Isolation, molecular characterisation and chromosomal location of repetitive DNA sequences in *Brassica*

Dissertation
zur Erlangung des Doktorgrades
der Fakultät für Agrarwissenschaften
der Georg-August-Universität Göttingen

vorgelegt von
Karla Galvão Bezerra dos Santos
geboren in Recife, Pernambuco, Brasilien

Göttingen, November 2004

D7

1. Referent: Prof. Dr. Heiko C. Becker
2. Korreferent: Prof. Dr. Wolfgang Köhler

Tag der mündlichen Prüfung: 18. November 2004

Table of Contents

List of Abreviations

A	adenine
bp	base pairs
BLAST	basic local alignment search tool
BSA	bovine serum albumin
C	cytosine
ca.	circa
Cot-1 DNA fraction	repetitive sequence fraction of genomic DNA
CTAB	cetryltrimethylammonium bromide
DAPI	4',6'-diamidino-2-phenylindole
DDBJ	DNA data bank of Japan
dH_2O	distilled water
DIG	digoxigenin
DMF	N,N-dimethylformamide
DNA	deoxyribonucleic acid
DNAse	deoxyribonuclease
dNTP	deoxyribonucleoside triphosphate
dTTP	deoxythymidine triphosphate
dUTP	deoxyuridine triphosphate
e.g.	*exempli gratia* (for example)
E. coli	*Escherichia coli*
EDTA	ethylenediaminetetraacetic acid
En/Spm	enhancer/supressor-mutator
et al.	*et alteri* (and others)
EMBL	European molecular biology laboratory
F	farad
FISH	fluorescent *in situ* hybridization
FITC	fluorescein isothiocyanate
g	gram
GenBank	genetic sequence database
G	guanine
GISH	genomic *in situ* hybridization
IPTG	isopropyl-ß-D-thiogalactopyranoside
h	hour
kb	kilo base pairs
l	litre
LB	Luria-Bertrani
M	molarity
m	meter
min	minute
M13 primer	universal promer

N	normality
2n	diploid chromosome number
NCBI	national center for biotechnology information
NOR	nucleolus organise region
OD_{600}	optical density at 600 nm
O/N	over night
ORF	open reading frame
P^{32}	radioactive phosphorus
pBo1.	plasmid library 1 from *B. oleracea* total DNA
pBo2.	plasmid library 2 from *B. oleracea* fractionated DNA
PBS	phosphate buffered saline
PCR	polymerase chain reaction
pH	negative logarithm of the hydrogen ion concentration in an aqueous solution
PVP	polyvinylpyrrolidone
rDNA	ribosomal DNA
RNAse	ribonuclease
rRNA	ribosomal RNA
rpm	rotations pro minute
RT	room temperature
S	Svedberg unity of sedimentation ($1S = 10^{-13}$ sec)
SDS	sodium dodecylsulfate
sec	second
SSC	standard saline citrate
TIGR	the institute for genomic research
T	thymidine
Tris	tris(hydroxymethyl) aminomethane
U	unit
V	volt
v/v	volume to volume
w/v	weight to volume
X-gal	5-bromo-4-chloro-3-indolyl-ß-galactosidase
YEB	yeast extract broth

Prefix for the units used in this work: p, pico (10^{-12}); n, nano (10^{-9}); µ, micro (10^{-6}); m, milli (10^{-3}); c, centi (10^{-2}); k, kilo (10^{3}); M, mega (10^{6}).

List of Figures

List of Tables

1. Introduction

The genus *Brassica* (tribe *Brassiceae*, family *Brassicacea*) includes six world-wide important cultivated species: *B*. *oleracea*, *B*. *rapa* (syn. *B*. *campestris*), *B*. *nigra*, *B*. *napus*, *B*. *carinata* and *B*. *juncea*. According to its use, they can be categorised in vegetable, oilseed (edible or industrial oils), condiment and forage crops, with some species presenting this whole range of use. Due in part to the many edible forms, *Brassica* crops are used in the cuisine of many cultures and are a valuable source of minerals, vitamins (particularly vitamin C), dietary fibres, and other possible salubrious factors such as anticancer compounds (Rosa, 1999). The seeds of *B*. *nigra* and *B*. *juncea* are utilised as condiment mustard and different cultivars of *B*. *oleracea*, *B*. *rapa* and *B*. *napus* are used in several countries as forage crops. *B*. *oleracea* and *B*. *rapa* comprise many of the vegetables in our daily diet. Several of these vegetables exhibit high variability in plant morphology, with root, stems, leaves, and terminal or axial buds being sometimes drastically modified (Becker *et al*., 1999). *Brassica* oil seed crops (*B*. *napus*, *B*. *rapa*, *B*. *juncea* and *B*. *carinata*) have become the world's most important edible oil source after soybean and palm, representing more than 11% of the global vegetable oil production in 2003 (FAO, 2003).

The cytogenetic relationships amongst the *Brassica* crops were established in 1935 by the Korean scientist U. This author showed that the three diploid species, *B*. *nigra* (2n= 16, genome BB), *B*. *oleracea* (2n= 18, genome CC) and *B*. *rapa* (2n= 20, genome AA) gave origin, through interspecific hybridisation, to the three amphidiploid species, *B*. *napus* (2n= 38, genome AACC), *B*. *juncea* (2n= 36, genome AABB), and *B*. *carinata* (2n= 34, genome BBCC) (U, 1935). U put these species in a triangle to explain their relationships, known after him as "U´s triangle" (see Figure 1). Comparative mapping in the *Brassica* genus has indicated that present-day diploid species are actually derived from hexaploid ancestors and that the chromosomal evolution in *Brassicaceae* seems to involve a high rate of chromosomal

rearrangements (Lagercrantz and Lydiate, 1996, Lagercrantz, 1998, Lukens *et al*., 2003).

Studies on genome organisation are important not only for our understanding of genome function and evolution, but also for the design of strategies for manipulating genomes (Lapitan, 1992, Heslop-Harrison, 2000). Cytogenetic techniques provide important information on genomic relationships based on karyotype and meiotic pairing analyses and on genome organisation at the chromosomal level. Analyses of meiotic pairing were, for example, one of the earliest tools used to assay the relationship between *Brassica* species (Moringa, 1928-1934, cited by Prakash and Tsunoda, 1980, U, 1935). Fluorescent *in situ* hybridisation (FISH) offers a powerful system to characterise the organisation of DNA sequences along the chromosomes. This technique has become a primary tool for the identification of chromosomes by using labelled DNA probe to hybridise to denatured DNA of chromosomes spread on microscope slides. A reliable identification of the chromosomes from crop species is important in several areas of plant breeding, as in the characterisation of substitution or addition lines, to assay chromosome variation following *in vitro* culture and regeneration and to integrate genetic and physical maps. Repetitive DNA sequences have been shown to be very useful for the chromosome identification and characterisation of genomes. Therefore, the investigation of DNA sequences that are repeated many hundred or thousand of times in the genome of crop plants may be a valuable tool for plant breeding.

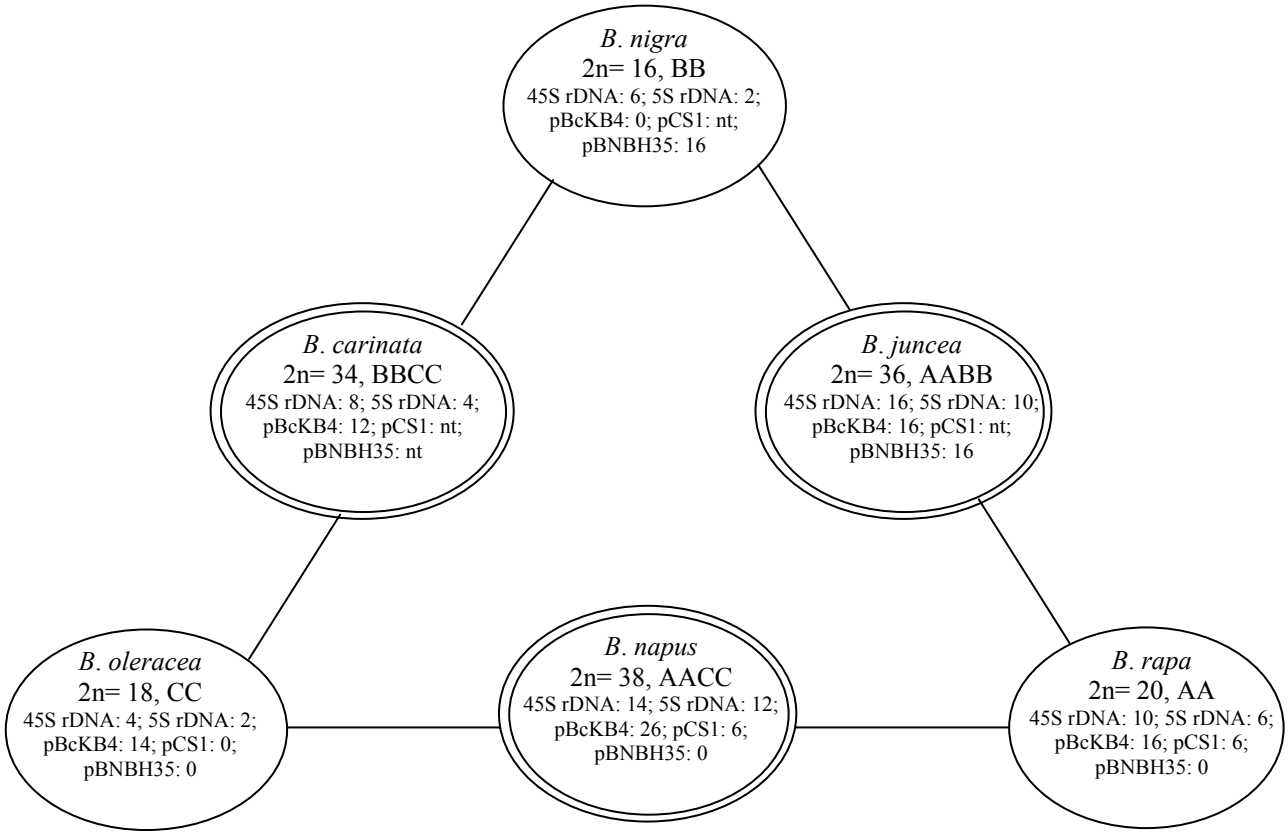


Figure 1. Summary of the relationships among cultivated *Brassica* species (after U, 1935, modified). The chromosome number, genome designations and number of chromosomes with hybridisation sites for repetitive sequences are given. The repetitive sequences indicated are: 5S and 45S rDNA (Hasterock *et al*., 2001), *Hind*III family (for repeat pBcKB4, Harrison and Heslop-Harrison, 1995), *B*. *rapa* species-specific repeat pCS1 (Iwabuchi *et al*., 1991 and Schrader, personal communication) and *B*. *nigra* species-specific repeat pBNBH35 (Schelfhout *et al*., 2004). nt= not tested or not mentioned.

1.1. Repetitive DNA sequences in plant genome

Several studies have shown that low copy sequences are generally conserved in the genomes of related species (e.g. Schmidt, 2002), whereas repeated sequences, that may be closely juxtaposed to such low copy sequences, are more rapidly changed during evolution (e.g. Li *et al*., 1995, Zhao *et al*., 1998a, Lapitan, 1992, Schmidt and Heslop-Harrison, 1998). In plants repetitive DNA sequences represent at least 20% and sometimes more than 90% of the total genome (Flavell, 1974, 1980).

In order to understand the structure and function of plant genomes, it is necessary to have a clear concept of the nature, topographical arrangements and genetic role of repetitive DNAs. Unlike in animal species, where repeated sequences have been studied since a long time, studies of higher-plant repetitive DNA are increasing in the last years and only in recent publications a detailed characterisation of these sequences, with respect to their features, genomic organisation, chromosomal localisation and evolutionary origins has been performed (e.g. Zhao *et al*., 1998a, Kubis *et al*., 1998, Nouzova *et al*., 1999). The biological function of most repetitive sequences seems to be complex and difficult to elucidate and is still a matter of debate. Although it has been suggested that they play no role (referred therefore as “junk DNA”) there is evidence that they may influence different aspects of cell development. While conclusive information about the function of repetitive DNA sequences is rare (e.g. Jasinska and Krzyzosiak, 2004), they may be involved in chromosome movement and pairing, recombination events, regulation of gene expression, determination of chromosome structure, genome response to environmental stimuli and physiological changes, cell proliferation and organism growth (see references in Lapitan, 1992, Xia *et al*., 1993 and Kubis *et al*., 1998). Repetitive sequences, in particular tandem arrays, are proposed to play a key role in stabilising DNA packaging and higher-order chromatin condensation (Heslop-Harrison, 2000) and were proposed to have a substantial impact on speciation (Bedbrook *et al*., 1980, Flavell, 1980). The molecular characterisation of repetitive DNA sequences and knowledge about their distribution should provide further insights into the organisation and evolution of plant genomes.

Many types of repetitive DNA sequences exist, some of them are ubiquitously distributed throughout the eukaryotic genomes, like microsatellites and transposable

elements (Kumar and Bennetzen, 1999, Bowen and Jordan, 2002, Zane *et al*., 2002). Microsatellites, also denominated as simple sequence repeats (SSR), are one of the most simple repetitive sequences, with only 1-6 bp repeat motifs (Zane *et al*., 2002). On the other hand, some repetitive sequences have very long repeat units, like the rDNA, that may be 6-40 kb long (Ahuja, 2001). However, a length of 160-180 bp, the extent of sequence wrapped around a single nucleosome, or 320-360 bp, corresponding to dinucleosomes, is frequently found for repeat units of repetitive DNA in both plants and animals (Harrison and Heslop-Harrison, 1995, Heslop-Harrison, 2000, Macas *et al*., 2002).

Individual sequence classes may represent 10% or even 50% of a genome (Kubis *et al*., 1998), but they vary widely in characteristics, like monomer size, GC content, copy number and localisation (see examples in PlantSat, 2004). Genomes of different species, even within the same family, can vary widely in genome size (see e.g. Bennett and Leicht, 1997, Bennett *et al*., 2000), reflecting the large variation in the proportion of repeated DNA sequences present in their genomes (Flavell, 1974, 1980, Heslop-Harrison, 2000, Ahuja, 2001). A particular sequence may be present in many species within a taxonomic family, or it can exhibit species-, genome- and even chromosome-specificity (e.g. Anamthawat-Jónsson and Heslop-Harrison, 1992, Bournival *et al*., 1994, Li *et al*., 1995, Wang *et al*., 1995, Zhao *et al*., 1998a, b, Ananiev *et al*., 2002) indicating that some repetitive sequences evolve rapidly, while others may be conserved.

Repetitive DNA elements in eukaryotic species are not uniformly distributed throughout the genome, some are arranged in tandem whereas others are widely dispersed. In Southern blot hybridisation, dispersed repetitive DNA sequences show hybridisation pattern typically as a smear (e.g. Santini *et al*., 2002, Shibata and Hizume, 2002, Dechyeva *et al*., 2003), whereas tandemly repeated sequences give typically a leader-like pattern of hybridisation (see e.g. Li *et al*., 1995, Houben *et al*., 2000, Dechyeva *et al*., 2003). Repetitive sequences, dispersed or tandemly repeated, including species-specific ones, with variable lengths of repeat units, have been extensively documented in plant genomes. A survey of the variety of plant DNA sequences tandemly repeated can be found in PlantSat (2004).

1.2. Isolation of repetitive DNA sequences

Presently the isolation of repetitive sequences, tandemly organised or dispersed in the genome, has been mostly obtained through the screening of plasmid DNA libraries. Normally genomic DNA is used as probe for colony or dot blot filter hybridisations, although satellite DNAs (Cot-1 DNA fraction) have also been utilised (for an example see Houben *et al*., 2000). In cases where species-specific sequences were searched, genomic DNAs from closely related species were used as probes on replica filters or on one filter with successive rounds of hybridisation. The DNA used to construct the library can be obtained from different sources. Genomic DNAs digested with a restriction endonuclease, or more frequently a specific size fraction of the digestion products, are the most frequent DNA sources. Several repetitive sequences were isolated in this way, for instance from *Beta procumbens* (Dechyeva *et al*., 2003), *Brassica* (Gupta *et al*., 1990, 1992, Kapila *et al*., 1996b, Itoh *et al*., 1991, Iwabuchi *et al*., 1991), *Gossypium* (Zhao *et al*., 1995, 1998a), *Helianthus* (Santini *et al*., 2002), *Medicago* (Calderini *et al*., 1997), *Musa* (Valárik *et al*., 2002) and *Vicia faba* (Frediani *et al*., 1999).

Prominent bands visible on agarose gel electrophoresis are very fruitful sources of repetitive DNA. In *Vicia faba*, for example, Nouzova *et al*. (1999) have cloned five new repetitive sequences by using DNA extracted from bands visible on agarose gels after digestion of genomic DNA with 34 different restriction enzymes. All newly described sequences were highly specific for *V. faba*, with little or no hybridisation to DNA of other *Vicia* species, and no hybridisation to DNA of other legumes tested. Several other examples can be found in the literature (e.g. *Beta*, Schmidt *et al*., 1991, 1998, Kubis *et al*., 1997, *Crepis capillaris*, Jamilena *et al*., 1993, *Ornithogalum longibracteatum*, Pedrosa *et al*., 2001).

In some cases, DNA obtained from chromosomes or chromosomal regions by microdissection were also used to construct plasmid DNA libraries, which were screened for repetitive sequences (e.g. Jamilena *et al*., 1995, Busch *et al*., 1995, Hizume *et al*., 2001, Shibata and Hizume, 2002). In *Allium cepa*, for instance, samples of chromosome 6 were dissected from metaphase plates and used as template in a DOP-PCR (degenerate oligonucleotide-primed polymerase chain reaction). The DNA obtained after this PCR (amplicon) was utilised to construct a

DNA library, which was screened for repetitive sequences. Genome-characterising sequences producing GISH (genomic *in situ* hybridisation)-like signals were isolated and investigated in *A*. *wakegi*, a natural hybrid between *A*. *cepa* and *A*. *fistulosum*. The chromosome organisation of these repetitive DNAs were analysed to gain an understanding of genome evolution in *Allium* (Shibata and Hizume, 2002).

There are, however, other ways to find repetitive sequences. In *Avena sativa*, for example, genome-specific repetitive DNA sequences were first identified in a random set of genomic DNA cosmid clones by gel-blot hybridisation, using labelled genomic DNAs from different *Avena* species as probes (Ananiev *et al*., 2002). RAPD markers specific for rye (*Secale cereale*) were identified as repetitive elements related to retrotransposons and could be used in FISH to characterise translocation lines from wheat-rye hybrids (Ko *et al*., 2002). In the genus *Vicia* two new families of tandem repeats were isolated by using a modified genomic self-priming PCR, a method that applies the genomic DNA itself as primer for PCR reactions (Macas *et al*., 2000).

1.3. Type and chromosomal localisation of repetitive DNA

Two classes of repeated DNA sequences can be easily recognised, as already mentioned: (1) tandemly repeated sequences, like telomeres and rDNA, and (2) dispersed repeated sequences, constituted mainly by transposable elements and remnants thereof, which can be thoroughly dispersed in the genome or enriched/depleted in some regions of the chromosomes. Intermediate forms of organisation were also found. Notably, telomeres, rDNA and retroelement sequences are all ancient - they are found in animals and plants, and might be considered as early derivatives of the 'RNA-world' from which DNA-based organisms evolved.

1.3.1. Tandemly repeated sequences

Tandem repeats consist of short repeat units arranged one after another and clustered in certain chromosomal domains, especially in centromeres and telomeres (Lapitan, 1992, Kubis *et al*., 1998). Sequences organised in this way were first isolated from genomic DNA by centrifugation, because they may form a distinct band (called "satellite band") separated from the main genomic DNA band in an equilibrium density gradient centrifugation, due to their GC-content which differs from the

average present in genomic DNA. Therefore, they are known as satellite DNA. Satellite DNAs consist of essentially identical repeating units, but some satellites contain more than one repeated sequence (Flavell, 1980). The lengths of the repeating units in satellite DNAs range from a few base pairs to thousands of base pairs, and the number of such repeats in the chromosomes ranges from a few hundred to over a million without interruption (Flavell, 1980). Satellite DNA is often localised in the constitutive heterochromatin, the permanently condensed and largely inactive portion of the chromatin, which is normally present in both homologous of a chromosome (Flavell, 1980, Avramova, 2002, Grewal and Elgin, 2002). However, satellite DNA in regions other than the heterochromatin was also found (Kamstra *et al*., 1997, Nouzova *et al*., 1999).

In plants the heterochromatic chromosomal regions are usually present in the subtelomere/telomere (telomere-associated sequences, TASs), around centromeres and, less frequently, at interstitial sites of the chromosomes. TASs are tandemly repeated sequences of high complexity, they are variable in nature and often exhibit a high variation of copy number even between closely related species or lines and are often species-specific. These repeats vary in length and degree of repetitiveness also between chromosomes in the same genome (see references in Chen *et al*., 1997 and Zhong *et al*., 1998).

Since they are closely associated with constitutive heterochromatin such tandemly repeated sequences are assumed to be important for chromosome structure (Heslop-Harrison, 2000, Vershinin and Heslop-Harrison, 1998). However, defined functions are only well established for two classes of tandemly organised repeats: the rDNA and telomeric sequences.

Ribosomal RNA genes (rDNA)

The ribosomal RNA genes (45S and 5S rDNA) encode the structural RNA components of ribosomes. The 45S rDNA loci consist of tandem arrays of repeating units containing the 18S, 5.8S and 25S rRNA genes and the transcribed and nontranscribed spacers, each unit being ca. 10 kb long in most plants. Hundreds or thousands of copies of these repeat units may be present, together representing up to ca. 10% of the genome (Heslop-Harrison, 2000). As a rule, 45S rDNA sequences are markers for the nucleolus organiser region (NOR), visible in mitosis as a

secondary constriction, whereas the 5S rDNA can be localised elsewhere in the genome. In some cases sites of 5S rDNA colocalise with those of 45S rDNA sequences. Loci for 45S rDNA are also found in chromosomes without secondary constriction. The rRNA genes are localised at one or more sites (loci) per chromosome set. The characteristic position of these genes along the chromosomes provides useful markers for chromosome identification (see section 1.4). The units themselves are highly conserved, and probes isolated originally from wheat can be used to localise the 45S and 5S rRNA genes in most eukaryotic species. Changes in chromosomal distribution of the units generally correlate with the rates of speciation, and they have been used to examine evolutionary relationships between related species (see section 1.4).

Telomeric sequences

Telomeres, the nucleoprotein structure localised at the very ends of eukaryotic chromosomes, play a critical role in maintaining chromosomal stability and function. The telomeric sequences, at the ends of most plant and animal chromosomes, allow a linear replication unit to be maintained, protect chromosome ends and overcome the 'end replication problem' (ends of linear DNA cannot be replicated completely during DNA synthesis, and chromosomes will naturally shorten with each successive round of DNA replication if no mechanism to avoid that is available). In addition to stabilising chromosomes, telomeres are proposed to play important roles in nuclear architecture and chromosome organisation (Biessmann and Manson, 1997, Pardue and DeBaryshe, 1999, McKnight *et al*., 2002, Riha and Shippen, 2003).

Telomere repeats are remarkably conserved among eukaryotes, and sequence variation among most of the major taxonomic groups does not exceed one or two nucleotides (Li *et al*., 2000). Angiosperm telomeric DNA was first isolated from *Arabidopsis thaliana* by Richards and Ausubel (1988), who showed that *Arabidopsis* telomeres consisted of repeats of a 7-bp motif, the 5´-TTTAGGG-3´. Since then it has been shown that the majority of plant species possess the *Arabidopsis*-type telomeric sequences at the chromosome ends, however repeat array length varies significantly among plants and even between different strains and different chromosomes within a species (Richards, 1995, Fuchs *et al*., 1995, Vershinin and Heslop-Harisson, 1998, Zhong *et al*., 1998), e.g. from 2-5 kb in *A*. *thaliana* (Richards and Ausubel, 1988), to 13-223 kb in tomato (Zhong *et al*., 1998). The presence of telomeric repeats at

nontelomeric locations has been detected in a variety of species, and in several cases this is associated with chromosome fusion-fission processes that occurred during evolution (Richards *et al*., 1991, Richards, 1995, Fuchs *et al*., 1995).

The *Arabidopsis*-type telomeric sequence was initially thought to be ubiquitous in plants, with the family *Alliaceae* being the only exception (Fuchs *et al*., 1995, Pich and Schubert, 1998). However, recently plants from several monocot genera (Adams *et al*., 2000a, 2001, Weiss and Scherthan, 2002, Sýkoravá *et al*., 2003b, Weiss-Schneeweiss *et al*., 2004) and in three dicotyledonous species, *Cestrum*, *Vestia* and *Sessea* (Sykorava *et al*., 2003a), were reported as lacking the TTTAGGG repeat. In several cases the typical *Arabidopsis*-type telomeric sequence has been partially or fully replaced by the human type telomeric sequence $(TTAGGG)_n$ (Sýkoravá *et al*., 2003b). This was, however, not the case in *Allium cepa* (Sýkoravá *et al*., 2003b). In this species a satellite DNA repeat, and possibly rDNA or transposon-like sequences, represent the most probable candidate sequences that may have acquired telomeric function(s) (Pich and Schubert, 1998).

1.3.2. Dispersed repeated sequences - Transposable elements

Transposable elements (TEs) are discrete segments of DNA that are distinguished by their ability to move and replicate within genomes. These sequences have been found in all genomes in which they have been sought and are ubiquitous in plants. TEs play a major role in plant gene and genome evolution, they contribute substantially to the structure, size and variability of plant genomes. TEs can comprise over 50% of the nuclear DNA content in many higher plants with large and complex genomes (Kumar and Bennetzen, 1999, Kidwell, 2002). In eukaryotes, TEs have been divided into two broad categories, depending on their mode of transposition. Class 1 (RNA) elements move via an RNA intermediate, which is reverse transcribed prior to its integration into the genome. RNA elements, widely known as retrotransposons, have either long terminal repeats (LTR retrotransposons) or terminate at one end with a poly(A) tract (non-LTR retrotransposons: long and short interspersed nuclear elements- LINEs and SINEs, respectively). LTR retrotransposons have been further classified as either *Ty1/copia*-like or *Ty3/gypsy*-like elements, based on the order of their encoded proteins that include a reverse transcriptase and integrase required for reverse transcription and integration. Class 2 (DNA) elements move via a DNA intermediate, by a cut-and-paste mechanism. They

are flanked by terminal inverted repeats and have subterminal repeated sequence motives, that are essential for their autonomous activity, whereas proteins encoded by the internal part of the element, for example transposase, may also be provided by independent transposon copies *in trans*. Based on the similarity of the transposases, the proteins that catalyse transposition and integration, DNA transposons have been grouped into superfamilies like *Tc/mariner*, *hAT*, *Enhancer/Supressor-mutator-* (*En/Spm*), *Mutator-* (*Mu*) and PIF/Pong elements (see references in Staginnus *et al*., 2001 and Zhang and Wessler, 2004).

Retrotransposons, or their remnants, represent a major fraction of interspersed repetitive DNA in eukaryotes, especially in plant species. Their amplification and dispersion contributed to genomic plasticity, perhaps allowing plants the adaptive response to environmental stresses made necessary by their inability to move (Kumar and Bennetzen, 1999).

1.4. Repetitive DNA sequences as cytogenetic tools

Repetitive sequences can assist in the construction of a karyotype with unambiguous identification of chromosomes, a requirement for the integration of genetic and physical maps.

45S and 5S rDNA sequences have since a long time been used as markers for chromosome identification, improving the construction of karyotypes of several plant species (e.g. *Picea glauca*, Brown *et al*., 1993, *Beta vulgaris*, Schmidt *et al*., 1994, *Pinacea*, Lubaretz *et al*., 1996, *Phaseolus*, Moscone *et al*., 1999, *Nicotiana*, Kitamura *et al*., 2000, *Lilium*, Lim *et al*., 2001). Together with information about chromosome length and centromere position these sequences can characterise the chromosome complement of some species with small numbers of chromosomes like *Arabidopsis thaliana* (2n= 10, Murata *et al*., 1997) and barley (2n= 14, Brown *et al*., 1999). The rDNA genes were also useful to gain insights into the phylogeny and chromosomal evolution of different groups of species, e.g. *Hordeum* (Molnar *et al*., 1989), *Triticum* (Jiang and Gill, 1994), *Vicia* (Raina and Ogihara, 1995, Raina *et al*., 2001), *Allium* (Lee *et al*., 1999), *Aloe* (Adams *et al*., 2000b), *Clivia* (Ran *et al*., 2001), and *Passiflora* (de Melo and Guerra, 2003). Actually, the repetitive DNA sequences most

widely used for chromosome identification and analyses of phylogeny and chromosomal evolution, both in plants and animals, are the rDNA genes.

Apart from the rDNA repeats, other repetitive DNAs have also been used as tools for the identification of chromosomes in different species, e.g. rye (Cuadrado *et al*., 1995), maize (Chen *et al*., 2000a, b, Sadder and Weber, 2001), pea (Neumann *et al*., 2001), lentil (Galasso *et al*., 2001) and *Vicia* (Navrátilová *et al*., 2003). In norway spruce, for example, the chromosomes are metacentric and the identification of homologous pairs is hampered by similar length and only slight differences at the morphological level. Vischi *et al*. (2003) were able to construct an unambiguous karyotype for this species by using three repetitive DNA sequences as probes in FISH combined with chromosome measurements.

The potential to characterise specific chromosomes makes repetitive sequences also very useful for the analyses of transgenic plants. Traditionally these plants are characterised using phenotypic and Southern analyses, but FISH has been efficiently used for visualisation, chromosomal localisation and estimation of numbers of transgene loci in transgenic plants (see for example Moscone *et al*., 1996, Dong *et al*., 2001, Snowdon *et al*., 2001). Recent results obtained in transformed plants from different species indicate that FISH, using the transgene sequence together with repetitive sequences as probes, is also a powerful tool for characterisation of transgene locus structure that significantly augments conventional Southern analysis in the evaluation of transgene plant material (Fransz *et al*., 1996, Svitashev *et al*., 2000, Jin *et al*., 2002, Kohli *et al*., 2003).

Karyotypic analysis using cloned repetitive sequences has been shown to be a powerful approach to the study of phylogeny and has some advantages over molecular studies based on sequence data, as it examines evolutionary changes over several different loci with little or no sequence similarities, and with potentially independent evolutionary constraints and cellular function (Lim *et al*., 2000). The analysis of a great number of repetitive sequences in the genome of related species can provide important informations concerning their relatedness. In the genus *Nicotiana*, for example, ten repetitive sequences, including the rRNA genes, some satellite DNAs and sequences related to geminiviral DNA, were used to establish a hypothesis about the phylogenetic relationships in this group (Lim *et al*., 2000). The analysis of the chromosomal distribution of these sequences in different species

allowed the characterisation of chromosomes, the identification of homeologous chromosomes and the construction of a phylogenetic tree for this genus.

Moreover, repetitive sequences are useful for studying the behaviour of plant genomes. Polyploidization, for instance, has played a major role in the evolution of many plant species like tobacco, cotton, and wheat. However, little is known regarding the subsequent evolution of DNA sequences after being combined in a common nucleous. Such subsequent processes were analysed in the polyploid cotton *(Gossypium barbadense* L., AD genome) by using several dispersed repetitive sequences specific to Old World (A genome) and New World (D genome) diploid ancestors (Zhao *et al*., 1998b). The presence of (otherwise) A genome specific repetitive sequences in the D genome of *G*. *gossypioides* and in the polyploid *G*. *barbadense* provided genome-wide support for the proposal that this diploid species, rather than its sister species *G*. *raimondii*, may be the closest living descendent of the New World D genome cotton ancestor (Zhao *et al*., 1998a, b).

A clear utility of repetitive DNA sequences in plant breeding is in the detection of chromatin introgression following sexual or somatic hybridization, especially in the form of chromosome addition lines, at a very early stage of development (e.g. *Triticeae*, Anamthawat-Jónsson and Heslop-Harrison, 1992, Bournival *et al*., 1994, *Solanaceas*, see references in Rokka *et al*., 1998, *Medicago*, Calderini *et al*., 1997, rye-wheat hybrids, Ko *et al*., 2002). Cao and Sleper (2001), for example, have successfully used the TF436 genome-specific repetitive DNA sequence in Southern blot hybridisation to monitor chromatin introgression from *Festuca mairei* into *Lolium perenne*.

GISH is one of the most popular and effective techniques for detecting alien chromatin introgression into breeding lines, however GISH analysis alone does not reveal the genetic identity of the alien chromosomes, whereas repetitive sequences that identify individual chromosomes can do.

Repetitive sequences may also be useful for other aims, as analyses of meiotic behaviour in hybrids (e.g. Jenkins *et al*., 2000), analyses of somaclonal variation caused by *in vitro* culture (e.g. Kubis *et al*., 2003), to study structural alterations, such as amplification and interchanges in chromosomes (see references in Rokka *et al*., 1998).

Repetitive sequences are frequently used as probe in Southern or *in situ* hybridisations, but they can also be transformed to PCR based markers, depending on the nature of the sequence and the objective of the study.

1.5. Molecular cytogenetics of *Brassica* species

Detailed cytogenetic analyses of *Brassica* species have been severely hindered by the small size and uniform morphology of their chromosomes. However, the development of molecular cytogenetics, allowing a better characterisation of chromosomes with small size and limited morphological differentiation, has opened the opportunity to better characterise the genome of *Brassica* crops at the cytological level. In the past 10 years *in situ* hybridisation has been successfully applied in *Brassica* species by using single and repetitive sequences as well as total genomic DNA as probes.

