GENOME INTERACTIONS IN POLYPLOID PLANT SPECIES

An investigation of *Poaceae* (tribe *Triticeae*)

Thesis submitted for the degree of

Doctor of Philosophy

at the University of Leicester

by

Alessandra Contento

Department of Biology

University of Leicester

February 2006

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Declaration

I hereby declare that no part of this thesis has been previously submitted to this or any other University as part of the requirements for a higher degree. The content of this thesis is the result of my own work unless otherwise acknowledged in the text or by reference.

The work was conducted in the Department of Biology, University of Leicester during the period October 2002 to October 2005.

Signed...A. Contento.....

Alessandra Contento, February 2006

Acknowledgments

I would like to thank Dr. Schwarzacher for her supervision during my PhD experience, for teaching me "the tricks" of *in situ* hybridization and immunostaining techniques and for her support during the writing up of my thesis, and Prof. Heslop-Harrison for helping me in the cloning experiments and for reading some chapters of the thesis.

My PhD work was founded primarily from National Environmental Research Council (NERC) to whom I am grateful. Additional funding was provided from Gatsby Foundation for participating to the International Polyploidy Conference (London, April 2003) and the International Chromosome Conference (London, September 2004).

Special thanks are due to Dr. Bailey for always being encouraging and positive and for helping me in the cytological experiments, and to Ms. Liggins for her kind daily support and technical help. I am also very grateful to Dr. Rosato for allow me to use the Real Time PCR machine in his lab and Dr. Bailhache for helping me in the Real Time PCR experiments; Dr. Royo for providing *Triticale* seeds; Dr. Kubis for useful advices; and Dr. Houben for an interesting discussion on DNA methylation changes in allopolyploid plant species.

I would like to thank my present and past colleagues for being very nice friends; in particular Saddiah, Teo, Azhar and Emmanuel for being collaborative and encouraging, especially in the last year of my PhD experience; Celia for many field-excursions and nice dinners; Olly for sharing with me concerts; Nestor, Manuel, Roxana, Rafael and Santiago for teaching me the Spanish language and the Paella recipe.

Special thanks are also due to technical and administrative staff of the Department of Biology for helping me in solving problems (money, forms, computers and printers!) and making my PhD experience easier.

Finally, I am grateful to the many people I have met in the Department of Biology during these three years for nice chats in the tea room or along corridors, and useful advices about my experiments and the writing up of my thesis.

GENOME INTERACTIONS IN POLYPLOID PLANT SPECIES

An investigation of *Poaceae* (tribe *Triticeae*)

by

Alessandra Contento

Abstract

The 120-bp repetitive sequence family, a widespread and old component of *Triticeae* genomes, has been used as a tool for investigating genome interactions in *Triticeae* allopolyploids. Changes in copy number of 120-bp repeat unit family were hypothesised to happen as a result of hybridization events generating *Triticeae* allopolyploids together with increase and decrease in DNA methylation of this major fraction of *Triticeae* genomes. In addition, by using 120-bp repeat unit family sequences for identifying chromosomes from different genomes in *Triticeae* allopolyploids whole-genome DNA methylation patterns were investigated on chromosome spreads. In all *Triticeae* allopolyploid species investigated in the present work, unevenly distributed whole-genome methylation patterns were present with enhanced and reduced methylation signal in different chromosomes pairs and/or different chromosomal regions, while diploid *Triticeae* species showed uniformly distributed whole-genome methylation patterns along their chromosomes. The role of repetitive sequences in relocating the overall DNA methylation through local increase and decrease in different chromatin regions and in contributing to coordinate the heterochromatin of different genome donors in a new allopolyploid species was discussed. Finally, two different examples of alteration of methylation patterns have been investigated in *Triticale* allopolyploids. In the former case, alterations in whole-genome methylation and methylation patterns of two different repetitive DNA fractions, 120-bp repeat unit family and 5S rDNA sequences were observed in the early generations obtained by crossing two advanced *Triticale* lines but the "memory" of methylation code was quickly re-established with no effect on plant growth. In the latter, treatments with 5-azacytidine, an inhibitor of DNA methylation, greatly reduced but not completely erased from chromosomes the methylation signal. Re-methylation processes appeared to be rapid and no genome-specific, however not able to recovery completely the 'memory' of methylation code and thus the methylation alteration permanently affected plant growth and development.

Chapter I: Introduction

In the present work an investigation of genome interactions in polyploid plant species belonging to *Poaceae* tribe *Triticeae* has been described by focusing on the heterochromatic portion of their genomes. A major DNA fraction of *Triticeae* species, the 120-bp repetitive DNA sequence family, present in the heterochromatic regions of many diploid and polyploid *Triticeae* genomes has been chosen as a subject of study. Thus, in order to characterize 120-bp repetitive DNA sequence family, investigations were carried out in diploid and polyploid *Triticeae* species in terms of sequence similarity and polymorphisms, copy number and chromosomal location, for example. In addition, the role of two major epigenetic marks of heterochromatin: DNA methylation and histone H3 methylation, in relation to the occurrence of genome interactions in polyploid *Triticeae* species has been investigated. In this respect, methylation patterns of 120-bp repetitive DNA sequence family and whole-genome DNA methylation patterns as well as whole-genome histone H3 methylation patterns were described in polyploid *Triticeae* species and compared with diploid *Triticeae* species.

Hereafter, the Introduction chapter is focused on two main topics: heterochromatin and polyploidy, which are described across four sections. The first two sections are focused on heterochromatin and present: the structural characteristics of heterochromatin, such as chromosomal location and sequence composition with special attention for repetitive DNA sequences; and the biochemical characteristics of heterochromatin, such as DNA methylation and histone methylation, not only in plant species but also in animals and lower eukaryotes. The last two sections concern polyploidy and describe: the occurrence of polyploidy and its consequences on phenotypes, genomes and genes, especially in plants where polyploidy has been recently considered a major force in evolution; and the role of epigenetic phenomena, involving DNA methylation and histone modifications for example, in relation to the formation of polyploid plant species.

1.1 Heterochromatin: structural characteristics

In this section the need for packing entire genomes into chromosomes as the genome size was increasing from prokaryotes to eukaryotes according to life complexity is presented. Nucleosomes, basic and high order chromatin structures are

described. The differences between euchromatin and heterochromatin features are discussed and in particular the structural characteristics of heterochromatin. Different kinds of repetitive DNA sequences present in heterochromatin are described, such as dispersed and tamdemly organized repetitive DNA sequences. Special attention is focused in tandemly repetitive DNA sequences, mainly satellites and telomeric repetitive DNA sequences, but centromeric repetitive DNA sequences and ribosomal DNA sequences are mentioned as well. Finally, an overview of the genome structure in grass species and its organization in euchromatic and heterochromatic regions is presented as cereals were chosen as a model of study in the present work.

1. 1. 1 Genome size and the chromosome model

Genome sequencing projects and analysis of open reading frames (ORFs) indicates that among prokaryotes the *Escherichia coli* genome is 4.6 Mb and thw *Salmonella enterica* 4.8 Mb carrying 4,300 and 4,600 protein-coding genes respectively (Blattner, Plunkett et al. 1997; Deng, Liou et al. 2003) while in lower eukaryote yeasts *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe* genome sizes are 12 Mb and 13.8 Mb with 6,217 and 5,002 genes respectively (Mewes, Albermann et al. 1997; Wood, Gwilliam et al. 2002). However, in higher eukaryotes much larger genomes have been described. Genome sizes of chicken and mouse are 1,060 Mb and 2,600 Mb respectively both containing a higher number of genes: 20,000-23,000 in chicken and about 30,000 in mouse (Waterston, Lindblad-Toh et al. 2002; Hillier, Miller et al. 2004). Human genome size is about 3,000 Mb containing approximately 30,000-40,000 protein-coding genes (Venter, Adams et al. 2001). In the plant kingdom, three organisms with small genomes have been sequenced so far: the model organism *Arabidopsis thaliana* (125 Mb) with about 25,500 genes (The Arabidopsis Genome Initiative 2000), and two subspecies of rice, which is the smallest cereal crop, *Oryza sativa* ssp. *indica* (466 Mb) and *Oryza sativa* ssp. *japonica* (420 Mb) with estimated number of genes 46,000-55,600 and 32,000-50,000 respectively (Goff, Ricke et al. 2002; Yu, Hu et al. 2002). Remarkably, both *Oryza* subspecies seem to have more genes than the human species in much smaller genomes. Currently, several other plant genomes are being sequenced and a challenging project has been launched for sequencing the huge genome (about 17,000 Mb) of hexaploid wheat *Triticum aestivum* 'Chinese Spring', one of the most important cereal crop species (Gill, Appels et al.

2004). It is evident that by increasing life complexity from prokaryotes to eukaryotes there was a need to package larger genomes into chromosomes and regulate the transcriptional activity of many genes (Bird 1995; Struhl 1999; Heslop-Harrison 2000).

How is the linear DNA molecule packed into chromatin in order to form chromosomes in eukaryotes? A complex structure made of multiple levels of folding is required to condense meters of DNA into the small nucleus. As shown in fig. 1. 1 about 147 bp of the double strand DNA molecule is wrapped in two turns around a core of histone proteins or "core histone octamer" (made of two copies of each of histones H2A, H2B, H3, H4) forming the "nucleosome" (Kornberg and Lorch 1999). Histone proteins consist of different domains: an internal domain, responsible of histone-histone and histone-DNA interactions within the nucleosome, and tails that are arranged radially from the nucleosome surface and contain aminoacids that can be modified at specific sites by methylation, acetylation and phosporylation processes for example, as will be mentioned in the next section of the introduction (Strahl and Allis 2000; Jenuwein and Allis 2001). Linker DNA molecules of variable length and linker histones (histone H1) are located between two nucleosomes (Kornberg and Lorch 1999). As shown in fig. 1. 1 the basic chromatin structure is made of arrays of nucleosomes forming the 10-nm chromatin fiber called "beads-on-a-string" and it is universally accepted, while different models for the secondary structure of the 30-nm chromatin fiber have been proposed according to different conformation of linker DNA molecules, different position of linker histones or different mass per unit length of the fiber (Horn and Peterson 2002). In the classical model of the "solenoid" six to eight nucleosomes per turn are arranged around a central cavity and linker DNA molecules connect two nucleosomes only (one-start helix model described in Dorigo *et al.* 2004). Recently, a two-start helix model has been proposed where linker DNA molecules are able to connect two "adjacent stacks of helically arranged nucleosome cores" resulting in a more compact structure (Dorigo, Schalch et al. 2004). Finally, the tertiary structure of chromatin is the "chromonema" fiber, present in interphase and mitotic chromosomes and ranging from 100 to 300-nm in diameter (Horn and Peterson 2002).

Recent results proved that nucleosomal organization and higher order chromatin structures are not homogeneous in all chromosome regions. Deviant nucleosomes, such as CENP-A in yeast and human centromeres and H2A.Z, a variant of histone H2A highly conserved from protozoa to humans, have been found (Wolffe and Pruss 1996). Chromatin fibers display high levels of heterogeneity according to the presence of such

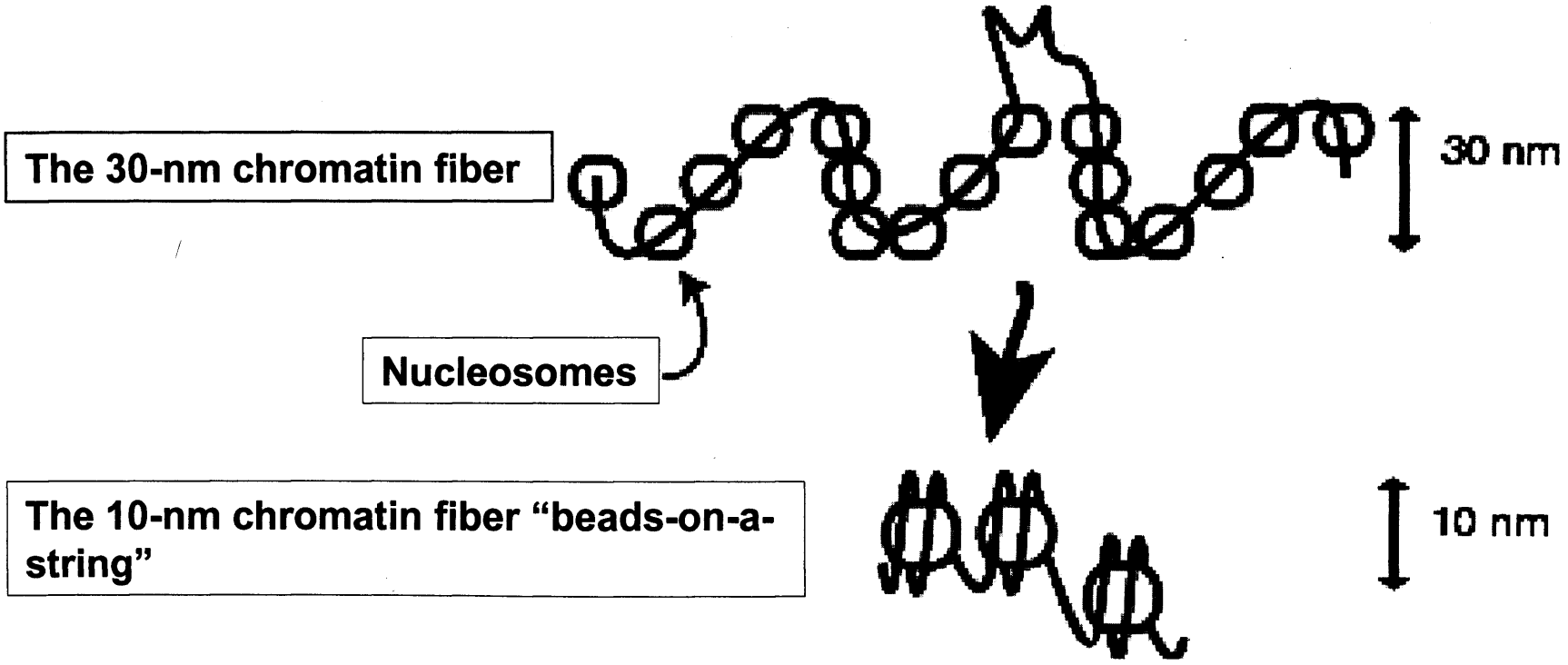


Figure 1. 1. The model of chromatin packing. DNA double helix is wrapped into two turns around nucleosomes, the arrays of nucleosomes form the 10-nm chromatin fiber called "beads-on-a-string" and the secondary structure of 30-nm chromatin fiber (modified from Heslop-Harrison, 2000).

histone variants, as well as linker histone variants (e.g. histone H5), and nonhistone proteins (e.g. HMG or high mobility group proteins) which are additional components of higher order chromatin structures (Horn and Peterson 2002). For example, differences in the nucleosome size of telomeric and bulk chromatin have been observed in human (Tommerup, Dousmanis et al. 1994). Similarly in plants, differences have been reported between telomeric and bulk chromatin in nucleosomal organization and linker DNA length (Vershinin and Heslop-Harrison 1998). In fact, by treating nuclear suspension of cereals with micrococcal nuclease (MNase) that cleaves the DNA linker between two nucleosomes the authors found that telomeric nucleosome size (160 bp) is shorter than bulk chromatin nucleosome size (175-185 bp).

1. 1. 2 Heterochromatin

Along chromosomes two types of chromatin states are observed: heterochromatin and euchromatin. Heterochromatin was originally defined as those parts of chromosomes that do not decondense at the end of telophase, in contrast to euchromatin and hence form chromocenters in interphase nuclei (Heitz 1932). Heterochromatin is a word coming from ancient Greek language literally meaning "different in colour" and refers to its staining properties compared with euchromatin. In fact, early cytological studies used to differentiate heterochromatin regions in eukaryotic metaphase chromosomes by means of different biochemical treatments and staining methods, with fluorescent agents (Q and H banding) or with Giemsa-based techniques (C and G banding) for example (de Jong 2003). More recently, heterochromatin has been defined as a highly condensed permanently inactive state of chromatin while euchromatin is a more open state of chromatin (Hennig 1999): according to the author there is no need to discriminate further between constitutive and facultative heterochromatin, heterochromatin being a "functional state of chromosomal regions or chromosomes" where transcription is repressed.

In table 1. 1 a comparison between the structural features of heterochromatin and euchromatin in complex genomes is presented (Henikoff 2000). The main differences between heterochromatin and euchromatin concern interphase appearance as mentioned above, and chromosomal location: telomeric and pericentromeric for heterochromatin while more distal for euchromatin. In addition, heterochromatin and euchromatin differ in sequence composition and gene density. Heterochromatin is

Table 1. 1. Differences between euchromatic and heterochromatic regions of complex genomes (modified from Henikoff, 2000).

Feature	Euchromatin	Heterochromatin
Interphase appearance	relatively uncondensed	condensed
Chromosomal location	distal	telomeric-pericentromeric
Sequence composition	mostly non-repetitive	repetitive
Gene density	high	low or absent
Meiotic recombination	normal	low or absent
Replication timing	throughout S phase	late S phase
Position-effect variegation	rare	frequent
Nucleosome spacing	variable	regular
Nucleosome accessability	variable	low

mainly made of different kinds of repetitive DNA sequences (DNA satellites, middle repetitive DNA sequences and retroelements) as will be described later and has low or absent gene density; whereas euchromatin is characterized by high gene density. Furthermore, heterochromatic regions are characterized by low or absent meiotic recombination, late replication timing in the S-phase, and frequently show the position-effect variegation (PEV) resulting in the silencing of genes close to heterochromatin/euchromatin boundaries or to telomeres (telomere position effect, TPE). Finally, differences between heterochromatin and euchromatin concern nucleosome spacing and accessibility, heterochromatin having a more regular nucleosome spacing with low accessibility to the activity of enzymes, such as micrococcal nuclease (MNase) for example.

Nevertheless, the heterochromatic portion of chromosomes appears to be less structurally and functionally uniform than expected. As observed in *Drosophila*, α heterochromatin totally lacks active genes whereas other heterochromatin regions called β heterochromatin contain active genes (Lohe and Hilliker 1995). Similarly, a series of recent papers showed that heterochromatic regions contain genes although only few in plant as well (Copenhaver, Nickel et al. 1999; Fransz, Armstrong et al. 2000). Copenhaver *et al.* (1999) showed that centromeres at chromosomes 2 and 4 of *A. thaliana* are made mainly of repetitive DNA sequence families and LTR-retrotransposons but some pseudogenes and few functional genes are also present. Similarly, by investigating the heterochromatic knob (hk4S) on the short arm of chromosome 4 of *A. thaliana*, Franz *et al.* (2000) found repetitive DNA sequences and retrotransposons but also 5S rDNA genes in a long array. Furthermore, to the structural characteristics of heterochromatin many biochemical ones have been recently added, such as DNA methylation (at cytosines) and histone modifications (methylation and deacetylation, for example) as will be described more extensively in the next section of the introduction. In conclusion, a new and more exhaustive definition of heterochromatin is urgently needed as suggested by Bennetzen (2000). A biological rather than technical definition of heterochromatin should take into account its many and different degrees of condensation that possibly "would form a continuum" from constitutive heterochromatin towards euchromatin, its variable and complex structural and biochemical features, and its different functions according to the presence/absence of active genes (Bennetzen, 2000 #323).

1. 1. 3 Repetitive DNA sequences

Eukaryotic genomes contain a high proportion of repetitive non-coding DNA sequences which are classified into two major classes: dispersed and tandemly organized repetitive DNA sequences, the former being interspersed along chromosomes between genes and/or other repetitive DNA sequences while the latter are grouped in tandem arrays of single units (Britten and Kohne 1968; Kubis, Schmidt et al. 1998; Kumar and Bennetzen 1999; Heslop-Harrison 2000).

The hypothesis of "selfish" or "junk" DNA was proposed as repetitive DNA sequences do not seem to have any phenotypic effect or evolutionary function (Doolittle and Sapienza 1980; Orgel and Crick 1980). At present it is generally accepted that repetitive DNA sequences are involved in chromatin and chromosomal packaging, telomere and centromere formation; in addition, repetitive DNA sequences are responsible for correct chromosome pairing during meiosis and event of recombination through crossing-over which can result in chromosomal rearrangements, as well as in regulating gene expression (Schmidt and Heslop-Harrison 1998; Heslop-Harrison 2000; Schwarzacher 2003). Furthermore, some repetitive DNA sequences are efficiently transcribed and for example account of about 3% of total cellular RNA as reported in rice (Wu, 1994 #19).

1. 1. 3. 1 Mobile genetic elements

Mobile genetic elements can be classified into two major groups: DNA retrotransposons and DNA transposons. The mechanisms of insertion of the former is via an RNA intermediate that is converted into DNA by a reverse trascriptase/RNase H enzyme and then inserted into a new location, according to a mechanism of "copy and paste" through which genome size can increase greatly, while the latter use an excision/repair system according to a mechanism of "cut and paste" without affecting genome size.

DNA retrotransposons can be further classified into LTR and non–LTR retrotransposons according to the presence or absence respectively of long terminal repeats (LTRs) that contain promoters and terminators associated with transcription of the LTR retrotransposon itself. The LTR retrotransposons also contain three major genes called *gag* (group specific antigene), *pol* (polyprotein) and *int* (integrase) genes

encoding proteins involved in processes of reverse transcription and insertion, and are subdivided into *Pseudoviridae* (Ty1-copia like elements) and *Metaviridae* (Ty3-gypsy like elements) groups that differ from each other in the gene order (Hull 1999; Kumar and Bennetzen 1999). The non–LTR retrotransposons are subdivided into LINEs (long interspersed repetitive elements) and SINEs (short interspersed repetitive elements) both having a poli (A) tail at the 3' terminus of the element. LINEs have *gag* and *pol* genes but lack the *int* gene and require alternative mechanisms of integration while SINEs are different from the other three classes of retrotransposons mentioned above, lacking of *gag*, *pol* and *int* genes, but having instead two genes similar to tRNA genes, called box A and box B (Kumar and Bennetzen 1999; Schmidt 1999).

In yeast *S. cerevisiae* five families of LTR-retrotransposons are present: Ty1, Ty2, Ty3, Ty4, Ty5 (Kim, Vanguri et al. 1998) but not LINEs and SINEs, conversely in the human genome LINEs and SINEs (e.g. L1 and ALU elements) are highly abundant reaching 35% of the total DNA (Smith 1996). In plants LTR and non-LTR retrotransposons are highly abundant in gymnosperm and angiosperm species and represent a large portion of their genomes (Heslop-Harrison 2000; Friesen, Brandes et al. 2001) especially in species with large genome sizes as will be discussed later for cereals. By using *in situ* hybridization methods and genome sequencing analyses, the organization and distribution of LTR and non-LTR retrotransposons have been investigated in many species of gymnosperms and angiosperms, revealing that generally mobile elements have dispersed distribution patterns along the euchromatin of all chromosomes, located more often between genes but also within or close to genes in some cases (Kumar and Bennetzen 1999; Heslop-Harrison 2000). However, exceptions have been observed for Ty1-copia elements close to telomeric or paracentromeric heterochromatic regions in *A. thaliana* (Brandes, Heslop-Harrison et al. 1997) and Ty3-gypsy localized at centromeres of grass species (Kumar and Bennetzen 1999), for example. According to their location retrotransposons can affect differently the host genome: generating mutations when inserted near or within genes and/or genome rearrangements through unequal or ectopic recombination, and having possibly a structural and/or functional role in centromeric and telomeric heterochromatic regions when present at these locations (Kumar and Bennetzen 1999; Heslop-Harrison 2000).

1. 1. 3. 2 Tandemly repeated DNA sequences

Tandemly repeated DNA sequences were defined as tandem arrays of monomers one following the other (Flavell 1980) and more recently tandem repeats were classified according to the size of their monomers in microsatellites or simple sequence repeats (SSRs), minisatellites and satellite DNA sequences (Charlesworth, Sniegowski et al. 1994).

Microsatellites and minisatellites have motifs of 2-6 and 10-100 bp respectively and represent a major component of the repeated DNA fraction present in a genome. Slippage replication mechanisms were indicated as responsible for the formation and expansion of microsatellites (Schlotterer and Tautz 1992; Hancock 1996). Microsatellites are ubiquitous in eukaryotic genomes (Tautz and Renz 1984; Lagercrantz, Ellegren et al. 1993) and have been largely used as molecular markers and for fingerprinting in fungal, plant and animal species as they evolve rapidly and form genome or species-specific patterns (Heslop-Harrison 2000). In plants, they have been used as molecular markers in many crop species. Furthermore, in plants synthetic oligonucleotide SSRs have been also used for *in situ* hybridization and showed different patterns on chromosomes according to the different motifs and/or the different species, allowing speculation about their evolution (Schmidt and Heslop-Harrison 1996; Cuadrado and Schwarzacher 1998). In satellite DNA sequences (the term satellite DNA refers to the shoulder or satellite visible in density gradient distribution curves) the monomer size ranges from 100 to 500 bp and they has been extensively characterized in both plant and animal species. In human the α-satellite sequences are present in all human chromosomes covering 10% of the genome (Willard 1989). Satellite DNA sequences are widely present in gymnosperms and angiosperms and represents a high proportion of nuclear DNA, especially in plants with large genomes (Kubis, Schmidt et al. 1998; Heslop-Harrison 2000).

Finally, megasatellite DNA sequences made of long monomers up to several kilobases were reported in human (Gondo, Okada et al. 1998) and *Drosophila* (O'Hare, Chadwick et al. 2002) for example. In plants, longer monomers are derived from rearrangements of smaller units and/or events of insertion and duplication as shown in some *Aveneae* species (Grebenstein, Grebenstein et al. 1996).

Currently, a database specialized for plant satellite repeats (PlantSat Database) has been established and about 160 different repeat unit families from gymnosperm and angiosperm species were analysed with respect to monomer length distribution (Macas, Meszaros et al. 2002). Overall, the range of monomer length was from 33 bp to 4 Kb,

but the majority of repeats (91%) was shorter than 600 bp and more frequently their length ranges between 135–175 or 315-375 bp. The most common monomer length was 165 bp (38.5% of the repeat unit families analysed). That is close to the length of the DNA wrapped around the nucleosome. Similarly, previous investigations showed nucleosome phasing (a strict defined arrangement around nucleosomes) of repetitive sequences. In tobacco two repeat unit families with monomers of about 180 and 215 bp long exhibited the property of forming stable curvatures around the nucleosomal core (Gazdova, Siroky et al. 1995; Matyasek, Gazdova et al. 1997) and in cereals the phasing of three different repeat unit families 120, 360 and 550 bp long was described (Vershinin and Heslop-Harrison 1998). It is likely that repetitive sequences have a "key role in stabilizing DNA packaging and higher order chromatin condensation" as suggested by Heslop-Harrison (2000).

Data about nucleotide composition and sequence motifs were also investigated in the repeat unit families present in PlantSat Database (Macas, Meszaros et al. 2002). The frequency of dinucleotide combinations showed that AA/TT was over-represented and TA/TA under-represented in the repeat unit families analysed, while CG/CG was under-represented in some families but not in others where it was relatively abundant. Furthermore, the pentanucleotide CAAAA motif involved in the breakage-reunion mechanism was present in 78.9% of repeat unit families suggesting the occurrence of recombination between repeats in satellite DNA sequences. However, the molecular mechanisms involved in satellite DNA amplification are still poorly characterized. Mechanisms of unequal and illegitimate crossing-over have been proposed (Charlesworth, Sniegowski et al. 1994; McAllister and Werren 1999). Evidences of breakage-fusion-bridge cycles (McClintock 1941) were found in head-to-head or tail-to tail orientation of repeat units in rye, for example (Vershinin, Schwarzacher et al. 1995). Recently, the variation in satellite DNA profiles was described by using the model of "library of satellite sequences" that represent independent evolutionary units (Ugarkovic and Plohl 2002). According to the authors changes in copy number due to unequal crossing-over or in nucleotide sequence due to different rates between mutation and homogeneization/fixation events can contribute to form species-specific satellite profiles. In addition, mutation can be spread unevenly between chromosomes by turnover mechanisms, generating also chromosome-specific satellite families and/or chromosomal rearrangements. Nevertheless, the authors admitted it is still difficult to explain why some repeat unit families are highly conserved for long evolutionary times

while others show rapid changes in copy number and/or nucleotide sequence even between closely related species. However, by noticing that some structural features such as monomer length or A+T content seem to be conserved in the majority of satellite families, as previously reported for about 160 repeat unit families from gymnosperm and angiosperm species (Macas, Meszaros et al. 2002), the authors agreed that repetitive DNA sequences are subject to similar constraints.

Finally, due to their rapid evolution and characteristic distribution patterns in plants, which can be species-, genome- and even chromosome-specific, repeated DNA sequences have been used for phylogenetic analysis of related species, for identification of chromosomes and genomes in hybrids, chomosome-addition and translocation lines, and for studying nuclear architecture (Schwarzacher 1996; Kubis, Schmidt et al. 1998; Heslop-Harrison 2000).

1. 1. 3. 3 Telomeric repetitive DNA sequences

Telomeres are the physical ends of chromosomes providing a range of different functions with the unifying aim of preserving chromosome integrity. Telomeres are involved in protecting chromosome ends from exonucleolitic degradation, end-to-end fusions and other rearrangements; in maintaining the length of the chromosomes by using a reverse-transcriptase-like enzyme, called telomerase, which uses an RNA template, as the replication machinery cannot provide semiconservative DNA replication at the ends of double strand DNA molecules; and finally in allowing chromosome segregation during cell division (Louis and Vershinin 2005).

Eukaryotic telomeric DNA sequences in most species consist of the consensus sequence ($dT/A_{1-4}dG_{1-8}$) organized in tandem arrays: in yeast TG_{1-3} repeats and in human (TTAGGG)n repeats are present for example (Louis and Vershinin 2005). The *Arabidopsis*-type telomere sequence (TTTAGGG)n is widespread in most plant species, with the exception of some species of the Asparagales (Adams, Hartman et al. 2001; Sykorova, Lim et al. 2003), and is characterized by a species-specific length, from 2 to 5 Kb in *A. thaliana* (Richards and Ausubel 1988) and from 60 to 160 Kb in tobacco (Fajkus, Kovarik et al. 1995) for example. Among cereals (Schwarzacher and Heslop-Harrison 1991; Kilian, Stiff et al. 1995; Vershinin and Heslop-Harrison 1998), different length of telomeric repeats have been described in rye (from 8 to 50 Kb), barley (from 20 to 80 Kb) and wheat (from 15 to 150 Kb), and furthermore a 25-fold variation (from

1.8 to 40 Kb) was reported in 22 inbred lines of maize (Burr, Burr et al. 1992). Plant telomeric repeats are added by the telomerase in tissues with dividing cells, such as root tips or floral organs, but not in quiescent tissues such as stems and leaves (Murata 2002).

Despite yeast telomeres being assembled in a non-nucleosomal structure (Wright, Gottschling et al. 1992), in higher eukaryotes a nucleosomal organization is present at the physical ends of all chromosomes, and differences in nucleosomal organization and nucleosome size between telomeric and bulk chromatin have been reported in human and cereals as previously mentioned (Tommerup, Dousmanis et al. 1994; Vershinin and Heslop-Harrison 1998).

Telomeric-associated sequences (TASs) are involved in creating a structure of transition between the telomere and the rest of the chromosome. In yeast subtelomeric regions are made of tandem repetitive DNA sequences (Y' element and core X) but also LTR-retroelements (Ty5) and several gene families (Louis 1995); in human among TAS sequences some members are highly polymorphic and associated with telomeres, some others are not exclusively localized at telomeres (Brown, MacKinnon et al. 1990).

In plants, subtelomeric regions have shown a variable organization according to different species and their genome organization. Single-copy TAS sequences (<1kb) have been found at the ends of chromosomes in *A. thaliana* (Richards, Chao et al. 1992), suggesting that in species with small genomes TAS sequences are not a relevant feature and telomere/subtelomere organization is quite simple. Furthermore, in *A. thaliana* unique DNA sequences have been described within the telomeric repeats (Richards, Chao et al. 1992) and rDNA genes close to telomeres at the end of chromosomes 2 and 4 (Copenhaver and Pikaard 1996). Some TAS sequences have been described in rice, such as Os48 repetitive sequence that is arranged in arrays of a 355 bp repeat unit and is located at or near the chromosome ends (Wu, Chung et al. 1991; Wu and Tanksley 1993). TAS sequences of rice seem to be not highly repeated but highly polymorphic as described in two *Oryza sativa* varieties by using RFLP analysis (Ashikawa, Kurata et al. 1994). In addition, a chromosome-specific organization of TAS sequences in rice has been reported, as a A-genome *Oryza* specific sequence, called TrsA, showed a localization restricted to the long arm of chromosomes 6 and 12 by using *in situ* hybridization on extended DNA fibers (Ohmido, Kijima et al. 2000; Ohmido, Kijima et al. 2001). In angiosperm species with large genomes highly or middle repeated TAS sequences are highly polymorphic even between closely related

species and can be located directly adjacent to telomeres or separated from them by a spacer region, as reported in tobacco and tomato respectively. Several TAS sequences have been isolated from different species of tobacco and appeared to be directly attached to telomeric repeats: the HRS60 sequence, with a 180 bp repeat unit, from *Nicotiana tabacum* (Fajkus, Kovarik et al. 1995; Fajkus, Kralovics et al. 1995); the NP3R and NP4R, with 165 and 180 bp repeat unit respectively, from *N. plumbaginifolia* (Chen, Wang et al. 1997); and the TAS49 dispersed repetitive sequence, not only restricted at subtelomeric locations, from *N. tomentosiformis* (Horakova and Fajkus 2000). In tomato (*Lycopersicon esculentum*), the satellite DNA sequence TRG1 is present on 20 out of 24 telomeres, forming large clusters of 100-1000 Kb in size at the end of chromosomes and representing about 2% of the tomato genome (Schweizer, Ganal et al. 1988; Lapitan, Ganal et al. 1989; Ganal, Lapitan et al. 1991). In addition, a spacer region of few hundred kilobases between the telomeric repeats and TRG1 sequence was found (Ganal, Lapitan et al. 1991) and the organization of tomato chromosome ends described in detail by using *in situ* hybridization on extended DNA fibers (Zhong, Fransz et al. 1998). Zhong *et al.* (1998) showed that the telomeric repeat (TR) and the TGR1 sequence are differently organized according to four different classes: (TR)-spacer-TGR1, (TR)-TGR1 without spacer, only TR and only TGR1, overall suggesting that chromosome ends of tomato have their own telomere/subtelomere organization.

1. 1. 3. 4 Centromeric repetitive DNA sequences

Centromeres are necessary for the attachment of microtubules to chromosomes, for the subsequent alignment of chromosomes at metaphase plate and finally for the separation of single chromatids at the two opposite poles during mitotic and meiotic cell division. In yeast *S. cerevisiae*, centromeres are made of 125 bp sequences containing three elements: CDEI, CDEII and CDEIII (Clarke 1990), while human centromeres contain α-satellite sequences with a monomer of 171 bp in length (Lee, Wevrick et al. 1997). In angiosperms species, centromere-specific repetitive DNA sequences have been isolated from monocots and dicots. In *A. thaliana,* centromeres of all five chromosome pairs are preferentially made of a 180 bp repeat unit family presenting variants at some chromosomes (Heslop-Harrison, Murata et al. 1999) together with several other repetitive DNA sequences (Brandes, Thompson et al. 1997). More detailed

analyses of centromeres of chromosomes 2 and 4 in *A. thaliana* describing repetitive DNA sequences and retrotransposons were mentioned above (Copenhaver, Nickel et al. 1999). In rice centromeres, complex structures made of different families of highly and middle repetitive DNA sequences have been found (Dong, Miller et al. 1998). In maize centromeres, three major families of tandemly organized DNA sequences, each one with a variable copy number in different chromosome pairs, often interrupted by insertions of retroelements, have been described (Ananiev, Phillips et al. 2000). In several species of cereals, Ty3-*gypsy* retrotransposons are present at centromeres suggesting an ancient event of insertion before their divergence from a common ancestor or that they have been preferentially amplified at centromere locations in different species as a consequence of a similar specialized function (Kumar and Bennetzen 1999).

1. 1. 3. 5 Ribosomal DNA sequences

In plants, the ribosomal RNA genes: 45S loci and 5S genes, encoding the structural RNA components of ribosomes, are tandemly organized, highly repetitive and localized at one or more sites per chromosome set (Heslop-Harrison 2000). The 45S loci are made of tandem arrays of the repeat unit containing rRNA genes 18S, 5.8S and 26S together with transcribed and non-transcribed spacers. Each unit is about 10 kb long and repeated hundreds or thousands times. The 5S rRNA genes are also organized in long tandem arrays. In *A. thaliana* the 45S loci is one major class of repetitive DNA representing about 8% of the total genome (Pruitt and Meyerowitz 1986). Similarly, the 5S rRNA genes have been shown to be highly repetitive when mapped to mitotic chromosomes in different ecotypes of *A. thaliana* (Murata, Heslop-Harrison et al. 1997). In plants, the ribosomal RNA genes appeared highly conserved in their units, but changes in their chromosomal distribution have been used to assess evolutionary trends, for example in *Triticeae* species (Castilho and Heslop-Harrison 1995; Taketa, Harrison et al. 1999).

1. 1. 4 Grass genomes

All grass species (*Poaceae* family) diverged from a common ancestor about 50-70 Mya (Kellogg 1998). As a result, recent studies of comparative mapping carried out

on several cereal species found that location and order of markers (co-linearity) are conserved across their different but related genomes. Co-linearity seems to be disrupted only by the occurrence of gross chromosomal rearrangements. For example, by comparing the seven chromosome pairs of diploid species of wheat, barley, rye and wild relatives (all with 2n=14), inversions and translocations have been observed, which appeared unrelated to phylogenetic distance or breeding system (Devos and Gale 2000). In addition, co-linearity does not always agree with microcolinearity (co-linearity at DNA sequence level) due to small local gene rearrangements as a consequence of deletions/insertions or translocations, gene duplications, and the existence of gene families (Devos and Gale 2000; Feuillet and Keller 2002). However, on the basis of generally high levels of co-linearity ten grass genomes have been described by using only about 30 linkage blocks from rice, the smallest cereal species (Devos and Gale 2000).

Despite gene order being generally conserved, cereal genomes mainly differ in their genome size. A 35-fold divergence in genome size has been measured by comparing the small genome of *O. sativa* (rice) with C-value (DNA content per haploid nucleus) of about 0.5 pg (corresponding to about 490 Mb) with the large genome of hexaploid *T. aestivum* (bread wheat) with C-value of about 17.33 pg (corresponding to about 17,000 Mb) (Bennett and Leitch 2004).

Data about gene distribution on genetic and physical maps described gene-rich regions and gene-poor regions in cereal species. The former comprise both high gene density clusters, where gene density is about one gene every 5-20 Kb, as well as individual genes separated by larger distances, while the latter comprise intergenic regions made of repetitive non-coding DNA sequences (Feuillet and Keller 2002). As a consequence of unequal gene distribution on chromosomes, a poor correlation has been found between genetic distances of genes and markers and their corresponding distances along chromosomes, in particular in the large chromosomes of cereals (Schwarzacher 1996) suggesting uneven distribution of recombination events possibly due to specific sequence motifs and/or chromatin configuration in the different chromosomal regions (Schwarzacher 2003). Feuillet and Keller (2002) reviewed examples where larger genomes had larger intergenic regions, for example in maize and barley compared to rice and sorghum, and concluded that intergenic regions were positively correlated with genome sizes. Similarly, by comparing rice chromosome 1 and a chromosome from wheat group 3 it has been shown that both display similar gene content and order but

the proximal regions of wheat chromosomes were made of blocks of repetitive DNA sequences which were absent in rice (Moore 2000).

Different mechanisms account for the marked differences in genome size, that mainly affect repetitive non-coding DNA sequences present in intergenic regions, observed in cereal species. Events of polyploidy, resulting from the duplication of a single complete genome or from the combination of two or more genomes, and mechanisms of genome expansion due to mobile DNA elements have been proposed (Feuillet and Keller 2002). In fact, polyploid species are common among cereals as will be described later, and mobile elements represent the 60% of genome or even more in species with large genomes such as maize, wheat and barley, but less than 50% in rice, whereas only 10% in species with very small genomes such as *A. thaliana* (Bennetzen 2002). In addition, the nested structure of LTR-retrotransposons inserted into other LTR-retrotransposons has been observed mainly in species with large genomes such as maize, wheat and barley (Feuillet and Keller 2002; Bennetzen, Ma et al. 2005).

As polyploidy and mobile DNA element amplification are responsible for genome expansion, the hypothesis that plants have a one-way ticket to genomic obesity has been proposed (Bennetzen and Kellogg 1997). However, some studies have reported mechanisms of genome contraction as well. Unequal intra-strand recombination occurs between LTR-retrotransposons leading to solo LTR elements, but similarly it can affect tandem repeats in direct orientation on the same chromatid (Bennetzen 2002). Furthermore, illegitimate recombination that does not require the regions of homology could be another mechanism responsible for decrease in genome size, as shown in *A. thaliana* and rice (Bennetzen, Ma et al. 2005). Finally, decrease in genome size can be due to the elimination of non-coding low-copy and repetitive DNA sequences after a polyplodization event as will be described later. Bennetzen *et al.* (2005) suggested that different species could use different mechanisms for decreasing genome size with variable efficiency, possibly explaining some of the differences observed between their genome sizes.

1. 2 Heterochromatin: biochemical characteristics

In this section biochemical characteristics of heterochromatin are presented: DNA methylation and histone modifications such as hypoacetylation of histones H3 and H4 and methylation of histone H3. The mechanisms and enzymes involved in DNA

methylation and demethylation processes are described together with the factors responsible for interpreting this epigenetic DNA feature of heterochromatin. The biological role of DNA methylation and its evolutionary consequences are discussed. Special attention is focused on DNA methylation in plant genomes, its amount and distribution at symmetrical and asymmetrical sites, and plant-specific functions of DNA methylation. Histone modifications and the "histone code" hypothesis are described. The link between DNA methylation and histone modifications is also discussed in the wider context of chromatin remodelling processes and silencing mechanisms involving RNA molecules.

1. 2. 1 Epigenetic phenomena and biochemical characteristics of heterochromatin

'Epigenetic phenomena are mitotically and meiotically heritable changes that do not involve changes in the nucleotide sequence but could have phenotypic and hence evolutionary effects' (Wu and Morris 2001). They comprise among others covalent modifications of DNA and histones such as acetylation, phosphorylation and methylation, for example. Heterochromatin is characterized by a series of epigenetic marks affecting DNA and histones. A first biochemical marker of heterochromatin is DNA methylation, the most common DNA modification in eukaryotic organisms, especially in mammals and plants (Richards and Elgin 2002). In addition, modifications at the N-terminal tails of histones H3 and H4 are also prominent features of heterochromatin: in particular, hypoacetylation of histones H3 and H4 and methylation of histone H3 at Lysine 9 (Richards and Elgin 2002).

1. 2. 2 DNA methylation: definition, methylation and demethylation mechanisms

DNA methylation is an important epigenetic feature in fungi, plants and animals. As shown in fig. 1. 2 (Singal and Ginder 1999) it is a post-replicative modification whereby a methyl group is added to the carbon position 5 in the cytosine ring by a DNA methyltransferase enzyme creating 5-methylcytosine (5-^{m}C).

DNA methyltransferase enzymes have been isolated and characterized from eukaryotes and also from prokarytes, and DNA methyltransferase genes represent a

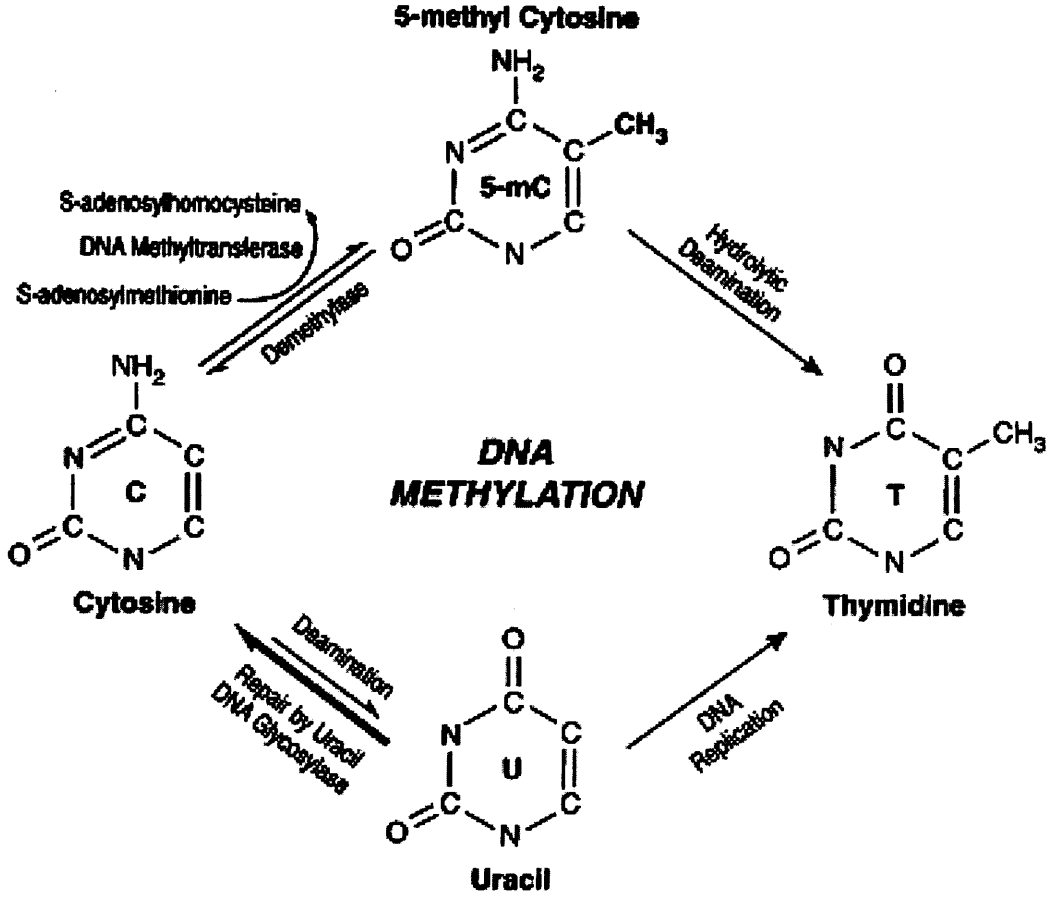


Figure 1. 2. Schematic representation of biochemical pathways for cytosine methylation, demethylation, and mutagenesis of cytosine and 5-methylcytosine. From Singal and Ginder, 1999.

highly conserved gene family indicating that DNA methylation is an ancestral form of DNA modification. In fact, by sequence comparison of the nine most conserved motifs (motifs I, II and IV-X) present in the catalytical domain of prokaryotic and eukaryotic methyltransferase genes, Colot and Rossignol (1999) built a phylogenetic tree showing five subfamilies. By focusing on fungi, plants and animals the three main subfamilies comprise: a group of maintenance DNA methyltransferase genes isolated from species belonging to each of the three kingdoms and coding for enzymes responsible for maintenance DNA methylation after DNA replication, a group of *de novo* DNA methyltransferase genes responsible for establishing DNA methylation at new cytosine sites, and a group of chromomethylases (CMTs) only present in plant species. The authors concluded that in fungi, plants and animals the diversity could have originated from a common ancestor during the diversification of life in the three kingdoms through cycles of duplication events and subsequent functional divergence.

In plants a higher number of DNA methyltransferases has been reported than in animals, and in *A. thaliana* there are at least 10 genes encoding DNA methyltransferases (The Arabidopsis Genome Initiative 2000). The MET group comprises genes responsible for maintenance DNA methylation, the group of Domain Rearranged Methyltrasferases (DRMs) seems to be involved in *de novo* DNA methylation, and the group of chromomethylases (CMTs), unique to plants and characterized by having a chromo-domain, responsible for non-CG methylation as will be described later (Finnegan and Kovac 2000; Martienssen and Colot 2001).

More difficulties have been encountered in isolating or indeed in proving the existence of DNA demethylase enzymes. As reviewed in Bird (2002) a glycosylase has been shown to be responsible for active DNA demethylation in chicken embryos and a mammalian protein has shown demethylase activity at CpG sites, but DNA demethylase enzymes still need to be fully characterized. In plants, a glycosylase has been shown to be involved in demethylating the maternal copies of MEDEA gene in the endosperm of *Arabidopsis* (Choi, Gehring et al. 2002). Another possibility for active DNA demethylation is a repair-like process that replaces methylcytosine with cytosine (Kress, Thomassin et al. 2001) whereas the breakage of the carbon-carbon bond between the pyrimidine ring and the methyl group is thermodynamically unlikely (Bird 2002). Conversely, passive DNA demethylation was though to be originated by affecting the maintenance DNA methylation mechanism after DNA replication, for example as a

consequence of unavailability or dysfunction of maintenance DNA methyltransferases (Jones and Takai 2001; Kress, Thomassin et al. 2001).

1. 2. 3 Interpreting the DNA methylation signal

Once a DNA methylation mark is present on defined DNA sequences a system for reading it is required. Two models have been proposed for interpreting the DNA methylation signal and repressing transcription: the repulsive and the attractive models. In the repulsive model DNA methylation blocks directly the binding between proteins, such as transcription factors, and their target DNA sequences (Tate and Bird 1993). On the contrary, the attractive model refers to proteins that are attract by methylated DNA, such as MeCP2 (Lewis, Meehan et al. 1992), the methyl-CpG binding domain proteins MBD1-MBD4 (Hendrich and Bird 1998) and the protein Kaiso (Prokhortchouk, Hendrich et al. 2001). In plants methyl-CpG binding domain proteins (MBD) did not contain most of the domains present in the mammalian MBD proteins, suggesting that despite the mechanisms for establishing and maintaining DNA methylation are similar between plants and animals, the ones for interpreting the DNA methylation signal evolved independently (Springer and Kaeppler 2005). In addition, among plants evolutionary divergence of MBD proteins has been reported as an ongoing process between monocots and dicots (Springer and Kaeppler 2005). Jaenisch and Bird (2003) claimed the attractive model made of many different proteins that can bind methylated DNA sequences to be extremely interesting. According to the authors the example of methyl-CpG binding domain proteins, performing multiple functions and thus showing the property of redundancy, suggest that gene repression could be "a multilayered process" involving many different factors and it is unlikely to happen by chance or mistake.

1. 2. 4 DNA methylation: biological role

Despite DNA methyltransferases being widespread enzymes in fungi, plants and animals, variable levels of DNA methylation in eukaryotic genomes belonging to each of the three kingdoms have been found. In fungi, DNA methylation is present in some species such as *Neurospora crassa* and *Candida albicans*, but not in *Saccharomyces cerevisiae*, which lacks DNA methyltransferases (Colot and Rossignol 1999). Among

invertebrates, in worm *Caenorhabditis elegans* 5-methylcytosine as well as DNA methyltransferases have not been detected, while fruitfly *Drosophila melanogaster* contains very low levels of DNA methylation and a DNA methyltransferase-like gene (Colot and Rossignol 1999; Tweedie, Ng et al. 1999). Other invertebrate genomes have fractions of methylated DNA separated by unmethylated DNA, as reported in sea squirt *Ciona intestinalis*, for example (Bird 2002). In an attempt at classification, Tweedie *et al.* (1997) suggests that genomic methylation patterns could be divided into three types: nonmethylated, partially methylated, and globally methylated. The last-one is typical of vertebrate genomes that show the highest levels of methylation in the animal kingdom. Vertebrate whole-genome methylation affects the coding and non-coding DNA (Yoder, Walsh et al. 1997; Walsh and Bestor 1999). Conversely in plants, methylation is a common and widespread DNA feature mainly affecting non-coding repetitive DNA sequences but is generally not present in genes (Finnegan, Genger et al. 1998). Some authors proposed that the difference could be due to a different distribution of repeats in the genomes of plants and animals: in plants transposable elements are mainly confined outside genes (Kumar and Bennetzen 1999) whereas in mammals a systematic colonization of introns by transposons is responsible for methylation of exons (Yoder, Walsh et al. 1997; Colot and Rossignol 1999; Walsh and Bestor 1999).

Although cytosine methylation seems to be an epigenetic mechanism already present in the common ancestor of eukaryotes, possibly it is not related to any essential function in eukaryotic organisms being absent in some species in each of the three kingdoms (Colot and Rossignol 1999). The authors proposed instead that cytosine methylation is a DNA mark to be used by different organisms "for setting up different functions which can coexist within the same type of organism or differ from one organism to another", for preventing or allowing gene expression, repressing homologous recombination, or being a signal of parental imprinting, for example.

Two contradicting hypotheses have been proposed for the evolutionary role of DNA methylation: the first one considers that DNA methylation arose as a consequence of the increase of gene number and biological complexity in higher organisms (Bird and Tweedie 1995; Bird 1995; Bird 1997), while the second one suggests it arose as a nuclear host-defence system against parasitic sequences such as retroelements (Yoder, Walsh et al. 1997; Walsh and Bestor 1999). Regarding the first hypothesis, Bird and collaborators compared the number of genes in many prokaryotic and eukaryotic organisms and noticed that during increasing life complexity gene number did not

increase uniformly through small increments. Conversely, two major changes in gene number happened at the prokaryote/eukaryote boundary and the vertebrate/invertebrate boundary. A much higher number of genes allowed the first organisms to acquire new functions and gain higher biological complexity (from unicellular to multicellular), but more precise mechanisms controlling gene expression were also required, especially for having cell differentiation in the different tissues and an efficient gene silencing of the non tissue-specific genes. In this respect, Bird and and collaborators argued that at the prokaryote/eukaryote boundary methylation was necessary in the first eukaryotes together with two other new features, the nuclear envelope and the nucleosomal chromatin, while at the vertebrate/invertebrate boundary high levels of DNA methylation provide an additional layer of gene repression and reduction of the transcriptional noise allowing the first vertebrates to use the genes derived from duplications as raw genetic material to acquire new functions. The role of DNA methylation related to the increase of gene numbers and biological complexity for setting up a variety of different functions and/or for remodeling polyploid genomesm, as will be discussed later, is generally accepted (Colot and Rossignol 1999; Martienssen and Colot 2001).

Regarding the second hypothesis, Yoder (1997) described the mammalian genome as an ecological niche for parasitic sequences such as transposable elements, (for example 35% of the human genome is made up of transposons) and DNA methylation as a mechanism of defence against mobile elements by preventing their transposition and trancription. However, the Yoder-hypothesis was debated (Bird 1997; Martienssen 1998). Bird (1997) argued against it, proving that DNA methylation does not control mobile element transposition in germ lines and embryonic cells where the danger and risk of transposition events should be more severe than in somatic tissues. Martienssen (1998) questioned to the Yoder-hypothesis arguing that transposons could be active even in heavily methylated genomes and furthermore are present in organisms that do not show methylation or at very low level, such as *S. cerevisiae* and *D. melanogaster* respectively (Colot and Rossignol 1999; Tweedie, Ng et al. 1999) where only transcription factors and chromatin proteins are responsible for transposon control. Martienssen (1998) proposed instead "most likely in higher organisms transposons are hidden from the genome by chromatin factors some of which could be mediated by methylation". Conversely, Walsh and Bestor (1999) agreed with Yoder-hypothesis

suggesting that cytosine methylation is "a biochemical specialization not only involved in regulating gene expression but also in silencing parasitic elements".

1. 2. 5 DNA methylation in plants

To assess DNA methylation amount, distribution and patterns, a wide range of methods has been used. Classical methods are High-Performance Liquid Chromatography (HLPC), Restriction Fragment Length Polymorphism (RFLP) with methylation-sensitive restriction enzymes and Southern hybridization, while more recently developed methods are Bisulfite Sequencing, Amplification Fragment Length Polymorphism (AFLP) and cDNA-AFLP, Methylation-Sensitive Amplification Polymorphism (MSAP), Immunostaining with anti-methylcytosine antibody and Microarray analysis.

1. 2. 5. 1 Cytosine methylation: amount

In flowering plants 5-methylcytosine generally accounts for 20-30% of the total cytosines (Shapiro 1976) with the exception of species with small genomes and low repetitive DNA content, such as *A. thaliana* where only 5-6% of the total cytosines has been shown to be methylated (Leutwiler, Hough-Evans et al. 1984). By using High-Performance Liquid Chromatography (HPLC) it has been found that 30-33% of all cytosines is methylated in the tobacco genome (Wagner and Capesius 1981) while 22% in the wheat genome (Wagner and Capesius 1981) and up to 25% in the maize genome (Papa, Springer et al. 2001). In addition, it was reported that about 23% of cytosine is methylated in the tomato genome but differences were recorded between different tissues: protoplasts (20%), mature tissues (25%), mature pollen (22%) and seeds (27%) (Messeguer, Ganal et al. 1991). Furthermore, by HPLC analysis of different nuclear DNA fractions obtained by centrifugation in Cs_2SO_4 density gradients, the amount of 5-methylcytosine has been compared in different plant species, dicots such as pea and tobacco and monocots such as wheat and maize (Montero, Filipski et al. 1992). The authors found that methylated cytosine amount is highly correlated with the G+C content of the different DNA fractions, while the unmethylated cytosine amount is the same despite the differences in G+C content between fractions.

In contrast to animals where cytosine methylation occurs mainly at the nucleotide CpG, in plants DNA methylation affects CpG and non-CpG sites as well: for example CpNpG, (more frequently in CpApG and CpTpG), and asymmetrical CNN (where N=A, T or G) sites (Martienssen and Colot 2001). The strand symmetry of CpG and CpNpG sites provides a mechanism for transmitting their methylation patterns through the semiconservative DNA replication; conversely methylation at asymmetrical sites is probably not transmitted through semiconservative DNA replication while it occurs through *de novo* DNA methylation after each cycle of DNA replication.

Data are accumulating about methylation at both CpG and CpNpG sites in many plant species. It was reported for example that approximately 80% of cytosines in CpG sites is methylated while CpNpG methylation only accounts for about 30 to 40% of total cytosine methylation in wheat seedlings (Gruenbaum, Naveh-Many et al. 1981). Conversely, by using methylation-sensitive restriction enzymes Messeguer *et al.* (1991) estimated that 55% of CpG and 85% of CpNpG sites are methylated in tomato nuclear genome. By using a combined approach of methylation-sensitive restriction enzymes *Msp*I-*Hpa*II and *Acc*III and thin-layer chromatography it was estimated that about 20-23% of internal and external cytosines is methylated at 5'-CCG-3' trinucleotides in the *Arabidopsis* genome, suggesting that the possibility to methylate the first and/or the second cytosine in 5'-CCG-3' trinucleotides gives additional flexibility to the DNA methylation system (Jeddeloh and Richards 1996). The presence of at least two independent methylation codes (CpG and CpNpG) might be related to the fact the plants are sessile organisms and require a fine adaptation of their genomes to environmental conditions (Jeltsch 2002).

Recently, it has been shown that methylation of cytosines in non-symmetrical sites occurs in plants, as in some fungal genomes (Martienssen and Colot 2001). Recent studies have shown that two plant specific DNA methyltransferase classes: the chromomethylases (CMTs) and the Domains Rearranged Methylases (DRMs) are responsible non-CG methylation (which comprises both CpNpG and asymmetric methylation) so widely abundant in plant genomes (Henikoff and Comai 1998; Bartee, Malagnac et al. 2001; Lindroth, Cao et al. 2001; Papa, Springer et al. 2001).

1. 2. 5. 2 Cytosine methylation: distribution

In plants DNA methylation affects the nuclear genome, rather than the chloroplast and mitochondrial genomes, and mainly repetitive DNA sequences, such as transposable elements and tandem repetitive DNA sequences (Finnegan, Genger et al. 1998).

In maize a genome-wide approach showed that methylation affects mainly transposable elements and other repeats (Rabinowicz, Schutz et al. 1999). A partial genomic library of maize nuclear DNA was made by using a methylation-restrictive *E. coli* strain, provided of *Mcr*BC restriction enzyme that can digest methylated DNA, and only 3.3 % of its clones contained retrotransposons and other annotated repeats, while 48.7% clones in a second genomic library made in a strain defective in the *Mcr*BC modification-restriction system. As a result, the first library had a higher number of clones containing known coding sequences (8.3%) versus the second library (1.4%). In a following paper, the authors investigated DNA methylation of transposons compared with exon sequences in maize using the McrPCR technique, where total genomic DNA was firstly digested with *Mcr*BC restriction enzyme and then amplified by PCR with primers specific for transposons and genes (Rabinowicz, Palmer et al. 2003). The results showed that while 95% of exons were amplified after complete digestion with *Mcr*BC restriction enzyme, most of transposons were not amplified or in very low amount, suggesting cytosine methylation was a common feature in their sequences.

As reported above, among DNA satellite sequences (Macas, Meszaros et al. 2002) the CG dinucleotide combination is under-represented in some repeat unit families, but not in others where it is relatively abundant. Data are accumulating in the literature about DNA methylation levels and sites of repetitive DNA sequence families isolated and characterized from different species. In rice, a repeated tandemly organized DNA sequence mainly located in the telomeric regions (Os48) is heavily methylated at CpG and CpNpG sites (Cheng, Stupar et al. 2001). Similarly, in *Nicotiana tomentosiformis* the NTRS repeat unit family contains CpG and CpNpG sites in different amounts, the latter being present in higher frequency (1 site per 2.5 monomeric units). However they appear heavily methylated at both sites (Matyasek, Gazdova et al. 1997). Two other repetitive DNA sequences of *N. tabacum* (HRS60, GRS), both having at least one CCGG site in their monomeric units, showed cytosines at the inner positions nearly completely methylated but not at the outer positions where differences in methylation levels between the two repeat unit families have been found (Kovarik, Koukalova et al. 2000).

Among ribosomal DNA sequences, DNA methylation levels at CpG and CpNpG sites of 25S rDNA gene were investigated in several species of angiosperms reporting that 40-70% of rDNA units were completely or nearly completely methylated at CpG sites while CpNpG methylation levels were more variable between species (Kovarik, Matyasek et al. 1997). In *A. thaliana* and *S. cereale* cytosine methylation levels of 5S rDNA sequences were measured by bisulfite sequencing (Fulnecek, Matyasek et al. 2002) and compared with previous analyses performed by the same authors in *N. tabacum* (Fulnecek, Matyasek et al. 1998) showing that CpG sites were highly methylated in all species analysed while CpNpG and asymmetrical sites showed variable levels of methylation between the three species. In fact, in *A. thaliana* and *S. cereale* the levels of methylated cytosines of 5S rDNA sequences were higher than whole-genome methylation level and represented 27% and 24% respectively of the total cytosine residues. In addition, in *A. thaliana* and *S. cereale* CpG sites (78% and 85% respectively) were more methylated than CpNpG sites (41% and 53% respectively) while asymmetrical methylation accounted only for 6% in *A. thaliana* and was almost negligible (1%) in *S. cereale*. On the other hand, in *N. tabacum* 5S rDNA sequences showed a level of methylation (44% of all cytosine residues) higher than whole-genome methylation, but CpG and CpNpG were affected to a similar extent (90% and 89% respectively), while methylation at asymmetrical sites was low (15%).

By PCR amplification and sequencing of bisulfite-treated samples the amount of methylated cytosines at CpG, CpNpG, and asymmetrical sites in the centromeric 180-bp repetitive sequence family was measured in *A. thaliana* (Luo and Preuss 2003). Interestingly, a marked strand-biased DNA methylation was found and related probably to centromeric activity, as CpG and CpNpG methylation accounted for 100% in one strand and only 82 and 60 % respectively in the corresponding strands, while at asymmetrical sites the difference between strands was even more dramatic (89% and 21% respectively).

1. 2. 5. 3 Cytosine methylation: function

Recent data showed that DNA methylation in plants plays a role in many important processes, such as parental imprinting and plant development.

As previously mentioned, in tomato differences in DNA methylation were recorded between different tissues and highest amount of methylated cytosines (27%)

was found in seeds (Messeguer, Ganal et al. 1991). Finnegan *et al.* (1998) reported a similar observation in *Arabidopsis* seeds. The authors noticed that it could be due to an artefact: the ratio between nuclear and plastid genomes being higher in seeds compared to leaves and nuclear genomes being the only methylated-ones, or it could truly represent an increased DNA methylation in seeds. The authors proposed as well that DNA methylation could provide a possible mechanism for having a differential expression of genes belonging to maternal and paternal genomes during seed development. So far, evidence that DNA methylation could be involved in parental imprinting, observed in the endosperm of developing seeds but not in the embryo, is accumulating in maize and *Arabidopsis* (Alleman and Doctor 2000; Vinkenoog, Bushell et al. 2003).

DNA methylation is a key factor for plant development. As reviewed by Richards (1997) by using DNA methylation inhibitors, such as 5-azacytidine, developmental abnormalities (e.g. dwarfing, reduced fertility) were induced in rice, *Triticale*, flax and tobacco. Similarly, in *Arabidopsis* a reduced level of 5-methylcytosines through 5-azacytidine treatment, mutation or methyltransferase antisense constructs, induced loss of apical dominance, reduced stature, alteration in leaf size and shape, reduced root length and homeotic transformation of floral organs as a consequence of ectopic expression of genes normally regulated by methylation (Richards 1997; Finnegan, Genger et al. 1998). Finally, reduced fertility was observed as a consequence of decreased levels of methylated DNA and self-pollinated *Arabidopsis* plants with 70% reduction in DNA methylation became infertile after few generations (Finnegan, Peacock et al. 1996; Finnegan, Genger et al. 1998).

1. 2. 6 Histone modifications and the hypothesis of the "histone code"

As described in the previous section, nucleosomes are made of histone proteins that possess aminoacidic tail domains that can be modified thorough acetylation, methylation and phosphorylation, for example. Phosporylation of serines of histones H1 and H3 is related to chromosome condensation during mitosis and meiosis, while acetylation of specific lysine residues at histones H3 and H4 is due to histone acetyltransferases and is related to the process of trascription (Strahl and Allis 2000). Methylation of histones H3 and H4 at residues of arginine (which can be mono- and dimethylated) and/or lysine (which can be mono-, di-, and trimethylated) is due to

histone methyltransferases and is associated with different processes: methylation of histone H3 at Lysines 4 and 79 relates to gene expression, at Lysines 9 and 27 to heterochromatin formation, for example (Wang, Wysocka et al. 2004).

During DNA replication parental histones are randomly distributed between the two new duplicated DNA molecules and new histones are rapidly assembled into nucleosomes and brought close to replicating DNA for being included into higher order chromatin structures, then after DNA replication is completed histone tails are eventually modified where required, by using acetyl- and methyltransferases for example (Strahl and Allis 2000; Wang, Wysocka et al. 2004). In addition, while histone deacetylases can remove acetyl groups, no histone demethylases have been found and semiconservative nucleosome distribution during DNA replication seems to be the only way to decrease histone methylation levels (Jenuwein and Allis 2001). Finally, interphase or transcription-associated histone replacement can also occur to a small extent, despite the majority of histones being assembled into chromatin during DNA replication (Wang, Wysocka et al. 2004).

As the number of sites and the range of possible modifications are very high, the hypothesis of an "histone code" due to a combination of various covalent modifications at the histone tails (e.g. acetylation, phophorylation, methylation) read by a set of chromatin-associated proteins that activate different pathways and/or responses was recently proposed (Strahl and Allis 2000; Jenuwein and Allis 2001). According to this hypothesis, Jenuwein and Allis (2001) stated that "modifications on the same or different histone tails may be interdependent", leading to cooperative interactions or inhibiting effects, or to competition events. For example, methylation of histone H3 at Lysine 9 is inhibited by acetylation at Lysine 4 and phosphorylation at Serine 10 and competes with acetylation at the same residue, while phosporylation of histone H3 at Serine 10 is inhibited by methylation at Lysine 9 and favoured by acetylation at Lysine 9 and/or Lysine 14 (Jenuwein and Allis 2001; Fransz and de Jong 2002).

Jenuwein and Allis (2001) reported that "the concentration and combination of differentially modified nucleosomes" are related to many processes, such as DNA replication, DNA damage and repair, recombination, chromosome condensation and segregation, and mainly to the formation of euchromatic and heterochromatic regions that then affect gene expression and silencing respectively. According to the authors, the "histone code" contributes to "on" (euchromatic) and "off" (heterochromatic) epigenetic states during the embryogenesis and development, for example. Finally, as in plants

more lysine residues can be methylated at histone H3 and occasionally Lysine 20 at histone H4 can be acetylated instead of being methylated, the idea of a "plant dialect of histone code" with more different combinations of histone modifications was recently proposed in order to have a wider range of responses to environmental conditions and/or stress (Loidl 2004).

1. 2. 7 DNA methylation, histone modifications and chromatin remodelling

Data are accumulating about how the DNA methylation code is related to the "histone code" and to other chromatin remodelling processes with the aim to find evidence of interdependence between them in affecting chromatin structure and function. A link between DNA methylation and histone H3 methylation, showing that histone methylation regulates DNA methylation, was firstly suggested in the fungus *Neurospora crassa*, as the loss of histone methyltransferases DIM5, responsible for histone H3 methylation at Lysine 9, also resulted in loss of DNA methylation (Tamaru and Selker 2001). Similarly, in *Arabidopsis* the loss of histone methyltransferases KRYPTONITE also affected DNA methylation at CpNpG sites as shown at the SUPERMAN locus, a gene involved in flower development (Jackson, Lindroth et al. 2002). On the other hand, data supported the model that DNA methylation regulates histone methylation as defects in maintenance DNA methyltransferase MET1 also affected histone H3 methylation at Lysine 9 in *Arabidopsis* (Soppe, Jasencakova et al. 2002). In addition, a link between DNA methylation and histone deacetylation was proved in vertebrates as methyl-cytosine binding domain proteins were able to form a complex with histone deacetylases (Dobosy and Selker 2001) and maintenance DNA methyltransferases were shown to physically interact with histone deacetylases (Robertson, Ait-Si-Ali et al. 2000; Rountree, Bachman et al. 2000). Finally, the possibility that DNA methylation and histone methylation patterns are affected by chromatin remodelling was shown in *Arabidopsis* by isolating the DDM1 gene (Vongs, Kakutani et al. 1993) that encodes a chromatin remodelling protein, member of SWI2/SNF2-like ATPase family. Mutations in DDM1 gene decreased genomic DNA methylation by 70% (Kakutani, Jeddeloh et al. 1995; Jeddeloh, Stokes et al. 1999) and affected histone H3 methylation patterns (Gendrel, Lippman et al. 2002).

Fransz and de Jong (2002) proposed the model of "chromatin dynamic" whereby in the presence of chromatin-remodeling complexes the open state of chromatin is associated with histone acetylation and permits transcription, while the more compact state of chromatin is associated with histone methylation leading to the formation of heterochromatin. In this respect, the authors noticed that plants differ from animals as epigenetic switches occur not only during embryogenesis and early development, but also at later stages, including the transition from vegetative to reproductive growth (with formation of floral organs) and the root formation directly from stem or leaves. Similarly, Jaenish and Bird (2003) stated that the switch between active (open) and silent (closed) chromatin is due to many factors and processes that "reinforce or inhibit themselves in a controlled and stable mechanism that unlikely will occur by chance" and "allow the genetic information of an organism to be differentially expressed in both time and space" during development and differentiation, and/or in response to environmental changes, stress or disease (Jaenisch and Bird 2003).

1. 2. 8 RNA-directed DNA Methylation and RNA Interference

A further level of complexity in establishing and maintaining DNA and histone methylation patterns is added by two silencing mechanisms involving RNA molecules: the RNA-directed DNA methylation and the RNA Interference. In RNA-directed DNA Methylation (RdDM) a trigger RNA can target a homologous DNA sequence forming a RNA-DNA heteroduplex that becomes a substrate for DNA methyltransferases and is then methylated (Bender 2004). On the contrary, RNA Interference (RNAi) is a eukaryotic defence system against viruses that acts through a double-stranded RNA intermediate (Bender 2001; Matzke, Matzke et al. 2001; Aufsatz, Mette et al. 2002). In fact, long double-stranded RNAs (dsRNAs) are digested into smaller molecules of about 25 bp in length (by a ribonuclease called "Dicer") and these resulting small interfering RNAs (siRNAs) or microRNAs (miRNAs) are then used by a multiprotein called RISC (RNA-induced silencing complex) or RITS (RNA-induced initiation of transcriptional gene silencing) to be guided to the RNA or DNA sequences respectively that show identity with the dsRNA. RNA Interference is a two-step process that not only digests dsRNA molecules but also targets the RNA or DNA sequences that are similar to them and seems to be a trigger for DNA methylation.

According to Bender (2004) it is not only RNA viruses and transgenes that produce signals for DNA methylation through RdDM and RNAi mechanisms but also endogenous sequences. In addition, evidence that RNA Interference in plants is a key mechanism for the silencing of DNA repeats and mobile DNA elements and maintenance of heterochromatin has been proposed (Madlung and Comai 2004; Chan, Henderson et al. 2005). Recently, a complex model was reported about how RNA Interference in *A. thaliana* can induce *de novo* DNA methylation by Domain Rearranged Methyltrasferases (DRMs) followed by the formation of a complex made of methyl-cytosine binding proteins (MBDs) and histone deacetylase (HDAC), that then result in deacetylation of histone H3 at Lysine 9 and subsequent histone H3 methylation at Lysine 9 (Tariq and Paszkowski 2004).

1. 3 Polyploidy

In this section the occurrence of polyploidy and the mechanisms to generate polyploid genomes in plants and animals are discussed. Its role as a major force in plant evolution is considered, and among polyploid crop species special attention is focused on the wheat allopolyploid complex. The occurrence of phenotypic effects, immediate effects at genomic and genetical level, and long-term genomic and genetic consequences of polyploidization events are discussed. A diversity of phenomena associated with polyploidy as consequence of the occurrence of duplicated genomes and duplicated genes are described, together with examples of nuclear-cytoplasm interactions and interactions between duplicated genes.

1. 3. 1 **Definition and mechanisms**

Polyploids have three or more complete sets of chromosomes in somatic and germline cells instead of the two present in diploid species (Stebbins 1950). Polyploidy has been defined as a "special class of mutation" that occurs through genomic doubling or gametic nonreduction (unreduced eggs in plants and animals, and unreduced pollen in plants, as unreduced sperm is not common in animals), involving a failure in cell division during mitosis or meiosis respectively (Otto and Whitton 2000). As shown in fig. 1. 3 it results in the duplication of a single complete genome (autopolyploidy) or in the combination of two or more genomes (allopolyploidy) (Stebbins 1950). Polyploid

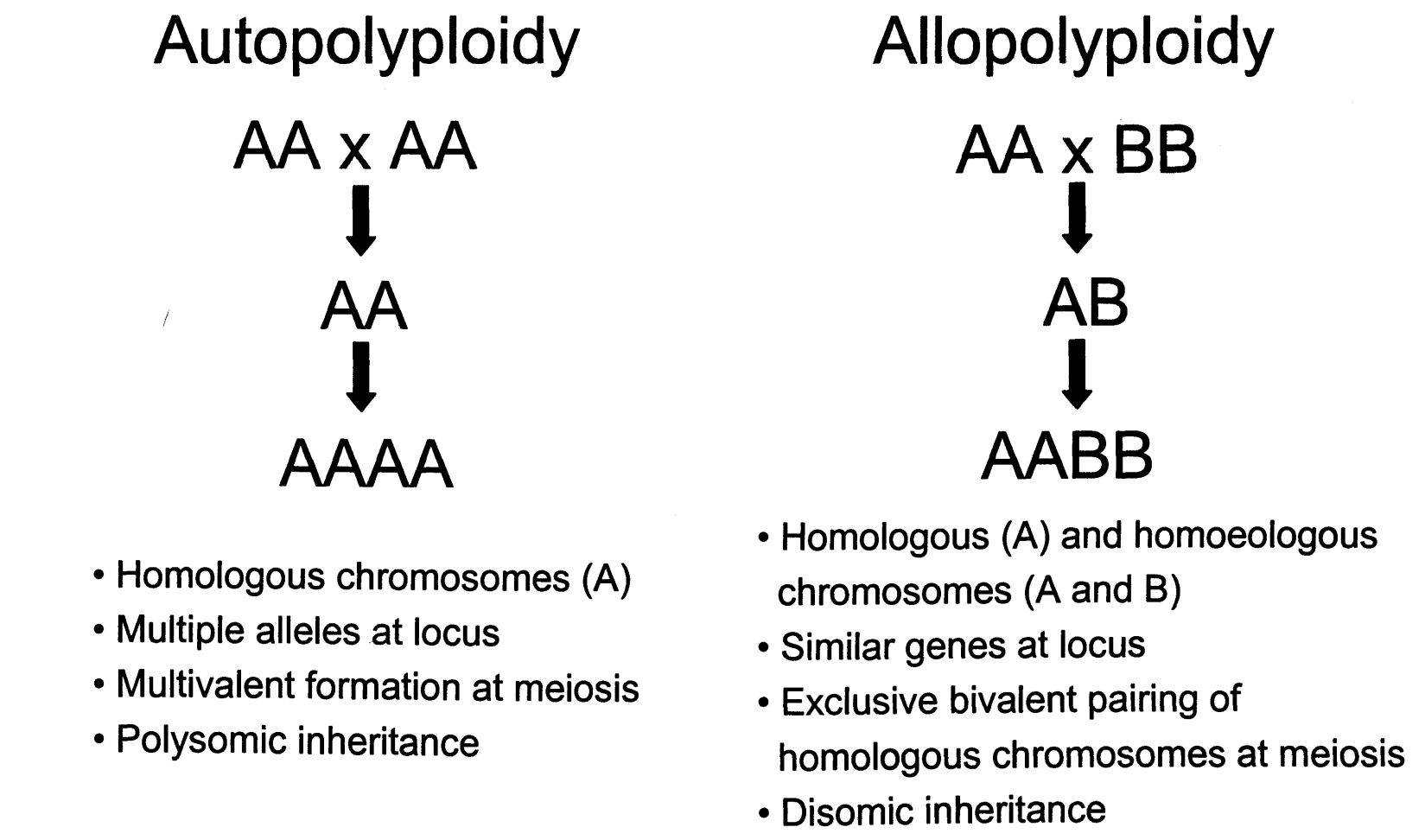


Figure 1. 3. Polyploidy formation is due to the duplication of a single complete genome (autopolyploidy) or to the combination of two or more genomes (allopolyploidy). Differences between auto- and allopolyploidy are presented.

formation was originally thought to be a single event leading to the formation of a new species that was genetically uniform. In this respect, a "genetic bottleneck" was supposed to happen as a consequence of the limited number of alleles present in the new species (Stebbins 1971). More recently many examples of polyploid plant species with multiple or recurrent origins have been described (Soltis and Soltis 1995; Soltis and Soltis 1999). Thus, a new auto- or allopolyploid species originated from multiple events from many and different individuals (parental genotypes) results in a group of polyploid genotypes with high levels of genetic diversity (Soltis and Soltis 1995; Soltis and Soltis 1999; Soltis and Soltis 2000).

Polyploidy is differently distributed in the plant and animal kingdoms: while it is a widespread event in plants, it is rare in animals, probably due to the presence of sex chromosomes and developmental constraints, and/or because interspecific hybridization is not a common event in animals (Otto and Whitton 2000). However, some polyploid species have been found in insects and vertebrates (fishes, amphibians, mammals) and two events of genomic doubling were supposed to happen during the evolutionary history of vertebrates (Ohno 1970).

1. 3. 2 **Estimates in plants and polyploidy as major force in evolution**

The widespread occurrence of polyploidy in plant indicates its role as a major force in evolution and speciation. It was reported that 70% of angiosperm species have undergone one or more events of polyploidization during their evolutionary history (Masterson 1994) and the amount of polyploids in ferns may be even higher, up to 95%-99% (Love, Love et al. 1977; Grant 1981). Through polyploidy certain species have generated high chromosome numbers (2n more than 1000) with the genetic consequence that genes are assorted in more combinations at meiosis (Heslop-Harrison 2000).

Ancient polyplodization events are difficult to detect because of chromosomal rearrangements and gene diversification for example, and other evolutionary changes following polyploidization (Wendel 2000). In fact, species traditionally considered as diploids are shown to be instead ancient polyploids or paleopolyploids, such as maize (Gaut and Doebley 1997) and *A. thaliana* (Grant, Cregan et al. 2000), as many duplicated regions have been found in their genomes. The hypothesis that the evolution of angiosperm genomes since their first appearance in the Cretaceous occurred through

cycles of polyploidization events followed by diploidization processes to restore a diploid-like behaviour seems to be likely (Wendel 2000). On the other hand, polyploidy is still an ongoing evolutionary process: among grasses, *Spartina anglica* is an example of a new allotetraploid species that appeared at the end of 19th century on the coasts of England and France (Baumel, Ainouche et al. 2001) and within the genus *Tragopogon* there are examples of recent and recurrent origin of polyploids, in the case of two different allopolyploids originated by the combination of three diploid parental species (Soltis, Soltis et al. 2004).

Many important crop plants are polyploids including tobacco, cotton, coffee, sugarcane, oat and wheat as they have undergone intentional hybridization and/or selective breeding (Otto and Whitton 2000; Wendel 2000). The wheat allopolyploid complex is a good model for evolutionary studies about polyploid formation as it includes diploid, tetraploid and hexaploid species as shown in fig. 1. 4. The diploid progenitors of wheat may have diverged from a common ancestor about 2.5-4.5 Mya and still have high level of synteny and homology (Huang, Sirikhachornkit et al. 2002). About 0.5 Mya, allotetraploid wheat *T. turgidum* ssp. *dicoccoides* (AB genome) arose from the hybridization of diploid *T. urartu* (A genome) with a diploid B genome-donor (an unknown species similar to *Aegilops speltoides*), followed by a chromosome doubling event, in geographical region of the Near- and Central-East called the Fertile Crescent, and it was domesticated about 10,500 years ago (Ozkan, Brandolini et al. 2005). Then, a second event of hybridization and chromosome doubling between domesticated allotetraploid wheat *T. turgidum* ssp. *durum* (AB genome) and diploid *T. tauschii* (D genome) gave rise to hexaploid wheat *T. aestivum* (ABD genomes) about 9,500 years ago (Salamini, Ozkan et al. 2002). At present, numerous other natural allopolyploids obtained by combining different species of *Aegilops* and *Triticum* enrich this model of study. Another interesting case of polyploidy is cotton: five allotetrapolyploid species of *Gossypium* originated about 1.5 Mya in the Americas, after the Asiatic A genome hybridized with the native D genome (Adams and Wendel 2004). The original diploid parents of *Gossypium* allopolyploids are extinct so that comparisons are only possible with modern diploid A and D genomes which represent the closest species to the ancient diploid donors. However, the allotetraploid species made of extinct A and D genomes showed additive genome size (2C=5.8 pg) when compared with the modern A and D diploid species with 2C=3.8 pg and 2 pg, respectively (Wendel 2000).

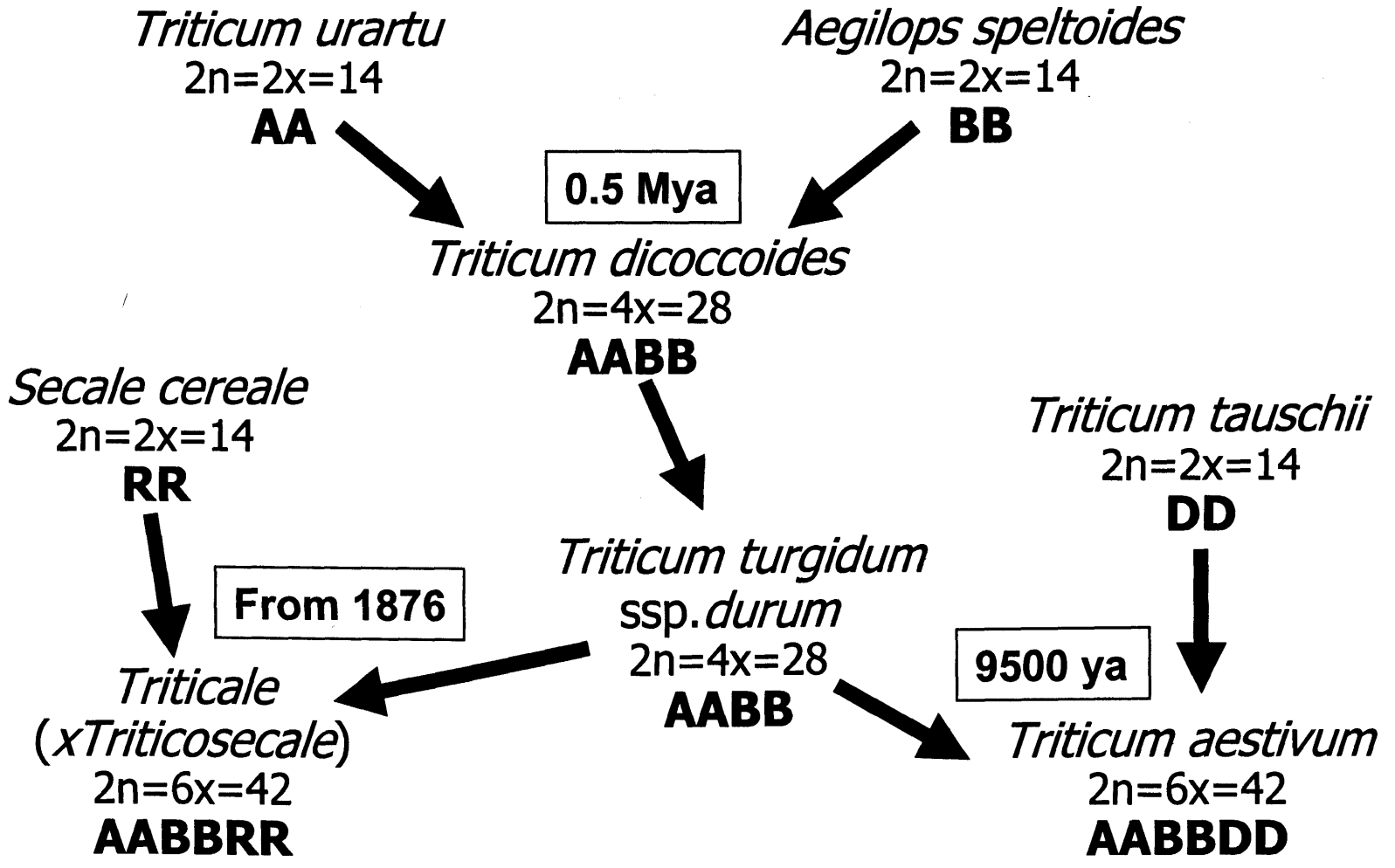


Figure 1. 4. Wheat evolution and *Triticeae* allopolyploid formation based on Ozkan *et al.* 2005, Salamini *et al.* 2002 and Wilson 1876.

1. 3. 3 **Polyploidy consequences**

1. 3. 3. 1 **Phenotypic effects**

An increase in cell volume is a common feature of polyploids. In fact cell size of stomata in fossil plants was used for an indirect estimate of polyploidy events in angiosperms (Masterson 1994). Changes in cell volume can affect metabolic and developmental processes, however changes at the cellular level can vary according to different species and/or environmental conditions (Otto and Whitton 2000). In addition, the increase in cell volume can affect size and shape of organs and size of whole organisms, as shown sometimes but not always in polyploid species of plants and insects (Otto and Whitton 2000).

1. 3. 3. 2 **Genomic and genetic changes**

According to McClintock (1984) "species crosses are a potent source of genomic modification" and bringing together two genomes into the same nucleus can be considered a source of stress. She defined "genomic shock" a pre-programmed response to stress following the interspecific hybridization event and concluded that "genomic modifications of some type would accompany the formation of new species", such as breakage-fusion-bridges cycles and mobile DNA element activation, in order to reorganize and restructure the genome of a newly formed allopolyploid species. More recently, Soltis and Soltis (1995, 2000) suggested that polyploids have extremely plastic genomes able to undergo many changes from molecular level to population level and defined them "dynamic polyploids".

In order to minimize problems and maximize the benefits of a polyploidization event, immediate and long-term effects of polyploidy have been observed as rapidity is a need for surviving and establishing as new species, while long-term modifications allow the new species to develop benefits from the polyplodization event.

