

# EVALUATION OF MORPHOLOGICAL DIVERSITY BASED ON LEAF AND FRUIT TRAITS IN ROSEHIP (*ROSA* SPP.) GENOTYPES USING PRINCIPAL COMPONENT ANALYSIS

Halit ALKAN<sup>1</sup>, Ercan YILDIZ<sup>1</sup>, Ahmet SÜMBÜL<sup>2</sup>, Mehmet YAMAN<sup>1\*</sup>

<sup>1</sup>Department of Horticulture, Faculty of Agriculture, Erciyes University, Kayseri, Türkiye

<sup>2</sup>Suğehri Timur Karabal Vocational School, Department of Plant and Animal Production, Sivas Cumhuriyet University, Sivas, Türkiye



## Abstract

This study evaluated morphological and biochemical variation among 31 rosehip (*Rosa canina* L.) genotypes using 28 leaf- and fruit-related traits. Data were analyzed through correlation analysis, principal component analysis (PCA), and hierarchical clustering (UPGMA) to examine trait relationships and genotype similarities. Correlation analysis revealed significant positive and negative associations, with strong positive correlations mainly among fruit and leaf size traits, whereas fruit color traits generally showed negative relationships with size-related traits. PCA identified 11 principal components with eigenvalues greater than 1, explaining 84.59% of total variation; the first three components accounted for 37.29% and represented the most discriminative traits. UPGMA grouped the genotypes into two main clusters, indicating clear morphological divergence. Overall, the results revealed substantial diversity among genotypes and showed that multivariate analyses are effective for identifying key traits and selecting promising material for breeding and genetic improvement.

**Keywords:** Rosehip, Genotype, Principal Component Analysis (PCA), UPGMA clustering

## 1. INTRODUCTION

The Rosaceae family is one of the most species-rich groups of plants, widely distributed across temperate and tropical regions of the Northern Hemisphere (Simpson, 2019). Among the members of this family, rosehip (*Rosa* spp.) stands out due to its widespread occurrence in natural flora, high adaptability, and the presence of bioactive compounds beneficial to human health (Doğan et al., 2007). One of the most common species, *Rosa canina* L., grows naturally in almost all regions of Türkiye and exhibits broad ecological tolerance. This species is distributed across a wide geographical range extending from Central and Western Asia to Europe, and from North Africa to the Caucasus, and is particularly notable for its resistance to high altitudes and harsh climatic conditions (Anşın, 1996).

Türkiye, owing to its diverse climate types, complex topography, and presence of microclimatic zones, is an important center of genetic diversity for rosehip. A large number of rosehip species and genotypes grow naturally in the country, indicating a significant potential for genetic variation. However, a considerable portion of these genetic resources has not yet been sufficiently characterized or economically utilized. In particular, the determination of morphological and

biochemical traits of genotypes within natural populations is of great importance for both the conservation of genetic resources and breeding studies (Şehirali et al., 2005; Salman-Özen, 2013).

Rosehip plants generally grow in shrub form, are deciduous during winter, and typically possess a thorny structure. However, it is also known that some genotypes exhibit reduced thorn density or are completely thornless. The plant can grow across a wide altitudinal range from sea level to high elevations and demonstrates high tolerance to environmental stresses such as low temperatures and drought. These characteristics enable rosehip to be cultivated under diverse ecological conditions (Perry, 1988).

Rosehip fruits are rich in ascorbic acid (vitamin C), phenolic compounds, carotenoids, and flavonoids, which makes them highly valuable as functional foods and natural health products. In addition to fresh consumption, rosehip fruits are processed into marmalade, jam, fruit juice, tea, and various dietary supplements (Doğan et al., 2007; Özdemir et al., 2022). This increases both the economic and nutritional value of rosehip and encourages its domestication and cultivation.

Significant variation exists among rosehip genotypes naturally distributed in wild populations in terms of fruit size, shape, color, and leaf characteristics. This variation provides an important opportunity for selection studies, allowing the identification of genotypes with superior traits (Kazankaya et al., 2005; Duran and Kılıç, 2023). However, these genotypes need to be systematically evaluated and compared using standardized data. At this point, morphological characterization studies emerge as one of the primary tools for the identification and classification of genotypes.

