DISS. ETH NO. 19402

Unraveling genetic factors controlling the restoration of fertility of C-type cytoplasmic male sterility in maize

A dissertation submitted to
ETH ZURICH

For the degree of
Doctor of Sciences

Presented by
Susanne Kohls
Master of Science, Universität Hohenheim

Born February 28, 1982
Citizen of Germany

Accepted on the recommendation of
Prof. Dr. Peter Stamp, examiner
Dr. Rainer Messmer, co-examiner
Dr. Carsten Knaak, co-examiner

Zurich, 2010

Contents

List of abbreviations

ASI	anthesis-silking interval
a.s.l.	above sea level
ATP	adenosine triphosphate
BAC	bacterial artificial chromosome
BLAST	basic local alignment search tool
BLUP	best linear unbiased predictor
bp	base pair
BSA	bulk segregant analysis
CAD	Cadenazzo
CIM	composite interval mapping
CTAB	cetyltrimethylammonium bromide
cM	centiMorgan
CMS	cytoplasmic male sterility
DEL	Delley
DNA	deoxyribonucleic acid
ESH	Eschikon
EST	expressed sequence tag
e-value	expectation value
dNTP	deoxynucleotide triphosphate
h^2	heritability
h^2_w	within-location heritability
indel	insertion/ deletion
K	potassium
Kb	kilobase pair
LOD	log likelihood of odds
Mb	megabase pair
mf	male-fertile
ms	male-sterile
N	nitrogen

npartR^2	normed partial correlation coefficient of determination
OPA	oligo pool assay
ORF	open reading frame
P	phosphorus
PCR	polymerase chain reaction
pmf	partially male-fertile
PPR protein	pentatricopeptide-repeat protein
rf	Restorer-of-fertility
QTL	quantitative trait locus
R^2	coefficient of determination
r	coefficient of correlation
RFLP	restriction fragment length polymorphism
RNA	ribonucleic acid
SNP	single nucleotide polymorphism
SSR	simple sequence repeat
T_m	melting temperature
TBE	Tris/Borate/EDTA
USDA	United States Department of Agriculture
UTR	untranslated region

Summary

Hybrid breeding in maize and many other crops led to large increases in productivity due to hybrid vigor and greater crop uniformity. The production of hybrid seeds requires a directional cross between a seed parent (female parent) and a pollen donor (male parent). The production of pollen by the female parent has to prevented to avoid female plants pollinating themselves. This can be achieved by manual emasculations or a genetic mechanism, among which cytoplasmic male sterility (CMS) is the most important non-transgenic mechanism in maize. Male sterility is induced by mutations of the mitochondrial DNA and is, therefore, maternally inherited. The trait can be compensated for by nuclear *restorer-of-fertility* genes (restorer genes), which restore male fertility in the first generation after crossing a CMS and a non-CMS plant carrying the restorer allele(s). This ensures shed of pollen and seed set in the hybrid. All restorer genes cloned thus far encode proteins of the pentatricopepetide-repeat family (PPR), except *Rf2* of maize, which encodes an aldehyde dehydrogenase. Three major types of male sterility inducing cytoplasms have been defined in maize (T, S and C). Breeders have made use of CMS since the 1950s to facilitate the production of hybrid seeds and to minimize costs. However, CMS-based systems were abandoned in the early 1970s after an epidemic of Southern Corn Leaf Blight in the US Corn Belt, which was caused by *Bipolaris maydis* race T, especially virulent on maize with T cytoplasm. European and US breeders are now becoming interested again in CMS as a convenient and cost-effective tool. They implement mainly C-CMS.

However, C-CMS has a number of shortcomings, which limit the use in the production of hybrid seeds. The system of restoration is complex and largely unknown. Restorer genes are very abundant throughout most gene pools, limiting the selection of maternal lines. Furthermore, partial restoration often occurs in maternal lines, which is variably expressed in different environments. Partial restorations can lead to self pollinations of the maternal parent and, thus, to impure hybrid seeds.

New inbred lines have usually been classified as C-CMS restorers or non-restorers

by means of test crosses. When lines classified as restorers are to be used as seed parents in a CMS-based system, they have to be converted to non-restorers by means of backcrosses. However, this process is time-consuming and not always successful. If molecular markers tightly linked to restorer loci were identified, this would accelerate the conversions and reduce the number of test crosses to be phenotyped. Furthermore, marker-assisted selection could improve the counter-selection of partial restorer alleles in maternal inbred lines, because partial restoration might not be expressed every year or at every location, depending on the environmental conditions. Therefore, the objectives of this study were to fine-map the major restorer gene *Rf4* and to locate QTLs involved in the partial restoration of male fertility.

The fine-mapping of *Rf4* was performed in an F_2 population (n=1,317) derived from a cross between the C-CMS inbred line B37C and the restorer inbred line K55. Male fertility was rated according to the number of anthers, the quality of the anthers and the anthesis-silking interval. These traits were combined to a fertility index to fine-map *Rf4*. Molecular markers tightly linked with the gene were developed by SSR (simple sequence repeat) mining and comparative sequencing of the parental lines. *Rf4* was mapped by selectively genotyping 40 % of the F_2 individuals showing an extreme phenotype with respect to male fertility. Candidate genes were called by annotating the B73 reference sequence flanking *Rf4*. QTLs for partial restoration were revealed by means of a Bulk Segregant Analysis and a subsequent QTL study. QTL mapping was carried out in a subset of the F_2 population used for the fine-mapping of *Rf4* (n=180), which was selected to carry the non-restoring allele *rf4*. This eliminated the masking effect of *Rf4* on the partial restorer genes. Phenotypic values were obtained from F_2BC_1 progenies grown at three locations in Switzerland.

The segregation of male fertility in F_2 suggested that it was governed by one fully restoring dominant gene (*Rf4*) and several further partially restoring genes. *Rf4* mapped 0.3 cM upstream of the closest markers in the chromosomal bin 8.00, which corresponds to a physical interval of 520 Kb comprising 12 predicted genes. *Rf4* might represent a new type of restorer gene, because none of the predicted genes is homologous to previously cloned restorer genes, which mainly encode proteins of the pentatricopeptide-repeat family. Six new PCR-based markers tightly linked to *Rf4* are now available for marker-assisted selection. However, testing these markers in a

set of restorer and non-restorer lines revealed that the *Rf4* allele alone is insufficient to predict the restorer property of an inbred line.

Partial restoration was inherited like an oligogenic trait. Major QTLs with high LOD scores and a consistent expression across environments were revealed in the chromosomal bins 2.09, 3.06 and 7.03. The low environmental influence on the expression of QTLs contrasts with the results of previous studies. The three major QTL regions collocated with other restorer genes of maize.

The maternal parent was an important source of restorer alleles. The presence of restorer alleles was indicated by an incomplete sterility of B37C and confirmed by QTLs in bins 2.09, 6.04 and 7.03, the restorer alleles of which originated from the maternal parent. Furthermore, B37C carries at least one complementary restorer gene at a locus other than *Rf4*, because F_1 progenies of crosses between two different C-CMS lines and the male-fertile version of B37 were fully restored.

In conclusion, a marker-assisted selection of restorer genes is promising. Although the restorer status of an inbred line can currently not be deduced based solely on the *Rf4* allele, marker-assisted selection will be useful if the abundance of *Rf4* in a certain gene pool is known through breeding experience. Counter-selection of partial restorer alleles by means of markers seems to be feasible, if the importance of the major QTL regions can be proven in different germplasm and under a broader range of environmental conditions. The selection and mapping of restorer genes has to consider the maternal parent as an important source of restorer alleles. Complementary interactions between restorer genes necessitate a selection even in CMS lines, which express a stable sterility. The cloning of *Rf4* promises new insights into the biology of CMS, because *Rf4* hypothetically encodes a new type of restorer protein. Studying the genetic structure of clusters of restorer genes in maize could result in a better understanding of the evolution and functioning of restorer genes in general. The long arm of chromosome 2 is in this respect particularly interesting, because it harbors restorer genes for all three major CMS systems in maize.

Zusammenfassung

Die Hybridzüchtung führte insbesondere durch die systematische Nutzung der Heterosis zu einer Steigerung und Stabilisierung der Ertragsleistung bei Mais und vielen anderen Kulturarten. Die Produktion des Hybridsaatguts erfordert eine kontrollierte Kreuzung zwischen einem Saatelter (mütterliche Inzuchtlinie) und einem Pollendonor (väterliche Inzuchtlinie). Dabei muss die Pollenproduktion des Saatelters unterbunden werden, um eine Selbstbestäubung zu verhindern. Dies kann durch manuelle Kastration oder aber mit Hilfe genetischer Systeme erreicht werden, unter denen die cytoplasmatisch-männliche Sterilität (CMS) das am weitesten verbreitete System bei Mais ist. Die Pollensterilität wird durch Mutationen der mitochondrialen DNA verursacht und vererbt sich deshalb maternal. Jedoch kann die Sterilität in CMS Pflanzen durch nukleäre *Restorer-of-fertility* Gene (Restorergene) aufgehoben werden. Eine Kreuzung zwischen einer CMS Pflanze und einer Pflanze, die die entsprechenden Restorergene trägt, ist daher fertil. Auf diese Weise wird die Pollenschüttung der Hybride sichergestellt. Die meisten Restorergene, die bisher kloniert worden sind, kodieren Proteine der Pentatricopeptid-repeat (PPR) Familie. Eine Ausnahme ist das *Rf2* Gen des Mais, welches eine Aldehyd Dehydrogenase kodiert. Bei Mais wurden drei CMS-Hauptgruppen (T, S und C) definiert, die sich anhand ihrer mitochondrialen Ursache sowie ihrer spezifischen Restorergene unterscheiden lassen. Schon in den fünfziger Jahren wurde T-CMS in den USA grossflächig für die Hybridsaatgutproduktion eingesetzt. Jedoch wurde nach einer verheerenden Blattdürre-Epidemie, die von einer T-CMS spezifischen Rasse des Schaderregers *Bipolaris maydis* ausgelöst wurde, wieder überwiegend manuell entfahnt. Erst seit einigen Jahren steigt das Interesse der Züchter wieder, CMS als kosteneffizientes Werkzeug der Hybridisierung einzusetzen, allerdings ist heute C-CMS die am häufigsten genutzte Form.

Das C-CMS System weist jedoch einige Schwächen auf, die den Einsatz in der Saatguterzeugung limitieren. Zum einen ist die Restauration genetisch komplex, und die beteiligten Gene sind weitgehend unbekannt. Zum anderen sind Restorergene

für C-CMS auch in mütterlichen Genpools stark verbreitet und erschweren daher die Selektion der Saatelterlinien. Ausserdem treten häufig partielle Restaurationen in den Mutterlinien auf, die zu Selbstungen während der Hybridsaatgutproduktion führen und so die Reinheit des Saatguts gefährden können.

Bisher wurden neue Inzuchtlinien meist aufgrund von Testkreuzungen als Restorer- oder Nichtrestorerlinien klassifiziert. Soll eine Linie, die Restorergene trägt, als CMS Saatelter eingesetzt werden, muss sie über mehrfache Rückkreuzungen in eine Nichtrestorerlinie konvertiert werden. Dieser Prozess ist zeitaufwändig und nicht immer erfolgreich. Wenn molekulare Marker identifiziert werden könnten, die eng mit Restorerloci gekoppelt sind, könnte dies die Rückkreuzungen beschleunigen und die Anzahl der Testkreuzungen reduzieren. Ausserdem könnten molekulare Marker die Selektion gegen partielle Restorergene beschleunigen, da deren Effekt in Abhängigkeit von den Umweltbedingungen nicht jedes Jahr oder an jedem Ort sichtbar ist. Das Ziel dieser Arbeit war es daher, eines der Hauptrestorergene für C-CMS, *Rf4*, feinzukartieren und QTLs zu lokalisieren, die in die partielle Restauration involviert sind.

Die genetische Kartierung des *Rf4* Gens wurde in einer F_2 Population (n=1,317) durchgeführt, die von einer Kreuzung der C-CMS Linie B37C und der Restorerlinie K55 abgeleitet worden war. Die männliche Fertilität wurde anhand der Antherenqualität, der Antherenmenge und des Blühintervalls bonitiert, die im Falle der *Rf4* Kartierung zu einem Fertilitätsindex kombiniert wurden. Durch vergleichende Sequenzierung der Elternlinien wurden sechs PCR basierte Marker entwickelt, die eng mit dem *Rf4* Gen gekoppelt sind. Für die Erstellung der genetischen Karte wurden 40 % der F_2 Individuen genotypisiert, die in Bezug auf die männliche Fertilität einen extremen Phänotyp aufwiesen. Die Sequenz des B73 Referenzgenoms diente dazu, Kandidatengene im flankierenden Sequenzbereich um *Rf4* zu definieren. Alle neuen Marker wurden in einem Satz Restorer- und Nichtrestorerlinien getestet.

Die Vererbung der partiellen Restauration wurde mit Hilfe einer QTL Studie analysiert, die in einer Fraktion (n=180) der beschriebenen F_2 Population durchgeführt wurde. Diese Fraktion umfasste ausschliesslich Pflanzen, die das nicht-restaurierende Allel (*rf4rf4*) trugen, so dass partielle Restorerloci ohne den maskierenden Effekt des *Rf4* Allels kartiert werden konnten. Die Phänotypisierung der männlichen Fertilität wurde in F_2BC_1 Nachkommenschaften an drei Orten in der Schweiz durchgeführt.

Die Segregation der männlichen Fertilität in F_2 deutete darauf hin, dass ein dominantes, vollständig restaurierendes Gen (*Rf4*) sowie mehrere partiell restaurierende Gene spalteten. *Rf4* kartierte 0.3 cM oberhalb des nächstgelegenen Markers im chromosomalen Abschnitt 8.00. Die Distanz zwischen diesem Marker und dem 5' Telomer des Chromosom 8 entspricht einem physikalischen Intervall von 520 Kb, welches 12 vorhergesagte Gene enthält. Da keines dieser Gene Homologie mit den vormals klonierten Restorergenen aufweist, repräsentiert *Rf4* möglicherweise einen neuen Restorergentypus. Das *Rf4* Allel verschiedener Restorer- und Nichtrestorerlinien entsprach meist nicht dem jeweiligen Restorerphänotyp, da es ausser *Rf4* offenbar weitere Restorergene für C-CMS gibt.

In den chromosomalen Abschnitten 2.09, 3.06 und 7.03 wurden Major-QTLs für partielle Restauration detektiert, die alle sehr hohe LOD Werte und eine stabile Expression in den Umwelten aufwiesen. Dies deutet darauf hin, dass die partielle Restauration ein oligogen vererbtes Merkmal ist. Alle drei Major-QTLs kolokalisierten mit anderen Restorergenen des Mais. Eine Gruppierung in bestimmten genomischen Bereichen scheint für Restorergene charakteristisch zu sein, da dies auch in anderen Pflanzenarten beobachtet wurde.

Der mütterliche Elter erwies sich in allen Experimenten als wichtige Quelle für Restorerallele. Zum einen war die Sterilität nie vollständig, und zum anderen trägt B37C mindestens ein komplementär wirkendes Restorergen, weil Kreuzungen zwischen der fertilen Variante von B37 (B37N) und zwei C-CMS Linien vollständig restauriert waren. Die Präsenz von Restorergenen wurde durch QTLs in den Abschnitten 2.09, 6.04 und 7.03 bestätigt, deren Restorerallele vom mütterlichen Elter stammten.

Insgesamt zeigte die Studie, dass eine Marker-gestützte Selektion von Restorergenen vielversprechend ist. Obwohl die Restorereigenschaft einer Linie derzeit nicht allein aufgrund des *Rf4* Allels vorhergesagt werden kann, so kann der Markereinsatz nützlich sein, wenn die Frequenz von *Rf4* im Zuchtmaterial bekannt ist. Eine Selektion gegen die Restorerallele der Major-QTLs wäre aufgrund dieser Studie sinnvoll, allerdings muss die Bedeutung der Major-QTLs in anderem Pflanzenmaterial und unter anderen Umweltbedingungen verifiziert werden. Die Selektion muss den mütterlichen Elter als Träger von Restorergenen berücksichtigen, und zwar selbst dann, wenn die Sterilität einer CMS Linie vollständig ist, da Komplementationen

auftreten können. Die Klonierung des *Rf4* Gens verspricht neue Erkenntisse über die Biologie der Restauration, da es möglicherweise einen neuen Restorergentyp kodiert. Ausserdem könnte eine Untersuchung der Restorergen-Cluster das Verständnis für die Evolution und Funktionsweise von Restorergenen vertiefen. In diesem Zusammenhang ist der lange Arm des Chromosoms 2 besonders interessant, weil die Region Restorergene für alle drei CMS Gruppen des Mais enthält.

1 General introduction

1.1 Importance of cytoplasmic male sterility (CMS)

Hybrid varieties of crops have led to large increases in productivity due to hybrid vigor and greater crop uniformity. The production of hybrid seeds requires a directional cross between a seed parent (female parent) and a pollen donor (male parent). The production of pollen by the female parent has to be prevented to avoid female plants pollinating themselves. This can be achieved by manual emasculations or a genetic sterility system, such as genic male sterility, self incompatibility, genetically engineered systems or cytoplasmic male sterility (CMS) (Chase et al. 2010). Among the genetic systems, CMS is the most important non-transgenic system of hybridization in maize. Male sterility is induced by mutations of the mitochondrial DNA disrupting the development of functional pollen (Schnable and Wise 1998). Female fertility is usually not affected (Laughnan and Gabay-Laughnan 1983), and plants can set seeds if viable pollen is provided. The maternally inherited trait can be compensated for by nuclear *restorer-of-fertility* genes (restorer genes), which restore fertility in the first generation progeny of a cross between a CMS and a non-CMS plant carrying the restorer allele(s). The production of hybrid seeds requires CMS, maintainer and restorer inbred lines. CMS plants carry the genotype *rfrf* and a CMS cytoplasm. Male fertility is maintained when such CMS plants are crossed with the maintainer plant (genotype *rfrf* and N-cytoplasm) as the pollen donor. When a restorer plant (genotype *RfRf*) is crossed with a CMS plant it restores the fertility of the CMS plant.

In the case of maize, the cost of seed production by means of CMS can be reduced by approximately 2 % (W. Schmidt, pers. comm., 2008). There are no official data on the effective use of CMS in today's production of hybrid seed; but most maize breeders in Europe develop CMS inbred lines for commercially important varieties, which are multiplied on areas larger than 150 ha (K.-H. Camp, pers. comm., 2010).

It is assumed that the rate of commercial hybrids, produced by means of CMS, will continue to increase (Munsch et al. 2008). There are three major types of cytoplasm that induce male sterility in maize: T cytoplasm (or **T**exas cytoplasm) (Rogers and Edwardson 1952), S cytoplasm (U**S**DA) (Jones and Manglesdorf 1957) and C cytoplasm (**C**harrua) (Beckett 1971). These three types of CMS are caused by different mitochondrial mutations and can be distinguished according to the specific nuclear restorer genes, which counteract the CMS trait. In the 1950s and 1960s, T-CMS was used extensively in the USA because of its reliability and the ease of finding suitable restorer genotypes although it was known that T-CMS confers specific susceptibility to the fungal pathogen *Bipolaris maydis* (Ullstrup 1972). A devastating epidemic of Southern Corn Leaf Blight in the early 1970s, caused by a new race (race T) of this pathogen, led to a yield loss of 50 %. As a consequence, seed producers went back to using normal male-fertile cytoplasm and/or S- and C-CMS. C-CMS is today the most important source of CMS, although a new strain of *B. maydis* (race C) was discovered in China, which affects C-CMS plants more severely than plants carrying a normal cytoplasm (Wei et al. 1988).

As well as hybrid seed production, CMS also seems to be a feasible solution to the problem of the coexistence of transgenic and non-transgenic crops or native plants. Transgenic CMS plants would not release pollen to the environment, so pollination of non-transgenic crops or native plants would not occur (Munsch et al. 2008). In this context, Feil and Stamp (2002) proposed the cultivation of cytoplasmic male-sterile hybrids for the production of pharmaceuticals without the risk of pollen drift.

The Plus-hybrid system (Weingartner et al. 2002) controls the release of pollen from GM maize and increases its grain yield by growing suitable mixtures of male-sterile (80 %) and unrelated male-fertile (20 %) plants. Several studies confirmed a yield increase of between 5 and 10 % (Weingartner et al. 2002; Munsch et al. 2008).

1.2 Overview of CMS systems in plants

1.2.1 Phenotypes of CMS

Since the discovery of CMS, it has been found in more than 150 plant species (Laser and Lersten 1972; Schnable and Wise 1998). CMS phenotypes encompass a large variety of reproductive abnormalities (Chase 2007). In so-called sporophytic systems, CMS occurs due to the degeneration of pollen-producing organs or tissues

such as the anthers or the tapetal cells, for example in T- and C-CMS of maize or PET1-CMS of sunflower (Warmke and Lee 1977; Smart et al. 1994). The fertility of the pollen is determined by the genotype of the mature plant; *Rfrf* plants shed 100 % viable pollen. In gametophytic CMS systems, such as S-CMS of maize, the disruption of pollen production depends solely on the genotype of the pollen, so that an *Rfrf* plant sheds only 50 % viable pollen (Hanson and Bentolila 2004). However, a continuum between sporophytic and gametophytic CMS systems was found in D2-CMS of cotton, where the *rf* allele is transmitted by the pollen but is less competitive than *Rf* pollen (Liu et al. 2003).

In some CMS systems, male sterility arises because of homeotic changes in the morphology of the flower. Male reproductive organs (stamens) can be homeotically converted into petals or female reproductive organs, as found in carrot (Linke et al. 2003), tobacco (Kofer et al. 1991) or wheat (Murai et al. 2002). This class of CMS is often found in alloplasmic CMS strains, which are derived from intergeneric crosses (Fujii and Toriyama 2008).

1.2.2 Mitochondrial causes of CMS

Common to all CMS systems is the mitochondrial inheritance of male sterility. At least 14 mitochondrial CMS-associated genes have been characterized thus far (Chase 2007). CMS genes often arise from rearrangement of mitochondrial DNA, which leads to new chimeric open reading frames (ORFs). They are considered to be gain-of-function mutations, because plants usually retain a complete set of normal mitochondrial genes (Chase 2007). CMS genes often affect ATP production when they are cotranscribed with other mitochondrial genes or fused to promotor regions of genes encoding subunits of the ATP synthase (Hanson and Bentolila 2004). In tobacco (Bergman et al. 2000) and wild beet (Ducos et al. 2001), rearrangements of the mitochondrial DNA do not lead to new peptides, but they alter expression of mitochondrial genes, which in turn affects ATP production. Furthermore, mitochondrial changes, associated with CMS, can lead to an impaired edition of RNA and an altered synthesis of mitochondrial proteins, for example in A3-CMS of sorghum (Howad and Kempken 1997).

