Investigation of the Influence of Genetic Profile on the Economic Characteristics of Lavender Fields

Mariya Zhelyazkova * and Veselina Badzhelova *

Posted Date: 29 April 2026

doi: 10.20944/preprints202604.2001.v1

Keywords: *Lavandula angustifolia* Mill.; SCoT markers; Field Quality Index (FQI); genetic diversity; Marker-Assisted Selection (MAS); essential oil quality and yield

 doi:10.20944/preprints202604.2001.v1

Article

Investigation of the Influence of Genetic Profile on the Economic Characteristics of Lavender Fields

Mariya Zhelyazkova [1,*] and Veselina Badzhelova [2,*]

[1] Trakia University, Faculty of Agriculture, Department of Fundamental Sciences in Animal Husbandry, Studentski Grad, Stara Zagora, 6000 Bulgaria

[2] Agricultural Academy, Institute for Roses and Aromatic Plants, Kazanlak, Boulevard „Osvobojdeni "49, 6100 Bulgaria

* Correspondence: mariya.zhelyazkova@trakia-uni.bg (M.Z.)

Abstract

Lavender (*Lavandula angustifolia* Mill.) is a globally significant crop, with Bulgaria maintaining a leading position in essential oil production. This study presents the first comprehensive, multi-regional analysis of commercial lavender plantations in Bulgaria, integrating phenotypic, biochemical, and genetic data. A novel Field Quality Index (FQI) was developed to objectively quantify production efficiency by balancing yield, essential oil quality, and intra-field homogeneity. Genetic profiling of 285 individual plants via Start Codon Targeted (SCoT) markers revealed significant genetic diversity and a population structure derived from two primary clusters (Delta K = 2), with high intra-field heterogeneity (64%). Our results demonstrate that peak FQI values are achieved in fields with moderate genetic diversity (Genetic Homogeneity Index HI = 0.6–0.7) and high polymorphic information content (PIC ≥ 0.35), whereas excessive clonal uniformity compromises both yield and phytochemical complexity. Regions in Northeastern Bulgaria (Shumen, Shabla) outperformed traditional areas, showing superior resilience to heat stress, which was found to suppress linalool biosynthesis while stimulating trans-ocimene accumulation. Association analysis identified six SCoT loci with high potential for Marker-Assisted Selection (MAS), explaining up to 33.97% of the variation in key terpenoids. These findings establish the FQI as a robust tool for genome-informed management and provide a strategic framework for the sustainable production of high-value lavender oil in a changing climate. The SCoT markers demonstrate substantial potential for practical use in assessing yield and quality, as well as for integration into breeding programs aimed at advancing lavender production.

Keywords: *Lavandula angustifolia* Mill.; SCoT markers; Field Quality Index (FQI); genetic diversity; Marker-Assisted Selection (MAS); essential oil quality and yield

1. Introduction

Lavender is a strategic crop for Bulgaria, providing high value-added raw materials for the perfume, cosmetics, and pharmaceutical industries [1–4]. Historically, cultivation was concentrated primarily in Central Bulgaria as a result of national zoning policies for oilseed crops [5]. However, over the past two decades, lavender acreage has expanded significantly across almost all regions of the country. Many of these new plantations were established haphazardly, using planting material of unknown origin and characteristics, leading to substantial deviations from established essential oil quality standards [6,7]. Such practices have negatively impacted the competitiveness and marketability of Bulgarian lavender oil on global markets [8,9].

In addition to the issues regarding the authenticity of planting material, contemporary lavender production in Bulgaria faces the challenges posed by global climate change. Recent studies indicate that extreme temperature fluctuations and prolonged droughts during the budding and flowering stages directly impact the chemical profile of the essential oil, altering the ratio between the primary

 doi:10.20944/preprints202604.2001.v1

components, linalool and linalyl acetate [9,10]. This necessitates a re-evaluation of the adaptive capacities of established Bulgarian cultivars under the conditions of new agro-climatic regions, where soil characteristics deviate from those traditional to the Sub-Balkan valleys [11].

Parallel to abiotic factors, the uncontrolled distribution of planting material from uncertified nurseries has led to an increased phytosanitary risk. There is a growing prevalence of economically hazardous pathogens, such as *Phoma lavandulae* and phytoplasma infections, which cause premature wilting of the shrubs and compromise entire plantations [12]. In this context, the resistance of Bulgarian cultivars to diseases and pests, combined with their genetic stability, has become a critical factor in maintaining the country's leading position in the international market [13,14].

To date, there is an absence of published data characterizing the genetic profile of these production areas. A critical challenge facing the sector is the lack of effective, routine quality control methods within the market for lavender planting material. Utilizing homogeneous planting material is essential for ensuring high yields and premium essential oil quality [11–15]. Nevertheless, scientific studies investigating how the heterogeneous genetic structure of lavender fields influences overall yield and quality remain scarce.

Essential oil derived from *Lavandula angustifolia* Mill. is widely recognized for its superior quality, primarily due to its significantly lower camphor content compared to *Lavandula latifolia* Medikus or the hybrid *Lavandula x intermedia* (LI) Emeric ex Loiseleur [16]. All Bulgarian lavender varieties are the result of rigorous selection within this species at the Institute of Roses, Essential and Medical Cultures (IREMC), aimed at developing varieties with optimal agronomic and technological traits. Despite these efforts, variations in oil yield and quality between different fields and seasons remain high, rendering production unpredictable and complicating export planning [11,15,17].

In this regard, Start Codon Targeted (SCoT) molecular markers represent a reliable, rapid, and cost-effective tool for assessing genetic diversity and population structure in aromatic and essential oil plants [18–23]. While modern Single Nucleotide Polymorphism (SNP) based methods offer higher resolution, SCoT analysis remains particularly effective for exploring the relationship between genetic variability and phenotypic or economic characteristics under real production conditions. Their application is especially valuable for local populations with limited genomic information [24], such as Bulgarian lavender. Previous studies have discussed the efficiency and capability of SCoT markers for assessing genetic diversity in *L. angustifolia* genotypes [25] and their advantages over other marker systems [26]. SCoT markers target functionally significant regions involved in the regulation of gene expression, demonstrating potential for the development of functional markers and gene characterization [27–30], as well as cultivar identification, phylogenetic analysis, and gene mapping [31]. To date, the construction of a Simple Sequence Repeat (SSR) based genetic linkage map for *L. angustifolia* has led to the mapping of 43 QTLs (Quantitative Trait Loci) associated with the accumulation of 25 different floral volatile compounds, identifying candidate genes for some of them [32]. In recent years, significant genomic milestones have been reached, including the assembly of the mitochondrial genome, the chloroplast genome, and a chromosome-level reference genome of *L. angustifolia* [33–36], along with large-scale SNP identification using Next-Generation Sequencing (NGS) technologies [37]. Furthermore, 207 members of the *WRKY* (*LaWRKY*) gene family have been characterized across the *L. angustifolia* genome; notably, 12 of these genes show high expression in flower buds and calyces in response to both intense light and low temperatures, playing a key role in terpenoid biosynthesis [38].

Despite this robust research foundation, an accessible and efficient marker system for the evaluation of planting material and the quality and yield control of commercial plantations has yet to be proposed.

The present study aims to establish the correlations between genetic structure and field indicators of productivity and essential oil quality. To this end, a new integrated methodology for field assessment has been developed—the Field Quality Index (FQI)—which enables a quantitative and comparable evaluation of the total production efficiency across various lavender plantations.

 doi:10.20944/preprints202604.2001.v1

The necessity for implementing an integrated indicator such as the FQI is dictated by the inherent complexity of secondary metabolism in *Lavandula angustifolia*, which cannot always be unambiguously predicted solely through genetic or agro-climatic data. While SCoT markers reveal polymorphism at the functional level related to translational start codons, field homogeneity and technological parameters reflect the adaptive capacity of the genotype to the specificities of the region. Synthesizing these data into a unified criterion provides a framework for establishing an objective link between genetic structure and the actual economic value of the plantations. This provides breeders and producers with a reliable tool for optimizing cultivar composition and forecasting the market potential of the production.

2. Materials and Methods

2.1. Plant Material

The study was conducted between 2023 and 2025. At the onset of the growing season in April–May 2023, a comprehensive survey of the primary lavender-producing regions in Bulgaria was performed. From each region, three representative commercial fields were selected for inclusion in the study (Figure 1).

To evaluate the quantity and quality of the essential oil from each site, three distinct types of samples were collected per field (s_1, s_2, and s_3). Specifically, s_1 represented a composite (bulk) sample of randomly selected plants reflecting the overall field average, while s_2 and s_3 were collected from individual plants to assess intra-field variability.

Inflorescences were harvested in July-August 2023 during the peak flowering stage (65–75% bloom) to serve as the exclusive material for both steam distillation and subsequent gas chromatography (GC) analysis, ensuring consistency in the evaluation of essential oil yield and chemical composition.

For the molecular analysis, DNA was extracted from 19 randomly selected individual plants per field (including the plants designated as s_2 and s_3) to characterize the genetic structure of the populations.

 doi:10.20944/preprints202604.2001.v1

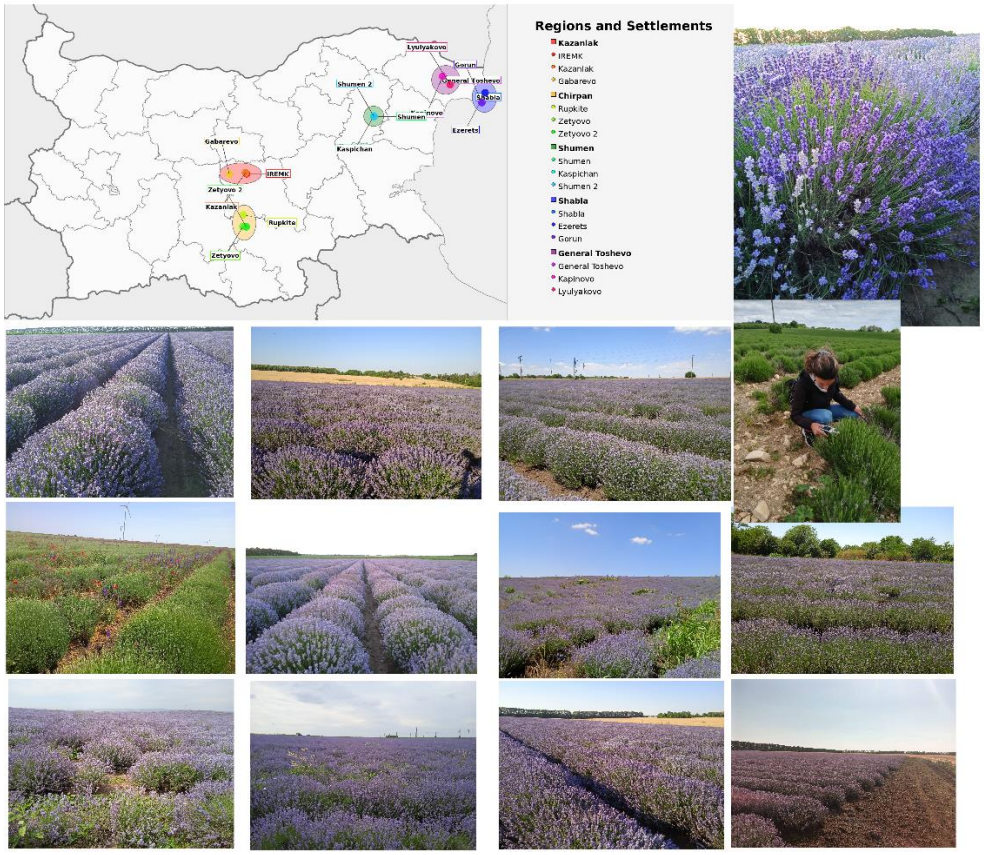


Figure 1. Location map and illustrative images of the sampling sites.

2.1. Essential Oil Distillation and Yield Determination

The study was conducted between July and August 2023, coinciding with the peak flowering stage of the plantations. Fresh inflorescences were harvested under dry and warm weather conditions during the midday hours to ensure optimal essential oil secretion. To determine the quantity of essential oil, the steam distillation method was used with the following parameters: raw material 350 g, hydromodule 1:8 (w/v), duration 1 hour.

2.2. Essential Oil Chemical Characterization

The chemical composition of the essential oils was analysed using an Agilent Technologies 7820A Gas Chromatograph (Agilent Technologies, Santa Clara, CA, USA) equipped with a flame ionization detector (FID) under the following operating parameters: A non-polar capillary column (30 m length × 320 μm internal diameter × 0.25 μm film thickness); The oven temperature was programmed from 60°C to 300°C at a heating rate of 5°C/min;The injector and detector temperatures were maintained at 250°C and 300°C, respectively and Carrier Gas: Hydrogen (H_2) at a constant flow rate of 0.5 mL/min.

Quantitative analysis was performed using the area normalization method, with the relative percentage of each component calculated based on the total peak area (100%). The sample injection volume was 0.1 μL.

2.3. Dna Extraction and Scot-Pcr Amplification

Total genomic DNA was isolated from frozen leaves using a Disruptor Genie® homogenizer (Scientific Industries, Inc., Bohemia, NY, USA) and the Quick-DNA Plant/Seed Miniprep Kit (Zymo Research, Irvine, CA, USA), following the manufacturer's protocol. The quality and concentration of the extracted DNA were checked using a NanoVue Plus spectrophotometer (GE Healthcare UK

 doi:10.20944/preprints202604.2001.v1

Limited, Amersham Place, Little Chalfont, Buckinghamshire, UK) and 1% (w/v) agarose gel electrophoresis.

A set of five SCoT primers were selected for DNA amplification (Table 1). The PCR reactions were performed in a total volume of 20 µL, containing: 1 µL (50 ng) genomic DNA, 10 µL Red Taq DNA Polymerase 2×Master Mix, 1 µL (10 pmol) primer, and 8 µL nuclease-free ddH_2O. The SCoT-PCR amplification was carried out using a Doppio Gradient 2 × 48 well thermal cycler (VWR®, Darmstadt, Germany), as follow: initial denaturation at 94°C for 5 min; 35 cycles of denaturation at 94°C for 45 s, primer- specific annealing at 56 – 61°C for 45 s, and extension at 72°C for 90 s; followed by a final extension at 72°C for10 min. The PCR products (10 µL mixed with 2 µL loading buffer) were separated on 2% (w/v) agarose gel electrophoresis in 1×TBE buffer for 130 minutes at 110 V. NZYDNA Ladder VI (50 - 1500 bp) was used as a molecular weight marker. Gels were stained with GelRed® (Biotium, Fremont, CA, USA) and documented using a UV transilluminator system (Bio-Imaging System, Modi'in-Maccabim-Re'ut, Israel).

Table 1. The SCoT primers used and their specific annealing temperature (sTa°).

Primer	Sequence	sTa°
SCoT 2	CAACAATGGCTACCACCC	56
SCoT 3	CAACAATGGCTACCACCG	56
SCoT 12	ACGACATGGCGACCAACG	61
SCoT 15	ACGACATGGCGACCGCGA	61
SCoT 24	CACCATGGCTACCACCAT	56

2.4. Data Analasis

2.4.1. Agronomic and Economic Performance Indices

To conduct a comparative analysis of the performance across different lavender fields, several integrated indices were developed and calculated:

Field Quality Index (FQI) - An integrated field assessment score composed of three primary components: Quality Index (Q), Yield Index (Y) and Intra-field Homogeneity Index (H).

Quality Index (Q):

The Q index quantifies the compliance of the oil's chemical profile with the reference limits defined by the ISO 3515:2002 standard [39] for each component measured in this study. For each field, an individual conformity index was calculated using the following formula:

$$Q = \prod_i q_i^{\frac{w_i}{\sum_j w_j}}, \quad (1)$$

where:

q_i – conformity of component i; w_i - weighting factor for component i. The weighting factors were normalized to satisfy the following condition:

$$\sum_i w_i = 1 \quad (2)$$

The resulting index (Q) resides in the range [0,1]. The weights ensure that components with higher assigned values exert a more substantial influence on the overall quality index. The specific weighting factors (w_i) used for the Q index calculation were:

Linalool (2.0), Linalyl acetate (2.0), Camphor (2.0), *cis*-Ocimene (1.5), *trans*-ocimene (1.5), Terpinen-4-ol (1.5), Lavandulyl acetate (1.5), 3-Octanone (1.5), Lavandulol (1.0), α-terpineol (1.0)

Yield Index (Y)

For each field, an average yield was calculated from three distinct samples (s_1, s_2, s_3) according to the formula:

 doi:10.20944/preprints202604.2001.v1

$$y_i = \frac{s_1 + \frac{s_2 + s_3}{2}}{2} \quad (3)$$

where:

y_i – average field yield; s_1 – representative bulk sample for the field; s_2, s_3 – individual plant samples. To allow for objective comparisons between fields, the yield values were normalized to a scale of [0,1] via min-max scaling against the observed sample values, using the formula:

$$Y_i = \frac{y_i - y_{min}}{y_{max} - y_{min}} \quad (4)$$

where:

Y_i - normalized field yield; y_{min}- minimum average sample yield observed; y_{max}- maximum average sample yield observed.

