TÜRKİYE CUMHURİYETİ
ADANA ALPARSLAN TÜRKEŞ SCIENCE AND TECHNOLOGY UNIVERSITY

GRADUATE SCHOOL
BIOENGINEERING DEPARTMENT

GENOMIC AND PHENOMIC EVALUATION OF SAINFOIN (*Onobrychis* spp.) ACCESSIONS AND ASSESSING THEIR FORAGE AND SEED YIELD POTENTIAL

EBRAR KARABULUT UZUN

Ph.D.

ADANA 2025

TÜRKİYE CUMHURİYETİ
ADANA ALPARSLAN TÜRKEŞ SCIENCE AND TECHNOLOGY UNIVERSITY

GRADUATE SCHOOL
BIOENGINEERING DEPARTMENT

GENOMIC AND PHENOMIC EVALUATION OF SAINFOIN (*Onobrychis* spp.) ACCESSIONS AND ASSESSING THEIR FORAGE AND SEED YIELD POTENTIAL

EBRAR KARABULUT UZUN

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 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]

Ebrar KARABULUT UZUN

ABSTRACT

GENOMIC AND PHENOMIC EVALUATION OF SAINFOIN (*Onobrychis* spp.) ACCESSIONS AND ASSESSING THEIR FORAGE AND SEED YIELD POTENTIAL

Ebrar KARABULUT UZUN
Ph.D., Department of Bioengineering

Supervisor: Prof. Dr. Muhammet ŞAKİROĞLU
Co-Supervisor: Dr. Brandon SCHLAUTMAN
October 2025, 116 pages

Sainfoin is an important forage legume that has been suggested to be used as an alternative perennial pulse crop, as the grain is a rich source of nutrients and is currently under domestication. Success in breeding as a pulse crop or as a grain/forage dual use would require the use of available genetic resources effectively. This research aimed to evaluate the forage and seed yield potential of sainfoin (*Onobrychis* spp.) genetic resources in Tekirdağ, Turkey, in 2023 and in Salina, Kansas, USA, in 2024. The study also included the use of cutting-edge phenomic tools to test their performance while assessing the grain yield potential of sainfoin germplasm and genomic tools to address the genetic relationship among available germplasms. Agronomic and morphological traits related to forage and legume fruit and seed morphological traits were evaluated. Genomic analyses, such as the population structure, inter- and intra-group diversity of sainfoin species and accessions, were deduced using a set of genome-wide SNP markers. Significant species and accession-level variation was observed in both locations. Phenotypic analyses and evaluations conducted in both environments did not reveal a clear separation among species, and phenotypic overlaps were observed. Phenomic tools were used to analyze sainfoin fruit and seed-related traits, and the results indicated that these methods could help with rapid and accurate phenotypic measurements in sainfoin and other perennial grain crops. Genomic tools also revealed no clear clustering among the sainfoin germplasm used in this study.

Keywords: forage, population structure, yield, sainfoin, seed

ÖZET

KORUNGA (*Onobrychis* spp.) GENETİK KAYNAKLARININ GENOMİK, FENOMİK YÖNTEMLERLE İNCELENMESİ VE YEM VE TOHUM VERİMİ POTANSİYELİNİN DEĞERLENDİRİLMESİ

Ebrar KARABULUT UZUN

Doktora, Biyomühendislik Anabilim Dalı

Danışman: Prof. Dr. Muhammet ŞAKİROĞLU

II. Danışman: Dr. Brandon SCHLAUTMAN

Ekim 2025, 116 sayfa

Korunga, zengin bir besin kaynağı olması ve günümüzde evcilleştirilme aşamasında olması nedeniyle alternatif çok yıllık baklagil bitkisi olarak kullanılması önerilen önemli bir yem bitkisidir. Baklagil bitkisi olarak veya tahıl/yem bitkisi olarak çift amaçlı ıslahta başarı, mevcut genetik kaynakların etkin bir şekilde kullanılmasını gerektirir. Bu araştırma, 2023 yılında Türkiye, Tekirdağ'da ve 2024 yılında ABD, Kansas, Salina'da korunga (*Onobrychis* spp.) genetik kaynaklarının yem ve tohum verim potansiyelini değerlendirmeyi amaçlamıştır. Çalışma ayrıca, korunga germplazmının tane verim potansiyelini değerlendirirken performanslarını test etmek için son teknoloji fenomik araçların ve mevcut germplazmlar arasındaki genetik ilişkiyi ele almak için genomik araçların kullanımını da içermiştir. Yem ve baklagil meyve ve tohum morfolojik özellikleriyle ilgili agronomik ve morfolojik özellikler değerlendirilmiştir. Popülasyon yapısı, korunga türlerinin ve örneklerinin grup içi ve gruplar arası çeşitliliği gibi genomik analizler, genom çapında bir SNP belirteci kümesi kullanılarak çıkarıldı. Her iki lokasyonda da önemli türler ve örnek düzeyinde çeşitlilik gözlemlendi. Her iki ortamda yürütülen fenotipik analizler ve değerlendirmeler türler arasında net bir ayrım ortaya koymadı ve fenotipik örtüşmeler gözlemlendi. Korunga meyvesi ve tohumuyla ilgili özellikleri analiz etmek için fenomik araçlar kullanıldı ve sonuçlar, bu yöntemlerin korunga ve diğer çok yıllık tahıl ürünlerinde hızlı ve doğru fenotipik ölçümlere yardımcı olabileceğini gösterdi. Genomik araçlar ayrıca, bu çalışmada kullanılan korunga germplazmında net bir kümelenme ortaya koymadı.

Anahtar Kelimeler: korunga, popülasyon yapısı, tohum, verim, yem

Dedication

I dedicate this thesis to all women and girls in the world and in our country who are exposed to gender inequality and who struggle for an equal future.

ACKNOWLEDGEMENTS

I extend my endless gratitude and appreciation to my esteemed advisor, Prof. Dr. Muhammet Şakiroğlu, for his guidance, direction, and assistance in every aspect. I would like to extend my sincere gratitude to my thesis monitoring committee members and thesis defense jury members, Prof. Dr. Zeynel Cebeci and Asst. Prof. Dr. Özgür Cem Erkin, for their advice and contributions throughout the writing of my thesis. I also extend my gratitude to the thesis defense jury members, Assoc. Prof. Dr. Esra Göv and Asst. Prof. Dr. Orhan Uluçay, for their contributions to the thesis and their valuable evaluations.

At my first location, Namık Kemal University (Tekirdağ/Türkiye), I also extend my gratitude to Prof. Dr. Metin Tuna for his assistance and support, and I would like to thank his invaluable students who helped me during the measurement process. At my second location, The Land Institute (Salina, Kansas, USA), I extend my endless gratitude to my second advisor, Dr. Brandon Schlautman, for providing me with unwavering support and assistance throughout my research. I also extend my gratitude to Spencer Barriball, who provided me with every assistance and guidance I needed. Furthermore, I would like to thank my other colleagues at The Land Institute, who supported and assisted me during my greenhouse and field work. I extend my sincere gratitude to my colleague Kübra Erkoç, who assisted me both in Tekirdağ and Kansas during my research.

I extend my endless gratitude to my precious mother, Hülya Karabulut; my father, Sinan Karabulut; and my dear sisters and brothers, who have always supported me throughout my education. I also extend my endless gratitude to my husband, my soulmate, and my dearest lover, Muharrem Uzun, for his unwavering support, assistance, and encouragement throughout my thesis writing process.

This thesis was produced as part of the project titled "Developing Methods and Infrastructure for Genetic Mapping in Autotetraploid and Allogamous Plants: The Model Plant Sainfoin" (project number 120C136) and was supported by TÜBİTAK.

TABLE OF CONTENTS

DECLARATION OF CONFORMITY i
ABSTRACT ii
ÖZET iii
Dedication iv
ACKNOWLEDGEMENTS v
TABLE OF CONTENTS vi
LIST OF FIGURES viii
LIST OF TABLES ix
LIST OF ABBREVIATIONS x
LIST OF SYMBOLS xiii
1. INTRODUCTION 1
2. LITERATURE REVIEW 3
2.1. Sainfoin 3
2.2. Origin and historical distribution 4
2.3. Taxonomy 5
2.4. Morphological properties 6
2.5. Polyploidy and genome size 6
2.6. Breeding sainfoin for forage and grain 9
2.7. Importance of yield in sainfoin 11
2.8. Phenomic studies of sainfoin 13
2.9. Genomic tools to accelerate breeding 14
2.10. Population genomics studies of sainfoin 16
3. MATERIAL AND METHOD 20
3.1. Plant Materials/Panel 20
3.2. Experimental Design of Field Trials 28
3.3. Phenotyping Data Collection 29
3.4. Phenomic Data Collection 40
3.5. Genotyping Data Collection 44
3.6. Data Analyses 45
3.6.1. Phenotyping Data Analyses 45
3.6.2. Phenomic Data Analyses 48

3.6.3. Genomic Data Analyses 50
4. RESULTS AND DISCUSSIONS 52
4.1. Evaluation of traits related to forage and seed yield 52
4.1.1. Effects of Environments 52
4.1.2. Trait means and Correlations among traits 55
4.1.3. Phenotypic Variation among Sainfoin Accessions 64
4.1.4. Heritability of traits 71
4.1.5. Accession and Species Clustering 72
4.2. Phenomic Analyses of the Sainfoin Diversity Panel 79
4.2.1. Effects of Environment 79
4.2.2. Trait means and Correlations among traits 79
4.2.3. Phenotypic Variation among Sainfoin Accessions 83
4.2.4. Heritability of the traits 85
4.2.5. Accession and Species Clustering 85
4.3. Genomic Analyses of the Sainfoin Diversity Panel 90
4.3.1. Genotyping Results 90
4.3.2. Population Structure 90
4.3.3. Phylogenetic and Distance Analyses 92
4.3.4. Molecular Variance Analysis (AMOVA) 94
5. CONCLUSION 96
REFERENCES 98
APPENDIX

LIST OF FIGURES

Figure 3.1. Taking measurement of phenotypic traits in the Tekirdağ, Türkiye 30
Figure 3.2. Transplantation of sainfoin seedlings to the field in the Salina, Kansas 38
Figure 3.3. Taking measurement of phenotypic traits in Salina, Kansas, USA 39
Figure 3.4. Legume fruit samples of three economically important sainfoin species and one of the cultivars during the threshing process in Salina, KS 41
Figure 4.1. Pearson correlation plot of agronomic traits in sainfoin accessions in the Tekirdağ-23 environment 61
Figure 4.2. Spearman correlation plot of agronomic traits in 80 sainfoin accessions in the Salina-24 environment 63
Figure 4.3. Clustering of dendrogram of sainfoin accessions/breeding populations/ cultivars in Tekirdağ-23 environment for agronomic traits 72
Figure 4.4. PCA plot of 47 sainfoin accessions/breeding populations/cultivars in the Tekirdağ-23 environment for agronomic traits 73
Figure 4.5. Clustering of dendrogram of sainfoin accessions/cultivars in the Salina-24 environment for agronomic traits 75
Figure 4.6. PCA plot of 80 sainfoin accessions/cultivars in the Salina-24 environment for agronomic traits 76
Figure 4.7. Performance comparison of 6 cultivars and 13 accessions in terms of two selected traits related to forage yield and seed yield 78
Figure 4.8. Pearson correlation plot of estimated seed traits in 80 sainfoin accessions in the Salina-24 environment 82
Figure 4.9. Clustering of dendrogram of sainfoin accessions/cultivars in the Salina-24 environment for estimated seed related traits 87
Figure 4.10. Principal Component Analysis (PCA) plot of 80 sainfoin accessions/cultivars in the Salina-24 environment for estimated seed related traits 89
Figure 4.11. Principal Component Analysis (PCA) plot of 43 sainfoin populations based on sainfoin species in the Tekirdağ-23 environment 91
Figure 4.12. Neighbor-joining tree of 43 sainfoin accessions obtained from genotyping data based on sainfoin species and breeding populations in the Tekirdağ-23 environment 93

LIST OF TABLES

Table 3.1. List of sainfoin accessions/cultivars/breeding populations in Tekirdağ, Türkiye location 21
Table 3.2. List of sainfoin accessions/cultivars in Salina, Kansas, USA 24
Table 3.3. Crop Ontology List of traits associated with forage and seed yield measured in Tekirdağ-23 environment 31
Table 3.4. Crop Ontology list of traits associated with forage yield measured in Salina-24 environment 34
Table 3.5. Crop Ontology list of traits associated with seed yield measured in Salina-24 environment 42
Table 4.1. F-values along with the significance from the analysis of variance (ANOVA) for common agronomic traits measured across two different environments 52
Table 4.2. F-values along with the significance from the analysis of variance (ANOVA) of agronomic traits in the Tekirdağ-23 environment 53
Table 4.3. F-values along with the significance from the analysis of variance (ANOVA) and H^2 heritability of forage-related traits in the Salina-24 environment 54
Table 4.4. The comparison of mean, standard deviation and their ranges of agronomic traits in both environments 55
Table 4.5. Descriptive statistics (mean, minimum, and maximum) of estimated marginal means (emmeans) with standard errors of agronomic performance based on sainfoin species and showing their statistical differences in the Tekirdağ-23 environment 56
Table 4.6. Descriptive statistics (mean, minimum, and maximum) of estimated marginal means (emmeans) with standard errors of agronomic performance based on sainfoin species and showing their statistical differences in the Salina-24 environment 58
Table 4.7. Means and standard deviations of estimated marginal means (emmeans) of the common traits in the common accessions across two environments 65
Table 4.8. Top performing sainfoin accessions and cultivars for forage and seed yield traits in the Salina-24 environment 67
Table 4.9. Ranking of sainfoin accessions/cultivars/breeding populations based on forage- and seed-related traits in both environments 68
Table 4.10 F-values along with the significance from the analysis of variance (ANOVA) and H^2 heritability of seed-related traits in the Salina-24 environment 79
Table 4.11. Descriptive statistics (mean, minimum, and maximum) of estimated marginal means (emmeans) with standard errors of agronomic performance of estimated seed traits based on sainfoin species and showing their statistical differences in the Salina-24 environment 80
Table 4.12. Ranking of 80 Sainfoin Accessions for Estimated Seed Traits in the Salina-24 environment 83
Table 4.13. Analysis of molecular variance (AMOVA) of 43 sainfoin accessions, including four sainfoin groups (*O. viciifolia, O. transcaucasica, O. arenaria*, cultivars/breeding populations) in the Tekirdag-23 environment 94

LIST OF ABBREVIATIONS

AF	: Amaranth Flower
AFLP	: Amplified Fragment Length Polymorphism
AI	: Artificial Intelligence
AMOVA	: Analysis of Molecular Variance
ANOVA	: Analysis of Variance
Bp	: Base pair
CanHt	: Canopy Height
CIS	: Commonwealth of Independent States
CLD	: Compact Letter Display
CO	: Crop Ontology
CSV	: Comma-Separated Values
CT	: Condensed Tannins
DEGs	: Differentially Expressed Genes
DM	: Dry Matter
DMY	: Dry Matter Yield
DNA	: Deoxyribonucleic Acid
Emmeans	: Estimated Marginal Means
EST-SSRs	: Expressed Sequence Tags derived Simple Sequence Repeats
FFY	: Fresh Forage Yield
FoDM	: Forage Dry Matter
FoSDM	: Forage Subsample Dry Matter
Fr1KK	: Legume Fruit Thousand Kernel Weight
FrAR	: Legume Fruit Aspect Ratio
FrEx	: Legume Fruit Extent Ratio
FrL	: Legume Fruit Length
FrM	: Legume Fruit Mass
FrPA	: Legume Fruit Projected Area
FrW	: Legume Fruit Width

GBS	: Genotyping By Sequencing
GWAS	: Genome-Wide Association Studies
HSD	: Honest Significant Difference
InfL	: Inflorescence Length
iPBS	: Inter-Primer Binding Site
ISSR	: Inter-Simple Sequence Repeats
KEGG	: Kyoto Encyclopedia of Genes and Genomes
MAB	: Marker-Assisted Breeding
MAS	: Marker-Assisted Selection
MCDM	: Multiple-Criteria Decision Making
NA	: Non-Available
NGS	: Next Generation Sequencing
NJ	: Neighbor Joining
PCA	: Principal Component Analysis
PlRGR	: Plant Regrowth Rating
PlStEr	: Plant Stem Erectness
QTL	: Quantitative Trait Loci
RaL	: Raceme Length
RaN	: Raceme Number
RAPD	: Random Amplified Polymorphic DNA
RNA	: Ribonucleic Acid
RT-qPCR	: Reverse Transcription Quantitative Real-Time Polymerase Chain Reaction
SD	: Standard Deviation
SdAR	: Seed Aspect Ratio
SdEx	: Seed Extent Ratio
SdL	: Seed Length
SdN	: Seed Number
SdPA	: Seed Projected Area
SdW	: Seed Width
SNP	: Single Nucleotide Polymorphism
SRAP	: Sequence-Related Amplified Polymorphism

SSR	: Simple Sequence Repeats
StL	: Stem Length
StN1	: First cutting Stem Number
StN2	: Second cutting Stem Number
StTh	: Stem Thickness
TAGEM	: General Directorate of Agricultural Research and Policies
TOPSIS	: Technique for Order Preference by Similarity to an Ideal Solution
TÜBİTAK	: The Scientific and Technological Research Council of Türkiye
TurkStat	: Turkish Statistical Institute
UK	: United Kingdom
USA	: United State of America
WF	: White Flower

LIST OF SYMBOLS

%	: Percent
°C	: Celsius
°F	: Fahrenheit
cm	: Centimeter
g	: Gram
H^2	: Heritability
kg/ha	: Kilograms per hectare
m/s	: Meter per second
mm	: Millimeter
mm^2	: Square millimeter
N	: Nitrogen
t ha-1	: Tons per hectare
µl	: Microliter
σ^2E	: Residual variance
σ^2G	: Genotypic variance
σ^2P	: Total phenotypic variance

1. INTRODUCTION

Food of animal origin is of great importance for humans to have a balanced, adequate, and healthy diet. Feed costs constitute about 70% of the total animal production cost and significantly affect the profitability of the enterprise. For this reason, forages are of great importance because they are relatively cheaper and positively affect the digestive activities of ruminant animals (S. Özkan & Bilgen, 2019; U. Özkan & Demirbağ, 2016).

Turkish forage and feed production fails to meet the demand (Cevher, 2023), so meadows and pastures are used to meet the demand gap. To achieve the goal, the cultivation areas of forage crops need to be expanded in Türkiye. It is necessary to develop forage crop varieties that can adapt to different ecological regions of Türkiye and provide efficient and high-quality forage. Sainfoin is an important plant that could help provide necessary protein to fam animals as it is resilient to various abiotic constraints (Avcı, 2010).

Sainfoin (*Onobrychis* ssp.) is a useful perennial forage plant that can be grown in different ecological conditions and several species are domesticated as forage including *Onobrychis viciifolia* Scop., *Onobrychis transcaucasica* Grossh., *Onobrychis altissma* Grossh., and *Onobrychis arenaria* (Kit.) DC. The genus *Onobrychis* is naturally distributed in Iran, Anatolia, and northern temperate regions of Europe and Asia (Ranjbar, Karamian, & Hajmoradi, 2010; Yıldız, Çıplak, & Aktoklu, 1999; Zarrabian, Majidi, & Ehtemam, 2013) but used as a forage legume in across the northern hemisphere (Cash & Ditterline, 1996). Sainfoin cultivation in Turkey is mostly done in Eastern Anatolia and Central Anatolia in Türkiye (Özbek, 2011; S. Özkan & Bilgen, 2019; Temel & Özalp, 2016).

Sainfoin has several advantages compare to other forages such as preventing rapid protein digestibility in cattle rumen, preventing bloating and limiting methane and ammonia emissions from cattle (Legesse et al., 2015; S. C. Sheppard & Bittman, 2012; Steve C. Sheppard et al., 2019). Sainfoin cultivars show great variation and variability in terms of winter hardiness, maturity, yield potential, and many other important factors. Nonetheless, research or breeding on sainfoin has been limited and the large existing variation in the geography and climate zones are largely unexplored (Carbonero, Mueller-Harvey, Brown, & Smith, 2011).

Moreover, genetic information of the crop and wild relatives is quite limited. Due to the decreasing cost of sequencing, sequence-based genome-wide genotyping techniques provide access to a high number of uniform SNP markers across the genome that could be assessed in a high number of individuals.

Recently, sainfoin has been suggested as a perennial pulse crop and being under domestication for human consumption. It is being domesticated as perennial pulse crop Baki Bean™(Craine, Barriball, Şakiroğlu, Peters, & Schlautman, 2024). However, improvement as forage or as a pulse crop require coordinated efforts for evaluation of a broad-based germplasm across all the cultivated species along with a thorough genomic evaluation of the germplasm. Particularly, seed yield potential of a broad range of available germplasm could help to understand the degree of variation and identify germplasm that has high potential for inclusion in breeding programs. Nonetheless, seed yield related traits are rather cumbersome to measure accurately. Utilization of the state-of-art techniques such as image analyses along with the artificial intelligence (AI) can help precisely estimate these difficult traits in many genotypes and accessions.

The aims of this study were to investigate the genetic diversity of the sainfoin populations of three species (*Onobrychis viciifolia, Onobrychis transcaucasica*, and *Onobrychis arenaria*) and to determine the population structure, to determine forage and seed yield potential of the selected germplasm panel in replicated field trials, and to test use of image analyses to collect fruit and seed yield related traits in sainfoin.

This thesis was conducted in two different years in two ecologically and geographically distinct environments to evaluate a broad-based global sainfoin germplasm for genomic diversity and forage and pulse related traits using genomic and phenomic tools.

2. LITERATURE REVIEW

2.1. Sainfoin

Sainfoin is the most suitable forage legume to be grown in barren and calcareous soils (Özbek, 2011; S. Özkan & Bilgen, 2019; Temel & Özalp, 2016; F. Tosun, 1992). Well-drained soils are preferred for sainfoin and do not grow well in heavy soils or stagnant water encountered under flood irrigation (Anderson, 2025; Wijekoon et al., 2021). Sainfoin can be used as a monoculture or mixed with grass or alfalfa (Acharya, 2015; Sottie et al., 2014).

Recent studies have shown that sainfoin has anthelmintic properties, the potential to control methane in ruminants, and the ability to protect protein against premature spoilage in the rumen when used as a forage crop in ruminant diets. These features mainly point to the tannin composition of the sainfoin in the leaf (Lorenz, Eriksson, & Udén, 2010; M. Mora-Ortiz & Smith, 2018; Novobilský, Mueller-Harvey, & Thamsborg, 2011; Pellikaan et al., 2011; Theodoridou et al., 2011).

Sainfoin is rich in potassium, calcium, magnesium, phosphorus, zinc, and iron minerals, especially in the stem elongation stage (Butkutė, Padarauskas, Ceseviciene, Taujenis, & Norkevičienė, 2018; Wijekoon et al., 2021). In addition, sainfoin is rich in secondary metabolites, which are beneficial compounds. These secondary metabolites are concentrated in various parts of the sainfoin and are known to benefit animal and human health (R. P. Barrett, 1990; Butkutė et al., 2018; Wijekoon et al., 2021). Cultivated sainfoin varieties are very important and tasty for ruminants and have a crude protein content (Acharya, 2015, 2018; Wijekoon et al., 2021). The nutritional quality parameters of sainfoin have been found to exceed all standard values recommended for ruminant animals (Turk, Albayrak, Tuzun, & Yuksel, 2011). Sainfoin is important not just for ruminants, but also it is important as a good source of nectar and pollen for many pollinator species such as honeybees and bumblebees, flying flies, and solitary bees (Hayot Carbonero, Mueller-Harvey, Brown, & Smith, 2011; Kells, 2001; M. Mora-Ortiz & Smith, 2018). Sainfoin starts vegetatively in early spring and remains green throughout the summer (Monsen, Stevens, & Shaw, 2004; Wijekoon et al., 2021). Because the sainfoin grows upright and the first-cut hay yield of the sainfoin is generally higher than that of clover, it is preferred for haying; however, later dry matter (DM) yield is mostly equal in new

cultivars or 20% to 30% less than alfalfa in old sainfoin cultivars. Therefore, sainfoin can be considered a short-lived forage crop and adapted for early-season harvest or grazing (Acharya, 2015; Steve C. Sheppard et al., 2019).

2.2. Origin and historical distribution

The origin center of the sainfoin is South Central Asia, and the sainfoin appears to be widespread in Asia Minor, the Anatolian Plateau of Türkiye, and as part of mixed grasses on the Caucasus and Caspian margins. It was also originally cultivated by Arabian cultures (De Candolle, 1883; M. Mora-Ortiz & Smith, 2018).

Sainfoin has a relatively long history, and various ethnobotanical sources suggest different domestication and cultivation history. Turkish authors suggest different times for the domestication of that crop. For instance, (Açıkgöz, 1991) suggest that sainfoin has been cultivated for more than 1000 years in Russian and the former British Empire hinterlands while (Elçi, 2005) reported that sainfoin farming started between 1450-1550. On the other hand, one paper citing an old Armenian botanical book (Alishan, 1895) dated the cultivation in Armenia in the 10th century (Akopian, 2009). It is understood that sainfoin has been used as a cultivated plant since the beginning of the last millennium. The exact cultivation dates in Europe are much more recent. It was first cultivated in 1582 in Southern France (Piper, 1914), and in 1560 in Southern Germany. Sainfoin was cultivated in the Netherlands and England in 1650 (Elçi, 2005; Genckan, 1992).

During the 18th century, sainfoin spread across Europe and became an important source of roughage for ruminants in countries such as France, Italy, the UK, Germany, and Switzerland and later was also introduced into the USA (Kölliker, Kempf, Malisch, & Lüscher, 2017; Piper, 1914). It is recorded that the sainfoin was taken to the United States at the beginning of the 19th century (Elçi, 2005; Genckan, 1992). In the middle of the 19th century, the two-harvest sainfoin was brought to England in 1884, and the giant sainfoin in 1850 (Elçi, 2005). It is said to have been grown on a central research farm in Ottawa in the late 19th century (Elçi, 2005). Later, since the alfalfa weevil (*Hypera variabilis* Hbst.) causes great damage to the alfalfa, importance has been given to the sainfoin, which is resistant to this pest. Today, sainfoin is grown in many

countries of Europe, the arid and semi-arid regions of the CIS, the USA, and Canada (Açıkgöz, 2001).

Its spread to Europe had a significant impact due to the profitable cultivation of very dry calcareous soils that were previously almost worthless. The culture of the sainfoin has been largely confined to chalky or other calcareous soils, especially when exposed to drought (Piper, 1914). Sainfoin is still grown as an important crop in parts of Asia, Türkiye, and Iran. In addition, sainfoin cultivation is currently continuing in parts of North America, Italy, and Spain (Koivisto & Lane, 2001; M. Mora-Ortiz & Smith, 2018).

2.3. Taxonomy

Onobrychis viciifolia belongs to the tribe Hedysareae of the subfamily Papilionoideae of the family Fabaceae, and the genus *Onobrychis* Mill (Hayot Carbonero et al., 2011). The phylogenetic analysis of a wide range species including *Arabidopsis thaliana*, soybean, common bean, *Lotus japonicus*, *Medicago truncatula*, sainfoin, and red clover based on the protein sequences of approximately 50 single copy number revealed that sainfoin was most closely related to red clover and *Medicago truncatula*, while it was more distantly related to *Lotus japonicus*, common bean, and soybean (Marina Mora-Ortiz et al., 2016).

There are many contradictions and inconsistencies in the taxonomy of *Onobrychis*, primarily due to different approaches to species delimitation, resulting in several recognized species (Emre, Turgut-Balik, Sahin, & Kursat, 2007; Hayot Carbonero et al., 2011). There are about 162 *Onobrychis* genus with annual and perennial species in the world (Aktoklu, 1995; Pulluk Uğur, 2022; Yücel, 2019). Also, another study suggested that, based on fruit morphology, the genus *Onobrychis* contains about 170 species. It is divided into two subspecies, Sisyrosema and *Onobrychis*, and eight parts (Hayot Carbonero et al., 2011; Yıldız et al., 1999). Although sainfoin is grown as a forage plant around the world, the main centers of species diversity are known as Anatolia (total 52 species), Iran (about 53 species), and Caucasus (up to 39 species) (Çelik Oğuz, Karakaya, Avcı, Sancak, & Özcan, 2011; Şakiroğlu, 2021).

2.4. Morphological properties

One of the important traits in sainfoin is the plant height and the earlier reports indicate a wide variation among species and accessions. For instance, *O. transcaucasica* reported to have a plant height of 60 cm whereas *O. altissima* reaches 90 cm (Akopian, 2009). Generally, a variation of 30-70 cm in height has been reported (Tarasenko, Butina, & Gerasimenko, 2015). The sainfoin has fleshy and hollow erect stems. As perennial plants have well developed root systems, sainfoin roots are thick, deep, and branched roots with many thin lateral roots (M. Mora-Ortiz & Smith, 2018). The depth of the taproot system could reach up to three meters (Tarasenko et al., 2015).

When considered as a temperate forage legume, leaves have utmost-paramount importance as they contain high protein for livestock. The compound sainfoin leaves alternate on the stem and consist of 11-29 leaflets that are narrow at the apex and broad at the base, with a single terminal leaflet at the apex and pairing the entire leaf except for a single terminal leaflet at the tip (M. Mora-Ortiz & Smith, 2018).

The sainfoin flower heads have about 80 rosy-pink flowers attached to short stems that connect them to the main stem (M. Mora-Ortiz & Smith, 2018; Teuber & Brick, 1988; Wijekoon et al., 2021). The sainfoin flowers have a purple, pink, whitish or yellowish color and the flowers are collected in auxiliary racemes or spikes (Tarasenko et al., 2015). The flowers in the panicle open from the bottom up (Açıkgöz, 2001). The fruit is a semicircular, small flat pod with a single seed. The fruit coat surrounds the seed. The semicircular edge of the fruit is thorny. Sainfoin seeds are bean-shaped and dark brown. 1000 grain weight of seeds is 17-32 g (Açıkgöz, 2001).

2.5. Polyploidy and genome size

Polyploidy is a mechanism in plant genome that results a genome-wide multiplication of chromosomes (Bretagnolle, Thompson, & Lumaret, 1995; Kempf, 2016). Polyploidy is common in the plant kingdom (Şakiroğlu, 2021; P. Soltis & Doyle, 2012). The nature and role of polyploidy have been extensively investigated, thanks to the availability of many diverse genomic tools (Z. J. Chen & Birchler, 2013; Şakiroğlu, 2021).

There are two different types of polyploidies: autopolyploidy and allopolyploidy (Kempf, 2016). Autopolyploids usually arise through the production of unreduced gametes resulting a genomic duplication within a single species (Harlan & DeWet, 1975). During meiosis, pairing and recombination can occur between all homologous copies of each chromosome. Therefore, autopolyploids exhibit polysomic inheritance (Bourke, Voorrips, Visser, & Maliepaard, 2018; Kempf, 2016). One of the best-studied examples of autotetraploid is the potato (*Solanum tuberosum* L.) (Bourke, Voorrips, et al., 2018). Allopolyploids are the product of hybridization between different species followed by genomic duplication also through the hybridization of unreduced gametes (Harlan & DeWet, 1975). Unlike autopolyploids, allopolyploids exhibit disomic inheritance, in which chromosome pairing and recombination occur within each genome during meiosis (Bourke, Voorrips, et al., 2018; Kempf, 2016). One of the best-studied allopolyploids is probably allohexaploid wheat (*Triticum aestivum* L.) (Bourke, Voorrips, et al., 2018).

Polyploid plants generally have an advantage over diploids in that they have larger flowers and seeds due to increased plant growth and vigor (Bretagnolle et al., 1995; Kempf, 2016). Perennial temperate legumes are largely polyploid. Autopolyploidy is predicted to evolve to some degree to behave like disomic inheritance (D. E. Soltis & Soltis, 1993).

Although potentially poses challenges in genomic studies and breeding efforts, autopolyploidy is reported to be a positive trait as it has the capacity to retain a higher number of beneficial mutant alleles, resulting in wider environmental adaptation (Annicchiarico, Barrett, Brummer, Julier, & Marshall, 2015; Bingham, Groose, Woodfield, & Kidwell, 1994; Gallais, 1981; Şakiroğlu, 2021).