In recent years, the use of multivariate statistical methods has become increasingly common in studies where multiple traits are evaluated simultaneously. In particular, Principal Component Analysis (PCA), correlation analysis, and clustering methods provide effective approaches for revealing relationships among genotypes and understanding variation. These methods simplify complex relationships within datasets, identify discriminating traits, and provide a more scientific basis for selecting genotypes to be used in breeding programs. In this context, the present study evaluates morphological characteristics related to leaf and fruit traits of different rosehip genotypes and analyzes the obtained data using multivariate statistical techniques. The aim of the study is to reveal morphological variation among genotypes, identify key discriminating traits, and determine genotypes that may contribute to breeding programs.

## 2. MATERIALS AND METHODS

This study was conducted in 2025 on naturally growing rosehip (*Rosa canina* L.) genotypes in the Kadışehri district of Yozgat, Türkiye. The plant material consisted of 31 individual genotypes growing in natural habitats in the district and its surrounding villages. Each selected plant was considered as a separate genotype based on observed differences in morphological characteristics and yield potential.

For characterization, 30 fruit and leaf samples were collected from each genotype. A total of 28 selected traits, including morphological and biochemical characteristics, were evaluated. Morphological traits included leaf width, leaf length, leaf area, petiole length and thickness, thorniness status, fruit weight, fruit width and length, fruit flesh thickness, fruit flesh ratio, seed weight, number of seeds, soluble solids content (SSC), and total dry matter content. In addition, biochemical traits such as total phenolic content, total flavonoid content, and antioxidant activity in both leaf and fruit samples were determined.

Statistical analyses were performed using JMP Pro 17 software. Relationships among traits were evaluated using correlation analysis. Principal Component Analysis (PCA) was conducted to identify the main sources of variation and to determine the most discriminative traits. Furthermore, hierarchical cluster analysis was performed using the Unweighted Pair Group Method with Arithmetic Mean (UPGMA) to classify genotypes based on their similarity.

### 3. RESULTS AND DISCUSSION

Principal Component Analysis (PCA) is applied to reveal variation among genotypes, determine relationships among numerous morphological variables in the dataset, and reduce the number of variables while minimizing information loss in order to identify discriminating traits (Iezzoni and Pritts, 1991). PCA generates new and independent components by considering the correlation structure among variables, particularly in multivariate datasets, thus allowing similarities and differences among genotypes to be more clearly demonstrated.

This method has been widely used in many fruit species such as grape (Sümbül and Yıldız, 2024), oleaster (Say et al., 2024), barberry (Yaman et al., 2024), blackberry (Doğan et al., 2025), walnut (Sümbül, 2025a), and mulberry (Sümbül, 2025b) to determine genetic diversity, distinguish important morphological and pomological traits, and guide selection studies. Similarly, it stands out as an effective and reliable statistical tool in the evaluation of morphological and agronomic variation in rosehip genotypes (Soare et al., 2015; Mertoğlu et al., 2024; Bozdoğan et al., 2025).

Within the scope of this study, PCA was performed on a total of 28 morphological traits representing leaf and fruit characteristics of 31 selected rosehip genotypes. The principal components obtained as a result of the analysis explained a large proportion of the variation among the examined traits, enabling the classification of genotypes, identification of similar groups, and determination of prominent characteristics for selection. In addition, by considering the variance explained by each component, the main morphological traits effective in the dataset were identified, and it was revealed which characteristics were responsible for the differences among genotypes (Table 1).

According to the results of the Principal Component Analysis (PCA), 11 principal components with eigenvalues greater than 1 were identified, and these components explained 84.59% of the total variation. The first three principal components accounted for 37.29% of the total variation. This result indicates that the traits contributing to the first three principal components exhibit the highest level of variability among genotypes and have the greatest influence on genotype differentiation (Iezzoni and Pritts, 1991). The effects of the examined traits differed across the first three principal components. In the first principal component, which explained 15.04% of the total variation, leaf width, leaf length, leaf area, petiole length, leaf b\* color value, fruit weight, fruit width, and fruit length had positive effects, whereas fruit L\* color value, fruit a\* color value, fruit b\* color value, and total phenolic content of the fruit had negative effects. In the second principal component, which explained 12.18% of the total variation, petiole length, petiole thickness, leaf b\* color value, fruit weight, and seed weight contributed positively, while fruit width, fruit length, fruit L\* color value, fruit b\* color value, fruit flesh thickness, and total phenolic content of the fruit had negative effects. In the third principal component, which explained 10.08% of the total variation, leaf width, leaf length, leaf area, fruit a\* color value, fruit flesh ratio, and total dry matter ratio had positive effects, whereas petiole thickness and leaf L\* and a\* color values had negative effects. Similar findings have been reported in previous studies conducted on rosehip (*Rosa canina* L.) in Türkiye

(Bozhuyuk et al., 2021; Bozdoğan et al., 2025) and Iran (Tafreshi et al., 2025), where comparable variables were found to contribute to the first three principal components.

