# Chemical partitioning and DNA fingerprinting of some pistachio (*Pistacia vera* L.) varieties of different geographical origin

**Giuseppe Mannino[a], Carla Gentile[b], Massimo E. Maffei[a]**

[a]Plant Physiology Unit, Dept. Life Sciences and Systems Biology. University of Turin, Via Quarello 15/a, 10135 Turin, Italy

[b]Department of Biological, Chemical and Pharmaceutical Sciences and Technologies, University of Palermo, Viale delle Scienze, 90128 Palermo, Italia

*Corresponding author: Massimo E. Maffei, Plant Physiology Unit, Dept. Life Sciences and Systems Biology. University of Turin, Via Quarello 15/a, 10135 Turin, Italy . E-mail: massimo.maffei@unito.it; Tel. +39 011 6705967; Fax: +39 011 2365967

## Abstract

The genus *Pistacia* (Anacardiaceae family) is represented by several species, of which only *P. vera* L. produces edible seeds (pistachio). Despite the different flavor and taste, a correct identification of pistachio varieties based on the sole phenotypic character is sometimes hard to achieve. Here we used a combination of chemical partitioning and molecular fingerprinting for the unequivocal identification of commercial pistachio seed varieties (Bronte, Kern, Kerman, Larnaka, Mateur and Mawardi) of different geographical origin. The total phenolic content was higher in the variety Bronte followed by Larnaka and Mawardi cultivars. The total anthocyanin content was higher in Bronte and Larnaka varieties, whereas the total proanthocyanidin content was higher in Bronte, followed by Mawardi and Larnaka varieties. HPLC-DAD-ESI-MS/MS analyses revealed significant ($P<0.05$) higher amounts of cyanidin-3-glucoside, idein, eryodictol-7-galactoside, quercetin-3-glucoside, luteolin-glucoside and marein in the variety Bronte, whereas higher amounts of peonidin-3-glucoside, okanin 4'-galactoside, hyperoside and quercetin-4'-glucoside were found in the variety Larnaka. The highest content of catechin was found in the Mawardi variety. A significantly ($P<0.05$) higher total amount of fatty acids was found in the varieties Mateur, Kern and Bronte, followed by the varieties Larnaka and Mawardi, whereas the variety Kerman showed the lowest total fatty acid content. GC-FID and GC-MS analyses revealed the presence of several polyunsaturated fatty acids. Kern and Mateur varieties showed a significantly ($P<0.05$) higher amount of linoleic acid, whereas the variety Bronte showed the highest amount of oleic acid. Molecular fingerprinting was achieved by ITS DNA PCR-RFLP analysis. Three different restriction enzymes (*RsaI*, *TaqαI* and *PstI*) were used to selectively cleave the resulting amplicons. A *TaqαI* site could be selectively found in the varieties Kerman, Larnaka and Mateur, whereas the digestion of the PCR products by *RsaI* gave specific patters exclusively on Bronte and Mawardi. Digestion by *PstI* gave specific patters exclusively on the Kern variety. The results showed that the Mediterranean varieties (Mateur, Bronte and Larnaka)

show similar chemical patterns and (particularly for Mateur and Larnaka) a close phylogenetic relationship, allowing a chemical and molecular partitioning with respect to the other varieties.

## Keywords

*Pistacia vera*; Anacardiaceae; anthocyanins; proanthocyanidins; flavonoids; fatty acids; Internal Transcribed Spacer (ITS).

## 1. Introduction

The genus *Pistacia* (Anacardiaceae) consists of at least 12 tree and shrub species, of which only *Pistacia vera* L. produces edible nuts (pistachio). Originating from the arid zones of Western Asia (especially Iran, Iraq, Syria and Turkey), *P. vera* cultivation has spread outside the traditional geographical regions (Khanazarov et al., 2009). In the Mediterranean area, local varieties were selected, including Bronte in Italy, Larnaka in Greece, and Mateur in Spain, which were commercialized all over the world. *P. vera* is also cultivated in the USA (California), because of the favourable climate, dry conditions and moderately cold winters (Benmoussa et al., 2017). The fruit is a drupe, containing an elongated seed, which is the edible portion. The seed has a mauve-coloured skin and light green flesh, with a distinctive flavour (Fabani et al., 2013). From a phytochemical point of view, several bioactive compounds have been identified in pistachio, including healthy lipids (Shahidi et al., 2007) and polyphenols (Fabani et al., 2013). *In vivo* studies showed a positive correlation between pistachio intake and reduced risk of cardiovascular disease (Gebauer et al., 2008; Tomaino et al., 2010). Moreover, pistachio consumption significantly improves oxidative stress of healthy individuals and lowers the levels of circulating inflammatory biomarkers, by ranking among the first 50 food products with the highest antioxidant potential (Sari et al., 2010). Some pistachio varieties contain substantial amounts of polyphenols that show radical-scavenging and anti-oxidative

properties and possess anti-inflammatory activities in *in vitro* models (Gentile et al., 2012; Gentile et al., 2015).

Pistachio seed kernels contain over 50% lipids, whereas polyphenols are mostly found in the seed skin, which is usually removed and treated as a waste (Aslan et al., 2002; Catalan et al., 2017). Despite the different flavour and taste, a correct identification of pistachio varieties simply based on the phenotypic parameters is not always possible. Therefore, the use of chemical and molecular profiling methods has been studied in pistachio, in order to help discrimination of varieties from different geographical origin. Chemical partitioning allowed pistachio geographical discrimination thought the identification of specific markers or entire metabolite profiling (Sobolev et al., 2017) using elemental analysis (Anderson and Smith, 2005), carbon and nitrogen isotope analyses (Anderson and Smith, 2006), heavy metals (Taghizadeh et al., 2017), phenolic profile (Saitta et al., 2014; Taghizadeh et al., 2018), essential oils (Dragull et al., 2010) and triacylglycerols (Ballistreri et al., 2010). Biomolecular characterization of pistachio also revealed to be a potent tool for variety discrimination through analysis of chloroplast DNA (Parfitt and Badenes, 1997; Sarra et al., 2015), RFLP analysis (Parfitt and Badenes, 1998), RAPD analysis (Hormaza et al., 1994), SSR-based genetic linkage map (Khodaeiaminjan et al., 2018) and retrotransposon markers (Kirdok and Ciftci, 2016). ITS is widely used in plant molecular systematics at the generic and species levels because of its potentially high resolution of inter- and intraspecific relationships (Cheng et al. 2016).

The aim of this study was to analyze the seed chemical composition and bimolecular profile of six pistachio commercial varieties (i.e., Bronte, Kerman, Kern, Larnaka, Mawardi and Mateur) from different geographical areas, rather than assessing the genetic variability among natural populations of *P. vera* cultivars. Chemical analyses included the characterization of phenolic compounds and fatty acids, whereas the DNA Restriction Fragment Length Polymorphism (PCR-RFLP) analysis was performed on the pistachio internal transcribed spacer (ITS). To our knowledge, there are no data on the ITS characterization and on the combined use of ITS and chemical data for the geographical partitioning of pistachio varieties. The combination of chemical and molecular data provided an

interesting integrated approach for the unequivocal identification of commercial pistachio seeds of different geographical origin.

## 2. Results and Discussion

### *2.1. The chemical partitioning of pistachio varieties from different geographical origin shows a significant differentiation in seed skin flavonoids and anthocyanins*

The total phenolic content (TPC) of the six pistachio varieties was quantified both in the seed flesh and skin. In general, significant differences were found among the six pistachio varieties. The skin TPC ranged between 91.37 (±01.04) and 363.75 (±16.50) mg $g^{-1}$ d.wt., whereas the TPC of seed flash was much lower (Table 1). The highest skin TPC was found in Bronte followed by Larnaka and Mawardi, whereas Kerman showed the lowest value. Similar TPC values have been reported for the Bronte (Martorana et al., 2013; Tsantili et al., 2011) and Kerman (Yang et al., 2009) varieties.

*P. vera* produces seeds containing anthocyanins (Bellomo and Fallico, 2007; Schulze-Kaysers et al., 2015), which are mainly stored in the seed skin (Tomaino et al., 2010). Table 1 shows the variability of the total anthocyanin content (TAC) of the six pistachio varieties under study. Significant differences were found in the skin among the six varieties, with the sole exception for Kern and Mawardi varieties. The highest TAC was found in Bronte and Larnaka varieties (Table 1). We found a positive correlation between TAC and TPC ($\rho = 0.86$), suggesting a possible contribution of TAC to the TPC. No anthocyanins were detected in the seed flash. Our results are consistent with previously reported data (Bellomo and Fallico, 2007; Liu et al., 2014; Seeram et al., 2006).

Proanthocyanidins (PACs) are the major polyphenolic compounds of some pistachio varieties (Gentile et al., 2015; Taghizadeh et al., 2018) and play a major role as bioactive component in *in vitro* inflammatory models (Gentile et al., 2012). High contents of total PACs (TPACs) were found in the skins of the pistachio varieties under study and were absent in the seed flesh (Table 1). The variety

121 Bronte showed the highest TPACs followed by Mawardi and Larnaka varieties, whereas the variety
122 Kern showed the lowest content. Intermediate values were shown by Kerman and Mateur varieties.
123 Supplementary Table S1 provides further information on statistical analyses.
124 Owing to the almost complete lack of phenolic compounds in the seed flesh in the six pistachio
125 varieties, we restricted their analysis to seed skins. In general, only small differences were detected
126 in the variety qualitative profile whereas a quantitative significant difference was found. In all
127 varieties, the most abundant compound was cyanidin-3-glucoside (**1**), followed by idein (**2**),
128 eriodictyol-7-glucoside (**3**), eriodictyol-7-galactoside (**4**) and catechin (**5**). Other common
129 compounds were peonidin-3-glucoside (**6**), hyperoside (**7**), quercetin-3-glucoside (**8**), quercetin-4'-
130 glucoside (**9**). Luteolin-glucoside (**10**) and marein (**11**) were absent in Mawardi and Larnaka varieties,
131 whereas okanin 4'-O-galactoside (**12**) was absent in the variety Kern (Table 2). Significant ($P<0.05$)
132 higher amounts of cyanidin-3-glucoside (**1**), idein (**2**), eryodictol-7-galactoside (**4**), quercetin-3-
133 glucoside (**8**), luteolin-glucoside (**10**) and marein (**11**) were found in the variety Bronte, in agreement
134 with literature data (Barreca et al., 2016; Martorana et al., 2013; Tomaino et al., 2010). The variety
135 Larnaka showed significantly ($P<0.05$) higher amounts of eryodictol-7-galactoside (**4**), peonidin-3-
136 glucoside (**6**), okanin 4'-galactoside (**12**), hyperoside (**7**) and quercetin-4'-glucoside (**9**). The highest
137 content of catechin (**5**) was found in the Mawardi variety (Table 2). A similar polyphenolic profile
138 has been previously reported in pistachio extracts (Erşan et al., 2017, 2018; Fabani et al., 2013; Goli
139 et al., 2005; Grace et al., 2016; Lalegani et al., 2018; Rodriguez-Bencomo et al., 2015; Sonmezdag et
140 al., 2018) and fruit skin (Tas and Gokmen, 2017). Figure 1 shows the chemical formulae of the
141 identified phenolic compounds.
142 The Principal Component Analysis (PCA) calculated on the data matrix of Tables 1 and 2 with
143 varimax rotation explained 57.59% and 20.57% of the total variance for PC1 and PC2, respectively.
144 Positive factor scores for PC1 discriminated the Mediterranean varieties Larnaka and Bronte because
145 of high TPC and the highest content of cyanidin-3-glucoside (**1**) and idein (**2**). Negative PC1 factors
146 scores separated all other varieties (Fig. 2). The varieties Kerman and Mawardi were separated by

147 both PC1 and PC2 negative factor scores because of the low content of luteolin-glucoside (**10**),
148 whereas the Bronte variety was separated by both PC1 and PC2 positive factor scores because of the
149 highest TAC, TPC and TPACs values. Supplementary Figure S1 shows the partitioning of the
150 different phenolic compounds based on PC1 and PC2 factor scores.
151
152 *2.2 Linoleic acid and oleic acid contribute to the chemical partitioning of pistachio seed flashes.*
153 The pistachio fatty acid composition has been used for the differentiation of varieties of different
154 geographical origin (Acar et al., 2008; Arena et al., 2007; Aslan et al., 2002; Chahed et al., 2008;
155 Rabadan et al., 2018; Rabadan et al., 2017), providing useful criteria for origin authentication of
156 pistachio seeds. As expected, the fatty acid content of the six pistachio variety was mainly present tin
157 the seed flash. In general, a significantly ($P<0.05$) higher total amount of the identified fatty acids
158 was found in the varieties Mateur, Kern and Bronte, followed by the varieties Larnaka and Mawardi,
159 whereas the variety Kerman showed the lowest total fatty acid amount (Table 3). The two main
160 identified fatty acids were linoleic acid (**13**) and oleic acid (**14**), in accordance with the literature data
161 (Catalan et al., 2017; Dreher, 2012; Pantano et al., 2016). With respect to the other varieties, Kern
162 and Mateur showed a significantly ($P<0.05$) higher amounts of linoleic acid (**13**), whereas the variety
163 Bronte showed the highest amount of oleic acid (**14**). Other minor fatty acids included mono and
164 polyunsaturated fatty acids (Table 3). Our results are in agreement with previously reported data
165 (Grace et al., 2016; Ling et al., 2016; Ojeda-Amador et al., 2018; Pantano et al., 2016; Rodriguez-
166 Bencomo et al., 2015).
167 The Principal Component Analysis (PCA) calculated on the data matrix of Table 3 with varimax
168 rotation explained 40.95% and 32.60% of the total variance for PC1 and PC2, respectively (Fig. 3).
169 Positive factor scores discriminated the Mediterranean varieties Larnaka and Bronte because of the
170 higher content of oleic acid (**14**), whereas negative factors scores separated the Californian variety
171 Kerman because of the lowest total fatty acid content. The Mawardi variety was separated by positive
172 PC1 and Negative PC2 factor scores because of the lowest content of linoleic acid (**13**) whereas Kern

and Mateur varieties were separated by positive PC2 and negative PC1 factor scores because of similar fatty acid contents. Supplementary Figure S2 shows the partitioning of the different fatty acids based on PC1 and PC2 factor scores.

