

Lehrstuhl und Institut für Genetik
der Technischen Universität München

Genetische und molekulare Charakterisierung
der Centromer 1-Region
von *Arabidopsis thaliana*

Wolfgang E. Haupt

Vollständiger Abdruck der von der Fakultät Wissenschaftszentrum Weihenstephan
für Ernährung, Landnutzung und Umwelt
der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr.rer.nat. A. Gierl

Prüfer der Dissertation:

1. Priv.-Doz. Dr.rer.nat., Dr.rer.nat.habil. R. A. Torres Ruiz
2. Univ.-Prof. Dr.rer.nat., Dr.rer.nat.habil. G. Wenzel

Die Dissertation wurde am 21.02.2001 bei der Technischen Universität München
eingereicht und durch die Fakultät Wissenschaftszentrums Weihenstephan
für Ernährung, Landnutzung und Umwelt am 17.08.2001 angenommen.

für Brigitte und Philipp

Danksagung:

Diese Arbeit wurde am Lehrstuhl und Institut für Genetik der technischen Universität München durchgeführt.

Meinem Betreuer PD Dr. Ramon Torres-Ruiz danke ich für die Bereitstellung des Themas, die Diskussionsbereitschaft und die Unterstützung während der Arbeit.

Ich möchte mich auch bei Prof. Dr. Alfons Gierl und Dr. Monika Frey für die Unterstützung dieser Arbeit bedanken.

Die Arbeit wurde von der DFG gefördert.

Mein besonderer Dank gilt Georg Haberer, der mit großer Gesprächs- und Hilfsbereitschaft jederzeit zur Seite stand.

Für eine gute Arbeitsatmosphäre und die Hilfsbereitschaft möchte ich allen Mitarbeitern des Lehrstuhls, besonders Hedi, Christine, Uta, Conny, Anita, Tim, Albert, Sebastian und ganz besonders Andrea, Parvin und unseren Mädels Sabine und Simin danken.

Meiner Familie danke ich ganz besonders für die große Geduld.

Inhaltsverzeichnis

1 Einleitung	1
2 Material & Methoden	5
2.1 Chemikalien und allg. molekularbiologische Materialien	5
2.2 spezielle Materialien für FISH-Analysen	5
2.3 Pflanzen Material	6
2.4 Bakterienstämme	6
2.5 λ -Phagen	7
2.6 Vektoren	7
2.7 Klone	7
2.8 Bibliotheken	7
2.8.1 BAC-Bibliotheken	7
2.8.2 cDNA-Bibliotheken	8
2.9 Allgemeine molekularbiologische Methoden	8
2.10 Gesamt RNA-Präparation aus Pflanzen	8
2.11 mRNA-Präparation (PolyAtract® IV Promega)	9
2.12 Formaldehyd Gelelektrophorese	10
2.13 Northern Blot	11
2.14 Labeling einer DNA-Probe	11
2.15 Hybridisierung geblotteter Filter	12
2.16 Fluorescent in situ hybridisation (FISH)	12
2.16.1 Präparation meiotischer Chromosomen	12
2.16.2 Vorbehandlung der Präparate	13
2.16.3 Nick-Translation	13
2.16.4 In situ-Hybridisierung und Waschschrte	13
2.16.5 Detektion und Amplifikation	14
2.16.6 spezielle Lösungen und Puffer	14
2.17 DNA Präparationen	15
2.17.1 Pflanzen	15
2.17.2 Plasmide	16
2.17.3 BACs	16
2.18 PCR	16
2.19 Kartierungen	19
2.19.1 morphologische Marker	19
2.19.2 RFLP	19
2.19.3 CAPS	19

2.19.4 HAMDE	20
2.19.5 Berechnung der genetischen Abstände	21
2.20 cDNA screens	23
2.21 DNA Sequenzierung	24
2.22 Computer Analysen	24
3 Ergebnisse	27
3.1 Genetische Charakterisierung der Centromer 1-Region	27
3.1.1 Kartierung von der GapB (rechte) Seite	27
3.1.2 Kartierung von der <i>tt1</i> (linke) Seite	28
3.1.3 Genetische Karte der Centromer 1-Region	29
3.2 Molekulare Charakterisierung der Centromer 1-Region	30
3.2.1 BAC-Subklone	31
3.2.1.1 Klonierung	32
3.2.1.2 Hybridisierung	32
3.2.1.3 Sequenzierung	37
3.2.2 Single copy Proben	38
3.2.2.1 Hybridisierungen auf IGF BAC-Filter	39
3.2.3 Repetitive Proben	39
3.2.3.1 Hybridisierungen auf IGF BAC-Filter	40
3.2.3.2 BAC Fingerprint Analysen	42
3.2.4 BAC contig	44
3.2.5 Vergleich der genetischen mit der physikalischen Karte	48
3.3 Histologische Charakterisierung der Centromer 1-Region	51
3.3.1 FISH mit BACs	52
3.3.2 FISH mit dem Klon a2-4	53
3.4 cDNA screens	54
3.4.1 Screening mit BACs F17A20 und F17G21	54
3.4.1.1 Gesamt-Bank	54
3.4.1.2 Grün-Bank	55
3.4.1.3 Blüten-Bank	55
3.4.2 Screening mit BAC F9D18, F6J23 und F3L22	55
3.4.2.1 Grün-Bank	55
3.4.2.2 Blüten-Bank	56
3.5 Northern-Analysen	57
3.6 Analyse sequenzierter BACs (aus dem contig)	59
3.6.1 BAC F28L22	60
3.6.2 BAC F9D18	61

3.6.3 weitere BACs auf dem 1 Chromosom	62
4 Diskussion	63
4.1 Vergleich des erstellten contigs (3.2.4) mit bestehenden contigs	63
4.2 Vergleich der genetischen mit der physikalischen KarteLage des Heterochromatins	64
4.2.1 Kompression bzw. Rekombinationsunterdrückung	65
4.2.2 Kompressionszunahme zwischen F1c13t7 und F4a3t7	66
4.2.3 Lage und Wirkung des Heterochromatins	67
4.2.4 Vorkommen von Genen in der CEN 1-Region	67
4.3 Struktur der CEN1 Region	68
5 Zusammenfassung	72
6 Literaturverzeichnis	73
7 Anhang	81
7.1 Genetische Kartierung	81
7.1.1 Rekombinante Pflanzen von der GapB-Seite	81
7.1.2 Rekombinante Pflanzen von der <i>tt1</i> -Seite	81
7.2 Single copy-Proben	83
7.3 Repetitive Proben	85
7.3.1 BACs	85
7.3.2 Sequenzen	89
7.3.2.1 a 2-4 (2777 bp)	89
7.3.2.2 as 31 (2569 bp)	90
7.3.2.3 gs 31 (2265 bp)	91
7.4 GENSCAN und BLAST-Analysen	92
7.4.1 BAC F28L22	92
7.4.2 BAC F9D18	97
7.5 <i>Arabidopsis</i> Physical Map close-up	100
7.5.1 links von CEN 1 (Nia2)	100
7.5.2 rechts von CEN 1 (Gap B)	101

Abkürzungen

μ	Micro-
A. thaliana	Arabidopsis thaliana
A.t.	Arabidopsis thaliana
Abb.	Abbildung
abs	absolut
BAC	Bacterial Artificial Chromosome
bp	Basenpaar
BPB	BromPhenolBlau
BSA	Rinderserumalbumin
C:IA	Chloroform:Isoamylalkohol (24:1)
CAPS	cleaved amplified polymorphic sequence
cDNA	komplementäre DNA
CEN	Centromer
Ci	Curie (1 Ci = 2.22×10^{12} Zerfälle/Minute)
cM	Centimorgan
DAPI	4'-Diamide-2-Phenylindole-Dihydrochloride
Col (C)	Columbia
CTAB	Hexadecyltrimethylammoniumbromid
dATP	Desoxyadenosin-5'-triphosphat
dCTP	Desoxycytidin-5'-triphosphat
DEPC	Diethylpyrocarbonat
dest	destilliert
dGTP	Deoxyguanosin-5'-triphosphat
DMSO	Dimethylsulfoxid
DNA	Desoxyribonukleinsäure (-'acid')
dNTP	Desoxynukleotidtriphosphat
DTT	Dithiothreitol
dTTP	Desoxythymidin-5'-triphosphat
E. coli	Escherichia coli
EDTA	Ethylendiamintetraacetat - Na ₂ -Salz
EMS	Ethylmethansulfonat
EtBr	Ethidiumbromid
EtOH	Ethanol
FA	Formaldehyd
FISH	Fluorescence In Situ Hybridisation
FITC	Fluorescein isothiocyanate isomer I
g	Gramm
<i>GAPB</i>	Glycerinaldehyd-3-phosphatdehydrogenase-B-Gen
GAPB	ebenso, Marker
<i>GK</i>	Gurke-Gen
<i>GURKE</i>	Gurke-Gen
GURKE	Gurke-Genprodukt
<i>gk</i>	<i>gurke</i> (Phänotyp)
GUS	β -Glucuronidase
<i>GUS</i>	β -Glucuronidase-Gen
h	Stunde
HAMDE	Heteroduplex Analysis Mutation Detection Enhancement

H ₂ O	Wasser
i.d.R.	in der Regel
KAc	Kaliumacetat
kb	Kilobasenpaare
l	Liter
Ler (L)	Landsberg erecta
Lsg.	Lösung
lsg.	Lösung
m	Milli-
min.	Minuten
mind.	mindestens
morpho.	morphologisch
mRNA	messenger RNA
MS-Salze	Murashige & Skoog-Salze
n	Nano-
NaAc	Natriumacetat
NBSB	Northen Blot Sample Buffer
Nd (N)	Niederzenz
NIA2	Nitratreduktase-2-Gen
NIA2	ebenso, Marker
OT	Objekträger
PCR	Polymerase Kettenreaktion ('Polymerase Chain Reaction')
RFLP	Restriction Fragment Length Polymorphism
RNA	Ribonukleinsäure
rpm	rounds per minute (Umdrehungen pro Minute)
RT	Raumtemperatur
SDS	Natriumdodecylsulfat
sek..	Sekunden
SSPE	Natrium-Phosphat-EDTA-Puffer
Tab.	Tabelle
TAIR	The Arabidopsis Information Ressource
Taq-Pol.	Thermus aquaticus-DNA-Polymerase
TBE	Tris-Borat-EDTA-Puffer
TE	Tris-EDTA-Puffer
Tris	Tris-(hydroxymethyl)-aminomethan
U	'Unit', Einheit der Enzymaktivität
üN	über Nacht
usw.	und so weiter
v/v	Volumen zu Volumen
vgl.	vergleiche
w/v	Gewicht zu Volumen
WB	Wasserbad
WT	Wildtyp
<i>wt</i>	Wildtyp
z.B.	zum Beispiel
z.T.	zum Teil

1 Einleitung

Die Verteilung der Erbinformation spielt bei der Zellteilung eine entscheidende Rolle. Die Centromere sind bei der gleichmäßigen Verteilung der Schwesterchromatiden auf die Tochterzellen während Mitose und Meiose essentiell beteiligt. Dabei dienen sie u.a. als Ansatzstellen für den Spindelapparat, entlang dessen die Verteilung stattfindet. (Karpen and Allshire, 1997; Richards and Dawe, 1998)

Arabidopsis thaliana (Ackerschmalwand) ist eine zweikeimblättrige, einjährige Blütenpflanze, die zu der Familie der Brassicaceae gehört. Aufgrund ihrer besonderen Eigenschaften wurde *Arabidopsis thaliana* als genetisches Studienobjekt vorgeschlagen (Redei, 1975; Meyerowitz und Pruitt, 1985). Neben ihrer geringen Größe und kurzen Generationszeit hat *Arabidopsis* ein relativ kleines Genom von ca. 100 Mbp, arm an dispers verteilten, repetitiven Sequenzen, das im haploiden Satz aus fünf Chromosomen besteht (Redei, 1970; Meyerowitz, 1989). In den letzten Jahren wurde *Arabidopsis* intensiv untersucht. Dies führte zu sehr genauen genetischen (Arabidopsis Genomic View; <http://genome-www3.stanford.edu/Arabidopsis/chromosomes/> ; <http://www.arabidopsis.org>) und physikalischen Karten (The MPI-MP-Map of the Arabidopsis thaliana Genome, Mozo et al., 1998; http://www.mpimp-golm.mpg.de/101/mpi_mp_map/access.html). Mittlerweile ist das Genom von Arabidopsis fast vollständig sequenziert (Genome Sequencing at the ATGC; <http://genome.bio.upenn.edu/SPP.html>; The Arabidopsis Genome Initiative, 2000). Komplexe DNA-Strukturen (Centromere können mehrere hundert Kb lang sein) und die vielen verschiedenen enthaltenen repetitive Elemente verhinderten bisher eine genaue Analyse der Centromer-Regionen, so daß dieser Teil des Genoms wenig erforscht ist.

Das Chromatin setzt sich im wesentlichen aus Euchromatin und Heterochromatin zusammen. Im Euchromatin findet man die meisten transkribierten Gene. Das stark kondensierte Heterochromatin wird dagegen nicht transkribiert (Griffiths, 1993). Heterochromatin setzt sich meist aus repetitiven DNA Sequenzen zusammen, wie man sie in den Telomeren und Centromeren findet (Franklin and Cande, 1999; Richards and Dawe, 1998; The Arabidopsis Genome Initiative, 2000).

Chromosomen stellen die Transporteinheiten der genetischen Information dar. Neben den Telomeren ist das Centromer ein wichtiges funktionelles Element. Das Centromer bezeichnet eine eingeschnürte Region eines Metaphasen-Chromosoms, die die Tochterchromatid Kohäsion unterstützt. Es ist auch die Ansatzstelle des Spindelapparates an die Chromosomen, mit dessen Hilfe während der Anaphase die Tochterchromatiden zu den entgegengesetzten Spindelpolen gezogen werden (Mc Kim and Hawley, 1995; Pluta et al., 1995; Allshire, 1997). Dabei interagieren die Mikrotubuli des Spindelapparates nicht direkt mit der DNA (CEN-DNA) bzw. dem Chromatin, der übergeordneten Struktur der Centromer Region, sondern mit dem sogenannten Kinetochor, einem Proteinkomplex (Allshire, 1997; Richards and Dawe, 1998). Dieses dient als Verbindungsstelle zwischen Chromosom und Spindelapparat. In diesem Zusammenhang wird der Begriff Centromer häufig für den Komplex aus CEN-DNA und Kinetochor benützt.

Kinetochore können holocentrisch oder polycentrisch organisiert sein (Pimpinelli and Goday, 1989; Traut, 1991; Pluta et al., 1995). *S. cerevisiae* hat mit 125 bp ein sehr kleines lokalisiertes Punktcentromer, das frei von repetitiven Sequenzen ist (Pluta et al., 1995; Clarke, 1998). *S. pombe*, ein naher Verwandter, weist dagegen ein 40 - 100 kb großes, regional begrenztes Centromer auf, welches viel komplexer aufgebaut ist und repetitive Elemente enthält. Regionale Centromere findet man auch bei Pflanzen. Diese können bei *Arabidopsis* größer als 1 Mb und sehr komplex aufgebaut sein (Round et al., 1997). Sie sind mit den Centromeren von Tieren, die 40 kb - mehrere Mb umfassen, vergleichbar (Richards and Dawe, 1998).

Das Vorkommen von repetitiven Elementen in den Centromeren von Pflanzen und Tieren führt zur Frage nach der Funktion dieser Elemente. Einige dieser repetitiven Elemente, z.B. alphoid-DNA bei Tieren, können die Bildung von Centromeren induzieren (Harrington et al., 1997). Andererseits konnte gezeigt werden, daß CEN-DNA alleine nicht immer ausreichend ist, um ein Centromer zu bilden. Manchmal reicht "nicht centromerische" DNA aus, um Centromer Funktionen zu übernehmen. Diese sogenannten Neocentromere, wie sie bei Mensch, *Drosophila* und auch in Pflanzen gefunden werden, übernehmen manchmal die Funktion, aber nicht immer alle Eigenschaften eines Centromers (Karpen and Allshire, 1997; Willard, 1998; Richards and Dawe, 1998). Diese Beobachtungen führten zu dem Vorschlag, daß die "normale" Centromer-Funktion epigenetisch reguliert ist. Epigenetisch bedeutet hier, daß vererbliche Veränderungen der Centromer-Funktion sich nicht in Veränderungen der entsprechenden primären DNA-Sequenz niederschlagen (Karpen and Allshire, 1997). Die Centromeren werden also von einer übergeordneten, höheren Ordnung des Chromatins bestimmt. Deshalb ist es von großer Wichtigkeit, sowohl die DNA als auch die Protein-Komponenten des Centromers zu kennen. Leider ist bei Pflanzen auf diesem Gebiet sehr wenig bekannt. Man muß auf Daten, die aus humanen bzw. tierischen Systemen abgeleitet sind, zurückgreifen. Deshalb besteht hier ein großer Bedarf an genuinen Daten aus der Pflanzenwelt.

Da der Frage nach der Position des funktionellen Centromers und seinen DNA-Bestandteilen eine große Bedeutung zukommt, wurde versucht, die Position repetitiver Elemente mit der Position von Centromeren zu korrelieren. Dazu wurde in Mais das B-Chromosom auf seinen Gehalt an B-repeats untersucht. Es konnte gezeigt werden, daß der Mais "B-repeat" ein 1,4 kb tandem array ist, der sich über 9 Mb erstreckt. Er ist über die gesamte Länge des Chromosom B-Centromers verteilt und Teil seiner funktionellen Region. Trotz signifikanter Veränderungen der Kopienzahl des "B-repeats" behält das B-Centromer seine Funktion bei (Alfenito and Birchler, 1993; Kaszas and Birchler 1996). Das "knob" Heterochromatin von Mais besteht hauptsächlich aus einem repetitiven 185 bp-Element. Dieses zeigt eine hohe Homologie zu dem "B-repeat" (Peacock et al., 1981; Alfenito and Birchler, 1993). Der "B repeat" kommt in Chromosom A-Centromeren nicht vor. Centromerische Chromosom A-Elemente sind Sau3A9 aus *Sorghum*, 745 bp lang und CSS1 aus *Brachipodium*, 250 bp lang aber nur teilweise homolog (Aragon-Alacide et al., 1996; Jiang et al., 1996). Dies zeigt, daß verschiedene repetitive Elemente innerhalb eines Organismus ein Centromer aufbauen können. Einige der Ty3 - gypsy Retrotransposon-Elemente von Gerste kommen hauptsächlich am Centromer vor und weisen Homologien zu Sau3A9 auf. Diese Elemente sind in verschiedenen Getreidearten - Mais, Reis, Weizen, Gerste, Roggen und Hafer - hoch konserviert (Presting et al., 1998). Die Sequenz von CSS1 ist homolog

zur tierischen CENP-B box konsensus Sequenz. Diese findet man auch im 180 bp knob repeat wieder (Richards and Dawe, 1998). CENP-B hat eine helix-turn-helix-Domäne und bindet an das Centromer (Iwahara et al., 1998). Es könnte ein zentrales Element des funktionellen Centromers darstellen.

In Dicotyledonen findet man im Gegensatz dazu ein 180 bp langes repetitives Element aus *Arabidopsis thaliana* (Martinez-Zapater et al., 1986; Simoens et al., 1988; Richards et al., 1991; Murata et al., 1994). Dieser repeat (pAL1) wurde mit verschiedenen Methoden in allen fünf *Arabidopsis*-Centromeren gefunden, und zwar mit FISH (Maluszynska and Heslop-Harrison, 1991; Franz et al., 1998), Segregationsanalyse von dedektierbaren Pulsfeld-Gelelektrophorese Polymorphismen (Round et al., 1997) und Tetradenanalyse mit der *quartet1*-Mutanten (Copenhaver et al., 1998). Das *Arabidopsis*-Centromer, welches mehr als 1 Mb umfassen könnte (Murata et al., 1994; Round et al., 1997), beinhaltet wahrscheinlich mehr als eine Klasse repetitiver Elemente. Mindestens ein weiteres Element, eine degenerierte Kopie des Athila Retroelements (Klon 106B), findet man am Centromer (Pellisier et al., 1996; Thompson et al., 1996a,b; Franz et al., 1998), wo es offensichtlich zwischen dem 180 bp repeat eingestreut vorkommt (Franz et al., 1998). Es wurde auch gezeigt, daß Sequenzen des BACs F17A20 auf beiden Seiten aller *Arabidopsis*-Centromeren vorkommen, wo sie den 180 bp und 106B repeat flankieren (Franz et al., 1998). Weitere repetitive Elemente könnten mit fortschreitendem *Arabidopsis* Genom-Projekt (<http://www.tair.org>) in pericentromerischen BACs gefunden werden. Drei Elemente, die Subklone der IGF-BACs F17A20 und F17G21 darstellen, werden in dieser Arbeit beschrieben. Der 180 bp *Arabidopsis thaliana* repeat scheint in nahe verwandten Arten nicht konserviert zu sein. Der repeat pAa214 von *Arabidopsis arenosa* zeigt nur schwache Ähnlichkeiten zum pAL1 repeat, wo hingegen andere *Arabidopsis* Arten, wie *A. wallichii*, *A. pumila* und *A. griffithiana*, in FISH-Analysen keine Sequenzhomologien zum pAL1 repeat aufzeigen (Maluszynska and Heslop-Harrison, 1993; Kamm et al., 1995).

Will man etwas über die Funktion des Centromers lernen, so muß man vorher seinen Aufbau und seine Struktur analysieren. Mittlerweile wird das *Arabidopsis* Genom auf 130 Mb geschätzt (Somerville, 1999; The Arabidopsis Genome Initiative, 2000). Die Centromeren der fünf *Arabidopsis*-Chromosomen wurden genetisch kartiert (Copenhaver et al., 1999; diese Arbeit), um die cytologische und genetische Karte zu integrieren. Auch zeigte sich, daß die Rekombination im Bereich der Centromeren unterdrückt wird (Koornneef et al., 1983; Copenhaver et al., 1999; diese Arbeit). Da diese Regionen aus Heterochromatin bestehen, wirken sich Insertionen von repetitiven Elementen nicht so kritisch aus und führen dazu, daß diese Elemente dort akkumulieren (Copenhaver et al., 1999; diese Arbeit). Dabei ist noch nicht geklärt, ob die repetitiven Sequenzen eine wahrscheinlich wichtige Voraussetzung für die Centromerfunktion darstellen oder sekundär in dieser Region "gefangen" werden.

Es ist von entscheidender Bedeutung, alle Elemente, die das Centromer ausmachen und für seine Funktion notwendig sind, zu finden. Man könnte dann das Zusammenspiel von CEN-DNA und Kinetochor weiter und eingehender untersuchen. Dabei ist es essentiell, das Centromer und Pericentromer genetisch und v.a. physikalisch einzugrenzen.

Mit diesen Daten wäre man in der Lage, ein minimales funktionelles Centromer zu definieren. Zusammen mit Telomerstrukturen und Replikationsorigins wäre man in der Lage, künstliche Pflanzenchromosomen herzustellen (Plant Artificial Chromosomes, PLAC).

Die weitere Untersuchung von Pflanzen würde enorm vom Vorhandensein von PLACs profitieren. Für weitere genetische Untersuchung von *Arabidopsis* wären PLACs sehr hilfreich, um eine größere Zahl von Genen gleichzeitig zu transformieren und um Positionseffekte bei Zufallstransformationen zu umgehen; z.B. um mehrere manipulierte Gene oder die vielen Gene eines komplexen Merkmals zu übertragen. Man könnte PLAC-Bibliotheken von verschiedensten Pflanzenarten erstellen und sogar YAC-PLAC Hybrid-Vektoren sind vorstellbar, wobei die YAC- Bibliotheken von verschiedenen Pflanzen in eine geeignete Wirtspflanze transformiert werden (Somerville, 1999).

Eine weitere offene Frage ist das Vorhandensein von funktionellen Genen in Centromer-Regionen von *Arabidopsis*, ähnlich wie bei *S. cerevisiae* (Copenhaver et al., 1999). Findet man Gene in dieser Region, so stellt sich zwangsläufig die Frage nach der Regulation und Expression dieser Gene in einer heterochromatischen Region, da Heterochromatin die Eigenschaft hat Genexpression zu reprimieren. Man spricht hier von "position-effect variegation" (Karpen and Allshire, 1997).

Ziele dieser Arbeit waren die genetische und molekulare Charakterisierung der Centromer 1 Region von *Arabidopsis*, um Einsicht in den Aufbau und die Struktur vom Centromer zu erhalten. Es sollte eine sehr präzise genetische und physikalische Karte der Region und ein Centromer 1 überspannender contig bzw. die Voraussetzungen dafür erstellt werden. Das an das Centromer angrenzende Pericentromer sollte dabei näher bezüglich seiner Struktur, der Anwesenheit von exprimierten Genen und seinen genetischen Eigenschaften charakterisiert werden.

Weiterhin sollte über Integration von Sequenzdaten und FISH-Experimenten ermittelt werden, welche Regionen am Centromer 1 heterochromatisch sind. Hiermit sollte eine mögliche Beziehung von (Sequenz-)Struktur und Chromatinverpackung näher untersucht werden. Gleichzeitig sollte dieser Vergleich eine Grundlage bieten die Größe von CEN 1 (bzw. von *Arabidopsis* Centromeren überhaupt) abzuschätzen. Zusammengefasst sollten alle Analysen eine Definition der Centromerregion ermöglichen.

2 Material & Methoden

2.1 Chemikalien und allg. molekularbiologische Materialien

Die verwendeten Chemikalien wurden wenn nicht anders erwähnt von den Firmen Sigma (USA), Merck (Darmstadt), Biorad (USA), Difco (USA), Serva (Heidelberg), USB (USA), Roth (Karlsruhe), Pharmacia (Schweden), Fluka (Schweiz) und Boehringer (Mannheim) bezogen.

Alle verwendeten Radioisotope wurden von Amersham (Braunschweig) bezogen. Membranen der Firmen Amersham und Schleicher&Schöll (Dassel) wurden in den Blotting- und Hybridisierungstechniken eingesetzt.

Enzyme wurden von den Firmen MBI Fermentas (Litauen), Boehringer, New England Biolabs (USA), Gibco BRL (USA) und Stratagene (USA) geliefert.

Die beschriebenen Oligonukleotide wurden von den Firmen MWG (Ebersberg), Eurogentec (Belgien) und Gibco BRL (USA) synthetisiert.

2.2 spezielle Materialien für FISH-Analysen

Enzyme:

Pectolyase (<i>A. japonicus</i>)	Sigma (USA)
Cytohelicase	Sigma (USA)
Cellulase	Boehringer Mannheim

Proben labeling Reagenzien:

dig-Nicktranslation Mix	Boehringer Mannheim
bio-Nicktranslation Mix	Boehringer Mannheim

Fluorochrom Antikörper, Konjugate:

Texas Red Avidin D	Vector Laboratories (USA)
anti-Avidin D~Biotin	Vector Laboratories (USA)
Mouse-anti-Digoxin	Sigma (USA)
Rabbit-anti-Mouse~FITC	Sigma (USA)
Goat-anti-Rabbit~Alexa ⁴⁸⁸	Molecular Probes (USA)
DAPI	Boehringer Mannheim
Vectashield	Vector Laboratories (USA)

Fotomaterial:

Es wurden Kleinbild- und Diafilme der Firmen Kodak und Fuji verwendet. (ISO 64, 100, 400, 1600)

2.3 Pflanzen Material

Stämme:

Die verwendeten Ökotypen Nd, Ler und Col wurden von Dr. R. A. Torres Ruiz erhalten und sind die ursprünglich in der AG Prof. Dr. G. Jürgens gebräuchlichen Linien.

Allele:

Die gurke - Allele (R216, U413, U409, G301 (EMS induziert) und X-58 (Röntgen induziert)) stammen aus den in der AG Prof. Dr. G. Jürgens durchgeführten Mutagenese-"screens" (Mayer et al., 1991) und späteren von Dr. R. A. Torres Ruiz durchgeführten "screens" und wurden in Ler induziert. Das gurke - Allel emb22 (EMS induziert) wurde von D. Meinke zur Verfügung gestellt und ist in Col induziert.

Kartierpopulationen:

Für *gk* heterozygote Pflanzen wurden mit Niederzenz Pflanzen gekreuzt. Die für *GK* heterozygoten F1-Pflanzen wurden durch die Segregation von $GK^{-/-}$ -F2-Samen identifiziert und zur Aufzucht der F2-Populationen verwendet. Diese wurden zur Kartierung eingesetzt.

Mit dem *gk* Allel emb22 wurden 2 weitere Kartierpopulationen generiert. Im Gegensatz zu den bisherigen Populationen wurden diese mit visuellen Markern (*tt1* und *ch1*) vorselektioniert. Dies bedeutet, daß die rekombinanten F2 Pflanzen leicht an der gelben Samenfarbe der F2 Samen (*tt1*) bzw. an der gelblichen Farbe der F3 Keimlinge (*ch1*) erkennbar sind. Derart kann man sehr schnell und ohne viel Aufwand (im Gegensatz zu RFLP- und CAPS-Markern) einen Großteil der nichtrekombinanten Pflanzen der Kartierpopulation von der Kartierung mit näheren Markern ausschließen. Allerdings braucht man für jede Seite eine eigene Population.

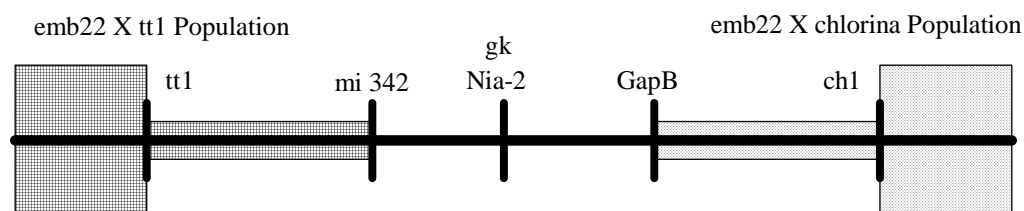


Abb. 1: Schematische Darstellung der Vorselektion großer Populationen.

2.4 Bakterienstämme

XL-1 Blue MRF	<i>E. coli</i> , Stratagene
SOLR	<i>E. coli</i> , Stratagene
DH5 α	<i>E. coli</i> , Sambrook et al. (1989)
C600	<i>E. coli</i> , Sambrook et al. (1989)

2.5 λ -Phagen

λ Zap II	Stratagene
ex assist Helfer Phage	Stratagene
λ gt 10	Sambrook et al. (1989)

2.6 Vektoren

pBluescript KS+	Stratagene
-----------------	------------

2.7 Klone

Die als Marker verwendeten Klone stammen, soweit nicht anders erwähnt, vom Arabidopsis Biological Resource Center, Ohio State University, Columbus, USA.

2.8 Bibliotheken

2.8.1 BAC-Bibliotheken

IGF BACs:

(Mozo, T., et al., 1998-1, -2)

Membranen und BAC-Klone der IGF BAC-Bibliothek wurden vom RZPD, (Ressourcen Zentrum Primär Datenbank). Ressourcenzentrum im Deutschen Humangenomprojekt am MPI für molekulare Genetik und am Deutschen Krebsforschungszentrum (<http://www.rzpd.de/>) bezogen.

TAMU BACs:

(Choi, S.D. et al., 1995-1, -2)

Membranen und BAC-Klone der TAMU BAC-Bibliothek wurden vom ABRC - AIMS (Arabidopsis Biological Resource Center, Arabidopsis Information Management System) Ohio State University, Columbus, USA, (<http://aims.cps.msu.edu/aims/>) bezogen.

2.8.2 cDNA-Bibliotheken

Gesamt Bank:

von	A. Bachmeier, MPI Köln
aus	A.t. cDNA, Ökotyp Columbia, junge blühende Pflanzen + Wurzel
Vektor	λ gt 10
Wirt	E. coli, C 600, hfl A
Phagentiter	$3,5 \cdot 10^8$ pfu/ml

Grün Bank:

von	vander Zaal, Leiden
aus	A.t. cDNA, Ökotyp Columbia, grüne Pflanzenteile, wenig Blüten + Wurzeln
Vektor	Uni Zap λ Phage (Stratagene), Eco RI, Xho II gerichtet kloniert (Adaptoren für cDNA´s Eco RI und Xho II)
Wirt	E. coli, XL1-Blue MRF (Stratagene)
Phagentiter	$3,2 \cdot 10^8$ pfu/ml

Blüten Bank:

von	D. Weigl, E. Meyerowitz (1991)
aus	A.t. RNA, Ökotyp Landsberg erecta, Infloreszenzen mit Blüten jünger als das 11.te Stadium
Vektor	λ Zap II (Uni Zap Kit, Stratagene), Eco RI, Xho II gerichtet kloniert (Adaptoren für cDNA´s Eco RI und Xho II)
Wirt	E. coli, XL1-Blue MRF (Stratagene)
Phagentiter	

2.9 Allgemeine molekularbiologische Methoden

Falls nicht anders erwähnt, wurden die allgemeinen molekularbiologischen Standardmethoden wie in Sambrook et al. (1989) und Gelvin & Schilperpoort (1994) beschrieben durchgeführt.

2.10 Gesamt RNA-Präparation aus Pflanzen

1-2 g Pflanzenmaterial (-70°C) werden in N₂ gemörsert. Das Pulver wird bevor es auftaut in einem Corex Röhrchen (50ml) mit 10 ml RNA Extraktions Puffer geschüttelt um eine klare Suspension zu erhalten. Suspension 10 min. bei 10000 rpm, 4°C (in JA 20, Beckman) zentrifugieren. Überstand in neues Corex Röhrchen, + 10 ml Phenol/Chloroform 1:1; gut schütteln, 10 min. bei 10000 rpm, 4°C zentrifugieren. Überstand in neues Corex Röhrchen, + 10 ml Chloroform; gut schütteln, 10 min. bei

10000 rpm, 4°C zentrifugieren. Überstand in neues Corex Röhrchen, + 1/10 Vol 3M NaAc pH 5,2 + 2 Vol EtOH, 1-2 h bei -20°C bzw. -70°C. 30 min. bei 10000 rpm, 4°C zentrifugieren. Pellet mit 20 ml 3M NaAc waschen, 10 min. bei 10000 rpm, RT zentrifugieren. Pellet mit 10 ml 3M NaAc waschen, 10 min. bei 10000 rpm, RT zentrifugieren. Pellet in 20 ml 70% EtOH resuspendieren und üN bei -20°C waschen. 10 min. bei 10000 rpm, RT zentrifugieren; Überstand abgießen, kurz anzentrifugieren restlichen Überstand abziehen, Pellet trocknen lassen. Pellet in 500 µl H₂O_{DEPC} aufnehmen.

Mengenabschätzung am Photometer:

50 µl RNA + 950 µl H₂O_{DEPC}, OD₂₆₀ und OD₂₈₀ bestimmen

OD₂₆₀ = 1 = 40 µg/ml RNA; OD₂₆₀/OD₂₈₀ >= 1,8 - 2,0 (gute - sehr gute RNA)

Ausbeute ca. 0,02 - 0,05% vom Ausgangsmaterial (zB. 500 µg aus 1 g)

Bemerkungen:

Material entweder frisch verwenden oder bei der Ernte sofort in flüssigem N₂ einfrieren und anschließend bei -70°C lagern. Fällung und anschließende Waschschrte kann auch in Falcons durchgeführt werden; JS 4.3 swing out bei max. 5000 rpm verwenden, länger zentrifugieren (Fällung auch 15 min. bei RT möglich). Guanidin HCl soll gewaschen werden; je länger desto besser!

Lagerung RNA: kurz: bei -70°C

länger: in 70% EtOH bei -70°C

alle Geräte und Lösungen müssen RNase frei sein

Reibschalen, Pistille und Corex Röhrchen werden mit Hypochlorid 1:10 einer 12% Lösung, 0,1 N NaOH/1 mM EDTA behandelt und gut mit H₂O_{DEPC} gespült. Plastik- und Glaswaren werden mit 0,1 N NaOH/1 mM EDTA behandelt und gut mit H₂O_{DEPC} gespült oder mit DEPC behandelt und autoklaviert.

RNA Extraktionspuffer:

8M Guanidin HCl, 20 mM EDTA, 20 mM MES (2-(N-Morpholino)ethansulfonic acid),

in H₂O_{DEPC} lösen, pH 7,0 einstellen, autoklavieren und bei RT lagern. Vor Gebrauch 50 mM 2-Mercaptoethanol dazupipettieren, bei 4°C lagern.

2.11 mRNA-Präparation (PolyA Tract® IV Promega)

0,1 - 1 mg gesamt RNA in 500 µl H₂O_{DEPC} 10 min. bei 65°C erhitzen + 3 µl biotinierte Oligo(dT) Probe + 13 µl 20XSSC_{DEPC}, vorsichtig umrühren und langsam auf RT abkühlen lassen (ca. 10 min.).

Waschen der Streptavidin-Paramagnetic Beads:

Beads vorsichtig resuspendieren (in original tube), auf Magnetständer fangen

Überstand vorsichtig abziehen. Beads 3 mal mit je 300 µl 0,5XSSC waschen, jedesmal auf Magnetständer fangen und Überstand ersetzen. Gewaschene Beads werden in 100 µl 0,1XSSC resuspendiert. Die Annealing Reaktion wird zu den Beads pipettiert.

Ansatz 10 min. bei RT inkubieren; alle 1-2 min. vorsichtig invertieren. Beads auf Magnetständer fangen, Überstand abziehen (Überstand nicht verwerfen, falls die markierte mRNA nicht ausreichend eluiert wurde). Beads 4 mal mit je 300 µl 0,1XSSC

waschen, dabei Beads vollständig durch Schnippen resuspendieren und auf Magnetständer fangen, Überstand ersetzen und nach dem letzten Waschschrift SSC Puffer möglichst vollständig entfernen. mRNA wird mit 100 μl $\text{H}_2\text{O}_{\text{DEPC}}$ eluiert; Beads resuspendieren und fangen; Überstand kommt in ein neues tube. Elution wird mit 150 μl $\text{H}_2\text{O}_{\text{DEPC}}$ wiederholt. Beide Eluate werden vereinigt; Σ 250 μl mRNAIsg.

Mengenabschätzung am Photometer:

25 μl RNA + 75 μl $\text{H}_2\text{O}_{\text{DEPC}}$, OD_{260} und OD_{280} bestimmen (Microküvette)

$\text{OD}_{260} = 1 = 40 \mu\text{g/ml RNA}$; $\text{OD}_{260}/\text{OD}_{280} \geq 1,8 - 2,0$ (gute - sehr gute RNA)

Ausbeute ca. 1 - 2% der Gesamt RNA (zB. 5 μg aus 500 μg)

Bemerkungen:

Lagerung mRNA:	kurz:	bei -70°C
	länger:	1/10 Vol 3M NaAc pH 5,2 + 2 Vol EtOH_{abs} bei -70°C

2.12 Formaldehyd Gelelektrophorese

Gelpräparation:

Für 100 ml 1,5% Gel werden 1,5 g Agarose in 85 ml $\text{H}_2\text{O}_{\text{DEPC}}$ + 10 ml 10X MOPS gelöst. Lösung auf 60°C abkühlen + 5,4 ml 37% Formaldehyd. Gel gießen, > 30 min. fest werden lassen. Gel mit 1X MOPS Puffer überschichten, Kamm entfernen, Taschen spülen.

RNA Präparation:

2 μg der gefällten mRNA 10 min. bei 14000 rpm. RT zentrifugieren. Überstand abziehen, Pellet in Speedvac ca. 5 min. trocknen. Pellet in 10 μl NBSB rücklösen, kurz zentrifugieren, auf Eis lagern. 2 μg RNA Längen-Marker in NBSB aufnehmen, (LTI 0,24 - 9,49 kb RNA ladder, Gibco). Proben 5 min. bei 65°C denaturieren. Sofort auf Eis stellen, kurz abzentrifugieren. RNA Längen-Marker separat auftragen da er vom Gel abgeschnitten wird. Gel bei 50 - 70 V laufen lassen bis BPB Front 10 cm erreicht hat. Längen-Marker wird vom restlichen Gel abgeschnitten; 5 - 10 min. in einer 0,1% Toluidin-Blau-Lösung gefärbt und anschließend mehrere Tage in 10 - 20 % EtOH -lösung entfärbt.

