Phytochemical data parallel morpho-colorimetric variation in *Polygala flavescens* DC.

Lorenzo Peruzzi [a], Francesco Roma-Marzio [b], David Dolci [a], Guido Flamini [c], Alessandra Braca [c], Marinella De Leo [c]

[a] Dipartimento di Biologia, Università di Pisa, via Derna 1, 56126 Pisa, Italy

[b] Sistema Museale di Ateneo, Orto e Museo Botanico, Università di Pisa, Via Luca Ghini 13, 856126, Pisa, Italy

[c] Dipartimento di Farmacia, Università di Pisa, Via Bonanno 33, 56126 Pisa, Italy

Abstract

Phytochemical data, integrated with other sources of information, represent a valuable tool helping to solve different kinds of taxonomic problems in plant systematics. In the present study, a comparative investigation, in order to clarify the systematic relationships of the three subspecies currently recognized within the Italian endemic *Polygala flavescens*, was carried out. Preliminarily, a morphometric and colorimetric analysis, in order to test the degree of distinctiveness among the taxa, was performed. Then, a phytochemical analysis based both on volatile and non-volatile compounds was obtained. Concerning the morpho-colorimetric analysis, our results confirm most of the characters as useful to discriminate the three subspecies. In addition, some volatile and non-volatile compounds are good taxonomic markers. Morpho-colorimetric variation is clearly paralleled by phytochemical results, confirming the value of this kind of data to infer relationships in plant systematics. Based on these results, we support a taxonomic treatment at subspecific level for the involved taxa. Finally, based on the most significant morphological characters, a revision of herbarium specimens allowed to redefine the distribution of the three subspecies. Accordingly, the range of *P. flavescens* subsp. *maremmana* is limited to Mt. Argentario (southern Tuscany) only. Finally, a key is reported for the identification of the three subspecies.

Keywords: *Polygala flavescens* subsp. *flavescens*, *Polygala flavescens* subsp. *maremmana*, *Polygala flavescens* subsp. *pisaurensis*, Polygalaceae, morphometrics, volatiles, saponins, flavonoids, oligosaccharides, Italy, identification key.

1. Introduction

Phytochemical data are valuable sources of comparative information, helping to solve different kinds of taxonomic problems in plant systematics (Stuessy 2009). However, as stressed also for karyology (Astuti et al. 2017), it is fundamental to integrate phytochemistry with other sources of information, to infer systematic relationships. Several recent studies highlighted congruence

between phytochemical data and other sources of information, concerning for instance *Juniperus oxycedrus* L. group - Cupressaceae (Roma-Marzio et al. 2017), *Crocus* L. ser. *Verni* B.Mathew - Iridaceae (Carta et al. 2015), *Lavandula* L. sect. *Lavandula* (Passalacqua et al. 2017) and *Salvia fruticosa* Mill. - Lamiaceae (Tundis et al. 2016). Recently, a phytochemical study on *Polygala flavescens* DC. subsp. *flavescens* (Polygalaceae), based on plants collected in Tuscany (De Leo et al. 2017), led to the isolation and structural characterization of 14 compounds, including six flavonol glycosides, four oligosaccharides, an apocarotenoid, and three triterpenoid saponins. Consequently, in order to clarify the systematic relationships of the three subspecies currently recognized within *P. flavescens* (Bartolucci et al. 2018), a comparative integrated phytochemical and morpho-colorimetric study was carried out.

The genus *Polygala* L. is the largest of the family Polygalaceae, comprising between 325 (Heywood et al. 2007) and 725 (Paiva 1998) species. This genus shows a high diversity of life forms and adaptive strategies, occupying a wide range of ecological niches and showing a nearly cosmopolitan distribution, with the exception of the Arctic, Antarctica, and New Zealand (Paiva 1998). The only comprehensive taxonomic treatment of *Polygala* was published by Chodat (1893), whereas there have been numerous regional treatments, suggesting various morphological traits for taxonomic use (e.g. Marques 1979; Paiva 1998; Bernardi 2000; Peruzzi et al. 2005; Arrigoni 2014). According to Conti et al. (2005), in Italy 28 taxa (including species and subspecies) occur, 14 of which are endemic to the country (Peruzzi et al. 2014). In a recent taxonomic revision of this genus in Italy (Arrigoni 2014), the number of taxa was raised to 35. One of these species, within *P.* subg. *Polygala* (McNeill 1968, is *Polygala flavescens* DC., which includes three taxonomically doubtful infraspecific taxa. *Polygala flavescens* is an Italian endemic species, originally described from Central Italy (Roma-Marzio and Peruzzi 2017), which is actually recorded all along the Italian peninsula, from Emilia-Romagna to Basilicata (Peruzzi et al. 2014). Most of the authors (Zangheri 1976; Pignatti 1982; Conti et al. 2005; Arrigoni 2014) treated this taxon at specific rank, with the exception of Fiori (1925), who considered it as a variety of *Polygala vulgaris* L. The latter author

also recorded *P. vulgaris* var. *flavescens* (DC.) Fiori f. *maremmana* (Fiori) Fiori, originally described as *P. flavescens* var. *maremmana* Fiori (Fiori et al. 1908). The latter taxon is currently recognized by Arrigoni (2014) as *P. flavescens* subsp. *maremmana* (Fiori) Arrigoni, a subspecies with a range putatively limited to the coasts of southern Tuscany, from the southern part of the Leghorn province to Mt. Argentario (Arrigoni 2014). *Polygala flavescens* subsp. *maremmana* is still recognised at varietal rank by Pignatti (2017), whereas Bartolucci et al. (2018) consider it as a taxonomically doubtful subspecies. Another species, *P. pisaurensis* Caldesi, was described based on plants collected in Marche near Pesaro, and it has always been considered very closely related to *P. flavescens* (McNeill 1968; Zangheri 1976; Pignatti 1982; 2017). Recently, Arrigoni (2014) and Bartolucci et al. (2018) treated *P. pisaurensis* as a subspecies of *P. flavescens*, i.e. *P. flavescens* subsp. *pisaurensis* (Caldesi) Arcang., while Pignatti (2017) is still considering it as a distinct species. However, these taxonomic changes were made in the absence of any quantitative observation. In addition, more recently, the three subspecies were shown to share the same chromosome number, i.e. $2n = 22$ (Peruzzi et al. 2017).

In order to quantitatively test the degree of morphological distinctiveness among the three taxa within the *Polygala flavescens* DC. group, morpho-colorimetric analyses were performed. Furthermore, their phytochemical composition was investigated, in order to test the congruence with morpho-colorimetric results, and to provide a more reliable taxonomic treatment. Finally, to update the distribution of the involved taxa, based on morphometric results, herbarium specimens were critically revised.

2. Materials & Methods

2.1 Plant material

Since the range of *Polygala flavescens* subsp. *flavescens* covers a large portion of the Italian Peninsula, we chose to sample three populations for this taxon, selected in order a) to cover a reasonable part of its distribution, b) to include its topotypical area. Concerning *P. flavescens* subsp.

105*maremmana* and *P. flavescens* subsp. *pisaurensis*, since both these subspecies are taxonomically
106doubtful and show very restricted ranges, we decided to limit their sampling to topotypical areas
107only, in order to have reliable results to be compared with the autonym subspecies. Type locality
108areas of the three taxa were identified based on the information published by Arrigoni (2004) and
109Roma-Marzio and Peruzzi (2017). The population sampled in Tuscany for *P. flavescens* subsp.
110*flavescens* (Polygalaceae) is the same already studied by De Leo et al. (2017) (Table 1). For each
111locality, a herbarium voucher was deposited at PI (herbarium acronyms follow Thiers 2017).
112Since all the sampling localities fall outside protected areas, and the studied taxa are not
113endangered, no specific permissions were required for our activities.
114
115*2.2 Chemicals*
116All solvents are HPLC grade and were purchased from VWR. HPLC grade water (18 MΩ) was
117prepared by a Mill- Ω50 purification system (Millipore Corp.). Standard flavonoids (5, 8, 10, 12,
11814, and 17), saponins (35 and 43), oligosaccharides (3, 6, 7, and 9), and an apocarotenoid (2) were
119previously isolated and fully characterized from *P. flavescens* subsp. *flavescens* DC. (Polygalaceae)
120(PFF-T) in our laboratory (De Leo et al. 2017).
121
122*2.3 Morphometric analyses*
123We sampled 20 individuals for each population. On each individual, we selected one stem, two well
124developed flowers, and one middle cauline leaf for the measurement of 10 characters (Table 2). The
125measurements obtained from the two flowers were averaged to a single value per individual. In
126addition, for each population, we measured 50 fruits and 20 seeds, for a total of 8 characters (Table
1272). Entire plants, fruits and seeds were scanned and then measured by means of ImageJ 1.47
128software (Rasband 1997). Three data matrices were built (S1.1; S1.2; S1.3 Tables): one for flower,
129leaf, and stem characters (dataset 1 in Table 2), one for fruits (dataset 2 in Table 2), and one for
130seeds (dataset 3 in Table 2).

131

132*2.4 Flower colorimetric analysis*

133Since the colour of the flowers could change with phenology (Weiss 1995), to quantitatively
134evaluate differences of this characters, pictures of 20 flowers at the same developmental stage (from
135plants in full blossom and without fruits) for each population were taken under the same light
136conditions. Then, using the image analysis software Gimp 2.8.14 (Kimball and Mattis 2014), the
137relative contributions of Red, Green and Blue (RGB) of flower wing, fringe and tube were
138measured, averaging the values obtained in an area of 300 pixels (S2 Table). While, in systematics,
139a RGB quantitative approach was previously used to compare diaspores (Bacchetta et al. 2008;
140Grillo et al. 2012), to the best of our knowledge this is the first time that this approach is used to
141quantify the differences in the colour of flowers.

142

143*2.5 Volatile Organic Compounds analysis*

144Volatile organic compounds (VOCs) were investigated separately in flowers, leaves, fruits and
145seeds from living plants collected in the field and temporarily cultivated in the Botanical Garden of
146the University of Pisa.

147SPME (Solid Phase Micro-extraction) sampling was performed for all the analyses using the same
148new fibre, preconditioned according to the manufacturer's instructions. Sampling was performed in
149an air-conditioned room (23 ± 1 °C) to guarantee a stable temperature during sampling. Supelco
150SPME devices coated with polydimethylsiloxane (PDMS, 100 μm) were used to collect the
151volatiles emitted by flowers, leaves, fruits and seeds inserted into a 12 ml glass septum vial, and
152allowed to equilibrate for 20 min. Subsequently, the fibre was exposed to the headspace for 25 min.
153Once sampling was finished, the fibre was withdrawn into the needle and transferred to the injection
154port of a GC/MS (Gas Chromatography-Mass Spectrometry) system.

