

3.2 Auswahl der Pflanzen-Stichproben

Die Probenentnahme erfolgte nach folgenden Kriterien: Es sollten viele Standorte ausgewählt und, wenn möglich, jeweils 5-10 Individuen pro Pflanzenart gesammelt werden. An einigen Standorten wurden für populationsgenetische Analysen jeweils 20 Individuen gesammelt. Neben der räumlichen Komponente (verschiedene Standorte) wurde zusätzlich die zeitliche Komponente untersucht, indem an ausgewählten Standorten im Zeitraum von drei Jahren Pflanzenproben entnommen wurden. Hier stellte sich die Frage, ob mit Hilfe von RAPD-Markern unterschiedliche Sammeljahre differenziert werden können.

Das Auswahlverfahren der Blattproben erfolgte, vorausgesetzt es waren ausreichend Pflanzen vor Ort, nach einer Zufallsmethode (GUARINO et al. 1996), wobei nach zwei Schritten rechts eine Blattprobe und nach weiteren zwei Schritten links eine Blattprobe entnommen wurde. Zusätzlich wurden von diesen Pflanzen, soweit vorhanden, Samen entnommen und für die Isozymanalyse (Bearbeitung I. Salifou) im ICRISAT, Niamey, bereitgestellt. Aufgrund des hohen Beweidungsdruckes und der Größe des Untersuchungsgebietes (2000 km Ost-West-Ausdehnung) war es nicht immer möglich, zum Zeitpunkt der Samenreife am Standort zu sein. Deshalb wurde in der Isozymanalyse neben Samen auch Blattmaterial derselben Pflanzen eingesetzt.

In Diskussion mit Prof. M. Saadou (Universität Niamey) wurden aus 12 im Sahel vorkommenden Futterpflanzen fünf ausgewählt, die folgende Kriterien erfüllten:

- sie stellen für die ansässige Bevölkerung wichtige Futterpflanzen dar
- sie sind weit verbreitet
- sie sind im Gelände leicht zu identifizieren

Im Zeitraum 1993-1996 wurden ca. 2000 Pflanzenproben der folgenden Futterpflanzen gesammelt, mit Silicagel getrocknet und nach Giessen gebracht:

Brachiaria spp.
Zornia glochidiata
Alysicarpus ovalifolius
Dactyloctenium aegyptium
Cenchrus biflorus

Neben molekulargenetischen und biochemischen Analysen wurden Erkenntnisse über die Art und Verteilung der Vegetation mit Schwerpunkt auf die krautigen Vertreter erarbeitet (Fr. Dr. Küppers, Uni Frankfurt). Zur Vegetationsklassifizierung dieser Gebiete, mit dem Ziel der Definition von Pflanzengesellschaften, wurde ein Pflanzenherbar angelegt. Die Auswertung von Luft- und Satellitenbilder von Pflanzengesellschaften bzw. die Untersuchung von Änderungen der Landoberfläche (Vegetation-Bodenkomplexe) sollte eine Abschätzung der Gefährdung dieser Gebiete ermöglichen. Fr. Dipl. Geogr. L. Cremers (FU Berlin) führte eine Befragung über die Weidepraktiken der einheimischen Bevölkerung im Hinblick auf eine *in situ*- Konservierung der genannten Zielarten durch.

3.3 Arbeitsmethoden

3.3.1 DNA - Extraktionsmethoden

Es wurden Vorversuche mit einheimischen Gräsern (*Lolium multiflorum*, *Echinochloa crus-galli*) und einer Leguminose (*Trifolium pratense*) durchgeführt. Hierbei sollte zunächst eine DNA- Extraktionsmethode etabliert werden, die sich auf die afrikanischen Pflanzenproben anwenden ließ. Dazu wurden die Testpflanzen mit Silicagel (ZHANG & TAO 1989) getrocknet. Nach 10 Tagen und dreimaligem Wechseln des Gels konnte das Blattmaterial mit verschiedenen Extraktionsmethoden bearbeitet werden. Dabei mußte die Methode schnell und problemlos durchzuführen sein, denn es sollten insgesamt 400 afrikanische Pflanzenproben extrahiert und für PCR-Versuche gereinigt werden. Zwei DNA-Extraktionsmethoden standen am Ende der Vorversuche zur Auswahl:

a) CTAB-Extraktionsmethode (DOYLE & DOYLE 1990), leicht modifiziert

Puffer:	0.1	M	Tris / HCl	pH 8.0
	1.4	M	NaCl	
	0.02	M	EDTA	pH 8.0
	4	%	CTAB	
	0.2	%	Mercaptoethanol	

b) SDS-Extraktionsmethode (EDWARDS et al. 1991), leicht modifiziert

Puffer:	0.2	M	Tris / HCl	pH 7.4
	0.25	M	NaCl	
	0.025	M	EDTA	pH 8.0
	0.5	%	SDS	

Versuchsdurchführung:

