DE GRUYTER
OPEN

herba polonica
From Botanical to Medical Research
Vol. 62 No. 2 2016
DOI: 10.1515/hepo-2016-0010

EXPERIMENTAL PAPER

Statistical analysis of the associations between phenolic monoterpenes and molecular markers, AFLPs and SAMPLs in the spice plant Oregano

ALI AZIZI[1], HAMIDREZA ARDALANI[2*], BERND HONERMEIER[3]

[1]Department of Horticultural Sciences
Faculty of Agriculture
Bu-Ali Sina University
Hamedan, Iran

[2]Department of Horticultural Sciences
Science and Research Branch
Islamic Azad University
Tehran, Iran

[3]Institute of Agronomy and Plant Breeding
Biomedical Research Centre Seltersberg (BFS)
Justus Liebig University Gießen
Schubert-Str. 81, D-35392 Gießen, Germany

[*]corresponding author: phone: +98 9128842797, fax: +98 8132532432,
e-mail: hrardalani@srbiau.ac.ir

Summary

Introduction: Molecular markers are the examples of the contribution of genome technology to medicinal plant breeding through marker-assisted selection (MAS) for pharmaceutical quality. **Objective:** Forty-two accessions of *Origanum vulgare* L. originating from Europe were evaluated to detect genomic and chemotypic polymorphisms and to discover possible associations between them. **Methods:** A total of 477 molecular polymorphisms including 214 AFLP (Amplified Fragment Length Polymorphism) and 263 SAMPL (Selectively Amplified Microsatellite Polymorphic Loci) were used for genotyping. Components in the essential oils were identified and quantified by gas chromatography (GC) and two major compounds (two economically important monoterpenes: carvacrol and thymol) were investigated. **Results:** Based on results, a relatively high correlation between chemotypic patterns and genetic markers was identified. Associations between traits of interest for

essential oils (carvacrol and thymol content) and genetic markers were tested using five statistical methods including three General Linear Model (GLM) and two unified Mixed Linear Model (MLM) approaches. Significant associations were found for 3 AFLP and 20 SAMPL with three key traits including essential oil yield, carvacrol and thymol content. **Conclusion:** These associations can constitute a useful starting point for marker-assisted selection. Therefore, the results provide the basis for molecular breeding of *O. vulgare* for pharmaceutical purposes.

Key words: *Origanum vulgare, essential oils, AFLP, SAMPL, marker-trait associations*

INTRODUCTION

The aromatic value and pharmaceutical properties of the spice and medicinal plant oregano (*Origanum vulgare* L.) is a consequence of the presence of its essential oil in the aerial parts which consists mostly of monoterpenes and sesquiterpenes [1]. *O. vulgare* is a perennial herb belonging to the family *Lamiaceae*. The species is naturally distributed all over Europe, North Africa and western Asia [2]. The biological activity of essential oils and herb extracts causes a high pharmaceutical interest in oregano, since antimicrobial, antifungal, insecticidal and antioxidative effects have been reported [3]. The species of *O. vulgare* has also a high phytochemical polymorphism with several chemotypes which shows marked spatial segregation in nature [4]. The essential oil of oregano is composed of carvacrol and/or thymol as dominant components, followed by γ-terpinene, *p*-cymene, linalool and terpinen-4-ol. A high level of carvacrol content in oregano is one of the most important goals of breeding, which has resulted in a number of cultivars in this species [5]. Characters most targeted in breeding of oregano include those related to spice and pharmaceutical properties such as essential oil yield and the content of two strong antimicrobial and phenolic monoterpenes, carvacrol and thymol [5]. DNA-based molecular markers have become increasingly important for surveying genetic fingerprinting and authentication of medicinal plants [5]. The chromosome number of *O. vulgare* was previously reported to be 2n = 2x = 32 [6]. Identification of SSRs (Simple Sequence Repeats) derived from ESTs (Expressed Sequence Tags) of epidermal glands on this commercial subspecies was also reported [7]. In a previous work, we have investigated genetic relationships between different populations of *O. vulgare* using two PCR-based marker approaches, Amplified Fragment Length Polymorphism (AFLP) and Selectively Amplified Microsatellite Polymorphic Loci (SAMPL), and we have also compared the relative efficiencies of these two marker systems for surveying intraspecific genetic diversity [8]. Most of the medicinal and aromatic plants species used commercially are collected from the wild nature. Domestication is a viable alternative and offers the opportunity to solve the problem. By bringing medicinal herbs into cultivation, conventional and biotechnological