Bemerkungen:

alle Geräte und Lösungen müssen RNase frei sein

Elektrophorese Kammer, Gelschiffchen und Kamm werden mit Spüli gewaschen, mit Hypochlorid 1:10 einer 12% Lösung, 0,1 N NaOH/1 mM EDTA behandelt und gut mit $\text{H}_2\text{O}_{\text{DEPC}}$ gespült. Plastik- und Glaswaren werden mit 0,1 N NaOH/1 mM EDTA behandelt und gut mit $\text{H}_2\text{O}_{\text{DEPC}}$ gespült oder mit DEPC behandelt und autoklaviert. Alle Lösungen sollten mit DEPC behandelt und anschließend autoklaviert werden.

AUSNAHME: SDS und Denharts; diese werden mit $\text{H}_2\text{O}_{\text{DEPC}}$ gemacht.

10X MOPS Puffer:

200 mM MOPS pH 7,0, 50 mM NaAc, 10 mM EDTA, H₂O_{DEPC}

Puffer autoklavieren, vor Licht schützen, bei 4°C lagern.

NBSB:

50% Formamid_{deionisiert}, 1X MOPS Puffer, 7% Formaldehyd (aus 37%), 0,04% Bromphenol Blau. NBSB aliquotieren, bei -20°C lagern.

2.13 Northern Blot

Gele > 1,3%: Gel 10 -20 min. vorsichtig in 0,05 M NaOH in 1X SSC schütteln. Gel 2X je 20 min. in 20X SSC vorsichtig schütteln. Membran (+ Nylon Membran) je 5 min. in H₂O_{DEPC} und 10X SSC einweichen. RNA mit 20X SSC blotten (Aufbau und Dauer des Blots wie bei einem Southern Blot). Membran >30 min. bei 80°C backen. Filter 1-3 h bei 42°C prähybridisieren. Filter üN. bei 42°C mit gelabelter Probe hybridisieren. Filter 20 min. mit 2X SSC, 0,5% SDS bei RT waschen. Filter 20 min. mit 2X SSC, 0,1% SDS bei RT waschen. Filter 4 h auf "storage phosphor screen" auflegen und einscannen. Filter >= 65 h (2-3 Tage) auf "storage phosphor screen" auflegen und einscannen. Filter können mit heißer 0,1% SDS-Lösung gestrippt werden; bei schwachen Signalen kann man den Filter auch ohne Strippen rehybridisieren.

Bemerkungen:

Man kann mit "normalen Lösungen" und Geräten (nicht RNase frei) weiterarbeiten. Bei Problemen sollte man auf RNase freie Lösungen und Geräte zurückgreifen. Filter sollte man nicht stripfen da man dabei zuviel Template verliert; schwache Signale nehmen nach erneutem Prähybridisieren und Hybridisieren signifikant ab. RNA Filter kann in einer 0,1% Toluidin-Blau-Lösung angefärbt werden (RNA wird blau gefärbt, Filter muß entfärbt werden).

Hybridisierungs Lösung:

50% Formamid_{deionisiert}, 6X SSC, 5X Denharts, 0,5% SDS

Prähybridisierungs Lösung:

Hybridisierungs Lösung + 0,1 mg/ml denaturiertes Heringssperma

H₂O_{DEPC}:

0,1 % DEPC in H₂O lösen; mehrere Stunden - üN einwirken lassen (schwenken); autoklavieren.

2.14 Labeling einer DNA-Probe

DNA-Proben wurden nach einer modifizierten Methode von Feinberg und Vogelsang (1983, 1984) radioaktiv gelabelt.

50 - 100 ng DNA in 15 µl H₂O, 5 min. im WB aufkochen + 5 µl 5X Oligomix, sofort

auf Eis, + 5 μ l α dCTP* 50 μ Ci (32 P-dCTP (3000 Ci/mmol, 10 μ Ci/ μ l)), + 1 μ l Klenow; fingertip mischen; > 1h bei RT inkubieren.

Reinigung: (Hochsalzfällung)

25 μ l Label + 25 μ l 2X STOP Puffer + 100 μ l EtOH_{abs.}

15 min. bei RT inkubieren, 15 min. bei 14.000 rpm (16-18°C) zentrifugieren, Pellet in 100 μ l H₂O aufnehmen, 5 min. im WB aufkochen, sofort auf Eis stellen.

Lösungen:

2X STOP Puffer	4 M Ammoniumacetat, 10 mM EDTA, 100 ng/ml Heringssperma-DNA
5X Oligomix	1 Vol. Lösung A, 2,5 Vol. Lösung B, 1,5 Vol. Lösung C
Lösung A	1 ml Lösung O, 18 μ l β -Mercaptoethanol, je 15 μ l 100 mM dATP/dGTP/dTTP
Lösung B	2 M HEPES-NaOH pH 6,6
Lösung C	90 OD ₂₆₀ -Einheiten random Hexanukleotide (Pharmacia) in 1 ml 1X TE
Lösung O	1,27 g MgCl ₂ *6H ₂ O in 50 ml 1,25 M Tris-HCl pH 8,0

2.15 Hybridisierung geblotteter Filter

Die Filter wurden 3 - 4 h bei 65°C prähybridisiert (6XSSC, 5X Denhardt's, 0,5% SDS und 100 ng/ml denaturiertes Heringssperma) und anschließend üN. bei 65°C hybridisiert (6XSSC, 5X Denhardt's, 0,5% SDS und gelabelte Probe). Die Filter wurden 15 min. in 2X SSC, 0,5% SDS bei RT und 15 min. in 0,5X SSC, 0,1% SDS bei 65°C (BAC Filter) bzw. 15 min. in 0,2X SSC, 0,1% SDS bei 65°C (Multi Enzym Filter, BAC Fingerprint Filter, RFLP Kartier Filter) gewaschen. Die feuchten Filter wurden in Frischhaltefolie eingepackt und 2 h - üN. auf "storage phosphor screens" aufgelegt und mit einem Phosphoimager (Molecular Dynamics, USA) eingescannt. Die Filter wurden bei -20°C gelagert. Vor erneutem Hybridisieren wurden die Filter mit heißer, 0,1% SDS-lsg. 20 min. bei 65°C gestrippt um noch vorhandene Signale zu entfernen.

2.16 Fluorescent in situ hybridisation (FISH)

(Fransz et al., 1996; Zhong et al., 1996; Zhong et al., 1997)

2.16.1 Präparation meiotischer Chromosomen

Von den Pflanzen werden die Infloreszenzen abgesammelt, dabei werden nur Knospen ohne sichtbare Blütenblätter verwendet. Die Infloreszenzen werden 1h bei RT in Carnoy'slsg. fixiert. Diese können dann bei -20°C mehrere Monate gelagert werden oder sofort weiterverarbeitet werden. Die Infloreszenzen werden 3x mit H₂O_{dest} gewaschen.

Protoplastierung:

Die Infloreszenzen werden in 1ml 3x pektolytische Enzymmischung 2h bei 37°C inkubiert. Die Infloreszenzen werden mit H₂Odest gewaschen und auf Eis gestellt. Eine Knospe wird auf einem OT mit 1 Tropfen H₂Odest mit einer spitzen Nadel zerkleinert.

Klärung:

Ein Tropfen 60% HAc (von gesamt 30 µl HAc) wird zu der Zellsuspension pipettiert, der OT wird auf einem 45°C warmen Heizblock ca. 30-60 sek. inkubiert.

Unter Rühren wird der Rest der HAc dazupipettiert. Der OT wird 2x mit eiskalter Carnoy'slsg. gewaschen; evtl. kurz mit EtOH gespült. Der OT wird mit Papier getrocknet. Das Präparat wird unter einem Phasenkontrastmikroskop auf entsprechende Meiosestadien (Pachytänstadien) untersucht. Das Präparat kann auch DAPI gefärbt werden. Das Präparat kann bei -20°C mehrere Monate lang gelagert werden (bei 4°C, RT mehrere Wochen lang).

2.16.2 Vorbehandlung der Präparate

Ausgewählte OT mind. 30 min. bei 60°C backen. OT 5 min. in 2x SSC waschen (kann ausgelassen werden). 100 µl RNase A (100 µg/ml in 2xSSC) werden auf jeden OT pipettiert; OT mit einem DG (24x50 mm) abdecken und 1 h bei 37°C inkubieren. OT 3x je 5 min. in 2x SSC waschen.

Proteasenbehandlung mit Pepsin (optional)

OT werden 2 min. in 0,01 M HCl inkubiert, OT trocknen (Unterseite), 100 µl Pepsin (5 µg/ml = 20 U/ml in 0,01 M HCl) werden auf jeden OT pipettiert; OT mit einem DG (24x50 mm) abdecken und maximal 5 min. bei 37°C inkubieren. OT 5 min. in 1x PBS waschen; evtl. DG entfernen.

OT 10 min. in 1% Formaldehyd in 1x PBS inkubieren, OT 5 min. in 1x PBS und 2x 5 min. in 2x SSC waschen.

OT werden in EtOH dehydriert; je 2 min. in 70%, 90%, 100% EtOH waschen; an der Luft trocknen lassen. Die Vorbehandlung der Präparate kann unter einem Phasenkontrastmikroskop untersucht werden.

2.16.3 Nick-Translation

1 µg Proben DNA in 16 µl sterilem H₂Odest. aufnehmen. 4 µl 5x Nick Translationsmix (Biotin bzw. Digoxin) dazupipettieren, vortexen und kurz abzentrifugieren. 90 min. bei 15°C inkubieren, die Reaktion wird mit 1 µl 0,5 M EDTA 10 min. bei 65°C gestoppt; kurz abzentrifugieren.

2.16.4 In situ-Hybridisierung und Waschschrte

Hybridisierungsmix:

S 20µl für einen OT

10 µl DS, 10 µl gelabelte Probe in HB50 (evtl. Kompetitor DNA), 10 µl DS + x1 µl

Probe 1 + x2 µl Probe 2 + (10-x1-x2) µl HB50, vortexen.

Denaturierung:

Probe und Target (OT) werden getrennt denaturiert.

Hybridisierungsmix wird 5 min. aufgekocht und sofort auf Eis gestellt, vortexen. 20 µl HB70 werden auf vorbehandelte OT pipettiert; OT mit DG zudecken 2 min. bei 75°C auf Heizplatte inkubieren, sofort eiskalt stellen. 2 min. in eiskaltem 70% EtOH waschen, je 2 min. in 90%, 100% EtOH waschen. OT an der Luft trocknen. Hybridisierungsmix wird auf den entsprechenden OT pipettiert, OT wird mit einem DG zugedeckt. OTs werden in einen Färbetrog reingeschlichtet (DG liegt oben), dieser wird mit Parafilm abgedichtet und in einer vorgewärmten feuchten Kammer bei 37°C üN inkubiert. OTs werden 3x je 5 min. in 50% Formamid / 2x SSC bei 42°C inkubiert. OTs kurz in 2x SSC waschen. OTs je 3x 5 min. in 0,1x SSC bei 60°C inkubieren (optional). OTs 5 min. in 2x SSC waschen.

2.16.5 Detektion und Amplifikation

OTs 5 min. in 4T waschen. 100 µl 4M werden auf jeden OT pipettiert, OTs mit DG (24x50 mm) zudecken, 30 min. bei 37°C inkubieren. OTs kurz in 4T waschen.

100 µl 2 µg/ml Avidin~Texas Red in 4M auf OT (1:1000 Verdünnung der Stocklsg.) + DG 30 min. bei 37°C inkubieren. 3x 5 min. in 4T waschen, 5 min. in TNT waschen.

100 µl 5 µg/ml Goat-anti-Avidin~Biotin in TNB (1:100 Verdünnung der Stocklsg.) und Mouse-anti-Diogxin in TNB auf OT (1:1000 Verdünnung der Stocklsg.) + DG 30 min. bei 37°C inkubieren. 3x 5 min. in TNT waschen.

100 µl 2 µg/ml Avidin~Texas Red in TNB (1:1000 Verdünnung der Stocklsg.) und Rabbit-anti-Mouse~FITC in TNB auf OT (1:1000 Verdünnung der Stocklsg.) + DG 30 min. bei 37°C inkubieren. 3x 5 min. in TNT waschen.

100 µl 10 µg/ml Goat-anti-Rabbit~Alexa in TNB auf OT (1:200 Verdünnung der Stocklsg.) + DG, 30 min. bei 37°C inkubieren. 3x 5 min. in TNT waschen.

OT werden in EtOH dehydriert; je 2 min. in 70%, 90%, 100% EtOH waschen; an der Luft trocknen lassen. 8-12 µl einer 5 µg/ml DAPI in Vectashield Antifadelsg. auf OT pipettieren, mit DG (24x32 mm bzw. 24x50 mm) abdecken. Ergebnisse unter einem Fluoreszenzmikroskop (Axiophot, Zeiss) visualisieren.

2.16.6 spezielle Lösungen und Puffer

10X Pectolytische Enzymlösung 1% Cellulase, 1% Pectolyase, 1% Cytohelicase,
in 100 mM Citratpuffer pH 4,5

Carnoy'slsg. EtOHabs. : Eisessig (3 : 1)

10 X PBS pH 7,0 100 mM NaPO₄ pH 7,0, 1,4 M NaCl

Lsg. a) 900 ml H₂O_{dest.} 16,02 g Na₂HPO₄*2H₂O + 73,84 g NaCl

Lsg. b) 200 ml H₂O_{dest.} 2,76 g NaH₂PO₄*H₂O + 16,56 g NaCl

Lsg. a mit Lsg. b auf pH 7,0 titrieren; autoklavieren.

Pepsin Stock 500 µg/ml in 0,01 M HCl

RNase A 100 µg/ml in 2XSSC

1% FA in PBS 87 ml H₂O_{dest.} + 10 ml 10XPBS + 3 ml 37% FA

HB50	50% Formamid, 2XSSC, 50 mM NaPhosphatpuffer pH 7,0
HB70	70% Formamid, 2XSSC, 50 mM NaPhosphatpuffer pH 7,0
20% Sodium Dextran in HB50	
DAPI in Vectashield	5 µg/ml DAPI in Vectashield lösen (5 µg/µl Stock)
Formamid Wash Puffer	50% Formamid, 2XSSC
10XTN pH 7,5	1 M Tris-HCl, 1,5 M, ad 1000 ml H ₂ O _{dest.} ; autoklavieren.
TNT Puffer	TN Puffer + 0,05% Tween 20
TNB Puffer	TN Puffer + 0,5% blocking reagent, bei -20°C lagern
4T Puffer	4XSSC + 0,05% Tween 20
4M Puffer	4XSSC + 5% Magermilchtrockenpulver

2.17 DNA Präparationen

2.17.1 Pflanzen

Genomische Pflanzen DNA wurde mit einer abgeänderten Methode nach Dellaporta et al. (1983) isoliert.

1g Pflanzenmaterial (Blätter) wird in einem Mörser mit flüssigem N₂ zu feinem Pulver zerrieben. Das Pulver wird in 50 ml Corex Röhrchen überführt, die je 7 ml EB Puffer (50 mM Tris, pH 8,0, 10 mM EDTA, pH 8,0, 100 mM NaCl, 1% SDS, 10 mM β-Mercaptoethanol) enthalten und gut geschüttelt. Die Suspension wird mind. 5 min. bei 65°C inkubiert. Dazu werden 2,5 ml 5 M KAc pipettiert, gut geschüttelt und mind. 15 min. auf Eis gestellt. Bei 8000 g (7800 rpm, JA 20, J2-H6, Beckman) und 4°C 10 min. zentrifugieren. Der Überstand wird über einen Verbandmull Filter in ein neues Corex Röhrchen, das 5 ml Isopropanol enthält, pipettiert und vorsichtig invertiert. Es wird bei 8000 g und 4°C 15 min. zentrifugiert. Das Pellet wird mit 2,5 ml 70% EtOH gewaschen und in 500 µl 1X TE rückgelöst. Nach dem Rücklösen kann die DNA zusätzlich mit RNase (10 µg) verdaut (30 min. bei 37°C). Anschließend wird die DNA phenolisiert (Phenol, C:IA), hochsalzgefällt und in 200 µl 1XTE rückgelöst.

Arabidopsis-DNA wurde für die PCR-Analysen nach einer modifizierten CTAB-Methode von Roger und Bendich (Rogers & Bendich, 1994) isoliert.

Ein bis zwei Blätter werden mit einer zugeschmolzenen blauen Pipettenspitze in einem Eppendorf-Reaktionsgefäß (1,5 ml) homogenisiert. Es werden 200 µl 2X CTAB-Puffer zugegeben und 15 sec. gevortext. Die Lösung wird für mindestens 30 min. bei 65°C inkubiert. Das CTAB wird mit 200 µl Chloroform extrahiert. Die Phasen werden durch eine Zentrifugation (14000 U/min., 15 min., RT) getrennt und die DNA wird aus 200 µl Überstand mit 140 µl Isopropanol durch eine erneute Zentrifugation gefällt. Das Pellet wird mit 70% EtOH_{abs} gewaschen und in 50 µl 1X TE (pH 8,0) gelöst. Für eine PCR-Reaktion wurden 2 µl der Lösung eingesetzt.

Teilweise wurde Arabidopsis-DNA für PCR-Analysen auch mit dem DNeasy Plant Mini Kit von Quiagen nach Vorschrift des Herstellers isoliert.

2.17.2 Plasmide

Plasmid-DNA aus *E. coli* wurde durch alkalische Lyse (Birnboim und Doly, 1979) nach einem modifiziertem Verfahren gewonnen.

Die Zellen einer Übernachtskultur (1,5 bis 3 ml) werden abzentrifugiert und in 400 µl P1/S1 Puffer (50 mM Tris-HCl, 10 mM EDTA, pH 8,0, 100 µg RNase) resuspendiert. Es werden 400 µl P2/S2 Puffer (0,2N NaOH, 1% SDS) zugegeben, invertiert und für max. 5 min. inkubiert. Proteine und genomische DNA werden durch Zugabe von eiskaltem P3/S3 (3 M KAc (pH 5,5)) gefällt. Die Lösung wird nach mind. 10 min. Inkubation auf Eis 15 min. bei 14000 rpm und 4°C zentrifugiert. Die Plasmid-DNA wird mit 0,7 Vol. Isopropanol aus dem Überstand gefällt und abzentrifugiert. Das Pellet wird mit 70% EtOH. gewaschen und in 50 µl 1X TE gelöst.

Wurde hochreine Plasmid-DNA benötigt so wurde die DNA über Ionenaustauscher Säulchen Tip 20, Quiagen und Nucleobond AX 20, Machery-Nagel (MN) gereinigt. Dazu wurde der Überstand nach der KAc Fällung auf die Säulchen aufgetragen und nach Vorschrift des Herstellers weitergearbeitet. Puffer P1-3, Quiagen; Puffer S1-3, Machery-Nagel

2.17.3 BACs

BAC-DNA wurde auch durch alkalische Lyse (Birnboim und Doly, 1979) gewonnen. Da BACs sich von low copy F-Plasmiden ableiten erfordert die BAC-DNA Präparation große Kulturvolumina (50 ml für Mini- Präparationen, 100-200 ml für Midi-Präparationen und 500 ml für Maxi- Präparationen).

Abweichend von der Plasmid-DNA Präparation wurden bei Mini-BAC-DNA Präparationen jeweils 800 µl der Puffer S 1-3 eingesetzt. Die BAC-DNA wurde entweder über Phenolisierung oder über Nucleobond AX 20 Säulchen (MN) nach Vorschrift des Herstellers gereinigt.

Bei Midi- und Maxi-BAC-DNA Präparationen wurden jeweils 8 ml der Puffer S 1-3 eingesetzt. Die BAC-DNA wurde entweder über Phenolisierung oder über Nucleobond AX 100 Säulchen (MN) nach Vorschrift des Herstellers gereinigt.

2.18 PCR

Alle PCR Reaktionen wurden in einem Peltier Thermoblock PTC MJ 225 von MJ Research, USA (Biozym) in dünnwandigen 0,2 ml Streifen ohne Paraffinöl durchgeführt. Das Standardreaktionsvolumen betrug 20 µl bestehend aus:

200µM dNTP Mix (je 200 µM dATP, dCTP, dGTP, dTTP), 1X Taq Puffer (Boehringer Mannheim), je 1 µM forward und reverse Primer, 1 U Taq Polymerase (Boehringer Mannheim), 50 ng DNA und H₂O_{Milipore}. Zur Optimierung der einzelnen PCR Reaktionen wurde ein Standardprogramm (2 min. 94°C, 40x(30 sek. 94°C, 30 sek. T°C, 1 min. 72°C), 5 min. 72°C) mit 3 annealing Temperaturen (T = 54°C, 58°C und 62°C) verwendet.

Name	Sequenz 5' - 3'	Name	Sequenz 5' - 3'
22a12-sp6-uni	gga tcg acc cag agt tca agt	22a12-sp6-rev	gac cat aca atc aca tcc atc tc
17a20-sp6-uni	cag tagctttcaaggcagcaaatg	17a20-sp6-rev	att ttc cag gga gtt cgg act c
3l22-sp6-uni	agg gag cta ttt ggt ttc t	3l22-sp6-rev	gat gcg tcc ttg ctt cac tt
11o6-sp6-uni	ctt tgg act tgt tctcctggacag	11o6-sp6-rev	gaa agt gtt cct ctg att gaagcc
11o6-t7-uni	gca agg gtgttgatgcacaactg	11o6-t7-rev	cag acg tca act caa ccttctgtc
4c4-t7-uni	gtg agctggaagtacaaaagattg	4c4-t7-rev	cga gga ctt ggt tct tct tga ag
3n5-t7-uni	cac cag ttt gtg gtt caa g	3n5-t7-rev	ccg gat aag tat gga tgt g
21b18-t7-uni	cga gga ctt ggt tcttctga ag	21b18-t7-rev	caa aaa ctg tcc acc tct ccc
t3h3-t7-f	agc agc tta cgt ccaaccgaag	t3h3-t7-r	gtg gat cat cgc ttg atg caa c
t2m19-sp6-f1	taa agt taa aatatagttcaga	t2m10-sp6-r	cca ttt gtt ctt gca tcc caa g
t2m19-sp6-f2	caa aat ctgatgagtccttgacc		
t4n16-t7-f	gac ttt gcg ggcttaaagctac	t4n16-t7-r	tta cct cct aca tca cgtctgtcc
t4n16-sp6-f	caagctacctgatgatgagaagtc	t4n16-sp6-r	cat gag agc gtt tgg att tggtag
t2l16-t7-f	cag ctt att acg atg gca cg	t2l16-t7-r	att ggc aag cga ctg ctt tg
t18h22-sp6-f	gag aag ccattagaaagttagcg	t18h22-sp6-r	tac ttt agg cta gcg aga aagac
t5b13-t7-f	gga atg ata gga agt ttc cac	t5b13-t7-r	cta ctc gcg ggt ttg aat gtg
t5b13-sp6-f	cgt aat gtt aaa cgaccggaac	t5b13-sp6-r	ggc caa tat gcg gat ttg att c
t10h24-sp6-f	gaa atg gat tacactgttctc	t10h24-sp6-r	ggg gaa agt gaa tcc aaa tatgg
t3l122-t7-f	atg gct gag act atc tgt atac	t3l122-t7-r	aaa ggc cta acc ata atg gct c
t6f19-sp6-f	gcc cca ttt ttc ttg gag tc	t6f19-sp6-r	ttg tcg tcc aga gca acc ttg
t3l13-t7-f	ttg tgt cgt gtc ttg ttg tgg	t3l13-t7-r	ttg gta tgg ttc cac cat ctc
t16m15-sp6-f	tgg gga cat gaa gtt gtt cac	t6m15-sp6-r	act gag ttt tgacaaattcactag
1c13-t7-f	gta tag aag gag gaa tag cc	1c13-t7-r	ttc gta aac gct caa acg g
11h6-t7-f	tca acc aaa tgt ttcagctccc	11h6-t7-r	gtg aactcaacgggtgcagacg
1n23-t7-f	gac agc ctt ggaactgtgtgg	1n23-t7-r	ttc gtg gtc ttc agg gct tag g
28b8-t7-f	tta tga aga gaaggtcatgccccg	28b8-t7-r	ttg ttg ggg ttc ttc aac tgt cg
6j23-t7-f	cat ttc ata ggg agc tat ttgg	6j23-t7-r	tcc aag ttg ttc ctc ggc tataag
4a3-t7-f	tgg aca aag gaatgaatgagcc	4a3-t7-r	atc aga cat gac aga agc tgc
9d18-sp6-f	ttc agc ggtacagagctgccattg	9d18-sp6-r	att tct ggt tat cct gactgcacc
9d18-t7-f	gct gta tct acccaataaacccc	9d18-t7-r	ttg tgg aag ggg aga tgt atc
9d18-orf1-f	gaccagcagatatgatgatcgac	9d18-orf1-r	ttg tgt ctt age ttc ccc ctg g
9d18-hyp prot 1-f	tct acc acg cat cag aat gct tc	9d18-hyp prot 1-r	gac ccc atc aag ccc ttg taa g
9d18-inte grase-f	gac att gtt ctg gtt tcc cca c	9d18-inte grase-r	agg tca aga gac aga gat ggt ccc

Name	Sequenz 5' - 3'	Name	Sequenz 5' - 3'
9d18-rev trans-f	cgc agt gaa ctg gag ata ccg aag	9d18-rev trans-r	tct ggc aat gaa ggt gaa gac ac
jcc3-1	ggc tac tgg tca aat cat tc	jcc3-2	gaa tct ttg caa acg agt gg
GapB 5-2	tgt tag agg aga gacagtggcg	GapB 3-2	ggt aga gac atc aga gca cctg
Nia2 5-2	cgg ttt aat agt tac cga tca	Nia2 3-2	ggc cgc cat gtt gga aag tg
a2-4-uni-1	gac cag tcg gtc aat atg tg	a2-4-rev-1	gtc taa act tcg cct cga tag
a2-4-uni-2	cga tcc gat ctt gaa caa ac	a2-4-rev-2	cga tct ctg gaa tct cac tct ac
a2-4-uni-3	tct tga cct acc cca aca c	a2-4-rev-3	cga gag tcc aga act tga cc
a2-4-uni-4	agg aca gct ttc atg gcg tg	a2-4-rev-4-uni	agc ttt gtg aag ggc atc
a2-4-uni-5	gat gcg ttt cgg gga gag gtg	a2-4-rev-5	cgc cgc aat tgt ggg tca gc
as31-uni-1	gaa gaa tgg aga ttg ctt g	as31-rev-1	gaa aac tcg ggc ttc atc
as31-uni-2	cac ctt cag gag cag aat g	as31-rev-2	gct tct tca agt gag aga tcc
as31-uni-3	ccc ttt att ccc ttg ctg	as31-rev-3	tgt tgc cac tca agt tgc cc
gs31-uni-1	ctc aaa ttg ctg cac ctg	gs31-rev-1	tct ccc ttg gtc act cga tc
gs31-uni-2	gcc ttt gtg att tgg gaa cat c	gs31-rev-2	tcg tct ctt gag ata tgg tag g
gs31-uni-3	gtc ata gat gtc agg aac gg	gs31-rev-3	tgc atg gtt cag tcc ttc

Tab. 1: Oligonukleotidsequenzen
(Sequenzieren, CAPS, BAC Enden für HAMDE und Hybridisierungsproben)

PCR Reaktion	Annealing Temp. (°C)	PCR Reaktion	Annealing Temp. (°C)
22a12-sp6	58	t5b13-sp6	62
17a20-sp6	62	t10h24-sp6	62
3l22-sp6	54	t31i22-t7	62
11o6-sp6	62	t6f19-sp6	54
11o6-t7	62	t31i13-t7	54
4c4-t7	62	t16m15-sp6	54
3n5-t7	54	1c13-t7	58
21b18-t7	62	11h6-t7	62
t3h4-t7	62	1n23-t7	62
t2m19-sp6	62	28b8-t7	62
t4n16-t7	62	6j23-t7	58
t4n16-sp6	62	4a3-t7	62
t21c16-t7	54	jcc3-1	58
t18h22-sp6	58	GapB 5-2	56
t5b13-t7-f	54	Nia2 5-2	54

Tab. 2: Annealingtemperaturen für PCR Reaktionen

2.19 Kartierungen

2.19.1 morphologische Marker

<i>tt1</i>	<i>transparent testa</i>	reife Samen sind gelb, da die Samenschale farblos ist
<i>chl</i>	<i>chlorina</i>	Keimlinge und junge Pflänzchen sind gelb - hellgrün

2.19.2 RFLP

RFLPs erlauben es F2 Populationen mit ökotyp-spezifischen Markern zu kartieren. Wegen eines Restriktionspolymorphismus treten polymorphe Banden auf. Die Ursprungsökotypen können anhand dieser polymorphen Banden durch eine Southernanalyse bestimmt werden.

Locus	Chromosom; Position (cM auf RI Karte)	Restriktionsenzym; Ökotyp, polymorphe Banden (Fragmentgröße kbp)	Referenz
mi133	1, 59	Hae III;	Liu et al. (1996)
mi342	1, 56,6	Hind III:	Liu et al. (1996)
17A20-sp6 ¹	1, 56,6 - 59	Dra I: Col, 2 (7 kb , 2kb), Ler, 2 (1,6 kb , 1,5 kb), Nd, 2 (6,8 kb, 1,7 kb)	

¹ Der Marker 17A20-sp6 erkennt wahrscheinlich 2 Loci; man kann dennoch eindeutig kartieren.

Tab. 3: RFLP Marker

2.19.3 CAPS

CAPS ist eine Methode Arabidopsis Mutanten mit co-dominanten ökotyp-spezifischen PCR Markern zu kartieren (Konieczny und Ausubel, 1993). Anhand bekannter DNA Sequenzen und den entsprechenden durch PCR amplifizierten DNA Fragmenten werden Marker etabliert. Wegen eines Restriktionspolymorphismus können die Ursprungsökotypen dieser Fragmente durch einen Restriktionsverdau mit Restriktionsendonucleasen bestimmt werden.

Bei dem Restriktionsverdau wurden 20 µl PCR Produkt in einem Volumen von 30 µl mit 5 U Restriktionsenzym, 1X NEB Restriktionspuffer und H₂O_{Milipore} angesetzt und 3 h bei der dem Restriktionsenzym entsprechenden Temperatur in einem PTC MJ 225 inkubiert. Dabei wurden die PCR Streifen weiterverwendet.

Die Restriktionsfragmente wurden auf einem 1 - 1,5 % Agarosegel bei 70V aufgetrennt und unter UV Licht (254 nm) sichtbar gemacht und mit einer Videokamera mit Videoprinter oder einem Computer (als Graphik) dokumentiert.

Locus	Chromosom; Position (cM auf RI Karte)	Fragment Größe (kb)	Restriktionsenzym; Ökotyp, Schnittstellen (Fragmentgröße kbp)	Referenz
GapB	1, 59,1	1,3	Sac I, Ler, 0, Nd, 1	Fischer (1997)
jcc3	1, 59,1	1,6	Nla III, Ler=Nd, 2, Col, 3	Chin, J.
Nia-2	1, 56,6 - 59	0,7	Hae III, Ler, 1, Nd, 0	Fischer (1997)

Tab. 4: CAPS Marker

2.19.4 HAMDE

HAMDE (Hauser et al., 1998) ist eine schnelle Methode co-dominante, polymorphe PCR-Marker zu etablieren. Dabei wird eine Heteroduplexanalyse mit einer hochauflösenden Gelmatrix kombiniert. So kann man einzelne Nukleotid Polymorphismen aufdecken und als Kartier-Marker verwenden. Für diese Methode können PCR Produkte einer Länge von 200 bp - 900 bp verwendet werden. Die entsprechenden Oligonukleotide wurden aus BAC Endsequenzen herausgesucht.

Um Polymorphismen zu finden wurden die amplifizierten Ler, Nd und Col PCR Fragmente und deren künstliche Heterozygoten (1:1 Mischung jeweils zweier Ökotypen) einer Heteroduplexanalyse unterzogen. Dabei werden die doppelsträngigen DNA Fragmente denaturiert und langsam reannealt. Die Homo- und Heteroduplexe wurden auf einem hochauflösenden (MDE) Gel aufgetrennt. Das Gel wurde anschließend in EtBr gefärbt um die DNA Banden zu visualisieren.

2/3 der getesteten PCR Fragmente konnten in der vorliegenden Arbeit als polymorphe Marker verwendet werden. Mit diesen wurden die F2 Populationen kartiert.

Für jedes Oligonucleotidpaar (potentieller Marker) werden zuerst die optimalen PCR Bedingungen ermittelt und DNA Fragmente mit den entsprechenden Ökotypen DNAs amplifiziert. Die PCR Produkte werden 7 min. in einem Thermoblock bei 95°C denaturiert und langsam (Thermoblock ausschalten) auf RT (mind. 30 min.) abgekühlt. 2 - 4 µl der Probe werden mit 1 µl Auftragspuffer (0,25 % Bromphenolblau, 0,25 % Xylencyanol, 66 % Succrose) gemischt und auf das MDE Gel aufgetragen.

Die hochauflösende Matrix ist ein natives 14% Acrylamid Gel mit 0,15% crosslinking (Bisacrylamid). Um die Auflösung evtl. zu verbessern kann man das Gel mit 2,5 M Harnstoff ansetzen. Es wurden 1 mm dicke 20 cm Gele verwendet, dafür wurden 40 ml Acrylamis-Lsg. benötigt. 20 ml 2X Stock (28 % : 0,3 % , Acrylamid : Bisacrylamid) + 2,4 ml 10XTBE + 17,6 ml H₂O + 125 µl 10% APS + 50 µl TEMED, 2-5 min. rühren. Das Gel wird sofort gegossen und ist nach ca. 1 h auspolymerisiert. Die benötigten Glasplatten müssen fettfrei sein (gut mit Spülmittel waschen und mit EtOH abwischen). Die Elektrophorese wird in 0,6X TBE bei 250 V durchgeführt bis der Blaumarker das untere Gelende erreicht hat, mind. aber 3 h. Bei größeren Fragmenten kann die Elektrophorese länger gefahren werden um die Auftrennung der einzelnen Banden zu verbessern. Das Gel wird 15 min. in einem EtBr Bad geschwenkt. Die Homo- und Heteroduplexe werden unter UV Licht (254 nm) sichtbar gemacht und mit einer Videokamera mit Videoprinter oder einem Computer (als Graphik) dokumentiert.

Locus	Chromosom; Position (cM auf RI Karte)	Polymorphismen <--> codominant ---> dominant
1c13-t7	1, 56,6 <- 59	Ler <--> Nd, Ler <--> Col
4a3-t7	1, 56,6 <- 59	Ler <--> Nd, Ler <--> Col
11h6-t7	1, 56,6 <- 59	Ler <--> Nd, Ler <--> Col
t21c16-t7	1, 56,6 - 59	Ler <--> Nd
t31i22-t7	1, 56,6 - 59	Ler <--> Nd
t18h22-sp6	1, 56,6 - 59	Ler ---> Col
6j23-t7	1, 56,6 -> 59	Ler ---> Nd, Ler <--> Col
3l22-sp6	1, 56,6 -> 59	Ler <--> Nd, Ler <--> Col

Tab. 5: HAMDE Marker

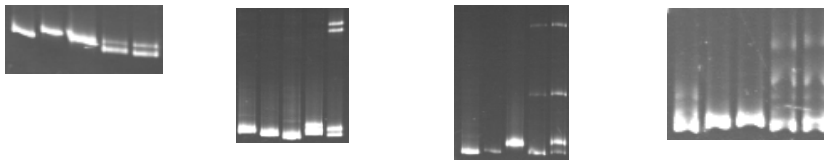


Abb. 2: Polymorphismen der Loci 1c13-t7, 4a3-t7, 6j23-t7, 3l22-sp6; die 5 Banden entsprechen jeweils Ler, Nd, Col, LerNd, LerCol





2.19.5 Berechnung der genetischen Abstände

Die aus der Selbstung der F1 Pflanzen hervorgehenden 16 möglichen Genotypen (Klassen) ergeben sich aus der Kombination der 4 Gameten. Zur Kartierung wurden in diesem Fall die Klassen mit einem eindeutigen Marker / Phänotypen ausgewertet. Dabei muß man berücksichtigen, daß *gk/gk* letal ist.

morphologische Kartierung:

		parental		rekombinant		
		(1-p)/2	(1-p)/2	p/2	p/2	
Gameten		+ / gk	+ / m	gk / m	+ / +	
parental	(1-p)/2	+ / gk	+ / + gk / gk	+ / m + / gk	+ / m gk / gk	+ / + gk / +
	(1-p)/2	+ / m	+ / m + / gk	m / m + / +	m / m + / gk	+ / m + / +
rekombinant	p/2	gk / m	+ / m gk / gk	m / m + / gk	m / m gk / gk	+ / m + / gk
	p/2	+ / +	+ / + gk / +	+ / m + / +	+ / m + / gk	+ / + + / +

Tab. 6: p = Anteil rekombinanter F1 Gameten, 1-p = Anteil parentaler F1 Gameten

	homozygot <i>gurke</i> Pflanzen		Pflanzen mit parentalem Genotyp
	rekombinante Pflanzen		rekombinante Pflanzen; nicht auswertbar

mathematische Formel zur Berechnung des genetischen Abstandes p:





$$RF = \frac{Rek}{Rek + Par} ; RF = X = \frac{Z}{N} ; X = \frac{2 * \left(\frac{p}{2} * \frac{1-p}{2} \right)}{2 * \left(\frac{p}{2} * \frac{1-p}{2} \right) + \left(\frac{1-p}{2} \right)^2} ; p = \frac{X}{2-X}$$

RF = Rekombinationsfrequenz; Rek = rekombinante Klassen; Par = parentale Klassen; Z = Zahl der beobachteten Rekombinanten; N = Zahl der auswertbaren Pflanzen

molekulare Kartierung:

			parental		rekombinant	
			(1-p)/2	(1-p)/2	p/2	p/2
Genotyp			Ler / gk	Nd / +	Nd / gk	Ler / +
parental	(1-p)/2	Ler / gk	Ler / Ler gk / gk	Ler / Nd gk / +	Ler / Nd gk / gk	Ler / Ler gk / +
	(1-p)/2	Nd / +	Ler / Nd gk / +	Nd / Nd + / +	Nd / Nd + / gk	Nd / Ler + / +
rekombinant	p/2	Nd / gk	Ler / Nd gk / gk	Nd / Nd + / gk	Nd / Nd gk / gk	Ler / Nd + / gk
	p/2	Ler / +	Ler / Ler gk / +	Nd / Ler + / +	Ler / Nd + / gk	Ler / Ler + / +

Tab. 7: p = Anteil rekombinanter F1 Gameten, 1-p = Anteil parentaler F1 Gameten

	homozygot <i>gurke</i> Pflanzen		Pflanzen mit parentalem Genotyp
	rekombinante Pflanzen		rekombinante Pflanzen; nicht auswertbar

mathematische Formel zur Berechnung des genetischen Abstandes p bei einer Letalkartierung:

$$RF = \frac{Rek}{\frac{3}{4}} ; RF = X = \frac{Z}{N} ; X = \frac{4}{3} * \left(6 * \left(\frac{1-p}{2} * \frac{p}{2} \right) + \left(\frac{p}{2} \right)^2 \right) ; p_{1/2} = \frac{3}{5} \pm \sqrt{\frac{9}{25} - \frac{3}{5}X}$$

RF = Rekombinationsfrequenz; Rek = rekombinante Klassen; Par = parentale Klassen; Z = Zahl der beobachteten Rekombinanten; N = Zahl der auswertbaren Pflanzen

Haldane Korrektur:

Für große Abstände auf der genetischen Karte müssen diese korrigiert werden.

Dazu kann man sich folgender Formel bedienen: $p = \frac{1}{2} * (1 - e^{-m})$ (Griffith et al., 1993)
 m entspricht dabei der Zahl der Crossovers. $m = -\ln*(1 - 2p)$
 m = 1 entspricht 50 korrigierten "map units" $p_{korrigiert} = \frac{1}{2}m$; $p \leq 0,5$

2.20 cDNA screens

Die Phagen wurden in einer Dichte von ca. 10000 pfu/Platte ausplattiert und solange bei 37°C inkubiert bis die Plaques einen Durchmesser von ca. 1 mm erreichten. Von jeder Platte wurden 2 Filter (Nitrocellulose, Ø 8,5 cm, Schleicher & Schüll) abgezogen, der erste wurde 2 min. und der zweite 4 min. aufgelegt. Die Filter wurden 2 min. denaturiert (1,5 M NaCl, 0,5 M NaOH), 2 min. neutralisiert (1,5 M NaCl, 0,5 M TrisHCl, pH 7,5) und 30 sek. gewaschen (2X SSC). Die Filter wurden an der Luft getrocknet und 2 h bei 80°C gebacken.

