



Universität für Bodenkultur Wien
Department für Integrative Biologie und Biodiversitätsforschung
Institut für Botanik

Mikrosatelliten-Entwicklung
für die einjährige Schlammlingsart *Cyperus fuscus*

Masterarbeit
Zur Erlangung des akademischen Grades
des Diplomingenieurs (Dipl. Ing.)

Betreuung:
Univ. Prof. Dipl. Geograph Dr. Karl Georg Bernhardt
Priv.-Doz. Mag. Dr. Karin Tremetsberger

Eingereicht am Institut für Botanik (Botany)
Department für Integrative Biologie und Biodiversitätsforschung
der Universität für Bodenkultur Wien

Eingereicht von David WIESER, Bakk. techn.
Matr. Nr.: 0640712

Eidesstattliche Erklärung

Ich versichere:

- dass ich diese Diplomarbeit selbstständig verfasst, keine anderen als die angegebenen Hilfsmittel und Quellen verwendet habe und mich aus sonst keiner unerlaubten Hilfe bedient habe.
- dass ich dieses Diplomarbeitsthema bisher weder im In- noch im Ausland in irgendeiner Form als Prüfungsarbeit vorgelegt habe.

Ort, Datum

Unterschrift

Abstract

The annual wetland plant species *Cyperus fuscus*, as a representative for the semi-aquatic communities of dwarf rushes (*Isoëto-Nanojuncetea*; DEIL, 2005) on periodically draining riverbanks, represents an ideal candidate for population genetic studies of natural and anthropogenic habitats and also serves to elucidate the role of the soil seed bank in the life cycle of the species itself. An overall of 21 microsatellite markers has been developed in this work making use of two different methods. 4 markers have been developed with the freeware programs MSATCOMMANDER (Faircloth, 2008) and primer3 (ROZEN & SKALETZKY, 2000) and 17 markers were created by ECOGENICS (Balgach, Switzerland). The methods used show significant differences in the quality and number of usable markers. The markers were tested on a natural as well as on an anthropogenically severely affected population to identify both the impact of ecosystem disturbances on the genetic diversity and determine the role of the seed bank in the soil as a genetic memory. Likewise, the yet not exactly known ploidy of the species should be clearly stated. For this purpose the *allelic richness*, the heterozygosity, the inbreeding coefficient (F_{IS}) as well as the *fixation index* (F_{ST}) between sub-populations of the soil seed bank and the overground sub-population have been determined. Both the *allelic richness* with 3.23 and the heterozygosity H_{OBS} 0.134 and 0.588 H_{EXP} of the natural population compared to the *allelic richness* 2.62 and the heterozygosity H_{OBS} 0.059 and 0.336 H_{EXP} of the anthropogenic population show, that the "ecosystem dynamics" (e. g. flooding) cause an increase in genetic diversity in natural populations. The *fixation index* could be calculated for the anthropogenic population with 0.114 (significance threshold = 0.013) and for the natural population with 0,013 (significance threshold = 0.63). This shows a significant differentiation between sub-populations of the anthropogenic population while the differentiation between the sub-populations of the natural population is not significant. *Cyperus fuscus* could unambiguously be determined to be a diploid species. The 21 newly developed microsatellite markers will be useful in further research to analyse the population structure and the role of the soil seed bank in the life-cycle of this species.

Kurzfassung

Die einjährige Riedgrasart *Cyperus fuscus* dient als Vertreter der ephemeren, an sommerlich trocken-fallenden Flussufern wachsenden Zwergbinsengesellschaften („Schlammfluren“), für populationsgenetische Untersuchungen natürlicher und anthropogener Habitats sowie zur Aufklärung der Rolle der Samenbank im Lebenszyklus der Art. Dazu sind in dieser Arbeit insgesamt 21 Mikrosatelliten-Marker mit zwei unterschiedlichen Methoden entwickelt worden. 4 Marker sind mit den Freeware-Programmen MSAT-COMMANDER (FAIRCLOTH, 2008) und PRIMER3 (ROZEN & SKALETZKY, 2000) entwickelt worden und 17 Marker wurden von ECOGENICS (Balgach, Schweiz) erstellt. Die verwendeten Methoden zeigen deutliche Unterschiede in der Qualität und Anzahl an verwendbaren Markern. Die Marker wurden an einer natürlichen und einer anthropogen stark beeinflussten Population getestet um einerseits die Auswirkungen von Ökosystem-Störungen auf die genetische Vielfalt zu ermitteln und andererseits die Funktion der Samenbank im Boden als genetisches Gedächtnis zu überprüfen. Ebenso soll der Ploidiegrad der Art eindeutig festgestellt werden. Es wurden dazu die *allelic richness*, der Heterozygotiegrad, der Inzuchtkoeffizient (F_{IS}) sowie der *fixation index* (F_{ST}) zwischen Subpopulationen im Boden (Samenbank) und oberirdisch entwickelter Subpopulationen ermittelt. Sowohl die *allelic richness* mit 3,23 als auch der Heterozygotiegrad mit H_{OBS} 0,134 und H_{EXP} 0,588 der natürlichen Population zeigen gegenüber der *allelic richness* von 2,62 und dem Heterozygotiegrad für H_{OBS} 0,059 und H_{EXP} 0,336 der anthropogenen Population, dass die „Ökosystem-Dynamik“ (Überflutung) eine Erhöhung der genetischen Vielfalt in der natürlichen Population bewirkt. Der *fixation index* konnte für die anthropogene Population mit 0,114 (Signifikanzschwelle = 0,013) und für die natürliche Population mit 0,013 (Signifikanzschwelle = 0,63) ermittelt werden. Dadurch wird für die anthropogene Population eine signifikante Differenzierung zwischen den Subpopulationen aufgezeigt während die Subpopulationen der natürlichen Population sich nicht signifikant differenzieren. *Cyperus fuscus* konnte zudem als eindeutig diploid festgelegt werden. Die 21 entwickelten Mikrosatelliten-Marker werden in weiteren Studien helfen, die Populationsstruktur sowie die Rolle der Samenbank im Lebenszyklus von *Cyperus fuscus* zu analysieren.

Inhaltsverzeichnis

1 Einführung.....	13
2 Hintergrund der Arbeit.....	16
2.1 FWF Projekt.....	16
2.2 Charakterart <i>Cyperus fuscus</i>	17
2.3 Geprüfte Hypothesen.....	18
3 Grundlagen.....	21
3.1 Mikrosatelliten in der Populationsgenetik.....	21
3.1.1 Entstehung und Entdeckungsgeschichte von Mikrosatelliten.....	22
3.1.2 Mutationsmechanismus - Entstehung von Allelen.....	22
3.1.3 Mikrosatelliten als molekulare Marker.....	23
3.1.3.1 Primer.....	24
3.1.3.2 Dominante vs. Co-dominante Marker.....	24
3.1.3.3 Mikrosatelliten-Marker für Populationsgenetische Untersuchungen.....	25
3.2 Suche nach Mikrosatelliten und Primerdesign.....	26
3.2.1 DNA-Extraktion.....	26
3.2.2 Sequenzdaten aus genomischer DNA	27
3.2.2.1 Pyrosequenzierung.....	27
3.2.2.2 Enrichment genomischer DNA.....	28
3.2.3 Grundlagen zur Isolation von Mikrosatelliten und Primerdesign.....	30
3.2.3.1 Die Verwendung von tag's und PIG-tails.....	31
3.2.3.2 Primer-PIG-tail.....	32
3.2.3.3 Programme zur Isolation von Mikrosatelliten und Primerdesign.....	32
3.2.3.3.1 MSATCOMMANDER.....	33
Auswahl der SSR Art.....	34
Flanking-regions auf Einzigartigkeit überprüfen - NCBI Blast.....	34
3.2.3.3.2 Primerdesign mittels PRIMER 3.....	35
Parameter für das Primerdesign.....	35
3.3 Methoden zum Testen von Mikrosatelliten-Marker.....	37
3.3.1 PCR Amplifikation.....	37
3.3.1.1 Funktionsweise.....	37
3.3.1.2 PCR Programme - Gradienten und Touchdown-PCR.....	39
3.3.2 Gelelektrophorese.....	40
3.3.3 Kapillarelektrophorese	41

3.3.3.1 Genetic Analyzer 3500.....	42
Fragment sizing am Genetic Analyzer 3500.....	43
3.3.4 Finden geeigneter SSR Marker.....	44
4 Methoden.....	46
4.1 Methoden zur Isolation der Mikrosatelliten und Primerdesign.....	46
Test-Individuen.....	47
4.1.1 Methode 1: Isolation von Mikrosatelliten mittels MSATCOMMANDER und Primerdesign mittels PRIMER3.....	48
4.1.1.1 Ausgangsdaten - Sequenzdaten.....	48
4.1.1.2 Primer Design.....	49
SSR Art.....	49
Parameter für das Primerdesign.....	50
Verwendete tag's und PIG-tails.....	50
4.1.1.3 Primer Tests und Auswahl geeigneter Mikrosatelliten-Marker.....	51
4.1.1.3.1 PCR Amplifikation.....	51
4.1.1.3.2 Test am Agarosegel.....	54
4.1.1.3.3 Fragmentanalyse am Genetic Analyzer 3500.....	55
4.1.1.3.4 Auswahl geeigneter SSR Marker.....	55
4.1.1.4 Zusammenfassung Methode 1.....	56
4.1.2 Methode 2: Isolation von Mikrosatelliten mit Enrichment und Primerdesign durch Ecogenics.....	57
4.1.2.1 Ausgangsdaten - Enrichment und 454 Pyrosequenzierung.....	57
4.1.2.2 Primerdesign.....	58
Suchen nach Mikrosatelliten und Primerdesign.....	58
Verwendete tag's und PIG-tails.....	58
4.1.2.3 Primer Tests und Auswahl geeigneter SSR Marker – Methode 2.....	59
4.1.2.3.1 PCR Amplifikation.....	59
4.1.2.3.2 Test am Agarosegel.....	60
4.1.2.3.3 Fragmentanalyse am Genetic Analyzer 3500.....	61
4.1.2.3.4 Auswahl geeigneter Mikrosatelliten-Marker.....	61
4.1.2.4 Zusammenfassung Methode 2.....	62
4.2 Anwendungstests der Marker.....	63
4.2.1 Populationen und Marker.....	63
4.2.2 Berechnete Werte.....	65
Messwerte im Detail.....	66
4.2.3 Verwendete Programme.....	68

5 Ergebnisse	69
5.1 Entwickelte Mikrosatelliten-Marker für <i>Cyperus fuscus</i>	69
Motivhäufigkeiten.....	72
5.2 Ergebnisse zu den Anwendungstests der Marker.....	74
Ploidiegrad.....	74
Anzahl Allele und allelic richness.....	75
Heterozygotie.....	76
Inzuchtkoeffizient.....	77
<i>F_{ST}</i> -Wert.....	78
6 Diskussion	80
6.1 Vergleich der verwendeten Methoden zum Primerdesign.....	80
Vergleich der Read-längen und Markerqualität.....	80
Verwendung freier Software.....	82
6.2 Beantwortung der Hypothesen.....	84
6.2.1 Ploidiegrad.....	84
6.2.2 Ökosystem-Störung erhöht die genetische Vielfalt.....	84
6.2.3 Oberflächenpopulation aus angepassten Genotypen.....	85
7 Literaturverzeichnis	87
8 Abbildungsverzeichnis	98
9 Tabellenverzeichnis	99

1 Einführung

Das Braune Zyperngras, *Cyperus fuscus*, ist eine einjährige, semiaquatische Riedgrasart, die im Mittelmeerraum sowie in gemäßigten Klimazonen Europas, Asiens und Nordamerikas vorzufinden ist. Es wächst bevorzugt an schlammigen, sandigen und kiesigen Untergründen an Fluss- und Seeufern, aber auch in anthropogenen Lebensräumen wie Kiesgruben, feuchten Wiesen und traditionell bewirtschafteten Fischteichen. (PHILIPPI, 1998) Letztere stellen für diese Art ein wichtiges Sekundärhabitat dar (TRAXLER, 1991; BERNHARDT et al., 2004). Sowohl Pflanzen mit $2n=36$ als auch Pflanzen mit $2n=72$ Chromosomen sind bekannt (KRAHULCOVÁ, 2003), jedoch konnte der Ploidiegrad der schnellwüchsigen Art bisher nicht eindeutig festgestellt werden. In der Roten Liste der gefährdeten Arten der IUCN als nicht gefährdet gelistet, kann *Cyperus fuscus* jedoch als Charakterart bedrohter und schwindender Zwergbinsengesellschaften (Anm.: auch bekannt als sogenannte „Schlammlingsfluren“. Klasse *Isoëto-Nanojuncetea*; DEIL, 2005) angesehen werden (NATURA 2000 FFH RICHTLINIE). Dieser Vegetationstyp ist durch periodische Wasserstandsschwankungen mit Wechsel zwischen Austrocknung und Überflutung des Bodens charakterisiert (BERNHARDT & POSCHLOD, 1993). An diese Gegebenheiten angepasste Arten weisen eine relativ kurze Entwicklungsdauer von Keimung bis Samenreife und eine hohe Samenproduktion auf, die es ermöglicht, ungünstige Verhältnisse im Boden zu überdauern. Im Lebenszyklus der schnellwüchsigen Art spielt die persistente Samenbank im Boden somit eine wichtige Rolle, weil die genetische Vielfalt über Jahre hinweg erhalten bleibt und daher als eine Art genetisches Gedächtnis fungiert. Untersuchungen von Diasporenbanken, wie sie beispielsweise in bebautem Ackerland vorkommen, haben gezeigt, dass nahezu die gesamte phänotypische Variation einer Art in der Diasporenbank im Boden gespeichert ist (BERNHARDT & HURKA, 1989; BERNHARDT, 1995.; u. a.). Bisher wurde die genetische Vielfalt in der Diasporenbank im Boden nur in wenigen Studien berücksichtigt (z. B.: TONSOR et al., 1993; CABIN et al., 1998; MCCUE & HOLTSFORD, 1998; MORRIS et al., 2002; KOCH et al., 2003) Im Zuge eines FWF-Projektes (siehe 2.1) soll anhand populationsgenetischer Untersuchungen die Rolle der Samenbank erhellt sowie die genetische Vielfalt derselben in verschiedenen natürlichen und anthropogenen Lebensräumen ermittelt und verglichen werden.

Ziel dieser Arbeit ist es, entsprechende Mikrosatelliten-Marker für die Art *Cyperus fuscus* zu entwickeln und zu testen. Damit sollen die populationsgenetischen Untersuchungen der Samenbank ermöglicht werden.

Die Entwicklung genetischer Marker wurde durch das Aufkommen neuer Technologien wie der 454-Pyrosequenzierung und frei zugänglicher Software erschwinglich und ist mit weitgehend unkomplizierten und gängigen Labortechniken und modernen Gerätschaften in relativ kurzer Zeit realisierbar. Für populationsgenetische Studien werden häufig Mikrosatelliten herangezogen, da diese eine hohe Anzahl an Allelen sowie hohe Heterozygotiegrade aufweisen und sich deren Isolation zudem einfach gestaltet (SCHLÖTTERER, 2004). In dieser Arbeit werden zwei unterschiedliche Methoden zur Markerentwicklung angewandt und präsentiert: zum einen die Isolation von Mikrosatelliten und die Markerentwicklung mithilfe der Freeware-Programme MSATCOMMANDER (FAIRCLOTH, 2008) und PRIMER3 (ROZEN & SKALETSKY, 2000) und zum anderen die Isolation von Mikrosatelliten und die Markerentwicklung durch den spezialisierten Betrieb ECOGENICS (Balgach, Schweiz). Es konnten durch diese beiden Methoden insgesamt 21 Mikrosatelliten-Marker für *Cyperus fuscus* optimiert werden. Um die Anwendbarkeit der entwickelten Marker zu bestätigen, werden 8 davon herangezogen, um zwei Hypothesen zu überprüfen. Dazu wird auf ein Subset von zwei ausgewählten Populationen zurückgegriffen, die sich hinsichtlich der Störung und der geografischen Lage stark unterscheiden (siehe 4.1). Zudem soll die bisher ungeklärte Frage nach der Ploidie der Art beantwortet werden. Die untersuchten Hypothesen beziehen sich auf die Rolle der Samenbank unterschiedlicher Ökosysteme, in denen *Cyperus fuscus* vorgefunden werden kann, und vergleichen dabei natürliche Habitate mit stark anthropogen beeinflussten Standorten. Beruhend auf der Annahme, dass die Ökosystemstörung einen starken Einfluss auf die genetische Vielfalt hat, wird erwartet, dass eine Korrelation zwischen Intensität der Störung natürlicher Habitate beziehungsweise des Eingriffs anthropogener Standorte und der genetischen Variation einer Population besteht. Die erste Hypothese besagt somit, dass Populationen natürlicher Standorte mit eher stochastischem Störungsregime, wie zeitweisen Überflutungen oder Trockenperioden, eine größere genetische Vielfalt aufweisen als Populationen traditionell bewirtschafteter Fischteiche mit regelmäßigen Managementregime. Dazu wird die genetische Vielfalt anhand der Anzahl der gefundenen Allele, der *allelic richness*, des Inzuchtkoeffizienten sowie des Heterozygotiegrades der untersuchten Populationen ermittelt und die Ergebnisse werden miteinander verglichen. Weiters wird überprüft, ob die Samenbank im Boden tatsächlich als genetisches Gedächtnis fungiert, welches zahlreiche Genotypen früherer Generationen speichert. Somit wird die zweite Hypothese überprüft, die besagt, dass die Oberflächenpopulation nicht zufällig entstanden ist, sondern aus Genotypen der Samenbank besteht, die an die vorherrschenden Lebensraumbedingungen angepasst sind. Es wird somit erwartet, dass sich die Ge-

notypenzusammensetzung der Samen aus den Bodenschichten deutlich von der Genotypenzusammensetzung der Oberflächenpopulation unterscheidet. Weiters sollen die genannten Unterschiede an den anthropogen beeinflussten Habitaten – mit ihren regelmäßigen Managementregimen – deutlicher in Erscheinung treten als an den natürlichen Standorten. Dazu werden die Subpopulationen aus den unterschiedlichen Bodentiefen sowie die Oberflächenpopulation der zwei ausgewählten Standorte analysiert. Mit Zuhilfenahme der entwickelten Mikrosatelliten-Marker soll für die Überprüfung dieser Hypothese der F_{ST} -Wert oder *fixation index* als Maß für die genetische Differenzierung zwischen Subpopulationen ermittelt und verglichen werden. Die technische Ausrichtung dieser Arbeit setzt ein gewisses Grundwissen über die Entstehung, Anwendung und Art genetischer Marker sowie über die zur Entwicklung von Mikrosatelliten-Markern nötigen Computerprogrammen und Labortechniken voraus. Diese werden im Kapitel Grundlagen dargelegt (Kapitel 3). Die in dieser Arbeit angewandten Methoden zur Entwicklung der Mikrosatelliten-Marker werden im Kapitel Methoden (4) im Detail besprochen. Ebenso werden die Anwendungstests der Marker zur Prüfung der Hypothesen im Punkt 4.2 aufgezeigt. Im Ergebniskapitel (5) werden die entwickelten Mikrosatelliten-Marker für *Cyperus fuscus* vorgestellt sowie die Ergebnisse der Anwendungstests präsentiert. Die Diskussion im Kapitel 6 zeigt die Stärken und Schwächen der beiden angewandten Methoden auf und vergleicht diese kritisch miteinander. Außerdem werden die Ergebnisse der Anwendungstests betrachtet und diskutiert, um mögliche Aussagen hinsichtlich der Hypothesen treffen zu können. Im nachfolgenden Punkt werden die Hintergründe der Arbeit beleuchtet, das Objekt dieser Arbeit, *Cyperus fuscus*, vorgestellt und die geprüften Hypothesen ausführlich präsentiert.

2 Hintergrund der Arbeit

2.1 FWF Projekt

Genetische und / oder phänotypische Variation sind für eine Population erforderlich, um sich an Umweltveränderungen anpassen zu können (FRANKHAM et al. 2010). In vielen einjährigen Pflanzenarten mit einer persistenten Bodensamenbank gibt es eine zeitliche Komponente der genetischen Variation. Samen von kurzlebigen Arten können sehr langlebig sein und bis zu 50 Jahre und länger im Boden überdauern, was zu einer komplexen Altersstruktur der Bodensamenbank führt (LEVIN, 1990). Für die Aufrechterhaltung der genetischen Variation haben diese Diasporenbanken im Boden eine fundamentale Bedeutung, weil sie als Puffer gegen genetische Drift agieren, die Auswirkungen von Selektion verzögern sowie aufgrund der erhöhten Mutationsrate in alternden Zellen auch direkt als Quelle genetischer Vielfalt dienen und weil in Jahren mit hoher Samenproduktion bestimmte Merkmale begünstigt werden (LEVIN, 1990).

Persistente Bodendiasporenbanken sind charakteristisch für einjährige, ephemere Arten (LECK et al., 1989), wie beispielsweise *Cyperus fuscus* und andere Vertreter von semi-aquatischen Zwergbinsengesellschaften. In Europa sind diese Gemeinschaften rückläufig. Wichtige Rückzugsgebiete von Populationen von Zwergbinsengesellschaften sind sekundäre, vom Menschen geschaffene Fischteiche und Fischlagerbecken (= Fischhälter). In Europa konzentriert sich die Verbreitung dieser Habitate vor allem im Bereich der Böhmisches Masse in Österreich und der Tschechischen Republik, wo Fischteiche und Fischhälter eine Jahrhunderte andauernde traditionelle Nutzung aufweisen (STANJEK, 1994; ŠUMBEROVÁ et al., 2005). Diese Gebiete sind zugleich der Verbreitungsschwerpunkt sekundärer Habitate von *Cyperus fuscus* (ŠUMBEROVÁ et al., 2006).

Das intensive Störungsregime in solchen Habitaten übt vermutlich einen starken Selektionsdruck auf die Populationen aus. Es gibt allerdings kaum empirische Belege für die Rolle und die Bedeutung der persistenten Bodendiasporenbanken für die Erhaltung genetischer Variation.

Das im Sommer 2012 gestartete FWF-Projekt soll die Rolle der persistenten Samenbank als integraler Bestandteil der Lebensgeschichte von kurzlebigen, ephemeren Pflanzenarten aufklären. Dazu werden verschiedene Glashauserversuche als auch populationsgenetische Untersuchungen mittels genetischer Marker durchgeführt. Im Vorfeld dieser Arbeit wurden dazu zahlreiche Samenproben der Art aus über 30 Flächen unterschiedlicher Nutzungs-

typen, aus den jeweiligen Oberflächenpopulationen sowie aus zwei verschiedenen Bodentiefen, entnommen. Im Sommer 2013 wurden aus den Bodenproben und den reifen Samen der Oberflächenpopulationen zahlreiche Individuen der Zielart im Glashaus angezogen sowie deren DNA extrahiert.

2.2 Charakterart *Cyperus fuscus*

Das Braune Zyperngras *Cyperus fuscus* ist ein Vertreter der heimischen Riedgräser und eine Charakterart für Zwergbinsengesellschaften (*Isoëto-Nanojuncetea*; DEIL, 2005) mit kurzer Lebensdauer und mit sich meist nicht überlappenden Generationen (z. B. PHILIPPI, 1968; POSCHLOD, 1996). *Cyperus fuscus* hat ein sehr weites Habitatsspektrum (TRAXLER, 1991). Die einjährige Pflanze ist bevorzugt an Sumpfrändern, offenen, teilweise verdichteten Sand- und Schlammböden, in feuchten Rinnen an Wegrändern sowie an Anlandungen und an Ufer extensiv genutzter Fischteiche zu finden. Wie die meisten Riedgräser hat auch das braune Zyperngras einen dreieckigen Stängel. Die Gesamthöhe der Pflanze variiert je nach Standortbedingungen zwischen 3 und 45 cm. Es wird charakterisiert durch die namensgebenden schwarzbraunen Deckblätter. Die braunen Ährchen bilden den köpfchenartigen Blütenstand, der oft noch zusätzliche, gestielte Köpfchen an den Seiten besitzt (FISCHER et al., 2008, AICHELE & SCHWEGLER, 2011). *Cyperus fuscus* ist windbestäubt und selbstkompatibel (EAST, 1940). Eine einzelne Pflanze produziert, je nach Standort und Bedingungen, zwischen 5 und 750 Achänen (VON LAMPE, M., 1996). Die Ausbreitung der Samen erfolgt sowohl anemochor, hydrochor und zoochor als auch durch menschliche Aktivitäten und Gerätschaften (VON LAMPE, M., 1996; HOHENSEE & FREY, 2001). Sind die Umweltbedingungen günstig, kann *Cyperus fuscus* in kurzer Zeit aus dem Boden keimen und innerhalb von 1 bis 2 Monaten die Blüte erreichen. Die Früchte reifen innerhalb von 2 bis 3 Monate nach der Keimung (VON LAMPE, M., 1996). Chromosomenzählungen ergaben die Chromosomenzahlen $2n = 36$ und $2n = 72$, die vermutlich diploiden und tetraploiden Pflanzen entsprechen (MURIN & MÁJOVSKÝ, 1987; KRAHULCOVÁ, 2003; ROALSON, 2008). Aufgrund der durchschnittlich sehr hohen Samenproduktion, der biologischen Eigenschaften und der weiten Verbreitung ist *Cyperus fuscus* als Stellvertreter der Zwergbinsengesellschaften ein geeignetes Objekt für populationsgenetische Untersuchungen. Zum Zeitpunkt des Verfassens dieser Arbeit waren keine molekulare Marker für *Cyperus fuscus* bekannt (NCBI, Stand 29.06.2015).



2.3 Geprüfte Hypothesen

Mit Hilfe der 21 in dieser Arbeit entwickelten Mikrosatelliten-Marker für *Cyperus fuscus* und zur Bestätigung deren Funktionsfähigkeit sollen, wie bereits in der Einleitung geschildert, zwei Hypothesen überprüft und die Frage nach dem Ploidiegrad der Art beantwortet werden. Die Hypothesen beziehen sich auf die genetische Variation der Populationen an verschiedenen Standorten mit verschiedenen Nutzungstypen. Ebenso soll die genetische Differenzierung zwischen den Bodenfraktionen und der oberirdischen Population verglichen werden, um Aussagen über die Funktion der Samenbank im Boden für die Anpassungsfähigkeit der Art an unterschiedliche Umweltbedingungen zu treffen.

Die entsprechenden Hypothesen lauten:

Hypothese 1: Die genetische Vielfalt wird mit steigender Ökosystem-Störung erhöht.

Aus den Annahmen der Hypothese 1 wird erwartet, dass die natürliche Population aufgrund diverser Faktoren wie beispielsweise saisonal stark schwankende Wasserstände, Temperaturunterschiede, periodische Austrocknung und andere Umwelteinflüsse eine größere genetische Vielfalt aufweist als die in vergleichsweise homogenen Verhältnissen lebende Population des Fischhälters. Zur Beantwortung dieser Hypothese bedarf es der Ermittlung der Heterozygotiewerte, der *allelic richness* und des Inzuchtkoeffizienten. Entsprechend der Hypothese sollten somit die Heterozygotiewerte sowie die *allelic richness* der Natürlichen Population höher sein als jene der Fischhälter-Population. Umgekehrt sollte der Inzuchtkoeffizient bei der Fischhälter-Population höher sein. Es wird aufgrund der relativen Isolation der Population erwartet, dass ein geringerer genetischer Austausch über diverse Ausbreitungsfaktoren (beispielsweise hydrochore Ausbreitung und Austausch über Flusswasser) stattfindet.

Hypothese 2: Die Genotypen an der Oberfläche sind nicht zufällig entstanden sondern sind an die vorherrschenden Lebensraumbedingungen angepasste Genotypen aus der Samenbank im Boden.