In PET1-CMS of sunflower and D2-CMS of cotton, a decreased supply of ATP due to mitochondrial rearrangements was suggested to trigger premature programmed cell death within the tapetum (Balk and Leaver 2001; Skibbe et al. 2008). This might lead to the collapse of the developing pollen grains. Premature degenera-

tion of tapetal cells has also been observed in T-CMS of maize and Ogura-CMS of rapeseed and radish. In those however, death of tapetal cells probably involves necrotic rather than programmed cell death. Necrotic cell death may be triggered by a severe decrease in available energy, when ATP levels are insufficient to support programmed cell death (Gonzalez-Melendi et al. 2008).

Although CMS-determining genes are expressed in the mitochondria of all tissues, the phenotypic effect is often limited to male gametogenesis. This has been attributed to the inability of CMS mitochondria to satisfy the high energy demands of male gametogenesis, indicated by a 20- to 40-fold increase in mitochondria per tapetal cell (Warmke and Lee 1977). However, the energy requirement of root respiration is as high as that of the tapetum (Balk and Leaver 2001). Therefore, a decrease in respiration might not be the only reason for the tissue specificity of CMS. Still unknown anther-specific factors might interact with the products of CMS-determining genes. Furthermore, natural selection might have restricted the effects of mitochondrial dysfunction to male gametogenesis. If mitochondrial mutations were to compromise plant growth or female reproduction, they would be eliminated quickly from the gene pool. In contrast, mitochondrial mutations, which affect male fertility only, can be passed on to subsequent generations by the maternal line (Chase 2007).

1.2.3 Restoration of fertility

All known restorer genes are located in the nucleus. Restorer genes have been cloned in maize (*Rf2*) (Cui et al. 1996), petunia (*Rf-PPR592*) (Bentolila et al. 2002), rice (*Rf1a*, *Rf1b*) (Akagi et al. 2004; Wang et al. 2006), Ogura/ Kosena radish (*Rfo, Rfk*) (Brown et al. 2003; Koizuka et al. 2003) and sorghum (*Rf1*) (Klein et al. 2005). The restorer allele at *Rf2* in maize encodes a functional mitochondrial aldehyde dehydrogenase, suggesting that restoration of fertility can occur through metabolic compensation to counteract the effects of a mitochondrial CMS-determining gene (Chase 2007). In contrast, all the other known restorer genes encode members of the pentatricopeptide-repeat (PPR) protein family. This family of proteins is very abundant in plants, comprising for example more than 450 proteins in *Arabidopsis thaliana* (Saha et al. 2007). They are believed to be involved in the regulation of mitochondrial and plastide genes by specifically binding to DNA or RNA sequences. Structurally, PPR proteins contain a characteristic 35 amino acid motif,

which occurs in tandem arrays of varying repetitions (Saha et al. 2007). All known restorer genes encoding PPR proteins target the mitochondria, thus preventing the accumulation of CMS-associated proteins or transcripts (Fujii and Toriyama 2009).

The three major CMS types in maize are distinguished by their specific restorer genes. In T-CMS, two jointly acting restorer genes, *Rf1* and *Rf2*, are essential for the complete restoration of fertility (Duvick 1965). These genes do however not restore fertility to S-CMS and C-CMS. A single major gene (*Rf3*) is required for fertile pollen in the S-CMS system (Duvick 1965). A second temperature sensitive restorer gene, *Rf9*, was reported by Gabay-Laughnan et al. (2009). The restoration of C-CMS, as far as known to date, is presented in detail in Chapter 1.3.3.

The restoration of CMS sometimes occurs spontaneously. This was found for T- and S-CMS of maize, in sorghum (Smith and Chowdhury 1989) and garden beans (Janska and Mackenzie 1993). In T-CMS of maize, the CMS-inducing gene, *urf13*, can be lost through recombination, resulting in restored fertility (Fauron et al. 1990). In S-CMS of maize, spontaneous reversion to fertility arises either through reorganization within the mitochondrial DNA (Small et al. 1988) or through mutations of the nuclear DNA resulting in new restorer genes (Gabay-Laughnan et al. 2004).

Apart from the restorer genes, which fully restore the fertility, some genes lead to a partial restoration of fertility (Duvick 1965). This occurs in all CMS systems of maize. Partially restored plants bear fewer anthers than male-fertile plants, and the emergence of the anthers is usually delayed. The anthers may be misshapen and have less pollen than a fully restored type (Tracy et al. 1991). The genetic basis of partial restoration is still unclear. It may governed by multiple genes, which may have an effect in the absence of fully restoring genes (Tracy et al. 1991; Has 2002). Furthermore, partial restoration might result from incomplete homozygosity if the inbred line has been converted into CMS by backcrossing rather than maternal induction of haploidy (Gontarovskii 1974).

1.2.4 Effects of the environment on CMS and restoration of fertility

It is of fundamental importance for the production of hybrid seeds that the CMS mechanism is environmentally robust. Reliable restoration of complete male fertility is required for the seed set of the hybrid, whereas complete and stable sterility of

the maternal parent is essential during the production of hybrid seeds to avoid self pollinations. Reliable maintenance of sterility would also be a prerequisite if CMS were to be employed for the containment of transgenic pollen.

The environment has strong effects on maize with regard to the partial restoration of male fertility (Duvick 1956; Tracy et al. 1991; Weider et al. 2009). Such plants may seem to be sterile at one location, but can dramatically change the degree of fertility at another (Duvick 1965; Tracy et al. 1991). Climatic factors such as temperature, photoperiod and light intensity are thought to be of primary importance (Kaul 1988). In general, humid and cool conditions are assumed to be conducive to restoration of fertility, whereas dry and hot conditions maintain sterility (Duvick 1956; Tracy et al. 1991). In maize, the response of S-CMS to environmental conditions is strongest, whereas that of C-CMS is between that of T- and S-CMS (Tracy 1982).

1.3 C-type CMS in maize

1.3.1 Types of C-CMS cytoplasm

Different independent strains of C-CMS were discovered in the past. Strains C, Rb, El Salvador, Pr and Bb were all assigned to the C group, because they give similar male-fertile or male-sterile responses in the presence of certain inbred line genotypes (Laughnan and Gabay-Laughnan 1983; Tracy et al. 1991). Heterogeneous patterns of endonuclease restriction as well as minor differences in the stability of CMS were found among strains of the C group (Pring et al. 1980).

1.3.2 Cause of C-CMS

C-CMS is a sporophytic system, in which the collapse of tapetal cells disrupts the production of pollen (Lee et al. 1979). The abortion of pollen in C-CMS anthers follows two distinct patterns, both of which are found in different anthers of the same plant. Lee et al. (1979) speculated that the origin for both C-CMS types is similar. Rather than the mechanism, the timing of tapetal disruption may distinguish the two types of C-CMS.

Despite comparisons of mitochondrial sequences and research on the expression of mitochondrial genes (Allen et al. 2007; Meyer and Newton 2009), the genetic cause of C-CMS has remained unknown. C-CMS is atypical compared to other

CMS types, because a unique chimeric transcript or protein, which correlates with the phenotype, was not found. However, a tassel-specific decrease in both ATP6 and ATP9 protein levels was discovered in C-CMS, which was not observed in presence of the restorer gene (Meyer and Newton 2009). Both proteins are mitochondrially encoded components of the ATP synthase F_0 subunit. The C-CMS genome has two copies of the *atp9* gene: One is the normal copy of *atp9* (*atp9-1*), which is also present in S, T and the normal cytoplasm, but which is transcribed to a lesser extent in C-CMS, possibly due to a mutation near the transcript start site. The other copy of *atp9* (*atp9-2*) is the predominant *atp9* transcript in C-CMS. The coding region is identical to *atp9-1*, but it has a chimeric 5'UTR, resulting in the use of a different promoter (cox2) and a larger transcript (4 Kb instead of 1 Kb). Therefore, Meyer and Newton (2009) hypothesized that the *atp9-2* transcript is not translated efficiently in C-CMS plants. When the demand for synthesis is high, as in the tapetal cells of the anther, the mitochondria might not be able to compensate, and ATP9 levels decrease. This could result in the formation of fewer F_0 subunits and complexes of ATP synthase, causing a reduction in levels of ATP and, ultimately, the abortion of developing pollen.

1.3.3 Restoration of C-CMS

Compared to the restoration of T- and S-CMS, the restoration of C-CMS is more complex and largely unknown. There is evidence of at least three nuclear restorer genes, called *Rf4*, *Rf5* and *Rf6*, which might interact in a complementary way. This trigenic system might be duplicated elsewhere in the genome (Vidakovic 1988; Vidakovic et al. 1997). *Rf4* and *Rf5* have been roughly mapped to chromosomes 8 (Sisco 1991) and 5 (Hu et al. 2006). Furthermore, Hu et al. (2006) mapped the gene *Rf-I*, which acts as an inhibitor of *Rf5* and is possibly identical to *Rf7* (Qin et al. 1990). Complementary interactions between paternal and maternal factors play an important role. This conclusion was drawn from crosses between C-CMS and maintainer lines in normal cytoplasm, the progeny of which was partially or fully restored (Vidakovic 1988; Sotchenko et al. 2007). The level of partial restoration is high. A C-CMS-specific form of partial restoration is the "late-break of sterility", in which pollen shedding can still occur several weeks after the emergence of the silks (Tracy et al. 1991; Kheyr-Pour et al. 1981; Vidakovic 1988).

1.3.4 Objectives of the doctoral study

The most important shortcomings of C-CMS are the high abundance of mainly unknown restorer genes throughout most gene pools in maize (Josephson et al. 1978; K.-H. Camp, pers. comm. 2007) and the high level of partial restoration, which can lead to self pollinations of the maternal inbred lines during the production of hybrid seeds (Sisco 1991; Tracy et al. 1991). Inbred lines have usually been classified as C-CMS restorers or non-restorers by means of test crosses. The inbred line is crossed to a male-sterile line, and the resulting F1 is scored for male sterility or fertility. When lines classified as restorers are to be used as seed parents in a CMS-based system, they have to be converted to non-restorers by means of time-consuming back crossings, which are sometime unsuccessful. If molecular markers tightly linked to major *rf* loci were identified, this would accelerate the conversions and reduce the number of test crosses to be phenotyped. Furthermore, marker-assisted selection could improve the counter-selection of partial restorer alleles in maternal inbred lines, because partial restoration might not be expressed every year or at every location, depending on the environmental conditions. Therefore, the objectives of this study were

1. to fine-map one of the major restorer genes of C-CMS, *Rf4*, by means of PCR-based markers.
2. to designate candidate genes for *Rf4*.
3. to map QTLs, which partially restore male fertility.
4. to assess the effects of the environment on the partial restoration of male fertility.

2 Fine-mapping of *Rf4*, a major restorer gene of C-type cytoplasmic male sterility in maize[1]

2.1 Abstract

C-type cytoplasmic male sterility (C-CMS) is today the most widely applied form of CMS for the production of hybrid seeds in maize. The inheritance of the restoration of fertility is complex and largely unknown. It is probably governed by several restorer genes, among which *Rf4* is considered to be a major one. Fine-mapping of *Rf4* was carried out in an F_2 population (n=1,317) derived from the C-CMS parent B37C and the restorer parent K55. Male fertility was rated according to an index derived from data on the quality and number of anthers as well as the anthesis-silking interval. Six new PCR-based markers linked to *Rf4* were developed by SSR (simple sequence repeat) mining and comparative sequencing of the parental lines. *Rf4* was mapped by selectively genotyping 40 % of the F_2 individuals, which had an extreme phenotype with regard to male fertility. *Rf4* mapped 0.3 cM upstream of the closest markers in the chromosomal bin 8.00. The region putatively flanking *Rf4* corresponds to a physical interval of 520 Kb, which begins at the 5' telomere of chromosome 8. This region comprises 12 predicted genes, none of which is homologous to previously cloned restorer genes mainly encoding proteins of the pentatricopeptide-repeat family. Therefore, *Rf4* might represent a new type of restorer gene.

[1]Related publication: Kohls S, Stamp P, Knaak C, Messmer R (2010) Fine-mapping of *Rf4*, a major restorer gene of C-type cytoplasmic male sterility in maize. Theor Appl Genet, submitted

2.2 Introduction

Hybrid breeding led to a rapid increase in productivity and to a stabilization of yields of maize and many other crops. The production of hybrid seeds requires a directional cross between a seed parent (female parent) and a pollen donor (male parent). The production of pollen by the female parent must be prevented to avoid female plants pollinating themselves. This can by achieved by manual detasseling or by a genetic mechanism such as cytoplasmic male sterility (CMS), which is the most important non-transgenic mechanism in maize (Chase et al. 2010). Male sterility is induced by mutations of the mitochondrial DNA disrupting the development of the pollen. Female fertility is usually unaffected (Laughnan and Gabay-Laughnan 1983), and male-sterile plants can set seeds if viable pollen is available. The maternally inherited trait can be compensated for by nuclear restorer-of-fertility genes (restorer genes), which restore fertility in the first generation after crossing a CMS and a non-CMS plant carrying the restorer allele(s) (Schnable and Wise 1998). All restorer genes cloned thus far encode proteins of the pentatricopepetide-repeat (PPR) family (Saha et al. 2007), with the exception of *Rf2* of maize, which encodes an aldehyde dehydrogenase (Cui et al. 1996). Three major types of male sterility-inducing cytoplasms are known in maize (T, C and S). They can be distinguished according to differences in the mitochondrial DNA as well as to the specific restorer genes that counteract CMS. Until 1970, more than 85 % of the hybrid seeds grown in the USA were produced by means of T-CMS. However, a sudden epidemic of Southern Corn Leaf Blight caused by a race of *Bipolaris maydis*, to which T-CMS plants were particulary susceptible, resulted in a rapid decline in the use of T-CMS (Ullstrup 1972). C-CMS then became the most widely used type of CMS in maize. C-CMS is a sporophytic system, in which the collapse of tapetal cells leads to an abortion of the developing pollen (Lee et al. 1979). The system of restoration is genetically complex and largely unknown. It is governed by at least three genes, which might complement each other (Vidakovic 1988). Furthermore, this trigenic system might be duplicated elsewhere in the genome (Vidakovic et al. 1997). Attempts to map restorer genes of C-CMS have, thus far, been limited to *Rf4*, *Rf5* and *Rf-I*, the latter of which is an inhibitor of *Rf5*. Sisco (1991) mapped *Rf4* 1.5 cM upstream of the RFLP marker npi114a on the short arm of chromosome 8. *Rf5* and *Rf-I* were roughly mapped to chromosomes 5 and 7 (Hu et al. 2006).

The frequency of restorer genes of C-CMS is high in most gene pools (Josephson

et al. 1978; K.-H. Camp, pers. comm., 2007). Inbred lines are usually classified as C-CMS restorers or non-restorers by means of test crosses. When using an inbred line, classified as a restorer, as a seed parent in a CMS-based system, it must first be converted to a non-restorer by means of backcrosses. However, this process is time-consuming and not always successful. If molecular markers tightly linked to major *rf* loci were identified, then this would accelerate the conversion and reduce the number of test crosses to be phenotyped. Therefore, the objective of this study was to fine-map one of the major restorer genes, *Rf4*, by means of PCR-based markers and to call candidate genes by means of an *in silico* annotation of the B73 reference sequence.

2.3 Materials and Methods

2.3.1 Mapping population and field experiments

Two F2 populations were derived from a cross between the C-CMS inbred line B37C and the restorer line K55. Seeds of the parental lines were obtained from the Maize Genetics Cooperation Stock Center at the University of Illinois (Urbana Champaign, IL, USA). The F2 populations were grown on the experimental station of the Swiss Federal Institute of Technology near Zurich (47°26'N, 8°40'E, 546 m a.s.l.) in 2007 (n_1=1,620 plants) and 2008 (n_2=1,302 plants), respectively. These two populations will be referred to as F_2-2007 and F_2-2008. The soil was a sandy loam eutric cambisol (IUSS Working Group WRB 2007). Kernels were sown at spacings of 20 cm within the row and 75 cm between the rows. 25 kg/ha P and 165 kg/ha K were applied before sowing, and 120 kg/ha N were applied in two portions of 60 kg/ha three and six weeks after sowing. F_2-2008 was germinated under a fleece to accelerate early development. Leaf samples were taken at the five-leaf stage of each F_2-2008 plant to extract DNA.

2.3.2 Phenotyping of male fertility

The rating of male fertility of each single plant was based on the quality and number of anthers as well as on the anthesis-silking interval (ASI). ASI was defined as the difference between the date of silking and anthesis. The quality of the anthers (anther quality) was rated on a scale of 1 to 6 (Fig. 2.3.1). 1: normal, dehiscent anthers, full shed of pollen; 2: some stunted anthers, full shed of pollen; 3: many stunted anthers and reduced shed of pollen; 4: only stunted anthers, nearly no shed of pollen; 5: empty anthers and no shed of pollen; 6: no emergence of anthers. The number of anthers (anther quantity) was rated on a scale of 1 to 6, too. 1: more than 75 % anther emergence; 2: more than 50 to 75 % anther emergence; 3: more than 25 to 50 % anther emergence; 4: more than 5 to 25 % anther emergence; 5: more than 0 to 5 % anther emergence; 6: no emergence of anthers (equivalent to a score of 6 for anther quality).

The observations took place in the morning when the anthers were fresh and filled with pollen. The tassel of each plant was examined every other day, and all the plants were phenotyped three or four times during male flowering. The final scores for anther quality and anther quantity corresponded to the average score of

anther quality and the last score of anther quantity. An index of male fertility was constructed by adding the scores for anther quality, anther quantity and the absolute value of ASI. When anthesis began before silking, ASI was set at zero. The higher the index, the lower the level of restoration. Male-sterile plants were given a separate score. A plant was considered to be fully fertile (i.e. fertility index = 2) when more than 75 % of the anthers had emerged, the appearance of the anthers was normal and anthesis started before or at silking.

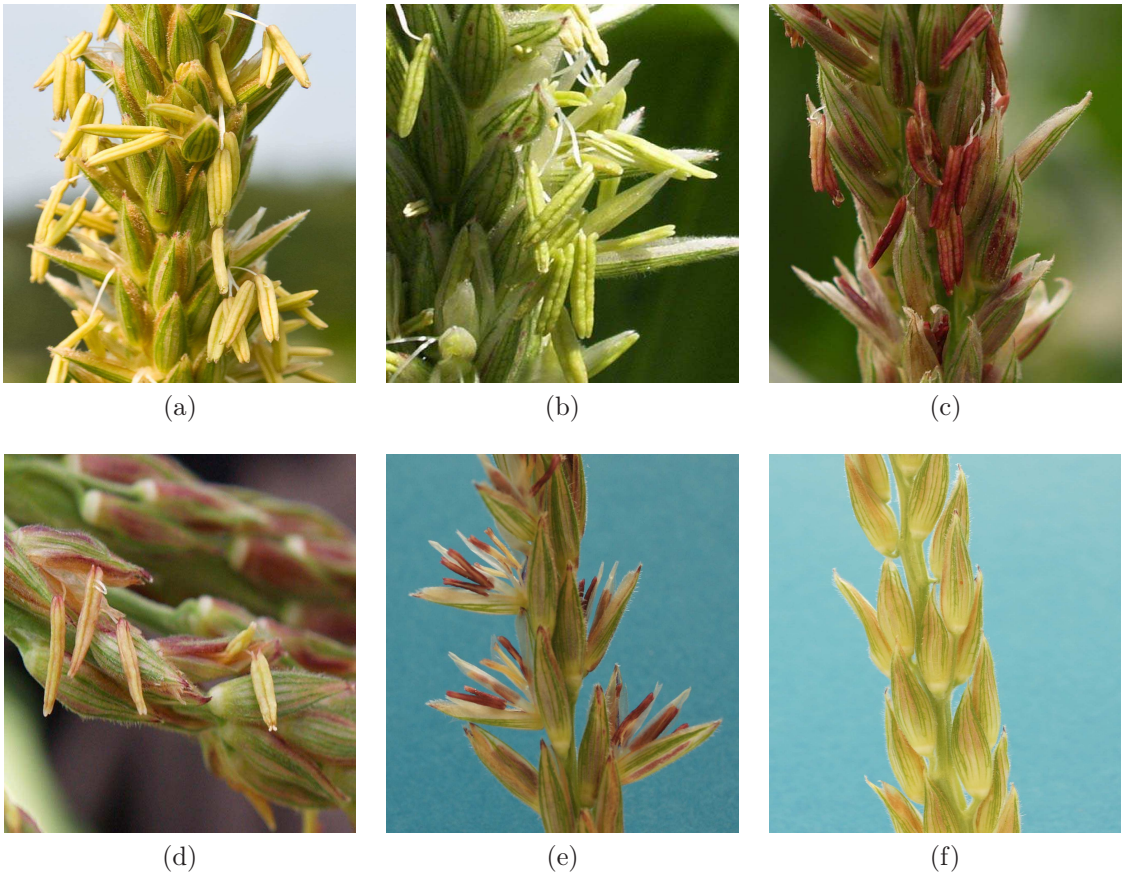

Fig. 2.3.1 Scale of anther quality: (a) 1, normal dehiscent anthers and full shed of pollen; (b) 2, some stunted anthers and nearly full shed of pollen; (c) 3, many stunted anthers and reduced shed of pollen; (d) 4, only stunted anthers, nearly no shed of pollen; (e) 5, completely empty anthers and no shed of pollen; (f) 6, no emergence of anthers

2.3.3 Analysis of field experiments

The Wilcoxon rank sum test (Mann and Whitney 1947) was used to assess whether the medians of anther quality, anther quantity, ASI and the fertility index were

significantly different between the two years. Spearman's rank coefficients of correlation between all traits were calculated (Kendall 1975). These analyses were done with the software R (R Development Core Team 2009). χ^2-square tests were calculated based on the fertility index to estimate the number of genes governing the restoration of fertility. A frequency distribution of 3:1 for male-fertile (*Rf4*) *vs* male-sterile plants (*rf4*) is expected in the event of a monogenic dominant inheritance. However, there were numerous partially restored plants, the phenotype of which was unclear with respect to the *Rf4* allele. Therefore, cut-off scores between five and 11 of the fertility index were applied, and χ^2-square tests were calculated for each splitting mode of the population.