1. 500 bis 1000 mg Blattmaterial werden mit etwas Quarzsand und Stickstoff gemörsert, das feine Pulver in 12 ml Greiner-Plastikröhrchen überführt und 4 ml vorgewärmten CTAB-Extraktionspuffer (bzw. SDS-Extraktionspuffer) sowie 40 μ l Proteinase K zugegeben. Dieser Ansatz wird für 45 min bei 65 °C im Wasserbad inkubiert. Proteinase K ist ein hitzebeständiges Enzym, das über einen großen pH-Bereich (7.4 -11.5) hinweg arbeitet und für die Denaturierung von DNAsen zuständig ist, die während der Extraktion aus den Zellen freigesetzt werden und die DNA empfindlich schädigen können.
2. Es werden 4 ml Chloroform / Isoamylalkohol (24:1) zugegeben und vorsichtig die beiden Phasen gemischt. Es folgt eine 15 minütliche Zentrifugation (4500 g Heraeus Minifuge), wobei sich die beiden Phasen wieder voneinander trennen. Dabei bildet sich in der Übergangszone eine sogenannte feste Interphase, in der sich denaturierte Proteine, Polysaccharide und Zellreste befinden. Die obere wäßrige Phase besteht aus Extraktionspuffer und DNA, die untere organische Phase aus Chloroform, gelösten bzw. denaturierten Proteinen und Polysacchariden. Am Röhrchenboden hat sich der Quarzsand und weiteres Zellmaterial abgelagert.
3. Die obere wäßrige Phase wird in ein neues Greiner-Röhrchen überführt und Punkt 2 wiederholt. Danach sollte die Interphase verschwunden bzw. nur noch sehr schwach zu sehen sein. Die obere wäßrige Phase hat nun ihre typische grüne Färbung verloren und wird vorsichtig in ein neues Greiner-Röhrchen überführt. Hierbei ist besonders darauf zu achten, daß keine Bestandteile aus der Interphase und keine Flüssigkeit aus der unteren organischen Phase mit übertragen werden.
4. Es werden nun 2 Volumen (ca. 8 ml) Ethanol (96 %) zugegeben, vorsichtig durchmischt und die DNA (evtl. über Nacht bei 4 °C) ausgefällt.
5. Durch Zentrifugation für 30 min (Heraeus Minifuge) wird sämtliche DNA als Pellet am Boden des Röhrchen konzentriert. Der Überstand wird dekantiert und das Pellet mit 4 ml Ethanol (70 %) vorsichtig gewaschen. Dabei sollte es sich vom Boden lösen.
6. Nach erneuter Zentrifugation für 10 min und Dekantieren des Ethanols läßt man die DNA etwas lufttrocknen, damit Reste des Ethanols noch verdunsten können. Das DNA- Pellet wird nun mit 1-2 ml TE-Puffer aufgenommen und über Nacht bei Raumtemperatur gelöst. Durch Zugabe von 3 μ l RNase/ml DNA-Lösung wird sämtliche RNA innerhalb einer Stunde verdaut. Die DNA kann nun eingefroren oder im Kühlschrank aufbewahrt werden.

7. Untersuchungen in der PCR- bzw. RAPD-Technik zeigten, daß das erzeugte Bandenmuster stark von der DNA-Konzentration abhängig sein kann, insbesondere zu hohe Konzentrationen führen oft zu nicht reproduzierbaren Ergebnissen (TINGEY et al. 1992).

Aus diesem Grund wurde die DNA entsprechend verdünnt und die Konzentration mit Hilfe eines DNA- Markers (λ -Marker) bestimmt. Für eine RAPD- PCR- Reaktion sind 10-15 ng DNA notwendig (WILLIAMS 1990).

Die SDS-Methode unterscheidet sich in der Versuchsdurchführung nicht von der CTAB-Methode. Letztere zeigte bei *Zornia glochidiata* keine zufriedenstellenden Ergebnisse, die SDS-Methode hingegen arbeitete für beide Pflanzen problemlos. Für alle weiteren Extraktionsversuche wurde deshalb die SDS-Methode verwendet.

3.3.2 Die Polymerase Ketten Reaktion (PCR)

Die Entwicklung der Polymerase Ketten Reaktion (polymerase chain reaction) durch MULLIS (1986) ist eine der bedeutendsten Erfindungen in der Molekularbiologie der letzten Jahrzehnte und hat mittlerweile weltweite Anwendung gefunden. Der wesentliche Fortschritt dieser Technik besteht in der Einfachheit und Schnelligkeit, jeden DNA- Bereich des Genoms in theoretisch unbegrenzter Menge amplifizieren zu können. Dabei sind nur geringe Mengen DNA notwendig und die technischen Voraussetzungen gering. Die Erbinformation DNA (Desoxyribonukleinsäure) besteht aus zwei komplementären Strängen, die zu einer Doppelhelix aufgewunden sind. Die vier beteiligten Basen Adenin, Guanin, Cytosin und Thymin sind über Wasserstoffbrückenbindungen mit den komplementären Basen des Gegenstranges verbunden. Die Information der DNA ist durch die Basenabfolge der DNA kodiert und die Größe kann dabei erheblich variieren. So haben Bakterien (*Escherichia coli*) eine Genomgröße von 4×10^6 Basenpaaren (bp), das haploide Genom des Menschen 3×10^9 bp und das der Pflanzen (*Allium cepa*) $1,5 \times 10^{10}$ bp (KRAWCZAK & SCHMIDTKE 1994).

Die PCR-Reaktion findet heute in sog. Thermocyclern statt, welche die Reaktionsproben in kürzester Zeit aufheizen und abkühlen können. Zunächst speichert man ein Temperaturprofil ab, das aus einem Zyklus mit drei Reaktionsschritten besteht. Der Reaktionsmix wird zu Beginn auf 94 °C erhitzt. Bei dieser Temperatur werden die Wasserstoffbrückenbindungen gelöst und die doppelsträngige DNA zerfällt in ihre beiden Einzelstränge (Denaturierung), das anschließende schnelle Abkühlen auf 36-60°C verhindert, daß die Stränge wieder miteinander verschmelzen. Bei dieser Temperatur (Annealing) kommen Primer zum Einsatz, die

