Dolce Vita in the Rice Paddy

Characterization of weedy rice groups in Northern Italy and investigation of their evolutionary origins

Zur Erlangung des akademischen Grades eines

DOKTORS DER NATURWISSENSCHAFTEN

(Dr. rer. nat.)

Fakultät für Chemie und Biowissenschaften

Karlsruher Institut für Technologie (KIT) – Universitätsbereich

genehmigte

DISSERTATION

von

Dipl. Biol. Annabelle Grimm

aus

Karlsruhe

Dekan: Prof. Dr. Peter Roesky

Referent: Prof. Dr. Peter Nick

Korreferent: Prof. Dr. Jörg Kämper

Tag der mündlichen Prüfung: 16. April 2014

„Nothing in biology makes sense except in the light of evolution”

Theodosius Dobzhansky

Die vorliegende Dissertation wurde am Botanischen Institut des Karlsruher Instituts für Technologie (KIT) von August 2010 bis März 2014 angefertigt.

Hiermit erkläre ich, dass ich die vorliegende Dissertation, abgesehen von der Benutzung der angegebenen Hilfsmittel, selbständig verfasst habe.

Alle Stellen, die gemäß Wortlaut oder Inhalt aus anderen Arbeiten entnommen sind, wurden durch Angabe der Quelle als Entlehnungen kenntlich gemacht.

Diese Dissertation liegt in gleicher oder ähnlicher Form keiner anderen Prüfungsbehörde vor.

Karlsruhe,

Danksagung

Ich bedanke mich bei Professor Peter Nick, dass er mir die Möglichkeit gegeben hat, diese Arbeit an seinem Institut anzufertigen und mich durch freies und selbstständiges Arbeiten als Wissenschaftler und Person weiterzuentwickeln.

Mein herzlicher Dank gilt Professor Jörg Kämper für das Korreferat und den weiteren Mitgliedern des Komitees Professor Horst Taraschewski und Professor Reinhard Fischer.

Mille grazie to Professor Aldo Ferrero, Francesco Vidotto and Silvia Fogliatto from the Torino University for the great cooperation and all the support.

Many thanks to professor Katrien Devos at the University of Georgia. She gave me the great opportunity to learn the technique of next-generation sequencing and to study abroad for six months. Plus I want to thank Stephan Schröder, Bochra Bahri, Guillaume Daverdin, Doug Eudy, Rajiv Parvathaneni and Trudi Thomas for all the support in the lab.

Vielen Dank an alle Mitarbeiter des botanischen Instituts, die mich während meiner Arbeit unterstützt haben. Ein ganz besonders herzliches Dankeschön an Sebastian für die Hilfe bei den Special-Effects.

Danke an meine Familie, meine Eltern und meinen Bruder, die mir Sicherheit geben, indem ich die Gewissheit habe, dass immer jemand hinter mir steht.

Ein riesiges Danke, von ganzem Herzen, an Daniel, für mehr selbstlose Unterstützung und Verständnis als ich je erhoffen konnte.

Special thanks to Alex Kidd, Nick Stom, Kaitlyn Rountree and Mel Spilman for being the best roommates that I could have wished for. You made me feel at home even though I was 4000 miles away from my family.

Ein riesiges Dankeschön an all die Autoren, Regisseure, Künstler und Musiker, die meinen Alltag zu Abenteuern gemacht und mich inspiriert haben.

Zusammenfassung

Weedy Rice oder auch roter Reis, aufgrund seiner charakteristischen roten Samenpigmentierung, ist ein Beispiel für ein aggressives Unkraut in der Landwirtschaft. Solche Getreide-Unkräuter entwickeln sich unter den Selektionsdrücken, welche durch den Menschen im Zuge der Landwirtschaft ausgeübt werden. Es wird vermutet, dass solche Unkräuter in der Landwirtschaft die gesamte Welt-Getreideproduktion um 10% reduzieren.
Roter Reis, welcher zur Spezies *Oryza sativa* gehört, ist ein Konkurrent von kultiviertem Reis, der Felder in der ganzen Welt infiziert und dadurch gravierende ökonomische Probleme in Form von Ernteverlusten und verminderter Saatgutqualität verursacht. Als Europas größter Reisproduzent leidet auch Italien unter schwerem Befall mit rotem Reis, welcher seit dem Wechsel zur Methode der Direktsaat in den 1960er Jahren dramatisch verstärkt wurde. Diese Arbeit konzentrierte sich auf roten Reis in Italien und hatte zum Ziel verschiedene Populationen zu charakterisieren und mögliche Ursprünge seiner Evolution zu untersuchen.

Eine Analyse von genomischen „simple sequence repeats“ (SSRs), welche für roten Reis und Kultivare aus Italien durchgeführt wurde, zeigte verschiedene Populationen, die nicht nur anhand von genetischen, sondern auch von morphologischen Merkmalen unterscheidbar waren. Außerdem konnten bestimmte Kultivare als mögliche Vorfahren von verschiedenen Gruppen von rotem Reis identifiziert werden. Manche dieser Kultivare repräsentieren alte Landrassen, andere moderne Kultivare. Die weitere Untersuchung von charakteristischen Merkmalen des roten Reises wie zum Beispiel die Pflanzenhöhe, die Färbung der Samen oder der Samenabwurf (Shattering) mit unterschiedlichen Methoden (Next-Generation Sequenzierung, morphologische Daten, physiologische Tests) lassen vermuten, dass manche Populationen sich aus den genetisch variableren alten Landrassen entwickelten während andere durch Mutation aus modernen Kultivaren hervorgingen. Untersuchungen der Pigmentierung der Samen konnten außerdem enthüllen, dass zumindest ein geringer Einfluss von wildem Reis auf bestimmte Populationen von rotem Reis stattgefunden hat. Da keine wilden Vertreter der Gattung *Oryza* in Europa heimisch sind, gehen diese Ereignisse vermutlich zurück auf die Anfangszeit des Reisanbaus in Europa, als Saatgut noch aus Asien importiert wurde. Die wilden *Oryza*-Spezies könnten möglicherweise durch kontaminiertes Saatgut nach Europa gelangt sein.
Basierend auf den Ergebnissen dieser Arbeit konnte ein Modell für die Evolution von rotem Reis in Italien über die Zeit rekonstruiert werden. Es zeigt, dass roter Reis in Italien verschiedene evolutionäre Ursprünge hat. Ein Einfluss von wildem Reis in der Anfangszeit

des Reisanbaus könnte die ersten Populationen in Italien hervorgebracht haben, ein Prozess der als „exoferality“ bezeichnet wird. Weitere Quellen können unter den unterschiedlichen Kultivaren gefunden werden, welche sich veränderten und so im Laufe der Zeit selbst zu Unkräutern wurden, auch bezeichnet als „endoferality“. Diese Abläufe sind kontinuierlich und machen so die Evolution von rotem Reis zu einem dynamischen und fortschreitenden Prozess, der im Laufe der Zeit neue Populationen hervorbringt, welche an die Bedingungen in der jeweiligen landwirtschaftlichen Umgebung angepasst sind.

Die Evolution von rotem Reis in Italien ist exemplarisch für die Evolution solcher landwirtschaftlichen Unkräuter im Allgemeinen. Da gezeigt wurde, dass deren Entwicklung ein fortschreitender Prozess ist, sollte in Zukunft nicht nur die Bekämpfung bereits existierender Unkräuter, sondern vielmehr der Analyse von Selektiondrücken, welche die Entwicklung neuer Unkräuter vorantreiben könnten, Ziel der Forschung sein. Eine Veränderung der Methoden, bevor sich die entsprechenden Unkräuter anpassen können, könnte helfen weiteren Befall mit Unkräutern in der Landwirtschaft zu reduzieren und so die Produktivität im Getreideanbau zu steigern.

Abstract

Weedy rice, or red rice due to the characteristic red pigmented grains, is an example for an aggressive agricultural weed. These crop weeds develop forced by selection pressures applied by human action during agriculture. Agricultural weeds are suspected to reduce the total world crop production by 10%.

Weedy rice, belonging to the *Oryza sativa* species, is a competitor of cultivated rice, infesting paddies worldwide and causing severe economic problems due to yield losses and reduced seed quality. Italy, as Europe's main rice producing country, also suffers from severe weedy rice infestations which dramatically increased since the switch to the direct sowing practice in the 1960s. This study focused on weedy rice in Italy, aimed to characterise different populations and investigate possible origins of its evolution.

An analysis of genomic simple sequence repeats (SSR), carried out on Italian weedy rice and cultigens, showed different populations that are not only distinct by genetic but also morphological matters. It also identified certain cultigens as possible ancestors for weedy rice groups. Of these candidate accessions some represent old landraces others contemporary cultivars. Further investigation of characteristic weedy rice traits (next-generation sequencing, morphology, physiology) suggest that some weedy rice groups might have evolved from the genetically more diverse landraces and others from modern cultivars by mutation. Studies on the seed pigmentation also revealed that at least a small influence of wild rice has shaped some weedy rice groups in Italy. Since no wild representatives of the genus *Oryza* are endemic in Europe, these events most likely date back to the early days of European rice agriculture when seeds were imported from Asia. The wild *Oryza* species might have been carried over to Europe by contaminated seed stocks.

Based on the results of this study, a model of the evolution of weedy rice in Italy over time could be reconstructed. It shows that Italian weedy rice populations have multiple origins. Wild rice influence in the early stages of rice cultivation might have created the first weedy rice populations in Italy, a process called exoferality. Further origins can be sought in different cultigens that changed and became weeds over time, also referred to as endoferality. These processes continue, making weedy rice evolution in Italy a dynamic ongoing process producing novel populations over time, which are adapted to the conditions in the corresponding agricultural environment.

The evolution of weedy rice in Italy is exemplarily for crop weed evolution in general. Since weed evolution was shown to be an ongoing process, not only the treatment of already existing weeds, but rather the analysis of the selection pressures that might generate new

weeds should be targets in future research. The change of methods before the weeds can adapt could help to reduce novel agricultural weed infestations of agricultural areas in the future, and increase crop productivity in systems.

Table of contents

List of Abbreviations

AZ	abscission zone
GMO	genetically modified organism
HPLC	high pressure liquid chromatography
KHYS	Karlsruhe House of Young Scientists
KIT	Karlsruhe Institute of Technology
NGS	next-generation sequencing
ORF	open reading frame
SNP	single nucleotide polymorphism
SSR	simple sequence repeat
SW	smoke water
UGA	University of Georgia
Unito	Universita degli studi di Torino
UPGMA	Unweighted Pair Group Method with Arithmetic Mean
wt	wild type

1 Introduction

Weedy rice or red rice, due to the reddish brown colour of the seed pericarp, is a major weed of the rice cultivation ecosystem and the object of interest in this study. Weedy rice is a cultigen parasite that grows in paddy fields in agro ecosystems worldwide. It competes with the rice cultigens for light, nutrients and other resources. Since cultigens are bred for seed production and thus less competitive, in the struggle for resources weedy rice has an advantage and outcompetes cultivated rice. This leads to yield losses and reduced seed quality. It is not easy to get hold of exact data of weedy rice infestation and yield losses. Most data reported in the literature are based on statistics in the US. Infestation rates for the total US rice growing area are estimated to have reached 30% (Gaily, 2005). The percentage of yield loss in a paddy caused by weedy rice are highly variable but are reported to reach up to 80% (Storminess et al., 2005). Since rice is one of the most important crop species for nutrition on earth, providing about 20% of the caloric intake worldwide (Thurber et al., 2011), these numbers are alarming. Weedy rice is only one example for an agricultural weed. Agricultural weeds are estimated to reduce the total world crop production by 10% (Coerce 2006). Therefore the understanding of the evolution of crop weeds in general is an important goal on the way to efficient treatment strategies. Weedy rice is one example for the evolution of such an agricultural weed. The good scientific background of its associated crop (rice) and the well documented history of rice cultivation in Italy make Italian weedy rice an ideal study system to trace the mechanisms of how crop weeds come into being and evolve in our agricultural systems. This knowledge might be crucial for the future to prevent the emergence of novel weed species and lead to new management strategies for already existing weeds.

Weedy rice used to be described as *Oryza sativa* f. *spontanea* (Cao et al., 2008) since it is a close relative of cultivated *Oryza sativa* but differs considerably against cultigens in morphological and physiological matters. However, nowadays weedy rice is accepted as the same species as cultivated rice, and therefore taxonomically described as *Oryza sativa.*

There are two major domesticated rice species in the world: *O. sativa* also referred to as Asian rice and *O. glaberrima* the African rice. The Asian rice was domesticated from *O. rufipogon*, its wild ancestor in South East Asia whereas *O. glaberrima*, the African variety has been bred from *O. barthii* (Gross et al., 2010). These two domestication lines are completely independent. This work focuses on weedy rice in Northern Italy, and therefore *Oryza sativa* weedy rice and cultigen accessions.

1.1 "Bad weeds grow tall" – morphological characteristics of Weedy Rice

Weedy Rice is characterised by numerous morphological and physiological traits in which it differs from rice cultivars. The following paragraph describes the major traits observed in weedy rice, which were also subject of the investigations in this study.

The characteristic, that is most eye-catching, is that weedy rice grows considerably taller that cultivated rice. Figure 1.1 shows two rice paddies next to each other. The left one is free from weedy rice whereas the one to the right shows a high rate of weedy rice infestation. The difference in the height of the plants is clearly visible. Most of today's cultivars are small, so called "semi dwarf" varieties (Asano et al., 2011). This phenotype is based on mutations in the SD1 gene, a gibberellin oxidase gene, which is also referred to as the "green revolution" gene. The green revolution designated a group of different traits achieved by breeding that led to a sudden strong increase in yields (Hargrove and Cabanilla 1979, Dalrymple 1986, Khush 1999). One big advantage of the smaller plants is the reduced breaking of the stems when the plants are exposed to windy weather conditions. But due to reduced height, the plants lose their ability to compete for light against rivals. Of course, in an artificial system like agriculture where only plants of the same height are gathered this is no disadvantage but in the competition with the taller weedy rice, cultivars are easily outcompeted.

Figure 1.1 Rice paddies in the Piemonte region. On the left field an herbicide resistant cultivar was sown, therefore no weedy rice infestations are visible. On right paddy a conventional cultivar was sown, showing high infestation with weedy rice. Picture by Annabelle Grimm, October 2010.

Weedy rice is also known by the term "red rice". This name derives from the red or brown colour of the pericarp typical for weedy rice accessions (figure 1.2). The red pericarp is nothing uncommon in rice species. Both ancestors of the two major rice cultivation events *O. rufipogon* and *O. barthii* originally have red grains. In African rice actually most of the *O. glaberrima* cultivars still have a red pericarp. The motif for selection on white pericarp

Figure 1.2 Rice seeds with red pericarp. Picture by Annabelle Grimm, January 2014.

is not fully understood but one hypothesis is that contaminations occurring during storage are more easily detected among white grains (Brooks et al., 2008). Other sources suggest that the proanthocyanidins in the grain, that determine the colour, are a factor involved in the dormancy of rice seeds (Sweeney et al., 2007). However, for agriculture dormancy is an unprofitable trait because a dependable level of germination is required to maintain stable agriculture. Weedy rice seeds often show high levels of dormancy. They depend on that trait due to their parasitic lifestyle and might therefore favour red pericarps even if it reduces the perfection of the mimicry of cultivated rice.

Seed dormancy is a trait often observed in wild plants. It describes a physiological state in which seed germination does not automatically occur in the next season but is delayed and triggered by certain environmental conditions. Dormant seeds can outlast for years in the soil until they finally germinate. After infestation of a paddy field, dormant weedy rice seeds can build up a seed bank in the soil. This makes the management of weedy rice extremely difficult because the dormant seeds will not be destroyed by the pest control methods and produce new weedy rice plants in the future. This makes dormancy an essential trait for weedy rice to ensure continued existence, despite weed treatment.

Another significant trait of weedy rice is the seed shattering, a natural mechanism of seed detachment after ripening. Shattering is a common feature in many wild plants because it ensures seed dispersal. The aim of cultivated crops is to harvest the mature grains; hence, shattering was reduced during domestication. For weedy rice, shattering is important to distribute the seeds in the environment. The problem for agriculture is that the shattered seeds reach the soil and build up a seed bank of weedy rice in the paddies.

Tall growth, pericarp pigmentation, dormancy and shattering are the most important characteristics of weedy rice. They facilitate the pest character of weedy rice and maintain its lifestyle. These key attributes are not the only peculiarities found in weedy rice. Other traits also discriminate weedy rice from cultigens and also different weedy rice populations among themselves.

One example is the awnedness of seeds. Awned seeds are characteristic for many grasses but are also known from other plants. Wild rice species usually have awned seeds. The exact function of the awn is not one hundred percent clear but most likely it is a structure important for the spread of the seeds e.g. by sticking to the fur of animals (Grundbacher, 1963). Most *O. sativa* cultigens are awnless, whereas most wild species have awns. In weedy rice many different groups can be distinguished by awn morphology (figure 1.3). The types range from short straw coloured awned to long black awned but also awnless varieties have been reported (Fogliatto et al., 2012). Not only the pericarp, visible at dehulled seeds, but also the seed hull can be of different colours. *O. sativa* cultigens mostly have a strawhull phenotype. Wild rice species often show brown or black seed hulls. In Weedy Rice both states can be found. The hull coloration has been an important trait to distinguish weedy rice populations in the US (Londo and Schaal 2007).

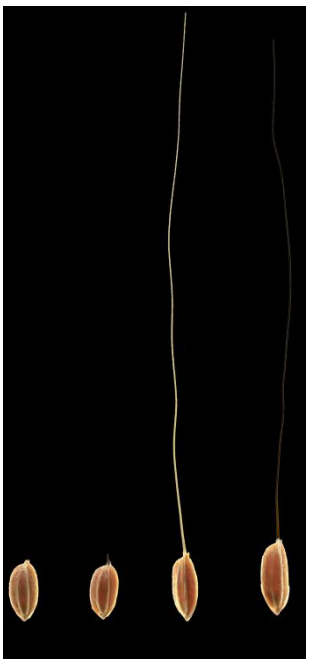

Figure 1.3 Awn types of weedy rice (l.t.r): awnless, mucronate, straw awned, black awned. Grimm at al. (2013)

1.2 Agricultural weeds – Crop wild relatives or crops gone wild?

Agricultural weeds occur in all cropping systems worldwide. It is widely accepted that the evolution of crop weeds is driven by human action (Dekker, 2011). The changes in nature made by man, turning areas into arable soils, provide what is referred to as opportunity space (Gressel and Valverde, 2009). The opportunity space is taken by weedy plants which then further evolve according to the selection pressures applied by human agriculture. Crops are species that were domesticated over centuries or even millennia by selecting for favoured traits. In most cases domestication is accompanied by the accumulation of rare recessive traits, also known as domestication syndrome (Gressel, 2005). The dedomestication of crops to weeds often can be achieved by dominant traits which are more effective in selection. Due to this fact the evolution of weeds is more rapid than the domestication process. The general mode of weed evolution was outlined by Gressel (2005), suggesting that the first step is the formation of a so called "volunteer weed", a crop that acquired one trait by mutation or hybridisation, that would provide an advantage in its environment, e.g. shattering. These volunteer weeds evolve by accumulating further fitness enhancing traits (e.g. tall growth, dormancy...) leading to a dedomestication of a cultigen and in the end of the process to a new weed.

The result of the process of dedomestication is often referred to as "ferality" (Gressel 2005). Two models for ferality are proposed for the evolution of crop weed species (figure 1.4). The model of endoferality describes weeds that directly descended from cultigen ancestors. In contrast, exoferality is a process in which hybrids of the crop and a wild species result in weedy descendants (Ellstrand et al., 2010). A special case is the exo-endoferality, where a hybridization of two crops initiates the formation of a weed. An example was shown by Ishikawa et al., (2005) for a population of weedy rice in China which most likely originated from a cross of *indica* and *japonica* rice.

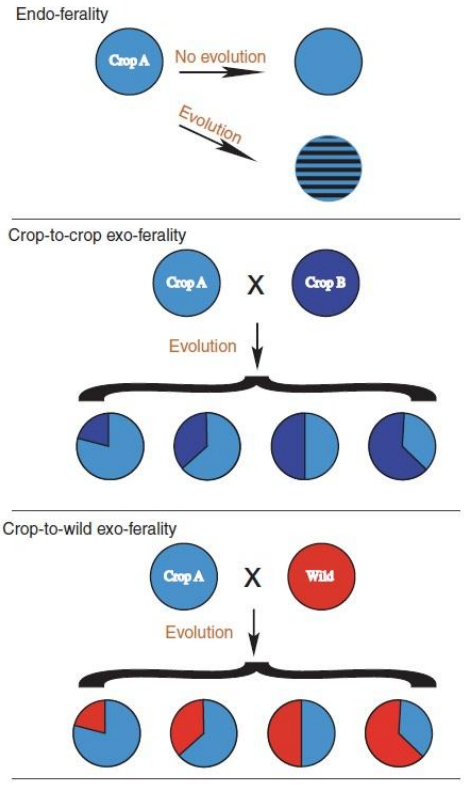

Figure 1.4: from Ellstrand et al., 2010
Models for the dedomestication of crops to weeds by endoferality and exoferality.

This section shows that agricultural weeds evolve in response to crop cultivation as a continuous process. The weeds adapt in response to weed management and agriculture (figure 1.5), a model also known as "weed management arms race" (Neve et al., 2009). This process could also been shown for weedy rice in recent studies. Gene flow between cultigens and weedy rice was reported, which might enable cultigens to pass on herbicide resistance (HR) genes to weedy rice (Shivrain et al., 2007, Zuo et al 2011). It also was reported that HR genes could provide beneficial effects, additional to the herbicide resistance, enhancing the fitness of weeds even in the absence of the herbicide (Wang et al., 2013).

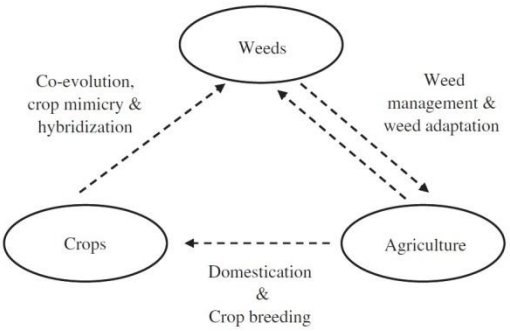

Figure 1.5: from Neve et al., 2009
Model for the "weed management arms race". Weeds evolve to adapt to weed management and the crop as a competitor. The crops are adapted by man in agriculture to achieve advantage over the weeds. This process is continuous, because a change in one of the components will result in the adaptation of the others in response.

1.3 Evolution of Weedy Rice

An insight in the evolution of crop weeds is important for their handling in the future. Rice is one of the best studied crops, and in some parts of the world has a good documented history. These preconditions make weedy rice an ideal model to study the parallel evolution of weeds and their crop ancestors.

Today it is commonly accepted that weedy rice is not a phenomenon with a single origin but evolved independently in different rice cultivation areas driven by human forces (see 1.2). Three models for weedy rice origins are discussed in the literature (DeWet and Harlan, 1975).

a) Weedy rice is a wild species that invaded the nutrient rich paddies and became competitor to cultivated rice. This model represents the case of exoferality.

This hypothesis suggests that weedy rice as a representative of the so called "crop wild relatives" These are species that are closely related to cultivated species and are found in the agricultural environment of the affected crop, benefiting from the favourable conditions in agricultural systems. In this case they act as pests, causing yield losses by competing with the crops. On the other hand they are often investigated as germplasm sources for improved crop species. Facts that support this hypothesis are surely that the main morphological and physiological characteristics of weedy rice (red pericarp, tall plants, shattering, dormancy) are also typical for wild rice species. In South and South East Asia *O. rufipogon*, the ancestor of most of today's rice cultigens is endemic to the surrounding landscape of rice cultivation areas and could easily invade paddies. In Europe and on the American continent no wild rice species are endemic but they could have been imported with contaminated seeds from Asian countries.

b) Weedy rice descends from hybrids of cultivated rice with wild species resulting in weedy forms of the domesticated crop. Another form of exoferality

Although rice usually has a high inbreeding nature geneflow between rice cultivars and wild rice species is possible. Studies with microsatellite markers revealed a tight relation between weedy rice and cultigens in some areas (Cao et al 2008, Grimm et al., 2013). Taking into account this similarity plus the morphological characteristics that resemble wild rice species hybrids offer a satisfying hypothesis for weedy rice evolution. Hybridisation of cultivars and wild rice could easily have taken place in Asia where both species are endemic in the same geographical areas. In other ecosystems where no wild rice is found, again contaminations in imported seeds might have provided the source for wild rice species in these areas.

c) Weedy Rice evolved from cultivars, an endoferality model.

The third possibility that is taken into account is that weedy rice could be a mutated and therefore "dedomesticated" form of cultivated rice. This theory would also explain the genetic similarity of weedy rice and cultigens shown in several cases. Since wild rice is the genetic ancestor of our recent cultigens, the traits that are characteristic for wild and weedy rice, might be readaptations achieved by (back) mutations. For the grain pigmentation (red pericarp) at least two examples in the literature show that cultivated rice reacquired a "wild trait", namely the red pericarp. In both examples cultigens with white grains produced forms with red pericarp by mutation. This was observed in the US and Italy, respectively (Brooks et al., 2008, Gulick et al., 2009). The red grained forms were not considered to be weedy rice, because no other characteristics for a weedy lifestyle were observed, but further mutations accumulating in the gene pool and enhancing the fitness over the original cultivar are imaginable.

1.4 Dolce Vita in the Rice Paddy – Weedy Rice and Agriculture

Rice has been domesticated by mankind for the last 8,000 to 10,000 years, starting in South and South East Asia with *O. rufipogon* as a common ancestor for most of today's *O. sativa* cultivars (Londo et al., 2006) splitting into *indica* and *japonica* subspecies. Weedy rice has been reported for a long time but was negligible due to an effective weed management in seedling state (Cao et al., 2008).

In former times, in rice agriculture, seedlings were grown until a certain stage and then transplanted into the paddy fields. Plants that had an unusual or suspect shape were discarded and never reached the paddy soil. This transplanting technique kept weeds auf of the rice paddies for a long time. In the last decades the sowing methods in many parts of the world changed from the transplanting method to direct seeding (Grimm et al., 2013). The spread of the seeds on the soil without effective control mechanisms promoted the colonisation of the paddy fields by weeds originating from contaminated seed stocks. Simultaneous with the shift in the sowing method, the reports of weedy rice infestations in the literature increase. In the year 2002 Europe reached an infestation of 70% of the total rice producing area (Catala et al., 2002). Another favourable fact for weedy rice is the likewise increasing rate of monocropping systems providing a stable environment of beneficial conditions for the weeds.