2.3.4 Molecular analyses

Total DNA was isolated from lyophilized leaves by means of the CTAB extraction method. New PCR-based markers were developed by searching for simple sequence repeats (SSRs), single nucleotide polymorphisms (SNPs) and insertions and deletions (indels) using the B73 genome sequence, version 4a.53 (Schnable et al. 2009), as a reference for the design of primers. Markers were developed within a 1.5 Mb interval, which started at the 5' telomere of chromosome 8. The SSRIT tool (Temnykh et al. 2001) was employed to identify putative SSRs. The parameters were set for the detection of di-, tri-, tetra- and penta-nucleotide motifs with a minimum of five repeats. Seven sequencing assays on the BAC (bacterial artificial chromosome) sequences AC198928.3 and AC193663.3 were requested from TraitGenetics GmbH (Gatersleben, Germany). Primer design was carried out by the software Primer3 (Rozen and Skaletsky 2000). The parameter for the range of the product size was set at 150-300 bp for SSRs and 500-800 bp for sequencing assays; other parameters were set according to the default. All primers were ordered from Microsynth AG (Balgach, Switzerland). The amount of DNA necessary for sequencing at Microsynth AG was adjusted according to concentrations measured at a NanoDrop ND-1000 UV-Vis spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA). PCR reactions were carried out according to CIMMYT (2005) at a primer annealing temperature of 56°C. The annealing temperature was adjusted by gradient PCR when necessary. Buffer and Taq polymerase were purchased from New England Biolabs Inc. (Ipswich, MA, USA) and dNTPs from Sigma-Aldrich (St. Louis, MI, USA). Genotyping of SSRs and indels in F_2-2008 was carried on an ABI3730xl DNA Analyzer (Applied Biosystems Inc., Wilmington, DE, USA).

One SNP was transformed into a marker detectable by pyrosequencing. Forward, reverse and sequencing primers were designed with the Pyrosequencing Assay Design Software, Version 1.0.6 (Qiagen GmbH, Hilden, Germany). PCR reactions were carried out in a volume of 40 μl containing 50 ng genomic DNA, 0.2 μM of forward and reverse primer, 1× ThermoPol PCR buffer, 5 vol % dimethyl sulfoxide (DMSO), 0.125 mM dNTP and 1 U Taq polymerase. PCR conditions were set as follows: Initial denaturation 95°C for 5 min, followed by 50 cycles of 95°C for 15 s, 60°C for 30 s, 72°C 15 s; final elongation at 72°C for 4 min. Optimal annealing temperatures of primers were revealed by gradient PCR. Genotyping of F2-2008 was carried out by KWS Saat AG on a PSQHSA96 Pyrosequencer.

2.3.5 Analysis of genetic linkage

New markers were mapped together with one proprietary marker of KWS Saat AG (Einbeck, Germany) (pm2) and three public SSR markers (umc1359, umc1327 and umc1483). The linkage map was constructed using Mapmaker 3.0 (Lander et al. 1987). A minimum LOD (log likelihood of the odds) score of 3.0 was applied, and the Kosambi function was used to obtain the genetic distances in centiMorgan (cM) (Kosambi 1944). The fertility indeces were transformed to dominant marker scores to map restoration of fertility as a dichotomous trait. Fertile plants were considered to carry the restorer allele (*Rf4Rf* or *Rf4rf4*), whereas partially restored and sterile plants were considered to carry the non-restoring allele (*rf4rf4*). Only phenotypically extreme plants, i.e. 20 % from the bottom and 20 % from the top part of the distribution of the fertility index, were selected for linkage mapping to avoid assigning an incorrect marker score to partially restored plants.

2.3.6 Annotation of the B73 reference sequence

The filtered gene set 4a.53 of the B73 reference sequence flanking *Rf4* (AC187051.4, AC192587.3, AC198928.3, AC193663.3 and AC226331.3; maizesequence.org) was supplemented with an *ab initio* annotation. Elements of repetitive sequence were masked from the B73 sequence by the RepeatMasker software open-3.2.9 (Smit et al. 1996-2010). The sequence was then submitted to the prediction programs Fgenesh 2.6 (Salamov and Solovyev 2000), using the training set of the monocotyledons, and GenScan 1.0 (Burge and Karlin 1997), using the training set of maize. Fgenesh was used to define the positions of introns and exons. The critera to define a gene were

(1) prediction from both prediction programs, (2) a minimal translational length of 50 amino acids and (3) similarity to a putative, hypothetical or known protein-coding sequence in the non-redundant protein database at an expectation value smaller than *e*-30 in BLASTX queries. The predicted open reading frame (ORF) was defined as a gene encoding an unknown protein when it matched an expressed sequence tag (EST) of maize in BLASTN queries with a minimum of 97 % identity between sequences. BLAST searches were supplemented with protein motif queries in the PFAM database (Finn et al. 2010). All predicted amino acid sequences were analyzed by Predotar 1.03 (Small et al. 2004), MitoProt 1.101 (Claros and Vincens 1996) and TargetP 1.1 (Emanuelsson et al. 2007), which predict the subcellular location of proteins.

2.3.7 Test of *Rf4* markers in a set of restorer and non-restorer lines

The most closely linked markers for *Rf4* were tested in a set of 12 restorer and 21 non-restorer lines. The restorer lines K55, M14, W64A, N6, Oh45 and W23 were obtained from the Maize Genetic Cooperation Stock Center and six lines from Delley Seeds and Plants Ltd (Delley, Switzerland). The non-restorer lines B37, Va58, KyS, Wf9, Mo17, B73, Tr, A619, A632, VA26, Hi27, Oh51, A634 and B73 were obtained from the Maize Genetics Cooperation Stock Center, V395/31, L70/09 from the European Maize Database at the Maize Research Institute Zemun Polje (Serbia and Montenegro) and one from Delley Seeds and Plants Ltd. Extraction of DNA and PCR was carried out as described above. Markers were visualized on 4 or 6 % high-resolution agarose gels (Carl Roth GmbH, Karlsruhe, Germany), which were prestained with Gelred™ (Biotium Inc., Hayward, CA, USA) and run in 1× TBE buffer (Sambrook et al. 1989) at 150 V for 90 min. A 20 or 100 bp ladder (Bio-Rad Laboratories, Hercules, CA, USA) was used as size standard.

2.3.8 Evaluation of restorer properties of B37

The fertile version of the maternal inbred line B37, carrying a normal cytoplasm (B37N), was crossed with the C-CMS inbred lines Mo17C, Wf9C, V395/31C and L70/9C in order to evaluate the presence of complementary restorer genes in B37. The resulting F_1 progenies were grown and phenotyped for the restoration of fertility at the same location as F_2-2008.

2.4 Results

2.4.1 Climatic conditions

The monthly mean temperatures in 2007 and 2008 were close to the long-term means, except in early spring. The mean temperature in April 2007 was 4 K higher than the long-term mean, but in 2008 it was 2 K lower (Table 2.4.1). Total precipitation during the vegetation period was above the long-term mean in both years, even though dry periods occurred in April 2007 and in May 2008. Male flowering lasted from the end of July to the end of August.

Table 2.4.1 Monthly mean temperatures and precipitation at Eschikon during the vegetation periods in 2007, 2008 and in the long-term (lt) mean.

	2007	2008	lt mean
Temperature [°C]			
April	14	8	10
May	15	16	14
June	18	18	18
July	18	19	19
August	17	18	18
September	13	13	15
Mean	**16**	**15**	**16**
Precipitation [mm]			
April	10	238	115
May	173	46	136
June	141	101	116
July	242	144	167
August	289	228	202
September	82	124	92
Sum	**937**	**881**	**828**

2.4.2 The maternal parent was partially restored

B37 carries at least one complementary restorer allele at a locus other than *Rf4*, because the F_1 progenies of the crosses between L70/09C and V395/31C with B37N were fully restored. Furthermore, B37C was not fully sterile. In 2007, 2008 and 2009 (Chapter 4), the fertility of approximately two thirds of the plants was partially restored.

2.4.3 Segregation of male fertility

The segregation of male fertility in the F_2 populations ranged from full fertility to complete sterility and included all levels of partial restoration (Fig. 2.4.2). The frequency distributions of anther quality, anther quantity, ASI and the fertility index reached maxima close to full fertility in both years. Male-sterile plants were extremely rare: 1.01 % in 2007 and 1.06 % in 2008. Medians of anther quality and anther quantity did not differ significantly between the years ($p>0.01$). However, significant differences were found for ASI and the fertility index due to a greater proportion of plants with delayed anthesis in 2007 than in 2008. The observed high positive ASI for some plants was due to unusually late silking because of developmental abnormalities rather than due to early anthesis. The skewness of the distributions to full fertility suggested that one dominant gene was responsible for complete restoration, whereas several other genes led to the high number of partially restored plants. χ^2 tests supported the presence of one fully restoring gene in both years. In 2007, χ^2 support was found at a cut-off of 10 of the fertility index (χ^2=0.23), whereas in 2008, it was found at 7 (χ^2=0.67). This was due to different distributions of ASI in 2007 and 2008.

The positive coefficients of correlation between anther quality and anther quantity in both years (Table 2.4.2) showed that a normal, fertile appearance of the anthers was accompanied by a high number of anthers. The negative correlation between these two traits and ASI indicates that fertile plants tended to start anthesis before silking (protandry), whereas the anthers of partially restored plants tended to emerge after silking (protogyny).

Table 2.4.2 Spearman's rank correlation coefficients among anther quality, anther quantity and the anthesis-silking interval (ASI) of the populations F_2-2007 and F_2-2008.

	F2-2007		F2-2008	
	Anther quality	Anther quantity	Anther quality	Anther quantity
ASI	-0.33**	-0.45**	-0.38**	-0.50**
Anther quantity	0.54**		0.60**	

** Significant at a significance level of $p \leq 0.01$

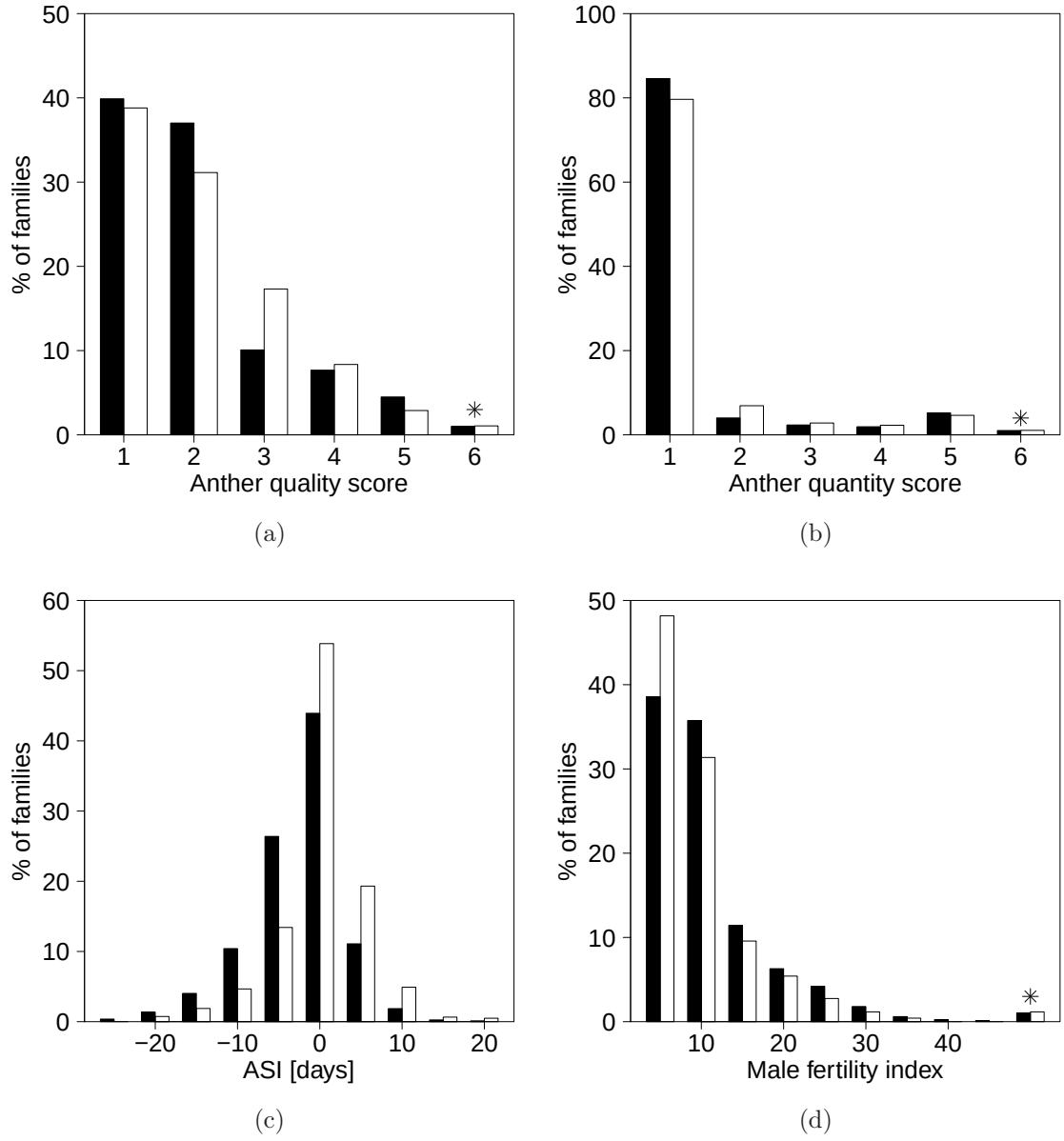

Fig. 2.4.2 Frequency distributions of (a) anther quality, (b) anther quantity, (c) the anthesis-silking interval (ASI) and (d) the fertility index of the populations F_2-2007 (black) and F_2-2008 (white). $*$ Male-sterile individuals.

2.4.4 Genetic map position of *Rf4*

Three codominant SSR markers (b0329-13.1, c0216-3.15 and b0440-16.2) and one dominant marker (b0640-3.4) were developed from 93 combinations of primers. Two indel markers (c0466-1.7, c0113-2.2) and one SNP marker (b0154-2.1) were derived from high-quality sequencing assays, which were developed from 134 combinations

of primers. Sequences of primers and further properties of all markers are listed in Table 2.4.3 and the appendix (5.4.1, 5.4.2, 5.4.3).

Table 2.4.3 Primer sequences of new PCR-based markers linked to *Rf4*.

Marker name	BAC	Marker type	Map position		Primer sequences (5'-3') forward/reverse
			cM	bp[a]	
b0329-13.1	AC187051.4	SSR	0.3	52,765-52,986	CTCCGAACCTGATCCGAGTA/ AGGGGAGAGGTCCCAGAATA
b0640-3.4	AC192587.3	PA[b]	0.3	12,095-12,415	AGCGAGGCAAAGAATTGAGA/ GCAGAGCTTAAAGACTCCAAAGAC
c0466-1.7	AC198928.3	indel	0.3	70,416-70,475	GACTCGATCGTTCCGGCCCAA/ CCGTCGACGAGCACCTTTTGGTA
c0113-2.2	AC226331.3	indel	1.5	58,895-59,163	GAAAGAGTTGGCAGCTTTCG/ TAGCGCGTGGTATGGTAGTG
c0216-3.15	AC188738.4	SSR	2.5	24,272-24,396	TCGCATCATTGCATAGTGCT/ GCGAACTATTTGTGCCGTCT
b0440-16.2	AC203872.3	SSR	3.1	91,951-92,146	AATGGTGGCTGGATTGGTT/ GGATGGGTGTCCTCGATATG

[a] Physical position on respective BAC
[b] Presence-absence polymorphism

All molecular markers mapped to the same linkage group, except b0154-2.1. A thorough retry in the alignment of sequences revealed that the sequence around this marker is duplicated twice on chromosome 4 (AC193391.3: 220,644,674 - 220,648,546 bp, AC204423.4: 122,712,521 - 122,713,133 bp), probably leading to a multiple amplification of loci in the genome. This marker was therefore not considered.

The map covered a segment 22.6 cM wide in the chromosomal bin 8.00 (Fig. 2.4.3). The genetic order of markers was in accordance with the physical order proposed at www.maizesequence.org. *Rf4* mapped 0.3 cM upstream of the co-segregating markers b0329-13.1, b0640-3.4 and c0466-1.7. This interval resulted from one recombination event. No marker was located upstream of *Rf4*. Downstream, the co-segregating group of markers was flanked by the marker c0113-2.2, which mapped 1.5 cM from the putative position of *Rf4*.

2.4.5 Annotation of the sequence flanking *Rf4*

The 1.5-cM interval between the 5' telomere of chromosome 8 and the marker c0113-2.2 corresponds to 520 Kb of the B73 reference sequence. It comprises the BACs AC187051.4, AC192587.3, AC198928.3, AC193663.3 and AC226331.3. AC187051.4

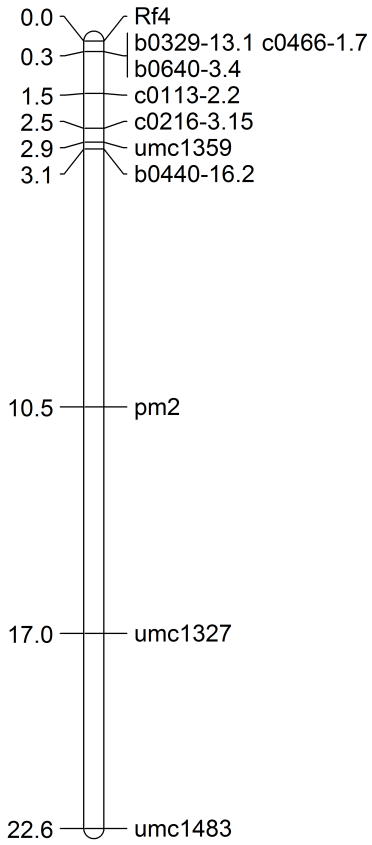

Fig. 2.4.3 Genetic map of *Rf4* in the chromosomal bin 8.00 based on the selectively genotyped population F_2-2008 (20 % from the bottom and 20 % from the top of the distribution of the fertility index).

is the first BAC on chromosome 8. This genetic region contains, on average, 76 % repetitive elements, the greatest proportion was harbored by BAC AC192587.3 (89 %). Twelve putative genes were predicted, which consist of seven genes from the filtered gene set and five *ab initio*-predicted genes (Table 2.4.4). No gene was predicted on the BAC AC193663.3. Putative functions were assigned to eight of 12 predicted genes by BLASTX and PFAM queries (Table 2.4.5). However, none of these genes was homologous to previously identified restorer genes: Patterns of the PPR and the aldehyde dehydrogenase protein family were not found. Categories of predicted genes included proteins involved in defense reactions and detoxification (peroxidase, glutathion S transferase), regulatory functions (bHLH protein, Squamosa promotor binding protein, phosphodiesterase), transport proteins (ABC transporter), proteases (peptidase M48) and several proteins with unknown function. One predicted gene did not align to a protein with a sufficiently low *e*-value.

Eight predicted genes were homologous to an EST of maize. None of the predicted proteins had a high probability of being exported to the mitochondria (data not shown).

Table 2.4.4 Genomic location of predicted genes in the 520 Kb region putatively flanking *Rf4*. Predicted genes belong to the "filtered gene set" (filtered gs) or were obtained from an *ab initio*-prediction by Fgenesh and GenScan. f: forward, r: reverse

BAC	Gene	Position [bp][a]	Prediction	Transcript ID	Strand
AC18705.4	1	119,668-121,420	filtered gs	GRMZM2G122853	r
	2	41,023-42,162	filtered gs	GRMZM2G021276	f
	3	62,162-62,925	*ab initio*	-	f
	4	87,717-94,701	*ab initio*	-	r
AC192587.3	5	78,411-84,917	filtered gs	GRMZM2G081127	r
AC198928.3	6	145,720-148,540	filtered gs	GRMZM2G384298	f
	7	143,294-145,137	filtered gs	GRMZM2G384296	f
	8	136,296-142,826	filtered gs	GRMZM2G085111	r
	9	903,83-102,882	filtered gs	GRMZM2G085038	f
AC226331.3	10	107,577-108,554	*ab initio*	-	f
	11	112,869-116,108	*ab initio*	-	r
	12	117,134-121,194	*ab initio*	-	r

[a] Physical position on respective BAC

2.4.6 Evaluation of markers linked to *Rf4* in other inbred lines

The restorer phenotype of a set of restorer and non-restorer lines was not reflected by the alleles of markers tightly linked to *Rf4*. The restorer allele of the marker c0466-1.7 occurred exclusively in K55, whereas all the other lines carried the allele of B37C. The marker b0329-13.1 showed at least six alleles, detected as bands between 200 and 280 bp in a 6 % agarose gel, which did not co-segregate with the restorer phenotype (data not shown).

Table 2.4.5 Putative functions and BLAST alignments of predicted genes in the genomic region putatively flanking *Rf4*.

