TÜRKİYE CUMHURİYETİ
ADANA ALPARSLAN TÜRKEŞ SCIENCE AND TECHNOLOGY UNIVERSITY

GRADUATE SCHOOL
BIOENGINEERING DEPARTMENT

REEVALUATION OF TAXA INCLUDED IN GENUS *Onobrychis* USING GENOMIC AND PHENOMIC APPROACHES

KÜBRA ERKOÇ

Ph.D.

ADANA 2025

TÜRKİYE CUMHURİYETİ
ADANA ALPARSLAN TÜRKEŞ SCIENCE AND TECHNOLOGY UNIVERSITY

GRADUATE SCHOOL
BIOENGINEERING DEPARTMENT

REEVALUATION OF TAXA INCLUDED IN GENUS *Onobrychis* USING GENOMIC AND PHENOMIC APPROACHES

KÜBRA ERKOÇ

Ph.D.

THESIS ADVISOR
Prof. Dr. MUHAMMET ŞAKİROĞLU

SECOND ADVISOR
Dr. BRANDON SCHLAUTMAN

ADANA, 2025

DECLARATION OF CONFORMITY

In this thesis study, which was prepared following the thesis writing rules of Adana Alparslan Türkeş Science and Technology University Institute of Graduate School, I declare that I provide all the information, documents, evaluations and results in accordance with scientific ethics and moral codes without resorting to any means or assistance that would be contrary to scientific ethics and traditions. I also declare that I refer to all of the articles I used in this study with appropriate references and accept all moral and legal consequences if a situation is found contrary to my statement regarding my work.

27/10/2025

[Signature]

Kübra ERKOÇ

ABSTRACT

REEVALUATION OF TAXA INCLUDED IN GENUS *Onobrychis* USING GENOMIC AND PHENOMIC APPROACHES

Kübra ERKOÇ

Ph.D., Department of Bioengineering

Supervisor: Prof. Dr. Muhammet ŞAKİROĞLU

Co-Supervisor: Dr. Brandon SCHLAUTMAN

October 2025, 111 pages

The genus *Onobrychis* is one of the most taxonomically complex genera of the Fabaceae family. Uncertainties in the number of species and high polymorphism in morphological characters make it necessary to re-evaluate this genus from a systematically. In this study, a multi-faceted approach was adopted to elucidate the taxonomic relationships of the genus *Onobrychis*. Firstly, the morphological diversity of the genus was measured manually and by digital image analysis on leaflet and fruit characters. Twelve traits were determined for the leaflets and 15 for the fruits. With the phenotypic data obtained, Principal Component Analyses (PCA) and cluster analyzes were performed and dendrograms were drawn. In addition, the assessment of species with common measurements from the phenotypic data obtained were classified using machine learning algorithms with an aim to reveal the distinctions between species more reliably. For molecular level evaluation, genetic diversity and taxonomic relationships were analyzed using Inter-Primer Binding Site (iPBS) markers. PCA analysis was performed by calculating the genetic distance from the data obtained by iPBS analysis and it was aimed to determine the taxonomic relationship by drawing a dendrogram. The iPBS marker results indicated that there was high polymorphism among the taxa included the genus and marker profiling could not clearly distinguish subgenera and sections. Leaf morphological analyses revealed that *Onobrychis* species largely reflect taxonomic distinctions at the subgenus level, but there is high morphological diversity in Sisyrosema subspecies. PCA and clustering analyses based on fruit morphology largely reflected subgenus-level distinctions in the genus *Onobrychis*, but higher morphological diversity was also detected, especially in the subgenus *Sisyrosema*.

Machine learning analyses have shown that fruit and leaf morphological data have high differenciation capacity among *Onobrychis* species. Classifications made with Random Forest (RF), Support Vector Machine (SVM), and K-Nearest Neighbors (KNN) algorithms achieved accuracy rates of up to 97%, demonstrating that classical morphological data carry a strong taxonomic signal when supported by modern computational methods. In conclusion, phenotypic analyses based on leaf and fruit morphology and molecular data obtained using iPBS markers revealed that taxonomic relationships can be largely reflected in the *genus Onobrychis*. Machine learning analyses have shown that morphological data have high differenciation power among species and that classical systematic approaches could be more reliable when supported by modern computational methods. However, it is thought that chloroplast and mitochondrial genome analyses, as well as further phenomic analyses could be employed to accurately reveal the boundaries at the subgenus.

Keywords: Genus *Onobrychis*, iPBS, Machine learning, Genetic diversity

ÖZET

GENOMİK VE FENOMİK YAKLAŞIMLAR KULLANILARAK *Onobrychis* CİNSİNE DAHİL EDİLEN TAKSONLARIN YENİDEN DEĞERLENDİRİLMESİ

Kübra ERKOÇ

Doktora, Biyomühendislik Anabilim Dalı

Danışman: Prof. Dr. Muhammet ŞAKİROĞLU

II. Danışman: Dr. Brandon SCHLAUTMAN

Ekim 2025, 111 sayfa

Onobrychis cinsi, Fabaceae familyasının taksonomik açıdan en karmaşık cinslerinden biridir. Tür sayısındaki belirsizlikler ve morfolojik karakterlerdeki yüksek polimorfizm, bu cinsin sistematik açıdan yeniden değerlendirilmesini gerekli kılmaktadır. Bu çalışmada, *Onobrychis* cinsinin taksonomik ilişkilerini aydınlatmak amacıyla çok yönlü bir yaklaşım benimsenmiştir. İlk olarak, cinsin morfolojik çeşitliliği yaprakçık ve meyve karakterleri üzerinden dijital görüntü analizi ile ve manuel olarak ölçümleri yapılmıştır. Yaprakçıklar için on iki özellik, meyveler için ise 15 özellik belirlenmiştir. Elde edilen fenotipik veriler ile Temel Bileşen Analizleri (Principal Component Analyses,PCA), kümeleme analizleri yapılarak dendrogramlar oluşturulmuştur. Ayrıca elde edilen fenotipik verilerden ortak ölçümlere sahip türlerin ölçümleri makine öğrenmesi algoritmaları kullanılarak sınıflandırılmış ve türler arasındaki ayrımların daha güvenilir biçimde ortaya konması hedeflenmiştir. Moleküler düzeyde değerlendirme için ise Primer Bağlanma Bölgeleri Arası (Inter-Primer Binding Site, iPBS) markörleri kullanılarak genetik çeşitlilik ve taksonomik ilişkiler analiz edilmiştir. iPBS markör analizi ile elde edilen verilerden genetik mesefa hesaplanarak PCA analizi yapılmış ve dendrogram oluşturularak taksonomik ilişkinin belirlenmesi amaçlanmıştır. iPBS markörüyle yapılan analizler sonucunda cinsin türleri arasında yüksek polimorfizmin olduğu ve genetik çeşitliliğin alt cinsleri ve seksiyonları net bir şekilde ayıramadığı belirlenmiştir. Yaprak morfolojik analizi sonucunda *Onobrychis* türünde alt cins düzeyinde taksonomik ayrımları büyük ölçüde yansıttığı ancak *Sisyrosema* alt cinsinde yüksek morfolojik çeşitlilik olduğu gözlemlenmiştir. Meyve morfolojisine dayalı PCA ve kümeleme analizleri, *Onobrychis* cinsinde alt cins düzeyindeki ayrımları büyük ölçüde yansıtmış, ancak yaprak morfilojileri ile

uyumlu şekilde özellikle *Sisyrosema* alt cinsinde daha yüksek morfolojik çeşitlilik saptanmıştır. Makine öğrenmesi analizleri, meyve ve yaprak morfolojik verilerinin *Onobrychis* türleri arasında yüksek ayırt edici güce sahip olduğunu göstermiştir. Rastgele Orman (Random Forest, RF), Destek Vektör Makinesi (Support Vector Machine, SVM) ve K-En Yakın Komşu (K-Nearest Neighbor, KNN) algoritmalarıyla yapılan sınıflandırmalarda %97'ye varan doğruluk oranlarına ulaşılmıştır, bu da klasik morfolojik verilerin modern hesaplamalı yöntemlerle desteklendiğinde güçlü bir taksonomik sinyal taşıdığını ortaya koymuştur. Sonuç olarak, yaprak ve meyve morfolojisine dayalı fenotipik analizler ile iPBS markörleri kullanılarak elde edilen moleküler veriler, *Onobrychis* cinsinde taksonomik ilişkilerin büyük ölçüde yansıtılabildiğini ortaya koymuştur. Makine öğrenmesi analizleri ise morfolojik verilerin türler arasında yüksek ayırt edici güce sahip olduğunu ve klasik sistematik yaklaşımların modern hesaplamalı yöntemlerle desteklendiğinde daha güvenilir hale geldiğini göstermiştir. Bununla birlikte, alt cins ve seksiyon düzeyindeki sınırların daha net biçimde ortaya konabilmesi için kloroplast ve mitokondriyal genom analizleri ile daha ileri fenomik analizlerle desteklenmesi gerektiği düşünülmektedir.

Anahtar Kelimeler: *Onobrychis* cinsi, iPBS, Makine öğrenmesi, Genetik çeşitlilik

Dedication

As a child of a family that experienced the earthquake in Hatay, I dedicate this thesis to all those who experienced this disaster in order to keep the memory of those we lost in the February 6 earthquake alive.

ACKNOWLEDGEMENTS

I would to extend my sincere gratute to my esteemed advisor Prof. Dr. Muhammet Şakiroğlu, who guided me with his knowledge, experience and mentorship at every stage of my diseration work, always shared his academic expertise and vision with me, and supported me with patience and understanding when I faced difficulties. I would like to also thank to my second advisor, Dr. Brandon Schlautman, who supported me and guided me during my studies in the United States and afterwards.

I would like to thank my jury members Prof. Dr. Zeynel Cebeci and Asst. Prof. Özgür Cem Erkin, who illuminated my path with their knowledge and experience throughout my diseration process, and my thesis defense jury members Assoc. Prof. Dr. Esra Göv and Asst. Prof. Dr. Orhan Uluçay.

I would like to thank to Prof. Dr. Allison Miller for helping me to carry out research work at the Missouri Botanical Garden, Emily J. Warschefsky, Ph.D. and to Dr. Jordan K. Teisher, Herbarium Curator and Director of the Missouri Botanical Garden for their support.

I also would like to thank Assoc. Prof. Dr. Aras Türkoğlu, who helped me conduct my research study in Konya.

A special thank to my dear parents and siblings Özlem, Aynur and Serdar, who have always been a great support to me and gave me strength, as well as to my nephews who add happiness and energy to my life.

I would like to thank my cohorts Mahmut Aydın for his support in the analysis, Ebrar Karabulut Uzun for her assistance in my studies in the United States and Turkey, and Özlem Gül Yastı for her contributions during my diseration work in Turkey. I would like to thank to the Land Instiute Staff and Missouri Botanical Garden Staff for their support during my visit.

This thesis was supported by Scientific and Technological Research Council of Turkey (TUBİTAK) project 120C136 and also within the scope of the Council of Higher Education (YÖK) 100/2000 Doctoral Scholarship Program in the field of ‘Digital Technologies in Agriculture and Livestock (Agricultural Technology, Artificial Intelligence in Agriculture, Smart Agriculture Applications). I would like to thank to TUBİTAK and YÖK for the support.

TABLE OF CONTENTS

DECLARATION OF CONFORMITY i
ABSTRACT ii
ÖZET iv
Dedication vi
TABLE OF CONTENTS viii
LIST OF FIGURES x
LIST OF TABLES xii
LIST OF ABBREVIATIONS xiii
LIST OF SYMBOLS xv
1. INTRODUCTION 1
2. THEORETICAL FOUNDATIONS AND LITERATURE REVIEW 5
2.1. Genus *Onobrychis* 5
2.2. Economic Importance of Member of *Onobrychis* and Sainfoin 10
2.2.1. Growth Conditions and Phenology 11
2.2.2. Morphology 12
2.2.3. Cytogenetics 14
2.2.4. Recent Additions of Species 16
2.2.5. Use of Morphological Traits in the Taxonomy of Genus 17
2.2.6. Use of Cytogenetic Tools in the Taxonomy 18
2.2.7. Use of Moleculer Tools in the Taxonomy 19
2.2.8. Use of Genomic Tools in the Taxonomic Classification 20
2.2.9. Inter-Primer Binding Site (iPBS) Markers 22
2.3. Machine Learning 24
3. MATERIALS AND METHODS 27
3.1. Species Taxonomically Evaluated in the Study 27
3.2. Plant Materials and DNA Analyses-iPBS Retrotransposon Markers 32
3.2.1. Plant Materials 32
3.2.2. DNA Extraction 34
3.2.2.1. DNA Extraction of Fresh Plants 34
3.2.2.2. DNA Extraction of Herbarium Specimen 35
3.2.3. Inter-Primer Binding Site (iPBS) Retrotranposon Markers Genotyping 35

3.2.3.1. iPBS-PCR Amplification 38
3.2.3.2. iPBS Data Analysis 39
3.3. Population Genetic Structure Analysis 39
3.4. Morphological Characterization and Phenotyping 40
3.4.1. Leaf Image Analyses 40
3.4.2. Fruit Image Analyses 44
3.4.3. Analysis of Leaf and Fruit Morphological Data 52
3.5. Machine Learning Analyses 53
3.5.1. Data Set 53
3.5.2. Preprocessing and Feature Selection 54
3.5.3. Machine Learning Algorithms 55
3.5.4. Model Evaluation 56
4. RESULTS AND DISCUSSION 57
4.1. iPBS Analyses 57
4.1.1. Genetic Structure 57
4.2. Morphological Characterization and Phenotyping 65
4.2.1. Efficiency of Different Traits in Sperately of Subgenus 65
4.2.1.1. Leaf Traits 65
4.2.1.2. Fruit Traits 82
4.3. Machine Learning 96
5. CONCLUSION 2
6. RECOMMENDATIONS 3
REFERENCES 4

LIST OF FIGURES

Figure 3.1. Different anneling temperatures a: iPBS 2415, b: iPBS 2389, c: iPBS 2222, d: iPBS 2401, e: iPBS 2394, f: iPBS 2395, g: iPBS 2383, h: iPBS 2228, i: iPBS 2398, j: iPBS 2491, k: iPBS 2392, l:iPBS 2393 [104]. 37
Figure 3.2. Measurement of morphological traits of leaves belonging to *Onobrychis* species using ImageJ. 44
Figure 3.3. Measurement of morphological traits of fruits belonging to *Onobrychis* species using ImageJ. 52
Figure 3.4. Correlation matrix (Pearson correlation analysis) created to examine the relationships between morphological traits. 54
Figure 3.5. Feature importance values calculated with the Random Forest algorithm. 55
Figure 4.1. Genetic structure plot obtained by Bayesian clustering analysis: STRUCTURE software results for 134 *Onobrychis* samples analyzed using 284 iPBS bands. Each individual is represented as a vertical bar, which is divided into different colors to show the probability (with % probability) that the individual belongs to different genetic clusters. (A) K = 2, (B) K = 3, (C) K=5, (D) K=6, (E) K=7, (F) K=9. 59
Figure 4.2. Two methods were used to determine the most appropriate value of K: (a) the ΔK statistic developed by [106] and (b) the ad hoc procedure proposed by [105]. 60
Figure 4.3. Distribution of 51 *Onobrychis* species and 6 different species of *Onobrychis* of unknown species based on the first two PCA obtained from the analysis of 284 iPBS bands. 61
Figure 4.4. Subgenus-based UPGMA dendrogram constructed using genetic distance data based on iPBS markers (blue: *Onobryhcis* subgenus, red: *Sisyrosema* subgenus). 63
Figure 4.5. Section-based UPGMA dendrogram constructed using genetic distance data based on iPBS markers (grey: *Afghanicae*, lilac: *Anthyllium*, yellow: *Heliobrychis*, blue: *Hymenobrychis*, purple: *Insignes*, red: *Onobrychis*, green: *Dendrobrychis*, pink: *Lophobrychis*, orange: *Hemicyclobrychis*, turquoise: *Laxiflora*). 64
Figure 4.6. Leaflet shapes. 66
Figure 4.7. Pairwise scatter matrix among leaf morphological traits. Lower diagonal parts show scatter plots between pairs of variables; diagonal lines show its variable distribution characteristics. Upper diagonal panels show Pearson distribution coefficients (r) and significance levels ($p < 0.05$ *, $p < 0.01$ **, $p < 0.001$ ***). 67
Figure 4.8. PCA analysis based on leaf morphological traits of *Onobrychis* species by subgenus. 69
Figure 4.9. PCA analysis based on leaf morphological traits of *Onobrychis* species by sections. 69
Figure 4.10. Hierarchical clustering dendrogram of leaf morphological traits from *Onobrychis* species colored by subgenus. 71
Figure 4.11. Hierarchical clustering dendrogram of leaf morphological traits from *Onobrychis* species colored by sections. 72
Figure 4.12. Determination of the optimum number of taxa clusters based on leaf morphological analysis using Gap (A), Elbow (B) and Silhouette (C) methods. 73

Figure 4.13. Correlation matrix between variables of fruit morphological traits. Shows Pearson correlation coefficients and significance levels ($p < 0.05$ *, $p < 0.01$ **, $p < 0.001$ ***)..... 83
Figure 4.14. PCA analysis based on mean fruit morphological traits of *Onobrychis* species by subgenus. 84
Figure 4.15. PCA analysis based on mean fruit morphological traits of *Onobrychis* species by section..... 85
Figure 4.16. Determination of the optimum number of taxa clusters based on fruit morphological analysis using Gap (A), Elbow (B) and Silhouette (C) methods. 86
Figure 4.17. Dendrogram according to fruit traits colored by subgenus..... 87
Figure 4.18. Dendrogram according to fruit traits colored by sections..... 88
Figure 4.19. Analysis results of ML algorithms using all morphological measurements of leaves and fruits..... 97
Figure 4.20. Analysis results of ML algorithms using selected morphological measurements of leaves and fruits. 98

LIST OF TABLES

Table 2.1. Taxonomic classification of the genus *Onobrychis* [8]. 8
Table 3.1. Compilation of the taxonomic information of *Onobrychis* species used in the current study. 27
Table 3.2. Species and information used in genetic analyses. 32
Table 3.3. Sequence and annealing temperature (°C) of iPBS markers used for genetic characterization of genus *Onobrychis* [104]. 38
Table 3.4. Traits measured in leaflets of taxa included in the genus *Onobrychis*. 42
Table 3.5. Representative fruit images of the *Onobrychis* species examined in the study. 45
Table 3.6. Traits measured in fruits of taxa included in the genus *Onobrychis*. 50
Table.4.1. Percentage of variation explained by the first 3 axes in Principal Component Analysis (PCA) 62
Table 4.2. Leaflet shape classification. 66
Table 4.3. Morphological traits of *Onobrychis* leaflets in the study. 75
Table 4.4. Morphological traits of *Onobrychis* fruits in the study. 90
Table 4.5. Comparison of results of all ML algorithms. 0

LIST OF ABBREVIATIONS

2-D	: Two dimensional
ADF	: Acid detergent fiber
AFLP	: Amplified Fragment Length Polymorphism
AMOVA	: Analysis of molecular variance
AR	: Aspect Ratio
AT	: Anneling temperature
bp	: Base pair
CAPS	: Cleaved Amplified Polymorphic Sequence
Circ.	: Circularity
Ctab	: Cetyltrimethylammonium Bromide
DNA	: Deoxyribonucleic Acid
DC.	: de Candolle
Gbs	: Genotyping-By-Sequencing
hcut	: Hierarchical clustering
Hrm	: High-Resolution Melting
iPBS	: Inter-Primer Binding Site
ISSR	: Inter Simple Sequence Repeats
I	: Shannon Index
KNN	: K-Nearest Neighbor
M	: marker
Mill.	: Miller
ML	: Machine Learning
NDF	: Neutral Detergent Fiber
O.	: Onobryhcis
nSSR	: nuclear Simple Sequence Repeat

PC	: Principal Component
PCA	: Principal Component Analysis
PCoA	: Principal Coordinates Analysis
PCR	: Polymerase Chain Reaction
Perim.	: Perimeter
pg	: picogram
PIC	: Phylogenetically Independent Contrasts
RAPD	: Random Amplified Polymorphic DNA
rDNA	: ribosomal DNA
RF	: Random Forest
Rech.	: Rechinger
rpm	: revolutions per minute
SCAR	: Sequence Characterized Amplified Region
Sect.	: Section
SNP	: Single Nucleotide Polymorphism
SSR	: Simple Sequence Repeats
SVM	: Support Vector Machines
Širj.	: Širjaev
UPGMA	: Unweighted Pair Group Method with Arithmetic Mean
var.	: varyete

LIST OF SYMBOLS

%	: Percent
≈	: Approximately
°C	: Celsius degree
cm	: Centimeter
h	: Hour
L	: Liter
m	: meter
mL	: Milliliter
mm	: Millimeter
mm^2	: Square millimeter
R	: Pearson correlation coefficient
µL	: Microliter
µm	: Micrometer

1. INTRODUCTION

Among the world's flowering plant families, the Fabaceae (Leguminosae) constitute the third largest plant group, with approximately 19,400 species and 730 genera. Cultivation of the Fabaceae family, which plays a crucial role in providing the energy, protein, and minerals needed by humans and animals, dates back to ancient times [1]. Therefore, the Fabaceae family is of great economic and medicinal importance. The family consists of the subfamilies Papilionoideae, Mimosoideae, and Caesalpinioideae, which are widespread in temperate and tropical climates [2].

Genus *Onobrychis* is one of the prominent genera of the Fabaceae (Leguminosae). The first information about the genus *Onobrychis* dates back to Linneus's (1753) inclusion of species of this genus within the genus *Hedysarum* in his *Species Plantarum* [3]. Later, Miller considered *Onobrychis* species as a separate group and used the genus name in this way for the first time [4]. Taxonomically, *Onobrychis* Miller, it is a genus belonging to the Angiospermae division, Dicotyledonae class, Rosales order, and Papilionoideae subfamily of the Fabaceae (Leguminosae) family in the plant kingdom. *Onobrychis* is included in the Hedysareae tribe, which is widespread in the temperate regions of North America, Europe, and the Middle East [5].

Onobrychis is considered one of the most difficult genera to identify and characterize within the Fabaceae family [6]. However, the lack of species identification keys for the genus also poses a significant problem, leading to conflicting results in different studies [4]. The boundaries and species content of this taxonomically critical genus have been addressed differently by different researchers and are not fully clarified [7]. *Onobrychis* species exhibit high polymorphism in their morphological characters, making it difficult to determine the boundaries of some species. Therefore, numerous taxonomic studies have been conducted based on morphological and cytological characteristics and phytochemical compounds [8].
Širjaev (1925) developed a classification based on floral characteristics, and this classification is still accepted by many researchers today. In this system, the genus is divided into two subgenera: *Onobrychis* and *Sisyrosema*. The *Onobrychis* subgenus includes four sections:

Dendrobrychis, *Lophobrychis*, *Onobrychis*, and *Hemicyclobrychis*, while the *Sisyrosema* subgenus consists of four sections: *Anthyllium*, *Afghanicae*, *Heliobrychis*, and *Hymenobrychis* [9]. Later, Rechinger (1984) reorganized these two subgenera and classified them into four and five sections, respectively [10].

The genus *Onobrychis* includes both perennial and annual members. However, the number of species varies significantly among different sources. The center of *Onobrychis* diversity is the Western Asia and Eastern Mediterranean regions, and approximately 130 species of this genus are distributed mostly in temperate regions [11]. According to the classification made by Širjaev (1925), the genus was accepted to contain 126 species [9]. Today, it is stated that the genus *Onobrychis* is represented by approximately 162 species worldwide, and 52 species and 60 taxa are found in Turkey [4]. Additionly, there are *Onobrychis* species in Russia, Afghanistan, Turkmenistan, Tajikistan, Iran and Iraq [12].

In plant systematics, molecular markers play a critical role in elucidating phylogenetic relationships among species and revealing genetic diversity. Molecular markers provide more reliable and objective results at the genetic level [13]. iPBS (Inter Primer Binding Site) markers are dominant markers that reveal polymorphisms in DNA using universal primers that bind to evolutionarily conserved regions of retrotransposons. Their PCR-based nature, their lack of prior knowledge of sequence, their high polymorphism, their reliable reproducibility, and their affordability make iPBS markers particularly advantageous for genetic diversity studies [14]. The use of iPBS markers in the genus *Onobrychis* allows the identification of similarities and differences between species at the nuclear DNA level [14]. These markers can contribute to the resolution of taxonomic uncertainties in the genus and provide important information in establishing species relationships.

Morphological traits have been used as fundamental criteria for plant identification and classification for many years. Fruit and leaf morphology are among the most widely used characters for distinguishing species. These characteristics are critical for both taxonomic studies and for identifying variation among ecotypes [15], [16]. With developing technologies, digital image processing methods have become widely used in morphological analyses. Thanks

to open-source software such as ImageJ [17], numerous morphometric measurements, such as length, width, area, perimeter, circularity, and aspect ratio, of leaves and fruits can be estimated or measured easily, quickly, and with high accuracy. This method offers more objective, reproducible, and scalable results compared to traditional measurement approaches. Furthermore, the multidimensional data obtained through digital measurements allows the development of new criteria for species identification. Thus, modern classification approaches based on digital criteria are emerging, beyond classical identification keys, and are contributing to a more detailed understanding of morphological diversity.

In recent years, machine learning has become an important tool in plant systematics and biodiversity studies. These methods facilitate the processing, classification, and pattern discovery of high-dimensional phenotypic and genetic datasets. Machine learning algorithms are widely used in the classification of phenotypic data. By combining different data sources, machine learning algorithms help to more confidently reveal relationships among species [15], [16], [18]. Applications of machine learning to the genus *Onobrychis* offer an innovative approach. Given the genus's morphological diversity and taxonomic uncertainties, classification of phenotypic data using machine learning contributes to species identification and allows for clearer delineation of species boundaries.

Based on data obtained from the literature review, the number and boundaries of species within the genus *Onobrychis* are unclear. Furthermore, the taxonomic classification within the genus is also uncertain. Based on this information, the primary objectives of this study are as follows:

i. To determine the genetic diversity of the *Onobrychis* genus and taxonomic relationships at nuclear DNA-level using inter-primer binding site (iPBS) markers.
ii. To reveal morphological variation among species by examining fruit and leaf morphological characteristics in detail and to develop a digital scale for species identification.
iii. To enable the classification of phenotypic data using machine learning methods.

This study adopted a multifaceted approach to better understand the taxonomic boundaries and diversity of species within the genus *Onobrychis*. First, genetic diversity and nuclear DNA relationships were examined using iPBS markers. Morphological diversity within the genus was then assessed using digital image analysis based on leaflet and fruit measurements. The resulting phenotypic data were then classified using machine learning algorithms to provide more reliable species distinctions. The combined application of these methods contributes to the resolution of taxonomic uncertainties within the genus *Onobrychis* and the development of more robust classification approaches.

2. THEORETICAL FOUNDATIONS AND LITERATURE REVIEW

2.1. Genus *Onobrychis*

Leguminosea or Fabaceae, which is among the flowering plant families in the world, is the third largest plant group with 19.400 species and about 730 genera are included in this group. The cultivation of the *Fabacea* family, which has an important place in providing the energy, protein and minerals needed by humans and animals, dates back to ancient times [1]. Therefore Fabaceae family is an economically and medically important plant family [19]. Commonly known as the pea family, Fabaceae are generally characterized by compound leaves and the production of fruits known as legumes [20]. The Fabaceae family consists of subfamilies Papilionoideae, Mimosoideae and Caesalpinioideae found in temperate and tropical climates. Fabaceae are characterized with trees, shrubs, vines, or herbs, with stipulate, often compound leaves and typically pentamerous flowers, usually with a single, unicarpellous pistil with marginal placentation and the fruit a legume [2]. The leaves are usually compound, sometimes simple. Flowering is variable but typically bracteate. Flowers are usually bisexual, sometimes unisexual, actinomorphic, or zygomorphic, petiolate or sessile, hypogynous or periginous. The fruit is a legume, sometimes divided into non-separating, winged, drupelike or transverse segments [2]. Family members show great differences in terms of these characteristic features [4]. The roots of members of the Fabaceae family have a symbiotic relationship with nitrogen-fixing bacteria that induce the formation of root nodules. Beside have an important place in agriculture with their ability to make nitrogen in the air usable by plants, thanks to the symbiosis they perform with the help of Rhizobium bacteria [1].

The first information about the genus *Onobrychis* started when Linneus (1753) included the species belonging to the genus *Onobrychis* in the genus *Hedysarum* in his work called *Species Plantarum* [3]. Miller was the first person to name this genus by collecting *Onobrychis* species under the genus *Onobrychis* [4]. Then, many species and studies were carried out in the genus *Onobrychis* by different researchers [3]. Taxonomically, *Onobrychis* Miller is a genus in the Dicoytledonae (dicotyledons) class of Angiospermae (angiosperms) in the plant kingdom, and

in the Papilionoidae subfamily of the Fabaceae (Leguminosae, Legume) family of the order Rosales. *Onobrychis* is a genus belonging to the family Fabaceae and the tribe *Hedysareae*, common in temperate regions of North America, Europe, and the Middle East [5]. *Onobrychis* is one of the most difficult genera in the Fabaceae family to diagnose and characterize [6]. In the literature, studies have been carried out on a small number of species, using a limited number of characters for the identification and separation of parts of the genus *Onobrychis*. Generally morphological and cytogenetic characters such as annual or perennial species, chromosome number and size, flower color, number of ovules, pod color, and flower and pod shape have been focused on for *Onobrychis* genus [6]. The *Onobrychis* family is a genus that has problems in collecting material. This difficulty is due to the fact that the flower color fades and changes during the herbarium formation of the genus, and the flowers and fruits fall off easily after drying [3]. In addition, the inadequacy of the species identification keys of the genus *Onobrychis* appears as another problem and causes differences in species studies [4].

Taxonomically, *Onobrychis* is a critical genus. The taxonomy of *Onobrychis* is not completely determined, with contradictory schemes for its internal structure and species content [7]. This genus has high polymorphism in its morphological characters, so the boundaries of some taxonomic species are very clear. For this reason, many studies performed regarding its taxonomy based on morphological, cytological features and the compounds it contains. Researchers have suggested different classification according to fruit morphology, chromosome number and flower color. Furthermore several biosystematics studies ranging from non-molecular data to DNA sequences have been carried out on *Onobrychis* genus [8]. Širjaev (1925) presented a classification for *Onobrychis* genus based on floral characteristics and the study is still accepted by many researchers. In their study, the genus consists of two subgenera, *Onobrychis* and *Sisyrosema*. Subgenus *Onobrychis* includes four sections *Dendrobrychis* DC, *Lophobrychis*, *Onobrychis*, and *Hemicyclobrychis* while the subgenus *Sisyrosema* comprises four sections *Anthyllium*, *Afghanicae*, *Heliobrychis*, *Hymenobrychis* (Table 2.1) [9]. Rechinger (1984) reclassified the two subgenera into four and five sections, namely *Onobrychis*, *Dendrobrychis* DC., *Lophobrychis* Hand.-Mazz., *Laxiflorae* (Širj.) Rech.f., *Anthyllium* Nábělek, *Afghanicae* Širj., *Heliobrychis* Bunge ex Boiss., *Hymenobrychis* DC. and *Insignes* (Širj.) Rech.f., respectively [10]. Širjaev (1925) divided the genus into two

subgenera, *Onobrychis* as *EuOnobrychis* (Bunge ex Boiss.) Širj. and *Sisyrosema* (Bunge ex Boiss.) Širj. each with four sections based on floral characteristics. However, the sectional positions of *O. splendida* Rech.f. & Podlech and *O. freitagii* Rech.f. (this species was not included in the present study) have remained hitherto uncertain [8].

A karyological study found that unlike *Onobrychis* subgenus, *Sisyrosema* subgenus is monophyletic [21]. About with 30 species *Heliobrychis* is one of the richest parts of *Onobrychis* in terms of species, which includes perennial species found in Iran, Turkiye and Caucasia [22]. 21 *Onobrychis* species found in Iran are included in the *Sisyrosema* subgenus class [23]. The flowers of *Onobrychis* subgenus are commonly pink, rarely cream or white. The flowers of the *Sisyrosema*e subgenus are commonly yellow, rarely cream or white [1]. *Onobrychis* genus has included perennial or annual members. The exact number of *Onobrychis* species is not known. The center of *Onobrychis* diversity is western Asia and the eastern Mediterranean regions, with about 130 species of this species mostly found in temperate regions [11]. According to the classification of *Onobrychis* developed by Širjaev (1925), it is accepted that there were originally 126 species of this genus [9]. The *Onobrychis* genus is represented throughout the world with 162 species and in Turkiye there are 52 species 60 taxon. 27 of these species found in Turkiye are endemic. *Onobrychis* is especially concentrated and diversified in the Anatolian-Iran-Caucasus triangle. In these regions, there are 53 species in Iran, 52 species in Turkiye and 39 species in the Caucasus [4]. There are also 62 species of *Onobrychis* in Russia, 17 species in Afghanistan, nine in Turkmenistan, nine in Tajikistan and 14 species in Iraq [12]. Approximately 173 species are existed by the International Legume Database and Information Service [24]. Yildiz et al. (1999) in study conducted according to fruit morphology, it was determined that there are 170 species, eight sections and two subgenera of *Onobrychis* genus [6]. Another study reported that *Onobrychis* Mill. genus *Hedysareae* (Fabaceae) is the second largest genus of the tribe and has more than 173 annual or perennial species [25], [26]. In another study, 342 perennial and annual species of *Onobrychis* were indicated [27]. In the Healthy Hay project, which is the European Union Legum Forage Crops Research Project, carried out in 2012, 362 different sainfoin species were identified across Europe [28].

Table 2.1. Taxonomic classification of the genus *Onobrychis* [8].