Als Proben für die Hybridisierung wurden BACs verwendet. Diese stellen mit einer durchschnittlichen Länge von ca. 100 kb eine Komplexe Probe dar die nicht ausreichend spezifisch gelabelt werden kann. Deshalb wurden die BACs mit Eco RI verdaut und die Banden auf einem 0,7% low melt Agarose Gel aufgetrennt. Die Vektorbande wurde aus dem Gel herausgeschnitten. Die restlichen Banden wurden in 4-5 Fraktionen (abhängig von der Größe des BACs) zusammengefaßt, so daß jede Fraktion ca. 20 - 25 kb ausmachte und die DNA aus dem Gel eluiert. Die DNA Fraktionen wurden einzeln gelabelt und zur Hybridisierung der Phagenfilter eingesetzt. I.d.R. wurden je 50 Platten ausplattiert und abgezogen. Entsprechend wurden jeweils 50 Filter 3 h bei 65°C zusammen in einem Volumen von 50 ml prähybridisiert (6XSSC, 5X Denhardt's, 0,5% SDS und 100 ng/ml denaturiertes Heringssperma), wobei jeder Filter einzeln in die Prähybridisierungslösung reingelegt und benetzt wurde. Die Filter wurden einzeln in die Hybridisierungslösung (6XSSC, 5X Denhardt's, 0,5% SDS und gelabelte Probe) überführt und üN. bei 65°C hybridisiert. Die Filter wurden 15 min. in 2X SSC, 0,5% SDS bei RT und 15 min. in 0,5X SSC, 0,1% SDS bei 65°C gewaschen. Aus der Hybridisierungslösung und den Waschlösungen wurden die Filter jeweils einzeln überführt. Die feuchten Filter wurden in Frischhaltefolie eingepackt und 2 h auf "storage phosphor screens" aufgelegt und mit einem Phosphoimager (Molecular Dynamics, USA) eingescannt. Die positiven Signale wurden auf Folie ausgedruckt und auf die entsprechenden Platten übertragen, so daß die positiven Plaques ausgestochen werden konnten. Diese werden in 200 µl SM + 40 µl Chloroform aufgenommen, gevortext und bei 4°C gelagert.

Die positiven Phagen wurden erneut wie oben beschrieben gescreent, wobei für jeden Klon jeweils 5 Platten einer Verdünnungsreihe (1:10 - 1:10⁵ in SM) ausplattiert wurden. Um positive Phagen zu vereinzeln muß evtl. ein zweites Mal regescreent werden. Vereinzelte positive Klone wurden nach dem in vivo excision Protokoll (Stratagene) gezappt und lagen als Plasmid vor.

Da die Filter feucht bleiben können sie mehrmals zum Hybridisieren verwendet werden. Starke Signale kann man mit heißer, 0,1% SDS-lsg. 20 min. bei 65°C stripfen.

2.21 DNA Sequenzierung

Sequenzierungen wurden nach der Methode des Cycle-Sequencing durchgeführt. Die Reaktionen erfolgten mit dem BigDye-Kit (Perkin Elmer, USA) bzw. DynamicET-Kit (Amersham). Für einen Ansatz wurden die Angaben des Herstellers jeweils halbiert, die weiteren Bedingungen blieben unverändert. Die Analyse der Reaktionen erfolgte mit einem ABI PRISM 377 (Perkin Elmer) nach den Richtlinien des Herstellers.

Für Sequenzreaktionen aus Plasmiden wurden der M13-universal- bzw. der M13-reverse-Primer (Stratagene, USA) eingesetzt. Weitere Sequenzierschritte wurden mit aus der Sequenz abgeleiteten Oligonukleotiden durchgeführt. Die Plasmid DNA wurde vor der Sequenzierung über Anionenaustauscher-Säulen gereinigt (Quiagen, Hilden und Macherey-Nagel, Düren).

2.22 Computer Analysen

Die Analyse und Weiterverarbeitung der Daten wurde an PC's, Macintosh Power PC's und einer Silicon Graphics Workstation vorgenommen. Zur Datenverarbeitung wurde MS Excel eingesetzt. Zur Erstellung der contigs wurde Corel Draw verwendet. Bilder wurden mit Adobe Photoshop bearbeitet. DNA-Sequenzen wurden lokal mit GCG und Lasergene analysiert und über das Internet mit Datenbanken verglichen (BLAST).

Folgende Internet Datenbanken und Ressourcen wurden abgefragt:

ABRC - Arabidopsis Information Management System
ARABIDOPSIS BIOLOGICAL RESOURCE CENTER Columbus, OH 43210 USA
<http://aims.cps.msu.edu/aims/>

Arabidopsis Genome Annotation Database (AGAD)
<http://www.tigr.org/tdb/at/agad/>

Arabidopsis Genomic View
<http://genome-www3.stanford.edu/Arabidopsis/chromosomes/>
<http://www.arabidopsis.org>

Arabidopsis thaliana Database (AtDB) at Stanford
<http://genome-www.stanford.edu/Arabidopsis/>

The Arabidopsis Information Ressource (TAIR)
<http://www.arabidopsis.org>

Arabidopsis thaliana WU-BLAST2 Search
<http://genome-www2.stanford.edu/cgi-bin/AtDB/nph-blast2atdb>
<http://www.arabidopsis.org>

Arabidopsis thaliana Genome Center at U.Penn
<http://cbil.humgen.upenn.edu/~atgc/ATGCUP.html>

Arabidopsis thaliana BAC (bacterial artificial chromosome) data

http://www.mpimp-golm.mpg.de/101/mpi_mp_map/bac.html

BAC CLONES CHOSEN BY THE ATGC FOR SEQUENCING BY THE SPP

http://genome.bio.upenn.edu/physical-mapping/BAC_data/SPPBACs.html

BAC tiling path of Chromosome 1

http://genome.bio.upenn.edu/physical-mapping/BAC_data/BAC_tiling_path.html

MIPS

<http://www.mips.biochem.mpg.de/>

BAC End Sequence Database Search

http://www.tigr.org/tdb/at/abe/bac_end_search.html

Entrez/PubMed - National Center for Biotechnology Information

PubMed, BLAST, GenBank

<http://www.ncbi.nlm.nih.gov/>

FPC Clone Search

<http://genome.wustl.edu/gsc/cgi-bin/fpcsearch.py>

Genome Sequencing at the ATGC

<http://genome.bio.upenn.edu/SPP.html>

Genome Sequencing Center at Washington U.

Search the Arabidopsis thaliana BAC fingerprint database

<http://genome.wustl.edu/gsc/arab/arabidopsis.html>

Lister & Dean RI Map Overview

<http://genome-www3.stanford.edu/cgi-bin/AtDB/RIintromap>

NASC - Nottingham Arabidopsis Stock Center

<http://nasc.nott.ac.uk/>

RZPD (RessourcenZentrumPrimärDatenbank).

Ressourcenzentrum im Deutschen Humangenomprojekt am MPI für molekulare Genetik
und am Deutschen Krebsforschungszentrum

<http://www.rzpd.de/>

Max Planck Institute of Molecular Genetics, Berlin

<http://www.mpimg-berlin-dahlem.mpg.de/>

TAMU BAC Library

<http://http.tamu.edu:8000/~creel/>

The MPI-MP-Map of the Arabidopsis thaliana Genome

Mozo et al., Plant J. 16, 377-384, 1998

http://www.mpimp-golm.mpg.de/101/mpi_mp_map/access.html

GENSCAN

<http://genome.dkfz-heidelberg.de/cgi-bin/GENSCAN/genscan.cgi>

<http://genes.mit.edu/GENSCAN.html>

(Burge, C. & Karlin, S., 1997, Prediction of complete gene structures in human genomic DNA. J.Mol. Biol. 268, 78-94)

Dotplots at Virtual Genome Centre at the University of Minnesota

<http://alces.med.umn.edu/rawdot.html>

3 Ergebnisse

Das *Arabidopsis thaliana*-Centromer auf Chromosom 1 sollte dahingehend untersucht und charakterisiert werden, das funktionelle CEN 1 in ein möglichst kleines genetisches und physikalisches Intervall einzuschließen. Die Position und der Aufbau des funktionellen CEN 1 sollte anhand molekularer und histologischer Charakterisierung genauer festgelegt werden.

3.1 Genetische Charakterisierung der Centromer 1-Region

Um die genetischen Abstände in der Centromerregion des Chromosoms 1 abzuschätzen, wurden die Kartierungsdaten des *gurke*-Gens verwendet, da *GURKE* in die Centromerregion kartiert (Fischer, 1997). Anhand einer Telotrisomanalyse (Patton et al., 1991) konnte gezeigt werden, daß *emb22 = gk* auf dem unteren Chromosomenarm des Chromosoms 1 rechts vom Centromer 1 liegt.

Als rechter Startpunkt der Kartierung war GapB vorgegeben (Fischer, 1997). Auf der linken Seite des Centromers fehlte ein naher, eindeutiger Marker (Nia-2 cosegregierte zunächst mit *gk*, Fischer, 1997). Als linken Startpunkt der Kartierung wurde deshalb der morphologische Marker *tt1* ausgewählt. Dieser liegt mit Sicherheit auf der linken Seite des CEN 1 (Koornneef, 1983-2) und hat den großen Vorteil, daß man die Kartierpopulation schnell und ohne großen Aufwand vorselektionieren kann.

Davon ausgehend wurde das genetische Intervall mit weiter innenliegenden Markern verkleinert. Die verwendeten Marker waren z.T. schon bekannt (mi342, mi133, Nia-2) bzw. wurden im Laufe eines "chromosomal walks" selber generiert (3l22-sp6, 6j23-t7, 17a20-sp6, 4a3-t7, 1c13-t7 und weitere HAMDE Marker; 2.19.4).

3.1.1 Kartierung von der GapB (rechte) Seite

Die Population aus 1000 F2 Pflanzen ($+^{Nd}/+^{Nd} \times gk^{Ler}/+^{Ler}$; Fischer, 1997) wurde mit einer Population aus 296 F2 Pflanzen ($+^{Nd}/+^{Nd} \times gk_{R216}^{Ler}/+^{Ler}$) erweitert und von der rechten Seite kartiert. Eine Liste der rekombinanten Pflanzen findet sich im Anhang (7.1.1).

GapB zeigt mit 0,35 cM eine eindeutige Kopplung zu *gk*. Mit mi133 konnte der genetische Abstand auf 0,19 cM reduziert werden. Die beiden cosegregierenden und codominanten Marker 3l22-sp6 und 6j23-t7 zeigen mit 0,08 cM eine sehr enge Kopplung zu *gk*.

Marker	"Art"	Zahl der Pflanzen	Zahl der Chromosomen	Zahl der rekombinanten Pflanzen	genet. Abstand (cM)
GapB	CAPS	1296	2592	9	0,35
mi133	RFLP	1296	2592	5	0,19
3L22-sp6	HAMDE	1296	2592	2	0,08
6J23-t7	HAMDE	1296	2592	2	0,08

Tab. 8: molekulare Marker die von der rechten Seite gegen *gk* kartiert wurden

3.1.2 Kartierung von der *tt1* (linke) Seite

Mit einer Population aus 2038 F2 Pflanzen ($+^{Ler} tt1^{Ler}/+^{Ler} tt1^{Ler} \times gk_{emb22}^{Col+Col/+^{Col} +^{Col}}$), die mit *tt1* vorselektioniert wurden, wurde von der linken Seite kartiert. Eine Liste der rekombinanten Pflanzen findet sich im Anhang (7.1.2).

Die Vorselektion der Population mit *tt1* ergab einen genetischen Abstand von 1,91 cM. Man erkennt eine genetische Kopplung mit *gk*. Die Zahl der Pflanzen, die weiter untersucht werden sollten, konnte erheblich von 2038 auf 78 *tt1*-rekombinante Pflanzen reduziert werden. *mi342* koppelt mit 0,49 cM an *gk*. Mit den Markern *1c13-t7* und *4a3-t7* konnte der genetische Abstand weiter auf 0,15 bzw. 0,12 cM reduziert werden. Die beiden Marker *17a20-sp6* und *Nia-2* kosegregieren mit *gk*.

Marker	"Art"	Zahl der Pflanzen	Zahl der Chromosomen	Zahl der rekombinanten Pflanzen	genet. Abstand (cM)
<i>tt1</i>	morpho.	2038	4076	78	1,91
<i>mi342</i>	RFLP	2038	4076	20	0,49
<i>1C13-t7</i>	HAMDE	2038	4076	6	0,15
<i>4A3-t7</i>	HAMDE	2038	4076	5	0,12
<i>17A20-sp6</i>	RFLP	2038	4076	0	0
<i>Nia-2</i>	CAPS	2038	4076	0	0

Tab. 9: Morphologische und molekulare Marker, die von der linken Seite gegen *gk* kartiert wurden

3.1.3 Genetische Karte der Centromer 1-Region

Die oben aufgeführten Kartierergebnisse ermöglichen es, daß Centromer 1 in ein sehr kleines genetisches Intervall von 0,20 cM (4A3-t7 - 6J23-t7) einzuschließen, und führen zu folgender sehr präzisen genetischen Karte der Centromer 1-Region.

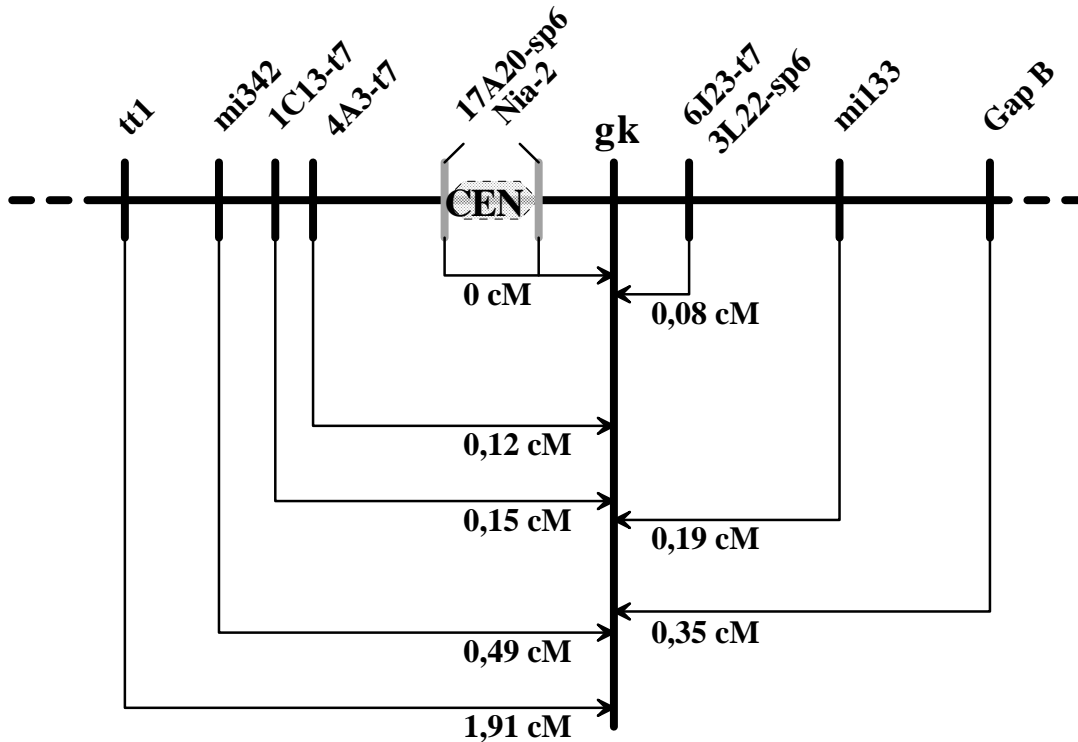


Abb. 3: genetische Karte der Centromer 1-Region

Die beiden Marker 17A20-sp6 und Nia-2 kosegregieren mit *gk*. Anhand der genetischen Daten kann man ihre Position am Centromer 1 (nördlich oder südlich) nicht bestimmen.

(Die Abstände in der Abbildung sind nicht maßstabsgerecht wiedergegeben.)

3.2 Molekulare Charakterisierung der Centromer 1-Region

Im Zuge der Kartierung des *gk*-Gens wurde eine YAC contig aufgestellt, der das Centromer 1 wahrscheinlich einschließt, aber nicht überspannt (Fischer, 1997). Auf der linken Seite des contigs wurden IGF BACs gefunden, die mit NIA2 hybridisieren (Fischer, 1997). Diese IGF BACs (17a20, 13j4, 22a12, 24p1, 7g10, 22c3, 6m21 und 17g21) und IGF BACs, die mit GapB hybridisieren, stellten die Startpunkte für die Erstellung eines BAC contigs um die Centromer 1-Region dar. Der BAC contig wurde als chromosomal walk über BAC-End-Hybridisierungen und "finger print"-Analysen aufgestellt. Enden von bekannten BACs wurden verwendet, um über Hybridisierungen auf BAC Filter überlappende BACs zu finden. Diese wurden anhand einer "finger print"- und Southern-Analyse untereinander und mit den bekannten BACs angeordnet. So konnten falsche BACs und BACs, die repetitive Elemente enthalten und nicht eindeutig zuzuordnen waren, aus dem "contig" ausgeschlossen werden.

Dabei wurden die Ergebnisse anderer Gruppen einbezogen, um BAC-Enden als Hybridisierungsproben zu generieren, die eigenen Daten zu verifizieren bzw. Lücken zu überbrücken.

Hybridisierungsproben: BAC End Sequence Database Search
http://www.tigr.org/tdb/at/abe/bac_end_search.html;

BAC-End-Hybridisierungen: Arabidopsis thaliana BAC data
http://www.mpimp-golm.mpg.de/101/mpi_mp_map/bac.html und
 BAC tiling path of Chromosome 1
http://genome.bio.upenn.edu/physical-mapping/BAC_data/BAC_tiling_path.html;

finger print Analysen: FPC Clone Search
<http://genome.wustl.edu/gsc/cgi-bin/fpcsearch.py>

Gegenwärtig ist der contig in der CEN 1 Region bis an die Ränder des CEN 1 Kerns entwickelt (The Arabidopsis Genome Initiative, 2000; Theologis et al., 2000). Der entscheidende Punkt beim Vergleich von physikalischen und genetischen Daten (und der Erstellung dieses BAC contigs) war es, eindeutige Hybridisierungsproben zu finden. Dies stellte sich als sehr schwierig dar, da man in den Centromer-Regionen von *Arabidopsis* gehäuft repetitive DNA findet (eigene Daten, Hemleben, et al., 2000), die für genetische und physikalische Kartierungen ungeeignet ist. Die Proben 3N5-t7 und 17A20-sp6 bilden eine Ausnahme, da sie zwar in drei bzw. zwei Kopien auf Chromosom 1 vorkommen, aber dennoch brauchbare Marker darstellen. Um eindeutige Proben zu finden, wurden zuerst zwei BACs (F17A20 und F17G21) subkloniert und die einzelnen Klone auf ihre Repetitivität untersucht. Als weniger aufwendig erwies es sich, sequenzierte BAC Enden zu verwenden (BAC End Sequence Database Search).

Die Enden wurden über BLAST-Vergleiche vorselektioniert (Arabidopsis thaliana WU-BLAST2 Search, <http://genome-www2.stanford.edu/cgi-bin/AtDB/nph-blast2atdb>, <http://www.arabidopsis.org> und BAC End Sequence Database Search).

Nur Enden, die ein bis fünf hits produzierten, wurden weiter verwendet. Alle anderen Enden wurden verworfen, da sie teilweise bis zu 100 hits produzierten, und dies als Indiz für eine repetitive Sequenz gewertet wurde. Anhand der Sequenz wurden PCR-Produkte generiert und als Probe für BAC Filter Hybridisierungen verwendet. Es zeigte

sich, daß 75 % der untersuchten Proben tatsächlich "single copy" sind. Die anderen erwiesen sich als repetitiv (zwei bis viele Kopien im Genom). Als Maßstab für die Zahl der Kopien einer Probe diente die Zahl der Hybridisierungssignale auf den BAC-Filtern. 10 - 15 Signale (= BACs) wurden als single copy eingestuft. Bei mehr Signalen muß man schon von zwei oder mehreren Kopien ausgehen (dies konnte anhand von "finger print"-Analysen gezeigt werden).

Nur $\frac{1}{3}$ der vorselektionierten BAC-Enden wurden weiter analysiert. $\frac{3}{4}$ dieser BAC-Enden erwiesen sich als single copy und konnten als Hybridisierungsproben und Marker verwendet werden.

Durch die Art und Weise, wie die IGF-BAC-Bibliothek konstruiert wurde (Mozo et al., 1998), kann es vorkommen, daß die Enden verschiedener BACs einer Ausgangsprobe identisch sind. Die identischen BAC-Enden wurden bei den BLAST-Vergleichen erkannt. Nur das Ende mit der längsten Sequenz wurde weiterverwendet.

Ausgangsprobe	BACs deren Enden analysiert wurden	Zahl der verwendbaren BAC Enden	s.c.
NIA2	24	7	6
f17a20sp6	15	10	8
f22a12sp6	4	3	0
t4n16t7	7	0	0
f11o6sp6	8	2	2
mi342	21	4	3
f3l22sp6	22	3	3
Summe	101	29	22

Tab. 10: BACs deren Enden auf die Tauglichkeit als Hybridisierungsproben untersucht wurden.

3.2.1 BAC-Subklone

Links vom CEN 1, um den Marker NIA2 waren keine weiteren Marker bekannt. Da Marker in dieser Region essentiell für die genetische und physikalische Kartierung sind, sollten sie aus Subklonen der BACs F17A20 und F17G21 generiert werden. Bei der Subklonierung der beiden BACs wurden zwei Strategien verfolgt. Zum einen wurden differentielle Banden (EcoRI Restriktionsverdau) kloniert, da diese Banden sich an den äußeren Enden der beiden BACs befinden und somit einen möglichst großen Abstand von NIA2 gewährleisten. Zum anderen wurden alle Restriktionsbanden beider BACs kloniert, um alle möglichen Marker erfassen zu können.

3.2.1.1 Klonierung

Die Inserts der BACs wurden nach einer Plasmidpräparation mit NotI aus dem Vektor herausgeschnitten und mit EcoRI restriktionsverdaut. Die differentiellen Banden wurden auf einem Agarosegel aufgetrennt und eluiert. Für die Einzelklonierungen wurden die Banden A1, A2 (F17A20) und G1 (F17G21) und für die "shotgun" Klonierung der gesamte Restriktionsverdau eingesetzt. Die Restriktionsfragmente wurden mit EcoRI linearisiertem, dephosphoriliertem pBlueskriptIIKS+ Vektor (Stratagen) ligiert und in Rubidiumchlorid-kompetente XL1 Blue-Zellen (Stratagen) hitzeschock transformiert. Die Transformanten erkennt man durch die Blau-Weiß-Selektion. Von den Fragmenten A1, A2 und G1 wurden je fünf positive Klone gepickt. Aus der "shotgun"-Klonierung wurden jeweils 40 Klone gepickt; as1-40 (F17A20) und gs1-gs40 (F17G21). Die Klone wurden alle präpariert und mit EcoRI restriktionsverdaut. Die Restriktionsfragmente wurden auf einem Agarosegel aufgetrennt.

Bei der "shotgun"-Klonierung wurden nur Fragmente kleiner 3 kb kloniert. Wie ein reverses Southernexperiment zeigte (3.2.1.2), wurden einige Fragmente mehrmals kloniert.

3.2.1.2 Hybridisierung

Um die geeigneten single copy-Fragmente zu identifizieren, wurden alle Klone mit *Arabidopsis* (Col) DNA revers hybridisiert. Bei diesem reversen Southern liefern alle unigen Klone ein schwaches bzw. kein Hybridisierungssignal. Repetitive Klone zeigen stärkere Signale. Bei diesem Hybridisierungsexperiment dienten GapB und mi342 als unique Kontrolle und pRib, ein ribosomaler Klon (G. Haberer), als mittelrepetitive Kontrolle.

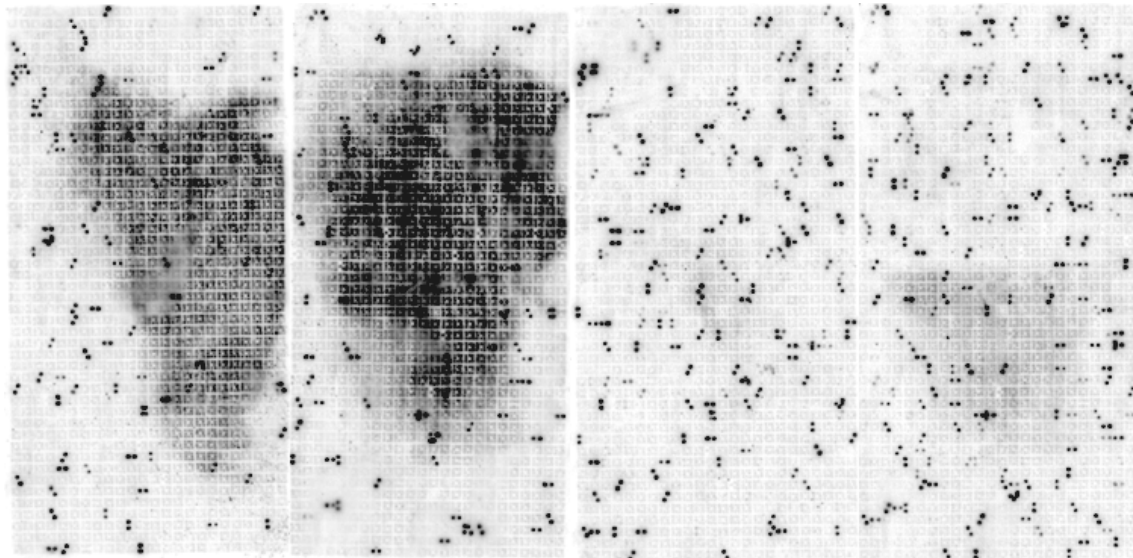
kein Signal	schwaches Signal	mittelstarkes Signal	starkes Signal
a2-3	as28	as9	a1-1
a2-4	gs10	as10	as20
as1	gs22	gs17	as29
as6	gs26	gs21	as31
as7	gs46	gs31	
as11		gs33	
as12		gs37	
as13		gs40	
as14			
as15			
as26			
as32			
gs7			
gs14			
gs20			
gs28			
gs29			
gs30			
gs38			

Tab. 11: Zusammenfassung der Hybridisierungsergebnisse.

Trotz dieser Vorselektion (kein bzw. nur ein schwaches Signal) erwiesen sich alle Klone als repetitiv und somit ungeeignet als Marker.

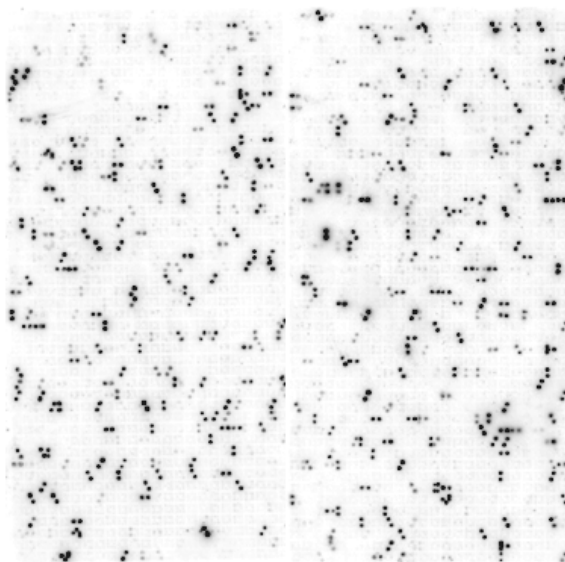
Die Klone konnten nicht für die genetische bzw. die physikalische Kartierung verwendet werden. Drei der Klone a2-4 (kein Signal), gs31 (schwaches Signal) und as31 (starkes Signal) wurden weiter untersucht, um mehr Informationen über die Qualität der repetitiven Elemente zu erhalten.

Die drei Klone wurden als Proben verwendet um die IGF BAC Filter zu hybridisieren.



IGF BACs / a2-4

IGF BAC / gs31



IGF BACs / as31

4	22	16	2	20	14	P
10		10	8		8	A
4	16	22	2	14	20	
1	19	13	6	24	18	P
7		7	12		12	A
1	13	19	6	18	24	
5	23	17	3	21	15	P
11		11	9		9	A
5	17	23	3	15	21	
48		25	24		1	

IGF BAC-Filter Spotanordnung

Abb. 4: Hybridisierungssignale der Klone a2-4, gs31 und as31. Anhand der Spotanordnung können die BAC-Filter ausgewertet werden. Auf einem Filter sind die BACs von 24 Platten mit je 16 Zeilen (A - P) und 24 Spalten gespottet. Zur sicheren Identifizierung ist jeder BAC zweimal aufgetragen. Das Spotmuster gibt die Plattenzahl an. Die Position auf dem Filter gibt die Lage des BACs auf der Platte an.

Die Auswertung der Signale findet man im Anhang 7.3.1

Chromosom	contig	a2-4	gs31	as31
chr1	ctg2	24	27	61
	ctg3	6	2	18
	ctg4	7	13	8
			4	1
	Σ	37	46	88
chr2	ctg1		5	4
	ctg2	23	37	42
				1
	Σ	23	42	47
chr3	ctg2			3
	ctg4		37	4
	ctg6			1
	Σ		37	8
chr4	ctg1	36	21	5
	ctg2	12	21	41
	ctg3	1		1
	Σ	49	43	47
chr5	ctg3		11	30
	ctg4	6	41	20
	ctg5		1	1
	Σ	6	53	51
ungeordnet	ctg1	24	25	26
	ctg2	5	8	23
	ctg3	2	11	
	Σ	34	44	49
nicht bek.		19	9	9
Summe		168	274	298

Tab. 12: Verteilung der IGF BACs im *Arabidopsis thaliana*-Genom, die mit den repetitiven Proben a2-4, gs31 und as31 hybridisieren (Abb. 4, 7.2.1).

Die Zahl der mit den drei Klonen hybridisierenden BACs (a2-4 168, gs31 274 und as31 298) zeigt, daß die Klone repetitiv sind. Des weiteren zeigt sich, daß die drei Klone ausschließlich um die fünf Centromerregionen zu finden sind. Man findet nicht alle drei Klone auf allen fünf Chromosomen und auf beiden Seiten der entsprechenden Centromeren (Tab. 12, Abb. 5). Allerdings bleibt festzustellen, daß gut 20% der gefundenen BACs sich nicht in einen contig auf einem der fünf Chromosomen einordnen lassen. Diese könnten die Verteilungslücken auffüllen.

Nur auf Chromosom 1 findet man alle drei Klone links und rechts vom CEN 1 vertreten. Das CEN 1 wird dabei in die contig Lücke zwischen NIA2 und mi133 positioniert, da die Lücke im BAC contig zwischen GapB und mi291a von einem YAC, CIC 11d19, überspannt und somit geschlossen wird (Abb. 5, Arabidopsis thaliana Genome Center at U.Penn, <http://cbil.humgen.upenn.edu/~atgc/ATGCUP.html>).

Auf Chromosom 2 findet man nur gs31 auf beiden Seiten des CEN 2. a2-4 und as31 findet man nur rechts von CEN2 (Abb. 5).

Auf Chromosom 3 findet man nur den Klon gs31 rechts vom CEN 3. Es bleibt festzustellen, daß auf dem 3 Chromosom links vom CEN 3 kein IGF BAC contig besteht und deshalb hier keine Aussage über BACs aus jener Region gemacht werden kann (Abb.5).

Auf Chromosom 4 findet man a2-4 und gs31 links und rechts vom CEN 4. as31 findet man nur auf der rechten Seite des CEN 4 (Abb.5).

Auf Chromosom 5 findet man as31 und gs31 links und rechts vom CEN 5. a2-4 findet man nur auf der rechten Seite des CEN 5 (Abb.5).

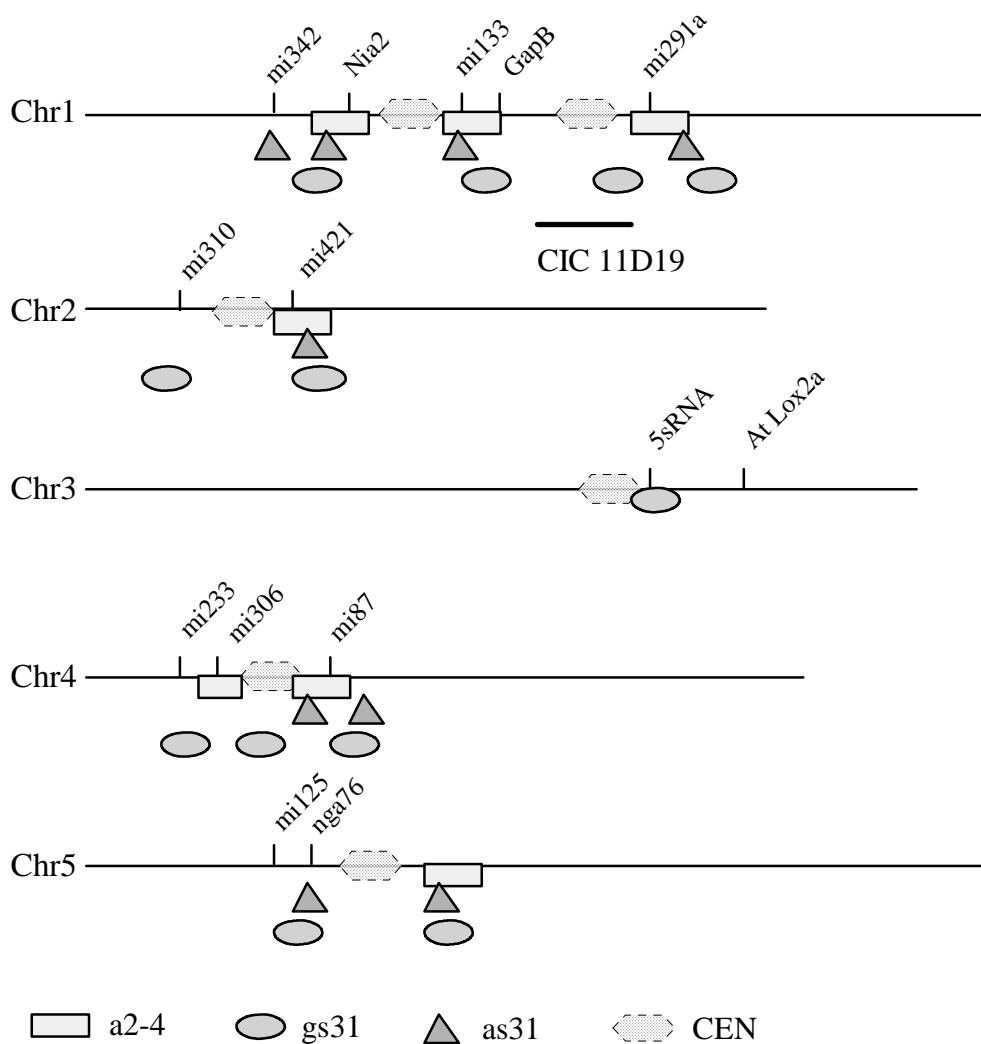


Abb. 5: Gruppen der IGF-BACs im *Arabidopsis thaliana*-Genom, die mit den repetitiven Proben a2-4, gs31 und as31 hybridisieren (7.2.1). Die Karte des Genoms entspricht der physikalischen *Arabidopsis*-Karte von TAIR.

Das zweite und vierte Chromosom sind schon durchsequenziert und annotiert. Die fehlenden Chromosomen werden voraussichtlich bis Ende 2000 fertigsequenziert werden. (*Arabidopsis thaliana* Database, <http://www.arabidopsis.org/cgi-bin/maps>, *Arabidopsis* Physical Map Close-up).

3.2.1.3 Sequenzierung

Die Identität der drei Klone a2-4, as31 und gs31 sollte anhand ihrer Sequenz näher charakterisiert werden. Die Sequenzierung der 3 Klone erfolgte schrittweise über Primer walking. Die einzelnen Sequenzen eines Klons wurden mit dem Lasergen-Paket zu einem contig angeordnet. Uneindeutige Nukleotide wurden anhand der Elektropherogramme zugeordnet. Die Sequenzen der drei Klone findet man im Anhang 7.3.2.1 - 3.

Die Sequenzen wurden verwendet, um sie mit bekannten Nukleotid- (BLAST N) und Aminosäuresequenzdatenbanken (BLAST X, TAIR) zu vergleichen.

Der Vergleich auf Nukleotidebene zeigt nur Homologien zu sequenzierten BAC-Enden bzw. zu sequenzierten BACs, nicht jedoch zu Genen. Die Zahl der BLAST hits läßt darauf schließen, daß die drei Klone mittelrepetitiv sind. Ihre Sequenzen findet man auch in dem sequenzierten BAC F28L22 wieder (3.6.1). Dieser BAC überlappt mit den beiden BACs F17A20 und F17G21.

Vergleicht man die Sequenzen mit BLAST X, so findet man für alle drei Klone sehr gute Homologien zu Retroelementen (siehe Tab. 13). Die drei Klone fallen zusätzlich mit GENSCAN vorhergesagten ORFs des BAC F28L22 zusammen (Abb. 20).

Damit konnte gezeigt werden, daß die drei Klone Retroelemente repräsentieren und somit als repetitiv einzustufen sind.

Die BLAST X - und Hybridisierungsergebnisse (3.2.1.2, Tab.11) und die Verteilung der mit den Klonen a2-4, gs31 und as31 hybridisierenden BACs um die CEN-Regionen 1-5 (3.2.1.2, Abb. 5) lassen den Schluß zu, daß man in den CEN-Regionen gehäuft Retroelemente findet.

Klon	Homologien zu
a2-4	gb AAD22283.1 AC006920_7 (AC006920) putative reverse transcriptase [Arabidopsis thaliana] Length = 1787 Score = 948 bits (2424), Expect = 0.0 gb AAD20433 (AC006446) putative pol polyprotein with a reverse transcriptase domain [Arabidopsis thaliana] Length = 889 Score = 808 bits (2064), Expect = 0.0
as31	gi 3377816 (AF076275) contains similarity to Arabidopsis thaliana retrotransposon Athila (GB:X81801) [Arabidopsis thaliana] Length = 940 Score = 657 bits (1676), Expect = 0.0 gb AAD15359 (AC006250) putative Athila retroelement ORF1 protein [Arabidopsis thaliana] Length = 532 Score = 427 bits (1085), Expect = e-118
gs31	gi 3377816 (AF076275) contains similarity to Arabidopsis thaliana retrotransposon Athila (GB:X81801) [Arabidopsis thaliana] Length = 940 Score = 657 bits (1676), Expect = 0.0 gb AAD15359 (AC006250) putative Athila retroelement ORF1 protein [Arabidopsis thaliana] Length = 532 Score = 427 bits (1085), Expect = e-118

Tab. 13: BLAST X-Homologien der Klone a2-4, as31 und gs31

3.2.2 Single copy Proben

Die Centromer 1-Region sollte mit einem BAC contig dargestellt werden.

Als Startpunkt diente auf der wahrscheinlich linken Seite des CEN 1 NIA2. Die mit NIA2 ermittelten BACs wurden mit einer finger print-Analyse angeordnet. Anschließend wurde der contig in beide Richtungen vorangetrieben. Zum Centromer hin (proximal) konnte der contig nicht verlängert werden, da nur repetitive Proben gefunden wurden. Das sp6-Ende des IGF BACs 22A12 schließt den contig zum CEN 1 von der linken Seite ab.

Vom Centromer weg (distal) wurde der contig über die Proben t3h3t7, t4n16sp6 und f17a20sp6 fortgeführt (3.2.4, 7.2). Die Verlängerung des contigs von f17a20sp6 über f11o6sp6, f3n5t7 zu mi342 erwies sich als falsch, da von f17a20sp6 zwei loci vorhanden sind, einer bei NIA2 und der andere zwischen mi342 und tt1. Mit TAMU-BACs konnte der richtige contig weitergeführt werden, von f17a20sp6 zu t4n16t7, t18h22sp6, t31i22t7 und t21c16t7 (3.2.4, 7.2). Ein Vergleich mit dem IGF BAC contig des MPI-MP (Mozo, T., et al., 1999, http://www.mpimp-golm.mpg.de/101/mpi_mp_map/access.html) zeigte, daß die beiden contigs ab diesen Proben übereinstimmten. Deshalb wurden zur Fertigstellung des contigs bis zu dem Marker mi342 und dem zweiten locus von f17a20sp6 die Hybridisierungsdaten des MPI-MP contigs übernommen und mit eigenen Daten ergänzt.

