

João Pedro Martins dos Reis

**Genetic diversity within a wild rocket
[*Diplotaxis tenuifolia* (L.) DC.]
germplasm collection and molecular
characterization of selected accessions**



Universidade do Algarve

Faculdade de Ciências e Tecnologias

2022

João Pedro Martins dos Reis

**Genetic diversity within a wild rocket
[*Diplotaxis tenuifolia* (L.) DC.] germplasm
collection and molecular characterization of
selected accessions**

Mestrado em Biologia Molecular e Microbiana

Trabalho efetuado sob a orientação de:

Orientador: Professor Doutor José Manuel Peixoto Teixeira Leitão,
Universidade do Algarve

Coorientadora: Doutora Maria Paula Mesquita dos Santos Coelho,
Instituto Nacional de Investigação Agrária e Veterinária, I.P.



Universidade do Algarve

Faculdade de Ciências e Tecnologias

2022

Genetic diversity within a wild rocket [*Diplotaxis tenuifolia* (L.) DC.] germplasm collection and molecular characterization of selected accessions

Declaração de autoria de trabalho

Declaro ser autor deste trabalho, que é original e inédito. Autores e trabalhos consultados estão devidamente citados no texto e constam da listagem de referências incluída.

João Pedro Martins dos Reis

Copyright © João Pedro Martins dos Reis

A Universidade do Algarve reserva para si o direito, em conformidade com o disposto no Código do Direito de Autor e dos Direitos Conexos, de arquivar, reproduzir e publicar a obra, independentemente do meio utilizado, bem como de a divulgar através de repositórios científicos e de admitir a sua cópia e distribuição para fins meramente educacionais ou de investigação e não comerciais, conquanto seja dado o devido crédito ao autor e editor respetivos.

Agradecimentos

Gostaria de expressar o meu agradecimento a todas as pessoas que contribuíram direta e indiretamente para a realização deste trabalho.

- Ao meu orientador, Professor Doutor José Leitão, pelo apoio, disponibilidade e todos os contributos na aquisição de conhecimentos essenciais para a realização deste trabalho, assim como toda a sua paciência, necessária em períodos de maior dificuldade;

-Um grande obrigada também à minha coorientadora, Doutora Paula Coelho, por toda a sua disponibilidade e paciência na execução desta dissertação;

-Aos meus pais, que me apoiaram e incentivaram diariamente a dar o melhor de mim, dizendo sempre para seguir o caminho que me trouxesse mais felicidade;

-Aos meus colegas de laboratório, pela ajuda em momentos de maior dificuldade e as horas intermináveis de trabalho efetuado em conjunto;

- Aos meus amigos, em especial ao Xavier Mestre, à Mariana Fernandes e Ana Rita Rodrigues por todo o apoio e carinho manifestado ao longo dos últimos anos;

A todos, o meu mais sincero obrigado!

Abstract

Wild rocket (*Diplotaxis tenuifolia* (L.) DC) is a species known for its growing consumption as “baby leaf” crop. Downy mildew, a disease caused by the pathogen *Hyaloperonospora* sp., is one cause of production losses in *D. tenuifolia* cultivation. To preserve elevated quality standards and sustainable production of wild rocket, the characterization of genetic diversity present in a germplasm collection is an important step towards the identification of genetic resources, fundamental to plant breeding. The objectives of this work encompass the characterization of the genetic diversity within a germplasm collection of *D. tenuifolia*, including a set of *E. sativa* samples, and the unequivocal identification of a group of selected accessions, using molecular markers. Genetic diversity analysis, performed using 5 Inter Simple-Sequence Repeats (ISSR) and 7 Random Amplified Polymorphic DNA (RAPD) primers, resulted in 110 scorable polymorphisms. Wild rocket accessions exhibited intraspecific genetic similarity with values ranging between 0.697 and 0.994. Tighter clusters aggregate accessions with greater genetic similarity provided by the same producing company. The unequivocal identification of 90 accessions was performed using Simple-Sequence Repeats (SSR) and Single Nucleotide Polymorphisms (SNP) markers, the latter analyzed as Cleaved Amplified Polymorphic Sequences (CAPS). These loci were mined from the *D. tenuifolia* genomic assembly (UAlgDiploT.01), uploaded to the genomic databases. Twenty SSR and 19 SNP loci (selected by the presence of a TaqI restriction region in the SNP site), were retrieved for analysis. From this, five SSR and nine SNP loci were selected, giving rise to 70 SSR alleles analyzed by capillary gel electrophoresis and 36 CAPS polymorphisms. The molecular characterization of selected accessions resulted in the identification of specific fingerprints for the accessions, excluding 1 trio and 6 pairs of wild rocket accessions, exhibiting full similarity (100%). Thus, confirming the presence of duplicate accessions provided by different breeding and commercial companies.

Key words: *Diplotaxis tenuifolia*, genetic diversity, RAPD, ISSR, microsatellite, CAPS.

Resumo

Rúcula selvagem (*Diplotaxis tenuifolia* (L.) DC) é uma espécie conhecida pelo seu crescente consumo como cultura “baby leaf”. O míldio, doença causada pelo patógeno *Hyaloperonospora* sp., é uma causa para as perdas de produção no cultivo de *D. tenuifolia*. Para preservar elevados padrões de qualidade e produção sustentável de rúcula selvagem, a caracterização da diversidade genética presente numa coleção de germoplasma é um passo importante na identificação de recursos genéticos, fundamentais para o melhoramento de plantas. Os objetivos deste trabalho englobam a caracterização da diversidade genética de uma coleção de germoplasma de *D. tenuifolia*, incluindo amostras de *E. sativa* e, a identificação inequívoca de um grupo de acessos selecionados, utilizando marcadores moleculares. A análise da diversidade genética, realizada utilizando 5 primers Inter Simple-Sequence Repeats (ISSR) e 7 Random Amplified Polymorphic DNA (RAPD), resultou em 110 polimorfismos contabilizados. Os acessos de rúcula selvagem exibiram elevada similaridade genética intraespecífica com valores entre 0.697 e 0.994. Grupos mais compactos agregam acessos com maior similaridade genética fornecidos pela mesma empresa produtora. A identificação inequívoca de acessos foi realizada utilizando marcadores Simple-Sequence Repeats (SSR) e Single Nucleotide Polymorphisms (SNP), estes últimos analisados como Cleaved Amplified Polymorphic Sequences (CAPS). Estes loci, extraídos do conjunto genômico *D. tenuifolia* (UAlgDiploT.01), foram carregados em bases de dados genômicas. Vinte loci SSR e 19 loci SNP (selecionados pela presença de uma região de restrição TaqI no sítio SNP) foram recuperados para análise. Destes, cinco loci SSR e nove SNP foram selecionados, dando origem a 70 alelos SSR analisados por eletroforese em gel capilar e 36 polimorfismos CAPS contabilizados. A caracterização molecular dos acessos selecionados resultou na identificação de “fingerprints” específicos para os acessos, excluindo 1 trio e 6 pares de acessos de rúcula selvagem, apresentando similaridade total (100%), confirmando a presença de acessos duplicados fornecidos por diferentes empresas produtoras.

Palavras-chave: *Diplotaxis tenuifolia*, diversidade genética, RAPD, ISSR, microsátélites, CAPS.

Resumo Extenso

Diplotaxis tenuifolia (L.) DC., comumente conhecida como rúcula selvagem é uma espécie pertencente à família Brassicaceae que engloba uma vasta diversidade de espécies economicamente importantes. Nativa da bacia Mediterrânea e de certas regiões da Ásia ocidental, a rúcula selvagem encontra-se neste momento dispersa globalmente, sendo conhecida pelo seu crescente uso como saladas “baby leaf”. Esta espécie é muitas vezes confundida com rúcula cultivada (*Eruca sativa* (L.) Cav.), com a qual apresenta algumas diferenças morfológicas. Em fases mais avançadas do desenvolvimento da planta, podem ser observadas claras diferenças, como pétalas amarelas, síliquas septadas e sementes de pequenas dimensões (~5000 por grama) na espécie *D. tenuifolia* vs. pétalas branca, síliquas simples e sementes de maiores dimensões (~500 por grama) de *E. sativa*.

A crescente popularidade no consumo de rúcula selvagem incitou também a necessidade de estabelecimento de uma coleção de germoplasma abrangendo diferentes genótipos de *D. tenuifolia*, permitindo identificação de materiais de interesse para uso em programas de melhoramento genético.

O projeto REMIRucula, em que este trabalho se insere, tem como focos principais: 1) o estabelecimento de uma coleção de germoplasma de rúcula selvagem sediada no Instituto Nacional de Investigação Agrária e Veterinária (INIAV); 2) aumentar a sustentabilidade na produção de *D. tenuifolia*, procedendo à seleção de genótipos resistentes ao míldio (*Hyaloperonospora* sp.) que podem potenciar, simultaneamente, maior proteção ambiental e, altos níveis de qualidade alimentar.

O míldio, uma doença causada pelo oomiceta *Hyaloperonospora* sp. é, atualmente, uma das maiores causas de perdas de produtividade na cultura de rúcula selvagem. Sendo, o uso de fungicidas sintéticos o método mais frequente na mitigação de *Hyaloperonospora* sp. a seleção de genótipos resistentes ao míldio é imprescindível para manter elevados padrões de segurança ambiental e alimentar.

A estimativa da diversidade genética através do uso de marcadores moleculares em *Diplotaxis tenuifolia* foi até recentemente limitada à análise de relações inter-específicas. Martín & Sánchez-Yélamo (2000) verificaram relações genéticas entre 10 espécies do género *Diplotaxis*, utilizando marcadores Inter Simple-Sequence Repeats (ISSR), salientando a proximidade genética de *D. tenuifolia* com as espécies *D. cretacea*, *D. simplex*, *D. viminea* e

D. muralis, todas agrupadas. A relação próxima destas mesmas cinco espécies, foi novamente observada por Eschmann-Grupe *et al.* (2003), através do uso de marcadores Random Amplified Polymorphic DNA (RAPD). Mais recentemente, Taranto *et al.* (2016) relataram o uso de marcadores ISSR para completar os dados reunidos a partir de análises fenotípicas. Confirmando a discriminação da espécie *E. sativa* das espécies *D. tenuifolia* e *D. muralis*.

O presente trabalho, tem como objetivos: 1) a avaliação da diversidade genética de um grupo representativo de acessos de uma coleção de germoplasma de *Diplotaxis tenuifolia*, onde se encontram incluídas amostras de *Eruca sativa*, utilizando marcadores moleculares RAPD e ISSR. 2) a identificação inequívoca de um grupo de acessos selecionados desta coleção de germoplasma, utilizando marcadores Simple-Sequence Repeats (SSR ou microsátélites) e Single Nucleotide Polymorphisms (SNP), sendo estes últimos analisados como marcadores CAPS (Cleaved Amplified Polymorphic Sequences).

Testes prévios num conjunto de 40 primers RAPD e 8 primers ISSR, permitiu a seleção de 7 primers RAPD e 5 primers ISSR eficientes na amplificação de bandas claras em 92 acessos de *D. tenuifolia* e 8 acessos de *E. sativa*. Esta análise permitiu a amplificação 110 polimorfismos, utilizados para calcular a similaridade genética entre os acessos e a construção de um dendrograma representativo desta similaridade. O menor valor de similaridade descrito na matriz de similaridade (0.126) ocorreu, como esperado, entre dois acessos de *D. tenuifolia* e *E. sativa* discriminados em dois aglomerados. Entre acessos de *D. tenuifolia*, o menor valor de similaridade foi observado entre os acessos 23 e 92 (0.697). O maior nível de similaridade genética ocorre entre os acessos 44 e 42 (0.994) e os acessos 44 e 43 (0.994), estritamente agrupados em conjunto com os acessos 6, 40, 41 45 e 46, todos provenientes da mesma empresa de melhoramento. Um mesmo caso ocorre entre os acessos 8, 10, 17 e 19, classificados como parcialmente resistentes ao míldio e, originários da mesma empresa. A presença de agrupamentos de acessos com resistência parcial ao míldio, indicam a hipótese de existir um locus de resistência comum entre acessos provenientes das mesmas empresas. Os resultados das análises moleculares realizadas numa coleção de germoplasma de *Diplotaxis tenuifolia* (e alguns acessos de *E. sativa*) revelaram uma ampla diversidade genética entre os acessos.

Dos resultados da “assembly” genómica de *D. tenuifolia* (UAlgDiploT.01) previamente inserida nas bases de dados genómicas do National Center for Biotechnology Information (NCBI), foram selecionados 20 de um total de 500 *loci* SSR. Após a otimização e verificação destes 20 *loci* SSR, foram mantidos 5 *loci*, sendo posteriormente desenhados primers

fluorescentes. A análise de fragmentos por eletroforese em gel de capilaridade levou à identificação de 70 alelos contabilizados, permitindo uma identificação primária das relações e similaridades genéticas dos 90 acessos analisados.

Dos dados da sequenciação, um total de 500 *loci* SNP foram identificados e introduzidos nas bases de dados do NCBI (GenBank OM144979 a OM145478). Destes, foram selecionados 19 *loci*, nomeadamente pela presença de uma região de restrição da enzima TaqI no local da variante nucleotídica, de modo a originar marcadores Cleaved Amplified Polymorphic Sequences (CAPS). A análise em gel de agarose dos marcadores SNP/CAPS com 9 pares de primers, levou à identificação de 36 polimorfismos para os acessos de *D. tenuifolia* e apenas 4 para *E. sativa*, discriminando claramente as espécies em dois grupos. Nesta análise, 20 acessos de *D. tenuifolia* foram completamente diferenciados por padrões distintos, sendo os restantes 67 acessos da mesma espécie, aglomerados em conjuntos de 2 a 14 acessos contendo similaridade total (100%).

Apesar da sua capacidade para estabelecer as relações genéticas entre os acessos a utilização de marcadores SSR impossibilita a quantificação das similaridades genéticas, visto que, para acessos da mesma espécie são obtidos baixos valores de similaridade genética. Por outro lado, os altos valores de similaridade utilizado marcadores SNP/CAPS refletem a quantidade reduzida de loci utilizados para esta análise. Com isto, uma alternativa à baixa discriminação dos acessos avaliados pelos marcadores SNP/CAPS foi a sua confluência dos dados obtidos na análise SSR.

Como expectado, o uso agrupado dos dados SSR+SNP/CAPS, permitiu novamente a divisão de ambas as espécies *D. tenuifolia* e *E. sativa* em dois “clusters” distintos. Para acessos de *D. tenuifolia* os valores de similaridade genética foram superiores a 0.6 em cerca de 96% dos casos. Com isto, os padrões moleculares obtidos a partir da confluência dos dados SSR+SNP/CAPS permitiu a discriminação inter e intraespecífica dos acessos da coleção de germoplasma de *D. tenuifolia* e de 3 acessos de *E. sativa* utilizados como “outgroup”. Em conjunto, os dados impediram a discriminação em 6 casos de similaridade total entre dois acessos e um caso de similaridade total num grupo de 3 acessos, discutidos com mais detalhe. Os resultados confirmam, que combinação dos marcadores SSR e SNP/CAPS são adequados para a identificação inequívoca dos acessos da coleção de germoplasma de *D. tenuifolia*.

Palavras-chave: *Diplotaxis tenuifolia*, diversidade genética, RAPD, ISSR, microsátélites, CAPS.

Abbreviation List

bp – “base pairs”

CAPS – “Cleaved Amplified Polymorphic Sequence”

DNA – “Deoxyribonucleic acid”

dNTPs – “deoxyribonucleoside triphosphate”

EDTA – “Ethylenediaminetetraacetic acid”

FW – “Fresh weight”

INIAV – “Instituto Nacional de Investigação Agrária e Veterinária”

ISSR – “Inter Simple-Sequence Repeats”

ITS – “Internal Transcribed Spacer”

LED – “Light-Emitting Diode”

LGGI – “Laboratory of Genomics and Genetic Improvement”

NCBI – “National Center for Biotechnology Information”

NGS – “Next-generation sequencing”

PCR – “Polymerase Chain Reaction”

QTL – “Quantitative trait loci ”

RAPD – “Random Amplified Polymorphic DNA”

RNA – “Ribonucleic acid”

SDS – “Sodium Dodecyl Sulfate”

SNP – “Single Nucleotide Polymorphism”

SSR – “Simple-Sequence Repeats”

TE – “Tris-EDTA”

Tris – “tris(hydroxymethyl)aminomethane”

USD – “United States dollar”

UV – “Ultraviolet”

Table List

Table 2.1. Analyzed <i>Diplotaxis tenuifolia</i> (L.) DC. and <i>Eruca</i> spp. accessions.....	8
Table 2.2. Tested and selected RAPD primers.....	10
Table 2.3. Tested and selected ISSR primers.....	10
Table 3.1. SSR primers used for accessions identification	19
Table 3.2. SNP/CAPS primers used for accessions identification	23
Table 3.3. Main results of the SNP/CAPS analyses.....	24

Figure List

- Figure 1.1.** (A) Wild rocket (*Diplotaxis tenuifolia*) yellow flowers; (B) Cultivated rocket (*Eruca sativa*) white flowers.....2
- Figure 1.2.** *Diplotaxis tenuifolia* cultivation in a polytunnel (Vitacress S.A., Odemira).....4
- Figure 1.3.** (A) Fourteen-days-old wild rocket (*Diplotaxis tenuifolia*) plant before pathogen inoculation. (B) A highly susceptible 21-days-old wild rocket plant infected with *Hyaloperonospora* sp. isolate D5.....6
- Figure 2.1.** RAPD amplification of *Diplotaxis tenuifolia* and *Eruca* spp. accessions with Operon Technologies primers OPU01(A) and OPAN16 (B). Notice the very different amplification pattern of the *Eruca* sp. accession 153.....11
- Figure 2.2.** ISSR amplification of *Diplotaxis tenuifolia* and *Eruca* spp. accessions with primers DiploArb_04 (A) and DiploArb_03 (B). Notice the very different amplification patterns of the *Eruca* spp. accessions 68, 71 and 105.....12
- Figure 2.3.** Dendrogram resulting from the UPGMA analysis of the genetic relationships assessed by RAPD and ISSR markers and using the coefficient of similarity DICE, of the 100 *Diplotaxis tenuifolia* and *Eruca* spp. accessions evaluated for their interaction with the pathogenic *Hyaloperonospora* sp. isolate D5 (HS – Highly Susceptible; S – Susceptible; PR – Partially Resistant; R – Resistant;).....13
- Figure 3.1.** *Diplotaxis tenuifolia* whole genome sequence assembly report.....16
- Figure 3.2.** Agarose gel electrophoresis of the preliminary amplification of the SSR locus MT317527 among 10 *D. tenuifolia* accessions.....19
- Figure 3.3.** Capillary polyacrylamide gel electropherograms of the locus MT317577 (A) Homozygous pattern of accession 15; (B) Heterozygous pattern of accession 37. X axis - Fragment size (bp); Y axis - Relative Fluorescence Unit (RFU). Notice the associated smaller “stutter” peaks.....20
- Figure 3.4.** (A) Dendrogram depicting the genetic relationships among the 90 germplasm accessions established based on SSR markers. The *E. sativa* accessions (Nos. 29, 68 and 105 at the bottom) are clustered clearly apart from all other (*D. tenuifolia*) accessions. (B) Dendrogram displaying the genetic relationships among the 90 germplasm accessions established based on SNP/CAPS markers.....22
- Figure 3.5.** Molecular patterns of 5 accessions for the OM145268 locus. Accessions 1 and 4 (Heterozygous Y/N); Accession 3 (Homozygous Y/Y) ; Accessions 6 and 9 (Homozygous N/N); M – DNA ladder VI (NZYTech). Y - one allele restricted; N – one allele not restricted.....24
- Figure 3.6.** Dendrogram depicting the genetic relationships among the 90 germplasm accessions established based on SSR and SNP/CAPS markers. The *Eruca* spp. accessions (Nos. 29, 68 and 105 at the bottom), used as outgroup, are clustered clearly apart from all the other (*D. tenuifolia*) accessions.....26

Index

Agradecimientos.....	i
Abstract.....	ii
Resumo.....	iii
Resumo Extenso.....	iv
Abbreviation List.....	vii
Table List.....	viii
Figure List.....	ix
1. Introduction.....	1
1.1. Metabolite Contents and Impacts on Human Health.....	3
1.2. Rocket Germplasm Collection Initiatives.....	4
1.3. Downy Mildew Disease.....	5
1.4. Objectives.....	6
2. Evaluation of the genetic relationships among a representative set of the wild rocket (<i>Diploaxis tenuifolia</i> L.) germplasm collection established at the INIAV.....	7
2.1. Material and Methods.....	8
2.1.1. Plant Germplasm Accessions.....	8
2.1.2. Plant Growing Conditions.....	8
2.1.3. Molecular Characterization of the Accessions.....	8
2.1.3.1. DNA Extraction and Evaluation.....	8
2.1.3.2. RAPD-PCR Analysis.....	9
2.1.3.3. ISSR Analysis.....	10
2.1.3.4. Data Analysis.....	11
2.2. Results.....	11
2.2.1. Molecular Diversity.....	11
2.3. Discussion.....	14
3. Unequivocal identification of germplasm accessions of wild rocket (<i>Diploaxis tenuifolia</i> (L.) DC.) by specific SSR and single SNP markers identified by previous genome NGS.....	15
3.1. Material and Methods.....	17
3.1.1. Plant Germplasm Accessions.....	17
3.1.2. Plant Growing Conditions.....	17
3.1.3. DNA Extraction Accession Analysis.....	17
3.1.4. Primer Design and Synthesis.....	17
3.1.5. Simple-Sequence Repeats (SSR) Marker Analysis.....	17
3.1.6. Single Nucleotide Polymorphisms (SNP) Marker Analysis.....	18
3.1.7. Data Analysis.....	18
3.2. Results.....	19
3.2.1. SSR Markers Analysis.....	19
3.2.2. SNP/CAPS Markers Analysis.....	21
3.2.3. Combined results of SSR and SNP/CAPS Markers.....	23
3.3. Discussion.....	25
4. Further prospects.....	29
5. References.....	30
6. Supplementary Material.....	34
7. Publications achieved in the scope of this dissertation.....	50

1. Introduction

Today, a fourth generation of vegetables, commercialized as pre-packed baby leaf and ready-to-eat salads are obtaining a crescent popularity based on consistent product quality, freshness, and increased shelf-life (Caruso *et al.*, 2018; Luca *et al.*, 2016). The market demand shows an increasing global trend for the packaged salad market, with a provisional growth from 8.6 billion USD in 2020, to 13.7 billion USD by the year 2028 (GlobeNewswire, 2021).

Accompanying this trend, the market suffered an expansion from pre-packed lettuce salads (*Lactuca sativa*) to products like rocket and other leafy vegetables. Out of the main edible rocket species that include (*Diplotaxis eruroides* and *Diplotaxis muralis*), only two, *Diplotaxis tenuifolia* and *Eruca sativa* are large-scale produced and commercialized as salad crops (Coelho *et al.*, 2021a; Martín & Sánchez-Yélamo, 2000).

Diplotaxis tenuifolia (L.) DC., commonly known as “wild rocket”, is an herbaceous perennial diploid ($2n = 22$) plant species whose commercial relevance has been steadily increasing throughout recent years. The genus *Diplotaxis*, belongs to the Brassicaceae family, that comprises around 400 genera and 4000 species, including important crops as *Brassica oleracea* (broccoli, cabbage, cauliflower, etc.), *Brassica napus* (canola), *Raphanus sativus* (radish) or *Brassica nigra* (black mustard), and the model plant *Arabidopsis thaliana* (Avato & Argentieri, 2015).