As soon as a new polyploid species is formed immediate effects also called "revolutionary changes" (Levy and Feldman 2004) are necessary to initiate the process of diploidization, especially in allopolyploids. At the cytological level, autopolyploids can have multivalent formation between homologous chromosomes at meiosis and polysomic inheritance as shown in fig. 1. 3. In contrast, for being fertile allopolyploids

need to recognize the differences between homeologous chromosomes (genetically related chromosomes of different genomes) and prevent their pairing in order to have a diploid-like behaviour at meiosis with exclusive bivalent pairing of homologous chromosomes and in order to have disomic inheritance at each locus as shown in fig. 1. 3. At genetical level, a new polyploid species faces the need to coordinate and balance the expression of two sets of genes to avoid over-expression and inefficiency. In fact, a polyploidization event involves the simultaneous duplication of all nuclear genes, regulatory and structural, originating multiple alleles per locus (in autopolyploids) or different alleles from similar genes per locus (in allopolyploids) as shown in fig. 1. 3. Two different genomes not always contribute equally to the transcriptome of a new polyploid species but more likely a preferential transcription of only one of the two genomes has been observed in polyploids. Furthermore, in allopolyploids it is necessary to minimize deleterious and negative interactions that possibly arise between genes belonging to two different genomes. In fact, Comai *et al.* (2003) noticed that despite some divergence, some protein subunits can still form dimers in allopolyploids, while other orthologous proteins may diverge rapidly in the parental donors and are then impaired and/or form heterodimers that do not function properly in allopolyploids. Finally, in polyploids intergenomic coordination is essential and achieved through a range of mechanisms such as chromosomal repatterning, amplification and reduction of repetitive sequences, for example, in order to synchronize different genomes during the cell cycle.

On the other hand, long-term effects of polyploidy also called "evolutionary changes" (Levy and Feldman 2004) are related to the opportunity of using raw genetic material to acquire new functions or new morphological traits not present in the parental species and being more successful in colonizing new ecological niches and environments.

Many polyploid species have been investigated such as *Nicotiana* (tobacco), *Spartina* and *Tragopogon*, but the majority of studies have concentrated on *Brassica*, *Gossypium* (cotton) and *Triticum/Aegilops* (wheat), because they are major crop species and it is easy to reproduce natural and non-natural genomic combinations in newly synthesized (synthetic) allopolyploids by crossing different parental species and by using colchicine for chromosome doubling. Further generations are then obtained by self-pollination. By analysing newly synthesized allopolyploids it is possible to detect genetic and genomic changes occurred soon after the hybridization or during the first

generations, because such changes are still not hidden by the evolution as happens in the case of established polyploid species. In synthetic allopolyploids the null hypothesis of genomic additivity is that both parental genomes contribute equally to a new allopolyploid species (Wendel 2000; Adams and Wendel 2004). Any deviation from additivity is an indication of phenomena associated with polyploidy in the attempt to coordinate different genomes, to gain stability and to evolve quickly.

In table 1. 2 revolutionary and evolutionary effects of polyploidy, especially of allopolyploidy, are presented. Mechanisms contributing to immediate genomic and genetic changes such as chromosomal repatterning, deletion of low-copy noncoding sequences, amplification or reduction of repetitive sequences and gene loss; to long-term genomic and genetic changes such as intergenomic colonization, intergenomic horizontal transfer, gene inactivation, and functional diversification of genes; together with nuclear-cytoplasm interactions and mechanisms involved in the interactions between duplicated genes (dosage compensation, concerted evolution, alteration of regulatory network) will be described here, while epigenetic mechanisms will be discussed later.

1. 3. 3. 3 **Changes in duplicated genomes**

a) Chromosomal repatterning (translocations)

In contrast to diploid hybrids that can largely and freely recombine their parental genomes as shown in sunflower (Ungerer, Baird et al. 1998), homeologous recombination is rare in established allopolyploids, but in newly formed allopolyploids it can contribute to genome restructuring (Comai 2000).

Intergenomic chromosome translocations that can occur in polyploid species are divided into two groups: "random translocations" involving different chromosomes in different populations of the same polyploid species and "species-specific translocations" involving always the same chromosomes in every populations of the same polyploid species (Jiang and Gill 1994). Among cereals, species-specific translocations were found in two tetraploid wheats: in *T. timophevii* involving chromosomes 6At, 1G and 4G and *T. turgidum* involving chromosomes 4A, 5A and 7B (Jiang and Gill 1994).

Controversial data about chromosome translocations have been reported in *Brassica* allopolyploids by two research groups (Song, Lu et al. 1995; Axelsson, Bowman et al. 2000). In *B. juncea* (AB genome) and in *B. napus* (AC genome) a

Table 1. 2. Immediate and long-term genetic and epigenetic effect of polyploidy (modified from Levy and Feldman, 2004).

Revolutionary changes
Genetic changes
Chromosomal repatterning
Deletion of low-copy noncoding sequences
Amplification and reduction of repetitive sequences
Gene loss
Epigenetic changes
Methylation repatterning
Mobile element activation
Alteration in gene expression due to transcriptional activation of retroelements
Gene silencing
Gene activation
Nucleolar dominance
Evolutionary changes
Genetic changes
Intergenomic colonization
Intergenomic horizontal transfer
Gene inactivation
Functional diversification of duplicated genes
Epigenetic changes
Functional diversification of duplicated genes

comparison between F2 and F5 synthetic allopolyploids and their parental diploid species *B. rapa* (A genome), *B. nigra* (B genome) and *B. oleracea* (C genome) was performed by using RFLP patterns and a wide range of probes comprising anonymous nuclear clones, cDNA clones, and nuclear genes of known functions (Song, Lu et al. 1995). Loss or gain of parental fragments was observed in higher frequency in AB than in AC *Brassica* synthetic allopolyploids suggesting the frequency of changes to be related to the divergence of parental diploid genomes: A and C genomes being more similar than A and B genomes. To explain these findings the authors proposed chromosome rearrangements due to intergenomic recombination at meiosis, as they found a high number of aberrant meiosis with chromosome bridges, chromosome lagging and multivalents in F2, F3 and F5 generations of *Brassica* synthetic allopolyploids. In contrast, by using 120 probes selected from *Brassica* genomic libraries (anonymous genomic DNA clones) for a wide RFLP study Axelsson *et al.* (2000) did not confirm the occurrence of rapid genomic changes showing on the contrary that linkage groups of natural and newly synthesized allotetraploid *B. juncea* (AB genome) were perfectly collinear with the linkage groups of diploid parental species *B. rapa* (A genome) and *B. nigra* (B genome). The authors concluded that the occurrence of chromosome rearrangements must be limited in number and size suggesting instead that the rapid genomic changes observed by Song (1995) might be related to processes other than homoeologous recombination.

By using genomic *in situ* hybridization technique (GISH.) inter-genomic chromosomal exchanges have been observed also in tobacco and among grasses in oats and synthetic hybrids between *Lolium multiflorum* and *Festuca* species, as reviewed by Wendel (2000). However, according to the author it is not probable that translocations, inversions or other types of chromosomal rearrangements can contribute significantly to diploidization and stabilization of allopolyploids as their frequency is not significantly different from the frequency of structural changes occurring in diploid species.

b) Deletion of low-copy noncoding DNA sequences

The existence of rapid and nonrandom elimination of low-copy noncoding DNA sequences during allopolyploid formation seems to be a key mechanism in *Triticeae* allopolyploids leading to an immediate diploidization process at cytological level, while the *Ph*1 and *Ph*2 genes have been proposed to evolve later in *Triticeae* species as a second system to suppress homeologous chromosome pairing (Feldman, Liu et al. 1997;

Liu, Vega et al. 1998). Early observations obtained by RFLP analysis showed that low-copy noncoding chromosome-specific (CSSs) and genome-specific (GSSs) sequences were present in all diploid progenitors but were found only in one genome of natural allotetraploid and allohexaploid wheats indicating elimination from the other genomes (Feldman, Liu et al. 1997). The same analyses were extended to a larger number of newly synthesized allopolyploids obtained by crossing *Triticum* and *Aegilops* species showing a rapid (occurring during the early generations after the hybridization event) and non-random (occurring from one parental genome only) elimination of sequences (Feldman, Liu et al. 1997; Liu, Vega et al. 1998). Ozkan *et al.* (2001) tested by RFLP analysis newly synthesized *Aegilops-Triticum* allopolyploids representing natural and nonnatural genomic combinations observing that genome-specific sequences (GSSs) were often eliminated before than chromosome-specific sequences (CSSs) and that sequence elimination was not related to ploidy level of parental species, divergence between parental genomes, type and origin (maternal or paternal) of the sequences and chromosomal location. In contrast, the phenomenon was related to genomic combination: in synthetic allopolyploids whose genomic combination was analogous to the natural ones elimination started earlier and was more rapid than in the ones representing nonnatural genomic combinations. Shaked *et al.* (2001) used a different approach to assess genomic changes in newly synthesized *Aegilops-Triticum* allopolyploids. By AFLP technique 3661 loci of unknown copy number and function were compared between newly synthesized *Aegilops-Triticum* allopolyploids and their parental species showing that among the polymorphic bands (the ones present in only one parental genome) 66 bands out of 619 (10.6%) were eliminated in the newly synthesized allopolyploids. The authors proposed a list of mechanisms that possibly were responsible of sequence elimination, including processes of excision and subsequent loss of a sequence, and recombination events. Similar results of fragment disappearance and deviation from additivity were recently obtained, by RFLP analysis and by using expressed sequence tags (ESTs) and retrotransposons as probes in Southern hybridizations, not only in newly synthesized *Aegilops-Triticum* allopolyploids, but also in newly synthesized *Triticum* hybrids and in newly synthesized *Triticale* species (Han, Fedak et al. 2003).

c) Changes in repetitive sequences and mobile elements

Amplification and reduction of repetitive sequences are likely to happen as a consequence of a polyplodization event. Elimination of repeats during allopolyploidization in cultivated tetraploid and hexaploid wheats has been observed (Petsova, Goncharov et al. 1998). The authors measured the amount in copy number of Spelt1, a tandem repetitive sequence associated with telomeric regions of *A. speltoides*, in about 300 species and accessions of diploid and polyploid *Triticum* and *Aegilops* species. While in *A. speltoides* accessions the copy number was in the range of 1.5-5.3 $x10^5$, in *T. durum* and *T. aestivum* it decreased to the range of 10^2-$1.2x10^4$, but not in two endemic species of tetraploid wheats *T. timopheevii* or *T. carthlicum* where it remained unaltered.

Repetitive sequences and mobile elements have been shown to move from one genome to the other in polyploid species. Intergenomic colonization by repetitive sequences from one of the two parental diploid genomes of both chromosome sets of allotetraploid *T. dicoccoides* have been described (Belyayev, Raskina et al. 2000). By using genomic *in situ* hybridization technique (GISH.) the authors showed that the A-genome chromosomes from *T. uratu* carried heterochromatic blocks from the B-genome of *A. speltoides* in allotetraploid *T. dicoccoides* (AB genome). To explain the occurrence of B-genome satellite DNA sequences in A-genome heterochromatin of allotetraploid *T. dicoccoides* (AB genome) the authors proposed mechanisms of interlocus concerted evolution and/or intergenomic colonization in order to stabilize the allotetraploid wheat genome. On the other hand, in natural allotetraploid cotton *G. hirsutum* (AD genome), the occurrence of intergenomic horizontal transfer was reported due to the colonization of D genome by A genome-specific mobile sequences after the allopolyploidization event (Hanson, Zhao et al. 1998; Zhao, Si et al. 1998; Hanson, Islam-Faridi et al. 2000). In fact, about 20 out of 83 dispersed repetitive families, made of transposons, were localized by *in situ* hybridization (ISH) in allotetraploid *G. hirsutum* (AD genome) and most families that at diploid level seemed to be specific for A genome, were found in D-genome as well as in allotetraploid cotton (Zhao, Si et al. 1998). Similarly, it was reported that Ty1 *copia*-like retrotransposons have been "horizontally" transferred between genomes, as they hybridized on both genomes of allotetraploid *G. hirsutum* (AD genome) but failed to have signal on chromosomes of diploid D genome (Hanson, Zhao et al. 1998; Hanson, Islam-Faridi et al. 2000). Wendel (2000) suggested that intergenomic horizontal transfer is likely to have a great impact not only in cotton allopolyploids, but more widely in polyploid evolution, due to the

ubiquity and abundance of retroelements in plants and due to the occurrence of mobile DNA element activation mechanisms following a polyplodization event as will be described later.

1. 3. 3. 4 **Nuclear-cytoplasm interactions**

As previously mentioned, in polyploids two different genomes are suddenly brought into the same nucleus, but in only one of two parental cytoplasms and therefore with only one set of mitochondrial and chloroplast genes. Then, mechanisms of cytonuclear stabilization after a polyploidization event are necessary, especially in allopolyploids. The cytoplasmic effect proposes that the genome from the maternal parent which contributes to the nucleus and the cytoplasm of the new allopolyploid species does not need to change as much as the genome from the paternal parent which only contributes to the nucleus.

By investigating reciprocal hybrids of newly synthesized *Brassica* allopolyploids, the cytoplasmic effect was observed in *B. juncea* (AB and BA genomes) but not in *B. napus* (AC and CA genomes) whose diploid parental A and C genomes are quite similar (Song, Lu et al. 1995). By analysing RFLP changes of *B. juncea*, in the allopolyploid AB plants containing A genome cytoplasm the loss of B genome fragments was greater than A genome fragments, while in allopolyploid BA plants containing B genome cytoplasm the opposite phenomenon was observed. The data confirmed that the paternally donated genome needed to undergo more changes than the maternally donated genome in newly synthesized *B. juncea* allopolyploids. In contrast, no cytoplasmic effect was observed in other allopolyploid species such as cotton (Brubaker, Paterson et al. 1999) and wheat (Liu, Vega et al. 1998; Liu, Vega et al. 1998).

1. 3. 3. 5 **Changes in duplicated genes**

In a polyploid species genes present in multiple copies (duplicated genes) are unlikely to retain their original function. One of the duplicated copies can be lost, inactivated through mutations, insertions and deletions, or undergo functional diversification (Lynch and Conery 2000; Adams and Wendel 2005; Comai 2005).

By comparing a newly formed allotetraploid obtained by crossing *Aegilops-Triticum* species and its diploid donors, gene loss was mentioned as one of the mechanisms responsible of fragment disappearance when 3072 transcribed loci were analysed by cDNA-AFLP technique (Kashkush, Feldman et al. 2002). The authors suggested gene loss as an immediate consequence of polyploidization that could happen soon or after chromosome doubling, and reviewed some previous studies of gene loss in wheat species concluding that they all possibly derived from the same kind of mechanism. According to Adams and Wendel (2005) differential gene loss in polyploids can affect some duplicated genes but not others, and thus is one of the factors contributing to deviation from co-linearity often observed among related species, for example in cereals. In addition, the authors believed that gene loss is not a random process but subject to natural selection. Similarly, recent analysis about genome duplication in *A. thaliana* showed that genes involved in transcription and signal transduction were preferentially retained and the ones coding organellar proteins and involved in DNA repair were preferentially lost (Blanc and Wolfe 2004).

From the analysis of isoenzymes in plants reviewed by Wendel (2000) and genes coding chlorophyll-binding proteins in ferns (Pichersky, Soltis et al. 1990) "gene silencing" was first suggested to happen as a consequence of point mutations, insertions and deletions that produce defective non-functional genes. More recently, epigenetic mechanisms have been proposed to be responsible for gene silencing as will be described later.

Duplicated genes belonging to different genomes may not evolve at the same rate in a polyploid species, leading to functional divergence: neofunctionalization when one of the duplicated genes acquires a new function, or subfunctionalization when the ancestral function is partitioned between the two duplicated genes (Lynch and Conery 2000; Adams and Wendel 2005; Comai 2005). Functional divergence allows a new polyploid species to have higher biochemical and physiological flexibility, to develop new morphological traits and finally to be able to colonize new environments (Stebbins 1950; Wendel 2000). In allotetraploid wheat *T. turgidum* ssp. *durum* (AB genome) several studies using a diversity of markers showed the B genome having higher diversity respect to A genome possibly as a consequence of external forces such as differential selection and interspecific introgression, or internal mechanisms such as mutation and recombination (Levy and Feldman 1988; Felsenburg, Levy et al. 1991; Siedler, Messmer et al. 1994). In cotton *Gossypium* allopolyploids evolutionary rates

seemed to be faster in the D than in the A genome as shown by measuring nucleotide diversity in *Adh* (alcohol dehydrogenase) and FAD2-1 (fatty acid desaturase) genes, as reviewed in Adams and Wendel (2004).

1. 3. 3. 6 **Interaction between duplicated genes**

Examples of interactions between genes directly at DNA level will be include dosage compensation and increased variation in dosage-regulated genes, intergenomic homogenization and concerted evolution; and among interactions between their protein products, the alteration of regulatory networks (made of transcription factors).

In polyploids the transcript levels are expected to be directly proportional to the number of copies of a gene, whereas dosage compensation has been observed in some cases, such as in maize where the expression of 18 genes was measured along a euploidy series, from haploid to tetraploid (Guo, Davis et al. 1996). A model of interactions between genes having allele-dosage effects, such as genes responsible for plant architecture or flowering time, has been described (Osborn, Pires et al. 2003). The authors proposed that in polyploids the presence of duplicated genes having allele-dosage effects leads to an increase in variability of gene expression levels due to their different combinations, possibly resulting in a wide range of phenotypic effects.

Several mechanisms, such as gene conversion and other recombination events, are responsible for intergenomic homogenization and concerted evolution whereby different loci become homogenized instead of evolving independently. For example in diploid and allopolyploid cotton species the analysis of the internal transcribed spacer (ITS) regions revealed minimal heterogeneity between repeats at diploid and polyploid level showing that concerted evolution happened not only within but also between arrays and genomes in both directions as some cotton allopolyploid species had D-genome origin repeats while one had A-genome origin repeats (Wendel, Schnabel et al. 1995). Inter-locus homogenization of rDNA repeats has been reported in many other polyploid species, such as of *Nicotiana* and *Glycine* (Doyle, Doyle et al. 2004; Kovarik, Matyasek et al. 2004). Intergenomic homogenization and concerted evolution for single-copy nuclear genes is still an open question. Possibly, gene conversion happened in allotetraploid tobacco in a gene encoding glucan endo-1,3-β-D-glucosidase soon after polyploidization, as in two recombined cDNA clones alternating blocks of sequences that had similarity to both parental genes from the diploid progenitors were found

(Sperisen, Ryals et al. 1991). Conversely, in allopolyploid cotton no recombination events and gene conversion were reported among 16 different loci as reviewed in Wendel (2000).

Alteration of regulatory networks in polyploids is an interesting hypothesis proposed by Osborn *et al.* (2003). According to the authors two sets of transcription factors, which regulate the expression of similar genes in a diploid genome, can interact producing an overall alteration of gene expression in polyploids. The effect was thought to be more prominent in allopolyploids than in autopolyploids due to a high number of alleles coming from different species. The authors assumed that alteration of regulatory networks could affect developmental timing or tissue specificity, for example. Furthermore, they proposed alteration of regulatory networks to be strongly related to the phenomenon of heterosis or hybrid vigor, whereby the heterozygous individual performs much better than the two parental ones, possibly conferring selective advantages to polyploids.

1. 3. 4 A variety of genomic and genetic changes in polyploids and the quiescent polyploids

Many authors noticed that genomic and genetic consequences after polyploidization might show different occurrence, timing and extent between plant systems revealing in polyploidy species a variety of genetic as well as epigenetic mechanisms that will be presented later (Wendel 2000; Levy and Feldman 2004; Comai 2005).

Different mechanisms have been described to cause the phenomenon of rapid genomic changes in newly formed allopolyploids of *Aegilops-Triticum* and *Brassica* for example. A rapid and nonrandom elimination of sequences was a common response to polyploidy in *Triticeae* species (Feldman, Liu et al. 1997; Liu, Vega et al. 1998; Ozkan, Levy et al. 2001; Shaked, Kashkush et al. 2001; Han, Fedak et al. 2003). Conversely, in *Brassica* allopolyploids chromosomal rearrangements seemed more likely to happen as a consequence of polyploidy despite the question is still debated (Song, Lu et al. 1995; Axelsson, Bowman et al. 2000). On the other hand, cotton seems to be a "quiescent allopolyploid" in contrast to *Brassica* and *Triticum/Aegilops* allopolyploids. In fact, by comparing nine newly synthesized *Gossypium* allotetra- and allohexapolyploids with their parentals AFLP analysis revealed that of 12,400 bands fragment loss was rare and

there was no appearance of new fragments showing that both parental genomes contributed equally to the new allopolyploid species (Liu, Brubaker et al. 2001) .

In addition, some consequences of polyploidy did not appear ubiquitous phenomena, as in the case of cytoplasmic effect reported in some *Brassica* allopolyploids (Song, Lu et al. 1995), but not in other allopolyploid species such as cotton (Brubaker, Paterson et al. 1999) and wheat (Liu, Vega et al. 1998; Liu, Vega et al. 1998). According to Osborn (2003) "changes in gene expression that occur immediately in all co-derived allopolyploids might indicate altered regulatory networks, whereas changes that occur in only a portion of the allopolyploids could indicate genetic and epigenetic changes."

Many authors claimed the importance of investigating the occurrence of genomic and genetic changes in allopolyploids, as well as their timing and frequency, by comparing natural and newly synthesized allopolyploids, in order to evaluate how widespread these changes are and to understand their evolutionary impact in new polyploids (Leitch and Bennett 1997; Soltis and Soltis 2000; Osborn, Pires et al. 2003).

1. 4 Epigenetic phenomena and Polyploidy

In this section epigenetic phenomena contributing to immediate and long-term effects on entire genomes and single genes in polyploids, especially allopolyploid species, are discussed. Genome-wide methylation repatterning, largely observed in *Brassica*, *Aegilops-Triticum* and *Arabidopsis-Cardaminopsis* synthetic allopolyploids, and mobile element activation are described. Examples of silencing of protein coding genes and ribosomal DNA genes (nucleolar dominance) are reported. In addition, the causes that directly or indirectly move epigenetic modifications are described. Furthermore, the links and interdependence between DNA methylation, histone modifications, chromatin remodelling and RNA interference in relation to the formation of a new allopolyploid species are considered. Finally, the role and the consequences of epigenetic phenomena in establishing polyploids, especially allopolyploid species, and in term of their evolution are discussed.

1. 4. 1 **Epigenetic phenomena in polyploids**

Epigenetic mechanisms can favor genomic and genetic diploidization, and intergenomic coordination overall contributing to genome stability of a new polyploid species through immediate effects such as genome-wide methylation repatterning and mechanisms of de-repression of dormant transposable elements (mobile DNA element activation) as reported in table 1. 2. Similarly, immediate changes to genes are mainly related to alteration of gene expression due to transcriptional activation of retroelements, silencing of protein coding genes and ribosomal DNA genes (nucleolar dominance), gene activation, while on a long-term scale heritable epigenetic alterations can lead to functional diversification of duplicated genes as reported in table 1. 2. In addition, it is known that methylated cytosines undergo spontaneous deamination to thymines at high frequency (Gonzalgo and Jones 1997) and thus methylation changes can originate novel genetic mutations in this way.

Interestingly, epigenetic phenomena do not represent permanent changes in a new polyploid species. Methylation repatterning and mobile element activation due to DNA methylation and/or chromatin remodelling processes lead to genomic reversible effects in comparison with the events of chromosomal repatterning and sequence elimination described earlier. Similarly, gene silencing and gene activation through epigenetic mechanisms are more effective and flexible in comparison to genetic mutations and allow the switching of genes on and off as necessary in a new polyploid species (Finnegan 2001; Lee and Chen 2001).

1. 4. 1. 1 **DNA methylation repatterning**

Genome-wide DNA methylation repatterning through mechanisms of *de novo* methylation and demethylation has been reported in *Brassica*, *Triticeae* and *Arabidopsis-Cardaminopsis* synthetic allopolyploids affecting the coding and noncoding fraction of their genomes.

Early evidence of DNA methylation changes at CpG and CpNpG sites were reported in *Brassica* allopolyploids and associated with the loss or gain of parental fragments found by analysing RFLP patterns obtained with methylation-sensitive restriction enzymes and a wide range of probes comprising anonymous nuclear clones and cDNA clones (Song, Lu et al. 1995).

Similarly, among cereals DNA methylation changes affecting repetitive and low-copy DNA sequences have been described by comparing newly synthesized

Triticeae allopolyploids and their parental species by using RFLP analysis with methylation-sensitive restriction enzymes together with genomic and cDNA sequences as probes (Liu, Vega et al. 1998). In a following study, alteration of cytosine methylation at CpG sites was proved to be a consequence of interspecific and intergeneric hybridization, between *Aegilops* species and *Aegilops-Triticum* species respectively (Shaked, Kashkush et al. 2001). Scoring the overall bands produced by MSAP analysis the authors found deviation from additivity in about 13% of genomic DNA loci examined. After isolation, cloning and sequencing of 12 MASP-isolated fragments that had shown DNA methylation alteration, most were similar to repetitive DNA sequences and highly repeated retrotransposons, while others having low-copy number did not show significant similarity to known sequences. In conclusion, alteration in cytosine methylation was shown to be a rapid response of the genome after a hybridization event. It could equally occurs in the F1 hybrid or in the newly formed allotetraploid (after chromosome doubling), affect both repetitive and low-copy DNA sequences, and differ in frequency between the two constituent genomes (for example, among 11 bands that showed DNA methylation changes, 10 came from one parent and only 1 from the other). Finally, DNA methylation changes were shown in newly synthesized *Aegilops-Triticum* hybrids, newly synthesized *Triticum* hybrids and newly synthesized *Triticale* hybrids by RFLP analysis with methylation-sensitive restriction enzymes and a set of expressed sequence tags (ESTs) and retrotransposons as probes (Han, Fedak et al. 2003). By using *Msp*I-*Hpa*II restriction enzymes deviation from additivity between newly formed allopolyploids and their parentals was observed at CpG sites of ESTs sequences, but not at CpNpG sites, while retrotransposons were always heavily methylated without showing major changes in their DNA methylation status.

In *Arabidopsis-Cardaminopsis* allopolyploids MSAP analysis was conducted on F3 individuals (Madlung, Masuelli et al. 2002) derived from a F2 plant that showed phenotypic variation (morphology, flowering time and fertility) and instability (Comai, Tyagi et al. 2000). Deviation from additivity was found in 52 out of 623 products corresponding to about 8%. More in detail, in 12 cases methylation polymorphisms were variable between F3 individuals and no further analyses were performed, while in 40 cases methylation polymorphisms were reproducible and changes in DNA methylation were attributed to hypermethylation or hypomethylation. After gel elution, re-amplification and sequencing it was shown that polymorphic bands mainly

represented genes, suggesting not only silencing but also activation through DNA methylation increases or decreases. In contrast, no overall hyper- or hypomethylation was observed in F3 individuals: by comparing percent methylation of *Taq*I sites (TCGA) no gross changes in overall CG methylation were observed and only a relocation of CG methylation suggested.

Conversely, a set of nine newly formed *Gossypium* allopolyploids showed complete additivity after the analysis of about 9700 bands obtained by the MSAP technique (Liu, Brubaker et al. 2001). Furthermore, on a subset of six *Gossypium* synthetic allopolyploids the authors performed RFLP analysis by using *Msp*I-*Hpa*II restriction enzymes and retrotransposons, mainly *copia*-like retroelements, as probes: complete additivity and no DNA methylation changes were observed. As mentioned in the previous section, cotton seems to be a quiescent polyploid species conversely to *Brassica*, *Triticeae* and *Arabidopsis-Cardaminopsis* allopolyploids, suggesting genomic changes and methylation repatterning are not always associated with allopolyploid formation.

1. 4. 1. 2 **Mobile DNA element activation**

In plants, mobile genetic elements are a source of genetic diversity: they can be activated and transpose as a response to stress and environmental conditions, for example, with a wide range of effects in host genomes (Wessler 1996; Grandbastien 1998; Wendel and Wessler 2000). Mc Clintock (1984) hypothesised that an interspecific hybridization event is a cause of "genomic shock", a programmed response to stress, that leads to mobile DNA element activation for example. More recently, it has been predicted that in polyploid species possible deleterious insertions (e.g. into genes) after mobile DNA element activation are minimized by the presence of two genomes and thus by gene redundancy (Matzke, Mittelsten Scheid et al. 1999; Wendel 2000). Finally, many authors agreed with McClintock's prediction that mobile DNA element activation is a consequence of allopolyploidization, possibly differing in extent and timing according to different species (Comai 2000; Liu and Wendel 2003).

Recently, examples of mobile DNA element activation were reported in *Triticeae* and *Arabidopsis-Cardaminopsis* allopolyploids. Among cereals, evidence of mobile DNA element activation was found as 12 new transcripts, showing similarity mainly to retroelements which appeared in a newly formed *Aegilops-Triticum*

allotetraploid species when 3072 transcribed loci from leaves and root tips were analysed by cDNA-AFLP technique (Kashkush, Feldman et al. 2002). In addition, activation of retrotransposons could have the consequence of altering the expression of adjacent genes leading to gene silencing or gene activation as reported in wheat (Kashkush, Feldman et al. 2003). In *Arabidopsis-Cardaminopsis* allopolyploids, a novel family of transposons, isolated in *A. thaliana* from a heterochromatic region of chromosome 4, seemed to be transcriptionally activated when performing microarray analysis with probes prepared from parental and allopolyploid mRNAs (Madlung, Tyagi et al. 2005). Furthermore, by RFLP analysis with methylation-sensitive restriction enzymes the authors showed that such transcriptional activation was a consequence of demethylation mechanisms.

1. 4. 1. 3 **Silencing of protein coding genes**

Few examples of epigenetic instability involving silencing or activation of genes have been reported for autopolyploids, while it was largely documented in synthetic and natural allopolyploids of *Triticeae* and *Arabidopsis-Cardaminopsis* by cDNA-AFLP analysis.

Among *Triticeae* allopolyploids, 48 out of 3,072 transcripts (1.6%), mainly involving genes but also retrotransposons, disappeared in newly synthesized *Aegilops-Triticum* allotetraploid compared to its diploid parents; detailed analyses were performed on 12 transcripts and in 4 cases this phenomenon was associated with an alteration in cytosine methylation at CpG sites (Kashkush, Feldman et al. 2002). Similarly, in synthetic and natural hexaploid wheat *T. aestivum* down-regulation of gene expression was reported (He, Friebe et al. 2003). According to the authors gene deletions were not the cause of such alteration of gene expression, involving in some genes both parental orthologues (aspecifically), while in other genes showing to be a non-random event as it happened in an orthologue-dependent manner.

In *Arabidopsis-Cardaminopsis* allopolyploids, phenotypic variation (morphology, flowering time and fertility) and instability were observed in individuals of F2 generation and related to rapid gene silencing. In fact, loss (in the majority of cases) or gain (only in two cases) of bands has been observed by comparing the allopolyploids with their parental plants. Silencing was confirmed in 3 out of 700 transcripts (about 0.4%) corresponding to a transposon and two genes (Comai, Tyagi et

al. 2000). In natural tetraploid *A. suecica* silencing of protein coding genes through cytosine methylation was reported; and among 4,428 cDNA fragments, 2.5% of polymorphic alleles were estimated to be differentially expressed between the natural allotetrapolyploid and its extant parental donors *A. thaliana* and *C. arenosa* (Lee and Chen 2001). A further characterization of a subset of these cDNA fragments showed they represent several categories of genes including four transcription factors and a transposase.

In natural allotetraploid *G. hirsutum* (AD genome), organ-specific silencing due to epigenetic mechanisms was observed by analysing the two homoeologous of many genes in different vegetative and floral organs through SSCP (Single-Stranded Conformation Polymorphism) by using RNA transcripts for producing single strand cDNAs to be separated electrophoretically and quantified by a software program (Adams, Cronn et al. 2003). In fact, silencing or biased expression of genes belonging to only one of the two extant parental genomes of *G. hirsutum* (AD genome) was proved for 11 genes in at least one organ. Furthermore, in some cases, such as *Adh*A gene (alcohol dehydrogenase A), biased expression was detected toward D genome in some organs (e.g. leaves and bracts) and toward A genome in others (e.g. cotyledons and roots). Furthermore, in floral organs the A-genome homeolog of *Adh*A gene is silenced in petals and stamens and the D-genome homeolog in carpels, suggesting subfunctionalization of the duplicated genes present in allotetraploid *G. hirsutum* (AD genome).

1. 4. 1. 4 **Nucleolar dominance**

Nucleolar dominance is a phenomenon, where nucleoli are present on chromosomes belonging to only one parental species, as a result of inactivation (silencing) of ribosomal genes from the other parental species, mainly observed in hybrids and allopolyploids (Pikaard 2000). In plants, nucleolar dominance was discovered about 70 years ago and described in many interspecific hybrids as reviewed in Pikaard (2000) together with intergeneric hybrids such as *Triticale* (Amado, Abranches et al. 1997; Neves, Silva et al. 1997) while in autopolyploids this phenomenon is still poorly documented. Recently it was studied at molecular level in *Brassica* and *Arabidopsis-Cardaminopsis* allopolyploids (Chen and Pikaard 1997; Chen and Pikaard 1997; Chen, Comai et al. 1998). Data obtained in natural and synthetic

Brassica allopolyploids gave evidence that epigenetic modifications such as cytosine methylation and histone deacetylation contribute to silence the ribosomal genes from the non–dominant genome, suggesting that the phenomenon is rapid and conserved by the evolution at the same time (Chen and Pikaard 1997a), and furthermore reversible as silenced rRNA genes were reactivated by using inhibitors of DNA methyltransferases and histone deacetylases, such as 5-azacytidine and trichostatin A, respectively (Chen and Pikaard 1997b). In addition, in hybrids obtained by crossing *A. thaliana* (A genome) with *C. arenosa* (C genome) the nucleolar dominance appeared to be an unstable phenomenon according to generations and genomic ratio. In newly formed allopolyploids with genome constitution AACC, F1 individuals displayed codominance of *C. arenosa* and *A. thaliana* genes, but also dominance of *C. arenosa* and silencing of *A. thaliana* genes; while F2 individuals displayed dominance of *C. arenosa* and silencing of *A. thaliana* genes. On the contrary, by changing the genomic ratio in favour to *A. thaliana*, that was creating a newly formed allopolyploid with genome constitution AAAC, it was possible to reverse the dominance effect: *A. thaliana* genes were expressed and *C. arenosa* rRNA genes were silenced (Chen, Comai et al. 1998).

1. 4. 2 **Direct and indirect causes of epigenetic modifications**

Much evidence has been reported about direct and indirect causes of epigenetic modifications in plants. In *Arabidopsis* has been shown that repeated sequences can pair ectopically with other repeats, inter- or intrachromosomally, and the aberrant paired regions are directly targeted by DNA methyltransferases leading to epigenetic silencing (Assaad, Tucker et al. 1993). Similarly, Liu and Wendel (2003) noticed that allelic interactions between homoeologous chromosomes and ectopic interactions between homologous or homoeologous chromosomal segments, for example as a result of intergenomic colonization and intergenomic horizontal transfer as reported in wheat and cotton earlier (Zhao, Si et al. 1998; Belyayev, Raskina et al. 2000; Hanson, Islam-Faridi et al. 2000), are resulting in aberrant pairing or non-pairing at meiosis and could be targeted directly by DNA methyltransferases in the attempt to silence foreign DNA sequences or by other chromatin remodelling processes. In addition, the presence of direct and inverted repeats can be indirectly a trigger for DNA methylation. As shown for *Arabidopsis PAI* gene family (*PAI*1, *PAI*2, *PAI*3) in some ecotypes an additional copy of *PAI*1 in inverted orientation leads to methylation of all four members of the

gene family, while in the ecotypes without the inverted repeat all genes are unmethylated (Luff, Pawlowski et al. 1999; Melquist, Luff et al. 1999). According to Bender (2004) in the case of *PAI* gene family RNA Interference can be responsible for DNA methylation. Furthermore, Liu and Wendel (2003) proposed that inverted repeats could be also generated by transposable elements, if reciprocal or unidirectional translocations between different genomes lead to the possibility that a transposable element is placed in inverted orientation to its homoeologue. Alternatively, two copies of transposable elements in tail-to-tail orientation can be inserted affecting flanking host sequences and/or altering the expression of nearby genes.

1. 4. 3 The role of epigenetic phenomena, chromatin remodelling and RNA Interference in polyploids

Recently, many authors have speculated about the role of DNA methylation, histone modifications, chromatin remodelling and RNA Interference in polyploid evolution (Liu and Wendel 2003; Osborn, Pires et al. 2003; Madlung and Comai 2004).

Osborn *et al.* (2003) proposed a model about the effect of polyploidy on chromatin restructuring whereby both diploid parental species have evolved a stable system of chromatin remodelling factors that regulate and organize their own chromatin structures. After merging two genomes in a new allopolyploid species, each set of chromatin remodelling factors can also affect the chromatin of the other genome, resulting in a wide alteration of chromatin structure of both genomes. Mechanisms due to RNA silencing are also likely to happen. As a result of these epigenetic changes in polyploids repression of gene expression or derepression of sequences that were repressed in the diploid parents could be induced, affecting the phenotype directly in the case of enzymes and structural proteins or indirectly in the case of regulatory genes.

Similarly, Liu and Wendel (2003) noticed that after an allopolyploidization event chromatin remodelling processes could be altered, firstly leading to histone modifications and then DNA methylation changes. To reinforce their speculation they observed that DNA methyltransferases do not show any sequence-specificity, although non–random DNA methylation changes were reported in newly synthesised *Arabidopsis-Cardaminopsis* and *Triticeae* allopolyploids (Liu, Vega et al. 1998; Shaked, Kashkush et al. 2001; Madlung, Masuelli et al. 2002; Han, Fedak et al. 2003), suggesting that higher order processes were responsible for such alteration in DNA

methylation patterns. Furthermore, the authors extended their speculation to the possibility that chromatin remodelling rather than DNA demethylation was responsible for transposable element activation in *Triticeae* (Kashkush, Feldman et al. 2002) and *Arabidopsis-Cardaminopsis* (Madlung, Tyagi et al. 2005). Finally, the authors hypothesised that alterations in chromatin folding patterns and position effects can be related to allopolyploid formation in cotton where no alteration in DNA methylation patterns was shown (Liu, Brubaker et al. 2001).

Madlung and Comai (2004) proposed that stress events, such as an interspecific hybridization, cause a relaxation of the existing "epigenetics imprints" and result in the occurrence of mobile DNA element activation or expression of sequences normally silenced, for example. As a consequence, a wide range of phenomena, such as RNA Interference and other "heterochromatin initiation and maintenance pathways", are involved in genome restructuring and chromatin remodelling in the attempt to regain genome stability and a new balance between "expressed and repressed chromatin".

1. 4. 4 **Consequences of epigenetic phenomena in polyploid evolution**

As mentioned earlier, new phenotypic traits are common consequences of polyploidy, affecting metabolic and developmental processes and possibly allowing polyploids to colonize new niches and habitats. It has been commonly accepted that especially allopolyploids are successful in colonizing new environments and well adapted to environmental changes as they can generate genetic and/or epigenetic variations resulting in novel phenotypes (Liu and Wendel 2003; Comai 2005). The case of *Spartina anglica* is an example of a recent allopolyploid speciation that happened at the end of the 19th century on the coasts of England and France where the new species showed an invasive nature in comparison to the parental species, possibly as a result of an increased heterosis or variability due to epigenetic phenomena (Baumel, Ainouche et al. 2001; Salmon, Ainouche et al. 2005). Similarly, many examples of allopolyploid species can be found in the arctic flora and have been shown to be very successful in invading new areas (Brochmann and *et.al.* 2004) possibly as a consequence of genetic and epigenetic changes. In newly synthesised *Arabidopsis-Cardaminopsis* allopolyploids, new phenotypes and phenotypic instability (e.g. morphology, flowering time) were observed due to epigenetic changes as well as poor fertility (Comai, Tyagi et al. 2000; Madlung, Masuelli et al. 2002). According to Comai (2005) these "sterility

bottlenecks" need to be rapidly solved to have a fertile and thus successful new allopolyploid species.

Furthermore, asexual reproduction, self-pollination, changes in morphology of floral organs (that can affect interactions with pollinators) and in flowering time were also reported in plants as consequence of polyploidy due to genetic or epigenetic mechanisms in order to create prezygotic barriers between a new polyploid species and its progenitors (Otto and Whitton 2000). In newly synthesised *Brassica* allopolyploids for example, changes in flowering time (e. g. early- and late-flowering lineages) arose as a consequence of changes in flowering genes due to alteration of gene structure and/or gene expression, as consequence of DNA methylation, chromatin modifications or alteration of regulatory networks for example (Schranz and Osborn 2000; Pires, Zhao et al. 2004). Furthermore, DNA methylation seems to have a role in vernalization, that is the occurrence of flowering after prolonged exposure at low temperatures, and represents an adaptive response of plants to high altitudes or latitudes in the attempt to delay flowering until the spring season. In fact, as reported in *Arabidopsis* and plant species sensitive to vernalization such as winter wheat, demethylation through 5-azacytidine treatment or by using a methyltransferase antisense construct was sufficient to cause early flowering, suggesting that a prolonged exposure at low temperatures could decrese DNA methylation levels and thus activate genes necessary for flowering (Finnegan, Genger et al. 1998; Finnegan, Peacock et al. 2000).

In addition, epigenetic mechanisms result in the formation of epialleles, defined as "different epigenetically regulated forms of a gene", randomly occurring without changing the DNA sequence of a gene and allowing successful adaptation in response to environmental and genomic stresses or originating novel phenotypes (Finnegan 2001; Kalisz and Purugganan 2004). For example, a natural mutant of flower symmetry (a radial symmetric flower instead of asymmetric one present in the wild type) originally described by Linnaeus in *Linaria vulgaris* was recently shown to be caused by hypermethylation of Lcyc gene that is responsible for the change in flower morphology from bilateral to radial (Cubas, Vincent et al. 1999).

Finally, Rapp and Wendel (2005) suggested that "new spatial and temporal patterns of gene expression may be achieved via a suite of epigenetic mechanisms, even in the complete absence of genetic variability". According to the authors, in contrast to the traditional genetic view where a bottleneck results in a permanent loss of genetic and phenotypic diversity, the epigenetic perspective suggests that novel phenotypes due

to epigenetic mechanisms can overcome a genetic bottleneck contributing to make polyploids highly successful in evolution.

1. 5 Aims

The overall aim of the present work was to investigate genome interactions in polyploid plant species of *Poaceae* tribe *Triticeae* by focusing on the heterochromatic portion of their genomes. The wheat allopolyploid complex made of several natural diploid and polyploid *Triticeae* species together with *Triticale* hybrids (between rye and wheat species) have been chosen as a model of study, and a major repetitive DNA sequence family of *Triticeae* genomes, the 120-bp repetitive DNA sequence family, present in the heterochromatic regions of many natural wheat and rye diploid species such as *T. monococcum, T. urartu, Ae. speltoides, T. tauschii* and *S. cereale* appeared an useful tool for investigating polyploid *Triticeae* species resulting from the combinations of diploid *Triticeae* genomes. In addition, the role of DNA methylation and histone H3 methylation, two major epigenetic marks of heterochromatin, in relation to genome interactions in *Triticeae* polyploid species and *Triticale* hybrids has been investigated.

A wide range of molecular and cytological techniques was used. Molecular techniques comprised PCR amplification, cloning and sequencing; Real Time PCR amplification; restriction with methylation-sensitive enzymes and Southern hybridization. The investigations at cytological level were performed by fluorescent and genomic *in situ* hybridization; immunostaining with anti-methylcytosine antibody and anti-di-methylated H3 histone at lysine 9 antibody.

More specifically, the aim of the present work was to characterize the 120-bp repeat unit family in diploid and polyploid *Triticeae* species according to the following points and in order to answer the following questions:

- Species distribution:
  How did 120-bp repetitive DNA sequence family spread in diploid and polyploid *Triticeae* species?
- Sequence similarity and polymorphisms:
  Does 120-bp repetitive DNA sequence family present species-specific variants?
- Phylogenetic relationships:

In which species or group of species 120-bp repetitive DNA sequence family originated and which are the phylogenetic relationships of *Triticeae* species based on this major repetitive fraction of their genomes?

- Molecular organization and sequence motifs:
  Does 120-bp repetitive DNA sequence family have a tandem array organization in diploid and polyploid *Triticeae* species?
- Chromosomal location and organization:
  Does 120-bp repetitive DNA sequence family have a preferential chromosomal location in diploid and polyploid *Triticeae* species?
- Amount and copy number:
  Does the amount and copy number of 120-bp repetitive DNA sequence family change between diploid *Triticeae* species and does it follow the additivity rule in polyploid *Triticeae* species?
- GC content:
  Is the GC content of 120-bp repetitive DNA sequence family high or low?
- Methylation levels and patterns at symmetrical and asymmetrical sites:
  Do DNA methylation levels and patterns at symmetrical and asymmetrical sites of 120-bp repetitive DNA sequence family change between diploid and polyploid *Triticeae* species?

In addition, by using 120-bp repeat unit family sequences as a tool for identifying chromosomes from different diploid *Triticeae* genomes in *Triticeae* polyploids, whole-genome DNA methylation patterns and whole-genome histone H3 at lysine 9 methylation patterns were investigated on chromosome spreads in order to detect if:

- Alterations or changes in whole-genome DNA methylation patterns accompanied the evolution of polyploid *Triticeae* species.
- Similar or different changes affect whole-genome DNA methylation patterns and whole-genome histone H3 at Lysine 9 methylation patterns in polyploid *Triticeae* species.

Finally, *Triticale* being not an established species but an allopolyploid species with a recent evolutionary history and showing phenotypic and genomic instability, two different examples of alteration of DNA methylation patterns were investigated on it, by

crossing two advanced *Triticale* lines and by using treatments with 5-azacytidine, an inhibitor of DNA methylation, in order to evaluate:

- Extent of DNA alterations in whole-genome DNA methylation patterns and methylation patterns of 120-bp repetitive DNA sequence family.
- How the "memory" of DNA methylation code is re-established after an alteration of DNA methylation.
- Whether an alteration of DNA methylation could affect plant growth and development.

Chapter 2: Materials and Methods

2. 1 Materials

2. 1. 1 List of species

As listed in table 2.1 diploid and polyploid *Triticeae* species as well as *Avena sativa* 'Titan' (tribe *Aveneae*) were investigated in the present work. Seeds of most species belonged He slop-Harrison Molecular Cytogenesis Laboratory Seed Stock; two advanced *Triticale* lines together with seeds of F1, F2 and F3 generation obtained by crossing them were kindly provided by EUCARPIA and IRTA Seed Banks; and *Triticale* 'Fidelio' and 'Lamberto' have been purchased by Semundo-Danko Company.

2. 1. 2 Seed germination and seedling growth

Seeds were germinated for two-three days at 20 °C in Petri dishes containing water-soaked filter paper. When roots were about 1-2 cm long synchronization of metaphases in seedling root tips was achieved by a cold treatment at 4°C for 24 h followed by a recovery at 20°C for other 24h. Finally, roots were collected, treated and fixed according to different immunostaining and *in situ* hybridization procedures: ice-treatment followed by fixation in ethanol:glacial acetic acid and fixation in freshly made paraformaldehyde solution respectively, as will be described later in this chapter.

After root collection step seedlings were put into soil and grown in greenhouse in order to obtain plants for extracting DNA from fresh young leaves, for collecting additional roots if necessary for immunostaining and *in situ* hybridization analyses, for observing plant growth and phenotype. For each diploid and polyploid *Triticeae* species the same individuals used for Real Time PCR experiments and Southern hybridization analyses were investigated by immunostaining with anti-methylcytosine antibody and subsequent *in situ* hybridization, as will be described later in chapters 4 and 5.

Table 2. 1. List of species investigated in the present work. Chromosome numbers, ploidy levels, genome constitutions and sources are indicated. References for chromosome number, ploidy level and genome constitution: A. Bennett and Leitch (1995), B. Castilho and Heslop-Harrison (1995), C.Taketa *et al*. (2000), D. Katsiotis *et al*. (1997), E. Orgaard and Heslop-Harrison (1994).

Species	Chromosome number	Ploidy level (all x=7)	Genome Constitution	Source*
Aegilops umbellulata [B]	2n=14	diploid (2x)	UU	1
Aegilops speltoides [A]	2n=14	diploid (2x)	BB	1
Leymus mollis [E]	2n=28	tetraploid (4x)	NNXX	1
Hordeum brachyantherum [E]	2n=28	tetraploid (4x)	HHHH	1
Hordeum bulbosum [C]	2n=14	diploid (2x)	II	1
Hordeum chilense [C]	2n=14	diploid (2x)	HH	1
Secale cereale 'Petkus Spring' [A]	2n=14	diploid (2x)	RR	1
Secale montanum [A]	2n=14	diploid (2x)	RR	1
Secale vavilovii [A]	2n=14	diploid (2x)	RR	1
Triticale 'Fidelio' [A]	2n=42	hexaploid (6x)	AABBRR	2
Triticale 'Lamberto' [A]	2n=42	hexaploid (6x)	AABBRR	2
Triticale SEC776	2n=42	hexaploid (6x)	AABBRR	3

Table 2. 1. continued

Species	Chromosome number (all x=7)	Ploidy level	Genome Constitution	Source*
Triticale ID-3013	2n=42	hexaploid (6x)	AABBRR	3
Triticale SEC776xID-3013 (F1)	2n=42	hexaploid (6x)	AABBRR	3
Triticale SEC776xID-3013 (F2)	2n=42	hexaploid (6x)	AABBRR	3
Triticale SEC776xID-3013 (F3)	2n=42	hexaploid (6x)	AABBRR	3
Triticum aestivum 'Chinese Spring' [A]	2n=42	hexaploid (6x)	AABBDD	1
Triticum monococcum [A]	2n=14	diploid (2x)	AA	1
Triticum tauschii [A]	2n=14	diploid (2x)	DD	1
Triticum turgidum ssp. *durum* [A]	2n=28	tetraploid (4x)	AABB	1
Triticum urartu [A]	2n=14	diploid (2x)	AA	1
Avena sativa 'Titan' [D]	2n=42	hexaploid (6x)	AACCDD	1

Source:1 Heslop-Harrison Molecular Cytogenetic Laboratory Seed Stock; 2 Semundo-Danko Company; 3 EUCARPIA and IRTA Seed Banks.

2. 1. 3 Standard solutions and media

The following standard solutions and media were used:

10x TE buffer (Tris-EDTA buffer)

100mM Tris (tris-hydroxymethylamino-methane)-HCL, pH8

10 mM EDTA (ethylene-diamine-tetra-acetic acid), pH8

TAE buffer

40 mM Tris-acetate

1mM EDTA

Gel-loading dye buffer 6X:

0.25% Bromophenol blue,

0.25% Xylene cyanol FF

30% Glycerol

LB medium

Bacto-Tryptone 1% (w/v)

Yeast Extract 0.5 % (w/v)

86mM NaCl; 20mM glucose

LB agar

Bacto-agar 1.5 % (w/v)

20mM glucose

20x SSC (saline sodium citrate or standard saline citrate)

0.3 M NaCl

0.03 M Na citrate

10x PBS (phosphate-buffered saline)

1.3 M NaCl

70 mM Na_2HPO_4,

30 mM NaH_2PO_4

10x KPBS

1.28 NaCl

20 mM KCl

80 mM Na_2HPO_4,

20 mM NaH_2PO_4

2. 2 Methods

2. 2. 1 Genomic DNA extraction

Total genomic DNA was extracted from leaves using the CTAB method (Gawel and Jarret 1991) with minor modifications. Fresh young leaves were harvested from single individuals. From 0.5 to 1g of material was collected. Leaves were quickly ground in liquid nitrogen to prevent enzymatic degradation, the powder put into a Falcon tube with 10 ml CTAB extraction buffer (CTAB Cetyltrimethylammonium bromide 2% (w/v), 100mM Tris-HCl, pH 8; 1.4M NaCl; 20mM EDTA, pH 8) pre-heated to 65°C and finally the tubes were shaken to suspend the powder. To each tube β-mercaptoethanol was added 1µl/10ml. Tubes were incubated for 30 min at 65°C agitating from time to time. 7.5 ml of chloroform-isoamylalcohol 24:1 (v/v) was then added to each tube and mixed by inversion at room temperature for 15 min. Samples were centrifuged at 5000 rpm for 5 min at room temperature. After centrifugation the content of the tubes contained three parts (aqueous, leaf debris, organic) and the top aqueous supernatants were transferred in another Falcon tubes using 1000-pipette with blue tips cut at the end. An equal volume of cold-isopropanol was added to each tube, tubes were mixed by inversion for 15-20 sec and left on ice for 1 h. A second centrifugation step at 3000 rpm for 5 min was performed, the supernatant discarded carefully, the pellet dried and re-suspended in 1x TE buffer (250-500 µl). RNAse treatment was performed with 10µg/ml for 30 min at room temperature and then DNA precipitated adding 1/10vol sodium acetate 3M (pH 6.8) and 2xvol ETOH 95% and leaving it overnight at -20°C. A first centrifugation step was performed at 14000 rpm at 4°C for 20 min, then the supernatant was discarded and 700 µl of cold-ETOH 70% added to purify the DNA. A second centrifugation step was done at 14000 rpm at 4°C for 10 min, finally the pellet was dried and dissolved in TE (100- 200µl). DNA samples were stored at -20°C. Genomic DNA was quantified by spectrophotometer (Helyos)

(sample dilutions 1:50) and the ratio O.D260/O.D.280 was used for checking DNA purity. DNA samples (1-2μl) were run in 1% agarose gel electrophoresis in 1× TAE running buffer for checking their integrity and purity by using Hyperladder I and II (Bioline) that contain bands of known length and concentration.

2. 2. 2 PCR amplification, cloning and sequencing

Sequences belonging to 120-bp repeat unit family were amplified from genomic DNA by PCR amplification (Mullis, Faloona et al. 1986). Amount of genomic DNA and primer pairs used in PCR reactions will be described in chapter 3. PCR reactions were performed in a TGradient Thermocycler (Biometra) in 30 μl reaction containing: 1× PCR buffer (16 mM $(NH4)_2SO_4$, 67 mM Tris-HCl, 0.1% Tween 20; Bioline), 1.5 mM $MgCl_2$, 300 μM of dNTPs (Roche), 10μM of primer, 0.75 U of Taq DNA Polymerase (Bioline) and 1μl of template DNA or water as control. Cycling conditions were: 94°C for 5 min; 35 cycles of 94°C 30 sec, X°C for 30 sec and 72°C for 90 sec, followed by 72°C for 5 min.

2. 3. 2. 1 Purification of PCR-amplificated DNA fragments

After PCR amplification, PCR products (30 μl) were mixed with 6x gel-loading dye buffer 1:5 (v/v) and separated by 2% agarose gel electrophoresis in 1× TAE running buffer. Gels were run at 5-8V/cm. Products were visualized by staining with ethidium bromide (0.5μg/ml) and the size of products was checked by comparison with Hyperladder II (Bioline) as markers of length and concentration. The most prominent bands were cut from the gel with a knife blade and placed in a 1.5 Eppendorf tubes. DNA was recovered from the agarose by using QIAquick Gel Extraction Kit (Qiagen) according to manifacture's instructions. DNA concentration and length of PCR fragments (5μl) was checked by 1% agarose gel electrophoresis in 1× TAE running buffer with Hyperladder II (Bioline) containing bands of known length and concentration.

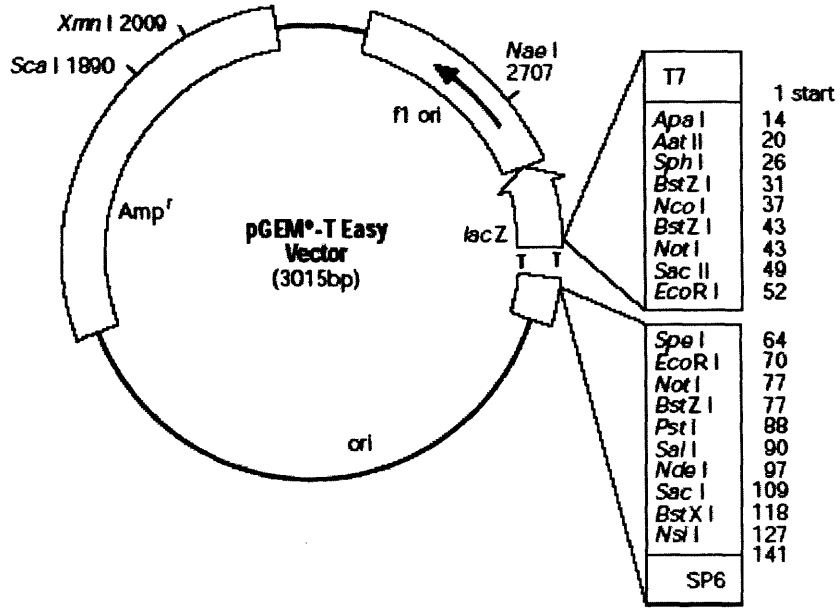
2. 3. 2. 2 Ligation of DNA fragments

PCR fragments were ligated into pGEM-T plasmids by using the kit pGEM-T

Easy Vector System I (Promega) according to manifacture's instructions with minor modifications. As shown in fig. 2. 1, pGEM-T plasmid has a 3' terminal thymidine to both ends in order to increase the ligation efficiency of blunt ended PCR fragments, by preventing the recircularization of the vector and creating compatible overhangs for PCR products. The ligation was performed in 10 µl reaction containing: 5 µl of 2x Rapid Ligation Buffer (60 mM Tris-HCl, pH 7.8; 20 mM MgCl2; 20 mM DTT; 2 mM ATP; 10% PEG from Promega); 1 µl of the pGEM-T plasmid; 1 µl T4 DNA Ligase; 4 µl DNA sample. A positive control reaction was obtained by performing the ligation of control insert DNA (542 bp) provided with the Promega kit. For DNA samples and control insert DNA overnight incubation at 4 °C was performed to maximize the number of recombinant plasmids.

2. 3. 2. 3 Competent cells of *E. coli*

Cells of *E. coli* are able to be transformed with DNA plasmids when they are in the condition of "competence'. Competent cells of *E. coli* were prepared from a glycerol stock of *E. coli* XL1-Blue. An LB plate with tetracycline (15µg/ml) was streaked of *E. coli* XL1-Blue by using a toothpick dipped into the glycerol stock and left at 37°C overnight. A single colony was picked, inoculated in 10 ml LB medium with tetracycline (15µg/ml) and grown overnight at 37°C in an orbital incubator (shaking at 250 rpm/min). Then, 50 ml of solution A (LB medium supplemented with 10mM $MgSO_4x7H_2O$ and 0.2% glucose) were inoculated with 0.5 ml of the overnight culture and incubated at 37°C in an orbital incubator (shaking at 250 rpm/min) for few hours until the optical density at 600 nm was 0.6 (OD_{600}=0.6). After that the culture was cooled for 10 min on ice and centrifugated for 5 min at 3000 rpm to pellet bacteria. Bacteria were carefully re-suspended in 0.5 ml of ice-cold solution A and afterwards 2.5 ml of storage solution B (LB medium supplemented with 12mM $MgSO_4x7H_2O$; Glycerol 36% (v/v); PEG polyethylene glycol 12% (w/v)) were added to the mixture and carefully mixed. Competent cells were then aliquoted (100µl) in pre-cooled Eppendorf tubes and stored at -80°C for two-three months.



M13 forward primer location: 2956-2972

M13 reverse primer location: 176-192

Figure 2. 1. Plasmid pGEM-T (Promega) used for cloning 120-bp repeat unit family DNA sequences. Vector map, sequence reference points and primer locations are indicated.

2. 3. 2. 4 Transformation of competent cells of *E. coli*

E. coli competent cells were transformed with ligation mixtures by heat shock. 2 µl of ligation mixture were added to 100 µl of competent cells, mixed briefly and then incubated for 30 min in ice. The tubes were then placed at 42°C for 90 sec (heat shock) and put on ice for 10 min. 500 µl of LB medium were added to each tube and the tubes incubated for 1 h at 37°C in an orbital incubator (shaking at 250 rpm/min). After that, 600 µl of the culture were plated in three LB plates (200 µl for each LB plate) with ampicillin (100µg/ml) X-gal 5-bromo-4-chloro-3-indolyl-β-D-galactosidase (40µg/ml) and IPTG isopropyl-β–Δ–thiogalactopyranoside (0.5mM) and incubated overnight at 37°C.

2. 3. 2. 5 Selection and storage of recombinant clones

The pGEM-T plasmid carries lacZ gene coding for a β-galactosidase that digest the chromogenic X-gal producing blue colonies, but if a DNA fragment is inserted into the plasmid it disrupt the lacZ gene, and thus the resulting white colonies are indication of successful transformed bacteria with recombinant plasmids. White colonies containing pGEM-T plasmids with 120-bp repeat unit family sequences were then picked, inoculated in 10 ml LB medium with ampicillin (40µg/ml) and grown overnight at 37°C in an orbital incubator (shaking at 250 rpm/min). Blue colonies as negative control and white colonies containing pGEM-T plasmids with control insert DNA as positive control were also inoculated in 10 ml LB medium with ampicillin (40µg/ml) and grown overnight at 37°C in an orbital incubator (shaking at 250 rpm/min).

To confirm the presence of 120-bp repeat unit family sequences into pGEM-T plasmids, PCR amplification was carried out on recombinant clones by using the universal M13 forward (5' GTA AAA CGA CGG CCA GT 3') and reverse (5' GGA AAC AGC TAT GAC CAT 3') primes located adjacent both sides of the insertion site (multiple cloning region) in pGEM-T plasmid (fig. 2. 1). PCR reactions were performed in a TGradient Thermocycler (Biometra) in 50 µl reaction containing: 1× PCR buffer (16 mM $(NH4)_2SO_4$, 67 mM Tris-HCl, 0.1% Tween 20 from Bioline), 1.5 mM $MgCl_2$, 300 µM of dNTPs (Roche), 5 µM of each M13 primer, 0.5 U of Taq DNA Polymerase (Bioline) and 3 µl of LB medium culture of recombinant clones and negative and positive controls (from a blue colony and a white colony with control insert DNA

respectively). PCR cycling conditions were: 94°C for 5 min; 35 cycles of 94°C 30 sec, 50°C for 30 sec and 72°C for 45 sec, followed by 72°C for 5 min. PCR products (5μl) were run on 1% agarose gel electrophoresis with Hyperladder II (Bioline) for checking length and concentration.

Recombinant clones were selected and their plasmids isolated by Wizard Plus Minipreps DNA Purification System (Promega) by using purification columns and buffers according to manifacture's instructions with minor modifications. Miniprep products (5μl) were run on 1% agarose gel electrophoresis with Hyperladder II (Bioline) for checking miniprep products length and concentration. Generally minipreps concentration ranged from 200 to 500ng/μl.

2. 3. 2. 6 DNA sequencing

All recombinant pGEM-T plasmids (200-500ng) containing 120-bp repeat unit family sequences were sequenced by the Protein and Nucleic Acid Chemistry Laboratory (PNACL) Sequencing Facility at the Leicester University by using M13 forward (5' TGT AAA ACG ACG GCC AGT 3') and reverse (5' CAG GAA ACA GCT ATG ACC 3') primers. Sequences and chromatograms were received by PNACL Sequencing Facility in computer format and analyzed by Chromas software.

2. 3. 2. 7 Sequence analysis and EMBL-EBI Database submission

For alignments of 120-bp tandemly repeated units (with no defined "start"), the sequences were first united head-to-tail and a start site was chosen arbitrarily. Software Jalview Multiple Alignment Editor V 1.3 (EMBL-EBI Version) was used for aligning the sequences by using default settings. The phylogeny of 120-bp repeat unit family sequences was analyzed by the Neighbour Joining method with Kimura's two parameters using Clustal W 1.8 software (DDBJ Version). 120-bp repeat unit family sequences were submitted to the EMBL-EBI Database (entries from AJ517227 to AJ517293).

2. 2. 3 DNA probe labelling

Since fluorescent detection methods have been used for *in situ* hybridization, two different approaches have been performed for nonradioactive labelling DNA probes: indirect labelling and direct fluorophore labelling (Schwarzacher and Heslop-Harrison 2000). In the first protocol, biotin or digoxigenin linked to uridine nucleotide are used in labelling procedures while the remaining nucleotides are unlabelled. Then avidin (or its derivatives) for biotin or anti-digoxigenin antibody for digoxigenin respectively, both conjugated with fluorophores, allow hybridization sites to be detected as shown in fig. 2. 2. Conversely, in the second protocol nucleotides directly conjugated to fluorophores are used in labelling procedures, then no further detection steps are required (fig. 2. 2).

2. 2. 3. 1 Indirect labelling Southern hybridization and fluorescent *in situ* hybridization (FISH)

The pSc119.2 clone (McIntyre, Pereira et al. 1990), a selection of 120-bp repetitive sequence family clones (Contento, Heslop-Harrison et al. 2005) and 5S rDNA pTa794 clone (Gerlach and Dyer 1980) were labelled by PCR amplification. The universal M13 forward (5' GTA AAA CGA CGG CCA GT 3') and reverse (5' GGA AAC AGC TAT GAC CAT 3') primers were used. Primers were synthesized by Life Technologies. Biotin 16-dUTP or digoxigenin 11-dUTP labelled nucleotides and an unlabelled nucleotide mixture of dATP, dTTP, dGTP, dCTP were used. PCR reactions were performed in a TGradient Thermocycler (Biometra) in 50 µl reaction containing: 1× PCR buffer (16 mM $(NH4)_2SO_4$, 67 mM Tris-HCl, 0.1% Tween 20 from Bioline), 1.5 mM $MgCl_2$, 300 µM of dNTPs (Roche), 10 µM biotin 16-dUTP (Roche) or digoxigenin 11-dUTP (Roche), 5 µM of each M13 primer, 1.5 U of Taq DNA Polymerase (Bioline) and 5-10 ng of template DNA or 2µl water as control. Template DNA was prepared by diluting 1/100 standard minipreparations (2-5 ng/µl). PCR cycling conditions were: 94°C for 5 min; 35 cycles of 94°C 30 sec, 45°C for 30 sec and 72°C for 45 sec, followed by 72°C for 5 min. The PCR products (3µl) were run on 1% agarose gel electrophoresis for checking product concentration, product size and label incorporation efficiency. After labelling procedure, a purification protocol was required to remove unincorporated nucleotides, enzyme and salts. Labelled DNA probes were precipitated with 1/10vol sodium acetate 3M (pH 6.8) and 2xvol ETOH 95% overnight at -20°C, then centrifugated at 12000xg for 30 min at 4°C, washed in 70% ETOH and

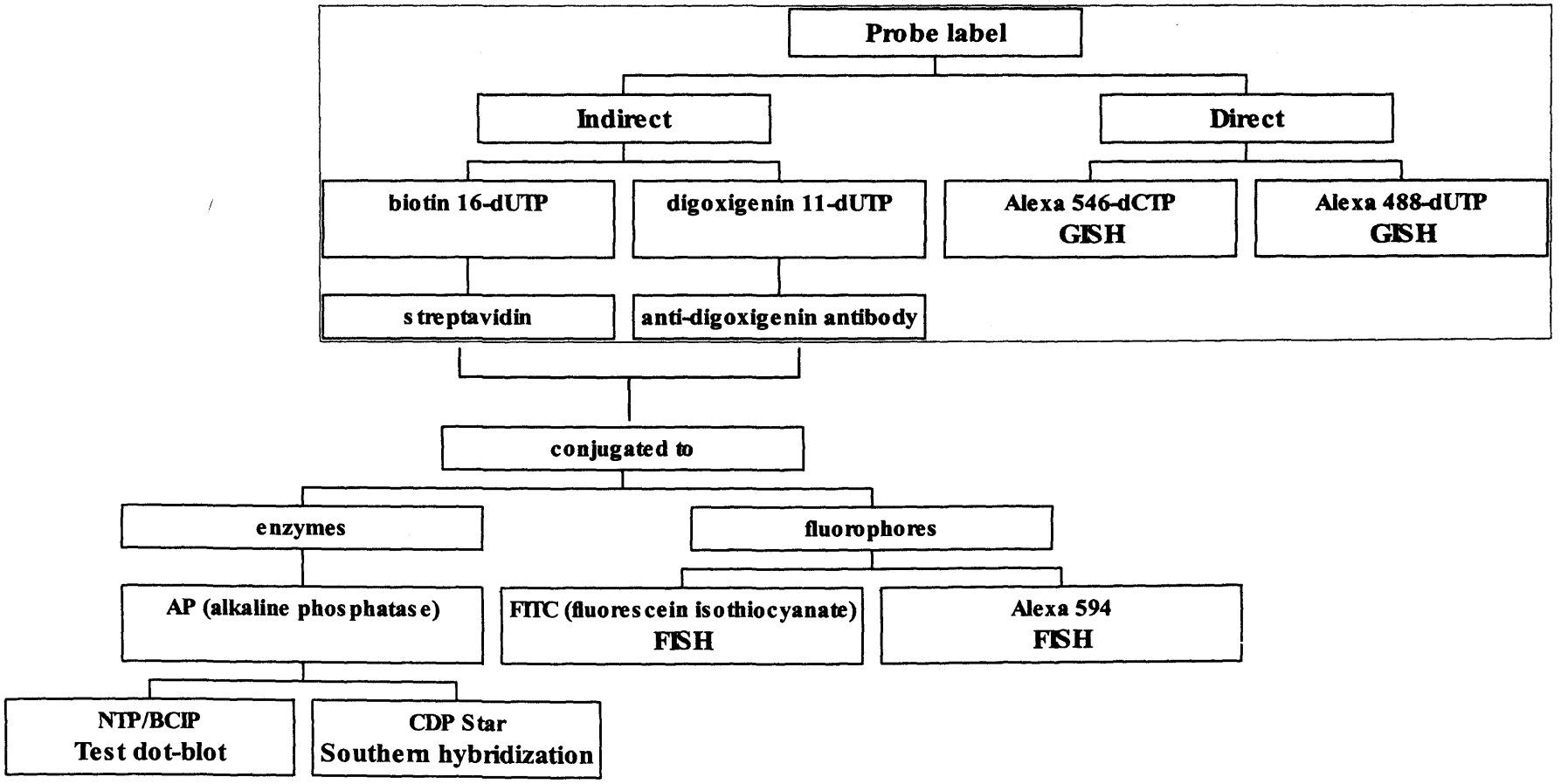


Figure 2. 2. Indirect labelling, direct labelling and signal generation system (according to Schwarzacher and Heslop-Harrison, 2000) used in Fluorescent *in situ* hybridization (FISH), Genomic *in situ* hybridization (GISH), Test dot-blot protocol and Southern hybridization.

centrifugated again at 12000xg for 5 min at 4°C. Finally, the dried pellet was dissolved in 30 µl TE.

To check labelling efficiency of probes a test dot-blot protocol was followed. A positive charged nylon membrane (Roche) was wetted in buffer 1 (100mM Tris-HCl, pH 7.5; 15mM NaCl) for 5 min and dried between filter paper. Labelled DNA probes (0.5-1µl) were micro-pipetted on the membrane and air-dried for 5 min. The membrane was then soaked in buffer 1 for 1 min, incubated in buffer 2 (Blocking Reagent (Roche) 0.5% (w/v) in buffer 1) for 30 min at room temperature in agitation, and then in antibody solution (anti-biotin-alkaline phosphatase (Roche) or anti-digoxigenin-alkaline phosphatase (Boehringer Mannheim) conjugated antibody diluted to 0.75U/ml in buffer 1) for 30 min at 37°C with agitation. Then, the membrane was washed in buffer 1 for 15 min at room temperature with agitation, equilibrated in buffer 3 (100mM Tris-HCl, pH 9.5; 100mM NaCl; 50 mM $MgCl_2$) for 2 min at room temperature with agitation and finally incubated in buffer 3 containing NBT (4-nitroblue tetrazolium chloride) and BCIP (5-bromo-4chloro-2-indolylphosphate) detection reagents (Life Technologies) (final concentration 75mg/ml and 50mg/ml respectively) for 5-10 min at room temperature in the dark. A colorimetric reaction developed and labelled probes showing strong purple-brown dots were chosen for Southern hybridization and fluorescent *in situ* hybridization (FISH) experiments.

As will be described later in this chapter, biotin- or digoxigenin-labelled probes were obtained by PCR amplification from pSc119.2 clone, a selection of 120-bp repetitive sequence family clones and 5S rDNA pTa794 clone and used to produce different signals in Southern hybridization and fluorescent *in situ* hybridization (FISH) as shown in fig. 2. 2. In fact, in Southern hybridization anti-biotin-alkaline phosphatase or anti-digoxigenin-alkaline phosphatase conjugated antibodies were used to develop chemiluminescent signal by alkaline phosphatase detection. Conversely, in fluorescent *in situ* hybridization (FISH) experiments streptavidin and anti-digoxigenin antibody conjugated to fluorophores were used to develop fluorescent signal on chromosome spreads.

2. 2. 3. 2 Direct labelling for genomic *in situ* hybridization (GISH.)

In Random Primer labelling method, single-stranded DNA is labelled by using a random mixture of oligonucleotides, fluorescently labelled nucleotides plus unlabelled nucleotides, and Klenow fragment of *E. coli* DNA polymerase I, which only has a 5' to 3' template-dependent DNA polymerase activity. Genomic DNA (1 μg) of *S. cereale* 'Petkus Spring' and *Ae. speltoides* was sheared into smaller fragments by autoclaving at 110°C it for 4 min. The fragments were run in 1% agarose gel electrophoresis and their size estimated to be in the range of 300-700 bp by using Hyperladder I (Bioline) as marker for length. The kit BioPrime Array CGH Genomic Labeling System (Invitrogen) was used according to manifacture's instructions with minor modifications. Briefly, 20 μl of 2.5 Random Primers Solution (containing random octamers) were added to 25 μl of shared genomic DNA (1μg), then denaturated at 95°C for 5 min and cooled on ice for 5 min. 1 μl of Klenow Fragment together with 2.5 μl of 10xdCTP Nucleotide Mix or 10xdUTP Nucleotide Mix were added and 1.5 μl of Alexa 546-dCTP (1 mM) or Alexa 488-dUTP (1mM) (Molecular Probes) as fluorescent nucleotides respectively. The reaction was incubated for 2 hours at 37°C, then labelled DNA probes were purified by using purification columns and buffers according to the manifacture's instructions.

2. 2. 4 Southern hybridization

2 .2. 4. 1 Restriction of DNA samples

Genomic DNA samples were digested with methylation-sensitive restrictions enzymes (McClelland, Nelson et al. 1994; Kubis, Castilho et al. 2003). Isoschizomers *Msp*I-*Hpa*II and *Bst*NI-*Scr*FI as well as *Mcr*BC restriction enzyme were used (New England BioLabs). Restriction tests were done to optimize conditions of DNA and enzyme amounts. For each digestion 7.5μg of genomic DNA has been used and 4U/μg of restriction enzyme in the appropriate enzyme buffer (1/10vol) in 25 μl of reaction. Digestion mixtures were incubated at 37°C or at the required temperature overnight to ensure complete digestion.

2 .2. 4. 2 Gel electrophoresis

Digestion mixtures were mixed with 6x gel-loading dye buffer 1:5 (v/v) and run on 2% agarose gel electrophoresis in 1x TAE running buffer in order to separate the DNA fragments according to their size. Gels were run at 5-8 V/cm. Products were visualised by staining with ethidium bromide (0.5μg/ml) and their size checked by Hyperladder II (Bioline) as a marker of length.

2 .2. 4. 3 Transfer of DNA into nylon membrane (Southern blotting)

After agarose gel separation, DNA fragments were transferred into a nylon membrane by upward capillary transfer (Southern 1975). The gel was treated with preliminary steps of depurination for 10 min (0.25N HCl), denaturation for 30 min (NaOH 0.5M, NaCl 1.5M) and neutralization for 30 min (Tris-HCl 0.5M pH7.5, NaCl 1.5M) to prepare single strand DNA fragments to be transferred. The gel was then placed upside down in the set-up for upward capillary transfer: in a tray containing transfer buffer a support was placed with a bridge on it, made of 3MM Whatman filter paper immerged in the transfer buffer from both sides. A positive charged nylon membrane (Roche) was placed on top of the gel followed by two pieces of 3MM Whatman filter paper and stack of paper towels. The 10x SSC transfer buffer moving from the tray through the bridge, the gel, the adjacent membrane and the pieces of 3MM Whatman filter paper into the stack of paper towels crated a flow of liquid that carries DNA fragments from the gel to the surface of the membrane. During the transfer the location of DNA fragments was preserved. As the rate of the transfer was due to the size of the DNA fragments but also depended on the concentration of agarose in the gel, transfers were carried out overnight to ensure the process was completed. To check transfer efficiency after blotting the agarose gel was re-stained in ethidium bromide solution (0.5μg/ml) for 10-15 min, washed in distilled water for 5 min and observed under UV light. After blotting the membranes were exposed to UV light for DNA crosslinking and stored at room temperature or 4°C before hybridization.

2 .2. 4. 4 Membrane hybridization

A non-radioactive hybridization method was used with biotin- or digoxigenin-labelled probes (Schwarzacher and Heslop-Harrison 2000). After wetting the membrane in 2xSSC it was pre-hybridized in a roller bottle at 42°C for at least 2 hours with 7ml of

pre-hybridization buffer containing: formamide 50% (v/v), 5 × SSC, Blocking Reagent (Roche) 2% (w/v), N-lauroylsarcosine 0.1% (w/v) and SDS (sodium dodecyl sulphate) 0.02% (w/v). Then, pre-hybridization buffer was replaced by hybridization solution containing formamide 50% (v/v), 5 × SSC, Blocking Reagent (Roche) 2% (w/v), N-lauroylsarcosine 0.1% (w/v), SDS 0.02% (w/v) and 30 μl (corresponding to 750 ng) of digoxigenin-labelled DNA probes (freshly denatured in boiling water for 10 min and then placed on ice to cool for 10 min). The amount of hybridization solution was calculated in accordance to the size of the membrane: 2.5 ml per 100cm^2 membrane were used. Overnight hybridization was performed in a roller bottle at 42°C.

High stringency washes were performed in agitation twice for 5 min each at room temperature with at least 100 ml of 2×SSC/SDS 0.1% (w/v) and then twice for 15 min each at 68°C with at least 50 ml 0.1×SSC/SDS 0.1% (w/v).

Detection procedure was carried out at room temperature as follows. The membrane was firstly rinsed for 5 min in 100 ml washing buffer 1 (0.1 M Maleic acid; 0.15 M NaCl, pH7.5 with Tween 20 0.3% (v/v)), then incubated for 30 min in 100 ml buffer 2 (Blocking Reagent (Roche) 1% (w/v) in buffer 1) and after that incubated for 30 min in 20 ml antibody solution (anti-biotin-alkaline phosphatase (Roche) or anti-digoxigenin- alkaline phosphatase (Boehringer Mannheim) conjugated antibody diluted to 37.5 mU/ml in buffer 2). Washes were performed twice for 15 min in 100 ml washing buffer 1, then the membrane was equilibrate for 5 min in 20 ml detection buffer 3 (0.1 M Tris-HCl; 0.1 M NaCl, pH 9.5) and finally, incubated for 5 min in CDP-Star solution (CDP-Star TROPIX (Roche) 0.25 mM in detection buffer 3). CDP-Star acts as a substrate for alkaline phosphatase and once dephosphorylated emits light at a specific wavelength that can be recorded on X-ray film. The excess of CDP-Star solution was dripped off before wrapping the membrane in cling film and exposing it to X-ray film (FUJI Medical X-Ray Film) in autoradiographic cassette in complete darkness. Multiple exposures were taken with films developed and new ones replaced from 5 min to few hours to detect the appropriate signals. Films were scanned by using EPSON 1600 Pro, and images acquired and processed by using Adobe Photoshop 6.0.

It was possible to remove probes from membranes after images were recorded and hybridize the membranes a second time with different probes. The following stripping procedure was followed. Membranes were rinsed briefly in sterile water, then washed twice for 15 min each a 37°C with 0.2M NaOH/SDS 0.1% (w/v), finally

washed for 5 min in 2xSSC, wrapped in film and stored at 4°C before pre-hybridizing and hybridizing with a biotin-labelled DNA probe.

2. 2. 5 *In situ* hybridization (ISH)

Methods for chromosome slide preparation and *in situ* hybridization (ISH) followed Schwarzacher and Heslop-Harrison (2000) with minor modifications.

2. 2. 5. 1 Ice-treatment and fixation of plant materials

In seedling root tips accumulation of metaphases was achieved by ice-treatment that halts any cell cycle at metaphase in cereals and temperate grasses (Schwarzacher and Heslop-Harrison 2000). Seedling root tips were put into tubes containing mineral water starting to freeze and the tubes placed in a mixture of ice and water at 4°C for 8-24 hours. Optimal results were obtained after sixteen hours: the chromatin was not heavily condensed and chromosomes could be stained and hybridised. After ice-treatment, roots were fixed in ethanol:glacial acetic acid (3:1) and stored at 4°C up to two-three months.

2. 2. 5. 2 Preparation of chromosome spreads

Seedling root tips were washed twice for 10 min in 1xEB buffer (4mM citric acid, 6mM tri-sodium acetate, pH 4.8) and then digested with proteolytic enzymes in enzymatic solution (Calbiochem cellulase 1.8% (w/v), Onozuka cellulase 0.2% (w/v) and Sigma pectinase 3 % (v/v) in 1xEB buffer) for 70-80 min at 37°C depending on root tip thickness. After digestion, root tips were washed in 1xEB buffer for at least 15 min and put onto glass slides (cleaned with sulphochromic acid at least 3 h and rinsed in distilled water). The root cap and other differentiated tissues were removed by using needles, then the meristematic tissue was dissected and squashed to one cell thickness in 60% acetic acid under coverslip by using light pressure. After freezing the slides on dry-ice for 5-10 min the coverslips were removed with a razor blade. Slides were air-dried and used within 24 h or stored at - 20°C for few months.

2. 2. 5. 3 *In situ* hybridization (ISH)

Slides were re-fixed in ethanol:glacial acetic acid (3:1) for 10 min, dehydrated in 96% ETOH twice for 10 min and air-dried. Chromosome preparations were pretreated with RNAse (100μg/ml) in 2xSSC for 1 h at 37°C in humid chamber, washed twice in 2xSSC for 5 min at room temperature and incubated in 10mM HCl for 5 min at room temperature. Then, slides were incubated with pepsin (5μg/ml) in 10mM HCl for 20-30 min at 37°C in humid chamber, washed twice in 2xSSC for 5 min and re-fixed in freshly made paraformaldehyde 4% (w/v) (cleared with few drops of 10M NaOH and pH 8 adjusted with 1 N H_2SO_4) for 10 min at room temperature. After that, slides were washed again twice in 2xSSC for 5 min at room temperature, dehydrated in an ethanol series (70%, 85% and 96% ETOH, 2 min each), and air-dried. Pretreatments of chromosome preparations were necessary to remove RNA and proteins that could increase the background when binding with probe or detection reagents, to remove the excess of cytoplasm covering chromosome spreads that can inhibit probe penetration, to fix the target material onto the slide in order not to loose it during the denaturation step at high temperature and finally to make target DNA sequences accessible to probes during the hybridization step.

For fluorescent *in situ* hybridization (FISH) the hybridization mixture consisted of 40 ml total volume made of 50–100 ng/slide of each probe, formamide 50 % (v/v), 2× SSC, dextran sulphate 10% (v/v), SDS 0.125% (w/v), 0.125 mM EDTA and 1 μg salmon sperm DNA. The above concentrations of formamide and salt at 37°C allowed sequences of 75-80% homology to form duplexes (80% stringency conditions) while salmon sperm DNA prevented hybridization at aspecific sites.

For genomic *in situ* hybridization (GISH.) the hybridization mixture consisted of 50–100 ng/slide of each genomic DNA, formamide 50 % (v/v), 2× SSC, dextran sulphate 10% (v/v), SDS 0.125% (w/v), 0.125 mM EDTA, 1 μg salmon sperm DNA and 500 ng/slide of blocking genomic DNA unlabelled in ratio 1:5 to 1:10 with the labelled genomic DNAs. The above concentrations of formamide and salt at 37°C allowed sequences of 75-80% homology to form duplexes (75-80% stringency conditions) while salmon sperm DNA prevented hybridization at aspecific sites.

The hybridization mixture was denatured at 70°C for ten min and placed on ice to cool for ten min before being added to chromosome preparations. Probes and chromosomes were denatured together at 70–80 °C for 6 min in ThermoHybaid HyPro-20 with vibration effect (vibration setting 3) as shaking increase the chance of probes to target chromosome sites and left to cool slowly to 37 ° C for overnight hybridization.

Post-hybridization washes were performed to remove the hybridization mixture and unbound probes. In addition, by using wash stringency slightly higher than the hybridization stringency probes not-specifically or weakly bounded to target sequences were removed in order to minimize the background signal. After overnight hybridization, slides were washed twice in formamide 20 % (v/v) in 0.1× SSC at 42°C for 5 min in agitation, twice in 0.1× SSC at 42°C for 5 min with agitation, twice in 2× SSC at 42°C for 3 min with agitation and finally left 5 min in detection buffer (4× SSC, Tween 20 0.1% (v/v)).

In fluorescent *in situ* hybridization (FISH) non-specific sites that could bind detection reagents were blocked by incubation with BSA (bovine serum albumin) 5% (w/v) (Sigma) in detection buffer for 5 min at room temperature then hybridization sites were detected by incubation with streptavidin conjugated to Alexa 594 (Molecular Probes) or FITC (fluorescein isothiocyanate)-conjugated anti-digoxigenin antibody (Roche) in detection buffer with BSA 5% (w/v) for 1 h at 37 ° C. Antibody final concentration ranged from 2-10 μg/ml. After three washes in detection buffer at 42°C in agitation for 5-10 min each chromosomes were counterstained with DAPI (4', 6-diamidino-2-phenylindole) (4 μg/ml (Sigma) in McIlvaine's buffer, made of 164mM Na_2HPO_4 and 18mM citric acid pH 7) for ten min at room temperature. DAPI binds DNA in two different ways: in the minor groove and between AT-rich regions (Wilson 90). The latter gives a high increase in fluorescence. Slides were then mounted in antifade solution (AF1, Agar Scientific) that prevents the fading of fluorescence when slides were observed.

In genomic *in situ* hybridization (GISH.) using direct labelled probes no detection steps were necessary so that after post-hybridization washes slides were counterstained with DAPI (4 μg/ml) for ten min at room temperature and mounted in antifade solution (AF1, Agar Scientific).

2. 2. 5. 4 Photography and imaging

Slides preparations were analyzed by using Axioplan 2 epifluorescence microscope with single band pass filters (Zeiss) equipped with a cooled CCD camera (Optronics, model S97790). Metaphases were recorded and photographed. Images were overlayed and analyzed by using Adobe Photoshop 6.0 software, and functions that treat all pixels equally were used for color balance and processing.

2. 2. 6 Immunostaining with anti-methylcytosine and anti-di-methylated histone H3 at Lysine 9 antibodies

2. 2. 6. 1 Fixation of plant materials and preparation of chromosome spreads

a) Immunostaining with anti-methylcytosine antibody

For immunostaining with anti-methylcytosine antibody (anti-5-mC) seedling root tips were ice-treated and fixed in ethanol:glacial acetic acid, digested with proteolytic enzymes, and squashed in 60% acetic acid as previously described for *in situ* hybridisation (ISH) protocol.

b) Immunostaining with anti-di-methylated histone H3 at Lysine 9 antibody

For immunostaining with anti-di-methylated histone H3 at Lysine 9 antibody (anti-di-methylated H3K9) different fixation and squashing procedures were required in order to preserve chromatin proteins. Seedling root tips were fixed under vacuum for 30 min with freshly polymerised paraformaldehyde 4% (w/v) in 1xKPBS buffer with Triton X-100 0.2% (v/v), then rinsed under vacuum twice in 1xKPBS buffer for 5 min and 30 min respectively and finally kept in 1xKPBS buffer overnight at 4°C. The following day, root tips were digested for 1h at 37°C in a mixture of Sigma pectinase 2.5% (v/v), Onozuka cellulase 5% (w/v) and Y-23 pectolyase 2.5% (w/v) in 1xKPBS buffer. Squashes were made under a coverslip in small drop of Triton X-100 0.5% (v/v) onto poly-L-lysine coated slides (Poly-Prep™ Slides, Sigma). Good chromosome spreads totally free of cytoplasm were difficult to get as acid fixation that would destroy proteins was avoided in this protocol and thus problems were encountered in totally removing cytoplasm. The coverslips were removed with a razor blade after freezing the slides on dry-ice for 5-10 min. Slides were not allowed to dry out and quickly transferred to 1x KPBS buffer for up to 4h before proceeding with immunostaining.