The genome structure of the amphidiploid species from the "U´s triangle" was analysed by GISH, applying genomic DNA from *B*. *nigra*, *B*. *oleracea* or *B*. *rapa* as a probe (Snowdon *et al*., 1997a, Bellin and dos Santos, 2002). The B genome was successfully differentiated from the A and C genomes in *B*. *juncea* and *B*. *carinata*, however, the differentiation of the A and C genomes in *B*. *napus* failed, due to the high degree of cross-hybridisation between the chromosomes and the genomic DNA of both genomes. GISH was also used to identify donor chromatin in hybrids between *B*. *napus* and *Eruca sativa* (Fahleson *et al*., 1997), *Raphanus sativus* (Snowdon *et al*., 1997a, 1999), *Lesquerella fendleri* (Skarzhinskaya *et al*., 1998), *Sinapis arvensis* (Snowdon *et al*., 2000b), synthetic rapeseed (*B*. *carinata* x *B*. *rapa*) (Li *et al*., 2004), *Crambe abyssinica* (Wang *et al*., 2004). In all these studies, the addition of donor chromosomes to the *B*. *napus* genome was successfully assessed by GISH, but only in few cases intergenomic translocations could be detected (Snowdon *et al*., 1999, Skarzhinskaya *et al*., 1998), although Southern blot hybridisation had revealed the presence of donor DNA also in hybrids with only 38 chromosomes (Skarzhinskaya *et al*., 1998). A small amount of donor DNA, probably too small to be detected by GISH, was given as possible reason for the absence of a GISH signal in the hybrids. However, the fact that *Brassica* chromosomes are frequently labelled only at the centromeric regions after GISH, which would hamper the detection of small

translocations in the distal regions of the chromosome arm, may also have contributed for the failure to detect alien DNA.

Besides the analysis of genome structure of the amphidiploid species and hybrids, analyses of genome evolution in *Brassica* were performed by using *Arabidopsis* BAC clones as FISH probe in *B*. *rapa*, an approach called comparative FISH mapping (Jackson *et al*., 2000). The results found by these authors, i. e. chromosomal duplication playing the major role in the evolution of the *B*. *rapa* genome, supported the hypothesis that *Brassica* diploids are secondary polyploids, as indicated by comparative genetic mapping results (Lagercrantz and Lydiate, 1996, Lagercrantz, 1998) and by cytogenetic analyses (Röbbelen, 1960, Armstrong and Keller, 1981, 1982).

In situ hybridisation has also been applied to localise repetitive and even single copy sequences on *Brassica* chromosomes. In *B*. *rapa* and *B*. *napus*, single copy genes from the self-incompatibility locus were visualised by FISH (Iwano *et al*., 1998, Kamisugi *et al*., 1998). In *B*. *oleracea* var. *alboglabra* an integration between the cytogenetic and linkage maps was performed by using FISH with different combinations of probes containing *Brassica* DNA sequences that have been genetically mapped (Howell *et al*., 2002). Most of the probes were large genomic DNA fragments inserted in bacterial artificial chromosomes (BACs) and required the use of repetitive DNA (Cot-1 DNA fraction) to reduce nonspecifc hybridisation. The use of such an approach for the localisation of a DNA sequence that cannot readily be mapped by genetic studies can be very time consuming. Several rounds of FISH, increasing the chance to loose important cells, may be required in some situations to identify the chromosome where the target sequence is localised. The availability of one or a few repetitive sequences, able to identify all the chromosomes, would be a simpler alternative for this purpose. Moreover, repetitive probe(s) from A and C genome should be more easily assigned to *B*. *napus* chromosomes.

Concerning the repetitive sequences analysed in *Brassica* crops, the localisation of rRNA genes, especially the 45S rDNA cistrons, has been reported by different authors (Maluszynska and Heslop-Harrison, 1993, Snowdon *et al*., 1997b, 2000a, 2002, Fukui *et al*., 1998, Armstrong *et al*., 1998, Hasterock and Maluszyska 2000a, b, Hasterock *et al*., 2001, Schrader *et al*., 2000, Ziolkowiski and Sadowski, 2002). In Figure 1 the number of rDNA loci in *Brassica* species from the "U´s triangle" is

indicated. The number and distribution of rDNA in the chromosomes of these *Brassica* species is highly variable. Furthermore, the number of rDNA loci of amphidiploid species is not always equal to the sum of loci in the ancestral species (Maluszynska and Heslop-Harrison, 1993, Snowdon *et al*., 1997b, Hasterock and Maluszyska 2000a) and intraspecific variation for the number of rRNA genes loci seems to exist (e.g. Armstrong *et al*., 1998, Ziolkowiski and Sadowski, 2002).

The identification, and in few cases the chromosomal localisation, of other repetitive DNA sequences has also been reported for *Brassica* crops. So far, only in *B*. *napus* a transposon-like DNA, the S1 SINE retrotransposon, has been localised by FISH (Goubely *et al*., 1999), although several transposon-like sequences have been identified in the genome of *B*. *oleracea* by database comparison (Zhang and Wessler, 2004). A highly repeated *Hind*III DNA sequence family was identified in the early 80ths and repeats from this family were isolated several times from *Brassica* species with the A and C genomes (Benslimane *et al*., 1986, Reddy *et al*., 1989, Harbinder and Lakshmikumaran, 1990, Lakshmikumaran and Ranade, 1990, Xia *et al*., 1993, 1994, Harrison and Heslop-Harrison, 1995). Harrison and Heslop-Harrison (1995) studied a representative of this family in the three diploid species and in their amphidiploids through Southern blot and *in situ* hybridisations. They found that this satellite DNA was present on a different number of chromosomes in *B*. *rapa* and *B*. *oleracea* (indicated in Figure 1), but was absent from the *B*. *nigra* genome. Xia *et al*. (1993), however, found a small amount of this sequence in the *B*. *nigra* genome through Southern blot hybridisation. Differences in the hybridisation procedure were appointed as responsible for this discrepancy (Harrison and Heslop-Harrison, 1995).

Five other repetitive sequences, one tandemly repeated (pBN34) and four dispersed (pBNBH35, pBNMbo5, pBN4, pBNE8), have been isolated from the *B*. *nigra* genome (Gupta *et al*., 1990, 1992, Kapila *et al*., 1996b). Southern blot hybridisations with different members of the *Brassicacea* were performed with these repeated DNA sequences. Two of them, pBNBH35, and pBN4, were found to be absent from the A and C genome in high stringency hybridisations, but hybridised to *Sinapis arvensis* DNA, showing that *B*. *nigra* has a closer relationship with the *S*. *arvensis* genome than with *B*. *oleracea* and *B*. *rapa*, which is in agreement with other studies (Prakash and Hinata, 1980, Song *et al*., 1988). Recently, one of these five *B*. *nigra* repetitive DNA sequences was analysed by FISH (Schelfhout *et al*., 2004). The subfragment of

pBNBH35 used as FISH probe in *Brassica* crops with the A, B and C genomes hybridised at the centromeric/pericentromeric regions of all B genome chromosomes, but was absent from A and C genome chromosomes.

In *B*. *rapa*, Iwabuchi *et al*. (1991) isolated a repetitive sequence, the pCS1, which was completely absent from the genome of the closely related *B*. *oleracea*. Using radioactive *in situ* hybridisation (ISH), this A genome specific sequence was localised at the centromeric regions of three pairs of *B*. *rapa* chromosomes. Itoh *et al*. (1991) used a variant form of pCS1, the pCT10, for ISH. This variant contains an additional sequence that hybridised also with *B*. *oleracea* DNA in Southern blots, but was specific for the A genome at the chromosomal level. By using a mixture of pCS1 and pCT10 as ISH probe four pairs of *B*. *rapa* chromosomes could be identified and the asymmetric nature of *B*. *rapa*/*B*. *oleracea* somatic hybrids could be demonstrated.

1.6. Objectives

The objectives of this work were to find repetitive sequences able to differentiate at the chromosomal level the highly homeologous A (*B*. *rapa*) and C (*B*. *oleracea*) genomes of *Brassica* in order to assist in the physical identification of *B*. *napus* chromosomes. For these purpose C genome specific repetitive sequences, tandemly organised or dispersed in the genome, were searched in genomic DNA libraries of *B*. *oleracea*. Selected sequences were analysed in the genome of *B*. *oleracea*, *B*. *rapa* and *B*. *napus* by Southern blot and fluorescent *in situ* hybridisations, and their sequences were compared with DNA databases.

The discovery of repetitive sequences specific to the C genomes of *Brassica*, besides providing information on chromosome and genome evolution in this group, might allow the identification of some or all C genome chromosomes from *B*. *napus*. The ability to identify the chromosomes in *B*. *napus* is not only important for the future integration of genetic and physical maps, but also for the genomic localisation of transgene inserts in genetically transformed oilseed rape.

2. Plant Materials

The *Brassica* species and cultivars used in this work, with their genome composition, source and accession number, are listed in Table 1.

Table 1. Plant material.

Species	Subspecies	Cultivar	Genome (2n)	Source/ accession number
B. oleracea	*italica*	Vitamina	CC (18)	IPK [a]/ BRA 1425/77
	capitata	Market Victor		IPbPz [b]/ 2457
	oleracea			IPbPz/ 3507
	alboglabra			IPbPz/ 165/83
B. incana			CC (18)	IPbPz/ 61
B. rapa		NPZ-35*	AA (20)	NPZ [c]
		Perko	AAAA (40)	KWS [d]
		Rex	AA (20)	NPZ
B. napus		Express	AACC (38)	NPZ
		Iris		NPZ
		Jockey		NPZ
B. nigra			BB (16)	IPK/ CR 2093/79- BRA 21
B. juncea			AABB (36)	IPK/ CR 2425/82- BRA 57
B. carinata			BBCC (34)	IPK/ BRA 489/77

*breeding material from NPZ[c]

[a] IPK- Institut für Pflanzengenetik und Kulturpflanzenforchung, Gatersleben

[b] IPbPz- Institut für Pflanzenbau und Pflanzenzüchtung, Göttingen

[c] Norddeutsche Pflanzenzucht Hans-Georg Lembke KG

[d] KWS Saat AG

3. Methods

3.1. Plant DNA extraction and quantification

Plant genomic DNA was isolated using a CTAB-extraction method modified from Rogers and Bendich (1988) (Section 3.1.1). The DNA concentration was measured with a Bio-Rad VersaFluor™ Fluorometer (Bio-Rad Laboratories, CA, USA) (Section 3.1.2).

3.1.1. Plant DNA extraction

- 5-10 g of fresh or frozen young leaf material was grounded with liquid nitrogen to a fine powder.
- 10 ml of warm (65°C) 2x CTAB buffer and proteinase K (final concentration 100 µg/ml) was added to the homogenate, mixed well and incubated under agitation for 30 min at 65°C.
- One volume of chloroform/isoamyl alcohol (24:1) was added to the homogenate in a 50 ml centrifuge tube, mixed well and the solution was centrifuged for 10 min at 4°C and 12000 rpm (Sigma centrifuge 4K 10, rotor n. 12166).
- The aqueous phase was transferred to a new 50 ml tube and 1/10 volume of 10% CTAB solution was added.
- One total volume of chloroform/isoamyl alcohol (24:1) was added and carefully mixed well.
- The mixture was centrifuged for 10 min at 4°C and 12000 rpm and the aqueous phase was again transferred to a new 50 ml tube.
- One volume of warm (65°C) CTAB precipitation buffer was added and carefully, but thoroughly mixed.
- The CTAB-DNA complex was pelleted by centrifugation for 10 min at 20°C and 12000 rpm and resuspended in 2 ml warm (65°C) High salt TE buffer.
- The DNA was precipitated with 2 volume of cold 96% (v/v) alcohol, mixed well but carefully and stored at –20°C overnight.
- The DNA solution was centrifuged for 15 min at 4°C and 12000 rpm (Biofuge, rotor n. 3754, Heraeus, Hanau) and the pellet was washed with 70% (v/v) alcohol,

recentrifuged for 5 min and dried in an excicator before resuspension in 1 ml 1x TE buffer with RNAse (10 µg/ml) 1 h at 37°C.

- The DNA solution was stored at 4-6°C until use.

Solutions used in plant DNA extraction:

2x CTAB buffer	100 mM Tris-HCl (pH 8) 20 mM EDTA 1.4 M NaCl 1% (w/v) PVP 40000 2% (w/v) CTAB
10% CTAB solution	10% (w/v) CTAB 0.7 M NaCl
CTAB precipitation buffer	50 mM Tris-HCl (pH 8) 10 mM EDTA 1% (w/v) CTAB
High salt TE buffer	10 mM Tris-HCl (pH 8) 1 mM EDTA 1 M NaCl
10x TE buffer stock solution	100 mM Tris-HCl (pH 8) 10 mM EDTA

3.1.2. Measurement of DNA concentration

- Bio-Rad VersaFluor^{TM} Fluorometer (Bio-Rad Laboratories, CA, USA) was used.
- The fluorometer was switched on 20 min before use.
- The gain was set to MED (medium) and the rage adjusted to read 00000.
- 2 ml of Hoechst working solution was added to 7 cuvettes (all solutions had room temperature before measurement).

- A standard curve was set up by adding 1000, 500, 200, 100, 50, 20 and 0 ng (blank) standard DNA, respectively to the seven 2 ml cuvettes.
- The instrument was zeroed with the blank cuvette and after 5-10 seconds the range of the instrument was set to 1000 with the highest concentration cuvette (1000).
- The cuvette with 500 ng standard DNA was placed into the instrument and after 5-10 seconds the relative fluorescence unit (RFU) was recorded.
- The former step was repeated until all standard samples were read. After determination of standard curve sample DNAs were measured by adding and mixing 2 µl of the DNAs to a cuvette with 2 ml of Hoechst working solution.
- The cuvette with the sample DNAs were placed into the instrument and after 5-10 seconds the relative fluorescence unit (RFU) was recorded.
- The sample concentration was determined by comparing its relative fluorescence unit values with the standard curve and by dividing the result by the volume of the DNA solution added to the cuvette.

Solutions used for measurement of the DNA concentration:

10x TEN buffer stock solution	100 mM Tris 2 M NaCl 10 mM EDTA pH 7.4
Hoechst 33258 stock solution (store in the dark at 4°C)	1 mg/ml Hoechst 33258 dye in sterile water
Hoechst work solution (store in dark)	0.1 mg/ml Hoechst 33258 dye in 1x TEN
Standard DNA	Calf thymus DNA at 100 µg/ml and 10 µg/ml in 1x TEN buffer

3.2. Phage libraries construction

The Phage libraries were constructed using the “ZAP Express^{R} Predigested Vector Kit" and "ZAP Express^{R} Predigested Gigapack^{R} Cloning Kit”, digested with *Eco*RI from the company Stratagene (Europe).

Genomic DNA from *B. oleracea italica* was digested with the enzyme *Tsp*509I, a four-base cutting enzyme that produces ends compatible to *Eco*RI, fractionated on a 1.5% (w/v) agarose gel and DNA size fractions were eluted from the gel (Section 3.2.1). Three libraries were constructed inserting 250 ng of digested *B. oleracea* DNA into the *Eco*RI site of the lambda phage vector: Library 1 with a DNA fraction of 200-1000 bp; Library 2 with a fraction of 1000-2000 bp; and Library 3 with unfractionated total DNA digested with *Tsp*509I (Section 3.2.2).

3.2.1. DNA digestion and fractionation

- 100 µg of *B. oleracea italica* was digested overnight with *Tsp*509I (3U/µg), in appropriate buffer, at 65°C.
- Digested DNA was precipitated with 0.1 volume of 3 M NaAc, pH 5.2, and 2.5 volume of 96% (v/v) cold ethanol for 1 h at –70°C.
- The precipitated DNA was centrifuged for 30 min at 13000 rpm (Biofuge, rotor n. 3754, Heraeus, Hanau), washed with cold 70% (v/v) ethanol, centrifuged 15 min at 13000 rpm, dried in an exsiccator and redissolved in 20 µl TE buffer for 4 h at 45°C.
- To isolate DNA fractions, the redissolved DNA was loaded on a 1.5% (w/v) agarose gel, run for 1 h, at 90V, in TAE buffer and stained with ethidium bromide (1 µg/ml in water) for 30 min.
- DNA fractions between 200-1000 bp and 1000-2000 bp were excised from the gel and eluted using the E.Z.N.A. gel extraction kit from PeqLab, following supplier’s instructions.

3.2.2. DNA ligation in phage lambda and packaging of DNA

Ligation of insert DNA and packaging of DNA were carried out according to the protocols provided in the Instruction Manual from "ZAP Express^R Predigested Vector Kit" and "ZAP Express^R Predigested Gigapack^R Cloning Kit", Stratagene (Europe).

3.3. Plasmid library construction

The plasmid library was constructed using the pUC19 plasmid vector digested with *Eco*RI (MBI Fermentas, St. Leon-Rot). Total DNA or a DNA fraction ranging from 1000 to 2000 bp from *B. oleracea italica* genomic DNA digested with the enzyme *Tsp*509I (Section 3.2.1) was ligated to pUC19 (Section 3.3.1). With the ligation, competent *E. coli* cell of strain XL1 were transformed by electroporation and cells were plated on appropriate medium (Sections 3.3.2 and 3.3.3). To extend the total DNA library the ligation and transformation were later repeated with unfractionated total DNA. A total of 205 recombinant white colonies from the fractionated library and 959 from the total DNA library (212 from the first ligation and 747 from the second ligation) were transferred with toothpicks to YT-ampicillin (60 µg/ml) medium and grow overnight. A sample of 350 µl from each clone was stored at –20°C after adding glycerol to a final concentration of 15% (v/v).

3.3.1. DNA ligation in plasmid vector

- The following components were added to a 1.5 ml reaction tube:
  250 ng of DNA digested with the enzyme *Tsp*509I (total DNA or 1-2 kb factioned DNA) from *B. oleracea italica*
  0.5 µl of pUC19 (500 ng/ul, MBI Fermentas, St. Leon-Rot)
  1.5 µl 10x ligase buffer (Stratagene, Europe)
  1.5 µl 10 mM rATP (pH 7.5)
  Water to a final volume of 14.5 µl
  0.5 µl T4 DNA ligase (4U/µl) (Stratagene, Europe)
- Components were mixed, spun and incubated overnight at 20°C.
- T4 DNA ligase was inactivated by heating to 65°C for 10 min and the ligation mixture was purified using a QIAprep column (Qiagen, Hilden), according to the

manufacturer´s specifications for purification of plasmid DNA prepared by other methods.

3.3.2. Preparation of competent *E. coli*

- LB agar plate containing tetracycline (12.5 µg/ml) was inoculated with *E. coli* XL1 and incubated overnight at 37°C.
- From a single colony 4 ml LB-tetracycline (12.5 µg/ml) liquid medium was inoculated and incubated for 5-8 h, at 37°C under vigorous shaking.
- 400 ml LB was inoculated with the 4 ml of bacterial culture and incubated at 37°C under vigorous shaking until the OD_{600} was at 0.5-0.9 (from here on buffer, tips, tubes, rotors, etc were kept ice-cold).
- The bacterial culture was than chilled in ice-water for 15-30 min followed by centrifugation at 4°C, at 3000 rpm (Sigma centrifuge 4K 10, rotor n. 12254), for 15 min to collect the bacteria.
- The bacterial pellet was carefully resuspended in 400 ml ice-cold sterile Millipore water and bacteria were again collected by centrifugation for 15 min at 4°C and 3000 rpm.
- The pellet was carefully resuspended in 200 ml ice-cold sterile Millipore water, and the bacteria were centrifuged at 3000 rpm for 15 min at 4°C and the pellet was carefully resuspended in 8 ml ice-cold sterile glycerine.
- Bacteria were centrifuged at 3000 rpm, for 15 min at 4°C and pellet was carefully resuspended in 0.8-1.2 ml ice-cold sterile glycerine.
- Aliquots of 40 µl were flash-frozen in liquid nitrogen and stored at -70°C until use.

3.3.3. Electroporation and library plating

- Competent cells were slowly thawed on ice, mixed with 4 µl of the ligation mixture and left for 1 min on ice.
- The transformation mix was transferred to the bottom of an ice-cold cuvette and an electric pulse was applied using the Gene Pulser^R II Electroporation system from BioRad (BioRad, USA) at the following settings: voltage of 2.5 kV; capacitance of 25 µF; impedance at Low ohms; Voltage Booster resistance at 200 ohms; impulse at ca. 5 msec.

- Immediately after electroporation 450 µl of YEB medium was added to the cuvette, mixed carefully with a Pasteur pipette and transferred to a 10 ml tube.
- The bacterial suspension was incubated under shaking at 200 rpm for 1 h at 37°C.
- 50-100 µl of transformed cells were plated on YT-agar plates containing ampicillin (200 µg/ml), X-Gal (60 µg/ml) and IPTG (60 µg/ml).

Solutions used in the Plasmid library construction:

YEB medium (pH 7.2)	1 g/l yeast extract 5 g/l beef extract 5 g/l Pepton (Casein) 0.5 g/l $MgSO_4$ x 7 H_2O 5 g/l sucrose
LB medium (pH 7.0)	10 g/l NaCl 10 g/l Tryptone 5 g/l yeast extract
LB agar medium	15 g/l bacto-agar in LB medium
YT medium (pH 7.0)	5 g/l NaCl 16 g/l Tryptone 10 g/l yeast extract
YT agar medium	15 g/l bacto-agar in YT medium
Ampicillin stock solution	100 mg/ml in water
Tetracycline stock solution	25 mg/ml in ethanol
IPTG stock solution	0.1 M (23.8 mg/ml in water)
X-Gal stock solution	50 mg/ml in DMF

3.4. Plating lambda libraries, plaque lifts and hybridisation procedure

To plate the packaged ligation product, 1µl of the final packaging reaction was mixed with 200 µl of *E. coli* XL1-Blue MRF´ cells at an OD_{600} of 0.5 and incubated at 37°C for 15 min. To this mixture 10 ml of NZY top agar at 48°C was added, mixed well and plated onto 150 mm NZY agar plates. Plates were incubated overnight at 37°C and chilled for at least 2 h at 4°C before plate lift.

Replica membranes with the plaques were prepared in order to screen the libraries. To lift the plaques, nylon membranes (Hybond N+ 132 mm in diameter, Amersham Pharmacia Biotech, Freiburg) were placed onto each NZY plate for 1-2 min for the first membrane and for 3-4 min for the second membrane. For orientation needle pricks were made through the membranes and agar. Membranes were denatured for 5 min in filter paper wetted with denaturation solution (1.5 M NaCl, 0.5 M NaOH), neutralised 2x for 3 min in filter paper wetted with neutralisation solution (1.5 M NaCl, 0.5 M Tris/HCl pH 7.5), washed in 2x SSC for 1 min, air dried and baked for 2 h at 80°C.

Total genomic DNA from *B. oleracea italica* or *B. rapa* NPZ-35 were used as probe for screening the libraries. Genomic DNAs digested with *Eco*RI were labelled with ^{32}P by random primed labelling, using the Rediprime II DNA labelling kit (Amersham Pharmacia Biotech, Freiburg) following supplier's instructions.

Hybridisation of the membranes followed the method of Bernatzky and Tanksley (1986) with some modifications. Membranes were prehybridised in hybridisation solution for at least 4 h in a glass dish at 65°C. Labelled probe, diluted in 300 µl of hybridisation solution, was added to the glass dish and hybridisation was carried out overnight at 65°C. After hybridisation, membranes were washed for 5 min in 2x SSC/0.1% (w/v) SDS followed by washing for 5 min in 1x SSC/0.05% (w/v) SDS. After washing, membranes were wrapped in plastic wrap and autoradiographed using Kodak X-OMAT AR film with a single intensifying screen at –70°C for 48-72 h.

The replica membranes were compared and plaques that showed hybridisation with only one genome were selected for a second screening.

For the second screening, selected plaques were picked up from the agar plates with a Pasteur pipette and resuspended in 300 µl of SM buffer. One to 2 µl of the SM-suspension was used with 100 µl of *E. coli* XL1-Blue MRF´ cells at an OD_{600} of 0.5 to be plated in the same way as described above, but here 87mm NZY agar plates and 4 ml of NZY top agar were used. Preparation of the membranes, prehybridisation, probe preparation, hybridisation and washes were performed as described above for the library screening.

Solutions used in plating the lambda libraries and hybridisation:

NZY broth medium (pH 7.0)	5 g/l NaCl 2 g/l $MgSO_4$ x 7 H_2O 5 g/l yeast extract 10 g/l NZ amine
NZY agar medium	15 g/l bacto-agar in NZY
NZY top agar medium	7 g/l agarose in NZY
SM buffer	100 mM NaCl 100 mM $MgSO_4$ x 7 H_2O 50 mM Tris-HCl (pH 7.5) 0.01% (w/v) gelatin Filter sterilised (0.2 µm)
20x SSC stock solution	3.0 M NaCl 0.3 M Na-citrate (pH 7.0)

100x Denhardt´s stock solution	2% (w/v) PVP-40 2% (w/v) Ficoll 2% (w/v) BSA (fraction V). Filter sterilised and stored at –20°C
Hybridisation buffer	5x SSC 0.6% (w/v) SDS 50 mM NaH_2PO_4 5x Denhardt´s 1% (w/v) blocking reagent filtered through a Miracloth filter (Calbiochem USA)
8% blocking reagent stock solution (stored at 4°C)	8g blocking reagent (Boehringer, Manheim) 100 ml 4x SSC heated for dilution

3.5. Dot-blot hybridisation procedure

Plasmid DNAs from individual clones of the library were prepared by the alkaline-lyses method modified from Birnboim and Doly (1979) (Section 3.5.1). Denatured plasmid DNA were dot blotted on nylon membranes (Hybond N+, Amersham Pharmacia Biotech, Freiburg), using a Dot-Blot equipment (S&S Minifold I, Schleicher and Schuell, Dassel) (Section 3.5.2). Total genomic DNAs from *B. oleracea italica* or *B. rapa* NPZ-35 were used as probes for screening for repetitive DNA sequences. Probe preparation, prehybridisation, hybridisation and stringent washing were performed as described for the phage library screening (Section 3.4).

3.5.1. Plasmid DNA isolation

- A white colony from the plasmid library was incubated overnight at 37°C in 3 ml of YT/ampicillin (50 µg/ml).
- 1.5 ml of the bacterial culture was transferred into a 2 ml tube and centrifuged at max. speed for 30 seconds (Microliter Hettich centrifuge, Tuttlingen).

- The medium was removed and the bacterial pellet was resuspended in 100 µl of ice-cold solution Sl by vigorous vortexing.
- 200 µl of freshly prepared solution SII was added to the tube and mixed carefully by inversion. Tubes were stored for 2-5 min on ice.
- 150 µl of ice-cold solution SIII was added to the tube and mixed carefully by inversion. Tubes were stored for 25 min on ice before centrifugation at max. speed for 15 min.
- The supernatant was transferred to a fresh tube with two volumes of 96% (v/v) ethanol, mixed and allowed to stand for 30 min for DNA precipitation.
- Tubes were than centrifuged at max. speed for 5 min and the supernatant was removed.
- The pellet was washed with 1 ml of 70% (v/v) ethanol, air dried, redissolved in 50 µl of TE containing 20 µg/ml of DNAse free RNAse, and stored at –20°C until use.

Solutions used in Plasmid DNA isolation:

Solution SI	50 mM Glucose 25 mM Tris/HCl (pH 8) 10 mM EDTA (pH 8)
Solution SII	0.2 N NaOH 1% (w/v) SDS
Solution SIII	3 M potassium acetate, pH 4.8

3.5.2. Dot blotting of plasmid DNA

- Plasmid DNAs were denatured in a water bath at 100°C for 5 min.
- After denaturation the DNA was rapidly transferred to an ice-cold water bath for 3-5 min, briefly centrifuged and kept on ice until use.
- Replica membranes were prepared by applying 5 µl of samples on two Hybond N+ membranes (Amersham pharmacia Biotech, Freiburg) fixed in a Dot-Blot equipment (S&S Minifold I from Schleicher and Schuell, Dassel).

- Suction was applied to the Dot-Blot equipment and after 30 min the membranes were taken out from the equipment and air dried.
- Dried membranes were placed on filter paper wetted in denaturation solution (1.5 M NaCl, 0.5 M NaOH) for 5 min, in neutralisation solution (1.5 M NaCl, 0.5 M Tris/HCl, 1 mM EDTA pH 7.2) for 1 min, and in fixation solution (0.4 M NaOH) for 20 min.
- Membranes were air dried again and wrapped in plastic wrap until use.

3.6. Southern-blot hybridisation procedure

After comparing the dot blot replica membranes, clones that showed different intensities of hybridisation with the A and C genomes were selected for Southern blot analysis. Between 2-10 µg of genomic DNAs digested with *Hind*III or *Eco*RI were separated on a 0.8-1% (w/v) agarose gel and blotted onto a Hybond N+ membrane (Amersham Pharmacia Biotech, Freiburg). Probe DNAs were prepared by PCR using M13 forward and reverse primers (Section 3.6.1) and labelled with ^{32}P with Rediprime II DNA labelling system, according to supplier's instructions (Amersham Pharmacia Biotech, Freiburg). After blotting the genomic DNA (Section 3.6.2), membranes were prehybridised, hybridised overnight and washed under stringent conditions, all at 65°C, before being autoradiographed using Kodak X-OMAT AR film (Section 3.6.3). Filters were stripped in order to be reused (Section 3.6.4).

3.6.1. Probe preparation

- Insert DNA was amplified from the respective plasmid by PCR as following:

  50 ng plasmid DNA
  10 µl 10x PCR buffer
  10 µl dNTP´s (2 mM each)
  3 µl forward M13 primer (10 µM)
  3 µl reverse M13 primer (10 µM)
  Water to complete 99.2 µl
  0.8 µl Taq DNA polymerase (5U/µl)

- The PCR- profile for amplification was 1x 5 min 94°C, 3 min 75°C, then 40x 30 sec 94°C, 1.5 min 55°C, 3 min 72°C.

- After PCR, the solution with the amplified DNA was mixed with one volume of chloroform-isoamylalcohol (24:1), and the supernatant was recovered after centrifugation for 15 min.
- The DNA was precipitated with 0.1 volume of 3 M NaAc, pH 5.2, and 2.5 volume of ice-cold 96% (v/v) ethanol for 2-4 h at –20°C.
- The precipitated DNA was centrifuged 30 min at 4°C and 13000 rpm, washed with cold 70% (v/v) ethanol, centrifuged 15 min at 13000 rpm, dried in excicator and redissolved in TE buffer.
- After at least 2 days in a refrigerator the DNA concentration was measured with a fluorometer as described (Section 3.1.2).
- The Probe DNAs were labelled with ^{32}P with the Rediprime II DNA labelling kit (Amersham Pharmacia Biotech, Freiburg), in accordance to supplier's instruction.

3.6.2. Southern blot

- Between 2-10 µg of genomic DNAs from *Brassica* species (Table 1) were digested with *Hind*III or *Eco*RI (3U/µg) in appropriate buffer overnight at 37°C.
- Digested DNAs and 100 ng of lambda *Eco*RI/*Hind*III marker were loaded with loading buffer on a 0.8-1% (w/v) agarose gel with ethidium bromide (0.125 µg/ml).
- Samples were run overnight at 40V in TAE buffer. Afterwards gels were photographed.
- Partial depurination, denaturation and neutralization of the gels was achieved through incubating the gel in 0.25 M HCl for 10 min, 1.5 M NaCl/0.5 M NaOH for 30 min and 1.5 M NaCl/0.5 M Tris/HCl pH 7.5 for 30 min, respectively. Between each treatment the gels were washed with water.
- Southern transfer was made by the classical capillary blotting technique (protocol provided by Amersham pharmacia biotech) for small gels, or by vacuum blotting (VacuGene^{TM}XL Vacuum blotting System from Parmacia) for large gels, using Hybond XL or N+ membranes (Amersham Pharmacia Biotech, Freiburg).
- After DNA transfer membranes were baked for 2 h at 80°C, cooled down and stored in refrigerator until use.

3.6.3. DNA-DNA hybridisation

- Before hybridisation the membranes were prehybridised for at least 4 h at 65°C in 45 ml of prehybridisation buffer.
- Prehybridisation buffer was replaced by 15 ml of hybridisation buffer and the probe, diluted in 300 µl of this buffer, was added.
- Membranes were hybridised overnight at 65°C in a hybridisation oven.
- After hybridisation membranes were washed with 200 ml of 2x SSC/0.1% (w/v) SDS and 1x SSC/0.05% (w/v) SDS, for 20 min each, at 65°C in the hybridisation tubes.
- The last wash step took place in a water bath with unrolled membranes in 500 ml of 0.5x SSC/0.025% (w/v) SDS, for 10-20 min at 65°C.
- After washing, the membranes were wrapped in plastic wrap, and autoradiographed using Kodak X-OMAT AR film with a single intensifying screen at –70°C for 24-72 h.

3.6.4. Stripping membranes

For removal of probes, moist membranes were incubated in 500 ml of 0.2 M NaOH for 30 min at 45°C, followed by an incubation in 500 ml of 0.2 M Tris/HCl pH 7.5/ 0.1x SSC/ 0.1% (w/v) SDS for 30 min at 45°C.

Solutions used in Southern blotting hybridisation:

6x Loading buffer stock solution	40% (w/v) sucrose 0.25% (w/v) xylene cyanol 0.25% (w/v) bromophenol blue
50x TAE buffer stock solution	2 M Tris 17.4 N acetic acid 50 mM EDTA
20x SSC	see solutions in 3.3.1

100x Denhardt´s	see solutions in 3.3.1
Prehybridisation buffer	5x SSC, 0.6% (w/v) SDS 50 mM NaH_2PO_4 5x Denhardt´s 200 µg/ml salmon sperm DNA
Hybridisation buffer	5x SSC 0.6% (w/v) SDS 50 mM NaH_2PO_4 5x Denhardt´s 100 µg/ml salmon sperm DNA 5% (w/v) dextran sulfate

3.7. DNA sequencing and computer analysis of sequence data

For sequencing all plasmid DNA were prepared using the QIAprep kit (Qiagen, Hilden), according to the supplier´s instructions.