*Table 1. Variation values of principal components (PCs) in rosehip genotypes*

| Characteristics                  | PCA1         | PCA2         | PCA3         | PCA11        |
|----------------------------------|--------------|--------------|--------------|--------------|
| Leaf width                       | <b>0.27</b>  | -0.10        | <b>0.33</b>  | 0.02         |
| Leaf length                      | <b>0.26</b>  | -0.12        | <b>0.32</b>  | 0.03         |
| Leaf area                        | <b>0.31</b>  | -0.17        | <b>0.20</b>  | 0.17         |
| Petiole length                   | <b>0.21</b>  | <b>0.25</b>  | 0.08         | -0.26        |
| Petiole thickness                | 0.09         | <b>0.24</b>  | <b>-0.32</b> | 0.13         |
| Leaf L* color value              | 0.02         | -0.07        | <b>-0.31</b> | -0.08        |
| Leaf a* color value              | 0.13         | 0.08         | <b>-0.32</b> | 0.05         |
| Leaf b* color value              | <b>0.26</b>  | 0.22         | -0.09        | 0.23         |
| Fruit weight                     | <b>0.20</b>  | -0.17        | -0.17        | 0.28         |
| Fruit width                      | <b>0.23</b>  | <b>-0.24</b> | -0.10        | -0.02        |
| Fruit length                     | <b>0.30</b>  | <b>-0.21</b> | 0.00         | -0.19        |
| Fruit L* color value             | <b>-0.27</b> | <b>-0.27</b> | -0.02        | 0.11         |
| Fruit a* color value             | <b>-0.21</b> | <b>0.23</b>  | <b>0.27</b>  | 0.11         |
| Fruit b* color value             | <b>-0.31</b> | <b>-0.27</b> | -0.02        | 0.10         |
| Fruit flesh thickness            | -0.05        | <b>-0.21</b> | 0.13         | 0.19         |
| Fruit flesh ratio                | -0.18        | 0.05         | <b>0.22</b>  | 0.18         |
| Seed weight                      | -0.09        | <b>0.28</b>  | -0.03        | -0.03        |
| Number of seeds                  | 0.08         | -0.07        | -0.16        | 0.04         |
| Soluble solids content           | 0.16         | -0.09        | 0.19         | 0.42         |
| Total dry matter content         | -0.11        | <b>0.20</b>  | <b>0.24</b>  | -0.18        |
| Thorniness status                | -0.03        | 0.15         | -0.17        | 0.14         |
| Productivity status              | 0.10         | -0.02        | -0.03        | -0.07        |
| Total phenolic content in fruit  | <b>-0.31</b> | <b>-0.27</b> | -0.02        | 0.10         |
| Total flavonoid content in fruit | -0.01        | 0.19         | -0.02        | 0.15         |
| Antioxidant activity of fruit    | 0.13         | 0.10         | 0.13         | -0.10        |
| Total phenolic content in leaf   | -0.03        | <b>0.27</b>  | 0.04         | 0.35         |
| Total flavonoid content in leaf  | -0.07        | -0.07        | 0.06         | 0.01         |
| Antioxidant activity of leaf     | 0.01         | <b>0.20</b>  | 0.16         | 0.30         |
| <b>Eigenvalue</b>                | <b>4.36</b>  | <b>3.53</b>  | <b>2.92</b>  | <b>1.00</b>  |
| <b>Percentage (%)</b>            | <b>15.04</b> | <b>12.18</b> | <b>10.08</b> | <b>3.45</b>  |
| <b>Cumulative percentage (%)</b> | <b>15.04</b> | <b>27.22</b> | <b>37.29</b> | <b>84.59</b> |

