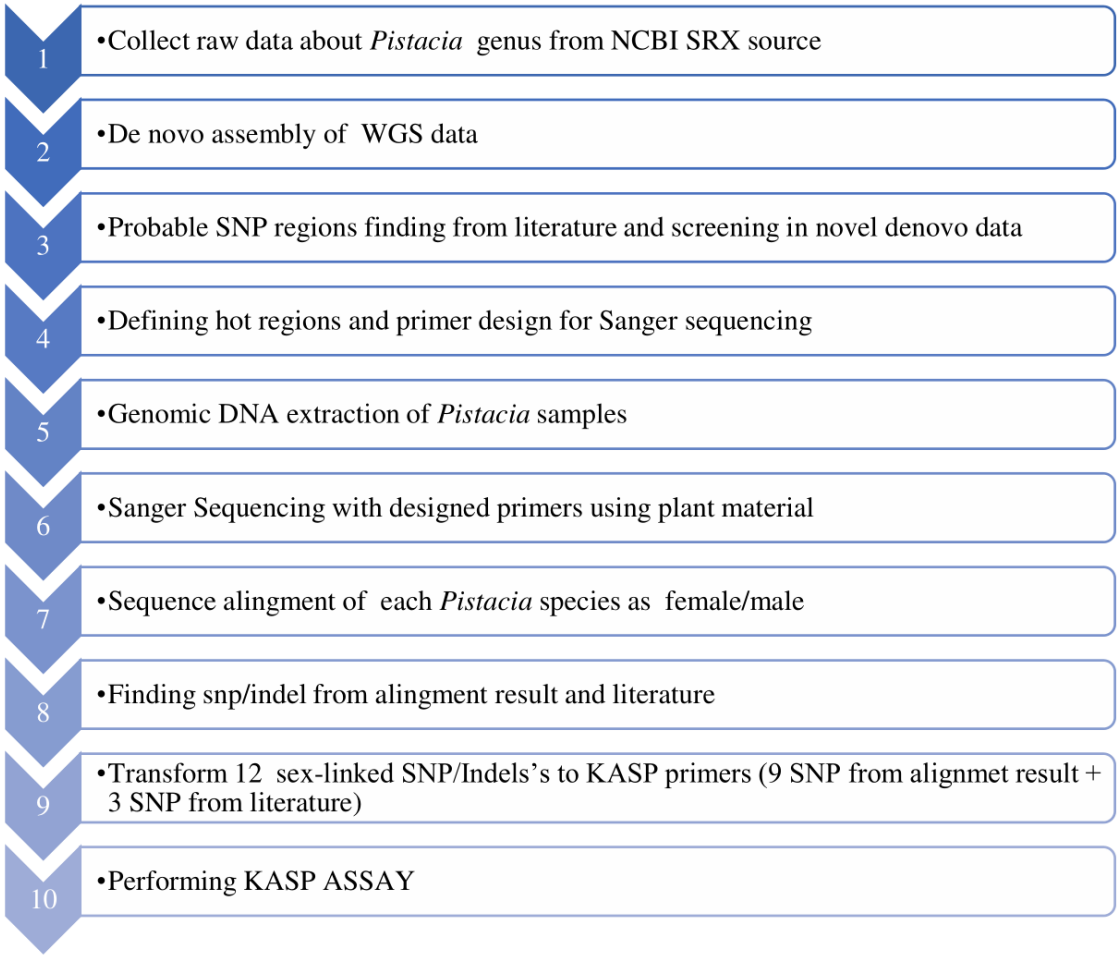


**Figure 1**

Working flow chart of this study.