plant breeding techniques can be applied to improve yield and uniformity, and to modify pharmaceutical properties [9]. Selection, hybridization, polyploidization and mutation are some effective strategies to improve medicinal and spice plants. A very useful tool for improving the efficiency of breeding programmes is the identification of genetic markers associated with important traits [10-11]. The methodology of association analyses, suitable for bi-allelic codominant marker types, mainly SSRs and single nucleotide polymorphisms (SNPs), has been well developed and used in a number of crop plants. The potential of dominant markers, such as RAPD and AFLP is poorly explored for association studies. However, many underrepresented plant species, such as most of medicinal plants with limited genomic information largely rely on dominant marker types, such as AFLPs [12-13]. These authors have recently investigated the use of dominant markers for estimating linkage disequilibrium in diploid species and developed an appropriate algorithm. Now, there are a number of reports on the use of AFLP markers for genome-wide linkage disequilibrium analyses and association studies in plants [14]. Studies on dominant markers suggested that they can be successfully applied to quantify population structure and assigning individuals to subpopulations (Q matrix) using a Bayesian approach when a large number of loci are genotyped. Dominant markers can also be a useful tool to estimate the kinship coefficients between individuals within populations. Many researcher incorporated the outcome of population structure (Q matrix) with the estimation of relatedness between individuals obtained through the marker-based kinship matrix (K) into a unified Mixed Linear Model (MLM) approach, that this approach effectively decreases error rates (false positives) and increases the power of the marker-trait association tests [15-16]. In a previous work, we have detected the correlations between genetic, morphological, and chemical diversities in a germplasm collection of the medicinal plant *O. vulgare*. In the present study, we use several models of statistical analysis to elucidate any marker-trait association using two kinds of molecular markers, AFLPs and SAMPLs together with pharmaceutically important traits in oregano [16].

MATERIAL AND METHODS

Plant material

A total of 42 accessions of *O. vulgare* were investigated, 39 accessions (including: ORI 2, ORI 6, ORI 7, ORI 8, ORI 10, ORI 11, ORI 12, ORI 13, ORI 14, ORI 15, ORI 16, ORI 17, ORI 18, ORI 19, ORI 20, ORI 21, ORI 23, ORI 24, ORI 25, ORI 26, ORI 27, ORI 28, ORI 29, ORI 30, ORI 31, ORI 33, ORI 34, ORI 35, ORI 36, ORI 37, ORI 39, ORI 40, ORI 41, ORI 42, ORI 43, ORI 45, ORI 47, ORI 49 and ORI 50) from the Gatersleben Genebank (IPK Gatersleben, Germany) along with three cultivated

types: 'Heracleoticum' from the seed company Pharmasaat (Artern, Germany), 'Creticum' and 'Samothrake' from the Syringa company (Hilzingen-Binningen, Germany). All accessions were grown in greenhouse. In March 2009, 10 individual plants of each accession (4-leaves stage) were transplanted from the seedling bed into pots (6 L). The soil mixture used in this experiment was based on a loess soil. The soil contained 7.8 mg P/100 g, 14.9 mg K/100 g and 1.42 mg N/100 g. The sieved soil was mixed with sand (soil: sand = 1:2 w/w) and fertilized with N, P, K, Mg and Ca to warrant optimal nutrient supply for plant growth. The mixed soil showed pH 6.7 (in H_2O). Plants were watered approximately twice a week by a controlled drip irrigation system. Finally, plants were harvested individually at the full flowering stage in July 2009.

Phytochemical evaluation

Samples of at least 20 g of dried leaves and inflorescences were hydro-distilled for 3h using a Clevenger-type apparatus. The essential oil yields (EOY) were gravimetrically (w/w) quantified. Each sample was analysed three times and the average yield of essential oil was used for statistic evaluation. The essential oil obtained was kept at 4°C until further analysis. The identification of individual components of the essential oils was realized by gas chromatography–mass spectrometry (GC–MS). For quantification purposes, percent values of peak areas were determined by gas chromatography–flame ionization detector (GC–FID). A Varian 3900 GC coupled with a Varian Saturn 2100T ion trap mass detector and a Varian CP-3800 GC–FID was employed. The chromatographic procedures have been previously described by other researcher. The identification of components of the essential oil was achieved on the basis of comparison of Kovat's retention indices (KI) with those of literature data and mass spectrometry by the comparing mass spectra of the unknown peaks with those stored in the Wiley 90 and NIST 98 mass libraries . Kovat's retention indices were calculated from the gas chromatogram by linear interpolation between bracketing *n*-alkanes.

Genotyping by AFLP and SAMPL analyses

Total genomic DNA was extracted from young leaves (100 mg per plant) of 5-week-old plants using a modified CTAB (Cetyltrimethyl Ammonium Bromide) procedure, according to other studies. After RNAse treatment, DNA content was quantified using Nano Drop ND-1000 UV-Vis Spectrophotometer (Labtech International, Ringmer, United Kingdom). Genomic DNA of 10 plants per accession was bulked and diluted to 25 ng/μl working solution.

