High copy sequences reveal the unique composition and evolution of the rye B chromosome

Dissertation
zur Erlangung des
Doktorgrades der Naturwissenschaften (Dr. rer. nat.)

der

Naturwissenschaftliche Fakultät I
Biowissenschaften
der Martin-Luther-Universität Halle-Wittenberg

vorgelegt von
Frau M.Sc. Sonja Klemme
geboren am 19.03.1984 in Schönebeck

Gutachter:

Dr Andreas Houben
Leibniz Institute of Plant Genetics and Crop Plant Research (IPK)
OT Gatersleben
Corrensstrasse 3
D-06466 Stadt Seeland
Germany

Prof. Dr. Neil Jones
Edward Llwyd Building, Penglais
Aberystwyth University
Aberystwyth, Ceredigion, SY23 3FL
Wales

Herr Prof. Dr. K. Humbeck
Institutsbereich Pflanzenphysiologie, Weinbergweg 10
06120 Halle (Saale)
Germany

Keywords: Chromosome Evolution, Fluorescence In Situ Hybridization, B chromosomes, Next Generation Sequencing, Genome Evolution, Chromosome Drive, Late Replication, Nondisjunction Control Region, satellite DNA, Transposable DNA Elements, organelle DNA, *Secale cereale* (rye), Chromosome mapping, Southern Hybridization, Sister Chromatid Exchange, Pericentromere

Stichworte: Chromosomenevolution, Fluoreszens-In-Situ-Hybridisierung, B Chromosomen, Next generation Sequenzierung, Genomevolution, Chromosomen Drive, Späte Replikation, Nondisjunktions-Kontroll-Region, Satelliten DNA, mobile Elemente, Organellen DNA, *Secale cereale* (Roggen), Chromosomenkartierung, Southern Hybridisierung, Schwesterchromatidaustausch, Perizentromer

Verteidigt am: 11.12.2013

Acknowledgements

This work was funded by the Deutsche Forschungsgemeinschaft (DFG) and carried out in the group Chromosome structure and function (CSF) of the Leibniz-Institut für Pflanzengenetik und Kulturpflanzenforschung (IPK) in Gatersleben.

My sincere thanks go out to my supervisor Dr. Andreas Houben for all his invaluable support and guidance during the course of my work. His help and encouragement were very appreciated.

I thank the whole group CSF for a truly enjoyable working atmosphere and inspiring discussions.

I would like to thank Dr. Ali Mohammad Banaei Moghaddam for his support, his excellent ideas and comments and for sharing his enthusiasm towards B chromosomes with me.

Many thanks to Dr. Jiri Macas and his group for his bioinformatic brilliance and giving me the opportunity to visit his group.

I thank the whole group of Dr. Klaus Mayer, in particular Mihaela Martis for bioinformatic assistance.

I would like to thank Dr. Veit Schubert for his work on the high resolution microscopy pictures.

Many thanks to Prof. Dr. Ingo Schubert for many enlightening and enjoyable talks as well as his helpful criticisms and interesting questions.

I wish to thank the technicians of the group, Oda Weiss, Karla Meier and especially Katrin Kumke for their great help and assistance as well as for allowing silly scientists to mess up their carefully ordered lab.

Ein besonderer Dank gilt meiner Familie und besonders meiner Mutter, die mich immer unterstützt und an mich geglaubt hat. Einen Dank auch an Florian Klemme für mathematische Unterstützung und Chauffeurdienste.

Index

Abstract

B chromosomes (Bs) are supernumerary chromosomes that vary in number between individuals of the same species. Because of the dispensable nature of the Bs and their non-Mendelian inheritance, one might assume the B followed a different evolutionary pathway than the A chromosomes (As). This is reflected by differences in their high-copy DNA constitution.

We provide a deep insight into the composition and distribution of rye (*Secale cereale*) B-located high-copy sequences. Although most mobile elements are similarly distributed along A and B chromosomes, several transposable sequences are either amplified or depleted on the B. A massive accumulation of B-enriched satellite repeats was found mostly in the region at the terminal part of the long arm regulating the nondisjunction of sister chromatids during the first pollen mitosis. This area is transcriptionally active and late replicating. Additionally, we found a B pericentromerically enriched sequence and large insertions of plastid and mitochondrial DNA. All B-enriched sequence components are not unique to the B but are also present in As of several *Secale* species, although often in lower copy numbers, suggesting the origin of the B from As of the same genus.

This work provides a cytogenetic map of the rye B chromosome. This map may serve as a tool for further research on this enigmatic chromosome type. We established the possibility to reliably identify Bs during interphase. Our findings highlight the differences between A and B chromosomes. Although Bs originated in As, they have taken a separate evolutionary pathway and are an invaluable research tool for studies on chromosome evolution.

Zusammenfassung

B Chromosomen (Bs) sind überzählige Chromosomen, die zwischen Individuen einer Spezies in ihrer Anzahl variieren können. Da die Bs entbehrlich sind und ihre Vererbung unabhängig von den Mendelschen Gesetzen ist, darf man annehmen, dass ihre Evolution einen anderen Weg beschreitet als die der A Chromosomen (As). Dies ist auch ersichtlich in der Zusammensetzung ihrer repetitiven DNA.

Wir geben Einsicht in die Zusammensetzung und Verteilung von stark amplifizierten Sequenzen auf dem B von Roggen (*Secale cereale* L.). Obwohl die meisten mobilen Elemente auf As und Bs gleich verteilt sind, fanden wir mehrere Sequenzen auf dem B vermehrt oder vermindert vor. Eine massive Anhäufung von B-angereicherten Satellitensequenzen wurde vor allem in der Endregion des langen Chromosomenarms gefunden, die für die Nicht-Teilung der Schwesterchromatiden während der ersten Pollenmitose zuständig ist. Diese Region ist transkriptionell aktiv und repliziert spät. Zusätzlich fanden wir eine Sequenz, die im Perizentromer des Bs angehäuft ist und große Insertionen von DNA, die ursprünglich aus Mitochondrien und Chloroplasten stammen. Alle auf dem B akkumulierten Sequenzen sind nicht B- spezifisch, sondern kommen auch in As mehrerer *Secale* Arten vor, wenn auch in geringerer Kopienzahl. Dies ist ein Hinweis auf den Ursprung des Bs von As derselben Familie.

Die auf dieser Arbeit beruhende zytologische Karte des Roggen-Bs liefert ein nützliches Werkzeug für zukünftige Studien zu diesem rätselhaften Chromosomentyp. Unter anderem haben wir die Möglichkeit etabliert, das B während der Interphase zu markieren. Unsere Erkenntnisse heben die Unterschiede zwischen A und B Chromosomen hervor. Obwohl Bs ursprünglich von As abstammen, haben sie seither einen eigenen Evolutionsweg eingeschlagen und sind daher ein wertvolles Werkzeug zur Erforschung der Chromosomenevolution.

Abbreviations

A	A chromosome
AFLP	amplified fragment length polymorphism
B	B chromosome
BAC	bacterial artificial chromosome
BLAST	basic local alignment search tool
BrdU	5-bromodeoxyuridine
bp	base pair
cDNA	complementary deoxyribonucleic acid
DAPI	4',6-diamidino-2-phenylindole
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide
EdU	5-ethynyl-2'-deoxyuridine
Fig	Figure
FISH	Fluorescence *in situ* hybridization
GAPDH	glyceraldehyde 3-phosphate dehydrogenase
GISH	Genomic *in situ* hybridization
ITS	internal transcribed spacer
kbp	kilo base pairs
LTR	long terminal repeat
Mbp	mega base pairs
MYA	million years ago
nm	nanometer
NOR	nucleolus organizer region
NUMT	nuclear insertion of mitochondrial DNA
NUPT	nuclear insertion of plastion DNA
PCR	polymerase chain reaction
RNA	ribonucleic acid
RT-PCR	reverse transcription - polymerase chain reaction
TE	transposable element
TIR	terminal inverted repeats

SCE	sister chromatid exchange
SIM	structural illumination microscopy
µm	micrometer

1. Introduction

1.1 B chromosome occurrence

Supernumerary chromosomes appear additionally to the standard chromosome set. These extra chromosomes have been termed B chromosomes (Bs) while the standard chromosomes are called auto- or A chromosomes (As). In most cases they do not convey any advantage for the host and can even be detrimental if Bs exceed a species-specific limit in number. To counter elimination, they do not follow Mendelian inheritance patterns, but instead often show selfish accumulation mechanisms, called 'drive' (reviewed in Jones & Rees, 1982; Jones, 1991; Jones & Houben, 2003; Jones, 2012).

Bs occur in plants, animals and fungi. There are over a thousand plant species and several hundred animal species reported with Bs. A comparative atlas of B-containing species has been given first by Battaglia, 1964 and later by Jones & Rees, 1982. After this, surveys have only been conducted throughout single taxa (Bedini *et al.*, 2011) found 5.3% of vascular plant species in Italy to contain Bs. One study looked for association between the presence of B chromosomes and properties of the genome such as acrocentricity of chromosomes, outbreeding, chromosome number and genome size (Palestis *et al.*, 2004b) and found correlations. In mammals in particular, B chromosomes seem to be more prevalent in acrocentric karyotypes (Palestis *et al.*, 2004a). The current known estimations of B occurrences are most likely underestimates, due to limited karyotypic information on many species, the fact that Bs are not present in all individuals of a species and some species have Bs only in certain organs (e. g. *Aegilops speltoides* harbours Bs only in aerial organs but not the root (Mendelson & Zohary, 1972)).

Due to the unusual nature of the B chromosome, it is often dismissed as inconsequential. This can mostly be attributed to the dispensable nature and the

apparent lack of adaptive significance. However, it is widely accepted that we can learn more from observing the abnormal, than from looking at the normal. This chromosome can teach us about evolutionary behaviour of chromosomes under varying selection pressure. Many Bs behave like a selfish entity and exploit cellular machinery to ensure their transmission and survival. There is still much to gain from these enigmatic chromosomes.

1.2 DNA composition and origin of B chromosomes

1.2.1 B chromosome origin

Molecular studies showed that in many species the Bs contain sequences that originated from one or different As (Houben *et al.*, 2001b; Page *et al.*, 2001; Cheng & Lin, 2003; Bugrov *et al.*, 2007; Martis *et al.*, 2012). Only few repetitive sequences are considered B-specific though they are also present on As in lower copy number. This points to an intraspecific origin of the B from the respective host genome. In some cases the origin of a B can be traced to a particular chromosome as in the case of *Locusta migratoria* where the B originated from chromosome 8 (Teruel *et al.*, 2010). The B of *Plantago lagopus* is a result of a rapidly altered trisomic chromosome 2 (Dhar *et al.*, 2002). In the genus *Phlox*, inversions are probably the origin of trisomic fragments that lead to the varieties of additional chromosomes observed (Meyer, 1944). Another possible origin for Bs is interspecific, as a result of hybridisation as proposed in the case of the fish *Poecilia formosa* (Schartl *et al.*, 1995).

Bs of a similar constitution and behaviour have formed in parallel but independent of each other in maize and rye. In contrast, two very different Bs have formed in *Brachycome dichromosomatica* (Carter & Smith-White, 1972). Interesting is that B emergence seems to correlate with large genomes, and low chromosome number (Palestis *et al.*, 2004b; Trivers *et al.*, 2004). This could hint at general events that lead to B formation for instance Bs as event during karyotype evolution or reducing the impact of aneuploidy.

Comparisons have often been made between Bs and sex chromosomes. Like Bs, Y chromosomes can be absent from the genome. It makes sense to assume similar evolution mechanics between both chromosome types. This also includes the emergence of a B from sex chromosomes. The B of the grasshopper *Eyprepocnemis plorans* has been shown to share many sequences with the X chromosome of the same species (Teruel *et al.*, 2009). A link between a B chromosome and the emergence of a neo-Y chromosome has been demonstrated in *Drosophila albomicans* (Zhou *et al.*, 2012). It has even been shown that the presence of a B chromosome can have influence on sex determination in cichlid fish (Yoshida *et al.*, 2011).

As the molecular processes that gave rise to Bs during evolution remain unclear, the characterisation of sequences residing on them might shed light on their origin and evolution. Until recently, little sequence data of Bs were available due to the problems of extracting the DNA of Bs from the rest of the genome. Earlier investigations on the DNA composition of Bs were mainly based on gradient density centrifugation (Timmis *et al.*, 1975), renaturation kinetics (Rimpau & Flavell, 1975), genomic *in situ* hybridization (Wilkes *et al.*, 1995), or comparative digestion of 0B and +B genomic DNA with restriction enzymes (Sandery *et al.*, 1990). Only in the last years, techniques like microdissection (Houben *et al.*, 2001a) and flow-sorting (Kubaláková *et al.*, 2003; Bartos *et al.*, 2008) have evolved enough to allow for confident isolation of B chromosome-derived DNA.

1.2.2 B chromosome translocations with As

Although Bs share sequences with the As, homologous pairing and recombination during meiosis are prevented. However, in rye, a translocation was found between the A chromosome 3R and the B (Schlegel & Pohler, 1994). Later it was shown that translocations with the As are rapidly counterselected and disappear from the population (Hasterok et al., 2002). Translocations between the rye B and wheat standard chromosomes have been utilised in studies about its nondisjunction

behaviour (Endo *et al.*, 2008). In maize, translocations between As and Bs are well documented (Birchler *et al.*, 1990) and have been used as a tool to study meiotic loss (Carlson & Roseman, 1992) and gene dosage compensation in hypoploid and hyperploid situations (Auger *et al.*, 2001).

1.2.3 Repetitive elements on Bs

Most Bs contain a high amount of repetitive elements, especially mobile elements such as DNA-transposons and retrotransposons. But also strong amplification of tandem repeats and satellite DNA has been observed, particularly in plants (Langdon et al., 2000; Dhar et al., 2002). The micro Bs of *B. dichromosomatica* are mainly composed of tandem repeats (Houben et al., 2001b). The high content of repetitive elements might facilitate the strong heterochromatinisation observed in many Bs. Since Bs are subjected to less selection pressure, mobile elements will accumulate to a greater proportion than on the As. Because the Bs are under relaxed selective pressure, they present a kind of safe spot for mobile elements. But as they take, by their definition, a separate way of evolution, the composition of the repetitive fraction is expected to differ from that of the As. This is especially apparent in species with already high content of mobile elements in the standard genome. An example are grasses, of which maize and rye have been of particular interest to the studies of Bs (Jones & Houben, 2003). Often, Bs contain repeats which are either specific to the supernumerary fraction and do not occur in the standard component of the genome, or are highly accumulated in the B fraction (see B-specific sequences).

1.2.4 Ribosomal DNA on Bs

Many Bs contain ribosomal DNA. Dhar et al. (2002) observed the formation of a new B in *Plantago lagopus* which involved a massive amplification of 5S rDNA. This specific B contains also active nucleolar organising regions (NOR) (Dhar et al., 2002). Inactive 18S rDNA was described in telomeric and centromeric regions of the Bs of the fish *Haplochromis obliquidens* (Poletto *et al.*, 2010). Even the tiny micro Bs of *B. dichromosomatica* contain 45S rDNA (Donald *et al.*, 1995; Houben *et al.*,

1997b). But due to the heterochromatic nature of Bs and their localization outside the nucleolus, these rRNA genes are in most cases inactive (Marschner *et al.*, 2007b), although also Bs with active NORs have been observed (Leach *et al.*, 2005). Bs do not obligatory carry rDNA, since there exist also Bs without rDNA, like those of rye (Puertas, 2002).

1.2.5 Single copy genes on Bs

Single copy genes have been found rarely on Bs. This could be due to a problem of finding them between the massively amplified mobile elements or due to the high sequence similarity between A- and B-located sequences. But many Bs have been suggested to contain even active genes due to visible phenotypical effects on the host plant (Pedro & Camacho, 2005) or active participation in a drive mechanism (Puertas, 2002). In maize, several B located sequences have been isolated that are homologous to non-coding regions of genes located on the As (Cheng & Lin, 2003). Histone genes have been found on the Bs of the migratory locust *Locusta migratoria*, although these genes appear in a repetitive fraction and often carried deleterious mutations (Teruel et al., 2010). Transcriptional activity of genes located on Bs has been shown in rye together with an alteration of the transcription of the As (Carchilan et al., 2009).

1.2.6 B-specific sequences

In many species sequences have been determined which are specific only to the Bs of that species. In *B. dichromosomatica*, the two repeats Bd49, which is specific to the centromere of the large B (John et al., 1991) and Bdm29, which is interspersed on the whole micro B of this species (Houben et al., 2001b) have been isolated. In rye the B-specific sequences D1100 (Sandery et al., 1990) and E3900 (Blunden et al., 1993) were found. Both are located at the end of the long arm of this B type (Wilkes et al., 1995) in a region supposed to contain one or more elements controlling the non-disjunction process of the rye Bs. Both sequences have been shown to be transcriptionally active (Carchilan et al., 2007). A high number of

rearrangements in the B-specific parts of the rye B has been reported, resulting in their distinctive current sequence organization (Langdon et al., 2000). The mechanism of accumulation of these different sequences is not yet known. In plants with high content of repetitive sequences it is conceivable that for example random sequences from the As get captured during the activity of mobile elements (Wicker *et al.*, 2010). Several sequences specific to the B have been isolated in maize (Alfenito & Birchler, 1993; Stark et al., 1996; Lin & Chou, 1997). Also in maize a sequence called starkB, which contains A derived as well as B specific sequences was found to be transcriptionally active (Lamb et al., 2007). Other organisms from which B-specific sequences have been identified are the wasp *Nasonia vitripennis* (Nur et al., 1988), the fish *Alburnus alburnus* (Ziegler et al., 2003), the raccoon dog *Nyctereutes procynoides* (Trifonov et al., 2002) and the greater glider *Petauroides volans* (McQuade et al., 1994).

1.3 Chromatin composition of B chromosomes

A strong tendency to heterochromatinisation is often observed for Bs. Strong condensation throughout the cell cycle is a feature frequently observed in many Bs. In the grasshopper *Eyprepocnemis plorans* the heterochromation was shown to consist of two repeats, one being AT-rich, the other GC-rich, just like the As of this species (Teruel et al., 2009). A nearly completely heterochromatic B was reported in *P. lagopus* (Dhar et al., 2002). But there are also Bs containing larger euchromatic regions, e.g. that of maize (Carlson, 1978). The large Bs of *B. dichromosomatica* are similar to their respective As in heterochromatin content (Houben *et al.*, 2003). A completely euchromatic B was reported in *Characidium cf. Zebra* (Venere et al., 1999). Recently, the heterochromatin content of the maize B has been examined closer in order to reveal the evolutionary development of three heterochromatic regions (Cheng, 2010). Heterochromatinization of Bs has been linked before in *Puschkinia libanotica* with a reduction in euchromatic histone marks rather than an increase in heterochromatic marks (Kumke *et al.*, 2008).

Epigenetic histone marks have not yet been intensively studied in Bs. However it has been shown that large Bs in *B.dichromosomatica* are less acetylated on histone H4 than the As (Houben et al., 1997a). This might be related to inactivation of ribosomal genes. In the same organism a reduction of euchromatin-typical methylation marks of histone H3 has been reported. The heterochromatic marks were similar between A and B which matches the observation of similar heterochromatin distribution (Marschner *et al.*, 2007a). In *P. libanotica* no difference was found regarding heterochromatic marks, but a reduction of euchromatic marks was found in the Bs (Kumke *et al.*, 2008). In rye the heterochromatic terminus of the long arm of the B, which also contains the two B-specific sequences, has been reported to contain both the heterochromatic histone marks H3K9/27me as well as the euchromatic mark H3K4me. However it is unclear whether these marks reside on the same nucleosome or are interspersed with each other (Carchilan et al., 2007).

1.3.1 B effects on DNA replication

Sequence composition and repetitive content can have an effect on DNA processing. Indeed DNA replication of Bs is delayed compared to As replication, as reported in black rat *Rattus rattus* (Raman & Sharma, 1974), the fox *Vulpes fulvus* (Świtoński *et al.*, 1987), the fish *Astyanax scabripinnis* (Maistro *et al.*, 1992) and the amphibian *Gastrotheca espeletia* (Schmid *et al.*, 2002) though it is not a rule. In *B. dichromosomatica* the replication of large Bs occurs late in S phase (Houben *et al.*, 1997a). In contrast, the replication of micro Bs, which are heterochromatic, occur during the entire S-phase (Marschner *et al.*, 2007a). Late replication was also shown for heterochromatic parts of the maize B and was even suspected to be responsible for the nondisjunction mechanism (Pryor *et al.*, 1980). For the rye B, it was shown that the heterochromatic part replicates latest within the chromosome (Lima-De-Faria & Jaworska, 1972; Ayonoadu & Rees, 1973).

1.3.2 B chromosomal structure: centromere and telomere

As most Bs are mitotically stable, they should include functional centromeres. A B was reported in the fly *Megaselia scalaris* which appears to be little more than an individual centromere (Wolf et al., 1991). The centromere of maize Bs has been particularly well researched. This led to the identification of repeats specific to the maize B centromere, called pZmBs (Alfenito & Birchler, 1993) and pBPC51 (Cheng & Lin, 2003). Only about 700 kbp of the ZmB sequence are sufficient for interaction with the centromeric histone variant CenH3 in order to create a fully functional centromere (Jin et al., 2005). In rye the pericentromeres appear to be of special importance. They harbour sequences which during the non-disjunction process provide chromatin adhesion sides activated *in trans* by an element situated at the end of the long arm of the B. The maize B appears to carry similar units acting in *trans* upon the sticking sides. Telomeres are required for a stable chromosome. The *de novo* addition of telomeres was reported by Dhar et al. (2002) while observing the formation of a new B. A sequence isolated from the maize B centromere included a conserved telomeric sequence and also shared low homology to centromeric sequences from maize chromosome 4 as well (Qi et al., 2002).

1.4 The B chromosome of rye

1.4.1 The distribution of Bs in the genus

One of the best models for a selfish chromosome is the B chromosome of rye (Jones & Puertas, 1993). *Secale cereale* L. Bs have been found in accessions from many parts of the world. They are most likely all of monophyletic origin and very stable, as they appear very similar even in other rye subspecies such as *S. cereale segetale*, which is very closely related to *S. cereale ancestrale* (Niwa & Sakamoto, 1995). However, they have not been reported in older rye species such as *S. strictum* or *S. sylvestre*. This hints at an origin close to the divergence point of the *S. cereale* clade from *S. strictum*, which is supported by molecular dating (Martis *et al.*, 2012). As Bs are often morphologically different to their respective As (Jones & Rees, 1982), the rye B is acrocentric while the standard chromosomes are metacentric. About half the size of an A, the rye B harbours a 1C content of around 560 Mbp (Martis *et al.*, 2012).

1.4.2 Secale evolution

The evolution of the genus *Secale* is very difficult to reconstruct. The whole genus is evolutionary very young and the species closely related. Several studies have tried to resolve the phylogenetic tree with different methods, such as rDNA ITS sequences (De Bustos & Jouve, 2002), AFLP marker (Chikmawati *et al.*, 2005) and microsatellite marker (Shang *et al.*, 2006). All studies show problems in distinguishing the different species. This is most likely due to the young age of the *Secale* genus of around 1.7 MYA. The split between the clades of *S. cereale* and *S. strictum* was estimated at 0.8 MYA (Martis *et al.*, 2012). In consistence with Frederiksen *et al.*, we decided to distinguish three major species: *S. cereale*, *S. strictum* and *S. sylvestre* with *S. cereale segetale* and *S. cereale ancestrale* as subspecies.

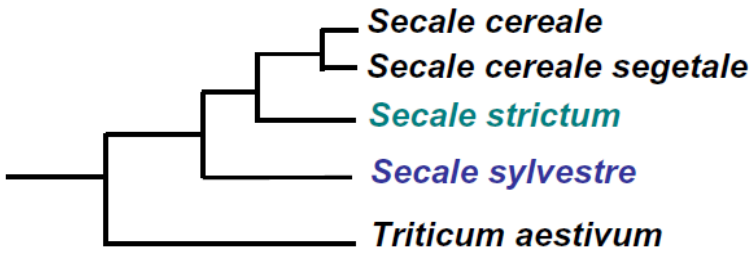


Fig. 1: Simplified phylogenetic tree of the genus *Secale* with wheat as an outgroup.
We distinguish three main species: *S. cereale, S. strictum,* and *S. sylvestre.* Comparative tree based on (Frederiksen & Petersen, 1998; De Bustos & Jouve, 2002; Chikmawati *et al.*, 2005; Shang *et al.*, 2006).

B chromosomes have until now only been found in the species *S. cereale*, as well as the subspecies *S. cereale segetale* and *S. cereale ancestrale* (Niwa & Sakamoto, 1995). So the common expectation is that the B chromosome arose during the split of *S. cereale* from *S. strictum*. Already early it was proposed, that rye has only one standard B (Müntzing, 1945) and that the reported deviations (e.g. (Müntzing, 1944)) are most likely isoforms of this standard B.

1.4.3 DNA composition of the rye B

In rye, early experiments showed that the ratio and heterogeneity of repeats in DNA of Bs do not differ from As. However, a slight increase in cytosine and guanine content was observed for DNA of B carrier plants (Rimpau & Flavell, 1975; Timmis *et al.*, 1975). Later, a B-specific sequence fragment (Tsujimoto & Niwa, 1992) and the B-specific sequences D1100 (Sandery *et al.*, 1990) and E3900 (Blunden *et al.*, 1993) were found. They serve typically as controls to identify Bs in microscopic pictures (Fig. 2). The terminal heterochromatin marked by these sequences also contains sequences shared with the As (Houben *et al.*, 1996) but is so strongly enriched in B specific sequences that it is unmarked in GISH experiments using 0B DNA (Tsujimoto & Niwa, 1992).

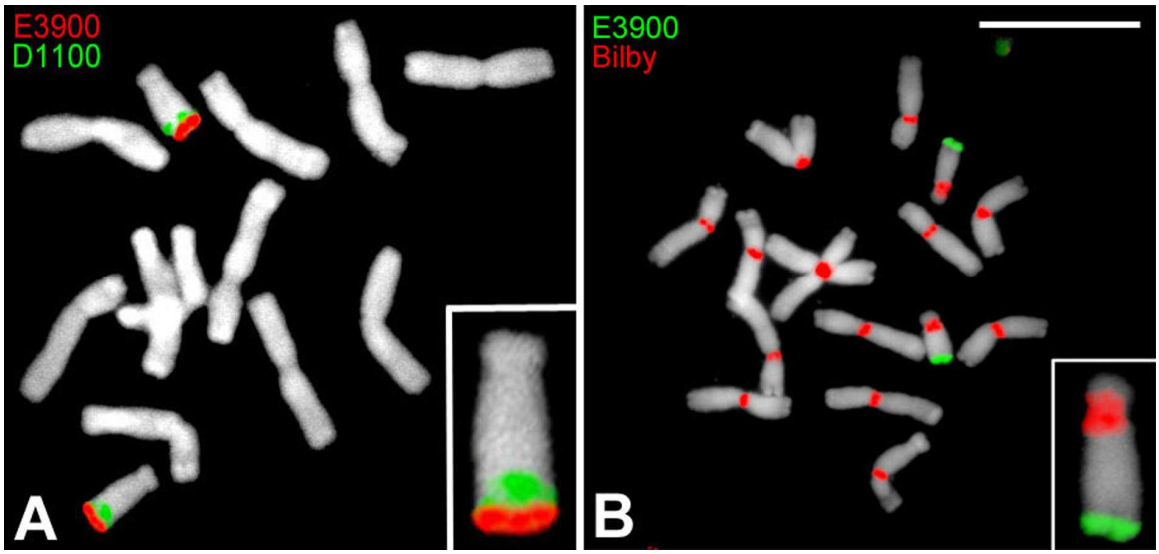


Fig. 2: Established sequences for the rye B.

A) The repeats E3900 (red) and D1100 (green) mark the terminal part of the long arm and make the B chromosome easily identifiable. B) The sequence E3900 (green) identifies the B chromosome, while the rye specific centromere marker *Bilby* (red) is present on both As and Bs. Bar indicates 10 µm.

The Giemsa-banding positive, heterochromatic distal part of the Bs in rye partly undergoes decondensation during interphase (Carchilan et al., 2007). The Giemsa-

banding positive structure might be caused by the large proportion of repetitive elements.

Recently, we were able to flow-sort the B of rye from its respective A chromosomes due to its considerably smaller size. Employing next generation sequencing we determined the DNA composition of Bs and the respective As (Martis *et al.*, 2012). The origin of rye B chromosomes was estimated to about 1.1-1.3 million years ago, thus overlapping in time with the time of radiation of the genus *Secale* (1.7 million years ago). Based on the obtained information we proposed that Bs originated of A chromosome rearrangement events during early evolution of rye. Multiple chromosomal rearrangements involving A chromosomes 3RS and 7R, accumulation of repeats and genic fragments derived from other A chromosomal region, as well as insertions of organellar DNA shaped the rye B (Martis *et al.*, 2012).

1.4.4 Drive mechanism of the rye B

The maintenance of Bs in natural populations is possible by their transmission at higher than Mendelian frequencies. The accumulation mechanism of the rye B by nondisjunction requires a factor located at the end of its long arm and the pericentromere (Müntzing, 1948; Håkanson, 1959; Endo *et al.*, 2008; Banaei-Moghaddam *et al.*, 2012). Two B-specific repeat families E3900 (Blunden *et al.*, 1993) and D1100 (Sandery *et al.*, 1990) reside in the long arm terminal nondisjunction control region. Expression analysis revealed that D1100 and E3900 are highly transcriptionally active in anthers. Since both repeats form noncoding RNA, it was proposed that these transcripts itself are involved in the nondisjunction process of the B. In addition, the distal heterochromatin is marked with the euchromatin-specific histone modification mark H3K4me3 (Carchilan *et al.*, 2007).

