

1 Einleitung

1.1 Molekulare Marker

In der Pflanzenzüchtung wurden und werden morphologische Merkmale als Marker zur Selektion von Populationen mit züchterisch wertvollen Eigenschaften verwendet. Der Nachteil morphologischer Merkmale ist jedoch, daß diese Marker in vielen Pflanzenarten nur in einer begrenzten Zahl vorhanden sind. Außerdem sind viele dieser Marker erst im adulten Zustand erkennbar und damit für Bonitierungen geeignet, wodurch die Züchtung einer Sorte sehr zeitaufwendig ist. In den vierziger Jahren wurden unter großem experimentellen Aufwand morphologische Genkarten bei Pflanzen konstruiert. Beispiele sind die Genkarten von Tomate (MACARTHUR 1934) und Mais (EMERSON et al. 1935).

Um die Selektionszeit zu verkürzen, wurden 1980 von TANKSLEY und RICK erstmals molekulare Marker in eine Kopplungskarte integriert. Bei diesen ersten molekularen Markern handelte es sich um Isoenzyme. Der Vorteil in der Verwendung von Isoenzymen als Marker war zunächst der, daß diese schon sehr früh, d. h. im Keimlingsstadium erkennbar waren. Außerdem handelte es sich um codominante Marker, was den Vererbungsvorgang der gesuchten Gene leichter nachweisbar machte. Nachteilig war hier nur, daß wenig bekannte Isoenzyme vorhanden waren, wodurch diese nur geringe Anwendungsmöglichkeiten in der Züchtung boten (TANKSLEY 1983). Außerdem werden die meisten Merkmale, die agronomisch von Interesse sind, durch mehrere Gene reguliert, sie sind also polygen und zeigen komplexe Vererbungen auf, was mehr Anforderungen an eine markergestützte Selektion erfordert (RIBAUT und HOISINGTON 1998).

Auf DNA-Ebene boten die von BOTSTEIN et al. 1980 entwickelten sog. Restriktions-Fragment-Längen-Polymorphismen (RFLPs) viel mehr Einsatzmöglichkeiten innerhalb der Pflanzenzüchtung. Dabei wird DNA verdaut und auf einem Agarosegel aufgetrennt. Durch Hybridisierung mit beliebigen, markierten DNA-Sonden werden Unterschiede im

Bandenmuster sichtbar gemacht. Diese Polymorphismen sind durch Punktmutationen, Inversionen, Insertionen und/ oder Deletionen entstanden.

RFLP-Marker erbrachten eine Vielzahl von Vorteilen: Es handelt sich hier um in der Regel codominante Marker, die multiple Allelien nachweisen können. Sie sind unabhängig von der Entwicklungsphysiologie und von Umweltfaktoren sowie phänotypisch neutral. Außerdem steht mit den RFLP-Markern eine unter praktischen Gesichtspunkten unbegrenzte Anzahl von Markern zur Verfügung.

Diese molekularen Marker fanden bisher zahlreiche Anwendung in der Züchtung wie z. B. bei der Sortenidentifizierung, der Berechnung genetischer Distanzen und der markergestützten Einkreuzung erwünschter Eigenschaften in Zuchtlinien wie z. B. das Einkreuzen von Wildtyp-Merkmalen in Zuchtlinien (TANKSLEY und MCCOUCH 1997). RFLP-Marker werden außerdem zum Pyramidisieren von Genen (HUANG et al. 1997), zur Überprüfung des Erfolges bei Einkreuzungen oder bei der somatischen Zellfusion eingesetzt. Weiterhin werden diese DNA-Marker bei der markergestützten Selektion qualitativer und quantitativer Merkmale, der Klonierung von Genen und der vergleichenden Genomanalyse z. B. bei Tomate und Kartoffel (BONIERBALE et al. 1988) oder bei Hirse und Mais (WHITKUS et al. 1992) eingesetzt.

Für viele dieser Anwendungen werden Kopplungskarten benötigt: Ein Beispiel sei hier die markergestützte Rückkreuzung (YOUNG und TRANKSLEY 1989). Hierfür wurden wiederum zahlreiche Kopplungskarten für Pflanzen und Tiere angelegt. Die erste Raps-RFLP-Karte wurde 1991 von LANDRY et al. aus der F₂ einer Kreuzung von zwei Sommerrapsorten erstellt. Für den Winterraps wurde von UZUNOVA et al. (1995) eine Kopplungskarte aus einer DH-Population von den Sorten „Mansholt's Hamburger Raps“ und „Samourai“ zusammengestellt. Diese Karte wird in der vorliegenden Arbeit ebenfalls verwendet. Eine Übersicht über erste RFLP-Karten pflanzlicher Genome ist von UZUNOVA (1994) zusammengefaßt worden. Mit Hilfe von RFLP-Karten können unter anderem Genom-Vergleiche durchgeführt werden (BONIERBALE et al. 1988, WHITKUS et al. 1992, SONG et al. 1995).