In diesem Fall wird erwartet, dass die Genotypenzusammensetzung der Oberflächenpopulation sich deutlich von der Genotypenzusammensetzung der Samen aus den Bodenschichten unterscheidet. Ist dies der Fall, so können Rückschlüsse darauf getätigt werden, dass die aktuelle Oberflächenpopulation nicht durch Zufall entstanden ist sondern aus Genotypen besteht, die an die momentanen Lebensraumbedingungen angepasst sind. Dabei sollten nicht nur Unterschiede in den Subpopulationen sondern ebenso Unterschiede zwischen den Populationen selbst sichtbar werden und zeigen, dass die Population im Fischhälter, aufgrund der eher homogenen Lebensraumbedingungen, aus entsprechend angepassten Genotypen aus der Samenbank im Boden hervorgegangen ist, während in der natürlichen Population die Genotypenunterschiede weniger eindeutig in Erscheinung treten. Dies kann sowohl auf die heterogenen Lebensraumbedingungen als auch auf den stärkeren genetischen Austausch mit anderen Populationen aufgrund verschiedener Ausbreitungsmechanismen zurückzuführen sein. Zur Beantwortung dieser Hypothese bedarf es der Berechnung des F_{ST} -Wertes oder *fixation index* zur Quantifizierung der genetischen Differenzierung von Subpopulationen.

Ploidiegrad:

Der Ploidiegrad von *Cyperus fuscus* konnte bisher nicht mit Sicherheit festgestellt werden. Eine Chromosomenzählung in der Slowakei im Jahr 1987 (MURIN & MÁJOVSKÝ, 1987) ergab eine Chromosomenzahl von $2n=72$. In ROALSON 2008 wurden zahlreiche Chromosomenzählungen der Gattung *Cyperus* aufgelistet. Dabei sind Chromosomenzahlen von 18, 36 und 54 besonders häufig aufgetreten (Abbildung 2) (ROALSON, 2008).

Eine weitere Chromosomenzählung im Jahr 2014 die an sieben Populationen aus Österreich, Tschechien, Ungarn und Slowakei durchgeführt wurde konnte die Chromosomenzahl eindeutig auf $2n=36$ festlegen (K. TREMETSBERGER, mündliche Mitteilung) Der Ploidiegrad konnte jedoch ohne genetische Markeranalysen nicht bestimmt werden. Durch die daraus resultierende Anzahl an Allele pro Locus und Individuum soll diese Frage eindeutig beantwortet werden. Es wird erwartet, dass einheitlich maximal 2 Allele je Locus auftreten und die Art somit als diploid zu bezeichnen ist.

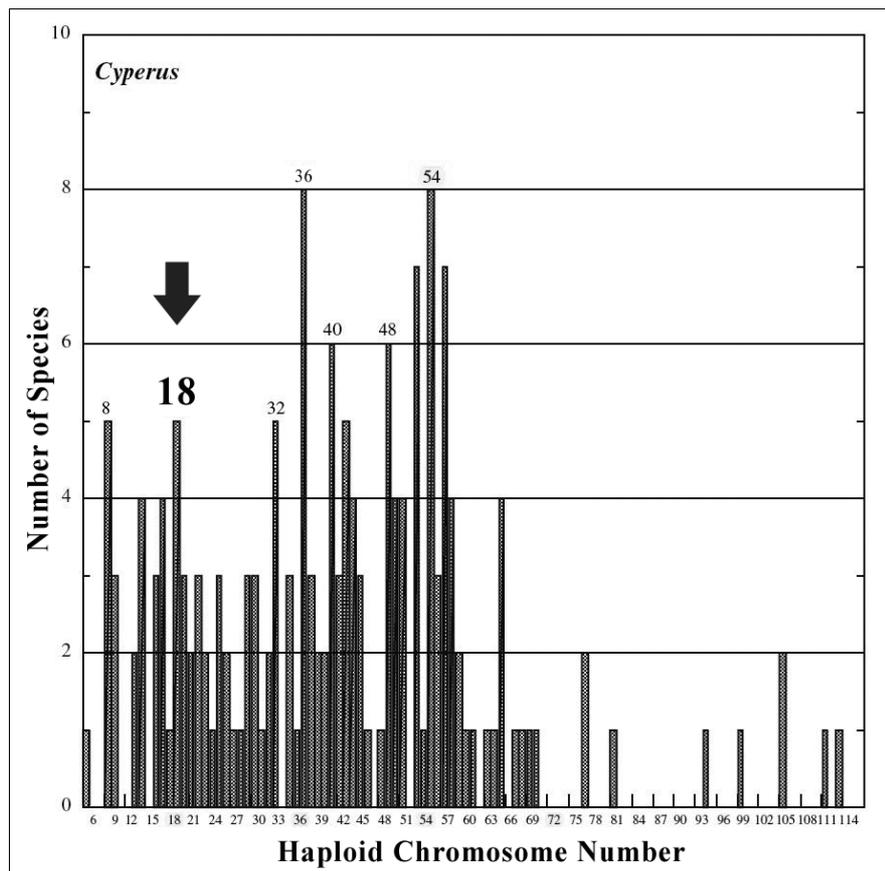


Abbildung 2: Chromosomenzählung der Gattung *Cyperus*. Modifiziert aus ROALSON 2008 (Haploid Chromosome Number variation in *Cyperus*)

3 Grundlagen

In diesem Punkt werden die zugrunde liegenden Prinzipien und Verfahren zur Erstellung spezifischer Mikrosatelliten-Marker für die Art *Cyperus fuscus* geschildert. Dabei werden zunächst die Mikrosatelliten und deren Verwendung im Allgemeinen besprochen um in weiterer Folge die Nutzung als Marker für populationsgenetische Untersuchungen aufzuzeigen. Anschließend werden das Erstellen von Primer und die dabei zu beachtenden Aspekte detailliert beschrieben. Ebenso werden auch die Testverfahren, mit denen die erstellten Primer getestet wurden, vorgestellt. Als erstes wird die Methode zur DNA in-vitro Vermehrung, die Polymerase Kettenreaktion (PCR), vorgestellt. Weiters werden die verwendeten analytischen Methoden zur Identifikation funktionierender Primerpaare, die Gel-Elektrophorese einerseits und die wesentlich genauere Kapillarelektrophorese andererseits, geschildert.

Die in dieser Arbeit konkret verwendeten Methoden zur Entwicklung von Mikrosatelliten-Marker werden im Punkt 4 ausführlich beschrieben.

3.1 Mikrosatelliten in der Populationsgenetik

Mikrosatelliten oder auch *simple sequence repeats* (SSR's) genannt, sind DNA Abschnitte bestehend aus hintereinander gereihten sich wiederholenden Einheiten von 1 bis 6 Basenpaaren (GOLDSTEIN & SCHLÖTTERER, 2000; FREELAND, 2005) und einer Gesamtlänge von maximal 100 Basenpaaren, die von längeren, einzigartigen Sequenzabschnitten flankiert werden. (BEEBEE & ROWE, 2008) Respektive werden diese Wiederholungen Mono-, Di-, Tri-, Tetra-, Penta- und Hexanukleotide genannt. Ein Beispiel eines Mikrosatelliten ist die dinukleotide Wiederholung (GT)₈, die aus 8 Wiederholungen der Einheit GT besteht (GTGTGTGTGTGTGTGT). In diesem Fall würde der komplementäre Abschnitt der DNA den Mikrosatelliten (AC)₈ darstellen (ACACACACACACACAC). (FREELAND, 2005) Mikrosatelliten kommen im Genom jedes eukaryotischen und prokaryotischen Lebewesen häufig vor, darin bevorzugt in nicht kodierenden Bereichen aber auch regelmäßig in *Promoter* Regionen kodierender Sequenzen. (FIELD & WILLS, 1996; GOLDSTEIN & SCHLÖTTERER, 2000; RUDOLPH, 2001; SCHLÖTTERER, 2004; BEEBEE & ROWE, 2008)

SSR's selbst werden zur Klasse von Sequenzen mit der Bezeichnung *variable number of tandem repeats* (VNTR) gezählt. (ELLENGREN, 2000)

Die Tatsache, dass SSR's praktisch in jedem Genom häufig und relativ gleichmäßig verteilt vorkommen sowie ihre unüblich hohe Polymorphismus-Rate macht sie zu attraktiven Kandidaten als genetische Marker. (SCHLÖTTERER, 2004)

3.1.1 Entstehung und Entdeckungsgeschichte von Mikrosatelliten

Wiederholungen mit längeren Einheiten als Mikrosatelliten werden Minisatelliten und die Steigerung davon Satelliten DNA genannt. Diese Satelliten DNA wurde erstmals 1960 beobachtet, als nach einer Dichtegradientenzentrifugation von Fragmenten genomischer DNA neben der Hauptbande mehrere, damals als *'satellite peaks'* bezeichnete Nebenbanden sichtbar wurden. Diese Nebenbanden wurden aus langen, sich wiederholenden Sequenzen gebildet. Als nachfolgend auch kürzere Wiederholungen entdeckt wurden, wurden diese respektive Minisatelliten und Mikrosatelliten genannt. (ELLENGREN, 2004; RUDOLPH, 2001)

3.1.2 Mutationsmechanismus - Entstehung von Allelen

Zu den Mutationsmechanismen von Mikrosatelliten gibt es mehrere Möglichkeiten. Zum einen die Schlaufenbildung bei der DNA Replikation durch die DNA-Polymerase (LEVINSON & GUTMAN, 1987) und zum anderen durch ungleiches *crossing-over* bei der chromosomalen Rekombination durch. Letzteres ist zwar theoretisch und für längere Sequenzen wie Minisatelliten möglich jedoch für Mikrosatelliten von untergeordneter Bedeutung. (GOLDSTEIN & SCHLÖTTERER, 2000; ELLEGREN, 2000). Die Schlaufenbildung bei der DNA-Replikation gilt daher allgemein als der grundlegende Mechanismus von Mutationen bei Mikrosatelliten. (ELLENGREN, 2000). Bei der DNA-Replikation kann es aufgrund der besonderen Struktur von Mikrosatelliten vermehrt zu einem sogenannten *slippage* oder *slipped-strand mispairing (SSM)* (LEVINSON & GUTMAN 1987), auf Deutsch *"Versetzungs Fehlpaarung"* kommen. Für nicht-repetitive Sequenzabschnitte stellt diese Fehlpaarung kein Problem dar, weil die komplementären Stränge sich nur auf eine Art und Weise wieder aneinander angliedern können. Bei repetitiven Sequenzen wie Mikrosatelliten kann es jedoch leicht vorkommen, dass der neugebildete Folgestrang sich phasenverschoben an den Leitstrang anheftet. (GOLDSTEIN & SCHLÖTTERER, 2000). Durch diese lückenhafte Wiederanordnung der komplementären Stränge entstehen Auswölbungen beziehungsweise Schleifen, sogenannte *hairpins*. (RICHARDS & SUTHERLAND, 1994). Wenn dabei Auswölbungen im Leitstrang entstehen, wird der Folgestrang kürzer sein als der Leitstrang und hingegen länger, wenn die Auswölbungen im Folgestrang entstehen (GOLDSTEIN & SCHLÖTTERER, 2000; ELLEGREN, 2000). (Abbildung 3) Solche Mutationen werden *insertion-deletion* Mutationen oder kurz *InDel's* bezeichnet. Diese entstehen, wenn die gebildeten Schleifen von den DNA Reparaturmechanismen aus der Sequenz entfernt werden. In den meisten Fällen wird dabei genau eine Wiederholungseinheit hinzugefügt beziehungsweise entfernt. Seltener sind mehr als

eine Einheit betroffen. Der Mutationsmechanismus von Mikrosatelliten folgt demnach dem *Stepwise Mutation Model* von Kimura und Ohta (KIMURA & OHTA, 1978).

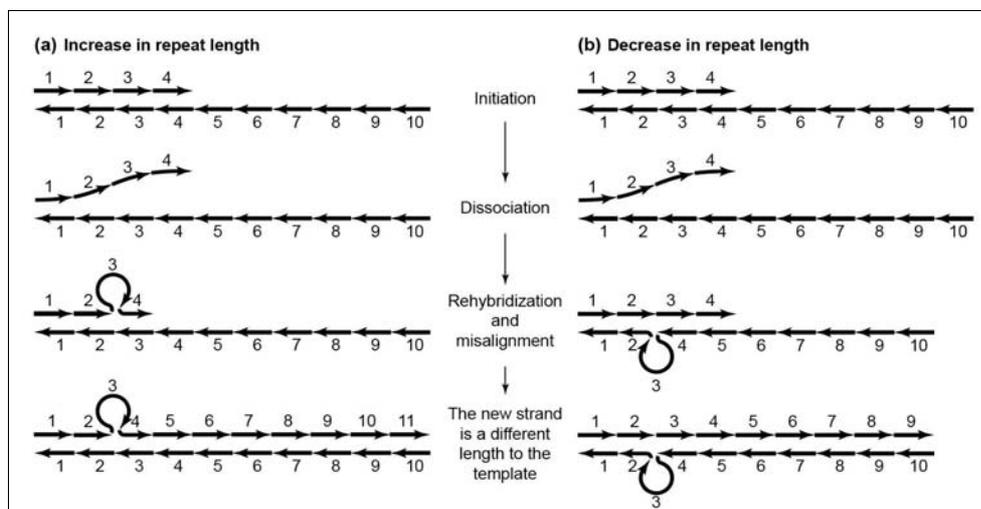


Abbildung 3: Schleifenbildung und Längenänderung der DNA stränge. (ELLEGRÉN, 2000; "Figure 1: Model of microsatellite mutation by replication slippage")

3.1.3 Mikrosatelliten als molekulare Marker

Mit dem Aufkommen neuer Technologien wie beispielsweise die 454 Pyrosequenzierung und frei zugänglicher Software können Sequenzdaten zu erschwinglichen Preisen erhalten und Marker relativ kostengünstig entwickelt werden. Dazu ist nur eine geringe Menge DNA nötig, die mit ebenso weitgehend unkomplizierten Labortechniken und modernen Gerätschaften in kurzer Zeit entnommen werden kann. Mikrosatelliten werden daher häufig als Marker eingesetzt. Wie alle co-dominanten, *single-locus* Marker liefern Mikrosatelliten Informationen über die Variation von Allelen eines bestimmten Locus (SCHLÖTTERER, 2004). Die Aussagekraft eines Locus ist jedoch begrenzt und es werden daher viele nicht miteinander verlinkte Loci benötigt, um eine ausreichende statistische Auflösung für Verwandtschaftsverhältnisse zu gewährleisten (BLOUIN et al., 1996). Durch die Entwicklung der Polymerase Kettenreaktion durch Kary Mullis im Jahre 1983 (MULLIS et al., 1986) konnten sich DNA-Basierte Molekulare Marker erst durchsetzen (siehe 3.3.1) Erstmals war es möglich mithilfe von oligonukleotiden Primerpaaren (siehe 3.1.3.1) jede beliebige Stelle im Genom zu amplifizieren ohne große Mengen an genomischer DNA bereitstellen zu müssen oder diese aufwendig in Vektoren zu klonen (SCHLÖTTERER, 2004). Die Vervielfältigung von Mikrosatelliten mittels PCR ist generell einfach und relativ zuverlässig. Es entstehen jedoch häufig aufgrund von Versetzungs-Fehlpaarungen wäh-

rend der PCR Amplifikation unerwünschte zusätzliche Produkte mit eigenen Banden, sogenannte *stutter bands*. Dabei entstehen Banden, die ein oder mehrere Wiederholungseinheiten kürzer sind als das untersuchte Allel. In seltenen Fällen können auch längere Banden entstehen. *Stutter bands* erschweren vor allem die Analyse von gemischten DNA-Proben. Tetranukleotide und generell längere Wiederholungseinheiten haben sich dabei als weniger anfällig für die Bildung von *stutter bands* erwiesen als beispielsweise kürzere dinukleotide. Es ist dabei anzumerken, dass dinukleotide Wiederholungen wiederum eine höhere Mutationsrate aufweisen und somit auch polymorpher sind als beispielsweise tetranukleotide Wiederholungen (GOLDSTEIN & SCHLÖTTERER, 2000).

3.1.3.1 Primer

Um Mikrosatelliten mittels PCR als molekulare Marker verwenden zu können bedarf es sogenannter Primer. Diese sind meistens 18 bis 30 Basenpaare lange, einzelsträngige DNA-Oligonukleotide die der DNA-Polymerase als Startpunkt für die DNA-Replikation dienen. Für die PCR werden Primer mit einer bestimmten Sequenz synthetisch hergestellt. Die Sequenz ist dabei ein „Spiegelbild“ jener Sequenz, die den zu amplifizierenden DNA-Abschnitt flankiert. Diese Primer binden somit an genau definierten Stellen im Genom und es kann auf diese Weise der DNA Abschnitt, der amplifiziert werden soll, wie beispielsweise ein Mikrosatellit, festgelegt werden (CHEMIE.DE, 2015).

3.1.3.2 Dominante vs. Co-dominante Marker

Gemäß den mendelschen Vererbungsgesetzen werden Merkmalsausbildungen diploider Organismen jeweils durch zwei homologe Genkopien beeinflusst. Individuen, bei denen die beiden homologen Gene gleichartig ausgebildet sind bzw. die, die selben Allele (=Genvarianten) aufweisen, werden homozygot, oder reinerbig, genannt. Im Gegensatz dazu werden Individuen mit unterschiedlichen Allelen als heterozygot, oder mischerbig, bezeichnet. Bei heterozygoten Individuen können sich die beiden Allele gleichermaßen in der Merkmalsausprägung durchsetzen und somit *intermediäres* Verhalten aufweisen, oder ein Allel kann die Wirkung des Anderen verdecken und somit *dominant* gegenüber dem *rezessiven* Allel sein. (SITTE et al., 1998) Diese Allelzustände sind nicht auf diploide Organismen beschränkt sondern treten bei Organismen höherer Ploidiestufen ebenso auf. (GOLDSTEIN & SCHLÖTTERER, 2000)

Entsprechend den Allelzuständen gibt es sowohl dominante als auch codominante Marker. Die wichtigsten Vertreter der dominanten Marker sind AFLP's (*Amplified Fragment Length Polymorphism*), RFLP's (*Restriction*

Fragment Length Polymorphism), RAPD's (*Randomly Amplified Polymorphic DNA's*) und ISSR's (*Inter Simple Sequence repeats*). Diese Marker basieren ebenso wie Mikrosatelliten auf Längenunterschiede von PCR-amplifizierten Teilen des Genoms, jedoch werden dabei unspezifische Primer verwendet, die an mehreren Stellen im Genom binden können. Der große Vorteil dieser Marker ist, dass keine a-priori Information benötigt wird. Diese Tatsache macht solche Marker attraktiv für Arten mit großen beziehungsweise wenig untersuchten Genomen. (SCHLÖTTERER, 2004)

Bei dominanten Markern sind jedoch nur die dominanten Allele sichtbar, rezessive Allele können nicht sichtbar gemacht werden, deshalb können heterozygote Individuen nicht von homozygoten Individuen unterschieden werden. Um auch heterozygot-rezessive Allele sichtbar zu machen, kann auf codominante Marker zurückgegriffen werden. Bei diesen sind, mit Ausnahme von sogenannten Nullallelen, alle Allele eines Genotyps sichtbar. Codominante Marker sind neben Mikrosatelliten auch SNP's (*Single Nucleotide Polymorphisms*).

3.1.3.3 Mikrosatelliten-Marker für Populationsgenetische Untersuchungen

Die Vorteile von Mikrosatelliten im Vergleich zu anderen molekularen Markern wie beispielsweise AFLP's oder RFLP's sind die hohe Anzahl an Allele und der hohe Heterozygotiegrad, Co-dominanz sowie die leichte Isolierung. Ebenso weisen Mikrosatelliten eine hohe Mutationsrate auf, die es zwar ermöglicht genetische Variation innerhalb der Population zu identifizieren aber gleichzeitig oft nur innerartliche Vergleiche ermöglicht. Als nachteilig zu bezeichnen sind die komplexen Mutationsmechanismen. Ebenso sind Mikrosatellitenanalysen auch aufgrund von *stutterbands* nicht leicht zu automatisieren und bedürfen daher eines höheren Arbeitsaufwandes. Auch die Anzahl an verwendbaren Mikrosatelliten kann unter Umständen zu gering sein, um Analysen, vor allem bei großen Populationen, sinnvoll zu gestalten. (SCHLÖTTERER, 2004)

3.2 Suche nach Mikrosatelliten und Primerdesign

Die Isolation von Mikrosatelliten beinhaltet die teilweise Sequenzierung des Genoms und das automatisierte Durchsuchen der dadurch erhaltenen Sequenzdaten auf die gewünschten Motive. Zunächst müssen dabei geeignete, repräsentative und vor allem gesunde Pflanzenindividuen gewählt werden, dessen DNA extrahiert und mittels Pyrosequenzierung, wie beispielsweise 454 Pyrosequenzierung von *Life sciences* (einem Roche- Unternehmen, Branford, Connecticut, USA), sequenziert werden. Das Primerdesign erfolgt danach entweder mithilfe von kommerzieller oder frei zugänglicher Software oder kann auch direkt von Betrieben durchgeführt werden. In diesem Projekt wurden zwei Methoden angewandt. Zum einen die Isolation von Mikrosatelliten und das Design von Primern mithilfe von online frei erhältlichen Freeware Programmen, nämlich MSATCOMMANDER (FAIRCLOTH, 2008) und PRIMER 3 (UNTERGASSER et al., 2012), und zum anderen die Isolation (inklusive Enrichment, siehe 3.2.2.2) mit anschließendem Primerdesign durch die Schweizer Firma ECOGENICS (Balgach, Schweiz).

3.2.1 DNA-Extraktion

Die DNA Extraktion erfolgte mit dem *DNeasy Plant Mini Kit* (QIAGEN, Valencia, California, USA). Dieses Verfahren wird häufig genutzt, um DNA aus Pflanzenmaterial zu erhalten (vgl. SHAMBLIN et al., 2007; CSENCISICS et al., 2010; BUEHLER et al., 2011 etc.) Zu beachten ist dabei, dass neben genomischer DNA auch RNA, Plastiden DNA etc. extrahiert werden. Dabei werden die frischen oder vorher getrockneten Pflanzenteile zunächst mittels einer *Retsch* Mühle mechanisch zermahlen und anschließend im *QIAcube* chemisch lysiert. Der *QIAcube* ist eine Roboter-Arbeitsstation, die DNA-Aufreinigung vollautomatisch durchführen kann. (QIAGEN, 2014). Im *QIAcube* wird die DNA isoliert, indem Zellteile, Proteine und Polysaccharide entfernt werden sowie auch RNA durch RNAsen beseitigt wird. Nach dem Homogenisieren der Probe durch die integrierte Zentrifuge wird die DNA an eine spezielle Siliciumdioxid-Membran gebunden und erneut kurz zentrifugiert, wobei alle nicht gebundenen und somit unerwünschten Bestandteile ungehindert durch die Membran entfernt werden. Nach einem abschließenden Waschprozess um überflüssige Chemikalien und Schmutzpartikel zu entfernen, wird die gewonnene DNA in einer Wasserlösung gelöst und bereitgestellt (QIAGEN, 2014). Als nächster Schritt muss die somit gewonnene genomische DNA sequenziert werden.

3.2.2 Sequenzdaten aus genomischer DNA

Das erste Ziel der meisten Genom-basierten Projekte ist es, die DNA Sequenz des Genoms oder einer großen Anzahl an Sequenzabschnitten zu ermitteln (GIBSON & MUSE, 2002). Auch für die Entwicklung spezifischer Primer ist das Wissen um die Sequenzen der Mikrosatelliten und der flankierenden Abschnitten unerlässlich. Um diese Daten zu erhalten können verschiedene Sequenziermethoden, wie die hier verwendete *454 Pyrosequenzierung* angewandt werden. Im nachfolgenden Punkt werden die Grundlagen dieses Verfahren geschildert.

3.2.2.1 Pyrosequenzierung

1996 entwickelten Pål Nyrén und Mostafa Ronaghi aus Schweden die Pyrosequenzierung (PFLANZENFORSCHUNG.DE, 2014). ein Verfahren zur DNA Sequenzierung, bei der während der DNA Synthese frei werdende Pyrophosphate registriert werden. In einer enzymatischen Kettenreaktion wird dabei sichtbares Licht generiert, das der Anzahl an eingebundenen Nukleotiden proportional ist. Bei der Einbindung eines Nukleotides wird nämlich inorganisches Pyrophosphat frei, das daraufhin durch ATP-Sulfurylase in ATP umgewandelt wird. Dieses ATP liefert dem Enzym Luziferase wiederum die nötige Energie um Luziferin in Oxyluziferin zu oxidieren und dabei ein Lichtsignal zu emittieren. Nachdem das der Reaktion zugeführte Nukleotid bekannt ist, kann die Sequenz ermittelt werden (RONAGHI, 2001). Die erste kommerzielle *Next Generation Sequencing (NGS)* Plattform, die mit Pyrosequenzierung arbeitet, ist die der Firma *454 Life Sciences* von *ROCHE* (Roche diagnostics, Basel, Schweiz). Diese Technologie bietet die längsten *reads*, (Stand: Juli 2015), also DNA-Abschnitte, unter den existierenden *NGS* Methoden (ABBOTT et al., 2010) und wird darum häufig für die Entwicklung von Mikrosatelliten-Marker verwendet (vgl. ABBOTT et al., 2010; CSENC-SICS et al., 2010; BUEHLER et al., 2011; TAKAYAMA et al., 2011).

454 Pyrosequenzierung

Bei der 454 Pyrosequenzierung werden zunächst durch das *Shotgun*-Verfahren zufällige Fragmente aus genomischer DNA gewonnen. Der Name bezieht sich auf die zufällige Streuung der entnommenen Fragmente im Genom, ähnlich der Streuung eines Schrotgewehrs. Die somit erhaltenen einzigartigen Fragmente werden abermals zerteilt und mit Adaptern ausgestattet, mit denen sie anschließend an winzige Perlen, den sogenannten *template-carrying Beads*, gebunden werden. An jede dieser 28 μ m großen Perlen wird ein einziges mit Adaptern ausgestattetes, DNA-Molekül an einem Ende kovalent gebunden. Diese werden anschließend von Tröpfchen einer *PCR-reaction-mixture-in-oil-emulsion* umschlossen (MARGULIES et al., 2005).

Konkret werden dabei die einzelnen Perlen in wässrige Tröpfchen, in denen die benötigten PCR Komponenten enthalten sind, eingeschlossen und durch eine öligen Substanz voneinander getrennt. Auf diese Weise können in einem Volumen von nur einem Mikroliter tausende PCR-Einzelreaktionen gleichzeitig stattfinden (GLENN, 2011). Als Ergebnis erhält man *DNA-carrying-beads*, also Perlen, an denen jeweils Millionen Kopien des Ausgangstranges haften (MARGULIES et al., 2005). Nach dem Auflösen der Emulsion werden Millionen dieser Perlen auf eine spezielle *picotitre-plate* geladen. Diese ist eine Scheibe aus Glasfaser mit ca. 1,6 Millionen winziger, $55 \mu\text{m}$ tiefen und $44 \mu\text{m}$ breiten, Vertiefungen (LEAMON et al., 2003). In jede dieser Vertiefungen lagert sich im Idealfall genau eine *DNA-carrying-bead* ab. Zusätzlich werden nun zahlreiche kleinere Perlen in die einzelnen Vertiefungen hinzugefügt, die für die Pyrosequenzier-Reaktion notwendige Enzyme transportieren. Nach dieser Vorbereitung wird die Platte senkrecht in die Durchflusskammer der Sequenziermaschine mit der Plattenrückseite an einem CCD Sensor positioniert. Die Durchflusskammer ermöglicht die zyklische Zugabe der einzelnen Nukleotide in definierter Abfolge. Diese werden von oben in die Kammer gepumpt und fließen über die Platte mit den Vertiefungen nach unten. Dadurch wird eine rasche Verteilung der Nukleotide auf die einzelnen Vertiefungen gewährleistet. Werden nun Nukleotide durch die Polymerase an die bestehenden DNA-Stränge auf einer Perle erfolgreich angefügt, entsteht, durch die zu Beginn erwähnte Enzymkette, ein Lichtblitz der vom Sensor registriert wird. Nachdem die Abfolge der Nukleotide, die durch die Kammer fließen, bekannt ist, kann die Sequenz der Stränge auf den Perlen auf diese Weise ermittelt werden (MARGULIES et al., 2005). Als Output dieses Prozesses wird eine Datei im *FASTA*-Format generiert, die alle Sequenzinformationen der Fragmente enthält (PEARSON & LIPMAN, 1988). Heute können bei einem Durchlauf über 1 Million *reads* mit *read*-Längen von bis zu 1000 Basenpaaren erstellt werden (Roche Diagnostics GmbH, 2014). Dabei lassen sich hunderte bis tausende Mikrosatelliten Loci aus einem Bruchteil eines Pyrosequenzier-Laufs identifizieren (vgl. CASTOE et al., 2010; BOOMER & STOW, 2010; SAARINEN & AUSTIN, 2010). 85% aller *reads* besitzen dabei eine Länge von mindestens 500 Basenpaaren (ROCHE DIAGNOSTICS, 2014).