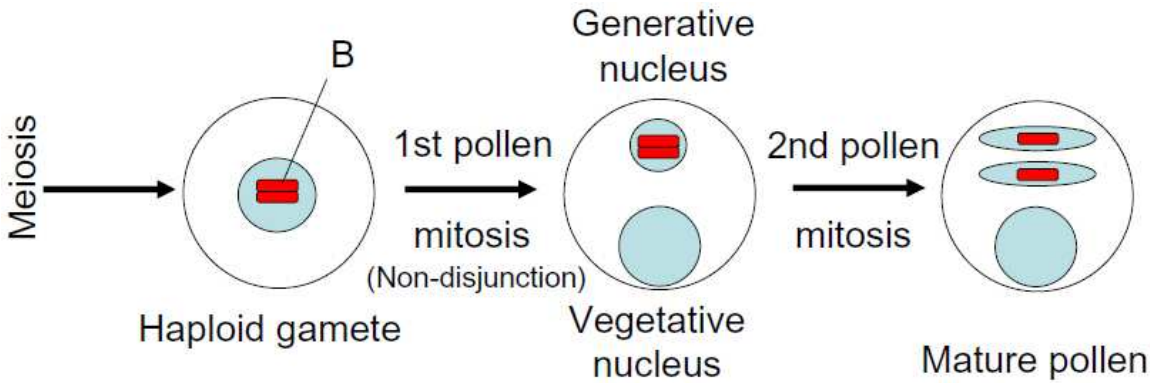


Fig. 3: Nondisjunction mechanism of the rye B

During the first pollen mitosis, the sister chromatids of the B chromosome do not separate and migrate preferentially to the generative nucleus. This drive mechanism allows active accumulation of Bs to counter selective pressure for loss.

1.5 Open questions and aim of the thesis

In this study we used the data from the sequencing of isolated rye Bs (Martis *et al.*, 2012) to gain insight into the high-copy DNA composition of the rye B. We searched for similarities and differences in mobile element, satellite repeat and nuclear organellar DNA integration content between As and Bs. Up to now, only broad surveys of the rye genome have been made. This is the first time we could take a close look at the make up of the rye B. Especially we looked at B-specific sequences, to establish a better cytological map of the B for future research. This map was then used to answer questions about structure and behaviour of the rye B, such as centromere composition and replication behaviour. Known B behaviours such as late replication were confirmed and mapped with the help of the novel B-specific sequences. Employing GISH with 0B DNA on +B plant material we determined the exact location of the border between B-specific and A-shared sequences.

2. Materials and Methods

2.1 Plant material and plant cultivation

Plants with Bs from the self-fertile inbred line 7415 of rye (*Secale cereale* L.) (Jimenez *et al.*, 1994) and the related wild rye *Secale strictum* (C. Presl) C. Presl subsp. *strictum* s. l. (Gatersleben Genbank accessions R548) and *Secale sylvestre Host* (Gatersleben Genbank R1116) as well as hexaploid wheat (*Triticum aestivum* L.) with added standard rye Bs (Lindström, 1965) were grown at the same temperature, humidity, and light conditions (16 h light, 22ºC day /16ºC night) in a green house.

2.2 Genome size measurement by flow cytometry

Relative DNA content of leaf nuclei was measured using a FACStar^{PLUS} flow cytometer (BD Biosciences, San Jose, CA, USA) equipped with an argon-ion laser INNOVA 90C (Coherent, Santa Clara, CA, USA). *Pisum sativum* 'Viktoria, Kifejtö Borsó' (Genbank Gatersleben accession number: PIS 630; 1C = 4.54 pg (Dolezel *et al.*, 2007)) corresponding to 4,445 Mbp (Dolezel *et al.*, 2003) was used as an internal reference standard. Small pieces of young leaf tissue of rye and *Pisum* were co-chopped with a razor blade in ice-cold nuclear isolation buffer (45 mM magnesium chloride, 30 mM sodium citrate, 20 mM 4-morpho-linepropane sulfonate, Triton X-100 $1^{mg}/_{ml}$) (Galbraith *et al.*, 1983) supplemented with 0.1% Triton X-100, 1% polyvinylpyrrolidone, DNase-free RNase (50 mg/ml) and propidium iodide (50 mg/ml) and filtered through a 35 µm mesh. Usually 10,000 nuclei per sample were analyzed in at least three replicates per line. The absolute DNA content of plants with 0B, 2Bs and 4Bs were calculated based on the ratios of the G1 peak means of sample and reference standard.

2.3 *In silico* identification of repeats

Transposable element (TE)-related sequences were identified by BLASTN search of individual 454 reads against the TREP repeat database (wheat.pw.usda.gov/ITMI/Repeats/) as previously described (Wicker *et al.*, 2009). In addition, *de novo* identification of repetitive sequences was performed (Martis *et al.*, 2012) using graph-based clustering of 454 sequence reads (accession number EBI-ENA ERP001061). The structure of cluster graphs was investigated using SeqGrapheR program (Novak *et al.*, 2010) and contig sequences, assembled from reads representing selected parts of cluster graphs, were used to design probes for FISH. These probes were PCR-amplified from genomic DNA, cloned and verified by sequencing. Probe sequences are available from GenBank under accession numbers KC243218 - KC243251 for the terminal satellites and KC243252 - KC243260 for pericentromeric sequences. The probes corresponding to novel repeats were named based on cluster and contig numbers (e.g. Sc9c11 is *S. cereale* repeat represented by contig 11 from cluster 9).

2.4 Probe generation for FISH

Probes for FISH were generated by PCR with *Taq* polymerase (Quiagen) from template DNA isolated from plants containing Bs. Annealing temperature was 58℃. The primers used are listed in table 1. PCR p roducts were purified using Fermentas GeneJet PCR Purification Kit (#K0702). For organellar DNA, BAC DNA was used with chloroplast (ChHB040G01) and mitochondrial (HVVMRXALLhA0) sequences from barley. Additionally, we isolated mitochondria from rye as described (Triboush *et al.*, 1998). For this purpose leaf material was homogenized in STE buffer and filtered before centrifugation for 20 min at 3,900 g. The supernatant was centrifuged twice at 18,000 g for 20 min and the pellet resuspended in ST buffer. After treatment for 20 min with 125 µg DNase I (Thermo Scientific) in 200 µl ST buffer (reaction stopped with EDTA), mitochondria were washed in NETF buffer and centrifuged again at 18,000 g for 20 min. The in TEN buffer resuspended mitochondria were lysed with 20% SDS

in H_2O and equilibrated with phenol. The resulting DNA was cleaned by chloroform/isoamylalcohol purification and amplified by templiPhi (illustra™ Code: 28-9642-86) before labeling.

For subtelomeric repeats we used established plasmids with pSc200 (Vershinin *et al.*, 1995) and pSc119.2 (McIntyre *et al.*, 1990). Isolation of genomic DNA was performed with Quiagen DNeasy Plant Maxi Kit (Cat. No. 68163) or CTAB extraction, BAC isolation with Quiagen-tip 100 (Cat. No. 10043) according to plasmid extraction protocol provided by the producer, plasmid DNA was isolated with the help of the Quiagen Plasmid Mini Kit (Cat. No. 12125).

2.5 Isolation of genomic DNA

In CTAB extraction about 1 g of leaf material was homogenized in liquid nitrogen and incubated with 15 ml of 2 x CTAB buffer for one hour in a water bath at 65℃. After letting the tube cool down to room temperature, 20 ml chloroform/ isoamylalcohol (24:1) were added and it was shaken for 15 min. The probe was then centrifuged for 15 min at 8000 rpm and the supernatant mixed with $^2/_3$ volumes of isopropanol to hook out the DNA. The DNA was washed twice with ice cold 70% ethanol, eluted in 500 µl water and treated with 5 µl RNase A (10 $^{mg}/_{ml}$). The resulting extract was cleaned with phenol, phenol/ chloroform (1:1) and chloroform and precipitated with ethanol over night at -20℃.

ST buffer	100 ml
Tris pH 7,8	0.606 g
Saccharose	13.69 g
BSA	100 mg

NETF buffer	500 ml
1.25 M NaCl	36.52 g
EDTA	50 ml of 0.5 M
50 mM TRIS	3.03 g

STE buffer	500 ml
Tris pH 7,8	3.03 g
Saccharose	68.46 g
EDTA	20 ml of 0.5 M
β-Mercaptoethanol	20 µl per 10 ml buffer
BSA	20 mg per 10 ml buffer

TEN buffer	100 ml
Tris pH 7,2	1.211 g
EDTA	10 ml of 0.5 M
NaCl	0.584 g
β-Mercaptoethanol	200 µl

2 x CTAB buffer	500 ml
100mM Tris pH 8.0	50 ml of 1M
20mM EDTA ph 8.0	50 ml of 0.5M
2% CTAB	10 g
1.4M NaCl	40.9 g
0.5% Na-bisulfide	2.5 g
1% β-mercaptoethanol	5 ml

2.6 Nick translation of FISH probes

Probes were labeled with ChromaTide Texas Red-12-dUTP, Alexa Fluor 488-5-dUTP (http://www.invitrogen.com) or Cy5-dUTP (GE Healthcare Life Sciences, http://www.gelifesciences.com) by nick translation. Nick translation was performed using 1 – 3 µg DNA, NT buffer, dNTP with decreased dTTP, β-Mercaptoethanol, labeled dUTP, 0,025 U DNase (Fermentas) and 20 U Polymerase I (Thermo Scientific). Nick translation was performed for 90 min at 15℃ followed by 10 min at 65℃.

10x NT-buffer	100 ml
0.5 M Tris-HCl pH 7,5	50 ml of 1 M
50mM $MgCl_2$	5 ml of 1 M
0.05% BSA	50 mg

dNTP mix

0.5 mM dATP	100 µl of 2 mM
0.5 mM dGTP	100 µl of 2 mM
0.5 mM dCTP	100 µl of 2 mM
0.1 mM dTTP	20 µl of 2 mM

Table 1: Primer sequences for FISH probe amplification

Primer sequence	Primer name
CAAGACATGCTCACGCTCAG	Sc26c38_F:
ACAGATGTAGAAGTATGTGGAGAAAC	Sc26c38_R:
CGCACTTCCGAGTAACCTGT	Sc26c38_R2
CCTTGGTGTGGCTGATGTTA	Sc36c82_F:
GCAGTCGGGCTCATATAGGT	Sc36c82_R:
TACTTGGAACCATGTTGTGG	Sc63c34_F:
TGCTTCGTACTAATCGTGCTTC	Sc63c34_R:
AATGTTCCTTAGCTTGCCAGA	ScCl11-1_F
TTTTCGTGCTTCCATCAACC	ScCl11-1_R
CGCATTAGGAACAGGAGGTC	ScCl11-2_F1
ATGACCCCAACAAGAGGTTG	ScCl11-2_R1
TATGCGAAGCTGGAAGATGA	Sc55c1_F
GCTTCCATGCATCCTCAAAT	Sc55c1_R
GACAATCCGATGGAAGAGGA	ScCl11-2_F2
AGGTGAGAGGGCCTCAAAAT	ScCl11-2_R2
CGAAGCCAACTTCAACCTTC	ScCl11-2_F3
TCCTCTTCCATCGGATTGTC	ScCl11-2_R3
TTAGATGGCTACAAATTATTGTAATGC	Sc9c11-F1
AGGCAATCAATGTCAGCTTCC	Sc9c11-R1
CGGAGATGGTGTAGTTTACAAGC	Sc9c11-F2
ACATGAAATCATGCCCAGAAG	Sc9c11-R2
GCCATCTTGATTCAGAGGTACAC	Sc9c15-F1
AAAGCACAAGTAGACCTCACCAC	Sc9c15-R1
GATACTTTGGACTCGCACATACTC	Sc9c15-F2
GTTGGTGTGGATATGATAATTGATG	Sc9c15-R2
CCATACATACCGTATTACACAAATAGG	Sc9c461-F1
TCACTTGAGAAAGAAGTGATGTGC	Sc9c461-R1

Primer sequence	Primer name
GCATGTCATCGGTAGGATAGG	Sc9c130-F1
ACCCCTTCCCTTTCGATCTAC	Sc9c130-R1
TGGCAAGCAAGAACTAGTAGACAG	Sc21c9-F1
TGAACCACTGAGCTTATCCTAGC	Sc21c9-R1
CAATGATTTGATTAGCATAACAGGTG	Sc21c67-F1
GATTTACAAATTAACCGTCATCTATGC	Sc21c67-R1
CAATGCTAGCTGCACCACCAACTG	GAPDH_F
CTAGCAGCCCTTCCACCTCTCCA	GAPDH_R
AGGAGAAGACTCACATCAACATC	EF1a_F
TGGGCTCGTTGATCTGGTCA	EF1a_R
ACATGGAATATGCTGGCAATG	Sc11c759-F1
AAGCTTGTGACATGTACCTTCTG	Sc11c759-R1
TTTGCGACAATGACTCAAGC	Bilby F1
TGTAGCTCATCGTGGAGTCG	Bilby R1
TCACGGGATAGAACACCACA	Bilby F2
ATCGGTTGGTGAGGTTTGAG	Bilby R2
CAGCGCATGTTAGAGGATCA	Caspar F1
GTTTCTTGCGTCATCTGCAA	Caspar R1
GCATTGAGGGGCTTCTACAG	Caspar F2
AGCCACCAAGATGGAATTTG	Caspar R2
GGCATCTGTCCTGGTTCATC	Clifford F
TTGATCCAACGGCATAATGA	Clifford R
GAGGCATTCGCCTACTGAAC	Jorge F
ACGGGACTAAACGGGCTACT	Jorge R
TACCGATGACAAAGGGAACA	Angela F
CCCCCTTCCATGTAGTAGGA	Angela R
AGCCATTGGAAATCTCATGC	Maximus F
TCACCATATTCCCCTTGAGC	Maximus R
CGACATTGGTGCTTTCATTG	Fatima F
CTAAAGGTGAGTGCCGGAAG	Fatima R
ATCACCTCTCCTCCGTTCCT	Revolver F1
GGAGGGAGATAGCCTGGAAG	Revolver R1
ATACTCGGGTCTTCCCACCT	Revolver F2
GTCGAAGCGTCACAAACAAA	Revolver R2
GTCCACCGTCGCAATTCTAT	Sabrina F
TCCTCGATGGCAACAATACA	Sabrina R
GCCGTGGTGTTGTATTTGTG	WHAM F
CAAGCACTAAGCATGGCAAA	WHAM R
GCGTGTCGTGAGAAACAAAA	5SrDNA_F1
TCCACCCCTCCCTCTATACC	5SrDNA_R1
GGTCTTGAAGGTGGCAAGAG	5SrDNA_F2
GGAATGCAACACGAGGACTT	5SrDNA_R2

2.7 Cloning and sequencing of PCR products

For cloning we ligated achieved PCR products into the pSC-A vector (StrataClone™ from Stratagene #240205) and, after amplification through transformation in *E. coli*(StrataClone™ SoloPack® Competent Cells), verified the insertion by enzymatic digestion with *Eco*RI. For every element, at least 5 clones were sequenced within the institute. Sequence analysis was performed with the Software SeqMan Pro from the DNASTAR Lasergene 10 Core Suite.

2.8 Preparation of mitotic and meiotic chromosomes for FISH

Mitotic chromosomes were prepared by a method modified after (Kato, 1997): washing of root meristems in ice cold water and washing in cold 1x citric buffer, two times each, then digestion in an enzyme cocktail (1% cellulose, 1% pectolyase Y-23, 1% cytohelicase in 1x citric buffer) for 1 h at 37℃. Afterwards root tips were washed again in citric buffer and in ice cold ethanol before transfer to an appropriate amount (depending on the number of root tips) of dropping solution (75% acidic acid, 25% methanol). Tissue was disrupted with a dissection needle and 7 µl cell solution was dropped on each slide. For dropping, slides were put on a heating plate at 45℃ with a wetted paper towel to ensure high air humidity.

For pachytene preparations, anthers were dissected and checked for appropriate meiotic stadium, then fixed with ethanol/acetic acid (3:1) for at least two days. After washing 3 times in water and 1x citric buffer, anthers were digested with enzymes (0.3% cytohelicase, 0.3% pectolyase, 0.3% cellulose in citrate buffer) at 37℃ before enzymes were washed off with citrate buffer. A single anther was used per slide and covered with 7 µl 60% acetic acid. After homogenizing the material with a needle, another 7 µl 60% acetic acid were added and incubated for 2 min. A further homogenization step was performed with another 7 µl 60% acetic acid and the slide was put on a 42℃ hot plate for 2 min. On the hot plate, the material is spread by hovering a needle over the drop without touching the slide, using the superficial tension. Afterwards the acetic acid drop was surrounded by ethanol/acetic acid (3:1) fixative at 4℃ to precipitate the nuclei

suspension on the slide. More fixative was added to precipitate the rest of the cells and the slide washed shortly in fixative, then washed 10 min in 60% acetic acid and finally washed in 100% ethanol. After air drying, slides were kept at 4℃ until hybridization.

2.9 Fluorescence in situ hybridization (FISH)

FISH and GISH was performed according to (Ma et al., 2010) and (Gernand et al., 2006). Chromosomal DNA was denatured at 80℃ for 3 min on untreated slides or for 8 min on slides prelabelled with EdU. GISH was performed with labeled genomic DNA from plants with 0B while blocking with genomic DNA from plants with 4B in a ratio of 1:20 to 1:40. Imaging was performed by using an Olympus BX61 microscope and an ORCA-ER CCD camera (Hamamatsu). Deconvolution microscopy was employed for superior optical resolution of globular structures. Projections (maximum intensity) were done with the program AnalySIS (Soft Imaging System). All images were collected in grey scale and pseudocoloured with Adobe Photoshop CS5 (Adobe). To achieve an optical resolution of >100 nm, we applied structural illumination microscopy (SIM) using a 363/1.40 objective of an Elyra super-resolution microscope system (Zeiss, Germany).

10x citric buffer

100 mM citric acid	40 ml
100 mM tri-sodium citrate	60 ml

2.10 DNA replication and sister chromatid exchange analysis

For studying late replication, roots of young seedlings were treated for 2 h with 15 µM EdU (5-ethynyl-2'-deoxyuridine, Click-it EdU Kit, Invitrogen), followed by water for 2.5 h. Fixation of roots was performed with ethanol/acetic acid (3:1) for several days.

For the sister chromatid exchange (SCE) analysis, the cut off roots were then treated for 7.5 h (8.30 AM to 4.00 PM) with 15 µM EdU. The roots were then put into water for 16 h (4.00 PM to 8.00AM), followed by treatment with 0.05% colchicine for maximal 3 h (8.00 AM to 11.00 AM). Fixation was performed with ethanol/acidic acid 3:1 for several days.

After preparation of chromosomes, the so-called click reaction was performed to detect EdU according to the kit protocol (baseclick GmbH). For FISH after replication analysis, slides were washed in 2x SSC for 10 min, then dehydrated in 70%, 90% and 100% ethanol for 3 - 5 min each, and fixed in ethanol/acetic acid (3:1) for at least 1 h. After washing in 2x SSC and fixing with 4% paraformaldehyde for 5 min, FISH was performed as described above.

20 x SSC (pH 7.0)	1 l
0,3 M NaCl	175.3 g
0,3 M Na-citrate	88.24 g

2.11 Southern Blotting

2.11.1 membrane preparation

Restriction enzyme-digested genomic DNA was run on a 1% agarose gel at 26 V over night to ensure well separated bands and reference photos were made. After, gels were rinsed in H_2O, denatured in denaturation buffer for 30 min, rinsed again and neutralized for 2 x 15 min in neutralization buffer. The blotting arrangement was assembled as follows: the gel was put face down on a blotting platform covered by filter paper with the ends submerged in 20 x SSC buffer, the membrane (Hybond-N from Amersham, Code RPN303S) on top, 3 more filter

paper (Whatmann) and 5 cm paper towels weighted with a 1 kg weight. DNA transfer to the membrane was performed over night. After rinsing agarose rests off the membrane in 2 x SSC, it was put face-up on a blotting platform with 0.4 M NaOH for 20 min for denaturation. The membrane was then rinsed shortly in 5 x SSC and stored in the fridge wrapped in plastic wrap.

Denaturation buffer	2 l
1.5 M NaCl	175.32 g NaCl
0.5 M NaOH	40.01 g NaOH

Neutralisation buffer	2 l
1.5 M NaCl	175.32 g NaCl
1M Tris HCl (pH7,2)	242.28 g Tris
0.001 M EDTA	3 ml of 0.5M EDTA

2.11.2 Southern blot hybridization

DNA probes were labeled with Decalabel DNA labeling Kit (Fermentas) according to protocol. Hybridisation was performed overnight at 65℃ in Church buffer and membranes were washed at 65℃ for 3 x 20 min in 0.5% SDS; 2x/ 1x/ 0.5x SSC respectively. Hybridization signals were detected with the Phosphoimager FLA-5100 from Fujifilm and the corresponding Software (Image Reader FLA5000 v3.0).

Church buffer	1 l
0.25 M Sodium Phosphate buffer	500 ml of 0.5 M
1 mM EDTA	2 ml of 0.5 M
1% BSA	10 g
7% SDS	70 g

2.12 Isolation of RNA

For RNA isolation a young leaf was homogenized in liquid nitrogen and incubated 5 min in 1 ml trizol before adding 200 µl chloroform and incubating for another 3 min. After centrifugation for 15 min at 1200 rpm and 4℃, the supernatant was mixed with 500 µl isopropanol and incubated min on ice. The RNA was centrifuged again for 10 min at 1200 rpm and 4℃ and the pellet washed with 70% ethanol. The RNA pellet was resuspended in DEPC water and treated with DNase (Ambion Turbo DNA-free Kit AM1907; Applied Biosystems).

2.13 RT-PCR

For transcription analysis, we performed RT-PCR on cDNA made (RevertAid H Minus first-strand cDNA synthesis kit, Fermentas) from total RNA extracted from root, leaf and anthers of plants with and without Bs. Annealing temperature was 58℃. The primers used are listed in table 1.

3. Results

3.1 The DNA content of the rye B chromosome

The size of the rye B has previously been measured with 5.5% of the total A chromosome size by Feulgen photometry (Jones & Rees, 1968). For a genome size of 8 Gbp this means a 440 Mbp DNA content for one unreplicated B (Jones & Houben, 2003). To confirm these results we determined the DNA content of plants with and without Bs flow cytometrically. Flow cytometry allows the analysis of thousands of nuclei within a short time and results in more precise values. From measurements of plants with 0B, 2Bs and 4Bs we extrapolated the 1C size of a single rye B to be ~570 Mbp (Fig. 4).

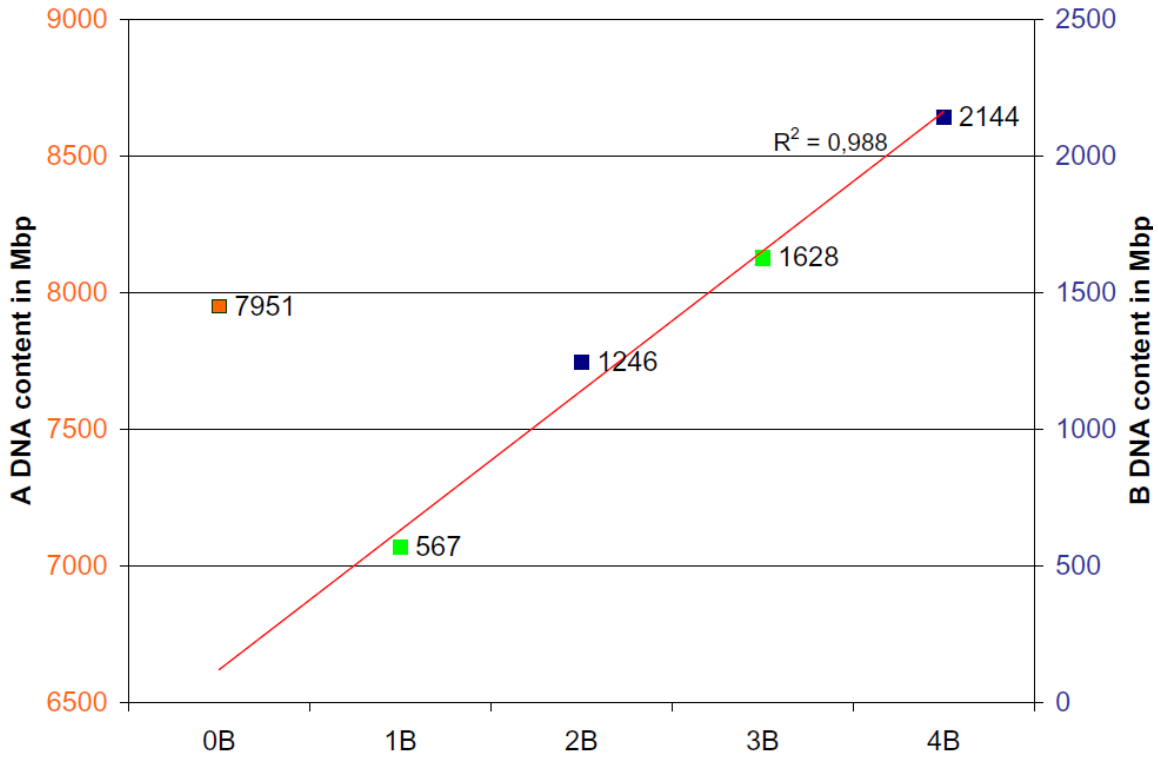


Fig 4: B size measurement by flow cytometry
The size of the rye B was estimated by measuring the DNA content of plants with 0B, 2Bs and 4Bs by flow cytometry. Green data points are estimations, while blue data points represent measurements of the size of 2 Bs and 4 Bs respectively. For comparison, the genome size without Bs is shown in orange on a separate axis. We estimated the 1C size of the B to be ~570Mbp.

3.2 The general composition of rye B-located high-copy repeats

To characterize the high copy DNA composition of the rye B chromosome, we first identified repeats (i.e. TE-related sequences) using BLAST of individual sequence reads against the TREP repeat database (wheat.pw.usda.gov/ITMI/Repeats/), combined with *de novo* repeat identification using graph-based clustering of sequence reads (Novak et al. 2010). The sequence data were obtained from previously 454-sequenced sorted Bs of rye (Martis *et al.*, 2012). As a reference, we used the sequence information from all A chromosomes (also purified by flow sorting) and genomic sequence information derived from plants both with (+B) and without (0B) B chromosomes. The whole genome extractions allowed to account for differences in sequence coverage (table 2).

Table 2: Sequence coverage of the rye B (560 Mbp) and the rye A genome (7950 Mbp) after 454-sequencing.

collection	reads	Total read size (bp)	Size (Mbp)	coverage
sorted Bs	1.518.884	521.476.205	560	0.93
sorted As	1.008.975	379.952.169	7950	0.05
gDNA 0B		416,597,333		0.05
gDNA 4B		663,341,461		1.14

Similarity-based clustering of sequence reads revealed that about 90% of the rye genome is composed of repetitive DNA, and 70% of the genome is represented by less than 60 different repeat families. The most abundant repetitive mobile elements on the B can be subdivided in 49% Ty3/Gypsy-, 9% Ty1/copia- and 4% DNA transposons-type repeats (table 3). The overall repetitive content appears similar in As and Bs (72% and 74% respectively), but the make-up of this repetitive fraction differs. Compared to As the B seems to have accumulated more *Copia* retrotransposons and less *Gypsy* retrotransposons. Notable is also the high amount of satellite repeats on the B (6%) compared to the As (1%).

Additionally we observed a large proportion of organellar DNA in the sorted B chromosome data.

Table 3: analysis of proportions of different repeat classes in the rye genome.

(1) Chromo-CR is a lineage of centromeric chromoviruses (including Bilby). It belongs to the Chromovir group but it is listed separately to show its amplification on Bs.

(2) A-type satellites are those mostly located on A chromosomes

(3) B-type satellites are mostly B-specific

proportion [%] in:	**sorted As**	**sorted Bs**	**genomic -Bs**	**genomic +Bs**
Retroelements	**62.47**	**57.85**	**64.12**	**65.01**
Ty3/Gypsy	***55.90***	***48.53***	***49.84***	***50.31***
Athila	*34.36*	*22.16*	*18.72*	*17.06*
Chromovir	*10.66*	*12.57*	*17.09*	*17.95*
Chromo-CR(1)	*5.56*	*9.30*	*3.49*	*4.27*
Tat/Ogre	*4.01*	*3.73*	*9.26*	*9.85*
unclassif.	*1.31*	*0.78*	*1.29*	*1.17*
Ty1/Copia	***6.46***	***9.07***	***14.15***	***14.54***
Angela	*2.62*	*3.07*	*9.18*	*9.12*
Maximus	*3.53*	*2.79*	*4.52*	*4.41*
Bianka	*0.05*	*2.44*	*0.06*	*0.46*
TAR	*0.12*	*0.36*	*0.26*	*0.39*
Tork	*0.08*	*0.33*	*0.04*	*0.08*
unclassif.	*0.06*	*0.08*	*0.08*	*0.09*
TRIM	***0.06***	***0.10***	***0.07***	***0.08***
non-LTR	***0.06***	***0.16***	***0.07***	***0.08***
DNA transp.	**4.74**	**3.76**	**4.20**	**4.02**
satellites	**0.94**	**5.56**	**2.77**	**2.98**
A-type(2)	*0.77*	*0.20*	*2.66*	*1.54*
B-type(3)	*0.17*	*5.36*	*0.11*	*1.43*
unclassif.	**3.31**	**3.29**	**2.27**	**2.28**
TOTAL in 98 CL:	71.78	74.47	76.71	76.88

3.3 The major part of the high-copy fraction of the rye genome is composed of mobile elements

3.3.1 Most mobile elements are similarly shared between As and Bs

The most abundant mobile element identified in rye is the class I retrotransposon *Angela* which is a member of the *Copia* superfamily (Ragupathy & Cloutier, 2008). Class I retrotransposons propagate via copy and paste mechanisms, frequently with an RNA intermediate product, and are often rapidly silenced (reviewed in Deragon *et al.*, 2008; Messing & Bennetzen, 2008). To determine whether the chromosomal distribution of *Angela* differs between A and B chromosomes of rye FISH was performed. The LTR (long terminal repeats) region of *Angela* used as FISH probe showed an even distribution of signals over all chromosome arms on both As and Bs but not in the pericentromere or subtelomeres (Fig. 5A). Another *Copia*-like retroelement, *Maximus*, similar to *Tork* showed the same chromosomal distribution pattern (Fig. 5B).