Intra-field Homogeneity Index (H)

To assess the intra-field (economic) homogeneity, Principal Component Analysis (PCA) was applied to the observations of the three samples (s_1, s_2, s_3) for each field, considering both yield and essential oil chemical composition. The index is calculated as follows.

$$H_k = 1 - \frac{(d_k - d_{min})}{(d_{max} - d_{min})}, H_k \in [0,1] \quad (5)$$

where:

H_k - Intra-field homogeneity index for field k, with values ranging from [0,1]. A value of 1 indicates a perfectly homogeneous field, while 0 represents maximum internal variability.

d_k- The average Euclidean distance between the three samples (s_1, s_2, s3) for the field k in the space of the first two principal components (PC1, PC2). This distance reflects the disparities between samples regarding yield and/or chemical profile.

d_{min} - The minimum average distance observed among all fields (reference for the most homogeneous field).

d_{max} - The maximum average distance observed among all fields (reference for the most homogeneous field).

Aggregation — Final Integrated Index FQI (Multiplicative)

The final integrated field score, Field Quality Index (FQI), is formed by the three normalized indices (Q, Y, H) through a multiplicative combination with exponential weighting factors.

$$FLQI = 100 * (Q^{alpha} * Y^{beta} * H^{gamma}) \quad (6)$$

where:

alpha, beta, gamma - Assigned weights, which for the purposes of this analysis were set as follows: Quality weight (α) = 0.6; Yield weight (β) = 0.3; Homogeneity weight (γ) = 0.1

2.4.2. Genetic Profiles and Genetic Diversity

The SCoT banding profiles were scored and converted into a binary data matrix based on the presence (1), absence (0), or missing/unclear (-1) bands. To evaluate the efficiency of the selected primers, the following parameters were calculated [25]: Polymorphic Information Content (PIC) by [40]; Effective Multiplex Ratio (EMR) by [41]; Marker Index (MI) by [42]; and Resolving Power (RP) by [43].

The genetic diversity represented by different (Na) and effective (Ne) number alleles, Shannon's Information Index (I), and expected heterozygosity (He) were calculated using GenALEx 6.5 [44]. Nei's genetic distance [45] and gene diversity [46] among all plants was determined in PopGen32. The genetic relationships were further explored by Principal Coordinate Analysis (PCoA) and Analysis of Molecular Variance (AMOVA), all performed in GenALEx v.6.5 [44]. Genetic differentiation among accessions was estimated using Φ_{PT} (*PhiPT*) (an F_{ST} analogue for dominant markers) via AMOVA. The first two principal coordinate axes (PC1 vs. PC2) of Genalex were used for visualization in GraphPad Prism 10.6.1 (GraphPad Software, Boston, MA, USA). An Unweighted Pair Group Method with Arithmetic Mean (UPGMA) cluster tree was constructed using MEGA12 software [47] and visualized using iTOL v.6 [48] to enhance the graphical presentation.

 doi:10.20944/preprints202604.2001.v1

The genetic structure of the population was analysed using STRUCTURE v.2.3.4, employing a Bayesian clustering model to estimate individual genetic ancestry (Q-values) [49]. The optimal number of clusters (K) was tested within a range of K= 2 to K= 15. Each K-value underwent 20 repetitions, with a burn-in iteration of 100,000 followed by 10,000 Markov Chain Monte Carlo (MCMC) replications post burn-in [50]. The resulting date were uploaded to the CLUMPAK web server, http://clumpak.tau.ac.il for determine the optimal K-value based on the Evanno method (ΔK), identifying the most probable number of genetic clusters [51].

2.4.3. Marker-Trait Association Analysis (Glm and Mlm)

Marker–trait associations (MTAs) were identified using TASSEL v.5.2 [52], employing both General Linear Model (GLM) and Mixed Linear Model (MLM) approaches to evaluate the relationships between 15 essential oil chemical components and 90 SCoT loci. For the GLM analysis, "location" was included as a fixed factor, while the Q-matrix (population structure) was used as a covariate to account for the chemotypic and genotypic variation among 30 selected plants (samples s_2 and s_3 from all 15 production fields). In the MLM analysis, a Kinship (K) matrix representing individual relationships was integrated alongside the Q-matrix as a covariate to minimize the detection of false-positive associations. The statistical significance of the associations between the chemical variations in lavender essential oil and the genetic profiles of the plants was evaluated. A significance threshold of $p < 0.05$ was employed to identify potential marker–trait associations, while a more stringent threshold of $p < 0.01$ was used to highlight the most robust and highly significant loci, ensuring the reliability of the findings after multiple testing considerations.

2.4.4. Correlations Between Economic and Genetic Characteristics

To evaluate the relationship between economic performance and genetic profile, a Genetic Homogeneity Index (HI) was developed. This index represents a normalized inversion of Nei's genetic distance [45] and is calculated as follows:

$$HI = 1 - \frac{(Nei_{mean} - Nei_{min})}{(Nei_{max} - Nei_{min})} \quad (7)$$

where:

Nei_{mean} - the average of the Nei's genetic distance for a given field,

Nei_{min}, Nei_{max} - the minimum and maximum Nei's genetic distance values observed across all plants in the same field.

The resulting *HI* values range from [0,1], where higher values indicate greater genetic homogeneity (i.e. lower intra-field variability).

All statistical computations and graphical visualizations were performed in Python v3.13 environment using standard scientific libraries (NumPy, Pandas, Matplotlib, and Seaborn) and software GenAlEx 6.5.

Correlation analysis between the calculated indices was performed using Pearson and Spearman rank correlation methods with two-tailed significance testing at $\alpha = 0.05$.

To account for multiple comparisons, Holm's correction was applied, while model robustness was verified using cross-validation [53,54].

2.4.5. Meteorological Data: Era5-Land Reanalysis

Given the spatial distribution of the experimental fields and the requirement for consistent agro-climatic indicators, meteorological data were retrieved from the ERA5-Land (ERA5L) reanalysis dataset [55,56]. ERA5L provides high-resolution hourly atmospheric and land-surface variables at a spatial resolution of ~9 km, which is highly suitable for assessing environmental impacts on agricultural crops [57–59]. For each experimental site (Kazanlak, Chirpan, Shumen, Shabla, and General Toshevo), hourly data were extracted for the grid cells corresponding to the fields' coordinates.

 doi:10.20944/preprints202604.2001.v1

The raw hourly data were processed to derive monthly aggregates for the critical vegetation period (May–July 2023). The analyzed variables included: mean 2-meter air temperature (t^2m), 2-meter dewpoint temperature (d^2m), total precipitation (tp), surface solar radiation downwards (ssrd), and wind speed at 10 meters (wind). Additionally, volumetric soil water (layer 1: 0–7 cm, swvl1) was included to evaluate the impact of soil moisture on lavender performance. All time-series data were adjusted for the local time zone (UTC+2). These variables were utilized to establish the environmental context for the Field Quality Index (FQI) and to perform the correlation analysis presented in the study.

Correlation matrices and heatmaps were further processed and refined using GraphPad Prism v11.0.0 (GraphPad Software, Boston, MA, USA).

3. Results

This section presents the findings of the integrated analysis of chemical, phenotypic, and genetic traits across fifteen lavender fields located in five major Bulgarian production regions (Table 2). By combining quantitative assessments of essential oil quality (*Q*), yield (*Y*), intra-field homogeneity (*H*), and genetic structure derived from SCoT markers, we investigated the correlation between genetic diversity and the agronomic performance of the plantations.

Table 2. Location details, regional distribution, and essential oil yield parameters of the studied lavender plantations.

Region	Location	Field ID	Coordinates	Quantity of essential oil (ml)		
				S_1	S_2	S_3
Kazanlak	Kazanlak	IREMK	42.637189, 25.388629	4.7	7	6.3
	Kazanlak	KS	42.646777, 25.375778	4.9	5.2	5.4
	Gabarevo	GB	42.634043, 25.173031	4.6	4	3.3
Chirpan	Zetyovo	ZA	42.1210759, 25.3431017	4.3	4	4.5
	Rupkite	Rup	42.2357990, 25.3477580	3	3.3	3.5
	Zetyovo	ZB	42.1190038, 25.3854098	4.9	3.7	4.5
General Toshevo	Lyulyakovo	MS	43.6577236, 28.0889702	10.4	8.6	10.5
	General Toshevo	MY	43.718357, 28.019853	8.7	2.4	9.6
	Kapinovo	KGT	43.739277, 27.991201	10.9	9.4	10.5
Shabla	Gorun	Gor	43.4837457, 28.4834800	8.7	7.4	8.8
	Ezerets	EZ	43.5806808, 28.5258270	8.1	4.1	8.1
	Shabla	Sh	43.5511740, 28.5400645	12.3	12.8	10.3
Shumen	Shumen	VP	43.2998906, 27.0798766	11.5	11.7	7.3
	Kaspichan	KSP	43.303020, 27.079809	13.2	12.4	11.6
	Shumen	VSP	43.308879, 27.085359	8.7	7.3	8.6

1 S1 - Composite Plant Sample, S2, S3 - Randomly Chosen Plant Samples.

3.1. Economic Characteristics of the Fields

3.1.1. New Integrated Indices

The comprehensive evaluation of the production fields was performed using integrated Field Quality Index (FQI). As defined in the methodology, the FQI is a multiplicative function of three primary components: The Quality Index (Q), the Yield Index (Y), and the Intra-field Homogeneity Index (H), with exponential weights of 0.6, 0.3, and 0.1, respectively. This integrated approach allows for a simultaneous capture of the chemical, agronomic, and technological performance of lavender plantations.

 doi:10.20944/preprints202604.2001.v1

The results reveal a clear regional differentiation in both quality and productivity (Table 3, Figure 2). Fields located in the Shumen, General Toshevo and Shabla regions (Northeastern Bulgaria) exhibited the highest FQI values, ranging from 61.76 to 87.45. These locations effectively combined high essential oil yields (Y range: 0.56 - 1.00) with excellent chemical profiles (Q range: 0.90 – 1.00). Notably, field KSP (from the Shumen region) achieved the highest overall score (FQI = 87.45), driven by a superior Yield index (Y = 1.00). Fields from the Shumen and Shabla regions (VP and Sh) are characterized by a near-optimal balance between quality (Q = 0.97 – 0.99) and yield (Y ≥ 0.82), reaching FQI value of 81.46 and 86.81, respectively. In contrast, fields from the Kazanlak and Chirpan regions exhibited lower yield indices (Y range: 0.25 - 0.45), despite maintaining high oil quality (Table 3). Across the study, the indices varied as follows: Q ranged from 0.76 (GB) to 1.00 (KS, ZA, Rup, ZB, KGT, Gor, EZ); Y ranged from 0.25 (Rup) to 1.00 (KSP); and H range from 0.10 to 0.31. Homogeneity index values exceeding 0.30 (e.g., KSP and Sh) indicate a stable and predictable chemical profile, whereas values below 0.25 reflect higher spatial heterogeneity. Such variability suggests a need for enhanced technological precision during harvesting and distillation to maintain batch consistency. The lowest FQI scores among the 15 studied fields were recorded for GB (FQI = 56.42), Rup (FQI = 56.95) and IREMK (FQI = 58.49).

Table 3. Economic assessment and integrated performance indices of the experimental lavender fields, Yield (Y), Quality (Q), Intra-field Homogeneity (H), and Field Quality Index (FQI).

Region	Field ID	Yield,mL	Y	Q	H	FQI
Kazanlak	IREMK	5.68	0.45	0.81	0.18	58.49
	KS	4.12	0.33	1.00	0.21	61.30
	GB	5.10	0.40	0.76	0.26	56.42
Chirpan	ZA	4.28	0.34	1.00	0.20	61.64
	Rup	3.20	0.25	1.00	0.22	56.95
	ZB	4.50	0.36	1.00	0.24	63.68
General Toshevo	MS	9.98	0.79	0.90	0.22	75.17
	MY	7.35	0.58	0.99	0.22	72.88
	KGT	10.42	0.83	1.00	0.01	61.76
Shabla	Gor	8.40	0.67	1.00	0.29	78.11
	EZ	7.10	0.56	1.00	0.16	70.2
	Sh	11.92	0.95	0.99	0.31	86.81
Shumen	VP	10.35	0.82	0.99	0.24	81.46
	KSP	12.60	1.00	0.97	0.31	87.45
	VSP	8.32	0.66	0.99	0.10	69.64

The mean FQI values, aggregated by production region, are illustrated in Figure 2. A clear spatial trend is observed, where fields located in the northeastern regions (Shumen and Shabla) exhibit the highest composite scores, whereas field in the southern regions (Kazanlak and Chirpan) show lower overall values, primarily due to higher within-field variability and lower yield indices, despite their stable essential oil quality, Supplementary Table S1.

 doi:10.20944/preprints202604.2001.v1

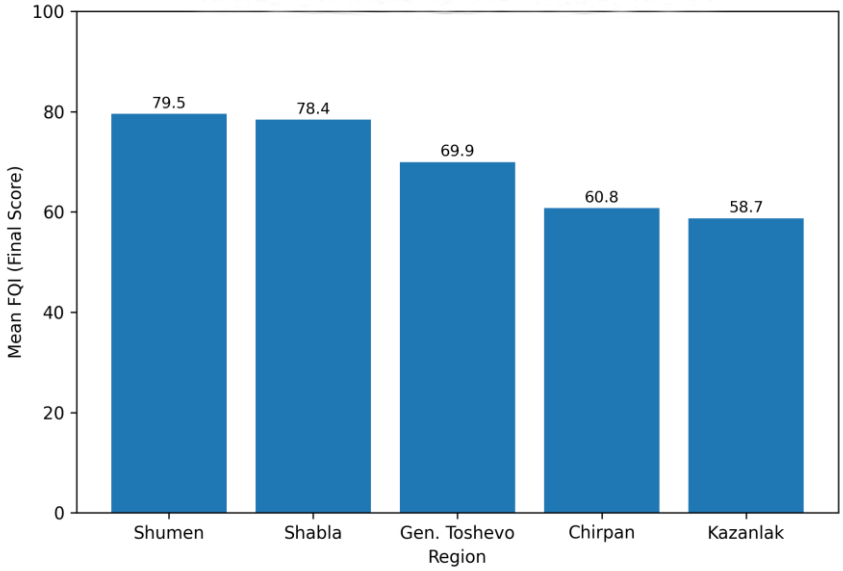


Figure 2. Regional evaluation of the integrated field quality index (FQI).

3.1.2. Analysis of Index Interrelationships and Integrated Performance

The correlation analysis among the three primary components (Q, Y, and H) revealed weak to moderate correlations ($|r| < 0.30$), confirming that these parameters describe largely independent aspects of production efficiency (Figure 3).

Essential oil quality (Q) and yield (Y) exhibited a weak positive relationship ($r = 0.15$), indicating that high chemical quality does not inherently guarantee high agronomic productivity. This independence is expected, as chemical composition and essential oil yield are governed by distinct physiological and genetic mechanisms. The Intra-field Homogeneity Index (H) showed a minimal association with yield ($r = 0.09$) and a slightly negative relationship with quality ($r = -0.12$), supporting its interpretation as a stabilizing factor rather than a direct driver of productivity. However, its influence on the Final Integrated Index (FQI) was moderate but statistically significant ($r = 0.48$), highlighting its role as a key indicator of predictability and technological reliability of the plantations (Figure 3). Fields with $H \geq 0.30$ demonstrated superior repeatability of results and lower internal variation in their chemical profile. In contrast, fields with $H < 0.25$ exhibited pronounced deviations among samples, increasing the environmental and technological risk of inconsistent harvesting outcomes.

