



MASTERARBEIT | MASTER'S THESIS

Titel | Title

Die Chalkonsynthese in der Gattung *Origanum*

verfasst von | submitted by

Vanessa Pfalz BSc

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of

Magistra pharmaciae (Mag. pharm.)

Wien | Vienna, 2024

Studienkennzahl lt. Studienblatt |
Degree programme code as it appears on the
student record sheet:

UA 066 605

Studienrichtung lt. Studienblatt | Degree
programme as it appears on the student
record sheet:

Masterstudium Pharmazie

Betreut von | Supervisor:

Dr. Johannes Novak

Abstract

Flavonoide stellen eine bedeutende Klasse sekundärer Metabolite dar, die eine Vielzahl pharmakologischer Wirkungen entfalten. Verschiedene *Origanum*-Arten, die seit Jahrhunderten in der traditionellen Medizin des östlichen Mittelmeerraums genutzt werden, gelten auch in der Krebsforschung als vielversprechend. Ihre heilenden Effekte werden teilweise bestimmten Flavonoid-Verbindungen und Komponenten des ätherischen Öls zugeschrieben. Die Biosynthese dieser Flavonoide ist weitgehend erforscht, insbesondere die Rolle der Chalkonsynthase (CHS) als erstes spezifisches Enzym in diesem Prozess. Jedoch wurden in der Gattung *Origanum* noch keine Flavonoidgene beschrieben.

Im Rahmen dieser Studie wurden die genomischen DNA-Sequenzen der CHS in fünf *Origanum*-Proben aus vier Arten (*Origanum majorana*, *Origanum dubium*, *Origanum syriacum* und *Origanum onites*) mithilfe des Gateway-Klonierungsverfahrens und Genexpressionsanalysen untersucht. Ziel war es, die genetische Variabilität sowohl innerhalb der Proben als auch zwischen den Arten zu erfassen. Die beiden Exons, mit Längen von 179 bp und 986 bp, wurden durch ein Intron variabler Länge (384–398 bp) getrennt. Im Vergleich zum Intron (2,50 SNPs pro 100 bp) wiesen die Exons eine geringere Polymorphismus-Rate auf (0,56 bzw. 0,70 SNPs pro 100 bp) auf. Insgesamt wurden 18 SNPs und 14 Haplotypen (= Multilocusstypen, MLGs) identifiziert. Von den 18 Polymorphismus-Stellen war nur einer nicht-synonym, wobei die Aminosäureveränderung keine kritischen Bereiche des CHS-Enzyms betraf.

Diese Untersuchung bestätigt bestehende Hypothesen zur genetischen Variabilität, zur Verteilung von Polymorphismen, erste Hinweise zur Genkopienanzahl, zum Selektionsdruck sowie zur evolutionären Bedeutung des CHS-Enzyms. Die Phylogenie basierend auf dem CHS-Gen bestätigte bisher bekannte Verwandtschaftsbeziehungen zwischen den Arten.

Die gewonnenen Sequenzdaten liefern wertvolle und zufriedenstellende Ergebnisse, die eine solide Grundlage für weiterführende Forschungsansätze bilden.

Abstract (English)

Flavonoids represent a significant class of secondary metabolites that exhibit a wide range of pharmacological effects. Various *Origanum* species, used as remedies in the traditional medicine of the Eastern Mediterranean region for centuries, are also considered promising in cancer research. Their healing effects partly rely on specific flavonoid compounds and components of the essential oil. The biosynthesis of these flavonoids is largely known, particularly the role of chalcone synthase (CHS) as the first specific enzyme in this process. In the genus *Origanum*, however, no flavonoid genes were described so far.

In this study, the genomic DNA sequences of CHS of five *Origanum* samples from four species (*Origanum majorana*, *Origanum dubium*, *Origanum syriacum*, and *Origanum onites*) were studied using the Gateway cloning method. The aim was to describe the genetic variability within the samples as well as between the species. The two exons, measuring 179 bp and 986 bp, were separated by an intron of variable length (384–398 bp). The exons showed a lower polymorphism rate (0.56 and 0.70 SNPs per 100 bp, respectively) compared to the intron, which showed higher genomic diversity (2.50 SNPs per 100 bp). A total of 18 SNPs and 14 haplotypes (= multilocus genotypes, MLGs) were identified. Of the 18 polymorphism sites, only one was non-synonymous, however the amino acid change did not affect critical regions of the CHS enzyme.

This investigation confirms existing hypotheses regarding genetic variability, the distribution of polymorphisms, gene copy number, selection pressure, and the evolutionary significance of the CHS enzyme. The phylogeny of the CHS genes confirms the known relationships between the *Origanum* species of our sample set.

The obtained sequence data provide valuable and satisfactory results, offering a solid foundation for further research approaches.

Abkürzungsverzeichnis

<i>OM</i> – <i>Origanum majorana</i>	NCBI – National Center for Biotechnology Information
<i>OD</i> – <i>Origanum dubium</i>	BLAST – Basic Local Alignment Search Tool
<i>OS</i> – <i>Origanum syriacum</i>	Expasy – Expert Protein Analysis System
<i>OO</i> – <i>Origanum onites</i>	<i>E. coli</i> – <i>Escherichia coli</i>
SLS – Natriumlaurylsulfat	MEGA – Molecular Evolutionary Genetics Analysis
CTAB – Cetyltrimethylammoniumbromid	CHS – Chalkonsynthase
U/min – Umdrehungen pro Minute	SNP – Single Nucleotide Polymorphism
SBS – Styrol-Butadien-Blockcopolymere	MLG – Multilocus-Genotyp = Haplotyp
TE – Tris-EDTA	DNA – Desoxyribonukleinsäure
EDTA – Ethylendiamintetraessigsäure	RNA – Ribonukleinsäure
SB – Natriumborat	mRNA – Boten-Ribonukleinsäure
TAE – Tris-Acetat-EDTA	AA – Amino acid Sequenz
PCR – Polymerase-Ketten-Reaktion	Indel – Insertion-Deletion
DMSO – Dimethylsulfoxid	T – Thymin
dNTP – Desoxyribo-Nukleotid-Triphosphat	U – Uracil
NIH – National Institute of Health	C – Cytosin
SNP – Single Nucleotid Polymorphism	A – Adenin
NTC – No-Template Control	G – Guanin
NaCl – Natriumchlorid	MLM – Maximum-likelihood Methode
LB – Luria-Bertani	LD – Linkage Disequilibrium = genetische Kopplung
CaCl ₂ – Calciumchlorid	bp – Basenpaare
OD – optische Dichte	AS – Aminosäure
S.O.C. – superoptimales Medium mit katabolischem Repressor	

Inhaltsverzeichnis

1 Einleitung	7
1.1 Gattung <i>Origanum</i> (<i>Lamiaceae</i>)	7
1.1.1 <i>Origanum majorana</i>	8
1.1.2 <i>Origanum dubium</i>	10
1.1.3 <i>Origanum syriacum</i>	11
1.1.4 <i>Origanum onites</i>	13
1.2 Grundlagen der Flavonoide	15
1.2.1 Flavonoid-Biosynthese	16
1.2.2 Chalkonsynthase (CHS)	19
1.3 Methodik-Wahl für die vorliegende Genexpressionsstudie	23
1.3.1 Proteinbiosynthese in Pflanzen	23
1.3.2 DNA-Extraktion	23
1.3.3 PCR	23
1.3.4 Agarose-Gelelektrophorese	24
1.3.5 Gateway-Klonierungsverfahren	24
1.4 Polymorphismus	25
1.5 Fragestellung	26
2 Materialien und Methoden	27
2.1 Probenmaterial	27
2.2 DNA-Extraktion	28
2.2.1 Probenvorbereitung	28
2.2.2 Extraktion der DNA	28
2.2.3 Kontrolle und Vermessung der DNA-Extrakte	29
2.3 Agarose-Gelelektrophorese	30
2.3.1 SB-Gel Herstellung	31
2.3.2 TAE-Gel Herstellung	31
2.4 PCR	32
2.4.1 Primer Design	32
2.4.2 Entry-PCR	35
2.4.2.1 SB-Gelbild	35

2.4.2.2 TAE-Gelbild	36
2.4.2.3 Gelextraktion und Aufreinigung der Amplifikationsprodukte	36
2.4.3 $\Delta 12$ -PCR	37
2.4.4 attB-PCR	37
2.5 Gateway-Klonierung	38
2.5.1 BP-Klonierungsreaktion	38
2.5.2 Vorbereitungen auf die Transformation	39
2.5.2.1 LB-Medium Herstellung	39
2.5.2.2 AB-Agar-Platten Herstellung	40
2.5.2.3 Herstellung einer 100 mM CaCl_2 -Lösung	40
2.5.2.4 <i>Escherichia coli</i> Bakterien Anzucht	41
2.5.3 Transformation	42
2.5.4 Kolonien picken.....	43
2.5.5 Plasmid-DNA-Extraktion	43
2.5.6 Kontrolle und Vermessung der Plasmid-DNA	44
2.6 Sequenzierung	44
2.7 Bioinformatik	45
3. Ergebnisse	46
3.1 Genanalyse	46
3.1.1 Genanalyse auf DNA-Ebene	47
3.1.2 Genanalyse auf mRNA-Ebene	56
3.1.3 Analyse der abgeleiteten Aminosäuresequenz	62
3.2 Homologie Modellierung und Voraussage der 3D-Struktur	65
4. Diskussion	70
5. Anhang	76
6. Literaturverzeichnis	80
7. Abbildungs- und Tabellenverzeichnis	82

1 Einleitung

1.1 Gattung *Origanum* (Lamiaceae)

Der Name *Origanum* bedeutet „Schmuck der Berge“ und leitet sich von den griechischen Wörtern *oros* (Berg) und *ganos* (Schmuck) ab (Alwafa, et al., 2021). Die Gattung *Origanum* L. gehört zur Familie der *Lamiaceae*, einer umfangreichen Pflanzenfamilie, die etwa 200 Gattungen umfasst, und innerhalb der Familie zum Stamm *Mentheae*. Die Gattung *Origanum* selbst besteht aus über 44 Arten, 6 Unterarten und 18 Hybriden (Abdelkahim Bouyahya, 2021). Gemäß der Klassifikation von Ietswaart (1980) können die in dieser Studie verwendeten vier Arten der Sektion *Majorana* zugeordnet werden (Ietswaart, 1980).

Die Chromosomenzahl der meisten *Origanum*-Arten beträgt $2n = 30$, mit Ausnahme einiger Abweichungen bei *O. vulgare* (Ietswaart, 1980).

Die Sektion *Majorana* erweist sich taxonomisch als besonders komplex, da viele ihrer Arten eine niedrige morphologische Variabilität aufweisen. Wichtige Bestimmungsmerkmale, wie etwa die Blütenstandstruktur und Blattmorphologie, zeigen oft fließende Übergänge zwischen den Arten. Besonders die Abgrenzung zwischen *Origanum majorana* und *Origanum dubium* wurde in zahlreichen Studien intensiv diskutiert. Diese beiden Arten wurden durch feine Unterschiede in der Form ihrer Blütenstände voneinander differenziert. Heute existieren sowohl das „Ein-Arten-Konzept“ als auch das „Zwei-Arten-Konzept“. Weitere Arten der Sektion *Majorana* sind *Origanum onites* und *Origanum syriacum*. Diese sind anhand des Gehalts an Arbutin gut voneinander trennbar – während *O. dubium* und *O. majorana* hohe Mengen an Arbutin akkumulieren, fehlt dieser Stoff in natürlichen Populationen von *O. syriacum* und *O. onites* oder ist nur in geringen Spuren nachweisbar (Brigitte Lukas, 2013).

Die Gattung ist vorwiegend im östlichen Mittelmeerraum verbreitet, wo sie aufgrund ihrer Komponenten im ätherischen Öl seit Jahrhunderten für kulinarische und medizinische Zwecke genutzt wird. Heute werden hauptsächlich zwei sensorische Typen von *Origanum* kommerziell gehandelt: Majoran, gewonnen aus *O. majorana*, und Oregano, hauptsächlich aus *O. onites* sowie *O. vulgare* (wird in dieser Arbeit nicht behandelt) (Brigitte Lukas, 2013).

Es wird angenommen, dass *O. onites* und *O. syriacum* mögliche Vorfahren von *O. dubium* darstellen, während *O. majorana* vermutlich von *O. syriacum* abstammt. Diese Erkenntnisse stützen die Hypothese, dass *O. onites* und *O. syriacum* als ursprüngliche Arten der Sektion *Majorana* betrachtet werden können (Brigitte Lukas, 2013).

Ein morphologisches Merkmal, das alle vier Arten innerhalb der Sektion *Majorana* aufweisen, ist ein hochblattartiger Kelch mit einer stark reduzierten Unterlippe. *O. onites* scheint die am besten definierte und homogenste Art zu sein, das durch korymbiforme Blütenstände und gezähnte Blätter charakterisiert werden kann. Morphologisch unterscheiden sich *O. syriacum*, *O. dubium* und *O. majorana* durch ihre rispigen Blütenstände und meist ganzrandigen Blätter (Tabelle 1) (Brigitte Lukas, 2013).

Tabelle 1: Morphologische Merkmale zur Differenzierung von *O. onites*, *O. syriacum*, *O. dubium* und *O. Majorana* (Ietswaart, 1980)

	<i>O. onites</i>	<i>O. syriacum</i>	<i>O. dubium</i>	<i>O. majorana</i>
Blütenstand	Traube	Rispe	Rispe	Rispe
Blätter	gezähnt	ganzrandig spitz Nerven abaxial	ganzrandig stumpf Nerven nicht abaxial	ganzrandig stumpf Nerven nicht abaxial
Indumentum	spitz behaart	spitz behaart	stark behaart	stark behaart
charakteristischer Chemotyp	Cymyl-Typ (Linalool-Typ)	Cymyl-Typ	Cymyl-Typ (Linalool-Typ)	Sabinyl-Typ
Arbutin	nicht vorhanden	nicht vorhanden	hoher Gehalt	hoher Gehalt

1.1.1 *Origanum majorana*

Taxonomie

Origanum majorana Linnaeus, bekannt als süßer Majoran, war früher unter dem Namen *Majorana hortensis* Moench sowie synonym als *Origanum dubium* Boiss. (*Cypriot oregano*) bekannt (1.1 Gattung *Origanum* (*Lamiaceae*)) (Abdelkahim Bouyahya, 2021).

Morphologie

O. majorana ist eine buschige, mehrjährige Pflanze, die eine Wuchshöhe von 30 bis 60 cm erreicht. Sie bildet herabhängende, stark verzweigte, rötlich gefärbte und quadratisch geformte Stängel aus, die schwach behaart und teils grün-rötlich gefleckt sind. Die kleinen, krautigen Blätter sind eiförmig bis länglich und messen 0,5 bis 1,5 cm in der Länge und 0,2 bis 0,8 cm in der Breite. Die Blätter sind oft dicht behaart, während die graugrünen Brakteen paarweise an den Stängeln sitzen. Die kleinen Blüten (0,3 cm) sind in kopfförmigen Blütenständen angeordnet, röhrenförmig mit einer zweilappigen oder einlappigen Oberlippe und weiß oder blassrosa gefärbt. Sie blühen von Juni bis September, und die ovalen, dunkelbraunen Samen reifen von August bis September. Die zylindrischen Wurzeln sind fein gefurcht, mit einem Durchmesser von 0,2 bis 0,6 mm (Abdelkahim Bouyahya, 2021).

Geographie

O. majorana ist in Zypern sowie der Region um Antalya (Türkei) heimisch und in verschiedenen Mittelmeerländern wie Serbien, Italien, Korsika, Südspanien, Portugal, Marokko und Algerien weit verbreitet. Die Pflanze wird zudem weltweit kultiviert, insbesondere in Europa (Frankreich, Ungarn), den USA sowie in Teilen Asiens, einschließlich Indien (Abdelkahim Bouyahya, 2021).

Sekundäre Inhaltsstoffe

O. majorana ist reich an sekundären Pflanzenstoffen, insbesondere an ätherischem Öl, das antibakterielle, antimykotische, antiparasitäre und entzündungshemmende Eigenschaften besitzt. Zu den Hauptkomponenten zählen Terpinen-4-ol, Carvacrol, Thymol und Linalool. Außerdem enthält die Pflanze bioaktive Substanzen wie Flavonoide, Phenolsäuren, Sterole, Alkaloide, Cumarine, Saponine und Tannine, die ihre therapeutischen Eigenschaften verstärken (Abdelkahim Bouyahya, 2021).

Wirkung und Anwendung

Die therapeutische Nutzung von *O. majorana* variiert je nach Pflanzenteil sowie geographische Lage und wird meist in Form von Aufgüssen oder Pulvern angewendet. Traditionell wird die

Pflanze in verschiedenen Kulturen bei Magenschmerzen, Erkältungen, Verdauungsproblemen, Menstruationsbeschwerden und neurologischen Störungen eingesetzt. In Europa nutzt man sie auch gegen Kopfschmerzen, Bluthochdruck und Asthma (Abdelkahir Bouyahya, 2021).

Toxikologische Studien an Ratten zeigen eine hohe Verträglichkeit und Sicherheit der Pflanze. Diese Ergebnisse unterstützen die Eignung der Pflanze für medizinische Anwendungen (Abdelkahir Bouyahya, 2021).

1.1.2 *Origanum dubium*

Taxonomie

Origanum dubium Boiss. ist eine ökonomisch bedeutende Wild-Oregano-Art und gehört zur Sektion *Majorana* innerhalb der Gattung *Origanum* (Familie *Lamiaceae*) (1.1 Gattung *Origanum (Lamiaceae)*) (Kenan Turgut, 2017).

Morphologie

O. dubium lässt sich morphologisch nur schwer von *O. majorana* unterscheiden. Untersuchungen zeigen jedoch feine Unterschiede in der Blütenstandsstruktur. So ist *O. dubium* stärker behaart als *O. majorana*. Ansonsten sind beide Pflanzen rispenartig aufgebaut, mit stumpfen, glattrandigen Blättern (Brigitte Lukas, 2013).

Geographie

O. dubium ist hauptsächlich in der Region Antalya (Türkei) verbreitet, wo sie intensiv aus wild-wachsenden Beständen in Bergregionen gesammelt wird. Weiters findet sie auch in Griechenland und Zypern Anwendung (Kenan Turgut, 2017).

Sekundäre Inhaltsstoffe

Regionale Unterschiede zeigen sich in der chemotypischen Zusammensetzung: In der Türkei wurden Carvacrol-, Linalool- und Thymol-Chemotypen dokumentiert, während auf Zypern fast ausschließlich der Carvacrol-Chemotyp vorkommt. Das ätherische Öl von *O. dubium* ist reich

an Carvacrol, das für seine antimikrobiellen, antioxidativen und fungiziden Eigenschaften bekannt ist, und daher von pharmazeutischem und industriellem Interesse. Neben Carvacrol sind auch Linalool und Thymol in verschiedenen Chemotypen vertreten, was den wirtschaftlichen Wert der Art erhöht (Kenan Turgut, 2017).

Wirkung und Anwendung

Mit einem ätherischen Öl-Gehalt von 6–8 % wird *O. dubium* hauptsächlich zur Ölgewinnung verwendet und weniger als Gewürz oder Teekraut. Traditionell werden die Blätter und Blüten als Aufguss bei Magen-Darm-Beschwerden und als Diuretikum genutzt, während das Öl äußerlich gegen rheumatische Beschwerden eingesetzt wird. Die Züchtung fokussiert sich auf die Entwicklung neuer Genotypen mit erhöhten Carvacrol-Konzentrationen, um den wachsenden Bedarf an natürlichen Antioxidantien und antimikrobiellen Substanzen zu decken (Kenan Turgut, 2017).

1.1.3 *Origanum syriacum*

Taxonomie

Origanum syriacum, auch bekannt als „Syrischer Oregano“ oder „Biblischer Ysop“, ist ein mehrjähriges, aromatisches Kraut aus der Mittelmeerregion und ein zentraler Bestandteil der Küche des Nahen Ostens. Besonders in der beliebten Zaatar-Mischung, eine Mischung aus *O. syriacum*, Sesam und Sumach (*Rhus coriaria*), findet die Pflanze weite Verwendung (Joelle Mesmar, 2022). Taxonomisch gehört *O. syriacum* zur Sektion *Majorana* innerhalb der Gattung *Origanum* (Familie *Lamiaceae*) (1.1 Gattung *Origanum* (*Lamiaceae*)).

Morphologie

O. syriacum erreicht 30 bis 60 cm Höhe, hat holzige Wurzeln und fein behaarte, verzweigte Stängel. Die Blätter und Drüsen können je nach Art gestielt oder aufsitzend sein. Die Blütenstände sind typischerweise in dichten, rispenartigen Ähren angeordnet (Alwafa, et al., 2021).

Geographie

O. syriacum ist in verschiedenen Regionen des östlichen Mittelmeerraums, Westasiens und Südeuropas heimisch und wird wegen ihrer Anpassungsfähigkeit häufig kultiviert (Alwafa, et al., 2021).

Sekundäre Inhaltsstoffe

Das ätherische Öl von *O. syriacum* weist regionale Unterschiede auf, abhängig von den klimatischen Bedingungen. Hauptbestandteile wie Carvacrol und Thymol dominieren in Proben aus dem Libanon, während in ägyptischen Proben andere Inhaltsstoffe wie cis-Sabinenhydrat und Terpinen-4-ol stärker vertreten sind. Proben aus Syrien weisen hingegen hohe Anteile an β -Myrcen und *p*-Cymol auf (Alwafa, et al., 2021). Weiters enthält die Pflanze auch Flavonoide und Phenole, die für ihre therapeutische Wirkung bekannt sind (Joelle Mesmar, 2022).

Wirkung und Anwendung

Carvacrol und Thymol verleihen der Pflanze nicht nur ihr charakteristisches Aroma, sondern auch starke antimikrobielle und antiparasitäre Eigenschaften. Darüber hinaus bieten die enthaltenen Flavonoide und Phenole antioxidative, entzündungshemmende und krebshemmende Effekte (Joelle Mesmar, 2022). Traditionell wird *O. syriacum* in der arabischen und türkischen Volksmedizin bei Magenbeschwerden, Erkältungen, Zahnschmerzen, Herzkrankheiten und zur Wundheilung eingesetzt (Alwafa, et al., 2021).

Dank der nachgewiesenen Verträglichkeit gilt die Pflanze als vielversprechender Kandidat für die Entwicklung neuer Arzneimittel (Joelle Mesmar, 2022).

1.1.4 *Origanum onites*

Taxonomie

Origanum onites L., auch bekannt als „Türkischer Oregano“, ist eine mehrjährige, krautige Pflanze, die in der Türkei weit verbreitet ist. Sie spielt eine wichtige Rolle in der traditionellen Medizin und Küche, wo sie unter lokalen Namen wie „kırkbaşı kekik“ oder „bilya kekik“ weit verbreitet ist (Bektas Tepe, 2016). Taxonomisch gehört *O. onites* zur Sektion *Majorana* innerhalb der Gattung *Origanum* (Familie *Lamiaceae*) (1.1 Gattung *Origanum* (*Lamiaceae*)).

Morphologie

O. onites zeichnet sich durch traubenförmige Blütenstände, gezähnte Blätter und behaarte, spitz zulaufende Stängel aus (Brigitte Lukas, 2013).

Geographie

Die Pflanze ist hauptsächlich entlang der Mittelmeerküste, insbesondere in der Türkei und Griechenland, beheimatet (Bektas Tepe, 2016).

Sekundäre Inhaltsstoffe

O. onites ist reich an sekundären Metaboliten wie Terpenoiden, Flavonoiden, phenolischen Säuren, Fettsäuren, Sterolen und Mineralstoffen (Bektas Tepe, 2016). Besonderes Interesse gilt dem ätherischen Öl, das bis zu 134 verschiedene Terpenoide enthält. Carvacrol und Thymol sind die Hauptbestandteile des Öls, deren Gehalt von Umweltbedingungen und Erntetechniken abhängt. Auch Linalool, α -Pinen und *p*-Cymol sind nachgewiesen, was auf sieben verschiedene Chemotypen hinweist (Bektas Tepe, 2016).

Wirkung und Anwendung

O. onites zeigt zahlreiche therapeutische Eigenschaften, darunter antimikrobielle, antioxidative, entzündungshemmende, hepatoprotektive, antidiabetische und krebshemmende Wirkungen. In der türkischen Küche wird die Pflanze häufig verwendet und

traditionell kommt sie als Aufguss oder Dekokt zur Behandlung von Magen-Darm-Beschwerden, Bronchitis, Bluthochdruck, Diabetes und Infektionen zum Einsatz. Sie wirkt auch krampflösend, antiseptisch und fördert die Wundheilung (Bektas Tepe, 2016).

Bisherige Studien zeigen, dass *O. onites* keine toxischen oder mutagenen Effekte aufweist, jedoch sind weitere Forschungen zur Sicherheit auf Säugetierzellen notwendig. Insgesamt birgt die Pflanze großes Potenzial für die Entwicklung von Phytopharmaka und kosmetischen Produkten (Bektas Tepe, 2016).