The control of weedy rice turns out to be problematic. Since weedy rice and cultigens are the same species (*Oryza sativa*) most herbicides are not applicable. A pre-treatment of the soils can show a short-term effect but is only semi-fruitful due to the dormant nature of weedy rice seeds. The dormant seeds are not destroyed by the herbicides and can germinate in following seasons producing new seeds that enter the seed bank. Approaches with herbicide resistant rice varieties such as BASF Clearfield® or genetically modified cultigens are the most effective method for control to date but can also do not fully recover infested soils. The crux is that geneflow between weedy and cultivated rice has been reported and recent studies showed that resistances from GMO (genetically modified organism) rice can be passed to weeds (Wang et al., 2013). Therefore most seed producing companies that offer herbicide resistant rice (GMO or classical bred lines) only allow the use of their varieties for one or two years in a row on the same paddy, which is greatly beaten by the dormancy abilities of weedy rice seeds.

At the present state, crop rotation, also with herbicide resistant varieties, and the pre-treatment of soils before seeding are the most efficient methods to at least limit the infestations with weedy rice and reduce the yield losses.

1.5 A short History of (Weedy) Rice in Italy

Rice cultivation in Italy was most likely introduced between the 13th and 15th century (Faivre-Rampant et al., 2011). With the expansion of trade of European countries with Asia, rice was imported and reached Europe at first in Italy. To date, Italy still is the main rice producer on the European continent (Fogliatto et al., 2010). For the first 200 – 300 years seed material was exclusively imported from Asia and sown in Italy. It was not before the early 19th century that Italy started to invent local breeding programs, producing the first Italian cultivars. This point in history coincides with the first citation of weedy rice in the literature by Brioli in 1807. Weedy rice remained a marginal problem for the next approximate 150 years. From the 1960s on the infestation rates of weedy rice exploded, reaching almost two-thirds of Europe's rice paddies until today (Fogliatto et al., 2010). This effect was caused by the shift in the sowing method described in section 1.3.

Since Italy, compared to other regions in the world, relatively recently joined the league of rice producing nations, the history of breeding and the cultigens is well documented. Two large review papers gather the information available on Italian rice cultivars and display their ancestry (Spada et al., 2004, Faivre-Rampant et al., 2011). Many of the ancient cultivars are still available today, and used in breeding, but they also are a valuable source for research with a historical background. Some of the above mentioned historical cultivars were available and used for this study.

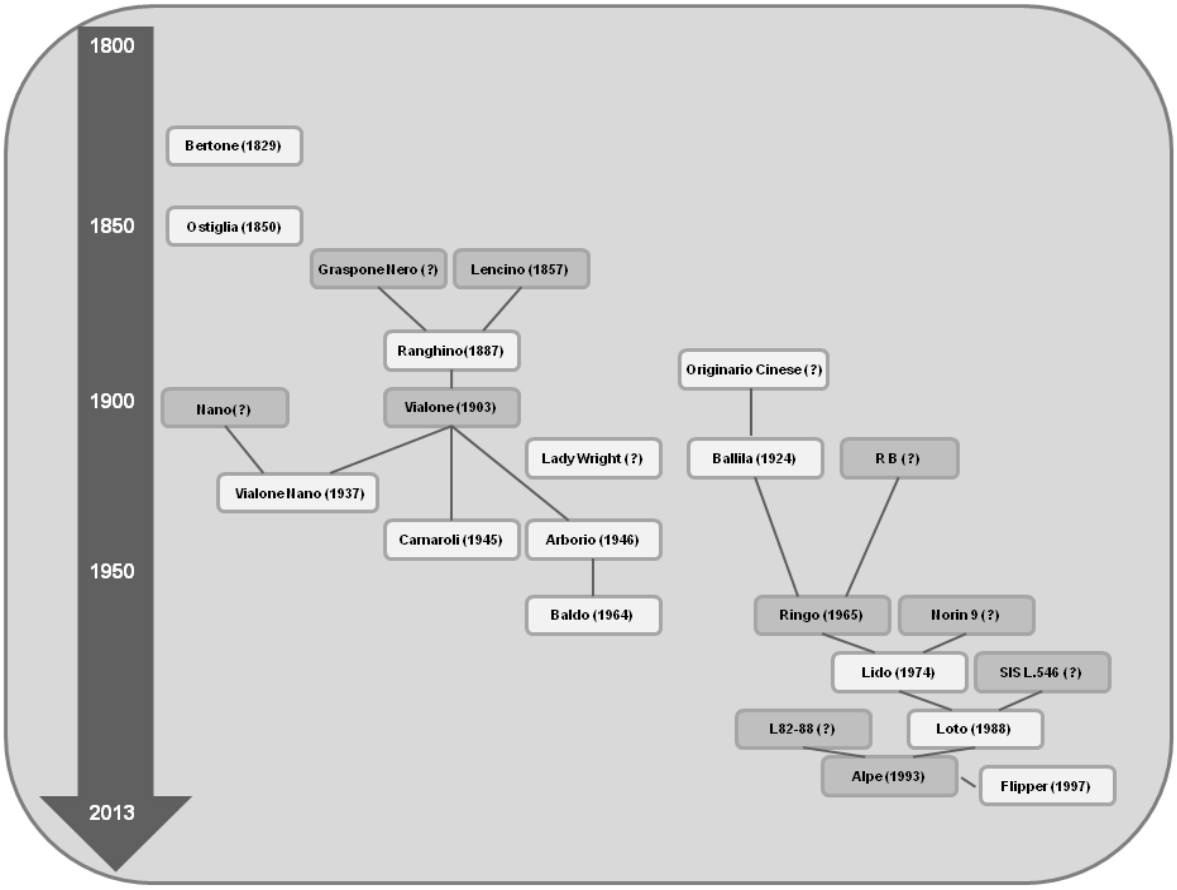

Figure 1.6 Timeline course from 1800 until today (arrow on the left side). Cultigens are shown in their ancestral context. Their date of first description is shown in brackets. Cultigens used in this study indicated by white boxes, accessions in grey boxes were not available but displayed for a better visualisation of the relations.

Figure 1.6 lists some of the Italian cultivars used in this study in a timeline, showing the year of release and their ancestry, if available. The oldest cultivars in our set are Bertone and Ostiglia, both dating back to the early 19th century and therefore marking the beginning of rice breeding in Italy. Most other varieties are separated in two clades: the "Ranghino clade" and the "Ballila clade". Ranghino and four of its later relatives from the late 19th century until the 1960s (Baldo) were investigated in this survey. The Ballila variety is the offspring of Originario cinese. The age of Originario cinese could not be extracted from the documents available but it is an old variety which was imported from China to be grown in Italy and hence one of the precursor varieties of Italian cultivars. The other varieties of the Ballila group are Lido, Loto and Flipper each representing one decade of the 20th century from the 70s to the 90s. The newest cultigen in the timeline is the Clearfield rice from BASF, released in 2008 (www.asiapacific.basf.com). This variety is herbicide resistant, but not GMO, to several BASF herbicides. Further varieties used in this study (table 2.2) could not be included in the timeline because the information required was not available.

1.6 Scope of the Dissertation

Italy is Europe's most important rice producing country. In the last decades the weedy rice infestations increased, leading to severe yield losses and large financial deficits. Due to the lack of efficient management strategies weedy rice cannot be successfully treated and the infestation rates further increase. Many studies have been conducted on weedy rice, addressing its characteristics and evolutionary origin in different geographical areas (Cao et al., 2006, Londo and Schaal 2007, Jiang et al., 2012). However, knowledge about weedy rice in Italy is limited. A more detailed characterisation of weedy rice, revelation of its origins and modes of evolution could be the foundation for efficient weed management strategies in the future. Furthermore this knowledge could be used as a general example for the evolution and structure of agricultural weeds and help to improve their management in the future. Therefore, the main aim of this study was to determine different groups of weedy rice in Northern Italy and investigate their evolutionary origins.
The study can be divided in three levels, building up on each other.

1.) SSR analysis

This first part of the study was carried out in collaboration with the group of Professor Aldo Ferrero from the Torino University in Italy.
An SSR approach was conducted to get a first impression of the structure of weedy rice populations in the study area and their relation to the cultigens from this geographical region in past and present. In total 150 accessions and 20 cultivars were investigated using 19 SSR markers described in Cao et al., (2006). With the data obtained in this analysis different values addressing diversity, phylogeny and population structure were calculated. The genetic diversity, the heterozygosity and other statistics were obtained from the SSR data to get an impression of the degree of differentiation of weedy rice samples among each other and compared to the cultigens. To investigate the differentiation of populations, the fixation index (F_{st}) was estimated. To get reinforce the insight in the population structure of our sampling; the most probable number of populations among the accessions was defined using Bayesian clustering. The phylogeny based on the genetic distance was computed to overview the genetic relationship between weedy rice and cultigens. The results were compared with morphological data from earlier studies of the Ferrero group to test if genetic and morphological distances are in context. A spatial structure analysis was used to check for a

possible link of genetic and geographic distance and eventually define a geographical origin of evolution for weedy rice in Italy.

2.) Genetic (next generation sequencing) and morpho-physiological investigations of characteristic weedy rice traits

The main focus still was on Italian weedy rice in this part but additionally we included weedy rice and cultigens from both, Brazil and Thailand. As a representative of wild rice *O. rufipogon* was also included in our accession setup. This enlargement of the sample spectrum aimed to provide a better comparability of the data. The goal was to gather information on the genetic and morpho-physiological distinctness for traits characteristic for weedy rice. Since these traits are not only typical for weedy rice but also important for their maintenance they are referred to as "weedy lifestyle traits".

In total 7 genes in 55 accessions were sequenced using the Illumina Miseq next generation sequencing technique. The results were grouped according to the traits encoded and compared with morphological and physiological data.

The dwarfing gene *SD1* was investigated for single nucleotide polymorphisms (SNPs) and the results were compared to data of fully grown plants of the related accessions.

The gene *Rc*, responsible for grain pigmentation in red rice, was investigated for functional mutations. Additionally seed extracts were analysed both qualitative, using HPLC, and quantitative (Vanillin Assay) for the pigments located in the pericarp.

To examine the shattering trait in our sampling, three genes (*qSH1*, *SHAT1*and *sh4*) were sequenced and checked for SNPs in the coding regions. For further analysis 2D and 3D microscopy of the abscission zone at the grains was performed on selected accessions.

Two genes associated with dormancy (*VP1* and *SDR4*) were chosen for the investigation of dormancy in weedy rice. Again the analysis of SNPs or other functional mutations was focused on the coding region. Furthermore the possibility of dormancy breaking in weedy rice with karrikins (dormancy breaking molecules) was screened by germination assays. The differences in germination in water (control) and karrikin containing solutions (smoke water) were determined.

3.) Reconstruction of a model for the evolution of weedy rice in Italy

Based on the data of the practical approaches (1 & 2) a model for the evolutionary history of weedy rice in Italy was constructed.

2 Material and Methods

2.1 Accessions used in this study

In this study, weedy rice accessions and cultigens from Italy, Thailand and Brazil, representing three different continents, were used. Since the focus of this study was on weedy rice in Italy, most accessions are from this area. The Italian weedy rice accessions were collected in the Piemonte and Lombardia region in northern Italy, the main rice growing area, in 2009 (Fogliatto et al., 2011). The sampling region was divided into 10 sections (see fig. 2.1) and 150 accessions were collected from this area in total.

Figure 2.1 Map of an overview and a detail of the Piemonte (right) and Lombardia (left) regions in Northern Italy. The collection areas for the samples of this study are located in the 10 squares of the grit in the border area between the regions (Grimm et al., 2013).

The tables 2.1 and 2.2 list the weedy rice accessions with the section of collection, and the cultivars from Italy used in this study. The SSR amplification and sequencing for all 150 accessions and the cultivars were carried out by our cooperation partners in Torino, Italy, with seeds from their own stock. For a detailed description see Grimm et al., (2013). For forty weedy rice accessions of the collection (table 2.1, accessions in bold) and for all of the cultivars (table 2.2) ten seeds were sown in the botanical garden of the Karlsruhe Institute of Technology (KIT) in October 2010. The seeds were harvested in June 2011 and stored at 7°C until use. These accessions were used for the investigation of genes linked to typical weedy rice traits with next generation sequencing (NGS) and the replenishing assays.

Table 2.3 shows all non-European rice accessions, weedy rice and cultigens, used for comparison in this study. Cultivars and weedy rice from both, Thailand and Brazil were used. The table shows the region of origin and the provider of the accessions.

Table 2.1 Accessions used in this study and their affiliation to the sampling locations (figure 2.1). All accessions were used in the SSR analysis, accessions in bold were also used in the Next Generation Sequencing approach and further experiments.

Number	Zone	Number	Zone	Number	Zone	Number	Zone
TO1	**4**	TO39	2	TO77	8	TO115	8
TO2	4	TO40	2	TO78	8	**TO116**	**8**
TO3	4	**TO41**	**2**	**TO79**	**8**	TO117	8
TO4	4	TO42	2	TO80	8	TO118	8
TO5	**7**	**TO43**	**1**	**TO81**	**8**	TO119	10
TO6	7	TO44	1	TO82	8	TO120	10
TO7	7	TO45	1	TO83	4	TO121	10
TO8	7	**TO46**	**2**	TO84	8	**TO122**	**10**
TO9	7	TO47	2	TO85	8	TO123	10
TO10	7	TO48	2	TO86	8	TO124	10
TO11	**7**	TO49	2	**TO87**	**8**	TO125	10
TO12	7	TO50	2	TO88	8	**TO126**	**10**
TO13	**7**	TO51	2	**TO89**	**8**	TO127	10
TO14	7	**TO52**	**2**	**TO90**	**8**	TO128	10
TO15	7	TO53	2	TO91	8	TO129	10
TO16	7	**TO54**	**2**	TO92	8	TO130	10
TO17	7	TO55	2	TO93	8	TO131	10
TO18	7	TO56	4	**TO94**	**9**	TO132	10
TO19	**7**	**TO57**	**4**	TO95	9	**TO133**	**10**
TO20	**3**	TO58	4	TO96	9	TO134	10
TO21	3	TO59	4	TO97	9	**TO135**	**10**
TO22	3	**TO60**	**5**	TO98	9	TO136	10
TO23	3	TO61	5	TO99	5	TO137	10
TO24	**3**	TO62	5	TO100	5	TO138	6
TO25	3	TO63	5	**TO101**	**9**	TO139	6
TO26	**3**	**TO64**	**5**	TO102	9	**TO140**	**6**
TO27	3	TO65	5	TO103	9	TO141	6
TO28	3	TO66	5	TO104	9	TO142	6
TO29	3	**TO67**	**5**	TO105	9	**TO143**	**6**
TO30	**3**	TO68	5	**TO106**	**9**	TO144	6
TO31	3	TO69	5	TO107	9	TO145	6
TO32	**1**	TO70	2	**TO108**	**9**	**TO146**	**6**
TO33	1	TO71	2	TO109	9	TO147	3
TO34	1	TO72	2	TO110	9	TO148	3
TO35	1	**TO73**	**2**	TO111	9	TO149	3
TO36	1	TO74	4	TO112	9	TO150	3
TO37	1	**TO75**	**4**	**TO113**	**9**		
TO38	**1**	TO76	4	TO114	8		

Table 2.2 Cultigen accessions used in this study and the year of first description if available.

Name	Year of Appearance	Name	Year of Appearance
Arborio	1946	Gladio	-
Artiglio	-	Lady Wright	(no longer in use)
Baldo	1964	Loto	1988
Ballila	1924	Originario Cinese	-
Bertone	1829 (no longer in use)	Ostiglia	1850 (no longer in use)
Carnaroli	1945	Prometeo	-
Centauro	-	Ranghino	1887
Clearfield	-	Selenio	-
Creso	-	Thaibonnet	-
Flipper	1997	Vialone Nano	1937

Table 2.3 List of the accessions from Brazil and Thailand used in this study. Cultigens and weedy rice are shown in the same table. The number represents the accession number of the botanical garden of the KIT (BG). The provider is the institution that donated the seed material for this accession.

Number	Identity	Region	Provider
BG2	Thai Cultigen	Thailand	Kasetsart University, Thailand
BG3	Thai Cultigen	Thailand	Kasetsart University, Thailand
BG4	Thai Cultigen	Thailand	Kasetsart University, Thailand
BG5	Thai Weedy Rice	Thailand	IRRI
BG6	Thai Weedy Rice	Thailand	IRRI
BG7	Thai Weedy Rice	Thailand	IRRI
BG49	Brazil Cultigen	Brazil	IRRI
BG51	Brazil Cultigen	Brazil	IRRI
BG54	Brazil Cultigen	Brazil	IRRI
BG57	Brazil Cultigen	Brazil	IRRI
BG58	Brazil Cultigen	Brazil	IRRI
BG67	Brazil Weedy Rice	Brazil	IRRI
BG68	Brazil Weedy Rice	Brazil	IRRI
BG69	Brazil Weedy Rice	Brazil	IRRI

2.2 SSR Analysis

The amplification and sequencing of the SSR markers were conducted by our cooperation partners from the Torino University. The data analysis was done at the KIT with different bioinformatics open source software applications. The paragraphs describing DNA extraction (2.2.1) and microsatellite PCR (2.2.2) are taken from Grimm et al., (2013).

2.2.1 DNA Extraction

Three seeds from each Weedy Rice accession and cultigen, were respectively sown on cell trays in the greenhouse, harvested at the 3-leaf stage, shock-frozen in liquid nitrogen, and stored at -80 °C. Seedlings were ground using a high-throughput disruptor (TissueLyser, Qiagen, Germany), and the DNA was extracted using the DNeasy Plant Mini Kit (Qiagen, Germany, 2000) as follows: after homogenisation, 5 ml of pre-heated (65 °C) extraction AP1 buffer, and 10 µl of RNase A were added to about 1 g of each sample and vortexed at high speed. The samples were incubated at 65 °C for 10 minutes and mixed periodically by inverting the tube to lyse the cells. Then, 1.8 ml of AP2 buffer were added to each sample, followed by mixing and 10 min of incubation on ice. Next, samples were centrifuged at 5000 g (Eppendorf, 5417 R) for 5 min at 20°C, and the supernatant filtered through a QIAshredder Mini spin column and placed in a collection tube and immediately spun at 5000 g for 5 min at 20°C. The lysate was transferred to a 2 ml-tube without disturbing the pellet containing some cell debris. A mixture of 1.5 volumes of AP3 buffer and ethanol was added to the samples and mixed. The samples were transferred into a DNeasy Mini Spin Column in a reaction tube, centrifuged at 5000 g for 2 min, the flow-through discarded, and 12 ml of Buffer AW (wash buffer) added to the DNeasy Mini Spin Columns. After centrifugation for 2 minutes at 5000 g to dry the membrane, the flow-through was discarded, and the DNeasy Mini Spin Column with the processed sample transferred to a new 50 ml-tube; 1 ml of AE buffer (elution buffer) was added to the column membrane and incubated at 20°C for 5 min. After centrifugation for 2 minutes at 5000 g for elution, further 750 µl of AE buffer were added and centrifuged again at the same speed. The two eluates were pooled and stored at -20°C.

2.2.2 Microsatellite PCR

Nineteen microsatellite markers (SSRs) spread over the different rice chromosomes and tested in previous studies (Cao *et al.,* 2006), were selected for this work. The corresponding oligonucleotide primers can be found in the appendix (p. 87, table 5.1). The PCR reaction mixture (25 µl) contained 20 mM TRIS, 50 mM KCl, 3 mM$MgCl_2$, 0.25 mM of dNTPs, 1 µM of each oligonucleotide primer, 1 unit Taq polymerase (Taq DNA polymerase, Qiagen), and 3 µl of a 1:10 dilution of ca. 10 ng/µl genomic DNA. PCR was conducted after 3 min at 94 °C in 40 cycles at 92 °C for 1 min, annealing at 56°C for 1 min, synthesis at 72 °C for 1.5 min, and a final cycle of 10 min at 72 °C. Amplificates of different sizes and fluorescent labels were multiplexed and run on a capillary electrophoresis sequencer/genotyper, ABI 3730xl (Applied Biosystems). GeneMapper® Software Version 4.0 (Applied Biosystems) was used to identify the alleles and the sizing was carried out with the internal size standard GS500LIZ (Applied Biosystems) containing 16 fragments with a size ranging from 35 to 500 bp (Chen 2011). Amplification anomalies were detected for the three markers RM14, RM44, RM55.

2.2.3 Analysis of the SSR data

The raw SSR data were first analysed using the open source software Identity (http://www.uni-graz.at/~sefck/). Allele number, expected and observed heterozygosity (H_e, H_o) were estimated using the software. Genetic distance of the accessions based on SSR allele lengths, the Chord distance (Cavalli-Sforza & Edwards 1967), was calculated using the Microsat software (http://hpgl.stanford.edu/projects/microsat/) using. The overall F_{st} values were calculated using the GENEPOP software version 4.0.10 (Raymond& Rousset 1995). The software Structure 2.2 (Pritchard *et al.,* 2000) was used to identify a model-based (Bayesian clustering) genetic structure in the SSR data. This program uses multi-locus genotype data to investigate population structure by a Markov-Chain Monte Carlo algorithm (MCMC), clustering individuals into K distinct populations by minimising Hardy-Weinberg disequilibrium and linkage disequilibrium between loci within groups. The program was run at the default settings with an initial burnin period of 20,000 followed by 100,000 MCMC repeats. K was estimated from 2 to 10 and calculations for each K were iterated 20 times. The best fitting K was calculated by the Evanno method using the website program STRUCTURE HARVESTER (Earl, 2012).

The phylogenetic tree was constructed using the MEGA 5.0 software (Tamura et al., 2007). A distance tree was calculated using the UPGMA algorithm on the genetic distances obtained from the SSR data for the 150 Weedy Rice, and the 20 cultigen accessions. The depiction in

circle topology was chosen for better visualization of the large number of accessions displayed in the tree.

2.2.4 Biogeography

A spatial structure analysis as described in Shivrain et al., (2010) was performed on the SSR data. This method compares genetic and geographical distance testing for correlation. The genetic distance was plotted against the geographic distance obtained from the GPS data of the collection sites provided by our cooperation partner from the Torino University, Italy. The autocorrelation coefficient, R, which is a measure for correlation of the data, was calculated (Microsoft Excel 2007).

2.3 Next generation sequencing of "weedy lifestyle genes"

This project was carried out at the University of Georgia (Athens GA, USA) in Prof. Katrien Devos' lab. The exchange was financially supported by the Karlsruhe House of Young Scientists (KHYS, KIT, Karlsruhe, Germany). Genes that are linked to traits that are characteristic for weedy rice and its lifestyle were sequenced in this part of the work. They are further referred to as "weedy lifestyle genes".

2.3.1 Growing of rice plants

To grow rice plants for tissue material required for DNA extraction, seeds were dehulled and sterilized. For sterilization dehulled seeds were covered with 70% ethanol (Roth, Karlsruhe, Germany) for one minute. The ethanol was removed and the seeds were rinsed with ddH_2O two times. A 6% sodiumhypochloride solution (Roth, Karlsruhe, Germany) was added and the seeds were incubated in the solution for 20 minutes at RT on a shaker running with 100 rpm. The hypochloride solution was removed and the seeds were washed four times with sterile ddH_2O. Twenty seeds of one accession were sown in one Magenta box (Sigma, Taufkirchen, Germany) containing 100 ml of 0.4% phytoagar (Duchefa Biochemie, Harleeem, Netherlands). The boxes were kept in a constant room at 29°C and constant light for 10 days until harvesting.

2.3.2 DNA Extraction

The DNA extractions in this study were carried out after the method of Doyle (1987) with modifications. One hundred mg of young leaf tissue were put in a reaction vessel with a 5 mm steal ball and frozen in liquid nitrogen. The plant tissue was powdered in the Tissue Lyzer (Qiagen, Hilden, Germany) for 20 sec at 21 Hz. 700 µl of warm CTAB buffer: 67°C, 3% w/v CTAB, 1.4 M NaCl, 0.3 M Tris-HCl (pH 8.0) and 25 mM EDTA (all Roth, Karlsruhe, Germany) and 1.75 µl proteinase K solution (20 mg/ml) were added and the samples were incubated at room temperature for 10 min on a shaker (Eppendorf Thermomixer 5436, Eppendorf, Hamburg, Germany). The samples were centrifuged for 5 min at 2,400 * g and 700µl of chloroform:isoamyl alcohol 24:1(v/v, Roth, Karlsruhe, Germany) were added to the upper phase in a new reaction vessel. The samples were mixed, incubated for 10 min at room temperature in a shaker and centrifuged for 10 min at 2,400 * g. Again the upper phase was used for further processing and 0.1 volumes 3M LiCl and 0.6 volumes of ice cold isopropyl alcohol (Roth, Karlsruhe, Germany) were added and mixed. The mixture was incubated for one hour on ice to precipitate the DNA and afterwards centrifuged for 10 min at 5,400 * g.

The pellets were washed with 70% ethanol and dried in a Speedvac (Eppendorf Concentrator 5301). The DNA was dissolved in 100 µl of millpore H_2O and the concentration was determined using the Nanodrop 1000 (Thermo Scientific, Karlsruhe, Germany).

2.3.3 PCR amplifications

Different types of PCR, standard and touchdown PCR, were carried out during this study. In a touchdown PCR the annealing temperature decreases with each cycle until the target temperature is reached. These cycles are followed by a standard PCR with the target temperature used for annealing. This method can be used to avoid unspecific products. The tables (2.4 – 2.7) list the composition of the reagents and the programs used. PCR amplification was carried out in the Tetrad 2 DNA Engine (Biorad, Hercules, CA, USA). Most primers in this study were designed using the Primer 3 software (Koressaar et al.,, 2007, Untergasser et al.,, 2012) using the *Oryza sativa* japonica genome from Rap-DB (Kawahara et al., 2013, Sakai et al., 2013). Primers from the literature are cited for their source. A detailed list with all the primers used in this study containing sequence, annealing temperature, product length and source can be found in the appendix (p.88, table 5.2). Primers were synthesized by Sigma-Aldrich (Saint Louis, MO, USA). The GoTaq Flexi DNA Polymerase (Promega, Madison, WI, USA) and the corresponding reaction ingredients were used for master mixes.

Table 2.4 Reagents used for one reaction of the standard PCR.

Reagent	Concentration	Volume
H_2O		3.72 µl
Buffer	5x	3 µl
Betaine	5 M	1.5 µl
Template	50 ng/µl	1.5 µl
Primer fw	10 pmol	1.2 µl
Primer rev	10 pmol	1.2 µl
dNTPs	25 mM each	1.2 µl
$MgCl_2$	25 mM	0.9 µl
DMSO		0.6 µl
Taq Polymerase	5u/µl	0.18 µl
Total volume:		**15.00 µl**

Table 2.5 Reagents used for one reaction of the touchdown PCR

Reagent	Concentration	Volume
H_2O		5.82 µl
Buffer	5x	3 µl
Template	50 ng/µl	1.5 µl
Primer fw	10 pmol	1.2 µl
Primer rev	10 pmol	1.2 µl
dNTPs	25 mM each	1.2 µl
$MgCl_2$	25 mM	0.9 µl
Taq Polymerase	5u/µl	0.18 µl
Total volume:		**15.00 µl**

Table 2.6 Program for the cycler for the standard PCR

Temperature	Time	Cycles
95°C	05:00 min	
95°C	00:30 min	35x
59°C	00:30 min	
72°C	02:00 min	
72°C	10:00 min	
04°C	hold	

Table 2.7 Program for the cycler for the touchdown PCR

Temperature	Time	Cycles
95°C	05:00 min	
95°C	00:30 min	10x
59°C – 52°C decreasing 0.7°C per cycle	00:30 min	
72°C	02:00 min	
95°C	00:30 min	35x
52°C	00:30 min	
72°C	02:00 min	
72°C	10:00 min	
04°C	hold	

2.3.4 Library preparation and Sequencing

Several genes referring to traits characteristic in weedy rice were amplified in various accessions and sequenced all together with a next generation sequencing approach in this study. Library preparation and sequencing are described in the following sections. All amplicons of one accession were pooled and pools were purified using the Agencourt AMPure system (Beckman Coulter, Brea, CA, USA). One library was constructed per accession. To be able to distinguish the different libraries after the sequencing, the Nextera XT indexing kit (Illumina, San Diego, CA, USA) was used. Libraries were pooled and all sequenced in one run on the MiSeq system (Illumina, San Diego, CA, USA).

