

DISS. ETH NO. 23568

**SELF-FERTILIZATION AND MARKER-TRAIT  
ASSOCIATIONS IN SAINFOIN (*ONOBRYCHIS  
VICIIFOLIA*)**

A thesis submitted to attain the degree of  
DOCTOR OF SCIENCES of ETH ZURICH  
(Dr. sc. ETH Zurich)

presented by  
**KATHARINA KEMPF**

Dipl.-Agr.Biol., University of Hohenheim

born on 17.04.1985

citizen of Germany

accepted on the recommendation of

Prof. Dr. Achim Walter, examiner  
Prof. Dr. Michael Kreuzer, co-examiner  
Prof. Dr. Bruno Studer, co-examiner  
Dr. Roland Kölliker, co-examiner

2016

## Summary

An extension of ruminant husbandry to satisfy the growing demand of milk- and meat products has negative consequences for the environment like e.g. increased methane gas emissions and wastewater. Additionally, higher production levels in animals are difficult to obtain because increased protein and energy levels must be provided within the fodder ration. There is also a continuous risk of diseases and parasite burdens in livestock due to resistance formation against conventional anthelmintic drugs. These problems could be tackled by the use of forage legumes with anthelmintic properties like sainfoin (*Onobrychis viciifolia*), which has health promoting characteristic mainly caused by the presence of polyphenols therein, especially condensed tannins. Condensed tannins are also known to reduce bloat and methane gas emissions in ruminants and improve protein absorption by protein-tannin-interactions in the rumen. Furthermore, the ability of sainfoin to fix atmospheric nitrogen by rhizobia bacteria association allows a reduction of nitrogen fertilizer input. However, sainfoin cultivation decreased during the last century, presumably due to cheap prices for inorganic fertilizer and competition from higher yielding forage crops. Consequently, sainfoin breeding attempts were reduced to a minimum. Existing varieties show poor and unstable biomass yields, reduced persistence and low seed yields. Breeding new, high performing varieties is indispensable to make sainfoin cultivation attractive to farmers.

Tetraploidy and, hence, complex inheritance patterns for traits are one important aspect breeders have to deal with in breeding sainfoin. Moreover, contradictory statements exist about the reproduction system in sainfoin, namely if sainfoin is strongly cross-fertilized or also able to self-fertilize. Here, molecular markers could help to unravel the breeding system in sainfoin and to clarify inheritance patterns on a molecular level. Detection of associations between molecular markers and traits would allow selection of individuals for desired traits on the basis of marker assisted selection (MAS). Thereby, localization of loci explaining variation in a quantitative trait (quantitative trait loci, QTL) within the genome could be improved with the establishment of a genetic linkage map which does not yet exist for sainfoin.

Due to the given necessity to assist sainfoin breeding, our study aimed to elucidate the breeding system of sainfoin as well as its consequences on plant performance (chapter 2), to test newly developed molecular markers which could be, later on, applied in genetic studies (chapter 3), to apply these markers to F<sub>1</sub> offspring of one biparental population to detect marker trait associations (chapter 4) and to establish a genetic linkage map based on this marker data (chapter 5). The present study was part of the European Marie Curie Initial Training Network 'LegumePlus' (PITN-GA-2011-289377) which aimed to conduct a holistic research of sainfoin.

Our first aim was to detect possible self-fertilization in sainfoin. Generally, sainfoin is described as mainly cross-fertilized, but few studies report high self-fertilization rates based on results from manual self-pollination by hand or using a recessive flower color marker as detection system. So far, information about self-fertilization under natural conditions is missing, as well as possible factors favoring self-fertilization. To address these questions, three biparental populations were established with bumble bee (*Bombus terrestris*) pollination in the greenhouse from five clones of two parental genotypes in each population. Using this approach, self-fertilization as well as cross-fertilization was possible, but the pollination was affected by the number of potential crossing partners and bumble bee movements. In addition to this approach, self-fertilization was assessed in three natural populations in the open field. For the first time, self-fertilization was assessed on the basis of dominant sequence related amplified polymorphism (SRAP) and co-dominant simple sequence repeat (SSR) molecular markers. In the greenhouse populations, high self-fertilization rates of 48.5, 61.5 and 64.8% were detected. Contrastingly, 0.0, 1.8 and 3.9% self-fertilization was detected in the three natural populations. The greenhouse populations were phenotyped over a period of three years to quantify possible inbreeding depression. It could be shown that seed yield per plant was highly affected (up to 79% reduction) by inbreeding depression, whereas plant height and vigor were less affected with a reduction of 23.8 and 17.2%, respectively.

