Authentication of olive oil based on DNA analysis

A. Batrinoua, I.F. Stratia, D. Houhoula, J. Tsaknis and V.J. Sinanoglou

Department of Food Science and Technology
Faculty of Food Sciences
University of West Attica, Athens, Greece

Corresponding authors: batrinou@uniwa.gr; estrati@uniwa.gr

SUMMARY: Olive oil, which has been produced mainly in the Mediterranean area since the ancient times, has a high nutritional value linked to many health benefits. Extra virgin, which is the purest form of olive oil, has excellent quality and premium prices. Many cases of adulteration and fraud necessitate the development of reliable and accurate methods for olive oil authentication. DNA-based methods analyze the residual DNA extracted from olive oil and use molecular markers for genetic identification of different species, subspecies or cultivars because these markers act as signs which reflect distinct genetic profiles. This study reviews the process by which DNA from olive oil is extracted and analyzed by the most recently used markers in the authentication of olive oil, such as Simple Sequence Repeats (SSR) or microsatellites and the single nucleotide polymorphisms (SNPs). Methods of analysis such as qPCR and digital PCR are also discussed with a special emphasis placed on the method of High-Resolution Melting (HRM), a post-polymerase chain reaction method, which enables rapid, high performing identification of genetic variants in the DNA regions of interest without sequencing, and may differentiate very similar cultivars which differ in only one nucleotide in a specific locus.

KEYWORDS: Authentication; Genetic Markers; (HRM); Olive oil; (SNPs); (SSR)

RESUMEN: Autenticación de aceite de oliva mediante análisis de ADN. El aceite de oliva, producido principalmente en el área mediterránea desde la antigüedad, tiene un alto valor nutricional vinculado a muchos beneficios para la salud. El aceite de oliva virgen extra, que es la forma más pura de aceite de oliva, tiene una excelente calidad y precios premium. Muchos casos de adulteraciones y fraudes requieren el desarrollo de métodos fiables y precisos para la autenticación del aceite de oliva. Los métodos basados en el ADN analizan el ADN residual extraído del aceite de oliva y usan marcadores moleculares para la identificación genética de diferentes especies, subespecies o cultivares, porque estos marcadores actúan como signos que producen perfiles genéticos distintos. Este estudio revisa el proceso mediante el cual el ADN del aceite de oliva es extraído y analizado por los marcadores utilizados más recientemente en la autenticación del aceite de oliva, como las repeticiones de secuencia simple (SSR) o los microsatélites y los polimorfismos de un solo nucleótido (SNP). Los métodos de análisis como qPCR y PCR digital también se analizan haciendo especial énfasis en el método de fusión de alta resolución (HRM), un método de reacción en cadena posterior a la polimerasa, que permite la identificación rápida y con alto rendimiento de variantes genéticas en regiones del ADN de interés sin secuenciación, y pueden diferenciar cultivares muy similares, que difieren en un solo nucleótido, en un lugar específico.

PALABRAS CLAVE: Aceite de oliva; Autenticación; (HRM); Marcadores genéticos; (SNPs); (SSR)

ORCID ID: Batrinou A https://orcid.org/0000-0001-8219-7537, Strati IF https://orcid.org/0000-0002-0656-3707, Houhoula D https://orcid.org/0000-0002-4692-2081, Tsaknis J https://orcid.org/0000-0002-9152-1228, Sinanoglou VJ https://orcid.org/0000-0001-8109-5363


Copyright: ©2020 CSIC. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International (CC BY 4.0) License.
1. INTRODUCTION

The cultivation of olive trees began thousand years ago in the Mediterranean area (fossil pollen was found in Macedonia, Greece and fossilized leaves were found on the island of Santorini, Greece, dated to 37,000 BP (NCBI)) and it has developed into the most important oil-producing crop in the Mediterranean region, covering more than 10 million hectares of land worldwide (FAOSTAT, 2016). The high nutritional value and health benefits of olive oil and in particular Virgin olive oil (VOO) represent a hallmark of the Mediterranean Diet, and have been established by various studies (Parkinson and Cicere, 2016; Martinez-Gonzalez and Martin-Calvo, 2016; Trichopoulou et al., 2014). The European Commission has recently permitted health claims concerning olive oil polyphenols, especially hydroxytyrosol and its derivatives (e.g. oleuropein complex and tyrosol) (Regulation EU 432/2012).