A number of different *Onobrychis* species with basic chromosome numbers of either x = 7 or x = 8 were reported (Akçelik, Avcı, Uzun, & Sancak, 2012; Somay Akçelik, Avcı, Uzun, & Sancak, 2012; Yucel et al., 2022) and mostly two ploidy levels are present in the genus as 2n=2x=14, 2n=4x=28, 2n=2x=16, 2n=4x=32 (Abou-El-Enain, 2002; Hayot Carbonero, Carbonero, Smith, & Brown, 2013; Somay Akçelik et al., 2012).

In a study conducted in Iran, cytogenetic analysis of wild sainfoin populations was performed. Among these agriculturally important species, *O. viciifolia* and *O. altissima* were tetraploid (2n=4x=28), while *O. transcaucasica* was found to be diploid (2n=2x=14) (Massoud, Karamian, & Hadadi, 2010).

Core DNA content was found to be more dynamic in diploid species than in polyploid species (Yucel et al., 2022). The average 2C core DNA contents of the 3 sainfoin species (*O. transcaucasica*, *O. viciifolia*, and *O. arenaria*) were determined as 2.65 (2.53 pg - 2.77 pg), 2.64 (2.49 pg - 2.78 pg), and 2.63 (2.52 pg - 2.75 pg) respectively (Şahin, Yılmaz, Gülsemin Savaş Tuna, Yücel, & Tuna, 2019). These three sainfoin species were determined as quite similar in terms of mean value and variance. When the nuclear DNA content of the accessions was correlated with the chromosome numbers, all the accessions analyzed in the study were determined tetraploid with 2n=4x=28 chromosome numbers (Şahin et al., 2019).

One study examined the evolution of chromosome number and ribosomal DNA (rDNA) loci number and localization in *Onobrychis* Mill. According to the analysis of the evolution of basic chromosome numbers, ancestral chromosome number was observed as x = 8 and decreasing dysploidy and polyploidization were observed as the main mechanisms of chromosome number evolution (Yucel et al., 2022).

The complexity of genome analysis in polyploids complicates the effective use of GWAS and genomic selection. The reasons for these difficulties are the fundamental differences in gene segregation and recombination patterns between diploids and tetraploids, and the high level of heterozygosity of autotetraploid due to the outbreeding nature (J. Chen et al., 2021; Şakiroğlu, 2021). Two strategies have been developed over time to solve these challenges (Sakiroglu & Brummer, 2017; Şakiroğlu, Doyle, & Charles Brummer, 2010).

Initially, before the widespread use of genomic tools, diploid counterparts or relatives of these polyploids were used to avoid polyploid complexes. Second, the widespread use of genomic tools and the development of specific statistical methods to study polyploid genome structure (Doerge & Craig, 2000; Wu, Ma, & Casella, 2002; Xie & Xu, 2000). However, many robust statistical methods and software tools have been developed and applied to enable appropriate

estimation of the allelic dosage of SNPs in polyploid organisms where the allelic dosage effect is important (M. Chen et al., 2021; Ferreira et al., 2019).

2.6. Breeding sainfoin for forage and grain

O. viciifolia generally has a self-incompatibility system and is considered an outbreeding species (Hayot Carbonero et al., 2011). *O. viciifolia* is one of the insect-pollinated species. It successfully is pollinated flowers by many insect species. The most important insect species in pollination are *Apis mellifera* (honeybee) and *Bombus* sp. (bumblebees) and the relatively less important insect species *Osmia* (solitary). It is possible to self-pollinate sainfoin plants under controlled conditions, but the resulting plants lack vigor and produce few seeds if the plants are viable. *Onobrychis*, like many other members of the Fabaceae, is described as an obligate insect-pollinating species (Hanley, Franco, Pichon, Darvill, & Goulson, 2008; Hayot Carbonero et al., 2011).

In a study of the use of sainfoin seeds in the diet of piglets, sainfoin seeds with and without hulls were added to a control diet for weaned piglets instead of using pea and soybean cake. Sainfoin seeds were found to provide an equally adequate source of amino acids as the control diet when added at 10% and 10-16%, respectively. Thus, sainfoin seeds can be a valuable feed component for animals owing to the higher protein content (Baldinger, Hagmüller, Minihuber, Matzner, & Zollitsch, 2016).

Previous studies have revealed that sainfoin seeds can be used as new food sources for humans (Butkutė et al., 2018; Wijekoon et al., 2021). Analyses of chemical composition of sainfoin seeds revealed that it contains 28.7% protein and 19.4% fiber, while the oil content did not exceed 8%. The total amount of amino acids in the sainfoin seed powder was 26.94 g/100 g of product, and the ratio of indispensable amino acids was 37.85%. The limiting amino acid for the protein complex of sainfoin powder was tryptophan, with a ratio of 48.0%. Emprical studies indicated that sainfoin seed powder has a more balanced amino acid composition than wheat flour and peanuts. For this reason, sainfoin seeds can be considered a very promising food ingredient and help to adjust the amino acid balance, which limits its use in food composition (Tarasenko et al., 2015).

There are two agricultural types of sainfoin: common sainfoin (*O. sativa* var. communis Ahlefed) and giant sainfoin (*O. sativa* var bifera Hort.). The common type is perennial, but less vigorous in its founding year, whereas in the UK the giant type is short-lived. The common type has more stems per plant than the giant type, but the plant height is generally shorter (Hayot Carbonero et al., 2011; Thomson, 1951). The common type in Europe is also known as the single-cut type, but the giant type is used for a double-cut system. The common type comes from Central Europe, while the giant type comes from the Middle East. Crops from Western Europe were mostly giant type and low yielding and poorly adapted to North America. With the introduction of the common type from the former Soviet Union and Türkiye, the common type performed relatively well in Western Canadian growing conditions (Goplen, Richards, & Moyer, 1991). Released in 1969, cv. Melrose is the first Canadian sainfoin cultivar, selected from accessions brought by the former Soviet Union (Bhattarai, 2017; Hanna, Cooke, Smoliak, & Goplen, 1972).

It is an important breeding goal to develop varieties with condensed tannin content in sainfoin to increase the current potential for ruminant animals. CT concentrations were measured in Communis and Bifera accessions in a study that tested whether accession and environment affected CT structures in the sainfoin. It was concluded that CTs in Bifera accessions have more stable and predictable properties and can be considered in breeding studies (Azuhnwi et al., 2013).

Numerous studies have been conducted both in Türkiye and worldwide to investigate the yield and quality characteristics of sainfoin (*Onobrychis sativa* L.), with the aim of developing genotypes with higher dry matter yield and improved forage quality (Çaçan, Kökten, & Koç, 2023; Sayar, Han, & Basbag, 2022; Turk et al., 2011). In a study of seed size in sainfoin, it was observed that seedlings from larger seeds emerged faster, had more nodules, and had higher N_2 fixation rates (Cash & Ditterline, 1996).

Forage yield and forage quality, which are crucial traits, are also important for sainfoin as they are for all forage crops. In a study where the forage yield and forage quality of sainfoin genotypes were measured according to years, it was reported that it was statistically significant ($P < 0.01$) only in terms of fresh forage and dry matter yield traits. Considering the three-year

range of the traits of the sainfoin genotypes, fresh forage yield (FFY) of 18.09- 37.70 t ha^{-1} and dry matter yield (DMY) of 4.74-9.63 t ha^{-1} were obtained (Sayar et al., 2022).

The genetic information of the sainfoin, which is one of the important forage crops, is quite limited. In a study evaluating the genetic diversity and trait relationships of 80 sainfoin species, it was stated that the anatomical traits were less affected by environmental constraints than the agro-morphological traits. It has been reported that most of the genetic variance of sainfoin species is found between accessions and less between geographical groups. In addition, the results of the study claimed that Asia and Eastern Europe could be the main diversity centers for the sainfoin (Zarrabian et al., 2013).

Environmental factors such as altitude, climate, and temperature changes have a significant impact on traits related to sainfoin yield (Çaçan et al., 2023). In addition to annual temperature changes, genotypic diversity has also been shown to have a significant impact on both forage yield and quality traits of sainfoin (Sayar et al., 2022).

2.7. Importance of yield in sainfoin

In the sainfoin fields planted with barren, very little flowering is seen in the first year. The main seed product in sainfoin is taken after the second year. Seed yield is quite high. 300-600 kg/ha of seeds can be taken in arid conditions. In very good fields, this yield can exceed 1000 kg/ha (Açıkgöz, 2001). In early sown, well fertilized and irrigated fields, seed crops can be obtained in the sowing year. In irrigated conditions, the yield increases up to 2000 kg/ha (Açıkgöz, 2001). Recently, sainfoin has been reevaluated as a forage crop due to its positive properties that lead to highly satisfactory animal performance (Guglielmelli et al., 2011).

Sainfoin production is done on aboveground dry and fresh biomass, and seed production. Forage dry matter (DM) yield is the most basic objective in the breeding programs of sainfoin. There was high variability in weed DM yield among sainfoin populations in many studies. A report was obtained a total annual forage yield of 10.79 tons ha^{-1} DM in three sections in the evaluation of 34 sainfoin populations under irrigation conditions in Iran (Nakhjavan, Bajolvand, Jafari, & Sepavand, 2011). A study made an evaluation for both seed and grass production of

the same plant material and found a 3.2-ton ha^{-1} DM yield in the first cut under turf conditions (Hosaininejadmir, Jafari, & Nakhjavan, 2011).

A research team achieved a total DM yield of 7.3- and 6.2-tons ha^{-1} in spaced plant and sward conditions, respectively (Mohajer, Jafari, & Taha, 2011). Mohajer et al. (2013) evaluated 72 sainfoin genotypes in another study and obtained a total annual DM yield of 13.5 tons ha^{-1} (Mohajer, Jafari, Taha, Yaacob, & Saleh, 2013). Another study achieved a yield of 3.96 tons ha^{-1} (DM) in Türkiye (M. Tosun, 1988). DM yield is a complex trait that is highly dependent on yield components and is highly influenced by many genetic and environmental factors. For this reason, evaluating the genotypic potential of sainfoin in different environments is an important step in breeding programs in terms of growing plants with commercially desired characteristics.

The sainfoin can be planted as a monoculture or in a mixture with other perennial grasses or alfalfa. The DM yield of sainfoin varies between 7-15 tons ha^{-1} in Western Canada, depending on the different growing conditions (Goplen et al., 1991). According to a five-year field study sainfoin produced DM yields of more than 6 tons ha^{-1} under rainfed conditions in Alberta, Canada (Bhattarai, 2017).

It was observed that the seed yield of Melrose sainfoin cultivar ranged between 307-1531 kg ha-1 in the studies conducted in six different locations in Western Canada (Bhattarai, 2017; Hanna et al., 1972). While the yield of sainfoin seed per hectare is generally 500-900 kg of clean seeds, 1100 kg/ha yield was obtained in some sainfoin cultivars in Canada (Goplen et al., 1991; Hayot Carbonero et al., 2011). It is reported to produce three to four seed crops per year in southern New Mexico, with sainfoin seed yields ranging from 56 to 1.344 kg/ha, averaging 336 kg/ha. (Lauriault, Marsalis, Mark A., & Contreras-Govea, Francisco E., 2009).

The sainfoin production area in Türkiye in the last five years is an average of 250.5 hectare. While sainfoin seed production was 248 tons in 2020, it was 113 tons in 2021. The amount of sainfoin seed production in the last five years is approximately 170 tons. While the sainfoin seed yield was 550 kg/ha in 2020, it was 400 kg/ha in 2021. The average yield over the last 5 years is 774 kg/ha (TurkStat, 2023). Considering the aboveground fresh and dry biomass

amounts of sainfoin in Türkiye, while the fresh biomass amount was approximately 1.93 million tons in 2020, the fresh biomass amount was approximately 1.55 million tons in 2021. The average amount of above-ground fresh biomass in the last five years is 1.8 million tons (TurkStat, 2023).

2.8. Phenomic studies of sainfoin

Phenomics is an interdisciplinary field that enables the systematic study of plants and plant parts through image taking and analyses (Bilder et al., 2009). Phenomics successfully integrates with many fields such as biology, bioinformatics, computer science, engineering, and statistics which can enable a detailed understanding of complex plant traits at both the physiological and molecular levels (Zhao et al., 2019). Phenomics combines high-performance computing and artificial intelligence technologies and provides comprehensive tools for analyzing various characteristics of plants, including agronomic, morphological, and biochemical traits (Awada, Phillips, & Bodan, 2024; Zhao et al., 2019).

Phenomic technologies offer numerous advantages over traditional methods such as collecting data from many phenotypic measurements in large populations quickly and accurately. Therefore, high-throughput data collection is one of the most significant advantages and also it minimizes human error and enables more accurate measurements using standardized procedures (Pasala & Pandey, 2020). Plant phenomics can also facilitate the measurement of plant development over time and the observation of change. It enables multi-scale analysis, from the plant cell to the entire plant, enabling the study of environmental, genetic, and phenotypic interactions (Zhao et al., 2019). It can improve selection decisions in plant breeding by allowing a relationship between genotypic data and phenotypic data (Lobos et al., 2017). It facilitates the study of environmental plasticity by providing phenotypic measurements under different ecological conditions (Tardieu, Cabrera-Bosquet, Pridmore, & Bennett, 2017). Finally, it can extract many complex and difficult-to-measure morphological traits, such as color and surface, from image data at a high resolution (Perez-Sanz, Navarro, & Egea-Cortines, 2017).

Image-based phenotyping systems are crucial in the field of plant phenomics. Various sensor systems, including RGB cameras, multispectral and hyperspectral cameras, infrared and

thermal cameras, and capture technologies are used for image acquisition. Image processing steps include preprocessing (noise reduction/contrast enhancement), segmentation (separating parts such as plants and seeds from the background), feature extraction (area, environment, etc.), modeling/classification/regression (phenotype prediction using machine learning and deep learning), and verification and quality control (Perez-Sanz et al., 2017).

Despite advances in phenomic tools and their implementation in many crops, phenomics-based studies on sainfoin (*Onobrychis* spp.) have recently begun to emerge and only focused on seed traits (Meyering, Barriball, & Schlautman, 2025).

2.9. Genomic tools to accelerate breeding

As the human population increases, the risk of encountering food shortages also increases. It is inevitable to increase the effort to develop high-yielding lines and varieties against various biotic and abiotic stresses. Although classical breeding techniques can help improve legume product characteristics in terms of quality, nutrition, and production, they do not help to the required extent. Cost-effective and improved molecular breeding techniques help scientists build better genomes of legume crops (Afzal et al., 2020).

Major limitations of sainfoin breeding are the paucity and inadequacy of genomic studies to identify genomic regions associated with important defining traits such as yield, quality, and resistance to biotic and abiotic stress factors. The use of GWAS and genomic selection for population genomic studies in sainfoin may be an important step in the future (Şakiroğlu, 2021). Indirect selection can be performed using traits that are easily identified and associated with the target trait or through genetic markers associated with the trait of interest [Marker-Assisted Selection (MAS)] and accelerate reproduction. The main prerequisite for MAS is the presence of markers that are tightly linked to genes or Quantitative Trait Loci (QTL). Such markers can be used for purposes such as selecting traits that are difficult to measure or dependent on developmental stage, maintaining recessive alleles, pyramiding monogenic traits, and accelerating the reproductive process in general. Regarding polygenic inherited polyploid species and traits, both factors hinder the efficient selection of genotypes based on their phenotypes and therefore the application of MAS would be particularly valuable (B. Barrett, Baird, & Woodfield, 2009). Therefore, detecting potential QTLs in sainfoin would be

particularly promising to improve breeding progress (Kempf, Malisch, Grieder, Widmer, & Kölliker, 2017).

For this reason, it is frequently used in fields such as gene mapping studies, genetic relationship studies, parental determination, detection of intraspecific variations, interspecies hybridization studies, population dynamics, and phylogeography (Hillis & Moritz, 1990; Kimmel et al., 1998; S. Özkan, 2017).

Sequence-related amplified polymorphism (SRAP) has been successfully used in the molecular investigation of allogamy and autogamy of sainfoin (Kempf et al., 2015, 2017). Using next-generation sequencing, RNA transcriptome sequencing has been successfully used to develop extensive sources of simple sequence repeats (SSR) markers that can be applied to assess genetic diversity in sainfoin (Kempf et al., 2017; Kempf, Mora-Ortiz, Smith, Kölliker, & Skøt, 2016; Marina Mora-Ortiz et al., 2016).

In a study that performed transcriptome assembly in sainfoin, the resulting 92K transcripts provide a rich source of SNP and SSR polymorphisms for future use in genetic studies of this crop. The sainfoin has an important trait such as condensed tannin, which annotation of genes involved in the condensed tannin biosynthesis pathway has provided the basis for further studies on the genetic control of this trait. At the stage of nodule initiation and development, 17 genes were identified whose homologues are known to be at the very beginning of nodule initiation and signaling processes (Marina Mora-Ortiz et al., 2016).

A transcriptomic study in sainfoin produced 24.630.711 clean reads from 14 different sainfoin tissues using Illumina double-ended sequencing technology. 77.764 unigene sequences were obtained and 6.752 EST-SSRs were identified using the de novo assembly. A total of 2.469 primer pairs were designed and 200 primer pairs were randomly selected to analyze the polymorphism in five sainfoin wild accessions. At the end of the study, further analysis of 40 sainfoin individuals from five wild populations using 61 EST-SSR loci showed that the number of alleles per locus ranged from 4 to 15, and accordingly the expected heterozygosity ranged from 0.55 to 0.91 (Shen et al., 2019).

In a study of a comprehensive analysis of the phenotype, metabolome, and transcriptome in sainfoin white flower (WF) and amaranth flower (AF), delphinidin, petunidin, and malvidin derivatives were detected as the main anthocyanin compounds in sainfoin AF, while these substances were not detected in sainfoin WF. In flowers, 10,166 DEGs in S1 stage and 15,334 DEGs in S3 stage were detected. KEGG pathway annotations showed that 6396 unigenes disclosed 120 pathways and contained 866 DEGs in the S1 stage, and 6396 unigenes disclosed 131 pathways while containing 1546 DEGs in the S3 stage. The nine DEGs of the "flavonoid biosynthesis" and "phenylpropanoid biosynthesis" pathways involved in flower color formation were identified and confirmed by RT-qPCR analyses. Among these DEGs, 4CL3, FLS, ANS, CHS, DFR, and CHI2 showed downregulated expression, while F3H showed upregulated expression in WF compared to AF. This caused a decrease in anthocyanin synthesis and the formation of WF in sainfoin (Qiao et al., 2021).

The sainfoin genomic studies have lagged other forage crops despite the superior agronomic traits of the sainfoin and its growing cultivation potential worldwide. Since there has been no interest in sainfoin genomic studies until recently, there is insufficient development of genomic resources for this species only (Marina Mora-Ortiz et al., 2016; Şakiroğlu, 2021). However, isozymes, RAPD, AFLP, SSR ISSR, iPBS, and SNP markers have been developed and used in many genomic studies only very recently.

Since it is rational to use model organism at the initial stages of the research, most genetic advances are made in self-pollinating diploid species resulting a vast array of diploid-specific tools and techniques in molecular genetic research (Bourke, van Geest, et al., 2018).

2.10. Population genomics studies of sainfoin

Diversity in plant genetic resources is important for plant breeders to develop new and improved varieties with desirable traits. Classification and identification of plant accessions help to understand genetic diversity (Majidi, Mirlohi, & Amini, 2009). Techniques such as Inter Simple Sequence Repeat (ISSR), Random Amplified Polymorphic DNA (RAPD), Amplified Fragment Length Polymorphism (AFLP), and Single Nucleotide Polymorphism (SNP) are molecular techniques used to understand genetic diversity for plant diversity studies (Bhattarai, 2017). A high within-population genetic variation, ranging from 78-89%, was found in different diversity

studies in sainfoin by using RAPD and ISSR markers. The percentage of polymorphic bands was observed to be in the range of 66-96%. Sainfoin diversity studies were reported to reveal higher levels of within-population genetic variation than among the populations that is a common pattern in outcrossing species (Bhattarai, 2017).

Population genomic studies have not yet been performed to understand the sainfoin's true center of diversity, its domestication pattern or the impact of domestication on the sainfoin genome. Large populations of sainfoin collected from natural growing areas were collected and stored in seed banks. These sainfoin populations provide an invaluable resource for genomic studies of the sainfoin. Targeting many broad-based populations with the presence of genomic tools such as GBS will provide an understanding of the organization of genetic diversity and the relationships between taxa (Şakiroğlu, 2021).

Genome-wide markers can be used to understand the landscape genomics of the sainfoin. Structural genomics studies targeting linkage maps and QTL mapping, or population genomic studies of association mapping and genomic selection in sainfoin are very rare. In one study, a single trait-marker association study in sainfoin was conducted to detect QTL, which controls important traits such as plant height, vigor, and seed yield in a biparental F1 population. Two independent loci associated with these three features were identified in viability in the sainfoin (Kempf, 2016; Şakiroğlu, 2021).

Genomic breeding values are improved with a high-throughput DNA sequencing method known as next-generation sequencing (NGS). NGS and other new technologies enable plant breeders to develop markers for studies such as diversity analyses, genotyping, high-density genetic maps, and population genetic studies, and these can be used together with existing breeding technologies to achieve desired goals (Afzal et al., 2020; Lorenz et al., 2010). Moreover, genomic approaches are necessary to deal with multigenic complex traits and have shown a high impact on the environment. In addition, genomic tools are essential for mapping quantitative trait loci (QTLs) of rare alleles that are not often detected in the gene pool (Afzal et al., 2020; Vaughan, Balázs, & Heslop-Harrison, 2007).

Draft genome sequencing allows for genomic gains and improves crops based on the selection of genes controlling favorable traits that increase yield and nutritional quality. It also provides the necessary information to understand the genome structure and identifies mutagenic changes in the genome resulting from deletions and insertions and uncovers pathways associated with the stress response. The availability of worldwide genetic resources provides plant community researchers the chance to identify unique alleles or combinations of genes required for crop development (Afzal et al., 2020; McCouch et al., 2016).

Various applications of DNA markers are used for breeding purposes. In addition, the use of molecular breeding for cultivar development has become more popular due to its advantages such as precision of the method, time-saving and maximum benefit. Also, simple traits are much more important when associated with complex or multiple genes. Molecular markers are also extremely important for the successful selection of varieties resistant to parasitic weeds and diseases (Afzal et al., 2020; Rietzschel & De Buyzere, 2012). These markers can also be easily used to find linked genes responsible for plant growth habit and quality (ie, vicine and tannin content) important to plants (Afzal et al., 2020; Torres et al., 2010).

MAB can be used in combination with traditional field breeding to accelerate the selection of desired traits (Collard & Mackill, 2008) and to increase selection efficiency (Afzal et al., 2020; Ragot & Lee, 2007). MAB transferred to the genotype of a specific marker system can be used to increase the efficiency and effectiveness of marker selection compared to classical breeding (Afzal et al., 2020; Collard & Mackill, 2008). Breeders can use these genes or alleles as diagnostic tools to identify desired genes (Afzal et al., 2020).

NGS and genotyping are important and cost-effective techniques that facilitate the identification of some important gene constructs. NGS also helps to gain insight into the basic mechanism of gene expression and metabolism. Furthermore, NGS facilitates the characterization of genomic resources and development, evolution and MAB even in the absence of sequence data. NGS also benefits studies to accelerate the development of transcriptome sequencing, de-novo assembly, and sequence-based marker sources for attachment development (Unamba, Nag, & Sharma, 2015). By editing the sequence, information is provided about the genes that control growth, adaptation to the environment, and

development (Afzal et al., 2020). These advanced technologies have the advantages such as being affordable, environmentally friendly, and less time-consuming and these technologies contribute to the development and improvement of pulses products and breeding programs (Afzal et al., 2020).

3. MATERIAL AND METHOD

3.1. Plant Materials/Panel

A comprehensive panel was assembled and analyzed to investigate the genetic diversity within sainfoin (*Onobrychis* spp.) across two distinct locations such as Tekirdağ, Türkiye, and Salina, Kansas, USA. These selected locations are geographically and environmentally distinct environments. The sainfoin diversity panel in Tekirdağ was established in 2022. The diversity panel in Tekirdağ was comprised a total of 47 sainfoin accessions, cultivars, and breeding populations, to represent three different sainfoin species. Especially, this panel included 3 accessions from *O. arenaria*, 19 from *O. transcaucasica*, and 12 from *O. viciifolia*. In addition, two modern Turkish sainfoin cultivars – Özerbey and Lütfübey – along with 13 breeding populations, were incorporated to enhance the genetic diversity of the panel. Detailed information regarding the plant materials used at the Tekirdağ, Türkiye site is presented in Table 3.1.

Table 3.1. List of sainfoin accessions/cultivars/breeding populations in Tekirdağ, Türkiye location

No	Species	USDA Accession No	Country of Origin/Breeding Station	Improvement status
1	*Onobrychis arenaria* subsp. *arenaria*	PI 312965	Former USSR	Accession
2	*Onobrychis arenaria* subsp. *arenaria*	PI 312967	Russian Federation	Accession
3	*Onobrychis arenaria* subsp. *arenaria*	PI 312969	Former USSR	Accession
4	*Onobrychis transcaucasica*	PI 251698	Former USSR	Accession
5	*Onobrychis transcaucasica*	PI 273758	Azerbaijan	Accession
6	*Onobrychis transcaucasica*	PI 273769	Voronezh, Russian Federation	Accession
7	*Onobrychis transcaucasica*	PI 273780	Armenia	Accession
8	*Onobrychis transcaucasica*	PI 273782	Ukraine	Accession
9	*Onobrychis transcaucasica*	PI 273783	Stavropol, Russian Federation	Accession
10	*Onobrychis transcaucasica*	PI 312979	Former USSR	Accession
11	*Onobrychis transcaucasica*	PI 312985	Former USSR	Accession
12	*Onobrychis transcaucasica*	PI 312991	Former USSR	Accession
13	*Onobrychis transcaucasica*	PI 312992	Former USSR	Accession
14	*Onobrychis transcaucasica*	PI 312994	Former USSR	Accession
15	*Onobrychis transcaucasica*	PI 313001	Former USSR	Accession
16	*Onobrychis transcaucasica*	PI 313003	Former USSR	Accession
17	*Onobrychis transcaucasica*	PI 313005	Former USSR	Accession
18	*Onobrychis transcaucasica*	PI 313006	Former USSR	Accession
19	*Onobrychis transcaucasica*	PI 313008	Former USSR	Accession
20	*Onobrychis transcaucasica*	PI 313023	Former USSR	Accession
21	*Onobrychis transcaucasica*	PI 313041	Former USSR	Accession
22	*Onobrychis transcaucasica*	PI 372816	Checkia	Accession
23	*Onobrychis viciifolia*	PI 206458	Türkiye	Accession
24	*Onobrychis viciifolia*	PI 223389	Iran	Accession
25	*Onobrychis viciifolia*	PI 251697	Former USSR	Accession
26	*Onobrychis viciifolia*	PI 273786	Stavropol, Russian Federation	Accession

No	Species	USDA Accession No	Country of Origin/Breeding Station	Improvement status
27	*Onobrychis viciifolia*	PI 273790	Krasnodar, Russian Federation	Accession
28	*Onobrychis viciifolia*	PI 302936	Spain	Accession
29	*Onobrychis viciifolia*	PI 311468	Spain	Accession
30	*Onobrychis viciifolia*	PI 313047	Germany	Accession
31	*Onobrychis viciifolia*	PI 313053	Former USSR	Accession
32	*Onobrychis viciifolia*	PI 313063	Former USSR	Accession
33	*Onobrychis viciifolia*	PI 319713	Romania	Accession
34	*Onobrychis viciifolia*	PI 494668	Romania	Accession
35	*Onobrychis viciifolia*	Özerbey	Turkiye	Cultivar
36	*Onobrychis viciifolia*	Lütfübey	Turkiye	Cultivar
37	*Onobrychis viciifolia*	NKU	Turkiye	Breeding Population
38	*Onobrychis viciifolia*	Kayı Elektrik Direği	Turkiye	Breeding Population
39	*Onobrychis viciifolia*	Kayı	Turkiye	Breeding Population
40	*Onobrychis viciifolia*	İsviçre	Turkiye	Breeding Population
41	*Onobrychis viciifolia*	AlfaK01	Turkiye	Breeding Population
42	*Onobrychis viciifolia*	AlfaK02	Turkiye	Breeding Population
43	*Onobrychis viciifolia*	AlfaK04	Turkiye	Breeding Population
44	*Onobrychis viciifolia*	AlfaK05	Turkiye	Breeding Population
45	*Onobrychis viciifolia*	Camialtı	Turkiye	Breeding Population
46	*Onobrychis viciifolia*	Vakıflar	Turkiye	Breeding Population
47	*Onobrychis viciifolia*	92 breeding line	Turkiye	Breeding Population
48	*Onobrychis viciifolia*	24 breeding line	Turkiye	Breeding Population
49	*Onobrychis viciifolia*	Klavuzlu	Turkiye	Breeding Population

A second sainfoin diversity panel was established in Salina, Kansas, USA in 2023, and the second panel consists of a total of 80 accessions and cultivars. This panel was represented a broad genetic base which included 21 accessions from *Onobrychis arenaria*, 33 from *O. transcaucasica*, and 26 from *O. viciifolia*. Additionally, six modern sainfoin cultivars- 'Shoshone', 'Delaney', 'Renumex', 'Eski', 'Rocky Mountain Remont', and 'AAC Mountainview' – were incorporated to strengthen the panel's agronomic relevance and genetic breadth. The complete list of plant materials used in Salina location is presented in Table 3.2.