The AFLP analysis was conducted as described by SAMPL procedure used here is according to. Pre-amplified AFLP library was used as template for selective amplification using fluorescent dye-labelled SAMPL and MseI+3 primers. Twenty four primer combinations were tested for both AFLP and SAMPL analysis. Out of them, three were selected for each marker (tab. 1) on the basis of their ability to generate informative data and values of resolving power. Selective amplification products were separated on 8% denaturing polyacrylamide gels using a Li-Cor 4200 DNA analyser. Their size was estimated in comparison to a 50-750bp labelled DNA-ladder. AFLP and SAMPL fragments were detected using the RFLP-scan 2.1 software package (Scan Analytics, Fairfax, USA). The bands were scored for their presence (1) or absence (0) across 42 accessions.

Table 1.

SAMPL and AFLP primer combinations used for genotyping 42 accessions of *O. vulgare* and total number (n), number of polymorphic bands (np) and resolving power (Rp) of primer combinations

Primer combination		Rp	n	np
AFLP-1	E-CAT × M-CAT	29,9	103	77
AFLP-2	E-ATG × M-CCG	28,3	76	65
AFLP-3	E-CAG × M-CTC	34,1	106	72
SAMPL-1	G(TG)4(AG)4A × M-ACG	30,6	99	92
SAMPL-2	G(TG)4(AG)4A × M-GTG	24,3	92	89
SAMPL-3	C(AC)4(AG)4A × M-CAG	24,8	96	82
Total			572	477

Comparison of distance matrixes

Euclidean distances were computed between accessions based on the chemotypic characters of the essential oils. In order to investigate the congruencies between chemotypic and genomic distances, the genetic distance matrix based on combined dataset of AFLP and SAMPL and the Euclidean phytochemical distance matrix were compared using Mantel tests by the MAXCOMP routine of NTSYS-pc software. The normalized Mantel statistic Z was used to determine the level of association between the matrices. Significance of Z was determined by comparing the observed Z values with a critical Z value obtained by calculating Z for one matrix with 1000 permuted variants of the second matrix.

Population inference and Principal component analysis

In order to infer population structure among accessions, the AFLP and SAMPL polymorphic markers were analysed by dominant-marker model of the STRUCTURE software (v. 2.3.1) used to assign accessions into subpopulations (K). Posterior probabilities of K (*Pr (X | K)*) were obtained for K=1 through K=10 clusters using the Admixture model, which allows for potential recombination between inferred clusters. Five runs were completed for each K, with 100,000 iterations, following a burn-in period of 50,000 iterations to find the optimal number of subpopulations and membership of each accession. We inferred the number of subpopulations according to other studies with posterior probabilities of K calculated assuming uniform priors on K and using for each K the maximum value of the probability of the data given K obtained over replicates. In order to provide an overall distance measure between the accessions based on major chemical components, Principal component analysis (PCA) was conducted on the accession means for each observed character using the NTSYS-pc.

Association analyses

Association tests between chemotypic traits and polymorphic AFLP and SAMPL markers were carried out across all accessions using the TASSEL (v. 3.0) software. This software determines association between genomic sites and phenotypes, while accounting for population structure and relative kinship. Five different approaches were used to control false-positive results in association tests (tab. 2). First, a General linear model (GLM) was tested to detect single marker effects on chemical traits. This model does not account for population structure as a potential cause of the genotype-chemotype relationship. In the second GLM model, Principal components (PC) 1 through 3 (PC_{1-3}) were used as quantitative covariates. For third GLM model estimates of the population structure obtained from the STRUCTURE software were incorporated into the model by using covariates that indicate percent contribution to each accession by a specific subpopulation (Q-matrix). A fourth and fifth model were tested using a unified mixed linear model (MLM). One contained the relative kinship matrix estimated from molecular marker data among all accessions, and the second contained the kinship matrix (K) plus the population structure (Q). P values for association tests were obtained from the F value of effects of each marker locus on trait values. Significance of F values was confirmed by 1000 permutations for each marker. The trait was considered to be significantly associated with a marker locus when both the P value from the F test and the experiment wise P value from the permutation test were <0.01. Chemotypic (phenotypic) variance values (partial R^2) were computed for the fixed marker effects.

Table 2.

Methods performed for association analyses between molecular markers and chemotypic traits of *Origanum vulgare*

Model	Data sets used in the analyses
1- GLM	(Chemotype) + (AFLP, SAMPL)
2- GLM	(Chemotype) + (AFLP, SAMPL) + (Covariates: PC_{1-3})
3- GLM	(Chemotype) + (AFLP, SAMPL) + (Covariates: Q_{1-5})
4- MLM	(Chemotype) + (AFLP, SAMPL) + (K)
5- MLM	(Chemotype) + (AFLP, SAMPL) + (K) + (Covariates: Q_{1-5})

Abbreviations:
GLM – General linear model, MLM – Mixed linear model, PC – Principal components
Q – population structure that results from the existence of subpopulations
K – Kinship matrix (general similarity in genetic background arising from shared kinship)

Ethical approval: The conducted research is not related to either human or animal use.