2. 2. 6. 2 Immunostaining

a) Immunostaining with anti-methylcytosine antibody

Immunostaining with primary anti-methylcytosine antibody and secondary antibody are illustrated in fig. 2. 3. For immunostaining with anti-methylcytosine antibody (anti-5-mC) slides were pretreated with RNAse and pepsin, then dehydrated in an ethanol series, and air-dried according to the *in situ* hybridization (ISH) protocol. Slides were blocked by BSA 1% (w/v) (Sigma) in 1xPBS buffer-Tween 20 0.5% (v/v) for 30 min at room temperature and then incubated with monoclonal anti-methylcytosine antibody (Oncogene) (5μg/ml per slide in 1xPBS buffer) overnight at 4°C in a humid chamber. Slides were washed in 1xPBS buffer-Tween 20 0.5% (v/v) twice for 5 min at room temperature with agitation and then incubated with Alexa 488-conjugated goat anti-mouse secondary antibody (Molecular Probes) (1: 200 in 1xPBS buffer-Tween 20 0.5% (v/v)) for 1 h at 37°C in a humid chamber. Finally, slides were washed in 1x PBS buffer-Tween 20 0.5% (v/v) twice for 5 min at room temperature in agitation, DAPI stained and mounted in antifade solution according to the *in situ* hybridization (ISH) protocol. Slide preparations were analysed under an epifluorescence microscope, metaphases recorded and photographed. The slides that need to be re-opened were put at 37°C for 10 min to reduce the viscosity of the antifade mountant and the coverslips were removed carefully with a razor blade. The methylation signal was removed by two washes in 4xSSC containing Tween 20 0.2% (v/v) for 30-60 min at room temperature, followed by two washes in 2xSSC for 5 min at room temperature, dehydrated in an ethanol series, air-dried and finally re-probed with 120-bp repeat unit family clones according to the *in situ* hybridization (ISH) protocol, allowing a comparative analysis of DNA methylation patterns and karyotypes (fig. 2. 3).

b) Immunostaining with anti-di-methylated histone H3 at Lysine 9 antibody

For immunostaining with anti-di-methylated histone H3 at Lysine 9 antibody (anti-di-methylated H3K9) slides previously kept in 1xKPBS buffer for up to 4h were re-fixed in freshly made paraformaldehyde 4% (w/v) in 1xKPBS buffer (pH 8 adjusted with 1M HCl) for 10 min at room temperature. Three washes in 1xKPBS buffer for 5 min each were performed. After that slide preparations were treated with freshly made sodium borohydride 0.1% (w/v) in 1xKPBS buffer for 5 min at room temperature to avoid autofluorescence, followed by four washes in 1xKBPS buffer 2 min each. Before antibody treatment non-specific sites were blocked by incubation with BSA 1-4% (w/v) (Sigma) in 1xKPBS buffer-Tween 20 0.1% (v/v) for 30 min at room temperature then

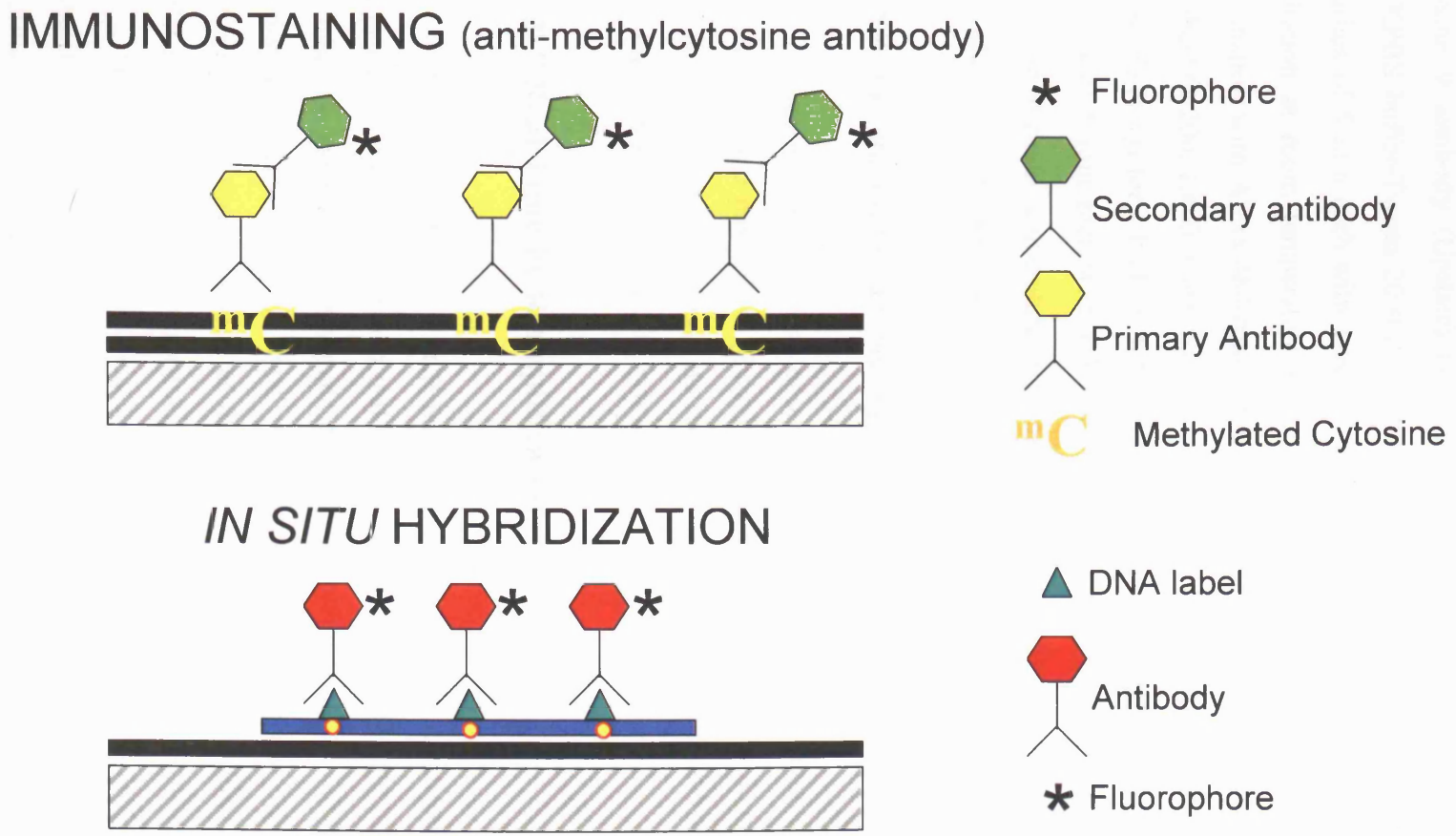


Figure 2. 3. Immunostaining for detecting DNA methylation with primary anti-methylcytosine antibody and secondary antibody is illustrated. Subsequent *in situ* hyridization allowed a comparative analysis of DNA methylation patterns and karyotypes in diploid and polyploid *Triticeae* species.

primary antibody incubation step was performed with anti-di-methylated histone H3 at Lysine 9 antibody (Upstate) (1:200 with acetylated BSA 0.1% (w/v) (Sigma) in 1xKPBS buffer-Tween 20 0.1% (v/v)), overnight at 4°C in a humid chamber. Three washes of 5 min each with 1xKPBS buffer-Tween 20 0.1% (v/v) were performed in agitation at room temperature to remove inbound primary antibodies followed by incubation with Alexa 488-conjugated goat anti-rabbit secondary antibody (Molecular Probes) (1:200, 1:400, 1:600 with 0.1% acetylated BSA (Sigma) 1xKPBS buffer-Tween 20 0.1% (v/v)) for 1 h at room temperature in a humid chamber was performed. Three washes of 5 min each with 1xKPBS buffer-Tween 20 0.1% (v/v) were performed in agitation at room temperature and finally followed by DAPI staining (4 μg/ml) for ten min at room temperature and mounting in antifade solution (AF1, Agar Scientific).

2. 2. 6. 3 Photography and imaging

Slide preparations were observed under Axioplan 2 epifluorescence microscope with single band pass filters (Zeiss) equipped with a cooled CCD camera (Optronics, model S97790). Metaphases were recorded and photographed, and images analysed as previously described for *in situ* hybridisation (ISH) protocol.

2. 2. 7 Real Time PCR amplification

Real Time PCR is a PCR-based system that can simultaneously amplify DNA sequences and quantify them in real time. A fluorescence detecting thermocycler is able to measure the fluorescent signals emitted by double-stranded DNA molecules when non-specific DNA-binding dyes, such as ethidium bromide or SYBR Green I, intercalate into them. As soon as the DNA molecules are amplified through cycles of PCR and generate the fluorescent signals, the equipment plot the fluorescence intensity in a Quantitation Curve, precisely quantifying the correspondent amount of DNA amplified. An exact quantification is obtained measuring DNA molecules amplified during the exponential phase of the PCR, when the reaction is not limited by any of its components. The threshold cycle (Ct) is the cycle of PCR required to get a certain amount of fluorescent signal that correspond to a certain amount of DNA. By comparing the Ct values obtained for standards with known amount of DNA it is possible to determine the amount of target sequences in samples of unknown

concentration. An absolute quantification is done by using a Standard Curve obtained by plotting the threshold cycle (Ct) against the logarithm of the amount of DNA for dilution series of standards with known concentration. After the PCR amplification has stopped, the equipment generates a Melting Curve by gradually increasing the temperature in order to confirm the specificity of the products amplified. As soon as the melting temperature is reached for the amplified products, a decrease of SYBR Green I is recorded as consequence of its release from the double-stranded DNA molecules following the denaturation. From the shape of the melting curve is it possible to gain information about the homogeneity of the amplified products, the presence of non-specific products and/or primer-dimers.

Real Time PCR experiments were performed in DNA Engine Opticon System made of 96 blue light emitting diodes (LEDs). Each LED beam focused onto its corresponding well in a plate exciting SYBR Green I. The system ensured minimal crosstalk and light scattering by a series of sequential steps of illuminating and detecting each well. A photomultiplier tube (PMT) detected and amplified the fluorescent signal from SYBR Green I, and a filter detection system separated the specific signal from the background. Data were collected using Opticon Monitor 2.03 Software.

Amount of total genomic DNA, plasmid dilutions of a 120-bp repeat unit family clone with known copy number (used as standards) and primer pairs used in a Two-Step Real Time PCR amplification protocol will be described in chapter 4. Real Time PCR reactions were performed in 20 µl reaction containing: 10 µl of 2x master mix from DyNAmo™ HS SYB Green qPCR kit from Finnzymes (containing a modified DyNAmo hot start DNA polymerase, SYBR Green I, optimized PCR buffer, 5mM $MgCl_2$, dNTP mix including dUTP), 1µl of 10 µM for each primer and 8 µl of template DNA or water as control. Cycling conditions of a Two-Step Real Time PCR amplification protocol were: 94°C for 3 min; 35 cycles of 94°C 10 sec, 57°C annealing temperature for 20 sec and plate read step; the melting curve was measured from 60°C to 95°C with increasing of 0.5°C and reading step of 1s. Additionally, a re-annealing step (at 70°C for 30 min) to obtain double-stranded molecules was performed and 2% agarose gel electrophoresis in 1x TAE running buffer carried out to estimate the size of the amplified products with Hyperladder II (Bioline) as a marker of length.

Chapter 3: Diversity of a major repetitive DNA sequence in diploid and polyploid Triticeae species

3. 1 Summary

In this chapter about 90 members of 120-bp repeat unit family, a major repetitive DNA sequence of *Triticeae* species originally described in *S. cereale* as pSc119.2, have been isolated from many diploid and polyploid *Triticeae* species by PCR amplification using primer combinations and single primers from along the length of the original pSc119.2 sequence. From sequence analysis, high levels of similarity were found in 120-bp repeat unit family sequences within and between *Triticeae* species with no species-specific variants. In addition, few insertions and deletions were present keeping the 120-bp repeat unit a constant length, but mainly single nucleotide mutations despite no hotspots of single-base variation were found. Phylogenetical analysis showed that no strong groupings of sequence families were evident independently of sequence copy number within a genome, ploidy level, cultivation and breeding system. Furthermore, fluorescent *in situ* hybridization (FISH) in rye, diploid and polyploid wheat species showed that chromosomal sites harboured many variants of the 120-bp repeat unit family. Finally, the evolutionary history of 120-bp repeat unit family sequences in *Triticeae* is discussed.

3. 2 Introduction

Species of the tribe *Triticeae*, including *Triticum* (wheat), *Secale* (rye), *Hordeum* (barley) species and their wild relatives, are characterized by large genomes (Bennett and Smith 1991) and the majority of their genomes is made up of repetitive DNA sequences (Flavell, Bennett et al. 1974; Flavell 1986).

Genome (or species)-specific repetitive DNA sequence families have been identified in *Triticeae* species together with universal repetitive DNA sequences found in a number of different genomes. For instance, a single widespread repetitive DNA sequence family has been isolated from several species of *Triticeae* and given different names by several authors: pAS1 from diploid wheat *Triticum tauschii* (Rayburn and

Gill 1986), pHcKB6 from *Hordeum chilense* (Anamthawat-Jónsson and Heslop-Harrison 1993), dpTa1 from hexaploid wheat *T. aestivum* (Vershinin, Svitashev et al. 1994) and the Afa family sequences (Nagaki, Tsujimoto et al. 1995; Nagaki, Tsujimoto et al. 1998) are all homologous. Another abundant repetitive DNA sequence family in *Triticeae* species, the TaiI family, showed a characteristic localization, either centromeric or subtelomeric, according to genus (Kishii, Nagaki et al. 1999; Kishii, Nagaki et al. 2001; Kishii and Tsujimoto 2002).

A major family of *Triticeae* repetitive DNA sequences, the 120-bp repeat unit family, was found in the major heterochromatic blocks at the ends of all short arms and many long arms of *S. cereale* chromosomes, consisting of short, tandemly repeated units which are present in hundreds or thousands of copies, and was among the first DNA sequences to be cloned from plants (Bedbrook, Jones et al. 1980). The original clone was named pSc119, but subsequent work showed that it was a chimera, with the sequence pSc119.2 made of three adjacent subunits having the characteristics described by Bedbrook *et al.* (1980) and being the major component of the rye heterochromatin (McIntyre, Pereira et al. 1990; Cuadrado, Ceoloni et al. 1995; Vershinin, Schwarzacher et al. 1995). Characterization by Southern hybridization, *in situ* hybridization and PCR amplification showed that 120-bp repeat unit family is present in some species of the *Triticeae* including hexaploid wheat *T. aestivum*, diploid and polyploid *Hordeum* species, and *Aegilops* species (Mukai, Nakara et al. 1993; Castilho and Heslop-Harrison 1995; Badaeva, Friebe et al. 1996; Taketa, Ando et al. 2000). Notably, homologous sequences have also been found in species belonging to the sister tribe *Aveneae* (Katsiotis, Hagidimitriou et al. 1997). While *in situ* hybridization with pSc119.2 sequence identified mainly sub-terminal sites in rye as the 120-bp repetitive sequence family is part of large subtelomeric sequence blocks, intercalary sites were detected in hexaploid wheat *T. aestivum* (Mukai, Nakara et al. 1993; Cuadrado, Ceoloni et al. 1995; Vershinin, Schwarzacher et al. 1995). Finally, the molecular organization of pSc119.2 in rye showed a trimer of 360 bp being defined by *Taq*I restriction sites and a monomer of 118 bp defined by *Hae*III (Vershinin, Schwarzacher et al. 1995).

The aim of the present chapter was to characterize in detail this major repetitive DNA sequence family of *Triticeae* genomes. Therefore, sequences related to 120-bp repeat unit family were isolated by PCR amplification, cloned and sequenced from many diploid and polyploid species of *Triticeae* and from *Avena sativa* 'Titan' (tribe *Aveneae*). Sequence analysis and phylogenetical analysis were carried out in order to

track the evolutionary history of these sequences through events involving speciation, hybridization and polyploidy. Using some clones as probes, fluorescent *in situ* hybridization (FISH) was performed in rye, diploid and polyploid wheat species in order to investigate their chromosomal distribution.

3. 3 Materials and Methods

3. 3. 1 Primer design and PCR amplification

Total genomic DNA from the following *Triticeae* species: *Aegilops umbellulata*, *Leymus mollis*, *Hordeum brachyantherum*, *H. chilense*, *H. bulbosum*, *Secale cereale* 'Petkus Spring', *S. montanum*, *S. vavilovii*, *Triticum aestivum* 'Chinese Spring', *T. monococcum*, *T. tauschii* and from *Avena sativa* 'Titan' (tribe *Aveneae*) belonging to Heslop-Harrison Molecular Cytogenetic Laboratory Genomic DNA Collection was diluted in water to approximately 10ng/µl.

Three forward and three reverse primers (table 3. 1) were designed along the original pSc119.2 sequence, which consists of three units of 118 bp (McIntyre, Pereira et al. 1990) using the Primer3 online primer design software (http://www.genome.wi.mit.edu/cgi-bin/primer/primer3www.cgi). Primer positions and expected products are indicated in fig. 3. 1a-b. Primers were synthesized and the melting temperatures calculated by Sigma-Genosys. PCR amplification was carried out with primer pairs and with single primers. As described in Material and Methods chapter, PCR reactions were performed in 30 µl-reaction containing: 1xPCR buffer, 1.5 mM $MgCl_2$, 300 µM of dNTPs, 10 µM of each primer, 0.75 U of Taq DNA Polymerase and 1 µl of template DNA or water as control. Cycling conditions for primer pairs were: 94°C for 5 min; 35 cycles of 94°C 30 sec, 60°C annealing temperature for 30 sec and 72°C for 90 sec, followed by 72°C for 5 min. When one single primer was tested, different annealing temperatures were used (see table 3. 1). After PCR amplification, PCR products were separated by 2% agarose gel electrophoresis in 1x TAE running buffer. PCR products were visualised by staining with ethidium bromide and the most prominent bands were recovered from gel by using QIAquick Gel Extraction Kit (Qiagen) according to manifacture's instructions.

Table 3. 1. Names, sequences, lenghts, melting temperatures (Tm) and annealing temperatures (Ta) of forward (F) and reverse (R) primers used in PCR amplification from diploid and polyploid *Triticeae* species and *Avena sativa* 'Titan'. Primers were synthesized and melting temperatures calculated by Sigma-Genosys.

Primer name	Sequence 5'- 3'	Length (bp)	Tm (°C)	Ta (°C)
F10	CGGCCAAAACTGCGAGTG	18	68.0°C	60-63°C
F25	GTGCTGATGACCGASACG	18	60.3°C	56-60°C
F106	CGGTGAGTGATAGTCCACG	19	61.8°C	56-60°C
R42	CCCCGGGGTGCGTTTACG	18	72.6°C	60-66°C
R147	GCACTCGCAGTTTTGGCCG	19	71.3°C	60-66°C
R208	GGGGTCCCGGAGTGATTTCC	20	70.8°C	60-66°C

3. 3. 2 Cloning, sequencing and sequence analysis

As described in Material and Methods chapter, PCR products were amplified with primer pairs from all *Triticeae* species, eluted from gels in adequate concentration (mainly containing monomers) and cloned into pGEM-T plasmids. In addition, the most intense bands amplified by PCR with single primers F25, R42 and R147 (at different annealing temperatures) from few *Triticeae* species were eluted from gel and cloned into pGEM-T plasmids. All clones were sequenced at the PNACL Sequencing Facility of Leicester University. The 120-bp repeat unit family sequences were submitted to EMBL-EBI Database (entries from AJ517227 to AJ517293). For alignments of the 120-bp tandemly repeated units (with no defined 'start'), the sequences were firstly united head-to-tail and a start site was chosen arbitrarily at the beginning of primer F25. Software Jalview Multiple Alignment Editor V 1.3 (EMBL-EBI Version) was used to align 120-bp repetitive sequences by using default settings. The phylogeny of 120-bp repetitive sequences was analysed by the Neighbour-Joining method with Kimura's two parameters using Clustal W 1.8 software (DDBJ Version). Sequences from different a *Triticeae* repetitive DNA sequence family: the pHcKB6 from *H. chilense* (Anamthawat-Jónsson and Heslop-Harrison 1993) and the Afacer1 from *S. cereale* (Nagaki, Tsujimoto et al. 1998), were chosen as outgroups.

3. 3. 3 *In situ* hybridization

As described in Material and Methods chapter, seedling root tips were treated with ice-water, fixed in ethanol-acetic acid, digested with proteolytic enzymes, and squashed in 60% acetic acid to obtain chromosome spreads. Slide preparations were pretreated with RNAse and pepsin, re-fixed in freshly made paraformaldehyde, dehydrated in an ethanol series, and air-dried. Biotin- and digoxigenin-labelled clones belonging to 120-bp repeat unit family sequences as well as pSc119.2 clone (McIntyre, Pereira et al. 1990), 5S rDNA pTa794 clone (Gerlach and Dyer 1980) were used as probes. The hybridization mixture consisted of 50 to 100 ng/slide of each probe, 50% formamide, 2xSSC, 10% dextran sulphate, 0.125% SDS, 0.125mM EDTA and 1μg salmon sperm DNA. Probes and chromosomes were denatured together at 70 to 80°C for 6 min and left to cool slowly to 37°C for overnight hybridization. Post-hybridization washes were performed in high stringency conditions and hybridization sites were

detected by streptavidin conjugated to Alexa 594 or FITC (fluorescein isothiocyanate) conjugated anti-digoxigenin antibody. Well spread metaphases were recorded and photographed. Three to nine metaphases were analysed for each tested probe combination and karyotypes established from selected metaphases of *S. cereale* 'Petkus Spring' following previous analyses (Cuadrado and Jouve 1994; Cuadrado, Ceoloni et al. 1995).

3. 4 Results

3. 4. 1 PCR amplification with primer pairs and cloning of 120-bp repetitive sequences

Primers for PCR amplification were designed from the original pSc119.2 sequence (McIntyre, Pereira et al. 1990) to be anchored at different parts of the 120-bp repeat unit and to allow amplification of fractions, complete 120-bp units or larger fragments (see fig. 3. 1a-b). The following primer combinations F25-R147, F25-R42, F25-R208, F106-R42 and F106-R208 were tested at the same annealing temperature (60°C) in all diploid and polyploid *Triticeae* species as well *A. sativa* 'Titan'. Primer pair F10-R42 did not amplify significant fragments with the chosen annealing temperature and primer combinations F106-R147, F10-R147 and F10-R208 were not tested. All PCR products obtained in diploid and polyploid *Triticeae* species and *A. sativa* with the above mentioned primer pairs are summarized in table 3. 2 and examples of gels are given in fig. 3. 2a. Monomers and dimers were mainly amplified in most species analysed. Primer combinations F25-R42, F25-R147, F25-R208, F106-R42 and F106-R208 amplified sequences from most *Triticeae* species and *A. sativa* 'Titan' (see table 3. 2), but primer pairs F106-R42 and F106-R208 showed less prominent bands than primer pairs F25-R42, F25-R147, F25-R208 (see fig. 3. 2). Primer combination F25-R147 did not show amplification in *T. monococcum*, *L. mollis* and *A. sativa* 'Titan' (see table 3. 2), and primer pair F106-R208 did not show amplification in *L. mollis* (see table 3. 2). As shown in gel pictures (see fig. 3. 2) PCR amplification resulted in distinct bands, but also showed some underlying products of variable lengths (smears) with all primer combinations tested. The sizes of the major fragments in all species were as expected from the primer positions within 120-bp repeat unit, with the most prominent bands corresponding to 115-200 bp and containing one complete unit (compare fig. 3. 2

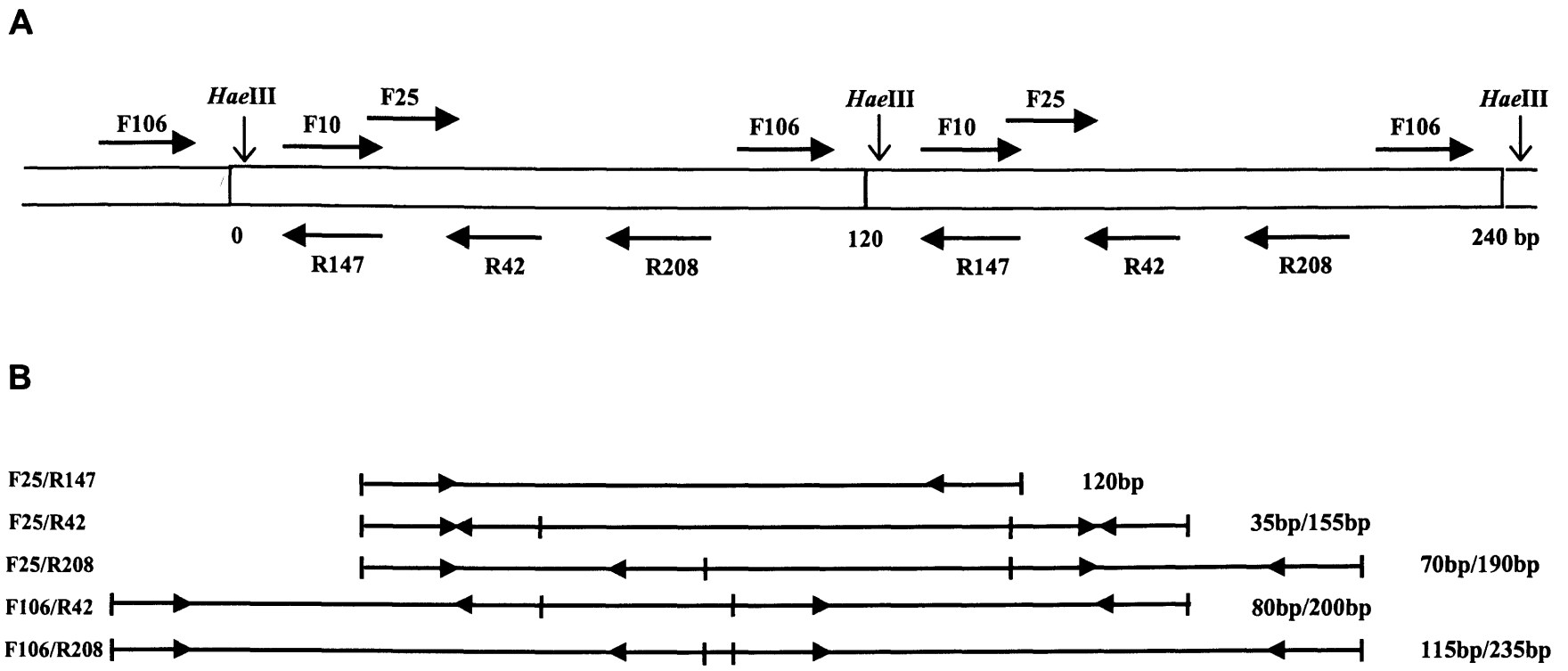


Figure 3. 1. Structure of 120-bp repeat unit family and primers used in PCR amplification. A) Diagram of 120-bp repeat with *Hae*III site defining the monomer of 118-bp (McIntyre et al. 1990; Vershinin and Heslop-Harrison, 1998). Position and orientation of primers designed from the original pSc119.2 sequence (McIntyre et al. 1990) are indicated. B) Primer pairs used to amplify by PCR 120-bp repeat family sequences from diploid and polyploid *Triticeae* species. The expected fragment lengths including at least one complete 120-bp unit are listed.

Table 3. 2. PCR products obtained in diploid and polyploid *Triticeae* species and *Avena sativa* 'Titan' using different primer pairs.

Species	F25-R42*	F25-R147*	F25-R208*	F106-R42	F106-R208
T. monococcum	Monomer and dimer	-	Monomer and dimer	Monomer and dimer	Monomer (weak band)
T. tauschii	Monomer and dimer	Monomer and dimer	Monomer and dimer	Monomer and dimer	Monomer (weak band)
Ae. umbellulata	Monomer and dimer	Monomer and dimer	Monomer and dimer	Monomer and dimer (no insert from elution)	Monomer and dimer
***T. aestivum* 'Chinese Spring'**	Monomer and dimer	Monomer and dimer	Monomer and dimer	Monomer and dimer	Monomer (no insert from elution)
***S. cereale* 'Petkus Spring'**	Monomer and dimer	Monomer and dimer	Monomer and dimer	Monomer and dimer (weak bands)	Monomer and dimer (weak bands)
S. vavilovii	Monomer and dimer	Monomer and dimer	Monomer and dimer	Monomer and dimer (weak bands)	Monomer and dimer
S. montanum	Monomer and dimer	Monomer and dimer	Monomer and dimer	Monomer and dimer (weak bands)	Monomer and dimer (weak band)

Table 3. 2. continued

Species	F25-R42*	F25-R147*	F25-R208*	F106-R42	F106-R208
L. mollis	Monomer	-	Monomer (no insert from elution)	Monomer (weak band)	-
H. bulbosum	Monomer and dimer	Monomer and dimer	Monomer and dimer	Monomer and dimer	Monomer and dimer (no insert from elution)
H. chilense	Monomer and dimer	Monomer and dimer	Monomer and dimer	Monomer	Monomer and dimer (no insert from elution)
H. brachyantherum	Monomer and dimer	Monomer and dimer	Monomer and dimer	Monomer and dimer	Monomer and dimer
***A. sativa* 'Titan'**	Monomer	-	Monomer and dimer	Monomer (weak band)	Monomer (weak band)

* Trimers of 360-430 bp lengths were visible with primer combination F25-R42, F25-R147 and F25-R208 only in some species but as weak bands.

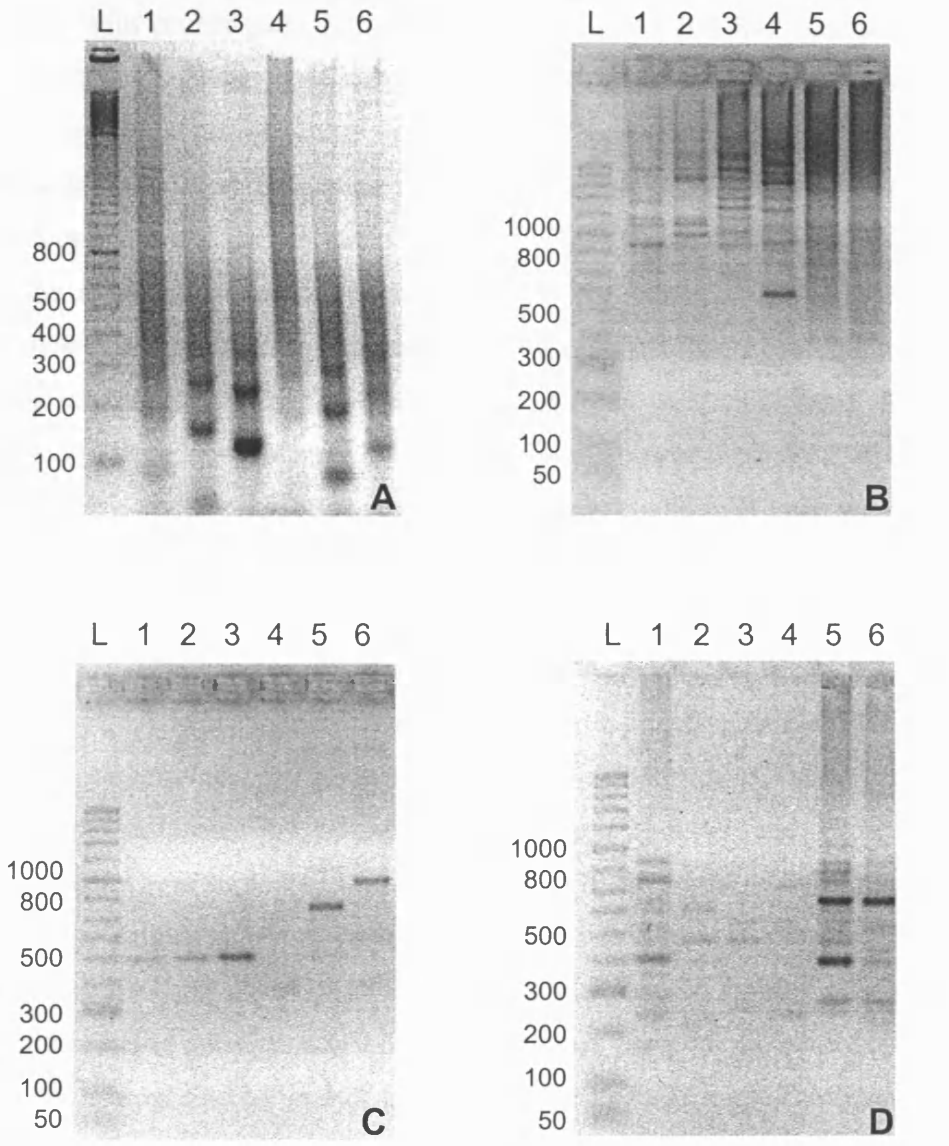


Figure 3. 2. Agarose gel separation of DNA fragments amplified by PCR. The images are inverted to show faint bands present in some of the species. A) Genomic DNA of *S. montanum* amplified with different primer pairs: lanes 1 F106/R42; 2 F25/R42; 3 F25/R147; 4 F10/R42; 5 F25/R208; 6 F106/R208. Prominent bands including incomplete and complete monomers, dimers or trimers are visible. Fragment profiles were similar for the other *Triticeae* species investigated. Genomic DNA of different *Triticeae* species amplified with single primers F25 (B) R42 (C) and R147 (D). Few fragments of varying length indicating rearranged or inverted repeat units were amplified by PCR from 1: *H. chilense*; 2: *H. brachyantherum*; 3: *T. monococcum*; 4: *T. tauschii*; 5: *T. aestivum* 'Chinese Spring'; 6: *S. cereale* 'Petkus Spring'.

with fig. 3. 1b). However, short DNA fragments corresponding to incomplete units were also visible with primer pairs F25-R42, F25-R208 and F106-R42 (compare fig. 3. 2 with fig. 3. 1b). Fragments containing two complete units (dimers) of lengths 235-320 bp were observed with all primer combinations (compare fig. 3. 2 with fig. 3. 1b). Trimers of lengths 360-430 bp were visible with primer pairs F25-R42, F25-R147 and F25-R208 only in some species but as weak bands (see fig. 3. 2).

The most intense bands corresponding to fragments containing monomers and dimers were eluted from gel and cloned into pGEM-T plasmids. According to table 3. 2, fragments corresponding to monomers and dimers were recovered from bands containing PCR products obtained by amplification with primer pairs F25-R42 and F25-R147 in twelve and nine species respectively, while monomers only were recovered from bands containing PCR products obtained by amplification with primer pair F25-R208 in eleven species. Monomers only were recovered with primer pair F106-R42 in six species (*T. monococcum*, *T. tauschii*, *T. aestivum 'Chinese Spring'* and all three *Hordeum* species) while monomers and dimers with primer pair F106-R208 in three species (*Ae. umbellulata*, *S. vavilovii* and *H. brachyantherum*). Cloning efficiency for PCR products amplified by primer combinations F25-R147 and F25-R208 in all *Triticeae* species was similar to the control plates (70-80 white colonies per plate). For PCR products amplified by primer combination F25-R42 the number of white colonies was very low (0-7 per plate) in all *Triticeae* species despite cloning efficiency in the control plates being quite good (40-50 white colonies per plate). Cloning efficiency for PCR products amplified by primer combination F106-R42 in six *Triticeae* species (*T. monococcum*, *T. tauschii*, *T. aestivum 'Chinese Spring'* and all three *Hordeum* species) was similar to the control plates (70-80 white colonies per plate) and cloning efficiency for PCR products amplified by primer combination F106-R208 was fine for *Ae. umbellulata*, and *S. vavilovii* in comparison to control plates (40-50 white colonies per plate) but not for *H. brachyantherum* (1-2 white colonies per plate). Two to three recombinant clones carrying monomers amplified with primer combinations F25-R147, F25-R208, F106-R42 and F106-R208 in most *Triticeae* species, and monomers amplified with primer pair F25/R42 only in some species, were sent to PNACL Sequencing Facility of University of Leicester for sequencing. In addition, five recombinant clones of *L. mollis* corresponding to fragments amplified by primer combination F25/R42 and two to three recombinant clones carrying dimers amplified

with primer pairs F25-R147 and F106-R208 only in few species were sent to PNACL Sequencing Facility of University of Leicester for sequencing.

Overall, 76 clones carrying PCR-fragments isolated from *Triticeae* species with primer pairs were successfully sequenced and homologous to pSc119.2, the majority of them containing mainly single units or incomplete dimers, and only 5 clones having more than one unit (see table 3. 3). Among 76 clones, 7 clones contained a consistently smaller insert (about 80 bp and 175 bp instead of the 200 bp and 320 bp respectively, the lengths expected for monomers and dimers amplified with primer pair F106/R42 or shorter than 155 bp, the length expected for monomers amplified with primer pair F25/R42) and one clone from *S. montanum* contained rearrangements (see table 3. 3). All 8 clones were considered not suitable of further sequence analyses.

3. 4. 2 PCR amplification with single primers and cloning of 120-bp repetitive sequences

To investigate the higher order structure of the 120-bp repeat unit, PCR amplification using single forward primers F25, F106, F10 and single reverse primer R208, R42 and R147 was performed. Each single primer was tested at different annealing temperatures (see table 3. 1). PCR amplification with forward primer F25 resulted in a weak irregular banding pattern variable among species (see fig. 3. 2b). Differences in band length and intensity for a same species were observed with single primer F25 when using 56°C, 60°C or a gradient of annealing temperatures (from 53.6°C, 55.8°C and 58°C). PCR amplification with single primers R42 and R147 resulted in differences in band length and intensity among species (see fig. 3. 2c-d) and within a same species when using 60°C or 66°C as annealing temperatures. When PCR amplifications with single primer F25, R42 and R147 were repeated twice different sizes of PCR products were observed within a same species probably because of a lack in sequence specificity during PCR amplifications with single primers. No matter which annealing temperature used, forward primer F106 and reverse primer R208 used on their own showed a uniform but very weak ladder pattern of bands with strong levels of smears in all species, and forward primer F10 showed only smears in all species.

The most intense bands observed in few species were eluted from gel and cloned into pGEM-T plasmids. Two to three recombinant clones were sent to PNACL Sequencing Facility of University of Leicester for sequencing. With forward primer F25

Table 3. 3. Clone names, fragment directions as sequenced, clone lengths, primer combinations and fragment descriptions of 120-bp repeat unit family clones isolated from diploid and polyploid *Triticeae* species and *Avena sativa* 'Titan'.

Species	Clone name	Fragment direction	Clone length	Primer combination	Fragment description
T. monococcum	T.mono106/42!133	5 -3'	200 bp	106/42	Incomplete dimer
	T.mono106/42!155	5' -3'	202 bp	106/42	Incomplete dimer
	T.mono25/208!1212	3' -5'	185 bp	25/208	Incomplete dimer
T. tauschii	T.tau25/42!2617	3' -5'	127 bp	25/42	Incomplete monomer*
	T.tau25/147!2524	5' -3'	241 bp	25/147	Dimer
	T.tau25/147!2625	3' -5'	123 bp	25/147	Monomer
	T.tau25/147!2726	3' -5'	123 bp	25/147	Monomer
	T.tau106/42!166	3' -5'	83 bp	106/42	Incomplete monomer*
	T.tau106/42!177	3' -5'	175 bp	106/42	Incomplete dimer**
	T.tau25/208!1315	5 -3'	185 bp	25/208	Incomplete dimer
	T.tau25/208!1416	5' -3'	185 bp	25/208	Incomplete dimer
	T.tau25/208!1517	5' -3'	185 bp	25/208	Incomplete dimer
Ae. umbellulata	Ae.umb25/42!091	5 -3'	153 bp	25/42	Incomplete dimer
	Ae.umb25/147!124	5' -3'	122 bp	25/147	Monomer
	Ae.umb25/147!135	5' -3'	122 bp	25/147	Monomer
	Ae.umb25/147!146	5 -3'	122 bp	25/147	Monomer
	Ae.umb25/208!157	5' -3'	185 bp	25/208	Incomplete dimer
	Ae.umb25/208!168	5' -3'	184 bp	25/208	Incomplete dimer
	Ae.umb25/208!179	5' -3'	185 bp	25/208	Incomplete dimer
	Ae.umb106/208!1810	3' -5'	115 bp	106/208	Monomer
	Ae.umb106/208!1911	5' -3'	115 bp	106/208	Monomer
	Ae.umb106/208!2012	3' -5'	115 bp	106/208	Monomer
T. aestivum	CS/3-25/147!2120	5 -3'	122 bp	25/147	Monomer
'Chinese Spring'	CS/3-25/147!2322	5' -3'	241 bp	25/147	Dimer

Table 3. 3. continued

Species	Clone name	Fragment direction	Clone length	Primer combination	Fragment description
T. aestivum **'Chinese Spring'**	CS/3-106/42!111	3' -5'	82 bp	106/42	Incomplete monomer*
	CS/3-106/42!122	3' -5'	82 bp	106/42	Incomplete monomer*
	CS/3-25/208!1618	5 -3'	185 bp	25/208	Incomplete dimer
	CS/3-25/208!1719	5' -3'	185 bp	25/208	Incomplete dimer
	CS/3-25/208!1820	5' -3'	185 bp	25/208	Incomplete dimer
S. cereale **'Petkus Spring'**	Pet22594-25/42!3324	5 -3'	152 bp	25/42	Incomplete dimer
	Pet22594-25/42!3425	5' -3'	152 bp	25/42	Incomplete dimer
	Pet22594-25/147!1918	3' -5'	123 bp	25/147	Monomer
	Petw-25/147!3231	5 -3'	123 bp	25/147	Monomer
	Petw-25/147!123	5' -3'	123 bp	25/147	Monomer
	Petw-25/208!077	3' -5'	185 bp	25/208	Incomplete dimer
	Petw-25/208!088	3' -5'	185 bp	25/208	Incomplete dimer
	Petw-25/208!099	5' -3'	175 bp	25/208	Incomplete dimer
	Pet22594-25/208!2325	5 -3'	191 bp	25/208	Incomplete dimer
	Pet22594-25/208!2426	5' -3'	185 bp	25/208	Incomplete dimer
	Pet22594-25!315	5 -3'	184 bp	F25	Incomplete dimer
	Pet22594-25!337	5' -3'	160 bp	F25	Incomplete dimer***
	Pet22594-25!348	5' -3'	158 bp	F25	Incomplete dimer***
S. vavilovii	S.vav25/208!182	5 -3'	185 bp	25/208	Incomplete dimer
	S.vav25/208!193	5' -3'	185 bp	25/208	Incomplete dimer
	S.vav106/208!204	5 -3'	234 bp	106/208	Dimer
	S.vav106/208!215	3' -5'	233 bp	106/208	Dimer
	S.vav42!237	3' -5'	239 bp	R42	Dimer
	S.vav42!248	3' -5'	225 bp	R42	Incomplete dimer***
	S.vav147!259	3' -5'	217 bp	R147	Incomplete dimer
	S.vav147!2711	3' -5'	228 bp	R147	Incomplete dimer

Table 3. 3. continued

Species	Clone name	Fragment direction	Clone length	Primer combination	Fragment description
S. montanum	S.mon25/147!081	3' -5'	123 bp	25/147	Monomer
	S.mon25/147!103	5 -3'	226 bp	25/147	Dimer***
	S.mon42!136	3' -5'	432 bp	R42	Trimer***
	S.mon147!1811	3' -5'	170 bp	R147	Incomplete dimer***
	S.mon147!169	3' -5'	170 bp	R147	Incomplete dimer***
	S.mon147!1710	3' -5'	170 bp	R147	Incomplete dimer***
L. mollis	L.moll25/42!156	5' -3'	152 bp	25/42	Incomplete dimer
	L.moll25/42!167	3' -5'	154 bp	25/42	Incomplete dimer
	L.moll25/42!189	3' -5'	153 bp	25/42	Incomplete dimer
	L.moll25/42!1910	3' -5'	129 bp	25/42	Incomplete monomer*
A. sativa **'Titan'**	Titan25/42!2920	5 -3'	150 bp	25/42	Incomplete dimer
	Titan25/42!3021	5' -3'	148 bp	25/42	Incomplete dimer
	Titan25/208!1921	5 -3'	185 bp	25/208	Incomplete dimer
	Titan25/208!2022	5' -3'	185 bp	25/208	Incomplete dimer
	Titan25/208!2123	5 -3'	185 bp	25/208	Incomplete dimer
H. bulbosum	H.bulb25/42!101	3' -5'	153 bp	25/42	Incomplete dimer
	H.bulb25/42!123	3' -5'	153 bp	25/42	Incomplete dimer
	H.bulb25/42!134	5' -3'	152 bp	25/42	Incomplete dimer
	H.bulb25/147!1312	3' -5'	125 bp	25/147	Monomer
	H.bulb25/147!1413	5' -3'	123 bp	25/147	Monomer
	H.bulb106/42!065	3' -5'	202 bp	106/42	Incomplete dimer
	H.bulb106/42!076	3' -5'	83 bp	106/42	Incomplete monomer*
	H.bulb147!3711	3' -5'	460 bp	R147	Trimer
	H.bulb25/208!021	5 -3'	119 bp	25/208	Monomer

Table 3. 3. continued

Species	Clone name	Fragment direction	Clone length	Primer combination	Fragment description
H. bulbosum	H.bulb25/208!032	5' -3'	185 bp	25/208	Incomplete dimer
	H.bulb25/208!043	5 -3'	183 bp	25/208	Incomplete dimer
H. chilense	H.chil25/42!145	3' -5'	153 bp	25/42	Incomplete dimer
	H.chil25/147!1514	5 -3'	123 bp	25/147	Monomer
	H.chil25/147!1615	3' -5'	122 bp	25/147	Monomer
	H.chil25/147!1716	5 -3'	123 bp	25/147	Monomer
	H.chil42!235	3' -5'	472 bp	R42	Direct and inverted units****
	H.chil42!352	3' -5'	473 bp	R42	Direct and inverted units****
	H.chil147!4115	3' -5'	484 bp	R147	Trimer
	H.chil147!1716	3' -5'	440 bp	R147	Trimer
H. brachyantherum	H.bra25/42!2112	3' -5'	153 bp	25/42	Incomplete dimer
	H.bra25/42!2213	3' -5'	153 bp	25/42	Incomplete dimer
	H.bra25/42!2415	3' -5'	153 bp	25/42	Incomplete dimer
	H.bra106/42!188	5 -3'	318 bp	106/42	Incomplete trimer
	H.bra106/42!199	3' -5'	201 bp	106/42	Incomplete dimer
	H.bra106/42!1110	3' -5'	201 bp	106/42	Incomplete dimer
	H.bra25/208!054	5 -3'	185 bp	25/208	Incomplete dimer
	H.bra25/208!066	5' -3'	186 bp	25/208	Incomplete dimer

* Incomplete monomers: about 80bp instead of the 200bp expected for monomers amplified with F106/R42; shorter than the 155bp expected for F25/R42.

** Incomplete dimer: about 175 bp instead of the 320 bp expected for dimers amplified with F106/R42.

***Sequences with rearrangements, degenerate repeats.

****Clones with a partial inverted unit, three direct but degenerate repeat units and a partial direct unit.

three clones of *S. cereale* 'Petkus Spring' individual 22594 were homologous to pSc119.2 (see table 3. 3). With reverse primers R147 sequences homologous to pSc119.2 were obtained from *S. vavilovii* (two clones), *S. montanum* (three clones), *H. bulbosum* (one clone containing three complete units) and *H. chilense* (two clones containing three complete units) (see table 3. 3). With reverse primers R42 sequences homologous to pSc119.2 were obtained from *S. vavilovii* (two clones) and *S. montanum* (one clone) and *H. chilense* (two clones) (see table 3. 3). Overall, 16 clones were successfully sequenced and homologous to pSc119.2. Among them, 7 clones (mainly from *S. vavilovii* and *S. montanum*) contained rearrangements or degenerate repeats (see table 3. 3); 2 clones isolated from *H. chilense* with reverse primer R42 (H.chil42!235 and H.chil42!352) have a partial inverted unit, three direct but degenerate repeat units and a partial direct unit (see table 3. 3 and fig. 3. 3). All 9 clones were considered not suitable of further sequence analyses.

Finally, additional 6 sequences amplified by PCR with forward primer F25 in *T. monococcum*, *T. tauschii* and *H. bulbosum* were found to be gypsy-related sequences and additional 8 sequences amplified by PCR with reverse primer R147 in *T. monococcum*, *S. cereale* 'Petkus Spring' and *T. aestivum* 'Chinese Spring' were mainly related to gypsy retroelements. The amplification of gypsy-related sequences with single primers F25 and R147 possibly explained the differences observed in their amplification patterns when PCR reactions were repeated twice or at different annealing temperatures.

3. 4. 3 Sequence analysis

A total of 75 clones homologous to the original pSc119.2 sequence (68 obtained by PCR amplification with primer pairs and 7 with single primers) containing a complete or near complete 120-bp repeat unit or in few cases more than one unit (dimers and trimers) were selected for sequence analyses. In total, 87 single 120-bp repeat units from *Hordeum* species, *Secale* species, *Triticum/Aegilops* species, *L. mollis*, and *A. sativa* 'Titan' were analysed in more detail. A total of 31 units originated from *Hordeum* species (fig. 3. 4a), 22 from *Secale* species (fig. 3. 4b) and 26 from *Triticum/Aegilops* species (fig. 3. 4c) were aligned by using Jalview Multiple Alignment Editor V 1.3 (EMBL-EBI Version). The lengths of 120-bp sequence monomers showed relatively little variation (*Secale*: 117-119 bp; *Hordeum*: 116-119 bp; *Aegilops/Triticum*:

H.chiR42!235

CCCCGGGGTGCGTTTACGTGTCCGTCGTCAACACTCACAGTTTTGGCCAATTTTGACCCGTTTCGTGGACCAAAATTAGTAGGTAATAGCATAGAAAACTGGCCAAAATTGACCAA
ATCTACGAGTTTTGACGAGGTCCCTGTAACCGGACCTCGGGGTTCCCCGAACGTACGGATCGCCCCGGGACCCAACATCAGTGAGTAATAGCATAGAAAACTGATCAGAATTGACC
AAATCTGCGAGTTTTGACGAGGTCCCCGTAATCGGACTATGGGGTTCCCCGGACGTTCGGATCGCCCCGGGACCAAAAATCAATGGGTAATAGAACAGAAAACTGGTCAGAATGGA
TCAAATCTGCGAGTTTTGACGACGTCCCCGTTATCGGACACCGGGGGGTTTNCCGAATGTTCGGAACGCCCTGGGGACCCAAATATGCGACTTTTTACGAGGTCCCCGTAAACGCAC
CCCNGGGG

Rep1 orientation 3'-5'
CCACGAAACGGGTCAAAATTGGCCAAAACTGTGAGTGTTGACGACGGACACGTAAACGCACCCCGGGG
Rep2 orientation 5'-3'
ACCAAAATTAGTAGGTAATAGCATAGAAAACTGGCCAAAATTGACCAAATCTACGAGTTTTGACGAGGTCCCTGTAACCGGACCTCGGGGTTCCCCGAACGTACGGATCGCCCCGGG
Rep3 orientation 5'-3'
ACCCAACATCAGTGAGTAATAGCATAGAAAACTGATCAGAATTGACCAAATCTGCGAGTTTTGACGAGGTCCCCGTAATCGGACTATGGGGTTCCCCGGACGTTCGGATCGCCCCGGG
Rep4 orientation 5'-3'
ACCAAAAATCAATGGGTAATAGAACAGAAAACTGGTCAGAATGGATCAAATCTGCGAGTTTTGACGACGTCCCCGTTATCGGACACCGGGGGGTTTNCCGAATGTTCGGAACGCCCTGGGG
Rep5 orientation 5'-3'
ACCCAAATATGCGACTTTTTACGAGGTCCCCGTAAACGCACCCCNGGGG

H.chiR42!352

CCCCGGGGTGCGTTTACGTGTCCGTCGTCAACACTCACAGTTTTGGCCAATTTTGACCCGTTTCGTGGACCAAAATTAGTAGGTAATAGCATAGAAAACCGGCCAAAATTGACCAA
ATCTACGAGTTTTGACGAGGTCCCTGTAACCGGACCTCGGGGTTCCCCGAACGTACGGATCGCCCNGGGACCCAACATCAGTGAGTAATAGCATANAAAACTGATCAGAATTGACC
AAATCTGCGAGTTTTGACGAGGTCCCCGTAATCGGACTATGGGGTTCCCCGAACGTTCGGATCGCCCCGGGACCAAAAATCAATGGGTAATAGAACAGAAAACTGGTCATNAATGG
ATCAAATCTGCGAGTATTGACNACGTCCCCGTTATCGGACACCGGGGGGTTTCCCGAATGTTCNGAACGCCCTGGGACCCAAATATGCGACTTTTTACGAGGTCCCCNGTAACGCA
CCCCNGGGG

Rep1 orientation 3'-5'
CCACGAAACGGGTCAAAATTGGCCAAAACTGTGAGTGTTGACGACGGACACGTAAACGCACCCCGGGG
Rep2 orientation 5'-3'
ACCAAAATTAGTAGGTAATAGCATAGAAAACCGGCCAAAATTGACCAAATCTACGAGTTTTGACGAGGTCCCTGTAACCGGACCTCGGGGTTCCCCGAACGTACGGATCGCCCNGGG
Rep3 orientation 5'-3'
ACCCAACATCAGTGAGTAATAGCATANAAAACTGATCAGAATTGACCAAATCTGCGAGTTTTGACGAGGTCCCCGTAATCGGACTATGGGGTTCCCCGAACGTTCGGATCGCCCCGGG
Rep4 orientation 5'-3'
ACCAAAAATCAATGGGTAATAGAACAGAAAACTGGTCATNAATGGATCAAATCTGCGAGTATTGACNACGTCCCCGTTATCGGACACCGGGGGGTTTCCCGAATGTTCNGAACGCCCTGGG
Rep5 orientation 5'-3'
ACCCAAATATGCGACTTTTTACGAGGTCCCCNGTAACGCACCCCNGGGG

Figure 3.3 120-bp repeat unit family sequences from two clones of *H. chilense* amplified by reverse primer R42 (H.chiR42!235 and H.chiR42!352). A partial inverted unit (red), three direct but degenerate repeat units (blue) and a partial direct unit (blue) are present in both clones. Primer location is underlined (red and blue).

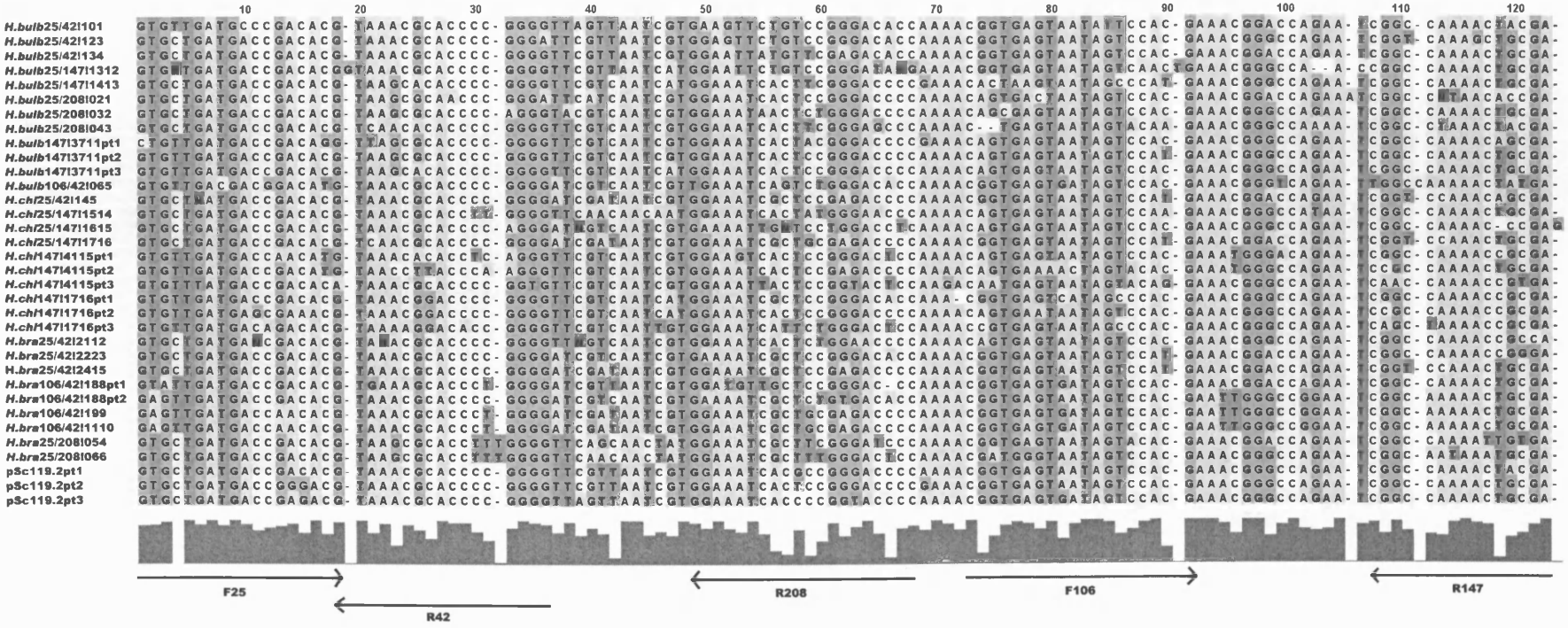


Figure 3. 4a. DNA sequence alignment of 120-bp repeat family sequences of *Hordeum* species using Jalview Multiple Alignment Editor. The three units of pSc119.2 sequence are shown in the alignment. The 5' end of primer F25 was used as arbitrary start site. Each nucleotide is shaded in a different grey level and gaps are introduced for optimal alignment or to show deletions. Identity scores for each base are given as bars at the bottom (large bar represents high similarity) and primers are indicated below the alignment. Clone names give abbreviated species names followed by primer combination and a sequential number.

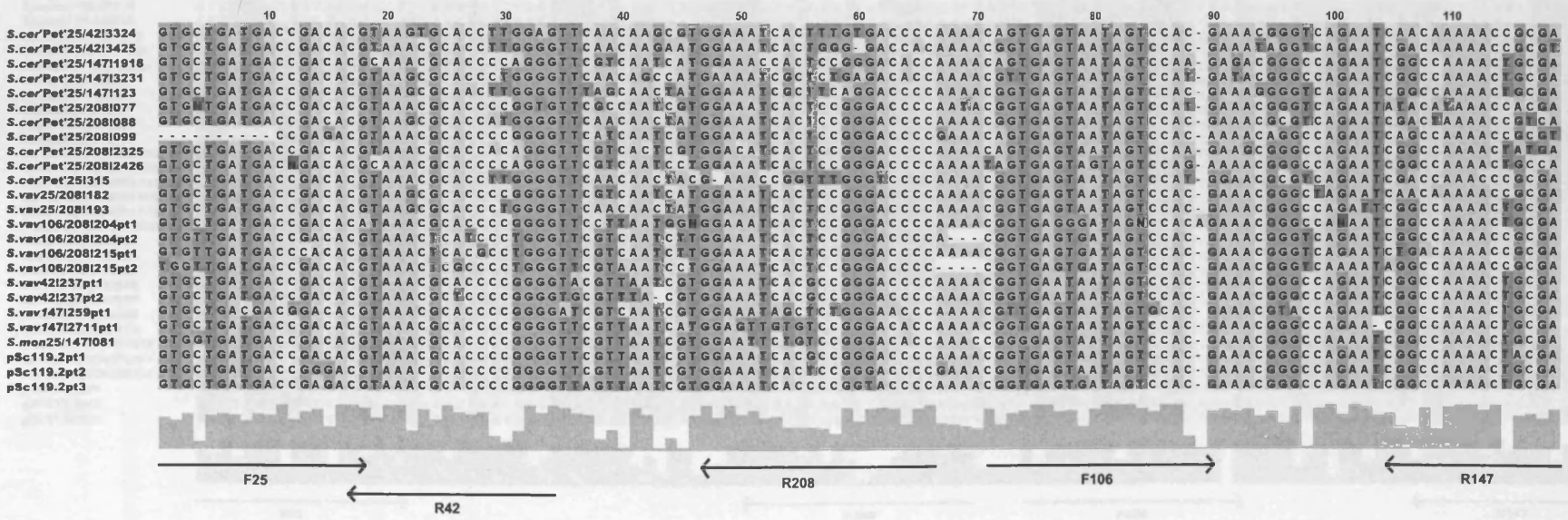


Figure 3. 4b. DNA sequence alignment of 120-bp repeat family sequences of *Secale* species using Jalview Multiple Alignment Editor. The three units of pSc119.2 sequence are shown in the alignment. The 5' end of primer F25 was used as arbitrary start site. Each nucleotide is shaded in a different grey level and gaps are introduced for optimal alignment or to show deletions. Identity scores for each base are given as bars at the bottom (large bar represents high similarity) and primers are indicated below the alignment. Clone names give abbreviated species names followed by primer combination and a sequential number.

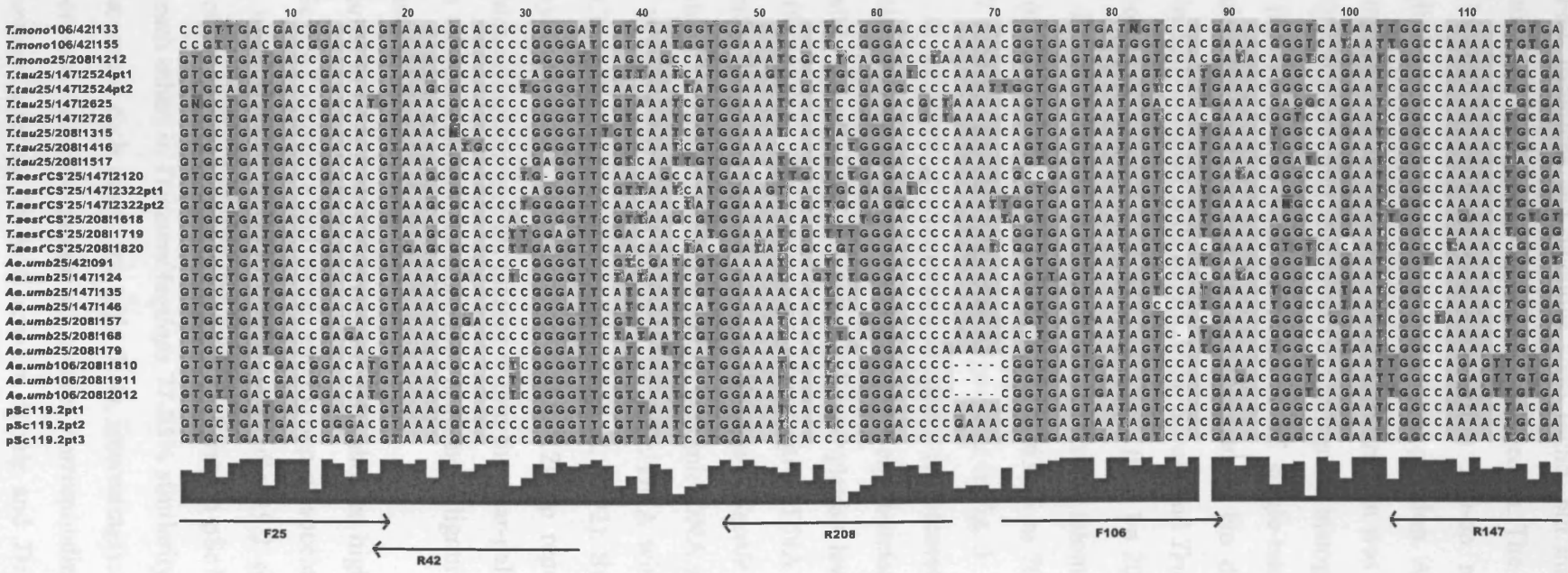


Figure 3. 4c. DNA sequence alignment of 120-bp repeat family sequences of *Triticum/Aegilops* species using Jalview Multiple Alignment Editor. The three units of pSc119.2 sequence are shown in the alignment. The 5' end of primer F25 was used as arbitrary start site. Each nucleotide is shaded in a different grey level and gaps show deletions. Identity scores for each base are given as bar at the bottom (large bar represents high similarity) and primers are indicated below the alignment. Clone names give abbreviated species names followed by primer combination and a sequential number.

117-118 bp). Few insertions and deletions were found, but mainly single nucleotide mutations keeping the 120-bp repeat unit a constant length. The use of multiple primer combinations located in different positions within 120-bp repeat unit minimized sequence selectivity because of primer sequence and position. As a result, the strategy used in PCR amplification showed that nucleotide variation was distributed throughout the length of 120-bp repeat unit sequences (see consensus histograms below individual sequences in fig. 3. 4a-b-c) with no hotspots of single-base variation, and few distributed nucleotide positions had no consensus base. No distinctive motifs were identified within 120-bp repeat units of *Hordeum*, *Secale* and *Triticum/Aegilops* groups, although a region of 20 bases $A_2CGCAC_4G_4T_2CGT_2$ from bp 20-21 to 40-42 represent an imperfect inverted repeat (palindrome) and two telomere-similar sequences CCCCAAAA or GGCCAAAA are located from bp 63-65 to 70-72 and from bp 105-109 to 112-117 respectively (see sequence alignments in fig. 3. 4a-b-c). As previously reported, CG dinucleotide combination is under-represented in some repetitive sequences families but not in others where it is relatively abundant (Macas, Meszaros et al. 2002). Possibly the resulting different DNA methylation levels are functionally or structurally differentiating the two groups of repetitive DNA sequences. In 120-bp repeat unit family the consensus sequence in *Hordeum*, *Secale* and *Triticum/Aegilops* groups is GC-rich (50-55%) compared to cereal genomic DNA with 43.5% GC amount (Swanson and Webster 1975) and to wheat genomic DNA with 45-48% GC amount (Wagner and Capesius 1981; Montero, Filipski et al. 1992). Symmetrical CCGG and CCNGG methylation sites were identified within 120-bp repeat units of *Hordeum*, *Secale* and *Triticum/Aegilops* groups located within the near-palindromic sequence and before the first telomere-similar sequence (see sequence alignments in fig. 3. 4a-b-c), and will be described more in detail in chapter 4.

As reported in table 3. 4, 120-bp repeat units showed high levels of similarity to the original pSc119.2 sequence and within each group of species. In *Hordeum* species 120-bp repeat units showed 80-93% similarity to pSc119.2 sequence and 73-100% amongst each other, in *Secale* species 77-97% similarity to pSc119.2 sequence and 74-97% amongst each other, in *Triticum/Aegilops* 77-95% similarity to pSc119.2 sequence and 70-100% amongst each other (see table 3. 4). Interestingly, homology levels were similar also between the three groups of species corresponding to 72-95% between *Secale* and *Hordeum* species, 71-96% between *Secale* and *Triticum/Aegilops* species and 69-94% between *Hordeum* and *Triticum/Aegilops* species (see table 3. 4). Finally,

Table 3. 4. Percentage of similarity of 120-bp repeat unit family sequences within and between species of *Hordeum, Secale, Triticum/Aegilops,* as well as within and between *Leymus mollis, Avena sativa* 'Titan' and all *Triticeae* species.

Species	*Hordeum*	*Secale*	*Triticum/Aegilops*	*L. mollis*	*A. sativa* 'Titan'
pSc119.2	80-93%	77-97%	77-95%	-	-
Hordeum	73-100%	72-95%	69-94%	69-94% to all *Triticeae* species	68-94% to all *Triticeae* species
Secale	-	74-97%	71-96%		
Triticum/Aegilops	-	-	70-100%		
L. mollis	69-94% to all *Triticeae* species			79-94%	-
A. sativa	68-94% to all *Triticeae* species			-	77-90 %

the similarity between three 120-bp repetitive sequences of *L. mollis* ranged from 79-94% and it was 69-94% to the *Triticeae* sequences and between five 120-bp repetitive sequences of *A. sativa* 'Titan' (not shown and not included in further phylogenetic analyses) ranged from 77-90 % and it was 68-94% to the *Triticeae* sequences (see table 3. 4). The overall results suggested similarity levels only slightly higher within than between species and no species-specific variants of 120-bp repeat unit family.

3. 4. 4 Phylogenetic analysis

The phylogeny of 120-bp repeat family was analysed by the Neighbour Joining method with Kimura's two parameters. The analysis of all 87 single 120-bp repeat units from *Hordeum* species, *Secale* species, *Triticum/Aegilops* species, *L. mollis*, and *A. sativa* 'Titan' revealed no taxonomically related clustering of 120-bp repeat unit family sequences (not shown). For illustrating purposes 120-bp repeat family sequences were separated into two groups: sequences from *Hordeum* species (fig. 3. 5a) and sequences from *Secale*, *Triticum* and *Aegilops* species together with *L. mollis* (fig. 3. 5b). Overall, no strong groupings of 120-bp repeat family sequences were evident in either of the two groups: the trees were deeply branched with low bootstrap support for the first nodes (see fig. 3. 5a-b). In addition, most clusters had low bootstrap values, only few clusters having bootstrap values higher than 800 (see fig. 3. 5a-b). Only few sequences clustered significantly, sometimes belonging to the same species and cloned as dimers, sometimes resulting from the same primer combination, but not always so. Due to the use of multiple primer combinations located in different positions within 120-bp repeat unit no notable groupings of sequences were based on primers. In general, sequences from a given species, primer pair or single primer were distributed throughout the trees with no relation to ploidy level (*T. tauschii, T. monoccoccum* versus *T. aestivum*), cultivation (cultivated rye *S. cereale* versus wild rye *S. vavilovii*), breeding system (rye being an obligate out-breeder, while barley and wheat are inbreeding and largely cleistogamous) or sequence copy number within a genome.

3. 4. 5 *In situ* hybridization analysis

A selection of 9 clones (see table 3. 5) with homology levels between 78-96% has been used for the *in situ* hybridization experiments discussed in this chapter, but

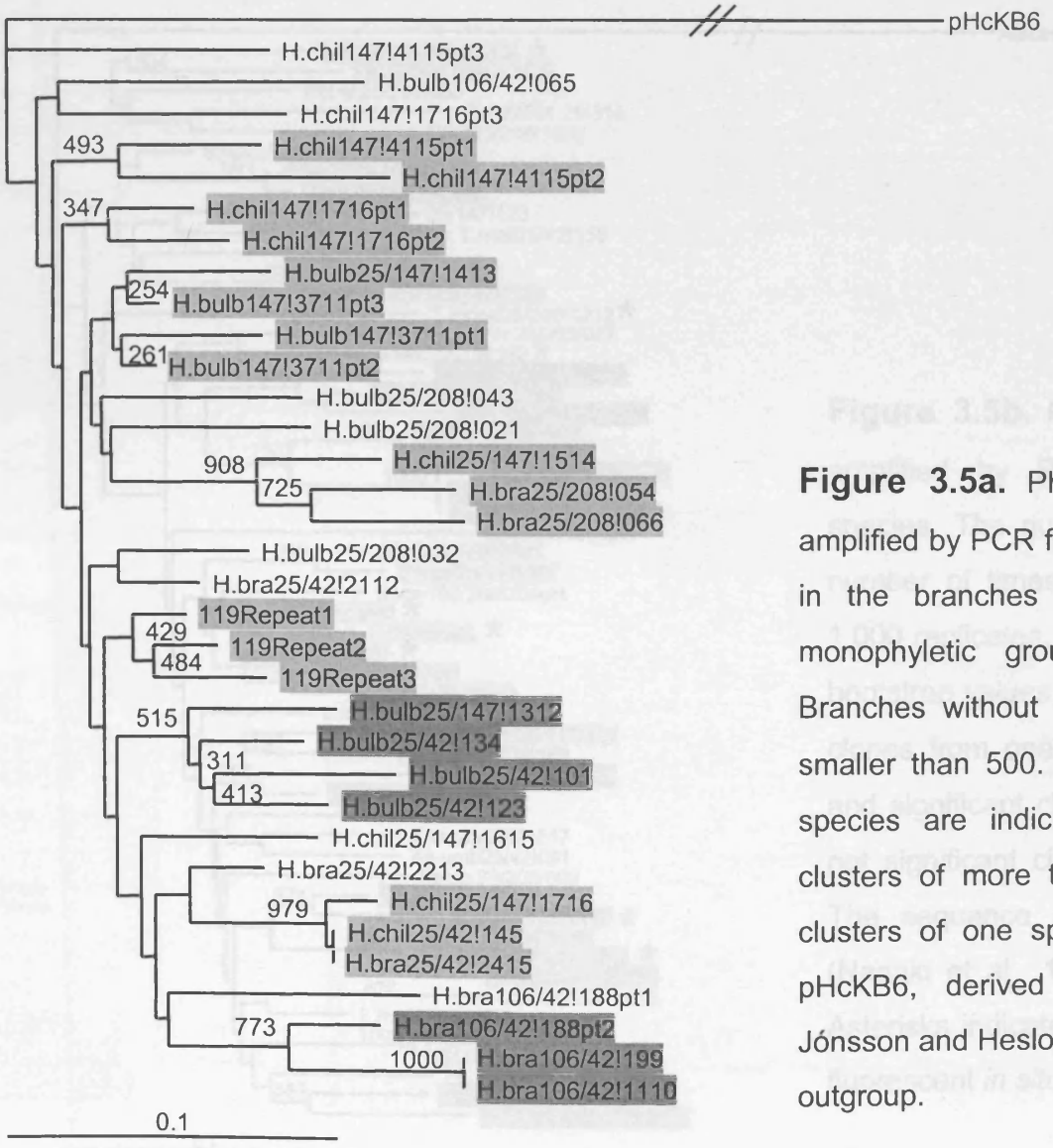


Figure 3.5a. Phylogenetic tree of 120-bp repeats amplified by PCR from *Hordeum* species. The numbers in the branches indicate the number of times a monophyletic group occurred in 1,000 replicates. Branches without numbers received bootstrap values smaller than 500. Clusters including clones from one species are indicated in dark grey and significant clusters of more than one species or not significant clusters of one species in light grey. The sequence pHcKB6, derived from *H. chilense* (Anamthawat-Jónsson and Heslop-Harrison, 1993) was chosen as the outgroup.

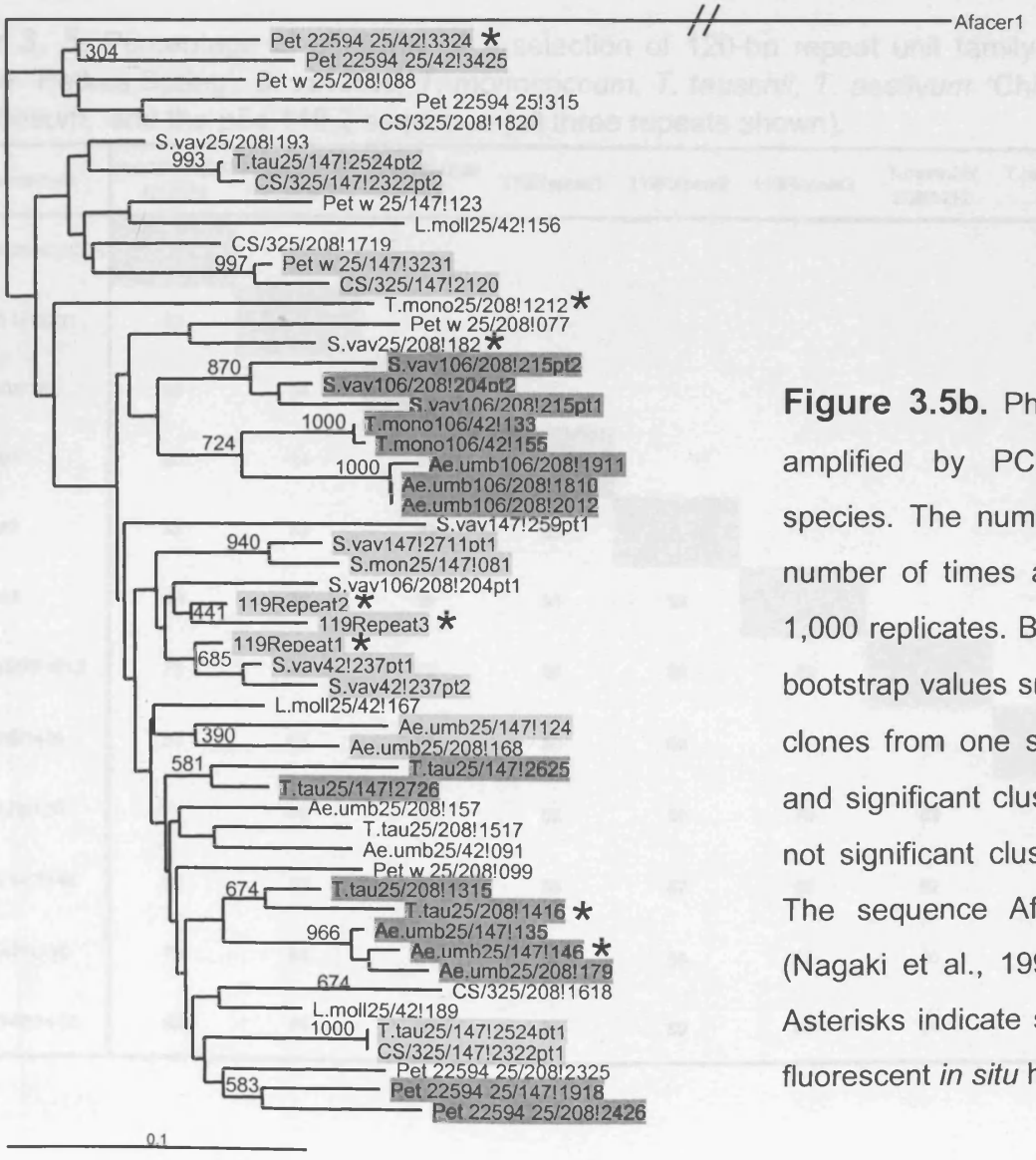


Figure 3.5b. Phylogenetic tree of 120-bp repeats amplified by PCR from *Secale/Triticum/Aegilops* species. The numbers in the branches indicate the number of times a monophyletic group occurred in 1,000 replicates. Branches without numbers received bootstrap values smaller than 500. Clusters including clones from one species are indicated in dark grey and significant clusters of more than one species or not significant clusters of one species in light grey. The sequence Afa-cer1, derived from *S. cereale* (Nagaki et al., 1998) was chosen as the outgroup. Asterisks indicate selected clones used as probes in fluorescent *in situ* hybridization (FISH).

Table 3. 5. Percentage of similarity of a selection of 120-bp repeat unit family clones within and between species of *S. cereale* 'Petkus Spring', *S. vavilovii, T. monococcum, T. tauschii, T. aestivum* 'Chinese Spring', *Ae. umbellulata, H. chilense, H. bulbosum,* and the pSc 119.2 sequence (all three repeats shown).

Clone names	Pet(22594)25/42!3324	Pet(w)25/147!3231	S.vav25/208!182	119Repeat1	119Repeat2	119Repeat3	T.mono25/208!1212	T.tau25/208!1416	CS/325/147!2120	Ae.umb25/147!146	H.chil25/147!1716	H.bulb25/147!1413
Pet(22594)25/42!3324												
Pet(w)25/147!3231	83											
S.vav25/208!182	88	84										
119Repeat1	83	84	91									
119Repeat2	82	83	90	95								
119Repeat3	80	82	88	94	94							
T.mono25/208!1212	78	86	82	86	83	83						
T.tau25/208!1416	81	83	86	89	88	87	79					
CS/325/147!2120	81	96	82	82	81	79	83	82				
Ae.umb25/147!146	80	83	86	88	87	85	82	89	82			
H.chil25/147!1716	79	84	84	89	88	87	80	84	82	83		
H.bulb25/147!1413	82	84	88	88	89	86	81	88	82	88	84	

also for Southern hybridization experiments and immunostaining experiments combined with *in situ* hybridization described in the next chapters. Clones belonging to 120-bp repeat unit family from *S. cereale* 'Petkus Spring', *S. vavilovii*, *Ae. umbellulata*, *T. monococcum* and *T. tauschii* (see asterisks in fig. 3. 5b and table 3. 5) were used as probes. In addition, pSc119.2 clone and 5S rDNA pTa794 clone were also used as probes. Double target *in situ* hybridization with different and reciprocal combinations of probes was carried out under high stringency conditions with digoxigenin-labelled probes detected by FITC (green fluorescence) and biotin-labelled probes detected by Alexa 594 (red fluorescence), chromosomes were counterstained with DAPI (blue fluorescence). Sites with overlapping green and red signals appeared yellow.

In *S. cereale* 'Petkus Spring', large and prominent bands corresponding to subtelomeric heterochromatic blocks were detected at most chromosome arms with all probes used (an example is given with *T. tauschii* and *S. cereale* 'Petkus Spring' clones, red and green signals respectively, in fig. 3. 6b-c). After chromosomes were identified and karyotypes established as described before (Cuadrado and Jouve 1994; Cuadrado, Ceoloni et al. 1995) the large subterminal sites were identified as chromosome arms 1RS, 1RL, 2RS, 2RL, 3RS, 4RS, 5RS, 6RS, 7RS (see karyotypes in figs. 3. 7 and in 3. 8). Karyotypes were used for detailed comparison of location and strength of hybridization sites detected with pSc119.2, *S. cereale* 'Petkus Spring' and *T. monococcum* clones (see figs. 3. 7 and 3. 8). Different 120-bp repeat unit family clones occupied the same chromosomal sites with only few differences in signal strength of minor intercalary bands independently of probe combinations used. For example, the intercalary band of 1RS is more green and the intercalary double band of 7RL is more red in fig. 3. 7, while the intercalary band of 1RS and the intercalary double band of 7RL are more red in fig. 3. 8. Overall it seemed that chromosomal sites harboured many variants of the 120-bp repeat unit family. In all experiments, additional weaker hybridization along the euchromatin was present possibly suggesting that sequences homologous to 120-bp repeat were dispersed throughout the genome of *S. cereale* 'Petkus Spring'.

In diploid wheat species *T. urartu* (A genome) and *T. tauschii* (D genome) 120-bp repeat unit family sequences showed telomeric locations. By using clones from *S. cereale* 'Petkus Spring', *T. monococcum* and *T. tauschii* on chromosome spreads of wild *T. urartu* (see fig. 3. 9) and *T. tauschii* (see fig. 3. 10) two small telomeric sites (with two additional weak telomeric sites not always visible) and on four telomeric sites

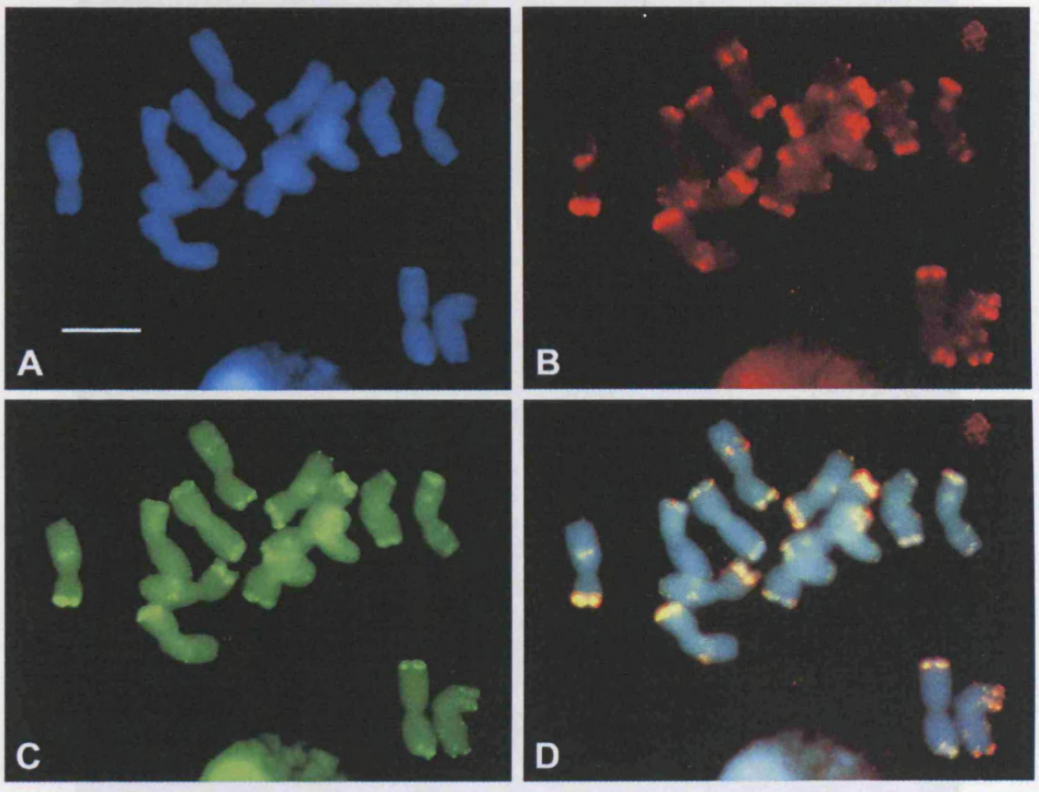


Figure 3. 6. Root tip metaphase of *S. cereale* 'Petkus Spring' (R genome, 2n=14). A) DAPI stained chromosomes. B) *In situ* hybridization with biotin-labelled Ttau25/208!1416 probe detected with Alexa 594 red fluorescence. C) *In situ* hybridization with digoxigenin -labelled Pet(22594)25/42!3324 probe detected with FITC green fluorescence. D) Overlay of A, B and C images. Yellow fluorescence indicates colocalization of green and red signals. Scale bar represents 10 μm.

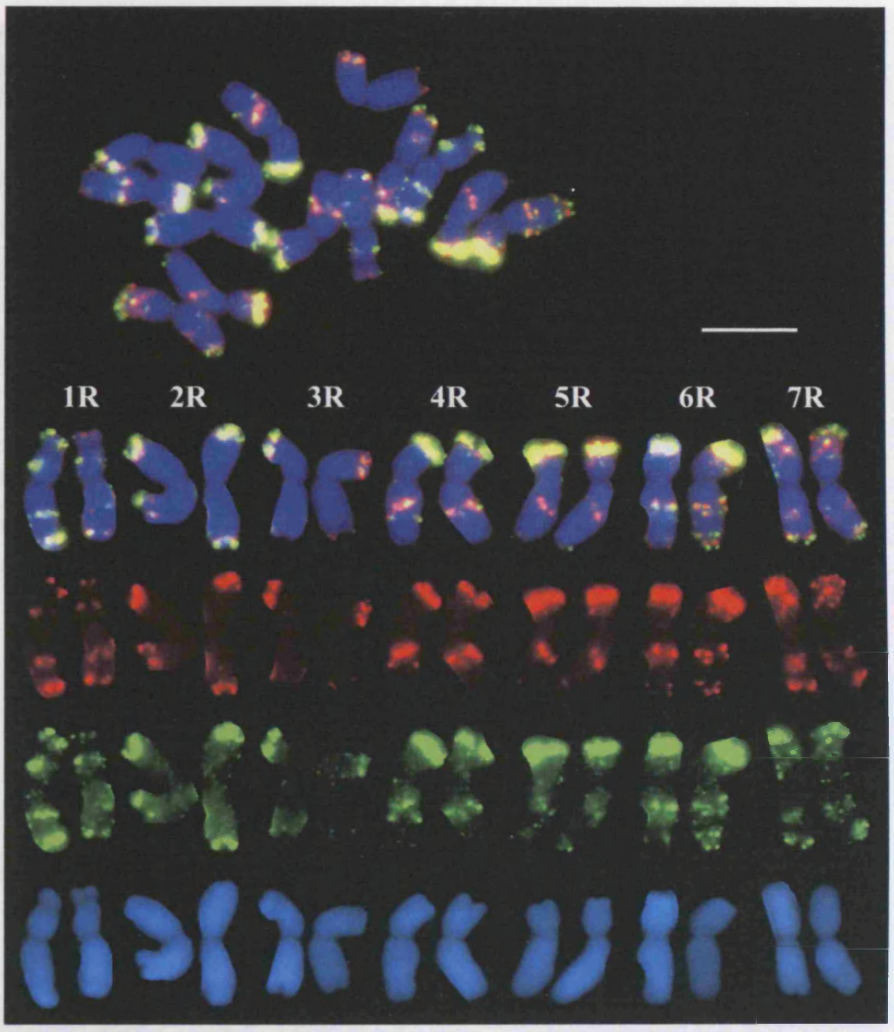


Figure 3. 7. Root tip metaphase of *S. cereale* 'Petkus Spring' (R genome, 2n=14). Chromosomes were identified according to Cuadrado and Jouve (1994) and Cuadrado *et al.* (1995), and arranged as karyotype with overlayed and separate channels displayed. I*n situ* hybridization with biotin-labelled Pet(22594)25 /42!3324 probe was detected with Alexa 594 red fluorescence and with digoxigenin-labelled pSc119.2 probe with FITC green fluorescence. Yellow fluorescence indicates co-localization of green and red signals. Scale bar represents 10 µm.