			BLASTX[b]		BLASTN[b]	
BAC	Gene	Homolog (*species*)[a]	*e*-value	Protein acc. no.	*e*-value (% identity)	EST acc. no. (*Species*)[a]
AC187051.4	1	peroxidase 1 (*Zm*)	1*e*-166	NP_001152255.1	0 (100 %)	DV171681.1 (*Zm*)
	2	basic-helix-loop-helix (bHLH) family protein (*Zm*)	3*e*-135	NP_001149921.1	0 (100 %)	EE288506.2 (*Zm*)[c]
	3	hypothetical protein (*Zm*)	1*e*-30	ACN36173.1	-	-
	4	hypothetical protein (*Zm*)	5*e*-108	AAQ06285.1	-	-
AC192587.3	5	Squamosa promoter binding protein-like 1 (*At*)	4*e*-145	NP_850468.1	0 (99 %)	DR817637.1 (*Zm*)[c]
AC198928.3	6	hypothetical protein (*Zm*)	3*e*-58	NP_001168125.1	0 (99 %)	EB162498.1 (*Zm*)
	7	unknown protein (*Zm*)	-	-	4*e*-139 (100 %)	FL470650.1 (*Zm*)
	8	MDR-like ABC transporter (*Os*)	0	CAD59590.1	-	-
	9	pto kinase interactor 1 (*Zm*)	7*e*-155	ACG37409.1	0 (100 %)	EB400110.1 (*Zm*)
AC226331.3	10	nucleotide pyrophosphatase/ phosphodiesterase (*Zm*)	2*e*-60	NP_001152048.1	1*e*-139 (98 %)	EB816248.1 (*Zm*)
	11	glutathione S-transferase, C-terminal-like (*Zm*)	2*e*-84	ACG32999.1	-	-
	12	peptidase M48, Ste24p (*Zm*)	1*e*-82	ACG33945.1	6*e*-167 (98 %)	DR803099.1 (*Zm*)

[a] *Zm Zea mays, At Arabidopsis thaliana, Os Oryza sativa*
[b] Dashes indicate homology values below threshold score
[c] Multiple EST hits at similar *e*-values

2.5 Discussion

2.5.1 Low environmental influence on the restoration of fertility

The scores of anther quality, anther quantity, ASI and the fertility index followed approximately similar frequency distributions in 2007 and 2008. This suggests a weak effect of the environment on the expression of male fertility. Furthermore, it shows the accuracy of the scoring method in the field. The distribution of ASI in 2007 indicated a stronger protogyny than in 2008. Protogyny is typically observed in partially restored plants (Gontarovskii 1974; Tracy et al. 1991) but this effect might have been intensified by an unusual drought stress early in the season in 2007, which affected non-CMS populations as well (K.-H. Camp, pers. comm., 2007). The proportion of sterile plants remained about the same, whereas the proportion of partially restored plants varied between years. A similar observation was made in subsets of T-CMS populations of maize, which usually have a similar percentage of male-sterile plants but often differ from each other with respect to the proportion of partially restored and fertile plants when grown at different locations (Duvick 1956). Apparently, the expression of partial restorer alleles is affected stronger by environmental conditions than the expression of fully restoring and non-restoring alleles (Tracy et al. 1991).

2.5.2 *Rf4* is located in bin 8.00

Rf4 mapped to the chromosomal bin 8.00, supporting the result of Sisco (1991). The map distance between *Rf4* and the closest marker is short enough for marker-assisted selection. The position of *Rf4* upstream of the uppermost markers may be caused by plants, the marker scores of which were incorrect because of unclear phenotypes. The fertility index, which took into account the data of anther quality, anther quantity and ASI, seemed to be less susceptible to such errors than any of these traits alone. When linkage mapping was performed with one of the traits instead of the index, the interval between *Rf4* and the closest markers increased from 0.3 cM to 1.7, 4.7 and 4.6 cM, respectively (data not shown). This was due to numerous plants, which partially fulfilled the criteria for full fertility (more than 75 % anther emergence, normal appearance of the anthers and anthesis before or at silking). Their allelic status may be misconceived when only one of the three traits

is used for mapping.

χ^2 analyses suggested that *Rf4* is inherited like a single dominant gene in the cross between B37C and K55. This corresponds to the results by Kheyr-Pour et al. (1981) and Sisco (1991) who found a similar pattern of segregation in crosses with other inbred lines. However, a 3:1 segregation in F_2 does not preclude the presence of more than one restorer gene or the possibility that *Rf4* is part of a complex of restorer genes, as postulated by Vidakovic (1988). The mode of action of *Rf4* could not be deduced from the results of this study, because the restorer properties of the maternal line B37C are not completely known. B37 carries at least one complementary restorer allele at a locus other than *Rf4*, because crosses with the C-CMS inbred lines V395/31C and L70/09C were fully restored in F_1. If this unknown second locus interacted with *Rf4* in a complementary way and if K55 carried the same allele as B37, then *Rf4* would nevertheless segregate like a single dominant gene. An association of *Rf4* to a complex of restorer genes was supported by the putative *Rf4* alleles of a set of inbred lines, which did not mirror their restorer or non-restorer phenotype. A C-CMS line may be male-sterile in spite of the restoring *Rf4* allele if it carries the non-restoring allele at a complementary locus or loci.

2.5.3 *Rf4* might encode a new type of restorer protein

Rf4 might represent a new type of restorer gene, because a PPR gene was not predicted in the region flanking *Rf4* (Table 2.4.5). C-CMS might involve a different mechanism of restoration, because, in contrast to other CMS systems restored by PPR proteins, there is no unique mitochondrial ORF, which correlates with the C-CMS phenotype (Allen et al. 2007; Meyer and Newton 2009). However, the possibility that *Rf4* encodes a PPR protein cannot be excluded until the parental lines are re-sequenced. The prediction of candidate genes may not fully represent the plant material used in this study, since it was based on the reference sequence of B73, which was still incomplete. Moreover, B73 might even lack the *Rf4* gene. This is a realistic possibility considering that B73 maintains C-CMS according to the Maize Genetics Cooperation Stock Center. Such a situation was described for the sequenced rice cultivar *Nipponbare*, which lacks the *Rf1a* gene restoring BT-CMS (Fujii and Toriyama 2008). Furthermore, the process of pollen abortion in A1-CMS of sorghum, which is indeed restored by a PPR protein (RF1, Klein et al. 2005), is morphologically similar to that of C-CMS (Lee et al. 1979).

Hypothetically, C-CMS is caused by the inefficient translation of an atp9 transcript. This might result in a dysfunctional F_0 unit of the ATP synthase and, ultimately, in abortion of pollen due to a lack of energy (Meyer and Newton 2009). Therefore, RF4 might be implicated in the restoration of the function of the ATP synthase or realize a translation of the atp9 mRNA. However, there may well be other physiological ways of restoration, and it is possible that RF4 does not directly impact the mitochondrial cause of C-CMS. The predicted genes in the region flanking *Rf4* comprise a variety of proteins (Table 2.4.5). Some are promising candidate genes, because they have been discussed in the context of mitochondrial-nucleic interactions before.

A putative basic helix-loop-helix (bHLH) protein was predicted on BAC AC18705.4. Proteins in this family are involved in various processes, for example in retrograde regulation, i.e. signaling from the mitochondria to the nucleus, which has been observed in cases of defective mitochondria in animals and yeast (Sekito et al. 2000). However, without further genetic mapping or expression studies, it remains speculative if and how the predicted bHLH protein could interfere with C-CMS.

A peroxidase was predicted on the same BAC as the putative bHLH transcription factor. A putative ascorbate peroxidase (APX) was hypothesized to be involved in the restoration of fertility in pol-CMS of rapeseed, because it interacts with the CMS-causing protein ATP6, and the expression of both ATP6 and APX is low in sterile as compared to fertile plants (Liu et al. 2010).

Another interesting candidate gene putatively encodes a M48 metalloprotease, located on AC226331.3. Matsuhira and Harada (2005) found a cluster of four mitochondria-targeted metalloprotease-like genes in the *Rf1* locus of sugar beet, which was severely truncated in the sterile *rf1* plant. However, there is no evidence yet as to whether this cluster of genes is associated with the restoration of fertility.

2.5.4 Conclusions

The fine-mapping of *Rf4* resulted in a set of tightly linked, codominant markers, which can be used for marker-assisted selection. Although it is still not possible to deduce the restorer status of an inbred line based solely on the *Rf4* allele, marker-assisted selection will be useful once the abundance of *Rf4* in a certain gene pool is known. The presence of at least one complementary restorer gene in the maternal parent B37C suggests the need to counter-select restorer alleles in maternal inbred lines, even if the expression of male sterility is stable. Mapping complementary

restorer genes could be achieved by crossing C-CMS lines with non-restoring lines carrying a normal cytoplasm. Populations that are appropriate for the genetic mapping of major restorer genes should not exhibit a high level of partial restoration, because plants with an unclear phenotype complicate the mapping process. If there is a constant variation in male fertility, then selecting phenotypically extreme plants for linkage mapping is a good way to alleviate this problem. The future cloning of *Rf4* promises novel insights into CMS and restoration of fertility, because *Rf4* hypothetically represents a new type of restorer protein.

3 Bulk Segregant Analysis of partial restorer genes of C-type cytoplasmic male sterility in maize

3.1 Abstract

Although partial restoration of male fertility limits the use of C-type cytoplasmic male sterility (C-CMS) for the production of hybrid seeds in maize, the genetic basis of the trait is still unknown. Therefore, a Bulk Segregant Analysis (BSA) was performed to identify major QTLs for a potential marker-assisted selection against partial restorer genes. BSA was carried out with plants from an F_2 population and 10 selected F_2BC_1 progenies, which were derived from the Corn Belt lines B37C and K55. Three major loci were detected in the chromosomal bins 3.05, 3.06 and 8.02 in the F_2 population, whereas numerous QTLs were found across the whole genome in the F_2BC_1 progenies. Both paternal and maternal factors were involved in the partial restorations. The presence of major loci is encouraging for a further QTL analysis to determine the inheritance of the partial restoration of C-CMS in maize.

3.2 Introduction

C-CMS became the most widely applied form of CMS for the production of hybrid seeds in maize. However, an important shortcoming of C-CMS is the frequent occurrence of partially restored plants, which bear fewer anthers and shed less pollen than fully fertile plants (Tracy et al. 1991). Partially restored inbred lines cannot be used as seed parents, because even a low level of pollen shedding can lead to self pollinations of the maternal parent and, consequently, to an admixture inbred to

hybrid seeds. Furthermore, if CMS were employed for the containment of transgenic pollen, then it would be necessary to prevent partial restoration. Even though the trait has important practical implications, its genetic basis is still unknown. Probably, the trait is governed by multiple factors, which have an effect in the absence of fully restoring genes. The expression of the partial restorer genes is apparently strongly influenced by the environmental conditions (Duvick 1956; Tracy et al. 1991; Weider et al. 2009).

To counter-select partial restorer alleles in the breeding germplasm, it is important to gain a better understanding of their inheritance. If major genomic loci were detected, then marker-assisted selection could be applied to enhance the efficiency of the development of maternal inbred lines, because the partial restoration might not be expressed every year or at every location due to the impact of environmental conditions. In this study, a Bulk Segregant Analysis (BSA) was performed with plants from an F_2 population and 10 selected F_2BC_1 progenies to identify major QTLs governing partial restoration.

3.3 Materials and Methods

3.3.1 Plant material and field experiments

The F_2 population (F_2-2007; n=1,620) was derived from a cross between the C-CMS inbred line B37C and the restorer inbred line K55. Seeds of the parental lines were obtained from the Maize Genetics Cooperation Stock Center at the University of Illinois (Urbana Champaign, IL, USA). The cross between B37C and K55 segregates for a major dominant restorer gene, *Rf4*, and an unknown number of partially restoring genes. F_2BC_1 progenies were generated by backcrossing 10 F_2 individuals with the male-fertile variant of B37 carrying a normal cytoplasm (B37N). By means of a marker, proprietary to KWS Saat AG (Einbeck, Germany), these F_2 plants were selected to carry the non-restoring allele *rf4* to eliminate the masking effect of the dominant *Rf4* allele over the partial restorer genes. Two selected F_2 individuals were sterile, whereas the others showed different levels of partial restoration. F_2BC_1 progenies of selected F_2 plants were expected to segregate into partially restored and sterile plants only. The progenies comprised 170 to 310 plants (Table 3.3.1).

In 2007 and 2008, respectively, F_2-2007 and F_2BC_1 families were grown on the experimental station of the Swiss Federal Institute of Technology near Zurich (47°26'N, 8°40'E, 546 m a.s.l.). The soil was a sandy loam eutric cambisol (IUSS Working Group WRB 2007). In both experiments, kernels were sown at intervals of 20 cm within a row and 75 cm between the rows. 25 kg/ha P and 165 kg/ha K were applied before sowing, and 120 kg/ha N were applied in two portions of 60 kg/ha three and six weeks after sowing. F_2BC_1 families were grown in plots of 15 plants, which were assembled in a randomized complete block design of 13 blocks. Two families occurred twice in each block. Leaf samples were taken from each F_2 and F_2BC_1 plant at the five leaf stage for the extraction of DNA.

3.3.2 Phenotyping and formation of bulks

The F_2 and F_2BC_1 plants were phenotyped for the number of anthers (anther quantity), the quality of the anthers (anther quality) and the anthesis-silking interval (ASI) (Chapter 2). Plant height was considered to be a general parameter of plant vitality. Anther quality and anther quantity were rated on scales from 1 to 6, 1 representing full male fertility and 6 representing male sterility. Plants were considered to be partially fertile if anther quality and quantity had scores equal or higher than

3. Sterile plants did not bear anthers within the first three weeks after silking. Table 3.3.1 gives the anther scores of F_2 plants, selected as parents for the generation of F_2BC_1 progenies.

In the BSA of F_2 individuals, two bulks with fully fertile, two bulks with fully sterile individuals and three bulks with differed levels of partially restored plants were formed. Within each F_2BC_1 family, two bulks were constructed containing partially restored and sterile plants, respectively. All bulks contained 12 plants whenever possible.

Table 3.3.1 Scores of anther quality, anther quantity and anthesis-silking interval (ASI) of the F_2 parents of F_2BC_1 progenies used for Bulk Segregant Analysis. pmf: partially male-fertile, ms: male-sterile, n (F_2BC_1): number of plants per F_2BC_1C1 family.

F_2 parent	Phenotype	Anther quality	Anther quantity	ASI [days]	n(F_2BC_1)
43-04	pmf	3.3	1	11	180
43-11	pmf	3.5	1	18	179
44-08	ms	6.0	6	-	173
73-12	pmf	4.0	3	15	314
73-19	pmf	3.0	1	15	152
74-10	ms	6.0	6	-	178
75-12	pmf	4.5	5	11	142
75-19	pmf	4.0	4	17	261
103-11	pmf	4.0	2	3	179
104-11	pmf	4.3	3	12	174

3.3.3 Statistical analysis

The randomized block design was analyzed using the R package ASReml (Butler et al. 2007). Missing values were estimated as a consequence of fitting the model. Analyses of variance were calculated for anther quality, anther quantity, ASI and plant height based on plot raw data. Best linear unbiased predictors (BLUPs) were estimated for each F_2BC_1 family. The distribution of residuals was tested for normality by a Shapiro-Wilk test (Shapiro and Wilk 1965). The mixed effects model used to analyze the experiment was:

$$y_{ij} = \mu + f_i + b_j + e_{ij} \tag{3.1}$$

where y_{ij} is the predictor of the i-th family in the j-th block; e_{ij} is the residual error variance and μ is the intercept. Blocks and families were considered to be random effects. Within-location heritabilities (sometimes referred to as repeatabilities) were calculated according to:

$$h_w^2 = \frac{\sigma_f^2}{\sigma_f^2 + \frac{\sigma_e^2}{N}} \tag{3.2}$$

where h_w^2 is the within-location heritability, σ_f^2 is the variance of the families, σ_e^2 is the residual error variance. N denotes the mean number of plants per plot.

3.3.4 Genotyping

All F_2 and F_2BC_1 bulks were genotyped with 1,536 SNP markers proprietary to KWS Saat AG (Einbeck, Germany) on the Illumina VeraCode BeadXpress platform multiplexing 384 markers (Illumina Inc., San Diego, CA, USA). The GoldenGate assay was performed according to the manufacturer's protocol and as described in Fan et al. (2003). The automatic allele calling for each locus was accomplished with the Illumina GenomeStudio Analysis Software. The two alleles of a SNP give different detectable signals due to an allele-specific ligation to fluorescent products. Therefore, a genotype that is homozygous for one or the other SNP allele will give one signal, whereas a heterozygous genotype will give both signals. All SNPs were checked for good separation of signals and were manually re-scored when errors in calling the homozygous or heterozygous signal were found. The genetic locations (cM) of the SNP markers correspond to the IBM neighbors 2008 reference map available at www.maizegdb.org (Schaeffer et al. 2008).

A QTL was recorded when a SNP was polymorphic between the fertile or partially fertile bulks *vs* the sterile bulks.

3.4 Results

3.4.1 Climatic conditions

The monthly mean temperatures in 2007 and 2008 were close to the long-term mean (Chapter 2, Table 2.4.1). However, in 2007 it was very warm in early spring, whereas the reverse was true in 2008. Total precipitation during the vegetation period was above the long-term mean in both years, even though dry periods occurred in April 2007 and in May 2008. The male flower lasted from end of July to end of August.

3.4.2 The maternal parent was partially restored

The maternal inbred line B37C was partially restored in both years of study as well as in 2009 (Chapter 4). Partial restoration occurred in approximately two thirds of the plants monitored. The level of restoration differed from plant to plant but never reached full restoration. As a spot check, pollen of partially restored B37C plants was examined under the microscope. Approximately half of the pollen grains had aberrant shapes and were not completely filled with starch; the rest seemed to be viable (data not shown).

3.4.3 Segregation of male fertility in F_2 and F_2BC_1

The population F_2-2007 segregated into a large proportion of fully restored (75 %), a small proportion of partially restored (approximately 25 %) and into very few sterile plants (approximately 1 %) (Chapter 2). The segregation ratio as well as subsequent molecular analyses suggest that restoration of complete fertility was governed by the major restorer gene *Rf4*. The low number of sterile plants indicates that partial restoration was inherited oligo- or polygenically.

Within the F_2BC_1 families, there were no fully fertile individuals; all families segregated exclusively into partially restored and sterile plants (Table 3.4.2). Thus, the masking effect of the *Rf4* allele was successfully eliminated by marker-assisted selection. Partial restoration was expressed as reduced shed of pollen due to stunted anthers, a lower number of anthers and a mean delay of anthesis by 12 days (Table 3.4.2). This is typical for partially restored plants. The segregation within F_2BC_1 families could not be explained by simple Mendelian patterns of inheritance involving a few unlinked genes. The fact that fewer than 10 % of the plants were sterile in half of the families suggests that multiple genes were involved in partial restoration.

Since F_2BC_1 families were generated by backcrossing F_2 plants to the non-restoring parent B37N, F_2BC_1 progenies were expected to exhibit an equal or lower level of restoration compared to their respective F_2 parent. However, three out of 10 F_2BC_1 families were, on average, more fertile (44-08, 74-10 and 75-12). F_2 plants 44-08 and 74-10 were sterile, but their respective F_2BC_1 progenies bore a considerable number of anthers (average anther quantity scores 1.3 and 4.4). The restoration of the F_2 plant 75-12 was very low, but its progeny were the most fertile with only 4 % sterile plants (Table 3.4.2).

The analysis of variance across F_2BC_1 families revealed highly significant differences among families for all traits. Whereas block effects were negligible with respect to anther quality, anther quantity as well as ASI, they were significant for plant height (data not shown). Consequently, within-location heritabilities were above 0.8 for all three fertility-related traits, but lower for plant height (0.72).

Table 3.4.2 Best linear unbiased predictors (BLUPs) of anther quality, anther quantity, the anthesis-silking interval (ASI) and plant height, and segregation of each family into partially male-fertile (pmf) and male-sterile (ms) plants in 2008.

F_2BC_1 family	Anther quality	Anther quantity	ASI [days]	Plant height [cm]	% pmf	% ms
43-04	2.0	2.5	11	225	93	7
43-11	1.6	3.8	13	208	86	14
44-08	5.2	1.3	15	226	69	31
73-12	4.0	2.3	10	214	94	6
73-19	4.0	2.8	11	219	91	9
74-10	4.9	4.4	15	220	63	37
75-12	3.7	1.9	9	225	96	4
75-19	4.4	3.6	9	223	80	20
103-11	4.8	3.4	12	210	93	7
104-11	4.4	3.6	13	216	86	14

3.4.4 BSA results

357 of 1,536 markers were scored as being polymorphic between the parental lines. Chromosomes were covered with an average number of 35 markers. Marker coverage was lowest (15 markers) on chromosome 9.

In F_2-2007, three markers deviated from the expected 1:1 ratio for unlinked loci.

Table 3.4.3 QTLs found in at least two of 10 F_2BC_1 families grown in 2008. The genetic position refers to the IBM neighbors 2008 reference map and covers the overlap of QTLs found in different families.

Bin	Genetic position [cM]	F_2BC_1 families	Origin of *Rf* allele
1.02	160-170	73-19, 75-19	K55
1.10	895	43-11, 75-19	K55
2.04	313	75-12,75-19	K55
3.05	343	73-19,75-19	B37C
4.06	334-417	103-11, 104-11, 43-11, 75-12	K55
5.02	150-188	104-11, 43-11, 75-12	K55
6.05	354-371	103-11, 104-11, 73-19	K55
8.03	247	73-12, 75-12	K55
8.06-8.08	451-535	43-11, 73-12, 75-12	K55
10.04	288	43-11, 75-12	K55

They revealed three putative QTLs for partial restoration in the chromosomal bins 3.05 (340 cM), 3.06 (480 cM) and 8.02 (132 cM). The sterile bulk always carried the allele of the maternal parent B37C, whereas fertile and partially fertile bulks were heterozygous. A large number of QTLs was revealed in F_2BC_1 families except in the families 44-08 and 74-10 (Fig. 3.4.1). Ten genomic regions harbored QTLs, which collocated in at least two families (Table 3.4.3). The QTLs often covered large chromosomal segments, which stretched across as many as five bins on the reference map, as in the case of family 75-12 on chromosome 4 (Fig. 3.4.1). The detected QTL in bin 8.02 in F_2-2007 was not found in any of the F_2BC_1 families. In two families, QTLs were detected in bins 3.05 and 3.06, which might correspond to the QTLs detected in F_2-2007. However, at one of these loci (bin 3.05, 343 cM), the restorer allele did not originate from the paternal parent K55 but from the maternal parent B37. Further QTLs, whose restorer alleles were provided by the maternal parent, were detected in bins 4.09 and 9.01.