The Nextera XT library preparation kit was used for construction of the libraries. The manufacturer's instructions were followed with some modifications. In general only 0.25x of the amount of reagents recommended by the manufacturer were used in this approach.
2.5 µl of Tagmentation DNA Buffer (TD) were added to each well of a PCR strip. 1.25 µl of input DNA (pooled amplicons of one accession) at 2.0 ng/µl were added to the prepared wells. The mix was completed with 1.25 µl of Amplicon Tagmentation Mix (ATM) and pipetted up and down five times to mix. Strips were sealed, put in a thermal cycler (Tetrad 2 DNA Engine, Biorad, Hercules, CA, USA) and kept at 55°C for 5 min then held at 10°C. In this step, the amplicons are broken down to fragments of ~300bp, which is the optimal size for sequencing with MiSeq, since 150 bp reads from each end are achieved. 1.25 µl of NT Buffer were added to each well and mixed by pipetting up and down five times. Samples were centrifuged quickly in a small bench centrifuge and incubated for 5 min at room temperature. The Nextera PCR Mastermix and indexing primers were thawed and mixed by inverting. 3.75 µl of Nextera PCR Mastermix (NPM) and 1.75µl of each primer (i5 and i7) were added to each well of the strip containing the neutralized tagmented DNA and mixed by pipetting up and down five times. Strips were spun down in a small bench centrifuge and placed in a thermal cycler and the following program was run:
The PCR with indexing Primers creates a tag for each library. The Nextera XT Indexing kit contains 24 primers (8*i5 and 12*i7 primers). Each primer contains the motif for the primers used in the sequencing later on and a variable sequence for the tag. A library tag is created of the combination of the i5 and i7 variable sequence. Thus, 96 tags can be created by using 20 primers.
About 28.75 µl ddH_2O were added to each well after PCR to add up the volume to ~40 µl.

40 µl of Agencourt AMPure beads (Beckman Coulter, Brea, CA, USA) were added to the tubes, mixed well and incubated for 5 min at room temperature. After incubation, tubes were placed on a magnet stand and let rest until the solution had cleared (~ 2 min). The supernatant was removed and discarded. 190 µl of freshly prepared 70% ethanol were added to each sample without removing the strips from the magnet and incubated for 30 sec. The ethanol was removed by pipetting and the step was repeated. Samples were air dried and 23µl ddH_2O was added to dissolve purified DNA. Tubes were placed on a magnet stand again (~2 min) until the solution cleared and the supernatant, now containing the purified library DNA was transferred to a new tube. 2 µl of each library were quantified on the Qubit (Invitrogen) using the high sensitivity kit (dsDNA HS Assay, Invitrogen) and 5 µl were run on a 1.5% agarose gel to visualize the fragment distribution. Libraries which centred around the 300 bp band in the marker were considered as successfully constructed and taken in for the further analysis. Libraries were pooled to equal amounts.

The libraries constructed in this study were sequenced together with other samples in one run on an Illumina MiSeq following the recommendations of Illumina.

2.3.5 Analysis of Next Generation Sequencing Data

NGS data obtained by the MiSeq sequencing were received as two .fastq files for each accession: One containing all forward reads, the other containing the reverse reads. In a first step the reads were aligned to the reference sequence obtained from RAP-DB (Kawahara et al., 2013, Sakai et al., 2013), representing *O. sativa ssp. japonica*. The alignment of the raw sequence data was conducted with Bowtie 2.0 (Langmead and Salzberg et al., 2012). The reference was built from fasta sequences of the genes investigated, taken from RAP-DB (Kawahara et al., 2013, Sakai et al., 2013). The reference was indexed following the instructions in the documentation with default settings of Bowtie 2.0.

The actual alignment was also conducted using default setting in Bowtie. The resulting .bam files, containing the aligned reads for each accession, were organized with SAMtools (Li et al., 2009) using the "sort" command. This process reduces the size of the file requiring less memory in downstream analysis steps. The aligned and sorted files were visualized with IGV (Thorvaldsdóttir et al., 2013). An alignment of all reads of one accession and a consensus sequence with the corresponding coverage for each base is displaxed in the output. SNPs and coverage for the genes of interest in each accession were read from this program and the haplotypes determined were displayed in Microsoft Excel 2007.

2.4 Morphological and physiological assays on "weedy lifestyle traits"

In the previous section, the sequencing and analysis of "weedy lifestyle genes" was described. An additional a morphological or physiological test was performed for each trait, to complete information on the specific characteristics. These assays are described in the following section.

2.4.1 RFLP analysis of *Rc*

We developed a "quick test", to investigate *Rc* genes for the wild type allele. For this assay a fragment of about 370 bp around the region of interest in exon 7 of *Rc* was amplified and digested with the BstUI restriction enzyme. The cutting site for BstUI (CGCG) is only present in the wild type allele and lost in all other forms due to a 14 bp deletion. Hence, the PCR product of a wild type (wt) allele will be cut and appear as two bands in an agarose gel whereas alleles with the 14 bp deletion do not contain a restriction site and will appear as single bands. The primers for the amplification of the region of interest were taken from Gross et al., (2010). Primer sequences are shown in table 5.2 (appendix, p.86). The PCR was carried out in 20 µl samples containing: 2 µl buffer (10x), 2 µl betaine (5M), 1 µl each primer (10 pmol), 0.8 µl DMSO, 0.4 µl dNTPs (10 mMol each), 0.1 µl Taq (New England Biolabs, Frankfurt a.M., Germany) and 2 µl template (50 ng/µl). The reactions were performed in a Primus 96 thermocycler by peqlab (Erlangen, Germany) with 95°C initial denaturation for 10 minutes, followed by 30 cycles of 1 min denaturation at 95°C, 1 min annealing at 56°C and 1 min elongation at 68°C. The reaction was terminated with a final elongation step of 68°C for 5 minutes and then cooled down to 4°C. The PCR products were purified using the PeqGOLD microspin cycle-pure kit (Peqlab, Erlangen, Germany) according to the manufacturer's instructions. The DNA was eluted in 20 µl sterile H_2O. 10 µl of the PCR product were mixed with 12.4 µl H_2O, 2.5 µl buffer no. 4 and 0.1 µl BstUI restriction (both New England Biolabs, Frankfurt a.M., Germany) and incubated for one hour at 60°C. The reaction was stopped by adding 5 µl of loading dye (H_2O, glycerol, bromophenol blue, xylene cyanol in equal parts). The samples were run on a 2% agarose gel (Roth, Karlsruhe, Germany) stained with 5 µl per 100 ml SYBRsafe (Life Technologies, Darmstadt, Germany) for 45 min at 100V.

2.4.2 HPLC analysis of seed extracts from Weedy Rice

To investigate the proanthocyanidin compounds contained in the pericarp of weedy rice seeds, methanolic extracts of weedy rice seeds were produced and analyzed using HPLC.

Seed extracts were prepared after Abdel-Aal & Hucl (2003) and Kim et al., (2008) with modifications. Seeds were dehulled and ground with an electric mill (Janke und Kunkel, Staufen, Germany). Ten ml of hexane were added to 10 g of powder and incubated for 3 h at 4°C while shaking. The hexane was removed and 100 ml methanol:HCl (85:15, v/v) were added, followed by another incubation step at 4°C for 4 h. The methanol:HCl step was repeated one time. The extracts were pooled, filtered through Whatman No. 3 paper and reduced to ~1 ml in a rotary evaporator (Büchi, Vacuubrand, Essen, Germany). The extracts were again filtered through a syringe filter (Chromafil; PET-20/15 MS, Macherey Nagel, Düren, Germany) and 50 µl of each extract were used for the HPLC analysis.

The HPLC analysis was carried out on an Agilent 1200 hplc (Agilent Technologies, Santa Clara CA, USA) with a Phenomenex Synergie Hydro RP column (4µM, 80 Å, 4.6x150 mm). The solvents acetonitrile (A), formic acid (B) and H_2O (C) were used. The program with the solvent gradient over time that was applied to our samples is shown in table 2.8.

The absorption spectrum at 280 was recorded to detect proanthocyanidins. The lab work for extracts and HPLC was performed by Kerstin Thellmann during her Bachelor Thesis (April – June 2013).

Table 2.8: Ratio of the solvents used in the HPLC run over time. A = acetonitrile, B = formic acid, C = H_2O

Time	A	B	C
0 min	5.8%	10%	84.2%
15 min	17.1%	10%	72.9%
30 min	26.5%	10%	63.5%
35 min	26.5%	10%	63.5%
41 min	5.8%	10%	84.2%

2.4.3 Vanillin Assay

The vanillin assay was performed to determine the amount of proanthocyanidins in the pericarp of Weedy Rice accessions. The assay was conducted as described in Price et al. (1980) with modifications.

One gram of dehulled, powdered rice seeds was extracted with 10 ml pure methanol for one hour, shaking at 100 rpm. The mixture was centrifuged for 10 min at 3,000 * g and the supernatant was used in the analysis.

For the standard curve, (+)-catechin solutions with the concentrations of 0, 0.06, 0.12, 0.18, 0.24 and 0.3 mg/ml in methanol were prepared.

The assay was conducted in a 30°C constant water bath to avoid temperature dependent variation in the data. Two sets of 1 ml of each standard and extract were prepared. One set was used as the blank values, the other for the vanillin reaction. For the blank values 5 ml 4% concentrated HCl in methanol were added to each sample, the other set was completed with 5 ml vanillin working solution (4% concentrated HCl and 0.5% vanillin in methanol) each. The vanillin reacts with the catechin residues in the proanthocyanidins, resulting in a red coloured complex that absorbs at 500 nm wavelength. All samples were incubated for exactly 20 min at 30°C and the absorption at 500 nm was measured. All values were normalized by subtracting the absorption of the blank value from the absorption of the vanillin sample.

The standard curve was determined by plotting the absorption at 500 nm against the catechin concentration. A linear regression was performed on the data and the equation was used to calculate the amount of catechin equivalents in the weedy rice seed extracts.

2.4.4 Microscopy of the abscission zone

Seed shattering is one of the characteristics of weedy rice. The process of abscission is initiated by the formation of an abscission layer between the grain and the pedicel. Programmed cell death in the abscission layer leads to a detachment of the seed from the pedicel. The abscission layer and the surrounding tissue are also called abscission zone (AZ). The surface of the AZ on the grain can be investigated by microscopy techniques and give information on the shattering habit of the investigated object.

Both, two and three dimensional images of the abscission zone from different rice accessions (weedy rice and cultigens) were taken to compare the texture of this area.

2.4.4.1 2D Microscopy

Pictures with 32 fold enlargement of the abscission zone of dehulled rice seeds were taken using the Leica M420 binocular microscope (Leica, Solms, Germany). Five seeds per accession were documented and the pictures were visually compared. The data were collected by Ramona Dries during a practical course in May 2013.

2.4.4.2 3D Microscopy

Rice seeds were cut close to the breaking point and glued to glass slides with clear nail polish. 3D images of five seeds per accession were taken with the Keyence VHX-600 digital microscope (Keyence, Neu-Isenburg, Germany). Between 20 and 35 optical sections of 5 µm thickness were taken for each seed and 3D images were calculated out of these data by the related software. The depth of the dimple formed at the abscission zone was measured with the software and the means of the five samples per accession were calculated.

2.4.5 Production of karrikin-containing smoke water and germination assay

Karrikins are a class of molecules recently discovered and shown to have a germination promoting activity in many plant species (Chiwocha et al., 2009, Nelson et al., 2009). They have been shown to be produced by the combustion of simple carbohydrates (Flematti et al., 2011). In this study the effect of karrikins on the dormancy of weedy rice seeds was examined, using smoke water (SW).

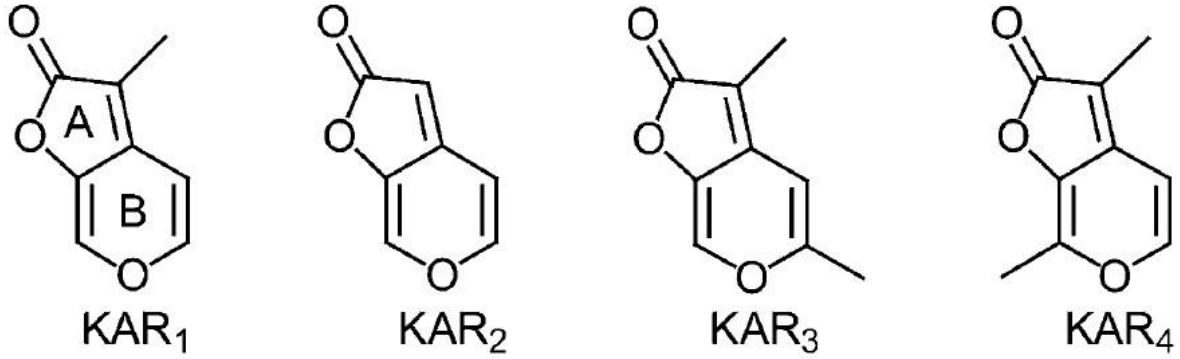

Figure 2.9: Representatives of the karrikin family. The germination stimulating activity is suggested to derive from the structural similarities with strigolactones. Figure by Nelson et al., (2009)

To produce karrikin containing smoke water, the method of Flematti et al., (2011) was followed with modifications. 2.4 g d-xylose (Merck, Darmstadt, Germany) were combusted for 15 min in a 250 ml round-bottom flask and the smoke was pumped through a wash bottle filled with 100 ml of Millipore H_2O. The resulting smoke water was diluted 1:10, 1:100 and 1:1000 for the use in germination experiments.

Weedy Rice accessions and cultigens of this study were tested for their germination behaviour in the presence of water and different concentrations of smoke water. The assay was performed in triplets for each accession. Twenty seeds (not dehulled) were sown in a glass Petri dish on tissue paper soaked with 5 ml water (control) or smoke water dilution (1:1000, 1:100 or 1:10). Dishes were sealed with parafilm (Roth, Karlsruhe, Germany) and kept in the dark at 20°C for three days. Germination rates in percent and the mean of all three approaches were calculated.

3 Results

The following chapter lists the results obtained by this study. The first part deals with the outcome of the SSR analysis conducted on 150 weedy rice accessions and 20 cultigens from Northern Italy. The results indicated a moderate genetic variability and differentiation among weedy rice samples. The comparison of genetic and geographic distance did not give any hints to an evolutionary origin. Several weedy rice groups could be distinguished by the markers. Some cultigens were identified to be genetically very closely related to weedy rice. The second part describes the results of the sequencing of genes linked to weedy rice characteristic traits and compares them with data from morphological and physiological assays. The results show a differentiation of Italian weedy rice from weedy rice samples of other geographic origins. The *japonica* identity of weedy rice could be shown. Furthermore wild rice specific alleles were discovered in the weedy rice gene pool. Studies on dormancy in weedy rice showed a trend of enhanced germination of weedy rice seeds, induced by karrikins.

A more detailed summary of the results can be found at the end of this section.

3.1 Characterization of Italian weedy rice and its relation to cultigens based on SSR studies

For the SSR analysis 150 weedy rice accessions from the 10 collection areas (Fig 2.1) and 20 cultivars (Table 2.2) were investigated using 19 SSR markers previously described by Cao et al. (2008). The statistical output of the analysis is displayed in table 3.1. In total 64 different alleles could be detected among the weedy rice samples with an average of 3.368 alleles per locus. The highest number of alleles per locus was detected in RM24 (7 alleles), three loci (RM84, RM211, RM289) were not polymorphic. The mean H_e was calculated with 0.295 which is relatively high; in comparison the H_o, with an average of 0.049, is considerably lower. Nine out of the nineteen loci tested were not heterozygous at all (H_o=0.000).

Table 3.1 Statistical output of the SSR analysis. SSR= name of the microsatellite, Chr= chromosome where the microsatellite is located, $N_{a=}$ number of different alleles found for a locus, H_e= expected heterozygosity, H_o= observed heterozygosity, bp range= lengths (no. of nucleotides) of the SSRs found in this study. For N_a, H_e and H_o the average was calculated.

SSR	Chr	N_a	H_e	H_o	bp range
RM 11	7	4	0.421	0.014	140-157
RM 14	1	7	0.300	0.333	192-200
RM 17	12	2	0.482	0.007	175-202
RM 19	12	2	0.014	0.000	231-262
RM 21	11	6	0.332	0.007	148-175
RM 44	8	4	0.360	0.269	124-128
RM 55	3	4	0.506	0.241	248-252
RM 84	1	1	0.000	0.000	126
RM 167	11	5	0.115	0.000	101-167
RM 180	7	2	0.059	0.007	123-126
RM 211	2	1	0.000	0.000	158
RM 212	1	2	0.343	0.000	131-133
RM 215	9	4	0.462	0.000	163-171
RM 219	9	4	0.640	0.020	212-220
RM 230	8	4	0.033	0.007	268-274
RM 253	6	4	0.595	0.000	148-153
RM 276	6	5	0.507	0.028	100-152
RM 280	4	2	0.425	0.000	186-188
RM 289	5	1	0.000	0.000	104
Average		3.368	0.295	0.049	

F-statistics were conducted to asses the differentiation of the weedy rice accessions. The pairwise F_{st} values vary between 0 and 0.685. This represents a range between no differentiation (0) and substantial differentiation (0.685). No clear trend to one direction could be observed in the dataset. The overall F_{st} was calculated as 0.179, which is in the range of 0.15 to 0.25 indicating substantial differentiation (Wright 1978).

A test for correlation between genetic and geographical distance was performed with a spatial structure analysis, as described in Shivrain et al. (2010). This type of analysis can indicate a centre of diversification pointing out origins of evolution for the study objects. The genetic distance based on the SSR values was plotted against the geographical distance calculated on the base of the GPS data of the collection points for each accession. The correlation coefficient R^2 was calculated. It is a measure for the degree of correlation between the compared dimensions. A value of 1 corresponds to 100% correlation whereas a value of 0 implies no correlation. The results are shown in figure 3.1 The value for R was calculated with $4*10^{-5}$for this dataset indicating that there is no correlation of genetic and geographic distance for the tested Italian weedy rice accessions.

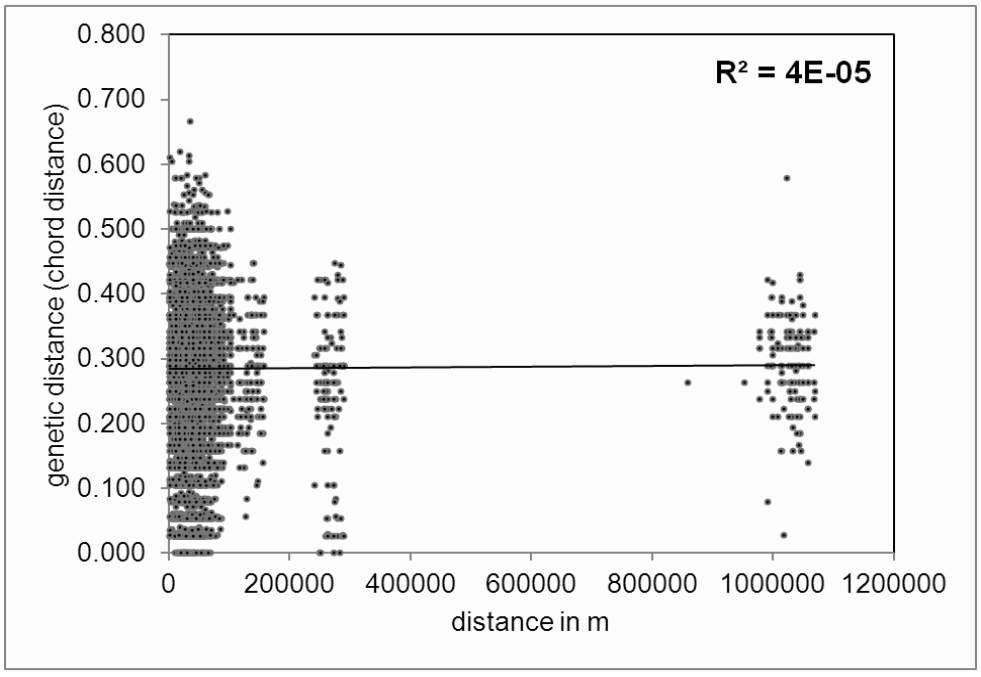

Figure 3.1 Plot for the spatial structure analysis. The genetic distance (chord distance) is plotted versus the geographic distance in m. Each point represents a single accession. The correlation coefficient was calculated and is displayed in the upper right corner of the diagram.

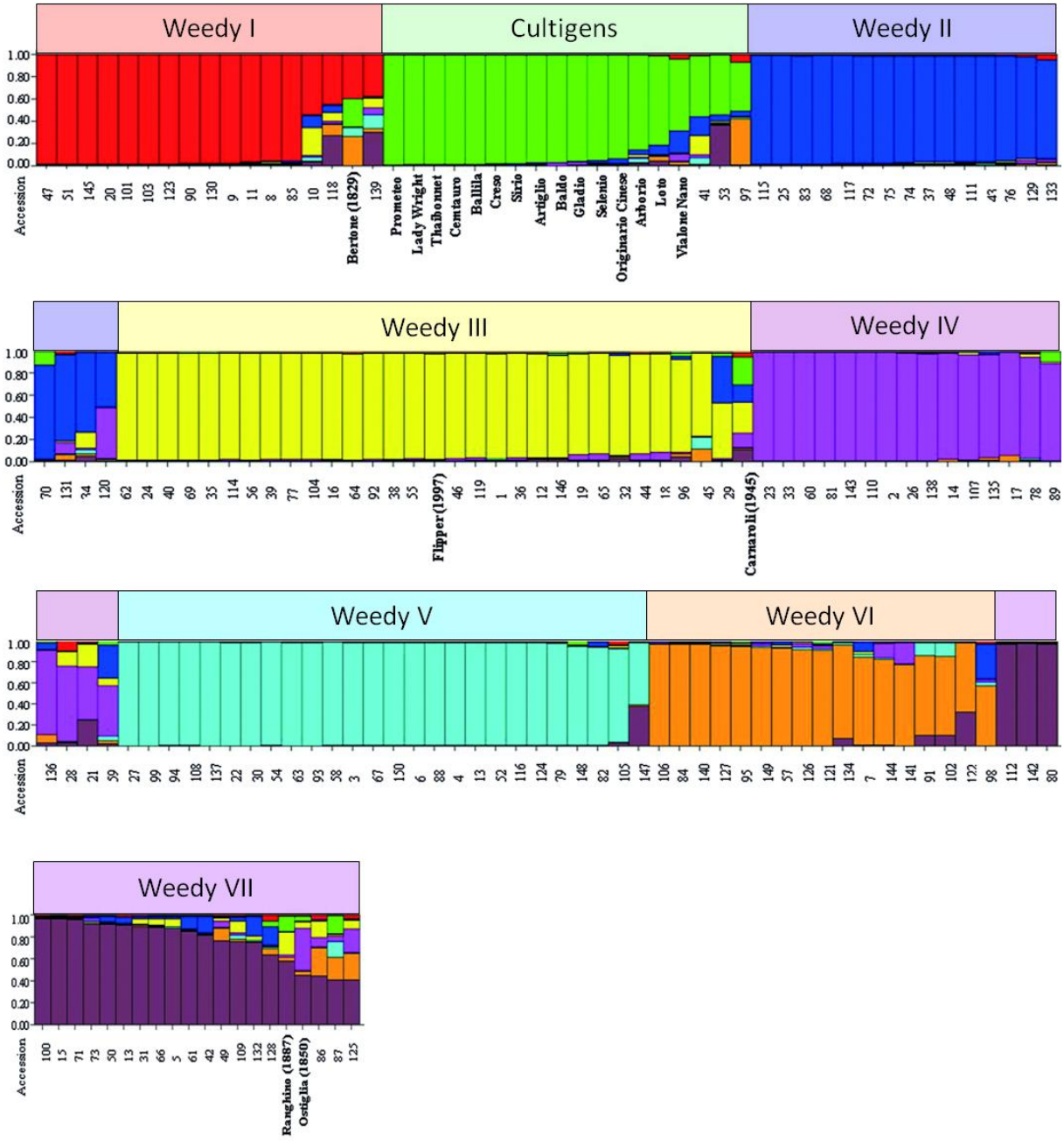

Figure 3.2 Bar plot of the population structure analysis obtained by Bayesian clustering. Each bar represents a single accession. The accession numbers and names in the case of the cultivars are given below each bar. The eight colours indicate the different populations. A full bar represents 100%, each coloured fragment of a bar shows the percentage of affiliation to the respective population. The population names are shown above the bars. For the cultivars that are interspersed in otherwise “weedy clusters” the release date is shown in brackets after the name.

A method of Bayesian clustering was used to infer the population structure of the weedy rice accessions in comparison with the cultivars. The calculations are based on the SSR length discovered by sequencing. A model that separated the accessions of this study in eight populations was best explained by the data (figure 3.2). The clusters representing the populations are shown each in a different colour. Each bar represents an individual accession and its affiliation to the different populations. An overall separation of weedy rice and cultigens can be observed in the population blot. Cultigens gather in the second population whereas the weedy rice accessions are distributed over the other seven clusters. However, the separation is not entirely strict. Exceptions of cultigens that are interspersed in other populations can be observed: Three weedy rice accessions (41, 53, 97) have their greatest affiliation with the cultigen cluster and are not grouped with other weedy rice accessions. Also cultigens scattered over weedy rice populations are recorded. Bertone is attached to the Weedy I population. In population weedy III the cultivars Flipper and Carnaroli can be found. And population Weedy VII includes Ostiglia and Ranghino. The dates for the first description are given in brackets (figure 3.2) for the cultigens associated with weedy rice. Flipper from 1997 and Carnaroli (1945) represent contemporary varieties. Flipper is still widely used in the collection area of our samples. Contrary to this, Bertone (1829), Ranghino (1887) and Ostiglia (1850) are old cultigens, dating back to the 19th century. These varieties are outdated and not cultivated anymore since several decades.

A comparison of the population structure with morphological data (Grimm et al. 2013) revealed a connection between populations and the awn state of weedy rice. (table 3.2). The separation of different awn types by the populations is not perfect but shows a clear trend for one awn state per population being the predominant type.

Table 3.2 Percentage of the different awn types found in the different populations of the Bayesian clustering.