The growing demand for olive oil and the high prices consumers are willing to pay for high quality products such as extra virgin olive oil (EVOO) have resulted in many cases of adulteration and fraud in recent years (Bajoub et al., 2018; Ramos-Gomez et al., 2016). The European Union has applied laws regulating the classification and labeling of olive oil in order to protect its geographical origin and quality. Commercial olive oils that are produced from a specific region and follow a particular traditional production process may be labeled as Protected Designation of Origin (PDO) or Protected Geographical Indication (PGI) (European Council Regulation, 2006).

The olive tree belongs to the species Olea europaea, which is actually a complex (the olive complex) comprising of six subspecies and more than 1250 cultivars that have different genotypes (differences in the sequence of DNA) (Bracci et al., 2011). The Olea europaea subsp. europaea, which has spread across the Mediterranean basin, has 46 chromosomes and is diploid (has two copies of each chromosome) (Besnard et al., 2008; Besnard et al., 2009). The high number of different olive genotypes which are currently cultivated necessitates the application of DNA-based methods that use genetic markers to assess varietal identification, genetic authenticity, traceability and adulteration (Sebastiani and Busconi, 2017).

DNA-based methods that test the authenticity of foods rely on the fact that the residual DNA extracted from any food matrix can be analyzed and provides information about the genetic identity of any species, subspecies or variety present in the matrix. In the case of olive oil, DNA analysis can reveal the presence of unique DNA sequences of different plant species (e.g. hazel nut oil in olive oil, Vietina et al., 2013), different varieties (e.g. the presence of non-PDO cultivars in a product claiming to be PDO monovarietal, Pasqualone et al., 2007) and can also resolve taxonomic issues such as synonymies (different names for the same variety), homonymies (same name for different variety) and misnaming (Muzzalupo et al., 2015), thus providing a powerful tool for discrimination analysis (Sebastiani and Busconi, 2017; Kalogianni et al., 2015; Madesis et al., 2014; Galimberti et al., 2013; Ben-Ayed et al., 2013).

The unique DNA sequences found in each organism (and thus in each plant), which are called genetic or molecular markers, are specific sequences of DNA with known physical locations on chromosomes and they constitute points of variation that can be used to identify individuals or species (NCI 2019, Nature 2019).

The common genetic markers which have been used for the identification of olive cultivars and the traceability of olive oil are the “simple sequence repeats” (called SSR or microsatellites, which are tandem repeats of di-, tri-, tetra-, and penta-nucleotides, Bazakos et al., 2016a; Kalogianni et al., 2015) and the “single nucleotide polymorphisms” (called SNPs, Abuzayed et al., 2018; Mousavi et al., 2017). Other markers have also been used as techniques to identify genetic variation, such as “randomly amplified polymorphic DNA” (called RAPD consisting of 10-mer primers that randomly anneal to several loci leading to the amplification of many sequences by PCR, providing many markers per assay, Belaj et al., 2003, Muzzalupo and Perri, 2002), “amplified fragment length polymorphism” (called AFLP which is based on the detection of restriction fragments by PCR amplification, Montemurro et al., 2008; Pafundo et al., 2005) and “inter-simple sequence repeat” (called ISSR, which are DNA fragments of about 100-3,000 bp present at an amplifiable distance in between two identical microsatellite repeat regions oriented in the opposite direction, Terzopoulos et al., 2005; Gemas et al., 2004; Pasqualone et al., 2001). RAPD, AFLP and ISSR have advantages when analyzing genetic relationships in olive trees mainly because the design of primers for the PCR do not require prior knowledge of the genome sequence. However, these markers have the drawback of low reproducibility in olive oil studies (reviewed by Pasqualone et al., 2016; Sebastiani and Busconi, 2017; Bazakos et al., 2016, Ben-Ayed et al., 2013). The genetic markers used by the most recent studies for olive oil authentication are simple sequence repeats (SSR) and single nucleotide polymorphisms (SNPs) (Pasqualone et al., 2016) and are located either in nuclear DNA or in the DNA found in organelles (chloroplast or mitochondrial DNA) (Pereira et al., 2018).