RESULTS

Genomic and chemical polymorphisms

The six selected AFLP and SAMPL primer combinations yielded a total of 572 scorable fragments across 42 accessions, of which 477 were found to be polymorphic (tab. 1) ranging in size from 50 to 470 bp. A total of 62 volatile compounds were detected in essential oils by GC and GC/MS analyses, however, the mean values of relative percentage amounts of these compounds were considered for chemotypic matrix construction. The chromatographic fingerprints showed the presence of high intraspecific diversity of chemical constituents in the essential oils from the accessions of *O. vulgare*. Dominant components in essential oils that determine different chemotypes were four monoterpenes, including carvacrol, thymol, *p*-cymene and γ-terpinene as well as two sesquiterpenes: β-caryophyllene and germacrene-D. In comparison to other reports on *O. vulgare*, relatively high content of germacrene-D was found in essential oils investigated in our study. Inferred population structure and Principal components analysis of 42 accessions with 477 AFLP and SAMPL markers identified six distinct subpopulations. The value of $Pr(X|K)$ was optimized at K = 6 for most numerical solutions. The membership of accessions in K = 6 subpopulations was highly consistent across multiple solutions. This resulting Q-matrix was used for third GLM model as covariates that indicate percent contribution to each accession by a specific subpopulation. In order to define and chemotypic relationships among the accessions, main values of essential oil data were elaborated to conduct a Principal component analysis. PCA revealed the existence of high genomic and phytochemical variations among *O. vulgare* accessions (fig. 1). The PCA based on the

concentration of 15 components in the essential oils were able to separate the accessions containing the desirable monoterpenes, carvacrol, thymol from the other accessions containing the undesirable sesquiterpene (fig. 1).

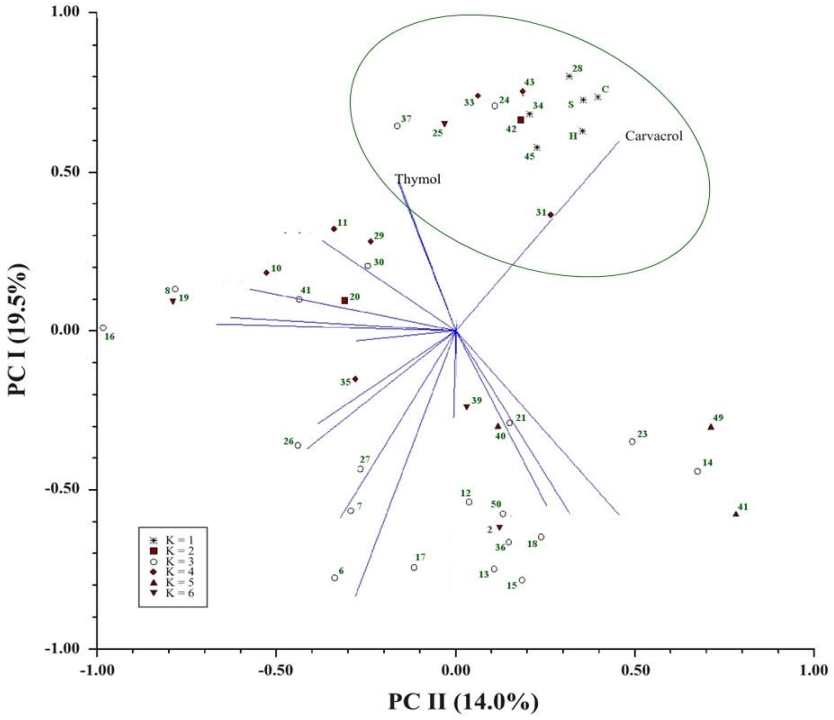

Figure 1.

Principal Component Analysis (PCA) showing the relationships among 42 accessions of *Origanum vulgare* based on chemotypic data (using 15 main components of essential oils, including carvacrol and thymol). K is subpopulation estimated by STRUCTURE analysis of 477 AFLP and SAMPL. The numbers represent the accession number and C, H and S represent "Creticum", "Heracleoticum", and "Samothrake", respectively.

Associations

The Mantel tests also showed that there was a relatively high, significant correlation between two distance matrices based on the genetic markers (AFLP and SAMPL) and the chemical compounds of essential oils ($r = 0.65$), confirming the congruence of intercession relationships in *O. vulgare* revealed by the chemical and molecular markers. A test of associations between 477 AFLP and SAMPL markers, essential oil yield and 2 chemotypic characters (two economically traits important for oregano, carvacrol and thymol content) detected significant marker-trait associations for these traits using at least one of the simple or population structure controlling models. A total of 1431 (477 markers × 3 traits) association tests

were performed by each of the five models. Of these, 91, 42, 32, 74 and 68 were significant at the nominal threshold of $p = 0.01$ based on models 1, 2, 3, 4 and 5 (tab. 2), respectively. The number of detected significant associations decreased in almost three traits when population structure was accounted for each GLM or MLM model. Associations between 23 markers and three major economically important traits are presented in table 3. These traits include essential oil yield (EOY), carvacrol content (CAC) and thymol content (THC). The table also shows the proportion of phenotypic variance (R^2) of the traits explained by markers detected by significance based on the five models. The effects of AFLP-2_31 on two traits, EOY and CAC, were significant by all five tested models. Twelve markers were found to be associated to THC, and association with the highest R^2 value for a trait was obtained for SAMPL-3_60 influencing THC (tab. 3).