Population	awnless	mucronate	straw awned	black awned
Weedy I	81,25%	6,25%	0%	12,5%
Weedy II	21%	5%	53%	21%
Weedy III	7%	7%	72%	14%
Weedy IV	79%	0%	5%	16%
Weedy V	4%	92%	0%	4%
Weedy VI	0%	0%	35%	65%
Weedy VII	5%	0%	14%	81%

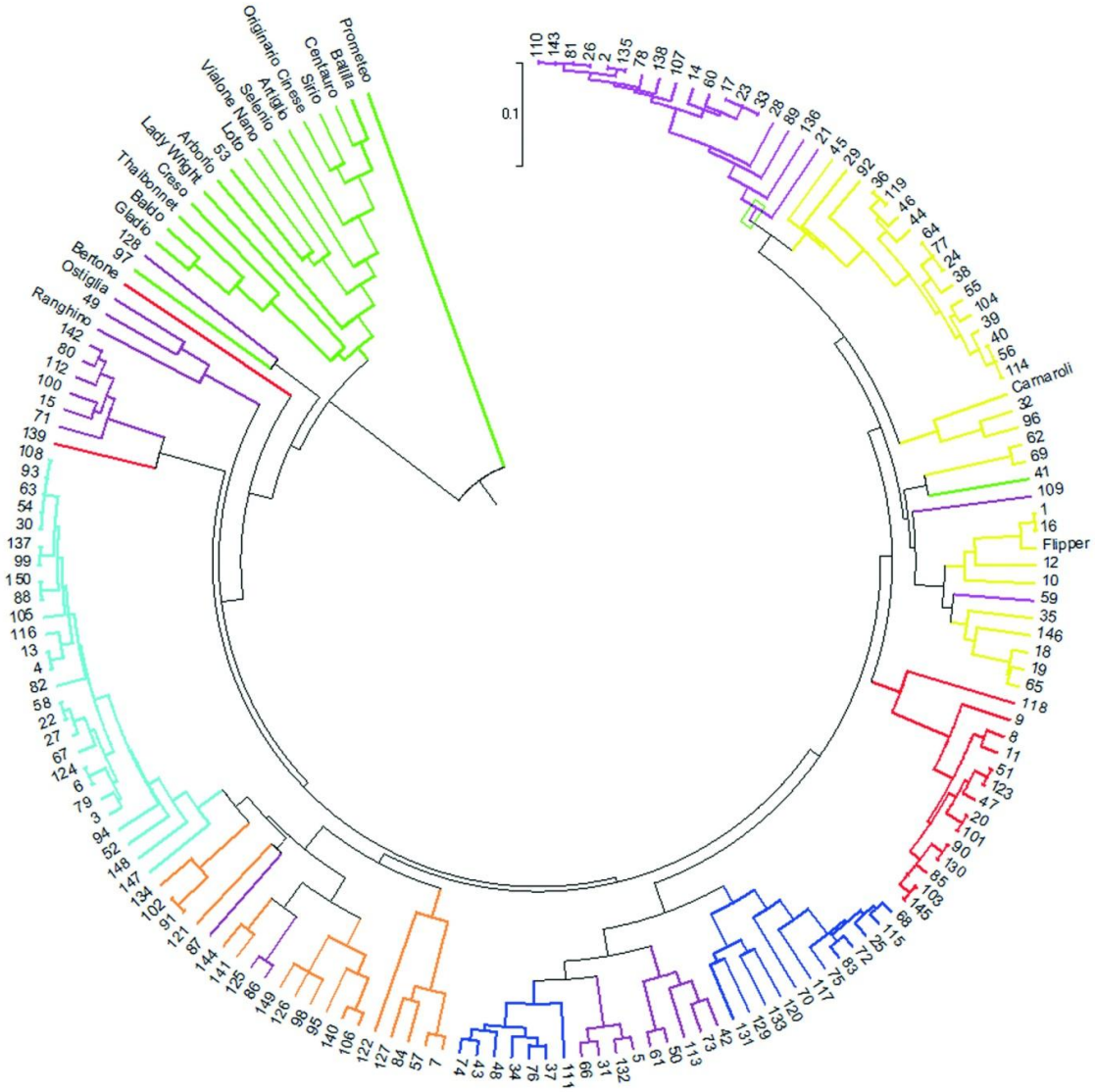

Figure 3.3 UPGMA based dendrogram of the 150 Weedy Rice accessions and 20 cultivars used in the SSR analysis. The cyclic representation was chosen for better visualisation. Each branch end represents one accession. The accession numbers of the weedy rice and the names of the cultivars are given for each branch. The branches are displayed in the colour code according to the Bayesian population the accession is affiliated with.

For a better and higher resolving view on the relations of the accessions (weedy rice and cultigens) in this study, a UPGMA dendrogram based on the genetic diversity was computed (figure 3.3). The affiliation to the respective Bayesian clusters is indicated by the same colour code for each accession. In general the clustering topology of the Bayesian analysis is maintained although occasional transitions can be observed. The cultigens being outliers in the structure analysis are also found interspersed the same weedy rice populations in the phylogenetic tree. The only exception is made by Bertone which now is closer to cluster VII than I. The cultivar Prometeo behaves as outgroup in the tree (figure 3.3) which is not the case in the Bayesian clustering (figure 3.2).

3.2 Characterization of "Weedy Lifestyle Traits"

Weedy rice seems to be closely related to cultivated rice but shows some significantly different traits linked to its lifestyle. These "weedy lifestyle traits" were investigated to better characterise weedy rice groups and collect information on their possible evolutionary origin. The traits that were chosen in this study are: Plant height, pericarp colouration, shattering and dormancy. The traits were on one hand investigated from the genetic point of view, by sequencing genes that are linked to them. On the other hand morphological and physiological approaches were conducted to complete the study. Weedy rice accessions for Italy were analysed in the groups obtained by the Bayesian clustering (3.1). For a better visualisation cultigens which clustered with weedy rice are displayed in the cultigen population in the following chapters. To further enlarge the sample spectrum and achieve a better comparability of the data, weedy rice and cultigens from each, Brazil and Thailand, and the wild rice *Oryza rufipogon* were included in the study from here on.

3.2.1 Plant Height

The enhanced growth is an important factor for weedy rice. In general weedy rice is considerably taller that cultivated rice outgrows it, and succeeds in the competition for light. In this study the aspect of plant height was studied by investigating the coding sequence regions of the SD1 gene (Os01t0883800-01), and comparing the results to the actual measured plant heights, obtained by our cooperation partners of the University of Turin (UNITO).

The SD1 gene is located on chromosome 1 in the region of about 40M. It is coding for a gibberellin oxidase (GA20OX2). The open reading frame (ORF) is composed of three exons of 557, 322 and 291 bases in length. Figure 3.4 shows the hapmap obtained form the next generation sequencing results. SNP identified in comparison with the *O. sativa ssp. japonica* sequence (bottom line) are visualised. Nucleotide characters separated by slashes represent heterozygosity for the corresponding site. 13 different haplotypes were identified. The haplotypes 1-10 are split in two major groups. The first group shows a sequence similar to the *O. sativa ssp. japonica* sequence used for comparison (1-5), and has most SNPs in the third exon. The second group (6-10) is characterized by a large deletion in exon 1 (381bp) and two common SNPs in exon 2 and 3. Actually, the deletion is missing in haplotype 6 but since it shares the other two SNPs it is incorporated in this group. *Oryza rufipogon* is found in a separate group but only differs from haplotype 6 by different SNPs in exon 1.

SD1																																			
	ex1					ex2			ex3									Population																		
Haplotype	262	299	318	387	456	7	58	97	70	77	79	93	98	101	107	110	140	C	I	II	III	IV	V	VI	VII	BC	BW	TC	TW	OR					
1	.	.	.	.	.	A	.	.	G/T	.	.	.	.	.	G/T	.	.	1							1										
2	.	.	.	.	.	.	.	.	G/T	.	.	G/C	G/A	.	G/T	G/T	.								1										
3	.	.	.	.	.	.	.	.	G/T	G/C	G/T	G/C	G/A	G/A	G/T	G/T	.						1												
4	.	.	.	.	.	.	.	.	.	.	.	.	G/A	.	G/T	.	.	4	3		2	3	1												
5	.	.	.	.	.	.	.	.	.	.	.	.	G/A	.	G/T	G/T	.	3		1			1												
6	.	G	.	.	.	.	.	T	.	.	.	.	G/A	.	G/T	.	G										1								
7	381bp deletion					.	.	T	.	.	.	.	.	.	.	.	G												1						
8	381bp deletion					.	.	T	.	.	.	.	G/A	.	.	.	G									1										
9	381bp deletion					.	.	T	.	.	.	.	G/A	.	G/T	.	G									1		1								
10	381bp deletion					.	.	T	.	.	.	.	G/A	.	G/T	G/T	G											1								
11	381bp deletion					.	.	T	G/T	.	.	G/C	G/A	.	G/T	G/T	G								1										
12	A	G	C	C	C	.	G	.	.	.	.	.	G/A	.	G/T	.	G													1					
O. sativa ssp. japonica	G	A	T	G	A	G	A	C	T	C	T	G	A	A	T	T	A	3			1	1			1										

Figure 3.4 Hapmap for the SD1 gene obtained from the NGS results. The nucleotide positions for the SNPs and the exons where the SNPs are located are displayed. Sites with no SNPs are indicated by dots. Each line shows one haplotype and the accessions representing it (Population). Italian accessions are grouped in C=cultigens and roman numbers = Weedy I-VII. Other accessions are summarized by: BC=Brazil cultigen, BW=Brazil weedy rice, TC=Thai cultigen, TW=Thai weedy rice and OR=*Oryza rufipogon*. For SNPs that were exclusively found to be heterozygous, both nucleotides are shown and separated by slash.

The first subgroup is exclusively represented by Italian weedy rice accessions and cultigens as well as the *japonica* database sequence. In the second subgroup the Thai and Brazilian weedy rice and cultivars are concentrated. No separation of weedy rice and cultigens in different haplotypes can be observed.

The deletion observed in the haplotypes 7-11 correspond to the original green revolution mutation of the "miracle rice" IR 8, the first semi-dwarf cultivar from the 1960s (Ashikari et al., 2002). The deletion results in a premature stop codon, that terminates the translation of the protein after amino acid 103 (Asano et al., 2007). In figure 3.5 the changes that are caused by the SNPs in the amino acid sequence of SD1 are shown.

Haplotype	aa no	88	100	106	129	152	188	205	218	317	319	320	324	326	326	329	330	340
1		.	.	-	.	.	V	.	.	A	.	.	.	.	.	G	.	.
2		.	.	.	.	.	.	.	.	A	.	.	P	G	.	G	G	.
3		.	.	.	.	.	.	.	.	A	G	V	P	G	G	G	G	.
4		.	.	.	.	.	.	.	.	.	.	.	.	G	.	G	.	.
5		.	.	.	.	.	.	.	.	.	.	.	.	G	.	G	G	.
6		.	G	.	.	.	.	.	Y	.	.	.	.	.	.	.	.	R
7		Stop codon after aa 103																
8		Stop codon after aa 103																
9		Stop codon after aa 103																
10		Stop codon after aa 103																
11		Stop codon after aa 103																
12		T	G	A	A	P	.	E	.	.	.	.	.	G	.	G	.	R
O. sativa ssp. japonica		A	E	A	A	P	V	E	Y	S	A	F	P	E	E	V	V	Q

Figure 3.5 The amino acids in the SD1 protein sequence, affected by SNPs from the hapmap, are shown for each haplotype. Sites where no SNP is found in the hapmap are indicated by dots.

The SNP at the site 140 in exon 3, which is characteristic for the second group, is one of two transitions that occurred during the differentiation process of *indica* and *japonica* varieties (Asano et al., 2011). The second one is the SNP at sit 299 in exon one. It can be observed in haplotype 6 and 11 (*O. rufipogon*). In the other accessions of the group this SNP is lost due to the deletion. Both SNPs cause amino acid transitions (E/G and R/Q) in the protein sequence (figure 3.5).

Most SNPs identified in this survey are located in exon 3, which is the substrate binding region of the gibberellin oxidase (Ashikari et al., 2002). These SNPs mainly cause the substitution of an amino acid with a glycine (figure 3.5). A striking observation is, that all SNPs, except for the evolutionary one at site 140, were found to be heterozygous (figure 3.4) in the analysed accessions. The primers were checked for alternative binding sites in the rice genome which might create false heterozygous results by binding to other copies of the gene

on the same chromosome or similar regions. For the primers used in this study only one binding each was found in the *O. sativa ssp. japonica* genome.

To be able to compare the data of the SD1 gene to actual plant height data, the average height (cm) of Italian weedy rice plants and cultigens used in this study was calculated (figure 3.6).

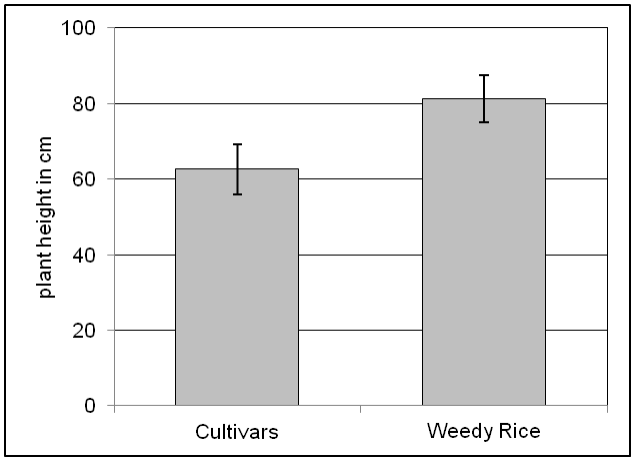

Figure 3.6 Average plant height in cm for Italian cultigens and weedy rice. Data obtained by Fogliatto et al., (2012). The standard deviation is indicated by error bars.

For the Thai and Brazilian accessions no exact measurements of the plant height are available but in green house experiments the weedy accessions grew taller than the cultigens. The plant height for the Italian weedy rice accessions and some cultigens were obtained by Fogliatto et al., (2012) and used for comparison in this study. Figure 3.6 shows the summary of the data on average plant height for Italian weedy rice and cultigens. The comparison shows that in average the weedy rice accessions a 20 cm more (81.21cm) in height compared to the cultigens (62.61cm). Since weedy rice and cultigens share the same haplotypes, but are clearly of different height, no direct link of the allele type to the plant height can be observed.

3.2.2 Pericarp Coloration

The red pericarp is the main characteristic of weedy rice and the reason why it is also called red rice. The genetic background of the pericarp pigmentation and the quality and quantity of the pigments in weedy rice were investigated in the following section.

The *Rc* gene (Os07t0211500-00) is located on chromosome 7 and encodes a bHLH protein which is involved in proanthocyanidin biosynthesis (Sweeney et al., 2006). It was shown to be responsible for the red grain pigmentation in wild and weedy rice (Furukawa et al., 2007, Gross et al., 2010). In this part of the study a section of exon 7 of *Rc* was sequenced. In this region the site of the mutation that caused the change from red pigmentation to no pigmentation during domestication is located. Furthermore the pigments found in weedy rice pericarps were qualitatively (HPLC) and quantitatively (vanillin assay) investigated.

In Figure (3.7) the results of the sequencing of *Rc* exon 7 are summarised. Two different groups could be identified additional to the *O. sativa ssp. japonica* sequence. The *japonica*-type allele shows a 14 bp deletion. Haplotype one has an additional 1 bp deletion plus the 14 bp cultivar-mutation (figure 3.7 and 3.8). Group 2 represents the "wild rice allele", found in *Oryza rufipogon* and other wild rice species, which is coding for the intact protein and results in the expression of a red pericarp.

rc															
	ex7		Population												
Haplotype	314	360/361	C	I	II	III	IV	V	VI	VII	BC	BW	TC	TW	OR
1	---	.		4		4	4	3		1					
2	.	ACGCGAAAAGTCGG			2	2			1	2		1		1	1
O. sativa ssp. japonica	G	---	11								1		1		

Figure 3.7 Hapmap for the Rc gene obtained from the NGS results. The nucleotide positions for the SNPs and the exons where the SNPs are located are displayed. Sites with no SNPs are indicated by dots. Each line shows one haplotype and the accessions representing it. Italian accessions are grouped in C=cultigens and roman numbers = Weedy I-VII. Other accessions are summarized by: BC=Brazil cultigen, BW=Brazil weedy rice, TC=Thai cultigen, TW=Thai weedy rice and OR=*Oryza rufipogon*.

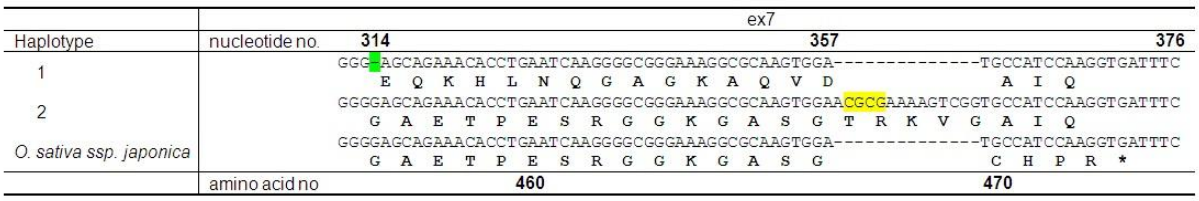

Figure 3.8 Alignment for the haplotypes detected by sequencing. The nucleotide alignments are shown in the upper line, the translated amino acid alignments below. Stop codons are indicated by *. The additional 1 bp deletion for haplotype one is highlighted in green. The target sequence for BstUI for the restriction assay is highlighted in yellow.

The consequences of the different haplotype groups on the protein sequence can be studied in the alignment of the alleles (figure 3.8). The 14 bp deletion in the japonica group results in a premature stop codon leading to the white pericarp in cultivated rice. All cultigens investigated in this study were found to have this allele type. No weedy rice accession with the *O. sativa ssp. japonica* sequence was detected. The weedy rice accessions spread over the other two groups, haplotype one and two. Haplotype one shows the 14 bp deletion typical for the cultigens but also an additional 1 bp deletion 43 base pairs upstream (figure 3.8 green box). As it can be seen from the translated amino acid sequence in this cultigen-like allele, the additional deletion restores the reading frame. The protein is shortened by five amino acids and the amino acid residues are substituted in the section between the two deletions. The reversion of the stop codon in this allele ensures a fully translated protein sequence. This allele has been reported in the literature for changing an Italian white pericarp cultivar in a type with red pericarp (Gulick et al., 2009). The haplotype 2 represents the wild type *Rc* gene as known from *O. rufipogon*. This version of the gene is the original sequence for the enzyme and ensures a correct translation for the protein. Italian accessions are split over the haplotypes one and two, whereas Thai and Brazilian weedy rice accessions are only found to have the wild type allele.

In figure 3.9, dehulled seeds of all three allele types are shown. The cultigen allele expresses the common white seed type. Both, the wild type allele and the cultigen-like allele show a red seed pericarp. In the photograph no differences in the coloration between the two alleles can be observed. Therefore, the pigments were investigated for their component profile (HPLC) and the quantitative content in the grains (vanillin assay). The results are described in the following parts.

Figure 3.7 Dehulled seeds of rice accessions representing the different allele types of Rc.

To invent a "quick test" for the *Rc* allele type, an assay based on a restriction digest was developed. The amplified region of exon 7 was digested with the restriction enzyme BstUI. The cutting site (figure 3.8 yellow box) is only present in wt alleles – in the other allele types the recognition site is removed by the 14 bp deletion. In case of cultigens or cultigen-like alleles the fragment is unaffected by restriction with BstUI. Amplicons of wt alleles are cut in two parts of 220 and 152 bp respectively. An example of the results is shown in figure 3.10. Six accessions were tested of which four (TO94, TO52, TO1, TO133) were not cut and are therefore haplotype group 1 or cultigens, whereas two (TO87, TO126) were digested which identifies them as wt alleles. Of course, this assay cannot distinguish cultigen or haplotype 1 alleles but it can definitely identify the wt allele.

The ratio of the haplotypes (1&2) for the Italian weedy rice accessions was calculated based on the results of the sequencing. The results are represented in figure 3.11. It shows that 75% of the Italian weedy rice accessions share the cultigen-like allele and 25% the wt allele.

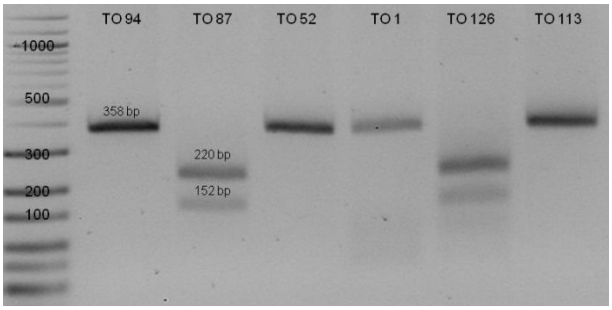

Figure 3.8 2% Agarose gel stained with SYBRsafe. It shows the results of the restriction digest of rc amplicons with BstUI. The fragment size of the marker is shown on the left. The fragment size of the bands is shown in examples for TO94 and TO87.

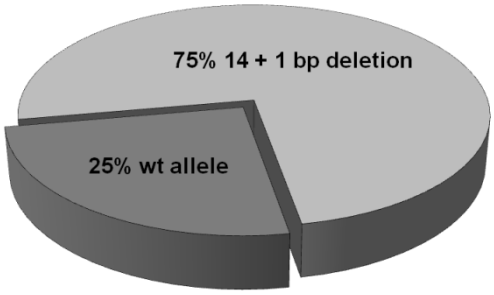

Figure 3.11 Ratio of the two allele types of Rc found in investigated weedy rice accessions.

To investigate not only the genetic aspect of pericarp coloration but also the physiological expression of the trait, extracts of weedy rice seeds were processed with HPLC to assess the pattern of proanthocyanidins produces. A summary of the results for the HPLC analysis is shown in figure 3.12. The spectra at 280 nm were recorded since this is the absorption maximum of proanthocyanidins (Lazarus 1999). Figure one C shows an example of an absorption spectrum for a peak of the pattern. The overall pattern for the proanthocyanidins computed by HPLC analysis was the same for all accessions tested in this assay. The first high and very broad peak represents the so called injection peak. At that retention time, remnants of the extraction are flushed through the column causing a peak. It is not relevant for the further analysis. The further pattern is mainly characterised by a group of three large peaks, the main components, followed by a number of considerably smaller peaks. The only difference that can be observed between the HPLC patterns of the accessions was the ratios of the main components in the extracts. They split the accessions in two subgroups: A and B (figure 3.12). In subgroup A the main component sticking out is represented by the first peak, in subgroup B by the second.

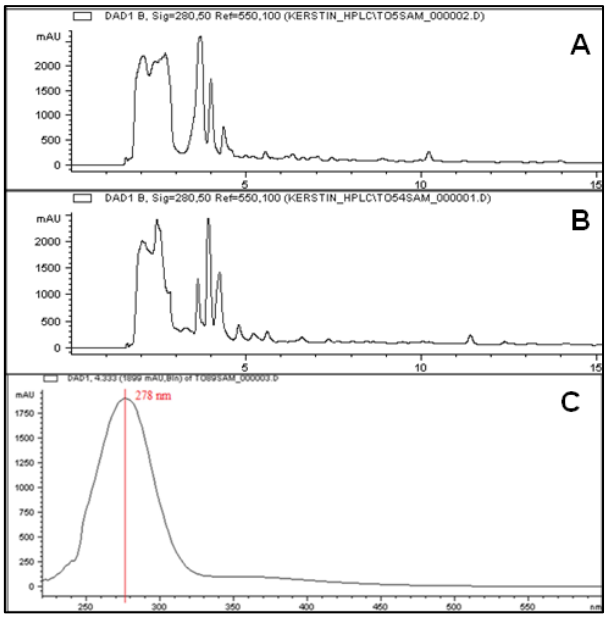

Figure 3.9: Example spectra for the HPLC analysis of weedy rice seed extracts. Retention times are shown on the x-axis, intensity in mAu on the y-axis. **A** = group 1, main component is represented by the first peak, **B** = group 2, main component is represented by the second peak. **C** = full absorption spectrum of a single peak. The intensity is plotted against the wavelength.

In addition the content of proanthocyanidins in the samples was determined using the vanillin assay (Butler et al., 1982). The measured values ranged from 14.693 to 38.81 µmol per gram grain (table 3.3), displaying a moderate level of variation within the sampling. The results of the quantitative analyses of the data from this study were compared with the allele type and population and are summarised in table 3.3. It can be drawn from the table that no allele type seems to be favoured by a certain population. Further, the allele type is not linked to the HPLC group. Both, A and B types can be observed in both alleles. A variation of the content of catechin equivalents in the samples is shown by the results (table 3.3) but the value does not correlate with a certain allele type or HPLC group.

Table 3.3 Population, Rc-allele type, HPLC group and proanthocyanidin group for each accession used in the study

Accession	Population	Allele type	HPLC group	Conc. proanthocyanidins µmol g^{-1}
TO41	Cultigens	cultigen-like	A	30.202
TO20	I	cultigen-like	B	28.496
TO101		cultigen-like	B	21.891
TO133	II	wild type	A	22.336
TO32	III	cultigen-like	B	29.460
TO64		cultigen-like	B	29.460
TO26	IV	cultigen-like	B	30.944
TO60		cultigen-like	B	25.527
TO81		cultigen-like	A	19.739
TO89		cultigen-like	A	33.616
TO135		cultigen-like	B	23.746
TO13	V	cultigen-like	A	38.810
TO54		cultigen-like	B	18.181
TO79		cultigen-like	B	28.644
TO108		cultigen-like	B	20.926
TO57	VI	cultigen-like	B	14.693
TO106		wild type	B	19.368
TO126		wild type	B	26.121
TO5	VII	cultigen-like	A	36.658
TO73		wild type	B	22.781

3.2.3 Shattering

Shattering is the natural way of seed dispersal in grasses. It was reduced during domestication in cultivated rice. Weedy rice however has a shattering phenotype. Different genes (qSH1, SHAT1 and sh4) involved in the formation of the abscission layer were investigated in weedy rice and cultigen accessions. Additionally the part of the abscission zone (AZ), that is still visible at the base of the grain after detachment, was investigated using different microscopy techniques.

The *qSH1* gene (Os01t0848400-0)1 is located on chromosome 1 and encodes a homeodomain transcription factor. The results of the sequencing are shown in figure 3.13. In total six SNPs were found, three in exon 1 and three in exon 3. Four haplotypes were detected within the sampling. Haplotype 1 and 2 show SNPs only in exon 3, group three has SNPs in exon 1 and 3. The SNPs in exon 3 at positions 202 and 210 are present in all groups except for the *japonica* haplotype.

qSH1																			
	ex1			ex3			Population												
Haplotype	6	122	192	202	210	379	C	I	II	III	IV	V	VI	VII	BC	BW	TC	TW	OR
1	.	.	.	C	C	.									3		2	2	
2	.	.	.	C	C	A										2			
3	A	G	C	C	C	.													1
O. sativa ssp. japonica	T	C	T	T	T	G	10	4	2	6	1	2	1	4					

Figure 3.13 Hapmap for the qSH1gene obtained from the NGS results. The nucleotide positions for the SNPs and the exons where the SNPs are located are displayed. Sites with no SNPs are indicated by dots. Each line shows one haplotype and the accessions representing it. Italian accessions are grouped in C=cultigens and roman numbers = Weedy I-VII. Other accessions are summarized by: BC=Brazil cultigen, BW=Brazil weedy rice, TC=Thai cultigen, TW=Thai weedy rice and OR=*Oryza rufipogon*.