Ein großer Nachteil der RFLP-Marker liegt jedoch darin, daß sie für Massenanalysen an Pflanzen nicht geeignet sind, weil die Nachweismethode sehr aufwendig ist. Sie

erfordern zudem größere Mengen an DNA (im µg-Bereich). Auf Polymerase-Ketten-Reaktionen (PCR) (SAIKI et al.1985, MULLIS und FALOONA 1987, SAIKI et al.1988) gestützte Nachweismethoden benötigen dagegen weitaus geringere Mengen an DNA. Mit Hilfe dieser Methoden können aus lediglich 20 bis 30 ng DNA gezielt polymorphe Loci vervielfältigt – also amplifiziert – werden. Dadurch wird es möglich, in sehr frühen Stadien Analysen durchzuführen, weil hier nur wenig Zellmaterial benötigt wird. Ein weiterer Vorteil liegt darin, daß schnelle DNA-Isolierungsmethoden gewählt werden können, gerade weil so wenig DNA für eine PCR benötigt wird. Wegen der hohen Effizienz der PCR-Technologie und Automatisierungsmöglichkeit der Analysen ist ein hoher Durchsatz an Proben möglich (KRESOVICH et al. 1995, POWELL et al. 1996, RIBAUTL et al. 1997, RIBAUTL und HOISINGTON 1998).

Auf PCR-Reaktion basierende Marker können in verschiedenen Varianten genutzt werden, die hier nur noch kurz erläutert werden sollen: Zu den PCR-Markern gehören in erster Linie die „random amplified polymorphic DNA“ (RAPD), die von WILLIAMS et al. 1990 entwickelt wurde. Dabei wird eine Auswahl willkürlicher, kurzer Primer (8 - 10 bp) verwendet, mit deren Hilfe zufällige Loci im Genom vervielfältigt und anschließend auf einem Agarosegel aufgetrennt werden. An diesen Loci entstandene Veränderungen im Genom werden durch das Auftreten unterschiedlicher DNA-Amplifikationsprodukte sichtbar. Diese Methode ist neben den oben aufgezählten Vorteilen zudem sehr einfach, schnell und kostengünstig und kann daher mit geringem Aufwand in jedem Labor ausgeführt werden. Es wurden unter anderem genetische Verwandtschaften mittels RAPD-Marker untersucht, so im Mais (WELSH et al. 1991) und der Ackerbohne (LINK et al. 1995), aber auch beim Raps (FÖRSTER und KNAAK 1995).

Mit Hilfe der RAPD-Marker können jedoch nur dominante Erbvorgänge nachgewiesen werden. Außerdem erwiesen sich die RAPD-Analysen oftmals als nicht reproduzierbar und weniger zuverlässig. Aus diesem Grund werden heute immer mehr RAPD-Marker in sog. „sequence characterized amplified regions“ (SCAR-Marker, PARAN und MICHELMORE 1993) umgeformt. Hierfür wird der Locus, der durch den RAPD-Marker amplifiziert wurde, sequenziert und für die Enden dieses DNA-Fragments längere, spezifische Primer synthetisiert. Diese SCAR-Marker erwiesen sich als reproduzierbarer und zuverlässiger als RAPD-Marker, aber genauso zeit- und

kostengünstig wie die RAPD-Marker, aus denen sie entstanden sind. Sie wurden z. B. für Massenanalysen in der markergestützten Selektion (marker-assisted selection = MAS) der Resistenz gegen Mehltau bei Weizen erfolgreich eingesetzt (LIU et al. 1999).

Ein weiterer PCR-Markertyp, der vielfach eingesetzt wird, ist der 1993 von ZABEAU und VOS patentierte „amplified fragment length polymorphism“ (AFLP), dessen Methode 1995 von VOS et al. ausführlich beschrieben wurde. Es handelt sich hier um eine Kombination aus Restriktions- und PCR-Methoden: Nach einer Restriktion der genomischen DNA werden an die Enden der Fragmente Adapter ligiert. Diese Fragmente werden als Template-DNA für eine PCR mit bestimmten Primern verwendet. Diese Primer besitzen die Sequenz der Adapter und der Restriktionschnittstelle. Durch zusätzliche selektive Basen wird die Spezifität des Primer-„Annealens“ bestimmt. Hier können mit wenigen Primern eine Vielzahl von Loci untersucht werden. Dieses bringt jedoch den Nachteil, daß die Auswertung der Analysen durch eine Vielzahl von Banden in Polyacrylamid-Gelen kompliziert ist. Außerdem ist diese Methode wieder zeitaufwendiger durch vorgeschaltete Restriktions-, sowie Ligations- und PCR-Reaktionen.