3.2.2.2 *Enrichment* genomischer DNA

Beim *Enrichment* werden DNA Fragmente angereichert, die bestimmte Repeat-Motive enthalten. Es gibt dazu zahlreiche Protokolle die zum Teil teuer und aufwendig sind sowie ein erhebliches Maß an Erfahrung von Seiten der durchführenden Nutzer erfordern (vgl. ZANE et al., 2002; SANTANA et al.,

2009; u.v.m.). Prinzipiell ist eine Vielfalt an Motiven wünschenswert, da unterschiedliche Motive in unterschiedlichen Genomregionen konzentriert sind, zum Beispiel ist AG bei weitem das häufigste Motiv bei Pflanzen generell und vor allem in transkribierten Regionen zu finden, genauso wie AAG. Letzteres findet man auch in kodierenden Regionen, da der Leserahmen nicht gestört wird (MORGANTE et al., 2002). Aus diesem Grund wurde das *Enrichment* von der Firma Ecogenics durchgeführt. Das *Enrichment* erfolgt dabei in mehreren Schritten. Zu Beginn wird die genomische DNA mittels Restriktionsenzymen in Fragmente von ungefähr 200 bis 1000 Basenpaaren Länge zerteilt. Die erhaltenen Fragmente werden anschließend an den stumpfen Enden (*blunt ends*) zunächst adenyliert und mit Adaptern, die als Primer dienen, ausgestattet. Nach einem Amplifikationsschritt mittels PCR werden Einzelstränge dieser Fragmente mit *biotinylierten* Oligos hybridisiert. Die Oligos sind dabei so designt, dass sie den gesuchten Repeat-Motiven entsprechen. Wenn also Fragmente mit den Oligos hybridisieren, steigt die Wahrscheinlichkeit, dass diese gesuchte SSR Motive enthalten. Auf diese Weise können bis zu 75% und mehr aller hybridisierten Fragmente SSR Motive enthalten (TECHEN et al., 2010). Im nächsten Schritt werden die hybridisierten Fragmente von den übrigen getrennt. Dies geschieht mit der Hilfe von magnetischen Streptavidin-Perlen. Das Biotin-Label der Oligos geht mit diesen Perlen eine Bindung ein. Im nachfolgenden Zentrifugations- und Waschschrift werden die Fragmente mit den magnetischen Perlen isoliert und die erhaltenen Einzelstränge eluiert. Als letztes erzeugt eine erneute PCR die ausreichende Menge der somit angereicherten Fragmente (TECHEN et al., 2010) (Abbildung 4). Es folgt darauf die 454 Pyrosequenzierung der Fragmente.

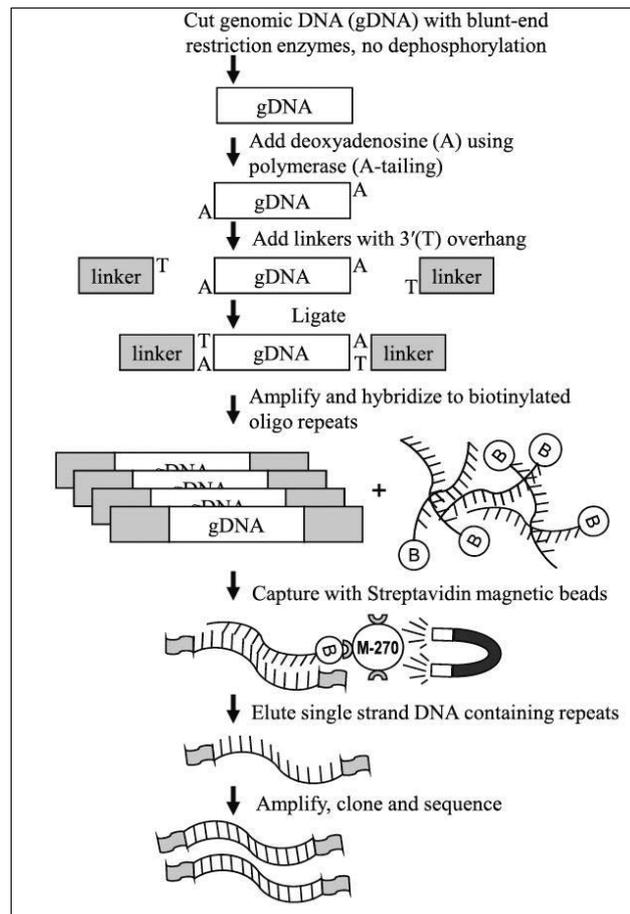


Abbildung 4: Übersicht der Schritte bei SSR-enrichment. (TECHEN, 2010; "Overview of genomic DNA SSR-enriched library protocol")

3.2.3 Grundlagen zur Isolation von Mikrosatelliten und Primerdesign

Die erhaltenen Daten der Pyrosequenzierung im *FASTA*-Format werden in weiterer Folge herangezogen, um passende SSR-Motive innerhalb der Sequenzdaten zu finden. Das Design der Primer selbst wird nach der Festlegung einiger Parameter, wie beispielsweise die Mindestanzahl an Wiederholungen oder die Größe der Wiederholungseinheiten, von geeigneter Software durchgeführt. In dieser Arbeit werden zwei Methoden eingesetzt. Zum einen die Isolation von Mikrosatelliten mittels MSATCOMMANDER (FAIRCLOTH, 2008) und Primerdesign mittels PRIMER3 (ROZEN & SKALLETZKY, 2000) und zum anderen die Isolation von Mikrosatelliten und Primerdesign durch ECOGENICS (Balgach, Schweiz). In beiden Methoden wurden sogenannte *Tags* und *PIG-tails* verwendet. In den nachfolgenden Punkten wird zunächst die Bedeutung und Verwendung von *Tags* und *PIG-tails* geschildert. Die Methoden werden im Kapitel 4.1 Methoden im Detail be-

sprochen.

3.2.3.1 Die Verwendung von *tag's* und PIG-tails

Eine Möglichkeit, Primer effektiver zu nutzen, bietet das sogenannte *Primer-Tailing*. Dabei werden Anhängsel aus kurzen DNA-Abschnitten mit bekannter und genau definierter Sequenz, sogenannte *tag's*, an die Primer geheftet. Im Verlauf einer Mikrosatellitenanalyse kann das Verwenden von Tags zu deutlichen Kostensenkungen und Arbeitserleichterungen führen (FAIRCLOTH, 2008). Die Tags werden dabei, wie aus Abbildung 5 ersichtlich, an das 5'-Ende von einem der beiden Primer angeheftet (BOUTIN-GANACHE et al., 2001). Damit ist gewährleistet, dass die eigentlich untersuchte Sequenz zwischen den Primern unverändert bleibt. Der große Vorteil der *tag's* bei Mikrosatellitenanalysen liegt darin, dass alle PCR-Amplifizierten Fragmente trotz unterschiedlicher Primerpaare die selben Endsequenzen besitzen. Werden nun in die Reaktion weitere *tag's* mit gleicher Sequenz hinzugefügt, jedoch mit Fluoreszenzlabeln am 5'Ende, können damit fluoreszenzgelabelte Fragmente für Analysen an Kapillarelektrophoresegeräten mit Fluoreszenzdetektoren oder ähnliche Anwendungen generiert werden (BOUTIN-GANACHE et al., 2001). In dieser Arbeit wurden die *tag's* M13R und CAG aus FAIRCLOTH 2008 und M13 aus SCHUELKE 2000 verwendet. Diese werden in den Punkten 4.1.1.2 und 4.1.2.2 ausführlich besprochen.

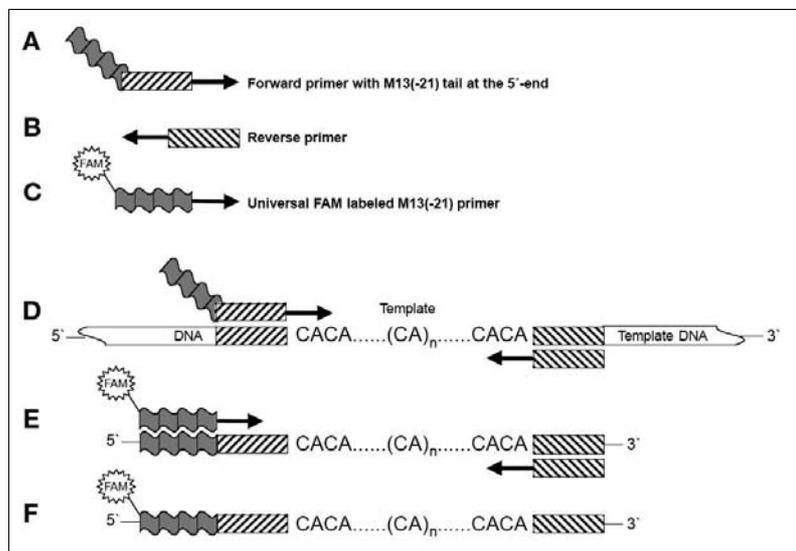


Abbildung 5: Schematische Gliederung der PCR Amplifikation mit Fluoreszenzlabeln. (SCHUELKE, 2000; "Figure 1: Amplification scheme for the one-tubed, single-reaction nested PCR method.")

3.2.3.2 Primer-*PIG-tail*

Eine weitere Möglichkeit Primer mit spezifischen Anhängseln zu optimieren bietet das sogenannte Primer-*PIG-tail*. Unter bestimmten Bedingungen kann die *Taq* DNA-Polymerase an das 3'Ende des zu bildenden Stranges eine Base, bevorzugt Adenin, hinzufügen. (HU et al., 1993) Konkret wird dabei an der letzten Base des Primers des neu-synthetisierten Stranges eine Base angehängt, obwohl diese nicht dem vorhandenen Primer entspricht. Ob eine Base angehängt wird oder nicht hängt hauptsächlich davon ab, welches der vier Nukleotide am Ende des Primers steht. Dabei gilt, dass bei Cytosin als End-Nukleotid bevorzugt eine Base angehängt und bei Adenin als End-Nukleotid meistens keine Base angehängt wird (BROWNSTEIN et al., 1996). Aus dieser Gegebenheit resultieren PCR Produkte unterschiedlicher Länge. Sowohl auf Agarosegelen als auch nach der Auftrennung mittels Kapillarelektrophorese können daraufhin sogenannte +A-Banden oder +A *peaks* erkannt werden. Nachdem dieser Prozess, je nach Struktur der Ausgangssequenz, unterschiedlich verlaufen kann, treten diese Banden unregelmäßig auf und erschweren somit das *scoring* der Allele. BROWNSTEIN et al. (1996) versuchten dieses Problem bei der PCR zu beseitigen. Es wurde versucht, zunächst die Primer-Enden so zu verändern, dass entweder Basen hinzugefügt oder weggelassen wurden. Sie fanden heraus, dass es um einiges schwieriger ist, Primer so zu verändern, dass keine Basen angehängt werden, wohingegen die Anheftung von Basen in den allermeisten Fällen durch relativ einfache und einheitliche Veränderungen der Primer herbeigeführt werden kann. Dabei wird an die Primersequenz am 5'Ende die von den Entwicklern *PIG-tail* genannte Sequenz GTTT angehängt. Dies bewirkt, dass in nahezu 100% der Fälle ein Adenin am 3' Ende des neu-synthetisierten Komplementärstranges hinzugefügt wird. (BROWNSTEIN et al., 1996) *PIG-tails* ermöglichen somit einen Störfaktor beim *Allel-scoring* zu eliminieren und werden in verschiedener Ausführung häufig bei SSR- Analysen eingesetzt (GLENN & SCHABLE, 2005; HILL et al., 2009; ABBOTT et al., 2010; GUICHOUX et al., 2011).

3.2.3.3 Programme zur Isolation von Mikrosatelliten und Primerdesign

In Methode 1 wurden die online frei zugänglichen Programme zur Isolation und zum Primerdesign MSATCOMMANDER (FAIRCLOTH, 2008) und PRIMER3 (ROZEN & SKALETZKY, 2000) verwendet. MSATCOMMANDER ist ein Programm zum Durchsuchen von Sequenzierungsdaten nach SSR Motiven. Es beinhaltet die PRIMER3-*core*. PRIMER3 ist ein Computerprogramm zum Designen von PCR Primer für eine Vielfalt von Anwendungen, darunter auch Mikrosatelliten (ROZEN & SKALETZKY, 2000). In den nachfolgenden Punk-

Auswahl der SSR Art

MSATCOMMANDER ermöglicht das Suchen nach *perfect* und *imperfect repeats* (BIOINFORMATICS.ORG). *Imperfect repeats* beinhalten Mutationen innerhalb der Mikrosatellitensequenz, wie beispielsweise *indels*, die diese unterbrechen. Deshalb ist zu beachten, dass bei *imperfect repeats* die Länge des Fragments von mehreren unterschiedlichen Sequenzen abhängig ist. *Imperfect repeats* müssen nämlich nicht zwangsweise dem *stepwise mutation model* (KIMURA & OHTA, 1978) folgen, während *perfect repeats* verlässlich diesem Modell entsprechend mutieren. (ESTOUP et al., 1995).

Größe der Repeat-units

Die Größe der Wiederholungseinheiten kann im Programm MSATCOMMANDER zwischen mononukleotide- und hexanukleotide-*repeats*, je nach Präferenz und Anforderungen, ausgewählt werden. Beim Verwenden größerer Wiederholungseinheiten als Trinukleotide können generell weniger Loci gleichzeitig, beispielsweise in Multiplex PCR, untersucht werden (GUICHOUX et al., 2011). Bei der Amplifikation von Dinukleotiden und in wesentlich geringerem Ausmaß auch bei Trinukleotiden, können häufig die im Punkt 3.1.3 genannten *stutter-bands* auftreten.

Anzahl Wiederholungen

Es ist nötig SSR's zu wählen, die eine ausreichende Anzahl an Wiederholungseinheiten besitzen um Polymorphismus zu gewährleisten (GUICHOUX et al., 2011). Es ist dabei einerseits zu beachten, dass SSR's mit vielen Wiederholungen eine höhere Mutationsrate aufweisen (KELKAR et al., 2008) und somit eine größere Anzahl an Allelen, aber genauso erhöhtes *stuttering* bewirken (HOFFMANN & AMOS, 2005). Eine weitere Schwierigkeit die sich durch lange *repeats* ergibt, tritt besonders dann auf, wenn die *Read*-Länge, also die Länge der Einzelsequenz die sequenziert wurde, eher gering ist. Bei langen SSR's kann es dabei geschehen, dass die flankierenden Regionen des SSR's, die *flanking regions* zu kurz sind, um geeignete Primer darauf zu designen (BOOMER & STOW, 2010). Es gilt demnach eine Anzahl an Wiederholungseinheiten zu wählen um die Vorteile der SSR's optimal zu nutzen und ihre Schwächen möglichst zu vermeiden.

Flanking-regions auf Einzigartigkeit überprüfen - NCBI Blast

Um die Locus-Spezifität der designten Primer zu gewährleisten und eventuelle Übereinstimmung mit bereits bekannten Genomabschnitten anderer Arten festzustellen (MADDEN, 2013) werden diese mittels dem *Basic Local Alignment Search Tool* "BLAST" (ALTSCHUL et al., 1990) mit der NCBI Nukleotid online-Datenbank abgeglichen.

3.2.3.3.2 Primerdesign mittels PRIMER 3

Das Primerdesign mittels PRIMER3 (ROZEN & SKALETZKY, 2000) erfolgt, bei entsprechender Auswahl im Menü von MSATCOMMANDER, automatisch. Dabei werden vordefinierte Parameter genutzt. Diese Parameter sind in der in MSATCOMMANDER integrierten Version von PRIMER3 nicht direkt modifizierbar.

Parameter für das Primerdesign

Primerpaare werden grundlegend nur dann in Betracht gezogen, wenn diese mindestens 10 Basenpaare vom Anfang und Ende der Mikrosatellitensequenz entfernt sind. Dies geschieht unter anderem aus dem Grund, dass es in der unmittelbaren Umgebung der SSR Motive vermehrt zu *indels* oder *SNP's* kommen kann und der Primer an dieser Stelle, aufgrund der veränderten Sequenz, nicht mehr bindet (MOGG et al., 2002). PRIMER3 versucht standardmäßig Primer zu entwickeln, die ein Amplikon von 100 bis 500 Basenpaaren Länge amplifizieren, eine optimale Schmelztemperatur von T_M 60,0° C besitzen (in einem Bereich von 57° C bis 62° C) und untereinander eine Abweichung in der T_M von maximal 5° C aufweisen, einen optimalen GC-Gehalt von 50% sowie mindestens ein Basenpaar *GC-clamp* besitzen (FAIRCLOTH, 2008). Dabei sucht PRIMER3 Primerpaare, die den genannten Optimalwerten am nächsten kommen und gleichzeitig eine ausgewogene Balance zwischen Primer Länge und Schmelztemperatur beibehalten (ROZEN & SKALETZKY, 2000). Eine weitere wichtige Funktion des Programms ist die Möglichkeit, durch thermodynamische Modelle nicht nur die Schmelztemperatur der Primer zu berechnen, sondern auch die Wahrscheinlichkeit, dass die Primer einer Primerkombination unerwünschte Sekundärstrukturen bilden beziehungsweise miteinander wechselwirken. Diese Sekundärstrukturen können die bereits erwähnten *Hairpins* sein, aber auch unspezifisches Binden der Primer an der falschen Stelle sowie das gegenseitige Binden der Primer selbst (Abbildung 7) (UNTERGASSER et al., 2012).

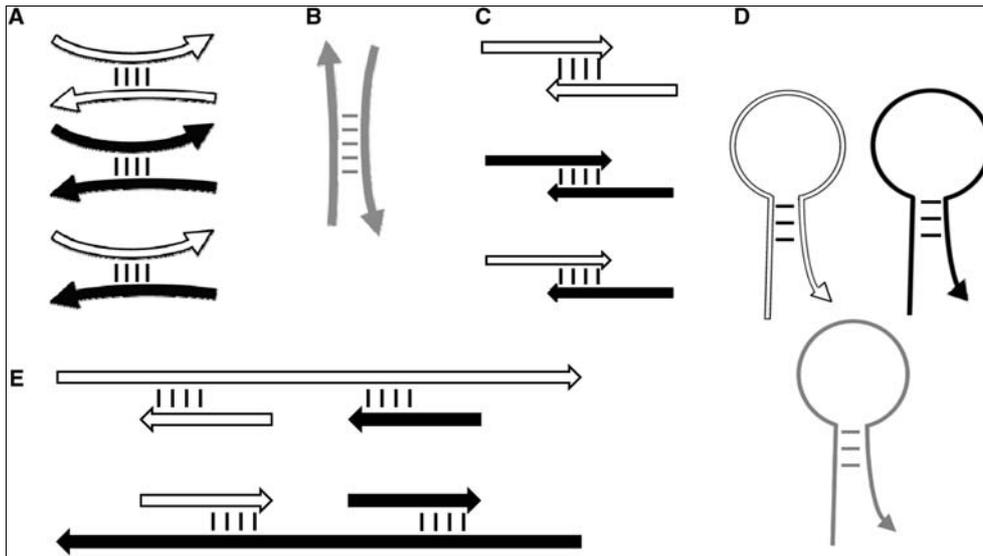


Abbildung 7: Übersicht über Hairpins und mögliche unerwünschte Sekundärstrukturen. (UNTERGASSER et al., 2012; "Figure 5: Primer3's thermodynamic models for predicting the stability of secondary structures")

Zum Auswählen der optimalen Primer berechnet PRIMER3, mittels einer *penalty function*, die Abweichung zu den Standardparametern und wählt jene Primer, die davon am geringsten abweichen (UNTERGASSER et al., 2012). Zusätzlich verwendet PRIMER3 eine sogenannte *objective function*. Diese ermöglicht die Gewichtung der einzelnen Parameter in der Berechnung zu beeinflussen und somit für die entsprechende Anwendung optimal gestaltete Primer zu finden. Beispielsweise kann in einer Versuchsanordnung die Schmelztemperatur der Primer von größerer Bedeutung sein als ihre Länge und wird demnach strenger bewertet (ROZEN & SKALETSKY, 2000). Es stehen über MSATCOMMANDER einige Auswahlmöglichkeiten bezüglich dem Primerdesign mittels PRIMER3 zur Verfügung. Ist die Funktion 'Tag Primers' ausgewählt, werden je nach Wunsch entweder der CAG-tag, der M13R-tag oder eine selbst definierte tag-Sequenz verwendet. Dabei ist wiederum zu beachten, dass PRIMER3 den für die jeweilige Sequenz besser geeigneten tag bevorzugt (Punkt 3.2.3.1). Es ist per Auswahl ebenso möglich die zuvor besprochenen *PIG-tails* an die Primer zu heften. Es steht außerdem zur Auswahl, ob kombinierte Mikrosatelliten beziehungsweise imperfekte Mikrosatelliten verwendet werden sollen. (FAIRCLOTH, 2008).

3.3 Methoden zum Testen von Mikrosatelliten-Marker

Das Testen der entwickelten Mikrosatelliten-Marker erfolgt in mehreren aufeinanderfolgenden Schritten. Zunächst werden die Marker mittels PCR (Polymerase Kettenreaktion) amplifiziert und nachher anhand der Gelelektrophorese sichtbar gemacht. In späteren Schritten werden die erfolgreich getesteten Individuen mittels Kapillarelektrophorese am *Genetic Analyzer 3500* (Applied Biosystems, Foster City, Kalifornien, USA) genauer untersucht und mit dem Programm *Genemarker* (SoftGenetics LLC, State College, PA, USA) auf Eignung als SSR-Marker überprüft. In den folgenden Punkten werden diese Verfahren detailliert geschildert.

3.3.1 PCR Amplifikation

Im Jahr 1983 entwickelte Dr. Kary Mullis die als *polymerase chain reaction* (PCR) bekannte Methode (MULLIS et al 1986). Diese Methode ermöglicht es, aus großen und komplexen Genomen bestimmte Abschnitte der DNA zu isolieren und zu amplifizieren. Dabei werden diese DNA-Abschnitte, begrenzt durch einzigartige Primer-Bindungssequenzen und mithilfe von oligonukleotiden Primerpaaren, exponentiell amplifiziert, bis eine Kopien-Anzahl erreicht ist, die eine weitere Charakterisierung und Weiterverarbeitung ermöglicht (FREELAND, 2005). Die PCR-Methode revolutionierte die Gen-Forschung und wird seither in verschiedensten Anwendungsbereichen in der Medizin und Biotechnologie, wie beispielsweise Sequenzierung, Genotypisierung, Gen-Expressions-Analysen, populationsgenetische Untersuchungen, Mutagenesen und viele mehr, eingesetzt (LIFE TECHNOLOGIES, 2011). Mullis wurde für die Entwicklung der PCR im Jahr 1993 der Nobelpreis für Chemie verliehen (NOBELPRIZE.ORG, 2014).

3.3.1.1 Funktionsweise

Grundlegend ist die PCR eine zyklische Aneinanderreihung von drei Grundschritten:

- Denaturierung, wobei sich die DNA Doppelstränge in Einzelstränge auflösen;
- Hybridisierung oder *annealing* der Primer an die Ausgangs-DNA (ab Schritt 2 an neu amplifizierte DNA) und
- Elongation, bei dem die Polymerase, ausgehend von den Primern, die dNTP's (Desoxyribonukleosidtriphosphate) entsprechend dem Ausgangsstrang hinzufügt.

Diese Abfolge wird mehrmals, meistens zwischen 20 und 30 mal, hintereinander abgewickelt (LIFE TECHNOLOGIES, 2011). Nachdem bei jedem Zyklus die Anzahl an neu gebildeten Strängen verdoppelt wird, weil auch die zuvor gebildeten neuen Stränge als Bindungsstelle für Primer dienen, können auf diese Weise, ausgehend von nur wenigen DNA-Molekülen, mehrere Millionen Kopien erstellt werden. (MULLIS et al., 1986; LIFE TECHNOLOGIES, 2011).

Um korrekt zu funktionieren werden dabei mehrere grundlegende Komponenten benötigt. Diese sind:

- Ausgangs DNA, in der die zu amplifizierende Sequenz enthalten ist;
- Zwei Primer, (*forward* und *reverse*) von denen die zu amplifizierende Sequenz jeweils links und rechts begrenzt wird;
- Thermostabile DNA Polymerase (*Taq*-Polymerase) zur Synthese der neuen Stränge;
- dNTP Lösung (Desoxiribonucleosidtriphosphate), die alle vier, zur Neubildung von DNA Strängen benötigten Nukleotide beinhaltet;
- Eine für die DNA-Polymerase geeignete Pufferlösung.

(SAMBROOK & RUSSELL, 2001)

Die einzelnen Komponenten werden in vorgegebene Konzentrationen in PCR-Reaktionsgefäße gegeben und in den *Thermocycler* platziert, der das gewünschte PCR-Programm durchführt. Es können zur Arbeitserleichterung, Zeiteinsparung und Vermeidung von Fehlerquellen auch *Readymixes* verwendet werden. Dies sind Pufferlösungen, die alle nötigen PCR-Komponenten und Reagentien (dNTP's sowie auch die thermostabile DNA-Polymerase) exklusive PCR-Wasser, DNA und Primer enthalten. Es sind zahlreiche für spezifische Reaktionen und diverse Anwendungen angepasste Varianten erhältlich. Sie bewirken eine deutliche Arbeitserleichterung und Zeitersparnis bei gleichzeitigem Eliminieren einiger Fehlerquellen, die beim Hantieren mit den verschiedenen Komponenten auftreten können und den Erfolg der PCR maßgeblich beeinflussen. Es müssen somit lediglich Wasser, die Primer sowie die Ausgangs DNA hinzugefügt werden um eine einsatzbereite PCR-Mixtur zu erhalten. Es gibt dazu mehrere Varianten, um spezifischen Anforderungen gerecht zu werden (Siehe nachfolgende Punkte). Ebenso sind auch unterschiedliche PCR-Programme möglich. Auch in dieser Arbeit ergibt sich, bedingt durch undefinierte Faktoren beim Primerdesign, (i.e. *annealing*-Temperatur) die Notwendigkeit unterschiedliche Varianten zu verwenden.