155GC/Electron Impact (EI)-MS analyses were performed with a Varian CP-3800 gas-chromatograph
156equipped with a DB-5 capillary column (30 m × 0.25 mm; coating thickness 0.25 μm) linked to a

157Varian Saturn 2000 ion trap mass detector. Analytical conditions: injector and transfer line
158temperatures 250 and 240 °C, respectively; oven temperature programmed from 60 to 240 °C at 3
159°C/min; carrier gas helium at 1 ml/min; splitless injection. Identification of the constituents was
160based on comparison of the retention times with those of authentic samples, comparing their linear
161retention indices relative to the series of *n*-alkanes, and on computer matching against commercial
162(NIST 14 and ADAMS) and home-made library mass spectra built up from pure substances and
163components of known mixtures and MS literature data (Stenhagen et al. 1974, Masada 1976,
164Jennings and Shibamoto 1980, Swigar et al. 1981, Davies 1990, Adams 2007). SPME sampling and
165desorption conditions were identical for all samples. Furthermore, blanks were performed before
166each first SPME extraction and randomly repeated during each series. Quantitative comparisons of
167relative peak areas were performed between the same chemicals in different samples. All analyses
168were performed at least in triplicate.
169
170*2.6 Non-volatile compounds analysis*
171Non-volatile compounds were obtained from dried and powdered aerial parts (5 mesh) of flowering
172plants from each population. Plants (367.4 g of PFF-T, 313.4 g of PFF-A, 311.3 g of PFF-M, 152.9
173g of PFP, and 184.6 g of PFM) were first defatted with *n*-hexane and successively extracted at room
174temperature with methanol (1 g of dried drug in 5 ml of solvent for three times, every 24 h). The
175obtained extracts were dried under vacuum at 38 °C to give 5.7 and 118.4 g (PFF-T), 5.0 and 93.9 g
176(PFF-A), 4.7 and 103.5 g (PFF-M), 1.8 and 40.3 g (PFP), and 1.8 and 40.4 g (PFM) of n-hexane
177and methanol residues, respectively. The dried methanol extracts (5 g each) were partitioned
178between ethyl acetate and *n*-butanol, and water. The obtained *n*-butanol extracts were dried and
179dissolved in methanol (2.0 mg/ml) and centrifugated. Finally, 20 µl of each supernatant solution
180were injected in the HPLC-PDA/UV-ESI-MS system.
181HPLC-PDA/UV-ESI-MS/MS analyses were performed using a Surveyor LC pump, a Surveyor
182autosampler, coupled with a Surveyor PDA detector, and a LCQ Advantage ion trap mass

183spectrometer (ThermoFinnigan) equipped with Xcalibur 3.1 software. Analyses were performed
184using a 4.6 × 250 mm, 4 µm, Synergi Fusion-RP column (Phenomenex).The eluent was a mixture
185of methanol (solvent A) and a 0.1% aqueous solution of formic acid (solvent B). The solvent
186gradient was as follows: 0–20 min, 35% A isocratic mode; 20–35 min, 35-50% A; 35–48 min, 50%
187A isocratic mode; 48–108 min, 50-80% A; 108–109 min, 80–100% A. The column was
188successively washed for 15 min with methanol and equilibrated with 35% A for 10 min. Elution
189was performed at a flow rate of 0.8 ml/min with a splitting system of 2:8 to MS detector (160
190µl/min) and PDA detector (640 µl/min), respectively. The volume of the injected methanol solution
191was 20 µl. Analyses were performed with an ESI interface in the negative mode. The ionization
192parameters used were optimized as previously reported by Abdallah et al. (2017). N_2 was used as
193the sheath and auxiliary gas. PDA data were recorded with 200-600 nm range, with preferential
194channels 254, 280, and 325 nm as the detection wavelengths. The identification of compounds was
195performed comparing their HPLC retention times (t_R), ESI-MS data, and UV with authentic
196reference compounds (0.5 mg/ml methanol solution).
197
198*2.7 Statistical analyses*
199The three matrices obtained for the morphometric studies (S1.1; S1.2; S1.3 Tables), and those
200obtained for the flower colorimetric study (S2 Table), were subjected to multivariate Discriminant
201Analysis (DA), an identification optimization procedure based on the probability of identification
202using *a priori* classification (Peruzzi et al. 2015, Tundis et al. 2016, Roma-Marzio et al. 2017).
203Furthermore, each morphological character, as well as the relative R, G, and B contribution was
204also subjected to univariate analysis (non parametric Kruskal-Wallis test with Bonferroni correction
205for multiple comparisons). Only P values < 0.01 have been considered significant. Concerning
206phytochemical data, we built a single matrix (S3 Table) including both volatile and non-volatile
207compounds. For this purpose, since for non-volatile compounds only relative comparisons among
208the same chemicals in the different samples (obtained by measuring the peaks area) were available,

209we followed the same approach also for volatile compounds. We assigned a default value of 1 to
210that population where a certain compound was detected in the highest amount whereas, in the other
211populations, the relative amount of the same compound was scaled proportionally. To evaluate the
212phytochemical relationships among the five populations, the matrix was subjected to Principal
213Component Analysis (PCA). All the statistical analyses have been carried out by means of the
214PAST version 3.15 (Hammer et al. 2001; Hammer 2017) and R version 3.3.1 (R Core Team 2016)
215software.
216
217*2.8 Morphological investigation (updating the geographic distribution)*
218Based on the morphometric results, we selected those morphological characters showing less
219overlapping values among taxa, for identification purposes on herbarium material. Then, using
220these characters, we performed a morphological analysis on specimens preserved in the following
221herbaria: APP, FI, HLUC, PI, RO, SIENA, UTV (see S4). Finally, using QGIS 2.18 software, we
222georeferenced all the specimens in order to draw an updated distribution map of the three taxa.
223
224**3. Results**
225*3.1 Morphometric analyses*
226Results of univariate analysis of morphological characters are summarized in Table 3. The states of
227five characters (LW25, CL, StL, SL, and SW) showed significant differences among the three taxa
228($P < 0.01$). The states of eight characters (WL, WW, BL, LW50, LW75, CW, CmA, and CA)
229resulted significantly different between *P. flavescens* subsp. *maremmana* and the other two taxa,
230whereas *P. flavescens* subsp. *pisaurensis* showed significant differences from other taxa concerning
231the character-states of SL and EL ($P < 0.01$). No significant difference among the three taxa in BW
232and L was found. Among the statistically significant characters, those with less overlapping among
233the three taxa were WL, BL, CL, and EL (Figure 1).
234

Figure 1. Boxplots showing the morphological character-states less overlapping among the three subspecies of *Polygala flavescens*. A = Length of flower wing (WL); B = Length of flower bracteole (BL); C = Length of the capsule stipe (StL); D = Length of the elaiosome (EL).

Discriminant Analysis (DA) based on the three datasets of morphological characters, resulted respectively in 85.0% (dataset 1), 67.6% (dataset 2), and 89.0% (dataset 3) of jackknifed correct classification of individuals, *a priori* attributed to the three subspecies (Figure 2).

Figure 2. Discriminant Analysis based on quantitative continuous morphological characters of dataset 1 (A) and dataset 2 (B). PFF = *Polygala flavescens* subsp. *flavescens* (squares = PFF-A; circles = PFF-M; triangles = PFF-T), PFM = *P. flavescens* subsp. *maremmana*, PFP = *P. flavescens* subsp. *pisaurensis*. In blue, the relative contribution of each variable is reported.

Concerning the dataset 1, the characters most contributing to the discriminant function (loading values higher than |0.2|) are WL, BL, LW25, and SL. In particular, high values of WL and BL contribute to neatly separate *Polygala flavescens* subsp. *maremmana* from the other two taxa, while high values of LW25 and low values of SL contribute to separate *P. flavescens* subsp. *pisaurensis* from *P. flavescens* subsp. *flavescens*. A very small overlapping among *P. flavescens* subsp. *maremmana* and *P. flavescens* subsp. *flavescens* was outlined (1 out of 120 individuals not correctly classified), whereas a certain degree of possible confusion between *P. flavescens* subsp. *pisaurensis* and *P. flavescens* subsp. *flavescens* was found (11 out of 120 individuals not correctly classified). Concerning the dataset 2, the characters most contributing to the discriminant function (loading values higher than |0.5|) are CmA and CA. Contrarily to what obtained for dataset 1, a higher overlapping degree among individuals of the three taxa could be observed. In particular, *P. flavescens* subsp. *flavescens* can be confused with the other two subspecies (21 out of 150

individuals wrongly attributed to *P. flavescens* subsp. *maremmana*, and 38 out of 150 individuals wrongly attributed to *P. flavescens* subsp. *pisaurensis*).

Finally, based on the dataset 3, the character most contributing to the discriminant function (loading value |0.2|) is EL: high values of this character contribute to separate *P. flavescens* subsp. *pisaurensis* from the other two taxa. Based on dataset 3, we found a small overlap among *P. flavescens* subsp. *flavescens* and *P. flavescens* subsp. *maremmana* (7 out of 80 individuals not correctly classified) and among *P. flavescens* subsp. *flavescens* and *P.* subsp. *pisaurensis* (4 out 80 individuals not correctly classified), whereas no overlap was found among *P. flavescens* subsp. *maremmana* and *P. flavescens* subsp. *pisaurensis*.

3.2 Flower colorimetric analysis

Results of univariate analysis of relative contribution of Red, Green and Blue are summarized in Table 4. Our results highlighted that *P. flavescens* subsp. *maremmana* is characterized by wings and fringed keel with significantly higher contribution of Red and lower contribution of Blue, resulting in flowers more markedly yellow-orange, whereas in *P. flavescens* s.str. and *P. flavescens* subsp. *pisaurensis* the flowers range from yellow-whitish to yellow (Figure 3).

DA based on the colorimetric characters, resulted in 91% of jackknifed correct classification of individuals, *a priori* attributed to the three subspecies (Figure 4).

Accordingly, *P. flavescens* subsp. *maremmana* was clearly separated from both *P. flavescens* s.str. and *P. flavescens* subsp. *pisaurensis*. On the contrary, a certain overlap degree (7 out of 80 individuals not correctly classified) can be observed among the latter two taxa.

Figure 3. Colour of wings, fringes keel and flower tube of *P. flavescens* subsp. *flavescens*, *P. flavescens* subsp. *maremmana*, and *P. flavescens* subsp. *pisaurensis*. For each photo, a pie chart with the relative mean contribution of R (Red), G (Green), and B (Blue) is reported.

Figure 4. Discriminant Analysis based colorimetric characters. In blue, the relative contribution of each R (Red), G (Green), and B (Blue) variable for fringed keels, wings and tube flowers. PFF = *Polygala flavescens* subsp. *flavescens* (squares = PFF-A; circles = PFF-M; triangles = PFF-T), PFM = *P. flavescens* subsp. *maremmana*, PFP = *P. flavescens* subsp. *pisaurensis*. In blue, the relative contribution of each variable is reported.

3.3 Phytochemical analysis

Concerning VOCs, 58, 76, 75, and 89 compounds have been identified in leaves, flowers, fruits and seeds, respectively (Tables 5–8), representing from 88.7% to 99.9% of the total emission. A comparison of the volatile profiles among the three populations of *Polygala flavescens* subsp. *flavescens* revealed that non-terpene derivatives represent the main class of compounds emitted by leaves of all populations, ranging from 95.7% to 83.7% (Tables 5). In flowers (Table 6), apocarotenes are the most abundant class in PFF-T (64.0%), oxygenated monoterpenes in PFF-A (44.1%), and monoterpene hydrocarbons in PFF-M (82.4%). In fruits (Table 7) and seeds (Table 8), non-terpene derivates resulted the most abundant class of compound emitted by PFF-T (66.5%) and PFF-M (65.0%), whereas in PFF-A mostly oxygenated monoterpenes are emitted by fruits (64.9%), and monoterpenes hydrocarbons by seeds (40.8%). In PFF-A (*E*)-3-hexen-1-ol (compound 1) is the most abundant compound emitted by the leaves of all the investigated populations, ranging from 62.8% in PFF-A to 85.6% in PFF-M. In flowers, cis-α-ambrinol (compound 97) prevails in PFF-T (62.2%), limonene (compound 16) in PFF-A (21.7%), and myrcene (compound 12) in PFF-M (77.6%); in fruits, nonanal (compound 29) prevails in PFF-T (18.1%), carvone (compound 63) in PFF-A (43.8%) and decanal (compound 55) in PFF-M (22.5%), whereas the most abundant compounds emitted by seeds are nonanal (compound 29) in PFF-T (29.8%), PFF-M (12.0%) and α-pinene (compound 3) in PFF-A (31.4%).

Comparing the three subspecies, non-terpene derivatives emitted by leaves (Table 5) are the most abundant class of compound also in PFM and PFP. In *P. flavescens* subsp. *maremmana* (PFM),

sesquiterpene hydrocarbons (49.9%) prevail in flowers (Table 6), oxygenated monoterpenes (53.3%) in fruits (Table 7) and monoterpene hydrocarbons (73.9%) in seeds (Table 8), whereas in *P. flavescens* subsp. *pisaurensis* (PFP) the most abundant classes of compounds are monoterpene hydrocarbons (47.0%), sesquiterpene hydrocarbons (47.0%), and non-terpene derivatives (52.0 %) in flowers (Table 6), fruits (Table 7), and seeds (Table 8), respectively.

Some VOCs are unique to only one of the three subspecies. Particularly, the following compounds were detected only in *P. flavescens* subsp. *pisaurensis*: methyl 4-nonenoate (compound 58, 1.7%) (Table 5); octanoic acid (compound 46, 1.8%) and 6-methyltridecane (compound 77, 0.6%) (Table 6); β-chamigrene (compound 109, 1,5%) and *(Z)*-γ-bisabolene (compound 128, 1.4%) (Table 7); *(E)*-2-octenal (compound 21, 1.6%) and *cis*-thujopsene (compound 94, 5.2%) (Table 8); aromadendrene (compound 99, 13.5%, Table 7; 7%, Table 8). In *P. flavescens* subsp. *maremmana*, the following unique compounds are found: α-longipinene (compound 79, 1.3%) (Table 6); *cis*-dihydrocarvone (compound 50, 1.6%) and *trans*-dihydrocarvone (compound 54, 5.7%) (Table 7); β-pinene (compound 10, 1.2%), 6-camphenone (compound 26, 0.7%) and pinocarvone (compound 41, 1.1%) (Table 8). Only β-Elemene (compound 85) is unique to *P. flavescens* subsp. *flavescens* (Table 8). In addition to unique compounds, compared to the other two subspecies *P. flavescens* subsp. *maremmana* shows high levels of β-caryophyllene (compound 92, 31.1%) and α-pinene (compound 3, 70.8%) in flowers (Table 6) and seeds (Table 8), respectively.