Revealing the relationships among the traits examined in studies is of great importance for the effectiveness and success of breeding programs. Particularly in studies where numerous morphological and pomological characteristics are evaluated, determining correlations among traits enables the development of indirect selection criteria. In this context, the correlation analysis graph of the traits examined in rosehip genotypes is presented in Figure 1. Significant positive and negative correlations were identified among the examined traits. The highest positive correlations were determined between total phenolic content of the fruit and fruit b\* color value ( $r = 1.00$ ), and between leaf length and leaf width ( $r = 0.94$ ). This suggests that these traits vary together and may

be controlled by similar biological processes. In contrast, the highest negative correlations were observed between fruit a\* color value and both fruit weight ( $r = -0.44$ ) and leaf L\* color value ( $r = -0.42$ ). Negative correlations indicate that the related traits vary in opposite directions, suggesting that an increase in one trait during the selection process may lead to a decrease in the other.

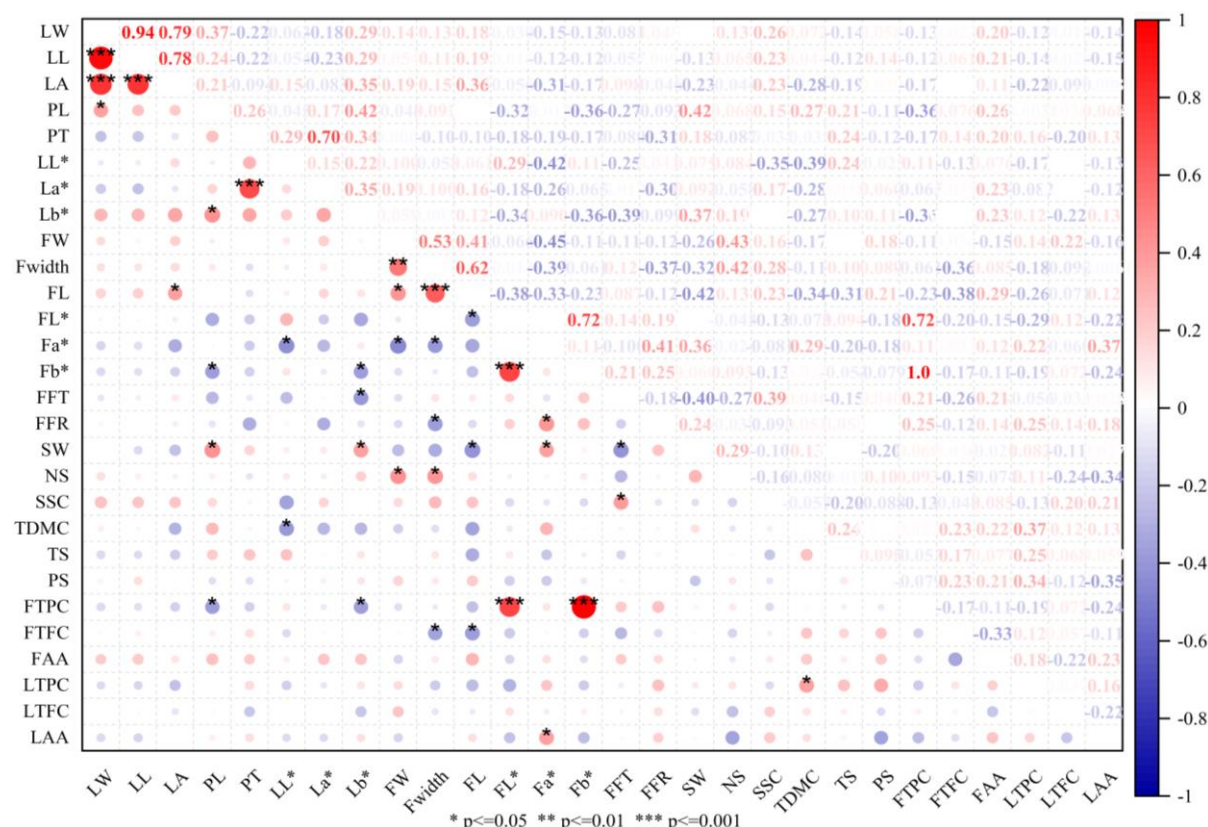


Figure 1. Correlation analysis of leaf and fruit traits of the examined rosehip genotypes