The PCA calculated on the overall data of Tables 1-3 with varimax rotation explained 41.80% and 28.35% of the total variance for PC1 and PC2, respectively (Fig. 4). The combination of phenolic compounds and fatty acids confirms the separation of the Mediterranean varieties Mateur, Bronte and Larnaka by positive factor scores of the PC1 and better separates the varieties Kern, Kerman and Mawardi by negative factor scores of PC1 (Fig. 3). Supplementary Figure S3 shows the distribution of the different chemical compounds on the two main PCs of the PCA.

*2.3. DNA fingerprinting using PCR-RFLP analysis reveals significant differences in pistachio varieties of different geographical origin*

In order to provide a molecular fingerprinting of the six pistachio varieties, ITS-1 coupled with ITS-4 was used for PCR amplification. Supplementary Figure S4 shows the nucleotide sequence of the ITS regions of the six varieties.

The ITS amplified sequences were 722bp long (Fig. 5 lanes 1-6) (**NCBI GenBank Accession Nos: MH444649, ITS1-4 Bronte; MH444689, ITS1-4 Kerman; MH444724, ITS1-4 Kern; MH444735, ITS1-4 Larnaka; MH444780, ITS1-4 Mateur; MH444793, ITS1-4 Mawardi**) and the alignment of the six verities sequences shows that 98.75% of the sites are conserved. In particular, out of the 1.25% variable sites, 0.83% provide little information and 0.42% are singleton sites. The ITS fragments were compared by BLAST alignment to other sequences deposited in GeneBank, and the analysis provided a match almost identical to *P. vera* (Sequence ID: AY677201.1) with a 99% query score.

In order to better characterize the varieties showing DNA fragments of similar size, a PCR–RFLP method was applied. Three different restriction enzymes (*RsaI*, *TaqαI* and *PstI*) were used to selectively cleave the resulting amplicons. From the identified sequences, a *TaqαI* site could be

199 selectively found in the varieties Kerman (Fig. 5 lane 7), Larnaka (Fig. 4 lane 8) and Mateur (Fig. 5
200 lane 9), giving five fragments of 76, 86, 90, 185 and 280 bp. Digestion of the PCR products by *RsaI*
201 gave specific patters exclusively on Bronte (Fig. 5 lane 10) and Mawardi (Fig. 5 lane 11) variety
202 sequences, by producing two fragments of 182 and 550 bp. Finally, PCR products from the different
203 varieties were digested by *PstI*, which produced two fragments of 92 and 630 bps exclusively on the
204 Kern variety (Fig. 5 lane 12). These results show that it is possible to differentiate among the six
205 species investigated, not exclusively by chemical characterization, but also by fingerprinting analysis.
206 Supplementary Table S2 provides the sequence of each ITS fragments generated after RFLP analysis
207 with *RsaI*, *TaqαI* and *PstI* restriction enzymes.
208 The sequences were further analyzed by the neighbour joining (NJ) method to infer phylogenetic
209 relationship among the pistachio varieties. Figure 6 shows the phylogenetic tree where the Mawardi
210 and Bronte varieties and Mateur and Larnaka form independent clusters, which robustness is
211 supported by high bootstrap scores. Our data are in agreement with DNA-RAPD markers on *P. vera*
212 phylogenetics (Hormaza et al., 1994).
213

214 ## 3. Conclusions

215 The combination of DNA analysis and phytochemical analyses is increasingly used to provide new
216 tools for the unequivocal identification of plants. The stability of DNA fingerprinting is a solid
217 method that supports the chemical partitioning. Despite some controversy exists over the value of
218 DNA barcoding, largely because of the perception that this method would diminish rather than
219 enhance traditional morphology-based taxonomy, an increasing number of gene sequences is now
220 available for DNA barcoding of flowering plants (Cheng et al., 2016).
221 In this work we showed that different varieties of pistachio, a plant with a high food value and
222 phytochemical potential, show a remarkable variability, both at the genomic and gene products
223 (phenolic compounds and fatty acids) levels. By using both molecular and chemical data it is possible

to partition the different pistachio varieties according to their geographical origin. In particular, the Mediterranean varieties (Mateur, Bornate and Larnaka) show similar chemical patterns and (in the case of Mateur and Larnaka) a close phylogenetic relationship.

Owing to the increased interest and relevance of *P. vera* as a food plant and as a source of interesting phytochemicals with pharmaceutical properties, the identification of bioactive phenolic compounds and specific gene sequences by PCR-RFLP described in this work offers a valuable tool for a rapid and unequivocal identification of pistachio varieties of different geographical origin.

## 4. Experimental

### *4.1. Plant material*

Seeds of different varieties of *Pistacia vera* L. (Bronte from Sicily, Mawardi from Turkey, Larnaka from Greece, Kern from Iran, Kerman from U.S.A., California and Mateur from Spain) were kindly provided by Pistacchio dell'Etna Srl (Bronte, Italy) and by Di Sano Srl (Rozzano, Italy). Seeds were stored in the dark at 4°C before extraction. At least three technical replicates were done for each lot of seeds.

### *4.2. Extraction of phenolic compounds*

The seed skin and flash of each variety was manually separated and extracted in 75:25 v/v ethanol:water solution, for 3 days in the dark at room temperature, using a 1:20 w/v extraction ratio. After centrifugation (10 min at 10,000 g, 4°C) and filtration through a Millex HV 0.45 μm filter (Millipore, Billerica, MA), the supernatants were recovered and stored at -80°C until analysis. For each variety, the extraction was performed in triplicate. Lipophilic extracts of seed flash were obtained by Soxhlet extraction by using cyclohexane (1:10, w/v). After extraction, the solvent was removed with a nitrogen flow.

249 *4.3. Total phenolic compounds content*

250 The total phenolic compounds content (TPC) was determined by the Folin-Ciocalteu's method
251 (Singleton et al., 1999). Gallic acid (GA) was used for the preparation of the calibration curve (see
252 Supplementary Table S3) and the results were expressed as mg GA $g^{-1}$ d.wt. All measurements were
253 repeated three times.

254

255 *4.4. Total anthocyanin content*

256 The total anthocyanin content (TAC) was measured using the differential pH method (Elisia et al.,
257 2007). Cyanidin chloride (CC) was used as standard and the total anthocyanin content was expressed
258 as mg CC $g^{-1}$ d.wt. (see Supplementary Table S3). All measurements were performed in triplicate.

259

260 *4.5. Total proanthocyanidin (PAC) content*

261 The 4-(dimethylamino)-cinnamaldehyde (DMAC) assay was used to evaluate the total amount of
262 PACs according to Prior et al. (2010) with minor modifications (Occhipinti et al., 2016). The total
263 PAC content was quantified via an external calibration curve made with a pure PAC-A2 standard and
264 was expressed as mg PAC-A2 $g^{-1}$ d.wt.. The measurements were performed in triplicate.

265

266 *4.6. HPLC-DAD-ESI-MS/MS analysis of phenolic compounds*

267 The HPLC system consisted of an Agilent Technologies 1200 coupled to a DAD and a 6330 Series
268 Ion Trap LC-MS System (Agilent Technologies, USA) equipped with an electrospray ionization
269 (ESI) source. The chromatographic separation was carried out at constant flow rate (0.2 ml $min^{-1}$).
270 The column was a reverse phase C18 Luna column (3.00 µm, 150 × 3.0 mm i.d., Phenomenex, USA).
271 maintained at 25°C by an Agilent 1100 HPLC G1316A Column Compartment. The UV–VIS spectra
272 were recorded between 220 and 650 nm and the chromatographic profiles were registered at 220, 280,

360 and 520 nm. Tandem mass spectrometry analyses were performed operating either in negative mode (for flavonoids) or in positive mode (for anthocyanins). The nitrogen flow rate was set at 5.0 ml $min^{-1}$ and maintained at 325°C, whereas the capillary voltage was set at 1.5 kV. Helium was used as a collision gas. Compound identification was carried out by comparison of the retention time and UV-VIS/MS spectra with those of authentic reference compounds or using literature data.

*4.6.1 Flavonoid analysis.* The binary solvent system for flavonoid analysis was MilliQ $H_2O$ acidified with 0.1% v/v (Solvent A) (Millipore, Billerica, MA, USA) and ACN acidified with 0.1% v/v formic acid (Solvent B). Samples were separated by the following gradient: 97% A and 3% B as initial conditions, 70% A and 30% B for 35 min, and then 2% A and 98% B for 5 min. The concentration of A was maintained at 2% for 5 min and eventually was raised to the initial condition before the next injection. Sample injection volume was 5 µl.

*4.6.2 Anthocyanin analysis*. The binary solvent system for anthocyanin analysis was MilliQ $H_2O$ acidified with 0.1% (v/v) formic acid (Solvent A) and MetOH 50% v/v acidified with 10% v/v formic acid (Sigma-Aldrich, USA) (Solvent B). The elution method involved a multistep linear solvent gradient changing from an initial concentration of 85% A and 15% B to 55% A and 45% B in 15 min. Finally, the gradient was 30% A and 70% B in 20 min. The concentration of solvent A was decreased to 2% and was maintained for 5 min before the next injection. Sample injection volume was 15 µl.

*4.7. Fatty acid analysis*

The Soxhlet extract was esterified with boron tri-fluoride (10% w/v in methanol). Fifty µg heptadecanoic acid (C17:0) were added as internal standard (Maffei and Peracino, 1993). Fatty acid methyl esters (FAME) were obtained by acid catalysis according to Christie and Han (2010) and were dehydrated with anhydrous $MgSO_4$. FAME identification and quantification was performed by GC-MS (5975T, AgilentTechnologies, USA) and by GC-FID (GC-2010 Plus, SHIMADZU, Japan), respectively. The GC carrier gas was helium with a constant flux of 1 ml $min^{-1}$, and separation was obtained with a non-polar capillary column ZB5-MS (30 m length, 250 µm diameter and stationary

299 phase thickness of 0,25 μm, 5% phenyl-arylene and 95% poly-dimethyl siloxane) (Phenomenex,
300 USA). The following temperature conditions was used: injector 250°C, oven initially at 60 °C, held
301 for 1 minute and raised to 180°C (10.0°C $min^{-1}$ and held for 1 minute). Then the temperature was
302 brought to 230 °C (1.0 °C $min^{-1}$ and held for 2 minutes) and to 320 °C (15 °C $min^{-1}$) held for 5
303 minutes. Same column and chromatographic condition were used for both GC-MS and GC-FID
304 analyses. MS parameters were: ionization energy of the ion source was set to 70 eV and the
305 acquisition mode was set to 50–350 m/z. Compounds were identified through comparison of mass
306 fragmentation spectra with reference NIST 98 spectra or by comparison of Kovats indexes and
307 internal standard co-injection of pure standards (Sigma-Aldrich, USA). FAME quantification was
308 obtained by internal standard. At least three technical replicates were run for each lot of pistachio
309 cultivars.
310
311 *4.8. DNA fingerprinting*
312 *4.8.1. DNA extraction, PCR amplification, subcloning and sequencing.* Whole pistachio seeds were
313 pulverized in liquid nitrogen using a mortar and pestle. Genomic DNA was extracted and quantified
314 according to Capuzzo and Maffei (2014). Briefly, twenty ng of genomic DNA were used as a template
315 for PCR amplification with specific primers for ITS1 (5'-TCCGTAGGTGAACCTGCGG-3') and
316 ITS4 (3'-CCGCAGGTTCACCTACGGA-5'). PCR products were separated by 1.0% (w/v) agarose
317 gel electrophoresis and visualized by GelRed (Biotium) staining under UV, and purified from the gel
318 using the NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel). The purified product was used
319 for subcloning using the TOPO-TA Cloning Kit (Thermo Fisher Scientific) and then transformed in
320 *Escherichia coli* Subcloning DH5α Efficiency Competent Cells (Invitrogen, Paisley, UK). Colonies
321 containing DNA inserts of the correct size were picked and grown overnight in 5 ml Luria-Bertani
322 liquid medium. The mini-preparation of plasmid DNAs was carried out using NucleoSpin Plasmid
323 Miniprep Kit (Macherey-Nage). Plasmid DNAs were used as a template for sequencing (Macrogen,
324 Wageningen, Holland). Both DNA strands were sequenced.