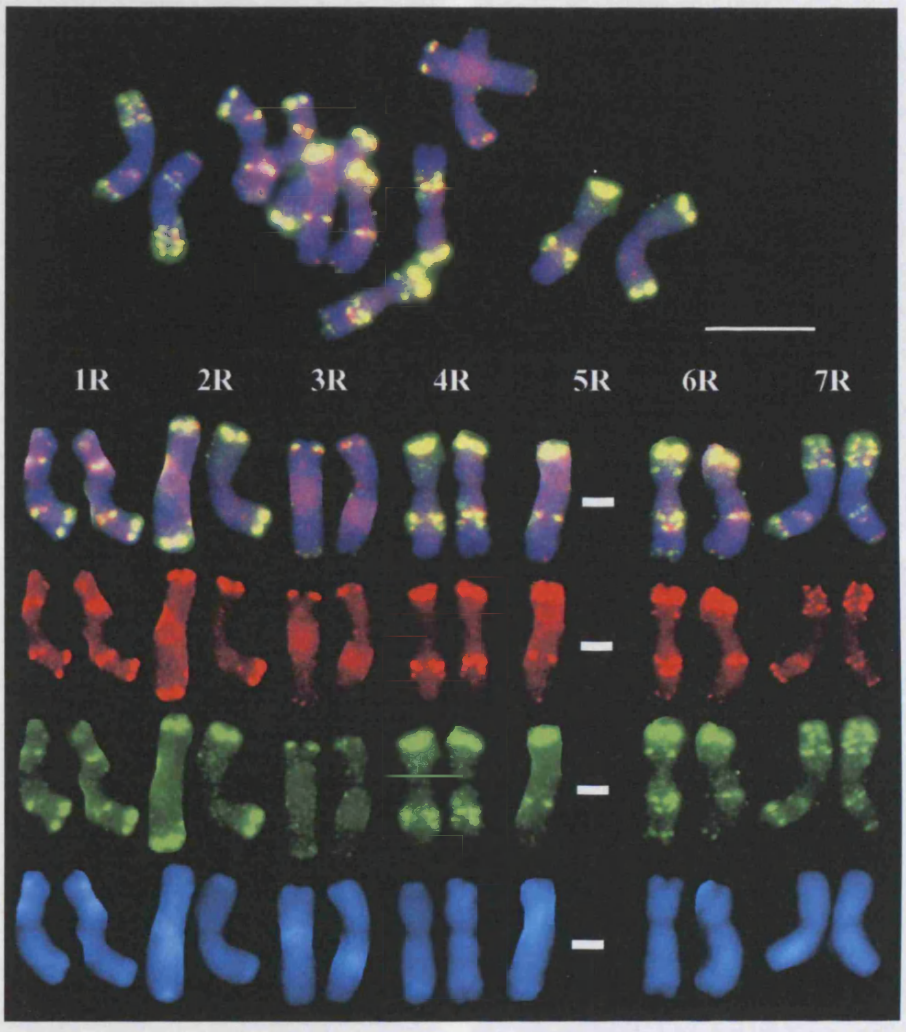


Figure 3. 8. Root tip metaphase of S. cereale 'Petkus Spring' (R genome, 2n=14; one chromosome 5R is missing). Chromosomes were identified according to Cuadrado and Jouve (1994) and Cuadrado *et al.* (1995), and arranged as karyotype with overlayed and separate channels displayed. *In situ* hybridization with biotin-labelled Tmo25/208!1212 probe was detected with Alexa 594 red fluorescence and with digoxigenin-labelled pSc119.2 probe with FITC green fluorescence. Yellow fluorescence indicates co-localization of green and red signals. Scale bar represents 10 μm.

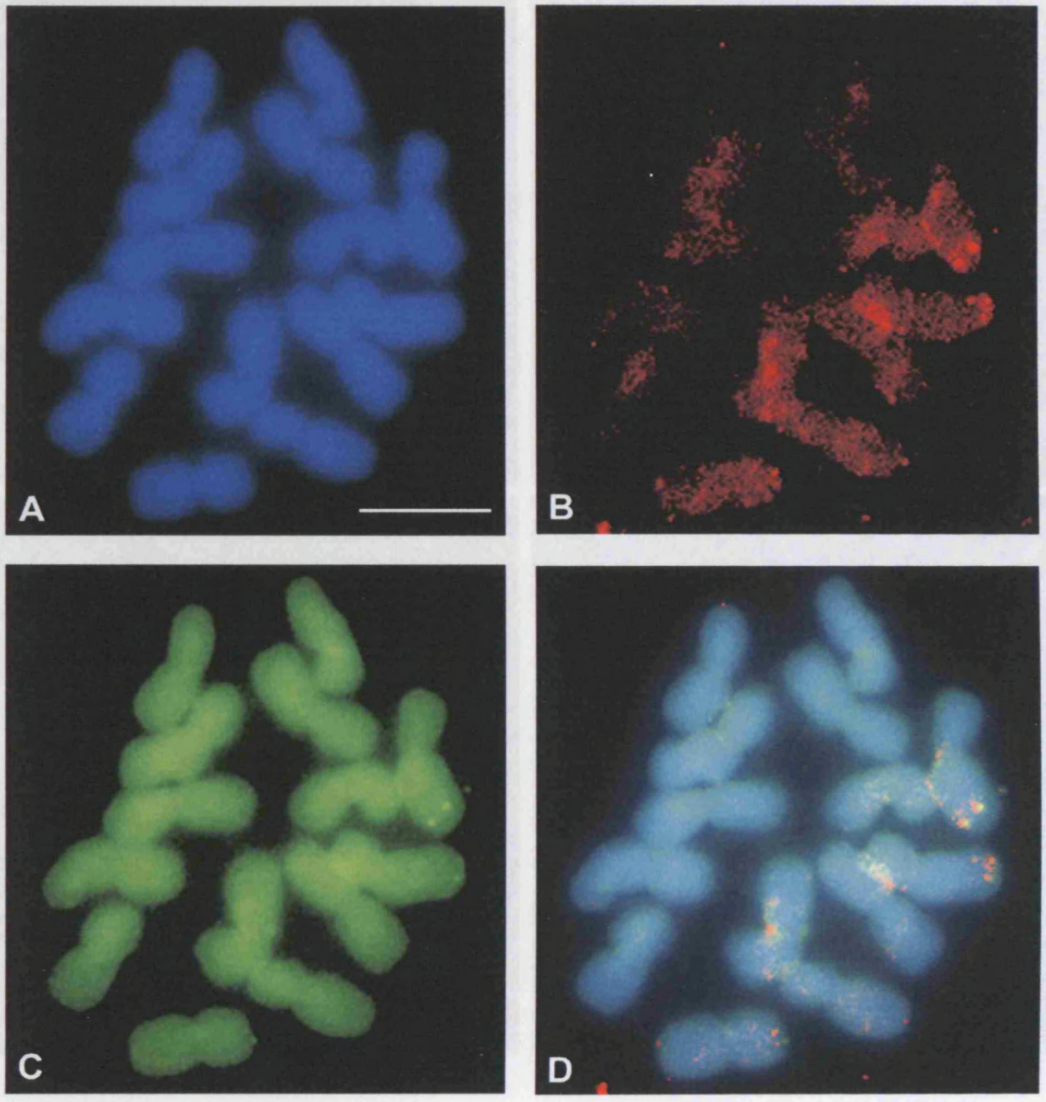


Figure 3. 9. Root tip metaphase of *T. urartu* (A genome, 2n=14). A) DAPI stained chromosomes. B) *In situ* hybridization with biotin -labelled Tmono25/208!1212 probe detected with Alexa 594 red fluorescence. C) *In situ* hybridization with digoxigenin-labelled Pet (22594)25/42!3324 probe detected with FITC green fluorescence. D) Overlay of A, B and C images. Yellow fluorescence indicates co-localization of green and red signals. Scale bar represents 10 μm.

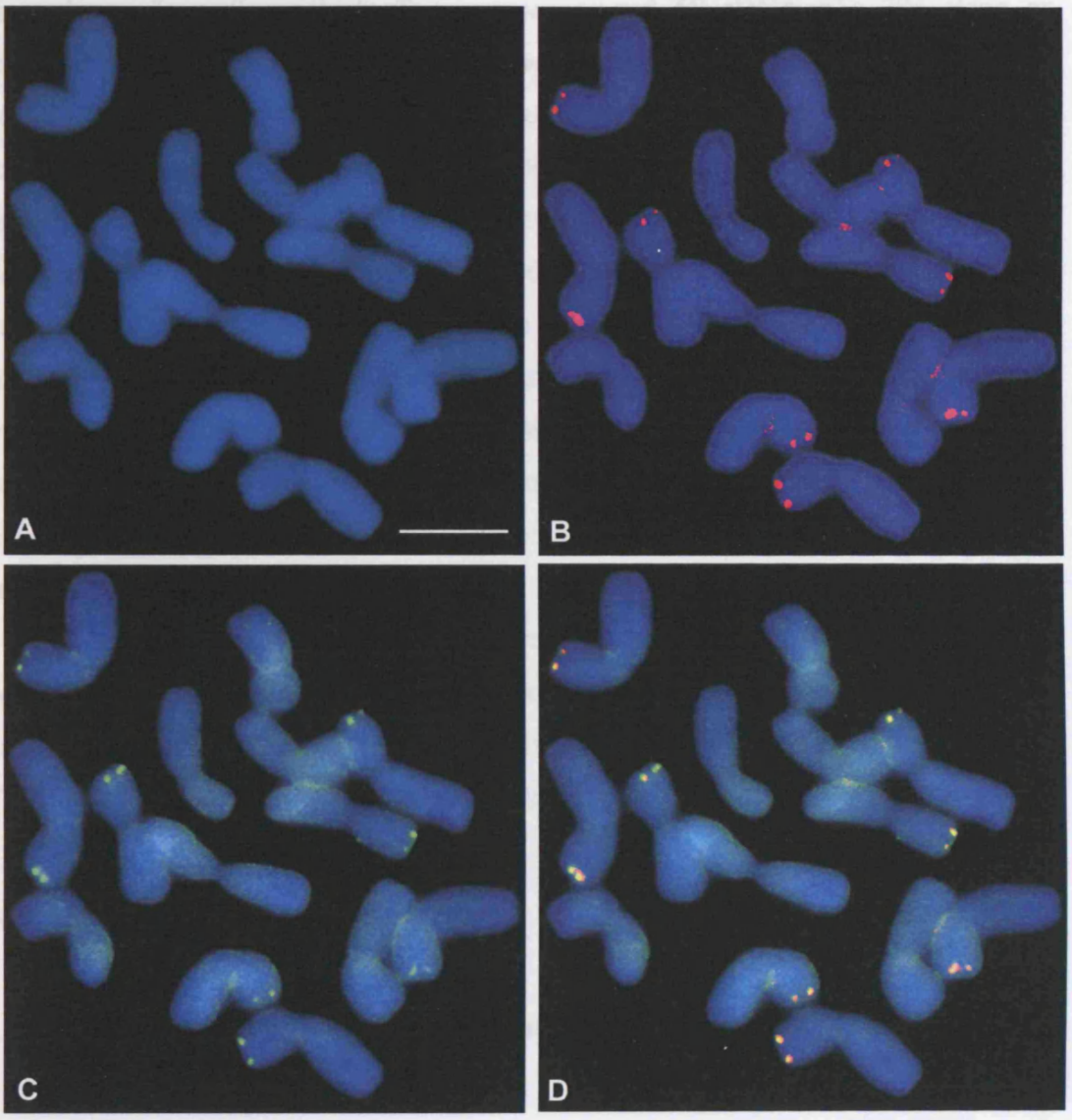


Figure 3. 10. Root tip metaphase of *T. tauschii* (D genome, 2n=14). A) DAPI stained chromosomes. B) *In situ* hybridization with biotin-labelled Pet(22594)25/42!3324 probe detected with Alexa 594 red fluorescence. C) *In situ* hybridization with digoxigenin-labelled Ttau25/208!1416 probe detected with FITC green fluorescence. D) Overlay of A, B and C images. Yellow fluorescence indicates co-localization of green and red signals. Scale bar represents 10 μm.

were observed respectively, according to previous descriptions of pSc119.2 clone in A and D chromosomes of hexaploid wheat *T. aestivum* (Mukai, Nakara et al. 1993). By using clones from *S. vavilovii*, *T. monococcum* and 5S rDNA pTa 794 clone on chromosome spreads of tetraploid *T. turgidum* ssp. *durum* (AB genome) (see figs. 3. 11 and 3. 12) and by using clones from *S. vavilovii*, *T. tauschii* and 5S rDNA pTa 794 clone on chromosome spreads of hexaploid wheat *T. aestivum* 'Chinese Spring' (ABD genome) (see figs. 3. 13 and 3. 14) 120-bp repeat unit family showed locations in A, B and D chromosomes similar to those occupied by pSc119.2 clone in hexaploid wheat *T. aestivum* (Mukai, Nakara et al. 1993). In both polyploid wheat species the 120-bp repeat unit family sequences formed interstitial bands together with subtelomeric ones in B genome while A and D genomes showed few telomeric sites as observed in diploid *T. urartu* and *T. tauschii* species. In all cases no prominent signal differences were detected on chromosomal sites by using different clones belonging to 120-bp repeat unit family as probes.

The overall results obtained by fluorescent *in situ* hybridization (FISH) in rye, diploid and polyploid wheat species suggested that possibly chromosomal sites contain many variants of the 120-bp repeat unit family independently of the varying abundance of 120-bp repeat unit family sequences in their genomes.

3. 5 Discussion

3. 5. 1 Diversity of 120-bp repeat unit family

By PCR amplification with primer pairs and single primers designed from the original pSc119.2 sequence of rye, about 90 members of 120-bp repeat unit family have been isolated, cloned and sequenced from diploid and polyploid *Triticeae* species. The lengths of 120-bp sequence monomers showed relatively little variation (*Secale*: 117-119 bp; *Hordeum*: 116-119 bp; *Aegilops/Triticum*: 117-118 bp) suggesting few insertions and deletions occurring throughout the sequence (fig. 3. 4). Nucleotide variation was distributed throughout the length of the sequence with no hotspots of single-base variation, and a few distributed nucleotide positions had no consensus base (fig. 3. 4). Sequence analysis of 31 units from *Hordeum* species, 22 from *Secale* species and 26 from *Triticum/Aegilops* species showed high levels of similarity to pSc119.2 sequence and within each group of species (73-100% in *Hordeum* species, 74-97% in

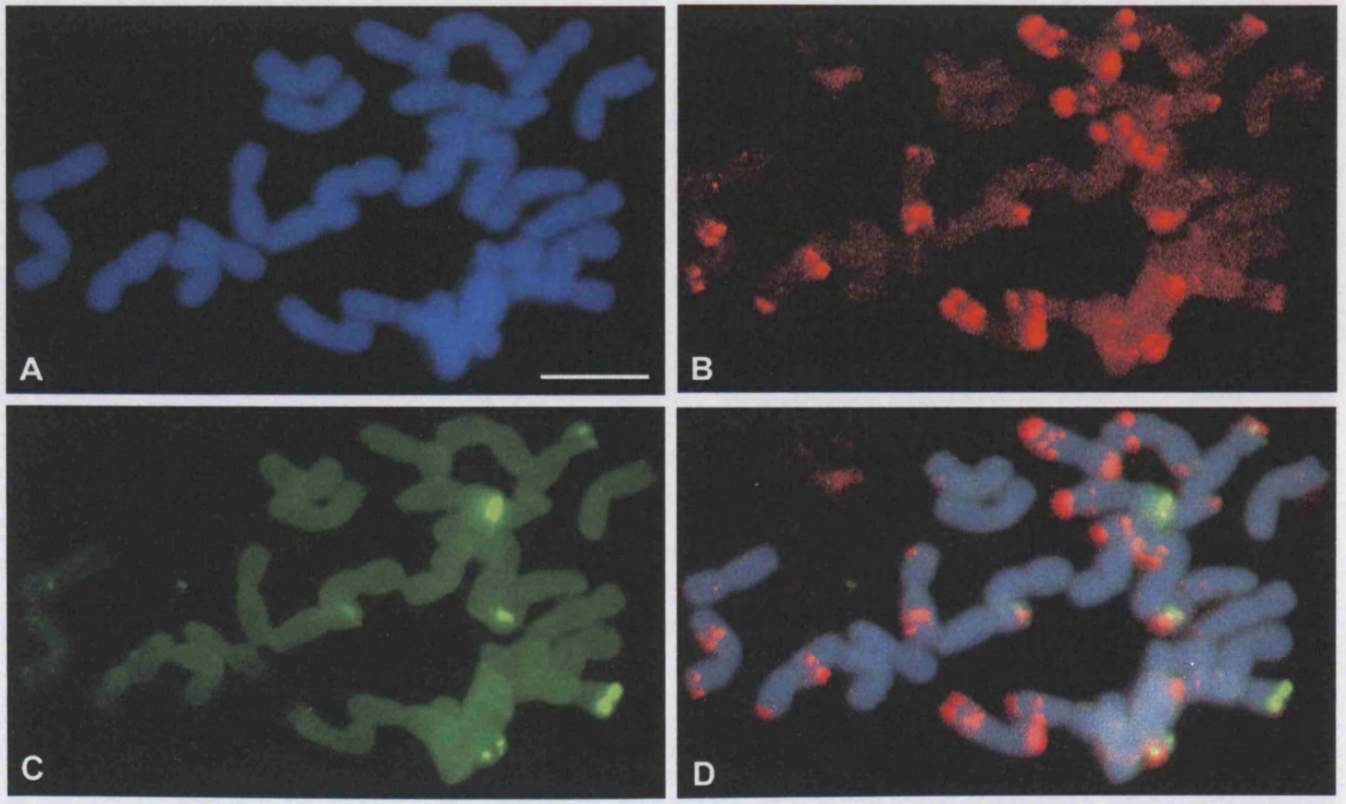


Figure 3. 11. Root tip metaphase of *T. turgidum* ssp. *durum* (AB genome, 2n=28). A) DAPI stained chromosomes. B) *In situ* hybridization with biotin-labelled Tmono25/208 !1212 probe detected with Alexa 594 red fluorescence. C) *In situ* hybridization with digoxigenin-labelled 5S rDNA pTa 794 probe detected with FITC green fluorescence. D) Overlay of A, B and C images. Scale bar represents 10 μm.

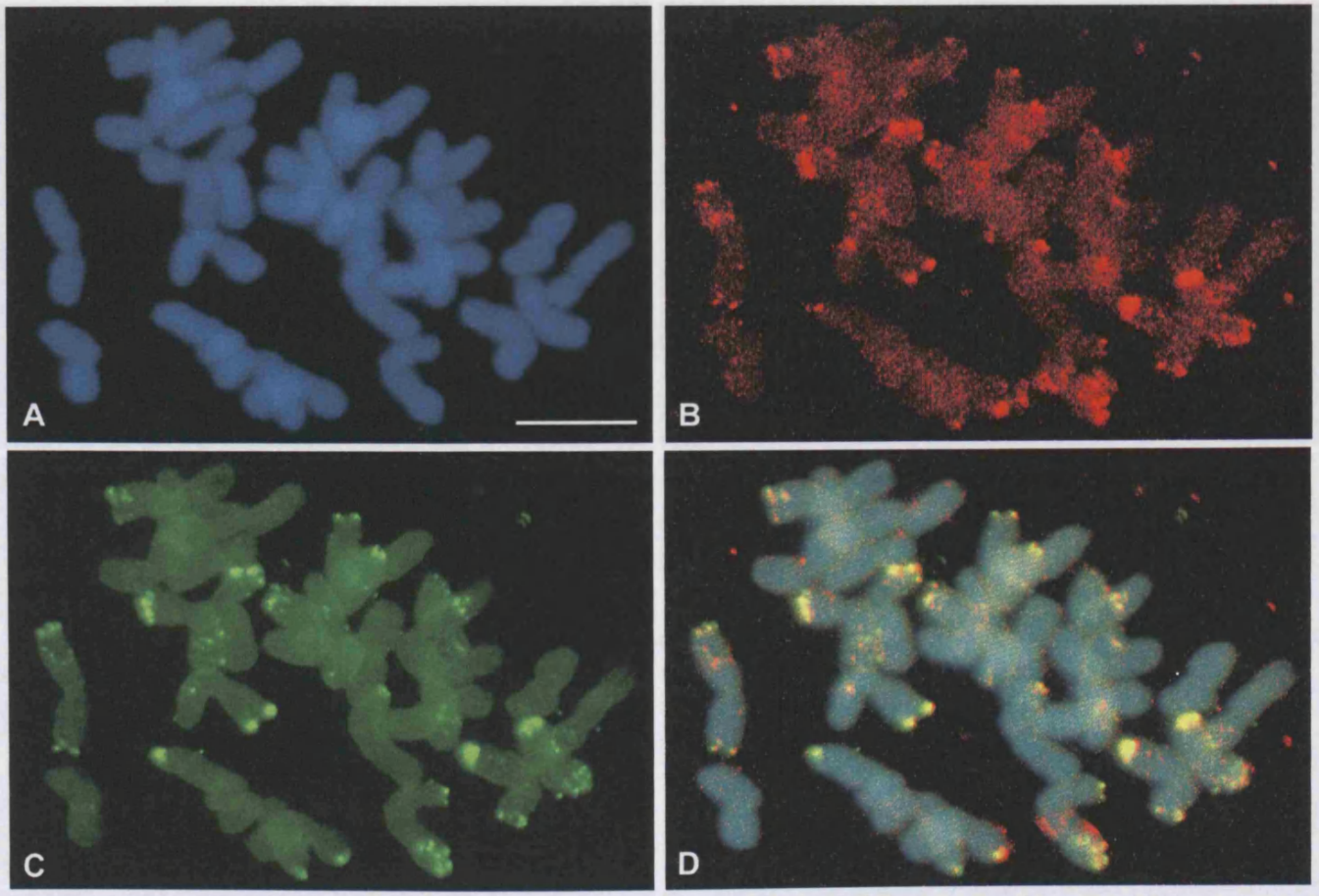


Figure 3. 12. Root tip metaphase of *T. turgidum* ssp. *durum* (AB genome, 2n=28). A) DAPI stained chromosomes. B) *In situ* hybridization with biotin-labelled Tmono25/208!1212 probe detected with Alexa 594 red fluorescence. C) *In situ* hybridization with digoxigenin-labelled Svav25/208!182 probe detected with FITC green fluorescence. D) Overlay of A, B and C images.Yellow fluorescence indicates co-localization of green and red signals. Scale bar represents 10 µm.

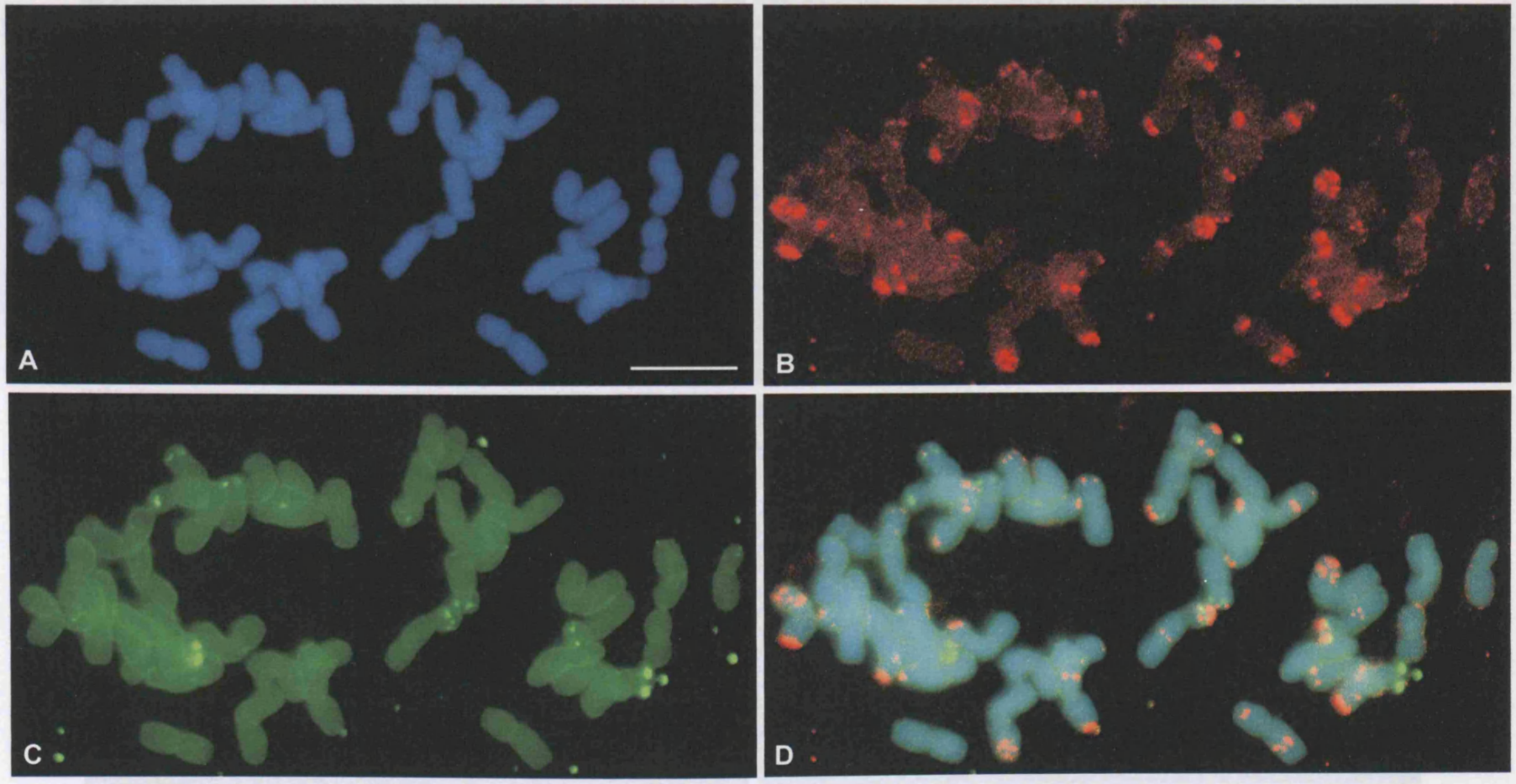


Figure 3. 13. Root tip metaphase of *T. aestivum* 'Chinese Spring' (ABD genome, 2n=42). A) DAPI stained chromosomes. B) *In situ* hybridization with biotin-labelled Ttau25/208!1416 probe detected with Alexa 594 red fluorescence. C) *In situ* hybridization with digoxigenin-labelled 5S rDNA pTa 794 probe detected with FITC green fluorescence. D) Overlay of A, B and C images. Scale bar represents 10 µm.

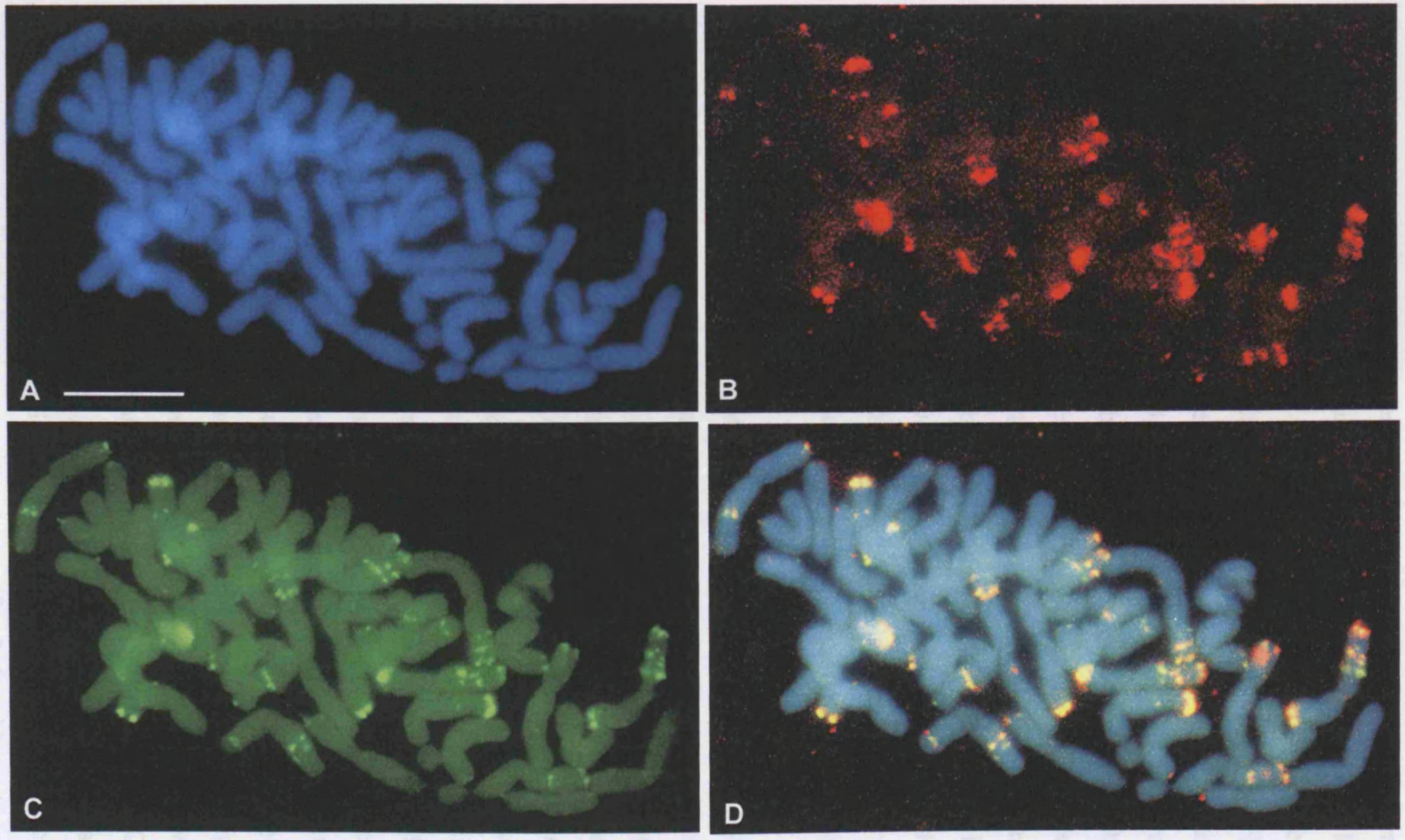


Figure 3. 14. Root tip metaphase of *T. aestivum* 'Chinese Spring' (ABD genome, 2n=42). A) DAPI stained chromosomes. B) *In situ* hybridization with biotin-labelled Svav25/208!182 probe detected with Alexa 594 red fluorescence. C) *In situ* hybridization with digoxigenin-labelled Ttau25/208!1416 probe detected with FITC green fluorescence. D) Overlay of A, B and C images. Yellow fluorescence indicates co-localization of green and red signals. Scale bar represents 10 μm.

Secale species and 70-100% in *Triticum/Aegilops*, table 3. 4). Interestingly, homology levels were similar also between the three groups of species corresponding to 72-95% between *Secale* and *Hordeum* species, 71-96% between *Secale* and *Triticum/Aegilops* species and 69-94% between *Hordeum* and *Triticum/Aegilops* species (table 3. 4). The overall results suggested similarity levels only slightly higher within than between species and no species-specific variants of 120-bp repeat unit family. Finally, from phylogenetic analysis no clear and well-supported groupings related to genus, species or primers used were observed (fig. 3. 5). 120-bp repeat unit family sequences were distributed throughout the trees with no relation to sequence copy number within a genome (fig. 3. 5). In fact, sequence diversity was observed for species such as *T. monococcum* and *T. tauschii* where 120-bp repetitive sequence family is not abundant (Mukai, Nakara et al. 1993; Katsiotis, Hagidimitriou et al. 1997) as well as for rye species, where it is a major repetitive DNA sequence family (McIntyre, Pereira et al. 1990; Cuadrado, Ceoloni et al. 1995; Vershinin, Schwarzacher et al. 1995). The variability of the 120-bp repeat units was also independent of ploidy level, cultivation or breeding system.

Fluorescent *in situ* hybridization (FISH) with 120-bp repetitive sequences used as probes showed that each chromosomal site, no matter if telomeric or interstitial, harboured many variants of the 120-bp repeat unit family. Presumably, the chromosomal sites made of 120-bp repeats were transferred as blocks and moved within the genome to give the variation in position and site number that was observed. Despite interstitial telomeric repeats not being observed in wheat and tomato (Schwarzacher and Heslop-Harrison 1991; Zhong, Fransz et al. 1998) 120-bp repeat unit family is present at interstitial sites of rye and wheat B genome, as previously reported for a TAS (Telomeric-associated sequence) isolated from tomato: the TGR1 sequence (Zhong, Fransz et al. 1998). Two different explanations have been proposed for explaining the interstitial location of TAS sequences in vertebrate chromosomes: they can be the remnants of telomeres after chromosome fusions or be generated through amplification of interstitial latent telomeres (Meyne, Baker et al. 1990). In agreement with previous data (Mukai, Nakara et al. 1993; Cuadrado and Jouve 1994; Cuadrado, Ceoloni et al. 1995; Vershinin, Schwarzacher et al. 1995) the *in situ* hybridization results showed that the 120-bp repeat unit family varied in both abundance and location of blocks of tandem repeats between genera in the *Triticeae* tribe. In rye (fig. 3. 6 and karyotypes in figs. 3. 7 and 3. 8) and wheat B genome (figs. 3. 11, 3. 12, 3. 13 and 3. 14) 120-bp repeat unit

family formed large subtelomeric heterochromatic blocks with minor intercalary bands, while few sites adjacent to telomeric repeats were present in the A and D wheat genomes (figs. 3. 9 and 3. 10).

Taken together, the results obtained suggested 120-bp repeat unit family to be an old component of *Triticeae* genomes, with no single 120-bp sequence being inherited by each genus or species from a single ancestor. Thus, the common ancestor of the *Triticeae* tribe had possibly multiple sequences of the 120-bp repeat unit family with a range of variation not unlike that seen within and between species today, and this diversity has been maintained in all species since their split from the common ancestor. More importantly, no homogenization events were apparently operative in 120-bp repeat unit family sequences within sites or within species since their evolutionary divergence, and therefore no characteristic genome or species-specific variants developed during evolution of the extant *Triticeae* genomes. In contrast, the 180-bp tandem repeat at the centromeres of *Arabidopsis* species showed similar levels of variation to that observed with the 120-bp repeat from rye, but variants were related to species (Heslop-Harrison, Brandes et al. 2003), a result which is most likely to be due to homogenization of the sequences after speciation. Furthermore, within the *Triticeae* tribe, the pAS1/dpTa1/AfaI/pHcKB6 repeat family showed variants characteristic of each genome where the sequences was found (Rayburn and Gill 1986; Anamthawat-Jónsson and Heslop-Harrison 1993; Vershinin, Svitashev et al. 1994; Nagaki, Tsujimoto et al. 1995; Nagaki, Tsujimoto et al. 1998) and similarly happened in the TaiI repeat family (Kishii, Nagaki et al. 1999; Kishii, Nagaki et al. 2001; Kishii and Tsujimoto 2002). These repetitive sequences showed evolutionary diversification in sequence and copy number, evidently under different constraints on their evolution compared to either chromosomal changes through translocations and duplications, or with respect to single-copy DNA sequences.

3. 5. 2 Molecular organization and sequence motifs of 120-bp repeat unit family

The *in situ* hybridization results and the PCR results, where all primer pairs gave fragments regardless of their position and direction in the sequence, demonstrated that the majority of the 120-bp repeat occurs in blocks of tandem arrays. However, PCR amplification of DNA fragments with single primers (F25, R42, R147) was an

indication of inverted junctions in some sequence arrays. Similarly, partial directed and partial inverted repeats in DNA fragments longer than three subrepeats were isolated using single primer R42 in *H. chilense* (fig. 3. 3), presumably due to the presence of some inverted repeats or degenerated sequences. In addition, smears (due to products of different sizes) were observed in PCRs with single primer and some primer pairs (fig. 3. 2), indicative of head-to-head junctions at multiple sites, non-tandem array organization, and/or presence of dispersed fragments of 120-bp repeats more widely along the genome. Finally, the *in situ* hybridization results showed weak hybridization along the euchromatin of chromosomes in *S. cereale* 'Petkus Spring' suggesting that sequences homologous to 120-bp repeat were dispersed throughout the genome of *S. cereale* 'Petkus Spring'.

Previous data about the molecular organization of pSc119.2 sequence analysed by restriction enzymes and Southern hybridization in rye (Vershinin, Schwarzacher et al. 1995) showed higher-order organization with a trimer of 360 bp defined by *Taq*I restriction site and a monomer of 118 bp defined by *Hae*III. Another non-homologous repetitive sequence family isolated from rye, pSc200, was found just adjacent to the telomeres, and occurred as a tandem array and with multiple reverse-orientation (head-to-head) junctions, always at the same point in the sequence, probably due breakage-fusion-bridge cycles (Vershinin, Schwarzacher et al. 1995). In rye head-to-head junctions were present in less extent in pSc250 repeat unit family that is further away from the telomeres (Vershinin, Schwarzacher et al. 1995) and possibly not very frequent in the 120-bp repeat unit family that is the most proximal repetitive sequence family located in the subtelomeric heterochromatin. A complex organization of satellite DNA sequences involving recombination and amplification steps was also shown in repetitive sequence families of other *Triticeae* species (Anamthawat-Jónsson and Heslop-Harrison 1993) and *Aveneae* species (Grebenstein, Grebenstein et al. 1996).

By focusing on the 120-bp repeat internal structure, a region of 18 bases was near-palindromic $A_2CGCAC_4G_4T_2CGT_2$ and two telomere-similar sequence motifs (CCCCAAAA or GGCCAAAA) were observed (fig. 3. 4). Imperfect inverted repeats and telomere-similar sequences have been previously found in telomeric-associated sequences (TASs) from tobacco, for example (Chen *et al.* 1997). However, no major length variants, duplications or other internal rearrangements were found during sequencing suggesting that sub-motifs were not conspicuous in the 120-bp repeat unit family although complex internal structures resulting from slippage replication and

other rearrangements are often common in many repetitive DNA sequences (Dechyeva *et al.* 2003).

Chapter 4: Changes in copy number and DNA methylation patterns of a major repetitive DNA sequence of Triticeae genomes

4. 1 Summary

In this chapter the copy number of 120-bp repeat unit family sequences was measured by using Real Time PCR in diploid and polyploid *Triticeae* species. T-test analysis showed significant differences in copy number of 120-bp repeat unit family sequences between allotetraploid wheat *T. turgidum* ssp. *durum*, allohexaploid *T. aestivum* 'Chinese Spring' and *Triticale* 'Fidelio' and their genome donors. Furthermore, changes in DNA methylation patterns of 120-bp repetitive sequence family were investigated in diploid and polyploid *Triticeae* genomes by restriction with enzymes sensitive to cytosine methylation and subsequent Southern hybridization, and discussed in the wider context of whole-genome DNA methylation and methylation patterns of another repetitive sequence family: the 5S ribosomal DNA sequences. Increase in DNA methylation at CpG sites and decrease in DNA methylation at CpNpG sites of 120-bp repetitive sequence family were observed in allopolyploid species *T. turgidum* ssp. *durum* and *Triticale* respectively. The evolutionary significance of modifications in copy number and methylation status of 120-bp repeat unit family sequences during events of allopolyploidization is discussed.

4. 2 Introduction

In higher plants repetitive DNA sequences generally correspond to 10^6-10^7 copies per unreplicated haploid genome as estimated in *Vicia faba*, *Olea* and *Pennisetum* species, for example (Kato, Yakura et al. 1984; Ingham, Hanna et al. 1993; Macas, Pozarkova et al. 2000; Contento, Ceccarelli et al. 2002). Changes in redundancy of repeated DNA sequences are considered an important factor in species evolution possibly due to different mechanisms such as unequal and illegitimate crossing-over (Charlesworth, Sniegowski et al. 1994; McAllister and Werren 1999) and breakage-

fusion-bridge cycles (McClintock 1941). Ugarkovic and Plohl (2002) reported that satellite DNA sequences vary between species according to changes in copy number due to unequal crossing-over and/or in nucleotide sequence due to different rates between mutation and homogeneization/fixation events. In addition, amplification and reduction of repetitive DNA sequences has been listed as one of the immediate consequences of a polyploidization event in the attempt to coordinate different genomes. Elimination of repeats during allopolyploidization has been observed in cultivated tetraploid and hexaploid wheats, but not in two endemic species of tetraploid wheats, *T. timopheevii* or *T. carthlicum*, where it remained unaltered (Petsova, Goncharov et al. 1998). Recently, sequence elimination mainly targeting rye genome and repetitive DNA sequences was reported in *Triticale* hybrids (Ma, Fang et al. 2004).

In this part of the work the copy number of 120-bp repeat unit family sequences was estimated by Real Time PCR in diploid and polyploid *Triticeae* species by using plasmid dilutions from a *S. cereale* 'Petkus' clone (corresponding from $5x10^5$ to 10^9 number of copies) as standards with the aim of evaluating if changes in redundancy in this major fraction of *Triticeae* genomes accompanied their evolution.

Cytosine methylation is an important feature of plant genomes affecting 20-30% of the total cytosines (Shapiro 1976), and among cereals 22% of the total cytosines in the wheat genome (Wagner and Capesius 1981). In contrast to animals where cytosine methylation mainly affects CpG sites, plant DNA methylation occurs not only at CpG but also at CpNpG sites (more frequently in CpApG and CpTpG), and asymmetrical CNN (where N=A, T or G) sites (Martienssen and Colot 2001). In this part of the work whole-genome DNA methylation in natural diploid and polyploid *Triticeae* species was analysed by using two combinations of restriction enzymes: *Msp*I-*Hpa*II and *Bst*NI-*Scr*FI for detecting methylation at CpG and CpNpG sites, as well as *Mcr*BC restriction enzyme that recognizes asymmetrical methylation, in order to have a complete representation of all three DNA methylation codes. In addition, since the 120-bp repeat unit family is GC-rich (GC content about 50-55%) methylation levels and methylation patterns at CpG, CpNpG and asymmetrical sites of this major DNA fraction of *Triticeae* genomes were investigated by Southern hybridization in natural diploid and polyploid *Triticeae* species. Finally, a comparison of methylation levels and methylation patterns at CpG, CpNpG and asymmetrical sites of 120-bp repetitive sequence family with whole-genome DNA methylation was performed, but also in respect to another repetitive sequence family: the 5S ribosomal DNA sequences. In fact,

previous data in literature showed 5S rDNA sequences to have high level of methylation, for example in tobacco with 44% of all cytosine residues methylated at CpG, CpNpG and asymmetrical sites (Fulnecek, Matyasek et al. 1998). Similarly, in *A. thaliana* and *S. cereale* methylation levels of 5S rDNA sequences measured by bysulfite sequencing (27% and 24% of all cytosine residues respectively) revealed that they are generally higher than whole-genome DNA methylation and affecting CpG, CpNpG and asymmetrical sites (Fulnecek, Matyasek et al. 2002).

Previous data reported changes in methylation at CpG sites in newly synthesized *Triticeae* allopolyploids as a rapid response of the genome to the allopolyploidization event and affecting its coding and non-coding fraction, mainly genes and mobile DNA elements (Shaked, Kashkush et al. 2001; Kashkush, Feldman et al. 2002; Han, Fedak et al. 2003). As tandemly organized DNA sequences have been related to nucleosome packaging (Heslop-Harrison 2000; Macas, Meszaros et al. 2002) it is essential to investigate their DNA methylation status in order to have informations about DNA methylation changes involved in chromatin organization processes that possibly happen after the polyplodization event. In this respect the 120-bp repeat unit family is an interesting model, being a widespread and probably ancient component of *Triticeae* genomes present in diploid and polyploid species. In conclusion, by comparing whole-genome DNA methylation levels and methylation patterns at CpG, CpNpG and asymmetrical sites with methylation levels and methylation patterns at CpG, CpNpG and asymmetrical sites of 120-bp repeat unit family sequences in natural diploid and polyploid *Triticeae* species, instead of the newly synthesized ones, it was possible to detect which DNA methylation changes were preserved by the evolution of *Triticeae* species.

4. 3 Materials and methods

4. 3. 1 Genomic DNA extraction

Total genomic DNA from *Triticeae* species: *S. cereale* 'Petkus Spring', *T. monococcum*, *Ae. speltoides*, *T. tauschii*, *T. turgidum* ssp. *durum*, *T. aestivum* 'Chinese Spring' belonging to Heslop-Harrison Molecular Cytogenetic Laboratory Seed Stock and *Triticale* 'Fidelio' and 'Lamberto' from Semundo-Danko Company, was extracted from fresh young leaves using the CTAB method (Gawel and Jarret, 1991). Genomic

DNA was quantified by gel and spectrophotometer as listed in table 4. 1. For the majority of samples the values of O.D.260/O.D.280 range from 1.8 to 2.0 indicating DNA purity.

4. 3. 2 Real Time PCR amplification

The amount of 120-bp repetitive sequence family in diploid and polyploid *Triticeae* species was estimated by developing a Two-Step Real Time PCR protocol. The protocol combines the annealing step with the extension step using the same temperature (57°C) for short time (20 seconds) so that during the cycles of amplification only very short fragments could be amplified. Two-Step Real Time PCR allows monomer amplification exclusively from 120-bp repeat unit family sequences, avoiding multiple products of amplification that could affect the measurements of sequence copy number.

Dilution series of genomic DNA (2.5 ng, 0.5 ng, 0.1 ng) were prepared for each species with the exception of *T. monococcum* and *T. tauschii* where higher DNA amount (62.5 ng, 12.5 ng, 2.5 ng) were necessary due to the low amount of 120-bp repetitive sequence family in their genomes. Two sets of replicates of genomic DNA dilution series from two individuals of each species were used. Plasmid dilutions from clone Pet-(22594)25/42-3324 were made corresponding to $5x10^5$, 10^6, 10^7, 10^8, 10^9 number of copies. Two sets of replicates of each dilution series were loaded and used as standars. A primer pair was synthesized by Sigma-Genosys:

F	5'	GCTGATGACCGACACGTAA	3' (Tm = 62.4°C)
R	5'	TGGACTATTACTCACTGTTTTGG	3' (Tm = 59.6 °C)

The locations of forward and reverse primers in 120-bp repetitive sequence family are indicated at the bottom of the aligned cloned sequences of *Secale* and *Triticum/Aegilops* species (fig. 4. 1a-b). Real Time PCR reactions were performed in DNA Engine Opticon System in 20 µl reaction containing: 10 µl of 2x master mix from DyNAmo™ HS SYB Green qPCR kit from Finnzymes (containing a modified DyNAmo hot start DNA polymerase, SYBR Green I, optimized PCR buffer, 5mM $MgCl_2$, dNTP mix including dUTP), 1µl of 10 µM for each primer and 8 µl of template DNA or water as control. Cycling conditions were: 94°C for 3 min; 35 cycles of 94°C 10 sec, 57°C annealing temperature for 20 sec and plate read step; the melting curve was measured

Table 4. 1. Genomic DNA concentrations (ng/μl) and O.D.260/O.D.280 ratios of individuals from diploid and polyploid *Triticeae* species.

Species	Individual	ng/μl	O.D.260/ O.D.280
S. cereale 'Petkus Spring'	Pt2	745.0	1.88
	Pt3	817.5	2.08
	Pt4	847.5	1.87
T. monococcum	Tm1	662.0	1.81
	Tm7	520.0	1.79
	Tm8	677.0	178
Ae. speltoides	Ae2	105.0	1.55
	Ae5	527.5	1.82
	Ae6	210.0	1.82
	AePool*	2783.0	1.42
T. tauschii	Tt9	850.0	1.88
	Tt11	700.0	1.86
	Tt12	1277.5	1.86
T. turgidum ssp.*durum*	Td1	1385.0	1.86
	Td2	1365.0	1.83
	Td4	737.0	2.54
T. aestivum 'Chinese Spring'	Cs3	1362.0	1.85
	Cs6	1410.0	1.84
	Cs7	1057.0	1.74
Triticale 'Fidelio'	Fi c	472.6	1.46
	Fi d	646.5	1.65
	Fi11	598.8	1.70
Triticale 'Lamberto'	La7	564.7	1.43
	La8	2989.2	1.53
	La11	2987.7	1.54

AePool* Total genomic DNA extracted from a pool of three individuals.

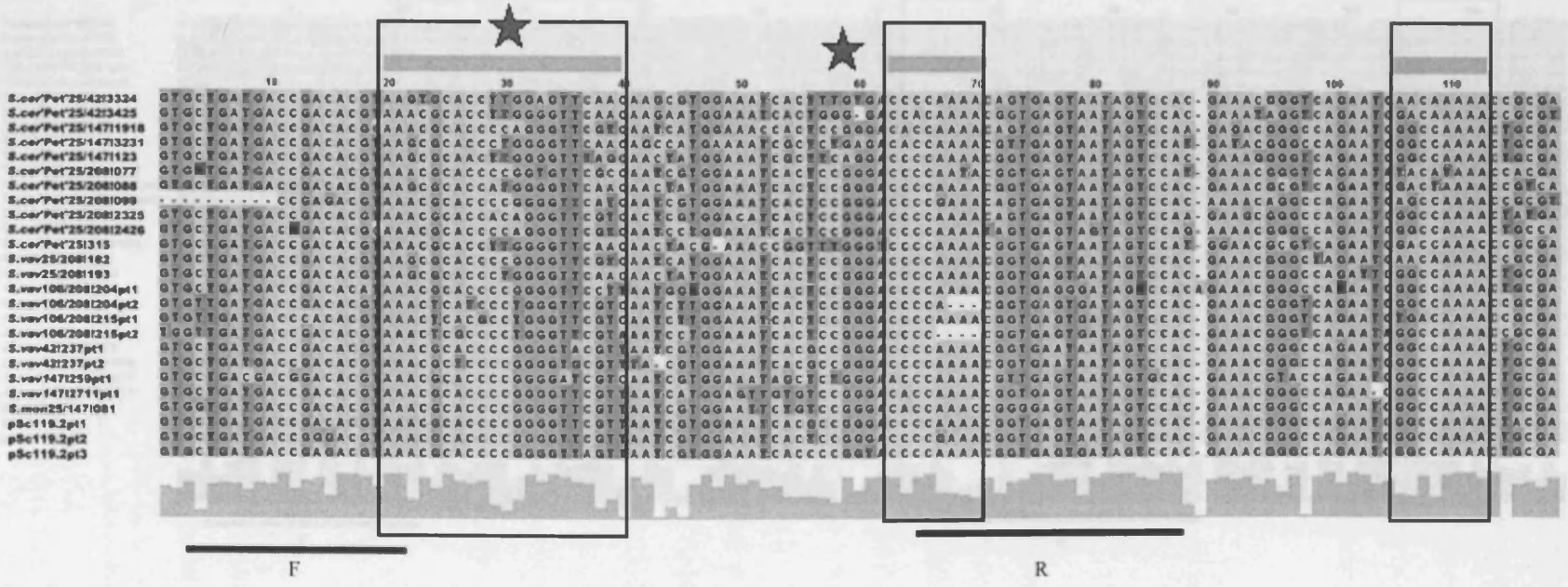


Figure 4. 1a. DNA sequence alignments of 120-bp repeat units of *Secale* species. The near-palindromic sequence (located from bp 20 to 40) and the two telomere-similar sequences (located from bp 63 to 70 and from bp 106 to 113 respectively) are boxed. Stars indicate symmetrical CpG and CpNpG sites.The locations of forward (F) and reverse (R) primers used in Real Time PCR experiments are indicated at the bottom of DNA sequence alignments.

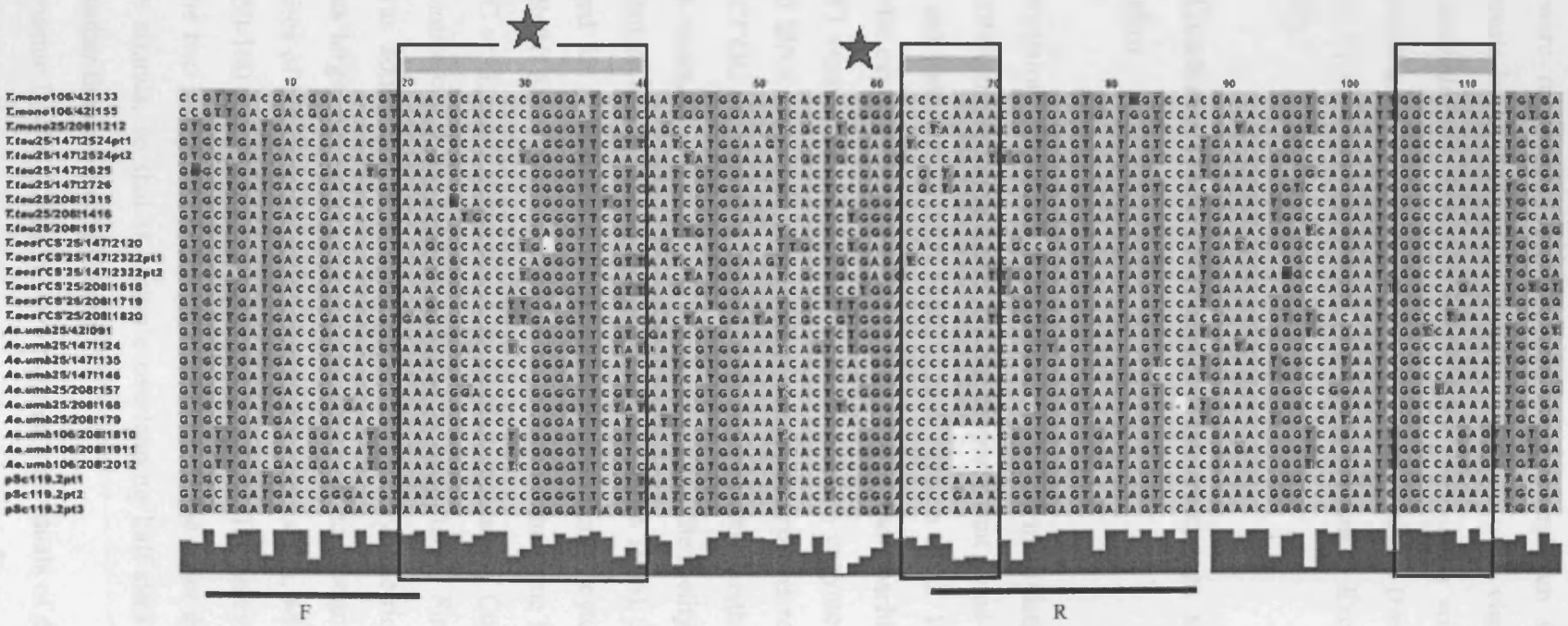


Figure 4. 1b. DNA sequence alignments of 120-bp repeat units of *Triticum/Aegilops* species. The near-palindromic sequence (located from bp 20 to 40) and the two telomere-similar sequences (located from bp 63 to 70 and from bp 106 to 113 respectively) are boxed. Stars indicate symmetrical CpG and CpNpG sites. The locations of forward (F) and reverse (R) primers used in Real Time PCR experiments are indicated at the bottom of DNA sequence alignments.

from 60°C to 95°C with increasing of 0.5°C and reading step of 1s. The Real Time PCR products were re-annealed at 70°C for 30 min and run into a 2% agarose gel electrophoresis in 1x TAE running buffer. Products were visualised by staining with ethidium bromide and their size checked by comparison with a size marker. Data analysis, melting curve, standard curve and T-test analysis (two-tailed distribution and two-sample equal variance) were done using Microsoft Excel Software (Microsoft Office 2000).

4. 3. 3 Genomic DNA restriction with enzymes sensitive to cytosine methylation

Restriction with enzymes sensitive to cytosine methylation is a classical method to assess DNA methylation based on the property that some restriction enzymes are unable to cut methylated DNA (McClelland, Nelson et al. 1994; Singal and Ginder 1999). In the experiments presented in this chapter the isoschizomers *Msp*I-*Hpa*II and *Bst*NI-*Scr*FI were used, as well as *Mcr*BC restriction enzyme that is able to cut only methylated DNA (fig. 4. 2). The *Msp*I-*Hpa*II isoschizomers recognize the same target sequence CCGG but *Hpa*II is sensitive to methylation at either cytosines while *Msp*I only at the external cytosine. Since it was shown that the methylation at ^{m}CCGG sites is not abundant in wheat species (Jeddeloh and Richards 1996), *Msp*I-*Hpa*II enzyme pair is then used for investigating methylation of the internal cytosine in CCGG sites. In addition, *Bst*NI-*Scr*FI restriction enzymes recognize the same target sequence CCNGG (N=A, T, C or G; more frequently A or T, Martienssen and Colot 2001) but the second enzyme cannot cut if the internal cytosine is methylated. Finally, *Mcr*BC restriction enzyme was added to the analysis for detecting asymmetrical methylation. *Mcr*BC enzyme has large, not stringently defined, asymmetric recognition sites: it recognizes two half-sites of the form (G/A)^{m}C separated by up to 3kb, with an optimal separation of about 50-100 bp (Kubis, Castilho et al. 2003). The enzyme is able to cut once between the two half-sites close to one half-site or the other and can cleave on one or both DNA strands, so that if multiple overlapping half-sites are present a smear is produced rather than a banding pattern.

Genomic DNAs (7.5 μg) of two to three individuals of diploid species *S. cereale* 'Petkus Spring', *T. monococcum*, *Ae. speltoides* and *T. tauschii*, and three to five individuals of polyploid species *T. turgidum* ssp. *durum*, *T. aestivum* 'Chinese Spring'

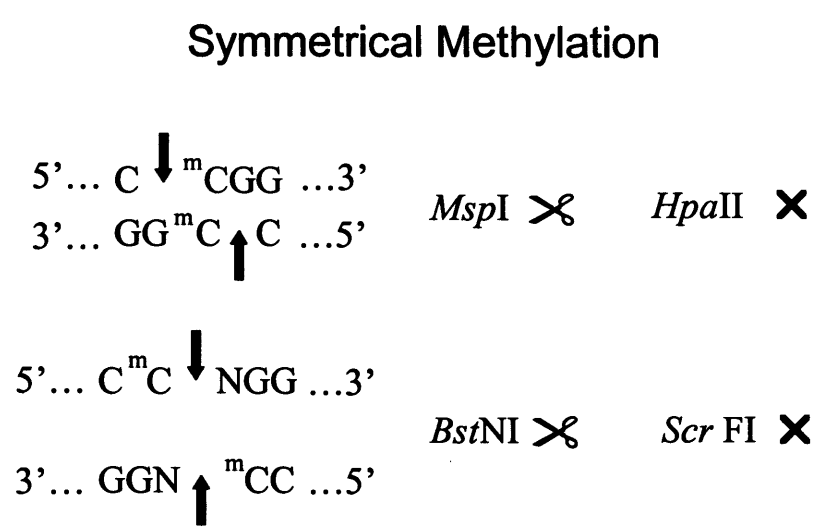


Figure 4. 2. Restriction enzymes used for detecting DNA methylation at symmetrical and asymmetrical sites, and their target sequences. Isoschizomers *Msp*I-*Hpa*II recognize the same target sequence CCGG, but *Hpa*II is sensitive to methylation at the internal cytosine. Isoschizomers *Bst*NI-*Scr*FI recognize the same target sequence CCNGG, but *Scr*FI is sensitive to methylation at the internal cytosine. *Mcr*BC recognizes two half-sites of the form (G/A)^{m}C separated by up to 3kb and cuts once between the two half-sites close to one half-site or the other.

and *Triticale* 'Fidelio' and 'Lamberto' were restricted. For each digestion 4U/μg of restriction enzyme have been used overnight at 37°C or the required temperature as described in Materials and Methods chapter. Enzyme digestions were performed twice for individuals of polyploid species *T. turgidum* ssp. *durum*, *T. aestivum* 'Chinese Spring' and *Triticale* 'Fidelio' and 'Lamberto', showing no major differences.

4. 3. 4 Agarose gel analysis and Southern hybridization

Due to the short length of the subunit of 120-bp repeat unit family the digested DNA samples were run in 2% agarose gel electrophoresis in 1x TAE running buffer in order to achieve a good separation of short DNA fragments, such as trimers (360 bp), dimers (240 bp) and monomers (120 bp). However, monomers (120 bp) were usually not visible, possibly because the digestions were not complete to the single units or because enzyme target sites are missing in some subunits of 120-bp repeat unit family but also as a result of technical reasons. 750 ng of undigested DNA samples were run together with the digested DNA samples and used as controls, in order to show that DNA was not degraded and to detect false cutting sites. DNA digests were visualised by staining the gel with ethidium bromide and the size of the fragments was checked by comparison with a size marker.

Southern hybridizations were carried out as described in Materials and Methods chapter. DNA blotting was performed overnight on positive charged nylon membranes. Non-radioactive hybridizations were carried out using 750 ng of digoxigenin-labelled clones belonging to 120-bp repeat unit family (see asterisks in fig. 3. 5b and table 3. 5) as probes to DNA transfers in high stringency conditions. Washes were performed in high stringency conditions. The signal was developed by chemiluminescence using alkaline phosphatase detection. Membranes were stripped and re-hybridised with biotin-labelled 5S rDNA pTa794 clone (Gerlach and Dyer 1980). Southern hybridizations with clones belonging to 120-bp repeat unit family and re-hybridizations with 5S rDNA pTa794 clone were performed twice for individuals of polyploid species *Triticum turgidum* ssp. *durum*, *T. aestivum* 'Chinese Spring' and *Triticale* 'Fidelio' and 'Lamberto', showing no major differences.

4. 4 Results

4. 4. 1 Copy number of 120-bp repetitive sequences in diploid and polyploid *Triticeae* species

Monomeric units of 120-bp repetitive sequences were amplified by using Two-Step Real Time PCR in diploid and polyploid *Triticeae* species. Melting curves were obtained by plotting the increasing temperature against the negative first derivative of the decrease of fluorescence and represented in fig. 4. 3a-b-c. The shape of the melting curves showing one major peak indicated that monomers only were exclusively amplified, and no primer–dimers were affecting the results since no additional peaks were located before the major one. Melting curves of individuals from *Ae. speltoides* (B genome) and *S. cereale* 'Petkus Spring' (R genome) shown in fig. 4. 3a were analysed separately from *T. tauschii* (D genome) individuals shown in fig. 4. 3b due to their marked differences in copy number, while individuals from polyploid species *T. turgidum* ssp. *durum* (AB genome), *T. aestivum* 'Chinese Spring' (ABD genome) and *Triticale* 'Fidelio' (ABR genome) were represented together in fig. 4. 3c. In all samples the length of products amplified by Two-Step Real Time PCR is about 90 bp as expected from primer pair location in the 120-bp repeats (previously shown in fig. 4. 1a-b). Gel analysis confirmed that monomers were exclusively amplified in diploid and polyploid *Triticeae* species (see fig. 4. 4a-b).

An absolute quantification of 120-bp repetitive sequences in diploid and polyploid *Triticeae* species was estimated by comparing samples with plasmid dilution series of clone Pet-(22594)25/42-3324. As shown in fig. 4. 5 a standard curve was obtained by plotting the threshold cycle (Ct) against the logarithm of the amount of DNA of plasmid dilution series (corresponding to $5x10^5$, 10^6, 10^7, 10^8, 10^9 number of copies). The standard curve was then used to calculate the copy number per nanogram of 120-bp repetitive sequences of two sets of replicates of genomic DNA dilution series in two individuals of diploid and polyploid *Triticeae* species.

As reported in table 4. 2, 120-bp repeat unit family sequences in diploid *S. cereale* 'Petkus Spring' (R genome) and *Ae. speltoides* (B genome) were in the range of 10^8 copies per nanogram of genomic DNA and no significant differences were found between the two individuals of each species having similar values ($3.36x10^8$ and $3.63x10^8$ in *S. cereale* 'Petkus Spring'; $2.52x10^8$ and $2.69x10^8$ in *Ae. speltoides*). In *T.*

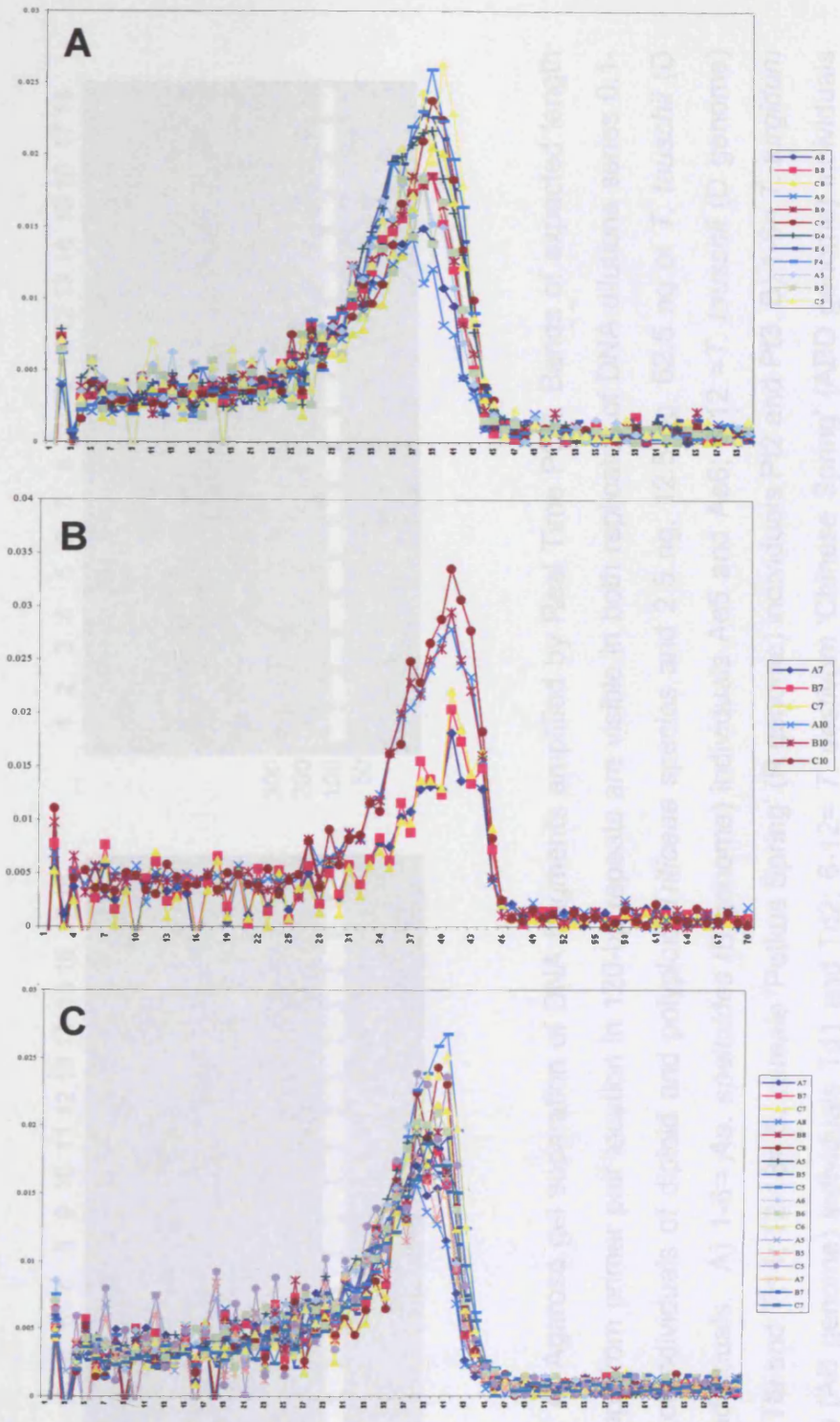


Figure 4. 3. A. Melting curves of *S. cereale* 'Petkus Spring' and *Ae.speltoides* (2.5, 0.5, 0.1 ng genomic DNA: replicates A8-C9 and D4-C5, respectively) B. Melting curves of *T. tauschii* (62.5, 12.5, 2.5 ng genomic DNA: replicates A7-C10) C. Melting curves of *T. turgidum* ssp. *durum*, *T. aestivum* 'Chinese Spring' and *Triticale* 'Fidelio' (2.5, 0.5, 0.1 ng genomic DNA: replicates A7-C8, A5-C6 and A5-C7, respectively).

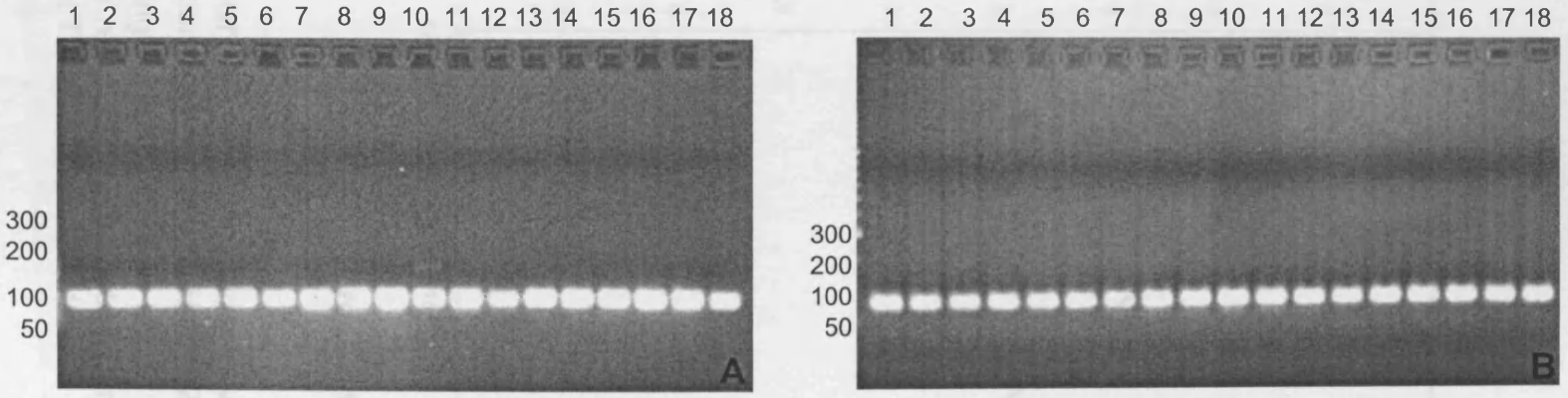


Figure 4. 4. Agarose gel separation of DNA fragments amplified by Real Time PCR. Bands of expected length (about 90 bp) from primer pair location in 120-bp repeats are visible in both replicates of DNA dilutions series 0.1-0.5-2.5 ng of individuals of diploid and polyploid *Triticeae* species and 2.5 ng, 12.5 ng, 62.5 ng of *T. tauschii* (D genome) individuals. A) 1-6= *Ae. speltoides* (B genome) individuals Ae5 and Ae6; 6-12 =*T. tauschii* (D genome) individuals Tt9 and Tt11; 12-18= *S. cereale* 'Petkus Spring' (R genome) individuals Pt2 and Pt3. B) 1-6= *T. turgidum* ssp. *durum* (AB genome) individuals Td1 and Td2; 6-12= *T. aestivum* 'Chinese Spring' (ABD genome) individuals Cs3 and Cs6; 12-18= *Triticale* 'Fidelio' (ABR genome) individuals Fi d and Fi11.

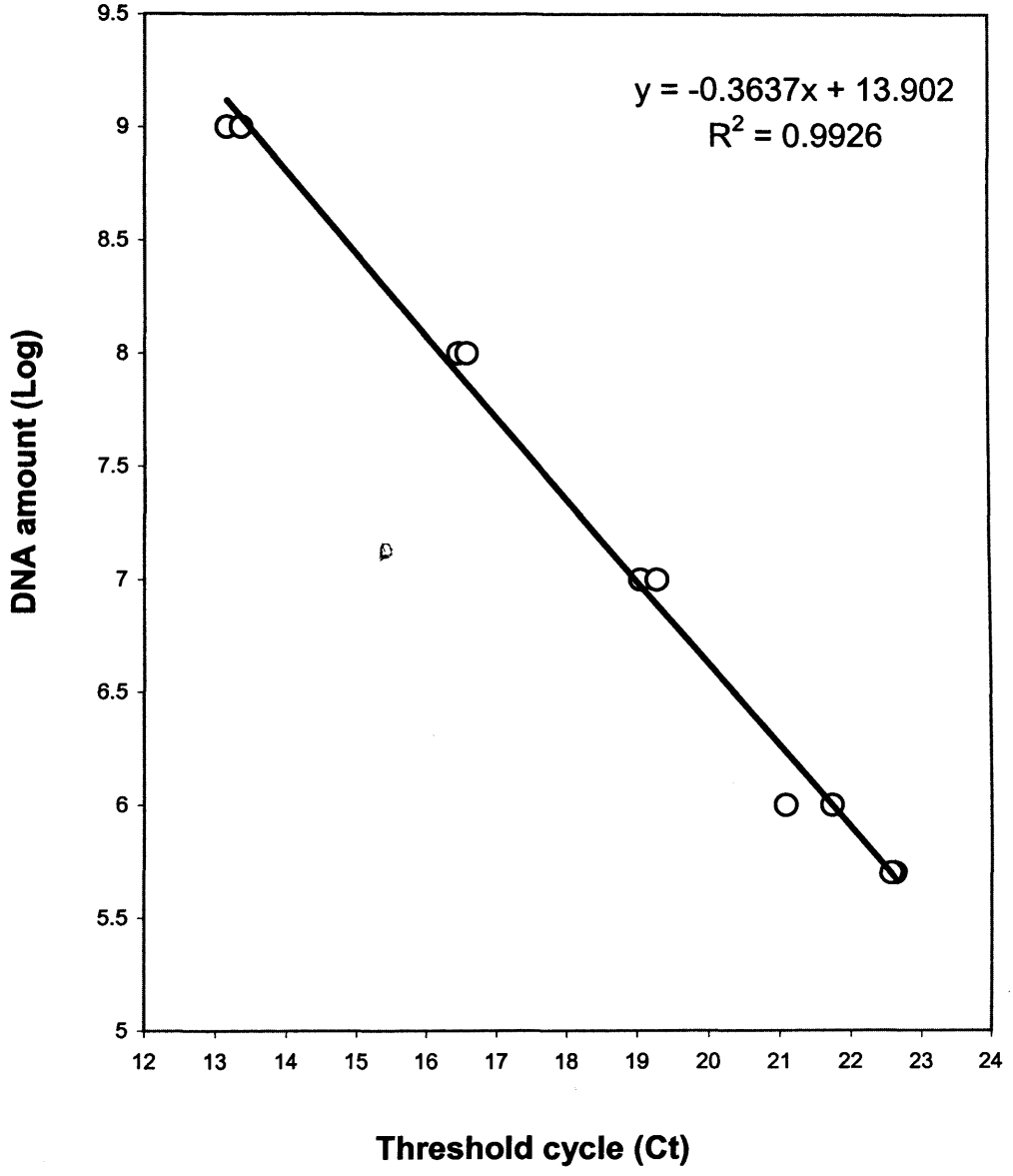


Figure 4. 5. Correlation between DNA amounts of plasmid dilutions of Pet(22594)25/42!3324 clone from *S. cereale* 'Petkus Spring' corresponding to $5x10^{5,}$ 10^6, 10^7, 10^8, 10^9 copy number (two replicates each dilution) and threshold cycles measured in Real Time PCR amplification.

Table 4. 2. Real time PCR estimates of copy number per nanogram and copy number per unreplicated haploid genome in individuals of diploid and polyploid *Triticeae* species.

Species	Genome size (pg/1C value)	Individual	Replicates (copy number /nanogram)		Average copy number /nanogram	Standard deviation (units)	Copy number /haploid genome
			I set of replicates	II set of replicates			
Secale cereale 'Petkus Spring' (R genome)	10.5 a)	Pt2	3.68×10^{8} 2.29×10^{8} 4.53×10^{8}	4.80×10^{8} 4.01×10^{8} 2.48×10^{8}	3.63×10^{8}	1.04×10^{8}	3.81×10^{6}
		Pt3	3.11×10^{8} 3.34×10^{8} 2.20×10^{8}	3.18×10^{8} 3.74×10^{8} 4.59×10^{8}	3.36×10^{8}	0.79×10^{8}	3.53×10^{6}
Aegilops speltoides (B genome)	5.1 a) 5 a)	Ae5	2.46×10^{8} 2.32×10^{8} 4.35×10^{8}	4.99×10^{8} 1.14×10^{8} 0.88×10^{8}	2.69×10^{8}	1.67×10^{8}	1.36×10^{6}
		Ae6	4.14×10^{8} 1.72×10^{8} 4.73×10^{8}	2.29×10^{8} 1.37×10^{8} 0.89×10^{8}	2.52×10^{8}	1.56×10^{8}	1.27×10^{6}
Triticum tauschii (D genome)	4.2 a) 5.08 b)	Tt9	1.90×10^{6} 1.14×10^{6} 0.47×10^{6}	0.44×10^{6} 1.84×10^{6} 0.75×10^{6}	1.09×10^{6}	0.66×10^{6}	0.51×10^{4}
		Tt11	3.09×10^{6} 6.96×10^{6} 6.08×10^{6}	7.36×10^{6} - -	5.87×10^{6}	1.93×10^{6}	2.72×10^{4}

Table 4. 2. continued

Species	Genome size (pg/1C value)	Individual	Replicates (copy number /nanogram) I set of replicates	Replicates (copy number /nanogram) II set of replicates	Average copy number /nanogram	Standard deviation (units)	Copy number /haploid genome
Triticum turgidum ssp. *durum* (AB genome)	12.9 a) 11.96 b)	Td1	2.27×10^7 2.15×10^7 1.41×10^7	2.91×10^7 3.26×10^7 2.34×10^7	2.39×10^7	0.06×10^7	2.97×10^5
		Td2	1.67×10^7 1.88×10^7 1.38×10^7	2.84×10^7 1.25×10^7 1.82×10^7	1.81×10^7	0.06×10^7	2.25×10^5
Triticum aestivum 'Chinese Spring' (ABD genome)	18.1 a) 17.8 a)	Cs3	8.23×10^7 4.33×10^7 3.33×10^7	10.4×10^7 10.3×10^7 6.61×10^7	7.20×10^7	2.98×10^7	1.29×10^6
		Cs6	4.12×10^7 3.17×10^7 2.2×10^7	7.42×10^7 6.24×10^7 3.78×10^7	4.49×10^7	1.96×10^7	0.81×10^6
Triticale 'Fidelio' (ABR genome)	19.8 a)	Fi d	- 3.59×10^7 6.22×10^7	- 7.67×10^7 8.73×10^7	6.55×10^7	2.23×10^7	1.30×10^6
		Fi11	1.05×10^7 3.3×10^7 3.58×10^7	1.65×10^7 6.41×10^7 6.04×10^7	3.67×10^7	2.2×10^7	0.73×10^6

a) Genome size (pg per 1C value) from Bennett and Smith (1991). If two estimates are present they refer to different methods and/or publications.

b) Genome size (pg per 1C value) from Ozkan *et al*. (2001).

tauschii (D genome) individuals 120-bp repeat unit family sequences had lower sequence copy number ($1.09x10^6$ and $5.87x10^6$) and were significantly different at P=0.0004. By using the genome size values (or the average when two values were present), previously reported for *Triticeae* species (Bennett and Smith 1991; Ozkan, Levy et al. 2001) and listed in table 4. 2, the copy number per unreplicated haploid genome of 120-bp repeat unit family sequences was calculated. The copy number per unreplicated haploid genome was in the range of 10^6 in *S. cereale* 'Petkus Spring' (R genome) and *Ae. speltoides* (B genome) species ($3.53x10^6$- $3.81x10^6$ and $1.27x10^6$-$1.36x10^6$, respectively), with *S. cereale* 'Petkus Spring' having the highest values, while in *T. tauschii* (D genome) was in the range of 10^4 ($0.51x10^4$- $2.72x10^4$). Conversely, the copy number of 120-bp repetitive sequences in A genome was estimated to be very low, in the range of 10^4 per nanogram of genomic DNA and 10^2 per unreplicated haploid genome.

As reported in table 4. 2, in tetraploid and hexaploid *Triticeae* species 120-bp repeat unit family sequences showed values in the range of 10^7 per nanogram of genomic DNA. No significant differences in the copy number of 120-bp repetitive sequences were found between individuals of tetraploid *T. turgidum* ssp. *durum* (AB genome) with $1.81x10^7$ and $2.39x10^7$ number of copies, nor in *T. aestivum* 'Chinese Spring' (ABD genome) individuals ($4.49 x10^7$ and $7.20 x10^7$), nor in *Triticale* 'Fidelio' (ABR genome) individuals ($3.67x10^7$ and $6.55x10^7$). The copy number per unreplicated haploid genome was in the range of 10^5 in *T. turgidum* ssp. *durum* (AB genome) individuals ($2.25x10^5$-$2.97x10^5$), while in *T. aestivum* 'Chinese Spring' (ABD genome) and *Triticale* 'Fidelio' (ABR genome) individuals was in the range of 10^6 ($0.81x10^6$-$1.29x10^6$ and $0.73x10^6$- $1.30x10^6$, respectively).

4. 4. 2 Whole-genome DNA methylation and methylation patterns of 120-bp repeat unit family

a) Whole-genome DNA methylation and methylation patterns of 120-bp repeat unit family in diploid and polyploid wheat species

Within the 120-bp repeat unit of *Secale* and *Triticum/Aegilops* a CCGG site (star in fig. 4. 1a-b) is present in the near-palindromic sequence (located from bp 20-21 to 40-41 and boxed in fig. 4. 1a-b). Remarkably, the CCGG site becomes in some clones a CCNGG site by the insertion of a single nucleotide. A second CCGG site (star in fig. 4.

1a-b) is present before the first telomere-similar sequence (located from bp 63 to 70 and boxed in fig. 4. 1a-b) in most *Secale* clones, but less frequently in *Triticum/Aegilops* clones and very rarely becomes a CCNGG site. The size of fragments expected in presence and absence of methylated cytosines at these symmetrical CpG and CpNpG sites is shown in fig. 4. 6. By considering two adjacent monomers (a dimer) if both restriction sites are digested within each monomer fragments of 30 and 90 bp are generated, while fragments of 120 bp if only one of the two corresponding sites is digested, and longer fragments of 150 bp if only one of the two sites is digested but not the corresponding ones. By considering three adjacent monomers (a trimer) fragments 210, 240 and 270 bp long are generated by different combinations of cutting sites within each monomer (note that not all combinations are possible), by considering four adjacent monomers fragments 330, 360 and 390 bp long are produced, and so on.

Genomic DNA restrictions of *T. monococcum* (A genome), *Ae. speltoides* (B genome), *T. tauschii* (D genome) and *T. turgidum* ssp. *durum* (AB genome) individuals were performed with the isoschizomers *Msp*I-*Hpa*II and *Bst*NI-*Scr*FI as well as *Mcr*BC restriction enzyme. Fragments were separated by 2% agarose gel and visualised through ethidium bromide staining (gels pictures in fig. 4. 7a and 4. 8a). With *Msp*I and *Bst*NI restriction enzymes that cut independently from DNA methylation, smears in all species investigated were detected, only to some extent with *Msp*I and marked with *Bst*NI. In contrast, different levels of DNA methylation were observed by using their isoschizomers, sensitive to DNA methylation, *Hpa*II and *Scr*FI respectively. Overall, A (*T. monococcum*) and D (*T. tauschii*) genomes seemed to be symmetrically methylated as *Hpa*II and *Scr*FI showed minimal digestion in their CCGG and CCNGG recognition sequences (see lanes 3 and 5, 15 and 17 in fig. 4. 7a), whereas in B (*A. speltoides*) and AB genomes (*T. turgidum* ssp. *durum*) *Hpa*II cut to a little extent (see lanes 9 in fig. 4. 7a and lanes 3, 9 and 15 in fig. 4. 8a) and *Scr*FI to some extent (see lane 11 in fig. 4. 7a and lanes 5, 11 and 17 in fig. 4. 8a). In addition, in B (*A. speltoides*) and in AB genomes (*T. turgidum* ssp. *durum*) CpG appeared more methylated than CpNpG sites (compare lanes 9 with 11 in fig. 4. 7a and lanes 3, 9 and 15 with lanes 5, 11 and 17 in fig. 4. 8a). Finally, A (*T. monococcum*) and D (*T. tauschii*) genomes appeared asymmetrically methylated (*Mcr*BC showed minimal digestion, see lanes 6 and 18 in fig. 4. 7a) while B (*A. speltoides*) and AB genomes (*T. turgidum* ssp. *durum*) genomes were digested by *Mcr*BC to some extent (see lane 12 in fig. 4. 7a and lanes 6, 12 and 18 in fig. 4. 8a).

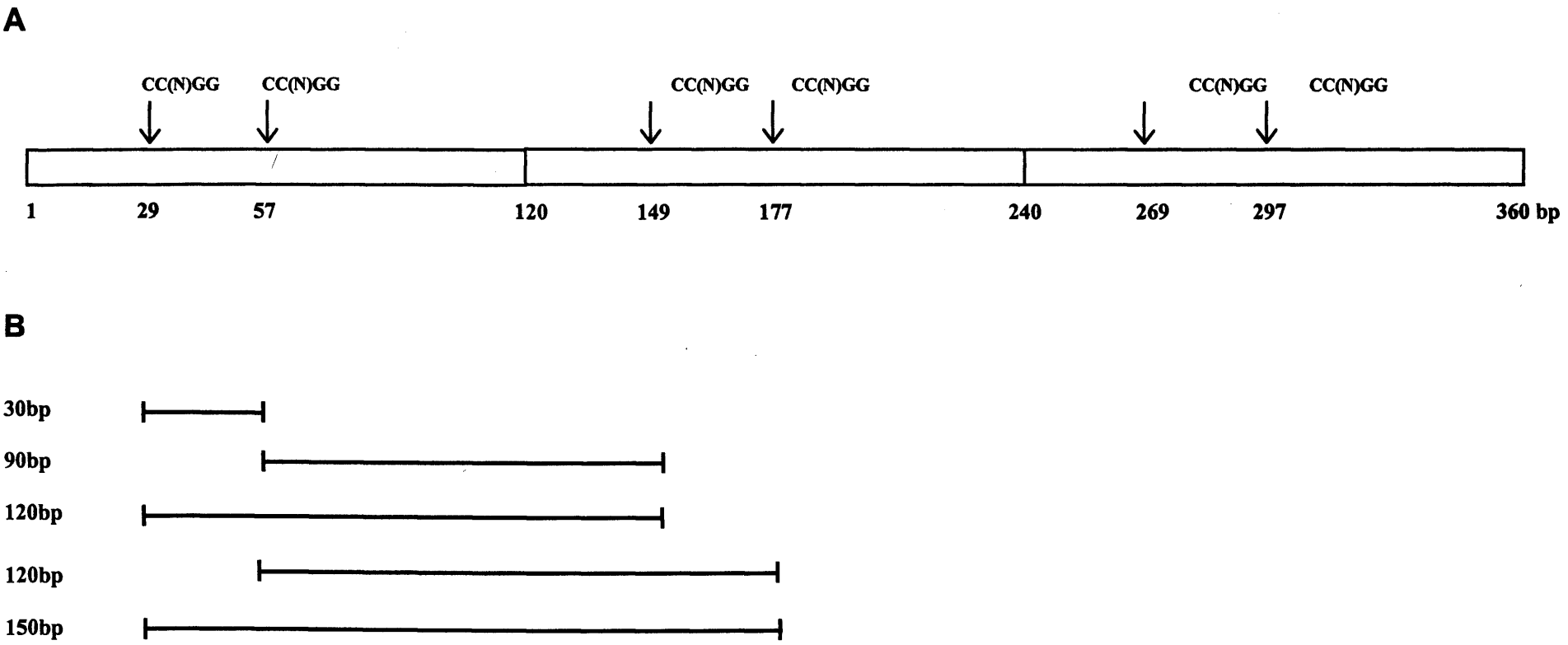


Figure 4. 6. The CCGG and CCNGG sites in 120-bp repeat unit family sequences. (A) Diagram of 120-bp repeat with target CCGG and CCNGG sites for isoschizomers *Msp*I-*Hpa*II and *Bst*NI-*Scr*FI. (B) Size of DNA fragments expected in presence and absence of methylated cytosines at symmetrical CpG and CpNpG sites corresponding to monomers of 120-bp repeat unit family. Fragments 210, 240 and 270 bp long are corresponding to dimers and fragments 330, 360 and 390 bp to trimers.