1.2 Grundlagen der Flavonoide

Flavonoide sind vielseitige sekundäre Pflanzenmetabolite, die eine bedeutende Rolle in der Pflanzenentwicklung sowie Verteidigung (Anthocyane, Chalkone, Aurone, Flavone, Flavonole), Pigmentfunktion (Stressschutz und Antioxidantien) und Signalübertragung (Bestäubung, Stoffwechselregulation) spielen und gleichzeitig umfassende Anwendungen in der Lebensmittel- und Medizinindustrie finden. Ihre Biosynthese ist komplex und erfordert ein koordiniertes Zusammenspiel zahlreicher Enzyme (Liu, et al., 2021).

Grundstruktur der Flavonoide

Flavonoide sind phenylpropanoide Verbindungen und fungieren als wasserlösliche Pigmente, die in den Vakuolen von Pflanzenzellen eingelagert werden. Die typische Struktur besteht aus einem C6-C3-C6-Gerüst (Abbildung 1), das zwei aromatische Benzolringe (Ringe A und B) über einen dazwischenliegenden dreigliedrigen Heterozyklus (Ring C) verbindet. Unterschiede im Oxidationszustand des Ringes C und zusätzliche Modifikationen wie Glykosylierung, Acylierung oder Polymerisation erzeugen eine breite Vielfalt an Flavonoid-Verbindungen. Bisher wurden über 9000 Flavonoide in Pflanzen identifiziert (Liu, et al., 2021).

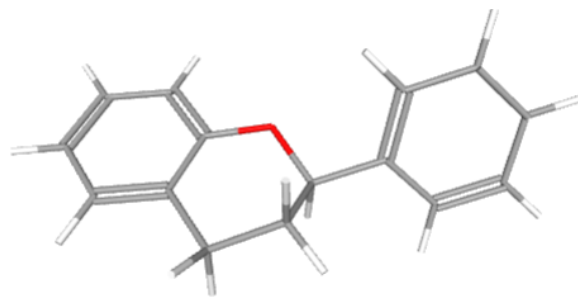


Abbildung 1: Flavan-Grundstruktur (PubChem, 3D Wire-Frame, National Center for Biotechnology Information. PubChem Compound Summary for CID 94156, Flavan. <https://pubchem.ncbi.nlm.nih.gov/compound/Flavan>. Accessed Oct. 28, 2024)

Bedeutung für Lebensmittel und Medizin

Flavonoide finden breite Anwendung in der Lebensmittel- und Pharmaindustrie. Anthocyane und Proanthocyanidine sind essbare Pigmente, die in Lebensmitteln und Wein zur Farbgebung und Geschmacksregulation eingesetzt werden. In der Medizin verzögern Flavonoide den

Alterungsprozess in verschiedenen Organsystemen (z. B. Nervensystem, Immunsystem, Haut) und helfen bei der Prävention von Erkrankungen wie Osteoporose, Diabetes, Herz-Kreislauf-Erkrankungen, Alzheimer und Brustkrebs (Liu, et al., 2021). Sie wirken stark entzündungshemmend, antiproliferativ, beruhigend, antidepressiv, krebshemmend und hepatoprotektiv (K.S. Sridevi Sangeetha, 2016).

Diese Effekte beruhen hauptsächlich auf den antioxidativen Eigenschaften der Flavonoide. Sie neutralisieren reaktive Sauerstoffspezies ROS, die durch verschiedene Stressfaktoren entstehen. Zudem beeinflussen sie Schlüsselenzyme wie Proteinkinasen, Alpha-Glucosidase sowie Aldose-Reduktase und regulieren so gestörte zelluläre Signalwege. Aufgrund ihrer vielseitigen Vorteile stehen Flavonoide im Fokus intensiver Forschung, und zahlreiche Produkte sind bereits als kostengünstige pharmazeutische Präparate erhältlich (K.S. Sridevi Sangeetha, 2016).

1.2.1 Flavonoid-Biosynthese

Die Biosynthese von Flavonoiden ist ein zentrales Thema der Pflanzenforschung, da noch nicht alle enzymatischen Schritte vollständig aufgeklärt sind. Flavonoide lassen sich in acht Unterklassen unterteilen: Stilbene, Aurone, Flavone, Isoflavone, Flavonole, Phlobaphene, Proanthocyanidine und Anthocyane. Die Synthese dieser Verbindungen erfolgt aus vier Hauptvorläufern: Chalkon, Flavanon, Leukoanthocyanidin und Dihydroflavonol (Liu, et al., 2021).

Der Prozess beginnt mit dem Phenylpropanoid-Weg, in dem Phenylalanin über drei Enzyme (Phenylalanin-Ammoniak-Lyase (PAL), Cinnaminsäure-4-Hydroxylase (C4H) und 4-Cumarat-CoA-Ligase (4CL)) zu p-Cumaroyl-CoA umgewandelt wird. Dies dient als Ausgangsstoff für die Flavonoid-Biosynthese (Liu, et al., 2021).

Der erste Schritt der Flavonoid-Biosynthese wird durch die Chalkonsynthase (CHS) katalysiert (Abbildung 2). Dabei werden p-Cumaroyl-CoA (aus dem Phenylpropanoid-Weg) und Malonyl-CoA (aus dem Acetat-Malonat-Weg) zu Naringenin-Chalkon umgewandelt. Im Anschluss erfolgt die zyklische, stereospezifische Umwandlung von Naringenin-Chalkon zu Naringenin durch das Enzym Chalkon-Isomerase (CHI), wobei dieser Schritt auch spontan

ablaufen kann. Naringenin fungiert als allgemeiner Vorläufer für mehrere Flavonoid-Unterklassen (Yonekura-Sakakibara K, 2019).

Enzyme wie Flavanon-3-Hydroxylase (F3H), Flavonol-Synthase (FLS) und Dihydroflavonol-4-Reduktase (DFR) leiten die Synthese spezifischer Flavonoid-Typen ein. Beispielsweise werden durch LDOX/ANS Anthocyanidine gebildet, die zu Flavan-3-olen wie Catechin weiterverarbeitet werden. Flavone und Isoflavone entstehen durch spezifische Synthasen wie FNS I/II und Isoflavon-Synthase (IFS) (Yonekura-Sakakibara K, 2019).

Untersuchungen deuten darauf hin, dass die Fähigkeit zur Flavonoid-Produktion entweder mehrfach unabhängig entstanden ist oder im Laufe der Evolution in bestimmten Pflanzenlinien verloren ging. Evolutionsstudien zeigen, dass die upstream-Gene der Anthocyan-Biosynthese, wie CHS, CHI und F3H, langsamer evolvieren als die downstream-Gene, etwa DFR und ANS. Dies deutet auf stärkere evolutionäre Zwänge hin, da die upstream-Gene eine zentrale Rolle im gesamten Stoffwechselweg spielen (Yonekura-Sakakibara K, 2019).

Interessanterweise zeigen Transkriptionsfaktoren, die die Anthocyan-Biosynthese regulieren, jedoch eine schnellere Evolution, was vermutlich auf eine Anpassung an Umweltveränderungen hinweist (Yonekura-Sakakibara K, 2019).

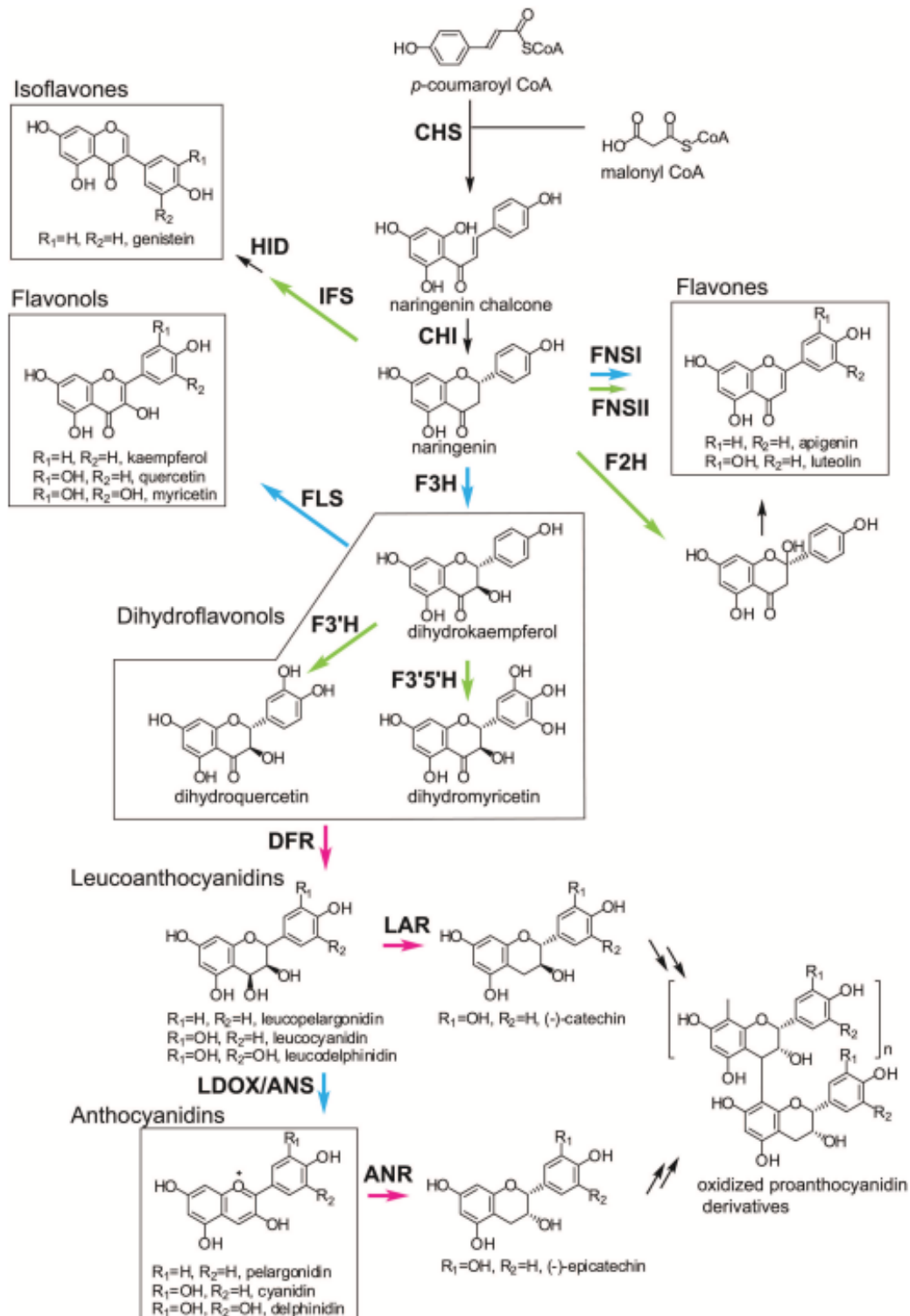


Abbildung 2: Allgemeine Flavonoid-Biosynthesewege in Pflanzen. (Die Pfeile in gr ün, blau und magenta kennzeichnen Enzyme der CYP-, 2OGD- und SDR-Superfamilien. Hier sind die beteiligten Enzyme und ihre Abkürzungen: ANR – Anthocyanidin-Reduktase; ANS – Anthocyanidin-Synthase; CHI – Chalcon-Isomerase; CHR – Chalcon-Reduktase; CHS – Chalcon-Synthase; DFR – Dihydroflavonol-4-Reduktase; F2H – Flavanon-2-Hydroxylase; F3H – Flavanon-3-Hydroxylase; F3'H – Flavonoid-3'-Hydroxylase; F3'5'H – Flavonoid-3'5'-Hydroxylase; FLS – Flavonol-Synthase; FNS – Flavon-Synthase; HID – 2-Hydroxyisoflavanon-Dehydratase; IFS – Isoflavon-Synthase; LAR – Leucoanthocyanidin-Reduktase; LDOX – Leucoanthocyanidin-Dioxygenase) (Yonekura-Sakakibara K, 2019)

1.2.2 Chalkonsynthase (CHS)

Die CHS gilt als der einfachste Vertreter der Typ-III-Polyketid-Synthasen (PKS) und katalysiert den ersten Schritt in der Flavonoid-Biosynthese. Sie fungiert dabei als ein essenzielles Schlüsselenzym für die Synthese sekundärer Pflanzenstoffe. Durch ihre hochkonservierte Struktur und die katalytische Triade (Cys164, His303 und Asn336) ermöglicht sie eine präzise Substratausrichtung, die zur Bildung komplexer aromatischer Verbindungen führt (Shahzad A. Pandith, 2020).

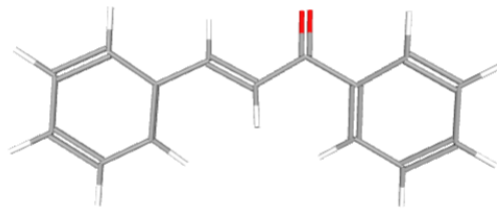


Abbildung 3: Chemische Struktur von Chalkon (PubChem, 3D Wire-Frame, National Center for Biotechnology Information. PubChem Compound Summary for CID 637760, Chalcone. <https://pubchem.ncbi.nlm.nih.gov/compound/Chalcone>. Accessed Oct. 28, 2024)

Verbreitung der CHS-Gene

Typ-III-PKS-Enzyme, wie die CHS, sind homodimere Proteine, gehören zur Thiolase-Superfamilie und umfassen etwa 20 funktionell unterschiedliche Vertreter, die entweder pflanzlichen oder bakteriellen Ursprungs sind. Alle sequenzierten Landpflanzengenome enthalten mindestens ein CHS-Gen, was auf die universelle Bedeutung dieses Enzyms für pflanzliche Stoffwechselwege hinweist. (Yonekura-Sakakibara K, 2019).

Eine Besonderheit der CHS-Gene liegt in ihrer Intronstruktur. Während die meisten CHS-Gene ein einzelnes Intron zwischen zwei Exons aufweisen, gibt es auch Ausnahmen. So besitzen Arten wie *Antirrhinum majus* (Plantaginaceae) und *Alpinia calcarata* (Zingiberaceae) zwei Introns. Weiters sind einige CHS-Gene intronlos oder enthalten mehr als zwei Introns, was auf eine bemerkenswerte evolutionäre Vielfalt hinweist (Shahzad A. Pandith, 2020).

Die Expression von CHS-Genen variiert stark in Abhängigkeit von Art, Gewebe und Entwicklungsstadium. Diese Variabilität beeinflusst die Synthese diverser sekundärer Metaboliten, die für die Pflanzenabwehr, die Pigmentierung und die Interaktionen mit der Umwelt entscheidend sind. Während einige CHS-Gene konstitutiv exprimiert werden, erfolgt

die Aktivierung anderer durch externe Faktoren wie UV-Licht, Verwundung oder Pathogenbefall. Die Akkumulation von Flavonoiden korreliert dabei eng mit der CHS-Expression (Shahzad A. Pandith, 2020).

Struktur und Funktion der CHS

Die Kristallstruktur der CHS aus *Medicago sativa* zeigt, dass das Enzym ein Homodimer bildet, wobei jedes Monomer eine Masse von 42–45 kDa besitzt. Jedes Monomer enthält ein tief im Inneren gelegenes aktives Zentrum, das aus der konservierten katalytischen Triade Cys164, His303 und Asn336 besteht (Abbildung 4). Diese katalytische Triade ist an einem $\alpha\beta\alpha$ -Pseudomotiv positioniert, das sich aus zwei pseudo-symmetrischen $\alpha\beta\alpha$ -Domänen zusammensetzt, die vermutlich aus einem Genduplikationsereignis hervorgegangen sind (Shahzad A. Pandith, 2020).

Das Dimerisierungsinterface besteht aus hydrophoben und hydrophilen Resten, die entlang einer flachen Oberfläche interagieren, welche durch zwei N-terminale Helices begrenzt wird. Die beiden funktional unabhängigen aktiven Zentren befinden sich an den unteren Enden der $\alpha\beta\alpha$ -Strukturen jedes Monomers. Die meisten katalytischen Reste stammen aus demselben Monomer, jedoch spielt auch Met137 des benachbarten Monomers eine entscheidende Rolle (Shahzad A. Pandith, 2020).

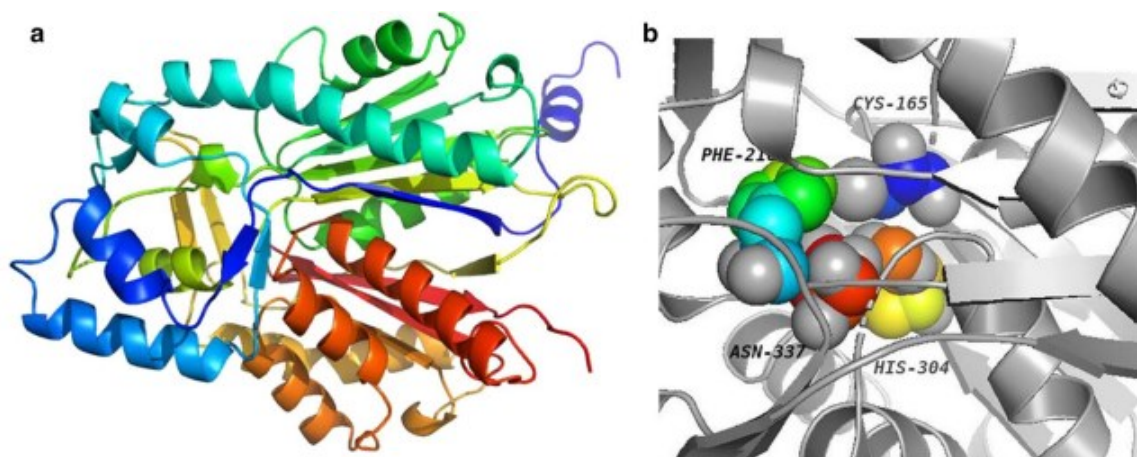


Abbildung 4: a, 3D-Struktur des ReCHS2-Proteins aus *Rheum australe* (Polygonaceae); b, Detailausschnitt: zeigt die katalytische Triade bestehend aus Cys164, His303 und Asn336 sowie Phe 216 (Nummerierung aus *Medicago sativa*), Cartoon-Modell (Shahzad A. Pandith, 2020)

Das aktive Zentrum enthält drei miteinander verbundene Bindungsstellen: einen CoA-Bindungstunnel, eine Cumaryl-Bindungstasche und eine Cyclisierungs-Tasche. Diese strukturierte Architektur ermöglicht die effiziente Bindung und Verarbeitung von *p*-Cumaroyl-CoA, dem Ausgangssubstrat für die Synthese von Flavonoiden. Der Reaktionsmechanismus beginnt mit der Übertragung der *p*-Cumaroyl-Gruppe auf Cys164, die durch His303 aktiviert wird. Asn336 stabilisiert den tetraedrischen Zwischenzustand durch Wasserstoffbrückenbindungen. Nach der Decarboxylierung von Malonyl-CoA entsteht ein stabiles Carbanion, das die Verlängerung der Polyketidkette ermöglicht. Dieser Prozess wiederholt sich in mehreren Zyklen, bis ein Tetraketid-Zwischenprodukt entsteht. Dabei wechseln die katalytischen Reste zyklisch zwischen verschiedenen Aktivierungszuständen. (Shahzad A. Pandith, 2020).

Studien zeigen, dass die aromatischen Reste wie Phe215 und Phe265 an der korrekten Substratausrichtung beteiligt sein könnten, was die effiziente Verlängerung der Polyketidkette unterstützt (Jean-Luc Ferrer, 1999).

Evolutionäre Bedeutung und funktionelle Diversität

Genverdopplungen spielen eine zentrale Rolle in der Evolution der Pflanzen, indem sie biochemische Flexibilität ermöglichen und die Entstehung neuer Funktionen begünstigen. Auch die CHS-Gene sind durch Genduplikationen entstanden, wodurch sich unterschiedliche funktionelle Varianten entwickeln konnten, ohne die Flavonoid-Biosynthese zu beeinträchtigen (Shahzad A. Pandith, 2020).

Eine bemerkenswerte Eigenschaft der CHS ist ihre Substratvielfalt gegenüber Startermolekülen. Das Enzym kann eine Vielzahl aliphatischer und aromatischer CoA-Ester verarbeiten und dadurch eine breite Palette an Polyketidprodukten synthetisieren. Diese Flexibilität verschafft den Pflanzen einen evolutionären Vorteil, da so neue bioaktive Verbindungen entstehen, die beispielsweise antimikrobiell wirken können (Shahzad A. Pandith, 2020).

Zudem reagiert die CHS-Expression stark auf Umweltfaktoren wie UV-Licht, Phytohormone und Pathogenbefall. Diese Regulation war vermutlich entscheidend für die Anpassung der Pflanzen an terrestrische Lebensräume, wo sie stärkerem Sonnenlicht ausgesetzt sind (Shahzad A. Pandith, 2020).

Fallstudien zu CHS-Genen in Salvia miltiorrhiza (Lamiaceae) und Matricaria chamomilla (Asteraceae)

Eine systematische Untersuchung des Genoms von *Salvia miltiorrhiza* (Lamiaceae) führte zur Identifizierung von 26 PKS-Genen, von denen 20 erstmals beschrieben wurden. Acht SmCHS-Gene konnten durch molekulares Klonieren bestätigt werden. Die Proteinsequenzen zeigen eine hohe Übereinstimmung mit den CHS-Proteinen anderer Pflanzenarten und enthalten die konservierten Chalkon-Domänen. Die offenen Leserahmen (ORFs) dieser Gene wurden auf 1173 bp festgelegt, was wichtige Anhaltspunkte für die vorliegende Arbeit liefert. Der SmCHS1-Primer aus der Studie kann auch weiterhin für diese Arbeit für die kultivierten Proben von *O. majorana* verwendet werden (Yuxing Deng, 2018).

Eine weitere Studie beschreibt die Analyse der CHS-Gene bei Kamillenarten. In sechs Genotypen von *Matricaria chamomilla* und *M. discoidea* wurden zwei Exons mit einer Länge von 188 bp bzw. 1011 bp identifiziert, die durch ein Intron von variabler Länge getrennt sind. Die Exons zeigen eine höhere Konservierung (5,3–6,2 SNPs pro 100 bp) im Vergleich zu den Introns, die eine höhere Polymorphismus-Rate aufweisen (11,5 SNPs pro 100 bp). Insgesamt wurden 96 SNPs in beiden Arten festgestellt, davon 70 verschiedene Haplotypen. Die Basenpaar-Länge lag bei rund 1390 (Melanie Novak, 2023).

1.3 Methodik-Wahl für die vorliegende Genexpressionsstudie

1.3.1 Proteinbiosynthese in Pflanzen

Pflanzen sind Eukaryoten, deren Erbgut in Genen organisiert ist. Diese bestehen aus codierenden Exons und nicht-codierenden Introns. Die DNA speichert die Informationen zur Synthese von RNA und Proteinen. Die Proteinbiosynthese erfolgt in zwei Hauptschritten: Zunächst findet die Transkription statt, bei der ein RNA-Abbild der DNA erstellt wird. Darauf folgt die Translation, bei der die RNA-Sequenz in eine spezifische Aminosäurekette übersetzt wird, wobei jedes Codon (Basentriplett) für eine bestimmte Aminosäure codiert. Nach der Translation können posttranslationale Modifikationen erfolgen. Ribosomen fungieren als zentrale „Maschinen“ der Proteinbiosynthese, die in drei Phasen – Initiation, Elongation und Termination – in der 5'-zu-3'-Richtung ablaufen (Schmidt, 2017).

1.3.2 DNA-Extraktion

Die Extraktion von DNA, RNA und Proteinen ist eine grundlegende Methode der Molekularbiologie. Diese Biomoleküle lassen sich aus biologischem Material isolieren und für analytische oder präparative Zwecke nutzen. Moderne, spezialisierte Verfahren, darunter säulenbasierte Protokolle, haben den Prozess vereinfacht (Siun Chee Tan, 2009). In dieser Arbeit kommt das InnuPREP Plant DNA Kit (IST Innuscreen, Berlin, DE) zum Einsatz, das einen CTAB-haltigen Lysepuffer (SLS) verwendet (2.2 DNA-Extraktion).

1.3.3 PCR

In dieser Genexpressionsstudie wird die Polymerase-Kettenreaktion (PCR) eingesetzt, um ein DNA-Fragment des Chalkonsynthase-Gens zu amplifizieren und für die Klonierung vorzubereiten. Die PCR durchläuft wiederholte Zyklen von Denaturierung (Aufspaltung der doppelsträngigen DNA), Annealing (Anlagerung spezifischer Primer) und Elongation (Verlängerung der Primer). Jede Wiederholung verdoppelt die DNA-Menge exponentiell. Zur Optimierung werden spezifische Primer entwickelt und ideale Bedingungen ermittelt.

Nach der ersten Amplifikation mittels Entry-PCR erfolgt ein zweiter Durchlauf mit um $\Delta 12$ -Sequenzen erweiterten Primern. Der letzte Schritt verwendet Primer mit attB-Stellen, die für die Integration des Zielgens im Gateway-Klonierungssystem von zentraler Bedeutung sind. Diese bieten spezifische Erkennungsstellen für Rekombinationsreaktionen (Mülhardt, 2013) (2.4 PCR).