Širjaev (1925, 1926)	**Boisser (1872)**	**Townsend (1974)**	**Rechinger (1984)**	**Aktoklu (1995)**	**Present study**
Subgenus *EuOnobrychis* (Bunge ex Boiss.) Širj.	*EuOnobrychis* Bunge	Subgenus *Onobrychis*	Subgenus *Onobrychis*	Subgenus *Onobrychis*	Subgenus *Onobrychis*
Sect. *Eubrychis* DC. (including subsects. Macropterae Hand.-Mazz., Macrosemiae Hand.-Mazz., Brachysemiae Hand.-Mazz., albae Hand.-Mazz., Hispanicae Širj. and Vulgatae Hand.-Mazz.)		Sect. *Onobrychis*	Sect. *Onobrychis*	Sect. *Onobrychis*	Sect. *Onobrychis*
Sect. *Dendrobrychis* DC. (including sers. Dielsianae Širj. and Litvinovianae Širj.)		Sect. *Dendrobrychis* DC.	Sect. *Dendrobrychis* DC.	Sect. *Dendrobrychis* DC.	
Sect. *Lophobrychis* Hand.-Mazz. (including subsects. Occidentales Širj. and Orientales Širj.)		Sect. *Lophobrychis* Hand.-Mazz.	Sect. *Lophobrychis* Hand.-Mazz.	Sect. *Lophobrychis* Hand.-Mazz.	
Sect. *Hemicyclobrychis* Širj.					Sect. *Hemicyclobrychis* Širj.
			Sect. Laxiflorae (Širj.) Rech.f.		
Subgenus *Sisyrosema* (Bunge ex Boiss.) Širj.	*Sisyrosema* Bunge	Subgenus *Sisyrosema* (Bunge ex Boiss.) Širj.	Subgenus *Sisyrosema* (Bunge) Grossha	Subgenus *Sisyrosema* (Bunge) Grossha	Subgenus *Sisyrosema* (Bunge ex Boiss.) Širj.

Sect. *Afghanicae* Širj.			Sect. *Afghanicae* Širj.	Sect. *Heliobrychis*	Sect. *Afghanicae* Širj.
Sect. *Anthyllium* Nábělek (including subsects. Fedtschenkoanae Širj., Lipskyanae Širj., Mirae Širj. and Nábělekianae Širj.)		Sect. *Anthyllium* Nábělek	Sect. *Anthyllium* Nábělek	Sect. *Hymenobrychis*	
					Sect. Lipskyanae (Širj.) Amirah. & Kaz. Osaloo
Sect. *Heliobrychis* Bunge (including subsects. Szovitsianae Širj., Boissierianae Širj. and Persicae Širj.)		Sect. *Heliobrychis* (Bunge ex Boiss.) Širj.	Sect. *Heliobrychis* Bunge ex Boiss.		Sect. *Heliobrychis* (Bunge ex Boiss.) Širj.
Sect. *Hymenobrychis* DC. (including subsects. *Insignes* Širj., Modestae Širj., Pulcherrimae Širj. and Laxiflorae Širj.)		Sect. *Hymenobrychis* DC.	Sect. *Hymenobrychis* DC.		Sect. *Hymenobrychis* DC.
			Sect. *Insignes* (Širj.) Rech.f.		Sect. *Insignes* (Širj.) Rech.f.
					Sect. Laxiflorae (Širj.) Rech.f.
					Sect. Litvinovianae (Širj.) Amirah. & Kaz. Osaloo

2.2. Economic Importance of Member of *Onobrychis* and Sainfoin

Genus *Onobrychis* has a large number of species. Agriculturally important *Onobrychis* species are *O. viciifolia* (Common Sainfoin), *O. arenaria* (Anatolian Sainfoin), *O. transcaucasia* (Caucasian Sainfoin), *O. montana* D.C., *O. sativa var. Persica* (Širjaev pro var.) [29]. Among these species, *O. viciifolia*, which is called common sainfoin and is widely cultivated around the world [30]. *Onobrychis* is an important forage legume crop with perennial and annual growing habbit [5].

The history of sainfoin as a forage extends for centuries in Europe and the Middle East, with reports of its use as early as the 1500s. The initial reported scientific trials about sainfoin in the UK were conducted as early as 1668 [31]. Sainfoin was widely grown in Euroasia before the widespread use of commercial fertilizers [5] and production spread to North America in the 19th and 20th centuries. It is a very tasty and nutritious legume fodder plant that produces hay and pasture [32]. Studies show that sainfoin has anthelmintic properties, methane control potential from ruminants, and protein protection ability against premature deterioration in the rumen when used as a fodder in the diet of ruminants owing to the tannin compound it contains [33]. The plant is resistant to drought, cold and low nutrient availability but is not resistant high soil salinity and high water tables [32]. Sainfoin is an important component in sustainable agriculture due to the ability to convert atmospheric nitrogen into bioavailable compounds, which reduce the dependency to regular input of inorganic nitrogen [7].

The genus *Onobrychis*, which grows naturally in Anatolia, consists of two subgenera, *Onobrychis* and *Sisyrosema*. *Onobrychis* subgenus divided into 3 different sections as *Dendobrychis*, *Laphobrychis* and *Onobrychis* [4]**.** Approximately 70 *Onobrychis* species naturally grow in Turkiye as the country is situated on the natural distribution range of the genus and is one of the centers of genetic diversity [34]. Sainfoin is also known as görüngen or korunga in [35]. Sainfoin is defined as the king of forage crops in Turkiye due to its superior agricultural characteristics, high yield and quality feed that can adapt to different ecological regions of Turkiye [36]. Sainfoin has a large cultivation area, especially in the Eastern provinces

of our country, and also found in many places in Central Anatolia [35]. Unlike alfalfa and some other legumes, sainfoin does not cause bloat in animals owing to the concentrated tannins [37].

Since 1990, the production of sainfoin has an increasing momentum in Turkiye. According to TÜİK data, sainfoin cultivation area has fluctuated in recent years. Sainfoin cultivation area increased from 957,590 hectares in 1990 to 1,792,534 hectares in 2020 and 1,814,737 hectares in 2021 [38]. However, it decreased in subsequent years, reaching 1,241,541 hectares in 2024. Similarly, sainfoin's green forage production was determined to be 1,781,789 tons in 2020 and 1,468,830 tons in 2024 [39]. Because of the high amount of nectar it contains, the sainfoin functions as a good bee pasture. It is seen that high yield and honey quality are obtained from beehives in areas where sainfoin is cultivated [35].

2.2.1. Growth Conditions and Phenology

The growth stage of the sainfoin is defined as follows: seedling emerge (Em), vegetative stage - no stem elongation (V1), vegetative stage - stem elongation (E1), early bud: 1 to 2 nodes with buds; no flowers or seed pods (B1), late bud: > 2 nodes with buds; no flowers or seed pods (B2), early flowering: one node with one open flower; no seed pods (F1), peak/late flowering: >= 2 nodes with open flower; no seed pods (F2), early seed pod: 1 to 3 nodes with green seed pods (P1), late seed pod: > 3 nodes with green seed pods (P2), ripe seed pod: nodes with mostly brown mature seed pods (P3), postharvest: after harvesting all seeds, fruit, or forage (PH1), dormant winter period (DO) [40].

Onobrychis species, which are highly resistant to drought and cold, can also be cultivated in barren and calcareous areas where other crop species cannot be cultivated as it is adapted to high temperature, rainfall, drought, salinity, diseases and pests [41], [42]. *Onobrychis* can be planted in spring in areas with harsh winters, and in autumn in regions with mild winters [1]. Sainfoin has many hollow stems which arise from basal buds to form a branched crown [7]. Pink, white, yellow or purple flowers of *Onobrychis* are highly attractive to pollinators. Sainfoin is an allagomous plant, and the female organ is longer than male organ. Sainfoin is generally cross-pollinated, mostly by honey bees (*Apis mellifera* L.) and leaf cutting bees (*Megachile rotundata* Fabricius.) [5]. In addition, bumble bees (*Bombus* spp.) are also frequent visitors to

sainfoin fields. *Bombus* spp. generally visit the sainfoin between morning and noon because there is not much wind, and flowers are ready for pollinators. Pollination success is relatively high in sainfoin. The duration between the opening of the first flower and the withering of the terminal flowers on a raceme is about 2-3 weeks. Sainfoin produces single-seeded indehiscent brown colored pods. Seed color varies from olive to brown or black. The seeds are kidney-shaped, with a smooth surface [43]. *Onobrychis* species have a deep taproot and in addition to this taproot, it has thick and many thin roots that come out sideways. The length of this pile root can reach 1-10 m depths, depending on the soil and climatic conditions where the sainfoin grows. Thanks to its deep roots, it aerates the soil and uses the nutrients and water from the depths [34]. There are less minerals such as calcium and sodium in sainfoin compared to other legumes. However, it has been determined that it is rich in phosphorus, magnesium and potassium [44], [45].

Fruit set depends upon the species and variety, but it can be correlated with the suitability and adaptability of its environment. The flowers open gradually over 24-h especially between night and sunrise [29], [46], [47].

2.2.2. Morphology

The genus *Onobrychis* consists of annual or perennial (most of the species are perennial) herbaceous species and a small number of spiny-radius species. The stem, which is ligneous or buttress in the lower parts, can be hairy or hairless. The stipules, which can be seen as free or compound, are usually membranous and have hairy edges. The leaflets are whole-edged, round to oblong in shape, and the apex ends with a hard tip. The leaflets are attached to the leaf axis at the base by a long stalk, and at the upper parts by a short stalk, and are rarely sessile. The inflorescence is axial and panicle. The calyx is bell-shaped, the teeth are unequal and usually lanceolate. Petals of flower can be yellow, pink, white, cream or mauve and have dark veins. Flag leaf has an elliptic structure and in some species the back is hairy. The paddles are generally shorter than the sepals, with auriculares and scapus. The boat is shorter than or equal to the flag leaf. The ovary has 1-3 ovules. Male organs are in diadelphous structure. The fruit is nearly round, does not open when dry and contains 1-2 seeds. It is soft hairy or hairless. Seeds are kidney or oblong shaped [4].

Onobrychis is closely related to *Hedysarum*, but the fruit always consists of only one segment and the ovary has 1-2 egg cell, and the crown of *Onobrychis* species is standard and shorter than the tip [48]. Hedge (1970) defined the characteristic features of sections of subgenera as follows: Characteristics of *Onobrychis* subgenera, *Dendobrychis* section: plant semi-bush, spiny, perennial [48]. Characteristics of *Onobrychis* subgenera, *Lophobrychis* section: annual herbaceous. *Onobrychis* subgenera, *Onobrychis* section, characteristic features are perennial herbs, fruit has no distinct junction and stem. Characteristics of section *Sisyrosema* subgenera, *Heliobrychis*: rarely annual, usually perennial herbs. Characteristics of the section *Sisyrosema* subgenera, *Hymenobrychis*: herbaceous perennials [48].

Sainfoin (*Onobrychis viciifolia*) axillary tillers possess inflorescences in an erect raceme, which contains approximately 70-80 flowers. The flowers have a pleasant scent. The sainfoin fruit is a semicircular, small flat pod with a single seed. The fruit pericarp surrounds the seed. The pericarps of the fruits are net-shaped, veined and form a protrusion on the semicircular edges [45]. The top of the fruit is veined in the form of a net and the semicircular edge is thorny in the form of a protrusion [34]. Fruit structure consists of bristly hairy or glabrous, rarely fine soft hairy, woolly or coarse hairy structure, two rows in the middle, honeycomb-like surface disc and margins of varying width [4]. Sainfoin is a deep-rooted perennial legume with erect or sub-erect hollow stems growing up to a height of 100 cm. It has a deep tap root system with few main branches and numerous fine lateral roots, which provide sites for root nodule development. Leaves are pinnately compounded with 5-14 pairs of leaflets and a terminal leaflet [43]. It can be approximately 100-120 cm in length, its surface is slightly hairy and there are 5-30 leaflets on its leaves. These, too, have a flat structure with a length of about 10-25 mm and a width of 3-8 mm [45]. Sainfoin has many hollow stems which arise from basal buds to form a branched crown [7]. Pink, white, yellow or purple flowers of sainfoin are highly attractive to pollinators [33].

Fruit morphology in the genus *Onobrychis* varies significantly among sections. According to Yildiz et al. (1999), in section *Dendrobrychis*, the fruit is single-segmented, sessile, and generally spineless, though spiny varieties are rarely observed. The fruit, lacking any differentiation between dorsal and ventral surfaces, exhibits a symmetrical structure and bears

a single seed (uniovulate) in the ovary. In section *Heliobrychis*, the fruit consists of a single disc with no apex. The fruits are curved and petiolate (stipitate), with a smooth stem and stiff hairs (setae) on the edges. Large, deep faveoles are arranged irregularly or in three or four rows on the fruit surface. The largest faveole is located apically and is usually spiny. The fruits of this section are densely hairy [6]. In section *Hymenobrychis*, the fruit is curved, broad-apposed, and often toothed or spiny. Deeply, large faveoles are arranged in two rows at the apex and on the disc, with linear to rectangular faveoles located at the apex. In this section, which usually contains two ovules, the ventral suture is fully curved, and the fruit tips are fused [6]. In Section *Lophobrychis*, the fruit is asymmetrical, with a toothed apex and two or three rows of deep, large faveoles on the disc. The fruits are sessile and often spiny. In Section *Onobrychis*, the fruit is sessile, erect, or slightly curved, and bears spiny or toothed crests on the disc and apex. There are no faveoles at the apex, while the disc has two regular rows of deep faveoles. The terminal faveole is longer than the others and generally spineless. The fruit is single-ovulated and has enlarged spines on its edges [6]. Zarrabian et al. grouped fruit types according to their morphological characteristics. Type 1 fruits are typical of *Onobrychis* section and are small, round, and slightly wrinkled; the oval to spherical seeds show no significant textural variation. Type 2 fruits, typical of *Sisyrosema* section, are characterized by long, cylindrical, and smooth pods; the seeds are more angular, smooth, and generally larger. Type 3 fruits, typical of *Lophobrychis* section, have flat, prominently ridged pods; the seeds are large, irregularly shaped, and have distinct surface textures. The recently described *O. crista-galli* type fruits are distinguished by long, curved pods bearing ridges or spines; the seeds are large and have a rough surface [49]. Another study on *Onobrychis* fruit morphology reported that most species are reniform, single-seeded, and characterized by lateral curvature. Some species exhibit round or slightly oval fruit shapes, while the edges may have spiny or wavy structures. Fruit surfaces may be completely smooth and flat, or they may be covered with fine hairs. Additionally, some species exhibit prominent veins or bubble-like structures on the fruit surface [50].

2.2.3. Cytogenetics

Cytological markers are variations in chromosome numbers, band patterns, sizes, shapes, order, and positions. These variations reveal differences in euchromatin and heterochromatin

distributions and can also be used in the identification of linkage groups and in physical mapping [51].

Most cytological studies in the *Onobrychis* genus have focused on chromosome number and length. *Onobrychis* species have basic chromosome number of x=7, x=8 and in the genus 2n=2x=14, 2n=4x=28, 2n=2x=16, 2n=4x=32 different chromosome numbers have been reported [52]. The *Onobrychis* genus has included diploid, tetraploid or octaploid forms but generally *Onobrychis* is diploid or tetraploid. While tetraploid species may be autopolyploid or allopolyploid, it is unclear whether the inheritance is tetrasomic or disomic [29]. This change in chromosome number has been thought to be the cause of the evolutionary change in *Onobrychis*. According to some researchers, x=8 is the ancestral chromosome number of the genus *Onobrychis*, and with aneuploidy, species with the basic chromosome number x=7 emerged, but according to some researchers, the evolution and change in the genus changed the basic chromosome number from x=7 to x=8 is thought to increase [53]. The genus *Onobrychis* has chromosomes ranging in lenght from 1.6 µm to 2.6 µm and in metacentric, submetacentric form [52].

O. viciifolia has seven basic chromosomes and includes both diploid and tetraploid varieties. *O. viciifolia* has an average length of 3.39 µm of chromosomes [54]. The scarcity of chromosome markers in *Onobrychis*, which complicates karyotype analysis, poses a problem. To overcome this problem, karyotypes were produced for comparative analysis by examining the chromosome number and the evolution of ribosomal DNA (rDNA) loci in 29 *Onobrychis* species. As a result of the analysis, two different basic chromosome numbers, x=7 and x=8, emerged in diploid and tetraploid species. In addition, it was determined that diploid *Onobrychis* species showed a high polymorphism in the number and localization of rDNA loci, and similar rDNA loci were present in polyploids. Chromosomal pattern analysis of rRNA loci in a phylogenetic context resulted in the restructuring of an ancestral 5S rDNA locus and a 35S rDNA locus in the interstitial chromosomal position within the genus, inferring that the ancestral chromosome number was x=8, and the presence of diminished dysploidy and polyploidization as the main mechanisms of chromosome number evolution [55].

Although the sainfoin is important from an agricultural point of view, its basic chromosome and centromere features have not been defined yet. Therefore, in order to determine the phylogenetic relationship between species in sainfoin, there is a need to analyze the centromeric histone CENH3 by performing genome and centromere analysis. In this study, four different CENH3 alleles were discovered in the sainfoin genome. The allelic diversity of OvCENH3s discovered in the sainfoin supports an allotetraploid origin that has not been clarified in the sainfoin [56].

2.2.4. Recent Additions of Species

The genus *Onobrychis* has a significant species richness as that is cross pollinated. However, it is a genus that is very difficult to define and determine species [3]. New species in *Onobrychis* have been determined according to such as characters: annual or perennial nature of the species, adnate or free stipules, ovule number; size, proportions, and indumentum of floral parts; and some morphological characters of the fruit (such as seed number, pubescence, with or without setae and spines, presence or absence of crest, stalked or stalkless fruit, and curvation of stalk). However, a limited number of characters are used in new species identification studies [6].

Taxonomic relationship determined and defined of *Onobrychis* species found in the northwest Alborz Mountain range in Iran with other species that is aimed. *O. alamutensis* has been determined that has different characteristics compared to other species and is a new species. *O. alamutensis*, *O. verae*, *O. ptychophylla* and *O. sosnowskyi* were found to be more closely related to long-winged species. The morphological features of *O. alamutensis* were described as a dorsoventrally flattened capsule, invertebrate disc, short toothed apex of 0.2-0.5 mm, and long wings 11 mm long [57].

Compared with Flora of Turkiye, Flora of URSS, Flora of Iranica, Flora of Iraq, Flora of Palaestina, and some general references, plant samples collected from Kars and Malatya were found to be two new species belonging to the genus *Onobrychis*. The sample collected from Malatya was named *Onobrychis fallax* Freyn & Sint. var. *longifolia* Aktoklu var. nova (Sect. *Onobrychis*), and the sample collected from Kars was named *Onobrychis atropatana* Boiss. var. *grandiflora* Aktoklu var. nova (Sect. *Heliobrychis*) [4].

2.2.5. Use of Morphological Traits in the Taxonomy of Genus

Morphological traits such as the presence and durability of sepals, fruit and seed size, shape, flower color, dorsal and abdominal ornamentation, wall thickness, pod opening, legume seed number play an important role in identifying species and revealing taxonomic relationships in *Onobrychis* genus and are frequently studied in studies [58]. Morphological traits, particularly those pertaining to the flowers, are the main diagnostic characters used to distinguish species. However, morphological traits have some advantages and disadvantages. So, the variability of these characters does not always allow a clear distinction to be made [59]. Although these traits are relatively easy to examine, the effect of location, climate and environment, being exposed to epistatic or pleiotropic effects on their formation reduces the reliability of the results. In addition, since *Onobrychis* is a cross-pollinated plant, it varies a lot in terms of morphological traits [60].

In the classification study performed on 77 species according to pollen morphology, 7 species *Dendrobrychis*, 5 species *Lophobrychis*, 14 species *Onobrychis*, 3 species *Laxiflorae*, 7 species *Anthyllium*, 3 species *Afghanicae*, 3 species *Insignes*, 21 species *Heliobrychis*, and 12 species *Hymenobrychis* with two unknown species determined to be included in the class [10]. In the study on pollen morphology in Iranian *Onobrychis* samples, it was concluded that the pollen morphology was homogeneous and that the pollen morphology could be defined by three pollen subtypes in terms of features such as pollen size, shape and ornamentation. It was concluded that there are some differences between these pollen subspecies, but these differences do not cause any change in the classification of the species [11]. Pollen size and morphology of 20 *Onobrychis* taxa found in Turkiye were investigated using both light and scanning electron microscopy. In *Onobrychis*, there is a correlation between ploidy level or chromosome number and pollen size at the taxon level; An increase in pollen grain size was found in species with higher ploidy levels [61]. In order to determine the level of difference between 5 different sainfoin species examined in *Heliobrychis* section, 27 morphological characters were examined, and variation was detected between species in terms of morphological characters. The reason for these variations is thought to be the high rate of cross-pollinated of these species and their growing in different ecological conditions [60]. The taxonomic evaluation of 44 populations of 36 *Onobrychis* taxa belonging to two subspecies and six divisions according to

their pollen morphology was made by electron and light microscopy. Prolate (dominant), perprolate and subprolate classes were determined according to the polar/equatorial axis length ratio. It has been observed that while the pollen grains differ according to the taxa in the equatorial view, the contour shapes in the polar view are almost the same for all taxa. Three types of exines were found in the studied species: reticulated, suprareticula and micro-netted. The obtained results revealed that the palynological features of the genus are relatively homogeneous and have limited taxonomic importance [62]. The genetic diversity between and within species was evaluated by evaluating the fruit, seed coat and seed storage protein counts of *O. arenaria*, *O. montana*, *O. viciifolia*, *O. alba*, *O. supina* and *O. caput-galli*. It was determined that other species showed similarity in terms of fruit morphometric characteristics, but the length, width, thickness and number of spines of the fruit of *O. caput-galli* were higher than those of other species. The highest values of seed coat thickness were recorded in diploid species and seed coats were determined to be highly variable. As a result of the evaluations made by electrophoretic analysis, it has been determined that the number and location of storage proteins vary between and within species [59]. In the study carried out to determine the species belonging to the genus *Onobrychis*, which grows naturally in Erzurum and its environs in Turkiye, and to determine the general prevalence, morphological and quality characteristics, species samples were collected from different locations during the flowering periods of 2007, 2008 and 2009. At the end of the study, a total of ten taxa representing eight species and two subspecies genotype was found. As a result of morphological evaluations, *O. atropatana* has the highest number of leaflets grandiflora, leaflet length and width are highest in *O. radiata*, number of main branches in *O. hajastana* has found. Furthermore raw ash ratio is highest in *O. stenostachya*, ADF and NDF ratio highest in *O. cornuta* and crude protein content in *O. stenostachya* has found [63].

2.2.6. Use of Cytogenetic Tools in the Taxonomy

Most of the earlier cytological studies in the genus are on the chromosome number [64]. Chromosome number, chromosome structure, size and behavior are the most widely used cytological characters. Chromosome structure and behavior in meiosis have important implications in explaining the degree of relatedness among populations and their evolution. In a karyological study within the *Onobrychis* clade, the subgenus *Sisyrosema* was found to be

monophyletic, unlike the *Onobrychis* subgenus [21]. Previous nrDNA ITS-based in a study for phylogenetic purposes showed that the *Onobrychis* subgenus was not monophyletic, while the *Sisyrosema* subgenus were monophyletic [65]. Another study suggested that neither of these two subgenra were monophyletic [8].

The morphological characters and meiotic chromosome numbers of *O. viciifolia*, *O. transcaucasica* and *O. altissima* populations in Iran were determined. It was stated that all taxa in the study had seven basic chromosome numbers and exhibited regular bivalent mating and chromosome separation in meiosis. In the study, it was determined that *O. viciifolia* and *O. altissima* were tetraploid (2n=4x=28), while *O. transcaucasica* was diploid (2n=2x=14). It was also determined that although *O. transcaucasica* is diploid, it has similar morphology with the tetraploid *O. transcaucasica* and *O. altissima* species [66]. Chromosome numbers and morphological characteristics of *O. caput-galli*, *O. fallax*, *O. cappadocica, O. aequidentata*, *O. lasiostachya*, *O. viciifolia*, *O. oxyodonta*, *O. hypargyrea* species were determined by karyological techniques. In the study, it was determined that *O. cappadocica* 2n=16, *O. viciifolia* 2n= 28 and the other species 2n=14 and the karyotypes consisted of submedian-centromeric (sm) chromosomes in *O. oxyodonta* and median-centromeric (m) chromosomes in other species [67].

In the cytological characterisation study on *O. viciifolia* varieties and related species, the 2C value of sainfoin was determined to be approximately 2.5 pg. *O. viciifolia* each chromosome length <5 μm and diploid (2n=2x=14) and tetraploid (2n=4x=28) varieties, *O. aequidentata* and *O. crista-galli*. The basic chromosome numbers were determined to be eight [54].

2.2.7. Use of Moleculer Tools in the Taxonomy

A considerable amount of work has been performed for taxonomic evaluation and delineating of genetic diversity in the genus *Onobrychis*. Molecular markers are polymorphisms that exist between the nucleotide sequences of different individuals. Insertion, deletion, point mutations, duplication and translocation are the basis of these polymorphisms. An ideal DNA marker type should be co-dominant, evenly distributed throughout the genome, highly reproducible, and

capable of detecting higher levels of polymorphism [51]. Many marker systems employed in the genus *Onobrychis* up until now.

The nuclear internally transcribed spacer region of the chloroplast genome and DNA sequences for the trnH-psbA and trnT-trnL intergenic spacers among *Onobrychis* species showed significant genetic diversity between *O. viciifolia* and other *Onobrychis* species with. Furthermore, a genetic distinction is defined between *O. viciifolia* accessions from western Europe and those from eastern Europe and Asia, reflecting a similar division based on agronomic properties [7]. Genetic diversity characterization was performed among a panel of genotypes that included a total of 29 *O. viciifolia*, 40 other accessions from sect. *Onobrychis*, seven from sect. *Hymenobrychis* DC., six from sect. *Lophobrychis*, and two from sect. *Heliobrychis* Bunge ex Boiss. DNA sequences were obtained for the nuclear internal transcribed spacer region and the trnH-psbA and trnT-trnL intergenic spacers of the chloroplast genome revealed that there was a clear distinction between *Heliobrychis* and *Hymenobrychis*, and similarities between *Lophobrychis* and *Onobrychis* were evident [7]. In the cytological characterization study on *O. viciifolia* varieties and related species, the 2C value of sainfoin was determined to be approximately 2.5 pg. with each chromosome length <5 µm and diploids having 2n=2x=14 and tetraploids having 2n=4x=28 chromosomes. *O. aequidentata* and *O. crista-galli* were reported to have the basic chromosome numbers of eight [54].

2.2.8. Use of Genomic Tools in the Taxonomic Classification

Genomic studies on sainfoin are very few in contrast to other forage crops such as alfalfa [68]. Researchers generally prefer polymerase chain reaction (PCR) based markers in their studies. These methods are Cleaved Amplified Polymorphic Sequence (CAPS), Amplified Fragment Length Polymorphism (AFLP), Simple Sequence Repeats (SSR), Inter Simple Sequence Repeats (ISSR), Sequence Characterized Amplified Region (SCAR), Random Amplified Polymorphic DNA (RAPD) and Single Nucleotide Polymorphism (SNP) [69]. RAPD, AFLP, SSR and ISSR have recently been used in genomic studies to determine population structure, taxonomic relationship and genetic diversity in sainfoin [68].

The intensity of population genomic studies on the genus *Onobrychis* has lagged behind other cultivated plants with high economic values. The first reason for this situation can be attributed to the fact that the number of researchers working on the *Onobrychis* genus is much less than other crops and the low interest of private companies in crops that do not need to be planted every year. Another reason is the scarcity of early genomics tools available for population genomic studies. Many molecular applications have restricted its use in population genomic studies, where the discovery and use of a dense and robust set of genetic markers in many individuals is required. Ideally, population genomic studies should be conducted with molecular tools that allow large numbers (hundreds to thousands) of polymorphic markers to adequately cover the entire genome in a single experiment. Because of polyploidy, obligatory cross-pollination, and large genome sizes, marker discovery and technical development in *Onobrychis* use to be very costly [68]. The first boring, unsound, and low-yield marker systems such as RFLP, AFLP, and RAPD were quickly replaced by more abundant and robust markers on the clover. As marker technology evolved towards low-cost and highly abundant SNP markers, high-resolution melting (HRM) technology based on transcriptome sequencing, and finally GBS-based SNPs were used in alfalfa genome mapping. These new and efficient population genomic tools are being used in alfalfa population genomic studies to understand patterns of genetic diversity, population structure, and association mapping [68].

Genetic diversity was determined in 80 sainfoin accessions using anatomical, morphological and molecular (by ISSR analysis) markers. As a result of the morphological evaluation were divided into three different groups that can be easily identified by features such as palatability, resistance to powdery mildew, and percentage of plant shoot and leaf. For anatomic features variation was found especially vessel diameter, sieve diameter, phloem width and xylem/phloem ratio. In the study, genetic diversity assessment was performed using 45 ISSR primers. As a result, it was concluded that the majority of genetic diversity is between accessions, there is less variation between geographic groups, and Asia and Eastern Europe may be the main diversity centers for this species [70]. RAPD markers and genetic diversity of 36 *O. sativa* populations cultivated in Iran were investigated. It was determined that 79 bands out of 5 of the 20 RAPD primers in the study were polymorphic and because of RAPD primer analysis, the genetic diversity value (uh) of 36 populations varied between 0.08 and 0.43, and

the mean uh value was 0.25. Genetic diversity with analysis of molecular variance (AMOVA) showed greater variability of 78% in populations. In addition, in the study, it was determined that the populations were divided into 3 groups according to the phylogenetic analysis [71]. In the study in which the Azerbaijani and Iranian sainfoin populations were analyzed using ISSR markers, it was determined that the rate of polymorphic ISSR loci was between 38.75% and 61.25% among the populations, and the neutral genetic diversity value of Nei varied between 0.118 and 1.179 as a result of ISSR analysis [69]. In another study, a total of 68 alleles were identified in 91 samples to determine the genetic diversity of two *O. viciifolia* cultivars (Özerbey and Lütfübey) and three local populations in Turkiye (Pleven, Kırşehir-1 and Kırşehir-2) using 10 nSSR loci specially designed for *Onobrychis* species. It was reported that a high level of genetic diversity was observed at the rate of %92 within the population and %8 between the populations. According to the UPGMA dendrogram obtained as a result of the study, it was determined that the samples were divided into 2 groups: Özerbey and Lütfübey were in one group, and Pleven and Kırşehir-2 were in the second group [73]. In the study in which the diversity of 83 sainfoin genotypes determined according to their agronomic performances was evaluated, ten nuclear simple sequence repeat (nSSR) primers were used, and it was concluded that all nSSR loci were polymorphic, and a total of 92 alleles were detected and the average allele number observed for each locus was determined as 9.2. In the study, Shannon Index I = 0.375, unbiased genetic diversity value uh = 0.243, and mean polymorphic information content PIC = 0.240. According to the data obtained from the studied sainfoin genotypes, it was concluded that they were divided into 2 main groups in the UPGMA clustering method analysis and into 3 main groups in the STRUCTURE analysis [74].

2.2.9. Inter-Primer Binding Site (iPBS) Markers

Inter-Primer Binding Site (iPBS) markers are DNA-based molecular markers that target binding sites specific to retrotransposons. These markers are very valuable for studying the contribution of retrotransposons to genome structure and their roles in evolutionary processes [75]. Since retrotransposons are widely found in many genomes, iPBS markers are effective in determining genetic diversity between species or populations [13]. The application area of iPBS markers is quite wide. In addition to the investigation of genetic diversity in agricultural and natural populations, these markers also contribute to the study of complex genetic processes such as

population structure, gene flow, and adaptation to environmental conditions [76]. They are also used in phylogenetic studies to establish evolutionary relationships through retrotransposon-based sequences [75]. In agricultural research, iPBS markers are used to determine important traits such as disease resistance, tolerance to environmental stresses, and yield. This makes these markers important in terms of breeding programs and the protection of genetic resources [77]. iPBS analyses are usually carried out based on PCR (Polymerase Chain Reaction). First, pure genomic DNA is extracted, and then DNA is amplified using primers specific to retrotransposon regions. These primers can be universal or species-specific depending on the purpose of the study [13]. The obtained PCR products are separated by agarose gel electrophoresis, and genetic differences are evaluated through banding patterns seen between different individuals [75]. These banding patterns are then subjected to statistical analysis to measure genetic distance and diversity [78]. iPBS markers have the advantage of showing high polymorphism, being cost-effective, and having a simpler protocol compared to other molecular marker systems [14]. They can be used in different organisms such as plants, animals, and microorganisms; therefore, they are evaluated in a wide range of studies such as taxonomy, phylogenetics, and genetic mapping [79]. However, the efficiency of iPBS analyses may vary depending on the accuracy of primer design and the genome structure of the organism studied [13]. Nevertheless, iPBS markers continue to be a powerful tool in genetic and evolutionary studies due to their capacity to reflect retrotransposon-induced genetic variation [80].

The iPBS retrotransposon marker system has been widely used to assess genetic diversity and population structure in legume species. In *Medicago sativa*, 45 local genotypes were evaluated using 10 iPBS primers and 267 polymorphic bands from 280 alleles were obtained, with an average polymorphism rate of 93.71% and an average PIC value of 0.14. Clustering and STRUCTURE analyses grouped the genotypes into four main clusters (Q1–Q4), indicating moderate within-population variability and the effect of local adaptation on genetic differentiation [81]. Similarly, 26 iPBS primers were applied to 40 Turkish *Phaseolus vulgaris* L. (bean) genotypes and 1,350 bands were generated with an average PIC value of 0.331, a Shannon index of 0.546 and a genetic diversity of 0.361. AMOVA determined that 67% of the total variation was within populations [82]. Furthermore, in field pea (*Pisum sativum* var. *arvense* L.) used 37 iPBS primers and detected 2,529 polymorphic bands with an average PIC

value of 0.345 and a Shannon information index of 0.299. Molecular analysis of variance showed that 74% of the genetic diversity was present within populations, while STRUCTURE and PCA analyses clearly separated plants into distinct clusters based on their geographical origin [83].