The FQI showed the strongest correlation with yield ($r = 0.84$), followed by homogeneity ($r = 0.48$) and quality ($r = 0.34$), Figure 3. This pattern reflects the broader dispersion of yield values within the dataset and identifies productivity as a primary economic driver in the studied regions. Nevertheless, the higher weight assigned to quality in the FQI formula ($\alpha = 0.60$) mathematically compensates for this statistical imbalance, ensuring that the final assessment prioritizes chemical excellence alongside agronomic performance.

 doi:10.20944/preprints202604.2001.v1

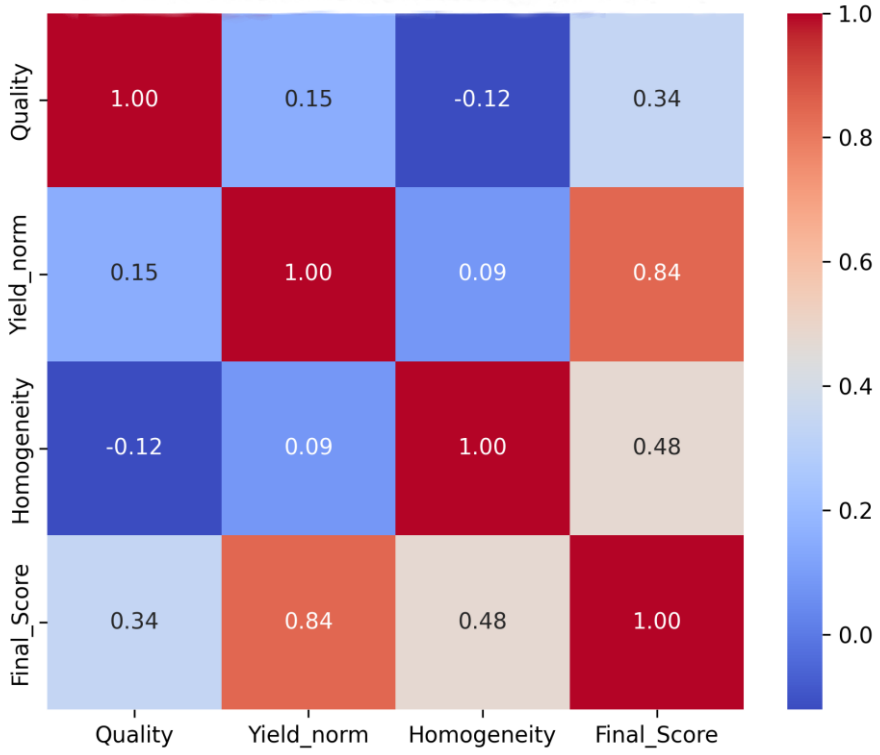


Figure 3. Heatmap of the correlation matrix between Quality (*Q*), Yield (*Y*), Homogeneity (*H*), and the integrated Field Quality Index (FQI).

3.2. Scot Marker Efficiency and Genetic Diversity

3.2.1. Scot-Amplification Results

The five SCoT markers used for 285 lavender plants successfully amplified a total of 90 bands, of which 89 were polymorphic, resulting in a high average polymorphism of 98.94% (Table 4). Across all primers, the number of polymorphism amplified bands ranged from 14 (SCoT 2) to 21 (SCoT 15). The Polymorphic Information Content (PIC) value ranged between 0.41 (SCoT 2) and 0.80 (SCoT 24), with an average value of 0.64. Furthermore, the calculated mean values for Effective Multiplex Ratio (EMR), Resolving Power (RP), and Marker Index (MI) were 147.28, 18.4, and 90.72, respectively. The maximum and minimum Shannon's information index (I) were 0.570 (SCoT 12) and 0.475 (SCoT 24), with average value of 0.503. The highest gene diversity (H) was recorded for SCoT 12 (0.409), while the minimum was observed for SCoT 2 (0.3111), and a mean value of 0.336 across all SCoT tested markers. These values confirm the high efficiency of the selected markers for detecting the genetic diversity of the studied lavender accessions. SCoT-PCR fingerprinting patterns generated with different SCoT markers in the lavender commercial fields are shown in Figure 4.

Table 4. The SCoT primers used, Specific annealing temperature (sTa°), Total bands (TB), Polymorphic bands (PB), Monomorphic bands (MB), Resolving power (Rp), Polymorphic information content (PIC), Effective multiplex ratio (EMR), Marker index (MI).

Primer	TB	PB	%PB	Rp	PIC	EMR	Mi	N	Ne	H	I
SCoT 2	14	14	100	20.77	0.41	211.4	87.66	268.92	1.506	0.3111	0.478
SCoT 3	19	18	94.7	19.70	0.64	140.0	89.81	278.84	1.543	0.319	0.480
SCoT 12	17	17	100	16.50	0.72	138.1	99.71	280.88	1.744	0.409	0.570
SCoT 15	21	21	100	21.5	0.65	145.7	95.10	278.00	1.567	0.376	0.493
SCoT 24	19	19	100	13.49	0.80	101.2	81.30	280.57	1.526	0.313	0.475

 doi:10.20944/preprints202604.2001.v1

Total	90	89								
Mean			98.94	18.39	0.64	147.28	90.72	1.577	0.35	0.50

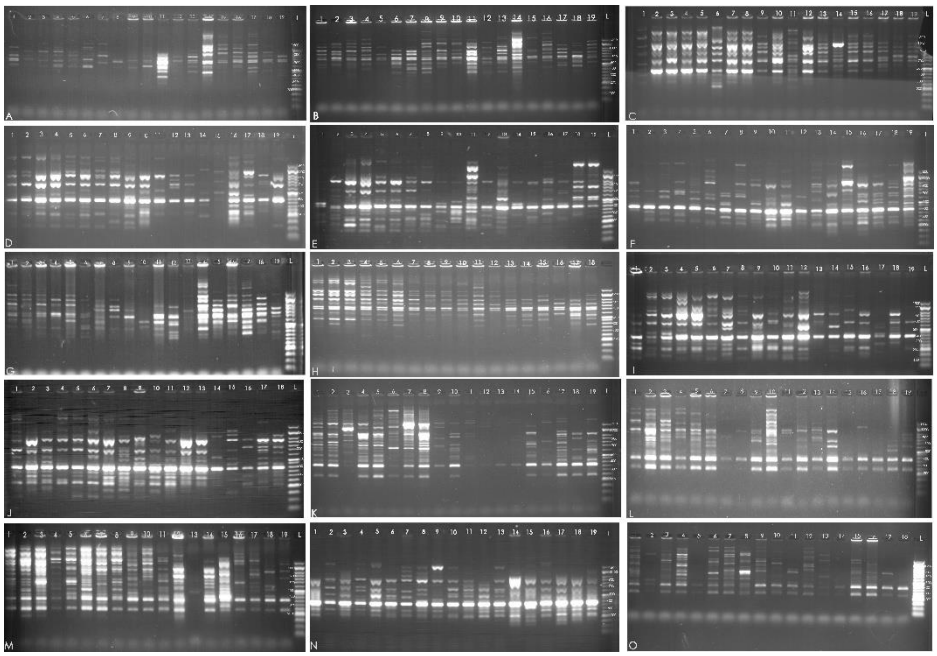


Figure 4. SCoT-PCR fingerprinting patterns generated by SCoT markers for 19 lavender plants in the 15 commercial fields from Bulgaria. The plants of the fields are presented as follow: A: IREMK1-19; B: KS1-19; C: GB1-19; D: ZA1-19, E: Rup1-19; F: ZB1-19, G: MS1-19; H: MY1-19, I: KGT1-19; J: Gor1-19; K: EZ1-19; L: Sh1-19, M: VP1-19; N: KSP1-19, O: VSP1-19; The markers are presented as follow: SCoT 3: K, L, M, O; SCoT 12: D, E, F, I, J, N; SCoT 24: A, B, C, G, H. For all gels NZYDNA Ladder VI (50–1500 bp) was sign as L.

3.2.2. Genetic Diversity and Differentiation Among Lavender Fields in Bulgaria

Genetic diversity indices were calculated for 19 individual plants per field across the 15 studied location (Table 5). The accessions from the General Toshevo region (MS, MY and KGT) exhibited the highest genetic uniformity, as evidenced by the lower values for the percentage of polymorphic loci (PB: 46.67% - 48.89%), Shannon's Information Index (I: 0.266 - 0.283), and expected heterozygosity (He: 0.169 - 0.193). In contrast, the highest levels of genetic diversity were recorded in the Shabla region, specifically in field EZ (PB = 71.11%, I = 0.384, He = 0.259). The lowest overall diversity parameters were observed in field GB (Gabarevo) within the Kazanlak region. Generally, lavender acessions in the Shabla and Shumen regions demonstrated higher genetic diversity (PB % ≥ 62.22, I = 0.320 - 0.384, He = 0.215 - 0.259) compared to those in the Kazanlak and Chirpan regions (PB ≤ 62.22%, I = 0.246 - 0,340, He = 0.167 - 0.229). The observed number (Na) and the effective number of alleles (Ne) across all fields ranged from 1.156 (field GB) to 1.489 (field EZ), and from 1.292 (field MY) to 1.454 (field EZ), respectively. These results, derived from the analysis of 285 individual plants, reveal significant variations in genetic diversity levels among the productive lavender fields in Bulgaria, Supplementary Table S2. The higher diversity in the Shabla/Shumen regions likely reflects a broader genetic base or the use of diverse seedling material, whereas the uniformity in General Toshevo suggests more standardized clonal propagation.

Table 5. Genetic diversity indices for the 15 lavender fields studied based on SCoT marker analysis. Number of individual (N), different (Na) and effective (Ne) alleles; Shannon's information index (I); Expected heterozygosity (He), Percentage of polymorphic loci (PB%) Standard errors are shown in parentheses.

Fields ID	N	Na	Ne	I	He	PB %
IREMK	18.700 (0.070)	1.400 (0.083)	1.367 (0.041)	0.315 (0.031)	0.212 (0.022)	58.89
KS	18.733 (0.079)	1.311 (0.082)	1.341 (0.042)	0.282 (0.032)	0.192 (0.022)	50.00

 doi:10.20944/preprints202604.2001.v1

GB	17.422 (0.501)	1.156 (0.089)	1.294 (0.040)	0.246 (0.031)	0.167 (0.022)	44.44
ZA	18.700 (0.100)	1.267 (0.089)	1.349 (0.041)	0.293 (0.032)	0.199 (0.022)	52.22
Rup	18.789 (0.068)	1.433 (0.084)	1.392 (0.039)	0.340 (0.031)	0.229 (0.021)	62.22
ZB	18.678 (0.075)	1.289 (0.087)	1.353 (0.042)	0.292 (0.032)	0.199 (0.022)	52.22
MS	18.456 (0.224)	1.256 (0.086)	1.338 (0.041)	0.283 (0.032)	0.193 (0.022)	48.89
MY	18.389 (0.231)	1.278 (0.084)	1.292 (0.040)	0.252 (0.030)	0.169 (0.021)	48.89
KGT	18.311 (0.333)	1.211 (0.087)	1.320 (0.041)	0.266 (0.032)	0.182 (0.022)	46.67
Gor	18.678 (0.075)	1.411 (0.088)	1.433 (0.043)	0.356 (0.032)	0.243 (0.022)	63.33
EZ	18.289 (0.214)	1.489 (0.088)	1.454 (0.041)	0.384 (0.030)	0.259 (0.021)	71.11
Sh	18.722 (0.067)	1.322 (0.096)	1.402 (0.041)	0.340 (0.031)	0.231 (0.022)	62.22
VP	18.444 (0.107)	1.389 (0.088)	1.373 (0.041)	0.320 (0.031)	0.215 (0.022)	61.11
KSP	18.822 (0.064)	1.456 (0.087)	1.432 (0.041)	0.364 (0.030)	0.247 (0.022)	66.67
VSP	18.711 (0.079)	1.378 (0.093)	1.385 (0.039)	0.339 (0.030)	0.227 (0.021)	64.44
Grand mean	18.523 (0.051)	1.336 (0.023)	1.368 (0.011)	0.311 (0.008)	0.211 (0.006)	56.89

3.2.3. Analysis of Molecular Variance (Amova)

To determine the genetic variation among and within the lavender fields studied, an analysis of molecular variance (AMOVA) was performed (Table 6). The results indicate that the most of total genetic variation (64%) is retained within the accessions in each field ($p < 0.001$), while a significant portion is attributable to genetic differences among fields (36%) ($p < 0.001$). These results demonstrated that a higher of genetic variation exsist at the intra-field level. This is further supported by the Φ_{PT} (*PhiPT*) value of 0.356, indicating a high level of population differentiation among the lavender accessions.

Table 6. Analysis of molecular variance (AMOVA) of studied lavender accessions, df = degrees of freedom; SS = sum of squares; MS = mean squares; Est. Var. = estimated variance; % Var. = percentage of total variance.

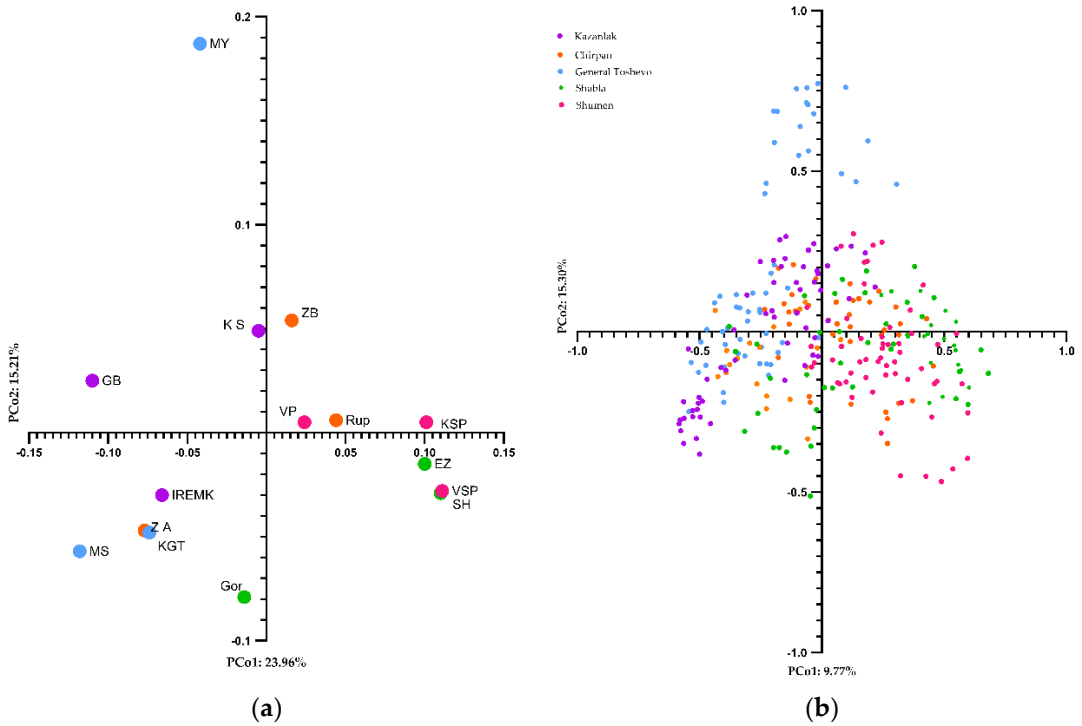
Source	df	SS	MS	Est. Var.	% Var	*p* Value
Among Fields	14	1749.754	124.982	6.006	36%	<0.001
Within Fields	270	2934.737	10.869	10.869	64%	<0.001
Total	284	4684.491		16.875	100%	

3.2.4. Principal Coordinate Analysis (Pcoa) and Genetic Structure

To assess the genetic relationships and population structure among the 285 lavender accessions from 15 fields, PCoA, UPGMA clustering, and STRUCTURE analysis were applied.

The first three axes of the PCoA based on genetic distances derived from SCoT-PCR profiles explained 52.98% of the total variance among 15 fields with individual contributions of: 23.96%, 15.21%, and 13.81% (Figure 5a). For the individual-level analysis of 285 plants, the first three axes accounted for 20.38% of the cumulative variation: 9.77%, 15.30%, and 20.38%, respectively (Figure 5b). The results demonstrate a significant genetic overlap among individual plants (Fig. 5b), indicating a shared and cohesive genetic pool across the studied Bulgarian lavender accessions. Despite this overlap, a clear regional trend emerges when observing the fields (Fig. 5a). This suggests that while the base germplasm is similar, local selection and environmental adaptation have led to the emergence of distinct regional genetic signatures. Specifically, accessions from the Kazanlak region (IREMK, KS and GB; purple) и and General Toshevo (MS, MY, and KGT; blue) were clearly separated from those in the Shumen (VP, KSP, and VSP; red) and Shabla (EZ and SH; green) regions. Fields from the Chirpan region (ZA, Rup, and ZB; orange) were positioned between the two main clusters. Notably, field MY (Malina Y) stood out as a genetic outlier, exhibiting a unique profile that separated it from other regional representatives along the upper part of the second axis.

 doi:10.20944/preprints202604.2001.v1



(**a**) (**b**)

Figure 5. Principal coordinate analysis (PCoA) biplot of the studied lavender accessions: (**a**) Distribution of 15 fields illustrating regional differentiation; (**b**) Distribution of 285 individual plants showing genetic overlap and intra-field variation. Colors represent different production regions: purple (Kazanlak), orange (Chirpan), blue (General Toshevo), red (Shumen) and green (Shabla).