The sequencing was carried out by the company MWG-Biotech, Comfort-Read service, using automated DNA sequencing. The DNA sequence was analysed using the BioEdit Sequence Alignment Editor program, version 4.7.8. All sequences were analysed for homology to known sequences using the BLAST service at NCBI for all organisms (http://www.ncbi.nlm.nih.gov) and at the TIGR service for the *B. oleracea* genome database (http://www.tigr.org/tdb/e2k1/bog1/index.shtml).

3.8. Fluorescent *in situ* hybridisation (FISH)

Fluorescent *in situ* hybridisation (FISH) was performed on *Brassica* chromosomes spreads obtained from root tips (Section 3.8.1), using as probes rDNA as control and insert DNAs from the clones selected in the plasmid library (Section 3.8.2). After pre-

treatment of the cytological material (Section 3.8.3), chromosomes were denatured and hybridised overnight to labelled probes (Section 3.8.4). Stringent washes were performed (Section 3.8.5) before probe detection (Section 3.8.6). Before and after FISH, slides were examined and chromosome spreads with good quality were recorded for further analyses (section 3.8.7).

3.8.1. Accumulation, fixation and squash preparation of plant chromosomes

- Seeds from *Brassica oleracea oleracea, B. rapa* Rex, *B. napus* Jockey and *B. carinata* BRA 489/77 were germinated on moist filter paper in a Petri dish, for 2-3 days at 24°C until the primary roots were 1-2 cm long. Cell division was synchronised by incubating the seedlings overnight at 4°C, followed by a further 24 h at 24°C.
- The roots were pre-treated with 2 mM 8-hydroxyquinoline for 2 h at RT under agitation, briefly blotted dry on paper tissue, fixed in Farmer fixative (3:1 ethanol : acetic acid) for 4-16 h at RT and stored in the fixative at –20°C until use.
- Before enzyme digestion roots were washed 2x for 5 min in water and for 15 min in enzyme buffer.
- Digestion with 2% (w/v) cellulase (Onozuka r-10; Merck, Darmstadt) and 20% (v/v) pectinase (Sigma-Aldrich, Munich) in enzyme buffer was carried out at 37°C for 45-60 min.
- 1-3 root tips were squashed per slide in a drop of 45% (v/v) acetic acid and coverslip was removed in liquid nitrogen. Slides were air dried and stored in a refrigerator until use.

3.8.2. Probe preparation and labelling

Two probes were used in FISH experiments as positive control: (1) one to detect the 45S rDNA site and another (2) to detect the 5S and 45S rDNA sites. The probe (1) was named here as 45S rDNA probe, and was generated by PCR using total genomic DNA from *Brassica* as template and primers from the 25S rDNA region (JF 09- 5´GCG AGC GAA CCG GGA TAA GCC C 3` and JF 10- 5´CGG AAT TTA CCG CCC GAT TGG GG 3`, Yokota et al., 1989), what generates an amplicon of about 220 bp long. The PCR- profile for amplification of this probe was 1x 5 min 94°C, then 34x 50 sec 94°C, 30 sec 55°C, 1.5 min 72°C and finally 1x 10 min 72°C. The probe

(2), was named here as "5+45S rDNA". This probe was a plasmid containing a 1735 bp *Eco*RI fragment of the 18S rDNA region from the 45S rDNA of *Aedes aegypti* and 509 bp of 5S rDNA from *Matthiola incana*. Plasmid DNA with "5+45S rDNA" and from selected clones were isolated using plasmid DNA extraction kit from Qiagen (Hilden), following supplier's instructions. Inserts from the clones selected in the plasmid libraries were amplified by PCR (Section 3.6.1) and used as probes after labelling.
Probes were labelled by one of the following methods: random priming, PCR or nick translation.

Random primed labelling

Inserts amplified by PCR (section 3.6.1) or plasmid DNA were labelled with digoxigenin-11-dUTP using DIG-High Prime DNA labelling kit (Roche, Mannheim), following supplier's instructions.

PCR labelling

- Inserted DNAs were amplified and labelled by PCR using Taq PCR core kit (Qiagen, Hilden) as follow:

  5.7 µl water + 20 ng selected plasmid DNA
  5 µl 5x Q sol
  2.5 µl 10x PCR buffer
  2.5 µl dNTP´s without dTTP (2 mM each)
  1.6 µl dTTP (2 mM)
  1.7 µl DIG-11- dUTP (1 mM) (Boehringer Manheim)
  2.5 µl forward M13 primer (10 µM)
  2.5 µl reverse M13 primer (10 µM)
  1 µl Taq DNA polymerase (1 U/µl)
- The PCR- profile for amplification was 1x 5 min 94°C, 3 min 75°C, then 40x 30 sec 94°C, 1.5 min 55°C, 3 min 72°C.
- For PCR labelling of 45S rDNA, primers, template and PCR profile used were the same as described in section 3.8.2.

Nick translation

Biotin-16-dUTP labelled probes were generated with the Biotin-Nick Translation Mix for *in situ* probes (Roche, Mannheim), following supplier's instructions, using either plasmid DNA or insert DNA amplified by PCR (section 3.6.1).

3.8.3. Pre-treatment of chromosome preparations for FISH

- Before FISH, slides were incubated for 30 min at 60°C, to fix cells to the glass slides, followed by incubation for 5-10 min at RT.
- Slides were treated with DNAse-free RNAse in 2x SSC (400 µg/ml) for 1 h at 37°C (100 µl per slide), washed 2x for 5 min with 2x SSC at RT and 1x for 5 min with 10 mN HCl at RT.
- Treatment with freshly diluted pepsin (1:100 dilution of 1 mg/ml pepsin in 0.01 N HCl) was carried out for 20 min at 37°C (100 µl per slide).
- Slides were than washed 2x for 5 min with 2x SSC, 1x for 5 min with 1x PBS and fixed in 1% (v/v) aqueous formaldehyde (Sigma-Aldrich, Munich) in 1x PBS at RT 10 min.
- Slides were washed 2x for 5 min in 1x PBS, dehydrated in an ethanol dilution series (2 min each in 70%, 90%, and 96% ethanol (v/v)) and air dried for 1 h.

3.8.4. *In situ* hybridisation

- Slides and probes were denatured separately before hybridisation.
- Slides were denatured in 200 µl of 70% (v/v) formamide in 2x SSC on a metal plate sitting with maximal heat contact in a water bath, at 70-72°C for 3.5-4 min. After denaturation slides were immediately dehydrated in an ice-could ethanol series of 70%, 90% and 96% (v/v) for 2 min each. Finally, the slides were air dried for at least 30 min at RT.
- Hybridisation mix (15 µl per slide) containing 50-150 ng of the probe, was denatured at 75°C for 10 min and rapidly cooled down 5 min on ice before being added to slides. The slides were hybridised at 37°C overnight.

3.8.5. Removing unbound and weakly hybridised probe

After hybridisation slides were washed stringently 2x for 5 min in 2x SSC, 2x for 5 min in 0.1x SSC and 2x for 5 min in 2x SSC, all at 42°C, followed by one wash for 5 min in 2x SSC at RT and one wash for 7 min in 4x SSC/ 0.1% (v/v) Tween 20 (probes labelled with biotin) or 1x PBS (probes labelled with digoxygenin) at RT.

3.8.6. Detection of hybridisation signals

Detection of biotin labelled probes

(Fluorescein-streptavidin, Biotinylated-anti-Streptavidin and Vectashield were from Vector Laboratories, Peterborough, England).

- To block unspecific antibody binding, 100 µl of 3% (w/v) BSA in 4x SSC/ 0.1% (v/v) Tween 20 was applied on the slide, covered with a plastic coverslip and incubated for 30 min at 37°C.
- The plastic coverslip was carefully removed with forceps. To detect the hybridised probe 100 µl of Fluorescein-streptavidin (10µg/ml in 1% (w/v) BSA) was added to the slide and incubated at 37°C for 20-60 min.
- Slides were then washed 3x for 2-10 min in 4x SSC/ 0.1% (v/v) Tween 20.
- Signals were amplified through incubation with 50 µl Biotinylated-anti-Streptavidin antibody (10µg/ml in 1% (w/v) BSA) for 20-60 min at 37°C, washed again 3x for 2-10 min in 4x SSC/ 0.1% (v/v) Tween 20 and amplified signals were detected again using 100 µl of Fluorescein-streptavidin (10µg/ml in 1% (w/v) BSA) for 20 min to 1 h at 37°C.
- After washing 3x 2-10 min in 4x SSC/ 0.1% (v/v) Tween 20, chromosomes were stained with 7-8 µl of DAPI (1µg/ml in Vectashield antifade) per slide.

Detection of digoxygenin labelled probes

(Fluorescent Antibody Enhance set for DIG detection (Boehringer, Mannheim) was used).

- Antibody stock solutions were freshly diluted before use to 1:25 in 1% (v/v) blocking solution (Boehringer, Mannheim) (final concentration 1ng/µl).
- Slides were immersed in 50 ml 1% (v/v) blocking solution in a coplin jar for 30 min at RT in order to block unspecific antibody binding.
- The probe was detected by incubating the slides at 37°C for 1 h in 50 µl mouse-anti-DIG antibody dilution (1ng/µl).
- Unspecifically bound antibodies were removed by washing the slides 3x for 2-5 min in 2x SSC/ 0.2% (v/v) Tween.
- Amplification of signals was made through incubation of slides at 37°C for 1 h in 50 µl anti-mouse-DIG antibody dilution (1ng/µl).

- Unspecifically bound antibodies were removed by washing the slides 3x for 2-5 min in 2x SSC/ 0.2% (v/v) Tween.
- Amplified signals were detected with 50 µl anti-DIG-fluorescein (1ng/µl) for 1 h at 37° C.
- Unspecifically bound antibodies were removed by washing the slides 3x for 2-5 min in 2x SSC/ 0.2% (v/v) Tween. Chromosomes were counterstained with 7-8 µl of DAPI (1µg/ml in Vectashield antifade) per slide.

Solutions used in FISH:

10x enzyme buffer stock solution	4 mM citric acid 6 mM tri-sodium-citrate pH 4.8
20x SSC	see solutions in 3.3.1
10x PBS stock solution	1.4 M NaCl 85 mM Na_2HPO_4 15 mM KH_2PO_4 30 mM KCl pH 7.4
Hybridisation mix	50% (v/v) formamide 10% (w/v) dextran sulfate 2x SSC
BSA	3% (w/v) BSA in 4x SSC/ 0.1% (v/v) Tween 20. Filter sterilised and stored at -20°C
DAPI stock solution	100µg/ ml in water

3.8.7. Microscopy and photography

The slides were examined with a Zeiss Axiolab Epifluorescent or a fluorescent Olympus BX-60 (Olympus Optical co. LTD, Japan) microscopes equipped with filter sets for DAPI (4,6-diamidino-2-phenylindole), and FITC (fluorescein isothiocyanate). Photographs were taken on Kodak Ektochrome P1600 professional reversal film for colour slides with a MC 80 camera attached to the Zeiss Axiolab microscope or with a CCD (Charged-Couple Device) camera (camera colour view 12) attached to the Olympus microscope using the analyses software Analysis 3.0 (Soft Imaging System GmbH, Münster). Images were processed uniformly before scoring the results by using image processing software (Adobe Photoshop version 5.5; Adobe System, San Jose, Calif.). For the karyotype construction colour images were then converted to grayscale mode (8 bits) and inverted. Black and white pictures were again uniformly processed before printing.

For the karyotype descriptions the total length and the ratio of the short and the long arms of the chromosomes from 3 metaphases stained with DAPI were measured. Karyotypes were constructed using the Adobe Photoshop software. The morphology of the chromosomes was classified according to the system suggested by Levan *et al*. (1964), using the terms metacentric and submetacentric. The number and distribution of signals of each probe were averaged in at least 3 metaphases for each species, when nothing else is mentioned.

4. Results

4.1. Searching for *Brassica oleracea* species-specific repetitive sequences

4.1.1. Phage library screening

The first library (library 1) constructed for searching for species-specific repetitive sequences in *B*. *oleracea* was a lambda phage library. A DNA fraction ranging from 200 to 1000 bp from *B*. *oleracea* total DNA digested with *Tsp*509I was used for this purpose. In order to find clones with repetitive DNA, replica filters with plaques from this library were made and screened using either *B*. *oleracea* or *B*. *rapa* total genomic DNA labelled with ^{32}P as probe. Figure 2 presents one autoradiogram from replica filters after the colony hybridisation. Strong spots represent clones with repetitive DNA ("positive clones") and are seen in few plaques, whereas most of the plaques show no or a very faint hybridisation signal ("negative clones"). Clones with repetitive DNA which hybridised strongly with only one of the genomic probe were selected for a second screening as putative species-specific. Two of such clones are represented in the detail of Figure 2.

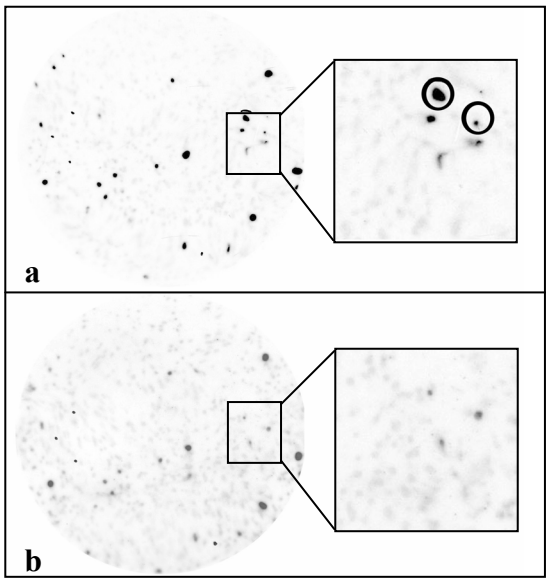


Figure 2. Autoradiogram from replica filters in the first screening of the *B*. *oleracea* lambda library 1 after hybridisation with total genomic DNA of (a) *B*. *oleracea* and (b) *B*. *rapa*. The insert on the right side shows a magnification of the boxed area in the

respective filters, representing the same region in both filters. Clones with repetitive DNA show strong hybridisation signals. Circled clones show differential hybridisation between the replica membranes and were collected to a second screening as candidates for species-specificity.

About 80,000 plaques of the lambda library 1 were screened, but only 350 clones showed strong hybridisation signals with total genomic DNA from *Brassica*, indicating the presence of repetitive DNA. This means that only 0.45% of the clones contained repetitive sequences. From the 350 positive clones in the first screening, ca. 60% hybridised with total DNA from both genomes (212 clones) and, accordingly, were repetitive but not species-specific, whereas about 40% (138 clones) hybridised either with *B*. *oleracea* (68 clones) or with *B*. *rapa* (70 clones), being putative species-specific clones.

To confirm the species-specificity of the 138 clones a second screening was performed. None of these clones was confirmed as containing species-specific repetitive DNA, they hybridised either with DNA of both genomic probes (90 clones) or showed no hybridisation signal (48 clones). Figure 3 shows an example of a clone where all plaques hybridised with DNA of both genomes. Sometimes, however, more than one plaque had been picked up from the Petri dish in the first screening and a mixture with "positive" and "negative" clones could be seen in the second screening.

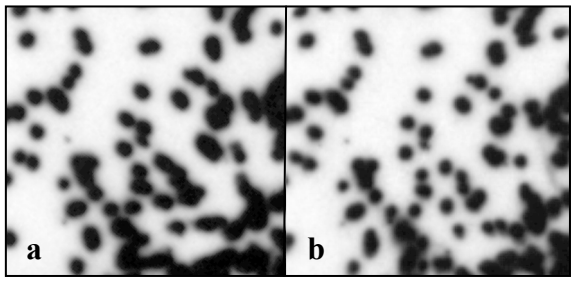


Figure 3. Autoradiogram of replica filters from a putative species-specific clone (lambda clone Bo-93) selected in the first screening, after hybridisation with *B*. *oleracea* (a) and *B*. *rapa* (b) total genomic DNA.

In order to see if the DNA fraction used as insert was responsible for the low amount of clones with repetitive DNA sequences observed in the first phage library (0.45%), two additional phage libraries were constructed. *B*. *oleracea* unfractionated total genomic DNA and a fraction between 1000-2000 bp, both digested with *Tsp*509I, were used to construct the second and the third library, respectively. About 12,000 plaques were screened for each library using only *B*. *oleracea* total DNA as probe. The amount of clones with repetitive DNA found was 2.4% for library 2 and 1.5% for library 3. These frequencies were somewhat higher than in the first library, but still much lower than would be expected taking in account that plant genomes have high amount of repetitive DNA. From these data it was concluded that the low frequency of repetitive DNA in the phage libraries was probably more related to the type of vector used in the library construction than to the DNA fraction used as insert. For this reason new libraries using plasmid DNA as vector were constructed.

4.1.2. Plasmid library screening

To construct *B*. *oleracea* plasmid libraries a DNA fraction ranging from 1000 to 2000 bp and unfractionated total genomic DNA digested with *Tsp*509I were cloned in the *Eco*RI site of pUC19. For simplification, the results of these two libraries will be presented together. Plasmid DNA from a total of 1164 white colonies, denominated as pBo2.1-205 (for clones from the fractionated library) and pBo1.1-959 (for clones from the total DNA library), were extracted, dot blotted on two replica filters and screened for repetitive DNA using either *B*. *oleracea* or *B*. *rapa* total DNA labelled with ^{32}P as probe. About 21% of the clones (246 clones) hybridised strongly with *B*. *oleracea* total DNA and were identified as containing repetitive DNA (Figure 4).

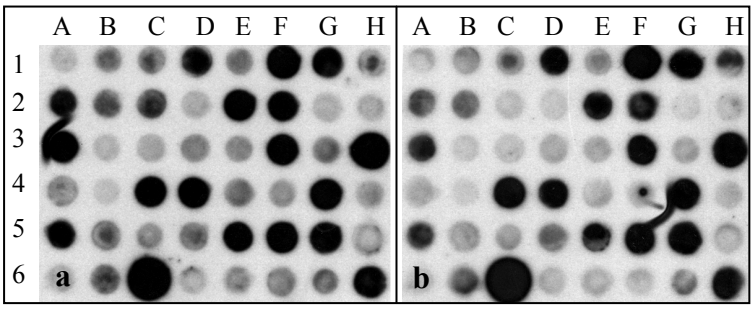


Figure 4. Dot-blot hybridisation of pBo clones with *B*. *oleracea* (a) and *B*. *rapa* (b) total genomic DNA. Clones with repetitive DNA (e.g. D1, E2, F2, C6) show strong hybridisation signals.

After comparison of the dot-blot hybridisation patterns with both probes, none of the 1164 clones was clearly identified as species-specific, but some exhibited differences in the intensity of hybridisation with both genomic probes. For exhibiting stronger hybridisation with C genome DNA probe, 15 clones were selected as putative C genome-specific or -enriched clones for further Southern blot analyses. Dot blots from the clones confirmed as enriched in C genome are shown in Figure 5.

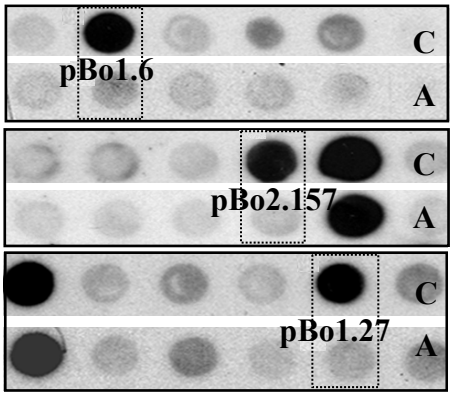


Figure 5. Dot-blot hybridisation of pBo clones confirmed as enriched in C genome after hybridisation with *B. oleracea* (C) and *B. rapa* (A) total genomic DNA.

4.2. Genomic organisation of the putative C genome-specific or enriched clones

The genomic organisation of the 15 selected clones was analysed by Southern blot hybridisation to genomic DNA of the A and C genomes of different *Brassica* species digested with *Eco*RI or *Hind*III.

Only three out of the 15 selected clones, pBo1.6, pBo1.27 and pBo2.157, were confirmed as enriched in C genome, showing differences in the number and intensity of bands hybridising with A and C genomic DNAs, independent of the enzyme used in DNA digestion. Ten out of the 15 selected clones, including the three confirmed as enriched in C genome, were hybridised to *B. nigra* DNA (B genome) and were also detected in this genome (data not shown). Figure 6 shows Southern blots hybridised with the three repetitive clones enriched in the C genome. When comparing the lanes of the blots, in Figure 6, where C genome DNA (lanes 1, 2, 7, 8), or A and C genome DNAs (lanes 3, 4) were loaded with that where only A genome DNA were loaded (lanes 5, 6), it is easily ascertained that clones pBo1.6, pBo1.27 and pBo2.157 gave

much stronger signals with C genome DNA. The repetitive sequence cloned in pBo1.6 was clearly less abundant in the A genome than the sequences of the two others clones, showing only a prominent discrete band of about 2 kb length in *Eco*RI and 10 kb in *Hind*III digested DNA of the A and also of the B genome (data not shown) and a very faint smear. Strong, smeared signals with some prominent bands were observed after hybridisation with pBo1.6 in all species with C genome. Clones pBo1.27 and pBo2.157 showed a smear of hybridisation signals with both genomes, in both enzyme treatments, although few prominent bands were also present. The three clones hybridised preferentially with fragments of the high molecular weight DNA fraction from genomic DNAs digested with *Hind*III, whereas fragments of the lower weight fraction were also seen when genomic DNAs digested with *Eco*RI were used.

From the remaining 12 clones, 10 exhibit similarly strong hybridisation signals on Southern blots with DNA of all genomes and two were not repetitive. The non-repetitive nature of these clones was concluded from the observation that Southern blots hybridised with these clones needed a much longer exposition to X-ray films (over seven days) than blots hybridised with the repetitive clones (two-three days). Figure 7 shows hybridisation patterns of some repetitive but not species-specific clones.

Hybridisation with the different probes produced two kinds of signals: (1) strong smeared signal, extending over several centimetre of the lane, indicating the occurrence of related sequences at many different genomic loci, and/or (2) two to six prominent discrete bands (see Figures 6 and 7). Such discrete bands were mostly present in genomes A, B and C as fragments of similar size, indicating their conservation in these different genomes.

It was not possible to determine if the sequences analysed were tandemly repeated or not, since a "ladder", typical from tandem repeated sequences, was absent on the blots with genomic DNA digested with *Eco*RI and *Hind*III.

Although none of the clones have been strictly species-specific, the three clones enriched in the C genome were sequenced and used in fluorescent *in situ* hybridisations, since they could give important information about the structure of the A and C genomes in *B. napus*.

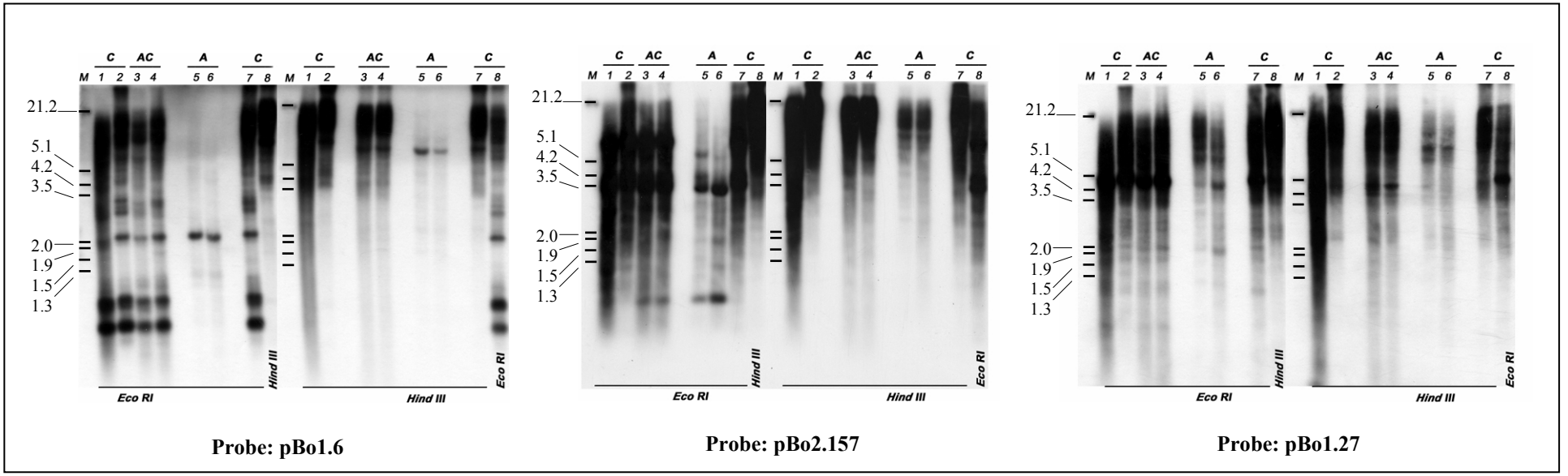


Figure 6. Southern blot hybridisation of genomic DNA digested with *Eco*RI or *Hind*III from different *Brassica* with C genome-enriched clones. *Brassica* species: *B*. *oleracea* Vitamina (lane 1) and Market Victor (lane 2); *B*. *napus* Express (lane 3) and Iris (lane 4); *B*. *rapa* NPZ-35 (lane 5) and Perko (lane 6); *B*. *incana* (lane 7) and *B*. *oleracea alboglabra* (lane 8). Molecular weight marker (Lambda *Eco*RI/*Hind*III, lane M) is given in kb.

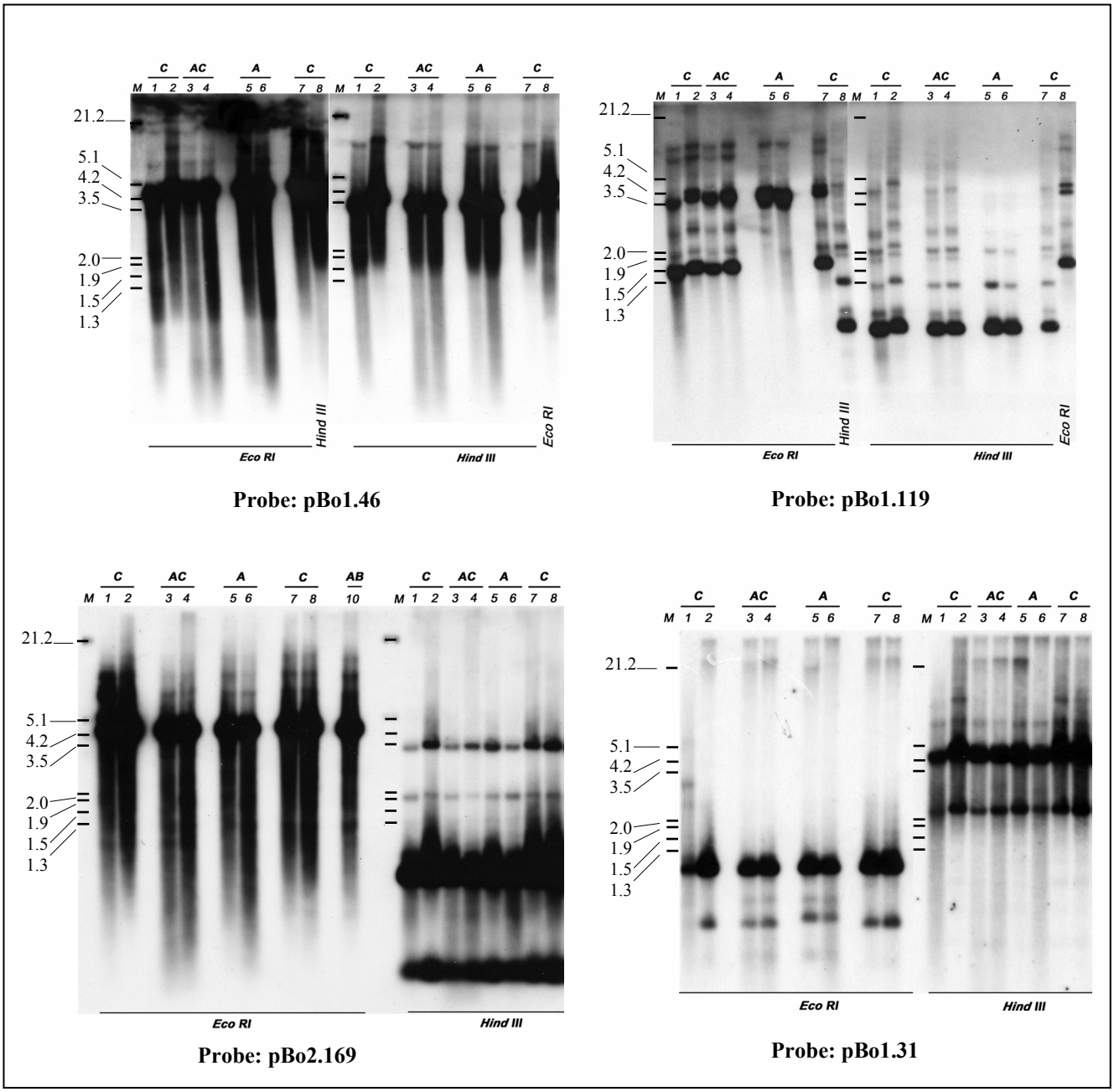


Figure 7. Southern blot hybridisation of genomic DNA digested with *Eco*RI or *Hind*III from different *Brassica* species with repetitive non species-specific clones. *Brassica* species: *B*. *oleracea* Vitamina (lane 1) and Market Victor (lane 2); *B*. *napus* Express (lane 3) and Iris (lane 4); *B*. *rapa* NPZ-35 (lane 5) and Perko (lane 6); *B*. *incana* (lane 7) and *B*. *oleracea alboglabra* (lane 8). Molecular weight marker (Lambda *Eco*RI/*Hind*III, lane M) is given in kb.

4.3. Sequence composition of repetitive DNA from *Brassica oleracea* and homology to published DNA sequences

The size and the base composition of the inserts from the C genome-enriched sequences, pBo1.6, pBo1.27 and pBo2.157, and two highly repetitive clones, pBo2.94 and pBo1.173, selected later as additional controls for *in situ* hybridisation, are presented in Table 2. The size of the cloned DNA fragments varied between 158 bp and 238 bp and the GC content varied from about 34.8% to 48.3%. Figure 8 presents the nucleotide sequences of these clones. In the inserts of the C genome-enriched clones no restriction site was found for the enzymes *Eco*RI and *Hind*III, used for the genomic organisation analyses of these clones. However, clone pBo2.157 contains five nucleotide sequences from which *Hind*III restriction site could arise by single base change (underlined in Figure 8).

Table 2. Length and GC content of the sequenced pBo clones.

pBo Clone	Length (bp)	G+C content (%)
1.6	203	48.28
1.27	182	43.41
2.157	206	37.38
2.94	238	37.82
1.173	158	34.81

All five clones showed similarity to previously described sequences of the EMBL/GenBank/DDBJ sequence databases. The data from the sequence alignments are summarised in Table 3, including up to 5 homologous sequences from different organisms. In a segment varying between 116 bp and 132 bp, the sequence pBo1.6 displayed between 85-89% identity with telomere-like DNA from many plant species. This sequence possessed 19 degenerated telomere motifs (15x TTT(G/C)GGG, 2x TATGGGG, 1x GTTGGGG, 1x TCTAGGG), and one perfect copy of the telomeric repeat of most plants (TTTAGGG), as shown in Figure 8 (see blocks).

pBo1.6 203bp; 27 A; 22 C; 78 T; 76 G.

AATTTCCTCG TAAAAAAAAC ACGGACCTTT GTATGGGGTT TGGGGGTTGG GGTTTCGGGT TTCGGGTTTC GGGTTTGGGG TTTCGGGTTT GGGGTTTCGG GTTTCGGGTT TGGGGTTTCG GGTTTAGGGT ATGGGGTTTG GGGTTTGGGG TTTCGGGTTT CGGGTTTCTG ATTCTAGGGA TTTAAACATA ACACTCGTT**A ATT**

pBo1.27 182bp; 56 A; 31 C; 47 T; 48 G.

AATTGGTTTA TAGTCTTGGG TTGCCTTCGG AGATGATTGA TGTTTGCATC GACAACTGCA TGATCTACTG GAAAGAAGAT GACAAGTTAG AAGAGTGTCG ATTCTGCAAA AAATCACGAT TCAAACCGCA AGGCCGTGGG AGGAATAGGG TACCGTACCA AAGGATGTGG TACCTACC**AA TT**

pBo2.157 206bp; 68 A; 24 C; 61 T; 53 G.

AATTAGAACG GATGTAGATT ATGGTGTAGG TACTGAGCAG ATGGTAAATG ATCATTTTAG AGGGGAAGAT TTACCCAATG CAGAAGCTAG GAGATTTTAT GATATGTTGG ATGCTGGAAA GCAACCATTG TACGAAGGTT GCAGAGATGG TCATTCAGCT TTATCATCTG CTACAAGATT GATGGGCATT AAAACAGATT AT**AATT**

pBo2.94 238bp; 67 A; 50 C; 81 T; 40 G.