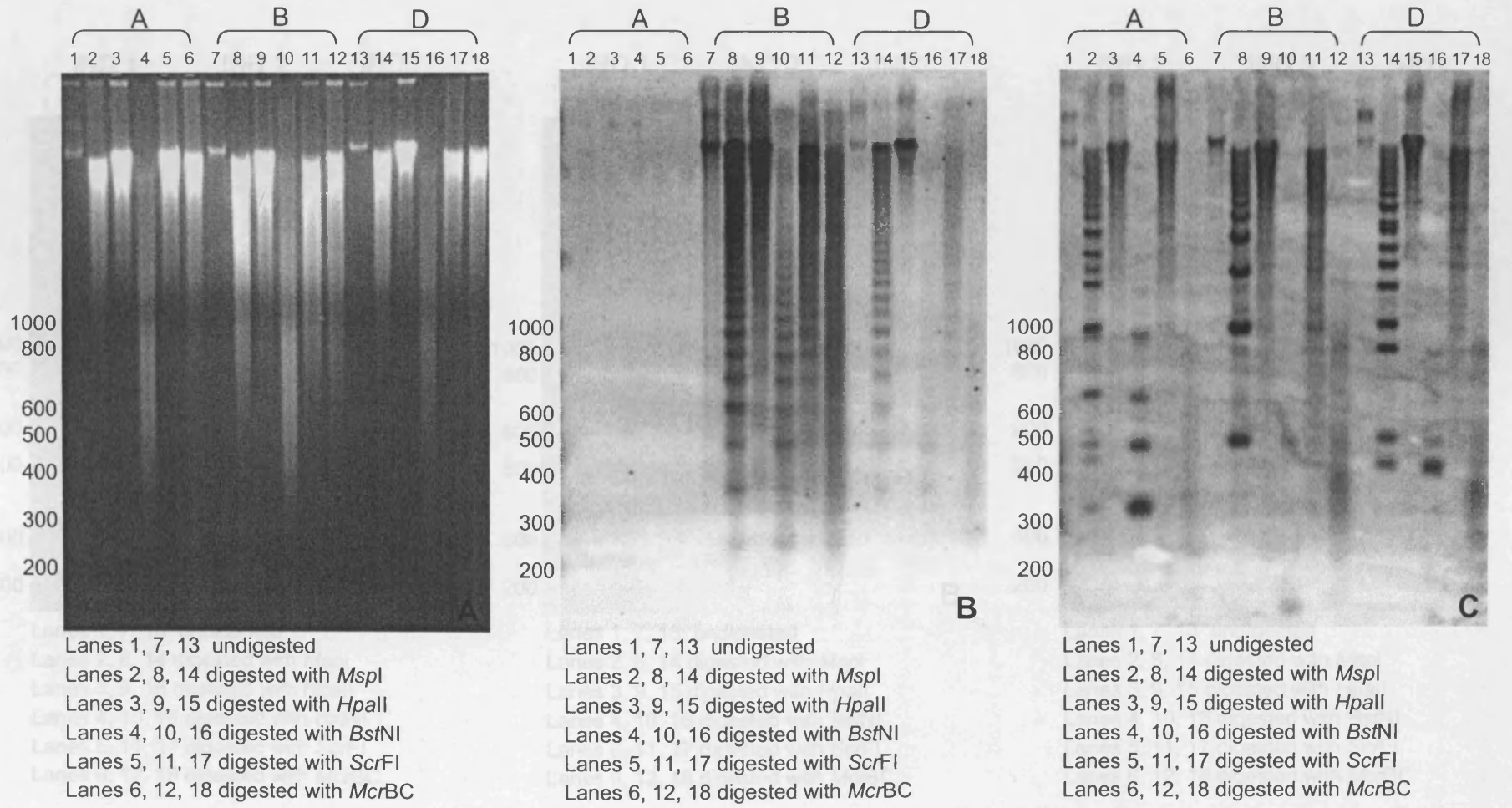


Figure 4. 7. A) Gel picture of genomic DNA of *T. monococcum* (A genome), *Ae. speltoides* (B genome), *T. tauschii* (D genome) individuals digested with isoschizomers *Msp*I-*Hpa*II and *Bst*NI-*Scr*FI as well as *Mcr*BC restriction enzyme. Southern hybridizations with digoxigenin-labelled Tmono25/208!1416 clone (B) and with biotin-labelled 5S rDNA pTa794 clone (C).

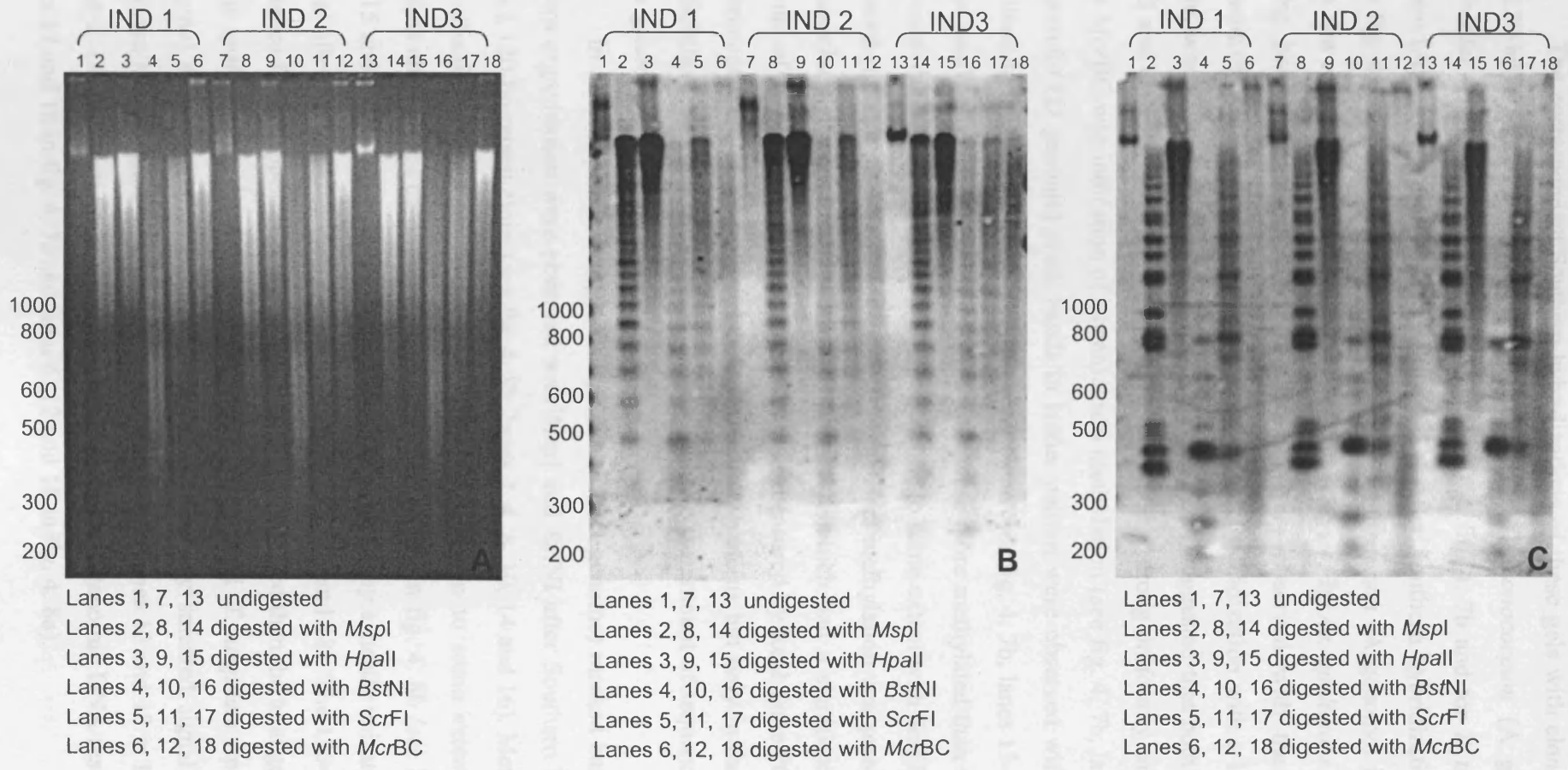


Figure 4. 8. A) Gel picture of genomic DNA of *T. turgidum* ssp. *durum* (AB genome) individuals digested with isoschizomers *Msp*I-*Hpa*II and *Bst*NI-*Scr*FI as well as *Mcr*BC restriction enzyme. Southern hybridizations with digoxigenin-labelled Pet(22594)25/42!3324 clone (B) and with biotin-labelled 5S rDNA pTa794 clone (C).

The correspondent Southern hybridizations of these gels with clones belonging to 120-bp repeat unit family are illustrated for *T. monococcum* (A genome), *Ae. speltoides* (B genome), *T. tauschii* (D genome) in fig. 4. 7b and for *T. turgidum* ssp. *durum* (AB genome) individuals in fig. 4. 8b. After Southern hybridization no signal (see fig. 4. 7b, lanes 1-6) was detected in *T. monococcum* (A genome) in agreement with the results obtained by Real Time PCR analysis. In *Ae. speltoides* (B genome) strong bands in ladder pattern organization were observed with the methylation insensitive *Msp*I and *Bst*NI enzymes after Southern hybridization with a 120-bp repeat clone, whereas marked smears with *Hpa*II and *Scr*FI suggested that both symmetrical CpG and CpNpG sites were poorly methylated and a strong uniformly smeared signal with *Mcr*BC was indication of asymmetrical methylation (see fig. 4. 7b, lanes 7-12). In *T. tauschii* (D genome) weak bands in ladder pattern were observed with *Msp*I after Southern hybridization with a 120-bp repeat clone (see fig. 4. 7b, lanes 13-18) and very weak bands after *Bst*NI digestion. CpG sites were more methylated than CpNpG sites (minimal digestion is observed with *Hpa*II while to some extent with *Scr*FI) and a weak smeared signal with *Mcr*BC indicated asymmetrical methylation was present. Between *T. tauschii* (D genome) individual Tt9 and Tt11, which had a significantly different amount of 120-bp repetitive sequence family as measured by Real Time PCR, no major differences were found in their DNA methylation patterns but only in the intensity of the Southern hybridization signals corresponding to the amount of sequences present in their genomes.

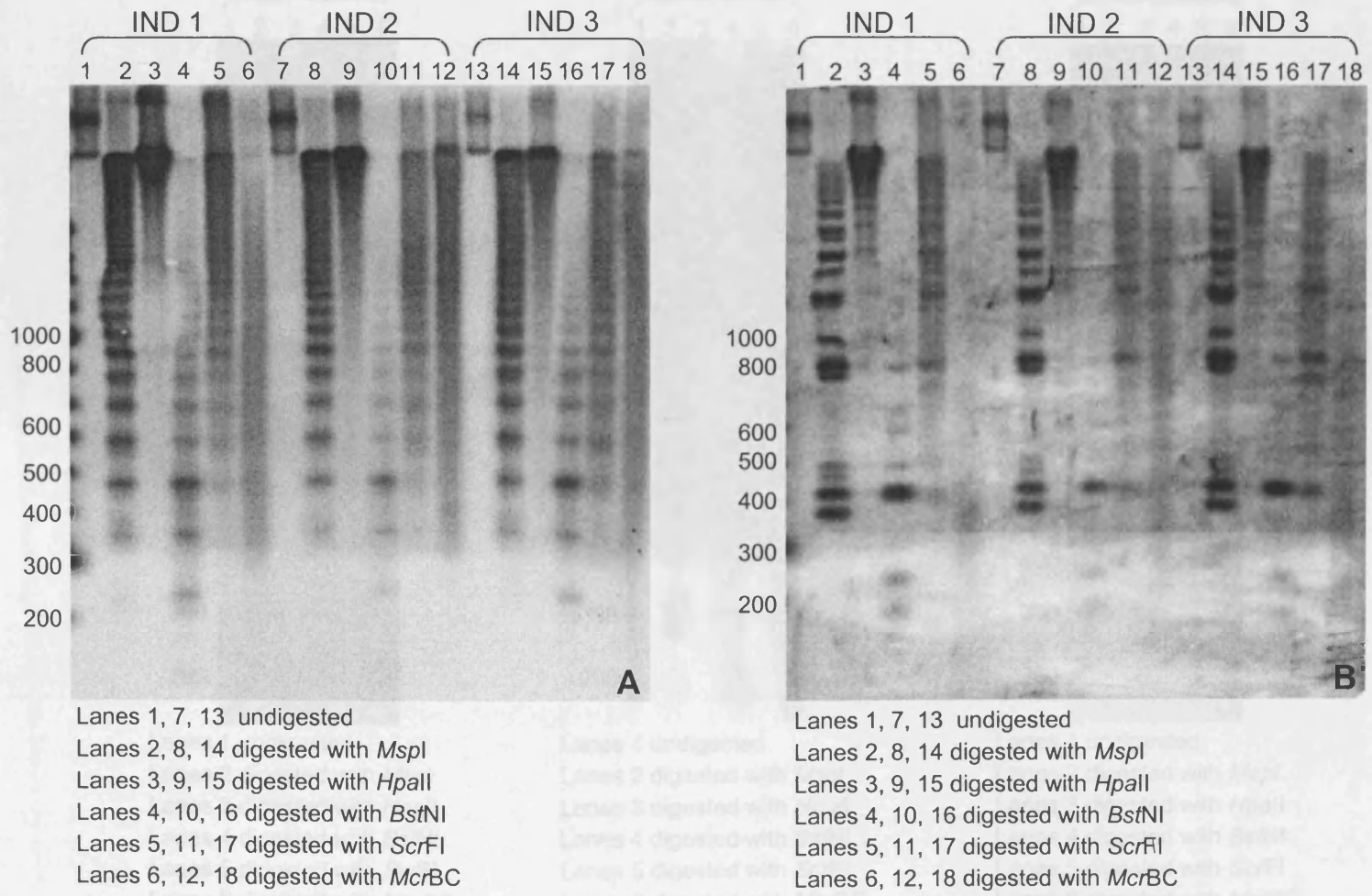
In allotetraploid *T. turgidum* ssp. *durum* (AB genome) marked bands in ladder pattern organization were observed with *Msp*I and *Bst*NI after Southern hybridization with a 120-bp repeat clone (see fig. 4. 8b, lanes 2, 4, 8, 10, 14 and 16). Methylation was not affecting some CpG sites (*Hpa*II showing digestion to some extent) and many CpNpG sites (*Scr*FI cutting to a large extent) as shown in fig. 4. 8b (see lanes 3, 5, 9, 11, 15 and 17). Asymmetrical methylation was revealed by a steady uniformly smeared signal after *Mcr*BC digestion (see fig. 4. 8b, lanes 6, 12 and 18). The 120-bp repetitive sequence family appeared more asymmetrically methylated than the total genomic DNA in *Ae. speltoides* (B genome), *T. tauschii* (D genome) and *T. turgidum* ssp. *durum* (AB genome) species as shown by comparing the uniformly smeared signal detected by Southern hybridization with 120-bp repeat clones (see lanes 12 and 18 in fig. 4. 7b and lanes 6, 12 and 18 in fig. 4. 8b) with the correspondent genomic DNA restrictions (see lanes 12 and 18 in fig. 4. 7a and lanes 6, 12 and 18 in fig. 4. 8a).

Genome-wide DNA methylation in hexaploid wheat *T. aestivum* 'Chinese Spring' (ABD genome) confirmed the features already observed for B (*Ae. speltoides*) and AB (*T. turgidum* ssp. *durum*) genomes: CpG were more methylated than CpNpG sites, asymmetrical methylation was present to some extent (not shown). After Southern hybridization with a 120-bp repeat strong bands in ladder pattern organization with *Msp*I and *Bst*NI were seen (see lanes 2, 4, 8, 10, 14 and 16 in fig. 4. 9a). Among methylation-sensitive enzymes, *Hpa*II cut only to some extent while digestion occurred with *Scr*FI to a large extent (see lanes 3, 5, 9, 11, 15 and 17 in fig. 4. 9a) Asymmetrical methylation was detected with *Mcr*BC as a marked uniformly smeared signal (see lane 6, 12 and 18 in fig. 4. 9a).

b) Whole-genome DNA methylation and methylation patterns of 120-bp repeat unit family in *Secale* and *Triticale* species

Genomic DNA digestions of individuals from *S. cereale* 'Petkus Spring' (R genome) and two varieties of *Triticale* (ABR genomes), 'Fidelio' and 'Lamberto', produced in the late '90s from Semundo-Danko Company and combining good traits of winter hardness and resistance to most diseases from parental species, were performed with isoschizomers *Msp*I-*Hpa*II and *Bst*NI-*Scr*FI for detecting methylation at CpG sites and CpNpG sites as well as *Mcr*BC restriction enzyme for detecting asymmetrical DNA methylation.

In *S. cereale* 'Petkus Spring' (R genome) *Msp*I and *Bst*NI cut to little and to some extent respectively (see lanes 2, 4 in fig. 4. 10a) and whole-genome methylation was detected at CpG sites, CpNpG sites but not at asymmetrical sites as the two methylation-sensitive enzymes *Hpa*II, *Scr*FI, and *Mcr*BC enzyme showed minimal digestion (see lanes 3, 5 and 6 in fig. 4. 10a). After Southern hybridization with a 120-bp repeat clone, very strong ladder patterns with *Msp*I and *Bst*NI were detected (see lanes 2 and 4 in fig. 4. 10b). Multiple bands resulting from different combinations of cutting sites within adjacent monomers (as shown in fig. 4. 6) were observed with both enzymes. Among methylation-sensitive enzymes, *Hpa*II cut to some extent, while *Scr*FI showed digestion to a large extent suggesting many CpNpG sites were not methylated (see lanes 3 and 5 in fig. 4. 10b). This was in contrast to genome-wide DNA methylation results where both symmetrical sites appeared to be methylated. A very strong uniformly smeared signal was detected after Southern hybridization with a 120-bp repeat clone with *Mcr*BC indicating that this repetitive family of sequences was



Lanes 1, 7, 13 undigested
Lanes 2, 8, 14 digested with *Msp*I
Lanes 3, 9, 15 digested with *Hpa*II
Lanes 4, 10, 16 digested with *Bst*NI
Lanes 5, 11, 17 digested with *Scr*FI
Lanes 6, 12, 18 digested with *Mcr*BC

Lanes 1, 7, 13 undigested
Lanes 2, 8, 14 digested with *Msp*I
Lanes 3, 9, 15 digested with *Hpa*II
Lanes 4, 10, 16 digested with *Bst*NI
Lanes 5, 11, 17 digested with *Scr*FI
Lanes 6, 12, 18 digested with *Mcr*BC

Figure 4. 9. Southern hybridizations with digoxigenin-labelled Pet(22594)25/42!3324 clone (A) and with biotin-labelled 5S rDNA pTa794 clone (B) of genomic DNA of *T. aestivum* 'Chinese Spring' (ABD genome) individuals digested with isoschizomers *Msp*I-*Hpa*II and *Bst*NI-*Scr*FI as well as *Mcr*BC restriction enzyme.

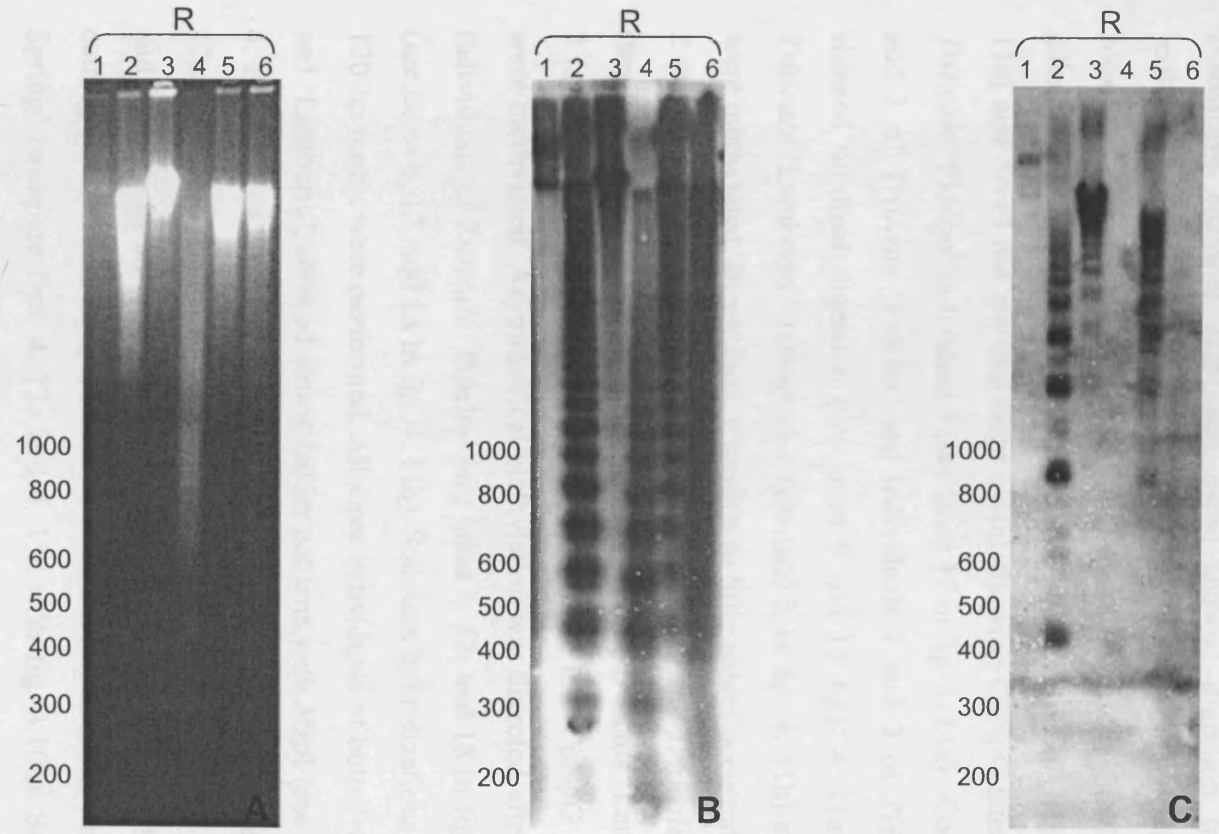
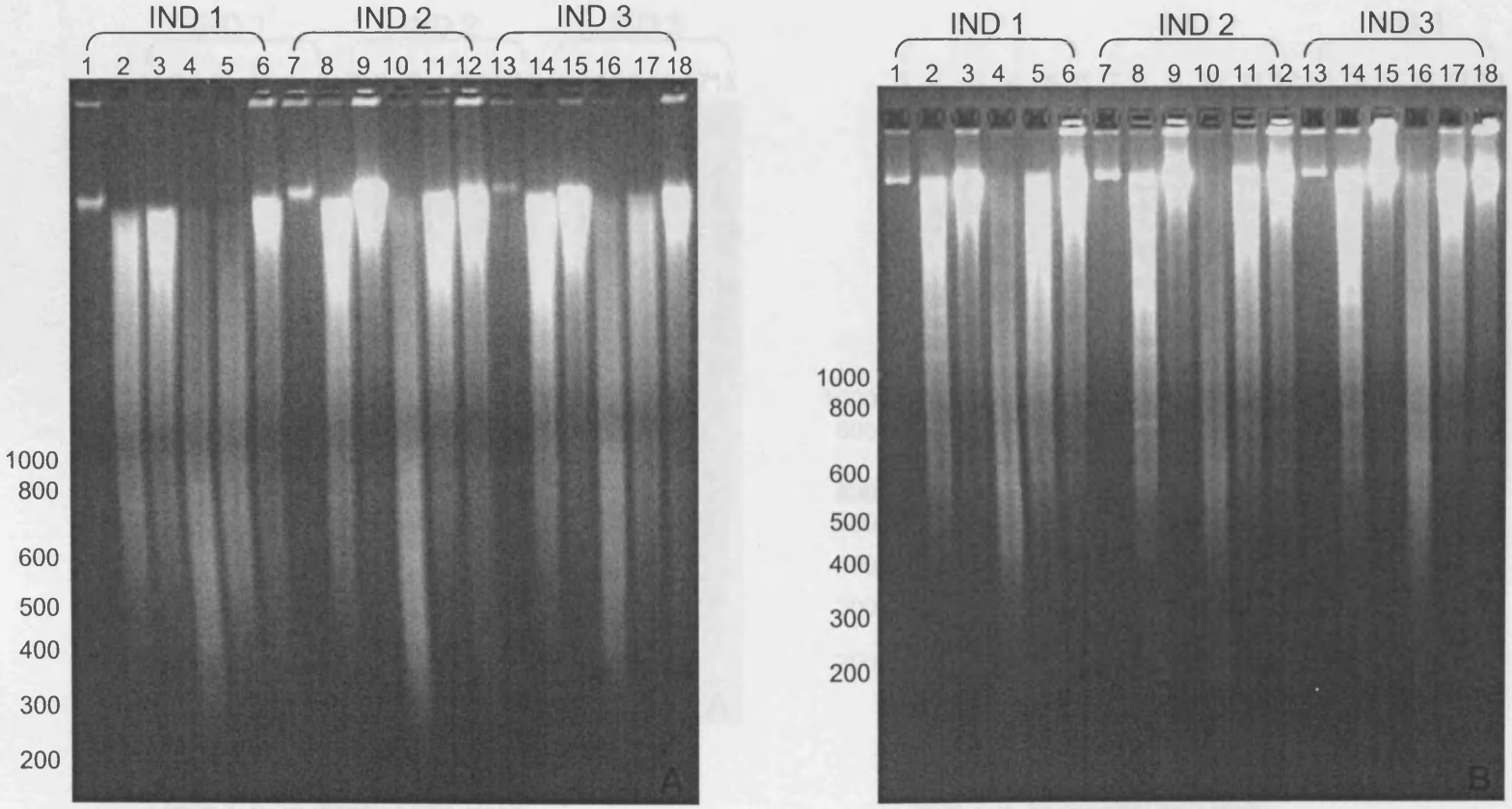


Figure 4. 10. A) Gel picture of genomic DNA of a *S. cereale* (R genome) individual digested with isoschizomers *Msp*I-*Hpa*II and *Bst*NI-*Scr*FI as well as *Mcr*BC restriction enzyme. Southern hybridizations with digoxigenin-labelled Pet(22594)25/42!3324 clone (B) and with biotin-labelled 5S rDNA pTa794 clone (C).

more asymmetrically methylated than the total genomic DNA (compare *Mcr*BC lanes in fig. 4. 10a-b).

Differences in whole-genome DNA methylation were observed between individuals of *Triticale* 'Fidelio' and 'Lamberto' (ABR genomes) at CpG, CpNpG and asymmetrical sites. In *Triticale* 'Fidelio and 'Lamberto' individuals methylation-insensitive enzymes *Msp*I and *Bst*NI showed digestion to a large extent in all individuals (see lanes 2, 4, 8, 10, 14 and 16 figs. 4. 11a-b). Among methylation-sensitive enzymes, *Hpa*II and *Scr*FI cut to a large extent showing that CpG and CpNpG sites were not methylated in *Triticale* 'Fidelio individual 1 (see lanes 3 and 5 in fig. 4. 11a) and *Scr*FI did cut to a large extent showing CpNpG sites were not methylated in *Triticale* 'Fidelio' individual 3 (see lanes 17 in fig. 4. 11a). Conversely, in individuals 2 and 3 of *Triticale* 'Fidelio' and individuals 2 and 3 of *Triticale* 'Lamberto' *Hpa*II showed minimal digestion (see lanes 9 and 15 figs. 4. 11a-b) or to little extent in *Triticale* 'Lamberto' individual 1 (see lane 3 in fig. 4. 11b) suggesting that CpG sites were methylated. In addition, digestion to little extent occurred with *Scr*FI in individual 2 of *Triticale* 'Fidelio' (see lane 11 in fig. 4. 11a) and *Triticale* 'Lamberto' individual 3 (see lane 17 in fig. 4. 11b) and to some extent in *Triticale* 'Lamberto' individuals 1 and 2 (see lanes 5 and 11 in fig. 4. 11b) suggesting that CpNpG sites in these individuals were methylated. Asymmetrical methylation was detected after *Mcr*BC digestion in all individuals of *Triticale* 'Fidelio' (see lanes 6, 12, and 18 in fig. 4. 11a) and 'Lamberto' (see lanes 6, 12, and 18 in fig. 4. 11b). Southern hybridizations of these gels with clones 120-bp repeat were performed. All three individuals of both *Triticale* varieties, 'Fidelio' and 'Lamberto', showed strong ladder patterns with *Msp*I (see lanes 2, 8 and 14 in figs. 4. 12a and 4. 13a), marked ladder patterns with *Bst*NI (see lanes 4, 10 and 16 in figs. 4. 12a and 4. 13a) and methylation-sensitive *Scr*FI (see lanes 5, 11 and 17 in figs. 4. 12a and 4. 13a). In all cases the multiple bands resulting from different combinations of cutting sites within adjacent monomers were less marked than in *S. cereale* 'Petkus Spring' (compare figs. 4. 12a and 4. 13a with fig. 4.10b). Smears of variable extent were present with methylation-sensitive *Hpa*II (see lanes 3, 9 and 15 in figs. 4. 12a and 4. 13a). In conclusion, in *Triticale* 'Fidelio' individual 1 and *Triticale* 'Lamberto' individual 1 methylation was not present in some CpG sites respectively, while in individuals of both *Triticale* varieties CpNpG sites were very weakly methylated or not methylated at all. Marked uniformly smeared signal with *Mcr*BC in both varieties was observed (see lanes 6, 12 and 18 in figs. 4. 12a and 4. 13a) while only digestion to



Lanes 1, 7, 13 undigested
Lanes 2, 8, 14 digested with *Msp*I
Lanes 3, 9, 15 digested with *Hpa*II
Lanes 4, 10, 16 digested with *Bst*NI
Lanes 5, 11, 17 digested with *Scr*FI
Lanes 6, 12, 18 digested with *Mcr*BC

Lanes 1, 7, 13 undigested
Lanes 2, 8, 14 digested with *Msp*I
Lanes 3, 9, 15 digested with *Hpa*II
Lanes 4, 10, 16 digested with *Bst*NI
Lanes 5, 11, 17 digested with *Scr*FI
Lanes 6, 12, 18 digested with *Mcr*BC

Figure 4. 11. Gel pictures of genomic DNA of *Triticale* (ABR genome) 'Fidelio' (A) and 'Lamberto' (B) individuals digested with isoschizomers *Msp*I-*Hpa*II and *Bst*NI-*Scr*FI as well as *Mcr*BC restriction enzyme.

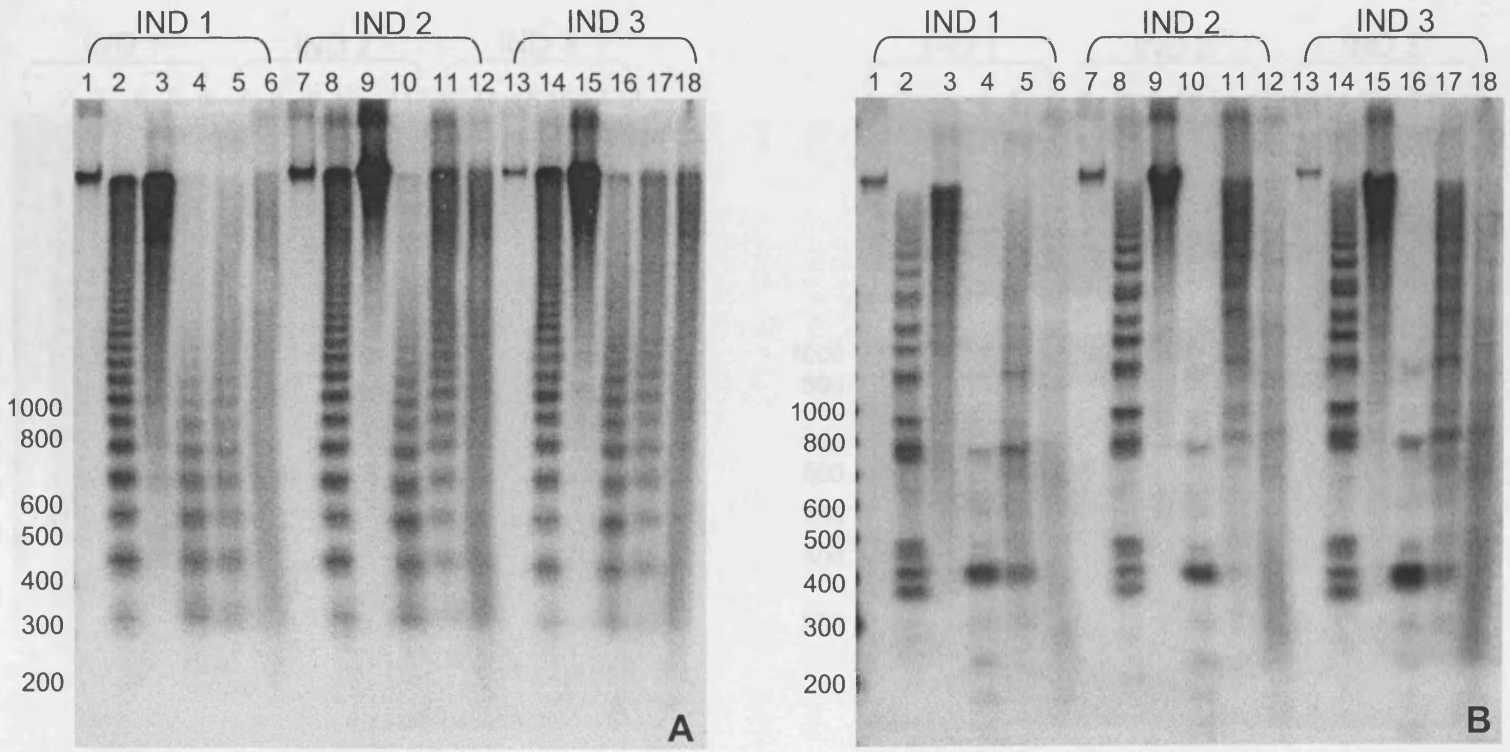
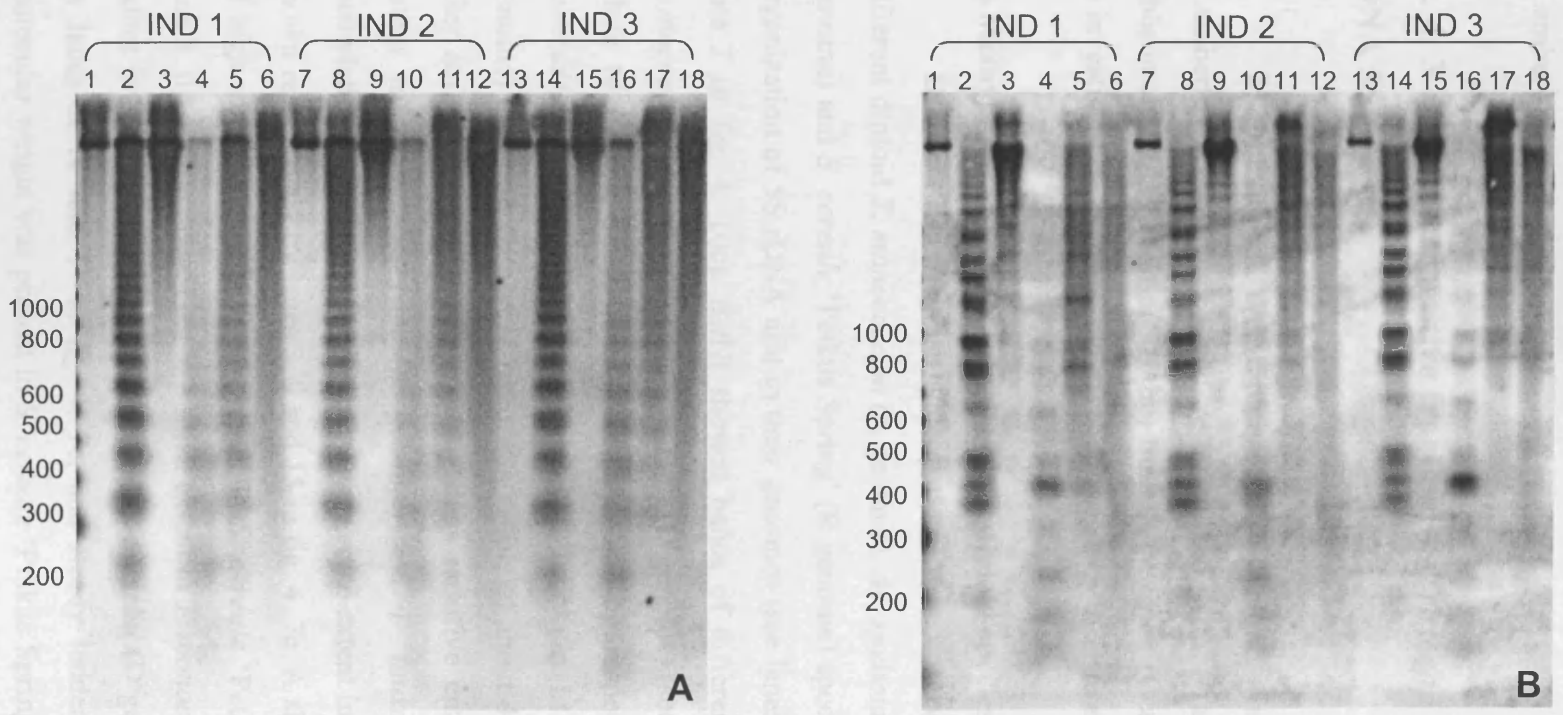


Figure 4. 12. Southern hybridizations with digoxigenin-labelled Tmono25/208!1416 clone (A) and with biotin-labelled 5S rDNA pTa794 clone (B) of genomic DNA of *Triticale* (ABR genome) 'Fidelio' individuals digested with isoschizomers *Msp*I-*Hpa*II and *Bst*NI-*Scr*FI as well as *Mcr*BC restriction enzyme.



Lanes 1, 7, 13 undigested
Lanes 2, 8, 14 digested with *Msp*I
Lanes 3, 9, 15 digested with *Hpa*II
Lanes 4, 10, 16 digested with *Bst*NI
Lanes 5, 11, 17 digested with *Scr*FI
Lanes 6, 12, 18 digested with *Mcr*BC

Lanes 1, 7, 13 undigested
Lanes 2, 8, 14 digested with *Msp*I
Lanes 3, 9, 15 digested with *Hpa*II
Lanes 4, 10, 16 digested with *Bst*NI
Lanes 5, 11, 17 digested with *Scr*FI
Lanes 6, 12, 18 digested with *Mcr*BC

Figure 4. 13. Southern hybridizations with digoxigenin-labelled Pet(22594)25/42!3324 clone (A) and with biotin-labelled 5S rDNA pTa794 clone (B) of genomic DNA of *Triticale* (ABR genome) 'Lamberto' individuals digested with isoschizomers *Msp*I-*Hpa*II and *Bst*NI-*Scr*FI as well as *Mcr*BC restriction enzyme.

variable extent in the correspondent genomic DNA restrictions (see lanes 6, 12 and 18 in figs. 4. 11a-b) confirming that 120-bp repetitive sequence family is more asymmetrically methylated than the total genomic DNA in *Triticale* 'Fidelio' and 'Lamberto' (ABR genomes).

4. 4. 3 Another repetitive fraction of *Triticeae* genomes: the ribosomal DNA 5S

Membranes of rye, diploid and polyploid wheats, *Triticale* 'Fidelio' and 'Lamberto' were stripped and re-hybridized with pTa794 clone that contains 5S ribosomal DNA (insert of 410 bp) from *T. aestivum* (Gerlach and Dyer 1980) in order to investigate another fraction of repetitive DNA in *Triticeae* genomes.

a) Methylation of 5S rDNA sequences in diploid species

Different and irregular ladder patterns were observed after *Msp*I digestion in the different diploid *T. monococcum* (A genome), *Ae. speltoides* (B genome), *T. tauschii* (D genome) and *S. cereale* 'Petkus Spring' (R genome) species corresponding to different organization of 5S rDNA unit in their genomes (see lanes 2, 8 and 14 in fig. 4. 7c and lane 2 in fig. 4. 10c). *Bst*NI showed bands of different length and intensity in *T. monococcum* (A genome) and *T. tauschii* (D genome) (see lanes 4 and 16 in figs. 4. 7c) while two weak bands (about 100 and 500 bp respectively) were present in *Ae. speltoides* (B genome) as shown in lane 10, fig. 4. 7c. In *S. cereale* 'Petkus Spring' (R genome) only one very weak band was visible of about 400 bp (see lane 4, fig. 4. 10c) after *Bst*NI digestion. Among methylation sensitive enzymes, *Hpa*II cut to a small extent in *T. monococcum* (A genome) and *T. tauschii* (D genome) suggesting methylation at CpG sites, while digested to some extent in *Ae. speltoides* (B genome) as shown respectively by lanes 3, 9 and 15 in fig. 4. 7c. A short ladder pattern with bands of high molecular weight was present in *S. cereale* 'Petkus Spring' (R genome) (see lane 3, fig. 4. 10c). *Scr*FI cut to a small extent in *T. monococcum* (A genome), to a large extent in *Ae. speltoides* (B genome) and *T. tauschii* (D genome) as shown respectively by lanes 5, 11 and 17 in fig. 4. 7c, and steady ladder pattern with bands of high molecular weight was present in *S. cereale* 'Petkus Spring' (R genome) (lane 5, fig. 4. 10c). Finally, asymmetrical methylation was present as a smeared signal in all species but different intensity: weak in *T. monococcum* (A genome), visible in *Ae. speltoides* (B

genome) and *T. tauschii* (D genome), and very weak in *S. cereale* 'Petkus Spring' (R genome) shown respectively in lanes 6, 12 and 18 in fig. 4. 7c and in lane 6 in fig. 4. 10c.

b) Methylation of 5S rDNA sequences in polyploid species

After re-hybridization with 5S rDNA pTa794 clone in *T. turgidum* ssp. *durum* (AB genome) and *T. aestivum* 'Chinese Spring' (ABD genome) a ladder pattern even if not regular was detected with *Msp*I (see lanes 2, 8 and 14 in figs. 4. 8c and 4. 9b), while only the 410 bp and a second band corresponding to about 800 bp were present with *Bst*NI, together with bands not always visible shorter than the monomer length (about 200 bp) (see lanes 4, 10 and 16 in figs. 4. 8c and 4. 9b). Among methylation-sensitive enzymes, after *Hpa*II digestion smears to some extent were detected in *T. turgidum* ssp. *durum* (AB genome) (see lanes 3, 9 and 15 in fig. 4. 8c) and to small extent in *T. aestivum* 'Chinese Spring' (ABD genome) (see lanes 3, 9 and 15 in fig. 4. 9b) indicating that some CpG sites were not methylated in 5S rDNA sequences. *Scr*FI cut bands corresponding to monomers, dimers and trimers (see lanes 5, 11 and 17 in figs. 4. 8c and 4. 9b) in both species suggesting that methylation was not present at many CpNpG sites of 5S rDNA subunits. After *Mcr*BC digestion, asymmetrical methylation was present with smears of variable intensity between individuals of both species (see lanes 6, 12 and 18 in figs. 4. 8c and 4. 9b).

Similarly, in *Triticale* 'Fidelio' individuals a ladder pattern even if not regular was detected with *Msp*I (lanes 2, 8 and 14 in fig. 4. 12b) and bands corresponding to about 400 bp (monomer), 800 bp (dimer) and 1200 bp (trimer) of different intensity together with bands not always visible shorter than the monomer length (about 200 bp) were observed with *Bst*NI (see lanes 4, 10 and 16 in fig. 4. 12b). Differences between individuals were observed at CpG sites (*Hpa*II shows digestion to large extent 1 and 3 compared to individual 2, see lanes 3, 9 and 15 in fig. 4. 12b). *Scr*FI in some individuals showed the 410-bp band together with additional bands corresponding to dimers and trimers with different intensity, confirming variable methylation patterns at CpNpG sites of 5S rDNA repeats (see lanes 5, 11 and 17 in fig. 4. 12b). Steady smeared signal with *Mcr*BC suggested asymmetrically methylation (see lanes 6, 12 and 18 in fig. 4. 12b). Similar results were obtained for 'Lamberto' (see fig. 4. 13b) with the only difference that *Hpa*II showed digestion to a small extension in all individuals (see lanes 3, 9 and 15 in fig. 4. 13b) compared to individuals 1 and 3 of *Triticale* 'Fidelio' (see

lanes 3, 9 and 15 in fig. 4. 12b) suggesting different levels of methylation at CpG sites between individuals of the two *Triticale* varieties.

4. 5 Discussion

4. 5. 1 Copy number of 120-bp repetitive sequences in diploid and polyploid *Triticeae* species

The copy number differences of 120-bp repetitive sequence family reported in diploid *Triticeae* species possibly represent a genomic diversification evolved during their evolution from a common ancestor (table 4. 2) and are in agreement with previous data reporting 120-bp repetitive sequence family a major DNA fraction in *S. cereale* (McIntyre, Pereira et al. 1990; Cuadrado, Ceoloni et al. 1995; Vershinin, Schwarzacher et al. 1995) and *Ae. speltoides* (Badaeva, Friebe et al. 1996), but present in much lower amount in other diploid species such as *T. monococcum* and *T. tauschii* (Mukai, Nakara et al. 1993). Overall, in most *Triticeae* species the copy number values of 120-bp repeat unit family estimated per nanogram of genomic DNA and per genome (table 4. 2) using Real Time PCR are in the range of 10^6-10^8/ng and 10^4-10^6/genome as previously reported for other repetitive sequence families, for example in *Vicia faba*, *Olea* and *Pennisetum* species (Kato, Yakura et al. 1984; Ingham, Hanna et al. 1993; Macas, Pozarkova et al. 2000; Contento, Ceccarelli et al. 2002).

T-test analysis showed significant differences in copy number of 120-bp repeat unit family sequences between allotetraploid *T. turgidum* ssp. *durum* (AB genome), allohexaploid *T. aestivum* 'Chinese Spring' (ABD genome), allohexaploid *Triticale* 'Fidelio' and their genome donors (table 4. 3). Firstly, 120-bp repetitive sequence copy number per haploid unreplicated genome in *Ae. speltoides* (B genome) individuals (1.27×10^6-1.36×10^6) and in *T. turgidum* ssp. *durum* (AB genome) individuals (2.25×10^5-2.97×10^5) were highly significantly different ($P=1.18 \times 10^{-4}$, see tables 4. 2 and 4. 3). By assuming that A genome-origin 120-bp repetitive sequences were not relevant in terms of copy number (in the range of 10^2 per unreplicated haploid genome) as estimated by Real Time PCR and confirmed by the absence of any detectable signal in Southern hybridization (fig. 4. 7b), the results suggested that some 120-bp repetitive sequences of B genome-origin were lost in allotetraploid *T. turgidum* ssp. *durum* (AB genome) after the allopolyploidization event, assuming that *Ae. speltoides* was the B genome donor of

Table 4. 3. P-values from T-test analysis of copy number per haploid unreplicated genome between diploid and polyploid *Triticeae* species.

Species A	Species B	P-value
Ae. speltoides	*T. turgidum* ssp. *durum*	1.18×10^{-4}
T. turgidum ssp. *durum*	*T. aestivum* 'Chinese Spring'	2.13×10^{-5}
S. cereale 'Petkus Spring'	*Triticale* 'Fidelio'	8.00×10^{-8}
T. turgidum ssp. *durum*	*Triticale* 'Fidelio'	1.39×10^{-4}

allotetraploid wheat. Secondly, there was also a highly significant difference ($P=2.13x10^{-5}$) in the copy number per haploid unreplicated genome of 120-bp repetitive sequences between *T. turgidum* ssp. *durum* (AB genome) individuals ($2.25\ x10^5$-$2.97x10^5$) and *T. aestivum* 'Chinese Spring' (ABD genome) individuals ($0.81\ x10^6$-$1.29x10^6$) (see tables 4. 2 and 4. 3). As *T. tauschii* (D genome) did not seem to largely contribute to 120-bp repetitive sequences of *T. aestivum* 'Chinese Spring' (ABD genome) due to its lower copy number (in the range of 10^4 per haploid unreplicated genome) the results suggested an amplification of AB genome-origin 120-bp repetitive sequences mainly as a consequence of the allopolyploidization event, but it cannot be excluded that some amplification also affected D genome-origin 120-bp repetitive sequences in *T. aestivum* 'Chinese Spring' (ABD genome). Finally, highly significant differences in copy number per unreplicated haploid genome of 120-bp repeat unit family sequences were found between *Triticale* 'Fidelio' (ABR genome) with $0.73x10^6$-$1.30x10^6$ and its parental species: *T. turgidum* ssp. *durum* (AB genome) with $2.25x10^5$-$2.97x10^5$ and *S. cereale* 'Petkus Spring' (R genome) with $3.53x10^6$- $3.81x10^6$ (see tables 4. 2 and 4. 3). Such differences could be the consequence of reduction mechanism of R genome-origin 120-bp repetitive sequences as previously reported for rye repetitive sequences in *Triticale* hybrids by *in situ* hybridization analysis (Seal and Bennet 1981) and by using AFLP and RFLP analysis (Ma, Fang et al. 2004) but also due to amplification of AB genome-origin 120-bp repetitive sequences, as both parental species are contributing largely to 120-bp repetitive sequence family of *Triticale* 'Fidelio' (ABR genome).

Changes in copy number of tandemly repetitive DNA sequences, mainly reduction, have been previously reported in polyploids. For example, a consistent reduction in copy number (from 50- to 1000-fold) of a telomeric repeated sequence family of *Ae. speltoides* (Spelt1) was observed in cultivated tetraploid and hexaploid wheats in the attempt to reorganize chromosome distribution in the nucleus, balance DNA content and synchronize DNA replication in hybrid genomes (Petsova, Goncharov et al. 1998). Elimination of repetitive DNA sequences together with the phenomenon of rapid, non-random elimination of non coding low-copy sequences (Feldman, Liu et al. 1997; Liu, Vega et al. 1998; Ozkan, Levy et al. 2001; Shaked, Kashkush et al. 2001) could contribute to a rapid decrease of genomic DNA in *Triticeae* allopolyploids. In fact, it has been reported that about 2 pg DNA at 2C, corresponding to 4-8% of genome size, were eliminated in synthetic tetraploid and hexaploid *Triticeae* species in order to

stabilize their genomes (Ozkan, Tuna et al. 2003). Similarly, a decrease in DNA content in hexaploid and octoploid *Triticale* was estimated to be about 28-30% and 9% respectively (Boyko, Badaev et al. 1984). It is now generally accepted that genome downsizing is a common event in many polyploids as evolutionary forces possibly select within a species those individuals with reduced amount of DNA because they have lost non-essential DNA sequences (Leitch and Bennett 2004).

4. 5. 2 Genome-wide DNA methylation in diploid and polyploid *Triticeae* species

DNA methylation affect 20-30% of the total cytosines in plant genomes (Shapiro 1976), and among cereals it was estimated that 22% of the total cytosines are methylated in wheat genome (Wagner and Capesius 1981). Presumably, the overall amount of methylated cytosines does not differ between *Triticeae* species as much as its partitioning among the three DNA methylation codes. In fact, in different *Triticeae* species, characteristic whole-genome DNA methylation levels and patterns were observed. For example diploid wheat species *T. monococcum* (A genome) and *T. tauschii* (D genome) showed similar high whole-genome DNA methylation levels at CpG, CpNpG and asymmetrical sites (fig. 4. 7a). On the other hand, in *Ae. speltoides* (B genome), *T. turgidum* ssp. *durum* (AB genome) and *T. aestivum* Chinese Spring' (ABD genome) CpG sites were more methylated than CpNpG sites and asymmetrical methylation was present to some extent (figs. 4. 7a and 4. 8a) independently from ploidy level. These results agree with previous surveys about genome-wide DNA methylation in wheat, reporting approximately 80% of cytosines in CpG sites methylated and CpNpG methylation being about 30 to 40% of total cytosine methylation (Gruenbaum, Naveh-Many et al. 1981). In addition, in *S. cereale* 'Petkus Spring' (R genome) whole-genome DNA methylation levels were high at CpG sites, CpNpG sites but not at asymmetrical sites (fig. 4. 10a). Finally, differences in whole-genome DNA methylation were observed between individuals of *Triticale* 'Fidelio' and 'Lamberto' (ABR genomes) at CpG, CpNpG and asymmetrical sites suggesting variable DNA methylation patterns in these allopolyploids (fig. 4. 11a-b).

4. 5. 3 Methylation of 120-bp repeat unit family sequences in diploid and polyploid *Triticeae* species

Within the 120-bp repeats of *Secale* and *Triticum/Aegilops* CCGG sites are not present in the same amount of CCNGG sites: a CCGG site (or less often a CCNGG site) is present in most clones of *Secale* and *Triticum/Aegilops*, whereas a second CCGG site is present in most *Secale* clones but less frequently in *Triticum/Aegilops* clones and very rarely becomes a CCNGG site as previously shown (fig. 4. 1a-b). In fact, in *S. cereale* 'Petkus Spring' and in *Triticale* 'Fidelio' multiple bands were visible in Southern hybridization with 120-bp repeat unit family clones resulting from different combinations of CpG and CpNpG cutting sites within adjacent monomers.

The results obtained by Southern hybridization with 120-bp repeat unit family clones showed that in diploid and polyploid wheat species: *T. tauschii* (D genome), *T. turgidum* ssp. *durum* (AB genome), and *T. aestivum* Chinese Spring' (ABD genome) (figs. 4. 7b, 4. 8b, 4. 9b), and in *S. cereale* 'Petkus Spring' (R genome) (fig. 4. 10b) high methylation levels are present at CpG sites, but lower at CpNpG sites possibly due to the unequal number of these sites in 120-bp repeats or truly representing differences between the two DNA methylation codes. Furthermore, in *Ae. speltoides* (B genome) both CpG and CpNpG sites appeared poorly methylated (fig. 4. 7b). Finally, in *Triticale* 'Fidelio' and 'Lamberto' variable methylation pattern at CpG and CpNpG sites were observed between individuals (figs. 4. 12b and 4. 13b): in some individuals of both varieties methylation was not present in some CpG sites while CpNpG sites were very weakly methylated or not methylated at all. Conversely, previous data for other species, such as rice and tobacco, reported high DNA methylation levels at symmetrical CpG and CpNpG sites of repetitive sequence families (Matyasek, Gazdova et al. 1997; Kovarik, Koukalova et al. 2000; Cheng, Stupar et al. 2001).

Smeared signals after *Mcr*BC digestion were present in all diploid and polyploid wheat species (figs. 4. 7b, 4. 8b and 4. 9b), rye (fig. 4. 10b) and *Triticale* varieties (figs. 4. 12b and 4. 13b) indicating that asymmetrical methylation is a common feature of 120-bp repeat unit family.

4. 5. 4 Alteration of methylation patterns of 120-bp repeat unit family sequences in diploid and polyploid *Triticeae* species

Methylation repatterning is a possible consequence of an allopolyplodization event as shown in *Brassica*, *Triticeae* and *Arabidopsis-Cardaminopsis* synthetic allopolyploids (Song, Lu et al. 1995; Liu, Vega et al. 1998; Comai, Tyagi et al. 2000; Shaked, Kashkush et al. 2001; Kashkush, Feldman et al. 2002; Madlung, Masuelli et al. 2002; Han, Fedak et al. 2003; Ma, Fang et al. 2004). In newly synthesised *Triticeae* allopolyploids DNA methylation changes following an allopolyploidization event were reported to happen in mobile DNA elements, such as retrotransposons, and 26S rDNA genes (Shaked, Kashkush et al. 2001; Kashkush, Feldman et al. 2002; Han, Fedak et al. 2003). In hexaploid and octoploid *Triticale* hybrids changes in DNA methylation patterns in comparison to their parental species were assessed by using AFLP and RFLP analysis and resulted in the appearance and disappearance of DNA fragments (Ma, Fang et al. 2004). Similarly, an alteration of DNA methylation patterns at symmetrical CpG and CpNpG sites of 120-bp repeat unit family was observed in allotetraploid wheat *T. turgidum* sub. *durum* (AB genome) and allohexaploid *Triticale* (ABR genome) by comparing these natural allopolyploids with their corresponding parental species. For example, ignoring the minimal contribution of A genome-origin 120-bp repetitive sequences to *T. turgidum* ssp. *durum* (AB genome), the B genome-origin 120-bp repetitive sequences appeared more methylated in *T. turgidum* ssp. *durum* (AB genome), especially at CpG sites (compare lane 9 in fig. 4. 7b with lanes 3, 9, 15 in fig. 4. 8b) but not at CpNpG sites, than in diploid B genome of *Ae. speltoides* assuming that this was the B genome donor of allotetraploid wheat. Conversely, no major changes seemed to happen in asymmetrical methylation as similar uniformly smeared signal was observed in *Ae. speltoides* (B genome) and *T. turgidum* ssp. *durum* (AB genome) suggesting similar distribution of multiple overlapping half-sites of the form $(G/A)^{m}C$. In *Triticale* 'Fidelio' and 'Lamberto' (ABR genomes) a possible decrease in DNA methylation of 120-bp repeat unit family sequences at CpNpG sites was observed (lanes 5, 11 and 17 in figs. 4. 12a and 4. 13a) in comparison to parental donors *S. cereale* 'Petkus Spring' (R genome) and *T. turgidum* ssp. *durum* (AB genome) (lane 5 fig. 4. 10b and lanes 5, 11 and 17 in fig. 4. 8b, respectively). From the results obtained it cannot be detected if a preferential decrease in DNA methylation of 120-bp repetitive sequences from only one of the two parental genomes or from both in equal extent has occurred when they merged in the allohexaploid *Triticale*.

Conversely, by comparing methylation levels at CpG, CpNpG and asymmetrical sites of 120-bp repeat unit family sequences of *T. aestivum* 'Chinese Spring' (ABD

genome) (fig. 4. 9a) with its parental donors: *T. turgidum* ssp. *durum* (AB genome) (fig. 4. 8b) and *T. tauschii* (D genome) (fig. 4. 7b), the three DNA methylation codes did not show major differences, indicating that no major changes in 120-bp repetitive sequence family resulted by the hybridization event that originated *T. aestivum* Chinese Spring' (ABD genome).

4. 5. 5 Changes in copy number and methylation patterns of 120-bp repeat unit family contribute differently to *Triticeae* allopolyploid genomes

The overall data obtained by Real Time PCR and Southern hybridization gave evidence that 120-bp repetitive sequence family underwent a series of modifications in copy number and DNA methylation status during the events of allopolyploidization originating tetra- and hexaploid wheats *T. turgidum* ssp. *durum* (AB genome), *T. aestivum* 'Chinese Spring' (ABD genome), and hexaploid *Triticale* (ABR genome). Reduction and amplification in copy number were not always accompanied by the same DNA methylation changes. In fact, significant copy number reduction and increase in DNA methylation at CpG sites were observed by comparing 120-bp repetitive sequences in *Ae. speltoides* (B genome) and the B genome-specific 120-bp repetitive sequences of *T. turgidum* ssp. *durum* (AB genome). Conversely, significant copy number amplification and no DNA methylation changes in any of three DNA methylation codes were registered in 120-bp repetitive sequences of *T. aestivum* 'Chinese Spring' (ABD genome) compared to 120-bp repetitive sequences of its donors *T. tauschii* (D genome) and *T. turgidum* ssp. *durum* (AB genome). Finally, decrease in DNA methylation at CpNpG sites possibly happened in 120-bp repetitive sequences of *S. cereale* 'Petkus Spring' (R genome) and/or *T. turgidum* ssp. *durum* (AB genome) that originated hexaploid *Triticale* (ABR genome), together with mechanisms of reduction of 120-bp repeat unit family sequences from *S. cereale* 'Petkus Spring' (R genome) and/or amplification of 120-bp repeat unit family sequences from *T. turgidum* ssp. *durum* (AB genome) respectively.

4. 5. 6 Methylation of 5S rDNA repetitive sequences in diploid and polyploid *Triticeae* species

Southern hybridization results showed variable DNA methylation levels and patterns of 5S rDNA repetitive sequences at CpG, CpNpG and asymmetrical sites in different diploid species, possibly related to differential gene expression. For example, methylation at CpG sites of 5S rDNA repetitive sequences in *T. monococcum* (A genome) and *T. tauschii* (D genome) was higher than *Ae. speltoides* (B genome), while CpNpG sites of *T. monococcum* (A genome) were more methylated than *Ae. speltoides* (B genome) and *T. tauschii* (D genome) (fig. 4. 7c). Asymmetrical methylation of *T. monococcum* (A genome) was higher than *Ae. speltoides* (B genome) and *T. tauschii* (D genome) (fig. 4. 7c). In *S. cereale* 'Petkus Spring' (R genome) 5S rDNA repetitive sequences were not methylated at all CpG sites and methylation did not affect many CpNpG sites (fig. 4. 10c) in agreement with previous estimates of 85% at CpG and 53% at CpNpG sites respectively being methylated in 5S rDNA sequences (Fulnecek, Matyasek et al. 2002). Asymmetrical methylation was negligible in *S. cereale* 'Petkus Spring' (R genome) (fig. 4. 10c) as previously reported (Fulnecek, Matyasek et al. 2002).

In *T. turgidum* ssp. *durum* (AB genome) and *T. aestivum* 'Chinese Spring' (ABD genome) methylation at CpG sites was stronger than CpNpG sites with no major differences between individuals (figs. 4. 8c and 4. 9c). Conversely, variable levels of DNA methylation were observed within and between *Triticale* 'Fidelio' and 'Lamberto' (ABR genome) at CpG and CpNpG sites of 5S rDNA repeats (figs. 4. 12c and 4. 13c). However, the 5S ribosomal DNA fraction in *T. aestivum* 'Chinese Spring' (ABD genome) and in *Triticale* 'Fidelio' and 'Lamberto' (ABR genome) showed DNA methylation patterns at symmetrical and asymmetrical sites more similar to *T. turgidum* ssp. *durum* (AB genome) than to the other diploid parental species, *T. tauschii* (D genome) and *S. cereale* (R genome) respectively.

By comparing DNA methylation levels and patterns of 5S rDNA repetitive sequences and 120-bp repeat unit family sequences differences were observed between *Triticeae* genomes. In *Ae. speltoides* (B genome) for example, 120-bp repeat family sequences were poorly methylated at CpG and CpNpG sites while methylation at CpG and CpNpG sites was present in 5S rDNA repetitive sequences; in both repetitive sequence families methylation at asymmetrical sites was stronger than whole-genome DNA methylation. In *S. cereale* 'Petkus Spring' (R genome) 120-bp repetitive sequence family was methylated at CpG sites, much less methylated at CpNpG sites, but strongly asymmetrically methylated, while in 5S rDNA sequences methylation levels at

symmetrical CpG and CpNpG sites were more similar and asymmetrical methylation was negligible. In conclusion, each repetitive sequence family characterized by specific methylation patterns at CpG, CpNpG and asymmetrical sites contributed to establish species-specific whole-genome DNA methylation patterns in *Triticeae* species.

Chapter 5: Methylation patterns of DNA and histone H3 on chromosomes of Triticeae species

5. 1 Summary

In this chapter methylation patterns of 120-bp repeat unit sequence family in diploid and polyploid *Triticeae* species have been analysed on chromosome spreads by immunostaining with anti-methylcytosine antibody combined with *in situ* hybridization and compared to whole-genome DNA methylation patterns. Along chromosomes, diploid species of rye and wheats showed uniform whole-genome cytosine methylation signals. In *S. cereale* (R genome) and *Ae. speltoides* (B genome) telomeric-sites that are part of large subtelomeric heterochromatic blocks, containing 120-bp repeat unit sequence family, were partially or not-methylated, while *T. monococcum* (A genome) and *T. tauschii* (D genome) chromosomes were fully methylated at their chromosome ends and thus their small telomeric-sites containing 120-bp repetitive sequences. Conversely, allopolyploid wheats *T. turgidum* sub. *durum* (AB genome) and *T. aestivum* 'Chinese Spring' (ABD genome), and allohexaploid *Triticale* (ABR genome) varieties had more unevenly distributed whole-genome and 120-bp repeat unit sequence family methylation patterns. In addition, histone methylation patterns of histone H3 at Lisine 9 were analysed, without observing major differences between diploid *S. cereale* and hexaploid *Triticale* both showing uniformly distributed histone methylation signal along chromosomes.

5. 2 Introduction

Whole-genome DNA methylation can be estimated with different methods. Restriction of total genomic DNA with enzyme pairs either sensitive or not to methylated cytosines, as shown in the previous chapter, can be informative about the overall amount of symmetrical and asymmetrical methylation sites present in a genome. Recently, a genome-wide approach was carried out by comparing two partial genomic libraries of maize, one made in a *E. coli* strain with *Mcr*BC enzyme activity that can

digest methylated DNA sequences and the second one in a strain defective in *Mcr*BC enzyme activity. In first library only 3.3 % of clones contain retrotransposons and other annotated repeats while the in second library 48.7% of clones, showing that DNA methylation mainly affects transposable elements and other repeats in a genome (Rabinowicz, Schutz et al. 1999). In this chapter, a cytological approach has been performed to visualize directly on chromosome spreads how whole-genome cytosine methylation is distributed in different genomes, chromosome pairs and chromosome regions of several *Triticeae* species by immunostaining with anti-methylcytosine antibody.

Immunostaining with anti-methylcytosine antibody has been performed in *A. thaliana* interphase nuclei and strong signals were observed in correspondence of chromocenters with the exception of rDNA containing regions (Jasencakova, Soppe et al. 2003). By using the same approach previous works investigated DNA methylation patterns on chromosome spreads of two species with large genomes such as *Allium cepa* and *Vicia faba*, as well as in *Melandrium album*, a dioecious species with heteromorphic sex chromosomes. In *A. cepa* metaphases methylated cytosines appeared at several chromosomal locations but strong signal were detected at telomeric sites that correspond to GC-rich satellite DNA sequences (Barnes, James et al. 1985; Castiglione, Giraldi et al. 1995). In *V. faba* DNA methylation was present at telomeric and subtelomeric regions as well as several intercalary sites so that the authors suggested a link with repetitive DNA sequences that were widespread along the chromosomes (Frediani, Giraldi et al. 1996). Finally, in *M. album* unevenly distributed DNA methylation signal was reported for sex chromosomes in male and female cells, but no prominent differences for autosomes in both male and female metaphases (Siroky, Castiglione et al. 1998). Among cereal species, whole-genome DNA methylation assessed by immunostaining with anti-methylcytosine antibody was reported for *Triticale* 'Lasko' (Castilho, Neves et al. 1999). In this hybrid species unevenly distributed cytosine methylation was found along chromosomes with enhanced or reduced signal in different chromosome regions, and after genomic *in situ* hybridization no differences in DNA methylation level or distribution were found between chromosomes belonging to R and AB genomes that are the genome donors of hexaploid *Triticale*. However, this work did not identify individual R and AB chromosomes of hexaploid *Triticale* and did not compare them to chromosomes of diploid rye and allotetraploid wheat species.

Together with DNA methylation, histone H3 methylation at Lysine 9 has been suggested as a biochemical mark of heterochromatin (Richards and Elgin 2002). Histone H3 lysines can be monomethylated, dimethylated and trimethylated. By immunostaining with anti-dimethylated histone H3 at Lysine 9 antibody Houben *et al.* (2003) analysed several monocot and dicot plant species of different genome size, showing that histone H3 methylation patterns at Lysine 9 were related to nuclear DNA content. In *A. thaliana* data obtained by immunostaining and chromatin immunoprecipitation assays showed that monomethylated and dimethylated histones H3 at Lysine 9, but not trimethylated ones, were present in the heterochromatin (Jackson, Johnson et al. 2004). More recently, it has been confirmed for *A. thaliana* the monomethylated and dimethylated status of histone H3 at Lysine 9 in the heterochromatin at pericentromeres and nucleolar organizing regions (NORs) of chromosomes together with chromocenters of interphase nuclei (Fuchs, Jasencakova et al. 2005). Fuchs *et al.* (2005) also extended to other species, such as *V. faba* and *H. vulgare*, the survey of chromosomal distribution of histone H3 methylation marks at Lysines 9 and 27 suggesting the possibility of species-specific patterns of histone H3 methylation.

The aim of the present work was to present an analysis of both features of heterochromatin, DNA methylation patterns and histone H3 at Lysine 9 methylation patterns, on chromosome spreads of several diploid and polyploid *Triticeae* species. Firstly, by using the cytological approach of performing immunostaining with anti-methylcytosine antibody followed by *in situ* hybridization with 120-bp repeat unit family clones necessary for karyotyping, DNA methylation levels has been detected in different genomes, chromosome pairs and/or chromosome regions of diploid and polyploid *Triticeae* species. Secondly, by comparing whole-genome and 120-bp repeat unit family methylation patterns any alterations or changes happened during the evolution has been evaluated at the entire whole-genome level as well as in such a widespread repetitive sequence family of *Triticeae* genomes. Finally, by the analysis of histone H3 at Lisine 9 methylation patterns in diploid *S. cereale* and hexaploid *Triticale* a comparison about how DNA and histone methylation patterns are affected by genome size and/or ploidy level has been discussed.

5. 3 Materials and methods

5. 3. 1 Immunostaining with anti-methylcytosine antibody and *in situ* hybridization

As described in Material and Methods chapter, seedling root tips of *T. monococcum, Ae. speltoides, T. tauschii, S. cereale* 'Petkus Spring', *T. turgidum* sub. *durum, T. aestivum* 'Chinese Spring', *Triticale* 'Fidelio' and 'Lamberto' were treated with ice-water; three to five root tips per seed were collected, fixed in ethanol-acetic acid, digested with proteolytic enzymes and squashed in 60% acetic acid to obtain chromosome spreads. Slides preparations were treated with RNAse and pepsin, dehydrated in an ethanol series, and air- dried. As described in Materials and Methods chapter, immunostaining with anti-methylcytosine antibody (5μg/ml per slide) has been performed and the signal detected with Alexa 488-conjugated secondary antibody. Well spread metaphases were recorded and photographed. Three to ten/fifteen metaphases were analysed.

Slides preparations of the following species *Ae. speltoides, S. cereale* 'Petkus Spring', *T. turgidum* sub. *durum, T. aestivum* 'Chinese Spring', *Triticale* 'Fidelio' were reopened, the methylation signal removed by washes, and *in situ* hybridization experiments performed with biotin-labelled 120-bp repeat unit family clones (see asterisks in fig. 3. 5b and table 3. 5) as described in Materials and Methods chapter. The hybridization mixture consisted of 50 to 100 ng/slide of each probe, 50% formamide, 2xSSC, 10% dextran sulphate, 0.125% SDS, 0.125mM EDTA and 1μg salmon sperm DNA. Probes and chromosomes were denatured together at 70 to 80°C for 6 min and left to cool slowly to 37°C for overnight hybridization. Post-hybridization washes were performed in high stringency conditions and hybridization sites were detected by streptavidin conjugated to Alexa 594. Metaphases were re-located and photographed. Two to three individuals per species were karyotyped according to wheat and rye karyotypes (Mukai, Nakara et al. 1993; Cuadrado and Jouve 1994; Cuadrado and Jouve 1995) and two slides per individual with about three to five complete metaphases each were analysed.

In conclusion, the possibility to combine immunostaing with anti-methylcytosine antibody and *in situ* hybridization with 120-bp repeat unit family clones

allowed a comparative analysis of methylation patterns and karyotypes on the same metaphases in individuals of diploid and polyploid *Triticeae* species.

5. 3. 2 Immunostaining with anti di-methylated histone H3 at Lysine 9 antibody

Immunostaining with anti-di-methylated histone H3 at Lysine 9 has been performed on seedling root tips of diploid *S. cereale* 'Petkus Spring' and hexaploid *Triticale* 'Fidelio' as described in Materials and Methods chapter. Three to five root tips per seed were collected, fixed in freshly made paraformaldehyde solution, digested with proteolytic enzymes and squashed in 0.5% Triton X-100 to obtain chromosome spreads. Slides preparations were not allowed to dry out and quickly transferred to 1x KPBS before a re-fixation step in freshly made paraformaldehyde and a following treatment in freshly made sodium borohydride solution. After non-specific sites were blocked by incubation with BSA, the primary antibody incubation step was performed with anti-di-methylated histone H3 at Lysine 9 antibody (1:200 dilution) followed by incubation with Alexa 488-conjugated secondary antibody. Variable concentrations of BSA and secondary antibody were tested on both species. Well spread metaphases were recorded and photographed. Two slides with three to five metaphases each from three individuals for each species were analyzed.

5. 4 Results

5. 4. 1 Whole-genome DNA methylation and methylation patterns of 120-bp repeat unit family in established *Triticeae* allopolyploids and their parental species

Whole-genome and 120-bp repeat unit family methylation patterns were analysed by immunostaining with anti-methylcytosine antibody and subsequent *in situ* hybridization in three individuals of *T. monococcum, Ae. speltoides, T. tauschii, S. cereale* 'Petkus Spring', *T. turgidum* sub. *durum, T. aestivum* 'Chinese Spring', *Triticale* 'Fidelio' and 'Lamberto' previously analysed by Real Time PCR and Southern

hybridization in chapter 4. In few cases additional individuals were used to complete karyological analyses.

Triticum monococcum* and *Triticum tauschii

Three individuals of the diploid species *T. monococcum* (A genome, 2n=14) and *T. tauschii* (D genome, 2n=14) were analysed by immunostaining with anti-methylcytosine antibody. As reported in table 5. 1, in the majority of analysable metaphases of all individuals of *T. tauschii* (87.5%, 87.5% and 80 %, respectively) and individual 1 of *T. monococcum* (90.9%) DNA methylation patterns were uniform along the euchromatin of all fourteen chromosomes including the centromeric and telomeric regions. No differences between chromosomes were detectable. An example for *T. tauschii* individual 11 is given in fig. 5. 1. However, individuals 7 and 8 of *T. monococcum* had unevenly distributed methylation signal in 50.0% and 63.6% of metaphases, respectively (see table 5. 1) as shown for individual 8 in fig. 5. 2a. In these individuals, one or two chromosome arms had enhanced signal, some chromosomes lacked methylation signal at centromeres but telomeres are generally fully methylated.

Subsequent *in situ* hybridization was not performed as identification of chromosomes showing uniform DNA methylation signal was not thought to be necessary. Hoeever, the 120-bp repeat unit sequence family present only in some chromosome ends of wheat D genome (in low copy number) and with only few copies in wheat A genome is presumably uniformly and fully methylated at its locations.

Aegilops speltoides

In diploid *Ae. speltoides* (B genome, 2n=14) individuals 2, 5 and 6 were analysed. As reported in table 5. 1, uniform distribution of methylated cytosines was observed in most metaphases (83.3%) of individual 2. As slide preparations of individuals 5 and 6 were not good enough, two additional individuals, A and B, were subsequently analysed confirming a quite uniform distribution of DNA methylation along chromosomes in the majority of their metaphases (66.7% in both individuals, see table 5. 1). An example is given for individual A in fig. 5. 3. I*n situ* hybridization was carried out on slides from individuals 2, A and B and three to five metaphases for each individual analysed. As shown in fig. 5. 3d, chromosomes were karyotyped according to previous analyses (Badaeva, Friebe et al. 1996). A variable number between 10 and 20 of telomeres (mode=16) did not show methylation signal in metaphases of individuals 2, A and B (see fig. 5. 3e-f). Often the signal was absent at one or both telomeres of the

Table 5. 1. DNA methylation patterns in individuals of diploid *Triticeae* species after immunostaining with anti-methylcytosine antibody.

Species	Individual	number of metaphases photographed	number of metaphases analysable*	number of metaphases (uniform)	number of metaphases (uneven)	% of analysed metaphases (uniform)**
Triticum monococcum	Tm 1	11	11	10	1	90.9%
	Tm 7	9	8	4	4	50.0%
	Tm 8	11	11	4	7	36.4%
(A genome, 2n=14)	TOT	31	30	18	12	60.0%
Species	**Individual**	**number of metaphases photographed**	**number of metaphases analysable***	**number of metaphases (uniform)**	**number of metaphases (uneven)**	**% of analysed metaphases (uniform)****
Aegilops speltoides	Ae2	6	6	5	1	83.3%
	Ae5	2	1	1	-	100.0%
	Ae6	2	-	-	-	-
	AeA	11	9	6	3	66.7%
(B genome, 2n=14)	AeB	15	12	8	4	66.7%
	TOT	36	28	20	8	71.4%
Species	**Individual**	**number of metaphases photographed**	**number of metaphases analysable***	**number of metaphases (uniform)**	**number of metaphases (uneven)**	**% of analysed metaphases (uniform)****
Triticum tauschii	Tt9	8	8	7	1	87.5%
	Tt11	10	8	7	1	87.5%
(D genome, 2n=14)	Tt12	5	5	4	1	80.0%
	TOT	23	21	18	3	85.7%

Table 5. 1. continued

Species	Individual	number of metaphases photographed	number of metaphases analysable*	number of metaphases (uniform)	number of metaphases (uneven)	% of analysed metaphases (uniform)**
Secale cereale 'Petkus Spring' (R genome, 2n=14)	Pt2	11	8	4	4	50.0%
	Pt3	5	5	4	1	80.0%
	Pt4	5	4	4	-	100.0%
	PtA	11	11	8	3	72.7%
	PtB	10	10	8	2	80.0%
	TOT	42	38	28	10	73.7%

* Cells with incomplete chromosome number, distorsion or staining artefacts were not analysed.

**100% equals to analysed cells.

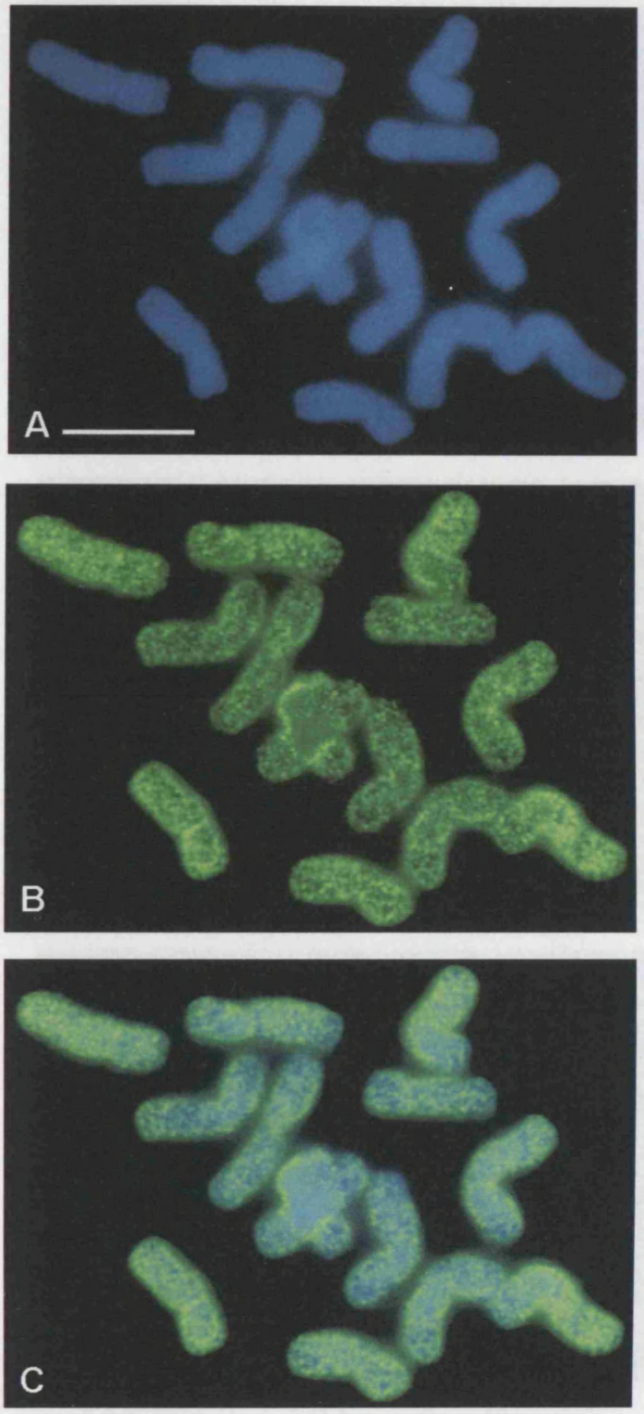


Figure 5. 1. Root tip metaphase of *T. tauschii* (D genome, 2n=14). A) DAPI stained chromosomes. B) Immunostaining with anti-methylcytosine antibody detected with Alexa 488 green fluorescence. C) Overlay of A and B images. Scale bar represents 10 μm.

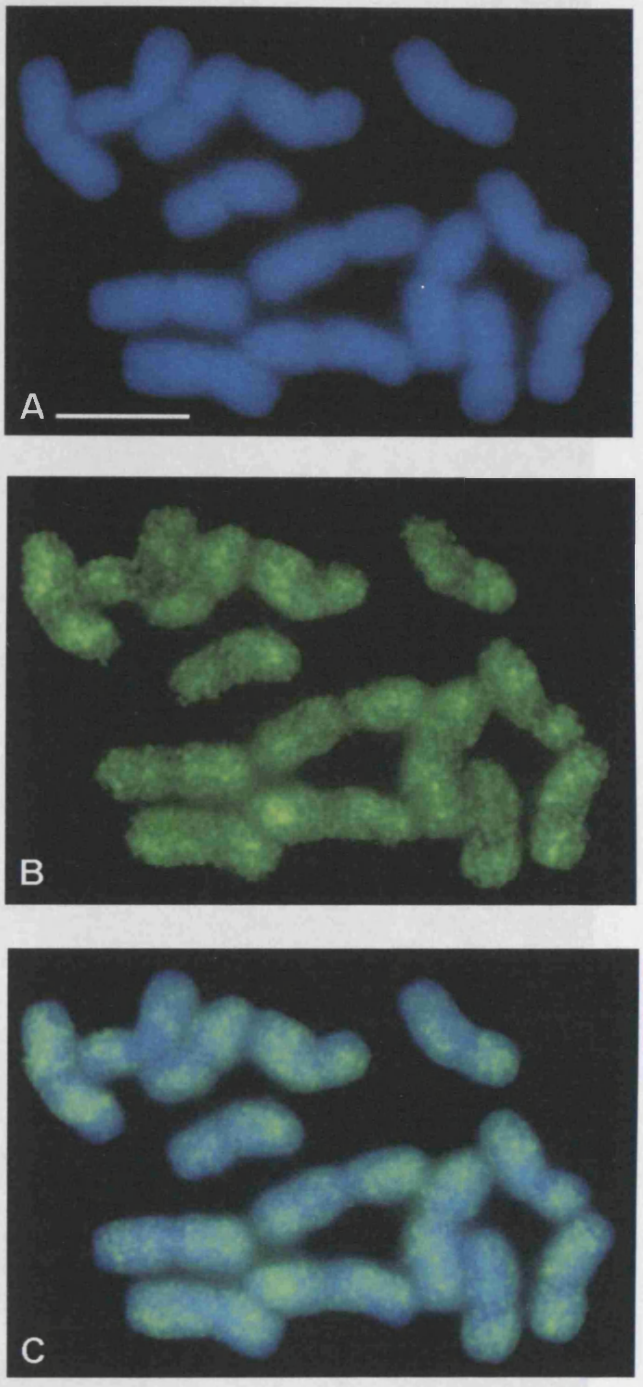


Figure 5. 2. Root tip metaphase of *T. monococcum* (A genome, 2n=14). A) DAPI stained chromosomes. B) Immunostaining with anti-methylcytosine antibody detected with Alexa 488 green fluorescence. C) Overlay of A and B images. Scale bar represents 10μm.