Weitere PCR-Marker setzen eine Kenntnis der zu untersuchenden DNA-Sequenz voraus, weil hierbei Primer an genomspezifische Loci binden. Hierzu gehören Bereiche, die nur einmal im Genom vorkommen und bereits sequenziert wurden. Diese „sequence tagged sites“ (STS) (MULLIS und FALOONA 1987, OLSON et al. 1989) sind codominante Marker, die ca. 100 bis 1.000 bp groß sind. Sie dienen als Erkennungsstellen für bestimmte Genomabschnitte und werden daher bei der physikalischen Kartierung eingesetzt (OLSON et al. 1989). Sie haben besonders in dem „human genome project“ (HGP) zur Aufklärung des menschlichen Genoms immer mehr an Bedeutung gewonnen (COLLINS und GALAS 1993).

Darüber hinaus können alle Sequenzen direkt auf Punktmutationen und dadurch auftretende Polymorphismen, sog. „single nucleotide polymorphisms“ (SNPs) (Beispiele: KONFORTOV et al. 1999, MEKSEM et al. 2001) untersucht werden. Für schnelle Massenanalysen sind allerdings automatische Sequenzierungen und

aufwendige Alignment-Software erforderlich, um diese Mutationen erkennen zu können.

Bei den sog. „direct amplifications of minisatellite DNA by PCR“ (DAMD-PCR-Markern) (SOMERS et al. 1996) dienen Minisatelliten als Grundlage für spezifische Primer. Dadurch werden aufwendige Hybridisierungsarbeiten in einfache PCR-Reaktionen umgewandelt.

Zuletzt soll hier die große Gruppe der Mikrosatelliten-Marker genannt werden, deren Identifizierung und Entwicklung aufgrund seiner vielseitigen Einsetzbarkeit immer weiter steigen.

1.2 Was sind „Mikrosatelliten“?

Genomische DNA erzeugt nach einer Zentrifugation im Cäsiumchlorid-Gradienten und Ethidiumbromid-Färbung (MESELSON et al. 1957) charakteristische, artspezifische Banden. Neben einer Hauptbande an genomischer DNA werden eine oder mehrere Nebenbanden sichtbar. Diese wurden 1962 von KIT als Begleit- oder Satelliten-DNA bezeichnet. Diese DNA war im wesentlichen auf die Centromere (HARRISON und HESLOP-HARRISON 1995) und Telomere (WESTMAN und KRESOVICH 1997) der Chromosomen begrenzt und bestand aus langen, sich tandemartig wiederholenden Sequenzbereichen, also repetitive DNA mit einem hohen Anteil an Guanin und Cytosin (G/C-Gehalt). Satelliten-DNA wurde aber auch in Chloroplasten (EDELMAN et al. 1965, RAY und HANAWALD 1965) und Mitochondrien (RABINOWITZ et al. 1965, CORNEO et al. 1966) gefunden, die ebenfalls einen hohen G/C-Gehalt in ihrer DNA aufwiesen. Die spätere Definition von Satelliten-DNA beschränkte sich auf diese langen, sich wiederholenden Sequenzen, egal ob diese nun in einem Dichtegradienten Satelliten-Banden bildeten oder nicht.

Durch Zufall wurden von BELL et al. 1982 durch Sequenzierungsarbeiten an menschlichen Insulingenen kurze DNA-Sequenzen entdeckt, die sich ebenfalls tandemartig wiederholten, die aber nur Größen von 15 Basenpaaren (bp) aufwiesen. Diese Core-Einheiten, sog. „Repeats“, die bis zu 100 bp lang sein konnten und

ca. 10- bis 100fach wiederholt werden, wurden 1985 von JEFFREYS et al. in der Klasse der Minisatelliten zusammengefaßt. Sie sind z.T. im Genom verstreut, kommen jedoch im telomeren (ROYLE et al. 1987, NAKAMURA et al. 1988) und subtelomeren Bereich (SIMMLER et al. 1987) gehäuft vor. Sie weisen Polymorphismen in ihren Längen auf, sog. „varial numbers of tandem repeats (VNTRs), die von Individuum zu Individuum variieren können und sich nach den Mendelschen Regeln vererben (JEFFREYS et al. 1985). Hieraus wurde von JEFFREYS und Mitarbeitern 1985 das genetische Fingerprinting entwickelt, bei der genomische DNA restringiert, gelelektrophoretisch aufgetrennt, auf Membranen fixiert und mit Minisatelliten-Sonden hybridisiert (SOUTHERN 1975) wird. Mit den daraus resultierenden individuell unterschiedlichen Bandenmustern ließen sich Verwandtschaftsbeziehungen nachweisen (JEFFREYS et al. 1985).