Auf der rechten Seite des CEN 1 wurde mit dem Marker GapB begonnen (3.2.4, 7.2). Die Lücken im contig zwischen den Proben jcc3, mi133 und f28b8t7 wurden mit Hilfe der Hybridisierungsdaten des MPI-MP contigs geschlossen. Von f28b8t7 konnte der contig mit zwei weiteren Proben, f3122sp6 und f6j23t7, verlängert werden (3.2.4). Weitere geeignete Proben konnten nicht gefunden werden. Das t7 Ende des IGF BACs 9d18 schließt den contig zum CEN 1 hin ab (3.2.4).

Marker	f22a12-sp6	t3h3-t7	t4n16-sp6	f17a20-sp6
BAC	F28L22	F28L22	F28L22	F28L22
Orientierung	revers	revers	revers	revers
Position	3643-3783 bp	73638-73326 bp	77656-78156 bp	97966-97791 bp

Marker	t4n16-t7	f18h22-sp6	t31i22-t7	t21c16-t7
BAC	T32E20	T32E20	T32E20	F1O3
Orientierung	revers	revers	revers	foreward
Position	52184-51855 bp	100397-9999 bp	103924-103240 bp	54497-54991 bp

Marker	f4a3-t7	f1c13-t7	mi342	TT1
BAC	T22A15	F10O5	F15O4	F11O6
Orientierung	revers	foreward	foreward	revers
Position	60614-61109 bp	8825-8486 bp	23781-23009 bp	* nicht bekannt

* (Weishaar, B., et al., persönliche Mitteilung)

Marker	f3l22-sp6	f6j23-t7	f28b8-t7	mi133
BAC	T4I21	T4I21	T4I21	F7F22
Orientierung	forward	forward	forward	revers
Position	69387-69110 bp	69392-68850 bp	89538-89055 bp	42860-44360 bp

Marker	jcc3	GAPB
BAC	T8D8	F13A11
Orientierung	forward	forward
Position	43688-45434 bp	64735-66884 bp

3.2.2.1 Hybridisierungen auf IGF BAC-Filter

Im Gegensatz zu den repetitiven Klonen a2-4, gs31 und as31, die mit 168, 274 und 298 IGF BACs hybridisieren, erkennen single copy-Proben weit weniger IGF BACs. Hier sollen exemplarisch die Hybridisierungsergebnisse der Marker mi342 (3 BACs) und jcc3 (10 BACs) dargestellt werden.

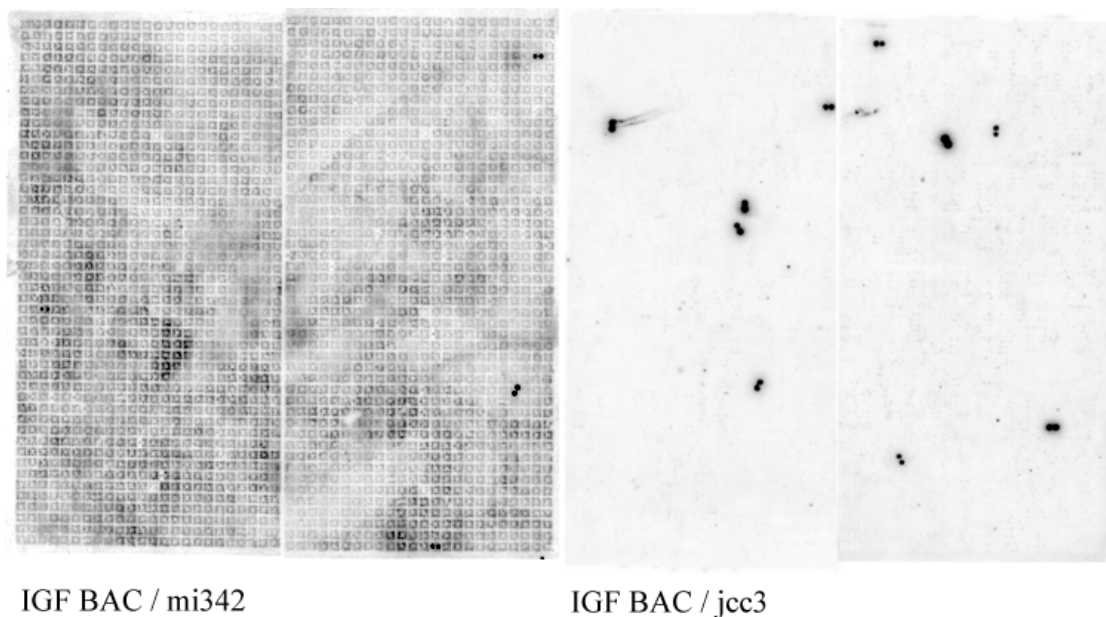


Abb. 6: Hybridisierungssignale der Marker mi342 und jcc3. Die Auswertung der Signale findet man im Anhang 7.2

3.2.3 Repetitive Proben

Die größte Schwierigkeit bei der Erstellung des contigs der CEN 1-Region war das Fehlen geeigneter single copy-Proben. $\frac{2}{3}$ der getesteten BAC-Enden erwies sich als mittel bis hoch repetitiv (s.o.), von dem verbleibenden Drittel erwies sich $\frac{1}{4}$ der Proben als "wenig" repetitiv mit zwei bis vier Kopien. Diese Proben erschwerten die Erstellung des BAC contigs, da ihre Hybridisierungsergebnisse zunächst auf einen einzigen locus zurückgeführt wurden. Erst bei Hybridisierungen auf fingerprints der gefundenen BACs konnte festgestellt werden, daß die BACs tatsächlich in mehrere Hybridisierungsgruppen fallen (3.2.3.2, 7.3.1). Dabei findet man entweder alle Kopien in der CEN 1-Region (f17a20sp6 und f3n5t7, 3.2.4) oder, wie bei f22a12sp6, eine Kopie in der CEN 1-Region und die restlichen drei Kopien im übrigen Genom. Die beiden Kopien der Probe f17a20sp6 (Abb. 8) findet man links vom CEN 1, die eine bei NIA2, die andere zwischen mi342 und tt1 (3.2.4). Die drei Kopien von f3n5t7 (Abb. 9) sind um das CEN 1 verteilt. Zwei Kopien findet man links vom CEN 1 zwischen mi342 und tt1 und eine Kopie rechts vom CEN 1 links von mi133. Die vier Kopien von f22a12sp6 (Abb. 10) findet man in den CEN-Regionen von drei Chromosomen. Eine Kopie liegt bei NIA2 auf Chr1, zwei Kopien findet man in der CEN 3-Region und die vierte Kopie in der CEN 2-Region.

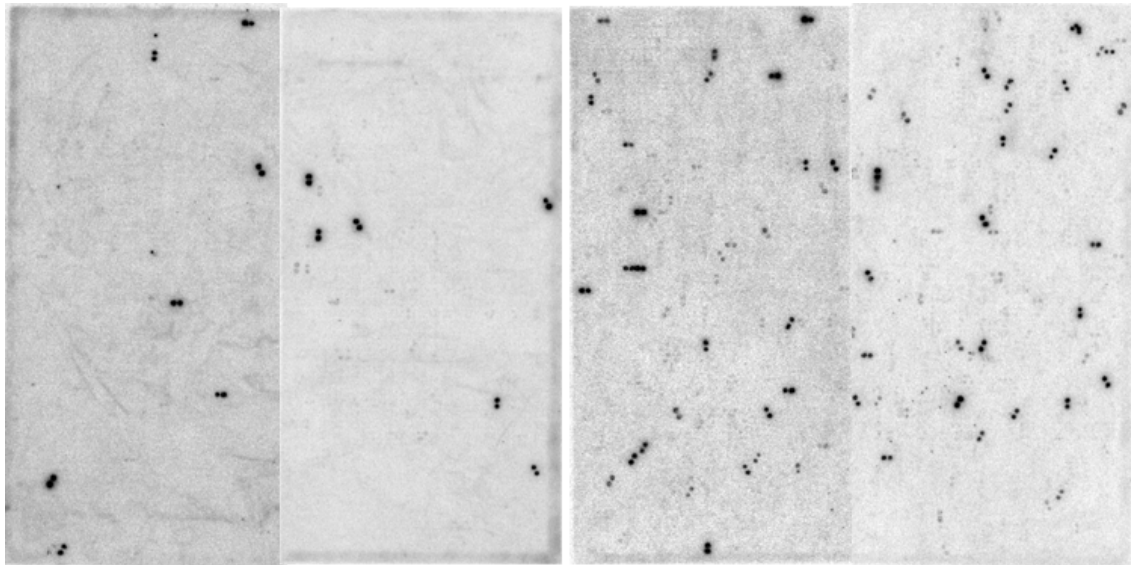
Die Vermutung liegt nahe, daß es sich bei diesen wenig repetitiven Proben um Centromer bzw. Pericentromer assoziierte Sequenzen handelt.

Probe	Zahl der Kopien
f17a20sp6	2
f3n5t7	3
f22a12sp6	4

Tab. 14: "wenig" repetitive Proben

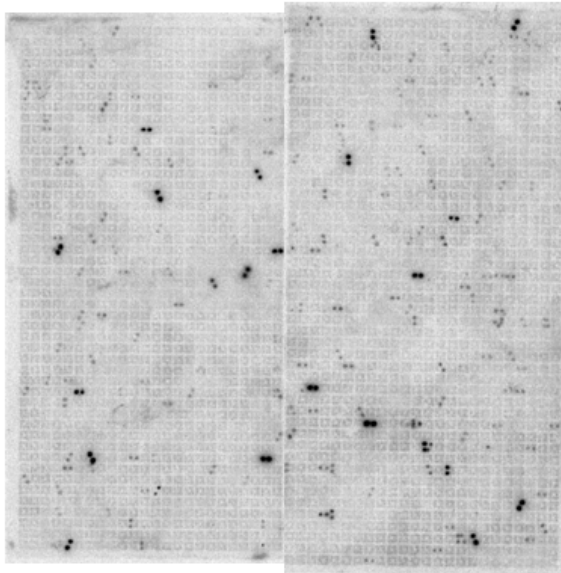
3.2.3.1 Hybridisierungen auf IGF BAC-Filter

Die "wenig" repetitiven Proben erkennen weit weniger IGF-BACs als die repetitiven Klone a2-4, gs31 und as31, die mit 168, 274 und 298 IGF BACs hybridisieren (Abb. 4, Tab 12). Im Vergleich zu den single copy-Markern, die meist mit 10 IGF BACs hybridisieren (7.2), erkennen die "wenig" repetitiven Proben teils gleich viele IGF BACs, F17a20sp6 (13 BACs), zum anderen Teil aber deutlich mehr IGF BACs, F3n5t7 (19 BACs) und F22a12sp6 (23 BACs) (Abb.7). Dies verdeutlicht die Schwierigkeit, die Zahl der Kopien einer Probe im Genom allein anhand der Zahl der hybridisierten BACs festzulegen. Im Grenzbereich, 10 - 20 BACs, werden wenig repetitive Proben leicht fälschlicherweise als single copy eingestuft. Diese Proben erschweren die Aufstellung des BACs contigs erheblich. Um gerade die "wenig" repetitiven Proben zu bestimmen, werden alle Hybridisierungsproben auf BAC "finger print"-Filter hybridisiert. Bei einer einzigen Kopie pro Genom erwartet man eine einzige Bande. Hat eine Probe mehrere Kopien pro Genom, so kann man auch mehrere Banden erwarten und so die tatsächlichen single copy Proben feststellen.



IGF BAC / F17A20sp6

IGF BAC / F3N5t7



IGF BAC / F22A12sp6

Abb. 7: Hybridisierungssignale der Proben F17a20sp6, F3n5t7 und F22a12sp6. Die Auswertung der Signale findet man im Anhang 7.3.1

3.2.3.2 BAC Fingerprint Analysen

Die BACs, die durch Hybridisieren einer Probe auf IGF BAC-Filter (Abb. 7) gefunden wurden, bedürfen weiterer Validierungsschritte. Repräsentiert eine Probe mehrere Loci im Genom, so findet man bei einer Hybridisierung auf die IGF BAC-Filter alle entsprechenden BACs. Die gefundenen BACs können anhand einer "finger print"-Analyse verschiedenen Gruppen zugeordnet werden.

Dazu wird von jedem BAC DNA präpariert und mit EcoRI restriktions-verdaut. Die Restriktionsbanden werden auf einem 0,7 % Agarosegel einer Feld-Inversions-Gel-Elektrophorese unterzogen (Pulszeiten: 5 sek. und 3sek.). Die Banden werden mit EtBr gefärbt und unter UV_{254nm} sichtbar gemacht. Alle BACs, die einem Locus entsprechen, werden mehr oder weniger (abhängig von der Größe des Überlapps) gemeinsame Banden aufweisen. Auf diese Art und Weise kann man die BACs in Gruppen zusammenfassen und sogar ihre Anordnung innerhalb einer Gruppe feststellen. Findet man nur eine BAC-Gruppe, kann man davon ausgehen, daß die entsprechende Probe "single copy" ist. Zusätzlich kann man die Probe auf den Southernblot der BACs hybridisieren und so die "finger print"-Analyse kontrollieren. Dabei sollte die Zahl der positiven Banden mindestens der Zahl der BAC-Gruppen entsprechen (Tab. 14, 3.2.3).

Für die Probe F17a20sp6 wurden zwei Banden, entsprechend zwei BAC-Gruppen (Abb. 8), für die Probe F3n5t7 zwei Bandengruppen, entsprechend zwei der drei BAC-Gruppen (Abb. 9) und für die Probe F22a12sp6 4 Banden, entsprechend vier BAC-Gruppen ermittelt (Abb. 10).

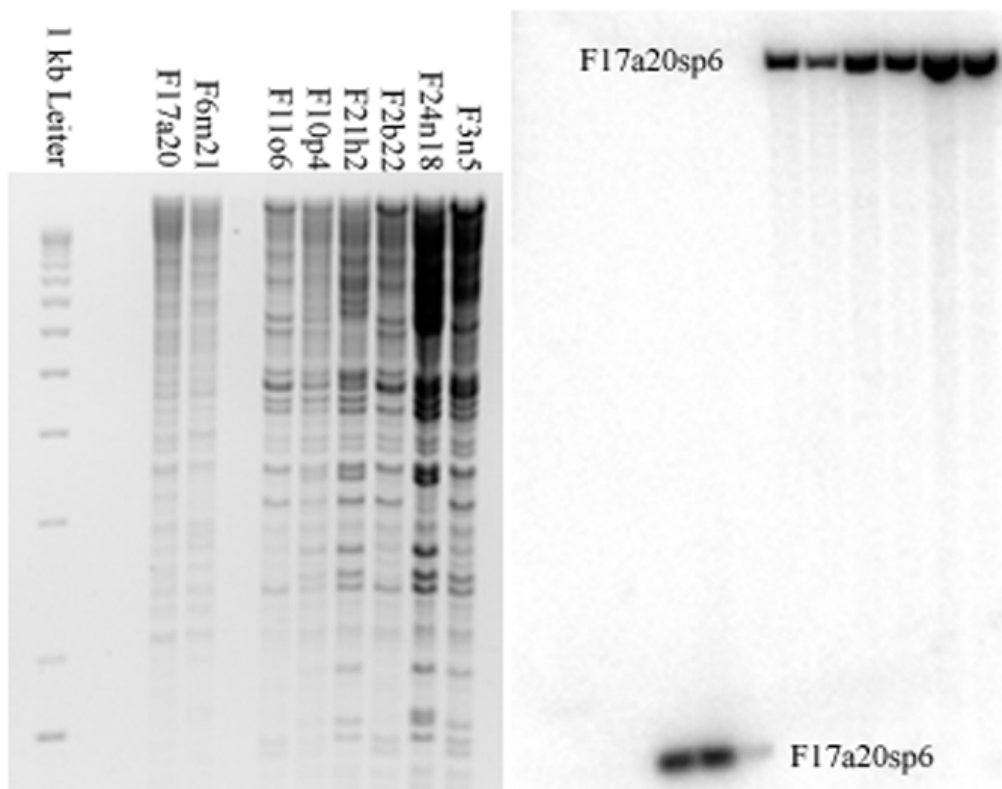


Abb. 8: finger print Analyse mit der Probe F17a20sp6. Das linke Bild stellt den finger print der BACs dar (Restriktionsverdau mit EcoRI). Das rechte Bild zeigt das Hybridisierungsergebnis mit der Probe F17a20sp6.

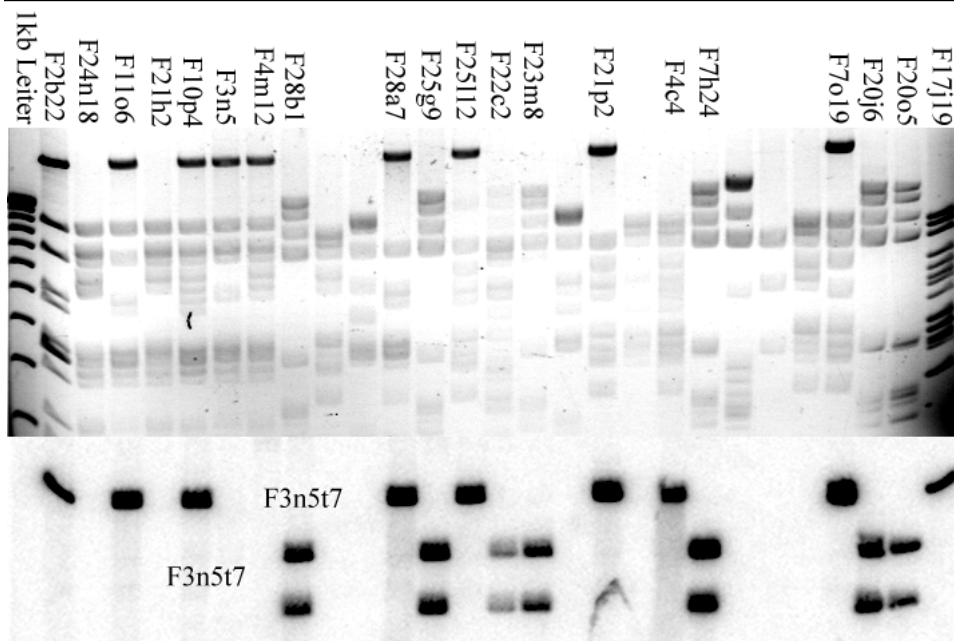


Abb. 9: finger print Analyse mit der Probe F3n5t7. Das obere Bild stellt den finger print der BACs dar (Restriktionsverdau mit EcoRI; EtBr.-Negativbild). Das untere Bild zeigt das Hybridisierungsergebnis mit der Probe F3n5t7.

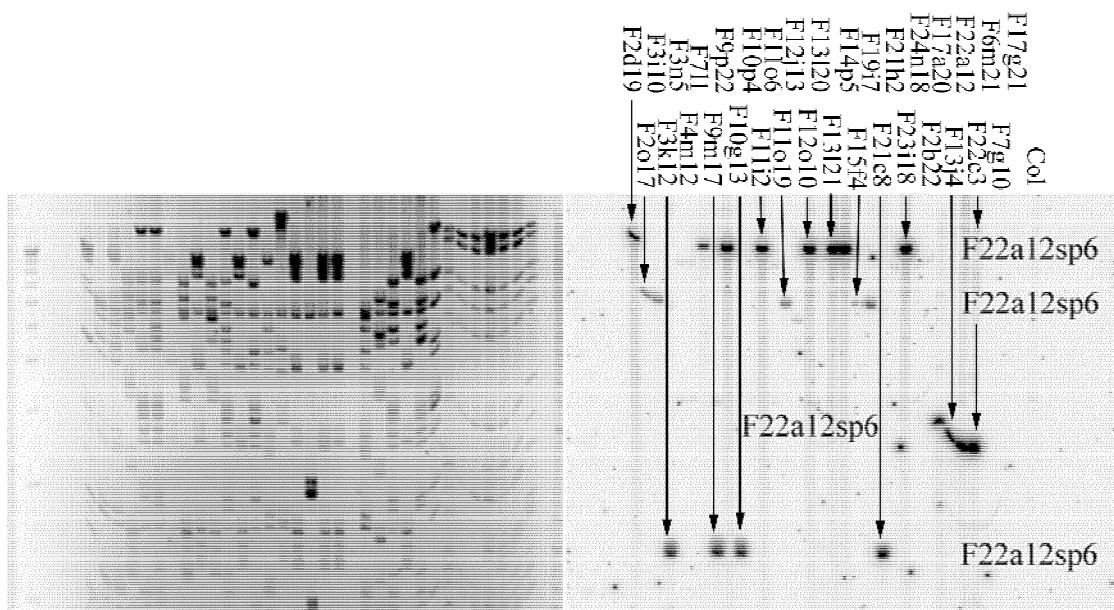
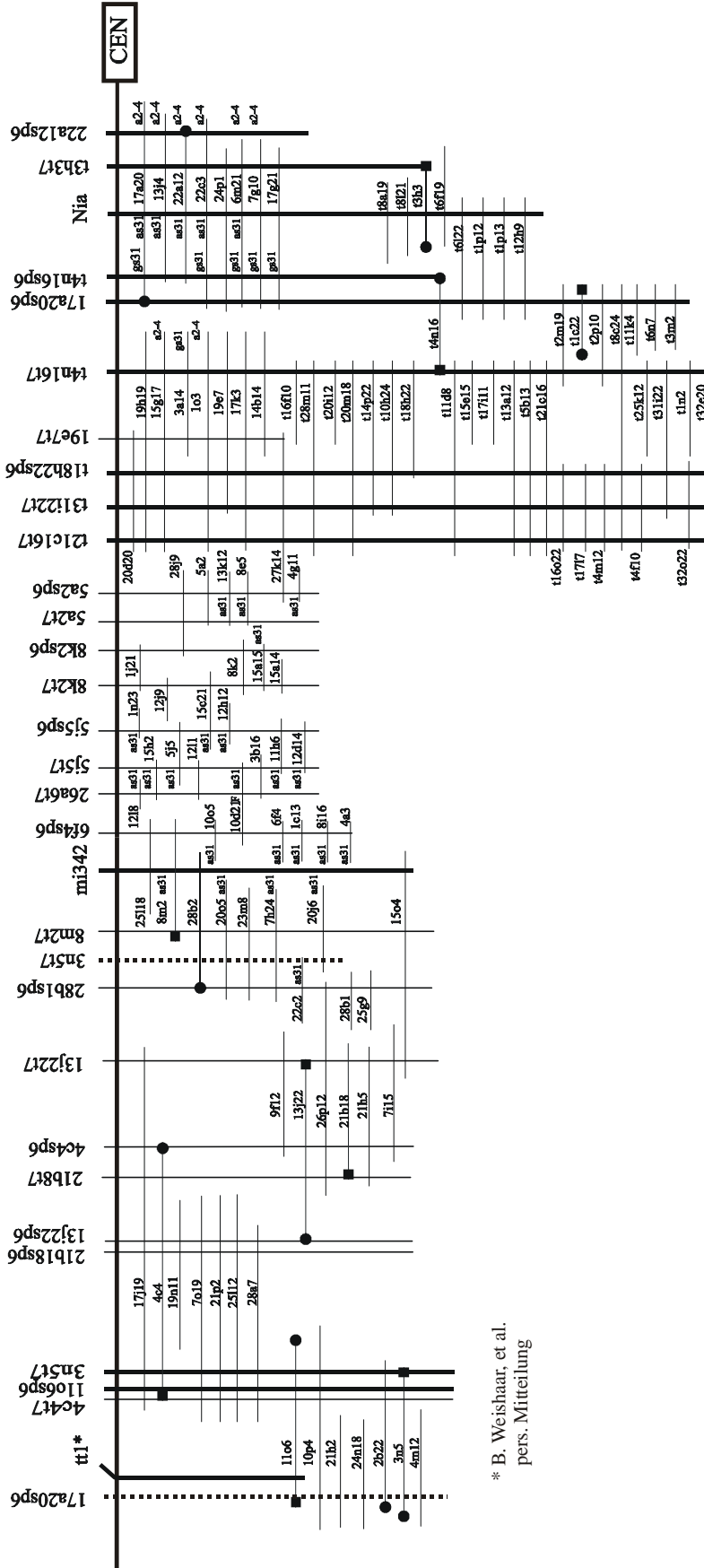


Abb. 10: finger print Analyse mit der Probe F22a12sp6. Das linke Bild stellt den finger print der BACs dar (Restriktionsverdau mit EcoRI). Das rechte Bild zeigt das Hybridisierungsergebnis mit der Probe F22a12sp6.

Die BAC-Spuren sind nur im rechten Bild benannt, die Pfeile dienen nur der leichteren Identifizierung einzelner BAC-Spuren, die Reihenfolge der BACs im linken Bild entspricht der im rechten Bild.

3.2.4 BAC contig



* B. Weishaar, et al.
pers. Mitteilung

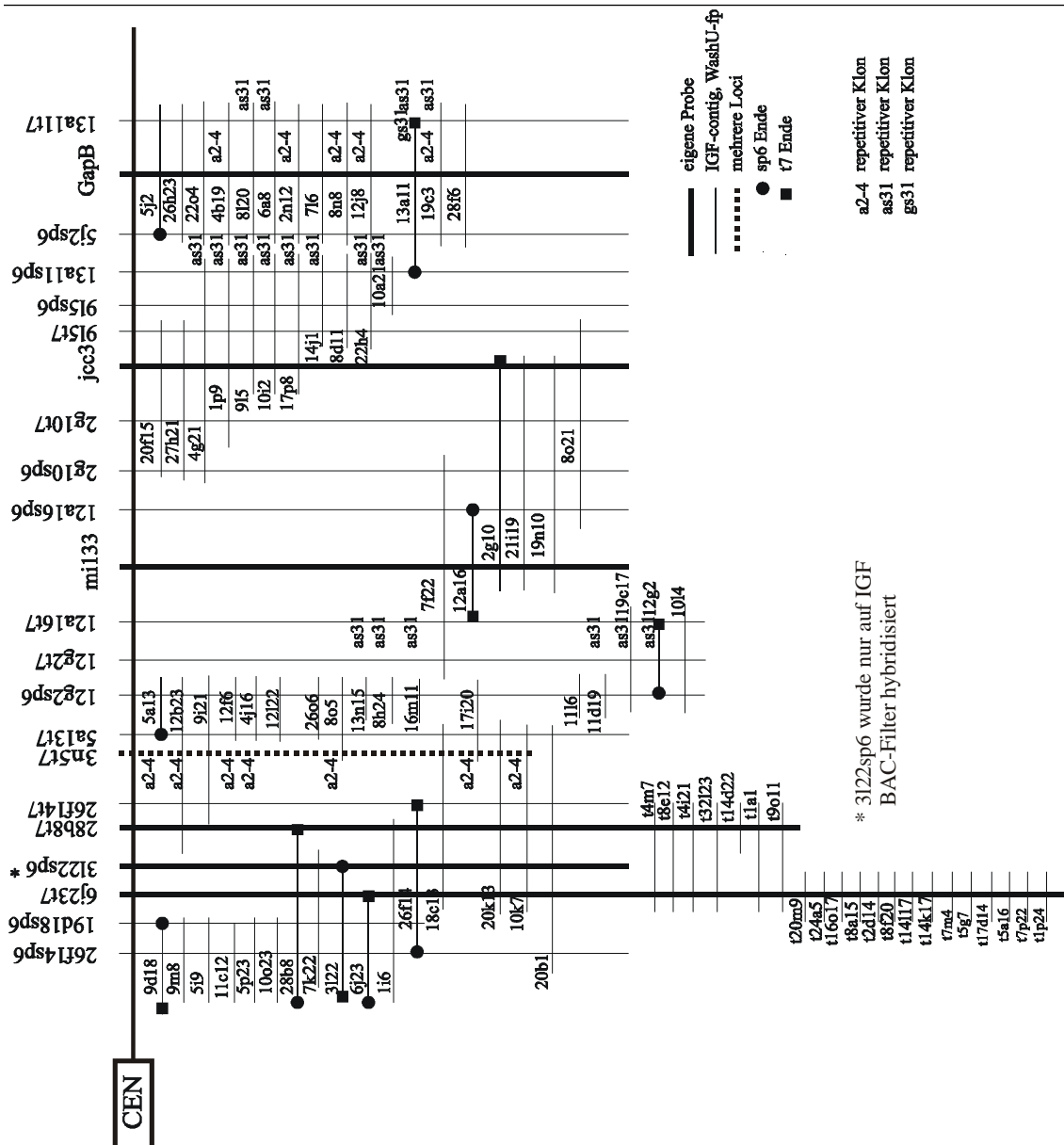


Abb. 11: BAC contig der Centromer 1-Region. Die Proben und BACs sind schematisch in ihrer physikalischen Reihenfolge dargestellt. Die Abstände der Proben und die Längen der BACs sind nicht maßstabsgetreu abgebildet.

Alle Proben, die durch eine dicke senkrechte Linie gekennzeichnet sind, wurden selbst auf IGF BAC-Filter hybridisiert. Die ermittelten BACs wurden anschließend über eine "finger print"-Analyse in den contig ein- und angeordnet. Die Anordnung der Proben wurde zusätzlich genetisch kontrolliert und bestätigt. Der Bereich um das CEN 1 konnte so sehr genau genetisch und physikalisch beschrieben werden. Der obige contig ist um das CEN 1 viel genauer als der IGF BAC contig des MPI-MP (Mozo, T., et al., 1999, http://www.mpimp-golm.mpg.de/101/mpi_mp_map/access.html), da dieser nur auf Hybridisierungsdaten beruht und in Centromernähe die meisten Proben repetitiv sind, was zu einer falschen Zuordnung der BACs führt. Mittlerweile ist diese Region bis nahe an die Grenze Centromerkern mit Hilfe von abwechselnd überlappenden BACs sequenziert (Theologis et al., 2000). Dennoch ist der obige contig sehr wertvoll, da er viele weitere BACs eindeutig positioniert. Dadurch wird verhindert, daß diese BACs irrtümlich anderen unsequenzierten Regionen zugeordnet und bei der weiteren Sequenzierung der *Arabidopsis* Centromeren berücksichtigt werden. Der erstellte BAC contig bildet die Grundlage für eine genaue physikalische Abstandsbestimmung zwischen den Proben und ermöglicht eine gute Integration der genetischen in die physikalischen Karte.

Um die komplexe Anordnung der repetitiven Elemente der pericentromerischen Region in Abhängigkeit von anderen repetitiven Elementen und Genen aufzuzeigen wurden BLAST DNA-Sequenzanalysen der BACs dieser Region durchgeführt. Dabei wurden alle BACs dieser Region mit dem BAC F28L22 (für F17G21) und dem BAC F7F22 (für F12A16) verglichen. Es zeigte sich, daß die BACs über weite Strecken verschiedene, dispers verteilte repetitive Elemente enthalten. Diese treten in mehr oder minder perfekten Kopien auf (Abb. 11aC).

Auf diese Art und Weise konnte der Bereich, der mit FISH-Experimenten der BACs F17G21 und F12A16 abgedeckt wurde "FISH-kompetente" Sequenz, physikalisch beschrieben werden (Abb. 14, 15). Man erkennt deutlich, daß auf dem linken Pericentromer die "Menge" an homologer "Fish-kompetenter" Sequenz stark abnimmt je weiter man sich vom Centromer Kern entfernt. Daraus folgt, daß die beiden BACs nach links hin nicht mehr richtig hybridisieren (Abb. 11aC).

Zwischen den beiden komplexen Bereichen der Pericentromeren fällt eine deutliche Lücke auf. FISH Experimente zeigen, daß diese Lücke hauptsächlich die repetitiven Elemente pAL1 und 106B zurückzuführen sind (Martinez-Zapater et al., 1986; Franz et al., 1998).

A: contig der CEN1-Region, alle verwendeten Marker, BACs und deren Positionen sind aufgetragen. Marker und BACs, die eingerahmt sind, wurden für FISH-Analysen verwendet. Die grau markierten BACs sind sequenziert.

B: Die Zahl der gefundenen Kandidaten für Retroelemente, Transposons, "unique", putative und unbekannte Gene ist über die Länge der entsprechenden BACs aufgetragen; dabei wurden Überlapps zweier BACs berücksichtigt (die Kandidaten im Überlapp wurden nur einem BAC zugerechnet)

C: Verteilung repetitiver Elemente der sequenzierten BACs F28L22 und F7F22 im Pericentromer des CEN1; diese liefern die FISH-Signale bei Hybridisierung mit den entsprechenden BACs. Der BAC F28L22 representiert dabei BAC F17g21 und BACF7F22 den BAC F12a16.

D: Die gemessenen Rekombinationsfrequenzen zwischen zwei Markern sind als n-fache Suppression der normalen *A. thaliana*-Rekombinationsfrequenz von 1 cM/200 kB angegeben.

3.2.5 Vergleich der genetischen mit der physikalischen Karte

Der Vergleich der genetischen mit der physikalischen Karte ermöglicht eine Aussage über die Zahl der Rekombinationsereignisse pro Sequenzabschnitt. Normalerweise geht man bei *Arabidopsis thaliana* von einem Verhältnis von 1cM pro 200 kB aus (Goodman et al., 1995). Nimmt die Zahl der Rekombinationsereignisse pro Sequenzabschnitt ab, wird die gleiche physikalische Strecke in eine kleinere genetische Distanz komprimiert. Je höher die Kompressionsrate ausfällt, desto weniger Rekombinationsereignisse finden in einem definierten physikalischen Abschnitt statt. Die Kompression ist deshalb ein Maß, um genomische Bereiche zu charakterisieren.

Für den Vergleich wurden die genetischen Daten der *gurke*-Kartierung (3.1.1 - 3.1.3, Tab. 8, 9, Abb. 3) und die physikalischen Daten des erstellten "contigs" (3.2.2, 3.2.4, Abb. 11a A) verwendet. Der "contig"-Bereich ist mittlerweile durchgehend sequenziert und die physikalischen Positionen der einzelnen Marker sind bekannt (Abb. 11a A). Aus der physikalischen Karte können die Abstände der einzelnen Marker zueinander abgelesen werden.

Bereich	genetischer Abstand (cM)	physikalischer Abstand (kB)	normiert auf 1 cM (kB)	Kompression
<i>ttl</i> - mi342	1,42	442	311	1:1,6
mi342 - 1c13-t7	0,34	161	473	1:2,4
1c13-t7 - 4a3-t7	0,03	26	867	1:4,3
4a3-t7 - Nia2	0,12	765	6375	1:31,9
CEN-1	0,08	2260	28250	1:141
6j23-t7 - gk *	0,08	>300	>3750	>1:18,7
mi133 - 6j23-t7	0,11	309	2809	1:14
GapB - mi133	0,16	240	1500	1:7,5

Tab. 15: Physikalischen Abstände; Berechnung der Kompression.

Der physikalische Abstand von *gurke* zu 6j23-t7 kann nur abgeschätzt werden (Tab. 15 *). Dabei wird von einer weiter ansteigenden Kompression ausgegangen. Der physikalische Abstand von *gk* zu 6j23-t7 beträgt mindestens 300 kb, da *gk* nicht auf dem BAC F9d18 liegt und dieser nicht mit dem BAC T4i21 überlappt (Abb. 11).

Auf der rechten Seite, zwischen *GapB* und *mi133*, findet man schon die 7,5fache Kompression und diese nimmt zum Centromer hin linear zu; zwischen *mi133* und 6j23-t7 14fach und darüber hinaus mindestens eine 18,7fache Kompression (Tab. 15, Abb. 12).

Auf der linken Seite verläuft die Kompression von *tt1* bis 1c13-t7 relativ "normal"; 1,6 und 2,4fache Kompression (Tab. 15, Abb. 12). Ab 1c13-t7 ist innerhalb eines kurzen physikalischen Intervalls von 26 kb ein signifikanter Anstieg von 2,4facher zu 31,9 facher Kompression zu verzeichnen (Tab. 15, Abb. 12).

Die Kompression nimmt von beiden Seiten her hin zum Centromer drastisch bis zum 18- bzw. 31fachen des Normalen zu. Auffällig ist dabei, daß die Kompression und damit die Suppression auf der linken Seite fast doppelt so hoch ausfällt wie auf der rechten Seite. Weiter fällt auf, daß die Rekombinationsfrequenz auf der linken Seite sprunghaft und auf der rechten linear abnimmt.

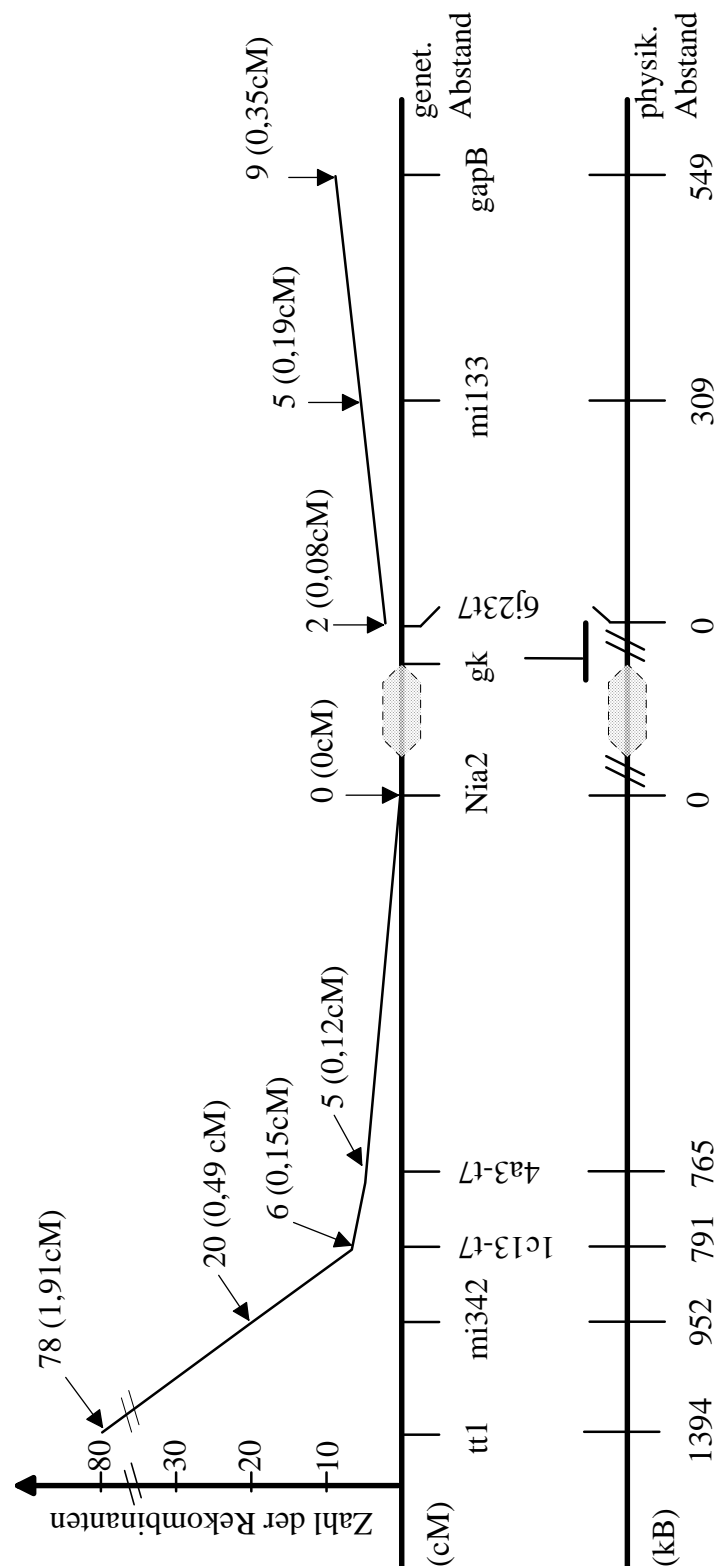


Abb. 12: Genetische-Physikalische Karte. Die genetische Karte bezieht sich auf *gurke* (Tab.8, 9, Abb. 3). Die physikalischen Abstände der Marker wurden aus dem BAC "contig" (Abb. 11a A, Tab. 15) abgelesen. Mit abnehmender Steigung der Kurve nimmt die Rekombinationsfrequenz ab und die Kompression zu.