Table 3.2. List of sainfoin accessions/cultivars in Salina, Kansas, USA

No	Species	USDA Accession No	Country of Origin/Breeding Station	Improvement status
1	*Onobrychis viciifolia*	PI 302936	Spain	Accession
2	*Onobrychis viciifolia*	PI 319713	Romania	Accession
3	*Onobrychis viciifolia*	PI 338651	Morocco	Accession
4	*Onobrychis viciifolia*	PI 273790	Russian Federation, Krasnoyarsk	Accession
5	*Onobrychis viciifolia*	PI 440576	Kazakhstan	Accession
6	*Onobrychis viciifolia*	PI 401468	Romania	Accession
7	*Onobrychis viciifolia*	PI 311468	Spain	Accession
8	*Onobrychis viciifolia*	PI 368036	Türkiye, Ankara	Accession
9	*Onobrychis viciifolia*	PI 313048	Soviet Union, Former	Accession
10	*Onobrychis viciifolia*	PI 494667	Romania	Accession
11	*Onobrychis viciifolia*	PI 600767	United States, New Mexico	Accession
12	*Onobrychis viciifolia*	PI 568209	Türkiye, Tokat	Accession
13	*Onobrychis viciifolia*	PI 313047	Germany	Accession
14	*Onobrychis viciifolia*	PI 313053	Soviet Union, Former	Accession
15	*Onobrychis viciifolia*	PI 302939	Spain	Accession
16	*Onobrychis viciifolia*	PI 494668	Romania	Accession
17	*Onobrychis viciifolia*	PI 223389	Iran	Accession
18	*Onobrychis viciifolia*	PI 273786	Russian Federation, Stavropol	Accession
19	*Onobrychis viciifolia*	PI 313063	Soviet Union, Former	Accession
20	*Onobrychis viciifolia*	PI 313065	Soviet Union, Former	Accession
21	*Onobrychis viciifolia*	PI 314099	Soviet Union, Former	Accession
22	*Onobrychis arenaria* subsp	PI 312949	Soviet Union, Former	Accession
23	*Onobrychis arenaria*	PI 312930	Soviet Union, Former	Accession
24	*Onobrychis arenaria*	PI 372798	Czech Republic, Central Bohemia	Accession
25	*Onobrychis arenaria*	PI 372799	Czech Republic, Central Bohemia	Accession

No	Species	USDA Accession No	Country of Origin/Breeding Station	Improvement status
26	*Onobrychis arenaria*	PI 273744	Kazakhstan	Accession
27	*Onobrychis arenaria*	PI 273748	Ukraine, Kiev	Accession
28	*Onobrychis arenaria*	PI 372803	Czech Republic, Central Bohemia	Accession
29	*Onobrychis arenaria*	PI 312910	Soviet Union, Former	Accession
30	*Onobrychis arenaria*	PI 312914	Soviet Union, Former	Accession
31	*Onobrychis arenaria*	PI 312916	Soviet Union, Former	Accession
32	*Onobrychis arenaria*	PI 273749	Ukraine	Accession
33	*Onobrychis arenaria* subsp.	PI 312965	Soviet Union, Former	Accession
34	*Onobrychis arenaria*	PI 273746	Russian Federation	Accession
35	*Onobrychis arenaria*	PI 273752	Russian Federation, Voror	Accession
36	*Onobrychis arenaria* subsp.	PI 312967	Soviet Union, Former	Accession
37	*Onobrychis arenaria*	PI 312919	Soviet Union, Former	Accession
38	*Onobrychis arenaria*	PI 312921	Soviet Union, Former	Accession
39	*Onobrychis arenaria*	PI 312922	Soviet Union, Former	Accession
40	*Onobrychis arenaria*	PI 312924	Soviet Union, Former	Accession
41	*Onobrychis arenaria*	PI 312925	Soviet Union, Former	Accession
42	*Onobrychis arenaria*	PI 273743	Ukraine	Accession
43	*Onobrychis arenaria*	PI 372800	Czech Republic, Central Bohemia	Accession
44	*Onobrychis arenaria*	PI 312926	Soviet Union, Former	Accession
45	*Onobrychis arenaria*	PI 505891	Soviet Union, Former	Accession
46	*Onobrychis arenaria*	PI 312920	Soviet Union, Former	Accession
47	*Onobrychis arenaria*	PI 372801	Czech Republic, Central Bohemia	Accession
48	*Onobrychis arenaria*	PI 372802	Czech Republic, Central Bohemia	Accession
49	*Onobrychis arenaria*	PI 312969	Soviet Union, Former	Accession
50	*Onobrychis arenaria*	PI 312927	Soviet Union, Former	Accession
51	*Onobrychis arenaria*	PI 372804	Czech Republic, Central Bohemia	Accession

No	Species	USDA Accession No	Country of Origin/Breeding Station	Improvement status
52	*Onobrychis arenaria* subsp.	PI 577666	Russian Federation	Accession
53	*Onobrychis arenaria*	PI 312928	Soviet Union, Former	Accession
54	*Onobrychis arenaria*	PI 312929	Soviet Union, Former	Accession
55	*Onobrychis transcaucasica*	PI 312991	Soviet Union, Former	Accession
56	*Onobrychis transcaucasica*	PI 312994	Soviet Union, Former	Accession
57	*Onobrychis transcaucasica*	PI 251698	Soviet Union, Former	Accession
58	*Onobrychis transcaucasica*	PI 372816	Czech Republic, Central Bohemia	Accession
59	*Onobrychis transcaucasica*	PI 273780	Armenia	Accession
60	*Onobrychis transcaucasica*	PI 313006	Soviet Union, Former	Accession
61	*Onobrychis transcaucasica*	PI 251697	Soviet Union, Former	Accession
62	*Onobrychis transcaucasica*	PI 312979	Soviet Union, Former	Accession
63	*Onobrychis transcaucasica*	PI 273769	Russian Federation, Voror	Accession
64	*Onobrychis transcaucasica*	PI 313023	Soviet Union, Former	Accession
65	*Onobrychis transcaucasica*	PI 312985	Soviet Union, Former	Accession
66	*Onobrychis transcaucasica*	PI 313005	Soviet Union, Former	Accession
67	*Onobrychis transcaucasica*	PI 273758	Azerbaijan	Accession
68	*Onobrychis transcaucasica*	PI 313001	Soviet Union, Former	Accession
69	*Onobrychis transcaucasica*	PI 273782	Ukraine	Accession
70	*Onobrychis transcaucasica*	PI 313003	Soviet Union, Former	Accession
71	*Onobrychis transcaucasica*	PI 273783	Russian Federation, Stavropol	Accession
72	*Onobrychis transcaucasica*	PI 313008	Soviet Union, Former	Accession
73	*Onobrychis transcaucasica*	PI 312992	Soviet Union, Former	Accession
74	*Onobrychis transcaucasica*	PI 313041	Soviet Union, Former	Accession
75	*Onobrychis viciifolia*	Shoshone	USA, Wyoming	Cultivar
76	*Onobrychis viciifolia*	Delaney	USA, Wyoming	Cultivar
77	*Onobrychis viciifolia*	Renumex	USA, New Mexico	Cultivar

No	Species	USDA Accession No	Country of Origin/Breeding Station	Improvement status
78	*Onobrychis viciifolia*	Eski	Türkiye, Eskişehir	Cultivar
79	*Onobrychis viciifolia*	Rocky Mountain Remont	USA, Montana	Cultivar
80	*Onobrychis viciifolia*	AAC Mountainview	Canada	Cultivar

3.2. Experimental Design of Field Trials

Tekirdağ-Türkiye location: A total of 47 sainfoin accessions/cultivars/breeding populations were sown to pots on November 7, 2021, and germinated. The plants were subsequently transplanted to the field using a randomized complete block design with three replications to the experimental field of Namık Kemal University in Tekirdağ/Türkiye (40.990012, 27.581711) on April 5, 2022. Entries consisted of three replications of 10 randomly selected plants (30 plants total). The field layout in Tekirdağ featured end-to-end and side-by-side row spacing of 75 cm with an intra-row plant spacing of 30 cm. To provide uniformity, all the plots were moved on August 25, 2022. Phenotypic measurements in the sainfoin experimental field at the Tekirdağ location were conducted on May 5-7, 2023, and again on 11 July 2023.

Salina-KS-USA location: A total of 2200 individual plants from 80 sainfoin accessions/cultivars were sown to pots in August 2023 and germinated. The seedlings were transplanted to The Land Institute Research Farm (38.770421, -97.591532) in a randomized complete design with three replications of 60 rows and 40 columns on September 19, 2023. Plots were spaced end-to-end and side-by-side row spacing of 75 cm with an intra-row plant spacing of 75 cm. After transplanting, irrigation was done twice, and fertilization wasn't applied. Phenotypic measurements in the sainfoin experimental field at the Salina location were conducted on different dates throughout the growing season. Canopy height and plant stem erectness were measured on May 14-15, 2024. Stem thickness, stem length, and raceme number per stem were assessed from May 28, 2024, to June 7, 2024. Stem number per plant and forage dry mass measurements began on June 20, 2024, and were completed on July 8, 2024. Plant regrowth rating was recorded for all plants on July 11, 2024. Legume fruit mass per plant was measured from July 30, 2024, to September 5, 2024. Legume Fruit 1000K weight was calculated on December 16, 2024. For legume fruit trait analysis, a 2-gram seed sample was taken from each seed lot, and high-resolution photographs were captured using a professional camera. This process started on July 23, 2024, and was completed on September 5, 2024. The seeds from each 2-gram sample were then dehulled using a threshing machine, and additional photographs were taken of the dehulled seeds. This dehulling process began on September 23, 2024, and was finalized on October 8, 2024. The collected images of both hulled and dehulled seeds were analyzed using image processing techniques to extract seed trait data.

3.3. Phenotyping Data Collection

The selected panel phenotyped for forage and seed yield and related traits in two locations as Tekirdağ, Türkiye and Salina, Kansas, USA. A total of 49 sainfoin accessions/cultivars/ breeding populations were arranged with three replications in Tekirdağ, Türkiye. A total of 80 sainfoin accessions and cultivars were arranged with three replications in Salina, Kansas. The trial in Tekirdağ was established in April 2022 whereas the one in Kansas, USA was established in August 2023.

A total of 440 individual plants from 49 sainfoin accessions/cultivars/populations were sown to pots on November 7, 2021, and germinated in Tekirdağ/Türkiye. The plants were subsequently transplanted to the field using a randomized complete block design with three replications on April 5, 2022, to the experimental field of Namık Kemal University in Tekirdağ (40.990012, 27.581711). Each accession consisted of three replications of 10 randomly selected plants (30 plants total). To provide uniformity of the data all the plots were mowed on August 25, 2022. Morphological measurements were not taken at the plots as it was the establishment year.

In 2023, measurements were taken at the Tekirdağ location in Türkiye for 48 different accessions of plants. The measurements were conducted in May and July of 2023, focusing on agronomic traits. These important traits measured in the Tekirdağ-23 environment are listed below, along with their abbreviations and trait numbers in the Sainfoin Crop Ontology: Canopy Height (CanHt) (CO_369:0001001), Stem Thickness (StTh) (CO_369:0001101), First Cutting Stem Number (StN1) (CO_369:0001082), Stem Length (StL) (CO_369:0001081), Forage Subsample Dry Mass (FoSDM) (CO_369:0001015), Second Cutting Stem Number (StN2) (CO_369:0001082), Raceme Number (RaN) (CO_369:0001055), Raceme Length (RaL) (CO_369:0001054), Inflorescence Length (InfL) (CO_369:0001018), and Seed Number per Raceme (SdN) (CO_369:0001070) (Table 3.3). Phenotypic measurements in the field in Tekirdağ were taken in May and July 2023 and the measurements were completed (Figure 3.1).

Figure 3.1. Taking measurement of phenotypic traits in the Tekirdağ, Türkiye

List with details of the traits whose measurements were taken, CO trait numbers, trait description, method number, method name, method description and the date the measurements were recorded (Table 3.3). Data was uploaded to the germinate instance of the sainfoinbase database (https://sainfoinconsortium.org/#/data/data-resources).

Table 3.3. Crop Ontology List of traits associated with forage and seed yield measured in Tekirdağ-23 environment

Trait name	Unit	Trait ID	Trait description	Method ID	Method name	Method description	Recording date
Canopy height	m	CO_369:0001001	The average canopy height of a plant, plot, or group of plants when measured from the soil surface to the growing apex of the plant(s), expressed in cm.	CO_369:0002050	Plant natural height measurement	Measure the natural height of plants	5.05.2023
Stem number	n	CO_369:0001082	The number of stems on a per plant basis.	CO_369:0002081	Stem number per plant computation	Record the total number of stems per plant	5.05.2023
Stem length	cm	CO_369:0001081	The length of the plant stem measured in cm when stretched out from end to end.	CO_369:0002079	Stem length measurement	Measure the length of one representative stem by stretching it out end to end.	5.05.2023/ 7.10.2023
Forage fresh mass	g	CO_369:0001012	The total fresh mass of harvested forage material expressed in kg.	CO_369:0002014	Forage mass measurement	Record the mass of forage material	5.05.2023
Forage subsample dry mass	g	CO_369:0001015	The total dry biomass of a subsample selected out of a larger portion of harvested biomass expressed in kg.	CO_369:0002015	Forage subsample mass measurement	Record the mass of a forage subsample	5.05.2023
Raceme number	n	CO_369:0001055	The number of racemes divided by the total number of stems sampled.	CO_369:0002056	Raceme number per stem computation	Compute the number of racemes per stem	7.10.2023

Trait name	Unit	Trait ID	Trait description	Method ID	Method name	Method description	Recording date
Raceme length	cm	CO_369:0001054	The length of the raceme measured in cm from end to end when laid out.	CO_369:0002053	Raceme length measurement	Measure the length of a raceme	7.10.2023
Inflorescence length	cm	CO_369:0001018	The overall length of the inflorescence when laid end to end expressed in mm.	CO_369:0002017	Inflorescence length measurement	Record the average inflorescence length	7.10.2023
Seed number	n	CO_369:0001070	The total number of seeds divided by the total number of racemes that were sampled.	CO_369:0002066	Seed number per raceme computation	Record the total number of seeds per raceme	7.10.2023
Stem thickness	mm	CO_369:0001101	The thickness of a stem as measured by digital calipers between the second and third nodes on a mature stem, expressed in mm.	CO_369:0002093	Stem thickness measurement	Measurement of stem thickness using digital calipers	5.05.2023
Plant winter survival	n	CO_369:0001052	The number of plants that survived the winter and regrow in the spring	CO_369:0002083	Winter survival estimation	Estimate the winter survival of a plot of plants	5.05.2023

Measurements of the similar important traits and more traits in Salina, KS were ended successfully in October 2024. These important traits measured in the Salina-24 environment are listed below, along with their abbreviations and trait numbers in the Sainfoin Crop Ontology: Canopy Height (CanHt) (CO_369:0001001), Stem Thickness per Plant (StTh) (CO_369:0001101), Stem Number per Plant (StN) (CO_369:0001082), Stem Length (StL) (CO_369:0001081), Forage Dry Mass (FoDM) (CO_369:0001010), Raceme Number (RaN) (CO_369:0001055), Raceme Length (RaL) (CO_369:0001054), Plant Stem Erectness (PlStEr) (CO_369:0003029), Plant Regrowth Rating (PlRGR), Legume Fruit Mass (FrM) (CO_369:0001032), Legume Fruit 1000K Weight (Fr1KK) (CO_369:0000135), (https://cropontology.org/ontology/CO_369) (Table 3.4).

Table 3.4. Crop Ontology list of traits associated with forage yield measured in Salina-24 environment

Trait name	Unit	Trait ID	Trait description	Method ID	Method name	Method description	Recording date
Canopy height	cm	CO_369:0001001	The average canopy height of a plant, plot, or group of plants when measured from the soil surface to the growing apex of the plant(s), expressed in cm.	CO_369:0002050	Plant natural height measurement	Measure the natural height of plants	May 14-15, 2024
Stem thickness	mm	CO_369:0001101	The thickness of a stem as measured by digital calipers between the second and third nodes on a mature stem, expressed in mm.	CO_369:0002093	Stem thickness measurement	Measurement of stem thickness using digital calipers	May 28, 2024
Stem length	cm	CO_369:0001081	The length of the plant stem measured in cm when stretched out from end to end.	CO_369:0002079	Stem length measurement	Measure the length of one representative stem by stretching it out end to end.	May 28, 2024
Forage dry mass	g	CO_369:0001015	The total dry biomass of a subsample selected out of a larger portion of harvested biomass expressed in kg.	CO_369:0002015	Forage subsample mass measurement	Record the mass of a forage subsample	June 20 - July 8, 2024
Stem number	n	CO_369:0001082	The number of stems on a per plant basis.	CO_369:0002081	Stem number per plant computation	Record the total number of stems per plant	June 20 - July 8, 2024
Raceme number	n	CO_369:0001055	The number of racemes divided by the total number of stems sampled.	CO_369:0002056	Raceme number per stem computation	Compute the number of racemes per stem	May 28, 2024

Trait name	Unit	Trait ID	Trait description	Method ID	Method name	Method description	Recording date
Plant stem erectness	0-9	CO_369:0001050	The proportion of plants within a sampled plot/sward that are resistant to stem lodging, i.e. have stem angles > 45 degrees.	CO_369:0002078	Stem erectness estimation	Estimate the erectness of stems in a plant or plot of plants	May 14-15, 2024
Plant regrowth rating	0-5	It didn’t have been added to Sainfoin Crop Ontology List yet.					July 11, 2024
Legume fruit mass	g	CO_369:0001116	The 'average' mass of a single legume fruit calculated from the total fruit mass or subsample mass at 15% moisture divided by the total number of fruits.	CO_369:0002108	Legume fruit mass ratio computation	Compute the legume fruit mass ratio for a sample of fruit	July 30 - Sep 5, 2024
Legume fruit 1000K weight	g	CO_369:0001117	The total mass of 1000 legume fruit at 15% moisture content.	CO_369:0002109	Legume fruit 1KK mass computation	Compute the 1KK legume fruit mass from a sample legume fruit mass ratio	Dec 16, 2024

Measurements of the accessions in Tekirdağ have been completed and measurements regarding forage and seed yield were taken in 2023. The measured traits are Canopy Height (CanHt), Stem Thickness (StTh), 1st Cutting Stem Number (StN1), Stem Length (StL), Forage Subsample Dry Mass (FoSDM), 2nd Cutting Stem Number (StN2), Raceme Number (RaN), Raceme Length (RaL), Inflorescence Length (InfL), and Seed Number Per Raceme (SdN), respectively. Plant canopy height (CanHt) is the height of the plant's natural stance without any support. Stem length (StL) is the length of the highest point of the plant, where all the stems are gathered. Stem number (SdN) is the counting each stem of the plant one by one. Stem thickness (StTh) is the measurement of the width of five randomly selected stems of the plant with the help of a caliper. During the measurement, the measurements were taken between the 2nd and 3rd nodes of the stem. For plant yield, the plants in each accession were harvested separately and their measurements were taken. A portion of plant samples were collected and separated to calculate forage subsample dry mass. For winter survival rate, since each accession is in its 2nd year, winter survival rates have been taken. The number of live plants in each accession was counted. The winter survival rate with its details was added to CO list (Table 3.3) but it wasn't included to any analyses (Karabulut et al., 2023).

The second phenotypic measurements in Tekirdağ were taken on July 10, 2023. For stem number (StN2), the lengths of three randomly selected stems were taken. The characters were measured as StN1 and StN2, that is, 1st cutting stem number and 2nd cutting stem number. For raceme number (RaN), the raceme number was taken from randomly selected stems. For inflorescence length (InfL), the part connecting to the stem from the top of the raceme was measured from 6 randomly selected racemes for each plant. For raceme length (RaL), the part between the top and bottom of the raceme was measured from the previous six randomly selected racemes for each plant. For seed number per raceme (SdN), the number of seeds available in the six selected racemes was calculated. The length of racemes whose flowering was sufficiently completed was measured. Some seed beds were also calculated. The fact that the seeds had fallen was ignored. According to the sainfoin crop ontology, all observable seed beds were calculated as the number of seeds in that raceme (Karabulut et al., 2023).

In the Salina-24 environment, sainfoin seeds were sown using IHORT Prefilled T50 Excel Propagation Trays in the growth chamber in August 2023 in Salina/Kansas. 76 accessions were

planted with 50 seeds in each tray. During planting, a pinch of superior legume inoculant-peat for sainfoin was added to the seeds of each accession. After their transplantation was completed, they were placed in the LED room. The LED room settings are 18 hours day/6 hours night, day heating target 70 F (21.11 °C), cooling target 76 F (24.44 °C), nighttime heating target 64 F (17.77 °C), cooling target 70 F (21.11 °C), the humidity was set to 60%. Then they were irrigated. It was watered regularly every day for the first week. In the following days, watering was done every other day. Since the seeds were disinfected during planting, gloves and masks were used. Seed planting was completed in four days. Germination of sainfoin seeds started on the 3rd day of sowing, and germination numbers were taken twice a week for the first month. Two of the 76 accessions were eliminated because they did not have sufficient germination. The remaining 74 accessions were allocated to the trays in one, two, or three replications according to the number of existing seedlings, according to a randomized complete block design. 60 sainfoin seedlings, each replication consisting of 20 trays, were prepared to be transferred to the field. For phenotypic comparison, seeds of six sainfoin cultivars were sown and included in the field design. These six sainfoin varieties were Renumex, Delaney, Eski, Shoshone, AAC Mountainview, and Rocky Mountain Remont in Salina/Kansas.

Currently, the transfer to the farmland of The Land Institute in the form of a randomized complete design of 60 rows and 40 columns was completed on September 19, 2023, in Salina/Kansas/USA. During transfer to the land, holes were opened on pre-drawn lines, plants were placed in order and planting was completed. There is a gap in 20 blocks because there are one or two replications of some accessions (Fig 3.2).

Figure 3.2. Transplantation of sainfoin seedlings to the field in the Salina, Kansas

There are five traits measured in common across both environments. In addition to the traits measured in the Tekirdağ-23 environment, a few more forage-related and seed-related traits were measured in the Salina-24 environment. Forage Dry Mass (FoDM) indicates to the total weight of the entire plant after drying following harvest. Plant stem erectness (PlStEr), scored on a scale from 0 to 9 based on the estimated percentage of stems with an average inclination angle higher than 45° relative to the ground. Plant Regrowth Rating (PlRGR) is a visual ranking from 0 to 5 that reflects the regrowth performance of each plant. This trait was evaluated 15-20 days after cutting, in comparison to other plants. Legume Fruit Mass (FrM) is the total seed weight obtained from each plant after threshing. Legume Fruit 1000 Kernel weight (Fr1KK) was calculated by dividing two grams of seeds per plant, photographing the sample, counting the number of seeds in two grams, and estimating the weight of 1000 seeds accordingly (https://cropontology.org/ontology/CO_369).

Figure 3.3. Taking measurement of phenotypic traits in Salina, Kansas, USA

In Tekirdağ-23 environment, agronomic traits were measured manually using tools such as calipers and measuring tapes (Figure 3.1). In the Salina-24 environment, forage- and seed-related traits were measured using similar manual methods (Figure 3.3).

3.4. Phenomic Data Collection

After completing the dry matter weight measurements for each plant, each plant was threshed in a threshing machine. The seeds collected in the machine were carefully packaged individually. The seeds were then stored in a cold storage facility for a few days before being weighed. The traits were measured as hulled and were named as legume fruit. In the hulled state, 2 g of the seed was taken from each packet, and photographs were taken. Then, using a dehulling machine, the seeds in each 2-g packet were separated from the hulls. The dehulled seeds were named as seeds, and photographs were taken. The photographs obtained in both states were used to obtain data on hulled and dehulled seeds using the seed imaging technique in Python (Meyering et al., 2025).

Seed-related traits were analyzed using seed imaging techniques in the Salina-24 environment. Thus, since the seed data obtained by the seed imaging method were not real measurements, they were referred to as estimated seed data, and the analyses performed in the previous chapter were similarly performed. Sample photographs were taken during the threshing process of sainfoin seeds of *Onobrychis viciifolia*, *Onobrychis transcaucasica*, and *Onobrychis arenaria* species, and one of the cultivars, are shown in Figure 3.4.

Onobrychis viciifolia

Onobrychis transcaucasica

Onobrychis arenaria

Renumex

Figure 3.4. Legume fruit samples of three economically important sainfoin species and one of the cultivars during the threshing process in Salina, KS

Important estimated seed-related traits measured in the Salina-24 environment are listed along with their abbreviations and trait numbers in the Sainfoin Crop Ontology: Legume Fruit Projected Area (FrPA) (CO_369:0001035), Legume Fruit Width (FrW) (CO_369:0001037), Legume Fruit Length (FrL) (CO_369:0001031), Legume Fruit Aspect Ratio (FrAR) (CO_369:0001026), Legume Fruit Extent (FrEx) (CO_369:0001030), Seed Projected Area (SdPA) (CO_369:0001073), Seed Width (SdW) (CO_369:0001058), and Seed Extent (SdEx) (CO_369:0001065) (https://cropontology.org/ontology/CO_369) (Table 3.5).

Table 3.5. Crop Ontology list of traits associated with seed yield measured in Salina-24 environment

Trait name	Unit	Trait ID	Trait description	Method ID	Method name	Method description	Recording date
Legume Fruit Projected Area	mm^2	CO_369:0001035	The area of a legume fruit as projected onto the 2-D surface of the image plane.	CO_369:0002042	Object projected area measurement	Calculate the projected area of an object (seed image, python-opencv)	Oct - Dec, 2024
Legume Fruit Width	mm	CO_369:0001037	The width of a legume fruit along its minor axis orthogonal to the main axis along the image plane.	CO_369:0002044	Object width measurement	Calculate the width of an object (seed image, python-opencv)	Oct - Dec, 2024
Legume Fruit Length	mm	CO_369:0001031	The length of a legume fruit along its main axis.	CO_369:0002040	Object length measurement	Calculate an objects length (seed image, python-opencv)	Oct - Dec, 2024
Legume Fruit Aspect Ratio	ratio	CO_369:0001026	The ratio of the legume fruit bounding box width to its height.	CO_369:0002035	Object aspect ratio measurement	Calculate the ratio of the object's bounding box width to its height (seed image, python-opencv)	Oct -Dec, 2024
Legume Fruit Extent	ratio	CO_369:0001030	The ratio of the legume fruit contour area to the bounding box area.	CO_369:0002039	Object extent measurement	Calculate the object extent (seed image, python-opencv)	Oct - Dec, 2024
Seed Projected Area	mm^2	CO_369:0001073	The area of a seed as projected onto the 2-D surface of the image plane.	CO_369:0002042	Object projected area measurement	Calculate the projected area of an object (seed image, python-opencv)	Oct -Dec, 2024
Seed Width	mm	CO_369:0001077	The width of a seed along its minor axis orthogonal to the main axis along the image plane.	CO_369:0002044	Object width measurement	Calculate the width of an object (seed image, python-opencv)	Oct -Dec, 2024
Seed Length	mm	CO_369:0001067	The length of a seed along its main axis.	CO_369:0002040	Object length measurement	Calculate an objects length (seed image, python-opencv)	Oct -Dec, 2024

Trait name	Unit	Trait ID	Trait description	Method ID	Method name	Method description	Recording date
Seed Aspect Ratio	ratio	CO_369:0001058	The ratio of the seed bounding box width to its height.	CO_369:0002035	Object aspect ratio measurement	Calculate the ratio of the object's bounding box width to its height (seed image, python-opencv)	Oct - Dec, 2024
Seed Extent	ratio	CO_369:0001065	The ratio of the seed contour area to the bounding box area.	CO_369:0002039	Object extent measurement	Calculate the object extent (seed image, python-opencv)	Oct - Dec, 2024

The legume fruit and seed traits were quantified using a seed imaging technique. Trait definitions and measurement protocols were applied according to the Sainfoin Crop Ontology (Table 3.5). Specifically, the evaluated fruit traits included: Legume Fruit Projected Area (FrPA) – the two-dimensional projected area of the fruit on the image plane; Fruit Width (FrW) – the width measured along the minor axis perpendicular to the main axis; Fruit Length (FrL) – the length along the main axis; Fruit Aspect Ratio (FrAR) – the ratio of fruit width to its height based on the bounding box; and Fruit Extent (FrEx) – the ratio of the contour area to the bounding box area. Corresponding seed traits included: Seed Projected Area (SdPA) – the two-dimensional projected area of the seed; Seed Width (SdW) and Seed Length (SdL) – the width and length measured along the minor and major axes, respectively; Seed Aspect Ratio (SdAR) – the ratio of bounding box width to height; and Seed Extent (SdEx) – the ratio of seed contour area to bounding box area.

3.5. Genotyping Data Collection

DNA isolation was carried out with samples taken from sainfoin accessions in Tekirdağ, Türkiye. The weights of the samples taken varied between 100-150 g. Plant samples were added to a 2 ml microcentrifuge tube containing 5 ceramic beads and 450 µl Plant DNA Isolation kit Lysis Buffer. 2 ml microcentrifuge tube was homogenized in Bead Ruptor 12 at 6 m/s speed for 2 minutes at room temperature. Then it was mini-spinned for 5 seconds in the centrifuge. The homogenate was incubated at 60 °C for 30 minutes and mixed by hand every 2-3 minutes.

After incubation, the homogenate was centrifuged at 14.000×g for 5 minutes at room temperature. After centrifugation, 350 µl of supernatant was taken and a sterile 1.5 ml microcentrifuge tube was added. Then, 5 µl of RNase was added to each tube and incubated at 37 °C for 20 minutes. Phenol:chloroform:isoamyl alcohol (25:24:1) was added to the tube as much as the supernatant (350µl) (1:1 ratio) after incubation. Then it was centrifuged for one minute at 14.000×g after 5 seconds of short vortex in Thermomac Vortex Mixer. In the phase separation that occurred after centrifugation, 200 µl of the upper phase was taken and added to the Spin Column & Collection Tube (DNA) and centrifuged at 10.000×g for one minute. After the liquid passing through the column was discarded, 500 µl Plant DNA Isolation kit Wash Buffer was added to the silica column and centrifuged at 10.000×g for one minute.

After removing the liquid passing through the column, 500 µl Plant DNA Isolation kit Wash Buffer was added to the silica column and centrifuged at 10.000×g for one minute. After the liquid passing through the column was discarded, the silica column was centrifuged empty at 15.000×g for one minute. After centrifugation, the silica column was transferred to a clean 1.5 ml microcentrifuge tube, then 50 µl of Plant DNA Isolation kit DNA Elution Buffer was added to the column matrix and centrifuged at 15.000×g for one minute. DNA samples obtained after centrifugation were stored at -20 °C. After DNA isolation, measurements were made with Qubit 3.0 Fluorometer as High sensitive or Broad range.

The genotyping diversity panel was genotyped using the Khufu Skim sequencing method. Equal concentrations of each accession will be collected in the DNA pool. Pools will be sequenced by Novogene using Illumina HiSeq 2 x 150 bp sequencing. After filtering, the total reads will be mapped. Analysis of sequenced bulks will be used by adapting the Khufu pipeline (https://www.hudsonalpha.org/khufudata). An "interval" filter will be applied, filtering out closely clustered (<100 bp) potential SNPs. This will show repetitive content and read from many overlapping places. The shiny app takes as input the R objects created by Khufu (Korani et al., 2021).

3.6. Data Analyses

The data analyses in this thesis are described in three main sections. The first chapter, phenotyping data analyses, describes the analyses of the phenotypic measurements and calculations. The second chapter is phenomic data analyses that describe the analyses of sainfoin seed data in the Salina-24 environment obtained using seed imaging. The third chapter is genomic data analyses that describe the analyses of sequence data obtained from sainfoin accessions collected from the Tekirdağ-23 environment.

3.6.1. Phenotyping Data Analyses

Phenotypic evaluations were conducted on sainfoin genotypes grown in two distinct locations -Tekirdağ, Türkiye, and Salina, Kansas, USA - across different years. Where the combination of year location could be considered as two environments. Initially, the accessions common in two environments were merged and analyzed to understand the effect of environment. The analyses of the same trait measured in both environments included canopy height, stem length,

stem thickness, raceme number, and stem number from 30 common accessions. Firstly, the Shapiro-Wilk normality test was conducted for both environments. The *dplyr, purrr, lme4, nortest, stringr,* and *tibble* packages were used to perform the normality test. Based on the results, log transformation was performed for CanHt (Salina-24 environment) and square root transformation (sqrt (x+0.5)) for RaN (Tekirdağ-23 environment). After the transformation, the *lme4* and *nortest* packages were used to test whether the residual error terms, after removing the effects of block and replication for each trait, conformed to a normal distribution. Then, F-values and statistical significance levels were obtained from the variance analysis of the five common traits using the *lmerTest* package.

To assess the normality of phenotypic data from both locations, the *nortest* package in R was used to perform normality test via the Anderson-Darling test (*ad.test*) and the Shapiro-Wilk test (*shapiro.test*). Traits that did not confirm to a normal distribution were transformed using Z-transformation with the *dplyr* package in R. Analysis of variance (ANOVA) were done by obtaining F-values and significance levels for each trait, based on species and accession in both locations, were analyzed using the *lmerTest* package in R.

To obtain the adjusted means, they were obtained to remove replication and block effects from the data. For this purpose, the raw dataset was analyzed using the *Matrix* and *lme4* packages in R to estimate the means for accessions and species. Furthermore, the *emmeans, pbkrtest*, and *lmerTest* packages in the R language program were used to compute the adjusted means for phenotypic traits based on accessions and species. Consequently, the means, standard deviations, minimum, and maximum values (their ranges) of the traits were derived from these adjusted mean values in both environments. The means, standard deviations, minimum and maximum values (ranges) of the traits were calculated for the species from these adjusted mean values in both environments. In addition, measurements were made for 30 common accessions grown in both environments, and the mean and standard deviation values were obtained by calculating the emmeans values among the five common traits, considering the environment, replication, and block effects.

For assessing the statistical significance of phenotypic differences among species, *emmeans* values of accessions in both locations were compared using Tukey's Honest Significant

Difference (HSD) test. Tukey's HSD test was conducted using the *Ime4, emmeans*, and *multcompView* packages in R. Group comparisons of accessions were visualized using compact letter display (CLD). That means groups sharing the same letter are not significantly different from each other at the 5% significance level.

The Anderson-Darling and Shapiro-Wilk normality tests were performed using the nortest R package. By comparing both normality tests, traits that did not conform to a normal distribution were identified for both environments. Z-transformations were performed on traits that did not conform to a normal distribution assumption using the *dplyr* R package. Residual values were obtained by removing replication and block effects using the *dplyr* package. For the Tekirdag-23 environment, pearson correlation plot was crated by using *GGally* and *ggplot2* in R. For the Salina-24 environment, since PlStEr and PlRGR traits are ordinal data, their medians of populations were used to obtain the value in the population for each individual and were included in the analysis. Then, *GGally* and *ggplot2* packages were used for the visualization of a Spearman correlation plot.

A table was created using traits thought to be highly correlated with forage and seed yield: canopy height, forage dry matter, raceme number, legume fruit mass, and legume fruit 1000-kernel weight. The performance of accessions and six cultivars grown around Salina-24 environment was examined with respect to these traits. Accessions with the highest average performance values were listed and included to illustrate the performance of commonly used cultivars in terms of forage and seed yield.