Concerning non-volatile compounds, 75 different constituents have been detected by HPLC-photodiode array (PDA)/UV-electrospray ionization (ESI)-MS/MS (Figure 5 and Table 9).

Figure 5. Comparison of the LC-ESI-MS profiles in the sampled populations. PFF-T = Cerbaie Hills (Pisa, Tuscany); PFF-A = Torano (Rieti, Lazio); PFF-M = Vallerotonda (Frosinone, Lazio); PFM = Monte Argentario (Grosseto, Tuscany); PFP = Fano (Pesaro e Urbino, Marche). For peak characteristics, see Table 9.

Compounds 2, 3, 5–10, 12, 14, 17, 35, and 43 were identified by comparison with retention times, ESI-MS data, and UV of authentic reference compounds, previously isolated from *P. flavescens* subsp. *flavescens* (De Leo et al. 2017). Compound 2 ($[M\text{-}HCOO]^-$ at *m/z* 565), was the only apocarotenoid identified in all analysed extracts, characterized by an α-ionol aglycon and a saccharide portion constituted by a glucose and an apiose units, as deduced by product ions at *m/z* 403 and 271, respectively. As revealed by ESI-MS/MS, compounds 5 ($[M–H]^-$ at *m/z* 741), 10 ($[M–H]^-$ at *m/z* 947), 12 ($[M–H]^-$ at *m/z* 917), 14 ($[M–H]^-$ at *m/z* 887), and 17 ($[M–H]^-$ at *m/z* 845), were flavonol triglycosides having quercetin as aglycon, as deduced from the presence of the product ion at *m/z* 301. ESI-MS/MS of all flavonols showed product ions due to the subsequent losses of apiose ($[M–H–132]^-$) and rhamnose-glucose ($[M–H–132–146–162]^-$), leading to characterize the trisaccharide chain. In four compounds, the C-5 of the apiose unit is linked to an aromatic acid identified as synapic acid ($[M–H–206]^-$), ferulic acid ($[M\text{-}H\text{-}176]^-$), coumaric acid ($[M–H–146]^-$), and benzoic acid ($[M–H–122]^-$) for compounds 10, 12, 14, and 17, respectively. Complete names of detected flavonoids are supplied in Table 9. The product ions at *m/z* 463 and 301 generated by ESI-MS/MS experiment of compound 8 ($[M–H]^-$ at *m/z* 609), showed the presence of one glucose and one rhamnose moiety linked to a quercetin skeleton. Thus, compound 5 was identified as rutin. Full MS of compound 9 showed a deprotonated molecule $[M–H]^-$ at *m/z* 915 and product ions at *m/z* 709, 503, and 341 corresponding to the subsequent losses of two sinapoyl and one hexose moieties, according with the structure of the oligosaccharide reiniose F. Other oligosaccharides detected in all extracts had a different behaviour in the ESI-MS, due to the formation of anionized molecules $[M+HCOO]^-$ at *m/z* 653 (compound 3), 695 (compound 6), and 799 (compound 7). All oligosaccharides were identified by comparison with reference standards (Table 9).

In addition to previous metabolites, two bidesmodic triterpenoid saponins, compounds 35 ($[M–H]^-$ at *m/z* 1469) and 43 ($[M–H]^-$ at *m/z* 1411) were identified in all extracts, excluding *P. flascensces* subsp. *pisaurensis* lacking compound 43. The fragmentation pathways of both compounds, registered in negative mode, are in agreement with data reported by De Leo et al. (2017). The

injection of reference standards led to establish both chemical structures (Table 9), characterized by six sugar units linked to the aglycons presegenin (compound 35) and medicagenic acid (compound 43).

Other 62 compounds remained unidentified, although 37 were tentatively characterized as saponins, due to the observation of diagnostic ion fragments in the ESI-MS/MS spectra of the parent ions, such as sugar fragments and product ions due to the losses of sugar portions from the aglycon. ESI-MS/MS of compounds 24, 25, 27, 28, 31–34, 37–42, 44–48, 50, 53, 54, and 57 showed a very similar fragmentation pattern of compound 35, with a base peak due to the loss of $–CH_2OH$ unit from the aglycon presegenin and characteristic product ions at *m/z* 937, 747, 455, 439, and 423 due to sugar product ions. On the contrary, MS/MS experiments for compounds 49, 51, 58–62, 66, 68, and 70–72 are similar to that of compound 43 with a base peak generated by the scission of one or two sugar portions and sugar fragments such as *m/z* 585, 455, 439, and 423. The exact identification of detected triterpenoid saponins was not achieved for the lack of reference standards.

Totally, the class of 48 compounds has been identified, resulting in 37 saponins, 4 oligosaccharides, 6 flavonoids, and 1 apocarotenoid. No compound is unique to *P. flavescens* subsp. *flavescens*. Contrarily, the compounds corresponding to the peak 73 (unidentified) and to the peak 57 (a saponin) are unique to *P. flavescens* subsp. *maremmana* and to *P. flavescens* subsp. *pisaurensis*, respectively. Furthermore, the compounds corresponding to the peak 3 (β-D-(6-*O*-benzoyl)-fructofuranosyl-(2→1)-[β-D-glucopyranosyl-(1→3)]-α-D-glucopyranoside) and to the peak 40 (a saponin) are more abundant in *P. flavescens* subsp. *maremmana*, compared to the other two subspecies (Table 9). Twelve compounds, all identified as saponins (peaks 35, 37, 39, 44, 48, 50, 53, 54, 66, 68, 71, and 72) are more abundant in *P. flavescens* subsp. *pisaurensis*, with respect to other subspecies (Table 9). According to the PCA results (Figure 6; 80.5% of variance explained by the first three axes), both *P. flavescens* subsp. *maremmana* and *P. flavescens* subsp. *pisaurensis* fall outside the overall (both volatile and non-volatile) chemical variability of *P. flavescens* subsp. *flavescens*.

Figure 6. PCA 3D scatter plot based on phytochemical data (Component 1: 32.2%, Component 2: 24.4%, Component 3: 23.8% of the observed variance). Yellow bubbles: *Polygala flavescens* subsp. *flavescens* (PFF-A; PFF-M; PFF-T); orange bubble: *P. flavescens* subsp. *maremmana* (PFM); green bubble: *P. flavescens* subsp. *pisaurensis* (PFP).

3.4 Morphological investigation (updating the geographic distribution)

By measuring WL, BL, StL, and EL on herbarium specimens, we were able to revise their identification. Consequently, the distribution of the three subspecies of *P. flavescens* in Italy was updated (Figure 7). *Polygala flavescens* subsp. *flavescens* is confirmed to occur in peninsular Italy from Emilia-Romagna northwards, to Puglia and Basilicata southwards. *Polygala flavescens* subsp. *maremmana* occurs only in Mt. Argentario (Tuscany), whereas *P. flavescens* subsp. *pisaurensis* can be found along the east side of central Italy in Emilia-Romagna and Marche, but it has never been recorded to co-occur at the same sites with *P. flavescens* subsp. *flavescens*.

Identification key to the three subspecies

1a. Length of the elaiosome 2.53 (± 0.31) mm; flowers of plants in full blossom yellow-whitish.....*P. flavescens* subsp. *pisaurensis*

1b. Length of the elaiosome < 2.1 mm; flowers of plants in full blossom yellow to yellow-orange.....2

2a. Flowers in full blossom yellow, showing wings 8.22 ± 0.83 mm long, bracteoles 3.88 ± 0.49 mm long.....*P. flavescens* subsp. *flavescens*

2b. Flowers in full blossom yellow-orange, showing wings 10.86 ± 0.94 mm long, bracteoles 5.10 ± 0.46 mm long.....*P. flavescens* subsp. *maremmana*

Figure 7. Distribution of *Polygala flavescens* subsp. *flavescens* (in yellow), *P. flavescens* subsp. *maremmana* (in orange), and *P. flavescens* subsp. *pisaurensis* (in green), as resulted by the revision of herbarium specimens. The map is implemented by dots in light colours, which are derived from literature (Fenaroli 1970; Gubellini et al. 2014; Del Guacchio 2010).

4. Discussion

4.1 Integrative systematics

Our morphometric analyses confirm that most of the characters reported in the literature (Fiori et al. 1908; Zangheri 1976; Pignatti 1982; Arrigoni 2014) are useful to discriminate the three taxa. Particularly, *P. flavescens* subsp. *maremmana* shows bracteoles, flower wings and capsule stipes longer than the other two subspecies, whereas the length of the elaiosome is the most reliable morphological character to discriminate *P. flavescens* subsp. *pisaurensis*. However, concerning the length of flower wings, considered the most important character to identify *P. flavescens* subsp. *maremmana*, we never found wings shorter than 9.9 mm, contrarily to the range values reported by Arrigoni (2004) for this subspecies (9–11.5 mm). In addition, we did not find significant differences in the angle formed by the apex of flower wings, which was also considered as a character useful to discriminate *P. flavescens* subsp. *pisaurensis* (apex putatively obtuse) from the other two subspecies (apex putatively acute) (Arrigoni 2004).

A quantitative colorimetric approach to flowers was never tried before in *Polygala*. However, according to Arrigoni (2004), flower wings vary from yellow-greenish in *P. flavescens* subsp. *flavescens*, to pale-yellow in *P. flavescens* subsp. *pisaurensis*, whereas in *P. flavescens* subsp. *maremmana* wings are generally purplish-tinged. Although we noticed that purplish wings can be observed in all the three subspecies at fruiting stage, we confirm a differentiation in colour profiles at flowering stage. Indeed, when evaluated at the same phenological stage (full blossom and without fruits), flowers of *P. flavescens* subsp. *maremmana* show higher Red and lower Blue contributions,

resulting in flowers more markedly yellow-orange, whereas *P. flavescens* subsp. *pisaurensis* shows flowers ranging from yellow-whitish to yellow.

Our phytochemical study revealed the occurrence of unique compounds, both VOCs and non-volatiles, that can be considered as molecular markers useful to discriminate the three subspecies of *Polygala flavescens*. This phytochemical differentiation is also confirmed by the results of multivariate analysis based on the overall phytochemical screening, showing a clear separation of the three subspecies, and pointing out the importance of phytochemical studies in plant systematics (Stuessy 2009; Astuti et al. 2017; and literature cited therein).

As far the geographic distribution is concerned, after herbarium revision we confirmed that the range of *P. flavescens* subsp. *flavescens* goes from Emilia-Romagna, in Northern Italy, to Apulia and Basilicata, in Southern Italy. On the other hand, *P. flavescens* subsp. *pisaurensis* is restricted to the east coast in Marche and Emilia-Romagna, while *P. flavescens* subsp. *maremmana* is limited to Mt. Argentario, despite Arrigoni (2004) quoted this subspecies for a larger area, i.e. from the southern part of the province of Leghorn (San Vincenzo) to the southern part of the province of Grosseto (Capalbio). In our opinion, this discrepancy is putatively due to differences in the phenological stage of the flowers measured by previous authors. Indeed, we noticed a marked elongation of wings from flower stage to fruit.

Considering that the three taxa show significant morphological and phytochemical differences, they are allopatric, and share the same chromosome number (Peruzzi et al. 2017), we deem appropriate their taxonomic treatment at subspecific level.

4.2 Possible ecological role of phytochemical and morpho-colorimetric variation

Among the unique compounds, the occurrence of *(E)*-2-octenal (compound 21) in seeds of *P. flavescens* subsp. *pisaurensis* could be discussed in the light of myrmecochory, the ant-mediated seed dispersal mechanism, whose occurrence in *P. flavescens* is suggested by the occurrence of elaiosomes. Since seed VOCs elicit ant-carrying behaviour of elaiosomes (Brew et al.1989;

Youngsteadt et al. 2008), the occurrence of unique compounds may deserve further studies concerning the ecological role of VOCs, particularly in plant-ant interactions (Willmer 2009). Incidentally, *(E)*-2-octenal was also identified in gland abdominal extract of two ant species (*Eurydema ventrale* and *E. oleraceum*) collected in Central Italy, suggesting that this compound plays a pheromonal role in ant communication (Aldrich et al. 2017). Interestingly, *P. flavescens* subsp. *pisaurensis* is also the subspecies showing the largest elaiosomes.