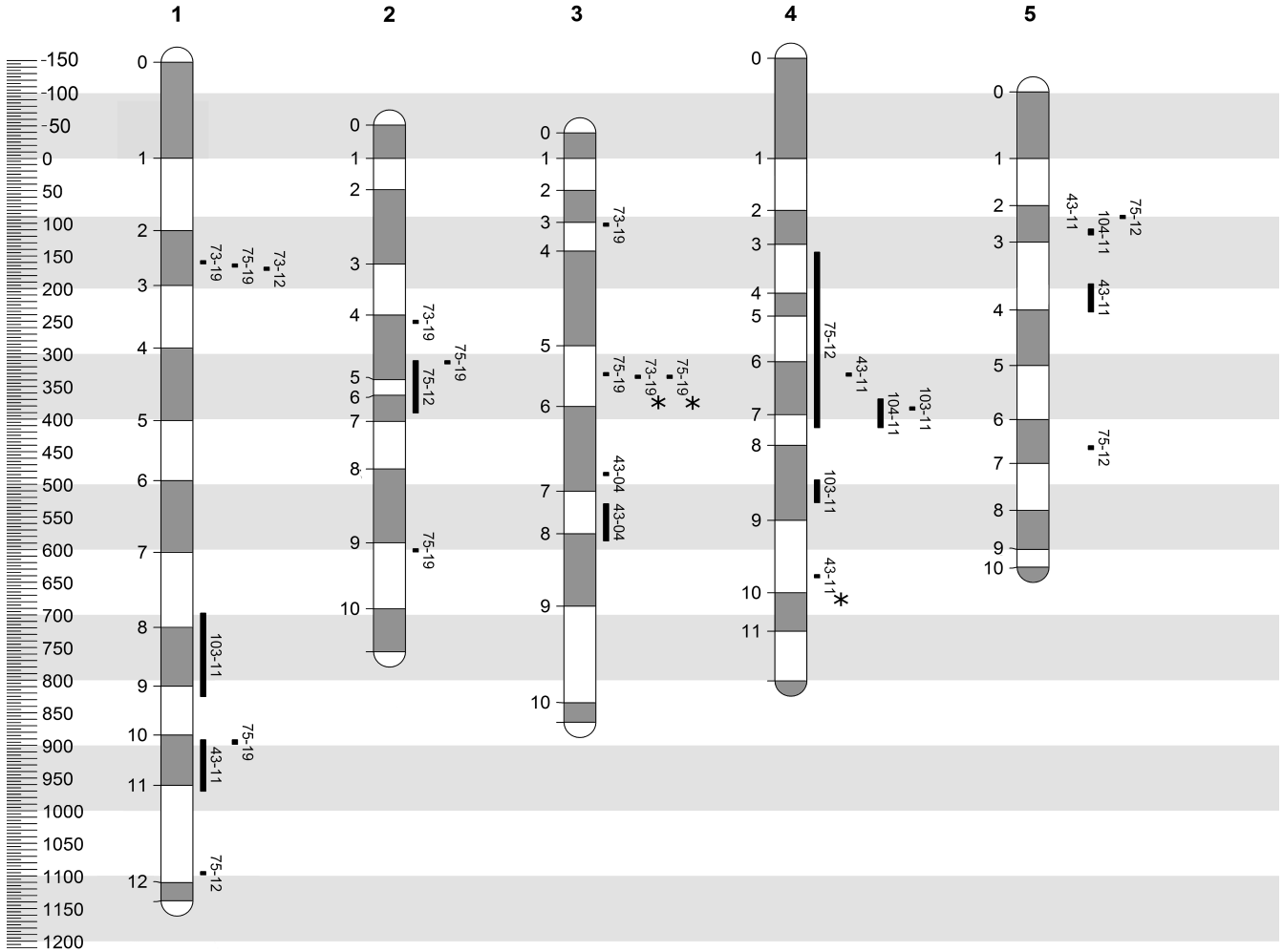

Fig. 3.4.1 QTLs detected by Bulk Segregant Analysis in F_2BC_1 families (43-04,43-11, 73-12, 73-19, 75-12, 75-19, 103-11, 104-11). * Restorer allele originates from male-sterile maternal parent B37C. The cM scale refers to the IBM neighbors 2008 reference map. The segmentation of the chromosomes refers to chromosomal bins.

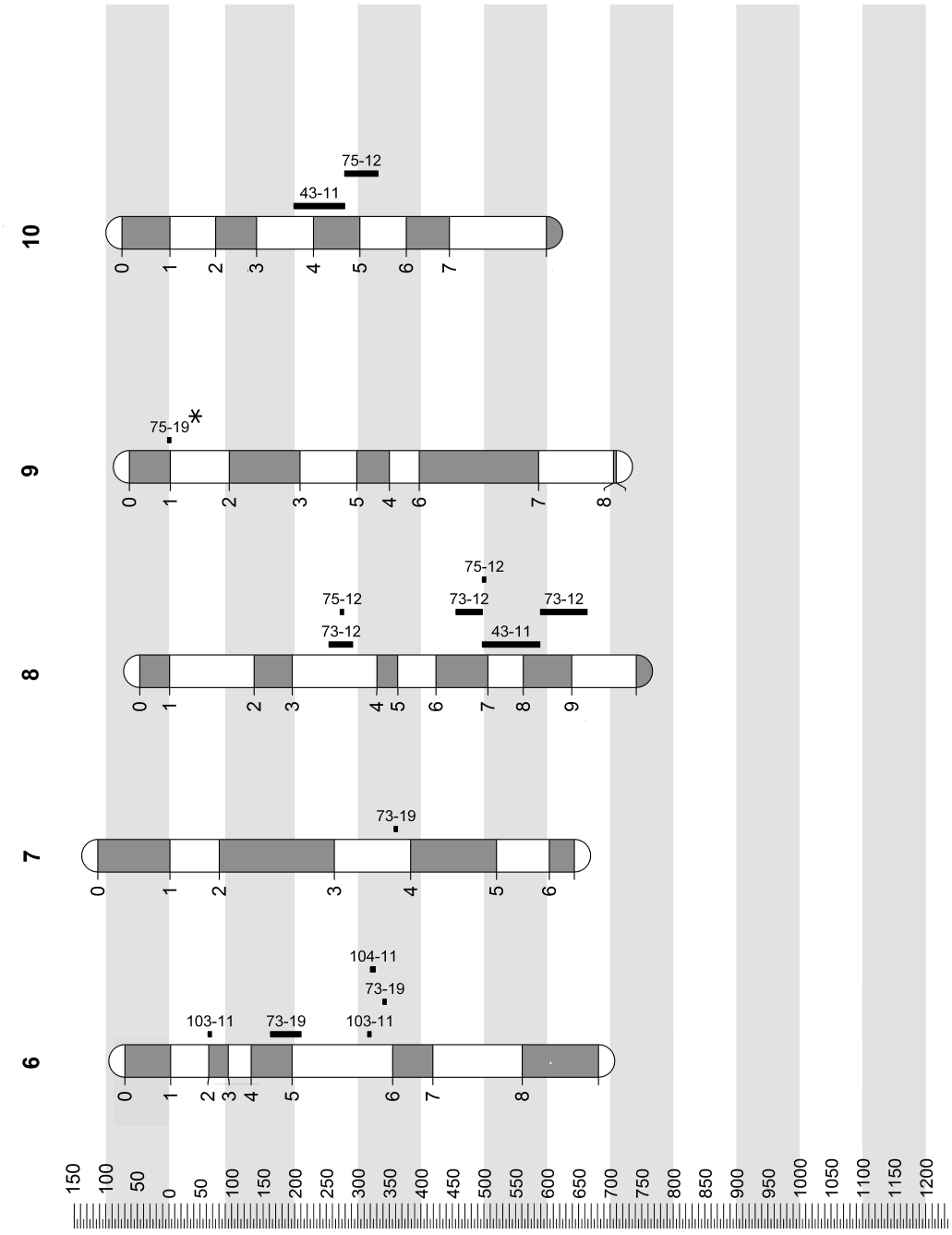

Fig. 3.4.1 Continued.

3.5 Discussion

3.5.1 High heritability of partial restoration of fertility

Within-location heritabilities of anther quality, anther quantity and ASI were high. This reflects the accuracy and reproducibility of the scoring method used to evaluate the partial restoration of male fertility in the field. In contrast to the fertility-related traits, heritability of plant height was lower. Hence, growing conditions were different in the field, but did not affect the restoration of fertility. This corroborates the results of Kaul (1988), who found that edaphic factors are less important for the expression of male fertility than climatic factors, particularly temperature and photoperiod, which are not expected to vary widely at a single location. The environmental impact would, thus, have to be measured in a multiple-location trial.

3.5.2 Multiple QTLs revealed by BSA

The high number of QTLs revealed by BSA in F_2 and F_2BC_1 families corresponds to the complex phenotypic segregation of partial restoration as well as to previous reports, which assumed a multigenic inheritance of the trait (Vidakovic 1988; Tracy et al. 1991). QTLs detectable by F_2 presumably had a major effect. The QTL in bin 8.02 might correspond to the gene *Rf4*, that fully restores fertility (Chapter 2). However, not only the fully fertile bulks, but also the partially restored bulks were heterozygous, although they were expected to carry the non-restoring allele *rf4*. The *Rf4* allele might be present in the partially restored bulks because of recombination, or the *Rf4* and the *rf4* alleles may not have been clearly separated in the bulks because of unclear phenotypes.

More loci were detected in F_2BC_1 families than in F_2, because the residual genetic variance is lower within QTL classes of F_2BC_1 families compared to F_2. Some QTLs, revealed in F_2BC_1, stretched over several bins, as was the case in the F_2BC_1 family 75-12. Hypothetically, such long intervals harbor several linked restorer genes, the single effects of which cannot be detected by BSA. Furthermore, the length of QTLs was increased, because the backcross decreased the chance of recombination by half.

3.5.3 The maternal parent is a source of restorer alleles

The genotypes of the sterile and partially fertile F_2BC_1 bulks revealed that not only the restorer parent K55, but also the maternal parent B37 possessed restorer alleles. Therefore, the backcross of F_2 plants with B37N must have again introduced restorer alleles. This provides a good explanation why three out of 10 F_2BC_1 families were on average more fertile than their respective F_2 parent. It remains unclear if partial restorations of B37C can be attributed to the QTLs in bins 3.05, 4.09 and 9.00. Partial restorations of B37C have never been reported by the Maize Genetics Cooperation Stock Center, but the emergence of some anthers filled with pollen was occasionally observed under field conditions in Serbia and Missouri (USA) (M. Vidakovic, K. Newton, pers. comm., 2009). Presumably, the expression of restorer alleles in B37C is usually suppressed in the environment of its origin, but supported by the cool and humid climatic conditions in Switzerland. Such conditions are assumed to be conducive to the expression of partial restoration (Duvick 1956; Gontarovskii 1974; Tracy et al. 1991).

3.5.4 Conclusions

The BSA yielded several major QTLs involved in the partial restoration. This is a promising result for further mapping studies on partial restorer genes. However, the approach could not give more than a general idea about the complexity of the inheritance. Nothing can be said about the relative importance of the QTLs found in F_2BC_1 families, because a random selection of QTLs segregated in the progenies. Furthermore, some QTLs might have remained undetected because of the small sample size of F_2BC_1 families. A QTL analysis would therefore be a way to decipher the inheritance of the partial restoration of male fertility.

4 QTLs involved in the partial restoration of male fertility of C-type cytoplasmic male sterility in maize[1]

4.1 Abstract

Partial restoration of male fertility limits the use of C-type cytoplasmic male sterility (C-CMS) for the production of hybrid seeds in maize. Nevertheless, the genetic basis of the trait is still unknown. Therefore, the aim to this study was to identify genomic regions that govern partial restoration by means of a QTL analysis carried out in an F_2 population (n=180). This population was derived from the Corn Belt inbred lines B37C and K55. F_2BC_1 progenies were phenotyped at three locations in Switzerland. Male fertility was rated according to the quality and number of anthers as well as the anthesis-silking interval. A weak effect of environment on the expression of partial restoration was reflected by high heritabilities of all fertility-related traits. Partial restoration was inherited like an oligogenic trait. Three major QTL regions were found consistently across environments in the chromosomal bins 2.09, 3.06 and 7.03. Therefore, a marker-assisted counter-selection of partial restoration is promising. The maternal parent was clearly involved in the partial restoration, because the restorer alleles at QTLs in bins 2.09, 6.04 and 7.03 originated from B37. The three major QTL regions collocated with other restorer genes of maize, a phenomenon, which seems to be typical for restorer genes. Therefore, a study of the clusters of restorer genes in maize could lead to a better understanding of the evolution

[1]Related publication: Kohls S, Stamp P, Knaak C, Messmer R (2010) QTLs involved in the partial restoration of male fertility of C-type cytoplasmic male sterility in maize. Theor Appl Genet, submitted

and function of restorer genes. In this respect, the long arm of chromosome 2 is particularly interesting, because it harbors restorer genes for the three major CMS systems (C, T and S) of maize.

4.2 Introduction

C-type cytoplasmic male sterility (CMS) is today the most widely applied form of CMS for the production of hybrid seeds in maize. However, an important shortcoming of C-CMS is the frequent occurrence of partially restored plants, which bear fewer anthers and shed less pollen than fully fertile plants. Anthers are often misshaped and usually emerge after silking (Tracy et al. 1991). Partially restored inbred lines cannot be used as seed parents, because even a low level of pollen shedding can lead to self pollination of the maternal parent and, consequently, to impure hybrid seeds. Furthermore, if CMS were employed to contain transgenic pollen, it would be necessary to prevent partial restoration (Munsch et al. 2009).

The environment apparently has a strong impact on the expression of partial restoration (Duvick 1956; Tracy et al. 1991; Weider et al. 2009). A C-CMS inbred line may be sterile at one location or in a particular year but may shed considerable amounts of pollen under different environmental conditions. In general, climatic factors like temperature and photoperiod are assumed to be of major importance (Kaul 1988). Humid and cool environments are assumed to be conducive to the expression of male fertility, whereas in dry and hot locations, sterility is maintained (Duvick 1965; Tracy et al. 1991). Even though partial restoration has important practical implications, its genetic basis is still unknown. In previous studies, the strong environmental effect made it impossible to draw conclusions about the number of involved genes (Tracy et al. 1991). However, partial restoration is clearly a heritable trait, because selection of progenies with a reduced partial restoration is usually successful over few generations (Gontarovskii 1974; Tracy et al. 1991). Partial restoration might be governed by multiple factors, which may have an effect in the absence of genes responsible for full restoration (Tracy et al. 1991; Has 2002). Complementary interactions between paternal and maternal factors probably play an important role (Gontarovskii 1974; Vidakovic 1988; Sotchenko et al. 2007).

To counter-select partial restorer alleles in the breeding germplasm, it is important to gain a better understanding of the mode of inheritance. If major genomic loci for partial restoration were revealed, marker-assisted selection would enhance the

development of maternal inbred lines, because the trait might not be expressed every year or at each location, depending on the environmental conditions. The aim of this study was, therefore, to conduct a QTL analysis to identify major genomic regions involved in the partial restoration of male fertility.

4.3 Materials and methods

4.3.1 Plant material

Mapping of QTLs was carried out in an F_2 mapping population, derived from a cross between the C-CMS inbred line B37C and the restorer inbred line K55. Seeds of the parental lines were obtained from the Maize Genetics Cooperation Stock Center at the University of Illinois (Urbana Champaign, IL, USA). The mapping population segregated for the fully restoring gene *Rf4* and an unknown number of genes, which partially restore male fertility. 198 F_2 plants, carrying the non-restoring *rf4* allele, were selected by means of a closely linked marker (proprietary to KWS Saat AG, Einbeck, Germany) to eliminate the masking effect of the dominant *Rf4* allele on the partial restorer genes. This marker could now be replaced by new, publicly available markers described in Chapter 2. Phenotypic values for QTL mapping were obtained from F_2BC_1 progenies, which were generated by backcrossing the selected F_2 plants to the male-fertile version of the maternal line B37 carrying a normal cytoplasm (B37N). This approach was comparable to the $F_{2:3}$ design described by Soller et al. (1990), the difference being that backcross progenies rather than selfing progenies were generated, because partially fertile and sterile plants produce little or no pollen.

4.3.2 Field experiments

The F_2 population was grown at the experimental station of the Swiss Federal Institute of Technology Zurich at Eschikon in 2008. 25 kg/ha P and 165 kg/ha K were applied before sowing, and 120 kg/ha N were applied in two portions of 60 kg/ha three and six weeks after sowing. Kernels were sown at intervals of 20 cm within a row and 75 cm between the rows. The plants were germinated under a fleece in order to accelerate early development. Leaf samples were taken at the five-leaf stage from each plant to extract DNA. In 2009, the F_2BC_1 progenies of 198 selected F_2 plants were grown in Cadenazzo (CAD), Delley (DEL) and Eschikon (ESH). Table 4.3.1 gives the descriptors for these environments. The environments were chosen because of their different climatic conditions. CAD, located on the Magadino plain of Ticino (southern Switzerland), is characterized by a higher mean temperature, higher humidity and more rainfall than DEL, which is located in the western part of the country. ESH is located in northern Switzerland and has a lower

Table 4.3.1 Description of the experimental locations Cadenazzo (CAD), Delley (DEL) and Eschikon (ESH).

	CAD	DEL	ESH
Coordinates	46°8'N, 8°55'E	46°55'N, 6°57'E	47°26'N, 8°40'E
Elevation [m a.s.l.]	203	500	546
Soil	sandy colluvisol	chromic luvisol	eutric cambisol
Sowing date	08.05.2009	22.04.2009	01.05.2009
Mean precipitation[a] [mm]	887	796	828
Mean temperature[a] [°C]	19	16	16

[a] Long-term mean during the vegetation period from April - September
[b] Soil classification according to IUSS Working Group WRB (2007)

mean temperature than CAD and more rainfall than DEL.

All the experiments were laid out as alpha (0,1) lattices with two replications, 20 blocks and 10 plots per block. The F_2BC_1 families were tested together with two entries for the parental lines B37C and K55. At ESH and CAD, 20 plants were sown in one plot, whereas at DEL, 16 plants were sown per plot. Seeds at DEL and ESH were sown with a plot seeder, whereas at CAD, the seeds were sown manually with jab planters. At DEL, plants were germinated under a fleece. At CAD, plants were irrigated with 30 mm of water at the six-leaf stage. Fertilization and plant protection were carried out according to the local practice.

4.3.3 Phenotyping male fertility

Male fertility of each F_2 and F_2BC_1 plant was rated according to the quality and number of anthers as well as the anthesis-silking interval (ASI) (Chapter 2). Plant height was considered to be a general parameter of plant vitality. ASI was defined as the difference between the date of silking and anthesis. The quality of the anthers (anther quality) was rated on a scale from 1 to 6, 1 representing normal, dehiscent anthers and 6 a lack of anthers (male sterility). The number of anthers (anther quantity) was rated on a scale from 1 to 6, with 1 representing more than 75 % anther emergence and 6 male sterility. Plants were considered to be partially fertile when anther quality and quantity had scores equal or higher than 3. Sterile plants did not bear anthers during the first three weeks after silking. In all experiments, only the 10 plants located in the center of a plot were phenotyped to minimize border effects on time of flowering. Observations were made in the morning when

the anthers were fresh and filled with pollen. The tassel of each plant was examined every other day, and all the plants were phenotyped three to four times during male flowering. The final scores for anther quality and quantity of a plant corresponded to the average quality scores and the last quantity score, respectively.

4.3.4 Statistical analysis of field experiments

The experimental designs were evaluated by means of the R package ASReml (Butler et al. 2007; R Development Core Team 2009). Missing values were estimated as a consequence of fitting the model. Analyses of variance and the best linear unbiased predictors (BLUPs) of each F_2BC_1 family were calculated from the plot raw data for anther quality, anther quantity, ASI and plant height in each individual trial as well as across environments. The mixed effects model used to analyze each individual trial was:

$$y_{ijk} = \mu + f_i + r_j + b(r)_{jk} + (fr)_{ij} + e_{ijk} \tag{4.1}$$

where y_{ijk} is the predictor of the i-th family (f) in the j-th replication (r) and the k-th block (b) within the j-th replication, μ is the intercept and e_{ijk} the residual error variance. Replications were assumed to be fixed effects and incomplete blocks and families were assumed to be random. Within-location heritabilities (sometimes referred to as "repeatabilities") were calculated according to:

$$h_w^2 = \frac{\sigma_f^2}{\sigma_f^2 + \frac{\sigma_{fr}^2}{R} + \frac{\sigma_e^2}{N}} \tag{4.2}$$

where h_w^2 is the within-location heritability, σ_f^2 is the variance of the families, σ_{fr}^2 the interaction variance of families $\times$ replications and σ_e^2 the residual error variance. R and N denote the number of replications and measured plants per plot. The model for the analysis across environments was:

$$y_{ijkl} = \mu + f_i + l_j + (fl)_{ij} + r(l)_{jk} + b(lr)_{jkl} + (flr)_{ijk} + e_{ijkl} \tag{4.3}$$

where y_{ijkl} is the predictor of the i-th family in the j-th environment (l), in the k-th replication and in the l-th block within the k-th replication, μ is the intercept and e_{ijkl} the residual error variance. Environments were assumed to be fixed, all the other factors were assumed to be random. The heritability was calculated according to:

$$h^2 = \frac{\sigma_f^2}{\sigma_f^2 + \frac{\sigma_{fl}^2}{L} + \frac{\sigma_{flr}^2}{LR} + \frac{\sigma_e^2}{N}} \tag{4.4}$$

where h^2 is the heritability, σ_f^2 the variance of the families, σ_{fl}^2 the interaction variance of families × environments, σ_{flr}^2 the interaction variance of families × environments × replications and σ_e^2 the residual error variance. L, R and N denote the number of environments, replications and measured plants per plot.

The phenotypic coefficients of correlation (Pearson's) and significance levels were determined by linear regressions in R based on standardized BLUPs (mean 0, standard variation 1).

4.3.5 Genotyping

180 of 198 F_2 plants were randomly selected for genotyping. A whole genome scan was carried out by Traitgenetics GmbH (Gatersleben, Germany) using a 1,536 SNP Illumina GoldenGate array (OPA GS0010903; Illumina Inc., San Diego, CA, USA) and whole-genome amplified DNA. The SNPs covered 512 loci that were evenly distributed over the maize genome (three SNPs per locus, belonging to different haplotypes). Furthermore, 672 SNP markers were analyzed on the Illumina Veracode BeadXpress platform multiplexing 384 SNPs in one reaction. The GoldenGate assays were performed according to the manufacturer's protocol and as described in Fan et al. (2003). The automatic allele calling for each locus was accomplished with the Illumina GenomeStudio Analysis Software. Both alleles of a SNP give different signals due to the allele-specific ligation to fluorescent products. Therefore, a genotype that is homozygous for one or the other SNP alleles gives one signal, whereas a heterozygous genotype gives both signals. All SNPs were checked for good separation of the homozygous and heterozygous signals and were manually re-scored when errors in calling the homozygous or heterozygous signals were found. 330 polymorphic high-quality SNP markers were chosen for linkage mapping.