325 *4.8.2 PCR-RFLP.* PCR products of the ITS gene were digested at 37°C for 15 min with either 10 U
326 *RsaI*, *PstI* (NEB, New England Biolabs, Ipswich, AM, USA) or *TaqI* (NEB, New England Biolabs,
327 Ipswich, AM, USA) at 65°C for 60 min. One microliter of each digestion reaction was analyzed by
328 capillary gel electrophoresis (CGE) using the Agilent 2100 Bioanalyzer (Agilent Technologies) and
329 the DNA 1000 LabChip Kit (Agilent Technologies) following the manufacturer's instructions.

330

331 **4.10. Statistical analyses**

332 Statistical analyses were performed in order to assess the errors related to the analytical procedures,
333 rather than assessing the internal variability among the different cultivars. Data are expressed as the
334 mean of three technical replicates for each lot of seeds. ANOVA followed by Tukey–Kramer's HSD
335 post-hoc test ($P < 0.05$) was used to determine significant differences. Principal Component Analysis
336 (PCA) was performed by using covariant matrix of extraction and varimax rotation. All statistical
337 analyses were performed by using the SYSTAT 10 software. The cladogram of gene sequences was
338 performed with ClustalX software by using the Neighbour Joining (NJ) method. Bootstrap values
339 were calculated from 100 resamplings of the alignment data.

340

341 ## ACKNOWLEDGMENTS

342 This work was partly supported by a grant from the PhD School of Pharmaceutical and Biomolecular
343 Sciences of the University of Turin. The authors wish to thank A. Occhipinti for technical support
344 and Pistacchio dell'Etna S.r.l. and Di Sano S.r.l. for kindly providing samples of pistachio seed lots.

345

346 ## REFERENCES

347

348 Acar, I., Kafkas, E., Ozogul, Y., Dogan, Y., Kafkas, S., 2008. Variation of fat and fatty acid
349 composition of some pistachio genotypes. Ital. J. Food Sci. 20, 273-279.

350 Anderson, K. A., Smith, B. W., 2005. Use of chemical profiling to differentiate geographic growing
351 origin of raw Pistachios. Journal of Agricultural and Food Chemistry 53, 410-418.
352 Anderson, K. A., Smith, B. W., 2006. Effect of season and variety on the differentiation of geographic
353 growing origin of pistachios by stable isotope profiling. Journal of Agricultural and Food
354 Chemistry 54, 1747-1752.
355 Arena, E., Campisi, S., Fallico, B., Maccarone, E., 2007. Distribution of fatty acids and phytosterols
356 as a criterion to discriminate geographic origin of pistachio seeds. Food Chemistry 104, 403-
357 408.
358 Aslan, M., Orhan, I., Sener, B., 2002. Comparison of the seed oils of Pistacia vera L. of different
359 origins with respect to fatty acids. International Journal of Food Science and Technology 37,
360 333-335.
361 Ballistreri, G., Arena, E., Fallico, B., 2010. Characterization of triacylglycerols in *Pistacia vera* L.
362 oils from different geographic origins. Ital. J. Food Sci. 22, 69-75.
363 Barreca, D., Lagana, G., Leuzzi, U., Smeriglio, A., Trombetta, D., Bellocco, E., 2016. Evaluation of
364 the nutraceutical, antioxidant and cytoprotective properties of ripe pistachio (Pistacia vera L.,
365 variety Bronte) hulls. Food Chemistry 196, 493-502.
366 Bellomo, M. G., Fallico, B., 2007. Anthocyanins, chlorophylls and xanthophylls in pistachio nuts
367 (Pistacia vera) of different geographic origin. J. Food Compos. Anal. 20, 352-359.
368 Benmoussa, H., Luedeling, E., Ghrab, M., Ben Yahmed, J., Ben Mimoun, M., 2017. Performance of
369 pistachio (Pistacia vera L.) in warming Mediterranean orchards. Environmental and
370 Experimental Botany 140, 76-85.
371 Capuzzo, A., Maffei, M. E., 2014. Molecular fingerprinting of some *Mentha* species by sequencing
372 and RFLP analysis of the 5S-rRNA non-transcribed spacer region. Plant Biosyst DOI:
373 10.1080/11263504.2013.790853.

374 Catalan, L., Alvarez-Orti, M., Pardo-Gimenez, A., Gomez, R., Rabadan, A., Pardo, J. E., 2017.
375 Pistachio oil: A review on its chemical composition, extraction systems, and uses. European
376 Journal of Lipid Science and Technology 119, 8.
377 Chahed, T., Bellila, A., Dhifi, W., Hamrouni, I., M'Hamdi, B., Kchouk, M. E., Marzouk, B., 2008.
378 Pistachio (Pistacia vera) seed oil composition: geographic situation and variety effects. Grasas
379 Y Aceites 59, 51-56.
380 Cheng, T., Xu, C., Lei, L., Li, C. H., Zhang, Y., Zhou, S. L., 2016. Barcoding the kingdom Plantae:
381 new PCR primers for ITS regions of plants with improved universality and specificity.
382 Molecular Ecology Resources 16, 138-149.
383 Christie, W. W., Han, X., 2010. Lipid Analysis, 4th Edition. Woodhead Publishing, Oxford.
384 Dragull, K., Beck, J. J., Merrill, G. B., 2010. Essential oil yield and composition of Pistacia vera
385 'Kerman' fruits, peduncles and leaves grown in California. Journal of the Science of Food and
386 Agriculture 90, 664-668.
387 Dreher, M. L., 2012. Pistachio nuts: composition and potential health benefits. Nutr. Rev. 70, 234-
388 240.
389 Elisia, I., Hu, C., Popovich, D. G., Kitts, D. D., 2007. Antioxidant assessment of an anthocyanin-
390 enriched blackberry extract. Food Chemistry 101, 1052-1058.
391 Erşan, S., Ustundag, O. G., Carle, R., Schweiggert, R. M., 2017. Determination of pistachio (Pistacia
392 vera L.) hull (exo- and mesocarp) phenolics by HPLC-DAD-ESI/MSn and UHPLC-DAD-
393 ELSD after ultrasound-assisted extraction. J. Food Compos. Anal. 62, 103-114.
394 Erşan, S., Ustundag, O. G., Carle, R., Schweiggert, R. M., 2018. Subcritical water extraction of
395 phenolic and antioxidant constituents from pistachio (Pistacia vera L.) hulls. Food Chemistry
396 253, 46-54.
397 Fabani, M. P., Luna, L., Baroni, M. V., Monferran, M. V., Ighani, M., Tapia, A., Wunderlin, D. A.,
398 Feresin, G. E., 2013. Pistachio (Pistacia vera var Kerman) from Argentinean cultivars. A

399 natural product with potential to improve human health. Journal of Functional Foods 5, 1347-
400 1356.
401 Gebauer, S. K., West, S. G., Kay, C. D., Alaupovic, P., Bagshaw, D., Kris-Etherton, P. M., 2008.
402 Effects of pistachios on cardiovascular disease risk factors and potential mechanisms of
403 action: a dose-response study. American Journal of Clinical Nutrition 88, 651-659.
404 Gentile, C., Allegra, M., Angileri, F., Pintaudi, A. M., Livrea, M. A., Tesoriere, L., 2012. Polymeric
405 proanthocyanidins from Sicilian pistachio (Pistacia vera L.) nut extract inhibit
406 lipopolysaccharide-induced inflammatory response in RAW 264.7 cells. European Journal of
407 Nutrition 51, 353-363.
408 Gentile, C., Perrone, A., Attanzio, A., Tesoriere, L., Livrea, M. A., 2015. Sicilian pistachio (Pistacia
409 vera L.) nut inhibits expression and release of inflammatory mediators and reverts the increase
410 of paracellular permeability in IL-1 beta-exposed human intestinal epithelial cells. European
411 Journal of Nutrition 54, 811-821.
412 Goli, A. H., Barzegar, M., Sahari, M. A., 2005. Antioxidant activity and total phenolic compounds of
413 pistachio (Pistachia vera) hull extracts. Food Chemistry 92, 521-525.
414 Grace, M. H., Esposito, D., Timmers, M. A., Xiong, J., Yousef, G., Komarnytsky, S., Lila, M. A.,
415 2016. Chemical composition, antioxidant and anti-inflammatory properties of pistachio hull
416 extracts. Food Chemistry 210, 85-95.
417 Hormaza, J. I., Dollo, L., Polito, V. S., 1994. Determination of relatedness and geographical
418 movements of *Pistacia vera* (Pistachio, Anacardiaceae) germplasm by RAPD analysis. Econ
419 Bot 48, 349-358.
420 Khanazarov, A. A., Chernova, G. M., Rakhmonov, A. M., Nikolyi, L. V., Ablaeva, E., Zaurov, D. E.,
421 Molnar, T. J., Eisenman, S. W., Funk, C. R., 2009. Genetic resources of Pistacia vera L. in
422 Central Asia. Genet. Resour. Crop Evol. 56, 429-443.

423 Khodaeiaminjan, M., Kafkas, S., Motalebipour, E. Z., Coban, N., 2018. In silico polymorphic novel
424 SSR marker development and the first SSR-based genetic linkage map in pistachio. Tree
425 Genetics & Genomes 14.
426 Kirdok, E., Ciftci, Y. O., 2016. Retrotransposon Marker Systems as a Tool to Analyze Molecular
427 Diversity of Mediterranean Pistacia Species. Int. J. Agric. Biol. 18, 601-606.
428 Lalegani, S., Gavlighi, H. A., Azizi, M. H., Sarteshnizi, R. A., 2018. Inhibitory activity of phenolic-
429 rich pistachio green hull extract-enriched pasta on key type 2 diabetes relevant enzymes and
430 glycemic index. Food Res. Int. 105, 94-101.
431 Ling, B., Yang, X. M., Li, R., Wang, S. J., 2016. Physicochemical properties, volatile compounds,
432 and oxidative stability of cold pressed kernel oils from raw and roasted pistachio (Pistacia
433 vera L. Var Kerman). European Journal of Lipid Science and Technology 118, 1368-1379.
434 Liu, Y. T., Blumberg, J. B., Chen, C. Y. O., 2014. Quantification and Bioaccessibility of California
435 Pistachio Bioactives. Journal of Agricultural and Food Chemistry 62, 1550-1556.
436 Maffei, M., Peracino, V., 1993. Fatty-Acids from Some *Lavandula* Hybrids Growing Spontaneously
437 in North-West Italy. Phytochemistry 33, 373-376.
438 Martorana, M., Arcoraci, T., Rizza, L., Cristani, M., Bonina, F. P., Saija, A., Trombetta, D., Tomaino,
439 A., 2013. In vitro antioxidant and in vivo photoprotective effect of pistachio (Pistacia vera L.,
440 variety Bronte) seed and skin extracts. Fitoterapia 85, 41-48.
441 Occhipinti, A., Germano, A., Maffei, M. E., 2016. Prevention of Urinary Tract Infection with
442 Oximacro®, a cranberry extract with a high content of A-type Proanthocyanidins (PAC-A).
443 A pre-clinical double-blind controlled study. Urol J In press.
444 Ojeda-Amador, R. M., Fregapane, G., Salvador, M. D., 2018. Composition and properties of virgin
445 pistachio oils and their by-products from different cultivars. Food Chemistry 240, 123-130.
446 Pantano, L., Lo Cascio, G., Alongi, A., Cammilleri, G., Vella, A., Macaluso, A., Cicero, N.,
447 Migliazzo, A., Ferrantelli, V., 2016. Fatty acids determination in Bronte pistachios by gas
448 chromatographic method. Natural Product Research 30, 2378-2382.

449 Parfitt, D. E., Badenes, M. L., 1997. Phylogeny of the genus Pistacia as determined from analysis of
450 the chloroplast genome. Proceedings of the National Academy of Sciences of the United
451 States of America 94, 7987-7992.
452 Parfitt, D. E., Badenes, M. L., 1998. Molecular phylogenetic analysis of the genus Pistacia. In:
453 Ferguson, L., Kester, D. (Eds.), Second International Symposium on Pistachios and Almonds.
454 International Society Horticultural Science, Leuven 1, pp. 143-151.
455 Prior, R. L., Fan, E., Ji, H. P., Howell, A., Nio, C., Payne, M. J., Reed, J., 2010. Multi-laboratory
456 validation of a standard method for quantifying proanthocyanidins in cranberry powders. J.
457 Sci. Food Agric 90, 1473-1478.
458 Rabadan, A., Alvarez-Orti, M., Gomez, R., de Miguel, C., Pardo, J. E., 2018. Influence of genotype
459 and crop year in the chemometrics of almond and pistachio oils. Journal of the Science of
460 Food and Agriculture 98, 2402-2410.
461 Rabadan, A., Pardo, J. E., Gomez, R., Alvarruiz, A., Alvarez-Orti, M., 2017. Usefulness of physical
462 parameters for pistachio cultivar differentiation. Scientia Horticulturae 222, 7-11.
463 Rodriguez-Bencomo, J. J., Kelebek, H., Sonmezdag, A. S., Rodriguez-Alcala, L. M., Fontecha, J.,
464 Selli, S., 2015. Characterization of the Aroma-Active, Phenolic, and Lipid Profiles of the
465 Pistachio (Pistacia vera L.) Nut as Affected by the Single and Double Roasting Process.
466 Journal of Agricultural and Food Chemistry 63, 7830-7839.
467 Saitta, M., La Torre, G. L., Potorti, A. G., Di Bella, G., Dugo, G., 2014. Polyphenols of Pistachio
468 (Pistacia vera L.) Oil Samples and Geographical Differentiation by Principal Component
469 Analysis. Journal of the American Oil Chemists Society 91, 1595-1603.
470 Sari, I., Baltaci, Y., Bagci, C., Davutoglu, V., Erel, O., Celik, H., Ozer, O., Aksoy, N., Aksoy, M.,
471 2010. Effect of pistachio diet on lipid parameters, endothelial function, inflammation, and
472 oxidative status: A prospective study. Nutrition 26, 399-404.