The ecological role of VOCs is also particularly relevant in floral scent that, in synergy with floral colour and shape, can act as signals for attraction of pollinators (Schiestl et al. 2013). Typically, floral scent is determined by volatile compounds and represents an important mode of communication among flowering plants, pollinators, and enemies (Knudsen et al. 2006; Raguso 2008). It has been observed that emissions rich in benzenoids or in linalool (and its oxides) seem to be an adaptation to butterflies or to generalist pollinators (Andersson et al. 2002). On the other hand, when the floral bouquet is dominated by a sole volatile in relatively large percentages, the pollination is often bee-mediated (Borg-Karlson et al 1996). The latter situation is experienced for the studied taxa, which emitted 1-2 main compounds in their floral bouquet. In particular, PFF-T mainly emitted *cis*-α-ambrinol (compound 97, 62.2%), while PFF-M and PFP emitted myrcene (compound 12) as their main volatile (77.6 and 46.6%, respectively). *Polygala flavescens* subsp. *maremmana* and the remaining population of *P. flavescens* subsp. *flavescens* (PFF-A) have two main volatiles in their flower emission: limonene (compound 16, 21.7%) and α-terpineol (compound 49, 20.5%) for PFF-A and β-caryophyllene (compound 92, 31.1%) and 1,8-cineole (compound 17, 15.3%) in the case of PFM. The hypothesis of bee-attraction is also in good agreement with the morphological requirements for such pollination (Faegry and van der Pijl 1979; Westerkamp 1997). Also the changes in flower colour quantified in this study, paralleled by change in floral scent, could reflect a change in pollinators, possibly leading to reproductive isolation of the three subspecies, as demonstrated for example in the two closely related species *Mimulus verbenaceus* Greene and *M. cardinalis* Douglas ex Benth. (Phrymaceae) (Vickery 1992). In

addition, genes controlling the flower colour might influence plant resistance to herbivory (Irwin et al. 2003), causing a synergism that may have a positive effect on reproductive fitness.

The flavonoid biosynthetic pathway, culminating in the production of anthocyanins, with carotenoids and betalains, are the main pigments responsible for the flower colour (Weiss 1995; Irwin et al. 2003; Borghi et al. 2017). In our study, we failed to find relevant differences in the overall flavonoid composition among the three subspecies, but further studies aimed to investigate specifically the non-volatile compounds occurring in flowers, as well as gene expression and biosynthetic pathways, could clarify the phytochemical basis of the documented differences in flower colour among the three taxa.

4.3 Potential pharmacological implications of our study

Besides their systematic value, our results concerning non-volatile compounds confirm that *Polygala* is a genus rich in flavonoids, saponins, and oligosaccharides (Clegg and Durbin 2000), supporting a pharmacological potential for all the three subspecies within *P. flavescens*. For example, due to the expectorant and anti-inflammatory effects, vaccine adjuvants or neurotrophic activity of the saponins (Klein et al. 2012 and literature therein), it is noteworthy that in *P. flavescens* subsp. *pisaurensis* we highlighted a unique saponin, and that further 12 saponins were more abundant than in other taxa. In addition, the two oligosaccharides (3,6'-di-*O*-sinapoylsucrose and reiniose F) isolated by De Leo et al.(2017) in *P. flavescens* subsp. *flavescens* as potential hLDH5 inhibitors, were also found in the other two subspecies, confirming their interest for the potential development of new anticancer agents (De Leo et al. 2017).

Table 1. Sampled localities used for the phytochemical, morphometric and colorimetric investigations.

Taxon	Acronym	Locality	Coordinate (WGS84)	Altitude m a.s.l.	Voucher numbers
P. flavescens subsp. *flavescens*	PFF-T	Cerbaie Hills (Pisa, Tuscany)	43.751228 N, 10.719234 E	55	PI n. 000455–58
P. flavescens subsp. *flavescens*	PFF-A	Torano (Rieti, Lazio)	42.157098 N, 13.270760 E	760	PI n. 000453–54
P. flavescens subsp. *flavescens*	PFF-M	Vallerotonda (Frosinone, Lazio)	41.588942 N, 14.007237 E	765	PI n. 000459–60
P. flavescens subsp. *maremmana*	PFM	Monte Argentario (Grosseto, Tuscany)	42.421952 N, 11.140779 E	130	PI n. 000466–69
P. flavescens subsp. *pisaurensis*	PFP	Fano (Pesaro e Urbino, Marche)	43.864231 N, 12.984113 E	25	PI n. 000461–62

519**Table 2.** Measured morphological characters.

Part of the plant	ID	Character
Dataset 1		
Flower	WL	Length of flower wings (mm)
Flower	WW	Width of flower wings (mm)
Flower	WA	Angle formed by the apex of flower wings (rad)
Flower	BL	Length of flower bracteole (mm)
Flower	BW	Width of flower bracteole (mm)
Leaf	LL	Leaf length (cm)
Leaf	LW25	Leaf width on 25% of leaf's length from base up (mm)
Leaf	LW50	Leaf width on 50% of leaf's length from base up (mm)
Leaf	LW75	Leaf width on 75% of leaf's length from base up (mm)
Stem	SL	Stem length (dm)
Dataset 2		
Fruit	CL	Capsule length (mm)
Fruit	CW	Capsule width (mm)
Fruit	StL	Length of the capsule stipe (mm)
Fruit	CmA	Area of the capsule membranous marginal part (mm^2)
Fruit	CA	Capsule area (mm^2)
Dataset 3		
Seed	SL	Seed length (mm)
Seed	SW	Seed width (mm)
Seed	EL	Length of the elaiosome (mm)

520

521

522

523

524

525

526

527

528

529

530

531

532

Table 3. Comparison of morphological features among the three subspecies of *Polygala flavescens*. Quantitative numerical values are expressed as mean ± SD. Character-states marked by different superscript letters are significantly different ($P < 0.01$). Characters in bold are also shown in Figure 1. For the meaning of the character acronyms, see Table 2.

ID character	***P. flavescens* subsp. *flavescens***	***P. flavescens* subsp. *maremmana***	***P. flavescens* subsp. *pisaurensis***
Dataset 1			
WL (mm)	**8.22 ± 0.83^{a}**	**10.86 ± 0.94^{b}**	**8.12 ± 0.70^{a}**
WW (mm)	3.07 ± 0.40^{a}	3.49 ± 0.51^{b}	3.31 ± 0.27^{a}
WA (rad)	1.17 ± 0.21^{a}	0.90 ± 0.12^{b}	1.08 ± 0.12^{b}
BL (mm)	**3.88 ± 0.49^{a}**	**5.10 ± 0.46^{b}**	**4.13 ± 0.43^{a}**
BW (mm)	0.48 ± 0.08^{a}	0.47 ± 0.00^{a}	0.55 ± 0.16^{a}
LL (cm)	2.23 ± 0.43^{a}	2.31 ± 0.39^{a}	2.17 ± 0.34^{a}
LW25 (mm)	1.91 ± 0.49^{a}	1.33 ± 0.42^{b}	2.38 ± 0.51^{c}
LW50 (mm)	2.31 ± 0.57^{a}	1.60 ± 0.48^{b}	2.57 ± 0.53^{a}
LW75 (mm)	1.82 ± 0.41^{a}	1.39 ± 0.36^{b}	1.93 ± 0.45^{a}
SL (dm)	2.00 ± 0.55^{a}	2.04 ± 0.46^{a}	1.52 ± 0.21^{b}
Dataset 2			
CL (mm)	5.55 ± 0.65^{a}	6.63 ± 0.66^{b}	5.21 ± 0.48^{c}
CW (mm)	4.09 ± 0.56^{a}	4.56 ± 0.54^{b}	4.15 ± 0.47^{a}
StL (mm)	**1.11 ± 0.20^{a}**	**1.52 ± 0.26^{b}**	**0.89 ± 0.17^{c}**
CmA (mm^2)	5.58 ± 1.78^{a}	7.04 ± 1.73^{b}	5.74 ± 1.30^{a}
CA (mm^2)	16.97 ± 4.36^{a}	21.16 ± 4.15^{b}	16.83 ± 3.44^{a}
Dataset 3			
SL (mm)	2.62 ± 0.19^{a}	3.13 ± 0.17^{b}	2.43 ± 0.14^{c}
SW (mm)	1.33 ± 0.09^{a}	1.48 ± 0.09^{b}	1.24 ± 0.12^{c}
EL (mm)	**1.83 ± 0.24^{a}**	**1.89 ± 0.16^{a}**	**2.53 ± 0.31^{b}**

Table 4. Comparison of colorimetric features among the three subspecies of *Polygala flavescens*. RGB (Red, Green, and Blue) values, expressed as mean ± SD, represent the coordinates in the RGB colour-space, where each value ranges from 0 to 255. Character states marked by different superscript letters are significantly different ($P < 0.01$).

ID character	*P. flavescens* subsp. *flavescens*	*P. flavescens* subsp. *maremmana*	*P. flavescens* subsp. *pisaurensis*
Wings			
R	126 ± 9^{a}	144 ± 9^{b}	128 ± 9^{a}
G	114 ± 9^{a}	116 ± 8^{a}	119 ± 9^{a}
B	64 ± 13^{a}	32 ± 4^{b}	74 ± 11^{c}
Fringes			
R	153 ± 14^{a}	172 ± 19^{b}	162 ± 20^{a}
G	126 ± 13^{a}	123 ± 14^{a}	145 ± 17^{b}
B	38 ± 14^{a}	22 ± 3^{b}	71 ± 17^{c}
Tube			
R	138 ± 10^{a}	160 ± 13^{b}	152 ± 13^{b}
G	127 ± 10^{a}	138 ± 11^{b}	140 ± 13^{b}
B	63 ± 9^{a}	54 ± 11^{b}	82 ± 14^{c}

Table 5. Comparison of the chemical composition (% of VOCs) emitted by the leaves of sampled populations. PFF-T = Cerbaie Hills (Pisa, Tuscany); PFF-A = Torano (Rieti, Lazio); PFF-M = Vallerotonda (Frosinone, Lazio); PFM = Monte Argentario (Grosseto, Tuscany); PFP = Fano (Pesaro e Urbino, Marche). ap: apocarotenes, mh: monoterpene hydrocarbons, ntp: non-terpene derivatives, om: oxygenated monoterpenes, os: oxygenated sesquiterpenes, ph: phenylpropanoids, sh: sesquiterpene hydrocarbons. Yellow columns = *P. flavescens* subsp. *flavescens*, orange column = *P. flavescens* subsp. *maremmana*, green column = *P. flavescens* subsp. *pisaurensis*.