Accessions grown in the Tekirdağ-23 and Salina-24 environments were ranked based on the mean values of forage agronomic traits, adjusted for replication and block effects. Ranking was performed using the TOPSIS method for multiple-criteria decision making (MCDM) with the *dplyr* and *topsis* R packages.

Broad-sense heritability (H^2) estimation was calculated to demonstrate the extent to which genetic variation accounts for the phenotypic variation in traits among sainfoin plants in the Tekirdağ-23 and Salina-24 environments. First, a mixed model was constructed to obtain the variance components. Access, replication, and block effects were added to the model as random

effects. The R packages *lme4* and *lmerTest* were used to calculate H^2, population-wise in the Tekirdağ-23 environment and individual-wise in the Salina-24 environment.

$$H^2 = \frac{\sigma_G^2}{\sigma_P^2} = \frac{\sigma_G^2}{\sigma_G^2 + \frac{\sigma_E^2}{r}}$$

σ^2G = genotypic variance

σ^2E = residual variance

σ^2P = total phenotypic variance (Hallauer, Filho, & Carena, 2010).

A colored cladogram was drawn for each trait in the accessions around Tekirdağ and Salina, according to species. The R package *missMDA* was used to apply a PCA-based imputation method for missing data in the data file. The R packages *dendextend* and *ape* were used to plot the cladogram. The distance matrix derived from the trait means was used for the resulting dendrogram. Black represents *O. viciifolia*, red represents *O. transcaucasica*, blue represents *O. arenaria*, and orange represents cultivars and breeding populations.

Principal Component Analysis (PCA) was performed to examine the distribution of sainfoin species for both environments. The R package *missMDA* was used to apply a PCA-based imputation method for missing data in the data file. After imputation, PCA results were obtained using the prcomp function, and the variance percentages for PC1 and PC2 were calculated for both environments. PCA plots were generated using the *dplyr* and *ggplot2* R packages.

Subsequently, a scatter plot was generated for forage dry mass (FoDM) and legume fruit mass (FrM) traits measured at the Salina-24 environment to observe the performance of 13 accessions and 6 cultivars in terms of both forage yield and seed yield. This scatter plot was created in R using the *ggplot2, dplyr,* and *ggrepel* packages.

3.6.2. Phenomic Data Analyses

Each individual sample was separated and measured as two grams. Seed images with pods and without pods were captured respectively for seed yield-related traits. Quantitative data obtained from the seed images was performed using a seed image processing method. Subsequently, CSV files were read and written by using the *csv* package in Python, while file and directory

operations were managed using the *os* package in Python. Numerical calculations were then conducted using the *numpy* package.

The analyses performed in the Phenotyping section were also performed for seed data obtained by imaging method. To assess the normality of phenotypic data from both locations, the nortest package in R was used to perform normality test via the Anderson-Darling test (*ad.test*) and the Shapiro-Wilk test (*shapiro.test*). Traits that did not confirm to a normal distribution were transformed using Z-transformation with the *dplyr* package in R program. Analysis of variance (ANOVA) were done by obtaining F-values and significance levels for each trait, based on species and accession in both locations, were analyzed using the *lmerTest* package in R program.

To obtain the adjusted means, they were obtained to remove replication and block effects from the data. For this purpose, the raw dataset was analyzed using the *Matrix* and *lme4* packages in R to estimate the means for accessions and species. Furthermore, the *emmeans, pbkrtest*, and *lmerTest* packages in the R language program were used to compute the adjusted means for phenotypic traits based on accessions and species. Consequently, the means, standard deviations, minimum, and maximum values (their ranges) of the traits were derived from these adjusted mean values in both environments. The means, standard deviations, minimum and maximum values (ranges) of the estimated seed traits were calculated for the species from these adjusted mean values in the Salina-24 environment.

To assess the statistical significance of phenotypic differences among species, *emmeans* values of accessions in both locations were compared using Tukey's Honest Significant Difference (HSD) test. Tukey's HSD test was conducted using the *lme4, emmeans*, and *multcompView* packages in R. Group comparisons of accessions were visualized using compact letter display (CLD). That means groups sharing the same letter are not significantly different from each other at the 5% significance level.

Accessions grown in the Salina-24 environment were ranked based on the mean values of estimated seed traits, adjusted for replication and block effects. Ranking was performed using

the TOPSIS method for multiple-criteria decision making (MCDM) with the *dplyr* and *topsis* R packages.

Broad-sense heritability (H^2) was calculated to demonstrate the extent to which genetic variation accounts for the phenotypic variation in traits among sainfoin plants in the Tekirdağ-23 and Salina-24 environments. First, a mixed model was constructed to obtain the variance components. Accession, replication, and block effects were added to the model as random effects. The R packages *lme4* and *lmerTest* were used to calculate H^2, population-wise in the Tekirdağ-23 and individual-wise in the Salina-24.

$$H^2 = \frac{\sigma_G^2}{\sigma_P^2} = \frac{\sigma_G^2}{\sigma_G^2 + \frac{\sigma_E^2}{r}}$$

σ^2G = genotypic variance
σ^2E = residual variance
σ^2P = total phenotypic variance (Hallauer et al., 2010)

A colored cladogram was drawn for each estimated seed trait in the accessions in Salina-24 environment, according to species. The R package *missMDA* was used to apply a PCA-based imputation method for missing data (NA) in the data file. The R packages *dendextend* and *ape* were used to plot the cladogram. The distance matrix derived from the trait means was used for the resulting dendrogram. Black represents *O. viciifolia*, red represents *O. transcaucasica*, blue represents *O. arenaria*, and orange represents cultivars and breeding populations.

Principal Component Analysis (PCA) was performed to examine the distribution of sainfoin species for Salina-24 environment. The R package *missMDA* was used to apply a PCA-based imputation method for missing data in the data file. After imputation, PCA results were obtained using the *prcomp* function, and the variance percentages for PC1 and PC2 were calculated for Salina-24 environment. PCA plots were generated using the *dplyr* and *ggplot2* R packages.

3.6.3. Genomic Data Analyses

A total of 430 individuals, 10 samples from 43 different populations in the field in Tekirdağ, were previously sequenced. After sequencing, a HapMap file containing 50K SNPs was

obtained. This HapMap file was analyzed by manually creating commands with the help of existing packages and commands in the R program. Each command used during the genotypic analysis will be explained step by step in the appendix of the thesis.

The SNP dataset, initially in HapMap format, comprised 46.613 SNP and 430 variables. For downstream analyses, two new columns, allele1 and allele2, were added using the *tidyverse* package in R. During the analysis of sequence data, allele frequencies of individuals were calculated on a population basis. To estimate SNP variation at each chromosomal position, the presence of SNPs in the allele1 column was counted and multiplied by 0.25. Since individuals were sampled randomly, the order of plants in the dataset was also random. To facilitate further analysis, the dataset was reorganized so that individuals from the same population were grouped together. Based on this new ordering, mean allele frequencies were calculated for each population. The resulting dataset was transposed and converted into both a matrix and a data frame. The average allele frequency on a population basis was visualized by Principal Component Analysis (PCA) in R program. PCA was performed using the allele frequency matrix, and the PCA plot was generated with the *ggplot2* package. With the PCA plot, the closeness, and relationships of the populations with each other on a plane were visualized. Finally, Euclidean distance matrix was computed from the same dataset, and a Neighbor Joining (NJ) tree was constructed using the *ape* package in R.

The mean allele frequencies of the populations were calculated, and the data file was transposed. The distance matrix was then calculated using the Euclidean method. Analysis of Molecular Variance (AMOVA) was performed at the population level based on the Euclidean distance matrix. For AMOVA, population species were defined as population groups. AMOVA was performed using the *pegas* package in the R package, taking the distance matrix and population groups into account.

4. RESULTS AND DISCUSSIONS

4.1. Evaluation of traits related to forage and seed yield

4.1.1. Effects of Environments

The environment effect was highly significant for all vegetative traits ($p < 0.001$). However, RaN, one of the traits related to flowering, did not show a significant environmental difference ($p > 0.05$). The species effect was particularly significant for StTh ($p < 0.001$) and RaN ($p < 0.01$), while variation was observed at the accession level for all traits ($p < 0.001$). The environment x species interaction was significant for StL ($p < 0.05$) and RaN ($p < 0.01$). The environment x accessions interaction was particularly significant for the StN, StL, and RaN traits ($p < 0.001$) (Table 4.1). Since the environmental effects are highly significant on vegetative traits, all further analyses were conducted separately for both environments.

Table 4.1. F-values along with the significance from the analysis of variance (ANOVA) for common agronomic traits measured across two different environments

	Df	CanHt†	StTh†	StN†	StL†	RaN†
Environment	1	107.76***	149.39***	38.26***	34.99***	0.14
Species	2	0.78	12.08***	0.20	1.31	5.17**
Accessions	27	1.66	3.16***	1.35	3.04***	2.06**
Environment*Species	2	0.77	2.72	2.04	4.15*	4.29*
Environment*Accessions	26	1.64	1.41	2.45***	3.88***	2.08**

†CanHt: Canopy height; StTh: Stem thickness; StN: Stem number; StL: Stem length; RaN: Raceme number. Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

In this study, genetic diversity and environmental effect were found to have a significant impact on phenotypic diversity in sainfoin species. The dominant influence of environmental factors on CanHt, StTh, StN, and StL traits indicates that the growing environment is a major factor for plant development. Therefore, the obtained variance components provide essential clues for determining target traits (such as StL and StTh) for breeding programs. Sayar et al. (2022) reported a study investigating forage-related yield and quality traits of genotypes of *Onobrychis viciifolia* in the Southeastern Anatolia region of Türkiye. They observed that both year and sainfoin genotypes had significant effects on all measured traits. They reported that climatic variation and environmental changes directly affected key productivity traits such as fresh forage yield and dry matter yield. Therefore, considering environmental factors such as year

and year-environment interactions that directly affect yield-related traits in breeding studies, and further investigating traits in multienvironment trials will contribute to breeding studies. Another study examined 44 accessions of *Onobrychis viciifolia* Scop. across three different environments. Location effects were found to be significant for most traits, while accession x location interactions were also significant for some traits (Delgado, Salvia, Buil, & Andrés, 2008).

Analysis of variance revealed significant species-level differences for RaL ($p < 0.01$) and InfL ($p < 0.05$), that indicate variation among species in terms of forage and seed traits. Accession effects were highly significant for StL ($p < 0.001$) and significant for traits including FoSDM, StN1, RaN, RaL, InfL, and SdN ($p < 0.01$) that reflect notable intra-specific variation. Unexpectedly, CanHt and StN2 showed no significant variation at either the species or accession level that would suggest similar performance in the Tekirdağ-23 environment (Table 4.2). Canopy height is the natural plant height and is indicated to be very environment depended in other forages such as alfalfa (Sakiroglu, Moore, & Brummer, 2011).

Table 4.2. F-values along with the significance from the analysis of variance (ANOVA) of agronomic traits in the Tekirdağ-23 environment

Source	Species	Accession	H^2
CanHt†	1.93	0.96	0.27
StTh	1.44	1.3	0.26
StL	1.35	2.43***	0.7
FoSDM	1.44	1.46	0.63
StN1	1.85	2.13**	0.52
RaN	1.73	1.44	0.41
StN2	0.55	0.81	0
RaL	4.08**	1.53	0.55
InfL	2.78*	2.10**	0.71
SdN	2.1	1.55	0.36

† CanHt: Canopy height; StTh: Stem thickness; StL: Stem length; FoSDM: Forage subsample dry mass; StN1: 1st stem number; RaN: Raceme number; StN2: 2nd stem number; RaL: Raceme length; InfL: Inflorescence length; SdN: Seed number per raceme. Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

In Tekirdağ-23, significant differences were observed only in raceme length and inflorescence length at the species level. While significant differences were observed in stem length at the accession level, some traits related to forage yield and seed yield (e.g., FoSDM, StN1, RaN, RaL, InfL, SdN) showed significant differences. Traits that differ at the species and accession

levels can be used in breeding studies to increase diversity in the traits of interest. In Mohajer et al (2011), data were collected and analyzed for seed yield, dry matter yield, plant height, stem count per square meter, flowering date, flower cluster length, and 1000-seed weight. Significant differences were observed among *O. viciifolia* populations for all traits except 1000-seed weight and stem count per square meter (Mohajer et al., 2011). This study is consistent with the observation of significant differences among the populations obtained in both environments.

The results of analysis of variance (ANOVA) for various agronomic traits revealed significant differences among both species and accessions (Table 4.3). CanHt, StL, StN, PlStEr, PlRGR, and Fr1KK were found to be highly statistically significant ($p < 0.01$ or $p < 0.001$), while StTh, RaN, and FrM traits showed lower levels of difference ($p < 0.05$) (Table 4.3).

Table 4.3. F-values along with the significance from the analysis of variance (ANOVA) and H^2 heritability of forage-related traits in the Salina-24 environment

Source	Species	Accession	H^2
CanHt†	4.26**	4.18***	0.597
StTh	3.05*	1.84**	0.064
StL	4.96**	2.22***	0.458
FoDM	2.99*	1.48*	0.386
StN	11.03***	2.76***	0.345
RaN	3.57*	2.03***	0.233
PlStEr	46.94***	15.36***	0.484
PlRGR	11.40***	3.50***	0.427
FrM	3.63*	2.08***	0.300
Fr1KK	6.26***	6.05***	0.685

† CanHt: Canopy height; StTh: Stem thickness; StN: Stem number; StL: Stem length; FoDM: Forage dry mass; RaN: Raceme number; PlStEr: Plant stem erectness; PlRGR: Plant regrowth rating; FrM: Legume fruit mass; Fr1KK: Legume fruit 1000K weight. Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Analysis of variance results indicate that genetic diversity is present for a significant portion of the agronomic traits measured in the Salina-24 environment, and that there are statistically significant differences between species and accessions. Another study found that variation among cultivars was statistically significant for most traits examined (Delgado et al., 2008), and the result is consistent with the findings of the present study. High F values and essential differences in characteristics such as canopy height, stem length, plant stem erectness, and

legume fruit 1000-kernel weight indicate the presence of a high level of genetic diversity in these traits.

4.1.2. Trait means and Correlations among traits

Mean, standard deviation (SD), and observed ranges of key agronomic traits measured in 47 sainfoin genotypes evaluated in the Tekirdağ-23 environment and 80 genotypes evaluated in the Salina-24 environment (Table 4.4). Significant variation was observed among the phenotypic traits across both environments. Generally, accessions grown in the Salina-24 environment performed higher mean values for most vegetative and reproductive traits compared to those grown in the Tekirdağ-23 environment. For example, canopy height averaged 66.47 cm in Salina-24 environment, whereas it was 50.53 cm in Tekirdağ-23 environment. In addition, other common traits like stem thickness, stem length, and stem number per plant had higher performance in the Salina-24 environment (Table 4.4).

Table 4.4. The comparison of mean, standard deviation and their ranges of agronomic traits in both environments

		2023-Tekirdağ/Türkiye		2024-Salina/KS/USA	
Traits	**Unit**	**Mean±SD**	**Range**	**Mean±SD**	**Range**
CanHt	cm	50.53±10.59	28.44-68.57	66.47±10.01	26.1-79.9
StTh	mm	4.67±1.00	0.21-5.93	6.78±1.75	3.71-20.69
StL	cm	64.18±12.37	23.8-86.08	78.99±7.72	55.7-98.5
StN1/StN	n	36.61±10.73	4.43-62.24	54.61±10.94	25.9-89.4
RaN	n	3.19±0.82	1.14-5.27	13.05±2.17	7.97-19.82
FoSDM	g	227.59±74.39	73.9-382.3	NA	NA
StN2	n	31.09±8.94	8.35-47.13	NA	NA
RaL	cm	9.29±2.26	3.58-12.25	NA	NA
InfL	cm	21.22±3.93	9.13-26.99	NA	NA
SdN	n	14.91±4.66	0.22-21.68	NA	NA
FoDM	g	NA	NA	324.72±79.13	35-447
FrM	g	NA	NA	47.8±14.97	2.67-83.33
PlStEr	1-9	NA	NA	3.24±1.07	1.13-7.89
PlRGR	0-5	NA	NA	2.18±0.6	0.44-3.92

† CanHt: Canopy height; StTh: Stem thickness; StL: Stem length; StN1: 1st stem number; RaN: Raceme number; FoSDM: Forage subsample dry mass; StN2: 2nd stem number; RaL: Raceme length; InfL: Inflorescence length; SdN: Seed number per raceme; FoDM: Forage dry mass; FrM: Legume fruit mass; PlStEr: Plant stem erectness; PlRGR: Plant regrowth rating. *NA: non-available

Comparative analysis of agronomic traits at two locations revealed significant differences, likely highlighting the importance of environmental conditions and genotype-environment interactions. These findings highlight the importance of multi-location trials in identifying stable, high-performing genotypes across diverse environments. Furthermore, the observed variation may provide critical information for selection strategies in sainfoin breeding programs to improve both vegetative vigor and reproductive output.

The mean values and observed ranges of various agronomic traits among sainfoin species grown in Tekirdağ-23 environment were presented in the Table 4.5. The analysis of variance using by Tukey's multiple comparison test found the significant differences among species and breeding populations/cultivars for few agronomic traits in Tekirdağ-23 environment. Such data are highly valuable for identifying interspecific agronomic differences and determining the potential traits for breeding purposes. The breeding populations/cultivars (b) was significantly taller than *O. arenaria* (ab) and *O. viciifolia* (a) and *O. transcaucasica* (a) in terms of canopy height. Stem length (StL) was the longest value in breeding populations/cultivars (b), while the modest value was observed in *O. viciifolia* (ab). *O. arenaria* (a) and *O. transcaucasica* (a) were statistically close to each other. Also, forage subsample dry matter (FoSDM) was highest in breeding populations/varieties (b) and significantly outperformed *O. arenaria* (a), *O. viciifolia* (a), and *O. transcaucasica* (a). StN1, RaN, RaL, StN2, InfL, and SdN traits, such as the StTh trait, did not differ significantly among species and breeding populations/cultivars (Table 4.5).

Table 4.5. Descriptive statistics (mean, minimum, and maximum) of estimated marginal means (emmeans) with standard errors of agronomic performance based on sainfoin species and showing their statistical differences in the Tekirdağ-23 environment

Species	Unit	*O. arenaria*	Breeding populations / cultivars	*O. transcaucasica*	*O. viciifolia*
CanHt†	cm	46.8±5.68^{ab} (28.2-45.7)	58.5±1.97^{b} (40.1-68.6)	49.5±1.84^{a} (36.4-61.3)	45.5±2.50^{a} (36.8-55.2)
StTh	mm	4.65±0.28^{a} (4.07-4.96)	4.94±0.11^{a} (4.25-5.49)	4.98±0.10^{a} (3.86-5.7)	4.66±0.13^{a} (3.64-5.75)
StL	cm	57.1±5.62^{a} (54.5-66.9)	72.1±2.45^{b} (51.2-86.2)	64.9±2.36^{a} (32.8-83.6)	65.8±2.86^{ab} (52.7-77.3)
FoSDM	g	171±46.70^{a} (160-193.9)	297±29.10^{b} (153-387.3)	217±28.70^{a} (156.2-313.1)	224±31^{a} (130.7-339)
StN1	n	45±5.06^{a} (39.03-49.11)	41.1±1.86^{a} (23.44-47.56)	35.4±1.76^{a} (4.29-52.11)	38.4±2.33^{a} (26-62.49)

Species	Unit	*O. arenaria*	Breeding populations / cultivars	*O. transcaucasica*	*O. viciifolia*
RaN	n	3.37±0.42[a] (2.65-3.94)	3.01±0.18[a] (2.22-3.72)	3.42±0.17[a] (2.61-4.56)	3.52±0.21[a] (2.39-5.28)
StN2	n	34.5±5.49[a] (27.5-35.6)	31.8±1.91[a] (20.96-41)	32.6±1.78[a] (8.37-42.89)	33.9±2.42[a] (19.73-47.13)
RaL	cm	11.07±0.95[a] (10.83-11.86)	9.85±0.50[a] (7.38-12.16)	9.88±0.49[a] (8.07-12.14)	8.71±0.55[a] (4.5-10.51)
InfL	cm	24.8±1.62[a] (26-26.6)	22.4±0.82[a] (19.4-24.4)	22±0.80[a] (16.6-27.4)	20.7±0.91[a] (15.4-25.1)
SdN	n/Ra	17.9±1.91[a] (15.65-19.97)	15.4±0.73[a] (11.57-21.06)	14.9±0.69[a] (10.21-20.96)	15.4±0.89[a] (8.73-22.31)

† CanHt: Canopy height; StTh: Stem thickness; StL: Stem length; FoSDM: Forage subsample dry mass; StN1: 1st stem number; RaN: Raceme number; StN2: 2nd stem number; RaL: Raceme length; InfL: Inflorescence length; SdN: Seed number per raceme

Comparative assessment of sainfoin species and breeding populations in the Tekirdağ-23 environment revealed notable differences in agronomic performance. These findings suggest that both breeding populations and selected varieties may have the potential to be valuable genetic resources for future sainfoin breeding programs. Breeding populations/cultivars in the Tekirdağ-23 environment generally outperformed sainfoin species in traits related to forage yield, such as plant canopy height, stem length, and dry mass of the forage subsample. This finding provides clear evidence that selection and breeding efforts effectively increased biomass potential over the time.

The mean values and observed minimum and maximum values of agronomic traits of different sainfoin species and cultivars grown in the Salina-24 environment were presented in Table 4.6. The analysis of variance using by Tukey's multiple comparison test found the significant differences among sainfoin species and cultivars for many forage-related traits in Salina-24 environment. Considerable variation among species was observed across almost all measured traits. For example, in terms of CanHt, individuals of O. transcaucasica (b) and O. arenaria (b) species had better performance than individuals of cultivars and O. viciifolia (a). StTh was highest in O. arenaria (b) and O. transcaucasica (b), cultivars (ab) had modest stems, O. viciifolia (a) also had lowest stems. StL was highest in O. arenaria (c), followed by O. transacuasica (b) and modest group was cultivars and O. viciifolia had the shortest stems. FoDM in O. arenaria was significantly highest, followed by O. transacaucasica, O. viciifolia and

cultivars had the lowest groups. StN was highest in O. transcaucasica (b), O. viciifolia (b) and O. arenaria (b) and cultivars (a) were significantly lowest. RaN were lower in O. viciifolia than in O. arenaria and cultivars. PlStEr was widely lower in O. viciifolia, while the other groups performed statistically similarly (3.41-3.90). PlRGR ranged from 2.17 to 2.39, but no significant differences were observed between species and cultivars in terms of PlRGR. In terms of FrM *O. transacaucasica* (c) had highest performance than *O. arenaria* (bc), *O. viciifolia* (a), and cultivars (ab) Cultivars (b) and *O. transacaucasica* (b) had same and highest performance than other two species (a) in terms of Fr1KK (Table 4.6).

Table 4.6. Descriptive statistics (mean, minimum, and maximum) of estimated marginal means (emmeans) with standard errors of agronomic performance based on sainfoin species and showing their statistical differences in the Salina-24 environment

Species	Unit	*O. arenaria*	Cultivars	*O. transcaucasica*	*O. viciifolia*
CanHt†	cm	69.8±2.53^{b} (58.3-79.6)	62.5±3.45^{a} (56.7-68.5)	71.3±2.63^{b} (64.2-76.8)	61.5±2.75^{a} (26.1-79.9)
StTh	mm	6.77±0.17^{b} (5.42-20.69)	6.54±0.26^{ab} (5.83-7.62)	6.78±0.18^{b} (5.86-8.75)	5.97±0.19^{a} (3.71-6.95)
StN	n	56.7±3.14^{b} (41.5-89.4)	43.3±4.34^{a} (25.9-49.2)	58.3±3.27^{b} (41.7-69.3)	57.2±3.43^{b} (29.6-82.5)
StL	cm	83.4±2.17^{c} (75.4-92.8)	73.3±30^{ab} (67.8-86.5)	79.1±2.26^{b} 70.7-98.5)	72.4±2.37^{a} (55.7-90.4)
RaN	n	14±0.98^{c} (9.25-18.24)	14.2±1.19^{bc} (10.73-19.82)	12.8±1.01^{b} (9.24-14.98)	11.2±1.03^{a} (7.97-14.95)
FoDM	g	381±21.90^{c} (317-447)	252±28.90^{a} (147-291)	350±22.70^{b} (271-386)	289±23.60^{a} (35-433)
PlStEr	0-9	3.41±0.14^{b} (1.94-6.73)	3.9±0.34^{b} (2.89-7.89)	3.42±0.17^{b} (2.65-5.66)	2.74±0.20^{a} (1.13-3.85)
PlRGR	1-5	2.25±0.07^{a} (1.46-3.2)	2.35±0.19^{a} (1.37-2.43)	2.39±0.09^{a} (1.35-3.07)	2.17±0.11^{a} (0.44-3.92)
FrM	g	52.9±1.98^{bc} (26.21-83.33)	43.6±3.97^{ab} (35.19-54.52)	56.1±2.37^{c} (44.43-75.74)	35±2.51^{a} (2.67-75.19)
Fr1KK	g	16.9±0.326^{a} (14.84-19.69)	20.5±0.59^{b} (18.67-22.08)	20±0.37^{b} (17.56-21.71)	17.5±0.39^{a} (8.51-21.05)

† CanHt: Canopy height; StTh: Stem thickness; StN: Stem number; StL: Stem length; FoDM: Forage dry mass; RaN: Raceme number; PlStEr: Plant stem erectness; PlRGR: Plant regrowth rating

Generally, it is observed that the *O. arenaria* and *O. transcaucasica* species performed better than *O. viciifolia* in terms of agronomic traits in the Salina-24 environment. The superior performance of *O. arenaria* in both forage dry mass (FoDM) and stem morphology suggests a

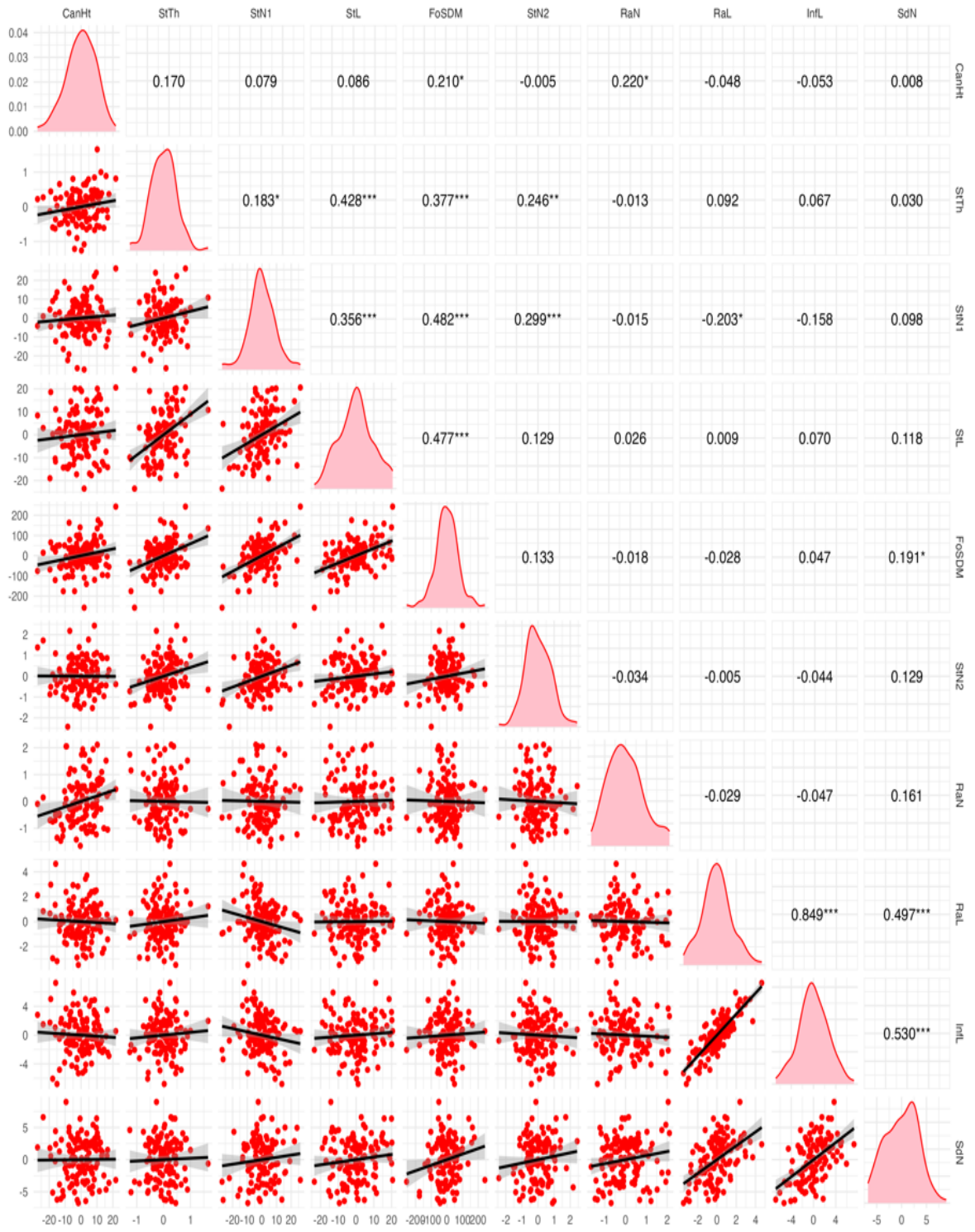
stronger adaptation to the Salina-24 environment. Therefore, *O. arenaria* and *O. transcaucasica*, which are wild *Onobrychis* species, exhibited significant forage performance, characterized by high biomass, long stems, and improved structural characteristics. Inclusion of these species in breeding programs can make substantial contributions to enhancing sainfoin forage yield. In terms of seed-related yield traits (FrM and Fr1KK), *O. transcaucasica* had better performance. Since there was no significant difference in plant regrowth rating, post-harvest regrowth potential can be considered consistent among species and cultivars.

Comparisons among sainfoin species in the Salina-24 environment reveal significant phenotypic diversity that can contribute to future sainfoin breeding studies. *O. arenaria* and *O. transcaucasica* exhibited superior vegetative growth characteristics, indicating their adaptation to the Salina/Kansas environmental conditions and the potential to serve as valuable genetic resources for biomass improvement studies. These observed variations highlight the importance of interspecific evaluation in sainfoin development, and *O. arenaria* and *O. transcaucasica* are promising sources for improving both forage efficiency and seed yield.

A study in southern Serbia investigated how row spacing affected seed yield and quality in plant populations of *O. viicifolia*. They emphasized that wider row spacings resulted in better 1000-kernel weights. They also observed the highest 1000-kernel weight at 40 cm (approximately 22.1 g) and the lowest 1000-kernel weight at 20 cm (20.6 g) (Beković, Stojković, Biberdžić, & Knežević, 2016). In the present study, a 75 cm row spacing was used, and 1000-kernel weights measured in the Salina/Kansas, ranged from 8.51 to 21.05 g, resulting in a wide range of variation. These findings suggest that seed size can be significantly affected by agronomic practices and growing conditions.

In another study, the agronomic characteristics of populations of different species (*O. viciifolia* and *O. altissima*) were investigated, and it was reported that canopy height in 20 populations varied between 70 and 100 cm (Cebeci, 2011). Although *O. altissima* populations were not examined in this study, canopy heights of *O. viciifolia* populations were measured between 26.1 and 79.9 cm.

Ten agro-morphological traits of sainfoin measured in the Tekirdağ-23 environment was to find out correlations among traits in Figure 4.1. Significant positive correlations were observed between several stem and reproductive traits. Especially, StTh exhibited a significant positive correlation with StL (r = 0.428*) and FoSDM (r = 0.377*) that means thicker stems were generally associated with taller plants and greater biomass production. In addition, StL was positive correlation with FoSDM (r = 0.482*) and SdN (r = 0.356*) that indicates longer stems may contribute to both vegetative and reproductive traits (Fig. 4.1). Among reproductive traits, InfL has a strong positive correlation with SdN (r = 0.849*). Also, RaL has a strong positive correlation with SdN (r = 0.497*) that indicates this kind of correlations highlights an important structural linkage between inflorescence length and seed number. Furthermore, RaL and InfL were positive correlated (r = 0.530*) each other, that reflect there is coordinated development within reproductive traits (Figure 4.1). In contrast, there are some traits like RaN that non-significant correlations with other measured traits (Figure 4.1).