1.3.4 Agarose-Gelelektrophorese

Die Agarose-Gelelektrophorese ermöglichte die Trennung und quantitative Abschätzung der Größe von DNA-Fragmenten basierend auf ihrer Wanderung durch ein elektrisches Feld in einem Agarosegel, von der Kathode zur Anode. Kleinere Fragmente bewegen sich schneller als größere. Nach der Elektrophorese werden die DNA-Banden mit einem fluoreszierenden Farbstoff wie PeqGreen unter UV-Licht sichtbar gemacht. Dies ermöglicht die Analyse und Dokumentation der DNA-Produkte. Für die Analyse wird SB-Puffer verwendet, während die DNA-Extraktion nach der Gelelektrophorese mit TAE-Puffer erfolgt, der durch seinen niedrigen Schmelzpunkt sowie konservierenden EDTA-Puffer eine schonende Isolierung der DNA ermöglicht (Mülhardt, 2013) (2.3 Agarose-Gelelektrophorese).

1.3.5 Gateway-Klonierungsverfahren

Das Gateway-Klonierungssystem nutzt Rekombinasen zur direkten Klonierung von PCR-Fragmenten in Donor-Vektoren mit attB-Sequenzen. Geklonte Gene können leicht in verschiedene Zielvektoren transferiert werden, was eine hohe Flexibilität für unterschiedliche Anwendungen bietet. Die verwendeten Vektoren sind Plasmide, die in Bakterien unabhängig vom Genom repliziert werden können (Mülhardt, 2013). Sie enthalten wichtige Elemente wie den Replikationsursprung (ori), ein Selektionsgen (meist ein Antibiotikaresistenz-Gen) und eine Klonierungsstelle (cloning site), in die fremde DNA eingefügt werden kann (Mansour Karimi, 2007).

Zur Klonierung werden die Bakterienzellen durch eine Behandlung mit Calciumchlorid transformationskompetent gemacht. Transformation beschreibt das Einbringen von DNA in Bakterienzellen, wobei *Escherichia coli* (*E. coli*) aufgrund ihrer einfachen Kultivierung und

hohen Vermehrungsrate als Wirt bevorzugt wird. Nach erfolgreicher Klonierung werden die DNA-Sequenzen bestimmt und mittels bioinformatischer Tools analysiert, um Rückschlüsse auf die genetische Variabilität sowie auf evolutionäre Beziehungen ziehen zu können (Mülhardt, 2013) (2.5 Gateway-Klonierung).

1.4 Polymorphismus

Ein besonderer Fokus dieser Studie liegt auf der Analyse von Polymorphismen im Chalkonsynthase-Gen. Im Gegensatz zu Mutationen, die seltener auftreten, sind Polymorphismen innerhalb einer Population häufiger anzutreffen, bleiben meist erhalten und werden durch Vererbung weitergegeben. Single-Nucleotide-Polymorphismen (SNPs) und Single-Nucleotide-Varianten (SNVs) entstehen durch Fehler in Prozessen wie Replikation, Reparatur oder Rekombination der DNA. Besonders bedeutend sind nicht-synonyme Mutationen, bei denen ein Codon für eine andere Aminosäure codiert, was die Funktion des resultierenden Proteins verändern kann (Schmidt, 2017).

SNPs – also Variationen an einer einzelnen DNA-Position – treten sowohl in codierenden als auch in nichtcodierenden Bereichen des Genoms auf. In nichtcodierenden Regionen sind sie jedoch häufiger, da sie dort keinem Selektionsdruck ausgesetzt sind. Wenn durch solche Varianten nur wenige Nukleotide hinzugefügt (Insertion) oder entfernt (Deletion) werden, spricht man von SNVs oder spezifischer von Indels (Schmidt, 2017).

Darüber hinaus werden in dieser Studie die Haplotypen (MLGs) analysiert. Ein Haplotyp beschreibt die Kombination von Allelen mehrerer eng gekoppelter Gene auf demselben Chromosom (Schmidt, 2017).

1.5 Fragestellung

Die Chalkonsynthase (CHS) ist ein Schlüsselenzym im ersten Schritt der Flavonoid-Biosynthese, welches eine zentrale Rolle in der Regulation des Flavonoidgehalts spielt. Diese Studie zielt darauf ab, die genetische Variabilität des CHS-Gens innerhalb der Gattung *Origanum*, genauer zwischen den Arten *O. majorana*, *O. onites*, *O. syriacum* und *O. dubium* zu untersuchen, um Grundlagen für zukünftige Forschungsprojekte zu schaffen.

Der Fokus liegt auf der Überprüfung früherer Annahmen zu Länge des Gens, Selektionsdruck, aktivem Zentrum und Polymorphismen-Verteilung im CHS-Gen. Dazu wird das CHS-Gen mittels Gateway-Klonierungsverfahren kloniert und sequenziert.

Im ersten Schritt werden die Genlängen und Exon-Intron-Grenzen definiert, gefolgt von einer Analyse der Polymorphismen, insbesondere SNPs und MLGs auf DNA- und mRNA-Ebene. Anschließend werden die Auswirkungen dieser SNPs auf die resultierenden Aminosäuresequenzen untersucht.

Erste Hinweise auf genetische Unterschiede zwischen kultivierten und wildwachsenden Proben können ebenfalls ansatzweise erfasst werden.

2 Materialien und Methoden

Um Kontaminationen zu vermeiden, zuverlässige Ergebnisse zu gewährleisten und die Sicherheit des Laborpersonals zu garantieren, sind strenge Hygiene- und Sicherheitsmaßnahmen in molekularbiologischen Laboren unerlässlich. Dazu zählen das Tragen geeigneter Laborbekleidung, gründliche Handhygiene, die Pflege sauberer Arbeitsflächen, eine sorgfältige Instrumentenpflege sowie der sichere Umgang mit toxischen Chemikalien und die ordnungsgemäße Entsorgung von Abfällen. Alle Proben und Materialien müssen exakt beschriftet und dokumentiert werden (Mülhardt, 2013).

Um optimale Sterilität zu erreichen, wurden sämtliche Gefäße, Glasflaschen und Instrumente autoklaviert. Dabei wurde darauf geachtet, dass die Behälter nicht vollständig verschlossen und der Autoklav (Wise Clave WAC47, Digital Fuzzy Control System, Wise Laboratory Instruments) nicht überfüllt war. Zur Kontrolle der Sterilität wurden die Behälter mit einem Indikatorband versehen. Das Sterilisationsprogramm wurde bei 121°C für 25 Minuten durchgeführt. Nach dem Druckabbau und dem Abkühlen auf 60°C wurde der Autoklav entladen. Während der gesamten Genotypisierung wurde ausschließlich mit autoklaviertem Labormaterial gearbeitet.

2.1 Probenmaterial

Als Ausgangsmaterial für diese Studie dienten sechs Proben aus vier verschiedenen Individuen der Gattung *Origanum* aus der Familie *Lamiaceae*. Die Proben umfassten:

- *Origanum majorana*: Proben 230271 (auch bekannt als MH01, im Gewächshaus kultiviert) und M85 (wild)
- *Origanum dubium*: Proben M105 (wild) und M114 (wild)
- *Origanum syriacum*: Probe Syr 18/12 (wild)
- *Origanum onites*: Probe 230283 (vom Jahrmarkt in Regensburg, wild)

Die Blätter der ausgewählten Exemplare wurden geerntet und bei Raumtemperatur auf Silicagel getrocknet. Die Probe 230271 wurde einbezogen, um potenzielle Unterschiede im Genotyp zwischen kultivierten und wildwachsenden Pflanzen zu untersuchen. Es wurde angenommen, dass kultivierte Proben einen sehr homogenen Genotyp mit geringerer Variation im Vergleich zu wildwachsenden Proben aufweisen.

2.2 DNA-Extraktion

2.2.1 Probenvorbereitung

Von dem getrockneten Blattmaterial wurden 2-3 kleine Blattstücke mit einer sterilen Pinzette entnommen und in 2 ml Eppendorfer Reaktionsgefäße, die mit einigen Glaskugeln gefüllt waren, überführt. Anschließend wurden die Proben mit einer Schwungmühle (Firma Retsch GmbH, Haan, Deutschland) bei 30 Hz für 1-3 Minuten vermahlen. Dieser Vorgang wurde so lange wiederholt, bis das Pulver fein vermahlen und homogenisiert war.

Der Thermoblock (Mixing Block MB-102, Thermoblock mit Schüttelfunktion, Bioer Technology, Hangzhou high tech, China) wurde auf 65 °C vorgeheizt. Alle benötigten Lösungen wurden im Vorfeld verdünnt und hergestellt. Die lyophilisierte Proteinase K wurde mit Milli-Q®-Wasser gelöst, auf eine Endkonzentration von 20 mg/ml gebracht und im Tiefkühler gelagert. Die anderen Reagenzien wurden bei Raumtemperatur aufbewahrt.

Für das Blattmaterial war das Protokoll 1 aus dem InnuPREP Plant DNA Kit (Firma IST Innuscreen GmbH, Berlin, Deutschland) mit dem Lysepuffer SLS am geeignetsten, da dieser CTAB als Detergens enthält.

2.2.2 Extraktion der DNA

Die DNA-Extraktion wurde gemäß den Anweisungen des Herstellers des InnuPREP Plant DNA Kits durchgeführt. Zum homogenisierten Blattmaterial wurden 400 µl Lysis-Lösung SLS und 20 µl Proteinase K hinzugefügt und für 5 Sekunden kräftig gevortext (Vortexer Vortex-Genie 2,

Scientific Industries, New York, USA). Die Lösung wurde dann bei 65 °C im Thermoblock mit kontinuierlichem Schütteln für 30 Minuten inkubiert.

Nach der Inkubation wurden die Proben auf ein Säulchen in einem Auffangröhrchen überführt und bei 11.000 U/min für 1 Minute zentrifugiert (Zentrifuge Micro 200, Hettich, Tuttlingen, Deutschland). Das Säulchen wurde entfernt, das Filtrat mit der DNA beibehalten. Den lysierten Proben (Filtrat) wurden 200 µl Bindungslösung SBS hinzugefügt und mehrfach auf- und abpipettiert, bis eine homogene Lösung entstand. Die Proben wurden auf ein neues Säulchen in einem neuen Auffangröhrchen überführt und erneut bei 11.000 U/min für 2 Minuten zentrifugiert. Bei diesem Schritt wurde die DNA an das Säulchen gebunden.

Beim nächsten Schritt wurde das Säulchen mit 650 µl Waschlösung MS befeuchtet, gefolgt von einer Zentrifugation bei 11.000 U/min für 1 Minute. Das Auffangröhrchen wurde verworfen. Dieser Vorgang wurde wiederholt. Um alle Spuren von Ethanol aus dem Säulchen zu entfernen, wurde zum Schluss nochmals bei maximaler Geschwindigkeit für 2 Minuten zentrifugiert.

Das Säulchen wurde in ein 1,5 ml Reaktionsgefäß überführt und 50-100 µl Elutionspuffer (TE) hinzugefügt. Nach einer einminütigen Inkubation bei Raumtemperatur wurde erneut bei 11.000 U/min für 1 Minute zentrifugiert. Dieser Elutionsschritt wurde wiederholt, um eine höhere Ausbeute an extrahierter DNA zu erzielen.

Die extrahierte DNA wurde bei -20 °C im Gefrierschrank aufbewahrt.

2.2.3 Kontrolle und Vermessung der DNA-Extrakte

Das Vorhandensein der extrahierten DNA wurde mittels SB-Agarosegel-Gelelektrophorese überprüft. Ein zuvor hergestelltes 1 %-iges SB-Gel ([2.3.1 SB-Gel Herstellung](#)) wurde in eine Elektrophoresekammer (Sub-Cell®, GT, Bio-Rad) eingelegt, die mit 1x SB-Puffer gefüllt war. Die Geltaschen wurden anschließend mit den DNA-Extrakten beladen. Hierzu wurden auf einem Stück Parafilm kleine Tropfen mit etwa 0,3 µl 6x DNA Loading Dye (CarlRoth GmbH & Co. KG, Karlsruhe, Deutschland) aufgetragen und nacheinander mit jeweils 2 µl der DNA-Extrakte vermischt. Die so vorbereiteten Proben wurden vorsichtig in die Geltaschen pipettiert. Als Referenz diente eine 1 kbp molekulare Leiter (Firma Carl Roth GmbH + Co. KG, Karlsruhe, Deutschland), von der jeweils 1 µl in die erste sowie letzte Geltasche pipettiert wurde.

Nach dem Aufsetzen des Kammerdeckels und dem Anschließen an das Netzgerät (PowerPac 300 Netzgerät, Bio-Rad) wurde das Gel bei 120 V für 50 Minuten elektrophoretisch aufgetrennt. Aufgrund ihrer negativen Ladung wandern die DNA-Fragmente im elektrischen Feld vom negativen zum positiven Pol, wobei die unterschiedlichen Längen der Fragmente eine Trennung nach Größe ermöglichen, die sich als Banden im Gel manifestieren. Größere Fragmente befinden sich dabei näher am oberen Ende des Gels.

Basierend auf der molekularen Leiter konnten die Fragmentlängen abgeschätzt werden. Das erwartete Produkt für das Chalkonsynthase-Gen sollte zwischen 1500 – 1600 bp sichtbar sein. (Yuxing Deng, 2018)

Nach Abschluss der Elektrophorese wurde das Gel in eine Dunkelhaube mit UV-Licht (Darkhood DH-50, Biostep, Burkhartsdorf, Deutschland) gelegt und mittels geeigneter Software (Biostep Argus X1 Gel Documentation System) visualisiert, abfotografiert und dokumentiert. Nach erfolgreicher Überprüfung konnte mit der PCR begonnen werden.

Die DNA-Konzentration und Reinheit der Extrakte wurden mittels Mikrovolumen-Spektralphotometer (DeNovix DS-11 Series) überprüft. Zunächst wurde das Gerät mit 1,3 µl Milli-Q®-Wasser (Merck KGaA, Darmstadt, Deutschland) gereinigt und mit 1,3 µl Elutionspuffer kalibriert. Aus jedem DNA-Extrakt wurden 1,3 µl auf den Sensorspot des Geräts pipettiert und gemäß den Herstelleranweisungen vermessen. Nach jeder Messung wurde der Sensor mit einem feinen Tuch gereinigt.

2.3 Agarose-Gelelektrophorese

Die Agarose-Gelelektrophorese ist eine etablierte Methode zur Qualitätskontrolle von DNA-Proben. In dieser Studie wurden zwei unterschiedliche Gele verwendet: ein SB-Gel (Natriumborat) zur Überprüfung der amplifizierten DNA-Produkte und ein TAE-Gel (Tris-Acetat-EDTA) zur anschließenden Extraktion der Produkte.

2.3.1 SB-Gel Herstellung

Für die Herstellung eines 1 % SB-Gels wurden 1,5 g Agarose (Carl Roth GmbH + Co. KG, Karlsruhe, Deutschland) in einer 500 ml Glasflasche mit 150 ml 1x SB-Puffer (Carl Roth GmbH + Co. KG, Karlsruhe, Deutschland) versetzt. Ein Magnetrührstäbchen wurde hinzugefügt. Die Flasche mit der Agarose-Lösung wurde mit leicht geöffnetem Deckel in der Mikrowelle (Daewoo KOR 63D7 CRS Concave reflex system, Bupjong, Südkorea) auf mittlerer Stufe erhitzt, bis die Agarose vollständig gelöst war. Zur Abkühlung wurde die Flasche in einem Becherglas mit kaltem Wasser auf einen Magnetrührer (IKA®, VAMED Heatin magnetic stirrer, VELP® Scientific, Staufen, Deutschland) gestellt. In der Zwischenzeit wurden die offenen Seiten des Gelschlittens mit Klebeband abgedichtet, die Kämme (meist 20er) eingesetzt und mittels Dosenlibelle für eine horizontale Ausrichtung des Schlittens gesorgt.

Nachdem die Agarosegel-Lösung etwas abgekühlt war, wurden 3 µl peqGREEN (Fluoreszenzfarbstoff) (VWR International GmbH, Pennsylvania, USA) hinzugefügt und durch kurzes Rühren am Magnetrührer gleichmäßig verteilt. Die Agaroselösung wurde in den vorbereiteten Gelschlitten gegossen, eventuell vorhandene Luftblasen wurden mit einer Pipettenspitze entfernt. Nach 20-30 Minuten Erstarrungszeit war das Agarosegel fest genug und die Klebebänder sowie Kämme wurden vorsichtig entfernt. Das SB-Gel wurde anschließend in drei gleich große SB-Gele geschnitten und im Kühlschrank gelagert, wo es bis zu drei Tage aufbewahrt werden konnte.

2.3.2 TAE-Gel Herstellung

Zur Herstellung des TAE-Gels wurde zunächst der Rotiphorese® 10x TAE-Puffer (Carl Roth GmbH + Co. KG, Karlsruhe, Deutschland) mit Milli-Q®-Wasser im Verhältnis 1:10 verdünnt. Die Herstellung des 1 % TAE-Gels erfolgte analog zur Beschreibung des SB-Gels ([2.3.1 SB-Gel Herstellung](#)), mit dem Unterschied, dass 1x TAE-Puffer anstelle des SB-Puffers verwendet wurde.

2.4 PCR

Die benötigten Reagenzien für die PCR wurden im Vorfeld vorbereitet und im Gefrierschrank gelagert. Für eine 5 % Dimethylsulfoxid (DMSO)-Lösung (Thermo Fischer Scientific, Waltham, USA) wurden 5 µl DMSO zu 95 µl Milli-Q®-Wasser pipettiert. Ein 20 mM dNTP-Mix (Solis BioDyne OÜ, Tartu, Estland) wurde mit Milli-Q®-Wasser auf 10 mM dNTP verdünnt (1.3.3 PCR).

2.4.1 Primer Design

Das Primer-Design erfolgte mithilfe der Software Primer3 innerhalb des Geneious Prime Programm (Version 2024.03). Als Vorlage diente Salbei (S1_Sm1_1F), da diese Pflanze aufgrund ihres Verwandtschaftsgrades dem Majoran (*Origanum majorana*) ähnlich ist. Die Plattform NIH (National Institutes of Health) erwies sich ebenfalls als nützlich für die Primer-Entwicklung, indem sie die nächstverwandte Pflanze anhand der Basenabfolge identifizierte. Die Parameter für das Primer-Design umfassten Basenlänge, GC-Gehalt, Schmelztemperatur (T_m), Primer-Dimer und Sekundärstruktur.

Um möglichst effiziente PCR-Produkte zu erzielen, wurden im Rahmen der Masterarbeit verschiedene genspezifische Primer-Kombinationen sowie unterschiedliche DNA-Verdünnungen (1:10, 1:20, 1:50, 1:100) getestet. Für die kultivierte Probe 230271 wurde der Primer S1_Sm1_1F (Tabelle 2) aus einer früheren Studie zur Genexpression von Salbei verwendet. Für die Untersuchung der wilden Proben mussten jedoch neue Primer entworfen werden.

Die Primer-Kombination S1_Sm1_1F und S1_Sm1_1165R erwies sich für die Probe 230271 als optimal. Die Kombination S1_Sm1_87F und S1_Sm1_1147R zeigte sich erfolgreich in den Proben M85, Syr 18/12 und 230283. Für die Proben M105 und M114 wurde die Kombination S1_Sm1_137F und S1_Sm1_1165R als ideal ermittelt. Die sortenspezifisch unterschiedlichen Bindungseffizienzen der Primer könnten durch SNPs (Single Nucleotide Polymorphismen) erklärt werden, die die Primerbindung beeinträchtigen können.

Die ausgewählten Primer (Merck KGaA, Darmstadt, Deutschland) wurden zunächst bei 10.000 U/min für 1 Minute zentrifugiert. Danach wurden sie mit TE-Puffer (Carl Roth GmbH + Co. KG, Karlsruhe, Deutschland) zu 100 µM (100 pmol/µl) Stocklösungen angesetzt. Die benötigte

Menge an TE-Puffer richtete sich nach der Molmasse der jeweiligen Primer. Nach einer fünfminütigen Inkubationszeit wurden die Stocklösungen für 5-8 Sekunden gevortext und anschließend aus diesen Stocklösungen, 10 μM (10 pmol/ μl) Verdünnungen mit Milli-Q®-Wasser für die PCR-Reaktionen hergestellt.

Tabelle 2: verwendete Primer

Primername	verwendete Probe	5'-3' Sequenz	Richtung	Schmelztemperatur [°C]	GC [%]	Primer Dimer	Sekundäre Strukturen	Annealing [°C]
S1_Sm1_1F	230271 (kultiviert)	ATGGTGACCGTGGAGGAGAT	forward	65,8	55	nein	keine	63
S1_Sm1_87F	M85 (wild); Syr18/12 (wild); 230283 (wild)	TGCGTCGACCAAGCGAATA	forward	68,7	50	nein	schwach	62
S1_Sm1_137F	M105 (wild); M114 (wild)	CGAGCACAAAGACCGAACTCA	forward	66,6	55	nein	schwach	63
S1_Sm1_1147R	M85 (wild); Syr18/12 (wild); 230283 (wild)	CACCGTCTCCACCGTGA	reverse	65	64,7	nein	schwach	62-63
S1_Sm1_1165R	230271 (kultiviert); M105 (wild); M114 (wild)	TAGCGACGCTGTGAAGGACC	reverse	67,7	60	nein	schwach	62-63
Δ12_S1_Sm1_1F	230271 (kultiviert)	AAAAAGCAGGCTATGGTGACCGTGGAGGAGAT	forward	77,7	50	nein	keine	63
Δ12_S1_Sm1_87F	M85 (wild); Syr18/12 (wild); 230283 (wild)	AAAAAGCAGGCTTTCGTCGACCAAGCGAATA	forward	79,7	46,8	nein	moderat	62
Δ12_S1_Sm1_137F	M85 (wild); Syr18/12 (wild); 230283 (wild)	AAAAAGCAGGCTCGACCAAGACCGAACTCA	forward	79,6	50	nein	schwach	63
Δ12_S1_Sm1_1147R	M105 (wild); M114 (wild)	AGAAAGCTGGGTTACCGTCTCCACCGTGA	reverse	80,4	58,6	nein	moderat	62-63
Δ12_S1_Sm1_1165R	230271 (kultiviert); M105 (wild); M114 (wild)	AGAAAGCTGGGTTAGCGACGCTGTGAAGGACC	reverse	79,7	56,2	nein	schwach	62-63
attB Adapter 1	alle Proben	GGGGACAAGTTTGTACAAAAAAGCAGGCT	forward	72,8	44,8	nein	sehr schwach	60
attB Adapter 2	alle Proben	GGGGACCACCTTTTGTACAAAGAAAGCTGGGT	reverse	74,6	51,7	nein	schwach	60

2.4.2 Entry-PCR

Für die 20 µl umfassende Entry-PCR wurde ein Mastermix vorbereitet, der pro Probe folgende Bestandteile enthielt: 11,76 µl autoklaviertes Milli-Q®-Wasser, 4 µl 5x Phusion HF-Puffer, 1 µl 5 % DMSO, 0,24 µl Phusion HS II (2 U/µl) (Thermo Fisher Scientific, Waltham, USA), 0,4 µl dNTP-Mix und je 0,8 µl Forward- sowie Reverse-Primer (10pmol/µl). Der Mastermix wurde kurz gevortext und aliquotiert, wobei jeweils 19 µl in 0,2 ml PCR-Reaktionsgefäße pipettiert und anschließend 1 µl DNA hinzugefügt wurde. Zudem wurde eine No-Template Control (NTC) mit denselben Reagenzien, jedoch ohne DNA, zur Erkennung von Kontaminationen sowie Prävention von Fehlinterpretationen angesetzt.

Die PCR-Reaktionsgefäße wurden kurz in einer Minizentrifuge (Sprout™, Biozym, Heathrow Scientific, Illinois, USA) zentrifugiert und anschließend im entsprechend programmierten Thermocycler (Proflex Cyclor, PCR System, Applied Biosystems) inkubiert. Die PCR-Bedingungen umfassten einen initialen Denaturierungsschritt bei 98 °C für 50 Sekunden, gefolgt von 40 Zyklen mit spezifischen Annealing-Temperaturen (98/62-63/72 °C) und -Zeiten (10/30/90 sec) innerhalb der Spezies sowie einem abschließenden Verlängerungsschritt bei 72 °C für 7 Minuten.

Die zugehörigen Primer zur entsprechenden Probe innerhalb der Spezies sind der Tabelle 2 zu entnehmen.

2.4.2.1 SB-Gelbild

Die PCR-Produkte wurden auf einem 1 %-igen SB-Agarosegel ([2.3.1 SB-Gel Herstellung](#)) überprüft. Dafür wurden jeweils 2 µl PCR-Produkt mit 1 µl 6x Loading Dye auf einem Parafilmstreifen vermischt und anschließend in die Geltasche pipettiert. Das Gel ist für 50 Minuten bei 120 Volt gelaufen. Sobald die Banden auf der erwarteten Höhe sichtbar waren, wurden die Produkte auf ein TAE-Gel aufgetragen ([2.3.2 TAE-Gel Herstellung](#)).

2.4.2.2 TAE-Gelbild

Die PCR-Produkte wurden auf ein 1 %-iges TAE-Agarosegel ([2.3.2 TAE-Gel Herstellung](#)) aufgetragen, wobei die Schritte gemäß [2.4.2.1 SB-Gelbild](#) befolgt wurden, mit dem Unterschied, dass hier die verbleibenden Volumina der PCR-Produkte mit einem entsprechenden Volumen an 6x Loading Dye vermischt und vorsichtig in die Geltaschen pipettiert wurden. Um eine klarere Auflösung zu gewährleisten und etwaige Überlappungen zu vermeiden, wurde empfohlen, eine Tasche zwischen den Proben freizulassen und je nach Probenanzahl ein oder mehrere Geltaschen mit einer 1x 1 kbp DNA-Leiter zu füllen. Für diese Gelelektrophorese wurden 120 V über 60 Minuten angewendet.