Haplotype.	aa no	2	41	64	68	169	226
1		.	.	.	L	T	.
2		.	.	.	L	T	S
3		I	G	A	L	T	.
O. sativa ssp. japonica		I	A	A	L	T	G

Figure 3.14 The amino acids in the qSH1 protein sequence, affected by SNPs from the hapmap, are shown for each haplotype. Sites where no SNP is found in the hapmap are indicated by dots.

Figure 3.14 shows the amino acid sequence for the SNPs. The first 2 SNPs in exon three (positions 202 and 210) do not cause changes in the amino acid sequence so the protein sequence of qSH1 in these accessions is the same as in the *japonica* haplotype. In haplotype 2 the amino acid 226 is changed from G to S and in haplotype 3 a substitution from A to G at

position 41 can be found. In total only two of the six SNPs alter the protein sequence. The accessions from Italy (cultigens and weedy rice) and the accessions from Thailand and Brazil are strictly separated by the haplotype groups. All Thai and Brazil accessions are distributed over the haplotypes 1 and 2. All Italian accessions share the *O. sativa ssp. japonica* haplotype. The *O. rufipogon* is haplotype (3) is separated from all other accessions (figure 3.13).

SHAT1 (Os04t0649100-02) is a gene involved in the shattering process in rice which has just recently been identified (Zhou et al., 2012). It is located on chromosome 4 and is composed of 4 exons. The results of the sequencing are shown in figure 3.15. Only two SNPs, both located in exon1 (sites 98 and 104) were discovered in the analysis. They built one haplotype (1), the other haplotype represents the *japonica* sequence. The SNPs lead to an altered amino acid sequence at two sites (figure 3.16). The amino acids 33 and 35 are changed from G to D and S to Y respectively by compared to the *japonica* sequence. All Italian weedy rice accessions and most of the Italian cultivars share the *japonica* type sequence. The haplotype 1 gathers all the Brazilian and Thai accessions – cultigen and weedy – and two Italian cultigens, of which one is the Clearfield® variety. As in *qSH1*, in *SHAT1* again a separation of the Italian sample group and the samples from other continents can be observed. In contrast to *qSH1*, *O. rufipogon* does not form a separate group in *SHAT1* but clusters, together with the Italian accessions in the *japonica* haplotype.

SHAT1															
	Exon1		Population												
Haplotype	98	104	C	I	II	III	IV	V	VI	VII	BC	BW	TC	TW	OR
1	A	A	2								1	1	1	1	
O. sativa ssp. japonica	G	C	13	2	1	6	3	2		2					1

Figure 3.15 Hapmap for the SHAT1 gene obtained from the NGS results. The nucleotide positions for the SNPs and the exons where the SNPs are located are displayed. Sites with no SNPs are indicated by dots. Each line shows one haplotype and the accessions representing it. Italian accessions are grouped in C=cultigens and roman numbers = Weedy I-VII. Other accessions are summarized by: BC=Brazil cultigen, BW=Brazil weedy rice, TC=Thai cultigen, TW=Thai weedy rice and OR=*Oryza rufipogon*.

Haplotype	aa no	33	35
1		D	Y
O. sativa ssp. japonica		G	S

Figure 3.16 The amino acids in the SHAT protein sequence, affected by SNPs from the hapmap, are shown for each haplotype. Sites where no SNP is found in the hapmap are indicated by dots.

The *sh4* gene (Os04t0670900-00) also referred to as *SHA1* in the past, is located at chromosome 4. It is composed of two exons and encodes a MYB-like transcription factor (Li et al., 2006). A SNP in exon one (position 237) of *sh4* is suggested to reduce shattering in cultigens (Li et al., 2006).

sh4																	
	ex1				Population												
Haplotype	237	539	611	620	C	I	II	III	IV	V	VI	VII	BC	BW	TC	TW	OR
1	G	C	.	.												1	
2	G	.	T	.		4	1	5	2	4		3		1			
3	G	.	.	C										1			
O. sativa ssp. japonica	G	G	G	G	13		1	1			1	1	1	1	1		1

Figure 3.17: Hapmap for the sh4 gene obtained from the NGS results. The nucleotide positions for the SNPs and the exons where the SNPs are located are displayed. Sites with no SNPs are indicated by dots. Each line shows one haplotype and the accessions representing it. Italian accessions are grouped in C=cultigens and roman numbers = Weedy I-VII. Other accessions are summarized by: BC=Brazil cultigen, BW=Brazil weedy rice, TC=Thai cultigen, TW=Thai weedy rice and OR=*Oryza rufipogon*.

Haplotype	aa no	180	204	207
1		L	.	.
2		.	L	.
3		.	.	P
O. sativa ssp. japonica		R	R	R

Figure 3.18 The amino acids in the sh4 protein sequence, affected by SNPs from the hapmap, are shown for each haplotype. Sites where no SNP is found in the hapmap are indicated by dots.

Four haplotypes were identified, including the *japonica* type (figure 3.17). In total three SNPs were found in the groups 1 to 3. In the past, a transition from T to G in exon one of sh4 (position 237) has been associated with extremely reduced shattering phenotypes. This non-shattering site is shown in the hapmap (figure 3.17). All accessions of this study have the guanine associated with reduced shattering. However, shattering could be observed of all weedy rice accessions in this study. All SNPs that were found in this study are also located in exon one. Each haplotype is characterized by a single SNP. The *japonica* haplotype is shared by all cultigens of this study (Brazilian, Thai and Italian), four Italian, one Brazilian weedy rice accession and *O. rufipogon*. Amino acid substitutions caused by the SNPs are shown in figure 3.18. Each SNP of the haplotypes 1-3 replaces an arginine residue with another amino acid. In the haplotypes 1 and 2, leucine, in haplotype 3 proline are translated for arginine in the *japonica* sequence.

2D and 3D imaging was performed to visualize the abscission zone (AZ) of detached seeds in weedy and cultivated rice.

The 2D imaging revealed different morphologies depending on the shattering ability of the accession. The accessions were categorized as rough, intermediate and smooth surface types representing low, moderate and easy shattering ability, respectively. Figure 3.19 shows example pictures for each of the AZ types are shown. For each accession five seeds were evaluated to determine the AZ type.

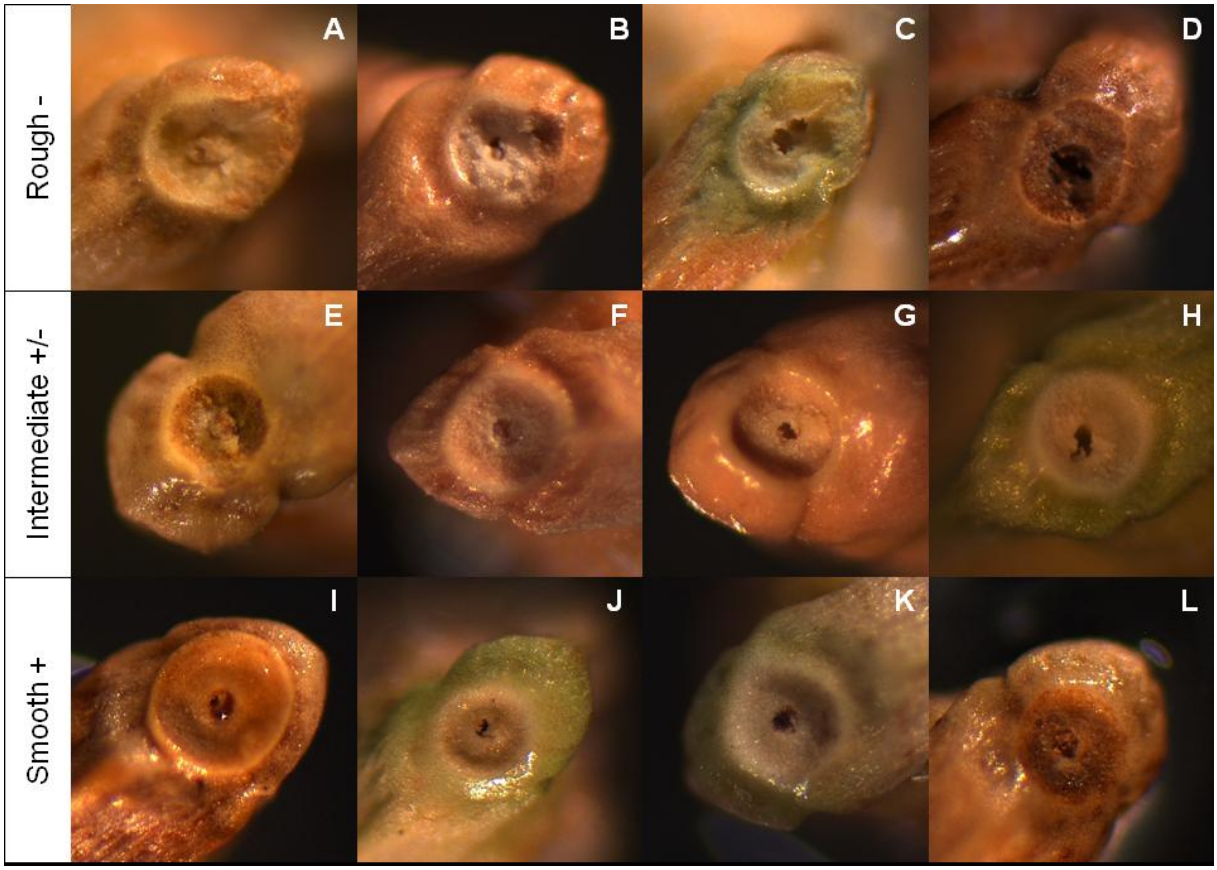

Figure 3.19 Images of the AZ of rice accessions with 32x magnification. Examples for each category are shown. A-D = rough (-) AZ, E-H = intermediate (+/-) AZ, I-L = smooth (+) AZ.

Table 3.4: Classification of the AZ-type of rice accessions based on 2D microscopy of the AZ. - = smooth,+/- = intermediate, + = rough. The left column shows the cultivated rice accessions, the left column represents weedy rice and wild rice.

Non-shattering	AZ phenotype	Shattering	AZ phenotype
O. sativa ssp. japonica (Nipponbare)	-	*Oryza barthii*	+
Arborio	+/-	TO11	+/-
Artiglio	+/-	TO13	+
Baldo	-	TO19	+
Ballila	+/-	TO20	+
Bertone	+	TO24	+
Carnaroli	+/-	TO26	+
Creso	+/-	TO30	+
Flipper	-	TO32	+
Gladio	+/-	TO64	+
Lady Wright	-	TO81	+
Originario Cinese	-	TO90	+
Ostiglia	+	TO101	+
Prometeo	-	TO106	+
Ranghino	+/-	TO108	+
Selenio	-	TO133	+
Thaibonnet	+/-		
Vialone Nano	-		

The results of the 2D imaging are shown in the table above. In the left column the results for the cultigen accessions are shown. Only one of the cultigen accessions (Ostiglia) showed a smooth AZ type. The other cultigens were equally split in rough and intermediate phenotypes. The right column shows the results for wild rice (O. barthii) and 15 Italian weedy rice accessions. Except for one (TO11), all showed a smooth (+) phenotype for the abscission zone. The candidate cultigens that were identified as possible ancestors of Italian weedy rice in the SSR analysis represent all three phenotypes: smooth (Ostiglia), intermediate (Bertone, Carnaroli, Ranghino) and rough (Flipper).

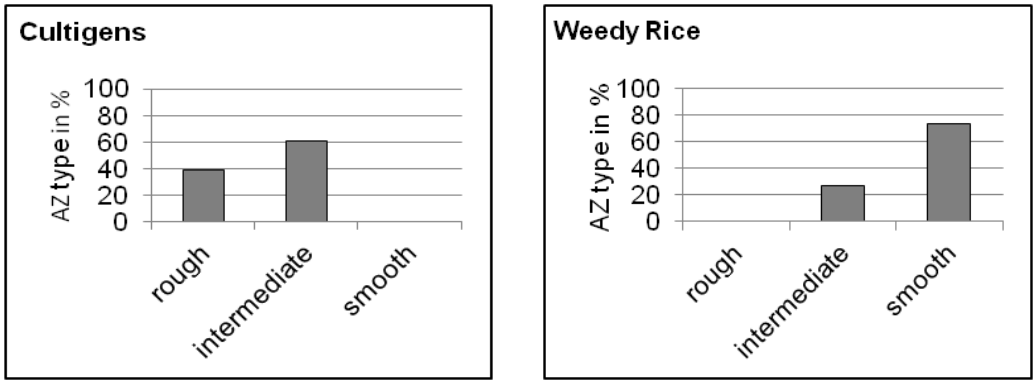

Figure 3.20 Percentage of AZ type for cultigens (left side) and weedy rice (right side). The AZ type (rough, smooth or intermediate) is shown on the x-axis, percentage on the y-axis.

The percentage of the AZ type for cultigens and weedy rice are shown in figure 3.20. The cultigens show values of 38.98% rough and 61.11% intermediate types indicating a low to moderate shattering ability. The weedy rice samples are distributed over 26.67% intermediate and 73.33% smooth types suggesting an easy shattering phenotype.

The 3D imaging was used to measure the dimple that is formed at the grain by detachment from the pedicel. The depth might depend on the force needed to detach the seed and therefore the level of shattering. Example pictures of the 3D microscopy are shown in figure 3.22. A summary of the results is plotted in figure 3.21. The measured values ranged from 37.902 µm (TO19) to 110.220 µm (Artiglio) showing considerable variation between accessions. The standard deviation also indicates high variability within the accessions. The average values were calculated with 84.587 µm for the cultigens and 76.685 µm for the weedy rice accessions, indicating a trend of greater depth for cultigens.

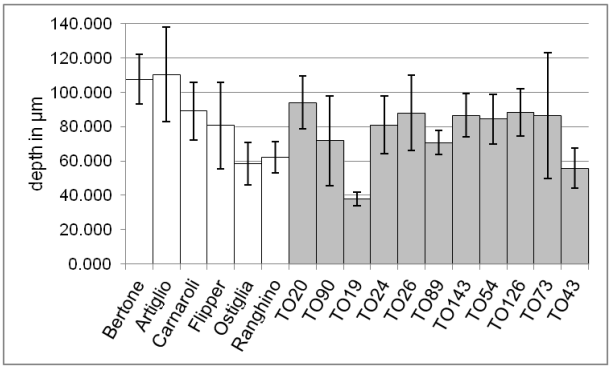

Figure 3.21 Average results for the dimple depth of the accessions. Cultigens are represented by white bars, weedy rice accessions by grey bars. Measures are shown in µm, the standard deviation is indicated by the error bars.

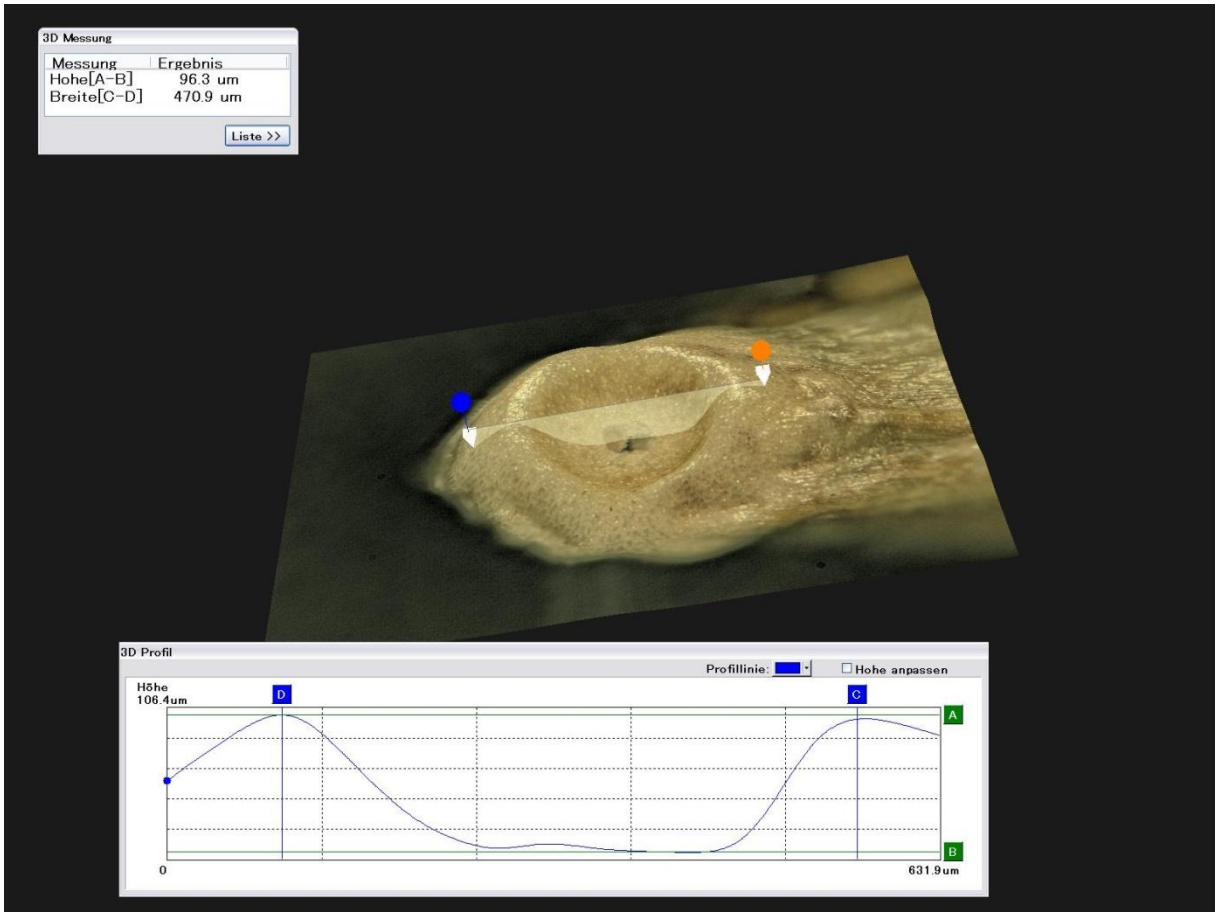

Figure 3.22 Example picture for the 3D imaging of AZ of rice grains. The pictures were taken at 200x magnification. The graph below the 3D depiction shows the profile of region between the blue and red mark. The depth was measured from the highest (A) to the lowest point (B) of the profile. Results for the measurements are shown in the box in the upper left corner.

3.2.4 Dormancy

Dormancy is the trait that makes the treatment of weedy rice infested soils extremely problematic. The induction of dormancy in seeds in a complex process with many factors involved. Two of these factors in rice are the genes VP1 (viviparous 1) and SDR4 (seed dormancy 4). They were included in this study to investigate dormancy in weedy rice on a molecular level. The physiological aspect – the germination behaviour of weedy rice – and the possibility of dormancy breaking for weedy rice treatment, were tested in a germination assay setup using karrikins in smoke water (SW).

VP1 (Os01t0911700-01) is assembled of seven exons, located on chromosome 1 and encodes a transcription factor. It positively regulates the expression of SDR4. The results of the sequencing of the VP1 exons are shown in figure 3.14. In total 12 SNPs were detected splitting in five haplotypes (figure 3.23). The SNPs are shattered over the exons 2, 3, 4 and 7. For the Thai and Brazilian accessions no usable data could be obtained from the sequencing. The Italian weedy rice accessions and cultigens are distributed over the haplotypes 1-3 and 5. The haplotype 4 is exclusively found in *O. rufipogon*. No separation of weedy rice and cultigens can be observed. Except for three (exon 2 position 110, exon 2 position 363 and exon 7 position 238) all SNPs are silent and do not cause amino acid substitutions in the protein sequence (figure 3.24). They are not suspected to have a crucial influence on the protein function. Unexpectedly, none of the Italian cultigens showed the *japonica* haplotype as found in the database, but many weedy rice accessions did.

VP1																										
	ex2							ex3	ex4		ex7		Population													
Haplotype	110	376	353	457	715	823	1102	63	6	60	27	238	C	I	II	III	IV	V	VI	VII	BC	BW	TC	TW	OR
1	G		.	.	.	.	.	.	.	.	.	.	7			2	3								
2	.	C	.	C	.	.	.	.	.	.	.	.						1							
3	.	C	.	C	C	.	.	.	.	G	T	C	3												
4	.	.	A	C	C	T	T	C	G	G	.	.													1
O. sativa ssp. japonica	A	T	G	T	A	C	C	G	A	A	C	T		4		4		3		2					

Figure 3.23: Hapmap for the VP1 gene obtained from the NGS results. The nucleotide positions for the SNPs and the exons where the SNPs are located are displayed. Sites with no SNPs are indicated by dots. Each line shows one haplotype and the accessions representing it. Italian accessions are grouped in C=cultigens and roman numbers = Weedy I-VII. Other accessions are summarized by: BC=Brazil cultigen, BW=Brazil weedy rice, TC=Thai cultigen, TW=Thai weedy rice and OR=*Oryza rufipogon*.

Haplotype	aa no	37	134	127	161	247	283	376	524	535	553	616	687
1		A	.	.	.	.	.	.	.	.	.	.	.
2		.	D	.	D	.	.	.	.	.	.	.	.
3		.	D	.	D	A	.	.	.	.	E	R	P
4		.	.	N	D	A	D	L	G	A	E	.	.
O. sativa ssp. japonica		T	D	D	D	A	D	L	G	A	E	R	S

Figure 3.24 The amino acids in the VP1 protein sequence, affected by SNPs from the hapmap, are shown for each haplotype. Sites where no SNP is found in the hapmap are indicated by dots.

The dormancy gene SDR4 (Os07t0585700-01) is located at chromosome 7 of the rice genome. Its expression is positively regulated by VP1 during the process of dormancy formation in the seeds (Sugimoto et al., 2010). Five haplotypes were found among the samples for the SDR4 coding region (figure 3.25). Two of them, the *O. sativa ssp. japonica* haplotype and haplotype 2, correspond to the alleles sdr4-n and sdr4-k described by Sugimoto et al., (2010), respectively. Sdr4-n represents an allele typical for *japonica* varieties, whereas sdr4-k is only found in *indica* cultivars. Haplotype 1 only differs by one SNP from sdr4-k', another allele described in the publication of Sugimoto et al. *O. rufipogon* shows a separate haplotype, similar to sdr4-k. One cultigen from Italy forms a separate haplotype one base different from sdr4-n. Most of the SNPs do not cause changes in the amino acid sequence. The substitutions that could be detected are shown in figure 3.26.

	SDR4																	Population												
Haplotype	88/89	271	284	295	353	507-509	514	516	517	519	520	521	590-595	596	606	642	820	C	I	II	III	IV	V	VI	VII	BC	BW	TC	TW	OR
1	CGG	.	.	.	G	-	G	C	A	C	T	C	.	T	T	.	C	1									1			
sdr4-k 3	.	T	.	.	G	-	G	C	A	C	T	C	.	.	T	.	C											1		
2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	1												
4	.	.	A	T	G	-	G	C	A	C	T	C	-	.	.	.	.													1
sdr4-n / *O. sativa ssp. japonica*	-	G	C	G	A	GCG	C	G	C	G	G	A	CCACCG	C	C	G	G	6	4	1	6	4	1	1	5	3			2	

Figure 3.25 Hapmap for the SDR4 gene obtained from the NGS results. The nucleotide positions for the SNPs and the exons where the SNPs are located are displayed. Sites with no SNPs are indicated by dots. Each line shows one haplotype and the accessions representing it. Italian accessions are grouped in C=cultigens and roman numbers = Weedy I-VII. Other accessions are summarized by: BC=Brazil cultigen, BW=Brazil weedy rice, TC=Thai cultigen, TW=Thai weedy rice and OR=*Oryza rufipogon*.

Haplotype	aa no	35	91	95	99	118	170	172	173	174	198-199	199	201	214	273
1		A	.	.	.	G	---	V	I	S	.	V	A	.	P
2		.	F	.	.	G	---	V	I	S	.	.	A	.	P
3		.	.	.	.	.	.	.	.	.	.	.	.	A	.
4		.	.	Q	S	G	---	V	I	S	------	.	.	.	.
O. sativa ssp. japonica		---	V	P	A	E	R	L	L	E	TA	A	A	A	A

Figure 3.26 The amino acids in the SDR4 protein sequence, affected by SNPs from the hapmap, are shown for each haplotype. Sites where no SNP is found in the hapmap are indicated by dots.

In the physiological part of the dormancy study the germination behaviour and dormancy breaking in weedy rice under the influence of karrikins was investigated. The dormancy of weedy rice seeds is a big unsolved problem in respect of the treatment. The dormant seeds remain in the soil for many years, and establish a seed bank that can produce new weedy rice plants after a period of treatment (e.g. with herbicides). Forcing dormant seeds to germinate all at the same time might be a step in the direction of a successful removal of weedy rice seed banks from infested fields.

Karrikins are a recently discovered class of molecules that are generated by the combustion of plant material i.e. sugars and that can break dormancy and promote germination in many plant species (Flematti et al., 2011). In this study smoke water (SW), which contains karrikins, was applied to weedy rice and cultigen seeds in different dilutions, and the effect on the germination rate was measured.

Figure 3.27 displays the results for the cultivars investigated in this experiment. The germination rate without treatment (water control) ranges between 75% and 90%. For most cultigens, except Flipper, a light increase or the same level of germination can be observed in the 1:1000 dilution of the SW. With increasing concentration of SW (1:100 and 1:10 dilution) the germination rates decrease in all accessions.

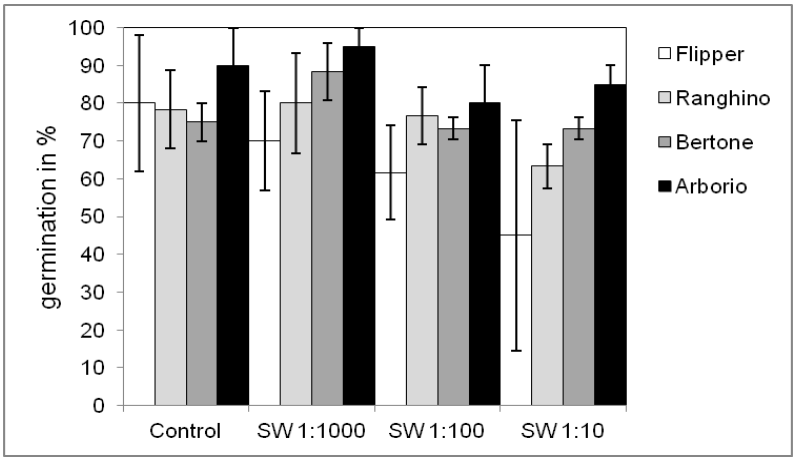

Figure 3.27 Germination rates in % for cultigen accessions at different concentrations of SW. The standard deviation is indicated by the error bars.