komplementär an die beiden Einzelstränge binden und sog. Hybridmoleküle bilden. Dazu wird vorher die richtige Annealing-Temperatur eingestellt, die wiederum von der Länge und Nukleotidabfolge des Primers abhängt. Ein Primer besteht aus einer kurzen einzelsträngigen DNA-Sequenz, die man entweder synthetisch herstellt oder experimentell durch Isolierung des gewünschten DNA-Abschnitts, z.B. aus einem Gen, erhält. Man verwendet zwei verschiedene Primer, die die Anfangs- bzw. Endsequenz des gewünschten Fragments kodieren. Nach der Annealing-Phase wird der Reaktionsmix auf 72°C erhitzt. Nun folgt der wichtigste Schritt der PCR, die Primerelongation. Der durch das Primer-Annealing entstandene, doppelsträngige Bereich dient nun als Startpunkt für die DNA-Polymerase, mit deren Hilfe der Einzelstrang zum Doppelstrang aufsynthetisiert wird. Die DNA-Polymerase ist ein Enzym, das aus dem hitzestabilen Bakterium *Thermus aquaticus* gewonnen wird, das bei 72°C sein Temperaturoptimum hat. Dieses Enzym baut Nukleotidtriphosphate (dNTPs) komplementär in 5'-3'-Richtung ein und vervollständigt somit den Doppelstrang. Als Produkt erhält man zwei neue doppelsträngige DNA-Moleküle, deren Länge durch die 5'-Enden der Primer festgelegt sind. Durch erneutes Erhitzen werden diese Doppelstränge in Einzelstränge zerlegt und die Primer können nach wiederholtem Abkühlen an die Enden dieser neuen Matrizen anbinden. Durch die Wiederholung dieses Zyklus (30-40 mal) aus Denaturierung, Annealing und DNA-Synthese kommt es jedesmal zur Verdopplung der DNA und damit in einer "Kettenreaktion" zu einer exponentiellen, selektiven etwa 10^6 - 10^7 fachen Anreicherung der durch die Oligonukleotide flankierten DNA-Sequenz (WINK und WEHRLE 1994). Damit erlaubt die PCR eine schnelle, hochsensitive Detektion und Amplifizierung von speziellen DNA-Abschnitten. Als Ausgangsmaterial ist theoretisch eine einzige Kopie der zu untersuchenden DNA ausreichend.

3.3.3 Random Amplified Polymorphic DNA (RAPD)

Eine abgewandelte Form der PCR ist die sogenannte RAPD-Methode, die 1990 gleichzeitig von WILLIAMS et al. und WELSH & MCCLELLAND entwickelt wurde. Diese genetischen Marker sind DNA-Fragmente, die mit nur einem Primer amplifiziert werden. Der Primer ist ein Oligonukleotid von 9-10 bp Länge und besteht aus einer zufälligen DNA-Sequenz, die nur durch den GC-Gehalt von über 50 % bestimmt ist. Da der Primer sehr kurz ist, gibt es im untersuchten Genom mehrere Bindungsstellen. Statistisch findet man daher etwa alle 10^6 Basen einen komplementären Bindungsort. Liegen nun zwei solcher Hybridisierungsstellen in entgegengesetzter Richtung und in einer Entfernung unterhalb 5 kb, kann mit Hilfe der PCR der dazwischen liegende DNA-Bereich amplifiziert werden. Befinden sich nun innerhalb des

Genoms mehrere dieser Bereiche, was in der Regel auch vorkommt, werden diese auch amplifiziert. Da diese DNA-Abschnitte zwischen den beiden Primerbindungsstellen meistens unterschiedlich lang sind, werden in einer PCR-Reaktion unterschiedlich lange DNA-Fragmente amplifiziert, die man im Anschluß auf einem Agarose-gel auftrennen und auswerten kann. Trotz der Einfachheit dieser Methode, sind die Fehlerquellen ungleich größer als bei der konventionellen PCR. RAPD-Marker verhalten sich in der Regel dominant im Vergleich zu codominanten Isozym- oder RFLP-Markern (LYNCH & MILLIGAN 1994). Dies bedeutet, daß sich eine Homo- oder Heterozygotie nicht ohne weiteres erkennen läßt. In einem diploiden Organismus können somit beide Allele nicht getrennt analysiert werden, da die genetische Information der amplifizierten DNA-Fragmente nicht bekannt ist, mit der Ausnahme, daß alle RAPD-Marker sogenannte „inverted repeats“ darstellen, d.h. am Anfang und Ende des Fragments besitzen sie die gleiche (komplementäre) Primersequenz (10 bp). Bei Homozygotie der Ziel-DNA wird von beiden Allelen das gleiche Amplifikationsprodukt gebildet, wohingegen bei Heterozygotie nur ein Produkt gebildet wird. Erst durch den Einsatz von RAPD-Sonden in der RFLP-Technik können auch verschiedene Allele differenziert werden. Dies gilt auch für die Komigration gleich großer DNA-Fragmente im Agarosegel. Vergleicht man das Bandenmuster zweier Individuen miteinander, so können Banden, die im Agarosegel als identisch bewertet werden, trotzdem unterschiedlicher genetischer Natur sein. Untersuchungen haben gezeigt, daß die Einhaltung der Solltemperatur und die Temperaturverteilung innerhalb des Heizblocks größeren Schwankungen unterliegen kann (HOELZEL 1990; LINZ 1990). Insbesondere bei der Verwendung von Mikrotiterplatten in der RAPD-Analyse ist eine homogene Temperaturverteilung notwendig. Durch den Einsatz von Deckelheizungen und hochwertigen Heizblöcken werden diese Anforderungen weitgehend erfüllt. Inzwischen werden für viele Thermocycler speziell geformte Reaktionsgefäße angefertigt, die durch ihr dünnwandiges Material einen optimalen Kontakt mit dem Heizblock gewährleisten. Trotz dieser Fortschritte findet man Variationen im RAPD-Muster, verwendet man das gleiche Temperaturprofil auf verschiedenen Thermocyclern (HE et al. 1994; MACPHERSON et al. 1993; MEUNIER & GRIMONT 1993; PENNER et al. 1993).