Nach der Elektrophorese wurden die ausgewählten Amplifikationsprodukte mithilfe einer Skalpellklinge entlang der Banden ausgeschnitten, in 2 ml Reaktionsgefäße überführt und extrahiert.

2.4.2.3 Gelextraktion und Aufreinigung der Amplifikationsprodukte

Die PCR-Amplifikationsprodukte wurden mit dem InnuPREP DOUBLEpure Kit (IST Innuscreen GmbH, Berlin, Deutschland) gemäß den Herstelleranweisungen extrahiert.

Vorab wurde der Waschlösung, die entsprechende Menge Ethanol hinzugefügt, der Thermoblock auf 50 °C vorgeheizt und die Zentrifuge eingeschaltet. Der Elutionspuffer (EB) wurde bis zur Verwendung im Thermoblock bei 50 °C gelagert, da der finale Elutionsschritt mit erhitztem Elutionspuffer die DNA-Ausbeute erhöhen kann.

In die 2 ml Reaktionsgefäße mit den DNA-Gelprodukten wurden 650 µl Gel-Solubilizer hinzugefügt und anschließend im Thermoblock für 10 Minuten bei 50 °C und kontinuierlichem Schütteln inkubiert, bis das Agarosegel vollständig aufgelöst war. Dann wurden jeweils 50 µl Binding Optimizer hinzugefügt und die Suspension gut gevortext. Die Proben wurden auf Säulchen übertragen, die sich in Auffanggefäßen befanden. Nach einer Zentrifugation bei 11.000 U/min für 1 Minute wurde das Filtrat verworfen.

Im nächsten Schritt wurden jeweils 700 µl Waschlösung LS hinzugefügt und erneut bei 11.000 U/min für 1 Minute zentrifugiert. Das Filtrat wurde verworfen, der Waschschrift wiederholt und zum vollständigen Entfernen des Ethanols bei maximaler Geschwindigkeit für

2 Minuten zentrifugiert. Das Auffanggefäß wurde verworfen und das Säulchen in ein 1,5 ml Reaktionsgefäß überführt.

Es wurden 15 µl Elutionspuffer vorsichtig auf das Säulchen pipettiert, bei Raumtemperatur für 1 Minute inkubiert und anschließend bei 11.000 U/min für 1 Minute zentrifugiert. Ein zweiter Elutionsschritt wurde durchgeführt, um die DNA-Ausbeute zu erhöhen. Die aufgereinigten Amplifikationsprodukte wurden bis zur weiteren Verwendung bei 4 °C im Kühlschrank aufbewahrt.

Die Qualität der Extraktionsprodukte wurde mittels SB-Gel (2.4.2.1 SB-Gelbild) überprüft und nach zufriedenstellenden Ergebnissen für die weiterführende $\Delta 12$ -PCR-Reaktion verwendet.

2.4.3 $\Delta 12$ -PCR

Für die $\Delta 12$ -PCR wurde ein 19 µl Mastermix vorbereitet (2.4.2 Entry-PCR) und anschließend mit 1 µl DNA, die durch die Aufreinigung der Entry-PCR gewonnen wurde, gemischt und kurz zentrifugiert.

Die PCR-Bedingungen ähneln denen der Entry-PCR, weisen jedoch eine Modifikation auf: eine Inkubation bei 72 °C für 60 Sekunden, die sich als besonders effizient erwiesen hat.

Anschließend wurde ein SB-Gelbild zur Überprüfung der $\Delta 12$ -PCR-Produkte erstellt (2.4.2.1 SB-Gelbild). Nach einem positiven Gelbild wurde ein TAE-Gel pipettiert (2.4.2.2 TAE-Gelbild), gefolgt von einer Gelextraktion (2.4.2.3 Gelextraktion und Aufreinigung der Amplifikationsprodukte).

Die gewonnenen DNA-Extrakte wurden für die weiterführende attB-PCR verwendet.

2.4.4 attB-PCR

Der finale PCR-Schritt wurde unter Verwendung von Primern mit den entsprechenden attB-Stellen (attB1 und attB2) durchgeführt. Die optimalen PCR-Bedingungen umfassen: 98 °C für 10 Sekunden, 60 °C für 30 Sekunden und 72 °C für 60 Sekunden. Alternativ können die Temperaturbedingungen der Entry-PCR beibehalten werden, wobei jedoch die

Verlängerungszeit bei 72 °C auf 60 Sekunden reduziert und die Primer-Konzentration verdoppelt wird.

Bei der Verwendung einer doppelten Konzentration der attB-Primer wurde die Menge des autoklavierten Milli-Q®-Wassers um 1,2 µl reduziert. Alle übrigen Reagenzienmengen wurden gemäß den Vorgaben in 2.4.2 Entry-PCR pipettiert, um ein Gesamtvolumen von 19 µl zu gewährleisten. Anschließend wurden noch jeweils 1 µl DNA, die durch die Aufreinigung der Δ12-PCR gewonnen wurde, hinzu pipettiert, vermischt und zentrifugiert.

Zur Überprüfung der Produkte wurde ein SB-Gelbild hergestellt (2.4.2.1 SB-Gelbild). Nach einem positiven Gelbild wurde ein TAE-Gelbild pipettiert (2.4.2.2 TAE-Gelbild), gefolgt von einer Gelextraktion (2.4.2.3 Gelextraktion und Aufreinigung der Amplifikationsprodukte).

Anschließend wurden die aufgereinigten DNA-Extrakte mittels SB-Gelbild und Mikrovolumen-Spektralphotometer (2.2.3 Kontrolle und Vermessung der DNA-Extrakte) auf Qualität und Quantität überprüft.

Sobald zufriedenstellende Ergebnisse erreicht wurden, konnte zum nächsten Schritt der Klonierung übergegangen werden. Die Extrakte wurden bei -20 °C im Gefrierschrank aufbewahrt.

2.5 Gateway-Klonierung

2.5.1 BP-Klonierungsreaktion

Die BP-Reaktion stellt einen wesentlichen Schritt der Gateway®-Klonierung dar. Um die korrekten Mengen der benötigten Reagenzien (DNA, Gateway® BP Clonase® II Enzym-Mix, pDONR™-Vektor, TE-Puffer) zu ermitteln, wurde die erforderliche Menge an DNA (in ng) basierend auf der Fragmentlänge berechnet. Das Chalkonsynthase-Gen besitzt eine Länge von 1551 bis zu 1565 Basenpaaren.

Zur Umrechnung von fmol-DNA in ng-DNA wurde die folgende Formel verwendet, wobei N die Größe des DNA-Fragments in Basenpaaren (bp) darstellt (Formel 1). Für die BP-Reaktion wurden 50 fmol des attB-PCR-Produkts benötigt, dessen Länge auf 1,6 kb gerundet wurde. Daraus ergab sich eine Gesamtmenge von 52,8 ng DNA.

Formel 1: Umrechnung von fmol in ng (Life Technologies, 2012, Cayla, SA)

$$\text{ng} = (\text{fmol})(N)\left(\frac{660 \text{ fg}}{\text{fmol}}\right)\left(\frac{1 \text{ ng}}{10^6 \text{ fg}}\right)$$

Auf Basis der ermittelten DNA-Konzentrationen sowie der berechneten Gesamtmenge von 52,8 ng DNA wurden die benötigten Mengen der Reagenzien für die BP-Reaktion festgelegt. Die Reaktion wurde gemäß den Herstellerangaben durchgeführt, wobei sie über einen Zeitraum von drei Stunden angesetzt wurde.

Beispielsweise wurde für die Probe M85 eine DNA-Konzentration von 13,922 ng/μl gemessen. Für einen Reaktionsansatz von 10 μl wurden 3,81 μl DNA, 1 μl des Gateway™ pDONR™ 221-Vektors, 2 μl des Gateway™ BP Clonase™ II Enzym-Mix (Thermo Fisher Scientific, Waltham, USA) sowie 3,19 μl TE-Puffer gemischt. Nach dem Pipettieren wurde der Ansatz durch Vortexen homogenisiert, der Reaktionsdeckel mit Parafilm verschlossen und die Mischung anschließend bei 25 °C für drei Stunden im Thermoblock inkubiert.

2.5.2 Vorbereitungen auf die Transformation

Das resultierende Plasmid wurde mittels Hitzeschock in chemisch kompetente *E. coli* transformiert und anschließend auf Antibiotika-haltigen Agarplatten ausgestrichen.

2.5.2.1 LB-Medium Herstellung

Für die Herstellung von 300 ml LB-Flüssigmedium wurden 3 g Trypton/Pepton, 1,5 g Hefeextrakt und 3 g NaCl (Natriumchlorid) (Carl Roth GmbH + Co. KG, Karlsruhe, Deutschland) in einem Messbecher eingewogen und anschließend in eine Glasflasche umgefüllt. Danach wurden 300 ml Milli-Q®-Wasser mit einem Messzylinder abgemessen und ebenfalls in die Glasflasche gegeben. Das LB-Medium wurde autoklaviert und anschließend im Kühlschrank gelagert.

2.5.2.2 AB-Agar-Platten Herstellung

Zur Herstellung von 600 ml LB-Agar-Lösung wurden 6 g Trypton/Pepton, 3 g Hefeextrakt, 6 g NaCl (Natriumchlorid) und 9 g Agar-Agar (Carl Roth GmbH + Co. KG, Karlsruhe, Deutschland) in einem Messbecher eingewogen und in eine Glasflasche überführt. 600 ml Milli-Q®-Wasser wurden mit einem Messzylinder abgemessen und in die Flasche hinzugegeben. Nachdem die Lösung auf 60 °C abgekühlt war, konnte die Arbeit fortgesetzt werden.

Als Antibiotikum wurde Kanamycinsulfat mit einer Endkonzentration von 50 mg/ml (entspricht 5 g/100 ml) (Carl Roth GmbH + Co. KG, Karlsruhe, Deutschland) verwendet. Das Antibiotikum wurde mit Milli-Q®-Wasser gelöst und durch eine sterile Nylonmembran mit 0,2 µm filtriert.

Die LB-Agar-Lösung mit dem Antibiotikum wurde anschließend gevortext und unter sterilen Bedingungen in einem Laminar-Air-Flow-Schrank auf Agarplatten gegossen. Nach dem Erstarren wurden die Agarplatten seitlich beschriftet und bei 4 °C im Kühlraum gelagert.

2.5.2.3 Herstellung einer 100 mM CaCl₂-Lösung

Zur Herstellung einer 300 ml 100 mM CaCl₂-Lösung wurden 4,41 g CaCl₂ (Carl Roth GmbH + Co. KG, Karlsruhe, Deutschland) auf einem Filterpapier eingewogen und in eine Glasflasche überführt. Anschließend wurden 300 ml Milli-Q®-Wasser hinzugegeben, die Lösung autoklaviert und danach im Kühlschrank gelagert.

Herstellung einer 100 mM CaCl₂-Lösung (300ml)

CaCl₂ x 2 H₂O -> MG = 147,02 g/mol

20 mM = 0,294 g für 100ml

100 mM = 1,47 für 100 ml

➔ **4,41 g für 300 ml**

2.5.2.4 *Escherichia coli* Bakterien Anzucht

Die Anzucht von kompetenten *E. coli* erfolgte an zwei aufeinanderfolgenden Tagen. Am ersten Tag wurden zweimal 50 ml LB-Medium mit Original One Shot™ TOP10 Chemically Competent *E. coli* (Thermo Fisher Scientific, Waltham, USA) inokuliert und in einem Schüttelwasserbad (Inkubationswasserbad, 1013 GFL®, VAMED Medizin und Labortechnik, Burgwedel, Deutschland) bei 37 °C und 220 U/min über Nacht inkubiert.

Am zweiten Tag erfolgte die Wachstumskontrolle und die Kultur mit der höheren Trübung wurde für die weiteren Schritte verwendet. Die optische Dichte (OD) der Kultur wurde mit einem Spektrophotometer (UV-2700, UV/Vis-Spektrophotometer, Shimadzu, Duisburg, Deutschland) bei 600 nm (OD₆₀₀) gemessen. Eine OD₆₀₀ von etwa 0,4 wurde angestrebt, da diese die höchste Potenz sowie optimale Bedingungen für effektives Wachstum gewährleistet. Unsere Kultur hatte anfangs eine OD₆₀₀ von 2,47 und wurde auf eine OD₆₀₀ von 0,402 mit LB-Medium runter verdünnt.

Nach Erreichen der optimalen OD wurden die *E. coli*-Kulturen in 50 ml Falcon™-Röhrchen überführt und für 10 Minuten auf Eis gelagert.

Die Proben wurden bei 5400 U/min und 4 °C für 10 Minuten zentrifugiert. Der Überstand wurde verworfen und zu der Hauptkultur wurden 25 ml der zuvor vorbereiteten eiskalten 100 mM Calciumchlorid-Lösung hinzugefügt sowie vorsichtig gevortext. Im nächsten Schritt wurde die Lösung für 20 Minuten auf Eis gestellt und erneut bei denselben Parametern für 10 Minuten zentrifugiert. Der Überstand wurde verworfen.

Danach wurden 5 ml 100 mM Calciumchlorid-Lösung mit der Hauptkultur vermischt und gevortext, gefolgt von einer 1,5-stündigen Inkubation auf Eis. Zur langfristigen Lagerung wurden 4 ml der Hauptkultur mit 1 ml 60 % Glycerin (wirkt hygroskopisch) vermischt und vorsichtig geschwenkt.

Anschließend wurden jeweils 50 µl der kompetenten *E. coli* in 2 ml Reaktionsgefäße aliquotiert, in flüssigem Stickstoff eingefroren und bei -80 °C gelagert, um eine Haltbarkeit von bis zu 6 Monaten zu gewährleisten.

2.5.3 Transformation

Die Transformation erfolgte gemäß den Angaben von Gateway Thermo Fisher Scientific.

Während der BP-Reaktion wurde Eis in einem geeigneten Behältnis bereitgestellt, das Wasserbad auf 42 °C und der Inkubator auf 37 °C vorgeheizt sowie das S.O.C.-Medium (Thermo Fisher Scientific, Waltham Deutschland) aus dem Kühlschrank genommen. Für jede Probe wurden zwei Kanamycin-50-Agarplatten ([2.5.2.2 AB-Agar-Platten Herstellung](#)) beschriftet und bei Raumtemperatur gelagert.

Um die BP-Reaktion zu stoppen, wurde 0,5 µl Proteinase K in das Reaktionsgefäß pipettiert, die Mischung gevortext und für 10 Minuten bei 37 °C im Thermoblock inkubiert. Während der Inkubation, wurden die chemisch kompetenten *E. coli* aus dem -80 °C Tiefkühler geholt und auf Eis gelegt. Damit der Ansatz nicht zu warm ist, wurde das Reaktionsgefäß auf Eis gelegt.

Die BP-Reaktionsmischung wurde zu den *E. coli* pipettiert und für 20 Minuten auf Eis inkubiert. Das Plasmid wurde durch Hitzeschock bei 42 °C im Wasserbad (Inkubationswasserbad, 1013 GFL®, VAMED Medizin und Labortechnik, Burgwedel, Deutschland) für 60 Sekunden, in die *E. coli* eingeführt, gefolgt von einer zweiminütigen Inkubation auf Eis.

Im weiteren Schritt wurden 250 µl S.O.C.-Medium in die *E. coli* gegeben, der Deckel mit einem Parafilm verschlossen und für 1 Stunde bei 37 °C im Inkubator (Mettler IN55, Mettler GmbH & Co. KG, Schwabach, Deutschland) unter kontinuierliches über Kopf Drehen (Mini Rotator, biosan, Bio RS-24, PRS-22, Lettland) inkubiert.

Je Probe wurden die *E. coli* auf zwei Kanamycin-50-Agarplatten (1x 150 µl und 1x die Restflüssigkeit) mit einem Spatel ausgestrichen und über Nacht bei 37 °C im Inkubator gelassen.

Am nächsten Tag erfolgte eine Wachstumskontrolle und die Platten wurden in den Kühlschrank gestellt. Bei sichtbaren Kolonien konnte der nächste Schritt eingeleitet werden.

2.5.4 Kolonien picken

Sterile Epprovetten mit 5 ml LB-Flüssigmedium und 5 µl Kanamycin (50 mg/ml) wurden vorbereitet und dokumentiert. Jede Epprovette wurde mit einem Stück Alufolie abgedeckt. Anschließend wurden ausgewählte Kolonien markiert und mithilfe einer sterilen Plastikimpföse vorsichtig von der Agarplatte in das Kanamycin-LB-Medium überführt. Die abgedeckten Epprovetten wurden dann über Nacht bei 37 °C und 220 U/min in einem Schüttelwasserbad inkubiert.

2.5.5 Plasmid-DNA-Extraktion

Am nächsten Tag wurden die Plasmide mit dem innuPREP Plasmid Mini Kit 2.0 (IST Innuscreen GmbH, Berlin, Deutschland) gemäß den Herstelleranweisungen extrahiert.

Zunächst wurde die entsprechende Menge an Ethanol zur Waschlösung B hinzugefügt. Von den 5 ml Ansatz wurden jeweils 4 x 1 ml Flüssigkultur bei 14.000 U/min für 1 Minute zentrifugiert und der Überstand verworfen. Das verbliebene bakterielle Pellet wurde mit 250 µl Resuspensionspuffer resuspendiert und bis zur vollständigen Auflösung gevortext.

Im weiteren Verlauf wurden 250 µl Lysepuffer hinzugefügt und 6- bis 8-mal invertiert. Dieser Schritt sollte nicht länger als 5 Minuten dauern. Anschließend wurden 350 µl Neutralisationspuffer hinzugegeben, 6- bis 8-mal invertiert und bei voller Geschwindigkeit für 8 Minuten zentrifugiert.

In der Zwischenzeit wurden die Auffanggefäße mit den Säulchen vorbereitet. Der klare Überstand wurde auf das Säulchen übertragen und bei voller Geschwindigkeit für 1 Minute zentrifugiert. Das Filtrat wurde verworfen.

Zur Entfernung von Verunreinigungen wurden 500 µl Waschlösung A auf das Säulchen pipettiert und bei 11.000 U/min für 1 Minute zentrifugiert. Das Filtrat wurde verworfen und der Waschschrift mit 700 µl Waschlösung B wiederholt. Anschließend wurde nochmals bei maximaler Geschwindigkeit für 2 Minuten zentrifugiert, um alle Ethanolspuren zu entfernen. Das Auffanggefäß wurde verworfen und das Säulchen in ein 1,5 ml Reaktionsgefäß gegeben.

Im letzten Schritt wurde das Sulchen vorsichtig mit 30 µl TE-Puffer angefeuchtet, fur 2 Minuten inkubiert und bei 11.000 U/min fur 1 Minute zentrifugiert. Das Sulchen wurde verworfen und das extrahierte Plasmid bei -20 °C im Gefrierschrank aufbewahrt.

2.5.6 Kontrolle und Vermessung der Plasmid-DNA

Um zu uberprufen, ob die Kolonien die Plasmide mit den gewunschten Inserts enthalten, wurde ein SB-Gelbild von den Plasmid DNA Extrakten angefertigt und die Konzentrationen mittels Mikrovolumen-Spektralphotometer vermessen (2.2.3 Kontrolle und Vermessung der DNA-Extrakte). Zur Kalibrierung diente der TE-Puffer.

2.6 Sequenzierung

Die Sequenzierung wurde durch Microsynth (Wien, osterreich) durchgefuhrt. Es wurde eine Gesamtmenge von 1200 ng DNA verwendet. Entsprechend der DNA-Konzentration wurde ein Ansatzvolumen von 15 µl in Mikro-Schraubrohrchen pipettiert, wobei die benotigte Menge an DNA und Milli-Q®-Wasser hinzugefugt und die Proben ordnungsgema etikettiert wurden.

2.7 Bioinformatik

Insgesamt wurden 134 Sequenzen erhalten, die mit geeigneter Software analysiert und identifiziert wurden. Die Auswertung und Assemblierung der DNA-Sequenzen erfolgte unter Verwendung von Geneious Prime (Version 2024.0.3). Zur Identifizierung und Abgrenzung der Sequenzen wurde BLAST (Basic Local Alignment Search Tool) eingesetzt, wobei zum Vergleich Sequenzen aus den NCBI-Datenbanken (National Center for Biotechnology Information) herangezogen wurden. Erste Analysen, insbesondere zur Untersuchung der SNP-Schnittstellen (Single Nucleotide Polymorphismen), wurden mit der Software MEGA (Version 11.0.13) durchgeführt.

Für weiterführende Analysen der erhaltenen DNA-Sequenzen kamen die Software R (Version 4.4.1) sowie Expasy (Expert Protein Analysis System, Swiss Institute of Bioinformatics) zum Einsatz. Um eine funktionelle Analyse zu ermöglichen, wurde versucht, die 3D-Struktur einer ausgewählten Sequenz vorherzusagen.

3. Ergebnisse

Das Ziel dieser Studie war es, einen umfassenden Datensatz von Sequenzkopien der Chalkonsynthase zu generieren, um evolutionäre Genveränderungen innerhalb und zwischen verschiedenen Individuen der Gattung *Origanum* zu untersuchen. Durch den Einsatz des Gateway-Klonierungsverfahrens sollten Rückschlüsse auf Mutationen, Single Nucleotide Polymorphismen (SNPs), Multilocus-Genotypen (MLGs) und Sequenzidentitäten gezogen werden. Zur Analyse der gewonnenen Daten wurden die Softwares Geneious, MEGA, R, Expasy (-SWISS) und BLAST verwendet.

Die erhaltenen Klone wurden vor Beginn der detaillierten Analyse mittels SB-Gelelektrophorese untersucht, um ihre Größe zu überprüfen. Auf dem SB-Gelbild waren Banden im Bereich von 5000–6000 bp sichtbar, was der erwarteten Größe des rekombinanten Produkts entsprach. Dieses setzt sich aus der Länge des Chalkonsynthase-Gens (ca. 1558 bp) und der Länge des pDONR™-Vektors (4000 bp) zusammen.

Nach der elektrophoretischen Überprüfung wurden die Sequenzen der Klone korrigiert und einer BLAST-Analyse unterzogen, um die erfolgreiche Amplifikation des gewünschten CHS-Zielgens zu bestätigen. Anschließend erfolgte eine multiple Sequenzalignierung aller erhaltenen Sequenzen, die nach einer Nachkorrektur eingehend analysiert wurden. Bei der Analyse lag der Schwerpunkt insbesondere auf dem Nachweis von Polymorphismen, die auf DNA-, mRNA- und Aminosäureebene charakterisiert und hinsichtlich ihrer potenziellen Auswirkungen auf die Genfunktion bewertet wurden.

3.1 Genanalyse

Insgesamt konnten 134 Sequenzkopien generiert werden, die sich wie folgt auf die verschiedenen *Origanum*-Arten verteilten: 19 Sequenzen von *Origanum onites*, 20 von *Origanum syriacum*, 58 von *Origanum majorana* und 37 von *Origanum dubium*. Die Qualität der extrahierten Sequenzen ermöglichte eine präzise und detaillierte Analyse.

3.1.1 Genanalyse auf DNA-Ebene

Mit Hilfe von Daten aus NCBI-Datenbanken und der Software Geneious konnten die Exon-Intron-Grenzen in zwei Gensequenzen präzise festgelegt werden. Zum Vergleich wurden die Chalkonsynthase-Gene von *Erythranthe lewisii* (GenBank: KJ011133.1) und *Camellia sinensis* (GenBank: KC357706.1) herangezogen.

Die Analyse der Sequenzen begann mit der Identifizierung der Exon-Intron-Grenzen sowie der Ermittlung der Verteilung von SNPs und Indels (Tabelle 3). Die Gesamtlänge des Chalkonsynthase-Gens variiert je nach Auftreten eines Indels im Intron und liegt im Bereich von 1551 bis 1565 bp. Das erste Exon erstreckte sich von Position 1 bis 179 bp, gefolgt von einem Intron, das von Position 180 bis 578 bp reichte und aufgrund von Indels eine variable Länge von 384 bis 398 bp aufwies. Das zweite Exon erstreckte sich von Position 579 bis 1565 bp, mit einer Gesamtlänge von 986 bp.

Der nicht-kodierende Abschnitt der DNA ist essenziell, um die evolutionäre Variabilität innerhalb und zwischen den verschiedenen Individuen besser verstehen und vergleichen zu können. Die Ergebnisse zeigten, dass die meisten Mutationen im Intron lokalisiert waren. Zur Quantifizierung der Variationen wurde die Anzahl der Mutationen pro 100 bp berechnet. Im ersten Exon wurden 0,56 SNPs pro 100 bp beobachtet, im Intron 2,50 SNPs und 0,75 Indels pro 100 bp, und im zweiten Exon 0,70 SNPs pro 100 bp. Insgesamt wurden 18 SNPs und 3 Indels beobachtet.