3.3 Histologische Charakterisierung der Centromer 1-Region

Cytologische Daten der Centromer 1-Region würden es erlauben, die molekularen Ergebnisse der übergeordneten Organisation von Chromosomen zuzuordnen. So könnte die Position von molekularen Markern und die Verteilung und Häufigkeit von repetitiven Elementen am Centromer 1 mit der FISH-Technik sichtbar gemacht werden. Zudem wäre es möglich, eine Größenabschätzung der Region, basierend auf Längenmessungen, vorzunehmen; dies scheint um so wichtiger, da sonst wegen einer fehlenden physikalischen Karte für diesen Bereich keine Größenangaben gemacht werden können.

Die geringe Größe der mitotischen und meiotischen Metaphasen-Chromosomen von *Arabidopsis* erschweren die cytologische Analyse erheblich, da nur Proben die 2 Mbp auseinanderliegen aufgelöst werden können (Murata und Motoyoshi, 1995). Meiotische Pachytänstadien sind dagegen viel länger; nur die heterochromatischen Bereiche um die Centromeren sind stark kondensiert, ansonsten ist man in der Lage, die einzelnen Chromosomenarme sichtbar zu machen (Abb. 13, Ross et al., 1996). Mittlerweile konnte FISH auf Pachytän-Chromosomen von Tomate (Zhong et al., 1996 a, b) und *Arabidopsis* (Franz et al., 1998) durchgeführt werden. Die Auflösung läßt sich durch die Verwendung von meiotischen Pachytänstadien erheblich verbessern. Es konnte gezeigt werden, daß der 180 bp repeat (pAL1) ins Zentrum der Heterochromatinstrukturen der fünf Chromosomen hybridisiert (Franz et al., 1998) und die BACs F17A20 und F17G21 alle fünf Centromer-Regionen beidseitig flankieren (Fischer, 1997, Franz et al., 1998).

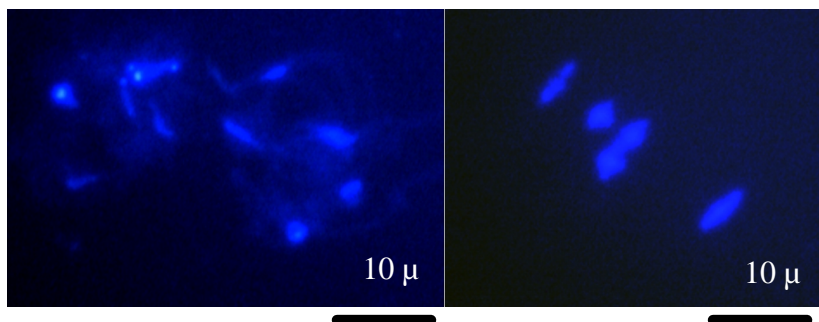


Abb. 13: Gespreitete und mit DAPI gefärbte Meiosestadien von *Arabidopsis thaliana*-Pollenmutterzellen; links Pachytän (2 Zellen), Heterochromatin des Peri-Centromers ist kondensiert, man erkennt eindeutig die fünf Centromerregionen; rechts Metaphase I, die fünf Chromosomen sind vollständig kondensiert.

3.3.1 FISH mit BACs

Um die Centromer 1-Region näher zu charakterisieren, wurde mit zwei BACs aus dieser Region je ein FISH-Experiment durchgeführt. Der BAC F17G21 liegt im linken und der BAC F12A16 im rechten Pericentromer-Bereich des Chromosom 1. Beide BACs wurden mit Texas Red gelabelt und auf Pachytänstadien hybridisiert.

Der BAC F17G21 enthält im Unterschied zum BAC F17A20, mit dem er zu 80% überlappt, nur eine Kopie von gs31, nicht jedoch von a2-4 und as31. Der BAC F12A16 enthält keines der drei Elemente (7.2.1). Die Verteilung der repetitiven Elemente dieser BACs um das CEN1 kann man der Abb. 11a C entnehmen.

Der BAC F17G21 liefert starke Signale auf allen pericentromerischen Bereichen links und rechts aller fünf Centromere (Abb. 14, deutlich im mittleren Bild, Pfeil, zu sehen). Gleichzeitig sieht man, daß der BAC nicht mit den NORs hybridisiert. Dies läßt den Schluß zu, daß der BAC ein oder mehrere repetitive Elemente enthalten muß, welche auf allen pericentromerischen Bereichen zu finden sind, nicht aber in telomerischen oder NOR-Bereichen.

Der BAC F12A16 liefert ein analoges Ergebnis; die Signale sind auch auf allen pericentromerischen Bereichen wiederzufinden, nicht aber in den NORs oder in telomerischen Bereichen (Abb. 15). Im Gegensatz zum BAC F17G21, bei dem die Signale kompakt erscheinen, sind die Signale des BACs F12A16 breiter gestreut; man erkennt sogar einzelne Unterbrechungen im Signal. Auch hier kommt man zu dem Schluß, daß der BAC ein oder mehrere repetitive Elemente enthält. Deren Verteilungsmuster unterscheidet sich jedoch eindeutig von dem Verteilungsmuster des BACs F17G21. Dies läßt sich auf eine unterschiedliche Zusammensetzung der repetitiven Elemente zurückführen.

Aus dem Vergleich der histologischen Ergebnisse mit der physikalischen Verteilung der repetitiven Elemente (Abb. 11a C, Abb. 5, Tab. 15) kann man ableiten, wie weit sich das Pericentromer erstreckt und wie die betreffenden physikalischen Regionen mit Heterochromatin besetzt sind.

Das Centromer, bestehend aus Kern- und Pericentromer Regionen definiert sich über seinen heterochromatischen Zustand. Eine DAPI-Färbung färbt die Centromeren und NORs besonders stark an da diese stark kondensierte DNA enthalten, d.h. diese Regionen stellen Heterochromatin dar (Abb. 13, 14, 15, 16). Vergleicht man die Signale der FISH Experimente der BACs F17A20, F17G21 und F12A16 mit der DAPI Färbung, erkennt man eine sehr gute Korrelation der Signale (Abb. 14, 15, 16; Fransz et al., 1998). Mit Hilfe der Sequenzdaten (Abb. 11aA) kann man die Größe der physikalischen Strecke bestimmen, die noch FISH-Signale mit den o.g. BACs liefert. Da die FISH-Signale zum einen mit den DAPI-Signalen und zum anderen mit den FISH-kompetenten Hybridisierungssignalen zusammenfallen (Abb. 11aC) kann man die Bereiche der Pericentromeren sehr genau festlegen. Das linke Pericentromer erstreckt sich vom BAC F2C1 bis zum BAC T22A15; das rechte Pericentromer mindestens vom BAC F9d18 bis zum BAC F1I21 und darüber hinaus wahrscheinlich zu der gleichen Länge wie das linke Pericentromer.

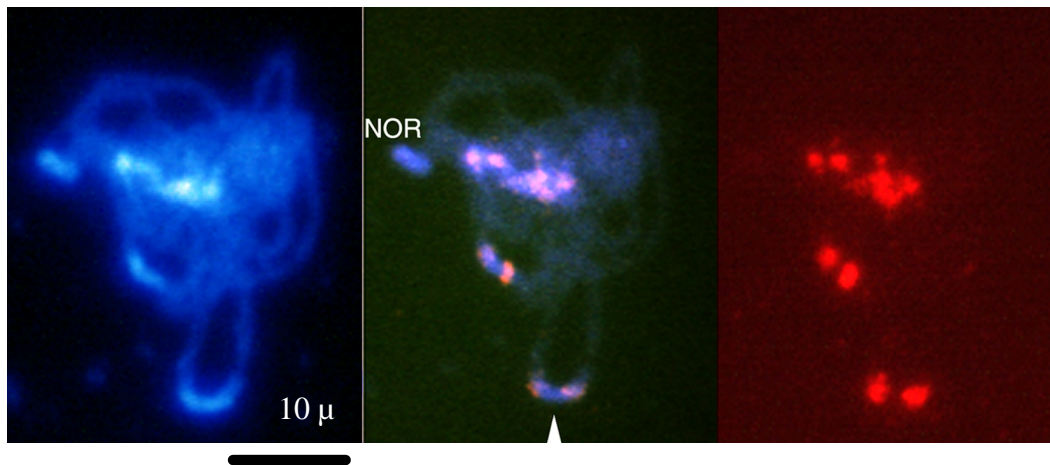


Abb. 14: FISH auf Pachytän-Chromosomen mit BAC F17G21 (Texas Red). Links ist die DAPI, in der mitte DAPI + Texas Red und rechts die Texas Red-Färbung.

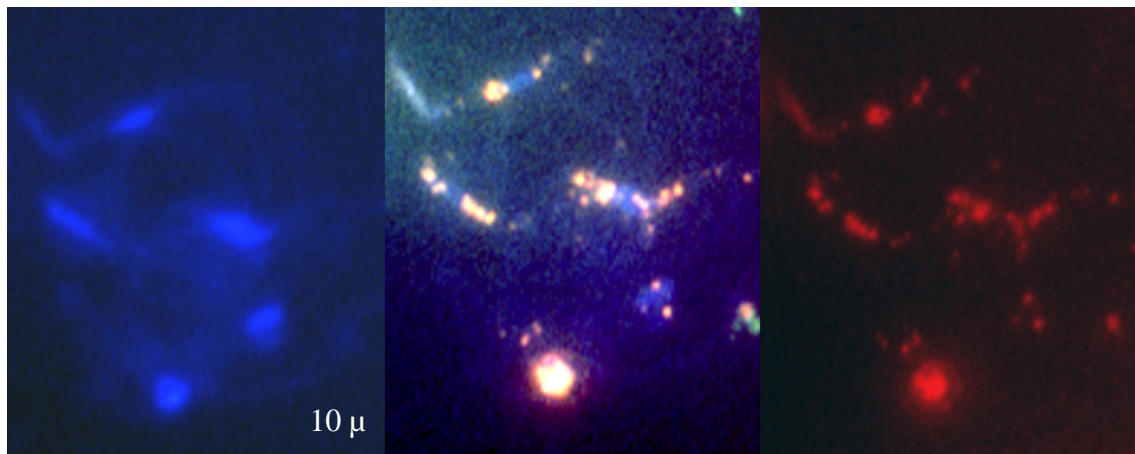


Abb. 15: FISH auf Pachytän-Chromosomen mit BAC F12A16 (Texas Red). Links ist die DAPI, in der mitte DAPI + Texas Red und rechts die Texas Red-Färbung.

3.3.2 FISH mit dem Klon a2-4

Um das Verteilungsmuster und die Häufigkeit einzelner repetitiver Elemente zu analysieren, wurden einzelne klonierte Elemente, die Teil der BACs F17G21 und F17A20 sind, gelabelt und auf Pachytän Chromosomen hybridisiert.

Der Klon a2-4 wurde mit Texas Red gelabelt und für FISH-Experimente verwendet. Man erkennt deutliche Signale von allen pericentromerischen Bereichen, nicht aber von den NORs und telomerischen Regionen. Diese sind kompakt, aber nicht so stark wie die Signale des BACs F17G21. Die Signale sind auf allen Chromosomenarmen vorhanden, jedoch nur in einer geringen Kopienzahl, sonst würde man stärkere Signale erwarten.

Aus der Signalestärke von a2-4 kann man auch schließen, daß der BAC F17G21 tatsächlich mehrere verschiedene repetitive Elemente enthält. a2-4 kommt im BAC F28L22 nur in einer Kopie vor (Abb. 20); das Signal von BAC F17G21 ist komplexer und va. länger, und dies kann nicht nur durch die Signale mehrerer verschiedener Elemente erklärt werden.

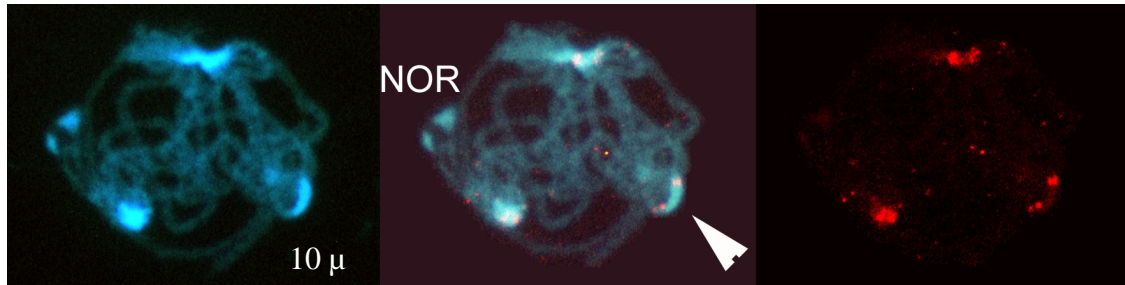


Abb. 16: FISH auf Pachytän-Chromosomen mit dem Klon a2-4 (Texas Red). Links ist die DAPI, in der mitte DAPI + Texas Red und rechts die Texas Red-Färbung.

3.4 cDNA screens

Durch das screening von λ -Bibliotheken sollten cDNAs identifiziert werden, die als Marker für die genetische und physikalische Kartierung dienen sollten. Zugleich erhält man auch schon Informationen über das Vorkommen und die Häufigkeit von Genen in der pericentromerischen Region von Chromosom 1. Um einen großen physikalischen Bereich abzudecken, wurden BACs als Hybridisierungsproben verwendet.

Die BACs stellen mit einer Größe von ca. 100 kb sehr komplexe Proben dar, die das Hybridisierungsergebnis stark beeinträchtigen (Sambrook et al., 1989). Um die Hybridisierung zu optimieren, wurden die BACs fraktioniert, und die einzelnen Fraktionen nacheinander auf die Filter hybridisiert. Für die Hybridisierungen wurden immer zwei identische Filtersätze verwendet. Die BACs wurden mit EcoRI restriktionsverdaut. Die aufgetrennten Restriktionsbanden wurden anhand ihrer Bandengröße in Fraktionen zusammengefaßt (4-5 Fraktionen).

3.4.1 Screening mit BACs F17A20 und F17G21

3.4.1.1 Gesamt-Bank

40 Platten der Gesamt λ -Bank mit ca. 2500 pfu/Platte wurden mit einer Probe aus je 100 ng BAC F17A20 und F17G21 DNA hybridisiert. Im rescreening erwiesen sich die putativ positiven Plaques als falsch Positive. Unter 100.000 Plaques der Gesamt λ -Bank wurden keine positiven Klone gefunden.

3.4.1.2 Grün-Bank

120 Platten der Grün λ -Bank mit ca. 2500 pfu/Platte wurden mit einer Probe aus je 100 ng BAC F17A20 und F17G21 DNA hybridisiert. Im rescreening erwiesen sich die putativ positiven Plaques als falsch Positive. Unter 300.000 Plaques der Grün λ -Bank wurden keine positiven Klone gefunden.

Mit der Grün λ -Bank wurden weitere 48 Platten mit einer Plaquezahl von ca. 5000 pfu gescreent. Zur Hybridisierung der Filter wurde der BAC F17A20 (Fraktionen a1-5) und Nia2 verwendet. Unter ca. 240.000 Plaques wurden mit dem BAC F17A20 239 positiven Plaques gefunden. 189 dieser Plaques hybridisieren auch mit Nia2 (Kontrolle). Die restlichen 50 positiven Signale konnten im rescreen nicht wiedergefunden werden. Mit dem BAC F17A20 lassen sich aus der Grün λ -Bank positive Klone ziehen. Diese repräsentieren alle Nia2.

3.4.1.3 Blüten-Bank

50 Platten der Blüten λ -Bank mit ca. 5000 pfu/Platte wurden mit dem BAC 17A20 (Fraktionen a1-5) und Nia2 hybridisiert. Unter ca. 250.000 Plaques wurden 97 Positive gefunden. 5 Plaques hybridisieren auch mit Nia2 (Kontrolle). Von den verbleibenden 92 Plaques wurden 71 gepickt und 64 rescreent. Im rescreening konnten keine positiven Signale gefunden werden. Aus der Blüten λ -Bank lassen sich mit dem BAC 17A20 auch nur Nia2 Klone ziehen.

Nia2 ist in der Grün λ -Bank deutlich höher als in der Blüten λ -Bank vertreten (189/240.000 bzw. 5/250.000 Klone).

Auf dem BAC 17A20 findet man außer Nia2 keine weiteren Gene bzw. diese sind so gering exprimiert, daß sie in den getesteten cDNA-Bibliotheken nicht oder unterrepräsentiert sind.

3.4.2 Screening mit BAC F9D18, F6J23 und F3L22

3.4.2.1 Grün-Bank

Mit der Grün λ -Bank wurden 100 Platten mit je 0,1 μ l einer $6,9 \cdot 10^7$ pfu/ml Phagensuspension ausplattiert. Nach der Inkubation erhielt man pro Platte eine Plaquedichte von ca. 7000 pfu/Platte.

Davon wurden 50 Platten mit ca. 350.000 Plaques mit dem BAC F9D18 (Fraktionen d1-4) gescreent. Es wurden keine positiven Plaques gefunden. Die übrigen 50 Platten wurden mit den BACs F6J23 (Fraktionen j1-4) und F3L22 (Fraktion l1) gescreent. Mit diesen Proben wurden auch keine positiven Plaques gefunden.

3.4.2.2 Blüten-Bank

Mit der Blüten λ -Bank wurden 48 Platten mit einer Plaquezahl von ca. 5000 pfu gescreent. Zur Hybridisierung der Filter wurden die BACs F6J23 (Fraktionen j1-4) und F3L22 (Fraktion 11) verwendet. Unter ca. 240.000 Plaques wurden keine positiven Plaques gefunden.

Auf den BACs F9D18, F6J23 und F3L22 findet man keine Gene bzw. diese sind so gering exprimiert, daß sie in den getesteten cDNA-Bibliotheken nicht oder unterrepräsentiert sind.

3.5 Northern-Analysen

Die Northern-Analysen sollten zeigen, wie häufig Gene im Pericentromer vorkommen bzw. transkribiert werden. Zusätzlich würde man gefundene Transkripte als Marker für die physikalische und genetische Kartierung verwenden können.

Für die Northern-Analysen wurde mRNA isoliert; aus Schoten und grünen Teilen von Ler-Pflanzen, aus grünen Teilen von WS-Pflanzen. Die mRNA wurde zusammen mit einem RNA-Längenmarker (Gibco LTI 0,24, 1,35, 2,37, 4,4, 7,46 und 9,49 kb) auf einem Agarosegel aufgetrennt. Die Laufstrecken der Längenmarkerbanden werden verwendet, um eine Eichkurve aufzustellen, mit deren Hilfe man die Größen der hybridisierten Banden ermittelt.

Hybridisiert man mit den 4 Fraktionen (d1-d4) des IGF BACs 9D18, erhält man 3 Hybridisierungssignale. Die entsprechenden Banden haben Größen von 3,3, 1,55 und 1,2 kb. Hybridisiert man die 4 Fraktionen einzeln, stellt man fest, daß nur die d4-Fraktion Hybridisierungssignale liefert (Abb. 17). Man erhält in der Schoten mRNA-Spur deutlich stärkere Signale als in der Grün mRNA-Spur. Dies deutet auf eine höhere Abundanz der gefundenen Transkripte in Schoten hin. Auf beide Spuren (Schote, Grün) wurde gleichviel mRNA aufgetragen. Die Kontrolle mit gapB zeigt stärkere Hybridisierungssignale in der Grün mRNA-Spur (Abb. 17).

Hybridisierungen mit den Fraktionen j1-j4 des IGF BACs 6F23 und der Fraktion I1 des IGF BACs 3L22 lieferten keine Hybridisierungssignale.

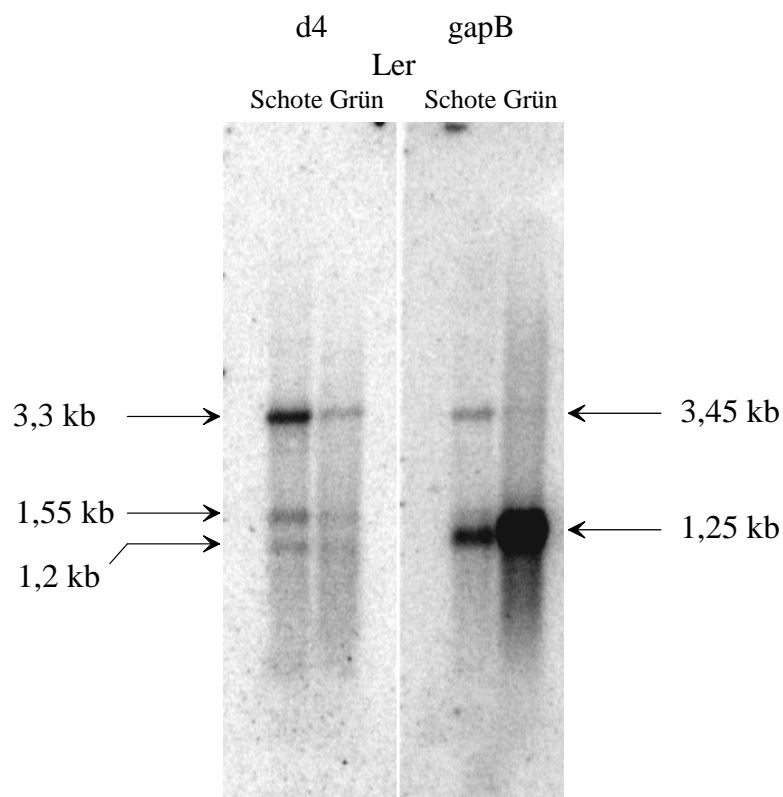


Abb. 17: Northern: Fraktion d4 des IGF BACs 9D18 und gapB hybridisiert auf mRNA aus Schoten und Grünteilen von Ler-Pflanzen.

Da der IGF BAC 9D18 viele putative Retroelemente enthält (3.6.2), wurde überprüft, ob die Signale nicht von diesen Elementen herrühren. Da sich die Retroelemente in 4 grobe Klassen unterteilen lassen (ORF1-Protein, hypothetisches Protein 1, Integrase und reverse Transcriptase), wurden die PCR-Produkte (r1-r4) jeweils eines typischen Vertreters gepoolt und für die Hybridisierung der Northernblots verwendet (Abb. 21). Dabei erhält man Signalbanden der Größen 3,3 kb, 1,55 kb und 1,2 kb (Abb. 18). Die erhaltenen Bandengrößen entsprechen den Bandengrößen, die man bei der Hybridisierung der d4 Fraktion des IGF BACs 9D18 erhält (Abb. 18).

Daraus kann man schließen, daß die Signale von Retroelement-Transkripten stammen und nicht von Genen. Dies bedeutet, daß man in der Strecke von ca. 100 kb kein abundant oder normal transkribierendes Gen findet.

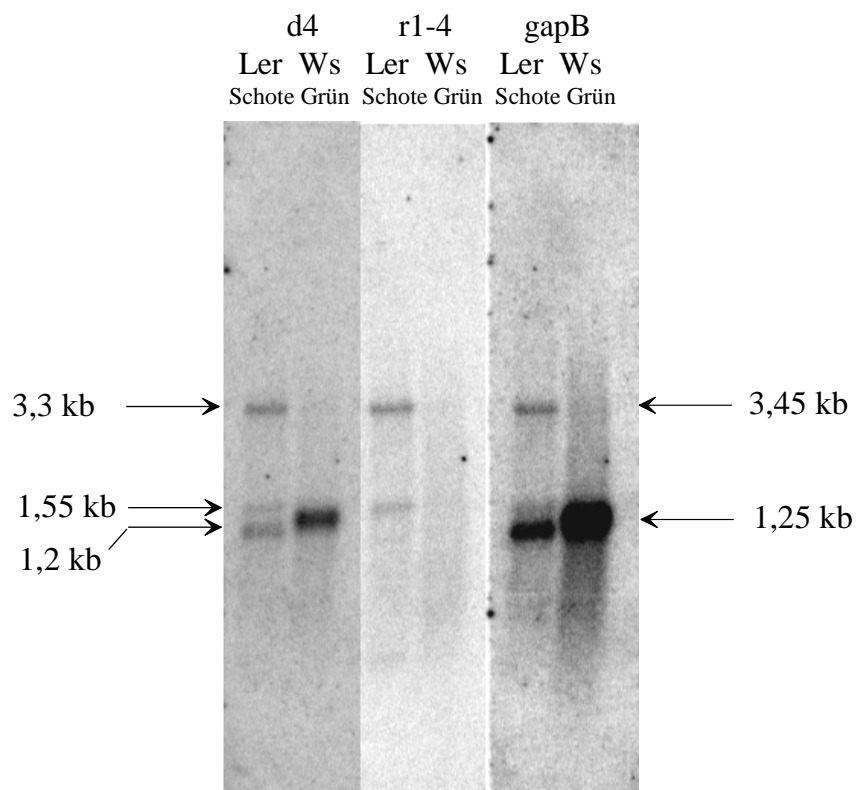


Abb. 18: Northern: Fraktion d4 des IGF BACs 9D18, PCR-Produkte von 4 putativen Retroelementen des IGF BACs 9D18 und gapB hybridisiert auf mRNA aus Schoten von Ler-Pflanzen und Grünteilen von Ws-Pflanzen.

Die einzelnen Proben wurden jeweils auf einen neuen Filterstreifen eines Filtersets, bestehend aus drei identischen Filterstreifen, hybridisiert. Damit sollte vermieden werden, daß "alte" Signale die neuen Hybridisierungsergebnisse verfälschen.

Hybridisiert man die Proben d4 und r1-r4 auf genomische Col-DNA, die mit EcoRI restriktionsverdaut wurde, zeigt es sich, daß diese, verglichen mit gapB, in höheren Kopienzahlen vorkommen (Abb. 19).

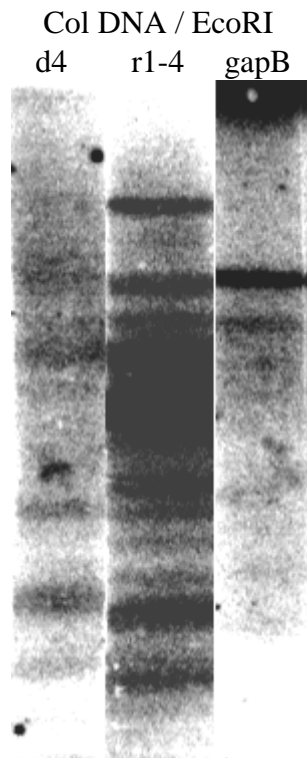


Abb. 19: Southern: Fraktion d4 des IGF BACs 9D18, PCR-Produkte von 4 putativen Retroelementen des IGF BACs 9D18 und gapB hybridisiert auf Col DNA/EcoRI.

3.6 Analyse sequenzierter BACs (aus dem contig)

Die beiden centromernahen BACs F28I22 und F9d18 wurden anhand ihrer Sequenz auf Gene analysiert mit dem Ziel, etwaige Gene als Marker für die genetische und physikalische Kartierung zu verwenden. Auch würde man über die Zahl und Art der Gene Einblick in die chromosomale Struktur im Pericentromer erhalten.

Dabei wurden sequenzierte BACs vom Chromosom 1 mit GENSCAN auf mögliche Gene analysiert. Die gefundenen Gene wurden mit Gen- und Proteindatenbanken auf Homologien verglichen (BLAST N, X; NCBI, TAIR). Dabei ließen sich die gefundenen Gene in 4 Klassen einteilen; mit repetitivem, unique, putativem und keinem Genprodukt.

3.6.1 BAC F28L22

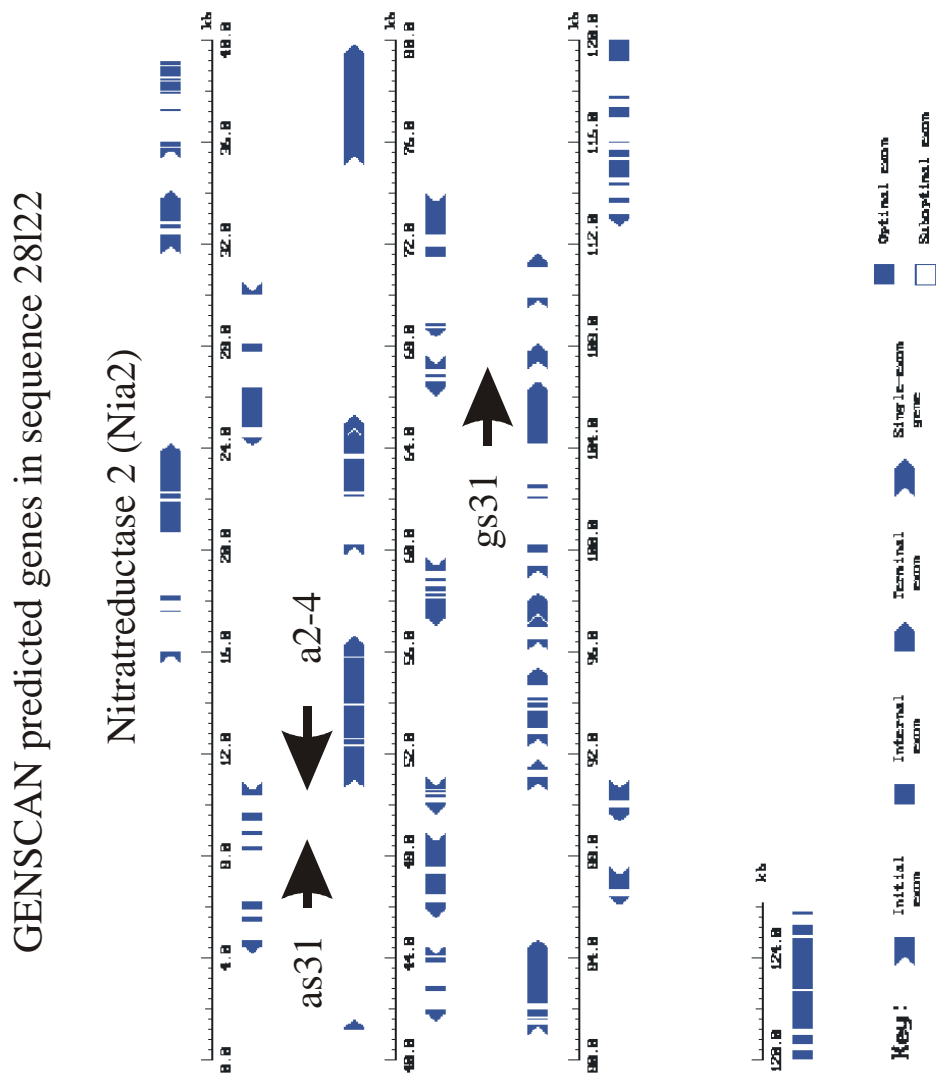


Abb. 20: BAC F28L22, mit GENSCAN vorhergesagte Gene.

Eine Liste der vorhergesagten Gene und Exonanordnung sowie die entsprechenden BLAST X-Homologien finden sich im Anhang (7.4.1)

Die BLAST X-Homologien der Klone as31, a2-4 und gs31 sind in Tab. 13 (3.2.1.3) aufgeführt. Die Pfeile zeigen die Lage der drei Klone im BAC F28L22 an; Nia2 (13368-24131 bp), as31 (7303-9745), a2-4 (14405-11776) und gs31 (113173-115355).

3.6.2 BAC F9D18

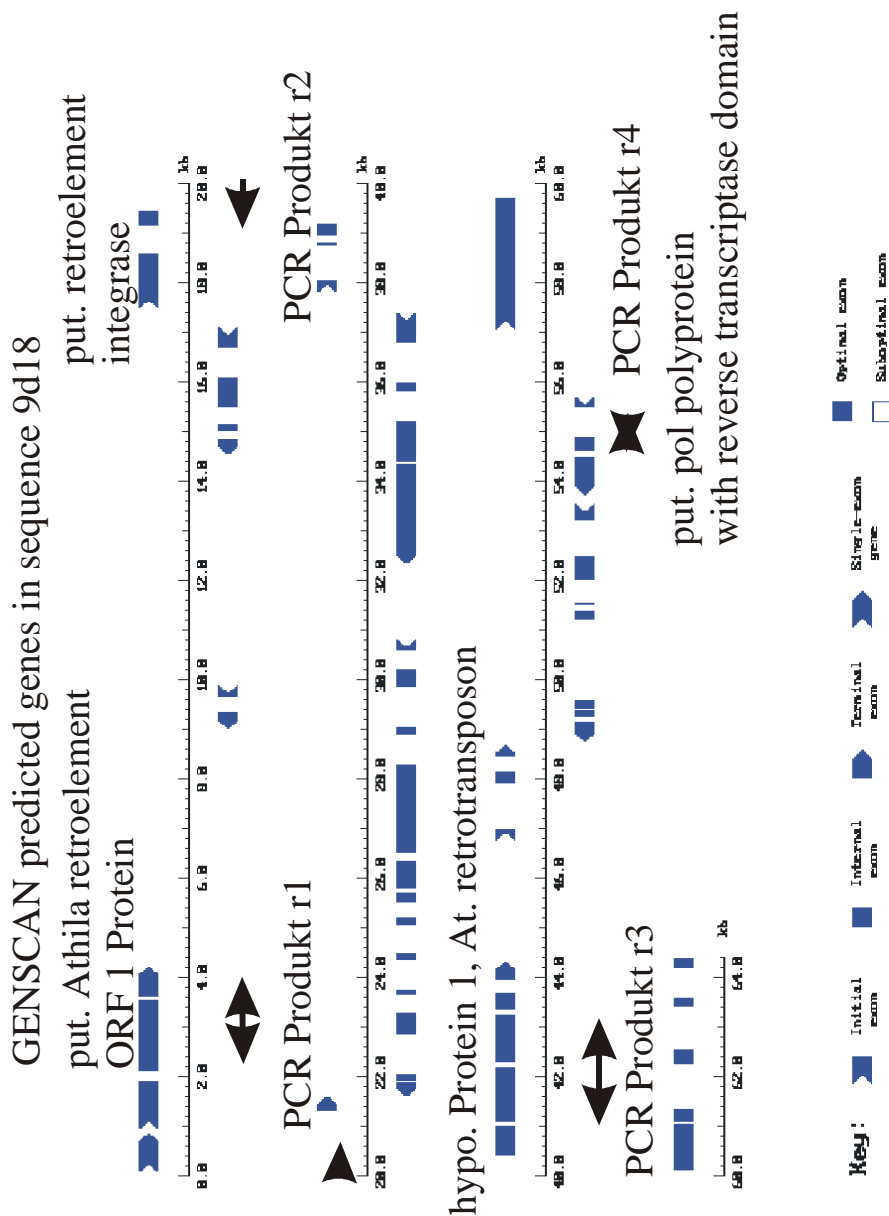


Abb. 21: BAC F9D18, mit GENSCAN vorhergesagte Gene.

Eine Liste der vorhergesagten Gene und Exonanordnung sowie die entsprechenden BLAST X-Homologien finden sich im Anhang (7.4.2)

Die Doppelpfeile stehen für PCR-Produkte, welche jeweils eine Retroelementklasse repräsentieren (vgl. 3.5 und 2.18); r1 (2304-3969 bp), r2 (19112-20791 bp), r3 (41076-43207 bp) und r4 (54612-55471 bp).

3.6.3 weitere BACs auf dem 1 Chromosom

Neben den BACs F28L22 und F9D18 wurden weitere BACs auf mögliche Gene (GENSCAN) und Proteine (BLASTX) analysiert.

BAC	ORFs	repetitiv		putativ	unique	unknown	physikalisc Position
		Trans- posons	Retro- elemente				
F17f8	20	1	1	13	5	0	UFO
F23m19	12	1		3	4	4	mi63
F21h2	20	5	2	4	2	7	
F11o6	24	3	5	8	5	3	tt1
T9i1	19	2	3	8	2	4	
F12a4	16		3	10	1	2	
F15o4	29	4	2	14	5	4	mi342
F14d7	12	2	4	4		2	
F10o5	17		4	7	4	2	
T22a15	14	1	6	5		2	
F5j5	15	3	5	5	1	1	
F15c21	17	6	3	4	3	1	
F7f23	15	1	4	4	3	3	
F28j9	21	2	5	1	1	12	
T15p17	22	1	5	8	1	7	
T32o22	9	1	5	3			
F1o3	16	1	8	3	1	3	
T32e20	24	5	7	9	2	1	
F28l22	28	2	11	3	3	9	Nia2
F9d18	13		10	2	1	0	
T4i21	18	2	9	4		3	
F5a13	17	5	6	4	1	1	
F7f22	13	2	5	5	0	1	mi133
T8d8	10	5		5			
F8d11	20	7	3	8	1	1	
F13a11	17	3	3	8	1	2	
F2h10	21	1	5	12		3	
F1i21	26	2	7	3	5	9	GapB
F28h19	27	3	8	2	4	10	mi72
F13f21	27	1		4	10	12	mi19
F14j16	19			5	14	0	mi103b

Tab. 16: Sequenzierte BACs auf Chromosom 1 (TAIR, <http://www.arabidopsis.org>). Anteil von bekannten repetitiven und unique Genprodukten an den mit GENSCAN gefundenen möglichen Proteinen. Der dunkelgraue Bereich stellt die Centromer-nächsten BACs dar, der hellgraue Bereich die Pericentromer-nahen BACs. Die gefundenen Gene wurden anhand von BLAST-Vergleichen in die angegebenen Gruppen eingeteilt: unique bekanntes Gen bzw. Funktion, 0 - $\leq 10^*E^{-50}$; putativ $\leq 10^*E^{-3}$; unbekannt $> 10^*E^{-3}$

4 Diskussion

Obwohl fast das gesamte Genom von *Arabidopsis* mittlerweile sequenziert und annotiert ist (The Arabidopsis Genome Initiative, 2000), stellen sich die Centromer-Regionen weiter als relativ weiße Flecken auf der genetischen und physikalischen Karte von *Arabidopsis* dar (ibid.). Repetitive Elemente behindern die Erstellung eines durchgehenden contigs und die Sequenzierung. Fehlende geeignete genetische und molekulare Marker erschweren es, die Lücken der genetischen und v.a. der physikalischen Karte zu schließen. Um BAC contigs vom *Arabidopsis*-Genom zu erstellen, wurden zwei Ansätze verfolgt. Zum einen wurde ein contig anhand von BAC End-Hybridisierungen auf BAC-Filter erstellt (Mozo et al., 1998-2; 1999). Zum anderen wurde eine Datenbank aufgebaut, die Daten von BAC Fingerprint-Analysen enthält (<http://genome.wustl.edu/gsc/arab/arabidopsis.html>; Marra et al., in the press (Genome Research)). Beide Methoden führten zu großen, zusammenhängenden contigs die sequenziert worden sind (The Arabidopsis Genome Initiative, 2000; Theologis et al., 2000). Viele weitere unsequenzierte BACs sind in diese contigs mit einbezogen.

In den Centromer-Regionen jedoch müssen diese contigs sehr vorsichtig betrachtet werden, da aufgrund zahlreicher repetitiver Elemente falsch positive Hybridisierungsergebnisse und Fingerprintanalysen falsche Verknüpfungen angeben und somit zu falschen contigs zum Centromer-Kern hin führen könnten. Trotz dieser Probleme sind diese Methoden, zusammen mit einer sehr präzisen genetischen Kartierung, die einzige Möglichkeit, einen contig zum Centromer hin weiterzuführen und evtl. ein Centromer mit einem contig zu überspannen. In dieser Arbeit wurden deshalb bestehende Daten aus der Centromer 1-Region als Ausgangspunkt benutzt, um anhand eigener Hybridisierungs- und Fingerprintanalysen (3.2.1.2; 3.2.3.2), zusammen mit der genetischen Kartierung der Region (3.1), einen sehr genauen BAC contig zu erstellen (3.2.4). Dabei wurden Abweichungen zu den bisher bekannten contigs festgestellt. Insbesondere sollten in diesem Bereich "unique" Proben ermittelt werden, um im Centromerbereich rekombinationsfrequenzen erheben zu können.

4.1 Vergleich des erstellten contigs (3.2.4) mit bestehenden contigs

Der BAC contig (3.2.4) endet auf der linken Seite des CEN 1 mit der Probe F22a12sp6 und den BACs F17G21, F7G10, F6M21, F24P1, F22C3, F22A12, F13J4, F28L22 und F17A20. Dabei zeigte sich, daß die Probe F22a12sp6 nicht single copy ist (3.2.3) und zu falschen BACs verzweigt und somit zu einem falschen contig führen würde. Alle weiteren BACs und Hybridisierungsproben, die im T. Altmann contig weiter rechts davon liegen (7.5.1), müssen als falsch positiv angesehen werden. Die Probe F22a12sp6 erkennt vier BAC-Hybridisierungsgruppen, und nur eine davon liegt tatsächlich links vom CEN 1 (3.2.3). Die anderen BACs gehören anderen Loci im Genom an und dürfen nicht im contig links vom CEN 1 aufgeführt werden, da sie den contig verfälschen. Da keines der anderen Enden der oben genannten BACs sich als Marker eignete, konnte der contig nicht weiter zum Centromer verlängert werden und bricht an dieser Stelle ab.