Once the DNA has been efficiently recovered from a sample, a specific region of DNA (a target sequence which is usually a genetic marker) is
amplified in many copies using the Polymerase Chain Reaction (PCR) so that an adequate amount of DNA is formed to permit further analysis. The next step is generally called “genotyping” and involves a number of methods by which differences in the DNA sequence of the sample are visualized and identified (Figure 1). Differences in the sequence of DNA (called variants or polymorphisms, which are any difference in the nucleotide sequence between individuals) can be single-base pair changes (nucleotide substitutions), deletions, insertions, or even variations in the number of copies of a given DNA sequence (Karki et al., 2015).

For genotyping (identifying the base sequence and thus revealing the polymorphisms), PCR amplicons (copies) of target-DNA can be sequenced, or analyzed by agarose gel electrophoresis or high-resolution capillary electrophoresis (CE). Capillary electrophoresis has a much lower limit of detection for DNA fragments. It can discriminate DNA sequences which differ by only few nucleotides in length and is more automated compared to agarose gel electrophoresis (Alvarez-Rivera et al., 2018; Lian and Zeng, 2017; Spaniolas et al., 2010; 2008). Techniques such as quantitative real-time PCR or high-resolution melting analysis (HRM) are increasingly used as they offer the advantage of simultaneous PCR amplification and genotyping, as well as the technique of Next-Generation Sequencing (NGS), which provides massive data of parallel short-read per run, lowering the analysis cost and revealing DNA fragments from a variety of species at the same time (Lo and Shaw, 2018; Madesis et al., 2014).

This review analyses the latest advances in the application of DNA-based methods in the authentication of olive oil with a special emphasis on the use of the more recently developed species-specific molecular markers such as simple sequence repeats

![Figure 1](https://doi.org/10.3989/gya.0682191)
(SSR) and single nucleotide polymorphisms (SNPs) and high-throughput methods such as High-Resolution Melting (HRM) as methods that present numerous advantages for varietal identification.

2. EXTRACTION OF DNA

Food genomics are commonly applied for assessing food safety, authenticity and traceability due to the many advantages of the DNA-based methodologies compared to other techniques based on the analysis of metabolites. The first step in all DNA-based authenticity tests is the extraction of DNA from the sample to be analyzed. DNA is a thermally stable molecule and can be detected even in small fragments in any type of raw, cooked or processed food; whereas proteins are denatured under high temperatures or when processed.

In olive oil, the main advantage to extracting DNA is that its sequence is not influenced by environmental factors; whereas the metabolic profile of olive oil such as polyphenol content and fatty acids is strongly affected by environmental factors (e.g. climate conditions and soil composition) as well as other factors such as stage of ripeness, harvesting time, storage period and extraction method (Ramos-Gómez et al., 2016; Kalaitzis and El-Zein, 2016; Ben-Ayed et al., 2013; Agrimonti et al., 2011; Montalegre et al., 2010; Heimler et al., 2004). Therefore, the base sequence of DNA extracted from olive oil can be used as a genetic fingerprint to reveal the identities of different oils that might be present in the oil. However, one of the major limiting steps of the process is that DNA extracted from olive oil can be very fragmented and may contain phenolic compounds or residual polysaccharides which can act as inhibitors of DNA polymerase which may affect the quality of the PCR reaction by yielding non-specific PCR amplifications (Raieta et al., 2015). However, recent studies have reported the development of a DNA polymerase enzyme which is more tolerant to PCR inhibitors and can be applied directly to crude tissues (Chiappetta et al., 2017; Muzzalupo et al., 2015). Another limitation may be the discrepancy between the results obtained from the DNA extracted from leaves as opposed to the DNA extracted from oils, which suggests differences in the quality of the extracted DNA (Ramos-Gómez et al., 2016; Pasqualone et al., 2007; Pasqualone et al., 2016). The correspondence between the profiles of microsatellites of oil and leaves can be significantly improved by shortening the target sequences for PCR amplification to as low as 80 bp (Ramos-Gomez et al., 2014; Ramos-Gomez et al., 2016; Bracci et al., 2011). The amount and quality of DNA isolated from oil samples remain critical parameters for the success of subsequent analyses (Sollo et al., 2016). DNA extraction from olive oil has been successfully achieved by various protocols such as CTAB and commercial kits such as Wizard kit (Promega), QIAamp DNA stool extraction kit (Qiagen), Nucleospin Food and Nucleospin Plant (Macherey Nagel), as reviewed by Ben-Ayed et al., (2013) and Agrimonti et al., (2011).