Tabelle 3: DNA-Annotation

Nr.	Variation	Anfangsposition	Endposition	Länge	Typ des Polymorphismus	Häufigkeit der Varianten	
1	C/T	126	126	1	SNP	93.3%/6.7%	Exon1
2	A/G	242	242	1	SNP	72.4%/27.6%	Intron
3	T/A	276	276	1	SNP	94.8%/5.2%	Intron
4	G/T	278	278	1	SNP	56.7%/43.3%	Intron
5	T/C	304	304	1	SNP	85.1%/14.9%	Intron
6	T/G	360	360	1	SNP	85.8%/14.2%	Intron
7	T/G	364	364	1	SNP	57.5%/42.5%	Intron
8	G/C	371	371	1	SNP	84.3%/14.9%	Intron
9	C/A	380	380	1	SNP	94.0%/6.0%	Intron
10	A/T	386	386	1	SNP	70.9%/29.1%	Intron
11	-/T	390	390	1	Indel	85.1%/14.9%	Intron
12	A/G	550	550	1	SNP	57.5%/42.5%	Intron
13	--/TT	557	558	2	Indel	85.8%/14.2%	Intron
	TTTAATTAAA						
14	TA/-----	559	571	12	Indel	72.4%/27.6%	Intron
15	C/A	933	933	1	SNP	71.6%/28.4%	Exon2
16	C/T	963	963	1	SNP	85.8%/14.2%	Exon2
17	C/T	1146	1146	1	SNP	85.1%/14.9%	Exon2
18	C/T	1186	1186	1	SNP	85.8%/14.2%	Exon2
19	C/A	1227	1227	1	SNP	71.6%/28.4%	Exon2
20	A/C	1250	1250	1	SNP	85.8%/14.2%	Exon2
21	C/G	1362	1362	1	SNP	93.3%/6.7%	Exon2

Für eine detaillierte Auswertung der SNP-Verteilung wurde die Software MEGA verwendet. Die Analyse ergab Durchschnittswerte von 24 % T(U), 27 % C, 22 % A und 27 % G, wobei keine signifikanten Unterschiede festzustellen waren. Darüber hinaus wurden die Substitutionsmuster der Nukleotide im Hinblick auf identische Paare (ii), Transitionspaare (si) und Transversionspaare (sv) untersucht. Anhand der Ergebnisse wurden 1.551 identische Basenpaare, 7 Transitionspaare und 11 Transversionspaare identifiziert. Der R-Wert, der das Verhältnis von Transitions- zu Transversionspaaren beschreibt, betrug 0.7, was darauf hindeutet, dass mehr sv-Paare vorhanden waren. Dabei umfassten die polymorphen Stellen die Basenpaare TG, CT, AT, AC und GA.

Die Maximum-Likelihood-Methode (MLM) wurde angewendet, um das am besten geeignete Substitutionsmodell zur Beschreibung der DNA-Sequenzevolution zu ermitteln. Das Jukes-Cantor-Modell erwies sich als das passendste (Anhang Tab.1). Dieses einfache Modell geht davon aus, dass Substitutionen zwischen den vier Nukleotiden (A, T, C, G) mit gleicher Wahrscheinlichkeit auftreten und unabhängig voneinander sind (Arenas, 2015).

Zur Bewertung der Zuverlässigkeit der phylogenetischen Distanzen wurde die Bootstrap-Methode mit 100 Wiederholungen durchgeführt. Die berechnete Distanz (d) innerhalb der Populationen betrug 0,003554 mit einer Standardabweichung von $\pm 0,000943$, was auf eine geringe genetische Variabilität schließen lässt.

Für die statistische Analyse der Daten sowie für die Erstellung von Grafiken wurde die Software R herangezogen. Dabei wurden die SNP-Stellen bezogen auf die genetische Vielfalt untersucht und zwei wichtige Kenngrößen betrachtet (Tabelle 4). An allen Positionen wurden zwei Zustandsformen (Allele) identifiziert.

Die erwartete Heterozygotie (H_{exp}) stellt ein Maß für die genetische Variabilität in einer Population dar. Sie gibt die Wahrscheinlichkeit an, dass zwei zufällig ausgewählte Allele an einem Locus unterschiedlich sind. Je höher der H_{exp} -Wert, desto höher die Variabilität. Die H_{exp} -Werte lagen unter 0.5, was auf eine moderate genetische Vielfalt hinweist. Der höchste H_{exp} -Wert wurde an Position 278 (mit 0.495), innerhalb des Introns, festgestellt, während der niedrigste Wert interessanterweise in unmittelbarer Nähe, an Position 276 (mit 0.100), gefunden wurde, ebenfalls im Intron.

Die Evenness, ein Maß für die Gleichverteilung der Allele, zeigt an, wie gleichmäßig die Allele zwischen den Individuen verteilt sind. Werte nahe 1, wie es an den Positionen 242, 278, 364, 386, 550, 933 und 1227 der Fall war, deuten auf eine fast gleichmäßige Verteilung der einzelnen SNPs hin. An Position 278 ist der höchste Wert mit 0.982 zu finden und an Position 276 mit 0.483 der niedrigste. Dieselben Positionen wurden auch bei den H_{exp} -Werten erörtert. Jedoch liegt die Mehrheit der Werte zwischen 0.5 und 0.8, was eine ungleichmäßige Verteilung unterstreicht.

Tabelle 4: Darstellung der SNPs bezogen auf die genetische Vielfalt (Hexp ... erwartete Heterozygotie, Evenness ... Wahrscheinlichkeit einer Gleichverteilung der Allele)

DNA locus	Hexp	Evenness
126	0.126	0.513 (Exon 1)
242	0.403	0.830 (Intron)
276	0.100	0.483 (Intron)
278	0.495	0.982 (Intron)
304	0.256	0.649 (Intron)
360	0.245	0.638 (Intron)
364	0.493	0.978 (Intron)
371	0.269	0.615 (Intron)
380	0.113	0.498 (Intron)
386	0.416	0.849 (Intron)
550	0.493	0.978 (Intron)
933	0.409	0.839 (Exon 2)
963	0.245	0.638 (Exon 2)
1146	0.256	0.649 (Exon 2)
1186	0.245	0.638 (Exon 2)
1227	0.409	0.839 (Exon 2)
1250	0.245	0.638 (Exon 2)
1362	0.126	0.710 (Exon 2)

Ein Multilocus-Genotyp (MLG) eines Individuums, auch als Haplotyp bezeichnet, beschreibt die spezifische Kombination von Allelen über alle untersuchten Loci hinweg (Schmidt, 2017) (Tabelle 5). Mittels Anzahl der MLGs lässt sich die genetische Diversität innerhalb einer Population bewerten: Je mehr unterschiedliche MLGs vorhanden sind, desto höher ist die Diversität der Population (Tianhua Niu, 2002). Der H-Wert misst die genetische Diversität in Bezug auf die Anzahl der Klone (N) und die entsprechenden MLGs dieser Population. Der λ -Wert gibt die Wahrscheinlichkeit an, dass zwei zufällig ausgewählte Individuen in einer Population den gleichen MLG aufweisen. Werte nahe 1 deuten darauf hin, dass es unwahrscheinlich ist, dass alle Individuen dasselbe MLG-Muster aufweisen. Der E.5-Wert quantifiziert die Gleichmäßigkeit der Verteilung von MLGs innerhalb der Population. Der H_{exp} -Wert, wie bereits oben beschrieben, repräsentiert die erwartete Heterozygotie. Der r_{barD} -Wert ist ein Maß für ein Linkage Disequilibrium (genetische Kopplung) zwischen Loci innerhalb einer Population. Alle genannten Kenngrößen beziehen sich auf die Diversität innerhalb einer Population und nicht auf die Unterschiede zwischen den Populationen (Slatkin, 2008).

Bei *O. majorana* wurde bei 57 Klonen nur ein einziger MLG festgestellt, ebenso zeigte *O. syriacum* bei 20 Klonen nur einen MLG, wobei alle anderen Diversitätsparameter vernachlässigbar waren. Im Gegensatz dazu wies *O. dubium* bei 38 Klonen vier unterschiedliche MLGs auf, was durch einen leicht erhöhten H- und λ -Wert unterstreicht wird. Ein E.5-Wert unter 0.5 deutet darauf hin, dass innerhalb dieser Population die MLGs ungleichmäßig verteilt sind, was darauf schließen lässt, dass nicht alle Individuen dasselbe genetische Muster aufwiesen. *O. onites* zeigte die höchste genetische Variabilität: Bei 10 Klonen wurden acht verschiedene MLGs identifiziert, was sich in einem hohen H-Wert von 1,910 und einem λ -Wert von 0,831 widerspiegelt. Diese Werte deuten auf eine signifikante genetische Diversität innerhalb dieser Population hin. Zudem könnten der hohe λ - und E.5-Wert auf eine ähnliche Verteilung der MLG-Muster hindeuten. In den Populationen von *O. dubium* und *O. onites* ist der rbarD-Wert leicht erhöht, was auf eine moderate genetische Kopplung zwischen den Loci hinweist.

Tabelle 5: Darstellung der genetischen Variabilität innerhalb der untersuchten *Origanum*-Arten (N ... Anzahl der Klone; MLG ... Multilocus-Genotypen; H ... Diversitätsindex zwischen N und MLG; λ ... Maß für die Diversität, bezogen auf die MLG-Muster-Verteilung; E.5... spezifischer Evenness-Index; Hexp... erwartete Heterozygotie; rbarD... standardisierte Linkage Disequilibrium; --- keine Verteilung einer Variation bestimmbar)

Art	Abkürzung	N	MLG	H	λ	E.5	Hexp	rbarD
<i>O. dubium</i>	OD	38	4	0.363	0.150	0.402	0.00877	0.315
<i>O. majorana</i>	OM	57	1	0.000	0.000	---	0.00000	---
<i>O. onites</i>	OO	19	8	1.910	0.831	0.855	0.10834	0.193
<i>O. syriacum</i>	OS	20	1	0.000	0.000	---	0.00000	---

Die MLG-Musterverteilung zeigte deutlich, dass jede untersuchte Art einzigartige Multilocus-Genotypen aufwies (Abbildung 5). Sowohl *O. majorana* als auch *O. syriacum* zeigten jeweils nur einen MLG, MLG.8 und MLG.10, und damit keine genetische Variabilität. Im Gegensatz dazu wiesen *O. dubium* und *O. onites* eine höhere genetische Variabilität auf. Von diesen beiden Populationen zeigte *O. onites* die größte genetische Vielfalt, mit insgesamt acht unterschiedlichen MLG-Mustern (MLG.1, MLG.2, MLG.3, MLG.4, MLG.11, MLG.12, MLG.13, MLG.14), wobei MLG.1 am häufigsten beobachtet wurde. *O. dubium* zeigte eine dominierende Verteilung des MLG.7-Musters, was durch den höheren E.5-Wert erklärt werden kann. Dieses Muster wird häufiger gemeinsam weitergegeben als die Muster MLG.5, MLG.6 und MLG.9. Die Ergebnisse stehen im Einklang mit den Angaben in Tabelle 5 und untermauern die beobachteten genetischen Muster.

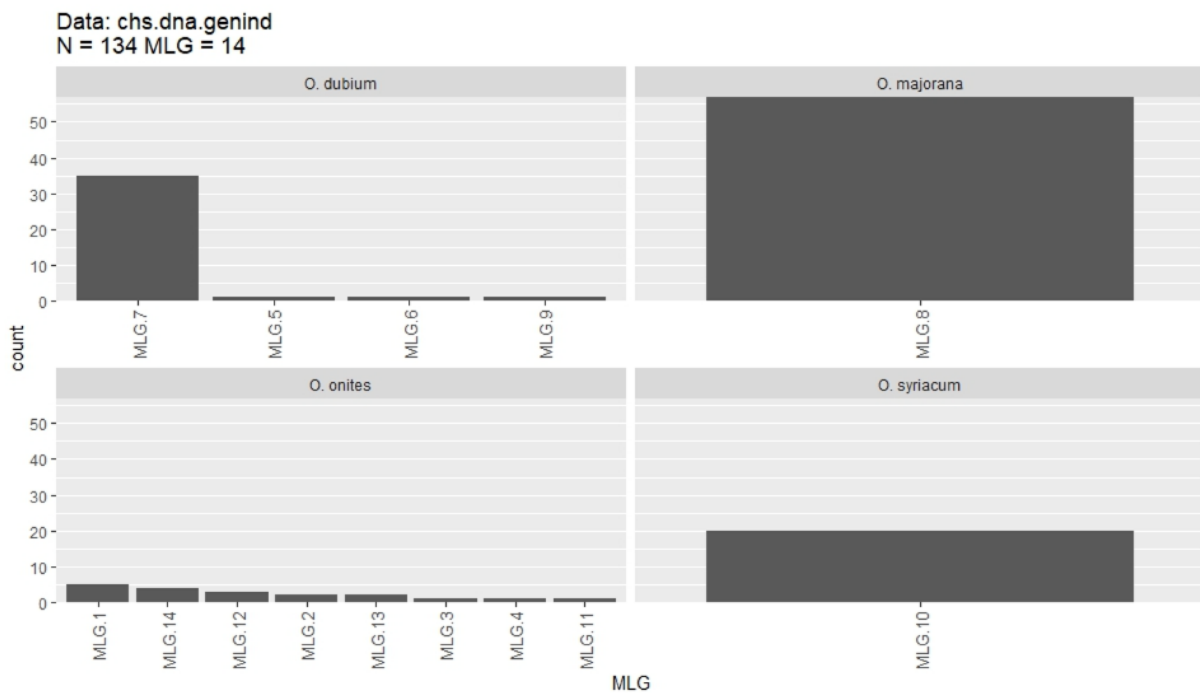


Abbildung 5: MLG-Muster-Verteilung innerhalb der *Origanum*-Populationen

Weiterhin wurde die Verknüpfung der Allele im Hinblick auf Linkage Disequilibrium (LD) graphisch dargestellt (Abbildung 6). Auf den x- und y-Achsen sind die jeweiligen Basenpaar-Positionen aufgetragen, während die wahrscheinlichen LD-Werte durch Farbcodes repräsentiert werden. Hellrote bis dunkelrote Felder zeigen eine starke LD an, was darauf hindeutet, dass die Allele nicht unabhängig voneinander vererbt werden, sondern in Kombination weitergegeben werden. Blaue Felder bestätigen hingegen die Nullhypothese, die besagt, dass keine LD vorliegt und dass die Allele an verschiedenen Loci unabhängig voneinander vererbt werden können (Livia M. de Souza, 2018).

Interessanterweise treten die stärksten Kopplungen, erkennbar an den intensiv roten Feldern, überwiegend im Intronbereich auf. Insgesamt sind deutlich mehr lila bis stark rote Kästchen zu erkennen, was auf eine signifikante genetische Kopplung zwischen den Loci hinweist. Blaue Felder, die keine Kopplung anzeigen, sind hingegen selten und treten lediglich an drei Stellen auf: Position 242 ist mit 559 nicht gekoppelt, Position 559 ist sowohl mit 1227 als auch mit 933 nicht gekoppelt. Die Grafik verdeutlicht somit das Vorhandensein starker genetischer Kopplung der Loci.

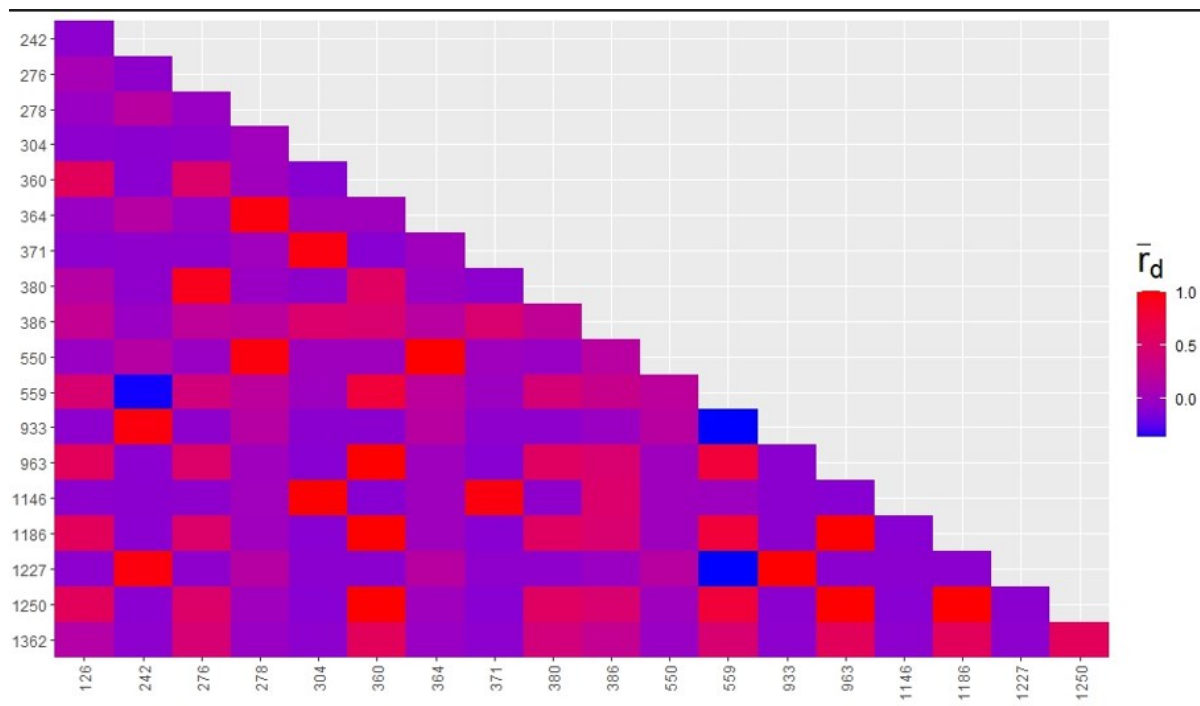


Abbildung 6: Graphische Darstellung von Linkage Disequilibrium (r_d) hinsichtlich der Verknüpfung der Marker (blau... keine Kopplung; lila... moderate Kopplung; hellrot... starke Kopplung; rot... sehr signifikante Kopplung)

Eine höhere genetische Kopplung könnte auch im Zusammenhang mit einem Selektionsdruck stehen. Dabei wurde der Tajima's D-Test mithilfe der Software MEGA durchgeführt. Die Neutralitätshypothese besagt, dass alle beobachteten genetischen Variationen in einer Population ausschließlich durch neutrale Mutationen entstanden sind, ohne Einfluss von Selektion, Populationsgröße oder anderen evolutionären Kräften ($D \approx 0$) (Schmidt, 2017). Ein negativer Tajima's D-Wert ($D < 0$) deutet darauf hin, dass mehr seltene Allele als erwartet vorkommen, was auf eine mögliche Selektion hinweist. Ein positiver Wert ($D > 0$) hingegen zeigt, dass weniger seltene Allele als erwartet vorhanden sind, was auf ein balancierendes Gleichgewicht oder einen Selektionsdruck auf bestimmte vorteilhafte oder nachteilige Polymorphismen hindeutet (Tajima, 1989). Je höher der Selektionsdruck, desto geringer ist die genetische Variation (Schmidt, 2017). In dieser Analyse lag der Fokus auf dem Tajima's D-Wert, der mit 1,627309 berechnet wurde. Dies lässt auf das Vorhandensein eines Selektionsdrucks innerhalb der Populationen schließen.

Zur Bewertung der Zuverlässigkeit der Abstände in den phylogenetischen Bäumen wurde die Bootstrap-Methode mit 100 Wiederholungen basierend auf dem Jukes-Cantor-Modell angewendet (Abbildung 7). Für Verzweigungen und eingezeichneten Distanzlängen mit Bootstrap-Werten unter 50 % besteht eine geringe Verlässlichkeit, weshalb diese Werte in der Abbildung nicht dargestellt wurden. Besonders entscheidend sind die Längen der Distanzlinien im Baum: Je länger die Linie, desto höher die genetische Variabilität.

Die Ergebnisse zeigen deutlich, dass *O. majorana* und *O. dubium* aus einer gemeinsamen Verzweigung hervorgehen, während *O. onites* in mehrere Subgruppen unterteilt ist. *O. syriacum* bildet hingegen eine eigene Verzweigung ohne Unterteilung in Subgruppen.

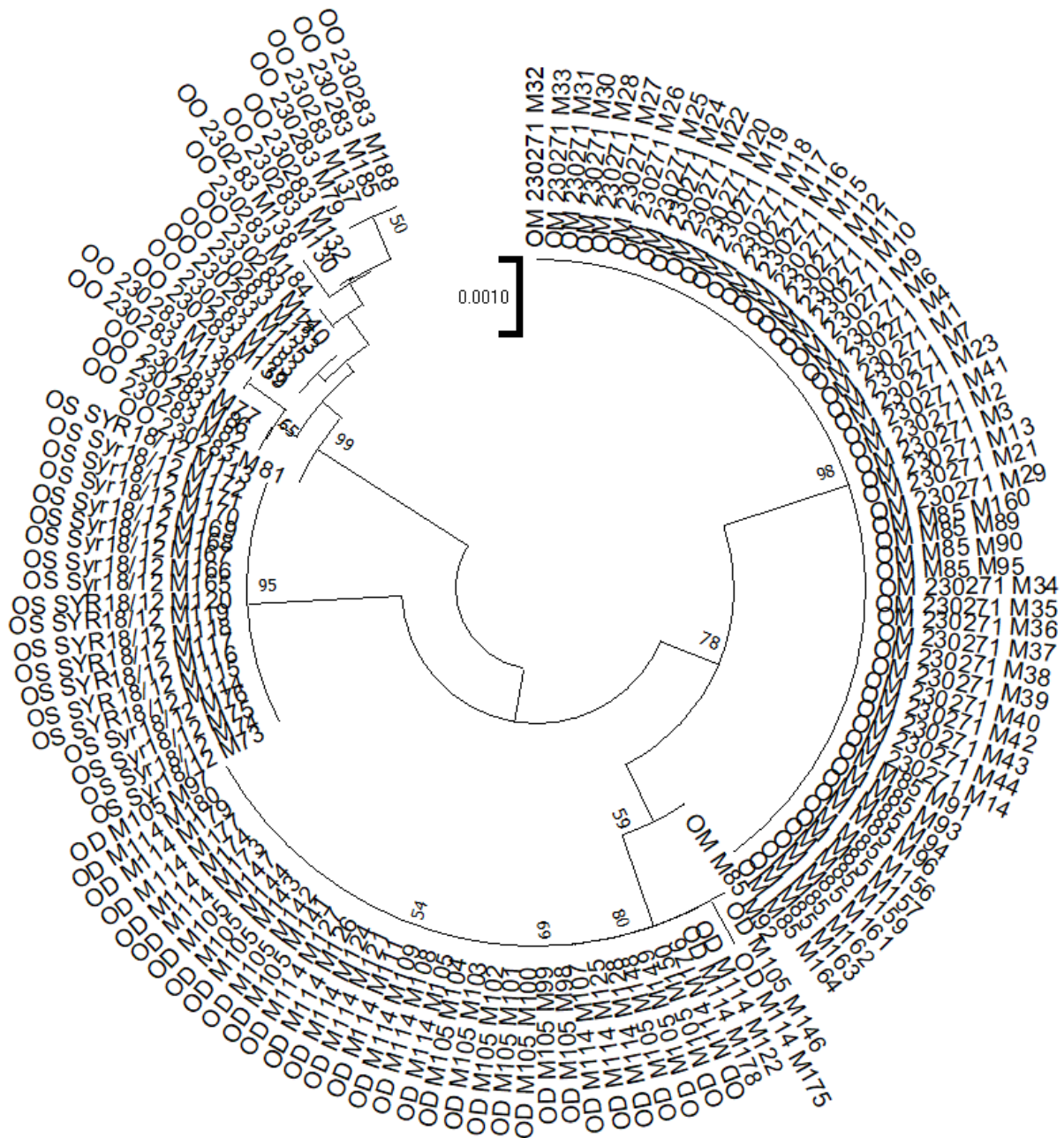


Abbildung 7: Der phylogenetische Baum wurde mithilfe der Neighbor-Joining-Methode abgeleitet und anhand der Bootstrap-Methode auf Basis von 100 Wiederholungen kalkuliert. Die Analyse umfasste 134 Klonsequenzen, mit insgesamt 1565 Basenpaarpositionen. Die evolutionären Abstände wurden mit dem Jukes-Cantor-Modell berechnet. Die Zahlen geben die Zuverlässigkeit der ausgewerteten Distanzen in % an. (OO... *Origanum onites*, OS... *Origanum syriacum*; OM... *Origanum majorana*; OD... *Origanum dubium*)

3.1.2 Genanalyse auf mRNA-Ebene

In diesem Kapitel werden die Kenngrößen sowie Methoden zur Analyse der genetischen Variabilität nicht weiter im Detail beschrieben, entsprechende Definitionen können in Kapitel [3.1.1 Genanalyse auf DNA-Ebene](#) nachgelesen werden.

Für die Genexpressionsanalyse wurden die Exons der DNA-Sequenzen mithilfe der Software analysiert. Die resultierende mRNA-Sequenz hatte bei allen Proben eine Länge von 1166 Basenpaaren. Die Nukleotide zeigten eine homogene Verteilung: 18 % U (T), 32 % C, 20 % A und 30 % G (Anhang Tab. 2). Von den 1166 Basenpaaren waren 1164 identisch, während acht polymorphe Stellen identifiziert wurden. Diese wiesen jeweils vier Transitionen (si) und vier Transversionen (sv) auf, was zu einem si/sv-Bias-Wert von $R = 1,0$ führte. Es wurden pro 100 Basenpaare 0,61 SNPs beobachtet.