AATTACCTGA TTTGAAAGTG GGATAACTTC TTCATGCCAA CTCCTATGAG ATTTATTCAA CTTCCTGGTG ATTCTCCACC ACTTTATGTA TCCAAATCAA CCTTTTCACA AAGTGATTCA TCCTGGTTTG ATTGGAACGA TGAAGAAGCT GTGCTATTCC CGAACTGGGA AACTGGAATC ACCTGATTTG AAAGTGGGAT AACTTCTTCT TGCCAACTCC TATGATATTT ATTC**AATT**

pBo1.173 158bp; 53 A; 28 C; 50 T; 27 G.

AATTTATTAA ATCTCATAGG AGTTGGGATG AAGAAGTTAT CCCACTTTCA AATCATGTGA TTCCAGTTTC CCAGTTTGGG AATAAGACAG CTTCTTCGTC GTTCCAATCA AACCAGGATG AATCACTTTA TAAGAAGCTT GATTTGGTTA CATA**AATT**

Figure 8. Nucleotide sequences, size and base composition of DNA fragments from *B. oleracea* cloned in pBo1.6, pBo1.27, pBo2.157, pBo2.94 and pBo1.173. The complete sequences are shown in groups of ten nucleotides from 5´-3´ with the *Tsp*509I sites bordering the fragments shown in bold. The variations of the telomeric repeat TTTAGGG found in pBo1.6 is shown as blocks and the nucleotide sequences from which *Hind*III restriction site could arise by single base change found in pBo2.157 are underlined.

The whole sequence of pBo2.157 showed high identity (98% of the whole sequence) with two sequences from *Brassica napus* in the databases. With one of them (accession number AJ245479) the similarity was in the region containing an Enhancer/Supressor-mutator (*En/Spm*)-type transposon like ORF and with the other (accession number X98373) the similarity was outside the coding region for the endo-polygalacturonase gene present in this sequence. When this last sequence was subjected as query for a search in the NCBI protein database high similarity also with *En/Spm*-type transposons could be found in the region including the sequence homologous to pBo2.157. The sequence pBo1.27 exhibited also high identity (87% of the whole sequence) with an *En*/*Spm*-type transposon like ORF from a clone of *Brassica rapa*.

Concerning the highly repetitive clones similarities varying from 89 to 98% were found with centromeric repeated DNA from *Brassica* species in the DNA databanks. In a pairwise sequence alignment, sequence pBo2.94 showed 90% homology with pBo1.173.

Moreover, an additional search in the *Brassica oleracea* sequence database from TIGR showed sequence identities of 99.5-100% with several sequences from this database, extending over the full length of the inserts, for all of the C genome-enriched sequences. The degree of identity of sequences of this database with the sequence of the two high repeated clones was slightly lower than that for the C genome-enriched sequences: 98.3% homology for pBo2.94 and 96.2% for pBo1.173.

Table 3. Comparison of the newly isolated repetitive DNA sequences from *Brassica* with sequences of the EMBL/GenBank/DDBJ databases.

Clone (size in bp)	Homologous DNA sequence	Database accession number	sequence identity in bp [a] (%)	E Value [b]
pBo1.6 (203 bp)	*Chlorella vulgaris* telomeric DNA	D26372	114/132 (86)	$2e^{-24}$
	Silene latifolia interstitial telomere-like repeat	AB085635	107/120 (89)	$2e^{-24}$
	Oryza sativa telomere sequence	AY367134	113/132 (85)	$5e^{-22}$
	Arabidopsis thaliana telomeric DNA	AC011621	113/132 (85)	$5e^{-22}$
	Nicotiana tabacum telomere sequence	D21066	113/132 (85)	$5e^{-22}$
pBo1.27 (182 bp)	*Brassica rapa* DNA for an *En/Spm*-type transposon like ORF	AB022082	160/182 (87)	$9e^{-45}$
pBo2.157 (206 bp)	*Brassica napus* DNA for an *En/Spm*-type transposon like ORF	AJ245479	202/206 (98)	$1e^{-102}$
		X98373	202/206 (98)	$1e^{-102}$

continued

Table 3- continued

Clone (size in bp)	Homologous DNA sequence	Database accession number	Sequence identity in bp [a] (%)	E Value [b]
pBo2.94 (238 bp)	Canrep centromeric repetitive DNA from *Brassica juncea*	X68786	224/230 (97)	$1e^{-109}$
	Canrep centromeric repetitive DNA from *B. napus*	X61583	124/129 (96)	$1e^{-53}$
	B. rapa centromeric repetitive DNA	Z22947	194/217 (89)	$4e^{-63}$
	B. oleracea satellite DNA inverted direct repeat	M31436	131/136 (96)	$5e^{-53}$
pBo1.173 (158 pb)	Canrep centromeric repetitive DNA from *Brassica juncea*	X68786	146/150 (97)	$1e^{-68}$
	Canrep centromeric repetitive DNA from *B. napus*	X61583	123/127 (96)	$6e^{-55}$
	B. campestris centromeric repetitive DNA	Z22947	129/133 (96)	$1e^{-58}$
	B. oleracea satellite DNA inverted direct repeat	M31434	120/122 (98)	$9e^{-57}$

[a] sequence identity= number (percentage) of base pairs that are identical in the query and the subject from the DNA databanks

[b] E Value- Describes the number of sequences expect to be found in a database just by chance and is the statistical significance threshold for reporting matches against database sequences. The default value used was 10, such that 10 matches were expected to be found merely by chance. The BLAST service at NBCI was used for the analyses.

4.4. Cytogenetic characterisation and chromosome organisation of repetitive DNA sequences in *Brassica napus* and its relatives *B. oleracea* and *B. rapa*

For the cytogenetical experiments the following *Brassica* varieties were used: *B. napus* Jockey, *B. oleracea oleracea*, *B. rapa* Rex, and *B. carinata* BRA 489/77.

4.4.1. Cytogenetic characterisation of *B. napus* and its relatives

The method for accumulation, fixation and squash preparation of *Brassica* chromosomes used in this work provided prometaphase and metaphase cells suitable for the cytogenetic characterisation of the three species analysed, namely, *B. napus*, *B. oleracea* and *B. rapa*.

Metaphase chromosome size and morphology after staining with the fluorochrome DAPI were used for the initial characterisation of the three *Brassica* species. In accordance with the literature, chromosome numbers of *B. oleracea* (CC) and *B. rapa* (AA) were 2n= 18 and 20, respectively, whereas the diploid karyotype of *B. napus* (AACC) showed 38 chromosomes. Chromosomes of all three species were very small and similar in size and morphology. Most of the chromosomes from *B. rapa* were smaller than those from *B. oleracea*. Karyotypes of each species are shown in Figure 9. The length of the chromosomes varied between 1.0-2.6 µm in *B. rapa*, 1.8-2.7 µm in *B. oleracea* and and 0.8-2.5 µm in *B. napus*. With respect to the centromeric position, only metacentric and submetacentric chromosomes were found. *B. rapa* presented 12 metacentric and 8 submetacentric chromosomes, whereas *B. oleracea* presented 6 metacentric and 12 submetacentric chromosomes. The karyotype of *B. napus* consisted of 20 metacentric and 18 submetacentric chromosomes. Unquestionable identification of individual chromosomes was in most case not possible, due to their very small and similar sizes. Determination of chromosome origin in *B. napus*, whether from the A or from the C genome, was not possible on the basis of morphometric analysis alone.

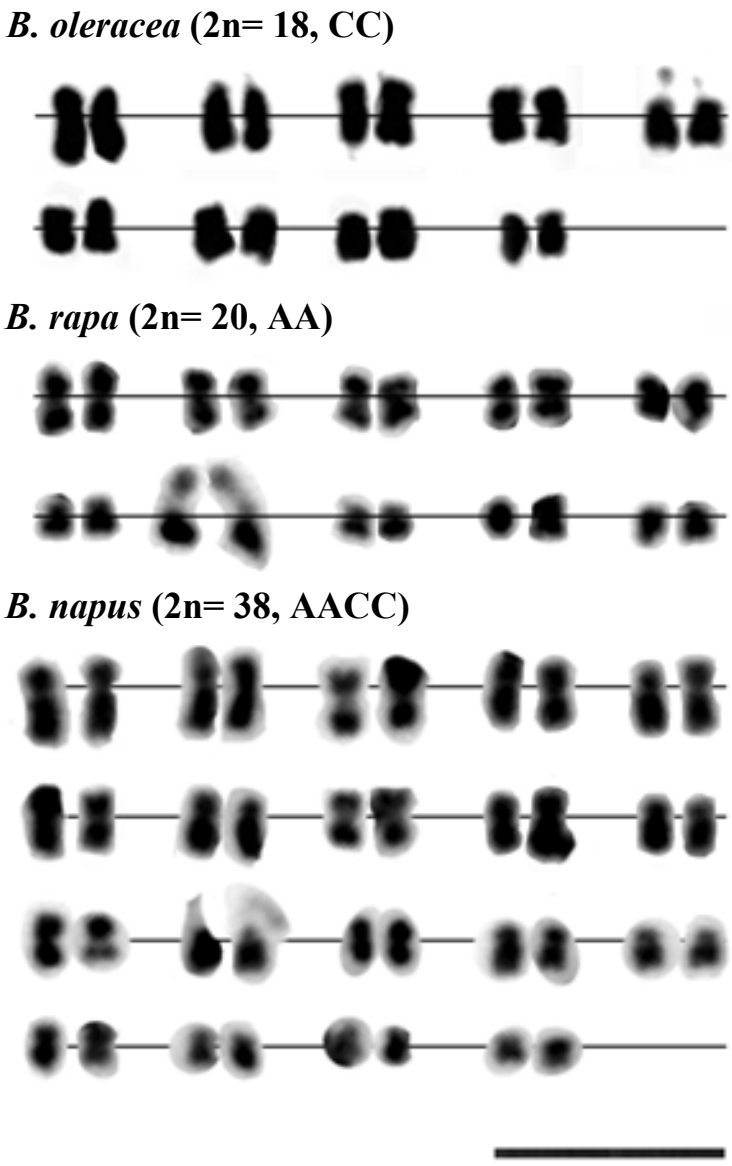


Figure 9. Karyotypes of *B*. *oleracea*, *B*. *rapa* and *B*. *napus* after DAPI staining, based on Figure 10 a, b and c, respectively. Scale bar= 10 µm.

Despite the presence of some blocks of chromatin strongly stained with DAPI in the interphase nuclei of all three species analysed, DAPI staining alone did not produce any $DAPI^+$ banding pattern in the metaphase chromosomes (Figure 10). However, after the FISH procedure bands strongly stained with DAPI were frequently observed in interphase nuclei and in the centromeric region of the pro- and metaphase chromosomes. Interphase nuclei and metaphase chromosomes of *B*. *oleracea* and *B*. *rapa* as well as prometaphase chromosomes of *B*. *napus* stained with DAPI after FISH are shown in Figure 11.

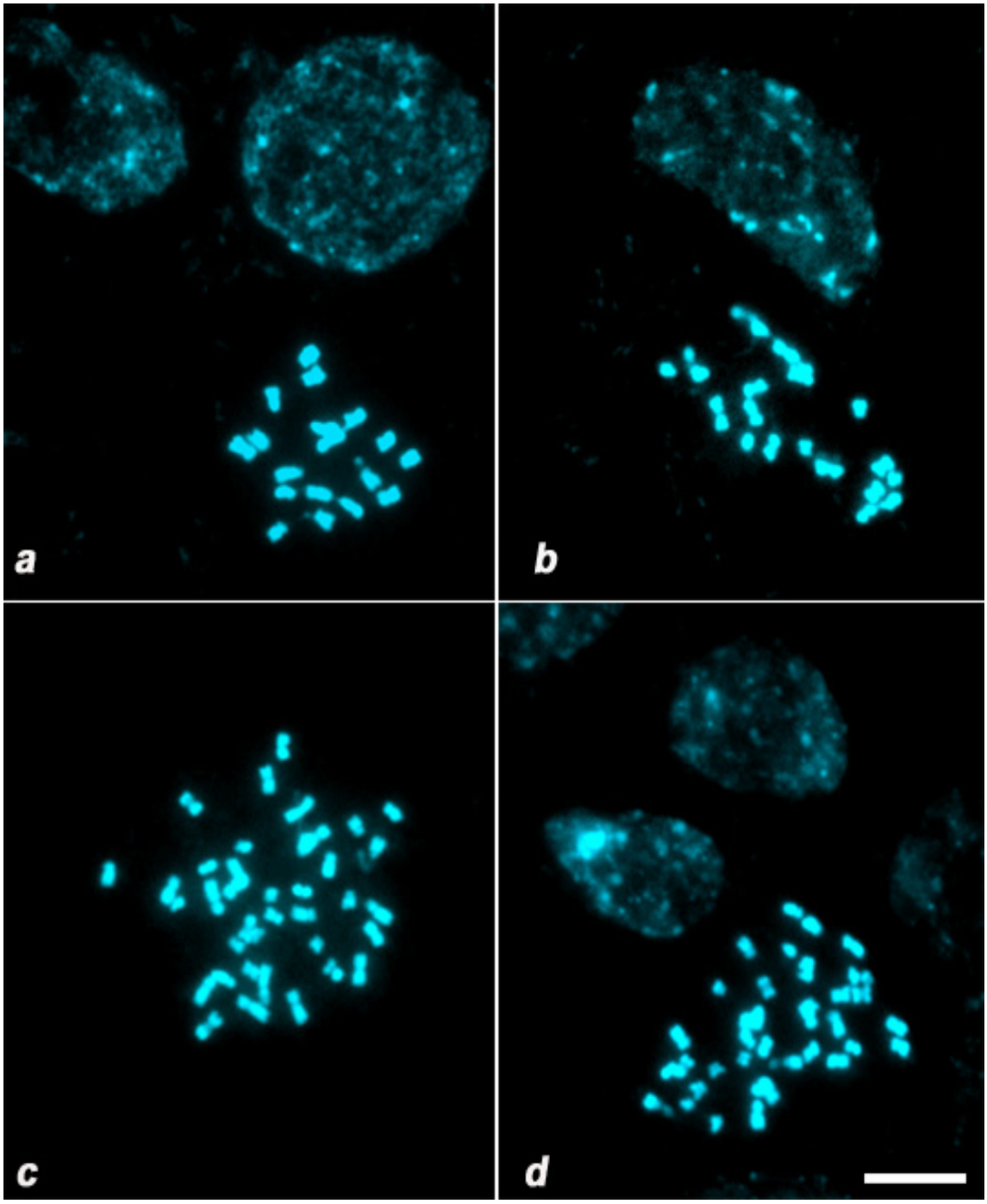


Figure 10. Interphase nuclei and mitotic metaphases of (a) *B*. *oleracea*, (b) *B*. *rapa* and (c, d) *B*. *napus* stained with DAPI. Scale bar= 10 µm for all panels.

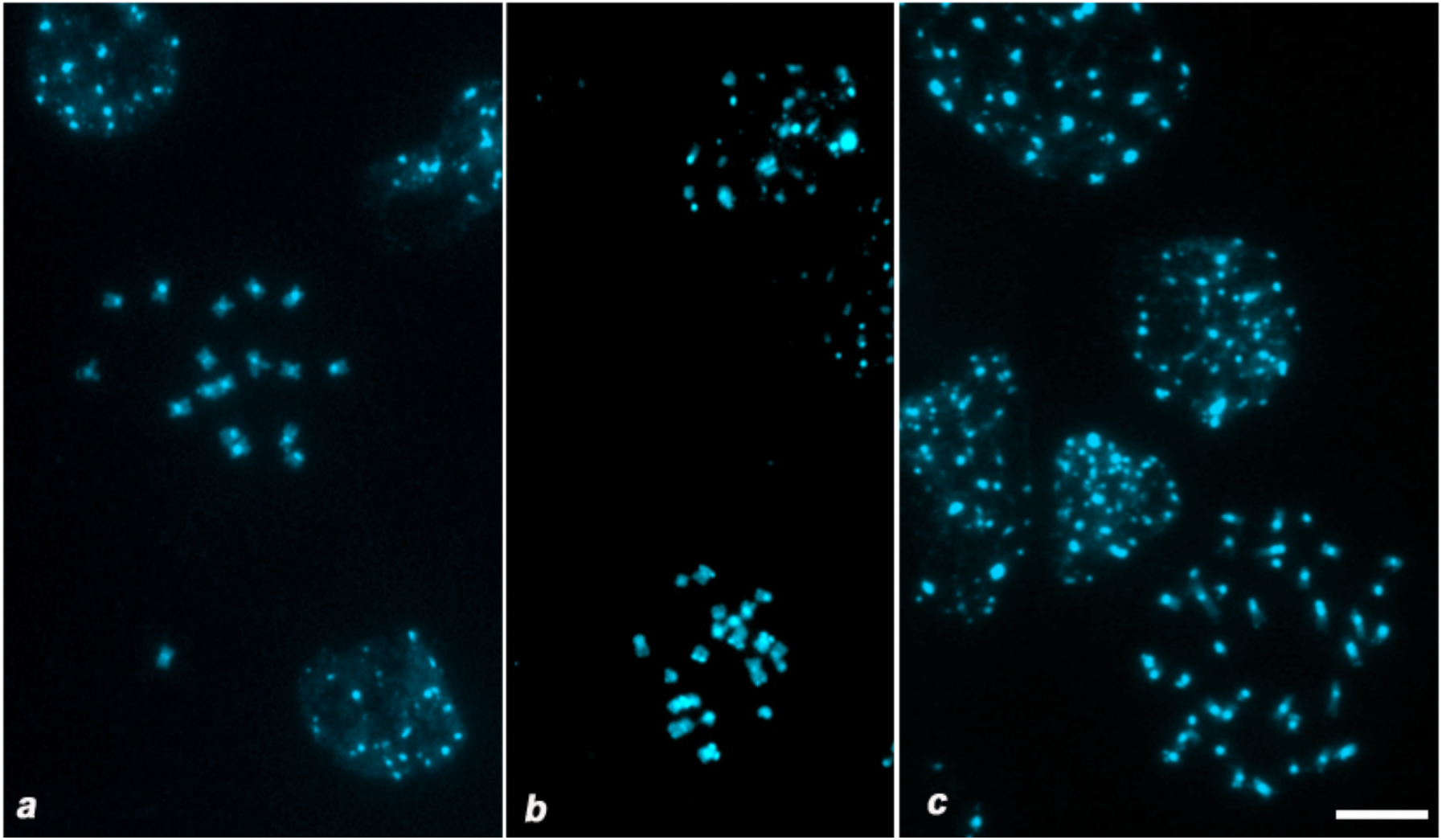


Figure 11. Interphase nuclei and mitotic metaphases of *B*. *oleracea* (a), *B*. *rapa* (b) and prometaphase of *B*. *napus* (c) stained with DAPI after FISH. Scale bar= 10 μm for all panels.

4.4.2. Chromosome organisation of repetitive DNA sequences in Brassica

For the establishment of the FISH procedure two probes were first used as controls: a 45S rDNA probe, generated by PCR, and the "5+45S rDNA" plasmid probe, containing part of both the 5S and the 45S rDNA genes. After labelling by PCR, nick translation or using the DIG-high prime DNA labelling kit, probes were tested in FISH experiments. Probes labelled with DIG-high prime DNA labelling kit were not suitable for FISH, giving high amount of background hybridisation in the cytological preparations, whereas probes labelled by PCR and nick translation provided suitable hybridisation signals. Therefore, PCR and nick translation were the labelling methods used in the subsequent experiments.

45S rDNA and "5+45S rDNA"

Figures 12 and 13 show the results of FISH with probe 45S rDNA and probe "5+45S rDNA", respectively, to the somatic metaphase chromosome complement of *B. oleracea*. In each image, the rDNA signals are large and discrete, and representative of those seen in a large number of cells analysed. With 45S rDNA probe *B. oleracea* showed four hybridisation sites, one pair of large loci localised in the Nucleolus Organiser Region (NOR) of the satellite-carrying chromosomes, whereas the other pair of loci localised near the telomeric region and showed heteromorphism in size, with one large and one small locus (arrows in Figure 12 and 13). With probe "5+45S rDNA" hybridisation sites were present in six chromosomes from the diploid complement. Four chromosomes presented hybridisation sites as described for 45S rDNA, whereas the additional chromosome pair with hybridisation sites showed two adjacent loci localised near the centromeric region (see arrowheads in Figure 13). With both probes hybridisation signals from the small 45S locus, localised near the telomeric region, were not always seen (Fig 13 e, g).

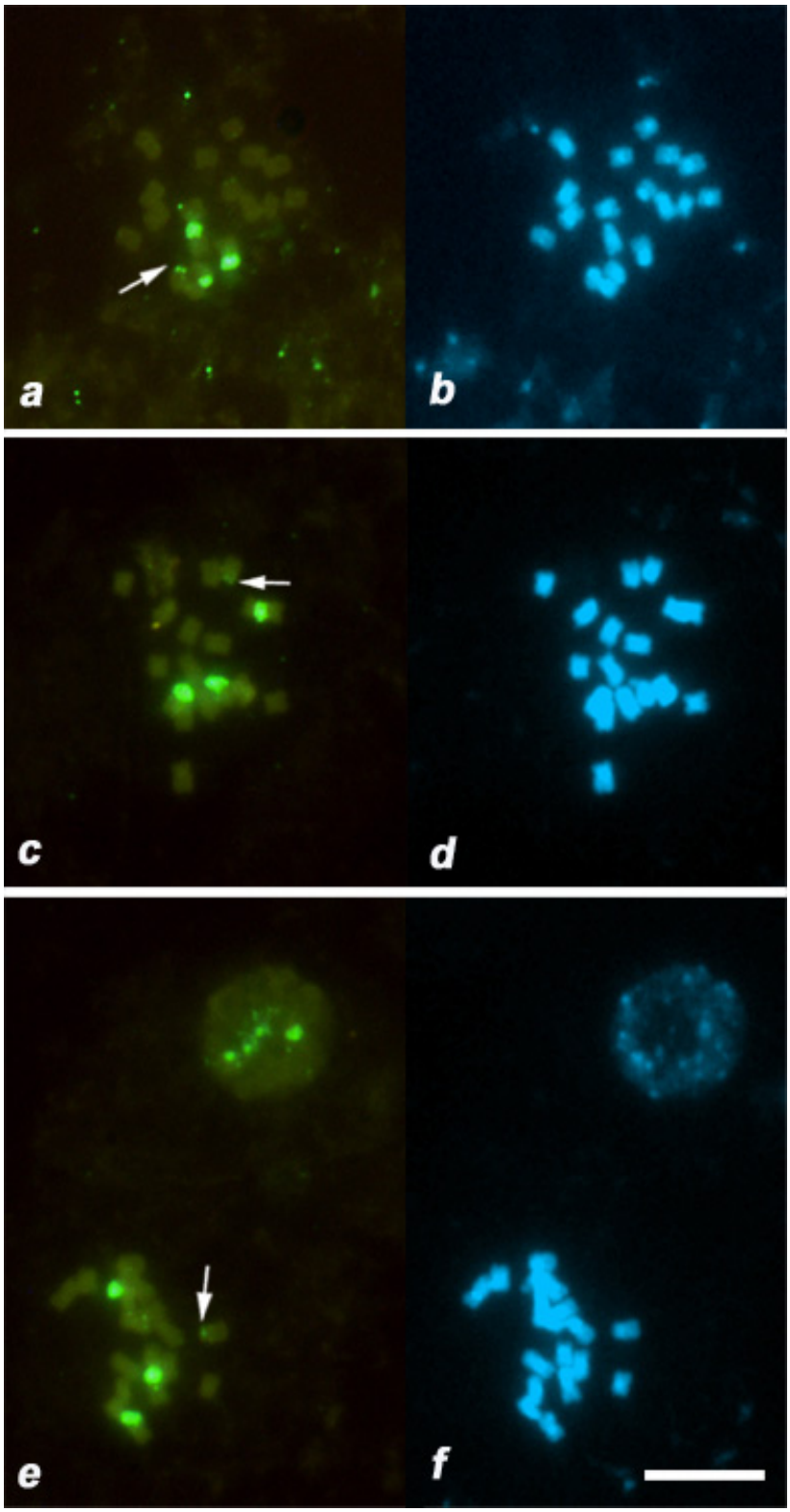


Figure 12. FISH with 45S rDNA probe hybridised to somatic metaphase chromosomes of *B. oleracea* (a, c, e). Panels b, d and f show the same cells counterstained with DAPI. Arrows indicate the small locus. Scale bar= 10 µm for all panels.

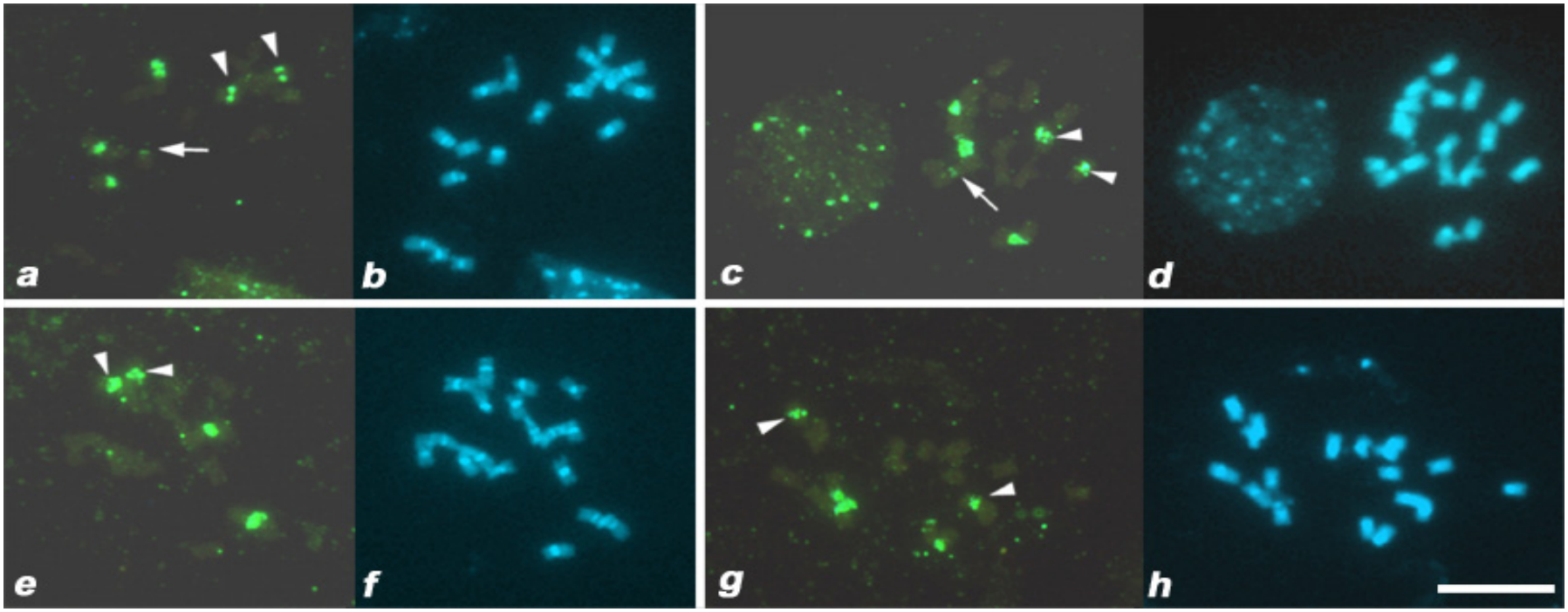
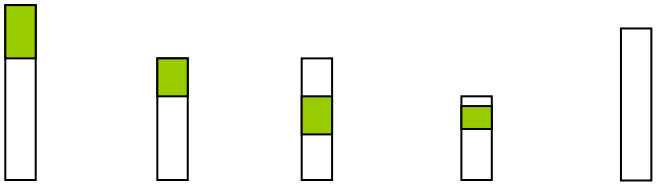


Figure 13. FISH with "5+45S rDNA" probe hybridised to somatic metaphase chromosomes of *B. oleracea* (a, c, e, g). Panels b, d, f and h show the same cells counterstained with DAPI. Arrows and arrowheads indicate the small 45S locus and the adjacent 5S loci, respectively. Scale bar= 10 μm for all panels.

B. *rapa* and *B*. *napus* chromosomes were only probed with "5+45S rDNA" and showed, as in *B*. *oleracea*, clear rDNA signals in several metaphases (Figure 14). This probe allowed the identification of five chromosome types (based on Hasterok *et al*., 2001, with modification), named chromosomal type I to V (see schema for chromosomal types in Table 4 and arrows in Figure 14 e). Chromosomes of type I were large with a terminal hybridisation site. Chromosomes of type II were of intermediate sizes with a terminal hybridisation site and chromosomes of type III presented different lengths with one (or two very close) interstitial hybridisation site. Chromosomes of type IV, containing a terminal-subterminal hybridisation site, were much smaller than the other types and chromosomes of type V, with varying sizes, showed no hybridisation site. Table 4 summarises the number of chromosomes from each type in the three species. Three, six and nine chromosome pairs were identified as containing rDNA genes in *B*. *oleracea*, *B*. *rapa* and *B*. *napus*, respectively.

Table 4. Number and types of chromosomes after FISH with probe "5+45S rDNA" in the diploid complement of *B*. *napus* and its progenitors.

			Chromosomal type				
							
Species	2n	Genome	I	II	III[a, b]	IV	V[a]
B. *oleracea*	18	CC	2	2	2	0	12
B. *rapa*	20	AA	0	2	8	2	8
B. *napus*	38	AACC	2	4	10	2	20

[a]Chromosomal types III and V represent chromosomes which can be of different length

[b]Chromosomal type III may have one or two closely localised interstitial hybridisation site

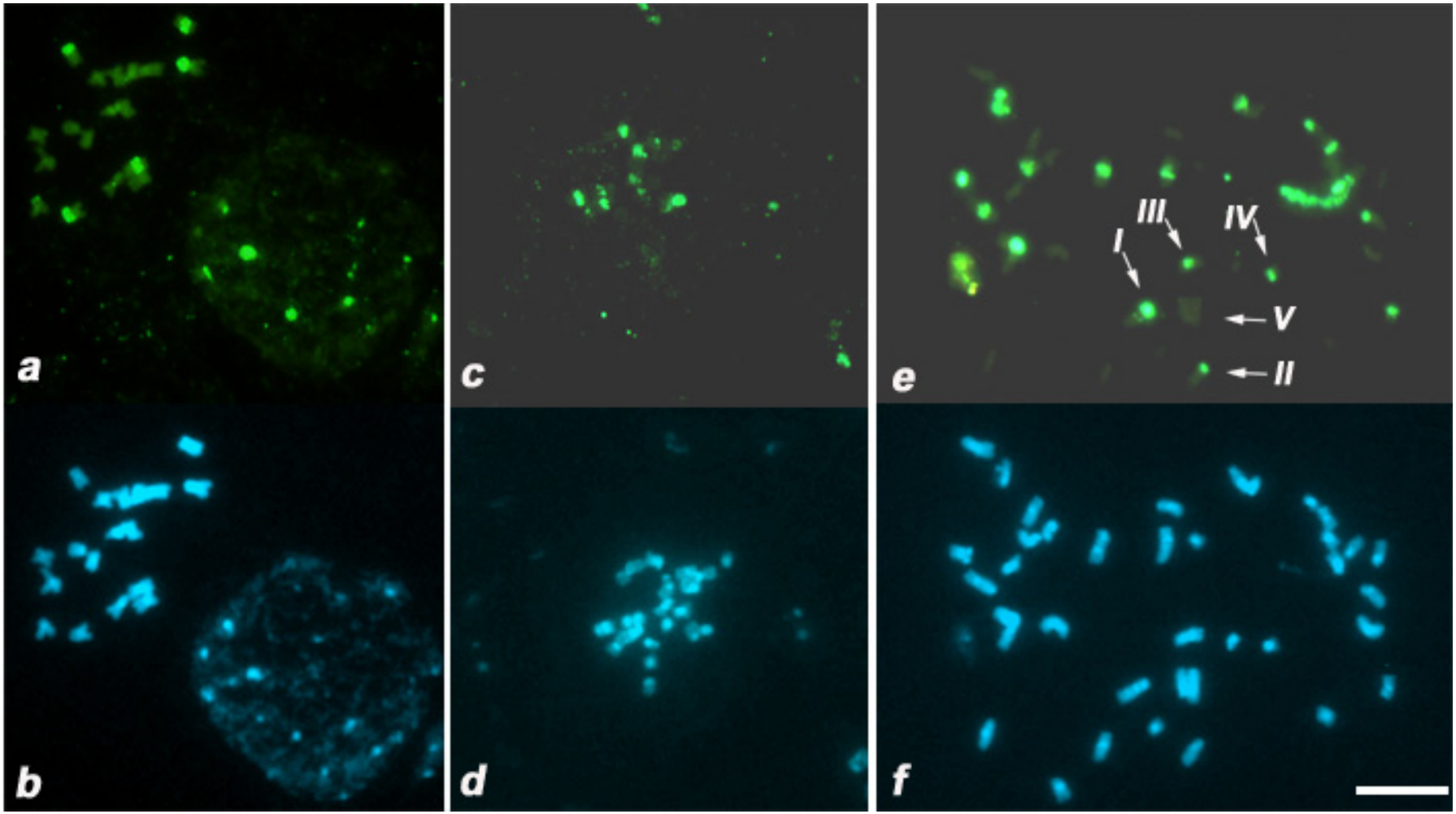


Figure 14. FISH with "5+45S rDNA" probe hybridised to somatic metaphase chromosomes of *B. oleracea* (a), *B. rapa* (c) and *B. napus* (e). Panels b, d and f show the same cells counterstained with DAPI. Chromosomal types I-V are indicated in e. Scale bar= 10 μm for all panels.

pBo2.94 and pBo1.173

Two additional controls were selected and investigated by FISH: pBo2.94 and pBo1.173. These clones were selected from the plasmid libraries, owing to the very strong hybridisation signals in the dot blot experiments with DNA from both the A and the C genomes, indicating a high degree of repetition (like clone C6 in Figure 4). Experiments using these sequences as probes were performed in four species of *Brassica*: *B*. *oleracea*, *B*. *rapa*, *B*. *napus* and *B*. *carinata*. Large hybridisation sites, characteristic for tandemly repeated sequences, were found in all four species at the centromeric/ pericentromeric region of several chromosomes. These signals colocalised with brightly DAPI stained heterochromatic blocks. Table 5 summarises the number of signals found in each species and Figures 15 and 16 show the results of FISH with probes pBo2.94 and pBo1.173, respectively, to the somatic chromosome complement of *Brassica* species.