Table 3.

Proportion of phenotypic variance (R^2) of three key traits explained by AFLP and SAMPL markers detected by significance of five models of association analysis: first numeral through fifth are R^2 based model 1 through 5 (see tab. 2)

Locus	EOY					CAC					THC				
AFLP-2_31	0.10	0.08	0.05	0.06	0.02	0.13	0.10	0.08	0.09	0.03					
AFLP-2_44						–	0.15	0.10	0.09	0.02	–	0.17	0.17	0.13	0.03
AFLP-3_6											0.19	0.12	–	0.10	0.10
SAMPL-1_3	–	0.26	0.11	–	–	–	–	–	0.10	–	0.14	–	–	0.04	–
SAMPL-1_25	0.23	0.13	–	0.03	–										
SAMPL-1_54	0.28	–	–	0.02	–	0.20	–	–	–	–	0.18	–	–	0.14	–
SAMPL-1_56	–	–	–	0.05	–	–	0.06	–	0.11	–					
SAMPL-1_65	–	0.05	–	0.08	–	–	0.05	–	0.08	–					
SAMPL-1_68	0.43	–	0.04	–	0.04	0.32	–	–	–	0.09					
SAMPL-1_81						0.16	–	0.07	0.06	0.02					
SAMPL-2_14	0.24	0.09	–	0.03	0.04										
SAMPL-2_29	0.32	–	–	0.04	–	0.22	–	–	–	–					
SAMPL-2_82	–	0.07	–	0.04	–	–	0.10	–	0.09	–	0.15	–	–	–	–
SAMPL-3_43											0.15	0.16	0.14	0.10	0.07
SAMPL-3_49											0.10	0.16	0.12	0.08	0.05
SAMPL-3_54	0.34	0.09	–	0.04	–	–	0.17	–	–	–	0.19	–	–	0.07	–
SAMPL-3_55											0.18	0.14	0.13	0.08	0.07
SAMPL-3_60											0.23	0.19	0.19	0.17	0.09
SAMPL-3_66	–	–	0.05	–	0.05						–	0.13	0.13	0.05	0.10
SAMPL-3_71	0.18	–	–	–	–	0.20	–	–	0.01	–					
SAMPL-3_73	–	0.09	–	0.04	–	–	0.20	0.07	–	0.06	0.24	0.12	0.09	0.05	0.02
SAMPL-3_78	0.17	–	–	–	–	0.20	–	–	–	–					
SAMPL-3_79	–	–	–	0.04	–	–	–	–	0.08	–					

The R^2 values shown correspond to the significance of marker at $p<0.01$and (–) indicates that the model represented by the numeral in this position did not meet the $p<0.01$ criterion.
Trait abbreviations: EOY – essential oil yield, CAC – carvacrol content, THC – thymol content

DISCUSSION

Polymorphisms and distance matrix congruencies

In the absence of SSR markers, as is still the case in *O. vulgare*, these two marker systems appear to be highly suitable for genetic diversity studies of the medicinal plants [17-18]. In applied breeding for medicinal plants' improvement, chemotypic and genetic distances between genotypes are expected to provide predictors for high heterosis effects on pharmaceutical qualities and yield performance of their hybrids [19-20]. In the present study, among the accessions based on thymol content and essential oil yield, ORI8, ORI25, ORI27, ORI29, and ORI37 showed promising performance that can be exploited in breeding programmes. The Mantel test showed that there was a relatively high significant correlation between the distance matrix based on the chemical compounds of essential oils (terpenes) and the distance matrix based on the combined data set of AFLP and SAMPL markers, confirming the congruence of inter-accession relationships in *O. vulgare* revealed by the chemical and molecular markers. This congruence supports the RAPD-terpenes correlation in a previous study on *Juniperus* spp. [21] and the ISSR-terpenes correlation reported for *Primula ovalifolia* [22]. High correlation between genetic and terpenoid distance matrices was also obtained in other aromatic and medicinal plants belonging to the *Lamiaceae* family, such as *Ocimum gratissimum*, *Thymus vulgaris, Melissa officinalis* and *O. vulgare* [23]. Genetic control of the chemical characteristics of thyme plants belonging to the *Lamiaceae* and fennel from *Apiaceae* has been proven, so the monoterpenes accumulated by the plant are controlled by a series of loci with epistatic relationships [24]. Therefore, it is necessary to analyse both Quantitative Trait Loci (QTL) and genetic markers to better explain the relationship between the two sets of variations.