Im weiteren Schritt wurde das maximale Likelihood-Substitutionsmodell bestimmt, wobei das Tamura 92 (T92) -Modell, als das am besten geeignete Modell identifiziert wurde (Anhang Tab. 3). Dieses Modell berücksichtigt Unterschiede in den Substitutionsraten zwischen Transitionen und Transversionen sowie in den Basenhäufigkeiten und ermöglicht so genauere Schätzungen der evolutionären Distanzen zwischen den Sequenzen (Tabelle 6) (Arenas, 2015). Transitionen waren häufiger zu finden als Transversionen, vor allem bei den Basen $A \leftrightarrow G$ und $T/U \leftrightarrow C$. Die häufigste Transition ist zwischen Guanin und Cytosin, mit 18.10 %. Transversionen wurden zwischen $A \leftrightarrow T/U$; $A \leftrightarrow C$; $T/U \leftrightarrow G$ und $C \leftrightarrow G$ beobachtet.

Tabelle 6: Wahrscheinlichkeit einer Substitution von einer Base zu einer anderen Base, berechnet nach dem Tamura-(92)-Modell. Die Analyse umfasste 134 Klonsequenzen, mit insgesamt 1166 Basenpaarpositionen. Die Zahlen geben die Wahrscheinlichkeiten in % an. Transitionen sind fett und Transversionen kursiv angezeigt. (A ... Adenin; T/U ... Thymin/Uracil; C ... Cytosin; G ... Guanin)

	A	T/U	C	G
A	-	<i>3.93</i>	<i>6.73</i>	18.10
T/U	<i>3.93</i>	-	18.10	<i>6.73</i>
C	<i>3.93</i>	10.57	-	<i>6.73</i>
G	10.57	<i>3.93</i>	<i>6.73</i>	-

Zur Bewertung der Zuverlässigkeit der phylogenetischen Distanzen wurde die Bootstrap-Methode mit 100 Wiederholungen nach dem Tamura-Modell angewendet. Innerhalb der Gruppen zeigte nur *O. onites* eine signifikante Distanz (d) von 0,000909 mit einer Standardabweichung von $\pm 0,000714$.

Die Analyse der Distanzen zwischen den Gruppen ergab, dass die größte Variabilität zwischen *O. dubium* und *O. onites* mit einer Distanz (d) von 0,005119 auftrat (Tabelle 7). Diese Werte weisen auf genetische Unterschiede zwischen diesen beiden Gruppen hin. Der niedrigste Wert liegt zwischen den Gruppen *O. majorana* und *O. syriacum* vor, was auf eine sehr geringe genetische Variation schließen lässt.

Tabelle 7: Darstellung der Distanzen zwischen den Arten, wobei die Anzahl der Änderungen betrachtet wurde.

	1	2	3	4
1. OO		0,001564	0,001348	0,001736
2. OS	0,004258		0,000847	0,001446
3. OM	0,003394	0,000858		0,001151
4. OD	0,005119	0,002577	0,001718	

Die Analyse der genetischen Vielfalt innerhalb der Populationen ergab eine Diversität von 0,000226 mit einer Standardabweichung von $\pm 0,000158$. Diese Ergebnisse zeigen eine insgesamt niedrige, aber dennoch vorhandene genetische Diversität innerhalb der untersuchten Populationen.

Nach der Transkription und der Entfernung des Introns blieben von den ursprünglich 18 identifizierten SNPs nur noch acht SNPs übrig (Tabelle 8). Die H_{exp} -Werte liegen zwischen 0.13 und 0.41, was auf eine moderate genetische Vielfalt hinweist. Die Werte für die Evenness sind überwiegend im mittleren Bereich anzufinden, mit Ausnahme der Positionen 534 und 828 mit dem Wert 0.84.

Tabelle 8: Darstellung der SNPs bezogen auf die genetische Vielfalt (H_{exp} ... erwartete Heterozygotie; Evenness ... Wahrscheinlichkeit einer Gleichverteilung der Allele)

mRNA locus	Hexp	Evenness
126	0.13	0.51
534	0.41	0.84
564	0.25	0.64
747	0.26	0.65
787	0.25	0.64
828	0.41	0.84
851	0.25	0.64
963	0.13	0.51

Bei *O. majorana* und *O. syriacum* wurde weiterhin jeweils nur ein MLG gefunden, MLG.5 und MLG.4 (Abbildung 8). Die MLG-Anzahl von den DNA- bis zu den mRNA-Sequenzen hatte sich bei *O. dubium* von vier auf ein MLG (MLG.3) reduziert. Somit zeigte sich, dass von den vier beobachteten MLG-Mustern nur eines in den Exon-Regionen vorhanden war, während die übrigen Variationen im Intron lokalisiert waren. *O. onites* zeigte die höchste Variabilität, was sich in vier unterschiedlichen MLG-Mustern widerspiegelte: MLG.1, MLG.2, MLG.6 und MLG.7. Aufgrund der hohen H- und λ -Werte wurde von einer signifikanten genetischen Diversität innerhalb dieser Population ausgegangen (Tabelle 9). Der hohe E.5-Wert deutete auf eine etwa gleichmäßige Verteilung der MLGs hin. Der rbarD-Wert ist leicht erhöht, was auf eine leichte genetische Kopplung zwischen den Loci hinweist. Zudem ist anzumerken, dass *O. onites* die Population mit dem kleinsten Datensatz und dennoch mit der höchsten Variabilität war.

Tabelle 9: Darstellung der genetischen Variabilität innerhalb der untersuchten *Origanum*-Populationen (N ... Anzahl der Klone; MLG ... Multilocus-Genotypen; H ... Diversitätsindex zwischen N und MLG; λ ... Maß für die Diversität, bezogen auf die MLG-Muster-Verteilung; E.5... spezifischer Evenness-Index; Hexp... erwartete Heterozygotie; rbarD... standardisierte Linkage Disequilibrium; --- keine Verteilung einer Variation bestimmbar)

Population	N	MLG	H	λ	E.5	Hexp	rbarD
<i>O. dubium</i>	38	1	0.000	0.000	---	0.000	---
<i>O. majorana</i>	57	1	0.000	0.000	---	0.000	---
<i>O. onites</i>	19	4	1.260	0.693	0.887	0.132	0.179
<i>O. syriacum</i>	20	1	0.000	0.000	---	0.000	---

Die Analyse ergab, dass in der Population *O. onites* MLG.1 und MLG.7 etwa dreimal häufiger vorkommen als MLG.6 und MLG.2. Alle anderen Populationen zeigten eindeutig einen individuellen MLG.

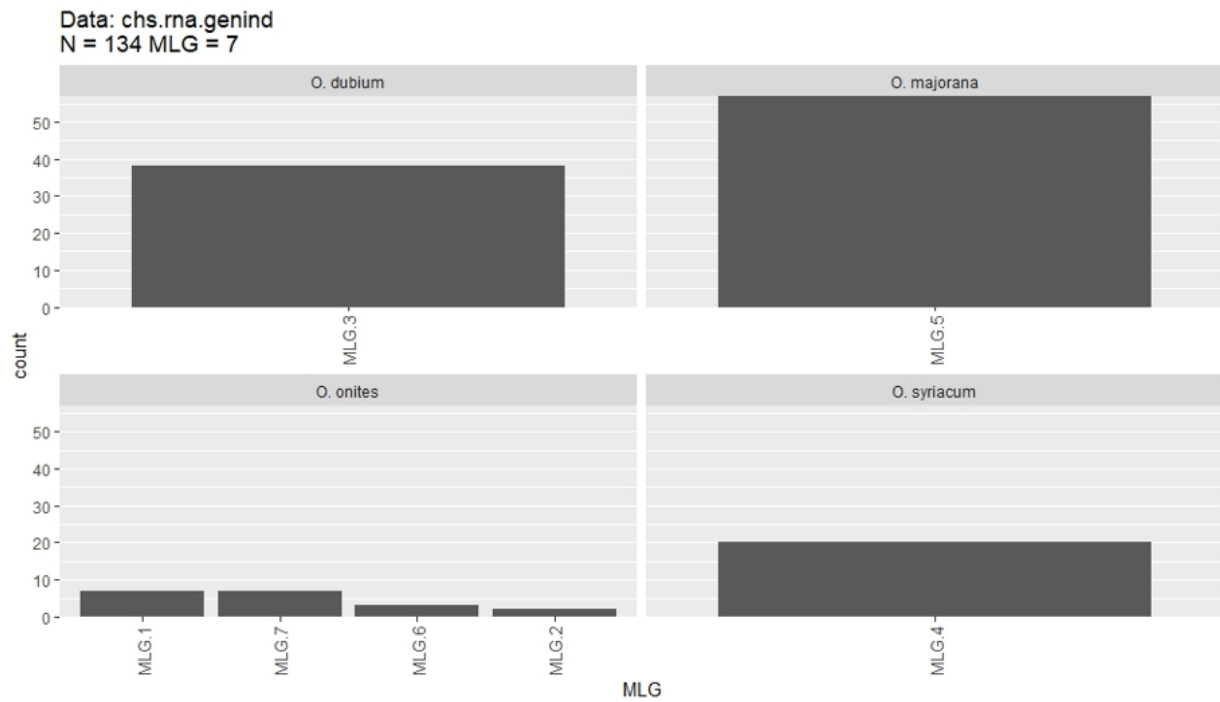


Abbildung 8: MLG-Muster-Verteilung innerhalb der *Origanum*-Populationen

Die Vielzahl roter Felder ist weg, was darauf hindeutet, dass die starke genetische Kopplung im Intron lokalisiert war (Abbildung 9). Interessanterweise wurden in diesem Abbild mehr blaue Felder in unmittelbarer Nähe zueinander festgestellt. Jedoch war ein gewisses Maß an genetischer Kopplung auch hier zu beobachten.

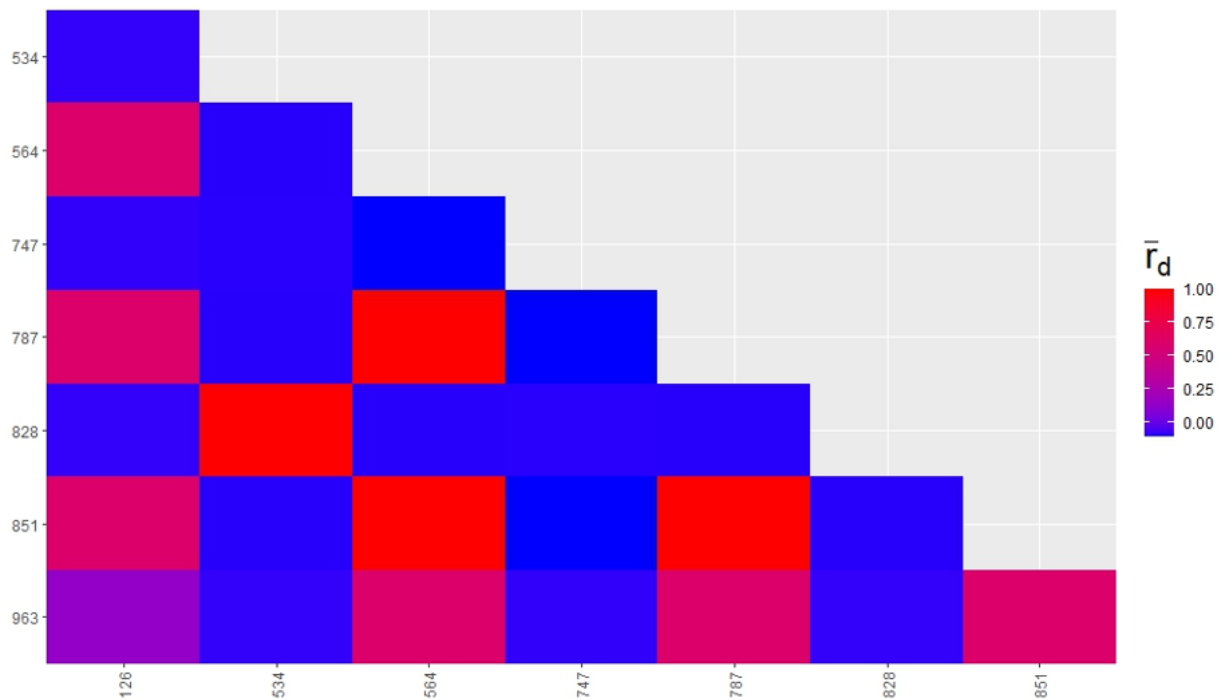


Abbildung 9: Graphische Darstellung von Linkage Disequilibrium (r_d) hinsichtlich der Verknüpfung der Marker (blau... keine Kopplung; lila... moderate Kopplung; hellrot... starke Kopplung; rot... sehr signifikante Kopplung)

Der Tajima's D-Wert wurde mit 0,954808003 berechnet, was auf das Vorliegen von Selektionsdruck hindeutet.

Ein phylogenetischer Baum wurde erstellt, um die evolutionären Beziehungen zwischen den Arten zu veranschaulichen (Abbildung 10). Die Arten *O. syriacum*, *O. majorana* und *O. dubium* gingen von einem gemeinsamen Vorfahren aus, während *O. onites* weiterhin eine separate Verzweigung aufwies. Interessanterweise hat sich die M92-Klonsequenz vom *O. majorana* (M85) bei den *O. dubium* eingeordnet.

Distanzwerte mit einer Zuverlässigkeit von unter 50 % wurden nicht angezeigt.

3.1.3 Analyse der abgeleiteten Aminosäuresequenz

Die mRNA-Sequenzen wurden mithilfe der Software MEGA in die 388 Aminosäuren (AS) umfassenden AS-Sequenzen übersetzt. Die Verteilung der AS ergab folgende Anteile: Alanin 8 %, Cystein 2 %, Asparaginsäure 5 %, Glutaminsäure 7 %, Phenylalanin 4 %, Glycin 9 %, Histidin 2 %, Isoleucin 5 %, Lysin 7 %, Leucin 10 %, Methionin 3 %, Asparagin 2 %, Prolin 5 % und Glutamin 3 %, Arginin 4 %, Serin 7 %, Threonin 5 %, Valin 9 %, Tryptophan 1 % und Tyrosin 2 %. Leucin war die am häufigsten vertretene Aminosäure, während Tryptophan am seltensten vorkam. Auffällig war, dass *O. onites* einen etwas höheren Anteil an Threonin und einen leicht reduzierten Anteil an Lysin aufwies.

An Position 284 der Sequenz wurde ein AS-Austausch identifiziert, der zum Wechsel von Threonin (T) zu Lysin (K) führte. Bis auf *O. onites* wiesen alle anderen drei Arten diesen einen Polymorphismus sehr wohl auf.

Aufgrund der begrenzten Anzahl von Variation auf AS-Ebene wurde auf eine weiterführende Analyse mit der Software R verzichtet, da deren Relevanz für die Darstellung dieser Ergebnisse als gering eingeschätzt wurde.

Zur Illustration der Polymorphismus-Position wurden zwei Klone ausgewählt, deren Sequenzen im Detail analysiert und die Mutationsstellen an Position 284 markiert wurden (Abbildung 11, Abbildung 12).

```
MVTVEEIRRAQRAEGPATVLAIGTAVPSNCVDQSEYPDYYFRITNSEHKTELKEKFMRMCEKSMIKKRYMHLTEEYLKENPNITAY
MAPSLDARQDIVVVEVPKLGKEAAQKAIKEWGQPKSKITHLVFCTTSGVDMPGADYQLTKLLGLRPSVKRFMMYQQGCFAGG
TVLRMAKDLAENNAGARVLVVCSEITAVTFRGPSDSLHDSLVGQALFGDGAAAVIVGSDPVVGVERPLFELVSAAQTILPDSGD
AIDGHLREVGLTFHLLKDVPLISKNIEKSLKAEAFAPMGISDWNSVFWIAHPGGPAILDQVELKLGLKPEKLRSTRHVLSEYGNM
SSACVLFILDEMRSKSAKEGMSSTGEGLDWGVLFGFGPGLTVETVVLHLSVA
```

Abbildung 11: Aminosäuren-Sequenz Klon M95 aus M85 (*O. majorana*) – markierte Stelle: 284 K

```
MVTVEEIRRAQRAEGPATVLAIGTAVPSNCVDQSEYPDYYFRITNSEHKTELKEKFMRMCEKSMIKKRYMHLTEEYLKENPNITAY
MAPSLDARQDIVVVEVPKLGKEAAQKAIKEWGQPKSKITHLVFCTTSGVDMPGADYQLTKLLGLRPSVKRFMMYQQGCFAGG
TVLRMAKDLAENNAGARVLVVCSEITAVTFRGPSDSLHDSLVGQALFGDGAAAVIVGSDPVVGVERPLFELVSAAQTILPDSGD
AIDGHLREVGLTFHLLKDVPLISKNIEKSLTAEAFAPMGISDWNSVFWIAHPGGPAILDQVELKLGLKPEKLRSTRHVLSEYGNM
SSACVLFILDEMRSKSAKEGMSSTGEGLDWGVLFGFGPGLTVETVVLHLSVA
```

Abbildung 12: Aminosäuren-Sequenz Klon M79 aus 230283 (*O. onites*) – markierte Stelle 284 T

Für die Analyse der Proteinsequenzen wurde das Whelan and Goldman (WAG)-Modell verwendet, da dieses als das am besten geeignete Substitutionsmodell anerkannt wurde. Dieses Modell wird häufig in der Phylogenetik eingesetzt. Es beschreibt die Substitutionsraten zwischen verschiedenen Aminosäuren und ermöglicht dadurch eine präzisere Darstellung der evolutionären Beziehungen zwischen Organismen. Zudem berücksichtigt das Modell die unterschiedlichen Wahrscheinlichkeiten, mit der verschiedene Aminosäuren durch Substitution ersetzt werden, sowie deren Verteilung innerhalb der Sequenz (Anhang Tab. 4) (Simon Whelan, 2001).

Die innerhalb der Arten ermittelte genetische Distanz betrug 0,000633 mit einer Standardabweichung von $\pm 0,000733$, was auf eine äußerst geringe Variabilität der Aminosäuresequenzen hindeutet.

Zwischen den Arten zeigte sich jedoch eine signifikante Distanz, insbesondere bei *O. onites*, der einen eigenen phylogenetischen Zweig bildet, mit einer Distanz von 0,00258065 (Tabelle 10). Die anderen Arten zeigten hingegen keine signifikanten Distanzen zueinander. Diese Beobachtung könnte mit dem zuvor erwähnten AS-Wechsel an Position 284 in Zusammenhang stehen.

Tabelle 10: Darstellung der Distanzen zwischen den Arten, wobei die Anzahl der Änderungen betrachtet wurde. Die Anzahl der AS-Substitutionen pro Position errechnet über das Mittel aller Sequenzpaare zwischen den Arten.

	1	2	3
1. OO			
2. OS	0,00258065		
3. OM	0,00258065	0,0000000	
4. OD	0,00258065	0,0000000	0,0000000

Zusätzlich wurde der Tajima's D-Wert berechnet, wobei der D-Wert 0,370236 betrug, was auf das Vorliegen eines Selektionsdrucks hindeutet.

Die phylogenetische Analyse ergab einen radialen Baum mit einer Zuverlässigkeit von 73 %, der zwei Hauptverzweigungen aufwies (Abbildung 13). *O. onites* bildete dabei eine eigene Hauptverzweigung, was auf den AS-Austausch an Position 284 zurückzuführen war. Interessanterweise gruppierte sich der Klon M92 aus *O. majorana* (M85) ebenfalls innerhalb der *O. dubium*-Sequenzen ein, wie es zuvor schon bemerkbar war.

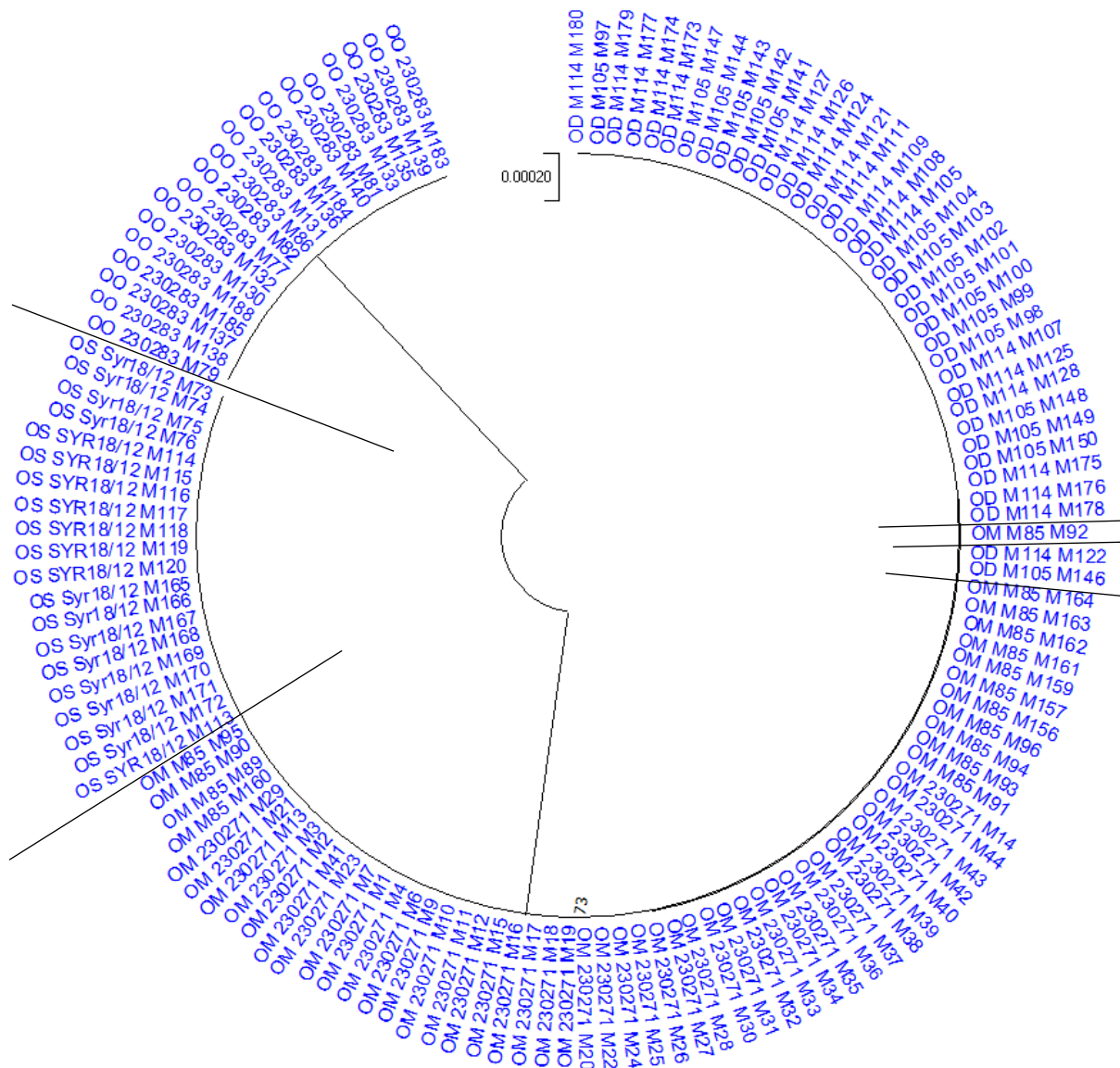


Abbildung 13: Der phylogenetische Baum wurde mithilfe der JTT-Methode abgeleitet und anhand der Bootstrap-Methode auf Basis von 100 Wiederholungen kalkuliert, mit einer Zuverlässigkeit von 73 %. Die Analyse umfasste 134 AA-Sequenzen, mit insgesamt jeweils 388 Positionen. Die evolutionären Abstände wurden mit der MLM-Methode (ausgehend vom höchsten log likelihood -1114,49) und des Whelan and Goldman-Modells berechnet. (OO ... *Origanum onites*, OS ... *Origanum syriacum*; OM ... *Origanum majorana*; OD ... *Origanum dubium*)

3.2 Homologie Modellierung und Voraussage der 3D-Struktur

Vorherige Studien haben gezeigt, dass die konservierten Stellen der Chalkonsynthase folgende Aminosäuren betreffen: Cys 164, His 303, Asn 336 sowie zur Unterstützung der Polyketidkette Phe 215 und Phe 265 (Jean-Luc Ferrer, 1999). Zur Darstellung der 3D-Struktur unseres extrahierten Gens wurde die Klonsequenz M95 von *O. majorana* (M85), die den AS-Austausch 284 (TK) enthält, genauer analysiert, (Abbildung 11).

Relevante Parameter dieser CHS-Sequenz konnten mithilfe der Expasy-Software ermittelt werden. Das Protein besteht aus 388 Aminosäuren, mit einem Molekulargewicht von 42.345,84 Da und einem isoelektrischen Punkt (pI) von 6,18. Die am häufigsten vorkommende Aminosäure ist Leucin (9,8 %), während Tryptophan mit 1,0 % am seltensten vertreten ist. Die Anzahl der negativ geladenen Aminosäuren (Aspartat, Glutamat) beträgt 47, wohingegen die positiv geladenen Aminosäuren (Arginin, Lysin) mit 43 vertreten sind.

Die berechnete Summenformel dieses CHS-Gens lautet $C_{1885} H_{3011} N_{507} O_{560} S_{19}$, was einer Gesamtzahl von 5982 Atomen entspricht. Die Halbwertszeit des Proteins wurde auf 30 Stunden in Säugetier-Retikulozyten *in vitro*, über 20 Stunden in Hefe *in vivo* und über 10 Stunden in *Escherichia coli in vivo* geschätzt.