3.3.1.2 PCR Programme - Gradienten und *Touchdown*-PCR

Nachdem die *annealing*-Temperatur der Primer in Methode 1 beim Design nicht berücksichtigt wurde, ist es notwendig die PCR-Reaktion so zu gestalten, dass ein Temperaturbereich, in dem die Ankopplung stattfinden sollte, abgedeckt wird. Durch die Berechnung der *annealing*-Temperaturen mittels diverser Programme ist eine Eingrenzung des Temperaturbereichs möglich (vgl.: ThermoFisher Tm Calculator; New England BioLabs Tm Calculator u.v.m.). Um dies zu bewerkstelligen wurden die PCR-Reaktionen sowohl als Gradienten- als auch als Touchdown-PCR ausgeführt. (DON et al., 1991; DARREN & MATTICK, 2008) Bei der Gradienten PCR wird im Probenblock der PCR-Maschine während des gesamten Laufs beim *annealing*-Schritt ein Temperaturgradient angewandt. Dabei wird der Probenblock nicht gleichmäßig sondern längsseitig unterschiedlich stark erhitzt. Der Temperaturunterschied zwischen den beiden Enden des Blocks kann dabei bis zu 20° C betragen (EPPENDORF, 1997). Auf ähnliche Weise wird bei der Touchdown PCR gearbeitet. Der Temperaturgradient wird dabei nicht im Probenblock selbst erzeugt sondern zeitlich, indem die *annealing*-Temperatur von einem Startpunkt aus in jedem Schritt um einen bestimmten Faktor bis zum Erreichen des Touchdown-Punkts sinkt (DON et al., 1991). Die PCR Reaktion erfolgt dabei in zwei Phasen. In Phase 1 findet die eigentliche Touchdown-PCR statt, wobei in mehreren Zyklen die *annealing*-Temperatur von einem deutlich über dem Optimalwert liegendem Startpunkt aus schrittweise reduziert wird. Phase 2 besteht aus einer Standard PCR-Amplifikation mit der tiefsten Touchdown-Temperatur als *annealing*-Temperatur (DARREN & MATTICK, 2008). Beide PCR-Varianten zielen darauf ab, bei nicht genau definierter *annealing*-Temperatur trotzdem mehrere unterschiedliche Primerpaare in einem einzigen PCR-Lauf erfolgreich zu amplifizieren. Es ist anzumerken, dass die inhärenten Eigenschaften der Gradienten-PCR es zwar ermöglichen, die optimale *annealing*-Temperatur eines Primers einzugrenzen, aber auch verlangen, dass dabei in jeder Spalte des Probenblocks der selbe Primer enthalten sein muss (QIAGEN, 2010).

Dies hat zur Folge, dass pro Lauf entweder nur wenige Individuen getestet werden können oder nur eine sehr begrenzte Anzahl an Primerkombinationen. Bei der Touchdown PCR kann, im Gegensatz dazu, die optimale *annealing*-Temperatur nur indirekt ermittelt werden (DON et al., 1991), dafür aber in jeder Reihe des Probenblocks eine andere Primerkombination mit maximal 8 Individuen getestet werden. Zur Optimierung der Touchdown- und Gradienten-PCR-Programme können die *annealing*-Temperaturen der Primer mit Programmen wie *NEB Tm Calculator v.1.7* (NEW ENGLAND BIOLABS) berechnet werden. Dies ermöglicht die Eingrenzung der Tempera-

turniveaus der PCR-Programme und eine Gruppierung der Primerpaare nach ihrer optimalen *annealing*-Temperatur.

3.3.2 Gelelektrophorese

Die Agarose Gelelektrophorese ist eine einfache und sehr übliche Methode um DNA-Fragmente in einem elektrischen Feld aufzutrennen. Das zugrundeliegende Prinzip ist die Elektrophorese oder elektrische Überführung. Geladene Teilchen, wie beispielsweise DNA-Moleküle, wandern in einer Lösung unter dem Einfluss eines elektrischen Stroms, je nach Ladung und Größe, mit unterschiedlicher Geschwindigkeit in unterschiedliche Richtungen (ENGELHARDT et al., 1994). Nachdem die DNA-Moleküle eine negative Ladung besitzen, wandern diese zur positiv geladenen Anode. Sichtbar gemacht wird die DNA mittels Farbstoffe, die sich an diese binden. Diese Farbstoffe sind fluoreszierend, sie absorbieren also unsichtbares UV-Licht und emittieren sichtbares Licht, wodurch die gefärbten DNA-Fragmente sichtbar werden. Die Agarose-Gele werden meistens, je nach Bedarf, mit Konzentrationen zwischen 0,7% bis 2% Agarose in TBE Elektrophoresepuffer (TRIS-Borat-EDTA) aufgelöst. Schwächere Konzentrationen resultieren in einer etwas weitmächtigeren Gel-Matrix und somit in eine raschere Auftrennung mit geringerer Auflösung (LEWIS, 2015). Die Auflösung liegt bei Konzentrationen um 1.3% bei ca. 200 Basenpaaren, während bei sehr konzentrierten Gelen mit 3% und darüber bei ca. 20 Basenpaaren liegt (QUIAGEN, 2010). Bei höheren Konzentrationen von ca. 2% und darüber wird die Gelmatrix engmaschiger und das Gel dichter. Damit kann eine höhere Auflösung der Banden bewirkt werden, jedoch dauert der Vorgang aufgrund der langsamen Durchflussgeschwindigkeit deutlich länger. Diese Konzentrationen sind vor allem bei geringeren Fragmentgrößen von Vorteil. Beim Laden der DNA auf das Gel werden meistens sogenannte *loading buffers* hinzugegeben. Diese bewirken durch ihre Dichte das zuverlässige Herabsinken der DNA in die dafür vorgesehene Tasche im Gel und sind meistens gefärbt um die Handhabung und das Laden zu erleichtern. Ebenso wandert die Farbe des *loading buffers* in die selbe Richtung wie die DNA-Fragmente und ermöglicht somit den Elektrophoreseprozess zu überwachen. Um die DNA-Fragmentlänge zu bestimmen müssen in einer gesonderten Tasche sogenannte Längenmarker oder Längenstandards (= *size marker*) auf das Gel geladen werden. Diese beinhalten eine Mischung aus mehreren Fragmenten bekannter Länge. Sie treten am Gel als Leiter in Erscheinung und ermöglichen durch Vergleich eine relativ grobe Bestimmung der DNA-Fragmentlänge (LEWIS, 2015).

3.3.3 Kapillarelektrophorese

Die Kapillarelektrophorese ist eine allgemein übliche Technik für die DNA Genotypisierung und wird häufig für Mikrosatellitenanalysen verwendet (DAVIDSON & CHIBA, 2003). Das zugrundeliegende Prinzip ist, wie bei der zuvor beschriebenen Gelelektrophorese, die Elektrophorese. Im Unterschied zur Gelelektrophorese wandern die Moleküle dabei nicht durch eine Gelmatrix in einem flach ausgegossenem Agarose- oder Polymergel, sondern durch ein Polymer in einer dünnen Kapillare. Der grundlegende Unterschied liegt jedoch in der Detektion der Fragmentlängenunterschiede. Während bei der Gelelektrophorese die entstandenen Produktbanden auf optischem Wege erfasst werden müssen, können diese Unterschiede in der Kapillarelektrophorese durch direkte photometrische Methoden vollautomatisch erfasst werden (ENGELHARDT et al., 1994).

Der grundlegende Aufbau der Apparatur gliedert sich in:

- Behälter für Anodenpuffer
- Behälter für Kathodenpuffer
- Probenbehälter
- Hochspannungsquelle
- Trennkapillare
- Thermostatisierungs Mechanismus (Ofen)
- Detektor

(LIFE TECHNOLOGIES, 2010; PHILIPPS UNIVERSITÄT MARBURG - Fachbereich Chemie, 2007)

Die Aufnahme der Proben in die Kapillare erfolgt an einem Ende der Kapillare. Dazu wird das Gefäß mit dem Kathodenpuffer für kurze Zeit durch das Gefäß ausgetauscht, in dem die Proben beinhalten sind. Als nächstes wird die Probe in die Kapillare befördert. Dazu sind mehrere Möglichkeiten gegeben, wie beispielsweise die Injektion durch Siphon-Effekt (Schwerkraft), Druckinjektion beziehungsweise hydrodynamische Injektion oder die elektrokinetische Injektion (PHILIPPS UNIVERSITÄT MARBURG - Fachbereich Chemie, 2007). Letztere Methode wird auch im *Genetic Analyzer 3500* angewendet, weshalb darauf näher eingegangen wird. Bei der elektrokinetischen Injektion wird für eine bestimmte Zeit eine Spannung zwischen den beiden Enden der Kapillare angelegt, wodurch eine Flüssigkeitsströmung in der Kapillare erzeugt wird. Daraufhin wird ein festgelegtes Volumen der Probenflüssigkeit in die Kapillare aufgenommen (PHILIPPS UNIVERSITÄT MARBURG - Fachbereich Chemie, 2007). Die Probenmenge, die dabei aufge-

nommen wird, ist dabei abhängig von der Höhe der angelegten Spannung, der Zeit und der Mobilität der Probenbestandteile (ENGELHARDT et al., 1994). Nachdem die Proben in die Kapillare aufgenommen worden sind, wird der Probenbehälter wieder mit dem Kathodenpuffer-Behälter ausgetauscht. Es wird nun an beiden Enden der Kapillare wiederum eine Spannung angelegt, wodurch die Elektrophoretische Auftrennung der Probe beginnt (PHILIPPS UNIVERSITÄT MARBURG - Fachbereich Chemie, 2007). Die Probenbestandteile wandern dabei, abhängig von ihrer Ladung und Größe, dem pH-Wert und die Konzentration des Puffers, der Temperatur und angelegten Spannung und dem verwendeten Polymer, unterschiedlich schnell durch die Kapillare. Im Fall von DNA-Molekülen bewegen sich diese, aufgrund ihrer negativen Ladung, in Richtung der Anode. Der Großteil der Kapillare verläuft dabei durch einen Ofen, der für eine konstante Temperatur und somit für eine optimale Auftrennung sorgt und gleichzeitig bewirkt, dass die enthaltene DNA denaturiert bleibt (LIFE TECHNOLOGIES, 2010). An dem, dem Anodenpuffer nahem, Kapillarende ist ein Detektor angebracht. Neben der Möglichkeit ein Massenspektrometer oder Leitfähigkeitsdetektor zu verwenden, sind vor allem optische Detektionsgeräte, wie UV- oder Fluoreszenzdetektoren zu nennen. Der *Genetic Analyzer 3500* (siehe unten) verwendet einen Fluoreszenzdetektor, weshalb nur diese Methode berücksichtigt wird. Die Fluoreszenzdetektion kann nur dann funktionieren, wenn fluoreszente Farbstoffe an die DNA-Moleküle geheftet werden. Diese werden vorab in einer entsprechenden PCR an die Fragmente gebunden. Die Kapillare verläuft im Detektor durch ein Laserstrahl. Dieser Laserstrahl regt die fluoreszenten Farbstoffe an, die daraufhin fluoreszente Strahlung abgeben. Nachdem die abgegebene Strahlung eine vom Laserstrahl verschiedene Wellenlänge besitzt, kann diese von einer angebrachten CCD Kamera detektiert werden. Dabei können mehrere Farbstoffe in einem einzigen Lauf aufgrund ihrer unterschiedlichen Wellenlänge verwendet und detektiert werden. Das erhaltene Signal wird als Datei am angeschlossenen PC gespeichert und steht für weitere Analysen zur Verfügung (ENGELHARDT et al., 1994; LIFE TECHNOLOGIES, 2014). Im nachfolgenden Punkt werden der *Genetic Analyzer 3500* und die Arbeitsschritte im Detail aufgezeigt.

3.3.3.1 Genetic Analyzer 3500

Für die Fragmentanalyse wird jeweils ein Primer eines Primerpaares mit einem Fluoreszenzfarbstoff ausgestattet und am *Genetic Analyzer 3500* von *Life technologies* mittels Kapillarelektrophorese aufgetrennt. Der *Genetic Analyzer 3500* von *Life technologies* ist ein fluoreszenz-basiertes Gerät zur DNA Sequenzierung und Fragmentlängenanalyse mittels Kapillarelektropho-

rese. Dabei können in den 8 vorhandenen Kapillaren 8 Proben gleichzeitig analysiert werden. Durch die angeschlossene Computer-Arbeitsstation können sowohl das Gerät selbst gesteuert werden als auch die Daten aus dem Gerät empfangen und gespeichert werden (APPLIED BIOSYSTEMS, 2010).

Fragment sizing am Genetic Analyzer 3500

Vorbereitende Schritte

Bevor die Proben in den *Genetic Analyzer 3500* eingegeben werden können, müssen diese entsprechend vorbehandelt werden. Im Falle von DNA-Fragmentlängen Analysen beinhaltet dies zunächst eine PCR-Reaktion in der entsprechende fluoreszente Label an die DNA Fragmente geheftet werden. Dabei wurden zunächst sowohl fluoreszenzgelabelte M13R-Primer verwendet als auch direkt mit Fluoreszenzlabeln ausgestattete Primer in späteren Multiplex Reaktionen. Nach einem Purifikationsschritt folgt die Zugabe verschiedener Reagentien inklusive einem Size-Standard zu den PCR-Produkten. Dieser besteht aus zahlreichen DNA-Fragmenten bekannter Länge die ebenfalls mit einem Fluoreszenzlabel ausgestattet sind. Ähnlich wie die Leiter bei der Gelelektrophorese wird dieser Size-Standard dazu verwendet, die genaue Größe der Fragmente zu ermitteln. Nachdem beim *Genetic Analyzer 3500* 8 Kapillaren vorhanden sind, bei denen zum Teil leicht unterschiedliche Ergebnisse nicht auszuschließen sind, wird der Size-Standard auch dazu verwendet, durch Kalibration der Ergebnisse auf die bekannten Fragmentlängen die Daten zu normalisieren und somit Fehlerquellen zu beseitigen (LIFE TECHNOLOGIES, 2010). Als letzter vorbereitender Schritt wird die DNA in einem Thermocycler für einige Minuten denaturiert. Die Zugabe chemischer Zusätze bewirkt, dass auch bei Raumtemperatur die Einzelstränge erhalten bleiben und zudem verhindern sie die Entstehung von Luftblasen. Die Proben werden in spezielle Platten zu 96 *wells* geladen und sind bereit für die Eingabe in das Gerät.

Sample sheet

Nach dem Positionieren der vorbehandelten Proben im *Genetic Analyzer 3500* wird im System eine vorher erstellte Datei, das *Sample sheet* geladen. Darauf sind die für das System wichtigen Informationen zu den einzelnen Proben enthalten, darunter die genaue Position der Proben auf der Platte, die enthaltenen Farbstoffe, Laufparameter sowie Dauer der Injektion und andere.

Start des Kapillarelektrophorese-Laufs

Das System startet nun mit dem Lauf, der in folgenden Schritten abläuft:

- Vorbereitung der Kapillaren: Frisches Polymer wird aus dem Poly

merbehälter mit hohem Druck in die Kapillaren gepumpt.

- Elektrokinetische Injektion der Proben aus der Probenplatte in die Kapillaren.
- Kapillaren werden vom Probenbehälter wieder in den Kathodenpufferbehälter eingetaucht.
- Graduelle Anlegung einer elektrischen Spannung zwischen den Kapillarenden: die DNA-Fragmente beginnen in Richtung der Anode zu wandern.
- Der integrierte Ofen hält die Kapillaren auf eine konstante Temperatur um die Denaturierung der DNA aufrecht zu erhalten und eine optimale Auftrennung zu gewährleisten.
- Im Detektor werden die Fluoreszenzlabel der ankommenden DNA-Fragmente durch ein Laserstrahl angeregt.
- Die angeregten Fluoreszenz-label emittieren Strahlung in alle Richtungen.
- Registrierung der emittierten Strahlung von einem CCD Sensor.
- Konvertierung der erfassten Signale zu *multi-dye data* (für den gesamten Lauf: Jedes Signal ein *peak* auf dem Elektropherogramm).

(LIFE TECHNOLOGIES, 2010)

Output - Elektropherogramm

Die Software des *Genetic Analyzer 3500* generiert aus den Daten sogenannte Elektropherogramme für jeden verwendeten Farbstoff. Darauf sind die, anhand des Size Standards normalisierten *peaks* der Signale, zu erkennen. Die gewonnenen Daten können zur weiteren Verwendung mit diverser Software, wie im nachfolgenden Punkt 3.3.4 geschildert, exportiert werden.

3.3.4 Finden geeigneter SSR Marker

Die Daten des *Genetic Analyzer 3500* werden mittels entsprechender Software ausgewertet. Dabei werden die erhaltenen *Peaks* einzelnen Allelen zugeordnet beziehungsweise als unerwünschte Nebenprodukte verworfen. In dieser Arbeit wurde das Programm *GeneMarker* von *Softgenetics* verwendet.

Genemarker

Genemarker ist eine präzise, schnelle und nutzerfreundliche Genotypisierungs-Software für DNA Fragmentanalysen (LIU et al., 2011). Zunächst werden die Daten in das Programm geladen, wo diese sowohl als Elektropherogramm als auch als synthetisches Gelbild dargestellt werden. Am Elektropherogramm sind die einzelnen Farbkanäle zu erkennen und, bei gelungener PCR-Amplifikation der gewünschten Marker, einzelne Peaks als Darstellung der Anhäufung an Fragmenten einer bestimmten Länge. Anhand des *size-*

standards werden die einzelnen Datensätze zunächst normalisiert, also ausgerichtet und zur Deckung gebracht. Im Anschluss werden die einzelnen Elektropherogramme auf Allele durchsucht. Dabei werden verschiedene Kriterien beachtet, um unerwünschte *Peaks*, wie unspezifische PCR-Produkte, Stutterbanden, +A-Banden, Überlappungen anderer Fluoreszenzlabels sowie Verunreinigungen und unregelmäßig versetzte Peaks, korrekt zu erkennen und von den gesuchten Allelen zu unterscheiden. Werden Allele gefunden, können diese im *Panel-Editor* für das jeweilige Primerpaar vermerkt werden. Bei einer großen Anzahl an Daten kann nämlich, bei der Verwendung des gleichen Primerpaares, das erstellte *Panel* verwendet werden, um automatisiert nach den Allelen zu suchen. Durch das Programm können somit SSR Marker, die sowohl homozygot als auch heterozygot in Erscheinung treten, erkannt werden. Für jeden Datensatz wird am Ende des *scoring* ein sogenannter *allele report* erstellt. Dieser beinhaltet neben allgemeinen Informationen zum Datensatz auch die gefundenen Allele der jeweiligen Proben.

Auswahl geeigneter Marker

Aus der Summe aller gefundenen Marker können nun jene Kandidaten ausgewählt werden, die sich am besten für weitere Analysen eignen. Die konkreten Parameter für geeignete Marker werden im nachfolgenden Methodenteil angegeben. Darunter zählen Polymorphie, Signalstärke bzw. Peak-Höhe, Multiplexfähigkeit, Anzahl an Allele und weitere Parameter.

4 Methoden

In diesem Punkt werden die angewandten Methoden zur Entwicklung von Mikrosatelliten-Marker für die Art *Cyperus fuscus* besprochen. Zunächst werden die beiden in dieser Arbeit verwendeten Methoden zur Isolation von Mikrosatelliten, zum Primerdesign und zur Auswahl geeigneter SSR-Marker vorgestellt und anschließend einzeln im Detail besprochen. Dabei werden die Sequenzierungsdaten der 454 Pyrosequenzierung sowie die benutzten *tag's* und *PIG-tails* dargelegt. Es werden die Parameter für das Primerdesign und die dazu verwendeten Programme angeführt. Ebenso werden die Methoden zu den erfolgten Tests der Primer und Marker, beginnend mit der PCR über Gelelektrophorese und schließlich Kapillarelektrophorese, mit den jeweiligen Parametern und den getesteten Individuen aufgezeigt. Ebenso werden erste Anwendungstests der entwickelten Marker besprochen. Diese Tests sollen die Anwendbarkeit der erstellten Marker bestätigen und erste Aussagen über die genetische Vielfalt der unterschiedlichen Standorte gemäß der Fragestellungen ermöglichen.

4.1 Methoden zur Isolation der Mikrosatelliten und Primerdesign

In dieser Arbeit wurden zwei Methoden zur Entwicklung von Mikrosatelliten-Marker durchgeführt:

- Methode 1: Isolation von Mikrosatelliten mittels MSATCOMMANDER und Primerdesign mittels PRIMER3
- Methode 2: Isolation von Mikrosatelliten und Primerdesign durch ECOGENICS

Für Methode 1 wurden 454 Pyrosequenzierungsdaten des Individuums Cf_23_II_A herangezogen. Das am 21.08.2012 gesammelte Individuum stammt von einem natürlichen Standort am Elbufer bei Záryby in der Tschechischen Republik (Herbarnummer WHB: 62957; Siehe Tabelle 1) Die Sequenzierung erfolgte bereits im Juli 2013. Darauf basierend wurden zwischen Februar und Mai 2014 Mikrosatelliten gesucht sowie Primer entwickelt und getestet. Anhand der Frequenzdaten der gefundenen SSR Motive in Methode 1, wurde im Juli 2014 eine weitere Pyrosequenzierung mit vorangehenden *Enrichment* und anschließendem Primerdesign durch die Firma Ecogenics (Balgach, Schweiz) in Auftrag gegeben (Siehe 4.1.2). Für die Pyrosequenzierung wurden die Individuen mit den Herbarnummern WHB 62959 und WHB 62987 herangezogen. Diese stammen aus Fischhältern in

Semovice und Tchořovice in der Tschechischen Republik und wurden zwischen 6 und 9.08.2012 gesammelt. Die Primer wurden zwischen Oktober und November 2014 getestet. In beiden Methoden wurden, wie bereits im Punkt 3.2.3 besprochen, *tags* und *PIG-tails* verwendet. In den nachfolgenden Punkten wird auf die einzelnen Schritte näher eingegangen. Zunächst werden die verwendeten *Cyperus fuscus* Individuen vorgestellt, die für die Tests der entwickelten Primer verwendet wurden.

Test-Individuen

Die getesteten Individuen unterscheiden sich hauptsächlich in der geografischen Lage ihrer Entnahme, dem Nutzungstyp des Gewässers, der räumlichen Trennung der Flusssysteme sowie dem Bodenprofil (Tabelle 6). Dadurch soll ein möglichst großes Genpool der Art erfasst werden, um möglichst viele polymorphe Allele finden zu können und somit die Eignung der SSR Marker für weitere populationsgenetische Untersuchungen zu ermitteln. Ausschlaggebend für die Auswahl geeigneter Kandidaten war außerdem eine ausreichende Menge an verfügbarem Pflanzenmaterial sowie hochwertige Extrakte. Aufgrund der verwendeten Eppendorf Mastercycler gradient PCR Cycler (Eppendorf, Hamburg, Germany) mit Heizplatten für 8x12 Mikroreaktionsgefäße (Anm.: *Eppi*) wurden 7 Individuen plus 1 Nullprobe gewählt. Für die Tests der Primer aus Methode 1 und Methode 2 mittels Gelelektrophorese wurden die in Tabelle 1 aufgelisteten Individuen verwendet.

Tabelle 1: Bezeichnung und Herkunft der Test-Individuen. *Bodenprofil Bezeichnet die Bodenschicht in der die Entnahme der Pflanzen erfolgte.

Individuum	Herkunft	Nutzungstyp	Bodenprofil*
WHB: 62957	Zaryby, Elbufer	Natürlicher Standort	Profil A (0-5 cm)
WHB: 62960	Hluboká nad Vltavou; Südböhmen	Fischhälter	Profil A (0-5 cm)
WHB: 62982	Lanzhot; Thayaufener	Natürlicher Standort	Oberirdisch
WHB: 62955	Cigacice; Seitenbecken der Oder	Natürlicher Standort	Oberirdisch
WHB: 62959	Semovice; Mittelböhmen	Fischhälter	Oberirdisch
WHB: 62958	Čejetice; Südböhmen	Fischhälter	Oberirdisch
WHB: 62956	Zwentendorf; Traisenkanal	Natürlicher Standort	Oberirdisch

4.1.1 Methode 1: Isolation von Mikrosatelliten mittels MSATCOMMANDER und Primerdesign mittels PRIMER3

4.1.1.1 Ausgangsdaten - Sequenzdaten

Die Sequenzdaten wurden mittels 454 Pyrosequenzierung an einem Genome Sequencer FLX Titanium Gerät (454; Roche diagnostics, Basel, Schweiz) durch LGC Genomics (Berlin, Deutschland) erstellt. Es wurden 143.027 *reads* mit einer durchschnittlichen Länge von 238 Basenpaaren für die Mikrosatellitenentwicklung erstellt (vgl. GenBank Sequence Read Archive: Bio-Project no. PRJNA275048) (Abbildung 8). Das kürzeste Fragment besitzt 7 Basen und das längste 762 (Tabelle 2).

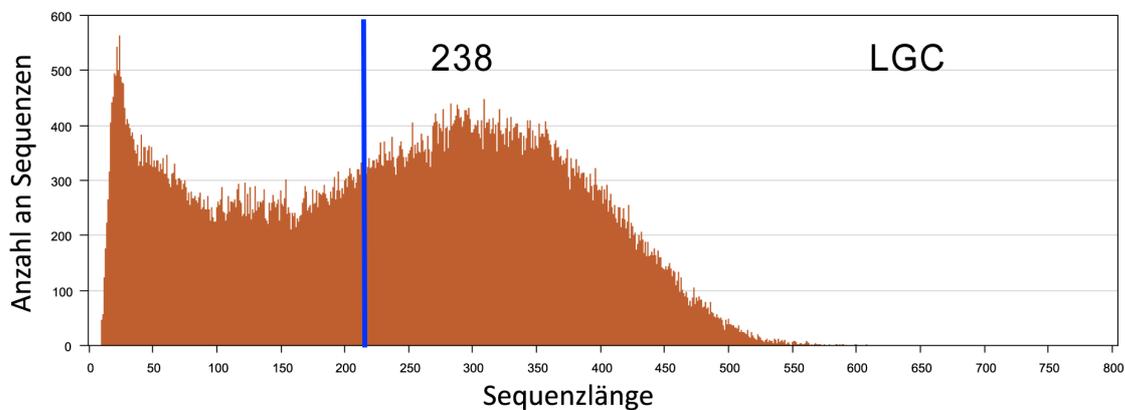


Abbildung 8: Length-graph von 143027 Sequenzen

Tabelle 2: Übersicht der 454 Pyrosequenzierungsdaten zu Methode 1

Methode 1 - LGC	
Sequenzierungsmethode	454 GS FLX+ Titanium
Read Gesamtanzahl	143027
Durchschn. Readlänge	238
Min. Länge	7
Max. Länge	762

Von diesen 143.027 *reads* beinhalteten 520 *reads* SSR-Sequenzen. Davon eigneten sich 101 Sequenzen zum Primerdesign.

4.1.1.2 Primer Design

SSR Art

Die verfügbaren 143.027 *reads* wurden, angelehnt an CSENSICS et al. (2010) und TAKAYAMA et al (2010), auf Mikrosatelliten durchsucht, die bestimmten Kriterien entsprechen.

Es wurden sowohl *perfect repeats* als auch *imperfect repeats* (BIOINFORMATICS.ORG) gesucht jedoch wurde das Hauptaugenmerk auf den *perfect repeats* gesetzt (Punkt 3.2.3.3.1).

Der Fokus wurde auf Di-, Tri- und Tetranukleotide-*repeats* gesetzt. Mononukleotide-*repeats* wurden ausgeschlossen, da diese bei der PCR Amplifikation zahlreiche *slippage* Fehler (Punkt 3.1.2) hervorrufen können, die wiederum zu falschen Ergebnissen führen (SUN et al., 2006).