ID compound	Compound	Class	LRI	PFF-T	PFF-A	PFF-M	PFM	PFP
1	(*E*)-3-Hexen-1-ol	ntp	850	83.1	62.8	85.6	70.1	43.9
6	Benzaldehyde	ntp	962	1.1	0.2	0	tr	0
7	1-Ethyl-3-methylbenzene	ntp	967	0	0	0.2	0	0
8	1-Octen-3-ol	ntp	976	0.2	0	0	0	0
11	6-Methyl-5-hepten-2-one	ntp	986	0	0.3	0	0	0
14	*(Z)*-3- Hexenyl acetate	ntp	1007	3.0	2.4	4.1	0	1.3
17	1,8-Cineole	om	1034	0	0.2	0	0	2.9
29	Nonanal	ntp	1104	0.8	0.5	0.5	1.0	1.9
38	Isobutyl hexanoate	ntp	1150	0	0	0	0	2.0
44	Menthol	om	1174	0	0.2	0	tr	0
45	4-terpineol	om	1179	0	0	0	0	3.3
47	(*Z*)-3- Hexenyl butyrate	ntp	1188	0	0.5	0	0	1.7
48	Methyl salicylate	ntp	1191	3.9	3.2	3.7	13.1	8.7
55	Decanal	ntp	1206	1.7	2.9	0.6	2.6	2.3
58	Methyl 4-nonenoate	ntp	1216	0	0	0	0	1.7
59	β-Cyclocitral	ap	1217	0	0	0	0.2	0
61	(*Z*)-3-Hexenyl isovalerate	ntp	1238	0	0	0	0	1.4
65	Ethyl salicylate	ntp	1267	0	0	0	0.2	0
67	Citronellyl formate	om	1275	0	0.1	0	0.3	0
70	Isobornyl acetate	om	1285	0.4	0	0	0.3	0
74	Undecanal	ntp	1306	0.3	0.2	0.1	0.3	0.6
80	Eugenol	ph	1358	0	0.9	0.4	0	1.5
85	β-Elemene	sh	1391	0	0.2	0	0	0
86	*(E)*-Jasmone	ntp	1392	0	0	0	0.5	0
88	*n*-Tetradecane	ntp	1399	0.1	0.2	0.2	0.4	1.8
89	Methyl eugenol	ph	1401	0	1.4	0	0	0
91	Dodecanal	ntp	1407	0.3	0	0.1	0.3	0.8
92	β-Caryophyllene	sh	1419	0	tr	0	0.2	0.6

95	β-Gurjunene	sh	1432	0	0	0	0.7	0
98	*trans*-α-Bergamotene	sh	1439	0	2.0	0	0	0
103	(*E*)-Geranylacetone	ap	1453	1.4	1.9	0.7	0.8	3.2
106	(*E*)-β-Farnesene	sh	1459	0	1.6	0	0	0
107	*cis*-Muurola-4(14),5-diene	sh	1460	0	0.2	0	0	0
111	Germacrene D	sh	1482	0	0.3	0	0	0
116	*n*-Pentadecane	ntp	1500	0	0.2	0	0.9	0
123	Tridecanal	ntp	1509	0.3	0	0	0	0
125	*trans*-γ-Cadinene	sh	1514	0	0.2	0	0	0
126	Geranyl isobutyrate	mh	1515	0.3	0	0	0	0
129	Benzoic acid, 4-ethoxyethyl ester	ntp	1522	0.5	0	0	0	3.8
133	(*Z*)-3-Hexenyl benzoate	ntp	1570	0	0	0.2	0	4.6
135	Caryophyllene oxide	os	1582	0	0	0	0	1.6
138	*n*-Hexadecane	ntp	1600	0	0.9	0	0	0
139	Tetradecanal	ntp	1613	0.2	0	0	0	0
140	1,10-di-*epi*-cubenol	os	1614	0	0.2	0	0	0
143	α-Muurolol	os	1645	0	2.5	0	0	0
144	β-Eudesmol	os	1649	0	0.4	0	0	0
147	*n*-Heptadecane	ntp	1700	0.2	tr	0.2	0	0
148	Benzyl benzoate	ntp	1762	0	tr	0	0	0
149	*n*-Octadecane	ntp	1800	0	0.8	0	0	0
150	2-Ethylhexyl salicylate	ntp	1807	0	7.6	0	0.3	0
152	Isopropyl tetradecanoate	ntp	1830	0	1.0	0	0	0
-	**Apocarotenes**	**ap**		**1.4**	**1.9**	**0.7**	**1.0**	**3.2**
-	**Monoterpene hydrocarbons**	**mh**		**0.3**	**0.0**	**0.0**	**0.0**	**0.0**
-	**Non-terpene derivatives**	**ntp**		**95.7**	**83.7**	**95.5**	**89.7**	**77**
-	**Oxygenated monoterpenes**	**om**		**0.4**	**0.5**	**0.0**	**0.6**	**6.2**
-	**Oxygenated sesquiterpenes**	**os**		**0.0**	**3.1**	**0.0**	**0.0**	**1.6**
-	**Phenylpropanoids**	**ph**		**0.0**	**2.3**	**0.4**	**0.0**	**1.5**
-	**Sesquiterpene hydrocarbons**	**sh**		**0.0**	**4.5**	**0.0**	**0.9**	**0.6**
-	**Total**			**97.8**	**96.0**	**96.6**	**92.2**	**90.1**

Table 6. Comparison of the chemical composition (% of VOCs) emitted by the flowers of sampled populations. PFF-T = Cerbaie Hills (Pisa, Tuscany); PFF-A = Torano (Rieti, Lazio); PFF-M = Vallerotonda (Frosinone, Lazio); PFM = Monte Argentario (Grosseto, Tuscany); PFP = Fano (Pesaro e Urbino, Marche). Ap: apocarotenes, mh: monoterpene hydrocarbons, ntp: non-terpene derivatives, om: oxygenated monoterpenes, os: oxygenated sesquiterpenes, sh: sesquiterpene hydrocarbons. Yellow columns = *P. flavescens* subsp. *flavescens*, orange column = *P. flavescens* subsp. *maremmana*, green column = *P. flavescens* subsp. *pisaurensis*.

ID compound	Compound	Class	LRI	PFF-T	PFF-A	PFF-M	PFM	PFP
3	α-Pinene	mh	941	0	6.0	0.3	0	0
4	Camphene	mh	955	0	1.1	tr	0	0
9	Sabinene	mh	977	0	0.8	tr	0	0
12	Myrcene	mh	993	0	4.6	77.6	0.7	46.6
16	Limonene	mh	1032	0	21.7	4.1	0	0
17	1,8-Cineole	om	1034	0	0	0	15.3	3.3
23	*cis*-Sabinene hydrate	om	1070	0	0.9	0	0	0
24	Fenchone	om	1088	0	0	0	0	1.0
25	Terpinolene	mh	1090	0	1.2	0.4	0	0
28	Linalool	om	1101	0	8.3	2.8	0	1.1
29	Nonanal	ntp	1104	2.3	0	0	6.4	2.6
30	(*Z*)-2-Undecene	ntp	1114	0	2.8	0	0	0
32	Nerol	om	1127	0	1.7	0	0	0
38	Isobutyl hexanoate	ntp	1150	0	0	0	0.5	0
35	Camphor	om	1145	0	0.6	0	0	0.4
45	4-Terpineol	om	1179	0	0.5	tr	0	0.5
46	Octanoic acid	ntp	1180	0	0	0	0	1.8
48	Methyl salicylate	ntp	1191	2.7	0	0	4	0
49	α-Terpineol	om	1192	0	20.5	0.9	0	0.7
55	Decanal	ntp	1206	1.9	3.3	0.5	6.1	2.9
61	(*Z*)-3-Hexenyl isovalerate	ntp	1238	0	0.5	0	0	0
63	Carvone	om	1244	0	0.9	0	0	0
64	(*E*)-2-Decenal	ntp	1263	0	tr	0	0.3	0.3
68	Methyl nerolate	om	1282	4.1	9	5.5	0	2.4
70	Isobornyl acetate	om	1285	0	0	0	0	3.3
72	*n*-Tridecane	ntp	1300	0	tr	0	0	6.2
74	Undecanal	ntp	1306	0	0.5	tr	1.0	0.6
76	Methyl geranate	om	1325	0	1.7	2.4	0	1.7
77	6-Methyltridecane	ntp	1346	0	0	0	0	0.6
79	α-Longipinene	sh	1352	0	0	0	1.3	0

82	α-Copaene	sh	1377	0	0	0.1	1.7	0.4
83	β-Bourbonene	sh	1385	0	0	0	0.6	0
84	β-Cubebene	sh	1391	0.8	tr	0	3.3	tr
87	(*Z*)-Jasmone	ntp	1395	0	0.9	0	0.5	0
88	*n*-Tetradecane	ntp	1400	0.9	0	0	0	0.5
91	Dodecanal	ntp	1407	0	0.7	0.2	1.3	1.3
92	β-Caryophyllene	sh	1419	8.4	2	0.6	31.1	6.1
96	γ-Elemene	sh	1434	0	0	0	2.4	0.2
97	*cis*-α-ambrinol	ap	1437	62.2	tr	0	1.9	0.2
102	5-Methyltetradecane	ntp	1452	0	1.3	0	0	0
103	(*E*)-Geranylacetone	ap	1456	1.8	5.3	1.3	4.6	2
104	α-Humulene	sh	1455	0.7	0	0.3	2.1	0
105	4-Methyltetradecane	ntp	1457	1.9	0	0	0	0
106	(*E*)-β-Farnesene	sh	1459	1.3	0	1.1	0	1.5
108	2-Methyltetradecane	ntp	1462	0	0	0	1.7	0.7
110	γ-Muurolene	sh	1478	0	0.5	0	6.3	0
111	Germacrene D	sh	1482	0.9	0	0.1	0	0.8
112	(*E*)-β-Ionone	ap	1486	0	0	0	0.9	0
113	β-Selinene	sh	1487	0	0	0	0	0.8
114	*cis*-β-Guaiene	sh	1491	0	0	0.2	0	0
117	Pentadecane	ntp	1500	0	0	0.3	0	0
118	δ-Decalactone	ntp	1501	0	0	0	0	0.8
121	(*E*,*E*)-α-Farnesene	sh	1508	0	0	0	0.6	0.4
123	Tridecanal	ntp	1510	0	0	0	0.5	0
125	*trans*-γ-Cadinene	sh	1514	0	0	0	0	0.7
127	10-*Epi*-italicene ether	os	1516	0.7	0	0	0	0
129	Benzoic acid 4-ethoxyethyl ester	ntp	1522	0	1.2	0	0	0
130	δ-Cadinene	sh	1524	0	0	0	0.5	0.8
133	(*Z*)-3-Hexenyl benzoate	ntp	1570	0	tr	0	0	0
134	Dendrolasin	os	1580	2.0	0	0	0	0
135	Caryophyllene oxide	os	1582	2.0	0	0	0.4	0
139	Tetradecanal	ntp	1613	0	tr	0	0.5	0.4
145	*Cis*-Methyl dihydrojasmonate	ntp	1655	0	0	0	0	0.5
147	*n*-Heptadecane	ntp	1700	0	0.4	0	0	0.7
149	*n*-Octadecane	ntp	1800	0	0	0	0	0.5
150	2-Ethylhexyl salicylate	ntp	1807	0	0	0	0	0.7
152	Isopropyl tetradecanoate	ntp	1830	0	0	0	0	0.8
-	**Apocarotenes**	**ap**		**64.0**	**5.3**	**1.3**	**7.4**	**2.2**
-	**Monoterpene hydrocarbons**	**mh**		**0.0**	**35.4**	**82.4**	**0.7**	**47.0**
-	**Non-terpene derivatives**	**ntp**		**9.7**	**11.6**	**1.0**	**22.8**	**22.0**
-	**Oxygenated monoterpenes**	**om**		**4.1**	**44.1**	**11.6**	**15.3**	**14.0**

-	**Oxygenated sesquiterpenes**	**os**	**4.7**	**0.0**	**0.0**	**0.4**	**0.0**
-	**Phenylpropanoids**	**ph**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**
-	**Sesquiterpene hydrocarbons**	**sh**	**12.1**	**2.5**	**2.4**	**49.9**	**12.0**
-	**Total**		**94.6**	**98.9**	**98.7**	**96.5**	**97.2**

578

579

580**Table 7.** Comparison of the chemical composition (% of VOCs) emitted by the fruits of sampled
581populations. PFF-T = Cerbaie Hills (Pisa, Tuscany); PFF-A = Torano (Rieti, Lazio); PFF-M =
582Vallerotonda (Frosinone, Lazio); PFM = Monte Argentario (Grosseto, Tuscany); PFP = Fano
583(Pesaro e Urbino, Marche). Ap: apocarotenes, ntp: non-terpene derivatives, om: oxygenated
584monoterpenes, os: oxygenated sesquiterpenes, ph: phenylpropanoids, sh: sesquiterpene
585hydrocarbons. Yellow columns = *P. flavescens* subsp. *flavescens*, orange column = *P. flavescens*
586subsp. *maremmana*, green column = *P. flavescens* subsp. *pisaurensis*.