Table 5. Number of chromosomes with strong pericentromeric hybridisation sites for probe pBo2.94 and pBo1.173 in four *Brassica* species.

Species	Genome (2n)	Probes	
		pBo2.94	pBo1.173
B. *oleracea*	CC (18)	12-14	12
B. *rapa*	AA (20)	(not tested)	16
B. *napus*	AACC (38)	28-30	28-30
B. *carinata*	BBCC (34)	12	12

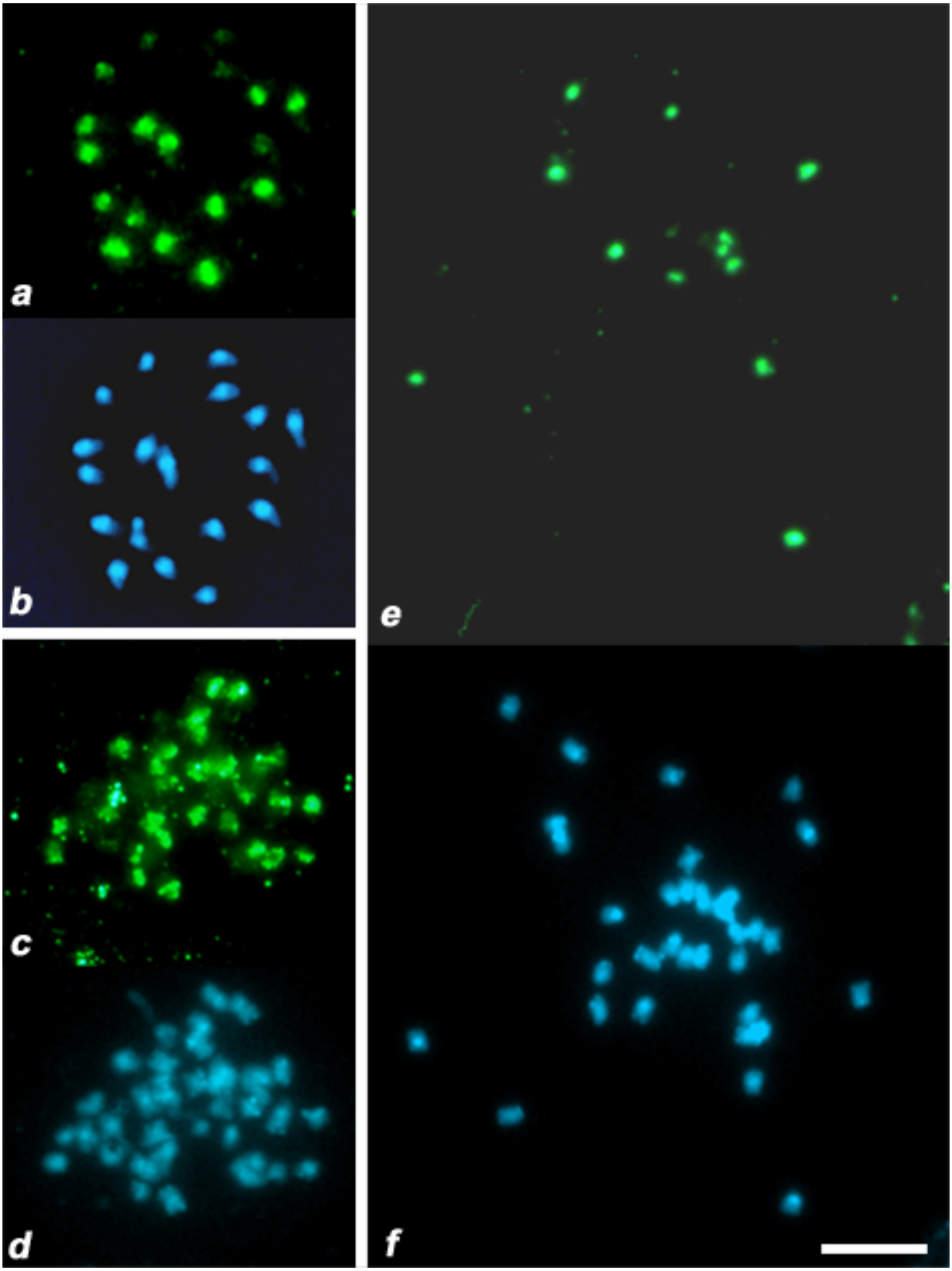


Figure 15. FISH with probe pBo2.94 hybridised to somatic metaphase chromosomes of *B. oleracea* (a), *B. napus* (c) and *B. carinata* (e). Panels b, d and f show the same cells counterstained with DAPI. Scale bar= 8 µm for a, b and 10 µm for c-f.

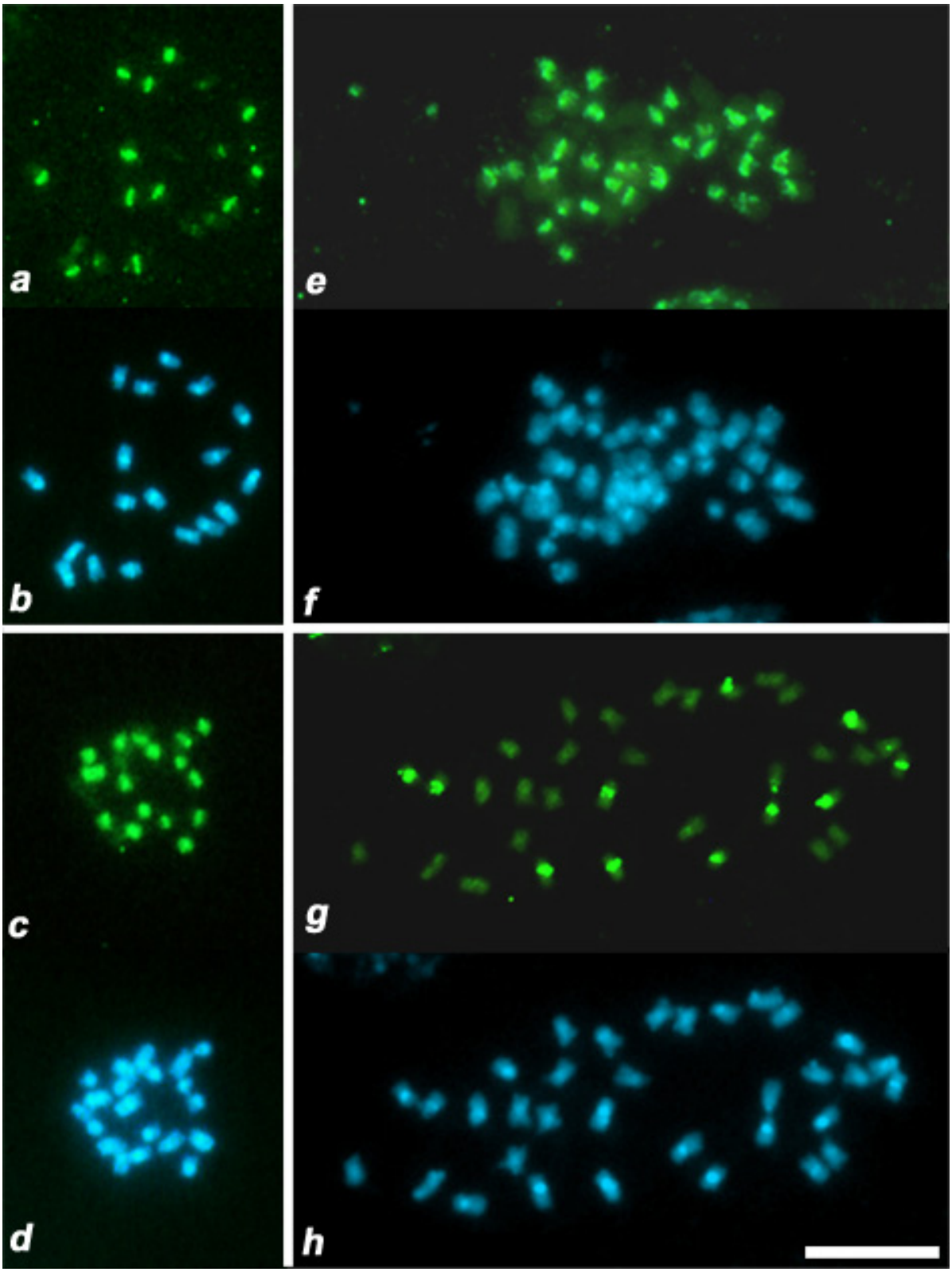


Figure 16. FISH with probe pBo1.173 hybridised to somatic metaphase chromosomes of *B. oleracea* (a), *B. rapa* (c), *B. napus* (e) and *B. carinata* (g). Panels b, d, f and h show the same cells counterstained with DAPI. Scale bar= 8 µm for a-d and 10 µm for e-h.

While hybridisation signals could be detected in FISH with highly repetitive sequences, initially no signal could be seen using the sequences enriched in the C genome. Since these probes were expected to produce signals weaker than that of pBo2.94 and pBo1.173, FISH experiments were repeated using pBo1.6, pBo1.27 and pBo2.157 as probes and analyses were carried out on a more sensitive microscope, in which even weak hybridisation signals could be visualised. The results are presented in Figures 17-20.

pBo1.6

As expected from the Southern blot analyses, probe pBo1.6 gave the most interesting results, differentiating the A and C genomes. Fluorescent *in situ* hybridisation showed that this repetitive sequence is present in all chromosomes of *B*. *oleracea*, but in none of the chromosomes of *B*. *rapa* for most cells. In *B*. *oleracea* signals were distributed in the interstitial and/or telomeric/subtelomeric region of all chromosomes (Figure 17 a). The amount of hybridisation signals varied among the chromosomes, with some chromosomes showing more hybridisation sites than others. In *B*. *rapa* no signal was detected in most of the cells and when detected they had a weaker intensity than in *B*. *oleracea* and were often detected as a single spot on only one chromatid in the telomeric/subtelomeric region of two to eight chromosomes (Figure 17 c).

B. *napus* showed also a variable number of chromosomes hybridising with the repetitive sequence pBo1.6. Eighteen to 24 chromosomes of *B*. *napus* presented hybridisation sites for pBo1.6 in the interstitial and/or telomeric/subtelomeric region. The amount of hybridisation signals varied, with larger chromosomes showing frequently more and stronger hybridisation signals than smaller ones, which showed few or no hybridisation signal at all (Figure 17 e). In some chromosomes of some cells, several hybridisation sites were detected along almost the whole chromosome length (see arrows in Figure 17 e).

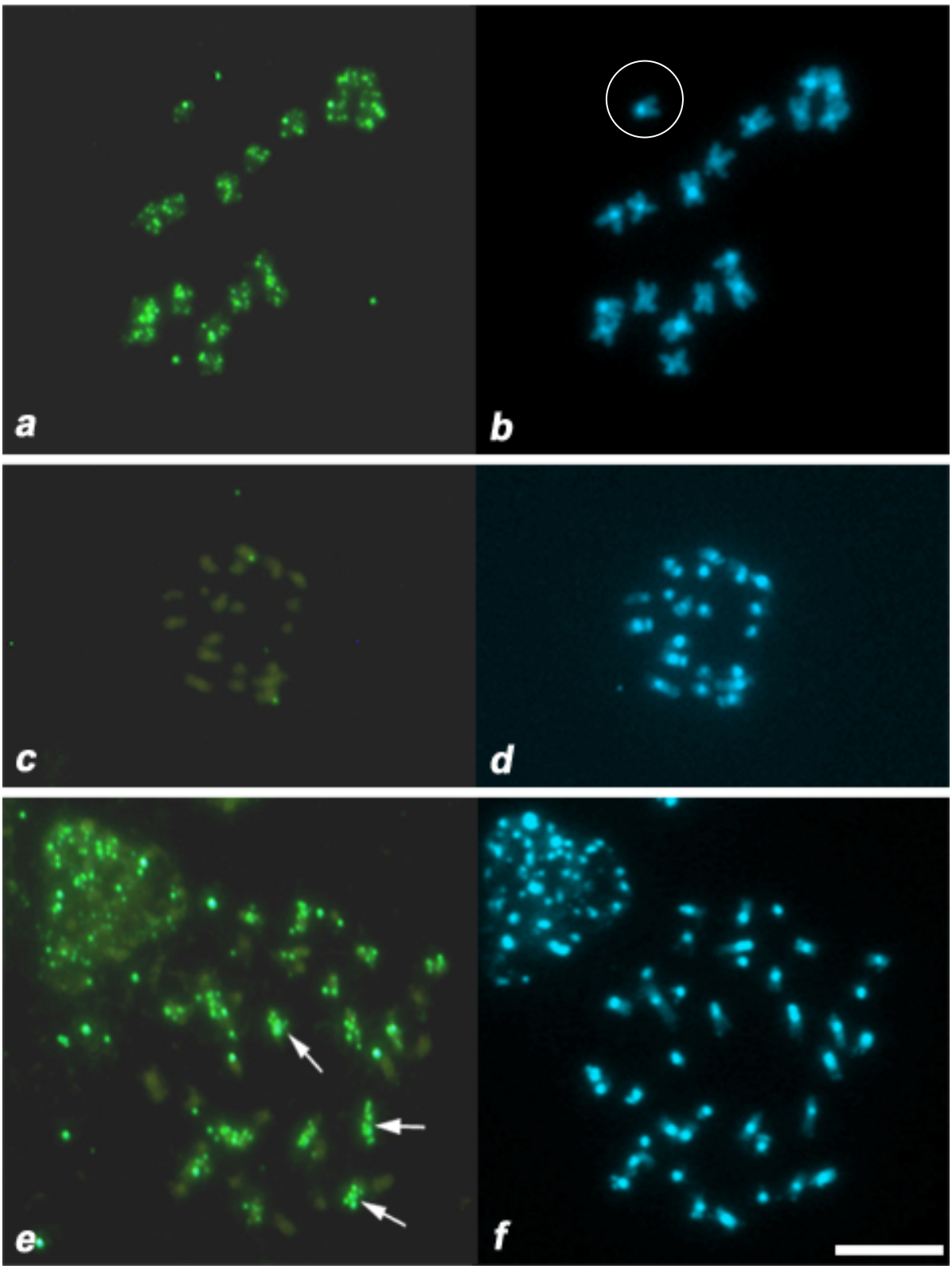


Figure 17. FISH with probe pBo1.6 hybridised to somatic metaphase chromosomes of *B. oleracea* (a), *B rapa* (c) and *B. napus* (e). Panels b, d, and f show the same cells counterstained with DAPI. Arrows in e indicate chromosomes labelled over its almost entire length. Circled chromosome in b was outside the metaphase plate. Scale bar= 10 µm for all panels.

pBo1.27

Probing *B*. *oleracea* chromosome spreads with labelled pBo1.27 resulted in dispersed hybridisation in all chromosomes giving signals of intermediate strength. Hybridisation sites were, however, not homogeneously distributed in the chromosome complement, with some chromosomes exhibiting more hybridisation sites than others (Figure 18 a). Probe pBo1.27 hybridised in *B*. *rapa* in the same way as for *B*. *oleracea* in all or almost all chromosomes, but the signals seemed to be weaker. Again some chromosomes showed more hybridisation sites than others (Figure 18 c). *B*. *napus* presented 14-16 chromosomes with dispersed signals of intermediate strength. The remaining chromosomes exhibited weak, very weak or no signal at all (Figure 18 e).

pBo2.157

The hybridisation pattern of probe pBo2.157 was very similar to that of pBo1.27, with dispersed hybridisation signals in the chromosomes of all three species. In *B*. *oleracea* all 18 chromosomes showed hybridisation signals of intermediate intensity (Figure 19 a), whereas in *B*. *rapa* signals seemed to be weaker and not in all chromosomes (Figure 19 c). Background hybridisation hampered the quantification of the chromosomes without signals in *B*. *rapa*. For *B*. *napus* it was also not possible to determine the number of chromosomes with hybridisation sites for pBo2.157, due to the low quality of the metaphases found after FISH with this sequence. However it was possible to identify several chromosomes labelled and also chromosomes with weak or no signal (Figure 19 e).

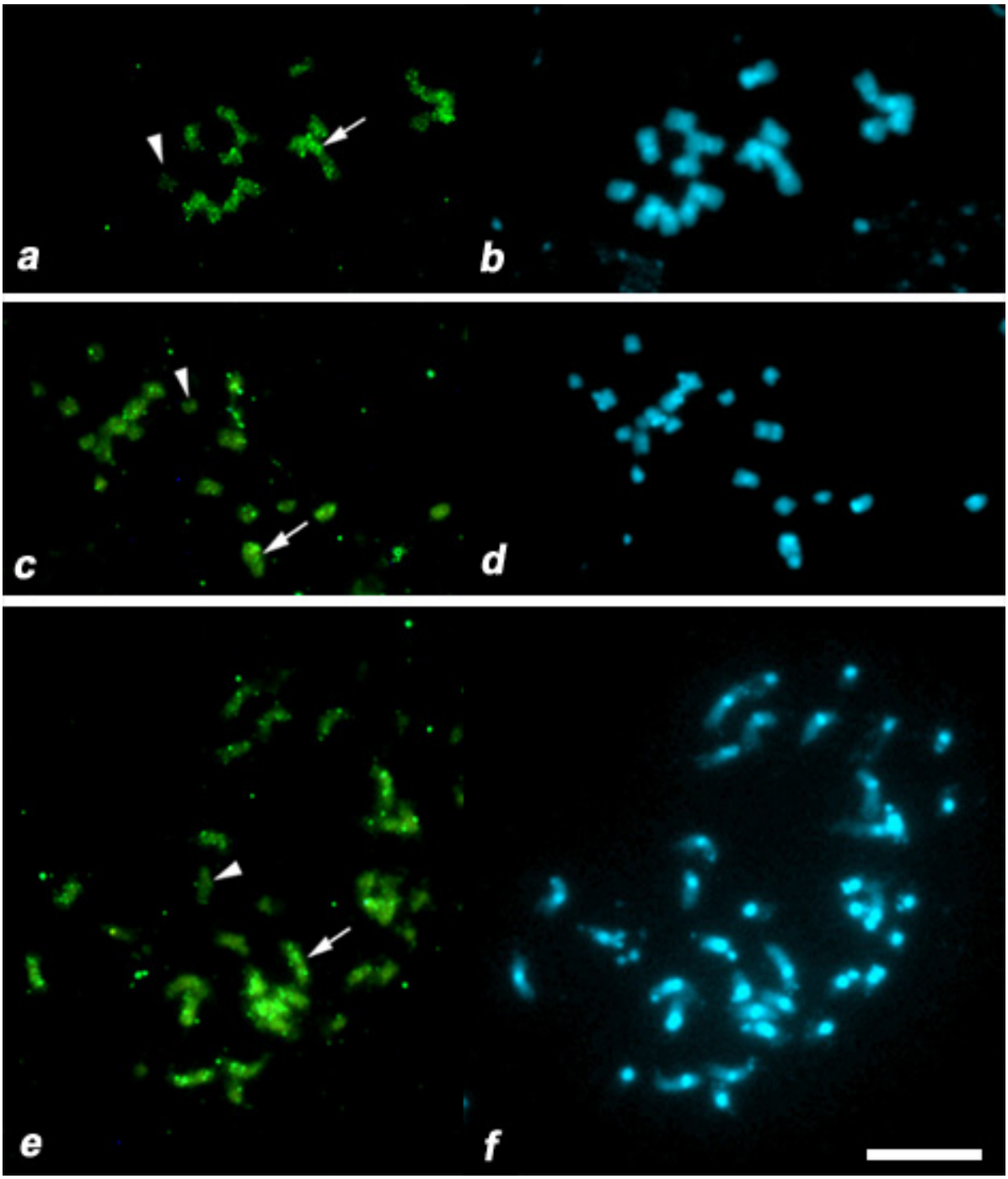


Figure 18. FISH with probe pBo1.27 hybridised to somatic metaphase chromosomes of *B. oleracea* (a), *B. rapa* (c) and *B. napus* (e). Panels b, d, and f show the same cells counterstained with DAPI. Chromosomes with less and more hybridisation sites are indicated in a, c and e by arrowheads and arrows, respectively. Scale bar= 10 µm for all panels.

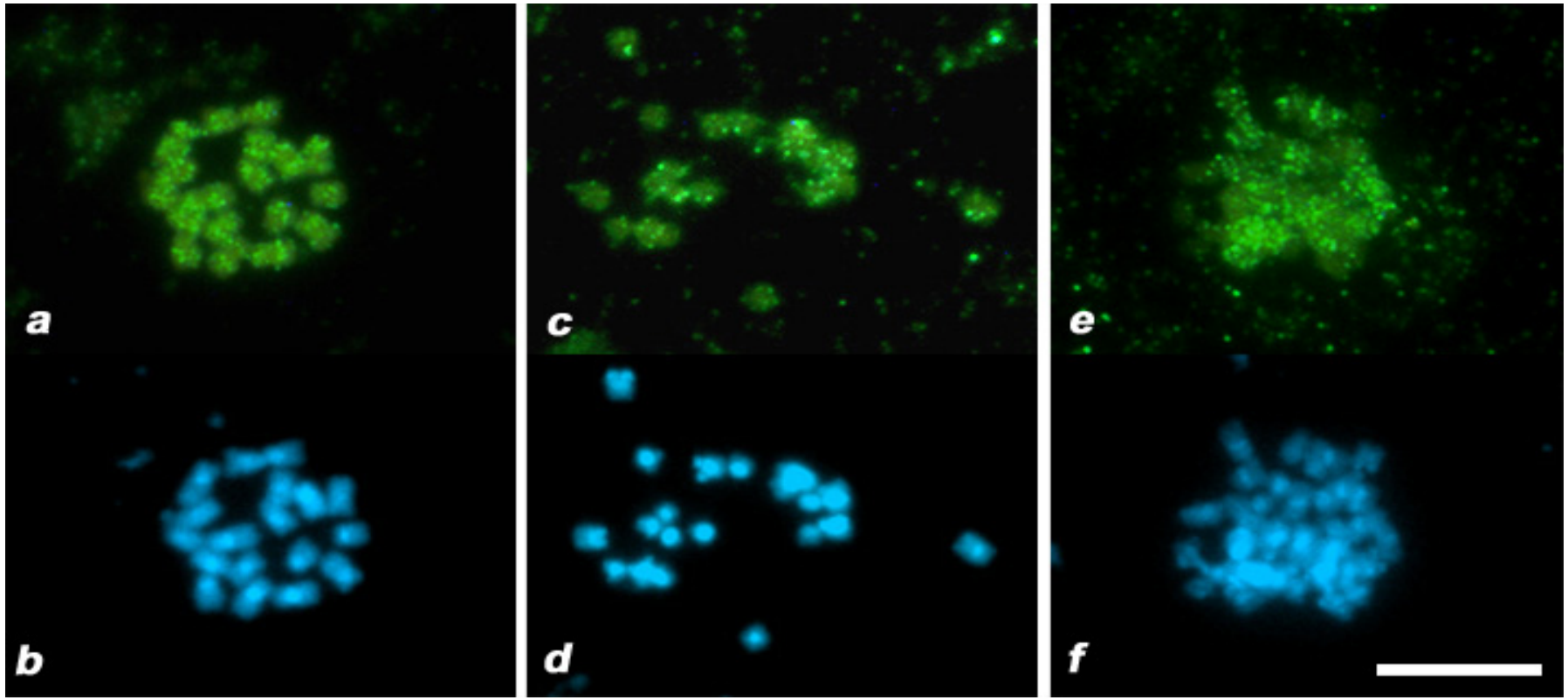


Figure 19. FISH with probe pBo2.157 hybridised to somatic metaphase chromosomes of *B*. *oleracea* (a), *B*. *rapa* (c) and *B*. *napus* (e). Panels b, d, and f show the same cells counterstained with DAPI. Scale bar= 10 µm for all panels.

In general, hybridisation signals after FISH using probes pBo1.6, pBo1.27 and pBo2.157 were absent on the centromeric regions of *Brassica* chromosomes. In interphase nuclei hybridisation sites for the sequences enriched in the C genome were dispersed throughout the whole nucleus, whereas signals from pBo2.94 and pBo1.173 were found clustered in some regions, coinciding with most of the $DAPI^+$ blocks from the interphase nuclei (Figure 20).

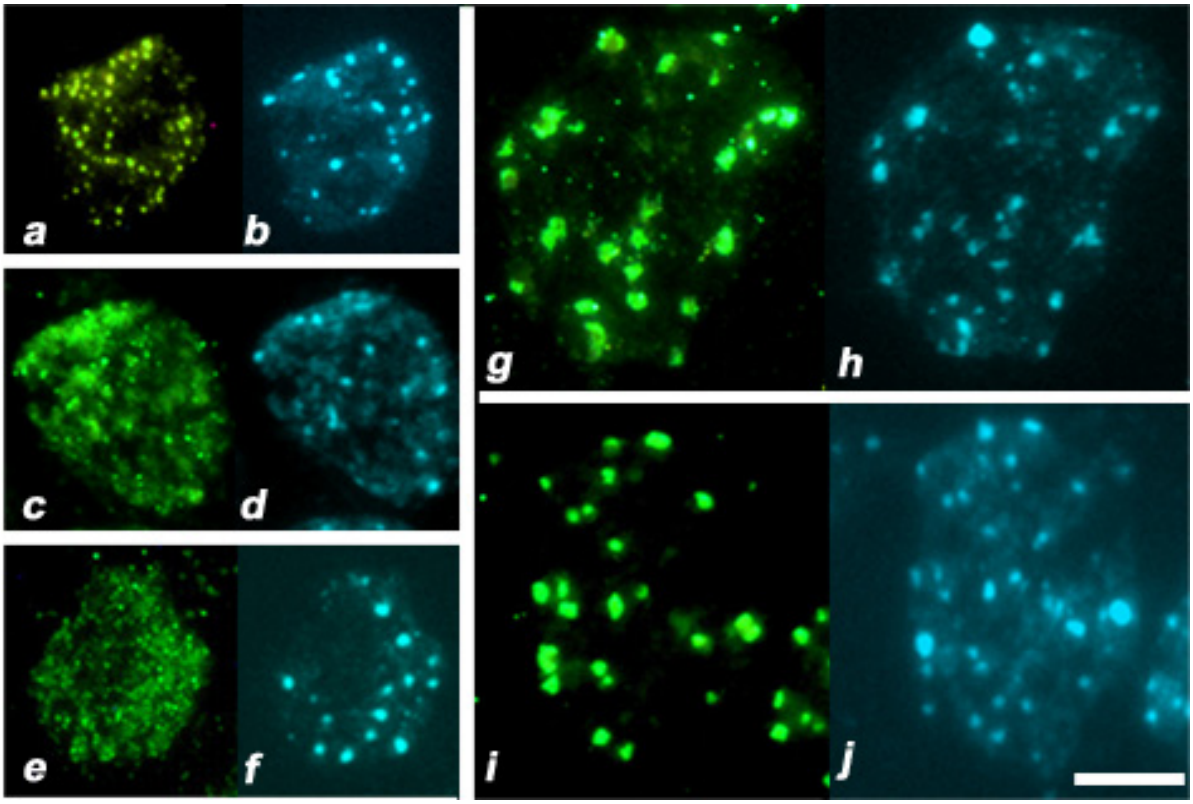


Figure 20. FISH with probes pBo1.6 (a), pBo1.27 (c) and pBo2.157 (e) hybridised to interphase nuclei of *B*. *oleracea* and with probes pBo2.94 (g) and pBo1.173 (i) hybridised to interphase nuclei of *B*. *napus*. Panels b, d, f, h and j show the same cells counterstained with DAPI. Scale bar= 10 µm for all panels.

A short overview of the results obtained with fluorescent *in situ* hybridisation using the newly isolated repetitive sequences as probes and the homology of these repeats to known sequences is shown in Table 6.

Table 6. Characteristics of newly isolated repetitive DNA sequences from *B*. *oleracea*. Homology to known sequences, chromosome location after FISH and frequency of signals on the chromosomes of *B*. *oleracea*, *B*. *rapa* and *B*. *napus*.

Clone (size in bp)	Homologous sequences in databases	Chromosomal location	Frequency of signals in the genomes of:		
			B. *oleracea* (2n= 18, CC)	*B*. *rapa* (2n= 20, AA)	*B*. *napus* (2n= 38, AACC)
pBo1.6 (203)	telomere-like DNA repeat	telomeric and interstitial	frequent in all chromosomes	absent (or only in the telomere from 2-8 chromosomes)	frequent in 18-24 chromosomes and rare or absent in the others
pBo1.27 (182)	*Brassica rapa* DNA for an *En/Spm*-type transposon-like ORF	dispersed over the chromosomes	frequent in most and rare in some chromosomes	Similar to C genome but less frequent	frequent in 14-16 chromosome and rare or absent in the others
pBo2.157 (206)	*Brassica napus* DNA for an *En/Spm*-type transposon-like ORF	dispersed over the chromosomes	frequent in all chromosomes	less frequent than in C genome in some chromosomes and absent in others	frequent in some chromosomes and less frequent, rare or absent in others
pBo2. 94 (238)	*Brassica* centromeric DNA repeat	in tandem in the centromeric region	present in 12-14 chromosomes	(not tested)	present in 28-30 chromosomes
pBo1.173 (158)	*Brassica* centromeric DNA repeat	in tandem in the centromeric region	present in 12 chromosomes	present in 16 chromosomes	present in 28-30 chromosomes

5. Discussion

5.1. Search for *B. oleracea* species-specific repetitive sequences

In the present work *B. oleracea* repetitive sequences useful as molecular cytogenetic markers were searched to assist in the physical identification of highly homeologous A and C genome chromosomes or chromosome segments in *B. napus* (2n= 36, AACC). It is well know that the species *B. oleracea* (2n= 18, CC) and *B. rapa* (2n= 20, AA) are very closely related. Several publications using molecular markers have shown the high homeology between these two genomes (Palmer *et al*., 1983, Song *et al*., 1988 and 1995, Parkin *et al*., 1995, Sharpe *et al*., 1995, Cheung *et al*., 1997). For instance, the nucleotide composition of a highly repetitive sequence family in *B. rapa* shows 98% sequence identity with that of *B. oleracea* (Lakshmikumaran and Ranade, 1990) and comparisons of rDNA intergenic spacer, which display extensive variation even among closely related species, revealed high sequence homology for these species (Bhatia *et al*., 1996). Meiotic studies in *B. campestris* (Syn: *B. rapa)- B. alboglabra* (Syn: *B. oleracea* var. *alboglabra*) addition lines have also revealed various degrees of homeology between the different chromosomes of the A and C genomes (Cheng *et al*., 1994). Moreover, the distinction between A and C genomes in *B. napus* through genomic *in situ* hybridisation was not possible, because of the high degree of cross-hybridisation between the chromosomes and the genomic DNA of both species used as probe (Snowdon *et al*., 1997a, Bellin and dos Santos, 2002). In view of this close relationship between *B. oleracea* and *B. rapa*, it was not expected to find a great number of C genome-specific repetitive sequences. For this reason a phage library, that enables the screening of many clones in a short period of time, was initially chosen to screen for *B. oleracea* specific repetitive DNA.

To find repetitive DNA sequences from *B. oleracea* the strategy of shot-gun cloning of DNA fragments after digestion with *Tsp*509I, a frequent cutter restriction enzyme, was used, followed by selection of clones with repetitive DNA through replica plaque hybridisation using total DNA from *B. oleracea* and *B. rapa* as probes. After the plaque hybridisation, however, only 0.45% to 2.4% of the clones from the three libraries analysed could be identified as containing repetitive DNA. A much higher amount of repetitive DNA was expected to be found, since repetitive sequences form an important part of the eukaryotic DNA and in higher plants they may account for 20

to more than 90% of the genome (Flavell, 1974, 1980). Also in *Brassica* species several families of repetitive sequences have been found (Ouyang and Buell, 2004).

In *B*. *nigra*, for example, Gupta *et al*. (1990) isolated a tandem repeated sequence that represents alone more than 2.5% of the haploid genome from this species and Beridze (1975) found, by gradient centrifugation, that the satellite DNA of *B*. *nigra* represents 37% of the nuclear genome. Uzunova (1994), isolating DNA probes for RFLP analyses in rapeseed, found that about 19% from 1650 clones of a *B*. *napus* plasmid library were of repetitive nature, indicated by the strong hybridisation signals with rapeseed total DNA. A *Hind*III family of highly repetitive DNA sequences from *B*. *napus* represents alone 3.9% of the genome of *B*. *oleracea* (Xia *et al*., 1993).

A possible explanation for detecting only a low proportion of clones with repetitive DNA in the phage libraries may lie on the small amount of DNA present in each phage plaque in combination with the type of DNA used as probe, i.e. total genomic DNA. Probably the quantity of middle and low repetitive sequences present in the probe was not sufficient to produce hybridisation signals strong enough to be visualised and only plaques with highly repetitive sequences could be identified. Moreover, the species-specificity from the plaques selected in the phage library 1 was not confirmed. Differences in the replica membranes, caused for example from small air bubbles between the membrane and the agar and background hybridisations, giving false positive signals, are possible explanations for that observation. In conclusion, the phage libraries, in the way they were used, were not suitable for identification of species-specific repetitive sequences from the *B*. *oleracea* genome.