Die Zahl der positiven BACs im contig nimmt in unmittelbarer Nähe zum CEN 1 hin stark zu. Dies kann man durch die Repetitivität der verwendeten Proben erklären. Diese Proben zeigen nicht nur die echt positiven BACs, sondern auch falsch positive BACs

von anderen Loci an, so daß der contig an dieser Stelle als falsch angesehen werden muß.

Das gleiche Bild findet man auch auf der rechten Seite vom CEN 1. Nahe des CEN 1 nimmt die Zahl der positiven BACs sprunghaft zu. Dies läßt sich nur durch falsch positive BACs erklären. Der CEN 1 nahe contig auf der rechten Seite muß daher auch als falsch angesehen werden. F6j23t7 und F9d18sp6 sind auf der rechten Seite die nächsten Proben zum CEN 1. Das BAC-Ende F9d18t7 stellte sich als hoch repetitiv heraus. Damit ist ein großer Teil der BACs, die mit dieser Probe hybridisieren, falsch positiv und verfälscht den contig an dieser Stelle (7.5.2). Der contig auf der rechten Seite des CEN 1 wird vom BAC F9D18 abgeschlossen (3.2.4). Es konnten keine geeigneten Proben gefunden, um den contig sicher in Richtung CEN 1 hin zu verlängern.

Das Aufstellen eines contigs um die CEN 1-Region zeigte sich als sehr aufwendig und voller Fallstricke, da viele der verwendeten Proben sich als repetitiv erwiesen. Diese repetitiven Proben konnten allerdings nur in der Kombination der BAC-Filter Hybridisierungen und Hybridisierung der Proben auf BAC-Fingerprint-Analysen ermittelt werden (3.2.3). Wendet man nur eine der Methoden an, um einen contig in einer Centromer Region zu erstellen, wird man zwangsläufig einen falschen contig aufbauen, je näher man dem Centromer kommt.

4.2 Vergleich der genetischen mit der physikalischen Karte Lage des Heterochromatins

Eine der wichtigen Fragen dieser Arbeit war die Frage nach der Größe der CEN 1-Region. Diese Frage läßt sich zunächst nur auf genetischer Ebene beantworten. Die CEN 1-Region läßt sich auf ein Intervall zwischen den Markern mi342 und GapB eingrenzen, wobei diese Region das Centromer und das Pericentromer umfaßt. Die Marker wurden beide zu *gurke* kartiert (3.1.3). Die Summe ihrer Abstände zu *gurke* ergibt einen genetischen Abstand von 0,84 cM zwischen den beiden Markern. Das Centromer selbst läßt sich auf ein sehr enges Intervall von 0,08 cM zwischen den Markern Nia2 und F6j23t7 eingrenzen.

Das physikalische Intervall der CEN 1-Region kann für die beiden pericentromerischen Bereiche um den Kern des Centromers anhand der Sequenzdaten des erstellten "contigs" ermittelt werden (3.2.5). Da der "contig" das Centromer 1 nicht überspannt, können physikalische Größenbestimmungen nur bis zu den Centromer nächsten Markern durchgeführt werden. Die Strecke zwischen diesen Markern, die die Lücke der beiden "contigs" schließt, kann anhand der genetischen Information bestimmt werden. Dabei muß allerdings die Kompressionsrate berücksichtigt werden.

Die hier erhobenen FISH Daten und Sequenzstrukturdaten, dh. FISH-kompetente Bereiche für F17G21/F17A20 ermöglichen aber in Kombination mit cytologischen Messungen eine relativ genaue Abschätzung der Centromer-1 Region.

Legt man eine normale Kompression zugrunde, würde das Centromer 1 eine physikalische Strecke von ca. 16 kb belegen und die gesamte CEN 1-Region ca. 170 kb. Aus histologischen Untersuchungen wurde für die CEN 1-Region eine Gesamtlänge von 4,17 μm und für den Kernbereich eine Länge von 2,15 μm ermittelt (P. Frasz, pers. Mitteilung). Vergleicht man die physikalische Größe von FISH-kompetenten Regionen mit der Länge der entsprechenden FISH-Signale erhält man ein Verhältnis von 1,05 Mb/ μm . Dieser Wert liegt zwischen den Werten für NOR-Regionen (2,2 Mb/ μm) und dem heterochromatischen knob hK4S (0,7 Mb/ μm) auf Chromosom 4 (Frasz et al., 2000). Den Wert von 1,05 Mb/ μm muß man auch zur Größenbestimmung des Centromerkerns mindestens einsetzen. So ergeben sich folgende Centromergrößen:

		Chr1	Chr2	Chr3	Chr4	Chr5	Gesamt
CEN Kern	μm	2,15	1,52	1,33	1,46	1,88	8,34
pAL1	Mb	2,26	1,6	1,4	1,53	1,97	8,76
CEN gesamt	μm	4,17	4,14	4,0	3,38	4,20	19,89
PeriCen + KernCen	Mb	4,4	4,35	4,2	3,55	4,4	20,88

Tab. 17: Geschätzte Größe der *Arabidopsis thaliana* Centromeren.
(P. Frasz, pers. Mitteilung)

Daraus kann man schließen, daß CEN1 das größte Centromer darstellt (4,4 Mb) und den größten Centromerkern hat (2,26 Mb). CEN4 stellt das kleinste Centromer dar (3,55 Mb) und CEN3 hat den kleinsten Centromerkern (1,4 Mb). Die Größe der Arabidopsis Centromere ist mit ca. 4 Mb charakteristisch für sogenannte regionale Centromere in vielzelligen Eukaryoten (Willard, 1998; Hemleben et al., 2000; Tyler-Smith and Florida, 2000).

4.2.1 Kompression bzw. Rekombinationsunterdrückung

Normalerweise entspricht 1 cM ca. 200 kb des *Arabidopsis*-Genoms (Goodman et al., 1995) und alle ca. 5 kb findet man durchschnittlich ein Gen. Man könnte dementsprechend in einem Intervall von 200 kb ca. 40 Gene erwarten.

Im Laufe der durchgeführten Untersuchungen zeigte sich, daß mit zunehmender Nähe zum Centromer 1 eine steigende Unterdrückung der Rekombinationsfrequenz zu beobachten war. Dies bedeutet, verglichen mit dem physikalischen Abstand, eine Kompression des genetischen Abstands zwischen zwei Markern.

Vergleicht man die genetische Karte (3.1.3) mit dem erstellten BAC contig (3.2.4), so stellt man eine erhöhte Kompressionsrate fest, die mit zunehmender Centromernähe dramatisch ansteigt (3.2.5). Auf der rechten Seite findet man bei GapB schon eine Kompression von 1:7,5, die über 1:14 (mi133) auf 1:18,7 (F6j23t7) steigt. Dies entspricht 3750 kb pro 1 cM (Tab. 15, Abb. 12). Auf der linken Seite findet man zunächst eine normale Kompression von 1:1,6 (tt1), die linear auf 1:2,4 (mi342) und 1:4,3 (F1c13t7) steigt (Abb. 12). Innerhalb des sehr kurzen Intervalls von 26 kb

zwischen den Markern F1c13t7 und F4a3t7 findet ein sehr drastischer Sprung der Kompression auf 1:31,9 (F4a3t7) statt. Diese Kompression entspricht mehr als 6375 kb pro 1 cM. Legt man mindestens diese Kompressionsrate für das Centromer zugrunde, würden 0,08 cM (Centromerintervall) 500 kb entsprechen. Diesen Wert muß man als Mindestwert betrachten, da die Kompression im Centromer nicht gleich bleibt, sondern zunimmt. Die Ergebnisse lassen sogar die Vermutung zu, daß die Rekombination im Kern des Centromers vollständig unterdrückt wird. Für die CEN 1-Region würde sich somit eine Mindestgröße von 3760 kb ergeben (952 kb mi342-CEN + 2260 CEN-Kern + 549 kb CEN-GAPB). Der Unterschied zwischen der physikalisch und der histologisch ermittelten Größe des CEN 1-Bereiches läßt sich vor allem auf den rechten Pericentromerbereich zurückführen. Dieser wurde nur bis zum Marker Gap B untersucht. Die Hybridisierungsergebnisse mit den Klonen a2-4, as31 und gs31 zeigen aber, daß sich der heterochromatische Bereich links vom BAC T22A15 bis zum BAC F2C1 und rechts weiter bis zu dem Marker mi72 erstreckt. Das linke Pericentromer erstreckt sich somit über eine Länge von 1070 kb. Fish und Sequenzdaten zeigen, daß das rechte Pericentromer die gleiche Länge erreicht. Daraus kann man für das CEN1 eine physikalische Größe von 4,4 Mb ableiten, sodaß sich physikalische und histologische Größe entsprechen (4.2).

4.2.2 Kompressionszunahme zwischen F1c13t7 und F4a3t7

Betrachtet man sich den Vergleich der genetischen mit der physikalischen Karte (3.2.5), so fällt links vom CEN 1 die drastische Zunahme der Kompression zwischen den beiden Markern F1c13t7 und F4a3t7 auf. Die Zunahme von 1:4,3 auf 1:31,9 ist so signifikant, daß sich die Frage nach der Ursache für eine so starke Rekombinationsunterdrückung stellt. Der BAC F14D7, der die beiden BACs F1C13 und F4a3 teilweise überlappt, wurde schon durchsequenziert. BLAST-Analysen zeigten Homologien zu repetitiven Elementen, wie sie auch bei anderen BACs in dieser Region gefunden wurden (3.6) z.B. ORF1-Protein, hypothetisches Protein 1, Integrase und Reverse Transcriptase. Die Abnahme der Rekombinationsereignisse könnte somit durch repetitive Elemente bedingt sein. Es wäre auch denkbar, daß Rekombinationsereignisse im Heterochromatin grundsätzlich unterdrückt werden (Copenhaver et al., 1999). Die repetitiven Elemente wären dann evtl. ein Anzeiger für das Heterochromatin. Dies würde mit dem Ergebnis der GENSCAN-Analyse von sequenzierten BACs im CEN 1 korrelieren (3.6.3). Es zeigt sich, daß BACs in CEN-Nähe mehr repetitive Elemente enthalten als BACs, die weiter entfernt sind. Im Bereich des Markers mi342 nimmt die Zahl der repetitiven Elemente ein "normales" Maß an, so daß ein Zusammenhang zwischen der Zahl der repetitiven Elemente und der Rekombinationsunterdrückung erkennbar wird.

In Abb. 11a B ist auch eine Abnahme der Retroelemente Richtung mi342 zu erkennen. Betrachtet man dagegen die FISH-kompetenten repetitiven Elemente (Abb. 11a C), die das Heterochromatin mitbeschreiben, erkennt man, daß deren gehäuftes Auftreten mit den BACs F1C13 und F4A3 zusammenfällt. Dies führt unweigerlich zu der Schlußfolgerung, daß das Heterochromatin für die drastische Rekombinationsunterdrückung (mit)verantwortlich ist.

Dieses Phänomen des drastischen Rekombinationsabfalls wird rechts vom CEN 1 nicht beobachtet. Dies liegt wahrscheinlich an der Tatsache, daß der contig und die genetische

Karte nicht über GapB hinaus erstellt wurden, so daß Daten diesbezüglich fehlen. Allerdings erkennt man bei der GENSCAN-Analyse von sequenzierten BACs rechts vom CEN 1 eine signifikante Abnahme der repetitiven Elemente im Bereich des Markers mi72. Die Kompression nimmt vom CEN1 zu GAPB hin langsam, linear ab; vergleichbar mit der linken Seite vom CEN1 zu 4A3-t7 hin (Abb. 12). Daraus könnte man schließen, daß man im Bereich von mi72 auch eine drastische Rekombinationsunterdrückung finden würde.

4.2.3 Lage und Wirkung des Heterochromatins

Betrachtet man nur die DAPI-Färbung der Arabidopsis CEN1 Region, erhält man nur ein undifferenziertes Bild des Heterochromatins (Abb. 13-16). FISH Experimente mit den BACs F17G21, F12A16 bzw. dem Klon a2-4 zeigen ein viel detaillierteres Bild (Abb. 14-16). Die Unterschiede die dabei zu Tage treten erinnern an *Drosophila* α - bzw. β -Heterochromatin. Bei der Diskussion über DNA-Bindestellen der Chromatin Proteine darf man nicht vergessen, daß Kern- und Pericentromer-Bereiche unterschiedlich aufgebaut sind (Fanti et al., 1998). Es ist nicht bekannt ob der Centromerkern und die Pericentromeren gleiche oder unterschiedliche Bindestellen für Heterochromatin-Proteine tragen. Retrotransposons die in allen centromerischen Regionen vorkommen, scheinen Heterochromatin zu induzieren (Mc Combie et al., 2000). In diesem Zusammenhang lohnt es sich die linke Grenze des CEN-1 näher zu betrachten. Hier geht ein Wandel der Sequenzzusammensetzung mit dem Übergang von Eu- nach Heterochromatin und einem enormen Rekombinationsrückgang einher (Abb. 12, Tab. 16). Es ist schon lange bekannt, daß sich Heterochromatin suppressiv auf die Genexpression und Rekombination auswirkt (Beadle, 1939; Baker, 1958).

Für das CEN-1 Heterochromatin konnte die Rekombinationsfrequenz anhand der Kartierung der *gurke* Allele bestimmt werden (Tab. 8, 9, Abb. 3). Dabei zeigt sich, daß das linke und rechte Pericentromer 1 die Rekombinationsfrequenz unterschiedlich hoch unterdrücken; 31,9-fach zu 18,7-fach (Tab.15). Da stellt sich die Frage ob die qualitativen und quantitativen Unterschiede der Sequenzzusammensetzung die Unterschiede im Heterochromatin bewirken, um so mehr da einzelne repetitive Elemente (a2-4, as31 und gs31) unterschiedlich auf dem linken und rechten Pericentromer 1 verteilt sind. Dies könnte bedeuten, daß die Sequenzzusammensetzung über das Heterochromatin die Rekombinationssuppression regelt. Besonders deutlich wird dies beim Vergleich der Pericentromeren mit dem CEN-1 Kern. Im CEN-1 Kern findet man eine 141-fache Suppression (Tab. 15). Da die Rekombinationsereignisse zwischen Nia-2 und 6j23t7 (Abb.12, Tab. 15) die zu diesem Ergebnis führen wahrscheinlich außerhalb des CEN-1 Kerns am Rand der Pericentromer-Regionen stattgefunden haben kann man davon, daß der Centromerkern die Rekombination vollständig unterdrückt

4.2.4 Vorkommen von Genen in der CEN 1-Region

Eine weitere wichtige Frage der Arbeit war, inwieweit Gene in der Nähe des Centromers zu finden sind. In der heterochromatischen, komplexen chromosomalen Struktur der

CEN 1-Region erwartet man definitionsgemäß keine oder nur sehr wenige "exotische" Gene.

Ausführliche cDNA screens (3.4) links und recht des CEN 1 ergaben bis auf das bekannte Gen *Nia-2* keine neuen Kandidaten. *Nia-2*, eine Nitratreduktase (Wilkinson et al., 1991), ist ein wichtiges Haushaltsgen. *Nia-2* kartiert sehr nahe an die linke Seite des CEN 1. Das *gurke*-Gen kartiert sehr nahe an die rechte Seite des CEN 1. *GURKE* spielt wahrscheinlich eine wesentliche Rolle bei der Cytokinin Signaltransduktion.

Offensichtlich befinden sich wichtige Gene in unmittelbarer Nachbarschaft des CEN 1, aber nur in einer sehr geringen Anzahl. Normal findet man alle 5 kb ein Gen, hier in der CEN 1-Region findet man ca. alle 100 - 200 kb ein Gen.

Eine GENSCAN-Analyse von BACs der Region (3.6) ergibt, abgesehen von vielen putativ repetitiven Elementen, keine oder nur sehr wenige putative Gene.

Es scheint sich zu bestätigen, daß in der Centromerregion nur sehr wenige Gene zu finden sind. Eine wichtige Frage, die sich dabei stellt, ist die nach der Expression und Regulation dieser Gene. Dies würde einen weiteren Schritt zum besseren Verständnis der Centromerregion beitragen.

4.3 Struktur der CEN1 Region

Das zentrale Element der *Arabidopsis*-Centromere scheint ein 180 bp langen repetitives Element (pAL 1 repeat) zu sein (Martinez-Zapater et al., 1986; Simoens et al., 1988; Richards et al., 1991; Murata et al., 1994, Maluszynska and Heslop-Harrison, 1991; Franz et al., 1998). Dieses Element erstreckt sich am CEN 1 über ca. 600 kb (4.2.1). Die Anordnung des pAL 1 repeats wird von mindestens einem weiteren Element unterbrochen (Pellisier et al., 1996; Thompson et al., 1996a,b; Franz et al., 1998). FISH-Analysen haben gezeigt, daß eine degenerierte Kopie des Athila Retroelements, Klon 106b, vereinzelt zwischen dem pAL repeat vorkommt (Franz et al., 1998). Dieser Kernbereich des Centromers wird vom Pericentromer beidseitig flankiert. Das Pericentromer zeichnet sich durch eine Vielzahl unterschiedlichster repetitiver Elemente aus. Durch FISH-Analysen mit BACs aus der Pericentromer-Region konnte gezeigt werden, daß repetitive pericentromerische Elemente alle fünf Centromeren von *Arabidopsis* flankieren. Dies gilt sowohl für die BACs F17A20 und F17G21 von dem linken CEN 1-Pericentromer (*Nia 2*) (Franz et al., 1998) als auch für den BAC F12A16 von dem rechten CEN 1- Pericentromer (*GapB*) (3.3.1). Die Hybridisierungen mit den BACs ergeben ein sehr grobes Bild, da man ihnen kaum Information über die Zusammensetzung und Verteilung der einzelnen repetitiven Elemente entnehmen kann.

Weitere Information sollten exemplarisch Subklone der untersuchten BACs liefern (3.2.1). Drei der Subklone, a2-4, as31 und gs31, wurden näher untersucht. Dabei zeigte sich, daß sich die Klone sowohl in ihrer Kopienzahl als auch ihrer Verteilung deutlich unterscheiden (3.2.1.2). Der Klon a2-4 läßt sich als wenig repetitiv einstufen, da er bei BAC- Filterhybridisierungen ca. 168 positive Signale zeigte, die sich in mindestens sieben Hybridisierungsgruppen zusammenfassen lassen (3.2.1.2). Die Klone as31 (274 Hybridisierungssignale) und gs31 (298 Hybridisierungssignale) sind wenig bis mittel repetitiv und lassen sich in mindestens neun bzw. 12 Hybridisierungsgruppen

zusammenfassen. Ähnlich der BACs F17A20 und F17G21 flankieren diese Klone auch alle fünf Centromeren von *Arabidopsis*. Allerdings lassen sie ein viel differenzierteres Bild der Verteilung solcher repeats zu. Bei einer FISH-Analyse mit dem Klon a2-4 (3.3.2) wurden wie mit den BACs F17A20, F17G21 und F12A16 (3.3.1) Centromer flankierende Signale gefunden. Im Gegensatz zu den BAC-Signalen sind die Signale vom Klon a2-4 sehr viel schärfer und lokal begrenzt (3.3.2). Dies führt soweit, daß man die Zahl der einzelnen Kopien in jedem Pericentromer bestimmen kann. Dabei zeigt sich eine Übereinstimmung der molekular bestimmten Kopienzahl mit der histologisch bestimmten Kopienzahl. Die Ermittlung von (Retro-)Transposons und Genen um CEN1 (Abb. 11a) verfeinert die Strukturanalysen noch weiter (Theologis et al., 2000; diese Arbeit). Demnach häufen sich Retro-, andere repetitive Elemente und Gene gehen verloren (4.2.4, Tab. 16, Abb. 20, 21).

Schließlich geben die Vergleichsanalysen mit den BACs F28L22 und, F7F22 zur Ermittlung FISH kompetenter Sequenzen, einen deutlichen Hinweis darauf wo (nicht funktionelle) repetitive Elemente sich besonders stark anhäufen, nämlich ab BAC T22A15 (Abb 11aC) in Richtung des Centromer Kerns.

Das CEN 1 läßt sich grob in zwei Bereiche unterteilen, den Kernbereich des Centromers, bestehend aus dem pAL 1 (180 bp repeat), und dem Klon 106B (degeneriertes Athila Retroelement) und dem flankierenden Pericentromer, bestehend aus zahlreichen Retro- bzw. repetitiven Elementen (Abb. 11a, Abb. 23). Ob der pAL 1 repeat ausschließlich für die Bildung des Kinetochors nötig ist oder ob die repetitiven Elemente auch eine Rolle spielen, ist noch völlig unklar und bedarf noch weiterer Untersuchungen.

Mit dem Klon a2-4 kann die Lage des CEN 1 endgültig sicher bestimmt werden. Die einzelnen Centromeren werden aufgrund ihres Aufbaus aus repetitiven Elementen in "contig Lücken" vermutet. Das CEN 1 kann anhand der molekularen und histologischen Daten für den Klon a2-4 genau lokalisiert werden. Die histologischen Daten zeigen, daß der Klon a2-4 das CEN 1 einrahmt. Demzufolge muß eine Lücke im contig vom Chromosom 1 gefunden werden, und die contigs links und rechts der Lücke müssen BACs enthalten, die mit dem Klon a2-4 positiv hybridisieren. Auf dem Chromosom 1 gibt es nur zwei Lücken im contig, die diese Bedingungen erfüllen (Lücke zwischen Nia2 und mi133; Lücke zwischen mi72 und mi291a, *Arabidopsis thaliana* Database, <http://www.arabidopsis.org/cgi-bin/maps>, *Arabidopsis Physical Map Close-up*). Die Lücke zwischen mi72 und mi291a stellt allerdings nur eine Lücke im BAC contig dar, die mit dem CIC YAC 11D19 überbrückt werden kann. Das CEN 1 muß also zwischen Nia2 und mi133 liegen.

Die Sequenzierung der untersuchten Klone zeigte, daß diese verschiedene putative Retroelemente darstellen könnten (3.2.1.3; a2-4 = putative Reverse Transcriptase, as31 und gs31 putatives Athila Retroelement ORF1 Protein). Auf DNA-Ebene konnten keine Homologien zwischen den einzelnen Klonen gefunden werden. Dies deutet aber zumindest darauf hin, daß im Pericentromer gehäuft Retroelemente vorkommen. Besonders Athila Retroelemente werden nur in den Pericentromeren gefunden. Es läßt sich vermuten, daß es sich dabei um inaktivierte Retroelemente handelt.

Dies wirft die Frage auf, ob diese Retroelemente eine funktionelle Struktur darstellen und für die Centromerfunktion essentiell sind, oder ob sie im Laufe der Evolution in den Pericentromeren "gefangen" wurden und aufgrund der Struktur des Centromers inaktiviert wurden, da sie nicht mehr transkribiert werden konnten.

Im Pericentromer sind nur die drei in dieser Arbeit untersuchten repetitiven Elemente dargestellt. Hier sei darauf verwiesen, daß noch Subklone vorhanden sind, die wahrscheinlich weitere repetitive Elemente darstellen (3.2.1.1, 3.2.1.2). Abgesehen davon gibt es wahrscheinlich noch eine Vielzahl anderer repetitiver Elemente, die bisher noch nicht bekannt sind.

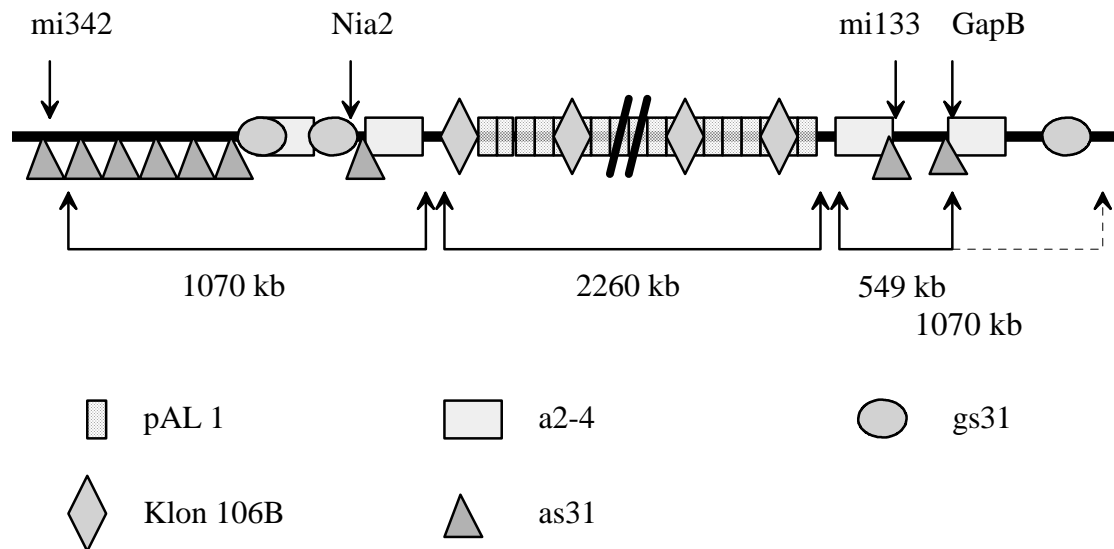


Abb. 23: schematischer Aufbau der Centromer 1-Region von *Arabidopsis thaliana*.

der zentrale Bereich enthält das pAL 1-Element (180bp), dazwischen findet man vereinzelte degenerierte Athila-Elemente (106B). Außerhalb des zentralen Bereiches findet man weitere repetitive Elemente (a2-4, as31, gs31) und komplexe DNA Abschnitte. Die physikalischen Abstände der Darstellung sind nicht maßstabsgerecht.

In dieser Arbeit wurde die Centromer 1-Region von *Arabidopsis thaliana* genetisch sehr präzise, in ein Intervall von nur 0,08 cM, kartiert. Es war auch möglich, das CEN 1 physikalisch genau zu lokalisieren und so seinen Platz auf den bestehenden BAC contigs zu bestimmen. Es wird vorgeschlagen das Centromer als die Einheit von linkem Pericentromer, Kern und rechtem Pericentromer zu betrachten. Pericentromere und Kern sind repetitiv aber werden aus verschiedenen Elementen aufgebaut. Erstere weisen noch aktive Gene auf. Alle drei sind heterochromatisch.

Die Größe der gesamten Centromer 1-Region läßt sich mindestens auf 2900 kb abschätzen. Der Kernbereich selbst umfaßt mindestens 1400 kb. Nähere Untersuchungen des Pericentromers zeigten, daß diese größtenteils aus repetitiven Elementen zusammengesetzt sind. Allerdings findet man vereinzelt Gene, die für wichtige Proteine kodieren (*Nia2*, *GURKE*).

Das Centromer konnte so weit eingegrenzt werden (2260 kb), daß es im Bereich des Möglichen erscheint, einen Klon zu finden, welcher die funktionelle Einheit des Centromers enthält, vorausgesetzt, das Pericentromer spielt dabei keine oder nur eine untergeordnete Rolle. Mit Hilfe dieses Klons wäre man in der Lage, ein künstliches Pflanzenchromosom (PLAC) zu generieren.

Parallel dazu muß die Struktur und der Aufbau des Centromers weiter untersucht werden, da nur so seine Funktion endgültig aufgeklärt werden kann.

5 Zusammenfassung

Bei der Zellteilung ist die gleichmäßige Verteilung der Schwesterchromatiden auf die Tochterzellen von entscheidender Bedeutung. Die Centromeren dienen dabei als Ansatzstellen für den Spindelapparat, der an dieser Verteilung beteiligt ist.

In dieser Arbeit wurde die Centromer 1-Region von *Arabidopsis thaliana* untersucht. Dazu wurde eine präzise genetische Karte um das Centromer 1 erstellt. Da *gurke* auf den unteren Arm des Chromosoms 1, nahe beim CEN 1, liegt, wurden die verwendeten Marker gegen *gurke* kartiert. Mit den genetischen Daten und weiteren molekularen Daten wurde ein verlässlicher contig der CEN 1-Region aufgebaut bzw. überprüft. Es gelang, die CEN 1-Region in ein genetisches Intervall von 0,84 cM und ein physikalisches Intervall von mind. 4400 kb einzugrenzen. Das Centromer 1 wurde auf 0,08 cM und mind. 2260 kb eingegrenzt.

Ein Vergleich der genetischen mit der physikalischen Karte zeigte eine enorme Kompressionszunahme in der CEN 1-Region bis zum 32-fachen der Normalwerte, wobei diese mit zunehmender Nähe zum Centromer größer wurde. Die Kompressionszunahme (Rekombinationsunterdrückung) läßt sich am linken Pericentromer auf wenige kb in der Nähe eines polymorphen Markers lokalisieren. Alle drei Regionen unterdrücken unterschiedlich stark. Das Auftreten repetitiver Elemente nimmt zum Centromer hin zu und korreliert mit der zunehmenden Kompression. Mit der Zunahme repetitiver Elemente geht die Abnahme komplexer DNA Abschnitte einher, man findet kaum noch Gene in der CEN 1-Region. Allerdings findet man auch welche, die eindeutig aktiv sind (*NIA 2*, *GAPB* und *GK*).

Einige dieser repetitiven Elemente wurden weiter untersucht. Diese stellen putative Retroelemente dar und kommen in allen Pericentromerbereichen von *Arabidopsis* vereinzelt vor und sind wenig bis mittel repetitiv. Es zeigte sich, daß nicht nur das Centromer aus hoch repetitiven Elementen besteht, sondern auch die pericentromerischen Bereiche. Vergleich von Sequenz (FISH kompetente Sequenzen) FISH mit BACs und DAPI zeigen, daß Pericentromer und Kern insgesamt heterochromatisch sind. Allerdings werden die Pericentromere durch andere repetitiven Elemente aufgebaut wie der Kern. Dies geht besonders deutlich aus den FISH-Analysen hervor.

6 Literaturverzeichnis

Alfenito, M.R., Birchler, J.A. (1993) Molecular characterization of a maize B chromosome centric sequence. *Genetics* **135**:589-597

Allshire, R.C. (1997) Centromeres, checkpoints and chromatid cohesion. *Curr Opin. Genet. Dev* **7**:264-273

Aragon-Alcaide, L., Miller, T., Schwarzacher, T., Reader, S., Moore, G. (1996) A cereal centromeric sequence. *Chromosoma* **105**:261-268

Baker, W.K. (1958) Crossing over in heterochromatin. *Amm. Nat.*, **92**, 59-60

Beadle, G.W. (1932) A possible influence of the spindle fibre on crossing-over in *Drosophila*. *PNAS*, **18**, 160-165

Bennetzen, J.L. (1996) The contribution of retroelements to plant genome organization, function and evolution. *Trends Microbiology* **4**, 347-353

Birnboim, H.C., Doly, J. (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nuc. Acids Res.* **7**: 1513-1523.

Brandes, A., Heslop-Harrison, J.S., Kamm, A., Kubis, S., Doudrick, R.L. and Schmidt, T. (1997) Comparative analysis of the chromosomal and genomic organization of Ty1-copia-like retrotransposons in pteridophytes, gymnosperms and angiosperms. *Plant Molecular Biology* **33**: 11-21

Casacuberta, E., Casacuberta, J.M., Puigdomenech, P. and Monfort, A. (1998) Presence of miniature inverted-repeat transposable elements (MITEs) in the genome of *Arabidopsis thaliana*: characterisation of the *Emigrant* family of elements. *The Plant Journal* **16**(1), 79-85

Choi, S.D., R. Creelman, J. Mullet, and R.A. Wing (1995-1) Construction and characterization of a bacterial artificial chromosome library from *Arabidopsis thaliana* *Weeds World*. **2**, 17-20.

Choi, S.D. et al. (1995-2) *PMBR* **13**, 124-128

Clarke, L. (1998) Centromeres: proteins, protein complexes and repeat domains at centromeres of simple eukaryotes. *Curr Opin Genet. Dev* **8**(2):212-8

Copenhaver, G.P., Browne, W.E., Preuss, D. (1998) Assaying genom wide recombination and centromere functions with *Arabidopsis* tetrads. *PNAS* **95**:247-252

Copenhaver, G.P, Nickel, K., Kuromori, T., Benito, M.-I., Kaul, S., Lin, X., Bevan, M., Murphy, G., Harris, B., Parnell, L.D., McCombie, W.R., Martienssen, R.A., Marra, M., Preuss, D. (1999) Genetic definition and sequence analysis of *Arabidopsis* Centromeres. *Science* **286**: 2468-2474

-
- Dellaporta, S.L., Wood, J., Hicks, J.B.** (1983) A plant miniprep: version II. *Plant Mol. Biol. Rep.* **1**:19-21
- du Sart, D., Cancilla, M.R., Earle, E., Mao, J.I., Saffery, R., Tainton, K.M., Kalitsis, P., Martyn, J., Barry, A.E. and Choo, K.H.** (1997) A functional neo-centromere formed through activation of a latent human centromere and consisting of non-alpha-satellite DNA. *Nat. Genet.*, **16**, 144-153
- Eberl, D.F., Duyf, B.J. and Hilliker, A.J.** (1993) The role of heterochromatin in the expression of a heterochromatic gene, the rolled locus of *Drosophila melanogaster*. *Genetics* **134**, 277-292
- Fanti, L., Dorer, D.R., Berloco, M., Henikoff, S. and Pimpinelli, S.** (1998) Heterochromatin protein 1 binds transgene arrays. *Chromosoma*, **107**, 286-292
- Feinberg, A.P., Vogelsang, B.** (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* **132**: 6-13.
- Feinberg, A.P. und Vogelsang, B.** (1984). Addendum: A technique for radiolabelling DNA restriction endonucleases fragments to high specificity. *Anal. Biochem.* **137**, 266
- Festenstein, R., Toliani, M., Corbella, P., Mamalaki, C., Parrington, J., Fox, M., Miliou, A., Jones, M. and Kioussis, D.** (1996) Locus control region function and heterochromatin-induced position effect variegation. *Science* **271**, 1123-1125
- Fischer, T.C.** (1997) Untersuchungen zur Positionsklonierung des *GURKE* - Gens und Charakterisierung der Centromerregion von *Arabidopsis thaliana*. Dissertation, Institut für Genetik, TU München
- Franklin, A.E. and Cande, W.Z.** (1999) Nuclear Organization and Chromosome Segregation. *The Plant Cell*, Vol. **11**, 523-534
- Fransz, P.F., Alonso-Blanco, C., Liharska, T.B., Peeters, A.J.M., Zabel, P. and de Jong, J.H.** (1996) High-resolution physical mapping in *Arabidopsis thaliana* and tomato by fluorescence in situ hybridisation to extended fibres. *The Plant Journal* **9**:421-430
- Fransz, P.F., Armstrong, S., Alonso-Blanco, C., Fischer, T.C., Torres-Ruiz, R.A. and Jones, G.** (1998) Cytogenetics for the model system *Arabidopsis thaliana*. *The Plant Journal* **13**(6), 867-876
- Fransz, P.F., Armstrong, S., de Jong, J.H., Parnell, L.D., van Drunen, C., Dean, C., Zabel, P., Bisseling, T. and Jones, G.H.** (2000) Integrated cytogenetic map of chromosome arm 4S of *Arabidopsis thaliana*: structural organization of heterochromatic knob and centromere region. *Cell*, **100**, 367-376
- Gelvin, S.B., Schilperoort, R.A.** (Hrsg.) (1994) Plant Molecular Biology Manual, 2nd edition. Kluwer Academic Publishers; Dordrecht, Niederlande
-

-
- Grancell, A., Sorger, P.K.** (1998) Chromosome movement: kinetochores motor along. *Curr Biol* **8**:R382-R385
- Griffith, A., Miller, J., Suzuki, D., Lewontin, R., Gelbart, W.** (1993) An Introduction to Genetic Analysis, fifth edition, W. H. Freeman and Company, New York
- Harrington, J.J., Bokkelen, G.V., Mays, R.W., Gustashaw, K., Willard, H.F.** (1997) Formation of de novo centromeres and construction of first generation human artificial chromosomes. *Nat Genet* **15**:345-355
- Hauser, M.-T., Adhami, F., Dörner, M., Fuchs, E. and Glössel, J.** (1998) Heteroduplex analysis combined with mutation detection enhancement (HAMDE) gel electrophoresis: an efficient method to generate codominant PCR based markers. *The Plant Journal* **16**(1), 117-125
- Hemleben, V., Torres-Ruiz, R. A., Schmidt, T. and Zentgraf, U.** (2000) Molecular Cell Biology: Role of Repetitive DNA in Nuclear Architecture and Chromosome Structure. Progress in Botany, Vol. 61, Springer-Verlag Berlin Heidelberg
- Iwahara, J., Kigawa, T., Kitagawa, K., Masumoto, H., Okazaki, T., Yokoyama, S.** (1998) A helix-turn-helix structure unit in human centromere protein B (CENP-B). *EMBO J* **17**:827-837
- Jiang, J., Nasuda, S., Dong, F., Scherrer, C.W., Woo, S.S., Wing, R.A., Gill, B.S., Ward, D.C.** (1996) A conserved repetitive DNA element located in the centromeres of cereal chromosomes. *PNAS* **93**:14210-14213
- Kamm, A., Galasso, I., Schmidt, T., Heslop-Harrison, J-S.** (1995) Analysis of a repetitive DNA family from *Arabidopsis arenosa* and relationships between *Arabidopsis* species. *Plant Mol Biol* **27**, 853-862
- Karpen, G.H., Allshire, R.C.** (1997) The case of epigenetic effects on centromere identity and function. *Trends Genet* **13**, 489-496
- Kaszas, E., Birchler, J.** (1996) Misdivision analysis of the centromere structure in maize. *EMBO J.* **15**, 5246-5255
- Kumar, A., Pearce, S.R., McLean, K., Harrison, G., Li, D., Heslop-Harrison, J.S., Waugh, R., Flavell, A.J.** (1997) The Ty1-copia group of retrotransposons in plants: genomic organization, evolution and use as molecular markers. *Genetica* **100**, 205-217
- Konieczny, A. und Ausubel, F. M.** (1993) A procedure for mapping *Arabidopsis* mutations using co-dominant ecotype-specific PCR-based markers. *The Plant Journal* **4**, 403-410
- Koornneef, M. & Van der Veen, J.H.** (1983-1) Trisomics in *Arabidopsis thaliana* and the location of linkage groups. *Genetica* **61**, 41-46
-

- Koornneef, M.** (1983-2) The use of telotrisomics for centromere mapping in *Arabidopsis thaliana* (L.) Heynh. *Genetica* **62**, 33-40
- Jones, R.N., Rees, H.** (1982) B Chromosomes. Academic Press, London
- Liu, Y.G., Mitsukawa, N., Lister C., Dean C., Whittier, R. F.** (1996) Isolation and mapping of a new set of 129 RFLP markers in *Arabidopsis thaliana* using recombinant inbred lines. *Plant J.* **10**(4), 733-6
- Maluszynska, J., Heslop-Harrison, J.S.** (1991) Localization of tandemly repeated DNA sequences in *Arabidopsis thaliana*. *Plant J* **1**, 159-166
- Maluszynska, J., Heslop-Harrison, J.S.** (1993) Molecular cytogenetics of the genus *Arabidopsis*: In situ localization of rDNA sites, chromosome numbers and diversity in centromeric heterochromatin. *Ann Bot* **71**, 479-484
- Martinez-Zapater, J.M., Estelle, M.A., Sommerville, C.R.** (1988) A highly repeated DNA sequence in *Arabidopsis thaliana*. *Mol Gen Genet* **204**, 417-423
- Masumoto, H., Yoda, K., Ikeno, M., Kitagawa, K., Muro, N. and Okazaki, T.** (1993) Properties of CENP-B and its target sequence in a satellite DNA. *NATO ASI Series Heidelberg, Series H* **72**, 31-43
- McKim, K.S., Hawley, R.S.** (1995) Chromosomal control of meiotic cell division. *Science* **270**, 1595-1601
- Meyerowitz, E.M. and Pruitt, R.E.** (1985) *Arabidopsis thaliana* and Plant Molecular Genetics. *Science* **229**, 1214-1218
- Meyerowitz, E.M.** (1989) *Arabidopsis thaliana*, a useful weed. *Cell* **56**, 263-269
- Miller, J.T., Jackson, S.A., Nasuda, S., Gill, B.S., Wing, R.A., Jiang, J.** (1998) Cloning and characterization of a centromere-specific repetitive DNA element from *Sorghum bicolor*. *Theor Appl Genet* **96**, 832-839
- Mozo, T., Fischer, S., Shizuya, H., Altmann, T.** (1998-1) Construction and characterization of the IGF *Arabidopsis* BAC library. *Mol. Gen. Genet.* **258**, 562-570
- Mozo, T., Fischer, S., Meier-Ewert, S., Lehrach, H., Altmann, T.** (1998-2) Use of the IGF BAC library for physical mapping of the *Arabidopsis thaliana* genome. *Plant J.* **16**, 377-384
- Mozo, T., Dewar, K., Dunn, P., Ecker, J.R., Fischer, S., Kloska, S., Lehrach, H., Marra, M., Martienssen, R., Meier-Ewert, S., Altmann, T.** (1999) A complete BAC-based physical map of the *Arabidopsis thaliana* genome. *Nature Genetics* **22**(3):271-5
- Murata, M., Ogura, Y., Motoyoshi, F.** (1994) Centromeric repetitive sequences in *Arabidopsis thaliana*. *Jpn J Genet* **69**:361-370