Fatima is a *Gypsy* retrotransposon of ~ 9kbp length. It contains relatively short LTRs and one ORF (Wicker *et al.*, 2001). Also this element is equally distributed over all chromosome arms with exception of the subtelomeres and centromeres. There was no differential hybridisation pattern noticeable between As and Bs (Fig. 5C).

WHAM (wheat abundant mobile DNA) is a *Gypsy* retrotransposon. It contains LTRs and was first isolated from wheat (SanMiguel *et al.*, 2002). This element shows a distribution pattern similar to that of *Fatima*, another *Gypsy* retrotransposon (Fig. 5D).

CACTA elements are members of a very abundant superfamily in Triticeae and are thought to contribute to the increase in genome size in this genus. The name stems from a terminal conserved CACTA motif. Three CACTA elements have been investigated (*Caspar*, *Clifford*, *Jorge*). CACTA elements are class II DNA

transposons. Class II type transposons propagate via a cut and paste mechanism and an increase in numbers happens mostly via gene conversion. These elements contain two open reading frames, a transposase and a protein of unknown function as well as terminal inverted repeats (TIRs) which are variable between the families. The whole superfamily is very variable. They often contain inserts of satellite DNA or tandem repeats, such as the Afa component which is frequent in the *Caspar* subfamily (Wicker *et al.*, 2003). The *Caspar* family is the most abundant CACTA element. It is large in size (10 to 13 kbp) and has been found in *Hordeum*, *Triticum*, *Secale* and *Aegilops* (Wicker *et al.*, 2003). The *Caspar* element is dispersed over the chromosome arms but shows an accumulation in the subtelomeric regions in the distal third of chromosome arms which matches the distribution described (Zhang *et al.*, 2004) for wheat. *Clifford* is a *Caspar*-like element. In wheat it has been proposed that *Clifford* propagated before the class I elements *Angela*, *Sabrina* and *Fatima*, which are the most abundant LTR retrotransposons (Charles *et al.*, 2008). *Jorge* is another CACTA element, also often found to contain deletions (Wicker *et al.*, 2003).

All three elements show similar distributions in rye (Fig. 5E-G). The distribution pattern observed matches the one previously reported in wheat (Sergeeva *et al.*, 2010). In Bs these elements are only found accumulated in the long arm, but not in the short arm. The distinct pattern of subtelomeric accumulation is also present in the B.

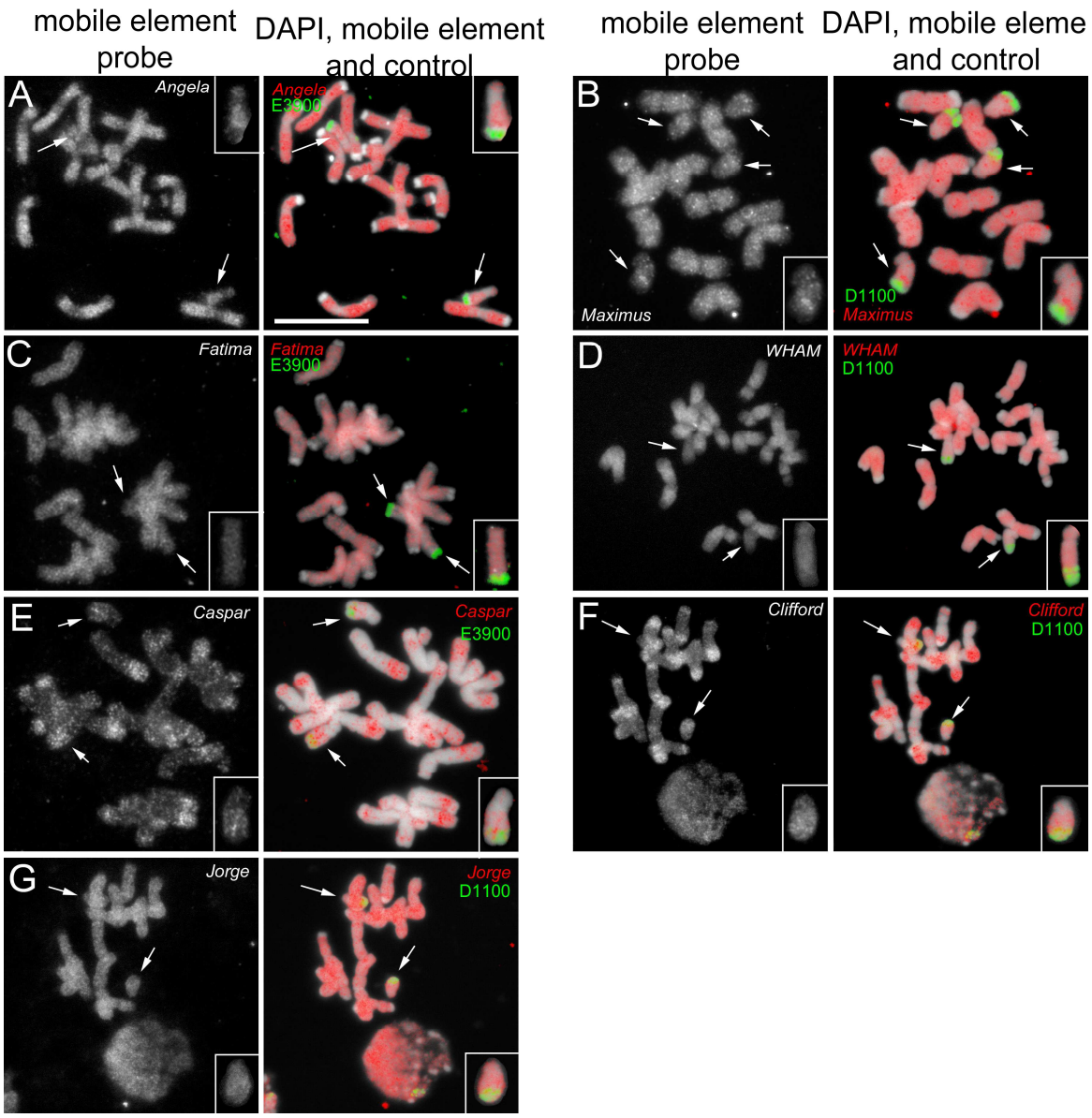


Fig. 5: FISH analysis of several mobile elements (class I and class II) that are similarly distributed on mitotic As and Bs of rye.

Mobile element probes are shown in red and the B identifying control (D1100 and E3900) are shown in green. Bar in (A) represents 10 µm. B chromosomes are indicated with arrows. A) *Copia* retrotransposon *Angela,* B) *Copia* retrotransposon *Maximus,* C) *Gypsy* retrotransposons *Fatima,* and D) *WHAM* E) DNA transposons of the CACTA family *Caspar*, F) *Clifford* and G) *Jorge*.

3.3.2 Three mobile elements were found accumulated on the B

Revolver has been published as a a transposon-like gene of ~3 kbp very common to rye with partial homology to retroelements (Tomita *et al.*, 2008; Tomita, 2010). Our sequencing data confirmed the retrotransposon nature. The hybridization pattern is also similar to that of other retroelements (such as *Sabrina*) and not to that of DNA transposons (e.g. *Caspar*) (Fig. 6D). Also the evolution of this element is noticeable. It is well conserved among *Triticeae*. While it is highly amplified in rye and can be found in medium numbers in the ancestors of wheat, this element was less frequent in extant hexaploid wheat possibly due to allopolyploidisation (Tomita, 2010). Another interesting feature is that it has not been subjected to silencing by epigenetic means or RNAi (Tomita *et al.*, 2008), and is highly transcribed in rye (Tomita *et al.*, 2008; Carchilan *et al.*, 2009). Aside from *S. cereale*, *Revolver* cDNAs have been cloned from *S. sylvestre*, *Dasypyrum villosum*, *Triticum monococcum* and *Aegilops tausschii*. No transcripts have been found in wheat (Tomita *et al.*, 2008; Tomita, 2010).

As previously described (Carchilan *et al.*, 2009), *Revolver* element is densely dispersed across all rye chromosome arms except at subtelomeres. On Bs, the element appears to be even stronger enriched than in the rest of the genome (Fig. 6D).

A B chromosome-specific amplification was found for the TAR-like (*Triticum aestivum* retrotransposon) (Matsuoka & Tsunewaki, 1997)) repeat Sc36c82 (Fig. 6B). TAR belongs to the *Copia* superfamily, contains LTRs and noncoding DNA and is related to Angela and Bianca (Wicker & Keller, 2007). The long arm of the B displayed an intense uniform labeling, while the As revealed disperse signals of less intensity. This element is transcriptionally active in roots and anthers and thus possibly still transposing (Fig. 9). The element is also present in *S. strictum* and *S. sylvestre* displaying a disperse distribution typical for mobile elements (Fig. 10B).

Also amplified on the B is the *Gypsy/chromovirus* retrotransposon represented by Sc11c759 (Fig. 6C). It shows two stronger bands particularly in the euchromatic part of the B chromosome, while exhibiting a characteristic dispersed labeling on the A chromosomes.

3.3.3 One retrotransposon type is depleted from Bs

Sabrina is one of the most abundant gypsy-type retroelements in *Triticeae* which was first described in barley (Shirasu *et al.*, 2000). This LTR element of approximately 1.6 kbp is part of the *Athila* superfamily (Ragupathy & Cloutier, 2008). It is often found to be disrupted and only few copies maintained a functional transposase. The element is supposedly ancient but became recently inactive (Shirasu *et al.*, 2000).

In rye *Sabrina* is dispersed evenly across the chromosome arms and less frequent in the pericentromeres and subtelomeres (Fig. 6A). Strikingly, this ancient retroelement, highly present on the As, is far less abundant on the Bs.

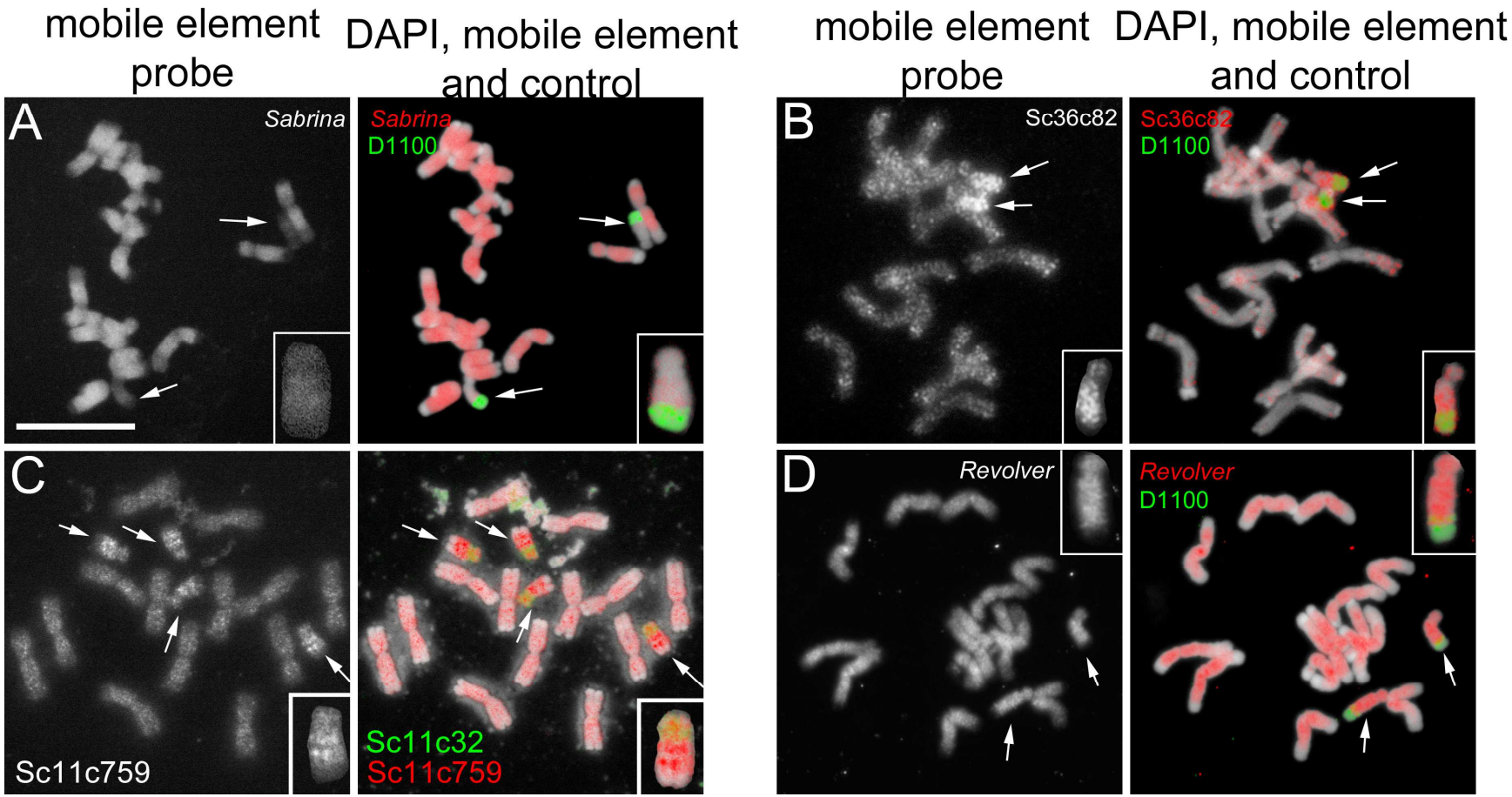


Fig 6: FISH analysis of several mobile elements that are either amplified on or depleted from Bs of rye compared to As. Mobile element probes are shown in red and the B identifying control (D1100 and Sc11c32) are shown in green. A) LTR-retrotransposon *Sabrina* is depleted from Bs while B) *Copia* predicted sequence Sc36c82 is accumulated. C) The LTR part of a retrotransposon Sc11c759 is accumulated on the B. D) Rye specific element *Revolver* is enriched on the Bs. Bar in (A) represents 10 µm. B chromosomes are indicated by arrows.

3.4 Identification of rye B-located satellite repeats

We identified 10 novel satellite repeats via similarity based *in silico* clustering (Novak *et al.*, 2010). After clustering, the relationship of reads inside the clusters can be visualized in a graph. In the graph shown in Fig. 7 representations of three clusters are given. In this graph, every dot represents one read. Laid together, they form a line along the consensus sequence. The closer a read is to that central consensus, the more similar it is. Areas of high variability can be seen as clouds emanating from the consensus line (e.g. the LTR region (blue) in Fig. 7A). A closed circle in the graph usually indicates a repetitive element that frequently borders to another copy of the same element (e.g. a tandem repeat). Parallel pathways (e.g. in Fig 7C, arrowheads) show deletions and insertions present together with the original element.

Next, primers were designed for these sequences (table 1) and the prediction of B-specificity was tested via FISH of the PCR products. In addition, the potential transcription activity of the satellite repeats was evaluated. To determine whether these repeats are specific for *S. secale* only, additonal members of the genus *Secale* were tested.

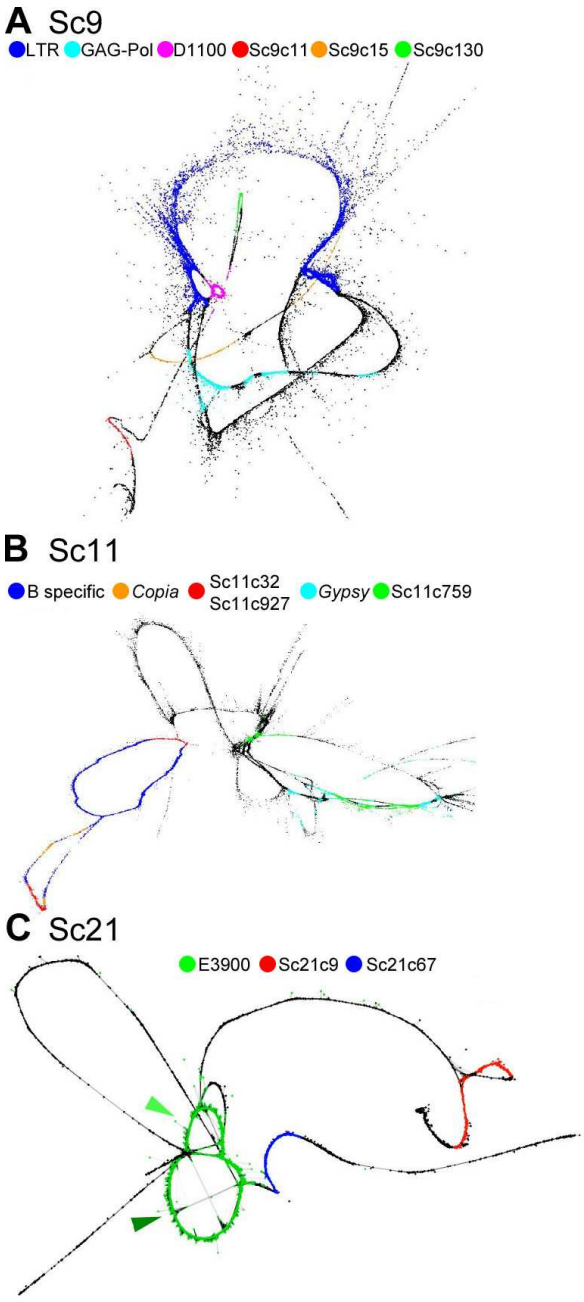


Fig 7: Visualization of clusters after similarity based clustering of 454 reads.

Each dot represents one read. Reads corresponding to identified sequences are colored as indicated. A) Visualization of cluster Sc9 including D1100, B) Sc11 including pericentromeric sequences C) and cluster Sc21, including the sequence E3900 (green). Arrowheads indicate the two versions of E3900, full and deleted. The deleted version runs around the large ring (dark green arrowhead), while the full version also includes the smaller ring on top (light green arrowhead).

3.4.1 The B nondisjunction area is characterized by locally amplified satellites

In addition, to the previously described satellite repeats E3900 and D1100 (Sandery *et al.*, 1990; Blunden *et al.*, 1993) there were several B-enriched repeats identified which generated discrete FISH signals indicative of satellite sequences. Most of the identified satellites were localized to the nondisjunction control region of the long B arm (Fig. 8 and Fig. 21). For fine mapping of these elements we employed pachytene chromosome preparation, which allow for higher resolution (Fig. 11).

Similar to E3900 and D1100, these sequences were characterized by long repeated units spanning up to several kilobases. The method of similarity-based read clustering, used to identify these repeats, revealed that most of them were parts of clusters with complex structure including LTR-retroelement-like sequences. Visualization of these structures using cluster graph projections (Fig. 7) combined with FISH localization of probes designed from selected regions of the graphs were applied to investigate genome distribution and possible origin of these repeats.

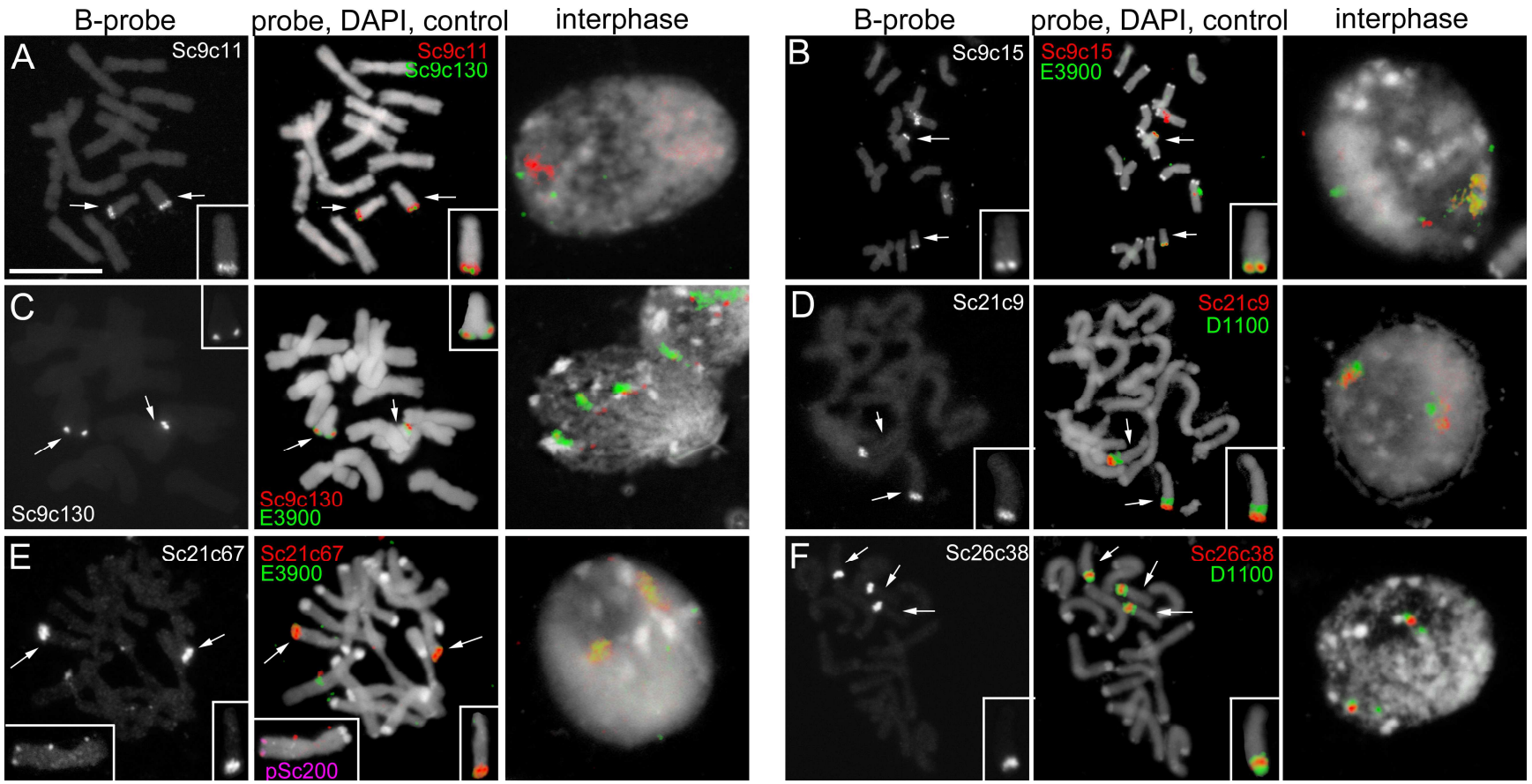


Fig 8: FISH analysis of satellites located in the nondisjunction area of the B chromosome.

New identified satellites are shown in red, B-identifying controls are shown in green. Satellites are shown during metaphase and interphase. Inserts show the B enlarged, B chromosomes in Metaphase are arrowed. A) Sc9c11, B) Sc9c15 and D) Sc21c9 colocalize with E3900.C) Sc9c130 marks the end of the long B arm. E) Sc21c67 colocalizes with E3900 and additionally shows signals on one A chromosome pair, where it colocalizes with pSc200 (magenta). F) Sc26c38 marks a subregion of D1100.

The graph of the cluster Sc9 which includes the satellite D1100 contained regions corresponding to gag-pol and LTR sequences of *Gypsy* element (Fig. 7A). The repeat D1100 forms ring-like structure on the junction of LTR/gag-pol regions (magenta), suggesting it originated by tandem amplification of a part of LTR sequence. There is another circular structure representing putative tandem repeat attached to D1100. The probe for this predicted 521bp-long sequence, Sc9c130 (green), produced FISH signals at a region more terminal than E3900 with a signal so precise that often both chromatids are distinguishable (Fig. 8C). When looking at other species of the genus *Secale*, we found slight amplifications in *Secale strictum* (Fig. 10D). Two additional probes representing linear parts of the graph (Sc9c11 and Sc9c15) mimicked the distribution of E3900 (Fig. 8A&B). When we tested for transcription in leaf, root and anther tissue, we could not show activity of Sc9c15 (Fig. 9). A small region of Sc9c11 (primers F2+R2, table 1) is transcribed in all tested tissues with and without Bs.

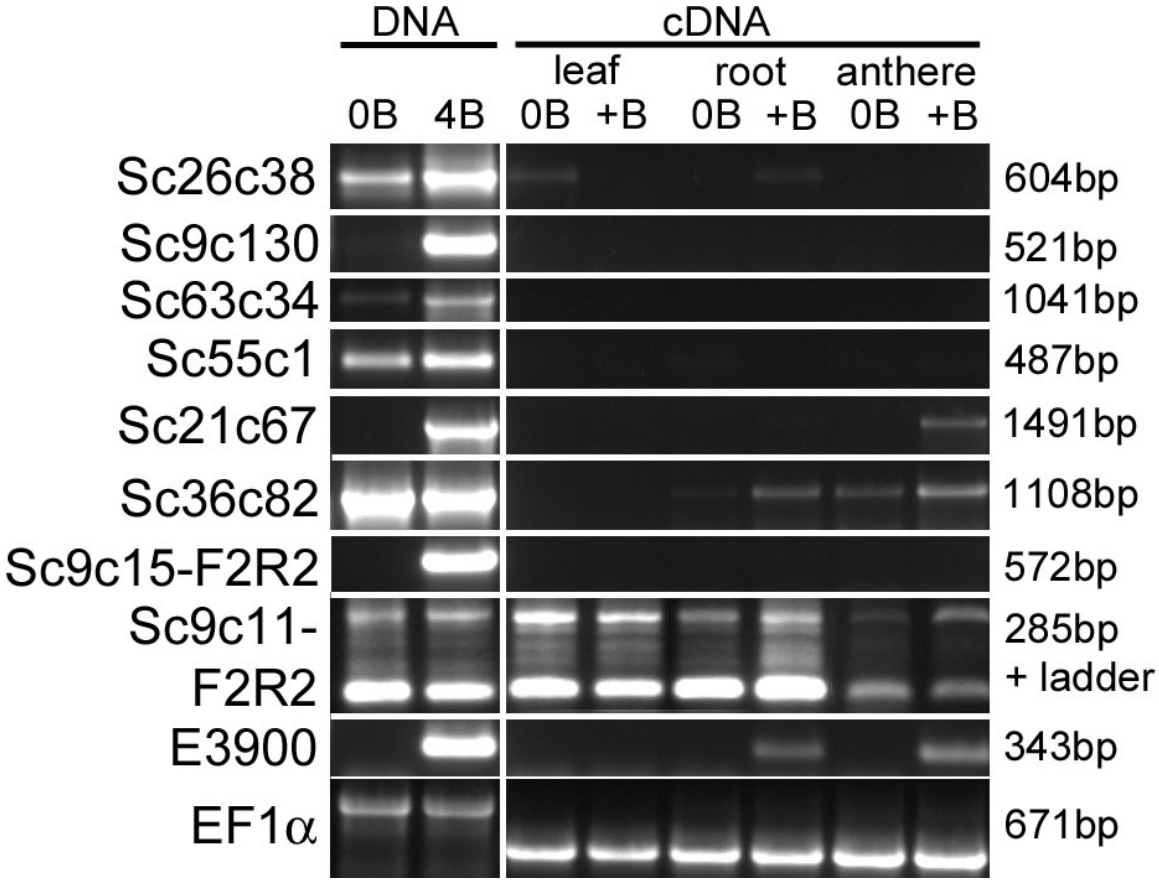


Fig 9: Transcription of B located repeats

PCR was performed on genomic DNA and cDNA (from leaf, root and anther tissue) of rye plants with and without Bs. Length of the achieved fragments are indicated. As a control, the gene EF1α and the known B-specific expressed E3900 were included.

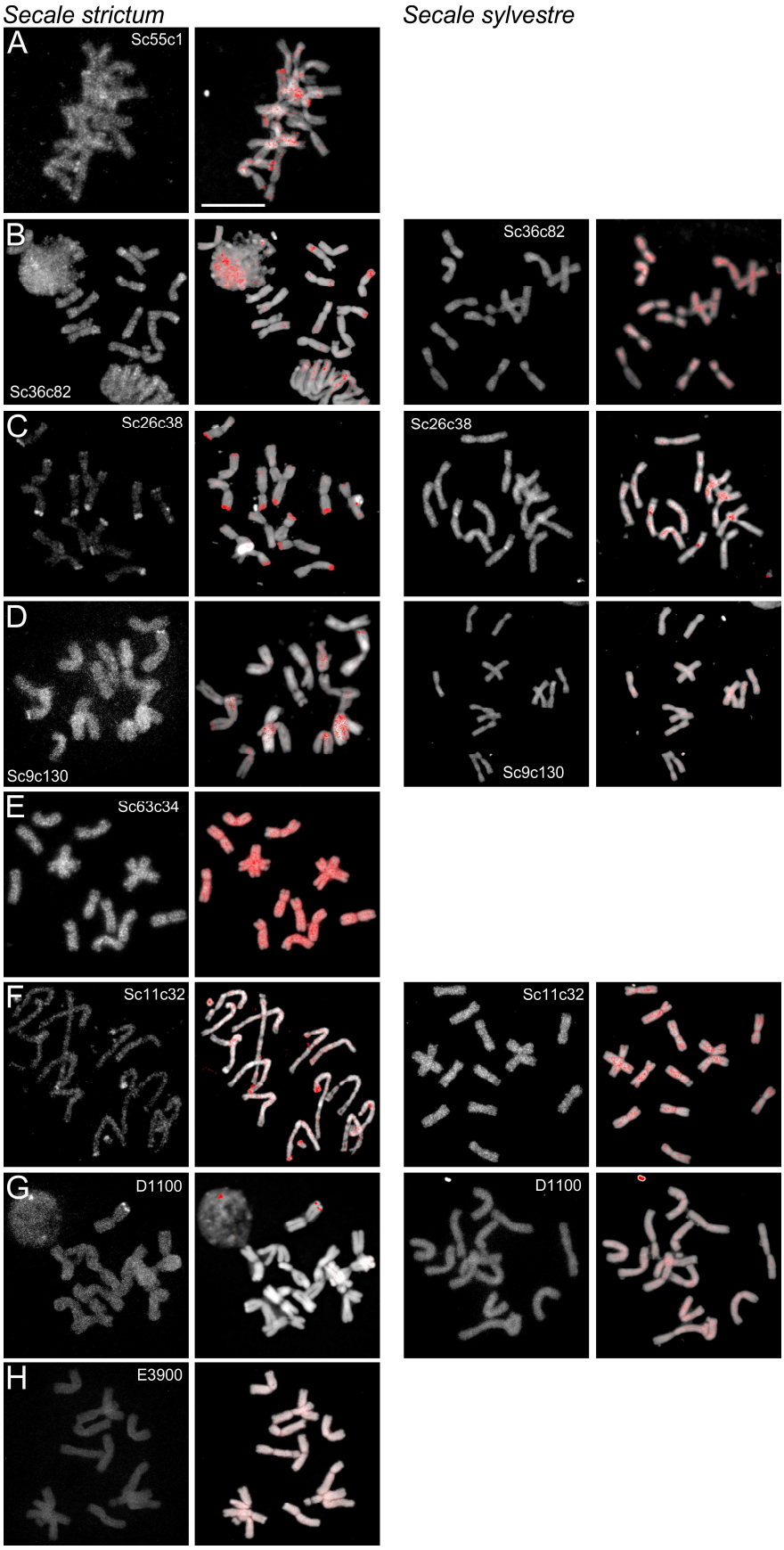

Secale strictum
Secale sylvestre
A
Sc55c1
B
Sc36c82
Sc36c82
C
Sc26c38
Sc26c38
D
Sc9c130
Sc9c130
E
Sc63c34
F
Sc11c32
Sc11c32
G
D1100
D1100
H
E3900

Figure 10: B enriched sequences are present in evolutionarily older *Secale* species. FISH of mitotic *Secale strictum* and *S. sylvestre* cells with rye B-enriched sequences (red). Bar in (A) represents 10 µm.
A) Sc55 shows partial accumulation in *S. strictum*. B) Sc36c82 is present in both species and shows slight accumulation in S. strictum. C) Sc26c38 shows local accumulations at the subtelomeres in S. strictum and close the centromere of one chromosome pair in S. sylvestre. D) Sc9c130 is accumulated on one chromosome pair of *S. strictum* but not in *S. sylvestre*. E) Sc63c34 is dispersed across the genome of *S. strictum*. F) Sc11c32 shows accumulation towards the subtelomeres in *S. strictum* but no visible accumulation in *S. sylvestre*. G) D1100 shows one band on one chromosome of *S. strictum* but not clustered signals on the genome of *S. sylvestre*. H) E3900 shows only background signals on a metaphase of *S. strictum*.