Weitere Fehlerquellen können das Ergebnis wesentlich beeinflussen. Wird z.B. die Annealing-Temperatur zu niedrig angesetzt, kann der Primer unspezifisch an nicht komplementäre Stellen (z.B. an nur 8 statt 10 komplementäre Basen) des Genoms binden und damit zur Amplifikation unerwünschter DNA-Fragmente führen. $MgCl_2$ verändert die Syntheseigenschaften der DNA-Polymerase, wobei zu hohe Konzentrationen zu unspezifischen DNA-Fragmenten führt. Gleiches gilt für die Primerkonzentration, hohe

Konzentrationen produzieren zusätzlich kleinere DNA-Fragmente. Abb. 15 zeigt den Einfluß unterschiedlicher Reaktionseinstellungen auf das RAPD-Bandenmuster.

Verunreinigte DNA kann die Aktivität der DNA-Polymerase einschränken. Dadurch sind u.U. bestimmte Bereiche des Genoms nicht zugänglich und werden deshalb bei der Amplifikation nicht berücksichtigt. Die dadurch fehlenden DNA-Fragmente führen dann bei der Auswertung zu Fehlinterpretationen. Eine andere Fehlerquelle stellt die DNA-Polymerase dar. Dieses Enzym wird heute synthetisch hergestellt und besitzt keine Reparaturfunktion, um den neu synthetisierten Strang noch einmal zu überprüfen und falsch eingebaute Basen auszutauschen. Tritt während einer PCR-Reaktion dieser Fehleinbau in den ersten Reaktionszyklen auf, so wird der Fehler nicht erkannt und pflanzt sich somit exponentiell fort. Das Ergebnis ist ein DNA-Produkt mit einer falsch eingebauten Base. Die Gesamtfehlerhäufigkeit der *Taq*-Polymerase beträgt bei 30 Zyklen etwa 0,25 %, das entspricht einer Fehlereinbaurrate von $1,2 \times 10^{-4}$ oder 1/12000 pro Nukleotid und Zyklus (SAIKI et al. 1988; SWATSCHEK 1992). Dieser Fehler kann z.B. durch den Einsatz mehrere DNA-Moleküle als Matrize minimiert werden. Weiterhin führt die Verwendung von DNA- Polymerasen verschiedener Anbieter zu unterschiedlichen Ergebnissen (Abb. 16). Das bedeutet, daß vor Beginn der RAPD-Analysen eine sorgfältige Auswahl des Puffersystems nach qualitativen und finanziellen Kriterien getroffen werden muß (SNOWDON & LANGSDORF 1998). Letzteres gilt insbesondere bei der Auswahl des Enzyms DNA-Polymerase und dem Kauf eines Thermocyclers. Die RAPD-Analysen von *Brachiaria spp.* und *Zornia glochidiata* wurden mit einem eigens hergestellten Puffer durchgeführt. Weiterhin konnte für beide zu untersuchenden Pflanzen der gleiche Reaktionsmix benutzt und die gleichen Temperaturprofile im Thermocycler eingestellt werden. Die RAPD-Versuche für *Brachiaria* wurden auf einem Thermocycler der Marke *Hybaid Omnigene* und für *Zornia glochidiata* auf dem Thermocycler der Marke *MJ Research PTC-200* durchgeführt.

Das Probenvolumen einer PCR-Reaktion betrug 25 µl und setzte sich folgendermaßen zusammen:

- 21 µl PCR-Puffer
- 1 µl Primer-Lösung
- 1 µl genomische DNA (in TE gelöst; 10-50 ng / µl)
- 1 µl dNTP-Mix
- 1 µl DNA-Polymerase (1 µl = 1U)

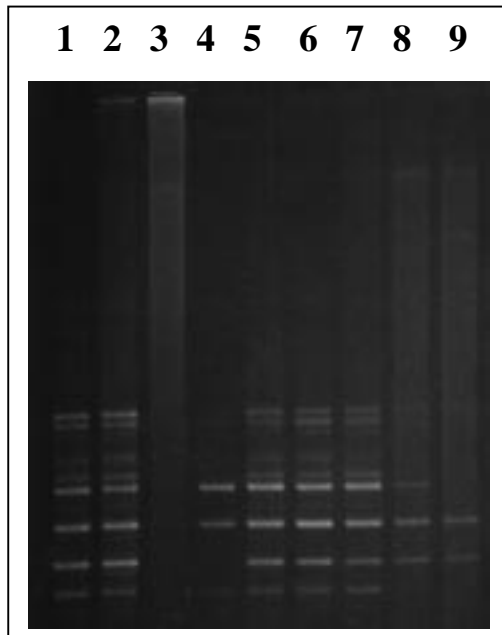


Abb. 15: Einfluß unterschiedlicher Reaktionsparameter auf das RAPD-Bandenmuster. Lane 1-3 und 7-9 wurden jeweils mit drei unterschiedlichen $MgCl_2$ -Konzentrationen (2.5, 3.5, 4.5 mM) amplifiziert, wobei in Lane 7-9 zusätzlich die Primer- und DNA-Konzentration halbiert wurde. Lane 4-6 zeigt den Einsatz verschiedener DNA-Polymerasekonzentrationen (0.5, 1.0, 1.5 U). Alle Reaktionen wurden mit der Pflanzenprobe Z-111 von *Zornia glochidiata* und dem Primer B-17 durchgeführt.