The growing importance of *D. tenuifolia* as a cultivated crop has been fuelled up by its characteristic astringent and bitter taste and freshness (Bell *et al.*, 2020; Caruso *et al.*, 2018; eFlorás, 2008). Native to the Mediterranean basin and western Asia region, wild rocket has disseminated worldwide (Caruso *et al.*, 2018). However, characteristics like the elongated taproot system (40 cm in depth), its natural adaptation to considerable ranges of differing edaphoclimatic conditions, and its tolerance to salt and alkaline soils, make *D. tenuifolia* a noxious weed in Europe, Australia, and South America (Caruso *et al.*, 2019; de Vos *et al.*, 2013; Kleemann *et al.*, 2007; Nicoletti *et al.*, 2007).

D. tenuifolia exhibits yellow flowers a characteristic used for distinction from some other *Diplotaxis* and other rocket species (Figure 1.1). The wild rocket juvenile leaves tend to exhibit different morphological shapes, compared to fully developed basal leaves, with oblong and deeply lobed shapes. The fruits are 2 – 5 cm long cylindrical erect siliqua, with two rows of multiple seeds separated by a false septum. The ovoid to ellipsoid seeds are very small (1 – 1.3 mm) and of light weight (~5000 seeds per gram). The basal adventitious buds present in the roots of in this perennial species make it particularly suitable for large-scale production and commercialization since this trait permits multiple rounds of leaf harvesting (eFlorAs, 2008; Hall *et al.*, 2012a; Pignone & Martínez-Laborde, 2011).



Figure 1.1. (A) Wild rocket (*Diplotaxis tenuifolia*) yellow flowers; (B) Cultivated rocket (*Eruca sativa*) white flowers.

Eruca sativa (L.) Cav. ($2n = 22$), commonly known as cultivated or annual garden rocket, is a species of the Brassicaceae family, long time consumed and widely produced baby leaf crop (Hall *et al.*, 2012b). Endemic to the Mediterranean basin and spread throughout Western Asia region and India (Hall *et al.*, 2012a), the annual garden rocket is widely recognized by its pharmacological properties, anciently used for treatment of digestive complications, and as a natural diuretic and aphrodisiac (Taffner *et al.*, 2019). In West Asia and Northern India, *Eruca sativa* seeds are used to manufacture “jamba” oil, mostly used in industrial processes, due to high contents of erucic acid (Garg & Sharma, 2014).

Despite the morphological similarities, white colour of the flowers (Figure 1.1), larger seed size (~500 seeds per gram) and simple siliqua structures clearly differentiates *E. sativa* from *D. tenuifolia* (Coelho *et al.*, 2021a; Hall *et al.*, 2012b).

1.1. Metabolite Contents and Impacts on Human Health

Salad rocket leaves, stems, and flowers are edible and harbour a considerable contents of fibers, minerals, and other beneficial secondary metabolites (glucosinolates, flavonoids, carotenoids, and ascorbic acid). Rocket salads are a beneficial source of non-degraded nutrients, consumed raw in salads and garnishes (Cavaiuolo & Ferrante, 2014).

As in other species, the phytochemical contents are highly linked to the surrounding edaphoclimatic conditions during plant development. Disparities in phytonutrient levels were evidenced in rocket plants, subjected to differing conditions of growth temperature (Jasper *et al.*, 2020), growth season (Bell *et al.*, 2020), light intensity (Jin *et al.*, 2009), soil system (Di Gioia *et al.*, 2018), harvesting rounds (Caruso *et al.*, 2018) and postharvest conditions (Martínez-Sánchez *et al.*, 2006).

Glucosinolates are secondary metabolites present in seeds, roots, leaves, stems, and flowers of multiple plant species, namely among species of the Brassicaceae family, that assure protection against exogenous organisms (Pignone & Martínez-Laborde, 2011; Vig *et al.*, 2009). The damage of the plant tissues caused by harvesting and ingestion leads to the hydrolysis of glucosinolates, catalyzed by the enzyme myrosinase, originating a large panoply of molecular byproducts linked to beneficial effects to human health (Avato & Argentieri, 2015; Bell & Wagstaff, 2014).

In general, the glucosinolate levels are responsible for the overall taste perception and consumer preference in rocket species. Hence, moderate levels of both bitterness and pungency caused by subtle divergences in these bioactive compounds, are also linked to consumer preferences (Bell *et al.*, 2020).

Rocket species, are considered nitrate hyper-accumulator species, reaching levels superior to 9.0 g kg⁻¹ FW, depending on the conditions of plant development (Signore *et al.*, 2020). However, despite the positive cardioprotective effects linked to the daily intake of nitrates (Bondonno *et al.*, 2015), the subsisting negative antinutritional effects caused by the ingestion of nitrites were previously linked to pathologies like non-Hodgkin Lymphoma (Yu *et al.*, 2020), and Methemoglobinemia (McNulty *et al.*, 2022).

Increased popularity and higher quota of wild rocket consumption promoted the incessant cultivation of this herbaceous plant throughout the full year. The impacts of low temperatures and decreased day period significantly impact wild rocket productivity in the autumn-winter cycles. Contrarily, optimal temperature values (16-18°C) and increased day lengths impel faster growth rates in the spring-summer cycles. Thus, the year-round production in protected greenhouse conditions (Figure 1.2) ensure the optimal temperature requirements for wild rocket cultivation (Hall *et al.*, 2012b; Hall *et al.*, 2015).



Figure 1.2. *Diplomatix tenuifolia* cultivation in a polytunnel (Vitacress S.A., Odemira).

Wild rocket can be produced directly under soil or soilless cultivation systems. For baby leaf production, seeds are sown in compact densities, ranging from 4 to 6 kg ha⁻¹ (2000 – 3000 seeds m⁻²), depending on the sowing season and method (Nicoletti *et al.*, 2007; Pimpini & Enzo, 1996).

Wild rocket yields depends on the cropping season, normally ranging from 1.5 to 1.6 kg m⁻² (Hall *et al.*, 2012b). Similar to the temperature, the number of harvests and the harvesting season also condition plant productivity and the nutrient availability (Caruso *et al.*, 2018; Hall *et al.*, 2012b).

1.2. Rocket Germplasm Collection Initiatives

Until 1994, rocket species were considered as underutilized Mediterranean crops. Efforts gathered by the International Plant Genetic Resources Institute (IPGRI) positively influenced the crescent importance of the genetic resources of the rocket species.

The first meeting of the Rocket Genetic Resources Network (Lisbon, Portugal, 1994) has decisively stimulated the establishment of germplasm collections of rocket species used and not used for human consumption (Padulosi, 1995).

The more traditional and since long time cultivation and commercialization of the cultivated annual rocket (*Eruca sativa*) and consequent longer germplasm conservation activities resulted in around 750 accessions so far gathered and conserved throughout worldwide genebanks. Currently, in the worldwide genebanks exist less than 70 *D. tenuifolia* available accessions (Genesys, <https://www.genesys-pgr.org/a/v26DkDz6bWZ>).

The main objectives of the REMIRucula project in the frame of which this work was performed, are: 1) the establishment of a wild rocket germplasm collection at the Instituto Nacional de Investigação Agrária e Veterinária (INIAV); 2) the expansion of production sustainability for this leaf crop, by the selection of downy mildew (*Hyaloperonospora* sp.) resistant genotypes, that could assure, simultaneously, high food quality and environmental protection.

The current management of the germplasm accessions, in particular in what concerns the identification of the conserved genetic variability, and the unequivocal identification of the accessions finds a helpful tool in the use of modern molecular marker technologies. The identification of molecular tools for the identification of both rocket species and cultivars is a main priority for genebanks, to improve the management of the collection accessions and to assist plant breeders and producers.

1.3. Downy Mildew Disease

Cultivation under protected conditions, with increased levels of humidity, promotes the spreading of fungal pathogens (Caruso *et al.*, 2019). Most of the species within the Brassicaceae family are hosts of *Hyaloperonospora* spp., the oomycete causative of the downy mildew disease (Lee *et al.*, 2017).

Downy mildew causes severe symptoms in young seedlings and adult plants, with destructive impact in both, yield and marketable quality of the product (Figure 1.3). Depending on the host genotype and environmental conditions, the downy mildew disease may imperil the total area of production, causing significant economic losses (Coelho *et al.*, 2012).

Lacking unequivocal identification, the pathogen responsible for downy mildew infection of *Diplotaxis tenuifolia* cultivars is currently classified generically as *Hyaloperonospora* sp. (Coelho *et al.*, 2021b).

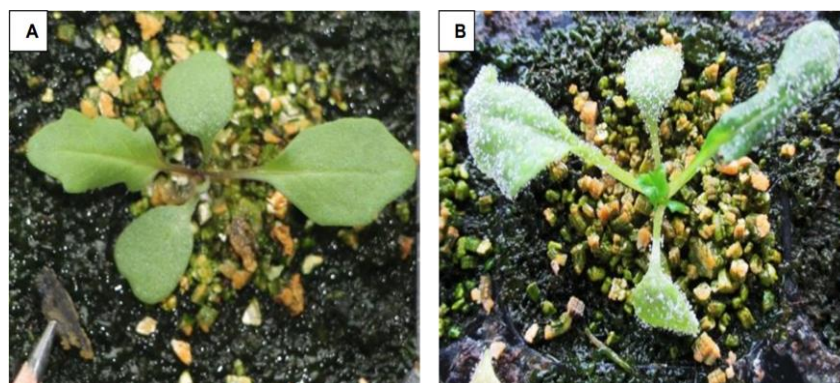


Figure 1.3. (A) Fourteen-days-old wild rocket (*Diplotaxis tenuifolia*) plant before pathogen inoculation; (B) A highly susceptible 21-days-old wild rocket plant infected with *Hyaloperonospora* sp. isolate D5.

The most common method for mitigation of the downy mildew negative impact production is the use of fungicides. However, beyond the increase of the production costs, this practice has severe impacts in the environment and human health (Coelho *et al.*, 2021a).

To avoid the abusive use of control agents, the cultivation of resistant varieties, and the implementation of good agricultural practices, can decrease the emergence and propagation of this pathogen. Breeding and selection of resistant varieties is the uppermost beneficial option towards large-scale production, maintaining lower impacts both, on the environment and on human health (Carlier *et al.*, 2011; Coelho *et al.*, 2021a; Farinhó *et al.*, 2004).

1.4. Objectives

This work has two main research objectives:

1) The assessment of the extant genetic variability within a representative group of accessions of the *Diplotaxis tenuifolia* germplasm collection established by the INIAV, using two random amplification based molecular markers techniques, Random Amplified Polymorphic DNA (RAPD) and Inter Single Sequence Repeats (ISSR).

2) The unequivocal identification of a selected set of *Diplotaxis tenuifolia* accessions by molecular fingerprinting. This objective will be pursued by the application of two species specific molecular marker techniques: Single Sequence Repeats (SSR) and Single Nucleotide Polymorphisms (SNPs). These last markers will be analysed as Cleaved Amplified Polymorphic Sequences (CAPS).

The results of this study are envisaged to complement the phenotypical evaluation of the studied germplasm accessions for their plant/downy mildew (*Hyaloperonospora* sp.) relationships and to create establish the starting point for further molecular identification of the phenotypically identified downy mildew resistance genes.

2. Evaluation of the genetic relationships among a representative set of the wild rocket (*Diplotaxis tenuifolia* L.) germplasm collection established at the INIAV

Studies estimating the genetic diversity in both rocket species were essentially focused on *E. sativa* molecular and phenotypical traits (Egea-Gilabert *et al.*, 2009; Guijarro-Real *et al.*, 2020). Similar studies including both salad rocket species evaluate the phenotypic diversity present in limited groups of *D. tenuifolia* accessions. Phenotypical diversity is commonly the uppermost direct and simple process towards the identification and of further genetic diversity in a set of germplasm accessions (Bozokalfa *et al.*, 2011).

Estimation of genetic relationships through the use of molecular markers in *Diplotaxis tenuifolia* is so far limited to the analysis of species-to-species relationships. Martín & Sánchez-Yélamo (2000) provided insights towards the genetic relationships of 10 *Diplotaxis* species, using ISSR markers, showing close relationship of *D. tenuifolia* with *D. cretacea*, *D. simplex*, *D. viminea* and *D. muralis*, all grouped in a cluster. The close relationship of the five species was once again observed by Eschmann-Grupe *et al.* (2003), using RAPD markers, combined with isozyme analysis, to compare the species relationships of 19 species within the *Diplotaxis* genus. Recently, Taranto *et al.* (2016) reported the use of ISSR markers to separate species using genetic profiles, further elucidating the data collected from the phenotypic analysis. Based on polymorphisms retrieved by the DNA markers, clusters separating *E. sativa* from *D. tenuifolia* and *D. muralis* were constructed.

Despite the scarce number of resources available, studies including *D. tenuifolia* demonstrate wide levels of genetic variability present within this species. Nevertheless, thorough molecular analysis to evaluate the genetic diversity within an established germplasm collection comprising several *Diplotaxis tenuifolia* species is one of the main objectives of this dissertation. Understanding the genetic diversity available in a large set of accessions is the primordial step for the selection and establishment of a plant breeding program.

2.1. Material and Methods

2.1.1. Plant Germplasm Accessions

A set of 100 wild (*D. tenuifolia*) and cultivated rocket (*Eruca* spp.) accessions, evaluated for their host-pathogen interaction with the isolate D5 of the downy mildew causal agent *Hyaloperonospora* sp., was selected for analysis of their genetic relationships (Table 2.1).

Table 2.1. Analyzed *Diplotaxis tenuifolia* (L.) DC. and *Eruca* spp. accessions.

Accession origin	Country of origin	Number of accessions
Australian Grains Genebank (AGG)	Australia	1
Leibniz Institute (IPK)	Germany	4
Breeding companies	FR, IT, NL, SP, USA, UK	83
Local trade	Portugal	10
Harvest missions	Portugal	2
Total		100

2.1.2. Plant Growing Conditions

The seeds were sown in plastic trays (3x3x5cm cells) containing a peat-based compost (Gramoflor GmbH & Co. KG, Vechta, Germany), covered with a layer of vermiculite, and watered by capillary matting. In the INIAV, the trays were placed in a growth chamber with a long-day lighting (19-h light, 5-h dark), 21°C daytime and 19°C night-time temperatures, and 70±10% relative humidity. The photoperiod was provided by LED lamps (LED ECOT814330F 14W 4000K, ROBLAN®) at an intensity of 250 $\mu\text{mol m}^{-2}\text{s}^{-1}$. The leaves of at least 5 plants per accession were collected and transferred in a refrigerated box to the UAIG. In multiple cases, replicates of the accessions were grown at the UAIG, in a glass greenhouse or at the laboratory, under environmental temperature and when needed, additional lightning.

2.1.3. Molecular Characterization of the Accessions

2.1.3.1. DNA Extraction and Evaluation

The genomic DNA extraction was performed as described in Farinhó *et al.* (2004) with minor modifications. One leaf from 5 plants of each accession was removed, washed with tap water, and wiped up with paper. The central nervures were removed using a scalpel and the leaves were ground under liquid nitrogen in a mortar with a pestle. An amount of the obtained fine powder was transferred to a microfuge tube containing 500 μl of extraction buffer (250

mM Tris-HCL, pH 8.0, 25 mM EDTA, 1% SDS) until the final volume reached approximately 750 μ l. Then, RNase A (20 μ g ml⁻¹) was added to the tube, and then transferred to a water bath at 65 °C for 15 minutes. Already at room temperature, 1 volume of phenol: chloroform : isoamyl alcohol (25:24:1) was added to the tube, which after successive inversions for 1 minute was centrifuged at 13,000 rpm for 3 min. The upper phase was transferred to a new tube and extracted with 1 volume of chloroform : isoamyl alcohol (24:1) as described in the previous step. This second extraction was repeated once or twice, until the interphase appeared completely transparent. The final upper phase was transferred to another tube and precipitated with 3 volumes of cold absolute ethanol and kept at -20 °C. After centrifugation at 13,000 rpm the DNA pellet was dried and resuspended in TE_{0.1} (10 mM Tris, 0.1 mM EDTA). The integrity of the extracted DNA and eventual contamination with RNA was assessed by agarose gel (1.4%) electrophoresis. The DNA concentration was determined by spectrophotometry (NanoDrop One, ThermoFisher) after a previous comparison via agarose gel electrophoresis with different amounts of DNA extracted from *Pisum sativum* roots, which do not contain chlorophyll or other pigments, that could bias the spectrophotometry results.

2.1.3.2. RAPD-PCR Analysis

The RAPD-PCR amplifications were performed as described in Elisiário *et al.* (1999) with minor modifications, in 15 μ l final volume reaction mixes containing: 10 ng genomic DNA, 1x gel load reaction buffer, 0.16 mM de dNTPs, 0.6 U de NZY[®]Taq DNA Polymerase, 1.33 μ M of each (single) primer. All RAPD primers used in this study were synthesized by Operon Technologies Inc. (Qiagen).

The PCR reactions were carried out in a Biometra UNO II (Thermoblock, Biotron) thermocycler according to the protocol: initial denaturation at 94 °C for 1 min and 30 sec, followed by 35 cycles of 30 sec at 94 °C, 30 sec at 36 °C and 1 min at 72 °C, and a final cycle of 10 min at 72 °C.

The amplification products are then analyzed by agarose gel (2%) electrophoresis. The gels are then immersed into an ethidium bromide solution, transferred to water for the diffusion of the stain from the gel and photographed under UV-transillumination with a digital camera Canon EOS 1300D. The tested and selected RAPD primers are displayed in Table 2.2.

Table 2.2. Tested and selected RAPD primers*.

Primer	Sequence	Primer	Sequence	Primer	Sequence
OPA02	TGCCGAGCTG	OPAL05	GACTGCGCCA	OPAN05	GGGTGCAGTT
OPA03	AGTCAGCCAC	OPAL07	CCGTCCATCC	OPAN14	AGCCGGGTAA
OPB03	CATCCCCCTG	OPAL12	CCCAGGCTAC	OPAN16	GTGTTCGAGTC
OPB09	TGGGGGACTC	OPAL15	AGGGGACACC	OPAN18	TGTCCTGCGT
OPB20	GGACCCTTAC	OPAL17	CCGCAAGTGT	OPAN20	GAGTCCTCAC
OPD14	CTTCCCCAAG	OPAL19	TCTGCCAGTG	OPAO01	AAGACGACGG
OPD19	CTGGGGACTT	OPAM03	CTTCCCTGTG	OPAO02	AATCCGCTGG
OPE13	CCCGATTCCG	OPAM04	GAGGGACCTC	OPAO08	ACTGGCTCTC
OPF15	CCAGTACTCC	OPAM08	ACCACGAGTG	OPAO10	GACATCGTCC
OPU01	ACGGACGTCA	OPAM13	CACGGCACAA	OPAO12	TCCCGGTCTC
OPAK06	TCACGTCCCT	OPAM14	TGGTTGCGGA	OPAO13	CCCACAGGTG
OPAK12	AGTGTAGCCC	OPAM16	TGGCGGTTTG	OPAO20	GGCTTGCCTG
OPAK14	CTGTCATGCC	OPAM19	CCAGGTCTTC		
OPAK20	TGATGGCGTC	OPAM20	ACCAACCAGG		

*The seven primers selected for further collection analysis are in bold.

2.1.3.3. ISSR Analysis

The ISSR amplifications were performed in the same thermocycler using a protocol identical to the used for RAPD markers amplification, except for the annealing temperatures that varied from primer to primer depending on the estimated specific melting temperature.

The sequences of the tested and selected ISSR primers are registered in Table 2.3. All primers are degenerated and a combination of two primers. As commonly accepted the annotation “R” stands for purines (A and G) and “Y” for pyrimidines (T or C).

Table 2.3 Tested and selected ISSR primers*

Primer	Sequence	T ^a annealing	Primer	Sequence	T ^a annealing
DiploArb_01	(CA) ₈ RG	-	DiploArb_05	(GA) ₈ YG	60 °C
DiploArb_02	(CA) ₈ RY	-	DiploArb_06	(AG) ₈ YT	58 °C
DiploArb_03	(GA) ₈ YT	58 °C	DiploArb_07	(AG) ₈ YC	-
DiploArb_04	(GA) ₈ YC	60 °C	DiploArb_08	(AC) ₈ YT	58 °C

* The five primers selected for further collection analysis are in bold.

2.1.3.4. Data Analysis

Only clear and distinct amplified bands were scored as '1' for presence and '0' for absence in the electrophoresis gels and included in the data binary matrix.

The NTSYS-pc program (Rohlf, 2009) was used for cluster analysis. The genetic similarity between the accessions was reckoned by pairwise comparisons based on the percentage of common fragments using the coefficient DICE (Nei & Li, 1979), according to the following equation: $\text{similarity} = 2N_{ab} / (N_a + N_b)$, where N_{ab} is the number of scored amplification products simultaneously present in accessions 'a' and 'b', N_a is the number of amplification products scored in accession 'a', and N_b is the number of scored fragments in accession 'b'. The unweighted pair-group method with arithmetic averages (UPGMA) was used to calculate the cophenetic matrix used for the dendrogram construction. The cophenetic correlation coefficient (r) was calculated by comparison of the similarity matrix with the UPGMA produced cophenetic matrix.

2.2. Results

2.2.1. Molecular Diversity

The previous selection among 40 RAPD and 8 ISSR primers, suitable for amplification of a higher number of clear polymorphic markers, resulted in the retaining of 7 RAPD primers (OPU01, OPAL07, OPAL17, OPAM13, OPAM20, OPAN05, OPAN16) and 5 ISSR primers (DiploArb_03, DiploArb_04, DiploArb_05, DiploArb_06, DiploArb_08) for further molecular analysis (Tables 2.2 and 2.3).

The molecular analysis of the 100 accessions with these 12 primers allowed the amplification of clear and easy to score 62 RAPD and 48 ISSR markers. The RAPD primers amplified an average number of ~8.9 scored markers per primer, relatively less than the ~9.6 average markers amplified by the ISSR markers (Figures 2.1 and 2.2).

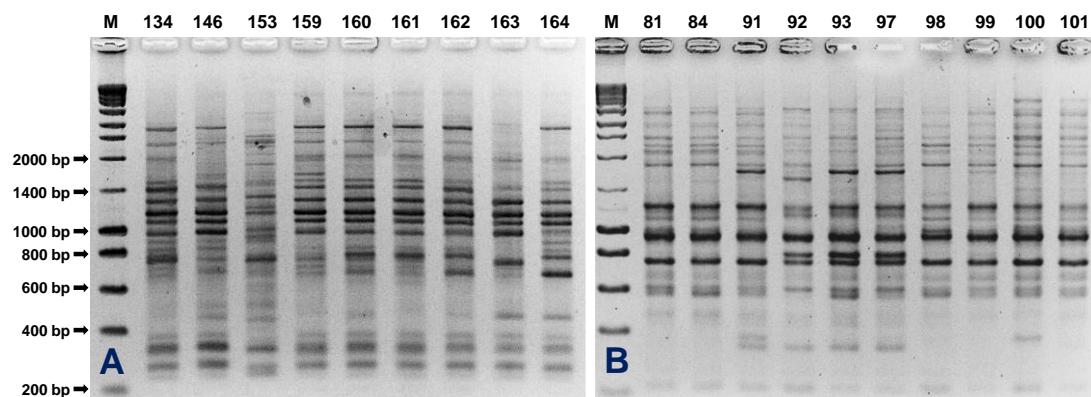


Figure 2.1. RAPD amplification of *Diplotaxis tenuifolia* and *Eruca* spp. accessions with Operon Technologies primers OPU01(A) and OPAN16 (B). Notice the very different amplification pattern of the *Eruca* sp. accession 153.