4.3.6 Construction of the genetic linkage map

The linkage map was constructed with the software Joinmap 4.0 (van Ooijen 2006). Five of 180 individuals were excluded from this procedure because of an elevated mean number of crossing overs per individual. This was probably due to genotyping errors. χ^2 values were calculated for all individual markers to detect deviations in

gametic segregation from the expected Mendelian 1:2:1 ratio for an F_2 population (α=0.05). Bonferroni correction was considered for multiple tests. A minimal LOD (log likelihood of the odds) of 3.0, a maximal recombination fraction of 0.5 and a maximal jump of 5 were used for clustering of markers. After construction of the map, the data files were screened for putative double recombinants based on the "genotype probabilities" calculated by Joinmap. Genotypes with a low probability were verified or corrected by reexamining the genotypic data. Once the most probable order of the markers was obtained, genetic distances were estimated for each linkage group applying the Kosambi mapping function (Kosambi 1944). The genetic positions of all SNP markers were projected onto the IBM neighbors 2008 reference map (Schaeffer et al. 2008).

4.3.7 QTL analysis

Composite interval mapping (CIM) (Zeng 1994) was performed with the software PLABQTL (Utz and Melchinger 2000) for each individual trait in each environment (individual QTL analyses) as well as in a combined approach across environments (joint analyses). The minimal LOD required to declare a QTL significant was obtained empirically for each trait by 1,000 permutation tests (Churchill and Doerge 1996). The resulting LOD thresholds (α=0.05) varied from 4.45 to 4.71, depending on the trait, with an average of 4.70. The latter was applied as the significance threshold for the detection of QTLs in all analyses. CIM was used with an "F-to-enter" value of 3.5 for the step-wise regression to preselect cofactors. Subsequently, the "cov SELECT" option was applied, which takes all the preselected markers as cofactors. However, if closely adjacent markers were selected as cofactors, then one of them was removed manually. The support interval of a QTL was defined as the length of the segment of the chromosome, over which the LOD at the peak decreased by half.

The effect of genotypic sampling on the estimation of QTLs was tested by 200 rounds of five-fold cross validation using the "cross-validate" option of PLABQTL. In each round, positions and effects of QTLs of 80 % of the genotypes were estimated, and a validation was performed with the remaining 20 %.

The percentage of phenotypic variance explained by all QTLs identified for a given trait was estimated by the adjusted coefficient of determination (adjusted R^2) (Hospital et al. 1997). The percentage of phenotypic variance explained by an individual QTL was calculated by the normed partial correlation coefficient of

determination (npartR^2). With such a normalization, the sum across all QTLs is equal to the adjusted R^2 (Zhu et al. 2004).

The additive effect of a QTL was calculated as the median of the additive effects obtained from 1,000 cross validation runs, taking into account only those QTLs, whose peaks were located within the LOD support interval obtained by CIM. The additive effects as such were calculated as half the difference between the mean of the homozygous *RfRf* class and the mean of the homozygous *rfrf* class. However, since the genetic structure of F_2 parents and their F_2BC_1 progenies differed, the additive effects were underestimated by half. The explanation is that the F_2BC_1 progeny of an F_2 plant with the genotype AA at a given locus is supposed to be 100 % heterozygous (Aa), and the progeny of an F_2 plant, heterozygous at a given locus (Aa), is supposed to segregate into 50 % heterozygous (Aa) and 50 % homozygous (aa) plants. Therefore, the additive effects obtained from PLABQTL were corrected by multiplying the values by two. Dominance effects could not be tested in the backcross design of this experiment. Epistatic interactions between additive effects were analyzed in ICIM (Li et al. 2007), which detects interactions between QTLs with insignificant main effects.

4.4 Results

4.4.1 Climatic conditions

Temperature during the vegetation period in 2009 was close to the long-term mean in all environments (Table 4.4.2). CAD received the highest temperature. Rainfall differed widely among locations. CAD had more rainfall than the long-term mean and three times more than at DEL, the driest location. At DEL and ESH, there was less rainfall in spring and during flowering than the long-term mean. The patterns of rainfall at DEL and ESH were similar, and rainfall was more evenly distributed than at CAD where rainfall was seldom but very intense. At CAD, a wet early season, resulting in loss of plants due to water logging, was followed by a dry period in May and June, during which irrigation was required.

Table 4.4.2 Monthly mean temperature and precipitation at Cadenazzo (CAD), Delley (DEL) and Eschikon (ESH) in 2009 and in the long-term (lt) mean.

	CAD		DEL		ESH	
	2009	lt mean	2009	lt mean	2009	lt mean
Temperature [°C]						
April	13	13	11	10	12	10
May	19	17	16	14	16	14
June	20	22	17	18	17	18
July	22	23	19	20	19	19
August	22	21	20	18	20	18
September	19	18	16	15	16	15
Mean	**19**	**19**	**17**	**16**	**17**	**16**
Precipitation [mm]						
April	296	136	18	78	22	115
May	45	114	30	83	102	136
June	213	124	84	68	144	116
July	251	185	98	82	176	167
August	240	182	53	108	99	202
September	93	147	36	77	58	92
Sum	**1138**	**887**	**319**	**496**	**601**	**828**

4.4.2 The maternal parent was partially restored

The maternal inbred line B37C was partially restored in both years of the study. The level of restoration differed from plant to plant but full restoration was never achieved. As a spot check, pollen of partially restored B37C plants was examined under the microscope. Approximately half the pollen grains had aberrant shapes; the rest seemed to be viable (data not shown).

4.4.3 Segregation of male fertility and heritabilities

Male flowering started at temperature sums between 600 and 700 and lasted for about four weeks in all environments. At CAD, anthesis started on July 19, in DEL on July 16 and in ESH on August 1. The majority of the 198 F_2BC_1 families segregated into partially restored and sterile plants. Therefore, the masking effect of the *Rf4* allele was in most cases successfully eliminated by marker-assisted selection. Ten families were fully fertile, due either to recombination between *Rf4* and the marker used for selection or due to other genetic constellations. None of the families was fully sterile.

The expression of partial restoration was similar among environments. The means of anther quality and anther quantity were insignificantly different among the individual trials, ranging from 3.4 to 3.7 for anther quality and from 2.6 to 3.1 for anther quantity (Fig. 4.4.1). Protogyny was strongly expressed in all environments, as it is a characteristic of partially restored plants. The average ASI was significantly longer at DEL (11 days) than at CAD and ESH (nine days), although the range of the values was similar. In contrast to the fertility-related traits, the mean plant height varied clearly among the environments, from 220 cm at DEL to 254 cm at ESH. The frequency distribution of all the traits followed approximate bell-shaped curves in all environments (Fig. 4.4.1). This suggests that multiple genes were involved in the partial restoration. h_w^2 of the fertility-related traits were higher than 0.80, except that of anther quality at DEL (0.68). h_w^2 of plant height was 0.77 at DEL and ESH but only 0.53 at CAD, probably resulting from an unequal development of the plants due to water logging after sowing.

The analysis of variance across families and environments gave highly significant differences among families for all traits. The variances of the location × family interaction and location × replication × family interaction were much smaller, however still significant (data not shown). Accordingly, h^2 of all the traits was high, but the

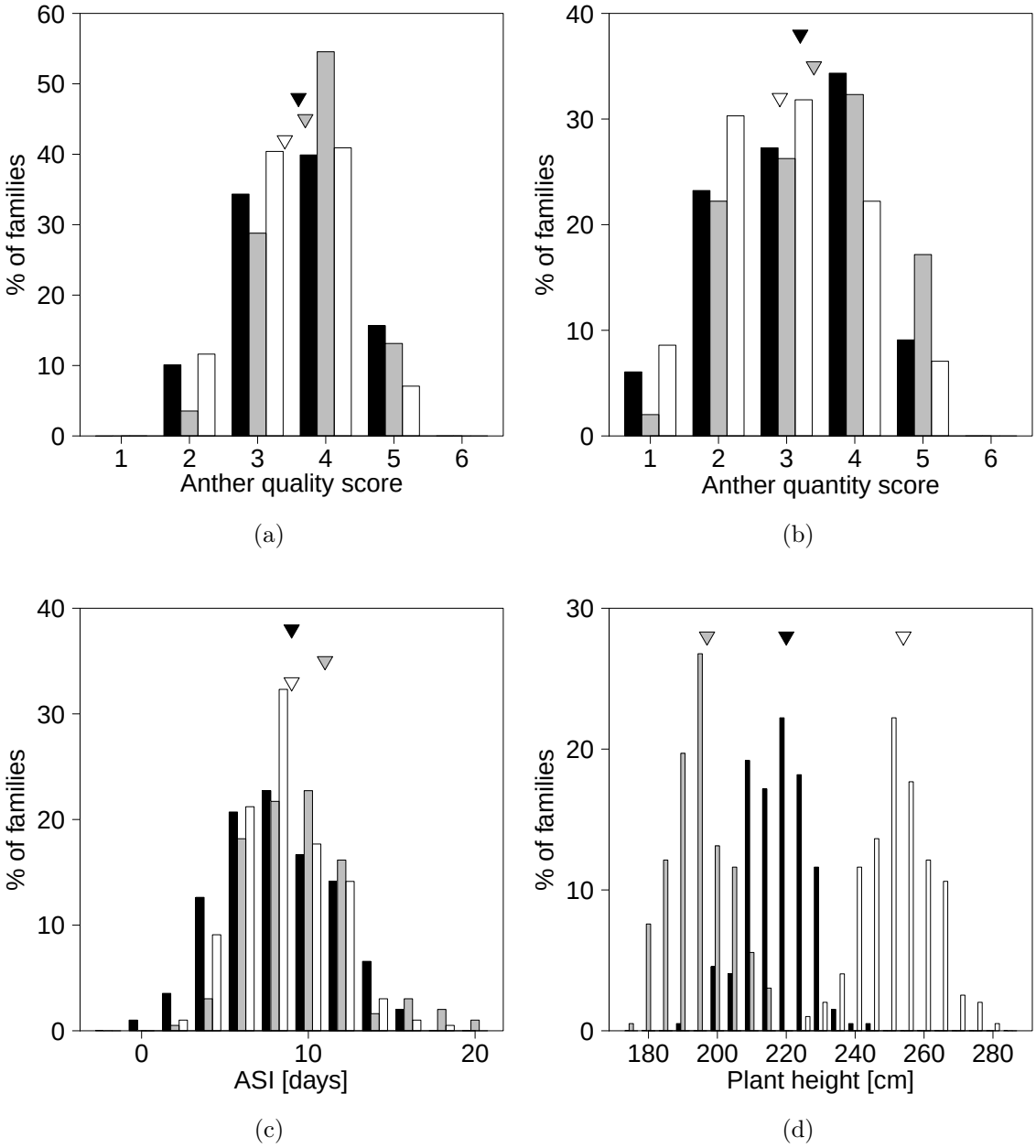

Fig. 4.4.1 Frequency distribution of F_2BC_1 family-based best linear unbiased predictors for (a) anther quality, (b) anther quantity, (c) the anthesis-silking interval (ASI) and (d) plant height at Cadenazzo (black), Delley (gray) and Eschikon (white). Means of individual trials: ∇.

h^2 of fertility-related traits (0.83-0.84) were higher than that of plant height (0.70).

Anther quality, anther quantity and ASI were closely correlated with each other in all environments. Whereas the phenotypic coefficients of correlation between anther

quality and anther quantity was above 0.9, the correlations with ASI were slightly lower but still above 0.8. The positive correlation between anther quality and anther quantity showed that a normal fertile appearance of anthers was accompanied by a high number of anthers. Partial fertility was associated with protogyny, as indicated by the negative sign of the correlations between ASI and anther quality and quantity.

4.4.4 Linkage map and QTL results

The genetic map covered a total length of 1,322 cM. The average inter-marker distance was 4.0 cM with a maximum distance of 28.3 cM. 3.8 % of the genotypic data were missing. Table 4.4.4 gives an overview over all detected QTLs. In the joint analysis, the QTLs for anther quality, anther quantity and ASI together explained 72.2, 70.2 and 62.3 % of the phenotypic variance (adjusted R^2). In the individual trials, the adjusted R^2 values were in the same range. The QTLs for anther quality and anther quantity usually explained more of the phenotypic variance than the QTLs for ASI (data not shown). The QTL analysis revealed three major QTL regions in the chromosomal bins 2.09, 3.06 and 7.03 (Fig. 4.4.2). In these genetic regions, QTLs collocated, with a few exceptions, across all three fertility-related traits and all three environments. In most cases they explained more than 10 % of the phenotypic variance (npartR^2) and exceeded the LOD threshold of 4.7 by far. LOD scores as high as 30 suggest that partial restoration was inherited oligogenically. The restorer alleles of the QTLs in bins 2.09 and 7.03 did not originate from the restorer parent K55 but from the maternal parent B37. The direction of effects was consistent across traits and environments (Table 4.4.3).

As well as the three major QTL regions, other QTLs for anther quality, anther quantity and the ASI were revealed on chromosomes 3, 4, 5, 6, 8 and 9 (Fig. 4.4.2). These QTLs are considered to be minor QTLs, because their effects usually explained less than 10 % of the phenotypic variance. Although the detection of minor QTLs was less consistent across traits and environments than that of the major QTLs, clustering was observed in bins 3.05, 4.06, 4.07 and 8.06. However, QTLs on chromosome 8 should be regarded with caution, because the fixation of the *rf4* allele might bias the QTL result. The restorer alleles of all the minor QTL regions, except that on chromosome 6, originated from the restorer parent K55 (Table 4.4.3).

Minor QTLs for anther quality were mapped to bins 3.05, 4.06 and 8.06 in the joint analysis and, with a few exceptions, in the individual trials as well. QTLs for anther quality found in individual experiments but not in the joint analysis were

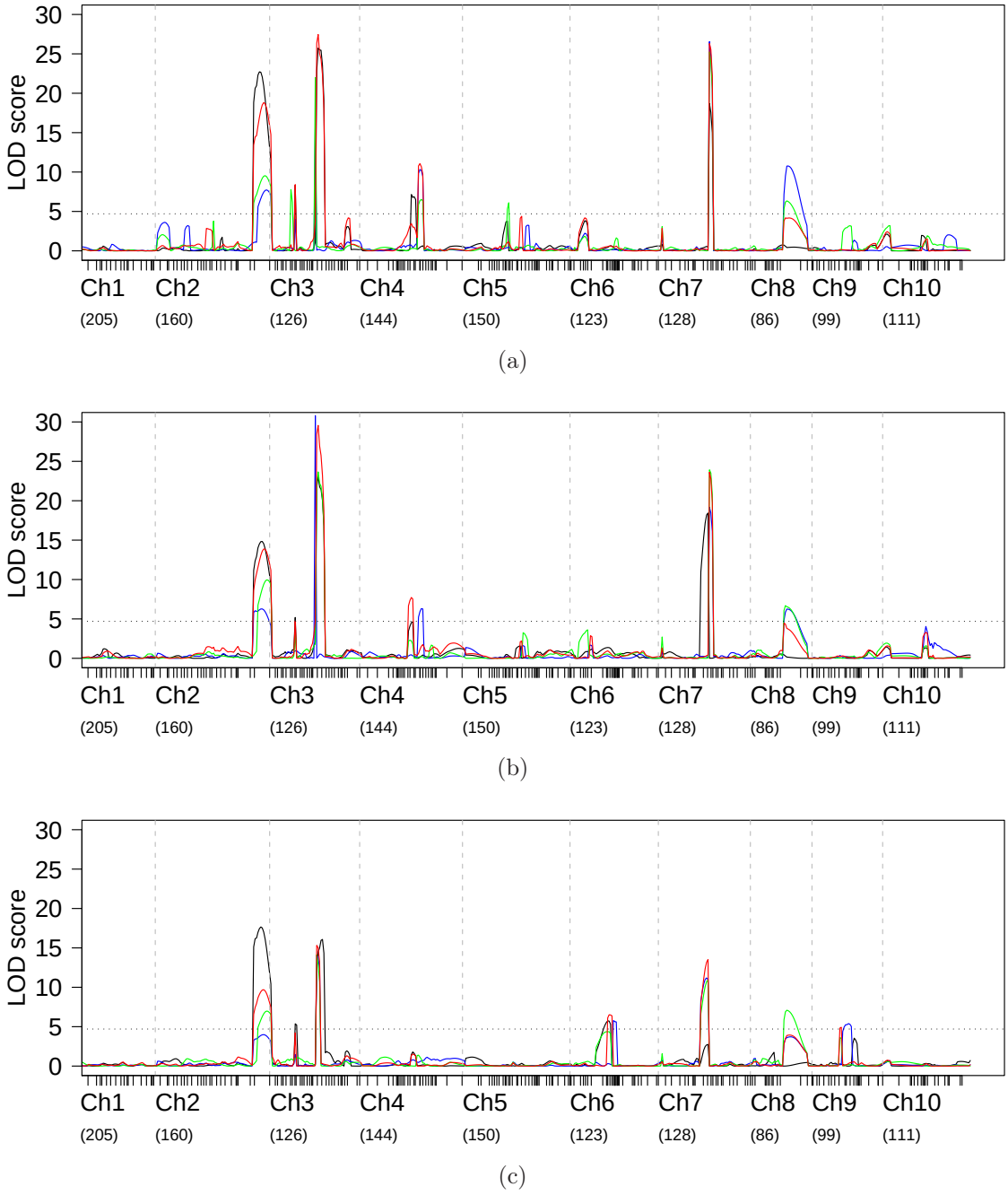

Fig. 4.4.2 LOD profiles for (a) anther quality, (b) anther quantity and (c) the anthesis-silking interval (ASI) at Cadenazzo (black), Delley (blue), Eschikon (green) and of the joint analysis (red). The dashed vertical lines indicate the significance threshold (LOD=4.7). The length (cM) of each chromosome (Ch) is given in parentheses.

located in bin 4.07 (DEL, ESH) and bin 5.03 (ESH). The latter QTL was specific for anther quality. QTLs for anther quantity were detected in the same regions as

for anther quality, except in bin 5.03. Bins 3.05 and 4.06 harbored QTLs only in the joint analysis, whereas QTLs in bin 8.06 were found at DEL and ESH but not in the joint analysis. The QTLs for ASI differed in part from the QTL regions found for anther quality and anther quantity. No QTLs were detected on chromosome 4, and QTLs in bins 6.04 and 9.03 were significant for ASI only (DEL and joint analysis). QTLs in bins 3.05 (CAD) and 8.06 (ESH) collocated with QTLs found for anther quality and anther quantity.

The LOD support intervals for most major and minor QTLs ranged from 8 to 18 cM around the respective peak. However, the support intervals for the QTLs in bin 8.06 were larger, covering 22 to 28 cM. Numerous digenic interactions among QTLs

Table 4.4.3 Estimated additive effects and origin of the restorer (*Rf*) allele of the QTLs found for anther quality, anther quantity and the anthesis-silking interval (ASI) in Cadenazzo (CAD), Delley (DEL), Eschikon (ESH) and in the joint analysis.

		Additive effect				Origin of
Trait	Bin	CAD	DEL	ESH	Joint analysis	*Rf* allele
Anther quality	2.09	1.1	0.5	0.6	0.8	B37
	3.05	-0.6	-	-0.5	-0.5	K55
	3.06	-1.3	-0.8	-0.8	-1.1	K55
	4.06	-0.5	-0.5	-0.5	-0.5	K55
	5.03	-	-	-0.4	-	K55
	7.03	0.9	0.8	1.0	0.9	B37
	8.06	-	-0.5	-0.5	-	K55
		-				
Anther quantity	2.09	1.1	0.7	0.9	1.0	B37
	3.05	-	-	-	0.6	K55
	3.06	-1.5	-1.6	-1.2	-1.6	K55
	4.06	-	-0.7	-	-0.6	K55
	7.03	1.2	1.3	1.3	1.3	B37
	8.06	-	0.7	0.7	-	K55
ASI	2.09	4.4	-	2.7	2.9	B37
	3.05	-2.6	-	-	-	K55
	3.06	-4.7	-4.2	-3.2	-3.7	K55
	6.04	2.1	2.3	-	2.1	B37
	7.03	-	3.3	2.7	3.2	B37
	8.06	-	-	-2.1	-	K55
	9.03	-	-2.4	-	-2.1	K55

were found throughout the genome, but all of them exceeded the LOD threshold only marginally, and the interactions across traits and environments were inconsistent. Therefore, digenic interactions were not considered. Significant QTL × environment interactions were detected for anther quality only. The interactions were significant for the QTLs in bins 3.06, 7.03 and 8.06, although these QTLs were detected consistently in most of the individual trials as well as in the joint analysis. The frequencies of QTL detection, revealed by cross validation, usually mirrored the pattern of LOD curves and confirmed the consistency of QTL detection across environments and fertility-related traits. QTL frequencies were higher for anther quality and anther quantity than for ASI. The highest QTL frequencies were found for the major QTLs in bin 7.03 (25 - 93 %, average 67 %), whereas the frequencies were lower in the other two major QTL regions (bin 2.09 15 - 34 %, average 25 %; bin 3.06 23 - 69 %, average 40 %). The QTL frequencies of minor QTLs ranged from 10 to 30 % in most cases; however, QTLs in bin 3.05 were detected with a frequencies up to 69 % (average 44 %).

Table 4.4.4 QTLs for anther quality, anther quantity and the anthesis-silking interval (ASI) at Cadenazzo (CAD), Delley (DEL) and Eschikon (ESH) as well as in the joint analysis (Joint): chromosomal bin, genetic position, LOD support interval (SI interval), approximate genomic position projected on the IBM neighbors 2008 reference map (peak IBM), LOD score and normed partial R^2 (npartR^2).