Etwa zeitgleich fanden MIESFELD et al. (1981) in der menschlichen Intergen-Region des δ - und β -Globins eine aus den Basen Guanin (G) und Thymin (T) bestehende „Core“-Einheit, die sich 17fach wiederholte (GT)₁₇. Kurz darauf wurden weitere kleinere bp-Motive entdeckt, die sich tandemartig wiederholten (u. a. HAMADA und KAKUNAGA 1982, TAUTZ und RENZ 1984, GROSS und GARRARD 1986, PARDUE et al. 1987 und BRAATEN et al. 1988). WEBER und MAY sowie LITT und LUTY nannten 1989 diese DNA-Sequenzen Mikrosatelliten. Sie sind auch unter „short tandem repeats“ (STRs) (WANG et al. 1994, SPRECHER et al. 1996) oder „simple sequence repeats“ (SSRs) (JACOB et al. 1991) bekannt. In der vorliegenden Arbeit werden im folgenden die Begriffe Mikrosatelliten und SSRs verwendet.

Die Definition der Mikrosatelliten ist bis heute noch umstritten: TAUTZ sprach 1989 von einer „Core“-Einheit aus Mono-, Di- oder Trinukleotiden, die wiederholt werden, wobei auch größere Motive eine Einheit bilden können. SPRECHER et al. (1996) definierten diese „Core“-Einheiten mit Größen von 3 bis 7 bp. LAVI et al. (1994) berichteten von Einheiten bestehend aus 2 bis 5 bp, und KOSTIA et al. schrieben 1995 von 2 bis 4 bp-Motiven. AYRES SIA et al. (1997) grenzen SSRs mit 1 bis 13 bp-Motiven gegen die Minisatelliten ab. Allgemein wird jedoch in der Literatur von 1 bis 6 bp gesprochen (Beispiele: WANG et al. 1994, MÖRCHEN et al. 1996, VAN DE VEN und MCNICOL 1996). Diese Einheit wurde auch für diese Arbeit als Grundlage gewählt. Auch die Anzahl an Wiederholungen ist nicht fest beschrieben. VAN DE VEN

und MCNICOL (1996) sprechen von mehr als 10 Wiederholungen. WEBER legte 1990 die Faustregel von ebenfalls mindestens 10 Wiederholungen bei Dinucleotiden fest, bei dem ein Polymorphismus entstehen kann. Spätere Untersuchungen zeigten, daß auch kürzere Mikrosatelliten durchaus polymorph sein können (SENIOR und HEUN, 1993). WEBER und MAY (1989) legten 15 – 30 Wiederholungen fest. Viele Autoren lassen diese Frage ganz offen. In dieser Arbeit wurde die Mindestzahl an Wiederholungen bei Dinucleotiden, wie bei SENIOR und HEUN (1993) beschrieben, auf 5 heruntersetzt.

Mikrosatelliten lassen sich in 3 Typen einteilen. Zu den einfachen Mikrosatelliten zählen die perfekten SSRs, deren Motive ohne Unterbrechungen wiederholt werden, und imperfekte Mikrosatelliten, deren Wiederholung von einem oder mehreren bp unterbrochen werden. Diese SSRs werden auch „interrupted SSRs“ genannt (LORIDON et al. 1998, BUTCHER et al. 2000) Zum dritten Typ gehören die zusammengesetzten SSRs, bei denen 2 oder mehrere Motive direkt aufeinander folgen. Dabei können perfekte sowie imperfekte Mikrosatelliten in einem SSR verbunden sein. Eine Übersicht gibt Tab. 1.1.

Tab. 1.1: Übersicht über die Einteilung der SSRs in Typen.

| SSR-Typ | Definition | Beispiel |
|-----------------|---|---|
| perfekt | Motivwiederholungen ohne Unterbrechungen | (GA) ₈ |
| imperfekt | Motive unterbrochen durch ein oder mehrere bp | (CA) ₄ GA(CA) ₁₉ |
| zusammengesetzt | Mehrere Motive folgen aufeinander | (TA) ₁₄ (CA) ₃ T(CA) ₆ |

SSRs kommen bei Prokaryonten (MOXON et al. 1994) sowie bei Eukaryonten (TAUTZ und RENZ 1984, LAGERCRANTZ et al. 1993, WANG et al. 1994) vor und sind dabei zufällig im Genom verteilt (WEBER und MAY 1989 sowie LITT und LUTY 1989). Mikrosatelliten kommen jedoch auch in Organellen-DNA vor. So fanden WANG et al. 1994 in der Genbank in Chloroplasten SSRs (cpSSRs). CATO und RICHARDSON konnten 1996 mit cpSSRs nachweisen, daß bei Coniferen die Chloroplasten größtenteils paternal vererbt werden.