2.3. Machine Learning

Machine learning (ML) is the field of artificial intelligence that enables computers to learn from data and make predictions or decisions without being explicitly programmed. In other words, ML algorithms recognize patterns based on past data and, based on these patterns, make predictions about new data [84]. Today, ML is widely used in many different fields, including healthcare, engineering, agriculture, economics, and biology. It yields particularly successful results in tasks such as classification, regression, clustering, and dimensionality reduction [85]. In biology and agriculture, it makes significant contributions to applications such as species identification, disease diagnosis, genetic data analysis, phenotyping, and the classification of morphological traits [18]. In recent years, ML has become a fundamental tool for species identification and modeling phenotypic variation. It is particularly notable for providing fast and accurate results in applications such as species identification, disease diagnosis, and high-throughput phenotyping from images of plant leaves, fruits, and other organs. Systematic studies demonstrate that both deep learning and classical ML methods consistently achieve high success in plant biology, and a standard workflow consisting of data collection, feature extraction, and classification is gradually becoming established [86]. Using morphological traits such as shape, vein pattern, and texture in combination with ML methods for leaf-based species identification yields reliable and highly accurate results, even under field conditions. Extensive studies highlight that extracting these features with image processing techniques and transferring them to ML models offers a powerful application framework [87]. Furthermore, morphological approaches are not limited to leaves; they are also being used in other plant organs, such as fruit shape, enabling the discrimination and classification of complex morphological variations [88]. Indeed, studies on species such as strawberries and tomatoes demonstrate that fruit shape can be successfully classified using ML, and that these methods improve the understanding of multidimensional morphological data [15], [16]. Furthermore,

machine learning methods such as SVM, KNN, RF, and others have also demonstrated high accuracy in the appearance of plant leaves [89].

K-Nearest Neighbor (KNN) is a simple yet effective algorithm that stores all existing cases and classifies new examples based on similarity metrics. KNN has been used as a nonparametric technique in predictability and pattern recognition since the early 1970s. The algorithm works as follows: A new case is classified based on the majority of its neighbors in the training set, and these neighbor relationships are typically determined by distance metrics such as Euclidean distance. In other words, an example is assigned to the class that is most common among its K nearest neighbors. Its nonparametric structure requires no prior assumptions about the data distribution, making it frequently preferred for biological datasets. This method, first described by Cover et al. [90], is still used today as a robust reference method for taxonomic datasets.

Support Vector Machines (SVM) is a popular algorithm developed in 1995 [91]. It generally involves splitting data into training and test sets. Each sample in the training set contains a "target value" and several "attributes." The goal of SVM is to produce a model that accurately predicts the target values of the test data (based on training), given only the attributes of the test data. SVM focuses on finding the optimal boundary (hyperplane) for separating data into different classes and maximizes the separation margin between classes. For data that cannot be separated linearly, kernel functions are used to move the data to a higher-dimensional space, enabling linear separation within this space. Furthermore, thanks to its ability to prevent overfitting, it can also achieve strong performance on small datasets [92], [93]. Studies on leaf datasets such as Flavia and MalayaKew have repeatedly demonstrated the robust performance of SVM models trained on shape, vein, and texture features [94].

Random Forest (RF) is an ensemble method consisting of multiple decision trees and has been widely used in classification problems since the 2000s [95]. In this method, each decision tree is constructed with randomly selected examples and variables from the training data. This diversity increases the model's generalization ability and reduces overfitting [96]. During classification, the predictions from all trees are combined, and the final class is determined by a majority vote. Its noise-resistance structure and the ability to naturally calculate feature

importance are key features that distinguish RF from other methods. This allows it to identify which morphological characters are more decisive in species discrimination. Frequently compared with KNN and SVM in the classification of leaf morphology and plant diseases, RF has attracted attention with its high accuracy rates in many studies [97].

KNN, SVM, and RF are among the most frequently used machine learning algorithms in biological classification studies. While KNN offers a rapid implementation thanks to its simple structure, its accuracy can decrease in large and complex datasets [90]. In contrast, SVM demonstrates high performance, especially in small datasets and cases with nonlinear distinctions, and offers flexible classification opportunities thanks to its diverse kernel functions [98]. RF, on the other hand, is robust to noise due to its ensemble structure, which consists of multiple decision trees, and stands out with both high accuracy rates and the ability to determine variable importance [99]. Comparative studies in the literature show that SVM and RF generally provide higher accuracy than KNN, and that RF is particularly successful in large and complex data sets, while SVM provides stable results in small data sets [100]. Using the Flavia leaf dataset, Wu et al. (2007) classified plant species and found that SVM performed better than KNN with high accuracy [101].

3. MATERIALS AND METHODS

3.1. Species Taxonomically Evaluated in the Study

In this study, species from different subgenera and sections of the genus *Onobrychis* were evaluated phenotypically and genotypically. Table 3.1. shows the names and taxonomic classifications of the species used in morphological and molecular analyses, based on previous taxonomic and revision studies [4], [6], [10], [48], [102]. This table represents the most up-to-date representation of the species along with the subgenus and section infirmation compiled prior to the study.

Table 3.1. Compilation of the taxonomic information of *Onobrychis* species used in the current study.

Species name	**Subgenus**	**Section**
O.aequidentata (Sm.) d'Urv.	*Onobrychis*	*Lophobrychis*
O.afghanica Širj. & Rech.f.	*Sisyrosema*	*Afghanicae*
O.africana Širj.	*Onobrychis*	*Lophobrychis*
O.alba Desv.	*Onobrychis*	*Onobrychis*
O.alba subp. laconica	*Onobrychis*	*Onobrychis*
O.allionii Ford.	*Onobrychis*	*Onobrychis*
O.altissima Grossh.	*Onobrychis*	*Onobrychis*
O.amoena Popov & Vved.	*Sisyrosema*	*Hymenobrychis*
O.arenaria (Kit.) DC.	*Onobrychis*	*Onobrychis*
O.arenaria subp. arenaria	*Onobrychis*	*Onobrychis*
O.argentea Boiss.	*Onobrychis*	*Onobrychis*
O.argyrea Boiss.	*Sisyrosema*	*Heliobrychis*
O.armena Boiss. & A. Huet	*Onobrychis*	*Onobrychis*
O.atropatana Boiss.	*Sisyrosema*	*Heliobrychis*
O.aucheri Boiss.	*Sisyrosema*	*Heliobrychis*

Species name	Subgenus	Section
O.baldzuanica Širj.		
O.bicolor Bornm.		
O.biebersteinii Sirj.	*Onobrychis*	*Onobrychis*
O.buhseana Boiss.	*Sisyrosema*	*Heliobrychis*
O.cadmea Boiss.	*Onobrychis*	*Onobrychis*
O.cappodocica Boiss.	*Sisyrosema*	*Hymenobrychis*
O.caput-galli (L.) Lam.	*Onobrychis*	*Lophobrychis*
O.chorossanica Boiss.	*Sisyrosema*	*Hymenobrychis*
O.conferta (Desf.) Desv.	*Onobrychis*	*Onobrychis*
O.conferta subp. argentea (Boiss.) Guitt. & Kerguélen	*Onobrychis*	*Onobrychis*
O.cornuta (L.) Desv.	*Onobrychis*	*Dendrobrychis*
O.crista-galli (L.) Lam.	*Onobrychis*	*Lophobrychis*
O.cyri Grossh.	*Onobrychis*	*Onobrychis*
O.dasycephala Baker		
O.dealbata Stocks	*Sisyrosema*	*Anthyllium*
O.decumbens Wallr.		
O.ebenoides Boiss. & Spruner	*Onobrychis*	*Onobrychis*
O.echidna Lipsky	*Onobrychis*	*Dendobrychis*
O.eriophora Desv.	*Onobrychis*	*Onobrychis*
O.eubrychidea Boiss.	*Sisyrosema*	*Insignes*
O.galegifolia Boiss.	*Sisyrosema*	*Hymenobrychis*
O.gaubae Bornm.	*Sisyrosema*	*Heliobrychis*
O.gracilis Besser	*Onobrychis*	*Onobrychis*
O.grandis Lipsky	*Sisyrosema*	*Anthyllium*
O.hajastana Grossh.	*Onobrychis*	*Onobrychis*

Species name	Subgenus	Section
O.hamata Vassilcz.	*Onobrychis*	*Onobrychis*
O.heliocarpa Boiss.	*Onobrychis*	
O.hemicycla C.I. Blanche ex Boiss.	*Onobrychis*	*Onobrychis*
O.heterophylla C.A.Mey.	*Sisyrosema*	*Heliobrychis*
O.hispanica Širj.		
O.humulis subsp. humulis (Loefl.) G.López	*Onobrychis*	*Lophobrychis*
O.hypargyrea Boiss.	*Sisyrosema*	*Hymenobrychis*
O.iberica Grossh.	*Onobrychis*	*Onobrychis*
O.inermis Steven	*Onobrychis*	*Onobrychis*
O.kachetica Boiss. & Buhse	*Sisyrosema*	*Hymenobrychis*
O.kemulariae Chinth.	*Sisyrosema*	*Hymenobrychis*
O.komarovii Grossh.	*Sisyrosema*	*Hymenobrychis*
O.kotschyana Fenzl	*Onobrychis*	*Onobrychis*
O.laconica Orph. ex Boiss.	*Onobrychis*	*Onobrychis*
O.lanata Boiss.	*Sisyrosema*	*Hymenobrychis*
O.lasiostachya Boiss.	*Onobrychis*	*Onobrychis*
O.laxiflora Baker	*Onobrychis*	*Laxiflorae*
O.lipskyi Korovin	*Onobrychis*	*Onobrychis*
O.lubrychidea		
O.lunata Boiss.	*Sisyrosema*	*Heliobrychis*
O.mairei Širj.		
O.megatrophos Boiss.	*Onobrychis*	*Onobrychis*
O.melanotricha Boiss.	*Sisyrosema*	*Heliobrychis*
O.meschetica Grossh.	*Sisyrosema*	*Hymenobrychis*
O.michauxii DC.	*Sisyrosema*	*Hymenobrychis*

Species name	Subgenus	Section
O.micrantha Schrenk	*Onobrychis*	*Lophobrychis*
O.miniata Steven	*Onobrychis*	*Onobrychis*
O.montana DC.	*Onobrychis*	*Onobrychis*
O.nummularia Boiss.	*Sisyrosema*	*Afghanicae*
O.Onobrychis (L.) H. Karst.	*Onobrychis*	*Onobrychis*
O.oxyodonta Boiss.	*Onobrychis*	*Onobrychis*
O.oxytropoides Bunge	*Onobrychis*	*Onobrychis*
O.pallasii (Willd.) M. Bieb.	*Sisyrosema*	*Hymenobrychis*
O.paucidentata Pomel		
O.peduncularis DC.	*Onobrychis*	*Onobrychis*
O.persica Sirj. & Rech.f.	*Onobrychis*	*Onobrychis*
O.petraea (Willd.) Fisch.	*Onobrychis*	*Onobrychis*
O.pindicola Hausskn.	*Onobrychis*	*Onobrychis*
O.psoraleifolia Boiss.	*Sisyrosema*	*Heliobrychis*
O.ptolemaica DC.	*Sisyrosema*	*Hymenobrychis*
O.puchella Schrenk	*Onobrychis*	*Lophobrychis*
O.radiata (Desf.) M.Bieb.	*Sisyrosema*	*Hymenobrychis*
O.saravschanica B.Fedtsch.	*Sisyrosema*	*Hymenobrychis*
O.sativa Lam.	*Onobrychis*	*Onobrychis*
O.saxatilis (L.) Lam.	*Onobrychis*	*Onobrychis*
O.scardica (Griseb.) Halaczy	*Onobrychis*	*Onobrychis*
O.scrobiculata Boiss.	*Sisyrosema*	*Hymenobrychis*
O.sibirica Turcz. ex Besser	*Onobrychis*	*Onobrychis*
O.sintenisii Bornm.	*Sisyrosema*	*Hymenobrychis*
O.sirinae Nábělek	*Sisyrosema*	*Anthyllium*

Species name	Subgenus	Section
O.sosnowskyi Grossh.	*Onobrychis*	*Onobrychis*
O.squarrosa Viv.	*Onobrychis*	*Onobrychis*
O.stenorhiza DC.	*Onobrychis*	*Onobrychis*
O.stewartii Baker	*Sisyrosema*	*Insignes*
O.subacaulis Boiss.	*Sisyrosema*	*Heliobrychis*
O.subnitens Bornm.	*Sisyrosema*	*Hymenobrychis*
O.supina (Vill.) DC.	*Onobrychis*	*Onobrychis*
O.tanaitica Fisch. ex Steud.	*Onobrychis*	*Onobrychis*
O.tommasinii Jord.	*Onobrychis*	*Onobrychis*
O.tournefortii (Willd.) Desv.	*Sisyrosema*	*Hymenobrychis*
O.transcaspica V.V.Nikitin	*Onobrychis*	*Onobrychis*
O.transcaucasica Grossh.	*Onobrychis*	*Onobrychis*
O.transsilvanica Simonk.	*Onobrychis*	*Onobrychis*
O.vaginalis C.A.Mey.	*Onobrychis*	*Onobrychis*
O.vassilczenkoi Grossh.	*Onobrychis*	*Lophobrychis*
O.venosa (Desf).Desv.	*Onobrychis*	*Onobrychis*
O.verae Širj.	*Onobrychis*	*Onobrychis*
O.viciifolia Scop.	*Onobrychis*	*Onobrychis*
Onobrychis **(ZC)**		
Onobrychis**-1**		
Onobrychis**-2**		
Onobrychis**-3**		
Onobrychis**-4**		
Onobrychis **(Lab)**		

3.2. Plant Materials and DNA Analyses-iPBS Retrotransposon Markers

3.2.1. Plant Materials

Both herbarium specimens and fresh leaf materials were used to determine the genetic diversity and taxonomic relationships of species belonging to the genus *Onobrychis*. Herbarium specimens were obtained from the *Onobrychis* collection of the Missouri Botanical Garden, and those suitable for DNA isolation were selected. Fresh leaf samples of *Onobrychis* species available in the USDA-ARS Germplasm Resources Information Network (GRIN) were collected and grown under controlled conditions (60% humidity and 25 °C temperature) in the growth chamber of Adana Alparslan Türkeş Science and Technology University, and then harvested for DNA isolation. The information regarding all specimens used in the study is presented in Table 3.2**.**

Table 3.2. Species and information used in genetic analyses.

No	Name of Species	Number of samples	Fresh Leaf/Herbarium
1	*O. afghanica*	1	Herbarium
2	*O. alba* subp. *laconica*	2	Fresh Leaf
3	*O. altissima*	4	Fresh Leaf/ Herbarium
4	*O. arenaria*	5	Fresh Leaf/ Herbarium
5	*O. arenaria* subp. *arenaria*	3	Fresh Leaf
6	*O. argyrea*	1	Fresh Leaf
7	*O. atropotana*	1	Herbarium
8	*O. biebersteinii*	5	Fresh Leaf/ Herbarium
9	*O. caput-galli*	6	Fresh Leaf/ Herbarium
10	*O. chorassanica*	4	Fresh Leaf/ Herbarium
11	*O. conferta* subp. *argentea*	1	Fresh Leaf
12	*O. cornuta*	1	Herbarium
13	*O. crista-galli*	6	Fresh Leaf

No	Name of Species	Number of samples	Fresh Leaf/Herbarium
14	*O. cyri*	5	Fresh Leaf/ Herbarium
15	*O. dealbata*	1	Herbarium
16	*O. echidna*	1	Herbarium
17	*O. eubrychidea*	1	Herbarium
18	*O. gracilis*	1	Herbarium
19	*O. grandis*	4	Fresh Leaf
20	*O. humilis* subp. *humilis*	2	Fresh Leaf
21	*O. hypargyrea*	2	Fresh Leaf
22	*O. iberica*	2	Fresh Leaf
23	*O. inermis*	3	Fresh Leaf
24	*O. kachetica*	4	Fresh Leaf/ Herbarium
25	*O. kemulariare*	1	Fresh Leaf
26	*O. laxiflora*	1	Herbarium
27	*O. lipskyi*	1	Herbarium
28	*O. megatrophos*	2	Fresh Leaf
29	*O. meschetica*	1	Herbarium
30	*O. michauxii*	4	Fresh Leaf/ Herbarium
31	*O. miniata*	1	Herbarium
32	*O. oxyodonta*	2	Fresh Leaf
33	*O. oxytropoides*	1	Herbarium
34	*O. pallasii*	5	Fresh Leaf
35	*O. petraea*	2	Fresh Leaf/ Herbarium
36	*O. pscrafolia*	1	Herbarium
37	*O. pulchella*	2	Herbarium
38	*O. radiata*	1	Fresh Leaf

No	Name of Species	Number of samples	Fresh Leaf/Herbarium
39	*O. scardica*	1	Herbarium
40	*O. scrbicolata*	1	Herbarium
41	*O. seravschanica*	1	Herbarium
42	*O. sintenisii*	2	Herbarium
43	*O. sosnowskyi*	1	Herbarium
44	*O. squarrosa*	1	Herbarium
45	*O. subnitens*	1	Herbarium
46	*O. supina*	1	Herbarium
47	*O. tanaitica*	2	Herbarium
48	*O. tournefortii*	3	Fresh Leaf
49	*O. vaginalis*	3	Fresh Leaf
50	*O. vassilczenko*	5	Fresh Leaf
51	*O. viciifolia*	10	Fresh Leaf
52	*Onobrychis*-1	1	Herbarium
53	*Onobrychis*-2	1	Herbarium
54	*Onobrychis*-3	1	Herbarium
55	*Onobrychis*-4	1	Herbarium
56	*Onobrychis* (Lab)	2	Fresh Leaf
57	*Onobrychis* (ZC)	5	Fresh Leaf

3.2.2. DNA Extraction

3.2.2.1. DNA Extraction of Fresh Plants

Total genomic DNA was extracted from all the taxa using fresh leaves from growth chamber. Genomic DNA was extracted following modifed the cetyltrimethylammonium bromide (CTAB) method [103]. Beads and young leaves from the plants were added to 2 ml tubes and

100 µl CTAB was added to the tubes. The plant samples were lysed in the Tissue Disintegrator for 1.5 to 3 minutes. When the process was completed, the beads in the tube were discarded in the waste bin and 900 µl of CTAB was added to the tube and turned over several times. The tubes were then placed in a water bath set at 65°C and incubated for 1 hour. The samples were removed from the water bath, inverted several times, uncovered and allowed to cool. After the samples were cooled, 400 µl of chloroform octanol solution was added and inverted for about 2 minutes. Centrifugation was performed at 10000 rpm for 10 minutes and at the end of centrifugation, 600 µl of the supernatant portion of the samples with phase separation in the tube was transferred to new tubes and 600 µl of 2 propanol was added and the samples were incubated at -20°C for 20 minutes. After incubation, centrifugation was performed at 10000 rpm for 10 minutes and the supernatant was poured at the end of centrifugation. 600 µl of ammonium acetate was added to the pellet remaining in the tube and left to mix in the shaker for 20 minutes. At the end of 20 minutes, centrifugation was performed again at 10000 rpm for 10 minutes and the supernatant was discarded again. The pellets remaining in the tube were allowed to dry overnight (˜16 h) and after drying, the pellets were dissolved with 300 µl autoclaved dH_20. The concentrations of the DNAs obtained were determined using the Qubit kit (QuantiFluor dsDNA System, E2670).

3.2.2.2. DNA Extraction of Herbarium Specimen

50 herbarium samples and 35 taxa ranging from 1965 to 2023 were selected from the herbarium samples and DNA isolation was performed (Table 3.2.). Leaf samples were taken from the herbarium samples to cover an area of 2 mm^2 and isolation was performed with DNeasy® Plant Mini Kit (Qiagen Inc., Valencia, CA, USA). The concentrations of the obtained DNAs were measured using the Qubit Flex Fluorometer (Invitrogen by Thermo Fisher Scientific) kit.

3.2.3. Inter-Primer Binding Site (iPBS) Retrotranposon Markers Genotyping

All 134 DNAs obtained by DNA isolation were used in iPBS analysis.The selection of iPBS markers used in the study for the genus *Onobrychis* was made by considering previous studies on *Medicago* species and genotypes in the same family as the genus *Onobrychis* and studies using iPBS markers [81]. Initially, a study on molecular markers used in alfalfa provided important information on the efficiency, reproducibility and polymorphism levels of iPBS markers in legume taxa. In addition, studies using iPBS markers in plant genetic analyses were

also included; these studies allowed the selection of markers with high discrimination and their effective use in population genetics and phylogenetic analyses [14]. Based on these evaluations, 13 polymorphic iPBS markers were selected for the genetic analysis of *Onobrychis* species. Then bulk DNA was prepared from *Onobrychis* species and used in the iPBS primer temperature optimization experiment. A total of seven different temperature values varying between 46-60 °C were tested using 13 different iPBS primers and the results were evaluated (Figure 3.1). Band formation was observed at almost all temperatures between 46 and 60 °C. Since iPBS primers are a dominant marker system, multiple band formation is expected. Accordingly, the optimum annealing temperature for each primer was determined by considering the temperature value that gives the most bands (Table 3.3). Multiple band formation was observed at each temperature in primers with annealing temperatures ranging between 48 °C and 60 °C. This situation was consistently detected at 7 different temperatures in the specified range. The most intense band formation was recorded at the annealing temperatures of 49.5 °C, 53 °C and 55 °C.

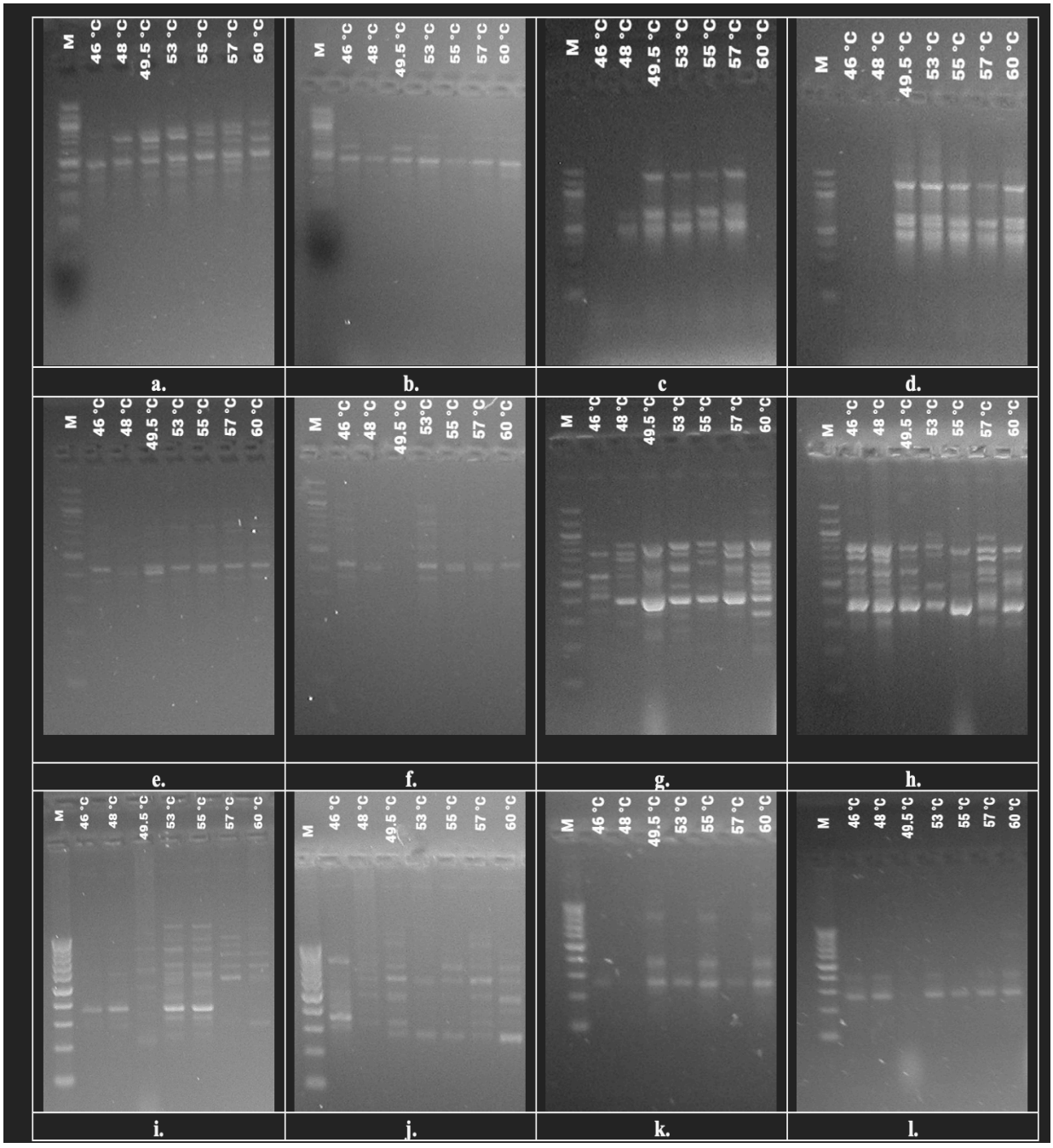


Figure 3.1. Different anneling temperatures a: iPBS 2415, b: iPBS 2389, c: iPBS 2222, d: iPBS 2401, e: iPBS 2394, f: iPBS 2395, g: iPBS 2383, h: iPBS 2228, i: iPBS 2398, j: iPBS 2491, k: iPBS 2392, l:iPBS 2393 [104].

Table 3.3. Sequence and annealing temperature (°C) of iPBS markers used for genetic characterization of genus *Onobrychis* [104].

Primer No	Primer Name	5'- 3' Sequence	Base length	AT (°C)
1	2222	ACTTGGATGCCGATACCA	18	57
2	2228	CATTGGCTCTTGATACCA	18	60
3	2383	GCATGGCCTCCA	12	49.5
4	2389	ACATCCTTCCCA	12	55
5	2391	ATCTGTCAGCCA	12	57
6	2392	TAGATGGTGCCA	12	55
7	2393	TACGGTACGCCA	12	55
8	2394	GAGCCTAGGCCA	12	60
9	2395	TCCCCAGCGGAGTCGCCA	18	60
10	2398	GAACCCTTGCCGATACCA	18	55
11	2401	AGTTAAGCTTTGATACCA	18	55
12	2402	TCTAAGCTCTTGATACCAA	19	53
13	2415	CATCGTAGGTGGGCGCCA	18	53

3.2.3.1. iPBS-PCR Amplification

PCR reactions were comprised of 10 µl volume containing: 1 µl primer (10 pmol), 1 µl 10X PCR buffer (750 mM Tris-HCI pH 8.8, 200 mM (NH4) 2SO4, %0.1 (v/v) Tween-20), 1 µl MgCl2 (25 mM), 0.25 µl dNTP (2 mM), 0.25 µl Taq DNA polymerase (5 U/µl) ve 2 µl DNA (about 20 ng/µl) and 4.5 µl dH2O. The PCR amplification was carried out under the following conditions: an initial denaturation at 95 °C for 5 minutes, followed by 38 cycles of denaturation at 95 °C for 60 seconds, annealing at 50-56 °C for 60 seconds, and extension at 72 °C for 2 minutes. This was concluded with a final extension step at 72 °C for 10 minutes. The amplified products were separated on a 1.5% agarose gel prepared in 1X SB buffer, electrophoresed at an electric field strength of 6 V/cm for 120 minutes. The gel was stained with ethidium bromide

(0.5 µg/mL) and visualized under a UV transilluminator. The sizes of the PCR products were estimated by comparison with a DNA ladder ranging from 100 to 1500 base pairs (ELK Biotechnology, Cat no.: EDL104-01/02).

3.2.3.2. iPBS Data Analysis

The evaluation of iPBS band patterns was analyzed using TotalLab TL120 software package (Nonlinear Dynamics Ltd., United Kingdom). The recording of iPBS amplification products was performed as either present (1) or absent (0). Only clear and strong bands were used and saved for further analysis.

3.3. Population Genetic Structure Analysis

Population structure analysis was performed using STRUCTURE software with 1-0 binary genetic data containing 284 dominant iPBS bands from 134 individuals of different *Onobrychis* species without any prior knowledge regarding geographical distribution or current subspecies classification [101]. STRUCTURE applies a model-based Bayesian approach using the provided iPBS marker data to determine K clusters and infer the population structure for each individual. Initially, K values ranging from 1 to 10 were evaluated for ten iterations to determine the optimal number of clusters. The burn-in length for Markov Chain Monte Carlo (MCMC) was set to 10,000 with 100,000 MCMC iterations for each replicate, based on previous studies suggesting that this parameter is sufficient [106]. The optimum K value was calculated using the LnP(D) values and determined by taking into account the ΔK method suggested by Evanno et al. (2005) and Pritchard et al. (2000) [105], [106].

A circular dendrogram was created to examine the genetic relationships between species based on iPBS marker data. For this purpose, a binary data matrix in the 1-0 (present/absent) format was prepared for each individual and the obtained data were analyzed according to Nei's genetic distance method using GenAlEx 6.51b2 software [107]. The resulting 134 × 134 genetic distance matrix was subjected to hierarchical clustering analysis using the UPGMA (Unweighted Pair Group Method with Arithmetic Mean) algorithm in R. The resulting dendrograms were visualized using Dendroscope software [108], and individuals were colored at the subgenera and section levels.

Principal component analysis (PCA) of iPBS data was performed using GenAlEx 6.51b2 software to visualize genetic differences between individuals [107]. The analysis was applied to the dominant iPBS marker data in 1-0 (present-absent) format, and the genotypic variation of each individual was represented on a two-dimensional plane through the first and second principal components (PC1 and PC2). The resulting PCA plot was used to reveal the inter-population distribution of tours, possible groupings, and genetic affinities. This analysis was also used as an additional visual assessment tool to support the STRUCTURE and dendrogram results.

3.4. Morphological Characterization and Phenotyping

3.4.1. Leaf Image Analyses

In this study, measurements were taken from digital images of leaves. The open-source ImageJ program was used [109]. The first objective was to identify simple yet reliable traits that could represent leaflet differences among species. Therefore, frequently used parameters such as length, width, area, circularity, roundness, solidity, and aspect ratio were calculated by ImageJ. Leaf images were obtained from two sources: (I) digital copies of herbarium specimens from the Tropicos database (https://www.tropicos.org/name/Search?name=*Onobrychis*) at the Missouri Botanical Garden and (II) fresh leaves from living plants grown in a growth chamber under controlled conditions. This allowed for comparison of both old herbarium material and specimens grown under standard conditions. Before measurements, all images were converted to cm using the "Set Scale" tool in ImageJ. The scale bars in the photographs were then converted to grayscale, thresholded to separate the leaves from the background, and converted to binary format. Leaflets that touched or overlapped were separated when necessary using the "Paintbrush" tool. Measurements were then performed in ImageJ using the "Particle Analysis" tool (Figure 3.2). A simple macro written in ImageJ was used for these operations. Manual checks were also performed in the program to ensure accuracy of measurements, in addition to automated processes. The obtained data were listed by species and averaged.. Area and perimeter measurements were also used to indicate leaf size and complexity. Leaflet types were also evaluated according to the definitions of Ellis et al. (2009) [109] and Hickey (1973) [110]. Acute, obtuse, rounded, mucronate, cuspidate, emarginate, and truncate tip types were

considered. Leaflet bases were classified as acute, rounded, cordate, oblique, and cuneate. Leaflet venation was also manually assessed from the images. In addition, indumentum absence (1/0) was determined manually. Finally, the leaflet shape of each species was classified based on the calculated AR and circularity values. Twelve traits were determined manually using ImageJ for the samples belonging to each taxon (Table 3.4). This morphological analyses allowed for a more objective comparison of leaf morphologies across species.

Table 3.4. Traits measured in leaflets of taxa included in the genus *Onobrychis*.

Trait name	Status	Trait description
Leaflet Circularity	number	Determining how smooth and circular the leaflet outlines are through image processing.
Leaflet Aspect Ratio	number	Determination of the ratio of the leaflet's major axis length to the minor axis by image processing.
Leaflet Round	number	The measure that shows how close the leaflet shape is to a circle; it is calculated by the ratio of the area to the square of the major axis length.
Leaflet Shape	Linear/Linear-oblong/Oblong/Oblong-elliptic/Elliptic/Elliptic-ovate/Ovate	General form of the leaflet determined by aspect ratio and circularity with image processing.
Indumentum Presence	Absent/Present (0-1	Identifying the presence of hair-like structures on the leaflet surface using image processing.
Leaflet Apex (type)	Acute/Acuminate/Rounded-obtuse/Acute-acuminate	The tip shape of a leaflet determined by image analysis based on the curvature and angle of the leaflet apex.
Leaflet Base	Cuneate/Attenuate/Rounded/Obtuse/Rounded-subcuneate/Rounded-cuneate/Rounded-subcordate	The basal shape of a leaflet determined by image analysis based on the curvature and contour of the lower margin.
Leaflet Solidity	number	The ratio of the leaflet area to the area of its convex hull measured by image analysis, indicating the compactness and smoothness of the leaflet shape.
Leaflet Area	mm^2	This is the total area of a single leaflet or the average area of a sample of leaflets expressed in mm^2.

Trait name	Status	Trait description
Leaflet Perimeter	mm	The total length of the outline of a leaflet measured by image analysis, expressed in millimeters.
Leaflet Length	mm	The average length of leaflets in a compound leaf measured by image processing is expressed in mm.
Leaflet Width	mm	The average width of the leaflets in a compound leaf measured by image processing is expressed in mm.