Trait	Location	Bin	Peak [cM]	SI interval [cM]	Peak IBM [cM]	LOD	npartR^2 [%]
Anther quality	CAD	2.09	142	130-158	623	22.7	16.5
		3.05	32	28-36	342	8.3	6.7
		3.06	64	60-76	515	25.7	20.1
		4.06	68	64-76	365	7.2	6.3
		7.03	64	60-72	298	18.7	17.1
	DEL	2.09	152	136-158	669	7.7	5.6
		3.06	60	56-64	494	16.1	22.4
		4.07	80	76-88	440	10.3	6.8
		7.03	64	60-72	298	26.6	22.0
		8.06	44	36-64	432	10.8	10.7
	ESH	2.09	150	130-158	661	10.0	5.8
		3.05	26	22-30	310	8.0	6.2

Trait	Location	Bin	Peak [cM]	SI interval [cM]	Peak IBM [cM]	LOD	npartR^2 [%]
		3.06	60	56-64	494	21.5	15.9
		4.07	82	74-86	446	7.4	5.1
		5.03	60	56-64	189	5.8	6.0
		7.03	64	60-72	298	27.1	20.6
		8.06	42	34-62	430	7.9	8.9
	Joint	2.09	148	132-158	651	19.8	11.2
		3.05	32	28-36	343	8.7	6.6
		3.06	64	60-76	515	30.1	18.2
		4.06	68	60-72	365	12.1	7.0
		7.03	64	60-72	298	26.9	21.7
		8.06	40	36-64	388	5.9	7.5
Anther quantity	CAD	2.09	146	130-158	643	14.8	15.1
		3.06	64	60-76	515	23.1	30.0
		7.03	62	50-66	298	18.4	17.0
	DEL	2.09	146	132-158	643	6.3	6.8
		3.06	60	56-64	494	30.8	28.0
		4.07	82	76-88	446	6.3	6.3
		7.03	64	60-72	298	19.2	21.2
		8.06	42	34-66	432	6.3	7.8
	ESH	2.09	152	136-158	669	10.0	7.1
		3.06	64	60-76	515	23.7	26.8
		7.03	64	60-72	298	23.9	22.8
		8.06	40	34-62	388	6.7	7.2
	Joint	2.09	148	132-158	651	13.9	10.7
		3.05	32	28-36	343	4.8	3.7
		3.06	64	60-76	515	29.6	25.1
		4.06	68	60-72	365	7.7	6.6
		7.03	64	60-72	298	23.6	24.1
ASI	CAD	2.09	144	132-158	634	17.7	20.7
		3.05	32	30-38	342	5.4	9.6
		3.06	70	58-74	566	16.1	16.5

Trait	Location	Bin	Peak [cM]	SI interval [cM]	Peak IBM [cM]	LOD	npartR^2 [%]
	DEL	3.06	62	58-70	498	14.1	23.0
		6.04	54	50-62	230	5.8	7.3
		7.03	60	48-64	295	11.2	14.8
		9.03	40	28-48	292	5.4	7.5
	ESH	2.09	152	136-158	651	7.0	7.8
		3.06	62	58-70	498	14.1	19.6
		7.03	62	50-66	298	10.9	15.1
		8.06	42	34-62	430	7.1	8.1
	Joint	2.09	148	132-158	651	9.7	10.9
		3.06	62	58-70	498	15.3	22.8
		6.04	48	44-56	210	6.5	6.9
		7.03	62	50-66	298	13.5	14.1
		9.03	30	26-34	256	4.9	7.6

4.5 Discussion

4.5.1 Weak effect of the environment on partial restoration

The effect of environment on partial restoration was weak. Accordingly, heritabilities for anther quality, anther quantity and ASI were high. This reflects the accuracy and reproducibility of the experimental conditions and the scoring method used to evaluate partial restoration in the field. Compared to the fertility-related traits, the heritability of plant height was lower, both in the joint analysis across environments as well as in the individual trials. The discrepancy indicates that growth conditions differed in the fields but did not impact the restoration of fertility. This corroborates the results of Kaul (1988), who found that edaphic factors are less important for the expression of male fertility than climatic factors, particularly temperature and photoperiod. The high heritabilities of fertility-related traits may be due to the relatively cool and moist climatic conditions in all the environments, because these conditions are assumed to be conducive to the expression of male fertility (Duvick 1965; Tracy et al. 1991). Therefore, Switzerland might be a suitable place to detect QTLs involved in partial restoration and to select male-sterile inbred lines. However, to quantify the effect of the environment on the expression of QTLs, locations with a continental climate and at a different latitude should be included.

There are no comparable values for the heritability of partial restoration reported in the literature. Most studies on restoration of CMS report a strong environmental impact (Tracy et al. 1991; Weider et al. 2009). In a study about the partial restoration of C-CMS in A632 background, Tracy et al. (1991) observed large differences among years, which made it impossible to infer the number of genes involved. In contrast, a study about C-CMS with a wide range of germplasm grown at two locations and in two seasons showed that the effect of environments was of minor importance compared to the nuclear background of CMS plants (Vidakovic 1988). Such inconsistent results concerning the effect of environment probably result from the specific choice of germplasm and test locations.

4.5.2 Partial restoration is inherited like an oligogenic trait

Major QTLs with very high LOD scores and stable expression across environments and fertility-related traits suggest that partial restoration was inherited oligogenically. Therefore, a marker-assisted selection of the QTLs in bins 2.09, 3.06 and 7.03

seems feasible if their importance for different germplasm and in a broader range of environments can be proven. The existence of major QTLs corroborates the field observations of Gontarovskii (1974) and Tracy et al. (1991), who reported that paired cross selection (Jones and Manglesdorf 1957) is often successful for reducing partial restoration in C- and T-CMS inbred lines over a few generations. It is assumed that this method will be successful only when few major genes govern the trait. As well as the major QTLs, there were numerous minor QTLs, which were expressed inconsistently across environments. Such loci are probably responsible for the differences in partial restoration in many studies.

All three major QTL regions found in this study map near other restorer genes of maize. The QTL region in bin 2.09 collocates with a cluster of multiple restorer genes for S-CMS, including the major restorer gene *Rf3* (Gabay-Laughnan et al. 2004) and *Rf8* and *Rf**, which restore T-CMS in presence of *Rf2* (Dill et al. 1997). The QTLs in bins 3.05 and 3.06 are linked to *Rf1* of T-CMS in bin 3.04. Sisco (1991) assumed restorer genes for C-CMS in this region, because the genomic area around the major restorer gene *Rf4* in bin 8.00 shows collinearity of the homologous sequences, as revealed by hybridization of RFLP probes (Gaut 2001). The QTL region in bin 7.03 maps near *Rf-I*, an inhibitor of *Rf5* of C-CMS and maybe identical to *Rf7*, which also acts as a partial restorer gene (Qin et al. 1990, Hu et al. 2006).

The clustering of restorer genes in maize corresponds to the complex genomic structure of restorer loci in other plants such as petunia, radish and rice, where restorer genes are found near highly homologous genes of the pentatricopeptide-repeat (PPR) family (Bentolila et al. 2002; Desloire et al. 2003; Wang et al. 2008). Furthermore, clusters of uncharacterized restorer genes were found in cotton (Zhang and Stewart 2001) and rapeseed (Li et al. 1998). Such clusters might consist of paralogous genes, which probably arise from repeated duplications through unequal crossing over (Touzet and Budar 2004). This process might facilitate the evolution of restorer genes in response to the spread of CMS in a hermaphroditic population of plants. Therefore, the high number of partial restorer genes in the maize genome might witness ancient constitutions of the mitochondrial genome, or these genes might have arisen as defective duplicates of existing restorer genes. Although linkage does not imply similarity of sequences (Osbourn 2010), linked restorer genes may have a common evolutionary origin and, thus, possibly a common mode of action. It has been hypothesized that paralogous restorer genes can diverge to restore different CMS systems (Bentolila et al. 2002; Touzet and Budar 2004; Wang et al. 2006).

This hypothesis could be tested by dissecting the genetic structure of the long arm of chromosome 2, which contains restorer genes for all three major CMS systems in maize (T, S and C).

4.5.3 The maternal parent is a source of restorer alleles

The maternal parent was clearly involved in the partial restoration of the mapping population, supporting previous results with different populations (Vidakovic 1988; Sotchenko et al. 2007). It is tempting to speculate that the QTLs in bins 2.09, 6.04 and 7.03 were responsible for the partial restoration of B37C. Whereas reversions to fertility in B37C have not been reported by the Maize Genetics Stock Center, some anthers filled with pollen were found occasionally under field conditions in Serbia and Missouri (USA) (M. Vidakovic, K. Newton, pers. comm., 2009). The expression of partial restorer alleles in B37C is probably usually suppressed at its site of origin, but favored by the climatic conditions in Switzerland.

4.5.4 Conclusions

Partial restoration of male fertility was inherited like an oligogenic trait. Therefore, a marker-assisted counter-selection of major QTLs in bins 2.09, 3.06 and 7.03 is promising, but the importance of these loci must be verified with other germplasm and under different environmental conditions. Elucidating the genetic structure of clusters of restorer genes could improve our understanding of the evolution and functioning of restorer genes. In this respect, the long arm of chromosome 2 is particularly interesting, because it contains restorer genes for all three major CMS systems in maize.

5 General conclusions

CMS is an important genetic mechanism for the production of hybrid seeds in maize. C-type CMS is today the most widely applied form of CMS, because plants carrying the T cytoplasm are susceptible to Southern Corn Leaf Blight and the sterility induced by the S cytoplasm is less stable. However, the major shortcomings of C-CMS are the unknown and complex system of restoration, a high abundance of restorer genes throughout most gene pools and a frequent partial restoration of maternal inbred lines, which is variably expressed in different environments. These factors limit the selection of maternal inbred lines. For an efficient use of C-CMS in the future, it would be desirable to select restorer genes in a way that the restoration is governed by only one fully restoring gene, which is absent in the maternal gene pools. This study offers new insights into the inheritance of the restoration of fertility of C-CMS, which will be useful for a marker-assisted selection of restorer genes.

Three experiments were conducted using mapping populations derived from the inbred lines B37C and K55. *Rf4* was fine-mapped to a 0.3 cM interval in the chromosomal bin 8.00 in a selectively genotyped F_2 population (Chapter 2). Several new, codominant and PCR-based markers are now available for the marker-assisted selection of *Rf4*. Hypothetically, *Rf4* encodes a new type of restorer protein, because predicted genes in the area flanking *Rf4* are not homologous to previously cloned restorer genes, which mainly encode PPR proteins. In Chapter 3 and 4, the inheritance of partial restoration was analyzed for the first time on the molecular level. A Bulk Segregant Analysis (BSA) in an F_2 and selected F_2BC_1 progenies was designed as a pilot experiment for a subsequent QTL study, which was carried out in 180 F_2 individuals. The QTL mapping population was selected to carry the non-restoring allele *rf4* to map partial restorer genes without the masking effect of the dominant *Rf4* allele. F_2BC_1 progenies were phenotyped at three locations in Switzerland. Major QTLs for partial restoration with stable effects across environments were revealed in bins 2.09, 3.06 and 7.03. Both paternal and maternal factors were involved in the partial restorations.

5.1 Counter-selection of partial restorer genes is promising

Counter-selection of partial restorer alleles by means of molecular markers is promising, because the inheritance was putatively oligogenic, and the environmental influence on the expression of male fertility was low in all experiments conducted from 2007 to 2009. This was revealed by a good accordance between the frequency distributions of anther quality, anther quantity and ASI between F_2-2007 and F_2-2008, by high heritabilities in the BSA study in 2008 and the three-location experiment in 2009 as well as by a lack of significant QTL × environment interactions. Contrasting results of previous studies are probably due to the specific choice of germplasm and the range of environmental conditions.

The most promising QTLs are those with the largest effects (QTLs in bins 2.09, 3.06 and 7.03), among which the locus in bin 3.06 may be the most prominent, because its effect was large enough to be detected by BSA in F_2. However, there were numerous minor QTLs, which were not consistently expressed across environments. Such loci are probably responsible for the versatile appearance of the partial restoration in many studies.

5.2 *Rf4* might interact with other restorer genes

This study found twofold evidence, yet not a proof, that *Rf4* interacts with other restorer genes, as suggested by Vidakovic (1988). *Rf4* might belong to a complex of complementary genes, because some inbred lines, which putatively carried the restoring *Rf4* allele, were male-sterile and *vice versa*. Furthermore, differences of the coefficients of correlation between anther quantity, anther quality and ASI in F_2-2008 compared to the QTL population can be interpreted as the result of an interaction between *Rf4* and partial restorer genes. Whereas correlations were very high in the QTL population, they were only medium in F_2-2008. The correlation was reduced by certain plants, which showed aspects of full fertility (for example 100 % anther emergence) but also aspects of partial restoration (for example a delayed anthesis or stunted anthers). This type of plant probably carried the *Rf4* allele, because it was absent in the QTL population fixed for *rf4*. The fully restoring effect of *Rf4* might have been reduced by interacting partial restorer genes, thereby decreasing the correlation between male fertility-related traits in F_2-2008.

5.3 Important aspects when working with CMS and restorer genes

Maternal factors involved in complete and partial restoration are wide-spread (Vidakovic 1988; Sotchenko et al. 2007). Therefore, the selection and mapping of restorer genes has to consider the maternal parent as an important source of restorer alleles. The incomplete sterility of B37C in all years of study (2007-2009) indicated the presence of restorer alleles, which was verified by the origin of restorer alleles at several QTLs as well as by the presence of at least one complementary restorer gene in B37C. Such genes could be mapped in populations derived from crosses between CMS and non-restorer lines carrying a normal cytoplasm.

The mapping of restorer genes requires a precise system to phenotype male fertility. Many previous studies on CMS used modifications of a scale proposed by Beckett (1971). This scale considers the quality and number of anthers like in this study, but, presumably in anticipation of a close correlation between the fertility-related traits, they are not assessed separately. However, the correlation between the traits in F_2-2007 and F_2-2008 was only medium. Therefore, a single, combined scale would not have described with full discriminatory precision the fertility status of the population used here.

Mapping studies about major restorer genes should be carried out in populations, which segregate in a simple fashion, because partial restorer genes complicate the mapping process. If a high level of partial restoration occurs, selective genotyping should be applied. Partial restorer genes should be mapped and counter-selected in environments, where the trait is strongly expressed.

5.4 Outlook

This study contributed new genetic facts about the restoration of fertility of C-CMS in maize. Furthermore, the results open interesting routes for further research, which promises insights into novel aspects about the restoration of C-CMS and of CMS in general.

Rf4 was fine-mapped to an interval in bin 8.00, which presumably does not contain PPR genes. Cloning will reveal if *Rf4* encodes a new type of restorer protein and if it functions according to a different mechanism than the previously identified restorer genes. A study about the interaction of *Rf4* with other restorer genes

is necessary, because the prospects of selecting restorer and non-restorer lines by means of markers are limited when based solely on the *Rf4* allele. Major QTLs for partial restoration of male fertility were found in bins 2.09, 3.06 and 7.03, but their importance has to be proven in different germplasm and in a broader range of environments. The major QTLs collocated with restorer genes of other CMS systems in maize, a phenomenon, which seems to be typical for restorer genes. Therefore, studying the genetic structure of the clusters in maize could improve the understanding about the evolution and functions of restorer genes. The long arm of chromosome 2 is particularly interesting, because it harbors restorer genes for all three major CMS types in maize.

References

Akagi H, Nakamura A, Yokozeki-Misono Y, Inakagi A, Takahashi H, Mori K, Fujimura T (2004) Positional cloning of the rice *Rf-1* gene, a restorer of BT-type cytoplasmic male sterility that encodes a mitochondria-targeting PPR protein. Theor Appl Genet 108:1449-1457

Allen J, Fauron C, Minx P, Roark L, Oddiraju S, Lin G, Meyer L, Sun H, Kim K, Wang C, Du F, Xu D, Gibson M, Cifrese J, Clifton J, Newton K (2007) Comparisons among two fertile and three male-sterile mitochondrial genomes of maize. Genetics 177:1173-1192

Balk J, Leaver CJ (2001) The PET1-CMS mitochondrial mutation in sunflower is associated with premature programmed cell death and cytochrome *c* release. Plant Cell 13:1803-1818

Beckett JB (1971) Classification of male-sterile cytoplasms in maize (Zea mays L.). Crop Sci 11:724-727

Bentolila S, Alfonso AA, Hanson MR (2002) A pentatricopeptide repeat-containing gene restores fertility to cytoplasmic male-sterile plants. Proc Natl Acad Sci USA 99:10887-10892

Bergman P, Edqvist J, Farbos I, Glimelius K (2000) Male-sterile tobacco displays abnormal mitochondrial *atp1* transcript accumulation and reduced floral ATP/ADP ratio. Plant Mol Biol 42:531-544

Brown GG, Formanova N, Jin H, Wargachuk R, Dendy C, Patil P, Laforest M, Zhang JF, Cheung WY, Landry BS (2003) The radish *Rfo* restorer gene of Ogura cytoplasmic male sterility encodes a protein with multiple pentatricopeptide repeats. Plant J 35: 262-272

Burge CB, Karlin S (1997) Prediction of complete gene structures in human genomic DNA. J Mol Biol 268:78-94 http://genes.mit.edu/GENSCAN.html. Accessed 10 November 2010

Butler DG, Cullis BR, Gilmour AR, Gogel BJ (2007) Analysis of mixed models for S language environments. ASReml-R reference manual.

Chase CD (2007) Cytoplasmic male sterility: A window to the world of plant mitochondrial-nuclear interactions. Trends Genet 23:81-90

Chase CD, Ribarits A, Heberle-Bors E (2010) Male sterility. In: Pua EC, Davey MR (eds.) Plant developmental biology - biotechnological perspectives. Springer-Verlag Berlin Heidelberg, pp 437-457

Churchill GA, Doerge RW (1996) Empirical threshold values for quantitative trait Mapping. Genetics 138:963-971

CIMMYT (2005) Laboratory protocols: CIMMYT applied molecular genetics laboratory. Mexico, D.F.: CIMMYT, 3rd edition

Claros MG, Vincens P (1996) Computational method to predict mitochondrially imported proteins and their targeting sequences. Eur J Biochem 241:779-786. http://ihg2.helmholtz-muenchen.de/ihg/mitoprot.html. Accessed 10 November 2010

Cui XQ, Wise RP, Schnable PS (1996) The *rf2* nuclear restorer gene of male-sterile T-cytoplasm maize. Science 272:1334-1336

Desloire S, Gherbi H, Laloui W, Marhadour S, Clouet V, Cattolico L, Falentin C, Giancola S, Renard M, Budar F, Small I, Caboche M, Delourme R, Bendahmane A (2003) Identification of the fertility restoration locus, *Rfo*, in radish, as a member of the pentatricopeptide-repeat protein family. EMBO Rep 4: 588-594

Dill CL, Wise RP, Schnable PS (1997) *Rf8* and *Rf** mediate unique T-*urf13*-transcript accumulation, revealing a conserved motiv associated with RNA processing and restoration of pollen fertility in T-cytoplasm maize. Genetics 147:1367-1379

Ducos E, Touzet P, Boutry M (2001) The male-sterile G cytoplasm of wild beet displays modified mitochondrial respiratory complexes. Plant J 26:171-180

Duvick DN (1956) Allelism and comparative genetics of fertility restoration of cytoplasmically pollen sterile maize. Genetics 41:544-565

Duvick DN (1965) Cytoplasmic pollen sterility in corn. Adv Genet 13:2-52

Emanuelsson O, Brunak S, von Heijne G, Nielsen H (2007) Locating proteins in the cell using TargetP, SignalP and related tools. Nat Protoc 2:953-971. http://www.cbs.dtu.dk/services/TargetP. Accessed 10 November 2010

Fan J, Oliphant A, Shen R, Kermani B, Garcia F, Gunderson K, Hansen M, Steemers F, Butler S, Deloukas P, Galver L, Hunt S, McBride C, Bibikova M, Rubano T, Chen J, Wickham E, Doucet D, Chang W, Campbell D, Zhang B, Krugylak S, Bentley D, Haas J, Rigault P, Zhou L, Stuelpnagel J, Chee M (2003) Highly parallel SNP genotyping. Cold Spring Harb Symp Quant Bio 68:6978

Fauron CMR, Havlik M, Brettell RIS (1990) The mitochondrial genome organization of a maize fertile CMS-T revertant line is generated through recombination between 2 sets of repeats. Genetics 124:423-428

Feil B, Stamp P (2002) The pollen-mediated flow of transgenes in maize can already be controlled by cytoplasmic male sterility. Ag Biotech Net 4:1-4

Finn RD, Mistry J, Tate J, Coggill P, Heger A, Pollington JE, Gavin OL, Gunasekaran P, Ceric G, Forslund K, Holm L, Sonnhammer ELL, Eddy SR, Bateman A (2010) The Pfam protein families database. Nucleic Acid Res 38:D211-D222. http://pfam.sanger.ac.uk. Accessed 10 November 2010

Fujii S, Toriyama K (2008) Genome barriers between nuclei and mitochondria exemplified by cytoplasmic male sterility. Plant Cell Physiol 49:1484-1494

Gabay-Laughnan S, Chase CD, Ortega VM, Zhao LM (2004) Molecular-genetic characterization of CMS-S *restorer-of-fertility* alleles identified in Mexican maize and teosinte. Genetics 166:959-970

Fujii S, Toriyama K (2009) Suppressed expression of retrograde-regulated male sterility restores pollen fertility in cytoplasmic male-sterile rice plants. Proc Natl Acad Sci USA 106:9513-9518

Fujimura-Kamada K, Nouvet FJ, Michaelis S (1997) A novel membrane-associated metalloprotease, Ste24p, is required for the first step of NH2-terminal processing of the yeast a-factor precursor. J Cell Biol 136:271-285

Gabay-Laughnan S, Chase CD, Ortega VM, Zhao LM (2004) Molecular-genetic characterization of CMS-S *restorer-of-fertility* alleles identified in Mexican maize and teosinte. Genetics 166:959-970

Gabay-Laughnan S, Kuzmin EV, Monroe J, Roark L, Newton KJ (2009) Characterization of a novel thermosensitive restorer of fertility for cytoplasmic male sterility in maize. Genetics 182:91-103

Gaut BS (2001) Patterns of chromosomal duplication in maize and their implications for comparative maps of the grasses. Genome Res 11:55-66

Geddy R, Brown GG (2007) Genes encoding pentatricopeptide repeat (PPR) proteins are not conserved in location in plant genomes and may be subject to diversifying selection. BMC Genomics 8:13

Gontarovskii V (1974) Role of supplementary genes in control of Texas type of CMS in maize. Genetika 10:15-24

Gonzalez-Melendi P, Uyttewaal M, Morcillo CN, Mora JRH, Jajardo S, Budar F, Lucas MM (2008) a light and electron microscopy analysis of the events leading to male sterility in Ogu-INRA CMS of rapeseed (*Brassica napus*). J Exp Bot 59:827-838

Hanson MR, Bentolila S (2004) Interactions of mitochondrial and nuclear genes that affect male gametophyte development. Plant Cell 16:S154-S169

Has V (2002) Evaluation of some maize inbred lines on fertility restoration patterns of male-sterile cytoplasms. Romanian Agricultural Research 17:1-7

Hospital F, Moreau L, Lacoudre F, Charcosset A, Gallais A (1997) More on the efficiency of marker-assisted selection. Theor Appl Genet 95:1181-1189

Howad W, Kempken F (1997) Cell type-specific loss of *atp6* RNA editing in cytoplasmic male sterile *Sorghum bicolor*. Proc Natl Acad Sci USA 94:11090-11095

Hu Y, Tang J, Yang H, Xie H, Lu X, Niu J, Chen W (2006) Identification and mapping of *Rf-I* an inhibitor of the *Rf5* restorer gene for Cms-C in maize (*Zea mays* L.). Theor Appl Genet 113:357-360

IUSS Working Group WRB (2007) World Reference Base for Soil Resources 2006, first update 2007. World Soil Resources Reports No. 103. FAO, Rome.