Der Instabilitätsindex (II) des Proteins wurde ebenfalls berechnet. Ein Wert unter 40 deutet auf ein stabiles Protein hin, während ein Wert über 40 auf Instabilität hindeutet. Für das Protein wurde ein Instabilitätsindex von 42,67 ermittelt, was auf eine (geringe) potenzielle Instabilität in wässriger Lösung hinweist. Der aliphatische Index beträgt 90,70, und der Grand Average of Hydropathicity (GRAVY) wurde mit -0,071 berechnet. Ein negativer GRAVY-Wert deutet auf ein hydrophiles Protein hin, wobei die Chalkonsynthase hier nur eine schwach hydrophile Eigenschaft aufweist (Kunchur Guruprasad, 1990).

Mithilfe der Swiss-Modellierungs-Funktion von Expasy konnten wir die Struktur detailliert betrachten und die konservierten Stellen (Cys 164, His 303, Asn 336, Phe 215 und Phe 265) auf Polymorphismen untersuchen.

Das nächstverwandte CHS-Monomer-Template stammt vom Organismus *Fagus sylvatica* (A0A2N9F4H0.1.A Chalcone synthase; AlphaFold DB model of A0A2N9F4H0_FAGSY) ab, mit einer Sequenzidentität von 90,46 % (Abbildung 14, Abbildung 15, Abbildung 16).

Besondere Aufmerksamkeit gilt den rot-markierten Positionen, die zum einen die AS- Reste umfassen, die für die Funktion des Enzyms von zentraler Bedeutung sind: Cystein 164, Histidin 303, Asparagin 336, Phenylalanin 215 und Phenylalanin 265 (Jean-Luc Ferrer, 1999). Zum anderen ist die N-terminale Domäne durch Methionin 137 gekennzeichnet, welche eine strukturell wichtige Rolle spielt.

Dieses Alignment verdeutlicht die evolutionäre Konservierung essentieller AS-Reste zwischen den beiden Arten und unterstreicht die funktionelle Bedeutung dieser konservierten Regionen innerhalb der Chalkonsynthase.



Abbildung 14: Das Model-Template Alignment der Klonsequenz M95 aus M85, *O. majorana* (blaue Sequenz) wurde mit dem CHS-Monomer-Template aus *Fagus sylvatica* (weiße Sequenz) verglichen. Identische Sequenzbereiche zwischen den beiden Proteinen sind durch Umrandungen hervorgehoben. Rot-markierte Positionen sind zum einen die konservierten Stellen: Cys 164, His 303, Asn 336, Phe 215 und Phe 265 und zum anderen die N-terminale Domäne: Met 137. (Pfeile ... Helices, Umkreisungen ... Faltblätter) (Swiss-Model Expaty)



Abbildung 15: Voraussage der 3D-Struktur anhand des Model-Template-Alignments (für die Funktion wichtige konservierte Aminosäurereste (Jean-Luc Ferrer, 1999) und die N-terminale Domäne wurden in der 3D-Struktur als Strukturformeln sichtbar gemacht, um sie herauszuheben) (Swiss-Model Expasy)

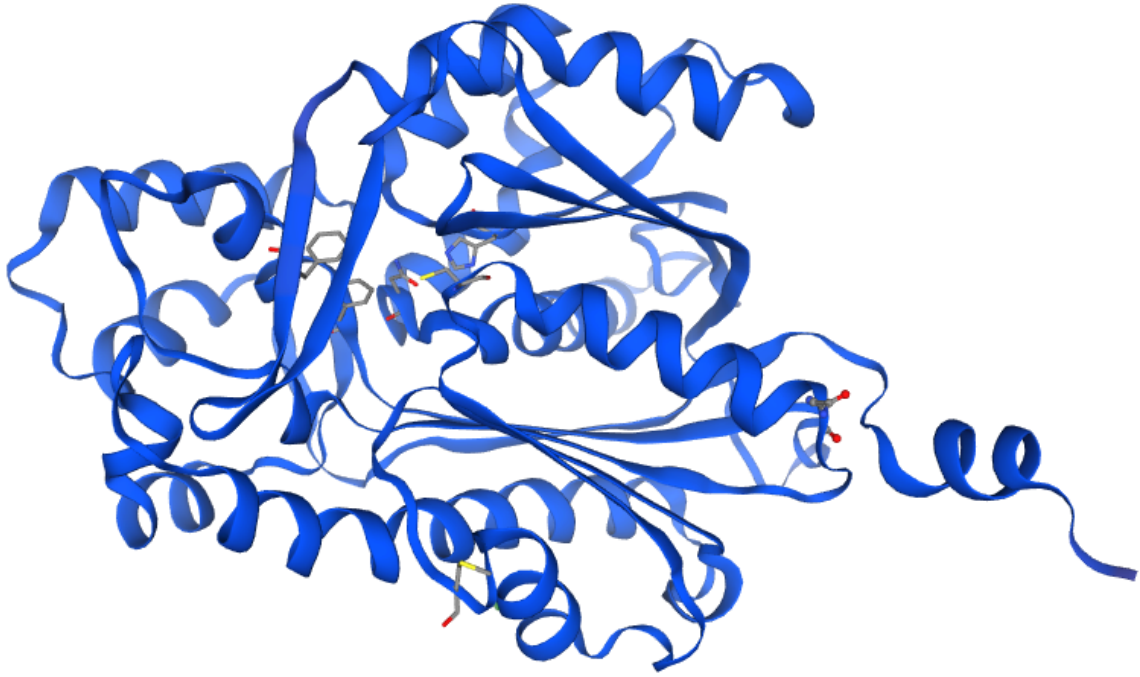


Abbildung 16: CHS 3D-Struktur (aus einem anderen Blickwinkel) (Swiss-Model Expasy)

Die fünf wichtigen Positionen wurden einzeln einer detaillierten Analyse unterzogen (Abbildung 17). Die Ergebnisse dieser Untersuchung bestätigten eindeutig, dass alle konservierten Reste in der untersuchten Gensequenz vorhanden und unverändert sind. Diese Konservierung untermauert die funktionelle Integrität des Gens und deutet darauf hin, dass die kritischen Aminosäurereste, die für die enzymatische Aktivität entscheidend sind, in ihrer ursprünglichen Form erhalten geblieben sind.

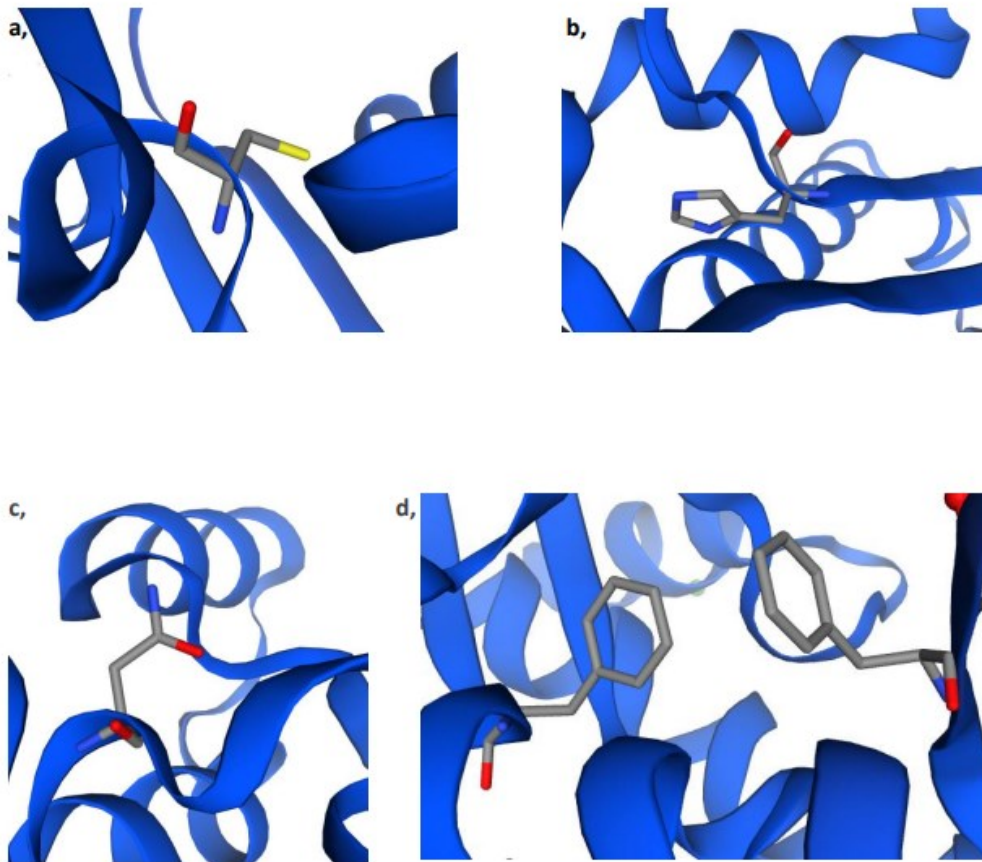


Abbildung 17: Konservierte Positionen des CHS-Gens – a Cys 164: entscheidend für die Aktivierung und Bindung des Coumaryl-CoA; b His 303: katalysiert die Aktivierungsreaktion von Cys; c Asn 336: spielt eine Rolle in der Decarboxylierungs-Reaktion von Malonyl-CoA und stabilisiert das Reaktionsintermediat; d Phe 215 und Phe 265: Unterstützung des Prozesses (Jean-Luc Ferrer, 1999) (Swiss-Model Expasy)

4. Diskussion

Ziel dieser Studie war es, die genetische Variabilität des Chalkonsynthase-Gens (CHS) zwischen und innerhalb von vier verschiedenen Arten der Gattung *Origanum* zu untersuchen: *Origanum majorana*, *Origanum syriacum*, *Origanum dubium* und *Origanum onites*. Die CHS ist das erste Schlüsselenzym in der Flavonoid-Biosynthese, die insbesondere unter Stressbedingungen verstärkt reguliert wird. Es ist bekannt, dass die meisten Pflanzen mindestens ein CHS-Gen besitzen (T. T. H. Dao, 2011), was die zentrale Rolle dieses Gens in der pflanzlichen Evolution und Anpassung verdeutlicht.

Die Identität der gewonnenen Sequenzen wurde durch NCBI-BLAST bestätigt und mit bereits in Datenbanken hinterlegten CHS-Sequenzen abgeglichen. Insgesamt konnten 134 Klonsequenzen erfolgreich generiert und analysiert werden. Diese Sequenzanalysen liefern wertvolle Erkenntnisse zur genetischen Variabilität und zu den evolutionären Veränderungen des CHS-Gens innerhalb der *Origanum*-Arten.

Die Sequenzierung des CHS-Gens war erfolgreich. Für die kultivierte Probe von *O. majorana* konnte ein Primer aus einer früheren Studie über Salbei (S1_Sm1_1F) verwendet werden, wohingegen für die Analyse aller anderen Proben neue Primer entwickelt und getestet werden mussten (Tabelle 2).

Das Gen selbst wies eine kodierende Sequenz von 1167 bp auf, die durch ein Intron in ein kürzeres Exon 1 und ein längeres Exon 2 unterteilt wurde. Die Gesamtlänge des Gens variierte zwischen 1551 und 1565 Basenpaaren, abhängig von den im Intron lokalisierten Indels. Somit konnte auch die bereits vorhandene These, dass fast alle CHS-Gene nur ein Intron besitzen, bestätigt werden (Shahzad A. Pandith, 2020). Diese Ergebnisse ähneln stark den ermittelten Werten von *Salvia miltiorrhiza* in Kapitel [1.2.2 Chalkonsynthase \(CHS\)](#).

SNPs

In den *Origanum*-Proben wurden 18 SNPs auf DNA-Ebene identifiziert (1,15 SNPs pro 100 bp), acht SNPs auf mRNA-Ebene (0,61 SNPs pro 100 bp) sowie ein einzelner Aminosäureaustausch an Position 284 von Threonin (T) zu Lysin (K). Diese Beobachtungen stützen die Hypothese, dass Polymorphismen überwiegend in nicht-kodierenden Regionen auftreten, während

kodierende Bereiche stark konserviert bleiben. Dies deutet auf einen möglichen Selektionsdruck auf das CHS-Gen hin, da abweichende Sequenzen in kodierenden Regionen vermieden werden (Melanie Novak, 2023). Die häufigsten SNP-Muster umfassten Transitionen von C zu T sowie C zu A. Im Vergleich zu einer ähnlichen Studie an *Matricaria chamomilla* wurden in dieser Art deutlich mehr SNPs (und Haplotypen) festgestellt.

MLGs

Die 134 analysierten Klonsequenzen gruppieren sich in 14 MLGs (Haplotypen). Die Anzahl der MLGs variierte zwischen den Arten: *O. majorana* und *O. syriacum* wiesen auf DNA-Ebene nur ein MLG auf, *O. dubium* vier und *O. onites* zeigte mit acht MLGs die größte genetische Diversität. Die Anzahl der MLGs nahm erwartungsgemäß von der DNA- zur AA-Sequenz deutlich ab, was auf eine niedrige genetische Variation in den kodierenden Regionen hinweist.

Genetische Kopplung

Eine moderate Linkage Disequilibrium (LD) wurde auf DNA-Ebene festgestellt, wobei die Kopplungen über das gesamte Gen verteilt waren und nicht, wie typischerweise erwartet, angehäuft auftraten (Abbildung 6, Abbildung 9). Auf mRNA-Ebene nahm das LD signifikant ab, was auf eine unabhängige Weitergabe der Allele hindeutet, möglicherweise bedingt durch Rekombination während des Crossing-Over. Dies könnte ein erster Hinweis auf das Vorhandensein von Genkopien im Genom sein (Schmidt, 2017).

Selektionsdruck

Die Ergebnisse des Tajima's D-Tests und der LD-Analyse unterstützen die Annahme eines starken Selektionsdrucks auf das CHS-Gen. CHS-Gene gehören bekanntlich zu den upstream-Genen, die einem höheren Selektionsdruck unterliegen (Yonekura-Sakakibara K, 2019). Ein D-Wert > 0 auf allen Ebenen lässt darauf schließen, dass das CHS-Gen stark zentral und konserviert ist und Polymorphismen stark selektiv gefiltert werden. Nur eine einzige nicht-synonyme Codon-Veränderung konnte bis zur Aminosäureebene nachverfolgt werden, während alle anderen Polymorphismen synonyme Veränderungen darstellten.

Phylogenetische Beziehungen

Die phylogenetischen Analysen in dieser Studie verdeutlichten die evolutionären Beziehungen zwischen den untersuchten *Origanum*-Arten (in einer zusammenfassenden Darstellung in Abbildung 18). Erwartungsgemäß nahmen die genetischen Distanzen sowie die Verzweigungen im phylogenetischen Baum ab, je näher man den Primärsequenzen kam. Alle vier untersuchten Arten, die der Sektion *Majorana* angehören, zeigten nur geringe Unterschiede in ihrer evolutionären Entwicklung, was ihre enge genetische Verwandtschaft bestätigt (Ietswaart, 1980). Diese Beobachtung steht im Einklang mit früheren Arbeiten und wird durch die vorliegende Untersuchung gestützt.

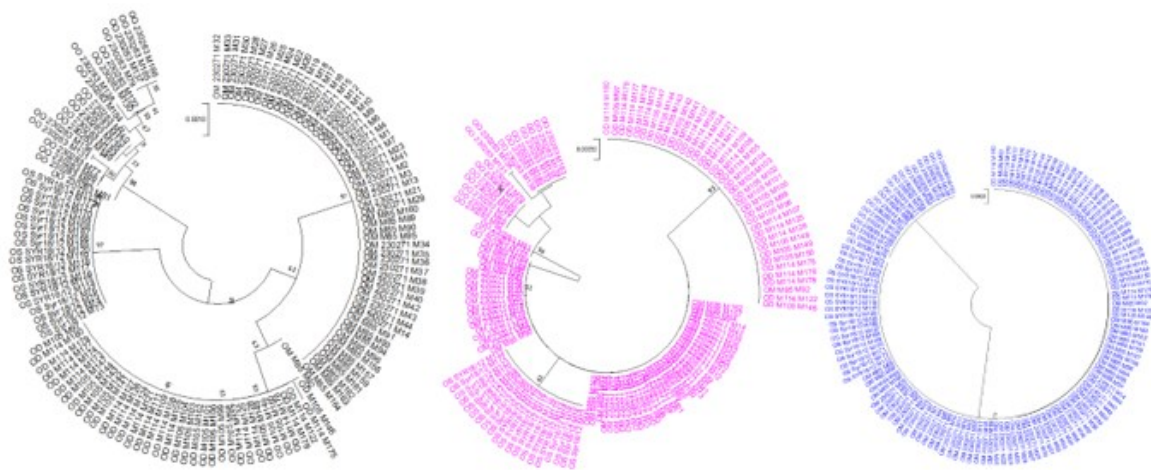


Abbildung 18: Phylogenetische Analyse von der DNA bis hin zur AS-Sequenz

Eine frühere Studie zu den Verwandtschaftsbeziehungen innerhalb der Gattung *Origanum* (Sektion *Majorana*) zeigte, dass sich alle Arten, mit Ausnahme von *O. dubium*, in einer eng zusammenhängenden Gruppe befanden (Brigitte Lukas, 2013). Interessanterweise weichen die Ergebnisse dieser Studie von dieser Darstellung ab. Hier wurde festgestellt, dass *O. dubium*, *O. majorana* und *O. syriacum* von einem gemeinsamen Vorfahren abstammen, während *O. onites* eine etwas größere genetische Distanz aufwies und eine eigenständige homogene Hauptverzweigung im phylogenetischen Baum bildete. Diese Beobachtung ist auf den Aminosäureaustausch an Position 284 zurückzuführen.

Der Grund für die Einordnung einer Klonsequenz (M92 aus M85) in den *O. dubium*-Zweig konnte in dieser Analyse nicht abschließend geklärt werden. Weitere Untersuchungen sind

erforderlich, um diese Diskrepanz in den phylogenetischen Mustern zu verstehen und die evolutionären Mechanismen, die diese Verzweigungen beeinflussen, zu beleuchten.

Auswirkung der SNPs auf die Funktion des Enzyms

Von allen identifizierten Polymorphismen führte nur eine einzelne SNP zu einem Aminosäureaustausch an Position 284, wobei Threonin durch Lysin ersetzt wurde (Abbildung 19). Demnach handelte es sich hierbei um eine nicht-synonyme Mutation (Schmidt, 2017). Dies wurde bei allen Arten außer *O. onites* beobachtet. Es ist bereits aus früheren Studien bekannt, dass *O. onites* morphologisch die am besten definierte und homogenste Art innerhalb der Gattung darstellt (Brigitte Lukas, 2013). Diese Beobachtung wird durch die vorliegenden Ergebnisse bestätigt.

Besonders bemerkenswert ist jedoch, dass *O. onites* sowohl auf DNA- als auch auf mRNA-Ebene die höchste genetische Variabilität unter den untersuchten Arten aufwies. Trotz dieser genetischen Diversität zeigte die Analyse der Aminosäuresequenzen, dass *O. onites* auf der Proteinebene die konservierteste Art war. Im Gegensatz zu den anderen untersuchten Arten, die an Position 284 einen Polymorphismus aufwiesen, blieb die Aminosäuresequenz bei *O. onites* unverändert. Dieses Ergebnis verdeutlicht, dass trotz hoher genetischer Variation auf DNA- und mRNA-Ebene der Selektionsdruck auf die kodierenden Regionen bei *O. onites* besonders stark ist, was zur Konservierung der Aminosäuresequenz führt.

Threonin ($C_4H_9NO_3$) besitzt ein kürzeres Kohlenstoffgerüst und zeigt polare sowie neutrale Eigenschaften, während Lysin ($C_6H_{14}N_2O_2$) eine längere Seitenkette mit zwei basischen Gruppen aufweist. Beim Vergleich des log P-Werts (Maß für die Hydrophilie) zeigte sich, dass in etwa beide gleich polar sind (Threonin: -2,94; Lysin: -3,05) (National Center for Biotechnology Information. PubChem Compound Summary for CID 6288, Threonine. <https://pubchem.ncbi.nlm.nih.gov/compound/Threonine>. Accessed Oct. 29, 2024.; National Center for Biotechnology Information. PubChem Compound Summary for CID 5962, Lysine. <https://pubchem.ncbi.nlm.nih.gov/compound/Lysine>. Accessed Oct. 29, 2024.).



Abbildung 19: Chemische Struktur von Threonin (T) und Lysin (K) (Pubchem; National Center for Biotechnology Information. PubChem Compound Summary for CID 6288, Threonine. <https://pubchem.ncbi.nlm.nih.gov/compound/Threonine>. Accessed Oct. 28, 2024; National Center for Biotechnology Information. PubChem Compound Summary for CID 5962, Lysine. <https://pubchem.ncbi.nlm.nih.gov/compound/Lysine>. Accessed Oct. 28, 2024)

Frühere Studien haben gezeigt, dass fünf Schlüsselaminosäuren (Cys 164, His 303, Asn 336, Phe 215 und Phe 265) für die Erhaltung der Enzymaktivität entscheidend sind (Jean-Luc Ferrer, 1999). Die Analyse der 3D-Struktur der M95-Klonsequenz mit dem Polymorphismus an Position 284 zeigte, dass dieser keine wesentliche konservierende Eigenschaft des CHS-Gens beeinträchtigt. Dennoch bleibt die vollständige funktionale Bedeutung dieser AS-Veränderung unklar und könnte zufällig aufgetreten sein.

Schlussfolgerung

Zusammenfassend lässt sich sagen, dass das CHS-Gen in den untersuchten *Origanum*-Arten stark konserviert ist und eine wesentliche zentrale Rolle in der Flavonoid-Biosynthese übernimmt. Die beobachteten Polymorphismen, mit Ausnahme der Position 284, hatten keinen Einfluss auf die kodierenden Regionen, was auf einen starken Selektionsdruck hinweist, der die genetische Integrität dieses Gens sicherstellt. Die Ergebnisse dieser Studie waren äußerst zufriedenstellend und bestätigten die aufgestellten Hypothesen.

O. onites stellt die homogenste und konservierteste Art dar – dies wurde bereits in einer früheren Studie festgestellt (Brigitte Lukas, 2013).

Ansätze für zukünftige Forschungsfragen

Diese Studie bietet eine solide Grundlage für weiterführende Forschungsansätze. Zukünftige Arbeiten könnten durch die Analyse einer größeren Stichprobe erweitert werden, wobei eine breitere Auswahl an kultivierten Proben einbezogen werden sollte, um die genetische Diversität in weiteren Bereichen des CHS-Gens detaillierter zu untersuchen.

Frühere genomweite Analysen haben bereits 26 Gene identifiziert, die in Zusammenhang mit der Flavonoid-Biosynthese stehen, von denen 20 neuartig sind (Yuxing Deng, 2018). Die Untersuchung der genetischen Variabilität dieser Gene könnte wertvolle Erkenntnisse über die Regulation der Biosynthese und deren Kontrolle durch Transkriptionsfaktoren liefern.

Darüber hinaus wäre eine Analyse der CHS-Expression unter verschiedenen Stressbedingungen von großem wissenschaftlichem Interesse, da dies neue Einsichten in die adaptive Regulation dieses Schlüsselenzyms bieten könnte. Ein weiterer vielversprechender Forschungsansatz bestünde im Vergleich der gewonnenen Daten mit Sequenzen anderer Pflanzenfamilien, um phylogenetische und funktionelle Gemeinsamkeiten sowie Unterschiede aufzuklären.

Die CHS-Aktivität korreliert stark mit der Akkumulation von Flavonoiden (Shahzad A. Pandith, 2020). Es wäre auch zielführend, den Zusammenhang zwischen dem Expressionsniveau des CHS-Gens und des Flavonoidgehalts zu vertiefen, um das Zusammenspiel von genetischen und umweltbedingten Faktoren in der Flavonoid-Biosynthese besser zu verstehen.

Des Weiteren heben immer mehr Studien das Potenzial von *Origanum*-Arten als Hoffnungsträger in der Krebsprävention und -behandlung hervor. Besonders *O. majorana*, mit seinen hohen Gehalten an bioaktiven Verbindungen, hat krebshemmende Aktivitäten gezeigt. Weitere Forschung in diesem Bereich könnte vielversprechende Ergebnisse liefern (Zuhal Emire, 2022).