The germination behaviour of Italian weedy rice is shown in figure 3.28. The germination levels in the water controls are more variable. The percentage of observed germination rates ranges between 18.33% and 75%. For almost all accessions in this setup, an increase in the germination rate from H_2O to SW 1:1000 can be observed (figure 3.28). The changes are distinctly larger than that one observed in the cultigens. For the 1:100 dilution of SW, in most accessions a further increase of the germination percentage can be constituted. Some accessions show a slight decrease with the increased SW concentration. With the highest concentration of smoke water a strong decrease of germination can be reported in all the study accessions. The standard deviation shows a certain kind of variation in the data for each accession. The comparison of the germination behaviour of cultigens and weedy ice under normal conditions (H_2O) shows that germination rates in average are lower. For the treatment with smoke water the same trend can be observed for both, cultigens and weedy rice. The germination rate rises lightly with the lowest concentration of smoke water but with increasing smoke water concentration decreases considerably. The changes of the germination rates in weedy rice were stronger than in cultigens.

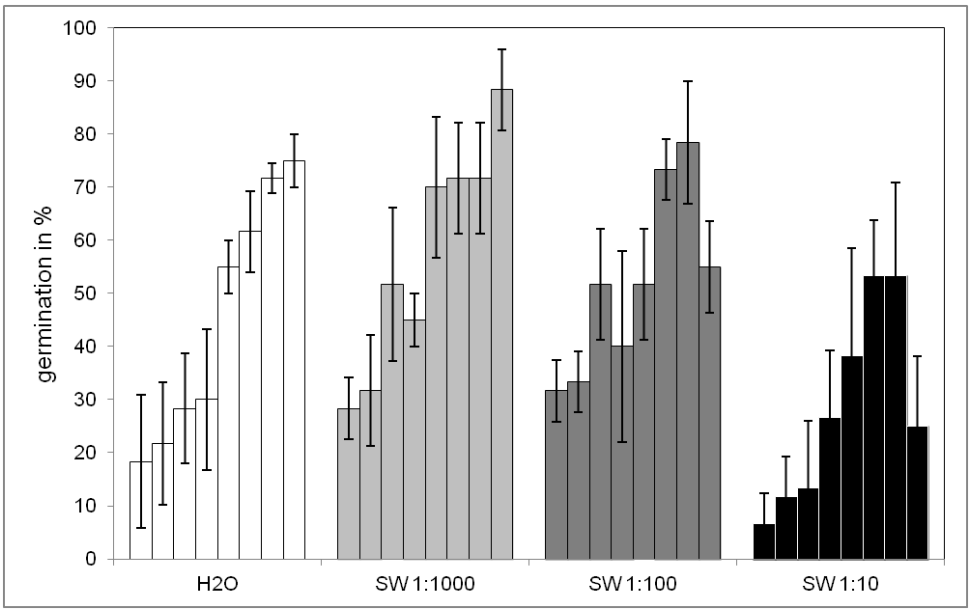

Figure 3.28 Germination rates in % for weedy rice accessions at different concentrations of SW. The standard deviation is indicated by the error bars.

The storage temperature of the seeds used in this study was 7°C. Some publications indicate that from temperatures of 5°C and increasing the dormancy can be reduced (Fogliatto et al., 2011). Considering that fact and the variability of the data the results must be seen as a preliminary trend for more detailed and focused assays to come.

3.3 Summary

The initial SSR analysis revealed high homozygosity and moderate genetic diversity for Italian weedy rice. A spatial structure analysis did not reveal any geographic origins for weedy rice in Italy. Bayesian clustering and phylogenetic analysis showed a separation of weedy rice from cultigens and an intra-weedy separation in seven distinct populations. These clusters correlated with the awn morphology of the accessions, showing also morphological separation of the genetic groups. Strikingly the clustering was not completely strict. Five cultigens were found interspersed in otherwise weedy populations. Two of these were of more recent date (1945 and 1997), the other three date back to the 19th century. These cultigens are interesting candidate accessions for possible ancestors of weedy rice in Italy.

In the following part "weedy lifestyle traits" were investigated genetically and morpho-physiologically.

The plant height was investigated by actual measurements of mature plants and the investigation of a dwarfing gene, known to cause reduced height in modern cultigens. The weedy rice accessions in average showed an increased height if compared to the cultigens. The sequence analysis of the SD1 gene did not explain the height difference of weedy rice and cultigens but it showed that *japonica* character for all Italian weedy rice accessions and *indica* character for the Thai and Brazilian accessions, each corresponding to the character of the adjacent cultigens.

The genetic and physiological aspects of the pericarp pigmentation in weedy rice were investigated. The genetic approach on the *Rc* gene, responsible for grain coloration, revealed two alleles responsible for a red pericarp in Italian weedy rice, one also found in wild rice and a cultigen-like allele that regains functionality by a 1bp deletion. The cultigen-like allele is the more frequent among accessions in of this study. The pattern of the proanthocyanidins produced was found to be very similar in all accessions, only varying in the ratio of the single components. The proanthocyanidin contents of the grains were slightly variable but no severe differences could be detected. Neither, proanthocyanidin content and pattern, was correlated with the allele type.

For the shattering trait, three transcription factor genes and the abscission zone (AZ) at the grains were examined. The *qSH1* and *SHAT1* did not reveal differences between weedy rice and cultigens but separated the Italian accessions (weedy and cultigens) from the samples of other geographic origin (Thailand, Brazil). The sh4 gene was investigated for a non-shattering SNP and other SNPs. Strikingly, all weedy rice accessions and *O. rufipogon* in this analysis were fixed for the non-Shattering SNP, even though shattering could be observed in all of

them. Other SNPs found in the study separated weedy rice accessions from cultigens. All of them resulted in substitutions of arginines to leucine or proline, respectively. The substitutions were found to be located in the same region of the protein (exon 1). The microscopy investigation of the AZ showed different types according to the shattering ability of the accessions. Cultigens expressed mainly rough and intermediate AZ phenotypes indicating no or moderate shattering. The weedy rice accessions predominantly showed smooth and intermediate types being associated with easy or moderate shattering. A 3D microscopy approach of the dimple that is formed at the AZ upon abscission showed slight differences between accessions. The dimple seems to be deeper in the less shattering cultigens than in the easy shattering weedy rice. However, great variation in the results occurred suggesting that further studies on this target should be performed to confirm the results.

For the aspect of dormancy two genes were investigated. The results for VP1 showed few haplotypes but did not allow to deduce any weedy-cultigen separations or to indicate functional changes in the protein. For SDR4 different haplotypes corresponding to alleles described in the literature were observed. These allowed separating the accessions according to *indica* and *japonica* subspecies. Again, the Italian weedy rice accessions could be shown to belong to the *japonica* tribe, like the associated cultigens, whereas the Brazilian and Thai accessions were all found to be *indica* type. As an application oriented approach the dormancy breaking potential of karrikins on weedy rice was tested in germination experiments. No significant effects could be obtained from the data but a trend of increased germination after karrikin treatment could be observed. Since the seeds in this study might have lost dormancy due to storage conditions the results must be seen as preliminary assay that should be carried further for stable results. But the trend that is indicated by the results is promising for a successful breaking of dormancy in weedy rice seeds and the approach should therefore be further pursued.

4 Discussion

Various studies aiming to reveal the origin of weedy rice in different regions of the world have been conducted (Cao et al., 2006, Londo and Schaal 2007, Jiang et al., 2012). Today it is widely accepted, that weedy rice has multiple origins in distinct geographical habitats. However, for most regions of the world the origins of weedy rice still remain unclear. The long history of rice cultivation in Asia, the high genetic variability of weedy rice and the lack of historical records often disguise the origin of weedy rice and make it difficult to reconstruct its evolutionary history. The history of rice cultivation and breeding in Italy is considerably younger than in other regions of the world and well documented. These preconditions make Italy a perfectly suiting model area to study the emergence and development of weedy rice in an agro-ecosystem. The aim of this study was to characterise weedy rice populations in Italy and seek information on their evolutionary origins. The goal was to reconstruct a model for the evolution of weedy rice in Italy from the beginning of rice agriculture until today, which could serve as an example for the formation and coevolution of agricultural weeds in cropping systems. The approach used to address these aims was to relate genetic, morphological and physiological data in a large dataset of weedy rice and cultigens. The focus of this study lies on Italy, so Italian accessions form the largest subset. The cultigens from Italy do not only represent contemporary varieties being cultivated in the collection area of our weedy rice, but also historical cultigens, dating back to the beginning of rice breeding in the early 19th century. For a better comparability, later in the study we included a small subset of weedy rice and cultigens from Thailand and Brazil, and *Oryza rufipogon*, as the wild ancestor of cultivated rice.

A close relation of weedy rice to both, ancient and contemporary cultigens, as well as a minor involvement of wild rice species in the evolution of weedy rice was assessed. The results revealed multiple origins for weedy rice in Italy that have shaped weedy rice populations over time. A small influence of wild rice species on Italian weedy rice can be concluded, most likely dating back to contaminated seed stocks in the early days of rice cultivation in Italy. However, the major source is most likely found in indigenous cultigens (past and present) which by mutations and selective sweeps developed weedy forms. The origins of weedy rice are as well anchored in the past but also found nowadays, suggesting that the formation of weedy rice is an ongoing process forced by adaptation to human agriculture.

4.1 SSR markers identify differentiated weedy rice populations and indicate multiple origins

To assess the genetic diversity and determine different groups of weedy rice in our set of Italian weedy rice and cultigens an SSR analysis with markers from Cao et al., (2008) was conducted. The data were analyzed for genetic diversity and population genetic aspects. Overall a substantial but not great genetic diversity could be drawn from the data (F_{st}, H_e and H_o). No correlation of genetic and geographic distance could be detected within our sampling (spatial structure analysis), failing to reveal geographical origins of Italian weedy rice. The Bayesian clustering revealed seven different populations for the weedy rice accessions, pointing to different origins for weedy rice and illustrated a separation of weedy rice from the cultigens. While the recent cultigens formed a separate cluster in the analysis, most of the ancient cultigens from the set were interspersed in weedy rice populations suggesting them to be possible ancestors for weedy rice in Italy.

Indicated by the F_{st} and H_e values a high a genetic variability can be found for weedy rice in our sampling area. On the other hand, very low H_o values and the number of non-polymorphic loci quench this observation. Other microsatellite studies of weedy rice, such as Cao et al. (2006), Andolfatto et al. (2011) and Jiang et al. (2012) show higher results for the same measures. The variation in the results might be an effect of the different designs of the studies. Cao et al. collected samples from an area much larger than in this study. The sampling area of Jiang et al. was also located in Northern Italy and of similar size as this study but the sampling was coarser. This way, the same set of markers reveals different levels of genetic diversity. The lower variability in our samples could be explained by the smaller sampling area and the self pollinating nature of rice (Jiang et al., 2012). The history of rice in the study area also must be taken into account. The history of cultigens and weedy rice in Italy is substantially younger than in the study of Cao et al.

A low genetic diversity in a young population of weedy rice points to a low genetic diversity in its ancestral populations. This is a type of genetic bottleneck also known as founder effect and can often be observed during neophytic invasions (Dlugosh and Parker 2008, Voss et al., 2012). The entire lack of correlation of genetic and geographic distance observed in our set of samples did not reveal defined geographical origins of the weedy rice populations. On the other hand one as to remember that this study is conducted in an artificial ecosystem. Seed dispersal of weedy rice is not achieved by natural mechanisms but mainly by agricultural practices. This could in fact disguise possible evolution centres in the sampling area.

The great morphological diversity in our weedy rice accessions stated by Fogliatto et al. (2012) and the moderate genetic diversity on the other hand seem paradox at first glance. Similar observations have been reported in several studies (Clements et al., 2004, Dlugosh and Parker 2008, Reagon et al 2011). In fact weeds often achieve a great morphological diversification out of a low genetic diversity (Reagon et al., 2011). This strategy is part of a successful weedy lifestyle and seems to apply for weedy rice in Northern Italy.

The Bayesian cluster analysis (STRUCTURE) and the distance based phylogeny (UPGMA) showed an overall separation of weedy rice and cultigens. These analyses also inferred a clustering of weedy rice in seven distinct populations. By taking a closer look it is obvious that the weedy rice and cultigen separation is not strictly maintained, and that in both, structure and UPGMA, five cultigens are interspersed in weedy rice populations. A comparison with morphological data from our cooperation partners also showed that weedy rice populations correlated with awn morphology (see also Grimm et al., 2013).

The cultigens that are genetically the most similar to weedy rice were the most promising candidates for our aim to identify evolutionary origins of weedy rice. The close relation might be due to a historical ancestry, or weedy rice was introduced from other regions and the relation was established by geneflow. Strikingly three of the five cultigens that were shown to be closely related to weedy rice, are ancient varieties: Bertone (1829), Ostiglia (1850), Ranghino (1887). Carnaroli is from 1945 but still outdated and no longer in use. Flipper (1997) is a recent cultivar which is still commonly used in the collection area.

Rice cultivation in Italy most likely began between the 13th and 15th century (Faivre-Rampant et al., 2011). Weedy rice in Italy was first mentioned in the literature in 1807 (Brioli). So a gap of about 300 years lies between the introduction of rice agriculture and the first serious infestations with weedy rice in Italy. This makes introduction from Asia as a source for weedy rice in Italy seem unlikely. Rice breeding in Italy was not performed until the beginning of the 19th century. This means, that the old cultigens of this study (Ostiglia, Bertone and Ranghino) are some of the first cultigens bred in Italy. The time points of the first reference of weedy rice and the beginning of rice breeding coincide, supporting the hypothesis that weedy rice developed from old cultigens and then spread over the agricultural landscape. Another point which supports this suggestion is, that no wild rice is native to Europe (Londo and Schaal 2008), making wild rice invasion as possible origin for weedy rice in Italy seem unlikely. These observations make the ancient candidate cultigens (Bertone, Ostiglia, Ranghino) even more interesting as probable ancestors for weedy rice. With respect to the dates of the emergence of weedy rice a possible model is that weedy rice evolved with the breeding of the

first cultigens in Italy. Due to the sowing habits in former times (transplanting) it remained a marginal problem until the shift to direct seeding in the 1960s.

A different case can be stated for Carnaroli and Flipper. They are not ancient cultigens, but more recent cultivars, Flipper is even still commonly used throughout the collection area today. Carnaroli is no longer in use, but from the 1940s and therefore not an oldie among the cultigens in this study. As it can be drawn from the rice timeline in the introduction (figure 1.4), Carnaroli is a descendant of the Ranghino tribe. Ranghino is one of the old cultigens that were identified to be progenitors of weedy rice groups. Still, Carnaroli could still be ancestor for a certain weedy rice group, but the close relation could as well be explained by the affiliation to Ranghino.

Flipper is the most recent of the candidate cultigens. It was introduced in 1997 and is also one of the cultigens clustering in a weedy rice population. This affiliation could be an effect of geneflow. Geneflow from cultivated to weedy rice has been shown in several studies (Gealy et al. 2003, Chen et al., 2004, Shivrain et al., 2007). In almost 20 years of "cocultivation", several genetic traits could have been passed on from the cultigens to weedy rice. In this case Flipper would not act as an ancestor of weedy rice but still would have an influence on its genetic constitution. However, a spontaneous transformation of the Flipper cultivar in weedy forms also cannot be ruled out. Brooks et al. (2008) showed that an American cultivar, Wells, recently evolved into a form with red grains, Red-Wells. If the reversion of the pericarp coloration can be adapted by a cultivar, then reversions of other traits can also be assumed. In this model, Flipper would act as a direct ancestor of weedy rice. These weedy rice groups would be a lot younger than the ones evolved from the historical cultivars.

The results of the SSR analysis on weedy rice in Northern Italy suggest multiple origins and different mechanisms in the evolution of weedy rice. The close relation to cultigens indicates that at least some of these origins are to be sought among the cultigens themselves. The limit of this investigation is, that no wild rice or (weedy) rice from other regions were included and therefore invasion and introgression from other sources that the native cultigens cannot be shown or excluded. Despite the considerable morphologic variation of weedy rice, certain core characteristics can be observed in all groups: plant height, pericarp colour, shattering and dormancy. To further address the question on the origins of weedy rice in Northern Italy, in the following parts these core characteristics are investigated for genetic, morphological and physiological aspects. To increase the comparability and also take in account the possibility of exterior origins, weedy rice and cultigens from other regions (Thailand and Brazil) and *Oryza rufipogon* as the wild ancestor of cultivated rice are included in the analyses.

4.2 A green revolution gene reveals *japonica* background for Italian weedy rice

The plant height is one of the most evident differences between weedy rice and cultivated rice. Weedy rice grows considerably taller than the modern semi-dwarf cultivars. These shorter varieties mark the beginning of the "green revolution" in the 1960s. The semi-dwarf varieties can achieve considerably larger yields that the taller ones. Mutations in the SD1 gene, a gibberellin oxidase, are responsible for the morpho-physiological changes in the plants. An investigation and comparison of this gene in weedy rice and cultigens therefore is a logical idea. Different haplotypes of the coding sequence were determined and compared with morphological plant height data. The comparison of the results of the absolute plant height of weedy rice and cultigens cannot be explained by a certain genetic type, pointing to alternative ways of plant height determination in weedy rice. However, the results could show the *indica* or *japonica* character of the investigated accessions, both cultigens as well as weedy rice.

The SD1 one gene was the most variable gene investigated in this study. In total 13 different haplotypes could be detected. This count is in the range of Reagon et al. (2011), who reported a total of 17 different haplotypes in a study on SD1 in U.S. weedy rice. SD1 is one of the most important genes in modern rice breeding. Compared to other domestication genes, e.g. the shattering genes, which were established in the early days of rice cultivation, SD1 gained its importance in breeding relatively recent (1960s) (Khush, 1999). Many different varieties of SD1 in cultigens are described in the literature, compared to the few fixed alleles of traits that were established earlier in domestication. This might explain the variability for SD1 demonstrated in this study.

A conspicuous observation is that most SNPs detected in exon three of SD1 in this study are heterozygous. The fact, that also homozygous types were detected with the same primers and that a BLAST search against the rice genome (RAP-DB) only yielded one hit for each primer, makes a sequence artefact unlikely. The heterozygote haplotypes are spread over almost all groups of weedy rice and cultigens and were also found in *O. rufipogon*. Reagon et al. (2011) also report heterozygotes in their study, but only in *O. rufipogon* and as a rare phenomenon. Regarding the translated protein sequence for many of the "heterozygous SNPs" no difference between the two alleles can be stated, but some of them do have an amino acid sequence difference from the reference. These SNPs or affected amino acids, respectively, are found in the middle of exon three. Referring to Ashikari et al. (2002) and the Uniprot database (Magrane, 2011), this region contains the domain for the binding of the substrate. These SNPs

might very well play a role in the function of the expressed protein. Regarding the distribution of the accessions over the different phenotypes it can be observed that there seem to be no haplotypes primarily found in either weedy rice or the cultigens. Since cultigens and weedy rice share the same haplotypes but the latter still grow considerably taller, a direct influence of SD1 on the plant height of Italian weedy rice is questionable pointing to alternative mechanisms for height adaptation. In fact, a second gibberellin oxidase (GA20OX1) is located on the rice genome. It is usually primarily expressed in the reproductive organs (Asano et al., 2008). The author of this study reported that an enhanced expression of GA20OX1 can partially compensate the lack of GA20OX2 and lead to taller plants. The sequencing of GA20OX1 was attempted in this study but could not be successfully achieved until now. Since a change in GA20OX1 sequence or expression is likely to be the explanation for the taller growth in Italian weedy rice, sequencing and expression profiling of GA20OX1 should be carried out in the future to further investigate this subject.

In total a considerable number of haplotypes could be detected for SD1. By taking a closer look it strikes the eye that many of the haplotypes are very similar to each other, often just differing by one SNP. Apart from small differences the haplotypes can be divided in two major groups. One of these groups is characterized by a big deletion in exon one, the other lacks this deletion. All accessions (weedy rice and cultigens) from Thailand and Brazil are found in the group with the deletion, all Italian accessions in the other group. This separation is strict and shows no exceptions. The deletion in exon one being characteristic for the Thai/Brazil group is also known as the "green revolution mutation". It was the first SD1 allele reported in the literature that was associated with a semi dwarf phenotype. It was found in the cultivar IR8, which is also known as "miracle rice". This cultigen marks the beginning of the green revolution in the 1960s (Khush, 1999). The large deletion leads to a frame shift resulting in a premature stop codon in exon two. It was the first allele for semi-dwarfing used in rice breeding. Since then many other varieties with different dwarfing alleles of SD1 have been reported. Interestingly a weedy rice shares the deletion in exon one whereas another one of the Thai/Brazil group, shares all other SNPs but the deletion with the cultigens. The latter could have regained the missing fragment by hybridising with *O. rufipogon* or other wild rice sources. This accession is a Brazilian weedy rice. No *O. rufipogon* is native to either North or South America so it seems unlikely as donor. But a look into history might explain this observation. Rice was introduced to Brazil by two different sources: the Portuguese and slaves from Africa (Carney, 2004). The rice brought by the Portuguese colonists was of Asian origin, old cultigens derived from *O. rufipogon*, the rice brought in by the slaves was African

rice, descended from *Oryza glaberrima*, African wild rice. Contaminations of both wild ancestors could have been introduced and affected weedy rice evolution in Brazil. This is highly speculative but a study with a large set of weedy rice from Brazil, Asia and Africa, including wild and cultivated rice, could provide an insight in the weedy rice evolution in South America.

The second weedy rice haplotype in the Thai/Brazil group has the green revolution deletion in exon one. Nevertheless, weedy rice from Thailand grew taller than Thai cultivars in our greenhouse (data not shown). Weedy rice groups with the "green revolution" mutation have been reported before in the U.S. and Japan (Kawasaki et al., 2009, Reagon et al., 2011). In Japan, accessions with the deletion in exon one have been shown, but those expressed the semi-dwarf phenotype then. In their study on SD1, Reagon et al., (2011) also report weedy rice accessions with the deletion. They explain their observation with introgression from cultivated rice. Geneflow from cultigens to weedy rice has recently been reported several times (Gealy et al. 2003, Chen et al., 2004, Shivrain et al., 2007). Since all the cultigens from Thailand in this study also harbour the deletion in their genome this model would as well apply for the case in this study.

The amino acids 100 and 340 in the GA20OX2 are of a certain importance, because they can tell about the *indica* or *japonica* subspecies affiliation of an accession. During rice domestication these two amino acid positions changed from G and R to E and Q in *japonica*. The SD1-GR form is maintained in *O. rufipogon* and *O. sativa ssp. indica* (Asano et al., 2011). This observation splits the accessions of this study in two groups: The SD1-GR group formed by *O. rufipogon* and all Thai and Brazilian accessions and the SD1-EQ group uniting all Italian accessions. For the Thai and Brazilian accessions with the deletion in exon one, the G at site 100 cannot be observed, because it is deleted, but the R at position 340 can be shown. To my knowledge and according to Asano et al. (2011) so far to accessions with intermediate alleles have been reported, which is not challenged by the results of this study. The GR allele of SD1 found in all Thai and Brazilian accessions (weedy and cultigens), shows their *indica* character, whereas the Italian samples clearly can be identified as *japonica* types. The fact that both, Thai and Brazil accessions, share the same group seems paradox, regarding their geographical and historical origins but is justified taking into account that they both belong to the *indica* subspecies. The comparison of the SD1 allele types again shows the close relationship of weedy rice and the endemic cultigens. No *indica* forms were found in the *japonica* background of Italian rice and the other way round for Thailand and Brazil. No

weedy rice could be observed sharing the exact genotype as *O. rufipogon*, making influence of wild rice on the plant height of weedy rice seem unlikely for all cases in this study.

The study on the SD1 gene in the set of weedy rice and cultigen samples showed a high diversity of haplotypes being subdivided in two main groups: and *indica* clade and a *japonica* clade. Accessions form Thailand and Brazil were all identified as *indica* varieties and so were the weedy rice accessions of the respective regions. All cultigens and weedy rice accessions for Northern Italy could be shown to have a *japonica* background. However, the differences in plant height of weedy and cultivated rice, being the motive to study this gene, are not explained by the sequence architecture of SD1. The key for the tall growth in weedy rice is most likely to be sought in the second gibberellin oxidase (GA20OX1) found in the rice genome. Further studies on this gene, such as expression profiles or localization studies, might unravel the weedy strategy for the tall weedy rice and provide new basic knowledge on the shoot growth regulation in rice.

4.3 The *Rc* gene unravels wild rice background in weedy rice

The red colouration of the pericarp is probably the most famous characteristic of weedy rice, leading to the term "red rice" as a synonym for weedy rice. The pigmented pericarp is also a characteristic of wild rice species (e.g. *O. rufipogon*) and has undergone transition to white during domestication. Most of the recent cultivars have white seeds. The differentiation in the seed colour between weedy rice and cultigens makes the *Rc* gene, and its downstream products, an interesting research object for the aims of this study. The results show a clear separation of weedy rice and cultigens not only by morphological but also by genetic means. Two different alleles of *Rc* were detected in Italian weedy rice suggesting different origins of the groups.

The sequencing of a part of exon 7 of the *Rc* gene showed three different allele types for the region of interest. One of them of course represented the typical cultigen allele with a 14 bp deletion, leading to a premature stop codon and a white pericarp. It was shared by all cultigens in this study regardless the geographical background. The other two allele types recognised display a red pericarp phenotype. One of them was the wild type allele which is found in wild *Oryza* with red pericarp. This type was represented by *O. rufipogon* and some weedy rice accessions. The other one also harbours the 14 bp deletion known from the cultigen allele plus an additional 1 bp deletion 43 bases upstream of the original mutation. For none of haplotypes heterozygous accessions could be detected.

The fact that all cultigens represent the 14 bp deletion was an expected result since this is the most common allele for a white pericarp in cultivated rice. The other cultigen-form of *rc* is an SNP also causing a premature stop codon. This allele is referred to as *Rc-s*, and reported to be found only in 3% of the cultivars (Sweeney et al., 2007) therefore it is not surprising that it is lacking in this study. Among all investigated weedy rice accessions none were detected to have the domestication-type allele and a white pericarp, respectively. This consequently suggests a great evolutionary importance for red grains in weedy rice.