The second previously identified B-specific satellite, E3900, was found to be a part of cluster Sc21 where it formed a ring-like structure (Fig. 7C). Interestingly, this graph also shows the presence of both versions of E3900, the complete sequence initially reported (Blunden *et al.*, 1993), and the truncated version (Pereira *et al.*, 2009) (Fig. 7C arrowheads). The probes Sc21c9 and Sc21c67, derived from adjacent graph regions but sharing no sequence similarity with E3900 were found to co-localize with E3900. Sc21c67 additionally marked a very small signal on one pair of the A chromosomes, possibly showing its origin before amplification on the B. This A signal coincides with that of pSc200, a highly variable marker for subterminal heterochromatin, on this specific chromosome only (Fig. 8E). Interestingly, Sc21c67 exhibits transcription in the anthers of rye plants with Bs only (Fig. 9) and undergoes decondensation during interphase (Fig. 8E). The B-specific transcription makes this sequence an additional candidate for the nondisjunction controlling function.

The known sequence E3900 is interspersed with other repeats, some of which also showed transcriptional activity. The euchromatin histone mark H3K4me3 is notably also enriched in this region (Carchilan *et al.*, 2007). The region is flanked on both sides with heterochromatic sequences (Sc26c38, Sc9c130) which

provide a natural border. D1100, which has also been shown to be transcribed, partially overlaps with this area but also spreads towards the interstitial region. As D1100 has a higher copy number than E3900 but is less transcribed, it is likely that only the D1100 copies colocalising with E3900 are transcribed.

Cluster Sc26c38 is colocalizing with a D1100-postive subregion (Fig. 8F). In condensed metaphase chromosomes this signal appears like one very precise band. However, in the higher resolution of pachytene preparation, the element reveals two distinctly separated bands (Fig. 11C). Unlike the decondensed D1100-positive interphase chromatin, Sc26c38-positive chromatin is highly condensed during interphase (Fig. 8F). Correlating with its heterochromatic nature, no transcription was detectable (Fig. 9). Interestingly, we found Sc26c38 to be also amplified in *S. strictum* at the terminal part of the chromosomes and in *S. sylvestre* near the centromere of one chromosome pair (Fig. 10C). Since we found no other amplification in the closer *S. strictum* varieties, it is likely that this is a secondary enrichment, independent of the amplification on the B.

Thus, the nondisjunction control region of the B is composed mainly of multiple B-specific enriched repeats. None of the repeats are truly B-specific as similar repeats, although of lower copy number, were also A-located within the genus *Secale,* suggesting that a specific amplification of repeats occurred during the evolution of Bs.

The distinct hybridization pattern of the nondisjunction control region shows many locally amplified satellites. This is in strong contrast to the homogenized appearance of the mobile element content. Local amplification of satellite sequences thus seems to be a feature of the entire B rather than just a particular region. Similar amplifications on the As of rye are only seen in the subterminal heterochromatic satellites.

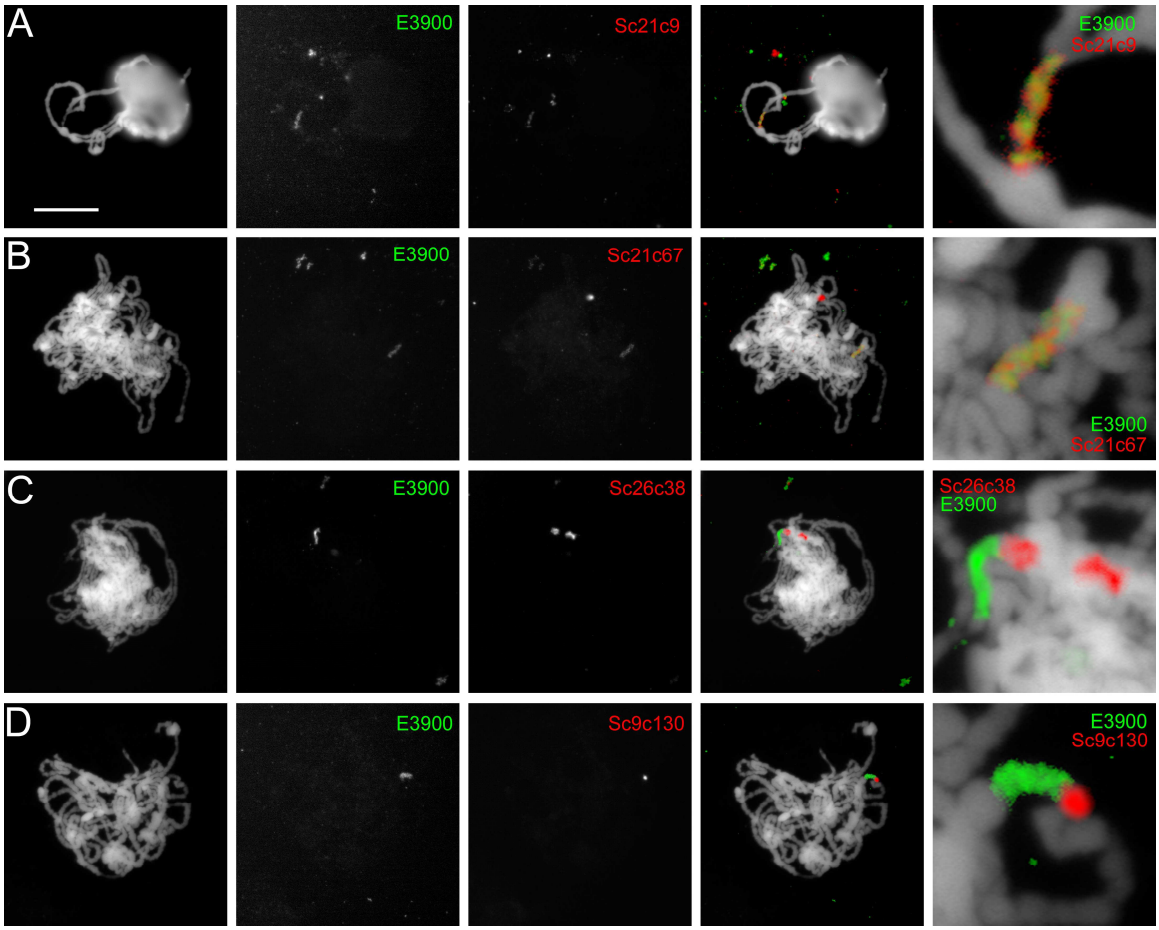


Fig. 11: FISH of some satellites in the nondisjunction region on pachytene spreads. As a control, E3900 is shown in green, satellites are shown in red. The last picture shows the nondisjunction region enlarged for better visibility. A) Sc21c9 completely overlaps with E3900, B) Sc21c67 completely overlaps with E3900, C) Sc26c38 does not colocalize with E3900 and shows two distinct signals, D) Sc9c130 lies distal to E3900.

3.4.2 The B centromere and pericentromere differ from their A counterparts

3.4.2.1 The B pericentromere shows amplification of a B specific sequence

The B pericentromere-located cluster Sc11 showed the most complex structure, including two rings of reads that were almost exclusively derived from B chromosomes attached to the rest of the graph which contained reads from both, As and Bs (Fig. 7B). Reads from different parts of the graph showed similarity to gag-pol protein domains of either *Gypsy* or *Copia* elements (Fig. 7C, green and orange respectively). On the other hand, *Copia* sequences were associated with one of the B-specific rings and were enriched specifically in the B pericentromere.

The probe Sc11c32 is derived from a junction of the second B-specific ring to the rest of the graph, while Sc11c927 corresponds to the middle of one of the B-specific *Copia* rings. No difference in localization was found between the two probes. Both *Copia* sequences showed diffuse signals overlapping the B pericentromeric region with exclusion of the transcriptionally active centromere (Fig. 12A, Fig. 13 A&C (Banaei-Moghaddam *et al.*, 2012)). Additionally, a secondary band in the middle of the long arm is visible (Fig. 14A-C&H). After longer exposure, we found weak labeling also on the subtelomeric parts of two A chromosome pairs (Fig. 12B), possibly reflecting the parental *Copia* elements before they became transferred to the B. This sequence gave only a dispersed signal in both species *S. strictum* and *S. sylvestre* (Fig. 10F).

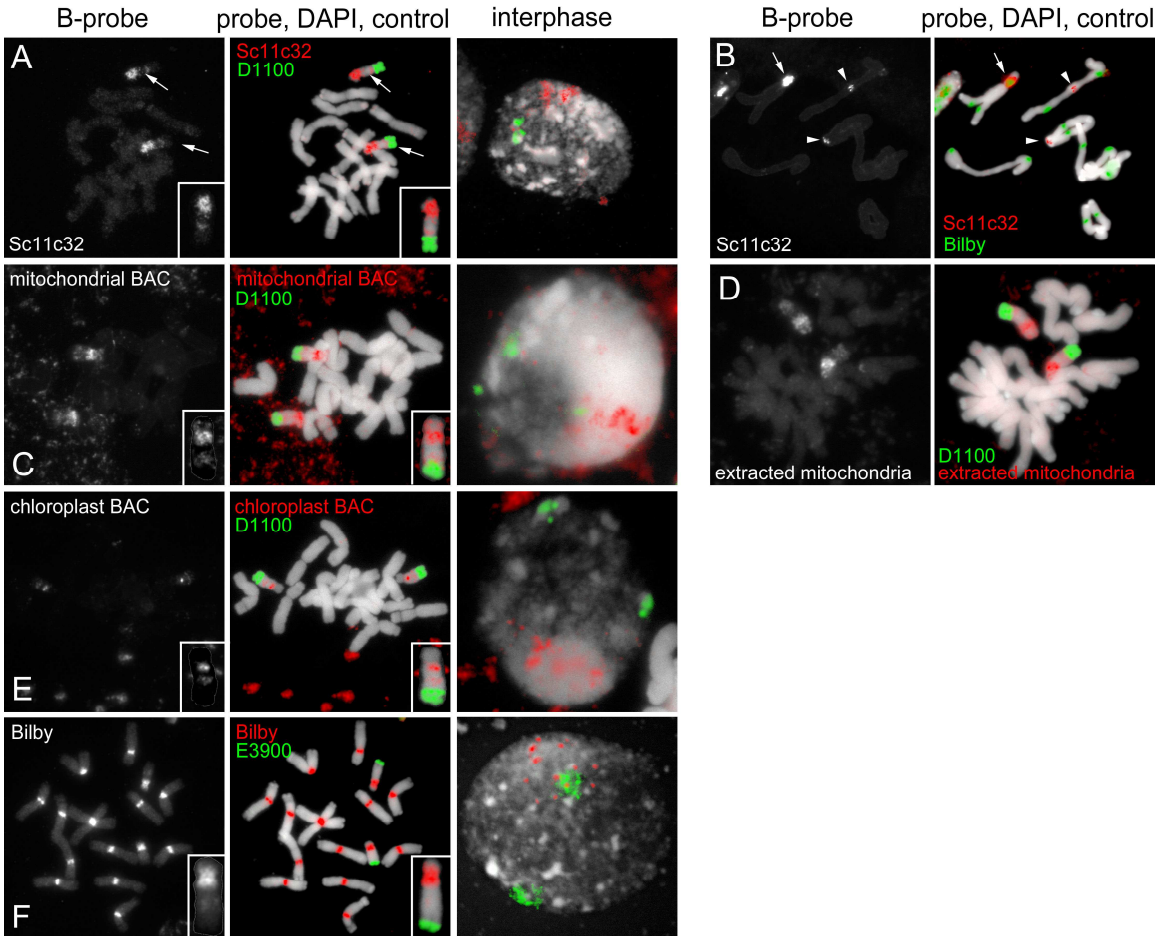


Fig 12: FISH showing B centromeric and pericentromeric signals in metaphase and interphase. Known control sequences are shown in green, probes are shown in red. A) Sc11c32 marks the pericentromeric area of the B as well as a second region proximal the B specific marker D1100. B) Meiotic rye bivalents. Arrowheads indicate A chromosome-located Sc11c32. C) Labeled BAC of mitochondrial DNA from barley marks the B pericentromere. Labeled mitochondria are visible on the slide. D) DNA from extracted mitochondria gives the same labeling as in C E) Labeled BAC with chloroplast sequences from barley gives a signal on the long arm adjacent to the pericentromere. F) The known rye centromeric sequence Bilby shows an extended signal in the B. (Banaei-Moghaddam *et al.*, 2012)

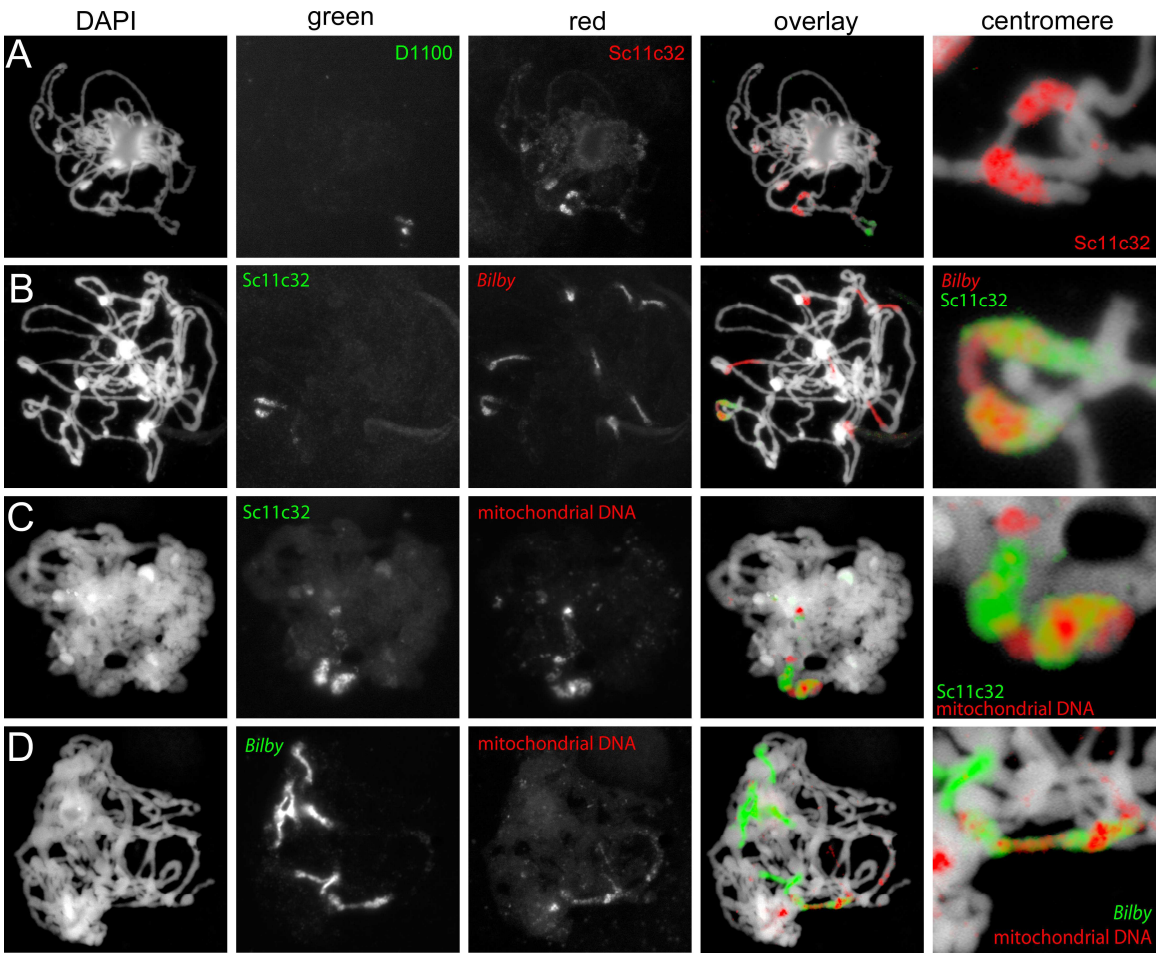


Figure 13: FISH on rye +B pachytene chromosomes with B-centromeric and pericentromeric sequences. The centromere area is enlarged.

A) Sc11c32 (green) is only present in the pericentromere B) The known rye centromeric sequence Bilby (green) is extended into the pericentromere on Bs. And slightly accumulated on the long arm. C) mitochondrial DNA is also present in the centromere and outside the pericentromere on the short arm. D) Mitochondrial DNA (red) is accumulated on the long arm.

3.4.2.2 Organellar DNA is enriched in the B pericentromere

To proof the *in silico* identified B-specific enrichment of organellar genome sequences, FISH was performed with labelled barley BACs encoding over 120 kbp-long regions of the mitochondrial (HVVMRXALLhA0) and chloroplast (ChHB040G01) genome. The high degree of sequence similarity of the organelle genome between related species (Huang *et al.*, 2002) allowed the application of BAC-probes derived from barley.

We found a great accumulation of organellar DNA-specific FISH signals on the rye Bs. Especially mitochondrial DNA probes gave a strong signal with a distinct pattern (Fig. 12C&D, Fig: 13C&D). The mitochondrial signal spans the centromere and pericentromere of the B chromosome. It colocalises with *Bilby* in the centromere but also gives additional bands in the short and long arm pericentromere (Fig. 13C&D). On the A chromosomes we found a band on one chromosome pair, but much more faint than the signal on the B. A secondary band of mitochondrial signal was found on the B chromosome in an interstitial position proximal of the nondisjunction area (Fig. 14E-H).

Additionally, as control, we isolated mitochondria from rye and labeled the DNA for a FISH experiment. Both types of mitochondria-specific probes gave the same FISH pattern (Fig. 12C&D).

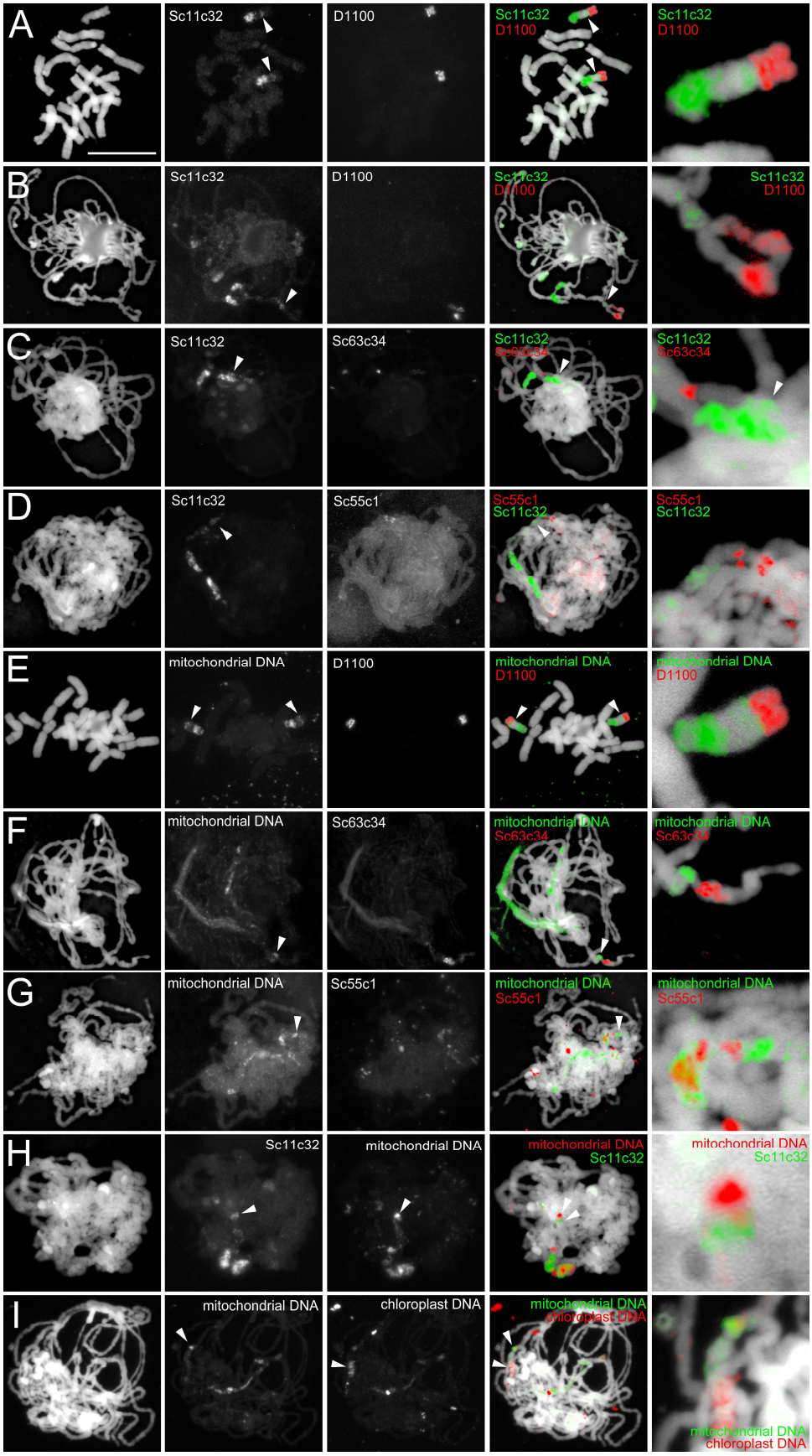

A
Sc11c32
D1100
Sc11c32
D1100
Sc11c32
D1100
B
Sc11c32
D1100
Sc11c32
D1100
Sc11c32
D1100
C
Sc11c32
Sc63c34
Sc11c32
Sc11c32
Sc63c34
D
Sc11c32
Sc55c1
Sc55c1
Sc11c32
Sc55c1
Sc11c32
E
mitochondrial DNA
D1100
mitochondrial DNA
D1100
mitochondrial DNA
D1100
F
mitochondrial DNA
Sc63c34
mitochondrial DNA
Sc63c34
mitochondrial DNA
Sc63c34
G
mitochondrial DNA
Sc55c1
mitochondrial DNA
Sc55c1
mitochondrial DNA
Sc55c1
H
Sc11c32
mitochondrial DNA
mitochondrial DNA
Sc11c32
mitochondrial DNA
Sc11c32
I
mitochondrial DNA
chloroplast DNA
mitochondrial DNA
chloroplast DNA
mitochondrial DNA
chloroplast DNA

Fig 14:: FISH on metaphase and pachytene chromosomes of rye +B for the fine mapping of the secondary interstitial bands of pericentromeric sequences. Important areas are enlarged. Arrowheads indicate the secondary band. Bar in (A) represents 10 µm. A-C) Sc11c32 (green) shows an interstitial signal that lies proximal to Sc63c34 and Sc55c1. D-G) Mitochondrial sequences show a second interstitial signal that lies proximal to Sc63c34 and adjacent to Sc55c1. H) The two second bands of Sc11c32 and mitochondrial DNA are adjacent to each other. I) The chloroplast interstitial signal is disperse and more proximal than the mitochondrial.

Also chloroplast DNA could be detected via FISH, indicating unusually large inserts or multiple small fragments adding up to a larger signal (Fig. 12E). Also chloroplast DNA inserted in the A-genome is present, but is only visible in FISH when longer exposed. This indicates a strong accumulation on the B. One hybridization signal was observed below the extended pericentromere on the long arm of the B (Fig. 12E, Fig15A) adjacent to the *Bilby* signal. A second signal was observed above the nondisjunction control area. This second band is in close proximity to the mitochondrial secondary band but more proximal (Fig. second band). Notably the chloroplast second band coincides with a faint signal of dispersed *Bilby* sequences (Fig. 15A). That means all mapped centromeric and pericentromeric sequences show an additional accumulation on the long arm. This might indicate an ancient inversion with breaks close to the pericentromere and in the interstitial region.

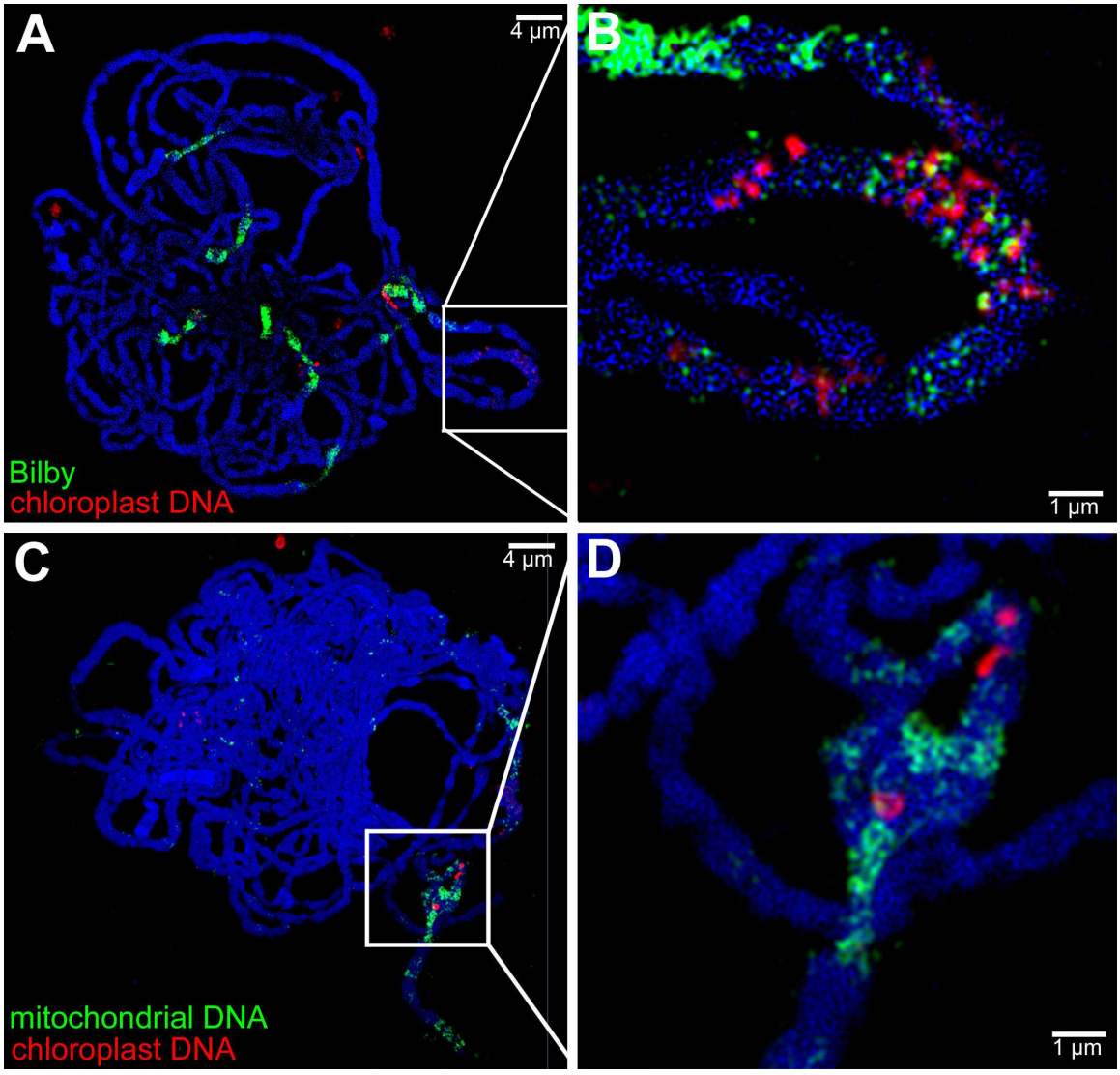


Fig 15: High resolution microscopy detection of organellar DNA on pachytene chromosomes. Important parts are enlarged. A) On the B, the pericentromeric chloroplast signal (red) lies adjacent to the Bilby signal (green). B) The interstitial second chloroplast band colocalizes with weak Bilby signals only visible in high resolution microscopy. Signals have been longer exposed compared to A due to intensity differences between pericentromerical and interstitial signals. C) Mitochondrial DNA (green) and chloroplast DNA (red) are enriched in the B pericentromere. D) the chloroplast signal locates to a region weaker in mitochondrial signals.