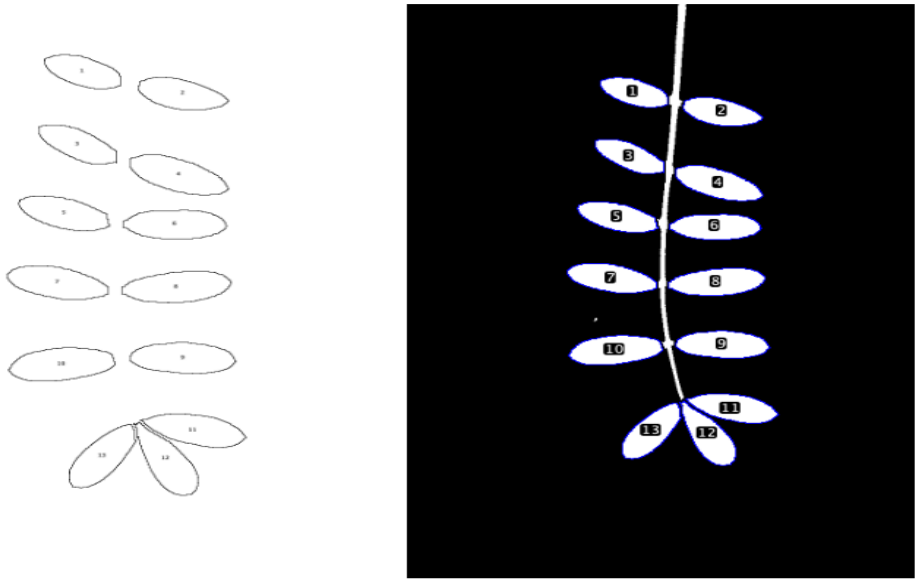
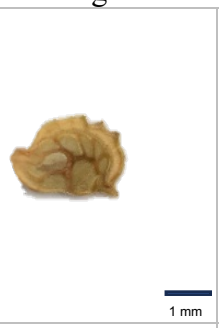
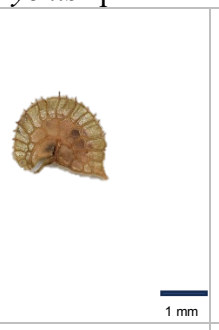
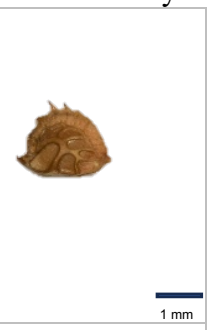
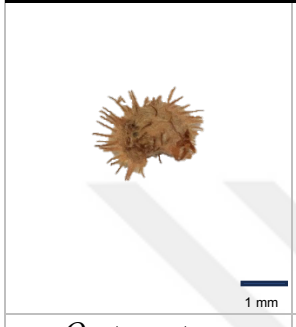
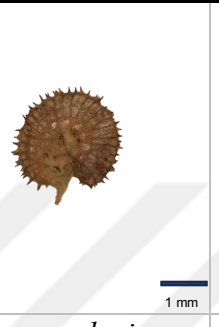
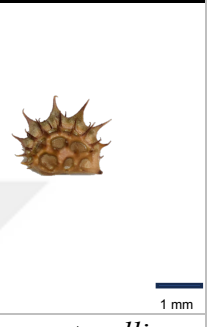
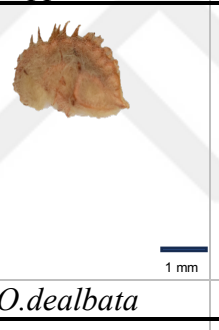
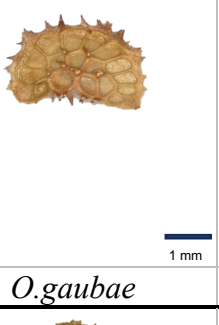
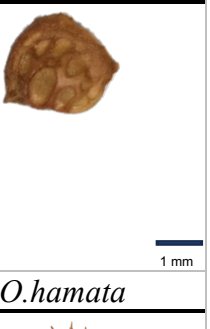
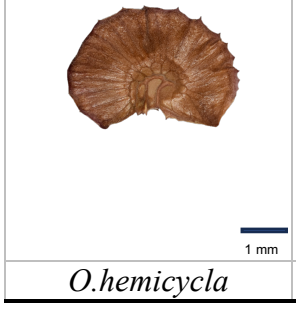
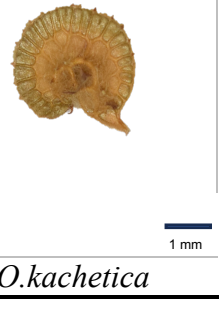


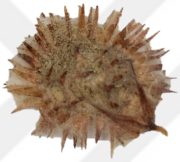
Figure 3.2. Measurement of morphological traits of leaves belonging to *Onobrychis* species using ImageJ.

3.4.2. Fruit Image Analyses

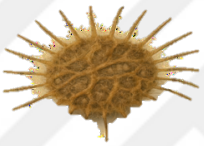
To determine interspecific variation in fruit morphology within the genus *Onobrychis*, high-resolution herbarium specimen images obtained from the Missouri Botanical Garden through the Tropicos online database (https://www.tropicos.org/name/Search?name=*Onobrychis*) and fruit images taken from herbarium specimens were used for morphological analyses (Table 3.5).

Table 3.5. Representative fruit images of the *Onobrychis* species examined in the study.

			
O.alba	*O.altissima*	*O.amoena*	*O.arenaria*
			
O.atropatana	*O.aucheri*	*O.cappodocica*	*O.caput-galli*
			
O.chorossanica	*O.crista-galli*	*O.dealbata*	*O.eubrychidea*
			
O.galegifolia	*O.gaubae*	*O.grandis*	*O.hamata*
			
O.hemicycla	*O.kachetica*	*O.laconica*	*O.lanata*

			
O.lasiostachya	*O.laxiflora*	*O.lubrychidea*	*O.lunata*
			
O.mairei	*O.melanotricha*	*O.meschetica*	*O.michauxii*
			
O.micrantha	*O.nummularia*	*O.oxyodonta*	*O.oxytropoides*
			
O.pallasii	*O.paucidentata*	*O.peduncularis*	*O.persica*
			
O.petraea	*O.psoraleifolia*	*O.ptolemaica*	*O.puchella*

			
O.radiata	*O.sativa*	*O.saxatilis*	*O.sibirica*
			
O.sintenisii	*O.sosnowskyi*	*O.squarrosa*	*O.stenorhiza*
			
O.stewartii	*O.subnitens*	*O.supina*	*O.transcaucasica*
			
O.verae	*O.venosa*	*O.viciifolia*	*O.aequidentata*
			
O.africana	*O.argyrea*	*O.buhseana*	*O.cadmea*

			
O.conferta	*O.cyri*	*O.eriphora*	*O.gracilis*
			
O.hajastana	*O.heliocarpa*	*O.hispanica*	*O.kotschyana*
			
O.pindicola	*O.saravschanica*	*O.scrobiculata*	*O.tommasinii*
			
O.transcaspica	*O.vaginalis*	*O.iberica*	

Fruit image analyses were performed using ImageJ software [17]. The dimensions of each image were calibrated by converting pixel measurements to centimeters using the "Set Scale" feature in ImageJ. This process ensured accurate measurements in the physical units (cm) required for comparative morphological studies. Images were first converted to 8-bit grayscale, and then automatic thresholding was applied using the "Default" method to separate fruits from the background. Binary transformation and morphological filtering (on-off) were performed to enhance fruit border definition. Fruits separated from herbarium specimens were manually marked with the "Paintbrush" tool and then individually selected with the "Wand (Tracing)

Tool." The area, length, width, perimeter, circularity, roundness, solidity, and aspect ratio (AR) of each fruit were measured using the Particle Analysis function in ImageJ software (Figure 3.3). Based on the obtained morphological parameters, fruit shapes were determined in R studio. Accordingly, the following were classified:

- Fruits with AR values between 0.95 and 1.05 and circularity >0.7 were classified as Circular;
- Fruits with AR values between 0.4 and 0.7 and circularity between 0.4 and 0.7 were classified as Elliptic;
- Fruits with AR values between ≥1.05 and circularity between 0.7 and 0.85 were classified as Semicircular; and
- Those with circularity <0.4 or AR <0.95 were classified as Reniform.

This analysis minimized subjective errors commonly encountered in manual classifications and provided repeatable, objective fruit shape categorization. Additionally, for each species, traits such as the presence and number of fruit spines, spines on spines, fruit trichomes, crest of fruit, and fruit veins were manually determined from fruit images. These traits were coded as 0 (absent) and 1 (present) in the data matrix. Morphological features were then determined species-wise by averaging the measured features for each taxon. A total of 15 fruit features were manually identified and measured using ImageJ (Table 3.6). All morphological and descriptive data were stored for use in further statistical analyses such as principal component analysis (PCA), dendrogram, and cluster analysis.

Table 3.6. Traits measured in fruits of taxa included in the genus *Onobrychis*.

Traits Name	Status	Trait description
Fruit shape	Eliptic/Circular/Semicircular/Reniform	The shape of a fruit according to the fruit shape rating scale.
Fruit length	mm	The length of a legume fruit along its main axis measured in mm or the length of the major axis of an ellipse fit to a legume fruit contour expressed in mm.
Fruit width	mm	The length of a legume fruit along its minor axis measured in mm or the length of the minor axis of an ellipse fit to a legume fruit contour expressed in mm.
Fruit aspect ratio	number	The aspect ratio of a legume fruit when measured by using basic image processing algorithms.
Fruit trichome presence	Absent/Present (0-1)	Binary outcome whether or not trichomes are present on a fruit surface.
Fruit spine presence	Absent/Present (0-1)	Binary outcome whether or not spines are present on a fruit surface.
Crest of fruit presence	Absent/Present (0-1)	The binary outcome of the presence/absence of a crest on the surface of the fruit.
Fruit vein presence	Absent/Present (0-1)	The binary outcome of the presence/absence of *Onobrychis* veins on the surface of the fruit.
Fruit area	mm^2	The area of a legume fruit as projected onto the 2-D surface of the image plane expressed in mm^2.
Fruit perimeter	mm	The perimeter of the 2-D projected contour of a legume fruit expressed in mm.

Traits Name	Status	Trait description
Fruit round	number	The measure that shows how close the shape of a fruit is to a perfect circle is calculated according to the ratio of the fruit area to the square of the major axis length.
Fruit circularity	number	The measurement that shows how smooth and circular the outer lines of the fruit are is calculated according to the ratio of the fruit area to the square of its perimeter.
Fruit spine number	number	The total number of spines on one side of a fruit or the projected surface of a fruit image.
Fruit spine on spine presence	Absent/Present (0-1)	Indicates whether there are thorns on the fruit surface.
Fruit solidity	number	The ratio of the fruit area to the convex wall area indicates how full the shape is or how smooth the surface is.

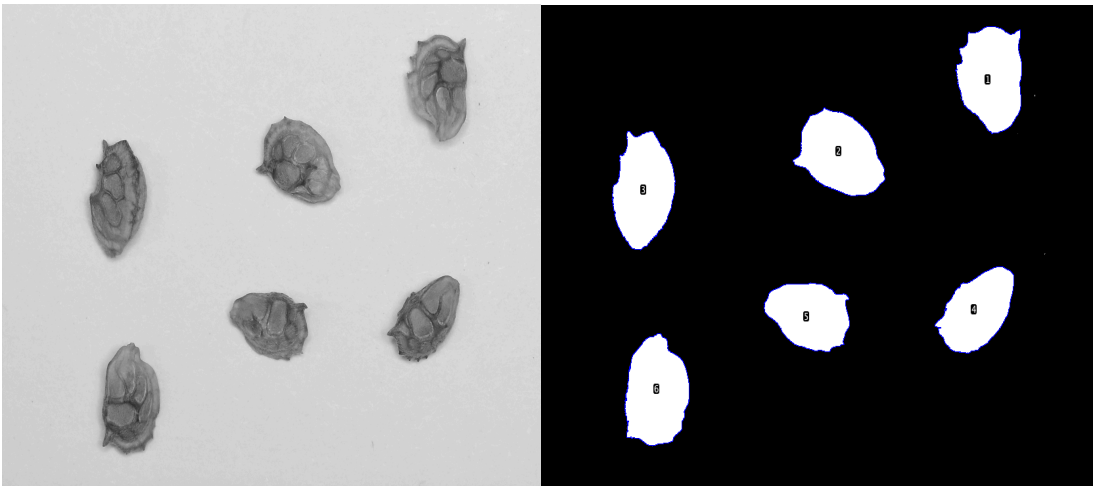


Figure 3.3. Measurement of morphological traits of fruits belonging to *Onobrychis* species using ImageJ.

3.4.3. Analysis of Leaf and Fruit Morphological Data

Leaf morphological parameters obtained as a result of measurements made with ImageJ were analyzed using R Version 2025.05.0+496 (R Project for Statistical Computing, www.r-project.org/). Leaf area, length, width and aspect ratio traits were used in these analyses. To determine the relationships between leaf morphological traits, pairwise correlation analyses of quantitative characters were performed. Pearson's correlation coefficient (r) was used in the analyses. All data were standardized prior to analysis, and only the quantitative variables obtained from image analysis circularity, AR, round, solidity, area, perimeter, length, and width were considered. The average of traits was determined for each species, then PCA analysis was performed with the obtained averages and the PCA graph of the species was drawn according to the information obtained from the classical classification and the subgenus and section information. Then, in order to standardize the extreme values and make the distribution suitable for the normal distribution, a new PCA drawing was made according to the subgenus and section information by providing log transformation of the values. Euclidean distance was calculated according to the log-transformed measurements of each species and according to these distances, the Ward.D2 algorithm was used to determine which group the species belonged to. A dendrogram was drawn according to the obtained data.

Fruit morphological parameters obtained as a result of measurements with ImageJ were analyzed using R Version 2025.05.0+496 (R Project for Statistical Computing, www.r-project.org/). In these analyses, fruit area, length, width, aspect ratio, Fruit trichome presence, Fruit spine presence, Crest of fruit presence, Fruit vein presence, Fruit perimeter, Fruit circularity traits were used. To perform species-based analysis, the averages of morphological traits belonging to the same species were taken. Then, the data was made suitable for normal distribution with Z-score transformation. Correlation analysis was conducted to evaluate linear relationships between fruit morphological traits. Pearson's correlation coefficient (r) was used in the analyses, and significance levels were set at $p < 0.05$, $p < 0.01$, and $p < 0.001$. All visualizations were performed in the R software environment using the GGally and ggplot2 packages. The correlation matrix includes the density distributions of the variables in the diagonal section, scatter plots in the lower section, and correlation coefficients in the upper section. Circularity, AR, round, solidity, area, perimeter, length, width, and fruit spine number variables were used in the analysis. PCA analysis was performed with the obtained data and two different visualizations were made according to subgenus and section of these PCA analysis results. The numerical distance of each species to other species according to its morphological data was determined with Euclidean Distance. Finally, a dendrogram was drawn to visualize the groups in a hierarchical tree structure for similar species according to the distance between them.

Elbow, Gap statistics, and Silhouette methods were used to determine the optimal number of clusters to be used in clustering analyses using morphological data from leaves and fruits. These analyses were conducted using R software.

3.5. Machine Learning Analyses

3.5.1. Data Set

Morphological measurements of fruit and leaflets of species belonging to the genus *Onobrychis* were compiled in Microsoft Excel and transferred to the Google Colab platform for analysis. Google Colab is a working environment that allows for online and free Python-based analyses. A total dataset of 967 leaflet and fruit samples from 48 different species, including fruit and leaflet measurements, was prepared for machine learning-based analyses. Morphological traits

for each sample were included, ensuring an equal number of samples from the same species to ensure data integrity. The compiled dataset was converted to a suitable format for statistical and machine learning analyses.

3.5.2. Preprocessing and Feature Selection

A total of 42 morphological features, initially obtained from fruit and leaflet measurements in the dataset, were used in the machine learning analyses. First, categorical variables were quantified using one-hot coding, and numerical variables were normalized using the StandardScaler method. Pearson correlation analysis was applied to examine multicollinearity among variables, and variables with high correlations ($r > 0.90$) were identified and removed from the dataset (Figure 3.4).

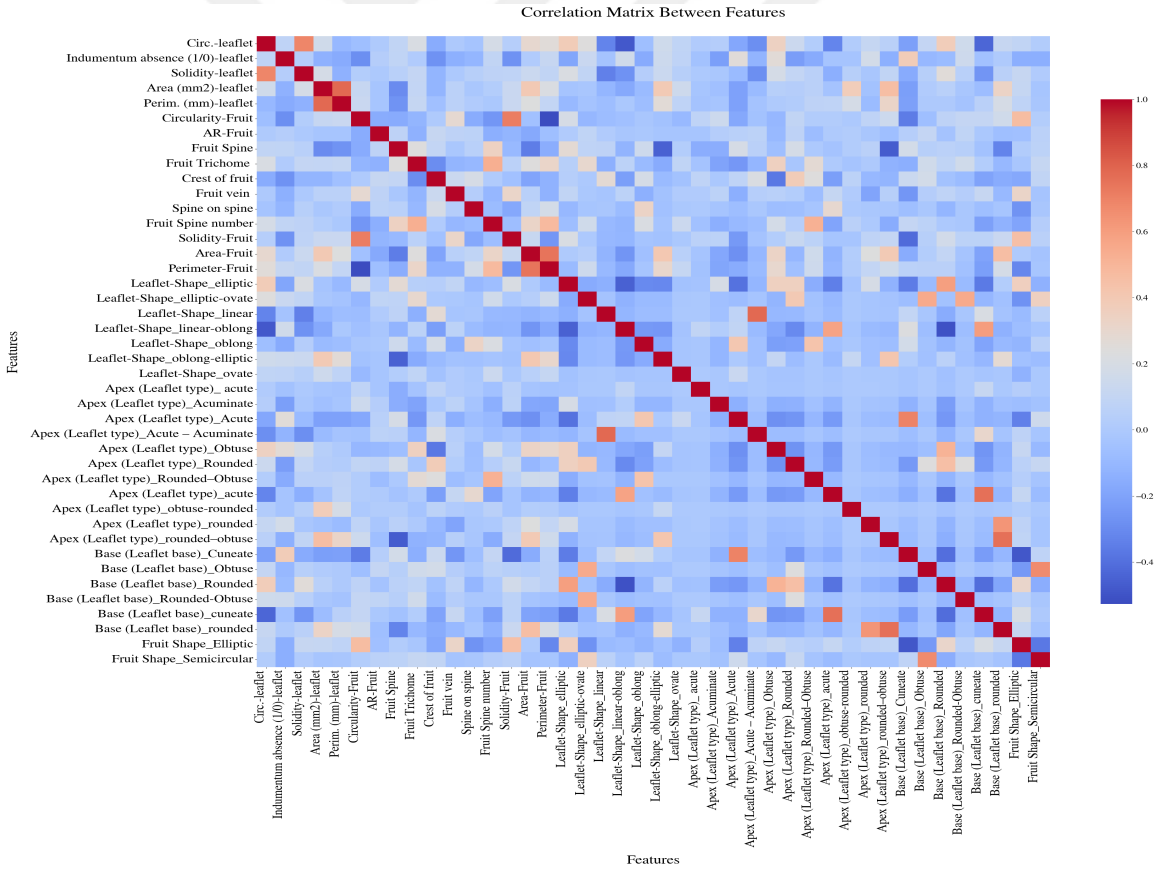


Figure 3.4. Correlation matrix (Pearson correlation analysis) created to examine the relationships between morphological traits.

Additionally, to assess the contribution of traits to classification, feature importance values were calculated using the Random Forest algorithm. Sixteen traits with an importance above 2.5 were selected, and the machine learning analyses were repeated using these variables (Figure 3.5). This approach contributed to the models producing more stable and generalizable results.

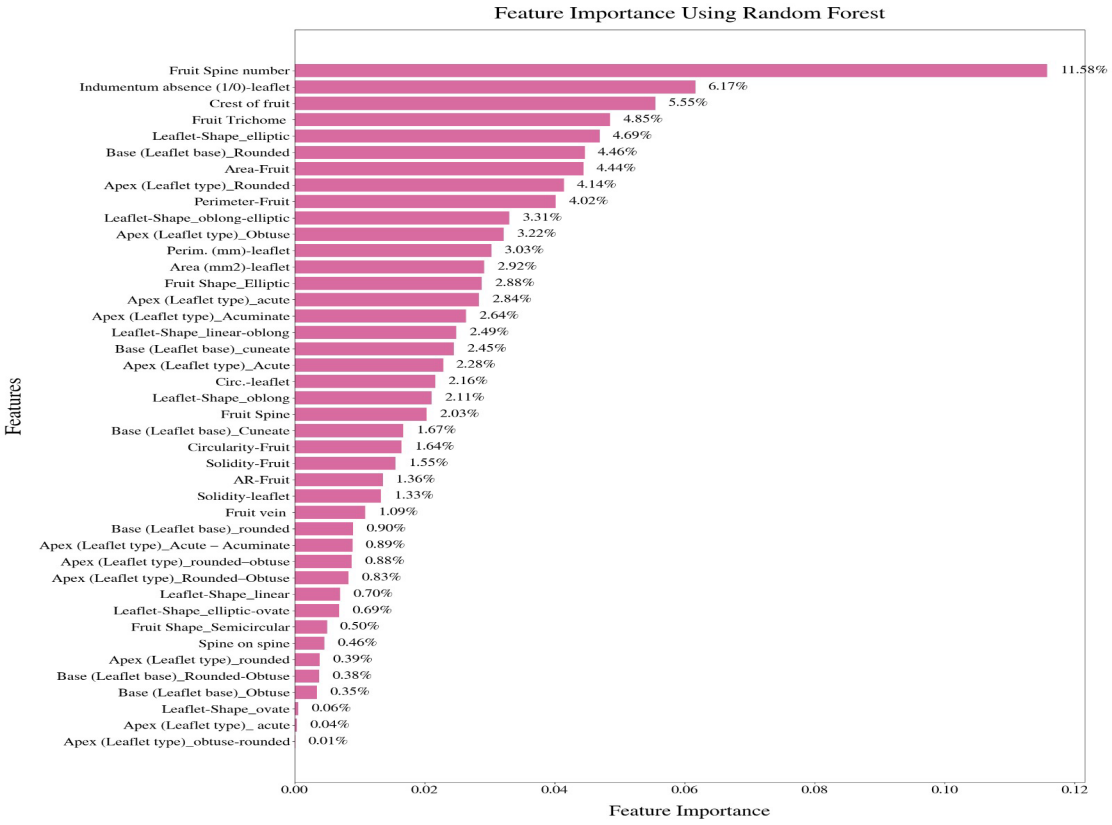


Figure 3.5. Feature importance values calculated with the Random Forest algorithm.

3.5.3. Machine Learning Algorithms

Three different ML algorithms were used to classify morphological data:

- Random Forest (RF): This method uses the ensemble approach of decision trees. It was used specifically in this analysis to determine the importance of variables. Thus, the morphological traits that contribute most to species discrimination were identified.

- K-Nearest Neighbors (KNN): This is a simple yet powerful method that performs classification based on distances between samples. The optimal k value was determined using the GridSearchCV method, and the best performance was achieved with a neighborhood value of k = 1. While the KNN algorithm achieved high accuracy for some species, classification errors were observed in species that were very similar morphologically.
- Support Vector Machines (SVM): This method, which has strong discrimination ability even with nonlinear boundaries, used the RBF kernel. The classification performance of the model was evaluated using the accuracy value and confusion matrix.

3.5.4. Model Evaluation

For all algorithms, the dataset was divided into two parts: 80% for training and 20% for testing. The performance of the models was measured using accuracy and confusion matrix. The model with the highest accuracy was accepted as the recommended method in the study.

4. RESULTS AND DISCUSSION

This study aimed to assess the taxonomic relationships and interspecific diversity of the genus *Onobrychis* using a set of genetic and phenotypic datasets and methods. First of all, the genetic diversity and phylogenetic relationships of the species at the nuclear DNA level was examined using iPBS markers. Fruit and leaf morphological traits of *Onobrychis* taxa were further measured in detail in herbarium and growth chamber specimens. Morphological diversity among taxa was determined, and a digital morphological scale for taxa identification was developed. Machine learning algorithms were applied to interpret the phenotypic data, demonstrating the potential of phenotypic traits for taxonomic classification and diversity. These findings provide a holistic taxonomic approach to the genus *Onobrychis* at the genetic and morphological levels and contribute significantly to future systematic studies.

4.1. iPBS Analyses

The 13 selected primers amplified 284 bands that were polymorphic among 51 *Onobrychis* species/taxa and the two unknown species. The amplified DNA ranged from 100 to 1500 bp.

4.1.1. Genetic Structure

The first set of analyses condcuted with the molecular markers was the Bayesian approach implemented in the STRUCTURE program. The analyses of iPBS markers revealed important findings by evaluating the genetic structure of 134 individuals of *Onobrychis* species using 284 bands. Structure analyses typically allow testing number of groups and infer the genmoic compostion of eack of the individual/genoype based on the atribution to each of the groups. Analyses for different K values (K=1-10) and tehir indepth evaluations provided information in understanding the complexity between the taxonomic structure of the genus and the genetic structure. For example, K=3 structure showed that the species cluster into three main genetic clusters. While this clustering was broadly in line with existing taxonomic patterns, the separate positioning of some species indicates the existence of a potentially new subgenus (Figure 4.1B). This suggests a deviation from classical morphological grouping. The genetic structure at K=5 was further investigated to see if the grouping would be in congurent with sections of the genus and results revealed significant genetic diversity not only among the existing subgenera but also

among the species belonging to the proposed new subgenus (Figure 4.1C). In particular, species belonging to this new subgenus differed only slightly from the other groups, supporting the clear genetic separation of this group. These findings provides a strong basis for a revised approach to the current *Onobrychis* species. On the other hand, the presence of genotypes belonging to more than one cluster in some species indicates that there may be hybridizations among species and even hybrid subspecies. A similar study for genus *Medicago* revealed the hybrid status of some of the taxa [111], [112], [113]. The extensive hybridization observed here sould also imply that speciation is ongoing or that some taxonomically defined species need to be re-examined. Overall, the analysis of STRUCTURE reveals that there is significant genetic differentiation among the subgenera that is also showing a hierarcical clustering pattern.

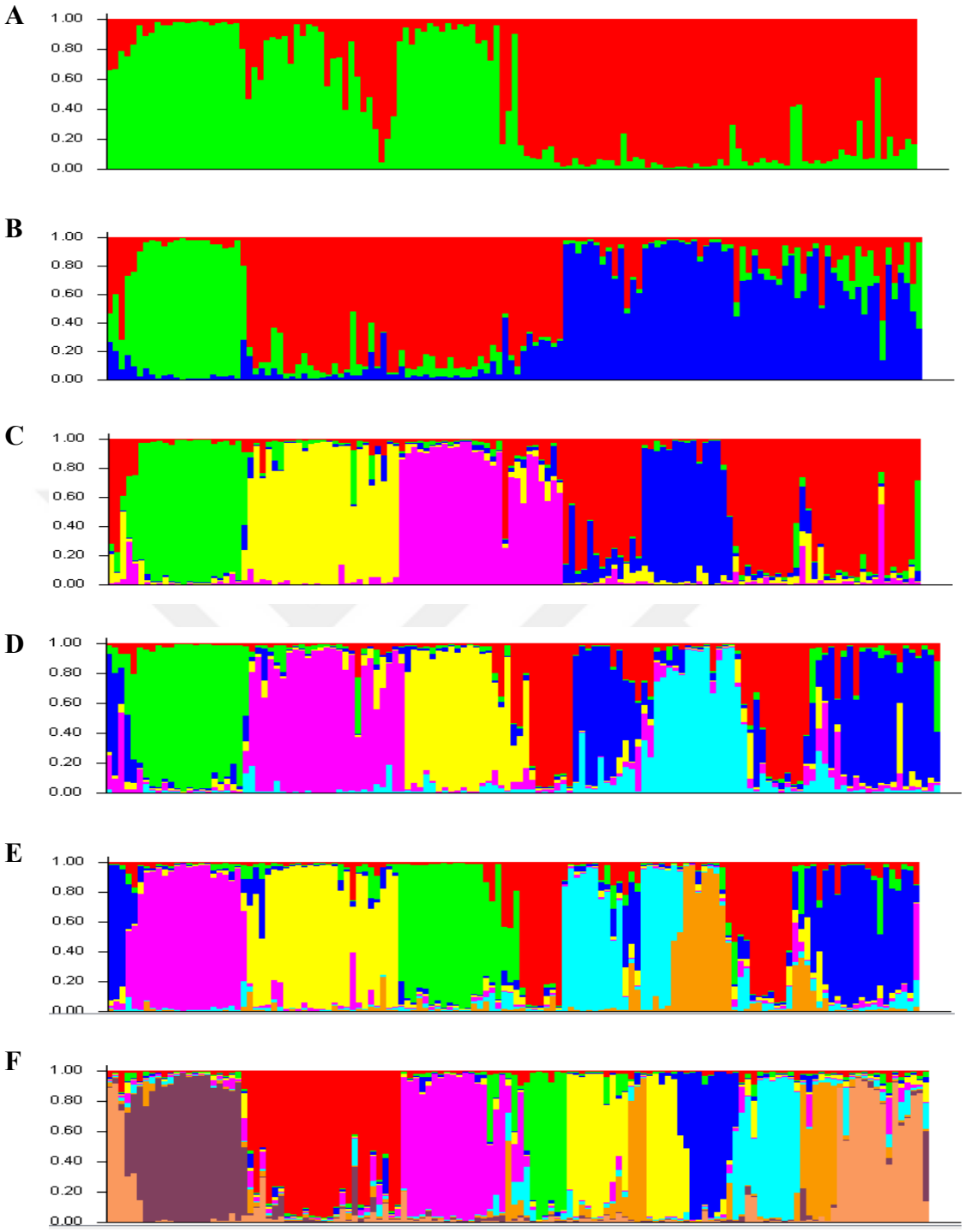


Figure 4.1. Genetic structure plot obtained by Bayesian clustering analysis: STRUCTURE software results for 134 *Onobrychis* samples analyzed using 284 iPBS bands. Each individual is represented as a vertical bar, which is divided into different colors to show the probability (with % probability) that the individual belongs to different genetic clusters. (A) K = 2, (B) K = 3, (C) K=5, (D) K=6, (E) K=7, (F) K=9.

The results of the STRUCTURE analysis were further analyzed for assessing the optimal number of groups based on the methods described by Evanno et al. and Pritchard et al. [105], [106]. Although values between seven and nine reflect the likely genetic structure, the Evanno method provides the strongest support for K = 9. Therefore, the *Onobryhcis* taxa examined in this study are thought to be divided into nine distinct genetic clusters (Figure 4.2). The result for K = 9 obtained by Evanno's method suggests that the four sections of the *Onobrychis* subgenus in classical classification are *Dendrobrychis*, *Lophobrychis*, *Onobrychis* and *Laxiflorae*, and the five sections of the *Sisyrosema* subgenus are *Anthyllium*, *Afghanicae*, *Heliobrychis*, *Hymenobrychis* and *Insignes* [4], [9], [10].

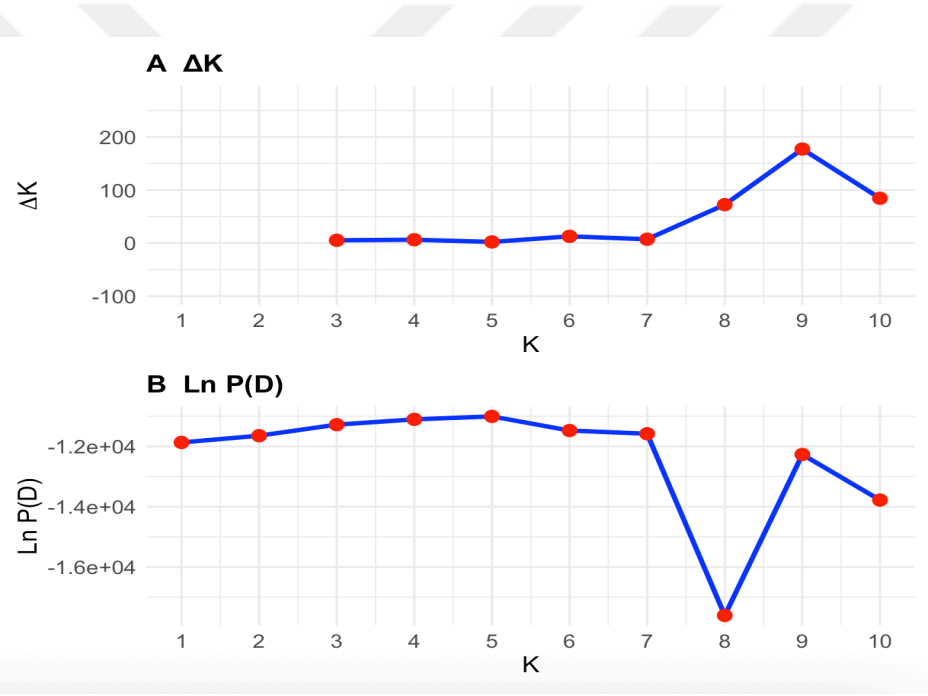


Figure 4.2. Two methods were used to determine the most appropriate value of K: (a) the ΔK statistic developed by [106] and (b) the ad hoc procedure proposed by [105].

The second approach to analyses of the genetic data was the use of Principal Component Analysis (PCA) of the iPBS data with GenAlEx (V 6.51b2) based on the genetic distance. PCA using iPBS markers showed that the first three axes explained 39.38% of the total genetic variance in the 134 individual genotypes. PC1 alone showed the most significant genetic differentiation, explaining 23.64% of the variation. The first two PC plots explained a

cumutaive variance of 33.07% and showed clear clustering patterns indicating distinct genetic groupings between populations. The results showed that individuals clustered into three apperent groups (Figure 4.3, Table.4.1) and these groups exhibit genetically distinct structures. In addition to the *Onobrychis*, *Sisyrosema* subgenera, individuals belonging to a separate group were observed to form a distinct cluster in the PCA. These indicates the existence of a genetic structure that is largely parallel to classical taxonomic classifications and somewhat different from the findings obtained from STRUCTURE analysis.