473 Sarra, C., Soumaya, R. C., Zined, M., Khaled, S., Noureddine, C., Khaled, C., 2015. Chloroplast
474 DNA analysis of Tunisian pistachio (Pistacia vera L.): Sequence variations of the intron trnL
475 (UAA). Scientia Horticulturae 191, 57-64.
476 Schulze-Kaysers, N., Feuereisen, M. M., Schieber, A., 2015. Phenolic compounds in edible species
477 of the Anacardiaceae family - a review. Rsc Adv 5, 73301-73314.
478 Seeram, N. P., Zhang, Y. J., Henning, S. M., Lee, R., Niu, Y. T., Lin, G., Heber, D., 2006. Pistachio
479 skin phenolics are destroyed by bleaching resulting in reduced antioxidative capacities.
480 Journal of Agricultural and Food Chemistry 54, 7036-7040.
481 Shahidi, F., Alasalvar, C., Liyana-Pathirana, C. M., 2007. Antioxidant phytochemicals in hazelnut
482 kernel (Corylus avellana L.) and hazelnut byproducts (vol 55, pg 1212, 2007). Journal of
483 Agricultural and Food Chemistry 55, 3232-3232.
484 Singleton, V. L., Orthofer, R., Lamuela-Raventos, R. M., 1999. Analysis of total phenols and other
485 oxidation substrates and antioxidants by means of Folin-Ciocalteu reagent. Method Enzymol
486 299, 152-178.
487 Sobolev, A. P., Circi, S., Capitani, D., Ingallina, C., Mannina, L., 2017. Molecular fingerprinting of
488 food authenticity. Curr. Opin. Food Sci. 16, 59-66.
489 Sonmezdag, A. S., Kelebek, H., Selli, S., 2018. Pistachio oil (Pistacia vera L. cv. Uzun):
490 Characterization of key odorants in a representative aromatic extract by GC-MS-olfactometry
491 and phenolic profile by LC-ESI-MS/MS. Food Chemistry 240, 24-31.
492 Taghizadeh, S. F., Davarynejad, G., Asili, J., Nemati, S. H., Karimi, G., 2018. Assessment of phenolic
493 profile and antioxidant power of five pistachio (Pistacia vera) cultivars collected from four
494 geographical regions of Iran. Avicenna J. Phytomedicine 8, 33-42.
495 Taghizadeh, S. F., Davarynejad, G., Asili, J., Nemati, S. H., Rezaee, R., Goumenou, M., Tsatsakis,
496 A. M., Karimi, G., 2017. Health risk assessment of heavy metals via dietary intake of five
497 pistachio (Pistacia vera L.) cultivars collected from different geographical sites of Iran. Food
498 and Chemical Toxicology 107, 99-107.

Tas, N. G., Gokmen, V., 2017. Phenolic compounds in natural and roasted nuts and their skins: a brief review. Curr. Opin. Food Sci. 14, 103-109.

Tomaino, A., Martorana, M., Arcoraci, T., Monteleone, D., Giovinazzo, C., Saija, A., 2010. Antioxidant activity and phenolic profile of pistachio (Pistacia vera L., variety Bronte) seeds and skins. Biochimie 92, 1115-1122.

Tsantili, E., Konstantinidis, K., Christopoulos, M. V., Roussos, P. A., 2011. Total phenolics and flavonoids and total antioxidant capacity in pistachio (Pistachia vera L.) nuts in relation to cultivars and storage conditions. Scientia Horticulturae 129, 694-701.

Yang, J., Liu, R. H., Halim, L., 2009. Antioxidant and antiproliferative activities of common edible nut seeds. Lwt-Food Science and Technology 42, 1-8.

## Figures legend

**Figure 1.** Structure formulae of the phenolic compounds and fatty acids characterizing the pistachio varieties under study.

**Figure 2.** Scatter plot of the principal components (PC1 and PC2) factor scores calculated on the PCA of phenolic compounds of the pistachio varieties of different geographical origin using the data matrix of Tables 1 and 2. A clear separation is obtained for the Mediterranean varieties Bronte and Larnaka, the Californian variety Kerman and the other varieties. See also Supplementary Figure S1 for the chemical partitioning of compounds.

**Figure 3.** Scatter plot of the principal components (PC1 and PC2) factor scores calculated on the PCA of fatty acids of the pistachio varieties of different geographical origin obtained from the data matrix of Table 3. A clear separation is obtained for the Mediterranean varieties Bronte and Larnaka, the Californian variety Kerman, the Turkish variety Mawardi and the other varieties. See also Supplementary Figure S2 for the chemical partitioning of compounds.

**Figure 4.** Scatter plot of the principal components (PC1 and PC2) factor scores calculated on the PCA of phenolic compounds and fatty acids of the pistachio varieties of different geographical origin using the data of Tables 1-3. A clear separation is obtained for the Mediterranean varieties Mateur, Bronte and Larnaka, the Californian variety Kerman and the other varieties. See also Supplementary Figure S3 for the chemical partitioning of compounds.

**Figure 5**. PCR products after capillary gel electrophoresis analysis of the ITS region of some *Pistacia vera* varieties of different geographical origin. Whole ITS sequence of Bronte (lane 1), Kerman (lane 2), Larnaka (lane 3), Kern (lane 4), Mateur (lane 5) and Mawardi (lane 6) varieties. All sequences have a length of about 720 bp. PCR–RFLP analysis using *TaqαI* pistachio digested PCR products produces five fragments of 75, 85, 90, 185 and 280 bp in Kerman (lane 7), Larnaka (lane 8) and Mateur (lane 9) varieties. Digestion of the PCR products from *RsaI* restriction enzyme activity on Bronte (lane 10) and Mawardi (lane 11) gives two fragments of 180 and 550 bp. Digestion of the

PCR products from *PstI* produces two fragments of 90 and 630 bp on the Kern (lane 12) variety. L = bp markers. The PCR products were separated by using the Agilent 2100 Bioanalyzer and the DNA 1000 LabChip Kit (Agilent Technologies). See Supplementary Table S2 for sequence data.

**Figure 6**. Cladogram of gene sequences performed with ClustalX software by using the Neighbour Joining (NJ) method of some *Pistacia vera* varieties of different geographical origin. A close phylogenetic relationship is present between the Mediterranean Mateur and Larnaka varieties. These two varieties are phylogenetically related to Kern and Kerman varieties. A close relationship is found between Bronte and Mawardi varieties. Bootstrap values were calculated from 100 resamplings of the alignment data.

546 **Table 1**. Total polyphenolic content (TPC), total proanthocyanidins content (TPACs) and total anthocyanin content (TAC) of some
547 pistachio varieties of different geographical origin. Mean values are expressed as mg $g^{-1}$ d.wt. (± SD). For each column, different letters
548 indicate significant ($P \leq 0.05$) differences.

| | TPC | | TPACs | | TAC | |
|---|---|---|---|---|---|---|
| **Variety** | **Seed flesh** | **Seed Skin** | **Seed flesh** | **Seed Skin** | **Seed flesh** | **Seed Skin** |
| Bronte | 1.55 (± 0.08)[a] | 363.75 (± 16.5)[a] | n.d. | 177.57 (± 0.40)[a] | n.d. | 27.31 (± 1.11)[a] |
| Kerman | 1.93 (± 0.03)[b] | 91.37 (± 1.04)[b] | n.d. | 88.51 (± 2.71)[b] | n.d. | 2.84 (± 0.12)[b] |
| Larnaka | 1.74 (± 0.04)[c] | 334.64 (± 15.41)[c] | n.d. | 155.09 (± 3.63)[c] | n.d. | 24.24 (± 0.24)[c] |
| Kern | 0.24 (± 0.01)[d] | 140.91 (± 11.6)[d] | n.d. | 54.48 (± 0.45)[d] | n.d. | 6.34 (± 0.36)[d] |
| Mateur | 0.18 (± 0.01)[d] | 181.55 (± 5.07)[e] | n.d. | 95.20 (± 3.35)[b] | n.d. | 9.79 (± 0.64)[e] |
| Mawardi | 0.18 (± 0.02)[d] | 290.28 (± 5.82)[f] | n.d. | 159.69 (± 2.35)[c] | n.d. | 6.74 (± 0.37)[d] |

549

550

551 **Table 2**. Qualitative and quantitative chemical analysis of the phenolic compounds present in the seed skin of some pistachio varieties of different

552 geographical origin. Mean values are expressed as μg $g^{-1}$ d.wt. (± SD). Within the same line, different letters indicate significant ($P \leq 0.05$) differences.

| Compound | RT | $[M-H]^-$ $[M-H]^+$ | m/z | λ | Varieties | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Bronte | Kerman | Larnaka | Kern | Mateur | Mawardi |
| Eriodictyol-7-galactoside | 16.4 | 449 | 288 | 360 | 71.71 (± 1.14)$^a$ | 43.67 (± 0.53)$^b$ | 366.67 (± 9.46)$^c$ | 135.14 (± 4.68)$^d$ | 116.9 (± 5.54)$^e$ | 88.28 (± 2.87)$^f$ |
| Idein | 18.7 | 449 | 286 | 520 | 1885.06 (± 23.58)$^a$ | 90.58 (± 1.72)$^b$ | 1774.73 (± 39.88)$^c$ | 661.66 (± 16.5)$^d$ | 739.3 (± 12.47)$^e$ | 416.85 (± 9.44)$^f$ |
| Cyanidin-3-glucoside* | 18.7 | 449 | 286 | 520 | 5297.52 (± 109.31)$^a$ | 737.62 (± 12.57)$^b$ | 5063.01 (± 97.01)$^c$ | 2219.66 (± 46.64)$^d$ | 2515.56 (± 45.92)$^e$ | 1675.92 (± 24.66)$^f$ |
| Eriodictyol-7-glucoside | 22.7 | 449 | 288 | 360 | 1194.42 ±( 27.91)$^a$ | 168.71 (± 1.11)$^b$ | 1116.88 (± 22.44)$^c$ | 425.76 (± 3.61)$^d$ | 562.18 (± 17.40)$^e$ | 347.34 (± 8.05)$^f$ |
| Peonidin-3-O-glycoside* | 23.9 | 463 | 301 | 520 | 120.03 (± 3.56)$^a$ | 23.46 (± 0.64)$^b$ | 244.31 (± 7.07)$^c$ | 103.06 (± 1.72)$^d$ | 82.32 (± 5.42)$^e$ | 21.23 (± 4.22)$^b$ |
| Catechin | 25.0 | | 289 | 280 | 1298.14 (± 35.78)$^a$ | 230.05 (± 5.68)$^b$ | 1931.68 (± 45.81)$^c$ | 172.61 (± 2.38)$^b$ | 204.57 (± 8.42)$^b$ | 2144.88 (± 22.11)$^d$ |
| Okanin 4'-O-galactoside | 26.8 | 449 | 288 | 280 | 325.42 (± 6.48)$^a$ | 66.55 (± 20)$^b$ | 398.75 (± 4.10)$^c$ | n.d. | 67.53 (± 2.46)$^b$ | 180.57 (± 4.72)$^d$ |
| Hyperoside | 26.8 | 463 | 302 | 360 | 314.47( ± 6.27)$^a$ | 82.35 (± 1.72)$^b$ | 533.95 (± 5.3)$^c$ | 131.47 (± 3.2)$^d$ | 151.8 (± 6.36)$^e$ | 131.85 (± 4.07)$^d$ |
| Quercetin-3-O-Glucoside | 29.8 | 463 | 302 | 360 | 248.6 (± 4.96)$^a$ | 96.4 (± 3.56)$^b$ | 195.04 (± 1.24)$^c$ | 139.15 (± 7.76)$^d$ | 179.48 (± 5.36)$^e$ | 89.52 (± 2.9)$^b$ |
| Marein | 32.8 | 449 | 288 | 360 | 221.23 (± 4.41)$^a$ | 107.91 (± 2.57)$^b$ | n.d. | 48.23 (± 0.82)$^c$ | 76.26 (± 6.32)$^d$ | n.d. |
| Luteolin-glucoside | 33.0 | 447 | 286 | 360 | 1029.3 (± 21.46)$^a$ | 19.24 (± 0.73)$^b$ | n.d. | 237.06 (± 6.32)$^c$ | 327.26 (± 8.75)$^d$ | n.d. |
| Quercetin-4'-O-Glucoside | 34.0 | 463 | 302 | 360 | 57.43 (± 0.1)$^a$ | 44.95 (± 1.08)$^b$ | 89.17 (± 2.23)$^c$ | 68.42 (± 3.21)$^d$ | 74.6 (± 4.87)$^d$ | 29.52 (± 1.70)$^e$ |