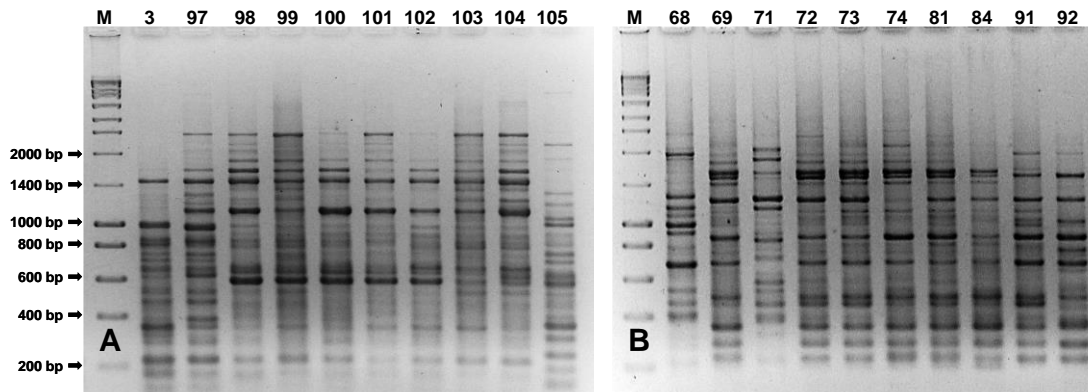


Figure 2.2. ISSR amplification of *Diplotaxis tenuifolia* and *Eruca* spp. accessions with primers DiploArb_04 (A) and DiploArb_03 (B). Notice the very different amplification patterns of the *Eruca* spp. accessions 68, 71 and 105.

The UPGMA analysis of the genetic similarity data resulted in a cophenetic matrix graphically displayed as a dendrogram (Figure 2.3), that exhibits a strong correlation ($r=0.996$) to the initial genetic similarity matrix (Table SM1). The first evident result of these analyses is the presence in the dendrogram of two main clusters: a major cluster (A) that gathers all *Diplotaxis tenuifolia* accessions, and a second major cluster (B) that reunites all *Eruca* spp. accessions.

This main discrimination confirms the rightness of information about the supplied seed accessions, while highlighting the rightness of the obtained host-pathogen interaction results and the accuracy of the performed DNA-markers analyses.

As expected, the lowest similarity value (~ 0.126) was found between a *D. tenuifolia* accession (No. 156) and an *E. sativa* accession (No. 30).

The lowest similarity value among the *Diplotaxis* accessions (0.697) was observed between the accessions 23 and 92. Together with the accession 93, these three accessions showed the lowest similarities with several accessions of the same species. An additional comparative analysis will be soon carried out for a finer morphological and molecular characterization of these three accessions.

The closest similarities among *D. tenuifolia* accessions were found between accession 44 and 42 (~ 0.994) and 44 and 43 (~ 0.994). These three accessions are gathered in a tight cluster that also includes accessions 45 and 46 in a wider cluster (of eleven accessions) that also includes accessions 6, 40 and 41. The accessions 40, 44 and 46 are partially resistant. It is worth of mention that all these cited accessions were provided by the same breeding company.

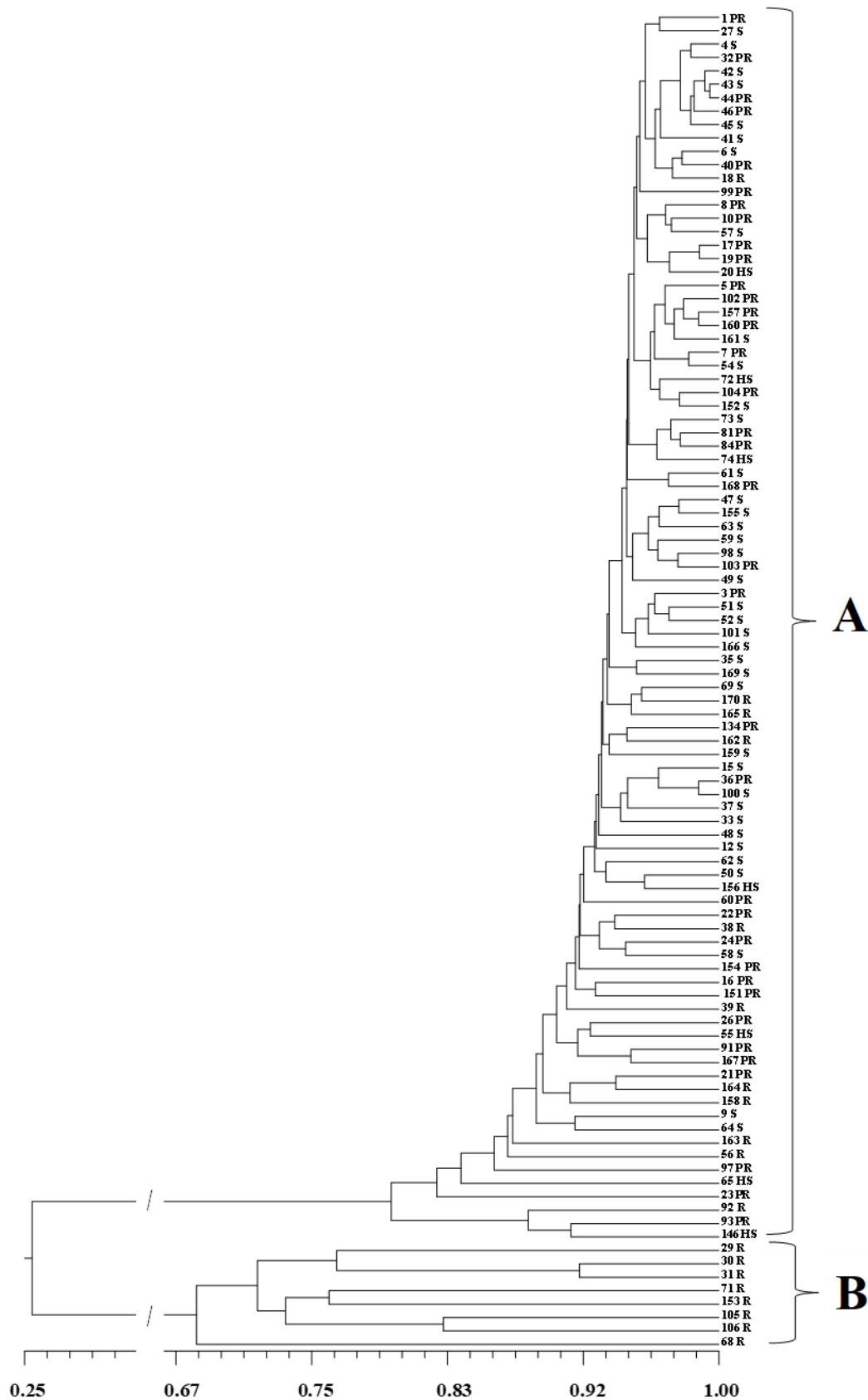


Figure 2.3. Dendrogram resulting from the UPGMA analysis of the genetic relationships assessed by RAPD and ISSR markers and using the coefficient of similarity DICE, of the 100 *Diplotaxis tenuifolia* and *Eruca* spp. accessions evaluated for their interaction with the pathogenic *Hyaloperonospora* sp. isolate D5 (**HS** – Highly Susceptible; **S** – Susceptible; **PR** – Partially Resistant; **R** – Resistant;).

Another example of closer clustering of accessions provided by the same donor is that of accessions 8, 10, 17 and 19 which, contrarily to other two accessions (20 and 57) in the same cluster, are partially resistant to downy mildew.

In the above two cases, aiming at the identification of the downy mildew resistance genes, we can establish as a working hypothesis the eventual sharing of the same resistance *locus* among the partially resistant accessions provided by the same company.

The *Eruca* spp. accessions exhibit a wide molecular diversity with genetic similarity values that vary from ~0.914 (accessions 30 and 31) to ~0.605 (accessions 31 and 68). The low genetic similarity among some *Eruca* accessions, needs to be further scrutinized in more detail. However, this is out of the scope of the present study.

In general terms, the molecular analyses revealed a wide genetic diversity within both kinds (*Diplotaxis* and *Eruca*) of accessions, and the absence of repeated genotypes. The genetic diversity was accompanied by a wide diversity of interactions with the *Hyaloperonospora* spp. isolate D5, which showed to be non-pathogenic to the *Eruca* spp. accessions.

2.3. Discussion

The RAPD and ISSR markers analyses have evidenced a wide genetic variability among the germplasm accessions of *D. tenuifolia* and *Eruca* spp. tested for plant-pathogen (*Hyaloperonospora* sp.) interaction.

The main obvious result is the clear separation between the two species (DICE coefficient ~0.25) and the closer genetic relationships within each species: DICE coefficient ~0.8 for *D. tenuifolia* and ~0.68 for the accessions of *Eruca* spp.

The accessions 42, 43 and 44 were found the most similar and very closely gathered to accessions 45 and 46, and slightly more apart from accessions 40 and 41. The fact that these accessions were provided by the same company attest for the accuracy of the performed molecular analyses. Although most of these accessions exhibited a susceptible phenotypic interaction with the *Hyaloperonospora* sp. isolate D5, the accessions 40, 44 and 46 displayed a partially resistant phenotype.

The accessions 8, 10, 17 and 19 constitute a second case of close clustered partially resistant accessions that are provided by the same breeding company.

These two cases of close clustering of partially resistant accessions provided by the same donor suggests that these partial resistances could be conferred by the same genetic features, a

hypothesis that will be taken into consideration in our further attempt of mapping the responsible genes.

The lowest similarity value (0.697) among the *Diplotaxis* accessions was observed between accessions 23 and 92. Together with the accession 93, the three accessions showed the lowest similarity with several *Diplotaxis* accessions. An additional comparative analysis focusing on these three accessions needs to be carried out for a finer morphological and molecular characterization, in the last case using different types of Sequence Tagged Site (STS) markers and genotyping by sequencing.

The *Eruca* spp. accessions exhibit a wide diversity with genetic similarity values that vary from ~0.914 (accessions 30 vs. 31) to ~0.605 (accessions 31 vs. 68). The low genetic similarity among some *Eruca* spp. accessions, particularly the low similarity showed by the accession 68 to the remaining accessions of the same genus, needs to be further scrutinized in more detail.

3. Unequivocal identification of germplasm accessions of wild rocket (*Diplotaxis tenuifolia* (L.) DC.) by specific SSR and single SNP markers identified by previous genome NGS

The genomic and transcriptomic data of *Diplotaxis tenuifolia* available at the NCBI database that could be used for the present work consisted of 83 gene and partial gene and Internal Transcribed Spacer (ITS) sequences, 500 microsatellite loci, the data of one RNAseq project, and of three genome sequencing projects with accession codes: PRJEB28010, PRJNA380551, and PRJNA624903.

The last genome project was implemented and uploaded to the NCBI by the Laboratory of Genomics and Genetic Improvement, MED, FCT, Universidade do Algarve in the frame of the project REMIRucula (Figure 3.1).

Termed “[UALgDiploT.01](#)” this genome assembly (scaffold), comprehends a total sequence length of ~ 424 Mbp assembled in ~170,000 contigs (N50 = 4,686; L50 = 23,088).

UALgDiploT.01

Organism name: [Diplotaxis tenuifolia \(eudicots\)](#)
Isolate: DT550
Sex: hermaphrodite
BioSample: [SAMN14589541](#)
BioProject: [PRJNA624903](#)
Submitter: Remirucula
Date: 2020/09/30
Assembly level: Scaffold
Genome representation: full
RefSeq category: representative genome
GenBank assembly accession: GCA_014822095.1 (latest)
RefSeq assembly accession: n/a
RefSeq assembly and GenBank assembly identical: n/a
WGS Project: [JABAYD01](#)
Assembly method: CLC Genomics Workbench v. 12.0.3
Expected final version: yes
Genome coverage: 60.0x
Sequencing technology: Illumina HiSeq

IDs: 8212281 [UID] 22406978 [GenBank]

Figure 3.1. *Diplotaxis tenuifolia* whole genome sequence assembly report.

The above mentioned 500 microsatellite loci (accessions MT317453 to MT317952) were retrieved from the contigs of this genome assembly project and subsequently imported to the NCBI database. A set of 500 SNP loci was also retrieved from the “UALgDiploT.01” whole-genome sequencing project data, and uploaded to the NCBI (GenBank OM144979 to OM145478).

The SSR and SNP *loci* used for the genetic analysis of the *D. tenuifolia*, and *E. sativa* accessions were selected among these two sets of 500 genomic sequences.

To the best of our knowledge, no prior studies using SSR, and SNP molecular markers have been performed in *Diplotaxis tenuifolia*.

3.1. Materials and Methods

3.1.1. Plant Germplasm Accessions

A set of 87 *D. tenuifolia* and 3 *E. sativa* accessions of the germplasm collection (INIAV), previously tested for their interaction with the isolate D5 of *Hyaloperonospora* sp., the causing agent of downy mildew disease (Coelho et al., 2022), were selected for unequivocal molecular characterization and identification by specific SSR and SNP patterns.

3.1.2. Plant Growing Conditions

The seeds were quickly washed with tap water and common detergent, and immersed during 1 min into a disinfection containing 0.5% of SDS and 10% of commonly commercialized bleach. After thorough washing with distilled water until the total removal of the disinfection solution, the seeds were transferred to petri dishes containing three layers of filter paper saturated with tap water. During the following days the consecutively germinated seedlings were transferred in groups of 3 to small 9 cm diameter pots containing a mix of 50% peat and 50% perlite. Five pots per accession were transferred to a glass greenhouse, and 3 weeks later only one well succeeded plant was left to grow per pot.

3.1.3. DNA Extraction Accessions Analysis

The extraction, quantification, quality evaluation and amplifiability of the genomic DNA samples extracted for molecular identification of the analyzed accessions was carried out as described in Chapter 2.

3.1.4. Primers Design and Synthesis

All primers were designed, their parameters calculated and their eventual self- or pair-annealing assessed, using the FastPCR 6.7 Software (Kalendar *et al.*, 2017). Common, non-labeled, primers were synthesized by the company Eurofins Genomics (Germany). The fluorescent labeled primers were ordered from the company STABVida.

3.1.5. Single Sequence Repeats (SSR) Marker Analysis

The amplification of the SSR (microsatellite) markers was performed in 30 µl reactions, starting with an initial denaturation at 94°C for 1 min and 30 sec, followed by 35 cycles of 30 sec denaturation at 94°C; 30 sec annealing, at different temperatures depending on the specific primer pair, and 1 min extension at 72°C, followed by a period of final extension at 72°C for 10 min.

The PCR products were analyzed in 3% agarose gel electrophoresis and the better amplified markers were selected for further, more accurate, analysis among the studied accessions. The amplifications were repeated with the same pair of primers with the forward primer labelled with a fluorochrome. Half of the amount of the amplified products were analyzed by agarose gel electrophoresis and the second half of the approved amplified samples was sent to the company STABVida for fragment analysis by capillary polyacrylamide gel electrophoresis.

The fragment analysis of the amplified by fluorescent SSR markers was performed in a 3730XL Genetic Analyzer platform using GeneScan™ 500 LIZ™ as the dye size standard. The resulting data were analyzed using the Peak Scanner™ Software v. 1.0 (Applied Biosystems, USA).

3.1.6. Single Nucleotide Polymorphisms (SNP) Markers Analysis

Five hundred SNP loci were identified using the Geneious Prime v.2021.2.5 software. Seven nucleotide sequences, 3 nucleotides from each side of the identified SNP, were analyzed by the NEBcutter V2.0 (<http://nc2.neb.com/NEBcutter2/>) software (Vincze *et al.*, 2003) for identification of restriction enzymes that differentially recognized the alternative SNP alleles.

Nineteen SNP markers harboring a TaqI restriction site encompassing the polymorphic nucleotide were selected for further work and amplified using the same protocol used for SSR markers. Fifteen microliters of the amplified products were analyzed by 3% agarose gel electrophoresis. The remaining 15 µl of well amplified samples were then cut with the TaqI restriction enzyme, and the samples analyzed as CAPS (Cleaved Amplified Polymorphic Sequences) markers in 3% agarose gels.

3.1.7. Data Analysis

The NTSYS-pc program (Rohlf, 2009) was used for cluster analysis. The genetic similarity between the accessions was reckoned using the coefficient DICE (Nei & Li, 1979) by pairwise comparisons based on the percentage of common fragments, according to the following equation: $\text{similarity} = 2N_{ab} / (N_a + N_b)$, where N_{ab} is the number of scored amplification products simultaneously present in accessions 'a' and 'b', N_a is the number of amplification products scored in accession 'a', and N_b is the number of scored fragments in accession 'b'. The unweighted pair-group method with arithmetic averages (UPGMA) was used to calculate the cophenetic matrix used for dendrogram construction. The cophenetic correlation coefficient (r)

was calculated by comparison of the similarity matrix with the UPGMA produced cophenetic matrix graphically represented as a dendrogram.

3.2. Results

3.2.1. SSR Markers Analysis

Five hundred SSR loci were previously identified among the new NGS-established genome assembly contigs and uploaded to the NCBI database (accessions: MT317952 to MT317453). Primer pairs were designed for 20 (Table SM2 and SM3) out of the 500 SSR loci and, after preliminary amplification tests (Figure 3.2), 5 out of the 20 SSR loci were selected for the unequivocal identification of 90 (87 wild rocket and 3 garden rocket) accessions of the rocket germplasm collection (Table 3.1).

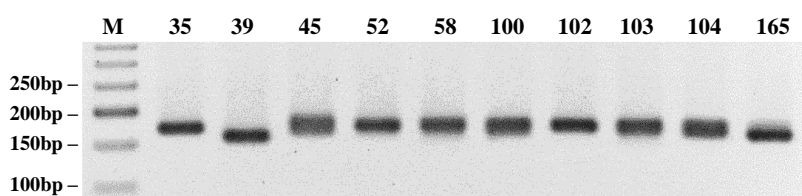


Figure 3.2. Agarose gel electrophoresis of the preliminary amplification of the SSR locus MT317527 among 10 *D. tenuifolia* accessions.

The forward primer of each pair of the 5 SSR loci retained for further analysis was labelled with a fluorophore (Table 3.1) and, after a previous quality test in agarose gels, the amplified products were sent for fragment analysis by capillary polyacrylamide gel electrophoresis to the company STABVida (Lisbon, Portugal).

Table 3.1. SSR primers used for accessions identification.

Locus		Primers	T (C°)**	Fluorophore
MT317577	FW	CGGATAAACATATCCGCTT*	57	Atto 565
	RV	GGTTAACATCACTAACGGT		
MT317610	FW	GTCTATTGATCTGATGCCG*	57	Atto 550
	RV	GCGAGAACCGATCTATGA		
MT317823	FW	CCAACATAGAAAGGTGCG*	57	Atto 550
	RV	GAGTTCCTCCAAAAGCTG		
MT317527	FW	ACTTTGACGAAACGAAGC*	58	HEX
	RV	CTCAGAACCAAAGAGAAGC		
MT317537	FW	CAGCTTTCTGTTAGTGGTC*	59	HEX
	RV	CCAACCAAAAACGTCAGA		

* Fluorophore labelled; ** Annealing temperature

The results of the fragment analysis were visualized graphically using the Peak Scanner™ Software v. 1.0 and the right alleles peaks were identified among the secondary peaks that result from the “band stuttering” (Cardoso et al., 2017), commonly associated to the SSR technique (Figure 3.3).

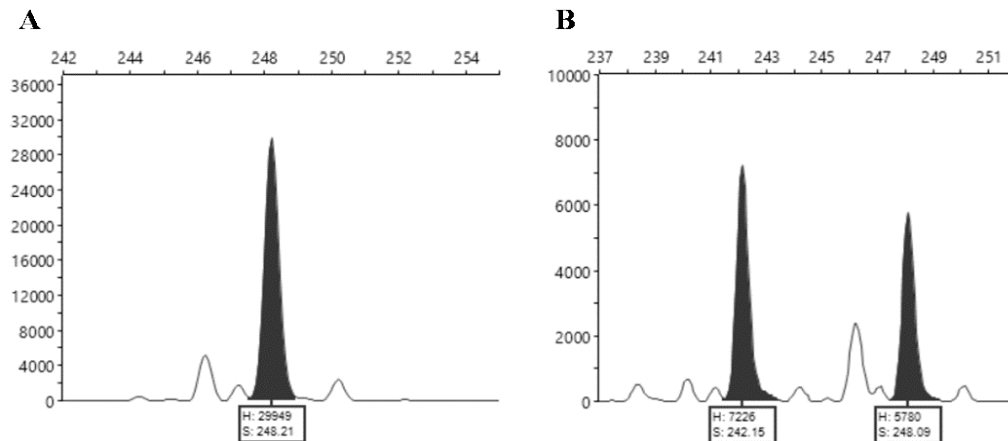


Figure 3.3. Capillary polyacrylamide gel electropherograms of the locus MT317577 (A) Homozygous pattern of accession 15; (B) Heterozygous pattern of accession 37. **X axis** - Fragment size (bp); **Y axis** - Relative Fluorescence Unit (RFU). Notice the associated smaller “stutter” peaks.

The fragment analysis allowed the identification of 70 different alleles for the 5 SSR loci that permitted to carry out the first assessment of the genetic similarities and genetic relationships between the analyzed accessions (Table SM4) and the establishment of specific SSR markers patterns (Table SM5) for identification of most of the *D. tenuifolia* accessions. Nevertheless, as it can be observed in the respective dendrogram (Figure 3.4A), 1 cluster of 8 accessions, 3 clusters of 3 accessions and 8 clusters of 2 accessions of this species shared identical SSR patterns, an issue that will be discussed below.

Although the primers were designed for the specific amplification of SSR loci identified in *D. tenuifolia* genome, the *E. sativa* accessions amplified for 3 out of the 5 microsatellite markers, that discriminated among the 3 accessions of this species and contributed for their drastic discrimination from the wild rocket accessions.

The SSR marker MT317577 did not amplify in the 3 *E. sativa* samples, and also not in six *D. tenuifolia* accessions: 56, 91, 92, 97, 134 and 146 (Table SM5). This is an interesting result, since, except for the samples 91 and 134, the other four accessions were previously found by RAPD and ISSR analyses to be relatively distanced genetically from the bulk of the accessions

of the same species and closer, although very apart, to the then analyzed *E. sativa* accessions (Coelho et al., 2022).

3.2.2. SNP/CAPS Markers Analysis

The *D. tenuifolia* genome assembly “UALgDiploT.01” was additionally mined using the software Geneious Prime v.2021.2.5, for the identification of sequences harbouring single nucleotide polymorphisms (SNP) (Table SM6).