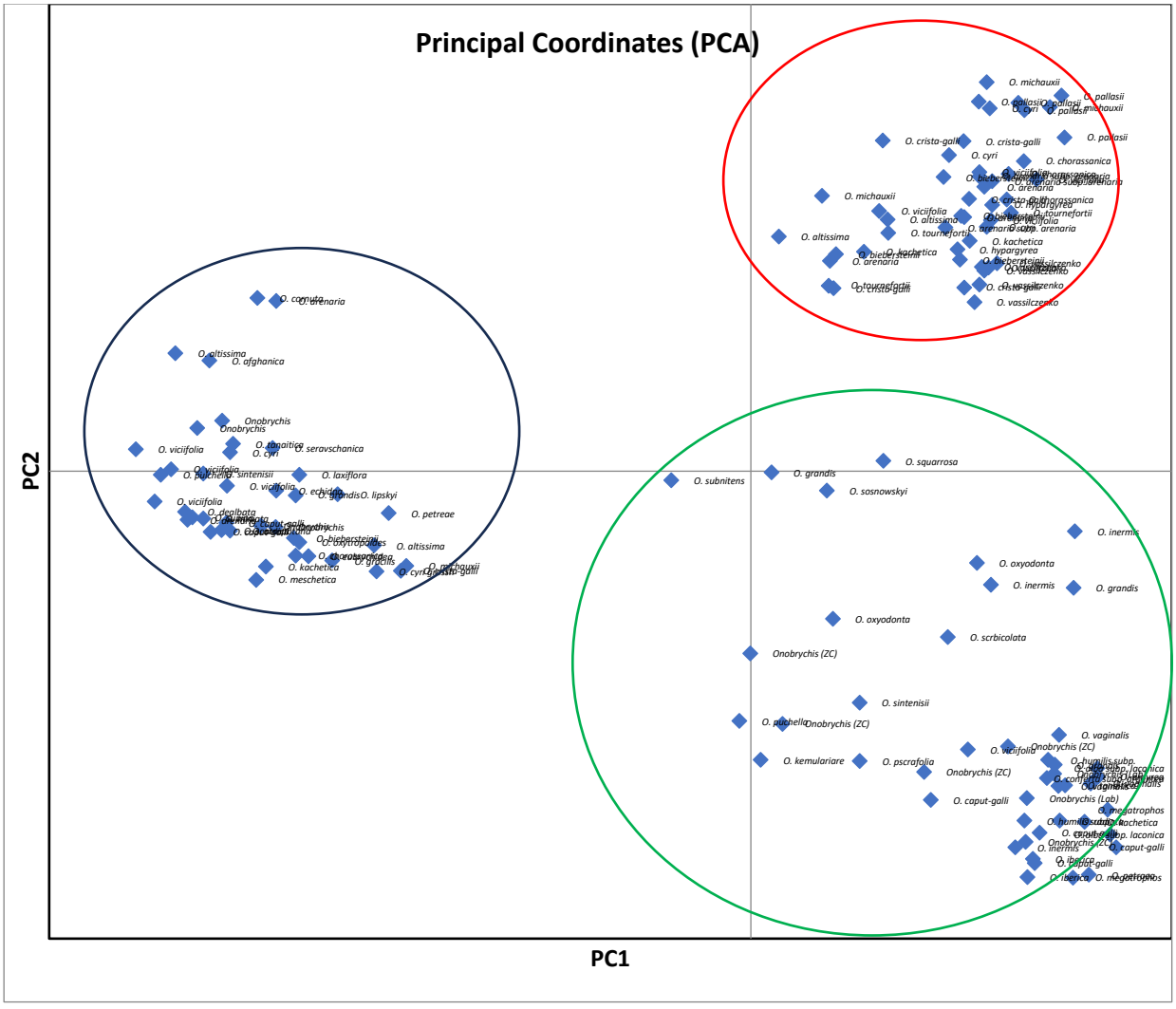


Figure 4.3. Distribution of 51 *Onobrychis* species and 6 different species of *Onobrychis* of unknown species based on the first two PCA obtained from the analysis of 284 iPBS bands.

Table.4.1. Percentage of variation explained by the first 3 axes in Principal Component Analysis (PCA)

Axis	**PC1**	**PC2**	**PC3**
%	23.64	9.43	6.31
Cum %	23.64	33.07	39.38

A dendrogram was aslo drawn according to the genetic distance results obtained from the GenAlEx program of the iPBS data (Figure 4.4 and Figure 4.5). The circular dendrogram created using iPBS markers to further infer the the genetic relationships among species visually. It was observed that individuals belonging to the same species generally formed close clusters in the dendrogram. This situation shows that most species have a genetically consistent structure. On the other hand, different individuals of some species are found to be in distant clades. The presence of individuals from some species in different clades suggests that intraspecific genetic variation may be high or that these individuals have acquired different genetic traits due to hybridization or misidentification.



Figure 4.4. Subgenus-based UPGMA dendrogram constructed using genetic distance data based on iPBS markers (blue: *Onobryhcis* subgenus, red: *Sisyrosema* subgenus).

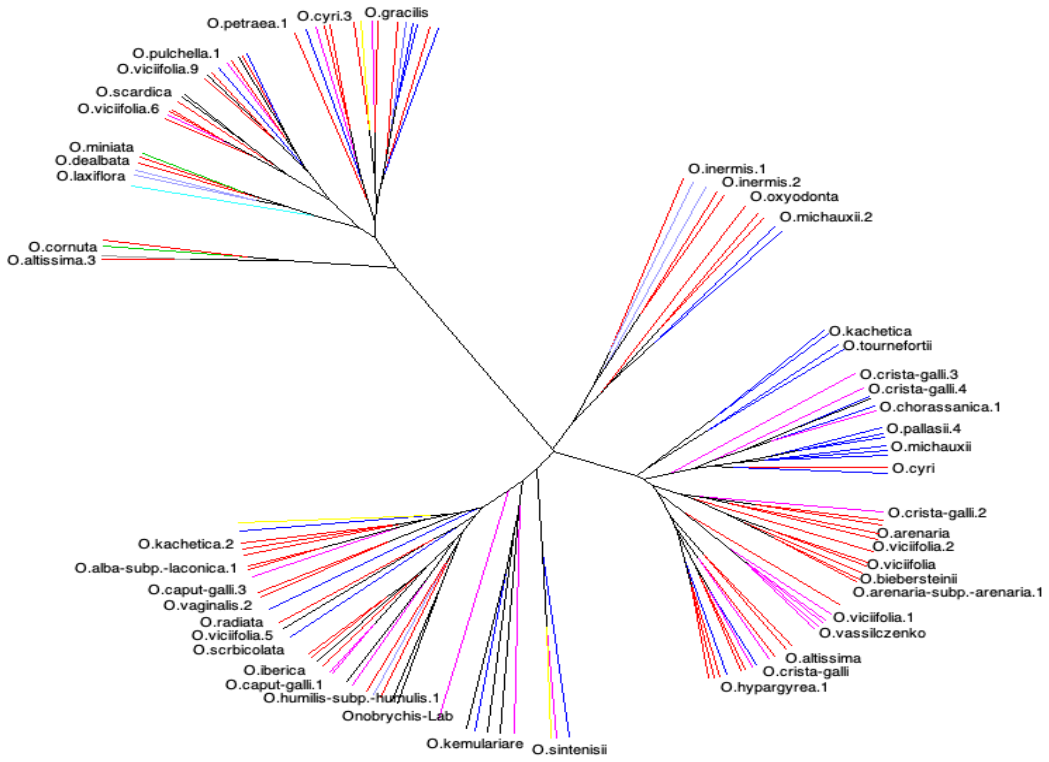


Figure 4.5. Section-based UPGMA dendrogram constructed using genetic distance data based on iPBS markers (grey: *Afghanicae*, lilac: *Anthyllium*, yellow: *Heliobrychis*, blue: *Hymenobrychis*, purple: *Insignes*, red: *Onobrychis*, green: *Dendrobrychis*, pink: *Lophobrychis*, orange: *Hemicyclobrychis*, turquoise: *Laxiflora*).

Genetic similarities among species within the *Onobrychis* genus were determined using retrotransposon-derived iPBS markers, and a circular hierarchical clustering dendrogram was constructed based on the obtained data. The resulting dendrogram revealed that the majority of species clustered within their respective subgenera (*Onobrychis* and *Sisyrosema*), consistent with existing taxonomic classifications. Genetic analysis revealed that species within the *Onobrychis* subgenus (shown in red) generally formed tighter clusters, suggesting a more genetically homogeneous structure within this subgenus. Species within the *Sisyrosema* subgenus (shown in blue) were found within larger, more dispersed clusters, yet formed consistent groups within subgenera, demonstrating that the taxonomic distinction is also evident at the genetic level. This suggests that subgeneric distinctions within the *Onobrychis* genus may

is based on both morphology and genetics [8], [114]. Some species shown in gray in the dendrogram are species that have not been clearly classified at the subgenus level in the existing literature or have been left undetermined due to lack of data. However, the positions of these species in the dendrogram provide important clues about the subgenera they may belong to. For example, species such as *Onobrychis humilis* subp. *humulis*, *O. raidata* were observed to be in the same clusters with largely *Onobrychis* subgenus. This suggests that these species likely belong to the *Onobrychis* subgenus. Similarly, considering that *O. megatraphos* are also located within the *Onobrychis* subgenus, a revision of the classification of these species may be recommended. However, the positioning of some species outside the current subgeneric structures suggests that either genetic affinities among species may differ from taxonomic boundaries or that species-level phylogenetic relationships should be reconsidered [20]. Because iPBS markers specifically target repetitive DNA regions, they can reflect recent evolutionary separations. Therefore, set of evolutionary conserved merker set would offer a offer significant complementary dat set in systematic and phylogenetic analyses [14].

4.2. Morphological Characterization and Phenotyping

4.2.1. Efficiency of Different Traits in Sperately of Subgenus

4.2.1.1. Leaf Traits

In this study, leaf morphologies of species belonging to the genus *Onobrychis* were measured digitally and evaluated based on basic morphological traits such as area, length, width and aspect ratio (AR) (Table 4.3). Traits such as leaflet shape, leaflet apex and leaflet base were determined manually for each sample (Table 4.3).

Leaflet morphology, particularly leaflet anatomical details, has frequently been emphasized in studies as supporting characters for taxonomic distinction in plants [115]. Developing a leaflet shape scale in the genus *Onobrychis* is particularly important because it offers the opportunity to objectively classify morphological variation. So a novel leaflet shape scale was developed to more systematically assess the leaflet morphology of *Onobrychis* species. Scale intervals were determined based on circularity and height-to-width ratio (Aspect Ratio, AR) values obtained using ImageJ software [17]. Analyses revealed that leaflet shapes were reliably classified based on these phenotypic measurements (Table 4.2, Figure 4.6). This approach provides an objective criterion for describing leaf morphology and making interspecies comparisons, particularly in *Onobrychis* taxa with high morphological diversity. It is anticipated that this scale will also contribute to future phylogenetic, taxonomic, and morphological studies.

Table 4.2. Leaflet shape classification.

Leaflet Shape Name	AR	Circularity
Linear	> 5.5	≤ 0.50
Linear-oblong	3.6-5.5	0.5-0.55
Oblong	2.6-3.5	0.55-0.6
Oblong-elliptic	2.1-2.5	0.6-0.65
Elliptic	1.7-2.2	0.65-0.75
Elliptic-ovate	1.4-1.6	0.75-0.80
Ovate	1.1-1.3	0.8-0.85

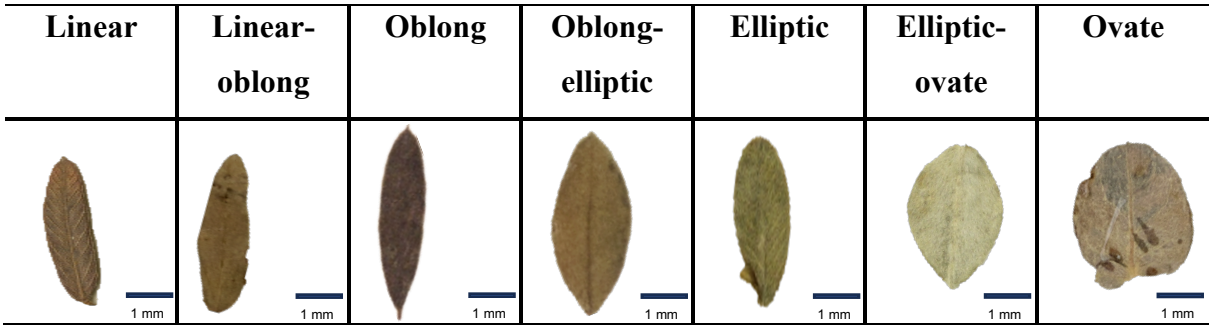


Figure 4.6. Leaflet shapes.

Correlation analysis between leaf morphological characters was conducted to identify variation patterns that may contribute to taxonomic differentiation among species. The analysis results revealed high correlations between some morphometric leaflet parameters, such as circularity, roundness, and aspect ratio. This suggests that geometric characters that determine the overall leaflet shape are strongly interrelated and represent structural integrity. In contrast, weaker correlations were found between characters reflecting surface integrity, such as solidity, and area, length, and width. These findings suggest that shape and size constitute partially independent axes of variation in leaf morphology, and therefore, considering these traits together can contribute to a more accurate interpretation of taxonomic distinctions. (Figure 4.7).

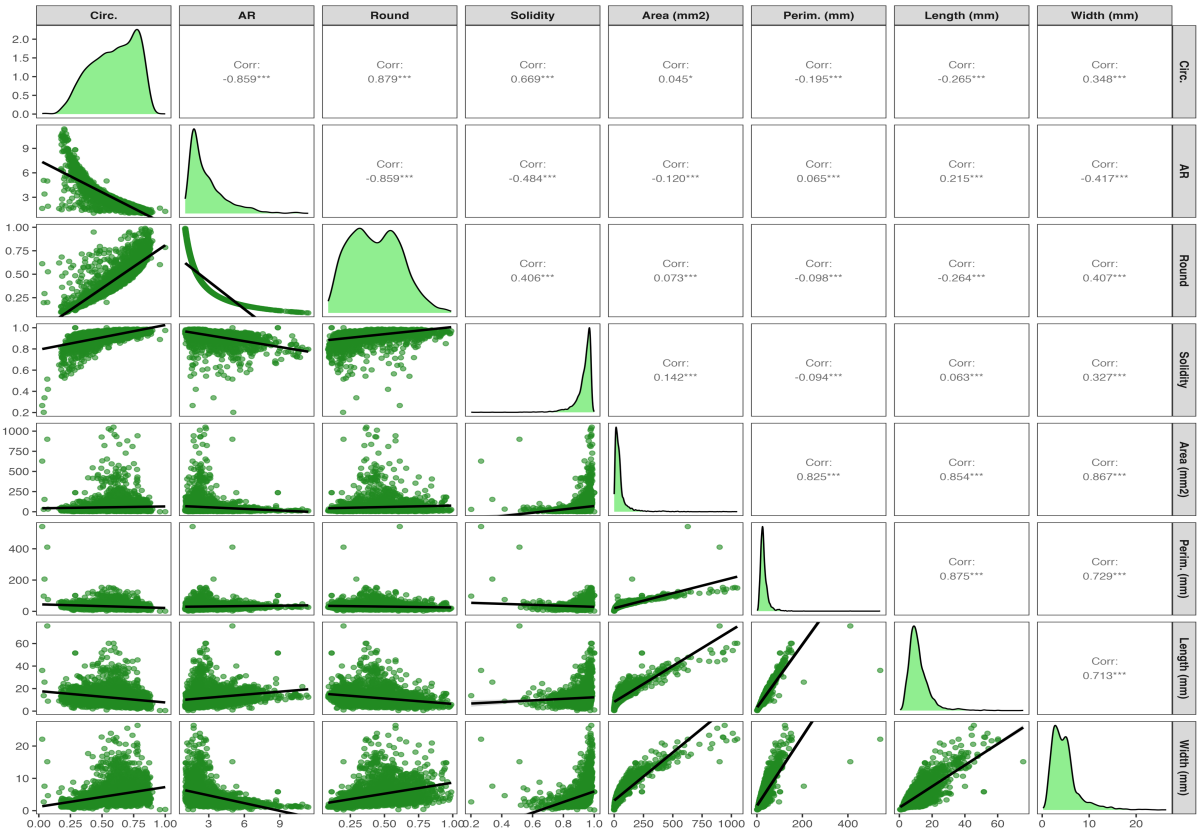


Figure 4.7. Pairwise scatter matrix among leaf morphological traits. Lower diagonal parts show scatter plots between pairs of variables; diagonal lines show its variable distribution characteristics. Upper diagonal panels show Pearson distribution coefficients (r) and significance levels ($p < 0.05$ *, $p < 0.01$ **, $p < 0.001$ ***).

Principal Component Analysis (PCA) based on quantified leaf morphological characteristics was used to visualize the differentiation of species. The first two components explained 93.1% of the total variance (PC1: 56.8%; PC2: 36.3%), indicating that leaf characteristics have a strong directional variation (Figure 4.8 and Figure 4.9). The subgenus Onobrychis (red) showed a more homogeneous structure in terms of leaf morphology, exhibiting a narrower and more dense clustering in the PCA graph (Figure 4.8). In contrast, the subgenus Sisyrosema (blue) showed a wider distribution; species such as O. saravschanica, O. galegifolia, and O. numularia were particularly distinct from others on the PC axes and showed a higher level of structural diversity within this subgenus. In addition to this distinction at the subgenus level, significant differences were also observed at the section level. For example, species in the *Hymenobrychis* section overlapped significantly and showed similar diversity in leaf morphology. In contrast, the wider distribution of species in the *Heliobrychis* section along the axes indicates greater structural diversity in this group. Morphological differentiation was more pronounced in sections with limited species such as *Lophobrychis* and *Dendobrychis* (Figure 4.9). Species such as *O. galegifolia* and *O. saravschanica*, particularly their distant positions in the PCA plane, indicate that these species possess unique leaf morphological characteristics. These findings indicate that leaf morphology is an important indicator for determining taxonomic boundaries within the genus *Onobrychis*, both at the subgenus and section levels. This demonstrates that leaf morphology provides valuable characters for taxonomy, as Viscosini pointed out [116]. Although classic taxonomic classifications appear largely consistent with morphological data, the morphological variation of some species suggests further examination of some species.

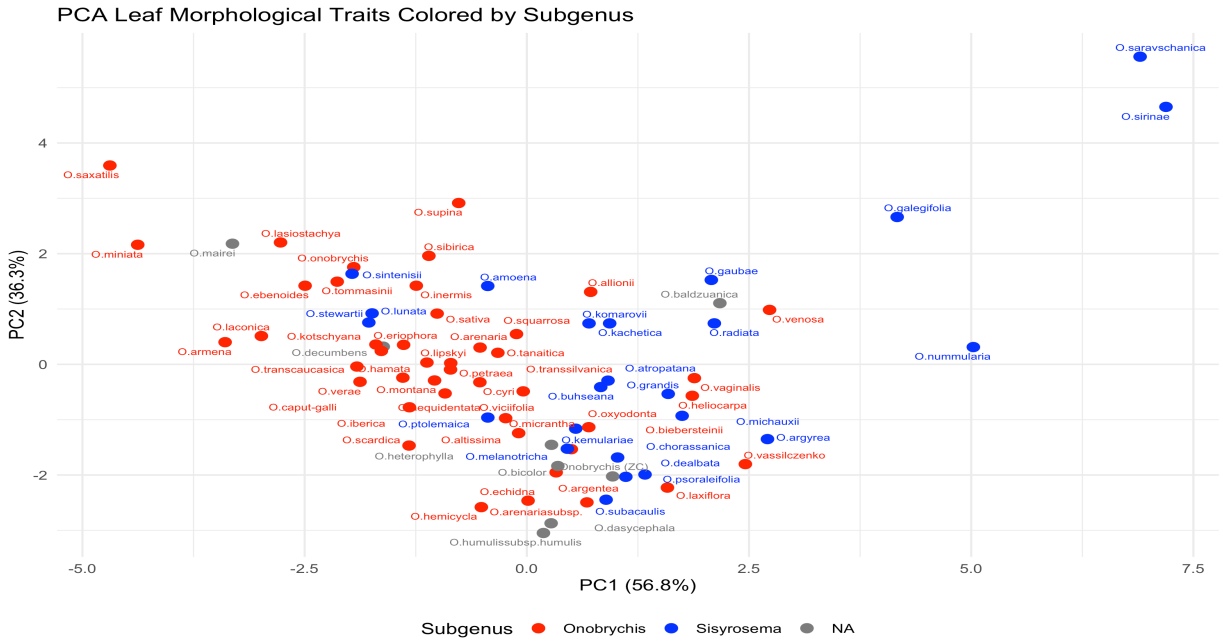


Figure 4.8. PCA analysis based on leaf morphological traits of *Onobrychis* species by subgenus.

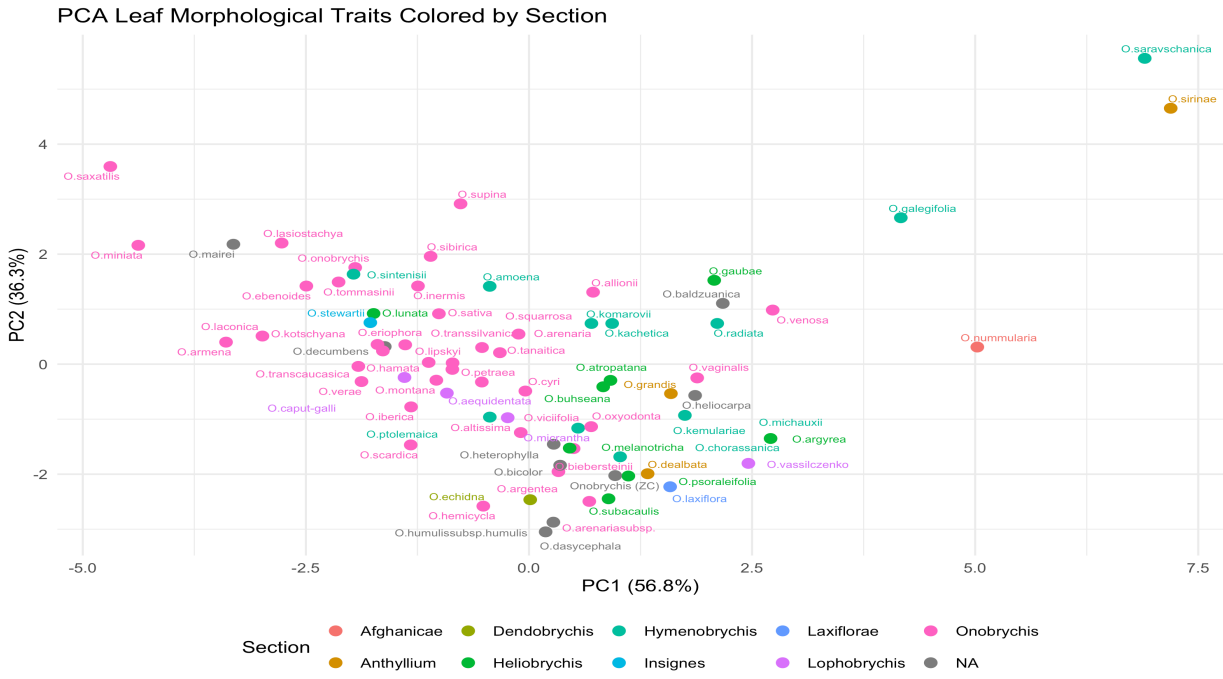


Figure 4.9. PCA analysis based on leaf morphological traits of *Onobrychis* species by sections.

Hierarchical clustering analysis using leaf morphological traits of species in the genus *Onobrychis* revealed similarities and differences among species (Figure 4.10). The analysis standardized trait data averaged across species, and clustering was performed using the Ward.D2 method based on Euclidean distance. In the resulting dendrogram, species are colored according to two proposed subgenera in the literature: species belonging to the subgenus *Sisyrosema* are shown in red, and species belonging to the subgenus *Onobrychis* are shown in blue. The overall clustering structure indicates that species belonging to the subgenus *Onobrychis* are largely clustered under the same branches, exhibiting a high level of morphological similarity. In contrast, species belonging to the subgenus *Sisyrosema* are more dispersed, with some species positioned closer to the *Onobrychis* group. This suggests that the subgenus *Sisyrosema* is more heterogeneous in terms of leaf morphology, and that there is no clear distinction between subgenera based on leaf morphological traits measured or estimated here. Similarly, morphological studies by Azizi et al. (2024) reported that species of the subgenus *Sisyrosema* exhibit a wide range of variation in morphological characters [50]. This suggests that morphological characters in this subgenus may have been influenced by ecological adaptation processes rather than phylogenetics. In conclusion, leaf morphological data generally reflect a distinction at the subgenus level, but morphological variation among species is particularly high, more so within the subgenus *Sisyrosema*. These findings demonstrate that classifications based solely on morphology could be insufficient and that morphological data should be supplemented with molecular/genomic data to reliably establish interspecific phylogenetic relationships.

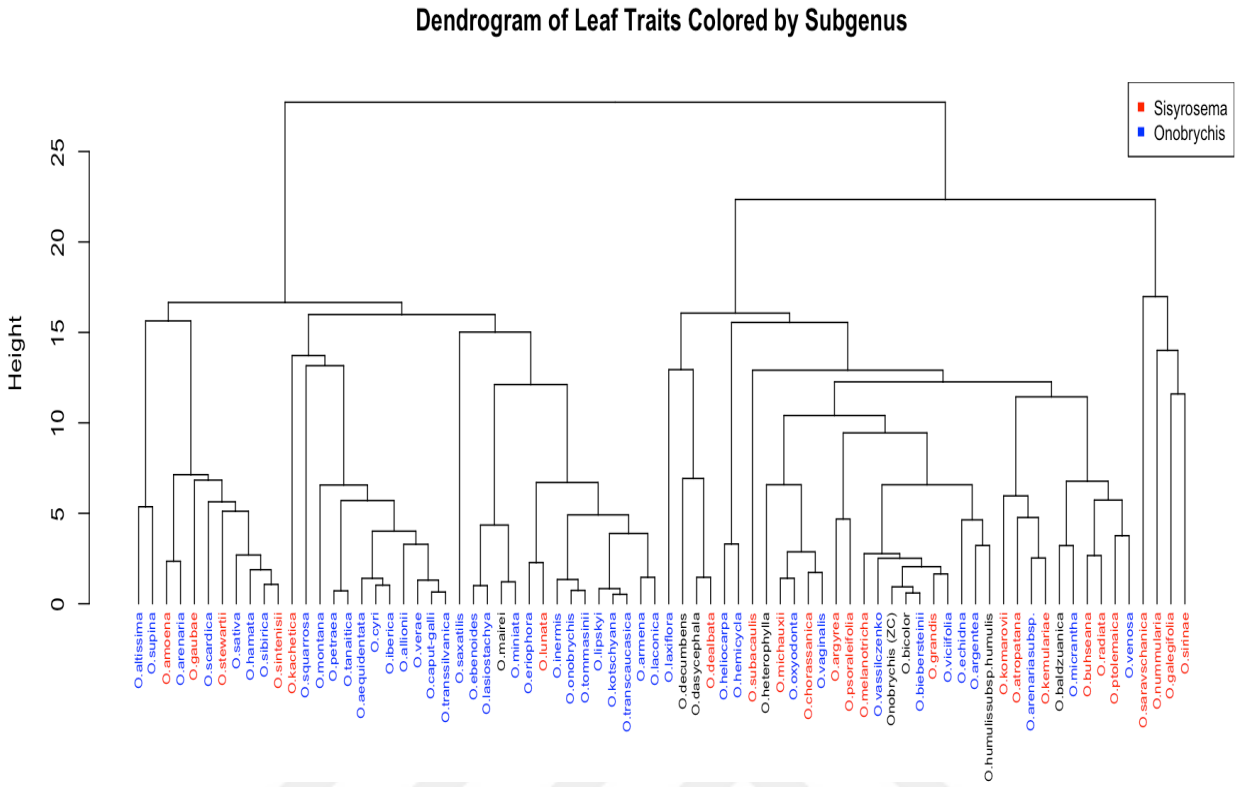


Figure 4.10. Hierarchical clustering dendrogram of leaf morphological traits from *Onobrychis* species colored by subgenus.

A hierarchical clustering dendrogram of the *Onobrychis* genus was color coded at the section level (Figure 4.11). In the resulting dendrogram, species were colored according to sections defined in the literature [4], [9], [10]. The overall clustering structure revealed that species from sections *Onobrychis* and *Hymenobrychis* were largely grouped under the same branches, and these sections exhibited a high degree of similarity in leaf morphology. Species from section *Heliobrychis* exhibited a broader distribution, indicating a more heterogeneous structure within this group. Furthermore, it was observed that sections with limited species, such as *Lophobrychis* and *Dendobrychis*, exhibited distinct species separation from other sections.

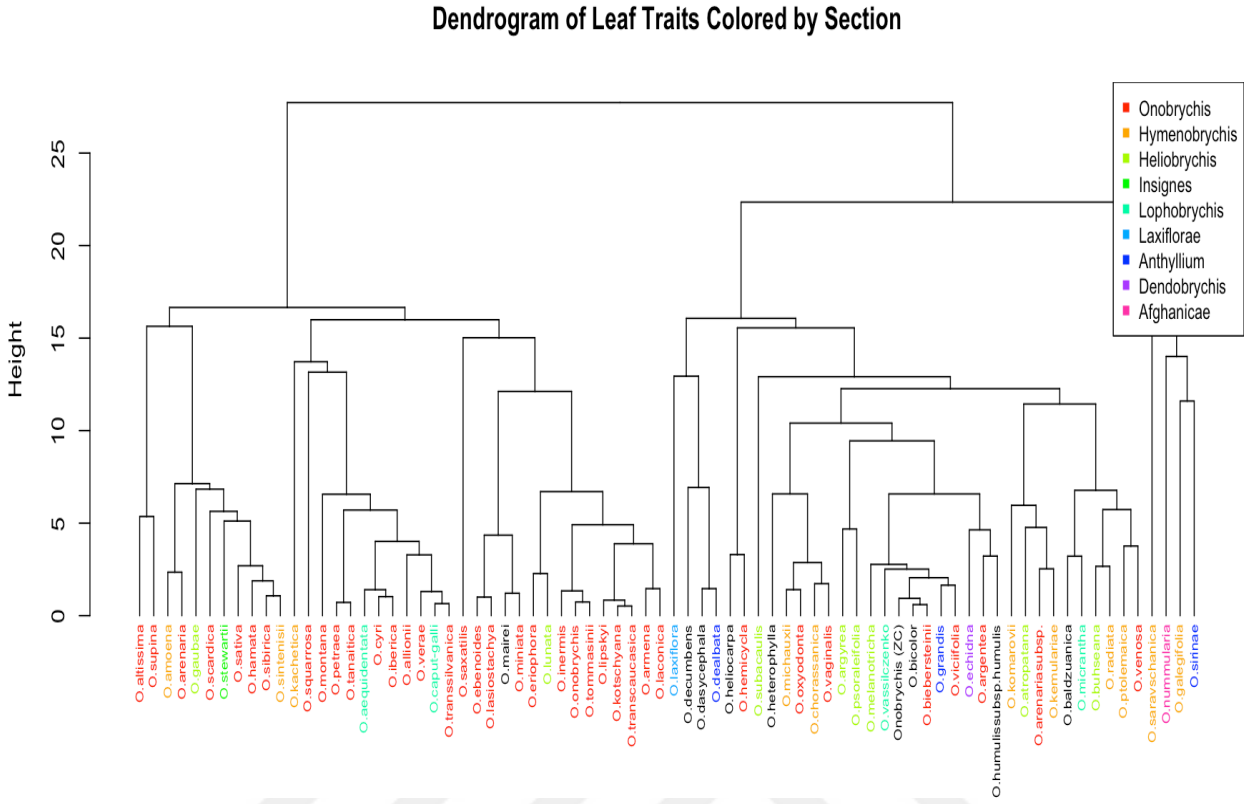


Figure 4.11. Hierarchical clustering dendrogram of leaf morphological traits from *Onobrychis* species colored by sections.

To evaluate the cluster structure of the species based on leaf morphological characteristics, the optimum cluster number was determined using three different methods (Figure 4.12). The Gap statistic (Figure 4.12A) indicated the optimum cluster number as K = 4. The Elbow method (Figure 4.12B) showed that the significant breakpoint in the total squared error (WSS) curve between clusters was in the range of K = 4–5. The Silhouette method (Figure 4.13C) revealed that the highest mean silhouette value was obtained at K = 8. When the results of the three different methods are evaluated together, it is seen that the optimum cluster number varies between 4 and 8 according to leaf morphological data. While the Gap and Elbow methods support lower cluster numbers (K = 4–5), the Silhouette method indicates a higher cluster number (K = 8). This shows that leaf morphology in the genus *Onobrychis* provides distinctive features both on a broad scale at the subgenus level and, in finer distinctions, at the section or species level. Therefore, although the optimum number of clusters varies depending on the

method used, it is concluded that leaf morphology exhibits a multidimensional variation and meaningful classifications can be made at different taxonomic levels.

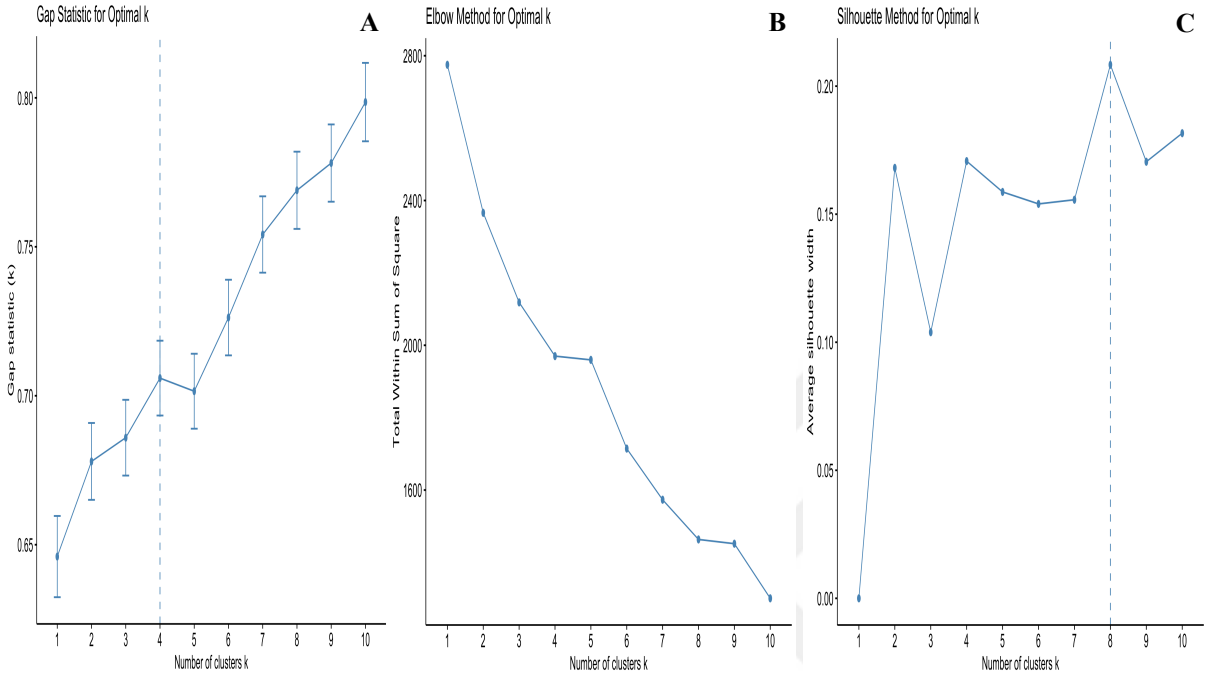


Figure 4.12. Determination of the optimum number of taxa clusters based on leaf morphological analysis using Gap (A), Elbow (B) and Silhouette (C) methods.