Another emphasis of the study was to test sainfoin specific co-dominant SSR markers (chapter 3). Such markers are essential to realize genetic studies in sainfoin or, further, could be used for marker assisted selection. For this reason,

newly developed SSR markers were tested on a set of 32 sainfoin individuals. Thereby, 101 amplified markers showed allelic polymorphisms between the individuals. The number of alleles per marker lay between two and 24 alleles, whereby only five of total alleles were present in all individuals. Summarizing marker data using multivariate statistics revealed two main clusters, the first one containing individuals from Western Europe, whereas the second one contained individuals from Southern- and Eastern Europe and the U.S. The discovery of genetic differences by SSR markers indicates that those markers are useful for further genetic studies in sainfoin.

A set of tested SSR, together with SRAP markers, was used to search for marker trait associations (chapter 4) in 122 F<sub>1</sub> offspring of one greenhouse population developed within the framework of chapter 2. Finding such markers would allow selecting promising individuals already in early stages of the breeding processes. For this reason, all marker alleles were scored as present (1) / absent (0) in the individuals and tested for trait associations via simple regression analysis. Trait associated markers could be found for plant height, seed yield and vigor, whereby a correlation analysis among associated markers indicates that these are linked to at least two different QTL. Furthermore, one associated marker was found for prodelphinidin-share of condensed tannins.

At last, a set of SSR and SRAP markers was used to create a genetic linkage map of sainfoin based on 122 F<sub>1</sub> offspring from the above mentioned populations (chapter 5). JoinMap (for diploid and allopolyploid species) and TetraploidMap (for autotetraploid species) software packages were used accounting for the unknown origin of polyploidy in sainfoin. It could be shown that none of the software produced a satisfactory linkage map, because most of the markers grouped to only one linkage group. Partially, certain linkages between markers could be found by both software, but, based on the used population, no definite map could be developed and no preference for one of the two software packages could be given.

The results of the present study showed high self-fertilization rates in sainfoin with accompanied inbreeding depression. Consequently, a self-incompatibility system can be excluded. These results are directly useable to optimize breeding methods. Furthermore, the usability of newly developed markers for further

genetic studies was proven for sainfoin. The detected marker trait associations for agronomic and compositional traits could be applied to select superior individuals in plant breeding programs to bred new sainfoin varieties.

## Zusammenfassung

Ein Ausbau der Wiederkäuerhaltung, aufgrund der wachsenden Nachfrage nach Milch- und Fleischprodukten, belastet die Umwelt durch steigende Methangaseinträge und anfallende Abwässer. Bezüglich der Tiergesundheit ist es eine Herausforderung genug Energie und Protein über die Fütterung zuzuführen um die Produktionsleistung zu steigern. In den Tierbeständen treten zudem Krankheiten und Parasiten auf, welche mit Medikamenten chemischen Ursprungs immer schwerer einzudämmen sind da z.B. Parasiten Resistenzen gegen gebräuchliche Anthelminthika entwickelt haben. Eine Alternative ist die Nutzung von Futterpflanzen mit anthelmintischen Eigenschaften wie die ausdauernde Leguminose Esparsette (*Onobrychis viciifolia*). Die Esparsette besitzt Polyphenole, darunter kondensierte Tannine, welche anthelminthisch wirken und zudem Blähungen und den Methangasausstoss reduzieren. Zusätzlich wird das in der Futtermittelration enthaltene Protein durch Interaktion mit diesen Tanninen besser vom Tier absorbiert.