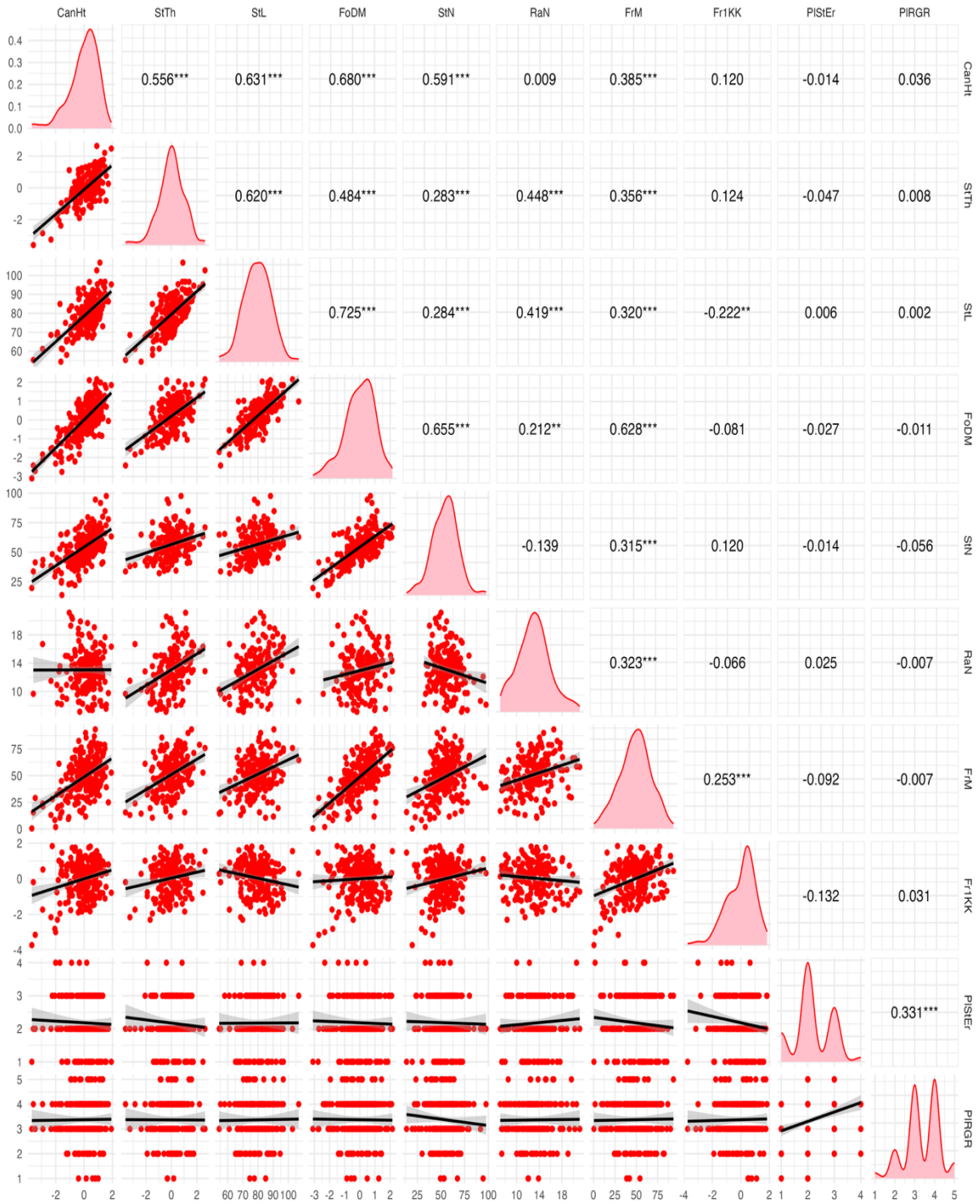


*CanHt: Canopy height; StTh: Stem thickness; StL: Stem length; FoSDM: Forage subsample dry mass; StN1: 1st stem number; RaN: Raceme number; StN2: 2nd stem number; RaL: Raceme length; InfL: Inflorescence length; SdN: Seed number per raceme

Figure 4.1. Pearson correlation plot of agronomic traits in sainfoin accessions in the Tekirdağ-23 environment

The most critical factors for increasing forage yield in the Tekirdağ-23 environment are thought to be stem length, stem thickness, and first cutting stem number. Furthermore, generative traits were significantly correlated with each other and were not directly linked to forage yield. Therefore, when forage yield is the primary goal in sainfoin breeding, it is recommended to use vegetative traits (CanHt, StL, FoSDM, and StTh). Conversely, when seed yield is the primary goal, generative traits (RaL, InfL, and SdN) may be used and preferred.

Statistically significant and mostly positive correlations were observed among the 11 agronomic traits examined in the Salina-24 environment (Figure 4.2). Strong positive correlations were observed between CanHt, StTh, StL, and FoDM, indicating a direct link between biomass yield and plant morphology. FoDM also showed a high correlation with seed yield (FrM) ($r = 0.63$, $p < 0.001$) and StN ($r = 0.42$), indicating that stem number and generative organ development support biomass production. Moderate correlations were found with RaN and FrM ($r = 0.32$), but RaN also has a weak correlation with FoDM ($r = 0.21$). Fr1KK showed a weak positive correlation with vegetative traits, indicating that big-size seeds may be under genetic control, rather than morphological. A moderately positive and significant correlation was observed between the ordinal traits PlStEr and PlRGR ($r = 0.33$). However, these two ordinal traits did not show significant correlation with other quantitative agronomic traits (Figure 4.2).



*CanHt: Canopy height; StTh: Stem thickness; StN: Stem number; StL: Stem length; FoDM: Forage dry mass; RaN: Raceme number; PlStEr: Plant stem erectness; PlRGR: Plant regrowth rating; FrM: Legume fruit mass; Fr1KK: Legume fruit 1000K weight

Figure 4.2. Spearman correlation plot of agronomic traits in 80 sainfoin accessions in the Salina-24 environment

Correlation analysis results revealed that biomass production in sainfoin (*Onobrychis* spp.) is largely dependent on vegetative traits in the Salina-24 environment. The high and significant correlation coefficients between CanHt, StTh, and StL indicate that plant height and stem number and structure are strong determinants for investigating dry matter yield. Furthermore, the relationships between FoDM and StN, and FrM may provide important links to yield. There are weak or moderate correlations between RaN, Fr1KK, and FrM and vegetative traits; this suggests that biomass yield and seed formation are partially independent processes. PlStEr and PlRGR did not show a significant correlation with other agronomic traits ($p > 0.05$), suggesting that these two traits are more related to plant architecture and physiology.

In an another correlation analysis conducted for 10 populations of *O. viciifolia*, dry matter yield was positively correlated with seed yield, plant height, stem number, and inflorescence length. Seed yield was positively and significantly correlated with 1000-grain weight, plant height, stem number, and inflorescence length (Mohajer et al., 2011). These results are consistent with the findings in this study that traits related to seed yield are directly related to forage yield.

4.1.3. Phenotypic Variation among Sainfoin Accessions

Means and standard deviations from two locations (Tekirdağ, Türkiye and Salina, Kansas) were evaluated in Table 4.7. Significant differences were observed among the sainfoin accessions for all morphological traits evaluated. Canopy height varied between 30.6 cm (PI 302936) and 66.4 cm (PI 312991). In addition, accessions PI 312991, PI 273783, and PI 273782 showed higher canopy height than 64 cm, which indicates the presence of accessions with strong vegetative growth potential. In contrast, PI 494668 and PI 302936 had the shortest canopy length. For stem length, PI 251697 showed the highest mean (84.51 cm), closely followed by PI 312965 (82 cm) and PI 273758 (79.47 cm). Interestingly, PI 494668 (31.99 cm) and PI 313023 (8.54 cm) performed quite short stems that indicate stem length varies significantly among accessions. For stem thickness measurement, PI 223389 (13.06 mm) showed the highest mean, PI 251697 (8.43 mm) and PI 273782 (8.35 mm) accessions followed. In general, most of the other accessions ranged from 5.49 to 7.24 mm. The number of racemes per plant varied greatly, with PI 223389 (4.17) and PI 312991 (3.65). In contrast, PI 302936 had the lowest mean raceme number (1.09) that referring to limited inflorescence production in both environments. Stem number ranged from 38.1 (PI 302936) to 57.4 (PI 273783), while PI

313041, PI 313008, and PI 313005 also showed higher stem numbers (> 50). These accessions may support the potential for vegetative propagation and forage biomass production (Table 4.7).

Table 4.7. Means and standard deviations of estimated marginal means (emmeans) of the common traits in the common accessions across two environments

Population	Canopy height	Stem length	Stem thickness	Raceme number	Stem number
PI 312965	50 ± 5.57	82 ± 22.4	8.22 ± 13.3	1.58 ± 3.83	45.7 ± 7.89
PI 312969	55.2 ± 4.56	69.9 ± 18.7	6.02 ± 11.1	2.74 ± 3.11	45.1 ± 6.42
PI 312967	51.4 ± 5.57	NA	NA	NA	11 ± 7.89
PI 313003	51 ± 4.56	67.29 ± 18.7	5.7 ± 11.1	2.79 ± 3.11	43.9 ± 6.42
PI 313008	57.2 ± 4.56	78.11 ± 18.7	5.73 ± 11.1	2.88 ± 3.11	50.5 ± 6.42
PI 313041	61.5 ± 4.56	44.08 ± 18.7	NA	2.84 ± 3.11	51 ± 6.42
PI 313023	59 ± 4.56	8.54 ± 18.7	NA	2.52 ± 3.11	47.5 ± 6.42
PI 312991	66.4 ± 4.56	78.73 ± 18.7	5.96 ± 11.1	3.65 ± 3.11	42 ± 6.42
PI 313001	60 ± 4.56	68.26 ± 18.7	5.82 ± 11.1	3.08 ± 3.11	50.1 ± 6.42
PI 312994	56 ± 4.56	70.21 ± 18.7	5.49 ± 11.1	2.6 ± 3.11	43.2 ± 6.42
PI 273769	63.7 ± 4.56	72.04 ± 18.7	6.43 ± 11.1	2.32 ± 3.11	43.6 ± 6.42
PI 273780	59.1 ± 4.56	35.5 ± 18.7	NA	2.89 ± 3.11	45.5 ± 6.42
PI 312985	64.3 ± 4.56	71.59 ± 18.7	5.98 ± 11.1	3.45 ± 3.11	44.5 ± 6.42
PI 372816	60.4 ± 4.56	40.62 ± 18.7	NA	2.85 ± 3.11	42.1 ± 6.42
PI 273783	66.2 ± 4.56	72.02 ± 18.7	5.74 ± 11.1	2.94 ± 3.11	57.4 ± 6.42
PI 251698	58.7 ± 4.56	74.04 ± 18.7	6.22 ± 11.1	2.56 ± 3.11	49.1 ± 6.42
PI 313006	51.9 ± 4.56	60.72 ± 18.7	5.5 ± 11.1	3.36 ± 3.11	49.1 ± 6.42
PI 251697	53.1 ± 5.57	84.51 ± 22.4	8.43 ± 13.3	2.91 ± 3.83	40.4 ± 7.89
PI 273782	64.6 ± 5.57	67.9 ± 22.4	8.35 ± 13.3	2.48 ± 3.83	45.7 ± 7.89
PI 313005	59.1 ± 4.99	71.79 ± 20.2	7.24 ± 12	2.5 ± 3.42	50.9 ± 7.05
PI 273758	63.7 ± 4.56	79.47 ± 18.7	5.84 ± 11.1	3.07 ± 3.11	53.7 ± 6.42
PI 273790	58.6 ± 4.56	43.77 ± 18.7	NA	2.8 ± 3.11	40 ± 6.42
PI 223389	56.3 ± 7.86	75.05 ± 31	13.06 ± 18.5	4.17 ± 5.45	49 ± 11.2
PI 313047	52.8 ± 4.99	72.21 ± 20.2	6.57 ± 12	2.39 ± 3.42	51.5 ± 7.05
PI 313063	44.9 ± 4.56	66.67 ± 18.7	5.2 ± 11.1	2.28 ± 3.11	42.7 ± 6.42
PI 313053	57.9 ± 4.56	46.06 ± 18.7	NA	2.31 ± 3.11	47 ± 6.42
PI 494668	40.5 ± 4.56	31.99 ± 18.7	NA	3.35 ± 3.11	45.5 ± 6.42
PI 319713	59.6 ± 4.56	42.42 ± 18.7	NA	2.62 ± 3.11	39.5 ± 6.42
PI 311468	55.5 ± 4.56	65.5 ± 18.7	5.12 ± 11.1	2.49 ± 3.11	48.5 ± 6.42
PI 302936	30.6 ± 5.57	NA	NA	1.09 ± 3.83	38.1 ± 7.89

*NA: non-available

When the variation observed among 30 sainfoin accessions grown in two different environments was examined, accessions PI 312991, PI 273783, PI 273758, and PI 223389 consistently performed well in multiple traits. The high performance of PI 312991 suggests that it may be a good candidate for sainfoin breeding studies aiming at both forage yield and seed yield. The high stem thickness performance and high raceme number observed in PI 223389 indicate a high potential for mechanical harvesting or grazing. In summary, the significant variation observed in phenotypic traits, such as canopy height and stem traits, provides a valuable resource for selecting genotypes with consistently superior performance under different growing conditions. In summary, future sainfoin breeding studies may prioritize the accessions PI 312991, PI 273783, PI 273758, and PI 223389 that I have mentioned earlier for trait and/or variety development.

Significant variation was observed among accessions in CanHt, FoDM, RaN, FrM, and Fr1KK in Salina-24 environment (Table 4.8). For CanHt trait, PI 600767 (79.9) cm) and PI312924 (79.6 cm) were ranked highest accessions, and PI 372804 (78.9 cm) and PI 273746 (78.3 cm) followed closely that they have indicated strong vegetative growth potential. Among the commercial sainfoin cultivars, AAC Mountainview (68.5 cm), Renumex (67.8 cm), and Shoshone (59.2 cm), performed moderate ranking rather than Eski (58.3 cm), Delaney (57.7 cm) and Rocky Mountain Remont (56.7 cm) (Table 3.10). A top list was created for the five main traits selected regarding the forage- and seed-related traits to determine six commercial cultivars among accessions. For forage dry mass, PI 312924 and PI 372798 (447 g) and PI 273786 (433 g) have highest values. Renumex (291 g), Shoshone (257 g), AAC Mountainview (236 g), and Eski (152 g) yielded performance among accessions. For raceme number, Delaney (19.82), PI 312967 (18.24), and PI 312969 (17.79) were the more efficient while Eski (13.51) and AAC Mountainview (10.73) performed less raceme number. Regarding legume fruit mass, PI 312919 (83.33 g), PI 313003 (75.74), and PI 273786 (75.19) accessions performed with high values than sainfoin cultivars. The best efficient cultivar was AAC Mountainview (54.52 g), while Delaney, Renumex, and Shoshone remained below 51 g in terms of legume fruit mass. Finally, for legume fruit 1000-kernel weight, Renumex (22.08 g), PI 312994 (21.71 g), and PI 313003 (21.28 g) had the highest values (Table 4.8).

Table 4.8. Top performing sainfoin accessions and cultivars for forage and seed yield traits in the Salina-24 environment

No	CanHt	CanHt_value (cm)	FoDM	FoDM_value (g)	RaN	RaN_value (n)	FrM	FrM_value (g)	Fr1KK	Fr1KK_value (g)
1	PI 600767 (1)	79.9	PI 312924 (1)	447	Delaney (1)	19.82	PI 312919 (1)	83.33	Renumex (1)	22.08
2	PI 312924 (2)	79.6	PI 372798 (2)	447	PI 312967 (2)	18.24	PI 313003 (2)	75.74	PI 312994 (2)	21.71
3	PI 372804 (3)	78.9	PI 273744 (3)	436	PI 312969 (3)	17.79	PI 273786 (3)	75.19	PI 313003 (3)	21.28
4	PI 273746 (4)	78.3	PI 273786 (4)	433	PI 312930 (4)	16.86	PI 312921 (4)	73.68	PI 313023 (4)	21.25
5	PI 312927 (5)	77	PI 312928 (5)	424	PI 312922 (5)	16.3	PI 312969 (5)	67.78	Rocky Mountain Remont (5)	21.16
6	AAC Mountainview (41)	68.5	Renumex (60)	291	PI 312965 (6)	15.9	AAC Mountainview (30)	54.52	PI 312992 (6)	21.1
7	Renumex (47)	67.8	Shoshone (67)	257	Renumex (16)	14.84	Renumex (36)	50.95	Delaney (7)	21.06
8	Shoshone (65)	59.2	AAC Mountainview (70)	236	Rocky Mountain Remont (21)	14.22	Shoshone (46)	45.59	PI 313048 (8)	21.05
9	Eski (66)	58.3	Rocky Mountain Remont (72)	225	Shoshone (22)	14.21	Rocky Mountain Remont (59)	39.95	Shoshone (15)	20.22
10	Delaney (69)	57.7	Eski (77)	152	Eski (31)	13.51	Delaney (65)	36.14	Eski (24)	19.72
11	Rocky Mountain Remont (70)	56.7	Delaney (78)	147	AAC Mountainview (66)	10.73	Eski (67)	35.19	AAC Mountainview (39)	18.67

*CanHt: Canopy height; FoDM: Forage dry mass; RaN: Raceme number; FrM: Legume fruit mass; Fr1KK: Legume fruit 1000K weight. * The numbers in parentheses next to the PI-numbered accessions and cultivars refer to the ranking of the accessions according to their performance in terms of the selected traits.

While wild accessions are significant in terms of morphological and biomass traits, cultivars are substantial in terms of reproduction and seed weight. Accessions in terms of higher biomass and morphological diversity, and cultivars in terms of reproductive performance and higher seed mass, complement each other. Therefore, the combined use of both wild genetic resources and cultivars plays a strategic role in improving both forage and seed yield in sainfoin breeding.

In the TOPSIS method analysis, all agronomic traits were included for both environments, and the performance ranking of all accessions was determined. However, an interpretation was made for the accessions in the top 10 for both locations (4.9). The TOPSIS method was used to rank the 44 sainfoin accessions/breeding populations/cultivars in the Tekirdağ-23 environment and the 76 sainfoin accessions/cultivars in the Salina-24 environment to identify the best-performing accessions in terms of all traits. Accessions PI 313047 and PI 273758 had the best performance, while Turkish modern cultivars Lütfübey and Özerbey were ranked 4th and 10th, respectively, in the Tekirdağ-23 environment. In terms of forage traits, the accessions numbered PI 273749 and PI 273748 showed the best performance, while AAC Mountainview, one of the cultivars, ranked fifth in the Salina-24 environment (Table 4.9).

Table 4.9. Ranking of sainfoin accessions/cultivars/breeding populations based on forage- and seed-related traits in both environments

	Tekirdag-23 environment		**Salina-24 environment**	
Rank	**Ranking based on traits**	**Score of traits**	**Ranking based on traits**	**Score of traits**
1	PI313047	0.65	PI273749	0.71
2	PI273758	0.63	PI273748	0.65
3	Breeding population6	0.63	PI273752	0.61
4	Lütfübey	0.63	PI273758	0.61
5	Breeding population11	0.62	AAC Mountainview	0.59
6	Breeding population13	0.61	PI273769	0.58
7	Breeding population8	0.61	PI312924	0.56
8	Breeding population9	0.60	PI372804	0.55
9	PI251698	0.57	PI372798	0.55
10	Özerbey	0.56	PI273786	0.55
11	PI273783	0.56	PI312922	0.55
12	Breeding population10	0.55	PI223389	0.54
13	Breeding population1	0.55	PI312927	0.54
14	PI313053	0.54	PI273746	0.53
15	PI273769	0.53	PI312919	0.53

	Tekirdag-23 environment		Salina-24 environment	
Rank	**Ranking based on traits**	**Score of traits**	**Ranking based on traits**	**Score of traits**
16	Breeding population2	0.52	PI313023	0.53
17	PI312994	0.52	PI273783	0.53
18	PI313008	0.52	PI251697	0.53
19	PI312991	0.51	PI312930	0.52
20	PI223389	0.51	PI372800	0.52
21	Breeding population3	0.50	PI313008	0.52
22	PI312969	0.49	PI372816	0.52
23	PI313023	0.48	PI273782	0.52
24	PI273781	0.46	PI313001	0.51
25	PI311468	0.46	PI312985	0.51
26	PI494668	0.46	PI368036	0.50
27	Breeding population4	0.45	PI372801	0.50
28	PI312965	0.43	PI600767	0.50
29	PI313001	0.43	PI273743	0.50
30	PI273790	0.42	PI251698	0.49
31	PI313063	0.42	PI505891	0.49
32	Breeding population5	0.42	PI312914	0.49
33	PI313041	0.42	PI312916	0.49
34	PI319713	0.42	PI313006	0.49
35	PI312985	0.40	PI313005	0.49
36	PI206458	0.39	PI372803	0.49
37	PI313005	0.39	PI313053	0.48
38	PI273780	0.37	PI273744	0.47
39	PI313003	0.37	PI312910	0.47
40	PI372816	0.36	PI312926	0.47
41	Breeding population12	0.35	PI312979	0.47
42	PI313006	0.34	PI401468	0.47
43	Breeding population7	0.29	PI312991	0.47
44	PI302936	0.25	PI312949	0.47
45	-	-	PI313041	0.46
46	-	-	PI313003	0.46
47	-	-	Renumex	0.46
48	-	-	PI312925	0.46
49	-	-	PI312929	0.45
50	-	-	PI313048	0.44
51	-	-	PI312921	0.44
52	-	-	PI312967	0.44
53	-	-	PI312928	0.43
54	-	-	PI273780	0.43

	Tekirdag-23 environment		Salina-24 environment	
Rank	**Ranking based on traits**	**Score of traits**	**Ranking based on traits**	**Score of traits**
55	-	-	PI273790	0.43
56	-	-	PI319713	0.43
57	-	-	PI338651	0.42
58	-	-	Delaney	0.42
59	-	-	PI312969	0.42
60	-	-	Shoshone	0.42
61	-	-	PI372802	0.41
62	-	-	PI312992	0.41
63	-	-	PI312965	0.41
64	-	-	PI312920	0.40
65	-	-	PI302939	0.36
66	-	-	PI312994	0.35
67	-	-	Eski	0.35
68	-	-	PI313065	0.35
69	-	-	PI313047	0.33
70	-	-	Rocky Mountain Remont	0.33
71	-	-	PI313063	0.32
72	-	-	PI311468	0.31
73	-	-	PI577666	0.29
74	-	-	PI494667	0.21
75	-	-	PI494668	0.19
76	-	-	PI302936	0.07

In the TOPSIS analysis for all measured agronomic traits, the genotypes prominent in Tekirdağ-23 and Salina-24 environments appear to be quite different from each other. The high rankings of different genotypes between the two environments indicate a strong environment × genotype interaction, indicating that the adaptability of genotypes may vary depending on environmental factors. Furthermore, the high scores of cultivars and breeding populations prominent in Tekirdağ-23 environment are immensely encouraging for breeding materials evaluated for multiple traits. This finding demonstrates that current breeding programs are successful in simultaneously improving targeted traits. Among the commercial varieties we included in our study in the Tekirdağ-23 environment, Lütfübey and Özerbey, ranked among the top in terms of performance in the thesis, were also identified as superior genotypes in a study conducted in Southeastern Anatolian climate and soil conditions due to their high forage yield and high-quality performance (Sayar et al., 2022). In general, it clearly demonstrates the necessity of

testing genotypes that are prominent in different environments in multiple locations and of developing environment-specific breeding strategies.

4.1.4. Heritability of traits

H^2 broad-sense heritability analysis was conducted for each trait in both environments (Table 4.2 and Table 4.3). Heritability analysis of sainfoin accessions was performed at the population level in the Tekirdağ-23 environment, and the agronomic traits StL (0.703) and InfL (0.715) showed high heritability. The traits FoSDM (0.635), StN1 (0.527), and RaL (0.558) showed medium to high heritability. The traits RaN (0.415) and SdN (0.369) showed medium to low heritability, indicating a significant environmental effect. CanHt, StTh, and StN2 had values less than 0.3, indicating that selection may have failed because these traits are largely due to environmental effects (Table 4.2).

In the Salina-24 environment, heritability analysis of sainfoin accessions was performed at the individual level (Table 4.3). In the H^2 calculation for agronomic traits, Fr1KK (0.675) and CanHt (0.604) exhibited high heritability, indicating strong genetic control of these traits and making them reliable for selection. StL (0.466), FoDM (0.390), StN (0.349), PlStEr (0.488), and PlRGR (0.435) traits were found to be of moderate heritability. StTh, RaN, and FrM showed low heritability values, and their H^2 value were less than that of an H^2 heritability of 0.3 (Table 4.3).

The broad-sense heritability (H^2) of agronomic traits studied in the Tekirdağ-23 and Salina-24 environments showed significant differences among the characteristics. Stem length ($H^2 = 0.70$) and inflorescence length ($H^2 = 0.71$) in the Tekirdağ-23 environment showed high heritability, indicating that these traits are under substantial genetic control and can be directly used in selection programs. Forage subsample dry matter ($H^2 = 0.63$) and Raceme length ($H^2 = 0.55$), which exhibited medium to high heritability, also showed potential for genetic improvement, although environmental effects partially influenced their expression. Legume fruit 1000-kernel weight and canopy height traits with high H^2 values in the Salina-24 environment are genetically controlled traits and can be directly used for selection in breeding programs. Traits with moderate heritability (Stem length, Plant stem erectness, Plant regrowth rating, Forage dry

matter), on the other hand, have the potential to contribute to genetic improvement even when influenced by environmental factors.

4.1.5. Accession and Species Clustering

The Neighbor-Joining tree revealed that populations were primarily located on branches close to each other, although some intraspecific clustering was also present. Accessions of *Onobrychis transcaucasica* and *Onobrychis viciifolia* tended to cluster into subclusters, despite being distributed on different branches. Cultivar/breeding populations were located on the same branches as interspecific examples (Figure 4.3).

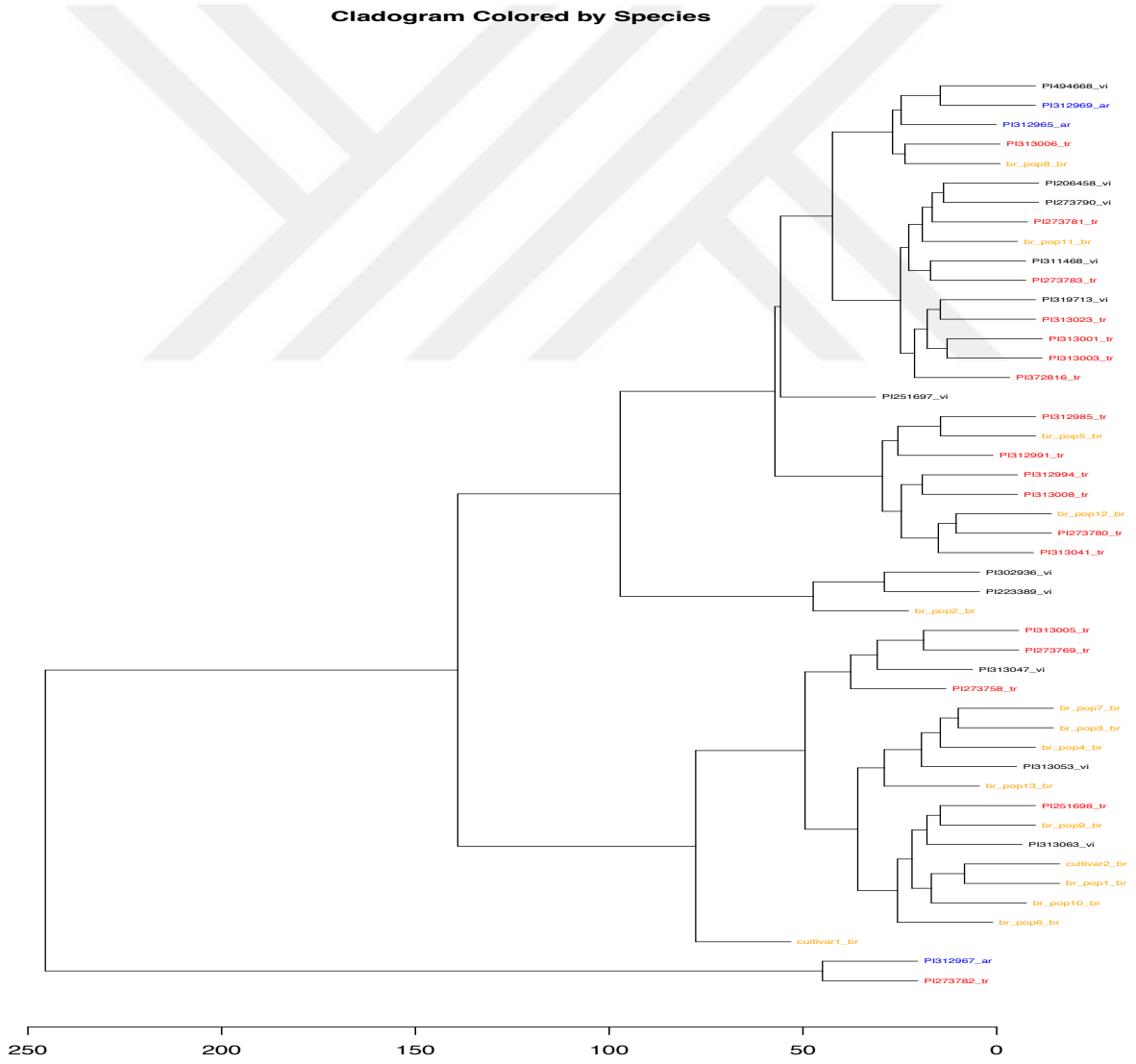


Figure 4.3. Clustering of dendrogram of sainfoin accessions/breeding populations/ cultivars in Tekirdağ-23 environment for agronomic traits

Principal Component Analysis (PCA) based on agronomic traits obtained around Tekirdağ-23 environment showed that most of the variation among the accessions was explained by PC1 (36.4%) and PC2 (21.6%) (approximately total variance = 58%). When the PCA distribution is examined, accessions belonging to *O. transcaucasica* (red) and *O. viciifolia* (black) species largely overlapped. In contrast, *O. arenaria*, which has a relatively limited number of accessions, showed partial separation. Breeding populations/cultivars (purple), although located in the central region with the species, did not exhibit independent clustering (Figure 4.4).