In summary, in addition to centromeric sequences shared by both types of chromosomes, the B harbors B-specific pericentromeric sequences. However, the pericentromere of the B appears to be more extended (Banaei-Moghaddam *et al.*, 2012). Because the rye B centromere is characterized by a prolonged sister chromatid cohesion (Schubert *et al.*, 2011) during pollen mitosis, that contributes to nondisjunction, we established a fine map for the B pericentromere (Fig. 16, based on Fig. 12 and Fig. 13).

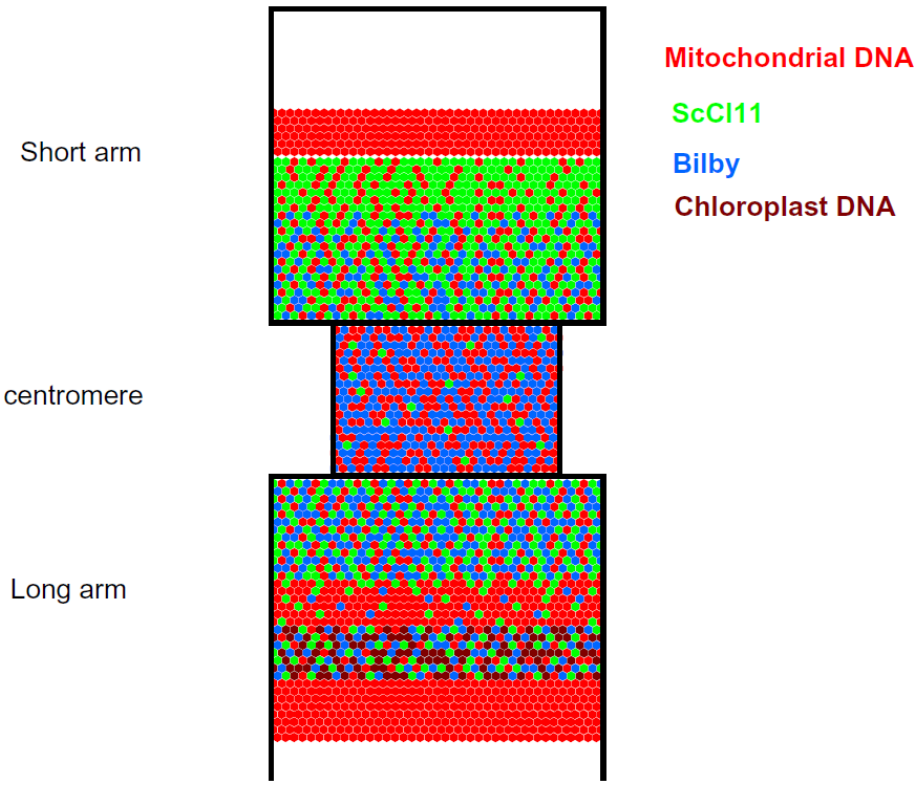


Fig 16: Diagram of the mapping for the rye B (peri)centromere.
Shown are the short arm, the enlarged centromere and the proximal part of the long arm. Present in the B pericentromere are Bilby (blue), mitochondrial DNA (red) and Sc11c32 (green). The chloroplast sequences (brown) are also included.

3.4.3 B-located interstitial satellites

In situ hybridization of Sc55c1 results in two interstitial bands on the long arm of the rye B. Sc55c1 displayed a FISH pattern on the As (Fig. 18A) that is reminiscent of the highly variable subterminal repeats characteristic for rye, such as pSc119.2 or pSc200 (Cuadrado & Jouve, 1994; Kubaláková *et al.*, 2003). These subterminal heterochromatic repeats are usually absent from the rye B (Tsujimoto & Niwa, 1992; Cuadrado & Jouve, 1994), while here we find interstitial B-located Sc55c1-signals, although of weak intensity. The presence of the subterminal signals on the As was variable even within our line. The signal on the B is constant. The distribution of Sc55c1 varies between *Secale cereale* and *S. strictum* where it shows a more dispersed distribution (Fig. 10A).

Southern hybridization of labeled Sc55c1 with partly digested +B genomic DNA revealed a ladder pattern indicative of tandemly organized repeats. (Fig. 17).

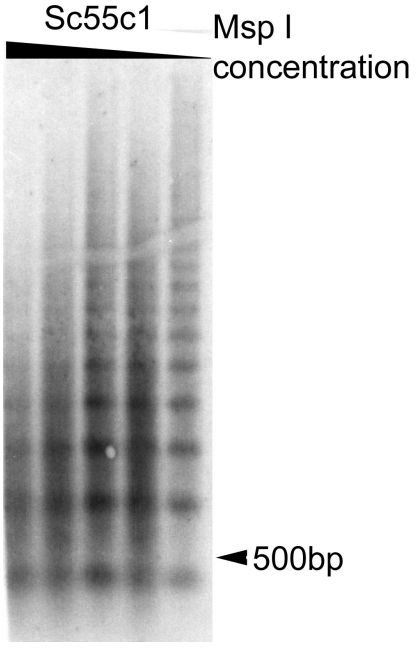


Fig 17: Tandem organization of the Sc55c1 repeat

Genomic rye +B DNA digested with *Msp* I with 5 U; 2.5 U; 1.25 U; 0.6 U; and 0.3U after Southern hybridization with labelled Sc55c1. A tandem structure with the predicted unit size of 487bp is visible.

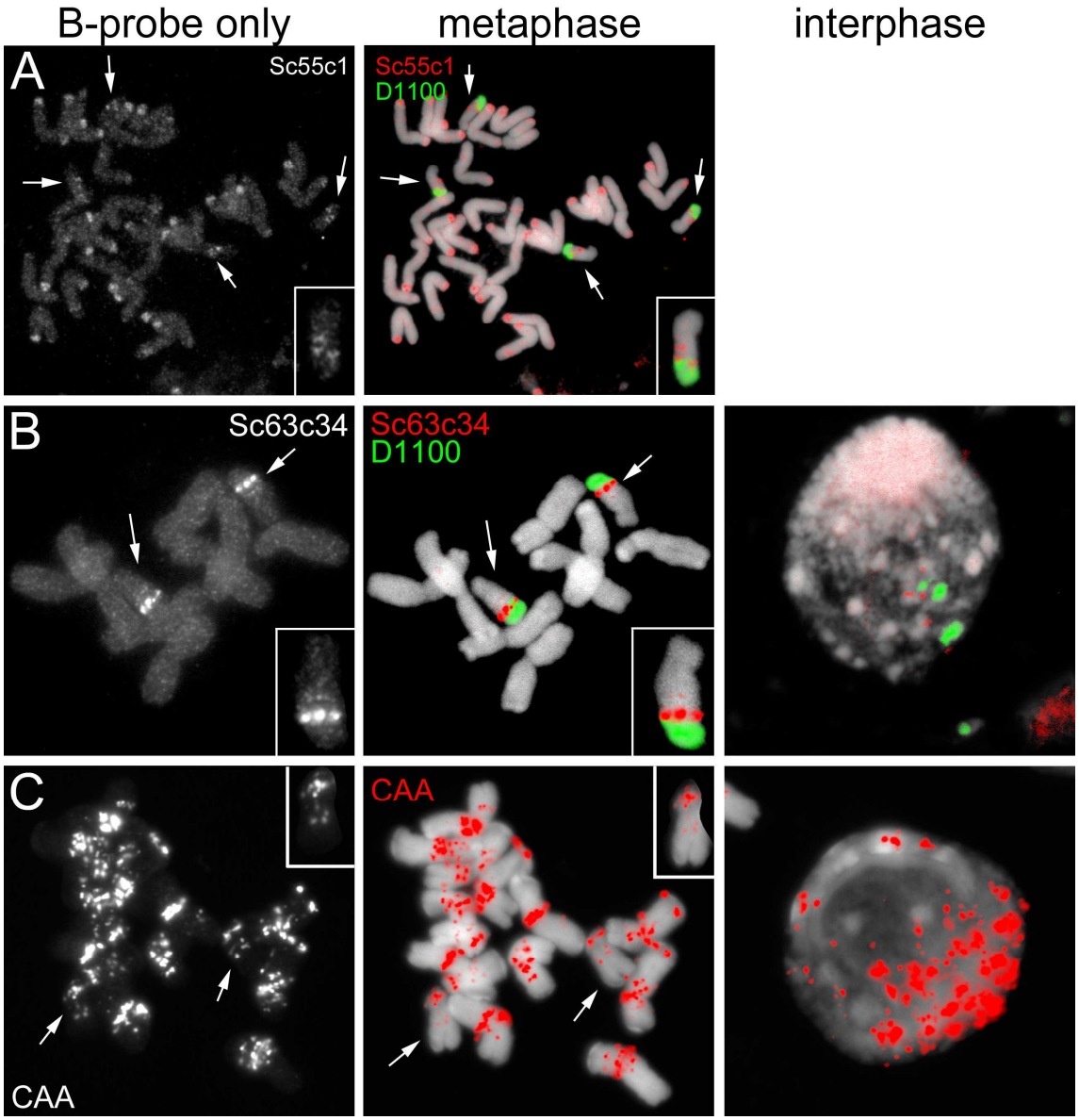


Fig 18: FISH of rye +Bs metaphase and interphase nuclei with interstitial repeats (red) and a B-specific marker (green). Insets show Bs further enlarged. B chromosomes are arrowed. A) Sc55c1 marks 2 interstitial bands on the B and subtelomeric signals on the As. B) SC63c34 shows one interstitial band on Bs. C) The microsatellite CAA has signals around the pericentromere on the As. On the Bs it shows one strong band on the short arm and a weak interstitial band on the long arm.

Sc63c34 shows one band on the long arm of the B chromosome right above the domain marked by D1100 and additionally faint, disperse signals along all chromosomes (Fig. 18B, Fig. 19A). Particularly, it marks the region between the two bands marked by Sc55c1 which is only detectable in the higher resolution provided by pachytene chromosomes (Fig. 19B). This sequence shows a dispersed FISH labeling also in *S. strictum* with no apparent clustering (Fig. 10E). Sc63c34-positive chromatin does not decondense during interphase (Fig. 18B) and does not form transcripts (Fig.9).

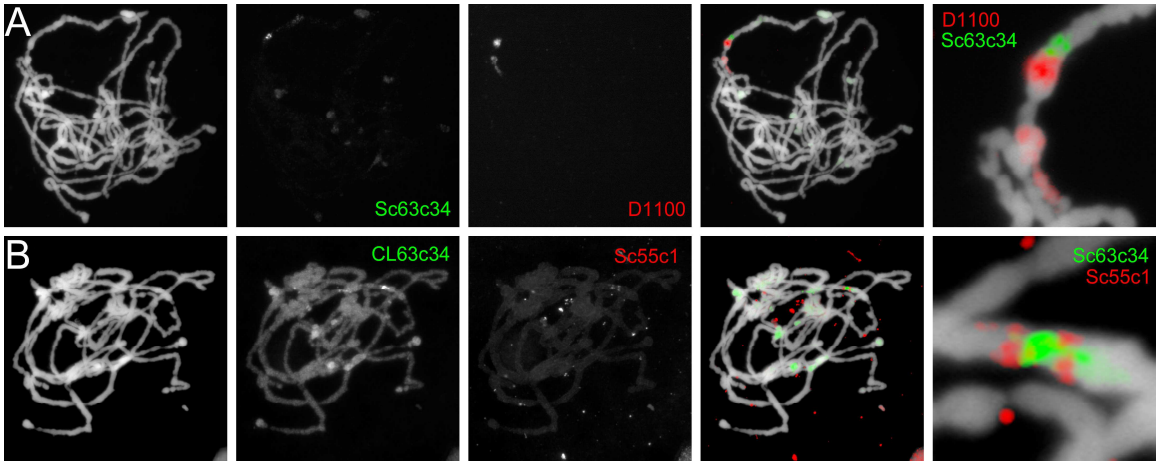


Fig 19 FISH of rye +B pachytene chromosomes with B interstitial sequences. Important parts of the pictures are enlarged. A) Sc63c34 localizes close to D1100 but does not colocalize, B) Sc63c34 is located between the two bands marked by Sc55c1.

3.4.4 Microsatellites: $(CAA)_n$ as marker for the short arm of the B chromosome

As amplification of microsatellites, also called simple sequence repeats (SSR), is documented in rye (Cuadrado *et al.*, 2008), we tested the chromosomal distribution of the following repeats: $(CA)_n$, $(GA)_n$ and $(CAA)_n$. Only $(CAA)_n$ gave distinct signals on the B (Fig. 18 C). Two signals were observed on the B. The smaller band is located interstitially in the euchromatic region of the long arm. The stronger band is notably on the short arm of the B, which makes it the only short arm marker for the rye B. On the A chromosomes the CAA microsatellite is located mainly in the pericentromeres.

3.5 B sequences are accumulated, but neither new nor exclusive

To test for B specificity of the newly identified satellite sequences we performed Southern hybridization on digested genomic DNA of plants with and without Bs. All newly identified B sequences we also found in traces on the A chromosomes. Even the formerly presumed B-specific sequences E3900 and D1100 clearly showed Southern signals in 0B genomic DNA (Fig. 20). For Sc11c32 we could show the A-located related sequences by FISH on meiotic chromosomes (Fig. 12B). However in most cases the A-located sequences showed a weaker and different hybridization pattern than the B located sequences, indicating missing or relocated restriction sites. This shows that the B-located satellite sequences have undergone restructuring evolutionary events after their transfer to the B chromosome before their amplification there.

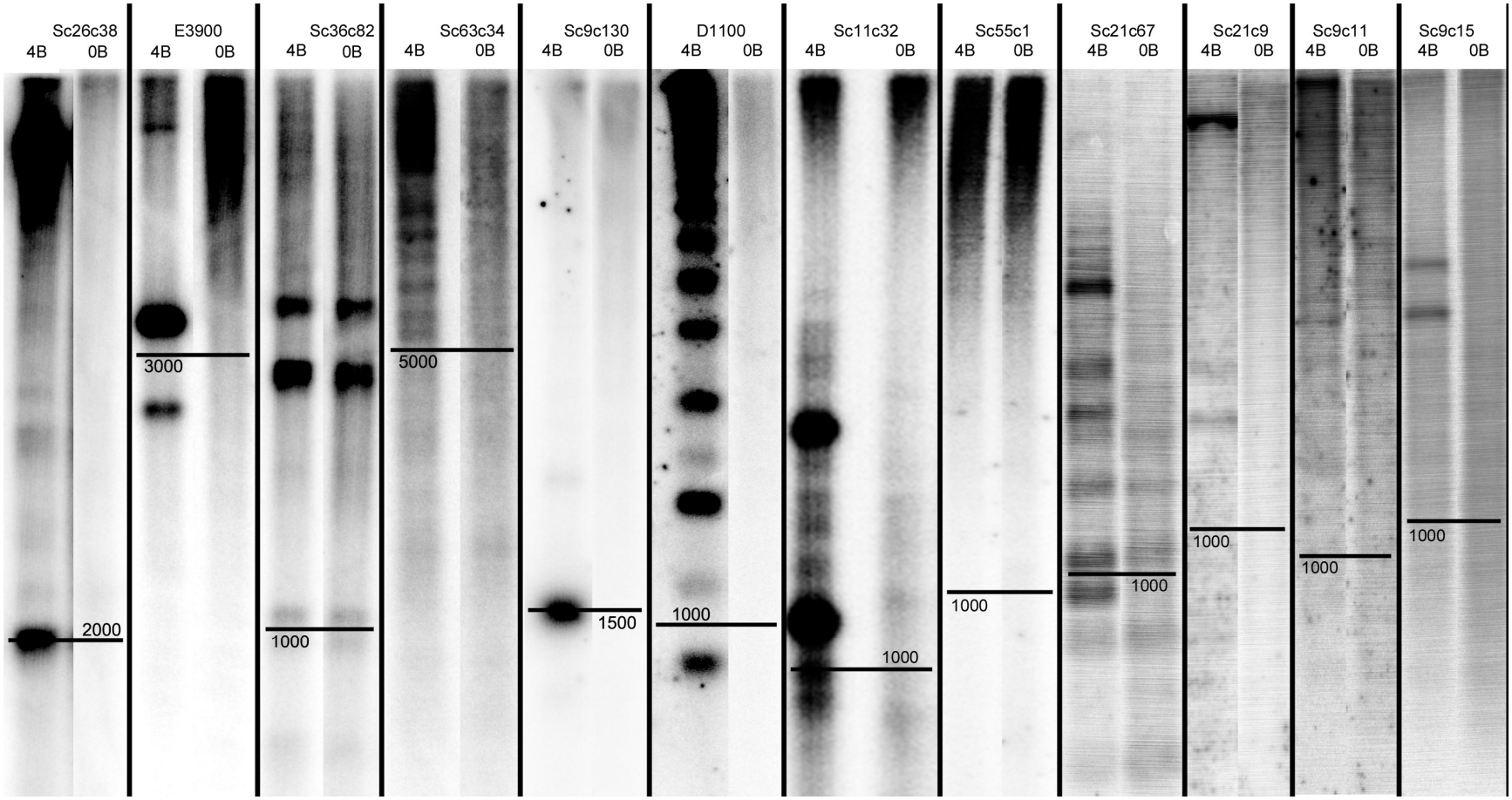


Fig 20: Southern hybridization was performed on digested genomic DNA of plants with and 4 Bs using radioactively labeled probes for the respective sequence. Molecular size is indicated in bp.
Enzymes used were: Sc26c38 *Hind*III; E3900 Eco RI; Sc36c82 *Eco*RI; Sc63c34 *Hind*III; Sc9c130 *Eco*RI; D1100 *Eco*RI; ScCl11 *Nmu*CI; Sc55c1 *Eco*RI; Sc21c67 *Eco*RI and *Hind*III; Sc21c9 *Eco*RI; Sc9c11 *Bgl*II; Sc9c15 *Bsm*FI; mitochondrial *Hpa*II. Marked bands show size in bp

3.6 New markers as a tools for research

3.6.1 The new markers provide a consensus map of the rye B

We identified new B enriched sequences spanning the whole chromosome. We then proceeded to finemap these sequences, producing a consensus map of the rye B (Fig. 21), as a useful tool for further research.

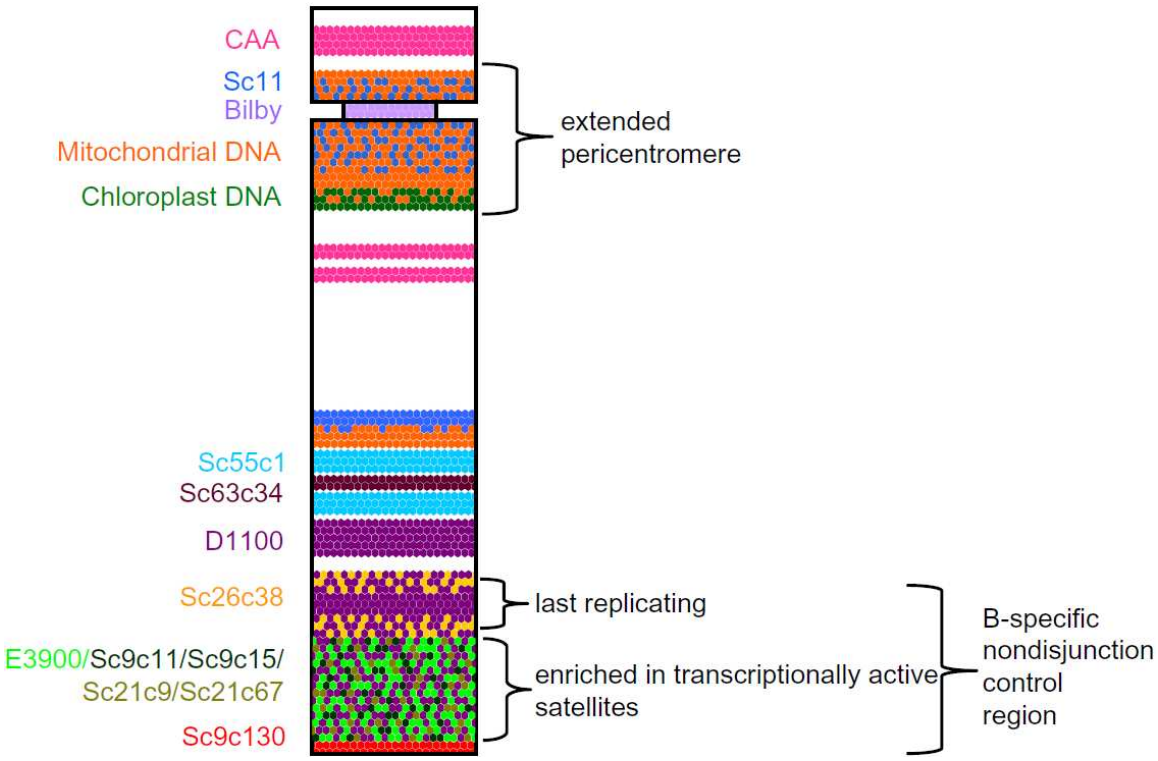


Fig 21: Model for the distribution of B marker sequences over the rye B chromosome. Distribution of 14 different B-enriched sequences on the rye B chromosome. Each sequence is represented by a color as indicated. Pericentromere, late replicating, transcriptionally active, and nondisjunction control regions are indicated. The subterminal region of the long arm encompassing the nondisjunction control region (indicated) is composed of mainly B-specific sequences.

Table 4: New sequences and accession numbers in the Genebank database

Sequence	Genebank number
Sc9c11	KC243218- KC243224
Sc9c15	KC243225- KC243229
Sc9c130	KC243230- KC243235
Sc21c67	KC243236- KC243239
Sc26c38	KC243240- KC243242
Sc36c82	KC243243- KC243245
Sc55c1	KC243246- KC243248
Sc63c34	KC243249- KC243251
Sc11c32	KC243252- KC243254
Sc11c927	KC243255- KC243260

3.6.2 Mapping of B specific area (GISH) with the new marker

To test whether beside the nondisjunction region another domain of the B chromosome is enriched in B-specific sequences, we performed a genomic *in situ* hybridisation (GISH) with labelled 0B DNA and unlabelled 4B DNA for blocking. As previously reported, (Tsujimoto & Niwa, 1992) after GISH the entire B revealed a disperse hybridization signal except at the end of the long arm (Fig. 22A). Hence, the nondisjunction region is uniquely enriched in B-specific sequences. To determine the border between A-shared and the B-specific region a subsequent FISH experiment was performed with the satellite Sc26c38. Figure 22B shows that this repeat marks the border of the B-specific nondisjunction region.

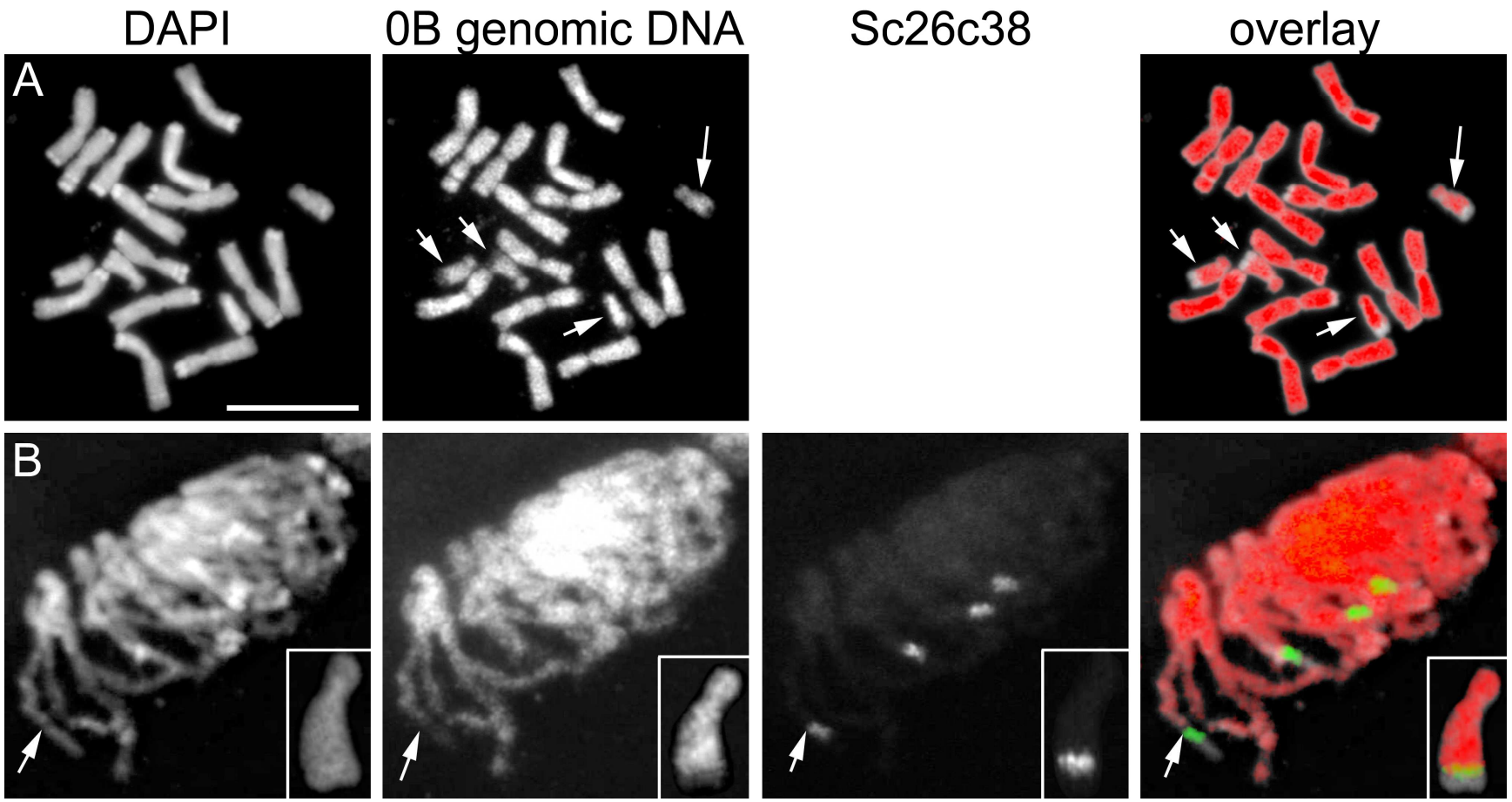


Fig 22: FISH with whole genomic DNA from a 0B plant (red) on metaphase and prophase chromosomes of a +B plant.

A) The B is completely marked by GISH except for the nondisjunction area on the terminal part of the long arm. B) The border between A shared and B exclusive sequences (arrow) is proximal to the band of Sc26c38. Insert shows an enlarged metaphase B chromosome.

3.6.3 One part of the nondisjunction region of the B chromosome replicates last

Next we asked whether the DNA replication behavior of A and B chromosomes differs. A replication gradient along the rye chromosomes was proposed before, and late replication for Bs was predicted (Lima-De-Faria & Jaworska, 1972). Therefore, 5-ethynyl-2'-deoxyuridine (EdU), a nucleoside analog of thymidine, was incorporated into DNA during replication. After detection of EdU-labeled DNA, FISH was conducted with labeled Sc26c38 and Sc9c130. As shown in figure 23, the nondisjunction control region of the B is undergoing replication last. The remaining B region exhibits a similar replication behavior as A chromosomes. Of all sequences tested, Sc26c38 closest colocalised with the signal of very late replicating areas (Fig. 23A). To test whether the replication behavior of the B is the same irrespective of the host genome we employed a wheat-rye B addition line. An overlapping of EdU- and of rye-specific signals was found after analysis of interphase nuclei (Fig. 23F). Thus, the terminal part of the long B-arm replicates last irrespective of the host genome.

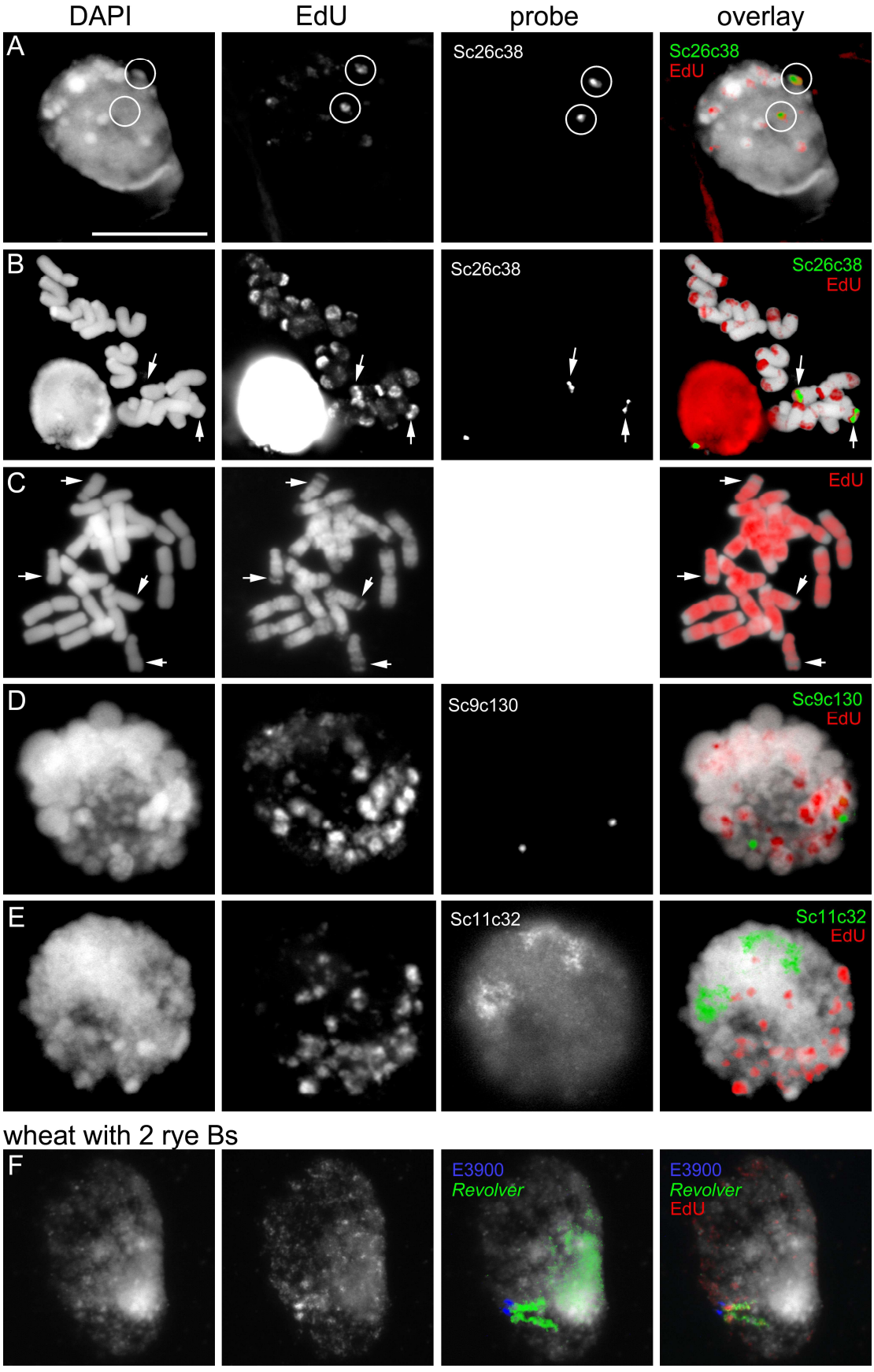

DAPI
EdU
probe
overlay
A
Sc26c38
Sc26c38
EdU
B
Sc26c38
Sc26c38
EdU
C
EdU
D
Sc9c130
Sc9c130
EdU
E
Sc11c32
Sc11c32
EdU
wheat with 2 rye Bs
F
E3900
Revolver
E3900
Revolver
EdU

Fig 23: Replication behaviour of rye A and B chromosomes.