3. SIMPLE SEQUENCE REPEATS (SSR)

Simple Sequence Repeats (SSR) or microsatellites are DNA sequences of 1 to 10 nucleotides long. They occur as tandem repeats (TRs) of di-, tri-, tetra-, and penta-nucleotides and are ubiquitous in eukaryotic genomes (Gupta et al., 1994). SSRs are widely distributed throughout the genome, especially in the euchromatin of eukaryotes, and in the coding and non-coding DNA of the nucleus and organelles. SSRs have been extensively used as markers for genotyping plants and olive oil varieties over the past 15 years (Table 1) because they are codominant (the alleles in a gene pair in a heterozygote are both expressed), multi-allele (multiple alleles are found in the population of the same species) and highly informative genetic markers which are experimentally reproducible and commonly found among related species (Ben-Ayed et al., 2013; Ben-Ayed et al., 2016). The occurrence of SSRs is usually lower in gene regions, probably due to the fact that SSRs have a high mutation rate that could compromise gene expression. SSRs exhibit repeat polymorphisms that usually result from the addition or deletion of the entire repeat units or motifs. Therefore, different individuals exhibit variations (polymorphisms) in SSRs as differences in the repeat numbers of specific DNA sequences (Vieira et al., 2016). An example in Olea europaea is the microsatellite (GA)_n, which is the dinucleotide GA repeated n times in a specific region of a specific chromosome (genetic locus). The polymorphism lies in the fact that different varieties of the same species may have a different number of repeats, e.g. one SSR locus is named ssrOeUA-DCA1 (Sefc et al., 2000) and has 22 repetitions of the dinucleotide GA symbolized as (GA)_22, but this allele is polymorphic, meaning that in some individuals (olive trees) a different number of repetitions is found on the same locus. Therefore, by identifying the number of repeats found in the DNA isolated from a sample, a process performed by amplifying the target sequence by PCR amplification using specific primers, discrimination of the variety can be achieved. Usually, a combination of SSR markers is used to provide a more powerful discriminatory analysis (Bazakos et al., 2016). Some applications include expressed sequence tag SSR (EST-SSR), which derive from regions of the genome that are expressed (sequence coding for a specific protein) and therefore these regions have greater transferability among species and their variation can be associated with the phenotype (Duran et al., 2009).
4. SINGLE NUCLEOTIDE POLYMORPHISMS (SNPS)

Single nucleotide polymorphisms (SNPs) are the most common type of DNA polymorphism in most organisms and they have gained popularity as forensic markers in all areas of biomedical and genetic research. SNPs can discriminate samples of the same species or variety which differ in only one nucleotide in a specific genetic locus (a locus could be any coding or non-coding region in the DNA sequence of an organism). An example of a SNP...
would be if a cytosine (C) nucleotide is present at a particular locus in one individual's DNA but a thymine (T) nucleotide occurs at the same locus in another individual's DNA.

The advantages to SNPs also apply to olive oil authentication: a) they are the most abundant genetic markers; b) they have lower mutation rates and are stably inherited; c) they can be detected within a low size range of amplicons (e.g. 100bp) so they are a better choice for DNA that has been degraded and extracted from a complex matrix such as olive oil, d) they are mostly biallelic (a specific locus in a genome contains two observed alleles), e) they are able to differentiate very similar cultivars that may differ in only one nucleotide in a specific locus (Mehta et al., 2017; Kalogianni et al., 2015; Bracci et al., 2011; Agrimonti et al., 2011).