The subsequent cluster analysis based on Nei's genetic distances reveals a complex, non-linear distribution of the studied accessions. Individuals from the same production fields do not form exclusive monophyletic groups. Instead, they are distributed across multiple sub-clusters, confirming the high intra-field genetic variation observed in the AMOVA and PCoA analyses. This mosaic structure underscores the shared genetic heritage of Bulgarian lavender and suggests that the current plantations are a result of significant germplasm exchange and potential multi-clonal compositions within individual fields. Plants from the fields in Chirpan (ZB), Shabla (Gor), and Kazanlak (KS) grouped together in Cluster 1. Cluster 2 integrated plants from Chirpan (Rup), Shumen (KSP, VS), and Shabla (EZ, Sh). Plants from General Toshevo (MY, MS, KGT), Kazanlak (IREMK, GB), and Chirpan (ZA) were primarily distributed in Cluster 3. Notably, plants from the VP field were identified as a distinct, independent clone. Furthermore, specific individual genotypes from one field were frequently found nested within groups of another field, originating from different geographic regions. Examples include genotypes ZB15, EZ2, EZ3, and EZ4; VS5, VS8, and VS9; Sh13; KSP9; Gor4, Gor5, and Gor 19; IREMK17, IREMK18, and IREMK 19; Gor17 and Gor 18.

 doi:10.20944/preprints202604.2001.v1

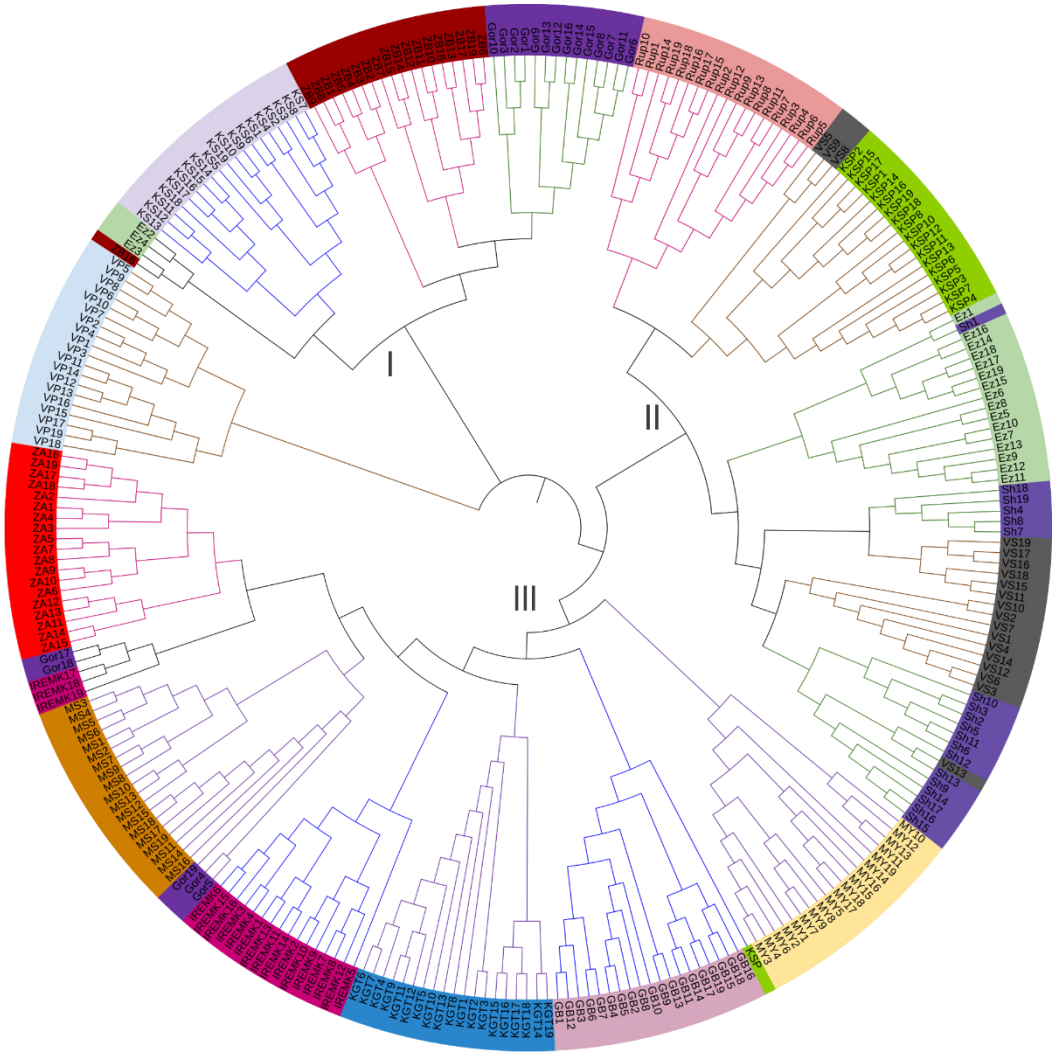


Figure 6. UPGMA dendrogram of 285 lavender plants based on Nei's genetic distance. The colors represent 15 different production fields, illustrating the high degree of genetic admixture and intra-field heterogeneity. The plants of the fields are presented as follow Kazanlak: IREMK1-19, KS1-19 and GB1-19; Chirpan: ZA1-19, Rup1-19 and ZB1-19; General Toshevo: KGT1-19, MS1-19 and MY1-19; Shumen: VS1-19, VP1-19 and KSP1-19; Shabla: Sh1-19, Gor1-19 and EZ1-19.

The population structure analysis was performed, and Evanno's ΔK method was used to determine the optimal number of genetic clusters (K). The highest ΔK value was observed at K = 2 (Figure 7). These results indicated that the 285 lavender accessions from 15 fields can be assigned into two major groups, based on their Q values. Some plants fall into a mixed group (Q <0.6), Suplementary table S3. The proportional membership of each individual to a given ancestral cluster is visualized in the STRUCTURE plot, with the two clusters represented in green and red (Figure 8). Cluster Q1 (in red) contains 118 accessions from fields in Kazanlak: IREMK (18), KS (3), GB (19); Chirpan: ZA (18), ZB (6); General Toshevo: KGT (19), MS (19), Shumen: VP (1) and Shabla: Gor (15). Cluster Q2 (in green) includes 135 accessions from Kazanlak: IREMK (1), KS (6), Chirpan: Rup (15), ZB (10), General Toshevo: MY (15), Shumen: VSP (19), VP (16), KSP (17) and Shabla: Sh (19), EZ (18). Notably, mixed ancestral origins exhibited 35 accessions from fields in IREMK (1), KS (13), ZA (1), Rup (4), ZB (3), MY (4), Gor (4), EZ (1), VP (2), KSP (2). This population structure results aligns with the UPGMA dendrogram and explains the significant overlap observed in the PCoA plots. The widespread distribution of both genetic groups across all geographic regions indicates a lack of strict isolation and suggests that the current industrial plantations are characterized by high intra-field heterogeneity, likely due to the historical use of diverse or admixed planting materials.

 doi:10.20944/preprints202604.2001.v1

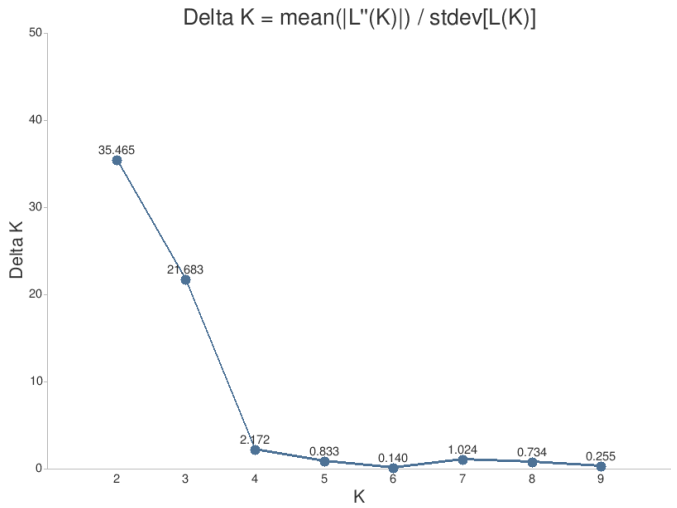


Figure 7. Delta K plot by Evanno based on SCoT-PCR profiles of 285 lavender accessions.

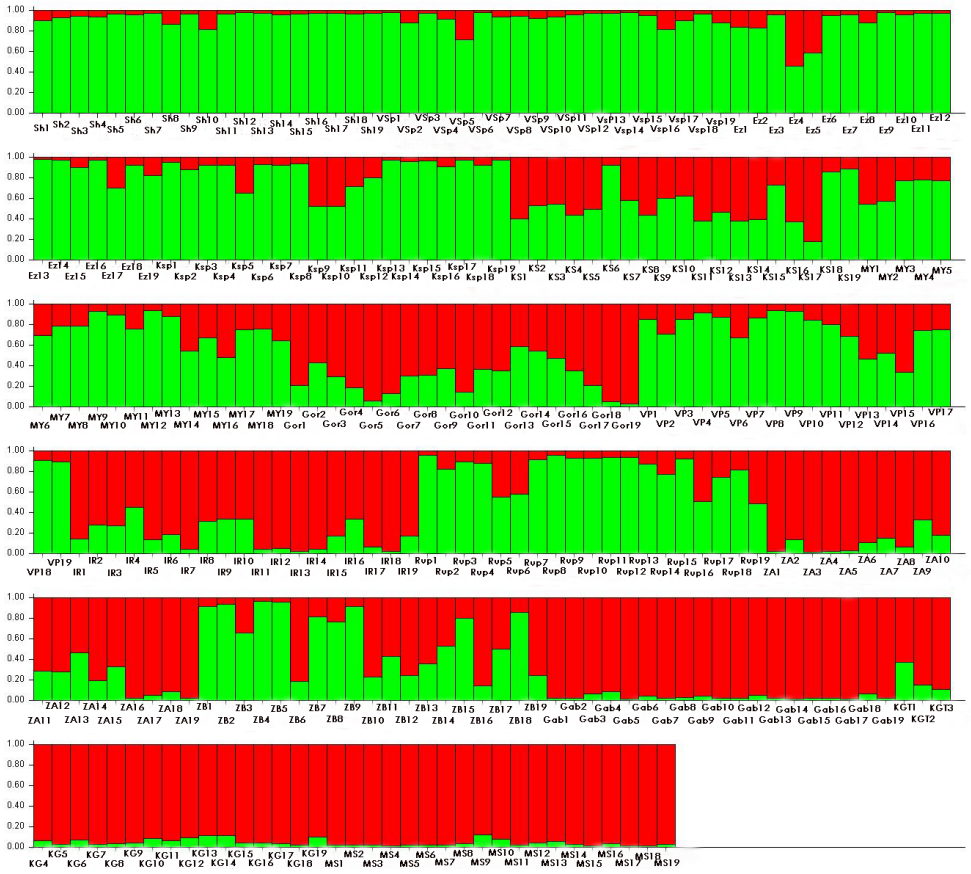


Figure 8. STRUCTURE bar plot representing the estimated genetic membership of 285 lavender accessions into two inferred clusters (K=2). Each vertical bar represents a single individual. The red color indicate membership in Cluster 1(Q1), while green color indicate membership in Cluster 2 (Q2). The length of each colored segment corresponds to the estimated proportion of the individual's genome assigned to that specific cluster. The plants of the fields are presented as follow: Kazanlak: IR1-19, KS1-19 and Gab1-19; Chirpan: ZA1-19, Rup1-19 and ZB1-19; General Toshevo: KG1-19, MS1-19 and MY1-19; Shumen: VS1-19, VSp1-19 and Ksp1-19; Shabla: Sh1-19, Gor1-19 and EZ1-19.

 doi:10.20944/preprints202604.2001.v1

3.3. Marker–Trait Association Analysis

To identify robust genetic markers associated with the chemical composition of lavender essential oil, General Linear Model (GLM) and Mixed Linear Model MLM analyses were performed. Association analysis identified seven significant marker-trait associations ($p < 0.01$) in MLM results and ten in the GLM results, involving ten distinct traits and explaining 19.57-34.19% of the observed chemotype variation (Table 7). Six of the investigated loci demonstrated high stability across both statistical analyses, indicating their superior potential for future marker-assisted selection (MAS): SCoT 12_550 bp, SCoT 24_1200 bp, SCoT 15_450 bp, SCoT 3_700 bp and SCoT 24_700 bp. Notably, the MLM model, which accounts for population structure (Q) and kinship (K), identified high-impact associaton for Camphene with SCoT 12_550 bp (33.97%), for 3-octanone with SCoT 24_1200 bp (25.13%), for Limonene+1.8 cineole with SCoT 24_1200 bp (21.37%), for Camphor with SCoT 12_400 bp (31.40%), for Borneol with SCoT 15_450 bp (31.87%), for Terpinen 4-ol with SCoT 3_700 bp (23.43%), for β-carphyllene wth SCoT 24_700 bp (32.95%).

Table 7. Association analysis results among SCoT markers and chemotypic traits based on MLM and GLM ($p < 0.01$**). R^2 percentage of phenotypic variance explained by the marker.

Trait	Marker	bp	MLM		GLM	
			p	R^2	*p*	R^2
Camphene (CampE)	SCoT 12	550	0.0063**	33.97%	0.0003**	36.12%
3-octanone (3-oct)	SCoT 12	450	0.0379*	14.08%	0.009**	17.86%
3-octanone (3-oct)	SCoT 24	1200	0.0065**	25.13%	0.0017**	25.13%
Cis-ocimene (Cis_oc)	SCoT 2	250	0.0291*	21.84%	0.0095**	21.84%
Limonene+1.8 cineole (Lim_c)	SCoT 24	1200	0.0057**	21.37%	0.0006**	22.77%
Linalool (Lin)	SCoT 15	350	0.0202*	19.47%	0.0033**	22.78%
Camphor (CampO)	SCoT 12	400	0.0072**	31.40%	0.0155*	19.57%
Borneol (Born)	SCoT 12	450	0.0309*	24.65%	0.0037**	31.45%
Borneol (Born)	SCoT 15	450	0.0094**	31.87%	0.0007**	89.11%
Borneol (Born)	SCoT 15	220	0.0471*	17.43%	0.0054**	26.88%
Lavandulool (Lavand)	SCoT 15	1000	0.0196*	24.78%	0.0079**	24.78%
Terpinen 4-ol (Terp)	SCoT 3	700	0.0096**	23.43%	0.0023**	23.43%
β-carphyllene (B_ca)	SCoT 24	700	0.0055**	32.95%	0.0071**	34.19%
β-carphyllene (B_ca)	SCoT 15	450	0.0209*	21.54%	0.0085**	22.17%

*GLM, General Linear Model; MLM, Mixed Linear Model; $p < 0.01$** indicates high statistical significance.

 doi:10.20944/preprints202604.2001.v1

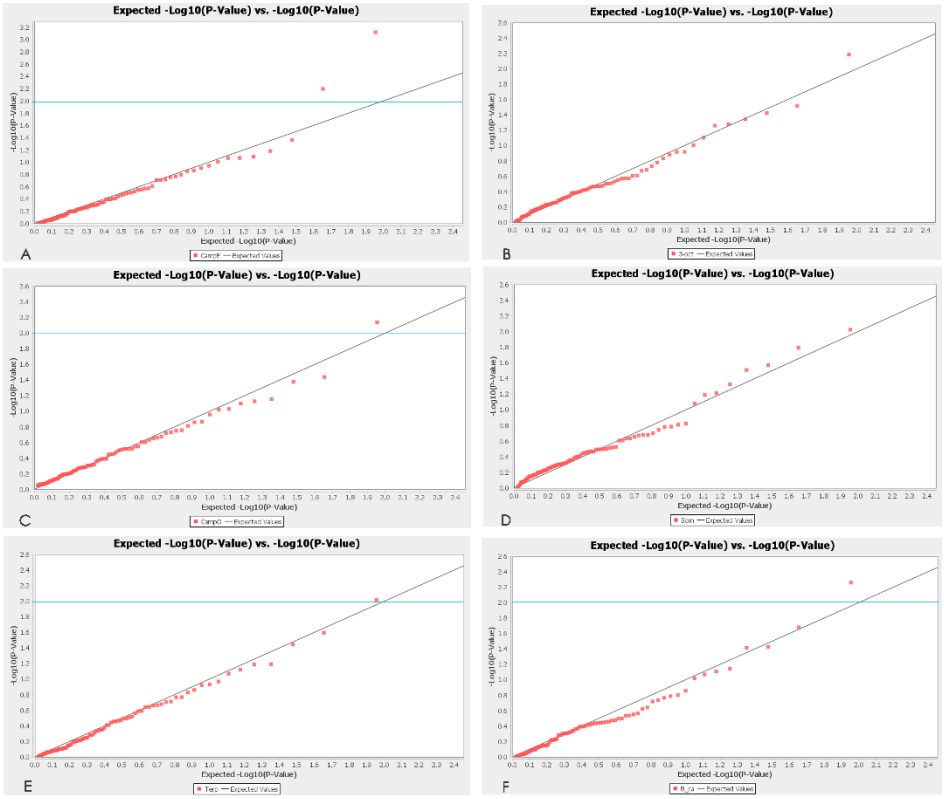


Figure 9. Q-Q plots of marker-trait associations for six chemotypic traits based on the Mixed Linear Model (MLM). The red dots represent individual SCoT loci. The horizontal blue lines indicate the threshold for high statistical significance ($p < 0.01$). Loci positioned above these lines are considered significantly associated with the respective traits. A: Camphene, B: 3-octanone, C: Camphor, D: Borneol, E: Terpinen 4-ol, F: β-carphyllene. .