Replicating DNA was labeled with EdU (red) and correlated to several B located probes (green). A) Interphase labeling shows 2 signals of very late replicating DNA (circled). These colocalized completely with Sc26c38. B) EdU labeling of metaphase chromosomes shows that heterochromatin also replicates late on the As. C) Metaphase of EdU labeled early replicating areas. The late replicating areas on the Bs are unlabelled as their replication is delayed. D) The sequence Sc9c130 which is also heterochromatic and more terminal than Sc26c38 does not colocalize with the late replicating areas. E) Sc11c32 replicates earlier. F) The late replication behaviour is also seen on Bs (marked with rye specific Revolver in green and B specific E3900 in blue) which are introgressed into wheat.

3.7 B chromosomes show similar rates of SCEs as A chromosomes

To analyze the sister chromatid exchange (SCE) behaviour of the B in comparison to the As we employed the newer method of 5-Ethynyl-2'-deoxyuridine (EdU) incorporation instead of the well-known 5-Bromdesoxyuridine (BrdU).

The frequency of SCEs can be determined by marking single chromatids with EdU. In this analysis, EdU is provided to the cell during S phase. Due to the semiconservative nature of DNA replication, this results in one strand of the double helix with marked DNA while the old strand is not marked. After another replication cycle without EdU, one chromatid per chromosome appears labelled (Fig 24A). Each shift of colour from one chromatid to the other marks one exchange event (Fig 24C).

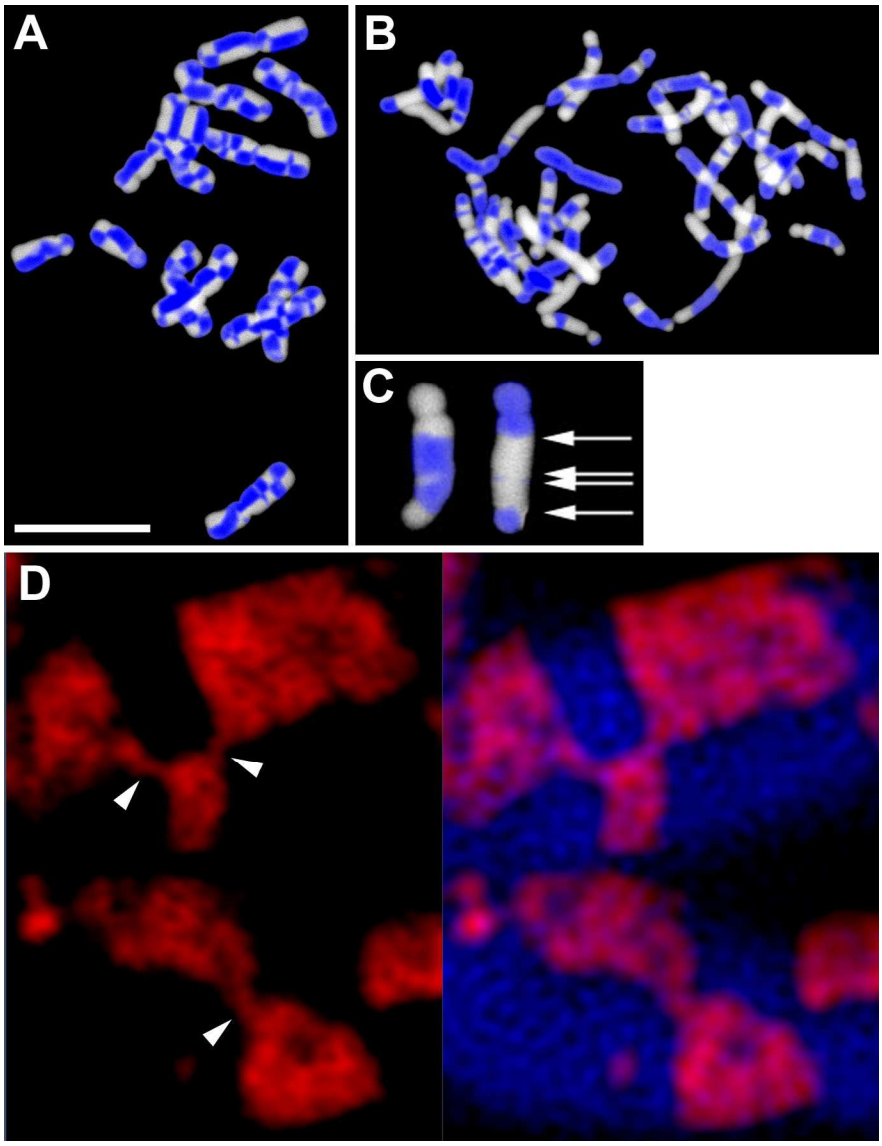


Fig 24: Sister chromatid exchanges (SCEs) of rye A and B chromosomes visualised by EdU labeling (blue) of one chromatid. A) Metaphase spread of chromosomes with one labelled chromatid. SCEs are marked by a switch of the labeling to the other chromatid. B) In a colchicine arrested metaphase the two corresponding chromatids are easily identifiable due to the complementary SCE patterns. C) Two corresponding B chromatids with 4 SCEs (arrows). D) High resolution image by structural illumination microscopy of a detail from a SCE labelled chromosome with DAPI in blue and EdU labelling in red. The breakpoints between chromatids are often connected by labelled fibers (arrowheads). Bar in (A) represents 10 µm.

To compare between the different regions (subtelomere, pericentromere, nondisjunction region and interstitial) of As and Bs the chromosomes were divided into different parts. The A chromosomes were sectioned into seven regions, while the shorter B was subdivided into five regions (Fig. 25).

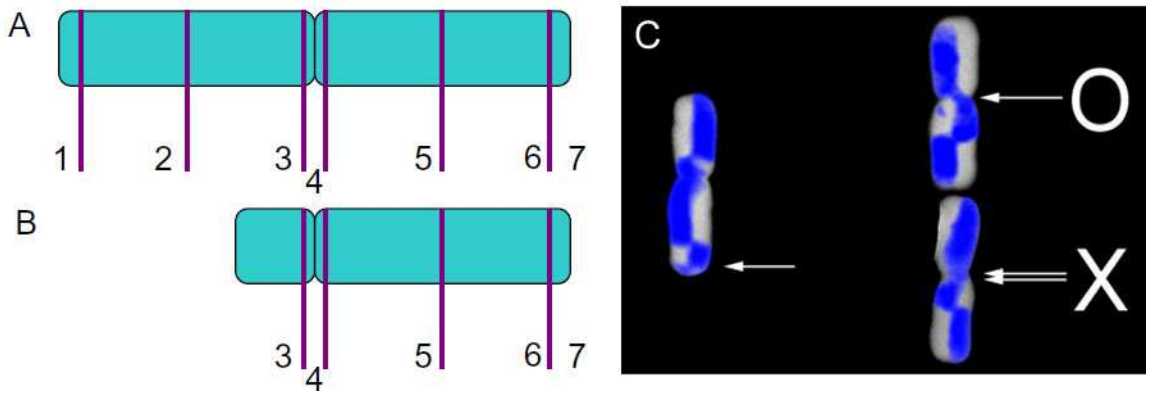


Fig. 25: SCEs were scored according to their place of occurrence. A) A chromosomes were divided into seven regions and B) B chromosomes were divided into five parts. C) Examples for scoring of SCEs. Breakpoint in the heterochromatic end and scoring for the centromere should show the break in the centromeric constriction (O) and not in the region close by (X). In the bottom example the break is also obscured by the three dimensional twisting of the chromatids.

SCE analysis was made based on three different plants. In total 120 B chromosomes and 319 A chromosomes were counted and scored for SCEs. The total number of SCEs per region divided by the total number of chromosomes counted gives the chance of a chromosome to have an SCE in that particular region or the rate of SCEs in this part of the chromosome (Fig.26). The total rate of SCEs for whole chromosomes were 3.4 SCEs per A chromosome and 2.5 SCEs (table 5) per B chromosomes (table 6). The strong difference is mainly due to the shorter length of the B chromosomes.

Table 5: Scoring of SCEs in A chromosomes.

part	total	rate per chromosome
1	74	0.23
2	193	0.61
3	217	0.68
4	15	0.05
5	251	0.79
6	252	0.79
7	96	0.30
Total chromosomes counted	319	
Rate (SCE per chromosome)	3.4	

Table 6: Scoring of SCEs in B chromosomes.

part	total	rate per chromosome
3	31	0.26
4	18	0.15
5	104	0.87
6	89	0.74
7	56	0.47
Total chromosomes counted	120	
Rate (SCE per chromosome)	2.5	

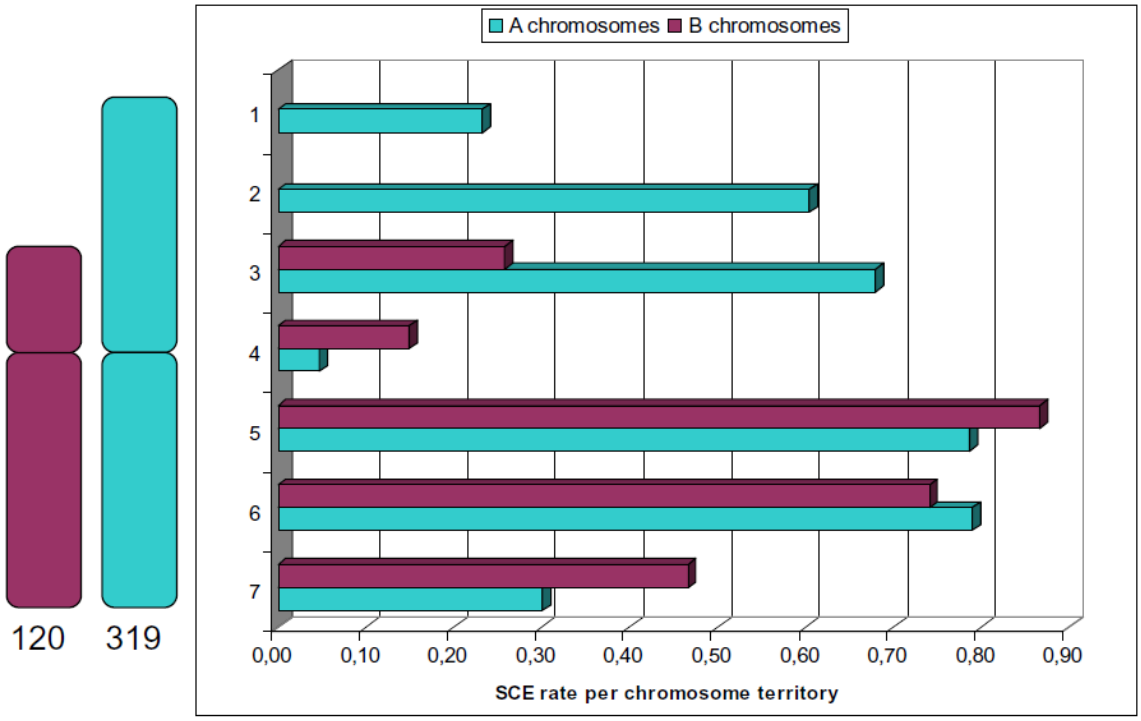


Fig. 26: Rate of SCEs per region in A and B chromosomes.

Rate is the chance of a particular chromosome to have a SCE in the particular region (as shown in figure 27): total number of SCEs in a region / total number of chromosomes counted. Total numbers of SCEs are indicated.

Compared to the previously determined rates of 0.79 SCE per chromosome (Friebe, 1978), the frequency of SCEs determined here is much larger (3.4 SCE per chromosome). Already from the visuals it is clear that every chromosome usually has several SCEs and chromosomes without any breaks are comparatively rare (Fig. 24B includes one chromosome without any visible SCE). Overall, B chromosomes show a similar rate of SCE as A chromosomes.

With the null hypothesis being that there is no connection between A and B SCE rates we employed the F test of the statistical analysis program Strata to determine a critical F value of 0.59 for a confidence interval of 5%. With F = 19 being much larger than the critical value, the null hypothesis is rejected. In our data, the overall SCE rates of A and B chromosomes are thus similar.

4. Discussion

4.1 Mobile elements appear similar in total but differ in constitution between A and B chromosomes

Based on similarity-based clustering of 454-sequence reads we could efficiently decipher the high copy DNA composition of the rye genome (Martis *et al.*, 2012). This method was also used successfully to review other large plant genomes, such as that of *P. sativum* (Macas *et al.*, 2007), Luzula elegans (Heckmann *et al.*, 2013), *Nicotiana tabacum* (Renny-Byfield *et al.*, 2011; Renny-Byfield *et al.*, 2012), *S. latifolia* (Macas *et al.*, 2011), *Rumex acetosa* (Steflova *et al.*, 2013), potato (Torres *et al.*, 2011) and bats (Pagan *et al.*, 2012). Particularly, the method has been employed to show differences between autosomes and sex chromosomes in *R. acetosa* (Steflova *et al.*, 2013).

Most mobile elements tested were similarly distributed along A and B chromosomes. This uniform pattern supports the assumption of an intraspecific origin of the Bs of rye. As the whole genome of rye appears homogenously regarding the pattern of mobile elements, it is no surprise to find that also for the Bs.

The chromosomal distribution pattern observed of CACTA type DNA transposons (*Caspar*, *Clifford*, *Jorge*) matches the one previously reported in wheat (Sergeeva *et al.*, 2010). Importantly, three mobile elements have been identified, that are differently distributed on As and Bs. The *Sabrina* element is an ancient, very abundant retroelement common to all *Triticeae*, which has been inactivated in rye (Shirasu *et al.*, 2000). There is much less of this retrotransposon on Bs than on As. A similar phenomenon has been seen in *Silene latifolia*, where an *Ogre*-like *Gypsy* element is less abundant on the Y chromosome than on other autosomes (Cermak *et al.*, 2008; Filatov *et al.*, 2009; Kejnovsky & Vyskot, 2010). *Revolver* on the other hand is a highly active element that was highly amplified on the B during the speciation of rye (Carchilan & Houben, 2007). The newly

identified element Sc36c82 that was predicted as *Copia* transposon also showed a high accumulation on the B chromosome specifically and was also found to be transcriptionally active.

The strong accumulation of these elements might have its cause in the relaxed selection on the Bs, where the integration of a mobile element does not interrupt genes of critical function As crossing-over has been proposed to contribute to the removal of mobile elements (Charlesworth *et al.*, 1994), the lower crossing-over frequency of the rye B might facilitate uncontrolled accumulation as proposed for several plant Y chromosomes (Charlesworth, 2008). For the B chromosome of rye, it has been shown that although they pair frequently with each other, or themselves when only one B is present, during pachytene (Diez *et al.*, 1993). However, the formation of bivalents is less frequent than in As (Jimenez *et al.*, 2000). For the accumulation of repetitive elements such as *Revolver* and Sc36c82, another, less likely model could be a directed transposition selectively to the B. Targeted transposition of mobile elements has been shown for yeast (Zhu *et al.*, 2003) and *Arabidopsis lyrata* (Tsukahara *et al.*, 2012).

A model that explains both the depletion of *Sabrina* elements as well as the accumulation of *Revolver* and Sc36c82 is possible just with the different transposing activities of both elements (Fig. 27). In the ancient predecessor of *Triticeae*, *Sabrina* was actively transposing and spread evenly over the entire genome. After the early inactivation (Shirasu *et al.*, 2000) before or during speciation of rye, the B chromosome was formed from *Sabrina* containing A chromosomes. The newly evolved *Revolver* then became active and transposed throughout the rye genome. Due to the dispensable nature of the B, the lower selective pressure allowed for stronger accumulation of *Revolver* on the B. Together with accumulation of B-specific sequence, this would dilute the originally abundant *Sabrina* element, which could no longer actively increase in copy number. This model would also fit for other transposing mobile elements such as the *Copia* repeat Sc36c82. Less likely would be a scenario where the

Sabrina element transposes predominantly in plants without Bs and stays inactive in plants with Bs. Such a behaviour has been proposed for the Y chromosome depleted *Ogre*-element in *Silene latifolia* (Kejnovsky & Vyskot, 2010). Another, unlikely possibility to explain the depletion of this element would be the predominant spreading of *Sabrina* via unequal crossing-over, as discussed in Cermak *et al.*, 2008. The reduced crossing-over frequency in the B would then prevent *Sabrina* from spreading on Bs.

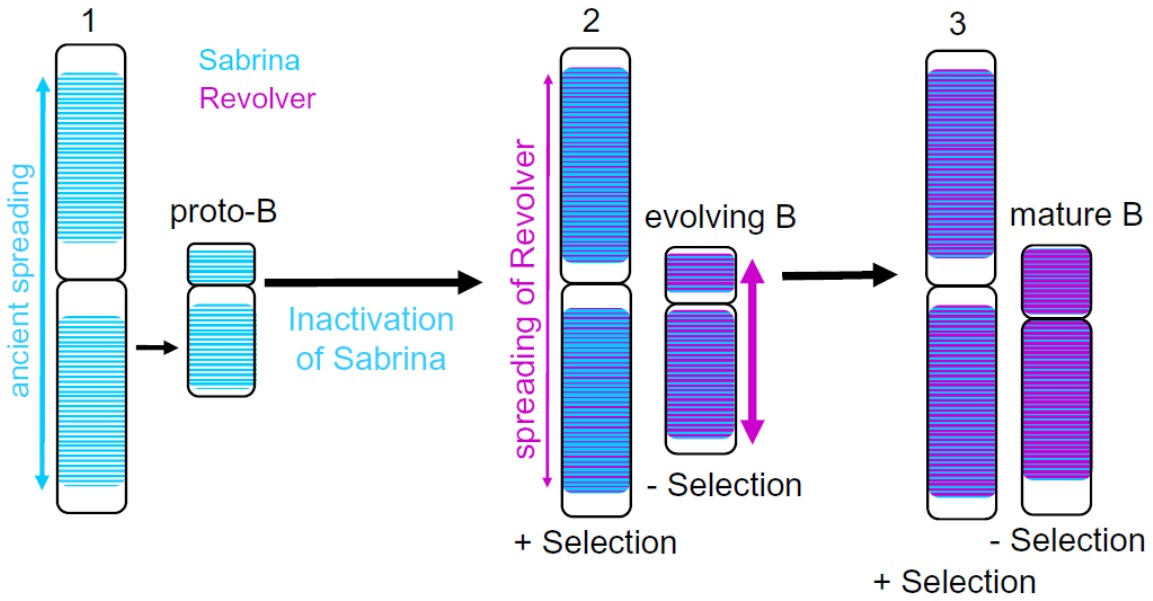


Fig 27: Model for the emergence of the specific mobile element make-up of the rye B, proposing the accumulation of the *Revolver* element with simultaneous depletion of the *Sabrina* element. The ancient *Sabrina* element spread through the early *Secale* species and was present on the B when it was formed. After inactivation, the young rye specific element *Revolver* started spreading throughout the genome. Lower selection pressure allowed for more insertion on the B compared to the A. Combined with a slight growth of the B during the accumulation of B marker sequences *Revolver* accumulates while simultaneously diluting *Sabrina*.

4.2 Organellar insertions hint emphasize the alien nature of the B compared to the rest genome

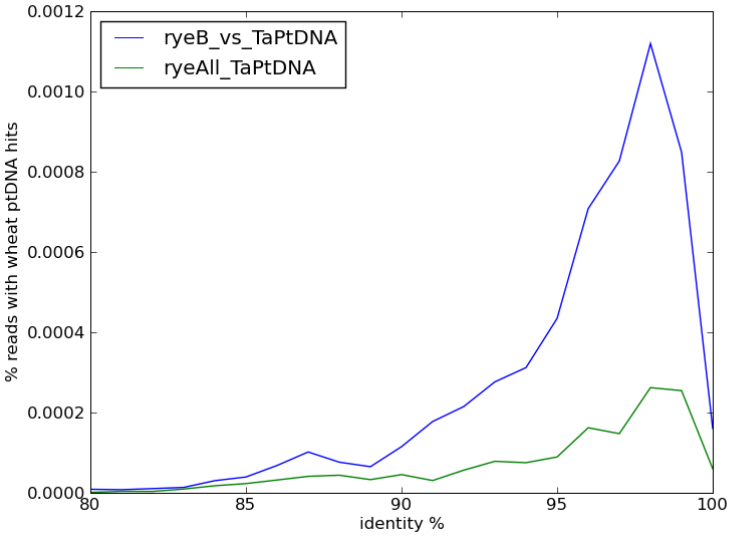
We found large insertions of organellar, specifically DNA on the B (Martis *et al.*, 2012). While plastid DNA is not absent from the As, the B localized sequences are considerably larger. They most likely stem from several independent insertion events.

Transfer of organellar DNA to the nucleus is a well known phenomenon (reviewed in (Timmis *et al.*, 2004; Kleine *et al.*, 2009; Lloyd *et al.*, 2012)) and nuclear insertions of plastid DNA (NUPTs) and mitochondrial DNA (NUMTs) have been shown to be involved in the origin of new functional genes (Lloyd & Timmis, 2011):The presence of organellar DNA in the B is not surprising. NUPTs have also been shown to accumulate on the *S. latifolia* Y chromosome (Kejnovsky *et al.*, 2006). Due to the dispensable nature of the B, the relaxed selective pressure provides a kind of save harbour for extra sequences without immediate benefit. However, the scale of accumulation here, particularly of mitochondrial DNA, is unexpected. It is also noticeable that the strongest accumulation happened in the pericentromere. Pericentric insertions of organellar DNA have also been shown in rice (Matsuo *et al.*, 2005). The confinement to mostly one region hints at two different possibilities for enrichment in mitochondrial DNA: directed repeated insertion into the surroundings of this area, or few insertions with subsequent local amplification of these sequences. Although the second explanation seems more likely, the diverse nature of the organellar reads suggests many independent insertion events instead of many copies derived from one incorporation.

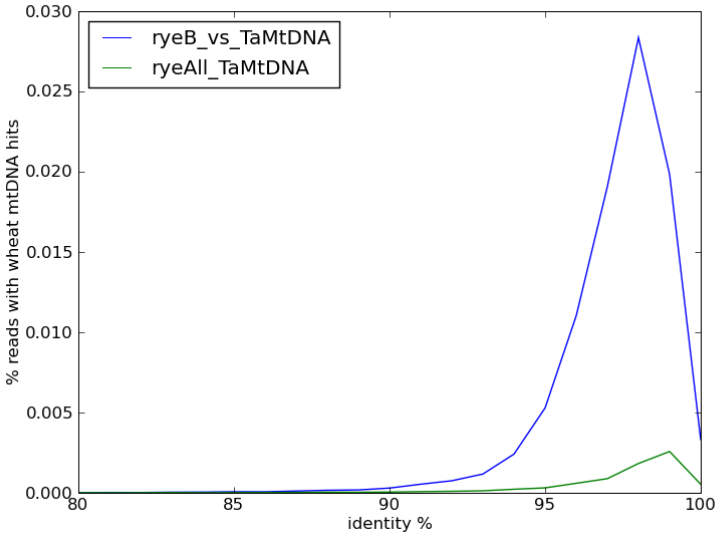
It is known that environmental stresses increase the incorporation of organellar DNA into the nucleus (Wang *et al.*, 2012). As the B, especially in higher numbers, can in itself be considered as a stress factor to the cell (Jones *et al.*, 2008), their presence might increase the basal rate of DNA transfer. Interestingly it has also been shown that organellar DNA often integrates several fragments into one location (Lloyd & Timmis, 2011). This would fit well with our data indicating many events rather than amplification of just one insertion (Martis *et al.*, 2012).

The location of the mitochondrial insertions in the pericentromere is also of importance, as this area is involved in the nondisjunction process that is vital to the survival of the B. Whether the mitochondrial DNA in this location has any effect on the preicentric stickyness, either by strengthening cohesion or simply by physically enlarging the pericentric area, remains to be seen as discussed in detail in Banaei-Moghaddam *et al.*, 2012.

When comparing A- and B-derived organellar reads to the original organelle genome sequence from wheat, we find that, additionally to the accumulation, the reads from the B show less similarity to the source genome (Fig. 28 (Martis *et al.*, 2012)). The inserts on the B therefore accumulated more mutations than A located inserts.



A



B

Fig 28: Comparison of rye A and rye B reads against organellar genomes.

The normalized results of comparisons of repeat-masked rye A and rye B against the complete genomes of wheat organelles are visualized. Comparisons against the (A) wheat chloroplast genome are indicated by TaPt (*Triticum aestivum* plastid genome) and against the (B) wheat mitochondrial genome by TaMt (*Triticum aestivum* mitochondrial genome). The observed identity profiles indicate that rye B carries a higher number of chloroplast and mitochondrial DNA insertions than rye A. (Martis *et al.*, 2012)

The higher divergence of organellar DNA on Bs than on As could either mean the sequences on the B are decaying faster than those on the As, or they are older. Some evidence suggests that the insertions are older. The sequences located on the A chromosomes might be more conserved due to hitchhiking conservation together with large loci of functional importance. The easiest explanation might be that the sequences on the B are indeed older, but not due to conservation of the B, but rather because of high turnover rates in the A genome. Transfer of organellar DNA to the nucleus is very frequent (Huang *et al.*, 2003; Sheppard *et al.*, 2008). But the inserts are also rapidly lost again (Sheppard & Timmis, 2009). The higher selection pressure on the As simply does not allow for permanent invasion. The high turnover rates that keep the sequences on the As from accumulating and degradation would be absent from the B and allow for sequence decay only on the B (Fig. 29).

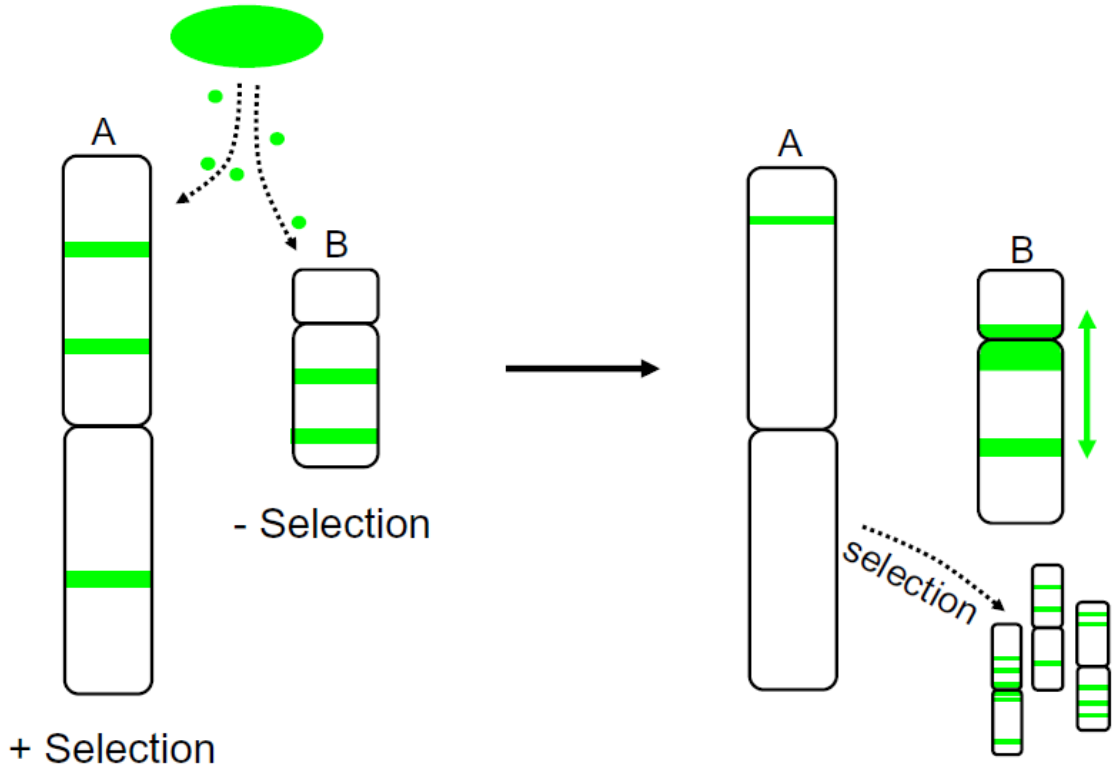


Fig: 29: Organellar DNA is incorporated into As and Bs in similar fashion. Due to selection pressure, the integrations on the As are lost while the ones located on Bs are allowed to accumulate.

4.3 The pericentromere of the B differs from that of the A

We have shown that the DNA composition of the pericentromere of Bs differs from that of As (Banaei-Moghaddam *et al.*, 2012). The extended nature of the B pericentromere has been noted before (Wilkes *et al.*, 1995). Aside from confirming this extended structure, we also found a difference in sequence content. There was a B-specific centromeric sequence proposed before. The sequence R56 contains 61bp (Long *et al.*, 2008) and was proposed to be a B centromeric specific high-copy sequence, but we were unable to detect it in our sequencing data. It might be that the sequence coverage of our analysis was not high enough to find this particular sequence. Also FISH analysis proved inconclusive. Different plant material is less likely the reason for the results due to high conservation between the monophyletic Bs of rye.