Um ausreichend Polymorphismus zu gewährleisten wurden bei der Suche mittels MSATCOMMANDER die Mindestanzahl an Wiederholungen folgendermaßen festgesetzt:

Tabelle 3: Untergrenze der Wiederholungen der SSR Motive

Motiv	<i>repeats</i>
Di-	10
Tri-	8
Tetra-	6
Combined	Di - 6 ; Tri - 4

Die Obergrenzen wurden in dieser Arbeit nicht beachtet, nicht zuletzt aus dem Grund, dass bei der verwendeten Methode eine relativ große Anzahl an potentiellen SSR-Motiven zu kurze *flanking-regions* für das Primerdesign aufwies und somit die gesamte Bandbreite der gefundenen SSR's in Betracht gezogen wurde (siehe Punkt 6). Von den 143.027 *reads* konnten 520 *reads* gefunden werden, die SSR's mit den genannten Kriterien enthielten.

Die *flanking-regions* wurden mittels BLAST (ALTSCHUL et al., 1990:) auf Einzigartigkeit überprüft (siehe S. 34: Flanking-regions auf Einzigartigkeit überprüfen - NCBI Blast).

Parameter für das Primerdesign

Es wurden Primerpaare nur dann in Betracht gezogen, wenn diese mindestens 10 Basenpaare vom Anfang und Ende der Mikrosatellitensequenz entfernt waren. PRIMER3 verwendet die nachfolgenden Standardparameter für das Primerdesign:

- Amplikon Länge: 100 -500 bp
- Schmelztemperatur 60° C (57° C - 62° C)
- Temperaturabweichungen zwischen Primer max. 5° C
- GC Gehalt optimal 50%

Aufgrund dieser Parameter wurden von den 520 verfügbaren *reads* 419 verworfen. In 101 Sequenzen konnten entsprechende Primer designt werden. Diese 101 Sequenzen und die entsprechenden Primer wurden, nach Betrachtung der Scorings der einzelnen Primerpaare sowie auch nach subjektiver Beurteilung, auf geeignete Kandidaten für die weiteren Tests durchsucht. Dabei wurden 42 Primerpaare ausgewählt (Siehe Ergebnisse 5.1).

Verwendete *tag's* und *PIG-tails*

MSATCOMMANDER und PRIMER 3 Tags

In MSATCOMMANDER und PRIMER3 werden zwei unterschiedliche Tags verwendet, nämlich M13 *reverse*, kurz M13R, mit der Sequenz 5' - GGA AAC AGC TAT GAC CAT - 3' und CAG mit der Sequenz 5' - CAG TCG GGC GTC ATC A - 3' (FAIRCLOTH, 2008). Dabei hängt MSATCOMMANDER an das 5'Ende jedes Primers beide Tags an und entfernt dabei jene Basen, die Primer und Tag an der Bindungsstelle gemeinsam haben. Es entstehen somit für ein Primerpaar vier verschiedene Kombinationen. Diese Kombinationen werden durch PRIMER3 anhand diverser, weiter unten im Punkt 4.1.2 geschilderten Kriterien bewertet. Es folgt die Auswahl jener Kombination mit der besten Bewertung (FAIRCLOTH, 2008).

Primer-*PIG-tail*

Die in Methode 1. erstellten Primer wurden direkt mit dem PIG-tail GTTT (BROWNSTEIN et al., 1996) ausgestattet. Bei Cf019_1 befindet sich das *PIG-tail* am Forward-Primer und das M13R Tag am Reverse-Primer.

In der nachfolgenden Tabelle werden die verwendeten *tag's* und PIG-tail von Methode 1 dargestellt (Tabelle 4).

Tabelle 4: tag's und PIG-tail - Methode 1

	Tag	Sequenz
MSATCOMMANDER	M13R	5' - <u>GGA AAC AGC TAT GAC CAT</u> - 3'
	CAG	5' - <u>CAG TCG GGC GTC ATCA</u> - 3'
PIG-tail	PIG-tail	5' - GTTT - 3'

4.1.1.3 Primer Tests und Auswahl geeigneter Mikrosatelliten-Marker

Die Primer wurden zwischen Februar und November 2014 getestet. Zunächst wurden diese Primer an 4 ausgewählten Individuen mittels Gelelektrophorese getestet (Tabelle 1, 1-4). Die somit ermittelten, positiv amplifizierten PCR Produkte wurden in weiterer Folge an insgesamt 7 Individuen (Tabelle 1) mittels Kapillarelektrophorese am *Genetic Analyzer 3500* getestet und mithilfe des Programms *Genemarker* auf Eignung als SSR-Marker überprüft.

4.1.1.3.1 PCR Amplifikation

PCR Komponenten

Die PCR Reaktionen wurden in einem Endvolumen von 25 µl durchgeführt. Dafür wurden 12.5 µl des Readymix *REDtaq® ReadyMix™* (Sigma-Aldrich, St. Louis, Missouri, USA) verwendet. Der Readymix ist mit einem inerten roten Farbstoff zur leichteren Handhabung ausgestattet. Dazu wurden 0.40 µM Fluoreszenz-gelabelter (FAM), universaler M13R- oder CAG-Primer, 0.40 µM der mit dem GTTT PIG-tail gelabelten Primer, 0.04 µM der M13R- oder CAG-gelabelten Primer und 1 µM verdünnter DNA der Testindividuen hinzugefügt und mit PCR-Wasser auf 25 µl Endvolumen verdünnt (Tabelle 5) (siehe auch 3.3.1)

Tabelle 5: PCR Komponenten und Reaktionsmengen Methode 1

Komponente	Reaktionsmenge
REDtaq® ReadyMix™	12,5 µl
FAM-Universalprimer (M13R; CAG)	0.4 µM
Primer mit GTTT PIG-tail	0.4 µM
Primer mit M13R- oder CAG-Tag	0.04 µM
DNA Testindividuum	1 µM
PCR-Wasser	ca. 9,5 µl
Endvolumen	25 µl

PCR Programme - Gradienten und *Touchdown*-PCR

In dieser Arbeit wurde eine abgeänderte Variante der in *Darren & Mattik 2008* beschriebenen *Touchdown*-PCR angewendet. Die *annealing*-Temperatur der Primer wurde mit dem online Programm *NEB Tm Calculator v.1.7* (NEW ENGLAND BIOLABS) berechnet. Die Berechnung erfolgt mithilfe von thermodynamischen Modellen unter Berücksichtigung der verwendeten Pufferlösungen sowie DNA-Polymerase (NEW ENGLAND BIOLABS). (Anm.: Im Laufe der Primertests wurden die entsprechenden PCR-Programme immer weiter angepasst und optimiert. Es wird das jeweils aktuellste Programm aufgezeigt).

PCR-Programm Gradient

Für die Gradienten PCR wurde, ausgehend von einem Temperaturniveau von 58° C, ein Gradient von $\pm 5^\circ$ C gewählt (Tabelle 6). Somit liegt die tiefste *annealing* Temperatur bei 53° C und die höchste bei 63° C. Insgesamt wurden 38 Zyklen durchgeführt.

Tabelle 6: PCR-Programm Gradient

1	95° - 5'	Melting
2	95° - 45"	
3	54° - 60"	<i>annealing</i>
4	58° ± 5° Gradient	(H:63° ; L:53°)
5	72° - 45"	Elongation
	GoTo step 2 * 38	Cycles
9	72° - 5'	Final Elong.
10	60° - 30'	
11	10° for ever	
	END	

PCR-Programm Touchdown

Die erste Phase der Touchdown PCR besteht aus 30 Zyklen. Ausgehend vom höchsten Temperaturniveau von 54° C, beziehungsweise einige Grad Celsius über der berechneten *annealing*-Temperatur, wurde in jedem Zyklus die Temperatur um 0,2° C gesenkt. Als Touchdown Temperatur ergibt sich daher nach 29 + 1 Zyklen 48° C, beziehungsweise knapp unter der berechneten *annealing*-Temperatur der Primer. Im Anschluss folgte eine kürzere Phase mit einer normalen PCR bei einer *annealing*-Temperatur um einige Grad Celsius tiefer als die Touchdown-Temperatur in Phase eins. Entgegen den Protokollen in DARREN & MATTICK, 2008, wurde diese Temperatur absichtlich um einige Grad Celsius herabgesenkt. Dadurch sollten auch tiefere *annealing*-Temperaturen, die deutlich außerhalb des berechneten Temperaturbereiches liegen, berücksichtigt werden. Sollte nämlich ein Primerpaar sehr schwache Produktbanden aufweisen, könnte das ein Hinweis darauf sein, dass die berechnete *annealing*-Temperatur für das Primerpaar zu hoch ist. Um die Zyklenanzahl trotzdem nicht übermäßig zu erhöhen, (Anm.: Empfehlung liegt bei maximal 30-35 Zyklen für eine Touchdown-PCR, DARREN & MATTICK, 2008) und den Großteil der Amplifikation im berechneten Temperaturbereich zu halten, wurde die tiefe *annealing*-Temperatur erst in Phase 2 integriert.

Tabelle 7: PCR-Programm Touchdown

1	95° - 5'		Phase 1
2	95° - 45"	Melting	
3	54° - 90"	<i>annealing</i>	
4	-0,2° / Cycle	(H:54° ; L:48°)	
5	72° - 60"	Elongation	
	GoTo step 2 * 29		
6	95° - 45"	Melting	Phase 2
7	45° - 90"	<i>annealing</i>	
8	72° - 60"	Elongation	
	GoTo step 6 * 9		
9	72° - 5'	Final Elong.	
10	60° - 30'		
11	10° for ever		
	END		

4.1.1.3.2 Test am Agarosegel

Von den 101 designten Primerpaaren wurden 42 vielversprechende mittels PCR amplifiziert und an Agarosegelen getestet. Nur bei vorhandenen Banden in den erwarteten Bereichen, d.h. bei Vorhandensein eines PCR-Produktes, wurden die entsprechenden Primerpaare für die Tests am Genetic Analyzer 3500 weiterverwendet. Die Agarose-Gelelektrophorese ermöglichte dabei auch die Erfassung der Qualität der vorangegangenen PCR-Reaktion. (Anm.: Bei nur schwach ausgeprägten Produktbanden oder ähnlichen Problemen konnten weitere PCR-Reaktionen entsprechend angepasst und optimiert werden). Es wurden sowohl *Biozym LE Agarose* (BIOZYM SCIENTIFIC GmbH, Hessisch Oldendorf, Deutschland) als auch *SeaKem® LE Agarose* (LONZA GROUP Ltd, Basel, Schweiz) in 1xTBE aufgelöst und bei 700 Watt für ca. 2 Minuten erhitzt. Nach Aushärtung des Gels im Gießstand wurde dieses in der Gel-Kammer mit den aufzutrennenden Proben versehen. Für die Tests wurden Gelkonzentrationen zwischen 1% und 2% gewählt. Nachdem das Vorhandensein eines Produktes im Vergleich zur Auflösung der Auftrennung im Vordergrund stand, musste die Auftrennung nur soweit erfolgen um eine annähernde Fragmentgrößen-Ermittlung anhand der beige-

fügten Leiter zu ermöglichen. Deshalb erfolgten die Auftrennungen bei Stromstärken im oberen Bereich mit durchschnittlich 100mA und einer Laufzeit zwischen 60 und 120 Minuten je nach Gelgröße und -konzentration. Als Pufferlösung für die Gelelektrophorese diente 1xTBE (Tris-Borat-EDTA) frisch hergestellt aus 5xTBE Stammlösung. Als Farbstoffe zur Sichtbarmachung der Produktbanden dienten sowohl GelRed (BIOTIUM, Hayward, Kalifornien, USA) als auch peqGreen (PEQLAB Biotechnologie GmbH, Erlangen, Deutschland). Für die ersten Primertests von Februar bis Mitte Mai 2014 wurde das Gel, nach dem abgeschlossenen Lauf, für ca. 15 Minuten in ein GelRed Bad unter Lichtausschluss getaucht. Alle nachfolgenden Tests wurden mittels peqGREEN gefärbt. Die Zugabe von peqGREEN erfolgte direkt in das noch flüssige Gel kurz vor dem Ausgießen. Von den 42 getesteten Primerpaaren zeigten 13 an allen Testindividuen klare Produktbanden und somit positive Amplifikation auf (Siehe Ergebnisse 5.1).

4.1.1.3.3 Fragmentanalyse am *Genetic Analyzer 3500*

Für die Analyse am Genetic Analyzer 3500 wurden jene 13 Primerpaare, die am Agarose-Gel eine positive Amplifikation der DNA-Fragmente aufwiesen, mit Fluoreszenzmarkierungen ausgestattet (Punkt 3.3.3). Dazu wurde eine erneute PCR-Reaktion mit den positiven Primer sowie Fluoreszenzgelabelter M13R- oder CAG-universal Primer durchgeführt. Es wurden die Fluoreszenz-Farbstoffe *ATTO 565* und *ATTO 550* (Sigma-Aldrich, St. Louis, Missouri, USA) sowie *FAMTM* (Applied Biosystems, Foster City, Kalifornien, USA) verwendet, die an eine entsprechende M13R- oder CAG-Sequenz angeheftet wurden. Als Size-Standard diente der *GeneScanTM 600 LIZ[®]* Size Standard (Applied Biosystems, Foster City, Kalifornien, USA). Die *sample sheets* (siehe 3.3.3.1) wurden vor jedem Lauf entsprechend für Fragmentanalysen erstellt und in das Gerät eingelesen. Es folgte die kapillarelektrophoretische Auftrennung der Proben. Die erhaltenen Elektropherogramme wurden in das Programm *Genemarker 2.4* (SoftGenetics, LLC, State college, PA) zum *fragment-sizing* übertragen.

4.1.1.3.4 Auswahl geeigneter SSR Marker

Es wurden alle monomorphen Marker sowie Marker, die an ein oder mehr Individuen keine klaren Peaks aufwiesen, verworfen. Es ist überprüft worden, ob die Signalintensität, d.h. die Höhe der Peaks und deshalb die Menge an PCR-Produkt, bei allen getesteten Individuen annähernd gleich groß war und, dass möglichst keine oder nur sehr geringe *stutter*-Banden sowie +A-Banden auftraten. Von den 13 getesteten Marker mussten 6 wegen Monomorphismus verworfen werden. Von den verbleibenden 7 Marker konnten 4

gefunden werden, die alle genannten Auswahlkriterien erfüllten (Punkt 5.1).

4.1.1.4 Zusammenfassung Methode 1

Nachfolgend sind die wichtigsten Punkte der Methode 1 tabellarisch zusammengefasst (Tabelle 8).

Tabelle 8: Zusammenfassung Methode 1

Ausgangsdaten	
Gesamtanzahl der <i>reads</i> aus 454 Pyrosequenzierung	143027
Anzahl an <i>reads</i> mit SSR-Sequenzen	520
Anzahl an geeigneter <i>reads</i> für Primerdesign	101
Getestete Primerpaare	42
Primer	
<i>tag's</i>	M13R: 5' - <u>GGA AAC AGC TAT GAC CAT</u> – 3' (FAIRCLOTH, 2008)
	CAG: 5' - <u>CAG TCG GGC GTC ATC A</u> – 3' (FAIRCLOTH, 2008)
PIG-tail	5'- GTTT - 3'
PCR	
Programme	Gradienten- und Touchdown-PCR
Ready Mix	<i>REDtaq® ReadyMix™</i> (Sigma-Aldrich)
Fragmentanalyse	
Fluoreszenzmarkierungen	<i>FAM</i> (Applied Biosystems)
	<i>ATTO 565</i> (Sigma-Aldrich)
	<i>ATTO 550</i> (Sigma-Aldrich)
Ausgewählte SSR Marker*	4

*Anm.: *Positive Amplifikation an allen Testindividuen sowie alle geforderten Kriterien erfüllt (Polymorphie, einheitliche peak-Höhe, Fehlen von stutter-banden/-A Banden).*

4.1.2 Methode 2: Isolation von Mikrosatelliten mit Enrichment und Primerdesign durch *Ecogenics*

4.1.2.1 Ausgangsdaten - Enrichment und 454 Pyrosequenzierung

Im Unterschied zu Methode 1 wurde in dieser Methode vor der Pyrosequenzierung ein *Enrichment* (Punkt 3.2.2.2) durchgeführt (Anm. zwischen Juli und August 2014). Es wurde extrahierte DNA der Individuen 62959 und 62987 sowie getrocknetes Frischmaterial als Backup an Ecogenics geschickt. Für die Art *Cyperus fuscus* wurden für das Enrichment die in Methode 1 am häufigsten vorkommenden Motive verwendet (Siehe 3.2.2.2). Bei der Auswahl der verwendeten Motive ist das Ziel, eine gewisse Vielfalt zu erreichen. Dinukleotide *repeats* sind grundsätzlich sehr polymorph, was für populationsgenetische Untersuchungen einen wichtigen Aspekt darstellt. Trinukleotide und tetranukleotide *repeats* zeigen hingegen weniger Stutter-Banden und vereinfachen somit das scoring (GARDNER et al., 2011). Es wurden die Motive CT und GT, die zugleich auch als Standardmotive von Ecogenics angeboten werden, sowie die Motive AAG und ATGT verwendet (Punkt 5.2.1.1 Tabelle 18: Motiv-Häufigkeiten). Nach dem Enrichment wurden die erhaltenen Fragmente, wie auch in Methode 1, mittels 454 Pyrosequenzierung an einem Genome Sequencer FLX Titanium Gerät (454; Roche diagnostics, Basel, Schweiz) sequenziert. Dabei wurden 4877 *reads* mit einer durchschnittlichen Länge von 415 Basenpaaren erstellt (vgl. GenBank Sequence Read Archive: BioProject no. PRJNA275048) (Abbildung 9). Das kürzeste Fragment besitzt 34 Basen und das längste 801 (Tabelle 9).

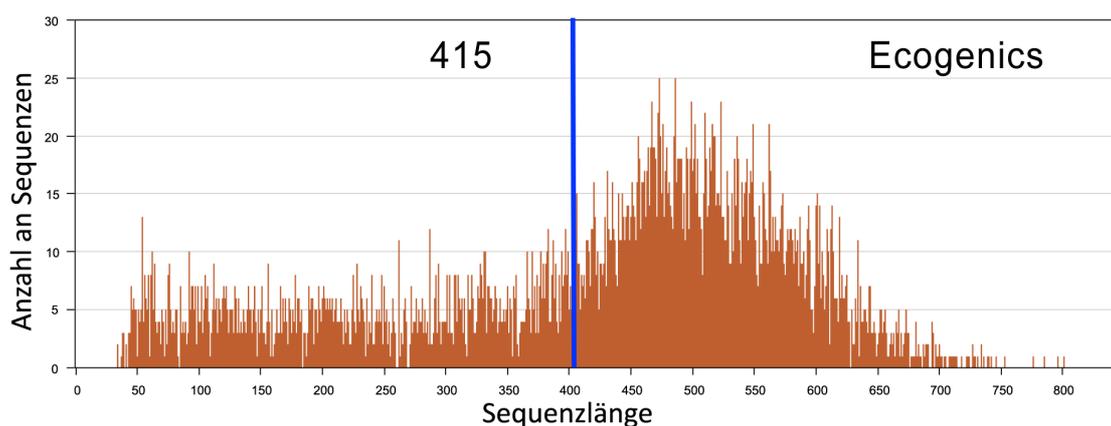


Abbildung 9: Sequenzlängen von 4877 Sequenzen

Tabelle 9: Übersicht der 454 Pyrosequenzierungsdaten zu Methode 2

Methode 2 - Ecogenics	
Sequenzierungs-methode	454 GS FLX+ Titanium + Enrichment
Read Geasmtanzahl	4877
Durchschn. Read Länge	415
Min. Länge	34
Max. Länge	801

4.1.2.2 Primerdesign

Suchen nach Mikrosatelliten und Primerdesign

Die 4877 erhaltenen Sequenzen wurden durch Ecogenics, analog zu Methode 1 mittels PRIMER 3 core, auf SSR's durchsucht. Von den 4877 *reads* enthielten 967 SSR Motive. Davon eigneten sich 494 zum Primerdesign (Siehe 5.1). Beim Design wurde eine einheitliche *annealing*-Temperatur von 56° C berücksichtigt. An den *forward* Primern wurde am 5' Ende das M13 Tag aus SCHUELKE 2000 angeheftet, ohne dabei die gemeinsamen Basen zwischen Primer und Tag zu entfernen. Die reversen Primer wurden ohne *PIG-tails* entworfen. Die 80 vielversprechendsten Primerpaare wurden mit dem M13 Tag aus Schuelke 2000 ausgestattet und zur Synthese an Microsynth (Microsynth AG, Balgach, Schweiz) weitergeleitet. Das *PIG-tail* GTTT (BROWN-STEIN et al., 1996) wurde in einem späteren Schritt angefügt.

Verwendete *tag's* und *PIG-tails*

ECOGENICS Tags

Ecogenics verwendet ausschließlich einen M13-Tag, jedoch besitzt dieser eine andere Sequenz als jener in MSATCOMMANDER. Alle forward Primer wurden am 5'Ende mit der M13 Sequenz 5'- TGT AAA ACG ACG GCC AGT - 3' aus SCHUELKE 2000 versehen.

Primer-PIG-tail

Wie auch in Methode 1 wurde ein GTTT *PIG-tail* verwendet. Aus den 80 gelieferten Primer wurden 48 polymorphe Marker mit einem *PIG-tail* versehen (Siehe weiter unten).

In der nachfolgenden Tabelle werden der verwendete *tag* und *PIG-tail* von Methode 2 dargestellt (Tabelle 10).

Tabelle 10: Tag und PIG-tail - Methode 2

	Tag	Sequenz
ECOGENICS	M13	5'- <u>TGT AAA ACG ACG GCC AGT</u> - 3'
PIG-tail	PIG-tail	5'- GTTT - 3'

4.1.2.3 Primer Tests und Auswahl geeigneter SSR Marker – Methode 2

Die 80 Ecogenics-Primerpaare wurden zwischen Oktober und November 2014 getestet (Anm.: 40 der 80 Primerpaare wurden von DI. J. BÖCKELMANN getestet) Die Primer wurden an 7 Individuen, analog zu Methode 1, getestet (Tabelle 1). Dazu wurden diese mittels PCR Amplifikation und anschließender Gelelektrophorese getestet. Die positiv amplifizierten PCR Produkte wurden wie in Methode 1 mittels Kapillarelektrophorese am *Genetic Analyzer 3500* getestet und mithilfe des Programms *Genemarker* auf Eignung als SSR-Marker überprüft.

4.1.2.3.1 PCR Amplifikation

PCR Komponenten

Für die Ecogenics-Primer wurde, auf Empfehlung von Ecogenics (BUSER A., persönliche Mitteilung), ein PCR-Mix mit einer Hotstart-DNA-Polymerase verwendet, nämlich *JumpStart™ REDTaq® ReadyMix™* (Sigma-Aldrich, St. Louis, Missouri, USA). Auch dieser Readymix ist mit einem inerten roten Farbstoff zur leichteren Handhabung ausgestattet. Die PCR Reaktionen wurden, analog zu Methode 1, in einem Endvolumen von 25 µl und selben Reaktionsmengen der einzelnen Komponenten durchgeführt.

Tabelle 11: PCR Komponenten und Reaktionsmengen Methode 2

Komponente	Reaktionsmenge
<i>JumpStart™ REDTaq® ReadyMix™</i>	12,5 µl
FAM-Universalprimer (M13; CAG)	0.4 µM
Primer mit GTTT PIG-tail	0.4 µM
Primer mit M13- oder CAG-Tag	0.04 µM
DNA Testindividuum	1 µM
PCR-Wasser	ca. 9,5 µl
Endvolumen	25 µl

PCR-Programm ECOGENICS

Nachdem alle Ecogenics-Primer mit einer vordefinierten *annealing*-Temperatur von 56° C designt worden sind, wurde das entsprechende PCR-Programm mit 38 Zyklen bei einer *annealing*-Temperatur von 56° C durchgeführt.

Tabelle 12: PCR-Programm 56° Ecogenics

1	95° - 5'	Melting
2	95° - 45"	
3	56° - 60"	<i>annealing</i>
4	72° - 60"	Elongation
	GoTo step 2 * 37	
6	72° - 5'	Final Elong.
7	60° - 30'	
8	10° for ever	
	END	

4.1.2.3.2 Test am Agarosegel

Die Tests am Agarosegel erfolgten analog zu Methode 1 (siehe 4.1.1.3.2). Die Gelkonzentration betrug dabei zwischen 1% und 1,5%. In dieser Methode wurde ausschließlich peqGREEN (PEQLAB Biotechnologie GmbH, Erlan-

gen, Deutschland) als Farbstoff verwendet. Von den 80 getesteten Primerpaaren zeigten 22 kein PCR Produkt beziehungsweise nicht an allen 7 Individuen eine Produktbande oder wiesen nur sehr schwache, kaum erkennbare Produktbanden auf. Die übrigen 58 Primerpaare wurden am Genetic Analyzer 3500 weiter getestet.

4.1.2.3.3 Fragmentanalyse am *Genetic Analyzer 3500*

Die Probenvorbereitung und Analyse der Primer erfolgte analog zu Methode 1 (siehe 4.1.1.3.3). Zusätzlich zu den Farbstoffen aus Methode 1 wurden *ATTO 532* (Sigma-Aldrich, St. Louis, Missouri, USA) und *Yakima Yellow* (Epoch Biosciences, Bothell, Washington, USA) verwendet.

4.1.2.3.4 Auswahl geeigneter Mikrosatelliten-Marker

Mittels Genemarker wurden die Elektropherogramme, analog zu Methode 1, auf Marker durchsucht. Aus der Summe aller gefundenen Marker wurden jene Kandidaten ausgewählt, die sich am besten für weitere Analysen eignen. Dabei wurden wiederum alle monomorphen Marker verworfen. Ebenso wurden alle Marker verworfen, die an ein oder mehr Individuen keine Peaks aufwiesen. Weiters ist überprüft worden, ob die Signalintensität beziehungsweise die Höhe der Peaks und somit Menge an PCR-Produkt, bei allen getesteten Individuen annähernd gleich groß ist. (Anm.: Schwache Peaks können auf Probleme bei der Amplifikation hindeuten. Deshalb werden, sofern ausreichend Kandidaten vorhanden sind, diese Marker vorerst ausgeschlossen und eventuell in späteren Tests erneut herangezogen). Ähnlich wurde bei Markern verfahren, die zwar an allen Individuen deutliche Peaks aufwiesen, jedoch auch starke Stutterbands oder -A-Banden besitzen. (Anm.: Diese können entsprechend optimiert und erneut getestet werden). All jene polymorphen Marker, die deutliche Peaks an allen Individuen aufwiesen und wenige bis keine Stutterbanden beziehungsweise keine -A-Banden besitzen, wurden als geeignete Marker angesehen. Von den 58 getesteten Primerpaaren konnten 10 aufgrund von fehlendem Polymorphismus ausgeschlossen werden. Von den übrigen 48 polymorphen Marker wurden 17 als Marker für *Cyperus fuscus* ausgewählt.

(Punkt 5.1). Diese wurden in weiteren Schritten direkt mit unterschiedlichen Fluoreszenzlabeln ausgestattet und in Gruppen zu jeweils 4 oder 5 Marker für Multiplex PCR-Reaktionen vorbereitet.

4.1.2.4 Zusammenfassung Methode 2

Nachfolgend sind die wichtigsten Punkte der Methode 1 tabellarisch zusammengefasst (Tabelle 13).