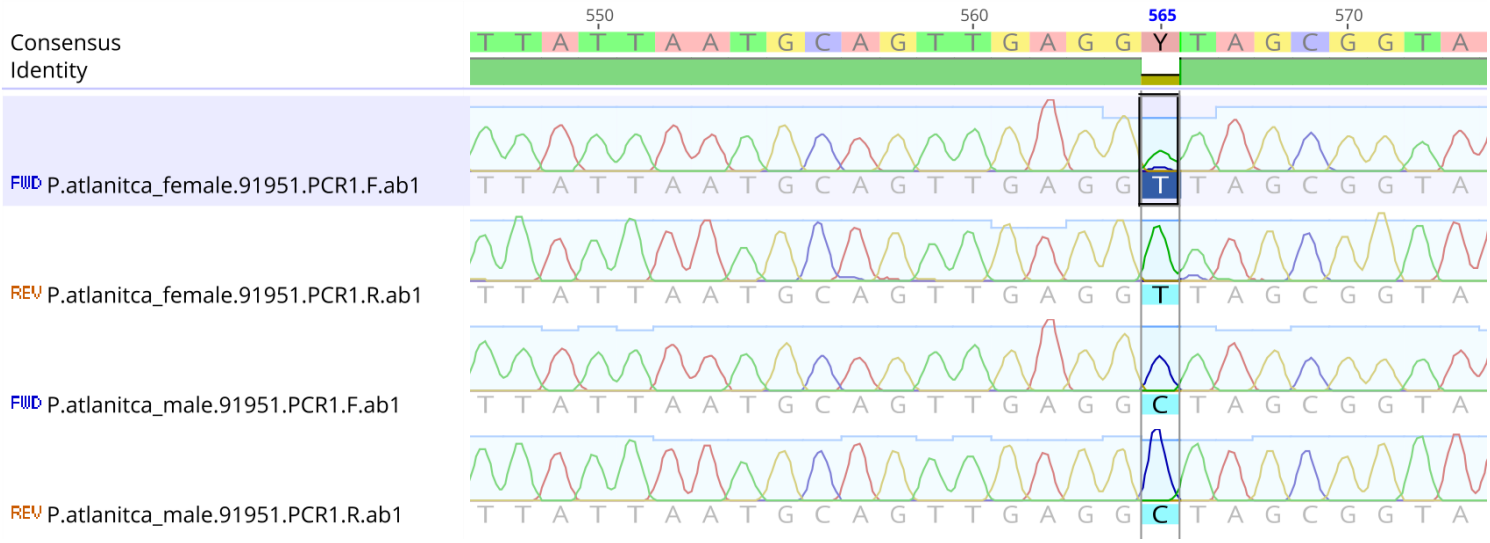


**Figure 2**

Variant detection followed by multiple alignment of Sanger sequencing results using Geneious Prime software: Example of a SNP mutation for P-ATL-91951-565.