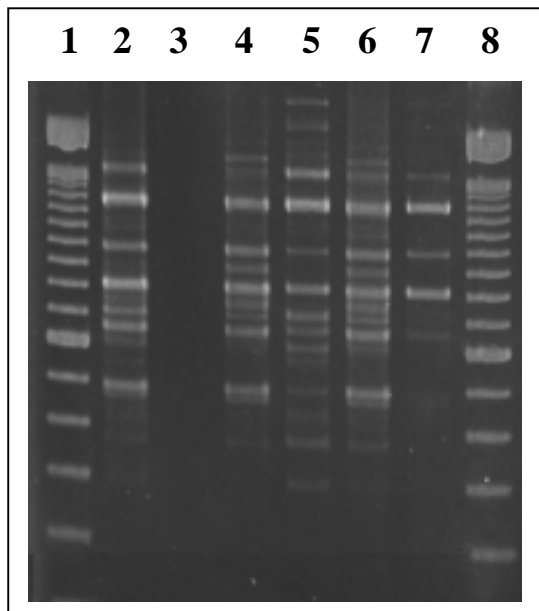


Abb. 16: RAPD-Bandenmuster bei *Zornia*, mit DNA-Polymerasen verschiedener Hersteller. Der Versuch wurde mit dem Standardprotokoll und den mitgelieferten Puffersystemen durchgeführt, mit Ausnahme von Lane 6 (Hier wurde der eigens angesetzte Puffer eingesetzt). Lane 1 und 8 sind Längenstandardmarker (BRL). Folgende Puffer- Polymerase- Systeme wurden verwendet: Lane 2 = Boehringer; Lane 3 = AGS; Lane 4 = BRL; Lane 5 = Pharmacia; Lane 6 = Pharmacia + eigener Puffer; Lane 7 = Supertaq.

Sämtliche Versuche wurden in Mikrotiterplatten (Fassungsvermögen für 96 Reaktionen) durchgeführt. Um eine Verdunstung des Puffers während der Aufheizphasen zu vermeiden, wurden alle Proben mit 50 μ l Mineralöl überschichtet.

Das Temperaturprofil setzte sich folgendermaßen zusammen:

- 1 min 94 °C (Denaturierung)
- 1 min 40 °C (Annealing)
- 2 min 72 °C (DNA-Synthese)
- 40 x wiederholen, dann
- 10 min 72 °C (DNA-Synthese)
- abkühlen auf 4 °C

Nach der PCR-Reaktion wurden die Proben im Kühlschrank aufbewahrt oder gleich auf ein Agarosegel aufgetragen. Dazu wurde der gesamte Probenansatz (25 μ l) mit Dye (1:6) angefärbt. Dieser Farbstoff enthält 30 % Glycerol, das sich an die DNA bindet und aufgrund der höheren Dichte beim Auftragen der Proben im Agarosegel ein Wegdiffundieren in den wäßrigen Puffer verhindert. Weiterhin befinden sich in diesem Farbstoff Bromphenolblau und Xylencyanol, die den Verlauf der Elektrophorese sichtbar machen. Hierbei bilden das Bromphenolblau eine dunkelblaue Elektrophoresefront, die etwa einem 400 bp großen DNA-Fragment entspricht und das Xylencyanol eine hellblaue Front von etwa 4000 bp. Da sich innerhalb einer RAPD-PCR die meisten DNA-Fragmente zwischen diesen beiden Molekulargrößen befinden, konnte so das Ende der Elektrophorese leichter abgeschätzt werden. Dies ist besonders dann wichtig, wenn unterschiedliche Agarosekonzentrationen zum Einsatz kommen und sich dadurch die Dauer der Elektrophorese verändert. In der Regel verwendet man in der RAPD-Analyse Agarosekonzentrationen von 1.4–2.0 % (1.4–2.0 g/100 ml), wobei eine höhere Konzentration eine längere Laufzeit bedeutet.

Aus dem Proben-Dye-Gemisch wurden nun jeweils 10 μ l auf ein Agarosegel aufgetragen, zusätzlich einen DNA-Längenstandard (BRL, 100 bp-Ladder), damit man nach Beendigung der Elektrophorese die Größe der einzelnen Fragmente bestimmen konnte. Dieser Marker wurde in der ersten und letzten Lane des Gels aufgetragen, denn die spätere Auswertung mit der dafür entwickelten Software RFLPscan benötigte pro Gel zwei Längenstandards. Nach der gelelektrophoretischen Auftrennung wurde das Gel mit Ethidiumbromid angefärbt, wobei sich der Farbstoff an die DNA anlagert und bei Bestrahlung mit UV-Licht orangefarbenes Licht emittiert. Sämtliche Agarosegele wurden nach der Anfärbung mit einer Polaroid S/W- Kamera fotografiert.

3.3.4 Datenauswertung

Die Photos wurden mit Hilfe eines Hand- bzw. Tischscanners bei 400 dpi Auflösung in den Computer eingescannt. Mit entsprechender Graphiksoftware wurden die Bilder für eine Auswertung mit der Spezialsoftware RFLPscan vorbereitet (Invertieren und Ausrichten).