3.4. Relationships Between Genetic Diversity Indices and Agronomic Traits

This section explores the correlations between genetic polymorphism—both at the level of individual SCoT markers and as an aggregate reflection of population consistency—and the agronomic characteristics of the studied lavender fields. This approach enables an assessment of how local genetic variation and the integrated effect of SCoT loci influence the chemotype and overall field productivity. To maintain focus on the most informative results, only relationships with statistical significance ($p < 0.05$) are presented. However, markers that did not strictly meet this threshold but showed a consistent trend (particularly with $0.05 < p < 0.10$) were also considered, as they may indicate biological dependencies within the complex environment of field trials [60,61].

The analysis determined both the impact of genetic distance between plants within an individual field—expressed as the Genetic Homogeneity Index (HI), calculated according to Formula (7)—and the influence of significant genetic diversity indices, such as the Polymorphic Information Content (PIC) (Table 8).

Table 8. Summary of genetic characteristics across the studied lavender fields, Polymorphic Information Content (PIC) and Genetic Homogeneity Index (HI).

Index	Field														
	Kazanlak			Chirpan		General Toshevo			Shabla			Shumen			
	IREMK	KS	GB	ZA	Rup	ZB	MS	MY	KGT	Gor	EZ	Sh	VP	KSP	VSP
PIC (mean)	0.29	0.36	0.30	0.36	0.39	0.37	0.31	0.35	0.36	0.37	0.40	0.39	0.32	0.39	0.39
HI	0.74	0.76	0.82	0.75	0.7	0.74	0.76	0.78	0.76	0.71	0.65	0.65	0.74	0.66	0.65

 doi:10.20944/preprints202604.2001.v1

3.4.1. Genetic Homogeneity Index (Hi)

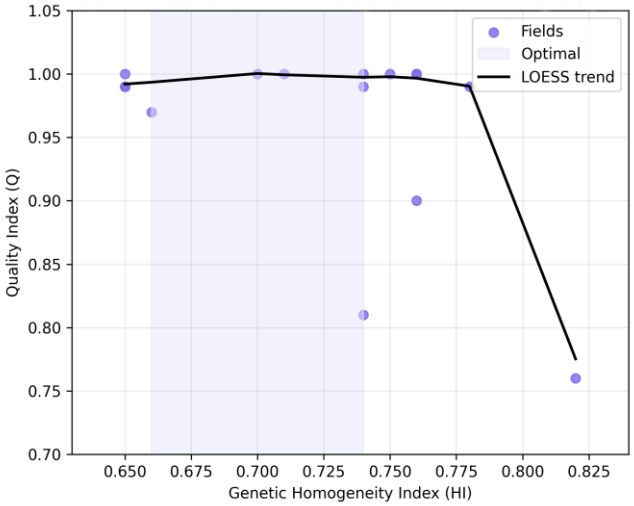
Genetic analysis based on Start Codon Targeted (SCoT) markers revealed significant correlations between intra-fields genetic structure and the production performance of the studied lavender plantations. The results indicate that the minimum genetic distance between plants (min_Nei) is positively correlated with the integrated Field Quality Index (FQI) (r = 0.68, p = 0.01) and normalized yield (Y) (r = 0.55, p = 0.04), Table 9. These findings suggest that optimal productivity is achieved under conditions of moderate, rather than extreme, genetic proximity between individual plants within a field.

The Genetic Homogeneity Index (HI), which reflects the degree of intra-field uniformity, indicates that higher HI values correspond to lower genetic diversity or reduced differentiation within the accessions in the fields. The analysis identified a negative correlation between HI and essential oil quality (Q) (r = -0.45, p = 0.09), Table 8. This suggests that while genetic uniformity is beneficial for yield stability, excessive homogeneity may slightly diminish the phytochemical complexity and overall quality of the lavender essential oil.

Table 9. Correlation analysis between genetic distance metrics and Genetic Homogeneity Index (HI), Field Quality Index (FQI), Essential oil quality (Q), Normalized yield (Y).

Nei_metric	Trait	*rho*	*p*
min_Nei	FQI	0.68	0.01
min_Nei	Y	0.55	0.04
Nei_cv	FQI	-0.46	0.08
Nei_mid	FQI	0.46	0.08
HI	Q	-0.45	0.09

The relationship between Genetic Homogeneity (HI) and essential oil quality (Q) exhibits a distinct non-linear trend (Figure 10). Peak quality values are observed at intermediate HI levels ranging from 0.65 to 0.74, suggesting that moderate genetic diversity provides an optimal balance between field stability and aromatic complexity. However, when the homogeneity exceeds the threshold of HI ≈ 0.72, quality parameters begin to fluctuate. A sharp decline in phytochemical indicators is observed once the index surpasses 0.78, indicating that extreme genetic uniformity may lead to a loss of essential oil richness.



 doi:10.20944/preprints202604.2001.v1

Figure 10. Non-linear relationship between the Genetic Homogeneity Index (HI) and essential oil quality (Q). The LOESS trend line illustrates the optimal range of genetic uniformity for maintaining high phytochemical quality.

Among all genetic indices, Polymorphic Information Content (PIC) stands out as the only statistically robust predictor of essential oil quality after applying the Holm adjustment ($p\_adj = 0.02$). Approximately 60% of the variation in the quality indicator (Q) can be explained by the mean PIC per field. This confirms that PIC serves as a highly reliable molecular indicator for assessing both the genetic fitness and phytochemical potential of lavender plantations (Table 10, Figure 11)

Table 10. Correlation between Polymorphic Information Content (PIC) and Essential Oil Quality (Q) after Holm correction.

Genetic	Pheno	Method	*r/rho/τ/D*	*p-raw*	*p-adj*	Signif
PIC	Quality	Pearson	0.77867	0.00063	0.02253	True

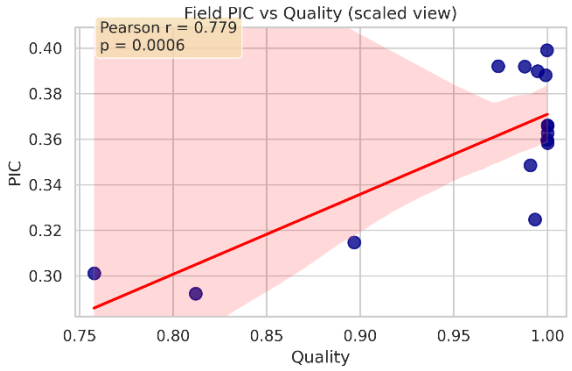


Figure 11. Regression analysis illustrating the predictive power of mean field PIC for essential oil quality. The high correlation coefficient ($r = 0.7787$) identifies PIC as a reliable molecular indicator for lavender field fitness.

Analysis of specific binary parameters identified SCoT-2 and SCoT-24 as the most significant markers associated with essential oil quality and yield, Figure 12. Both markers exhibited strong positive statistical correlations ($r > 0.71$ for ScoT 2) и ($r > 0.71$ за ScoT 24). Notably, both linear and monotonic models demonstrate a consistent trend, positioning SCoT-24 as a balanced molecular marker with high potential for marker-assisted selection (MAS) to identify genotypes with superior productivity and a stable aroma profile.

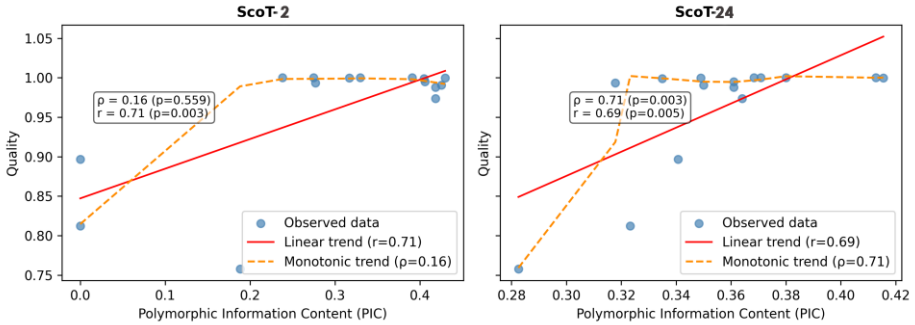


Figure 12. Linear and monotonic Spearman and Pearson correlations between Polymorphic Information Content (PIC) and essential oil quality (Q). The high correlation coefficients for SCoT-2 and SCoT-24 identify these loci as robust molecular indicators for predicting lavender phytochemical profiles.

 doi:10.20944/preprints202604.2001.v1

Results for the Expected heterozygosity (He) in the studied plantation revealed that marker SCoT-3 demonstrates a stable negative correlation with production indicators ($r \approx -0.61$; $p < 0.05$), At this stage, this marker serves as a molecular indicator of risk regarding overall productivity (Figure 13). Conversely, SCoT-2 showed a positive correlation with quality ($r \approx 0.72$; $p < 0.05$) identifying it as a key indicator for improving essential oil phytochemical characteristics.

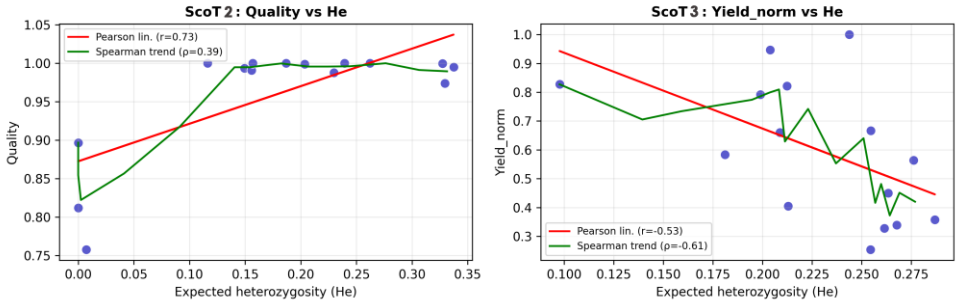


Figure 13. Impact of expected heterozygosity (He) on lavender essential oil quality (Q) and normalized yield (Y). The contrasting trends observed for SCoT-2 and SCoT-3 illustrate the trade-off between genetic diversity and agronomic performance.

The integrated correlation analysis, visualized through a heatmap (Figure 14), confirms the distinct genetic architecture underlying oil quality and yield. Genetic diversity indices for SCoT-2 and SCoT-24, specifically Allelic Diversity (AD), Polymorphic Information Content (PIC), and Percentage of Polymorphic Bands (PB%), show strong positive correlations ($r = 0.63$ to 0.74) with both quality and yield. Notably, the Relative Multi-Locus Genotype index (relMLG) which reflects the proportion of unique genetic profiles within a plants exhibits a significant negative correlation with normalized yield (r = -0.53 to -0.58) for markers SCoT 2 and SCoT 24. This suggests that while high allelic diversity (PIC and AD) is essential for essential oil quality, an excessive increase in unique genotypes (high relMLG) may negatively impact field productivity.

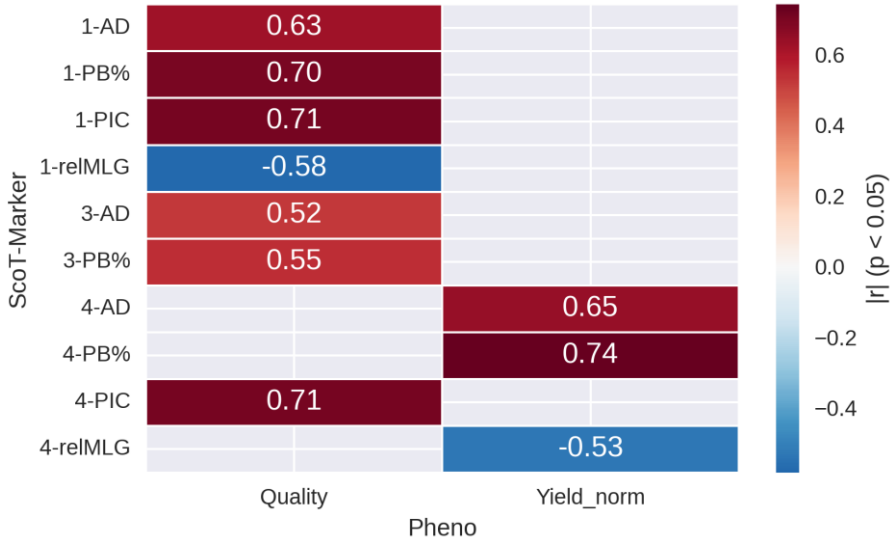


Figure 14. Heatmap of significant Pearson and Spearman correlations ($p < 0.05$) between SCoT-marker genetic diversity indices Allelic Diversity (AD), Polymorphic Information Content (PIC), Percentage of Polymorphic Bands (PB%), Multi-Locus Genotype index (relMLG) and phenotypic traits – Quality (Q) and Normalized Yield

 doi:10.20944/preprints202604.2001.v1

(Y) for 1 – SCoT 2; 3 – SCoT 12; 4 – SCoT 24. Red color indicates positive correlations, while blue indicates negative correlations, highlighting the diverging genetic drivers for oil quality and field productivity.

3.5. Agro-Climatic Characterization and Environmental Drivers of the Lavender Cultivation Sites

To ensure a precise and in-depth interpretation of the results, meteorological data for the lavender vegetation period were analyzed across five distinct regions in Bulgaria by ERA5-Land (ERA5L) reanalysis dataset (Table 11). The 2023 data indicate that the highest mean temperatures (t^2m) were recorded in the Chirpan region, reaching their peak in July (27.49 °C). Conversely, the coastal region of Shabla exhibited the highest solar radiation levels (*ssrd*) during June (743.56 MJ/m²) and significantly stronger wind speeds throughout the season compared to the other sites (4.54 - 4.60 m/s). Regarding soil moisture (*swvl1* at 0-7 cm depth), the lowest values were consistently observed in Shabla (0.22 m³/m³), General Toshevo (0.28 m³/m³) and Shumen (0.29 m³/m³), whereas the Kazanlak and Chirpan fields maintained higher soil water content during the early growth stages (0.37 - 0.41 m³/m³). These contrasting agro-climatic profiles between the experimental sites provide the necessary environmental context for understanding the variations in essential oil yield and quality indices.

Table 11. Summary of the monthly meteorological conditions across the five experimental sites: Kazanlak, Chirpan, General Toshevo, Shabla, and Shumen for the period May–July 2023. Mean 2-meter dewpoint temperature (d^2m, °C), Mean 2-meter air temperature (t^2m, °C), Total precipitation (*tp*, mm), Surface solar radiation downwards (*ssrd*, MJ/m²), Volumetric soil water layer 1 (*vswl1*, 0–7 cm depth, m³/m³), Mean wind speed at 10 meters height (m/s).