Tabelle 13: Zusammenfassung Methode 2

Ausgangsdaten	
Gesamtanzahl der <i>reads</i> aus 454 Pyrosequenzierung + Enrichment	4877
Anzahl an <i>reads</i> mit SSR-Sequenzen	967
Anzahl an geeigneter <i>reads</i> für Primerdesign	494
Getestete Primerpaare	80
Primer	
TAG	M13: 5' - <u>TGT AAA ACG ACG GCC AGT</u> – 3' (SCHUELKE, 2000)
PIG-tail	5'- GTTT - 3'
PCR	
Programme	PCR-Programm „Ecogenics“, <i>annealing</i> 56°
Ready Mix	<i>REDtaq® ReadyMix™</i> (Sigma-Aldrich) <i>Jumpstart™ REDtaq® ReadyMix™</i> (Sigma-Aldrich)
Fragmentanalyse	
Fluoreszenzmarkierungen	<i>FAM</i> (Applied Biosystems)
	<i>ATTO 565</i> (Sigma-Aldrich)
	<i>ATTO 550</i> (Sigma-Aldrich)
	<i>ATTO 532</i> (Sigma Aldrich)
	<i>Yakima Yellow</i> (Epoch Biosciences)
Ausgewählte SSR Marker*	17

Anm.: *Positive Amplifikation an allen Testindividuen sowie alle geforderten Kriterien erfüllt (Polymorphie, einheitliche peak-Höhe, Fehlen von stutter-banden/-A-Bande).

4.2 Anwendungstests der Marker

Mit den insgesamt 21 entwickelten, polymorphen Marker für *Cyperus fuscus* (4 aus Methode 1, 17 aus Methode 2), können nun die Hypothesen der Arbeit beantwortet werden (siehe Einleitung und 2.3). Dazu werden die ersten 8 Marker an zwei unterschiedlichen Populationen getestet. Dabei werden mithilfe der Computer-Programme *MSA - Microsatellite Analyser* (DIERINGER & SCHLÖTTERER, 2003) und *Arlequin 3.5* (EXCOFFIER & LISCHER, 2010) aus den Markerdaten die populationsgenetisch relevanten Werte *allelic richness*, Heterozygotie, Inzuchtkoeffizient und Fixationsindex (F_{ST}) errechnet und ausgewertet (Punkt 4.3.3 & Punkt 5.3 ff). Zunächst werden in Kürze die Populationen vorgestellt. Im Weiteren werden die gemessenen Werte und Programme kurz besprochen.

4.2.1 Populationen und Marker

Für die nachfolgenden Versuche werden die oberflächlichen Individuen (*O*) und die Individuen aus der Bodenfraktion zwischen 5 und 15 cm Tiefe (*B*) der Populationen mit den WHB-Nummern 62981 (Břeclav) und 62969 (Doksy) herangezogen. Ab hier werden die Populationen einfachheitshalber als Population A (für WHB: 62981) und Population B (für WHB: 62969) bezeichnet. Diese Populationen wurden ausgewählt, da sie sich grundlegend in ihrer Nutzung unterscheiden und geographisch weit entfernt gelegen sind. Außerdem sind ihre Populationsgrößen mit mehreren tausend Individuen und sehr hoher Populationsdichte (ca. 200-300 Pflanzen /m²) annähernd gleich beziehungsweise konnten annähernd gleich viele Individuen je Fraktion erfolgreich mit den ersten 8 Markern getestet werden (Punkt 5.3)(Tabelle 15). Bei Population A handelt es sich um eine natürliche Population an einem mit dem Hauptarm verbundenen Nebenarm der Thaya nahe der Österreichisch-Tschechischen Grenze bei Rabensburg im Bezirk Břeclav. Population B hingegen ist eine Population aus einem Fischhälter bei Doksy nördlich von Prag. Die beiden Populationen sind rund 270 Kilometer voneinander entfernt.



Abbildung 10: Fotos der Standorte. Linke Seite: Natürlicher Standort der Population A. Rechte Seite: Fischhälter der Population B. (Foto: J. Böckelmann)

Tabelle 14: Standort- und Nutzungsinformationen der Populationen A und B

	Population A	Population B
Nutzungstyp	Natürlicher Standort	Fischhälter
Standort	Břeclav, nahe Rabensburg	Doksy, N von Prag
Koordinaten	N 48° 40.354' E 16° 55.442'	N 50° 33.830' E 14° 39,532'
Populationsgröße	Mehrere 1000 Individuen	
Populationsdichte	ca. 200-300 Individuen/m ²	
Anzahl getestete Individuen		
Oberflächenfraktion	15	15
Bodenfraktion	16	23
Total	31	38

Die 8 Marker wurden mit Fluoreszenzlabel ausgestattet (Siehe Tabelle 15). Die Marker 1-4 aus Methode 1 wurden in einzelnen PCR-Reaktionen mit gelabelten M13R-Primern amplifiziert, während für die Marker 5-8 aus Methode 2 eine Multiplex-PCR-Reaktion mit direkt gelabelten Primern durchgeführt wurde. (Anm.: Die Multiplex PCR Reaktionen wurden von DI. J. BÖCKELMANN durchgeführt).

Tabelle 15: Verwendete Marker für vorläufige Anwendungstests.

Locus	Marker ID	Multiplex Kombination	Fluoreszenz-Farbstoff	Fragmentlängen Spannweite
1	Cf008_1_M13R	-	ATTO 565	312-344
2	Cf017_1_M13R	-	ATTO 550	218-244
3	Cf019_1_M13R	-	FAM	184-205
4	Cf104_1_M13R	-	Yakima Yellow	180-184
5	Cypfus_0173_M13ecog	1	ATTO 532	189-201*
6	Cypfus_0551_M13ecog	1	ATTO 565	205-229*
7	Cypfus_1207_M13ecog	1	FAM	138-150*
8	Cypfus_2506_M13ecog	1	FAM	218-245*

* Fragmentlängen ohne PIG-tail, mit M13 Tag.

Die Daten aus dem *Genetic Analyzer 3500* wurden mit freeware-Programmen ausgewertet und daraus die gewünschten Messwerte errechnet (Punkt 4.3.3).

4.2.2 Berechnete Werte

Zur Überprüfung der Hypothesen wurden unter Zuhilfenahme der Freeware-Programme *MSA - Microsatellite Analyser* (DIERINGER & SCHLÖTTERER, 2003) und *Arlequin 3.5* (EXCOFFIER & LISCHER, 2010) zur Auswertung populationsgenetischer Daten (Punkt 4.3.4) die folgenden Werte ermittelt:

- *allelic richness*
- Heterozygotiegrad
- Inzuchtkoeffizient
- F_{ST} Wert oder *fixation index*

Diese Werte werden häufig herangezogen um Aussagen über die genetische Vielfalt einer Art beziehungsweise von Subpopulationen einer Art zu treffen (vgl. BALLOUX & LUGON, 2002; BOOMER & STOW, 2010; BUEHLER et al., 2011; u.v.m.).

Messwerte im Detail

allelic richness

allelic richness ist ein Maß für die Anzahl der Allele pro Locus (ALLENDORF & LUIKART, 2007) und somit ein Maß für die Gesamtanzahl an Allele in einer Population. Es ist von der tatsächlich beobachteten Anzahl an Allele zu unterscheiden und stellt ein häufig gebrauchtes Maß für die genetische Vielfalt dar (vgl. LEBERG, 2002; KALINOWSKI 2005; u.v.m.). Vor allem SSR-Marker weisen eine hohe *allelic richness* auf, wodurch dieses Maß bei SSR Analysen häufig betrachtet wird (GUICHOUX et al., 2011). Nachdem die beobachtete Anzahl der Allele abhängig von der Größe der untersuchten Population ist, muss bei ungleich großen Populationen das vorhandene Stichprobenungleichgewicht ausgeglichen werden. (RODRIGÁÑEZ et al., 2008; KALINOWSKI, 2005; FOULLEY & OLLIVIER, 2006). Dazu können zwei statistische Methoden verwendet werden. Diese sind die *rarefaction method* (KALINOWSKI, 2005) und die *extrapolation method* (FOULLEY & OLLIVIER, 2006). Daraus resultiert eine statistisch berechnete Schätzung der insgesamt erwarteten Allele pro Locus bezogen auf eine bestimmte, zuvor festgelegte Anzahl an Individuen in der Population. Als Indikator für die genetische Vielfalt wird nun die über alle Loci gemittelte *allelic richness* der Populationen verglichen. Dabei gilt grundlegend, dass höhere Werte der *allelic richness*, also viele Allele pro Locus, auf eine höhere genetische Vielfalt hindeuten und umgekehrt (INTERNATIONAL LIVESTOCK RESEARCH INSTITUTE, 2011).

Heterozygotiegrad / Heterozygotie

Der Heterozygotiegrad ist ein häufig verwendetes Maß für die genetische Varianz von Populationen (LEBERG, 2002). Dieser ist wie die *allelic richness*, nicht von der Stichprobengröße der untersuchten Population beeinflusst (KALINOWSKI, 2005). Der Heterozygotiegrad bezeichnet die Anzahl an Loci an denen ein Individuum heterozygot ist. Es wird zwischen der beobachteten Heterozygotie (H_{OBS}) und der erwarteten Heterozygotie (H_{EXP}) unterschieden. Letztere wird aus der H_{OBS} unter Zuhilfenahme des *Hardy-Weinberg Gleichgewichts* berechnet (OJANGO et al., 2011). Die Werte schwanken zwischen 0 und 1 wobei 0 völlige Homozygotie und 1 vollständige Heterozygotie aufzeigt. Geringe Heterozygotie kann einerseits durch Isolation einer Population oder durch eine geringe Populationsgröße und Inzuchteffekte erklärt werden (OJANGO et al., 2011; McDONALD, 2008). Hohe Werte können hingegen einen Genaustausch zwischen Populationen zeigen.

Inzuchtkoeffizient [F_{IS}]

Dieser Wert gibt den Verlust an Heterozygotie aufgrund nicht zufälliger Paarung von Individuen einer Subpopulation wieder und zeigt das Ausmaß genetischer Inzucht auf (SHANE'S SIMPLE GUIDE TO F STATISTICS). Er ist definiert als die Wahrscheinlichkeit, dass zwei Allele an einem Locus aufgrund von Vererbung identisch sind (identical by descent) also die Wahrscheinlichkeit, dass zwei Individuen an einem Locus die gleichen Allele wie ein gemeinsamer Vorfahre besitzen (WRIGHT, 1922; LEUTENEGGER, et al., 2003). Die Werte reichen von -1 (alle beobachtete Individuen sind heterozygot), bis +1 (alle Individuen sind vollständig homozygot). Ebenso können mit dem Inzuchtkoeffizienten Aussagen über die Art der Befruchtung getroffen werden, wobei hohe Werte auf Selbstbefruchtung und steigende Homozygotie hindeuten und umgekehrt. Nicht außer Acht zu lassen ist dabei jedoch, dass *Cyperus fuscus* selbstkompatibel ist (EAST, 1940) und möglicherweise von sich aus bereits hohe Inzuchtwerte aufweist.

F_{ST} Wert / fixation index

Der F_{ST} -Wert oder Fixationsindex (*fixation index*) zeigt die Reduktion der durchschnittlichen Heterozygotie von Subpopulationen innerhalb der Gesamtpopulation relativ zur durchschnittlichen Heterozygotie der Gesamtpopulation (SHANE'S SIMPLE GUIDE TO F STATISTICS). Dies bedeutet, dass mit dem F_{ST} -Wert Inzuchteffekte aufgrund der Korrelation von Allelen innerhalb der Subpopulation gemessen werden können (BALLOUX & LUGON, 2002). Es wird folgendermaßen berechnet:

$$F_{ST} = \frac{(H_T - H_S)}{H_T}$$

H= Heterozygotie

S= Subpopulation

T= Total

Es ist somit ein weiteres Maß für die genetische Differenzierung von Subpopulationen. Es können Werte zwischen 0 (keine genetische Differenzierung) und theoretisch 1 (vollständige genetische Differenzierung) auftreten. Dabei ist jedoch zu beachten, dass die maximale Differenzierung nicht größer als die beobachtete Homozygotie sein kann. Es werden folgende Werte aus BALLOUX & LUGON, 2003 zur Beurteilung der genetischen Differenzierung herangezogen:

0,0 bis 0,05	schwache gen. Diff.
0,05 bis 0,15	mäßige gen. Diff.
0,15 bis 0,25	starke gen. Diff.
über 0,25	sehr starke gen. Diff.

(BALLOUX & LUGON, 2003)

Die genannten Richtwerte sind jedoch nur sehr allgemein zu betrachten. Je nachdem welche Marker verwendet wurden, können die Werte deutlich von den Richtwerten abweichen.

4.2.3 Verwendete Programme

Zur Berechnung wurden die Programme *MSA - Microsatellite Analyser* (DIERINGER & SCHLÖTTERER, 2003) und *Arlequin 3.5* (EXCOFFIER & LISCHER, 2010) verwendet. Diese können aus den unterschieden der Fragmentlängen der gefundenen Allele, unter Angabe der Größe der Wiederholungseinheit des betroffenen Mikrosatelliten, die Allele und folgend die oben genannten Werte errechnen. Die dazu erforderlichen Daten werden direkt aus den Daten aus *Genemarker* übernommen. Für jedes getestete Individuum werden die Fragmentlängen der beiden Allele in Basen angegeben. Bei homozygoten Individuen werden dabei zwei gleiche Werte und bei heterozygoten entsprechend zwei unterschiedliche Fragmentlängen angegeben. Nach der Dateneingabe berechnen die Programme die gewünschten Werte:

<i>allelic richness</i>	Anzahl Allele (standardisiert auf 10-15 Standardindividuen)
Heterozygotiegrad:	beobachtete Heterozygotie [H_{OBS}] erwartete Heterozygotie [H_{EXP}]
Inzuchtkoeffizient	[F_{IS}]
<i>fixation index</i>	[F_{ST}]

5 Ergebnisse

In diesem Kapitel werden die im Zuge der Arbeit entwickelten Mikrosatelliten-Marker für *Cyperus fuscus* präsentiert, die Motivhäufigkeiten in den gefundenen Mikrosatelliten aufgezeigt sowie die Ergebnisse der Anwendungstests dargelegt.

5.1 Entwickelte Mikrosatelliten-Marker für *Cyperus fuscus*

Im Zuge dieser Arbeit wurden insgesamt 21 polymorphe Mikrosatelliten-Marker für *Cyperus fuscus* entwickelt (Tabelle 16). (Anm.: Davon wurden 7 von DI. J. BÖCKELMANN entwickelt). Von den 21 Markern wurden 4 durch Methode 1 und 17 durch Methode 2 entwickelt.

Marker Methode 1

Es konnten mit MSATCOMMANDER insgesamt 67 Primerpaare gefunden werden. Davon wurden 42 für weitere Tests ausgewählt. Bei 15 Primerpaaren konnte dabei eine positive Amplifikation an allen getesteten Individuen festgestellt werden. Bei den nachfolgenden Tests an 7 Individuen am *Genetic Analyzer 3500* konnten 11 Primerpaare aufgrund fehlendem Polymorphismus ausgeschlossen werden. Es verbleiben somit die 4 Marker Cf_008, Cf_017, Cf_019 und Cf_104. (Tabelle 16, Nr.1-4). Diese sind mit dem M13R Tag (GGAAACAGCTATGACCAT) aus MSATCOMMANDER sowie mit *PIG-tails* (GTTT) ausgestattet.

Marker Methode 2

Aus den insgesamt 4877 *reads* konnten 967 Sequenzen gefunden werden die Mikrosatelliten mit di- tri- oder tetranukleotiden *repeats* enthielten, mit jeweils mindestens 6 *repeats* für tri- und tetranukleotide und 10 *repeats* für dinukleotide Motive. 494 dieser *reads* eigneten sich für ein Primerdesign. Davon wurden 80 selektiert und von Microsynth (Microsynth AG, Balgach, Schweiz) synthetisiert. Bei den nachfolgenden Tests an 7 Individuen am *Genetic Analyzer 3500* konnten 48 polymorphe Loci gefunden werden. Davon wurden 17 als SSR Marker optimiert (Tabelle 16, Nr.5-21). Die *forward*-Primer aus Methode 2 besitzen an ihrem 5'Ende ausnahmslos den M13-*tag* (TGTAACACGACGGCCAGT) aus SCHUELKE 2000. Alle Reverse-Primer wurden mit *PIG-tails* (GTTT) ausgestattet.

Tabelle 16: Übersicht über die 21 entwickelten Mikrosatelliten-Marker für *Cyperus fuscus*

Nr.	Methode	Bezeichnung	Sequenz (5' → 3')
1	1	Cf_008	F: <u>GGAAACAGCTATGACCATAGATAATTAACGGATCAGGGACG</u> R: <u>GTTT</u> GAGACAGATTACTCACCTCTCAAG
2	1	Cf_017	F: <u>GGAAACAGCTATGACCATGAGGCAATAGAAATTGTTGGAG</u> R: <u>GTTT</u> ACGAAATGAGGAGCCATAACTG
3	1	Cf_019	F: <u>GTTT</u> AATTGTCAGGCCACATGCC R: <u>GGAAACAGCTATGACCATA</u> CAGGGAGCAACCTGAGC
4	1	Cf_104	F: <u>GGAAACAGCTATGACCATGACAGAAATGAATTAAGGCCAC</u> R: <u>GTTT</u> CGATGACAGTTTAAAGGTCCAG
5	2	Cypfus_0173	F: <u>TGTA AACGACGGCCAGT</u> CGCCAAAGGAGAATGAGGTG R: <u>GTTT</u> ATCGAACAATCCGATCTCGC
6	2	Cypfus_0551	F: <u>TGTA AACGACGGCCAGT</u> TTCACATTGACGCACAC R: <u>GTTT</u> AGCGTGCTATTACAACCTTGG
7	2	Cypfus_1207	F: <u>TGTA AACGACGGCCAGT</u> ATCTTCTCACTCCCGCCATC R: <u>GTTT</u> GGAGTAAACCACGGACTCG
8	2	Cypfus_2506	F: <u>TGTA AACGACGGCCAGT</u> ACCCTAACGACTGCATCACC R: <u>GTTT</u> AAATCTTGCCGTCTTCACCG
9	2	Cypfus_3114	F: <u>TGTA AACGACGGCCAGT</u> TCCCGACTTCTCCCAATTC R: <u>GTTT</u> AGCTCGCAGCATACTAGAC
10	2	Cypfus_3300	F: <u>TGTA AACGACGGCCAGT</u> TTTTGTCTGGTTCACGGG R: <u>GTTT</u> AGGTCCTCATTCTCTTCACCG
11	2	Cypfus_4093	F: <u>TGTA AACGACGGCCAGT</u> GTCTCTCCAAACAGGAGGGC R: <u>GTTT</u> GTACAGGTAAGCGCAAGAGC
12	2	Cypfus_4216	F: <u>TGTA AACGACGGCCAGT</u> GTTGTGAAAACCTAGGCGG R: <u>GTTT</u> ATTGAGGCCAGCACAAAC
13	2	Cypfus_4666	F: <u>TGTA AACGACGGCCAGT</u> GGGTGTTGCATGACTGTAGC R: <u>GTTT</u> CGTAAGGGTACATAAGTCGATCC
14	2	Cypfus_2663	F: <u>TGTA AACGACGGCCAGT</u> TGCAATTAAGCCGTCCAG R: <u>GTTT</u> ACCTCCCTATGAGGTTCTTAGC
15	2	Cypfus_2987	F: <u>TGTA AACGACGGCCAGT</u> ACGGATTCTTCTCACACCC R: <u>GTTT</u> GCACGATGCTGCCTATACTTG
16	2	Cypfus_3212	F: <u>TGTA AACGACGGCCAGT</u> ACACCTAAAAGCGAAAGCGG R: <u>GTTT</u> GACCGAAAGACGCTTGAAC
17	2	Cypfus_3921	F: <u>TGTA AACGACGGCCAGT</u> ATGGATGACGAGGAGGTTGG R: <u>GTTT</u> GTAGAGGGAGGTTGGTAGCG
18	2	Cypfus_2257	F: <u>TGTA AACGACGGCCAGT</u> AACCAGAGAAGTCCAGGTGC R: <u>GTTT</u> GGGTCCCAGTCTCTGACATC
19	2	Cypfus_2993	F: <u>TGTA AACGACGGCCAGT</u> ATCGACTGCAAAGCATAGGG R: <u>GTTT</u> GGCCTCGGTCAGTTCTAC
20	2	Cypfus_3218	F: <u>TGTA AACGACGGCCAGT</u> TGCTCTCTCCAACAAGC R: <u>GTTT</u> GAAATTCAACGGAGAGCGGG
21	2	Cypfus_4236	F: <u>TGTA AACGACGGCCAGT</u> GCTGTACGTGGAGAGAGGAG R: <u>GTTT</u> AAATCCACCGTCGCAAATCC

Anm. zu Tabelle 16: *F* = forward-primer; *R*: reverse-primer; Die unterstrichenen Sequenzabschnitte unterlegen angefügte M13R- bzw. M13- tag's sowie PIG-tails.

Die gefundenen Marker amplifizierten Fragmente zwischen 94 und 344 Basenpaaren Länge. Die Durchschnittslänge aller amplifizierten Fragmente der 7 Testindividuen beträgt 207 Basenpaare. An den 7 getesteten Individuen konnten dabei zwischen 2 und 6 Allele je Marker gefunden werden. Durchschnittlich wurden 3 Allele pro Marker gefunden (Tabelle 17).

Tabelle 17: Repeat Motive, Spannweite der Fragmentlängen (Allele size range) und Anzahl gefundener Allele je Marker der 21 entwickelten Mikrosatelliten-Marker

Nr.	Bezeichnung	Repeat Motiv	Fragmentlänge – Spannweite (min-max) [Bp]	Anzahl Allele an 7 Testindividuen
1	Cf_008	(AG) ₁₁	312-344	4
2	Cf_017	(CTTT) ₁₃	218-242	3
3	Cf_019	(CTT) ₇ +(CTT) ₆	184-205	2
4	Cf_104	(GT) ₁₄	185-189	2
5	Cypfus_0173	(GAA) ₉	189-201*	3
6	Cypfus_0551	(TGTA) ₉	205-229*	2
7	Cypfus_1207	(CAG) ₇	138-150*	3
8	Cypfus_2506	(TTC) ₁₂	218-245*	4
9	Cypfus_3114	(CT) ₁₅	160-180*	4
10	Cypfus_3300	(GTAT) ₁₃	232-249*	3
11	Cypfus_4093	(GA) ₁₃	94-98*	2
12	Cypfus_4216	(TTC) ₂₀	183-214*	6
13	Cypfus_4666	(TATG) ₇	189-221*	3
14	Cypfus_2663	(CATA) ₇	230-242*	3
15	Cypfus_2987	(CTT) ₉	248-264*	4
16	Cypfus_3212	(AAG) ₈	209-226*	3
17	Cypfus_3921	(CGC) ₈	261-270*	3
18	Cypfus_2257	(CT) ₁₃	230-236*	3
19	Cypfus_2993	(GAA) ₈	142-162*	3
20	Cypfus_3218	(CTT) ₉	163-194*	3
21	Cypfus_4236	(AG) ₁₂	177-185*	3

Anm. zu Tabelle 17: * = Fragmentlängen *inklusive* M13 Tag und *ohne* PIG-tail.

Die Repeat Motive der Mikrosatelliten sind di-, tri- und tetranukleotide mit mindestens 7 und maximal 20 *repeats* (Durchschnitt 10).

Motivhäufigkeiten

Bei der Suche nach geeigneten Mikrosatelliten in Methode 1 konnten die relativen Häufigkeiten der Mikrosatellitenmotive ermittelt werden (Tabelle 18). Es zeigt sich, dass dinukleotide Motive etwa doppelt so häufig sind wie trinukleotide Motive und dass tetranukleotide Motive erwartungsgemäß eher selten vorkommen. Das bei weitem häufigste Motiv bei mindestens 6 Wiederholungseinheiten ist AG mit insgesamt 663 gefundenen Mikrosatelliten, gefolgt von AC mit 373 Mikrosatelliten. Die trinukleotiden Motive werden bei mindestens 4 Wiederholungseinheiten angeführt von CTT mit 249 Mikrosatelliten und GTT mit 210 Mikrosatelliten. Das tetranukleotide Motiv ATGT ist bei mindestens 4 Wiederholungseinheiten mit insgesamt 28 Mikrosatelliten deutlich häufiger als alle anderen gefundenen tetranukleotiden Motive.

Repeat-Art	Motiv (reverse Motiv)	Häufigkeit
dinukleotide	AG (CT)	663
	AC (GT)	373
	AT (AT)	296
trinukleotide	CTT (AAG)	249
	GTT (AAC)	210
	AAT (ATT)	179
	ATC (GAT)	163
	GGT (ACC)	155
	GCT (AGC)	101
	AGG (CCT)	64
	ACT (AGT)	54
	CGG (CCG)	29
	ACG (CGT)	7
	tetranukleotide	ATGT (ACAT)
ATTT (AAAT)		12
CTTT (AAAG)		12
CATT (AATG)		7
ACAG (CTGT)		7

Tabelle 18: Motivhäufigkeiten der Mikrosatelliten aus Methode 1

5.2 Ergebnisse zu den Anwendungstests der Marker

Nachfolgend werden die Ergebnisse der Anwendungstests der Marker dargestellt. Zunächst soll die Frage zum Ploidiegrad der Art geklärt werden. Die berechneten Werte zur *allelic richness*, dem Heterozygotiegrad, dem Inzuchtkoeffizienten und den F_{ST} -Wert werden präsentiert.

Ploidiegrad

Während der Markerentwicklung und der Marker-Tests am *Genetic Analyzer 3500* konnte festgestellt werden, dass bei *Cyperus fuscus* ausnahmslos maximal zwei Allele je Locus und Individuum auftreten. Auch an den Elektropherogrammen, wie beispielsweise in Abbildung 11, ist bei allen getesteten Individuen das für Diploide typische Muster ersichtlich, dass bei zwei Allelen jeweils das kürzere Allel eine etwas höhere Signalintensität aufweist, wie das längere Allel. Dies beruht darauf, dass bei der PCR-Amplifikation ein kürzeres DNA-Fragment einen geringen Vorteil gegenüber einem längeren hat, indem es etwa schneller amplifiziert wird. Im Gegensatz dazu würden bei Tetraploiden mit 2 Allelen unterschiedliche Verhältnisse der Allel-Intensitäten in Erscheinung treten und somit vom obigen Muster der Signal-Intensitäten abweichen. Es könnten beispielsweise 3 Kopien des kurzen Allels und 1 Kopie des langen Allels im Verhältnis 3:1 auftreten und somit deutlich unterschiedliche Allel-Intensitäten aufzeigen. Daneben wären die Verhältnisse 2:2 und 1:3 genauso erkennbar, womit 4 Chromosomensätze anzunehmen wären.

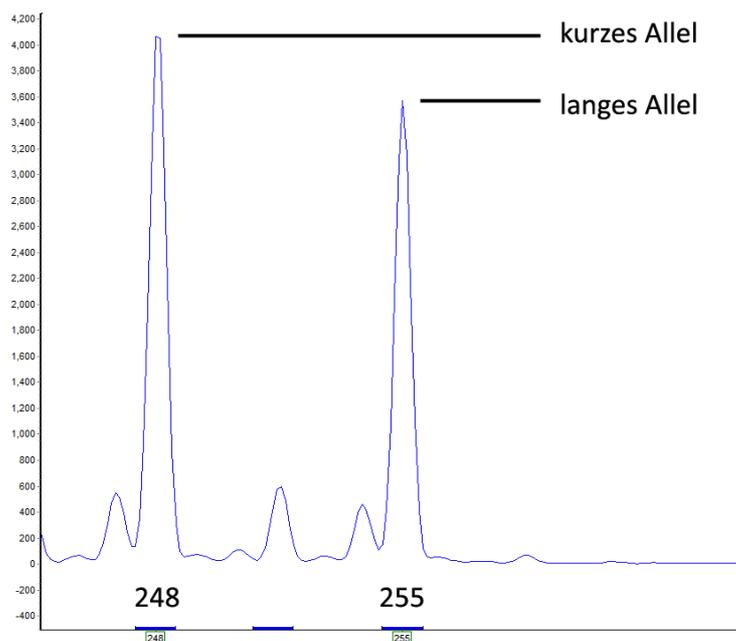


Abbildung 11: Typische Unterschiede der Signalintensität diploider Organismen

Anzahl Allele und *allelic richness*

In Tabelle 19 ist die Anzahl der gefundenen Allele je Locus der Oberflächenfraktionen und der Bodenfraktionen 5-15 cm Tiefe der Populationen A und B aufgezeigt.