3.3.4.1 RFLPscan

Das Programm RFLPscan (SCANALYTICS 1994) wurde für die Auswertung von RFLP-, RAPD- und Isozymdaten entwickelt. Hierbei werden zunächst das Photo eingelesen und das RAPD-Bandenmuster gescreent. Das Programm erkennt die Anordnung der Banden, markiert diese und ordnet jeder Bande ein Molekulargewicht zu (Abb. 17). Der Anwender muß nun entscheiden, welche dieser Banden er auswerten will, denn aufgrund der hohen Erkennungsrate des Programmes werden meist zu viele Banden identifiziert. Die Ursache dafür ist, daß die Auftrennung des Bandenmusters im Agarosegel nicht immer „optimal“ verläuft. Oft werden sog. „smiling lanes“ gebildet oder die Banden werden unterschiedlich stark amplifiziert. Agarosegele unterscheiden sich in der Regel durch ihre unterschiedliche Lauflänge des Bandenmusters. Erst durch das Auftragen eines Molekulargewicht-Längenstandards ist eine genaue Größenbestimmung der Banden möglich. Ein weiterer Vorteil des Programmes RFLPscan ist die Erstellung einer 1/0-Matrix. Die Auswertung dieser Daten erfolgt auf der Basis der An- oder Abwesenheit der entsprechenden Molekulargewichtsgrößen. Ist die Bande bei Individuum A vorhanden, wird sie mit 1 gewertet, fehlt sie hingegen bei Individuum B, wird eine 0 zugeordnet. Um Verwechslungen mit gleich großen Molekulargewichtsgrößen aus RAPD-Analysen anderer Primer zu vermeiden, wurde pro Agarosegel nur das RAPD-Bandenmuster eines Primers aufgetragen.

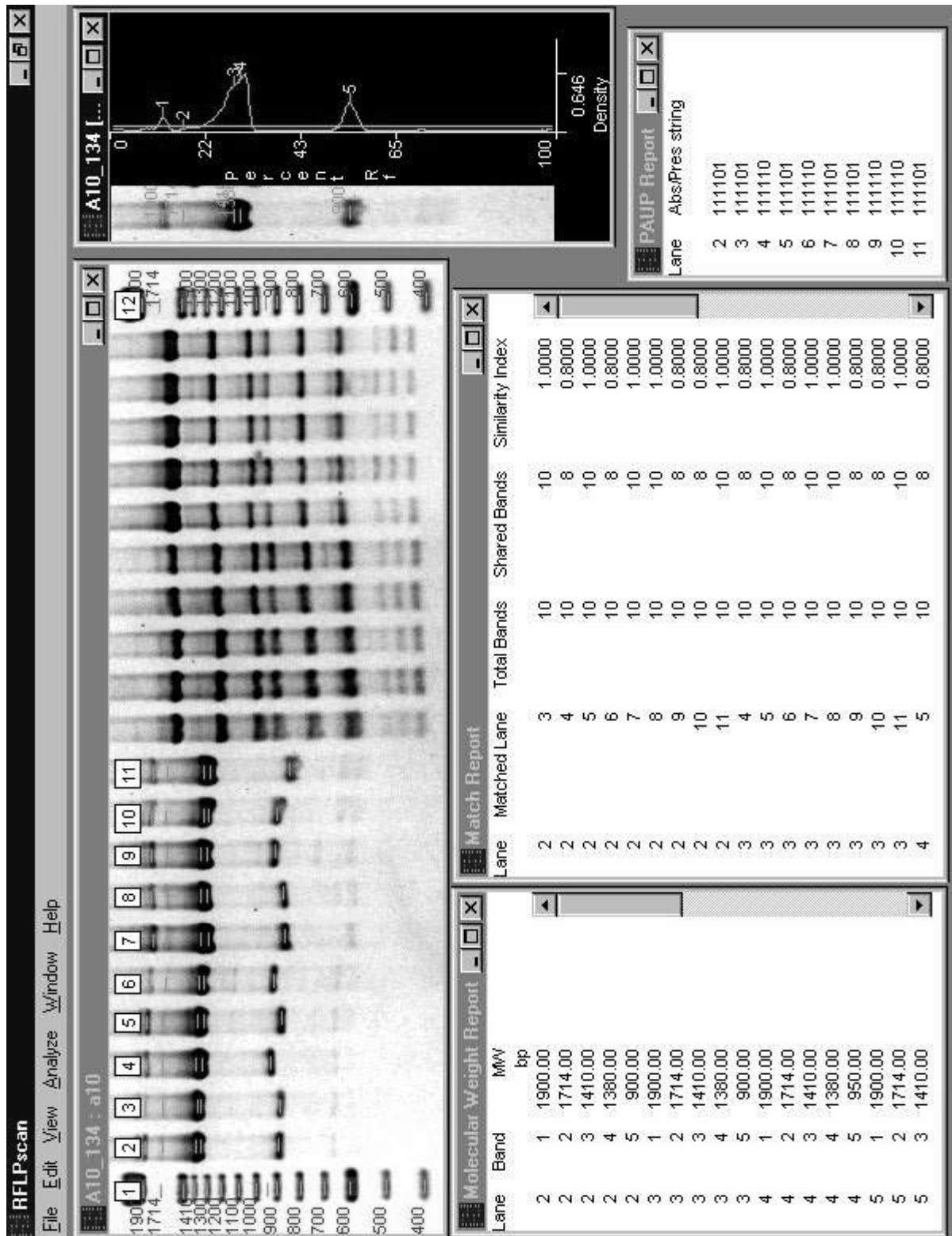


Abb. 17: RFLPscan Software zur Auswertung der RAPD-Daten.

Abgebildet ist ein Screenshot des Programmes. Bild oben links: RAPD-Bandenmuster mit zwei Molekulargewichtsmarkern (Lane 1 und 12); Zur besseren Auswertung wurde das Polaroid-Photo invertiert. Bild oben rechts: Angabe des RF-Wertes und der Bandenintensität. Bild unten links: Für jede Bande wird das Molekulargewicht berechnet. Bild unten Mitte: Berechnung des Similarity-Index (simple matching) und der Anzahl polymorpher Banden. Bild unten rechts: Berechnung einer 1/0-Matrix.