Marker-trait associations

The aim of marker-trait association analyses in present work was to constitute a starting point for marker-assisted selection in *O. vulgare* using AFLP, SAMPL and chemical polymorphisms. The whole-genome association analyses by AFLP, SAMPL markers dispersed throughout the genome could lead to the identification of a number of markers with significant associations to some economically and pharmaceutically important traits [25]. We have identified three AFLP and 20 SAMPL markers associated to traits, essential oil yield, carvacrol content and thymol content (tab. 3). SAMPL markers seem to be more effective for association analyses as in our previous study, this marker system was found to be useful for studies on intraspecific diversity and relationships among *O. vulgare* subspecies [26-27]. Considering that mapping data for the AFLP and SAMPL markers were not available, we were not able to examine the extent of disequilibrium among associated

markers, so that it is not possible to speculate on the degree of disequilibrium between identified markers. However, the identified novel allelic variation for these important traits should be of considerable interest for breeding purposes, since the gene-linked SSRs and locus-specific SNPs are still not developed for *O. vulgare*. The significant associations resulting from an unrecognized population structure are considered to be false positives. Therefore, we have used five different approaches for marker-trait association analysing to correct the effect of population structure. In the five models performed in the present study, using inferred population structure 'Q' (proportion of membership of accessions to sub-populations) or kinship matrix 'K' (general similarity based on shared kinship) for GLM or MLM models, decreased the number of detected significant associations and also reduced R^2 values (proportion of phenotypic variance explained) for most of the traits. In agreement with other studies, which have analysed associations of AFLP markers with quantitative traits among oat varieties, the lowest R^2 values were obtained when the combination of both 'Q' and 'K' was used in MLM model (tab. 3, fifth numeral). Marker-trait associations found in the present study were supported by 1,000-times permutated *p*-values of the GLM and MLM models. However, it is still possible that some of the marker-trait associations identified in our study are false positives, therefore further validation is required. Further mapping studies in segregating populations will help confirm whether the associated markers are linked to QTLs influencing the traits and whether any of these QTL effects are caused by orthologs of known genetic factors. The validation of QTLs for the traits of interest can also be assisted by functional genomic studies. Studies on QTLs linked to synthesis pathways of different monoterpenes in aromatic plants of the *Lamiaceae* family are very rare. Although there is no information on the mode of inheritance of carvacrol and thymol contents in *O. vulgare*, it has been reported that biosynthesis of these two phenolic monoterpene in *Thymus vulgaris* is governed by an epistatic series of several biosynthetic loci. However, with regard to broad variation in the essential oil of *O. vulgare*, the biosynthetic pathway of carvacrol and thymol seems to be different and more complicated. This makes it difficult to detect QTLs for this pathway and also to identify individual genes because specific pathway branches control the synthesis of different monoterpenes. However, the markers showing strongest effects on three economically and pharmaceutically important traits investigated in this study, could be starting points for further studies, marker assisted selection and practical breeding [28-29]. Among the markers listed in table 3, based on the size of R^2 values, and the co-association with more than one trait, there are two markers that are most interesting candidates for further work: I) Marker AFLP-2_31 may be used in breeding for pharmaceutical quality because it was co-associated with two key traits: essential oil yield and carvacrol content. However, the marker effects (R^2 values) were relatively low (tab. 3). II) The relatively strong effect of SAMPL-3_60 on thymol content (R^2 values: 0.23, 0.19, 0.19, 0.17 and 0.09) could make it useful for marker assisted selection of this very important antimicrobial compound (tab. 3).

The identified markers can potentially help to improve the polygenic, complex quantitative traits related to pharmaceutical quality of *O. vulgare*.

CONCLUSIONS

Significant correlation between two distance matrices based on the genetic marker systems (AFLP and SAMPL) and the chemical compounds of essential oils indicated the congruence of inter-accession relationships in *O. vulgare* revealed by the chemical and molecular markers. Finally, analyses of marker-trait associations showed that two molecular markers, AFLP-2_31 and SAMPL-3_60 may be included in marker-assisted programmes to improve breeding efficiency of pharmaceutical properties for the spice plant, oregano.

ACKNOWLEDGEMENTS

The authors are also grateful to the IPK Gatersleben, Germany, for providing seed stocks.

Conflict of interest: Authors declare no conflict of interest.