Analysis of leaf morphological trait data revealed that species belonging to the subgenus Sisyrosema exhibited a wider variation. In contrast, the subgenus *Onobrychis* showed a more homogeneous distribution. Similarly, hierarchical clustering analysis showed that species belonging to the subgenus *Onobrychis* were mostly clustered under the same branches, while the subgenus *Sisyrosema* exhibited a more heterogeneous and dispersed structure. This is largely consistent with previous morphological and molecular studies [7], [54], [117]. Clustering methods (Gap, Elbow, Silhouette) revealed that the optimal number of clusters could be at different values (K = 4–8). This difference suggests that leaf morphology can support both general discrimination at the subgenus level and finer discrimination at the section or species level. Similarly, Jensen (2003) [118] emphasized that leaf morphology carries a high level of taxonomic signal, particularly in the Fabaceae family, but that boundaries can become blurred in some groups due to environmental plasticity [116], [118].

The use of leaf morphology in taxonomy has long been debated. Jensen (2003) stated that morphometric traits are a powerful tool for taxonomic distinctions, but are also susceptible to plasticity due to environmental factors [118], [119]. The findings in this study also suggest that leaf morphology reflects distinctions at the subgenus and section levels in the *Onobrychis* genus, but that it must be supported by molecular data to reliably delineate interspecific boundaries.

In conclusion, this study demonstrates that leaf morphology is a strong indicator for taxonomic distinctions at the subgenus and section levels in the *Onobrychis* genus, but that morphological data must be supported by molecular analyses to reliably delineate interspecific boundaries.

Table 4.3. Morphological traits of *Onobrychis* leaflets in the study.

Species name	Circ.	AR	Round	Solidity	Area (mm^2)	Perim. (mm)	Length (mm)	Width (mm)	Leaflet Shape	Leaflet Apex	Leaflet Base	Indumentum absence (1/0)
O.aequidentata	0.54	2.61	0.45	0.89	28.83	25.95	9.32	3.69	Oblong	Acute	Cuneate	0
O.allionii	0.50	3.12	0.32	0.95	107.98	51.71	20.70	6.59	Oblong	Acute	Cuneate	1
O.altissima	0.66	2.38	0.46	0.94	31.04	24.06	9.17	4.18	Elliptic	Acuminate	Attenuate	0
O.amoena	0.43	3.87	0.26	0.94	68.07	44.54	18.30	4.74	Linear-Oblong	Acute	Cuneate	0
O.arenaria	0.52	3.67	0.36	0.94	46.21	34.10	13.64	4.39	Linear-Oblong	Acute	Cuneate	1
O.arenaria subsp. arenaria	0.78	1.62	0.63	0.96	24.63	19.92	7.10	4.40	Elliptic	Rounded-Obtuse	Rounded	0
O.argentea	0.72	1.84	0.55	0.95	26.70	21.66	7.88	4.32	Elliptic	Acute	Cuneate	0
O.argyrea	0.77	1.39	0.72	0.97	114.00	42.99	14.16	10.20	Elliptic-Ovate	Acute	Cuneate	1
O.armena	0.42	3.86	0.27	0.80	8.33	13.75	5.49	1.51	Linear-Oblong	Acute	Cuneate	1
O.atropatana	0.62	2.20	0.48	0.95	82.70	39.67	14.56	6.84	Oblong-Elliptic	Rounded-Obtuse	Rounded	0
O.baldzuanica	0.60	2.23	0.46	0.95	176.61	59.73	21.55	10.08	Oblong-Elliptic	Acute	Cuneate	1

Species name	Circ.	AR	Round	Solidity	Area (mm^2)	Perim. (mm)	Length (mm)	Width (mm)	Leaflet Shape	Leaflet Apex	Leaflet Base	Indumentum absence (1/0)
O.bicolor	0.70	1.85	0.59	0.94	31.15	23.38	8.27	4.61	Elliptic	Rounded	Rounded	0
O.biebersteinii	0.71	2.07	0.53	0.95	39.70	26.16	9.77	4.98	Elliptic	Rounded	Rounded	0
O.buhseana	0.62	2.07	0.49	0.94	75.86	38.65	13.88	6.76	Oblong-Elliptic	Rounded	Rounded	1
O.caput-galli	0.48	3.42	0.32	0.93	20.23	22.15	8.98	2.61	Oblong	Acute	Cuneate	1
O.chorassanica	0.73	1.70	0.60	0.95	56.78	28.84	9.93	6.18	Elliptic	Obtuse	Rounded	0
O.cyri	0.58	3.24	0.37	0.94	35.09	28.30	11.54	3.93	Oblong	Acute	Cuneate	0
O.dasycephala	0.75	1.43	0.73	0.93	13.39	14.90	4.86	3.50	Elliptic-Ovate	Rounded	Obtuse	0
O.dealbata	0.75	1.49	0.70	0.95	58.19	29.30	9.94	6.37	Elliptic-Ovate	Rounded	Obtuse	0
O.decumbens	0.44	3.61	0.28	0.91	23.77	26.04	10.39	2.90	Linear-Oblong	Rounded	Obtuse	0
O.ebenoides	0.37	5.53	0.23	0.88	24.08	28.56	12.05	2.50	Linear	Acute	Cuneate	0
O.echidna	0.74	1.70	0.59	0.94	14.25	15.42	5.52	3.25	Elliptic	Rounded	Cuneate	1
O.eriophora	0.47	3.62	0.30	0.91	29.26	28.31	11.32	3.31	Linear-Oblong	Obtuse	Cuneate	0

Species name	Circ.	AR	Round	Solidity	Area (mm^2)	Perim. (mm)	Length (mm)	Width (mm)	Leaflet Shape	Leaflet Apex	Leaflet Base	Indumentum absence (1/0)
O.galegifolia	0.61	2.16	0.48	0.96	345.64	83.23	30.08	14.12	Oblong-Elliptic	Rounded-Obtuse	Rounded	0
O.gaubae	0.60	2.43	0.43	0.93	229.08	60.46	23.12	9.07	Oblong-Elliptic	Acute	Cuneate	0
O.grandis	0.71	2.11	0.51	0.96	119.14	40.98	15.49	7.23	Elliptic	Rounded	Rounded	0
O.hamata	0.47	3.56	0.30	0.90	23.47	25.39	10.17	2.91	Linear-Oblong	Acute	Cuneate	0
O.heliocarpa	0.64	1.60	0.66	0.94	110.53	45.40	14.35	9.49	Elliptic	Rounded	Rounded	1
O.hemicycla	0.72	1.76	0.60	0.92	6.66	10.69	3.78	2.22	Elliptic	Rounded	Rounded	1
O.heterophylla	0.67	1.88	0.55	0.93	38.81	25.83	9.04	5.09	Elliptic	Obtuse	Rounded	1
O.humulis **subsp. *humulis***	0.81	1.59	0.64	0.95	9.15	11.84	4.25	2.68	Elliptic	Acute	Rounded	0
O.iberica	0.55	3.01	0.34	0.93	14.27	18.09	7.33	2.47	Oblong	Acute	Cuneate	0
O.inermis	0.40	4.56	0.23	0.93	45.89	38.12	16.16	3.63	Linear-Oblong	Acute	Cuneate	0
O.kachetica	0.57	2.73	0.38	0.95	110.81	47.96	18.75	7.03	Oblong	Rounded-Obtuse	Rounded-Subcuneate	0

Species name	Circ.	AR	Round	Solidity	Area (mm^2)	Perim. (mm)	Length (mm)	Width (mm)	Leaflet Shape	Leaflet Apex	Leaflet Base	Indumentum absence (1/0)
O.kemulariae	0.69	2.26	0.45	0.97	45.45	28.72	11.38	5.07	Elliptic	Rounded-Obtuse	Rounded	1
O.komarovii	0.56	2.91	0.39	0.94	99.25	46.46	18.07	6.91	Oblong	Rounded-Obtuse	Cuneate	0
O.kotschyana	0.45	4.15	0.27	0.92	22.01	24.78	10.36	2.71	Linear-Oblong	Acute	Cuneate	1
O.laconica	0.40	4.42	0.24	0.85	8.87	16.74	6.89	1.61	Linear-Oblong	Acute	Cuneate	1
O.lasiostachya	0.31	6.52	0.17	0.89	26.55	32.79	14.39	2.32	Linear	Acute	Cuneate	0
O.laxiflora	0.78	1.36	0.75	0.95	53.99	29.26	9.51	7.13	Elliptic-Ovate	Rounded	Rounded-Obtuse	1
O.lipskyi	0.50	3.57	0.31	0.93	27.95	26.78	11.01	3.19	Linear-Oblong	Acute	Cuneate	1
O.lunata	0.44	3.74	0.28	0.86	37.36	31.27	12.52	3.58	Linear-Oblong	Obtuse	Cuneate	1
O.mairei	0.31	5.56	0.19	0.81	43.29	34.60	10.92	2.07	Linear	Acute	Cuneate	1
O.melanotricha	0.69	1.86	0.56	0.94	47.15	26.12	9.13	5.26	Elliptic	Rounded	Rounded	1
O.michauxii	0.70	1.75	0.60	0.96	92.52	40.13	13.95	8.30	Elliptic	Obtuse	Rounded	1

Species name	Circ.	AR	Round	Solidity	Area (mm^2)	Perim. (mm)	Length (mm)	Width (mm)	Leaflet Shape	Leaflet Apex	Leaflet Base	Indumentum absence (1/0)
O.micrantha	0.62	2.48	0.46	0.93	34.51	25.38	9.48	4.44	Oblong-Elliptic	Acute	Cuneate	1
O.miniata	0.27	6.23	0.17	0.78	10.11	21.70	8.80	1.44	Linear	Acute	Cuneate	1
O.montana	0.52	3.06	0.35	0.92	27.43	25.92	9.66	3.23	Oblong	Acute	Attenuate	1
O.nummularia	0.78	1.28	0.79	0.97	308.20	67.70	21.36	17.00	Ovate	Acute	Attenuate	1
O.Onobrychis	0.38	4.94	0.24	0.87	40.47	37.08	15.37	3.43	Linear-Oblong	Acute	Cuneate	0
O.oxyodonta	0.69	2.21	0.47	0.97	50.43	30.15	11.84	5.33	Elliptic	Obtuse	Rounded	1
O.petraea	0.53	3.27	0.33	0.93	33.43	27.81	11.21	3.64	Oblong	Acute	Rounded	0
O.psoraleifolia	0.74	1.53	0.66	0.95	45.90	27.23	9.20	6.04	Elliptic-Ovate	Rounded	Rounded	1
O.ptolemaica	0.61	2.47	0.42	0.93	29.18	23.87	9.23	3.78	Oblong-Elliptic	Obtuse	Rounded	1
O.radiata	0.59	2.01	0.52	0.94	155.67	57.84	19.73	10.09	Oblong-Elliptic	Rounded	Rounded	0
O.saravschanica	0.61	2.42	0.42	0.97	675.49	116.43	44.73	18.41	Oblong-Elliptic	Obtuse-Rounded	Rounded-Cuneate	0

Species name	Circ.	AR	Round	Solidity	Area (mm^2)	Perim. (mm)	Length (mm)	Width (mm)	Leaflet Shape	Leaflet Apex	Leaflet Base	Indumentum absence (1/0)
O.sativa	0.45	3.91	0.27	0.92	46.38	35.61	14.67	3.91	Linear-Oblong	Acute	Cuneate	1
O.saxatilis	0.22	9.51	0.11	0.81	16.04	29.98	13.70	1.48	Linear	Acute - Acuminate	Cuneate	0
O.scardica	0.61	2.64	0.44	0.91	8.42	13.53	5.00	1.92	Elliptic	Acute	Cuneate	1
O.sibirica	0.40	4.35	0.25	0.89	67.65	49.92	17.07	4.46	Linear-Oblong	Acute	Cuneate	0
O.sintenisii	0.38	4.89	0.22	0.88	46.55	34.53	14.36	3.37	Linear-Oblong	Acute	Cuneate	0
O.sirinae	0.68	2.02	0.51	0.97	683.10	107.14	39.55	20.56	Elliptic	Rounded	Rounded	0
O.squarrosa	0.51	3.19	0.32	0.95	61.87	39.01	15.71	4.97	Oblong	Acute	Attenuate	0
O.stewartii	0.42	4.12	0.25	0.91	27.12	28.21	11.75	2.86	Linear-Oblong	Acute	Cuneate	1
O.subacaulis	0.77	1.54	0.68	0.95	32.40	22.18	7.62	5.07	Elliptic	Rounded	Rounded-Subcordate	1
O.supina	0.38	4.96	0.24	0.88	95.55	57.39	23.60	5.22	Linear-Oblong	Acuminate	Attenuate	0
O.tanaitica	0.54	3.19	0.32	0.95	49.59	33.92	14.06	4.46	Oblong	Acute	Rounded	0

Species name	Circ.	AR	Round	Solidity	Area (mm^2)	Perim. (mm)	Length (mm)	Width (mm)	Leaflet Shape	Leaflet Apex	Leaflet Base	Indumentum absence (1/0)
O.tommasinii	0.36	4.94	0.21	0.90	29.76	32.24	13.60	2.77	Linear-Oblong	Acute	Cuneate	0
O.transcaucasica	0.47	3.90	0.28	0.91	14.44	19.46	8.12	2.22	Linear-Oblong	Acute	Cuneate	1
O.transsilvanica	0.51	3.29	0.32	0.94	33.54	28.79	11.60	3.63	Oblong	Acute	Cuneate	1
O.vaginalis	0.68	1.91	0.53	0.96	116.57	46.28	16.70	8.80	Elliptic	Obtuse	Rounded	0
O.vassilczenko	0.83	1.43	0.71	0.98	94.94	37.16	12.86	9.00	Elliptic	Rounded	Rounded	0
O.venosa	0.64	2.04	0.50	0.96	196.16	61.78	22.20	11.08	Oblong-Elliptic	Obtuse	Rounded	1
O.verae	0.50	3.42	0.32	0.90	13.37	18.22	7.34	2.28	Oblong	Acute	Cuneate	1
O.viciifolia	0.61	3.07	0.45	0.94	45.07	31.15	12.22	4.67	Elliptic	Rounded	Rounded	0
***Onobrychis* (ZC)**	0.78	1.77	0.57	0.97	39.58	25.23	9.39	5.31	Elliptic	Rounded	Rounded	0

4.2.1.2. Fruit Traits

In the present study, a number of fruit related morphometric measurements were recorded and used in downstream analyses. All the fruit morphological traits were averaged at the species level using sample-based values for each species and a morphological profile was obtained for each species (Table.4.4).

Correlation analysis revealed a strong negative correlation between circularity and AR among fruit morphological traits, indicating that circularity decreases as the fruit elongates, indicating a form change. A high positive correlation was found between Roundness, AR, and Circularity. Strong positive correlations were found between the dimensional parameters Area, Perimeter, Length, and Width. Fruit spine number also showed positive correlations with fruit size parameters such as Area, Perimeter, Width, and Length. This suggests that spine number may increase with fruit size. Furthermore, weak negative correlations ($r \approx -0.22$) with Circularity and AR suggest that spines may affect fruit shape (Figure 4.13).

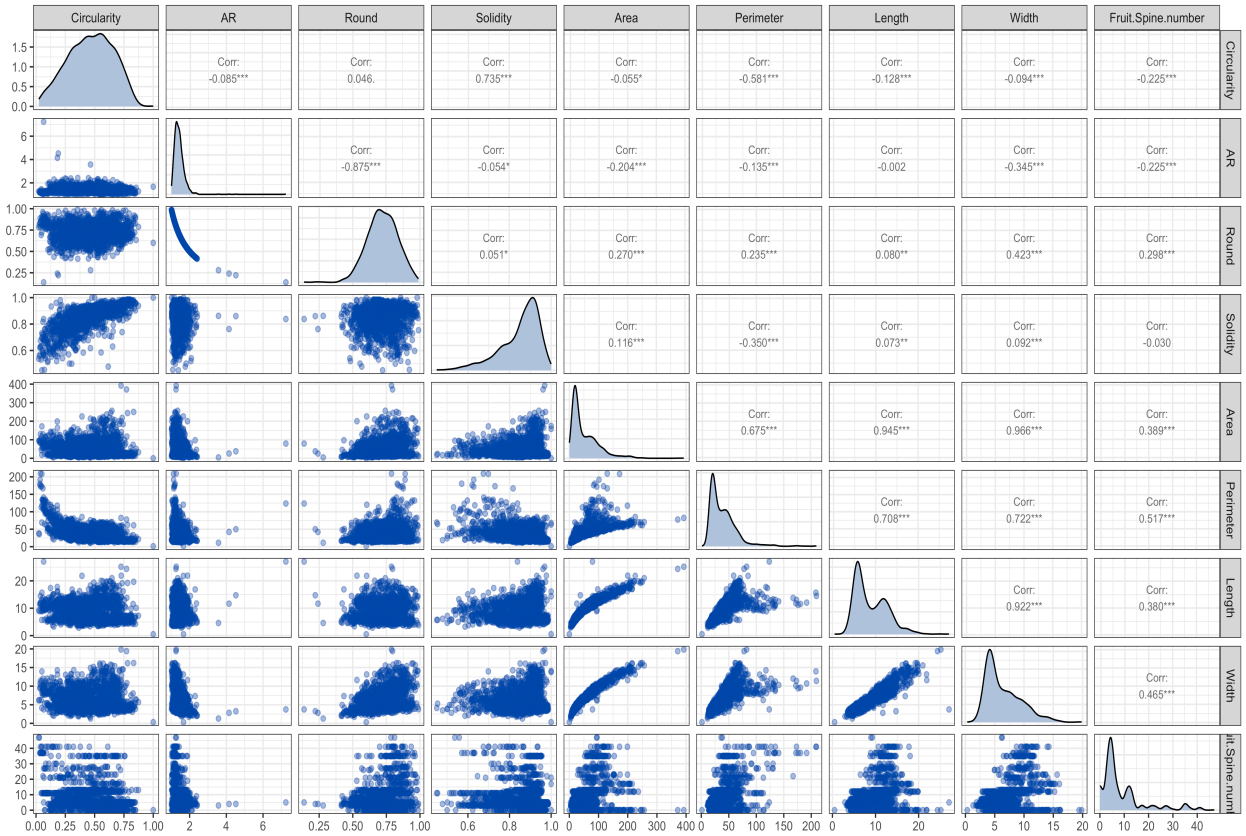


Figure 4.13. Correlation matrix between variables of fruit morphological traits. Shows Pearson correlation coefficients and significance levels ($p < 0.05$ *, $p < 0.01$ **, $p < 0.001$ ***).

Principal Component Analysis (PCA) based on fruit morphological traits of species was performed to evaluate interspecific variation and to reveal differences at the taxonomic level. As a result of the PCA analysis, the first two main components (PC1 and PC2) explain approximately 54.2% of the total variance (PC1: 34.7%, PC2: 19.5%). This ratio shows that the the first two PCA could be good tools to test distinction among taxa based on the selected fruit morphology characters. In the PCA graph color coded based on subgenera, it is evident that there is a morphological differentiation between *Onobrychis* and *Sisyrosema* (Figure 4.14). Species belonging to the *Sisyrosema* subgenus are generally grouped on the left side of the graph by positioning them at more negative values along PC1. This indicates that species belonging to this subgenus may have structural characteristics such as smaller fruit area and lower perimeter. Species belonging to the *Onobrychis* subgenus, although showing a wider distribution, are positioned on the positive PC1 axis. This situation shows that the taxonomic

distinction at the subgenus level is with fruit morphology. Similar results were obtained to previous studies on fruits belonging to taxa in the *Onobrychis* genus [6], [49].

When the taxonomic distinction is examined at the section level, it is observed that the species have a more complex structure on the PCA plane but still exhibit partial separations (Figure 4.15). Especially species belonging to sect. Lophobrychis, sect. *Heliobrychis* and sect. *Hymenobrychis* are grouped on the PCA plane. Species belonging to *Heliobrychis* such as O. melanotricha, *O. aucheri* are located at more extreme positions on the PC1 and PC2 axes and attract attention with their fruit morphologies that are more different from other species. Sect. *Hymenobrychis* is generally seen to have distinguishable fruit morphology by being positioned at more negative PC2 values. However, the spread of species belonging to *Onobrychis* over a wide PCA area shows that this section is a morphologically heterogeneous group (Figure 4.15).

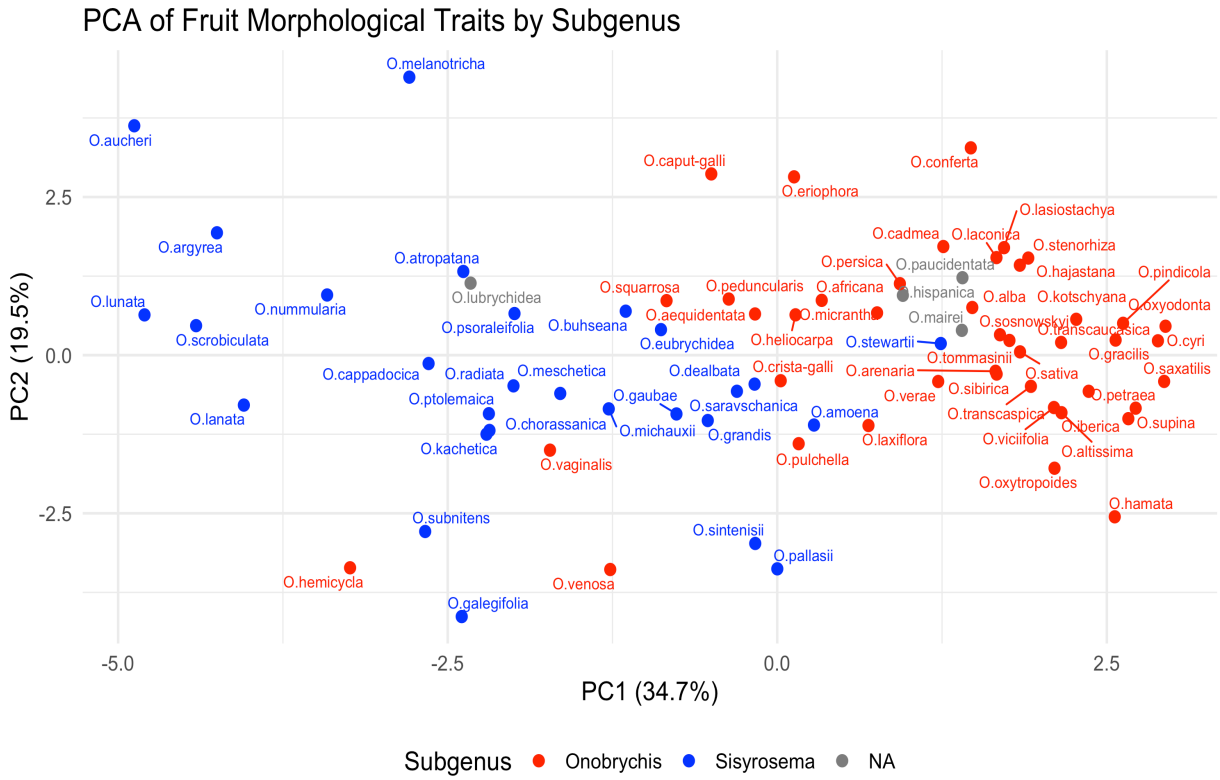


Figure 4.14. PCA analysis based on mean fruit morphological traits of *Onobrychis* species by subgenus.

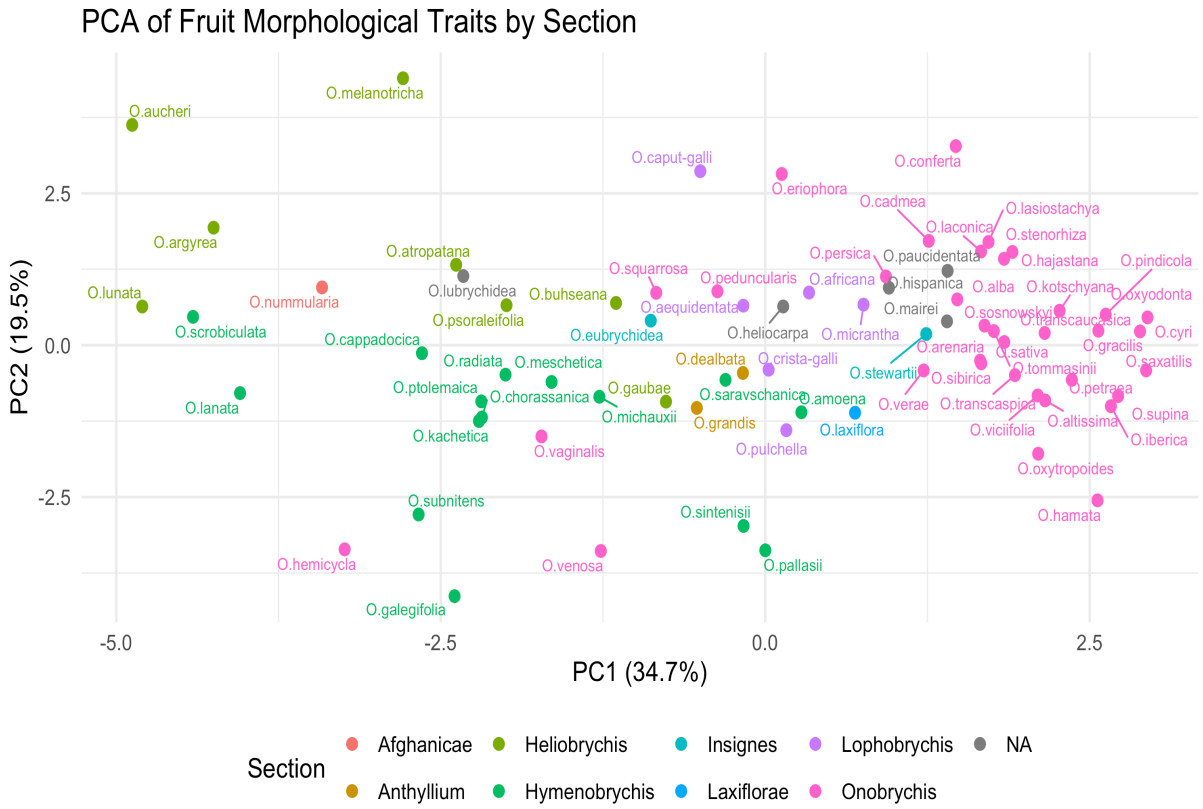


Figure 4.15. PCA analysis based on mean fruit morphological traits of *Onobrychis* species by section.

Elbow, Silhouette, and Gap methods were applied to determine the appropriate cluster number based on fruit morphological characteristics of *Onobrychis* species (Figure 4.16). According to the Silhouette method, the highest value was obtained at K=3. Similarly, the Elbow method showed a clear breakpoint in the range of K=3–4. The Gap statistic indicated the optimum cluster number as K=2. When these results are evaluated together, it is clear that natural segregation in the data occurs particularly at K=2 and K=3. Considering that the *Onobrychis* genus is divided into two subgenera, *Onobrychis* and *Sisyrosema* in the literature [4], [6], [9], the K=2 value obtained in this study was found to be biologically more significant and was accepted as the optimum value.

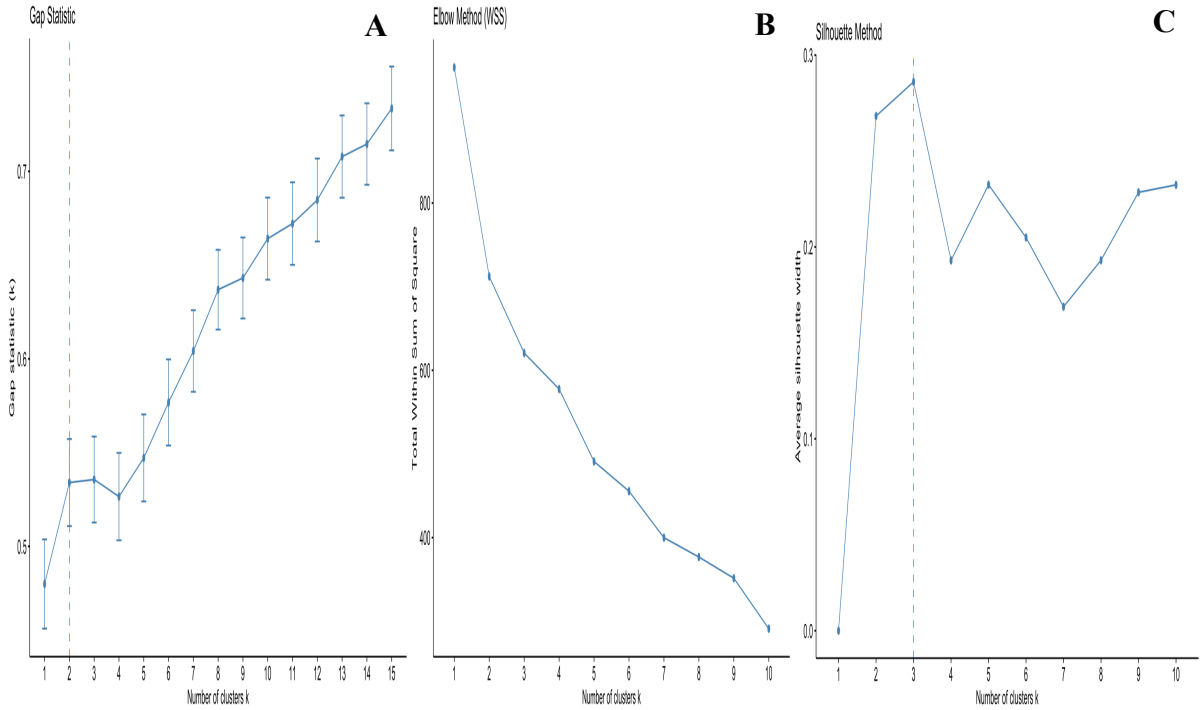


Figure 4.16. Determination of the optimum number of taxa clusters based on fruit morphological analysis using Gap (A), Elbow (B) and Silhouette (C) methods.

The dendrogram obtained as a result of hierarchical clustering analysis based on fruit morphological traits of species belonging to the genus *Onobrychis* was evaluated in order to reveal similarities and differences between species at the subgenus level (Figure 4.17). Species names on the dendrogram are shown in colors according to the subgenus they belong to (*Onobrychis*: red, *Sisyrosema*: blue). In general, most of the species are divided into similar clusters in accordance with the subgenus they belong to. Species belonging to the *Sisyrosema* subgenus in particular are located on the right side of the dendrogram, which shows a more homogeneous structure. This suggests that the *Sisyrosema* subgenus has a more consistent fruit morphology. Similar results have been reported in morphological studies on *Sisyrosema* species originating from Iran [50]. On the other hand, species belonging to *Onobrychis* showed a wider distribution and were grouped with *Sisyrosema* members in some cases. This situation reveals that morphological characters have a certain distinctiveness at the subgenus level, and the fact

that some species do not fully comply with the current taxonomic classifications indicates morphological variation. It also points out the need to re-evaluate the current taxonomic classification. In conclusion, the data on fruit morphology largely support the taxonomic structure of the *Onobrychis* genus at the subgenus level, but the deviations observed in some species emphasize the necessity of a holistic evaluation of these characters together with molecular data.

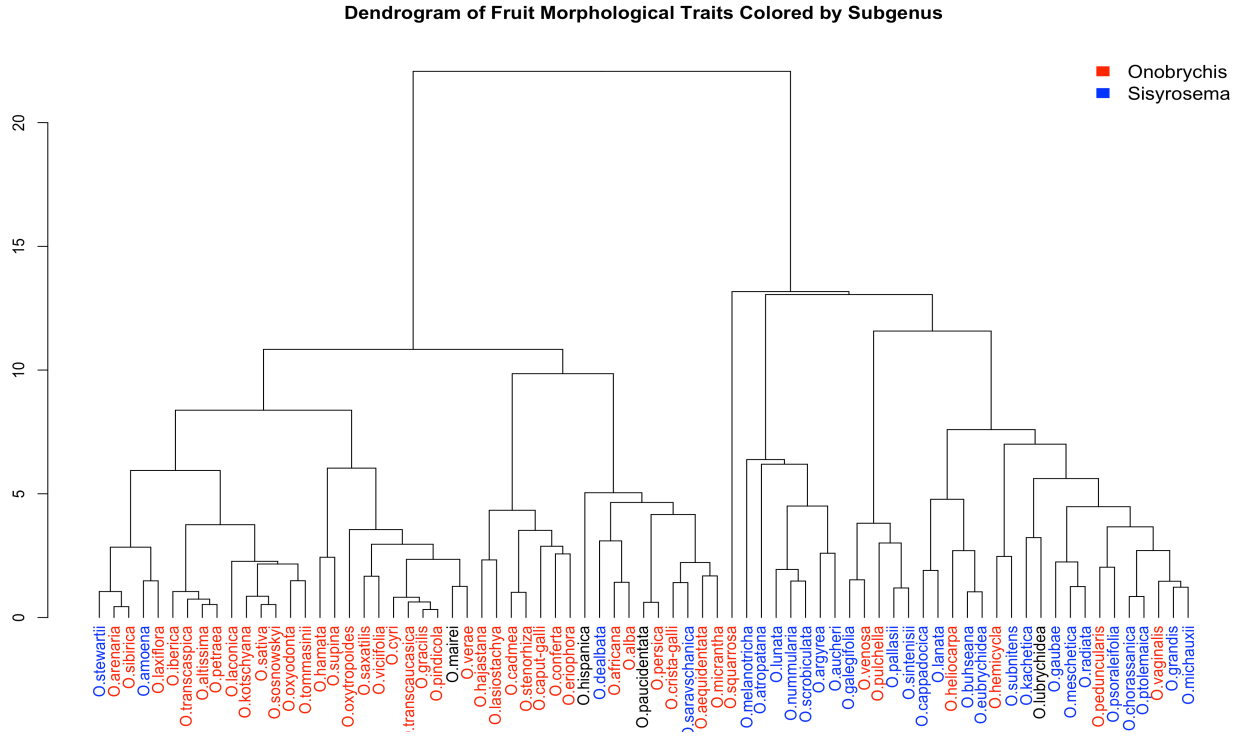


Figure 4.17. Dendrogram according to fruit traits colored by subgenus.