553 RT, retention time; λ, wavelength expressed in nm.

554

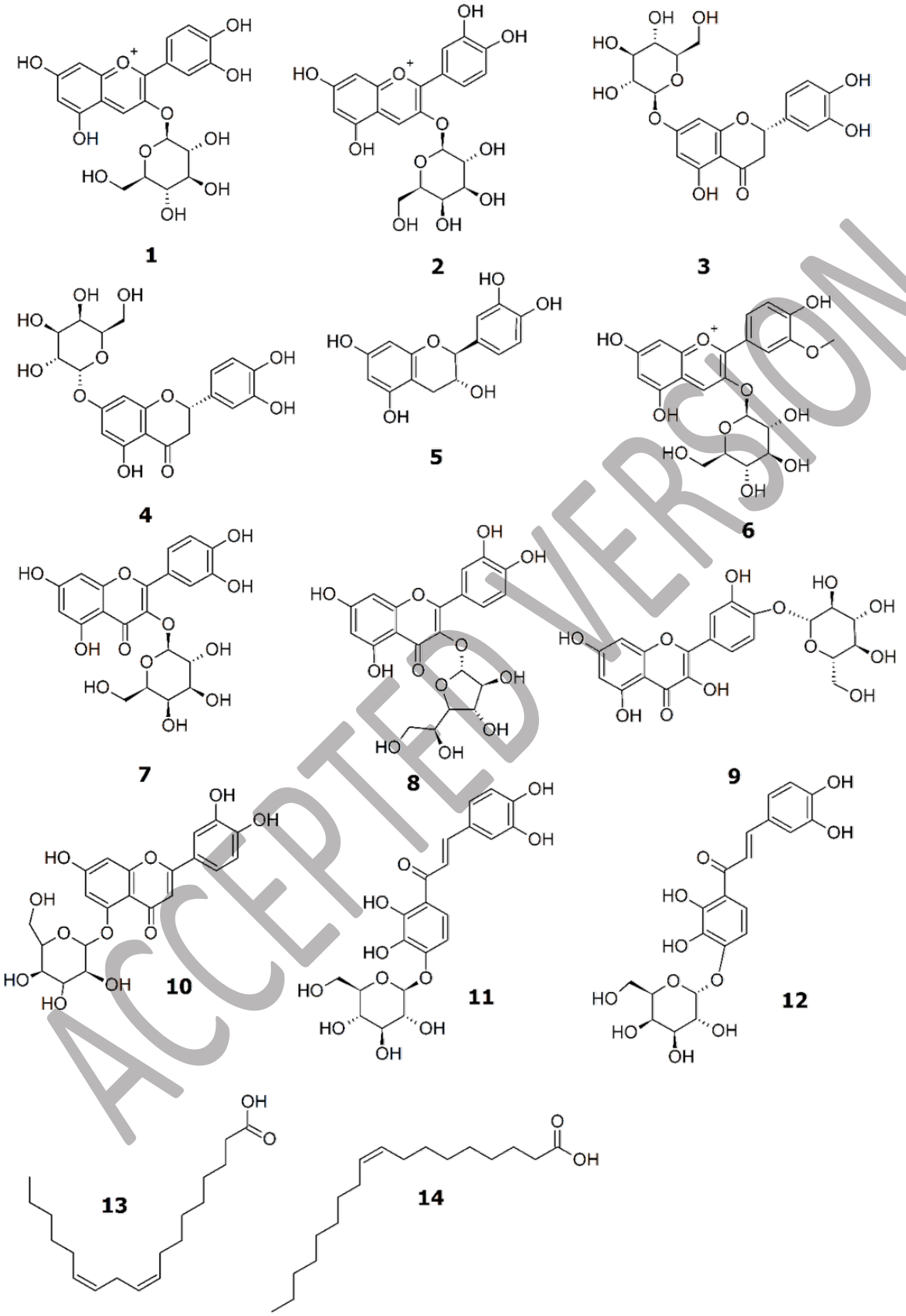
555 **Table 3**. Fatty acid composition of some pistachio varieties of different geographical origin. Mean values are expressed as mg $g^{-1}$ d.wt. (± SD). In the

556 same line, different letters indicate significant ($P<0.05$) differences. Ki, Kovats index.

| Compound | Ki | Variety | | | | | |
|---|---|---|---|---|---|---|---|
| | | **Bronte** | **Kerman** | **Larnaka** | **Kern** | **Mateur** | **Mawardi** |
| Palmitoleic acid | 1878 | 4.83 (± 0.05)$^{a}$ | 3.88 (± 0.09)$^{b}$ | 4.88 (± 0.01)$^{a}$ | 4.74 (± 0.02)$^{a}$ | 4.53 (± 0.01)$^{c}$ | 2.43 (± 0.08)$^{d}$ |
| Palmitic Acid | 1886 | 58.58 (± 0.42$^{c}$ | 48.24 (± 0.83)$^{a}$ | 57.41 (± 0.32)$^{b}$ | 60.92 (± 1.31)$^{c}$ | 64.74 (± 0.52)$^{d}$ | 46.14 (± 0.52)$^{e}$ |
| Linoleic acid | 2082 | 97.26 (± 2.74)$^{a,b}$ | 117.48 (± 19.3)$^{a}$ | 73.06 (± 4.16)$^{b}$ | 179.08 (± 17.01)$^{c}$ | 168.72 (± 3.6c)$^{a}$ | 78.92 (± 3.47)$^{b}$ |
| Oleic acid | 2085 | 431.86 (± 15.26)$^{a}$ | 245.69 (± 30.11)$^{b}$ | 408.15 (± 20.77)$^{a,c}$ | 355.3 (± 26.74)$^{b,c}$ | 378.18 (± 9.00)$^{a,c}$ | 384.69 (± 8.06)$^{a,c}$ |
| Elaidic acid | 2093 | 16.46 (± 0.14)$^{a}$ | 13.89 (± 0.02)$^{b}$ | 14.52 (± 0.15)$^{c}$ | 18.35 (± 0.22)$^{d}$ | 17.28 (± 0.03)$^{e}$ | 12.33 (± 0.04)$^{f}$ |
| Stearic Acid | 2133 | 9.35 (± 0.06)$^{a}$ | 5.41 (± 0.05)$^{b}$ | 8.83 (± 0.07)$^{c}$ | 6.99 (± 0.05)$^{d}$ | 8.68 (± 0.06)$^{c}$ | 10.95 (± 0.03)$^{d}$ |
| γ-Linolenic acid | 2220 | 1.32 (± 0.03)$^{a,b}$ | 1.05 (± 0.05)$^{c}$ | 1.26 (± 0.06)$^{b,d}$ | 0.89 (± 0.01)$^{e}$ | 1.42 (± 0.09)$^{a}$ | 1.18 (± 0.02)$^{c,f}$ |
| Arachidic Acid | 2284 | 0.72 (± 0.07)$^{a,b}$ | 0.42 (± 0.02)$^{c}$ | n.d. | n.d. | 0.66 (± 0.01)$^{b}$ | 0.77 (± 0.01)$^{a}$ |
| Total | | 620.61 (± 15.38)$^{a}$ | 435.46 (± 0.55)$^{b}$ | 568.14 (± 4.94)$^{c}$ | 626.46 (± 8.16)$^{a,d}$ | 643.58 (± 3.93)$^{d,e}$ | 537.05 (± 4.39)$^{e}$ |