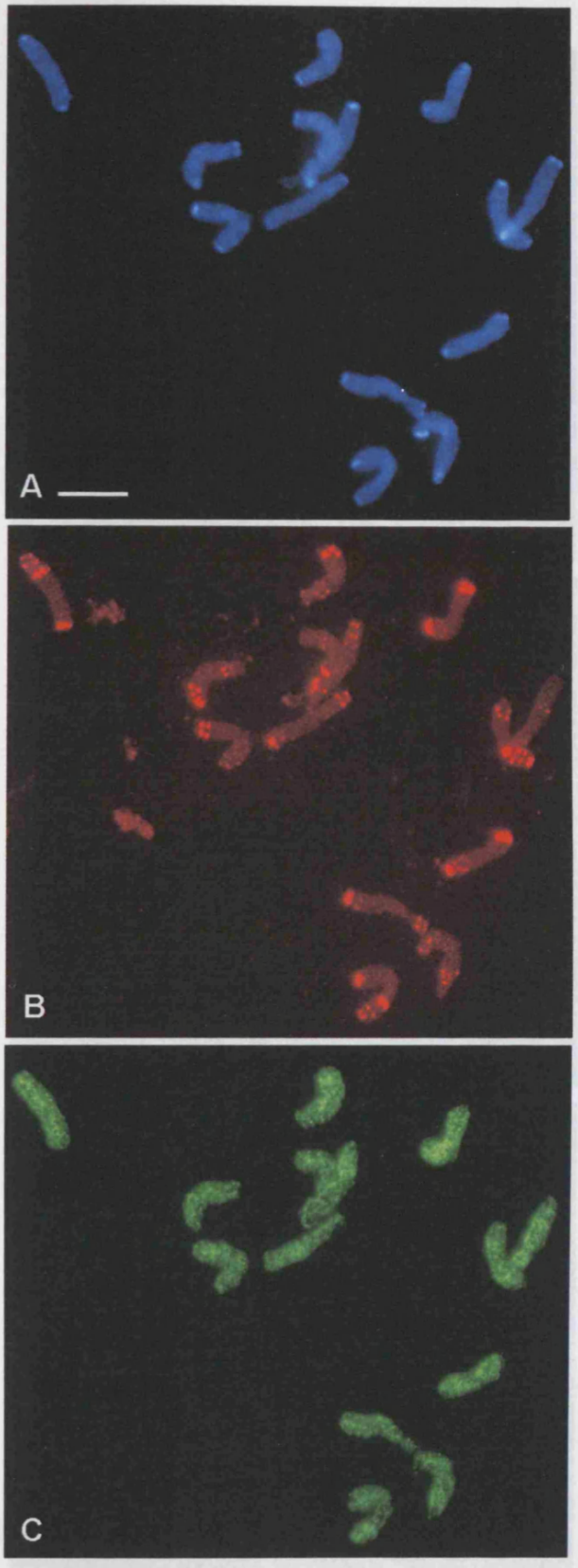


Figure 5. 3. Root tip metaphase of *Ae. speltoides* (B genome, 2n=14). A) DAPI stained chromosomes. B) *In situ* hybridization with biotin-labelled Pet(22594)25/42 !3324 probe detected with Alexa 594 red fluorescence. C) Immunostaining with anti-methylcytosine antibody detected with Alexa 488 green fluorescence. Scale bar represents 10 μm

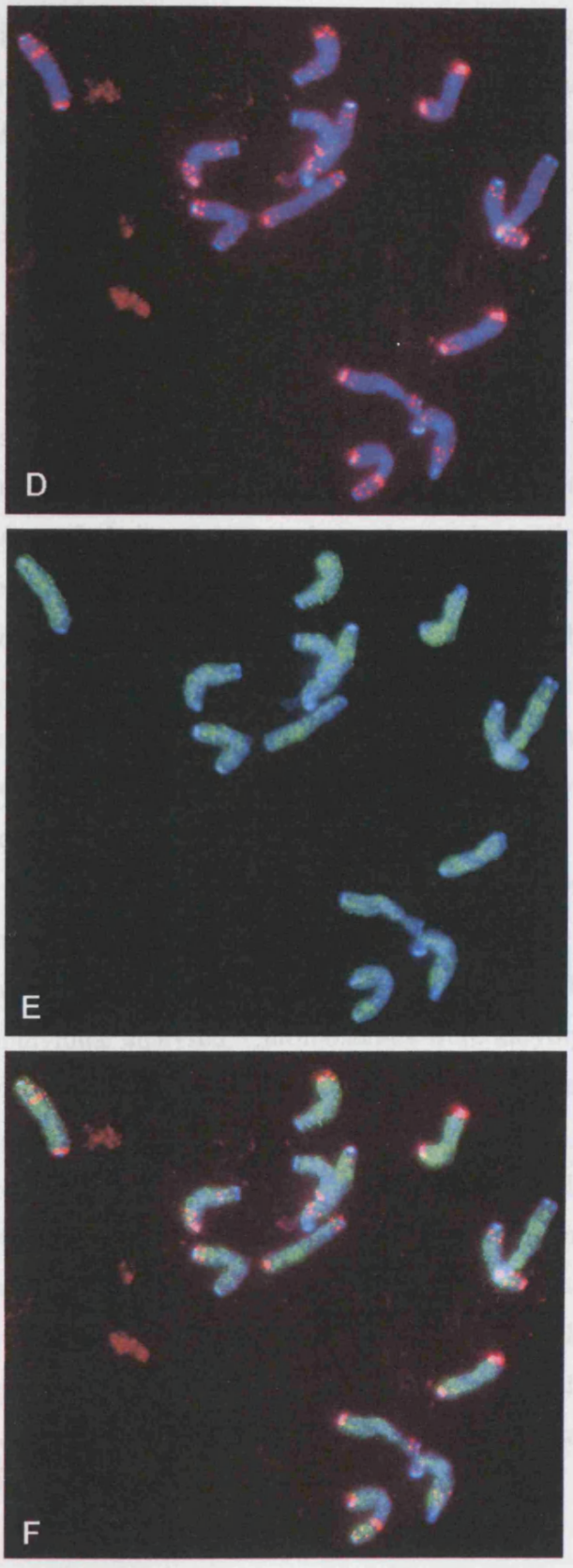


Figure 5. 3 continued. Root tip metaphase of *Ae. speltoides* (B genome, 2n=14). D) Overlay of A and B images. E) Overlay of A and C images. F) Overlay of A, B and C images.

same chromosome pair and differences in the number of methylated telomeres were observed. As shown in figs. 5. 3e-f 120-bp repeat unit sequence family that is mainly present in large blocks at the end of some wheat B genome chromosomes with additional intercalary bands did not appear to be fully methylated in all its sites. Whereas intercalary and small telomeric bands made of 120-bp repeat unit family were fully methylated in all chromosomes, the large sub-telomeric heterochromatic blocks on some chromosome pairs were partially methylated or not methylated.

***Secale cereale* 'Petkus Spring'**

As reported in table 5. 1 for diploid *S. cereale* 'Petkus Spring' (R genome, 2n=14), uniform DNA methylation patterns along the euchromatin of the fourteen chromosomes were observed in the 80% and 100% of analysable metaphases in slides preparations of individuals 3 and 4, conversely individual 2 showed uniformly methylated patterns only in 50% of chromosome spreads. Additional individuals A (fig. 5. 4) and B were analysed and high percentages of metaphases with uniform distribution of methylated cytosines were counted (72.7% and 80%, respectively in table 5. 1). A variable number between 12 and 16 of telomeres (mode=14) did not show methylation signal in metaphases of individuals 2, 3, 4, A and B. Sixteen telomeres showed no signal also in prometaphase where chromosomes are less condensed (see fig. 5. 5). I*n situ* hybridization was carried out on slides from individuals A and B and six to eight metaphases for each individual analysed. Chromosomes were karyotyped according to previous data (Cuadrado and Jouve 1994; Cuadrado, Ceoloni et al. 1995) and the analyses presented in chapter 3 (see karyotypes in figs. 3. 7 and in 3. 8). An example is given in fig. 5. 4d. Intercalary and small telomeric bands made of 120-bp repeat unit family were fully methylated in all chromosomes whereas the telomeric-sites that are part of the large sub-telomeric heterochromatic blocks on chromosome arms 1RS, 1RL, 2RS, 2RL, 3RS, 4RS, 5RS, 6RS were partially methylated or not-methylated (see fig. 5. 4d-e-f). Methylation patterns seemed to be very uniform between cells and individuals of rye and only differences in the number of methylated telomeres were found between metaphases of the same slide or between individuals.

Triticum turgidum* ssp. *durum

After immunostaining with anti-methylcytosine antibody, tetraploid wheat *T. turgidum* sub. *durum* (AB genome, 2n=4x=28) showed differences between

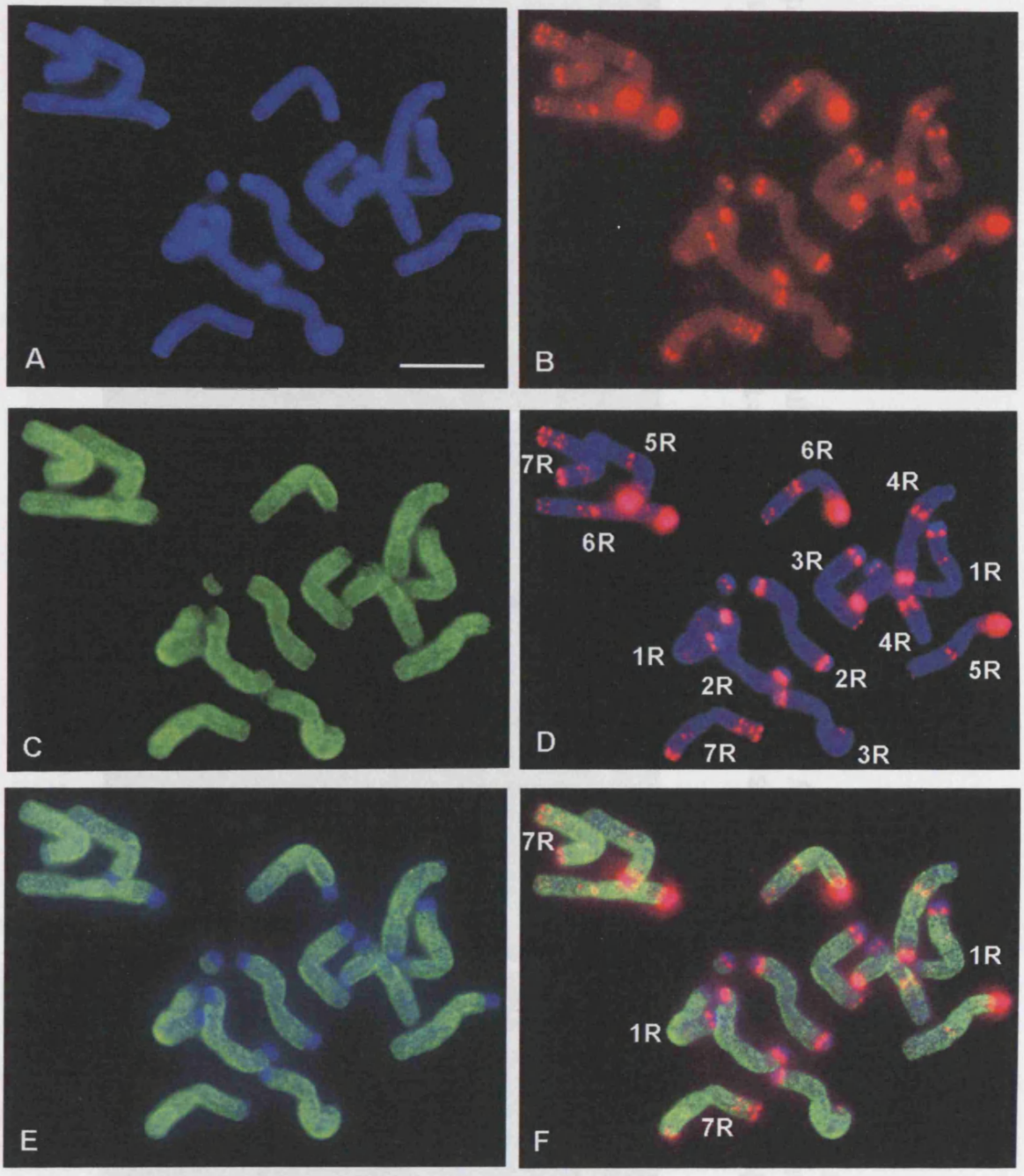


Figure 5. 4. Root tip metaphase of *S. cereale* 'Petkus Spring' (R genome, 2n=14). Chromosomes were identified according to Cuadrado and Jouve (1994) and Cuadrado *et al.* (1995). A) DAPI stained chromosomes. B) *In situ* hybridization with biotin-labelled Pet(22594)25/42!3324 probe detected with Alexa 594 red fluorescence. C) Immunostaining with anti-methylcytosine antibody detected with Alexa 488 green fluorescence. D) Overlay of A and B images. E) Overlay of A and C images. F) Overlay of A, B and C images. Scale bar represents 10μm.

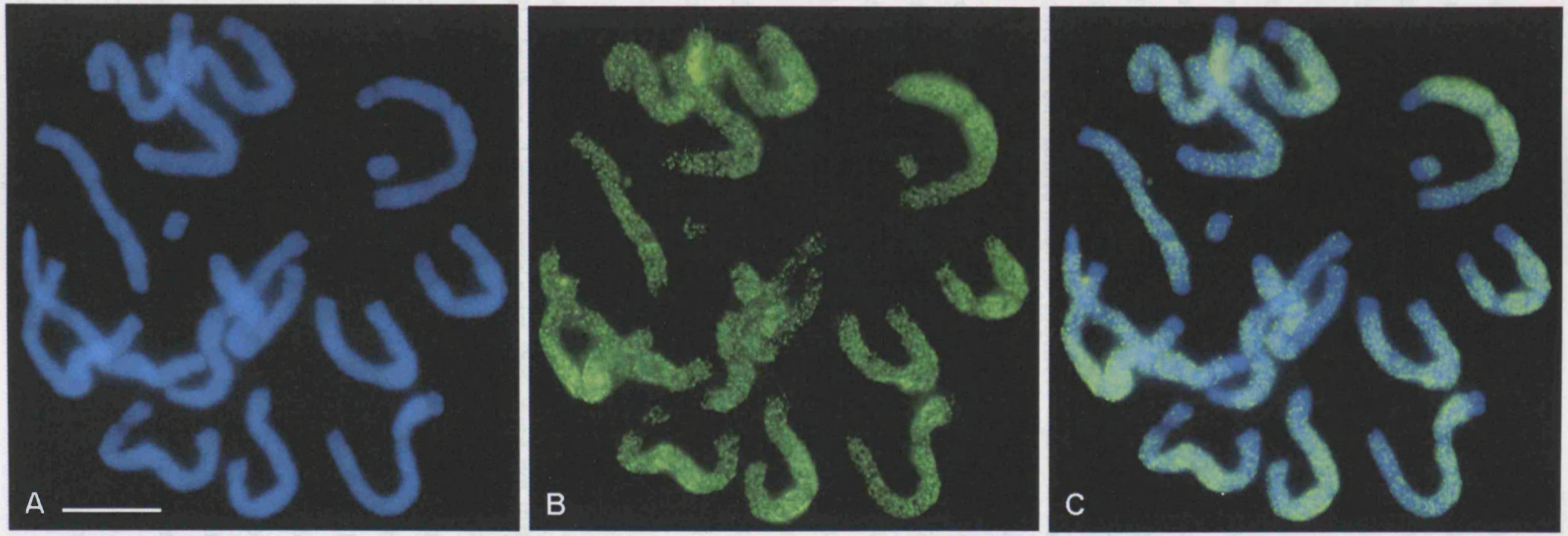


Figure 5. 5. Prometaphase of *S. cereale* 'Petkus Spring' (R genome, 2n=14). A) DAPI stained chromosomes. B) Immunostaining with anti-methylcytosine antibody detected with Alexa 488 green fluorescence. C) Overlay of A and B images. Scale bar represents 10 μm.

chromosomes and chromosome regions. As difficulties were encountered in preparing chromosome spreads, slides preparations of the same individuals previously analysed by Real Time PCR and Southern hybridization were not good enough to be used.

Additional individuals A, B and C were analysed instead by immunostaining with anti-methylcytosine antibody and karyotyped by *in situ* hybridization. In most metaphases of individuals A, B and C (100%, 100% and 61.5% respectively, see table 5. 2) chromosomes showed less uniformly distributed DNA methylation signal than in the majority of metaphases of the diploid donor species *T. monococcum* and *Ae. speltoides* (compare fig. 5. 6 with table 5. 1 and fig. 5. 3). As shown in fig. 5. 6 for individual C, DNA methylation signal along the euchromatin of all chromosomes was patchy: some chromosomes showed enhanced signal at one or two arms, close to the telomeric or centromeric region or at intercalary location, others lacked DNA methylation signal at telomers, centromeres or chromosome arms. After *in situ* hybridization (an example is given in fig. 5. 6d for individual C) chromosomes were karyotyped according to previous analysis (Mukai, Nakara et al. 1993) and data presented in chapter 3 (see figs. 3. 11 and 3.12), and five metaphases for each individual analysed. Karyotyping analysis after immunostaining with anti-methylcytosine antibody were performed in individuals B and C and are summarized in fig. 5. 7. Analysing A genome-origin chromosomes, chromosome pairs 4A and 5A with telomeric 120-bp repetitive sequences at opposite chromosome arms showed enhanced methylation signal at one or both chromosome ends with additional intercalary sites. Other A chromosomes showed unevenly distributed methylation signal but due to the absence of 120-bp repeat unit family in these chromosome pairs they were not distinguished. However, 2 to 4 chromosomes had enhanced signal at one chromosome arm, and 1 to 4 chromosomes at both chromosome arms, and finally 1 to 2 chromosomes appeared heavily methylated. Overall, a variable number of A chromosomes from 4 to a maximum of 8 (more often 4, 5, 6, 8 in similar frequencies) were unevenly or heavily methylated. Analysing B genome-origin chromosomes, enhanced signal was detected at the same chromosome arm with an additional intercalary site in chromosome pairs 1B or 6B (not distinguishable, as both having the same banding pattern after *in situ* hybridization with 120-bp repetitive sequences, see fig. 5. 6f); at one chromosome end with additional subtelomeric and intercalary sites in chromosome pair 2B; at both chromosome ends in chromosome pairs 3B and 4B, despite in some cells enhanced DNA methylation signal was detected also at opposite arms of chromosome pair 4B (see fig. 5. 6f); at one or two

Table 5. 2. DNA methylation patterns in individuals of polyploid *Triticeae* species after immunostaining with anti-methylcytosine antibody.

Species	Individual	number of metaphases photographed	number of metaphases analysable*	number of metaphases (uniform)	number of metaphases (uneven)	% of analysed metaphases (uneven)**
Triticum turgidum ssp. *durum* (AB genome, 2n=4x=28)	TdA	8	4	-	4	100.0%
	TdB	9	7	-	7	100.0%
	TdC	16	13	5	8	61.5%
	TOT	33	24	5	19	79.2%

Species	Individual	number of metaphases photographed	number of metaphases analysable*	number of metaphases (uniform)	number of metaphases (uneven)	% of analysed metaphases (uneven)**
Triticum aestivum 'Chinese Spring' (ABD genome, 2n=6x=42)	Cs3	12	5	-	5	100.0%
	Cs6	7	4	1	3	75.0%
	Cs7	7	3	-	3	100.0%
	TOT	26	12	1	11	91.7%

Species	Individual	number of metaphases photographed	number of metaphases analysable*	number of metaphases (uniform)	number of metaphases (uneven)	% of analysed metaphases (uneven)**
Triticale 'Lamberto' (ABR genome, 2n=6x=42)	La7	2	2	2	-	0.0%
	La8	3	-	-	-	-
	LaA	9	3	-	3	100.0%
	LaB	5	2	-	2	100.0%
	LaC	4	1	-	1	100.0%
	TOT	23	8	2	6	75.0%

Table 5. 2. continued

Species	Individual	number of metaphases photographed	number of metaphases analysable*	number of metaphases (uniform)	number of metaphases (uneven)	% of analysed metaphases (uneven)**
Triticale 'Fidelio' (ABR genome, 2n=6x=42)	Fi c	18	13	1	12	92.3%
	Fi d	8	6	1	5	83.3%
	Fi11	8	2	2	-	0.0%
	Fi A	10	8	-	8	100.0%
	Fi B	12	9	-	9	100.0%
	TOT	56	38	4	34	89.5%

* Cells with incomplete chromosome number, distorsion or staining artefacts were not analysed.

**100% equals to analysed cells.

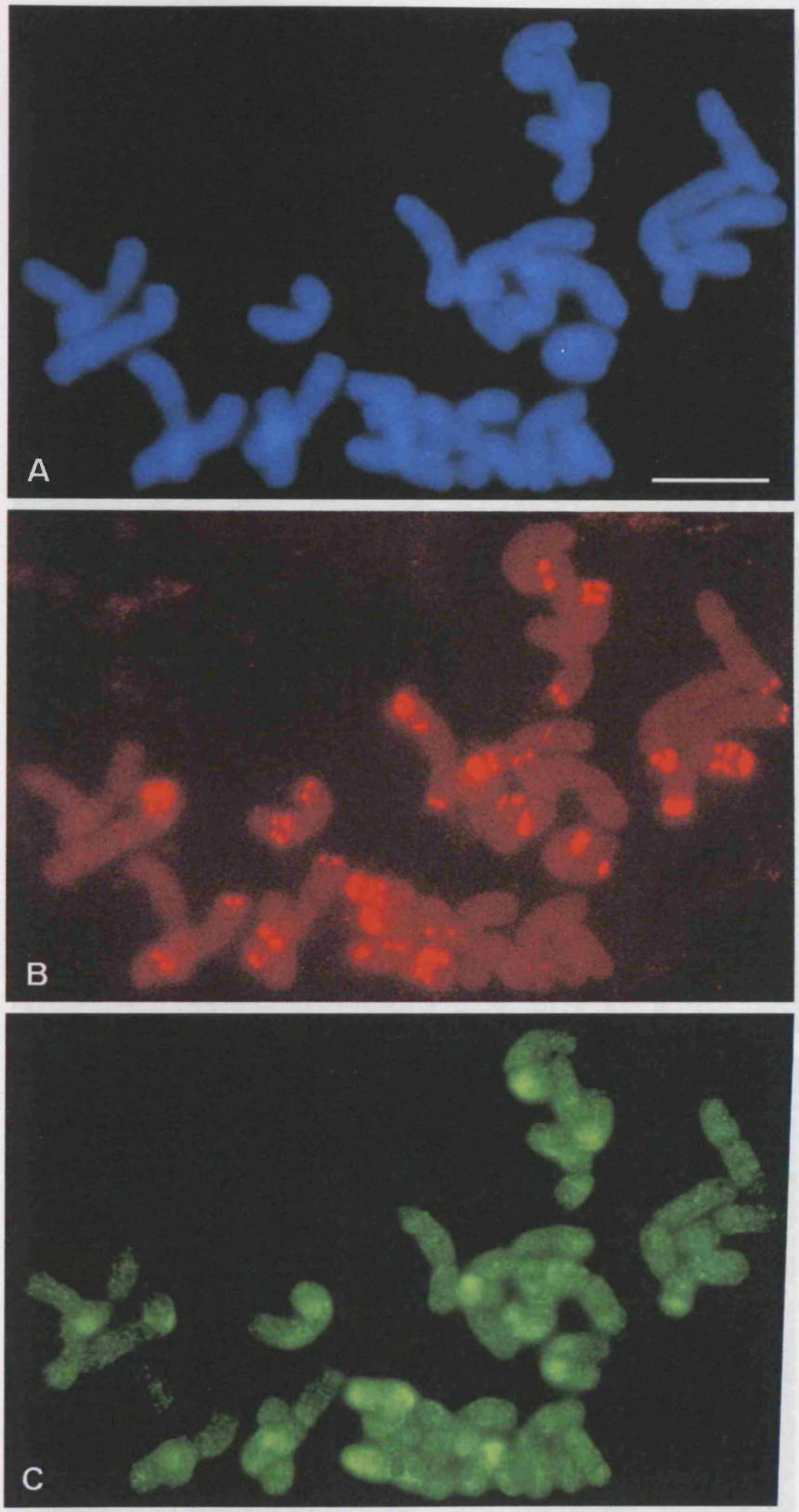


Figure 5. 6. Root tip metaphase of *T. turgidum* ssp. *durum* (AB genome, 2n=28). A) DAPI stained chromosomes. B) *In situ* hybridization with Tmono25/208!1212 probe detected with Alexa 594 red fluorescence. C) Immunostaining with anti-methylcytosine antibody detected with Alexa 488 green fluorescence. Scale bar represents 10 μm.

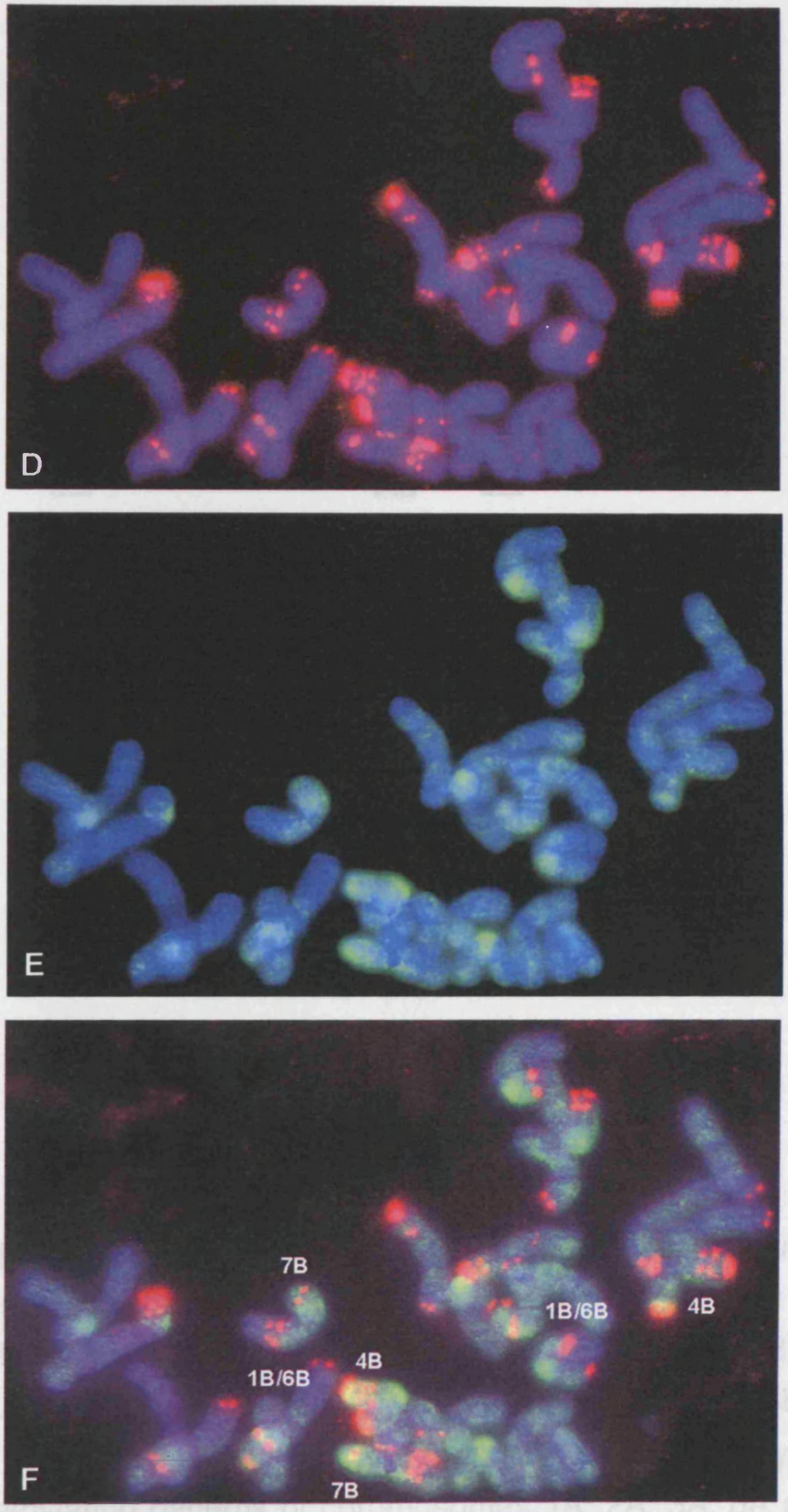


Figure 5. 6 continued. Root tip metaphase of *T. turgidum* ssp. *durum* (AB genome, 2n=28). D) Overlay of A and B images. E) Overlay of A and C images. F) Overlay of A, B and C images.

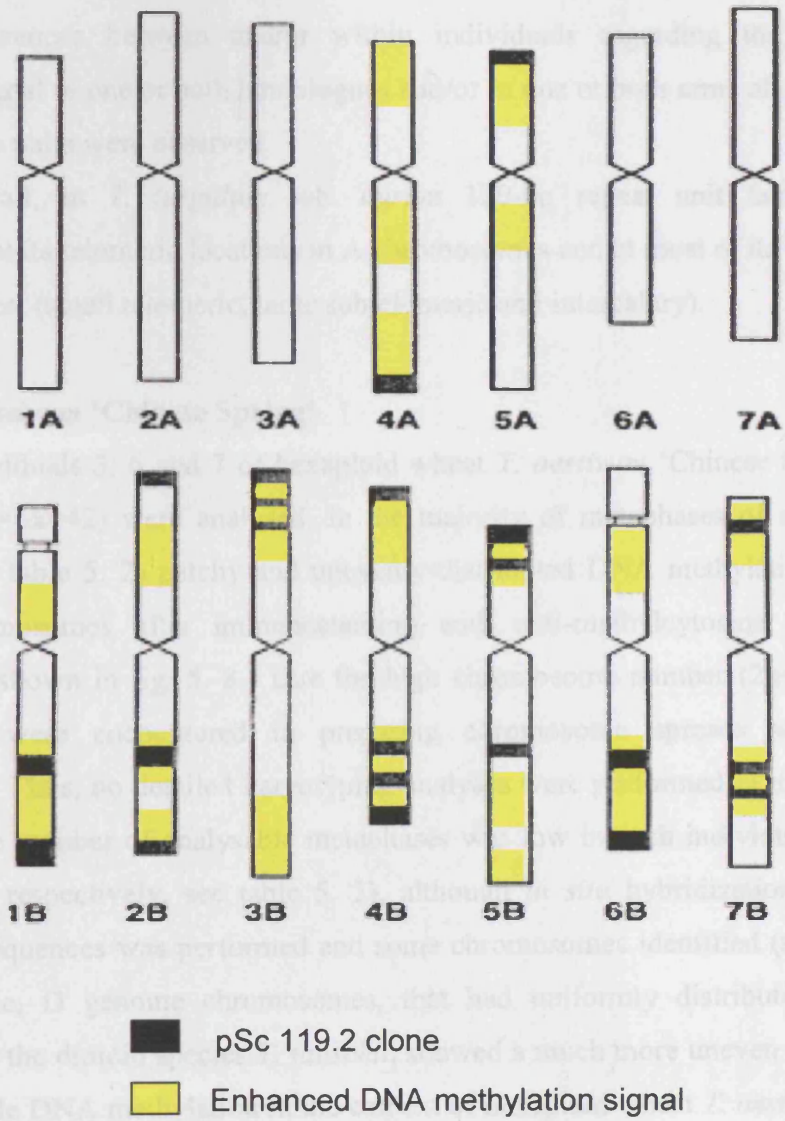


Figure 5. 7. Karyotype of *T. turgidum* ssp. *durum* (AB genome) representing banding patterns obtained after *in situ* hybridization with pSc 119.2 clone (figure modified from Mukai *et al.*, 1993) and DNA methylation patterns obtained by immunostaining with anti-methylcytosine antibody. DNA methylation patterns of chromosomes 1A, 2A, 3A, 6A and 7A are not shown as the chromosomes could not be identified.

chromosome ends with an additional intercalary location in chromosome pair 5B and at one chromosome end with an additional intercalary location in chromosome pair 7B (see fig. 5. 6f).

Differences between and/or within individuals regarding the presence of enhanced signal in one or both homologues and/or in one or both arms of some A and B chromosome pairs were observed.

Overall, in *T. turgidum* sub. *durum* 120-bp repeat unit family appeared methylated at its telomeric locations in A chromosomes and at most of its locations in B chromosomes (small telomeric, large subtelomeric and intercalary).

***Triticum aestivum* 'Chinese Spring'**

Individuals 3, 6 and 7 of hexaploid wheat *T. aestivum* 'Chinese Spring' (ABD genome, 2n=6x=42) were analysed. In the majority of metaphases of all individuals (91.7%, see table 5. 2) patchy and unevenly distributed DNA methylation was found along chromosomes after immunostaining with anti-methylcytosine antibody. An example is shown in fig. 5. 8. Due the high chromosome number (2n=6x=42) more difficulties were encountered in preparing chromosome spreads with complete metaphases. Thus, no detailed karyotyping analyses were performed in individuals 3, 6 and 7 as the number of analysable metaphases was low in each individual (5, 4 and 3 metaphases respectively, see table 5. 2), although *in situ* hybridization with 120-bp repetitive sequences was performed and some chromosomes identified (see fig. 5. 8d). For example, D genome chromosomes, that had uniformly distributed methylated cytosines in the diploid species *T. tauschii*, showed a much more uneven distribution of genome-wide DNA methylation in the context of hexaploid wheat *T. aestivum* 'Chinese Spring', presenting chromosomal regions with enhanced and reduced DNA methylation signal (compare fig. 5. 8 and fig. 5. 1). In chromosome pair 5D (having the same banding pattern of chromosome pair 5A: one telomeric band at the short arm) as well as in chromosome pairs 2D, 3D and 4D (all with similar arm length and/or arm ratio so it was not possible to distinguish them in detail) D genome-origin 120-bp repetitive sequences seemed to be fully methylated in the majority of their telomeric sites (see fig. 5. 8f). As chromosome ends were methylated also in the diploid donor species *T. tauschii* (see fig. 5. 1) and thus presumably the 120-bp repetitive sequences present at these locations, no major changes in DNA methylation affected 120-bp repeat unit

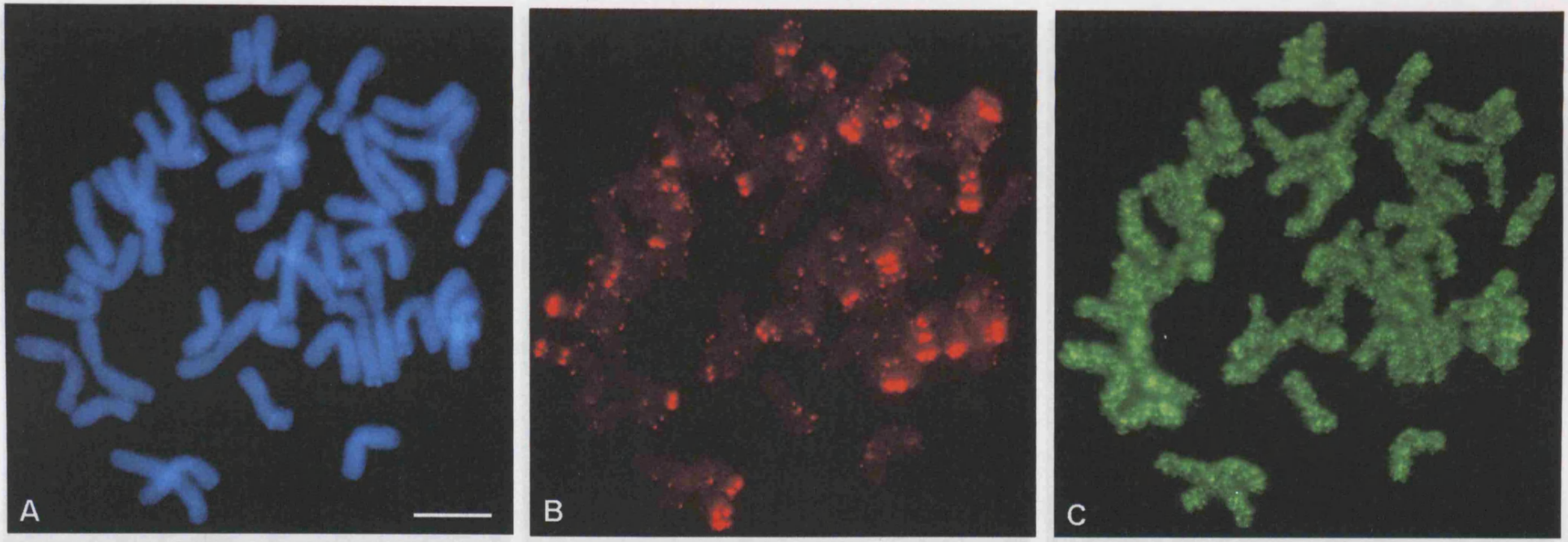


Figure 5. 8. Root tip metaphase of *T. aestivum* 'Chinese Spring' (ABD genome, 2n=42). A) DAPI stained chromosomes. B) *In situ* hybridization with Svav25/208!182 probe detected with Alexa 594 red fluorescence. C) Immunostaining with anti-methylcytosine antibody detected with Alexa 488 green fluorescence. Scale bar represents 10 μm.

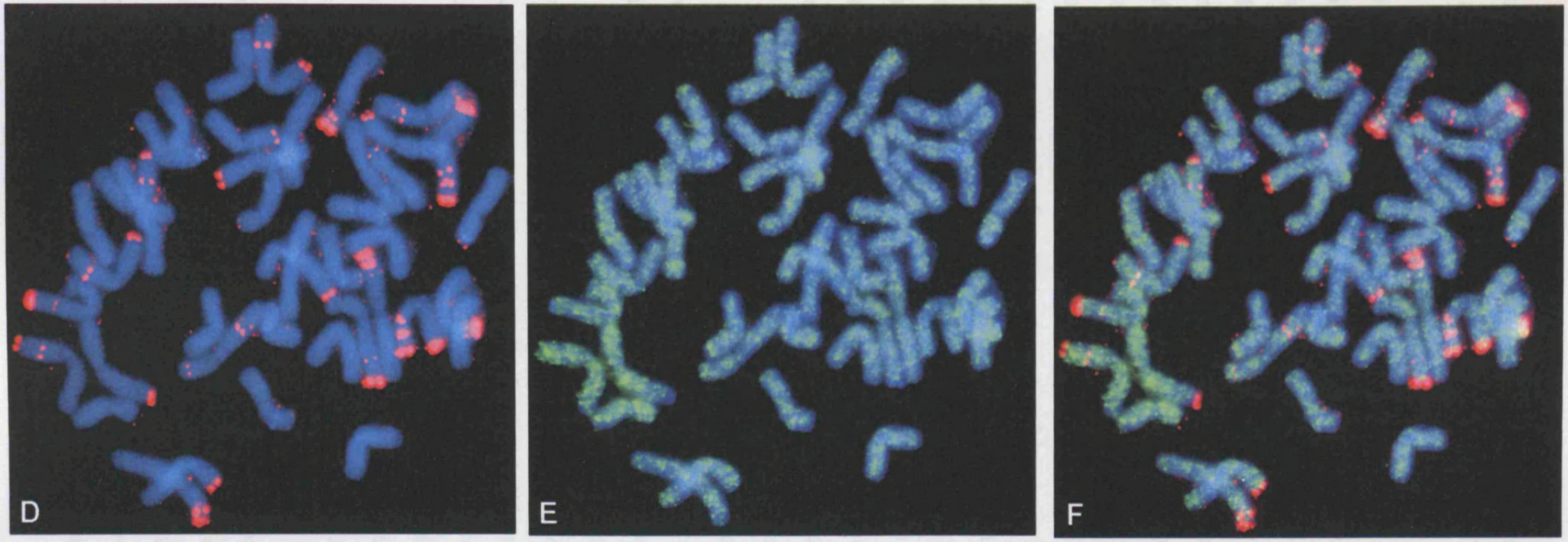


Figure 5. 8 continued. Root tip metaphase of *T. aestivum* 'Chinese Spring' (ABD genome, 2n=42). D) Overlay A and B images. E) Overlay A and C images. F) Overlay A, B and C images.

family sequences at telomeric locations of D chromosomes as a consequence of the polyploidization event originating *T. aestivum* 'Chinese Spring'.

***Triticale* 'Fidelio' and 'Lamberto'**

Individuals from *Triticale* 'Fidelio' and 'Lamberto' (ABR genome, 2n=6x=42) were analysed. Chromosome spreads were made for *Triticale* 'Lamberto' individuals 7 and 8. A quite uniform DNA methylation signal but in only two complete or nearly complete metaphases for individual 7, while no analysable metaphases for individual 8 were found (see table 5. 2). Slides from three additional individuals were prepared showing a differentiated DNA methylation signal between chromosomes in the majority of analysable metaphases (6 out of 8 metaphases from all individuals, see table 5. 2). However, due to technical difficulties encountered in removing cytoplasm, spreading 'sticky' chromosomes, preserving chromosome morphology during the staining and due to low metaphase index no further analyses were performed in *Triticale* 'Lamberto'.
After immunostaining with anti-methylcytosine antibody in *Triticale* 'Fidelio' individuals C and D patchy and unevenly distributed DNA methylation signal between chromosomes and/or chromosomal regions was found in most metaphases (92.3% and 83.3% of metaphases respectively, see table 5. 2) with stronger intensity at some telomeres or along one or two chromosome arms in some chromosomes while other chromosomes were lacking DNA methylation signal at centromeres, as previously observed in *T. turgidum* sub. *durum*. Subsequent *in situ* hybridization with 120-bp repetitive sequences distinguished all forty-two chromosomes in *Triticale* 'Fidelio' individual C, but not in individual D. Furthermore, in *Triticale* 'Fidelio' individual 11 a low number of metaphases and weak antibody signal did not allow further analyses. Two additional *Triticale* 'Fidelio' individuals, A and B, gave good slide preparations and were then analysed with immunostaining with anti-methylcytosine antibody. In both individuals DNA methylation signal was unevenly distributed in all metaphases analysable (100% of chromosome spreads, see table 5. 2). Overall, *Triticale* 'Fidelio' differentiated DNA methylation patterns were in agreement with unevenly distributed DNA methylation patterns reported on chromosome spreads of *Triticale* 'Lasko' (Castilho, Neves et al. 1999). Successful *in situ* hybridization with 120-bp repetitive sequences allowed complete karyotyping analyses for individuals A, B (shown in fig. 5. 9) and C. *Triticale* 'Fidelio' individuals A, B and C were karyotyped according to wheat and rye karyotypes (Mukai, Nakara et al. 1993; Cuadrado and Jouve 1994; Cuadrado

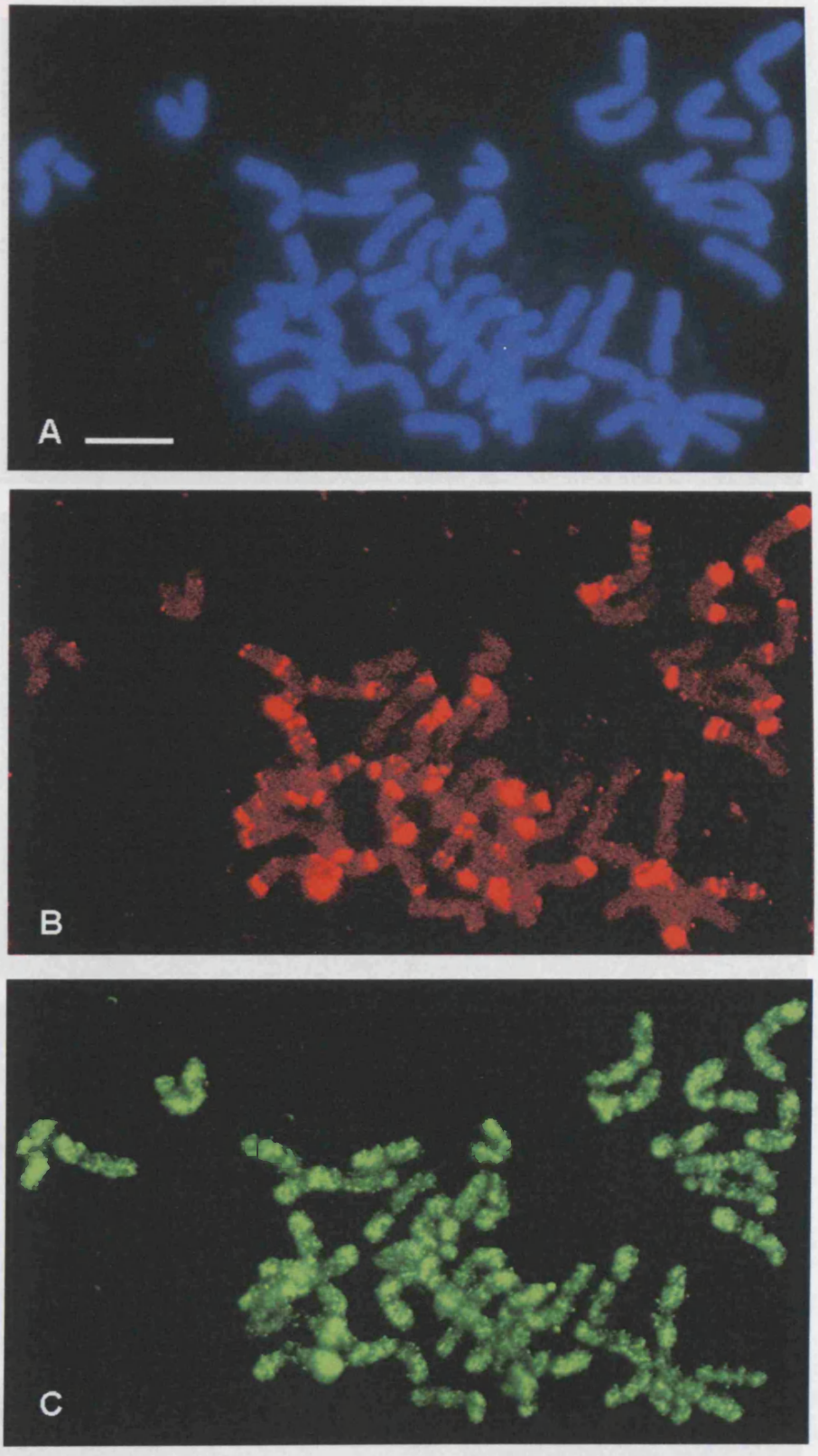


Figure 5. 9. Root tip metaphase of *Triticale* ' Fidelio' (ABR genome, 2n=42). A) DAPI stained chromosomes. B) *In situ* hybridization with Pet(22594)25/42!3324 probe detected with Alexa 594 red fluorescence. C) Immunostaining with anti-methyl cytosine antibody detected with Alexa 488 green fluorescence. Scale bar represents 10 μm.

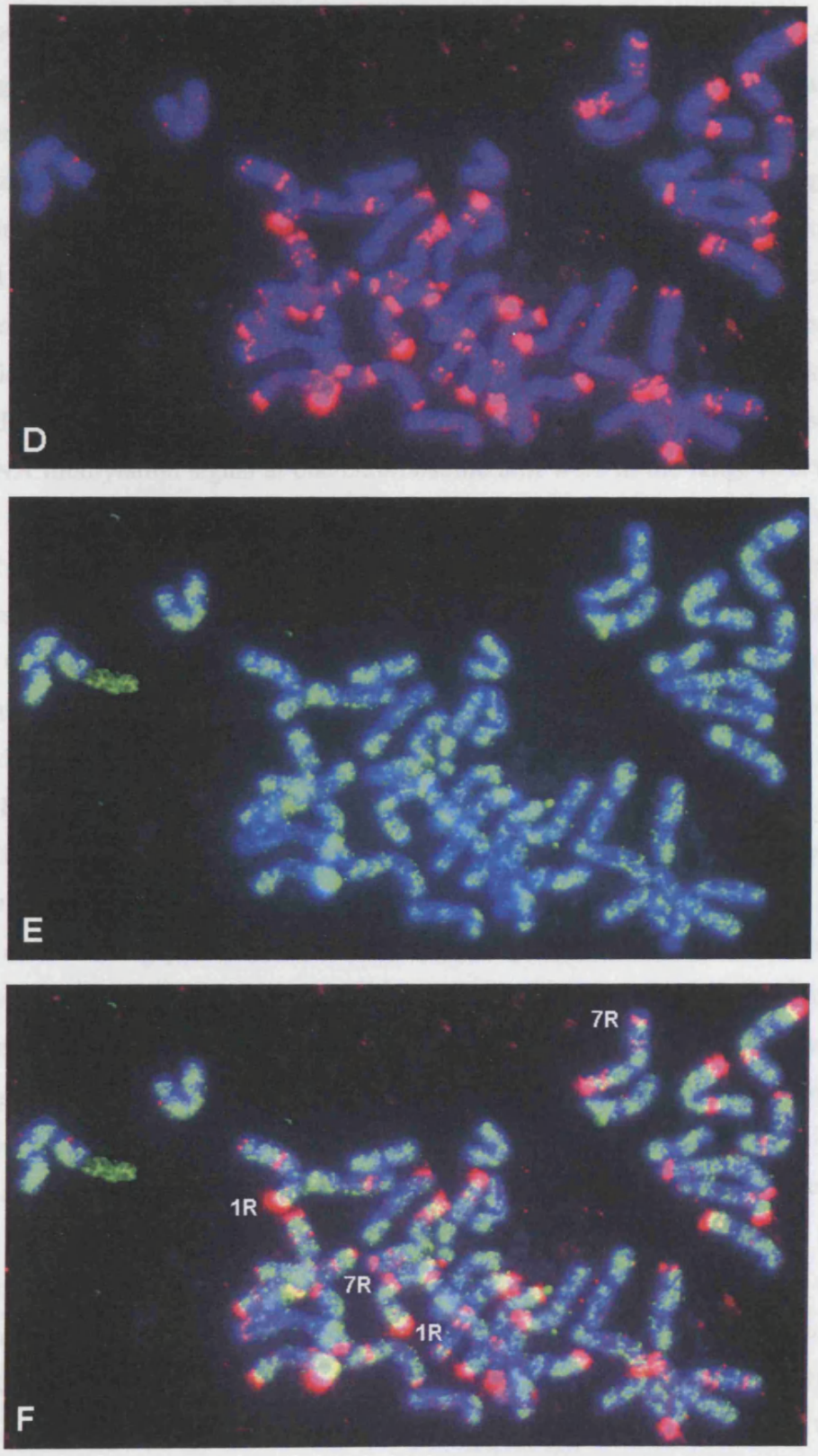


Figure 5. 9 continued. Root tip metaphase of *Triticale* 'Fidelio' (ABR genome, 2n=42). D) Overlay A and B images. E) Overlay A and C images. F) Overlay A, B and C images.

and Jouve 1995) and previous analyses in chapter 3 (see karyotypes in figs. 3. 7 and in 3. 8, and figs. 3. 11 and 3. 12); three to six metaphases for each individual were analysed. Karyotyping analysis after immunostaining with anti-methylcytosine antibody are summarized in fig. 5. 10 and compared to *T. turgidum* sub. *durum* karyotype (fig. 5. 7). Comparing DNA methylation patterns of A-origin chromosomes between *Triticale* 'Fidelio' and *T. turgidum* sub. *durum* (compare figs. 5. 10 and 5. 7), chromosome pairs 4A and 5A maintained the same DNA methylation patterns with additional enhanced DNA methylation signal at the telomeric region of chromosome arm 5AL in *Triticale* 'Fidelio' (indicated by an asterisk in fig. 5. 10). In *Triticale* 'Fidelio' chromosomes with enhanced DNA methylation signal at one chromosome arm were in the range of 2 to 5 similarly to *T. turgidum* sub. *durum*, but more chromosomes (range 3 to 7) showed enhanced DNA methylation signal at both chromosome arms and a higher number of chromosomes (range 2 to 4) appeared fully methylated in comparison to *T. turgidum* sub. *durum*. Overall, a variable number of A chromosomes (from 8 to a maximum of 11) higher than in *T. turgidum* sub. *durum* were unevenly or fully methylated in *Triticale* 'Fidelio'. Generally, a more uneven distribution of DNA methylation was also observed in B-origin chromosomes of *Triticale* 'Fidelio' compared to *T. turgidum* sub. *durum*: additional enhanced DNA methylation signal at intercalary location in chromosome pairs 2B, 5B and 7B (indicated by asterisks in fig. 5. 10) and at two satellites of chromosome pairs 1B and 6B (indicated by asterisks in fig. 5. 10) were observed, while only chromosome pairs 3B and 4B retained the distribution of methylated cytosines previously observed in tetraploid wheat *T. turgidum* sub. *durum* (compare figs. 5. 10 and 5. 7). Finally, in *Triticale* 'Fidelio' R chromosomes showed a more differentiated DNA methylation signal than in diploid *S. cereale* 'Petkus Spring'. In addition, a variable number of rye telomeres from 8 to a maximum of 12 (mode=8), lower than in diploid *S. cereale* 'Petkus Spring', were methylation free. In fact, differently to diploid *S. cereale* 'Petkus Spring', in chromosome pairs 1R and 2R only one telomere instead of two were methylation free and telomeres from chromosome pairs 3R, 4R and 5R were often fully methylated. Furthermore, in about half of overall *Triticale* 'Fidelio' metaphases (5 out 12) chromosome pairs 1R and 7R changed their methylation patterns compared to diploid *S. cereale* 'Petkus Spring' (see fig. 5. 4), the first-one becoming methylated at one or both chromosome ends and the latter lacking DNA methylation at telomere of chromosome arm 7RL (see fig. 5. 9f).

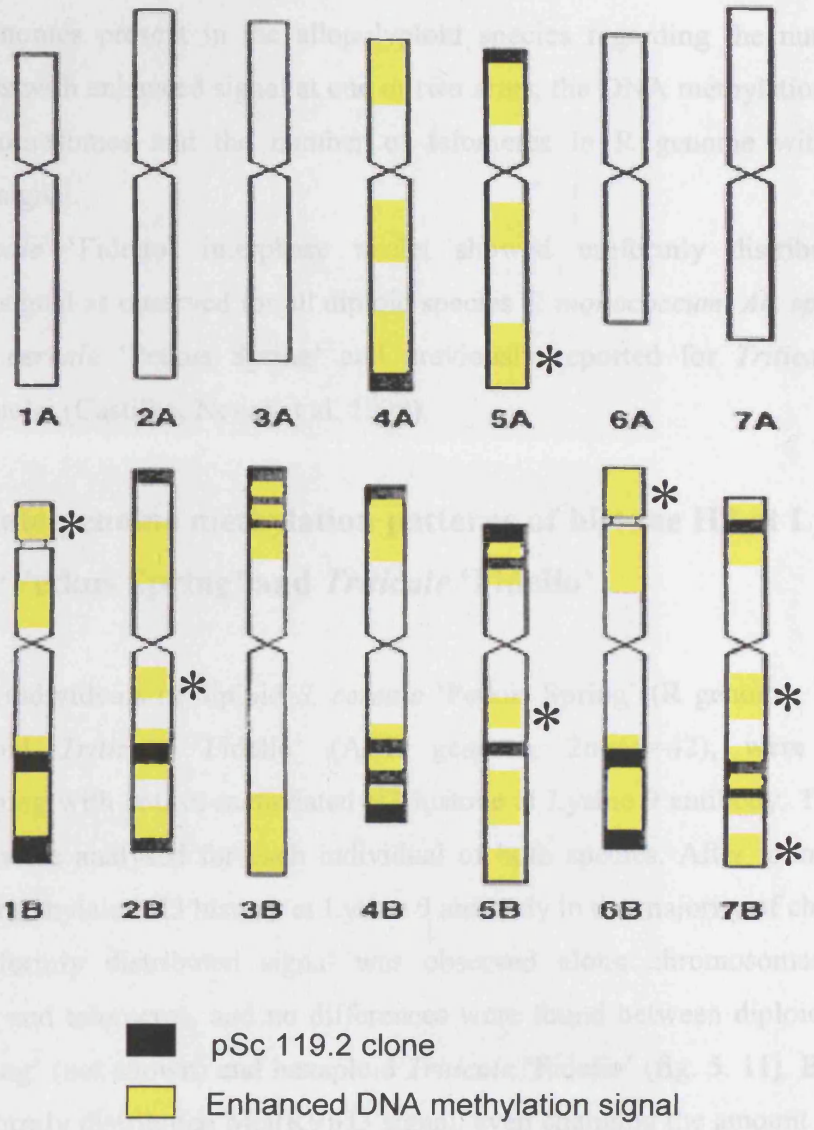


Figure 5. 10. Karyotype of A and B chromosomes of *Triticale* 'Fidelio' (ABR genome) representing banding patterns obtained after *in situ* hybridization with pSc 119.2 clone (figure modified from Mukai *et al.*, 1993) and DNA methylation patterns obtained by immunostaining with anti-methylcytosine antibody. DNA methylation patterns of chromosomes 1A, 2A, 3A, 6A and 7A are not shown as the chromosomes could not be identified.

Differences of DNA methylation patterns were found between and within *Triticale* 'Fidelio' individuals as well as between homologous chromosomes in each of the three genomes present in the allopolyploid species regarding the number of A chromosomes with enhanced signal at one or two arms, the DNA methylation pattern at most B chromosomes and the number of telomeres in R genome without DNA methylation signal.

Triticale 'Fidelio' interphase nuclei showed uniformly distributed DNA methylation signal as observed for all diploid species *T. monococcum*, *Ae. speltoides*, *T. tauschii*, *S. cereale* 'Petkus Spring' and previously reported for *Triticale* 'Lasko' interphase nuclei (Castilho, Neves et al. 1999).

5. 4. 2 Whole-genome methylation patterns of histone H3 at Lysine 9 in *S. cereale* 'Petkus Spring' and *Triticale* 'Fidelio'

Five individuals of diploid *S. cereale* 'Petkus Spring' (R genome, 2n=14) and the hexaploid *Triticale* 'Fidelio' (ABR genome, 2n=6x=42), were tested by immunostaining with anti di-methylated H3 histone at Lysine 9 antibody. Two to three metaphases were analysed for each individual of both species. After immunostaining with anti di-methylated H3 histone at Lysine 9 antibody in the majority of chromosomes spreads uniformly distributed signal was observed along chromosomes including centromeres and telomeres, and no differences were found between diploid *S. cereale* 'Petkus Spring' (not shown) and hexaploid *Triticale* 'Fidelio' (fig. 5. 11). Both species display uniformly distributed Met(K9)H3 signal, even changing the amount of blocking BSA (1, 2 or 4%) and secondary antibody Alexa 488 (dilutions 1:200, 1:400 or 1:600).

5. 5 Discussion

5. 5. 1 Whole-genome DNA methylation and methylation patterns of 120-bp repeat unit family in diploid *Triticeae* species

The majority of metaphases of diploid species *T. monococcum* (2n=14, A genome), *Ae. speltoides* (2n=14, B genome), *T. tauschii* (2n=14, D genome) and *S.*

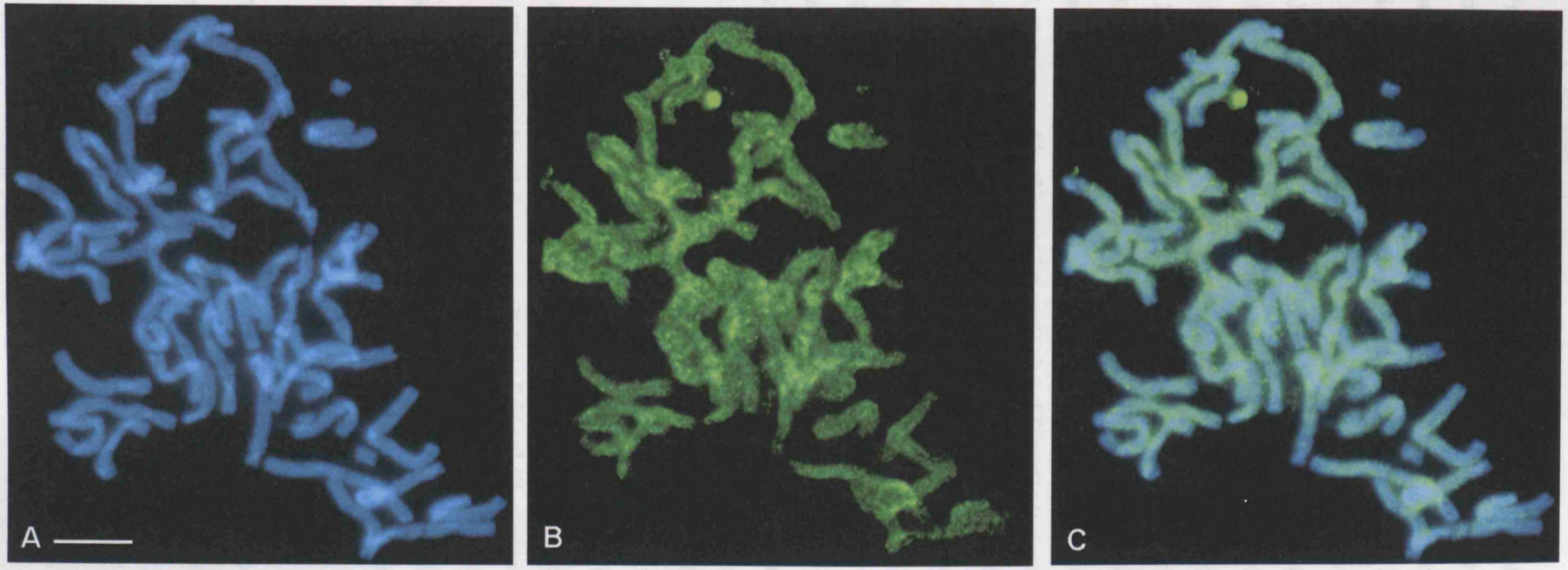


Figure 5. 11. Prometaphase of *Triticale* 'Fidelio' (ABR genome, 2n=42). A) DAPI stained chromosomes. B) Immuno-staining with anti-di-methylated H3 histone at Lysine 9 antibody detected with Alexa 488 green fluorescence. C) Overlay of A and B images. Scale bar represents 10 μm.

cereale 'Petkus Spring' (2n=14, R genome) analysed by immunostaining with anti-methylcytosine antibody showed uniformly distributed whole-genome cytosine methylation along the euchromatin of all fourteen chromosomes including the centromeric regions with no differences between chromosomes detectable (table 5. 1). In A and D genomes DNA methylation extend till all chromosome ends with fully methylated telomeres (figs. 5. 1 and 5. 2). Conversely, a variable number of telomeres, between 10 and 20 (mode=16) in *Ae. speltoides* (B genome) and between 12 and 16 (mode=14) in *S. cereale* 'Petkus Spring' (R genome), were methylation free (figs. 5. 3, 5. 4 and 5. 5). Often the signal was absent at one or both telomeres of the same chromosome pair and differences in the number of methylated telomeres were observed, possibly as a result of technical reasons (presence of cytoplasm, chromosomes not totally flat or enhanced background signal) or a different accessibility of the antibody to DNA at subtelomeric heterochromatin (Frediani, Giraldi et al. 1996). However, it cannot be excluded that methylation patterns at telomeric regions are highly variable in both species.

120-bp repeat unit sequence family, that is located in large blocks at the end of most *S. cereale* 'Petkus Spring' (R genome) chromosomes and some *Ae. speltoides* (B genome) chromosomes whereas present in lower copy number in some chromosome ends of *T. tauschii* (D genome) and only in few copies in *T. monococcum* (A genome), showed a different DNA methylation status according to the different diploid species analysed. Where present only at telomeric regions, such as in *T. monococcum* (A genome) and *T. tauschii* (D genome), the 120-bp repeat unit sequence family was presumably uniformly and fully methylated at its locations (figs. 5. 1 and 5. 2). In *Ae. speltoides* (B genome) and *S. cereale* 'Petkus Spring' (R genome) intercalary and small telomeric bands made of 120-bp repeat unit family were fully methylated in all chromosomes, whereas telomeric-sites that are part of large sub-telomeric heterochromatic blocks were partially methylated or not-methylated (figs. 5. 3, 5. 4 and 5. 5).

5. 5. 2 Whole-genome DNA methylation and methylation patterns of 120-bp repeat unit family in polyploid *Triticeae* species

Overall, the results obtained *T. turgidum* sub. *durum* (AB genome, 2n=4x=28), *T. aestivum* 'Chinese Spring' (ABD genome, 2n=6x=42) and *Triticale* 'Fidelio' (ABR

genome, 2n=6x=42) showed that after immunostaining with anti-methylcytosine antibody all allopolyploid species analysed in the present work had more unevenly distributed whole-genome DNA methylation patterns on chromosome spreads (table 5. 2) if compared with the uniform DNA methylation signal along chromosomes found in the majority of metaphases in diploid species *T. monococcum* (2n=14, A genome), *Ae. speltoides* (2n=14, B genome), *T. tauschii* (2n=14, D genome) and *S. cereale* 'Petkus Spring' (2n=14, R genome). The differentiated DNA methylation patterns observed in *Triticale* 'Fidelio' were in agreement with the observations made in *Triticale* 'Lasko' (Castilho, Neves et al. 1999), while the results obtained in *T. turgidum* sub. *durum* and *T. aestivum* 'Chinese Spring' extended the survey of whole-genome DNA methylation patterns to other natural allopolyploid *Triticeae* species.

Furthermore, when diploid A and B genomes contributed to tetraploid wheat *T. turgidum* sub. *durum* (AB genome, 2n=4x=28), as well as D genome to hexaploid wheat *T. aestivum* 'Chinese Spring' (ABD genome, 2n=6x=42) or R genome to hexaploid *Triticale* 'Fidelio' (ABR genome, 2n=6x=42), methylated cytosines showed enhanced and reduced DNA methylation signal in different chromosomes pairs and/or different chromosome regions (figs. 5. 6, 5. 8 and 5. 9). Previous data on *A. cepa* and *V. faba*, two species with large genomes, also found unevenly distributed cytosine methylation along chromosomes after immunostaining with anti-methylcytosine antibody (Castiglione, Giraldi et al. 1995; Frediani, Giraldi et al. 1996). Conversely, in *M. album* a species with a small genome (1C=2.9 pg), unevenly distributed DNA methylation signal was reported only for sex chromosomes in male and female cells, but no prominent differences for autosomes in both male and female metaphases (Siroky, Castiglione et al. 1998). In conclusion, the results present in this chapter together with previous data about whole-genome DNA methylation in species with large and small genome sizes (*A. cepa*, *V. faba*, and *M. album* respectively), suggest that diploid *Triticeae* species with relatively smaller genomes have uniform DNA methylation patterns and if contributing to the larger genomes of polyploid *Triticeae* species inevitably undergo DNA methylation changes resulting in differentiated DNA methylation patterns.

Detailed karyotyping analyses by using *in situ* hybridization with 120-bp repetitive sequences were performed in *T. turgidum* sub. *durum* and *Triticale* 'Fidelio'. In *T. turgidum* sub. *durum* (AB genome, 2n=4x=28) a variable number of A chromosomes showed unevenly distributed whole-genome DNA methylation patterns at

one or both chromosome arms, while B chromosomes showed reproducible differentiated whole-genome DNA methylation patterns (fig. 5. 6) compared to diploid species *T. monococcum* (2n=14, A genome) and *Ae. speltoides* (2n=14, B genome). Overall, it seemed that A genome-origin 120-bp repeat unit family sequences retained their fully DNA methylated status in *T. turgidum* sub. *durum*. Despite *Ae. speltoides* is thought to be the diploid *Aegilops* species most similar to B genome donors of cultivated and wild allotetraploid wheats (Badaeva, Friebe et al. 1996; Belyayev, Raskina et al. 2000), B genome-origin 120-bp repetitive sequences possibly underwent some chromosomal rearrangements as the B genome-karyotype in *T. turgidum* sub. *durum* differed from the one of *Ae. speltoides* after *in situ* hybridization with 120-bp repetitive sequences. As a consequence, DNA methylation changes possibly affected B genome-origin 120-bp repetitive sequences at telomeric, subtelomeric and/or intercalary locations in the new context of allotetraploid wheat *T. turgidum* sub. *durum*.

In *Triticale* 'Fidelio' (ABR genome, 2n=6x=42) both A and B genomes maintained unevenly distributed whole-genome DNA methylation patterns (fig. 5. 9) as already observed in *T. turgidum* sub. *durum* (2n=4x=28, AB genome) with a higher number of A chromosomes showing enhanced and reduced DNA methylation signal, and additional intercalary sites and satellites being methylated in most B chromosomes. However, no major changes in 120-bp repeat unit family DNA methylation levels seemed to happen at telomeric and intercalary sites of B-origin chromosomes and at telomeres of chromosome pairs 4A and 5A in *Triticale* 'Fidelio' if compared to their DNA methylation status in *T. turgidum* sub. *durum*. In addition, in the allohexaploid context of *Triticale* 'Fidelio' R-origin chromosomes had differentiated DNA methylation signal compared to diploid *S. cereale* 'Petkus Spring' (2n=14, R genome). In *Triticale* 'Fidelio' more telomeric regions were methylated (2R, 3R, 4R and 5R) in comparison to *S. cereale* 'Petkus Spring' and thus presumably their 120-bp repetitive sequences present in large blocks at subtelomeric location. Furthermore, in *Triticale* 'Fidelio' chromosomes pairs 1R and 7R changed DNA methylation patterns at telomeres in comparison to *S. cereale* 'Petkus Spring'. As a result, 120-bp repeat unit family sequences underwent some DNA methylation changes at the ends of rye chromosome pairs 1R and 7R: the large sub-telomeric heterochromatic blocks on chromosome arms 7RL being not methylated and the ones of 1RS and/or 1RL being methylated.

Minor differences in DNA methylation patterns between and/or within individuals were observed at only one or both homologues and/or arms of A and B chromosomes in *T. turgidum* sub. *durum*; and between and within individuals as well as between homologues for each of the three genomes present in *Triticale* 'Fidelio'. Minor differences could be due to technical reasons (presence of cytoplasm, chromosomes not totally flat or enhanced background signal) or truly representing the variability of methylation patterns in both allopolyplod species. As already reported in *V. faba* after immunostaining with anti-methylcytosine antibody (Frediani, Giraldi et al. 1996) differences between corresponding regions of homologues could be due to differential DNA methylation of the same region in a chromosome pair or to a different accessibility of the antibody to DNA, maybe as a consequence of local differences in chromatin organization due to histone modifications or interaction with proteins, for example.

5. 5. 3 Whole-genome methylation patterns of histone H3 at Lysine 9 in *S. cereale* 'Petkus Spring' and *Triticale* 'Fidelio'

Ploidy level does not affect histone H3 methylation patterns at Lysine 9 as both diploid *S. cereale* 'Petkus Spring' (not shown) and hexaploid *Triticale* 'Fidelio' (fig. 5. 11) showed similar uniformly distributed signal on their interphase nuclei and along their chromosomes till chromosome ends after immunostaining with anti di-methylated H3 histone at Lysine 9 antibody. These results confirmed previous data reported by Houben *et al.* (2003) for diploid species, such as *S. cereale* and *H. vulgare*, and species with a high nuclear DNA content, such as *V. faba* or hexaploid *T. aestivum*. Conversely to DNA methylation patterns, no variability of histone H3 methylation patterns was reported for hexaploid *Triticale* 'Fidelio'.

Chapter 6: Alteration of DNA methylation patterns in allopolyploid Triticale

6. 1 Summary

In this chapter two examples of alteration of DNA methylation patterns have been presented in hexaploid *Triticale* species. Firstly, seeds of F1, F2 and F3 generation obtained by crossing two advanced *Triticale* lines were analysed by Southern hybridization using 120-bp repeat unit family and 5S rDNA clones. DNA methylation changes were observed in F1, F2 and F3 individuals at symmetrical sites of 120-bp repetitive sequences but in less extent in 5S rDNA sequences. Conversely, asymmetrical DNA methylation seemed to be a variable feature of 5S rDNA sequences but not of 120-bp repetitive sequences F1, F2 and F3 individuals. As a possible consequence of DNA methylation changes, a trend from uniform to differentiated DNA methylation patterns has been found on chromosome spreads from F1 to F3 individuals by immunostaining with anti-methylcytosine antibody and subsequent *in situ* hybridization 120-bp repeat unit family clones. In addition, alteration of DNA methylation was induced by 5-azacytidine treatments in *Triticale* 'Fidelio'. The effects on germination and root emission were observed together with genome re-methylation mechanisms that arise as soon as 5-azacytidine treatments have stopped. DNA re-methylation appeared to be a quick process: in fact only one cell cycle was necessary to reach up to 60-80% of re-methylated metaphases, and it was not genome-specific as rye and wheat genomes were both affected.

6. 2 Introduction

Triticale hybrids are human-made cereal species with a very recent evolutionary history and good model for studying genome interactions in allopolyploids. The first artificial hybridization between wheat and rye dated back to the end of the 19th century (Wilson 1876) but many hexaploid and octoploid *Triticale* primary and secondary lines were not developed until the 19-sixties. Phenotipic instability has been reported for primary *Triticale* lines (Voylokov and Tikhenko 2002). A decrease in DNA content was

estimated to be about 28-30% and 9% in hexaploid and octoploid *Triticale* respectively (Boyko, Badaev et al. 1984). In addition, cytological instability of rye chromosomes into a wheat background was suggested by polymorphisms in the presence and the size of sites of two repetitive sequence families isolated from rye, pSc200 and pSc250, detected at chromosome ends of rye chromosomes in wheat-rye addition and substitution lines (Alkhimova, Heslop-Harrison et al. 1999). Furthermore, changes in rye gene expression are common in *Triticale* hybrids as reviewed by Ma *et al.* (2004). In hexaploid and octoploid *Triticale* hybrids changes in cytosine methylation patterns in comparison to their parental species were assessed by using AFLP and RFLP and resulted in the appearance and disappearance of DNA fragments (Ma, Fang et al. 2004). Ma *et al.* (2004) reported sequence elimination much more frequent than the appearance of novel DNA fragments and mainly targeting rye genome and repetitive DNA sequences rather than low-copy DNA sequences including genes.

The aim of the present work was to determine if DNA methylation instability is present in the progeny obtained by crossing two advanced *Triticale* lines. Alterations in whole-genome DNA methylation and methylation patterns 120-bp repeat unit family and 5S rDNA sequences were investigated in order to compare these results with the observations made in established *Triticale* 'Fidelio' and 'Lamberto' and presented in chapters 4 and 5.

5-azacytidine is an inhibitor of DNA methyltransferases. It acts by being incorporated in place of cytosines during DNA synthesis and then the modified base cannot be methylated anymore by maintenance DNA methyltransferases (Santi, Norment et al. 1984). Previous data in literature have shown the effect of 5-azacytidine treatments on DNA methylation and chromatin condensation. By investigating *T. aestivum* 'Beaver' carrying a translocation between the short arm of rye chromosome 1 (1RS) which has 25S rDNA gene, onto the long arm of wheat chromosome 1 (1BL), hypomethylation of CG dinucleotides was reported in 25S rDNA gene after 5-azacytidine treatment together with a complex response of chromatin condensation (Glyn, Egertova et al. 1997). In fact, opposite effects of chromatin condensation and decondensation were observed in rye chromosome arm (1RS) as a response to different concentration of 5-azacytidine so that the authors concluded that the drug affects chromatin structure depending on its local genome composition, either rich in genes or repetitive noncoding DNA sequences (Glyn, Egertova et al. 1997). In a following paper

Castilho *et al.* (1999) reported that weak and dispersed DNA methylation signal was present on chromosome spreads and interphase nuclei of *Triticale* 'Lasko' after 5-azacytidine treatments. In addition, the demethylation process did not appear to be genome-specific as no differences were observed between rye and wheat genomes in *Triticale* 'Lasko' (Castilho, Neves et al. 1999).

As it is proven that 5-azacytidine treatments induce genome demethylation, the aim of the present work was to investigate the process of re-methylation after such a treatment. By using immunostaining with antimethyl-cytosine antibody followed by *in situ* hybridization with 120-bp repeat unit family clones used as probes it was possible to evaluate if there was genome coordination of heterochromatin and euchromatin in re-methylation and whether some genomes or genome portions were preferentially re-methylated in respect to the others. Finally, these observations have been compared with whole-genome DNA methylation and methylation patterns 120-bp repeat unit family sequences of untreated *Triticale* 'Fidelio' previously analysed in chapter 5.

6. 3 Materials and methods

6. 3. 1 Genomic DNA extraction

F1, F2 and F3 generation of seeds were obtained by crossing two advanced *Triticale* lines, SEC776 from EUCARPIA Seed Bank and ID-3013 from IRTA Seed Bank, in ABR x ABR crosses (both lines were used as female and male partners) and subsequent self-pollination, and were kindly provided from IRTA Seed Bank (www.irta.es). Seeds were germinated at 20 °C in Petri dishes. All seeds from F2 and F3 generation germinated well while some seeds from F1 and ID-3013 did not. Conversely, lower germination rates were observed in the parental line SEC776. 10% bleach treatment to avoid contamination was performed and/or germination at different temperatures but very few seeds of parental line SEC776 germinated in a batch of twenty. Total genomic DNA from seedlings of three individuals of *Triticale* F1, F2 and F3 generation and two individuals of their parental lines ID-3013 and SEC776 was extracted from fresh young leaves using the CTAB method (Gawel and Jarret 1991). Genomic DNA was quantified by gel and spectrophotometer as listed in table 6. 1. The values of O.D.260/O.D.280 ranged from 1.59 to 2.16 (table 6. 1). Despite being exploited for DNA extraction and root tips, plant growth and phenotype seemed to be

Table 6. 1. Genomic DNA concentrations (ng/μl) and O.D.260/O.D.280 ratios of individuals from *Triticale* F1, F2, F3 generation and their parental lines SEC776 and ID-3013.

Species	Individual	ng/μl	O.D.260/ O.D.280
Triticale F1	1	2529.9	1.63
	2	2415.9	1.62
	3	2631.8	1.60
Triticale F2	1	1441.0	1.59
	2	1400.8	1.61
	3	1922.6	1.63
Triticale F3	1	1665.5	1.61
	2	1668.2	1.64
	3	1618.5	1.59
Triticale SEC776	1	1380.0	1.90
	2	-	-
Triticale ID-3013	1	3032.0	2.16
	2	1855.0	2.06

normal in all individuals, and some emitted spikelets with few seeds, with the exception of parental line SEC776, whose seedlings were not perfectly healthy.

6. 3. 2 Genomic DNA restriction with enzymes sensitive to cytosine methylation

Restrictions with isoschizomers *Msp*I-*Hpa*II and *Bst*NI-*Scr*FI were performed for detecting methylation at symmetrical CpG and CpNpG sites as well as with *Mcr*BC restriction enzyme for detecting asymmetrical methylation (fig. 4. 2). Their target sequence within the 120-bp repeats of *Secale* and *Triticum/Aegilops*, and the size of fragments expected in presence and absence of methylated cytosines at symmetrical CpG and CpNpG sites were previously showed in chapter 4 (figs. 4. 1a-b and 4. 6).

Genomic DNA (7.5μg) of three individuals of F1, F2 and F3 generation respectively and one to two individuals of their parental lines ID-3013 and SEC776 were restricted. For each enzyme digestion 4U/μg of restriction enzyme was used overnight at 37°C or the required temperature as described in Materials and Methods chapter. Enzyme digestions were performed twice for F1, F2 and F3 individuals showing no major differences.

6. 3. 3 Agarose gel analysis and Southern hybridization

As previously described in chapter 4 due to the short length of the subunit of 120-bp repeat unit family the digested DNA samples were run in 2% agarose gel electrophoresis in 1x TAE running buffer in order to have a good separation of short DNA fragments, together with 750 ng of undigested DNA samples used as controls. Southern hybridizations were carried out as described in Materials and Methods chapter. 750ng of digoxigenin-labelled clones belonging to 120-bp repeat unit family (see asterisks in fig. 3. 5b and table 3. 5) were used as probes. Membranes were stripped and re-hybridised with biotin-labelled 5S rDNA pTa794 clone (Gerlach and Dyer 1980).

6. 3. 4 Treatments with 5-azacytidine in *Triticale* 'Fidelio'

Seeds from *Triticale* 'Fidelio' were germinated for three days at 20 °C in Petri dishes containing water-soaked filter paper (control seeds) or filter paper soaked with 0.004 M 5-azacytidine solution freshly prepared every day. All germinated seeds were then washed in water, five seeds were fixed in ethanol-acetic acid (T0), while the others transferred in Petri dishes with water-soaked filter paper and fixed in ethanol-acetic acid in number of five at several times: at T1 (16 hrs), T2 (40 hrs), T3 (64 hrs), T4 (88 hrs) and T5 (112 hrs), corresponding approximately to 1, 3, 5, 7 and 9 cell cycles respectively, according to previous data about the cell cycle length in hexaploid cereals at 20°C which is approximately 12.5 h (Glyn, Egertova et al. 1997). After fixation seeds were planted into soil to observe their growth and phenotype.

6. 3. 5 Fixation of plant materials and preparation of chromosome spreads

As described in Material and Methods chapter, seedling root tips from of *Triticale* F1, F2 and F3 individuals and their parental lines SEC776 and ID-3013 were treated with ice-water, fixed in ethanol-acetic acid, digested with proteolytic enzymes, and squashed in 60% acetic acid to obtain chromosome spreads. Seedling root tips from 5-azacytidine treated *Triticale* 'Fidelio' fixed in ethanol-acetic acid at different times (from T0 to T5) after the 5-azacytidine treatments have stopped, were digested with proteolytic enzymes, and squashed in 60% acetic acid to obtain chromosome spreads. All slides from *Triticale* F1, F2 and F3 individuals, their parental lines SEC776 and ID-3013, and azacytidine-treated *Triticale* 'Fidelio' were treated with RNAse and pepsin, dehydrated in an ethanol series, and air-dried.

6. 3. 6 Immunostaining with anti-methylcytosine antibody and *in situ* hybridization

Slides preparations from three individuals of *Triticale* F1, F2, F3 generation, one to two individuals of parental lines SEC776 and ID-3013, and three to four individuals azacytidine-treated *Triticale* 'Fidelio' were analysed by using immunostaining with anti-methylcytosine antibody as described in Materials and Methods chapter. Immunostaining with anti-methylcytosine antibody was performed (5μg/ml per slide)

and the signal detected with Alexa 488-conjugated secondary antibody. Well spread metaphases recorded and photographed. Five to fifteen metaphases were analysed of *Triticale* F1, F2 and F3 individuals and their parental lines SEC776 and ID-3013, while eight to fifteen metaphases of azacytidine-treated *Triticale* 'Fidelio' individuals. Slide preparations of *Triticale* F1 and F3 individuals together with slides of azacytidine-treated *Triticale* 'Fidelio' were reopened, the methylation signal removed by washes, fluorescent *in situ* hybridization (FISH) was performed with biotin-labelled 120-bp repeat unit family clones (see asterisks in fig. 3. 5b and table 3. 5) and genomic *in situ* hybridization experiments (GISH.) were performed with Alexa 488-labelled and Alexa 546-labelled genomic DNA of *S. cereale* 'Petkus Spring' and *Ae. speltoides* respectively, and genomic DNA of *T. monococcum* as blocking DNA, as described in Materials and Methods chapter. Hybridization sites were detected indirectly using streptavidin conjugated to fluorochrome Alexa 594 and directly by Alexa 488-labelled and Alexa 546-labelled for *S. cereale* 'Petkus Spring' and *Ae. speltoides* genomes, respectively. Metaphases were re-located and photographed.

Karyotyping analyses of *Triticale* F1 and F3 individuals and azacytidine-treated *Triticale* 'Fidelio' were performed according to wheat and rye karyotypes (Mukai, Nakara et al. 1993; Cuadrado and Jouve 1994; Cuadrado and Jouve 1995) and the analyses presented in chapter 3 (see karyotypes in figs. 3. 7 and in 3. 8, and figs. 3. 11 and 3.12). Three to five metaphases of F1 and F3 individuals and two to three metaphases for 5-azacytidine treated *Triticale* 'Fidelio' (individuals fixed at T1) were analysed.

6. 4 Results

6. 4. 1 Crosses of *Triticale* lines: whole-genome DNA methylation and methylation patterns of 120-bp repeat unit family

Genomic DNA restrictions of individuals from *Triticale* F1, F2 and F3 generation and their parental lines SEC776 and ID-3013 were tested with isoschizomers *Msp*I-*Hpa*II and *Bst*NI-*Scr*FI as well as with *Mcr*BC restriction enzyme. Differences in whole-genome DNA methylation were observed at CpG and CpNpG sites between individuals of parental lines SEC776 and ID-3013. In parental line ID-3013 *Msp*I showed digestion to some extent while marked digestion with *Bst*NI is observed (lanes

2 and 4 in fig. 6. 1a). Among methylation-sensitive enzymes, *Hpa*II showed digestion to little extent at CpG sites while some digestion occurred with *Scr*FI at CpNpG sites (lanes 3 and 5 in fig. 6. 1a). In parental line SEC776 both enzyme pair *Msp*I-*Hpa*II and *Bst*NI-*Scr*FI showed digestion to a large extent suggesting that CpG and CpNpG sites were not methylated (lanes 8, 9, 10 and 11 in fig. 6. 1a). Asymmetrical methylation was detected after *Mcr*BC digestion, minimal in parental line ID-3013 and to a little extent in parental line SEC776 (lanes 6 and 12 respectively, in fig. 6. 1a).

After hybridization with 120-bp repeat unit family clones a strong ladder pattern with methylation-insensitive enzymes *Msp*I and *Bst*NI was clearly detected in both *Triticale* parental lines SEC776 and ID-3013 (lanes 2, 4, 8 and 10 in fig. 6. 1b). However, parental line SEC776 showed more intense signal with *Hpa*II and *Scr*FI compared to parental line ID-3013 (compare lanes 9 and 11 with lanes 3 and 5 respectively in fig. 6. 1b) suggesting lower methylation at CpG and CpNpG sites or as a consequence of some degraded DNA (notice the uncut DNA in lane 7 of fig. 6. 1b). Uniformly smeared signal was present in both parental species after *Mcr*BC digestion, but was stronger in parental line SEC776 (compare lanes 6 and 12 in fig. 6. 1b).

Differences in whole-genome methylation at CpG and CpNpG sites were observed between individuals of *Triticale* F1, F2 and F3 generation. Methylation-insensitive enzymes *Msp*I and *Bst*NI showed digestion to a large or some extent in all individuals (lanes 1, 3, 5, 7, 9, 11, 13, 15 and 17 in fig. 6. 2a and 6. 3a). Conversely, differences between individuals were observed with methylation-sensitive enzymes *Hpa*II and *Scr*FI. In fact, in F1 individual 2 and F2 individual 1 *Hpa*II showed little digestion (lanes 4 and 8 in fig. 6. 2a) while *Scr*FI showed digestion to some extent (lanes 4 and 8 in fig. 6. 3a), but F1 individuals 1 and 3 together with F2 individuals 2 and 3 appear digested to larger extent with *Hpa*II (lanes 2, 6, 10 and 12 in fig. 6. 2a) and *Scr*FI (lanes 2, 6, 10 and 12 in fig. 6. 3a) showing that in these individuals CpG and CpNpG sites were poorly or not methylated. However, all three individuals of the F3 generation gave more uniform results with *Hpa*II showing little digestion (lanes 14, 16 and 18 in fig. 6. 2a) and with *Scr*FI to some extent, suggesting that CpG and CpNpG sites were methylated (lanes 14, 16, and 18 in fig. 6. 3a). By using *Mcr*BC restriction enzyme no major differences in asymmetrical methylation were found between individuals: all showed minimal digestion (fig. 6. 4a), with the exception of F1 individual 1 which showed digestion to a some extent (lane 1 in fig. 6. 4a) as previously shown for parental line SEC776 (lane 12 in fig. 6. 1a).