When the same analysis was color coded based in the known sections to deduce if the hierarchical cluster analysis using fruit morphological traits among species of the genus *Onobrychis* would follow the current taxonomic distinction at the section level (Figure 4.18). The analysis resulted in a large number of species clustering according to sections defined in the literature. Species from sections *Onobrychis* and *Hymenobrychis*, in particular, concentrated on the same branches, demonstrating their morphological similarities. In contrast, sections such as *Heliobrychis* and *Lophobrychis* exhibited a more scattered distribution, with some species

located close to other section groups. Species such as *O. stenostachya*, *O. caput-galli*, and *O. radiata* differed from other species in the dendrogram, as seen in PCA analysis, demonstrating their distinctive fruit characters. Furthermore, the separation of a limited number of species from sections *Anthyllium* and *Afghanicae* from other groups suggests that these groups are defined by more distinct morphological characters. In conclusion, dendrogram analysis based on fruit morphology was found to largely support the taxonomic classification at the section level in the genus *Onobrychis*.

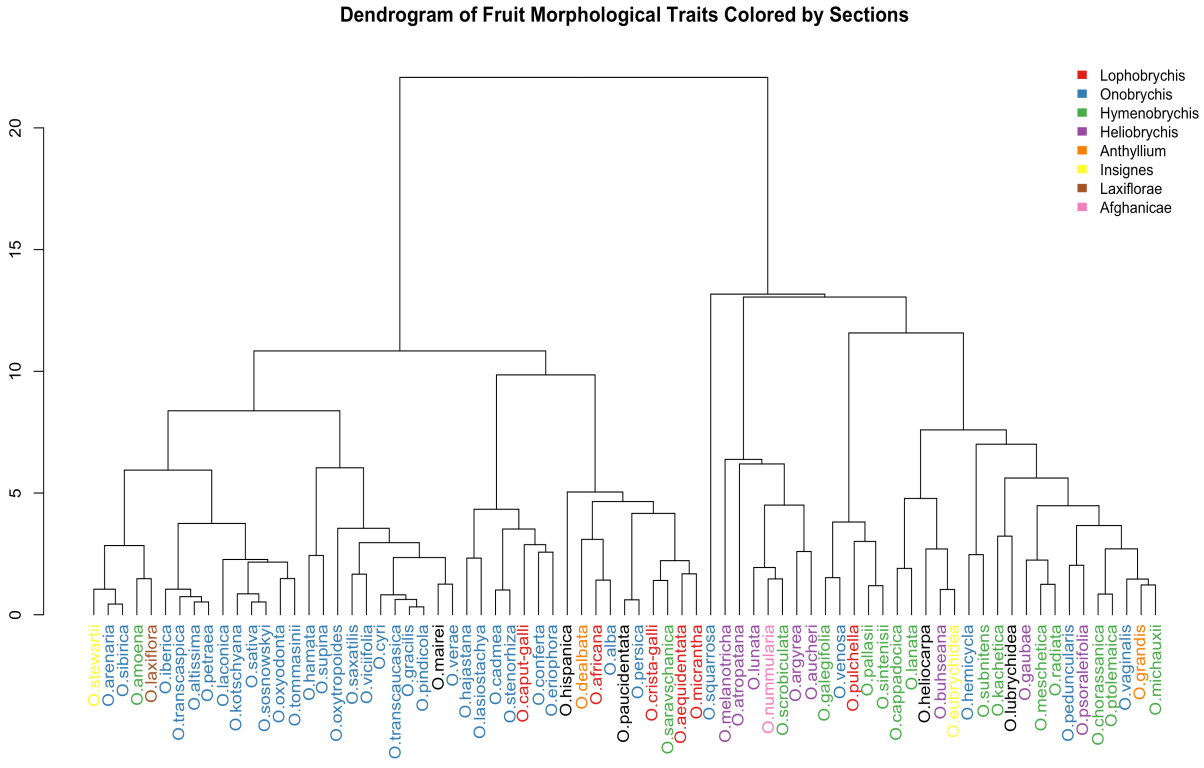


Figure 4.18. Dendrogram according to fruit traits colored by sections.

In the fruit morphology-based dendrogram, the distinctions at the section level are generally similar to the classical classifications in the literature [4]. However, in some sections, such as *Heliobrychis* and *Lophobrychis*, species are distributed intermixed with other groups. This suggests that, as emphasized in Yildiz et al. (1999) study on *Onobrychis* fruit morphology, fruit

characters are distinctive in some taxonomic groups, but environmental influences and adaptive traits can blur the boundaries between species [6]. Macro- and micromorphological analyses by Zarrabian et al. (2025) demonstrated that fruit morphology carries a taxonomically significant signal, but that convergent character evolution in some species can complicate morphological distinction [120]. Similarly, in another study emphasized that Onobrychis species originating from Pakistan exhibit distinct differences in fruit shape and surface characteristics, but that these differences are also related to the species' geographical distribution and habitat conditions [121]. This findings show explain why some species deviate from their classical taxonomic positions in the present study and are placed in different clusters. When PCA, clustering, and dendrogram analyses are evaluated together, it is clear that there is distinction at the subgenus level within the *Onobrychis* genus, but there is a high level of variation, particularly within the subgenus *Sisyrosema*. These findings demonstrate that morphology could be a powerful tool in taxonomic classification, but it must be obtained through accurate estimation of the traits using image analyses techniques [117], [118].

Table 4.4. Morphological traits of *Onobrychis* fruits in the study.

Species name	Circ.	AR	Round	Solidity	Area (mm^2)	Perim. (mm)	Length (mm)	Width (mm)	Fruit Shape	Fruit Spine	Fruit Trichome	Crest of Fruit	Fruit Vein	Spine on Spine	Fruit Spine Number
O.aeuquantative	0.27	1.44	0.70	0.74	66.20	54.34	10.71	7.54	Reniform	1	0	1	1	0	7
O.africana	0.29	1.45	0.70	0.77	43.38	43.66	8.87	6.17	Reniform	1	1	1	1	0	3
O.alba	0.37	1.46	0.69	0.81	23.28	28.79	6.46	4.47	Reniform	1	1	1	1	0	3
O.altissima	0.64	1.45	0.70	0.92	22.44	21.29	6.30	4.36	Elliptic	1	0	0	1	0	5
O.amonea	0.68	1.40	0.72	0.91	46.58	29.84	9.05	6.47	Semicircular	1	1	0	1	0	13
O.arenaria	0.57	1.43	0.70	0.90	19.36	20.77	5.91	4.16	Elliptic	1	1	0	1	0	4
O.argyrea	0.15	1.17	0.85	0.76	101.97	96.08	12.31	10.48	Reniform	1	1	0	0	0	30
O.atropatana	0.11	1.26	0.80	0.72	70.82	93.12	10.62	8.47	Reniform	0	1	0	0	0	0
O.aucheri	0.06	1.16	0.87	0.65	80.56	140.86	10.84	9.39	Reniform	1	1	0	0	0	37
O.buhseana	0.49	1.39	0.74	0.86	58.13	40.21	9.69	7.36	Elliptic	1	1	0	0	0	26
O.cadmea	0.29	1.61	0.64	0.80	25.73	33.95	7.13	4.59	Reniform	1	0	0	0	0	6
O.cappodocica	0.51	1.17	0.86	0.90	92.69	48.32	11.63	9.97	Elliptic	1	1	0	0	0	28
O.caput-galli	0.18	1.26	0.80	0.66	38.45	58.89	7.68	6.11	Reniform	1	0	0	0	0	11
O.chorossanica	0.47	1.26	0.80	0.89	103.38	53.55	12.76	10.19	Elliptic	1	1	0	1	0	13

Species name	Circ.	AR	Round	Solidity	Area (mm^2)	Perim. (mm)	Length (mm)	Width (mm)	Fruit Shape	Fruit Spine	Fruit Trichome	Crest of Fruit	Fruit Vein	Spine on Spine	Fruit Spine Number
O.conferta	0.19	1.64	0.63	0.64	18.11	36.68	6.00	3.66	Reniform	1	0	0	0	0	4
O.crista-galli	0.36	1.63	0.62	0.83	73.84	52.23	12.32	7.60	Semicircular	1	0	1	1	0	7
O.cyri	0.49	1.45	0.71	0.85	10.85	16.51	4.36	3.07	Elliptic	1	0	1	1	0	5
O.dealbata	0.40	1.68	0.60	0.90	65.36	46.47	11.74	6.99	Semicircular	1	1	1	1	0	13
O.eriophora	0.21	1.50	0.69	0.72	26.93	40.36	7.00	4.80	Reniform	1	1	0	0	0	11
O.eubrychidea	0.49	1.28	0.79	0.92	49.71	36.17	8.97	6.95	Elliptic	1	1	0	0	0	23
O.galegifolia	0.64	1.32	0.77	0.92	168.52	56.60	16.02	12.34	Elliptic	0	0	0	1	0	0
O.gaubae	0.43	1.53	0.66	0.88	83.76	51.35	12.74	8.33	Elliptic	1	0	0	1	0	13
O.gracilis	0.49	1.40	0.73	0.82	14.98	19.85	5.13	3.72	Elliptic	1	0	1	1	0	3
O.grandis	0.52	1.55	0.66	0.88	75.92	43.07	12.08	7.85	Elliptic	1	1	0	1	0	5
O.hajastana	0.43	1.22	0.83	0.82	11.84	18.84	4.25	3.53	Elliptic	1	0	0	0	0	3
O.hamata	0.74	1.28	0.78	0.95	21.03	18.86	5.85	4.57	Elliptic	0	0	1	1	0	0
O.heliocarpa	0.45	1.35	0.75	0.87	41.32	37.12	8.29	6.18	Elliptic	1	1	0	0	0	0
O.hemicycla	0.65	1.36	0.73	0.93	186.08	59.68	17.84	13.10	Elliptic	1	0	0	1	0	12
O.hispanica	0.24	1.99	0.52	0.77	44.50	48.36	10.58	5.41	Reniform	1	0	0	1	0	9

Species name	Circ.	AR	Round	Solidity	Area (mm^2)	Perim. (mm)	Length (mm)	Width (mm)	Fruit Shape	Fruit Spine	Fruit Trichome	Crest of Fruit	Fruit Vein	Spine on Spine	Fruit Spine Number
O.iberica	0.70	1.52	0.67	0.92	16.94	17.29	5.65	3.76	Semicircular	1	0	0	1	0	4
O.kachetica	0.55	1.30	0.78	0.92	99.11	48.36	12.68	9.87	Elliptic	1	1	1	1	0	35
O.kotschyana	0.44	1.45	0.71	0.81	14.39	20.47	5.05	3.56	Elliptic	1	0	0	1	0	4
O.laconica	0.27	1.48	0.68	0.72	19.53	30.27	6.05	4.10	Reniform	1	0	0	1	0	4
O.lanata	0.45	1.12	0.89	0.92	139.34	64.60	14.02	12.53	Elliptic	1	1	0	0	0	23
O.lasiostachya	0.41	1.27	0.79	0.76	18.30	23.81	5.44	4.29	Semicircular	1	0	1	0	0	4
O.laxiflora	0.63	1.61	0.63	0.92	48.79	30.68	9.75	6.10	Elliptic	1	1	0	1	0	7
O.lubrychidea	0.16	1.13	0.89	0.84	61.80	70.62	9.42	8.37	Reniform	1	1	1	1	0	35
O.lunata	0.20	1.23	0.82	0.80	120.82	106.80	13.66	11.12	Reniform	1	1	0	1	0	41
O.mairei	0.39	1.26	0.80	0.80	28.68	32.40	6.74	5.37	Reniform	1	0	1	1	0	4
O.melanotricha	0.03	1.26	0.80	0.55	124.20	247.08	14.09	11.23	Reniform	1	1	0	0	0	47
O.meschetica	0.42	1.36	0.74	0.85	90.37	52.27	12.31	9.15	Elliptic	1	0	0	1	0	27
O.michauxii	0.45	1.40	0.73	0.89	81.61	48.25	11.95	8.65	Elliptic	1	1	0	1	0	11
O.micrantha	0.32	1.69	0.60	0.75	51.21	46.65	10.42	6.23	Reniform	1	0	1	1	0	8
O.nummularia	0.15	1.25	0.80	0.80	89.08	96.07	11.89	9.52	Reniform	1	1	0	1	0	27

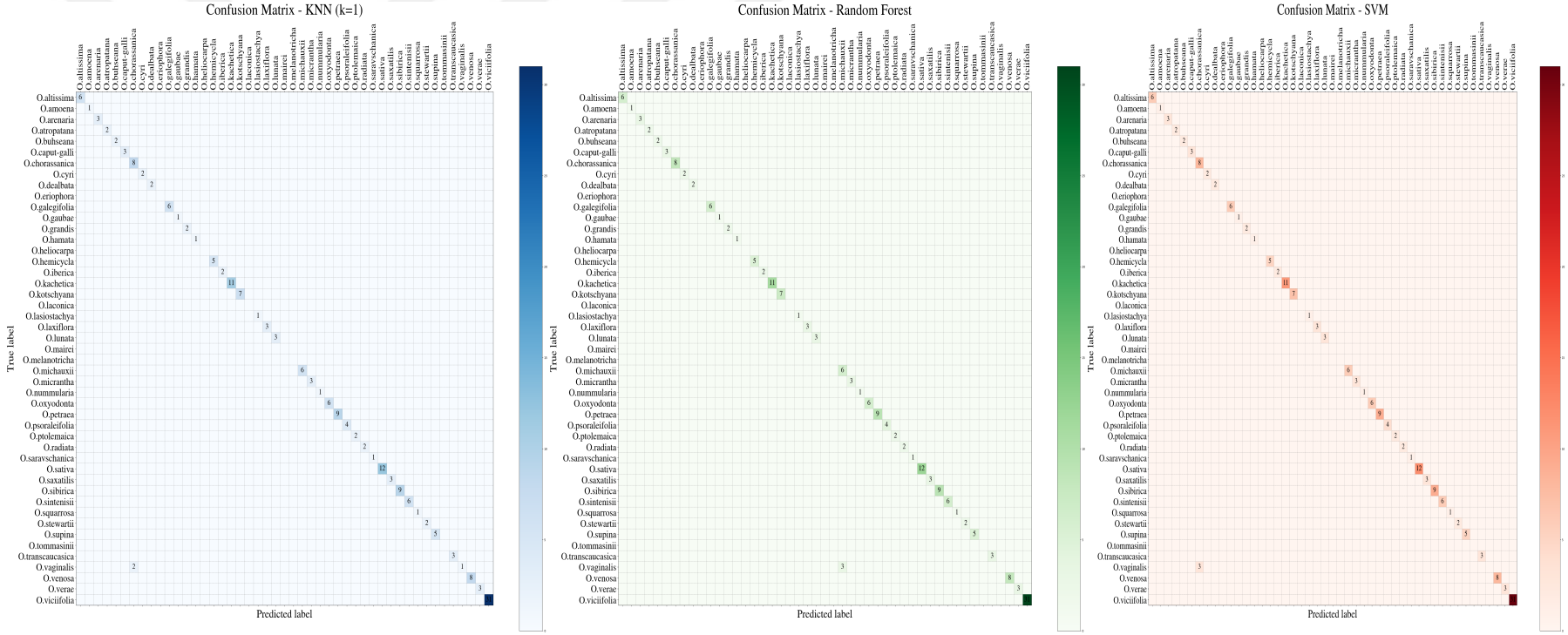
Species name	Circ.	AR	Round	Solidity	Area (mm^2)	Perim. (mm)	Length (mm)	Width (mm)	Fruit Shape	Fruit Spine	Fruit Trichome	Crest of Fruit	Fruit Vein	Spine on Spine	Fruit Spine Number
O.oxyodonta	0.48	1.66	0.63	0.83	10.48	16.90	4.58	2.85	Elliptic	1	0	0	1	0	3
O.oxytropoides	0.75	1.17	0.85	0.96	27.10	21.31	6.35	5.41	Semicircular	1	0	1	1	0	1
O.pallasii	0.70	1.22	0.82	0.95	78.18	36.27	10.68	8.74	Semicircular	0	0	0	1	0	0
O.paucidentata	0.33	1.36	0.74	0.70	29.25	33.72	7.12	5.22	Reniform	1	0	1	1	0	3
O.peduncularis	0.31	1.36	0.75	0.77	47.54	45.07	8.90	6.60	Reniform	1	1	0	1	0	7
O.persica	0.34	1.34	0.76	0.68	38.97	38.43	8.10	6.11	Reniform	1	0	1	1	0	3
O.petraea	0.58	1.47	0.68	0.91	17.10	19.40	5.64	3.85	Elliptic	1	0	0	1	0	5
O.pindicola	0.46	1.40	0.73	0.81	12.90	18.77	4.70	3.45	Elliptic	1	0	1	1	0	4
O.psoraleifolia	0.25	1.38	0.74	0.81	72.82	67.25	11.26	8.24	Reniform	1	1	0	1	0	21
O.ptolemaica	0.44	1.18	0.86	0.90	92.72	53.05	11.53	9.95	Elliptic	1	1	0	1	0	18
O.pulchella	0.37	1.53	0.66	0.85	63.16	50.08	10.96	7.21	Reniform	0	0	0	1	0	0
O.radiata	0.36	1.21	0.83	0.84	94.93	59.47	11.86	9.88	Elliptic	1	0	0	1	0	21
O.saravschanica	0.48	1.57	0.66	0.82	78.40	44.96	12.27	8.13	Elliptic	1	0	1	1	0	19
O.sativa	0.46	1.44	0.71	0.85	21.89	25.37	6.28	4.41	Elliptic	1	0	0	1	0	5
O.saxatilis	0.57	1.72	0.61	0.88	18.76	21.47	6.32	3.76	Elliptic	1	0	1	1	0	4

Species name	Circ.	AR	Round	Solidity	Area (mm^2)	Perim. (mm)	Length (mm)	Width (mm)	Fruit Shape	Fruit Spine	Fruit Trichome	Crest of Fruit	Fruit Vein	Spine on Spine	Fruit Spine Number
O.scrobiculata	0.15	1.14	0.88	0.81	116.40	112.73	13.00	11.38	Reniform	1	1	0	1	0	22
O.sibirica	0.59	1.37	0.75	0.89	18.83	19.98	5.59	4.21	Elliptic	1	1	0	1	0	3
O.sintenisii	0.63	1.37	0.74	0.91	83.36	40.73	11.88	8.83	Elliptic	0	0	0	1	0	0
O.sosnowskyi	0.42	1.40	0.72	0.85	20.03	24.95	5.94	4.28	Reniform	1	0	0	1	0	9
O.squarrosa	0.24	1.57	0.64	0.72	83.14	69.28	12.71	8.16	Reniform	1	0	1	1	1	3
O.stenorhiza	0.40	1.57	0.66	0.82	16.21	23.75	5.62	3.65	Elliptic	1	0	0	0	0	7
O.stewartii	0.52	1.41	0.72	0.87	19.98	22.95	5.93	4.25	Elliptic	1	1	0	1	0	12
O.subnitens	0.60	1.26	0.81	0.92	158.35	57.72	15.66	12.68	Elliptic	1	0	1	1	0	17
O.supina	0.50	1.48	0.70	0.83	15.54	22.77	5.33	3.63	Reniform	0	0	1	1	0	0
O.tommasinii	0.37	1.65	0.62	0.86	26.78	31.40	7.44	4.58	Reniform	1	0	0	1	0	4
O.transcaspica	0.58	1.40	0.73	0.88	22.94	23.02	6.33	4.59	Elliptic	1	0	0	1	0	5
O.transcaucasica	0.45	1.39	0.73	0.84	19.91	24.18	5.89	4.29	Elliptic	1	0	1	1	0	5
O.vaginalis	0.58	1.42	0.71	0.87	107.30	47.45	13.70	9.65	Elliptic	1	1	0	1	0	11
O.venosa	0.60	1.36	0.75	0.91	123.41	49.96	13.96	10.41	Elliptic	0	0	0	1	0	0
O.verae	0.46	1.38	0.73	0.86	40.57	33.36	8.31	6.10	Elliptic	1	0	1	1	0	5

Species name	Circ.	AR	Round	Solidity	Area (mm^2)	Perim. (mm)	Length (mm)	Width (mm)	Fruit Shape	Fruit Spine	Fruit Trichome	Crest of Fruit	Fruit Vein	Spine on Spine	Fruit Spine Number
O.viciifolia	0.59	1.47	0.71	0.90	28.12	26.15	7.15	4.94	Elliptic	1	0	1	1	0	5

4.3. Machine Learning

Morphological traits have long been used as a primary data source in plant systematics, particularly for distinguishing between species. Furthermore, testing the extent to which these traits provide discriminatory information for species classification using statistical and computational methods is one of the most important contributions of modern bioinformatics approaches. In this context, machine learning analyses performed on data obtained from fruit and leaf measurements have yielded remarkable results. Initially, classification was performed using all morphological traits using three different algorithms (RF, SVM, and KNN), achieving an accuracy rate of up to 97% (Figure 4.17). This result demonstrates that fruit and leaf morphology carry strong discriminatory signals among *Onobrychis* species. Variable selection was then conducted based on correlation analysis and feature importance values, and 16 traits with correlations above 2.5 were identified. The selectec traits are as follows Fruit Spine number, Indumentum absence (1/0)-leaflet, Crest of fruit, Fruit Trichome, Leaflet-Shape_elliptic, Base (Leaflet base)_Rounded, Area-Fruit, Apex (Leaflet type)_Rounded, Perimeter-Fruit, Leaflet-Shape_oblong-elliptic, Apex (Leaflet type)_Obtuse, Perim. (mm)-leaflet, Area (mm2)-leaflet, Fruit Shape_Elliptic, Apex (Leaflet type)_acute, Apex (Leaflet type)_Acuminate. It has been determined that these traits are important in taxonomic evaluation. Repeated analyses with this subset maintained accuracy rates between 93% and 97%, demonstrating that these traits provide essential information for classification and that their biological significance was not lost despite the reduced data size (Figure 4.19). In particular, the RF algorithm maintained a 97% accuracy rate in both cases, demonstrating both the robustness of the algorithm and the biological representativeness of the selected traits. The findings demonstrate that fruit and leaf measurements are not only descriptive but also highly informative for numerical classification. These traits can be used as reliable parameters in assessing interspecies relationships, taxonomic revisions, and potentially in breeding studies. Furthermore, this validation with machine learning methods demonstrates the robustness of classical morphological approaches when supported by modern statistical methods.

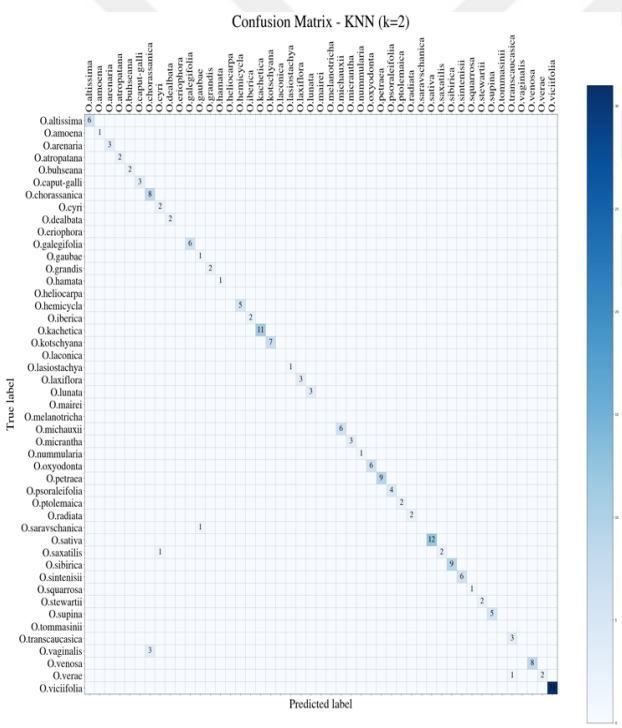


accuracy			0.97	194
macro avg	0.96	0.96	0.96	194
weighted avg	0.97	0.97	0.97	194

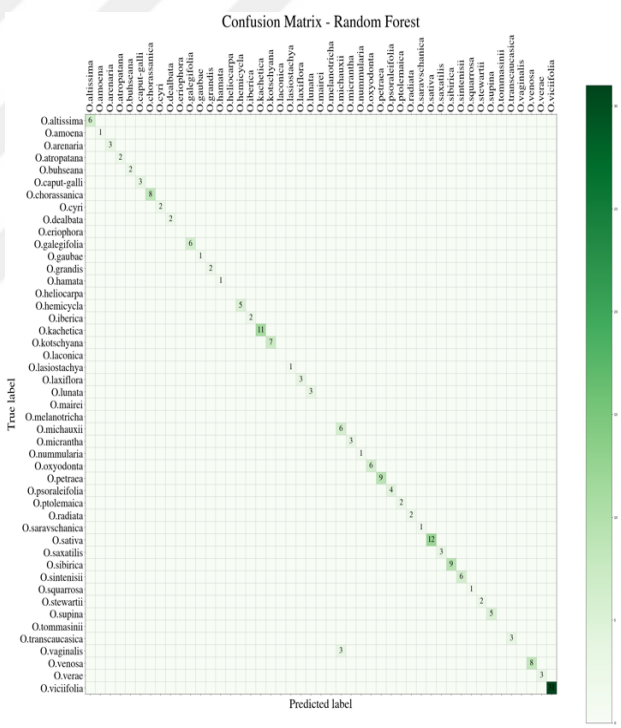
accuracy			0.97	194
macro avg	0.93	0.95	0.93	194
weighted avg	0.95	0.97	0.96	194

accuracy			0.97	194
macro avg	0.94	0.95	0.94	194
weighted avg	0.95	0.97	0.96	194

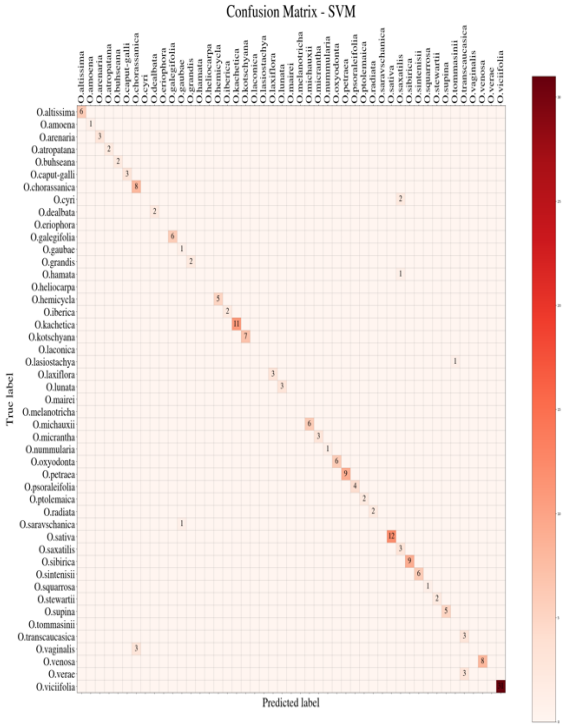
Figure 4.19. Analysis results of ML algorithms using all morphological measurements of leaves and fruits.



accuracy			0.95	194
macro avg	0.88	0.91	0.89	194
weighted avg	0.94	0.95	0.94	194



accuracy			0.97	194
macro avg	0.93	0.95	0.93	194
weighted avg	0.95	0.97	0.96	194



accuracy			0.93	194
macro avg	0.75	0.81	0.77	194
weighted avg	0.89	0.93	0.90	194

Figure 4.20. Analysis results of ML algorithms using selected morphological measurements of leaves and fruits.

When run with all features, all three models yielded very high accuracy (≈ 0.97). When running the ML algorithms with the 16 selected features, SVM's performance was significantly lower, while KNN's performance was generally stable, albeit with a slight decrease, while RF's performance was maintained. Feature selection reduced model complexity but also reduced SVM's generalization power (Table 4.5).

Table 4.5. Comparison of results of all ML algorithms.

Model	Accuracy	Macro Avg. (Precision/Recall/F1)	Weigted Avg (Precision/Recall/F1)	Total Samples	Data set (Leaf+Fruit)
KNN	0.97	0.96 / 0.96 / 0.96	0.97 / 0.97 / 0.97	194	All traits
RF	0.97	0.93 / 0.95 / 0.93	0.95 / 0.97 / 0.96	194	All traits
SVM	0.97	0.94 / 0.95 / 0.94	0.95 / 0.97 / 0.96	194	All traits
KNN	0.95	0.88 / 0.91 / 0.89	0.94 / 0.95 / 0.94	194	Selected traits
RF	0.97	0.93 / 0.95 / 0.93	0.95 / 0.97 / 0.96	194	Selected traits
SVM	0.93	0.75 / 0.81 / 0.77	0.89 / 0.93 / 0.90	194	Selected traits

In plant systematics, morphological traits have emerged as a key source of information in recent years. They frequently yield accurate results when it comes to species differentiation [122]. However, the contributions of contemporary bioinformatics and machine learning techniques have made it feasible to test objectively and quantifiably the actual amount of discriminatory information these traits convey. Accuracy of up to 97% was attained in this study when three distinct machine learning algorithms (RF, SVM, and KNN) were used to assess the classification performance of measurements based on leaf and fruit morphology. This finding supports traditional systematic approaches and shows that morphological traits among *Onobrychis* species carry strong phylogenetic signals. Especially in analyses based on trait selection, the maintenance of accuracies between 93% and 97%, even when only 16 morphological variables were used, demonstrates the biological importance of these characters in classification. The RF algorithm's 97% accuracy in both analyses highlights both the robust nature of the algorithm and the high representativeness of the selected traits. Similar to this, recent research on a variety of plant species has demonstrated that morphological measurements in conjunction with machine learning techniques yield high classification accuracy [123], [124]. Similarly, Karbstein et al. (2024) emphasized that machine learning-based species delimitation methods complement classical morphological analyses, reducing subjectivity in systematic botany [125]. Another study showed that AI-based classification algorithms significantly

increased accuracy rates in integrated morphological and genetic analyses of species [126]. These results show that, when combined with contemporary data-driven techniques, morphology offers a potent classification and validation mechanism in addition to being a descriptive tool. Consequently, fruit and leaf morphology can be regarded as trustworthy metrics that can be applied to taxonomic revisions as well as breeding in the context of identificatio of distinct germplasm and also in the form of parent selection. Additionally, these machine learning algorithm-derived results substantiate the resilience of traditional morphological methods and unequivocally show how bioinformatics techniques contribute to systematic botanical research.

5. CONCLUSION

The Genus *Onobychis* has unarguably one of the most controvercial taxonomy and ther is an urgent need to devote a set of extensive efforts to deliante the true status of taxa included in. The current study is an effort to understand the distinction among a number of taxa included in the genus along with testing the use of image analyses of fruit and leaf related traits. With the advances in the image capture and and furher processing including machine learling tecniques, it is now possible to proccess a large number of images taken froma larger set of taxa very accurately to deduce the relationship among taxa. Here, the image analyses were coupled with the marker based data to understand the relatioinship among the taxa included in the genus. The results showed that leaf and fruit morphology of the genus *Onobrychis* estimated using image analyses techniques provide a powerful taxonomic signal for identifying interspecies differences. High morphological diversity was observed in the subgenus *Sisyrosema*, but PCA, clustering, and dendrogram analyses mainly supported distinctions, especially at the subgenus level. However, they are not very effective to provide a clear resoultion among species in the same subgenus or section. High polymorphism was found by molecular analyses using iPBS markers; however, genetic diversity failed to draw distinct divisions at the subgenus or section level. Despite the lack of possibility to deduce the true species status, morphological data can distinguish species with high accuracy (up to 97%), according to machine learning algorithms. This shows that the reliability of classical morphological approaches is increased when they are combined with contemporary computational techniques. All things considered, these results imply that a more thorough delineation of the taxonomic boundaries of the genus *Onobrychis* are needed particularly for identifying hybridization patterns for closely related taxa and the degree of gene flow among species. The genus would still benefit from results of the combined use of morphological and molecular data. In addition to PCR based molecular data, a set of more conserved DNA polymorphism from chloroplast and mitocondria would be very helpful to discern the evoution of the taxa included in the genus and would be an ambitious research task to accomplish.

6. RECOMMENDATIONS

- In addition to leaflet and fruit morphology, it is believed that examining root and flower structures using digital measurement techniques could contribute to the taxonomic analysis of the genus. Digital measurements should be used to make species identification keys more reliable.
- In addition to iPBS markers, analyses of chloroplast and mitochondrial genomes could be more accurate to provide a taxonomic relationships.
- The relationship among cultivated taxa and their speices status has not been set and should be tested based on hybridization patterns.
- According to results, it is suggested that the taxonomic level of some infraspecific taxa should be updated according to species.
- The accuracy of species classifications can be increased by developing integrated machine learning models that process morphological and molecular data together.
- The use of deep learning and machine learning algorithms, particularly in image-based automatic species identification systems, is recommended.

REFERENCES

[1] S. Y. Tan, “Agricultural Characteristics, Economic Importance, Taxonomies and General Structural Characteristics of Leguminous Forage Crops,” *Forage Crops Legum. Forage Crops*, vol. 2, pp. 277–289, 2009.