Obwohl die Esparsette wertvoll für die Tierernährung ist und ihr Anbau Ressourcen spart (aufgrund atmosphärischer Stickstofffixierung durch assoziierte Knöllchenbakterien), ging der Anbau im letzten Jahrhundert stetig zurück. Gründe hierfür sind günstige Preise für anorganischen Dünger und Konkurrenz durch andere, ertragreichere Kulturarten. Die Esparsette wird in Folge dessen kaum noch gezüchtet. Vorhandene Sorten haben oft geringe und schwankende Biomasseerträge über die Anbaujahre, eine reduzierte Persistenz und wenig Saatgutertrag. Die Zucht neuer, leistungsstarker Sorten ist unabdingbar, um den Esparsettenanbau für Landwirte attraktiv zu machen. Die züchterische Bearbeitung der Esparsette stellt nicht nur aufgrund der Tetraploidie und daraus resultierender komplexer Vererbungsgänge von Merkmalen eine Herausforderung dar. Des Weiteren existieren widersprüchliche Aussagen darüber, ob die Esparsette ausschliesslich fremdbefruchtet wird. Die Unterstützung der Züchtung mit Hilfe molekularer Marker würde es ermöglichen, sowohl die Frage der Befruchtungsart in der Esparsette zu klären, als auch Vererbungsgänge auf molekularer Ebene zu beschreiben. Durch Nutzung dieser Marker für das Auffinden von Marker-Merkmalsassoziationen

könnte gezielt auf gewünschte Merkmale selektiert werden. Um die Effektivität der Selektion zu verbessern ist die Erstellung einer genetischen Kopplungskarte von Bedeutung, da es so möglich ist die Chromosomenregion welche für das quantitative Merkmal (QTL) verantwortlich ist genauer zu lokalisieren.

Aufgrund der Notwendigkeit die Esparsetten Züchtung zu unterstützen, ergaben sich folgende Ziele für unserer Studie: die Aufklärung des Befruchtungssystems der Esparsette und dessen Auswirkung auf die Leistung der Pflanze (Kapitel 2), molekulare Marker für genetische Studien zu testen (Kapitel 3), diese molekularen Marker zum Auffinden von Merkmalsassoziationen in einer Kreuzungspopulation anzuwenden (Kapitel 4) und eine genetische Kopplungskarte der Esparsette zu erstellen (Kapitel 5). Die Studie wurde dabei im Rahmen des europäischen Marie Curie Initial Training Network, 'LegumePlus' (PITN-GA-2011-289377) durchgeführt, welches sich mit der ganzheitlich Erforschung der Esparsette befasst.

Die erste Fragestellung der Studie bezüglich des Befruchtungssystems (Kapitel 2) gründet sich auf widersprüchliche Angaben der Literatur zu diesem Thema. Wird die Esparsette zumeist als überwiegend fremdbefruchtet beschrieben, so finden sich wenige Studien die von geringen bis teilweise hohen Selbstbefruchtungsraten sprechen. Diese Ergebnisse stützen sich dabei auf morphologische Marker (Blütenfarbe) und Samenansatz durch gezielten Handbestäubungen mit eigenem Pollen. Eine Erfassung der Selbstbestäubungsrate unter natürlichen Bedingungen und unter Berücksichtigung von Aspekten welche Fremd- oder Selbstbefruchtung fördern wurde noch nicht erbracht. Um dieser Frage nachzugehen, wurden drei Kreuzungen mit je zwei fünffach verklonten Elternpflanzen im Gewächshaus mittels Hummelbestäubung (*Bombus terrestris*) durchgeführt. Hierbei war es möglich, dass die Pflanzen mit dem eigenen oder fremden Pollen bestäubt wurden. Der Bestäubungsvorgang war durch die Anzahl an vorhandenen Bestäubungspartnern und Hummeln eingeschränkt. Zusätzlich wurde die Bestäubungsrate auf drei natürlichen Flächen im Freiland ohne Einschränkungen erfasst. Die Selbstbefruchtungsraten wurden in unserer Studie erstmals anhand dominanter sequence related amplified polymorphism (SRAP) und zwei co-dominanter simple sequence repeat (SSR) molekularer Marker

bestimmt. In den Gewächshauspopulationen konnten hohe Selbstbefruchtungsraten von 48.5, 61.5 und 64.8% erzielt werden. Im Gegensatz dazu waren die Selbstbefruchtungsraten auf den natürlichen Flächen vernachlässigbar gering mit 0.0, 1.8 und 3.9%. Um eine mögliche Inzuchtdepression in der Esparsette zu quantifizieren, wurden die drei Gewächshauspopulationen über drei Jahre phenotypisiert. Es konnte dabei gezeigt werden, dass die Saatguterträge pro Einzelpflanze in den Selbstungsnachkommen stark zurückgehen (bis zu 79%), wobei die Pflanzenhöhe und die allgemeine Wüchsigkeit von einer etwas geringeren Inzuchtdepression betroffen waren (23.8% und 17.2%).