	Kazanlak: IREMK, KS, Gab			**Chirpan: ZA, Rup, ZB**			**General Toshevo: KGT, MS, MY**			**Shabla: Sh, Gor, EZ**			**Shumen: VS, VSP, KSP**		
	May	Juny	July	May	Juni	July	May	Juni	July	May	Juni	July	May	Juni	July
d^2m	9.87	14.31	16.75	10.37	13.58	15.15	9.38	13.60	16.40	11.33	16.06	18.88	9.38	13.46	15.75
t^2m	13.93	19.36	24.29	15.88	21.92	27.49	14.70	19.87	24.42	15.18	20.90	24.34	14.98	20.17	25.47
tp	95.16	102.30	33.77	65.52	39.48	20.30	87.83	56.49	48.94	29.41	19.83	28.68	87.81	46.08	44.29
ssrd	586.06	661.56	722.51	585.44	674.48	734.26	636.92	700.97	694.32	694.20	743.56	714.55	604.54	680.66	706.11
swvl1	0.37	0.37	0.26	0.41	0.34	0.29	0.28	0.25	0.21	0.22	0.16	0.16	0.29	0.25	0.20
wind	1.63	1.31	1.46	2.20	1.91	1.99	3.25	2.82	3.27	4.54	3.88	4.60	2.47	2.12	2.37

The correlation analysis (Figure 15, Suplementary Table S4) revealed a complex interplay between environmental conditions and the phenotypic expression of lavender genotypes. A dominant climatic signal was observed during June and July. Notably, solar radiation (ssrd-07) and soil moisture (swvl1-07) in July exhibited strong negative correlations with essential oil yield indices ($r < -0.70$, $p < 0.001$). This suggests that excessive solar exposure and high soil moisture levels—as observed in the Chirpan and Kazanlak fields—act as primary limiting factors for oil yields. Weak to moderate correlations were also identified between yield and July precipitation ($r = 0.63$, $p = 0.01$), as well as wind speed throughout the summer months ($r < 0.53$, $p < 0.04$). Interestingly, lavender oil quality (Q) remained remarkably stable, showing only negative correlations with June precipitation ($r = -0.69$, $p = 0.004$). In contrast, the Field Quality Index (FQI) was moderately influenced by June solar radiation ($r = 0.57$, $p = 0.03$) and wind speed ($r = 0.59$, $p = 0.02$), while soil moisture during both months emerged as a decisive factor for the overall field performance ($r \leq -0.74$, $p < 0.001$). Furthermore, the percentage of polymorphic bands (PB%) demonstrated a moderate negative correlation with precipitation and soil moisture ($r \leq -0.54$, $p < 0.02$), suggesting that genetic polymorphism in the studied genotypes might be sensitive to high hydrothermal stress.

 doi:10.20944/preprints202604.2001.v1

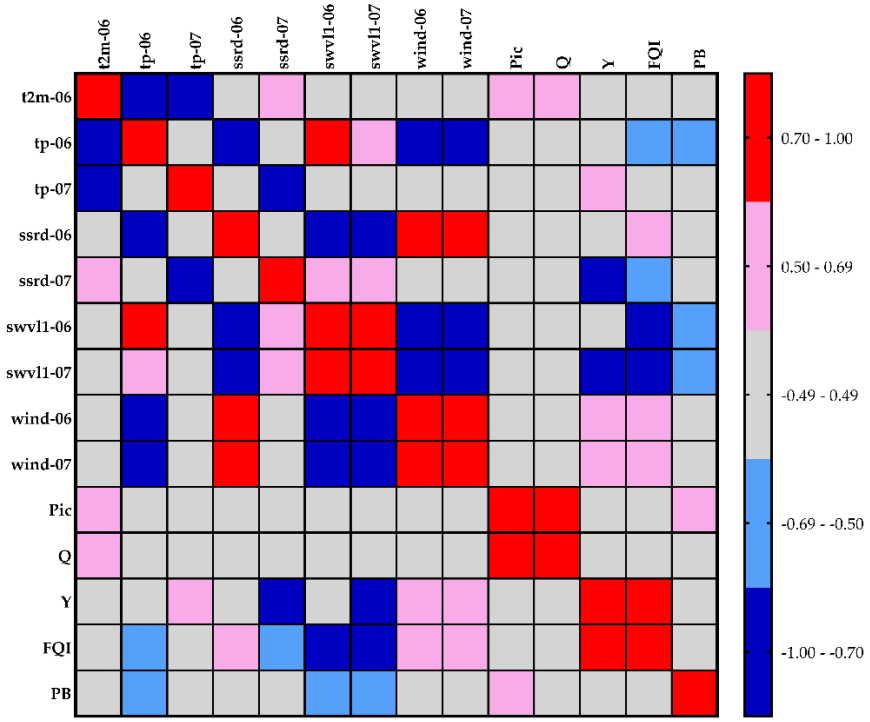


Figure 15. Heatmap of the Pearson correlation coefficients (*r*) between genetic diversity indices: Polymorphic Information Content (PIC), Percentage of Polymorphic Bands (PB%), Climatic variables (temperature (t^2m), precipitation (*tp*), solar radiation (*ssrd*), soil moisture – 0-7 sm (*swvl1*), wind for June and July), and Phenotypic traits: Quality (Q), Yield (Y), Final Quality Indices (FQI) of lavender genotypes in 15 fields. The color scale represents the strength of the correlation, ranging from dark blue (strong negative, $r \leq -0.70$) to dark red (strong positive, $r \geq 0.70$). The light-colored cells indicate non-significant or weak to moderate relationships ($|r| < 0.50$).

Regarding the chemical composition of the essential oil across the different regions, mean daily temperatures throughout the entire vegetation period were found to be strongly correlated with increased levels of trans-ocimene ($r < 0.71$, $p < 0.001$), and moderately correlated with camphor ($r < 0.54 – 0.62$, $p < 0.03$) and terpinen-4-ol ($r < 0.58 – 0.68$, $p < 0.02$), Figure 16, Supplementary Table S5. This confirms that heat stress stimulates the biosynthesis of compounds often associated with lower perfumery quality. A moderate negative correlation was observed between high summer temperatures for June and July and linalool content ($r < -0.54$, $p < 0.04$), highlighting that oil quality may be compromised in hotter regions. Furthermore, high temperatures in May appear to suppress the accumulation of limonene + 1,8-cineole ($r = -0.75$, $p = 0.001$), likely shifting metabolic pathways toward competitive biochemical routes (specifically trans-ocimene ($r = 0.7$, $p = 0.002$)) involved in early-season heat adaptation. This trend persists through May and June, leading to a moderate suppression of 3-octanone levels as well ($r < -0.64$, $p < 0.01$).

 doi:10.20944/preprints202604.2001.v1

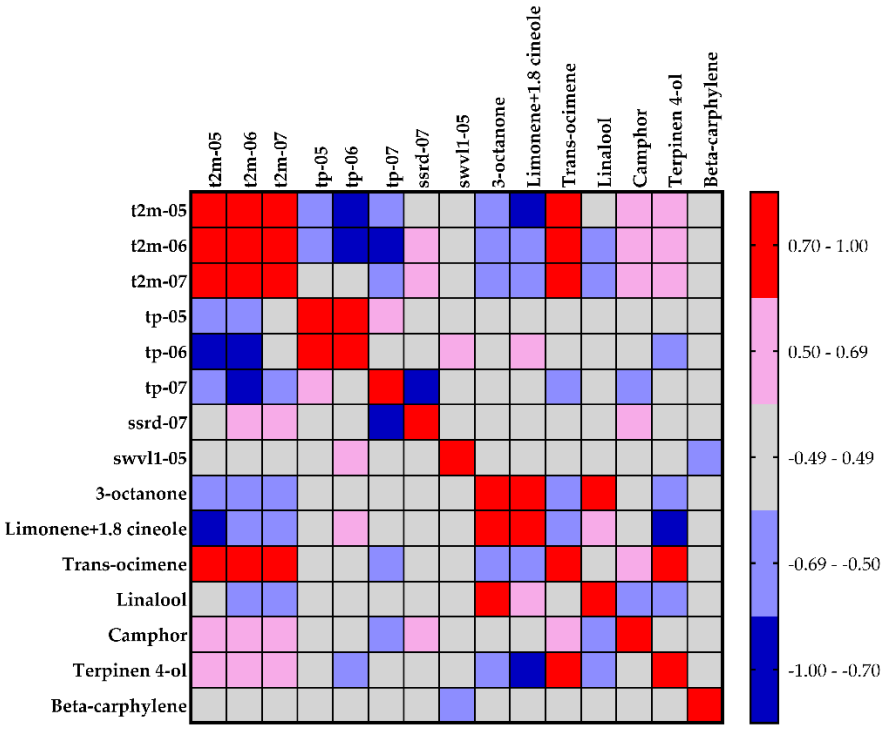


Figure 16. Heatmap of the Pearson correlation coefficients (r) between Climatic variables (temperature (t^2m), precipitation (*tp*), solar radiation (*ssrd*), soil moisture – 0-7 sm (*swvl1*) for May, June and July), and Phenotypic traits of 30 lavender genotypes in 15 fields from five region-Kazanlak, Chirpan, General Toshevo, Shabla and Shumen. The color scale represents the strength of the correlation, ranging from dark blue (strong negative, $r \leq -0.70$) to dark red (strong positive, $r \geq 0.70$). The light-colored cells indicate non-significant or weak to moderate relationships ($|r| < 0.50$).

4. Discussion

The yield and quality of lavender essential oil (*Lavandula angustifolia* Mill.) are profoundly influenced by a multitude of factors [62]. Variations in the chemical profile are frequently observed, even within the same genotype grown under identical conditions and subjected to the same harvesting and extraction methods [63–65]. According to numerous studies, these factors include the harvest timing—specifically the flowering stage and time of day [66,67]; plant age [68–70]; the specific genotype utilized [64,65]; cultivation technology, such as fertilization and irrigation [71–73]; the role of pollinators [74,75]; microclimatic conditions, including temperature and precipitation distribution [64,66,70]; soil structure [76]; and altitude [77,78].

We conducted the first comparative genetic and phenotypic study of commercial lavender plantations across the five primary cultivation regions in Bulgaria for the 2023 season. For the first time, an integrated methodology for evaluating lavender plantations has been developed and applied, unifying chemical, phenotypic, and genetic parameters. The newly introduced Field Quality Index (FQI) allows for a quantitative expression and comparability of production efficiency by simultaneously reflecting essential oil quality, yield, and intra-field homogeneity. Overall, our findings confirm that high essential oil quality alone cannot fully offset low productivity. The balanced contribution of all three components (Q, Y, and H) ultimately determines the true economic value of a field. The FQI structure provides an objective and robust framework where high productivity is a necessary but insufficient condition for success; true field excellence is attained only through the synergy of superior quality and intra-field stability.

 doi:10.20944/preprints202604.2001.v1

4.1. Characteristics of the Bulgarian Lavender Regions

Our study demonstrates that the regions in Northeastern Bulgaria—Shumen, General Toshevo, and Shabla—which are considered relatively new and non-traditional for commercial lavender cultivation [65], achieve significantly higher yields than traditional regions like Kazanlak and Chirpan, without compromising the quality of the essential oil (Table 3, Figure 2). Specifically, fields from these regions, such as Kaspichan (FQI = 87.45) and Shabla (FQI = 86.81), received the highest Field Quality Index (FQI) scores, reflecting their superior performance in terms of yield, homogeneity, and oil quality. These findings align with previously reported high quality of the Bulgarian cultivars 'Hemus', 'Sevtopolis', 'Yubileyna', and 'Druzhba' in field trials, which identified 'Sevtopolis' as the highest-yielding for the region, while 'Hemus' exhibited the best chemical composition [65]. The authors concluded that these cultivars successfully realize their productive potential in this new region, confirming that essential oil yield depends not only on meteorological conditions during harvest but also significantly on the genotype [65]. In a similar study conducted in Eastern Bulgaria (Aytos), it was emphasized that the qualitative and quantitative parameters of the essential oil for 'Hemus', 'Sevtopolis', and 'Yubileyna' are determined by both precipitation during the vegetation period and the specific genotype, with 'Yubileyna' emerging as the most productive for that specific area [79].

Due to the field-based nature of this study, the strict varietal identity of the 285 individual plants sampled across 15 locations cannot be definitively guaranteed. Consequently, the specimens were categorized based on the results of the clustering analyses. Data from PCoA, UPGMA, and STRUCTURE population analysis consistently identified two primary genetic clusters. In the regions of Northeastern Bulgaria, plants cultivated in Shabla (Sh, EZ) and Shumen (VP, VSP, KSP) predominantly belong to the second genetic group (Q2), making this genotype prevalent for the area. Notably, however, a field from Rupkite (Chirpan, Southern Bulgaria) also clustered within this group yet recorded the lowest overall yield. This clearly emphasizes that yield is more significantly influenced by environmental conditions than by genotype alone, suggesting that the strategic synchronization of both factors is essential for optimal production. Existing research highlights that abundant rainfall during May and June typically leads to a decline in essential oil yield [65,79]. This is consistent with the higher humidity levels recorded in our study for the Kazanlak and Chirpan regions, which showed a strong correlation with lower yields (Figure 15, Table 9). Furthermore, higher mean monthly temperatures during the harvest period—particularly in Chirpan—demonstrated a significant negative correlation with yield across the studied regions.

In contrast to yield, essential oil quality remained stable across all sites. Quality showed only a moderate positive correlation with mean June temperatures, but a strong correlation with the measured field polymorphism (PIC). This suggests that quality is significantly more dependent on the genotypes present in the fields than on environmental fluctuations. Recent analysis of gene expression patterns across 10 genes noted that expression in flower buds is primarily stimulated by intense light and low temperatures, whereas gene expression in leaves is triggered by drought [38]. The authors identified two specific genes (*LaWRKY57* and *LaWRKY75*) as potential negative regulators in terpenoid biosynthesis [38], providing a molecular basis for the observed metabolic shifts under environmental stress. In the present study, the negative correlation confirms that high temperature acts as a limiting factor. This suggests that the more moderate temperature regimes in Northeastern Bulgaria favor the gene expression responsible for linalool biosynthesis, which explains the preservation of high essential oil quality in these regions, Figure 16.

The combined analysis of phenotypic and genetic parameters in this study indicates that an excessively narrow genetic base limits both the yield and the phytochemical complexity of the oil. Fields with moderate genetic diversity (HI = 0.66–0.74) demonstrate the highest values for quality and productivity. Conversely, when fields become overly genetically homogeneous (HI > 0.80), the quality index drops drastically (falling below 0.80 at HI ≈ 0.82), suggesting that extreme clonal purity may act as a limiting factor for lavender oil quality. The observed decline in FQI at high homogeneity levels (HI > 0.80) may be attributed to the loss of heterosis or genetic segregation, often resulting from

 doi:10.20944/preprints202604.2001.v1

uncontrolled farm-level propagation. Such practices likely lead to the erosion of essential agronomic traits, compromising the plant's ability to maintain a complex phytochemical profile under environmental stress. The presence of outliers outside the main trend suggests that while genetics is a primary driver, other local factors—likely agro-climatic or edaphic—can also diminish quality (Figure 10).

Furthermore, while certain levels of genetic diversity positively influence quality, they may contribute to a slight decline in total yield (Figure 13). Based on these findings, it is recommended that new plantations maintain a minimum Nei's distance of ≥ 0.25 between plants and utilize vegetatively propagated planting material from at least three different genotypes. To avoid the risks of genetic segregation and loss of heterosis associated with informal farm-level propagation, it is crucial that the recommended genetic diversity (at least three genotypes) be established using certified, vegetatively propagated starting material. This ensures that the field maintains both the necessary robustness and the high phytochemical standards required by the industry. This strategy aims to maintain optimal genetic diversity and ensure the stability of production indicators. Finally, the SCoT-1 and SCoT-4 markers can serve as genetic indicators for selecting high-quality, stable plantations, while SCoT-2 could be used to identify areas with lower productivity, which should be harvested separately (Figures 11, 12, 13). The PIC (Polymorphic Information Content) values derived from SCoT markers serve as effective tools for the early diagnosis of plantation quality. While both SCoT-1 and SCoT-4 markers proved valuable, SCoT-4 exhibited a stronger monotonic relationship (ρ = 0.71), making it more precise for detecting non-linear shifts in quality. Correlations for SCoT-1 and SCoT-4 (Figures 12, 14) and the generalized regression (Figure 11) demonstrate a robust positive correlation between PIC and quality index Q ($r \approx 0.71 - 0.78$). Furthermore, the non-linear analysis (Figure 10) confirms that field homogeneity significantly impacts quality only after reaching a specific critical threshold. Moreover, while overall homogeneity tends to increase yield (Figure 12), the percentage of polymorphic bands (PB%)—another metric of genetic diversity—also correlates positively with yield (r = 0.74, Figure 14). This suggests that genetic diversity is not inherently detrimental to productivity; rather, the presence of specific polymorphic loci (particularly those identified by SCoT-4) may indicate the presence of high-yielding alleles. These data confirm that while heterozygosity in the planting material must be maintained, an excessive prevalence of certain multi-locus genotypes (MLGs) can negatively affect both yield (for 4-relMLG, r = -0.53) and quality (for 1-relMLG, r = -0.58), as shown in Figure 14.