The genomic sequences, of approximately 500 nucleotides, surrounding the first identified 500 SNPs (GenBank OM144979 to OM145478) were retrieved for further analysis (Table SM6). Then, the identified SNP loci were analyzed using the software NEBcutter V2.0 for identification of restriction enzymes that differently digested the alternative SNP alleles, allowing these markers to be assessed as CAPS (Cleaved Amplified Polymorphic Sequences) markers. Reflecting the method of analysis, the SNP markers are further on referred to as SNP/CAPS.

Nineteen out of the selected 500 SNP loci were identified as harbouring the polymorphic nucleotide within the sequence 5'-TCGA-3' differentially recognized by the restriction enzyme TaqI, allowing the use of this enzyme for allele discrimination within these loci. Primers were designed for 9 loci (Table 3.2; Table SM3) which, after amplification and analysis by agarose gel electrophoresis (Figure 3.5), showed to be suitable for molecular discrimination and were used to assess the 90 germplasm accessions.

The SNP marker loci, the specifically designed primers, amplification annealing conditions, and the respective resulting fragments after TaqI digestion are displayed in Table 3.2. The results of the molecular analysis of the accessions by SNP/CAPS markers can be fully consulted in Table SM5 and are summarized in Table 3.3.

The three *Eruca* spp. accessions were amplified uniquely for the locus OM145023 which, as for the large majority of the *Diplotaxis* accessions showed to be homozygous for the restriction by TaqI. Nevertheless, this circumstance did not hamper the *Eruca* spp. accessions (29, 68, 105) to cluster clearly apart from the *Diplotaxis* accessions (Figure 3.4B).

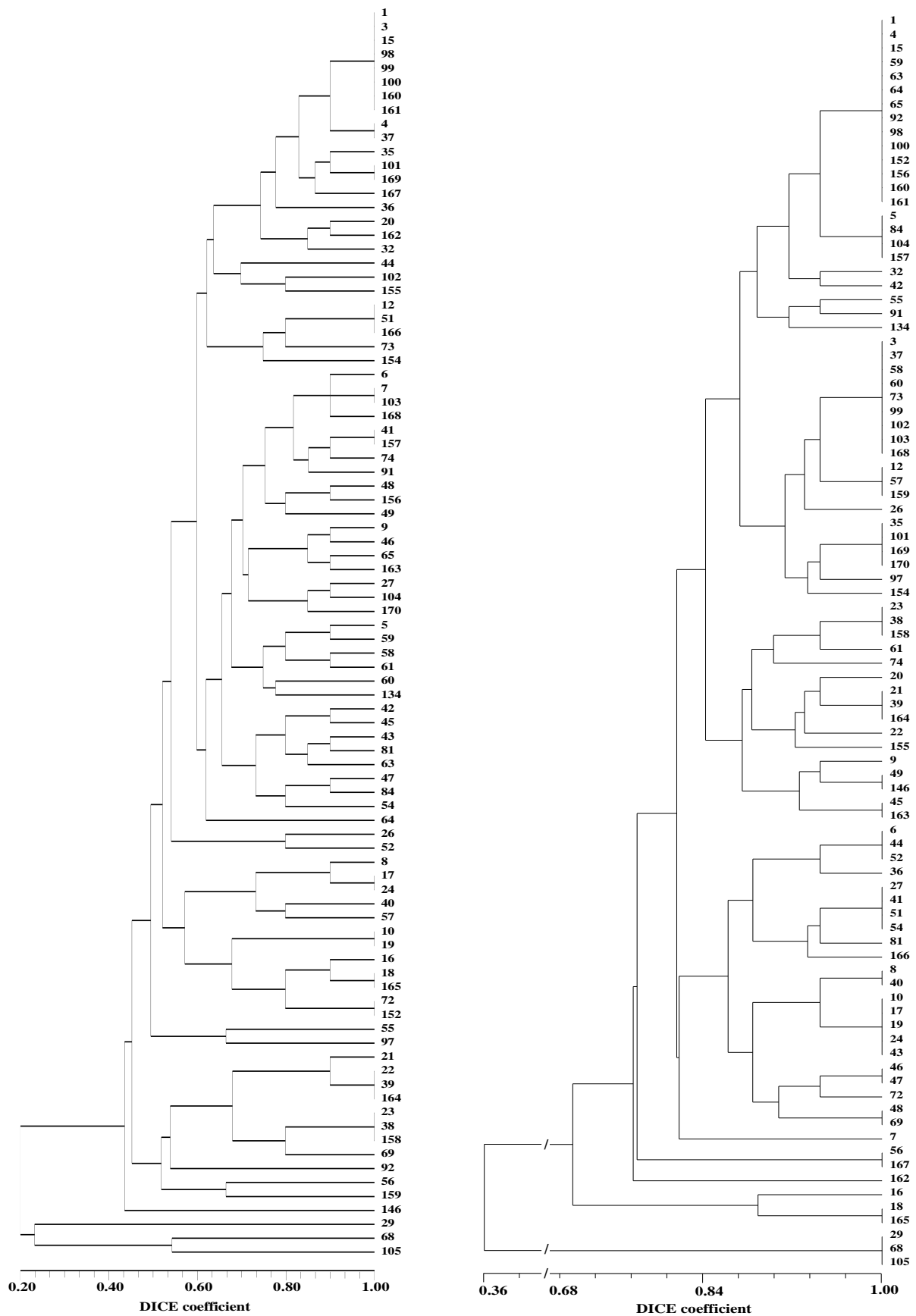


Figure 3.4. (A) Dendrogram depicting the genetic relationships among the 90 germplasm accessions established based on SSR markers. The *E. sativa* accessions (Nos. 29, 68 and 105 at the bottom) are clustered clearly apart from all other (*D. tenuifolia*) accessions. (B) Dendrogram displaying the genetic relationships among the 90 germplasm accessions established based on SNP/CAPS markers.

Table 3.2. SNP-CAPS primers used for accessions identification.

Locus		Primers	T° (°C)*	SNP	Restriction Fragments
OM145023	FW	GTCCCATGATTAGATATGGT	58	G/T	143/80
	RV	AGTCTTTAAGTGACAAGCG			
OM145040	FW	GCGTTTGAATGGTTTCTG	56	T/C	205/75
	RV	CTTGCTCATCCGTCAA			
OM145068	FW	CATCTGTGCCTGATCTC	58	G/T	191/54
	RV	CATGTCAAGTCGTGATTCA			
OM145116	FW	GAGCCTTGTAACAGATCC	59	T/G	194/94
	RV	CTGGGTTTGTGACTGTTG			
OM145191	FW	TGGCATCCTACTCTTTCTC	60	C/T	170/61
	RV	CCGTGTGTTTCATGATATGG			
OM145198	FW	TCTACGATGCCAAGAGTT	58	T/C	144/94
	RV	TGTACTGGAAAGGTATCCC			
OM145248	FW	CTAATGAGTTCGTCGTTTCG	60	A/G	216/108
	RV	AGCTTATCTCTTCCTGGTG			
OM145268	FW	CGAAAGAGATGCTATCGG	59	C/T	167/76
	RV	TACGTCCTCAGTTCTAACC			
OM145426	FW	GGCTGGATATGTGTCAGA	59	A/C	155/127
	RV	GTCATGTTCTTCTGGAAC			

* Annealing temperature

The analyzed SNP/CAPS markers were not enough for the full discrimination of the *D. tenuifolia* accessions, since only 20 accessions were characterized by individual patterns. The remaining 67 accessions remained gathered in clusters of full identity (DICE coefficient = 1.000) comprehending from 2 or 3 accessions to larger groups of 9 or 14 accessions (Table SM7; Figure 3.4B).

3.2.3. Combined results of SSR and SNP-CAPS markers

The combination of the results of the analyses performed with both types of markers allowed the analysis of the genetic similarities and genetic relationships between the 90 accessions to be refined (Table SM8; Figure 3.6).

Individual specific molecular fingerprints were identified for almost all accessions, except for 1 trio and 6 pairs of wild rocket accessions that exhibited full similarity, an issue that

will be discussed below. The immediately below high genetic similarity (0.964) was exhibited by 17 pairs of accessions of this species.

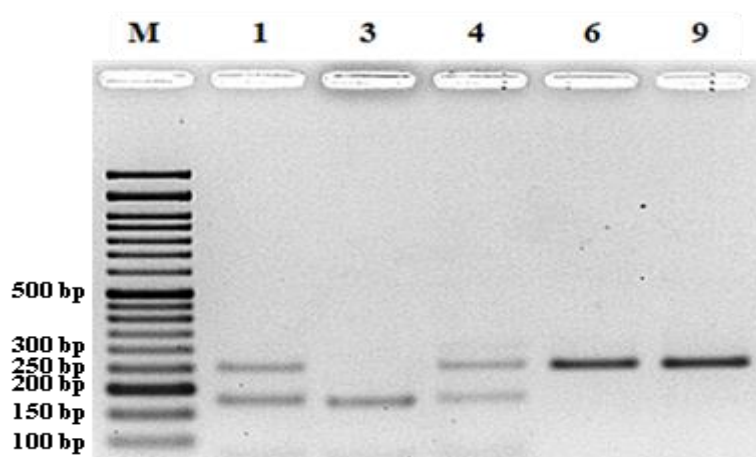


Figure 3.5. Molecular patterns of 5 accessions for the OM145268 locus. Accessions 1 and 4 (Heterozygous Y/N); Accession 3 (Homozygous Y/Y); Accessions 6 and 9 (Homozygous N/N); **M** – DNA ladder VI (NZYTech). **Y** - one allele restricted; **N** – one allele not restricted.

The lowest genetic similarity reckoned between *D. tenuifolia* accessions were 0.464, between the accession 58 and accessions 18 and 165, and 0.481 between the accessions 56 and accessions 16, 18 and 165. However, the calculated genetic similarity values between two accessions of this species were over 0.6 in ~96% of the cases. As expected, the lowest coefficient of similarity (0.111) was registered between accessions of both species.

The established molecular fingerprints (Table SM5) allowed the clear discrimination between the two rocket species and between the accessions within each species, except for the below discussed cases of identical molecular patterns of some *D. tenuifolia* accessions.

Table 3.3. Main results of the SNP/CAPS analyses

Locus	Homozygous restricted	Heterozygous	Homozygous unrestricted
OM145023	86	4	0
OM145040*	87	0	0
OM145068*	58	29	0
OM145116*	66	18	3
OM145191*	49	34	4
OM145198*	55	29	3
OM145248*	0	5	82
OM145268*	7	44	36
OM145426*	0	14	73

* The 3 *Eruca* sp. accessions did not amplify.

The tested and validated set of SSR and SNP/CAPS markers constitute a useful tool for the unequivocal identification of all present and future accessions of the *D. tenuifolia* germplasm collection and for multiple other intraspecific and interspecific discrimination studies.

3.3. Discussion

In a previous assessment of the genetic relationships among a large set of accessions of the same germplasm collection using RAPD and ISSR markers, only in one case the genetic similarity (DICE coefficient) between two *Diploaxis* accessions fallen below 0.7, and in 95% of the cases this parameter was over 0.8.

This level of genetic similarity agrees with the obtained in our lab during the last decades whenever randomly amplified DNA markers (RAPD, ISSR, AFLP) were used to the assess the genetic similarity and genetic relationships within multiple plant species, e.g., pear (Monte-Corvo *et al.*, 2000), apple (Goulão *et al.*, 2001), fig (Cabrita *et al.*, 2001) or common beans (Svetleva *et al.*, 2006).

In the present study, as expected, the calculated genetic similarity values based on SSR markers were relatively low, a consequence of their hyper-polymorphism, a feature that make these markers highly useful for molecular identification of individuals and establishment of genetic relationships, but not the most suitable markers for quantification of the genetic similarities, since these can be zero between two individuals of the same species, a result that is, obviously, absurd. This circumstance is also clear in the present study (Table SM4) as the reckoned genetic similarity between accessions often fallen drastically below 0.8, reaching values as low as 0.3 (e.g., accessions 12 and 16 of *D. tenuifolia*).

On the other hand, the intraspecific (*D. tenuifolia*) genetic similarity values calculated based on of the SNP/CAPS markers were clearly higher, with the lowest value (0.556) registered in 8 cases (Table SM7).

Apparently, in accordance with the results found in our lab using randomly amplified markers (Coelho *et al.*, 2022), the present results of SNP/CAPS markers analysis are, in fact, a consequence of the low number of analyzed loci. This situation could be improved by the assessment of a larger number of SNP loci or by complementation with results of other molecular markers analyses. We chose the second option and the combination of the results of the SNP/CAPS and SSR markers analyses allowed the identification of specific molecular fingerprints for most of the accessions (Table SM5; Figure 3.6).

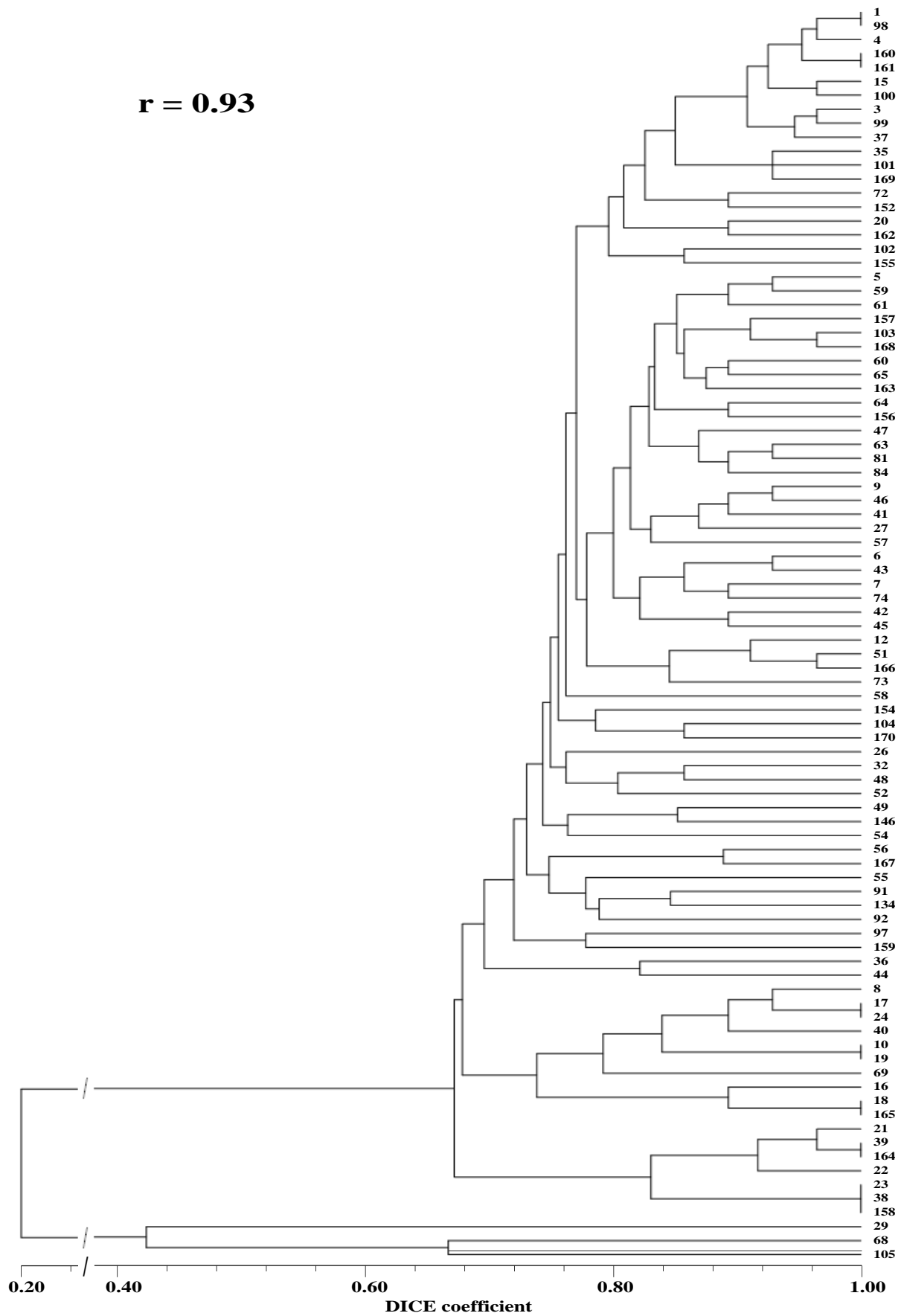


Figure 3.6. Dendrogram depicting the genetic relationships among the 90 germplasm accessions established based on SSR and SNP/CAPS markers. The *Eruca* spp. accessions (Nos. 29, 68 and 105 at the bottom), used as outgroup, are clustered clearly apart from all the other (*D. tenuifolia*) accessions.

The 7 cases of non-discrimination among *D. tenuifolia* accessions deserve a detailed analysis and discussion, as the identification of specific fingerprints for all accessions of the germplasm collection is one of the main objectives of this work.

The accession 1 was provided in 2015 identified as a commercial variety by a wild rocket producer, while the accession 98 was provided in 2019 as the same cultivar by the breeding company.

The accessions 160 and 161 appeared as identical in our analysis. However, they were supposed to correspond to two different cultivars provided by two different companies. For that reason, an accurate comparative analysis of multiple phenotypic traits of these two accessions will be performed, new samples will be requested from the original providers and new molecular analysis will be carried out. Then the needed correction will be introduced in the germplasm collection data.

The accessions 17 and 24 were provided in different years by the same donor, under the same name. These accessions were assumed as an internal control of the performed analyses.

The accessions 10 and 19 were provided, respectively, by a wild rocket producer and a breeding company as the same, relatively resistant to downy mildew cultivar.

The accessions 18 and 165 were provided by two different wild rocket producers, one in Portugal and the other in USA, as the same downy mildew resistant cultivar .

The accession 39 was registered (in 2019) in the germplasm collection as a trial sample of a new cultivar provided by a breeding company. The accession 164 was received later (in 2020) identified as a commercially available downy mildew resistant cultivar. Both accessions are among the very few that, in our studies (Coelho *et al.*, 2022) have exhibited strong resistance to the used *Hyaloperonospora* isolate.

The accessions 23, 38 and 158 were provided by three different donors. The first from a wild rocket leaf producer. The second as a new cultivar under assay (identified by a company code). The third accession is a very well identified, highly resistant to downy mildew, cultivar. In fact, the three accessions were also among the very few *D. tenuifolia* accessions that exhibited downy mildew resistance in our previous studies (Coelho *et al.*, 2022). This relative rare phenotypic trait shared by the three accessions, also contributes to confirm these three accessions as the same cultivar, despite the different providers.

Except for the case of the molecular identity of samples 160 and 161, that need further explanation, the other cases of accessions that exhibit the same molecular pattern reinforces our assumption that the set of molecular markers used in this study can be further utilized for unequivocal identification and register of all accessions of the germplasm collection.

4. Further prospects

The work carried in the frame of this dissertation opens the way to a further line of study: the genome mapping and further identification of the genetic features that confer downy mildew resistance to *D. tenuifolia*.

Based on the previous experience of the Laboratory of Genomics and Genetic Improvement (UAlg), in collaboration with the INIAV and the Instituto Superior de Agronomia (ISA) in the genome localization of a downy mildew resistance gene in *Brassica oleracea* (Carlier *et al.*, 2011; Carlier *et al.*, 2012; Carlier *et al.*, 2016; Farinhó *et al.*, 2004; Farinhó *et al.*, 2007), the first crosses between highly resistant and highly susceptible wild rocket accessions have been already performed to obtain the necessary mapping populations.

The F1 plants have been already characterized for their interaction with the *Hyaloperonospora* sp. isolate D5 and self-pollinated to generate the F2 progenies. Since the accessions selected for initial progenitors are relatively heterozygous, both generations (F1 and F2) can be used for the identification of markers linked to the molecular features, most likely QTLs, that determine the exhibited resistances.

5. References

- Avato, P., & Argentieri, M. P. (2015). Brassicaceae: a rich source of health improving phytochemicals. *Phytochemistry Reviews*, 14(6), 1019–1033. <https://doi.org/10.1007/s11101-015-9414-4>
- Bell, L., & Wagstaff, C. (2014). Glucosinolates, Myrosinase Hydrolysis Products, and Flavonols Found in Rocket (*Eruca sativa* and *Diplotaxis tenuifolia*). *Journal of Agricultural and Food Chemistry*, 62(20), 4481–4492. <https://doi.org/10.1021/jf501096x>
- Bell, L., & Wagstaff, C. (2019). Rocket science: A review of phytochemical & health-related research in *Eruca* & *Diplotaxis* species. *Food Chemistry: X*, 1, 100002. <https://doi.org/10.1016/j.fochx.2018.100002>
- Bell, L., Lignou, S., & Wagstaff, C. (2020). High Glucosinolate Content in Rocket Leaves (*Diplotaxis tenuifolia* and *Eruca sativa*) after Multiple Harvests Is Associated with Increased Bitterness, Pungency, and Reduced Consumer Liking. *Foods*, 9(12), 1799. <https://doi.org/10.3390/foods9121799>
- Bondonno, C. P., Liu, A. H., Croft, K. D., Ward, N. C., Puddey, I. B., Woodman, R. J., & Hodgson, J. M. (2015). Short-Term Effects of a High Nitrate Diet on Nitrate Metabolism in Healthy Individuals. *Nutrients*, 7(3), 1906–1915. <https://doi.org/10.3390/nu7031906>
- Bozokalfa, M. K., Eşiyok, D., İlbi, H., Kavak S., & Aşçıoğlu, T. K. (2011). Evaluation of phenotypic diversity and geographical variation of cultivated (*Eruca sativa* L.) and wild (*Diplotaxis tenuifolia* L.) rocket plant. *Plant Genetic Resources*, 9(4), 454-463. <https://doi.org/10.1017/S1479262111000657>
- Cabrita, L. F., Aksoy, U., Hepaksoy, S., & Leitão, J. M. (2001). Suitability of isozyme, RAPD and AFLP markers to assess genetic differences and relatedness among fig (*Ficus carica* L.) clones. *Scientia horticultrae*, 87(4), 261-273.
- Cardoso, A., Pereira, R., Fonseca, M., & Leitão, J. (2017). A microsatellite sequence in the fifth intron provides a broad-spectrum SSR marker for multiple alleles of the *er1/PsMLO1* powdery mildew resistance gene in *Pisum sativum* L. *Molecular Breeding*, 37(7), 1-5. <https://doi.org/10.1007/s11032-017-0685-x>
- Carrier, J. D., Alabaça, C. S., Sousa, N. H., Coelho, P. S., Monteiro, A. A., Paterson, A. H., & Leitão, J. M. (2011). Physical mapping in a triplicated genome: mapping the downy mildew resistance locus Pp523 in *Brassica oleracea* L. *G3: Genes/ Genomes/ Genetics*, 1(7), 593-601. <https://doi.org/10.1534/g3.111.001099>
- Carrier, J. D., Alabaça, C. A., Coelho, P. S., Monteiro, A. A., & Leitão, J. M. (2012). The downy mildew resistance locus Pp523 is located on chromosome C8 of *Brassica oleracea* L. *Plant breeding*, 131(1), 170-175. <https://doi.org/10.1111/j.1439-0523.2011.01904.x>
- Carrier, J., Alabaça, C., Rodrigues, C., & Leitão, J. (2012, November). On the Way to the Identification of a Downy Mildew Resistance Gene in *Brassica oleracea* L. In *VI International Symposium on Brassicas and XVIII Crucifer Genetics Workshop* 1005 (pp. 233-237). <https://doi.org/10.17660/ActaHortic.2013.1005.25>
- Caruso, G., Parrella, G., Giorgini, M., & Nicoletti, R. (2018). Crop Systems, Quality and Protection of *Diplotaxis tenuifolia*. *Agriculture*, 8(4), 55. <https://doi.org/10.3390/agriculture8040055>
- Caruso, G., De Pascale, S., Nicoletti, R., Cozzolino, E., & Roupael, Y. (2019). Productivity, nutritional and functional qualities of perennial wall-rocket: Effects of pre-harvest factors. *Folia Horticulturae*, 31(1), 71-80. <https://doi.org/10.2478/fhort-2019-0004>
- Cavaiuolo, M., & Ferrante, A. (2014). Nitrates and Glucosinolates as Strong Determinants of the Nutritional Quality in Rocket Leafy Salads. *Nutrients*, 6(4), 1519–1538. <https://doi.org/10.3390/nu6041519>