557

558 FIGURE 1

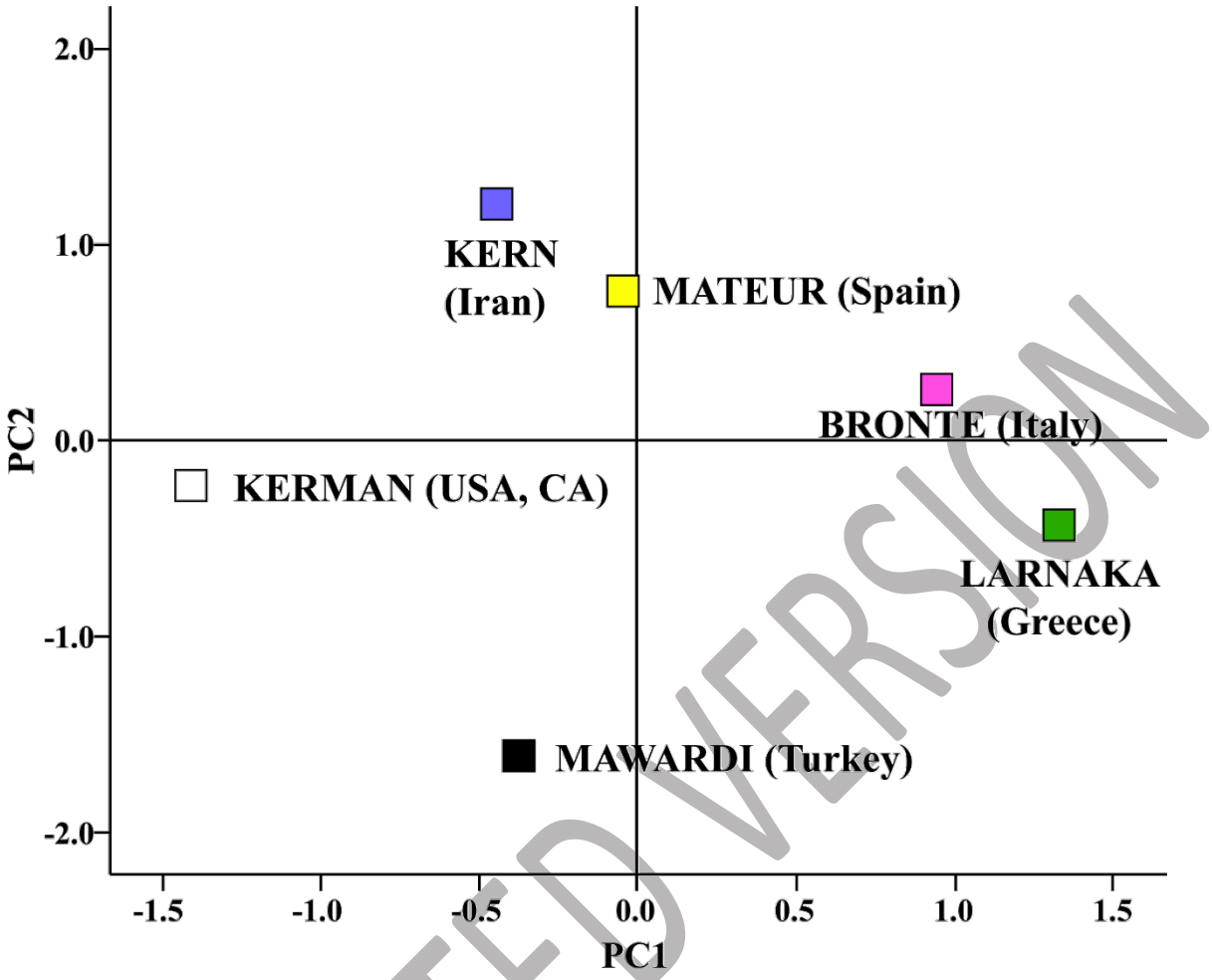





559

560

561 FIGURE 2

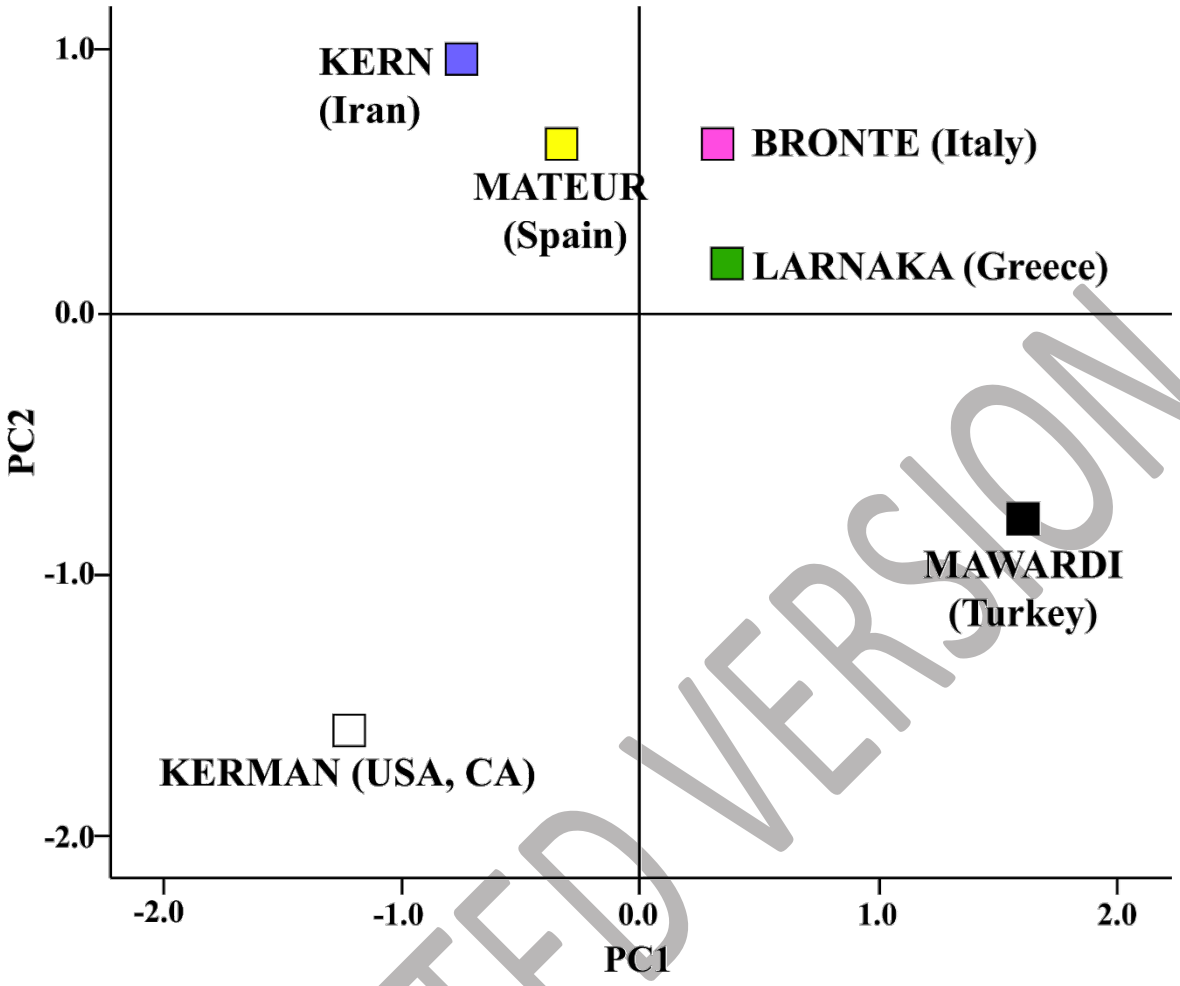



562

563

564

565 FIGURE 3



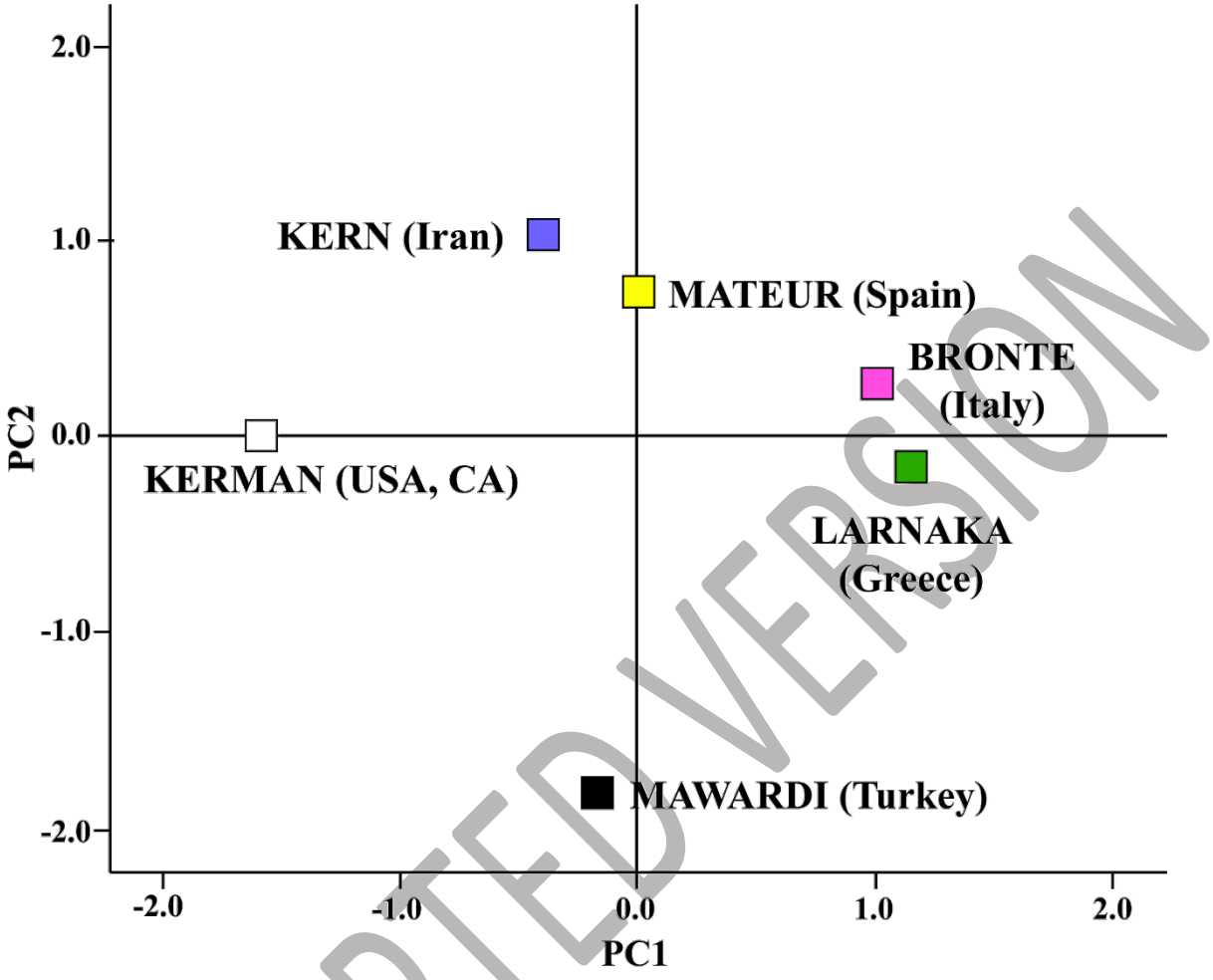

566

567

568

569 FIGURE 4

570

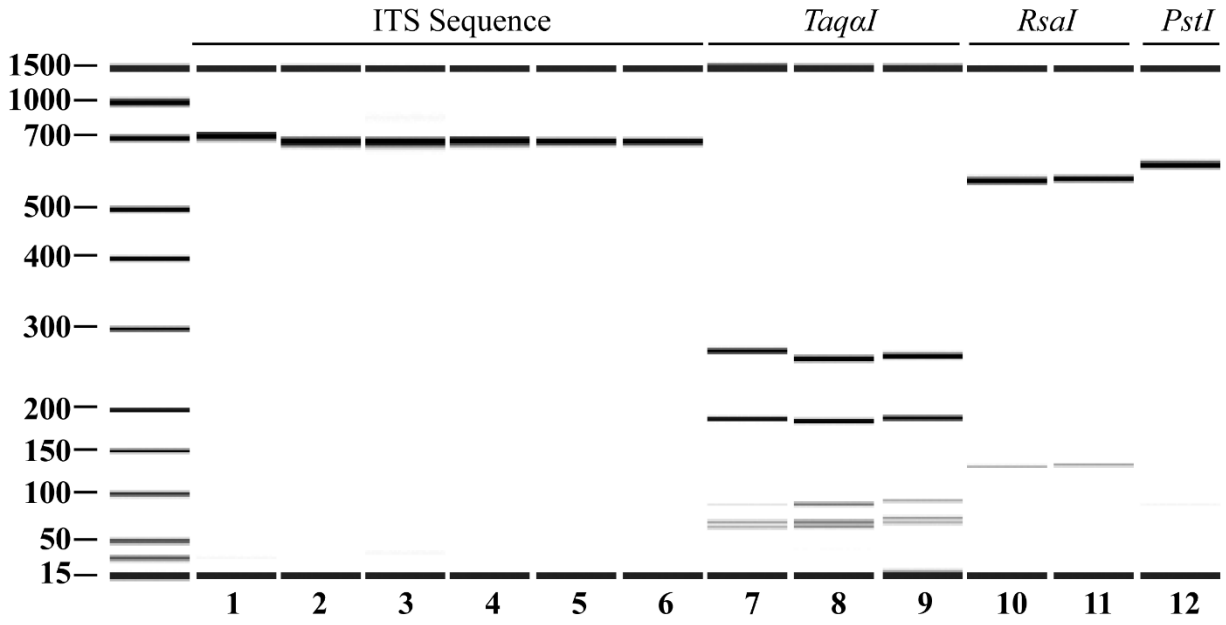



571

572

573

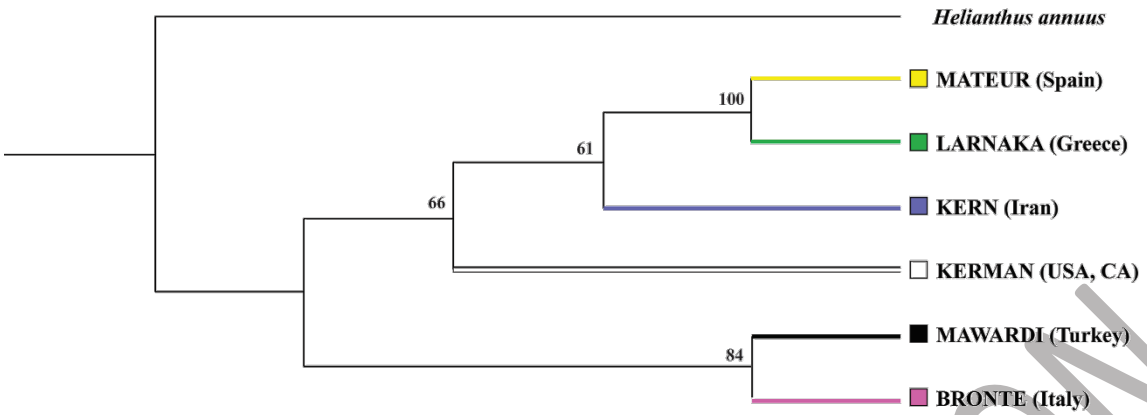
574 FIGURE 5




575

576

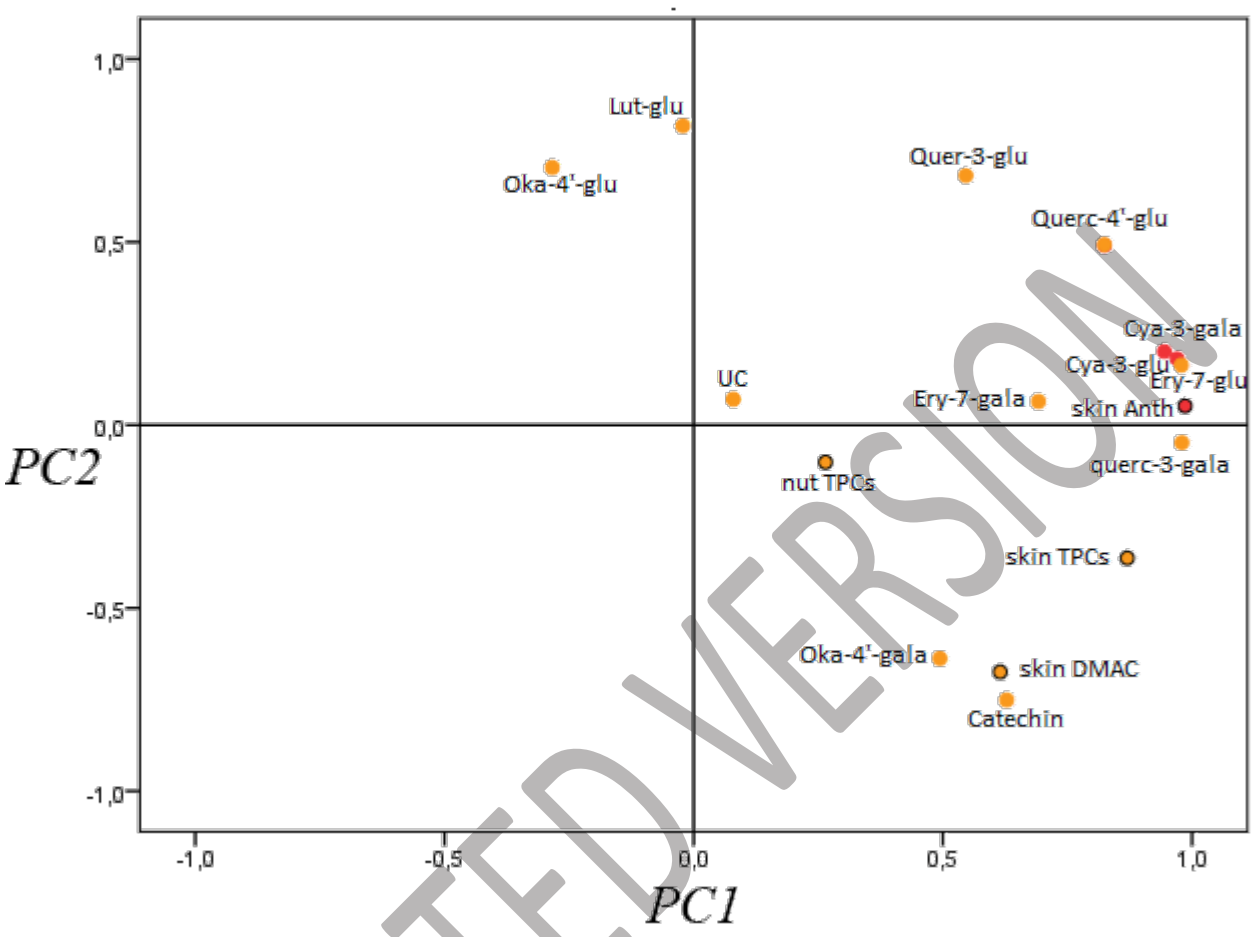
577 FIGURE 6




578

579

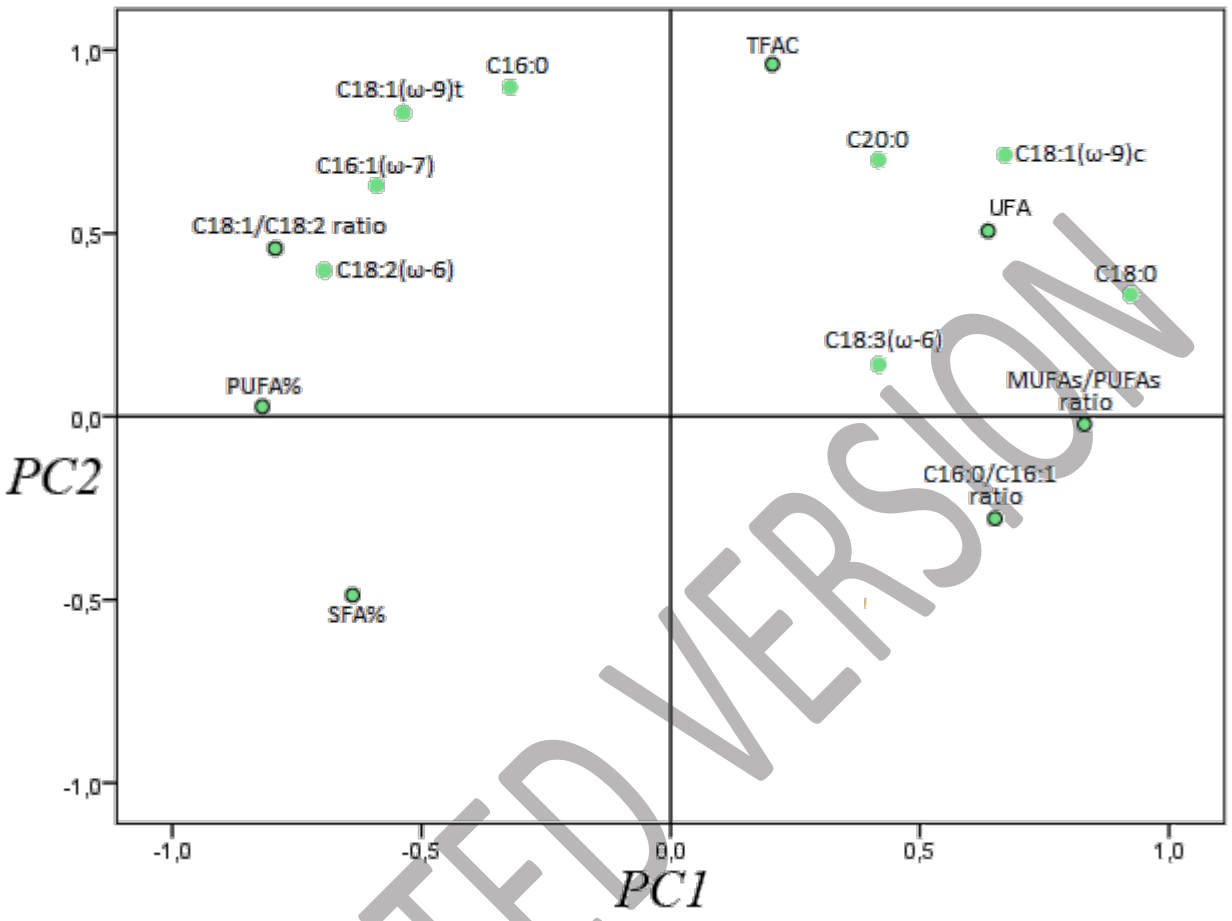
580 SUPPLEMENTARY FIGURE S1




581

582

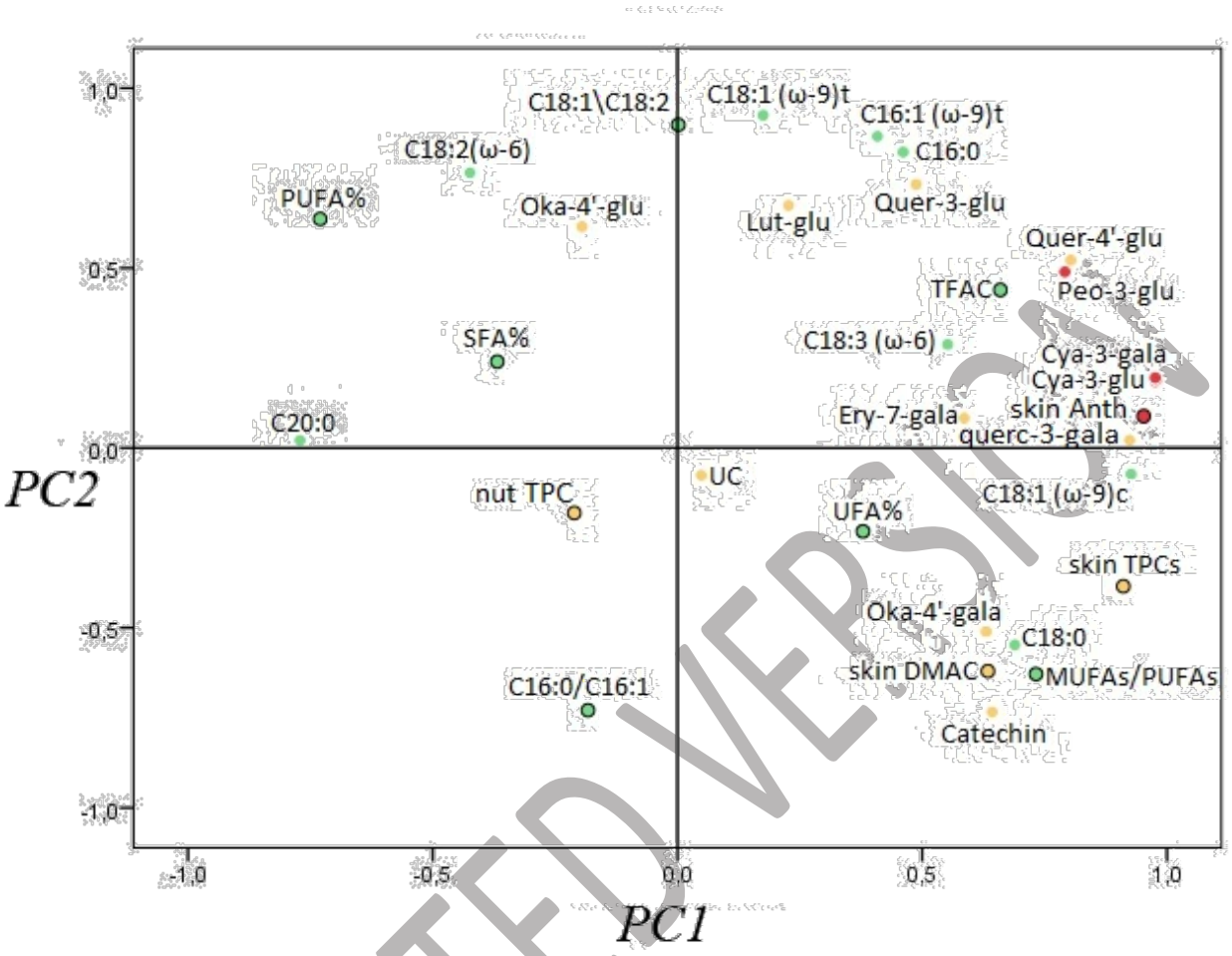
583 SUPPLEMENTARY FIGURE S2




584

585

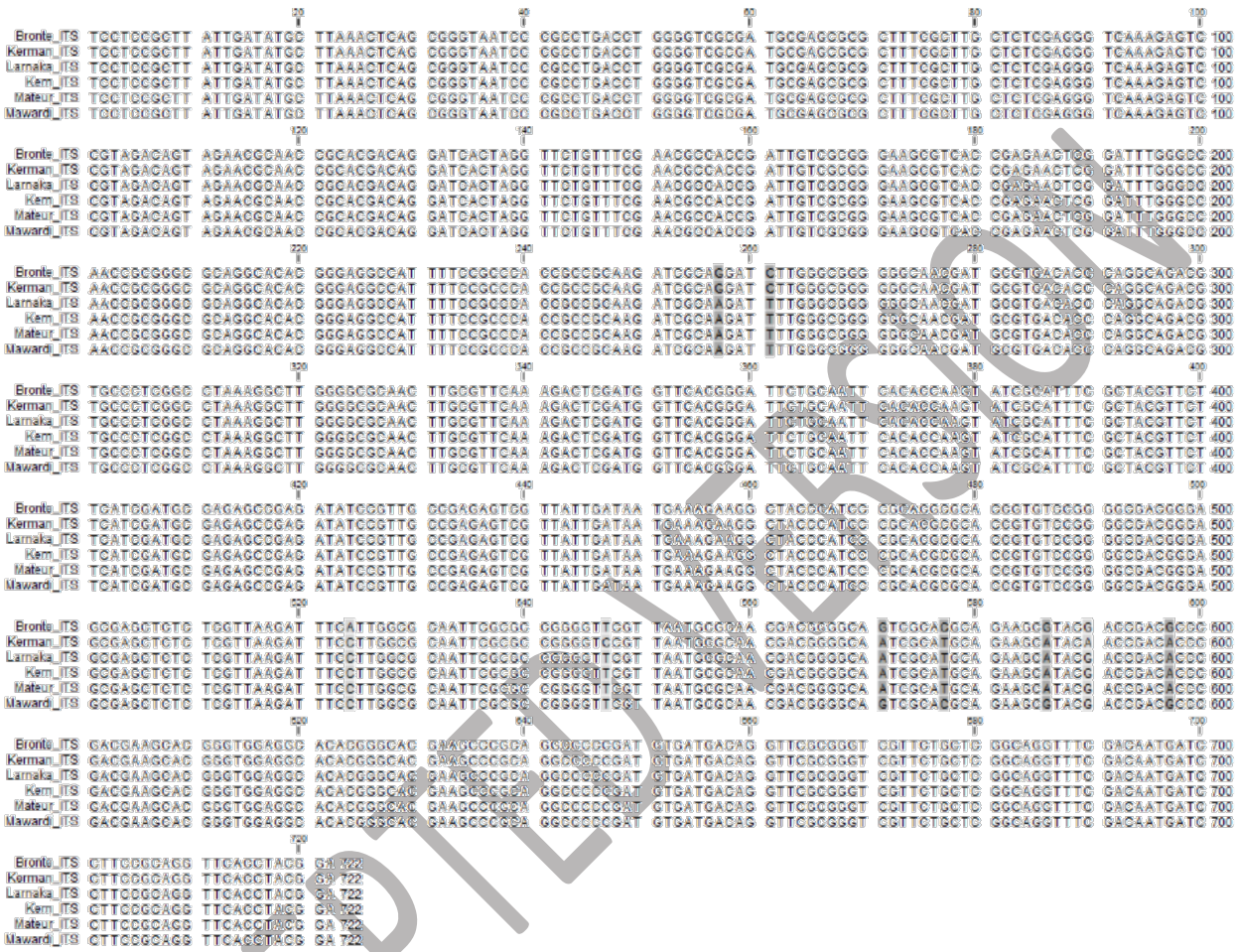
586 SUPPLEMENTARY FIGURE S3




587

588

SUPPLEMENTARY FIGURE S4



**Supplementary Table S1**: *Tukey's HSD post hoc differences in total polyphenols content (TPC), total anthocyanins content (TAC) and total proanthocyaninds content (t-PAC) among the six skin extracts of pistachio. *P<0.05; **P<0.01; ***P<0.001.*

| | | TPC | TAC | TPAC |
|---|---|---|---|---|
| Bronte | Kerman | 272.38*** | 24.47*** | 89.06*** |
| | Larnaka | 29.11 | 3.07*** | 22.48*** |
| | Kern | 222.84*** | 20.97*** | 123.09*** |
| | Mateur | 182.20*** | 17.52*** | 82.37*** |
| | Mawardi | 73.47*** | 20.57*** | 17.88*** |
| Kerman | Larnaka | -243.27*** | -21.40*** | -66.58*** |
| | Kern | -49.54** | ì-3.50*** | 34.03*** |
| | Mateur | -90.18*** | -6.95*** | -6.69 |
| | Mawardi | -198.91*** | -3.90*** | -71.18*** |
| Larnaka | Kern | 193.73*** | 17.90*** | 100.61*** |
| | Mateur | 153.09*** | 14.45*** | 59.89*** |
| | Mawardi | 44.36** | 17.50*** | -4.60 |
| Kern | Mateur | -40.64** | -3.45*** | -40.72*** |
| | Mawardi | -149.37*** | -0.4 | -105.21*** |
| Mateur | Mawardi | -108.73*** | 3.05*** | -64.49*** |

599 **Supplementary Table S2**: *Sequences of each* ITS fragments generated after RFLP analysis with