REFERENCES

1. Crocoll C, Asbach J, Novak J, Gershenzon J, Degenhardt J. Terpene synthases of oregano (*Origanum vulgare* L.) and their roles in the pathway and regulation of terpene biosynthesis. Plant Mol Biol 2010; 73 (6):587-603. doi: http://dx.doi.org/10.1007/s11103-010-9636-1
2. Skoula M, Harborne J.B. The taxonomy and chemistry of *Origanum*. In: Oregano, The genera *Origanum* and *Lippia*. Taylor & Francis Publication 2002; 65-108.
3. Bakkali F, Averbeck S, Averbeck D, Idaomar M. Biological effects of essential oils – a review. Food Chem Toxicol 2008; 46(2):446-75. doi: http://dx.doi.org/10.1016/j.fct.2007.09.106
4. D'Antuono LF, Galletti GC, Bocchini P. Variability of essential oil content and composition of *Origanum vulgare* L. populations from a North Mediterranean area (Liguria region, Northern Italy). Ann Bot 2000; 86 (3):471-8. doi: http://dx.doi.org/10.1006/anbo.2000.1205
5. Skoula M, Gotsiou P, Naxakis G, Johnson CB. A chemosystematic investigation on the mono-and sesquiterpenoids in the genus *Origanum* (*Labiatae*). Phytochem 1999; 52(4):649-57. doi: http://dx.doi.org/10.1016/S0031-9422(99)00268-X
6. Scheerer H. Chromosomenzahlen aus der schleswig-holsteinischen Flora. II. Planta. 1940; 30(5):716-25. doi: http://dx.doi.org/10.1007/BF01917180
7. Novak J, Lukas B, Bolzer K, Grausgruber-Gröger S, Degenhardt J. Identification and characterization of simple sequence repeat markers from a glandular *Origanum vulgare* expressed sequence tag. Mol Ecol Resour 2008; 8(3):599-601. doi: http://dx.doi.org/10.1111/j.1471-8286.2007.02059.x
8. Azizi A, Wagner C, Honermeier B, Friedt W. Intraspecific diversity and relationship between subspecies of *Origanum vulgare* revealed by comparative AFLP and SAMPL marker analysis. Plant Syst Evol 2009; 281(1-4):151-60. doi: http://dx.doi.org/10.1007/s00606-009-0197-1
9. Vines G. Herbal harvests with a future: towards sustainable sources for medicinal plants. Plantlife International 2004:3.

10. Canter PH, Thomas H, Ernst E. Bringing medicinal plants into cultivation: opportunities and challenges for biotechnology. Trends Biotechnol 2005; 23(4):180-5. doi: http://dx.doi.org/10.1016/j.tibtech.2005.02.002
11. Bernath J, (ed.). Strategies and recent achievements in selection of medicinal and aromatic plants. International Conference on Medicinal and Aromatic Plants Possibilities and Limitations of Medicinal and Aromatic Plant 2001:576. doi: http://dx.doi.org/10.17660/ActaHortic.2002.576.19
12. Moose SP, Mumm RH. Molecular plant breeding as the foundation for 21st century crop improvement. Plant Physiol 2008; 147(3):969-977. doi: http://dx.doi.org/10.1104/pp.108.118232
13. Gupta PK, Rustgi S, Kulwal PL. Linkage disequilibrium and association studies in higher plants: present status and future prospects. Plant Mol Biol 2005; 57(4):461-85. doi: http://dx.doi.org/10.1007/s11103-005-0257-z
14. Li Y, Li Y, Wu S, Han K, Wang Z, Hou W et al. Estimation of multilocus linkage disequilibria in diploid populations with dominant markers. Genetics 2007; 176(3):1811-21. doi: http://dx.doi.org/10.1534/genetics.106.068890
15. Achleitner A, Tinker NA, Zechner E, Buerstmayr H. Genetic diversity among oat varieties of worldwide origin and associations of AFLP markers with quantitative traits. Theor Appl Genet 2008; 117(7):1041-53. doi: http://dx.doi.org/10.1007/s00122-008-0843-y
16. Falush D, Stephens M, Pritchard JK. Inference of population structure using multilocus genotype data: dominant markers and null alleles. Mol Ecol Notes. 2007; 7(4):574-8. doi: http://dx.doi.org/10.1111/j.1471-8286.2007.01758.x
17. Hardy OJ. Estimation of pairwise relatedness between individuals and characterization of isolation-by-distance processes using dominant genetic markers. Mol Ecol. 2003; 12(6):1577-88. doi: http://dx.doi.org/10.1046/j.1365-294.x
18. Yu J, Pressoir G, Briggs WH, Bi IV, Yamasaki M, Doebley JF et al. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat Genet 2006; 38(2):203-8. doi: http://dx.doi.org/10.1038/ng1702
19. Azizi A, Hadian J, Gholami M, Friedt W, Honermeier B. Correlations between genetic, morphological, and chemical diversities in a germplasm collection of the medicinal plant *Origanum vulgare* L. Chem Biodivers 2012; 9(12):2784-801. doi: http://dx.doi.org/10.1002/cbdv.201200125
20. Azizi A, Yan F, Honermeier B. Herbage yield, essential oil content and composition of three oregano (*Origanum vulgare* L.) populations as affected by soil moisture regimes and nitrogen supply. Ind Crop Prod 2009; 29(2):554-61. doi: http://dx.doi.org/10.1016/j.indcrop.2008.11.001
21. Adams RP. Systematics of the one seeded *Juniperus* of the eastern hemisphere based on leaf essential oils and random amplified polymorphic DNAs (RAPDs). Biochem Syst Ecol 2000; 28(6):529-43. doi: http://dx.doi.org/10.1016/S0305-1978(99)00096-4
22. Nan P, Peng S, Shi S, Ren H, Yang J, Zhong Y. Interpopulation congruence in Chinese *Primula ovalifolia* revealed by chemical and molecular markers using essential oils and ISSRs. Zeitschrift für Naturforschung C. 2003; 58(1-2):57-61. doi: http://dx.doi.org/10.1515/znc-2003-1-210
23. Ardalani H, Eradatmand Asli D, Moradi P. Physiological and morphological response of lemon balm (*Melissa officinalis* L.) to prime application of salicylic hydroxamic acid. Electr J Biol 2014; 10(3): 93-97.
24. Abdossi V, Ghahremani A, Hadipanah A, Ardalani H, Aghaee K. Quantitative and qualitative responses in chemical composition of three ecotypes of fennel (*Foeniculum vulgare* Mill.) cultivated in Iran climatic conditions. J Biodivers Environ Sci 2015; 6(3):401-407.
25. Vogel JM, Scolnik PA. Direct amplification from microsatellites: detection of simple sequence repeat-based polymorphisms without cloning. DNA Markers Protocols, Applications, and Overviews, Wiley 1997:133-150.
26. Doyle JJ. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull 1987; 19:11-5.
27. Singh A, Chaudhury A, Srivastava P, Lakshmikumaran M. Comparison of AFLP and SAMPL markers for assessment of intra-population genetic variation in *Azadirachta indica* A. Juss Plant Science 2002; 162(1):17-25. doi: http://dx.doi.org/10.1016/S0168-9452(01)00503-9
28. Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES. TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 2007; 23(19):2633-5. doi: http://dx.doi.org/10.1093/bioinformatics/btm308