In contrast to the phage libraries, the plasmid libraries used in this work proved to be suitable for the screening of repetitive sequences present in the genome of *B*. *oleracea*. These libraries were constructed and analysed in the same way as the phage libraries but using pUC19 plasmid as vector. The proportion of the clones representing repetitive sequences (21%) was in accordance with the expectation (see above). C genome-enriched clones were found in these libraries, but none of them was strictly C genome-specific.

It could be asked whether the number of clones screened was high enough to find species-specific repetitive sequences. The number of clones that should be screened

for this purpose can not be determined a priori and is rarely mentioned in other publications. However, there are some instances where this number was reported and in this case they are widely variable. For example, De Jeu *et al*. (1997) found species-specific repetitive DNA from *Alstroemeria aurea* by screening only 180 clones of transformed *E*. *coli*. Li *et al*. (1995) screened about 500 white colonies to find one specific repeated sequence from *Haynaldia villosa*, an annual relative of wheat. On the other hand, Calderini *et al*. (1997) screened about 800 clones from a partial plasmid DNA library to find species-specific repetitive DNA in *Medicago* and libraries consisting of 1000 and 700 recombinant clones were screened to find species-specific sequences from *Solanum brevidens* and *S*. *tuberosum*, respectively (Pehu *et al*., 1990). For *Medicago*, Calderini *et al*. (1997) found eight clones that showed hybridisation signals with *M*. *corulea* DNA but not with *M*. *arborea*, two closely related species. However, all *Medicago* clones bore the same sequence, showing that closely related species do not possess a high number of genome-specific repetitive DNA. In *Avena sativa* (2n= 6x=42, AACCDD genome), for instance, no repetitive sequences were identified that could distinguish between the A and D genomes, although A/D and C genome-specific DNA sequences were found (Ananiev *et al*, 2002).

For *Brassica* species, at least two B genome-specific repetitive sequences were found (Gupta *et al*., 1992, Kapila *et al*., 1996b), whereas only one paper reported an A genome-specific middle repetitive DNA (Iwabuchi *et al*., 1991). This *B*. *rapa*-specific repeat was isolated by screening about 1100 recombinant plasmids.

Considering that the genome size of *B*. *oleracea* is 0.630 x 10^9 bp/1C (Arumuganathan and Earle, 1991) and that the average size of the cloned DNA was 800 bp[1], the plasmid libraries together represent about 0.1% of this genome. The number of clones screened in the present work would probably be enough to find species-specific repetitive sequences if the homology of the genomes analysed would have been lower. As shown above, for several species, a smaller number of clones was sufficient for this purpose.

After the dot blot hybridisation several clones showed some differences in the intensity of hybridisation with the A and the C genome of *Brassica*. Fifteen such

[1]Inset size average is bases on the estimated length of the 15 selected clones, which varied from 150-1500 bp.

clones were further analysed by Southern blot, but for only 3 the difference found in the first screening could be confirmed. The remaining clones were either repetitive in both genomes (10 clones) or not repetitive at all (2 clones). This result was not very surprising, since plasmid DNAs extracted with a rapid alkaline procedure were used in the dot blot experiments and DNAs extracted in this way are not free of contaminations, like proteins and RNAs. Differences in the hybridisations with dot blotted replica membranes should be expected and more precise and sensitive methods, like Southern blot hybridisation, are necessary to confirm possible positive results from the dot blot. However, possible experimental errors, like differences in the amount of plasmid DNA loaded in the replica membranes, mistakes in the identification of the loaded clones etc., can not be ruled out as possible causes for the high number of false positive clones found by the dot blot hybridisation.

An analysis of the hybridisation patterns for the 13 clones that were repetitive in the Southern blot hybridisation gives the impression that these clones are organised as dispersed sequence in the *B. oleracea* genome, since none of them showed a pattern of hybridisation like a ladder, typical for tandem repeated sequences. Sequences repeated in tandem usually show this ladder-like pattern of hybridisation when an enzyme that cuts the sequence at least one time is used to digest the genomic DNA investigated. This happens because the different units of the sequence, repeated in tandem, usually exhibit some nucleotide differences, caused by mutations during the evolution of the sequence since its amplification. The repeated units are then not exactly the same. If mutations occur in the recognition site of an enzyme, the enzyme will not cut the modified sequence and fragments of the genomic DNA with different sizes will hybridise with the probed sequence, causing a pattern of hybridisation similar to a ladder. Although the ladder-like pattern of hybridisation have not been found for the 13 repetitive clones analysed by Southern blot hybridisation, it is not possible to determine how these sequences are organised in the genome, since the organisation of an unknown DNA sequence can only be determined by Southern blot if several restriction enzymes are used to digest the genomic DNA. Another way to determine the organisation of repeated sequences in the genome is through *in situ* hybridisation. In this case the location can be seen directly on the chromosomes. Since the objective of the present work was the *in situ* localisation of selected sequences in *Brassica*, digestion with several enzymes, in order to find their genomic organisation by Southern blot, would not be required.

The clear differences found in the intensity of hybridisation with the A and C genomes for the three C genome-enriched clones show that these sequences are present in different copy-number in the genomes of *B*. *oleracea* and *B*. *rapa*. As high stringency conditions were used in the experiments (post-hybridisation washes were made in 0.5x SSC/0.05% SDS at 65°C, giving stringencies of 84-88.6%[2] for the different C genome-enriched sequences) the hybridisation patterns found can not be attributed to a high amount of imperfect DNA-DNA hybrids. The organisation of the C genome-enriched clones in the two genomes, whether in tandem or interspersed in the chromatin throughout the genome, could not be determined based only on the Southern blot analyses for the clones pBo1.6 and pBo1.27, since no recognition site for the enzymes *Eco*RI and *Hind*III were found in their sequences. On the other hand, for pBo2.157 five recognition sites for *Hind*III could occur as the result of a single base pair change, making it likely that many copies of this repetitive sequence in the C genome should contain a *Hind*III recognition site. If pBo2.157 sequences were organised in tandem, ladder-like bands should consequently appear in genomic DNA digested with this enzyme. This was the case, for example, for tandemly repeated sequences from *Glycine max*, *Nicotiana plumbaginifolia* and *Lens culinaris* (Kolchinsky and Gresshoff, 1995, Chung-Mong *et al*., 1997, Galasso *et al*., 2001). These sequences showed a ladder-like hybridisation pattern after Southern blot hybridisation with genomic DNA digested with endonucleases that had only similar restriction site in the cloned sequences. However this was not the case for pBo2.157, indicating that this sequence is dispersed in the genome of *Brassica*. The subsequent *in situ* hybridisation using this sequence as probe confirmed its dispersed organisation. A more detailed discussion of the molecular data of the C genome-enriched clones (pBo1.6, pBo1.27 and pBo2.157) is presented together with the analyses of the chromosomal organisation of these sequences.

5.2. Cytogenetic characterisation of *B. napus* and its relatives

Morphological analysis is a fundamental method of chromosome identification. In some plant species, e.g. *Crepis capillaris* (Jamilena *et al*., 1994), the determination of chromosome length, centromere position and location of the secondary constriction is sufficient for the identification of all chromosome pairs in the karyotype. Mitotic metaphase chromosomes of *Brassica* species, however, do not have enough

[2]Calculated after Meinkoth and Wahl, 1984.

morphological features for their identification since they are very small and similar in size. Nevertheless, the chromosome number, size and morphology were analysed in *B*. *napus* and its relatives, *B*. *oleracea* and *B*. *rapa*, in order to provide an initial karyotypic characterisation of these species.

B. *rapa* (2n= 20, AA) presented a more asymmetrical karyotype than *B*. *oleracea* (2n= 18, CC). Chromosome size in *B*. *rapa* varied from 1.0 to 2.6 µm, whereas in *B*. *oleracea* they varied from 1.8 - 2.7 µm. Additionally, most of the chromosomes of *B*. *rapa* were smaller than that of B. *oleracea* (see Figure 9). Cheng *et al*. (1995) also observed that chromosomes of the C genome are larger than those of the A genome and the inspection of chromosome sizes in several publications also shows differences in size between A and C genome chromosomes (see for example pictures in Hasterock *et al*., 2001 and Snowdon *et al*., 2002). This is in accordance with the differences in genome size between these species (Arumuganathan and Earle, 1991; Narayan, 1998). Since the genome size of *B*. *oleracea* is larger (630 Mbp or 0.97pg/1C) than that of *B*. *rapa* (506 Mpb or 0.84 pg /1C) and the chromosome number is smaller (2n= 18 in *B*. *oleracea* versus 2n= 20 in *B*. *rapa*), then the chromosomes from the A genome should be smaller than those of the C genome, as described in the present work. In *B*. *napus* differences in chromosome size were also conspicuous, with chromosomes length varying between 0.8 to 2.5 µm. It seems to be evident that larger chromosomes originate from the C genome whereas smaller ones originate from the A genome.

Concerning the morphology, only metacentric and submetacentric chromosomes were found in the metaphase of the three *Brassica* species analysed. This is consistent with the results reported on *B*. *oleracea* by Wang *et al*. (1989). However, subterminal and, in few cases, terminal chromosomes were reported for *B*. *napus*, *B*. *rapa* and *B*. *oleracea* (Olin-Fatih and Heneen, 1992, Cheng *et al*., 1995, Hasterock and Maluszynska, 2000a). These discrepancies might be the result of differences in varieties and/or methods used. The use of later prophases and prometaphases in several karyotype descriptions, in contrast with metaphases used in the present work, is probably the most important difference influencing the chromosome morphology, since *Brassica* chromosomes present distal arms not homogeneously condensed at later prophases/ prometaphase and differences in the arm ratio are probable strongly influenced by the total condensation of the chromosome arms.

The metaphases of *Brassica* were analysed after staining with the fluorochrom DAPI. This dye stains preferentially AT rich regions of the chromosomal DNA (see Schweizer 1980, Kapuscinski, 1995) and in many species (e.g. *Pinus densiflora*, Hizume *et al*., 2001, *Vicia*, Navrátilová *et al*., 2003), it produces a characteristic banding pattern in the chromosomes that assist in their identification. In *Brassica*, however, no $DAPI^+$ band was produced in untreated chromosomes, in spite of the presence of DAPI positive chromocenters in interphase nuclei. However, after the FISH procedure the pericentromeric regions of most chromosomes and the telomeric regions of some chromosomes were strongly stained with DAPI in several cells (compare Figures 10 and 11), producing a kind of banding pattern. These banding patterns closely resembled the C-banding patterns described by Olin-Fatih and Heneen (1992), and were also observed in *Brassica* by Maluszynska and Heslop-Harisson (1993) and Snowdon (1997) after ISH experiments. In some preparations the steps of the ISH procedure may remove some DNA from the chromosomes, or modify their organisation, in the same way as the C-banding procedure does.

It was demonstrated, previously and in the present work, that the pericentromeric region of many chromosomes of *Brassica* hybridises with probes that are A-T rich (ex. pBo2.94, 62% of bases AT and pBo1.173, 65% of bases AT) showing that at least pericentromeric regions of *Brassica* chromosomal DNA are in fact AT rich. Why the chromosomes of *Brassica* are uniformly stained with DAPI, in spite of the presence of AT rich regions in many of its chromosomes is perhaps related to the presence of AT-rich dispersed sequences in the *Brassica* genome, what may lead to an even DAPI staining and therefore mask other AT-rich heterochromatic segments in metaphase chromosomes. The absence of DAPI bands in condensed chromosomes in spite of evidences of AT-rich regions were also reported for other species, like in *Vignia* (Galasso *et al*., 1995), *Citrus*, *Poncirus* and *Fortunella* (Miranda *et al*., 1997, Guerra *et al*., 2000).

5.3. Chromosome organisation of rDNA and highly repeated sequences in *Brassica*

45S rDNA and "5+45S rDNA"

The localisation of rDNA genes in *Brassica* was reported in several publications (Maluszynska and Heslop-Harrison, 1993, Snowdon *et al*., 1997b, Fukui *et al*., 1998, Armstrong *et al*., 1998, Hasterock and Maluszyska 2000a, b, Schrader *et al*., 2000, Snowdon *et al*., 2002, Ziolkowiski and Sadowski, 2002), therefore rDNA probes were used as control to establish the FISH procedure in our laboratory, before using the selected clones as FISH probes. Frequently, formamide is used in the stringent washes to produce high stringency under low temperature. Snowdon *et al*. (2000a, 2002) have used stringent washes without formamide to localise the rDNA genes in *Brassica* and found the same result as previously when using formamide (Snowdon, 1997). To avoid the use of this harmful chemical all stringent washes used in the FISH experiments in the present work were performed in 0.1 x SSC at 42°C, giving a stringency of about 73-78%[3] for the C genome-enriched sequences.

Chromosomes of *B*. *oleracea*, *B*. *rapa* and *B*. *napus* were hybridised with a plasmid probe, named here as "5+45S rDNA", containing a 1735 bp *Eco*RI fragment of the 18S rDNA region from the 45S rDNA of *Aedes aegypti* and a 509 bp of 5S rDNA from *Matthiola incana*. These sequences were used as heterologous probes in barley chromosomes with excellent results (Brown *et al*., 1999). *B*. *oleracea* was additionally probed with the 25S region of the 45S rDNA, generated by PCR, using genomic DNA of *Brassica* as template. The results obtained here are comparable with those reporter by Hasterock *et al*. (2001), who described the simultaneous localisation of 25S and 5S rDNA sites, visualised with different fluorophores, in *Brassica* diploids (*B*. *nigra*, *B*. *oleracea* and *B*. *rapa*) and amphidiploids (*B*. *carinata*, *B*. *juncea* and *B*. *napus*). In the present study, however, the 5S and 45S rDNA loci were labelled with the same fluorophore and could not be accessed independently. Only in the case of *B*. *oleracea* the chromosome pair with two adjacent loci for 5S rDNA was easily identified by comparing the results of FISH with both probes (compare Figures 12 and 13). Regardless of the impossibility to distinguish the 5S and 45S rDNA loci, probe "5+45S rDNA" gave excellent results and was, in all three species, useful to identify four-five chromosome types carrying rDNA genes (see Table 4). Three, six

[3]Calculated after Meinkoth and Wahl, 1984.

and nine chromosome pairs were identified as containing rDNA genes in *B*. *oleracea*, *B*. *rapa* and *B*. *napus*, respectively.

There are some differences concerning the number of chromosomes carrying these genes among the different reports, but the results presented here matched completely well with those described previously by Hasterock *et al*. (2001) if the probes for 5S and 45S rDNA are considered together. The different results found in the literature are attributed to differences between the varieties analysed or to differences in sensitivity to detect minor sites either due to the FISH protocol or to the microscope used in the studies.

The variety of *B*. *oleracea* used in our FISH experiments showed a heteromorphism in the size of the hybridisation site for 45S rDNA for one pair of chromosomes. *In situ* hybridisation, like Southern blot hybridisation, is not a fully quantitative technique, but large differences in hybridisation signal strength at different loci correspond to major differences in copy number of the tandem repeat unit (Maluszynska and Heslop-Harrison, 1993). In the *B*. *oleracea* variety used in this work there was an obvious variation in the copy number of the 45S rDNA in one pair of homologous chromosomes carrying this gene. The hybridisation signals at this locus were sometimes even missed in one chromosome of the pair, probably because of difficulties to detect the small number of repeats from the rDNA gene array at this locus. Ribosomal RNA genes in plants are highly variable both in copy number and in intergenic spacer length. Analysis of inheritance indicates that copy number change is rapid, occurring even among somatic cells of individual plants (Rogers and Bendich, 1987). Moreover, rapid changes in rDNA-repeat copy number at single loci were demonstrated in tissue culture from wheat (Leitch *et al*., 1993). Variation in the size of the 45S rDNA signal in a locus was found in cytotypes from *Phaseolus vulgaris* (Pedrosa *et al*., 2003) and heteromorphism in the size of the signal for rDNA between homologous chromosomes have also been reported (e.g. 5S rDNA in *Passiflora glandulosa*, de Melo and Guerra, 2003; 45S rDNA in aphid *Acyrthosiphon pisum*, Mandrioli *et al*., 1999). As *B*. *oleracea* is normally strictly cross pollinated (Becker *et al*., 1999), and the variety *oleracea* is its wild form, heterozygosity for different loci may be expected.

pBo2.94 and pBo1.173

After the establishment of the FISH procedure with the rDNA probes, the repetitive clones selected in this work (pBo1.6, pBo1.27 and pBo2.157) were tested by *in situ* hybridisation. However when using the microscope Zeiss Axiolab Epiflorescence, no signal could be visualised with none of these probes. The absence of signal was interpreted as possibly being related to the degree of repetition of the sequences in the genome. Therefore, two high repetitive clones (pBo2.94 and pBo1.173), that showed very strong hybridisation signals with both the A and C genomes in the dot blot experiments, were selected and tested by FISH. *In situ* hybridisation showed that both probes hybridise with the pericentromeric region of several chromosomes giving very bright signals. Based on these results, it was concluded that the hybridisation signals produced with the C genome-enriched sequences were not strong enough to be detected by the available microscope, probable because of its degree of repetition and/or organisation in the genomes. A more sensitive microscope (Olympus BX60, from the Institute of Human Genetics, Göttingen) was subsequently used and this interpretation was confirmed. Small hybridisation signals, dispersed along the chromosomes, could be detected for all sequences analysed, although unspecific background hybridisation signals also became more evident.

Since clones pBo2.94 and pBo1.173 hybridised with the pericentromeric region of several chromosomes from all three *Brassica* species (see Table 5), we were interested to know whether they were members of a family of highly repetitive sequences from *Brassica* described in the literature or were representatives of a new tandemly repeated centromeric sequence. Therefore, these clones were sequenced and subjected to a DNA database comparison. The results showed that they were members of the *Brassica Hind*III repeat family, related to tRNA genes, that is represented in the GenBank/EMBL/DDBJ/PDB database by more than 20 accessions from *B. oleracea*, *B. rapa*, *B. napus*, *B. juncea*, *B. carinata* and *Diplotaxis erucoides* (Benslimane *et al*., 1986, Reddy *et al*., 1989, Harbinder and Lakshmikumaran, 1990, Lakshmikumaran and Ranade, 1990, Xia *et al*., 1993, 1994, Harrison and Heslop-Harrison, 1995).

The chromosomal location of representatives from this *Hind*III repeat family was first investigated by Iwabuchi *et al*. (1991), in *B. rapa*, by Xia *et al*. (1993) in *B. napus* and by Harrison and Heslop-Harrison (1995) in all species from the "U´s triangle".

Concerning the number of chromosomes containing representatives of this repeat there are some differences between our results and that previously reported. For *B*. *rapa*, pBo1.173 hybridised to the centromeric region of eight chromosomes pairs, in accordance with the results reported by Harrison and Heslop-Harrison (1995) for the repeat isolated from *B*. *rapa* (pBcKB4, 350 bp) under low stringency conditions. Iwabuchi *et al*. (1991) detected, however, homologous of this repeat (also isolated from *B*. *rapa* - the BT11, 175 bp) at all 10 pairs of *B*. *rapa* chromosomes. According to Harrison and Heslop-Harrison (1995) this difference results from the use of lower stringency conditions in hybridisation and washing by Iwabuchi and co-workers. This may also explain the higher number of *B*. *napus* chromosomes hybridising with pBo1.173 and pBo2.94 in the present work (28-30 centromeric regions), and with the homologous sequence pxc3h2 (ca. 350 bp) reported by Xia *et al*. (1993, 24-30 centromeric regions), when compared with those 26 centromeric regions reported by Harrison and Heslop-Harrison (1995). For *B*. *carinata* 12 sites of hybridisation were found, in accordance with Harrison and Heslop-Harrison (1995). In *B*. *oleracea* we have detected 12-14 chromosomes hybridising strongly with pBo2.94 and only 12 with pBo1.173, whereas Harrison and Heslop-Harrison (1995) reported 14 hybridisation sites for pBcKB4. Armstrong *et al*. (1998), using the same clone as Harrison and Heslop-Harrison (pBcKB4) to hybridise *B*. *oleracea* var. *alboglabra* found only 12 chromosomes strongly labelled. These results show that differences between varieties, and not only the hybridisation condition used, are probably involved in the differences found among the reports. Moreover, it is possible that different members of this family of repeats are better represented in some chromosomes. Harrison and Heslop-Harrison (1995) showed that different hybridisation conditions altered the number of labelled chromosomes, indicating that the sequences present on different chromosomes have somewhat diverged.

It is interesting to notice that the highest homology found for clones pBo2.94 and pBo1.173 in the database was with a representative isolated from *B*. *juncea* (AABB) (97% similarity in a region of 224 bp and 146 bp respectively for pBo2.94 and pBo1.173). This indicates that variants from this family of repeated DNA with very similar sequences are present in both the A and C genome, since our repeats were isolated from the C genome and the sequence from *B*. *juncea* originated almost certainly from the A genome (see below).

Also interesting to notice is that Harrison and Heslop-Harrison (1995) found no hybridisation of pBcKB4 in the B genome of *Brassicacea*, either by Southern blot nor by *in situ* hybridisation, whereas Xia *et al*. (1993) have detected a low amount of pxc3h2 in *B*. *nigra* DNA by Southern and dot blot hybridisation. Compared pairwise, pBcKB4 and *canrep* (the consensus sequence including pxc3h2) share 88% sequence identity. According to Harrison and Heslop-Harrison (1995) differences in Southern blot results may lie in the methods used by the authors. We have not tested pBo2.94 and pBo1.173 in Southern blot hybridisations, but the presence of only 12 chromosomes with hybridisation sites for these sequences in *B*. *carinata* (BBCC) is in accordance with that reported by Harrison and Heslop-Harrison (1995) and indicates the absence of this sequence in the B genome. If this repeat is really present in *B*. *nigra* its copies are probable scattered in the B genome and are not recognised as FISH signals. If so, than this sequences probably amplified in the genome after the divergence of *B*. *nigra* but before the separation of *B*. *oleracea* and *B*. *rapa*, and supports that the *Brassica* A and C genomes are more closely related than both with the B genome.

As shown above centromeres of A and C chromosomes from *Brassica* contain the AT rich *Hind*III family of highly repeated DNA. However, this repeat is not the only one present in *Brassica* A and C genome centromeres, since some chromosomes with centromeric heterochromatic blocks did not possess this sequence and three pairs of chromosomes from *B*. *rapa* were shown to be enriched with a species-specific middle repetitive G-C rich (58%) DNA sequence (Iwabuchi *et al*., 1991). This G-C rich sequence shows no similarity to the members of the family that includes pBo2.94 and pBo1.173. It is possible that other sequences, so far not identified, enrich the centromeres of some *Brassica* chromosomes. However, the three repetitive sequences cloned in our pBo libraries were dispersed in the chromosomes from *Brassica* and frequently it was possible to observe that they were less represented or absent in the centromeric region.

5.4. Characterisation of newly isolated dispersed repetitive sequences of *Brassica*

pBo1.27 and pBo2.157

Sequence analyses of the DNA inserted in clones pBo1.27 (182 bp) and pBo2.157 (205 pb) showed high similarity with open reading frames (ORFs) similar to Enhancer/Supressor-mutator (*En/Spm*)-type transposon. This plant transposon superfamily was first described in the maize genome (reviewed by Kunze *et al*., 1997). *En/Spm*-like transposon sequences were found in a variety of distantly related plant species (Staginnus *et al*., 2001). Both sequences, pBo1.27 and pBo2.157, possess an intact ORF, coding for 60 and 68 aminoacids, respectively. Search in the protein database confirmed the homology of these sequences with *En/Spm*-like transposons at the protein level (data not shown). The sequences of the two clones were, however, not homologous to one another and in the Southern blot experiments they showed different patterns of hybridisation with genomic DNA from *Brassica* species. This suggests that pBo1.27 and pBo2.157 represent different groups of *En/Spm*-like elements in *Brassica*. Evolutionary diverse lineages from this transposon superfamily were also found in chickpea (*Cicer arietinum*), where *En/Spm*-like transposon sequences were investigated in detail (Staginnus *et al*., 2001). The chromosomal localisation of these sequences in *Brassica*, determined through FISH, showed a dispersed distribution of these elements in all three genomes investigated. Dispersed genomic organisation is a very common feature of plant repeats, especially those which originate from transposable elements (for references see Nouzova *et al*., 1999).

Southern blot hybridisations showed that pBo1.27 and pBo2.157 are more frequent in the C genome than in the A genome of *Brassica*. At the chromosomal level, however, this difference was less evident, probably because of the high degree of amplification of the hybridisation sites necessary for the visualisation of these sequences in FISH experiments. In relation to the evolution of these sequences, it is possible that ancestor sequences of pBo1.27 and pBo2.157 were already present in the progenitor of *B*. *oleracea* and *B*. *rapa* and the amplification or elimination of these ancestor sequences took place in a different degree in each species after their separation, resulting in different copy numbers of both elements in the A and C genomes. To our knowledge, this is the first report about the chromosomal

localisation by FISH of sequence-like DNA transposon elements (class 2 elements) in *Brassica*.

pBo1.6

Among the clones selected in this work, pBo1.6 was the most informative. Southern blot hybridisation using this sequence as probe showed the presence of only one discrete band in *B*. *rapa* in contrast to a smear of hybridisation signals with several discrete bands in *B*. *oleracea* and in *B*. *napus*. This repeat is clearly underrepresented in the A genome and the *in situ* hybridisation showed the C genome-specificity of this sequence at the chromosomal level (see Figure 17). Species-specificity only at the chromosomal level, but not in Southern blot hybridisations, was shown for a sequence, pT10, isolated from *B*. *rapa* by Itho *et al*. (1991). Clone pT10 was a chimera of two distinct sequences, one species-specific for the A genome and the other also present in the C genome. When hybridised *in situ*, however, pT10 hybridised only with chromosomes from the A genome, showing that, in spite of the presence of a part of its DNA in the C genome, at the chromosomal level this sequence was specific for *B*. *rapa* chromosomes.

All chromosomes from *B*. *oleracea* hybridised with pBo1.6, whereas in most cells of *B*. *rapa* no hybridisation signal was detected. Interestingly in *B*. *napus* frequently more than 18 chromosomes hybridised with pBo1.6, suggesting that this sequence may have either spread to A genome chromosomes and amplified after the formation of this natural hybrid or that sequences present in a small amount in the A chromosomes have amplified in the polyploid genome. In tetraploid cotton (*G*. *barbadense*, AD genome) FISH analysis have also showed that some A genome dispersed repeats appear to have spread to D genome chromosomes (Zhao *et al*., 1998b). Alternatively, the ancestral *B*. *rapa* involved in the origin of rapeseed could contain more copies from this sequence than the varieties used in the present work. However, a large variation in copy number between varieties is not to be expected for this sequence, since different varieties of *B*. *rapa* were used for Southern blot and *in situ* hybridisations and the same result was found. Noteworthily, larger chromosomes in *B*. *napus* were more strongly labelled than smaller ones. As *B*. *oleracea* has larger chromosomes than *B*. *rapa*, the chromosomes strongly labelled with pBo1.6 must be originated from the C genome. Since *B*. *rapa* and *B*. *oleracea* are very closely

related, we assume that the pBo1.6 is a relatively young sequence and that its amplification in the C genome occurred after the separation of these two species.

The C genome specificity at the chromosomal level makes pBo1.6 a useful sequence for the analyses of transgenic rapeseed. It has been suggested that transgene introgressed into the C genome of *B*. *napus* would be considerably less likely to be introgressed into wild *Brassica* populations than those present on A genome chromosomes, because *B*. *napus* (x) *B*. *oleracea* hybrids are extremely rare and transgene-carrying C genome chromosomes will be eliminated from wild *B*. *napus* (x) *B*. *rapa* backcross offspring due to the absence of homologues (for references see Snowdon *et al*., 2002). If this is right, the selection of *B*. *napus* plants with transgene inserted on C genome chromosomes can be assisted by FISH using the transgene and pBo1.6 as probes.

The molecular analyses of pBo1.6 showed high similarity between this sequence and telomeric-like DNA from several species (see Table 3). This sequence possessed 19 degenerated telomere motifs, and one perfect copy of the telomeric repeat of most plant species (see blocks in Figure 8), first isolated from *Arabidopsis* (Richards and Ausubel, 1988). Interestingly, *Sinapsis arvensis*, another species from the family *Brassicaceae*, presented also a species-specific tandem repeated DNA family (pSA) which exhibited high similarity with *Arabidopsis*- telomeric repeat. However, pSA was not detected in *B*. *oleracea* DNA in Southern blot hybridisation under high stringency conditions (Kapila *et al*.,1996a), and when compared with pBo1.6 similarity in only a very small region (55bp) is found (data not shown).

Since pBo1.6 is highly homologous to the telomeric sequence of most plants (85% similarity with a TTTAGGG tract of the same size) it should be expected that the telomeric repeats at the very ends of the *Brassica* chromosomes would hybridise with pBo1.6 under hybridisation conditions that allow more than 15% of mismatches, as the one used in this work. When genomic DNA of an organism with long telomeres are digested with six-base cutting restriction enzymes and hybridised with oligonucleotides corresponding to perfect telomeric repeats a smear of hybridisation signals is obtained, since no restriction site for enzymes is present in the tandem array of telomere repeats and the number of repeats units is variable between different chromosomes of a cell. Besides, degenerated telomeric repeats are present in several organisms (Richards, 1995, Riethman *et al*., 2004). In our Southern

hybridisation using pBo1.6 as probe, after 72 h of exposition of the membranes to X-ray films, it was possible to see a smear of DNA fragments in the high-molecular-weight range in species with the C genome. In addition, several smaller fragments of different sizes were also labelled with this probe. However, in species with only the A or B genomes (data not shown), just one discrete band, of about 2 kb in DNA digested with *Eco*RI or 10 kb in DNA digested with *Hind*III, and only a very faint smear were present. After one week of exposition, however, this smear became stronger and some additional weakly hybridising discrete bands appeared in A and B genomes (data not shown). In the FISH experiments for most cells of *B*. *rapa*, as well as in *B*. *oleracea* and *B*. *napus*, where pBo1.6 hybridised in several other sites, no signal was consistently detected at the telomeres. There are two possible explanations for the absence of hybridisation signals at the very ends of *Brassica* chromosomes in FISH with pBo1.6 under low stringency, or even if the telomeric repeat of *Arabidopsis* would be used: (1) presence of a different type of telomeric repeat or (2) short telomeres with only a small number of telomeric repeat units, that, together with technical limitations, would produce no FISH signal.

(1) It could be speculated that *Brassica* chromosomes do not possess a telomeric repeat sequence similar to the *Arabidopsis* repeat, as has been reported for monocotyledonous species of the families *Alliaceae*, *Asparagales* and *Asphodelaceae* (Fuchs *et al*., 1995, Pich and Schubert, 1998, Adams *et al*., 2000a, 2001, Weiss and Scherthan, 2002, Sýkoravá *et al*., 2003b, Weiss-Schneeweiss *et al*., 2004) and in three dicotyledonous species from the family *Solanaceae*: *Cestrum*, *Vestia* and *Sessea* (Sykorava *et al*., 2003a). However, Richards and Ausubel (1988) found that pAtT4, the telomeric repeat isolated from *A*. *thaliana*, hybridised to genomic DNA from *B*. *rapa* and *B*. *carinata* at high stringency conditions, whereas Uzunova *et al*. (1995) reported the hybridisation of this probe in *B*. *napus*, with an unscorable banding pattern. Moreover telomerase expression was found in *B*. *oleracea*, using primers complementary to *Arabidopsis* telomeric sequences in a telomere repeat amplification protocol (TRAP) (Fitzgerald *et al*., 1996). These data together show that the *Arabidopsis*-telomeric repeat (TTTAGGG) is also present in *Brassica* species, as would be expected since they are member of the same family.

(2) A small number of the telomeric repeats at the chromosome ends would produce very small hybridisation sites for a telomeric specific probe at this region and FISH

procedure may be unadequate to consistently detect such sites. The average length of telomeric repeat arrays are species and cell type specific and varies significantly among plants, ranging from 2-5 kb in *A*. *thaliana* (Richards and Ausubel, 1988), 60-160 kb in tobacco (Fajkus *et al*, 1995) and 13-223 kb in tomato (Zhong *et al*., 1998). In *Lotus japonicus* Pedrosa *et al*. (2002) suggested that the telomeric tracts were relatively short, since FISH with an *Arabidopsis*-like telomeric probe showed signals only in some chromosomes which, in addition, were very weak. Also, in *A*. *thaliana* only weak hybridisation signals were visible when the *Arabidopsis*-type telomeric sequence was used as probe (Weiss-Schneeweiss, personal communication). When hybridising pBo1.6 to *B*. *rapa* chromosome spreads no hybridisation site was detected in most of the cells. When signals were detected they were weaker in intensity than in *B*. *oleracea* and were often present as a spot in only one chromatid (see Figure 17c). Moreover, variable numbers of chromosomes (two to eight) showed labelling in the different metaphases analysed. Because of these inconsistent results the possibility that such signals represent background hybridisations can not be ruled out. However, labelling of only one chromatid with some probes has already been reported in different FISH studies and it was proposed that the size of the hybridisation site in the genome (Clark *et al*., 1989, Gustafson *et al*., 1990, Benabdelmouna *et al*., 1999), or technical artefacts (Lorite *et al*., 2002) were responsible for that observation.