- Coelho, P. S., Vicente, J. G., Monteiro, A. A., & Holub, E. B. (2012). Pathotypic diversity of *Hyaloperonospora brassicae* collected from *Brassica oleracea*. *European Journal of Plant Pathology*, 134(4), 763-771. <https://doi.org/10.1007/s10658-012-0052-z>
- Coelho, P. S., Pereira, A. L., Carranca, C., Scotti, P., Lopes, V. R., Reis, J., & Leitão, J. M. (2021). Míldio na rúcula selvagem – O porquê de investigar esta doença. *Vida Rural*, 44–50. <https://doi.org/10.13140/RG.2.2.22412.23686>
- Coelho, P. S., Pereira, A. L., Carranca, C., Scotti, P., Gaspar, C., Lopes, V., & Leitão, J. (2021b). O Míldio é um problema grave na produção de rúcula selvagem (*Diplotaxis tenuifolia* L.). *Voz do Campo*, 40–42. <https://doi.org/10.13140/RG.2.2.32868.50563>
- Coelho, P. S., Reis, J. M., Pereira, A. L., Vairinhos, A., Lopes, V., & Leitão, J. M. (2022). Downy Mildew Resistance and Genetic Variability in a Wild Rocket Germplasm Collection. *Agronomy Journal*, 1-13. <https://doi.org/10.1002/agj2.21190>
- Di Gioia, F., Avato, P., Serio, F., Argentieri, M. P. (2018). Glucosinolate profile of *Eruca Sativa*, *Diplotaxis tenuifolia* and *Diplotaxis erucooides* grown in soil and soilless systems. *J. Food Composition and Analysis*, 69, 197–204. <https://doi.org/10.1016/j.jfca.2018.01.022>
- eFloras (2008). *Flora of North America. Diplotaxis tenuifolia*. Retrieved August 8, 2022, from http://www.efloras.org/florataxon.aspx?flora_id=1&taxon_id=250009668
- Egea-Gilabert, C., Fernández, J. A., Migliaro, D., Martínez-Sánchez, J. J., & Vicente, M. J. (2009). Genetic variability in wild vs. cultivated *Eruca vesicaria* populations as assessed by morphological, agronomical and molecular analyses. *Scientia Horticulturae*, 121(3), 260-266. <https://doi.org/10.1016/j.scienta.2009.02.020>
- Elisiário, P. J., Justo, E. M., & Leitão, J. M. (1999). Identification of mandarin hybrids by isozyme and RAPD analysis. *Scientia Horticulturae*, 81(3), 287-299. [https://doi.org/10.1016/S0304-4238\(99\)00013-8](https://doi.org/10.1016/S0304-4238(99)00013-8)
- Eschmann-Grupe, G., Hurka, H., & Neuffer, B. (2003). Species relationships within *Diplotaxis* (Brassicaceae) and the phylogenetic origin of *D. muralis*. *Plant Systematics and Evolution*, 243(1), 13-29. <https://doi.org/10.1007/S00606-003-0047-5>
- Farinhó, M., Coelho, P., Carlier, J., Svetleva, D., Monteiro, A., & Leitao, J. (2004). Mapping of a locus for adult plant resistance to downy mildew in broccoli (*Brassica oleracea* convar. *italica*). *Theoretical and Applied Genetics*, 109(7), 1392-1398. <https://doi.org/10.1007/s00122-004-1747-0>
- Farinhó, M., Coelho, P., Monteiro, A., & Leitão, J. (2007). SCAR and CAPS markers flanking the *Brassica oleracea* L. Pp523 downy mildew resistance locus demarcate a genomic region syntenic to the top arm end of *Arabidopsis thaliana* L. chromosome 1. *Euphytica*, 157(1), 215-221. <https://doi.org/10.1007/s10681-007-9414-6>
- Garg, G., & Sharma, V. (2014). *Eruca sativa* (L.): Botanical description, crop improvement, and medicinal properties. *Journal of Herbs, Spices & Medicinal Plants*, 20(2), 171-182. <https://doi.org/10.1080/10496475.2013.848254>
- Genesys Team. (n.d.). *Diplotaxis tenuifolia* accession browser. *Homepage*. Retrieved August 12, 2022, from <https://www.genesys-pgr.org/a/v26DkDz6bWZ>
- GlobeNewswire. (2021, November 17). *Global packaged salad market is expected to reach USD 13.7 billion by 2028 : Fior Markets*. GlobeNewswire News Room. Retrieved July 20, 2022, from <https://www.globenewswire.com/news-release/2021/11/17/2336038/0/en/Global-Packaged-Salad-Market-Is-Expected-to-Rach-USD-13-7-billion-by-2028-Fior-Markets.html>

- Goulão, L., Cabrita, L., Oliveira, C. M., & Leitão, J. M. (2001). Comparing RAPD and AFLP™ analysis in discrimination and estimation of genetic similarities among apple (*Malus domestica* Borkh.) cultivars. *Euphytica*, 119(3), 259-270. <https://doi.org/10.1023/A:1017519920447>
- Guijarro-Real, C., Navarro, A., Esposito, S., Festa, G., Macellaro, R., Di Cesare, C., ... & Tripodi, P. (2020). Large scale phenotyping and molecular analysis in a germplasm collection of rocket salad (*Eruca vesicaria*) reveal a differentiation of the gene pool by geographical origin. *Euphytica*, 216(3), 1-20.
- Hall, M., Jobling, J., & Rogers, G. (2012a). Some perspectives on rocket as a vegetable crop: A review. *Journal of Fruit and Ornamental Plant Research*, 76(1), 21-41. <https://doi.org/10.2478/v10032-012-0002-5>
- Hall, M. K. D., Jobling, J. J., & Rogers, G. S. (2012b). Factors affecting growth of perennial wall rocket and annual garden rocket. *International Journal of Vegetable Science*, 18(4), 393-411. <https://doi.org/10.1080/19315260.2012.660565>
- Hall, M. K., Jobling, J. J., & Rogers, G. S. (2015). Fundamental differences between perennial wall rocket and annual garden rocket influence the commercial year-round supply of these crops. *Journal of Agricultural Science*, 7(3), 1. <https://doi.org/10.5539/jas.v7n3p1>
- Jasper, J., Wagstaff, C., & Bell, L. (2020). Growth temperature influences postharvest glucosinolate concentrations and hydrolysis product formation in first and second cuts of rocket salad. *Postharvest Biology and Technology*, 163, 111157. <https://doi.org/10.1016/j.postharvbio.2020.111157>
- Jin, J., Koroleva, O. A., Gibson, T., Swanston, J., Magan, J., Zhang, Y., ... & Wagstaff, C. (2009). Analysis of phytochemical composition and chemoprotective capacity of rocket (*Eruca sativa* and *Diplotaxis tenuifolia*) leafy salad following cultivation in different environments. *Journal of Agricultural and Food Chemistry*, 57(12), 5227-5234. <https://doi.org/10.1021/jf9002973>
- Kalendar, R., Khassenov, B., Ramankulov, Y., Samuilova, O., & Ivanov, K. I. (2017). FastPCR: An in silico tool for fast primer and probe design and advanced sequence analysis. *Genomics*, 109(3-4), 312-319. <https://doi.org/10.1016/j.ygeno.2017.05.005>
- Kleemann, S. G., Chauhan, B. S., & Gill, G. S. (2007). Factors affecting seed germination of perennial wall rocket (*Diplotaxis tenuifolia*) in southern Australia. *Weed Science*, 55(5), 481-485. <https://doi.org/10.1614/WS-06-197.1>
- Lee, J. S., Lee, H. B., Shin, H. D., & Choi, Y. J. (2017). Diversity, phylogeny, and host-specialization of *Hyaloperonospora* species in Korea. *Mycobiology*, 45(3), 139-149. [doi:10.5941/MYCO.2017.45.3.139](https://doi.org/10.5941/MYCO.2017.45.3.139)
- Luca, A., Mahajan, P. V., & Edelenbos, M. (2016). Changes in volatile organic compounds from wild rocket (*Diplotaxis tenuifolia* L.) during modified atmosphere storage. *Postharvest Biology and Technology*, 114, 1-9. <https://doi.org/10.1016/j.postharvbio.2015.11.018>
- Martin, J. P., & Sánchez-Yélamo, M. D. (2000). Genetic relationships among species of the genus *Diplotaxis* (Brassicaceae) using inter-simple sequence repeat markers. *Theoretical and Applied Genetics*, 101(8), 1234-1241. <https://doi.org/10.1007/S001220051602>
- Martínez-Sánchez, A., Marín, A., Llorach, R., Ferreres, F., & Gil, M. I. (2006). Controlled atmosphere preserves quality and phytonutrients in wild rocket (*Diplotaxis tenuifolia*). *Postharvest Biology and Technology*, 40(1), 26-33. <https://doi.org/10.1016/j.postharvbio.2005.12.015>
- McNulty, R., Kuchi, N., Xu, E., & Gunja, N. (2022). Food-induced methemoglobinemia: A systematic review. *Journal of Food Science*, 87(4), 1423-1448. <https://doi.org/10.1111/1750-3841.16090>
- Milne, I., Stephen, G., Bayer, M., Cock, P. J., Pritchard, L., Cardle, L., ... & Marshall, D. (2013). Using Tablet for visual exploration of second-generation sequencing data. *Briefings in bioinformatics*, 14(2), 193-202. <https://doi.org/10.1093/bib/bbs012>

- Monte-Corvo, L., Cabrita, L., Oliveira, C., & Leitão, J. (2000). Assessment of genetic relationships among *Pyrus* species and cultivars using AFLP and RAPD markers. *Genetic Resources and Crop Evolution*, 47(3), 257-265. <https://doi.org/10.1023/A:1008794809807>
- Nei, M., & Li, W. H. Mathematical model for studying genetic variation in terms of restriction endonucleases. (1979). *Proceedings of the National Academy of Sciences*, 76(10), 5269–5273. <https://doi.org/10.1073/pnas.76.10.5269>
- Nicoletti, R., Raimo, F., & Miccio, G. (2007). *Diplotaxis tenuifolia*: biology, production and properties. *European Journal of Plant Science and Biotechnology*, 1(1), 36-43.
- Padulosi, S. (1995). Rocket Genetic Resources Network. *Report of the First Meeting, 13-15 November 1994, Lisbon, Portugal*. International Plant Genetic Resources Institute, Rome, Italy. IPGRI Via delle Sette Chiese, 142(00145).
- Pignone, D., & Martínez-Laborde, J. B. (2011). *Diplotaxis*. In *Wild Crop Relatives: Genomic and Breeding Resources* (pp. 137-147). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-14871-2_7
- Pimpini, F., & Enzo, M. (1996, December). Present status and prospects for rocket cultivation in the Veneto region. In *Rocket: A Mediterranean crop for the world*. Report of a workshop (Vol. 1314, pp. 51-66).
- Rohlf, F. J. (1993). Numeric taxonomy and multivariate analysis system. *NTSYS-pc*. Setauket, N.Y., Applied Biostatistics, Inc. ISBN: 0925031313 9780925031310. 2009
- Signore, A., Bell, L., Santamaria, P., Wagstaff, C., & Van Labeke, M. C. (2020). Red light is effective in reducing nitrate concentration in rocket by increasing nitrate reductase activity, and contributes to increased total glucosinolates content. *Frontiers in Plant Science*, 11, 604. <https://doi.org/10.3389/fpls.2020.00604>
- Svetleva, D., Pereira, G., Carlier, J., Cabrita, L., Leitão, J., & Genchev, D. (2006). Molecular characterization of *Phaseolus vulgaris* L. genotypes included in Bulgarian collection by ISSR and AFLP™ analyses. *Scientia Horticulturae*, 109(3), 198-206. <https://doi.org/10.1016/j.scienta.2006.04.001>
- Taffner, J., Cernava, T., Erlacher, A., & Berg, G. (2019). Novel insights into plant-associated archaea and their functioning in arugula (*Eruca sativa* Mill.). *Journal of Advanced Research*, 19, 39-48. <https://doi.org/10.1016/j.jare.2019.04.008>
- Taranto, F., Francese, G., Di Dato, F., D'Alessandro, A., Greco, B., Onofaro Sanajà, V., ... & Tripodi, P. (2016). Leaf metabolic, genetic, and morphophysiological profiles of cultivated and wild rocket salad (*Eruca* and *Diplotaxis* spp.). *Journal of Agricultural and Food Chemistry*, 64(29), 5824-5836. <https://doi.org/10.1021/acs.jafc.6b01737>
- Vig, A. P., Rampal, G., Thind, T. S., & Arora, S. (2009). Bio-protective effects of glucosinolates—A review. *LWT-Food Science and Technology*, 42(10), 1561-1572. <https://doi.org/10.1016/j.lwt.2009.05.023>
- Vincze, T., Posfai, J., & Roberts, R. J. (2003). NEBcutter: a program to cleave DNA with restriction enzymes. *Nucleic Acids Research*, 31(13), 3688-3691. <https://doi.org/10.1093/nar/gkg526>
- de Vos, A. C., Broekman, R., de Almeida Guerra, C. C., van Rijsselberghe, M., & Rozema, J. (2013). Developing and testing new halophyte crops: A case study of salt tolerance of two species of the Brassicaceae, *Diplotaxis tenuifolia* and *Cochlearia officinalis*. *Environmental and Experimental Botany*, 92, 154-164. <https://doi.org/10.1016/j.envexpbot.2012.08.003>
- Yu, M., Li, C., Hu, C., Jin, J., Qian, S., & Jin, J. (2020). The relationship between consumption of nitrite or nitrate and risk of non-Hodgkin lymphoma. *Scientific reports*, 10(1), 1-11. <https://doi.org/10.1038/s41598-020-57453-5>

Table SM2. Twenty tested SSR loci.

Locus		Sequence	Locus		Sequence
MT317459	FW	CCTTTGATCTCACCAGTC	MT317571	FW	CTTGCTGTTTGAAGTTTGC
	RV	AGCAGCAGTAATAGTGGA		RV	CGTAGCAATGACTTGCC
MT317469	FW	GCGTATGTCACCAAAGAT	MT317762	FW	CACAAGAGGCAACAGAA
	RV	CTGTGAGTAGCTGGGAA		RV	GCCTCATTTCATGTTACCTC
MT317471	FW	CATTGATTGTCCCTAAGCT	MT317541	FW	GTGTGGGAGCATTTCAC
	RV	ATGCTAGGGCTCATAAGT		RV	CCTGGTGTCTTTGCAG
MT317520	FW	CCTTAATAGCCACCTGC	MT317527	FW	ACTTTGACGAAACGAAGC
	RV	TTCTCTGCAACATCTGTG		RV	CTCAGAACCAAAGAGAAGC
MT317482	FW	CACCACGCAGACAATATG	MT317537	FW	CAGCTTTCTGTTAGTGGTC
	RV	CATATCCGTGGGTAAGGA		RV	CCAACCAAAAACGTCAGA
MT317461	FW	TGAACCTGTAAAGACTCCA	MT317610	FW	GTCTATTGATCTGATGCCG
	RV	CTCCAAAACATACCTAGACC		RV	GCGAGAACCGATCTATGA
MT317478	FW	CGATGTCTTATCCATGAGC	MT317613	FW	GTAATAAAGCGACCCTGC
	RV	TTGCAGAGTGATGGGAA		RV	CTGGTGGGAAGGTATCA
MT317577	FW	CGGATAAACATATCCGCTT	MT317734	FW	GACTAACTCTGTGTGTGC
	RV	GGTTAACATCACTAACGGT		RV	TTAATTCAGTTCCAGGCG
MT317578	FW	GCTGAACAAACCATACGA	MT317741	FW	CAAGTAATTCAGTCTCGGG
	RV	TCTAACTACGATCAGGACAG		RV	GGTAAAGCAGGGTCTTG
MT317582	FW	GCAAACAAAGCCAATCC	MT317823	FW	CCAACATAGAAAGGTGCG
	RV	CTTCTCGTCTTCTTCAGGA		RV	GAGTTCCTCCAAAAGCTG

Table SM3. (A) Twenty microsatellite loci retrieved from *D. tenuifolia* genomic data. The 5 loci selected and used for unequivocal analysis have their title in bold and underlined. Primer loci are in bold and underlined. The SSR loci highlighted in green. Grey highlight represents the PCR fragment amplifications. **(B)** Nineteen SNP loci containing a TaqI restriction site, retrieved from the genomic data. The 9 loci selected and used for unequivocal analysis have their title in bold and underlined. Primer loci are in bold and underlined. The SNP variations are highlighted in yellow. Grey highlight represents the putative PCR fragment amplifications.

(A) SSR Loci

MT317577

AAAAAAGTTATTCTAGTTTGGTTTTGGTTCAGTTATTTCAAGTAATTTAGAAAAAATAACGA
 TATTTTAGATAAAATATTGAGTATTTTGAAAATTATCAAATTGTTAGAATGATTTAGATAAAAA
 GAGTTCAGATAGATAGGATTGTTTTGGATATTT **CGGATAAACATATCCGCTT**ACTTTCAGGATT
 GATGATTGATAATTTTTTTAAAATATTTTCGAAGACTTTAATAGATTTTTAAAT **TATATATAT**
ATATATATATATATATTTTGGCTATGCTATGTATATAATTAATATTTGGGTTTTTAGTGATTTAA
 ACTCCCAACTATTTTTAAA **TCGGAAGAAAAACCTCCA**ACTAAAATTTTATGTTTTTAAACCTTG
 AACTATTTAA **ACCGTTAGTGATGTTAACC**TCCGTTAAAAAATTCGTTTTTCCCTAAACAGACATC
 CGTTAAATTGAAACGCGGTGTTTCATATAATCACATAAAGGACACGCAATGTTTATAATCAACA
 AATATTTTAGTGATTTAAACCTCTAACTATTTTTAAATAGGATGAAAAACC

MT317527

TGATGAACAAAAAATTCACCATAATTTTTTTCGTAAACATGGTGTAGCAGTCGTTGTTAATT
 GCCTTTACGAAAGATACTTTATTTCCACGGTTTATGAAATGACCAGATAAGTCCGATATATATA
 GTCGGATCAAACATTTTAGAATATTTCTATTTACATGATCTGTACAACATAGTCAGGAATGGAG
 AAGGAAAATGAGAATAAATCCCAACTGCTATTATCTGATATGAGATAATTAGTCACTAAGATCA
 TATCAGAGGGTCAACCCGCGTG **ACTTTGACGAAACGAAGCCC** **GTGTGTGTGTGTGTGTGTGTGT**
GTGAGTGAAGCTGGTTAGGGTTTTGTCAGAAGGCTCGACAGAGTGTATTAGTCGTGGTCAAGTT
 AGAAGAGTGTGTGCTGTTCAATGGGGCTGCTGCTTGTGGGACTACAGATTCT **GCTTCTCTTTGGT**
TCTGAGATCCTCACGTTGGAGATTGCAACCACTAGCACATCTACCACACCGCTCTCATCTTCTT
 CTTCTCTCTCTGTGTTGGAGAATGAATCAGATGTTGGTTTCT

MT317537

GATGAAGGTTCCGAGAATAACCAGCTTGGTGTCAAACCAAAGATGGGTTATACTATAAATTC
 ACTGGCTTTTCGTGACCAGGTTTGTGTTTTGCAGACAAGATTCCTTTGCTTTTGCTTGAAGCTTA
 TGTGTGTGTTCTTTTGGTGGTGTAAATAGGATGTTGCGAGCCTGACAAGTTTTTTCCAAAGTGCG
 TTTGGGAAAACACCTGAAGAGAAA **CAGCTTCTGTTAGTGGTCC**GGAATTGGGGAGAAGTGGATT
 TGAATG **GTGTGTGTGTGTGTGTGTGTGT**TGTCCTTCGCATTTAAGCTTCTTATGTGGATAAAG
 TCTTAGCTTTATTTCTTTTTTGTGTTGTGTGTGTGAAGGGAATAA **TCTGACGTTTTTGGTTGG**
 GGGAAAGCAAGCTTTTGGAGTGTCTCTGGCTGATGTTTACAGACTCAGCTTCAAGGAAAAAT
 GATGTTTTTGTGGAGTTTCATGTTGATGACACTGCTGGTGCCAATGAGGTATGATATTGTATGT
 AAGTGTGTGTTATGTTAGCTGAGTTGCTGATTCTGTTGTTA

MT317610

TTAGTGATTATGCTAGGTTGAAACGGTTTTAAAAGTAAATTGCAGGTTGTGAGAAATGGTAAT
 TTGCTCGATTGAATCGGTTTTGGTATTAACAATAAGGAAAGGTAGTTAGATTTAAGATTACTA
 TTCAGATAATCAGGATTATAATGCTATATATGCCTAATAAGTTGCATGCATGATATTAAGA
 TTGAAATATGCTAATATGCTCCTCGCGTTTTAGACTC **GTCTATTGATCTGATGCCG**AAAAATATCG
 ATCGATGTTTTGACAGCGATATCGA **TATATATATATATATATATA**TTTGTCGTCGATCGGTT
 ACTCTATACGAGTATCGATCGATATAGCTCAGATCGAGTTGATGTGGATGCTCACCAAGA **TCAT**
AGATCGGTTCTCGCGTATCCCTAGAGGATCTTAGCGTATGCAAGATGATGTTAGGATGATGATT
 AACGAGTTAATTCATTCCAAGATCATATTCTAGTTTGTACTCTAGGATAAACATTAGAGTTCA
 TCATGATAATGAATATCACAACCTCAACAATCTATAGTTGGGGCT

MT317582

TTAAAATGAGCATAACAAGAACCAAGAAAAAGTAGCTAAAATTGACTCGTGTAACCTTCAACGT
ACCAAAACAAAAACCCCTTTTGTATTATATCCCCAAATCCTCCTCCTCTCTCTCATCTTCTCAC
TTCAAACCTCTCTATCTCTCTCTCTCGCCTCATATCAAGAGCTAAATATTTCTCTGAAAGCTG
ACTTCAAAGCTTTTTCTAAA **GCAAACAAAGCCAATCC**TTCTCTAATTGAGATTTTATT **TATATA**
TATATATATATATATATAAATTTAAAAAGATGGGGAGACATTCATGCTGTTACAAACAGAAGCTG
AGAAAAGGACTTTGGTC **TCCTGAAGAAGACGAGAAG**CTTCTTAGGTACATCCTAAGTACGGCC
ATGGTTGCTGGAGCTCTGTCCCTAAACAAGCTGGTAATATAATTTTCTTCTATTTATCCGTTTA
TTATAGCTTTATTGTGTTAGATTAATCCTCCTTCTAATTCCTGAATTATCATTTAAAAATATT
TGTAGGTTTGCAGAGGTGTGGAAAGAGCTGTAGATTAAGATGGATAAACTA