Most of the methods used to identify the points of variations, therefore genotyping SNPs, rely on amplification by the PCR of the genomic regions that span the SNPs before the actual genotyping reaction. Also, the PCR is sensitive and specific enough to distinguish between heterozygous and homozygous SNP genotypes in large, complex genomes (Syvanen, 2001).

5. QUANTITATIVE REAL-TIME PCR (qPCR) IN AUTHENTICATION OF OLIVE OIL

The quantification of an ingredient is essential in order to verify an adulteration, whether intended or unintended, or to evaluate if a limit set by law has been exceeded (Madesis et al., 2014). Quantitative Real-time PCR (qPCR) measures the number of copies of the marker gene (the gene that codes for the ingredient or adulterant) and compares it to the copy number of a ‘normalizing’ gene, thus providing a percentage of the ingredient or adulterant (Primrose et al., 2010). qPCR is more sensitive than conventional PCR and enables the detection and quantification of DNA in very low amounts. Another advantage is that the molecular markers used in qPCR are designed to amplify small fragments of DNA; therefore, the specific SNPs before the actual genotyping reaction. Also, the PCR is sensitive and specific enough to distinguish between heterozygous and homozygous SNP genotypes in large, complex genomes (Syvanen, 2001).

5. QUANTITATIVE REAL-TIME PCR (qPCR) IN AUTHENTICATION OF OLIVE OIL

The quantification of an ingredient is essential in order to verify an adulteration, whether intended or unintended, or to evaluate if a limit set by law has been exceeded (Madesis et al., 2014). Quantitative Real-time PCR (qPCR) measures the number of copies of the marker gene (the gene that codes for the ingredient or adulterant) and compares it to the copy number of a normalizing gene, thus providing a percentage of the ingredient or adulterant (Primrose et al., 2010). qPCR is more sensitive than conventional PCR and enables the detection and quantification of DNA in very low amounts. Another advantage is that the molecular markers used in qPCR are designed to amplify small fragments of DNA; therefore, the specific SNPs before the actual genotyping reaction. Also, the PCR is sensitive and specific enough to distinguish between heterozygous and homozygous SNP genotypes in large, complex genomes (Syvanen, 2001).

5. QUANTITATIVE REAL-TIME PCR (qPCR) IN AUTHENTICATION OF OLIVE OIL

The quantification of an ingredient is essential in order to verify an adulteration, whether intended or unintended, or to evaluate if a limit set by law has been exceeded (Madesis et al., 2014). Quantitative Real-time PCR (qPCR) measures the number of copies of the marker gene (the gene that codes for the ingredient or adulterant) and compares it to the copy number of a normalizing gene, thus providing a percentage of the ingredient or adulterant (Primrose et al., 2010). qPCR is more sensitive than conventional PCR and enables the detection and quantification of DNA in very low amounts. Another advantage is that the molecular markers used in qPCR are designed to amplify small fragments of DNA; therefore, the specific SNPs before the actual genotyping reaction. Also, the PCR is sensitive and specific enough to distinguish between heterozygous and homozygous SNP genotypes in large, complex genomes (Syvanen, 2001).

5. QUANTITATIVE REAL-TIME PCR (qPCR) IN AUTHENTICATION OF OLIVE OIL

The quantification of an ingredient is essential in order to verify an adulteration, whether intended or unintended, or to evaluate if a limit set by law has been exceeded (Madesis et al., 2014). Quantitative Real-time PCR (qPCR) measures the number of copies of the marker gene (the gene that codes for the ingredient or adulterant) and compares it to the copy number of a normalizing gene, thus providing a percentage of the ingredient or adulterant (Primrose et al., 2010). qPCR is more sensitive than conventional PCR and enables the detection and quantification of DNA in very low amounts. Another advantage is that the molecular markers used in qPCR are designed to amplify small fragments of DNA; therefore, the specific SNPs before the actual genotyping reaction. Also, the PCR is sensitive and specific enough to distinguish between heterozygous and homozygous SNP genotypes in large, complex genomes (Syvanen, 2001).