Leaf width: LW, Leaf length: LL, Leaf area: LA, Petiole length: PL, Petiole thickness: PT, Leaf L\* color value: LL\*, Leaf a\* color value: La\*, Leaf b\* color value: Lb\*, Fruit weight: FW, Fruit width: Fwidth, Fruit length: FL, Fruit L\* color value: FL\*, Fruit a\* color value: Fa\*, Fruit b\* color value: Fb\*, Fruit flesh thickness: FFT, Fruit flesh ratio: FFR, Seed weight: SW, Number of seeds: NS, Soluble solids content: SSC, Total dry matter content: TDMC, Thorniness status: TS, Productivity status: PS, Total phenolic content in fruit: FTPC, Total flavonoid content in fruit: FTFC, Antioxidant activity of fruit: FAA, Total phenolic content in leaf: LTPC, Total flavonoid content in leaf: LTFC, Antioxidant activity of leaf: LAA

Overall, it was determined that traits related to leaf size showed positive and significant correlations among themselves, and similarly, traits related to fruit size also exhibited positive and significant relationships within their own group. In contrast, fruit color parameters were generally found to display negative correlations with both fruit size traits and leaf color characteristics. This suggests the possible existence of a physiological or genetic balance mechanism between morphological size and color traits. The correlation relationships among the examined parameters contribute to understanding the direct or indirect effects of one trait on others. Strong and significant

relationships among traits provide important advantages for reeders in determining selection criteria, thereby enabling breeding programs to be conducted more rapidly, efficiently, and in a goal-oriented manner (Garazhian et al., 2020). The correlation findings obtained in this study are generally consistent with previous studies conducted on rosehip (*Rosa canina* L.) (Bozdoğan et al., 2025; Tafreshi et al., 2025), supporting the reliability of the present study and its consistency with the existing literature.

The dendrogram obtained from the UPGMA (Unweighted Pair Group Method with Arithmetic Mean) cluster analysis based on fruit and leaf characteristics of rosehip genotypes is presented in Figure 2. Examination of the dendrogram shows that the genotypes are generally grouped into two main clusters. In the first main cluster (Group A), genotypes G6, G10, and G23 are included, while all remaining genotypes form the second main cluster (Group B).

This indicates that the genotypes in Group A are clearly separated from the others in terms of the studied morphological traits and exhibit a distinct morphological structure. Group B is further divided into subgroups, and within Group B1, genotypes G5 and G12 were found to be the most closely related genotypes. This proximity indicates a high degree of similarity between these genotypes in terms of fruit and leaf characteristics. Similarly, in subgroup B2, genotypes G18 and G31 were also determined to have very similar traits.

This clustering pattern provides important information regarding genetic similarity and diversity based on morphological characteristics. In particular, closely clustered genotypes may have similar genetic structures, whereas genotypes located in different clusters may offer greater variation potential for breeding studies. In this context, selecting genotypes from different clusters as parents may contribute to obtaining new combinations with higher variability.