- Murata, M., Heslop-Harrison, J.S. and Motoyoshi, F.** (1997) Physical mapping of the 5S ribosomal RNA genes in *Arabidopsis thaliana* by multi-color fluorescence in situ hybridization with cosmid clones. *The Plant Journal* **12**(1), 31-37
- Murphy, T.D. and Karpen, G.H.** (1995) Localization of centromere function in a *Drosophila* minichromosome. *Cell* **82**, 599-609
- Nicklas, R.B.** (1997) How cells get the right chromosomes. *Science* **275**:632-637
- Patton, D.A., Franzmann, L.H., Meinke, D.W.** (1991) Mapping genes essential for embryo development in *Arabidopsis thaliana*. *Mol. Gen. Genet.* **227**(3), 337-347
- Peacock, W.J., Dennis, E.S., Rhoades, M.M., Pryor, A.J.** (1981) Highly repeated DNA sequence limited to knob heterochromatin in maize. *PNAS* **78**(7):4490-4494
- Pearce, S.R., Harrison, G., Li, D., Heslop-Harrison, J.S., Kumar, A., Flavell, A.J.** (1996) The Ty1-copia group retrotransposons in *Vicia* species: copy number, sequence heterogeneity and chromosomal localization. *Mol Gen Genet* **250**, 305-315
- Pellisier, T., Tutois, S., Deragon, J.M., Tourmente, S., Genestier, S., Picard, G.** (1995) *Athila*, a new retroelement from *Arabidopsis thaliana*. *Plant Molecular Biology* **29**, 441-452
- Pellisier, T., Tutois, S., Tourmente, S., Deragon, J.M., Picard, G.** (1996) DNA regions flanking the major *Arabidopsis thaliana* satellite are principally enriched in *Athila* retroelement sequences. *Genetica* **97**:141-151
- Pidoux, A.L. and Allshire, R.C.** (2000) Centromeres: getting a grip of chromosomes. *Curr. Opin. Cell Biol.* **12**, 308-319
- Pimpinelli, S., Goday, C.** (1989) Unusual kinetochores and chromatin diminution in *Parascaris*. *Trends Genetic* **5**:310-315
- Pluta, A.F., Mackay, A.M., Ainzstein, A.M., Goldbegr, I.G., Earnshaw, W.C.** (1995) The centromere: hub of chromosomal activities. *Science* **270**:1591-1594
- Presting, G.G., Malysheva, L., Fuchs, J., Schubert, I.** (1998) A Ty3/gypsy retrotransposon-like sequences localizes to the centromeric regions of cereal chromosomes. *Plant J.* **16**:721-728
- Pruitt, R.E. and Meyerowitz, E.M.** (1986) Characterization of the genome of *Arabidopsis thaliana*. *Journal of Molecular Biology* **187**: 169-183
- Rédei, G.P.** (1970) *Arabidopsis thaliana* (L.) Heynh. A review of the genetics and biology. *Bibl. Genet.* **20**: 1-151
- Rédei, G.P.** (1975) *Arabidopsis* as a genetic tool. *Annual Review of Genetics* **9**: 111-127

- Richards, E.J., Goodman, H.M., Ausubel, F.M.** (1991) The centromere region of *Arabidopsis thaliana* chromosome 1 contains telomere-similar sequences. *Nucleic Acids Res* **19**:3351-3357
- Richards, E.J., Dawe, R.K.** (1998) Plant centromeres: structure and control. *Curr opin Plant Biol* **1**:130-135
- Rogers, S.O., Bendich, A.J.** (1994) Extraction of total cellular DNA from plants, algae and fungi. *Plant Molecular Biology Manual*. Gelvin, S.B., Schilperoort, R.A. (Hrsg.), pp. D1, 1-8. Kluwer Academic Publishers; Dordrecht, Niederlande
- Ross, K.J., Fransz, P. & Jones, G.H.** (1996) a light microscopic atlas of meiosis in *Arabidopsis thaliana*. *Chromosome Research* **4**, 507-516
- Round, E.K., Flowers, S.K., Richards, E.J.** (1997) *Arabidopsis thaliana* centromere regions: genetic map positions and repetitive DNA structure. *Genome Res* **7**:1045-1053
- Sambrook, J., Fritsch, E.F. and Maniatis, T.** (1989) *Molecular Cloning. A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, New York
- Schmidt, T., Kubis, S., Heslop-Harrison, J.S.** (1995) Analysis and chromosomal localization of retrotransposons in sugar beet (*Beta vulgaris* L.): LINEs and Ty1-copia-like elements as major components of the genome. *Chromosome Res* **3**, 335-345
- Schmidt, T., Heslop-Harrison, J.S.** (1996) High resolution mapping of repetitive DNA by in situ hybridisation: molecular and chromosomal features of prominent dispersed and discretely localized DNA families from wild beet species *Beta procumbens*. *Plant Mol Biol* **30** 1099-1113
- Schmidt, T., Heslop-Harrison, J.S.** (1998) Genomes, genes and junk: the large-scale organization of plant chromosomes. *Trends Plant Sci* **3**, 195-199
- Simoens, C.R., Gielen, J., Van Montagu, M. Inze, D.** (1988) Charakterization of highly repetitive sequences of *Arabidopsis thaliana*. *Nucleic Acis Res* **16**:6753-6766
- Sommerville, C. and Sommerville S.** (1999) Plant functional genomics. *Science* **285**: 380-383
- Sun, X., Wahlstrom, J., Karpen, G.** (1997) Molecular structure of a functional *Drosophila* centromere. *Cell* **91**, 1007-1019
- Sunkel, C.E., Sussex, I.M.** (1995) The elusive centromere: sequence divergence and functional conservation. *Current Opinion in Genetics and Develonment* **5**, 756-767
- Suoniemi, A., Tanskanen, J., Schulman, A.H.** (1998) Gypsy-like retrotransposons are widespread in the plant kingdom. *Plant Journal* **13**, 699-705

-
- The Arabidopsis Genome Initiative** (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408**, 796-815
- Theologis, A., Ecker, J.R. et al.** (2000) Sequence and analysis of the chromosome 1 of the plant *Arabidopsis thaliana*. *Nature* **408**, 816-820
- Thompson, H.L., Schmidt, R., Dean, C.** (1996 a) Analysis of the occurrence and nature of repeated DNA in an 850 kb region of *Arabidopsis thaliana* chromosome 4. *Plant Mol Biol* **32**:365-376
- Thompson, H.L., Schmidt, R., Dean, C.** (1996 b) Identification and distribution of seven classes of middle-repetitive DNA in the *Arabidopsis thaliana* genome. *Nucleic Acid Res* **24**:3017-3022
- Thompson, H.L., Schmidt, R., Brandes, A., Heslop-Harrison, J.S., Dean, C.** (1996 c) A novel repetitive sequence associated with the centromeric regions of *Arabidopsis thaliana* chromosomes. *Mol. Gen. Genet.* **253**, 247-252
- Traut, W.,** (1991) Chromosomen. Springer, Berlin Heidelberg New York
- Tyler-Smith, C. and Florida, G.** (2000) Many path to the top of the mountain: diverse evolutionary solutions to centromere structure. *Cell*, **102**, 5-8
- Vershinin, A.V. and Heslop-Harrison, J.S.** (1998) Comparative analysis of the nucleosomal structure of rye, wheat and their relatives. *Plant Molecular Biology* **36**, 149-161
- Voytas, D.F., Cummings, M.P., Konieczny, A., Ausubel, F.M.** (1992) copia-like retrotransposons are ubiquitous among plant. *Proc Natl Acad Sci USA* **89**, 7124-7128
- Wessler, S.R., Bureau, T.E., White, S.E.** (1995) LTR-retrotransposons and MITEs: important players in the evolution of plant genomes. *Current Opinion in Genetics and Development* **5**, 814-821
- Wessler, S.R.** (1996) Plant retrotransposons: turned on by stress. *Curr Biol* **6**, 959-961
- White, S.E., Habera, L.F., Wessler, S.R.** (1994) Retrotransposons in the flanking regions of normal plant genes: a role for copia-like elements in the evolution of gene structure and expression. *Proc. Natl. Acad. Sci. USA* **91**, 11792-11796
- Wilkinson, J.Q., Crawford, N.M.** (1991) Identification of the *CHL3* gene as the nitrate reductase structural gene *NIA2*. *The Plant Cell* **3**, 461-471
- Willard, H.F.** (1990) Centromeres of mammalian chromosomes. *Trends Genet.*, **6**, 410-416
- Willard, H.F.** (1998) Centromeres: the missing link in the development of human artificial chromosomes. *Curr Opin Genet Dev* **8**, 219-225
-

Williams, B.C., Murphy, T.D., Goldberg, M.L. and Karpen, G.H. (1998) Neocentromere activity of structurally acentric mini-chromosomes in *Drosophila*. *Nat. Genet.* **18**, 30-37

Zhong, X.B., de Jong, J.H., and Zabel, P (1996a) Localization of repetitive sequences on metaphase and pachytene chromosomes of tomato using fluorescence in situ hybridisation. *Chromosome Res.* **4**, 24-28

Zhong, X.B., Fransz, F.P., Wennekes-van Eden, J., Zabel, P., van Kammen, A. and de Jong, J.H. (1996a) High resolution mapping on pachytene chromosomes and extended DNA fibres by fluorescence in situ hybridisation. *Plant Molecular Biology Reporter* **14**, 56-65

Zhong, X.B., Fransz, F.P., Wennekes-van Eden, J., de Jong, J.H., and Zabel, P. (1997) Fluorescence in situ Hybridisation in Tomato. Laboratory Protocols, Tomato Genome Group, Department of Biomolecular Sciences, Wageningen Agricultural University

Zhong, X.B., Fransz, F.P., Wennekes-van Eden, J., Ramanna, M.S., van Kammen, A., Zabel, P. and de Jong, J.H. (1998) FISH studies reveal the molecular and chromosomal organisation of individual telomere domains in tomato. *The Plant Journal* **13**(4), 507-517

Zinkowski, R.P., Meyne, J., Brinkley, B.R. (1991) The centromere-kinetochore complex: a repeat subunit model. *J Cell Biol* **113**, 1091-1110

7 Anhang

7.1 Genetische Kartierung

7.1.1 Rekombinante Pflanzen von der GapB-Seite

Pflanze	Elter	gurke	GapB	mi133	3l22-sp6	6j23-t7	Nia2	t21c16-t7	t31i22-t7
33	U413XNd	gk	N	LN	LN	LN	LN	LN	LN
513	R216XNd	gk	L	L	LN		LN	LN	LN
573	R216XNd	wt	LN	N	N		N	N	N
667	R216XNd	wt	LN	N	N		N	N	N
1153	R216XNd	gk	N	LN	LN		LN	LN	LN
1370	G301XNd	wt	LN	LN	N	N (L)	N	N	N
1380	G301XNd	wt	LN	LN	LN	LN	N	N	N
1743	R216XNd	gk	N	N	LN	LN		LN	
1939	R216XNd	gk	L	L	L	L (N)	LN	LN	LN

■ rekombinant; rekombinante Klassen: L/gk; N/gk; LN/wt; L/wt

L = Landsberg homozygot

N = Niederzenz homozygot

LN = Landsberg, Niederzenz heterozygot

gk = heterozygot gk für gurke

wt = homozygot wt für gurke

7.1.2 Rekombinante Pflanzen von der *tt1*-Seite

Pflanze	gurke	tt1	mi342	lc13-t7	4a3-t7	17A20-	Nia2	icc3
1	gk	tt1	LC				LC	
4	gk	tt1	L	LC	LC	LC	LC	LC
10	gk	tt1	L	L	L	LC	LC	LC
17	gk	tt1	L	LC	LC	LC	LC	LC
20	gk	tt1	LC		LC		LC	
21	gk	tt1	LC		LC			
27	gk	tt1	LC				LC	
54	wt	wt	L					
55	wt	wt	LC	L	L	L	L	L
56	wt	wt	LC	L	L	L	L	L
57	wt	wt	L					
58	wt	wt	L					
59	wt	wt	LC	L	L	L	L	L
60	wt	wt	L					
61	wt	wt	L					
62	wt	wt	LC	L	L	L	L	L
63	wt	wt	L		L			

Pflanze	gurke	tt1	mi342	lc13-t7	4a3-t7	17A20-	Nia2	icc3
64	wt	wt	L					
65	wt	wt	LC	L	L	L	L	L
66	wt	wt	L					
67	wt	wt	L					
68	wt	wt	L					
69	wt	wt	L					
70	wt	wt	LC	L	L	L	L	L
71	wt	wt	L					
72	wt	wt	L					
73	wt	wt	LC	L	L	L	L	L
74	wt	wt	L					
75	wt	wt	L					
76	wt	wt	L					
77	wt	wt	LC	LC	L	L	L	L
78	wt	wt	L					
79	wt	wt	L					
80	wt	wt	L					
81	wt	wt	LC		L	L	L	L
82	wt	wt	L					
83	wt	wt	L					
84	wt	wt	L					
88	gk	tt1	LC	LC	LC			
89	gk	tt1	LC	LC	LC			
90	gk	tt1	LC	LC	LC			
91	gk	tt1	LC	LC	LC			
92	gk	tt1	LC	LC	LC			
93	gk	tt1	LC	LC	LC			
94	gk	tt1	LC	LC	LC			
96	gk	tt1	LC	LC	LC			
97	gk	tt1	L	LC	LC			
99	gk	tt1	LC	LC	LC			
105	gk	tt1						
107	wt	wt						
108	wt	wt						
109	wt	wt						
110	wt	wt	L		L			
111	wt	wt	L	L	L		L	
112	wt	wt	L	L	L			
113	wt	wt		LC	LC			
114	wt	wt	L	L	L			
115	wt	wt		L				
116	wt	wt		L	L			
117	wt	wt	LC	L	L			
118	wt	wt	LC	L	L			
119	wt	wt	L	L	L			
120	wt	wt	L	L	L			
121	wt	wt	LC	L	L			
122	wt	wt	L	L	L			
123	wt	wt	L	L	L			
124	wt	wt	L	L	L			
125	wt	wt	LC	L	L			
126	wt	wt	L	L	L			
127	wt	wt	LC	LC	LC	L	L	
128	wt	wt	LC	LC	LC			
129	wt	wt	LC	LC	LC			
131	wt	wt						

Pflanze	gurke	tt1	mi342	lc13-t7	4a3-t7	17A20-	Nia2	icc3
132	gk	tt1						
133	gk	tt1	LC	LC	LC			
134	wt	wt						
135	wt	wt						
136	wt	wt						

■ rekombinant; rekombinante Klassen: L/gk; C/gk; LC/wt; C/wt

L = Landsberg homozygot

C = Columbia (emb22) homozygot

LC = Landsberg, Columbia heterozygot

gk = heterozygot gk für gurke

wt = homozygot wt für gurke

tt1 = homozygot tt1 für transparent testa

wt = heterozygot tt1 oder wt für transparent testa

7.2 Single copy-Proben

f11o6sp6	mi342	t21c16t7	t31i22t7	t18h22sp6	t4n16t7	f17a20sp6	t4n16sp6	t3h3t7	NIA2
17j19	F8M2	15G17	14B14	14B14	14B14	F17A20	F17A20	F17A20	F17A20
4c4	F15O4	17K3	15G17	15G17	15G17	F24P1	F13J4	F13J4	F13J4
7o19	F25L18	19H19	17K3	17K3	17K3	F7G10	F22A12	F22A12	F22A12
21p2	F28B2	1O3	19E7	19E7	19E7	F22C3	F24P1	F24P1	F24P1
25112	T3I18	20D20	19H19	19H19	19H19	F6M21	F7G10	F7G10	F7G10
28a7	T7A21	T11D8	1O3	1O3	1O3	F17G21	F22C3	F22C3	F22C3
11o6	T8A20	T13A12	20D20	20D20	3A14	T4n16	F6M21	F6M21	F6M21
10p4	T11D4	T16O22	T10H24	T10H24	T10H24	T6I22	F17G21	F17G21	F17G21
2b22	T22J21	T17L7	T11D8	T11D8	T11D8	T1p12	T4n16	T3h3	T8a19
3n5		T20M18	T13A12	T13A12	T13A12	T2M19			T8I21
		T21C16	T14P22	T14P22	T14P22	T2P10			T3h3
		T23N15	T16O22	T16O22	T15E15	T8C24			T6f19
		T28M11	T17L7	T17L7	T16F10	T1c22			T6I22
		T30P8	T20M18	T18H22	T17I11				T1p12
		T32O22	T21C16	T20M18	T18H22				
		T4F10	T23N15	T21C16	T1N2				
		T4G9	T28M11	T28M11	T20I12				
		T4M12	T30P8	T30P8	T20M18				
		T5B13	T31I22	T31I22	T21C16				
		T8C24	T32E20	T32E20	T25K12				
		T9K1	T32O22	T32O22	T28M11				

f11o6sp6	mi342	t21c16t7	t31i22t7	t18h22sp6	t4n16t7	f17a20sp6	t4n16sp6	t3h3t7	NIA2
			T4F10	T4F10	T2M19				
			T4M12	T4M12	T2P10				
			T5B13	T5B13	T31I22				
			T8C24	T8C24	T32E20				
			T9K1	12L8	T4N16				
				15P16	T5B13				
				18I21	T8C24				
				20M14					
				20N8					
				2L6					
				3P21					
				T12A5					
				T12D22					
				T14E8					
				T14K14					
				T15I16					
				T1O5					
				T26J12					
				T27K3					
				T28D8					
				T28L9					
				T2M22					
				T4D21					
				T6B7					
				T5A5					

BAC die mit den single copy Proben links vom CEN 1 hybridisieren. BACs ohne Kennzeichnung oder mit einem F gekennzeichnete BACs stammen aus der IGF BAC Bibliothek. Mit einem T (t) gekennzeichnete BACs stammen aus der TAMU BAC Bibliothek.

f6j23t7	f3122sp6	f28b8t7	mi133	jcc3	GapB
28b8	28b8	28b8	F2G10	20f15	F2N12
7k22	7k22	1i6	F7F22	27h21	F4B19
3122	3122	26f14	F12A16	4g21	F5J2
6j23	1i6	20b1	F19N10	1p9	F8L20

f6j23t7	f3l22sp6	f28b8t7	mi133	jcc3	GapB
i16	26f14	18c13	F21I19	915	F8N8
26f14	20b1	20k13		10i2	F12J8
18c13	20k13	10k7		17p8	F19C3
20k13	10k7	12b23		2g10	F22O4
10k7		t4m7		21i19	F26H23
t20m9		t8e12		19n10	F28F6
t24a5		t4i21		8o21	13a11
t16o17		t32l23			7l6
t8a15		t14d22			6a8
t2d14		t1a1			
t8f20		t9o11			
t14l17					
t14k17					
t7m4					
t5g7					
t17d14					
t5a16					
t7p22					
t1p24					

BACs die mit den single copy Proben rechts vom CEN 1 hybridisieren.bridisieren.
 BACs ohne Kennzeichnung oder mit einem F gekennzeichnete BACs stammen aus der
 IGF BAC Bibliothek. Mit einem T (t) gekennzeichnete BACs stammen aus der TAMU
 BAC Bibliothek.

7.3 Repetitive Proben

7.3.1 BACs

a2-4		gs31		as31		a2-4		gs31		as31	
23F21	chr1	6C2	chr1	15H2	chr1			10H11	chr5	24B20	chr4
4J16	chr1	6J1	chr1	15P3	chr1	168		15O1	chr5	24H16	chr4
5A13	chr1	8L2	chr1	16J24	chr1			16K2	chr5	24M19	chr4
6M21	chr1	9G18	chr1	16K23	chr1			1D21	chr5	2A21	chr4
7G10	chr1	17A20	chr1	16K24	chr1			21C15	chr5	2E14	chr4
8O5	chr1	17G21	chr1	17A20	chr1			22D21	chr5	3E16	chr4
11I14	chr1	22C3	chr1	19D10	chr1			24D21	chr5	3H6	chr4
16P16	chr1	6M21	chr1	19M18	chr1			2B23	chr5	3I20	chr4
24B11	chr1	11K13	chr1	1C13	chr1			4C18	chr5	4E4	chr4

a2-4		gs31		as31		a2-4		gs31		as31	
2G17	chr1	13A11	chr1	1F24	chr1			4I4	chr5	4N3	chr4
4F16	chr1	7L6	chr1	1H9	chr1			7E23	chr5	5J21	chr4
9I21	chr1	22K23	chr1	1N23	chr1			10D6	chr5	5P10	chr4
11N11	chr1	2D19	chr1	20J6	chr1			10F13	chr5	6I23	chr4
14P23	chr1	2E11	chr1	20K7	chr1			11I23	chr5	6M14	chr4
16I17	chr1	11J10	chr1	20O13	chr1			11N13	chr5	7F17	chr4
20K8	chr1	13A15	chr1	20O5	chr1			11P10	chr5	7M21	chr4
21I12	chr1	16O8	chr1	20O7	chr1			12K19	chr5	8O18	chr4
22K23	chr1	1N1	chr1	20P22	chr1			12N11	chr5	9B6	chr4
2D19	chr1	21A17	chr1	21F1	chr1			12P2	chr5	9F11	chr4
10G22	chr2	7L21	chr1	21I23	chr1			14B13	chr5	9M13	chr4
12J5	chr2	9D23	chr1	21M2	chr1			14B17	chr5	13G23	chr4
14O4	chr2	22O2	chr1	22A12	chr1			14K10	chr5	10H11	chr5
15O11	chr2	4C5	chr1	22C2	chr1			16I11	chr5	11A20	chr5
17G4	chr2	6G20	chr1	22E15	chr1			18A12	chr5	13D8	chr5
17K1	chr2	11O20	chr1	23F21	chr1			18F18	chr5	13I23	chr5
18M14	chr2	1P7	chr1	23N13	chr1			18G23	chr5	13K20	chr5
18O8	chr2	5L4	chr1	24A15	chr1			19H2	chr5	15O1	chr5
19J11	chr2	12D2	chr2	24O6	chr1			19N2	chr5	16K2	chr5
21C8	chr2	12O11	chr2	2C1	chr1			1C19	chr5	17N17	chr5
21D2	chr2	15G19	chr2	3A6	chr1			1E16	chr5	17N21	chr5
23D21	chr2	23N9	chr2	4A3	chr1			1E17	chr5	17P24	chr5
24C20	chr2	4M17	chr2	5J5	chr1			1H18	chr5	18F14	chr5
24F2	chr2	11C19	chr2	6C2	chr1			1L6	chr5	19A12	chr5
3I24	chr2	11C23	chr2	6F4	chr1			20E13	chr5	21C15	chr5
3K12	chr2	12C2	chr2	6J1	chr1			20J2	chr5	21D16	chr5
4E20	chr2	12J5	chr2	6M21	chr1			20J3	chr5	22D21	chr5
5C2	chr2	12N14	chr2	7H24	chr1			20N5	chr5	22D5	chr5
5N20	chr2	14F1	chr2	8E5	chr1			22D11	chr5	23C8	chr5
6G23	chr2	14J17	chr2	8I16	chr1			22E14	chr5	23P1	chr5
9B11	chr2	14K23	chr2	8L2	chr1			23N15	chr5	3B6	chr5
9D10	chr2	14O11	chr2	8M2	chr1			24M17	chr5	3F24	chr5
12J6	chr4	15E22	chr2	9G18	chr1			2A20	chr5	3J17	chr5
13E19	chr4	15J21	chr2	10A21	chr1			4M5	chr5	4C18	chr5
13I2	chr4	15K19	chr2	10I2	chr1			5C21	chr5	4I4	chr5
14E20	chr4	16B23	chr2	11D19	chr1			5K15	chr5	5N11	chr5
15K12	chr4	16L7	chr2	12G2	chr1			5N22	chr5	6C14	chr5
16H22	chr4	17C8	chr2	13N15	chr1			6L6	chr5	7E23	chr5
17M13	chr4	17D8	chr2	14J1	chr1			7N22	chr5	8D17	chr5
17O19	chr4	17F17	chr2	16M11	chr1			7O16	chr5	8J18	chr5
17P22	chr4	17K1	chr2	17P8	chr1			8G23	chr5	14A22	chr5
18F6	chr4	18K15	chr2	19C17	chr1			14I1	chr5	15G12	chr5
18N13	chr4	18O8	chr2	19C3	chr1			11C7	u ctg1	15I15	chr5
19B18	chr4	19M7	chr2	1P9	chr1			12C4	u ctg1	15L21	chr5
20B12	chr4	20E7	chr2	22H4	chr1			12J7	u ctg1	15O18	chr5
20F20	chr4	22B6	chr2	4G21	chr1			13D14	u ctg1	17C3	chr5
20J15	chr4	24C20	chr2	6A8	chr1			13O17	u ctg1	17O22	chr5
20N11	chr4	2B2	chr2	8H24	chr1			15D14	u ctg1	17O24	chr5
21C9	chr4	2I13	chr2	8L20	chr1			17N2	u ctg1	18F18	chr5
22O19	chr4	3E9	chr2	9L5	chr1			18L5	u ctg1	18M12	chr5
23G24	chr4	3M13	chr2	10C15	chr1			19M5	u ctg1	1D3	chr5

a2-4		gs31		as31		a2-4		gs31		as31	
23O1	chr4	3M7	chr2	13A15	chr1			19O5	u ctg1	1E16	chr5
24F4	chr4	4L18	chr2	13L15	chr1			1D1	u ctg1	1L6	chr5
24M11	chr4	5C2	chr2	16G5	chr1			1D7	u ctg1	22C19	chr5
24N24	chr4	6M23	chr2	19A24	chr1			20H9	u ctg1	23H9	chr5
4A7	chr4	7E22	chr2	20H5	chr1			21G12	u ctg1	2A20	chr5
4I21	chr4	9D10	chr2	21D6	chr1			22G7	u ctg1	3D18	chr5
8A16	chr4	9G11	chr2	6A10	chr1			23H18	u ctg1	3K18	chr5
8K14	chr4	9P16	chr2	1G9	chr1			2G2	u ctg1	6A2	chr5
9C3	chr4	10P16	chr3	18E1	chr2			3E20	u ctg1	8I24	chr5
9G2	chr4	12A3	chr3	17N6	chr2			6F13	u ctg1	17I9	chr5
9K3	chr4	12F3	chr3	5H15	chr2			6P19	u ctg1	10B16	u ctg1
9M1	chr4	12G5	chr3	9K16	chr2			8G20	u ctg1	10H3	u ctg1
9M6	chr4	12O5	chr3	9L6	chr2			8O13	u ctg1	10H7	u ctg1
9P4	chr4	12P9	chr3	10G13	chr2			9E21	u ctg1	10I4	u ctg1
2K24	chr4	14E21	chr3	12C2	chr2			9N16	u ctg1	12J7	u ctg1
10E2	chr4	14F15	chr3	12N8	chr2			12P1	u ctg2	12O10	u ctg1
12B19	chr4	14J15	chr3	12P23	chr2			13D7	u ctg2	15A1	u ctg1
12G20	chr4	15H12	chr3	13I7	chr2			19I11	u ctg2	17A15	u ctg1
12H15	chr4	16G10	chr3	13J11	chr2			23F23	u ctg2	17D23	u ctg1
17O9	chr4	17H6	chr3	14G14	chr2			2C13	u ctg2	17N2	u ctg1
20I1	chr4	18I17	chr3	14L1	chr2			4B15	u ctg2	18K14	u ctg1
20N17	chr4	18P9	chr3	14P14	chr2			4C11	u ctg2	18L5	u ctg1
22E19	chr4	19C22	chr3	15J21	chr2			7C4	u ctg2	19M5	u ctg1
2A21	chr4	19J12	chr3	15K19	chr2			10J2	u ctg3	1D1	u ctg1
3I20	chr4	19K7	chr3	15N6	chr2			13J3	u ctg3	21G12	u ctg1
7F17	chr4	20C5	chr3	17I15	chr2			13K14	u ctg3	22G7	u ctg1
7M21	chr4	20I17	chr3	17M6	chr2			13K8	u ctg3	23H12	u ctg1
10K23	chr4	21G10	chr3	19J10	chr2			16L6	u ctg3	23H18	u ctg1
13C21	chr5	21L16	chr3	20A5	chr2			1K22	u ctg3	24O11	u ctg1
17O24	chr5	21L22	chr3	20E7	chr2			20F5	u ctg3	2F20	u ctg1
20E13	chr5	22C9	chr3	20H14	chr2			23M13	u ctg3	3E20	u ctg1
22D11	chr5	23C24	chr3	20K15	chr2			3I18	u ctg3	8O13	u ctg1
22E14	chr5	23J4	chr3	21C8	chr2			7C16	u ctg3	9E21	u ctg1
24M17	chr5	24G17	chr3	21D2	chr2			8M14	u ctg3	9I1	u ctg1
11C7	u ctg1	3G7	chr3	22C24	chr2			12C7	n.b.	9M14	u ctg1
11P20	u ctg1	4K4	chr3	23C19	chr2			13N12	n.b.	10A6	u ctg2
13D14	u ctg1	6G11	chr3	23L3	chr2			14C2	n.b.	11F13	u ctg2
13F19	u ctg1	6O4	chr3	23M2	chr2			23G5	n.b.	11P4	u ctg2
13O17	u ctg1	7B6	chr3	2B2	chr2			4I16	n.b.	13D7	u ctg2
15L24	u ctg1	7C6	chr3	2H9	chr2			5I12	n.b.	14J11	u ctg2
16M18	u ctg1	7M19	chr3	3K12	chr2			6P6	n.b.	17G22	u ctg2
17H7	u ctg1	7N10	chr3	3M7	chr2			9M18	n.b.	17M7	u ctg2
18F2	u ctg1	8C23	chr3	3M8	chr2			9N22	n.b.	19I11	u ctg2
19M5	u ctg1	8C3	chr3	5N20	chr2					1M22	u ctg2
20H9	u ctg1	11N2	chr4	6H22	chr2			274		20C16	u ctg2
22G7	u ctg1	13E19	chr4	7B19	chr2					20G20	u ctg2
22M24	u ctg1	14E20	chr4	7E22	chr2					21N24	u ctg2
22N18	u ctg1	17M13	chr4	7K9	chr2					23F23	u ctg2
3C9	u ctg1	17P22	chr4	7P20	chr2					24C5	u ctg2
3E20	u ctg1	18J23	chr4	9G11	chr2					24I23	u ctg2
4H22	u ctg1	1C4	chr4	9J6	chr2					3M22	u ctg2
4K23	u ctg1	20F20	chr4	9M17	chr2					4B15	u ctg2
5C14	u ctg1	21C9	chr4	9O11	chr2					4C11	u ctg2

a2-4		gs31		as31		a2-4		gs31		as31	
6F13	u ctg1	21D24	chr4	11C19	chr3					5O15	u ctg2
7J15	u ctg1	21M4	chr4	3G14	chr3					7J19	u ctg2
8H12	u ctg1	21N19	chr4	8D7	chr3					8G15	u ctg2
8N17	u ctg1	22O19	chr4	15B16	chr3					9A17	u ctg2
9E21	u ctg1	24N24	chr4	23G13	chr3					9K2	u ctg2
11F13	u ctg2	2O6	chr4	4K4	chr3					11H5	n.b.
20C16	u ctg2	4F2	chr4	8C3	chr3					12G10	n.b.
21N24	u ctg2	6A16	chr4	1J4	chr3					12H11	n.b.
24C5	u ctg2	6L20	chr4	15K12	chr4					22E6	n.b.
5O15	u ctg2	7M13	chr4	19B18	chr4					24E12	n.b.
10J2	u ctg3	8N15	chr4	21I2	chr4					5I12	n.b.
16L6	u ctg3	9G2	chr4	7M13	chr4					7J20	n.b.
1K22	u ctg3	10A2	chr4	8N15	chr4					8A15	n.b.
7C16	u ctg3	10B19	chr4	10E2	chr4					8D12	n.b.
8M14	u ctg3	10E2	chr4	12B19	chr4						
10D20	n.b.	10J6	chr4	12G20	chr4					298	

IGF BACs die mit den repetitiven Proben a2-4, gs31 und as31 hybridisieren. Die erste Spalte zeigt die Probe an, die zweite Spalte gibt an auf welches Chromosom und in welchen contig die einzelnen BACs lokalisiert wurden (Mozo et al., 1998-1; http://www.mpimp-golm.mpg.de/101/mpi_mp_map/access.html).

f17a20sp6	f3n5t7	f22a12sp6
17a20	12b23	17a20
22c3	9i21	13j4
24p1	18c13	22a12
6m21	17i20	22c3
7g10	20k13	11i2
17g21	10k7	12o10
t4n16	28b2	13l21
t6l22	20o5	14p5
t1p12	23m8	23i18
t2m19	7h24	2d19
t1c22	22c2	7l1
t2p10	20j6	9p22
t8c24	17j19	11o19
11o6	4c4	19i7
10p4	7o19	2o17
21h2	21p2	3i10
24n18	25l12	10g13
2b22	28a7	21c8

f17a20sp6	f3n5t7	f22a12sp6
3n5	11o6	3k12
4m12	10p4	9m17
nicht zuzuordnen	2b22	nicht zuzuordnen
t1p13	3n5	12j13
t12h9		13l20
t11k4		15f4
t6n7		
t3m2		

BACs die mit "wenig" repetitiven Proben hybridisieren.