3.3.4.2 Distanz- und Ähnlichkeitsmaße

Die genetische Diversität bezeichnet die Vielfalt bzw. Mannigfaltigkeit in der genetischen Zusammensetzung von Organismengruppen auf derselben oder auf unterschiedlichen Organisationsstufen. Sie ist das Ergebnis der Evolution und daher die Konsequenz der Wirkung der verschiedenen Evolutionsfaktoren, insbesondere von Mutation, Selektion, Zufallsdrift und Migration sowie der Vermehrungsstruktur von Populationen. Auch der Einfluß des Menschen durch Domestizierung und gezielte Auswahl in der Züchtung spielt dabei eine wichtige Rolle (SPERLICH 1988; HINTUM 1995). Die Charakterisierung der genetischen Diversität kann auf der Ebene von Individuen, Populationen, Arten oder auch hinsichtlich höherer taxonomischer Einheiten erfolgen (PONS et al 1998). Die genetische Diversität beschreibt das Ausmaß an Unterschieden im gesamten genetischen Material (genomische und zytoplasmatische DNA) der untersuchten Organismengruppe. Bei der eingeschränkten Materialauswahl ist die Erfassung und Erhaltung der genetischen Variation für eine Genbank oder für einen Pflanzenzüchter von besonderer Bedeutung. Ziel der Konservierung genetischer Ressourcen ist eine vollständige, zumindest eine repräsentative Auswahl der vorhandenen genetischen Variation einer Pflanzenart.

Um einer solchen Zielvorstellung gerecht zu werden, genügt es nicht, die genetische Variation innerhalb und zwischen den einzelnen Gruppen aufzuzeigen. Eine Struktur ist oft nicht von vornherein bekannt und zudem kann nicht vorausgesetzt werden, daß eine vermutete Gruppierung den „wahren“ genetischen Strukturen im Untersuchungsmaterial entspricht.

Es ist also sinnvoll und notwendig, bei der Analyse der experimentellen Ergebnisse Methoden anzuwenden, die eine Charakterisierung des Materials entsprechend der Zielvorstellung gestatten.

Genetische Distanzen drücken genetische Unterschiede / Ähnlichkeiten zwischen Individuen oder Populationen durch eine einzige Zahl aus, die in der Regel zwischen 0 (kein Unterschied) und 1 (maximaler Unterschied) liegt (DICE 1945). Die Analyse der genetischen Diversität anhand von RAPD-Marker wird auf der Basis von binären Daten durchgeführt. Hierbei liegt für jedes Individuum ein Vektor von Nullen und Einsen vor, der das Vorhandensein (1) oder Fehlen (0) des beobachteten Merkmals, z.B. einzelner Banden auf einem Agarosegel, charakterisiert. Für die Berechnung von Ähnlichkeiten bzw. Distanzen wurden zwei Koeffizienten verwendet.

Der Ähnlichkeitskoeffizient S_{xy}

S_{xy} bezeichnet den Ähnlichkeitskoeffizienten (Dice-Index, DICE 1945), der auf der Anzahl gemeinsamer Banden beruht, dessen Wertebereich zwischen 0 (keine gemeinsamen Banden) und 1 (gemeinsame Banden) liegt. Da jeder Primer pro Individuum eine bestimmte Anzahl an Banden produziert, können diese Daten als Binärdatei bearbeitet werden. Der Dice-Index ist identisch mit der genetischen Distanz nach NEI & LI (1979). Somit berechnet sich der Ähnlichkeitskoeffizient S_{xy} wie folgt:

$$S_{xy} = \frac{2n_{xy}}{n_x + n_y + 2n_{xy}}$$

x und y bezeichnen die beiden zu vergleichenden Individuen
 n_{xy} Anzahl gemeinsamer Banden
 n_x und n_y Anzahl Banden der Einzelindividuen

Im Unterschied zum Tanimoto Ähnlichkeitsmaß (Jaccard-Index) werden beim Dice-Index die Anzahl gemeinsamer Banden doppelt gewertet. Damit stellt S_{xy} ein qualitatives Ähnlichkeitsmaß dar, das besonders das Ausmaß der genetischen Konservierung widerspiegelt. Zudem ist der Dice-Algorithmus identisch mit der genetischen Distanz $(1 - S_{xy})$ nach NEI & LI (1979), die in der statistischen Auswertung von RAPD-Daten breite Anwendung gefunden hat (HUFF et al. 1993, MÖSGES & FRIEDT 1994, SKROCH et al. 1992, OROZCO-CASTILLO et al. 1994, LYNCH & MILLIGAN 1994, TAO et al. 1993, TOHME et al. 1996, M'RIBU & HILU 1994, CLARK & LANIGAN 1993, YU & PAUL 1993, SHAH et al. 1994).

Die quadrierte euklidische Distanz D

Der Koeffizient D bezeichnet die quadrierte euklidische Distanz, d.h. der Abstand zwischen zwei Individuen ist die Quadratwurzel aus der Summe der quadrierten Wertedifferenzen über jeder Variablen. Hierbei werden gemeinsame Banden nicht berücksichtigt. Nach HUFF et al. (1993) sind die genetische Distanz nach NEI & LI und die euklidische Distanz D nahezu austauschbar. D stellt ein quantitatives Maß für die Unähnlichkeit zweier Individuen dar und ist wie folgt definiert:

$$D = \sum_k (n_{kx} - n_{ky})^2$$

x, y stehen für die beiden zu vergleichenden Individuen,
 k Anzahl vergleichbarer Banden jedes Individuums
 n_x, n_y Anzahl Banden jedes Individuums

Die quadrierte euklidische Distanz D wurde zur Errechnung der Varianzkomponenten mit dem WINAMOVA Programm von EXCOFFIER (1992) verwendet.