ID Compound	Compound	Class	LRI	PFF-T	PFF-A	PFF-M	PFM	PFP
15	1-Hexyl acetate	ntp	1010	0	0	3.5	0	tr
17	1,8-Cineole	om	1034	5.7	0.7	0	tr	0
19	Benzyl alcohol	ntp	1044	5.2	0	tr	tr	0
24	Fenchone	om	1088	3.9	tr	0	0	0
29	Nonanal	ntp	1104	18.1	0	12.2	3.5	15.2
35	Camphor	om	1145	1.1	0.3	0.9	tr	0
36	Ethyl 2-heptenoate	ntp	1146	0	0	0	0	0.5
40	(*E*)-2-nonenal	ntp	1163	1.1	0.2	0	tr	0.9
43	*Neo*-menthol	om	1167	0	0.2	1.2	0	0
44	Menthol	om	1174	0	0	0	0.6	0
45	4-Terpineol	om	1179	4	11.5	0	18.2	0
48	Methyl salicylate	ntp	1191	0	0	3.8	0	0
49	α-Terpineol	om	1192	4.7	7.4	0	3.7	0
50	*cis*-Dihydrocarvone	om	1194	0	0	0	1.6	0
52	Dihydrocitronellol	om	1196	0	0.6	0	0	0
54	*trans*-Dihydrocarvone	om	1204	0	0	0	5.7	0
55	Decanal	ntp	1206	15.5	8.8	22.5	12.7	18.6
57	1-Octyl acetate	ntp	1213	0	0	0.7	0	0.5
60	*cis-p*-Mentha 1(7)-8-dien-2-ol	om	1231	0	0	0	0	0.6
62	Cumin aldehyde	om	1241	0	0.2	0	0	0
63	Carvone	om	1244	0	43.8	0	23.5	0

57

58

63	(*E*)-2-decenal	ntp	1263	0.7	0.5	0.3	0.5	0
67	Citronellyl formate	om	1275	0	0	1.5	0	0
69	(*E*)-Anethole	ph	1284	0	11.1	0	12.5	0
71	10-Undecenal	ntp	1299	0	0	1.2	0	0
72	*n*-Tridecane	ntp	1300	1.7	0.2	0	tr	0
73	Carvacrol	om	1301	0	0.2	0	0	0
74	Undecanal	ntp	1306	2.3	0.6	3.2	tr	1.7
83	β-Bourbonene	sh	1385	0	0	0	0	9.8
88	*n*-Tetradecane	ntp	1400	2.3	tr	1.5	tr	0
91	Dodecanal	ntp	1407	5.5	0.5	4.3	1.1	0
92	β-Caryophyllene	sh	1419	tr	0	4.4	tr	4.6
99	Aromadendrene	sh	1440	0	0	0	0	13.5
101	α-Himachalene	sh	1450	0	0	1.4	0.6	0
103	(*E*)-Geranylacetone	ap	1453	4.1	2.3	21.9	3	4.8
106	(*E*)-β-Farnesene	sh	1459	0	0	0	0	2.4
109	β-Chamigrene	sh	1476	0	0	0	0	1.5
111	Germacrene D	sh	1482	0	0	0	0	1.2
112	(*E*)-β-Ionone	ap	1486	2.6	0.2	0	tr	1
116	*n*-Pentadecane	ntp	1500	3.4	0	0	0	0
118	δ-Decalactone	ntp	1501	0	0	2.4	0	0
122	β-Bisabolene	sh	1508	0	0	0	0	4.7
123	Tridecanal	ntp	1509	1.7	tr	0	tr	0
124	α-Alaskene	sh	1511	0	0	0	tr	7.7
126	Geranyl isobutyrate	om	1515	0	0	0	0	1.5
128	(*Z*)-γ-Bisabolene	sh	1517	0	0	0	0	1.4
132	Dihydroactinidiolide	ap	1536	4.8	0.6	0	0	3.5
135	Caryophyllene oxide	os	1582	0	6.3	0	8.7	0
136	Viridiflorol	os	1591	0	0.4	0	0	0
137	Carotol	os	1595	0	0	0	0.6	0
138	*n*-Hexadecane	ntp	1600	0	0	6.6	0	0.5
139	Tetradecanal	ntp	1613	1.7	0	tr	0	0.9
142	Selina-3,11-dien-6-α-ol	os	1644	0	0.7	0	0	0
146	Cadalene	os	1675	0	0	0	1.1	0
147	*n*-Heptadecane	ntp	1700	2.7	0.2	1.6	0.8	1.2
149	*n*-Octadecane	ntp	1800	2.4	0.1	1.2	tr	1.6
152	Isopropyl tetradecanoate	ntp	1830	2.2	0	0	0	0
153	(*E*,*E*)-α-Farnesyl acetate	os	1843	0	0.4	0	1.1	0
-	**Apocarotenes**	**ap**		**11.5**	**3.1**	**21.9**	**3.0**	**9.3**
-	**Non-terpene derivatives**	**ntp**		**66.5**	**11.1**	**65.0**	**18.6**	**42.0**
-	**Oxygenated monoterpenes**	**om**		**19.4**	**64.9**	**3.6**	**53.3**	**2.1**
-	**Oxygenated sesquiterpenes**	**os**		**0.0**	**7.8**	**0.0**	**12.5**	**0.0**

-	**Phenylpropanoids**	**ph**	**0.0**	**11.1**	**0.0**	**12.5**	**0.0**
-	**Sesquiterpene hydrocarbons**	**sh**	**0.0**	**0.0**	**5.8**	**0.6**	**47.0**
-	**Total**		**97.4**	**98.0**	**96.3**	**100**	**100**

587

588**Table 8.** Comparison of the chemical composition (% of VOCs) emitted by the seeds of sampled
589populations. PFF-T = Cerbaie Hills (Pisa, Tuscany); PFF-A = Torano (Rieti, Lazio); PFF-M =
590Vallerotonda (Frosinone, Lazio); PFM = Monte Argentario (Grosseto, Tuscany); PFP = Fano
591(Pesaro e Urbino, Marche). Ap: apocarotenes, mh: monoterpene hydrocarbons, ntp: non-terpene
592derivatives, om: oxygenated monoterpenes, os: oxygenated sesquiterpenes, ph: phenylpropanoids,
593sh: sesquiterpene hydrocarbons. Yellow columns = *P. flavescens* subsp. *flavescens*, orange column
594= *P. flavescens* subsp. *maremmana*, green column = *P. flavescens* subsp. *pisaurensis*.

ID Compound	Compound	Class	LRI	PFF-T	PFF-A	PFF-M	PFM	PFP
2	Heptanal	ntp	901	0	0	0	0	2.1
3	α-Pinene	mh	941	19.1	31.4	1.0	70.8	0
4	Camphene	mh	955	0	0	0	0.9	0
5	Thuja-2,4(10)-diene	mh	959	0.8	0	0	0	0
9	Sabinene	mh	977	0	0.4	0	0	0
10	β-Pinene	mh	982	0	0	0	1.2	0
13	Octanal	ntp	1002	0	0	0	0	11.1
15	1-Hexyl acetate	ntp	1010	0	0	tr	0	0.2
16	Limonene	mh	1032	0	0	0	1.0	0
18	3-Octen-2-one	ntp	1043	0	0	0	0	0.8
20	(*E*)-β-Ocimene	mh	1052	0	9	0	0	0
21	(*E*)-2-Octenal	ntp	1062	0	0	0	0	1.6
22	γ-Terpinene	mh	1063	0	0	1.2	0	0
23	*cis*-Sabinene hydrate	om	1070	0	0	3.7	0	0
26	6-Camphenone	om	1091	0	0	0	0.7	0
27	*trans*-Sabinene hydrate	om	1099	0	0	3.6	0	0
28	Linalool	om	1101	0	0	0	0	0.6
29	Nonanal	ntp	1104	29.8	11.5	12	0.4	25
31	*trans-p*-Mentha-2,8-dien-1-ol	om	1126	0	0	0	1.2	0
33	α-Campholenal	om	1127	1.8	0.9	0	1.7	0
34	*cis*-Verbenol	om	1144	0.8	0	0	0	0
35	Camphor	om	1145	1.1	0	0	1.8	0
37	*trans*-Verbenol	om	1147	0.2	0	0	3.5	0
39	*trans*-Pinocamphone	om	1162	0	0	0	0.4	0
41	Pinocarvone	om	1164	0	0	0	1.1	0
42	Benzyl acetate	ntp	1165	0	0	4.4	0	0
48	Methyl salicylate	ntp	1191	0	0	1.3	0	0
51	Myrtenal	om	1194	1.1	2.0	0	0.7	0
53	*n*-Dodecane	ntp	1200	0.3	0	0	0	0.2
55	Decanal	ntp	1206	6.1	2.4	7.7	0.4	4.3

56	Verbenone	om	1207	7.6	5.1	0	4.9	0
62	Cumin aldehyde	om	1241	0	0	0	0.2	0.1
64	(*E*)-2-Decenal	ntp	1263	0.6	0	0	0.1	0
66	*Neo*-menthyl acetate	om	1273	0.5	0	0	0	0
67	Citronellyl formate	om	1275	0.8	0	0	0	0
69	(*E*)-Anethole	ph	1284	8.6	9.5	11.9	0	tr
70	Isobornyl acetate	om	1285	0	0	0	1.0	0
72	*n*-Tridecane	ntp	1300	0.8	0	0	0	0
74	Undecanal	ntp	1306	1.5	1.2	2.3	tr	0
75	Methyl 4-decenoate	ntp	1310	0	0	0	0	2.4
78	α-Cubebene	sh	1352	0	0	0	0.5	0
81	2-Methylundecanal	ntp	1368	0	0	0	0	4.2
82	α-Copaene	sh	1377	0.2	0	0	0.1	0.2
83	β-Bourbonene	sh	1385	0	0	0	0.2	0
85	β-Elemene	sh	1391	1.5	2.9	3.3	0	0
90	(*Z*)-Caryophyllene	sh	1406	0	0	0	0	1.0
91	Dodecanal	ntp	1407	1.2	0	2.2	0.1	0
92	β-Caryophyllene	sh	1419	0.3	1.9	2.0	tr	0.7
93	β-Copaene	sh	1430	tr	tr	tr	0.1	0.4
94	*cis*-Thujopsene	sh	1431	0	0	0	0	5.2
99	Aromadendrene	sh	1441	0	0	0	0	7.0
100	α-Guaiene	sh	1440	0	4.3	0	0	0
101	α-Himachalene	sh	1450	1.9	0	4.8	0.2	0
103	(*E*)-Geranylacetone	ap	1456	0.8	0	tr	tr	0.8
104	α-Humulene	sh	1455	0	0.5	0.1	0.2	tr
110	γ-Muurolene	sh	1478	1.1	1.2	1.5	0.3	1.1
111	Germacrene D	sh	1482	1.1	6.6	11.1	1.1	3.1
115	α-Muurolene	sh	1499	3.2	0	0	0	0
119	(*Z*)-α-Bisabolene	sh	1504	1.1	0.7	4.2	0	0
120	α-Bulnesene	sh	1507	0	0	2.1	0	7.4
122	β-Bisabolene	sh	1508	0	0	0	0	4
125	*trans*-γ-cadinene	sh	1514	1.2	1.1	2.9	0.7	2.5
130	δ-Cadinene	sh	1524	0.7	1.6	2.1	0.2	0
131	*trans*-Cadina-1(2),4-diene	sh	1534	0	0	0	0	1.9
139	Tetradecanal	ntp	1613	0	0	0.5	0	0
141	Eremoligenol	os	1630	0	0	0	0	0.5
145	*cis*-Methyl dihydrojasmonate	ntp	1655	0	0	0.4	0	0
146	Cadalene	os	1675	0.7	1.0	0	tr	1.6
147	*n*-Heptadecane	ntp	1700	0	0.5	1.4	tr	0
149	*n*-Octadecane	ntp	1800	0.5	0	1.2	tr	0
151	Hexadecanal	ntp	1817	0	1.5	0	0	0
154	Hexahydrofarnesylacetone	ap	1843	0	0	0.7	0	0
155	Cyclohexadecanolide	ntp	1930	0	0	1.3	0	0
156	Hexadecanoic acid	ntp	1960	0	0	2.6	0	0

-	Apocarotenes	ap	0.8	0.0	0.7	0.0	0.8
-	Monoterpene hydrocarbons	mh	19.9	40.8	2.2	73.9	0.0
-	Non-terpene derivatives	ntp	40.8	17.1	37.3	1.0	52.0
-	Oxygenated monoterpenes	om	13.9	8.0	3.6	17.2	0.7
-	Oxygenated sesquiterpenes	os	0.7	1.0	0.0	0.0	2.1
-	Phenylpropanoids	ph	8.6	9.5	11.9	0.0	0.0
-	Sesquiterpene hydrocarbons	sh	12.3	20.8	34.1	3.6	35.0
-	Total		97.0	97.2	89.8	95.7	90.6

595

596

Table 9. Comparison of the non-volatile compounds (relative amounts) in the sampled populations. PFF-T = Cerbaie Hills (Pisa, Tuscany); PFF-A = Torano (Rieti, Lazio); PFF-M = Vallerotonda (Frosinone, Lazio); PFM = Monte Argentario (Grosseto, Tuscany); PFP = Fano (Pesaro e Urbino, Marche). F: flavonoid, A: apocarotenoid, O: oligosaccharide, S: saponin, un = unidentified. Yellow columns = *P. flavescens* subsp. *flavescens*, orange column = *P. flavescens* subsp. *maremmana*, green column = *P. flavescens* subsp. *pisaurensis*. Compound numbers correspond to peak numbers indicated in Figure 5.