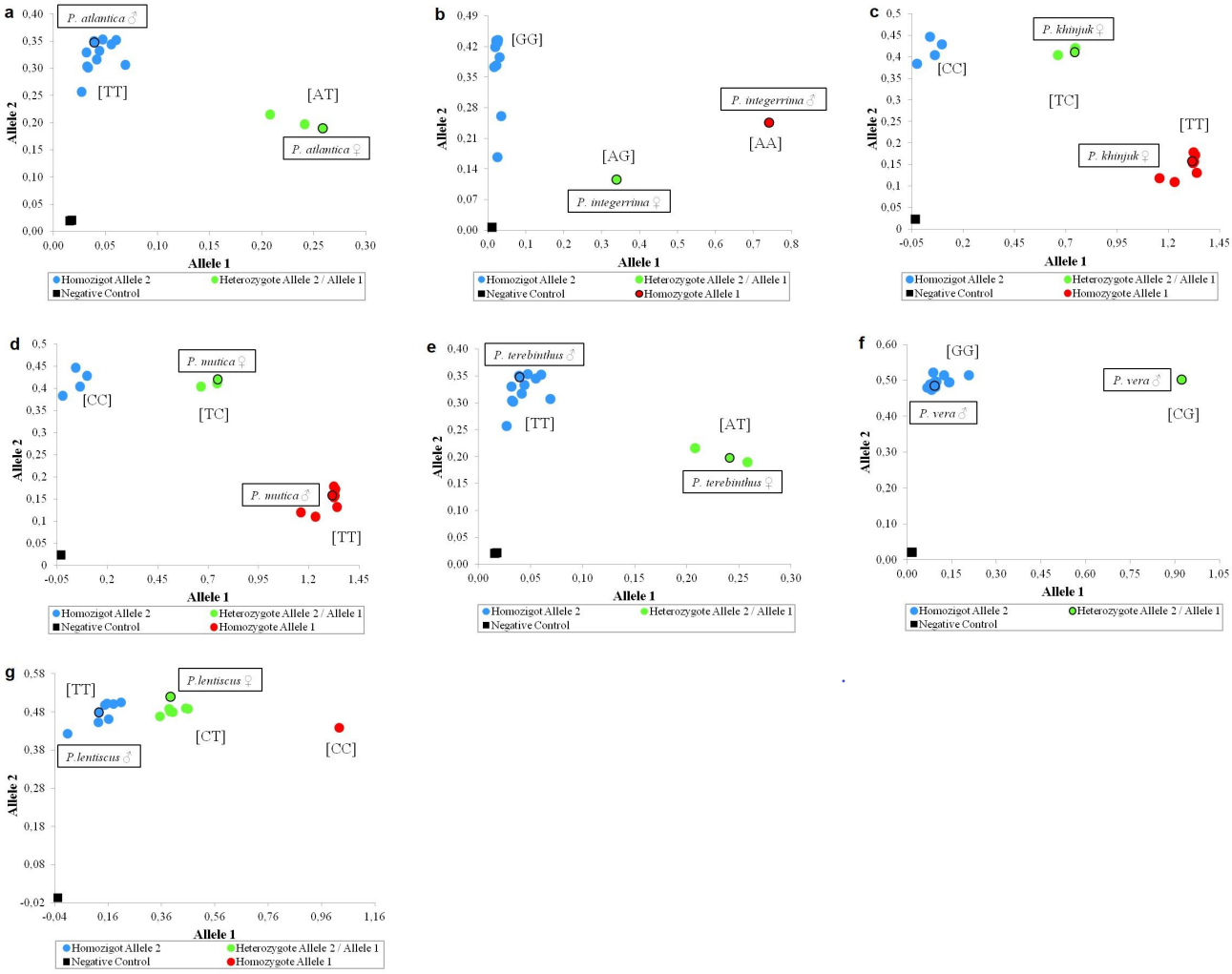


**Figure 3**

KASP genotyping results of P. atlantica (a), P. integerrima (b), P. khinjuk (c), P. mutica (d), P. terebinthus (e). P. vera (f) and P. lentiscus (g) species with 5 SNP marker identified gender–based SNP differences.

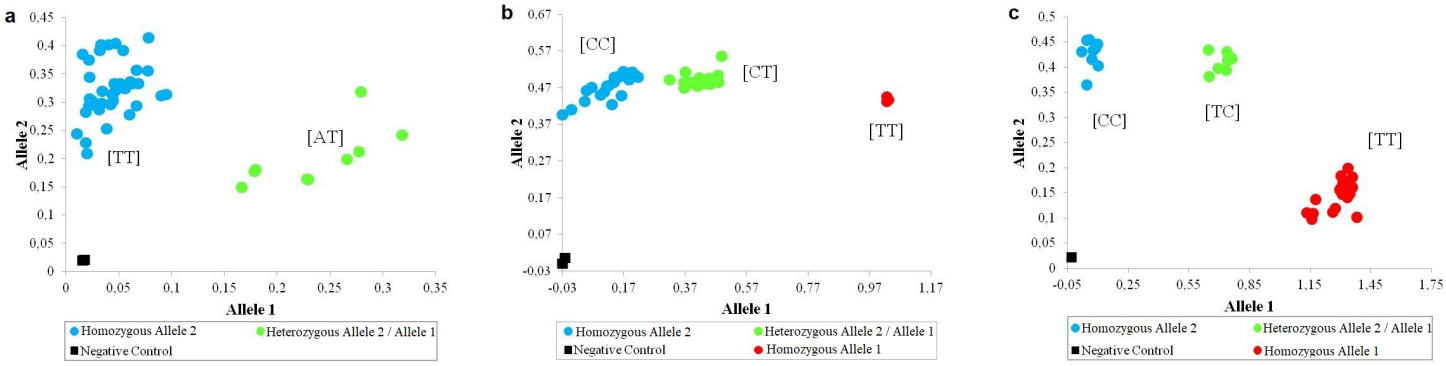


**Figure 4**

4 KASP genotyping results for all species used in this study (a) SNP-PIS-167992 (b) P-KHIN-91951-115 (c) P-ATL-91951-565 assays.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Appendices.docx

- CoverLetterforMolecularBiologyReports.pdf