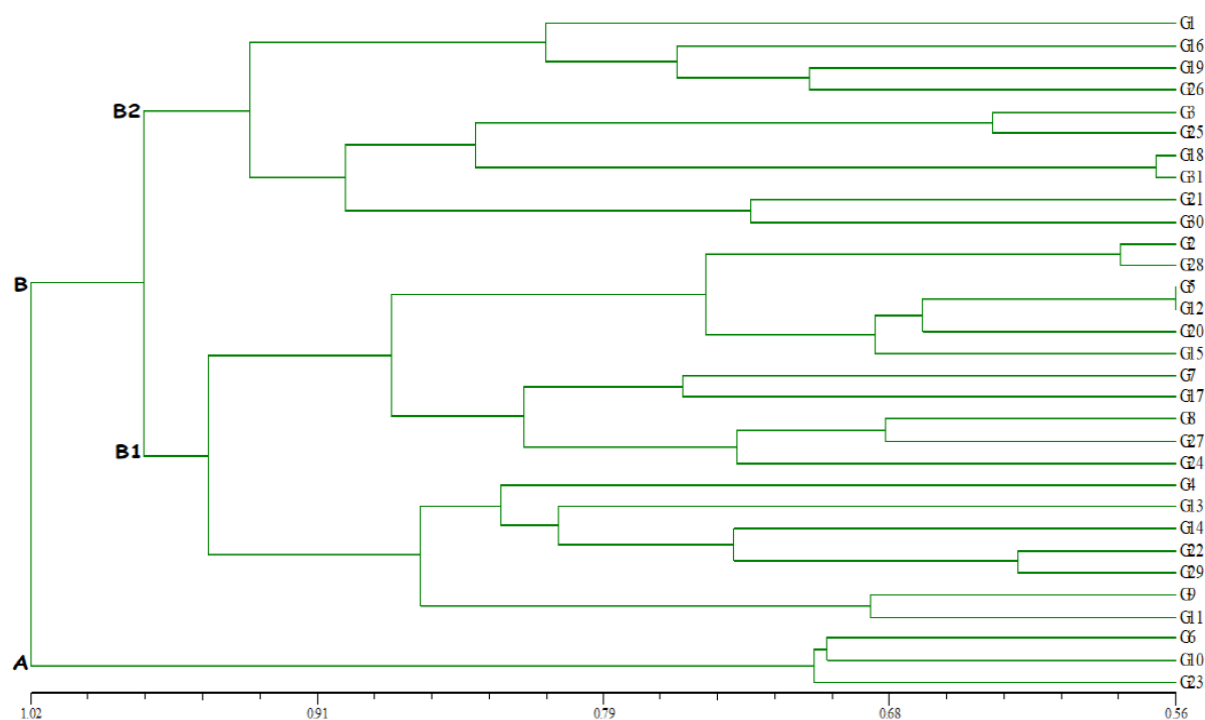


Figure 2. UPGMA cluster analysis of rosehip genotypes based on leaf and fruit traits

#### 4. CONCLUSION

The results of this study clearly demonstrated a high level of morphological and biochemical diversity among the evaluated rosehip (*Rosa canina* L.) genotypes based on leaf and fruit traits. The combined application of correlation analysis, Principal Component Analysis (PCA), and UPGMA clustering provided a comprehensive understanding of the relationships among traits and the extent of variation among genotypes.

PCA revealed that a limited number of principal components explained a substantial proportion of the total variation, indicating that specific morphological and biochemical traits play a key role in genotype differentiation. In addition, the clustering analysis grouped the genotypes into two main clusters, highlighting distinct genetic and phenotypic divergence within the studied population.

The significant correlations identified among traits suggest that certain characteristics can be effectively used as indirect selection criteria in breeding programs. In particular, the clear separation of genotypes into different clusters indicates that the selection of parents from genetically distant groups may enhance variability and breeding efficiency.

Overall, the findings of this study provide valuable insights for the conservation, characterization, and utilization of rosehip genetic resources. The identified variability and key discriminative traits offer a strong scientific basis for future selection and breeding strategies aimed at developing superior genotypes with desirable agronomic and biochemical properties.