3.3.4.3 Clusteranalyse

Mit dem Programm SPSS für Windows 6.1.3 wurden verschiedene hierarchische agglomerative Clusterverfahren gerechnet und miteinander verglichen (BAVERAGE, WAVERAGE). In beiden Fällen erhält man die gleiche Clusteranordnung, so daß die Darstellung aller Clusteranalysen auf den am weitestverbreiteten UPGMA-Algorithmus (BAVERAGE, „unweighted pair-group method using arithmetic averages“) beschränkt werden kann (SNEATH & SOKAL 1973). Mit Hilfe dieses Algorithmus wird während einer Clusteranalyse versucht, relativ homogene Fallgruppen oder Variablen zu identifizieren und diese dann in separate Cluster eingeordnet. Diese Cluster werden dann agglomerativ zu einem hierarchisches Cluster aufgebaut und in einem Dendrogramm grafisch dargestellt. Individuen mit ähnlichem genetischen Datensatz können dann in einem gemeinsamen Cluster angeordnet werden. Die Position bzw. die mehr oder weniger starke Abgrenzung dieser Cluster innerhalb des Dendrogramms beschreibt dann die genetische Struktur bzw. die genetische Diversität von Populationen.

3.3.4.4 Varianzanalyse

Nähere Informationen über den Grad der genetischen Variabilität innerhalb und zwischen Populationen wurden über eine AMOVA (Analysis of Molecular Variance) ermittelt. Dazu wurde das Softwareprogramm WINAMOVA Version 1.55 (EXCOFFIER 1992) verwendet. Als Datenbasis diente die quadrierte euklidische Distanz D . Die Varianz s_x^2 ist die Summe der Abweichungsquadrate (SQ = sum of squares) aller Meßwerte einer Verteilung von ihrem Mittelwert, dividiert durch den Freiheitsgrad (n-1):

$$s_x^2 = \frac{SQ}{n-1} = \frac{1}{n-1} \cdot \sum (x_i - \bar{x})^2$$

wobei	n	= Stichprobenumfang
	n-1	= Freiheitsgrad
	x_i	= Meßwert
	\bar{x}	= arithmetisches Mittel der Meßwerte

Die Varianz beschreibt eine Methode zum Prüfen der Nullhypothese mehrerer gleicher Gruppenmittelwerte in der Grundgesamtheit (Gesamtstichprobe). Je größer die Streuung zwischen den Gruppenmittelwerten bzw. je kleiner die Streuung innerhalb der Gruppen ist, desto größer ist die Vermutung, daß es einen Unterschied der Gruppenmittelwerte in der Grundgesamtheit gibt. Dabei wird die für die Gruppenmittelwerte geschätzte Stichprobenvarianz mit derjenigen innerhalb der Gruppen verglichen.

3.3.4.5 Hauptkoordinatenanalyse

Ausgangspunkt einer Hauptkoordinatenanalyse (principle coordinates analysis) ist eine Vielzahl von Variablen, von denen nicht a priori bekannt ist, ob und in welcher Weise sie miteinander korreliert sind. Mit Hilfe der Hauptkoordinatenanalyse werden Gruppen von Variablen untersucht, die durch eine „Hintergrundvariable“ (BROSIUS 1989) charakterisiert werden können. Solche „Hintergrundvariablen“ werden auch Faktoren genannt. Ziel der Hauptkoordinatenanalyse ist es, den Grad der Komplexität, der durch eine Vielzahl von Variablen dargestellt wird, durch möglichst wenige Faktoren wiedergeben zu können.

Ausgangsbasis für eine Hauptkoordinatenanalyse ist die Distanzmatrix, die über den Dice-Index (S_{xy}) für alle Individuen berechnet wurde. Der Vorteil der Hauptkoordinatenanalyse liegt in der besseren graphischen Darstellung großer Individuenzahlen, wohingegen individuelle genetische Unterschiede in einer Clusteranalyse deutlicher werden.

Die Interpretation der Faktoren ergibt sich aus der Relation, die sie zu den Beobachtungsvariablen haben, für die sie die Hintergrundvariable darstellen. Diese Daten lassen sich aus der Faktorenmatrix entnehmen. Hierbei zeigen große Faktorladungen eine große, kleine eine kleine Bedeutung eines Faktors für die entsprechende Variable an. Ein Faktor läßt sich dann leicht interpretieren, wenn einige Variablen, die gleiche Parameter aufweisen, (Standort, Niederschlag, Sammeljahr) hoch auf ihn laden und die anderen

gleichzeitig nur gering. Umgekehrt ist ein Faktor schwer zu interpretieren, wenn er zu allen Variablen eine gleich hohe Korrelation zeigt.

3.3.4.6 Diskriminanzanalyse

Ausgehend von der Herkunft der Individuen besteht grundsätzlich die Möglichkeit, jedes Individuum einer geographischen Gruppe zuzuordnen. Diese Gruppierung kann anschließend mit einer Diskriminanzanalyse der genetischen Marker geprüft werden.

In der vorliegenden Arbeit wurden unterschiedliche Standorte derselben Pflanzenart mit RAPD-Markern untersucht. Da einige dieser Standorte sehr nah, andere hingegen sehr weit voneinander entfernt waren und verschiedenen ökoklimatischen Zonen angehören, wurde geprüft, ob die Individuen eines oder mehrerer Standorte einem gemeinsamen Ökotyp zugeordnet werden können. Als Datenbasis wurde eine 1/0-Matrix der RAPD-Daten erstellt und mit den Softwareprogrammen SPSS 7.5 und NTSYS 2.0 die Diskriminanzfunktionen berechnet. Es wurde keine Schätzung der Diskriminanzfunktion mit Hilfe einer zweiten bzw. einer Teilstichprobe durchgeführt, da aufgrund der geringen Individuenzahlen ein beträchtlicher Informationsverlust entstehen und damit das Ergebnis stark beeinflussen könnte.