600 RsaI, TaqαI and PstI restriction enzyme. Lowercase letter indicate the band reported in Figure 4.

601 Letter "a" denotes the highest band.

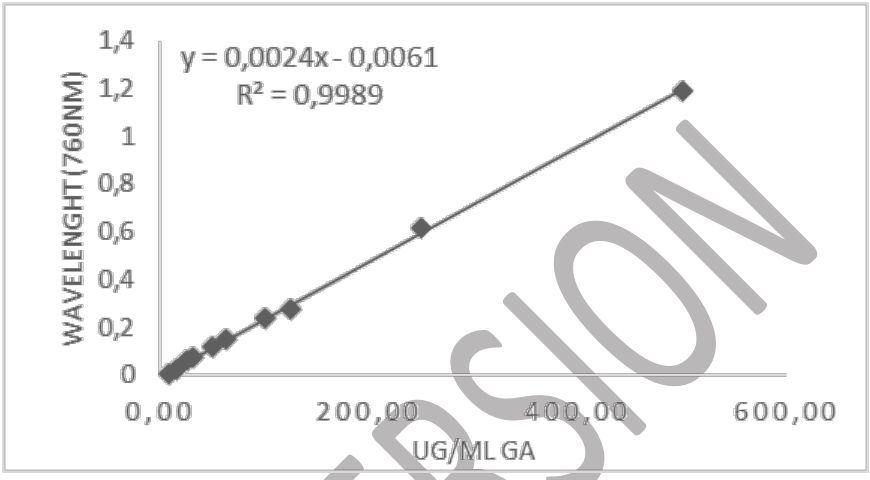
| *Type of Sequence* | *Lenght* | *Sequence* |
|---|---|---|
| ITS1-4 Bronte -RsaI-a | 587 | ACGCTTCTGCGTGCGACTGCCCCGTCGTTGCGCATTAACGAACCCCGGCGCGAATTGCGCCAAGGAAATCTTAACGAGAGAGCTCGCTCCCGTCGCCCCGGACACG GTGCGCGTGCGGGATGGGTAGCCTTCTTTCATTATCAATAACGACTCTCGGCAACGGATATCTCGGCTCTCGCATCGATGAAGAACGTAGCGAAATGCGATACTTGG TGTGAATTGCAGAATCCCGTGAACCATCGAGTCTTTGAACGCAAGTTGCGCCCCAAGCCTTTAGGCCGAGGGCACGTCTGCCTGGGTGTCACGCATCGTTGCCCCCC GCCCAAGATCCTGCGATCTTGCGGCGGTGGGCGGAAAATGGCCTCCCGTGTGCCTGCGCCCGCGGTTGGCCCAAATCCGAGTTCTCGGTGACGCTTCCCGCGACAAT CGGTGGCGTTCGAAACAGAACCTAGTGATCCTGTCGTGCGGTTGCGTTCTACTGTCTACGGACTCTTTGACCCTCGAGAGCAAGCGAAAGCGCGCTCGCATCGCGAC CCCAGGTCAGGCGGGATTACCCGCTGAGTTTAGGCATATCAATAAGCGGAGGA |
| ITS1-4 Bronte -RsaI-b | 136 | TCCGTAGGTGAACCTGCGGAAGGATCATTCGTCGAAACCTGCCGAGCAGAACGACCCGCGAACCTGTCATCACATCGGGGGCCTGCGGGCTTCGTGCCCGTGTGCCC TCCACCCGTGCTTCGTCGGGCGTCGGTCGT |
| ITS1-4 Kerman-TaqαI-a | 285 | CGAAACCTGCCGAGCAGAACGACCCGCGAACCTGTCATCACATCGGGGGCCTGCGGGCTTCGTGCCCGTGTGCCTCCACCCGTGCTTCGTCGGGTGTCGGTTGTATG CTTCTGCATGCGATTGCCCCGTCGTTGCGCATTAACGGACCCCGGCGCGAATTGCGCCAAGGAAATCTTAACGAGAGAGCTCGCTCCCGTCGCCCCGGACACGGTGC GCGTGCGGGATGGGTAGCCTTCTTTCATTATCAATAACGACTCTCGGCAACGGATATCTCGGCTCTCGCAT |
| ITS1-4 Kerman-TaqαI-b | 197 | CGAGTCTTTGAACGCAAGTTGCGCCCCAAGCCTTTAGGCCGAGGGCACGTCTGCCTGGGTGTCACGCATCGTTGCCCCCCGCCCAAGATCCTGCGATCTTGCGGCGG TGGGCGGAAAATGGCCTCCCGTGTGCCTGCGCCCGCGGTTGGCCCAAATCCGAGTTCTCGGTGACGCTTCCCGCGACAATCGGTGGCGTT |
| ITS1-4 Kerman-TaqαI-c | 66 | CGAGAGCAAGCGAAAGCGCGCTCGCATCGCGACCCCAGGTCAGGCGGGATTACCCGCTGAGTTTAA |
| ITS1-4 Kerman-TaqαI-d | 64 | CGAAACAGAACCTAGTGATCCTGTCGTGCGGTTGCGTTCTACTGTCTACGGACTCTTTGACCCT |
| ITS1-4 Kerman-TaqαI-e | 59 | CGATGAAGAACGTAGCGAAATGCGATACTTGGTGTGAATTGCAGAATCCCGTGAACCAT |
| ITS1-4 Kerman-TaqαI-f | 31 | TCCGTAGGTGAACCTGCGGAAGGATCATTGT |
| ITS1-4 Larnaka-TaqαI-a | 285 | CGAAACCTGCCGAGCAGAACGACCCGCGAACCTGTCATCACATCGGGGGCCTGCGGGCTTCGTGCCCGTGTGCCTCCACCCGTGCTTCGTCGGGTGTCGGTCGTATG CTTCTGCATGCGATTGCCCCGTCGTTGCGCATTAACGAACCCCGGCGCGAATTGCGCCAAGGAAATCTTAACGAGAGAGCTCGCTCCCGTCGCCCCGGACACGGTGC GCGTGCGGGATGGGTAGCCTTCTTTCATTATCAATAACGACTCTCGGCAACGGATATCTCGGCTCTCGCAT |
| ITS1-4 Larnaka-TaqαI-b | 197 | CGAGTCTTTGAACGCAAGTTGCGCCCCAAGCCTTTAGGCCGAGGGCACGTCTGCCTGGGTGTCACGCATCGTTGCCCCCCGCCCAAAATCTTGCGATCTTGCGGCGG TGGGCGGAAAATGGCCTCCCGTGTGCCTGCGCCCGCGGTTGGCCCAAATCCGAGTTCTCGGTGACGCTTCCCGCGACAATCGGTGGCGTT |
| ITS1-4 Larnaka-TaqαI-c | 86 | CGAGAGCAAGCGAAAGCGCGCTCGCATCGCGACCCCAGGTCAGGCGGGATTACCCGCTGAGTTTAAGCATATCAATAAGCGGAGGA |
| ITS1-4 Larnaka-TaqαI-d | 64 | CGAAACAGAACCTAGTGATCCTGTCGTGCGGTTGCGTTCTACTGTCTACGGACTCTTTGACCCT |
| ITS1-4 Larnaka-TaqαI-e | 59 | CGATGAAGAACGTAGCGAAATGCGATACTTGGTGTGAATTGCAGAATCCCGTGAACCAT |
| ITS1-4 Larnaka-TaqαI-f | 31 | TCCGTAGGTGAACCTGCGGAAGGATCATTGT |
| ITS1-4 Kern -PstI-a | 640 | TCCTCCGCTTATTGATATGCTTAAACTCAGCGGGTAATCCCGCCTGACCTGGGGTCGCGATGCGAGCGCGCTTTCGCTTGCTCTCGAGGGTCAAAGAGTCCGTAGACA GTAGAACGCAACCGCACGACAGGATCACTAGGTTCTGTTTCGAACGCCACCGATTGTCGCGGGAAGCGTCACCGAGAACTCGGATTTGGGCCAACCGCGGGCGCA GGCACACGGGAGGCCATTTTCCGCCCACCGCCGCAAGATCGCAGGATTTTGGGCGGGGGGCAACGATGCGTGACACCCAGGCAGACGTGCCCTCGGCCTAAAGGC TTGGGGCGCAACTTGCGTTCAAAGACTCGATGGTTCACGGGATTCTGCAATTCACACCAAGTATCGCATTTCGCTACGTTCTTCATCGATGCGAGAGCCGAGATATCC GTTGCCGAGAGTCGTTATTGATAATGAAAGAAGGCTACCCATCCCGCACGCGCACCGTGTCCGGGGCGACGGGAGCGAGCTCTCTCGTTAAGATTTCCTTGGCGCA ATTCGCGCCGGGGTTCGTTAATGCGCAACGACGGGGCAATCGCATGCAGAAGCATACGACCGACACCCGACGAAGCACGGGTGGAGGCACACGGGCACGAAGCC TGCA |
| ITS1-4 Kern -PstI-b | 82 | GGCCCCCGATGTGATGACAGGTTCGCGGGTCGTTCTGCTCGGCAGGTTTCGACAATGATCCTTCCGCAGGTTCACCTACGGA |
| ITS1-4 Mateur-TaqαI-a | 84 | TCCTCCGCTTATTGATATGCTTAAACTCAGCGGGTAATCCCGCCTGACCTGGGGTCGCGATGCGAGCGCGCTTTCGCTTGCTCT |
| ITS1-4 Mateur-TaqαI-b | 64 | CGAGGGTCAAAGAGTCCGTAGACAGTAGAACGCAACCGCACGACAGGATCACTAGGTTCTGTTT |
| ITS1-4 Mateur-TaqαI-c | 197 | CGAACGCCACCGATTGTCGCGGGAAGCGTCACCGAGAACTCGGATTTGGGCCAACCGCGGGCGCAGGCACACGGGAGGCCATTTTCCGCCCACCGCCGCAAGATC GCAAGATTTTGGGCGGGGGGCAACGATGCGTGACACCCAGGCAGACGTGCCCTCGGCCTAAAGGCTTGGGGCGCAACTTGCGTTCAAAGACT |
| ITS1-4 Mateur-TaqαI-d | 59 | CGATGGTTCACGGGATTCTGCAATTCACACCAAGTATCGCATTTCGCTACGTTCTTCAT |
| ITS1-4 Mateur-TaqαI-e | 285 | CGATGCGAGAGCCGAGATATCCGTTGCCGAGAGTCGTTATTGATAATGAAAGAAGGCTACCCATCCCGCACGCGCACCGTGTCCGGGGCGACGGGAGCGAGCTCT CTCGTTAAGATTTCCTTGGCGCAATTCGCGCCGGGGTTCGTTAATGCGCAACGACGGGGCAATCGCATGCAGAAGCATACGACCGACACCCGACGAAGCACGGGTG GAGGCACACGGGCACGAAGCCCGCAGGCCCCCGATGTGATGACAGGTTCGCGGGTCGTTCTGCTCGGCAGGTTT |
| ITS1-4 Mateur-TaqαI-f | 33 | CGACAATGATCCTTCCGCAGGTTCACCTACGGA |
| ITS1-4 Mawardi -RsaI-a | 587 | ACGCTTCTGCGTGCGACTGCCCCGTCGTTGCGCATTAACGAACCCCGGCGCGAATTGCGCCAAGGAAATCTTAACGAGAGAGCTCGCTCCCGTCGCCCCGGACACG GTGCGCGTGCGGGATGGGTAGCCTTCTTTCATTATCAATAACGACTCTCGGCAACGGATATCTCGGCTCTCGCATCGATGAAGAACGTAGCGAAATGCGATACTTGG TGTGAATTGCAGAATCCCGTGAACCATCGAGTCTTTGAACGCAAGTTGCGCCCCAAGCCTTTAGGCCGAGGGCACGTCTGCCTGGGTGTCACGCATCGTTGCCCCCC GCCCAAAATCTTGCGATCTTGCGGCGGTGGGCGGAAAATGGCCTCCCGTGTGCCTGCGCCCGCGGTTGGCCCAAATCCGAGTTCTCGGTGACGCTTCCCGCGACAAT CGGTGGCGTTCGAAACAGAACCTAGTGATCCTGTCGTGCGGTTGCGTTCTACTGTCTACGGACTCTTTGACCCTCGAGAGCAAGCGAAAGCGCGCTCGCATCGCGAC CCCAGGTCAGGCGGGATTACCCGCTGAGTTTAAGCATATCAATAAGCGGAGGA |
| ITS1-4 Mawardi -RsaI-b | 135 | TCCGTAGGTGAACCTGCGGAAGGATCATTGTCGAAACCTGCCGAGCAGAACGACCCGCGAACCTGTCATCACATCGGGGGCCTGCGGGCTTCGTGCCCGTGTGCCT CCACCCGTGCTTCGTCGGGCGTCGGTCGT |

602

603

**Supplementary Table S3**: C*alibration curve of Gallic Acid (GA) and proanthocyaninds A-type dimers (PAC-A) used for the quantification of total polyphenol content (TPC) and total proanthocyanidins content (t-PACs) in pistachio extracts.*

| µg/mL GA | TPC |
|---|---|
| 500.00 | 1.1881 |
| 250.00 | 0.6189 |
| 125.00 | 0.2736 |
| 100.00 | 0.2383 |
| 62.50 | 0.1517 |
| 50.00 | 0.1159 |
| 31.25 | 0.0728 |
| 25.00 | 0.0611 |
| 15.63 | 0.0215 |
| 7.81 | 0.0046 |




| µg/mL | t-PAC |
|---|---|
| 50.00 | 1.0857 |
| 40.00 | 0.9251 |
| 31.25 | 0.7175 |
| 25.00 | 0.5682 |
| 15.63 | 0.3441 |
| 12.50 | 0.2842 |
| 10.00 | 0.2072 |
| 7.81 | 0.1574 |
| 5.00 | 0.1022 |
| 2.50 | 0.0505 |