5. Anhang

Anhang Tab. 1: Wahl des passendsten Substitutionsmodells durch die Maximum-likelihood Methode (DNA)

Model	#Param	BIC	AICc	InL	Invariant	Gamma	R	Freq A	Freq T	Freq C	Freq G
JC	265	8006,171748	5290,672597	-2379,99847	n/a	n/a	0,5	0,25	0,25	0,25	0,25
JC+G	266	8006,175378	5280,431614	-2373,875427	n/a	0,05	0,5	0,25	0,25	0,25	0,25
JC+G+I	267	8012,555094	5276,566735	-2370,940427	0,49392971	0,05	0,5	0,25	0,25	0,25	0,25
K2+G	267	8014,583929	5278,595571	-2371,954845	n/a	0,05	0,715061823	0,25	0,25	0,25	0,25
JC+I	266	8014,768522	5289,024758	-2378,171999	0,49392971	n/a	0,5	0,25	0,25	0,25	0,25
K2	266	8020,237424	5294,493661	-2380,90645	n/a	n/a	0,714442042	0,25	0,25	0,25	0,25
T92	267	8021,60902	5285,620662	-2375,46739	n/a	n/a	0,714283626	0,230298868	0,230298868	0,269701132	0,269701132
T92+G	268	8021,776552	5275,543618	-2369,426298	n/a	0,05	0,714710514	0,230298868	0,230298868	0,269701132	0,269701132
K2+G+I	268	8024,229107	5277,996173	-2370,652575	0,49392971	0,05	0,715852534	0,25	0,25	0,25	0,25
K2+I	267	8026,844948	5290,85659	-2378,085354	0,49392971	n/a	0,714602789	0,25	0,25	0,25	0,25
T92+I	268	8027,303451	5281,070517	-2372,189747	0,49392971	n/a	0,714411387	0,230298868	0,230298868	0,269701132	0,269701132
T92+G+I	269	8028,204975	5271,727486	-2366,515651	0,49392971	0,05	0,715449757	0,230298868	0,230298868	0,269701132	0,269701132
HKY+G	270	8041,887906	5275,16588	-2367,232259	n/a	0,05	0,71419247	0,221853132	0,238744603	0,274030499	0,265371766
HKY	269	8047,854499	5291,37701	-2376,340413	n/a	n/a	0,714246441	0,221853132	0,238744603	0,274030499	0,265371766
HKY+G+I	271	8051,294928	5274,328384	-2365,810912	0,49392971	0,05	0,714110152	0,221853132	0,238744603	0,274030499	0,265371766
TN93+G	271	8052,782994	5275,81645	-2366,554944	n/a	0,05	0,713443258	0,221853132	0,238744603	0,274030499	0,265371766
HKY+I	270	8053,486745	5286,764718	-2373,031678	0,49392971	n/a	0,714220524	0,221853132	0,238744603	0,274030499	0,265371766
TN93	270	8058,455404	5291,733378	-2375,516008	n/a	n/a	0,714133978	0,221853132	0,238744603	0,274030499	0,265371766
TN93+G+	272	8062,314965	5275,103923	-2365,196072	0,49392971	0,05	0,712516906	0,221853132	0,238744603	0,274030499	0,265371766
TN93+I	271	8064,795538	5287,828994	-2372,561216	0,49392971	n/a	0,714003097	0,221853132	0,238744603	0,274030499	0,265371766
GTR+G	274	8089,013501	5281,313519	-2366,295623	n/a	0,05	0,713550946	0,221853132	0,238744603	0,274030499	0,265371766
GTR	273	8095,490386	5298,034864	-2375,658924	n/a	n/a	0,714338759	0,221853132	0,238744603	0,274030499	0,265371766
GTR+G+I	275	8098,878812	5280,934389	-2365,103421	0,49392971	0,05	0,712834745	0,221853132	0,238744603	0,274030499	0,265371766
GTR+I	274	8101,242008	5293,542026	-2372,409877	0,49392971	n/a	0,714728004	0,221853132	0,238744603	0,274030499	0,265371766

Anhang Tab. 2: mRNA-Codon-Verteilung

UUU	0.0		UCU	2.0		UAU	1.0		UGU	1.0
UUC	14.0		UCC	7.7		UAC	8.0		UGC	5.0
UUA	0.0		UCA	0.3		UAA	0.0		UGA	0.0
UUG	1.1		UCG	7.0		UAG	0.0		UGG	4.0
CUU	3.0		CCU	4.0		CAU	0.0		CGU	0.1
CUC	17.0		CCC	7.9		CAC	9.0		CGC	9.9
CUA	3.0		CCA	1.0		CAA	1.0		CGA	1.0
CUG	13.9		CCG	6.1		CAG	10.0		CGG	3.0
AUU	1.0		ACU	0.0		AAU	0.0		AGU	0.0
AUC	14.0		ACC	15.0		AAC	9.0		AGC	11.0
AUA	3.0		ACA	0.0		AAA	6.0		AGA	0.0
AUG	13.0		ACG	5.1		AAG	19.9		AGG	3.0
GUU	2.1		GCU	6.0		GAU	4.1		GGU	1.0
GUC	18.9		GCC	18.7		GAC	14.9		GGC	16.0
GUA	2.0		GCA	1.3		GAA	2.0		GGA	2.0
GUG	12.0		GCG	6.0		GAG	26.0		GGG	14.0

Anhang Tab. 3: Wahl des passendsten Substitutionsmodells durch die MLM (mRNA)

Model	#Param	BIC	AICc	InL	Invariant	Gamma	R	Freq A	Freq T	Freq C	Freq G
T92	267	6501,606119	3843,424142	-1654,253308	n/a	n/a	0,800344522	0,184359079	0,184359079	0,315640921	0,315640921
T92+G	268	6510,383235	3842,248963	-1652,662279	n/a	0,05	0,804376352	0,184359079	0,184359079	0,315640921	0,315640921
T92+I	268	6512,71374	3844,579468	-1653,827532	0,496569468	n/a	0,801869742	0,184359079	0,184359079	0,315640921	0,315640921
HKY	269	6521,695922	3843,60938	-1652,339036	n/a	n/a	0,800302671	0,201908553	0,166809605	0,320761117	0,310520724
T92+G+I	269	6521,919001	3843,832459	-1652,450575	0,496569468	0,1	0,821507727	0,184359079	0,184359079	0,315640921	0,315640921
TN93	270	6527,747064	3839,708278	-1649,385019	n/a	n/a	0,802544647	0,201908553	0,166809605	0,320761117	0,310520724
HKY+G	270	6530,610439	3842,571653	-1650,816707	n/a	0,05	0,805951703	0,201908553	0,166809605	0,320761117	0,310520724
HKY+I	270	6532,926798	3844,888013	-1651,974887	0,496569468	n/a	0,801126579	0,201908553	0,166809605	0,320761117	0,310520724
TN93+G	271	6537,160998	3839,169994	-1648,112399	n/a	0,05	0,807698478	0,201908553	0,166809605	0,320761117	0,310520724
TN93+I	271	6538,71317	3840,722166	-1648,888485	0,496569468	n/a	0,805551719	0,201908553	0,166809605	0,320761117	0,310520724
HKY+G+I	271	6542,122419	3844,131415	-1650,59311	0,496569468	0,1	0,820746384	0,201908553	0,166809605	0,320761117	0,310520724
TN93+G+	272	6548,074307	3840,13111	-1647,589467	0,496569468	0,1	0,871108764	0,201908553	0,166809605	0,320761117	0,310520724
JC	265	6558,549963	3920,272654	-1694,684404	n/a	n/a	0,5	0,25	0,25	0,25	0,25
GTR	273	6562,15704	3844,261677	-1648,651246	n/a	n/a	0,693666644	0,201908553	0,166809605	0,320761117	0,310520724
JC+G	266	6568,381754	3920,152098	-1693,620713	n/a	0,1	0,5	0,25	0,25	0,25	0,25
JC+I	266	6569,92724	3921,697584	-1694,393456	0,496569468	n/a	0,5	0,25	0,25	0,25	0,25
K2	266	6570,197311	3921,967655	-1694,528491	n/a	n/a	0,80028633	0,25	0,25	0,25	0,25
GTR+G	274	6570,249352	3842,401848	-1646,717815	n/a	0,05	0,462787366	0,201908553	0,166809605	0,320761117	0,310520724
GTR+I	274	6574,590693	3846,743188	-1648,888485	0,496569468	n/a	0,805551719	0,201908553	0,166809605	0,320761117	0,310520724
K2+G	267	6579,630945	3921,448969	-1693,265721	n/a	0,05	0,801377706	0,25	0,25	0,25	0,25
K2+I	267	6581,420305	3923,238329	-1694,160401	0,496569468	n/a	0,861243606	0,25	0,25	0,25	0,25
GTR+G+I	275	6583,951829	3846,15221	-1647,589467	0,496569468	0,1	0,871108764	0,201908553	0,166809605	0,320761117	0,310520724
K2+G+I	268	6590,662538	3922,528266	-1692,801931	0,496569468	0,1	0,882335673	0,25	0,25	0,25	0,25
JC+G+I	267	6666,83161	4008,649633	-1736,866054	0,496569468	0,05	0,5	0,25	0,25	0,25	0,25

6. Literaturverzeichnis

- Abdelkahim Bouyahya Imane Chamkhi, Taoufiq Benali, Fatima-Ezzahrae Guaouguaou, Abdelaali Balahbib, Nasreddine El Omari, Douae Taha, Omar Belmehdi, Zengin Ghokkan, Naoual El Menyiy** Traditional use, phytochemistry, toxicology, and pharmacology of *Origanum majorana* L. [Journal] // Journal of Ethnopharmacology . - 2021. - S. 1-31.
- Alwafa R.A., Mudalal S. und Mauriello G.** *Origanum syriacum* L. (Za'atar), from Raw to Go: A Review [Journal] // Plants. - 2021. - S. 1-17.
- Arenas Miguel** Trends in substitution models of molecular evolution [Journal] // Frontiers in Genetics . - 2015. - Article 319.
- Bektas Tepe Ahmet Cakir, Arzuhan Sihoglu Tepe** Medicinal Uses, Phytochemistry, and Pharmacology of *Origanum onites* (L.): A Review [Journal] // Chem. Biodiversity . - 2016. - S. 504 - 520 .
- Brigitte Lukas Rosabelle Samuel, Eduard Mader, Kemal Hüsnü can Baser, Hayri Duman, Johannes Novak** Complex evolutionary relationships in *Origanum* section *Majorana* (Lamiaceae) [Journal] // Botanical Journal of the Linnean Society. - 2013. - S. 667-686.
- Brigitte Lukas Rosabelle Samuel, Eduard Mader, Kemal Hüsnü Can Baser, Hayri Duman, Johannes Novak** Complex evolutionary relationships in *Origanum* section *Majorana* (Lamiaceae) [Journal] // Botanical Journal of the Linnean Society. - 2013. - S. 667-686.
- Ietswaart J. H.** A taxonomic revision of the genus *Origanum* (Labiatae) [Buch]. - Netherlands : Leiden Botanical Series , 1980.
- Jean-Luc Ferrer Joseph M. Jez, Marianne E. Bowman, Richard A. Dixon, Joseph P. Noel** Structure of chalcone synthase and the molecular basis of plant polyketide biosynthesis [Journal] // Nature Structural Biology . - 1999. - S. 775-784.
- Joelle Mesmar Rola Abdallah, Adnan Badran, Marc Maresca, Elias Baydoun** *Origanum syriacum* Phytochemistry and Pharmacological Properties: A Comprehensive Review [Journal]. - 2022. - S. 1-23.
- K.S. Sridevi Sangeetha S. Umamaheswari, C. Uma Maheswara Reddy, S. Narayana Kalkura** Flavonoids: Therapeutic Potential of natural pharmacological agents [Journal] // International Journal of pharmaceutical Sciences and research . - 2016. - S. 3924-3930.
- Kenan Turgut Yasar Özyigit, Begüm Tütüncü, Esra Ucar Sözmen** Agronomic and chemical performance of selected *Origanum dubium* Boiss. clones for industrial use [Journal] // Turkish Journal of Agriculture and Forestry. - 2017. - S. 272-277.
- Kunchur Guruprasad B. V. Bhasker Reddy, Madhusudan W. Pandit** Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence [Journal] // Protein Engineering, Design and Selection . - 1990. - Volume 4. - S. 155-161.
- Liu W. [et al.]** The Flavonoid Biosynthesis Network in Plants [Journal] // International Journal of Molecular Science . - 2021 . - S. 1-18.

Livia M. de Souza Luciano H. B. dos Santos, Joao R. B. F. Rosa, Carla C. da Silva, Camila C. Mantello, André R. O. Conson, Erivaldo J. Scaloppi Jr., Josefino de F. Fialho, Mario Luiz T. de Moraes, Paulo de S. Goncalves, Gabriel R. A. Margarido, Antonio A Linkage Disequilibrium and Population Structure in Wild and Cultivated Populations of Rubber Tree (*Hevea brasiliensis*) [Journal] // *frontiers in Plant Science* . - 2018. - Volume 9.

Mansour Karimi Ann Depicker, Pierre Hilson Recombinational Cloning with Plant Gateway Vectors [Journal] // *Plan Physiology*. - 2007. - S. 1144-1154.

Melanie Novak Dijana Jovanovic, Johannes Novak Variability of Chalcone Synthase in Chamomile Accessions (*Matricaria chamomilla*) [Journal] // *Planta Med.* - 2023. - S. 534-545.

Mülhardt Cornel Der Experimentator - Molekularbiologie/Genomics [Buch]. - Heidelberg : Springer-Verlag Berlin, 2013.

Schmidt Olaf G. Genetik und Molekularbiologie [Buch]. - Essen : Springer-Verlag, 2017.

Shahzad A. Pandith Salika Ramazan, Mohd Isfaq Khan, Zafar A. Reshi, Manzoor A. Shah Chalcone synthases (CHSs): the symbolic type III polyketide synthases [Journal] // *Springer nature* . - 2020. - S. 1-29.

Simon Whelan Nick Goldman A General Empirical Model of Protein Evolution Derived from Multiple Protein Families Using a Maximum-Likelihood Approach [Journal] // *Society for Molecular Biology and Evolution* . - 2001. - S. 691-699.

Siun Chee Tan Beow Chin Yiap DNA, RNA, and Protein Extraction: The Past and The Present [Journal]. - 2009.

Slatkin Montgomery Linkage disequilibrium — understanding the evolutionary past and mapping the medical future [Journal]. - California : *Nature Reviews Genetics* , 2008. - S. 477-485 .

T. T. H. Dao H. J. M. Linthorst, R. Verpoorte Chalcone synthase and its functions in plant resistance [Journal] // *Phytochem Rev.* . - 2011. - S. 397-412.

Tajima Fumio Statistical Method for Testing the Neutral Mutation Hypothesis by DNA [Journal] // *Genetics Society of America* . - 1989. - S. 585-595.

Tianhua Niu Zhaohui S. Qin, Xiping Xu, Jun S. Liu Bayesian Haplotype Inference for Multiple Linked Single-Nucleotide Polymorphisms [Journal] // *The American Society of Human Genetics* . - 2002. - S. 157-169.

Yonekura-Sakakibara K Higashi Y, Nakabayashi R The Origin and Evolution of Plant Flavonoid Metabolism [Journal] // *Frontiers in Plant Science* . - 2019. - Volume 10.

Yuxing Deng Caili Li, Heqin Li, Shanfa Lu Identification and Characterization of Flavonoid Biosynthetic Enzyme Genes in *Salvia miltiorrhiza* (Lamiaceae) [Journal] // *Molecules*. - 2018. - S. 1-20.

Zuhal Emire Erdal Yabalak Can *Origanum* be a hope for cancer treatment? A review on the potential of *Origanum* species in preventing and treating cancers [Journal] // *International Journal of Environmental Health Research*. - 2022. - S. 894-910.

7. Abbildungs- und Tabellenverzeichnis

Abbildung 1: Flavan-Grundstruktur (PubChem, 3D Wire-Frame, National Center for Biotechnology Information. PubChem Compound Summary for CID 94156, Flavan. https://pubchem.ncbi.nlm.nih.gov/compound/Flavan . Accessed Oct. 28, 2024)	15
Abbildung 2: Allgemeine Flavonoid-Biosynthesewege in Pflanzen. (Die Pfeile in gr ün, blau und magenta kennzeichnen Enzyme der CYP-, 2OGD- und SDR-Superfamilien. Hier sind die beteiligten Enzyme und ihre Abkürzungen: ANR – Anthocyanidin-Reduktase; ANS – Anthocyanidin-Synthase; CHI – Chalcon-Isomerase; CHR – Chalcon-Reduktase; CHS – Chalcon-Synthase; DFR – Dihydroflavonol-4-Reduktase; F2H – Flavanon-2-Hydroxylase; F3H – Flavanon-3-Hydroxylase; F3'H – Flavonoid-3'-Hydroxylase; F3'5'H – Flavonoid-3'5'-Hydroxylase; FLS – Flavonol-Synthase; FNS – Flavon-Synthase; HID – 2-Hydroxyisoflavanon-Dehydratase; IFS – Isoflavon-Synthase; LAR – Leucoanthocyanidin-Reduktase; LDOX – Leucoanthocyanidin-Dioxygenase) (Yonekura-Sakakibara K, 2019)	18
Abbildung 3: Chemische Struktur von Chalkon (PubChem, 3D Wire-Frame, National Center for Biotechnology Information. PubChem Compound Summary for CID 637760, Chalcone. https://pubchem.ncbi.nlm.nih.gov/compound/Chalcone . Accessed Oct. 28, 2024)	19
Abbildung 4: a, 3D-Struktur des ReCHS2-Proteins aus Rheum australe (Polygonaceae); b, Detailausschnitt: zeigt die katalytische Triade bestehend aus Cys164, His303 und Asn336 sowie Phe 216 (Nummerierung aus Medicago sativa), Cartoon-Modell (Shahzad A. Pandith, 2020)	20
Abbildung 5: MLG-Muster-Verteilung innerhalb der Origanum-Populationen	52
Abbildung 6: Graphische Darstellung von Linkage Disequilibrium (r_d) hinsichtlich der Verknüpfung der Marker (blau... keine Kopplung; lila... moderate Kopplung; hellrot... starke Kopplung; rot... sehr signifikante Kopplung)	53
Abbildung 7: Der phylogenetische Baum wurde mithilfe der Neighbor-Joining-Methode abgeleitet und anhand der Bootstrap-Methode auf Basis von 100 Wiederholungen kalkuliert. Die Analyse umfasste 134 Klonsequenzen, mit insgesamt 1565 Basenpaarpositionen. Die evolutionären Abstände wurden mit dem Jukes-Cantor-Modell berechnet. Die Zahlen geben die Zuverlässigkeit der ausgewerteten Distanzen in % an. (OO... Origanum onites, OS... Origanum syriacum; OM... Origanum majorana; OD... Origanum dubium)	55
Abbildung 8: MLG-Muster-Verteilung innerhalb der Origanum-Populationen	59
Abbildung 9: Graphische Darstellung von Linkage Disequilibrium (r_d) hinsichtlich der Verknüpfung der Marker (blau... keine Kopplung; lila... moderate Kopplung; hellrot... starke Kopplung; rot... sehr signifikante Kopplung)	60
Abbildung 10: Der phylogenetische Baum wurde mithilfe der Neighbor-Joining-Methode abgeleitet und anhand der Bootstrap-Methode auf Basis von 100 Wiederholungen kalkuliert. Die Analyse umfasste 134 Klonsequenzen, mit insgesamt 1166 Basenpaarpositionen. Die evolutionären Abstände wurden mit dem Tamura-3-Parameter-Modell berechnet. Die Zahlen geben die Zuverlässigkeit der ausgewerteten Distanzen in % an. (OO ... Origanum onites, OS... Origanum syriacum; OM... Origanum majorana; OD... Origanum dubium)	61
Abbildung 11: Aminosäuren-Sequenz Klon M95 aus M85 (O. majorana) – markierte Stelle: 284 K	62
Abbildung 12: Aminosäuren-Sequenz Klon M79 aus 230283 (O. onites) – markierte Stelle 284 T	62

Abbildung 13: Der phylogenetische Baum wurde mithilfe der JTT-Methode abgeleitet und anhand der Bootstrap-Methode auf Basis von 100 Wiederholungen kalkuliert, mit einer Zuverlässigkeit von 73 %. Die Analyse umfasste 134 AA-Sequenzen, mit insgesamt jeweils 388 Positionen. Die evolutionären Abstände wurden mit der MLM-Methode (ausgehend vom höchsten log likelihood -1114,49) und des Whelan and Goldman-Modells berechnet. (OO ... *Origanum onites*, OS ... *Origanum syriacum*; OM ... *Origanum majorana*; OD ... *Origanum dubium*) _____ 64

Abbildung 14: Das Model-Template Alignment der Klonsequenz M95 aus M85, *O. majorana* (blaue Sequenz) wurde mit dem CHS-Monomer-Template aus *Fagus sylvatica* (weiße Sequenz) verglichen. Identische Sequenzbereiche zwischen den beiden Proteinen sind durch Umrandungen hervorgehoben. Rot-markierte Positionen sind zum einen die konservierten Stellen: Cys 164, His 303, Asn 336, Phe 215 und Phe 265 und zum anderen die N-terminale Domäne: Met 137. (Pfeile ... Helices, Umkreisungen ... Faltblätter) (Swiss-Model Expasy) _____ 66

Abbildung 15: Voraussage der 3D-Struktur anhand des Model-Template-Alignments (für die Funktion wichtige konservierte Aminosäurereste (Jean-Luc Ferrer, 1999) und die N-terminale Domäne wurden in der 3D-Struktur als Strukturformeln sichtbar gemacht, um sie herauszuheben) (Swiss-Model Expasy) _____ 67

Abbildung 16: CHS 3D-Struktur (aus einem anderen Blickwinkel) (Swiss-Model Expasy) _____ 68

Abbildung 17: Konservierte Positionen des CHS-Gens – a Cys 164: entscheidend für die Aktivierung und Bindung des Coumaryl-CoA; b His 303: katalysiert die Aktivierungsreaktion von Cys; c Asn 336: spielt eine Rolle in der Decarboxylierungs-Reaktion von Malonyl-CoA und stabilisiert das Reaktionsintermediat; d Phe 215 und Phe 265: Unterstützung des Prozesses (Jean-Luc Ferrer, 1999) (Swiss-Model Expasy) _____ 69

Abbildung 18: Phylogenetische Analyse von der DNA bis hin zur AS-Sequenz _____ 72

Abbildung 19: Chemische Struktur von Threonin (T) und Lysin (K) (Pubchem; National Center for Biotechnology Information. PubChem Compound Summary for CID 6288, Threonine. <https://pubchem.ncbi.nlm.nih.gov/compound/Threonine>. Accessed Oct. 28, 2024; National Center for Biotechnology Information. PubChem Compound Summary for CID 5962, Lysine. <https://pubchem.ncbi.nlm.nih.gov/compound/Lysine>. Accessed Oct. 28, 2024) _____ 74

Tabelle 1: Morphologische Merkmale zur Differenzierung von *Origanum onites*, *O. syriacum*, *O. dubium* und *O. Majorana* (Ietswaart) _____ 8

Tabelle 2: verwendete Primer _____ 34

Tabelle 3: DNA-Annotation _____ 48

Tabelle 4: Darstellung der SNPs bezogen auf die genetische Vielfalt (Hexp ... erwartete Heterozygotie, Evenness ... Wahrscheinlichkeit einer Gleichverteilung der Allele) _____ 50

Tabelle 5: Darstellung der genetischen Variabilität innerhalb der untersuchten *Origanum*-Arten (N ... Anzahl der Klone; MLG ... Multilocus-Genotypen; H ... Diversitätsindex zwischen N und MLG; λ ... Maß für die Diversität, bezogen auf die MLG-Muster-Verteilung; E.5... spezifischer Evenness-Index; Hexp... erwartete Heterozygotie; rbarD... standardisierte Linkage Disequilibrium; --- keine Verteilung einer Variation bestimmbar) _____ 51

<i>Tabelle 6: Wahrscheinlichkeit einer Substitution von einer Base zu einer anderen Base, berechnet nach dem Tamura-(92)-Modell. Die Analyse umfasste 134 Klonsequenzen, mit insgesamt 1166 Basenpaarpositionen. Die Zahlen geben die Wahrscheinlichkeiten in % an. Transitionen sind fett und Transversionen kursiv angezeigt. (A ... Adenin; T/U ... Thymin/Uracil; C ... Cytosin; G ... Guanin)</i>	56
<i>Tabelle 7: Darstellung der Distanzen zwischen den Arten, wobei die Anzahl der Änderungen betrachtet wurde.</i>	57
<i>Tabelle 8: Darstellung der SNPs bezogen auf die genetische Vielfalt (H_{exp} ... erwartete Heterozygotie; Evenness ... Wahrscheinlichkeit einer Gleichverteilung der Allele)</i>	58
<i>Tabelle 9: Darstellung der genetischen Variabilität innerhalb der untersuchten Origanum-Populationen (N ... Anzahl der Klone; MLG ... Multilocus-Genotypen; H ... Diversitätsindex zwischen N und MLG; λ ... Maß für die Diversität, bezogen auf die MLG-Muster-Verteilung; E.5... spezifischer Evenness-Index; H_{exp}... erwartete Heterozygotie; r_{barD}... standardisierte Linkage Disequilibrium; --- keine Verteilung einer Variation bestimmbar)</i>	58
<i>Tabelle 10: Darstellung der Distanzen zwischen den Arten, wobei die Anzahl der Änderungen betrachtet wurde. Die Anzahl der AS-Substitutionen pro Position errechnet über das Mittel aller Sequenzpaare zwischen den Arten.</i>	63
<i>Formel 1: Umrechnung von fmol in ng (Life Technologies, 2012, Cayla, SA)</i>	39

"Ich habe mich bemüht, sämtliche Inhaber*innen der Bildrechte ausfindig zu machen und ihre Zustimmung zur Verwendung der Bilder in dieser Arbeit eingeholt. Sollte dennoch eine Urheberrechtsverletzung bekannt werden, ersuche ich um Meldung bei mir."

Datum

Unterschrift