Compound	Class	t_R (min)	Parent ion	Product ions	PFF-T	PFF-A	PFF-M	PFM	PFP
1	un	2.0	m		0.688	0.122	0.874	0.691	1.000
2	A	5.7	565[b]	519, 403, 337, 261	1.000	0.787	0.773	0.786	0.973
3	O	6.9	653[b]	635, 585, 517	0.645	0.444	0.409	1.000	0.469
4	un	11.7	213[a]	169, 125, 11	0.751	0.608	0.577	1.000	0.836
5	F	15.5	741[a]	723, 609, 591, 475, 343, 301	1.000	0.390	0.610	0.441	0.382
6	O	18.2	695[b]	635, 529, 491	0.932	0.389	0.371	1.000	0.773
7	O	21.0	799[b]	783, 731, 623, 551, 371	1.000	0.221	0.018	0.082	0.360
8	F	23.7	609[a]	591, 463, 343, 301, 271, 179	1.000	0.458	0.518	0.401	0.646
9	O	27.0	915[a]	900, 723, 709, 691, 503, 341	1.000	0.639	0.511	0.212	0.939
10	F	30.5	947[a]	741, 723, 609, 591, 475, 301	1.000	0.499	0.479	0.385	0.519
11	un	31.8	783[a]	661, 607, 485	0.543	0.503	0.809	1.000	0.781
12	F	33.1	917[a]	899, 741, 723, 609, 591, 475, 301	0.973	0.745	0.613	1.000	0.767
13	un	33.7	669[b]		0.637	0.347	0.246	0.851	1.000
14	F	34.6	887[a]	869, 741, 723, 609, 475, 301	1.000	0.390	0.176	0.245	0.546
15	un	35.6	783[a]	661, 607, 485	1.000	0.453	0.591	0.733	0.691
16	un	36.7	813[a]	691, 607, 485	1.000	0.137	0.010	0.468	0.710
17	F	37.9	845[a]	827, 723, 609, 591, 457, 301	1.000	0.318	0.391	0.387	0.233
18	un	38.8	1121[a]	929, 915, 897, 691, 529	0.463	1.000	0.662	0.284	0.512
19	un	39.5	1121[a]	929, 915, 897, 691, 529	0.894	1.000	0.800	0.672	0.767
20	un	40.5	825[a]	783, 703, 649, 631, 527	1.000	0.126	0.347	0.270	0.339
21	un	41.1	855[a]	733, 649, 631, 527	1.000	0.090	0.168	0.211	0.232
22	un	42.5	1121[a]	915, 897, 691, 673	0.474	0.858	1.000	0.281	0.660
23	un	43.6	1121[a]	915, 897, 691, 529	0.316	1.000	0.775	0.447	0.246
24	S	45.3	1589[a]	1559[c], 1427, 747[d], 455[d]	1.000	0.169	0.000	0.133	0.000
25	S	46.2	1427[a]	1397[c], 1203, 937, 747[d], 455[d], 439[d]	1.000	0.924	0.000	0.204	0.000
26	un	47.6	1121[a]	915, 897, 691	0.699	0.987	1.000	0.617	0.992
27	S	49.6	1427[a]	1397[c], 1203, 937, 747[d], 455[d], 439[d]	1.000	0.631	0.033	0.731	0.275
28	S	51.1	1265[a]	1235[c], 1011, 937, 455[d], 439[d]	0.608	1.000	0.256	0.397	0.244
29	un	52.8	1695[b]	1649, 1611, 1044, 949, 683	0.733	0.695	0.774	1.000	0.856

30	un	53.9	1695[b]	1573, 1259, 1045, 965	0.951	0.278	0.935	1.000	0.949
31	S	56.4	1235[a]	1205[c], 1011, 981, 455[d], 423[d]	1.000	0.570	0.000	0.206	0.646
32	S	57.8	1103[a]	1073[c], 879, 455[d], 439[d]	0.541	1.000	0.223	0.368	0.352
33	S	60.6	1235[a]	1205[c], 1011, 981, 455[d]	1.000	0.169	0.000	0.638	0.105
34	S	62.3	1631[a]	1601[c], 1471, 747[d], 455[d]	1.000	0.115	0.146	0.794	0.291
35	S	63.5	1469[a]	1439[c], 937[d], 747[d], 455[d], 439[d]	0.373	0.140	0.300	0.207	1.000
36	un	64.7	1173[a]	741, 723, 609, 547, 343	1.000	0.387	0.162	0.144	0.000
37	S	64.9	1469[a]	1439[c], 747[d], 455[d], 423[d]	0.000	0.000	0.061	0.016	1.000
38	S	66.0	1601[a]	1571[c], 937[d], 747[d], 455[d]	0.956	0.620	0.344	1.000	0.663
39	S	67.5	1439[a]	1409[c], 455[d], 439[d]	0.392	0.041	0.070	0.291	1.000
40	S	67.7	1615[a]	1585[c], 1291, 1277, 1173, 747[d], 455[d]	0.407	0.414	0.182	1.000	0.252
41	S	68.8	1453[a]	1423[c], 1291, 937[d], 455[d]	1.000	0.447	0.561	0.532	0.000
42	S	69.8	1307[a]	1277[c], 1175, 1011, 455[d],	0.534	0.000	0.353	0.180	1.000
43	S	70.3	1411[a]	1249, 1187, 1025, 747[d], 585[d], 439[d]	0.532	1.000	0.050	0.252	0.000
44	S	70.8	1439[a]	1409[c], 1215, 1173, 1277, 455[d], 423[d]	0.118	0.000	0.000	0.000	1.000
45	S	71.3	1145[a]	1115[c], 849, 747[d], 455[d]	0.000	0.617	0.491	1.000	0.688
46	S	72.4	1277[a]	1247[c], 1115, 1055, 455[d]	1.000	0.415	0.198	0.445	0.592
47	S	73.6	1239[a]	1501[c], 1369, 455[d]	0.000	1.000	0.000	0.000	0.000
48	S	74.0	1511[a]	1481[c], 1349, 937[d], 747[d], 455[d]	0.092	0.042	0.419	0.093	1.000
49	S	74.7	1381[a]	1219, 1157, 995, 585[d], 439[d]	0.851	1.000	0.000	0.352	0.000
50	S	75.1	1349[a]	1319[c], 1187, 937[d], 455[d]	0.050	0.147	0.438	0.067	1.000
51	S	75.8	1395[a]	1233, 1171, 1025, 585[d], 439[d]	0.000	1.000	0.000	0.192	0.000
52	un	76.8	647[a]	579	1.000	0.723	0.195	0.000	0.000
53	S	76.4	1319[a]	1289[c], 1187, 1011, 455[d]	0.000	0.000	0.127	0.000	1.000
54	S	77.1	1481[a]	1451[c], 1319, 1173, 455[d]	0.000	0.000	0.024	0.000	1.000
55	un	77.9	1087[a]	863, 759, 627, 423	1.000	0.684	0.000	0.277	0.000
56	un	79.3	566[a]	581, 543, 499	1.000	0.492	0.211	0.322	0.000
57	S	79.3	1349[a]	1319[c], 1187, 1083, 1041, 455[d]	0.000	0.000	0.000	0.000	1.000
58	S	80.2	1219[a]	1087, 995, 863, 439[d]	0.000	0.000	1.000	0.000	0.000
59	S	81.0	1453[a]	1291, 1187, 747[d], 455[d], 439[d]	0.595	0.746	0.623	0.376	1.000
60	S	81.7	1531[a]	1291, 1231, 1187, 747[d], 439[d]	1.000	0.096	0.177	0.253	0.262
61	S	82.5	1501[a]	1471, 1177, 455[d]	0.414	0.577	0.766	0.318	1.000
62	S	83.5	1423[a]	1199, 1157, 995, 439[d] 1171, 1125, 439[d]	0.710	0.382	0.491	0.380	1.000
63	un	84.4	668[a]	436, 396	0.759	0.181	1.000	0.184	0.656
64	un	85.1	602[a]	579, 549, 273	1.000	0.000	0.000	0.116	0.415
65	un	85.4	675[a]	652, 631, 355	1.000	0.382	0.162	0.215	0.000
66	S	85.9	1261[a]	1219, 995, 585[d], 423[d]	0.137	0.070	0.131	0.102	1.000
67	un	86.7	587[a]	375	0.536	0.150	1.000	0.445	0.607
68	S	87.2	1291[a]	1025, 863, 523[d], 439[d]	0.000	0.169	0.000	0.000	1.000
69	un	87.8	1275[a]	1051, 1009, 991, 863, 669	1.000	0.162	0.068	0.263	0.000
70	S	88.1	1333[a]	1171, 1125, 585[d], 439[d]	0.000	0.000	1.000	0.000	0.000
71	S	88.6	1261[a]	1219, 995, 585[d], 423[d]	0.308	0.174	0.000	0.000	1.000

72	S	90.3	1303[a]	995, 977, 439[d]	0.137	0.106	0.262	0.064	1.000
73	un	93.2	929[a]	883, 817, 725, 657, 493	0.000	0.000	0.000	1.000	0.000
74	un	99.4	883[a]	845, 747, 575	0.982	1.000	0.530	0.304	0.222
75	un	102.7	721[a]	675, 637, 585, 415, 235	0.993	0.769	0.503	0.769	1.000

Compound 2 = 3β-hydroxy-5,6-epoxy-β-ionol 9-*O*-β-D-apiofuranosyl (1→6)-β-D-glucopyranoside; **compound 3** = β-D-(6-*O*-benzoyl)-fructofuranosyl-(2→1)-[β-D-glucopyranosyl-(1→3)]-α-D-glucopyranoside; **compound 5** = quercetin 3-*O*-β-D-apiofuranosyl(1→2)-*O*-[α-L-rhamnopyranosyl-(1→6)]-β-D-glucopyranoside]; **compound 6** = β-D-(6-*O*-benzoyl) fructofuranosyl-(2→1)-[β-D-glucopyranosyl-(1→3)]-6-acetyl-α-D-glucopyranoside; **compound 7** = 3,6'-di-*O*-sinapoylsucrose; **compound 8** = rutin; **compound 9** = reiniose F; **compound 10** = quercetin 3-(5-*O*-sinapoyl)-β-D-apiofuranosyl(1→2)-*O*-[α-L-rhamnopyranosyl-(1→6)]-β-D-glucopyranoside]; **compound 12** = quercetin 3-(5-*O*-*t*-feruloyl)-β-D-apiofuranosyl(1→2)-*O*-[α-L-rhamnopyranosyl-(1→6)]-β-D-glucopyranoside]; **compound 14** = quercetin 3-(5-*O*-coumaroyl)-β-D-apiofuranosyl(1→2)-*O*-[α-L-rhamnopyranosyl-(1→6)]-β-D-glucopyranoside]; **compound 17** = quercetin 3-(5-*O*-benzoyl)-β-D-apiofuranosyl(1→2)-*O*-[α-L-rhamnopyranosyl-(1→6)]-β-D-glucopyranoside];
compound 35 = 3-*O*-β-D-glucopyranosyl presegenin 28-*O*-{β-D-galactopyranosyl-(1→4)-β-D-xylopyranosyl-(1→4)-α-L-rhamnopyranosyl-(1→2)-[β-D-glucopyranosyl-(1→3)]-[4-*O*- acetyl]}-β-D-fucopyranosyl ester; **compound 43** = 3-*O*-β-D-glucopyranosyl medicagenic acid 28-*O*-{β-D-galactopyranosyl-(1→4)-β-D-xylopyranosyl-(1→4)-α-L-rhamnopyranosyl-(1→2)-[β-D-glucopyranosyl-(1→3)]}-β-D-fucopyranosyl ester.
[a] $[M-H]^-$. [b] $[M+HCOO]^-$. [c] $[M-H-30]^-$. [d] Sugar fragments.