4.2. Genetic Diversity Characteristics of Lavender Genotypes in Bulgaria

The population structure analysis performed via STRUCTURE software (Figure 8) is fundamental to understanding the origin of the studied plantations. The Delta K method (Figure 7) showed a distinct peak at K = 2, indicating that the entire sampled population originates from two primary genetic sources. A significant level of uncontrolled mixing of planting material was observed, with almost no field being genetically uniform. This is further corroborated by the Analysis of Molecular Variance (AMOVA), which revealed high intra-field heterogeneity (64%). Previous studies of the Bulgarian lavender gene pool using 13 SCoT markers reported genetic diversity levels (He = 0.27, I = 0.39) comparable to the results obtained here for the 285 individual plants (He = 0.35, I = 0.50). Most modern cultivars are derived from the 'Hemus' variety, which has consistently proven to be of the highest quality [25]. This shared lineage likely explains the significant overlap and admixture of genetic components observed in our study; most commercial fields in Bulgaria utilize cultivars that are closely genetically related.

In the PCoA analysis at the individual plant level (Figure 5b), this trend of overlap persists; however, at the field level, a clear separation along the horizontal axis (PCo1 = 23.96%) is observed, suggesting a potential for regional genetic profiling of lavender in Bulgaria. The UPGMA dendrogram further illustrates this heterogeneity by grouping plants from the same field into distinct genetic clades across different regions.

 doi:10.20944/preprints202604.2001.v1

This explains the low Homogeneity Index (HI) observed in certain fields, which are likely composed of genetically distant individuals—a factor that diminishes both yield and quality, thereby reducing their overall Field Quality Index (FQI). Based on the clustering analyses and genetic diversity data for the General Toshevo sites, it can be concluded that the planting material used there possesses a more homogeneous but distinct origin compared to other northern locations. The results indicate that the genetic foundation of lavender in General Toshevo is more closely aligned with certain southern sites or represents distinct, purer lines that differ from the predominant genetic cluster (Q2) characteristic of Northeastern Bulgaria. This distinct genetic profile likely accounts for the intermediate position that General Toshevo occupies between the high-productivity and low-productivity fields in the studied regions.

4.3. Effectively of Scot Markers Vl Mas

SCoT markers demonstrate a high capacity for revealing genetic diversity and differentiating closely related individuals within the studied samples. This is evidenced by the robust average values calculated for Marker Index (MI = 90.72), Resolving Power (Rp = 18.39), Polymorphic Information Content (PIC = 0.64), and Effective Multiplex Ratio (EMR = 147.28). The analysis of 285 individuals established significant genetic diversity (He = 0.35, I = 0.50), and Φ_{PT} (*PhiPT*) value of 0.356 indicating substantial potential for the improvement and strategic utilization of existing germplasm. High intraspecific diversity is fundamental to plant adaptation and evolution. Understanding the relative contributions of genetic and environmental factors to chemical variation is essential for identifying adaptive traits, thereby enhancing cultivar selection and the development of locally adapted varieties [80]. In the present study, an association analysis between 90 loci and 30 individuals identified 6 loci with high potential ($p < 0.01$) for Marker-Assisted Selection (MAS). These loci explain between 21.37% and 33.97% of the phenotypic variation for camphene, 3-octanone, limonene + 1.8-cineole, borneol, terpinen-4-ol, and β-caryophyllene, Table 7, Figure 9. In comparison, other studies utilizing SCoT markers and the General Linear Model (GLM) accounted for only 8.24% to 12.30% of the variation in quantitative gene effects controlling rust in orchardgrass; notably, these effects remained consistent across two sampling years, further validating the reliability of this marker system [81]. Similarly, researchers using the Mixed Linear Model (MLM) identified eight SCoT markers in Chrysanthemum morifolium Ramat associated with six key traits (e.g., flowering duration, plant architecture, and pest resistance), explaining 11.18–20.92% of the phenotypic variation [82]. The superior explanatory power observed in our study highlights the precision of SCoT markers for lavender metabolic profiling.

Studies examining the impact of climate change on the chemical composition of essential oils from *Salvia officinalis* and *Inula viscosa* have shown a sharp increase in secondary metabolites concurrent with rising temperatures and progressive water stress during an initial two-year period; however, this trend plateaued or declined in subsequent years under prolonged stress [83]. Our data comparing different regions and oil compositions align with these observations, suggesting that rising temperatures as early as May trigger plants to accumulate higher levels of trans-ocimene while suppressing the limonene + 1,8-cineole pathway (Figure 16).

The inherent complexity and variability of essential oils, the risk of adulteration, and the challenges in producing consistent high-quality products lead to undesirable heterogeneity in the lavender oils available on the market [16]. Comparative analyses of various lavender species (*L. angustifolia*, *L. x intermedia*, *L. spica*, *L. latifolia*, *L. stoechas* from Egypt, Australia, and France) have revealed significant variations in both oil quality and yield [84]. Even within *L. angustifolia* alone, reported ranges for linalool and linalyl acetate vary widely, from 10.0–57.5% and 4.0–55.0%, respectively [85]. In our study, the levels of linalool and linalyl acetate across 45 samples from 15 commercial fields in Bulgaria's primary production regions ranged from 19.92–40.32% and 20.29–39.39%, respectively. These results confirm that while Bulgarian lavender oil maintains a competitive profile, regional environmental pressures and genetic heterogeneity significantly influence its chemical consistency.

 doi:10.20944/preprints202604.2001.v1

5. Conclusions

The findings of this study demonstrate that the highest integrated FQI values are achieved in lavender fields characterized by moderate genetic diversity (HI = 0.6–0.7) and high polymorphic information content (PIC ≥ 0.35). The established correlations confirm that excessive genetic uniformity limits both the yield and the phytochemical complexity of the essential oil, whereas moderate differentiation ensures production stability. To maintain this balance, it is recommended that new plantations establish a minimum Nei's genetic distance of ≥ 0.25 between individuals and utilize certified planting material from at least three distinct genotypes.

Fields in the Shumen region and Northeastern Bulgaria demonstrated an optimal balance between productivity and quality, highlighting the profound influence of regional agro-climatic conditions. Specifically, more moderate temperature regimes in these areas appear to favor the gene expression responsible for linalool biosynthesis, preventing the quality degradation associated with high-temperature stress.

Among the molecular tools tested, markers ScoT 2 and ScoT 24 emerged as the most reliable indicators for breeding programs targeting enhanced quality and yield, while ScoT 12 can serve as a diagnostic tool for identifying lower-efficiency production areas. The integration of genetic and phenotypic parameters within the FQI framework provides a robust model for genome-informed management, offering a strategic pathway for the sustainable production of lavender oil with high and predictable economic value in the face of shifting climatic conditions.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/doi/s1, Table S1: Chemical composition and essential oil yields across 15 commercial lavender fields in Bulgaria.;. Table S2. Intra-field pairwise Nei's genetic distances among individual plants across the studied lavender populations; Table S3. Population structure and individual ancestry coefficients (Q-values) for 285 lavender plants as determined by SCoT markers.; Table S4. Pearson correlation matrix between genetic diversity indices, climatic variables, and phenotypic traits.; Table S5. Pearson correlation matrix between climatic parameters and the chemical profile of lavender essential oil. .

Author Contributions: Conceptualization, M.Z. and V.B.; methodology, M.Z. and V.B.; software, M.Z.; validation, M.Z., and V.B.; formal analysis, M.Z. and V.B.; investigation, M.Z. and V.B; resources, M.Z. and V.B.; data curation, M.Z. and V.B.; writing—original draft preparation, M.Z. and V.B.; writing—review and editing, M.Z. and V.B.; visualization, M.Z. and V.B.; supervision, M.Z.; project administration, M.Z.; funding acquisition, M.Z. All authors have read and agreed to the published version of the manuscript.

Funding: The APC was funded by the Bulgarian Ministry of Education and Science (MES) in the framework of the Bulgarian National Recovery and Resilience Plan, Component "Innovative Bulgaria," Project No BG-RRP-2.004-0006-C02 "Development of research and innovation at Trakia University in service of health and sustainable well-being". The study was supported by the Bulgarian Ministry of Education and Science under the National Science Fund grant number: КП 06 М61/7 15.12.22, Project: "Genetic Profile of Lavender Plantations from Different Regions of Bulgaria. Determining the Authenticity and Uniformity of the Initial Planting Material and Its Influence on the Quality of the Obtained Essential Oil.".

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed at the corresponding author.

Acknowledgments: The authors would like to express their sincere gratitude to Eng. Dinko Vasilev for the development of the integrated methodology for field evaluation—the Field Quality Index (FQI)—and for his invaluable contribution to the associated data analysis. Gratitude is also expressed to master Vanesa Vasileva and PHD Zornitsa Zherkova for their valuable assistance in the collection and transportation of the plant material. The author further acknowledges the lavender growers for their cooperation and for granting access to their fields for sampling.

Conflicts of Interest: The authors declare no conflicts of interest.

 doi:10.20944/preprints202604.2001.v1

References

1. Shellie, R.; Marriott, P.; Huie, C. Comprehensive two-dimensional gas chromatography (GC×GC) and the characterization of lavender essential oils. J. High Resolut. Chromatogr. **2002**, 25, 403–408. https://doi.org/10.1002/1521-4168(20020701)25:7<403::AID-JHRC403>3.0.CO;2-9
2. Prusinowska, R.; Śmigielski, K.B. Composition, biological properties and therapeutic effects of lavender (*Lavandula angustifolia* L.) – A review. *Herba Pol.* **2014**, *60*, 56–66. https://doi.org/10.2478/hepo-2014-0010
3. Stoyanova, A.; Balinova, A. Composition and quality of Bulgarian lavender oil. *Ind. Crops Prod.* **2019**, *138*, 111427. https://doi.org/10.1016/j.indcrop.2019.111427
4. Minev, M. Lavender cultivation and essential oil yield in Bulgaria: Current status and challenges. *Bulg. J. Agric. Sci.* **2020**, *26*, 445–452.
5. Atanasova, T.; Nedkov, N. Regional specialization and production zones of essential oil crops in Bulgaria. *Proc. Inst. Roses Essent. Oil Crops* **2004**, *21*, 45–50.
6. Stanev, S.; Dzhurmanski, A. Variability in quality parameters of lavender oil from different regions in Bulgaria. *J. Essent. Oil Res.* **2011**, *23*, 45–52. https://doi.org/10.1080/10412905.2011.9700451
7. Kozuharova, E.; Dimitrova, V.; Vassileva, T. Genetic and chemical diversity of Bulgarian lavender populations. *Plants* **2023**, *12*, 955. https://doi.org/10.3390/plants12040955
8. Inteli Agro. Market report: Bulgarian lavender production and export potential. Sofia, Bulgaria, **2018**.
9. Badzhelova, V.; Zhelyazkova, M.; Nedeva, D.; Donchev, I. Quantitative and qualitative characteristics of lavender oil from different regions of Bulgaria. *J. Mt. Agric. Balk.* **2024**, *27*, 527–544.
10. Semerdjieva, I.B.; Zheljazkov, V.; Cantrell, C.L.; Astatkie, T.; Ali, A. Essential oil yield and composition of the Balkan endemic *Satureja pilosa* Velen. (Lamiaceae). *Molecules* **2020**, *25*, 827.
11. Stanev, S.; Angelova, V.; Zagorcheva, T. Yield and quality characteristics of lavender oil under different agroecological conditions. *Bulg. J. Agric. Sci.* **2016**, *22*, 471–478.
12. Kozuharova, E.; Simeonov, V.; Batovska, D.; Stoycheva, C.; Valchev, H.; Benbassat, N. Chemical composition and comparative analysis of lavender essential oil samples from Bulgaria in relation to the pharmacological effects. *Pharmacia* **2023**, *70*, 395-403.
13. Zheljazkov, V.D.; Cantrell, C.L.; Astatkie, T.; Jeliazkova, E. Distillation time effect on lavender essential oil yield and composition. *J. Oleo Sci.* **2013**, *62*, 195-199.
14. Yankov, P.; Drumeva, M. Effects of different main soil tillage methods on the vertical distribution of sunflower seeds in the soil layer and plant development. *Yuzuncu Yıl Univ. J. Agric. Sci.* **2021**, *31*, 396-407.
15. Zagorcheva, T.; Stanev, S.; Angelova, V. Variation in essential oil yield and composition among lavender fields in Bulgaria. *Ind. Crops Prod.* **2020**, *145*, 111975. https://doi.org/10.1016/j.indcrop.2019.111975
16. Pokajewicz, K.; Białoń, M.; Svydenko, L.; Hudz, N.; Balwierz, R.; Marciniak, D.; Wieczorek, P.P. Comparative Evaluation of the Essential Oil of the New Ukrainian *Lavandula angustifolia* and *Lavandula x intermedia* Cultivars Grown on the Same Plots. *Molecules* **2022**, *27*, 2152. https://doi.org/10.3390/molecules27072152
17. Stanev, S.; Angelova, V. Assessment of productivity and oil composition in Bulgarian lavender populations. *J. Essent. Oil Bear. Plants* **2022**, *25*, 23–31. https://doi.org/10.1080/0972060X.2022.2044183
18. Collard, B.C.Y.; Mackill, D.J. Start codon targeted (SCoT) polymorphism: A simple, novel DNA marker technique for generating gene-targeted markers in plants. *Plant Mol. Biol. Rep.* **2009**, *27*, 86–93. https://doi.org/10.1007/s11105-008-0060-5
19. Koohdar, F.; Sheidai, M. Genetic diversity and population structure in medicinal plant Melissa officinalis L. (Lamiaceae). *Genet. Resour. Crop Evol.* **2022**, *69*, 1753-1758.
20. Beicu, R.; Popescu, S.; Neacsu, A.; Imbrea, I.M. Applicability of SCoT (Start Codon Targeted) markers in evaluation of Thymus genetic variability. *Res. J. Agric. Sci.* **2020**, *52*, 24–30.
21. Karagöz, H.; Hosseinpour, A.; Karagöz, F.P.; Cakmakci, R.; Haliloglu, K. Dissection of genetic diversity and population structure in oregano (*Origanum acutidens* L.) genotypes based on agro-morphological properties and SCoT markers. *Biologia* **2022**, *77*, 1231–1247.
22. Salama, A.M.; Osman, E.A.; El-Tantawy, A.A. Taxonomical studies on four *Mentha* species grown in Egypt through morpho-anatomical characters and SCoT genetic markers. *Plant Arch.* **2019**, *19*, 2273–2286.