MT317571

ATATGATTTATGTGGCTAGCTGGTTCATGCAGTTTTTTAGCCAACCTATCTATTGATTATTGCATC
AATCTTATTGTTTAAAGCAAAGTCCTTACAAATTGCACAACAAAAAATAGTCATATTGCAAGAA
TGATATATTGTTGCAATATGCTCATATATTTTTGCAAATATGAAACTGCAATAGAATTGTAACA
AATTTTAGTTT **CTTGCTGTTTGAATGTTGCA**AATTCATATTTGTTGCATAGACGTAGTAAAATT
AGCAACAAATTGAAATGCTGT **TATATATATATATATATATATA**TCATATTTTGGCAACCATT
TTAATAGTGATTTTGTGTATATTTACAACGATTTTCTAACGATAAATGGT **GGCAAGTCATT**
GCTACGTAAAATATATAGTAAAACACATTGCAAACCTGGTAATTATAGTTTTTTTAAATTAATAA
TTTTATTGCAAAAAATATTTATATATAATTTTTCCGTACGATTGCAACGATTGTTGCATTTT
TGAAGTTATTACAAAATTTGCTATATAAAATATATAGCAAATGTTGTTGCAA

MT317762

CATATTAGAATTTTTAAGAGGGAAAACAGCAAACCATTTATTTTAGTCAGTGAGCTTCCGATGT
TATGTATGATTTATGAGGCCATGATAACTGAAATATCAATTGATATGACTCGTTTTGCTACTG
GCATCATAAATCTCTAGTTCTACTACAATAGAGAGAA **CACAAGAGGCAACAGAAAACTTAAAA**
CCAACTCTATGCCAAGAGAAAACATAAAGGAAAGCCA **GAGAGAGAGAGAGAGAGAGAGAGAG**
AGAAACCTTTCTGGCTTCATCTTCTTTCAACTTAGCAGCATGGATCTTCTCAGTACAGACTTCA
ATTTCTTTTCTCAAGTGTTCCAGTGCTTACAATATTTTCAACAAAAGAAGA **GAGGTAACATGAA**
TGAGGCGTTCCCTGAGTAAAAAAGTGAGACACTTGCTTGCTAATACTGTACCTGCTTTCTTA
GGACCAGCAGTGAGTTTCATTTCCATGGTGAGGTTACCAAACATGCTTTCTTAGGACCAGCAG
TGAGTTTCATTTCCATGGTGAGGTTACCAAACACCCCTTA

MT317541

TAATCTTATCTGTAAAGCTTTACGTATGCTCTATTGTTTTCCCTCTTTCTATCACGCATGCTAT
ATTGTTTTCAATCTCATCTGTGAAGCTAATAATATCTTTTTTTACTCTTGCAGATGGCATCCTC
TGCTTTTCGATAACGAACAAAACGATATTGATATGCATTGTGAAGACCAAGAGTACCAAGAGACA
AGTGCTAATGTGAATGCAACATCCAGAGCATCCTCTCAGGCGAGTAAGGTGAAGAAGTTGGTTG
TGCCAAGGTCTCGT **GTGTGGGAGCATTTCACA**AGAACC AAAGCGACCCGAGACAA **GTGTGTGTG**
TGTGTGTGTGTGTGTGTGTCACCATTGCCACAATATCTTTTCATGCGCATCTAAGTCAGGGA
CTTCGAACCTGAAGCAGCATGCAAAAT **CTGCAAAGAACCAGG**CGTTCTTAATTGGTCAGAAAA
CGGATCAACAACAGCTTGATAATGAAGGAAAGCTGAAGCCATCTAGAGTTTCCGAGTCTTTTT
CAAAGAAGCGTGTAATGAATTGATGGTGTAGCAGAGTTGTC

MT317613

AAAATAAAATCTAACTCTTACCAACAAAAGGTAACCTCAGGTAATTTGATAAGCTAATGTGT
AATCCACCATTAAGATTTTCATTTTTTATCTTTAGGTGAAGAAATTACGAAAAAGCATGAGACC
GTACATAATACTTATGGAGCTTCTATTTCAATTAATCTATATTAGTACCTTAGCAGTTGCTCAA
AACCAGGATTACATGACATGCAGTACATGTTA **GTAATAAAGCGACCTGC**ATGCACCAAGTATA
ATAATCTTAATTTTACA **TATATATATATATATATATA**ACTAATAAATTTGTGCAAAAAAACT
TCTCAATAAAATAAATCTTTAAATTTGGATAGCAAAACAA **TGATACCTTCCCACCAG**TTTCCCTC
TTGATGCGCTTTTGGAAATCGCCTTCCCTCTCCATCTTTCTTGTACTGCATAAGTTAAAGATAAAG
AAGAAAAGCAGAATAATTTTGAATTAATAATACATATAGGGACTACACGTACATGTAAGTAT
ATAAGTTTTGTTACGTTACTCACATAGGTGTTATGGGTATGATAGGTAC

MT317734

TCTTGCTTCAGTTCATGAGATACAAAATGATTTGGGGAATCCATCATATATTATCAACGGTTTC
TCTCAAGTTCCTCCAGATATGCGGATATGAGGTTGTTCTGTTGATAGTAAGGTCAAAGCTCACAC
AGAAGGTGCTCTACATGATTCTGTAATGTTGATGTGGTAGGGATTTAGAGCAAGAATT **GACTAA**
CTCTGTGTGTGCACCTTCTGTGCATATATATCTTTTTACTATCTCAGCAAGAAGA **TATATA**
TATATATATATATATATATATATCTTGTACTTCCGCTGATCATATTGCTCTGTTTCAGAGATT
ACAAAGCAACAACCTTTTTAATTCTTCCA **CGCCTGGAAC****TGAATTAA**AACTGAAGCATGCTGAG
ATGTTTTGTTACCAAAGCGGGGGCTTGTACCGACTGTTTCCTGACTTTTCTAATTATAAAAAGA
TGCGCATAAATGCATCGAACTTGATGATACTTCTCAAAACATAGTAAAAAGTTTTTACTATAAA
TAAAGTGAATAGTGAAAGATGAAGCTATCACCAATCAATTCTCAG

MT317741

AGCACCAAAGACTTCGAGGGGAGAGAGTTCAGCCAGTAGCACAAGCTTGCTGAATATTATACT
GATGGTGAGGAAAAAGCACTTGATTGTAAGCCTTAACTGCTGCATCCTTTAGAGAGTCATCACC
TCTGTGAATCATATTTAAAACACGATTTAAACCCCGATAAAATTACAACAAGCCACGGTCAAAC
TATA **CAAGTAATTCAGTCTCGGG**TATCTTGTATGGCTATATAA **TATATATATATATATATATAT**
ATATATATCTCTGCTTGTGTGTGTTAACATTTTTTCGAGAGAGTCTATATAAGAAAACCTTACTCA
GTAGCCAATAACTCTTCGCACAGAACCTTGATCATTT **CAAGACCCTGCTTTACC**TTCAACAGGT
TTCTTGTATGGCTACCAAATTTCTTTACACAGCCTACCTCAATGTCTTTGTCAATCATTGCCCTC
CAATGTGCATATTGAACGCGATGCCACGGCCAGATCATTGAGCTGGAAAAATAAGTTTGTCTTCT
CAACGCTTGATATATAAACAAACACTCAATTTATAAATCGTCATGAT

(B) SNP Loci

OM145023

GAAAAAAATTTCTAAACCGTTCAATATTTTTCTTTTTGATAATCCTTTTCGGCCCATGTCATTCCTCTTTTCA
CTCGAGAATGATCCATTTGTTAGTAGACTTTCTATTAACCATTATAAATATTGTATATAATGTCTTCTTT
TGAGCACA **GTCCCATGATTAGATATGGT**TGTATATGCTAATTAATTTTTCTTCTCCATCGTCTCAGTGT
AATAGATATGCTTGCCTTACTTGGTGGAAATATGAATGGACTATGGTTCCTGAATGTTTCCGAGCC
TACAAACTTTC **(G/T)**AAAGGAAGTTACTTTGGAGAACCAAAAGAAAGAGATCTCGTGGTTTTATATCT
TGGAAAA **CGCTTGTCACTTAAAGACT**GTAACAATCTCGGTGAGAGACTGTAATGATGTTCAAAAACCTTAA
GATGATCAAAGAGTTGGCACTTTCTTCTCGAACTTCAACTGCATGTCAACTCATATTCGAATAATCTTAT
GTTTATATTTTTATGATTTGAACTTTGCTTCTATATGTATCCGA

OM145040

AGCTCAGAATCCTATCTGTGACATGCCTTGCCACCATGCTTGTGTCTGATACTTTATCAAAAACAGAAAA
TGTTTGGGAAAATACATGTGAGTTTTTTCGAAAAGACATTCAATATAAGGAAACATTATAATAGACCATGT
AACTTATCTCAAATTTTTTTCTTTGGTTCATAT **GCCTTTGAATGGTTTCTG**GAAATTTTAGGTATAAC
CATGACTATGTTGGTTTATAGGATTATGTCTAAGTGATG **(T/C)**CGAAAAAGAAGGAGCCCTACTGAAA
ATCAAATAATTTTTGAGAAGTAATGGAACCTCATTAGATAAATTGGACATCAATGCTTAAAGCCCTGCA
ACTTTAATGACTCTAAAAATGTTTTGATCATGAATGAGCTGAGTTATAACCAAAAAGAGTTTAGGGAGA
AAATGATAGAGACATCTTAAAA **TTGACGGATGAGCAAAG**GAAGATATACAAAGAAATTATTGATGATGTT
TTAGAAAAAGAAGGGTGTATTTTTTCGTTTAT

OM145068

GTTTATTACAAGTTTAGTTATTAGAATAATATGGTTTAAAGAGGCTTGATGAATGATCAAAACGTATTTTA
TTGAGAGAAGTAACATACAAACATGCGTATAGAAGTACTCTTAAACGAAGACTATCATCTGTACGTGTTTC
CTGGACATCTTTCATGTCACTGTCACGGTG **CATCTGTGCCTGATCTC**CAGTGATTTTCAGACACTGCCTG
TTCCCTTGCCCTTTTC **(G/T)**ATACCTTGAGCGCTTGAGAAGAGCTCTACAGGCTCAGACTCTGCGATCTT
GCCTCCTTCAAGTGGTGTCCATGTTTGTATCTCACTCGTATAGCCTCGTCTTCAGACTGCTTCGCTTGA
GCTGTTGCGTAGTTTATGTCGTCTTTGTCTTTTGCCATCTTTAGGGTTTTAA **TGAATCAGACTTGACATG**
GGACCGTATATAAACAGAGTTAGGTTCTTCGTGTCAGCTCTGGACTCTCGTGTGTTGCTTACTTCTCTT
TGTTGTGTGTTCTCTTTTTGTTGACAC

OM145116

ATCCACAATAACTTCTCAAATATGTGATTCTTCTGTTTATAGAAATAGCATTACTCTCTTTAACCTAGT
ACCTTCGATGTACACATTATATAAAATACTAGTTGTATTGTCTTCTTCAATAATAAGAAACACATTACATT
TACATGGTATCAA**GAGCCTTGTAACAGATCC**AAAAATTTCTTTCTCTCTCTTTCTTCTCACTACACACCAA
CTCCATTATCATGTCTACGACCTCTGCTGAAACAG **(T/G)** CGACACAAATCCACAAACCCTCCTCAATC
TGAACATGACAAACGTACGAAACTAACCTCCAGCAACTACATGATGTGGAGCCTCCAAGTTCAAGCTCT
CCTTGACGGCTATGATCTCAGCAGCCACCTTGACAACCTCCACTCCACCACCGCCACCAACCATAACCCT
GACGGAA**CAACAGTCACAAACCCAG**AGTTTCGCCCTCTGAAACGTCAAGACCGGCTAATCTTTAGCGGTC
TCATCGGAGCCATCTCAGTTCCCA

OM145191

GAATCCCTACTGTCTAAGCACCGTCTACTTCTCATTTCCATTTGAACAGCAAATTGGACAATACTTAAC
CTGAATTCACCGCGAGA**TGGCATCCTACTCTTTCTC**GGCCTACCTGGTGGGCGGCGTGTGCTGGCGGGA
AGAGCAACCCATTGTCTCCTTCCACACCCATCTTAATTGATGTTATGTTGGTGTGGTCAACGGGATTGAT
GCTTCCACTGTAAGCCATTGCCAGTATGCGGTGCTGTAAACTGGGT **(C/T)** GACCTTACTCTCCACAGA
CTCTTTAGCATAAAATGCCCGCAG**CCATATCATGAACACACGG**TATGGCTAGCGCTTCAAACCTCTTGCAA
GACCATTTTTTGTTCACCAGATCGACTCTTGAATACATTCCATTTTCATCCACAACCTCGTACTCTCCCT
CTATGATGTGGTGGACACCATATTCTGAACCTCGTAAAGTTTTTTGCCAAGATCTCATTACCTTTGG
TGTTAGCTTTCCAACGTTTTGTGCTACCGCTCATGTCTT

OM145198

AAATGTACCCGTGCAATGCAAGATATTGATCTTGGTAGCAATGATGCTCCTTTTGTGCTACCTGTTGAT
GTGGTGAATCGGGCTGCTGAAGAAAAATATTCATTCTGGTGGGGAGTCCAACGATGCCACGGAAACAGA
ATCTCAGAGGCATCATCT**TCTACGATGCCAAGAGTT**TGGGGGTTAGAAGGGATTGTTCTGGGCAAATTAT
GGAAGGAAGAAGATCCAATTCATCTTTCCAAGTGAGGAAT **(T/C)** GATGGACAACGTACTCAGACGTGG
TCCGTGGGCTTATGCAGATAGGATGATCACGCTTCAGAAGTGGACTCCGTTGATGGATTTGGCTTTGTTG
AACTTCATTCCTTTTGGATTCAAGTTTCGT**GGGATACCTTTCCAGTACA**TGAATCGTTAAGTGGTGTATCA
ATATTGCAAGATTGATGGGGCAATACATCCAAGTGGATTACAATGAAGAAGTTGTAGGGCGTTTGGAGTT
TGTTCTGATCAGACTTAATTGGGACATCACTCAACC

OM145248

AGAAGTGTAAAAATCGTGTAATAAGAGGTGGACCAATAAATGTAAAAGTAACGGATGTGTCAGGCCACA
AAATAGGTGGCTTATCAGTTTTTTCACACTCCACTTGTGTCAGGCACACATTATAAGCATCGAGTCACATG
CCCTGACTCAGTT**CTAATGAGTTCGTCGTTCC**CCATTAGCTTTTTTACATCTATAAAATATATTAAGTGACA
TCTCAATGTTTATGCAAACTAGTTTTTTTTTTTCATAAGGCTTCTTTTC **(A/G)** AGTAGTTTGCACAAA
GTAGAGAGTGAATAAAATATAATAGAATGTATGCATTAGTAGAATTTGACTTGGTTCTCCATTACAATGG
CCACGCTGCATGTGCAATATCTTTCAAGAACGAAGGTGATGATGACGATGATGTAGTCTATGATTACGCT
CCAGCTGCATGAGTCAATCATCTTTAGTTATGACGCTTTCAC**CACCAGGAAGAGATAAGCT**TTTTCTAGCT
AGCTAGATAACAACCTAATTAGCAATATATGTATTTTTTAGCAATATATGTAT

OM145268

CCACCTTAATGTGTGCTGAGCTGGATTAGTGTAAAGATGAGTGTATATCAAAATTAACCTAGTAAAAATG
TG**CGAAAAGAGATGCTATCGG**CTTTTCATGAACCTTATTGTGTAATCAATCATAAAAGGCTTTACTGCACAA
TGTATACAAAATTAGAAAAGCTGATGAGAAAGTAAACAATTTGCAATACAAACTAAAGAACTGGTAAAGAAA
AATAGTGGTCTTTGATATAGGAATCTACG **(C/T)** CGATTATGTGCCCATCATATCCCCATCTGTTAATAA
CTTCAACACATCAGCTCAT**CGTTAGAAGTGGAGACGTA**CCATCCTCGTCTTTTTTTATCATTGCAATT
GCATTGTTACCGGCACATCAGCTCATCGGTTGTTCACTGATCCGAAATTAATATATACTGTACTACTTTT
TTAAAATGTCGGCAAGGAAATAGTCATCTGAATTAATTTTATGGTTTCAGAACTGGGGAGAAGCTATTAGAC
AATGATTGGAGGAAAATAGATGAGTGTGTGGCTACCTTGT

OM145426

CAGCAGATTTAGTCGGTTGTTGTATGTGTCGGTTTTCTCTATAGTGCAGGGGGTTTTGGATGGTTGCTCG
ACCTCATGAAATGCCGATCTGGAGGTTGATCTGGTTCTCGAGTTT**GGCTGGATATGTGTCAGA**TCTACTT
CGGAGAGTTTTTCAAAAATGTCAAAGTGTCACTTCTTTTTCATTTCTGGATAGCTACACTGCATCCTCTTG
TCTCCATCTCTCATTTTTTCGGTTTTTTTCGT **(A/C)** GAGGTTTGGAGCTTTCATCAGGAACATCTTTGC
TTATCATCTTGCATAAATTTAGCATGACGATCTGCATGATTGGACTTCTTTTCAACTACCTCGGCTGT
TTTTGGAGCTTTTGCTCCTCTAATTTTTTATAG**TTCCAGAAGGAACATGAC**TGATGTTTTACTTTGCATT
TTTTCTATTTTCTGAAGCTGTAGTCATTTATGAATGTTTTTCTCGAGTGAAGGGACCGATGTTGAGTGCCT
AAAACCTCTTCA

OM145047

TATCAACATGCAACATCTCGATACTAGGACTCTAAAGTACAAGATCGTGCTTTTCTCGGAGACATAGTCT
TGACTCTTGCCCATGGTCTTCGACAATACAAAAAAAACGAGGAACCAGCAAGATAGGAACCAATGGGA
AGGGTCTTACAAAATATCAAAGATAGTTTGTTCGGAGTAAACAAGCTCGAAAAATGTAAAAC TAGATTC
CCGAGAGTTGACCTCGGAATACTTACAACCTTTAAAAGATACTACCCTGAAAGTT (C/T) GATTTGCAA
GATCCAACCTCACATCTCGAACTAGGGGGAGGCTAACATTTCCCTGTCACTTTGTGAATCAGGATCGGCAC
CCCAAACCTTCCAAATAGAGCAATAAATTCAAAACAACCTTAAAGCGAAAAAGAAAAGAAAACCCGACAT
AGAACTACCAAAGACTTGATCCTTCAAAAAGGGCACATAGACATCTCAACCTCAAGCGCAACTCAAATA
AAATAAAAAATGCCCTCAAGCCGTTA

OM145130

CGTCTCCTGCCTCCTCTCCGCATCTTCTGCAGCGTAATCCAAACTCAACATCTCCTTGCTATTCTTCAG
AAGTTGTACCGTCGACGACTATGTGTGTAGTCTTATTCGGCAGCTTCTCTACACCTCAATGCAATGGGA
GATCAATGCACCTTCAAAGACAGTTTTTTAGCTGTTCTCAAACCTCATTTATTGGGTTCTTTAACGT (C/T
) GATTTGACTTCAGCGCGTTCTTTCGAATGCAGGCTTAGGATCACAAGTGAAGCTTCTCTATGAGTCT
CTTCTCTCCTCAGCCGTATATATCCTCTGTTATGTTTTAGTTGTTCTAATTTAGTTTACCCTCGAGTTT
CTCTCCAGTTGTATTCACCTTAGCCTTTTCGGGCTTAGTTTCTAATATAATCAGTTGTTCAAAAAAAA
GTATATAACCTATACATACATGCCACTTGAATATTTAGCTTTATTAATA

OM145178

AAAGTCGAGAAGGCATTTCTCATCCTAACTCGAATTCCTGGTCCATGTATTCCTCCGTTAGTTGACCTTGT
TCTGATATAGAATTGATGGCGTAGTTATGCGCCCTCGAGATCATGTCCCTCGTGATGGACGTACTTGTCT
GTTGTGAGGGTTTTTCTTTTTAGATATTTGCTCGGATGCAGCGCTTTTCGCAGATGTCTTTGTTGGCTTA
TGCTGCTCCGACTGGATTTTAATCTCCTCTT (T/C) GATGATTATGTAATCGTTAGCTTTGTGGAGTGCG
TCCTGGATCGTACACAATTTCTCGAGAGTTATCCACTTTCTGAATTTGACTTTGAACCAAAGTGACTTTC
TTAGTACGTCGATGGCTTCTCTGTGCTTATTCGGCTAACCTTGCCATGATGAGTTTGAACCTATTCATG
AAGTCGCGGAGAGGCTCTTCTCCTTCTTAGGACAACTCCAGAGGTCAGCGTCTGAATTTTCTCTGTTTA
TGAACATCGAGTACTGTTTGAGAAACTA

OM145294

ACATAGCCTAACCTCTCCCAATCACCTTCGCCTTCTCCTCCACCTCGACGGCCGAAACCTTCGCCGTAA
ACCTCCTCTCTCTCTTCTTCTTACTCTCAAACAATCCAGAGAGATCCAACCCCGACGAAAGCGAGATCAG
ATCAAAAAGCAGTTATCCAACCTCGATTTCTCCATAACCGGTTCCGGTTCAAATTTAAATTCAGAACCGGTT
AACGATTTCTTAAACCAGGCGGTTTTTCATAACCGCCTC (G/T) AACTCATCCTCGTACAGGATTCGGA
TCAAGCATCTGATATATAATCGATCTCGCTTGCTTCGAGATCCAGCTCGGGAACCTATAATCCCTCCGTT
GAATCTTCCGATACATCGAAGCAATATTTGAATCATCAAACGGGACCTCTCCGGCGAGGAGAACGAAGAG
AATCACTCCGCAAGACCACGCGTCAGCCTTCGCTCCGTCGTATCCCCGCCGCGATATAACCTCCGGAGCC
GTGTAAGCCGGCTACCGCACGCGGTGTGGAG

OM145310

AAAAACAAGACAAGTCTGCCGAGGAGTAGTCACGCTTCTCAGCTACTCTGTTTACCAGACAGACGCACTA
CCAAGGTGCTACCTCCGAGATGTTTACGATGCCATACGCCCTCGCTTAGCAACACGTCTCAAGCTACGAGC
CATGTACTGGCACTAGCGTTACGTCGACTAGCACGCTCTGATCAGGCAAGCCAGCGACGGACCTCGTCTA
ACCGACCTACCACCGGTTTACGAAAAGACGATCATTTT (T/C) GAACTCATAATTTTACGAGTGTGGG
TTGGCCGACAAAAAGGCTCTGTAAAAATACTTTATGACTTATCAAATACCTTGATTGCTCTAAAATCATC
GTACTCACATGCATTTATGTTGGTTTCGAAATCGAGCAACTATAGAAACGTACGAGATTAGACCGTTACAT
TCGATAACCTTAAGTCTAGTACCTTACGAAAAATGACTTACTGAGTTACTTCGATCAAACGAGTTAAAAC
CTAAGATGGTTGTCAAACCTCAAATC