The *Copia* element Sc11c32, specifically accumulated in the pericentromere of Bs, might be the result of a chromosome-specific amplification event in that area. Also the centromeric sequences (like *Bilby*, *ScCCS1*) show an extended pericentromeric hybridisation pattern on the Bs compared to the centromeres of As (Banaei-Moghaddam *et al.*, 2012). This indicates that the originally small pericentromere of the proto Bs may have undergone a secondary local amplification that expanded the pre-existent sequences in this area. The possibly ancestral *Copia* elements of Sc11c32 are still detectable on the A chromosomes, but have a subterminal localisation there, further indicating restructuring of the B. Notably, a comparable observation has been made for the B of maize. Here, a sequence has been identified specific to the B centromere, which shows homology over a 90 bp stretch to the maize knob sequences (Alfenito & Birchler, 1993).

Additional to the pericentromeric location, Sc11c32 shows a second locus near the nondisjunction controlling region of the rye B. The same distribution pattern is true for the B pericentric accumulation of mitochondrial sequences and chloroplast sequences on Bs (Martis *et al.*, 2012). Notably the second interstitial chloroplast signal colocalizes with a faint signal of the centromeric sequence

Bilby. The occurrence of these sequences also outside the pericentromere might indicate an ancient inversion with breakpoints at the edge of the pericentromere on the long arm and above the B-specific end, which would create the situation we find now on the estant B (Fig. 30).

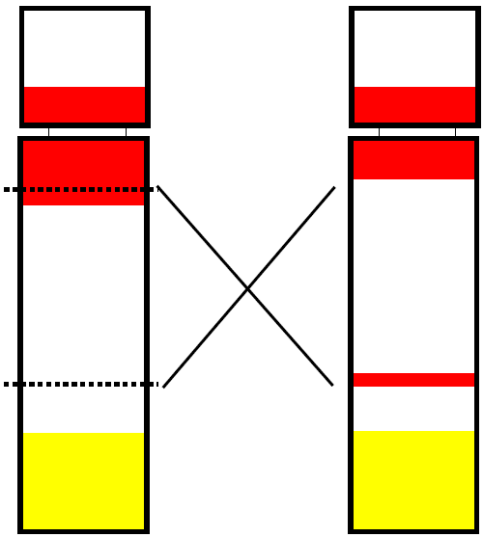


Fig. 30: An ancient inversion event with breakpoints in the pericentromere (red) on the long arm and above the nondisjunction control region (yellow) created a second and interstitial signal of pericentric sequences (Sc11c32, *Bilby*, organellar DNA).

4.4 The B interstitial region hints at frequent restructuring

Several of the B-enriched repeats which were found on the rye B show a tandem-like structure. The micro Bs of *Brachycome dichromosomatica* are mainly composed of tandem repeats (Houben *et al.*, 2001b). Accumulation of tandem repeats was also reported for sex chromosomes e.g. the *Silene latifolia* Y chromosome (Hobza *et al.*, 2006) and it has been proposed that tandem repeats can expand through unequal crossover (Smith, 1976).

Sc55c1 is the first A chromosome subterminal repeat shared by Bs of rye although on the B this sequence is interstitially located. Generally, subtelomeric sequences common to the As are absent from the Bs, e.g. pSc74 (Tsujimoto & Niwa, 1992; Manzanero & Puertas, 2003), pSc119 (Tsujimoto & Niwa, 1992;

Wilkes *et al.*, 1995; Manzanero & Puertas, 2003), pSc200 (Hasterok *et al.*, 2002; Kubaláková *et al.*, 2003; Manzanero & Puertas, 2003; Zhou *et al.*, 2010) and pSc250 (Kubaláková *et al.*, 2003). Other authors have claimed a potential polymorphism on rye Bs, that carry traces of pSc74 and pSc119 in the subtelomeric end of the long arm (Cuadrado & Jouve, 1994; Kubaláková *et al.*, 2003). All the subterminal repeats are extremely variable in rye (Gonzalez-Garcia *et al.*, 2006). Sc55c1 is in close vicinity to several other B-enriched sequences such as Sc63c34 and the second bands of Sc11c32 and mitochondria derived sequences (Martis *et al.*, 2012). Possibly, this region adjacent to the nondisjunction controlling region was frequently restructured. The subtelomeric repeats might be a product of an independent accumulation process, which is supported by the simultaneous increase of Sc63c34, or represents a rudiment of the area from one of the A chromosomes involved in producing the proto-type of the estant B (Martis *et al.*, 2012). Rearrangements and amplifications point to the B-specific terminal part of the long arm as a highly dynamic region. Langdon *et al.*, 2000 showed that E3900 and D1100 repeats evolved from ancestral A-located elements on the B through rearrangements indicating instability of this specific domain.

4.5 The nondisjunction controlling region is composed of multiple B-enriched, partly trancriptionally active satellite sequences

The majority of our newly identified B-enriched satellite repeats map to the nondisjunction controlling region. This region is truly B-specific, as after genomic *in situ* hybridization with labeled 0B DNA this part of the B did not display cross-hybridization. We were able to map the border between the A and B sequence-shared region the subterminal B sequence-enriched region to the late replicating sequence Sc26c38. Similar results are described for maize, where the B also has been shown via renaturation studies to consist mostly of the same repetitive DNA as the As (Chilton and McCarthy, 1973) except for a small terminal chromosome region (Stark et al., 1996). The terminal part of the long

arm of the maize B too harbours a region important to nondisjunction (Ward, 1973).

Four sequences (Sc9c11, Sc9c15, Sc21c9, Sc21c67) colocalized directly with the previously identified B-specific repeat E3900, while no newly identified sequence identified colocalized completely with the D1100-enriched region which spans the whole nondisjunction control region. FISH with either Sc26c38 or Sc9c130 allows reliable determination of B number even in interphase.

We found Sc9c11 to be constitutively transcribed in all three tested tissues in plants with and without Bs, while Sc21c67 showed transcription only in anthers with B chromosomes. Anthers, along with the embryo sacks are the tissue where post meiotic nondisjunction of rye Bs takes place. Previously, the B-enriched repeats D1100 and E3900 were found to be transcriptionally active (Carchilan *et al.*, 2007). It is notable that some newly identified transcriptionally active sequences colocalise with E3900 and partly with D1100. Furthermore, this area is rich in the euchromatic histone mark H3K4me3 and decondenses during interphase (Carchilan *et al.*, 2007). It is flanked at both sides by strongly heterochromatic sequences (Sc26c38 and Sc9c130). We presume that the satellite repeats located in terminal part of the rye B are transcriptionally active only this context. This presumption fits with the finding that D1100, although more abundant, is weaker transcribed than E3900 (Carchilan *et al.*, 2007). It is not known whether these transcripts are involved in the process of B accumulation. We could not show transcription for other identified elements by RT-PCR. However, since we used primers designed on reads stemming from genomic DNA, absence of proof of transcription does not necessarily mean proof of absence of transcription.

When checking the evolutionary older *Secale* species *S. strictum* and *S. sylvestre* (De Bustos & Jouve, 2002; Chikmawati *et al.*, 2005; Shang *et al.*, 2006) for the presence of rye B-enriched sequences, we found several of them also

present. Most notable was a local amplification of Sc26c38 in the elder *Secale* species, *S*. *sylvestre*. We assume this amplification happened independently of that on the B. Because this amplification is close to the centromere in *S. sylvestre* and terminal in the B, and we find no such local amplification in *S. strictum*, which is closer to *S. cereale* than *S. sylvestre*. All in all, our research indicates that the sequences on the B are not new, but evolved by amplification from preexisting sequences after being separated from the standard chromosome complement.

We demonstrated that the Sc26c38-enriched part of the nondisjunction control region is the last component of the genome to replicate in S-phase and therefore undergoes replication even later than all heterochromatic regions of the A chromosomes. The late replication of the B-specific end of the supernumerary chromosome has been reported previously (Lima-De-Faria & Jaworska, 1972; Ayonoadu & Rees, 1973). It has been suggested that late replication areas are poor in replication origins (Debatisse *et al.*, 2012). It is conceivable that Sc26c38 lacks these origins due to its sequence. The sequence has no apparent similarities to known elements. Late replication was also shown for heterochromatic parts of the maize B and also there suspected to be involved in the nondisjunction mechanism (Pryor *et al.*, 1980). Whether functional interrelationship exists between delayed replication and nondisjunction of rye Bs during first pollen mitosis is not known, but appears to be unlikely as the nondisjunction regions also acts *in trans* (Lima-de-Faria, 1962; Endo *et al.*, 2008). Physical interaction of the chromatids at this position during nondisjunction is not present as the sticking sites reside at the pericentromere. However, it is still possible that an up to now unknown transcript is involved in the nondisjunction process. It is however noticeable that this strongly heterochromatic Sc26c38-positive region directly borders one side of the transcriptionally active region of the nondisjunction control region, possibly as a physical barrier for transcription.

4.6 The new B markers provide a valuable tool for future research

Previously, the interphase territory of rye Bs has been visualized by GISH using a wheat line carrying rye Bs and a GISH-probe derived from rye. This analysis allowed to show that Bs exhibit similar interphase organization as As except for prolonged centromere association during interphase when the nondisjunction region is present (Schubert *et al.*, 2011). However, it was unknown whether the B shows the same interphase behavior in the rye background.

The identification of a set of B-enriched high copy sequences provides a tool for visualizing interphase Bs in rye. For this we employed the newly identified probe Sc36c82 which is highly enriched on B chromosomes. This enrichment is also visible during interphase. Additional use of mitochondrial DNA to mark the centromere and D1100 as a large marker of the long arm, allowed to identify orientation of the interphase B. Thus we are for the first time able to reliably distinguish The B chromosome in interphase in its native genomic background (Fig. 31). This should open new possibilities to extend the research conducted up to now, as we before were limited to either metaphase stages or the nondisjunction area, which was the only part easily identifiable. In combination with immunodetection it could help to identify key proteins differently distributed to As and Bs.

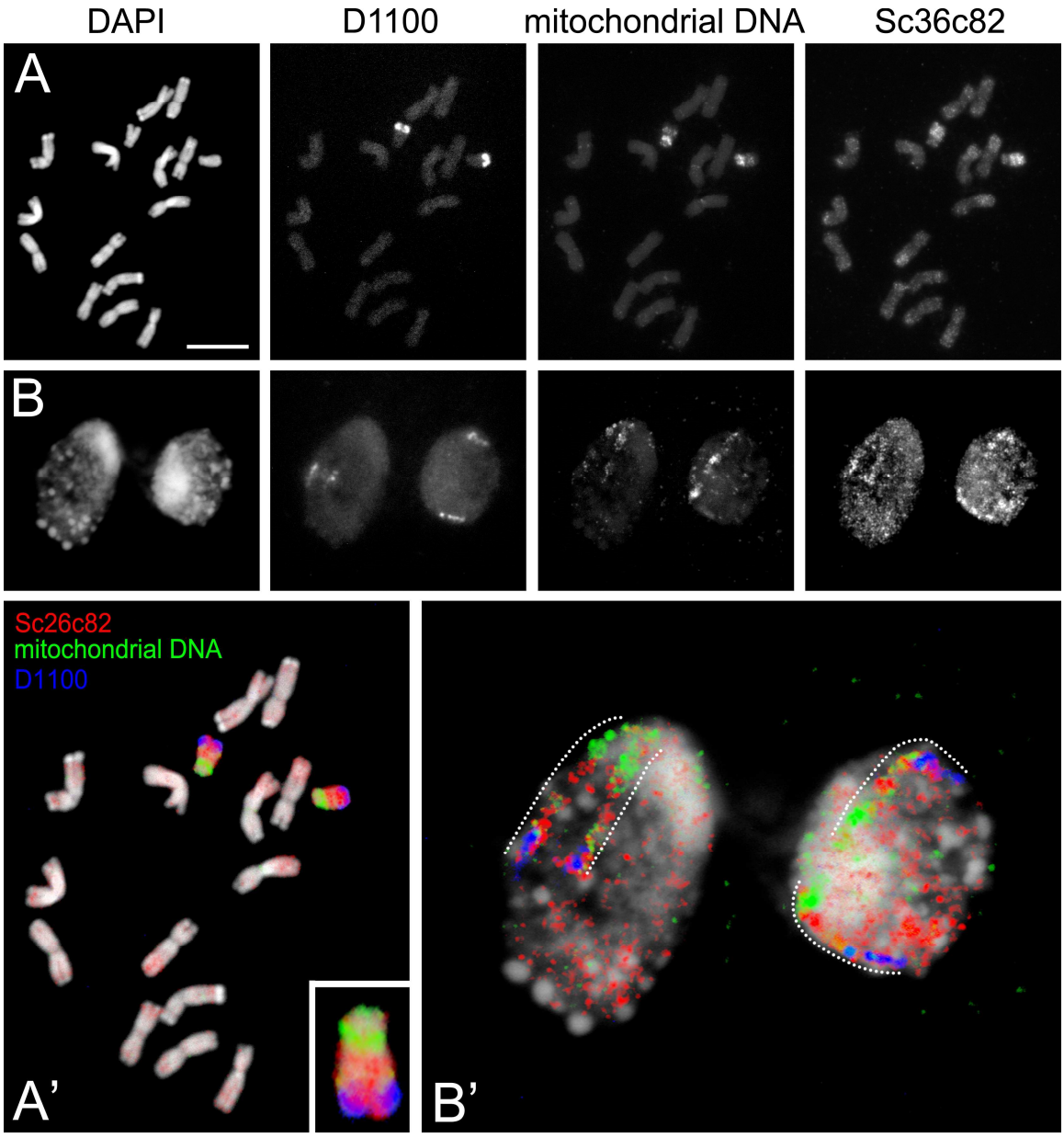


Fig. 31: interphase painting of the rye B chromosome in native genomic background. The B can be reliably marked over its entire length in A) a metaphase and B) an interphase nucleus with the B amplified probe Sc36c82 (red), mitochondrial DNA (green) for the pericentromere and D1100 (yellow) for the the long arm.

The two strongly heterochromatic satellites Sc26c38 and Sc9c130 do not disperse during interphase. This allows for easier counting of Bs in interphase nuclei than E3900 or D1100. Since the number of B chromosomes vary between individuals, each plant has to be evaluated for B chromosome number. The new markers are more reliable and comfortable in determining the number of Bs per cell.

4.7 Sister chromatid exchange rates of As and Bs of rye

To check for possible difference in somatic homologous recombination frequency between Bs and As, the rate of sister chromatid exchanges (SCEs) in rye has been determined. Previously, 0.79 SCEs per chromosome were reported (Friebe, 1978). The frequency of SCEs in chromosome 1R was determined separately as it is easily distinguishable from the rest of the rye chromosomes. The SCE frequency on 1R with 0.76 SCEs per chromosome is comparable to the whole genome, indicating that most likely all chromosomes have a similar exchange rate (Friebe, 1978). The presence of B chromosomes appeared to have no effect on the frequency of SCEs in the A chromosomes. Also, B chromosomes had a similar SCE rate (Friebe, 1980).

Compared to the previously determined rates of 0.79 SCE per chromosome (Friebe, 1978), the frequency of SCEs determined here is much higher (3.4 SCE per chromosome). Already from the visuals every chromosome usually displays several SCEs and chromosomes without SCE are comparatively rare. Fig. 26B showes a single chromosome without visible SCE. The great discrepancy of the two methods might be due to the great advances achieved in microscopy. The increased visibility of SCEs naturally leads to a higher number in measurement. A high variation due to different experimental conditions has been shown even between similar experiments (Schvartzman, 1987). A difference of SCE frequency between accessions is possible, although no significant differences between three accessions of rye were observed (Friebe, 1980). We also found no difference in SCE distribution between the plants tested. A last possibility to

achieve heightened rates of SCE would be the use of another base analogue. Indeed, 5-Bromdesoxyuridin (BrdU) is supposed to induce SCEs (Friebe, 1978). However, since both studies employed a similar method, this should not cause major differences in number of SCEs. If at all, EdU supposedly causes less SCEs than BrdU.

The rates in the long arms of the chromosomes are similar. The frequency on the short arm of the B chromosome seems to be much lower than that on the corresponding area on the A chromosome (region 3). This is most likely due to the considerably shorter length of the B chromosome arm, resulting in proportionally less DNA damage.

In our data, none of the observed increased frequencies are statistically significant. This fits with the previous observation that there is no non-random involvement of heterochromatin in spontaneous SCEs (Schubert & Rieger, 1981).

5. Outlook

We showed that although rye Bs are similar to As in overall DNA content, they do vary in their repetitive composition. The regions most important to the vital nondisjunction mechanism (pericentromere and terminal part of the long arm) are particularly outstanding. Clear differences in replication behaviour and mobile element content highlight the different evolutionary pathways of As and Bs. Combined with the young age of the rye B of around 1.1-1.3 million years (Martis *et al.*, 2012) these conditions make the Bs of rye a valuable tool to research the basic forces that act in chromosome evolution. For this future research, we established a comprehensive map of marker sequences that indicate the B-specific pericentromere, the interstitial region and the nondisjunction area.

Recent advances in sequencing and bioinformatic tools have enabled the snapshot sequencing of the B chromosome of rye (Martis *et al.*, 2012). This data yielded not only knowledge of composition, but also new tools. New sequences have been identified, which enable finemapping of the rye B with respect to functional properties such as replication behaviour (Klemme *et al.*, 2013). They also allowed closer comparison between variants of different origins (Marques *et al.*, 2013). Another resulting tool is interphase painting of B chromosomes in their native environment (Fig 33). Due to the high homogenous distribution of the mobile elements within the rye genome, this was previously only possible with rye Bs introgressed into wheat (Morais-Cecilio *et al.*, 1997; Schubert *et al.*, 2011). Using GISH the number of B-specific signals in nuclei with two, three and four Bs was often lower than the maximum B number. In the background of wheat, Bs may be grouped together in different patters rather than being randomly distributed throughout the nucleus (Morais-Cecilio *et al.*, 1997). Whether the same occurs in the background of its native host genome is not known. Applying a combination of B-specific repeats provides new opportunities as it opens up possibilities to research B behaviour in all cell cycle stages.

As said before, Bs of rye and maize share many similarities, although they arose independently. Thus the event dynamics shaping these extra chromosomes might be similar in different species. It will be interesting to see how much of the intricacies found on the rye Bs are applicable to other B chromosomes. Finding similarities and differences between diverse types of Bs of different species will give inside into general evolutionary forces acting during chromosome evolution. For example, it seems to be a general feature of Bs to accumulate sequences from different sources and further studies may hint at the mechanisms involved in their incorporation and multiplication. To this aim, it would be interesting to test, if also Bs of other species accumulate organellar DNA, because this type of DNA is readily available due to organelle decay. Another interesting aspect of karyotype evolution is the formation of new chromosomes, that can no longer undergo recombination with their parental units. Comparative studies in many species hint at a connection of B formation to large genome size and low chromosome number (Palestis *et al.*, 2004b; Trivers *et al.*, 2004). Differences in mobile element distribution between As and Bs could give insight to general evolutionary constraints of transposition. Comparison of features shared by many Bs and special features of single B types could differentiate between evolutionary forces and genomic background constraints. For this it would be especially interesting to compare Bs from smaller genomes and Bs from very large genomes.

Due to their unusual features, B chromosomes are often dismissed as inconsequential. This can mostly be attributed to the dispensable nature and the apparent lack of adaptive significance. However, it is widely accepted that we can learn more from observing the abnormal, than from looking at the normal. This chromosome can teach us about evolutionary behaviour of chromosomes under varying amounts of selection pressure. Many Bs behave like a selfish entity and exploit the cellular machinery to ensure their transmission and survival. There is still much to learn from these enigmatic chromosomes.

References

Alfenito MR, Birchler JA. 1993. Molecular characterization of a maize B chromosome centric sequence. *Genetics* **135**(2): 589-597.

Auger DL, Newton KJ, Birchler JA. 2001. Nuclear gene dosage effects upon the expression of maize mitochondrial genes. *Genetics* **157**(4): 1711-1721.

Ayonoadu U, Rees H. 1973. DNA synthesis in rye chromosomes. *Heredity* **30**(2): 233-240.

Banaei-Moghaddam AM, Schubert V, Kumke K, Weiβ O, Klemme S, Nagaki K, Macas J, González-Sánchez M, Heredia V, Gómez-Revilla D, González-García M, Vega JM, Puertas MJ, Houben A. 2012. Nondisjunction in Favor of a Chromosome: The Mechanism of Rye B Chromosome Drive during Pollen Mitosis. *The Plant Cell Online* **24**(10): 4124-4134.

Bartos J, Paux E, Kofler R, Havrankova M, Kopecky D, Suchankova P, Safar J, Simkova H, Town CD, Lelley T, Feuillet C, Dolezel J. 2008. A first survey of the rye (Secale cereale) genome composition through BAC end sequencing of the short arm of chromosome 1R. *Bmc Plant Biology* **8**: -.

Battaglia E. 1964. Cytogenetics of B-chromosomes. *Caryologia* **17**(1): 245-299.

Bedini G, Garbari F, Peruzzi L. 2011. Karyological knowledge of the Italian vascular flora as inferred by the analysis of "Chrobase.it". *Plant Biosystems - An International Journal Dealing with all Aspects of Plant Biology* **146**(4): 889-899.

Birchler JA, Chalfoun DJ, Levin DM. 1990. Recombination in the B chromosome of maize to produce A-B-A chromosomes. *Genetics* **126**(3): 723-733.

Blunden R, Wilkes TJ, Forster JW, Jimenez MM, Sandery MJ, Karp A, Jones RN. 1993. Identification of the E3900 Family, a 2nd Family of Rye Chromosome-B Specific Repeated Sequences. *Genome* **36**(4): 706-711.

Bugrov AG, Karamysheva TV, Perepelov EA, Elisaphenko EA, Rubtsov DN, Warchalowska-Sliwa E, Tatsuta H, Rubtsov NB. 2007. DNA content of the B chromosomes in grasshopper Podisma kanoi Storozh. (Orthoptera, Acrididae). *Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology* **15**(3): 315-325.

Carchilan M, Delgado M, Ribeiro T, Costa-Nunes P, Caperta A, Morais-Cecilio L, Jones RN, Viegas W, Houben A. 2007. Transcriptionally active heterochromatin in rye B chromosomes. *Plant Cell* **19**(6): 1738-1749.

Carchilan M, Houben A. 2007. Transcription analysis of rye B chromosomes. *Chromosome Research* **15**: 72-72.

Carchilan M, Kumke K, Mikolajewski S, Houben A. 2009. Rye B chromosomes are weakly transcribed and might alter the transcriptional activity of A chromosome sequences. *Chromosoma* **118**(5): 607-616.

Carlson WR. 1978. The B chromosome of corn. *Annual review of genetics* **12**: 5-23.

Carlson WR, Roseman RR. 1992. A new property of the maize B chromosome. *Genetics* **131**(1): 211-223.

Carter CR, Smith-White S. 1972. The cytology of Brachycome lineariloba. *Chromosoma* **39**(4): 361-379.

Cermak T, Kubat Z, Hobza R, Koblizkova A, Widmer A, Macas J, Vyskot B, Kejnovsky E. 2008. Survey of repetitive sequences in Silene latifolia with respect to their distribution on sex chromosomes. *Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology* **16**(7): 961-976.

Charles M, Belcram H, Just J, Huneau C, Viollet A, Couloux A, Segurens B, Carter M, Huteau V, Coriton O, Appels R, Samain S, Chalhoub B. 2008. Dynamics and differential proliferation of transposable elements during the evolution of the B and A genomes of wheat. *Genetics* **180**(2): 1071-1086.

Charlesworth B, Sniegowski P, Stephan W. 1994. The evolutionary dynamics of repetitive DNA in eukaryotes. *Nature* **371**(6494): 215-220.

Charlesworth D. 2008. Plant sex chromosomes. *Genome Dyn* **4**: 83-94.

Cheng YM. 2010. Evolution of the heterochromatic regions on maize B long arm based on the sequence structure of CL-repeat variants. *Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology* **18**(5): 605-619.

Cheng YM, Lin BY. 2003. Cloning and characterization of maize B chromosome sequences derived from microdissection. *Genetics* **164**(1): 299-310.

Chikmawati T, Skovmand B, Gustafson JP. 2005. Phylogenetic relationships among Secale species revealed by amplified fragment length polymorphisms. *Genome* **48**(5): 792-801.

Cuadrado A, Cardoso M, Jouve N. 2008. Physical organisation of simple sequence repeats (SSRs) in Triticeae: structural, functional and evolutionary implications. *Cytogenetic and Genome Research* **120**(3-4): 210-219.

Cuadrado A, Jouve N. 1994. Highly repetitive sequences in B chromosomes of Secale cereale revealed by fluorescence in situ hybridization. *Genome* **37**(4): 709-712.

De Bustos A, Jouve N. 2002. Phylogenetic relationships of the genus Secale based on the characterisation of rDNA ITS sequences. *Plant Systematics and Evolution* **235**(1-4): 147-154.

Debatisse M, Le Tallec B, Letessier A, Dutrillaux B, Brison O. 2012. Common fragile sites: mechanisms of instability revisited. *Trends in Genetics* **28**(1): 22-32.

Deragon JM, Casacuberta JM, Panaud O. 2008. Plant transposable elements. *Genome Dyn* **4**: 69-82.

Dhar MK, Friebe B, Koul AK, Gill BS. 2002. Origin of an apparent B chromosome by mutation, chromosome fragmentation and specific DNA sequence amplification. *Chromosoma* **111**(5): 332-340.

Diez M, Jiménez MM, Santos JL. 1993. Synaptic patterns of rye B chromosomes. II. The effect of the standard B chromosomes on the pairing of the A set. *TAG Theoretical and Applied Genetics* **87**(1): 17-21.

Dolezel J, Bartos J, Voglmayr H, Greilhuber J. 2003. Nuclear DNA content and genome size of trout and human. *Cytometry. Part A: The Journal of the International Society for Analytical Cytology* **51**(2): 127.

Dolezel J, Greilhuber J, Suda J. 2007. Estimation of nuclear DNA content in plants using flow cytometry. *Nat. Protocols* **2**(9): 2233-2244.

Donald TM, Leach CR, Clough A, Timmis JN. 1995. Ribosomal RNA genes and the B chromosome of Brachycome dichromosomatica. *Heredity* **74**(5): 556-561.

Endo TR, Nasuda S, Jones N, Dou Q, Akahori A, Wakimoto M, Tanaka H, Niwa K, Tsujimoto H. 2008. Dissection of rye B chromosomes, and nondisjunction properties of the dissected segments in a common wheat background. *Genes & Genetic Systems* **83**(1): 23-30.

Filatov DA, Howell EC, Groutides C, Armstrong SJ. 2009. Recent spread of a retrotransposon in the Silene latifolia genome, apart from the Y chromosome. *Genetics* **181**(2): 811-817.

Frederiksen S, Petersen G. 1998. A taxonomic revision of Secale (Triticeae, Poaceae). *Nordic Journal of Botany* **18**(4): 399-420.

Friebe B. 1978. Untersuchungen zum Schwesterchromatidenaustausch bei Secale cereale. *Microscopica Acta* **81**(2): 159-165.

Friebe B. 1980. Lack of effect of B chromosomes on sister chromatid exchanges in Secale cereale. *Environmental and Experimental Botany* **20**(1): 21-26.

Galbraith DW, Harkins KR, Maddox JM, Ayres NM, Sharma DP, Firoozabady E. 1983. Rapid Flow Cytometric Analysis of the Cell Cycle in Intact Plant Tissues. *Science* **220**(4601): 1049-1051.

Gernand D, Rutten T, Pickering R, Houben A. 2006. Elimination of chromosomes in Hordeum vulgare x H. bulbosum crosses at mitosis and interphase involves micronucleus formation and progressive heterochromatinization. *Cytogenet Genome Res* **114**(2): 169-174.

Gonzalez-Garcia M, Gonzalez-Sanchez M, Puertas MJ. 2006. The high variability of subtelomeric heterochromatin and connections between nonhomologous chromosomes, suggest frequent ectopic recombination in rye meiocytes. *Cytogenetic and Genome Research* **115**(2): 179-185.

Håkanson A. 1959. Behaviour of different small accessry rye chromosomes at pollen mitosis. . *Hereditas* **45**: 623-631.

Hasterok R, Jenkins G, Langdon T, Jones RN. 2002. The nature and destiny of translocated B-chromosome-specific satellite DNA of rye. *Chromosome Research* **10**(1): 83-86.

Heckmann S, Macas J, Kumke K, Fuchs J, Schubert V, Ma L, Novák P, Neumann P, Taudien S, Platzer M, Houben A. 2013. The holocentric species Luzula elegans shows interplay between centromere and large-scale genome organization. *The Plant Journal* **73**(4): 555-565.

Hobza R, Lengerova M, Svoboda J, Kubekova H, Kejnovsky E, Vyskot B. 2006. An accumulation of tandem DNA repeats on the Y chromosome in Silene latifolia during early stages of sex chromosome evolution. *Chromosoma* **115**(5): 376-382.

Houben A, Banaei-Moghaddam A, Klemme S 2013a. Biology and Evolution of B Chromosomes. In: Greilhuber J, Dolezel J, Wendel JF eds. *Plant Genome Diversity Volume 2*: Springer Vienna, 149-165.

Houben A, Banaei-Moghaddam AM, Klemme S, Timmis JN. 2013b. Evolution and biology of supernumerary B chromosomes. *submitted*.

Houben A, Belyaev ND, Leach CR, Timmis JN. 1997a. Differences of histone H4 acetylation and replication timing between A and B chromosomes of brachycome dichromosomatica. *Chromosome Res* **5**(4): 233-237.

Houben A, Demidov D, Gernand D, Meister A, Leach CR, Schubert I. 2003. Methylation of histone H3 in euchromatin of plant chromosomes depends on

basic nuclear DNA content. *The Plant journal : for cell and molecular biology* **33**(6): 967-973.

Houben A, Field BL, Saunders VA. 2001a. Microdissection and chromosome painting of plant B chromosomes. *Methods Cell Sci* **23**(1-3): 115-124.