7.3.2 Sequenzen

7.3.2.1 a 2-4 (2777 bp)

TCCTGAGGGAACTCCGAAAGAACATTTCTCTGGGTTTAGCTTCATCTCGAATTTGTTTA
GGATTTTGAAGCATTCTCTTAATTGCGGGAGGTGGTCTTTTTTCGTGAGCTGACTTGACC
AGCATGTCGTGCATATAGAGCTCCATCGTTTTACCTAGTTGGTCAGCGAACATTCTGTT
CACCAATCTTTGGTAAGTAGCTCCAGTGTTTTTTAAACCAAAGGGCATTACTTTATAAC
AGTCAGTAATAAAAAGCTGTTTTTTCTTGGTCATCTGGGCGCATAAGTATTTGGTTGTAC
CCGGAGAATGCATCCATGAAGGACAATAGTTCTGTGTCCCGCGGTGGCTTTGACCAGTCG
GTCAATATGTGGTAGTGGGAACTGTCTTTAAGACATGCCTTATTCAGGTCCATAAAGT
CGATGCATACTCTCCACTTCCCGTTTTTGTTTTTTAACGAACGAACGGGTTGGCGAACCA
GTCAGGGTATTTTACTTCTCTTATCGATCCGATCTTGAGCAAACGAACGACCTCATCCT
GCACAGCTTTAGCCCTTTCCGGACCTAGCTTGCAGCGTTTTCTGCTTTATAGGTCCGGTAG
GTCGGATCGACGTTGAATTCGTGTGAGGTTACCTCGAAGCTTATTCCTTGGAGGTTTCGC
ACTTGACCAGGCGAACTGTCTTTGTTTTCTTTGAGAAAGGTTATGAAATCCTCTCGTAT
CGCTGGGTCGAGGTCTTGACCTACCCCAACACATTGTTTGGGATCTGATTCTGCTATGT
TAACCTGCGTTATCGATGATTTTTGAAGTCCGCATTGAGAAAGATCAGAAGATCTGACT
GTTTTTCAATTTCCGCGAATTTGTCTTCTATCATTAGTACTACGAATTCAGCTTTTTTGTAT
AAGCTCGTATCTGCCATATAGCATCTTCCGGAGCTGCGTTGGCTGCCGTATATGACTG
CTATCCCTTTTTCCGAAGGGAATTTTAATGCATTGGTGGTATGTTGATAGGACAGCCGT
TTCATGGCGTGTATCCATGGTCTTCCAGTATTGCGTTGAAAGGAGCTTTCTTGTCTAC
TACAAGAAAACCCGGTAAGTGGCGTTTTTCTGCCTCTGCCAGCTCGGAATCGAAGTGCC
CCATTTTTTTGAAAGCGTCGTAAAAGAATTTTTAACGTCAACTGAGCTGCCTGTTGTCT
ATCATTATACGAGAGAGTTTGTAGTTGCCTACATCGATTGAAATAACAATAGCGTCATC
GTGAGGTTTATCGAGGTCCGGTGGTTTTCGCTTTCCATAATGTGATCTTGGTGGTCCAGG
TCCCATCATCAGGCTCCTTGCCTTTTTGTGTAGGAATCTGCCTTCTCTGATGAGTTTTG
ATCGAATTGATTGAGTCGTTGCAGAGCTGAGACCCTCCCATGATGAAATCGATGCGTTT
CGGGGAGAGGTGCTCAGTTGCGCTTCTGCAGAAATTTGAGAGGCGTAATGGAAACCTCA
TGTTTTTTCTTATGCGGGGCGGAGGGAGGTGTGGTGAACCTGACCTTTCCTGTCTTGCCA
CATAGATCGAGTTCCGTAAAGAATAGGTATATGGCGTTCTCTCCTAGCGGGTGGGGGAG
AGTCACCTCCCTCCAGTGTTTCGTTCCCGTTTTCTTTGGTGGGGGAGTTGGCTCCTCCTTT
CCCTGTTTCGTGCGGAGGAGGGTTTTCAACTTTAGGGTTTGAACCTTTTTTTAGTTTTTGCC
CCCAGCTGCGATTAAATTTTTTTTACGGCGCGACACTCCTTCCGGGAATGTCTTCTTTT

GTTGTGGAACCTTGCAATATTTTCCGAGGTCAAGTTCTGGACTCTCGTCTTCAGCTGCTT
 CTTCTATTGCGAATAAGGTATGCTGCCCTTGCCTTTGGTTTTCCTTTGGGAAGGGGGTT
 TTAATTGCAAGAGGAGCATTTTTCGTTTTAGATTCTTTGAACTTTAAGGCCAGCACAGC
 GAGCTCTTCTTCGCCCTTTGGCGAAGTGGAGAGCTTTGTGAAGGGCATCGTATAAAGAGA
 TTGGTTGCCTAACAGTGAGTTCTCTCACGGAATGTAGAGGAAAACCACAGACCGTTACGA
 AGGGCTGCTATGGCTACCTCCTCGTTTTAGTTTCGCAATCTTCGATTTAATTGCTTTGGAA
 TTTTTCAACGTAAGATTTCGATGGGATTTCGTTTTGGAAGTAGAGTGAGACTCCAGAGATC
 GGCTTCAGACTCCTTTGTTCGATAAGTACAGAATCCTGTTTTAAGAAAGCGACCTGCCA
 GTTGTTAGGAGGTGTGACGTCTATCGAGGCGAAGTCTAGACGAATCTTTGGAGTAGGAT
 CGGTCGGGGAGACGGTACCTGTCAAATCTATGGGAACTGGGCAGGAATCCTAGTTTCGC
 GGGGTGGGAACTTCCTTACCCGCGGTATCGCGCTCTACCTCATGTCCAAGCGAGTGGA
 GTCGCATCTTTGCTGACCCACAATTGCGGCGAGGTTGCGATGCGCAATTTCTTGTTTT
 TGGTTAAGTTGATCAAGTTCATGATCATTGCGCGGAGCTCAGCCAATTCGCTGGCAGGT
 CGCGGAGGGACCTGATTAGACGGGTTGACCGGTTATCTCCATTTGGAGGTTGTATTGT
 GACGTTTTGTTCTCCGCCCTTGTTCGCGAGGGGAGTGATTTGGTCAGATCCAGAA

7.3.2.2 as 31 (2569 bp)

GATTCGTTTGGGAACGATCCGCGAAAAAGAAGAGTGATTCTAGTCAGTTGTGCACTTAG
 CAGCATGAATCTCCTCCTTAAAAAGACGTGTCTCGGATAAGGGACAAGAAAAGTGTA
 GAAAACCTCGGGCTTCATCGAGATTCGCGCCTCATATGGTGCAGTCCGGACCCAGTGTCT
 ATCGACACTAGAGTGTATCGATCGACACCCTCGTGAGGAAGTGTTTTTGAAAATAATGT
 TCGATACGGTGTGATCGACACTGGTTGGTGTGCGGTCGAAACCAACTCTGCAGCAGCGA
 GCTCCAACCTTGCTTTCAGTGTCCAATTCCTCCCAAACATCATCTCCTTCTCTGGATCAA
 GATGGCATTACAGCTCTTCCTTGGGTTAGCATCAGGATTACCAGGAAGAAAGCCCTCTT
 GTCTCGTAATGGAATTGTGAGTTGAGCTACCTGTACATCTAGCTTCTTCAAGTGAGAG
 ATCCAATTGTTGATCTTCCCATTAAGATCTATGTAGACAACATCAAACCTTGCCATTAAA
 CTCCGTA AACGAGACTTTGTTTTGGCTCTCAAGAATCTGCTCAAGCATGGATTCTATTCT
 GGTCTCTGATGAGAGGGTGGATGGGGGGCTTGGTAAACTGAATTACCATAGGTTCTCTG
 GAAGTTGCTCTAAGGAAATGTCTTCTAAAAAAGCAGGCTTTGGAGTGTAATTTGAGGAA
 CTCTAGTTCTTGTAAACCTGTTGCCACTCAAGTTGCCCTATGGCTGTCCAAACTTCTT
 GTAGCCCTGAACATCTATGTAAAAGACTCCCTTTTCTAACTCAGTCTCAGGCTCAATGG
 TCTCATCCTCAGCTGCAAGGTGGACATGCTTCTTCCCTGTAAGCAGGTTGTGAACCGAA
 TACAGCTTAGCATGTACCTCTGCCATTTGATTACCTGAAACAGTTCCAACAATCTTTTT
 CCTCTCAAAGTCAGCATTCTTGGTGTCTTACTACAAGCTAGGTTCTCTATCAGAGCAG
 TAGCATCAACTGGGATATCTCGTGTGGAAGTTACCATTGCTCGCTGCATCTAGAGCCAT
 TTGGTATCTCTAATCAACTCCATTGAAAAAGGTACCAAGAAGTTGTACCTCATTGAAGC
 CATGATGTGGGCAGTCCCGCTTGATACTCCTTCGAATCTAACCCATGCGGCTTTGAAGT
 CCTCAACTGGACCTTGTGTGAAGGTAGACAACCTTTGTTTCTTCAGCTTCTCAGACTTAG
 CATCATCATAGAAATTGTTCAAAAAGGCTATCTTGATGCTTCTTCAAACCTTGAAGGG
 AACCTATTTTGAGTTGTTTTAGCCACGAAGCAGAATCTCCAGCAAGTTGAATAAGGGAA
 GAGCTTGCACAGTAGGTAGTCTTCTGAGACTCCATTAACCTTTGATGCTCAGAACAAGT
 CTTCAAATGCTCGATATGGTCCATTGGCTGCTCATGAGACAAGCCATGGAAAGGATGC
 TGGCTTACTAAAGTAAAGTAGCCAGGCTTCAGCTCGTAATTGTTCTTCTCGAAGGGTGG
 GGATACGATCGCTGATCGATTAGCATATAACAGGTGAGGCCTGTTGAAACCCCAAGAA
 TACGACGTGGTTCCACAGCTGGAGGAATAGCTGCAGAAATCTGTGGTTGGTGTGATCG
 ACCCTCTGTTGTGTCGATGTACACCGATCACGTGTGCATTCTGCTCCTGAAGGTGTGCG
 ATCGACACCTAGTTGGTGTGCTTGGACACCGAGCCCTGCTGCGTCTXTAGCAGCTAATT
 GCGCATTGAGGCGGCCAGATCCTCATTGGCGTCTCAGGTAGGATGAGTGTATGATCA
 TCCAATTTCTGACCAGCCTCATTGTAGGCTTGCCCATCCTCATCTCTCAAACCTCGTTC
 TGCATCAGCTCGGATGACTCTTTGAGCATTGCTAACCAATATCACTAGTATATATAT
 AATTTTAATATAATTTTTAAATAGGGCCATTTAAAGTAATAAACTTTTTAATATAATAC
 TAATTAAGATCTAAATTTTATAACATTTTCAACAAGCAATCTCCATTTCTCATATAAC
 TAATATATTTTATATATTAACAAATAAAATTTAAATTTTCTATTTTCATGACTCAATTAC
 ACACATGTATATATATGAGGTCTATTAGCTGACAAAAAATTATTAATATATGATGAAAT

TGAATTAATAATAAGTATTATTATTATTTATATATGGACTTTTATGTTTTGTACGTATG
 GATTAAGATAAACAAAATTTATAATGTATGAGAATATAAAAATATATAATATAACATATA
 GTTTTTCTTTATAATATTTCTATTAACCAAATTGTAAGAGCTATATATATATATATAAT
 ACGGTTCAAATTATCCTAAGAACTACTCCCTCAAATAAGAGATCAATTGTAGTATTTAA
 GGATCGATTTCCACAAAGTTCTTTTGTTCACACAATGAATTCTGCAGCCCGGGGGATCC
 CAGGCCTCTTAGAGCGGCCCGCCTCCC

7.3.2.3 gs 31 (2265 bp)

AAAAATCAAGTATGTGGAGAGTCAATTTGCATCAACCTCTGCTCTGGAGCATCCTCAAC
 AACTTCCAGGGGAAAGCAGTCCAAAATCCTAAGGACTATGCTACTGGTAATGCTATCAC
 CATCCACCAAGAAGATGAGTCTCCACCAAGTCGTCGAGACTCCCATACTGAGGAAAATA
 TGATTCAAGAAGGGGGGGGGGATTCAACTCAAATTGCTGCACCTGCCACCCAGCAATTC
 ATCTGGACTATAAAGCACACACCTTTGCTGAATACTCACCATCCTGGAAAAAAGTYCAA
 AATTGAAAAACTCAAGAAGGATCAGAATCAAGAGATGATCAGCAGTGCCTCTACCCAAG
 AAACATACAAGAAGAAAAATCATCCAAGAAAAGCTTGATGATCCCGGTTCAATCACTCTG
 CCATGCTCTCTAAGACCATTGACATTCAATAGGTGCCTTTGTGATTTGGGAGCATCTGT
 TAGTCTTATGCCTCTCTCTACAGCTAAAAGGCTAAGGATTATGGAATACAAGTTCTGCA
 ACTTAGCTTTACTCTTAGCTGACGGTTCTGTAGCACATCCTCATGGTTTAATAGAAAAT
 CTGCCGTGTCAAGATTGAAAATGTTGAGATACCTACAGATTTTGTGGTATTGGACGTAGA
 TGAAAAAGGCAAGGATCCCCTGATCCTAGGCAGACCATTCTTAGCTTCAGCTGGAGCAG
 TCATAGATGTCAGGAACGGAAAAATTAATCTGAACCTTGAAGGCATCAAGATGAAGTTT
 GACATCAGAGAATCTTCATGGAAATCAACAACAGGAGTTCAGAATTTTGGGGTTCAGAA
 CATGGATGTTGATGAAGAAACAGAAGCTGGAACCCACCAAGAGTTAACTATACTTCTC
 AAGTTAGCAAGCTGAAGAGAACTTTTGACCACATGAAGAGAACAACAGAAAGGCTAGCT
 CAGACTGAAGATCCTACAGAAGATGACTGGTATGAGATGAGAAAAATAAATAAATGGCA
 AAGTAGAGTCATAGAAGGATTGTCTTCAAGGGTGATGAAGCTGAAAGATCAACTTTGGA
 TGTTTGAAAAAAGAGTTAAAAATTTTTCAAATGGTATTGATCTAGAGAAGATGGAAACT
 GCACTAGCTATCAAGGAGGACGAAAATTTCAACAACAGAATGGTTTGCACCAGAAGATCA
 GAAAGTAGTTCACTTAGAGGAAAAGAAGCCTAGAGAGCAAGTTGTTGAGTGAGGAGGAAA
 CTCCTATTCTGATTTATTCTCCAGGAGTGCCTAAACAGCCCACCAACACCTCTGCTGAA
 GAAGAATACTTGATTCATGATTCTGTTGCTGAAGAACTTCTGAAACTCTAAGTTGGATA
 CAAAGCTCAAGAGGATACCATTGGGGAGCTAACTTACAAGATGGAAGGACTGAACCATG
 CAATTCTTGAACTTCAAGAAATGGTTAAAGGCTACCCAGTTTCGAATTTTAGGATTG
 GTCAAAACAGCCTGGTTTCAGAAGAAGAAGACTACTCTGCTGATCAGAAAGAAGCTTATT
 TTGAAGAAAAATCAAATAAGTACTCATCCTACCATATCTCAAGAGACGATGCAGAGTAT
 GATTC AAGCACTGAAAATTCGACCACTGAAGATGAGGACTTCTCTGTTCTCTCTTCCA
 TCTCTTTTCTGACTAAGCTCTGTAAGAGTCAAGCTTAGTGACTTAAAAAAGCTCACTTG
 GGAGGAAGTCCCATGTCTATCCATGTATATAACTTTCTATTTTTCTTGTATTTTTGATG
 TTTTTGTTTAAGTGTTCAGGTTTATTTCAATTATTCAATTATGAGTCAACGGTTCGAGTA
 AATTC AACC AAGTCAAGAAAAATATCAACCGTTGATTGAGTAGAATGAGGAGGAGTAT
 TTTGAGCCATTGATTTGAATGACTAAATGTGGGAAATCACCTCACCGTTTGATCGAGT
 GACCAAGGGAGAAGTATTCTCCACCGTTGATGAAGTATTTCACTTCAAGAAGATAC
 TCGATTCTTTACTCGAGCTCTCATATTTCGAAGACCAAGGCATAACTTTGTGATCCAAC
 TTGCTCCCCAACACTCCTTTACTCAACGGACAAAGAGATTACCCATGCCCTCACCAT
 TTACAATCACTTTTTTATCCTTTCCCAAAGAAACCAAACACCATTTCTCTCACCCAT
 TG

7.4 GENSCAN und BLAST-Analysen

7.4.1 BAC F28L22

Predicted genes/exons; BLAST hits

Gn.Ex	Type	S	.Begin	...End	.Len	Fr	Ph	I/Ac	Do/T	CodRg	P....	Tscr..
1.08	PlyA	-	1133	1128	6							1.05
1.07	Term	-	4631	4455	177	2	0	55	45	212	0.994	15.30
1.06	Intr	-	5591	5447	145	1	1	63	72	136	0.904	13.86
1.05	Intr	-	6158	5858	301	0	1	106	40	230	0.968	19.87
1.04	Intr	-	8320	8219	102	2	0	49	20	76	0.574	1.13
1.03	Intr	-	8913	8841	73	0	1	2	79	67	0.926	0.26
1.02	Intr	-	9667	9334	334	0	1	84	86	169	0.637	15.95
1.01	Init	-	10549	10353	197	1	2	50	-44	194	0.550	5.85
1.00	Prom	-	11263	11224	40							-6.95

gb|AAC95175.1| (AC005970) putative reverse transcriptase
 [Arabidopsis thaliana] Length = 1352 Score = 106 bits (263),
 Expect = 3e-22 Identities = 52/93 (55%), Positives = 68/93 (72%)

2.00	Prom	+	13368	13407	40							-5.55
2.01	Init	+	15849	16007	159	2	0	63	97	92	0.602	12.47
2.02	Intr	+	17528	17616	89	1	2	51	27	61	0.028	-0.85
2.03	Intr	+	18014	18154	141	2	0	38	57	60	0.021	1.45
2.04	Intr	+	20704	21830	1127	1	2	95	83	1331	0.032	126.55
2.05	Intr	+	22017	22157	141	1	0	73	37	77	0.961	5.50
2.06	Term	+	22254	23821	1568	1	2	-12	34	1451	0.639	124.22
2.07	PlyA	+	24126	24131	6							1.05

Nia 2

3.05	PlyA	-	24312	24307	6							-3.64
3.04	Term	-	24413	24354	60	2	0	58	37	47	0.381	-1.57
3.03	Intr	-	26318	24819	1500	2	0	52	59	1090	0.285	96.45
3.02	Intr	-	28072	27780	293	2	2	11	53	279	0.413	17.63
3.01	Init	-	30138	30036	103	0	1	43	3	155	0.791	7.95
3.00	Prom	-	30338	30299	40							-10.25

gb|AAB81877|AAB81877 (AC002983) predicted transposon protein
 [Arabidopsis thaliana] Length = 907 Score = 877 bits (2242),
 Expect = 0.0

4.00	Prom	+	30955	30994	40							-7.15
4.01	Init	+	31881	32340	460	2	1	83	86	393	0.504	39.56
4.02	Intr	+	32623	32718	96	2	0	25	62	69	0.097	2.06
4.03	Term	+	32899	33749	851	2	2	-84	51	893	0.118	65.42
4.04	PlyA	+	34172	34177	6							1.05

gb|AAD23708.1|AC006918_6 (AC006918) putative Athila retroelement
 ORF1 protein [Arabidopsis thaliana] Length = 555 Score = 467
 bits (1189), Expect = e-131

5.00	Prom	+	34742	34781	40							-13.68
5.01	Init	+	35648	35709	62	1	2	37	15	88	0.636	2.17
5.02	Intr	+	35795	35984	190	2	1	48	-25	185	0.278	6.87
5.03	Intr	+	37199	37258	60	1	0	104	62	12	0.646	3.21
5.04	Intr	+	37906	37921	16	0	1	95	89	-7	0.354	-0.70
5.05	Intr	+	38046	38319	274	1	1	47	63	139	0.399	7.97

5.06	Intr +	38402	38463	62	2	2	40	107	25	0.978	2.06
5.07	Intr +	38555	38703	149	0	2	59	95	111	0.984	12.93
5.08	Intr +	38781	38901	121	2	1	57	18	5	0.824	-5.45
5.09	Intr +	38995	39150	156	2	0	89	90	130	0.834	17.36
5.10	Term +	41120	41241	122	0	2	58	33	110	0.259	5.16
5.11	PlyA +	41548	41553	6							-3.94

gb|AAD31371.1|AC006053_13 (AC006053) biotin holocarboxylase synthetase [Arabidopsis thaliana]^Length = 355 Score = 423 bits (1076), Expect = e-117

6.05	PlyA -	41634	41629	6							-1.75
6.04	Term -	41937	41784	154	2	1	38	36	145	0.125	5.71
6.03	Intr -	42880	42684	197	1	2	83	11	138	0.348	7.89
6.02	Intr -	43978	43808	171	1	0	39	69	66	0.938	4.02
6.01	Init -	44127	44056	72	0	0	80	53	80	0.850	9.82
6.00	Prom -	44300	44261	40							-10.35
7.07	PlyA -	45375	45370	6							1.05
7.06	Term -	46143	45869	275	1	2	61	34	264	0.808	17.95
7.05	Intr -	46679	46486	194	1	2	-33	75	209	0.683	10.71
7.04	Intr -	47009	46696	314	2	2	-69	-38	473	0.876	18.06
7.03	Intr -	47232	47036	197	1	2	54	49	157	0.435	11.61
7.02	Intr -	47685	47554	132	1	0	1	89	79	0.634	4.00
7.01	Init -	48544	47731	814	1	1	68	10	899	0.858	80.02
7.00	Prom -	48847	48808	40							-7.55

gi|3377816 (AF076275) contains similarity to Arabidopsis thaliana retrotransposon Athila (GB:X81801) [Arabidopsis thaliana] Length = 940 Score = 657 bits (1676), Expect = 0.0
gb|AAD15359| (AC006250) putative Athila retroelement ORF1 prote in [Arabidopsis thaliana] Length = 532 Score = 427 bits (1085), Expect = e-118

8.05	PlyA -	49412	49407	6							1.05
8.04	Term -	50077	49932	146	2	2	30	46	170	0.995	9.09
8.03	Intr -	50411	50256	156	0	0	81	94	28	0.986	6.76
8.02	Intr -	50527	50456	72	2	0	85	48	34	0.503	2.56
8.01	Init -	50731	50611	121	1	1	16	82	115	0.993	9.10
8.00	Prom -	50856	50817	40							-17.98
9.00	Prom +	50874	50913	40							-14.62
9.01	Init +	50955	51390	436	2	1	83	90	454	0.988	46.47
9.02	Intr +	51433	52243	811	2	1	-19	86	767	0.302	60.37
9.03	Intr +	52319	52544	226	2	1	36	19	309	0.511	20.86
9.04	Intr +	52585	53868	1284	0	0	44	30	1182	0.840	99.36
9.05	Intr +	53922	54046	125	2	2	71	-76	180	0.222	3.36
9.06	Intr +	54077	55708	1632	2	0	-38	92	1585	0.263	137.66
9.07	Intr +	55806	55956	151	0	1	55	-12	152	0.595	6.14
9.08	Term +	55964	56275	312	1	0	-40	43	385	0.884	20.22
9.09	PlyA +	56623	56628	6							1.05

gb|AAD22283.1|AC006920_7 (AC006920) putative reverse transcriptase [Arabidopsis thaliana] Length = 1787 Score = 948 bits (2424), Expect = 0.0

gb|AAD20433| (AC006446) putative pol polyprotein with a reverse transcriptase domain [Arabidopsis thaliana] Length = 889 Score = 808 bits (2064), Expect = 0.0

gi|4206306 (AF049110) prpol [Zea mays] Length = 1317 Score = 637 bits (1625), Expect = 0.0

10.07	PlyA -	57274	57269	6							-1.75
10.06	Term -	57766	57321	446	2	2	-7	43	592	0.265	44.61
10.05	Intr -	58062	57795	268	0	1	73	59	332	0.344	29.98
10.04	Intr -	58228	58134	95	2	2	32	-8	106	0.586	-0.84

10.03	Intr	-	58501	58353	149	0	2	33	68	202	0.445	16.66
10.02	Intr	-	58825	58717	109	2	1	48	-16	124	0.361	1.42
10.01	Init	-	59348	59156	193	2	1	80	25	307	0.981	27.78
10.00	Prom	-	59587	59548	40							-8.55

gi|3377830 (AF075598) contains similarity to *Aplysia californica* intermediate filament protein A (GB:X60297) [*Arabidopsis thaliana*] Length = 628 Score = 313 bits (794), Expect = 1e-84
 gb|AAD38224.1|AF147264_7 (AF147264) contains similarity to intermediate filament protein A; may be a pseudogene [*Arabidopsis thaliana*] Length = 786 Score = 263 bits (664), Expect = 2e-69

11.00	Prom	+	59831	59870	40							-11.74
11.01	Init	+	60075	60160	86	2	2	39	66	126	0.674	10.84
11.02	Intr	+	62071	62101	31	1	1	79	32	20	0.400	-2.29
11.03	Intr	+	62255	62388	134	1	2	71	79	82	0.609	9.12
11.04	Intr	+	62426	63521	1096	2	1	55	-12	1067	0.596	87.44
11.05	Term	+	63771	64358	588	2	0	79	54	583	0.808	52.53
11.06	PlyA	+	64496	64501	6							-3.64

gb|AAD30632.1|AC006085_5 (AC006085) Hypothetical protein [*Arabidopsis thaliana*] Length = 1295 Score = 470 bits (1196), Expect = e-131
 gb|AAD22505.1|AC007045_5 (AC007045) putative Tat1 pol polyprotein [*Arabidopsis thaliana*] Length = 1212 Score = 372 bits (946), Expect = e-102
 gi|4115365 (AC005957) reverse transcriptase-like protein [*Arabidopsis thaliana*]

12.00	Prom	+	64588	64627	40							-13.01
12.01	Sngl	+	64733	64939	207	1	0	31	54	377	0.898	28.54
12.02	PlyA	+	65026	65031	6							1.05
13.04	PlyA	-	65115	65110	6							1.05
13.03	Term	-	66614	66331	284	0	2	34	37	315	0.830	20.60
13.02	Intr	-	66845	66731	115	2	1	70	22	84	0.945	4.10
13.01	Init	-	67266	67066	201	0	0	66	113	49	0.534	9.12
13.00	Prom	-	68003	67964	40							-5.95

gb|AAB81877|AAB81877 (AC002983) predicted transposon protein [*Arabidopsis thaliana*] Length = 907 Score = 188 bits (472), Expect = 3e-47

14.06	PlyA	-	68080	68075	6							-0.45
14.05	Term	-	68629	68619	11	2	2	114	44	1	0.837	0.68
14.04	Intr	-	68858	68749	110	1	2	71	64	129	0.950	12.81
14.03	Intr	-	71874	71481	394	1	1	60	116	326	0.561	30.39
14.02	Intr	-	73491	72378	1114	0	1	29	28	689	0.331	50.37
14.01	Init	-	73619	73560	60	2	0	51	46	-10	0.360	-2.60
14.00	Prom	-	74700	74661	40							-5.25

emb|CAA57897| (X82547) putative movement protein [*Cherry callitovirus A*] Length = 463 Score = 44.1 bits (102), Expect = 0.003
 sp|P09520|VMP_FMVD MOVEMENT PROTEIN (CELL-TO-CELL TRANSPORT PROTEIN)
 gi|76779|pir||S01279 hypothetical protein 1 - figwort mosaic virus >gi|58809|emb|CAA29523| (X06166) ORF I (AA 1-323) [*Figwort mosaic virus*] Length = 323 Score = 43.4 bits (100), Expect = 0.004

15.00	Prom	+	75032	75071	40							-3.15
15.01	Sngl	+	75400	79470	4071	0	0	55	44	2924	0.944	279.79
15.02	PlyA	+	79502	79507	6							-6.75

pir||S05465 retrovirus-related polyprotein - *Arabidopsis thaliana* retrotransposon

Tal-3 >gi|16534|emb|CAA31653| (X13291) polyprotein
 [Arabidopsis thaliana] Length = 1291
 Score = 1885 bits (4828), Expect = 0.0
 emb|CAA20201.1| (AL031187) putative transposable element [Arabidopsis thaliana] Length=1308 Score = 1799 bits (4609), Expect = 0.0
 pir||S23319 hypothetical protein 2 - Arabidopsis thaliana retrotransposon Tal-2(strain Landsberg) (fragment)
 gi|16384|emb|CAA37924|(X53976) orf 2 [Arabidopsis thaliana]
 Length = 1084 Score = 1603 bits (4106), Expect = 0.0

16.00	Prom +	79714	79753	40							-3.15
16.01	Init +	81267	81338	72	2	0	63	79	58	0.543	8.52
16.02	Intr +	81519	81590	72	2	0	87	12	143	0.049	10.28
16.03	Intr +	81678	81893	216	2	0	-40	4	197	0.022	1.28
16.04	Intr +	82190	82633	444	1	0	29	77	421	0.031	32.77
16.05	Intr +	82684	83712	1029	0	0	20	-86	848	0.099	55.42
16.06	Term +	83773	84384	612	0	0	-36	46	552	0.093	37.09
16.07	PlyA +	85022	85027	6							-0.45

gb|AAD15357| (AC006250) putative Athila retroelement ORF1 protein [Arabidopsis thaliana] Length = 929 Score = 462 bits (1177), Expect = e-129

17.04	PlyA -	86336	86331	6							-1.75
17.03	Term -	86406	86338	69	0	0	34	48	55	0.463	-1.94
17.02	Intr -	86829	86652	178	2	1	38	38	171	0.705	11.10
17.01	Init -	87214	86892	323	1	2	19	-30	303	0.028	13.45
17.00	Prom -	88436	88397	40							-5.35

18.03	PlyA -	88655	88650	6							-0.45
18.02	Term -	89878	89617	262	0	1	53	49	219	0.657	13.51
18.01	Init -	90629	90154	476	2	2	90	67	139	0.823	11.99
18.00	Prom -	90749	90710	40							-10.55

gi|331554 (M90543) reverse transcriptase [Cauliflower mosaic virus] Length = 680 Score = 34.4 bits (77), Expect = 0.82

19.00	Prom +	90780	90819	40							-13.01
19.01	Init +	90832	91075	244	0	1	35	-38	280	0.976	13.04
19.02	Term +	91328	91413	86	0	2	51	44	227	0.651	16.44
19.03	PlyA +	92191	92196	6							1.05

20.00	Prom +	92510	92549	40							-11.93
20.01	Init +	92568	92703	136	2	1	54	-42	145	0.944	3.45
20.02	Intr +	92994	93664	671	1	2	88	-16	465	0.515	31.59
20.03	Intr +	93799	94002	204	0	0	8	-1	212	0.299	7.77
20.04	Intr +	94095	94182	88	2	1	36	41	122	0.783	5.92
20.05	Term +	94665	95050	386	1	2	46	40	277	0.399	17.77
20.06	PlyA +	96153	96158	6							-0.45

21.00	Prom +	96305	96344	40							-5.65
21.01	Init +	96350	96418	69	1	0	76	-3	123	0.499	8.10
21.02	Term +	96924	97055	132	2	0	107	38	77	0.580	6.71
21.03	PlyA +	97247	97252	6							-1.75

sp|P92969|RPOM_ARATH DNA-DIRECTED RNA POLYMERASE MITOCHONDRIAL PRECURSOR
 gi|1742971|emb|CAA70583| (Y09432) mitochondrial single subunit DNA-directed RNA polymerase [Arabidopsis thaliana]
 gi|2330566|emb|CAA70210| (Y09006) mitochondrial single-subunit DNA-dependent RNA polymerase [Arabidopsis thaliana]
 Length = 976 Score = 51.2 bits (120), Expect = 9e-07
 emb|CAA69331| (Y08137) mitochondrial single-subunit DNA-dependent RNA polymerase [Arabidopsis thaliana] Length = 976 Score = 51.2 bits (120), Expect = 9e-07

28.03	Intr	-	124716	122730	1987	1	1	42	33	1823	0.287	163.57
28.02	Intr	-	125241	124862	380	2	2	78	42	416	0.750	34.68
28.01	Intr	-	125801	125666	136	0	1	62	54	31	0.348	0.81

gb|AAD19780| (AC006528) putative pol polyprotein with reverse transcriptase domain (Pfam: PF00078) [Arabidopsis thaliana]
 Length = 841 Score = 779 bits (1991), Expect = 0.0
 pir|S66306 hypothetical protein 1 - Arabidopsis thaliana retrotransposon Athila
 gi|806535|emb|CAA57397| (X81801) Athila ORF 1 [Arabidopsis thaliana] Length = 935 Score = 724 bits (1848), Expect = 0.0

7.4.2 BAC F9D18

Predicted genes/exons; BLAST hits

Gn.Ex	Type	S	.Begin	...End	.Len	Fr	Ph	I/Ac	Do/T	CodRg	P....	Tscr..
-------	------	---	--------	--------	------	----	----	------	------	-------	-------	--------

1.01	Sngl	+	203	691	489	1	0	67	39	503	0.985	44.11
1.02	PlyA	+	763	768	6							1.05

put. Athila retroelement ORF1 protein 3e-58 gb/AAD15358/AC006250
 hypoth. protein 1 -At retrotransposon2e-38 pir/S66306/gi/806535

2.00	Prom	+	895	934	40							-17.34
2.01	Init	+	1069	1878	810	0	0	48	39	494	0.922	40.25
2.02	Intr	+	2113	3514	1402	0	1	68	98	1412	0.863	133.07
2.03	Term	+	3620	4044	425	0	2	96	51	338	0.875	30.37
2.04	PlyA	+	4446	4451	6							1.05

put. Athila retroelement ORF1 protein e-155gb/AAD23708/AC0069186
 similar to Vf retrotransposon like gene e-119gi/3319360/AF077408

3.03	PlyA	-	5076	5071	6							1.05
3.02	Term	-	9321	9146	176	1	2	-2	33	144	0.761	1.84
3.01	Init	-	9714	9645	70	0	1	89	66	85	0.952	12.66
3.00	Prom	-	10187	10148	40							-9.55

vacuolatin cytotoxin A; VacA Helicobact. pylori 3,9
 gi/4103994/4103996

4.05	PlyA	-	10357	10352	6							1.05
4.04	Term	-	14831	14669	163	1	1	19	36	195	0.697	8.83
4.03	Intr	-	15111	14997	115	1	1	69	30	115	0.463	7.39
4.02	Intr	-	16064	15468	597	0	0	80	59	285	0.631	21.38
4.01	Init	-	16934	16691	244	2	1	36	37	189	0.596	11.54
4.00	Prom	-	17046	17007	40							-7.95

putative protein At 7e-25 emb/CAA18234/AL022223
 contains revers transcriptase domain rvt PF/00078 3e-21
 gb/AAD21699.1/AC004793
 non-ltr retrotransposon reverse transcriptase-like protein At
 1e-21 emb/CAA66812/X98130

5.00	Prom	+	17380	17419	40							-10.84
5.01	Init	+	17621	17643	23	1	2	20	53	25	0.905	-4.64
5.02	Intr	+	17649	18551	903	0	0	14	-43	645	0.783	38.47
5.03	Intr	+	19140	19422	283	0	1	78	99	162	0.977	17.80
5.04	Term	+	21293	21427	135	1	0	97	48	113	0.974	10.24
5.05	PlyA	+	21537	21542	6							1.05

putative pol polyprotein At e-164 gb/AAD23707.1/AC006918
 putative retroelement integrase At 4e-82 gb/AAD15533/AC006217

6.13	PlyA	-	21550	21545	6							-1.95
6.12	Term	-	21865	21735	131	2	2	44	34	107	0.964	3.26
6.11	Intr	-	22038	21903	136	0	1	71	54	75	0.729	6.62
6.10	Intr	-	23256	22850	407	1	2	59	-29	431	0.640	26.64
6.09	Intr	-	23720	23635	86	1	2	60	25	27	0.521	-2.86
6.08	Intr	-	24451	24333	119	1	2	46	65	32	0.703	0.04
6.07	Intr	-	25201	25050	152	2	2	2	69	135	0.008	6.96
6.06	Intr	-	25694	25520	175	2	1	-17	37	75	0.009	-4.51
6.05	Intr	-	26330	25771	560	0	2	52	79	532	0.010	45.32
6.04	Intr	-	28268	26511	1758	0	0	7	2	1550	0.069	129.67
6.03	Intr	-	29020	28885	136	1	1	71	54	76	0.941	6.72
6.02	Intr	-	30194	29833	362	0	2	59	-29	341	0.902	18.51
6.01	Init	-	30639	30573	67	0	1	51	25	42	0.400	0.49
6.00	Prom	-	30789	30750	40							-11.44

put. Athila retroelement ORF1 protein 0,0
 gb/AAD15357/AAD15358/AC006250
 hypoth. protein 1 - At retrotransposon 0,0 pir/S66306/gi/806535
 put. Athila retroelement ORF1 protein 0,0 gb/AAD19759/AC006413

7.05	PlyA	-	30996	30991	6							-0.45
7.04	Term	-	34336	32471	1866	1	0	-4	41	1676	0.958	143.04
7.03	Intr	-	35208	34382	827	1	2	7	14	738	0.095	53.50
7.02	Intr	-	35960	35825	136	2	1	78	54	89	0.941	8.72
7.01	Init	-	37228	36773	456	1	0	10	-29	393	0.650	20.65
7.00	Prom	-	37339	37300	40							-10.35

hypoth. protein 1 - At retrotransposon 0,0 pir/S66306/gi/806535
 put. Athila retroelement ORF1 protein 0,0 gb/AAD15357/AC006250
 put. Athila retroelement ORF1 protein 0,0 gb/AAD19759/AC006413

8.00	Prom	+	37594	37633	40							-8.85
8.01	Init	+	37939	38039	101	0	2	64	103	77	0.514	11.58
8.02	Intr	+	38747	38793	47	2	2	26	30	51	0.514	-4.77
8.03	Intr	+	38958	39160	203	1	2	76	-35	191	0.646	8.68
8.04	Intr	+	40400	40992	593	1	2	45	29	567	0.879	41.97
8.05	Intr	+	41073	42169	1097	0	2	4	-31	1428	0.509	116.18
8.06	Intr	+	42263	43230	968	0	2	33	100	864	0.695	76.90
8.07	Intr	+	43339	43643	305	0	2	-70	-10	375	0.683	11.36
8.08	Term	+	43951	44146	196	1	1	37	51	327	0.902	24.80
8.09	PlyA	+	44571	44576	6							1.05

hypoth. protein 1 - At retrotransposon e-159pir/S66306/gi/806535
 put. Athila retroelement ORF1 protein e-147
 gb/AAD23706.1/AC006918
 put. Athila retroelement ORF1 protein e-140 gb/AAD19759/AC006413

9.00	Prom	+	46408	46447	40							-10.45
9.01	Init	+	46892	46950	59	1	2	68	67	70	0.915	8.73
9.02	Intr	+	47900	48114	215	2	2	44	59	179	0.776	13.04
9.03	Term	+	48445	48506	62	2	2	43	39	106	0.844	3.19
9.04	PlyA	+	48624	48629	6							1.05

uroporphyrinogen decarboxylase UDP 0,12 sp/p32347/DCup Yeast
 replication protein a 70 kd DNA-binding subunit RP-A RF-A 3,0

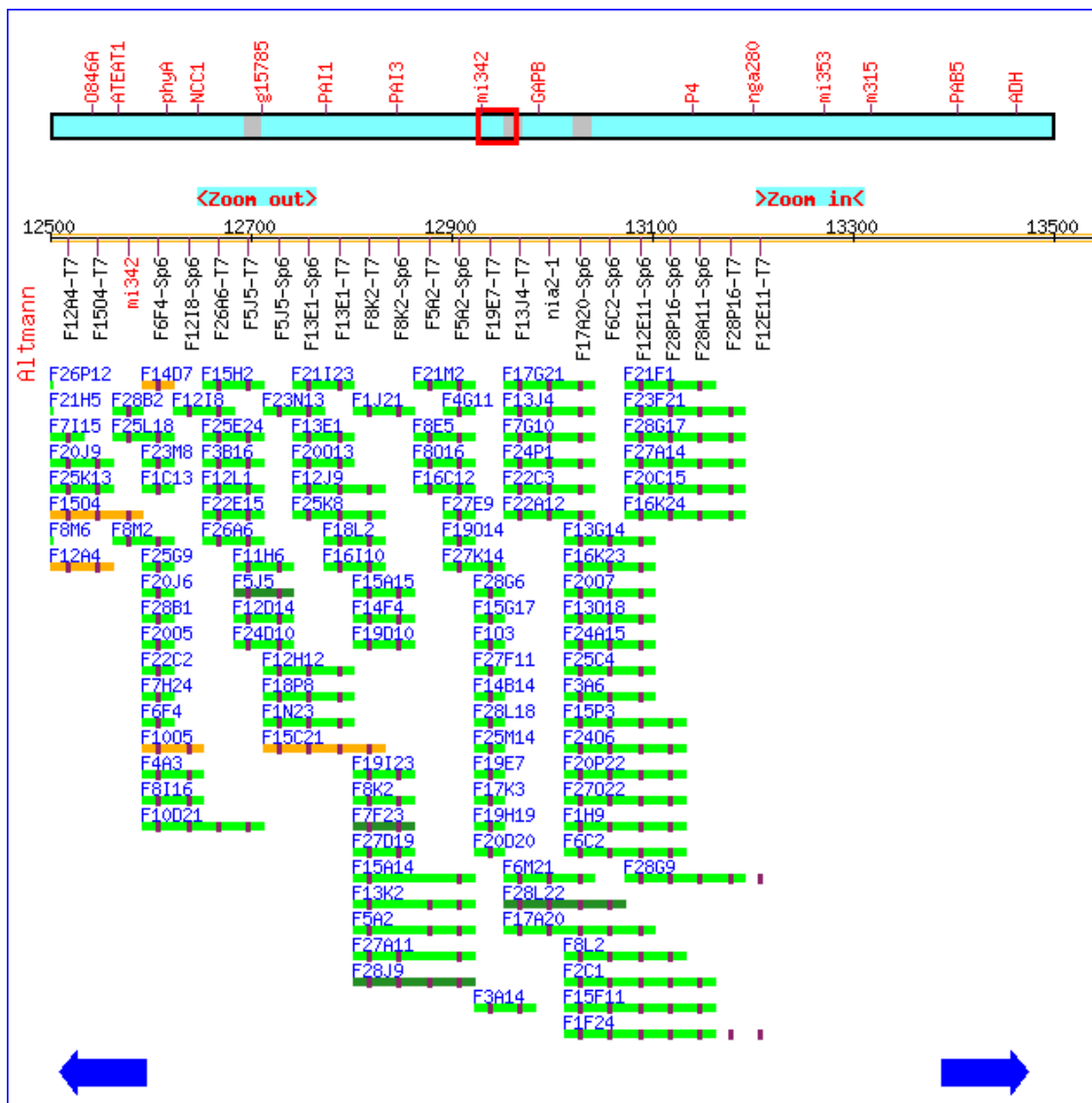
10.08	PlyA	-	48661	48656	6							1.05
10.07	Term	-	49136	48888	249	2	0	102	39	241	0.977	20.32
10.06	Intr	-	49360	49235	126	1	0	18	60	126	0.982	7.76
10.05	Intr	-	49550	49413	138	2	0	29	27	169	0.956	9.64
10.04	Intr	-	51371	51200	172	1	1	38	5	203	0.505	11.02
10.03	Intr	-	51531	51498	34	1	1	55	37	52	0.058	-1.84

10.02	Intr -	52447	52010	438	2	0	-11	-8	349	0.029	12.96
10.01	Init -	53375	53195	181	2	1	64	40	210	0.923	18.19
10.00	Prom -	53473	53434	40							-11.74
	putative reversetranscriptase At. 8e-43 gb/AAD15474/AC006067										
	putative Tat1 pol polyprotein At. 2e-42 gb/AAD22505.1/AC007045 5										
11.04	PlyA -	53719	53714	6							-3.44
11.03	Term -	54459	53866	594	0	0	29	45	294	0.298	17.74
11.02	Intr -	54868	54597	272	2	2	70	69	103	0.471	7.94
11.01	Init -	55530	55491	40	0	1	46	80	63	0.668	6.70
11.00	Prom -	55576	55537	40							-18.04
	putative pol polyprotein with reverse transcriptase domain Pfam: PF00078 At 2e-75 gb/AAD19780/AC006528										
	reverse transcriptase-like protein Vb. 2e-67 dbj/BAA22787/AB007466										
	put. Athila retroelement ORF1 protein 4e-58 gb/AAD15356/AC006250										
12.00	Prom +	55662	55701	40							-6.15
12.01	Sngl +	57199	59442	2244	0	0	38	43	1951	0.902	183.85
12.02	PlyA +	59458	59463	6							-3.94
	hypoth. protein 1 - At retrotransposon 0,0 pir/S66306/gi/806535										
	put. Athila retroelement ORF1 protein 0,0 gb/AAD15357/AC006250										
	put. Athila retroelement ORF1 protein 0,0 gb/AAD19759/AC006413										
13.00	Prom +	59466	59505	40							-16.41
13.01	Init +	59518	59696	179	0	2	48	107	231	0.353	24.78
13.02	Intr +	60098	61028	931	2	1	16	88	440	0.457	32.08
13.03	Intr +	61070	61314	245	1	2	-13	68	156	0.540	3.87
13.04	Intr +	62231	62528	298	2	1	-10	46	224	0.336	9.35
13.05	Intr +	63414	63546	133	2	1	50	17	119	0.399	5.30
13.06	Intr +	64160	64352	193	0	1	54	18	132	0.395	5.43
	put. Athila retroelement ORF1 protein e-155 gb/AAD15356/AC006250										
	reverse transcriptase-like protein Vb. e-148 dbj/BAA22787/AB007466										
	putative pol polyprotein with reverse transcriptase domain Pfam: PF00078 At e-131										

7.5 *Arabidopsis* Physical Map close-up

(<http://www.arabidopsis.org/cgi-bin/maps/>)

7.5.1 links von CEN 1 (Nia2)



7.5.2 rechts von CEN 1 (Gap B)