The variation in the number of FISH signals for pBo1.6 and the often labelling of only one chromatid in *Brassica* may be related to the size of the telomere in the species analysed. As far as we know, the length of telomeric repeats has not been determined for *Brassica*. However, preliminary data indicate that they may be similar to those of *A*. *thaliana* (Richards, personal communication). If so, than a consistent detection of hybridisation sites will depend on the FISH procedure used. Hybridisation and detection efficiency can be influenced by several factors during FISH experiments, including mechanical damage or degradation of the homologous site, as proposed for *Brassica* pachytene chromosomes (Ziolkowiski and Sadowski, 2002), or inappropriate amplification of the small hybridisation sites, causing technical artefacts. The FISH procedure applied in the present work, where chromosomal and probe DNA were denaturated separately, may often cause mechanical damage of small sites. In addition, the small size of *Brassica* chromosomes may disturb the

chromosomal analyses, making the interpretation of hybridisation patterns more difficult than in species with large chromosomes.

At the chromosomal level pBo1.6 presented several hybridisation sites, at telomeric/subtelomeric and also at interstitial positions, in the C genome. The presence of telomeric or telomere-like repeats in non-telomeric locations has been detected in a great number of species (for reference see Richards, 1995, Lorite *et al*., 2002). In *Arabidopsis*, *Chlorella vulgaris*, *Silena latifolia*, *Nicotiana tomentosiformis*, among others, variant or degenerated units of the TTTAGGG repeats were found adjacent to cloned telomeres or in subtelomeric regions (Richards *et al*., 1992, Higashiyama *et al*., 1995, Buzek *et al*., 1997, Hokáková and Fajkus, 2000). The explanations for the occurrence of interstitial repeats are diverse. In many eukaryotes, interstitial telomeric sequences may be remnants of chromosome rearrangements that occurred during genome evolution. Interstitial telomeric sequences may also arise from a random short sequence array with fortuitous homology to telomeric repeats, which may become extend by slippage during replication (for references see Biessmann and Mason, 1994). The replication slippage is a commonly observed replication error, which occurs at repetitive sequences when the new strand mispairs with the template strand. The slipping of DNA polymerase III from the DNA template strand at the repeat region and the subsequent reattachment at a more distant site can cause the newly created DNA strand to contain an expanded section of DNA. Microsatellite polymorphisms, for example, are mainly caused by replication slippage (Dieringer and Schlotterer, 2003). Other possibilities to explain the emergence of telomere-like interstitial repeats are that telomeric repeats are attached by telomerase or recombination to extrachromosomal linear fragments, which may then integrate into the genome. Telomeric repeats can also be attached to transposable elements and are distributed by them to interstitial regions (for references see Biessmann and Mason, 1994). Which mechanism was responsible for the high amount of telomere-like sequences in the C genome remains to be investigated.

As far as we are aware, pBo1.6 is the first clone characterised as containing a sequence that hybridises preferentially to all chromosomes of the C genome. Snowdon (1997) used a RAPD fragment of 2.2 kb, that could only be amplified from species with the C genome, as FISH probe and found hybridisation on all *B*. *oleracea*

chromosomes in an interspersed pattern, whereas in *B*. *rapa* hybridisation was restrict to only a small number of chromosomes. Unfortunately, this fragment was neither characterised by Southern blot nor sequenced making impossible to establish a relationship between this sequence and pBo1.6. So far only one sequence that differentiate the A and C genomes, isolated from *B*. *rapa*, was characterised at the molecular level and localised *in situ* (Iwabuch *et al*., 1991).

The discovery of a sequence highly enriched in the C genome of *Brassica* opens the opportunity for detailed studies regarding the subsequent evolution of DNA sequences in polyploids with the C genome and may enable the reliable identification of the genomic location of transgene inserts in genetically modified oilseed rape.

6. Summary

"Isolation, molecular characterisation and chromosomal location of repetitive DNA sequences in *Brassica*"

Plant genomes contain high amounts of repetitive DNA sequences, of which several are present in many species within a taxonomic family, whereas others can exhibit species- or genome-specificity. Genome-specific repetitive sequences can for example be used as chromosome landmarkers for the discrimination of genomes in natural or artificial hybrids at a very early stage of development.

Rapeseed, *B. napus* (2n=38, AACC), is a natural hybrid derived from the closely related species *B*. *rapa* (2n=20, AA), turnip rape, and *B*. *oleracea* (2n=18, CC), cabbage. The discovery of repetitive sequences specific to the C genomes of *Brassica*, besides providing information on chromosome and genome evolution in this group, might allow the identification of some or all C genome chromosomes from *B*. *napus*. The aim of this work was to find repetitive sequences able to differentiate at the chromosomal level the highly homeologous A (*B*. *rapa*) and C (*B*. *oleracea*) genomes of *Brassica* in order to assist in the physical identification of *B*. *napus* chromosomes. For this purpose, C genome specific repetitive sequences were searched in genomic DNA libraries from *B*. *oleracea*.

Phage and plasmid libraries were screened using either *B*. *oleracea* or *B*. *rapa* total genomic DNA as probe in plaque and dot-blot hybridisations of replica filters. In the screening of more than 100,000 plaques from the phage libraries only 0.45%-2.40% of the clones exhibited strong hybridisation signals, indicating the presence of repetitive DNA in their inserts. Since only a very small amount of the clones could be identified as containing repetitive DNA, phage libraries seem not to be suitable for the identification of species-specific repetitive sequences from *B*. *oleracea*, when total genomic DNA is used as probe.

In the plasmid libraries 1164 clones were screened by dot-blot hybridisation of replica filters. In these libraries ca. 21% of the clones exhibited strong hybridisation signals with *B*. *oleracea* genomic DNA and could be identified as containing repetitive sequences. Fifteen clones, which showed stronger hybridisation signals with the C- than with the A- genome DNA, were selected for further Southern blot analyses, to verify the dot-blot results. Three clones (pBo1.6, pBo1.27 and pBo2.157) were

confirmed as more frequent in the C genome ("C genome enriched clones"). These clones, as well as two highly repetitive clones (pBo2.94 und pBo1.173), which showed very strong hybridisation signals with both the A and the C genome DNAs in the dot blot experiments, were sequenced, compared with DNA sequence databases and localised through FISH on the chromosomes of *B. oleracea*, *B. rapa* and *B. napus*. The highly repetitive clones were also tested in *B. carinata* (Ethiopian mustard). In addition, the ribosomal RNA loci of the A and C genomes were localised by FISH on *Brassica* chromosomes using different rDNA probes.

In accordance with the literature, three, six and nine chromosome pairs were identified as containing 5S and/or 45S rDNA loci in *B. oleracea*, *B. rapa* and *B. napus*, respectively. The highly repetitive clones, pBo2.94 (238 bp) und pBo1.173 (158 bp), were identified as homologous to already described centromeric sequences from *Brassica* (up to 97% sequence identity). Accordingly, these clones were localised in the centromeric regions of 12, 12-14, 16 and 28-30 chromosomes from *B. carinata*, *B. oleracea*, *B. rapa* and *B. napus*, respectively.

Sequence analyses of the C genome enriched clones showed that two of them, pBo1.27 (182 bp) and pBo2.157 (205 bp), have high similarity with *En/Spm*-transposon-like sequences (87% and 98% sequence identity, respectively). The chromosomal localisation of these sequences in *Brassica* by FISH showed a dispersed distribution of these elements in the A and the C genome. The higher frequency of these sequences in the C genome was, however, more evident in the Southern blot hybridisations than in the FISH, probably because of the high degree of amplification necessary for the visualisation of the hybridisation sites in FISH.

The third clone, pBo1.6 (203 bp), displayed in a segment varying between 116 bp and 132 bp up to 89% sequence identity with telomere-like DNA from many plant species. This sequence was localised through FISH at telomeric/subtelomeric and interstitial regions of all chromosomes from *B. oleracea*, whereas in *B. rapa* no signal was detected in most of the cells. In *B. napus* chromosomes with and without hybridisation signals were found. Frequently more than 18 chromosomes hybridised with pBo1.6 in this polyploid, suggesting that the sequence may have either spread to A genome chromosomes and amplified after the formation of this hybrid or that sequences present in a small amount in the A chromosomes have been amplified in the polyploid genome.

The discovery of a sequence highly enriched in the C genome of *Brassica* opens the opportunity for detailed studies regarding the subsequent evolution of DNA sequences in polyploid genomes. Moreover, pBo1.6 may be useful in applied genetics for the determination of the chromosomal location of transgene DNA in genetically modified oilseed rape.

7. Zusammenfassung

"Isolierung, molekulare Charakterisierung und chromosomale Lokalisierung von repetitiven DNA Sequenzen in *Brassica*"

Pflanzengenome enthalten einen großen Anteil repetitiver DNA. Verwandte Arten haben viele identische repetitive Sequenzen, aber einige von ihnen sind auch art- oder genomspezifisch. Diese genomspezifischen repetitiven Sequenzen können z.B. als Chromosomen-Marker zur Unterscheidung von Genomen in etablierten oder neu hergestellten Arthybriden in sehr frühen Entwicklungsstadien verwendet werden.

Raps, *B. napus* (2n=38, AACC), ist ein natürlicher Hybrid der sehr eng verwandten Arten Rübsen, *B. rapa* (2n=20, AA), und Kohl, *B. oleracea* (2n=18, CC). Um eine bessere Charakterisierung der Rapschromosomen zu erreichen, wurden artspezifische repetitive DNA Sequenzen von *B. oleracea* gesucht, welche das A und C Genome im Raps unterscheiden können. Diese Sequenzen sollen bei der Fluoreszenz *in situ* Hybridisierung (FISH) als genom- oder chromosomenspezifische Marker dienen.

Phagen- und Plasmid-Bibliotheken wurden jeweils durch Plaque- und Dot-Blot-Hybridisierung untersucht, wobei die Gesamt DNAs des A- und des C- Genoms als Sonde verwendet wurden. In den Phagen-Bibliotheken wurden mehr als 100.000 Plaques getestet. Nach der Hybridisierung von Replicafiltern mit *B. oleracea* oder *B. rapa* Gesamt-DNA zeigten nur 0,45%-2,40% aller Klone der Bibliotheken starke Hybridisierungssignale und konnten als repetitive Sequenzen erkannt werden. Da nur ein sehr geringer Anteil der Klone in den Bibliotheken repetitive Sequenzen aufwies, sind Phagen-Bibliotheken offensichtlich schlecht für die Suche nach repetitiven artspezifischen Sequenzen geeignet, wenn Gesamt DNA als Sonde verwendet wird.

In den Plasmid-Bibliotheken wurden 1164 Klone mit Hilfe der Dot-Blot-Hybridisierung getestet. Nach der Hybridisierung von Replicafiltern zeigten ca. 21% der Klone starke Hybridisierungsignale mit *B. oleracea* Gesamt-DNA, und wurden als repetitiven Ursprungs identifiziert. Die Hybridisierungsmuster der Dot-Blot-Filter mit beiden Sonden wurde verglichen und 15 Klone, welche stärkere Hybridisierungsintensitäten mit dem C-Genom zeigten, wurden weiter über eine Southern-Blot Hybridisierung analysiert, um die Dot-Blot Ergebnisse zu prüfen. Drei Klone (pBo1.6, pBo1.27 und

pBo2.157) konnten in dieser Untersuchung als häufiger im C-Genom vorkommend bestätigt werden („C-Genom-angereicherte Klone“). Diese Klone, sowie zwei hoch-repetitive Klone (pBo2.94 und pBo1.173), welche in der Dot-Blot-Hybridisierung gleichstarke Signale für das A- und C- Genom gezeigt haben, wurden sequenziert, mit DNA-Datenbanken verglichen und mit Hilfe der FISH in den Chromosomen von *B*. *oleracea*, *B*. *rapa* und *B*. *napus* lokalisiert. Die hoch-repetitiven Klone wurden darüber hinaus auch in *B*. *carinata* (Abessinischer Senf) untersucht. Weiterhin wurde durch FISH, mit Hilfe ribosomaler DNA (rDNA) Sonden, die ribosomalen RNA Loci in den Chromosomen der A und C Genome von *Brassica* lokalisiert.

In Übereinstimmung mit der Literatur wurden in *B*. *oleracea*, *B*. *rapa* und *B*. *napus* jeweils drei, sechs und neun Chromosomenpaare mit 5S und/oder 45S rDNA Loci gefunden. Die hoch-repetitiven Klone, pBo2.94 (238 Bp) und pBo1.173 (158 Bp), waren homolog zu schon beschriebenen Zentromer-Sequenzen von *Brassica* (bis zu 97% Sequenzidentität) und wurden durch FISH in den Zentromeren von 12, 12-14, 16 beziehungsweise 28-30 Chromosomen in *B*. *carinata*, *B*. *oleracea*, *B*. *rapa* und *B*. *napus* lokalisiert.

Von den C-Genom-angereicherten Klonen zeigten zwei, pBo1.27 (182 Bp) und pBo2.157 (206 Bp), Homologien mit dem *En-Spm*-Transposon-ähnlichen Element (jeweils 87% und 98% Sequenzidentität). Durch FISH konnte gezeigt werden, dass sie verstreut auf verschiedenen Chromosomen des A- und C-Genoms von *Brassica* liegen. Die größere Häufigkeit dieser Sequenzen im C-Genom war aber in der Southern-Blot Hybridisierung deutlicher erkennbar als in der FISH, möglicherweise weil die Signale in der FISH Methode mehrfach verstärkt werden mussten.

Der dritte Klon, pBo1.6 (203 Bp), zeigte in einer Region von 116-132 Bp bis zu 89% Sequenzidentität mit Telomer-ähnlichen Sequenzen. In der FISH hybridisierte dieser Klon mit den Telomeren/Subtelomeren und den interstitiellen Regionen von allen Chromosomen von *B*. *oleracea*, während diese Sequenz in *B*. *rapa* in den meisten Zellen keine Hybridisierungssignale zeigte. In *B*. *napus* sind beide Typen von Chromosomen (mit und ohne Hybridisierungssignale) gefunden worden. In den meisten Zellen von *B*. *napus* waren aber mehr als 18 Chromosomen markiert. Das weist darauf hin, dass pBo1.6 entweder in einige Chromosomen des A-Genoms des Polyploides übertragen und amplifiziert wurde oder dass die schon in geringer Zahl vorhandenen pBo1.6 Sequenzen des A-Genoms selbst amplifiziert wurden.

Die Entdeckung einer Sequenz, die im C-Genom von *Brassica* häufiger vorkommt als im A-Genom, eröffnet eine Gelegenheit zu detaillierten Untersuchungen der auf die interspezifische Bastardierung folgende Evolution der DNA Sequenzen in solchen Polyploiden. In der Angewandten Genetik wird diese Entdeckung möglicherweise die zuverlässige genomische Lokalisierung von transgener DNA in genetisch verändertem Raps ermöglichen.

8. References

ADAMS, S. P., LEITCH, I. J., BENNETT, M. D., LEITCH, A. R. (2000a). *Aloe* L.- a second plant family without (TTTAGGG)n telomeres. *Chromosoma* 109: 201-205.

ADAMS, S. P., LEITCH, I. J., BENNETT, M. D., CHASE, M. W., LEITCH, A. R. (2000b). Ribosomal DNA evolution and phylogeny in *Aloe* (*Asphodelaceae*). *Am*. *J*. *Bot*. 87: 1578-1583.

ADAMS, S. P., HARTMAN, T. P. V., LIM, K. Y., CHASE, M. W., BENNETT, M. D., LEITCH, I. J., LEITCH, A. R. (2001). Loss and recovery of *Arabidopsis*-type telomere repeat sequences 5'-(TTTAGGG)n-3' in the evolution of a major radiation of flowering plants. *Proc*. *R*. *Soc*. *London Ser*. *B* 268: 1541-1546.

AHUJA, M. R. (2001). Recent advances in molecular genetics of forest trees. *Euphytica* 121: 173-195.

ANAMTHAWAT-JÓNSSON, K., HESLOP-HARRISON, J. S. (1992). Species specific DNA sequences in the *Triticeae*. *Hereditas* 116: 49-54.

ANANIEV, E. V., VALES, M. I., PHILLIPS, R. L., RINES, H. W. (2002). Isolation of A/D and C genome specific dispersed and clustered repetitive DNA sequences from *Avena sativa*. *Genome* 45: 431-441.

ARMSTRONG, K. C., KELLER, W. A. (1981). Chromosome pairing in haploids of *Brassica campestris*. *Theor*. *Appl*. *Genet*. 59: 49-52.

ARMSTRONG, K.C., KELLER, W.A. (1982). Chromosome pairing in haploids of *Brassica oleracea*. *Can*. *J*. *Gen*. *Cytol*. 24: 735-739.

ARMSTRONG, S. J., FRANSZ, P., MARSHALL, D. F., JONES, G. H. (1998). Physical mapping of DNA repetitive sequences to mitotic and meiotic chromosomes of *Brassica oleracea* var. *alboglabra* by fluorescence *in situ* hybridisation. *Heredity* 81: 666-673.

ARUMUGANATHAN, K., EARLE, E. D. (1991). Nuclear DNA content of some important plant species. *Plant Mol*. *Biol*. *Rep*. 9: 208-219.

AVRAMOVA, Z. V. (2002). Heterochromatin in animals and plants. Similarities and differences. *Plant Physiol*. 129: 40-49.

BECKER, H. C., LÖPTIEN, H., RÖBBELEN, G. (1999). Breeding: an overview. In: GÓMEZ-CAMPO, C. (Ed.) Biology of *Brassica* Coenospecies. Developments in Plant Genetics and Breeding. pp. 413-460. Elsevier Science, Amsterdam, Netherlands.

BEDBROOK, J. R., O'DELL, M., FLAVELL, R. B. (1980). Amplification of rearranged repeated DNA sequences in cereal plants. *Nature* 288: 133-137.

BELLIN, U., DOS SANTOS, K. G. B. (2002). Beurteilung der Rekombinationsfähigkeit von *Brassica*-Arthybriden durch Unterscheidung der Genome mit Hilfe von GISH. *Vortr*. *Pflanzenzüchtung*. 54: 373-376.

BENABDELMOUNA, A., PELTIER, D., HUMBERT, C., ABIRACHED-DARMENCY, M. (1999). Southern and fluorescent *in situ* hybridization detect three RAPD-generated PCR products useful as introgression markers in *Petunia*. *Theor*. *Appl*. *Genet*. 98: 10-17.

BENNETT, M. D., LEITCH, I. J. (1997). Nuclear DNA amounts in angiosperms-583 new estimates. *Ann*. *Bot*. *(London)* 80, 169-196

BENNETT, M. D, BHANDOL, P., LEITCH, I. J. (2000). Nuclear DNA amounts in angiosperms and their modern uses - 807 new estimates. *Ann*. *Bot*. *(London)* 86: 859-909.

BENSLIMANE, A. A., DRON, M., HARTMANN, C., RODE, A. (1986). Small tandemly repeated DNA sequences of higher plants likely originate from a tRNA gene ancestor. *Nucleic Acids Res*. 14: 8111-8119.

BERIDZE, T. (1975). DNA nuclear satellites of the genus *Brassica*: Variation between species. *Biochim*. *Biophys*. *Acta* 395: 274-279.

BERNATZKY, R., TANKSLEY, S. D. (1986). Toward a saturated linkage map in tomato based on isozymes and random cDNA sequences. *Genetics* 111: 887-898.

BHATIA, S., NEGI, M. S., LAKSHMIKUMARAN, M. (1996). Structural analysis of the rDNA intergenic spacer of *Brassica nigra*: evolutionary divergence of the spacers of the three diploid *Brassica* species. *J*. *Mol*. *Evol*. 43: 460-468.

BIESSMANN, H., MASON, J. M. (1994). Telomeric repeat sequences. *Chromosoma* 103: 154-161.

BIESSMANN, H., MASON, J. M. (1997). Telomere maintenance without telomerase. *Chromosoma* 106: 63-69.

BIRNBOIM, H. C., DOLY, J. (1979). A rapid alkaline extraction procedure for screening recombinat plasmid DNA. *Nuclei Acids Res*. 7: 1513-1523.

BOURNIVAL, B., OBANNI, M., ABAD, A., OHM, H., MACKENZIE, S. (1994). Isolation of a new species-specific repetitive sequence from *Thinopyrum elongatum* and its use in the studies of alien translocations. *Genome* 37: 97-104.

BOWEN, N. J. , JORDAN, I. K. (2002). Transposable elements and the evolution of eukaryotic complexity. *Curr*. *Issues Mol*. *Biol*. 4: 65-76.

BROWN, G. B., AMARASINGHE, V., KISS, G., CARLSON, J. E. (1993). Preliminary karyotype and chromosomal localization of rDNA loci for white spruce by fluorescence *in situ* hybridization. *Genome* 36: 310-316.

BROWN, S., STEPHENS, J. L., LAPITAN, N. L. V., KNUDSON, D. L. (1999). FISH landmarks for barley chromosomes (*Hordeum vulgare* L.). *Genome* 42: 274-281.

BUSCH, W., MARTIN, R., HERRMANN, R. G., HOHMANN, U. (1995). Repeated DNA sequences isolated by microdissection: I. Karyotyping of barley (*Hordeum vulgare* L.). *Genome* 38: 1082-1090.

BUZEK, J., KOUTNIKOVA, H., HOUBEN, A., RIHA, K., JANOUSEK, B., SIROKY, J., GRANT, S., VYSKOT, B. (1997). Isolation and characterization of X chromosome-derived DNA sequences from a dioecious plant *Melandrium album*. *Chromosome Res*. 5: 57-65.

CALDERINI, O., PUPILLI, F., PAOLOCCI, F., ARCIONI, S. (1997). A repetitive and species-specific sequence as a tool for detecting the genome contribution in somatic hybrids of the genus *Medicago*. *Theor*. *Appl*. *Genet*. 95: 734-740.

CAO, M., SLEPER, D. A. (2001). Use of genome-specific repetitive DNA sequences to monitor chromatin introgression from *Festuca mairei* into *Lolium perenn*e. *Theor*. *Appl*. *Genet*. 103: 248-253.

CHEN, C. M., WANG, C. T., WANG, C. J., HO, C. H., KAO, Y. Y., CHEN, C. C. (1997). Two tandemly repeated telomere-associated sequences in *Nicotiana plumbaginifolia*. *Chromosome Res.* 5: 561-568.

CHEN, C. C., CHEN, C. M., HSU, F. C., WANG, C. J., YANG, J. T., KOA, Y. Y. (2000a). The pachytene chromosomes of maize as revealed by fluorescence *in situ* hybridization with repetitive DNA sequences. *Theor*. *Appl*. *Genet*. 101: 30-36.

CHEN, C., YAN, H., ZHAI, W., ZHU, L., SUN, J. (2000b). Identification and chromosomal location of a new tandemly repeated DNA in maize. *Genome* 43: 181-184.

CHENG, B. F., HENEEN, W. K., CHEN, B. Y. (1994). Meiotic studies on a *Brassica campestris-alboglabra* monosomic addition line and derived *B. campestris* primary trisomics. *Genome* 37: 584-589.

CHEUNG, W. Y., CHAMPAGNE, G., HUBERT, N., LANDRY, B. S. (1997). Comparison of the genetic maps of *Brassica napus* and *Brassica oleracea*. *Theor*. *Appl*. *Genet*. 94: 569-582.

CHENG, B. F., HENEEN, W. K., CHEN, B. Y. (1995). Mitotic kyryotypes of *Brassica campestris* and *Brassica alboglabra* and identification of the *B. alboglabra* chromosome in an addition line. *Genome* 38: 313-319.

CHUNG-MONG, C., CHI-TING, W., CHUNG-JU, W., CHIA-HSING, H., YEN-YU, K., CHI-CHANG, C. (1997). Two tandemly repeated telomere-associated sequences in *Nicotiana plumbaginifolia*. *Chromosome Res*. 5: 561-568.

CLARK, M., KARP, A., ARCHER, S. (1989). Physical mapping of the B-hordein loci on barley chromosome 5 by *in situ* hybridization. *Genome* 32: 925-929.

CUADRADO, A., CEOLONI, C., JOUVE, N. (1995). Variation in highly repetitive DNA composition of heterochromatin in rye studied by fuorescence *in situ* hybridization. *Genome* 38: 1061-1069.

DE JEU, M. J., LASSCHUIT, J., KUIPERS, A. G. J., KAMSTRA S. A., VISSER, R. G. F. (1997). Characterization and localization of repetitive DNA sequences in the ornamental *Alstroemeria aurea* Graham. *Theor*. *Appl*. *Genet*. 94: 982-990.

DE MELO N. F., GUERRA M. (2003). Variability of the 5S and 45S rDNA sites in *Passiflora* L. species with distinct base chromosome numbers. *Ann*. *Bot*. *(London)* 92: 309-316.

DECHYEVA, D., GINDULLIS, F., SCHMIDT, T. (2003). Divergence of satellite DNA and interspersion of dispersed repeats in the genome of the wild beet *Beta procumbens*. *Chromosome Res*. 11: 3-21.

DIERINGER, D., SCHLOTTERER, C. (2003). Two distinct modes of microsatellite mutation processes: Evidence from the complete genomic sequences of nine species. *Genome Res*. 13: 2242-2251.

DONG, J., KHARB, P., CERVERA, M., HALL, T. C. (2001). The use of FISH in chromosomal localization of transgenes in rice. *Methods in Cell Sci*. 23: 105-113.

FAHLESON, J., LAGERCRANTZ, U., MOURAS, A., GLIMELIUS, K. (1997). Characterisation of somatic hybrids between *Brassica napus* and *Eruca sativa* using species-specific repetitive sequences and genomic *in situ* hybridisation. *Plant Sci*. 123: 133-142.

FAJKUS, J., KOVARIK, A., KRALOVICS, R., BEZDEK, M. (1995). Organization of telomeric and subtelomeric chromatin in the higher plant *Nicotiana tabacu*m. *Mol*. *Gen*. *Genet*. 247: 633-638.

FAO. http://faostat.fao.org/.

FITZGERALD, M. A. S., MCKNIGHT, T. D., SHIPPEN D. E. (1996). Characterization and developmental patterns of telomerase expression in plants. *Proc*. *Natl*. *Acad*. *Sci*. *U*.*S*.*A*. 93: 14422-14427.

FLAVELL, R. (1980). Molecular characterization and organization of plant chromosomal DNA sequences. *Ann*. *Rev*. *Plant Physiol*. 31: 569-596.

FLAVELL, R. B., BENNETT, M. D., SMITH, J. B., SMITH, D. B. (1974). Genome size and the proportion of repeated nucleotide sequence DNA in plants. *Biochem*. *Genet*. 12: 257-269.

FREDIANI, M., GELATI, M. T., MAGGINI, F., GALASSO, I., MINELLI, S., CECCARELLI, M., CIONINI, P. G. (1999). A family of dispersed repeats in the genome of *Vicia faba*: structure, chromosomal organization, redundancy modulation, and evolution. *Chromosoma* 108: 317-324.

FRANSZ, P. F., STAM, M., MONTIJN, B., TEN HOOPEN, R., WIEGANT, J., KOOTER, J. M., OUD, O., NANNINGA, N. (1996). Detection of single- copy genes and chromosome rearrangements in *Petunia hybrida* by fluorescence *in situ* hybridization. *Plant J*. 9: 767-774.

FUCHS, J., BRANDES, A., SCHUBERT, I. (1995). Telomere sequence localization and karyotype evolution in higher plants. *Plant Syst*. *Evol*. 196: 227-241.

FUKUI, K., NAKAYAMA, S., OHMIDO, N., YOSHIAKI, H., YAMABE, M. (1998). Quantitative karyotyping of three diploid *Brassica* species by imaging methods and localisation of 45S rDNA loci on the identified chromosomes. *Theor*. *Appl*. *Genet*. 96: 325-330.

GALASSO, I., SCHMIDT, T., PIGNONE, D., HESLOP-HARRISON, J. S. (1995). The molecular cytogenetics of *Vigna unguiculata* (L.) Walp: the physical organization and characterization of 18S-5.8S-25S rRNA genes, 5S rRNA genes, telomere-like sequences, and a family of centromeric repetitive DNA sequences. *Theor*. *Appl*. *Genet*. 91: 928-935.

GALASSO, I., SCHMIDT, T., PIGNONE, D. (2001). Identification of *Lens culinaris* ssp. *culinaris* chromosomes by physical mapping of repetitive DNA sequences. *Chromosome Res*. 9: 199-209.

GOUBELY, C., ARNAUD, P., TATOUT, C., HESLOP-HARRISON. J.S., DERAGON, J.-M. (1999). S1 SINE retroelements are methylated at symmetrical and non-symmetrical positions in *Brassica napus*: identification of a preferred target site for asymmetrical methylation. *Plant Mol*. *Biol*. 39: 243-255.

GREWAL, S. I. S., ELGIN, S. C. R. (2002). Heterochromatin: New possibilities for the inheritance of structure. *Curr*. *Opin*. *Genet*. *Dev*. 12: 178-187.

GUERRA, M., SANTOS, K. G. B., SILVA, A. E. B., EHRENDORFER, F. (2000). Heterochromatin banding patterns in *Rutaceae-Aurantioideae* - a case of parallel chromosomal evolution. *Am. J. Bot*. 87: 735-747.

GUPTA, V., JAGANNATHAN, V., LAKSHMIKUMARAN, M. S. (1990). A novel AT-rich tandem repeat of *Brassica nigra*. *Plant Science* 68: 223-229.

GUPTA, V., LAKSHMISITA, G., SHAILA, M. S., JAGANNATHAN, V., LAKSHMIKUMARAN, M. S. (1992). Characterization of species-specific repeated DNA sequences from *Brassica nigra*. *Theor*. *Appl*. *Genet*. 90: 397-402.

GUSTAFSON, J. P., BUTLER, E., MCINTYRE, C. L. (1990). Physical mapping of a low-copy DNA sequence in rye (*Secale cereale* L.). *Proc*. *Natl*. *Acad*. *Sci*. *U*.*S*.*A*. 87: 1899-1902.

HARBINDER, S., LAKSHMIKUMARAN, M. (1990). A repetitive sequence from *Diplotaxis erucoides* is highly homologous to that of *Brassica campestris* and *B*. *oleracea*. *Plant Mol*. *Biol*. 15 : 155-156.

HARRISON, G. E., HESLOP-HARRISON, J. S. (1995). Centromeric repetitive DNA in the genus *Brassica*. *Theor*. *Appl*. *Genet*. 90, 157-165.

HASTEROK, R., MALUSZYNSKA, J. (2000a). Cytogenetic analysis of diploid *Brassica* species. *Acta Biol*. *Cracov*. *Ser*. *Bot*. 42: 145-163.

HASTEROK, R., MALUSZYNSKA, J. (2000b). Nucleolar dominance does not occur in root tip cells of allotetraploid *Brassica* species. *Genome* 43: 574-579.

HASTEROK, R., JENKINS, G., LANGDON, T., JONES, R. N., MALUSZYNSKA, J. (2001). Ribosomal DNA is an effective marker of *Brassica* chromosomes. *Theor*. *Appl*. *Genet*. 103: 486-490.

HESLOP-HARRISON, J. S. (2000). Comparative genome organization in plants: from sequence and markers to chromatin and chromosomes. *Plant Cell* 12: 617-635.

HIGASHIYAMA, T., MAKI, S., YAMADA, T. (1995). Molecular organization of *Chlorella vulgaris* chromosome I: presence of telomeric repeats that are conserved in higher plants. *Mol*. *Gen*. *Genet*. 246: 29-36.

HIZUME, M., SHIBATA, F., MARUYAMA, Y., KONDO, T. (2001). Cloning of DNA sequences localized on proximal fluorescent chromosome bands by microdissection in *Pinus densiflora*. *Chromosoma* 110: 345-351.

HORÁKOVÁ, M., FAJKUS, J. (2000). TAS49- a dispersed repetitive sequence isolated from subtelomeric regions of *Nicotiana tomentosiformis* chromosomes. *Genome* 43: 273-84.

HOUBEN, A., WANNER, G., HANSON, L., VERLIN, D., LEACH, C. R., TIMMIS, J. N. (2000). Cloning and characterisation of polymorphic heterochromatic segments of *Brachycome dichromosomatica*. *Chromosoma* 109: 433-433.

HOWELL, E. C., BARKER, G. C., JONES, G. H., KEARSEY, M. J., KING, G. J., KOP, E. P., RYDER, C. D., TEAKLE, G. R., VICENTE, J. G., ARMSTRONG, S. J. (2002). Integration of the cytogenetic and genetic linkage maps of *Brassica oleracea*. *Genetics* 161: 1225-1234.

ITOH, K., IWABUCHI, M., SHIMAMOTO, K. (1991). *In situ* hybridization with species-specific DNA probes gives evidence for asymmetric nature of *Brassica* hybrids obtained by X-ray fusion. *Theor*. *Appl*. *Genet*. 81: 356-362.

IWABUCHI, M., ITOH, K., SHIMAMOTO, K. (1991). Molecular and cytological characterization of repetitive DNA sequences in *Brassica*. *Theor*. *Appl*. *Genet*. 81: 349-355.

IWANO, M., SAKAMOTO, K., SUZUKI, G., WATANABE, M., TAKAYAMA, S., FUKUI, K., HINATA, K., ISOGAI, A. (1998). Visualization of a self-incompatibility gene in *Brassica campestris* L. by multicolor FISH. *Theor*. *Appl*. *Genet*. 96: 751-757.

JACKSON, S. A., CHENG, Z., WANG, M.-L., GOODMAN, H. M., JIANG, J. (2000). Comparative FISH mapping of a 431-kb *Arabidopsis thaliana* BAC contig reveals the role of chromosomal duplications in the expansion of the *Brassica rapa* genome. *Genetics* 156: 833-838.

JAMILENA, M., REJON, C. R., REJON, M. R. (1993). Repetitive DNA sequence families in *Crepis capillaris*. *Chromosoma* 102: 272-278.