5. QUANTITATIVE REAL-TIME PCR (qPCR) IN AUTHENTICATION OF OLIVE OIL

The quantification of an ingredient is essential in order to verify an adulteration, whether intended or unintended, or to evaluate if a limit set by law has been exceeded (Madesis et al., 2014). Quantitative Real-time PCR (qPCR) measures the number of copies of the marker gene (the gene that codes for the ingredient or adulterant) and compares it to the copy number of a normalizing gene, thus providing a percentage of the ingredient or adulterant (Primrose et al., 2010). qPCR is more sensitive than conventional PCR and enables the detection and quantification of DNA in very low amounts. Another advantage is that the molecular markers used in qPCR are designed to amplify small fragments of DNA; therefore, the specific SNPs before the actual genotyping reaction. Also, the PCR is sensitive and specific enough to distinguish between heterozygous and homozygous SNP genotypes in large, complex genomes (Syvanen, 2001).

5. QUANTITATIVE REAL-TIME PCR (qPCR) IN AUTHENTICATION OF OLIVE OIL

The quantification of an ingredient is essential in order to verify an adulteration, whether intended or unintended, or to evaluate if a limit set by law has been exceeded (Madesis et al., 2014). Quantitative Real-time PCR (qPCR) measures the number of copies of the marker gene (the gene that codes for the ingredient or adulterant) and compares it to the copy number of a normalizing gene, thus providing a percentage of the ingredient or adulterant (Primrose et al., 2010). qPCR is more sensitive than conventional PCR and enables the detection and quantification of DNA in very low amounts. Another advantage is that the molecular markers used in qPCR are designed to amplify small fragments of DNA; therefore, the specific SNPs before the actual genotyping reaction. Also, the PCR is sensitive and specific enough to distinguish between heterozygous and homozygous SNP genotypes in large, complex genomes (Syvanen, 2001).

6. HIGH RESOLUTION MELTING (HRM)

High-resolution melting (HRM) analysis is a post-polymerase chain reaction (PCR) method that enables rapid, high-throughput identification of genetic polymorphisms in the target-DNA without sequencing. In the first step, PCR is performed to amplify the target DNA (region of interest) in many copies in the presence of a fluorescent dye that binds to double-stranded DNA (dsDNA). This dye emits fluorescence at a high intensity when it is bound in the double-stranded DNA but when DNA is thermally denatured and is converted to single-stranded random coils the dye shows low levels of fluorescence (unbound state). The thermal denaturation step occurs after PCR is completed and the amplicon (usually a 50 to 500 bp long dsDNA sequence amplified in many copies) is denatured by gradually increasing temperature (ca. by 0.01 °C to 0.2 °C). During this step, termed «melting analysis» (melting meaning thermal denaturation of dsDNA), the gradually denaturing amplicon releases the fluorescent dye.

The melting temperature (T_m = the temperature at which 50% of the DNA is single-stranded) varies among DNA molecules and depends on the length of the DNA sequence and its Guanine and Cytosine (GC) base content (in the double-stranded DNA Guanine and Cytosine bases are bound with three hydrogen bonds and are more stable than Adenine and Thymine, which are bound with 2 hydrogen bonds). Therefore, DNA sequences which have a high GC content demonstrate a higher T_m than DNA sequences that contain a low number

of GC base pairs. A curve, called melting curve, is obtained from plotting the decreasing fluorescent emission values against the increasing temperature (Figure 2). This curve has a unique profile which allows the detection of polymorphisms between DNA samples based on the differences in the melting temperature of PCR products, as low as in a single nucleotide such as SNPs (Simco, 2016; Muleo et al., 2009). The rate of fluorescent decrease is generally greatest near the melting temperature ($T_m$) of the PCR product. Therefore, a DNA sequence which has a unique combination of length, GC content and pattern of complementary bases will exhibit a unique melting curve or melting profile (Xanthopoulou et al., 2014).