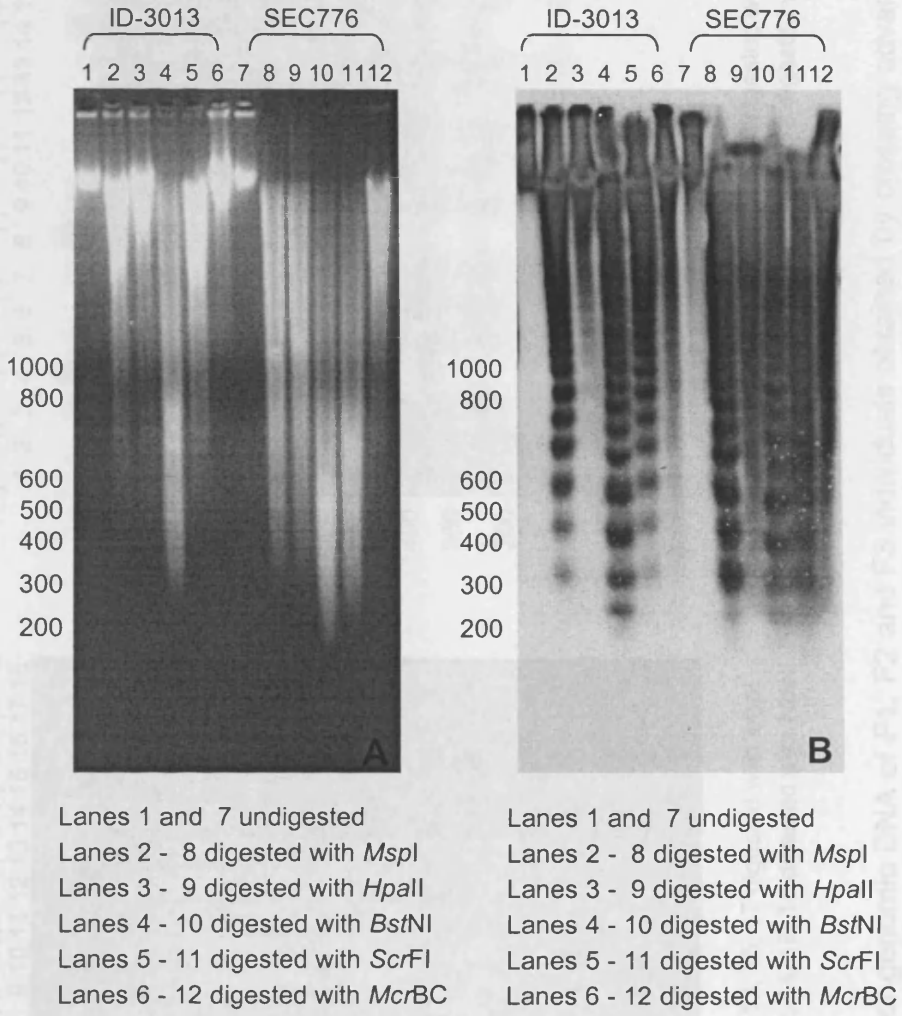
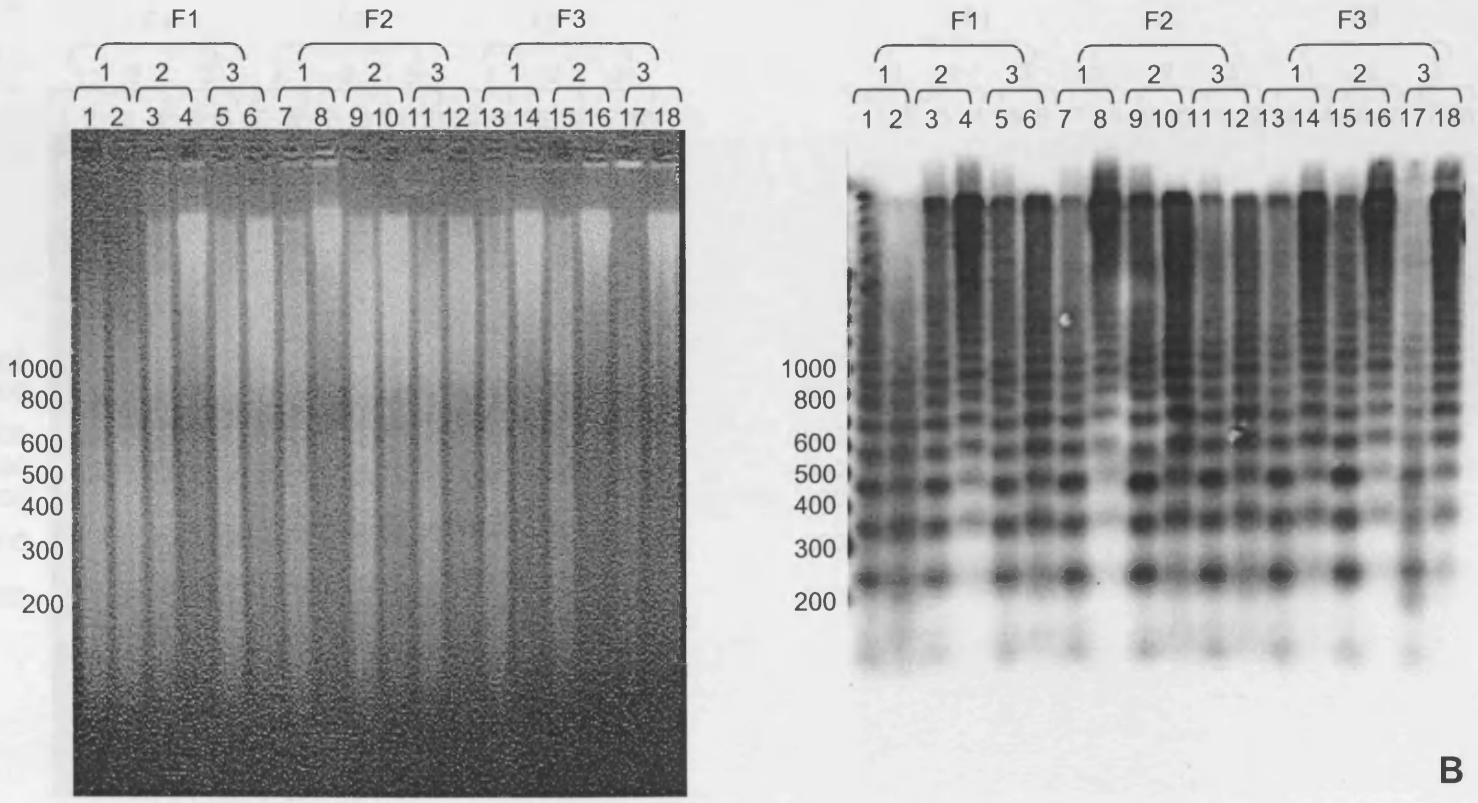


Figure 6. 1. A) Gel picture of genomic DNA of advanced *Triticale* (ABR genome) lines ID-3013 and SEC776 digested with isoschizomers *Msp*I-*Hpa*II and *Bst*NI-*Scr*FI as well as *Mcr*BC restriction enzyme. Southern hybridization with digoxigenin-labelled Svav25/208!182 clone (B).



Lanes 1, 3, 5, 7, 9, 11, 13, 15, 17 digested with *Msp*I
Lanes 2, 4, 6, 8, 10, 12, 14, 16, 18 digested with *Hpa*II

Lanes 1, 3, 5, 7, 9, 11, 13, 15, 17 digested with *Msp*I
Lanes 2, 4, 6, 8, 10, 12, 14, 16, 18 digested with *Hpa*II

Figure 6. 2. A) Gel picture of genomic DNA of F1, F2 and F3 individuals obtained by crossing advanced *Triticale* (ABR genome) lines ID-3013 and SEC776, digested with *Msp*I-*Hpa*II restriction enzymes. Southern hybridization with digoxigenin-labelled Tmono25/208!1212 clone (B).

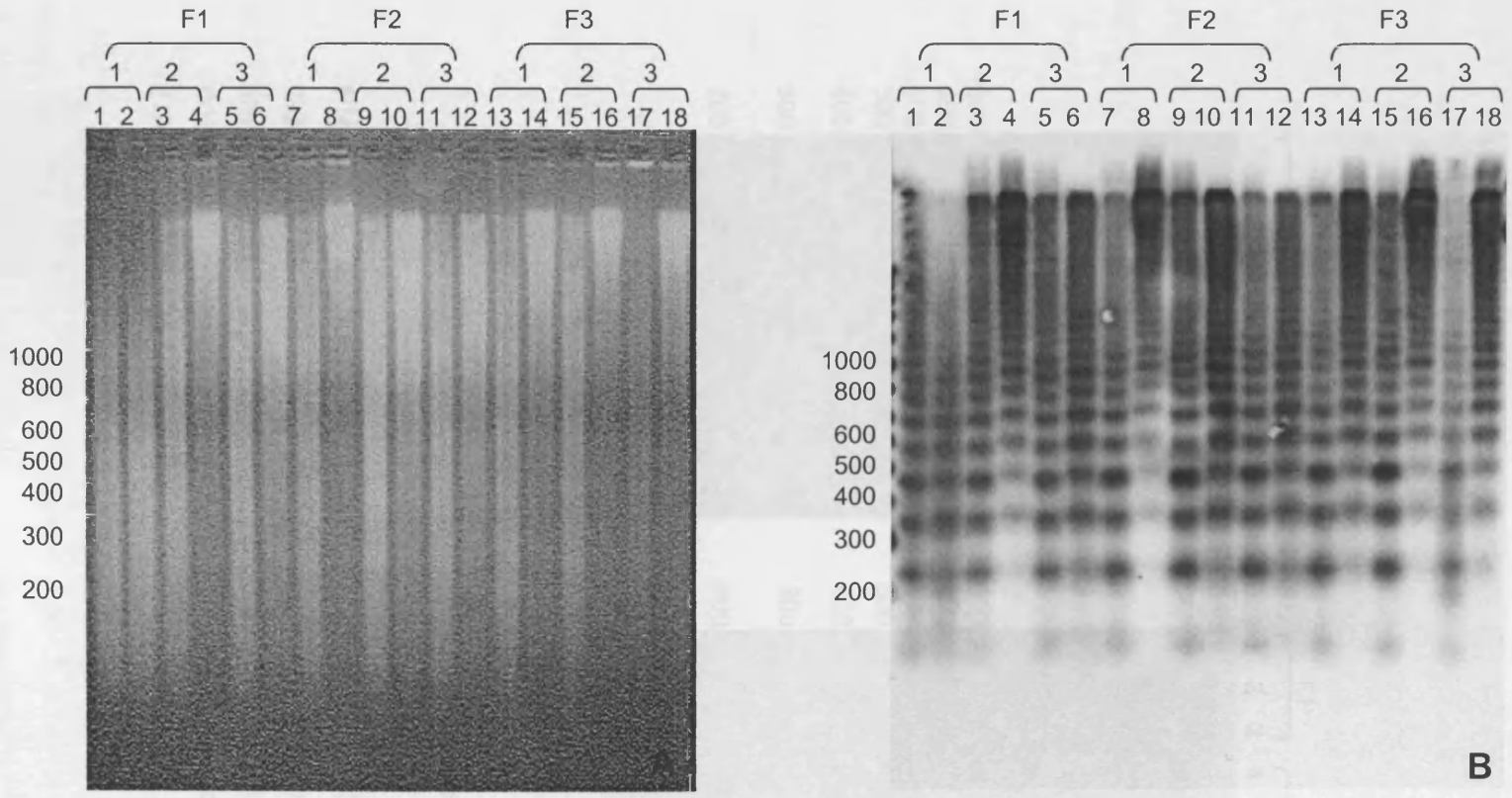


Figure 6. 3. A) Gel picture of genomic DNA of F1, F2 and F3 individuals obtained by crossing advanced *Triticale* (ABR genome) lines ID-3013 and SEC776, digested with *Bst*NI-*Scr*FI restriction enzymes. Southern hybridization with digoxigenin-labelled Pet(22594)25/42!3324 clone (B).

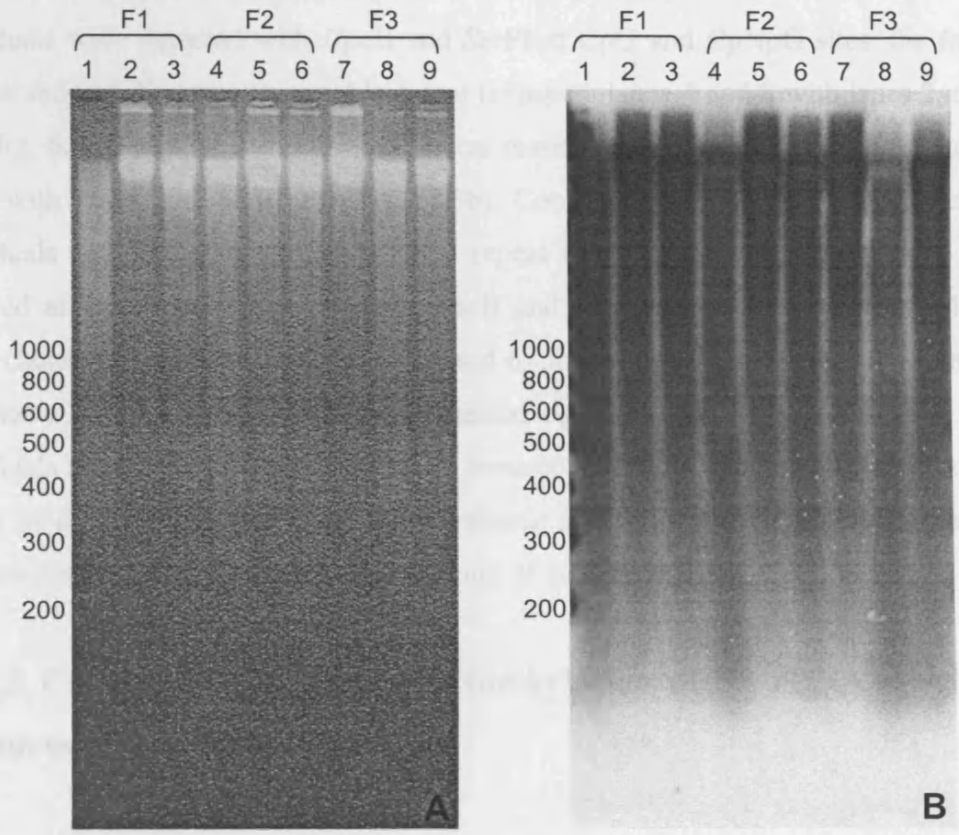


Figure 6. 4. A) Gel picture of genomic DNA of F1, F2 and F3 individuals obtained by crossing advanced *Triticale* (ABR genome) lines ID-3013 and SEC776, digested with *Mcr*BC restriction enzyme. Southern hybridization with digoxigenin-labelled Tmono25/208!1212 clone (B).

After hybridization with 120-bp repeat unit family clones *Triticale* F1, F2, and F3 individuals showed strong ladder patterns with methylation-insensitive enzymes *Msp*I and *Bst*NI (lanes 1, 3, 5, 7, 9, 11, 13, 15 and 17 in fig. 6. 2b and 6. 3b). Multiple bands resulting from different combinations of cutting sites within adjacent monomers as previously shown in chapter 4 (figs. 4. 1a-b and 4. 6) were observed with both enzymes. Differences in 120-bp repeat unit family methylation between F1 and F2 individuals were detected with *Hpa*II and *Scr*FI at CpG and CpNpG sites: the former enzyme showed digestion in variable extent (compare lanes 4 and 8 with lanes 2, 6 and 10 in fig. 6. 2b) and the latter a more or less marked ladder patterns (compare lanes 4 and 8 with lanes 2, 6 and 10 in fig. 6. 3b). Conversely, no differences between F3 individuals and throughout higher 120-bp repeat unit family methylation levels were observed at CpG and CpNpG sites as *Hpa*II and *Scr*FI showed digestion to a lesser extent (lanes 14, 16, and 18 in fig. 6. 2b and 6. 3b) than in F1 individuals 1-3 and F2 individuals 2-3. A consistent uniformly smeared signal was detected with *Mcr*BC in all individuals (fig. 6. 4b) as reported in their parental lines SEC776 and ID-3013 (lanes 6 and 12 in fig. 6. 1b) suggesting that in *Triticale* 120-bp repetitive sequence family is more asymmetrically methylated than the bulk of genomic DNA.

6. 4. 2 Crosses of *Triticale* lines: methylation of 5S rDNA repetitive sequences

To investigate methylation changes in another repetitive fraction of the genomes, Southern membranes of F1, F2 and F3 individuals were re-hybridized with 5S rDNA pTa794 clone. A ladder pattern even if not regular was detected with *Msp*I (lanes 1, 3, 5, 7, 9, 11, 13, 15, 17 in fig. 6. 5a) and bands corresponding to about 400 bp (monomer), 800 bp (dimer) and 1200 bp (trimer) of different intensity together with bands shorter than the monomer length (about 200 bp although not always visible) were observed with *Bst*NI (lanes 1, 3, 5, 7, 9, 11, 13, 15, 17 in fig. 6. 5b). Differences were found between individuals from F1 and F2 generation at methylated CpG sites as *Hpa*II showed digestion to a large extent in F1 individuals 1 and 3 compared to F1 individual 2 (compare lanes 2 and 6 with lane 4 in fig. 6. 5a) and F2 individuals 2 and 3 compared to F2 individual 1 (compare lanes 10 and 12 with lane 8 in fig. 6. 5a) while in all F3 individuals *Hpa*II did show digestion to a little extent (lanes 14, 16 and 18 in fig. 6. 5a).

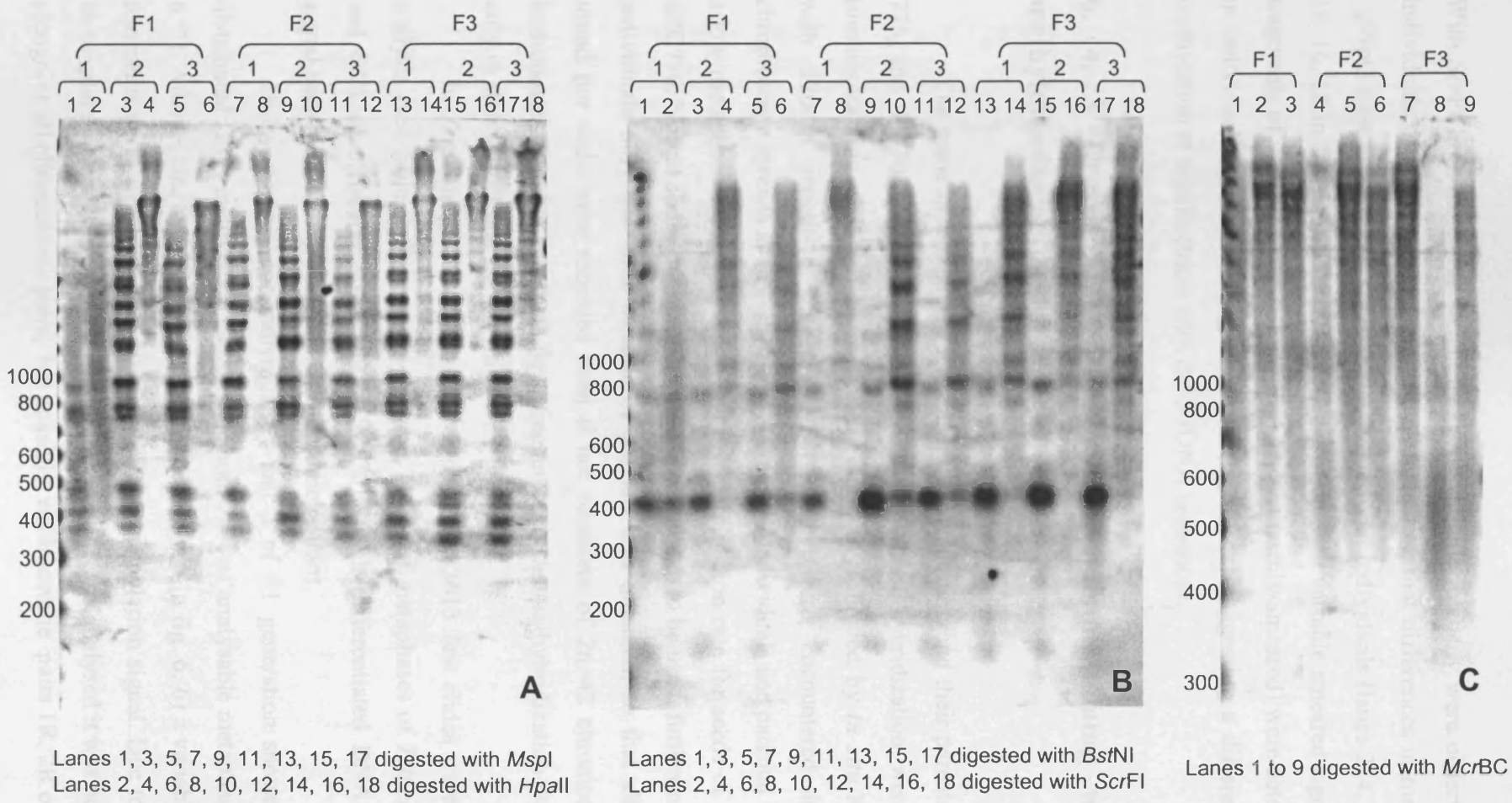


Figure 6. 5. Southern hybridizations performed with biotin-labelled 5S rDNA pTa 794 clone of genomic DNA of F1, F2 and F3 individuals obtained by crossing advanced *Triticale* (ABR genome) lines ID-3013 and SEC776, digested with isoschizomers *Msp*I-*Hpa*II (A) and *Bst*NI-*Scr*FI (B) as well as *Mcr*BC (C) restriction enzyme.

With *Scr*FI high-molecular weight bands of variable intensity were observed in some individuals of F1, F2 and F3 generations indicating that differences in methylation of CpNpG sites in 5S rDNA sequences occur between individuals (lanes 2, 4, 6, 8, 10, 12, 14, 16, 18 in fig. 6. 5b). After *Mcr*BC digestion high variable smeared signals made of fragments of variable lengths (high- low- and uniformly-smeared) were detected mainly in individuals of F2 and F3 generations (fig. 6. 5c) suggesting a different extent of methylation at asymmetrical sites in 5S rDNA sequences.

6. 4. 3 Crosses of *Triticale* lines: immunostaining with anti-methylcytosine antibody

The same individuals of F1, F2 and F3 generation and their parental lines SEC 776 and ID-3013 previously analysed by Southern hybridization were tested by immunostaining with anti-methylcytosine antibody followed by *in situ* hybridization with 120-bp repetitive sequences. Difficulties were encountered in preparing chromosome spreads as in *Triticale* F1, F2 and F3 individuals and parental line ID-3013 a low metaphase index was found and low germination rate for seeds of parental line SEC776. Not all slides preparations were good enough to be used, furthermore for some individuals only one slide suitable of analysis was obtained, so that all metaphases found per slide were recorded even if not complete of 2n=42 chromosomes. After immunostaining with anti-methylcytosine antibody, *in situ* hybridization was performed only in slides of F1 and F3 individuals.

In *Triticale* parental lines SEC 776 and ID-3013 few slides were suitable for analysis (see table 6. 2) however showing 71.4% of metaphases of *Triticale* SEC 776 and 80% of metaphases in *Triticale* ID-3013 with differentiated DNA methylation signal between chromosomes and chromosome regions.

Slides obtained for individuals 1 and 2 of F1 generation showed uniformly distributed DNA methylation patterns in about 70% of analysable metaphases (see table 6. 2). After *in situ* hybridization (an example is given in fig. 6. 6) a variable number of telomeres up to a maximum of 10 had no DNA methylation signal. Due to presence of incomplete metaphases among the chromosomes spreads analysed it was not possible to recognize all chromosome pairs, but possibly chromosome pairs 1R, 2R or 3R (having

Table 6. 2. DNA methylation patterns in individuals of *Triticale* F1, F2, F3 generation and their parental lines ID-3013 and SEC776 after immunostaining with anti-methylcytosine antibody.

Species	Generation	Individual	Number of metaphases photographed	Number of metaphases analysable*	Number of metaphases (uniform)	Number of metaphases (uneven)	Total % of metaphases (uneven)**
Triticale crosses (F1, F2, F3 generation) (ABR genome, 2n=6x=42)	F1	1	5	5	4	1	20.0%
		2	2	1	-	1	100.0%
		3	-	-	-	-	-
	TOT		7	6	4	2	33.3%
	F2	1	6	5	-	5	100.0%
		2	7	4		4	100.0%
		3	2	2	-	2	100.0%
	TOT		15	11		11	100.0%
	F3	1	-	-	-	-	-
		2	6	4	-	4	100.0%
		3	15	11	1	10	90.9%
	TOT		21	15	1	14	93.3%
Triticale parental lines (ABR genome, 2n=6x=42)	ID-3013	1	3	3	1	2	50.0%
		2	2	2	-	2	100.0%
	TOT		5	5	1	4	80.0%
	SEC776	1	7	7	2	5	71.40%

* Cells with chromosome distorsion or staining artefacts were not analysed.

**100% equals to analysed cells.

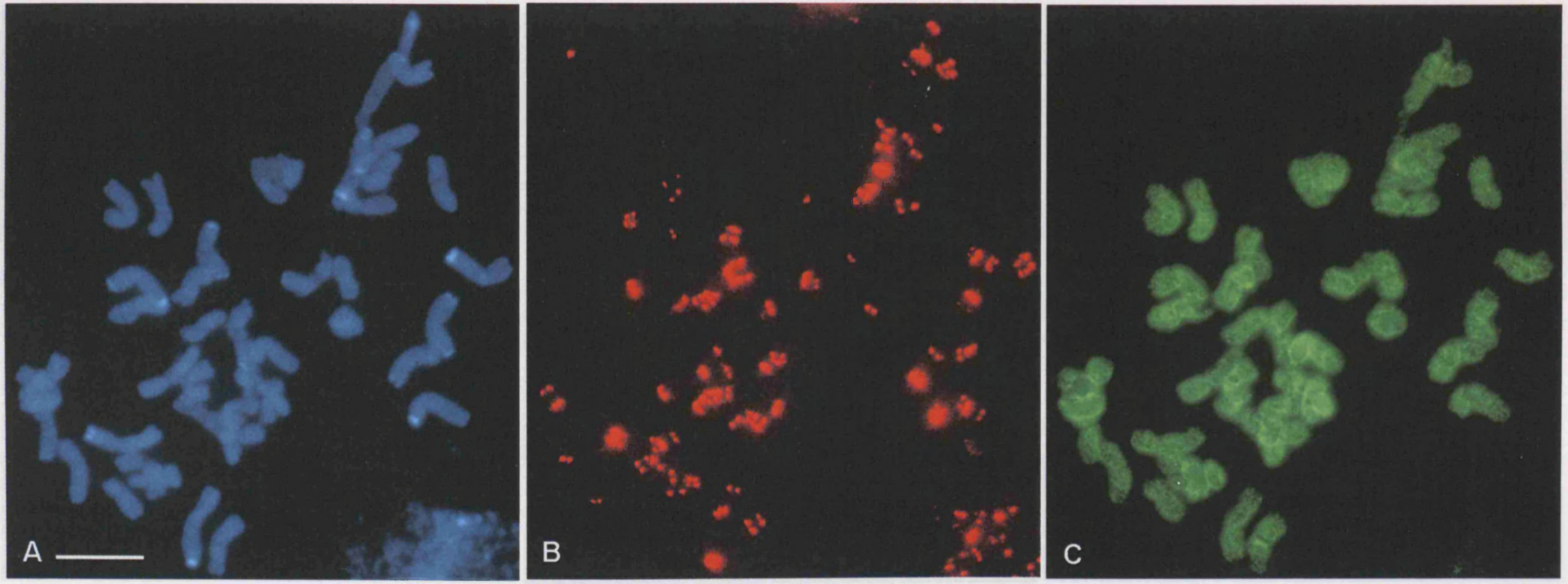


Figure 6. 6. Root tip metaphase of *Triticale* F1 individual 2 (ABR genome, 2n=42). A) DAPI stained chromosomes. B) *In situ* hybridization with biotin-labelled Tmono25/208!1212 probe detected with Alexa 594 red fluorescence. C) Immunostaining with anti-methylcytosine antibody detected with Alexa 488 green fluorescence. Scale bar represents 10 μm.

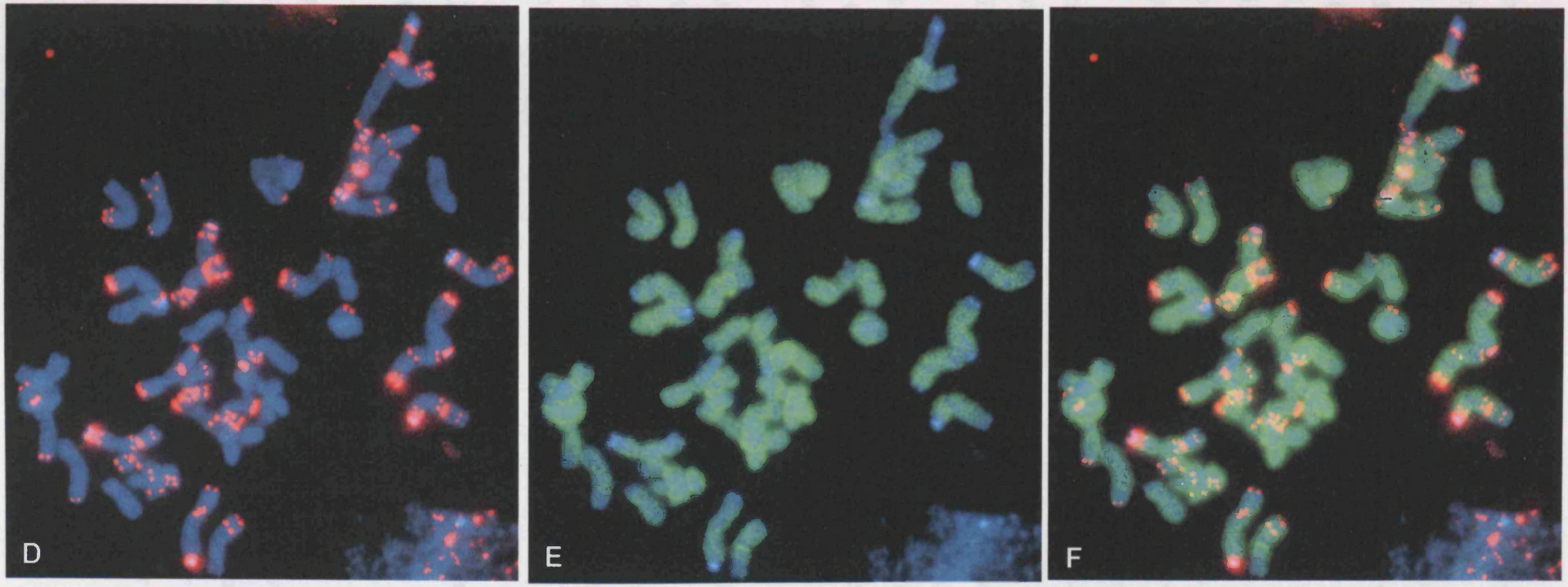


Figure 6. 6 continued. Root tip metaphase of *Triticale* F1 individual 2 (ABR genome, 2n=42). D) Overlay of A and B images. E) Overlay of A and C images. F) Overlay of A, B and C images.

the same banding pattern with 120-bp repeat unit family clones), 6R and 7R had no DNA methylation signal at telomeres (fig. 6. 6).

All metaphases analysed from individuals of F2 generation showed a more unevenly distributed DNA methylation along chromosomes with enhanced or reduced signal in different chromosome regions (see table 6. 2). A variable number of telomeres lacked DNA methylation signal but it was not possible to identify them as no further *in situ* hybridization was performed due to the low quality of chromosome spreads.

From individuals 2 and 3 of F3 generation good preparations were obtained (see table 6. 2). Both F3 individuals showed unevenly distributed DNA methylation with enhanced or reduced signal in different chromosome regions of A, B and R chromosomes in most metaphases analysed (93% of metaphases, see table 6. 2). An example of unevenly distributed DNA methylation patterns in F3 individuals is shown in fig. 6. 7. Some chromosomes showed enhanced DNA methylation signal at one or two arms, close to the telomeric region or in intercalary location, other chromosomes lacked DNA methylation at centromeres. In addition, a variable number of rye telomeres (a maximum of 10) had no DNA methylation, at chromosome pairs 2R or 3R (having the same banding pattern with 120-bp repeat unit family clones), 4R, 5R, and 6R.

Taking together these results gave evidence of high variability and instability of DNA methylation patterns in F1, F2 and F3 individuals even if the number of slides suitable for analyses was limited. It seemed that a trend was present from a uniform DNA methylation patterns in F1 individuals to more differentiated DNA methylation patterns in F3 individuals.

6. 4. 4 *Triticale* 'Fidelio' after 5-azacytidine treatments: germination and growth

As mentioned above, 5-azacytidine can be incorporated in place of cytosines during DNA synthesis and cannot be methylated anymore by maintenance DNA methyltransferases (Santi, Norment et al. 1984). The concentration of 5-azacytidine used in the experiments described here was 100 times higher than it was used in previous works, where it was reported to be not enough for a complete genome demethylation (Glyn, Egertova et al. 1997; Castilho, Neves et al. 1999), in the attempt to obtain complete demethylated DNA sequences.

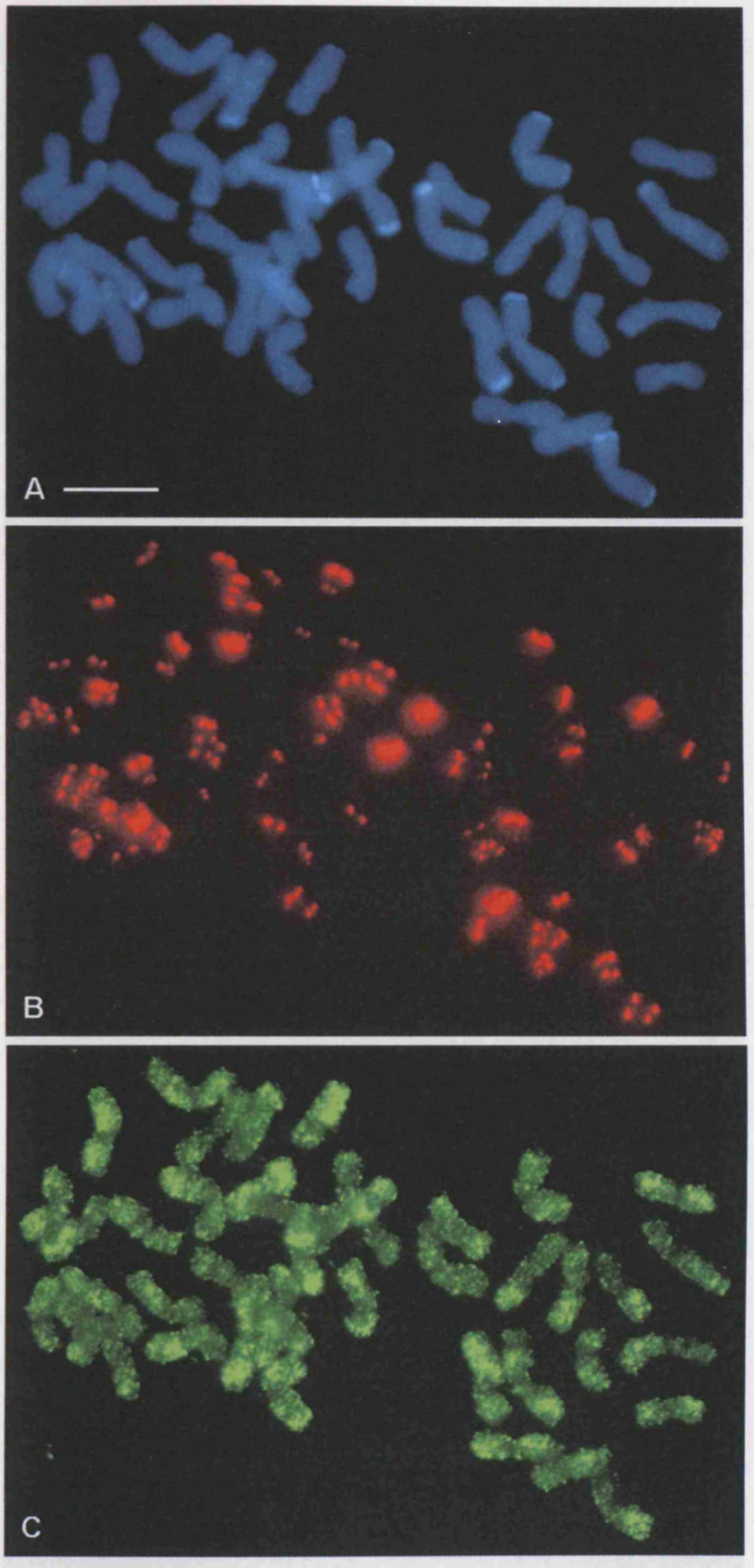


Figure 6. 7. Root tip metaphase of *Triticale* F3 individual 3 (ABR genome, 2n=42). A) DAPI stained chromosomes. B) *In situ* hybridization with biotin-labelled Tmono25/208!1212 probe detected with Alexa 594 red fluorescence. C) Immunostaining with anti-methylcytosine antibody detected with Alexa 488 green fluorescence. Scale bar represents 10 μm.

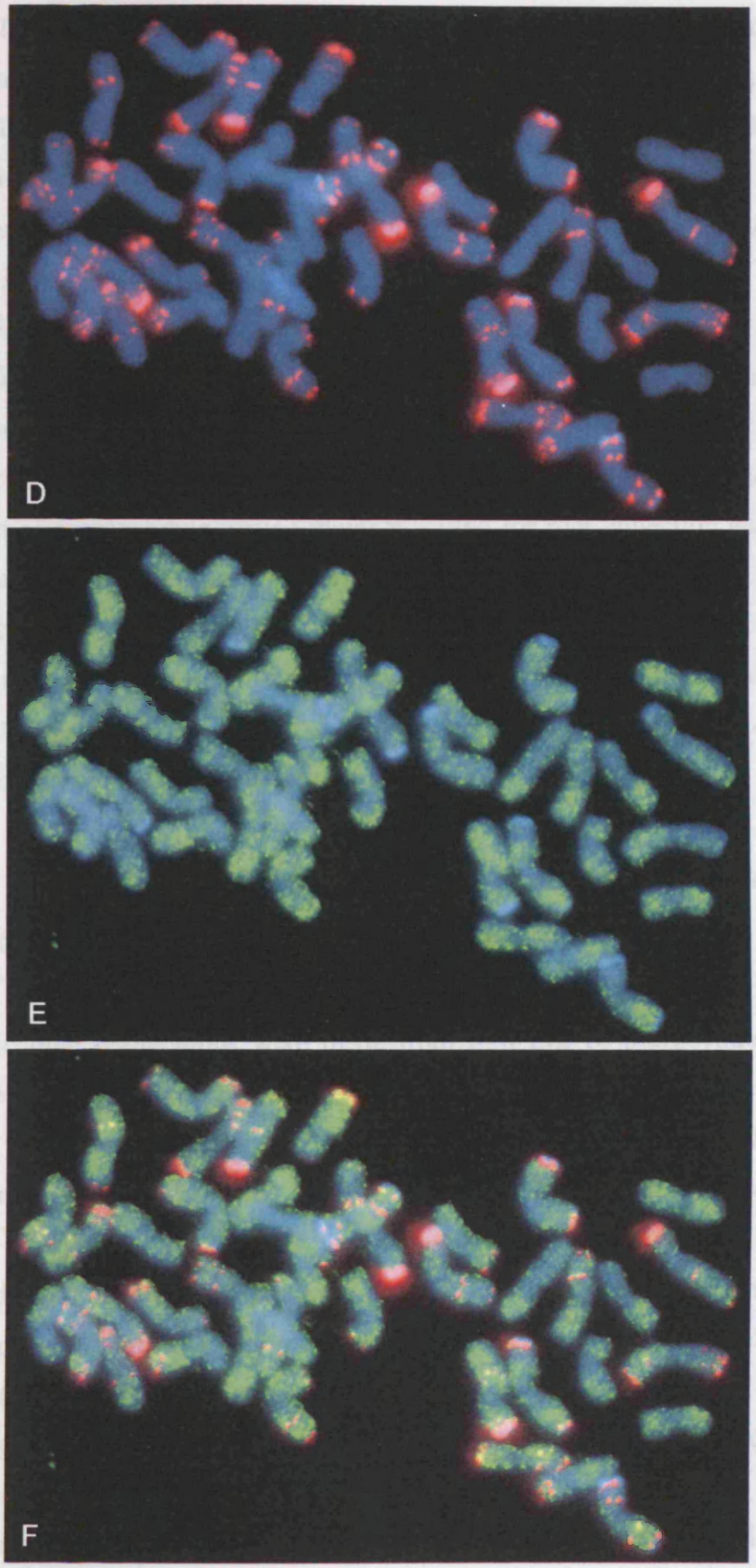


Figure 6. 7 continued. Root tip metaphase of *Triticale* F3 individual 3 (ABR genome, 2n=42). D) Overlay of A and B images. E) Overlay of A and C images. F) Overlay of A, B and C images.

While *Triticale* 'Fidelio' seeds germinated in water had 100% germination rate, the 5-azacytidine treated seeds had lower germination rate: 13 from 17 seeds (76.5%) and 16 from 21 seeds (76.2%) analysed in two different experiments. In addition, roots from 5-azacytidine treated seeds showed slower growth and were present in lower number per seed (only two to three instead of four/six roots present in untreated seeds) indicating that DNA methylation is an important factor for seed germination.

Treated plants showed very slow growth in greenhouse compared to untreated controls and died at the stage of apical shoot emergence so that it was not possible to observe phenotypic features (fig. 6. 8). To evaluate if being exploited for roots necessary for cytological investigations possibly contributed to the fate of the seedlings, a control batch of 5-azacytidine treated plants not used for root collection was grown. Germination rate (12 from 15 seeds, 80 %) and number of roots emerging (in the range of one to three) were in agreement to previous observations. After three weeks 5-azacytidine treated plants showed very slow growth compared to the untreated controls and all died at the stage of first apical shoot emergence, showing that root collection was not affecting growth and development of 5-azacytidine treated plants.

6. 4. 5 *Triticale* 'Fidelio' after 5-azacytidine treatments: immunostaining with anti-methylcytosine antibody

Chromosome preparations were made from root tips of 5-azacytidine treated *Triticale* 'Fidelio' fixed immediately after the 5-azacytidine treatments have stopped (at T0), and after 16 hrs (at T1), 40 hrs (at T2), 64 hrs (at T3), 88 hrs (at T4) and 112 hrs (at T5), corresponding approximately to 1, 3, 5, 7 and 9 cell cycles respectively (Glyn, Egertova et al. 1997). Immunostaining with antimethyl-cytosine antibody was performed and more detailed analyses were obtained on chromosome spreads by using *in situ* hybridization with 120-bp repeat unit family clones and genomic *in situ* hybridization (GISH.) with Alexa 488- and Alexa 546-labelled genomes from *S. cereale* 'Petkus Spring' (R genome) and *Ae. speltoides* (B genome) respectively and genomic DNA of *T. monococcum* (A genome) used as blocking DNA .

In slide preparations from root tips of *Triticale* 'Fidelio' fixed immediately after 5-azacytidine treatments have stopped (at T0) no DNA methylation signal was detected in the majority of metaphases after immunostaining with antimethyl-cytosine antibody (see table 6. 3). The percentage of chromosome spreads with DNA methylation signal

Figure 6. 8. Growth and phenotype of 5-azacytidine treated *Triticale* 'Fidelio' individuals (A) and untreated *Triticale* 'Fidelio' individuals (B) after being planted into soil and grown in greenhouse for three weeks.

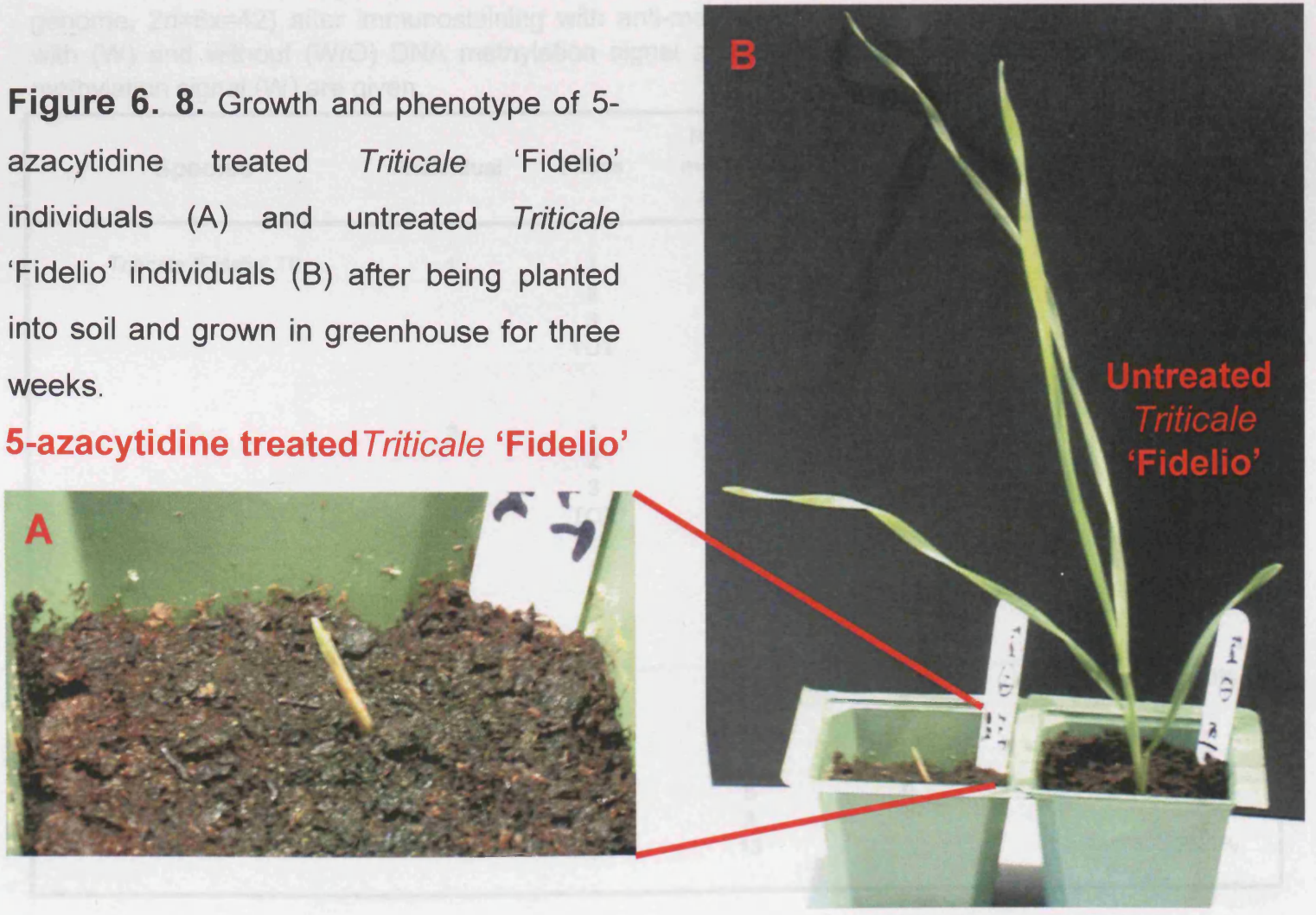


Table 6. 3. DNA methylation patterns in individuals of 5-azacytidine treated *Triticale* 'Fidelio' (ABR genome, 2n=6x=42) after immunostaining with anti-methylcytosine antibody. Number of metaphases with (W) and without (W/O) DNA methylation signal and total percentage of metaphases with DNA methylation signal (W) are given.

Species	Individual	Slides	Number of metaphases analysed	Number of metaphases (W)	Number of metaphases (W/O)	Total % of metaphases *(W)
Triticale 'Fidelio' T0	1	1	5	0	5	
		2	5	1	4	
		3	4	3	1	
		TOT	14	4	10	28.6%
	2	1	5	2	3	
		2	5	0	5	
		3	-	-	-	
		TOT	10	2	8	20.0%
	3	1	5	0	5	
		2	-	-	-	
		3	5	1	4	
		TOT	10	1	9	10.0%
Triticale 'Fidelio' T1	1	1	5	0	5	
		2	5	0	5	
		3	3	3	0	
		TOT	13	3	10	23.1%

Table 6. 3. continued

Species	Individual	Slides	Number of metaphases analysed	Number of metaphases (W)	Number of metaphases (W/O)	Total % of metaphases *(W)
Triticale 'Fidelio' T1	2	1	5	0	5	
		2	5	4	1	
		3	5	5	0	
		TOT	15	9	6	60.0%
	3	1	5	5	0	
		2	-	-	-	
		3	5	3	2	
		TOT	10	8	2	80.0%
	4	1	-	-	-	
		2	3	0	3	
		3	5	2	3	
		TOT	8	2	6	25.0%

*100% equals to analysed cells.

ranged from 10% to 28.6% in individuals fixed at time T0 and where present the signal was weak and patchy along chromosomes. Furthermore, micronuclei (fig. 6. 9) were present. In addition, by using genomic *in situ* hybridization (GISH.) intermixed nuclei with no spatial separation between rye and wheat genomes were observed (fig. 6. 10) in contrast with the structure of interphase nuclei in untreated *Triticale* 'Fidelio' where rye and wheat genomes occupied different nuclear regions (fig. 6. 11).

At time of fixation T1 (16 hrs) approximately one cell cycle could have occurred. Between four individuals high variability in the number of metaphases with DNA methylation signal was observed (see table 6. 3). The percentage of chromosome spreads with DNA methylation signal ranged from 23.1% to 80% showing major differences between individuals. DNA methylation was present in individuals 1 and 4 in low percentages of metaphases (23.1% and 25% respectively), whereas in individuals 2 and 3 in high percentages of metaphases (60% and 80% respectively). Where present on chromosome spreads of individuals fixed at T1, the DNA methylation signal was more intense (examples are given in figs. 6. 12b and 6. 13e) if compared to chromosome spreads of individuals fixed at T0. After genomic *in situ* hybridization (GISH.) using Alexa 488-labelled and Alexa 546-labelled genomic DNAs of *S. cereale* 'Petkus Spring' and *Ae. speltoides* respectively, DNA methylation appeared on chromosomes belonging to rye and wheat parental genomes as shown in individual 3 (fig. 6. 12). More detailed analysis were performed in individual 2 by *in situ* hybridization with 120-bp repeat unit family clones revealing one chromosome arm or both showing enhanced DNA methylation signal on chromosomes belonging to R and AB genomes, while other chromosomes belonging to R and AB genomes appeared to be weakly methylated (fig. 6. 13). In addition, a variable number of rye telomeres (maximum of 10) possibly belonging to chromosome pairs 2R or 3R (having the same banding pattern with 120-bp repeat unit family clones), 4R and 5R were not methylated.

In chromosome spreads made from root tips fixed from T2 (40 hrs), T3 (64 hrs) and T4 (88 hrs), when 3, 5 and 7 cell cycles could have occurred respectively, DNA methylation was present in interphase nuclei at the level of normal untreated roots, but meristems were scarcely present, low or no mitotic index was counted and only interphase nuclei were observed suggesting that from fixation times T2 to T4 roots have stopped their growth. At fixation time T5, after about 10 cell cycles could have occurred, few metaphases were obtained and differentiated DNA methylation patterns were encountered after immunostaining with anti-methylcytosine antibody. However,

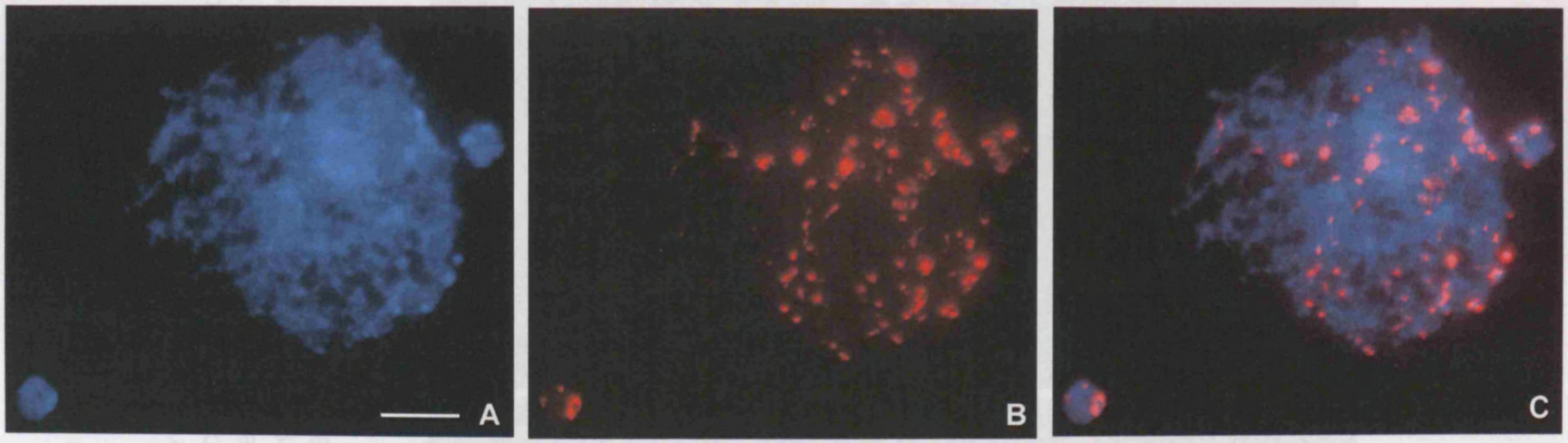


Figure 6. 9. Interphase nucleus with micronuclei of 5-azacytidine treated *Triticale* 'Fidelio' (ABR genome, 2n=42). A) DAPI staining. B) *In situ* hybridization with biotin-labelled Tmono25/208!1212 probe detected with Alexa 594 red fluorescence. C) Overlay of A and B images. Scale bar represents 10 μm.

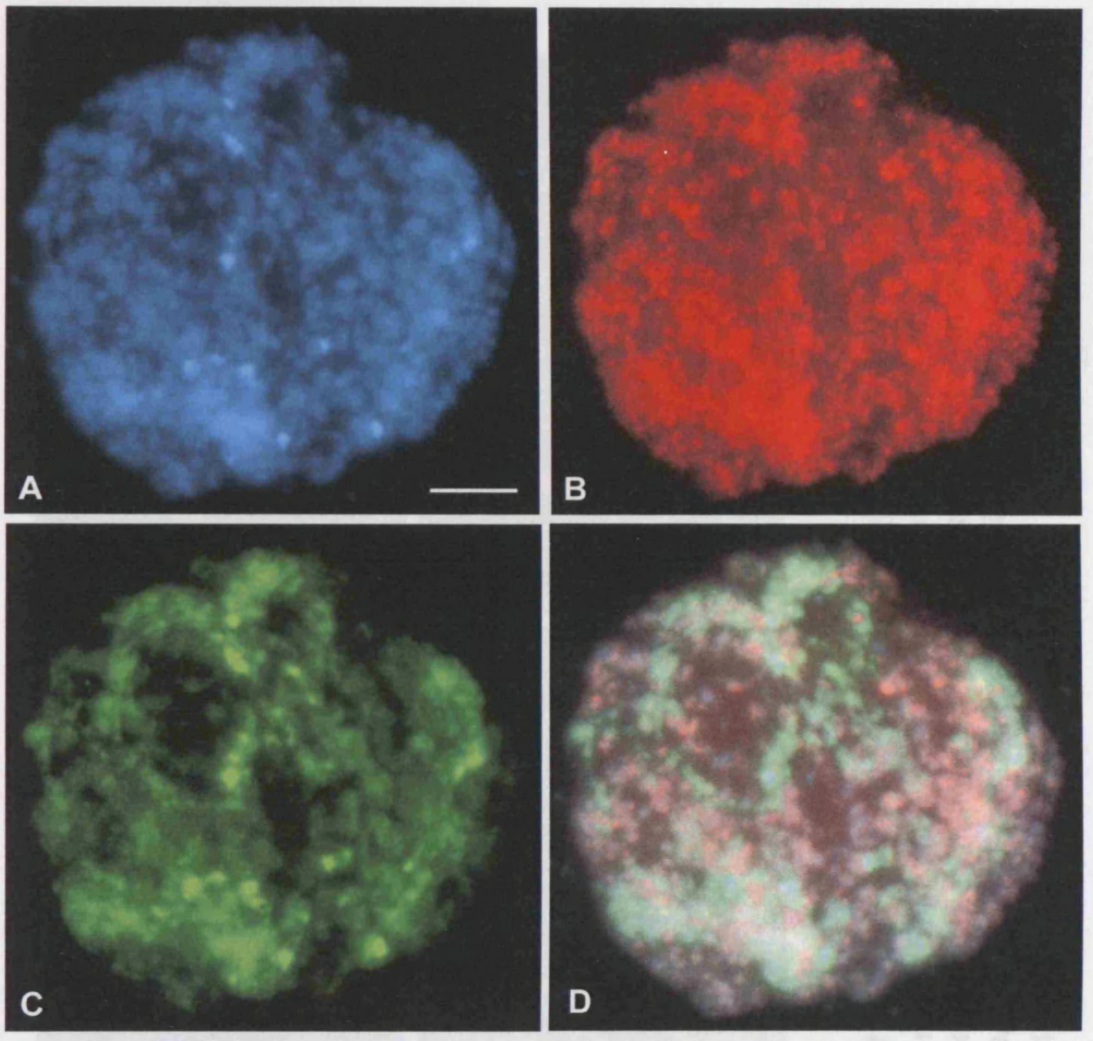


Figure 6. 10. Interphase nucleus of 5-azacytidine treated *Triticale* 'Fidelio' (ABR genome, 2n=42). A) DAPI staining. B) *In situ* hybridization with genomic DNA of *S. cereale* 'Petkus Spring' (Alexa 488 green fluorescence). C) *In situ* hybridization with genomic DNA of *Ae. speltoides* (Alexa 546 red fluorescence). D) Overlay of A, B and C images. Scale bar represents 10 μm.

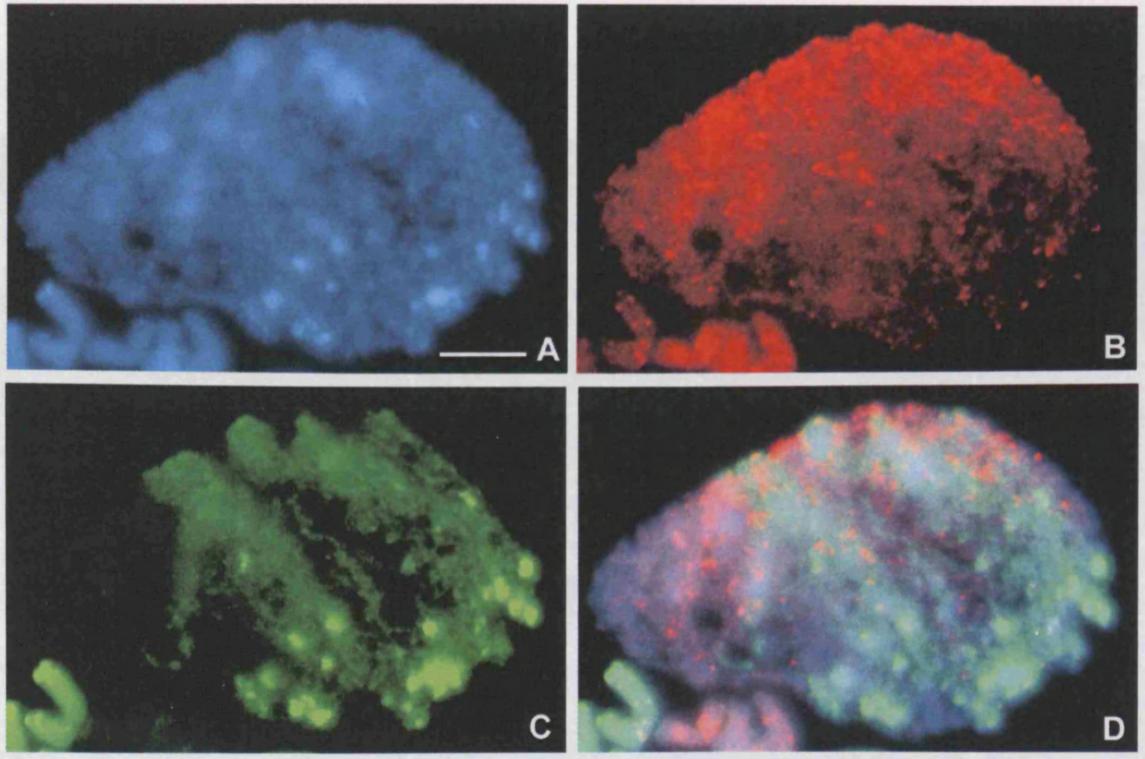


Figure 6. 11. Interphase nucleus of *Triticale* 'Fidelio' (ABR genome, 2n=42). A) DAPI staining. B) *In situ* hybridization with genomic DNA of *S. cereale* 'Petkus Spring' (Alexa 488 green fluorescence). C) *In situ* hybridization with genomic DNA of *Ae. speltoides* (Alexa 546 red fluorescence). D) Overlay of A, B and C images. Scale bar represents 10 μm.

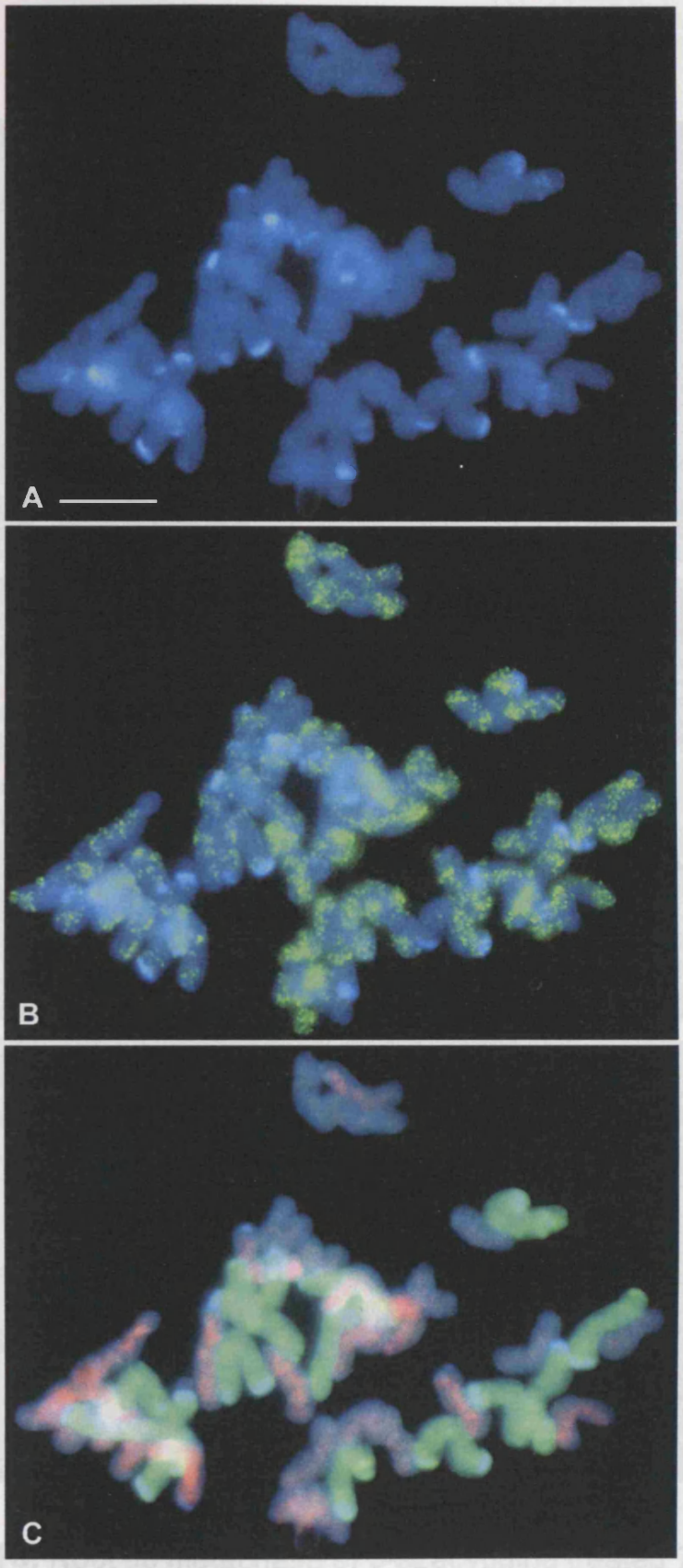


Figure 6. 12. Root tip metaphase of 5-azacytidine treated *Triticale* 'Fidelio' individual 3 (ABR genome, 2n=42). A) DAPI stained chromosomes. B) Immunostaining with anti-methylcytosine antibody detected with green Alexa 488 fluorescence. C) I*n situ* hybridization with genomic DNA of *S. cereale* 'Petkus Spring' (Alexa 488 green fluorescence) and *Ae. speltoides* (Alexa 546 red fluorescence). Scale bar represents 10 µm.

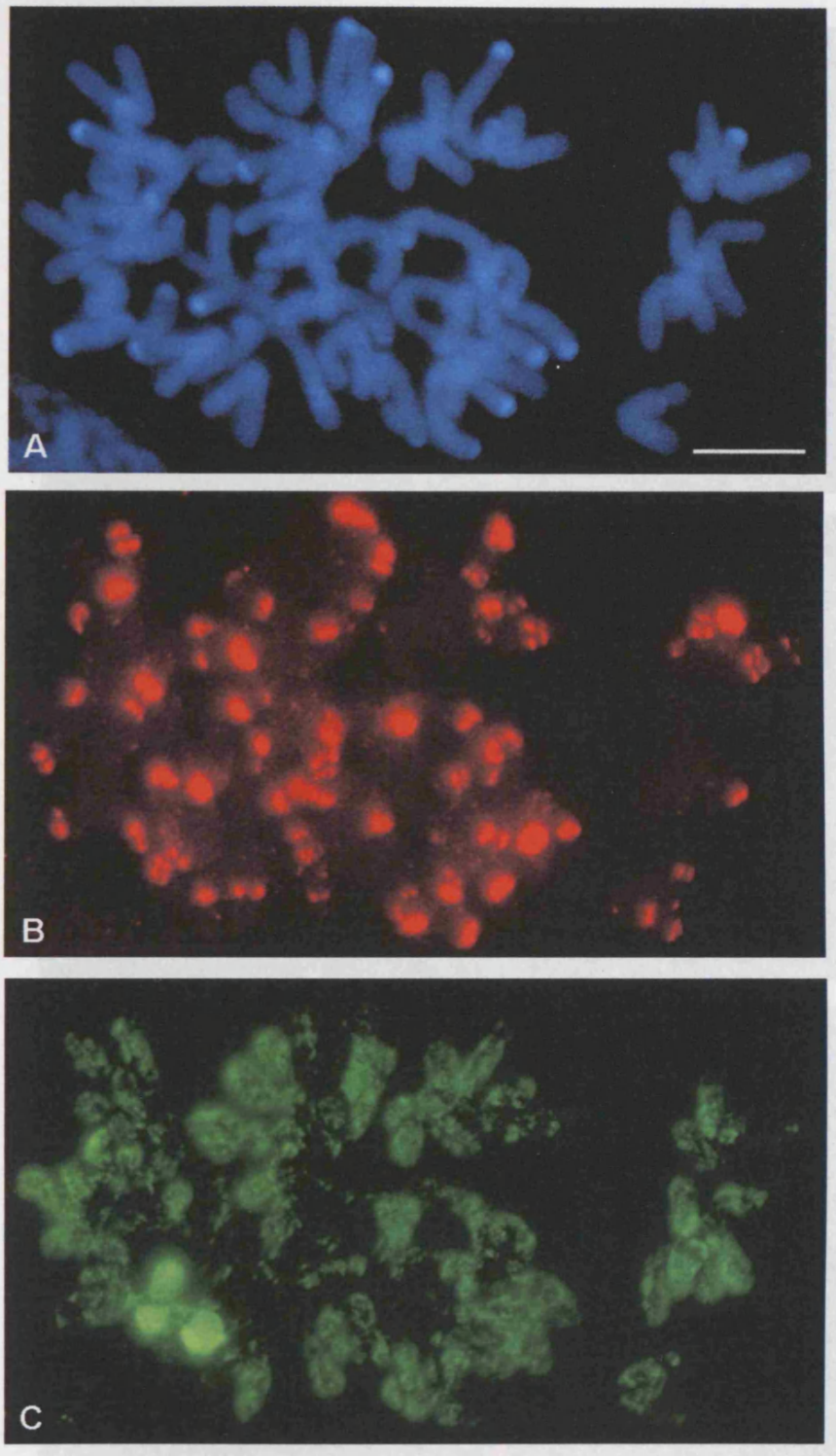


Figure 6.13. Root tip metaphase of 5-azacytidine treated *Triticale* 'Fidelio' individual 2 (ABR genome, 2n=42). A) DAPI stained chromosomes. B) *In situ* hybridization with biotin-labelled Tmono25/208!1212 probe detected with Alexa 594 red fluorescence. C) Immunostaining with anti-methylcytosine antibody detected with Alexa 488 green fluorescence. Scale bar represents 10 μm.

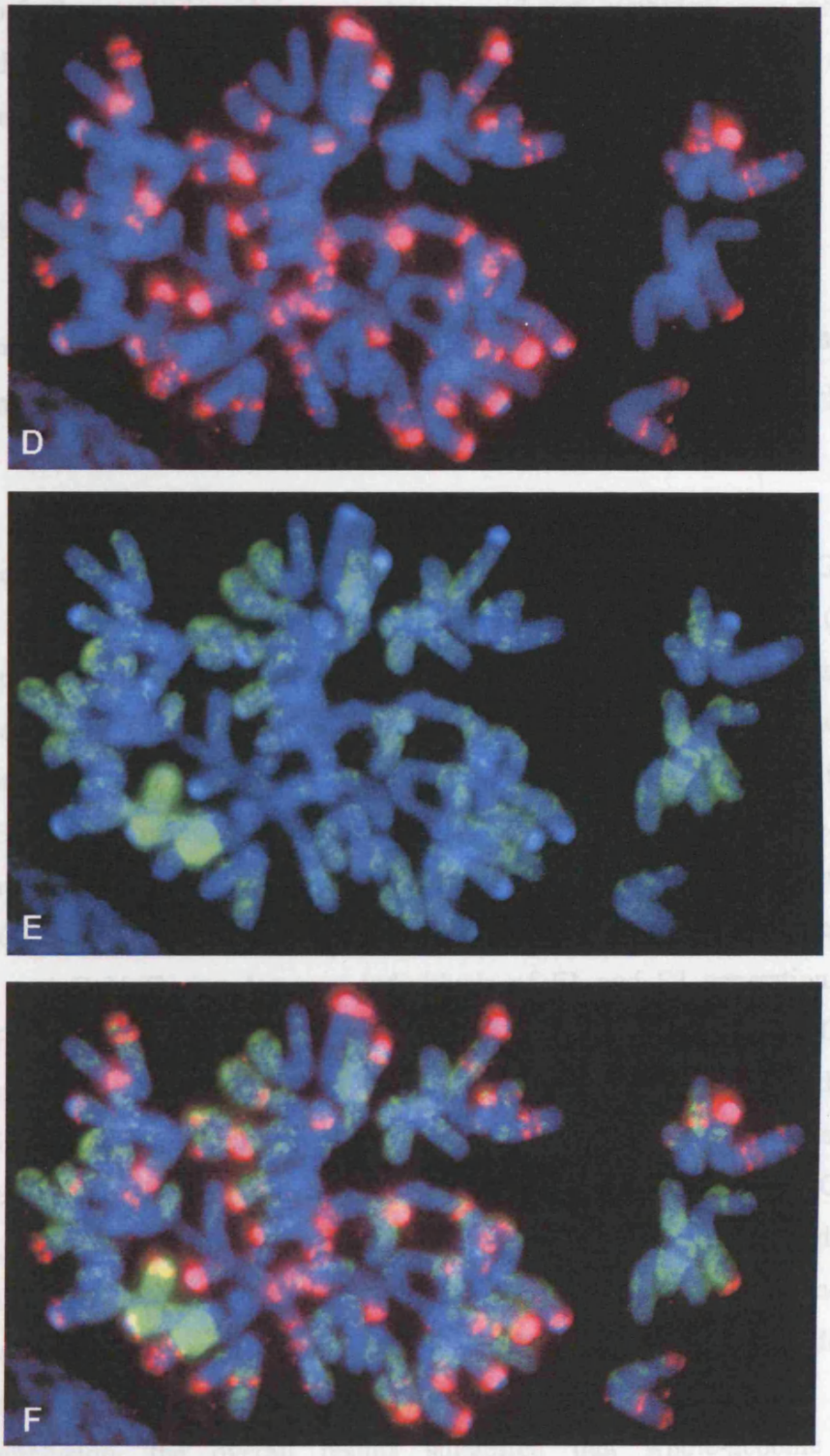


Figure 6. 13 continued. Root tip metaphase of 5-azacytidine treated *Triticale* 'Fidelio' individual 2 (ABR genome, 2n=42). D) Overlay of A and B images. E) Overlay of A and C images. F) Overlay of A, B and C images.

due to the very low number of metaphases no further *in situ* hybridization was performed. No further analysis with immunostaining with anti-methylcytosine antibody was performed on root tips from older seedlings as they died at the stage of first apical shoot emergence.

6. 5 Discussion

6. 5. 1 Crosses of *Triticale* lines: whole-genome DNA methylation and methylation patterns of 120-bp repeat unit family and 5S rDNA sequences

Differences in whole-genome DNA methylation at CpG, CpNpG and asymmetrical sites were observed between advanced lines ID-3013 and SEC776 (fig. 6. 1a) with comparable degree of variation as previously observed in some individuals of *Triticale* 'Fidelio' and 'Lamberto' (chapter 4, fig. 4. 11). Similarly, differences between the two parental lines were observed in methylation levels of 120-bp repeat unit family, parental line SEC776 showing lower methylation at CpG and CpNpG sites (fig. 6. 1b).

By analysing the progenies obtained by crossing *Triticale* advanced lines ID-3013 and SEC776 marked differences in whole-genome DNA methylation were found at CpG sites and CpNpG sites between individuals of F1 and F2 generation, some individuals (F1 individuals 1 and 3, F2 individuals 2 and 3) having CpG and CpNpG sites poorly methylated or not methylated at all and similarly not heavily methylated 120-bp repetitive sequence family (figs. 6. 2 and 6. 3). Conversely, individuals of F3 generation became more similar with higher levels of methylation at CpG and CpNpG sites. Finally, a consistent uniformly smeared signal after *Mcr*BC digestion indicating asymmetrical methylation was observed in all individuals of F1, F2 and F3 generation (fig. 6. 4b) as reported in parental lines SEC776 and ID-3013 and previously detected for 'Fidelio' and 'Lamberto' (chapter 4, figs. 4.12 and 4.13).

In conclusion, the overall results suggested that whole-genome DNA methylation and methylation patterns of the 120-bp repeat unit family were possibly affected by the cross between two advanced lines of *Triticale*, in different extent and timing, and that the three DNA methylation codes reacted differently to the crossing event. In fact, whole-genome DNA methylation and methylation patterns of the 120-bp repeat unit family showed differences between individuals of F1 and F2 generation,

some individuals being poorly methylated or not methylated at CpG and CpNpG sites, and became more similar in individuals of F3 generation that showed overall higher levels of methylation at CpG and CpNpG sites. In contrast, at asymmetrical sites no major changes between individuals of F1, F2 and F3 generation were detected in whole-genome DNA methylation as well as in methylation patterns of the 120-bp repeat unit family.

By comparing 5S rDNA sequences with 120-bp repetitive sequence family it was evident that different repetitive sequence families were differentially affected by the crossing event between two advanced lines of *Triticale*. In fact, in individuals of F1, F2 and F3 generation DNA methylation changes at CpG and CpNpG sites were not as marked in 5S rDNA sequences (fig. 6. 5a-b) than in 120-bp repetitive sequence family. Conversely, asymmetrical methylation showed variable smears in 5S rDNA sequences of individuals of F2 and F3 generations (fig. 6. 5c) suggesting a different organization of multiple overlapping half-sites of the form $(G/A)^{m}C$, whereas similar and uniform smears were detected for 120-bp repetitive sequence family in F1, F2 and F3 individuals. However, both repetitive sequence families appeared to be more asymmetrically methylated than the bulk of DNA.

As a possible consequence of DNA methylation changes affecting different repetitive sequence families of a genome (such as changes in methylation patterns at CpG and CpNpG sites in 120-bp family and 5S rDNA sequences as well as changes in asymmetrical methylation of 5S rDNA sequences) and thus producing changes in whole-genome DNA methylation at CpG, CpNpG and symmetrical sites, a trend from uniform to differentiated DNA methylation patterns has been found on chromosome spreads from F1 (table 6. 2 and fig. 6. 6) to F3 individuals (table 6. 2 and fig. 6. 7).

6. 5. 2 DNA re-methylation in *Triticale* 'Fidelio' after 5-azacytidine treatments

After 5-azacytidine treatments (at time T0) DNA methylation was not completely erased from chromosomes as DNA methylation signal ranged from 10% to 28.57% in metaphases of three different individuals (table 6. 3) and where present the signal was weak along chromosomes, suggesting that 5-azacytidine treatments even in high concentrations were not enough for a complete genome de-methylation. By investigating DNA methylation in metaphases from root tips fixed 16 hrs after the

treatment stopped (at time T1), when approximately one cell cycle could have occurred, re-methylation appeared to be a quick process: in fact one cell cycle was necessary to reach up to 60-80% of re-methylated metaphases but differences between individuals with variable percentage of metaphases already re-methylated were observed (table 6. 3). In fact, while in individual 1 any re-methylation process had begun, for individual three one cell cycle was necessary to remethylated 80% of its chromosome spreads (table 6. 3). It can be argued that differences in the extension of remethylation were due to the fact that different individuals were in different stages of cell cycle: possibly in individual 1 and 4 no cell cycle had started yet after 5-azacytidine treatment had stopped and individual 2 was slightly behind individual 3 in completing its cell cycle. However, it can be possible as well that the re-methylation process itself varied in its efficiency between different individuals.

No genome-specific de-methylation was observed after 5-azacytidine treatments by using genomic *in situ* hybridization (GISH.) in *Triticale* 'Lasko' (Castilho, Neves et al. 1999) and similarly no genome specific re-methylation was observed in *Triticale* 'Fidelio'. In fact, using *in situ* hybridization with genomic DNA of *S. cereale* 'Petkus Spring' and *Ae. speltoides* or with 120-bp repetitive sequence family clones re-methylation processes did not appeared genome specific as chromosomes from both rye and wheat genomes showed variable (weak and strong) DNA methylation signal in metaphases from root tips fixed at time T1 (figs. 6. 12 and 6. 13).

In slide preparations from root tips fixed at time T0 micronuclei and atypical genome intermixing in interphases were observed (figs. 6. 9 and 6. 10). Micronuclei and intermixed nuclei were already reported to be present in *Triticale* 'Lasko' after 5-azacytidine treatments (Castilho, Neves et al. 1999). In contrast, multiple chromosomal rearrangements involving different types of wheat-rye translocations (terminal, intercalary, centromeric and near centromaric) were not observed in 5-azacytidine treated *Triticale* 'Fidelio' as previously reported for 5-azacytidine treated *Triticale* 'Lasko' (Castilho, Neves et al. 1999). *Triticale* being a hybrid with a very recent evolutionary history, its different varieties could have different degree of genome stability and thus respond differently to events of genome destabilization such as 5-azacytidine treatments.

As 5-azacytidine treatments cancel the 'memory' of DNA methylation code and maintainance DNA methylation mechanisms are not likely to operate after the treatment is arrested (Santi, Norment et al. 1984), then *de novo* DNA methylation mechanisms are

candidate for re-methylation processes. In a recent review about different genetic pathways for establishing and maintaining DNA methylation in *A. thaliana*, *de novo* DNA methylation mechanisms depend on a particular class of DNA methyltransferases, the Domain Rearranged Methyltransferases (DRM) which seems to be guided by RNA Interference (Chan, Henderson et al. 2005). However, after 5-azacytine treatments it is questionable if *de novo* DNA methylation mechanisms are guided or random processes and if additional excision-repair processes occur in order to remove the modified bases and replace them with cytosines.

Despite the low number of metaphases analysed in 5-azacytidine treated *Triticale* 'Fidelio' individual 2 (fixed at time T1) it seemed that its DNA methylation patterns in most chromosome pairs of A and B genomes were similar of the ones of untreated *Triticale* 'Fidelio' individuals analysed in chapter 5 and resumed in fig. 5. 13. Similarly, the number of telomeres of rye genome not methylated was in agreement with previous observations reported for untreated *Triticale* 'Fidelio' individuals analysed in chapter 5. Overall re-methylation seemed to have occurred in both rye and wheat genomes and new DNA methylation patterns to be similar to the ones of untreated *Triticale* 'Fidelio' but cell division was arrested from fixation times T2 to T4. However, at time of fixation T5 cell cycle seemed to have started again and metaphases although few were encountered with differentiated DNA methylation signal. Neverthless, seedlings died at the stage of first apical shoot emergence.

Overall, from the results obtained it seemed likely the hypothesis that after 5-azacytidine treatments in *Triticale* 'Fidelio' re-methylation mechanisms happened in a random way so that DNA methylation patterns were similar but not exactly reproducing the DNA methylation patterns of untreated *Triticale* 'Fidelio' and this affected cell division and plant growth. The results obtained in *Triticale* 'Fidelio' after 5-azacytidine treatments are in agreement with previous data in rice, *Triticale*, flax, tobacco and *A. thaliana* showing that genome-wide demethylation through 5-azacytidine treatments, mutation or methyltransferase antisense constructs had effects on morphology, development and fertility (Finnegan, Peacock et al. 1996; Richards 1997; Finnegan, Genger et al. 1998). However, it cannot be excluded that the cytotoxicity of 5-azacytidine treatments had cytostatic effects on *Triticale* 'Fidelio' individuals (Plagemann, Behrens et al. 1978).

Chapter 7: General discussion

7. 1 A major fraction of *Triticeae* genomes: 120-bp repeat unit family

7. 1. 1 Diversity of 120-bp repeat unit family

The results obtained by PCR amplification, cloning and sequencing together with *in situ* hybridization experiments in many diploid and polyploid species of *Triticeae* and in *Avena sativa* demonstrated that 120-bp repeat unit family is widely distributed within *Triticeae* and *Aveneae* species, and forms a large component of their genomes in agreement with data previously reported for *S. cereale*, hexaploid wheat *T. aestivum*, diploid and polyploid *Hordeum* species, *Aegilops* species (Mukai, Nakara et al. 1993; Castilho and Heslop-Harrison 1995; Badaeva, Friebe et al. 1996; Taketa, Ando et al. 2000) and for species belonging to the *Avenae* (Katsiotis, Hagidimitriou et al. 1997).

The phylogenetic (Neighbour Joining) trees (see chapter 3, fig. 3. 5) showed no clear and well-supported groupings that could be related to genus, species or primers used, and branches were deep, such that 120-bp repeat unit family appeared to be an ancient component of *Triticeae* genomes with no single 120-bp sequence to be inherited by each genus or species from a single ancestor. The common ancestor of *Triticeae* tribe possibly had multiple sequences of 120-bp repeat unit family with a range of variation not unlike that seen within and between species today. In addition, homogenization events were unlikely to be operative in 120-bp repeat unit family sequences since their evolutionary divergence, and therefore no characteristic genome or species-specific variants developed during evolution of the extant *Triticeae* genomes.

This evolutionary model for the 120-bp repeat unit family differs from that for other tandemly repeated DNA sequences, such as the pAS1/dpTa1/AfaI/pHcKB6 repeat family (Rayburn and Gill 1986; Anamthawat-Jónsson and Heslop-Harrison 1993; Vershinin, Svitashev et al. 1994; Nagaki, Tsujimoto et al. 1995; Nagaki, Tsujimoto et al. 1998) and TaiI repeat family (Kishii, Nagaki et al. 1999; Kishii, Nagaki et al. 2001; Kishii and Tsujimoto 2002), where species-specific variants are recognized, and where sequence homogenization occurs within a species.

7. 1. 2 120-bp repeat unit family is part of the chromosome ends in *Triticeae* species

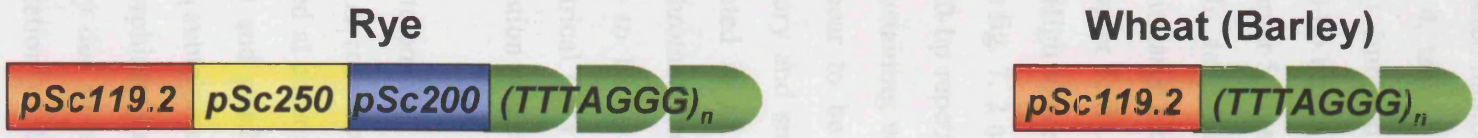
All eukaryotic chromosomes have been shown to contain subtelomeric DNA repeats adjacent to telomeres and these have been referred to as telomeric-associated sequences or TASs (Louis and Vershinin 2005). In plants, TASs are often structurally more variable and polymorphic than any other region of the genome even in species where they were described as not highly repeated, as in rice and maize for example (Burr, Burr et al. 1992; Ashikawa, Kurata et al. 1994). The overall results obtained in the present work showed that 120-bp repeat unit family has many characteristics of TASs: telomere-similar motifs were found (see chapter 3, fig. 3. 4), and among 87 repeat units from *Secale, Hordeum* and *Triticum/Aegilops* groups of species high levels of similarity within and between groups were found suggesting similarity levels only slightly higher within than between species (see chapter 3, table 3. 4) with no species-specific variants.

In different species, the *Triticeae* chromosome ends showed a different organization, from simple to more complex structures, and 120-bp repeat unit family had different locations, as observed in barley (*H. vulgare*), wheat (*T. aestivum* 'Chinese Spring') and rye (*S. cereale*) telomeres and shown in fig. 7. 1.

a) Barley and Wheat telomeres

In *H. vulgare* and *T. aestivum* 'Chinese Spring' 120-bp repeat unit family is directly adjacent to the telomeres (fig. 7. 1). In barley, previous authors in searches for telomeric-associated sequences isolated a highly repeated sequence with a monomer of about 118 bp in length and found it to be 71% homologous to pSc119.2 sequence (Belostotsky and Ananiev 1990; Kilian and Kleinhofs 1992). On chromosome spreads, it was shown to be located on 12 or 13 of the 14 chromosome ends depending on the barley variety analysed (Roder, Lapitan et al. 1993; Brandes, Roder et al. 1995) and to be hypervariable as found by polymorphisms in four different cultivars (Kilian, Kudrna et al. 1999). Similarly, a highly repeated telomeric-associated sequence with a monomer about 118 bp in length, was isolated in wheat, shown to be 88% homologous to pSc119.2 sequence and having polymorphisms between two different varieties (Mao, Devos et al. 1997), but there has been no detailed description of its location on chromosome spreads.

Chromosome ends organization



Chromosome ends DNA methylation

Rye

- Variable number of telomeres (from 12 to 16; mode=14) not methylated in R chromosomes
- 120-bp repeat unit family sequences at large sub-telomeric heterochromatic blocks partially methylated or not-methylated in most R chromosomes

Wheat

- Telomeres are fully methylated in A and D chromosomes
- Variable number of telomeres (from 10 to 20; mode=16) not methylated in B chromosomes
- 120-bp repeat unit family sequences fully methylated in A and D chromosomes
- 120-bp repeat unit family sequences at large sub-telomeric heterochromatic blocks partially methylated or not-methylated in some B chromosomes

Figure 7. 1. Organization and localization of 120-bp repeat unit family on chromosome ends of rye (*S. cereale*), wheat (*T. aestivum* 'Chinese Spring') and barley (*H. vulgare*) are presented according to Vershinin *et al.* (1995), Mao *et al.* (1997) and Kilian and Kleinhofs (1992). DNA methylation of chromosome ends and DNA methylation of 120-bp repeat unit family sequences in *S. cereale* (R genome), *T. monococcum* (A genome), *Ae. speltoides* (B genome) and *T. tauschii* (D genome) are listed.

From the *in situ* hybridization results obtained in the present work in wild *T. urartu* (A genome) and *T. tauschii* (D genome), 120-bp repeat unit family sequences showed telomeric locations but not on all chromosomes (see chapter 3, figs. 3. 9 and 3. 10). The low number of sites agreed with low copy number of 120-bp sequences per unreplicated haploid genome, in the range of 10^2 and 10^4 respectively (see fig. 7. 2 and chapter 4, table 4. 2) in cultivated *T. monococcum* (A genome) and *T. tauschii* (D genome). Immunostaining with anti-methylcytosine antibody indicated in both species that 120-bp repeat unit sequence family was fully methylated at its telomeric locations (see chapter 5, figs. 5. 1 and 5. 2).

In *Ae. speltoides* (B genome) 120-bp repeat unit family sequences were mainly present in large blocks at the end of some chromosomes with additional intercalary sites (see chapter 5, fig. 5. 3) in agreement with previous karyotypes (Badaeva, Friebe et al. 1996). High values of copy number per unreplicated haploid genome, in the range of 10^6 (see fig. 7. 2 and chapter 4, table 4. 2) and strong signals in Southern hybridization with 120-bp repetitive sequences (see chapter 4, fig. 4. 7) were reported. Interestingly, immunostaining with anti-methylcytosine antibody showed that 120-bp sequences did not appear to be fully methylated in all sites (see chapter 5, fig. 5. 3). Whereas intercalary and small telomeric bands made of 120-bp repeat unit family were fully methylated in all chromosomes, the large sub-telomeric heterochromatic blocks on some chromosome pairs were partially methylated or not methylated. These findings seemed to be in agreement with Southern hybridization results showing that both symmetrical CpG and CpNpG sites were poorly methylated, while only asymmetrical methylation was present (see chapter 4, fig. 4. 7).

b) Rye telomeres

A complex structure made of three different repetitive sequence families was observed at chromosome ends of *S. cereale* (fig. 7. 1). Two of these repetitive families, pSc200 and pSc250, have long monomers of 379 and 550 bp respectively and are strictly subtelomeric located (Vershinin, Schwarzacher et al. 1995). They showed polymorphisms between rye varieties by *in situ* hybridization on chromosome spreads possibly due to a rapid evolution (Alkhimova, Heslop-Harrison et al. 1999) and the organization of the arrays of monomers was detailed described in both repetitive sequence families by *in situ* hybridization on extended chromatin fibers (Alkhimova, 2004 #436).