Tabelle 19: Gefundene Allele je Locus und Fraktion der Populationen A und B

	Population A		Population B	
Locus*	Fraktion O	Fraktion B	Fraktion O	Fraktion B
1	3	4	2	1
2	3	2	2	4
3	4	3	3	2
4	2	2	2	2
5	5	3	1	3
6	3	3	3	3
7	3	3	2	3
8	3	4	3	3
Summe	26	22	18	21
Durchschnitt**	3,25	3	2,5	2,62

Anm: * (siehe Tabelle 15); ** Allele pro Bodenfraktion

Es konnten für die Population A insgesamt 27 Allele und bei Population B 24 Allele gefunden werden. Dies ergibt eine *allelic richness* von 3,23 bei Population A und 2,62 für Population B (Tabelle 20).

Tabelle 20: *allelic richness* und Gesamtanzahl gefundener Allele der Populationen A und B

	Population A	Population B
<i>allelic richness</i> *	3,23	2,62
s. d.**	0,1	0,2
Σ gefundener Allele	27	24

Anm. zu Tabelle 20: * basierend auf 10 - 15 Standardindividuen; ** Standardabweichung über alle Loci.

Heterozygotie

Es wurden für die Populationen A und B sowohl die beobachtete Heterozygotie (H_{OBS}) als auch die erwartete Heterozygotie (H_{EXP}) berechnet. Dabei wurden die Heterozygotiewerte für 8 Loci (Fehler: Referenz nicht gefunden) jeweils an der Oberflächenfraktion O und der Bodenfraktionen B berechnet (Tabelle 21).

Tabelle 21: Werte für beobachtete (H_{OBS}) und erwartete (H_{EXP}) Heterozygotie für die Oberflächen Fraktion O und Bodenfraktion B der Populationen A und B an 8 Loci

		Population A		Population B	
Locus*		H_{OBS}	H_{EXP}	H_{OBS}	H_{EXP}
Fraktion O	1	0,214	0,626	0,076	0,076
	2	0,153	0,578	0	0,415
	3	0,266	0,613	0	0,370
	4	0,200	0,526	0,071	0,452
	5	0,071	0,716	0	0
	6	0,133	0,570	0,066	0,197
	7	0	0,634	0,200	0,434
	8	0,066	0,544	0,266	0,452
Durchschnitt		0,137	0,600	0,084	0,299
s. d.**		0,088	0,060	0,099	0,147
Fraktion B	1	0,267	0,742	0	0
	2	0,125	0,508	0,052	0,698
	3	0,071	0,658	0,045	0,206
	4	0,133	0,514	0,086	0,394
	5	0	0,548	0	0,479
	6	0,125	0,522	0	0,166
	7	0,200	0,627	0,043	0,505
	8	0,125	0,483	0,043	0,528
Durchschnitt		0,130	0,575	0,033	0,372
s. d.**		0,079	0,090	0,030	0,187
Gesamtdurchschnitt		0,134	0,588	0,059	0,336

Anm. zu (Tabelle 21): * (siehe Tabelle 15); ** Standardabweichung über alle Loci; $H_{OBS} = 0$: am Locus x sind an keinem getesteten Individuum Heterozygote Allele vorgefunden worden. $H_{EXP} = 0$: Alle Allele des Locus x sind homozygot und besitzen an allen getesteten Individuen den selben Wert, d.h. die Berechnung der H_{EXP} ergibt 0.

Die in der Tabelle aufgezeigten Werte für H_{OBS} und H_{EXP} zeigen teilweise große Unterschiede. Diese rühren daher, dass die Größendifferenz der Allele eines Locus unterschiedlich in die Berechnung der Heterozygotiewerte einfließen. Ein kleiner Wert für H_{OBS} zeigt, dass an diesem Locus in einer Population nur wenige heterozygote Allele vorhanden sind. Beispielsweise sind an Population A, Fraktion B, Locus 8 lediglich 2 von 16 Individuen heterozygot und die Differenz beträgt in beiden Individuen 3 Basenpaare bzw. 1 repeat. Dadurch ergibt sich hier ein relativ geringer beobachteter Heterozygotiewert von 0,125. Der relativ kleine Wert 0,483 für H_{EXP} zeigt in diesem Fall eine geringe Größendifferenz aller Allele der Fraktion. Betrachtet man jedoch beispielsweise den Locus 5 der Fraktion O von Population A sind die Unterschiede zwischen H_{OBS} und H_{EXP} sehr deutlich. Ein einziges heterozygoten Individuum bewirkt 0,071 für H_{OBS} , jedoch unterscheiden sich die homozygoten Individuen untereinander in ihrer Allelgröße sehr stark, wodurch der Wert 0,716 für H_{EXP} zustande kommt.

Inzuchtkoeffizient

Bezüglich der Inzuchtkoeffizienten der beiden Populationen A und B wurden wiederum 8 Loci (Tabelle 15) jeweils an der Oberflächenfraktion O und der Bodenfraktion B berechnet (Tabelle 22).

Tabelle 22: Inzuchtkoeffizienten der Populationen A und B an 8 Loci und zwei Bodenfraktionen O und B

		Population A	Population B
	Locus*	F_{IS}	F_{IS}
Fraktion O	1	0,656	-0,020**
	2	0,733	1
	3	0,562	1
	4	0,616	0,842
	5	0,900	n. d.
	6	0,765	0,660
	7	1	0,536
	8	0,877	0,405
Durchschnitt		0,763	0,557
s. d. ***		0,152	0,353
Fraktion B	1	0,639	n. d.
	2	0,753	0,925
	3	0,891	0,779
	4	0,740	0,779
	5	1	1
	6	0,760	1
	7	0,680	0,914
	8	0,741	0,918
Durchschnitt		0,775	0,789
s. d. ***		0,116	0,091
Gesamtdurchschnitt		0,769	0,770

Anm. zu Tabelle 22: *(siehe Tabelle 15); ** negativ weil H_{OBS} größer als H_{EXP} , in diesem Fall mit 0 gleichzusetzen; *** Standardabweichung über alle Loci; n. d. = F_{IS} nicht definierbar ($H_{EXP} = 0$, d. h. die Population ist an diesem Locus monomorph = alle Individuen sind homozygot für dasselbe Allel); 1 = vollständige Inzucht (alle Individuen sind homozygot).

F_{ST}-Wert

Mittels Arlequin 3.5 (EXCOFFIER & LISCHER, 2010) wurden für die beiden

Populationen A und B die jeweiligen F_{ST} -Werte errechnet. Dazu wurden die Bodenfraktionen O und B der einzelnen Populationen als Subpopulationen in die Berechnung integriert (Tabelle 23).

Tabelle 23: Berechnete F_{ST} -Werte für die Subpopulationen der Populationen A und B

	Population A	Population B
F_{ST}	0,013	0,114
Signifikanzschwelle*	0,63	0,013

Anm.: * Berechnung basierend auf 10100 Permutationen. Allowed level of missing data = 0,5 (=50%)

6 Diskussion

6.1 Vergleich der verwendeten Methoden zum Primerdesign

Vergleich der Read-längen und Markerqualität

Die beiden Methoden weisen deutliche Unterschiede in der Anzahl an gewonnenen SSR Marker auf. Dies ist im wesentlichen auf zwei Punkte zurückzuführen:

- Länge der *reads* der Pyrosequenzierung
- SSR Enrichment in Methode 2

Obwohl die verwendeten Methoden zum Primerdesign sehr ähnlich sind, konnten wesentlich mehr Primer mit Methode 2 entwickelt werden. Nachdem beide Methoden das Programm PRIMER3 verwenden, kann davon ausgegangen werden, dass der Grund für diese Unterschiede nicht im Primerdesign selbst liegt. Mit sehr großer Wahrscheinlichkeit ist die unterschiedliche Länge der *reads* der Pyrosequenzierungen in den Methoden ein Grund für die unterschiedliche Erfolgsrate.

Tabelle 24: Übersicht der 454 Pyrosequenzierungsdaten der Methoden 1 und 2

Methode	Sequenzierungsmethode	Read Gesamtanzahl	Durchschn. Read Länge	Min. Länge	Max. Länge
Methode 1 - LGC	454 GS FLX+ Titanium	143027	238	7	762
Methode 2 - Ecogenics	454 GS FLX+ Titanium + Enrichment	4877	415	34	801

Beim Betrachten der Sequenzierungsdaten (Tabelle 24 & Abbildung 12) ist ersichtlich, dass ein Großteil der Fragmente in Methode 1 mit durchschnittlich 238 Basenpaaren Länge deutlich unter den Angaben des Herstellers (Roche diagnostics, Basel, Schweiz) liegen. Laut diesen sollten mindestens 85% aller *reads* eine Länge von mehr als 500 Basenpaare sowie 45% eine Länge von mindestens 700 Basenpaare aufweisen. Als Folge davon ist eine erhebliche Anzahl an *reads* zu kurz um ausreichend große flankierende Re-

gionen für die Entwicklung von Primer zu bieten. In Methode 2 hingegen konnte annähernd eine durchschnittliche Länge erreicht werden, die den Angaben der Hersteller entspricht. (Abbildung 12)

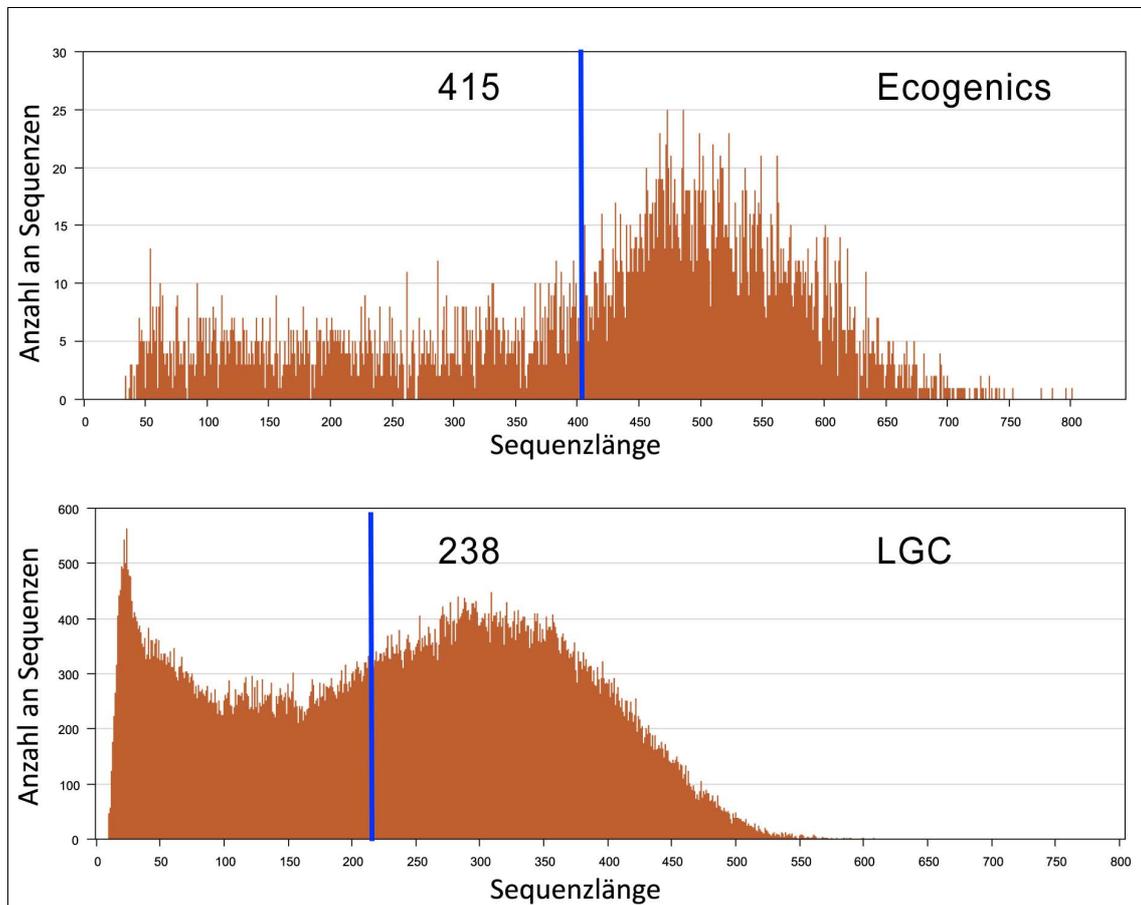


Abbildung 12: Length-graphs und durchschnittliche Read-Länge der beiden Pyrosequenzierungen durch Ecogenics und LGC

Anm.: Die blaue Linie markiert die durchschnittliche Read-Länge.

Ein weiterer, ausschlaggebender Punkt ist das Enrichment in Methode 2, das von großem Vorteil beim Suchen nach SSR Motiven im Genom von *Cyperus fuscus*, wie auch bei anderen *non-model species*, zu sein scheint (vgl. GUICHOUX et al., 2011). Dies bestätigt auch die Analyse der Pyrosequenzierungsdaten (Abbildung 12). Dabei wird klar, dass aufgrund des Enrichments auch bei wesentlich geringerer Anzahl an Gesamtreads in Methode 2 eine weitaus höhere Dichte an SSR Motiven zu finden ist. Das Enrichment ist somit ein äußerst wichtiges Instrument für das Entwickeln von SSR Marker bei *Cyperus fuscus*. Ein weiterer Unterschied der Methoden ist die Verwendung unterschiedlicher Tags. Während die unterschiedlichen M13, M13R oder CAG Tags kaum Unterschiede in Ihrer Funktion beziehungsweise Wirkung

erkennen lassen, ist die Verwendung von *PIG-tails* in Methode 1 mit wesentlichen Qualitätssteigerungen der Primer verbunden. Dadurch konnten -A-Banden erfolgreich beseitigt und das eindeutige Identifizieren von Allelen deutlich erleichtert werden. Dies wird vor allem dann klar, wenn die Graphen am *Genetic Analyzer 3500* verglichen werden. Bei Methode 2 wurden dabei nicht selten -A-Banden sichtbar (Abbildung 13). Aus diesem Grund wurden jene Primer, bei denen -A-Banden auftraten mit einem *PIG-tail* versehen um dadurch eine wesentliche Steigerung der Markerqualität zu erzielen.

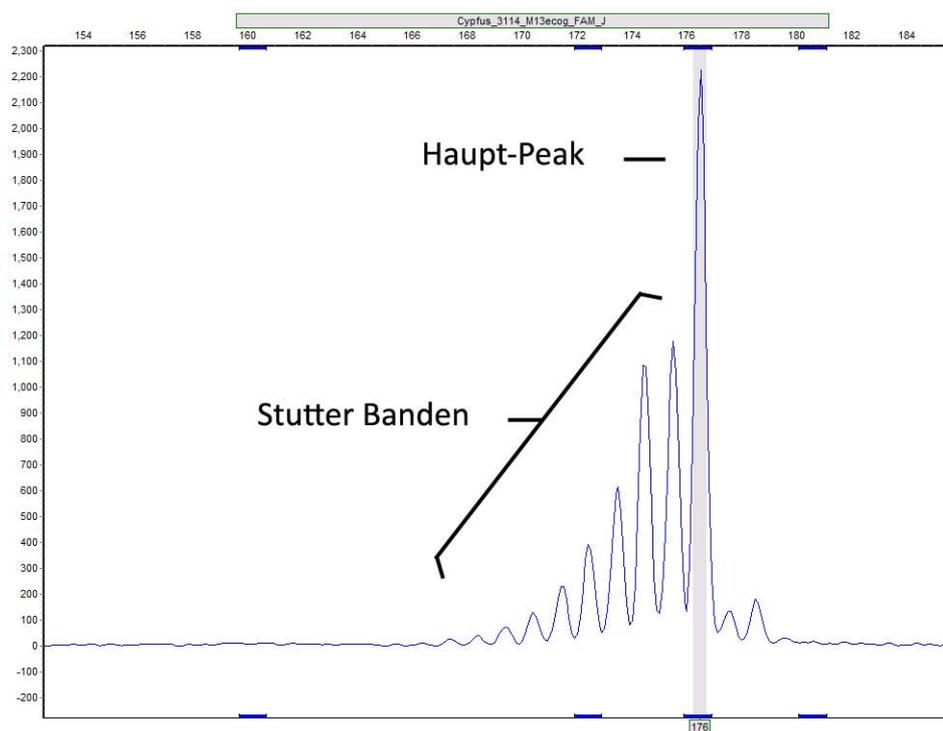


Abbildung 13: Length graph des Markers Cypfus_3114 ohne PIG-tail: Deutliche stutter und -A-Banden hinter Haupt-peak sichtbar.

Verwendung freier Software

Obwohl die Verwendung freier Software durchaus eine Möglichkeit darstellt, um SSR Marker für wenig untersuchte Arten herzustellen, sprechen die Ergebnisse klar für die Verwendung professionell entwickelter Marker mit vorangehendem Enrichment. Der Zeit- und Arbeitsaufwand sind bei Methode 1 um ein Vielfaches größer als in Methode 2. Es ist dabei zu bedenken, dass die Entwicklung von Primern mittels MSATCOMMANDER und PRIMER3 in zahlreiche Schritte gegliedert ist. Dazu zählen beispielsweise das manuelle Durchsuchen der PRIMER3 output-Dateien auf potentiell interessante Primer u. v. m.. Ebenso müssen die gefundenen Primer zur Synthese

an spezialisierte Betriebe gesendet werden und nachfolgend mit ihrer jeweiligen *annealing* Temperatur in mehreren PCR Reaktionen getestet werden. Diese sind somit zahlreiche Schritte, die den Prozess verlangsamen und den Arbeitsaufwand steigen lassen. Ein weiterer wichtiger Nachteil der Markerentwicklung mit freier Software ist die Tatsache, dass mit jedem Schritt die Fehlerquote durch diverse Faktoren, beispielsweise bei der teilweise subjektiven Auswahl von Primerkandidaten oder simple Tippfehler bei der Dateneingabe, erhöht wird. Diese Fehler beziehungsweise subjektiven Einschätzungen sind im Nachhinein nur schwer festzustellen oder zu quantifizieren. Vergleichsweise sind die Arbeitsschritte bei Methode 2 deutlich geringer. Ein großer Vorteil ist dabei das Designen der Primer mit einer vorher definierten, einheitlichen *annealing* Temperatur. Somit können alle Primer mit dem selben PCR-Programm verwendet werden. Das heißt, dass bei eventuellen Problemen bei der PCR Reaktion, alle Primer in gleicher Weise betroffen sind und alle nachfolgenden Programmkorrekturen das Ergebnis aller Primer im gleichen Maße beeinflussen.

Die wohl bedeutsamsten Unterschiede der zwei Methoden liegt in der Qualität der Sequenzdaten aus 454 Pyrosequenzierung, wie aus Tabelle 24 und Abbildung 12 ersichtlich, sind nicht nur die Anzahl an brauchbarer *reads* im Verhältnis der Gesamtanzahl an *reads* deutlich verschieden sondern auch die durchschnittliche Länge der *reads*. Diese ist ausschlaggebend für eine erfolgreiche Markerentwicklung. Aus welchem Grund die Pyrosequenzierung in Methode 1 nicht die erwünschten Ergebnisse erbrachte ist nicht bekannt, es können jedoch auch bei der durch die Firma durchgeführten Vorbereitungen der Probe für die Pyrosequenzierung Fehler aufgetreten sein. Es ist wahrscheinlich, dass die relativ geringe Ausbeute von lediglich 4 Marker in Methode 1 den qualitativ mäßigen Sequenzierungsdaten zu verschulden sind.

Ein weiterer wesentlicher Unterschied ist das Enrichment in Methode 2. Dafür wurden von ECOGENICS angebotene Standardmotive CT und GT verwendet und zusätzlich die häufigsten Motive aus Methode 1 AAG und ATGT. Dadurch konnte ein effizientes Enrichment mit einer zufriedenstellenden Pyrosequenzierung durchgeführt werden. Es zeigt sich beim Betrachten der daraus gewonnenen Marker, dass die verwendeten Motive tatsächlich die beste Markerausbeute bewirkten. Es wurden jedoch zusätzlich auch andere Motive gefunden, für die nicht speziell angereichert worden ist.

Im Zuge der Markerentwicklung zeigte sich deutlich, dass die Verwendung unterschiedlicher Tags keine Vorteile gegenüber der Verwendung eines einzigen Tags bietet. Alle 21 entwickelten Marker wurden letztlich mit zwei

Varianten des M13 Tags (M13R und M13) getestet. Der *CAG-tag* aus MSATCOMMANDER hat sich in dieser Hinsicht nicht bewährt. Es ist ebenso zu erwähnen, dass das Anheften eines Primer-*PIG-tails* eine deutliche Qualitätssteigerung zur Folge hatte. Das Anheften der +A-Banden konnte erfolgreich bewerkstelligt werden und bewirkte eine wesentliche Erleichterung beim *scoring* der Allele.

Abschließend kann behauptet werden, dass ein vorangehendes Enrichment, auch ohne dem Wissen über Motivhäufigkeiten, deutliche Vorteile bei der Entwicklung von Markern bringen kann. Ebenso bewirkt ein einheitliches Design mit einer definierten *annealing*-Temperatur und einheitlichen *tags* sowie *PIG-tails* eine deutliche Ersparnis des Zeit- und Arbeitsaufwandes und gleichzeitig die Reduktion möglicher Fehlerquellen.

6.2 Beantwortung der Hypothesen

In den folgenden Punkten werden die Ergebnisse der Arbeit in Hinblick auf die Beantwortung der Hypothesen und die verwendeten Methoden diskutiert.

6.2.1 Ploidiegrad

Wie bereits in den Ergebnissen geschildert, traten bei der Art *Cyperus fuscus* während der Markerentwicklung und den durchgeführten Versuche ausnahmslos maximal zwei Allele pro Locus auf. Obwohl die nachgewiesene Chromosomenanzahl $2n = 36$ auch höhere Ploidiegrade zulassen würde, wird durch die Markeranalyse zweifelsfrei bestätigt, dass es sich bei *Cyperus fuscus* um eine diploide Art handelt. Abweichungen von diesen Ergebnissen könnten demnach auf seltene Mutanten oder wahrscheinlicher auf verunreinigte Proben hinweisen.

6.2.2 Ökosystem-Störung erhöht die genetische Vielfalt

Untersucht wurde die Hypothese, dass die genetische Vielfalt mit steigender Ökosystem-Störung ebenso erhöht wird (Siehe Punkt 1 Einleitung und Punkt 2.3 Hypothesen).

Die natürliche Population A weist mit insgesamt 27 gefundenen Allelen im Vergleich zu den 24 Allelen der Fischhälter-Population B, in Relation zur Populationsgröße, geringfügig mehr Allele auf. Ebenso ist die *allelic richness* bei Population A mit 3,23 deutlich höher als bei Population B mit 2,62 (Tabelle 20). Die Anzahl der gefundenen Allele und die *allelic richness* stützen somit die Hypothese, dass die genetische Vielfalt bei der natürlichen Po-

population höher ist. Ähnlich verhält es sich, wenn die Heterozygotie der Populationen betrachtet wird. Obwohl mit einem Durchschnittswert von 0,134 die beobachtete Heterozygotie (H_{OBS}) der natürlichen Population A relativ gering ist, zeigt dieser im Vergleich zur Fischhälter-Population B mit 0,059 deutliche Unterschiede. Die Werte für die beobachtete Heterozygotie zeigen einerseits, dass die natürliche Population über höhere genetische Vielfalt verfügt und andererseits bestätigen sie die Annahme, dass die Fischhälter-Population B genetisch isolierter ist. Die berechneten Werte für die erwartete Heterozygotie (H_{EXP}) zeigen mit 0,588 für Population A und 0,336 für Population B, ebenfalls diesen Unterschied (Tabelle 21). In beiden Fällen ist somit die natürliche Population A aufgrund höherer Heterozygotiewerte als genetisch vielfältiger als die Fischhälter-Population B zu betrachten. Die Inzucht-koeffizienten der beiden Populationen unterscheiden sich nicht und sind mit 0,769 bei Population A und Population B relativ hoch. Daraus lässt sich schließen, dass die Art einen hohen Selbstungsgrad aufweist und sich häufig durch Selbstbefruchtung vermehrt. Mit jeder Generation in der Selbstbefruchtung vorherrscht, steigt somit der Homozygotiegrad und umgekehrt verringert sich die Heterozygotie. Aus den Ergebnissen zeigt sich jedoch nicht wie erwartet ein höherer Inzucht-koeffizient bei der Fischhälter-Population B.

Die untersuchten Populationen weisen bezüglich ihrer genetischen Vielfalt tatsächlich deutliche Unterschiede auf. Entsprechend den Erwartungen ist die genetische Vielfalt in der natürlichen Population A größer als in der Fischhälter-Population B. Somit kann die Hypothese, dass steigende Ökosystem-Störung eine Erhöhung der genetischen Vielfalt zur Folge hat, jedenfalls bezüglich *allelic richness* und Heterozygotie, als bestätigt betrachtet werden. Ebenso ist zu beachten, dass das hohe Ausbreitungsvermögen den Genaustausch zwischen natürlichen Standorten an Flüssen über Ausbreitungsvektoren wie Fließwasser oder Vögel ermöglicht und somit ebenfalls eine Erhöhung der genetischen Vielfalt zur Folge hat. Weitere Untersuchungen der Art mithilfe der hier entwickelten Mikrosatelliten sollen die Gültigkeit und Aussagekraft dieser Ergebnisse bekräftigen.

6.2.3 Oberflächenpopulation aus angepassten Genotypen

Es wurde untersucht, ob die Genotypen an der Oberfläche zufällig entstanden sind oder an die vorherrschenden Lebensraumbedingungen angepasste Genotypen aus der Samenbank im Boden darstellen (Siehe Punkt 1 Einleitung und Punkt 2.3 Hypothesen).

Dazu wurden mittels *Arlequin 3.5* (EXCOFFIER & LISCHER, 2010) die F_{ST} .

Werte für beide Populationen berechnet. Es konnte für die natürliche Population A der Wert 0,013 (Signifikanzschwelle 0,63) und für die Fischhälter-Population B der Wert 0,114 (Signifikanzschwelle 0,013) errechnet werden. Anhand der Richtwerte für genetische Differenzierung durch F_{ST} -Werte aus BALLOUX & LUGON, 2003 konnten die Werte eingeteilt werden. Für Population A mit 0,013 würde somit eine sehr geringe, nicht signifikante genetische Differenzierung zwischen Oberflächen und Bodenpopulation gelten, während bei Population B mit 0,114 eine signifikante genetische Differenzierung erkennbar ist. Es könnte somit vorsichtig angenommen werden, dass die Oberflächenpopulation bei Population B aus den Umweltgegebenheiten besser angepassten Genotypen besteht. kann folgendermaßen interpretiert werden: Die natürliche Population besteht auf einem über Generationen hinweg relativ stabilen Standort wodurch keine Änderung von Allelfrequenzen erfolgt, während die Fischhälter-Population einer evolutiv bedingten Änderung der Allelfrequenzen durch Selektion und Adaptation an die unnatürlichen Standortbedingungen unterliegt. Genetische Drift kann als Ursache eher ausgeschlossen werden, nachdem diese hauptsächlich durch Flaschenhalseffekte an kleinen Populationen mit wenigen Individuen auftritt und die Populationsgrößen im Sammeljahr mehrere tausend Individuen je Population umfassten. Es ist jedoch nicht auszuschließen, dass ebensolche Drift in vorangehenden Jahren erfolgte und deren Effekte hier noch sichtbar sind.