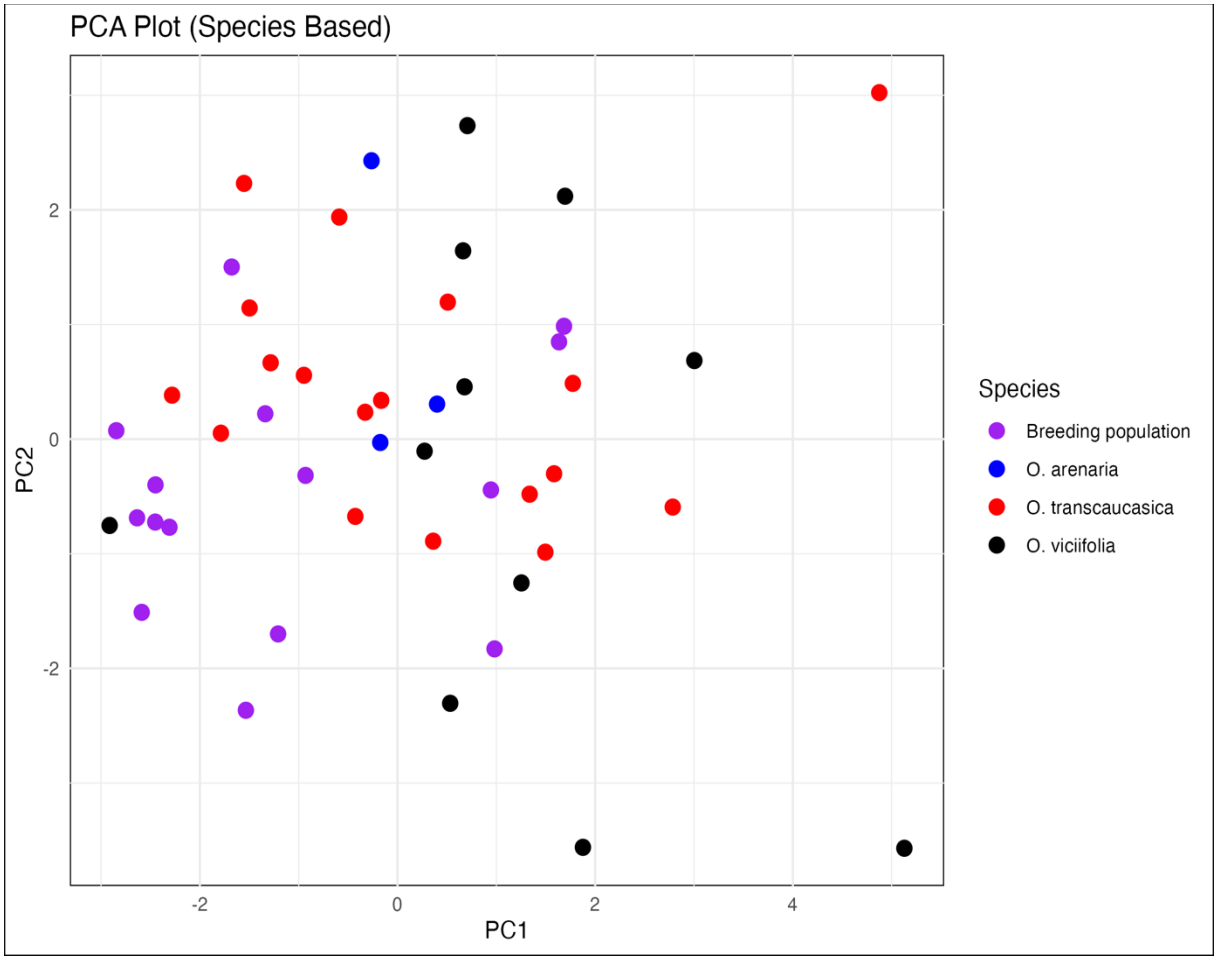


Figure 4.4. PCA plot of 47 sainfoin accessions/breeding populations/cultivars in the Tekirdağ-23 environment for agronomic traits

When PCA and NJ analyses on accessions in the Tekirdağ-23 environment revealed that agronomic traits in sainfoin accessions did not form clear boundaries between species, species did not create distinct clusters on the PCA plane, suggesting that agronomic traits can vary under the influence of environmental conditions and that phenotypic overlap between species

is widespread. Although a tendency for partial clustering between species was observed in the NJ tree, the branches generally exhibited a mixed distribution. The presence of cultivars/breeding populations on the same branches as different species suggests that the genetic base of the breeding materials is broad and complex, incorporating phenotypic diversity from other species.

The NJ tree, constructed using the Euclidean distance matrix, exhibited a mixed branching structure. Although some intraspecific clustering was observed, the branches mostly contained accessions from different species. The cultivar was mainly located on close branches with accessions from *O. viciifolia* and *O. transcaucasica* in Salina-24 environment (Figure 4.5).

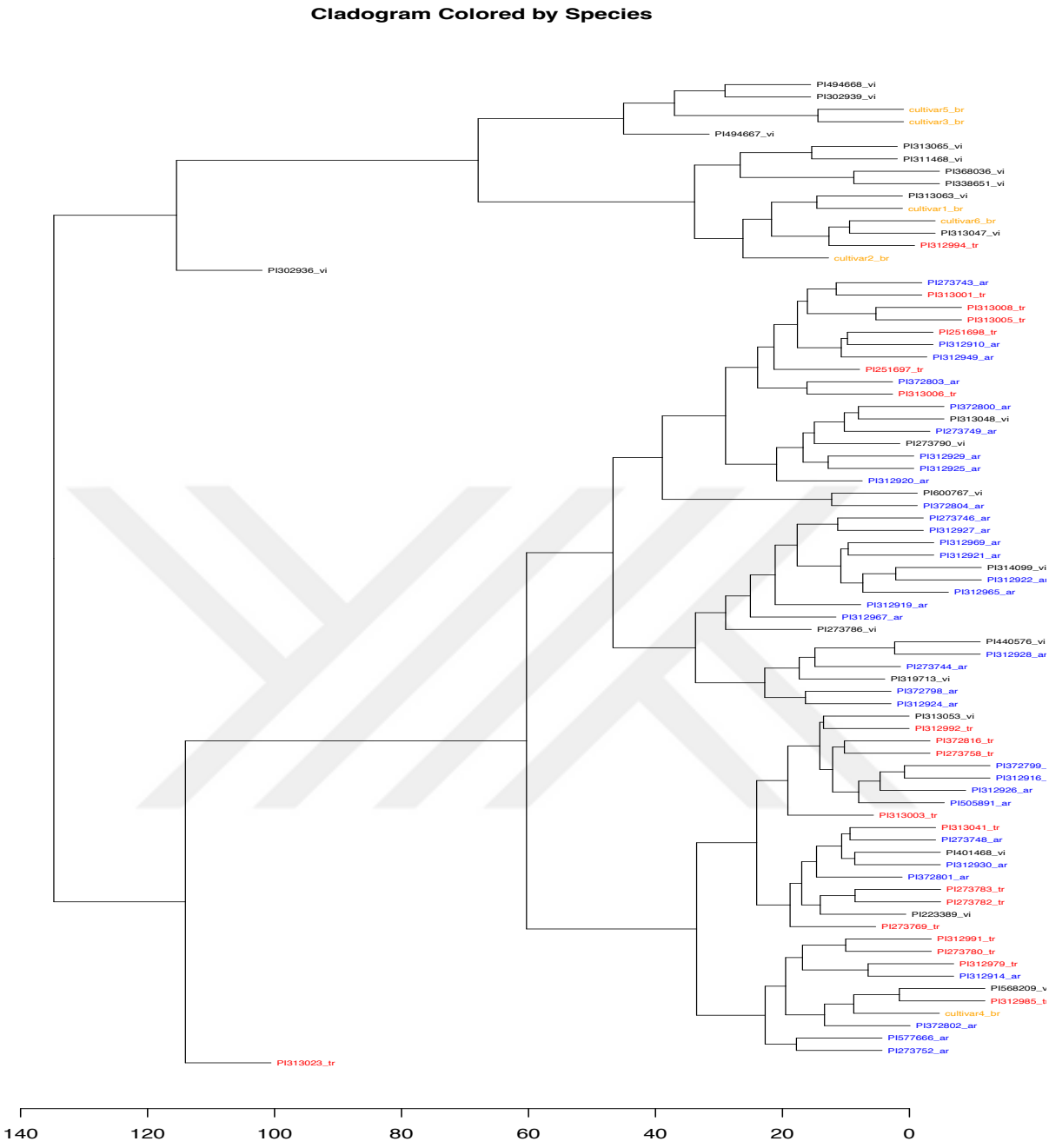


Figure 4.5. Clustering of dendrogram of sainfoin accessions/cultivars in the Salina-24 environment for agronomic traits

In the Principal Component Analysis obtained using the distance matrix obtained from agronomic traits in Salina-24 environment, PC1 explained 56.5% of the variance, while PC2 explained 14.1% of the variance (approximately total variance = 70%). In the PCA distribution, it was observed that accessions belonging to *O. arenaria* (blue) and *O. viciifolia* (black) partially overlapped. In contrast, accessions belonging to *O. transcaucasica* (red) were more

widespread in the center. The cultivars/breeding population (purple) were nested with accessions from these three species and did not differentiate (Figure 4.6).

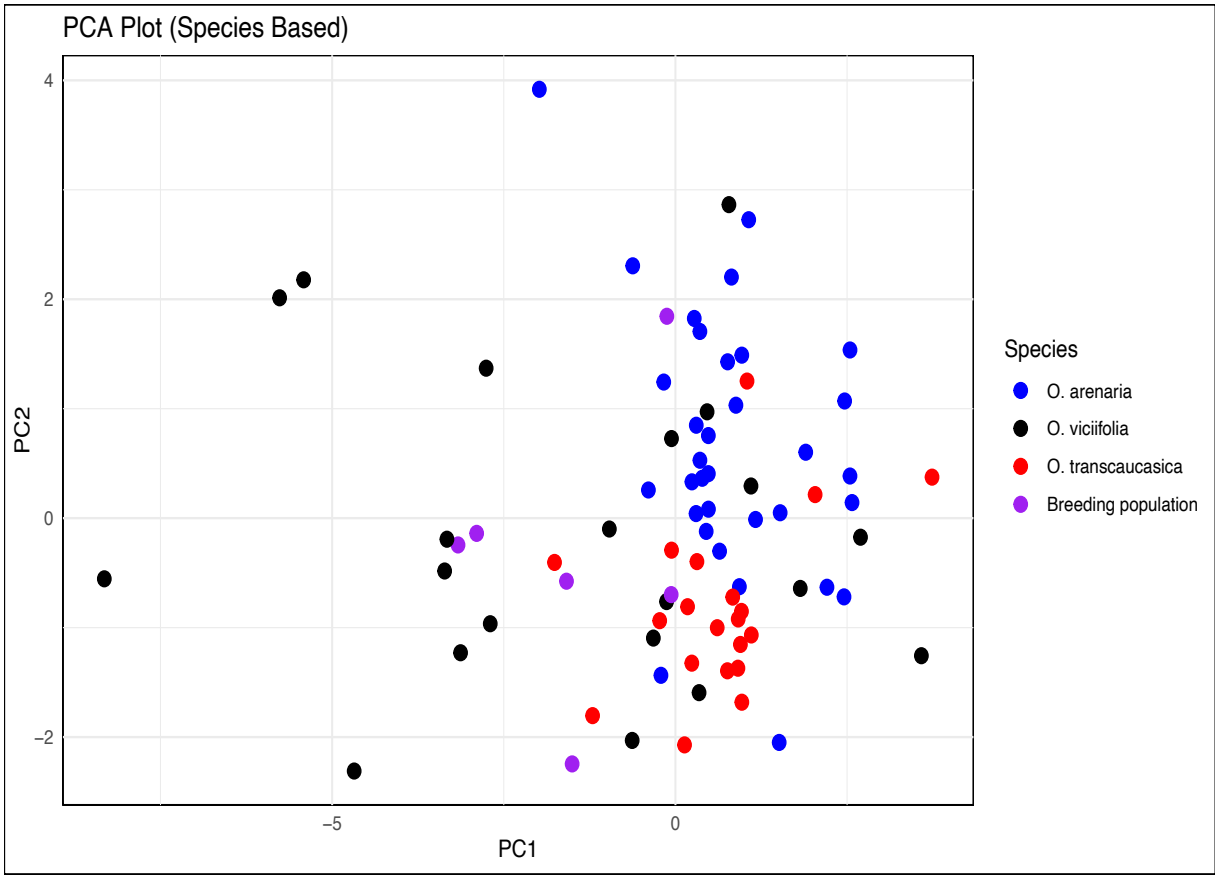


Figure 4.6. PCA plot of 80 sainfoin accessions/cultivars in the Salina-24 environment for agronomic traits

PCA and Neighbor Joining tree analyses on accessions in the Salina-24 environment revealed that agronomic traits in these accessions did not form clear boundaries between species, like those in the Tekirdağ-23 environment. The phenotypic overlap among species in the NJ tree highlights the role of environmental influences and phenotypic plasticity. Furthermore, the interbreeding of cultivars with different species indicates a high level of genetic mixing in modern breeding materials, which is reflected in the diversity of agronomic traits. In the PCA plot, PC1 had a high variance explanation rate (56.5%), indicating that a significant portion of the variation in agronomic traits was concentrated on a single axis. Although PC2 had a lower variance explanation rate, it enabled the separation of some genotypes. In general, species did not show any separation in the PCA plane, indicating that agronomic traits did not draw decisive boundaries between species.

A study was conducted to examine the diversity and relationship between morphological and genetic traits in 12 sainfoin accessions of *Onobrychis sativa* by (Mohajer et al., 2013). They found significant genetic variation among the populations. Significant differences were found in all traits between years. Their PCA analysis revealed that the first four components explained 88 percent of the total variance. Ward cluster analysis observed that the 12 populations were divided into three groups (Mohajer et al., 2013). In the PCA analysis conducted for the sainfoin accessions around Tekirdağ-23, the first two components explained approximately 58% of the total variance. For the accessions around Salina-24, the total variance of the first two components was approximately 70%, which is similar to the present study. Cluster analyses, including three different species in this study, did not reveal clear groupings in either environment.

The relationship between Forage Dry Matter (FoDM), selected for forage yield, and Legume Fruit Mass (FrM), selected for seed yield, was evaluated across six cultivars and 13 top-ranked accessions (Figure 4.7). A significant positive relationship was observed between both traits. Delaney and Eski cultivars showed low values for both forage and seed yield. AAC Mountainview, Rocky Mountain Remont, and Shoshone cultivars performed moderately. Renumex and PI 313048 accession approached higher values. The highest performance was observed in accessions PI 312924, PI 273744, and PI 372798, which exhibited superior performance in both forage yield and seed yield. In general, a significant portion of the accessions showed higher performance than the cultivars (Figure 4.7).

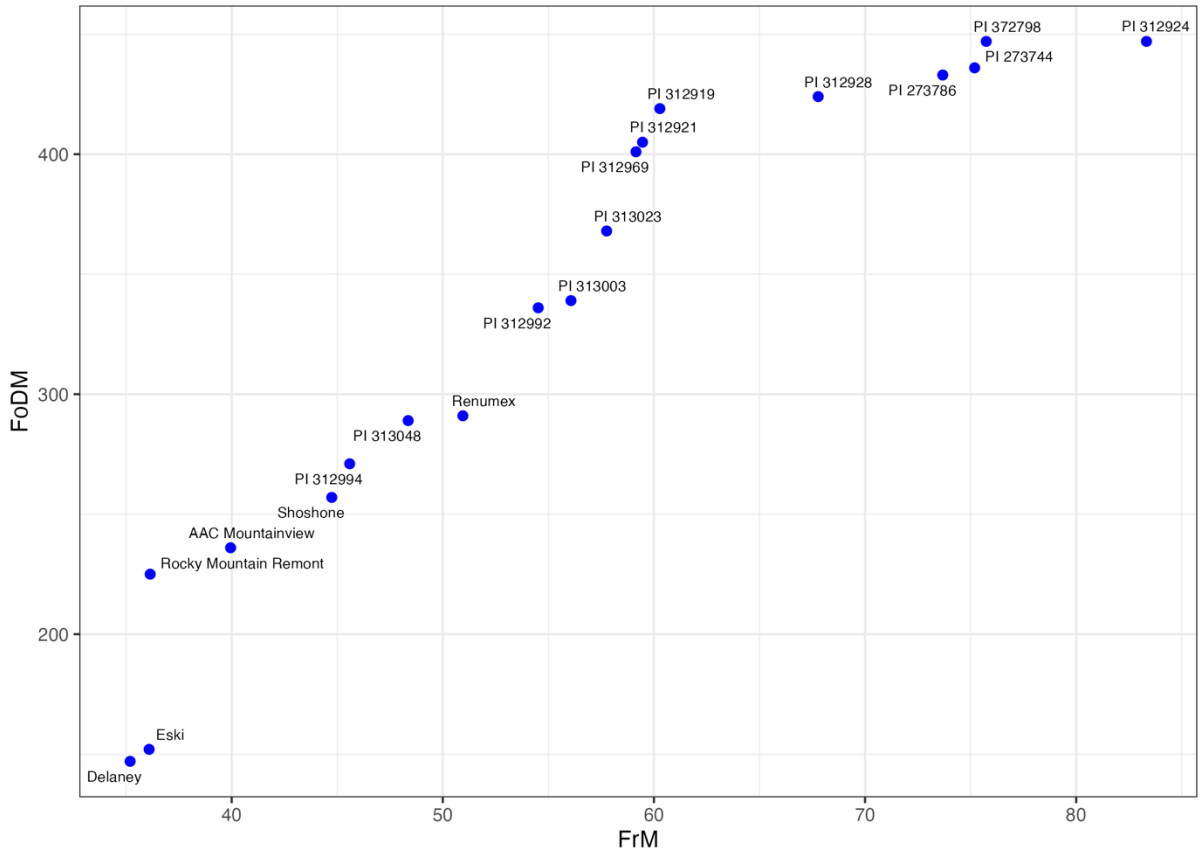


Figure 4.7. Performance comparison of 6 cultivars and 13 accessions in terms of two selected traits related to forage yield and seed yield

This study demonstrates that both forage and seed yields of sainfoin can be increased. It was determined that accessions with high forage yield also have the potential for high seed yields. This suggests that the same materials can be used for both biomass and seed production in the sainfoin breeding program. Furthermore, the generally lower performance of the cultivars suggests that attention may have been paid to yield stability in the improved varieties. However, accessions such as PI 312994 and PI 273744, which are already sources of genetic variation, should be evaluated because they offer significant potential in terms of forage and seed yield.

Given the seed yield results and the overall goal to breed sainfoin as perennial pulse crop, we could conclude that the selection for seed yield could be speeded by the integration of high yielding material. Particularly PI 312919 and PI 313003 could possess favorable alleles for seed yield.

4.2. Phenomic Analyses of the Sainfoin Diversity Panel

4.2.1. Effects of Environment

The results of analysis of variance (ANOVA) performed to investigate predicted seed morphological traits showed significant differences among species and accessions for most traits. Especially, FrPA, SdPA, FrW, SdW, and FrL showed significant differences ($p < 0.001$). However, regarding fruit and seed aspect ratios, FrAR and SdAR traits did not show significant differences among species and accessions (Table 4.10). Variance analysis data obtained from estimated seed data indicate significant genetic diversity among sainfoin species.

Table 4.10 F-values along with the significance from the analysis of variance (ANOVA) and H^2 heritability of seed-related traits in the Salina-24 environment

Source	Species	Accession	H^2
FrPA†	6.73***	6.02***	0.51
FrW	5.66**	5.82***	0.51
FrL	6.63***	5.23***	0.48
FrAR	2.55	0.93	0.00
FrEx	3.30*	2.32***	0.32
SdPA	6.72***	5.08***	0.64
SdW	6.13***	6.35***	0.70
SdL	5.87***	3.82***	0.21
SdAR	0.47	1.39	0.02
SdEx	1.15	2.74***	0.30

†FrPA: Legume fruit projected area; FrW: Legume fruit width; FrL: Legume fruit; FrAR: Legume fruit Aspect ratio; FrEx Legume fruit extent; SdPA: Seed projected area; SdW: Seed width; SdL: Seed length; SdAR: Seed Aspect ratio; SdEx: Seed extent. Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

4.2.2. Trait means and Correlations among traits

Observed variation had the significant among the sainfoin species and cultivars about seed morphological traits under Salina-24 environment (Table 4.11). The analysis of variance using by Tukey's multiple comparison test found the significant differences among sainfoin species and cultivars for many seed-related traits in Salina-24 environment. For legume fruit projected area (FrPA), where cultivars had the largest average projected area, when it's compared *to O. arenaria, O. transcaucasica* and *O. viciifolia.* In terms of fruit appearances, the length (FrL) was highest in cultivars, whereas fruit width (FrW) remained almost consistent across species, ranging between 4.2- and 4.47-mm. Legume fruit aspect ratio (FrAR) and extent (FrEx) displayed relatively minimal differences, in terms of conserved shape characteristics (Table

4.11). Seed traits followed a similar pattern. Seed projected area (SdPA) was highest in *O. transcaucasica*, and cultivars followed closely. Seed length (SdL) and width (SdW) performed medium variation, with the longest seeds observed in *O. transcaucasica*, and the widest seeds in cultivars. Seed aspect ratio (SdAR) remained almost the same across species, suggesting a stable seed shape rather than size differences (Table 4.11). In short, in terms of legume fruit morphology, FrPA, FrW, and FrL traits were found at widely higher values in cultivars and *O. transacaucasica*, while they were at the lowest value in *O. arenaria*. FrAR, FrEx, SdAR, and SdEx did not show any significant differences among the species. In terms of seed morphology, SdPA, SdW, and SdL traits had the highest values in *O. transcaucasica*, while *O. arenaria* had the lowest values (Table 4.11).

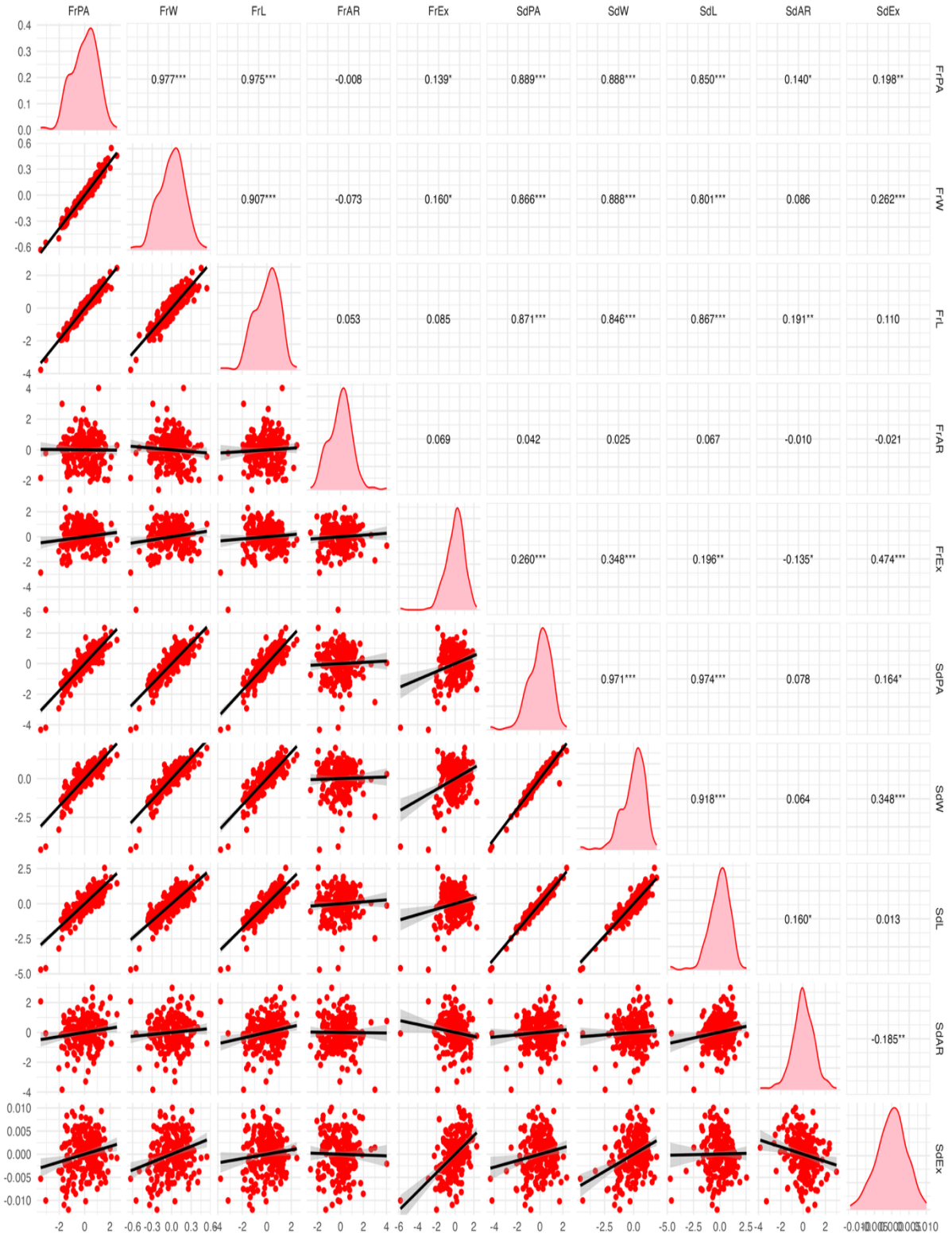
Table 4.11. Descriptive statistics (mean, minimum, and maximum) of estimated marginal means (emmeans) with standard errors of agronomic performance of estimated seed traits based on sainfoin species and showing their statistical differences in the Salina-24 environment

Species	Unit	*O. arenaria*	Cultivars	*O. transcaucasica*	*O. viciifolia*
FrPA†	mm^2	19±0.16^{a} (17.2-21.4)	21.3±0.37^{c} (19.4-22.9)	20.9±0.21^{c} (18.5-22.1)	19.8±0.22^{b} (14-22.6)
FrW	mm	4.20±0.01^{a} (4.01-4.4)	4.47±0.04^{c} (4.29-4.62)	4.40±0.02^{c} (4.21-4.53)	4.31±0.02^{b} (3.71-4.71)
FrL	mm	5.81±0.02^{a} (5.5-6.25)	6.13±0.06^{b} (5.79-6.35)	6.09±0.03^{b} (5.62-6.3)	5.87±0.03^{a} (4.93-6.27)
FrAR	ratio	1.01±0.00^{a} (1.01-1.03)	1.01±0.00^{a} (1.01-1.03)	1.01±0.00^{a} (1-1.02)	1.01±0.00^{a} (1.01-1.02)
FrEx	ratio	0.68±0.00^{a} (0.673-0.706)	0.69±0.00^{a} (0.68-0.69)	0.68±0.00^{a} (0.67-0.69)	0.68±0.00^{a} (0.64-0.70)
SdPA	mm^2	7.77±0.08^{a} (6.69-9.02)	8.52±0.18^{bc} (8.07-8.97)	8.61±0.10^{c} (7.96-9.08)	8.17±0.11^{b} (4.65-9.36)
SdW	mm	2.70±0.01^{a} (2.5-2.91)	2.87±0.03^{c} (2.81-2.94)	2.86±0.01^{c} (2.77-2.93)	2.76±0.01^{b} (2.1-2.97)
SdL	mm	3.66±0.03^{a} (3.41-3.96)	3.78±0.06^{ab} (3.65-3.88)	3.83±0.04^{b} (3.6-3.96)	3.78±0.04^{b} (2.83-4.73)
SdAR	ratio	1.00±0.00^{a} (1-1.02)	1.01±0.00^{a} (1.00-1.01)	1.00±0.00^{a} (0.99-1.02)	1.01±0.00^{a} (0.98-1.02)
SdEx	ratio	0.74±0.00^{a} (0.73-0.74)	0.74±0.00^{a} (0.74-0.74)	0.74±0.00^{a} (0.74-0.74)	0.74±0.00^{a} (0.73-0.74)

† FrPA: Legume fruit projected area; FrW: Legume fruit width; FrL: Legume fruit; FrAR: Legume fruit Aspect Ratio; FrEx Legume fruit extent; SdPA: Seed projected area; SdW: Seed width; SdL: Seed length; SdAR: Seed Aspect Ratio; SdEx: Seed extent

Cultivars and *O. transcaucasica* have larger and higher mass in terms of both fruit and seed morphology. In contrast, *O. arenaria* has smaller and more compact seeds and fruits in the Salina-24 environment. These differences highlight the importance of wild species as genetic resources. Furthermore, differences in cultivars suggest that previous selection was primarily focused on yield during the breeding process. In addition, the conservation of seed shape and proportions among species suggests that these morphological seed characteristics look like more conserved.

Then, correlation analysis was performed to observe the relationships among the predicted seed morphological traits (Figure 4.8). Highly positive correlations were observed among FrPA, FrW, and FrL ($r > 0.90$, $p < 0.001$). SdPA, SdW, and SdL also exhibited strong positive correlations with each other ($r > 0.90$, $p < 0.001$). Insignificant or weak correlations were found between FrAR and SdAR traits and other traits. In addition, SdAR showed negative correlations with some traits, such as SdW ($r = -0.18$, $p < 0.05$) (Figure 4.8).



FrPA: Legume fruit projected area; FrW: Legume fruit width; FrL: Legume fruit; FrAR: Legume fruit Aspect Ratio; FrEx Legume fruit extent; SdPA: Seed projected area; SdW: Seed width; SdL: Seed length; SdAR: Seed Aspect Ratio; SdEx: Seed extent

Figure 4.8. Pearson correlation plot of estimated seed traits in 80 sainfoin accessions in the Salina-24 environment

Because the fundamental dimension measurements of seed morphology—length, width, and area—vary significantly, these traits are thought to be genetically and developmentally interconnected, with their developmental processes controlled by common genetic factors. The high positive correlation between traits related to fruit and seed size suggests that these traits can be considered together in selection programs and that even selecting for one trait can indirectly influence the others.

4.2.3. Phenotypic Variation among Sainfoin Accessions

Ranking analysis was performed for all estimated seed traits obtained by image processing for the Salina-24 environment, and the performance of 80 accessions was ranked. However, an interpretation was made for the top 10 accessions for the Salina-24 environment (Table 4.12). In the estimated seed traits, the accessions numbered PI 313003 and PI 273786 showed the best performance, while none of the cultivars were in the top 10.

Table 4.12. Ranking of 80 Sainfoin Accessions for Estimated Seed Traits in the Salina-24 environment

Rank	Ranking based on seed traits	Score of traits
1	PI313003	0.87
2	PI273786	0.85
3	PI312919	0.83
4	PI313006	0.79
5	PI372816	0.77
6	PI273758	0.76
7	PI251697	0.74
8	PI312927	0.73
9	PI313023	0.73
10	PI312992	0.72
11	PI312979	0.71
12	PI273769	0.71
13	PI313063	0.70
14	PI372803	0.69
15	PI312965	0.69
16	PI313008	0.69
17	Renumex	0.68
18	PI312921	0.68
19	PI313005	0.68
20	PI312910	0.66
21	PI313053	0.66

Rank	Ranking based on seed traits	Score of traits
22	PI313048	0.65
23	PI313001	0.65
24	PI372798	0.65
25	PI312922	0.65
26	PI273782	0.65
27	PI312914	0.65
28	PI312969	0.65
29	PI312991	0.64
30	PI273746	0.64
31	PI273780	0.64
32	PI505891	0.64
33	PI312967	0.64
34	AAC Mountainview	0.63
35	PI312994	0.63
36	PI312985	0.62
37	PI273743	0.62
38	PI312926	0.62
39	PI312916	0.61
40	PI312924	0.61
41	PI313047	0.60
42	PI312949	0.60
43	PI251698	0.60
44	Shoshone	0.60
45	PI273744	0.58
46	PI273783	0.58
47	Rocky Mountain Remont	0.57
48	PI372800	0.57
49	PI313041	0.56
50	Delaney	0.55
51	PI372804	0.55
52	PI273748	0.54
53	PI372801	0.54
54	PI273749	0.53
55	Eski	0.52
56	PI312925	0.52
57	PI372802	0.51
58	PI401468	0.51
59	PI311468	0.49
60	PI223389	0.48
61	PI302936	0.48
62	PI368036	0.47

Rank	Ranking based on seed traits	Score of traits
63	PI313065	0.47
64	PI302939	0.46
65	PI312930	0.45
66	PI273790	0.45
67	PI273752	0.44
68	PI312928	0.43
69	PI338651	0.43
70	PI312929	0.42
71	PI600767	0.38
72	PI319713	0.32
73	PI312920	0.31
74	PI494668	0.29
75	PI577666	0.24
76	PI494667	0.23

When the performance ranking of predicted seed traits across 80 sainfoin accessions was examined, the accessions numbered PI 313003 and PI 273786 showed the best performance, while no sainfoin cultivars were ranked in the top 10. Therefore, it is necessary to emphasize that these sainfoin accessions, which have been insufficiently studied or have received very little attention, are valuable genetic resources. It is recommended that they be included in current breeding studies.

4.2.4. Heritability of the traits

When broad-sense heritability (H^2) estimates were examined, SdPA ($H^2 = 0.64$), SdW ($H^2 = 0.70$), and FrPA ($H^2 = 0.51$) were particularly notable for their high heritability values. However, SdL ($H^2 = 0.21$), FrAR ($H^2 = 0.005$), and SdAR ($H^2 = 0.02$) showed low heritability values. These findings suggest that traits such as width and area are under genetic control, while environmental effects are more dominant in ratio-based traits such as FrAR and SdAR (Table 4.10). Broad-sense heritability analysis reveals high heritability values for SdW, SdPA, and FrPA, suggesting that selection can be made more effectively in breeding programs due to their strong genetic control and high selection potential.