The evolutionary original form of *Rc* was determined of course in the representative of wild *Oryza*, *O. rufipogon*, and several weedy rice accessions from each Thailand, Italy and Brazil. Thai and Brazilian weedy rice were shown to only display the wild type allele and no other form of *Rc* in this study. The appearance of wild genetic material in weedy rice definitely shows the involvement of wild rice in weedy rice background. These DNA segments could have been recently introgressed into weedy rice populations by hybridization or represent relicts from the past, indicating an ancestry of wild rice for weedy rice. Both, the model of introgression and ancestry are suitable for the Thai weedy rice group. Wild *Oryza* are native

in Thailand and could therefore serve as donors for genetic material to weedy rice in regions with rice agriculture. However, the case for Brazil is different. No wild *Oryza* are originally found in South America but slaves from Africa brought wild rice (*Oryza glaberrima*) and grew it close to their habitations (Carney, 2004). It is conceivable that some rice plants escaped to the wilderness where they grew close to agricultural spaces and therefore contributed to traits like seed coloration in the historical background of weedy rice in Brazil.
Since in Europe also no wild *Oryza* are indigenous, recent geneflow between wild and weedy rice in Italy is unlikely. But considering Italy's history of rice cultivation, in particular, the long periods of importing seeds from Asia between the 15th and 19th century (Faivre-Rampant et al., 2011), contaminations with wild *Oryza* species in agricultural areas, caused by wild rice in seed stocks, are possible. Thus, wild rice could have contributed to weedy rice evolution in Italy and played a role in the ancestry of weedy rice groups in Italy.
The third type of *Rc* allele depicted in this study is the cultigen-like form with the common 14 bp deletion and an additional deleted site upstream the original mutation. This further deletion reverses the stop codon and restores the reading frame. The gene product is five amino acids shorter and is altered in the region between the two deletions, but seems to be functional since representatives of this allele express a red pericarp. It was found only in Italian weedy rice accessions in this study. This exact form of *Rc* has been described by Gulick et al. (2009) as a mutation in the Italian cultivar Perla transforming it in a form with red pericarp, Perla-Rosso. A similar observation has been made in the U.S. by Brooks et al. (2008) where the cultivar Wells mutated in Red-Wells also expressing a red grain phenotype. The mechanism for this transformation is the same as in the previous variant, but the additional 1 bp deletion is found 19 bases upstream of the deletion (43 in Italian samples). This allele is referred to as *Rc-g* in the literature. Both authors point out that the observed accessions with red pericarp are not weedy rice but cultigens. The proof of such an allele in weedy rice populations suggests a model of dedomestication of cultigens as one of the forces leading to the evolution of weedy crop varieties and supports the presumption of cultigens being direct ancestors of weedy rice in Northern Italy. The readaptation of seed pigmentation as shown for Perla and Wells (Brooks et al., 2008, Gulick et al., 2009) could be followed by other alterations enhancing the compatibility against the original cultigens leading to evolutionary stable weedy forms.
Since both, the cultigen-like allele and the wild form are shown in Italian weedy rice multiple origins can be assumed. Some of them clearly must be sought among the endemic cultigens but also an involvement of wild rice is shown by the investigations on the *Rc* gene. The ratio of the cultigen-like allele and the wild form is 75% to 25% in this study. Based on this

numbers, dedomestication of cultigens can be hypothesised as being the more important force in weedy rice evolution in Northern Italy.

The importance of the seed pigmentation for weedy rice is accentuated by the fact that this trait is common in all weedy rice accessions examined and that several mechanisms for the achievement of this trait are shown. The fixation of the red pericarp in weedy rice seems paradox, since a white pericarp would be a better mimicry of the competitor and protect from selection by human powers. However, the proanthocyanidins in the rice pericarp are suspected to play a role in dormancy and protect dormant seeds in the soil from destruction by bacteria and fungi. Since dormancy is another important trait of weedy rice, because it enables seeds to outlast periods of weed treatment in the soil and therefore guarantees the continuance of its existence.

All samples investigated in this study displayed homozygous genotypes for their *Rc* allele. The defective allele with the 14bp deletion, *rc*, is recessive and therefore homozgyotes for the white cultivars is expectable. The importance of the pigments produced in case of a functional *Rc* gene product for weedy rice is indicated by the results of this study. Hence, heterozgyotes with one functional and one dysfunctional allele would not be favourable for weedy rice and therefore be extremely rare or non existent.

The pigmentation of the pericarp with proanthocyanidins has been shown to be important in weedy rice. The quality and content of proanthocyanidins in the grains of different weedy rice accessions was analysed using HPLC and a vanillin assay.

The qualitative analysis (HPLC) showed overall the same pattern for all samples of the study. The only variation could be detected in the ratios between the components, which split the samples in two subgroups. Comparing the allele types of individuals of the two subgroups no correlation of allele type and HPLC pattern could be asserted. So the pattern of the proanthocyanidins expressed is the same in both cases, indicating no limitation for the altered protein from the cultigen-like allele.

Furukawa et al. (2006), who investigated the involvement of *Rc* and another gene (*Rd*) in grain pigmentation of rice, also generated HPLC profiles of grain extracts from rice accessions with a red pericarp. The pattern of these samples is different from our analysis. These differences could be the result of the different extraction and HPLC methods conducted by Furukawa et al. But, they might as well be the effect of dissimilar environmental conditions in which the plants were grown or different conditions of seed storage and seed age, respectively.

The content of proanthocyanidins in pigmented grains, calculated by a vanillin assay, showed minor variation between accessions. The variations are not linked to certain weedy rice groups (Populations I-VII) or the rc allele type. They therefore are most likely to be seen as results of environmental conditions of the donor plants or as inaccuracies of the method. However, comparing the values of this study (14.69 – 38.81 µmol g^{-1}) to the data obtained by Furukawa et al. (2006) who also investigated the content of pigmented rice (1.27-37.74 µmol g^{-1}) they cover a similar range. This leads to the assumption that, independent of the allele type and minor changes in the environmental conditions, the amount of proanthocyanidins produced by red rice phenotypes is relatively constant.

The investigation of the grain pigmentation in weedy rice unravelled different mechanisms for obtaining this trait. Two different alleles were found to result in the same phenotype and also in the same physiological expression of red pericarps. The two alleles had different origins, one to be found in cultigens, achieved by mutation and the other being the ancestral form found in wild rice. Based on these results multiple origins for weedy rice in Northern Italy can be assumed of which one must be sought in the cultigens, hypothesising a dedomestication model (endoferality) while the other suggests the involvement of wild *Oryza*, most likely *O. rufipogon* in its evolution (exoferality). The induction of *O. rufipogon*, being not native to Europe, has most likely been caused by contaminations in seeds imported from Asia. The import of sowing material was a common habit between the 15th and 19th century in Italy. Considering the ratio of the alleles, a cultigen-based way of weedy rice evolution seems to be more common. Weedy rice groups with wild rice ancestors are more likely to represent relicts of former times. Hereby the results of the SSR analysis, that show a close relation of certain cultigens to weedy rice is further supported. The remains of wild rice genetic material in the genepool plus the more recent cultigen-like allele in *Rc* to the point of lately ongoing mutations in *Rc* observed in modern cultivars, shows that weedy rice is not only to be contemplated as the result of an evolutionary event in the past but as an ongoing process generating further weedy rice groups until today.

4.4 Was shattering in weedy rice reacquired by mutations in a cultigen transcription factor?

The reduction of seed shattering is a milestone in the domestication of crop plants (Purugganan and Fuller 2009). On the other hand increased seed shattering is a hallmark for weedy species, since seed dispersal guarantees their persistence. It is therefore one of the "weed adaptive traits" depicted by Ellstrand (2010), that are obtained due to similar selection pressures to weedy plants are exposed. In this study three identified shattering genes (*qSH1*, *SHAT1* and *sh4*) and the abscission zone after seed detachment have been investigated. The results suggest a model by which the seed shattering ability in Italian weedy rice could have derived from moderate shattering cultivars of the region, rather than an involvement of wild rice.

Overall a low variability was detected in the shattering genes investigated (*qSH1, SHAT1, sh4*). The *qSH1* gene is a homeodomain-type transcription factor (Thurber et al., 2010). A SNP in the 5' regulatory region of *qSH1* is suspected to cause a reduced shattering phenotype. This site is not covered by the sequences in this study. However, four haplotypes for the coding sequence could be obtained for the data. They split the samples in the previously identified *indica* group (all Thai and Brazilian accessions), *japonica* group (all Italian accessions) and *O. rufipogon* as wild rice representative. Most of the SNPs do not cause alterations in the amino acid sequence, so functional consequences are assumed to be unlikely.

The *SHAT1* gene also is a transcription factor, belonging to the APETALA2-group (Zhou et al., 2012) showed only two groups based on the sequence. One represents the *indica* group, the other the *japonica* group and *O. rufipogon*. Interestingly, the *indica-japonica* separation does not reach total agreement because two Italian cultivars, definitely being *japonica* species are grouped with *indica*. This overlapping might be an artefact from breeding, often using crossing. Both cultigens might have *indica* heredity derived from hybridisation. Matching the observations in *qSH1*, for *SHAT1* no differences between weedy rice and cultivated rice can be stated, deducing that this gene as well does not affect the shattering ability in weedy rice.

The *sh4* gene, representing a myb3 transcription factor (Li et al., 2006), is one of the most thoroughly studied genes involved in shattering. A SNP, converting a T to a G in exon one of *sh4*, was long suspected to be the major event in the achievement of non-shattering varieties during rice domestication. In the meantime, many studies showed that also shattering accessions are fixed for the G allele, challenging that hypothesis (Thurber et al., 2010, Zhu et

al., 2012). This study confirms these observations since all accessions show the "G-allele" despite the shattering phenotype. However, four haplotypes for *sh4* could be obtained from sequencing: one representing the *O. sativa ssp. japonica* sequence and three others each characterized by a single SNP in exon one. The haplotype two has also been described by Zhu et al. (2012) as a fixed SNP for weedy rice accessions from Italy and Spain. The accessions were parted in cultigen accessions, sharing the *japonica*-type sequence and weedy rice, distributed over the remaining three types. Regarding the derived amino acid sequence each of the single SNPs replaces an arginine with another amino acid (leucine or proline). All substitutions are located in a section between the trihelix DNA binding domain and a proline rich region (Lin et al., 2007). The location of the SNPs and the fact that they target the same amino acid suggest a functional polymorphism for the haplotypes one to three. Since these haplotypes are solely found in the easy shattering weedy rice accessions, the functional polymorphisms might be involved in the enhanced shattering ability. On the other hand few weedy rice accessions and the also shattering O. *rufipogon* share the *japonic*a haplotype. Like *O. rufipogon*, they all share the wild type allele for *Rc*, whereas in the types1-3, both *Rc* alleles are found. The linkage of shattering and *Rc* has been indicated by Sweeney et al. (2006). Considering the results, two models for shattering in weedy rice can be assumed, depending on the combination of *Rc* and *sh4* alleles. Multiple ways of shattering further support the hypothesised multiple origins for weedy rice. Further these results challenge the hypothesis postulated by Zhu et al. (2012) proposing a minor role of *sh4* in the shattering process. In this study, the SNPs in *sh4* were the only differences in the investigated shattering genes that could explain the difference in the shattering behaviour. Three SNPs from geographically and evolutionary independent, shattering weedy rice groups, were found in the exact same region of the gene. These results indicate certain significance for the *sh4* allele type in conjunction with shattering.

Microscopy based investigation of the AZ of detached seeds in Italian accessions revealed different shattering types. 3D measurements of the dimple formed on the grain when detached from the pedicel showed great variability within accessions. Therefore clear statements are difficult to make but trends can be observed. Hereby cultigens seem to tend towards a deeper dimple compared to weedy rice possibly caused by the larger forces that must apply for detaching the seeds. Interestingly most of the candidate cultigens, suspected to be ancestors of weedy rice, showed a flat dimple, suggesting increased shattering ability.

The 2D visualisation of the AZ helped to classify accessions in three groups by the surface pattern. Rough, intermediate and smooth types were identified, indicating non-, intermediate

and easy shattering phenotypes. The cultigens showed patterns typical for non- and moderate shattering, weedy rice displayed moderate and easy shattering patterns. The considerable number of intermediate types found among the cultigens, suggest a moderate shattering ability for the majority of rice cultivars in Italy. All candidate cultigens for weedy rice ancestry (Ostiglia, Bertone, Ranghino, Carnaroli and Flipper) showed a moderate shattering AZ phenotype. Moderate shattering varieties were favoured in the past because hand threshing, was easier performed on those varieties (Ji et al., 2006, Li et al 2006). The fact that all historical cultigens (Ostiglia, Bertone, Ranghino) in this study are found to be moderate or even easy shattering, is not surprising. All weedy rice accessions showed an intermediate or smooth type, depicting the severe shattering ability. The easy shattering habit of weedy rice might originate from either wild rice, or the moderate shattering ability of cultigens that was increased again. Thurber et al. (2011) showed that the timing of the degradation of the abscission layer in weedy rice and wild rice is different and therefore conclude a parallel evolution of shattering rather than an evolutionary link. Considering that, and the observation that all candidate varieties are most likely moderate shattering, the origin for shattering in Italian weedy rice is most likely to be sought in a cultigen background. Even though the indications are strong, a quantitative analysis and comparison of the shattering ability in weedy rice and cultigens of Italy as shown for other weedy rice groups in different studies (Ji et al., 2006, Thurber et al., 2011), might help to get a more complete and well founded impression of the development of this trait in weedy rice.

The investigation of shattering by molecular and microscopy based methods, suggest a role of the transcription factor *sh4* in the shattering ability in weedy rice. Single SNPs in the section between two domains could be responsible for the increased shattering ability in weedy rice. The sequence changes could have been derived from cultigens as well as wild rice. The evaluation of the AZ morphology suggests a cultigen origin rather than a wild ancestry for shattering in weedy rice. All candidate cultigens identified in the SSR analysis fit the preconditions for this hypothesis. Since *O. rufipogon* and some weedy rice accessions in this study shared that *sh4* haplotype of cultivated, reduced shattering rice, the involvement of other factors in the shattering process is highly promoted. Considering that the variability in all investigated shattering genes was very low and single SNP is suggested to cause functional changes increasing the shattering ability, a selective sweep, as observed for the Rc gene, can be assumed for shattering.

4.5 A dormancy gene shows the *indica* or *japonica* character of weedy rice groups

Field studies and laboratory experiments on the weedy rice accessions used in this study, conducted by Fogliatto et al. (2010), showed considerable levels of seed dormancy in Italian weedy rice. In this study two genes, VP1 and SDR4, associated with seed dormancy were investigated. The coding sequences did not reveal an explanation in the difference in germination levels between weedy rice and cultigens but again confirmed the *japonica* character of Italian weedy rice. A germination experiment with smoke water shows a trend that could lead to dormancy breaking in weedy rice.

The two genes, VP1 and SDR4, are nuclear transcription factors involved in the establishment of dormancy in seeds (Fan et al., 2007, Sugimoto et al., 2020). VP1, a positive regulator of SDR4 (Sugimoto et al, 2010), showed a relatively low variability in this study. Five groups were revealed by sequencing of the exons, each differing in only few SNPs from the *O. sativa ssp. japonica* sequence that was used for comparison. Most of the SNPs were silent in the protein sequence i.e. no alterations in the amino acid sequence were caused. The four amino acids that are affected by SNPs in the VP1 sequence are not suspected to have an effect on the function of the gene product. These findings suggest if at all only a minor role for VP1 in weedy rice seed dormancy. SDR4 acts downstream VP1 and is also involved in the formation of dormancy in rice seeds. In the haplotype analysis, five haplotypes could be detected, three of them being consistent with the results of Sugimoto et al. (2010). In a cloning study on SDR4 they showed three haplotypes sdr4-n, sdr4-k and sdr4-k'. The allele sdr4-n is represented by the *japonica* sequence in this study, sdr4-k by a Thai cultigen (haplotype 2). The third, sdr4-k', matches with haplotype 1, except for one SNP (position 271) that was not found in this study. The original study also states sdr4-n as the *japonica* allele, which is associated with reduced dormancy. But this study also grouped five *indica* accessions in the sdr4-n haplotype. Sdr4-k represents the *indica* subspecies and *O. rufipogon*. In this study sdr4-k was found in a Thai cultigen, which also is of *indica* type. *O. rufipogon*, in this study, formed a separate group but differed only in a few sites from sdr4-k, which could be accession specific. Sdr4-k' could be shown in an Italian cultigen and a Brazilian weedy rice accession. The sdr4-n allele in *indica* accessions and the presence of an *indica* allele in an Italian cultigen (*japonica*) seems paradox, but Sugimoto et al. (2010) report both alleles (sdr4-n and sdr4-k) in *indica* varieties most likely as a result of introgression. The same explanation might apply in the case for this study. The fact that all Italian weedy rice accessions are found

to have the sdr4 allele for *japonica* rice (sdr4-n), again shows its affiliation to this subspecies. A direct involvement of these two genes in the dormancy of Italian weedy rice cannot be concluded from these data. However they suggest a role of upstream regulators or alternative mechanisms for dormancy in weedy rice and further promote our hypothesis that weedy rice emerged from cultigens that dedomesticated.

The distinctive seed dormancy in weedy rice is one of the traits that make treatment of infested soils extremely difficult. Dormancy breaking therefore might be a step in the direction of an effective management for weedy rice. Fogliatto et al. (2010) showed for the same accessions as used in this study that dormancy could be reduced by winter flooding of the paddies. The reduction of dormancy by application of karrikins is another approach and was addressed in this study. Karrikins have been shown to have a dormancy breaking ability in most plant groups and are produced by burning sugars, e.g. the cellulose in the case of bushfires (Nelson et al., 2009). The effect of different concentrations of smoke water (SW) on the germination rates of rice seeds showed an increasing trend for low concentrations and a germination inhibiting effect for the highest concentrations. This decrease might be caused by other substances accumulated in the water during combustion. However, these results were extremely variable and cannot be taken for safe conclusions. They can indicate a trend that should be investigated and confirmed in further experiments. The germination rates in the water controls in both, cultigens and weedy rice, were relatively high, indicating no significant dormancy levels for weedy rice, although Fogliatto et al. (2010) had shown dormancy for the same accession panel as used in this study. The loss of dormancy might have occurred during storage. The seeds for this study were stored at 7°C. It is shown in the literature that temperatures from +5°C increasing can promote dormancy breaking (Gianinetti and Cohn 2008, Fogliatto et al., 2010). The data of the germination assay can be seen as a preliminary assay for experiments to come with truly dormant seeds. Since a germination enhancing trend could be seen in the seeds of this study, a dormancy breaking effect of karrikins for weedy rice is likely.

4.6 Exoferality and Endoferality shaped the evolution process of Italian weedy rice

This study aimed to investigate weedy rice groups in Northern Italy and seek information on their origins and evolution. An initial SSR analysis identified five cultigens of the sample set as possible progenitors of weedy rice in the Piemonte region. Three of them represented old cultigens dating back to the beginning of rice breeding in the early 19th century (Bertone, Ostiglia and Ranghino) the other two (Carnaroli, Flipper) are more recent (20th century).

The further investigation of characteristic weedy traits on the genetic and morpho-physiological level gave further information on the origins and evolution of weedy rice in the study region. Two genes (SD1 and SDR4) showed that Italian weedy rice, like the endemic cultigens, belong to the *japonica* subspecies of rice, supporting the hypothesis of cultigens as ancestors for certain weedy rice groups. Assays on the pericarp pigmentation additionally revealed wild rice background for weedy rice in Italy. The shattering ability of weedy rice in the study region seems to be linked to a SNP in a transcription factor and most likely has derived from the moderate shattering ability of the previously identified candidate ancestors. Both, pericarp colour and shattering ability are thought to represent a selective sweep, as reported for weedy rice groups in other geographical regions (Brooks et al., 2008).

The results of this study reveal multiple origins for Italian weedy rice. Some ancestry in wild rice was shown, probably dating back to the long period of seed import from Asia (~15th – 18th century). Wild rice species might have been introduced as contaminations in sowing material reached the paddies and played a role in the evolution of some weedy rice groups. Old cultigens also have been shown to be involved in weedy rice progeny. These cultigens date back to the 19th century, when cultigens were not genetically identical like cultivars today but rather landraces, bred for desired morphological and physiological characteristics. Individuals of a landrace are not genetically identical and therefore provide a greater level of genetic variability than cultivars. This diversity might have provided the background for the evolution of weedy rice under the selection pressures in the agricultural system. Additional factors like the switch from transplanting to direct sowing in the 1960s favoured the further spread and propagation of weedy rice in Italy. But not all origins of weedy rice in Italy must be sought in the past. Contemporary cultigens also were proven to be closely related to weedy rice in this study and are therefore suggested to have played a role in the development of weedy rice. The close relation might as well be explained by geneflow from cultigens to weedy rice as reported several times (Gealy et al. 2003, Chen et al., 2004, Shivrain et al.,

2007) but the results and other studies (Cao et al., 2006, Jiang et al., 2012) also suggest "dedomestication" of modern cultigens as the motor for weedy rice evolution. In this case weedy rice populations have developed from spontaneous mutations. Quantitative comparison in this study (*Rc* alleles) indicates that Italian weedy rice derived from cultigens seems to be more common than contamination with wild *Oryza*. The multiple origins (wild ancestry, dedomestication), for weedy rice in Northern Italy can be traced back to the past but also circumstantial indications for the development of weedy rice groups until the present day were shown. This shows that the formation and evolution of weedy rice is not only a product of an event that took place in the past but also an ongoing process until today, forced by selection pressures applied by agricultural practices.

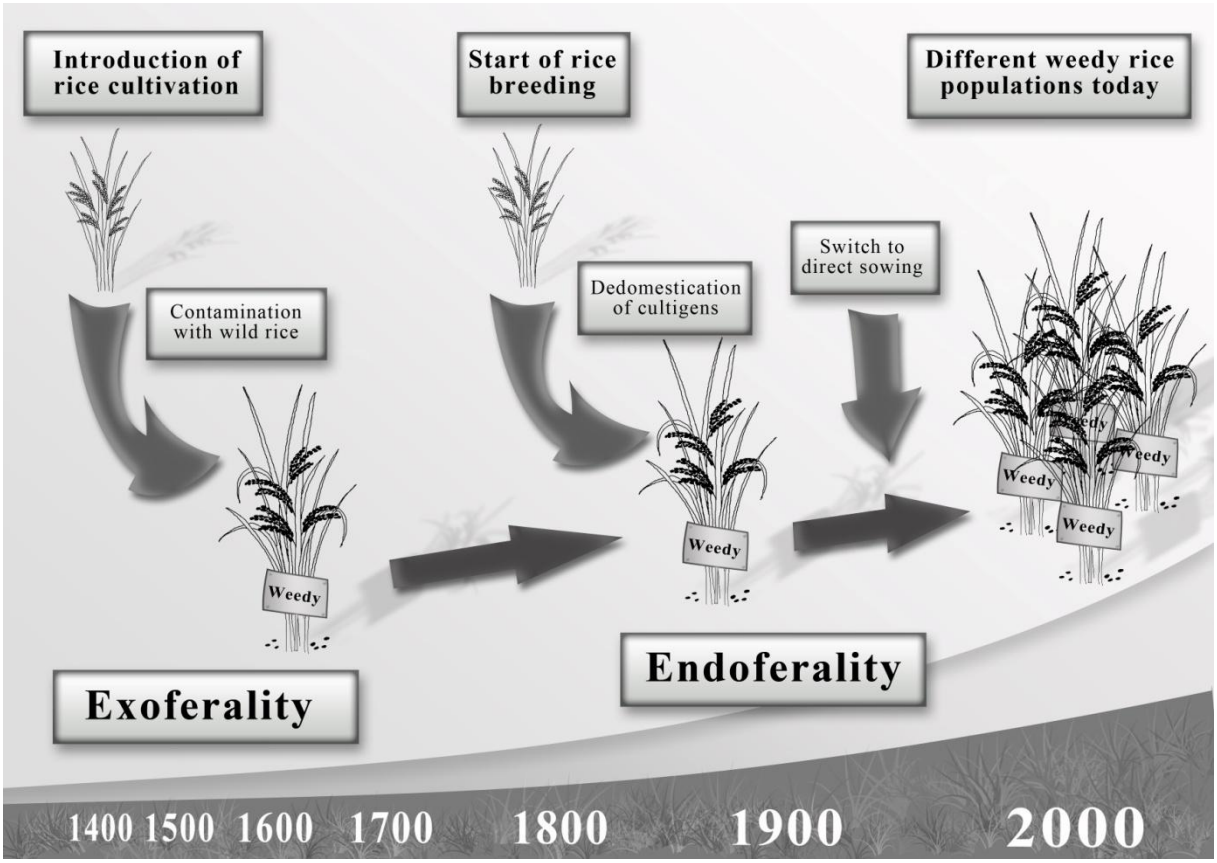

Figure 4.1: Timeline of weedy rice evolution in Northern Italy from the beginning of rice agriculture until today. Induction of wild rice from contaminated seed stocks most likely lead to the first weedy rice populations around 1500, representing an exoferality model. From the start point of local breeding programs (~1800) until today, the dedomestication of cultigens produced further weedy rice populations by endoferality. With the switch to the direct sowing practice around the 1960s, the infestations increased dramatically, leading to a large number of weedy rice population in the agricultural landscape of Italy today.

Based on the results of this study a timeline model for the evolutionary process of weedy rice in Northern Italy can be deduced (Fig. 4.1). The beginning of rice cultivation in Italy most likely took place around the 15th century (Faivre-Rampant et al., 2011). In this early stage of rice agriculture no breeding was performed in Italy, instead seeds were imported from Asia and then cultivated in the paddies. The influence of wild rice in Italy's weedy rice populations, that was revealed by the *Rc* gene in this study, most likely happened in this period of about 400 years. Seeds, contaminated with wild rice from Asia, might have been imported and wild Oryza might have entered Italian paddies via that pathway. The mode when interactions of a wild species with a crop result in the formation of a weed is defined as exoferality (Ellstrand et al., 2010). The exoferality events in Italy most likely generated the first weedy rice populations in this geographical area.

The practice of seed import and cultivation in Italy did not change until the early 19th century when the first breeding efforts were made. The first Italian rice cultivars were most likely landraces, meaning individuals were not genetically identical, like in cultivars , but shared a certain level of genetic variability. These landraces like for example Ostiglia and Bertone

might have developed weedy forms by readaptation of shattering and the secondary adaptation of a pigmented pericarp by spontaneous mutation. This idea is supported by the *japonica* identity (*SD1*, *SDR4*) and the close genetic relation of Italian weedy rice to the cultigens (SSR analysis). The process of a crop changing into a weed is also called endoferality. The results of this study suggest that endoferality also occurred, but later in time than exoferality. Contemporary cultigens with an extremely close relation to weedy rice indicate that dedomestication is still taking place and thereby produces new weedy rice lines. The change to direct sowing in the 1960s increased the number of weedy rice infestations and made it a serious threat of rice agriculture in Italy until today.

Another interesting aspect of the model is the time. Rice has been domesticated and bred by humans for millennia to achieve today's cultivated varieties. In comparison to the domestication that spanned thousands of years, the evolution of several populations of feral weeds in only a few centuries and even decades is an extremely rapid process. The ability to quickly adapt to environmental conditions most likely is one of the strategies that makes weed colonization successful.

This model illustrated that weedy rice in Italy has multiple origins and evolved in different modes of action (exoferality and endoferality), some of which can be directly linked to events in agriculture. It also shows that formation and evolution weedy rice is an ongoing process until today and will be further going on in the future. The goals for the future therefore cannot only be the efficient treatment of already existing weedy rice populations but also the avoidance of the formation of new weedy rice populations by adapting the agricultural practices.