Der zweite Schwerpunkt der vorliegenden Studie befasste sich mit dem Testen von spezifischen co-dominanten molekularen SSR Markern für die Esparsette (Kapitel 3). Solche molekularen Marker sind für die Durchführung von genetischen Studien in der Esparsette unerlässlich. Zudem können sie, falls sie mit einem phänotypischen Merkmal assoziiert sind in der praktischen Züchtung verwendet werden. Zu diesem Zweck wurden neue SSR Marker an einem Pool aus 32 diversen Esparsetten Pflanzen verschiedener Herkunft angewendet. Dabei konnten 101 amplifizierte Marker Polymorphismen zwischen den Individuen aufzeigen. Die Anzahl von Allelen pro Marker reichte dabei von zwei bis 24 über alle Individuen. Zudem fanden sich nur fünf Allele die in allen Individuen vorkamen. Die genetische Charakterisierung der Individuen zeigte, dass diese sich in zwei Gruppen aufspalten. Gruppe 1 beinhaltet Individuen aus Westeuropa, wohingegen Gruppe 2 Individuen aus Süd- und Osteuropa und den USA aufweist. Die durch die SSR Marker aufgezeigten genetischen Unterschiede und Gemeinsamkeiten in den Individuen zeigen, dass sich diese Marker auch für weitere genetische Studien in der Esparsette nutzen lassen

Die Nutzung eines Sets dieser charakterisierten SSR Marker und dominanter SRAP Marker zur Erfassung von Marker-Merkmalassoziationen war ein weiteres Teilziel unserer Studie (Kapitel 4). Das Auffinden solcher Marker würde es in Zukunft erlauben, Individuen gezielt, schon in frühen Züchtungsstadien, nach erwünschten Merkmalen zu selektieren. Für die Suche nach solchen Assoziationen wurden 122 F1 Kreuzungsnachkommen aus einer der drei Gewächshauspopulationen (Kapitel 2) untersucht. Zur Detektion dieser



Assoziationen wurden alle Marker Allele als anwesend (1) / abwesend (0) in den Individuen erfasst und mittels einfacher Regressionsanalyse auf vorhandene Assoziationen zu phenotypisierten Merkmalen getestet. Merkmalasoziierte Marker konnten dabei für Pflanzenhöhe, Saatgutertrag und Wüchsigkeit gefunden werden. Des Weiteren wurde ein assoziierter Marker für den Prodelphinidin-Anteil in kondensierten Tanninen gefunden. Die Korrelation unter den Markern deutet zudem darauf hin, dass mindestens zwei QTL mit diesen Markern gekoppelt sind und sich für eine gezielte Merkmalsselektion eignen.

In einem letzten Schritt sollte aus einem Set von SSR und SRAP Markern eine genetische Kopplungskarte für 122 Kreuzungsnachkommen aus der oben genannten Kreuzungspopulation erstellt werden (Kapitel 5). Aufgrund des ungeklärten Ursprungs der Polyploidie in der Esparsette wurden die Software JoinMap (für diploide und allopolyploide Arten) und TetraploidMap (für autotetraploide Arten) angewandt. Es zeigte sich, dass mit keiner der beiden Softwares eine umfassende Kopplungskarte für die Esparsette erstellt werden konnte, da die meisten Marker Allele in eine Kopplungsgruppe gruppieren. Die aufgezeigten Marker Kopplungen waren dabei nur teilweise in beiden Softwares zu finden. Letztendlich liess sich auf Basis der verwendeten Population keine definitive Kopplungskarte erstellen und keine Aussage konnte darüber getroffen werden, welche Software die geeignete ist, um eine Kopplungskarte für die Esparsette zu erstellen.

Die Resultate unserer Studie zeigen dass hohe Selbstbefruchtungsraten in der Esparsette möglich sind und diese von einer Inzuchtdepression begleitet werden. Somit lässt sich ein Selbstinkompatibilitätssystem ausschliessen. Diese Erkenntnisse sind direkt nutzbar um praktische Züchtungsmethoden zu optimieren. Des Weiteren konnte die Nutzbarkeit neu bereitgestellter Marker für die Esparsette bewiesen werden, was zukünftige genetische Studien in der Esparsette fördern könnte. Die Assoziation einzelner SSR Marker mit agronomischen Merkmalen und Inhaltsstoffen kann zur Selektion von Individuen für Pflanzenzuchtprogramme genutzt werden und somit die Züchtung neuer Esparsettensorten unterstützen.