Um eine generelle Aussage bezüglich der genetischen Differenzierung von natürlichen und anthropogenen Populationen treffen zu können müssen jedoch die Ergebnisse weiterer Untersuchungen mit allen Markern und an mehreren Populationen abgewartet werden.

7 Literaturverzeichnis

- ABBOTT, C. L. et al. (2010): Twelve microsatellite markers in the invasive tunicate, *Didemnum vexillum*, isolated from low genome coverage 454 pyrosequencing *reads*. *Conservation Genetics Resources*, volume 3 (2011): 79-81.
- AICHELE, D.; SCHWEGLER, H.-W. (2011): *Unsere Gräser*. 12. Aufl., Stuttgart: Franckh-Kosmos Verlags-GmbH Co.KG.
- ALLENBORG, F. W.; LUIKART, G. (2007): *Conservation and the Genetics of Populations*. Oxford: Blackwell Publishing.
- ALTSCHUL, S. F. et al. (1990): Basic Local Alignment Search Tool. *Journal of Molecular Biology*, volume 215: 403-410.
- BALLOUX, F.; LUGON-MOULIN, N. (2002): The estimation of population differentiation with microsatellite markers. *Molecular ecology*, volume 11: 155-165.
- BEEBEE, T. J. C.; ROWE, G. (2008): *An introduction to Molecular Ecology*. 2. Aufl., New York: Oxford University Press.
- BERNHARDT, K-G.; HURKA, H. (1989): Dynamik des Samenspeichers in einigen mediterranen Kulturböden. *Weed Research*, volume 24: 247-254.
- BERNHARDT, K-G.; POSCHLOD, P. (1993): Zur Biologie semi-aquatischer Lebensräume aus botanischer Sicht. In: Bernhardt K-G, Hurka H, Poschlod P (eds.): *Vegetationsdynamik amphibischer Standorte und spezielle Aspekte der Populationsbiologie*. Natur und Wissenschaft, Solingen. S: 5-17.
- BERNHARDT, K-G. (1995): Untersuchungen zur phänotypischen Plastizität des klonalen Wachstums von *Juncus articulatus*. *Verhandlungen der Gesellschaft für Ökologie, Göttingen*, volume 24: 93-96.
- BERNHARDT, K-G. et al. (2004): The soil seed bank as a resource for in situ and ex situ conservation of extinct species. *Scripta Botanica Belgica*, volume 29: 135-139.
- BLOUIN, M. S. et al. (1996): Use of microsatellite loci to classify individuals

by relatedness. *Molecular Ecology*, volume 5: 393-401.

BOOMER, J. J.; STOW, A. J. (2010): Rapid isolation of the first set of polymorphic microsatellite loci from the Australian gummy shark, *Mustelus antarcticus* and their utility across divergent shark taxa. *Conservation Genetics Resources*, volume 1: 393-395.

BOUTIN-GANACHE, I. et al. (2001): M13-Tailed Primers Improve the Readability and Usability of Microsatellite Analyses Performed with Two Different Allele- Sizing Methods. *BioTechniques*, volume 31: 24-28.

BROWNSTEIN, M. J. et al. (1996): Modulation of Non-Templated Nucleotide Addition by *Taq*DNA Polymerase: Primer Modifications that Facilitate Genotyping. *BioTechniques* 20: 1004-1010.

BUEHLER, D. et al. (2011): Using the 454 pyrosequencing-based technique in the development of nuclear microsatellite loci in the alpine plant *Arabis alpina* (Brassicaceae). *American Journal of Botany*, volume 98 (5): e103-e105

CABIN, R.J.; MITCHELL, R.J.; MARSHALL, D.L. (1998): Do surface plant and soil seed bank populations differ genetically?: A multipopulation study of the desert mustard *Lesquerella fendleri* (Brassicaceae). *American Journal of Botany*, volume 85: 1098-1109.

CASTOE, T. A. et al. (2010): Rapid identification of thousands of copperhead snake (*Agkistrodon contortrix*) microsatellite loci from modest amounts of 454 shotgun genome sequence. *Molecular Ecology Resources*, volume 10: 341-347.

CSENCICS, D.; BRODBECK, S.; HOLDBERGER, R. (2010): Cost-Effective, Species-Specific Microatellite Development for the Endangered Dwarf Bulrush (*Typha minima*) Using Next-Generation Sequencing Technology. *Journal of Heredity* 101(6): 789-793.

DARREN, J. K.; MATTICK, J. S. (2008): Touchdown PCR for increased specificity and sensitivity in PCR amplification. *Nature Protocols*, volume 3: 1452-1456.

DAVIDSON, A. & CHIBA, S. (2003): Laboratory temperature variation is a

previously unrecognized source of genotyping error during capillary electrophoresis. *Molecular Ecology Notes*, volume 3: 321-323.

DEIL, U. (2005): A review on habitats, plant traits and vegetation of ephemeral wetlands – a global perspective. *Phytocoenologia*, volume 35: 533–706.

DIERINGER, D.; SCHLÖTTERER, C. (2003): Microsatellite analyser (MSA): a platform independent analysis tool for large microsatellite data sets. *Molecular Ecology Notes*, volume 3: 167-169.

DON, R. H. et al. (1991): 'Touchdown' PCR to circumvent spurious priming during gene amplification. *Nucleic Acids Research*, volume 19: 4008.

EAST, E. M. (1940): The distribution of self-sterility in the flowering plants. *Proceedings of the American Philosophical Society*, volume 82: 449-518.

ELLEGREN, H. (2000): Microsatellite mutations in the germline: implications for evolutionary inference. *Trends in Genetics*, volume 16, No. 12.: 551-558.

ELLEGREN, H. (2004): Microsatellites: simple sequences with complex evolution. *Nature reviews genetics*, volume 5: 425-445.

ENGELHARDT, H.; BECK, W.; SCHMITT, T. (1994): *Kapillarelektrophorese: Methoden und Möglichkeiten*. Braunschweig, Wiesbaden, Friedrich Vieweg & Sohn Verlagsgesellschaft mbH.

EPPENDORF (1997): *Mastercycler Gradient: Operating manual*. Eppendorf, Hamburg.

ESTOUP, A. et al. (1995): Microsatellite Variation in Honey Bee (*Apis Mellifera* L.) Populations: Hierarchical Genetic Structure and Test of the Infinite Allele and Stepwise Mutation Models. *Genetics*, volume 140: 679-695.

EXCOFFIER, L.; LISCHER, H.E. L. (2010): Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, volume 10: 564-567.

FAIRCLOTH, B. C. (2008): MSATCOMMANDER: detection of microsatellite repeat arrays and automated, locus specific primer design. *Molecular Ecology Resources* 8: 92-94.

FIELD, D.; WILLS, C. (1996): Long, Polymorphic Microsatellites in Simple Organisms. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, volume 263: 209-215.

FISCHER, M.; OSWALD, K.; ADLER, W. (2008): *Exkursionsflora für Österreich, Liechtenstein und Südtirol*. 3. Aufl., Linz: Land Oberösterreich, Biologiezentrum der Oberösterr. Landesmuseen.

FOULLEY, J. L.; OLLIVIER, L. (2006): Estimating *allelic richness* and its diversity. *Livestock Science*, volume 101: 150-158.

FRANKHAM, R.; BALLOU, J.D.; BRISCOE, D.A. (2001): *Introduction to conservation genetics*. Cambridge: Cambridge University Press, volume 2.

FREELAND, J. R. (2005): *Molecular Ecology*. England: John Wiley & Sons Ltd.

GARDNER, M. G. et al. (2011): Rise of the machines – recommendations for ecologists when using next generation sequencing for microsatellite development. *Molecular Ecology Resources*, volume 11: 1093-1101.

GLENN, T. C. (2011): Field guide to next-generation DNA sequences. *Molecular Ecology Resources*, volume 11: 759-769.

GIBSON, G.; MUSE, S. V. (2002): *A Primer of Genome Science*. Sunderland Massachusetts: Sinauer Associates.

GOLDSTEIN, D. B.; SCHLÖTTERER, C. (2000): *Microsatellites: Evolution and Applications*. 2. Aufl., New York: Oxford University Press 11: 3-11.

GUICHOUX, E. et al. (2011): Current trends in microsatellite genotyping. *Molecular Ecology Resources* 11: 591-611.

HILL, C. R.; BUTLER, J. M.; VALLONE, P. M. (2009): A 26plex Autosomal STR Assay to Aid Human Identity Testing. *Journal of Forensic Sciences*, 54: 1008–1015

HOFFMANN, J. I.; AMOS, W. (2005): Microsatellite genotyping errors: detection approaches, common sources and consequences for paternal exclusion. *Molecular Ecology*, volume 14: 599-612.

HOHENSEE, C.D.; FREY, W. (2001): Experiments on epizoochorous dispersal by mallards (*Anas platyrhynchos*). *Botanische Jahrbücher für Systematik, Pflanzengeschichte und Pflanzengeographie*, volume 123: 209-216.

HU, G. et al. (1993): DNA Polymerase-Catalyzed Addition of Nontemplated Extra Nucleotides to the 3' End of a DNA Fragment. *DNA and Cell Biology*, volume 12: 763-770.

KALINOWSKI, S. T. (2005): HP-RARE 1.0: a computer program for performing rarefaction on measures of *allelic richness*. *Molecular Ecology Notes*, volume 5: 187-189.

KELKAR, Y. D. et al. (2008): The genome-wide determinants of human and chimpanzee microsatellite evolution. *Genome Research*, volume 18: 20-38.

KIMURA, M.; OHTA, T. (1978): Stepwise mutation model and distribution of allelic frequencies in a finite population. *Proceedings of the National Academy of Sciences of the USA*, volume 75: 2868-2872.

KOCH, M.; HUTHMANN, M.; BERNHARDT, K-G. (2003): *Cardamine amara* L. (Brassicaceae) in dynamic habitats: Genetic composition and diversity of seed bank and established populations. *Basic and Applied Ecology*, volume 4: 339-348.

KRAHULCOVÁ, A (2003): Chromosome numbers in selected monocotyledons (Czech Republic, Hungary, and Slovakia). *Preslia*, volume 75: 97-113.

LEAMON, J. H. et al. (2003): A massively parallel PicoTiterPlate™ based platform for discrete picoliter-scale polymerase chain reactions. *Electrophoresis*, volume 24: 3769-3777.

LEBERG, P. L. (2002): Estimating *allelic richness*: Effects of sample size and bottlenecks. *Molecular Ecology*, volume 11: 2445-2449.

LECK, M.A.; PARKER, V.T.; SIMPSON, R.L. (1989): Ecology of soil seed banks.

Academic Press, San Diego, California, S: 283-305.

LEUTENEGGER, A. L. et al. (2003): Estimation of the Inbreeding Coefficient through Use of Genomic Data. *The American Journal of Human Genetics*, volume 73 (3): 516-523.

LEVIN, D.A. (1990): The seed bank as a source of genetic novelty in plants. *The American Naturalist*, volume 135: 563-572.

LEVINSON, G.; GUTMAN, G. A. (1987): Slipped-Strand Mispairing: A Major Mechanism for DNA Sequence Evolution. *Molecular Biology and Evolution*, volume 4: 203-221.

LIU, C. S. J. et al. (2011): GeneMarker® Genotyping Software: Tools to Increase the Statistical Power of DNA Fragment Analysis. *Journal Of Biomolecular Techniques*, volume 22: 35-36.

MADDEN, T. (2013): *The NCBI Handbook: The BLAST Sequence Analysis Tool*. National Center for Biotechnology Information. Second edition.

MARGULIES, M. et al. (2005): Genome sequencing in microfabricated high-density picolitre reactors. *Nature*, volume 437: 376-380.

MCCUE, K.A.; HOLTSFORD, T.P. (1998): Seed bank influences on genetic diversity in the rare annual *Clarkia springvillensis* (Onagraceae). *American Journal of Botany*, volume 85: 30-36.

MOGG, R. et al. (2002): Characterization of the flanking regions of *Zea mays* microsatellites reveals a large number of useful sequence polymorphisms. *Theoretical and Applied Genetics*, volume 105: 532-543.

MORRIS, A.B.; BAUCOM, R.S.; CRUZAN, M.B. (2002): Stratified analysis of the soil seed bank in the cedar glade endemic *Astragalus bibullatus*: evidence for historical changes in genetic structure. *American Journal of Botany*, volume 89: 29-36.

MORGANTE, M., et al. (2002): Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. *Nature Genetics*, volume 30: 194-200.

MULLIS, K. et al. (1986): Specific Enzymatic Amplification of DNA In Vitro: The Polymerase Chain Reaction. Cold Spring Harbor Symposia on Quantitative Biology, volume 1: 263-273.

MURIN, A.; MÁJOVSKÝ, J. (1987): Karyological study of the Slovak flora XIX. Acta Facultatis Rerum Naturalium Universitatis Comenianae, Botanica. Bratislava. volume 34: 3-20.

OJANGO, J. M., et al. (2011): Quantitative methods to improve the understanding and utilisation of animal genetic resources. Animal Genetics Training Resource. International Livestock Research institute, Nairobi, Kenya, and Swedish University of Agricultural Sciences, Uppsala, Sweden. Version 3, 2011.

PEARSON, W. R. & LIPMAN, D. J. (1988): Improved tools for biological sequence comparison. Proceedings of the National Academy of Sciences of the United States of America, volume 85: 2444-2448

PHILIPPI, G. (1998): Cyperaceae, In: SEBALD, O. et al. (1998): Die Farn- und Blütenpflanzen Baden-Württembergs, Band 8: Spezieller Teil - Juncaceae bis Orchidaceae. Ulmer, Stuttgart. S: 47-98.

POSCHLOD, P. (1996): Population biology and dynamics of a rare short-lived pond mud plant, *Carex bohemica* Schreber. Verhandlungen der Gesellschaft für Ökologie, volume 25: 321-337

QIAGEN (2010): Multiplex PCR Handbook: for fast and efficient multiplex PCR without optimization. QIAGEN.

ROALSON, E. H. (2008): A Synopsis of Chromosome Number Variation in the Cyperaceae. The Botanical Review, volume 72 (2): 209-393.

RODRIGÁÑEZ, J. et al. (2008): Genetic diversity and *allelic richness* in Spanish wild and domestic pig population estimated from microsatellite markers. Spanish Journal of Agricultural Research, volume 6: 107-115.

RONAGHI, M. (2001): Pyrosequencing Sheds Light on DNA Sequencing. Genome Research Cold Spring Harbor Laboratory Press

ROZEN, S.; SKALETSKY, H. (2000): Primer3 on the WWW for General Users and for Biologist Programmers. *Methods in Molecular Biology*, volume 132: Bioinformatics Methods and Protocols, 365-386.

RUDOLPH, B. (2001): Entwicklung, Charakterisierung und genetische Kartierung von Mikrosatelliten-Markern beim Raps (*Brassica napus* L.). 1. Aufl., Göttingen: Cuvillier Verlag.

SAARINEN, E. V.; AUSTIN, J. D. (2010): When Technology Meets Conservation: Increased Microsatellite Marker Production Using 454 Genome Sequencing on the Endangered Okaloosa Darter (*Etheostoma okaloosae*). *Journal of Heredity*, volume 101(6): 784-788.

SAMBROOK, J.; RUSSELL, D. W. (2001): *Molecular Cloning: A Laboratory Manual*. 3. Aufl., New York: Cold Spring Harbor Laboratory Press.

SANTANA, Q. C. et al. (2009): Microsatellite discovery by deep sequencing of enriched genomic libraries. *Bio Techniques*, volume 46: 217-223.

SCHLÖTTERER, C. (2004): The evolution of molecular markers: just a matter of fashion?. *Nature reviews genetics*, volume 5: 63-69.

SCHUELKE, M. (2000): An economic method for the fluorescent labeling of PCR fragments: A poor man's approach to genotyping for research and high-throughput diagnostics. *Nature Biotechnology Resources*, volume 18: 233-234.

SITTE, P. et al. (1998): *Strasburger Lehrbuch der Botanik*. 34. Aufl., Stuttgart: G. Fischer.

STANJEK, K. (1994): Teichbrachen und Teichwüstungen. *Die Oberpfalz* 82: 257-265.

ŠUMBEROVÁ, K.; HORÁKOVÁ, V.; LOSOSOVÁ, Z. (2005): Vegetation dynamics on exposed pond bottoms in the Českobudějovická basin (Czech Republic). *Phytocoenologia*, volume 35: 421-448.

ŠUMBEROVÁ, K.; LOSOSOVÁ, Z.; FABŠIČOVÁ, M.; HORÁKOVÁ, V. (2006) Variability of vegetation of exposed pond bottoms in relation to management and environmental factors. *Preslia*, volume 78: 235-252.

SUN, X. et al. (2006): Detection of Mononucleotide Repeat Sequence Alterations in a Large Background of Normal DNA for Screening High-Frequency Microsatellite Instability Cancers. *Clinical Cancer Resources*, volume 12: 454-459.

TECHEN, N. et al. (2010): Optimized construction of microsatellite-enriched libraries. *Molecular Ecology Resources*, volume 10: 508-515.

TONSOR, S.J. et al. (1993): A life-history based study of population genetic structure: seed bank to adults in *Plantago lanceolata*. *Evolution*, volume 47: 833-843.

TRAXLER, A. (1991): Gefährdung und Förderung von Isoëto-Nanojuncetea-Gesellschaften unter intensiver Teichbewirtschaftung im österreichischen Waldviertel. *KTB Martin Luther-Universität Halle-W. Heuberg 91*: 347-350.

UNTERGASSER, A. et al. (2012): Primer3: new capabilities and interfaces. *Nucleic Acids Research*, volume 40 (15): e115.

VON LAMPE, M. (1996): Wuchsform, Wuchsrhythmus und Verbreitung der Arten der Zwergbinsengesellschaften. *Dissertationes Botanicae*, Band 266. J. Cramer, Berlin, Stuttgart.

WRIGHT, S. (1922): Coefficients of Inbreeding and Relationship. *The American Naturalist*, volume 56, Nr. 654: 330-338

ZANE, L.; BARGELLONI, L.; PATARNELLO, T. (2002): Strategies for microsatellite isolation: a review. *Molecular Ecology*, volume 11: 1-16.

Internetquellen

ANONYM, (nach 2000): Shane's simple guide to F-statistics. Permalink: University of Auckland. URL: <http://www.library.auckland.ac.nz/subject-guides/bio/pdfs/733Pop-g-stats2.pdf> (15.12.2015)

APPLIED BIOSYSTEMS (2010): Applied Biosystems 3500/3500xL Genetic Analyzer User Guide. URL:<http://tools.lifetechnologies.com/content/sfs/manuals/4401661.pdf>. (15.12.2015)

BIOINFORMATICS.ORG (2006): URL: Microsatellite Suite. URL:
<http://www.bioinformatics.org/microsatellite/wiki/Main/HomePage>.
(15.12.2015)

CHEMIE.DE (2015): Information Service GmbH, Berlin, Deutschland. URL:
<http://www.chemie.de/lexikon/Primer.html> (15.12.2015)

INTERNATIONAL LIVESTOCK RESEARCH INSTITUTE (2011): Animal
Genetics Training Resource: Measuring genetic structure from molecular
data. URL: [http://agtr.ilri.cgiar.org/index.php?](http://agtr.ilri.cgiar.org/index.php?option=com_content&view=article&id=267&Itemid=339)
[option=com_content&view=article&id=267&Itemid=339](http://agtr.ilri.cgiar.org/index.php?option=com_content&view=article&id=267&Itemid=339) (15.12.2015)

LEWIS, M. (2015): Agarose gel electrophoresis: basic method. URL:
<http://www.methodbook.net/dna/agarogel.html>. (15.12.2015)

LIFE TECHNOLOGIES (2011): Molecular Biology Essentials.
URL:[http://vivogen.rs/wp-content/uploads/2012/03/Molecular-Biology-](http://vivogen.rs/wp-content/uploads/2012/03/Molecular-Biology-Essential-Catalog.pdf)
[Essential-Catalog.pdf](http://vivogen.rs/wp-content/uploads/2012/03/Molecular-Biology-Essential-Catalog.pdf) (15.12.2015)

LIFE TECHNOLOGIES (2014): DNA Fragment Analysis by Capillary
Electrophoresis. User Guide.
URL:<http://tools.lifetechnologies.com/content/sfs/manuals/4474504.pdf>.
(15.12.2015)

McDONALD, D., (2008): Genetic Markers: lecture notes. University of
Wyoming. URL: <http://www.uwyo.edu/dbmcd/molmark/lect04/lect4.html>
(15.12.2015)

NEW ENGLAND BIOLABS (2015): NEB Tm Calculator v.1.7.
URL:<http://tmcalculator.neb.com/#/>. (15.12.2015)

NOBELPRIZE.ORG (2014): "The Nobel Prize in Chemistry 1993". Nobel
Media AB 2014.
URL:http://www.nobelprize.org/nobel_prizes/chemistry/laureates/1993/
(15.12.2015)

PFLANZENFORSCHUNG.DE (2014): Sequenzierung: Die Pyrosequenzierung.
URL:[http://www.pflanzenforschung.de/de/themen/lexikon/sequenzierung-](http://www.pflanzenforschung.de/de/themen/lexikon/sequenzierung-die-pyrosequenzierung-667)
[die-pyrosequenzierung-667](http://www.pflanzenforschung.de/de/themen/lexikon/sequenzierung-die-pyrosequenzierung-667). (15.12.2015)

PHILIPPS UNIVERSITÄT MARBURG - Fachbereich Chemie (2007): Dateien zur Forschung: Kapillarelektrophorese. URL: <https://www.uni-marburg.de/fb15/ag-pyell/forschung/Dateien/CE>. (15.12.2015)

QIAGEN (2014): DNeasy Plant Mini Kit: Product Details. URL:<http://www.qiagen.com/at/products/catalog/sample-technologies/dna-sample-technologies/genomic-dna/dneasy-plant-mini-kit/#productdetails>. (15.12.2015)

QIAGEN (2014): QIAcube: Product Details. URL:<http://www.qiagen.com/at/products/catalog/automated-solutions/sample-prep/qiacube/#productdetails>. (15.12.2015)

ROCHE DIAGNOSTICS GMBH, (2014): GSFLX Application Flyer: URL:http://454.com/downloads/GSFLXApplicationFlyer_FINALv2.pdf. (15.12.2015)

Software und Programme

MSATCOMMANDER: FAIRCLOTH, B. C. (2008): MSATCOMMANDER: detection of microsatellite repeat arrays and automated, locus-specific primer design. *Molecular Ecology Resources*, volume 8:92-94. doi:10.1111/j.1471-8286.2007.01884.x.

PRIMER3: ROZEN, S.; SKALETSKY, H. J. (2000): Primer3 on the WWW for general users and for biologist programmers. In: *Bioinformatics Methods and Protocols: Methods in Molecular Biology* (eds Krawetz S, Misener S). Humana Press, Totowa, NJ.

GENEMARKER 2.4: SoftGenetics LLC, State College, PA, USA.

MSA: DIERINGER, D.; SCHLÖTTERER, C. (2003): Microsatellite analyser (MSA): a platform independent analysis tool for large microsatellite data sets. *Molecular Ecology Notes*, volume 3 (1), 167-169.

ARLEQUIN 3.5: EXCOFFIER, L.; LISCHER, H.E. L. (2010): Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, volume 10: 564-567.

8 Abbildungsverzeichnis

Abbildung 1: <i>Cyperus fuscus</i> mit Blüte, Fruchtstand und Samen (KOPS, J. et al., (1872): Flora Batava).....	18
Abbildung 2: Chromosomenzählung der Gattung <i>Cyperus</i> . Modifiziert aus ROALSON 2008 (Haploid Chromosome Number variation in <i>Cyperus</i>).....	20
Abbildung 3: Schleifenbildung und Längenänderung der DNA stränge. (ELLEGRÉN, 2000; "Figure 1: Model of microsatellite mutation by replication slippage")	23
Abbildung 4: Übersicht der Schritte bei SSR-enrichment. (TECHEN, 2010; "Overview of genomic DNA SSR-enriched library protocol")	30
Abbildung 5: Schematische Gliederung der PCR Amplifikation mit Fluoreszenzlabeln. (SCHUELKE, 2000; "Figure 1: Amplification sceme for the one-tubed, single-reaction nested PCR method."	31
Abbildung 6: Beispiel einer PRIMER 3 output-Datei (PRIMER 3).....	33
Abbildung 7: Übersicht über Hairpins und mögliche unerwünschte Sekundärstrukturen. (UNTERGASSER et al., 2012; "Figure 5: Primer3's thermodynamic models for predicting the stability of secondary structures").....	36
Abbildung 8: Length-graph von 143027 Sequenzen.....	48
Abbildung 9: Sequenzlängen von 4877 Sequenzen.....	57
Abbildung 10: Fotos der Standorte. Linke Seite: Natürlicher Standort der Population A. Rechte Seite: Fischhälter der Population B. (Foto: J. Böckelmann).....	64
Abbildung 11: Typische Unterschiede der Signalintensität diploider Organismen.....	74
Abbildung 12: Length-graphs und durchschnittliche Read-Länge der beiden Pyrosequenzierungen durch Ecogenics und LGC.....	81
Abbildung 13: Length graph des Markers <i>Cypfus_3114</i> ohne PIG-tail: Deutliche stutter und -A-Banden hinter Haupt-peak sichtbar.....	82

Tabellenverzeichnis

Tabelle 1: Bezeichnung und Herkunft der Test-Individuen. *Bodenprofil Bezeichnet die Bodenschicht in der die Entnahme der Pflanzen erfolgte.....	47
Tabelle 2: Übersicht der 454 Pyrosequenzierungsdaten zu Methode 1.....	48
Tabelle 3: Untergrenze der Wiederholungen der SSR Motive.....	49
Tabelle 4: tag's und PIG-tail - Methode 1.....	51
Tabelle 5: PCR Komponenten und Reaktionsmengen Methode 1.....	52
Tabelle 6: PCR-Programm Gradient.....	53
Tabelle 7: PCR-Programm Touchdown.....	54
Tabelle 8: Zusammenfassung Methode 1.....	56
Tabelle 9: Übersicht der 454 Pyrosequenzierungsdaten zu Methode 2.....	58
Tabelle 10: Tag und PIG-tail - Methode 2.....	59
Tabelle 11: PCR Komponenten und Reaktionsmengen Methode 2.....	60
Tabelle 12: PCR-Programm 56° Ecogenics.....	60
Tabelle 13: Zusammenfassung Methode 2.....	62
Tabelle 14: Standort- und Nutzungsinformationen der Populationen A und B.....	64
Tabelle 15: Verwendete Marker für vorläufige Anwendungstests.	65
Tabelle 16: Übersicht über die 21 entwickelten Mikrosatelliten-Marker für <i>Cyperus fuscus</i>	70
Tabelle 17: Repeat Motive, Spannweite der Fragmentlängen (Allele size range) und Anzahl gefundener Allele je Marker der 21 entwickelten Mikrosatelliten-Marker.....	71
Tabelle 18: Motivhäufigkeiten der Mikrosatelliten aus Methode 1.....	73
Tabelle 19: Gefundene Allele je Locus und Fraktion der Populationen A und B.....	75
Tabelle 20: allelic richness und Gesamtanzahl gefundener Allele der Populationen A und B.....	75
Tabelle 21: Werte für beobachtete (H_{OBS}) und erwartete (H_{EXP}) Heterozygotie für die Oberflächen Fraktion O und Bodenfraktion B der Populationen A und B an 8 Loci.....	76
Tabelle 22: Inzuchtkoeffizienten der Populationen A und B an 8 Loci und zwei Bodenfraktionen O und B.....	78
Tabelle 23: Berechnete F_{ST} -Werte für die Subpopulationen der Populationen A und B...79	79
Tabelle 24: Übersicht der 454 Pyrosequenzierungsdaten der Methoden 1 und 2.....	80