29. Sarwat M, Das S, Srivastava P. Analysis of genetic diversity through AFLP, SAMPL, ISSR and RAPD markers in *Tribulus terrestris*, a medicinal herb. Plant Cell Rep 2008; 27(3):519-28. doi: http://dx.doi.org/10.1007/s00299-007-0478-5

ANALIZA STATYSTYCZNA ZWIĄZKU MONOTERPENÓW FENOLOWYCH I MARKERÓW MOLEKULARNYCH AFLP I SAMPL LEBIODKI POSPOLITEJ (OREGANO)

ALI AZIZI[1], HAMIDREZA ARDALANI[2*], BERND HONERMEIER[3]

[1]Department of Horticultural Sciences
Faculty of Agriculture
Bu-Ali Sina University
Hamedan, Iran

[2]Department of Horticultural Sciences
Science and Research Branch
Islamic Azad University
Tehran, Iran

[3]Institute of Agronomy and Plant Breeding
Biomedical Research Centre Seltersberg (BFS)
Justus Liebig University Giessen
Schubert-Str. 81
D-35392 Gießen, Germany

[*] autor, do którego należy kierować korespondencję: tel.: +98 9128842797,
faks: +988132532432, e-mail: hrardalani@srbiau.ac.ir

S t r e s z c z e n i e

Wstęp: Markery molekularne stanowią przykład udziału technologii genomowej w hodowli roślin leczniczych za pomocą selekcji typu MAS, prowadzonej na potrzeby przemysłu farmaceutycznego. **Cel:** Zbadano 42 rośliny *Origanum vulgare* L. pochodzące z Europy w celu wykrycia zmienności genomu i chemotypu oraz określenia potencjalnych zależności między nimi. **Metody:** Oceniono 477 polimorfizmów molekularnych stosując technikę AFLP (Amplified Fragment Length Polymorphism) dla 214 polimorfizmów i metodę SAMPL (Selectively Amplified Microsatellite Polymorphic Loci) dla 263 polimorfizmów. Za pomocą chromatografii gazowej (GC) zidentyfikowano i określono składniki olejków eterycznych oraz zbadano ich dwa główne komponenty, monoterpeny o znaczeniu ekonomicznym: karwakrol i tymol). **Wyniki:** Na podstawie uzyskanych wyników

stwierdzono stosunkowo wysoką korelację pomiędzy wzorem chemotypowym a markerami genetycznymi. Związek między badanymi cechami olejków eterycznych (zawartością karwakrolu i tymolu) a markerami genetycznymi oceniono za pomocą pięciu metod statystycznych stosując trzy ogólne modele liniowe (General Linear Model – GLM) i dwa liniowe modele mieszane (Mixed Linear Model – MLM). Wykazano istotne zależności między markerami genetycznymi (3 AFLP i 20 SAMPL) a trzema kluczowymi cechami surowca: zawartością olejku eterycznego, karwakrolu i tymolu. **Wnioski:** Stwierdzone zależności mogą wskazywać na przydatność badań molekularnych w hodowli roślin. Uzyskane wyniki stanowią podstawę dla zastosowania markerów molekularnych w selekcji *O. vulgare* do celów farmaceutycznych.

Słowa kluczowe: *Origanum vulgare, olejki eteryczne, AFLP, SAMPL, sprzężenie marker-cecha*