Janska H, Mackenzie SA (1993) Unusual mitochondrial genome organization in cytoplasmic male-sterile common bean and the nature of cytoplasmic reversion to fertility. Genetics 135:869-879

Jones D, Manglesdorf P (1957) The production of hybrid seed corn without detasseling. Bull Connecticut Agric Expt Sta, 550

Jones DF, Stinson HT, Khoo U (1957) Pollen restoring genes. Bull Connecticut Agric Expt Sta 610:1-43

Josephson L, Morgan T, Arnold J (1978) Genetics and inheritance of fertility restoration of male-sterile cytoplasms in corn. In: Proc. 33rd Ann Corn Sorghum Res Conf 33:4859

Kaul M (1988) Male sterility in higher plants. Springer-Verlag Berlin Heidelberg

Kendall M (1975) Rank correlation methods. Charles W. Griffin, London

Kheyr-Pour A, Gracen VE, Everett HL (1981) Genetics of fertility restoration in the C-group of cytoplasmic male-sterility in maize. Genetics 98:379-388

Klein RR, Klein PE, Mullet JE, Minx P, Rooney WL, Schertz KF (2005) Fertility restorer locus *Rf1* of sorghum (*Sorghum bicolor* L.) encodes a pentatricopeptide repeat protein not present in the colinear region of rice chromosome 12. Theor Appl Genet 111:994-1012

Kofer W, Glimelius K, Bonnett H (1991) Modifications of mitochondrial DNA cause changes in floral development in homeotic-like mutants of tobacco. Plant Cell 3:759-769

Koizuka N, Imai R, Fujimoto H, Hayakawa T, Kimura Y, Kohno-Murase J, Sakai T, Kawasaki S, Imamura J (2003) Genetic characterization of a pentatricopeptide repeat proetin gene, orf687, that restores fertility in the cytoplasmic male-sterile Kosena radish. Plant J 34:407-415

Koppen M, Langer T (2007) Protein degradation within mitochondria: versatile activities of AAA proteases and other peptidases. Crit Rev Biochem Mol Biol 42: 221-242

Kosambi DD (1944) The estimation of map distances from recombination values. Ann Eugenics 12: 172-175

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newburg L (1987). MAPMAKER: An interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174-181

Laser KD, Lersten NR (1972) Anatomy and cytology of microsporogenesis in cytoplasmic male sterile angiosperms. Bot Rev 38: 425-454

Laughnan JR, Gabay-Laughnan S (1983) Cytoplasmic male sterility in maize. Annu Rev Genet 17:27-48

Lee SJ, Gracen VE, Earle ED (1979) Cytology of pollen abortion in C-cytoplasmic male-sterile corn anthers. Am J Bot 66:656-667

Li HH, Ye GY, Wang JK (2007) A modified algorithm for the improvement of composite interval mapping. Genetics 175:361-374

Li XQ, Jean M, Landry BS, Brown GG (1998) Restorer genes for different forms of *Brassica* cytoplasmic male sterility map to a single nuclear locus that modifies transcripts of several mitochondrial genes. Proc Natl Acad Sci USA 95:10032-10037

Linke B, Nothnagel T, Börner T (2003) Flower development in carrot CMS plants: mitochondria affect the expression of MADS box genes homologous to GLOBOSA and DEFICIENS. Plant J 34:27-37

Liu Z, Yue Y, Xiang J, Wang J, Wu J, Li X, Yang Y (2010) Isolation of an ascorbate peroxidase in *Brassica napus* and analysis of its specific interaction with ATP6. Afr J Biotechnol 9:2050-2055

Liu ZY, Peter SO, Long MH, Weingartner U, Stamp P, Kaeser O (2002) A PCR assay for rapid discrimination of sterile cytoplasm types in maize. Crop Sci 42: 566-569

Mann HB, Whitney DR (1947) On a test whether one of 2 random variables is stochastically larger than the other. Ann Math Stat 18:50-60

Matsuhira H, Harada M (2005) Structural polymorphism of mitochondrial metalloprotease-like gene cluster in the *restorer-of-fertility*, *Rf1*, locus of sugar beet. Genes Genet Syst 80:477

Meyer L, Newton K (2009) Pollen abortion in CMS-C is associated with reduction of F_0 ATP synthase proteins. In: Maize Genetics Conference Abstracts 51:P55

Michelmore RW, Meyers BC (1998) Clusters of resistance genes in plants evolve by divergent selection and a birth-and-death process. Genome Res 8:1113-1130

Munsch M, Camp KH, Stamp P, Weider C (2008) Modern maize hybrids can improve grain yields as plus-hybrids by the combined effects of cytoplasmic male sterility and allopollination. Maydica 53:261-268

Munsch MA, Stamp P, Christov NK, Foueillassar XM, Hüsken A, Camp KH, Weider C (2009) Grain yield increase and pollen containment by Plus-Hybrids could improve acceptance of transgenic maize. Crop Sci 50:909-919

Murai K, Hirabayashi C, Hama E, Takumi S, Ogihara Y, Ishikawa M, Hirano H (2002) Pistillody, homeotic transformation of stamens into pistil-like structures, in alloplasmic wheat is associated with alterations to expression pattern of floral organ identity genes. Plant J 29:169-181

Osbourn A (2010) Secondary metabolic gene clusters: evolutionary toolkits for chemical innovation. Trends Genet 26:449-457

Pring DR, Conde MF, Levings CS (1980) DNA heterogeneity within the C group of maize male-sterile cytoplasms. Crop Sci 20: 159-162

Qin T, Xu M, Dun D (1990) Cytoplasmic male sterility: identification of the number of the restorer genes. MNL 64:124

R Development Core Team (2009) A language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria. http://www.R-project.org. Accessed 10 November 2010

Rogers JS, Edwardson JR (1952) The utilization of cytoplasmic male-sterile inbreds in the production of corn hybrids. Agron J 44: 8-13

Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol 132:365-386

Saha D, Prasad AM, Srinivasan R (2007) Pentatricopeptide repeat proteins and their emerging roles in plants. Plant Physiol Bioch 45:521-534

Salamov AA, Solovyev VV (2000) *Ab initio* gene finding in *Drosophila* genomic DNA. Genome Res 10:516-522 http://www.softberry.com. Accessed 10 November 2010

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: A laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, USA

Schaeffer (Polacco) M, Sanchez-Villeda H, Coe E (2008) IBM neighbors 2008 map. Available at http://www.maizegdb.org. Maize Genetics and Genomics Database. Accessed 10 November 2010

Schnable PS, Wise R (1998) The molecular basis of cytoplasmic male sterility and fertility restoration. Trends Plant Sci 3:175-180

Schnable PS, Ware D, Fulton RS, Stein JC, Wei F et al. (2009) B73 maize genome: complexity, diversity and dynamics. Science 326: 1112 - 1115. http://www.maizesequence.org. Accessed 10 November 2010

Sekito T, Thornton J, Butow RA (2000) Mitochondria-to-nuclear signaling is regulated by the subcellular localization of the transcription factors Rtg1p and Rtg3p. Mol Biol Cell 11:2103Ű2115

Shapiro SS, Wilk MB (1965) An analysis of variance test for normality (complete samples). Biometrika 52:591-611

Sisco P, Gracen V, Manchester C, Everett H (1982) Revised restoration ratings for Cornell's cytoplasm bank. MNL 56:80-82

Sisco PH (1991) Duplications complicate genetic mapping of *Rf4*, a restorer for CMS-C cytoplasmic male sterility in corn. Crop Sci 31:1263-1266

Skibbe DS, Wang XJ, Borsuk LA, Ashlock DA, Nettleton D, Schnable PS (2008) Floret-specific differences in gene expression and support for the hypothesis

that tapetal degeneration of *Zea mays* L. occurs via programmed cell death. J Genet Genom 35:603-616

Small I, Peeters N, Legeai F, Lurin C (2004) Predotar: A tool for rapidly screening proteomes for N-terminal targeting sequences. Proteomics 4:1581-1590. http://urgi.versailles.inra.fr/predotar/predotar.html. Accessed 10 November 2010

Small ID, Earle ED, Escotecarlson LJ, Gabay-Laughnan S, Laughnan JR, Leaver CJ (1988) A comparison of cytoplasmic revertants to fertility from different CMS-S maize sources. Theor Appl Genet 76:609-618

Smart C, Moneger F, Leaver C (1994) Cell-specific regulation of gene expression in mitochondria during anther development in sunflower. Plant Cell 6:811-825

Smit A, Hubley R, Green P (1996-2010) RepeatMasker Open-3.0. http://www.repeatmasker.org. Accessed 10 November 2010

Smith RL, Chowdhury MKU (1989) Mitochondrial DNA polymorphism in male-sterile and fertile cytoplasms of pearl millet. Crop Sci 29:809-814

Soller M, Beckmann JS (1990) Marker-based mapping of quantitative trait loci using replicated progenies. Theor Appl Genet 80:205-208

Sotchenko V, Gorbacheva A, Kosogorova N (2007) C-type cytoplasmic male sterility in corn. Russian Agricultural Sciences 33:83-86

Temnykh S, DeClerck G, Lukashova A, Lipovich L, Cartinhour S, McCouch S (2001) Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): frequency, length variation, transposon associations, and genetic marker potential. Genome Res 11:1441-1452. http://www.gramene.org/db/markers/ssrtool. Accessed 10 November 2010

Touzet P, Budar F (2004) Unveiling the molecular arms race between two conflicting genomes in cytoplasmic male sterility? Trends Plant Sci 9:568-570

Tracy WF (1982) The inheritance of late-breaking partial male fertility in the inbred Minn. A632 in PR cytoplasm. Dissertation, Cornell University

Tracy WF, Everett HL, Gracen VE (1991) Inheritance, environmental effects, and partial male-fertility in C-type CMS in a maize inbred. J Hered 82:343-346

Ullstrup AJ (1972) The impacts of the Southern Corn Leaf Blight epidemics of 1970-1971. Annu Rev Phytopathol 10:37-50

Utz H, Melchinger A (2000) PLABQTL: A computer program for statistical analysis of plant breeding experiments. Release version 1.1

Van Ooijen J (2006) JoinMap4, Software for the calculation of genetic linkage maps in experimental populations. Kyazma, B.V. Wageningen, Netherlands

Vidakovic M (1988) Genetics of fertility restoration in cytoplasmic male-sterility of the c-type (cmsC) in maize (*Zea mays* L.) Maydica 33:51-64

Vidakovic M, Vancetovic J, Vidakovic M (1997) The existence of a duplicated or parallel genetic system for fertility restoration in cmsC of maize (*Zea mays* L.). Maydica 42:313-316

Wang Z, Zou Y, Li X, Zhang Q, Chen L, Wu H, Su D, Chen Y, Guo J, Luo D, Long Y, Zhong Y, Liu Y (2006) Cytoplasmic male sterility of rice with Boro II cytoplasm is caused by a cytotoxic peptide and is restored by two related PPR motif genes via distinct modes of mRNA silencing. Plant Cell 18:676-687

Wang ZW, Zhang YJ, Xiang CP, Mei SY, Zhou Y, Chen GP, Wang T (2008) A new fertility restorer locus linked closely to the *Rfo* locus for cytoplasmic male sterility in radish. Theor Appl Genet 117:313-320

Warmke H, Lee S (1977) Mitochondrial degeneration in Texas cytoplasmic male-sterile corn anthers. J Hered 68:213-222

Wei JK, Liu KM, Chen JP, Luo PC, Lee-Stadelman OY (1988) Pathological and physiological identification of race C *Bipolaris maydis* in China. Phytopathology 78:550-554

Weider C, Stamp P, Christov N, Hüsken A, Foueillassar X, Camp KH, Munsch M (2009) Stability of cytoplasmic male sterility in maize under different environmental conditions. Crop Sci 49:77-84

Weingartner U, Kaeser O, Long M, Stamp P (2002) Combining cytoplasmic male sterility and xenia increases grain yield of maize hybrids. Crop Sci 42:1848-1856

Zeng ZB (1994) Precision mapping of quantitative trait loci. Genetics 136:1457-1468

Zhang JF, Stewart JM (2001) Inheritance and genetic relationships of the D8 and D2-2 restorer genes for cotton cytoplasmic male sterility. Crop Sci 41:289-294

Zhu S, Rossnagel BG, Kaeppler HF (2004) Genetic analysis of quantitative trait loci for groat protein and oil content in oat. Crop Sci 44:254-260

Zubko MK (2004) Mitochondrial tuning fork in nuclear homeotic functions. Trends Plant Sci 9:61-64

Appendix

Table 5.4.1 Primer sequences of comparative sequencing assays in the genomic region flanking *Rf4* polymorphic between B37C and K55.

Sequence name	BAC	Primer sequences 5'-3' (forward/reverse)	Amplicon length [bp]	T_m [°C]	Position[a] [bp]
cs-c046-1.1	AC198928.3	AAGGCGGGGTGACACATCAATAAT TAAGCGACACAGAAGCAGGACGAT	502	60	37,734
cs-c046-1.2	AC198928.3	CCGCCGCCTATGCCACCAA GGGAATGGGATATGGGGAGGAAGA	507	60	37,571
cs-c046-1.3	AC198928.3	GGCGGGGTGACACATCAATAATGG CTCCGGATGTCTCCTTCCTTCACC	563	60	37,731

[a] Physical position on respective BAC

Table 5.4.2 Primer sequences of comparative sequencing assays in the genomic region flanking *Rf4* monomorphic between B37C and K55.

Sequence name	BAC	Primer sequences 5'-3' (forward/reverse)	amplicon length [bp]	T_m [°C]	Position[a] [bp]
cs-b640-3.1	AC192587.3	ATGGTAAGAAGAGCTGCCGA CCTGACAGGAATTGACCGTA	452	60	80,436
cs-b640-3.2	AC192587.3	AACGACATGCAATCCAGACA TACCTCGGCTTCTTCAATGC	423	60	80,641
cs-b640-1.5	AC192587.3	GATCCTAAACTGGTTGTCGCA TTGTCACCAATTACCGGTTG	458	60	82,348
cs-b640-2.2	AC192587.3	CTCCTGCGATGACTGAACTG GCCAGACAGCAAATGTTTCA	498	60	83,706
cs-c0113-12.2	AC226331.3	AGAACTGAGATCATGGCGAAA TGTATGCATGGTTCTTGAGCA	605	60	84,068
cs-c0113-13.1	AC226331.3	GCTCAAGAACCATGCATACAA CGAACTATCCGTTTGGAGGA	721	60	84,653
cs-c0113-13.3	AC226331.3	GCTGAAATGTACATCAAGGGC CCATGGGACTCACAAAGACC	732	60	84,762
cs-c0113-7.1	AC226331.3	TGCAGTGGTAGCCACAAAAG CATCCCGTAGTCCGTTGAGT	480	60	83,107
cs-c0216-10.1	AC188738.4	CTGGGATTTTGCTACGGTGT GGCTACTATTGGGGGAGAGC	774	60	117,861
cs-c0216-11.2	AC188738.4	TACCTGTGCCAACCAAACA CTCATGCATTTTTCACACGG	668	60	86,314
cs-c0216-6.1	AC188738.4	GCCAGCGAAGATCGTTTTCC TCGCTCTGTCACCCTTCAGC	540	60	119,494
cs-c0216-7.1	AC188738.4	TCGCGAAACGTATCAGTGAG TTTCCACTAGCGGTTGTGTT	792	60	118,672
cs-c0216-8.2	AC188738.4	ACGGTCCACCAAACAATGAT TTCGCTGGCGGTTATTTTAG	722	60	118,781
cs-c0216-8.3	AC188738.4	ACCGCCTTTGATGGTTGTAG TTCGCTGGCGGTTATTTTAG	556	60	118,947
cs-c0216-9.2	AC188738.4	CTTCATGGCGTGTAGGGATT ACTTGACCCTTTTGACACCG	782	60	117,113

[a] Physical position on respective BAC

Table 5.4.3 Properties of newly developed markers linked to *Rf4*: Annnealing temperatures (T_a) of the primers, amplicon lengths and sequence polymorphisms.

Marker name	T_m [°C]	Parent	Amplicon length [bp]	Polymorphism (5'-3')
b0329-13.1	60	B37C	203	$(AG)_3GG(AG)_4$
		K55	251	$(AG)_{24}CGCT(AG)_5$
b640-3.4	65	B37C	-	
		K55	321	
c046-1.7	60	B37C	75	GAG****GAC
		K55	79	GAG**CGAG**GAC
c0113-2.2	60	B37C	275	$(AG)_7$**AAAAGGCAAGCT--TGAGACGGGGAA**$(AG)_8$
		K55	221	$(AG)_{13}$
c0216-3.15	60	B37C	58	$(TAA)_4$
		K55	115	$(TAA)_{22}$
b0440-16.2	60	B37C	187	$(CGG)_4$
		K55	196	$(CGG)_7$

Acknowledgments

This study was conducted in collaboration with KWS Saat AG (Einbeck, Germany) supporting laboratory analyses and Delley Seeds and Plants Ltd. (Delley, Switzerland) supporting field trial work.

The field trial in Cadenazzo was conducted with the support of Agroscope Changins-Wädenswil in Cadenazzo.

Pyrosequencing reactions were carried out with the support of Dr. Aria Minder and the Genetic Diversity Centre, ETH Zürich.

Dr. Roland Kölliker (Agroscope Reckenholz-Tänikon) kindly provided access to software for genetic mapping.

Danksagung

Besonderer Dank richtet sich an meinen Doktorvater Professor Peter Stamp. Ich danke Dir für Deine Unterstützung - nicht nur fachlich, sondern auch moralisch - die Freiheit, die Du mir gelassen hast und Deine Ehrlichkeit.

Dr. Rainer Messmer danke ich für sein herausragendes Engagement als Betreuer. Auf Dich war immer Verlass, sei es bei Engpässen im Feld, Kopf oder Computer! Ein grosses Dankeschön gilt auch Deiner Freundin Diana, die sich mehrmals zu Wochenendeinsätzen auf dem Feld einspannen liess.

Dr. Carsten Knaak von der KWS Saat AG danke ich für die Übernahme der Prüferschaft. Ausserdem möchte ich Dir meinen Dank für Deinen kompetenten und pragmatischen Einsatz aussprechen, der das Projekt in der entscheidenden Planungsphase mitgeformt hat und ohne den das Projekt in diesem Umfang nicht realisierbar gewesen wäre.

Dr. Karl-Heinz Camp von der DSP AG möchte ich für die Ermöglichung der Feldversuche in Delley und die organisatorische Hilfe danken. Ausserdem schätzte ich die Diskussionen über alle möglichen Aspekte der Maiszüchtung und den Einblick in Euren Zuchtgarten sehr. Dem ganzen DSP Team danke ich für Ihre freundliche Aufnahme.

Dr. Mario Bertossa danke ich für die Bereitstellung der Versuchsfläche beim ACW Cadenazzo, die zuverlässige Betreuung des Feldes und seine Hilfsbereitschaft während der Versuchszeit. Dank gilt auch all den Mitarbeitern vom ACW, die bei der Handaussaat mitgeholfen haben.

Marianne Wettstein, unserem guten Geist im Labor, danke ich für ihren unermüdlichen und fröhlichen Einsatz bei allen Laborarbeiten und diversen Feldein-

sätzen. Dein schallendes "Guten Morgen, Susanne" vermisse ich schon, seit die Laborversuche abgeschlossen sind! Danken möchte ich ausserdem den Auszubildenden David, Sabrina und Luca für ihre Mithilfe bei der Laborarbeit.

Die umfangreichen Bonituren im Feld wären nicht möglich gewesen ohne Euch fleissigen Boniteure! Marylène, Adrian, Eveline, Angela, Salome und Flavio, Ihr habt wirklich ganze Arbeit geleistet! Weitere unermüdliche Hilfskräfte waren bei diversen Aufgaben beteiligt, und dafür möchte ich Euch herzlich danken: Yelena, Grace, Ana Lorena, Rebecca, Youpi, Ernest, Adolphe, Mounir, Seraina und Olivia.

Die technische Arbeit auf dem Feld wurde 2008 mit grossem Sachverstand und einer Energie, die ihresgleichen sucht, von Helen Rast organisiert. Ernst Merz danke ich für die Betreuung des Feldes 2007.

Ernst Weiss, Landwirt in Lindau, danke ich für die Bereitstellung einer Versuchsfläche. Du hast so manche spontane Änderung im Versuchsplan ohne zu Zögern mitgetragen und für einen gesunden und schönen Maisbestand gesorgt.

Brigitta Herzog, Urs Moser und André Imboden danke ich für die Betreuung der Gewächshauskammern während der Saatgutvermehrungen im Winter.

Prof. Hartwig H. Geiger von der Universität Hohenheim danke ich für seinen kompetenten Rat in der Planungsphase des Projektes.

In bester Erinnerung bleibt die ganze Stamp-Truppe, bleiben die Apéros, die Ausflüge zum Letten, die Dissfeiern mit Euch... David, Chi, Quanjai, Regina, Dominique, Magali, Andi, Ravi, Niclas, Moritz, Jörg, Christophe, Debbie, Juan, Sämi, Chaminda, Kathrin, André, Susanne S. , Susanne H., Brigitta, Markus, Inge und Madeleine, ich werde Euch vermissen!

Zum Schluss möchte ich meinem Freund Sebastian einen speziellen Gruss und Dank senden. Du hast alles von Anfang bis Ende mitgetragen. Ausserdem hast Du dafür gesorgt, dass nicht auch auf unserem Balkon noch Mais wuchs!

Curriculum vitae

Name	Susanne Kohls
Date of birth	February 28, 1982
Place of birth	Itzehoe, Germany
Nationality	German
2007 - 2010	P.h.D. student at the Institute of Plant, Animal and Agroecosystem Sciences, Group Agronomy and Plant Breeding, ETH Zurich, Switzerland
2007	M.Sc. degree in Agricultural Sciences, Major Crop Sciences
	M.Sc. thesis: Einfluss einer Saatgutinkrustierung mit *Glomus intraradices* auf die Ertragsleistung von Mais. Universität Hohenheim, Germany
2001 - 2007	Studies in Agricultural Sciences, Universität Hohenheim, Germany
2001	Abitur, Sophie-Scholl-Gymnasium Itzehoe, Germany