Houben A, Kynast RG, Heim U, Hermann H, Jones RN, Forster JW. 1996. Molecular cytogenetic characterisation of the terminal heterochromatic segment of the B-chromosome of rye (Secale cereale). *Chromosoma* **105**(2): 97-103.

Houben A, Leach CR, Verlin D, Rofe R, Timmis JN. 1997b. A repetitive DNA sequence common to the different B chromosomes of the genus Brachycome. *Chromosoma* **106**(8): 513-519.

Houben A, Verlin D, Leach CR, Timmis JN. 2001b. The genomic complexity of micro B chromosomes of Brachycome dichromosomatica. *Chromosoma* **110**(7): 451-459.

Huang CY, Ayliffe MA, Timmis JN. 2003. Direct measurement of the transfer rate of chloroplast DNA into the nucleus. *Nature* **422**(6927): 72-76.

Huang S, Sirikhachornkit A, Su X, Faris J, Gill B, Haselkorn R, Gornicki P. 2002. Genes encoding plastid acetyl-CoA carboxylase and 3-phosphoglycerate kinase of the Triticum/Aegilops complex and the evolutionary history of polyploid wheat. *Proceedings of the National Academy of Sciences* **99**(12): 8133-8138.

Jimenez G, Manzanero S, Puertas M. 2000. Relationship between pachytene synapsis, metaphase I associations, and transmission of 2B and 4B chromosomes in rye. *Genome* **43**(2): 232-239.

Jimenez MM, Romera F, Puertas MJ, Jones RN. 1994. B-Chromosomes in inbred lines of rye (Secale cereale L). 1. Vigor and fertility. *genetica* **92**(3): 149-154.

Jin W, Lamb JC, Vega JM, Dawe RK, Birchler JA, Jiang J. 2005. Molecular and functional dissection of the maize B chromosome centromere. *Plant Cell* **17**(5): 1412-1423.

John UP, Leach CR, Timmis JN. 1991. A sequence specific to B chromosomes of Brachycome dichromosomatica. *Genome* **34**(5): 739-744.

Jones N. 2012. B chromosomes in plants. *Plant Biosystems - An International Journal Dealing with all Aspects of Plant Biology* **146**(3): 727-737.

Jones N, Houben A. 2003. B chromosomes in plants: escapees from the A chromosome genome? *Trends in Plant Science* **8**(9): 417-423.

Jones N, Puertas MJ. 1993. The B-Chromosomes of Rye
(Secale cereale L.). *Frontiers in Plant Science Research*: 81-112.

Jones RN. 1991. B-Chromosome Drive. *American Naturalist* **137**(3): 430-442.

Jones RN, Gonzalez-Sanchez M, Gonzalez-Garcia M, Vega JM, Puertas MJ. 2008. Chromosomes with a life of their own. *Cytogenetic and Genome Research* **120**(3-4): 265-280.

Jones RN, Rees H. 1968. The influence of B-chromosomes upon the nuclear phenotype in rye. *Chromosoma* **24**(2): 158-176.

Jones RN, Rees H. 1982. *B chromosomes*. London; New York: Academic Press.

Kato A. 1997. An improved method for chromosome counting in maize. *Biotech Histochem* **72**(5): 249-252.

Kejnovsky E, Kubat Z, Hobza R, Lengerova M, Sato S, Tabata S, Fukui K, Matsunaga S, Vyskot B. 2006. Accumulation of chloroplast DNA sequences on the Y chromosome of Silene latifolia. *Genetica* **128**(1-3): 167-175.

Kejnovsky E, Vyskot B. 2010. Silene latifolia: the classical model to study heteromorphic sex chromosomes. *Cytogenetic and Genome Research* **129**(1-3): 250-262.

Kleine T, Maier UG, Leister D. 2009. DNA transfer from organelles to the nucleus: the idiosyncratic genetics of endosymbiosis. *Annu Rev Plant Biol* **60**: 115-138.

Klemme S, Banaei-Moghaddam AM, Macas J, Wicker T, Novák P, Houben A. 2013. High-copy sequences reveal distinct evolution of the rye B chromosome. *New Phytologist* **199**(2): 550-558.

Kubaláková M, Valárik M, Bartoš J, Vrána J, Číhalíková J, Molnár-Láng M, Doležel J. 2003. Analysis and sorting of rye (Secale cereale L.) chromosomes using flow cytometry. *genome* **46**(5): 893-905.

Kumke K, Jones RN, Houben A. 2008. B chromosomes of Puschkinia libanotica are characterized by a reduced level of euchromatic histone H3 methylation marks. *Cytogenetic and Genome Research* **121**(3-4): 266-270.

Lamb JC, Riddle NC, Cheng YM, Theuri J, Birchler JA. 2007. Localization and transcription of a retrotransposon-derived element on the maize B chromosome. *Chromosome Research* **15**(3): 383-398.

Langdon T, Seago C, Jones RN, Ougham H, Thomas H, Forster JW, Jenkins G. 2000. De novo evolution of satellite DNA on the rye B chromosome. *Genetics* **154**(2): 869-884.

Leach CR, Houben A, Field B, Pistrick K, Demidov D, Timmis JN. 2005. Molecular evidence for transcription of genes on a B chromosome in Crepis capillaris. *Genetics*.

Lima-de-Faria A. 1962. Genetic interaction in rye expressed at chromosome phenotype. *Genetics* **47**: 1455-1462.

Lima-De-Faria A, Jaworska H. 1972. The relation between the chromosome size gradient and the sequence of DNA replication in rye. *Hereditas* **70**(1): 39-57.

Lin BY, Chou HP. 1997. Physical mapping of four RAPDs in the B chromosome of maize. *Theoretical and Applied Genetics* **94**(3-4): 534-538.

Lindström J. 1965. Transfer to wheat of accessory chromosomes from rye. *Hereditas* **54**(2): 149-155.

Lloyd AH, Rousseau-Gueutin M, Timmis JN, Sheppard AE, Ayliffe MA 2012. Promiscuous Organellar DNA
Genomics of Chloroplasts and Mitochondria. In: Bock R, Knoop V eds.: Springer Netherlands, 201-221.

Lloyd AH, Timmis JN. 2011. The Origin and Characterization of New Nuclear Genes Originating from a Cytoplasmic Organellar Genome. *Molecular Biology and Evolution* **28**(7): 2019-2028.

Long H, Qi ZX, Sun XM, Chen CB, Lie XL, Song WQ, Chen RY. 2008. Characters of DNA constitution in the rye B chromosome. *Journal of Integrative Plant Biology* **50**(2): 183-189.

Ma L, Vu GT, Schubert V, Watanabe K, Stein N, Houben A, Schubert I. 2010. Synteny between Brachypodium distachyon and Hordeum vulgare as revealed by

FISH. *Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology* **18**(7): 841-850.

Macas J, Kejnovsky E, Neumann P, Novak P, Koblizkova A, Vyskot B. 2011. Next Generation Sequencing-Based Analysis of Repetitive DNA in the Model Dioceous Plant Silene latifolia. *PLoS One* **6**(11): e27335.

Macas J, Neumann P, Navratilova A. 2007. Repetitive DNA in the pea (Pisum sativum L.) genome: comprehensive characterization using 454 sequencing and comparison to soybean and Medicago truncatula. *Bmc Genomics* **8**: -.

Maistro EL, Foresti F, Oliveira C, Almeida Toledo LF. 1992. Occurrence of macro B chromosomes in Astyanax scabripinnis paranae (Pisces, Characiformes, Characidae). *Genetica* **87**(2): 101-106.

Manzanero S, Puertas MJ. 2003. Rye terminal neocentromeres: characterisation of the underlying DNA and chromatin structure. *Chromosoma* **111**(6): 408-415.

Marques A, Banaei-Moghaddam AM, Klemme S, Blattner FR, Niwa K, Guerra M, Houben A. 2013. B chromosomes of rye are highly conserved and accompanied the development of early agriculture. *Annals of Botany*.

Marques A, Klemme S, Guerra M, Houben A. 2012. Cytomolecular characterization of de novo formed rye B chromosome variants. *Molecular cytogenetics* **5**(1): 34.

Marschner S, Kumke K, Houben A. 2007a. B chromosomes of B-dichromosomatica show a reduced level of euchromatic histone H3 methylation marks. *Chromosome Research* **15**(2): 215-222.

Marschner S, Meister A, Blattner FR, Houben A. 2007b. Evolution and function of B chromosome 45S rDNA sequences in Brachycome dichromosomatica. *genome* **50**(7): 638-644.

Martis MM, Klemme S, Banaei-Moghaddam AM, Blattner FR, Macas J, Schmutzer T, Scholz U, Gundlach H, Wicker T, Šimková H, Novák P, Neumann P, Kubaláková M, Bauer E, Haseneyer G, Fuchs J, Doležel J, Stein N, Mayer KFX, Houben A. 2012. Selfish supernumerary chromosome reveals its origin as a mosaic of host genome and organellar sequences. *Proceedings of the National Academy of Sciences* **109**(33): 13343-13346.

Matsuo M, Ito Y, Yamauchi R, Obokata J. 2005. The rice nuclear genome continuously integrates, shuffles, and eliminates the chloroplast genome to cause chloroplast-nuclear DNA flux. *The Plant cell* **17**(3): 665-675.

Matsuoka Y, Tsunewaki K. 1997. Presence of wheat retrotransposons in Gramineae species and the origin of wheat retrotransposon families. *Genes & Genetic Systems* **72**(6): 335-343.

McIntyre CL, Pereira S, Moran LB, Appels R. 1990. New Secale cereale (rye) DNA derivatives for the detection of rye chromosome segments in wheat. *Genome* **33**(5): 635-640.

McQuade LR, Hill RJ, Francis D. 1994. B-chromosome systems in the greater glider, Petauroides volans (Marsupialia: Pseudocheiridae). II. Investigation of B-chromosome DNA sequences isolated by micromanipulation and PCR. *Cytogenet Cell Genet* **66**(3): 155-161.

Mendelson D, Zohary D. 1972. Behaviour and transmission of supernumerary chromosomes in Aegilops speltoides. *Heredity* **29**(3): 329-339.

Messing J, Bennetzen JL. 2008. Grass genome structure and evolution. *Genome Dyn* **4**: 41-56.

Meyer JR. 1944. Chromosome Studies of Phlox. *Genetics* **29**(2): 199-216.

Morais-Cecilio L, Delgado M, Jones RN, Viegas W. 1997. Interphase arrangement of rye B chromosomes in rye and wheat. *Chromosome Res* **5**(3): 177-181.

Müntzing A. 1944. Cytological studies of extra fragment chromosomes in rye: I. Iso-fragments produced by misdivision. *Hereditas* **30**(1-2): 231-248.

Müntzing A. 1945. Cytological studies of extra fragment chromosomes in rye II. Transmission and multiplication of standard fragments and iso-fragments. *Hereditas* **31**(3-4): 457-477.

Müntzing A. 1948. Cytological studies of extra fragment chromosomes in rye. V. A new fragment type arisen by deletion. *Hereditas* **34**: 435-442.

Niwa K, Sakamoto S. 1995. Origin of B-Chromosomes in Cultivated Rye. *Genome* **38**(2): 307-312.

Novak P, Neumann P, Macas J. 2010. Graph-based clustering and characterization of repetitive sequences in next-generation sequencing data. *BMC bioinformatics* **11**: 378.

Nur U, Werren JH, Eickbush DG, Burke WD, Eickbush TH. 1988. A "selfish" B chromosome that enhances its transmission by eliminating the paternal genome. *Science* **240**(4851): 512-514.

Pagan HJ, Macas J, Novak P, McCulloch ES, Stevens RD, Ray DA. 2012. Survey sequencing reveals elevated DNA transposon activity, novel elements, and variation in repetitive landscapes among vesper bats. *Genome biology and evolution* **4**(4): 575-585.

Page BT, Wanous MK, Birchler JA. 2001. Characterization of a maize chromosome 4 centromeric sequence: evidence for an evolutionary relationship with the B chromosome centromere. *Genetics* **159**(1): 291-302.

Palestis BG, Burt A, Jones RN, Trivers R. 2004a. B chromosomes are more frequent in mammals with acrocentric karyotypes: support for the theory of centromeric drive. *Proceedings. Biological sciences / The Royal Society* **271 Suppl 3**: S22-24.

Palestis BG, Trivers R, Burt A, Jones RN. 2004b. The distribution of B chromosomes across species. *Cytogenetic and Genome Research* **106**(2-4): 151-158.

Pedro J, Camacho M 2005. B Chromosomes. In: Gregory TR ed. *The Evolution of the Genome*: Elsevier/Academic Press.

Pereira HS, Barao A, Caperta A, Rocha J, Viegas W, Delgado M. 2009. Rye Bs disclose ancestral sequences in cereal genomes with a potential role in gametophyte chromatid segregation. *Molecular Biology and Evolution* **26**(8): 1683-1697.

Poletto AB, Ferreira IA, Martins C. 2010. The B chromosomes of the African cichlid fish Haplochromis obliquidens harbour 18S rRNA gene copies. *BMC Genet* **11**: 1.

Pryor A, Faulkner K, Rhoades MM, Peacock WJ. 1980. Asynchronous replication of heterochromatin in maize. *Proceedings of the National Academy of Sciences of the United States of America* **77**(11): 6705-6709.

Puertas MJ. 2002. Nature and evolution of B chromosomes in plants: A non-coding but information-rich part of plant genomes. *Cytogenetic and Genome Research* **96**(1-4): 198-205.

Qi ZX, Zeng H, Li XL, Chen CB, Song WQ, Chen RY. 2002. The molecular characterization of maize B chromosome specific AFLPs. *Cell Research* **12**(1): 63-68.

Ragupathy R, Cloutier S. 2008. Genome organisation and retrotransposon driven molecular evolution of the endosperm Hardness (Ha) locus in Triticum aestivum cv Glenlea. *Mol Genet Genomics* **280**(6): 467-481.

Raman R, Sharma T. 1974. DNA replication, G- and C-bands and meiotic behaviour of supernumerary chromosomes of Rattus rattus (Linn.). *Chromosoma* **45**(1): 111-119.

Renny-Byfield S, Chester M, Kovařík A, Le Comber SC, Grandbastien M-A, Deloger M, Nichols RA, Macas J, Novák P, Chase MW, Leitch AR. 2011. Next Generation Sequencing Reveals Genome Downsizing in Allotetraploid Nicotiana tabacum, Predominantly through the Elimination of Paternally Derived Repetitive DNAs. *Molecular Biology and Evolution* **28**(10): 2843-2854.

Renny-Byfield S, Kovarik A, Chester M, Nichols RA, Macas J, Novak P, Leitch AR. 2012. Independent, rapid and targeted loss of highly repetitive DNA in natural and synthetic allopolyploids of Nicotiana tabacum. *PLoS One* **7**(5): e36963.

Rimpau J, Flavell RB. 1975. Characterization of Rye B Chromosome DNA by DNA-DNA Hybridization. *Chromosoma* **52**(3): 207-217.

Sandery MJ, Forster JW, Blunden R, Jones RN. 1990. Identification of a Family of Repeated Sequences on the Rye B-Chromosome. *Genome* **33**(6): 908-913.

SanMiguel PJ, Ramakrishna W, Bennetzen JL, Busso CS, Dubcovsky J. 2002. Transposable elements, genes and recombination in a 215-kb contig from wheat chromosome 5A(m). *Funct Integr Genomics* **2**(1-2): 70-80.

Schartl M, Nanda I, Schlupp I, Wilde B, Epplen JT, Schmid M, Parzefall J. 1995. Incorporation of subgenomic amounts of DNA as compensation for mutational load in a gynogenetic fish. *Nature* **373**(6509): 68-71.

Schlegel R, Pohler W. 1994. Identification of an a-B Chromosome-Translocation in Diploid Rye (Secale-Cereale L). *Breeding Science* **44**(3): 279-283.

Schmid M, Ziegler CG, Steinlein C, Nanda I, Haaf T. 2002. *Chromosome banding in Amphibia: XXIV. The B chromosomes of Gastrotheca espeletia (Anura, Hylidae)*. Basel, SUISSE: Karger.

Schubert I, Rieger R. 1981. Sister chromatid exchanges and heterochromatin. *Human Genetics* **57**(2): 119-130.

Schubert V, Meister A, Tsujimoto H, Endo TR, Houben A. 2011. Similar rye A and B chromosome organization in meristematic and differentiated interphase nuclei. *Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology* **19**(5): 645-655.

Schvartzman JB. 1987. Sister-chromatid exchanges in higher plant cells: Past and perspectives. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **181**(1): 127-145.

Sergeeva EM, Salina EA, Adonina IG, Chalhoub B. 2010. Evolutionary analysis of the CACTA DNA-transposon Caspar across wheat species using sequence comparison and in situ hybridization. *Mol Genet Genomics* **284**(1): 11-23.

Shang H-Y, Wei Y-M, Wang X-R, Zheng Y-L 2006. Genetic diversity and phylogenetic relationships in the rye genus Secale L. (rye) based on Secale cereale microsatellite markers.In: scielo. 685-691.

Sheppard AE, Ayliffe MA, Blatch L, Day A, Delaney SK, Khairul-Fahmy N, Li Y, Madesis P, Pryor AJ, Timmis JN. 2008. Transfer of plastid DNA to the nucleus is elevated during male gametogenesis in tobacco. *Plant Physiology* **148**(1): 328-336.

Sheppard AE, Timmis JN. 2009. Instability of plastid DNA in the nuclear genome. *PLoS Genet* **5**(1): e1000323.

Shirasu K, Schulman AH, Lahaye T, Schulze-Lefert P. 2000. A contiguous 66-kb barley DNA sequence provides evidence for reversible genome expansion. *Genome Res* **10**(7): 908-915.

Smith GP. 1976. Evolution of repeated DNA sequences by unequal crossover. *Science* **191**(4227): 528-535.

Stark EA, Connerton I, Bennett ST, Barnes SR, Parker JS, Forster JW. 1996. Molecular analysis of the structure of the maize B-chromosome. *Chromosome Res* **4**(1): 15-23.

Steflova P, Tokan V, Vogel I, Lexa M, Macas J, Novak P, Hobza R, Vyskot B, Kejnovsky E. 2013. Contrasting patterns of transposable element and satellite distribution on sex chromosomes (XY1Y2) in the dioecious plant Rumex acetosa. *Genome biology and evolution* **5**(4): 769-782.

Świtoński M, Gustavsson I, Höjer K, Plöen L. 1987. Synaptonemal complex analysis of the B-chromosomes in spermatocytes of the silver fox (Vulpes fulvus Desm.). *Cytogenetic and Genome Research* **45**(2): 84-92.

Teruel M, Cabrero J, Perfectti F, Acosta MJ, Sanchez A, Camacho JP. 2009. Microdissection and chromosome painting of X and B chromosomes in the grasshopper Eyprepocnemis plorans. *Cytogenet Genome Res* **125**(4): 286-291.

Teruel M, Cabrero J, Perfectti F, Camacho JPM. 2010. B chromosome ancestry revealed by histone genes in the migratory locust. *Chromosoma* **119**(2): 217-225.

Timmis JN, Ayliffe MA, Huang CY, Martin W. 2004. Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. *Nature reviews. Genetics* **5**(2): 123-135.

Timmis JN, Ingle J, Sinclair J, Jones N. 1975. The Genomic Quality of Rye B Chromosomes. *Journal of Experimental Botany* **26**(3): 367-378.

Tomita M. 2010. Revolver and superior: novel transposon-like gene families of the plant kingdom. *Curr Genomics* **11**(1): 62-69.

Tomita M, Shinohara K, Morimoto M. 2008. Revolver is a new class of transposon-like gene composing the triticeae genome. *DNA Res* **15**(1): 49-62.

Torres GA, Gong Z, Iovene M, Hirsch CD, Buell CR, Bryan GJ, Novak P, Macas J, Jiang J. 2011. Organization and evolution of subtelomeric satellite repeats in the potato genome. *G3* **1**(2): 85-92.

Triboush SO, Danilenko NG, Davydenko OG. 1998. A Method for Isolation of Chloroplast DNA and Mitochondrial DNA from Sunflower. *Plant Molecular Biology Reporter* **16**(2): 183-183.

Trifonov VA, Perelman PL, Kawada SI, Iwasa MA, Oda SI, Graphodatsky AS. 2002. Complex structure of B-chromosomes in two mammalian species:

Apodemus peninsulae (Rodentia) and Nyctereutes procyonoides (Carnivora). *Chromosome Res* **10**(2): 109-116.

Trivers R, Burt A, Palestis BG. 2004. B chromosomes and genome size in flowering plants. *Genome / National Research Council Canada = Genome / Conseil national de recherches Canada* **47**(1): 1-8.

Tsujimoto H, Niwa K. 1992. DNA-Structure of the B-Chromosome of Rye Revealed by Insitu Hybridization Using Repetitive Sequences. *Japanese Journal of Genetics* **67**(3): 233-241.

Tsukahara S, Kawabe A, Kobayashi A, Ito T, Aizu T, Shin-i T, Toyoda A, Fujiyama A, Tarutani Y, Kakutani T. 2012. Centromere-targeted de novo integrations of an LTR retrotransposon of Arabidopsis lyrata. *Genes & development* **26**(7): 705-713.

Venere PC, Miyazawa CS, Galetti PM. 1999. New cases of supernumerary chromosomes in characiform fishes. *Genetics and Molecular Biology* **22**(3): 345-349.

Vershinin AV, Schwarzacher T, Heslop-Harrison JS. 1995. The large-scale genomic organization of repetitive DNA families at the telomeres of rye chromosomes. *Plant Cell* **7**(11): 1823-1833.

Wang D, Lloyd AH, Timmis JN. 2012. Environmental stress increases the entry of cytoplasmic organellar DNA into the nucleus in plants. *Proceedings of the National Academy of Sciences of the United States of America* **109**(7): 2444-2448.

Ward EJ. 1973. Nondisjunction: localization of the controlling site in the maize B chromosome. *Genetics* **73**(3): 387-391.

Wicker T, Buchmann JP, Keller B. 2010. Patching gaps in plant genomes results in gene movement and erosion of colinearity. *Genome Res* **20**(9): 1229-1237.

Wicker T, Guyot R, Yahiaoui N, Keller B. 2003. CACTA transposons in Triticeae. A diverse family of high-copy repetitive elements. *Plant Physiol* **132**(1): 52-63.

Wicker T, Keller B. 2007. Genome-wide comparative analysis of copia retrotransposons in Triticeae, rice, and Arabidopsis reveals conserved ancient evolutionary lineages and distinct dynamics of individual copia families. *Genome Res* **17**(7): 1072-1081.

Wicker T, Stein N, Albar L, Feuillet C, Schlagenhauf E, Keller B. 2001. Analysis of a contiguous 211 kb sequence in diploid wheat (Triticum monococcum L.) reveals multiple mechanisms of genome evolution. *Plant J* **26**(3): 307-316.

Wicker T, Taudien S, Houben A, Keller B, Graner A, Platzer M, Stein N. 2009. A whole-genome snapshot of 454 sequences exposes the composition of the barley genome and provides evidence for parallel evolution of genome size in wheat and barley. *Plant Journal* **59**(5): 712-722.

Wilkes TM, Francki MG, Langridge P, Karp A, Jones RN, Forster JW. 1995. Analysis of Rye B-Chromosome Structure Using Fluorescence in-Situ Hybridization (Fish). *Chromosome Research* **3**(8): 466-472.

Wolf KW, Mertl HG, Traut W. 1991. Structure, mitotic and meiotic behaviour, and stability of centromere-like elements devoid of chromosome arms in the fly Megaselia scalaris (Phoridae). *Chromosoma* **101**(2): 99-108.

Yoshida K, Terai Y, Mizoiri S, Aibara M, Nishihara H, Watanabe M, Kuroiwa A, Hirai H, Hirai Y, Matsuda Y, Okada N. 2011. B Chromosomes Have a

Functional Effect on Female Sex Determination in Lake Victoria Cichlid Fishes. *PLoS Genet* **7**(8): e1002203.

Zhang P, Li W, Fellers J, Friebe B, Gill BS. 2004. BAC-FISH in wheat identifies chromosome landmarks consisting of different types of transposable elements. *Chromosoma* **112**(6): 288-299.

Zhou JP, Yang ZJ, Li GR, Liu C, Tang ZX, Zhang Y, Ren ZL. 2010. Diversified chromosomal distribution of tandemly repeated sequences revealed evolutionary trends in Secale (Poaceae). *Plant Systematics and Evolution* **287**(1-2): 49-56.

Zhou Q, Zhu HM, Huang QF, Zhao L, Zhang GJ, Roy SW, Vicoso B, Xuan ZL, Ruan J, Zhang Y, Zhao RP, Ye C, Zhang XQ, Wang J, Wang W, Bachtrog D. 2012. Deciphering neo-sex and B chromosome evolution by the draft genome of Drosophila albomicans. *Bmc Genomics* **13**: 109.

Zhu Y, Dai J, Fuerst PG, Voytas DF. 2003. Controlling integration specificity of a yeast retrotransposon. *Proceedings of the National Academy of Sciences* **100**(10): 5891-5895.

Ziegler CG, Lamatsch DK, Steinlein C, Engel W, Schartl M, Schmid M. 2003. The giant B chromosome of the cyprinid fish Alburnus alburnus harbours a retrotransposon-derived repetitive DNA sequence. *Chromosome Res* **11**(1): 23-35.

List of publications related to this thesis

Articles:

Klemme S, Banaei-Moghaddam AM, Macas J, Wicker T, Novák P, Houben A. 2013. High-copy sequences reveal distinct evolution of the rye B chromosome. *New Phytologist* **199**(2): 550-558.

Martis MM, Klemme S, Banaei-Moghaddam AM, Blattner FR, Macas J, Schmutzer T, Scholz U, Gundlach H, Wicker T, Šimková H, Novák P, Neumann P, Kubaláková M, Bauer E, Haseneyer G, Fuchs J, Doležel J, Stein N, Mayer KFX, Houben A. 2012. Selfish supernumerary chromosome reveals its origin as a mosaic of host genome and organellar sequences. *Proceedings of the National Academy of Sciences* **109**(33): 13343-13346.

Marques A, Klemme S, Guerra M, Houben A. 2012. Cytomolecular characterization of de novo formed rye B chromosome variants. *Molecular cytogenetics* **5**(1): 34.

Marques A, Banaei-Moghaddam AM, Klemme S, Blattner FR, Niwa K, Guerra M, Houben A. 2013. B chromosomes of rye are highly conserved and accompanied the development of early agriculture. *Annals of Botany*.

Banaei-Moghaddam AM, Schubert V, Kumke K, Weiβ O, Klemme S, Nagaki K, Macas J, González-Sánchez M, Heredia V, Gómez-Revilla D, González-García M, Vega JM, Puertas MJ, Houben A. 2012. Nondisjunction in Favor of a Chromosome: The Mechanism of Rye B Chromosome Drive during Pollen Mitosis. *The Plant Cell Online* **24**(10): 4124-4134.

Houben, A., Banaei-Moghaddam, A.M., Klemme, S. & Timmis, J.N. 2013 Evolution and biology of supernumerary B chromosomes. *submitted* .

Book chapter:
Houben, A., Banaei-Moghaddam, A. & Klemme, S. Biology and Evolution of B Chromosomes. in *Plant Genome Diversity Volume 2* (eds. Greilhuber, J., Dolezel, J. & Wendel, J.F.) 149-165 (Springer Vienna, 2013).

Curriculum Vitae

Personal Data

Name: Sonja Klemme

Adress: Valentin-Feldmann Straße 7
39218 Schönebeck, Germany

Telephone: +49 (0)176 70271577

E-Mail: Klemme@ipk-gatersleben.de

Birthdate and -place: 19.03.1984 in Schönebeck/Elbe

marital status: unmarried

School Education

1990 –1994	„Grundschule Dr. Salvador Allende“ in Schönebeck/Elbe
1994 –2003	„Gymnasium am Malzmühlenfeld“ in Schönebeck/Elbe, Certification: Abitur (1.5), 04.07.03

University Education

2003 – 2006	University of Bremen Course of studies: Diplom Biology
2006 – 2008	University of Cologne Course of studies: Bachelor Biology Certification: Bachelor of Science in Biology (2.1), 20.08.08
2008 – 2009	University of Cologne Course of studies: Master of Biology Certification: Master of Science in Biology (1.9), 21.12.09

Research Experiences

2008	Bachelor Thesis at the University of Cologne Group: Bacterial Genetics (Prof. Karin Schnetz) Topic: Mapping of the Interaction domains of LuxR type transcription factors RcsA, RcsB, YjjQ and BglJ in *Escherichia coli*.
2009	Master Thesis at the Max Plack Institute for Plant Breeding in Cologne, Germany

	Group: Transcription Control in Flowering Time (Dr. Franziska Turck) Topic: Analysis of expression of FT under various inductive photoperiods in *Arabidopsis*.
2013	PhD Thesis at the Leibniz Institute for Plant Genetics and Crop Plant Research (IPK) in Gatersleben, Germany Group: Chromosome Structure and Function (Dr. Andreas Houben) Topic: High-copy sequences reveal a distinct evolution of the rye B chromosome

Prizes:

1st prize best talk: Student Scientific Conference on Biotechnology and Biomedicine 2012, Brno, Czech Republic

1st prize best talk: Plant Science Student Conference 2012, Leibniz Institute of Plant Genetics and Crop Plant Research, Gatersleben, Germany

Audience price best talk: Plant Science Student Conference 2012, Leibniz Institute of Plant Genetics and Crop Plant Research, Gatersleben, Germany

2nd price poster: International Chromosome Conference 2012, Manchester, England

Erklärung an Eides statt

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Arbeit selbstständig und ohne fremde Hilfe verfasst habe. Es wurden keine anderen als die in der Arbeit angegebenen Quellen und Hilfsmittel benutzt. Die den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen sind als solche kenntlich gemacht.

Hiermit erkläre ich, dass ich noch keine vergeblichen Promotionsversuche unternommen habe und die vorliegende Dissertation nicht in der gegenwärtigen bzw. in einer anderen Fassung bereits einer anderen Fakultät / anderen wissenschaftlichen Einrichtung vorgelegt habe.

Halle/Saale, ……

Unterschrift