OM145328

TGGATTTTTTCCAAATTTCCCATACATCCTGACGATCAAGAAAAAACTACCTTCACATGCCCTTTTCCAAAT
TCCCATACATCCTGACGATCAAGAAAAAACTACCTTCACATGCCCTTATGGGACATTCGCATATCGTAGG
ATGCCTTTTGGCATGTGTAATGCCCTGCAACTTTCCAACATTTGTATGATGTGAGTTTTCACGTATATGA
TCGACAATTGCTTGGAAAGTATTCATGGATGATTTTTCCGCTATGGATCGAGTTTT (C/G) GAAAATGTC
TCGACAACCTGTGCAAGGTTCTCGCCAGGTGCGAAGAAAAGAACCCTCGTGTGAATTTGGGAGAAATGCTA
CTTCATGGTCAATAATGGTATAGTTTTGGGACACAAGATCTCTGCAGCTGGTATAGAAGTAGTCCGAGCT
AAAATCGAAGTAATGACAGGATTGCCAGCACCCACGAACGTCAAAGATGTTAGATGTGTTCTCGGGCATG
CAGGATTTCTATAGGAGTTTTATAAAGAAATTCAGCAAAA

OM145355

ACGTTTTCTGTAGACTGCCTTTTTCTTTCCCTTCAATTCAGTGTCCATAGGTCCCAACCTGATCACAAAA
CACTTGTAAAAACAGAACATTTCTCAGCTTTTAGCAAAGCAAGACAATGAGAACTCACATAGTGAACAA
CCACATGATACAAAAGGCAGTGGAACAGACAGCAAGCCCAAGATCCTTCCATTTGATAGATACAGGGGCAG
ATTCATCAGCATCGATCCTT (T/C) GAGAAGGCTGACCAAACCCATTGACCAAAGCATTAACAAACTCAG
CAGGAGAAACACCATTAGTAGACTGAGACTTGACAGAGGACAACATGGTGTGGTATATCCAGAAACGC
CTCGGCATCTGCAATCTGCTCCCTTGGCTTCTGAACTGTATAACCAAACCAAACAAAAAAGCAG
CCTTTTCAGCAAAAGACATAATTAGCCTCGAGTCATTGCTTATAACAATCAAAGAGTTTACCTTTCTGGTG
CAAGTTCTCAAATTCATTGAAAAAT

OM145362

AGGAGAAAATATTCGGCCCAGAAGCGGAGAAAGAAAGCTCAAACCAACCTACGGAGGAGTATATAAAGGAT
GCTAAGGCCAGAGGAGAGGGGAGACTCTTTTAAACCTAGACAATTCCTGTTAGCTTAGACGAATTTAGGA
CTTAGGTGGCTAGACTAGCAAGTCGAGTACTTATAACGTTTGGTGATCTGAGCCTCGAATTTCTTTGT
GTAACAACCTGCAGTCTTTC (A/G) ATTCATATACTTATTAATAATAGCCTTTACATAATTTCTCATCTA
TATTATGTTTTTCTCTTTGCCGTTTTTGTGTGTTTGTGCAGAGATTGGAGACCTCTGAAAAATTAGAATT
TTCTTAAGTTTCTTTATTTTACAGAAAATCGACAGTACGAAATTTATGTTCTCACAATGTAGTCCCTCGTC
TTTTTCTAGATCAGTCTTTAGAAAATTTGATCACCGGGCATAACATTGATTGAAGTGAATATTATTGTAG
GTATTAGTTTACATAACCCCTCTAA

OM145386

TCCTCCCTTATAATCACCCCTCCTTCCCCTTCATTTTCATCAGGTATGAAGTGCCACTCTTCTTTCAGCAT
TCTTCTTCCACTCACCTCTCCAAAGTTTTTACACATATCGTCATCTGTGAAAACAAAACAAATTTACGAAAT
GTGTCATTATTTATCCGATTGTGGAATAAAGCTTTTAAACCAAAAAAATTAACAATACAAACAGTTTCGGT
TTCATATGAAAAGAGAGGTTAGGGTTTTAATAAGACTC (A/G) AATAATTAAGATGATTGACGATCATTAC
ATGTTGAAACCTTAACAACCTATCTCGAGTAGAAAAGTTTTTGTGGTTACCTTTCCACCTCAGCAGAA
AGCCACCTTCTTTGATTCTGATTTTCGGACCGAAGGGTTGAGGAAAAGGTTTATGGATATTGTATTTTATA
GACAGAACCATAAAGAAGATGAACCACCGCCTCTGCTTACGTTATTCTTGTATGTATAGGCCGTAGATC
TTCCGATATCTGTGTGGGATATGCAAATTTGG

OM145463

CTCTCACAGGTTTCAGAACTCGAGTGGGAGATCATGGAGAAGAAGACTAGAGTGTGGTCACGAACGGTGA
GAAAAGCCGTTACGACTCTGTTTTACGGCGAACGGATCCTCTCCGACCACGCTTCTCTTCCCTCCGCCGC
GATTCGAGAATCTTCATTCGCTGAGATCACTTACAAAGCGCTTTGGCTTTGTTCTTTTCCCGGAAAC
ATGGCGAAATCGCGGAAAACGCCTGAGAAAATCTTCCCTCACTCTCGACGTTTGCCAGACGATCG (T/C) C
GAGCTTATGCCGAGAAATCGAAGAGGTTTTTTCAGTTACGAATCGACGCTTCCGTCAAATCGCTGATCGCCG
GAACCTAGCGAACCTCGAAGAAAACGGTGATTTTCGATGATCGATGAGTTCGAGTCGTCGATTTCAAGGA
GTCTCGAAAATCGATAATCTCCAACGGCGGAATCCACCAGCTTACAAGATACGTCATGAACTACATCGTC
TTCTCGCCGATTACAGCGACACGCTCGCTAAAAAT

Table SM5. Molecular fingerprints of the 90 accessions (SSR and SNP/CAPS markers).

ACCESSIONS	MT317577	MT317527	MT317610	MT317537	MT317823	DiploTSNP.045	DiploTSNP.062	DiploTSNP.090	DiploTSNP.138	DiploTSNP.213	DiploTSNP.220	DiploTSNP.270	DiploTSNP.290	DiploTSNP.448	ACCESSIONS	MT317577	MT317527	MT317610	MT317537	MT317823	DiploTSNP.045	DiploTSNP.062	DiploTSNP.090	DiploTSNP.138	DiploTSNP.213	DiploTSNP.220	DiploTSNP.270	DiploTSNP.290	DiploTSNP.448
1	248;248	179;181	157;157	163;175	182;184	Y/Y	Y/Y	Y/Y	Y/Y	Y/Y	Y/Y	N/N	Y/N	N/N	58	242;246	179;181	157;157	163;171	184;188	Y/Y	Y/Y	Y/Y	Y/Y	Y/N	Y/N	Y/N	Y/N	N/N
3	248;248	179;181	157;157	163;175	182;184	Y/Y	Y/Y	Y/Y	Y/Y	Y/N	Y/N	N/N	Y/Y	N/N	59	242;252	173;181	157;157	163;171	184;188	Y/Y	Y/Y	Y/Y	Y/Y	Y/Y	Y/Y	N/N	Y/N	N/N
4	242;248	179;181	157;157	163;175	182;184	Y/Y	Y/Y	Y/Y	Y/Y	Y/Y	Y/Y	N/N	Y/N	N/N	60	242;242	167;173	157;157	163;171	184;188	Y/Y	Y/Y	Y/Y	Y/Y	Y/N	Y/Y	N/N	Y/N	N/N
5	242;252	173;181	157;157	163;175	184;188	Y/Y	Y/Y	Y/Y	Y/Y	Y/N	Y/Y	N/N	Y/N	N/N	61	242;246	173;181	157;157	163;171	184;188	Y/Y	Y/Y	Y/Y	Y/Y	Y/Y	Y/Y	N/N	N/N	N/N
6	242;242	173;181	157;157	163;175	184;184	Y/Y	Y/Y	Y/N	Y/N	Y/Y	Y/Y	N/N	N/N	N/N	63	242;246	173;181	157;157	163;173	182;184	Y/Y	Y/Y	Y/Y	Y/Y	Y/Y	Y/Y	N/N	Y/N	N/N
7	242;242	173;181	157;157	163;175	182;184	Y/N	Y/Y	Y/Y	Y/N	Y/Y	Y/Y	N/N	Y/N	N/N	64	242;242	161;161	157;157	163;175	184;184	Y/Y	Y/Y	Y/Y	Y/Y	Y/Y	Y/Y	N/N	Y/N	N/N
8	244;248	173;173	157;157	163;171	156;184	Y/N	Y/Y	Y/N	Y/N	Y/Y	Y/Y	N/N	Y/N	N/N	65	242;242	167;181	157;157	163;171	182;184	Y/Y	Y/Y	Y/Y	Y/Y	Y/Y	Y/Y	N/N	Y/N	N/N
9	242;242	179;181	157;157	163;171	182;184	Y/Y	Y/Y	Y/N	Y/Y	Y/Y	Y/Y	N/N	N/N	N/N	68	X	170;170	157;157	168;300	X	Y/Y	X	X	X	X	X	X	X	X
10	248;248	173;181	157;157	163;171	156;188	Y/Y	Y/Y	Y/N	Y/N	Y/Y	Y/Y	N/N	Y/N	N/N	69	238;248	171;173	157;157	163;171	184;188	Y/Y	Y/Y	Y/N	Y/N	Y/N	Y/N	N/N	Y/N	N/N
12	242;252	179;181	157;157	163;163	182;184	Y/Y	Y/Y	Y/Y	Y/Y	Y/Y	Y/Y	N/N	Y/N	N/N	72	248;248	173;173	157;157	163;175	182;184	Y/Y	Y/Y	Y/N	Y/N	Y/N	Y/Y	N/N	Y/N	N/N
15	248;248	179;181	157;157	163;175	182;184	Y/Y	Y/Y	Y/Y	Y/Y	Y/Y	Y/Y	N/N	N/N	N/N	73	252;252	161;181	157;157	163;163	182;184	Y/Y	Y/Y	Y/Y	Y/Y	Y/N	Y/N	N/N	Y/N	N/N
16	248;248	173;173	157;157	163;175	156;188	Y/Y	Y/Y	Y/N	N/N	Y/Y	Y/N	N/N	Y/N	Y/N	74	242;242	173;173	157;157	163;175	182;184	Y/Y	Y/Y	Y/Y	Y/N	Y/Y	Y/Y	N/N	N/N	N/N
17	244;244	173;173	157;157	163;171	156;184	Y/Y	Y/Y	Y/N	Y/N	Y/Y	Y/Y	N/N	Y/N	N/N	81	242;246	173;181	157;157	163;175	182;184	Y/Y	Y/Y	Y/N	Y/Y	Y/Y	Y/Y	N/N	Y/N	N/N
18	248;248	173;173	157;157	163;175	156;156	Y/Y	Y/Y	Y/N	N/N	Y/Y	Y/Y	N/N	N/N	Y/N	84	242;246	161;181	157;157	163;175	182;184	Y/Y	Y/Y	Y/Y	Y/Y	Y/N	Y/Y	N/N	Y/N	N/N
19	248;248	173;181	157;157	163;171	156;188	Y/Y	Y/Y	Y/N	Y/N	Y/Y	Y/Y	N/N	Y/N	N/N	91	X	173;179	157;157	163;175	182;184	Y/Y	Y/Y	Y/N	Y/Y	Y/Y	Y/Y	N/N	Y/Y	N/N
20	242;248	177;179	157;157	163;175	182;184	Y/Y	Y/Y	Y/Y	Y/Y	Y/N	Y/N	N/N	N/N	Y/N	92	X	161;173	157;157	163;163	186;188	Y/Y	Y/Y	Y/Y	Y/Y	Y/Y	Y/Y	N/N	Y/N	N/N
21	238;242	167;167	157;157	163;171	184;188	Y/Y	Y/Y	Y/Y	Y/Y	Y/N	Y/Y	N/N	N/N	Y/N	97	X	177;179	157;157	163;163	156;184	Y/Y	Y/Y	Y/Y	Y/Y	Y/Y	N/N	N/N	N/N	N/N
22	238;238	167;167	157;157	163;171	184;188	Y/N	Y/Y	Y/Y	Y/Y	Y/N	Y/Y	Y/N	N/N	Y/N	98	248;248	179;181	157;157	163;175	182;184	Y/Y	Y/Y	Y/Y	Y/Y	Y/Y	Y/Y	N/N	Y/N	N/N
23	238;238	171;173	157;157	163;171	186;188	Y/Y	Y/Y	Y/Y	Y/Y	Y/Y	Y/Y	N/N	N/N	Y/N	99	248;248	179;181	157;157	163;175	182;184	Y/Y	Y/Y	Y/Y	Y/Y	Y/N	Y/N	N/N	Y/N	N/N
24	244;244	173;173	157;157	163;171	156;184	Y/Y	Y/Y	Y/N	Y/N	Y/Y	Y/Y	N/N	Y/N	N/N	100	248;248	179;181	157;157	163;175	182;184	Y/Y	Y/Y	Y/Y	Y/Y	Y/Y	Y/N	N/N	N/N	N/N

Table SM5. (Continue.)

26	242;242	181;181	157;157	163;175	211;213	Y/Y	Y/Y	Y/Y	Y/Y	N/N	Y/N	N/N	Y/N	N/N	101	248;248	179;179	157;157	163;175	182;184	Y/Y	Y/Y	Y/Y	Y/Y	N/N	Y/N	N/N	N/N	N/N	
27	242;242	173;181	157;157	163;171	184;184	Y/Y	Y/Y	Y/N	Y/Y	Y/Y	Y/N	N/N	Y/N	N/N	102	242;248	181;181	157;157	163;175	182;184	Y/Y	Y/Y	Y/Y	Y/Y	Y/N	Y/N	N/N	N/N	N/N	
29	X	169;173	135;157	167;171	X	Y/Y	X	X	X	X	X	X	X	X	103	242;242	173;181	157;157	163;175	182;184	Y/Y	Y/Y	Y/Y	Y/Y	Y/N	Y/N	N/N	Y/N	N/N	
32	242;248	177;181	157;157	163;175	184;188	Y/Y	Y/Y	Y/Y	Y/N	Y/N	Y/Y	N/N	Y/N	N/N	104	242;242	173;177	157;157	163;171	184;184	Y/Y	Y/Y	Y/Y	Y/Y	Y/N	Y/Y	N/N	Y/Y	N/N	
35	248;248	179;179	157;157	163;175	184;188	Y/Y	Y/Y	Y/Y	Y/Y	Y/N	Y/N	N/N	N/N	N/N	105	X	162;170	135;157	172;300	X	Y/Y	X	X	X	X	X	X	X	X	X
36	248;248	167;179	157;157	163;175	184;184	Y/Y	Y/Y	Y/N	Y/Y	Y/Y	Y/N	Y/N	N/N	N/N	134	X	173;179	157;157	163;171	184;188	Y/Y	Y/Y	Y/Y	Y/Y	Y/Y	Y/N	N/N	Y/Y	N/N	
37	242;248	179;181	157;157	163;175	182;184	Y/Y	Y/Y	Y/Y	Y/Y	Y/N	Y/N	N/N	Y/N	N/N	146	X	173;179	157;157	163;177	211;213	Y/Y	Y/Y	Y/N	Y/Y	Y/N	Y/Y	N/N	N/N	N/N	
38	238;238	171;173	157;157	163;171	186;188	Y/Y	Y/Y	Y/Y	Y/Y	Y/Y	Y/Y	N/N	N/N	Y/N	152	248;248	173;173	157;157	163;175	182;184	Y/Y	Y/Y	Y/Y	Y/Y	Y/Y	Y/Y	N/N	Y/N	N/N	
39	238;238	167;167	157;157	163;171	184;188	Y/Y	Y/Y	Y/Y	Y/Y	Y/N	Y/Y	N/N	N/N	Y/N	154	240;242	173;179	157;157	163;163	182;184	Y/Y	Y/Y	Y/Y	Y/N	Y/N	Y/N	N/N	N/N	N/N	
40	244;244	173;179	157;157	163;171	182;184	Y/N	Y/Y	Y/N	Y/N	Y/Y	Y/Y	N/N	Y/N	N/N	155	248;248	181;181	157;157	163;171	184;184	Y/Y	Y/Y	Y/Y	Y/Y	Y/N	Y/Y	N/N	N/N	Y/N	
41	242;242	173;179	157;157	163;175	182;184	Y/Y	Y/Y	Y/N	Y/Y	Y/Y	Y/N	N/N	N/N	N/N	156	242;242	177;179	157;157	163;175	184;188	Y/Y	Y/Y	Y/Y	Y/Y	Y/Y	Y/Y	N/N	Y/N	N/N	
42	246;248	173;181	157;157	163;175	184;188	Y/Y	Y/Y	Y/Y	Y/N	Y/Y	Y/Y	N/N	N/N	N/N	157	242;242	173;179	157;157	163;175	182;184	Y/Y	Y/Y	Y/Y	Y/Y	Y/N	Y/Y	N/N	Y/N	N/N	
43	242;246	173;181	157;157	163;175	184;184	Y/Y	Y/Y	Y/N	Y/N	Y/Y	Y/Y	N/N	Y/N	N/N	158	238;238	171;173	157;157	163;171	186;188	Y/Y	Y/Y	Y/Y	Y/Y	Y/Y	Y/Y	N/N	N/N	Y/N	
44	240;248	173;181	157;157	163;163	184;184	Y/Y	Y/Y	Y/N	Y/N	Y/Y	Y/N	Y/N	N/N	N/N	159	242;242	171;179	157;157	163;171	186;186	Y/Y	Y/Y	Y/Y	Y/Y	Y/Y	Y/N	N/N	N/N	N/N	
45	246;248	173;181	157;157	163;175	182;184	Y/Y	Y/Y	Y/Y	Y/Y	Y/N	Y/Y	N/N	N/N	N/N	160	248;248	179;181	157;157	163;175	182;184	Y/Y	Y/Y	Y/Y	Y/Y	Y/N	Y/Y	N/N	Y/N	N/N	
46	242;242	173;179	157;157	163;171	182;184	Y/Y	Y/Y	Y/N	Y/Y	Y/Y	Y/Y	N/N	Y/N	N/N	161	248;248	179;181	157;157	163;175	182;184	Y/Y	Y/Y	Y/Y	Y/Y	Y/N	Y/Y	N/N	Y/N	N/N	
47	242;246	161;181	157;157	163;171	182;184	Y/Y	Y/Y	Y/N	Y/Y	Y/Y	Y/Y	Y/N	Y/N	N/N	162	242;248	177;181	157;157	163;175	182;184	Y/Y	Y/Y	Y/Y	Y/Y	Y/N	N/N	N/N	Y/N	Y/N	
48	242;242	179;181	157;157	163;175	184;188	Y/Y	Y/Y	Y/N	Y/N	Y/N	Y/N	N/N	Y/N	N/N	163	242;242	181;181	157;157	163;171	182;184	Y/Y	Y/Y	Y/Y	Y/Y	Y/N	Y/Y	N/N	N/N	N/N	
49	242;246	173;179	157;157	163;175	184;188	Y/Y	Y/Y	Y/N	Y/Y	Y/N	Y/Y	N/N	N/N	N/N	164	238;238	167;167	157;157	163;171	184;188	Y/Y	Y/Y	Y/Y	Y/Y	Y/N	Y/Y	N/N	N/N	Y/N	
51	242;252	179;181	157;157	163;163	182;184	Y/Y	Y/Y	Y/N	Y/Y	Y/Y	Y/N	N/N	Y/N	N/N	165	248;248	173;173	157;157	163;175	156;156	Y/Y	Y/Y	Y/N	N/N	Y/Y	Y/Y	N/N	N/N	Y/N	
52	242;242	181;181	157;157	163;175	186;188	Y/Y	Y/Y	Y/N	Y/N	Y/Y	Y/N	N/N	N/N	N/N	166	242;252	179;181	157;157	163;163	182;184	Y/Y	Y/Y	Y/N	Y/Y	Y/Y	Y/N	N/N	Y/N	Y/N	
54	242;246	161;181	157;157	163;173	184;213	Y/Y	Y/Y	Y/N	Y/Y	Y/Y	Y/N	N/N	N/N	N/N	167	248;248	179;179	157;157	163;163	182;184	Y/Y	Y/Y	Y/Y	Y/Y	N/N	Y/Y	N/N	Y/Y	N/N	
55	238;242	161;177	157;157	163;173	156;184	Y/Y	Y/Y	Y/Y	Y/Y	Y/Y	Y/Y	N/N	Y/Y	N/N	168	242;242	173;181	157;157	163;175	184;188	Y/Y	Y/Y	Y/Y	Y/Y	Y/N	Y/N	N/N	Y/N	N/N	
56	X	179;179	157;157	163;171	184;186	Y/Y	Y/Y	Y/Y	Y/Y	N/N	Y/Y	N/N	Y/Y	N/N	169	248;248	179;179	157;157	163;175	182;184	Y/Y	Y/Y	Y/Y	Y/Y	Y/N	N/N	N/N	N/N	N/N	
57	242;244	161;173	157;157	163;171	182;184	Y/Y	Y/Y	Y/Y	Y/Y	Y/Y	Y/N	N/N	N/N	N/N	170	242;246	173;177	157;157	163;163	184;184	Y/Y	Y/Y	Y/Y	Y/Y	Y/N	Y/N	N/N	N/N	N/N	

Table SM6. Five hundred SNP loci retrieved from the *D. tenuifolia* genome assembly “UAlgDiploT.01” published in the NCBI database (GenBank 144979 to 145478).