 doi:10.20944/preprints202604.2001.v1

23. Liu, S.; Wang, Y.; Song, Y.; Khayatnezhad, M.; Minaeifar, A.A. Genetic variations and interspecific relationships in *Salvia* (Lamiaceae) using SCoT molecular markers. *Caryologia* **2021**, *74*, 77–89.
24. Bidyananda, N.; Jamir, I.; Nowakowska, K.; Varte, V.; Vendrame, W.A.; Devi, R.S.; Nongdam, P. Plant Genetic Diversity Studies: Insights from DNA Marker Analyses. *Int. J. Plant Biol.* **2024**, *15*, 607-640. https://doi.org/10.3390/ijpb15030046
25. Zhelyazkova, M.; Badzhelova, V.; Stanev, S. Assessment of Genetic Diversity Among Bulgarian Lavender Varieties Using Scot Markers. *Agronomy* **2025**, *15*, 2889. https://doi.org/10.3390/agronomy15122889
26. Zhelyazkova, M.T. Capability of ISSR, SCoT and CEAP Markers for Genetic Diversity Assessment of Lavender (*Lavandula angustifolia* Mill.) Genotypes. *Int. J. Plant Biol.* **2026**, *17*, 8. https://doi.org/10.3390/ijpb17010008
27. Al-Khayri, J.M.; Mahdy, E.; Taha, H.; Eldomiaty, A.S.; Abd-Elfattah, M.A.; Abdel, L.A.; Rezk, A.A.; Shehata, W.F.; Almaghasla, M.I.; Shalaby, T.A.; et al. Genetic and morphological diversity assessment of five kalanchoe genotypes by SCoT, ISSR and RAPD-PCR markers. *Plants* **2022**, *11*, 1722.
28. Deng, L.; Liang, Q.; He, X.; Luo, C.; Chen, H.; Qin, Z. Investigation and analysis of genetic diversity of diospyros germplasms using SCoT molecular markers in guangxi. *PLoS ONE* **2015**, *10*, e136510.
29. Rai, M.K. Start codon targeted (SCoT) polymorphism marker in plant genome analysis: Current status and prospects. *Planta* **2023**, *257**, 34.
30. Shaban, A.S.; Arab, S.A.; Basuoni, M.M.; Abozahra, M.S.; Abdelkawy, A.M.; Mohamed, M.M.; Merah, O. SCoT, ISSR, and SDS-PAGE investigation of genetic diversity in several egyptian wheat genotypes under normal and drought conditions. *Int. J. Agron.* **2022**, *2022*, 7024028.
31. Poczai, P.; Varga, I.; Laos, M.; Cseh, A.; Bell, N.; Valkonen, J.P.; Hyvönen, J. Advances in plant gene-targeted and functional markers: A review. *Plant Methods* **2013**, *9*, 6.
32. Georgieva, P.; Rusanov, K.; Rusanova, M.; Kitanova, M.; Atanassov, I. Construction of Simple Sequence Repeat-Based Genetic Linkage Map and Identification of QTLs for Accumulation of Floral Volatiles in Lavender (*Lavandula angustifolia* Mill.). *Int. J. Mol. Sci.* **2025**, *26*, 3705. https://doi.org/10.3390/ijms26083705
33. Ma, L. The complete chloroplast genome sequence of the fragrant plant *Lavandula angustifolia* (Lamiaceae). *Mitochondr. DNA Part B* **2018**, *3*, 135–136.
34. Wang, J.; Liu, X.; Zhang, M.; Liu, R. The mitochondrial genome of *Lavandula angustifolia* Mill. (Lamiaceae) sheds light on its genome structure and gene transfer between organelles. *BMC Genom.* **2024**, *25*, 929.
35. Alp Furan, M.; Yildiz, F.; Kaya, O. Exploring the Complete Chloroplast Genomes of Key Lamiaceae Species Uncovers the Secrets of Evolutionary Dynamics and Phylogenetic Relationships. *J. Plant Growth Regul.* **2025**, *44*, 3974–3987.
36. Li, J.; Wang, Y.; Dong, Y.; Zhang, W.; Wang, D.; Bai, H.; ...; Shi, L. The chromosome-based lavender genome provides new insights into Lamiaceae evolution and terpenoid biosynthesis. *Hortic. Res.* **2021**, *8*.
37. Fopa Fomeju, B.; Brunel, D.; Bérard, A.; Rivoal, J.B.; Gallois, P.; Le Paslier, M.C.; Bouverat-Bernier, J.P. Quick and efficient approach to develop genomic resources in orphan species: Application in *Lavandula angustifolia*. *PLoS ONE* **2020**, *15*, e0243853.
38. Kelimujiang, K.; Zhang, W.; Zhang, X.; Waili, A.; Tang, X.; Chen, Y.; Chen, L. Genome-wide investigation of WRKY gene family in *Lavandula angustifolia* and potential role of LaWRKY57 and LaWRKY75 in the regulation of terpenoid biosynthesis. *Front. Plant Sci.* **2024**, *15*, 1449299.
39. ISO 3515:2002. Oil of lavender (Lavandula angustifolia Mill.) – Specification. International Organization for Standardization: Geneva, Switzerland, **2002**.
40. Roldan-Ruiz, I.; Dendauw, J.; Vanbockstaele, E.; Depicker, A.; De Loose, M. AFLP markers reveal high polymorphic rates in ryegrasses (*Lolium* spp.). *Mol. Breed.* **2000**, *6*, 125–134.
41. Nagaraju, J.; Damodar, R.K.; Nagaraja, G.M.; Sethuraman, B.N. Comparison of multilocus RFLPs and PCR-based marker systems for genetic analysis of the silkworm, *Bombyx mori*. *Heredity* **2001**, *86*, 588–597.
42. Varshney, R.K.; Chabane, K.; Hendre, P.S.; Aggarwal, R.K.; Graner, A. Comparative assessment of EST-SSR, EST-SNP and AFLP markers for evaluation of genetic diversity and conservation of genetic resources using wild, cultivated and elite barleys. *Plant Sci.* **2007**, *173*, 638–649.
43. Prevost, A.; Wilkinson, M.J. A new system of comparing PCR primers applied to ISSR fingerprinting of potato cultivars. *Theor. Appl. Genet.* **1999**, *98*, 107–112.

 doi:10.20944/preprints202604.2001.v1

44. Peakall, R.; Smouse, P.E. Genalex 6.5: Genetic analysis in excel. Population genetic software for teaching and research—An update. *Bioinformatics* **2012**, *28*, 2537–2539.
45. Nei, M. Genetic distance between populations. *Am. Nat.* **1972**, *106*, 283–292. https://doi.org/10.1086/282771
46. Nei, M. Estimation of average heterozigosity and genetic distance from small number of individuals. *Genetics* **1978**, *89*, 583-590.
47. Kumar, S.; Stecher, G.; Suleski, M.; Sanderford, M.; Sharma, S.; Tamura, K. MEGA12: Molecular Evolutionary Genetic Analysis version 12 for adaptive and green computing. *Mol. Biol. Evol.* **2024**, *41*, msae263.
48. Letunic, I.; Bork, P. Interactive Tree of Life (iTOL) v6: Recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Res.* **2024**, *52*, W78–W82.
49. Pritchard, J.K.; Stephens, M.; Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics* **2000**, *155*, 945–959.
50. Pan, H.; Deng, L.; Zhu, K.; Shi, D.; Wang, F.; Cui, G. Evaluation of genetic diversity and population structure of *Annamocarya sinensis* using SCoT markers. *PLoS ONE* **2024**, *19*, e0309283.
51. Kopelman, N.M.; Mayzel, J.; Jakobsson, M.; Rosenberg, N.A.; Mayrose, I. Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Mol. Ecol. Resour.* **2015**, *15*, 1179-1191.
52. Bradbury, P.J.; Zhang, Z.; Kroon, D.E.; Casstevens, T.M.; Ramdoss, Y.; Buckler, E.S. Tassel: Software for association mapping of complex traits in diverse samples. *Bioinformatics* **2007**, *23*, 2633–2635.
53. Holm, S. A simple sequentially rejective multiple test procedure. *Scand. J. Stat.* **1979**, *6*, 65–70.
54. Arlot, S.; Celisse, A. A survey of cross-validation procedures for model selection. *Stat. Surv.* **2010**, *4*, 40–79. https://doi.org/10.1214/09-SS054
55. Copernicus Climate Change Service, Climate Data Store. ERA5 Land hourly time-series data on single levels from 1950 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). Available online: https://cds.climate.copernicus.eu/datasets/reanalysis-era5-land-timeseries (accessed on 9 April 2026).
56. Muñoz-Sabater, J.; Dutra, E.; Agustí-Panareda, A.; Albergel, C.; Arduini, G.; Balsamo, G.; Boussetta, S.; Choulga, M.; Harrigan, S.; Hersbach, H.; et al. ERA5-Land: a state-of-the-art global reanalysis dataset for land applications. *Earth Syst. Sci. Data* **2021**, *13*, 4349–4383. https://doi.org/10.5194/essd-13-4349-2021
57. Pelosi, A.; Belfiore, O.R.; D'Urso, G.; Chirico, G.B. Assessing crop water requirement and yield by combining ERA5-land reanalysis data with CM-SAF satellite-based radiation data and Sentinel-2 satellite imagery. *Remote Sens.* **2022**, *14*, 6233. https://doi.org/10.3390/rs14246233
58. Pelosi, A.; Terribile, F.; D'Urso, G.; Chirico, G. Comparison of ERA5-land and UERRA MESCAN-SURFEX reanalysis data with spatially interpolated weather observations for the regional assessment of reference evapotranspiration. *Water* **2020**, *12*, 1669. https://doi.org/10.3390/w12061669
59. Croci, M.; Ragazzi, M.; Grassi, A.; Impollonia, G.; Amaducci, S. Assessing the temporal transferability of machine learning models for predicting processing pea yield and quality using Sentinel-2 and ERA5-land data. *Smart Agric. Technol.* **2025**, *12*, 1-20.
60. Ohio State University Extension. Understanding and interpreting p-values in research (ANR-40). The Ohio State University, **2020**. Available online: https://ohioline.osu.edu/factsheet/anr-40 (accessed on 26 April 2025).
61. Shrestha, J. P-value test of significance in agricultural research. LinkedIn, **2022**. Available online: https://www.linkedin.com/pulse/p-value-test-significance-agricultural-research-jiban-shrestha (accessed on 26 April 2025).
62. Crișan, I.; Ona, A.; Vârban, D.; Muntean, L.; Vârban, R.; Stoie, A.; Mihăiescu, T.; Morea, A. Current Trends for Lavender (*Lavandula angustifolia* Mill.) Crops and Products with Emphasis on Essential Oil Quality. *Plants* **2023**, *12*, 357. https://doi.org/10.3390/plants12020357
63. Bogdan, M.; Bungau, S.; Tit, D.M.; Copolovici, L.; Behl, T.; Otrisal, P.; ...; Copolovici, D.M. Variations in the chemical composition of the essential oil of *Lavandula angustifolia* Mill., Moldoveanca 4 Romanian variety. *Rev. Chim.* **2020**, *71*, 307-315.

 doi:10.20944/preprints202604.2001.v1

64. Stanev, S. Evaluation of the stability and adaptability of the Bulgarian lavender (*Lavandula angustifolia* Mill.) cultivars' yield. *Agric. Sci. Technol.* **2010**, *2*, 121–123.
65. Georgieva, R.; Kırchev, H.; Delıbaltova, V.; Matev, A.; Chavdarov, P.; Raycheva, T. Essential oil chemical composition of lavender varieties cultivated in an untraditional agro-ecological region. *Yuzuncu Yıl Univ. J. Agric. Sci.* **2022**, *32*, 98-105.
66. Lee, Y.L.; Ding, P. Production of Essential Oil in Plants: Ontogeny, Secretory Structures and Seasonal Variations. *Pertanika J. Sch. Res. Rev.* **2016**, *2*, 1–10.
67. Hassiotis, C.N.; Lazari, D.; Vlachonasios, K. The Effects of Habitat Type and Diurnal Harvest on Essential Oil Yield and Composition of *Lavandula angustifolia* Mill. *Fresenius Environ. Bull.* **2010**, *19*, 1491–1498.
68. Barbieri, C.; Borsotto, P. Essential Oils: Market and Legislation. IntechOpen: London, UK, **2018**. ISBN 978-1-78923-780-1.
69. Najar, B.; Pistelli, L.; Fratini, F. Exploitation of Marginal Hilly Land in Tuscany through the Cultivation of *Lavandula angustifolia* Mill.: Characterization of Its Essential Oil and Antibacterial Activity. *Molecules* **2022**, *27*, 3216.
70. Détár, E.; Németh, É.Z.; Gosztola, B.; Demján, I.; Pluhár, Z. Effects of Variety and Growth Year on the Essential Oil Properties of Lavender (*Lavandula angustifolia* Mill.) and Lavandin (*Lavandula x Intermedia* Emeric Ex Loisel.). *Biochem. Syst. Ecol.* **2020**, *90*, 104020.
71. Sałata, A.; Buczkowska, H.; Nurzyńska-Wierdak, R. Yield, Essential Oil Content, and Quality Performance of *Lavandula angustifolia* Leaves, as Affected by Supplementary Irrigation and Drying Methods. *Agriculture* **2020**, *10*, 590. https://doi.org/10.3390/agriculture10120590
72. Mihalascu, C.; Tudor, V.; Bolohan, C.; Mihalache, M.; Ionuț, R. The Effect of Different Fertilization upon the Growth and Yield of Some *Lavandula angustifolia* (Mill.) Varieties Grown in South East Romania. *Sci. Pap. Ser. B Hortic.* **2020**, *64*, 685–692.
73. Mavandi, P.; Abbaszadeh, B.; Emami Bistgani, Z.; Barker, A.V.; Hashemi, M. Biomass, Nutrient Concentration and the Essential Oil Composition of Lavender (*Lavandula angustifolia* Mill.) Grown with Organic Fertilizers. *J. Plant Nutr.* **2021**, *44*, 3061–3071. https://doi.org/10.1080/01904167.2021.1927083
74. Valchev, H.; Kolev, Z.; Stoykova, B.; Kozuharova, E. Pollinators of *Lavandula angustifolia* Mill., an Important Factor for Optimal Production of Lavender Essential Oil. *BioRisk* **2022**, *17*, 297–307.
75. Radev, Z. Honey Bee (Apis mellifera L.) Pollination as an ecological method to increase the quality of lavender essential oil. *Agric. Conspectus Sci.* **2023**, *88*, 85-88.
76. Çolak, İ.A.; Çelik, S.A. Determination of essential oil yields and composition lavender and lavandin cultivars (*Lavandula* sp.) cultivated in Tuzlukçu, Konya. *Selcuk J. Agric. Food Sci.* **2023**, *37*, 582-588.
77. Demasi, S.; Caser, M.; Lonati, M.; Cioni, P.L.; Pistelli, L.; Najar, B.; Scariot, V. Latitude and Altitude Influence Secondary Metabolite Production in Peripheral Alpine Populations of the Mediterranean Species *Lavandula angustifolia* Mill. *Front. Plant Sci.* **2018**, *9*, 1–12.
78. Aprotosoaie, A.C.; Gille, E.; Trifan, A.; Luca, V.S.; Miron, A. Essential Oils of Lavandula Genus: A Systematic Review of Their Chemistry. *Phytochem. Rev.* **2017**, *16*, 761–799.
79. Delibaltova, V.; Manhart, S.; Stoychev, I.; Nedyalkov, M. Comparative research of productive and qualitative indicators in lavender varieties cultivated in Eastern Bulgaria. *Sci. Pap. Ser. A Agron.* **2024**, *67*.
80. Despinasse, Y.; Moja, S.; Soler, C.; Jullien, F.; Pasquier, B.; Bessière, J.M.; ...; Nicolè, F. Structure of the chemical and genetic diversity of the true lavender over its natural range. *Plants* **2020**, *9*, 1640.
81. Yan, H.; Zhang, Y.; Zeng, B.; Yin, G.; Zhang, X.; Ji, Y.; ...; Yan, Y. Genetic diversity and association of EST-SSR and SCoT markers with rust traits in orchardgrass (Dactylis glomerata L.). *Molecules* **2016**, *21*, 66.
82. Wang, F.; Chen, X.; Huang, Z.; Wei, L.; Wang, J.; Wen, S.; ...; Zhou, Y. Phenotypic characterization and marker–trait association analysis using SCOT markers in chrysanthemum (*Chrysanthemum morifolium* Ramat.) germplasm. *Genes* **2025**, *16*, 664.
83. Laftouhi, M.S.; Bouhrim, M.; Zerkani, H. Physiological Adaptation of Plants Under Stressful Climatic Conditions. *Korean J. Young Sch.* **2023**, *38*, 1893-1907.
84. Eldeghedy, H.I.; El-Gendy, A.E.N.G.; Nassrallah, A.A.; Aboul-Enein, A.M.; Omer, E.A. Comparative chemical profiles of lavandula species essential oils grown in Egypt and others from France and Australia: Evidence from chemometric analysis. *J. Essent. Oil Bear. Plants* **2022**, *25*, 52-63.

 doi:10.20944/preprints202604.2001.v1

85. Smigielski, K.; Raj, A.; Krosowiak, K.; Gruska, R. Chemical composition of the essential oil of *Lavandula angustifolia* cultivated in Poland. *J. Essent. Oil Bear. Plants* **2009**, *12*, 338-347.