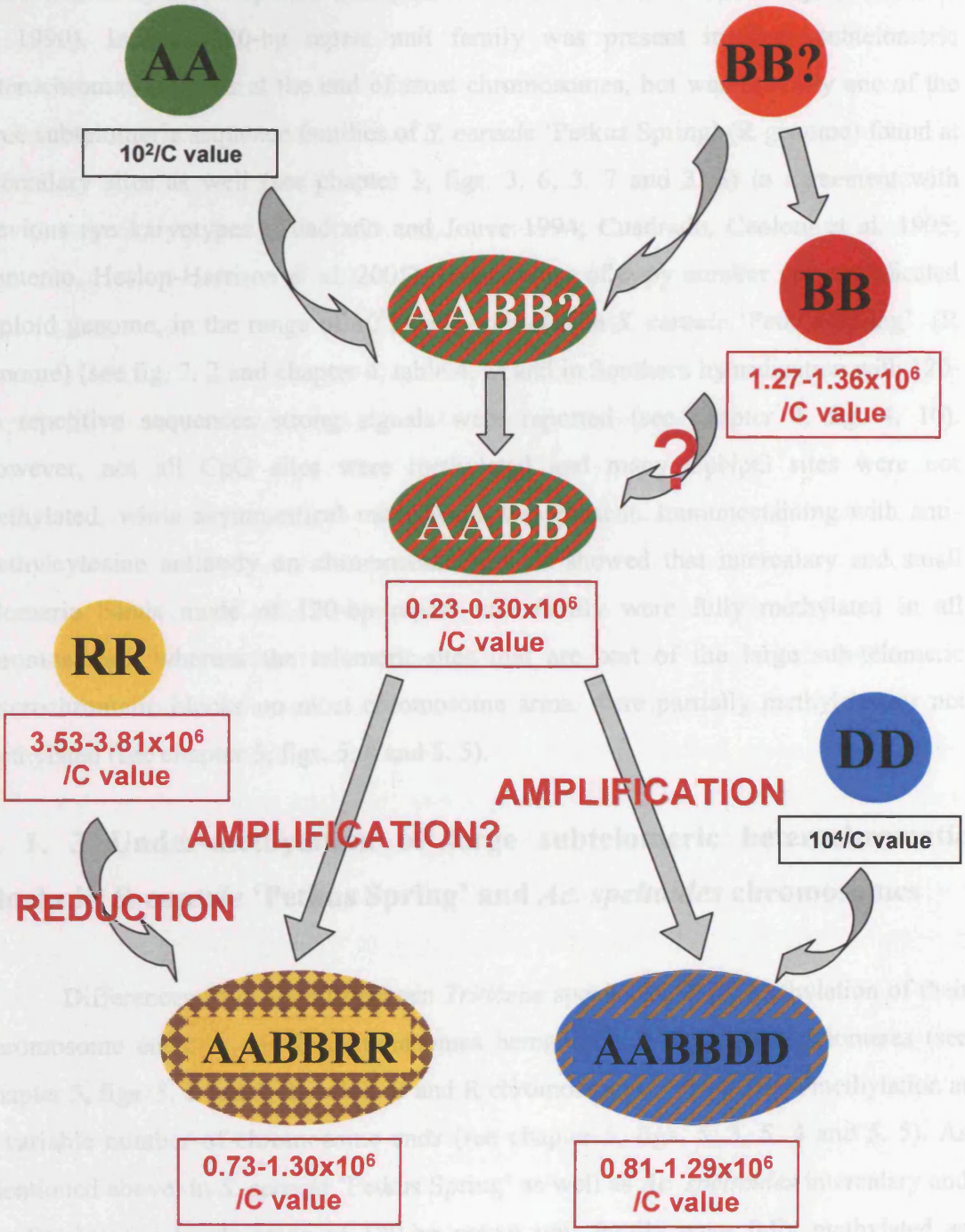


Figure. 7. 2. Copy number per haploid unreplicated genome of 120-bp repeat unit family sequences in diploid and polyploid *Triticeae* genomes. Reduction and amplification mechanisms of 120-bp repeat unit family sequences during hybridization events generating *Triticeae* allopolyploids are presented. *Triticeae* evolutionary tree is based on Ozkan *et al.* 2005, Salamini *et al.* 2002 and Wilson 1876.

The third one of these rye repetitive sequence families is the most proximal and characterized by a 118 bp monomer (Bedbrook, Jones et al. 1980; McIntyre, Pereira et al. 1990). In rye, 120-bp repeat unit family was present in large subtelomeric heterochromatic blocks at the end of most chromosomes, but was the only one of the three subtelomeric sequence families of *S. cereale* 'Petkus Spring' (R genome) found at intercalary sites as well (see chapter 3, figs. 3. 6, 3. 7 and 3. 8) in agreement with previous rye karyotypes (Cuadrado and Jouve 1994; Cuadrado, Ceoloni et al. 1995; Contento, Heslop-Harrison et al. 2005). High values of copy number per unreplicated haploid genome, in the range of 10^6, were estimated in *S. cereale* 'Petkus Spring' (R genome) (see fig. 7. 2 and chapter 4, table 4. 2) and in Southern hybridization with 120-bp repetitive sequences strong signals were reported (see chapter 4, fig. 4. 10). However, not all CpG sites were methylated and many CpNpG sites were not methylated, while asymmetrical methylation was present. Immunostaining with anti-methylcytosine antibody on chromosome spreads showed that intercalary and small telomeric bands made of 120-bp repeat unit family were fully methylated in all chromosomes, whereas the telomeric-sites that are part of the large sub-telomeric heterochromatic blocks on most chromosome arms, were partially methylated or not methylated (see chapter 5, figs. 5. 4 and 5. 5).

7. 1. 3 Under-methylation of large subtelomeric heterochromatic blocks in *S. cereale* 'Petkus Spring' and *Ae. speltoides* chromosomes

Differences were found between *Triticeae* species in DNA methylation of their chromosome ends: A and D chromosomes being fully methylated at telomeres (see chapter 5, figs. 5. 1 and 5. 2), while B and R chromosomes lacking DNA methylation at a variable number of chromosome ends (see chapter 5, figs. 5. 3, 5. 4 and 5. 5). As mentioned above, in *S. cereale* 'Petkus Spring' as well as *Ae. speltoides* intercalary and small telomeric bands made of 120-bp repeat unit family were fully methylated as generally predicted for highly repeated DNA sequences, from rice and tobacco for example that appear heavily methylated at CpG and CpNpG sites (Matyasek, Gazdova et al. 1997; Cheng, Stupar et al. 2001). In fact, it is probably necessary that these interstitial noncoding DNA sequences located amongst active chromatin maintain their silencing due to DNA methylation. Conversely, the large sub-telomeric heterochromatic blocks of 120-bp repeat unit family were partially methylated or not methylated in *S.*

cereale 'Petkus Spring' and *Ae. speltoides*. A possible explanation for that is that these large heterochromatic blocks are heavily compact regions, as previously shown in rye: in Feulgen stained nuclei where chromocenters were not decondensed (Anamthawat-Jónsson and Heslop-Harrison 1990) and by a very short synaptonemal complex and no recombination events (Albini and Schwarzacher 1992). Large heterochromatic blocks are so heavily compact, possibly due to histone modifications and/or other chromatin remodelling factors, that DNA methylation is not required for inhibiting transcription of non-coding DNA sequences. Alternatively, the anti-methylcytosine antibody cannot have access to its target sites as proposed by Frediani *et al.* (1996), which would indirectly confirm the tight compaction degree of heterochromatin at subterminal sites in chromosome ends of *Ae. speltoides* and *S. cereale* 'Petkus Spring'.

In addition, histone H3 methylation patterns were analysed in *S. cereale* 'Petkus Spring' and uniformly distributed signal on interphase nuclei and along chromosomes until their chromosome ends was found by immunostaining with anti di-methylated H3 histone at Lysine 9 antibody, as previously reported by Houben *et al.* (2003). The results gave evidence that in *S. cereale* 'Petkus Spring' chromosome ends were equally methylated at Lysine 9 of histone H3, while DNA methylation was present only in some telomeres but was lacking in others, overall demonstrating that methylation of histone H3 at Lysine 9 and DNA was not always coupled. Conversely, telomeres of cultivated *T. monococcum* and *T. tauschii* were fully DNA methylated and presumably methylated as well at histone H3 at Lysine 9 as suggested by Houben *et al.* (2003). In conclusion, in addition to previous data reporting that telomeric nucleosome size (160 bp) was shorter than bulk chromatin nucleosome size (175-185 bp) in *Triticeae* species (Vershinin and Heslop-Harrison 1998), here a species-specific and/or a chromosome-specific organization of telomeric and subtelomeric nucleosome is hypothesised, as a consequence of variable degrees of heterochromatin condensation at chromosome ends due to variable DNA and histone methylation patterns.

7. 2 120-bp repeat unit family as a tool for investigating genome interactions in *Triticeae* polyploids

7. 2. 1 Genomic and epigenetic changes in 120-bp repeat unit family in *Triticeae* polyploids

120-bp repeat unit family being a widespread and old component of *Triticeae* genomes, hence was used for investigating genome interactions in *Triticeae* polyploid species resulting from the combinations of *T. monococcum* (A genome), *Ae. speltoides* (B genome), *T. tauschii* (D genome) and *S. cereale* (R genome).

Changes in copy number of 120-bp repeat unit family were hypothesised to happen as a result of hybridization events generating polyploids across *Triticeae* evolutionary history as shown in fig. 7. 2, suggesting that 120-bp repeat unit family sequences needed to be adjusted through events of reduction and amplification. For example, by comparing the extant genomes of *Ae. speltoides* (B genome) and *T. turgidum* ssp. *durum* (AB genome) a reduction of B genome-origin 120-bp repeat unit family sequences could be hypothesised as a consequence of polyplodization assuming that *Ae. speltoides* was the B genome donor of *T. turgidum* ssp. *durum* and that A genome-origin 120-bp repeat unit family sequences were not relevant in terms of copy number (fig. 7. 2). In addition, events of reduction and/or amplification were estimated to happen when *S. cereale* (R genome) and *T. turgidum* ssp. *durum* (AB genome) contributed to *Triticale* (ABR genome). Finally, an amplification of AB genome-origin 120-bp repeat unit family sequences in *T. aestivum* 'Chinese Spring' (ABD genome) was reported assuming that D genome-origin 120-bp repeat unit family sequences were not relevant in terms of copy number (fig. 7. 2).

Increase and decrease in DNA methylation of 120-bp repeat unit family sequences were hypothesised to happen by comparing natural *Triticeae* polyploids with their parental species. In fact, an increase in CpG methylation of B genome-origin 120-bp repeat unit family sequences was observed by comparing the extant genomes of *Ae. speltoides* (B genome) and *T. turgidum* ssp. *durum* (AB genome) (fig. 7. 3) and a decrease in CpNpG methylation of R and/or AB genome-origin 120-bp repeat unit family sequences by comparing the extant genomes of *S. cereale* (R genome) and *T. turgidum* ssp. *durum* (AB genome) with *Triticale* (ABR genome) (fig. 7. 3).

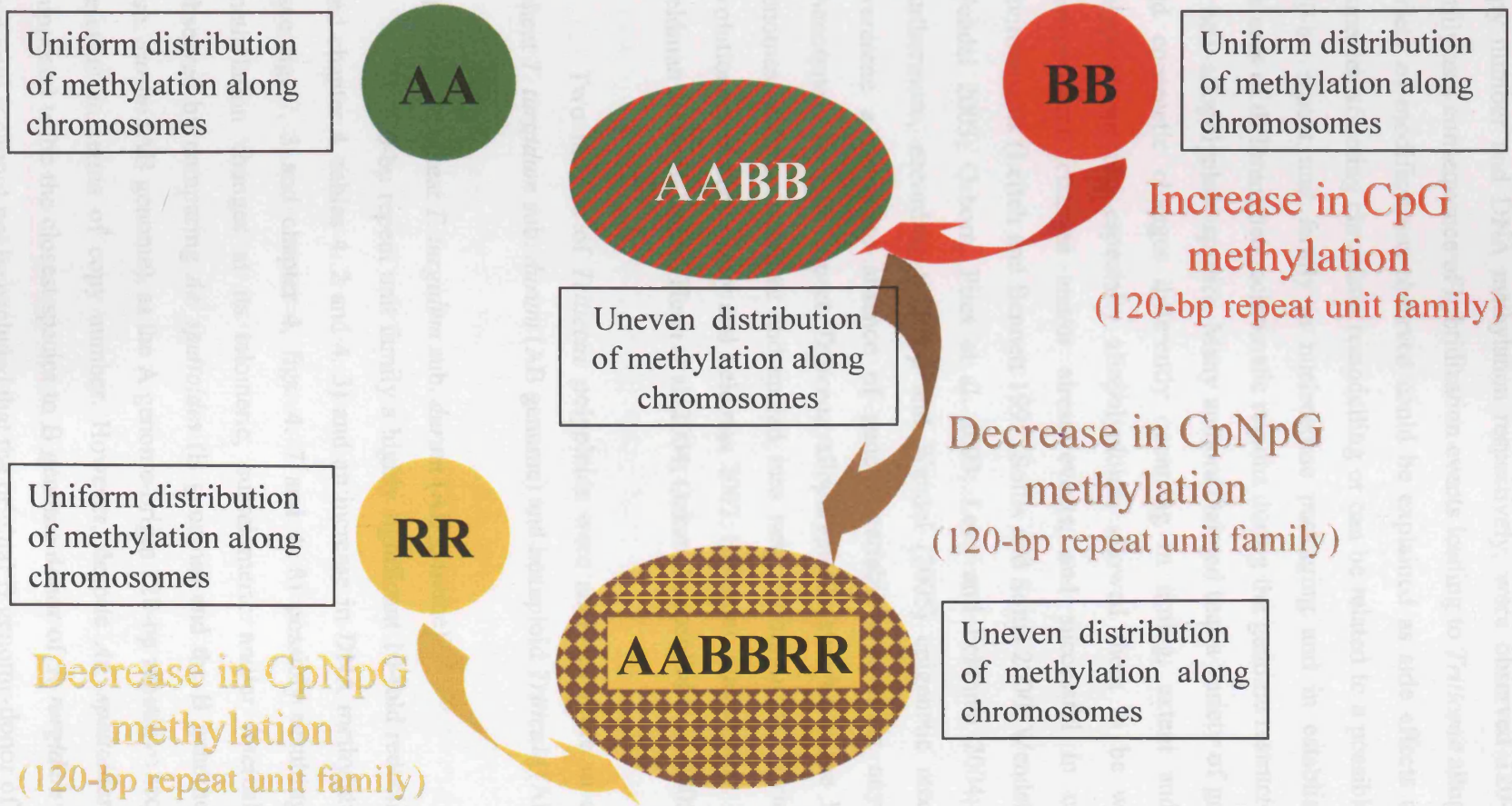


Figure. 7. 3. Model of whole-genome DNA methylation patterns and DNA methylation changes proposed for 120-bp repeat unit family in *T. turgidum* ssp. *durum* (AB genome) and *Triticale* 'Fidelio' (ABR genome). Increase in CpG methylation (red arrow) and decrease in CpNpG DNA methylation (yellow and brown arrows) contribute to differentiate DNA methylation patterns on chromosome spreads of *T. turgidum* ssp. *durum* (AB genome) and *Triticale* 'Fidelio' (ABR genome).

In conclusion, a complex picture of genomic and epigenetic changes, affecting copy number and DNA methylation respectively, were observed 120-bp repeat unit family as a consequence of hybridization events leading to *Triticeae* allopolyploids. The variety of modifications observed could be explained as side effects of higher order processes affecting chromatin remodelling or can be related to a possible active role of 120-bp repeat unit family in nucleosome packaging and in establishing the local features of different heterochromatic regions during the genome restructuring processes in new allopolyploid species. Many authors claimed that a variety of genomic, genetic and epigenetic changes differently occurring in timing, extent and frequency in polyploid species, especially allopolyploids, allowed them to be well adapted to environmental changes and/or stress events, and successful in colonizing new environments (Leitch and Bennett 1997; Soltis and Soltis 2000; Wendel 2000; Liu and Wendel 2003; Osborn, Pires et al. 2003; Levy and Feldman 2004; Comai 2005). Furthermore, according to Rapp and Wendel (2005) epigenetic mechanisms could overcome a complete absence of genetic variability resulting anyways in novel phenotypes. In this respect *Triticeae* allopolyploids showed to be highly dynamic genomes, able to colonize and spread into new environments, having a successful evolutionary history (Levy and Feldman 2002; Salamini, Ozkan et al. 2002; Levy and Feldman 2004; Salamini, Heun et al. 2004; Ozkan, Brandolini et al. 2005).

Two models of *Triticeae* polyploids were analysed more in detail: tetraploid wheat *T. turgidum* sub. *durum* (AB genome) and hexaploid *Triticale* (ABR genome).

a) tetraploid wheat *T. turgidum* sub. *durum* (AB genome)

In 120-bp repeat unit family a highly significant 10-fold reduction (see fig. 7. 2 and chapter 4, tables 4. 2 and 4. 3) and an increase in DNA methylation at CpG sites (see fig. 7. 3 and chapter 4, figs. 4. 7 and 4. 8) possibly resulting in local DNA methylation changes at its telomeric, subtelomeric and/or intercalary sites, were observed by comparing *Ae. speltoides* (B genome) and the B genome of *T. turgidum* ssp. *durum* (AB genome), as the A genome-origin 120-bp repetitive sequences were not relevant in terms of copy number. However, despite *Ae. speltoides* (B genome) is supposed to be the closest species to B genome donor of *T. turgidum* sub. *durum* (AB genome) it cannot not be excluded that the original B genome-donor of *T. turgidum* ssp. *durum* (AB genome) differed from the extant species *Ae. speltoides* (B genome) and

thus comparisons between the two species are not exactly reproducing the evolutionary history of *T. turgidum* sub. *durum* (AB genome). In fact, some chromosomal rearrangements were hypothesised to happen in *Ae. speltoides* (B genome) after the polyploidization event originating *T. turgidum* sub. *durum* (AB genome) (Badaeva, Friebe et al. 1996; Belyayev, Raskina et al. 2000). Similarly, *in situ* hybridization with 120-bp repetitive sequences showed some differences between the B genome-karyotype in *T. turgidum* sub. *durum* (AB genome) (see chapter 3, figs. 3.11 and 3.12 and chapter 5, fig. 5. 6) and the karyotype of *Ae. speltoides* (B genome) (see chapter 5, fig. 5. 3).

b) hexaploid *Triticale* 'Fidelio' (ABR genome)

A highly significant reduction of R genome-origin 120-bp repetitive sequences (see fig. 7. 2 and chapter 4, tables 4. 2 and 4. 3) was estimated to happen in *Triticale* 'Fidelio' (ABR genome) and/or a possible amplification of AB genome-origin 120-bp repetitive sequences (see fig. 7. 2 and chapter 4, tables 4. 2 and 4. 3). Furthermore, a decrease in DNA methylation at CpNpG sites has been reported in *Triticale* 'Fidelio' (ABR genome) in comparison with its parental species (see fig. 7. 3 and chapter 4, figs. 4. 8, 4. 10 and 4. 12), mainly affecting R genome-origin rather than in AB genome-origin in 120-bp repeat unit family sequences as observed on chromosome spreads (see chapter 5, fig. 5. 9). In fact, despite a higher number of A chromosomes showing enhanced and reduced DNA methylation signal as well as additional intercalary sites and satellites of most B chromosomes being methylated in *Triticale* 'Fidelio' (ABR genome) in comparison with *T. turgidum* sub. *durum* (AB genome), the A and B-origin 120-bp repeat unit family sequences did not show major differences in their DNA methylation status (see chapter 5, figs. 5. 7 and 5.10). It is likely that any decrease in DNA methylation at CpNpG sites affects B-origin 120-bp repetitive sequences at telomeric and/or intercalary sites in very little extent and thus was not visible as a major change in DNA methylation patterns on chromosomes spreads, whereas A genome-origin 120-bp repetitive sequences were not relevant in terms of copy number and thus in terms of DNA methylation levels. Conversely, in *Triticale* 'Fidelio' (ABR genome) patchy DNA methylation signal was present along euchromatin of all R chromosomes (see chapter 5, fig. 5. 9) compared to the uniformly distributed DNA methylation present in *S. cereale* 'Petkus Spring' (R genome) (see chapter 5, fig. 5. 4) probably affecting DNA methylation levels of 120-bp repetitive sequences at many intercalary sites. In addition, in *Triticale* 'Fidelio' (ABR genome) higher number of telomeres

appeared methylated than in *S. cereale* 'Petkus Spring' (R genome) and thus possibly the 120-bp repetitive sequences present at these locations (see chapter 5, figs. 5. 4 and 5. 9). In conclusion, a wide redistribution of DNA methylation patterns was observed in 120-bp repeat unit family of R chromosomes in *Triticale* 'Fidelio' (ABR genome) in comparison with *S. cereale* 'Petkus Spring' (R genome), overall resulting in a decrease in DNA methylation at CpNpG sites.

7. 2. 2 Genome-wide DNA methylation in diploid and polyploid *Triticeae* genomes and the role of repetitive sequences

In *T. monococcum* (A genome), *Ae. speltoides* (B genome), *T. tauschii* (D genome) and *S. cereale* (R genome) uniformly distributed whole-genome cytosine methylation along the euchromatin of all fourteen chromosomes including the centromeric regions was observed (see chapter 5, table 5. 1). Conversely, all allopolyploid species investigated in the present work: *T. turgidum* sub. *durum* (AB genome), *T. aestivum* 'Chinese Spring' (ABD genome) and *Triticale* 'Fidelio' (ABR genome), have more unevenly distributed whole-genome DNA methylation patterns with enhanced and reduced DNA methylation signal in different chromosomes pairs and/or different chromosome regions (see chapter 5, table 5. 2). Previous data on *Triticale* 'Lasko' and other species with large genomes such as *V. faba* and *A. cepa* showed similar differentiated DNA methylation patterns (Castiglione, Giraldi et al. 1995; Frediani, Giraldi et al. 1996; Castilho, Neves et al. 1999).

How have differentiated methylation patterns originated in *Triticeae* allopolyploids? Examples of DNA methylation repatterning following an allopolyplodization event were previously reported in literature in newly synthesised *Triticeae* allopolyploids and *Triticale* hybrids affecting mobile DNA elements and genes (Liu, Vega et al. 1998; Shaked, Kashkush et al. 2001; Kashkush, Feldman et al. 2002; Han, Fedak et al. 2003; Ma, Fang et al. 2004). In the present work methylation changes at CpG and CpNpG sites of 120-bp repeat unit family sequences were reported in *T. turgidum* ssp. *durum* (AB genome) and *Triticale* 'Fidelio' (ABR genome) respectively (see chapter 4, figs. 4. 8 and 4. 12) together with species-specific DNA methylation levels and patterns of 5S rDNA sequences in diploid and polyploid *Triticeae* species (see chapter 4, figs. 4. 7, 4. 8, 4. 9 and 4. 10) and variable DNA methylation patterns of 5S rDNA sequences in *Triticale* varieties (see chapter 4, figs. 4.

12 and 4. 13). It has been reported in literature that DNA methylation changes could be originated by direct and indirect causes. For example, aberrant pairing between repetitive sequences is a direct cause of DNA methylation (Assaad, Tucker et al. 1993; Liu and Wendel 2003) and RNA Interference indirectly lead to DNA methylation in presence of inverted repeats as reported for the *Arabidopsis PAI* gene family (Bender 2004). Interestingly, it has been noticed in newly synthesised *Arabidopsis* allopolyploids that DNA methylation changes did not affect the overall hyper- or hypomethylation of CG methylation, suggesting only a relocation of CG methylation after a polyploidization event (Madlung, Masuelli et al. 2002).

By using the wheat allopolyploid complex and *Triticale* hybrids as a model of study it has been shown in the present work that each repetitive sequence family, characterized by species-specific methylation patterns at CpG, CpNpG and asymmetrical sites, contributes to species-specific whole-genome DNA methylation patterns in diploid *Triticeae* species and to differentiate whole-genome DNA methylation patterns in allopolyploid *Triticeae* species, relocating the overall DNA methylation through local increase and decrease as a result of structural and functional differences in different chromatin regions. Examples of increase in CpG methylation and decrease in CpNpG methylation of 120-bp repeat unit family sequences contributing to differentiate whole-genome DNA methylation patterns in *T. turgidum* ssp. *durum* (AB genome) and *Triticale* 'Fidelio' (ABR genome) respectively are given in fig. 7. 3. In conclusion, diploid *Triticeae* species with uniformly distributed DNA methylation patterns contributed to the larger genomes of polyploid *Triticeae* species and underwent DNA methylation changes, affecting largely the non-coding fraction of their genomes (repetitive sequences as well as mobile DNA elements) and resulting in differentiated DNA methylation patterns, in the attempt to coordinate the heterochromatin of different genome donors in a new allopolyploid species. The term "heteromes" has been used for referring to the heterochromatic portions of the different genomes that need to be coordinate in new allopolyploid species (Comai, Madlung et al. 2003). Recently, many authors speculated about the links and interdependence between DNA methylation, histone modifications, chromatin remodelling processes and RNA Interference occurring in a new allopolyploid species in order to restructure heterochromatic and euchromatic regions and regain genome stability after the hybridization event (Liu and Wendel 2003; Osborn, Pires et al. 2003; Madlung and Comai 2004).

7. 3 Cytosine methylation as instable epigenetic mark of heterochromatin in *Triticale* allopolyploids

7. 3. 1 Differences between individuals of *Triticale* 'Fidelio' and 'Lamberto' in whole-genome DNA methylation, 120-bp repeat unit family and 5S rDNA methylation patterns

By investigating whole-genome DNA methylation and methylation patterns of 120-bp repeat unit family of *Triticale* 'Fidelio' and 'Lamberto' (ABR genomes) major differences between individuals were observed at CpG and CpNpG sites, some individuals being weakly methylated or not methylated (see chapter 4, fig. 4. 11). Similarly, variable DNA methylation levels were observed within and between *Triticale* 'Fidelio' and 'Lamberto' (ABR genome) at CpG sites and at CpNpG sites of 5S rDNA repetitive sequences (see chapter 4, figs. 4. 12 and 4. 13). Finally, by immunostaining with anti-methylcytosine antibody on chromosome spreads of hexaploid *Triticale* 'Fidelio' differences were detected within individuals and between homologues in DNA methylation patterns of A, B and R chromosomes (see chapter 5, fig. 5. 9 and 5.10). In conclusion, individual variant of DNA methylation patterns and/or chromosome-specific methylation patterns arose possibly a consequence of genome instability in *Triticale* allopolyploids, hybrids of much recent origin than established allopolyploids species *T. turgidum* ssp. *durum* (AB genome), and *T. aestivum* (ABD genome).

After immunostaining with anti di-methylated H3 histone at Lysine 9 antibody in *Triticale* 'Fidelio' a uniformly distributed signal on interphase nuclei and along chromosomes till their chromosome ends was found (see chapter 5, fig. 5. 11). Houben *et al.* (2003) proposed that uniformly distributed Met(K9)H3 signals along chromosomes are due to the high amount of repetitive sequences not only present in constitutive heterochromatin but also interspersed along euchromatic regions, whereas only in species with very small genome sizes (<500Mb) and a very low repetitive sequence content, such as *A. thaliana*, Met(K9)H3 signal is present in DAPI-stained heterochromatin but not in the rest of genome. In conclusion, it seemed that DNA and histone H3 methylation patterns were differently affected by genome composition in *Triticale* allopolyploids: DNA methylation patterns were more variable and presented

enhanced and reduced signals according to nature of chromosomal regions, while no variability nor differentiation of histone H3 methylation patterns was observed. Finally, differences in the organization of *Triticale* 'Fidelio' telomeres were observed as methylation was rather equally distributed at histones H3 Lysine 9, but not at DNA level with some telomeres being DNA methylation free.

7. 3. 2 Two cases of alteration of DNA methylation patterns in *Triticale* and the "memory" of DNA methylation code

Triticale being a hybrid with a recent evolutionary history and not an established species, the careful balance between its different genomes can be easily altered. Two different examples of alteration of DNA methylation patterns were investigated in *Triticale* allopolyploids, by crossing two advanced *Triticale* lines and by using 5-azacytidine treatments on *Triticale* 'Fidelio'.

Greater differences than previously observed in established *Triticale* 'Fidelio' and 'Lamberto' (ABR genomes) were observed between in F1, F2 and F3 individuals obtained by crossing two advanced *Triticale* lines. DNA methylation changes were observed in a different extent in whole-genome DNA methylation and methylation patterns of two different repetitive DNA fractions, 120-bp repeat unit family and 5S rDNA sequences, affecting differently the three DNA methylation codes (see chapter 6, figs. 6. 2, 6. 3, 6. 4 and 6. 5). As a result, high variability and instability of DNA methylation patterns were observed on chromosomes spreads of individuals from F1, F2 and F3 generation after immunostaining with anti-methylcytosine antibody (see chapter 6, table 6. 2). A trend from a uniformly distributed DNA methylation to a more differentiated DNA methylation pattern from F1 to F3 individuals was observed (see chapter 6, figs. 6. 6 and 6. 7). The alteration of DNA methylation patterns could have been produced by the crossing event, maybe indirectly as a consequence of higher order chromatin alterations. According to Liu and Wendel (2003) chromatin remodelling processes are altered by an allopolyploidization event leading to histone modifications and then DNA methylation changes that represent side effects of higher order processes, as DNA methyltransferases do not have any sequence specificity. Similarly, Madlung and Comai (2004) proposed that as a consequence of stress events, RNA Interference is involved in chromatin remodelling leading to genome restructuring. However, despite the alteration of DNA methylation patterns in the early generations obtained by crossing

two advanced *Triticale* lines the "memory" of DNA methylation code was quickly re-established with no effect on plant growth and development in F1, F2 and F3 individuals.

Conversely, 5-azacytidine treatments directly affect DNA methylation in *Triticale* 'Fidelio' roots. After 5-azacytidine treatments, it was observed that DNA methylation signal was greatly reduced but not completely erased from chromosomes in a low percentage of metaphases (10-30%) (see chapter 6, table 6. 3). Re-methylation appeared to be quick process as DNA methylation signal was present after only one cell cycle in most metaphases (60-80%) (see chapter 6, table 6. 3), through no genome specific re-methylation processes, with new DNA methylation pattern similar to the ones of untreated *Triticale* 'Fidelio' (see chapter 6, figs. 6. 12 and 6. 13). However, re-methylation mechanisms seemed not to be able to recover completely the 'memory' of DNA methylation code, and thus in the second example the DNA methylation alteration becomes established and affects plant growth and development. The results obtained in *Triticale* 'Fidelio' after 5-azacytidine treatments are in agreement with previous data obtained in other plant species on genome-wide demethylation and its effects on morphology, development and fertility. They overall suggest that plants do not reset their global patterns of DNA methylation during development, in contrast to mammals where demethylation and re-methylation mechanisms during development produce cell-type-specific DNA methylation patterns and contribute to genomic imprinting (Reik, Dean et al. 2001; Surani 2001).

7. 4 Future prospects

By using the wheat allopolyploid complex and *Triticale* hybrids as a model of study and by focusing on a major repetitive DNA fraction of *Triticeae* genomes, the 120-bp repeat unit family, some consequences of polyploidization events on heterochromatin have been presented in the present work.

By comparing natural diploid and polyploid *Triticeae* species, genomic and epigenetic changes have been described to happen in 120-bp repeat unit family, such as changes in copy number and DNA methylation patterns, possibly as a result of genome restructuring mechanisms in polyploid *Triticeae* species. The variety of modifications observed suggested that repetitive DNA sequences, such as 120-bp repeat unit family investigated here, could have either an active role in nucleosome packaging and in

establishing the local features of different heterochromatin regions or could be affected by side effects of higher order processes in polyploids. Future investigations can contribute to solve these open questions by analysing the link between DNA methylation and RNA Interference. In fact, RNA Interference has been recently proposed as a key mechanism for silencing DNA repeats and mobile elements and in general for maintenance of heterochromatic regions (Madlung and Comai 2004; Chan, Henderson et al. 2005).

In the present work the analyses carried out on *Triticale* hybrids revealed DNA methylation to be an instable epigenetic mark of heterochromatin, as individuals variant of DNA methylation patterns and/or chromosome-specific methylation patterns were described in establish *Triticale* varieties, as well as alterations of DNA methylation patterns when crossing two advanced *Triticale* lines. Overall, the results obtained suggested genome instability in these human-made hybrids of recent origin where the careful balance between different genomes can be easily altered. An open question concerns the role of chromatin remodelling processes in affecting DNA methylation and producing alterations of DNA methylation patterns in hybrids, as it was recently suggested by Osborn *et al.* (2003) and Liu and Wendel (2003).

The results obtained in the present work also showed that 120-bp repeat unit family is directly adjacent to telomeres in some diploid wheat species and fully methylated at these locations, but it is located in subtelomeric regions of *S. cereale* and *Ae. speltoides* where it is only partially or not methylated. Future investigations could explain why large subtelomeric heterochromatic blocks can differ in DNA methylation levels, while being equally methylated at Lysine 9 of histone H3. In addition, future work could clarify if a chromosome-specific organization of telomeric and subtelomeric regions is present in some plant species, due possibly to variable DNA and/or histone methylation patterns affecting nucleosomal organization and heterochromatin condensation.

Bibliography

(2004). "Finishing the euchromatic sequence of the human genome." Nature **431**(7011): 931-45.

Adams, K. L., R. Cronn, et al. (2003). "Genes duplicated by polyploidy show unequal contributions to the transcriptome and organ-specific reciprocal silencing." Proc Natl Acad Sci U S A **100**(8): 4649-54.

Adams, K. L. and J. F. Wendel (2004). " Exploring the genomic mysteries of polyploidy in cotton." Biological Journal of the Linnean Society **82**: 573 - 581.

Adams, K. L. and J. F. Wendel (2005). "Polyploidy and genome evolution in plants." Curr Opin Plant Biol **8**(2): 135-41.

Adams, S. P., T. P. Hartman, et al. (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 Biol Sci **268**(1476): 1541-6.

Albini, S. M. and T. Schwarzacher (1992). "In situ localization of two repetitive DNA sequences to surface-spread pachytene chromosomes of rye." Genome **35**: 551-559.

Alkhimova, A. G., J. S. Heslop-Harrison, et al. (1999). "Rye chromosome variability in wheat-rye addition and substitution lines." Chromosome Res **7**(3): 205-12.

Alleman, M. and J. Doctor (2000). "Genomic imprinting in plants: observations and evolutionary implications." Plant Mol Biol **43**(2-3): 147-61.

Amado, L., R. Abranches, et al. (1997). "Development-dependent inheritance of 5-azacytidine-induced epimutations in triticale: analysis of rDNA expression patterns." Chromosome Res **5**(7): 445-50.

Anamthawat-Jónsson, K. and J. S. Heslop-Harrison (1990). "Centromeres, telomeres and chromatin in the interphase nucleus of cereals." Caryologia **43**: 205-213.

Anamthawat-Jónsson, K. and J. S. Heslop-Harrison (1993). " Isolation and characterization of genome-specific DNA sequences in Triticeae species." Mol Gen Genet **240**: 151-158.

Ananiev, E. V., R. L. Phillips, et al. (2000). "Complex structure of knobs and centromeric regions in maize chromosomes." Tsitol Genet **34**(2): 11-5.

Ashikawa, I., N. Kurata, et al. (1994). "Cloning and mapping of telomere-associated sequences from rice." DNA Res **1**(2): 67-76.

Assaad, F. F., K. L. Tucker, et al. (1993). "Epigenetic repeat-induced gene silencing (RIGS) in Arabidopsis." Plant Mol Biol **22**(6): 1067-85.

Aufsatz, W., M. F. Mette, et al. (2002). "RNA-directed DNA methylation in Arabidopsis." Proc Natl Acad Sci U S A **99 Suppl 4**: 16499-506.

Axelsson, T., C. M. Bowman, et al. (2000). "Amphidiploid Brassica juncea contains conserved progenitor genomes." Genome **43**(4): 679-88.

Badaeva, E. D., B. Friebe, et al. (1996). "Genome differentiation in Aegilops. 1. Distribution of highly repetitive DNA sequences on chromosomes of diploid species." Genome **39**: 293-306.

Barnes, S. R., A. M. James, et al. (1985). "The organisation, nucleotide sequence and chromosomal distribution of a satellite DNA from Allium cepa." Chromosoma **92**: 184-192.

Bartee, L., F. Malagnac, et al. (2001). "Arabidopsis cmt3 chromomethylase mutations block non-CG methylation and silencing of an endogenous gene." Genes Dev **15**(14): 1753-8.

Baumel, A., M. L. Ainouche, et al. (2001). "Molecular investigations in populations of Spartina anglica C.E. Hubbard (Poaceae) invading coastal Brittany (France)." Mol Ecol **10**(7): 1689-701.

Bedbrook, J. R., J. Jones, et al. (1980). "A molecular description of telomeric heterochromatin in Secale species." Cell **19**(2): 545-60.

Belostotsky, D. A. and E. V. Ananiev (1990). "Characterization of relic DNA from barley genome." Theor. Appl. Genet. **80**: 374-380.

Belyayev, A., O. Raskina, et al. (2000). "Coevolution of A and B genomes in allotetraploid Triticum dicoccoides." Genome **43**(6): 1021-6.

Bender, J. (2001). "A vicious cycle: RNA silencing and DNA methylation in plants." Cell **106**(2): 129-32.

Bender, J. (2004). "DNA methylation and epigenetics." Annu Rev Plant Biol **55**: 41-68.

Bennett, M. D. and I. J. Leitch (2004). "Plant DNA C-values Database (release 3.0, December 2004)." http://www.rbgkew.org.uk/cval/homepage.html.

Bennett, M. D. and J. B. Smith (1991). "Nuclear DNA Amounts in Angiosperms." Philos. Trans. R. Soc. London Ser. B **334**: 309-45.

Bennetzen, J. L. (2002). "Mechanisms and rates of genome expansion and contraction in flowering plants." Genetica **115**(1): 29-36.

Bennetzen, J. L. and E. A. Kellogg (1997). "Do Plants Have a One-Way Ticket to Genomic Obesity?" Plant Cell **9**(9): 1509-1514.

Bennetzen, J. L., J. Ma, et al. (2005). "Mechanisms of recent genome size variation in flowering plants." Ann Bot (Lond) **95**(1): 127-32.

Bird, A. (1997). "Does DNA methylation control transposition of selfish elements in the germline?" Trends Genet **13**(12): 469-72.

Bird, A. (2002). "DNA methylation patterns and epigenetic memory." Genes Dev **16**(1): 6-21.

Bird, A. and S. Tweedie (1995). "Transcriptional noise and the evolution of gene number." Philos Trans R Soc Lond B Biol Sci **349**(1329): 249-53.

Bird, A. P. (1995). "Gene number, noise reduction and biological complexity." Trends Genet **11**(3): 94-100.

Blanc, G. and K. H. Wolfe (2004). "Functional divergence of duplicated genes formed by polyploidy during Arabidopsis evolution." Plant Cell **16**(7): 1679-91.

Blattner, F. R., G. Plunkett, 3rd, et al. (1997). "The complete genome sequence of Escherichia coli K-12." Science **277**(5331): 1453-74.

Boyko, E. V., N. S. Badaev, et al. (1984). "Does DNA content change in the course of Triticale breeding?" Cereal Res. Commun. **12**: 99-100.

Brandes, A., J. S. Heslop-Harrison, et al. (1997). "Comparative analysis of the chromosomal and genomic organization of Ty1-copia-like retrotransposons in pteridophytes, gymnosperms and angiosperms." Plant Mol Biol **33**(1): 11-21.

Brandes, A., M. S. Roder, et al. (1995). "Barley telomeres are associated with two different types of satellite DNA sequences." Chromosome Res **3**(5): 315-20.

Brandes, A., H. Thompson, et al. (1997). "Multiple repetitive DNA sequences in the paracentromeric regions of Arabidopsis thaliana L." Chromosome Res **5**(4): 238-46.

Britten, R. J. and D. E. Kohne (1968). "Repeated sequences in DNA. Hundreds of thousands of copies of DNA sequences have been incorporated into the genomes of higher organisms." Science **161**(841): 529-40.

Brochmann, C. and *et.al.* (2004). "Polyploidy in artic plants." Biological Journal of the Linnean Society **82**: 521-536.

Brown, W. R., P. J. MacKinnon, et al. (1990). "Structure and polymorphism of human telomere-associated DNA." Cell **63**(1): 119-32.

Brubaker, C. L., A. H. Paterson, et al. (1999). "Comparative genetic mapping of allotetraploid cotton and its diploid progenitors." Genome **42**: 184-203.

Burr, B., F. A. Burr, et al. (1992). "Pinning down loose ends: mapping telomeres and factors affecting their length." Plant Cell **4**(8): 953-60.

Castiglione, M. R., E. Giraldi, et al. (1995). "The DNA methylation pattern of Allium cepa metaphase chromosomes." Biol Zentbl **114**: 57-66.

Castilho, A. and J. S. Heslop-Harrison (1995). "Physical mapping of 5S and 18S-25S rDNA and repetitive DNA sequences in Aegilops umbellulata." Genome **38**: 91-96.

Castilho, A., N. Neves, et al. (1999). "5-Methylcytosine distribution and genome organization in triticale before and after treatment with 5-azacytidine." J Cell Sci **112 (Pt 23)**: 4397-404.

Chan, S. W., I. R. Henderson, et al. (2005). "Gardening the genome: DNA methylation in Arabidopsis thaliana." Nat Rev Genet **6**(5): 351-60.

Charlesworth, B., P. Sniegowski, et al. (1994). "The evolutionary dynamics of repetitive DNA in eukaryotes." Nature **371**(6494): 215-20.

Chen, C. M., C. T. Wang, et al. (1997). "Two tandemly repeated telomere-associated sequences in Nicotiana plumbaginifolia." Chromosome Res **5**(8): 561-8.

Chen, Z. J., L. Comai, et al. (1998). "Gene dosage and stochastic effects determine the severity and direction of uniparental ribosomal RNA gene silencing (nucleolar dominance) in Arabidopsis allopolyploids." Proc Natl Acad Sci U S A **95**(25): 14891-6.

Chen, Z. J. and C. S. Pikaard (1997). "Epigenetic silencing of RNA polymerase I transcription: a role for DNA methylation and histone modification in nucleolar dominance." Genes Dev **11**(16): 2124-36.

Chen, Z. J. and C. S. Pikaard (1997). "Transcriptional analysis of nucleolar dominance in polyploid plants: biased expression/silencing of progenitor rRNA genes is developmentally regulated in Brassica." Proc Natl Acad Sci U S A **94**(7): 3442-7.

Cheng, Z., R. M. Stupar, et al. (2001). "A tandemly repeated DNA sequence is associated with both knob-like heterochromatin and a highly decondensed structure in the meiotic pachytene chromosomes of rice." Chromosoma **110**(1): 24-31.

Choi, Y., M. Gehring, et al. (2002). "DEMETER, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in arabidopsis." Cell **110**(1): 33-42.

Clarke, L. (1990). "Centromeres of budding and fission yeasts." Trends Genet **6**(5): 150-4.

Colot, V. and J. L. Rossignol (1999). "Eukaryotic DNA methylation as an evolutionary device." Bioessays **21**(5): 402-11.

Comai, L. (2000). "Genetic and epigenetic interactions in allopolyploid plants." Plant Mol Biol **43**(2-3): 387-99.

Comai, L. (2005). "The advantages and disadvantages of being polyploid." Nature Reviews Genetics **6**: 836-846.

Comai, L., A. Madlung, et al. (2003). "Do the different parental 'heteromes' cause genomic shock in newly formed allopolyploids?" Philos Trans R Soc Lond B Biol Sci **358**(1434): 1149-55.

Comai, L., A. P. Tyagi, et al. (2000). "Phenotypic instability and rapid gene silencing in newly formed arabidopsis allotetraploids." Plant Cell **12**(9): 1551-68.

Contento, A., M. Ceccarelli, et al. (2002). "Diversity of Olea genotypes and the origin of cultivated olives." Theor Appl Genet **104**(8): 1229-1238.

Contento, A., J. S. Heslop-Harrison, et al. (2005). "Diversity of a major repetitive DNA sequence in diploid and polyploid Triticeae." Cytogenet Genome Res **109**(1-3): 34-42.

Copenhaver, G. P., K. Nickel, et al. (1999). "Genetic definition and sequence analysis of Arabidopsis centromeres." Science **286**(5449): 2468-74.

Copenhaver, G. P. and C. S. Pikaard (1996). "RFLP and physical mapping with an rDNA-specific endonuclease reveals that nucleolus organizer regions of Arabidopsis thaliana adjoin the telomeres on chromosomes 2 and 4." Plant J **9**(2): 259-72.

Cuadrado, A., C. Ceoloni, et al. (1995). "Variation in highly repetitive DNA composition of heterochromatin in rye studied by fluorescence in situ hybridization." Genome **38**: 1061-1069.

Cuadrado, A. and N. Jouve (1994). "Mapping and organization of highly-repeated DNA sequences by means of simultaneous and sequential FISH and C-banding in 6x-triticale." Chromosome Res **2**(4): 331-8.

Cuadrado, A. and N. Jouve (1995). "Fluorescent in situ hybridization and C-banding analyses of highly repetitive DNA sequences in the heterochromatin of rye (Secale montanum Guss.) and wheat incorporating S. montanum chromosome segments." Genome **38**(4): 795-802.

Cuadrado, A. and T. Schwarzacher (1998). "The chromosomal organization of simple sequence repeats in wheat and rye genomes." Chromosoma **107**: 587-594.

Cubas, P., C. Vincent, et al. (1999). "An epigenetic mutation responsible for natural variation in floral symmetry." Nature **401**(6749): 157-61.

de Jong, J. H. (2003). "Visualizing DNA domains and sequecens by microscopy: a fifty-year history of molecular cytogenetics." Genome **46**: 943-946.

Deng, W., S. R. Liou, et al. (2003). "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2 and CT18." J Bacteriol **185**(7): 2330-7.

Devos, K. M. and M. D. Gale (2000). "Genome relationships: the grass model in current research." Plant Cell **12**(5): 637-46.

Dobosy, J. R. and E. U. Selker (2001). "Emerging connections between DNA methylation and histone acetylation." Cell Mol Life Sci **58**(5-6): 721-7.

Dong, F., J. T. Miller, et al. (1998). "Rice (Oryza sativa) centromeric regions consist of complex DNA." Proc Natl Acad Sci U S A **95**(14): 8135-40.

Doolittle, W. F. and C. Sapienza (1980). "Selfish genes, the phenotype paradigm and genome evolution." Nature **284**(5757): 601-3.

Dorigo, B., T. Schalch, et al. (2004). "Nucleosome arrays reveal the two-start organization of the chromatin fiber." Science **306**(5701): 1571-3.

Doyle, J. J., J. L. Doyle, et al. (2004). "Evolution of the perenial soybean complex (Glycine subgenus Glycine): a study of contrasts." Biological Journal of the Linnean Society **82**: 583-597.

Fajkus, J., A. Kovarik, et al. (1995). "Organization of telomeric and subtelomeric chromatin in the higher plant Nicotiana tabacum." Mol Gen Genet **247**(5): 633-8.

Fajkus, J., R. Kralovics, et al. (1995). "The telomeric sequence is directly attached to the HRS60 subtelomeric tandem repeat in tobacco chromosomes." FEBS Lett **364**(1): 33-5.

Feldman, M., B. Liu, et al. (1997). "Rapid elimination of low-copy DNA sequences in polyploid wheat: a possible mechanism for differentiation of homoeologous chromosomes." Genetics **147**(3): 1381-7.

Felsenburg, T., A. A. Levy, et al. (1991). "Polymorphism of high molecular weight glutenins in wild tetraploid wheat: spatial and temporal variation in a native site." Isr J Bot **40**: 451-479.

Feuillet, C. and B. Keller (2002). "Comparative genomics in the grass family: molecular characterization of grass genome structure and evolution." Ann Bot (Lond) **89**(1): 3-10.

Finnegan, E. J. (2001). "Epialleles - a source of random variation in times of stress." Curr Opin Plant Biol **5**: 101-106.

Finnegan, E. J., R. K. Genger, et al. (1998). "DNA Methylation in Plants." Annu Rev Plant Physiol Plant Mol Biol **49**: 223-247.

Finnegan, E. J. and K. A. Kovac (2000). "Plant DNA methyltransferases." Plant Mol Biol **43**(2-3): 189-201.

Finnegan, E. J., W. J. Peacock, et al. (1996). "Reduced DNA methylation in Arabidopsis thaliana results in abnormal plant development." Proc Natl Acad Sci U S A **93**(16): 8449-54.

Finnegan, E. J., W. J. Peacock, et al. (2000). "DNA methylation, a key regulator of plant development and other processes." Curr Opin Genet Dev **10**(2): 217-23.

Flavell, R. (1980). "The molecular characterization and organization of plant chromosomal DNA sequences." Annu Rev Plant Physiol **31**: 569-596.

Flavell, R. B. (1986). "Repetitive DNA and chromosome evolution in plants." Philos Trans R Soc Lond B Biol Sci **312**(1154): 227-42.

Flavell, R. B., M. D. Bennett, et al. (1974). "Genome size and the proportion of repeated nucleotide sequence DNA in plants." Biochem Genet **12**(4): 257-69.

Fransz, P. F., S. Armstrong, et al. (2000). "Integrated cytogenetic map of chromosome arm 4S of A. thaliana: structural organization of heterochromatic knob and centromere region." Cell **100**(3): 367-76.

Fransz, P. F. and J. H. de Jong (2002). "Chromatin dynamics in plants." Curr Opin Plant Biol **5**(6): 560-7.

Frediani, M., E. Giraldi, et al. (1996). "Distribution of 5-methylcytosine-rich regions in the metaphase chromosomes of Vicia faba." Chromosome Res **4**(2): 141-6.

Friesen, N., A. Brandes, et al. (2001). "Diversity, origin, and distribution of retrotransposons (gypsy and copia) in conifers." Mol Biol Evol **18**(7): 1176-88.

Fuchs, J., Z. Jasencakova, et al. (2005). "On chromatin modification and heterochromatin." Chromosome Res **13**(Supplement 1): 8-9.

Fulnecek, J., R. Matyasek, et al. (2002). "Distribution of 5-methylcytosine residues in 5S rRNA genes in Arabidopsis thaliana and Secale cereale." Mol Genet Genomics **268**(4): 510-7.

Fulnecek, J., R. Matyasek, et al. (1998). "Mapping of 5-methylcytosine residues in Nicotiana tabacum 5S rRNA genes by genomic sequencing." Mol Gen Genet **259**(2): 133-41.

Ganal, M. W., N. L. Lapitan, et al. (1991). "Macrostructure of the tomato telomeres." Plant Cell **3**(1): 87-94.

Gaut, B. S. and J. F. Doebley (1997). "DNA sequence evidence for the segmental allotetraploid origin of maize." Proc Natl Acad Sci U S A **94**(13): 6809-6814.

Gawel, N. J. and R. L. Jarret (1991). "A modified CTAB DNA extraction procedure for Musa and Ipomoea." Plant Molecular Biology Reporter **9**(3): 262-266.

Gazdova, B., J. Siroky, et al. (1995). "Characterization of a new family of tobacco highly repetitive DNA, GRS, specific for the Nicotiana tomentosiformis genomic component." Chromosome Res **3**(4): 245-54.

Gendrel, A. V., Z. Lippman, et al. (2002). "Dependence of heterochromatic histone H3 methylation patterns on the Arabidopsis gene DDM1." Science **297**(5588): 1871-3.

Gerlach, W. L. and T. A. Dyer (1980). "Sequence organization of the repeating units in the nucleus of wheat which contain 5S rRNA genes." Nucleic Acids Res **8**(21): 4851-65.

Gill, B. S., R. Appels, et al. (2004). "A workshop report on wheat genome sequencing: International Genome Research on Wheat Consortium." Genetics **168**(2): 1087-96.

Glyn, M. C., M. Egertova, et al. (1997). "The influence of 5-azacytidine on the condensation of the short arm of rye chromosome 1R in Triticum aestivum L. root tip meristematic nuclei." Chromosoma **106**(8): 485-92.

Goff, S. A., D. Ricke, et al. (2002). "A draft sequence of the rice genome (Oryza sativa L. ssp. japonica)." Science **296**(5565): 92-100.

Gondo, Y., T. Okada, et al. (1998). "Human megasatellite DNA RS447: copy-number polymorphisms and interspecies conservation." Genomics **54**(1): 39-49.

Gonzalgo, M. L. and P. A. Jones (1997). "Mutagenic and epigenetic effects of DNA methylation." Mutat Res **386**(2): 107-18.

Grandbastien, M. (1998). "Activation of plant retrotransposons under stress conditions." Trends in Plant Science **3**: 181-187.

Grant, D., P. Cregan, et al. (2000). "Genome organization in dicots: genome duplication in Arabidopsis and synteny between soybean and Arabidopsis." Proc Natl Acad Sci U S A **97**(8): 4168-73.

Grant, V. (1981). "Plant Speciation." Columbia University Press.

Grebenstein, B., O. Grebenstein, et al. (1996). "Distribution and complex organization of satellite DNA sequences in Aveneae species." Genome **39**(6): 1045-50.

Gruenbaum, Y., T. Naveh-Many, et al. (1981). "Sequence specificity of methylation in higher plant DNA." Nature **292**(5826): 860-2.

Guo, M., D. Davis, et al. (1996). "Dosage effects on gene expression in a maize ploidy series." Genetics **142**(4): 1349-55.

Han, F. P., G. Fedak, et al. (2003). "Rapid genomic changes in interspecific and intergeneric hybrids and allopolyploids of Triticeae." Genome **46**(4): 716-23.

Hancock, J. M. (1996). "Simple sequences and the expanding genome." Bioessays **18**(5): 421-5.

Hanson, R. E., M. N. Islam-Faridi, et al. (2000). "Ty1-copia-retrotransposon behavior in a polyploid cotton." Chromosome Res **8**(1): 73-6.

Hanson, R. E., X.-P. Zhao, et al. (1998). "Evolution of interspersed repetitive elements in Gossypium (Malvaceae)." Am J Bot **85**: 1364-1368.

He, P., B. R. Friebe, et al. (2003). "Allopolyploidy alters gene expression in the highly stable hexaploid wheat." Plant Mol Biol **52**(2): 401-14.

Heitz, E. (1932). "Die Herkunft der Chromozentren." Planta **18**: 571-636.

Hendrich, B. and A. Bird (1998). "Identification and characterization of a family of mammalian methyl-CpG binding proteins." Mol Cell Biol **18**(11): 6538-47.

Henikoff, S. (2000). "Heterochromatin function in complex genomes." Biochim Biophys Acta **1470**(1): O1-8.

Henikoff, S. and L. Comai (1998). "A DNA methyltransferase homolog with a chromodomain exists in multiple polymorphic forms in Arabidopsis." Genetics **149**(1): 307-18.

Hennig, W. (1999). "Heterochromatin." Chromosoma **108**(1): 1-9.

Heslop-Harrison, J. S. (2000). "Comparative genome organization in plants: from sequence and markers to chromatin and chromosomes." Plant Cell **12**(5): 617-36.

Heslop-Harrison, J. S., A. Brandes, et al. (2003). "Tandemly repeated DNA sequences and centromeric chromosomal regions of Arabidopsis species." Chromosome Research **11**: 241-253.

Heslop-Harrison, J. S., M. Murata, et al. (1999). "Polymorphisms and genomic organization of repetitive DNA from centromeric regions of Arabidopsis chromosomes." Plant Cell **11**(1): 31-42.

Hillier, L. W., W. Miller, et al. (2004). "Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution." Nature **432**(7018): 695-716.

Horakova, M. and J. Fajkus (2000). "TAS49--a dispersed repetitive sequence isolated from subtelomeric regions of Nicotiana tomentosiformis chromosomes." Genome **43**(2): 273-84.

Horn, P. J. and C. L. Peterson (2002). "Molecular biology. Chromatin higher order folding--wrapping up transcription." Science **297**(5588): 1824-7.

Huang, S., A. Sirikhachornkit, et al. (2002). "Genes encoding plastid acetyl-CoA carboxylase and 3-phosphoglycerate kinase of the Triticum/Aegilops complex and the evolutionary history of polyploid wheat." Proc Natl Acad Sci U S A **99**(12): 8133-8.

Hull, R. (1999). "Classification of reverse transcribing elements: a discussion document." Arch Virol **144**(1): 209-13; discussion 213-4.

Ingham, L. D., W. W. Hanna, et al. (1993). "Origin of the main class of repetitive DNA within selected Pennisetum species." Mol Gen Genet **238**(3): 350-6.

Jackson, J. P., L. Johnson, et al. (2004). "Dimethylation of histone H3 lysine 9 is a critical mark for DNA methylation and gene silencing in Arabidopsis thaliana." Chromosoma **112**(6): 308-15.

Jackson, J. P., A. M. Lindroth, et al. (2002). "Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase." Nature **416**(6880): 556-60.

Jaenisch, R. and A. Bird (2003). "Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals." Nat Genet **33 Suppl**: 245-54.

Jasencakova, Z., W. J. Soppe, et al. (2003). "Histone modifications in Arabidopsis- high methylation of H3 lysine 9 is dispensable for constitutive heterochromatin." Plant J **33**(3): 471-80.

Jeddeloh, J. A. and E. J. Richards (1996). "mCCG methylation in angiosperms." Plant J **9**(5): 579-86.

Jeddeloh, J. A., T. L. Stokes, et al. (1999). "Maintenance of genomic methylation requires a SWI2/SNF2-like protein." Nat Genet **22**(1): 94-7.

Jeltsch, A. (2002). "Beyond Watson and Crick: DNA methylation and molecular enzymology of DNA methyltransferases." Chembiochem **3**(4): 274-93.

Jenuwein, T. and C. D. Allis (2001). "Translating the histone code." Science **293**(5532): 1074-80.

Jiang, J. and B. S. Gill (1994). "Different species-specific chromosome translocations in Triticum timopheevii and T. turgidum support the diphyletic origin of polyploid wheats." Chromosome Res **2**(1): 59-64.

Jones, P. A. and D. Takai (2001). "The role of DNA methylation in mammalian epigenetics." Science **293**(5532): 1068-70.

Kakutani, T., J. A. Jeddeloh, et al. (1995). "Characterization of an Arabidopsis thaliana DNA hypomethylation mutant." Nucleic Acids Res **23**(1): 130-7.

Kalisz, S. and M. D. Purugganan (2004). "Epialleles via DNA methylation:consequences for plant evolution." Trends Ecol Evol **19**: 309-314.

Kashkush, K., M. Feldman, et al. (2002). "Gene loss, silencing and activation in a newly synthesized wheat allotetraploid." Genetics **160**(4): 1651-9.

Kashkush, K., M. Feldman, et al. (2003). "Transcriptional activation of retrotransposons alters the expression of adjacent genes in wheat." Nat Genet **33**(1): 102-6.

Kato, A., K. Yakura, et al. (1984). "Sequence analysis of Vicia faba repeated DNA, the FokI repeat element." Nucleic Acids Res **12**(16): 6415-26.

Katsiotis, A., M. Hagidimitriou, et al. (1997). "The close relationship between the A and B genomes in Avena L. (Poaceae) determined by molecular cytogenetic analysis of total genomic, tandemly and dispersed repetitive DNA." Ann Bot **79**: 103-107.

Kellogg, E. A. (1998). "Relationships of cereal crops and other grasses." Proc Natl Acad Sci U S A **95**(5): 2005-10.

Kilian, A. and A. Kleinhofs (1992). "Cloning and mapping of telomere-associated sequences from Hordeum vulgare L." Mol Gen Genet **235**: 153-6.

Kilian, A., D. Kudrna, et al. (1999). "Genetic and molecular characterization of barley chromosome telomeres." Genome(42): 412-419.

Kilian, A., C. Stiff, et al. (1995). "Barley telomeres shorten during differentiation but grow in callus culture." Proc Natl Acad Sci U S A **92**(21): 9555-9.

Kim, J. M., S. Vanguri, et al. (1998). "Transposable elements and genome organization: a comprehensive survey of retrotransposons revealed by the complete Saccharomyces cerevisiae genome sequence." Genome Res **8**(5): 464-78.

Kishii, M., K. Nagaki, et al. (2001). "A tandem repetitive sequence located in the centromeric region of common wheat (Triticum aestivum) chromosomes." Chromosome Res **9**(5): 417-28.

Kishii, M., K. Nagaki, et al. (1999). "Exclusive localization of tandem repetitive sequences in subtelomeric heterochromatin regions of Leymus racemosus (Poaceae, Triticeae)." Chromosome Res **7**(7): 519-29.

Kishii, M. and H. Tsujimoto (2002). "Genus-specific localization of the TaiI family of tandem-repetitive sequences in either the centromeric or subtelomeric regions in Triticeae species (Poaceae) and its evolution in wheat." Genome **45**(5): 946-55.

Kornberg, R. D. and Y. Lorch (1999). "Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome." Cell **98**(3): 285-94.

Kovarik, A., B. Koukalova, et al. (2000). "Comparative analysis of DNA methylation in tobacco heterochromatic sequences." Chromosome Res **8**(6): 527-41.

Kovarik, A., R. Matyasek, et al. (1997). "Variability in CpNpG methylation in higher plant genomes." Gene **204**(1-2): 25-33.

Kovarik, A., R. Matyasek, et al. (2004). "Concerted evolution of 18-5.8-26S rDNA repeats in Nicotiana allotetraploids." Biological Journal of the Linnean Society **82**: 615-625.

Kress, C., H. Thomassin, et al. (2001). "Local DNA demethylation in vertebrates: how could it be performed and targeted?" FEBS Lett **494**(3): 135-40.

Kubis, S. E., A. M. Castilho, et al. (2003). "Retroelements, transposons and methylation status in the genome of oil palm (Elaeis guineensis) and the relationship to somaclonal variation." Plant Mol Biol **52**(1): 69-79.

Kubis, S. E., T. Schmidt, et al. (1998). "Repetitive DNA elements as a major component of plant genomes." Ann Bot **82**: 45-55.

Kumar, A. and J. L. Bennetzen (1999). "Plant retrotransposons." Annu Rev Genet **33**: 479-532.

Lagercrantz, U., H. Ellegren, et al. (1993). "The abundance of various polymorphic microsatellite motifs differs between plants and vertebrates." Nucleic Acids Res **21**(5): 1111-5.

Lapitan, N. L., M. W. Ganal, et al. (1989). "Somatic chromosome karyotype of tomato based on in situ hybridization of the TGR1 satellite repeat." Genome **32**: 992-998.

Lee, C., R. Wevrick, et al. (1997). "Human centromeric DNAs." Hum Genet **100**(3-4): 291-304.

Lee, H. S. and Z. J. Chen (2001). "Protein-coding genes are epigenetically regulated in Arabidopsis polyploids." Proc Natl Acad Sci U S A **98**(12): 6753-8.

Leitch, I. J. and M. D. Bennett (1997). "Polyploidy in Angiosperms." Trends Plant Sci **2**(12): 470-476.

Leitch, I. J. and M. D. Bennett (2004). "Genome downsizing in polyploid plants." Biological Journal of the Linnean Society **82**: 651-663.

Leutwiler, L. S., B. R. Hough-Evans, et al. (1984). "The DNA of Arabidopsis thaliana." Mol Gen Genet **194**: 15-23.

Levy, A. A. and M. Feldman (1988). "Ecogeographical distribution of HMW glutenin alleles in populations of the wild tetraploid wheat Triticum turgidum var. dicoccoides." Theor Appl Genet **75**: 651-658.

Levy, A. A. and M. Feldman (2002). "The impact of polyploidy on grass genome evolution." Plant Physiol **130**(4): 1587-93.

Levy, A. A. and M. Feldman (2004). "Genetic and epigenetic reprogramming of the wheat genome upon allopolyploidization." Biological Journal of the Linnean Society **82**: 607-613.

Lewis, J. D., R. R. Meehan, et al. (1992). "Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA." Cell **69**(6): 905-14.

Lindroth, A. M., X. Cao, et al. (2001). "Requirement of CHROMOMETHYLASE3 for maintenance of CpXpG methylation." Science **292**(5524): 2077-80.

Liu, B., C. L. Brubaker, et al. (2001). "Polyploid formation in cotton is not accompanied by rapid genomic changes." Genome **44**(3): 321-30.

Liu, B., J. M. Vega, et al. (1998). "Rapid genomic changes in newly synthesized amphiploids of Triticum and Aegilops. II. Changes in low-copy coding DNA sequences." Genome **41**(4): 535-42.

Liu, B., J. M. Vega, et al. (1998). "Rapid genomic changes in newly synthesized amphiploids of Triticum and Aegilops. I. Changes in low-copy non-coding DNA sequences." Genome **41**: 272-277.

Liu, B. and J. F. Wendel (2003). "Epigenetic phenomena and the evolution of plant allopolyploids." Mol Phylogenet Evol **29**(3): 365-79.

Lohe, A. R. and A. J. Hilliker (1995). "Return of the H-word (heterochromatin)." Curr Opin Genet Dev **5**(6): 746-55.

Loidl, P. (2004). "A plant dialect of the histone language." Trends Plant Sci **9**(2): 84-90.

Louis, E. and A. Vershinin (2005). "Telomeres: a diversity of solutions to the problem of chromosome ends." Chromosome Res **13**(5): 425-9.

Louis, E. J. (1995). "The chromosome ends of Saccharomyces cerevisiae." Yeast **11**(16): 1553-73.

Love, A., D. Love, et al. (1977). "Cytotaxonomical Atlas of the Pteridophyta." Vaduz, Germany: Cramer.

Luff, B., L. Pawlowski, et al. (1999). "An inverted repeat triggers cytosine methylation of identical sequences in Arabidopsis." Mol Cell **3**(4): 505-11.

Luo, S. and D. Preuss (2003). "Strand-biased DNA methylation associated with centromeric regions in Arabidopsis." Proc Natl Acad Sci U S A **100**(19): 11133-8.

Lynch, M. and J. S. Conery (2000). "The evolutionary fate and consequences of duplicate genes." Science **290**(5494): 1151-5.

Ma, X. F., P. Fang, et al. (2004). "Polyploidization-induced genome variation in triticale." Genome **47**(5): 839-48.

Macas, J., T. Meszaros, et al. (2002). "PlantSat: a specialized database for plant satellite repeats." Bioinformatics **18**(1): 28-35.

Macas, J., D. Pozarkova, et al. (2000). "Two new families of tandem repeats isolated from genus Vicia using genomic self-priming PCR." Mol Gen Genet **263**(5): 741-51.

Madlung, A. and L. Comai (2004). "The effect of stress on genome regulation and structure." Ann Bot (Lond) **94**(4): 481-95.

Madlung, A., R. W. Masuelli, et al. (2002). "Remodeling of DNA methylation and phenotypic and transcriptional changes in synthetic Arabidopsis allotetraploids." Plant Physiol **129**(2): 733-46.

Madlung, A., A. P. Tyagi, et al. (2005). "Genomic changes in synthetic Arabidopsis polyploids." Plant J **41**(2): 221-30.

Mao, L., K. M. Devos, et al. (1997). "Cloning and genetic mapping of wheat telomere-associated sequences." Mol Gen Genet **254**: 584-591.

Martienssen, R. (1998). "Transposons, DNA methylation and gene control." Trends Genet **14**(7): 263-4.

Martienssen, R. A. and V. Colot (2001). "DNA methylation and epigenetic inheritance in plants and filamentous fungi." Science **293**(5532): 1070-4.

Masterson, J. (1994). "Stomatal Size in Fossil Plants:Evidence for Polyploidy in Majority of Angiosperms." Science **264**: 421-423.

Matyasek, R., B. Gazdova, et al. (1997). "NTRS, a new family of highly repetitive DNAs specific for the T1 chromosome of tobacco." Chromosoma **106**(6): 369-79.

Matzke, M., A. J. Matzke, et al. (2001). "RNA: guiding gene silencing." Science **293**(5532): 1080-3.

Matzke, M. A., O. Mittelsten Scheid, et al. (1999). "Rapid structural and epigenetic changes in polyploid and aneuploid genomes." Bioessays **21**(9): 761-7.

McAllister, B. F. and J. H. Werren (1999). "Evolution of tandemly repeated sequences: What happens at the end of an array?" J Mol Evol **48**(4): 469-81.

McClelland, M., M. Nelson, et al. (1994). "Effect of site-specific modification on restriction endonucleases and DNA modification methyltransferases." Nucleic Acids Res **22**(17): 3640-59.

McClintock, B. (1941). "The stability of broken ends of chromosomes in Zea mays." Genetics **26**: 279-282.

McIntyre, C. L., S. Pereira, et al. (1990). "New Secale cereale (rye) DNA derivatives for the detection of rye chromosome segments in wheat." Genome **33**(5): 635-40.

Melquist, S., B. Luff, et al. (1999). "Arabidopsis PAI gene arrangements, cytosine methylation and expression." Genetics **153**(1): 401-13.

Messeguer, R., M. W. Ganal, et al. (1991). "Characterization of the level, target sites and inheritance of cytosine methylation in tomato nuclear DNA." Plant Mol Biol **16**(5): 753-70.

Mewes, H. W., K. Albermann, et al. (1997). "Overview of the yeast genome." Nature **387**(6632 Suppl): 7-65.

Meyne, J., R. J. Baker, et al. (1990). "Distribution of non-telomeric sites of the (TTAGGG)n telomeric sequence in vertebrate chromosomes." Chromosoma **99**(1): 3-10.

Montero, L. M., J. Filipski, et al. (1992). "The distribution of 5-methylcytosine in the nuclear genome of plants." Nucleic Acids Res **20**(12): 3207-10.

Moore, G. (2000). "Cereal Chromosome Structure, Evolution, and Pairing." Annu Rev Plant Physiol.Plant Mol Biol **51**: 195-222.

Mukai, Y., Y. Nakara, et al. (1993). "Simultaneus discrimination of three genomes in hexaploid wheat by multicolor fluorescence in situ hybridization using total genomic and highly repetitive DNA probes." Genome **36**: 489-494.

Mullis, K., F. Faloona, et al. (1986). "Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction." Cold Spring Harb Symp Quant Biol **51 Pt 1**: 263-73.

Murata, M. (2002). "Telomeres and centromeres in plants." Current Genomics(3): 527-538.

Murata, M., J. S. Heslop-Harrison, et al. (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**(1): 31-7.

Nagaki, K., H. Tsujimoto, et al. (1995). "Molecular characterization of a tandem repeat, Afa family, and its distribution among Triticeae." Theor Appl Genet **38**: 479-486.

Nagaki, K., H. Tsujimoto, et al. (1998). "Dynamics of tandem repetitive Afa-family sequences in Triticeae, wheat-related species." J Mol Evol **47**(2): 183-9.

Neves, N., M. Silva, et al. (1997). "Nucleolar dominance in triticales: control by unlinked genes." Chromosome Res **5**(2): 125-31.

O'Hare, K., B. P. Chadwick, et al. (2002). "A 5.9-kb tandem repeat at the euchromatin-heterochromatin boundary of the X chromosome of Drosophila melanogaster." Mol Genet Genomics **267**(5): 647-55.

Ohmido, N., K. Kijima, et al. (2000). "Quantification of total genomic DNA and selected repetitive sequences reveals concurrent changes in different DNA families in indica and japonica rice." Mol Gen Genet **263**(3): 388-94.

Ohmido, N., K. Kijima, et al. (2001). "Visualization of the terminal structure of rice chromosomes 6 and 12 with multicolor FISH to chromosomes and extended DNA fibers." Plant Mol Biol **47**(3): 413-21.

Ohno, S. (1970). "Evolution by gene duplication." New York:Springer-Verlag.

Orgel, L. E. and F. H. Crick (1980). "Selfish DNA: the ultimate parasite." Nature **284**(5757): 604-7.

Osborn, T. C., J. C. Pires, et al. (2003). "Understanding mechanisms of novel gene expression in polyploids." Trends Genet **19**(3): 141-7.

Otto, S. P. and J. Whitton (2000). "Polyploid incidence and evolution." Annu Rev Genet **34**: 401-437.

Ozkan, H., A. Brandolini, et al. (2005). "A reconsideration of the domestication geography of tetraploid wheats." Theor Appl Genet **110**(6): 1052-60.

Ozkan, H., A. A. Levy, et al. (2001). "Allopolyploidy-induced rapid genome evolution in the wheat (Aegilops-Triticum) group." Plant Cell **13**(8): 1735-47.

Ozkan, H., M. Tuna, et al. (2003). "Nonadditive changes in genome size during allopolyploidization in the wheat (aegilops-triticum) group." J Hered **94**(3): 260-4.

Papa, C. M., N. M. Springer, et al. (2001). "Maize chromomethylase Zea methyltransferase2 is required for CpNpG methylation." Plant Cell **13**(8): 1919-28.

Petsova, E. G., N. P. Goncharov, et al. (1998). "Elimination of tandem repeat of telomeric heterochromatin during the evolution of wheat." Theor Appl Genet **97**: 1380-1386.

Pichersky, E., D. Soltis, et al. (1990). "Defective chlorophyll a/b-binding protein genes in the genome of a homosporous fern." Proc Natl Acad Sci U S A **87**(1): 195-9.

Pikaard, C. S. (2000). "Nucleolar dominance: uniparental gene silencing on a multi-megabase scale in genetic hybrids." Plant Mol Biol **43**(2-3): 163-77.

Pires, J. C., J. Zhao, et al. (2004). "Flowering time divergence and genomic rearrangements in resynthesized Brassica polyploids." Biological Journal of the Linnean Society **82**: 675-688.

Plagemann, P. G., M. Behrens, et al. (1978). "Metabolism and cytotoxicity of 5-azacytidine in cultured Novikoff rat hepatoma and P388 mouse leukemia cells and their enhancement by preincubation with pyrazofurin." Cancer Res **38**(8): 2458-66.

Prokhortchouk, A., B. Hendrich, et al. (2001). "The p120 catenin partner Kaiso is a DNA methylation-dependent transcriptional repressor." Genes Dev **15**(13): 1613-8.

Pruitt, R. E. and E. M. Meyerowitz (1986). "Characterization of the genome of Arabidopsis thaliana." J Mol Biol **187**(2): 169-83.

Rabinowicz, P. D., L. E. Palmer, et al. (2003). "Genes and transposons are differentially methylated in plants, but not in mammals." Genome Res **13**(12): 2658-64.

Rabinowicz, P. D., K. Schutz, et al. (1999). "Differential methylation of genes and retrotransposons facilitates shotgun sequencing of the maize genome." Nat Genet **23**(3): 305-8.

Rapp, R. A. and J. F. Wendel (2005). "Epigenetics and plant evolution." New Phytologist **165**: 1-11.

Rayburn, A. L. and B. S. Gill (1986). " Isolation of a D-genome specific repeated DNA sequence from Aegilops squarrosa." Plant Molecular Biology Reporter **4**: 102-109.

Reik, W., W. Dean, et al. (2001). "Epigenetic reprogramming in mammalian development." Science **293**(5532): 1089-93.

Richards, E. J. (1997). "DNA methylation and plant development." Trends Genet **13**(8): 319-323.

Richards, E. J. and F. M. Ausubel (1988). "Isolation of a higher eukaryotic telomere from Arabidopsis thaliana." Cell **53**(1): 127-36.

Richards, E. J., S. Chao, et al. (1992). "Characterization of Arabidopsis thaliana telomeres isolated in yeast." Nucleic Acids Res **20**(15): 4039-46.

Richards, E. J. and S. C. Elgin (2002). "Epigenetic codes for heterochromatin formation and silencing: rounding up the usual suspects." Cell **108**(4): 489-500.

Robertson, K. D., S. Ait-Si-Ali, et al. (2000). "DNMT1 forms a complex with Rb, E2F1 and HDAC1 and represses transcription from E2F-responsive promoters." Nat Genet **25**(3): 338-42.

Roder, M. S., N. L. Lapitan, et al. (1993). "Genetic and physical mapping of barley telomeres." Mol Gen Genet **238**(1-2): 294-303.

Rountree, M. R., K. E. Bachman, et al. (2000). "DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci." Nat Genet **25**(3): 269-77.

Salamini, F., M. Heun, et al. (2004). "Comment on "AFLP data and the origins of domesticated crops"." Genome **47**(3): 615-20; discussion 621-2.

Salamini, F., H. Ozkan, et al. (2002). "Genetics and geography of wild cereal domestication in the near east." Nat Rev Genet **3**(6): 429-41.

Salmon, A., M. L. Ainouche, et al. (2005). "Genetic and epigenetic consequences of recent hybridization and polyploidy in Spartina (Poaceae)." Mol Ecol **14**(4): 1163-75.

Santi, D. V., A. Norment, et al. (1984). "Covalent bond formation between a DNA-cytosine methyltransferase and DNA containing 5-azacytosine." Proc Natl Acad Sci U S A **81**(22): 6993-7.

Schlotterer, C. and D. Tautz (1992). "Slippage synthesis of simple sequence DNA." Nucleic Acids Res **20**: 211-215.

Schmidt, T. (1999). "LINEs, SINEs and repetitive DNA: non-LTR retrotransposons in plant genomes." Plant Mol Biol **40**(6): 903-10.

Schmidt, T. and J. S. Heslop-Harrison (1996). "The physical and genomic organization of microsatellites in sugar beet." Proc Natl Acad Sci U S A **93**(16): 8761-5.

Schmidt, T. and J. S. Heslop-Harrison (1998). "Genomes, genes and junk: the large-scale organization of plant chromosomes." Trends Plant Sci **3**(5): 195-199.

Schranz, M. E. and T. C. Osborn (2000). "Novel flowering time variation in the resynthesized polyploid Brassica napus." J Hered **91**(3): 242-6.

Schwarzacher, T. (1996). "The physical organization of Triticeae chromosomes." The Society of Experimental Biology: 71-76.

Schwarzacher , T. (2003). "Meiosis, recombination and chromosomes: a review of gene isolation and fluorescent in situ hybridization data in plants." Journal of Experimental Botany **54**: 11-23.

Schwarzacher, T. and J. S. Heslop-Harrison (1991). "In situ hybridization to plant telomeres using synthetic oligomers." Genome **34**: 317-323.

Schwarzacher, T. and J. S. Heslop-Harrison (2000). Practical in situ hybridization. Oxford.

Schweizer, G., M. W. Ganal, et al. (1988). "Species-specific DNA sequences for the identification of somatic hybrids between Lycopersicon esculentum and Solanum acaule." Theor Appl Genet **75**: 679-684.

Seal, A. G. and M. D. Bennet (1981). "The rye genome in winter hexaploid Triticales." Genome **23**: 647-653.

Shaked, H., K. Kashkush, et al. (2001). "Sequence elimination and cytosine methylation are rapid and reproducible responses of the genome to wide hybridization and allopolyploidy in wheat." Plant Cell **13**(8): 1749-59.

Shapiro (1976). "Distribution of purines and pyrimidines in deoxyribonucleic acids." Handbook of Biochemistry and Molecular Biology **CRC Press**: 241-281.

Siedler, H., M. M. Messmer, et al. (1994). "Genetic diversity in European wheat and spelt breeding material based on RLFP data." Theor Appl Genet **88**: 994-1003.

Singal, R. and G. D. Ginder (1999). "DNA methylation." Blood **93**(12): 4059-70.

Siroky, J., M. R. Castiglione, et al. (1998). "DNA methylation patterns of Melandrium album chromosomes." Chromosome Res **6**(6): 441-6.

Smith, A. F. A. (1996). "The origin of interspersed repeats in the human genome." Curr Opin Genet Dev **6**: 743-748.

Soltis, D. E. and P. S. Soltis (1995). "The dynamic nature of polyploid genomes." Proc Natl Acad Sci U S A **92**(18): 8089-91.

Soltis, D. E. and P. S. Soltis (1999). "Polyploidy: recurrent formation and genome evolution." Trends in Ecology and Evolution **14**(9): 348-352.

Soltis, D. E., P. S. Soltis, et al. (2004). "Recent and recurrent polyploidy in Tragopogon (Asteraceae):cytogenetic, genomic and genetic comparisons." Biological Journal of the Linnean Society **82**: 485-501.

Soltis, P. S. and D. E. Soltis (2000). "The role of genetic and genomic attributes in the success of polyploids." Proc Natl Acad Sci U S A **97**(13): 7051-7.

Song, K., P. Lu, et al. (1995). "Rapid genome change in synthetic polyploids of Brassica and its implications for polyploid evolution." Proc Natl Acad Sci U S A **92**(17): 7719-23.

Soppe, W. J., Z. Jasencakova, et al. (2002). "DNA methylation controls histone H3 lysine 9 methylation and heterochromatin assembly in Arabidopsis." Embo J **21**(23): 6549-59.

Southern, E. M. (1975). "Detection of specific sequences among DNA fragments separated by gel electrophoresis." J Mol Biol **98**(503-517).

Sperisen, C., J. Ryals, et al. (1991). "Comparison of cloned genes provides evidence for intergenomic exchange of DNA in the evolution of a tobacco glucan endo-1,3-beta-glucosidase gene family." Proc Natl Acad Sci U S A **88**(5): 1820-4.

Springer, N. M. and S. M. Kaeppler (2005). "Evolutionary divergence of monocot and dicot methyl-CpG-binding domain proteins." Plant Physiol **138**(1): 92-104.

Stebbins, G. L. (1950). "Variation and Evolution in Plants." Columbia University Press, New York.

Stebbins, G. L. (1971). "Chromosomal Evolution in Higher Plants." Edward Arnold, London.

Strahl, B. D. and C. D. Allis (2000). "The language of covalent histone modifications." Nature **403**(6765): 41-5.

Struhl, K. (1999). "Fundamentally different logic of gene regulation in eukaryotes and prokaryotes." Cell **98**(1): 1-4.

Surani, M. A. (2001). "Reprogramming of genome function through epigenetic inheritance." Nature **414**(6859): 122-8.

Swanson, C. P. and P. L. Webster (1975). The cell, 4th edn. Englewood Cliffs, New Jersey.

Sykorova, E., K. Y. Lim, et al. (2003). "Telomere variability in the monocotyledonous plant order Asparagales." Proc Biol Sci **270**(1527): 1893-904.

Taketa, S., H. Ando, et al. (2000). "The distribution, organization and evolution of two abundant and widespread repetitive DNA sequences in the genus Hordeum." Theor Appl Genet **100**: 169-176.

Taketa, S., G. Harrison, et al. (1999). "Comparative physical mapping of the 5S and 18S-25S rDNA in nine wild *Hordeum* species and cytotypes." Theor. Appl. Genet. **98**: 1-9.

Tariq, M. and J. Paszkowski (2004). "DNA and histone methylation in plants." Trends Genet **20**(6): 244-51.

Tate, P. H. and A. P. Bird (1993). "Effects of DNA methylation on DNA-binding proteins and gene expression." Curr Opin Genet Dev **3**(2): 226-31.

Tautz, D. and M. Renz (1984). "Simple sequences are ubiquitous repetitive components of eukaryotic genomes." Nucleic Acids Res **12**(10): 4127-38.

The Arabidopsis Genome Initiative (2000). "Analysis of the genome sequence of the flowering plant Arabidopsis thaliana." Nature **408**(6814): 796-815.

Tommerup, H., A. Dousmanis, et al. (1994). "Unusual chromatin in human telomeres." Mol Cell Biol **14**(9): 5777-85.

Tweedie, S., H. H. Ng, et al. (1999). "Vestiges of a DNA methylation system in Drosophila melanogaster?" Nat Genet **23**(4): 389-90.

Ugarkovic, D. and M. Plohl (2002). "Variation in satellite DNA profiles--causes and effects." Embo J **21**(22): 5955-9.

Ungerer, M. C., S. J. Baird, et al. (1998). "Rapid hybrid speciation in wild sunflowers." Proc Natl Acad Sci U S A **95**(20): 11757-62.

Venter, J. C., M. D. Adams, et al. (2001). "The sequence of the human genome." Science **291**(5507): 1304-51.

Vershinin, A., S. Svitashev, et al. (1994). "Characterization of a family of tandemly repeated DNA sequence in Triticeae." Theor Appl Genet **89**: 217-225.

Vershinin, A. V. and J. S. Heslop-Harrison (1998). "Comparative analysis of the nucleosomal structure of rye, wheat and their relatives." Plant Mol Biol **36**(1): 149-61.

Vershinin, A. V., T. Schwarzacher, et al. (1995). "The large-scale genomic organization of repetitive DNA families at the telomeres of rye chromosomes." Plant Cell **7**(11): 1823-33.

Vinkenoog, R., C. Bushell, et al. (2003). "Genomic imprinting and endosperm development in flowering plants." Mol Biotechnol **25**(2): 149-84.

Vongs, A., T. Kakutani, et al. (1993). "Arabidopsis thaliana DNA methylation mutants." Science **260**(5116): 1926-8.

Voylokov, V. and N. D. Tikhenko (2002). "Triticale as a model for study genome interaction and genome evolution in allopolyploid plants." Proceedings 5th International Triticale Symposium **Vol.1 Edited by Aniol *et al.* Radzikow, Poland**: 63-70.

Wagner, I. and I. Capesius (1981). "Determination of 5-methylcytosine from plant DNA by high-performance liquid chromatography." Biochem Biophys Acta **654**: 52-56.

Walsh, C. P. and T. H. Bestor (1999). "Cytosine methylation and mammalian development." Genes Dev **13**(1): 26-34.

Wang, Y., J. Wysocka, et al. (2004). "Linking covalent histone modifications to epigenetics: the rigidity and plasticity of the marks." Cold Spring Harb Symp Quant Biol **69**: 161-9.

Waterston, R. H., K. Lindblad-Toh, et al. (2002). "Initial sequencing and comparative analysis of the mouse genome." Nature **420**(6915): 520-62.

Wendel, J. F. (2000). "Genome evolution in polyploids." Plant Mol Biol **42**(1): 225-49.

Wendel, J. F., A. Schnabel, et al. (1995). "Bidirectional interlocus concerted evolution following allopolyploid speciation in cotton (Gossypium)." Proc Natl Acad Sci U S A **92**(1): 280-4.

Wendel, J. F. and S. R. Wessler (2000). "Retrotransposon-mediated genome evolution on a local ecological scale." Proc Natl Acad Sci U S A **97**(12): 6250-2.

Wessler, S. R. (1996). "Turned on by stress. Plant retrotransposons." Curr Biol **6**(8): 959-61.

Willard, H. F. (1989). "The genomics of long tandem arrays of satellite DNA in the human genome." Genome **31**(2): 737-44.

Wilson, A. S. (1876). "Wheat and rye hybrids." Edimburgh Bat. Sac. Trans **12**: 286-288.

Wolffe, A. P. and D. Pruss (1996). "Deviant nucleosomes: the functional specialization of chromatin." Trends Genet **12**(2): 58-62.

Wood, V., R. Gwilliam, et al. (2002). "The genome sequence of Schizosaccharomyces pombe." Nature **415**(6874): 871-80.

Wright, J. H., D. E. Gottschling, et al. (1992). "Saccharomyces telomeres assume a non-nucleosomal chromatin structure." Genes Dev **6**(2): 197-210.

Wu, C. and J. R. Morris (2001). "Genes, genetics, and epigenetics: a correspondence." Science **293**(5532): 1103-5.

Wu, H. K., M. C. Chung, et al. (1991). "Localization of specific repetitive DNA sequences in individual rice chromosomes." Chromosoma **100**(5): 330-8.

Wu, K. S. and S. D. Tanksley (1993). "Genetic and physical mapping of telomeres and macrosatellites of rice." Plant Mol Biol **22**(5): 861-72.

Yoder, J. A., C. P. Walsh, et al. (1997). "Cytosine methylation and the ecology of intragenomic parasites." Trends Genet **13**(8): 335-40.

Yu, J., S. Hu, et al. (2002). "A draft sequence of the rice genome (Oryza sativa L. ssp. indica)." Science **296**(5565): 79-92.

Zhao, X. P., Y. Si, et al. (1998). "Dispersed repetitive DNA has spread to new genomes since polyploid formation in cotton." Genome Res **8**(5): 479-92.

Zhong, X. B., P. F. Fransz, et al. (1998). "FISH studies reveal the molecular and chromosomal organization of individual telomere domains in tomato." Plant J **13**(4): 507-17.