4.7 Future challenges for agricultural weed management

The previous chapter showed that the evolution of weedy rice in Italy is a continuous process over time that produced various populations. This process is exemplarily for the development of agricultural weeds. The introgression of wild species and/or mutation of crops produce the first crop weed populations in a new agricultural area. Others follow over the time by adapting to the selection pressures that are applied through human action. Compared to the domestication process, weed evolution is substantially more rapid. As for the case of weedy rice in Italy, crop weed development is a dynamic process that continues over time leading to the formation of novel populations. The new weeds will have adapted to the methods of treatment, which act as selection pressures. This course reveals new challenges for the management of agricultural weeds in the future. The evolution model shows, that it cannot only be the goal to treat agricultural areas for already existing weeds, but also strategies for the prevention of the development of new feral weeds must be considered.

Regarding the process of crop weed formation, the weeds act according to the "red queen hypothesis" postulated by Van Valen (1974). This evolution theory refers to the character of the red queen in Lewis Carrolls book "Through the looking glass" (1871), who has to run constantly as fast as she can just to stay in one place. In his theory, Van Valen transfers this metaphor to co-evolving species, which have to constantly adapt to not become extinct. By applying this theory to the case of weedy rice in Italy it confirms the observation that new weedy rice populations emerge over time and predicts new, better adapted weedy rice populations in the future. The counterpart of weedy rice are the cultivars. They do not evolve by natural selection, but by human selection applied through breeding and genetic modifications. Weedy rice adapts to compete with the cultigens and to overcome the weed control methods. One example for the adaptation of weedy rice to cultigens can be seen by the differences in plant height. Cultivars are bred for semi-dwarf phenotypes to increase productivity. Weedy rice populations adapted by expressing taller phenotypes to overgrow the cultivars. As an adaptation from the agricultural point of view, herbicide resistant crop varieties were introduced to get rid of weedy rice population. In reaction to that, resistant weedy rice populations, generated by geneflow from the cultigens were found (Gealy et al. 2003). This example for coevolution of a feral weed and the corresponding crop illustrates the arms race of two species over time. Since evolution and natural selection is a process that cannot be stopped, there is no reason to assume that the development of feral weeds will come to a halt in the future. With regard to the red queen hypothesis, for the weed management

strategies this means that the strategy of choice might be to be “one step ahead” of the weeds. Treatment strategies that only focus on the destruction of already existing weeds will only result in new selection pressures that will generate new weeds. Instead, the focus should be on the prediction of selection pressures that might lead to new weeds and avoid those, or change the strategies before the weed has the time to adapt. Studies like the one described in this thesis might be the first step in the direction of an evolutionary approach of weed management in the future. Similar considerations are made by Neve (2009) and Vigueira et al. (2013). By predicting the direction that weed evolution might take and then change practices to avoid adaptation, agriculture might take one step ahead of the weeds.

5 Appendix

A 5.1 Primer pairs used for the SSR analysis in this study. All primers were taken from Cao et al. (2006).

SSR	Chr	Fw 5‘-3‘	Rev 5‘-3‘
RM 11	7	tctcctcttcccccgatc	atagcgggcgaggcttag
RM 14	1	ccgaggagaggagttcgac	gtgccaatttcctcgaaaaa
RM 17	12	tgccctgttattttcttctctc	ggtgatcctttcccatttca
RM 19	12	caaaaacagagcagatgac	ctcaagatggacgccaaga
RM 21	11	acagtattccgtaggcacgg	gctccatgagggtggtagag
RM 44	8	acgggcaatccgaacaacc	tcgggaaaacctaccctacc
RM 55	3	ccgtcgccgtagtagagaag	tcccggttattttaaggcg
RM 84	1	taagggtccatccacaagatg	tgcaaatgcagctagagtac
RM 167	11	gatccagcgtgaggaacacgt	agtccgaccacaaggtgcgttgtc
RM 180	7	ctacatcggcttaggtgtagcaacacg	acttgctctacttgtggtgagggactg
RM 211	2	ccgatctcatcaaccaactg	cttcacgaggatctcaaagg
RM 212	1	ccactttcagctactaccag	cacccatttgtctctcattatg
RM 215	9	caaaatggagcagcaagagc	tgagcacctccttctctgtag
RM 219	9	cgtcggatgatgtaaagcct	catatcggcattcgcctg
RM 230	8	gccagaccgtggatgttc	caccgcagtcacttttcaag
RM 253	6	tccttcaagagtgcaaaacc	gcattgtcatgtcgaagcc
RM 276	6	ctcaacgttgacacctcgtg	tcctccatcgagcagtatca
RM 280	4	acacgatccactttgcgc	tgtgtcttgagcagccagg
RM 289	5	ttccatggcacacaagcc	ctgtgcacgaacttccaaag

A 5.2 Primers used for the NGS approach in this study. Names and sequences of all primers are provided. The %GC content, T_m and amplicon size are shown in separate columns. All primers designed for this work were calculated using the Primer3 software (Untergasser et al., 2010). Primers obtained from publications are cited in

Name	5' - 3' sequence	%GC	T_m	Amplicon size	Reference
qSH1_1_fw	ACCGGCAGTACTACCAGCAG	60	60.33		designed using Primer3
qSH1_1_rev	CTGATGATGCACGCTATGCT	50	60	1790 bp	designed using Primer3
qSH1_2_fw	CACGCAACCAGGTAAATAGAAA	40.91	59.18		designed using Primer3
qSH1_2_rev	GCTAAGCCCATTTCGTCATC	50	59.67	1008 bp	designed using Primer3
rc_015_for	CTGAAGGAAGTGATGACAACAAGACC	46	67		designed using Primer3
rc_015.2_rev	TTAAGTATGACTTATATTTTACATATTTGCAC	21	59	570 bp	Gross et al. (2010)
SD1_1_fw	CAACACAGCGCTCACTTCTC	55	59.78		Gross et al. (2010)
SD1_1_rev	AATCACGTCAGGTCGGTTTC	50	59.97	1181 bp	designed using Primer3
SD1_2_fw	GGGAATTGTTGTGTGTGCAG	50	60.01		designed using Primer3
SD1_2_rev	GTACAGCGGTAGGGTCCAAA	55	59.99	468 bp	designed using Primer3
SDR4_fw	GCCTTCTTAACCCCACCAC	57.98	59.4		designed using Primer3
SDR4_rev	TTAGAACCTGGCCTTGCATC	50	60.21	1276 bp	designed using Primer3
SH4_1_fw	CGCTCGGTTGATTAGGAGAG	55	59.97		designed using Primer3
SH4_1_rev	CACACTGCACGCAGCTTTAT	50	60.08	1357 bp	designed using Primer3
SH4_2_fw	ATTGCGAAATCACTCGCTTT	40	59.85		designed using Primer3
SH4_2_rev	TGCAGCCATTCCAAACAATA	40	60.07	861 bp	designed using Primer3
SHAT1_fw	TTGCAGATGAGCAACCTGAC	50	59.99		designed using Primer3
SHAT1_rev	GATGAATGCAGCGATCTTGA	45	59.91	1449 bp	designed using Primer3
VP1_1_fw	ATAAGTGGGCCCAGAGGAAA	50	60.82		designed using Primer3
VP1_1_rev	TCTTCTGGAGGTGGTGGTTC	55	60.09	1070 bp	designed using Primer3
VP1_2_fw	CTCACGAGCAACCGTGAGTA	55	60.05		designed using Primer3
VP1_2_rev	GCTCTGCTTCAGCACCTTCT	55	59.9	1251 bp	designed using Primer3
VP1_3_fw	GCATGCAGACGATTGACATC	50	60.24		designed using Primer3
VP1_3_rev	TAGCGCTACGATTCACATGC	50	60.01	1961 bp	designed using Primer3

6 Acknowledgements

This work was supported by the Landesgraduiertenförderung PhD fellowship of the state Baden-Württemberg for two years.

The research exchange with the University of Georgia (UGA) was funded by the "Auslandsstipendium" of the Karlsruhe House Of Young Scientists (KHYS) of the KIT.

I acknowledge the support of my work by Prof. Katrien Devos of the University of Georgia (UGA).

Further I would like to thank the Deutsche Forschungsgemeinschaft and Open Access Publishing Fund of Karlsruhe Institute of Technology for funding my publication "Microsatellite markers reveal multiple origins for Italian weedy rice".

7 References

Abdel-Aal, E.-S.M., and Hucl, P. (2003). Composition and stability of anthocyanins in blue-grained wheat. Journal of Agricultural and food Chemistry *51*, 2174-2180.

Andolfatto, P., Davison, D., Erezyilmaz, D., Hu, T.T., Mast, J., Sunayama-Morita, T., and Stern, D.L. (2011). Multiplexed shotgun genotyping for rapid and efficient genetic mapping. Genome research *21*, 610-617.

Asano, K., Takashi, T., Miura, K., Qian, Q., Kitano, H., Matsuoka, M., and Ashikari, M. (2007). Genetic and molecular analysis of utility of sd1 alleles in rice breeding. Breeding science *57*, 53-58.

Asano, K., Yamasaki, M., Takuno, S., Miura, K., Katagiri, S., Ito, T., Doi, K., Wu, J., Ebana, K., and Matsumoto, T. (2011). Artificial selection for a green revolution gene during japonica rice domestication. Proceedings of the National Academy of Sciences *108*, 11034-11039.

Ashikari, M., Sasaki, A., Ueguchi-Tanaka, M., Itoh, H., Nishimura, A., Datta, S., Ishiyama, K., Saito, T., Kobayashi, M., and Khush, G.S. (2002). Loss-of-function of a rice gibberellin biosynthetic gene, GA20 oxidase (GA20ox-2), led to the rice 'green revolution'. Breeding science *52*, 143-150.

Biroli, G. 1807. Del riso. Tipografia Giovanni Silvestri, Milan, Italy.

Brooks, S.A., Yan, W., Jackson, A.K., and Deren, C.W. (2008). A natural mutation in rc reverts white-rice-pericarp to red and results in a new, dominant, wild-type allele: Rc-g. Theoretical and Applied Genetics *117*, 575-580.

Butler, L. G., Price, M. L., and Brotherton, J. E. (1982). Vanillin assay for proanthocyanidins (condensed tannins): modification of the solvent for estimation of the degree of polymerization. Journal of Agricultural and Food Chemistry *30*, 1087-1089.

Cao, Q., Lu, B.-R., Xia, H., Rong, J., Sala, F., Spada, A., and Grassi, F. (2006). Genetic diversity and origin of weedy rice (Oryza sativa f. spontanea) populations found in north-eastern China revealed by simple sequence repeat (SSR) markers. Annals of botany *98*, 1241-1252.

Carney, Judith. (2004). With Grains in Her Hair: Rice in Colonial Brazil. Slavery and Abolition, Vol. 25, No. 1, April 2004, p. 1-27

Catala, M., Eleftherohorinos, I., Martins, L., Vidotto, F., Ferrero, A., Dhima, K., Feougier, G., and Mouret, J. (2002). Red rice control in Southern Europe with pre-and postemergence applied herbicides.

Cavalli-Sforza, L.L., and Edwards, A.W. (1967). Phylogenetic analysis. Models and estimation procedures. American journal of human genetics *19*, 233.

Chen, L. J., Lee, D. S., Song, Z. P., Suh, H. S., and LU, B. R. (2004). Gene flow from cultivated rice (Oryza sativa) to its weedy and wild relatives. Annals of botany *93*, 67-73.

Chen, S., S. Karudapuram, C. Cesena, L. Joe, A. Wheaton, and C. Heiner. (2011). Study of the new and improved size standard in fragment analysis. Applied Biosystems, Foster City, CA.

Chiwocha, S.D., Dixon, K.W., Flematti, G.R., Ghisalberti, E.L., Merritt, D.J., Nelson, D.C., Riseborough, J.-A.M., Smith, S.M., and Stevens, J.C. (2009). Karrikins: a new family of plant growth regulators in smoke. Plant Science *177*, 252-256.

Clements, D.R., DiTommaso, A., Jordan, N., Booth, B.D., Cardina, J., Doohan, D., Mohler, C.L., Murphy, S.D., and Swanton, C.J. (2004). Adaptability of plants invading North American cropland. Agriculture, ecosystems & environment *104*, 379-398.

Dalrymple, D.G. (1985). The development and adoption of high-yielding varieties of wheat and rice in developing countries. American Journal of Agricultural Economics *67*, 1067-1073.

De Wet, J., and Harlan, J. R. (1975). Weeds and domesticates: evolution in the man-made habitat. Economic Botany *29*, 99-108.

Dekker J. (2011). Evolutionary ecology of weeds. http://agron-www.agron.iastate.edu/~weeds/ag517/517Course/EEWbook.html

Dlugosch, K., and Parker, I. (2008). Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. Molecular Ecology *17*, 431-449.

Doyle, J.J. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem bull *19*, 11-15.

Earl, D.A. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources *4*, 359-361.

Ellstrand, N.C., Heredia, S.M., Leak-Garcia, J.A., Heraty, J.M., Burger, J.C., Yao, L., Nohzadeh-Malakshah, S., and Ridley, C.E. (2010). Crops gone wild: evolution of weeds and invasives from domesticated ancestors. Evolutionary Applications *3*, 494-504.

Estorninos Jr, L. E., Gealy, D. R., Gbur, E. E., Talbert, R. E., and McClelland, M. R. (2009). Rice and red rice interference. II. Rice response to population densities of three red rice (Oryza sativa) ecotypes.

Faivre-Rampant, O., Bruschi, G., Abbruscato, P., Cavigiolo, S., Picco, A.M., Borgo, L., Lupotto, E., and Piffanelli, P. (2011). Assessment of genetic diversity in Italian rice germplasm related to agronomic traits and blast resistance (Magnaporthe oryzae). Molecular Breeding *27*, 233-246.

Fan, J., Niu, X., Wang, Y., Ren, G., Zhuo, T., Yang, Y., Lu, B.-R., and Liu, Y. (2007). Short, direct repeats (SDRs)-mediated post-transcriptional processing of a transcription factor gene OsVP1 in rice (Oryza sativa). Journal of experimental botany *58*, 3811-3817.

Flematti, G.R., Scaffidi, A., Dixon, K.W., Smith, S.M., and Ghisalberti, E.L. (2011). Production of the seed germination stimulant karrikinolide from combustion of simple carbohydrates. Journal of agricultural and food chemistry *59*, 1195-1198.

Fogliatto, S., Vidotto, F., and Ferrero, A. (2010). Effects of winter flooding on weedy rice (Oryza sativa L.). Crop Protection *29*, 1232-1240.

Fogliatto, S., Vidotto, F., and Ferrero, A. (2011). Germination of weedy rice in response to field conditions during winter. Weed Technology *25*, 252-261.

Fogliatto, S., Vidotto, F., and Ferrero, A. (2012). Morphological characterisation of Italian weedy rice (Oryza sativa) populations. Weed Research *52*, 60-69.

Furukawa, T., Maekawa, M., Oki, T., Suda, I., Iida, S., Shimada, H., Takamure, I., and Kadowaki, K.i. (2007). The Rc and Rd genes are involved in proanthocyanidin synthesis in rice pericarp. The Plant Journal *49*, 91-102.

Gealy DR, (2005). Gene movement between rice (Oryza sativa) and weedy rice(Oryza sativa) – a U.S. temperate rice perspective, in Crop Ferality and Volunteerism, ed. by Gressel J. CRC Press, Boca Raton, FL, pp. 323–354

Gealy, D. R., Mitten, D. H., and Rutger, J. N. (2003). Gene flow between red rice (Oryza sativa) and herbicide-resistant rice (O. sativa): implications for weed management 1. Weed technology *17*, 627-645.

Gianinetti, A., and Cohn, M. A. (2008). Seed dormancy in red rice. XIII: Interaction of dry-afterripening and hydration temperature. Seed science research *18*, 151-160.

Gressel, J. (2005). Introduction-the challenges of ferality. Crop ferality and volunteerism, 1-7.

Gressel, J., and Valverde, B. E. (2009). A strategy to provide long-term control of weedy rice while mitigating herbicide resistance transgene flow, and its potential use for other crops with related weeds. Pest management science *65*, 723-731.

Grimm, A., Fogliatto, S., Nick, P., Ferrero, A., and Vidotto, F. (2013). Microsatellite markers reveal multiple origins for Italian weedy rice. Ecology and evolution *3*, 4786-4798.

Gross, B.L., Steffen, F.T., and Olsen, K.M. (2010). The molecular basis of white pericarps in African domesticated rice: novel mutations at the Rc gene. Journal of evolutionary biology *23*, 2747-2753.

Grundbacher, F. (1963). The physiological function of the cereal awn. The Botanical Review *29*, 366-381.

Gulick, P., Lee, D., Lupotto, E., and Powell, W. (2009). G-string slippage turns white rice red. Genome *52*, 490-493.

Hargrove, T.R., and Cabanilla, V.L. (1979). The impact of semidwarf varieties on Asian rice-breeding programs. BioScience, 731-735.

Ishikawa, R., Toki, N., Imai, K., Sato, Y., Yamagishi, H., Shimamoto, Y., Ueno, K., Morishima, H., and Sato, T. (2005). Origin of weedy rice grown in Bhutan and the force of genetic diversity. Genetic Resources and Crop Evolution *52*, 395-403.

Ji, H.-S., Chu, S.-H., Jiang, W., Cho, Y.-I., Hahn, J.-H., Eun, M.-Y., McCouch, S. R., and Koh, H.-J. (2006). Characterization and mapping of a shattering mutant in rice that corresponds to a block of domestication genes. Genetics *173*, 995-1005.

Ji, H.-S., Chu, S.-H., Jiang, W., Cho, Y.-I., Hahn, J.-H., Eun, M.-Y., McCouch, S.R., and Koh, H.-J. (2006). Characterization and mapping of a shattering mutant in rice that corresponds to a block of domestication genes. Genetics *173*, 995-1005.

Jiang, Z., Xia, H., Basso, B., and Lu, B.-R. (2012). Introgression from cultivated rice influences genetic differentiation of weedy rice populations at a local spatial scale. Theoretical and Applied Genetics *124*, 309-322.

Kawahara, Y., de la Bastide, M., Hamilton, J.P., Kanamori, H., McCombie, W.R., Ouyang, S., Schwartz, D.C., Tanaka, T., Wu, J., and Zhou, S. (2013). Improvement of the Oryza sativa Nipponbare reference genome using next generation sequence and optical map data. Rice *6*, 1-10.

Kawasaki, A., Imai, K., Ushiki, J., Ishii, T., and Ishikawa, R. (2009). Molecular constitution of weedy rice (Oryza sativa L.) found in Okayama prefecture, Japan. Breeding science *59*, 229-236.

Khush, G.S. (1999). Green revolution: preparing for the 21st century. Genome *42*, 646-655.

Kim, M.-K., Kim, H.-a., Koh, K., Kim, H.-S., Lee, Y.S., and Kim, Y.H. (2008). Identification and quantification of anthocyanin pigments in colored rice. Nutrition research and practice *2*, 46-49.

Koressaar, T., and Remm, M. (2007). Enhancements and modifications of primer design program Primer3. Bioinformatics *23*, 1289-1291.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nature methods *9*, 357-359.

Lazarus, S.A., Adamson, G.E., Hammerstone, J.F., and Schmitz, H.H. (1999). High-performance liquid chromatography/mass spectrometry analysis of proanthocyanidins in foods and beverages. Journal of Agricultural and Food Chemistry *47*, 3693-3701.

Li, C., Zhou, A., and Sang, T. (2006). Rice domestication by reducing shattering. science *311*, 1936-1939.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). The sequence alignment/map format and SAMtools. Bioinformatics *25*, 2078-2079.

Lin, Z., Griffith, M.E., Li, X., Zhu, Z., Tan, L., Fu, Y., Zhang, W., Wang, X., Xie, D., and Sun, C. (2007). Origin of seed shattering in rice (Oryza sativa L.). Planta *226*, 11-20.

Londo, J. P., Chiang, Y.-C., Hung, K.-H., Chiang, T.-Y., and Schaal, B. A. (2006). Phylogeography of Asian wild rice, Oryza rufipogon, reveals multiple independent domestications of cultivated rice, Oryza sativa. Proceedings of the National Academy of Sciences *103*, 9578-9583.

Londo, J., and Schaal, B. (2007). Origins and population genetics of weedy red rice in the USA. Molecular Ecology *16*, 4523-4535.

Londo, J.P., Chiang, Y.-C., Hung, K.-H., Chiang, T.-Y., and Schaal, B.A. (2006). Phylogeography of Asian wild rice, Oryza rufipogon, reveals multiple independent domestications of cultivated rice, Oryza sativa. Proceedings of the National Academy of Sciences *103*, 9578-9583.

Magrane, M. (2011). UniProt Knowledgebase: a hub of integrated protein data. Database: the journal of biological databases and curation *2011*.

Minch, E., Ruiz-Linares, A., Goldstein, D., Feldman, M., and Cavalli-Sforza, L. (1995). Microsat (version 1.5): A computer program for calculating various statistics on microsatellite allele data. WWW: http://lotka. stanford. edu/microsat. html.

Nelson, D.C., Riseborough, J.-A., Flematti, G.R., Stevens, J., Ghisalberti, E.L., Dixon, K.W., and Smith, S.M. (2009). Karrikins discovered in smoke trigger Arabidopsis seed germination by a mechanism requiring gibberellic acid synthesis and light. Plant Physiology *149*, 863-873.

Neve, P., Vila-Aiub, M., and Roux, F. (2009). Evolutionary-thinking in agricultural weed management. New Phytologist *184*, 783-793.

Oerke, E.-C. (2006). Crop losses to pests. The Journal of Agricultural Science *144*, 31-43.

Price, M.L., Hagerman, A.E., and Butler, L.G. (1980). Tannin content of cowpeas, chickpeas, pigeon peas, and mung beans. Journal of Agricultural and Food Chemistry *28*, 459-461.

Pritchard, J.K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. Genetics *155*, 945-959.

Purugganan, M.D., and Fuller, D.Q. (2009). The nature of selection during plant domestication. Nature *457*, 843-848.

Raymond, M., and Rousset, F. (1995). GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. Journal of heredity *86*, 248-249.

Reagon, M., Thurber, C.S., Gross, B.L., Olsen, K.M., Jia, Y., and Caicedo, A.L. (2010). Genomic patterns of nucleotide diversity in divergent populations of US weedy rice. BMC evolutionary biology *10*, 180.

Sakai, H., Lee, S.S., Tanaka, T., Numa, H., Kim, J., Kawahara, Y., Wakimoto, H., Yang, C.-c., Iwamoto, M., and Abe, T. (2013). Rice annotation project database (RAP-DB): an integrative and interactive database for rice genomics. Plant and Cell Physiology *54*, e6-e6.

Shivrain, V. K., Burgos, N. R., Anders, M. M., Rajguru, S. N., Moore, J., and Sales, M. A. (2007). Gene flow between Clearfield™ rice and red rice. Crop Protection *26*, 349-356.

Shivrain, V., Burgos, N., Agrama, H., Lawton-Rauh, A., Lu, B., Sales, M., Boyett, V., Gealy, D., and Moldenhauer, K. (2010). Genetic diversity of weedy red rice (Oryza sativa) in Arkansas, USA. Weed Research *50*, 289-302.

Spada, A., Mantegazza, R., Biloni, M., Caporali, E., and Sala, F. (2004). Italian rice varieties: historical data, molecular markers and pedigrees to reveal their genetic relationships. Plant breeding *123*, 105-111.

Sugimoto, K., Takeuchi, Y., Ebana, K., Miyao, A., Hirochika, H., Hara, N., Ishiyama, K., Kobayashi, M., Ban, Y., and Hattori, T. (2010). Molecular cloning of Sdr4, a regulator involved in seed dormancy and domestication of rice. Proceedings of the National Academy of Sciences *107*, 5792-5797.

Sweeney, M. T., Thomson, M. J., Pfeil, B. E., and McCouch, S. (2006). Caught red-handed: Rc encodes a basic helix-loop-helix protein conditioning red pericarp in rice. The Plant Cell Online *18*, 283-294.

Sweeney, M.T., Thomson, M.J., Cho, Y.G., Park, Y.J., Williamson, S.H., Bustamante, C.D., and McCouch, S.R. (2007). Global dissemination of a single mutation conferring white pericarp in rice. PLoS Genetics *3*, e133.

Tamura, K., Dudley, J., Nei, M., and Kumar, S. (2007). MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Molecular biology and evolution *24*, 1596-1599.

Thorvaldsdóttir, H., Robinson, J.T., and Mesirov, J.P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Briefings in bioinformatics *14*, 178-192.

Thurber, C. S., Reagon, M., Gross, B. L., Olsen, K. M., Jia, Y., and Caicedo, A. L. (2010). Molecular evolution of shattering loci in US weedy rice. Molecular ecology *19*, 3271-3284.

Thurber, C.S., Hepler, P.K., and Caicedo, A.L. (2011). Timing is everything: early degradation of abscission layer is associated with increased seed shattering in US weedy rice. BMC plant biology *11*, 14.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., and Rozen, S.G. (2012). Primer3—new capabilities and interfaces. Nucleic acids research *40*, e115-e115.

Van Valen, L. (1974). Molecular evolution as predicted by natural selection. Journal of molecular evolution *3*, 89-101.

Vigueira, C., Olsen, K., and Caicedo, A. (2013). The red queen in the corn: agricultural weeds as models of rapid adaptive evolution. Heredity *110*, 303-311.

Voss, N., Eckstein, R.L., and Durka, W. (2012). Range expansion of a selfing polyploid plant despite widespread genetic uniformity. Annals of botany *110*, 585-593.

Wang, W., Xia, H., Yang, X., Xu, T., Si, H.J., Cai, X.X., Wang, F., Su, J., Snow, A.A., and Lu, B.R. (2013). A novel 5-enolpyruvoylshikimate-3-phosphate (EPSP) synthase transgene for glyphosate resistance stimulates growth and fecundity in weedy rice (Oryza sativa) without herbicide. New Phytologist.

Wright, S. (1965). The interpretation of population structure by F-statistics with special regard to systems of mating. Evolution, 395-420.

Zhou, Y., Lu, D., Li, C., Luo, J., Zhu, B.-F., Zhu, J., Shangguan, Y., Wang, Z., Sang, T., and Zhou, B. (2012). Genetic control of seed shattering in rice by the APETALA2 transcription factor SHATTERING ABORTION1. The Plant Cell Online *24*, 1034-1048.

Zhu, Y., Ellstrand, N.C., and Lu, B.R. (2012). Sequence polymorphisms in wild, weedy, and cultivated rice suggest seed-shattering locus sh4 played a minor role in Asian rice domestication. Ecology and evolution *2*, 2106-2113.

Zuo, J., Zhang, L., Song, X., Dai, W., and Qiang, S. (2011). Innate factors causing differences in gene flow frequency from transgenic rice to different weedy rice biotypes. Pest management science *67*, 677-690.

Publikationen

Teile dieser Arbeit wurden publiziert in:

“Microsatellite markers reveal multiple origins for Italian weedy rice” (2013)
Annabelle Grimm, Silvia Fogliatto, Peter Nick, Aldo Ferrero & Francesco Vidotto. Ecology and Evolution DOI 10.1002/ece3.848