Small differences between individual melting curves are better visualized by the calculation of the HRM difference (or differential) plot provided by the HRM software (Pereira et al., 2018; Song et al., 2016). This plot is constructed by setting the melting profile of one of the known genotypes as a baseline reference and then by subtracting the normalized fluorescent data by this reference melting profile as shown in Figure 3.

![Figure 2](image1.png)

**Figure 2.** HRM normalized melting curve obtained from the diminishing fluorescent emission values plotted against increasing temperature. In this hypothetical example, 5 samples of extracted DNA from the olive oil of different olive cultivars were amplified by PCR using specific primers (targeting a SNP) and analyzed by HRM. The samples were distinguished by their unique melting curves (each sample corresponds to a different color line) [adapted from Song et al., 2016].

![Figure 3](image2.png)

**Figure 3.** A HRM difference plot based on the analysis of 5 samples. One of the samples was used as a reference (reference genotype, the red line in this graph) and the fluorescence of the samples minus the fluorescence of the reference was calculated (vertical axis). This graph is used to better visualize the differences between individual melting curves [adapted from Song et al., 2016].
In general, the melting curve analysis software defines the $T_m$ of a PCR product as the inflection point of the melting curve. To visualize the $T_m$ more clearly, the negative first derivatives ($-\frac{dF}{dT}$) are often plotted, making the $T_m$ of the products appear as peaks, as seen in Figure 4.

In the case of diploid organisms (whose genome has two copies of each chromosome) such as Olea plants, two alleles (variant forms of the same gene or genetic locus) may differ in only one base and can be used as molecular markers. Individual organisms can be either homozygous for an allele (having two copies of exactly the same sequence) or heterozygous (each copy has a slight difference e.g. one different base) (Nature, 2019).

In a heterozygote (assuming that allele A1 and allele A2 differ in only one base and therefore have a single nucleotide polymorphism, SNP), after PCR amplification of the selected target sequence where the SNP is located, four duplexes are formed: two homoduplexes (allele A1/complement of allele A1 and allele A2/complement of allele A2) and two heteroduplexes (allele A1/complement of allele A2 and allele A2/complement of allele A1). The melting curve analysis exhibits the melting profile of each duplex and reveals the presence of heteroduplexes (derived from heterozygotes) and homoduplexes (derived from homozygotes) because they have different melting curve shapes (Mehta et al., 2017; Liew et al., 2004; Gundry et al., 2003). High-resolution melting analysis can reliably detect the differences between homozygous samples which exhibit the sharpest melting transitions and heterozygous samples in which a SNP is present which shows broader transitions because the heteroduplexes formed are heterogeneous and lower-melting. In addition, heterozygotes that have different variations such as SNPs will produce different melting curve shapes because different heteroduplexes are formed (Reed and Whittwer, 2004). Therefore, the presence or absence of specific alleles in a sample provide information about the cultivars present or absent from that sample and these data could be used to justify fraud.

HRM analysis is also combined with an analysis named DNA barcoding which is based on a short, standardized fragment of the genome called “DNA barcode” (Hebert et al., 2003). The DNA barcode could be a locus of the chloroplast, mitochondrial or nuclear genome and is used to identify organisms at the species level (Pereira et al., 2018). The data derived from the analysis of DNA barcodes of genomes combined with the analysis of genetic markers such as microsatellites or Simple Sequence Repeat (SSR) and SNPs markers can be used for verification, identification and classification of cultivars, germplasm and species and also in the authentication of food (Lo and Shaw, 2018). HRM has some main advantages such as diminishing post PCR manipulation, closed-tube analysis, cost effectiveness, and the relatively rapid obtaining of results (Pereira et al., 2018) which have contributed to its wide application in plant research and, in particular, in the food sector as shown in Table 3.