4.2.5. Accession and Species Clustering

A cladogram, colored by species, was constructed using the Euclidean distance matrix for the estimated seed traits in the Salina-24 environment. This cladogram plot shows that species do not completely segregate sainfoin genotypes, but they do observe a tendency to cluster to a

certain extent (Figure 4.9). While accessions from *O. viciifolia* (blue) and *O. arenaria* (orange) mainly were clustered within themselves, some accessions were also observed nested within other species. Accessions from *O. transcaucasica* (red) exhibited a more scattered distribution and were located on close branches with the other two species. The cultivar/breeding group (black) clustered together with accessions from different species and did not form a separate group. The overall branching structure suggests that there are no distinct species boundaries between populations, although a tendency for partial clustering at the species level is evident (Figure 4.9).

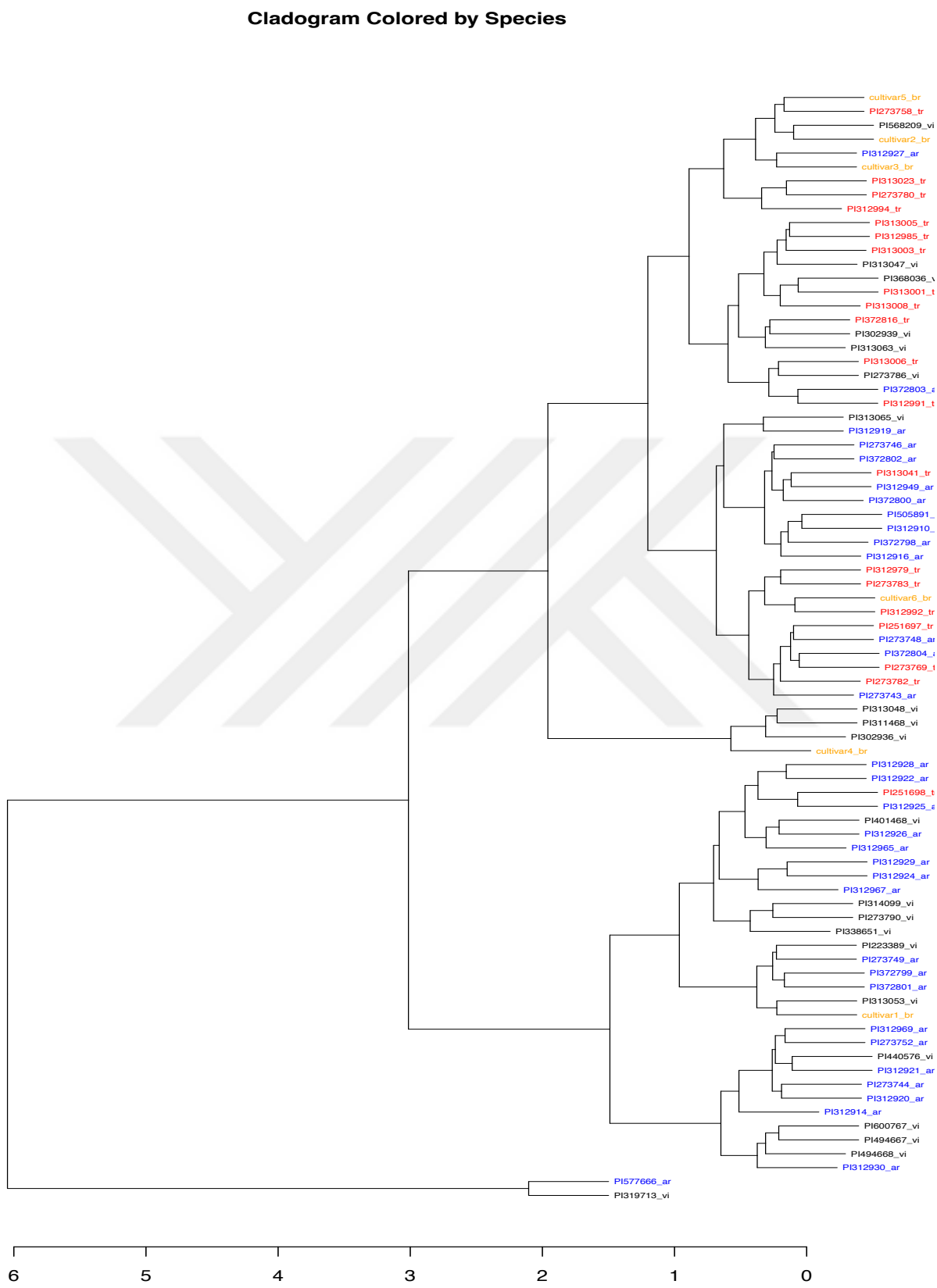


Figure 4.9. Clustering of dendrogram of sainfoin accessions/cultivars in the Salina-24 environment for estimated seed related traits

The branching structure in the cladogram indicates significant genetic overlap and mixing among species. Despite the tendency for intraspecific clustering, particularly observed in the accessions of *O. viciifolia* and *O. arenaria*, the presence of interbreeding accessions with other species suggests that genetic permeability between species may be high. The distribution of *O. transcaucasica* across different clades suggests high genetic heterogeneity for this species or the possibility of historical introgression. The scattered distribution of cultivar/breeding populations within the same clades, along with accessions from different species, suggests that allelic diversity from more than one species is present in the breeding material.

Principal component analysis (PCA) was performed to determine the morphological variation of four sainfoin groups by species for estimated seed-related traits in the Salina-24 environment (Figure 4.10). The first two components (PC1 & PC2) explained 75% of the total variance (PC1 = 59.5%, PC2 = 15.5%). PC1 explained most of the variation of fruit/seed morphology in the populations, while PC2 explained a smaller amount of additional variation. According to the distribution in the PCA plot, most of the accessions belonging to *O. arenaria* (blue) and *O. viciifolia* (black) exhibited an overlapping pattern. Accessions belonging to *O. transcaucasica* (red) were concentrated in the central region and were significantly mixed with other species. Cultivars/Breeding populations (purple) spread with species and didn't form independent clusters/grouping design (Figure 4.10).

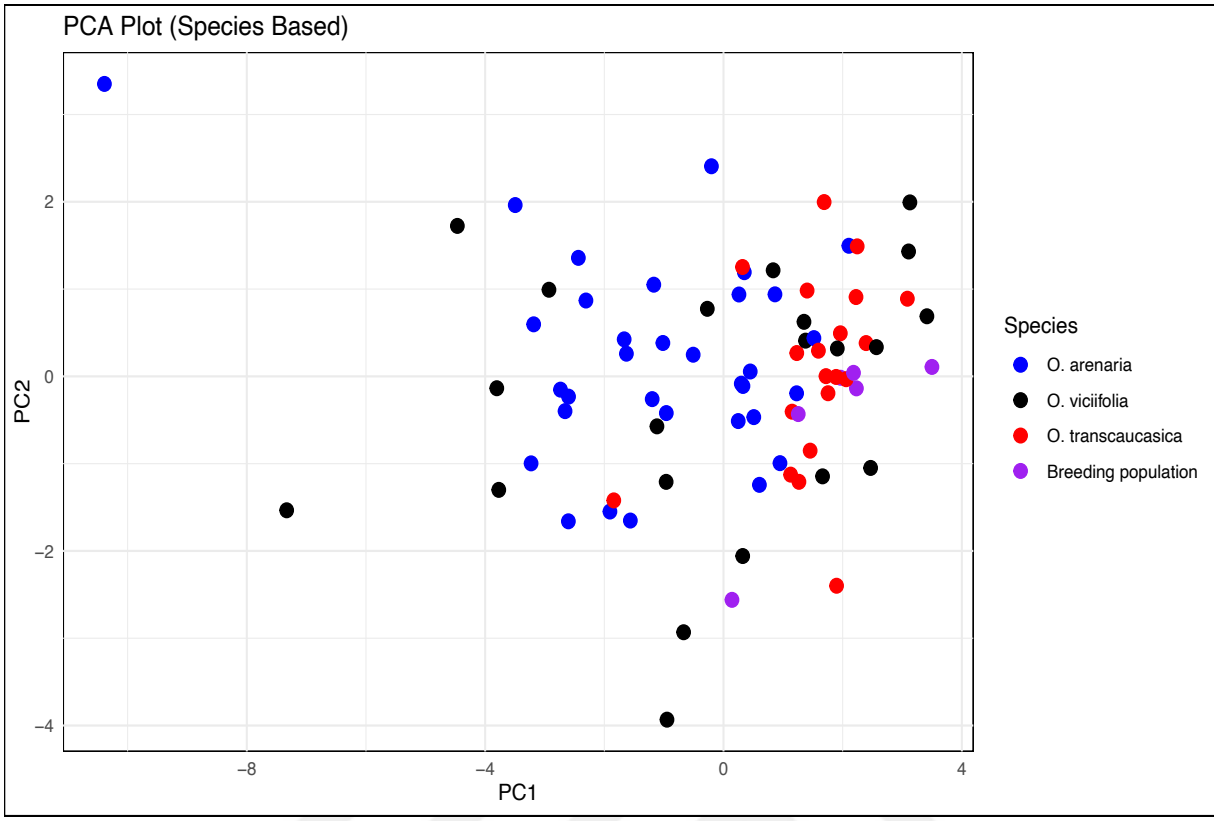


Figure 4.10. Principal Component Analysis (PCA) plot of 80 sainfoin accessions/cultivars in the Salina-24 environment for estimated seed related traits

PCA results indicate that seed morphology does not clearly diverge the sainfoin species, and there is a high level of phenotypic overlap. The significant intermixing of sainfoin species in the PCA plane suggests that seed morphological characteristics are unlikely to be substantial in distinguishing between species. Furthermore, it is thought that overlap between species is widespread due to environmental influences or plasticity. In particular, the concentration of *O. transcaucasica* in the center of the PCA plane suggests that this species may overlap with the other two species (*O. viciifolia* and *O. arenaria*) in terms of seed morphology.

Collections of samples of *O. arenaria* and *O. transcaucasica*, originating from Kazakhstan, were selected in a study. The seeds were processed to obtain protein extractions, and bands were obtained by running SDS-PAGE gels. Then, these data were analyzed using hierarchical clustering. Although clustering analyses revealed clear distinctions between *O. arenaria* and *O. transcaucasica*, some samples were reported to have intermediate profiles, suggesting areas of genetic transition (Bulatova, Massonichich-Shotunova, Meiirman, & Mazkirat, 2015).

Processes of gene flow can explain this situation, variation based on common ancestry, or ecological adaptation.

In summary, both the seed storage protein-based study conducted in Kazakhstan and the phenotypic and phenomic analyses conducted in the Tekirdağ/Türkiye and Salina/Kansas, within the scope of this study, indicate that sainfoin possesses a rich and robust genetic diversity. They also highlight that sainfoin may be a genetically heterogeneous species complex.

4.3. Genomic Analyses of the Sainfoin Diversity Panel

4.3.1. Genotyping Results

Skim sequencing-based results were aligned for tetraploid phased genotyping was conducted using a house draft sainfoin genome, and a total of 46,613 SNP markers were generated. The markers were subsequently imputed, and the imputed genotypes were converted to allele frequencies.

4.3.2. Population Structure

Principal Component Analysis (PCA) was performed on a Euclidean distance matrix generated based on mean allele frequencies for 43 sainfoin populations to separate genetic groups. But it didn't reveal a significant separation between groups (Figure 4.11). PC1 (8.83%) and PC2 (7.06%) explained only a minimal portion of the total molecular variance, resulting in a widespread distribution of sainfoin populations across the plot. Populations belonging to *O. transcaucasica* and *O. viciifolia* overlapped with each other, while cultivar/breeding populations are partially overlapped with these two species. *O. arenaria*, being a limited population number, was partially isolated from the others but still did not form a clear clustering pattern. An overview of the PCA plot reveals that the species do not cluster separately and that there is extensive genetic overlap between the populations (Figure 4.11).

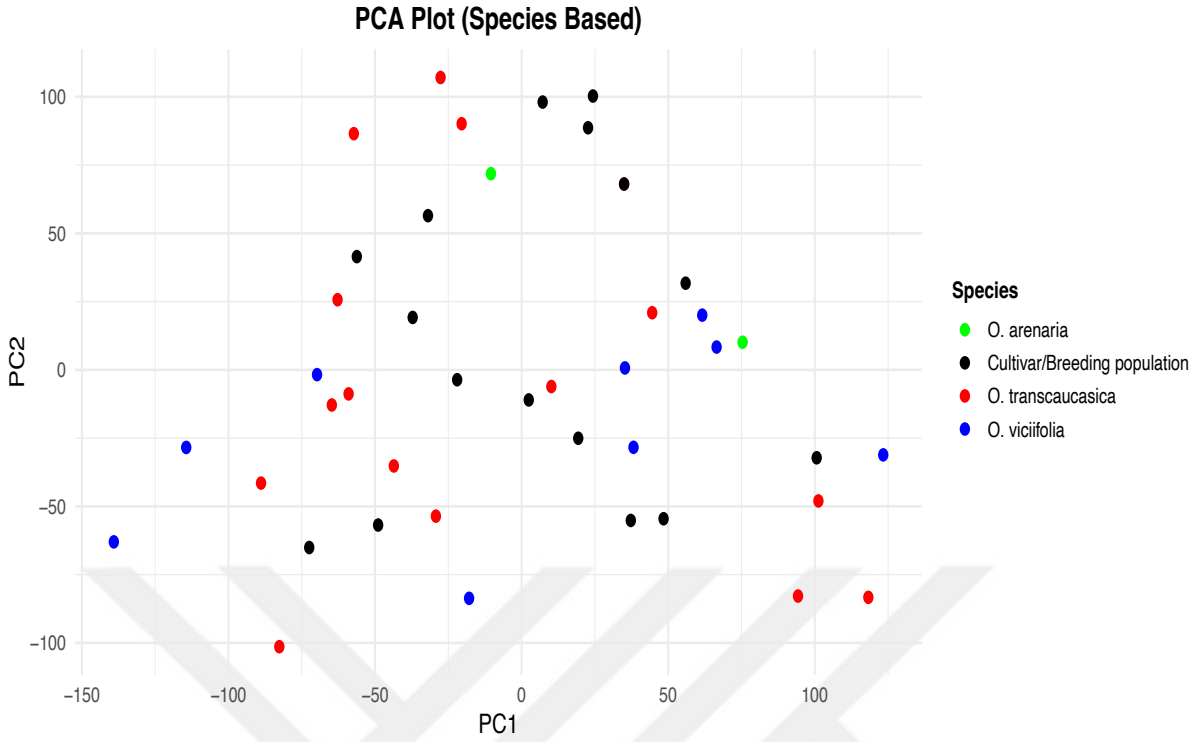


Figure 4.11. Principal Component Analysis (PCA) plot of 43 sainfoin populations based on sainfoin species in the Tekirdağ-23 environment

PCA results indicate that genetic differentiation among species is not well established. When populations are evaluated within taxonomic boundaries, they show no distinct structure. This suggests the possibility of significant gene flow and/or historical introgression among *O. transcaucasica, O. viciifolia*, and cultivars/breeding groups. The scattered distribution of *O. arenaria* populations suggests that they do not form a distinct cluster and that extensive hybridizations with other species during the seed bank seed increase may have contributed to allele sharing and ultimately lack of genetic distinction. In the present study approximately 47,000 SNP markers were obtained from 430 individuals of three sainfoin species, PCA analysis was performed using the Euclidean distance matrix. PC1 (8.83%) and PC2 (7.06%) explained only a small portion of the total variation in the PCA analysis. Kempf et al. (2016) conducted a study using 400 SSR primers in a plant population consisting of 32 individuals and 29 different plant populations of *O. viciifolia*. In the PCA analysis, PC1 (10.3%) and PC2 (4.9%) explained only a small portion of the total variance (Kempf et al., 2016). These percentages were similar to the PCA variance percentages in the present study.

4.3.3. Phylogenetic and Distance Analyses

A NJ tree was constructed using Euclidean distance metrics based on mean allele frequencies. No clear clustering was observed at the species level among the 43 sainfoin populations. *Onobrychis transcaucasica* and *Onobrychis viciifolia* populations were distributed across different branches of the phylogenetic tree and did not show any species-specific grouping. Similarly, *O. arenaria* populations were distributed across different branches of the phylogenetic tree. Cultivar/breeding populations were nested within genotypes of sainfoin species and showed close relationships with both *O. transcaucasica* and *O. viciifolia*. When the Neighbor-Joining tree was evaluated in general, the branching structure of the tree revealed that the populations contained a high level of genetic mixing and that there were no clear separations between species (Figure 4.12).

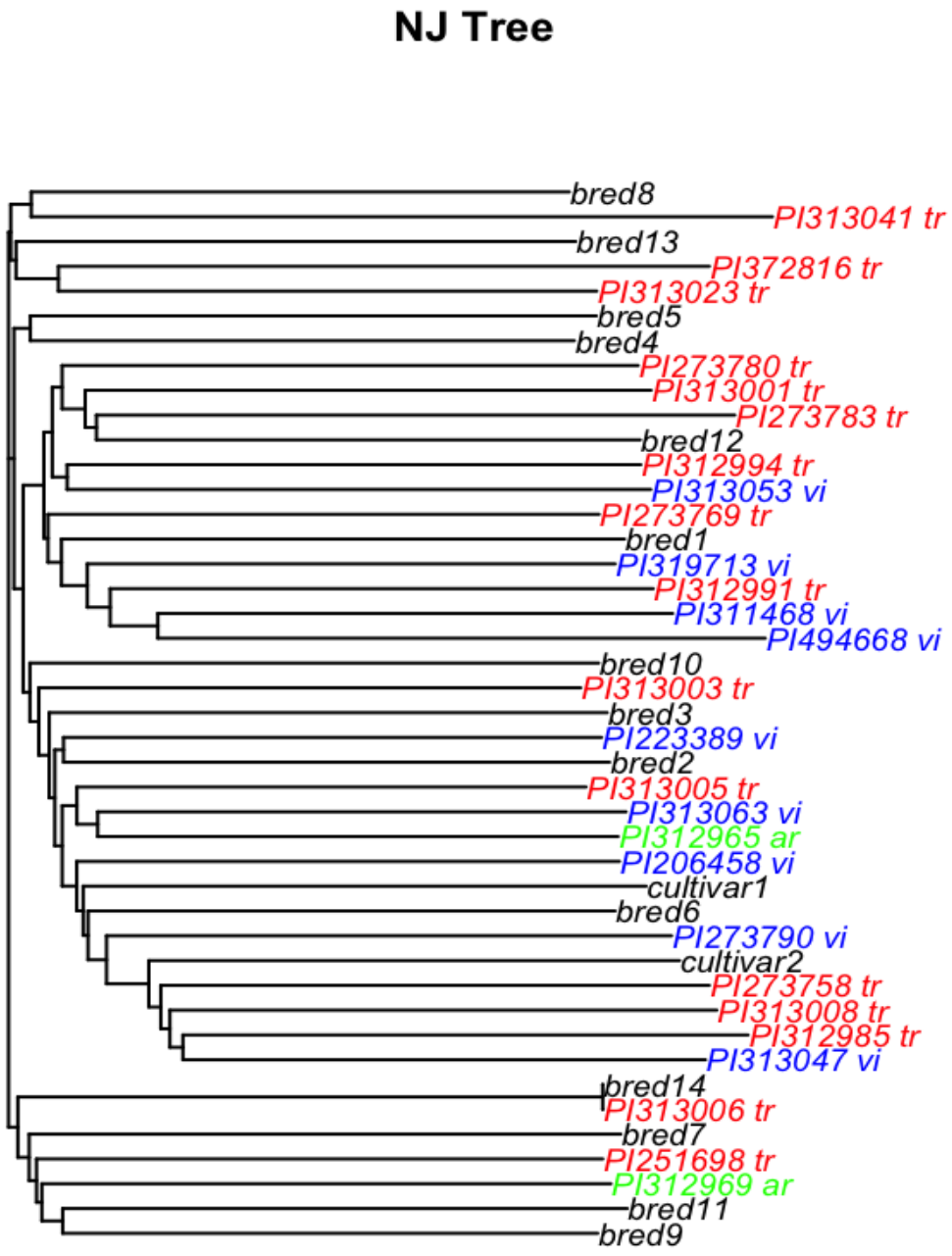


Figure 4.12. Neighbor-joining tree of 43 sainfoin accessions obtained from genotyping data based on sainfoin species and breeding populations in the Tekirdağ-23 environment

Neighbor-joining tree results showed that there was no clear separation between the sainfoin species. The lack of separation could be tested in different marker systems particularly more evolutionary conserved DNA markers such as non-nuclear variations.

4.3.4. Molecular Variance Analysis (AMOVA)

Molecular analysis of variance (AMOVA) was performed to calculate the percentage distribution of variance among sainfoin populations using allele frequency data (Table 4.13). The variance between groups was found to be -0.25. However, because biologically negative values are considered 0, the variance between groups cannot be explained. The within-population variance was 93.03%, explaining almost all the variance within the population. The p value was 0.566, making the between-group variance statistically insignificant. The φ-statistic between population groups yielded a negative value (-0.0026), but biologically it is considered 0. This indicates that there is no genetic differentiation between species (Table 4.13).

Table 4.13. Analysis of molecular variance (AMOVA) of 43 sainfoin accessions, including four sainfoin groups (*O. viciifolia, O. transcaucasica, O. arenaria*, cultivars/breeding populations) in the Tekirdag-23 environment

Source of variation	Df	Sum of squares (SS)	Mean Squares (MS)	Variance components (σ^2)	P-value
Among groups	3	271.95	90.65	-0.25	0.566
Within populations	39	3628.23	93.03	93.03	-
Total	42	3900.18	92.86	92.78	-

AMOVA results also revealed no significant genetic differences among *Onobrychis* species (*O. viciifolia, O. transcaucasica*, and *O. arenaria*) and cultivar/breeding populations and most of the genetic variation was explained within the population (93.03%) in populations of these three different species. Majority of genetic variation in RAPD markers in five populations of the *O. viciifolia* species was reported to reside within the population (89.03%) (H. Nosrati, Feizi, Tarrah, & Haghighi, 2012). The plant material used in the previous study was re-examined using the ISSR marker. As a result of the AMOVA analysis, the intra-population diversity (67.5%) was found to be relatively higher, and this was associated with high heterogeneity (Houshang Nosrati, Feizi, Latifian, & Haghighi, 2016). A study of 83 genotypes, mostly *O. viciifolia* but also including *O. arenaria*, *O. altissima*, and other *Onobrychis* species, was conducted using 10 nSSRs for diversity analysis. This study found rich diversity among genotypes and reported that each genotype may exhibit high heterogeneity (Sutcu, Bilgen, & Tuna, 2022). According to the AMOVA results of 31 species, including *Onobrychis viciifolia*, *Onobrychis transcaucasica*, and *Onobrychis arenaria*, 68.41% of the genetic variance was found to be within the species (Zarrabian, Dehkordi, Ehtemam, & Majidi, 2022).

In one study, genetic diversity and relationships were analyzed in 102 accessions belonging to 33 *Onobrychis* species using 22 ISSR primers. AMOVA and genetic variance analyses revealed that a significant portion of the variance was within species. It has been reported that some species, such as *O. altissima*, exchange genes with *O. viciifolia*, which may lead to interspecific confusion. Furthermore, it has been suggested that the genetic boundaries between these three species may not be clear, that interspecific gene flow may occur, and that intrapopulation genetic diversity may be dominant (Zarrabian & Majidi, 2015).

In another study, a genetic diversity and genetic relationship investigation was conducted in Türkiye using 23 landraces and one registered variety (Lütfübey) of *Onobrychis vicifolia* with five RAPD and four ISSR primers. The study reported a wide genetic diversity among the populations (Ertuş, 2021).

To sum up, previous sainfoin studies reported that genetic variation was largely within the population, with limited interspecific differences. The PCA, NJ, and AMOVA results obtained in this study supported the same conclusion, and it was observed that no clear genetic boundaries could be discerned among sainfoin species, and that the genetic variation was largely contained within population. All these results implied significant genetic admixture and gene flow among the sainfoin germplasm. Therefore, it appears that the cultivars and breeding populations harbor allelic diversity from multiple species.

In this study, the combined evaluation of genetic diversity observed at the molecular level with diversity based on morphological and agronomic traits related to forage and seed yield will contribute to a more effective evaluation of sainfoin species and their accessions in terms of breeding.

5. CONCLUSION

Significant natural variation has been observed in sainfoin germplasm evaluated in this study. The variation was observed among species (*Onobrychis viciifolia*, *Onobrychis transcaucasica*, and *Onobrychis arenaria*), cultivars, improved germplasm, and wild accessions. Both phenotypic assessments and phenomic analyses revealed a considerable variation in traits related to forage and seed yield. Correlation analyses disclosed significant positive correlations among several traits suggesting focusing easy to assess traits during selection. The analyses also revealed that fruits and seeds with significantly larger grain sizes and a wider length-to-width ratio. The sufficient variation in sainfoin seeds in this regard may allow them to be evaluated and grown as legumes in the future, and this variation could significantly increase seed yield. Results suggest that sainfoin can be grown both as a forage legume and as a perennial grain crop and there is a good natural variation to perform long term selection program to improve grain yield to a level that is economically viable for farmers.

The top three accessions ranked in terms of performance in the Salina-24 environment were accessions of *Onobrychis arenaria* (Kit) DC (PI 273749, PI 273748, PI 273752) originated from different areas in Ukraine. The fourth-ranked accession is *Onobrychis transcaucasica* Grossh. (PI 273758) originating in Azerbaijan with unknown improvement status. This accession also ranked second in terms of performance in Tekirdağ-23 environment. The first-ranked accession in the Tekirdağ-23 environment was PI 313047 (*Onobrychis viciifolia*) originating from Germany with unknown improvement status.

Some traits with high broad-sense heritability values were identified in both regions. These traits include stem length, inflorescence length, and forage subsample dry mass traits in the Tekirdağ-23 environment, and legume fruit 1000K weight and canopy height traits in the Salina-2024 environment. These traits with higher broad sense heritability (H^2) values can have a strong response to selection during breeding.

Phenomic tools provide rapid and accurate trait estimation and were used to analyze sainfoin fruit and seed related traits in the current study. Although phenomics is a relatively new approach, it has already been implemented in many crops to accurately generate data quickly, and reliably, and has just been applied to seed images in sainfoin. The perennial grain crop

domestication depends on accurate estimation of seed yield and quality related traits in a large number of individuals. Our results indicate phenomics tools provide precise and accurate estimates while obtaining data from traits that are difficult to measure quickly and easily in sainfoin and could be estimated to other perennial crops.

Access to sequence-based SNPs in the order of hundreds of thousands have helped to delineate the nature genetic variation in crop gene pools as well as the hierarchical structure of the genetic variation. We obtained many SNP markers through skim sequencing in a pool of 430 individuals from 43 accessions/population/cultivars. Genotypic analyses in the present study revealed no clear distinction among the germplasm from the three cultivated species a pattern that was also observed with phenotype-based clustering. The results here suggest reconsidering the species status of the cultivated germplasm. Given the fact that a generous species status offered in a closer genus *Medicago* revoked after extensive genetic studies (Şakiroğlu & İlhan, 2021), it could be argued that the current species status needs to be further investigated. Analyses using evolutionarily conserved DNA markers could be conducted to complement the current results.

Despite the general overall conclusion, this study did not evaluate seed quality traits of the germplasm as protein yield and amino acid composition of the germplasm. Evaluating seed nutritional quality and economic analyses of sainfoin seeds, as protein source for food and grain feed, would be other interesting research venues to pursue.

REFERENCES

Abou-El-Enain, M. M. (2002). Chromosomal criteria and their phylogenetic implications in the genus Onobrychis Mill. Sect. Lophobrychis (Leguminosae), with special reference to Egyptian species. *Botanical Journal of the Linnean Society*, *139*(4), 409–414. https://doi.org/10.1046/j.1095-8339.2002.00075.x

Acharya, S. N. (2015). AAC Mountainview sainfoin (Onobrychis viciifoila subsp. Viciifolia). *Canadian Journal of Plant Science*, *95*(3), 603–607. https://doi.org/10.4141/cjps-2014-332

Acharya, S. N. (2018). AAC Glenview sainfoin (Onobrychis viciifoila subsp. Viciifolia). *Canadian Journal of Plant Science*, *98*(4), 980–984. https://doi.org/10.1139/cjps-2017-0331

Açıkgöz, E. (1991). *Yem Bitkileri Tarımı*. Bursa: Uludağ Üniversitesi Yayınevi.

Açıkgöz, E. (2001). Yem Bitkileri. *Uludağ Üniversitesi Ziraat Fakültesi Yayınları*.

Afzal, M., Alghamdi, S. S., Migdadi, H. H., Khan, M. A., Nurmansyah, Mirza, S. B., & El-Harty, E. (2020). Legume genomics and transcriptomics: From classic breeding to modern technologies. *Saudi Journal of Biological Sciences*, *27*(1), 543–555. https://doi.org/10.1016/j.sjbs.2019.11.018

Akçelik, E., Avcı, S., Uzun, S., & Sancak, C. (2012). Karyotype analysis of some Onobrychis (sainfoin) species in Turkey. *Archives of Biological Sciences*, *64*. https://doi.org/10.2298/ABS1202567A

Akopian, J. (2009). J.A. Akopian. On some wild relatives of cultivated sainfoin (Onobrychis L.) from the flora of Armenia. *Crop Wild Relative. The Newsletter of the CWRSG*, 17–18.

Aktoklu, E. (1995). Türkiye'de yetişen Onobrychis Miller (Fabaceae) türlerinin revizyonu. *İnönü Üniversitesi Fen Bilimleri Enstitüsü Biyoloji Anabilim Dalı*.

Alishan, G. (1895). *Haybousak kam haykakan busabaratyun (Armenian botany or botanic glossary). Surb Ghazar, Venice. 624 pp. [In Armenian]*. 624.

Anderson, C. (2025). Sainfoin, Fenugreek: Not Your Everyday Feed. Retrieved September 25, 2025, from DTN Progressive Farmer website: https://www.dtnpf.com/agriculture/web/ag/news/article/2016/03/23/sainfoin-fenugreek-everyday-feed

Annicchiarico, P., Barrett, B., Brummer, E. C., Julier, B., & Marshall, A. H. (2015). Achievements and Challenges in Improving Temperate Perennial Forage Legumes. *Critical Reviews in Plant Sciences*. (world). Retrieved from https://www.tandfonline.com/doi/abs/10.1080/07352689.2014.898462

Avcı, S. (2010). *Türkiye'de doğal olarak yetişen yabani korunga (Onobrychis sp.) türlerinin toplanması ve morfolojik özelliklerinin belirlenmesi* (DoctoralThesis, Fen Bilimleri Enstitüsü). Fen Bilimleri Enstitüsü. Retrieved from https://dspace.ankara.edu.tr/xmlui/handle/20.500.12575/34207

Awada, L., Phillips, P. W. B., & Bodan, A. M. (2024). The evolution of plant phenomics: Global insights, trends, and collaborations (2000-2021). *Frontiers in Plant Science*, *15*. https://doi.org/10.3389/fpls.2024.1410738

Azuhnwi, B. N., Boller, B., Dohme-Meier, F., Hess, H. D., Kreuzer, M., Stringano, E., & Mueller-Harvey, I. (2013). Exploring variation in proanthocyanidin composition and content of sainfoin (Onobrychis viciifolia). *Journal of the Science of Food and Agriculture*, *93*(9), 2102–2109. https://doi.org/10.1002/jsfa.6119

Baldinger, L., Hagmüller, W., Minihuber, U., Matzner, M., & Zollitsch, W. (2016). Sainfoin seeds in organic diets for weaned piglets—Utilizing the protein-rich grains of a long-known forage legume. *Renewable Agriculture and Food Systems*, *31*(1), 12–21. https://doi.org/10.1017/S1742170514000386

Barrett, B., Baird, I., & Woodfield, D. (2009). White Clover Seed Yield: A Case Study in Marker-Assisted Selection. In T. Yamada & G. Spangenberg (Eds.), *Molecular Breeding of Forage and Turf* (pp. 241–250). New York, NY: Springer. https://doi.org/10.1007/978-0-387-79144-9_22

Barrett, R. P. (1990). Legume species as leaf vegetables. *Advances in New Crops. Proceedings of the First National Symposium "New Crops: Research, Development, Economics", Indianapolis, Indiana, USA, 23-26 October 1988*, 391–396.

Beković, D., Stojković, S., Biberdžić, M., & Knežević, J. (2016). *The Influence of Row Spacing on Yield and Quality Levels of Sainfoin (Onobrychis viciifolia Scop.) Seed in Agroecological Conditions of South Serbia*.