Diplotaxis tenuifolia clone DiploTSNP.001 genomic sequence

GenBank: OM144979.1

[FASTA](#) [Graphics](#)

[Go to:](#)

LOCUS OM144979 509 bp DNA linear PLN 28-NOV-2022
DEFINITION Diplotaxis tenuifolia clone DiploTSNP.001 genomic sequence.
ACCESSION OM144979
VERSION OM144979.1
KEYWORDS .
SOURCE Diplotaxis tenuifolia
ORGANISM [Diplotaxis tenuifolia](#)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliopsida; eudicotyledons; Gunneridae;
Pentapetalae; rosids; malvids; Brassicales; Brassicaceae;
Brassicaceae; Diplotaxis.
REFERENCE 1 (bases 1 to 509)
AUTHORS Reis,J.P., Pereira,R.J., Coelho,P.S. and Leitao,J.M.
TITLE Direct Submission
JOURNAL Submitted (28-DEC-2021) FCT, Universidade Do Algarve, Campus de
Gambelas, Faro 8005-139 Faro, Portugal
COMMENT ##Assembly-Data-START##
Sequencing Technology :: Sanger dideoxy sequencing
##Assembly-Data-END##
FEATURES Location/Qualifiers
source 1..509
/organism="Diplotaxis tenuifolia"
/mol_type="genomic DNA"
/db_xref="taxon:264416"
/clone="DiploTSNP.001"

Diplotaxis tenuifolia clone DiploTSNP.500 genomic sequence

GenBank: OM145478.1

[FASTA](#) [Graphics](#)

[Go to:](#)

LOCUS OM145478 516 bp DNA linear PLN 28-NOV-2022
DEFINITION Diplotaxis tenuifolia clone DiploTSNP.500 genomic sequence.
ACCESSION OM145478
VERSION OM145478.1
KEYWORDS .
SOURCE Diplotaxis tenuifolia
ORGANISM [Diplotaxis tenuifolia](#)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliopsida; eudicotyledons; Gunneridae;
Pentapetalae; rosids; malvids; Brassicales; Brassicaceae;
Brassicaceae; Diplotaxis.
REFERENCE 1 (bases 1 to 516)
AUTHORS Reis,J.P., Pereira,R.J., Coelho,P.S. and Leitao,J.M.
TITLE Direct Submission
JOURNAL Submitted (28-DEC-2021) FCT, Universidade Do Algarve, Campus de
Gambelas, Faro 8005-139 Faro, Portugal
COMMENT ##Assembly-Data-START##
Sequencing Technology :: Sanger dideoxy sequencing
##Assembly-Data-END##
FEATURES Location/Qualifiers
source 1..516
/organism="Diplotaxis tenuifolia"
/mol_type="genomic DNA"
/db_xref="taxon:264416"
/clone="DiploTSNP.500"

7. Publications achieved in the scope of this dissertation

Technical Report

MÍLDIO NA RÚCULA SELVAGEM – O PORQUÊ DE INVESTIGAR ESTA DOENÇA

A rúcula (*Diplotaxis* spp. e *Eruca* spp.) é uma hortícola de folhas “baby-leaf” muito apreciada pelas suas características organolépticas e benefícios para a saúde. Pertence à família das Brassicáceas e a rúcula selvagem (*Diplotaxis tenuifolia* (L.) DC.) é a mais apreciada pelos consumidores. Nas últimas décadas, tem tido uma grande expansão e, atualmente, apresenta uma relevância económica crescente a nível mundial devido à utilização na indústria de saladas minimamente processadas. Apesar de ser uma cultura bem-adaptada ao clima mediterrânico, a sua produção intensiva, ao ar livre ou em estufa, favorece a ocorrência de míldio. Esta doença causa perdas de produção e prejuízos avultados, pelo que a identificação de variedades resistentes à infeção e a adoção de práticas agronómicas sustentáveis são importantes para o aumento de rendimento dos produtores.

Paula S. Coelho¹, Ana L. Pereira¹, Corina Carranca¹, Paula Scott¹, Violeta Lopes¹, Clarisse Boto², João Reis³, José Leitão³

¹ Instituto Nacional de Investigação Agrária e Veterinária



² Vitacress Portugal SA



³ Universidade do Algarve



Introdução

O mercado mundial de hortícolas de folha, colhidas ainda jovens e conhecidas por “baby leaf”, cresceu substancialmente desde os anos 90 devido a um maior interesse por uma alimentação saudável, com acesso a alimentos diversificados, seguros, nutritivos e de qualidade. A rúcula é uma planta espontânea e é um caso recente de sucesso de domesticação de espécies vegetais. A rúcula selvagem foi utilizada desde a antiguidade para diferentes fins, como alimento, na cosmética e na medicina tradicional^[1].

O aumento da procura de saladas pré-embaladas e prontas a comer, exigentes em produtos inovadores e de elevada qualidade, fez disparar, numa primeira fase, o consumo de alface, a que se sucederam outras hortícolas, nomeadamente a rúcula selvagem^[2] cujas folhas, inconfundíveis pelo aroma característico e sabor picante, podem ser consumidas cruas ou cozidas. Nas últimas décadas, devido ao aumento da procura, esta espécie vegetal adquiriu importância económica na indústria das hortícolas a nível mundial, assistindo-se ao aumento exponencial do número de produtores e da área cultivada. Porém, a produção intensiva, tanto em estufa como ao ar livre, potenciou o surgimento de várias doenças, nomeadamente do míldio, colocando em risco o rendimento dos produtores.

A rúcula – O que é?

São designadas de rúcula as diferentes espécies dos géneros *Diplotaxis* e *Eruca* que pertencem à família das Brassicáceas. Para consumo humano são usadas quatro espécies, mas apenas duas são comercializadas em larga escala: a rúcula selvagem (*Diplotaxis tenuifolia* (L.) DC.) e a rúcula cultivada (*Eruca sativa* (L.) Cav.). São plantas nativas da região mediterrânica, mas diferentes espécies do género *Diplotaxis* foram encontradas em países como a Índia, Paquistão, Cabo Verde e Nepal^[3]. A espécie *D. tenuifolia* é economicamente mais importante e a mais consumida na Europa, América do Norte e Austrália, enquanto que a *E. sativa* é amplamente cultivada no Médio Oriente e Sul da Ásia^[4].

Posters

II International Meeting of the Portuguese Society of Genetics (Online)

II INTERNATIONAL MEETING OF THE PORTUGUESE SOCIETY OF GENETICS

Abstract book

IMPSG 2021
1st – 2nd July 2021

Universidade de Trás-os-Montes e Alto Douro
Vila Real, Portugal

utad UNIVERSIDADE
DE TRÁS-OS-MONTES
E ALTO DOURO



Molecular characterization of a wild rocket (*Diplotaxis tenuifolia* L. DC) germplasm collection

J Reis¹, P Coelho³, R Pereira^{1,2}, J Leitão^{1*}

¹Laboratory of Genomics and Genetic Improvement (LGGI), FCT, UAIG, Campus de Gambelas, Edif 8, 8005-139, Faro, Portugal. ²Corte velada, Monte Ruivo CX 552X 8600-237 Odiáxere, Portugal; ³ INIAV, Av. Da República, Quinta do Marquês, 2780-157 Oeiras, Portugal.

*jleitao@ualg.pt

A germplasm collection of wild rocket (*Diplotaxis tenuifolia* L. DC) was established in the frame of the FCT funded research project: "PTDC/ASP-PLA/28963/2017 Remirúcula - Resistance characterization to downy mildew in wild rocket crop", aimed at the gathering of a good representation of the genetic diversity of the species, while allowing the identification of genotypes resistant to downy mildew (DM), a devastating disease of this popular baby leaf crop, caused by the oomycete *Hyaloperonospora* sp. Twenty out of 500 microsatellite loci previously uploaded by the LGGI to the genomic databases (www.ncbi.nlm.nih.gov/nuccore/?term=Diplotaxis%20microsatellite) were selected for establishment of SSR markers. After preliminary tests, 10 SSR markers were retained for further use, in combination with 7 RAPD primers, to assess the genetic diversity of the collection and the genetic relationships among the 100 germplasm accessions. The use of capillary gel electrophoresis for the SSR analysis allows the identification of an unequivocal individual molecular fingerprint for all accessions.

III International Meeting of the Portuguese Society of Genetics



III International Meeting of the Portuguese Society of Genetics

27th and 28th June 2022

Auditorium of the Colégio Espírito Santo, University of Évora

Book of abstracts

(Draft Version)

Assessment of Downy Mildew Resistance and Genetic Relationships Among Wild Rocket (*Diplotaxis tenuifolia* (L.) DC) Accessions

Paula Coelho¹, João Reis², Ana Pereira¹, Aliana Vairinhos², Violeta Lopes³, José Leitão²

¹INIAV - Instituto Nacional de Investigação Agrária e Veterinária, Quinta do Marquês, Av. da República, 2784-505 Oeiras, Portugal

²MED - Mediterranean Institute for Agriculture, Environment and Development & CHANGE - Global Change and Sustainability Institute, FCT, Universidade do Algarve, Campus de Gambelas, 8005-139 Faro, Portugal

³BPGV/INIAV - Banco Português de Germoplasma Vegetal, Polo de Inovação de Braga, 4700-859 Braga, Portugal

Email: jleitao@ualg.pt, paula.coelho@iniav.pt, a61179@ualg.pt

Wild rocket (*Diplotaxis tenuifolia* (L.) DC.) is an herbaceous species produced as baby leaf salad, which importance is increasing as a response to the growing preference of the consumers over the "cultivated" rocket (*Eruca* spp.). Downy mildew (DM) disease, elicited by the oomycete *Hyaloperonospora* sp., is one of the main causes of yield losses in wild rocket cultivation. From a point of view of food safety, environmental concerns, and viability of organic farming production the identification of resistant genotypes is a priority to this industry. A set of 29 *D. tenuifolia* accessions, and 1 *Eruca* spp. accession, were screened at the seedling stage, for resistance to the *Hyaloperonospora* sp. isolate DS, collected on *D. tenuifolia* infected plants. The plant/pathogen interaction phenotype (IP) was assessed at the cotyledon and first two leaves stage, and the accessions were classified as resistant (R), partially resistant (PR), susceptible (S) and highly susceptible (HS) in both stages. The *Eruca* spp. accession was resistant at both stages (R/R), while the *Diplotaxis* accessions were, respectively, 1 (R/R), 3 (PR/R), 8 (S/PR), 6 (HS/PR), 9 (HS/PR) and 2 (HS/HS). A significant correlation ($r=0,865896$, $P<0.001$) was found between the disease index values exhibited in cotyledons vs. young leaves. No plants R or PR in cotyledons became susceptible in the first leaves. Molecular analyses were performed using 7 RAPD and 5 ISSR markers, which have amplified 110 clearly scorable markers. As expected for genotypes from the same species, the genetic similarity among the *Diplotaxis* accessions varied from 0,751 to 0,994. The *Eruca* accession exhibited very low genetic similarity (0,171) to the wild rocket accessions. Some of the accessions that exhibited high genetic similarity were provided by the same breeding company.

This work was funded by the project: PTDC/ASP-PLA/28963/2017- REMIRucula - Resistance characterization to downy mildew in wild rocket crop.

Unequivocal identification of *Diplotaxis tenuifolia* (L.) DC accessions by species specific SSR and SNP markers

João Reis¹, Ricardo Pereira², Paula Coelho³, José Leitão¹

¹MED - Mediterranean Institute for Agriculture, Environment and Development & CHANGE - Global Change and Sustainability Institute, FCT, Universidade do Algarve, Campus de Gambelas, 8005-139 Faro, Portugal

²Corte Velada Investimentos, Odiáxere, Portugal

³INIAV - Instituto Nacional de Investigação Agrária e Veterinária, Quinta do Marquês, Av. da República, 2784-505 Oeiras, Portugal

Email: jleitao@ualg.pt, paula.coelho@iniav.pt, a61179@ualg.pt

Wild rocket (*Diplotaxis tenuifolia* (L.) DC) is currently commercialized as baby leaf salad in individual packages or as a major component of salad leaf mixtures. The high content of glucosinolates and the release of volatile isothiocyanates, provides this crop with a spicy and bitter taste, and a strong aroma, that determines the growing consumer preference over the "cultivated" rocket (*Eruca sativa* L.). A germplasm collection of wild rocket, that includes some accessions of *Eruca* spp., is being established at the INIAV (Oeiras), which requires unequivocal identification of the gathered accessions. For this purpose, the most appropriated are the DNA markers technologies, particularly, genome specific Single Sequence Repeats (SSR) and SNPs (Single Nucleotide Polymorphisms) markers. Here, we describe the first attempt of unequivocal identification of 29 wild rocket and 1. Twenty out of 500 SSR loci and 20 out of 500 SNP loci, mined out from the first genome assembly of *D. tenuifolia* (ncbi.nlm.nih.gov/assembly/GCA_014822095.1), were selected and tested as SSR and SNP markers. A total of 5 SSR and 9 SNP markers were retained for further accession identification. The amplified 80 SSR markers were analyzed by capillary gel electrophoresis and the amplified SNP markers were analyzed as 36 TaqI generated CAPS (cleaved amplified polymorphic sequence) markers, revealed by agarose gel electrophoresis. As expected, small groups of accessions displayed identical DNA fingerprints (100% similarity), confirming those accessions as the same commercial variety, although provided by different producers. The *Eruca* spp. accession exhibited very different and unique molecular patterns, which resulted in the lowest (0,118) genetic similarity towards the remaining accessions.

This work was funded by the project: PTDC/ASP-PLA/28963/2017- REMIRucula - Resistance characterization to downy mildew in wild rocket crop.



Downy mildew evaluation in wild rocket genotypes [*Diplomatix tenuifolia* (L.) DC.] under field and controlled conditions

P.S. Coelho¹, A.L. Pereira¹, J. Reis², C. Carranca¹, V.R. Lopes³ and J.M. Leilão²
¹INIAV, Oeiras, Portugal; ²MED & CHANGE, FCT/UAIG, Faro, Portugal; ³INIAV/BPGV, Braga, Portugal

Contact: paula.coelho@iniav.pt



Introduction

Wild rocket (*Diplomatix tenuifolia*) belongs to the Brassicaceae family is a much appreciated and healthy baby-leaf native to the Mediterranean region. Wild rocket downy mildew (DM) is a devastating foliar disease, especially in temperate climates, caused by the oomycete *Hyaloperonospora* spp. that can affect plants at all stages. The objective of our study is to compare seedling and adult plants DM resistance in a group of *D. tenuifolia* accessions, to determine whether seedling resistance tested in controlled environment is a good indicator of adult plant behaviour in the field.

Materials and Methods

Plant material: Forty *D. tenuifolia* accessions from different origins were tested in a commercial producer located at Odemira (Vilaverde Portugal, SA). **Growth conditions:** The plants were grown in a field trial (polytunnel) during the period from October to December 2021 and were naturally infected with *Hyaloperonospora* spp. (Fig. 1). **Screening method:** Seedlings were evaluated in a previous study (data under publication) with 14 days using 0-6 classes in controlled environment. Adult plants were assessed 71 days after sowing using 0-5 classes, taking into account the number of infected leaves per plant and the size of sporulating *Hyaloperonospora* spp. lesions (small $\leq 0.1\text{cm}$ and large >math>\geq 0.1\text{cm}</math>) (Fig. 2).

Results

- Twenty-eight wild rocket commercial varieties and 12 collected accessions from other origin (prospecting missions and gene banks) were evaluated (Table 1).
- Sixteen accessions were identified with an interesting downy mildew resistance response in adult phase.

Table 1. Adult plant evaluation to downy mildew disease in the field.

Phenotypic categories of DM evaluation	Commercial variety	Collected accession	Total no. of accessions
Resistant (R)	11	5	16
Partially Resistant (PR)	6	2	8
Susceptible (S)	8	4	12
Highly Susceptible (HS)	3	1	4



Fig. 2. Downy mildew in wild rocket (Highly Susceptible).



Fig. 1. Adult wild rocket DM evaluation in a polytunnel.

- A significant coefficient of correlation was observed ($r=0.628$, $P=0.000$) between seedling and adult plant DM resistance.

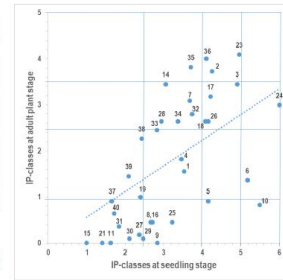


Fig. 2. Comparison of seedling and adult plant DM resistance.

- In general, similarly to other brassicas, wild rocket accessions were more susceptible at seedling than at adult stage. No resistant accessions at seedling stage and susceptible at adult plant were identified.
- On the contrary, two accessions with a highly susceptible response at seedling stage presented a resistant and a partially resistant response at adult phase, respectively, accessions no. 10 and 6.

Conclusions and perspectives

Our results suggest that downy mildew resistance observed at the seedling stage is a good indication of wild rocket resistance behaviour at field adult plant. It is important to increase the resistance of varieties to pests and diseases, ensuring the production of healthy and safe food, less dependent on chemical products. The accessions classified as resistant at seedling stage (14 days) exhibited a good DM resistance as adult plants and can be further exploited in breeding programmes to downy mildew resistance.



Funding: This research was supported by project REMIUCULA - "Resistance characterization to downy mildew in wild rocket crop" (PTDC/ ASP-PLA/28963/2017) and FCT I.P., through the R&D Unit "GREEN-IT - Bioresources for Sustainability" (UIDB/04551/2020 and UIDP/04551/2020).



ARTICLE

Agronomic Application of Genetic Resources

Downy mildew resistance and genetic variability in a wild rocket germplasm collection

Paula S. Coelho¹ | João M. Reis² | Ana L. Pereira¹ | Aliana Vairinhos² |
Violeta Lopes¹ | José M. Leitão²¹INIAV – Instituto Nacional de Investigação Agrária e Veterinária, Av. da República, 2784-505 Oeiras, Portugal²MED – Mediterranean Institute for Agriculture, Environment and Development & CHANGE – Global Change and Sustainability Institute, Faculdade de Ciências e Tecnologia, Universidade do Algarve, Campus de Gambelas, 8005-139 Faro, Portugal

Correspondence

Paula S. Coelho, INIAV – Instituto Nacional de Investigação Agrária e Veterinária, Av. da República, 2784-505 Oeiras, Portugal.
Email: paula.coelho@iniav.ptJosé M. Leitão, MED – Mediterranean Institute for Agriculture, Environment and Development & CHANGE – Global Change and Sustainability Institute, Faculdade de Ciências e Tecnologia, Universidade do Algarve, Campus de Gambelas, 8005-139 Faro, Portugal.
Email: jleitao@ualg.pt

Assigned to Associate Editor Faheem Baloch.

Funding information

REMIRUCULA project "Resistance characterization to downy mildew in wild rocket crop", Grant/Award Number: PTDC/ASP-PLA/28963/2017; FCT - Fundação para a Ciência e a Tecnologia, I.P., through the R&D Unit "GREEN-IT - Bioresources for Sustainability", Grant/Award Numbers: UIDB/04551/2020, UIDP/04551/2020

Abstract

One hundred accessions of a "core collection" of *Diplotaxis tenuifolia* (L.) DC. and *Eruca* spp. were screened at seedling stage for resistance to downy mildew. Accessions tested at the seedling stage were assigned to 0–6 interaction phenotypes. All cultivated rocket (*Eruca* spp.) accessions exhibited a resistant (R) response both in cotyledons and in young leaves. The wild rocket (*D. tenuifolia*) accessions exhibited higher susceptibility in cotyledons than in the 1st and 2nd leaves, with 16 and 47 accessions classified as resistant or partially resistant (PR) in the cotyledon and in leaves stages, respectively. Only three wild rocket accessions displayed an R phenotype in cotyledons and leaves. The most frequent response in cotyledons vs. leaves was the highly susceptible/susceptible (HS/S) combination (33 accessions), followed by the S/PR combination (18 accessions). A significant correlation ($r = 0.917$, $P < .000$) was observed between the disease index in cotyledons and leaves. The molecular markers analyses revealed a wide genetic distance between *Diplotaxis* and *Eruca*, which gather in two clearly separated species clusters. The molecular variability is accompanied by a wide diversity of interactions with the pathogen isolate. The closest similarities among *D. tenuifolia* accessions were found in accessions provided by the same breeding company. Future studies will be focused on two main objectives: (a) the assessment of the accessions behavior that have evidenced an R/R, S/PR, and HS/PR cotyledon and leaf response under greenhouse or field production and (b) the genome mapping of genetic features that provide downy mildew resistance.

Abbreviations: DI, disease severity index; HS, highly susceptible; ISSR, inter-simple sequence repeats; PCR, polymerase chain reaction; PR, partially resistant; R, resistant; RAPD, random amplified polymorphic DNA; S, susceptible; UPGMA, unweighted pair-group method with arithmetic averages.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2022 The Authors. *Agronomy Journal* published by Wiley Periodicals LLC on behalf of American Society of Agronomy.

Downy mildew evaluation in wild rocket genotypes [*Diplotaxis tenuifolia* (L.) DC.] under field and controlled conditions

P.S. Coelho¹, A.L. Pereira¹, J. Reis², C. Carranca¹, V.R. Lopes³ and J.M. Leitão²

¹INIAV - Instituto Nacional de Investigação Agrária e Veterinária, Av. da República, Quinta do Marquês, 2784-505 Oeiras, Portugal (E-mail: paula.coelho@iniav.pt); ²MED - Mediterranean Institute for Agriculture, Environment and Development & CHANGE - Global Change and Sustainability Institute, Faculdade de Ciências e Tecnologia, Universidade do Algarve, Campus de Gambelas, 8005-139 Faro, Portugal (E-mail: jleitao@ualg.pt); ³INIAV/BPGV, Banco Português de Germoplasma Vegetal, S. Pedro de Merelim, 4700-859 Braga, Portugal

Abstract

Wild rocket downy mildew (DM) is a foliar disease caused by the oomycete *Hyaloperonospora* spp. that limits the production of brassicas, especially in temperate climates, infecting plants at all growth stages. During 2021 autumn season, a field trial (polytunnel) with a set of forty wild rocket accessions from different origins (harvest missions, gene banks, commercial varieties) was installed in a commercial enterprise located in Odemira, southern Portugal. The plants were naturally infected by *Hyaloperonospora* spp. and were visually evaluated 71 days after sowing. Field observations of adult plants evidenced differences among accessions concerning the DM resistance and agronomic traits (e.g. flowering date, vigour, plant habit, leaf serration). Sixteen wild rocket accessions were identified with an interesting resistance response to DM in adult phase. In a previous study, the wild rocket accessions were tested to *Hyaloperonospora* isolate D5 at seedling stage under controlled conditions. A significant coefficient of correlation was observed ($r=0.628$, $P=0.000$, $n=36$) by comparing the plants in both growth phases. In general, the accessions were more resistant at adult than at seedling stage. No accessions were resistant at seedling and susceptible at adult stage. However, two accessions with a highly susceptible response at seedling stage presented a resistant and a partially resistant response at adult phase under field conditions. Our results suggest that DM resistance observed at the seedling stage is a good indication of wild rocket resistance behaviour at field adult plant. The most promising accessions identified as resistant at seedling stage may be further exploited in breeding programmes.

Keywords: field resistance, germplasm collection, oomycete pathogen, seedling and adult plant resistance

INTRODUCTION

The genus *Diplotaxis* is native to the Mediterranean region and covers about 30 different species (Gómez-Campo, 1999). Wild rocket grows in the Mediterranean basin up to 1000 m altitude, preferably on sandy and calcareous soils along the edges of the roads, in cultivated or abandoned areas (Bianco, 1995). Since ancient times, several species of *Diplotaxis* genus have been consumed as vegetables in fresh salads within the Mediterranean diet, apart from other purposes like cosmetic and medical uses (Cavaiuolo and Ferrante, 2014).

Rocket baby leaf gained economic importance in the horticultural industry in the last decades, with an exponential increase of the cultivated area and number of producers, with wild rocket (*D. tenuifolia*) as the most consumed species in Europe, North America and Australia (Løkke et al., 2012). *D. tenuifolia* leaves have a spicy, strong and bitter taste due to the high levels of glucosinolates, and their peculiar pungent aroma depends on the release of volatile isothiocyanates (Pasini et al., 2011). The consumer's acceptance is greatly

Paula Santos Coelho

De: symposiacontributions@ishs.org
Enviado: 11 de junho de 2022 09:20
Para: Paula Santos Coelho
Assunto: Acta Horticulturae - fulltext accepted : Downy mildew...

Dear author,

This is to confirm that your article:

Downy mildew evaluation in wild rocket genotypes [*Diplotaxis tenuifolia* (L.) DC.] under field and controlled condition has been reviewed and is accepted by the editorial board of:
International Symposium on Sustainable Control of Pests and Diseases for publication in *Acta Horticulturae*

Presenting Author: Dr. Paula S Coelho paula.coelho@iniav.pt

Symposium details + contact information are available from:
<https://www.ishs.org/symposium/758>