JAMILENA, M., REJÓN, C. R., REJÓN, M. R. (1994). A molecular analysis of the origin of the *Crepis capillaris* B chromosome. *J*. *of Cell Sci*. 107: 703-708.

JAMILENA, M., GARRIDO-RAMOS, M., REJÓN, M. R., REJÓN, C. R., PARKER, J. S. (1995). Chracterisation of repeated sequences from microdissected B chromosomes of *Crepis capillaris*. *Chromosoma* 104: 113-120.

JASINSKA, A., KRZYZOSIAK, W. J. (2004). Repetitive sequences that shape the human transcriptome. *FEBS Lett*. 567: 136-141.

JENKINS, G., HEAD, J., FORSTER, J. W. (2000). Probing meiosis in hybrids of *Lolium* (*Poaceae*) with a discriminatory repetitive genomic sequence. *Chromosoma* 109: 280-286.

JIANG, J., GILL, B. S. (1994). New 18S·26S ribosomal RNA gene loci: Chromosomal landmarks for the evolution of polyploid wheats. *Chromosoma* 103: 179-185.

JIN, W. W., LI, Z. Y., FANG, Q., ALTOSAAR, I., LIU, L. H., SONG, Y. C. (2002). Fluorescence *in situ* hybridization analysis of alien genes in *Agrobacterium*-mediated *Cry1A(b)*-transformed Rice. *Ann*. *Bot*. *(London)* 90: 31-36.

KAMISUGI, Y., NAKAYAMA, S., O'NEIL, C. M., MATHIAS, R. J., TRICK, M., FUKUI, K. (1998). Visualization of the *Brassica* self-incompatibility S-locus on identified oilseed rape chromosomes. *Plant Mol*. *Biol*. 38: 1081-1087.

KAMSTRA, S. A., KUIPERS, A. G. J., DE JEU, M. J., RAMANNA, M. S., JACOBSEN, E. (1997). Physical localisation of repetitive DNA sequences in *Alstroemeria*: karyotyping of two species with species-specific and ribosomal DNA. *Genome* 40: 652-658.

KAPILA, R., DAS, S., SRIVASTAVA, P. S., LAKSHMIKUMARAN, M. (1996a). A novel species-specific tandem repeat DNA family from *Sinapis arvensis*: Detection of telomere-like sequences. *Genome* 39: 758-766.

KAPILA, R., NEGI, M. S., THISEY, P., DELSENY, M., SRIVASTAVA, P. S., LAKSHMIKUMARAN, M. (1996b). A new family of dispersed repeats from *Brassica nigra*: characterization and localization *Theor*. *Appl*. *Genet*. 93 : 1123-1129.

KAPUSCINSKI, J. (1995). DAPI: a DNA-specific fluorescent probe. *Biotech*. *Histochem*. 70. 220-233.

KIDWELL, M. G. (2002). Transposable elements and the evolution of genome size in eukaryotes. *Genetica* 115: 49-63.

KITAMURA, S., INOUE, M., OHMIDO, N., FUKUI, K. (2000). Quantitative chromosome maps and rDNA localization in the T subgenome of *Nicotiana tabacum* L. and its putative progenitors. *Theor*. *Appl*. *Genet*. 101: 1180-1188.

KO, J.-M., DO, G.-S., SUH, D.-Y., SEO, B.-B., SHIN, D.-C., MOON, H.-P. (2002). Identification and chromosomal organization of two rye genome-specific RAPD products useful as introgression markers in wheat. *Genome* 45: 157-164.

KOHLI, A., TWYMAN, R. M., ABRANCHES, R., WEGEL, E., STOGER, E., CHRISTOU, P. (2003). Transgene integration, organization and interaction in plants. *Plant Mol*. *Biol*. 53: 247-258.

KOLCHINSKY, A., GRESSHOFF, P. M. (1995). A major satellite DNA of soybean is a 92- base pairs tandem repeat. *Theor*. *Appl*. *Genet*. 90: 621-626.

KUBIS, S., HESLOP-HARRISON J. S., SCHMIDT, T. (1997). A family of differentially amplified repetitive DNA sequences in the genus *Beta* reveals genetic variation in *Beta vulgaris* subspecies and cultivars. *J. Mol. Evol*. 44: 310-320.

KUBIS, S., SCHMIDT, T., HESLOP-HARRISON, J. S. (1998). Repetitive DNA as a major component of plant genomes. *Ann*. *Bot*. *(London)* 82: 45-55.

KUBIS, S. E., CASTILHO, A. M. M. F., VERSHININ, A. V., HESLOP-HARRISON, J. S. (2003). Retroelements, transposons and methylation status in the genome of oil palm (*Elaeis guineensis*) and the relationship to somaclonal variation. *Plant Mol*. *Biol*. 52: 69-79.

KUMAR, A., BENNETZEN, J. L. (1999). Plant retrotransposons. *Annu*. *Rev*. *Genet*. 33: 479-532.

KUNZE, R., SAEDLER, H., LÖNNIG, W. (1997). Plant transposable elements. *Adv*. *Bot*. *Res*. 27: 332-470.

LAGERCRANTZ, U. (1998). Comparative mapping between *Arabidopsis thaliana* and *Brassica nigra* indicates that *Brassica* genomes have evolved through extensive genome replication accompanied by chromosome fusions and frequent rearrangements. *Genetics* 150: 1217-1228.

LAGERCRANTZ, U., LYDIATE, D. J. (1996). Comparative genome mapping in *Brassica*. *Genetics* 144: 1903-1910.

LAKSHMIKUMARAN, M., RANADE, S. A. (1990). Isolation and characterization of a highly repetitive DNA of *Brassica campestris*. *Plant Mol*. *Biol*. 14: 447-448.

LAPITAN, N. L. V. (1992). Organization and evolution of higher plant nuclear genomes. *Genome* 35: 171-181.

LEE, S. H., DO, G. S., SEO, B. B. (1999). Chromosomal localization of 5S rRNA gene loci and the implications for relationships within the *Allium* complex. *Chromosome Res*. 7: 89-93.

LEITCH, A. R., SCHWARZACHER, T., WANG, M. L. (1993). Molecular cytogenetic analysis of repeated sequences in a long wheat suspension culture. *Plant Cell Tissue Organ Cult*. 33: 287-296.

LEVAN, A., FREDGA, K., SANDBERG, A. A. (1964). Nomenclature for centromeric position on chromosomes. *Hereditas* 52: 201-220.

LI, B., OESTREICH, S., DE LANGE, T. (2000). Identification of human Rap1: implications for telomere evolution. *Cell* 101: 471-83.

LI, M., QIAN, W., MENG, J., LI, Z. (2004). Construction of novel *Brassica napus* genotypes through chromosomal substitution and elimination using interploid species hybridization. *Chromosome Res*. 12: 417-426.

LI, W. L., CHEN, P. D., QI, L. L., LIU, D. J. (1995). Isolation, characterization and application of a species-specific repeated sequence from *Haynaldia villosa*. *Theor*. *Appl*. *Genet*. 90: 526-533.

LIM, K.-B., WENNEKES, J., DE JONG, J. H., JACOBSEN, E., VAN TUYL, J. M. (2001). Karyotype analysis of *Lilium longiflorum* and *Lilium rubellum* by chromosome banding and fluorescence *in situ* hybridisation. *Genome* 44: 911-918.

LIM, K. Y., MATYASEK, R., LICHTENSTEIN, C. P., LEITCH, A. R. (2000). Molecular cytogenetic analyses and phylogeny of the *Nicotiana* section *Tomentosae*. *Chromosoma* 109: 245-258.

LORITE, P., CARRILLO, J. A., PALOMEQUE, T. (2002). Conservation of (TTAGG)n telomeric sequences among ants (*Hymenoptera*, *Formicidae*). *J*. *Hered*. 93: 282-285.

LUBARETZ, O., FUCHS, J., AHNE, R., MEISTER, A. (1996). Karyotyping of three *Pinaceae* species via fluorescent *in situ* hybridization and computer-aided chromosome analysis. *Theor*. *Appl*. *Genet*. 92: 411-416.

LUKENS, L., ZOU, F., LYDIATE, D., PARKIN, I., OSBORN, T. (2003). Comparison of a *Brassica oleracea* genetic map with the genome of *Arabidopsis thaliana*. *Genetics* 164: 359-372.

MACAS, J., POZARKOVA, D., NAVRATILOVA, A., NOUZOVA, M., NEUMANN, P. (2000). Two new families of tandem repeats isolated from genus *Vicia* using genomic self-priming PCR. *Molecular Gen*. *Genet*. 263: 741-751.

MACAS, J. Í, MÉSZÁROS, T., NOUZOVÁ, M. (2002). PlantSat: a specialized database for plant satellite repeats. *Bioinformatics* 18: 28-35.

MANDRIOLI, M., BIZZARO, D., GIUSTI, M., MANICARDI, G. C., BIANCHI, U. (1999). The role of rDNA genes in X chromosome association in the aphid *Acyrthosiphon pisum*. *Genome* 42: 318-386.

MALUSZYNSKA, J., HESLOP-HARRISON, P. (1993). Physical mapping of rDNA loci in *Brassica* species. *Genome* 36: 774-781.

MCKNIGHT, T. D., RIHA, K., SHIPPEN, D. E. (2002). Telomeres, telomerase, and stability of the plant genome. *Plant Mol*. *Biol*. 48: 331-337.

MEINKOTH, J., WAHL, G. (1984). Hybridization of nucleic acids immobilized on solid supports. *Anal*. *Biochem*. 138: 267-284.

MIRANDA, M., IKEDA, F., ENDO, T., MORIGUCHI, T., OMURA, M. (1997). Comparative analysis on the distribution of heterochromatin in *Citrus*, *Poncirus* and *Fortunella* chromosomes. *Chromosome Res*. 5: 86-92.

MOLNAR, S. J., GUPTA, P. K., FEDAK, G., WHEATCROFT, R. (1989). Ribosomal DNA repeat unit polymorphism in 25 *Hordeum* species. *Theor*. *Appl*. *Genet*. 78: 387-392.

MOSCONE, E. A., MATZKE, M. A., MATZKE, A. J. M. (1996). The use of combined FISH/GISH in conjunction with DAPI counterstaining to identify chromosomes containing transgene inserts in amphidiploid tobacco. *Chromosoma* 105: 231-236.

MOSCONE, E. A., KLEIN, F., LAMBROU, M., FUCHS, J., SCHWEIZER, D. (1999). Quantitative karyotyping and dual-color FISH mapping of 5S and 18S-25S rDNA probes in the cultivated *Phaseolus* species (*Leguminosae*). *Genome* 42: 1224-1233.

MURATA, M., HESLOP-HARRISON, J. S., MOTOYOSHI, F. (1997). Physical mapping of the 5S ribosomal RNA genes in *Arabidopsis thaliana* by multi-color fluorescence *in situ* hybridization with cosmid clones. *Plant J*. 12: 31-37.

NARAYAN, R. K. J. (1998). The role of genomic constraints upon evolutionary changes in genome size and chromosome organization. *Ann*. *Bot*. *(London)* 82: 57-66.

NAVRÁTILOVÁ, A., NEUMANN, P., MACAS, J. (2003). Karyotype analysis of four *Vicia* species using *in situ* hybridization with repetitive sequences. *Ann*. *Bot*. *(London)* 91: 921-926.

NEUMANN, P., NOUZOVA, M., MACAS, J. (2001). Molecular and cytogenetic analysis of repetitive DNA in pea (*Pisum sativum* L.). *Genome* 44: 716-728.

NOUZOVA, M., KUBALAKOVA, M., DOLEZELOVA, M., KOBLIZKOVA, A., NEUMANN, P., DOLEZEL, J., MACAS, J. (1999). Cloning and characterization of new repetitive sequences in field bean (*Vicia faba* L.). *Ann*. *Bot*. *(London)* 83: 535-541.

OLIN-FATIH, M., HENEEN, W. K. (1992). C-banded karyotypes of *Brassica campestris*, *B*. *oleracea* and *B*. *napus*. *Genome* 35: 583-589.

OUYANG, S., BUELL, C. R. (2004). The TIGR Plant Repeat Databases: a collective resource for the identification of repetitive sequences in plants. *Nucleic Acids Res*. 32: 360-363.

PALMER, J. D., SHIELDS, C. R., COHEN, D. B., ORTON, T. J. (1983). Chloroplast DNA evolution and the origin of amphidiploid *Brassica* species. *Theor*. *Appl*. *Genet*. 65: 181-189.

PARDUE, M.-L., DEBARYSHE, P. G. (1999). Telomeres and telomerase: more than the end of the line. *Chromosoma* 108: 73-82.

PARKIN, I. A. G., SHARPE, A. G., KEITH, D. J., LYDIATE, D. J. (1995). Identification of the A and C genomes of amphidiploid *Brassica napus* (oilseed rape). *Genome* 38: 1122-1131.

PEHU, E., THOMAS, M., POUTALA, T., KARP, A., JONES, M. G. K. (1990). Species-specific sequences in the genus *Solanum*: identification, characterization, and application to study somatic hybrids of *S*. *brevidens* and *S*. *tuberosum*. *Theor*. *Appl*. *Genet*. 80: 693-698.

PEDROSA, A., JANTSCH, M. F., MOSCONE, E. A., AMBROS, P. F., SCHWEIZER, D. (2001). Characterisation of pericentromeric and sticky intercalary heterochromatin in *Ornithogalum longibracteatum* (Hyacinthaceae). *Chromosoma* 110: 203-213.

PEDROSA, A., SANDAL, N., STOUGAARD, J., SCHWEIZER, D., BACHMAIR, A. (2002). Chromosomal map of the model legume *Lotus japonicus*. *Genetics* 161: 1661-1672.

PEDROSA, A., VALLEJOS, C., BACHMAIR, A., SCHWEIZER, D. (2003). Integration of common bean (*Phaseolus vulgaris* L.) linkage and chromosomal maps. *Theor*. *Appl*. *Genet*. 106: 205-212.

PICH, U., SCHUBERT, I. (1998). Terminal heterochromatin and alternative telomeric sequences in *Allium cepa*. *Chromosome Res*. 6: 315-321.

PLANTSAT. (2004). http://w3lamc.umbr.cas.cz/PlantSat/.

PRAKASH, S., HINATA, K. (1980). Taxonomy, cytogenetics and origin of crop *Brassicas* , a review. *Opera*. *Bot*. 55: 3-57.

PRAKASH, S., TSUNODA, S. (1980). Cytogenetics of *Brassica*. In: Swaminathan, M. S., Gupta, P. K., Sinha, U. (Eds.) Cytogenetics of crop plants. pp 481-513. Macmillan India Limited, New Delhi, India.

RAINA, S. N., OGIHARA, Y. (1995). Ribosomal DNA repeat unit polymorphism in 49 *Vicia* species. *Theor*. *Appl*. *Genet*. 90: 477-486.

RAINA, S. N., MUKAI, Y., KAWAGUCHI, K., GOEL, S., JAIN, A. (2001). Physical mapping of 18S-5.8S-26S and 5S ribosomal RNA gene families in three important vetches (*Vicia* species) and their allied taxa constituting three species complexes. *Theor*. *Appl*. *Genet*. 103: 839-845.

RAN, Y., HAMMETT, K. R. W., MURRAY, B. G. (2001). Phylogenetic analysis and karyotype evolution in the genus *Clivia* (*Amaryllidaceae*). *Ann*. *Bot*. *(London)* 87: 823-830.

REDDY, A. S., SRIVASTAVA, V., GUHA-MUKHERJEE, S. (1989). A tandemly repeated DNA sequence from *Brassica juncea*. *Nucleic Acids Res*. 17: 5849.

RICHARDS, E. J. (1995). Plant Telomeres. In: GREIDER, C., BLACKBURN, E.H. (Eds.) Telomeres. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

RICHARDS, E. J., AUSUBEL, F. M. (1988). Isolation of a higher eukaryotic telomere from *Arabidopsis thaliana*. *Cell* 53: 127-136.

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., CHAO, S., VONGS, A., YANG, J. (1992). Characterization of *Arabidopsis thaliana* telomeres isolated in yeast. *Nucleic Acids Res.* 20: 4039-4046.

RIETHMAN, H., AMBROSINI, A., CASTANEDA, C., FINKLESTEIN, J., HU, X.-L., MUDUNURI, U., PAUL, S., WEI, J. (2004). Mapping and Initial Analysis of Human Subtelomeric Sequence Assemblies. *Genome Res*. 14: 18-28.

RIHA, K., SHIPPEN, D. E. (2003). Telomere structure, function, and maintenance in *Arabidopsis*. *Chromosome Res*. 11: 263-275.

RÖBBELEN, G. 1960. Beitrage zur Analyse des *Brassica* Genome. *Chromosoma* 11: 205-228.

ROGERS, S. O., BENDICH, A. J. (1987). Ribosomal RNA genes in plants: variability in copy number and in the intergenic spacer. *Plant Mol*. *Biol*. 9: 509-520.

ROGERS, S. O., BENDICH, A. J. (1988). Extraction of DNA from plant tissues. *Plant Mol*. *Biol*. *Manual* A6: 1-10.

ROKKA, V.-M., CLARK, M. S., KNUDSON, D. L., PEHU, E., LAPITAN, N. L. V. (1998). Cytological and molecular characterization of repetitive DNA sequences of *Solanum brevidans* and *Solanum tuberosum*. *Genome* 41: 487-494.

ROSA, E. A. S. (1999). Chemical Composition. In: GÓMEZ-CAMPO, C. (Ed.) Biology of *Brassica* Coenospecies. Developments in Plant Genetics and Breeding. pp. 315-346. Elsevier Science, Amsterdam, Netherlands.

SADDER, M.T., WEBER, G. (2001). Karyotype of maize *(Zea mays* L.) mitotic metaphase chromosomes as revealed by fluorescence *in situ* hybridization (FISH) with cytogenetic DNA markers. *Plant Mol*. *Biol*. *Rep*. 19: 117-123.

SANTINI, S., CAVALLINI, A., NATALI, L., MINELLI, S., MAGGINI, F., CIONINI, P. (2002). *Ty1/copia*- and *Ty3/gypsy*-like DNA sequences in *Helianthus* species. *Chromosoma* 111: 192-200.

SCHELFHOUT, C. J., SNOWDON, R. J., COWLING, W. A., WROTH, J. M. (2004). A PCR based B-genome specific marker in *Brassica* species. *Theor*. *Appl*. *Genet*. 190: 917-921.

SCHMIDT, R. (2002). Plant genome evolution: lessons from comparative genomics at the DNA level. *Plant Mol*. *Biol*. 48: 21-37.

SCHMIDT, T., JUNG, C., METZLAFF, M. (1991). Distribution and evolution of two satellite DNAs in the genus *Beta*. *Theor*. *Appl*. *Genet*. 82: 793-799

SCHMIDT, T., HESLOP-HARRISON, J. S. (1994). Variability and evolution of highly repeated DNA sequences in the genus *Beta*. *Genome* 36: 1074-1079.

SCHMIDT, T., HESLOP-HARRISON, J. S. (1998). Genomes, genes and junk: the large scale organization of plant chromosomes. *Trends Plant Sci.* 3: 195-199.

SCHRADER, O., BUDAHN, H., AHNE, R. (2000). Detection of 5S and 25S rRNA genes in *Sinapis alba*, *Raphanus sativus* and *Brassica napus* by double fluorescence *in situ* hybridization. *Theor*. *Appl*. *Genet*. 100: 665-669.

SCHWEIZER, D. (1980). Fluorescent chromosome banding in plants: application, mechanisms, and implications for chromosome structure. In: Davies D. R.,

Hopwood D. A. (Eds.) The plant genome. John Innes Charity, Norwich, pp. 61-72.

SHARPE, A. G., PARKIN, I. A. G., KEITH, D. J., LYDIATE, D. J. (1995). Frequent nonreciprocal translocations in the amphidiploid genome of oilseed rape (*Brassica napus*). *Genome* 38: 1122-1131.

SHIBATA, F., HIZUME, M. (2002). The identification and analysis of the sequences that allow the detection of *Allium cepa* chromosomes by GISH in the allodiploid *A. wakegii*. *Chromosoma* 111: 184-191.

SKARZHINSKAYA, M., FAHLESON, J., GLIMELIUS, K., MOURAS, A. (1998). Genome organization of *Brassica napus* and *Lesquerella fendleri* and analysis of their somatic hybrids using genomic *in situ* hybridization. *Genome* 41: 691-701.

SNOWDON, R.J. (1997). Fluorescence *in situ* hybridisation techniques for *Brassica* species: Methodological development and practical applications. Dissertation, Justus-Liebig-Universität Gießen, 1997.

SNOWDON, R. J. , KÖHLER, W., FRIEDT, W., KÖHLER, A. (1997a). Genomic *in situ* hybridisation in *Brassica* amphidiploids and hybrids. *Theor*. *Appl*. *Genet*. 95: 1320-1324.

SNOWDON, R. J, KÖHLER, W., KÖHLER, A. (1997b). Chromosomal localisation and characterisation of rDNA loci in the *Brassica* A and C genomes. *Genome* 40: 582-587.

SNOWDON, R. J., KÖHLER, A.; KÖHLER, W., FRIEDT, W. (1999). FISHing for new rapeseed lines: the application of molecular cytogenetic techniques to *Brassica* breeding. Proceedings, 10th International Rapeseed Congress, September 1999, Canberra, Australia.

SNOWDON, R. J., FRIEDT, W., KÖHLER A., KÖHLER W. (2000a). Molecular cytogenetic localisation and characterisation of 5S and 25S rDNA loci for chromosome identification in oilseed rape (*Brassica napus* L.). *Ann*. *Bot*. *(London)* 86: 201-204.

SNOWDON, R. J., WINTER, H., DIESTEL, A., SACRISTÁN, M. D. (2000b). Development and characterisation of *Brassica napus*-*Sinapis arvensis* addition lines exhibiting resistance to *Leptosphaeria maculans*. *Theor*. *Appl*. *Genet*. 101: 1008-1014.

SNOWDON, R. J., BÖTTINGER, P., PICKARDT, T., KÖHLER, F., FRIEDT, W. (2001). Physical localisation of transgenes on *Vicia faba* chromosomes. *Chromosome Res*. 9: 607-610.

SNOWDON, R. J., FRIEDRICH, T., FRIEDT, W., KÖHLER, W. (2002). Identifying the chromosomes of the A and C genome diploid Brassica species *B*. *rapa* (syn. *campestris*) and *B*. *oleracea* in their amphidiploid *B*. *napus*. *Theor*. *Appl*. *Genet*. 104: 533-538.

SONG, K., OSBORN, T., WILLIAMS, P. H. (1988). *Brassica* taxonomy based on nuclear restriction fragment length polymorphysms (RFLPs). 1. Genome evolution of diploid and amphidiploid species. *Theor*. *Appl*. *Genet*. 75: 784-794.

SONG, K., TANG, K., OSBORN, T. C. (1995). Rapid genome change in synthetic polyploids of *Brassica* and its implications for polyploid evolution. *Proc*. *Natl*. *Acad*. *Sci*. *U*.*S*.*A*. 92: 7719-7723.

STAGINNUS, C., HUETTEL, B., DESSEL, C., SCHMIDT, T., KAHL, G. (2001). A PCR-based assay to detect *En/Spm*-like transposon in plants. *Chromosome Res*. 9: 591-605.

SVITASHEV, S., BRYNGELSSON, T., LI, X.-M., WANG, R. R.-C. (1998). Genome-specific repetitive DNA and RAPD markers for genome identification in *Elymus* and *Hordelymus*. *Genome* 41: 120-128.

SVITASHEV, S., ANANIEV, E., PAWLOWSKI, W. P., SOMERS, D. A. (2000). Association of transgene integration sites with chromosome rearrangements in hexaploid oat. *Theor*. *Appl*. *Genet*. 100: 872-880.

SYKOROVA, E., LIM, K. Y., CHASE, M. W., KNAPP, S., LEITCH, I. J., LEITCH, A. R., FAJKUS, J. (2003a). The absence of *Arabidopsis*-type telomeres in

Cestrum and closely related genera *Vestia* and *Sessea* (*Solanaceae*): first evidence from eudicots. *Plant J*. 34: 283-291.

SÝKOROVÁ, E., LIM, K. Y., KUNICKÁ, Z., CHASE, M. W., BENNETT, M. D., FAJKUS, J., LEITCH, A. R. (2003b). Telomere variability in the monocotyledonous plant order *Asparagales*. *Proc*. *R*. *Soc*. *London Ser*. *B* 270: 1893-1904.

U, N. (1935). Genome analysis in *Brassica* with special reference to the experimental formation of *B*. *napus* and peculiar mode of fertilization. *Jpn*. *J*. *Bot*. 7: 389-452.

UZUNOVA, M. (1994). Erstellung einer RFLP- Karte von Raps (*Brassica napus* L.) und Kartierung erster züchterisch wichtiger Gene. Cuvillier Verlag Göttingen. Zugl. Diss. Univ. Göttingen.

UZUNOVA, M., ECKE, W., WEISSLEDER, K., RÖBBELEN, G., (1995). Mapping the genome of rapeseed (*Brassica napus* L.). I. Construction of an RFLP linkage map and localization of QTLs for seed glucosinolate content. *Theor*. *Appl*. *Genet*. 90: 194-204.

VALÁRIK, M., ŠIMKOVÁ, H., ŠAFÁŘ, J., DOLEŽELOVÁ, M., DOLEŽEL, J. (2002). Isolation, characterization and chromosome localization of repetitive DNA sequences in bananas (*Musa* spp.). *Chromosome Res*. 10: 89-100.

VERSHININ, A. V., HESLOP-HARRISON, J. S. (1998). Comparative analysis of the nucleosomal structure of rye, wheat and their relatives. *Plant Mol*. *Biol*. 36: 149-161.

VISCHI, M., JURMAN, I., BIANCHI, G., MORGANTE, M. (2003). Karyotype of Norway spruce by multicolor FISH. *Theor*. *Appl*. *Genet*. 107: 591-597.

WANG, X.-H., LUO, P., SHU, J.-J. (1989). Giemsa N-banding pattern in cabbage and Chinese kale. *Euphytica* 41: 17-21.

WANG, Z.-X., KURATA, N., SAJI, S., KATAYOSE, Y., MINOBE, Y. (1995). A chromosome 5-specific repetitive DNA sequence in rice (*Oryza sativa* L.). *Theor*. *Appl*. *Genet*. 90: 907-913.

WANG, Y. P., SNOWDON, R. J., RUDLOFF, E., WEHLING, P., FRIEDT, W., SONNTAG, K. (2004). Cytogenetic characterization and fae1 gene variation in progenies from asymmetric somatic hybrids between *Brassica napus* and *Crambe abyssinica*. *Genome* 47: 724-731.

WEISS, H., SCHERTHAN, H. (2002). *Aloe* ssp.- plants with vertebrate-like telomeric sequences. *Chromosome Res*. 10: 155-164.

WEISS-SCHNEEWEISS, H., RIHA, K., JANG, C. G., PUIZINA, J., SCHERTHAN, H., SCHWEIZER, D. (2004). Chromosome termini of the monocot plant *Othocallis siberica* are maintained by telomerase, which specifically synthesises vertebrate-type telomere sequences. *Plant J*. 37: 484-493.

XIA, X., SELVARAJ, G., BERTRAND, H. (1993). Structure and evolution of a highly repetitive DNA sequence from *Brassica napus*. *Plant Mol*. *Biol*. 21: 213-224.

XIA, X., ROCHA, P.S., SELVARAJ, G., BERTRAND, H. (1994). Genomic organization of the canrep repetitive DNA in *Brassica juncea*. *Plant Mol*. *Biol*. 26: 817-832.

ZANE, L., BARGELLONI, L., PATARNELLO, T. (2002). Strategies for microsatellite isolation: a review. *Mol*. *Ecol*. 11: 1-16.

ZHANG, X., WESSLER, S. R. (2004). Genome-wide comparative analysis of the transposable elements in the related species *Arabidopsis thaliana* and *Brassica oleracea*. *Proc*. *Natl*. *Acad*. *Sci*. *U.S.A*. 101: 5589-5594.

ZHAO, X., WING, R. A., PATERSON, A. H. (1995). Cloning and characterization of the majority of repetitive DNA in cotton (*Gossypium* L.). *Genome* 38: 1177-1188.

ZHAO, X., JI, Y., DING, X., STELLY, D. M., PATERSON, A. H. (1998a). Macromolecular organization and genetic mapping of a rapidly evolving chromosome-specific tandem repeat family (B77) in cotton (*Gossypium*). *Plant*. *Mol*. *Biol*. 38: 1031-1042.

ZHAO, X.-P., SI, Y., HANSON, R. E., CRANE, C. F., PRICE, H. J., STELLY, D. M., WENDEL, J. F., PATERSON, A. H. (1998b). Dispersed repetitive DNA has spread to new genomes since polyploid formation in cotton. *Genome Res*. 8: 479-492.

ZHONG, X. B., FRANSZ, P. F., WENNEKES-EDEN, J., RAMANNA, M. S., VAN KAMMEN, A., ZABEL, P., DE JONG, J. H. (1998). FISH studies reveal the molecular and chromosomal organization of individual telomere domains in tomato. *Plant J*. 13: 507-517.

ZIOLKOWSKI, P. A., SADOWSKI, J. (2002). FISH-mapping of rDNAs and *Arabidopsis* BACs on pachytene complements of selected *Brassicas*. *Genome* 45: 189-197.

Acknowledgement

I am very grateful to Prof. Heiko Becker and Dr. Ulrike Bellin for providing me the opportunity to conduct this project and for the encouragement and support during all these years. Many thanks for the interest to read the manuscripts of this thesis and for the very useful suggestions and discussions. I wish to express my gratitude to "Ulli" for the supervision, friendship and many helps during my work and to Heiko for his patience, gentleness and respectful way of dealing with us students.

My gratitude extends to Dr. W. Ecke for his suggestions during the work with molecular genetic techniques and corrections/suggestions of the manuscript, very important for the improvement of this thesis.

My sincere thanks goes to Prof. Marcelo Guerra (Depto. de Botânica, UFPE, Brazil), for his enthusiasm to critically read the manuscript, in spite of his very busy schedule, and for the many valuable suggestions.

My sincere gratitude to Prof. Köhler, for kindly accept to be the second referee of this thesis and Profs. Karlovisky and Finkeldey for their interest to be my examiners.

Many thanks to Dr. J. Neesen (Institut für Humangenetik, Göttingen), for making available the Olympus microscope, indispensable for the conclusion of the FISH experiments.

My thanks also to Prof. Dr. A. Graner (Genbank- IPK, Gaterleben) for supplying part of the seeds used in this work, and to Prof. Dr. Knudson (Department of Entomology, Colorado State University, USA) for providing the plasmid probe "5+45SrDNA".

The financial support of the DAAD (German Academic Exchange Service), during the language course (1998), and the scholarship provided by CNPq (Brazilian National Council for Scientific and Technological Development) in 1999 are acknowledged. The Deutsche Forschungsgemeinschaft is acknowledged for the financial support for the research project.

Many thanks are given to all friends and colleges of "our institute" for the friendship, advices and good work atmosphere. I am particularly grateful to Anja

Kramer and Lilli Nehlin, for their interest to read the manuscript, Mahmoud Zeid and Rubens Marschalek, for the many helps, Kerstin Diekmann, for the indispensable technical assistance during the FISH experiments, J. Hippe and U. Ammermann, for solving many of my "computer problems", H. Heise, for some home assembled instruments used in this work, W. Link and C. Möllers, for correcting the Zusammenfassung and abstracts.

My thanks also to Dr. Andrea Pedrosa-Harand (Institute of Botany, University of Vienna, Austria) for helpful discussions concerning the FISH results and Dr. Weiss-Schneeweiss (Institute of Botany, University of Vienna, Austria) and Prof. E. Richards (Department of Biology, Washington University, USA) for personal communications.

My especial thanks goes to my mother, my family and my friends, especially those from the "Freudeskreis-Goethe-Institut-98" and from our small "Brazilian community in Göttingen" (particularly families Cattanio, Dahn-Batista, Ide, Marschalek), also to families Nehlin and Kramer/Stenzel, for the moral/emotional support, indispensable when we are living abroad.

Finally, my most especial thanks goes to my beloved daughter Isabelle and husband Martin for patiencie, motivation, inspiration, and exceptional capacity to show me every day how exciting, fun and beautiful life can be...

Curriculum Vitae

Personal data

Name:	Karla Galvão Bezerra dos Santos.
Nationality:	Brazilian.
Date of birth:	04. November 1970.
Place of birth:	Olinda, Pernambuco (PE), Brazil.

Education

1975 - 1985:	Primary school at the Colégio da Sagrada Família, Recife, PE, Brazil.
1986 - 1988:	Secondary school at the Colégio Contato, Recife, PE, Brazil.
1989 - 1992:	Course of Biological Sciences at the Federal Univerity of Pernambuco, Recife, PE, Brazil.
1994 - 1996:	Master Degree in Genetics and Molecular Biology at the Federal University of Rio Grande do Sul, Porto Alegre, Rio Grande do Sul, Brazil.
2000 - 2004:	Ph. D. in Plant Breeding at the Institute of Agronomy and Plant Breeding, Georg-August University, Göttingen, Germany.

Professional Career

1993:	Assistant scientist at the Federal University of Pernambuco, Recife, PE, Brazil.
1997 - 1998:	Lecturer for Human Genetics, Genetics and Cytogenetics at the Department of Genetic, Federal University of Pernambuco, Recife, PE, Brazil.
1999:	Assistant scientist at the Karyotype Evolution Department, Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben, Germany.
2000 - 2004:	Assistant scientist and lecturer for the Practical Course of Plant Cytogenetics at the Institute of Agronomy and Plant Breeding, Georg-August University, Göttingen, Germany.