#### 5. REFERENCES

- Anşin, R. (1996). Naturally growing *Rosa* L. taxa in the Eastern Black Sea region. Rosehip Symposium. 5–6 September 1996, Gümüşhane/Türkiye, 85-95.
- Bozdoğan, N., Yıldız, E., Tunç, Y., Khadivi, A., Yaman, M., Sümbül, A., Yılmaz, K. U., Ateş, U., Öztürk, B. (2025). Morphological, biochemical, and molecular characterizations of rosehip (*Rosa canina* L.) genotypes naturally grown in the Eastern Central Anatolia region of Türkiye. *Scientia Horticulturae*, 350, 114375.
- Bozhuyuk, M. R., Ercisli, S., Karatas, N., Ekiert, H., Elansary, H. O., Szopa, A. (2021). Morphological and biochemical diversity in fruits of unsprayed *Rosa canina* and *Rosa dumalis* ecotypes found in different agroecological conditions. *Sustainability*, 13(14), 8060.
- Doğan, A., Kazankaya, A., Çelik, F., Ekinçialp, A. (2007). Determination of fruit characteristics of naturally growing rosehips (*Rosa* spp.) in Yüksekova (Hakkari) region. *Proceedings of the 5th National Horticultural Congress*, 4–7 September 2007, Erzurum/Türkiye, 174-179.
- Dogan, H., Aglar, E., Ozturk, B., Tekin, O., Alan, D., Sumbul, A. (2025). Pomological and Biochemical Properties of Blackberry (*Rubus fruticosus*) Genotypes. *Food Science & Nutrition*, 13(7), e70591.
- Duran, F., Kılıç, D.D. (2023). Phenological and pomological characteristics of *Rosa canina* L. Species cultivated and naturally distributed in Amasya Province. *Black Sea J. Agric.* 6 (6), 622–630.
- Garazhian M, Gharaghani A, Eshghi S. (2020). Genetic diversity and inter-relationships of fruit bio-chemicals and antioxidant activity in Iranian wild blackberry species. *Scientific Reports* 10, 18983.
- Iezzoni, A.F., Pritts, M.P. (1991). Applications of Principal Component Analysis to horticultural research. *HortScience*, 26(4), 334-338.
- Kazankaya, A., Turkoglu, N., Yilmaz, M., Balta, M.F. (2005). Pomological description of *Rosa canina* selections from Eastern Anatolia, Turkey. *Int. J. Bot.* 1 (1), 100–102.
- Mertoğlu, K., Sulusoglu-Durul, M., Korkmaz, N., Polat, M., Bulduk, İ., Esatbeyoglu, T. (2024). Screening and classification of rosehip (*Rosa canina* L.) genotypes based on horticultural characteristics. *BMC Plant Biology*, 24: 345.
- Özdemir, N., Pashazadeh, H., Zannou, O., Koca, I. (2022). Phytochemical content, and antioxidant activity, and volatile compounds associated with the aromatic property, of the vinegar produced from rosehip fruit (*Rosa canina* L.). *LWT-Food Science and Technology*, 154, 112716.
- Perry, F. (1988). *Guide to Plants and Flowers*. Simon and Schuster Inc. Spain.
- Salman-Özen, M. (2013). Selection of rosehip (*Rosa* spp.) genetic resources and determination of antioxidant activities in Bolu Central District. MSc Thesis, Selçuk University, Konya/Türkiye, 78 pp.

- Say, A., Sümbül, A., Dirim, E., Yaman, M., Yıldız, E. (2024). Unravelling the genetic diversity of oleaster (*Elaeagnus Angustifolia* L.) with multivariate analysis. *Applied Fruit Science*, 66(2), 719-730.
- Simpson, M.G. (2019). *Plant Systematics* (3rd ed.). Elsevier Academic Press.
- Soare, R., Babeanu, C., Bonea, D., Panita, O. (2015). The content of total phenols, flavonoids and antioxidant activity in rosehip from the spontaneous flora from South Romania. *Scientific Papers. Series A. Agronomy*, 58, 307-314.
- Sümbül, A., Yıldız, E. (2024). Explanation of morphological and biochemical diversity of autochthonous grapes grown in Türkiye (Kelkit Basin) using multivariate analysis. *Turkish Journal of Food and Agriculture Sciences*, 6(2), 160-172.
- Sümbül, A. (2025a). Investigation of fruit trait diversity in walnut (*Juglans regia* L.) genotypes of the Middle Kelkit Valley using multivariate analyses. *Meyve Bilimi*, 12(1), 31-38.
- Sümbül, A. (2025b). Characterization of genetic diversity of mulberry (*Morus alba*) genotypes growing naturally in Northeastern Türkiye (Kelkit Valley) using morphological, biochemical, and molecular markers. *Horticulturae*, 11(3), 298.
- Şehirali, S., Özgen, M., Karagöz, A., Sürek, M., Adak, S., Güvenç, İ., Tan, A., Burak, M., Kaymak, H.Ç. (2005). Conservation and utilization of plant genetic resources. *Proceedings of the VI Technical Congress of the Chamber of Agricultural Engineers, Ankara/Türkiye*, 253-273.
- Tafreshi, Y. M., Eghlima, G., Hatami, M., Vafadar, M. (2025). Exploring the potential impact of salicylic acid and jasmonic acid in promoting seed oil content, vitamin C and antioxidant activity in rosehip (*Rosa canina* L.). *BMC Plant Biology*, 25(1), 249.
- Yaman, M., Sun, M., Sümbül, A., Demirel, F., Tunç, Y., Khadivi, A., Yılmaz, K. U. (2024). Multivariate analysis of morphological variation, biochemical content, antioxidant activity, and molecular characteristics of autochthonous *Berberis crataegina* DC. genotypes in Central Türkiye. *BMC Plant Biology*, 24(1), 1155.