References

Abdallah HMI, Elshamy AI, El Gendy AEG, Abd El-Gawad AM, Omer EA, De Leo M, Pistelli L. 2017. Anti-inflammatory, antipyretic, and antinociceptive effects of a *Cressa cretica* aqueous extract. Planta Med. 83: 1313–1320.

Adams RP. 2007). Identification of essential oil components by gas chromatography/mass spectrometry, 4th Ed. Carol Stream, Allured.

Aldrich JR, Mosi MC, Bin F. 2017. Behavioral compounds correlates for minor volatile from stink bugs (Heteroptera: Pentatomidae). J Chem Ecol. 21: 1907–1920.

Andersson S, Nilsson LA, Groth I, Bergström G. 2002. Floral scent in butterfly-pollinated plants: possible convergence in chemical composition. Bot J Linn Soc. 140: 129–153.

Arrigoni PV. 2014. Revisione tassonomica e corologica del genere *Polygala* in Italia. Inform Bot Ital. 46: 235–263.

Astuti G, Roma-Marzio F, Peruzzi L. 2017. Traditional karyomorphological studies: can they still provide a solid basis in plant systematics? Flora Medit. 27: 91–98.

Bacchetta G, Grillo O, Mattana E, Venora G. 2008. Morpho-colorimetric characterization by image analysis to identify diaspores of wild plant species. Flora 203(8): 669–682.

Bartolucci F, Peruzzi L, Galasso G, Albano A, Alessandrini A, Ardenghi NMG, Astuti G, Bacchetta G, Ballelli S, Banfi E, Barberis G, Bernardo L, Bouvet D, Bovio M, Cecchi L, Di Pietro R, Domina G, Fascetti S, Fenu G, Festi F, Foggi F, Gallo L, Gottschlich G, Gubellini L, Iamonico D, Iberite M, Jiménez-Mejías P, Lattanzi E, Marchetti D, Martinetto E, Masin RR, Medagli P, Passalacqua NG, Peccenini S, Pennesi R, Pierini B, Poldini L, Prosser F, Raimondo FM, Roma-Marzio F, Rosati L, Santangelo A, Scoppola A, Scortegagna S, Selvaggi A, Selvi F, Soldano A, Stinca A, Wagensommer RP, Wilhalm T, Conti F. 2018. An updated checklist of the vascular flora native to Italy. Pl Biosyst. 152(2): 179–303.

Bernardi LF. 2000. Consideraciones taxonómicas y fitogeográficas acerca de 101 Polygalae Americanas. Cavanillesia Altera 1: 1–456.

648Borghi M, Fernie AR, Schiestl FP, Bouwmeester HJ. 2017. The sexual advantage of looking,
649 smelling, and tasting good: the metabolic network that produces signals for pollinators. Trends Pl
650 Sci. 22: 338–350.
651Borg-Karlson AK, Unelius CR, Valterova I, Nilsson A. 1996. Floral fragrance chemistry in the
652 early flowering shrub *Daphne mezereum*. Phytochemistry. 41: 1477–1483.
653Brew C, O'Dowd DJ, Rae ID. 1989. Seed dispersal by ants: behaviour releasing compounds in
654 elaiosomes. Oecologia. 80: 490–497.
655Carta A, Flamini G, Cioni PL, Pistelli L, Peruzzi L. 2015. Flower bouquet variation in four species
656 of *Crocus* ser. *Verni*. J Chem Ecol. 41: 105–110.
657Chodat R. 1893. Monographia Polygalacearum. Mem Soc Phys Geneve. 31: 1–500.
658Clegg MT, Durbin M. 2000. Flower color variation: A model for the experimental study of
659 evolution. PNAS. 97: 7016–7023.
660Conti F, Abbate G, Alessandrini A, Blasi C. 2005. An annotated checklist of the Italian vascular
661 flora. Palombi Editori, Roma.
662Dal Piaz F, De Leo M, Braca A, De Simone F, Morelli I, De Tommasi, N. 2005. Electrospray
663 ionization mass spectrometry for identification and structural characterization of pregnane
664 glycosides. Rapid Commun. Mass Spectrom. 19: 1041–1052.
665Davies NW. 1990. Gas chromatographic retention indexes of monoterpenes and sesquiterpenes on
666 methyl silicone and Carbowax 20 M phases. J Chromatogr. 503: 1–24.
667De Leo M, Peruzzi L, Granchi C, Tuccinardi T, Minutolo F, De Tommasi N, Braca A. 2017.
668 Constituents of *Polygala flavescens* ssp. *flavescens* and their activity as inhibitors of human
669 lactate dehydrogenase. J Nat Prod. 80: 2077–2087.
670Del Guacchio E. 2010. Appunti di floristica campana: novità e precisazioni. Inform Bot Ital. 42, 35–
671 46.
672Faegri K, van der Pijl L. 1979. The principles of pollination ecology, 3rd edn. Pergamon Press,
673 Oxford.

674Fenaroli L. 1970. Florae garganicae prodromus. Pars altera. Webbia. 24: 435–578.

675Fiori A. 1925. *Polygala* L. In: Fiori A (eds.) Nuova Flora Analitica d'Italia 2. Tip. Ricci, Firenze,

676 pp. 120–126.

677Fiori A, Béguinot A, Pampanini R. 1908. Schedae ad Floram Italicam Exsiccatam. Centuriae IX–X.

678 N. Giorn Bot Ital n.s. 15: 445–532.

679Grillo O, Draper D, Venora G, Martínez-Laborde JB. 2012. Seed image analysis and taxonomy of

680 *Diplotaxis* DC. (Brassicaceae, Brassiceae). Syst Biodiv. 10(1): 57–70.

681Gubellini L, Hoffmann N, Pinzi M. 2014. Contributo alla conoscenza della flora vascolare delle

682 Marche di alcune regioni limitrofe. Inform Bot Ital. 46: 17–26.

683Hammer Ø. 2017. PAST Version 3.15. Available from: http://folk.uio.no/ohammer/past/.

684Hammer Ø, Harper DAT, Ryan PD. 2001. PAST: paleontological statistics software package for

685 education and data analysis. Palaeontol Electron. 4: 1–9.

686Heywood VH, Brummitt RK, Culham A, Seberg O. 2007. Flowering plant family of the World.

687 Firefly Books, Ontario, Canada.

688Irwin RE, Strauss SY, Storz S, Emerson A, Guibert G. 2003. The role of herbivores in the

689 maintenance of a flower color polymorphism in wild radish. Ecology. 84: 1733–1743.

690Jennings W, Shibamoto T. 1980. Qualitative analysis of flavor and fragrance volatiles by glass

691 capillary chromatography. Academic, New York.

692Kimball S, Mattis, P. 2014. Gimp software 2.8.14. Available at www.gimp.org.

693Klein LC, Andrade de SF, Filho VC. 2012. A pharmacognostic approach to the *Polygala* genus:

694 phytochemical and pharmacological aspects. Chem Biodiv. 9: 181−209.

695Knudsen JT, Eriksson R, Gershenzon J, Ståhl B. 2006. Diversity and distribution of floral scent. Bot

696 Rev. 72: 1–120.

697Marques MCM. 1979. Revisão das espécies do gênero *Polygala* L. (Polygalaceae) do estado do Rio

698 de Janeiro. Rodriguésia. 31: 69–339.

699Masada Y. 1976. Analysis of essential oils by gas chromatography and mass spectrometry. Wiley,
700 New York.
701McNeill J. 1968. *Polygala* L. – Pp. 231-236 in: Tutin, T.G., Burges, N.A., Chater, A.O.,
702 Edmondson, J.R., Heywood, V.H., Moore, D.M., Valentine, D.H., Walters, S.M., Webb, D.A.
703 (ed.), Flora europaea 2. Cambridge University Press.
704Paiva JAR. 1998. Polygalarum Africanarum et Madagascariensium prodromus atque gerontogaei
705 generis *Heterosamara* Kuntze, a genere *Polygala* L. segregati et a nobis denuo recepti, synopsis
706 monographica. Fontqueria. 50: 1–346.
707Passalacqua NG, Tundis R, Upson TM. 2017. A new species of *Lavandula* sect. *Lavandula*
708 (Lamiaceae) and review of species boundaries in *Lavandula angustifolia*. Phytotaxa. 292: 161–
709 170.
710Peruzzi L, Passalacqua NG, Gargano D. 2005. Re-evaluation of *Polygala apiculata* (Polygalaceae),
711 a rare endemic of S Italy. Willdenowia. 35: 65–68.
712Peruzzi L, Conti F, Bartolucci F. 2014. An inventory of vascular plants endemic to Italy. Phytotaxa.
713 168(1): 1–75.
714Peruzzi L, Astuti G, Carta A, Roma-Marzio F, Dolci D, Caldararo F, Bartolucci F, Bernardo L.
715 2015. Nomenclature, morphometry, karyology and SEM cypselae analysis of *Carduus brutius*
716 (Asteraceae) and its relatives. Phytotaxa. 202: 237–249.
717Peruzzi L, Chiavegatto RB, D'Antraccoli M, Dolci D, Roma-Marzio F. 2017. Chromosome
718 numbers for the Italian flora: 4. Italian Botanist. 4: 53–59.
719Pignatti S. 1982. Flora d'Italia vol. 2. Edagricole (Ed.), Bologna.
720Pignatti S. 2017. Flora d'Italia Ed. 2, vol. 2. Edagricole (Ed.), Bologna.
721R Core Team. 2016. R: a Language and Environment for Statistical Computing. R Foundation for
722 Statistical Computing, Vienna, Austria. https://www.R-project.org/.
723Raguso RA. 2008. Wake up and smell the roses: the ecology and evolution of floral scent. Annu
724 Rev Ecol Evol Syst. 39: 549–569.

725Rasband WS. 1997. ImageJ 1.47v. U.S. National Institutes of Health, Bethesda, Maryland, U.S.A.
726 Available from: http://imagej.nih.gov/ij/ (Accessed 23 October 2017).
727Roma-Marzio F, Peruzzi L. 2017. Typification of the name *Polygala flavescens* DC.
728 (Polygalaceae). Phytotaxa. 331(2): 299–300.
729Roma-Marzio F, Najar B, Alessandri J, Pistelli L, Peruzzi L. 2017. Taxonomy of prickly juniper
730 (*Juniperus oxycedrus* group): a phytochemical-morphometric combined approach at the contact
731 zone of two cryptospecies. Phytochemistry. 141: 48–60.
732Schiestl FP, Johnson SD. 2013. Pollinator-mediated evolution of floral signals. Trends Ecol Evol.
733 28: 307–315.
734Stenhagen E, Abrahamsson S, McLafferty FW. 1974. Registry of mass spectral data. Wiley, New
735 York.
736Stuessy TF. 2009. Plant Taxonomy. The Systematic Evaluation of Comparative Data. Columbia
737 University Press.
738Swigar AA, Silverstein RM. 1981. Monoterpenes. Aldrich Chem Comp, Milwaukee.
739Thiers B. 2017. Index Herbariorum: a Global Directory of Public Herbaria and Associated Staff.
740 Available from: http://sweetgum.nybg.org/ih/. Accessed 7 October 2017.
741Tundis R, Loizzo MR, Bonesi M, Leporini M, Menichini F, Passalacqua NG. 2016. A study of
742 *Salvia fruticosa* Mill. subsp. *thomasii* (Lacaita) Brullo, Guglielmo, Pavone & Terrasi, an
743 endemic sage of southern Italy. Pl Biosyst. 152: 130–141.
744Vickery RK Jr. 1992. Pollinator preferences for yellow, orange, and red flowers of *Mimulus*
745 *verbenaceus* and *M. cardinalis*. Great Basin Naturalist. 52: 145–148.
746Weiss MR. 1995. Floral colour change: a widespread functional convergence. Am J Bot. 82: 167–
747 185.
748Westerkamp C. 1997. Keel blossoms: bee flowers with adaptations against bees. Flora. 192, 125–
749 132.

750Willmer PG, Nuttman CV, Raine NE, Stone GN, Pattrick JG, Henson K, Stillman P, McIlroy L,
751 Potts SG, Knudsen JT. 2009. Floral volatiles controlling ant behaviour. Funct Ecol. 23: 888–900.
752Youngsteadt E, Nojima S, Häberlein C, Schulz S, Schal C. 2008. Seed odor mediates an obligate
753 ant–plant mutualism in Amazonian rainforests. PNAS. 105(12): 4571–4575.
754Zangheri P. 1976. Flora Italica. Cedam, Padova.