**Figure 4.** HRM derivative plot. This plot is used to visualize the melting temperatures $T_m$ of the samples analyzed more clearly ($T_m$ = the temperature at which 50% of the DNA is single stranded). The plot uses the negative first derivatives of fluorescence over temperature ($-\frac{dF}{dT}$). The $T_{m}$s of the products appear as peaks [adapted from Liew et al., 2004].
### TABLE 3. Applications of High Resolution Melting (HRM) in the authentication of olive oil

<table>
<thead>
<tr>
<th>Year of publication</th>
<th>Main Findings</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>2018</td>
<td>HRM melting curve analysis based on nuclear microsatellites (gSSR) and plastid DNA (cpDNA) generated 14 HRM profiles sufficient to genotype all oil varieties present in oil samples.</td>
<td>Gómez et al., 2018</td>
</tr>
<tr>
<td>2016</td>
<td>HRM curve profiles of DNA microsatellites of crude olive-pomace oil revealed the alteration of the declared varietal profile of a virgin olive oil.</td>
<td>Pasqualone et al., 2016b</td>
</tr>
<tr>
<td>2015</td>
<td>HRM with 17 SSR markers was used for discriminating the olive cultivars used in the “Terra di Bari” PDO extra virgin olive oil showing different melting temperatures; and also for quantitative detection of the different levels of adulteration (10–20–30–40–50%) between the monovarietal oil Coratina and non-PDO Aeleh.</td>
<td>Montemuro et al., 2015</td>
</tr>
<tr>
<td>2014</td>
<td>High-resolution melting (HRM) analysis, coupled with five microsatellite markers was used to characterize 47 <em>O. europaea</em> cultivars collected from the National Olive Tree Germplasm Collection established in Chania, Greece by generating a unique melting curve profile for each of the cultivars. Three microsatellite markers (DCA03, DCA09 and DCA17), which generated 29 HRM profiles, were sufficient to genotype all the olive cultivars studied.</td>
<td>Xanthopoulou et al., 2014</td>
</tr>
<tr>
<td>2013</td>
<td>HRM analysis was performed with PCR primers for the rbcL gene in order to detect the presence of maize and sunflower oils in artificial mixtures of olive oils and particularly in olive maize oil mix and olive-sunflower oil mix in ratios (v/v) 50/50, 70/30, 80/20 and 90/10. The HRM results showed that both maize oil and sunflower oil could be detected down to 10% limit of detection.</td>
<td>Vicetina et al., 2013</td>
</tr>
<tr>
<td>2013</td>
<td>The adulteration of olive (<em>Olea europaea</em>) oil with canola (<em>Brassica napus</em>) oil was analyzed with HRM and the results showed that the universal rbcL region is efficient enough to discriminate plant oil species and to detect the existence of 1% canola oil admixed into olive oil.</td>
<td>Ganopoulos et al., 2013</td>
</tr>
<tr>
<td>2009</td>
<td>HRM analysis was performed on PCR products of a fragment of phytochrome A gene (<em>phyA</em>) from six cultivars of <em>O. europaea</em> L. The presence of SNPs was identified in all cultivars tested and enabled the grouping of the cultivars according to their genotype.</td>
<td>Muleo et al., 2009</td>
</tr>
</tbody>
</table>

### 7. CONCLUSIONS

There is an increasing global demand for high quality olive oil due to its established beneficial properties. Extra virgin olive oil produced in limited quantities by certain varieties and in specific geographical areas has premium prices that reflect its excellent quality and for that reason it is subjected to a high degree of adulteration. It is imperative to establish reliable and effective methods to assess varietal identification, genetic authenticity, traceability and adulteration. The use of DNA based methods aims to develop discriminatory tools that could provide accurate authentication since residual DNA is present in olive oil and its analysis (genotyping) could provide detailed information about its genetic origin (variety, subspecies, species) which in some cases is more accurate than biochemical methods that measure metabolites, which are strongly influenced by environmental factors. To this end, a number of specific biomarkers such as SSRs and SNPs found in the genome of *Olea europaea* have been used in many recent studies to detect the identity of DNA extracted from olive oil. High resolution melting (HRM), a method widely used in biomedical studies, enables a rapid, high-throughput identification of genetic variants in olive oil. In the future, the application of next-generation sequencing in food genomics and the development of quick protocols for on-site authentication will contribute substantially to the detection of genetic authenticity of olive oil and other valuable food products.

### REFERENCES