Bhattarai, S. (2017). *Characterization of diverse germplasm of sainfoin (Onobrychis viciifolia Scop.) using agro-morphological traits and AFLP molecular markers* (Thesis, University of Saskatchewan). University of Saskatchewan. Retrieved from https://harvest.usask.ca/handle/10388/7977

Bilder, R. M., Sabb, F. W., Cannon, T. D., London, E. D., Jentsch, J. D., Parker, D. S., … Freimer, N. B. (2009). Phenomics: The systematic study of phenotypes on a genome-wide scale. *Neuroscience*, *164*(1), 30–42. https://doi.org/10.1016/j.neuroscience.2009.01.027

Bingham, E. T., Groose, R. W., Woodfield, D. R., & Kidwell, K. K. (1994). *Complementary Gene Interactions in Alfalfa are Greater in Autotetraploids than Diploids*. https://doi.org/10.2135/cropsci1994.0011183X003400040001x

Bourke, P. M., van Geest, G., Voorrips, R. E., Jansen, J., Kranenburg, T., Shahin, A., … Maliepaard, C. (2018). polymapR—linkage analysis and genetic map construction from F1 populations of outcrossing polyploids. *Bioinformatics*, *34*(20), 3496–3502. https://doi.org/10.1093/bioinformatics/bty371

Bourke, P. M., Voorrips, R. E., Visser, R. G. F., & Maliepaard, C. (2018). Tools for Genetic Studies in Experimental Populations of Polyploids. *Frontiers in Plant Science*, *9*, 513. https://doi.org/10.3389/fpls.2018.00513

Bretagnolle, F., Thompson, J. D., & Lumaret, R. (1995). The Influence of Seed Size Variation on Seed Germination and Seedling Vigour in Diploid and Tetraploid Dactylis glomerata L. *Annals of Botany*, *76*(6), 607–615. https://doi.org/10.1006/anbo.1995.1138

Bulatova, K., Massonichich-Shotunova, R., Meiirman, G., & Mazkirat, S. (2015). Diversity of sainfoin (Onobrychis Mill.) collection samples according to the spectra of storage proteins. *Genetika*, *47*(2), 375–382. https://doi.org/10.2298/GENSR1502375B

Butkutė, B., Padarauskas, A., Ceseviciene, J., Taujenis, L., & Norkevičienė, E. (2018). Phytochemical composition of temperate perennial legumes. *Crop and Pasture Science*, *69*, 1020–1030. https://doi.org/10.1071/CP18206

Çaçan, E., Kökten, K., & Koç, A. (2023). Determination of High Yield and Quality Sainfoin Genotypes (Onobrychis viciifolia Scop.) for the Bingöl Province of Turkey. *Kahramanmaraş Sütçü İmam Üniversitesi Tarım ve Doğa Dergisi*, *26*(3), 619–628. https://doi.org/10.18016/ksutarimdoga.vi.1084441

Carbonero, C., Mueller-Harvey, I., Brown, T., & Smith, L. (2011). Sainfoin (Onobrychis viciifolia): A beneficial forage legume. Plant Genet. *Plant Genetic Resources*, *9*, 70–85. https://doi.org/10.1017/S1479262110000328

Cash, S. D., & Ditterline, R. L. (1996). Seed size effects on growth and N2 fixation of juvenile sainfoin. *Field Crops Research*, *46*(1–3), 145–151. https://doi.org/10.1016/0378-4290(95)00096-8

Cebeci, H. (2011). *Farklı kökenli korunga (Onobrychis viciifolia Scop. Ve Onobrychis altissima Grossh) Populasyonlarının tarımsal özelliklerinin belirlenmesi. Ankara, Türkiye* (Yüksek Lisans Tezi). Ankara Üniversitesi Fen Bilimleri Enstitusu Tarla Bitkileri Anabilim Dalı,.

Çelik Oğuz, A., Karakaya, A., Avcı, S., Sancak, C., & Özcan, S. (2011). Powdery mildews observed on Onobrychis spp. In Turkey. *Australasian Plant Disease Notes*, *6*. https://doi.org/10.1007/s13314-011-0017-7

Cevher, C. (2023). Socioeconomic Factors Affecting Sustainable Management of Improved Rangelands in Kayseri, Turkey. *Rangeland Ecology & Management*, *87*, 44–54. https://doi.org/10.1016/j.rama.2022.11.009

Chen, J., Leach, L., Yang, J., Zhang, F., Tao, Q., Dang, Z., … Luo, Z. (2021). A tetrasomic inheritance model and likelihood-based method for mapping quantitative trait loci in autotetraploid species. *New Phytologist*, *230*(1), 387–398. https://doi.org/10.1111/nph.16413

Chen, M., Fan, W., Ji, F., Hua, H., Liu, J., Yan, M., … Huang, X. (2021). Genome-wide identification of agronomically important genes in outcrossing crops using OutcrossSeq. *Molecular Plant*, *14*(4), 556–570. https://doi.org/10.1016/j.molp.2021.01.003

Chen, Z. J., & Birchler, J. A. (2013). *Polyploid and Hybrid Genomics*. John Wiley & Sons.

Collard, B. C. Y., & Mackill, D. J. (2008). Marker-assisted selection: An approach for precision plant breeding in the twenty-first century. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1491), 557–572. https://doi.org/10.1098/rstb.2007.2170

Craine, E. B., Barriball, S., Şakiroğlu, M., Peters, T., & Schlautman, B. (2024). Amino acid and fatty acid profiles of perennial Baki™ bean. *Frontiers in Nutrition*, *10*. https://doi.org/10.3389/fnut.2023.1292628

De Candolle, A. (1883). *Origine des plantes cultivées*. G. Baillière et cie.

Delgado, I., Salvia, J., Buil, I., & Andrés, C. (2008). The agronomic variability of a collection of sainfoin accessions. *Spanish Journal of Agricultural Research, ISSN 1695-971X, Nº. 3, 2008, Pags. 401-407*, *6*. https://doi.org/10.5424/sjar/2008063-333

Doerge, R. W., & Craig, B. A. (2000). Model selection for quantitative trait locus analysis in polyploids. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(14), 7951–7956. https://doi.org/10.1073/pnas.97.14.7951

Elçi, Ş. (2005). Baklagil ve buğdaygil yem bitkileri.

Emre, I., Turgut-Balik, D., Sahin, A., & Kursat, M. (2007). Total electrophoretic band patterns of some Onobrychis species growing in Turkey. *American-Eurasian Journal of Agricultural and Environmental Sciences*, *2*(2), 123–126.

Ertuş, M. M. (2021). Genetic Relationship in Sainfoin (Onobrychis viciifolia) Landraces Cultivated East Anatolia by using RAPD and ISSR Markers. *Legume Research*, *44*(8), 888–893. https://doi.org/Doi:%252010.18805/LR-614

Ferreira, R. C. U., Lara, L. A. de C., Chiari, L., Barrios, S. C. L., do Valle, C. B., Valério, J. R., … de Souza, A. P. (2019). Genetic Mapping With Allele Dosage Information in

Tetraploid Urochloa decumbens (Stapf) R. D. Webster Reveals Insights Into Spittlebug (Notozulia entreriana Berg) Resistance. *Frontiers in Plant Science*, *10*. Retrieved from https://www.frontiersin.org/articles/10.3389/fpls.2019.00092

Gallais, A. (1981). Quantitative genetics and breeding theory of autopolyploid plants. Retrieved September 30, 2025, from https://www.cabidigitallibrary.org/doi/full/10.5555/19821610476

Genckan, M. S. (1992). Culture of forage crops. *Aegean University, Agriculture Faculty Publication*, (467), 519.

Goplen, B. P., Richards, K. W., & Moyer, J. R. (1991). *Sainfoin for western Canada* (Rev). Ottawa: Minister of Supply and Services Canada.

Guglielmelli, A., Calabro`, S., Primi, R., Carone, F., Cutrignelli, M. I., Tudisco, R., … Danieli, P. P. (2011). In vitro fermentation patterns and methane production of sainfoin (Onobrychis viciifolia Scop.) hay with different condensed tannin contents: In vitro fermentation, methane production and tannin content of sainfoin. *Grass and Forage Science*, *66*(4), 488–500. https://doi.org/10.1111/j.1365-2494.2011.00805.x

Hallauer, A. R., Filho, J. B. M., & Carena, M. J. (2010). Hereditary Variance: Mating Designs. In M. J. Carena, A. R. Hallauer, & J. B. Miranda Filho (Eds.), *Quantitative Genetics in Maize Breeding* (pp. 81–167). New York, NY: Springer. https://doi.org/10.1007/978-1-4419-0766-0_4

Hanley, M. E., Franco, M., Pichon, S., Darvill, B., & Goulson, D. (2008). *Breeding system, pollinator choice and variation in pollen quality in British herbaceous plants*. Retrieved from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/j.1365-2435.2008.01415.x

Hanna, M. R., Cooke, D. A., Smoliak, S., & Goplen, B. P. (1972). Sainfoin for western Canada. Retrieved October 1, 2025, from CABI Databases website: https://www.cabidigitallibrary.org/doi/full/10.5555/19760740409

Harlan, J. R., & DeWet, J. M. J. (1975). *The Origins of Polyploidy*.

Hayot Carbonero, C., Carbonero, F., Smith, L. M. J., & Brown, T. A. (2013). Cytological characterisation of the underutilized forage crop Onobrychis viciifolia Scop. And other members of the Onobrychis genus. *Genetic Resources and Crop Evolution*, *60*(7), 1987–1996. https://doi.org/10.1007/s10722-013-9967-2

Hayot Carbonero, C., Mueller-Harvey, I., Brown, T. A., & Smith, L. (2011). Sainfoin (Onobrychis viciifolia): A beneficial forage legume. *Plant Genetic Resources*, *9*(1), 70–85. https://doi.org/10.1017/S1479262110000328

Hillis, D. M., & Moritz, C. (1990). Molecular Systematics. *Annals of the Entomological Society of America*, *85*, 536–537. https://doi.org/10.1093/aesa/85.4.536a

Hosaininejadmir, F., Jafari, A. A., & Nakhjavan, S. (2011). Seed and forage yield in populations of sainfoin (Onobrychis sativa) grown as swards. *Journal of Food, Agriculture & Environment*, *9*(1), 404–408.

Karabulut, E., Erkoç, K., Acı, M., Aydın, M., Barriball, S., Braley, J., … Şakiroğlu, M. (2023). Sainfoin (Onobrychis spp.) crop ontology: Supporting germplasm characterization and international research collaborations. *Frontiers in Plant Science*, *14*. Retrieved from https://www.frontiersin.org/articles/10.3389/fpls.2023.1177406

Kells, A. (2001). Sainfoin: An alternative forage crop for bees: Bee World: Vol 82, No 4. Retrieved January 12, 2023, from https://www.tandfonline.com/doi/abs/10.1080/0005772X.2001.11099526

Kempf, K. (2016). *Self-fertilization and marker-trait associations in sainfoin (Onobrychis viciifolia)* (ETH Zurich; p. 1 Band) [Application/pdf]. ETH Zurich. https://doi.org/10.3929/ETHZ-A-010852406

Kempf, K., Grieder, C., Walter, A., Widmer, F., Reinhard, S., & Kölliker, R. (2015). Evidence and consequences of self-fertilisation in the predominantly outbreeding forage legume Onobrychis viciifolia. *BMC Genetics*, *16*(1), 117. https://doi.org/10.1186/s12863-015-0275-z

Kempf, K., Malisch, C. S., Grieder, C., Widmer, F., & Kölliker, R. (2017). Marker-trait association analysis for agronomic and compositional traits in sainfoin (Onobrychis viciifolia). *Genetics and Molecular Research*, *16*(1), gmr16019483. https://doi.org/10.4238/gmr16019483

Kempf, K., Mora-Ortiz, M., Smith, L. M. J., Kölliker, R., & Skøt, L. (2016). Characterization of novel SSR markers in diverse sainfoin (Onobrychis viciifolia) germplasm. *BMC Genetics*, *17*(1), 124. https://doi.org/10.1186/s12863-016-0431-0

Kimmel, M., Chakraborty, R., King, J. P., Bamshad, M., Watkins, W. S., & Jorde, L. B. (1998). Signatures of Population Expansion in Microsatellite Repeat Data. *Genetics*, *148*(4), 1921–1930. https://doi.org/10.1093/genetics/148.4.1921

Koivisto, J. M., & Lane, G. P. E. (2001). Royal Agricultural College on behalf of the BGS Forage Legumes Special Interest Group. Retrieved September 30, 2025, from https://www.feedipedia.org/node/1727

Kölliker, R., Kempf, K., Malisch, C. S., & Lüscher, A. (2017). Promising options for improving performance and proanthocyanidins of the forage legume sainfoin (Onobrychis viciifolia Scop.). *Euphytica*, *213*(8), 179. https://doi.org/10.1007/s10681-017-1965-6

Korani, W., O'Connor, D., Chu, Y., Chavarro, C., Ballen, C., Ozias-Akins, P., … Clevenger, J. (2021, March 14). *Accurate analysis of short read sequencing in complex genomes: A case study using QTL-seq to target blanchability in peanut (Arachis hypogaea)* (p. 2021.03.13.435236). p. 2021.03.13.435236. bioRxiv. https://doi.org/10.1101/2021.03.13.435236

Lauriault, L., Marsalis, Mark A., & Contreras-Govea, Francisco E. (2009). *Sainfoin Production in New Mexico*. Retrieved from https://www.academia.edu/22405961/Sainfoin_Production_in_New_Mexico

Legesse, G., Beauchemin, K. A., Ominski, K. H., McGeough, E. J., Kroebel, R., MacDonald, D., … McAllister, T. A. (2015). Greenhouse gas emissions of Canadian beef production in 1981 as compared with 2011. *Animal Production Science*, *56*(3), 153–168. https://doi.org/10.1071/AN15386

Lobos, G. A., Camargo, A. V., del Pozo, A., Araus, J. L., Ortiz, R., & Doonan, J. H. (2017). Editorial: Plant Phenotyping and Phenomics for Plant Breeding. *Frontiers in Plant Science*, *8*, 2181. https://doi.org/10.3389/fpls.2017.02181

Lorenz, M. M., Eriksson, T., & Udén, P. (2010). Effect of wilting, silage additive, PEG treatment and tannin content on the distribution of N between different fractions after ensiling of three different sainfoin (Onobrychis viciifolia) varieties: Tannins and N-fractions in ensiledsainfoin. *Grass and Forage Science*, *65*(2), 175–184. https://doi.org/10.1111/j.1365-2494.2010.00736.x

Majidi, M. M., Mirlohi, A., & Amini, F. (2009). Genetic variation, heritability and correlations of agro-morphological traits in tall fescue (Festuca arundinacea Schreb.). *Euphytica*, *167*(3), 323–331. https://doi.org/10.1007/s10681-009-9887-6

Massoud, R., Karamian, R., & Hadadi, A. (2010). Cytosystematics of three Onobrychis species (Fabaceae) in Iran. *Caryologia*. (world). Retrieved from https://www.tandfonline.com/doi/abs/10.1080/00087114.2010.10589733

McCouch, S. R., Wright, M. H., Tung, C.-W., Maron, L. G., McNally, K. L., Fitzgerald, M., … Mezey, J. (2016). Open access resources for genome-wide association mapping in rice. *Nature Communications*, *7*(1), 10532. https://doi.org/10.1038/ncomms10532

Meyering, B., Barriball, S., & Schlautman, B. (2025). Deep learning driven, image-based phenotyping of seed processing efficiency in sainfoin (Onobrychis viciifolia). *Frontiers in Plant Science*, *16*. https://doi.org/10.3389/fpls.2025.1655350

Mohajer, S., Jafari, A. A., & Taha, R. (2011). Studies on seed and forage yield in 10 populations of sainfoin (Onobrychis sativa) grown as spaced plants and swards. *Journal of Food Agriculture and Environment*, *9*, 222–227.

Mohajer, S., Jafari, A. A., Taha, R. M., Yaacob, J. S., & Saleh, A. (2013). Genetic diversity analysis of agro-morphological and quality traits in populations of sainfoin (Onobrychis sativa). *Australian Journal of Crop Science*, *7*(7), 1024–1031.

Monsen, S. B., Stevens, R., & Shaw, N. L. (2004). Restoring Western Ranges and Wildlands, vol. 2. *Gen. Tech. Rep. RMRS-GTR-136-Vol-2. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. Pages 295-698 plus Index*, *136*. https://doi.org/10.2737/RMRS-GTR-136-V2

Mora-Ortiz, M., & Smith, L. M. J. (2018). Onobrychis viciifolia; a comprehensive literature review of its history, etymology, taxonomy, genetics, agronomy and botany. *Plant Genetic Resources*, *16*(5), 403–418. https://doi.org/10.1017/S1479262118000230

Mora-Ortiz, Marina, Swain, M. T., Vickers, M. J., Hegarty, M. J., Kelly, R., Smith, L. M. J., & Skøt, L. (2016). De-novo transcriptome assembly for gene identification, analysis,

annotation, and molecular marker discovery in Onobrychis viciifolia. *BMC Genomics*, *17*(1), 756. https://doi.org/10.1186/s12864-016-3083-6

Nakhjavan, S., Bajolvand, M., Jafari, A. A., & Sepavand, K. (2011). *Variation for Yield and Quality Traits in Populations of Sainfoin (Onobrychis sativa)*.

Nosrati, H., Feizi, M. A. H., Tarrah, S. S., & Haghighi, A. R. (2012). *Population Genetic Variation in Sainfoin (Fabacae) revealed by RAPD Markers*. (1), 11–16.

Nosrati, Houshang, Feizi, M. A. H., Latifian, F., & Haghighi, A. R. (2016). *Eco-geographical variations of ISSRS among populations of Onobrychis viciifolia (Sainfoin, Fabacae)*. (2).

Novobilský, A., Mueller-Harvey, I., & Thamsborg, S. M. (2011). Condensed tannins act against cattle nematodes. *Veterinary Parasitology*, *182*(2), 213–220. https://doi.org/10.1016/j.vetpar.2011.06.003

Özbek, H. (2011). Korunga (Onobrychis Viciifolia Scop.): Önemli Bir Arı Bitkisi. *Uludağ Arıcılık Dergisi*, *11*(2), 51–62.

Özkan, S. (2017). *Bazı korunga (onobrychis viciifolia scop.) çeşit ve populasyonlarında mikrosatellit belirteçleri kullanılarak genetik çeşitliliğin belirlenmesi* (masterThesis, Namık Kemal Üniversitesi). Namık Kemal Üniversitesi. Retrieved from http://acikerisim.nku.edu.tr/xmlui/handle/20.500.11776/2508

Özkan, S., & Bilgen, B. B. (2019). Bazı Korunga (Onobrychis viciifolia) Çeşit ve Popülasyonlarının Mikrosatellit Belirteçleri Kullanılarak Genetik Karakterizasyonu. *Tekirdağ Ziraat Fakültesi Dergisi*, 45–54. https://doi.org/10.33462/jotaf.516991

Özkan, U., & Demirbağ, N. (2016). Türkiyede Kaliteli Kaba Yem Kaynaklarını Mevcut Durumu. *Türk Bilimsel Derlemeler Dergisi*, *9*, 23–27.

Pasala, R., & Pandey, B. B. (2020). Plant phenomics: High-throughput technology for accelerating genomics. *Journal of Biosciences*, *45*, 111.

Pellikaan, W. F., Stringano, E., Leenaars, J., Bongers, D. J. G. M., Schuppen, S. van L., Plant, J., & Mueller-Harvey, I. (2011). Evaluating effects of tannins on extent and rate of *in vitro* gas and CH4 production using an automated pressure evaluation system (APES). *Animal Feed Science and Technology*, *166–167*, 377–390. https://doi.org/10.1016/j.anifeedsci.2011.04.072

Perez-Sanz, F., Navarro, P. J., & Egea-Cortines, M. (2017). Plant phenomics: An overview of image acquisition technologies and image data analysis algorithms. *GigaScience*, *6*(11), 1–18. https://doi.org/10.1093/gigascience/gix092

Piper, C. V. (1914). *Forage Plants and their culture*. The Macmillan Company.

Pulluk Uğur, B. (2022). *RDNA its region and ISSRs markers-based genetic characterization of some sainfoin (Onobrychis L.) genetic resources from Central Anatolia region* (doctoralThesis, Ankara University). Ankara University. Retrieved from https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp

Qiao, Y., Cheng, Q., Zhang, Y., Yan, W., Yi, F., & Shi, F. (2021). Transcriptomic and chemical analyses to identify candidate genes involved in color variation of sainfoin flowers. *BMC Plant Biology*, *21*(1), 61. https://doi.org/10.1186/s12870-021-02827-8

Ragot, M., & Lee, M. (2007). Marker-assisted selection in maize: Current status, potential, limitations and perspectives from the private and public sectors. In *Marker-assisted selection. Current status and future perspectives in crops, livestock, forestry and fish* (pp. 117–150). Retrieved from https://inis.iaea.org/records/gcg5f-96y70

Ranjbar, M., Karamian, R., & Hajmoradi, F. (2010). Chromosome number and meiotic behaviour of two populations of Onobrychis chorassanica Bunge (O. sect.

Hymenobrychis) in Iran. *Journal of Cell and Molecular Research*, *2*(1), 49–55. https://doi.org/10.22067/jcmr.v2i1.7449

Rietzschel, E., & De Buyzere, M. (2012). High-sensitive C-reactive protein: Universal prognostic and causative biomarker in heart disease? *Biomarkers in Medicine*, *6*(1), 19–34. https://doi.org/10.2217/bmm.11.108

Şahin, B., Yılmaz, C., Gülsemin Savaş Tuna, G., Yücel, G., & Tuna, M. (2019). *Determination of Nuclear DNA Content and Ploidy of Onobrychis Miller Accessions Maintained in USDA Gene Banks by Flow Cytometer.* Retrieved from https://avesis.omu.edu.tr/yayin/a09dfc7d-e1aa-428f-94f8-415b6faf66a8/determination-of-nuclear-dna-content-and-ploidy-of-onobrychis-miller-accessions-maintained-in-usda-gene-banks-by-flow-cytometer

Şakiroğlu, M. (2021). *Population Genomics of Perennial Temperate Forage Legumes*. In *Population Genomics*. Cham: Springer International Publishing. https://doi.org/10.1007/13836_2021_90

Sakiroglu, M., & Brummer, E. C. (2017). Identification of loci controlling forage yield and nutritive value in diploid alfalfa using GBS-GWAS. *TAG. Theoretical and Applied Genetics. Theoretische Und Angewandte Genetik*, *130*(2), 261–268. https://doi.org/10.1007/s00122-016-2782-3

Şakiroğlu, M., Doyle, J. J., & Charles Brummer, E. (2010). Inferring population structure and genetic diversity of broad range of wild diploid alfalfa (Medicago sativa L.) accessions using SSR markers. *Theoretical and Applied Genetics*, *121*(3), 403–415. https://doi.org/10.1007/s00122-010-1319-4

Şakiroğlu, M., & İlhan, D. (2021). *Medicago sativa* species complex: Revisiting the century-old problem in the light of molecular tools. *Crop Science*, *61*(2), 827–838. https://doi.org/10.1002/csc2.20316

Sakiroglu, M., Moore, K. J., & Brummer, E. C. (2011). Variation in Biomass Yield, Cell Wall Components, and Agronomic Traits in a Broad Range of Diploid Alfalfa Accessions. *Crop Science*, *51*(5), 1956–1964. https://doi.org/10.2135/cropsci2010.11.0658

Sayar, M., Han, Y., & Basbag, M. (2022). Forage Yield and Forage Quality Traits of Sainfoin (Onobrychis viciifolia Scop.) Genotypes and Evaluations with Biplot Analysis. *Fresenius Environmental Bulletin*, *31*, 4009–4017.

Shen, S., Chai, X., Zhou, Q., Luo, D., Wang, Y., & Liu, Z. (2019). Development of polymorphic EST-SSR markers and characterization of the autotetraploid genome of sainfoin (Onobrychis viciifolia). *PeerJ*, *7*, e6542. https://doi.org/10.7717/peerj.6542

Sheppard, S. C., & Bittman, S. (2012). Farm practices as they affect NH3 emissions from beef cattle. *Canadian Journal of Animal Science*, *92*(4), 525–543. https://doi.org/10.4141/cjas2012-055

Sheppard, Steve C., Cattani, D. J., Ominski, K. H., Biligetu, B., Bittman, S., & McGeough, E. J. (2019). Sainfoin production in western Canada: A review of agronomic potential and environmental benefits. *Grass and Forage Science*, *74*(1), 6–18. https://doi.org/10.1111/gfs.12403

Soltis, D. E., & Soltis, P. S. (1993). *Molecular data facilitate a reevaluation... - Google Scholar*. Retrieved from https://scholar.google.com/scholar_lookup?hl=en&volume=12&publication_year=1993&pages=243-273&journal=Crit.+Rev.+Pl.+Sci.&author=D.E.+Soltis&author=P.S.+Soltis&title=M

olecular+data+facilitate+a+reevaluation+of+traditional+tenets+of+polyploid+evolution

Soltis, P., & Doyle, J. J. (2012). *Molecular Systematics of Plants II: DNA Sequencing*. Springer Science & Business Media.

Somay Akçelik, E., Avcı, S., Uzun, S., & Sancak, C. (2012). Karyotype analysis of some Onobrychis (sainfoin) species in Turkey. *Archives of Biological Sciences (Serbia)*.

Sottie, E. T., Acharya, S. N., McAllister, T., Thomas, J., Wang, Y., & Iwaasa, A. (2014). Alfalfa Pasture Bloat Can Be Eliminated by Intermixing with Newly-Developed Sainfoin Population. *Agronomy Journal*, *106*(4), 1470–1478. https://doi.org/10.2134/agronj13.0378

Sutcu, T., Bilgen, B. B., & Tuna, M. (2022). Analysis of genetic diversity among Onobrychis accessions with high agronomic performance by simple sequence repeat (SSR) markers. *Molecular Biology Reports*, *49*(6), 5659–5668. https://doi.org/10.1007/s11033-022-07584-x

Tarasenko, N., Butina, E., & Gerasimenko, E. (2015). Peculiarities of Chemical Composition of Sainfoin Seeds Powder. *Oriental Journal of Chemistry*, *31*(3), 1673–1682. https://doi.org/10.13005/ojc/310346

Tardieu, F., Cabrera-Bosquet, L., Pridmore, T., & Bennett, M. (2017). Plant Phenomics, From Sensors to Knowledge. *Current Biology*, *27*(15), R770–R783. https://doi.org/10.1016/j.cub.2017.05.055

Temel, O., & Özalp, M. (2016). Artvin'in Şavşat ilçesinde yetiştirilen korunga (Onobrychis sativa Scop.) yem bitkisinin verimi ve kalitesi üzerine yükseltinin ve bazı toprak özelliklerinin etkisi. *Anadolu Tarım Bilimleri Dergisi*, *31*(1), 106–116. https://doi.org/10.7161/anajas.2016.31.1.106-116

Teuber, L. R., & Brick, M. A. (1988). Morphology and Anatomy. In *Alfalfa and Alfalfa Improvement* (pp. 125–162). John Wiley & Sons, Ltd. https://doi.org/10.2134/agronmonogr29.c4

Theodoridou, K., Aufrère, J., Andueza, D., Morvan, A. L., Picard, F., Stringano, E., … Baumont, R. (2011). *Effect of plant development during first and second growth cycle on chemical composition, condensed tannins and nutritive value of three sainfoin (Onobrychis viciifolia) varieties and lucerne*. *66*(3). Retrieved from https://onlinelibrary.wiley.com/doi/10.1111/j.1365-2494.2011.00798.x

Thomson, J. R. (1951). Seed Studies in Sainfoin. *Grass and Forage Science*, *6*(3), 147–159. https://doi.org/10.1111/j.1365-2494.1951.tb00923.x

Torres, A. M., Avila, C. M., Gutierrez, N., Palomino, C., Moreno, M. T., & Cubero, J. I. (2010). Marker-assisted selection in faba bean (*Vicia faba* L.). *Field Crops Research*, *115*(3), 243–252. https://doi.org/10.1016/j.fcr.2008.12.002

Tosun, F. (1992). Bitki Yetiştiriciliğinin Fizyolojik Esasları.

Tosun, M. (1988). *Kuru ve sulu koş ullarda değişik sıra arası mesafelerin ve değişik fosfor dozlarının korunganin tohum verimi ve diğer agronomik özelliklerine etkileri üzerinde araştırmalar.* (doctoralThesis). Ege Üniv. Ziraat Fak. Tarla Bitkileri Bölümü.

Turk, M., Albayrak, S., Tuzun, C. G., & Yuksel, O. (2011). Effects of Fertilisation and Harvesting Stages on Forage Yield and Quality of Sainfoin (Onobrychis sativa L.). *Bulgarian Journal of Agricultural Science*, (17: 789-794), 6.

TurkStat. (2023). *Yem Bitkileri Üretimi*.

Unamba, C. I. N., Nag, A., & Sharma, R. K. (2015). Next Generation Sequencing Technologies: The Doorway to the Unexplored Genomics of Non-Model Plants. *Frontiers in Plant Science*, *6*, 1074. https://doi.org/10.3389/fpls.2015.01074

Vaughan, D. A., Balázs, E., & Heslop-Harrison, J. S. (2007). From Crop Domestication to Super-domestication. *Annals of Botany*, *100*(5), 893–901. https://doi.org/10.1093/aob/mcm224

Wijekoon, C., Acharya, S. N., Siow, Y. L., Sura, S., Thandapilly, S., & Sabra, A. (2021). Canadian sainfoin and fenugreek as forage and functional foods. *Crop Science*, *61*(1), 1–20. https://doi.org/10.1002/csc2.20280

Wu, R., Ma, C.-X., & Casella, G. (2002). A bivalent polyploid model for linkage analysis in outcrossing tetraploids. *Theoretical Population Biology*, *62*(2), 129–151. https://doi.org/10.1006/tpbi.2002.1608

Xie, C., & Xu, S. (2000). Mapping quantitative trait loci in tetraploid populations. *Genetical Research*, *76*(1), 105–115. https://doi.org/10.1017/s0016672399004395

Yıldız, B., Çıplak, B., & Aktoklu, E. (1999). *Fruit morphology of sections of the genus onobrychis miller (fabaceae) and ITS phylogenetic implications*. Retrieved from http://dspace.balikesir.edu.tr/xmlui/handle/20.500.12462/7709

Yücel, G. (2019). *Kültürü yapılan korunga (onobrychis mill., baklagiller) taksonları ve bazı yabani akrabalarının moleküler sitogenetik yöntemler ile karaktezasyonu* (doctoralThesis, Namık Kemal Üniversitesi). Namık Kemal Üniversitesi. Retrieved from http://acikerisim.nku.edu.tr/xmlui/handle/20.500.11776/3406

Yucel, G., Betekhtin, A., Cabi, E., Tuna, M., Hasterok, R., & Kolano, B. (2022). The Chromosome Number and rDNA Loci Evolution in Onobrychis (Fabaceae). Retrieved October 1, 2025, from https://www.mdpi.com/1422-0067/23/19/11033

Zarrabian, M., Dehkordi, A. N., Ehtemam, M. H., & Majidi, M. M. (2022, May 3). *Relationships of Worldwide Onobrychis species: From seed & pod morphology*

and molecular perspective (p. 2022.05.02.490301). p. 2022.05.02.490301. bioRxiv. https://doi.org/10.1101/2022.05.02.490301

Zarrabian, M., & Majidi, M. M. (2015). Genetic diversity and relationships within and among Onobrychisspecies using molecular markers. *TURKISH JOURNAL OF BOTANY*, *39*, 681–692. https://doi.org/10.3906/bot-1406-47

Zarrabian, M., Majidi, M. M., & Ehtemam, M. H. (2013). Genetic Diversity in a Worldwide Collection of Sainfoin Using Morphological, Anatomical, and Molecular Markers. *Crop Science*, *53*(6), 2483–2496. https://doi.org/10.2135/cropsci2013.03.0130

Zhao, C., Zhang, Y., Du, J., Guo, X., Wen, W., Gu, S., … Fan, J. (2019). Crop Phenomics: Current Status and Perspectives. *Frontiers in Plant Science*, *10*. https://doi.org/10.3389/fpls.2019.00714

APPENDIX

The analyses performed within the scope of the thesis were made in the R program, and all R commands used are available at the GitHub link.
https://github.com/EKarabulut01/All_thesis_analysis_R/tree/main