[2] M. G. Simpson, “Diversity and Classification of Flowering Plants: Eudicots,” in *Plant Systematics (Second Edition)*, M. G. Simpson, Ed., San Diego: Academic Press, 2010, pp. 275–448. doi: 10.1016/B978-0-12-374380-0.50008-7.

[3] M. Okcu, “Identification of wild sainfoins (Onobrychis spp.) in Eastern Anatolia region and determination of their some characteristics,” Atatürk University, Graduate School of Natural and Applied Sciences, 2009.

[4] E. Aktoklu, “Revision of Onobrychis Miller (Fabaceae) species grown in Turkey,” *İnönü Univ.-Grad. Sch. Nat. Appl. Sci.-Dep. Biol.*, 1995.

[5] C. H. Carbonero, I. Mueller-Harvey, T. A. Brown, and L. Smith, “Sainfoin (Onobrychis viciifolia): a beneficial forage legume,” *Plant Genet. Resour.*, vol. 9, no. 1, pp. 70–85, Apr. 2011, doi: 10.1017/S1479262110000328.

[6] B. Yildiz, B. Çiplak, and E. Aktoklu, “Fruit morphology of sections of the genus Onobrychis miller (fabaceae) and its phylogenetic implications,” *Isr. J. Plant Sci.*, vol. 47, no. 4, pp. 269–282, May 1999, doi: 10.1080/07929978.1999.10676784.

[7] C. Hayot Carbonero, F. Carbonero, L. M. J. Smith, and T. A. Brown, “Phylogenetic characterisation of Onobrychis species with special focus on the forage crop Onobrychis viciifolia Scop.,” *Genet. Resour. Crop Evol.*, vol. 59, no. 8, pp. 1777–1788, Dec. 2012, doi: 10.1007/s10722-012-9800-3.

[8] A. Amirahmadi, S. Kazempour-Osaloo, A. Kaveh, A. A. Maassoumi, and R. Naderi, “The phylogeny and new classification of the genus Onobrychis (Fabaceae-Hedysareae): evidence from molecular data,” *Plant Syst. Evol.*, vol. 302, no. 10, pp. 1445–1456, Dec. 2016, doi: 10.1007/s00606-016-1343-1.

[9] Sirjaev, “Onobrychis generis revisio critica,” *Publ. Fac. Sci. Univ. Masaryk*, vol. 56, pp. 96–97, 1925.

[10] H. Akhani, “Flora Iranica: Facts and figures and a list of publications by K. H. Rechinger on Iran and adjacent areas,” *Rostaniha*, vol. 7, pp. 19–61, Jan. 2006.

[11] S. N. Semnani, A. Majd, H. Ghassempour, S. N. Lesani, and N. Javar, “Pollen Morphology of the Genera Onobrychis (Hedysareae, Fabaceae) in Iran,” p. 6, 2014.

[12] H. Amirabadizadeh, A. Jafari, and H. Mahmoodzadeh, “Comparative morphology, anatomy and palynological studies of perennial species of Onobrychis (Fabaceae) in northeast Iran,” *Nord. J. Bot.*, vol. 33, no. 2, pp. 159–169, 2015, doi: 10.1111/njb.00580.

[13] D.-L. Guo, M.-X. Guo, X.-G. Hou, and G.-H. Zhang, “Molecular diversity analysis of grape varieties based on iPBS markers,” *Biochem. Syst. Ecol.*, vol. 52, pp. 27–32, Feb. 2014, doi: 10.1016/j.bse.2013.10.008.

[14] R. Kalendar, K. Antonius, P. Smýkal, and A. H. Schulman, “iPBS: a universal method for DNA fingerprinting and retrotransposon isolation,” *TAG Theor. Appl. Genet. Theor. Angew. Genet.*, vol. 121, no. 8, pp. 1419–1430, Nov. 2010, doi: 10.1007/s00122-010-1398-2.

[15] D. V. Vazquez, F. E. Spetale, A. N. Nankar, S. Grozeva, and G. R. Rodríguez, “Machine Learning-Based Tomato Fruit Shape Classification System,” *Plants*, vol. 13, no. 17, p. 2357, Aug. 2024, doi: 10.3390/plants13172357.

[16] T. Ishikawa *et al.*, "Classification of Strawberry Fruit Shape by Machine Learning," *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*, vol. XLII–2, pp. 463–470, May 2018, doi: 10.5194/isprs-archives-XLII-2-463-2018.
[17] C. A. Schneider, W. S. Rasband, and K. W. Eliceiri, "NIH Image to ImageJ: 25 years of image analysis," *Nat. Methods*, vol. 9, no. 7, pp. 671–675, July 2012, doi: 10.1038/nmeth.2089.
[18] A. D. J. van Dijk, G. Kootstra, W. Kruijer, and D. de Ridder, "Machine learning in plant science and plant breeding," *iScience*, vol. 24, no. 1, Jan. 2021, doi: 10.1016/j.isci.2020.101890.
[19] V. Kuete, K. Viertel, and T. Efferth, "Antiproliferative Potential of African Medicinal Plants," in *Medicinal Plant Research in Africa*, V. Kuete, Ed., Oxford: Elsevier, 2013, pp. 711–724. doi: 10.1016/B978-0-12-405927-6.00018-7.
[20] Ahmad, Anwar, F., and Hira, S., "Review on Medicinal Importance of Fabaceae Family," *Pharmacol. OnLine*, vol. 3, pp. 151–157, 2016.
[21] H. H. S. Mohsen and M. Z. Nasab, "Cytotaxonomy of some Onobrychis (Fabaceae) species and populations in Iran," *Caryologia*, vol. 63, no. 1, pp. 18–31, Jan. 2010, doi: 10.1080/00087114.2010.589705.
[22] A. Kaveh, S. Kazempour-Osaloo, A. Amirahmadi, A. Maassoumi, and G. M. Schneeweiss, "Systematics of Onobrychis sect. Heliobrychis (Fabaceae): morphology and molecular phylogeny revisited," *Plant Syst. Evol.*, vol. 305, no. 1, pp. 33–48, Jan. 2019, doi: 10.1007/s00606-018-1549-5.
[23] M. Ranjbar, R. Karamian, and S. Afsari, "Meiotic chromosome number and behaviour of Onobrychis avajensis (Fabaceae): a new species from western Iran," *Plant Ecol. Evol.*, vol. 143, no. 2, pp. 170–175, July 2010, doi: 10.5091/plecevo.2010.431.
[24] I. ILDIS, "International Legume Database & Information Service (ILDIS)," ILDIS. [Online]. Available: https://ildis.org/cgi-bin/Araneus.pl?version~10.01
[25] A. Amirahmadi and M. Khoshsokhan, "Molecular and morphological evidence of Onobrychis avanakensis, a new species from Iran," *Phytotaxa*, vol. 487, pp. 75–82, Feb. 2021, doi: 10.11646/phytotaxa.487.1.6.
[26] P. Polhill, R. M. R. M., "Advances in Legume Systematics.," *Kew R. Bot. Gard.*, vol. Part 1., 1981.
[27] M. Tekin and G. Yilmaz, "Comparative Root and Stem Anatomy of Four Rare Onobrychis Mill. (Fabaceae) Taxa Endemic in Turkey," *Not. Sci. Biol.*, vol. 7, no. 3, Art. no. 3, Sept. 2015, doi: 10.15835/nsb739595.
[28] "HealthyHay Project | Legume Plus." Accessed: Dec. 05, 2022. [Online]. Available: https://cotswoldseeds.com/legumeplus-healthyhay-project.asp
[29] M. Mora Ortiz and L. M. J. Smith, "Onobrychis viciifolia; a comprehensive literature review of its history, etymology, taxonomy, genetics, agronomy and botany," *Plant Genet. Resour.*, vol. 16, no. 5, pp. 403–418, Oct. 2018, doi: 10.1017/S1479262118000230.
[30] A. Kök, "General structural (morphological) characteristics of legumes feed plants 1. Vegetative features," p. 120, 2019.
[31] C. Wijekoon, S. N. Acharya, Y. L. Siow, S. Sura, S. Thandapilly, and A. Sabra, "Canadian sainfoin and fenugreek as forage and functional foods," *Crop Sci.*, vol. 61, no. 1, pp. 1–20, 2021, doi: 10.1002/csc2.20280.
[32] S. Bhattarai, B. Coulman, and B. Biligetu, "Sainfoin (Onobrychis viciifolia Scop.): renewed interest as a forage legume for western Canada," *Can. J. Plant Sci.*, vol. 96, no. 5, pp. 748–756, Oct. 2016, doi: 10.1139/cjps-2015-0378.

[33] M. Mora-Ortiz and L. M. J. Smith, "*Onobrychis viciifolia* ; a comprehensive literature review of its history, etymology, taxonomy, genetics, agronomy and botany," *Plant Genet. Resour. Charact. Util.*, vol. 16, no. 5, pp. 403–418, Oct. 2018, doi: 10.1017/S1479262118000230.
[34] A. Açıkgöz E., "Forage Crops." Bursa Uludağ Unıversity, Ziraat Fakültesi Yayınları, Bursa., 2001.
[35] M. Deveci, R. Sıralı, and Ş. Cınbırtoğlu, "The Importance Of Sainfoin (Onobrychis sp.) Cultivation For Beekeeping," *Arıcılık Araşt. Derg.*, vol. 8(2), pp. 16–19, Dec. 2012.
[36] M. Okcu and S. Şengül, "A study on the determination of the morphological, Yield and quality characteristics of some sainfoin species (Onobrychis spp.) native to East Anatolia," *Pak. J. Bot.*, vol. 46, pp. 1837–1842, Oct. 2014.
[37] Y. Wang, T. A. McAllister, and S. Acharya, "Condensed Tannins in Sainfoin: Composition, Concentration, and Effects on Nutritive and Feeding Value of Sainfoin Forage," *Crop Sci.*, vol. 55, no. 1, pp. 13–22, 2015, doi: 10.2135/cropsci2014.07.0489.
[38] "TÜİK," Tüik. Accessed: Nov. 06, 2022. [Online]. Available: https://data.tuik.gov.tr/Kategori/GetKategori?p=tarim-111&dil=1
[39] "TÜİK - Veri Portalı." Accessed: Sept. 11, 2025. [Online]. Available: https://data.tuik.gov.tr/Kategori/GetKategori?p=tarim-111
[40] E. Karabulut, K. Erkoç, and M. Şakiroğlu, "Sainfoin (Onobrychis sp.) crop ontology: supporting germplasm characterization and international research collaborations," 2023.
[41] Akçelik, "Bazı yabani korunga (Onobrychis sp.) türlerinin kromozom sayılarının tespit ve karyotip analizi (Doctoral dissertation, Yüksek lisans tezi, Ankara Üniversitesi, Fen Bilimleri Enstitüsü, Ankara).".
[42] I. Avci E., Erayman, M., and Sancak, C, "Analysis of onobrychis genetic diversity using SSR markers from related legume species," *J. Anim. Plant Sci.*, vol. 24, pp. 556–566, Apr. 2014.
[43] S. 1989- Bhattarai, "Characterization of diverse germplasm of sainfoin (Onobrychis viciifolia Scop.) using agro-morphological traits and AFLP molecular markers," Thesis, University of Saskatchewan, 2017. Accessed: Jan. 09, 2022. [Online]. Available: https://harvest.usask.ca/handle/10388/7977
[44] D. Altindal and N. Altindal, "Allelopathic effects of olive oil mill wastewater (omw) on sainfoin (Onobrychis Viciifolia Scop.) germination," *Uluslar. Tarım Orman. Ve Yaşam Bilim. Derg.*, vol. 2, no. 2, Art. no. 2, Dec. 2018.
[45] Z. Dumanoğlu, E. Çaçan, and K. Kökten, "A Research on Determination of Physical Properties of Sainfoin (Onobrychis viciifolia Scop.) Genotypes Seeds," *J. Anatol. Environ. Anim. Sci.*, vol. 6, pp. 18–24, 2021, doi: 10.35229/jaes.802858.
[46] S. Demdoum, "Caracterización agronómica, genética y composición química de una colección de variedades de esparceta," Ph.D. Thesis, Universitat de Lleida, 2012. Accessed: Oct. 17, 2022. [Online]. Available: http://www.tdx.cat/handle/10803/124106
[47] S. E. McGregor, *Insect Pollination of Cultivated Crop Plants*. Agricultural Research Service, U.S. Department of Agriculture, 1976.
[48] Hedge, "Onobrychis Adans.," *Flora Turk. East Aegean Isl.*, vol. 3, pp. 560–589, 1970.
[49] M. Zarrabian, A. N. Dehkordi, M. H. Ehtemam, and M. M. Majidi, "Relationships of Worldwide Onobrychis species: from seed & pod morphology and molecular perspective," May 03, 2022, *bioRxiv*. doi: 10.1101/2022.05.02.490301.
[50] N. Azizi, S. M. Talebi, R. Tabaripour, and M. Hatami, "Characterization of wild Onobrychis Mill. Subgenus Sisyrosema bunge (Fabaceae) germplasm resources from

Iran using macro and micro-morphological traits," *Genet. Resour. Crop Evol.*, Sept. 2023, doi: 10.1007/s10722-023-01734-1.
[51] M. A. Nadeem *et al.*, "DNA molecular markers in plant breeding: current status and recent advancements in genomic selection and genome editing," *Biotechnol. Biotechnol. Equip.*, vol. 32, no. 2, pp. 261–285, Mar. 2018, doi: 10.1080/13102818.2017.1400401.
[52] M. M. Abou-El-Enain, "Chromosomal criteria and their phylogenetic implications in the genus Onobrychis Mill. sect. Lophobrychis (Leguminosae), with special reference to Egyptian species," *Bot. J. Linn. Soc.*, vol. 139, no. 4, pp. 409–414, Aug. 2002, doi: 10.1046/j.1095-8339.2002.00075.x.
[53] M. Zarrabian and M. M. Majidi, "Genetic diversity and relationships within and among Onobrychis species using molecular markers," *Turk. J. Bot.*, vol. 39, no. 4, pp. 681–692, Jan. 2015, doi: 10.3906/bot-1406-47.
[54] C. Hayot Carbonero, F. Carbonero, L. M. J. Smith, and T. A. Brown, "Cytological characterisation of the underutilized forage crop Onobrychis viciifolia Scop. and other members of the Onobrychis genus," *Genet. Resour. Crop Evol.*, vol. 60, no. 7, pp. 1987–1996, Oct. 2013, doi: 10.1007/s10722-013-9967-2.
[55] G. Yucel, A. Betekhtin, E. Cabi, M. Tuna, R. Hasterok, and B. Kolano, "The Chromosome Number and rDNA Loci Evolution in Onobrychis (Fabaceae)," *Int. J. Mol. Sci.*, vol. 23, no. 19, Art. no. 19, Jan. 2022, doi: 10.3390/ijms231911033.
[56] A. L. Tek and S. D. Kara Öztürk, "High allelic diversity of the centromere-specific histone H3 (CENH3) in the legume sainfoin (Onobrychis viciifolia)," *Mol. Biol. Rep.*, vol. 47, no. 11, pp. 8789–8795, Nov. 2020, doi: 10.1007/s11033-020-05926-1.
[57] A. Amirahmadi, S. OSALOO, M. MOZAFFAR, and M. CHARKHCHIAN, "A new species of Onobrychis sect. Onobrychis (Fabaceae) from Iran," *Turk. J. Bot.*, vol. 38, no. 4, pp. 658–664, Jan. 2014, doi: 10.3906/bot-1309-54.
[58] M. Noori, M. Dehshiri, and M. Sharifi, "Numerical Taxonomy of Onobrychis Miller (Hedysareae, Fabaceae) from Markazi Province, Iran Using Pod and Seed Morphological Characters," *Int. J. Mod. Bot.*, vol. 4, pp. 40–47, Jan. 2014, doi: 10.5923/j.ijmb.20140402.02.
[59] C. A. Cenci, G. Bassi, F. Ferranti, and B. Romano, "Some morphometric, anatomical and biochemical characteristics of fruits and seeds of *Onobrychis* spp. in Italy," *Plant Biosyst. - Int. J. Deal. Asp. Plant Biol.*, vol. 134, no. 1, pp. 91–98, Jan. 2000, doi: 10.1080/11263500012331350395.
[60] S. Avci, "Heliobrychis Seksiyonuna Ait Bazı Korunga (Onobrychis sp.) Türleri Üzerinde Morfolojik Araştırmalar," *J. Field Crops Cent. Res. Inst.*, p. 6, 2010.
[61] S. Avci, C. Sancak, A. Can, A. Acar, and N. Pınar, "Pollen morphology of the genus Onobrychis (Fabaceae) in Turkey," *Turk. J. Bot.*, vol. 37, no. 4, pp. 669–681, Jan. 2013, doi: 10.3906/bot-1207-52.
[62] S. M. Talebi, N. Azizi, P. Yadegari, and A. Matsyura, "Analysis of pollen morphological characteristics in Iranian Onobrychis Miller (Fabaceae) taxa," *Braz. J. Bot.*, vol. 43, no. 3, pp. 609–632, Sept. 2020, doi: 10.1007/s40415-020-00623-6.
[63] M. Okcu and S. Şengül, "A study on the determination of the morphological, Yield and quality characteristics of some sainfoin species (Onobrychis spp.) native to East Anatolia," *Pak. J. Bot.*, vol. 46, pp. 1837–1842, Oct. 2014.
[64] R. Massoud, R. Karamian, and A. Hadadi, "Cytosystematics of three Onobrychis species (Fabaceae) in Iran," *Caryologia*, vol. 63, no. 3, pp. 237–249, July 2010, doi: 10.1080/00087114.2010.10589733.

[65] Ghanavati, F., Mozafari, J., Naghavi, M. R., Amirabadizadeh, H., & Darvish, F., "Phylogenetic relationships of Onobrychis Mill.(Fabaceae: Papilionoideae) based on ITS sequences of nuclear ribosomal DNA and morphological traits.," pp. 91–99, 2012.
[66] R. Massoud, R. Karamian, and A. Hadadi, "Cytosystematics of three Onobrychis species (Fabaceae) in Iran," *Caryologia*, vol. 63, no. 3, pp. 237–249, July 2010, doi: 10.1080/00087114.2010.10589733.
[67] H. Sepet, İ. Emre, Y. Kiran, M. Kursat, and A. Sahin, "Karyological studies on eight species of Onobrychis genus in Turkey," *Biologia (Bratisl.)*, vol. 66, no. 6, pp. 996–1002, Dec. 2011, doi: 10.2478/s11756-011-0112-4.
[68] M. Şakiroğlu, "Population Genomics of Perennial Temperate Forage Legumes," in Population Genomics. , Cham: Springer International Publishing, 2021, pp. 1–39. doi: 10.1007/13836_2021_90.
[69] S. M. Zargar *et al.*, "Recent advances in molecular marker techniques: Insight into QTL mapping, GWAS and genomic selection in plants," *J. Crop Sci. Biotechnol.*, vol. 18, no. 5, pp. 293–308, Dec. 2015, doi: 10.1007/s12892-015-0037-5.
[70] M. Zarrabian, M. M. Majidi, and M. H. Ehtemam, "Genetic Diversity in a Worldwide Collection of Sainfoin Using Morphological, Anatomical, and Molecular Markers," *Crop Sci.*, vol. 53, no. 6, pp. 2483–2496, 2013, doi: 10.2135/cropsci2013.03.0130.
[71] M. R. F. Darvish, "Assessment of genetic variability of 36 populations of sainfoin (Onobrychis sativa) based on RAPD markers.," *Int. J. Biosci.*, vol. 3, pp. 15–26, 2013, doi: 10.12692/ijb/3.10.15-26.
[72] H. Nosrati, M. Feizi, F. Latifian, and A. Razban Haghighi, "Eco-geographical variations of issrs among populations of onobrychis viciifolia (Sainfoin, Fabacae)," *Ann. Univ. Oradea Fasc. Biol.*, vol. 23, pp. 62–66, July 2016.
[73] S. Özkan and B. B. Bilgen, "Genetic characterization of sainfoin (Onobrychis viciifolia) varieties and populations using microsatellite markers.," *J. Tekirdag Agric. Fac.*, vol. 16, no. 1, pp. 51–60, 2019.
[74] T. Sutcu, B. B. Bilgen, and M. Tuna, "Analysis of genetic diversity among Onobrychis accessions with high agronomic performance by simple sequence repeat (SSR) markers," *Mol. Biol. Rep.*, vol. 49, no. 6, pp. 5659–5668, June 2022, doi: 10.1007/s11033-022-07584-x.
[75] R. Kalendar and A. H. Schulman, "IRAP and REMAP for retrotransposon-based genotyping and fingerprinting," *Nat. Protoc.*, vol. 1, no. 5, pp. 2478–2484, 2006, doi: 10.1038/nprot.2006.377.
[76] K. R. Oliver, J. A. McComb, and W. K. Greene, "Transposable elements: powerful contributors to angiosperm evolution and diversity," *Genome Biol. Evol.*, vol. 5, no. 10, pp. 1886–1901, 2013, doi: 10.1093/gbe/evt141.
[77] N. Hasan, S. Choudhary, N. Naaz, N. Sharma, and R. A. Laskar, "Recent advancements in molecular marker-assisted selection and applications in plant breeding programmes," *J. Genet. Eng. Biotechnol.*, vol. 19, no. 1, p. 128, Aug. 2021, doi: 10.1186/s43141-021-00231-1.
[78] P. Nongdam and T. Amom, "The Use of Molecular Marker Methods in Plants: A Review," July 2017, doi: 10.7324/IJCRR.2017.9171.
[79] U. Kumar *et al.*, "Association Mapping, Principles and Techniques," *J. Biol. Environ. Eng.*, vol. 2, pp. 1–9, Jan. 2017.
[80] J. Zou, H. Gong, T.-J. Yang, and J. Meng, "Retrotransposons - a major driving force in plant genome evolution and a useful tool for genome analysis," *J. Crop Sci. Biotechnol.*, vol. 12, no. 1, pp. 1–8, Mar. 2009, doi: 10.1007/s12892-009-0070-3.

[81] B. Eren *et al.*, “Assessment of genetic diversity and population structure in local alfalfa genotypes using iPBS molecular markers,” *Genet. Resour. Crop Evol.*, vol. 70, no. 2, pp. 617–628, Feb. 2023, doi: 10.1007/s10722-022-01450-2.
[82] K. Haliloğlu, A. Türkoğlu, H. I. Öztürk, G. Özkan, E. Elkoca, and P. Poczai, “iPBS-Retrotransposon Markers in the Analysis of Genetic Diversity among Common Bean (Phaseolus vulgaris L.) Germplasm from Türkiye,” *Genes*, vol. 13, no. 7, p. 1147, June 2022, doi: 10.3390/genes13071147.
[83] M. İ. Işık *et al.*, “Understanding genetic diversity and population structure in forage pea (Pisum sativum var. arvense L.) using inter-primer binding site (iPBS) retrotransposon marker,” *Genet. Resour. Crop Evol.*, vol. 72, no. 7, pp. 8389–8403, Oct. 2025, doi: 10.1007/s10722-025-02469-x.
[84] P. P. Shinde and S. Shah, “A Review of Machine Learning and Deep Learning Applications,” in *2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)*, Aug. 2018, pp. 1–6. doi: 10.1109/ICCUBEA.2018.8697857.
[85] M. I. Jordan and T. M. Mitchell, “Machine learning: Trends, perspectives, and prospects,” *Science*, vol. 349, no. 6245, pp. 255–260, July 2015, doi: 10.1126/science.aaa8415.
[86] U. Lee, S. Chang, G. A. Putra, H. Kim, and D. H. Kim, “An automated, high-throughput plant phenotyping system using machine learning-based plant segmentation and image analysis,” *PLOS ONE*, vol. 13, no. 4, p. e0196615, Apr. 2018, doi: 10.1371/journal.pone.0196615.
[87] P. Wilf, S. Zhang, S. Chikkerur, S. A. Little, S. L. Wing, and T. Serre, “Computer vision cracks the leaf code,” *Proc. Natl. Acad. Sci.*, vol. 113, no. 12, pp. 3305–3310, Mar. 2016, doi: 10.1073/pnas.1524473113.
[88] J. Wäldchen, M. Rzanny, M. Seeland, and P. Mäder, “Automated plant species identification—Trends and future directions,” *PLoS Comput. Biol.*, vol. 14, no. 4, p. e1005993, Apr. 2018, doi: 10.1371/journal.pcbi.1005993.
[89] B. K. Hatuwal, A. Shakya, and B. Joshi, “Plant Leaf Disease Recognition Using Random Forest, KNN, SVM and CNN”.
[90] T. Cover and P. Hart, “Nearest neighbor pattern classification,” *IEEE Trans. Inf. Theory*, vol. 13, no. 1, pp. 21–27, Jan. 1967, doi: 10.1109/TIT.1967.1053964.
[91] C. Cortes and V. Vapnik, “Support-vector networks,” *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, Sept. 1995, doi: 10.1007/BF00994018.
[92] V. Kecman, “Support Vector Machines – An Introduction,” in *Support Vector Machines: Theory and Applications*, L. Wang, Ed., Berlin, Heidelberg: Springer, 2005, pp. 1–47. doi: 10.1007/10984697_1.
[93] W. S. Noble, “A biologist’s introduction to support vector machines,” 2006.
[94] S. Ghosh, A. Singh, Kavita, N. Z. Jhanjhi, M. Masud, and S. Aljahdali, “SVM and KNN Based CNN Architectures for Plant Classification,” *Comput. Mater. Contin.*, vol. 71, no. 3, pp. 4257–4274, 2022, doi: 10.32604/cmc.2022.023414.
[95] G. Biau and E. Scornet, “A random forest guided tour,” *TEST*, vol. 25, no. 2, pp. 197–227, June 2016, doi: 10.1007/s11749-016-0481-7.
[96] M. Schonlau and R. Y. Zou, “The random forest algorithm for statistical learning,” *Stata J.*, vol. 20, no. 1, pp. 3–29, Mar. 2020, doi: 10.1177/1536867X20909688.
[97] B. K. Hatuwal, A. Shakya, and B. Joshi, “Plant Leaf Disease Recognition Using Random Forest, KNN, SVM and CNN”.

[98] C. Cortes and V. Vapnik, “Support-vector networks,” *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, Sept. 1995, doi: 10.1007/BF00994018.
[99] L. Breiman, “Random Forests,” *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, Oct. 2001, doi: 10.1023/A:1010933404324.
[100] R. A. Nugrahaeni and K. Mutijarsa, “Comparative analysis of machine learning KNN, SVM, and random forests algorithm for facial expression classification,” in *2016 International Seminar on Application for Technology of Information and Communication (ISemantic)*, Aug. 2016, pp. 163–168. doi: 10.1109/ISEMANTIC.2016.7873831.
[101] S. G. Wu, F. S. Bao, E. Y. Xu, Y.-X. Wang, Y.-F. Chang, and Q.-L. Xiang, “A Leaf Recognition Algorithm for Plant Classification Using Probabilistic Neural Network,” in *2007 IEEE International Symposium on Signal Processing and Information Technology*, Dec. 2007, pp. 11–16. doi: 10.1109/ISSPIT.2007.4458016.
[102] P. H. Davis, *Flora of Turkey*, vol. 3. Edinburgh: Edinburgh University Press, 1970.
[103] C. J. Doyle, D. J. Thomson, and J. E. Sheehy, “The future of sainfoin in British agriculture: an economic assessment,” *Grass Forage Sci.*, vol. 39, no. 1, pp. 43–51, Mar. 1984, doi: 10.1111/j.1365-2494.1984.tb01663.x.
[104] K. Erkoç, M. Işık, Ş. Erkol, A. Türkoğlu, and M. Sakiroglu, *Onobrychis Cinsinin Genetik Varyasyonunun Tanımlanması için Primer Arası Bağlanma Yerine (iPBS İşaretleyicileri) Dayalı PCR Analizinin Optimizasyonu*. 2024.
[105] J. K. Pritchard, M. Stephens, and P. Donnelly, “Inference of Population Structure Using Multilocus Genotype Data,” *Genetics*, vol. 155, no. 2, pp. 945–959, June 2000, doi: 10.1093/genetics/155.2.945.
[106] G. Evanno, S. Regnaut, and J. Goudet, “Detecting the number of clusters of individuals using the software structure: a simulation study,” *Mol. Ecol.*, vol. 14, no. 8, pp. 2611–2620, 2005, doi: 10.1111/j.1365-294X.2005.02553.x.
[107] R. Peakall and P. E. Smouse, “genalex 6: genetic analysis in Excel. Population genetic software for teaching and research,” *Mol. Ecol. Notes*, vol. 6, no. 1, pp. 288–295, 2006, doi: 10.1111/j.1471-8286.2005.01155.x.
[108] D. H. Huson and C. Scornavacca, “Dendroscope 3: An Interactive Tool for Rooted Phylogenetic Trees and Networks,” *Syst. Biol.*, vol. 61, no. 6, pp. 1061–1067, Dec. 2012, doi: 10.1093/sysbio/sys062.
[109] B. Ellis *et al.*, *Manual of Leaf Architecture*. GB: Cornell University Press CABI, 2009. doi: 10.1079/9781845935849.0000.
[110] L. J. Hickey, “Classification of the Architecture of Dicotyledonous Leaves,” *Am. J. Bot.*, vol. 60, no. 1, pp. 17–33, 1973, doi: 10.1002/j.1537-2197.1973.tb10192.x.
[111] M. Şakiroğlu and D. İlhan, “Medicago sativa species complex: Revisiting the century-old problem in the light of molecular tools,” *Crop Sci.*, vol. 61, no. 2, pp. 827–838, 2021, doi: 10.1002/csc2.20316.
[112] M. Şakiroğlu, J. J. Doyle, and E. Charles Brummer, “Inferring population structure and genetic diversity of broad range of wild diploid alfalfa (Medicago sativa L.) accessions using SSR markers,” *Theor. Appl. Genet.*, vol. 121, no. 3, pp. 403–415, Aug. 2010, doi: 10.1007/s00122-010-1319-4.
[113] D. İlhan, X. Li, E. C. Brummer, and M. Şakiroğlu, “Genetic Diversity and Population Structure of Tetraploid Accessions of the Medicago sativa–falcata Complex,” *Crop Sci.*, vol. 56, no. 3, pp. 1146–1156, 2016, doi: 10.2135/cropsci2015.12.0750.
[114] S. S. C. Kar, F. Ghanavati, M. R. Naghavi, H. Amirabadi-zade, and R. Rabiee, “Molecular phylogenetics of the Onobrychis genus (Fabaceae: Papilionoideae) using

ITS and trnL–trnF DNA sequence data," *Aust. J. Bot.*, vol. 62, no. 3, pp. 235–250, July 2014, doi: 10.1071/BT13279.
[115] R. B. Pinto, M. G. Lusa, V. de F. Mansano, A. M. G. de A. Tozzi, and J. L. S. Mayer, "Morphoanatomy of the leaflets of the Hymenaea clade (Fabaceae: Detarioideae) reveals their potential for taxonomic and phylogenetic studies," *Bot. J. Linn. Soc.*, vol. 187, no. 1, pp. 87–98, Apr. 2018, doi: 10.1093/botlinnean/boy007.
[116] V. Viscosi and A. Cardini, "Leaf Morphology, Taxonomy and Geometric Morphometrics: A Simplified Protocol for Beginners," *PLOS ONE*, vol. 6, no. 10, p. e25630, Oct. 2011, doi: 10.1371/journal.pone.0025630.
[117] A. Amirahmadi, S. Kazempour Osaloo, F. Moein, A. Kaveh, and A. A. Maassoumi, "Molecular systematics of the tribe Hedysareae (Fabaceae) based on nrDNA ITS and plastid trnL-F and matK sequences," *Plant Syst. Evol.*, vol. 300, pp. 729–747, Apr. 2014, doi: 10.1007/s00606-013-0916-5.
[118] R. J. Jensen, "The Conundrum of Morphometrics," *Taxon*, vol. 52, no. 4, pp. 663–671, 2003, doi: 10.2307/3647340.
[119] C. P. Klingenberg, "Morphological Integration and Developmental Modularity," *Annu. Rev. Ecol. Evol. Syst.*, vol. 39, no. Volume 39, 2008, pp. 115–132, Dec. 2008, doi: 10.1146/annurev.ecolsys.37.091305.110054.
[120] M. Zarrabian, M. H. Ehtemam, M. M. Majidi, and A. N. Dehkordi, "Assessment of the Relationship between Onobrychis Species through Morphological and Molecular Perspectives," *J. Rangel. Sci.*, vol. 15, no. 1, pp. 1–11, Jan. 2025, doi: 10.57647/j.jrs.2025.1501.07.
[121] A. Khan *et al.*, "Morpho-anatomical and palynotaxonomic study of the genus Onobrychis Miller (Hedysareae—Fabaceae) in Pakistan, and its systematic significance," *Phytotaxa*, vol. 650, pp. 121–147, May 2024, doi: 10.11646/phytotaxa.650.2.1.
[122] T. F. Stuessy, *Plant Taxonomy: The Systematic Evaluation of Comparative Data*. Columbia University Press, 2009.
[123] İ. Çuğu *et al.*, "Treelogy: A Novel Tree Classifier Utilizing Deep and Hand-crafted Representations," Jan. 28, 2017, *arXiv*: arXiv:1701.08291. doi: 10.48550/arXiv.1701.08291.
[124] S. Purohit, R. Viroja, S. Gandhi, and N. Chaudhary, "Automatic plant species recognition technique using machine learning approaches," in *2015 International Conference on Computing and Network Communications (CoCoNet)*, Dec. 2015, pp. 710–719. doi: 10.1109/CoCoNet.2015.7411268.
[125] K. Karbstein *et al.*, "Species delimitation 4.0: integrative taxonomy meets artificial intelligence," *Trends Ecol. Evol.*, vol. 39, no. 8, pp. 771–784, Aug. 2024, doi: 10.1016/j.tree.2023.11.002.
[126] M. M. A. Salles and F. M. C. B. Domingos, "Towards the next generation of species delimitation methods: an overview of machine learning applications," *Mol. Phylogenet. Evol.*, vol. 210, p. 108368, Sept. 2025, doi: 10.1016/j.ympev.2025.108368.